



**VNiVERSiDAD D SALAMANCA**

*Departamento de Informática y Automática*

**ANALÍTICA VISUAL EN *E*LEARNING**

**TESIS DOCTORAL**

D. DIEGO ALONSO GÓMEZ AGUILAR

**Directores:**

DR. D. FRANCISCO JOSÉ GARCÍA PEÑALVO

DR. D. ROBERTO THERÓN SÁNCHEZ

Abril 2015





**VNiVERSiDAD D SALAMANCA**

*Departamento de Informática y Automática*

## **ANALÍTICA VISUAL EN ELEARNING**

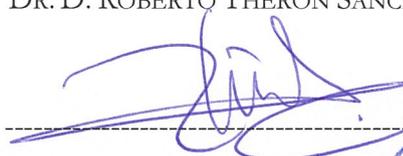
**TESIS DOCTORAL**

**Directores:**

DR. D. FRANCISCO JOSÉ GARCÍA  
PEÑALVO

DR. D. ROBERTO THERÓN SÁNCHEZ

  
-----

  
-----

**Doctorando:**

D. DIEGO ALONSO GÓMEZ AGUILAR

  
-----

Marzo 2015







Dpto. de Informática y Automática  
Facultad de Ciencias  
Plaza de los Caídos s/n 37008  
Salamanca, España

**D. Francisco José García Peñalvo** y **Roberto Therón Sánchez**, profesores Titulares de Universidad del Departamento de Informática y Automática de la Universidad de Salamanca, en calidad de directores del trabajo de tesis doctoral titulado “**Analítica Visual en eLearning**” y realizado por Diego Alonso Gómez Aguilar

HACEN CONSTAR

Que dicho trabajo tiene suficientes méritos teóricos contrastados adecuadamente mediante las validaciones oportunas, publicaciones relacionadas y aportaciones novedosas. Por todo ello consideran que procede su defensa pública.

En Salamanca, a 16 de Marzo de 2015.

Dr. D. Francisco José García Peñalvo  
Dpto. Informática y Automática  
Universidad de Salamanca



Dr. Roberto Therón Sánchez  
Dpto. Informática y Automática  
Universidad de Salamanca



## Agradecimientos

*Quiero agradecer en primer lugar a mis directores de tesis, Francisco José García Peñalvo y Roberto Therón Sánchez, por sus enseñanzas, su ayuda y dedicación. He aprendido mucho y seguro que seguiré haciéndolo.*

*Muchas gracias!*

*También quiero agradecer a todo el grupo de investigación GRLAL y VIS-USAL, porque es un placer teneros como compañeros, y aún mas en especial a los que han contribuido de algún modo en la realización de esta Tesis, gracias Alicia, Paco, Valentina, Miguel Ángel, Rodrigo y Carlos. Tampoco quiero olvidar la ayuda recibida del resto de compañeros de la USAL, en especial de José Antonio, Antonio, Silvana, Hicham y Sergio.*

*También quiero agradecer especialmente a L. Jessica C-H, F.G.M. Patricia C., B.S.F. Lucia R., P. Giulia P.M., W.P. Lydia S., T.G. Manuel Guarneros, Stefan H., C. P. Juan F.G.N., N. Carla R. M. V., Ch. Mibaela O. M., C. C. Carlos A.G.P., L. S. Lara C. y C. U. Laura G. la ayuda con el mantenimiento de los ánimos y con todo, gracias siempre, ya que sin su apoyo y cariño no hubiese sido posible y además por sus apoyos a C. Ángel H., M. Damiana M. S., M.P. Cristina S. G., Luis Zamora C., R. Chain y S. A. Fernando Montes.*

*Quiero agradecer a M.B. Ignacio Marcué y a C.I. Ronal Bermudes el tiempo y la dedicación a este trabajo de investigación durante mi trayecto de revisiones y correcciones.*

*Me gustaría agradecer, además, a mis amigos colaboradores A. Andrea K., O. Valeria R., T. Teresa L., T. Marcos N., D. Daniel I., E. Leticia N., F. David R., Ch. Jesús C., Mariu, VB. Vanesa y M. Raúl. Sin olvidar a Carlos R., G. Diego C. M., C. Catalina G. G-H. y L. Andren, CH. Edgar O. E. B.*

*Agradezco de todo corazón a mi familia, porque me han apoyado y animado constantemente, sin esas muestras de apoyo todo hubiera sido mucho más duro, gracias Papa, Ch. Marisol, Abuela, G. Guille, Mamá, B. Miguel Ángel, C. León, Ch. Cecilia, C. Luis, tíos y familiares*

*¡cuánto te tengo que agradecer!*

*Y especialmente a los dioses y a Diego, mi amigo y compañero en la vida, a ellos por su amor, apoyo incondicional y ayuda en todo. Gracias dioses y gracias Diego.*



## Resumen

La docencia universitaria ha experimentado cambios espectaculares en los últimos años debido al impacto de la tecnología en diferentes actividades cotidianas. El *eLearning* (o aprendizaje electrónico) y el *bLearning* (*Blended Learning* o aprendizaje mixto) nacieron gracias a este fenómeno, ya fuera como alternativa o complemento a la enseñanza tradicional. Esto ha producido cambios de paradigma en los últimos años en la docencia universitaria donde muchos profesores se han apoyado o han sustituido las clases presenciales de la enseñanza tradicional, entendiendo la enseñanza tradicional como cara a cara (del inglés *Face to Face*, F2F) por el aula virtual.

La Web 2.0 ha abierto nuevas posibilidades, que incluyen numerosas formas de aprendizaje y colaboración entre estudiantes y profesores. Por ello, ha surgido la necesidad, por parte de los educadores, de adoptar diferentes estrategias para obtener información sobre el rendimiento de sus estudiantes, así como plantear nuevas formas de evaluación basadas en el análisis de información educativa, que sean capaces de medir y cuantificar la cantidad de trabajo, así como el número y la calidad de las habilidades que han adquirido. Tales estrategias de comunicación requieren de nuevos métodos analíticos que hagan posible comprender y analizar la propia plataforma y el aprendizaje. Estas consideraciones arrojan nueva luz sobre la evaluación de los estudiantes, que ya no puede basarse únicamente en los resultados de los exámenes finales convencionales, sino en un proceso educativo integral que considere y evalúe otras competencias más allá de las académicas.

La disponibilidad y facilidad de uso de los recursos web ha permitido el uso extendido de los *Learning Management Systems*, o plataformas de *eLearning*. Sin embargo, los educadores que usan estos entornos se encuentran con graves limitaciones a la hora de evaluar las actividades de los estudiantes, de discriminar sus comportamientos *online* y de evaluar la propia plataforma y la utilidad de esta. Por ello, es necesario encontrar y desarrollar técnicas novedosas para obtener información sobre las pautas de aprendizaje y comportamiento de los estudiantes en un entorno electrónico.

En esta tesis doctoral se propone un modelo de visualización analítica en *eLearning* como base para construir una estrategia de seguimiento y evaluación de la información que proporciona, no solo a los profesores, sino también a gestores académicos y estudiantes, información necesaria para entender el proceso de aprendizaje de los estudiantes en una plataforma de *eLearning*, que sirva de guía para el alumnado y que proporcione métricas para los gestores sobre la plataforma y el desempeño, además de tomarse como base para desarrollo de futuros sistemas de analítica visual en *eLearning*. El modelo proporciona los elementos para crear un sistema de analítica visual en *eLearning* encaminado a perfeccionar el proceso de enseñanza/aprendizaje.

Este sistema se ha diseñado mediante una arquitectura constituida por distintas capas. La capa inferior está sustentada en un conjunto de servicios web que permiten la extracción de los datos a analizar del servidor. La siguiente capa, contiene la lógica de pre-procesamiento,

estandarización y análisis de los datos. Por último, una tercera capa, en la que se realiza el proceso de analítica visual, que permite al profesor, estudiante o gestor académico llevar a cabo un análisis más exhaustivo, completo e interactivo. Con la finalidad de poner en práctica y realizar una prueba de los alcances de este sistema, se ha desarrollado un prototipo plenamente funcional del mismo.

El desarrollo del prototipo se realizó por medio de un conjunto de iteraciones de investigación-acción para la mejora del alcance de las capacidades de análisis del sistema y de la usabilidad del prototipo de visualización analítica visual en *eLearning*, con el objetivo último de soportar el proceso de aprendizaje, el rendimiento académico y, a su vez, y como ya se mencionó, estas aplicaciones al tomarse como base para desarrollo de futuros sistemas de analítica visual en *eLearning*. Se utiliza como fuente de datos el sistema de gestión de aprendizaje Moodle. Los resultados obtenidos se complementaron y probaron con un estudio de los patrones de uso de las plataformas de *eLearning* en dos universidades con distintos contextos pedagógicos y sociales: la Universidad Politécnica de Madrid y la Universidad de Salamanca.

Esta valiosa experiencia produjo un caudal de nueva información y conocimiento y, por tanto, una importante fuente de realimentación que han contribuido a la mejora notable de las capacidades de análisis que ofrece la plataforma y cubre adecuadamente las necesidades y funcionalidades que se requieren en el modelo propuesto y descrito en esta Tesis Doctoral.

## **Palabras clave:**

Visualización, *eLearning*, Análisis, Analítica Visual, Analítica Académica, Analítica del Aprendizaje, VeLA.

## Abstract

University teaching has changed dramatically in recent years due to the impact of technology on different daily activities. eLearning (e-learning) and *bLearning* (semipresencial learning or blended learning) were born thanks to this phenomenon, either as an alternative or complement to traditional teaching. This paradigm shift has occurred in recent years in university teaching and many teachers have supported or have replaced the face-to-face classes of the traditional teaching with the virtual classroom.

The Web 2.0 has opened new possibilities, including numerous forms of learning and collaboration among students and teachers. Therefore, a need has arisen on the part of educators to adopt different strategies for gathering information about student performance and propose new forms of assessment based on the analysis of educational information, which enables the measurement and quantification of the amount of work, and the number and quality of the skills they have acquired. Such communication strategies require new analytical methods that make it possible to understand and analyze both the platform and the learning process. These considerations shed new light on the evaluation of the students, who can no longer rely solely on the results of conventional final exams, but in a comprehensive educational process to consider and evaluate other skills beyond academics.

The availability and usability of Web resources have enabled widespread use of Learning Management Systems, or eLearning platforms. However, educators using these environments have serious limitations when evaluating the activities of students, to discriminate their online behavior, assess the platform and its usefulness. Therefore, it is necessary to find and develop new techniques to obtain information on patterns of learning and behavior of students in an electronic environment.

In this doctoral thesis a model of visual analytics on eLearning is proposed as a basis for building a strategy for monitoring and evaluation of the information to provide, not only to teachers but also to academic managers and students, information necessary to understand the learning process of students in a eLearning platform which serves as a guide for students and provide metrics for academic managers about the platform and its performance, in addition to become the basis for future development of visual analytics systems in eLearning. The model provides the elements to create a visual analytics system in eLearning aimed at improving the teaching / learning processes.

The system has been specifically designed to achieve the mentioned goals, which consists of different layers. In the lower layer, the system is supported by a set of web services that enable the extraction of data to analyze from a server. The next layer contains the preprocessing, standardizing and data analysis logic. Finally, the third layer, where the process of visual analytics is performed, allowing the teacher, student or manager to conduct a more thorough, comprehensive and interactive analysis. In order to assess and run a test of the scope of the proposed model, a prototype of a visual eLearning analytics tool has been developed.

The development of the prototype was carried out by a set of action research iterations to improve the analytical capacity of the system and the usability of the prototype, with the aim to enhance the learning process, academic performance and in turn, as it was mentioned, these applications at the moment that these are taken into account as a base of development of the future Visual eLearning Analytics Systems. Used as the data source the learning management system Moodle. The results were complemented and proved with a study of the patterns of use of eLearning platforms at two universities with different educational and social contexts; the Polytechnic University of Madrid and the University of Salamanca.

This valuable experience produced a wealth of new information and knowledge, and therefore an important source of feedback, which has enabled a significant improvement of the analytical capabilities that the platform provides and meets adequately the needs and functionalities, which are required to implement the model that is presented and is described in this doctoral thesis.

## **Keywords:**

Visualization, eLearning, Analysis, Visual Analytics, Academic Analytics, Learning Analytics, VeLA.

# Índice general

<b>Capítulo I</b>	<b>Introducción</b> .....	<b>23</b>
I.1	Motivación.....	24
I.2	Problemática.....	34
I.2.1	Limitaciones del proceso de enseñanza/aprendizaje.....	35
I.2.2	Retos y necesidades del proceso de enseñanza/aprendizaje.....	37
I.3	Objetivos y Preguntas de Investigación.....	41
I.3.1	Preguntas de investigación.....	42
I.3.2	Objetivo principal.....	42
I.3.3	Objetivos específicos.....	42
I.4	Metodología de investigación.....	43
I.5	Líneas de investigación de la tesis.....	45
I.6	Contenido.....	48
<b>Capítulo II</b>	<b>El aprendizaje y la tecnología</b> .....	<b>51</b>
II.1	Inicios del Aprendizaje Electrónico.....	52
II.1.1	Formación Presencial.....	55
II.1.2	Formación mixta.....	56
II.1.3	Formación virtual.....	56
II.2	Definición de <i>eLearning</i> .....	57
II.3	Sistemas de administración de aprendizaje.....	62
<b>Capítulo III</b>	<b>Análítica en la educación</b> .....	<b>67</b>
III.1	Análítica.....	68
III.2	Análítica Académica.....	76
III.3	Análítica del Aprendizaje.....	81
III.3.1	Proceso de la Análítica del Aprendizaje.....	83
III.3.2	Modelo de referencia para la Análítica del Aprendizaje.....	88
III.4	Minería de datos para la educación.....	92
III.5	Análítica social del aprendizaje.....	94
III.6	Big data.....	97
<b>Capítulo IV</b>	<b>Visualización de la Información y Análítica Visual</b> .....	<b>101</b>
IV.1	Visualización.....	103
IV.2	Fundamentos de la psicología de la educación.....	105
IV.3	Visualización de la información.....	107
IV.3.1	Tipos de información.....	108
IV.3.2	Interacción.....	109
IV.4	Análítica Visual.....	117
IV.5	Técnicas y estrategias de <i>infoVis/VA</i> .....	121
IV.5.1	Estrategias de diseño de visualización.....	122
IV.5.2	Estrategias de diseño de navegación.....	125
IV.5.3	Técnicas y estrategias de interacción.....	126
IV.6	Alfabetización de datos e información.....	127

<b>Capítulo V Aportaciones al estado del conocimiento actual de Analítica Educativa, Analítica Académica y Analítica Visual en eLearning</b>	<b>131</b>
V.1 Las representaciones visuales: Técnicas de interacción .....	132
V.1.1 Contribuciones visuales que apoyan al <i>eLearning</i> .....	136
V.1.2 Contribuciones sin representación visual que apoyan al <i>eLearning</i> .....	156
V.1.3 Contribuciones visuales independientes al <i>eLearning</i> .....	160
V.2 Análisis, problemas encontrados y posibles soluciones.....	181
V.2.1 Análisis .....	182
V.2.2 Problemas y soluciones posibles .....	198
<b>Capítulo VI Modelo de Analítica Visual en eLearning .....</b>	<b>203</b>
VI.1 Diseño del modelo de analítica visual en <i>eLearning</i> .....	204
VI.1.1 Analítica Visual aplicada al <i>eLearning</i> ( <i>Visual eLearning Analytics</i> ).....	205
VI.1.2 Definición del modelo de analítica visual aplicada al <i>eLearning</i> .....	206
VI.2 Proceso del diseño y evaluación del prototipo de VeLA .....	212
VI.2.1 Implementación y acceso a las fuentes de datos .....	212
VI.2.2 Proceso de evaluación del diseño del prototipo.....	233
<b>Capítulo VII Casos de estudio.....</b>	<b>241</b>
VII.1 Descripción del prototipo a través de casos de estudio.....	242
VII.1.1 Caso de estudio 0.....	242
VII.1.2 Caso de estudio 1.....	244
VII.1.3 Caso de estudio 2.....	253
VII.1.4 Caso de estudio 3.....	258
VII.1.5 Caso de estudio 4.....	264
VII.2 Evaluación de usabilidad .....	268
<b>Capítulo VIII Conclusiones y trabajos futuros.....</b>	<b>271</b>
VIII.1 Conclusiones .....	272
VIII.2 Trabajos futuros .....	276
VIII.3 Publicaciones relacionadas con la tesis .....	276
VIII.3.1 Revistas.....	276
VIII.3.2 Congresos.....	277
VIII.3.3 Capítulos de libro y Postproceedings .....	280
<b>Capítulo IX Anexos.....</b>	<b>281</b>
IX.1 Encuesta.....	281
IX.2 Acrónimos .....	283

## Índice de Figuras

Fig. 1. Espectro de los usos de Internet como herramienta docente en el contexto universitario. Fuente: adaptado de (Duart Montoliu, Gil, Pujol, & Castaño, 2008) .....	55
Fig. 2. Proceso de la analítica del aprendizaje, modificado de (Chatti et al., 2012). .....	83
Fig. 3. Analítica del aprendizaje y la pedagogía; modificado de (Greller & Drachsler, 2012). .....	85
Fig. 4. Dimensiones críticas de la analítica del aprendizaje (Greller & Drachsler, 2012). .....	87
Fig. 5. Modelo de referencia para la analítica del aprendizaje, modificado de (Chatti et al., 2012). .....	89
Fig. 6. Discurso analítico implica la retroalimentación mutua entre los objetivos y las tareas del usuario con las posibles comunicaciones y usos de las visualizaciones interactivas, adaptado de (Pike et al., 2009). .....	110
Fig. 7. La analítica visual integra visualización como núcleo en conjunto con disciplinas adyacentes y depende de la disponibilidad de instalaciones de infraestructura y evaluación adecuados (D. Keim et al., 2010). .....	117
Fig. 8. La analítica visual integra disciplinas científicas para mejorar la división del trabajo entre el hombre y la máquina, modificado de (D. Keim, Andrienko, et al., 2008). .....	118
Fig. 9. Las dos clases de problemas. Se pueden observar los problemas generales de aplicación del área de TI (blancos) y los problemas analíticos (gris claro); ambos se pueden resolver mediante el Análisis Automático (negro), técnicas de Visualización (gris) y Analítica Visual (gris oscuro). Debe tenerse en cuenta que no todos los problemas de un análisis automático o análisis visual se pueden resolver por medio de la analítica visual, si existen otras maneras efectivas y eficientes de resolver el problema. Modificado de (Daniel A. Keim et al., 2010). ..	119
Fig. 10. Potencial de la analítica visual. Modificado de (Daniel A. Keim et al., 2010). .....	120
Fig. 11. Proceso de VA propuesto por Keim et al. Presentado y modificado de (D. Keim, Andrienko, et al., 2008). .....	121
Fig. 12. Diseño y representación ontológica del escenario de <i>eLearning</i> de Guo and Chen (2006). Diseño y representación ontológica del escenario de <i>eLearning</i> de Guo and Chen (2006). .....	137
Fig. 13. Captura de pantalla del visor de TM4L (Dicheva & Dichev, 2006; Dicheva et al., 2005). .....	138
Fig. 14. Mapa del explorador de conocimiento (enfoque abierto de modelado del estudiante) y un botón para lanzar el maximizador de conocimiento (enfoque de secuenciación de problemas) (P. Brusilovsky, Baishya, Hosseini, Guerra, & Miner, 2013). .....	139

- Fig. 15. Esta figura es resultado de hacer *zoom* sobre el nodo *Expressions* (esquina superior izquierda como se marca en la Fig. 14). Revela el siguiente nivel de la jerarquía de dicho concepto. Ahora el usuario puede, ver el todo, que consta de varios conceptos desconocidos (coloreados en rojo) y varios bien aprendidos (coloreados en verde), que tiene conocimiento intermedio del concepto *Logic Expression* nodo que se muestra en color amarillo (P. Brusilovsky et al., 2013)..... 140
- Fig. 16. Visualización de la interacción. Se presenta un gráfico que representa la interacción entre los participantes de un curso de ejemplo. Los nodos en color cian representan los maestros, los nodos en naranja indican los estudiantes y un nodo en color verde representa a todos los participantes en el curso. Por tanto, ya que es posible enviar mensajes a todos los participantes, en esta visualización se pone en evidencia las personas que han enviado mensajes a todos los participantes. Utilizando la técnica de consultas dinámicas, los usuarios pueden modificar dinámicamente la visualización a través de controles en la parte superior de la ventana (Romani & da Rocha, 2000). ..... 140
- Fig. 17. Flujo de temas (Feipeng et al., 2013). ..... 141
- Fig. 18. Representación de HEODAR (Morales Morgado et al., 2009). Se divide en tres partes: 1) la representación de la media de los resultados de todas las evaluaciones realizadas de los LO (esta es la evaluación del LO); 2) el valor multiplicado por el porcentaje de profesores que han evaluado el LO con respecto al número total de profesores, lo que resulta en la evaluación ponderada y, por último, 3) el porcentaje de profesores que han evaluado el LO con respecto al número total de docentes. .... 142
- Fig. 19. Prototipo del analizador de narrativa (Williams & Conlan, 2007). 142
- Fig. 20. Cliente de escritorio. El tiempo se representa de izquierda a derecha, desplazable desde el principio hasta el final del proyecto (Ganoë et al., 2003). ..... 144
- Fig. 21. Un ejemplo de un menú de operación que aparece en un nodo particular que muestra todas las acciones, propiedades y atributos asociados del objeto correspondiente. En *ENCCON* los nodos representan los objetos (es decir, actividades, artefactos, grupos, eventos, agentes, perfiles, etc.), mientras que las líneas se utilizan para presentar relaciones entre los objetos, en este caso se limita a las jerarquías relacionales (Q. V. Nguyen et al., 2004)..... 145
- Fig. 22. Un ejemplo de dos diseños de gráficos de la misma red: a) mediante un *Spring Embedded* con particiones de *ENCCON*, y b) utilizando un algoritmo de diseño circular (Q. Nguyen & Huang, 2010). ..... 145
- Fig. 23. Un ejemplo de las diferentes visualizaciones de redes: a) visión abstracta de un gráfico jerárquico y b) vista de detalle asociada a un elemento (Q. Nguyen & Huang, 2010). ..... 146
- Fig. 24. Visualización del modelo basado en la línea. Cada fila representa la actividad de escritura de un usuario. Los puntos están conectados por líneas y el grosor de una línea indica la intensidad de la conducta del usuario durante un período de tiempo. Esta figura muestra un ejemplo de gráficos generados usando esta visualización: a) muestra toda la

- actividad, donde los puntos están conectados a las líneas en cada sub-actividad (se definió cada línea como una serie, que tiene diferente color y grosor); b) es una figura complementaria (*zoom* sobre la figura 2a), lo que da más detalles acerca de la conducta del estudiante sobre la actividad propuesta del proyecto de escritura (Ming et al., 2013)...... 146
- Fig. 25. Representación de las discusiones de la herramienta de (Mazza & Milani, 2004). ..... 147
- Fig. 26. El gráfico de Comportamiento representa el comportamiento del estudiante (Mazza & Dimitrova, 2005). ..... 149
- Fig. 27. A la izquierda la vista desde arriba de *CrystalChat* y a la derecha, una vista mirando hacia abajo en la estructura de las conversaciones en *CrystalChat* (Tat & Carpendale, 2006). ..... 150
- Fig. 28. Presentación del mapa de discusiones de *MTRDS*, los ejes X e Y denotan el día de la fecha y hora, respectivamente. Cada mensaje de discusión (*post*) está representado por un círculo con un código de color (nodo del mensaje) y una línea (enlace) que conecta un nodo respuesta al mensaje originario (Gibbs et al., 2006). ..... 151
- Fig. 29. Patrones de interacción entre los ocho participantes en las diferentes fases de una tarea de aprendizaje (De Laat et al., 2007). ..... 152
- Fig. 30. El uso de NAT para visualizar y filtrar los lazos sociales de la persona, el tipo de lazo, y tema. Debido a limitaciones de espacio, se trata de una imagen compuesta que muestra toda la red sin filtrar, pero cuando se selecciona un usuario de esta, se filtra el panel izquierdo, como se muestra, para mostrar solo su ego-red y temas (Schreurs et al., 2013). ..... 153
- Fig. 31. Ejemplo de la representación de (Silva & Figueira, 2012). a) Una red centralizada; b) Detección de *click*; c) Gráfico del foro “*Moodle creates success*”. ..... 154
- Fig. 32. La interfaz de usuario con tres visualizaciones (Govaerts et al., 2012). ..... 155
- Fig. 33. A) Una pantalla espiral de porcentajes de consumo mensuales para los 112 alimentos durante el período 1980-1988. Cada alimento se ha asignado uno de los doce colores. B) El ajuste de una vista espiral de datos de sonido de cinco instrumentos. De arriba hacia abajo la estructura del sonido se devela por sí sola. Esta muestra tres puntos de vista diferentes sobre la base de tiempo diferente por los valores de vuelta. Los datos provienen de una grabación de un disco compacto de cinco instrumentos (Carlis & Konstan, 1998). ..... 162
- Fig. 34. Dos visualizaciones de la intensidad del sol utilizando aproximadamente el mismo espacio de la pantalla y el mismo esquema de codificación por color, en este caso en escala de grises. En la visualización de espiral es mucho más fácil comparar los días para detectar períodos nublados, o para ver eventos como el amanecer y la puesta del sol (Weber et al., 2001). ..... 163
- Fig. 35. El reloj de la conversación representa cada participante en un color diferente. La longitud de cada rectángulo es proporcional al volumen capturado en ese punto en el tiempo. Los puntos a lo largo de la circunferencia de cada anillo implican que la conversación está activa

- pero nadie está hablando. Cada anillo representa un minuto de conversación (Karrie G. Karahalios & Bergstrom, 2009). ..... 164
- Fig. 36. La representación de los votos de la conversación es compatible con un botón de votación para contribuir de manera significativa, y este voto está representado por medio de los puntos en los bordes estrechos de las barras rectangulares (Karrie G. Karahalios & Bergstrom, 2009). ..... 164
- Fig. 37. Una vista de la visualización de agrupaciones de la conversación. Las palabras encerradas están a una distancia común dada su similitud. Las palabras fuera de las áreas cerradas están por encima del umbral de similitud de ese clúster tema. Las palabras aparecen en tiempo real conforme se habla (Karrie G. Karahalios & Bergstrom, 2009). ..... 165
- Fig. 38. *ThemeRiver* utiliza una metáfora del río para representar los temas de una colección de discursos de Fidel Castro, entrevistas y artículos desde finales de 1959 hasta mediados de 1961 (S. Havre et al., 2000; Susan Havre et al., 2002). ..... 166
- Fig. 39. 1) *TimeZoom* con dos regiones de atención adicionales que muestran diferentes niveles de detalle. 2) Definición de las diferentes zonas de las regiones de enfoque. 3) Las partes de la interfaz de usuario, (Dachsel & Weiland, 2006). ..... 167
- Fig. 40. Arriba, 1) panel de la visión general línea de tiempo, 2) panel de la vista cronológica de detalles y 3) el panel del filtro de dimensión. Los tres paneles tienen atributos que han sido diseñados para hacer frente a los problemas de escala, la jerarquía y las relaciones entre estos y el valor temporal. Abajo, el panel de información general continuo que muestra los datos temporales como un histograma, para ampliar grandes conjuntos de datos sin perder el contexto -visión detallada del panel (1)- (André et al., 2007). ..... 167
- Fig. 41. Aplicación *elastic tag maps* (Stefaner, 2007). [http://well-formed-data.net/experiments/tag\\_maps\\_v5/](http://well-formed-data.net/experiments/tag_maps_v5/). ..... 170
- Fig. 42. Un PTC que revela las diferencias en la prevalencia de drogas entre las decisiones escritas de los Tribunales de Estados Unidos del Circuito de Apelaciones. Los Tribunales de Circuito de Apelación se componen de las 12 divisiones judiciales regionales (numeradas de primero a undécimo, más el Circuito DC) (Collins et al., 2009). ..... 171
- Fig. 43. Alternativas de diseño para unificar las nubes de etiquetas y los mini gráficos (Bongshin et al., 2010). ..... 171
- Fig. 44. En el centro de la parte superior de la figura se presenta un visor gráfico de tendencias que muestra una curva de importancia extraído de una colección de documentos con diferentes puntos de tiempo (Weiwei et al., 2010). ..... 172
- Fig. 45. En la figura se muestran las actrices que han ganado premios como actriz principal o secundaria (en la parte central y abajo). Algunas de ellas son Julia Roberts, Nicole Kidman, Jennifer Connelly y Renee Zellweger. Se puede ver el perfil de sus películas en las coordenadas paralelas (arriba a la derecha) y el género más relevante de sus películas en la nube de palabras (arriba a la izquierda) (Therón, Santamaria, et al., 2007). ..... 172

- Fig. 46. La figura representa los eventos en el año 2007 de tres diferentes barcos utilizados para inmigración. Es muy obvio que no hay tendencias y/o patrones de uso del tipo de barco, y también que el número de barcos de algún tipo no son lo mismo (Matkovic, Ammer, Gracanin, Purgathofer, & Lez, 2010). ..... 173
- Fig. 47. El éxito de los inmigrantes. El punto de vista de la izquierda muestra fracaso (contexto) como gris y éxito (cepillado) en rojo. Marzo, mayo y junio fueron meses de mayor éxito para los inmigrantes. La vista derecha muestra lo mismo pero con una única opción seleccionada (por lo que no se muestra el contexto) (Matkovic et al., 2010). ..... 174
- Fig. 48. Árbol de palabras de los discursos de campaña presidencial de Obama, con la distancia de Jaccard, y la coloración cronológica. El rojo corresponde al comienzo de la campaña ("niños", "Irak", "guerra", "mundo"), mientras que azul corresponde a la final ("McCain", "Wall Street", "crisis", "impuestos") (Gambette & Véronis, 2010). ..... 174
- Fig. 49. Tres ejemplos de los tipos de redes que son el tema de esta revisión. (a) Una visualización de la estructura de la red de Internet a nivel de "sistemas autónomos" - grupos locales de computadoras cada uno en representación de cientos o miles de máquinas. (b) Una red social, en este caso de los contactos sexuales. (c) Una red trófica de las interacciones depredador-presa entre las especies en un lago de agua dulce (Newman, 2003). ..... 176
- Fig. 50. A) Un diseño de árbol para un modelo de datos que genera un gráfico grande. B) Disposición H-árbol. C) Vista Globo. D) Vista Radial. E) *TreeMap*: rectángulos con color pertenecen al mismo nivel de la (árbol) jerarquía ((Herman, Melancon, & Marshall, 2000). ..... 177
- Fig. 51. A) Versión en 3D de un algoritmo radial. B) Información Cube. C) Un árbol de conos. D) Navegador de archivos de sistema en SGI. E) Vista hiperbólica de un árbol en 3D (Herman et al., 2000). ..... 178
- Fig. 52. Una visión general de una la red semántica de términos sobre la base de las 100 dimensiones semánticas latentes en el conjunto de datos VAST más significativas. La lista de dimensiones conceptuales se filtra a través de una vista de ojo de pez para una fácil selección. Cada término está representado por un nodo rojo. El tamaño de los nodos refleja la contribución global del término a todo el espacio concepto (Zhu & Chen, 2007). ..... 179
- Fig. 53. La idea de las coordenadas paralelas es el uso de ejes desplegados paralelamente en lugar de ortogonalmente. Los cuatro puntos de la izquierda están representados por las cuatro líneas de la derecha (Siirtola, Laivo, Heimonen, & Raiha, 2009). ..... 179
- Fig. 54. Consultas de selección en coordenadas paralelas (Heer & Shneiderman, 2012). ..... 180
- Fig. 55. (a) Muestra un esquema en el que toda la información se muestra al mismo tiempo, lo que complica en demasía el proceso de abstracción. (b) Muestra un enfoque diferente, en el que el usuario sigue un flujo de trabajo ordenado para ir conociendo parte del sistema modelado. .... 202
- Fig. 56. Modelo teórico de la Visualización Analítica en *eLearning* (Gómez Aguilar, García Peñalvo, & Therón, 2014). ..... 207

Fig. 57. Proceso de Analítica visual en <i>eLearning</i> (Gómez Aguilar et al., 2014). .....	209
Fig. 58. Vista general del prototipo de VeLA.....	213
Fig. 59. Esquema de flujo de trabajo del prototipo de VeLA.....	213
Fig. 60. Vista global del campus completo de una universidad representado con una red social, en esta se muestran los cursos en color naranja, sus foros en color marrón y sus discusiones en color verde únicamente. 214	214
Fig. 61. Vista global del campus completo de una universidad representado con una red social, en esta se muestran los cursos en color naranja y sus foros en color marrón únicamente. ....	215
Fig. 62. Vista Global del campus completo de una universidad en coordenadas paralelas. En las columnas, de izquierda a derecha, se expresan los cursos, sus foros, sus discusiones, sus <i>posts</i> , quién escribió el <i>post</i> , y quién contestó. <a href="http://analiticavisual.es/video-gallery/item/12-vela-vimoodle-sep-2013">http://analiticavisual.es/video- gallery/item/12-vela-vimoodle-sep-2013</a> . ....	216
Fig. 63. Vista del subconjunto de las actividades realizadas durante todo el curso de un estudiante con una nota de más de 8. ....	217
Fig. 64. Elementos representados en la visualización de línea de tiempo en espiral. En esta se aprecian las tres representaciones del tiempo: en la esquina izquierda de abajo la espiral, a la derecha la línea de tiempo en vertical del contexto y arriba la línea de tiempo horizontal de los datos seleccionados. Además, se aprecian, indicados con líneas puntadas, los tres elementos gráficos que esta visualización aporta. ....	220
Fig. 65. Visión global de la línea de tiempo en espiral. Esta visualización muestra el aspecto general del patrón temporal del curso expresando el patrón semanal. Cuando se coloca sobre cualquier representación de una actividad, esta se resalta en blanco. Ténganse en cuenta los círculos (A, B y C) colocados sobre la imagen, que se han puesto para indicar cuál es la última actividad que se puede ver en la representación espiral sin deformación. Todos los círculos ponen de relieve el mismo momento de tiempo, pero se puede ver en la representación lineal en la parte superior (ver el círculo B de ese punto) que con la espiral solo puede verse la mitad de los datos, de modo que la deformación de la espiral es una técnica útil. ....	221
Fig. 66. Distorsión de la espiral, las barras rojas son las actividades, los polígonos grises representan los períodos de inactividad, que forman un polígono envolvente para cada vuelta de la espiral de 360° y las barras amarillas son las actividades del siguiente turno. Por último, la distancia total del espacio donde se despliegan las barras es <b>T</b> , y este está formado por la distancia de la barra más alta, denominada <b>t</b> más una sangría <b>s</b> . ....	224
Fig. 67. Visión global de la representación de la línea de tiempo, esta visualización muestra el aspecto general del patrón temporal semanal del curso ya deformado en la espiral. Obsérvese la tabla de la derecha que muestra la agrupación de las actividades seleccionadas, donde se han excluido todas las actividades sin categoría agrupadas bajo la etiqueta <i>uncategorized</i> . ....	225

Fig. 68. Representaciones de las coordenadas paralelas y la nube de palabras/etiquetas. En esta imagen se puede observar la comparación de las nubes de palabras de dos cursos diferentes, en la derecha están representadas más de cuatro mil palabras y en la izquierda casi cinco mil palabras.....	227
Fig. 69. En esta imagen se puede notar que la etiqueta “práctica” fue escrita al inicio del curso (expresado por la onda en azul sobre la palabra) y, sin embargo, fue leída hacia el final del curso (expresado por la onda en naranja).....	228
Fig. 70. En esta figura se pueden observar las etiquetas informativas sobre el período que corresponde únicamente a la letra sobre la que el ratón se encuentra y la tecla Ctrl presionada. ....	229
Fig. 71. Codificación gráfica de los nodos que representan los elementos de la red social (Gómez Aguilar, Theron, & García Peñalvo, 2008).....	230
Fig. 72. Representación de la red social. A la izquierda se representa un curso completo y se muestran todos sus elementos sin activar los enlaces internos y externos (procesándose junto con los elementos invisibles más de novecientos elementos entre enlaces y nodos) y a la derecha un curso en el cual se han ocultado todos los elementos con excepción de las discusiones y los participantes, con el fin de resaltar una métrica de LA (cuanta más participación en discusiones, mejor rendimiento) por medio de los enlaces internos y externos (procesándose junto con los elementos invisibles más de dos mil elementos entre enlaces y nodos). Dado que las distancias en representaciones que no son distribuidos los elementos en una línea, sino depende de las fuerzas de atracción y repulsión, se pueden desplegar, o no, los círculos concéntricos que se ven en la figura, siendo estos una ayuda visual que permite la comparación de las distancias entre dos nodos. ....	231
Fig. 73. Representación de la red social, dotada de funcionalidad de búsqueda, un cuadro de texto en la base de la visualización. En esta representación se muestran dos cursos diferentes, ocultando todos los nodos con excepción de los participantes y mostrando, por medio de las líneas, las relaciones de comunicación entre los participantes. La distancia de estos representa la frecuencia de participación entre ellos. ....	232
Fig. 74. Visualización global de las relaciones entre los usuarios. ....	242
Fig. 75. Filtro de Actividad (izquierda) y el filtro de la Palabra (derecha)...	243
Fig. 76. Francisco (izquierda) y Patricia (derecha) redes sociales con filtro de conectividad "6".....	244
Fig. 77. Esta figura muestra el subconjunto de los datos de un curso resaltado por medio de la interacción de la caja de texto de búsqueda y de selección, localizados en el panel central inferior.....	245
Fig. 78. Se muestra solo el periodo de tiempo seleccionado a la derecha de la visualización de SST. ....	247
Fig. 79. Patrón temporal de las actividades semanales de un curso.....	248
Fig. 80. La figura muestra el patrón de la actividad semanal de una persona. ....	249

Fig. 81. Se muestra el patrón de actividad semanal de la acción “agregar post” en la plataforma.....	250
Fig. 82. Un ejemplo de la utilidad del uso de ambas representaciones de línea de tiempo, tanto la espiral como la lineal.....	251
Fig. 83. La figura representa el patrón temporal semanal de la actividad “view”.....	251
Fig. 84. Se representa el patrón temporal diario de las actividades de un curso específico.....	252
Fig. 85. Se representa el patrón temporal diario de las actividades de una persona específica.....	253
Fig. 86. Vista global de la red social. Esta visualización muestra el aspecto general de la estructura del curso. Obsérvese las líneas de conversación, formadas por un grupo de <i>post</i> en un foro, también algunas de ellas están relacionadas con sus discusiones, las cuales a su vez están relacionadas con su foro. Además, nótese el grupo de triángulos azules que son recursos de la asignatura, que están alrededor de la gente en el curso. ....	254
Fig. 87. Vista de la red social filtrada previamente con el fin de mostrar solo la relación entre la lectura de los recursos, el número de personas con quien se relaciona y gente que lo hizo. Es importante mencionar que el recuadro rojo es un modo de interacción para la selección de un grupo de personas a analizar.....	255
Fig. 88. Vista de la red social filtrada previamente con el fin de mostrar solo la relación entre la lectura de los debates, los mensajes del foro, el número de personas con quien se relaciona y gente que lo hizo. Obsérvese también el marco rojo, que es un modo de interacción para la selección de un grupo de personas para su posterior análisis.....	255
Fig. 89. Vista de la red social que se filtró previamente con el fin de mostrar la relación entre la lectura de los recursos, la lectura de las discusiones, los mensajes en los foros, el número de personas con las que se relaciona cada participante y las personas que lo hicieron. Obsérvese también que existe un cuadro rojo, este es un modo de interacción para la selección de un grupo de personas para analizar posteriormente...	256
Fig. 90. Distribución de las agrupaciones D:1, D:2, D:3; R:1, R:2, R:3 y DR:1, DR:2, DR:3, correspondientes a los tres grupos de personas seleccionados de los tres casos a analizar con la finalidad de comprobar la hipótesis. ....	257
Fig. 91. Vista global de las coordenadas paralelas, donde se percibe por medio de las diferentes oscuridades de las figuras formadas por las líneas que se entrecruzan y correlacionan de eje en eje, información a primera vista. Por ejemplo se puede observar a la derecha del eje en rojo en la parte central de la imagen, cuatro bifurcaciones más oscuras que el resto, estas se corresponden a las actividades más realizadas (“view resource”, “view assignment”, “view quiz”, “view course”). ....	260
Fig. 92. Vista del subconjunto de las actividades comunes con la interacción en foros como: <i>search forum</i> , <i>update post forum</i> , <i>add post forum</i> , etc.....	261

Fig. 93. Vista del subconjunto de las actividades realizadas durante todo el curso de un estudiante con una nota de 5.2. ....	261
Fig. 94. Vista del subconjunto de las actividades realizadas durante todo el curso de un estudiante con una nota de 6.5. ....	262
Fig. 95. Vista del subconjunto de las actividades realizadas durante todo el curso de un estudiante con una nota de más de 8. ....	262
Fig. 96. Vista del subconjunto de las actividades de " <i>view resource</i> " realizadas durante todo el curso de un estudiante con una nota de 5.2. ....	263
Fig. 97. Vista del subconjunto de las actividades de " <i>view resource</i> " realizadas durante todo el curso de un estudiante con una nota de 6.5. ....	263
Fig. 98. Vista del subconjunto de las actividades de " <i>view resource</i> " realizadas durante todo el curso de un estudiante con una nota de 9.6. ....	264
Fig. 99. Se representan los días de la semana codificados por número (1 es el lunes 2 es martes, y así sucesivamente) en el eje vertical, y en el eje horizontal, son los diferentes grados de desempeño de los estudiantes en los cursos. ....	265
Fig. 100. Vista del conjunto de 3 cursos, que se filtró previamente con el fin de mostrar la relación entre la agrupación de las actividades y el desempeño de los estudiantes de los datos. El eje vertical es el porcentaje de actividades que fue hecho en total y en el horizontal es la agrupación de las actividades para cada grado de desempeño de los estudiantes para cada curso. ....	267
Fig. 101. Vista de la SST que se filtró previamente con el fin de mostrar solo las actividades de los estudiantes con nota 9 y sus actividades pasivas de la agrupación por propósito. Obsérvese también que la zona de color naranja en una de las secciones de la representación deformada (que representa martes) es el más grande, por tanto, el día más activo. Los datos de esta imagen son del Curso 1. ( <a href="http://analiticavisual.es/images/dialgoag/TEEM2013/fig4-ERM-9-Tuesday-Passive.png">http://analiticavisual.es/images/dialgoag/TEEM2013/fig4-ERM-9-Tuesday-Passive.png</a> ). ....	267
Fig. 102. Promedio de las Encuestas realizadas de los ciclos de investigación acción. Los valores van de 1 al 5, donde el 1 una respuesta "muy mal" y el 5 como "excelente". ....	269



## Índice de Tablas

Tabla 1. Preguntas cuyas respuestas son fuentes de información y conocimiento, mediante un análisis direccional (Thomas H. Davenport et al., 2010).....	70
Tabla 2. Definiciones conceptuales y funcionales de diversos tipos de analítica adaptado de (Van Barneveld et al., 2012).....	74
Tabla 3. Analítica del aprendizaje y analítica académica - Enfoque de las actividades analíticas y de los beneficiarios. Modificado de (Long & Siemens, 2011). ....	91
Tabla 4. Modificado de las taxonomías de importancia sobre las técnicas de interacción (Yi et al., 2007).....	115
Tabla 5. Tabla comparativa de los tipos de datos educativos que se utilizan en las herramientas visuales que soportan la mejora y/o la comprensión del proceso de aprendizaje. ....	188
Tabla 6. Tabla comparativa de las técnicas de analítica educativa que se utilizan en las herramientas visuales que soportan la mejora y/o la comprensión del proceso de aprendizaje.....	190
Tabla 7. Tabla comparativa de las técnicas de LA que se encontraron en los trabajos no visuales indicadas con ✓ e indicando con doble símbolo (✓★) las que se comprobó que existe una relación con el rendimiento del estudiante. Los trabajos marcados con un asterisco * al inicio de la fila, se refieren a que han categorizado las actividades pero solo aquellas, indicadas en la tabla, que han estudiado. En el trabajo de (Giovannella et al., 2013), marcado al inicio de la fila con @, es de relevancia mencionar que aunque se marcan estas técnicas de búsqueda de correlación, en este trabajo se analizan no solo las del LMS sino también las de otras herramientas sociales como <i>twitter</i> y <i>Facebook</i> . La caracterización de las actividades se realiza en un número mayor a 30 actividades categorizando les desde diferentes enfoques en el trabajo de (Agudo-Peregrina et al., 2014), marcado con ** en el inicio de la fila. ....	192
Tabla 8. Tabla comparativa de las representaciones que se utilizan en las herramientas visuales que soportan la mejora y/o la comprensión del proceso de aprendizaje.....	194
Tabla 9. Tabla comparativa de las técnicas de analítica visual que se utilizan en las herramientas que soportan la mejora y/o la comprensión del proceso de aprendizaje por medio de representaciones visuales. ....	196
Tabla 10. Tabla comparativa de la forma de obtención de los datos en las herramientas y frecuencia simultanea de: tipo de datos, representación visual, número de técnicas de analítica del aprendizaje implementadas y por último, el número de técnicas de visualización analítica implementadas en las herramientas. Coloreadas desde el color rojo para un valor menor pasando por el amarillo hasta el verde para un valor mayor con el objetivo de ayudar a su lectura. ....	197
Tabla 11 Cuadro general del proceso de diseño del prototipo del modelo VeLA y las principales técnicas implementadas en cada versión. En esta	

tabla se puede observar el objetivo a lograr en cada mejora del prototipo y el aumento de técnicas de visualización en cada una de las diferentes versiones (léase la última columna de la derecha de arriba hacia abajo, es importante mencionar que solo se ha escrito lo que se ha incrementado o eliminado, el resto de características continúan igual que en la versión anterior) y en cada una de estas se ha indicado el nombre de la o las técnicas que se han implementado específicamente para la visualización por medio de superíndices, por ejemplo *zoom* semántico en la nube de palabras estaría como nube de palabras<sup>zs</sup>. ...235

Tabla 12 Cuadro general de iteraciones de diseño del prototipo del modelo VeLA y los principales resultados de la evaluación.....237

# Capítulo I Introducción

*“Education, beyond all other devices of human origin, is the great equalizer of the conditions of men, the balance-wheel of the social machinery.”*

— *Horace Mann, 19th-century American*

*educator*



La presente tesis doctoral aborda el análisis visual de la información que los estudiantes y otras personas involucradas en el proceso de aprendizaje generan en las plataformas de *eLearning*, con el objetivo de comprender y mejorar tal proceso. En este capítulo se presenta, a modo de introducción al trabajo de investigación que se ha llevado a cabo, la motivación; la hipótesis y los objetivos; la metodología de trabajo; las principales aportaciones y, por último, la estructura de este documento.

Así, en la primera sección, la motivación, se plantea la justificación de este trabajo en sus aspectos principales. En la sección segunda se presenta la problemática que da lugar al desarrollo del mismo, para exponer las limitaciones actuales de la enseñanza y delimitar los aspectos que aborda esta investigación. En la tercera sección se describen los objetivos principales y las preguntas de investigación planteadas al inicio de este trabajo. En la sección cuatro se presenta la metodología usada. En la quinta sección se resumen las líneas de investigación desarrolladas en esta tesis doctoral. Por último, en la sección seis se describe la organización de los restantes capítulos.

## I.1 Motivación

En la última década la tecnología y las telecomunicaciones en todas sus formas han estado cambiando la forma de vivir, de trabajar, de producir, de comunicarnos, de convivir, etc. Todos estos cambios han influido en todos los aspectos de la vida cotidiana. Ante toda esta revolución la educación tiene un reto importante en el cual debe cuestionarse como replantearse sus principios y objetivos, reinventar sus metodologías docentes, como por ejemplo la interdisciplinariedad de la educación avanzada y la utilización de las nuevas tecnologías de manera apropiada en el proceso de aprendizaje. Aprender es la fuente de riqueza y de bienestar más importante, es el punto de partida para obtener la capacidad de competencia y de cooperación.

Hoy en día, y como siempre, el esfuerzo debe dedicarse al diseño de instituciones capaces de evolucionar con el objetivo de adaptar sus medios a las nuevas necesidades individuales y de la sociedad en general. Con el apoyo de las tecnologías de la información y la comunicación se puede ofrecer educación de calidad a un mayor segmento de personas, educación que la Institución debe estar preparada a ofrecer, con un equipo humano, físico, técnico y tecnológico capaz de afrontar los nuevos retos. Hay siempre diferentes formas de conocimiento y cultura que intervienen en la edificación de las sociedades, comprendidas entre estas las influidas con fortaleza por el progreso científico y técnico. No se puede admitir que la revolución de las Tecnologías de la Información y la Comunicación (del inglés, *Information and Communications Technology*, ICT ) conduzca a prever una forma única de sociedad posible (Afgan & Carvalho, 2010).

“El conocimiento se asienta en dos pilares fundamentales, uno son los procesos de *Investigación, Desarrollo e Innovación (I + D + i)*, que crea nuevos conocimientos y mejoras en otros ya establecidos, y el otro es la *Educación*, que transmite los conocimientos existentes. La educación y la calidad de la misma es pieza crítica y clave en la evolución y desarrollo de toda sociedad. Hoy como ayer, el dominio del conocimiento puede ir acompañado de un cúmulo importante de desigualdades, exclusiones y luchas sociales. Durante mucho tiempo el conocimiento fue acaparado por círculos de sabios o iniciados, donde el principio rector de estos círculos era el secreto. Desde el Siglo de las Luces, los progresos de la exigencia democrática –basada en un principio de apertura y en la lenta aparición de un ámbito público del conocimiento– permitieron la difusión de las ideas de universalidad, libertad e igualdad. Esta evolución histórica fue unida a la propagación de conocimientos por intermedio del libro, y luego de la imprenta, y también a la difusión de una educación para todos en la escuela y la universidad” (Bindé, 2005; Mateo, 2006).

Actualmente la sociedad encontró en las tecnologías de la información y la telecomunicación una manera diferente y rápida de comunicarse, de transportar información, de adquirir conocimientos, de intercambiar productos y de acceder a diversas fuentes de conocimientos disponibles. Cambiando, por tanto, la manera no solo de compartir y socializar, sino de aprender colocando conexiones directas entre profesor y estudiante ocultando edades y presentándolos como homólogos, por lo menos virtualmente. Este hecho de trascendencia inconmensurable traspasó la era

de la revolución industrial para colocarnos en la era de las telecomunicaciones mucho más allá de la revolución de la información. “La noción de sociedad de la información se basa en los progresos tecnológicos. En cambio, el concepto de sociedades del conocimiento comprende dimensiones sociales, éticas y políticas mucho más vastas. El hecho de referirse a sociedades, en plural, no se debe al azar, sino a la intención de rechazar la unicidad de un modelo “listo para su uso” que no tenga suficientemente en cuenta la diversidad cultural y lingüística, único elemento que permite a cualquiera reconocerse en los cambios que se están produciendo actualmente” (I. M. d. I. UNESCO, 2005).

A este respecto, uno puede preguntarse si se poseen ya los medios que permitan un acceso igual y universal al conocimiento, así como un auténtico aprovechamiento compartido de este. Actualmente, la difusión de las tecnologías y la aparición de la red pública Internet parecen abrir nuevas perspectivas a la ampliación del espacio público del conocimiento. El nacimiento de una sociedad mundial de la información como consecuencia de esta revolución no debe hacer perder de vista que se trata solo de un instrumento para la realización de auténticas sociedades del conocimiento. El desarrollo de las redes no puede por sí solo sentar las bases de la sociedad del conocimiento (Mansell, 1999; Mansell & Wehn, 1998).

La información es efectivamente un instrumento del conocimiento, pero no es el conocimiento en sí. Educar en la sociedad de la información, concepto que evolucionó a las sociedades del conocimiento, plantea, para todos los que viven en ella, múltiples desafíos. El conocimiento crece, se modifica y se hace más complejo; las fuentes del saber se multiplican y diversifican. Enseñar y aprender están lejos de ser una simple transferencia o transmisión de saberes.

Hoy en día, uno de las formas más comunes de archivar/almacenar la información no solo educativa sino personal es en formato digital. Esto crea un vasto acervo de conocimientos que solo se puede acceder, por razones obvias, a través de herramientas digitales. Los usuarios de esta información tendrán la fluidez en las herramientas de acceso digital, exploración, visualización, análisis y colaboración. Estas herramientas influyen tanto en la creación y el análisis de la información.

Los analistas basan su proceso de producción de los datos, hipótesis y las preguntas a resolver dependiendo de las posibilidades que sus herramientas les ofrecen. El almacenamiento digital y la interconectividad de la Web ofrecen nuevos retos en términos de cantidad y calidad de la información. Crean un nuevo medio para la presentación, así como una base para la colaboración, que es independiente de la ubicación física.

La información en forma digital ofrece oportunidad sin igual para combinar, extraer, presente y compartir ideas complejas. Existen grandes retos en la era del conocimiento: el desarrollo de nuevos esquemas y paradigmas para la representación de la compleja y vasta información actual que puede ser compartida, analizada y comparada; la creación de una alfabetización en el análisis de información y visualización que tiene el mismo rigor y la riqueza como la actual; y utilizar eficazmente la inmensa cantidad de datos obtenidos, computados y almacenados por los sistemas de información actuales, con el propósito de obtener nuevo conocimiento.

Este panorama es común a prácticamente todas las actividades que desarrollan las personas en sociedad, y los procesos de aprendizaje están cada vez más inmersos en esta vorágine generadora de datos. En este contexto, el razonamiento analítico facilitado por las interfaces visuales interactivas y el uso de métodos computacionales adecuados se revelan como herramientas indispensables.

La constante actualización de los conocimientos es un requisito de la sociedad en la que se está inmerso, el uso del Internet para la capacitación de las personas no es una moda, sino una herramienta que facilita esta constante necesidad.

Así, un entorno virtual de aprendizaje puede incluir cualquier tipo de herramienta que se utilice durante el proceso de aprendizaje, por ejemplo vídeos de *YouTube*, artículos de la *Wikipedia*, repositorios de contenidos de una institución, comunidades virtuales de grupos de estudiantes, etc. Esto supone que todo sistema de administración de contenidos de aprendizaje (del inglés, *Learning Content Management System*, LCMS) o sistema de administración de cursos en línea (del inglés, *Learning Management System*, LMS) siempre sea un Entorno de Aprendizaje Virtual (del inglés, *Virtual Learning Environment*, EVA o VLE); sin embargo, no todo VLE es un LMS (García Peñalvo & García Carrasco, 2002). Un LMS esencialmente ayuda a gestionar las actividades y competencias de aprendizaje de una organización. Las actividades gestionadas por el LMS podrían ser diferentes a las de un aula de formación dirigida por un instructor de seminarios educativos para la formación en línea basado en la Web. Este último es aquel en la que la carga lectiva de la asignatura o curso o parte de este se desarrolla completamente en un formato asíncrono, esto es, sin coincidencia espacial ni temporal, por medio de un entorno de interacción y comunicación virtual ubicado en Internet y que recibe el nombre de entorno virtual de aprendizaje o campus virtual, y estos pueden ser una página web estática asociada a una asignatura. Por el contrario, y desde el punto de vista del usuario final, un LMS proporciona una forma eficaz de hacer un seguimiento de las habilidades y competencias individuales, y proporciona un medio de fácil localización y registro de las actividades de aprendizaje relevantes para mejorar aún más los niveles de habilidad del estudiante.

Un LMS también proporciona acceso a cursos en línea en los que el usuario puede registrarse. Administrativamente, un LMS hace que sea fácil entrar, rastrear, administrar e informar sobre las actividades y competencias de aprendizaje en una organización. En esencia, un LMS se centra principalmente en las competencias, actividades de aprendizaje y la logística de la entrega de las actividades de aprendizaje. Un LMS no se centra en la creación, reutilización, gestión, o la mejora de los contenidos en sí. Los Sistemas de Administración de Cursos (del inglés, *Course Management Systems*, CMS) son un tipo de *software* dedicado principalmente a la gestión de contenidos web ya sea en Internet o Intranet. Sin embargo, en el ámbito de la formación se especializa para facilitar la gestión de contenidos educativos, documentos, resultados, etc. El objetivo de los CMS en este ámbito es, por tanto, la creación y gestión de información en línea de diferente tipo, sin incluir actividades ni herramientas que fomenten la colaboración de los usuarios.

En contraste, un LCMS ayuda a crear, reutilizar, localizar, entregar, administrar y mejorar el contenido de aprendizaje. El contenido se mantiene típicamente en un repositorio de contenidos centralizado en forma de objetos pequeños que se auto-describen y son unívocamente identificables. Estos pequeños objetos, Objeto de Aprendizaje (OA), son una mínima expresión de contenido formativo con entidad por sí mismo, etiquetado con metadatos para permitir su búsqueda, descripción, flujo, contenido y recuperación, y que tiene como características principales el ser de formato digital, tener un propósito pedagógico, contener interacción, ser independiente e indivisible y por último reutilizable.

Dadas a estas características los OA puede ser enlazados a otros para crear unidades de instrucción de mayor tamaño y con un flujo determinado y nivel pedagógico diferente, donde cada uno de los cuales satisface uno o más los objetivos de aprendizaje definidos. Cada objeto de aprendizaje puede haber sido creado desde cero o mediante la reutilización de documentos de conocimiento existentes en otros formatos. (Morales Morgado, 2010; Morales Morgado, García Pealvo, & Barrón, 2007; Morales Morgado, Gómez Aguilar, García Peñalvo, & Therón Sánchez, 2009).

Un LCMS puede encontrar y entregar un objeto de aprendizaje para el usuario final como una unidad individual para satisfacer una necesidad específica de la tarea o entregar el objeto de aprendizaje como parte de un curso, programa o actividad de aprendizaje más amplio definido en un LMS. Un LCMS avanzado rastrea las interacciones del usuario con cada objeto de aprendizaje y utiliza esta información detallada para ofrecer experiencias de aprendizaje altamente personalizadas al tiempo que proporciona autores con informes completos para el análisis de la claridad, relevancia y efectividad de los contenidos, por lo que se puede mejorar de manera continua. Un LCMS se centra esencialmente en la creación, la reutilización, la localización, la entrega, la gestión y la mejora de contenidos (Süral, 2010).

La facilidad de acceso a la educación que abren estas herramientas se ha reflejado en el aumento del uso de los LMS y LCMS (Francisco José García Peñalvo, 2008; Rengarajan, 2001). Este tipo de ambientes de formación, usan grandes cantidades de datos, presentan una interactividad compleja, amplia cobertura y una ilimitada restricción de espacio y tiempo. Surge entonces la interrogante ¿cómo analizar dicha información sacando el mayor provecho de esta para mejorar el proceso educativo?

Esta gran cantidad de información presenta, para los diseñadores de elementos de aprendizaje, profesores, gestores, tutores y/o usuarios, el desafío de cómo elegir, evaluar, analizar y obtener nueva información útil, sin olvidar que, en la mayoría de los casos, la información se encuentra oculta detrás de la plataforma de administración de cursos.

Actualmente se presentan tres retos generales en el uso de la tecnología en la educación:

- 1) La integración de los diversos sistemas *eLearning* (Conde González, 2012; Conde González, García Peñalvo, Casany Guerrero, & Alier Forment, 2009; Luo, Li, Cao, & Ge, 2006);

- 2) La estandarización de los objetos de aprendizaje (Anane, Crowther, Beadle, & Theodoropoulos, 2004);
- 3) El análisis de la información generada en un LMS, como la problemática de evaluación del estudiante e identificación de contenido por parte del docente frente a un LCMS (Drigas & Vrettaros, 2006; Mallinson & Sewry, 2004; Zenha-Rela & Carvalho, 2006).

Consecuentemente, existen propuestas en los distintos campos de la teoría educativa: los factores que influyen en la adopción de herramientas de analítica (Ali, Asadi, Gašević, Jovanović, & Hatala, 2013), aplicación de sistemas semánticos en la educación (Guo & Chen, 2006), estudios de satisfacción y actividad en relación con la heterogeneidad de los estudiantes (Ke & Kwak, 2013), influencia del tamaño de los grupos de clase en la participación estudiantil online (Kim, 2013), estudios de predicción del éxito estudiantil (Kruger-Ross & Waters, 2013), integración de sistemas de aprendizaje en línea heterogéneos (Luo et al., 2006), así como estándares (Rosenberg, 2001) y ontologías en apoyo a la creación, organización y clasificación de los objetos de aprendizaje (Guo & Chen, 2006; Morales Morgado et al., 2009; Wiley, 2002), además de la estructuración y clasificación de contenidos educativos en los LCMS y/o LMS (Feipeng, Yanyan, & Zhiqiang, 2013; Milikic, Krcadinac, Jovanovic, Brankov, & Keca, 2013; Snodin, 2013). También las estructuras sociales que se forman entre los participantes han sido un foco de interés para la realización de investigaciones sobre patrones en la interacción en las redes de aprendizaje y la visualización de la estructura de estas (Bakharia & Dawson, 2011; Dradilova, Martinovic, Slaninova, & Snasel, 2008; Ferguson & Buckingham Shum, 2012a; Fournier, Kop, & Sitlia, 2011; Hou & Wu, 2011; Kim, 2013; Reis Figueira & Bras Laranjeiro, 2007).

Además, se han estudiado las estructuras narrativas y del estilo de aprendizaje de los estudiantes en los sistemas de *eLearning* (Leighton Álvarez, Prieto Ferraro, & García Peñalvo, 2012; Prieto Ferraro, 2006; Williams & Conlan, 2007). Recientemente se han aplicado algunas técnicas de visualización, minería de datos (E. García, Romero, Ventura, de Castro, & Calders, 2010) y de analítica visual (del inglés, *Visual Analytics*, VA) a los datos de las plataformas *eLearning* que se describirán en la sección III.3 y las secciones IV.3, IV.4 y IV.5.

Actualmente, con la introducción de las ICT en la educación, se ha fomentado el uso de la tecnología con fines de apoyo al proceso de aprendizaje y enseñanza. Este fenómeno, junto con la aplicación de nuevas técnicas de procesamiento de los datos, que se han ido adaptando desde otros campos (por ejemplo el empresarial), han propiciado la necesidad de un análisis de datos avanzado en la educación, dando origen al término analítica académica (Philip J. Goldstein & Richard N. Katz, 2005).

La Analítica del Aprendizaje o Analítica Educativa (del inglés, *Learning Analytics*, LA) responde a la necesidad de llevar a cabo el seguimiento y control de la actividad en el campus para la toma de decisiones estratégicas, mediante el aprovechamiento de la gran cantidad de datos producidos por los estudiantes en actividades académicas (Fournier et al., 2011).

La mejora de la educación apoyada en las ICT ha centrado sus objetivos en la progresión académica de los estudiantes, esto es, predecir su comportamiento e identificar elementos problemáticos. La analítica del aprendizaje tiene como principio el análisis, estudio, cuestionamiento e interpretación de un amplio rango de datos producidos y recogidos acerca de los estudiantes con el fin de llegar a estos objetivos (Fournier et al., 2011).

Una vez recolectados, registrados, analizados y presentados, se busca permitir a los profesores la adaptación de manera eficaz y rápida de las estrategias educativas a la vez que se tienen en cuenta las necesidades y capacidades de cada estudiante.

La información aportada por la analítica del aprendizaje permite personalizar la acción formativa y diseñar entornos de aprendizaje acordes con las necesidades, intereses y formas de interacción entre profesores y estudiantes. El registro de los procesos de aprendizaje posibilita que los docentes no centren toda la atención en los materiales y la enfoquen en el diseño y análisis de los procesos formativos. Por otro lado, el registro estadístico de la actividad de los estudiantes y docentes permite identificar puntos conflictivos dentro del proceso de enseñanza-aprendizaje y asistir a su optimización continua (Van Barneveld, Arnold, & Campbell, 2012). Por tanto, la gestión de la información es de suma relevancia para la docencia, el aprendizaje y la investigación. En este sentido, el aporte teórico de este trabajo es un modelo que soporta el proceso educativo mediante el uso de técnicas de analítica visual y diversas técnicas de analítica del aprendizaje y analítica académica, que al ser utilizado por los diferentes agentes del proceso, permita la mejora del proceso formativo. Además, el alcance de aplicación de este modelo está justificado por distintos factores que se describen a continuación.

El índice de educación es uno de los componentes del Índice de Desarrollo Humano anual de las Naciones Unidas, el cual se calcula mediante la combinación de las tasas de alfabetización con la tasa bruta de participación en la educación primaria, secundaria y terciaria (ONU, 2013). Este índice representa el hecho de que la necesidad de educación en general, y educación superior en particular, es un problema mundial cada vez mayor, sobre todo para los países en vía de desarrollo donde el índice de educación es muy bajo. Existen diversas razones por las cuales un estudiante no accede a la educación, a continuación se mencionan algunos ejemplos: la pobreza económica que conlleva deficiencia de infraestructura necesaria en los procesos de formación (África) (I. f. S. UNESCO, 2010), los factores culturales como extracción social, machismo, determinismo social que generan exclusión (Instituto de Estadística, 2010). Se suman a estos factores: los desastres naturales, conflictos políticos, o revoluciones sociales, que afectan severamente a los colegios y universidades, que dejan pocas o ninguna opción para los estudiantes que desean continuar su educación.

Por otra parte, los costes elevados a consecuencia de políticas neoliberales, partidarias de la sustracción de la responsabilidad del Estado en la financiación de instituciones educativas de carácter público (Lynch, Engle, & L. Cruz, 2011), crean cada vez más unas circunstancias difíciles para costearse la educación.

Estos factores se han reducido o eliminado, en algunos casos, con la llegada del movimiento de Conocimiento Abierto (J. F. García Peñalvo, C. García de Figuerola, & J. A. Merlo, 2010) y su objetivo de promover y difundir el conocimiento científico. Quizás sea exagerado considerar que para encontrar el primer paso en la evolución histórica de este movimiento haya que retroceder hasta 1667, con la aparición de la Real Sociedad de Londres para la Mejora del Conocimiento Natural (también conocida como Sociedad Real), que es probablemente la más antigua de una serie de instituciones similares basadas en Europa durante los siglos XVII y XVIII. No obstante, en tiempos más cercanos, diferentes iniciativas, todas con semejanzas notables al movimiento del *software* libre, trataron de promover el intercambio de conocimientos en diferentes campos con el menor número posible de restricciones. Así, en 1998 David Wiley fundó el Proyecto *Open Content*, junto con Eric Raymond, Tim O'Reilly y otros, inspirado en los conceptos de *software* de código abierto y de otros elementos, como las licencias de GNU. El Proyecto *Open Content* tenía por objeto el mundo académico y propuso una licencia (la licencia *Open Publication*) que hizo que el proceso de compartir las creaciones intelectuales fuera más fácil. Fue David Wiley quien acuñó el término "contenidos abiertos", aunque está claro que la idea se basa en el movimiento del *Software Libre* que había surgido unos años antes.

En 1999 la Universidad de Rice inició el proyecto *Connexions*<sup>1</sup>, un proyecto para compartir los recursos educativos de forma gratuita y para poner en disposición de cualquier persona bajo las licencias de contenido libre (del inglés, *Open-Content*) y código abierto (del inglés, *Open-source*) (Baraniuk et al., 2002). La UNESCO (2002) acuñó el término Recursos Educativos Abiertos (REA) para referirse a la disposición abierta de los recursos educativos, habilitados por las tecnologías de la información y la comunicación, para la consulta, uso y adaptación por parte de la comunidad de usuarios con fines no comerciales (Fidalgo Blanco, Sein-Echaluce Lacleta, Borrás Gené, & García Peñalvo, 2014; F. J. García Peñalvo, C. García de Figuerola, & J. A. Merlo, 2010; J. F. García Peñalvo et al., 2010; Laal, 2011).

Es evidente que con el acceso abierto a la producción científica se está multiplicando el volumen de la documentación disponible y se están reduciendo los obstáculos temporales y económicos para el acceso a los artículos científicos y otros resultados de la investigación. Con raíces en el movimiento del *software* libre, la filosofía del acceso abierto ha tenido un profundo impacto en la información digital, con un énfasis especial en la educación y la cultura. Así, se produjo un salto cualitativo en lo que se conoce como la ciencia abierta, con un importante apoyo de las instituciones gubernamentales, como la Unión Europea; finalmente, es transferido al sector de la producción de la innovación abierta (Harnad, 2005). Dentro de este movimiento se han realizado de una serie de iniciativas que han tratado de proporcionar la tecnología para apoyar la práctica de la publicación y reutilización: normas y especificaciones de

---

<sup>1</sup> <http://cnx.org>

metadatos, repositorios de objetos de aprendizaje (LORs) (Morales Morgado et al., 2009), especificaciones de envasado de contenidos (Gil et al., 2009; Henze, Dolog, & Nejdil, 2004) y se han desarrollado motores de diseño de aprendizaje del tiempo de ejecución que permiten a los usuarios la reproducción de una secuencia específica a través de conjuntos de recursos, donde para cada unidad de aprendizaje se crea una instancia para crear un plazo para esa unidad de aprendizaje (Anane et al., 2004; Koper, 2006). Así, la infraestructura que la comunidad había construido para apoyar la reutilización puede no haber sido la apropiada para las necesidades de la labor académica diaria, hasta la llegada de las tecnologías de la información, en particular con el uso de los VLE como por ejemplo, LMS, LCMS o CMS, los cuales se describirán detalladamente en el Capítulo II, y que ha ampliado la posibilidad de acceso a la educación (Davis et al., 2010). Actualmente se realizan esfuerzos con éxito para ayudar a llegar a millones de estudiantes a las escuelas o permitirles el acceso a la educación alternativa. Además, está aumentando la conciencia de que la educación es aprender, no solo la escolarización. En algunos casos, esto comienza con un cambio en el diálogo social sobre la educación con la intención de hacer una modificación desde el solo ofrecer el acceso a esta hacia más que solo el acceso, esto es, ofrecer el aprendizaje.

En este contexto, la medición de los resultados del aprendizaje es un paso crucial para asegurar que todos, en todas partes, sean capaces de ejercer su derecho a una educación de calidad y convertirse en un ciudadano global productivo. Para los interesados en la medición y seguimiento adecuado de la educación a nivel global todavía hay mucho trabajo por hacer. El informe del Instituto de Estadística de la UNESCO (LMTF, 2013) ofrece una visión global de cómo se debe medir el aprendizaje en todo el mundo. La educación es responsabilidad de todos, por tanto, la medición de los resultados del aprendizaje debe dirigirse y apoyarse por todos los que tienen interés en asegurar que las personas aprendan. Esto incluye, además del gobierno, a los estudiantes, profesores, padres de familia, la sociedad civil, el sector privado y una variedad de otros actores. Independientemente de la "cultura de la evaluación" en un determinado país, hay pasos que los gestores de la educación pueden dar para mejorar la medición de los resultados del aprendizaje y, en última instancia, para mejorar los niveles de aprendizaje. En este sentido, en el proyecto LMTF (del Inglés, *Learning Metrics Task Force*) del LMTF (2013), se aborda esta cuestión que es crucial: ¿Cómo puede la medición de aprendizaje traducirse en políticas y prácticas que se reflejan en mejoras en el aprendizaje?

Teniendo en cuenta las ideas anteriormente mencionadas, los procesos de aprendizaje a través de la historia se han visto influidos por las diferentes tendencias culturales, políticas, económicas, sociales, etc., y, hoy en día, hay que agregar un nuevo tipo de influencia: "la influencia tecnológica" (García Peñalvo, 2005; García Peñalvo & Seoane Pardo, 2015). Como consecuencia, los procesos de aprendizaje están en continua evolución y estos dependen de las nuevas tendencias sociológicas, pedagógicas y/o tecnológicas (García Peñalvo, 2005; Francisco José García Peñalvo, 2008), apareciendo nuevos enfoques en la universidad hacia la docencia basada en medios electrónicos y sociales (Berlanga, Peñalvo, & Sloep, 2010).

Consecuentemente, la aplicabilidad de la investigación de esta tesis, su proyección en la sociedad, quiénes se benefician de esta, ya sea una organización o un grupo social, se convierte casi en evidente. Este trabajo persigue la propuesta de un modelo que soporta al *eLearning* y aporta mejoras en el análisis académico, la medición y autoevaluación del aprendizaje en línea por parte del estudiante, así como del profesor y gestor, de forma que, se reducen, y en algunos casos se eliminan, las problemáticas antes mencionadas.

La aplicación de dicho modelo conlleva un cambio en la metodología actual de enseñanza y, por tanto, una justificación de esta, la cual se describe a continuación. Con el auge y desarrollo de la Informática y las Telecomunicaciones, los procesos de educación se han visto, en algunos casos, subordinados a los soportes tecnológicos que facilitan el acceso al conocimiento. Antes de los años 70 ya se preveía la utilización de herramientas audiovisuales como medios de aplicación de la tecnología a la educación y desde esa década se conoce el término Educación Basada en el Uso del Ordenador (del inglés, *Computer-Based Training* - CBT), como uno de los primeros conceptos que fundamentan lo que hoy se conoce como *eLearning* (este concepto se desarrollará en la sección II.1).

Los sistemas CBT evolucionaron en los años 80 a través de sistemas tutores inteligentes, los cuales usaban técnicas de Inteligencia Artificial (IA) para representar el conocimiento, análisis y predicción; estos sistemas fueron llamados Entornos de Aprendizaje Inteligentes (del inglés, *Intelligent Learning Environments* - ILEs) (E. García et al., 2010). Poco después en los comienzos de la década de los 90 hubo una auténtica revolución en los CBT, gracias al desarrollo y masificación de Internet (United States, Kerrey, & Isakson, 2000); motivo por el que actualmente no se deben separar los conceptos de *eLearning* de los de *web-based learning*.

No hay que olvidar que un estudiante en un contexto de *eLearning* está actuando en un entorno virtual de aprendizaje y requiere de la posibilidad y la necesidad de representar la dinámica de la interacción del proceso de aprendizaje, que son apenas visibles en su forma orgánica. A diferencia de la educación tradicional cara a cara, todas las acciones o intervenciones que realizan los individuos pueden ser reutilizadas y representadas en la misma plataforma y se quedan registradas en la base de datos, y esos datos son necesarios para entender las relaciones establecidas entre las personas implicadas, tales como el reconocimiento de sus respectivas actuaciones en la plataforma para la mejora del aprendizaje y estrategias de enseñanza.

La incorporación de las ICT en el campo educativo ha permitido rediseñar los escenarios donde se producen los procesos de enseñanza-aprendizaje, lo que ha provocado que sea cada vez menos nítida la frontera entre educación presencial, semipresencial y virtual e igualmente entre educación formal, no formal e informal (conceptos que se ampliarán en el Capítulo II) como menciona (2010); Adell and Castañeda (2013); donde los roles de docente y estudiante deben cambiar como afirma Salinas (2008).

Bromham, Lindell y Oprandi, Paolo (2006), mediante el uso de herramientas de *eLearning*, prueban que se puede mejorar el aprendizaje del estudiante, además del tiempo que el personal pierde en la fase de desarrollo inicial, sin un costo considerable. Demuestra así que un recurso de Internet puede ofrecer una verdadera mejora a un curso universitario,

gracias al aporte de algo que no se entrega fácilmente a través de las clases presenciales: la oportunidad para la auto-evaluación periódica y relevante, dirigida por el profesor y alineada con el plan de estudios.

En el mismo sentido, es evidente también, que el aprendizaje tiene una dimensión individual, ya que cada ser humano es único, con sus propios intereses, limitaciones y capacidades, de forma que requiere que el proceso de asimilación de conocimientos sea personalizado, por tanto, surge la necesidad de propiciar los llamados Entornos Personales de Aprendizaje (del inglés, *Personal Learning Environment*, PLE), que se definen como “el conjunto de herramientas, fuentes de información, conexiones y actividades que cada persona utiliza de forma asidua para aprender” (Adell & Castañeda, 2010).

A pesar de la necesidad individual y personalizada de la educación, en cada individuo se encuentra, en la percepción de los sentidos, una vía genérica o global de transmisión. La comparación de la percepción visual con los demás sistemas sensoriales da como resultado que la visión es el canal sensorial primario en la construcción de representaciones mentales. Según Villalba Simón et al. (2000) esto es así porque "se extiende al ser humano más allá de su propio cuerpo, es el mediador de otras impresiones sensoriales, y actúa como un estabilizador entre la persona y el mundo exterior".

Por tanto, entre todos los estímulos visuales, los que añaden un mayor aporte a los conocimientos y/o modelo mental del estimulado, en un menor período de tiempo, son las representaciones gráficas visuales; si se tiene en cuenta lo antes mencionado se supone que las representaciones visuales son la vía más rápida y efectiva de entendimiento y mejora del proceso y por tal motivo, son de gran importancia en el proceso de aprendizaje y su optimización. Sin embargo, a pesar de la promesa y el entusiasmo por las visualizaciones para el aprendizaje, se sabe relativamente poco acerca de cómo las personas piensan o aprenden con visualizaciones y multimedia. La representación de la realidad mediante tecnología es, en la mayoría de los casos, una determinante en la aceptación y la comprensión de la solución a problemas complejos. La visualización amplía las capacidades cognitivas humanas (J. J. Thomas & Cook, 2006; Villalba Simón et al., 2000).

Como se ha mencionado anteriormente la analítica del aprendizaje y la analítica académica son de vital importancia para la mejora del aprendizaje, sin embargo, esta perspectiva tiene una limitación importante para su aplicación en los LMS: la extracción de información de plataformas de aprendizaje no es una problemática trivial y en la mayoría de los casos significa modificar directamente el código fuente el LMS y/o herramienta de visualización (*hard-coding*). Al realizar *hard-coding* normalmente se establece un vínculo con la plataforma y, peor aún, con una versión específica de esta y/o la de la herramienta, de forma que para versiones futuras de cualquiera de ambas partes peligrará el vínculo establecido y, por tanto, el buen funcionamiento en la comunicación. Siguiendo esta idea y como solución en el ámbito educativo la aplicación de una arquitectura orientada a servicios (del inglés, *Service-Oriented Architecture*, SOA) (Dagger, O'Connor, Lawless, Walsh, & Wade, 2007) es útil para conseguir adaptar los LMS a las tecnologías emergentes, a los nuevos modelos de interacción,

a los *frameworks* o a las especificaciones, para transformar de este modo estos sistemas legados en plataformas de *eLearning* basadas en servicios. Es evidente, por tanto, que aplicar este tipo de arquitecturas posibilita la comunicación con las plataformas de aprendizaje de una manera transparente y segura.

Existen diferentes enfoques como solución a este problema. El presente trabajo de investigación toma en consideración arquitecturas orientadas a servicios (Jeff A. Estefan, Propulsion, Ken Laskey, Francis G. McCabe, & Danny Thornton, 2006; Pätzold, Rathmayer, & Graf, 2008), tal como lo han especificado los trabajos de Conde González (2012); Conde González, Gómez Aguilar, Del Pozo de Dios, and García Pealvo (2011) que proponen un *framework* de servicios de interoperabilidad entre los contextos institucionales y los PLE. Dicho *framework* tiene una estructura básica, formada por una serie de componentes, servicios e interfaces, que facilita el intercambio de interacción e información entre estos dos contextos educativos con dos formas muy diferentes de entender las actividades de aprendizaje. Dicho *framework* se basa en el uso de especificaciones de interoperabilidad y de servicios web. En concreto, utiliza la capa de servicios web de Moodle (M. Á. Conde González et al., 2009) como un *proxy* para el intercambio de información.

## I.2 Problemática

La enseñanza en la educación superior ha sufrido cambios notables en los últimos años, ya que muchos profesores han pasado de la enseñanza tradicional cara a cara a la enseñanza virtual (o aprendizaje virtual), ya sea de una forma completa o combinada con los métodos tradicionales. El *eLearning* (o aprendizaje electrónico) (García Peñalvo, 2005; García Peñalvo & Seoane Pardo, 2015; Francisco José García-Peñalvo, 2008) y el *Blended Learning* (o aprendizaje combinado) nacieron gracias a este proceso.

El contexto también ha cambiado; por un lado permite formar potencialmente a un mucho mayor número de estudiantes, que tienen una mayor diversidad de conocimientos previos, habilidades relevantes y objetivos futuros. La tecnología claramente ha proporcionado algunos beneficios (Suriá, 2010). Por otro lado, los estudiantes de hoy no siempre tienen que estar físicamente presentes en el aula, en lugar de esto, se utiliza la tecnología que les proporciona "en cualquier momento y en cualquier lugar" el acceso a las conferencias grabadas en vídeos, libros electrónicos, o redifusiones multimedia (del inglés, *podcasts*) de audio. A pesar de esto, en el centro de todos estos usos de la tecnología sigue estando el mismo modelo subyacente: el principal modo de transmisión del conocimiento es el de un estudiante sentado y escuchando a un profesor dando una conferencia (Seoane Pardo, 2014; Seoane Pardo & García Peñalvo, 2014; United States et al., 2000).

Desde los inicios de la "educación a distancia", entendiéndose como no más que un circuito cerrado de televisión, se puede resumir que con frecuencia se ha tratado de utilizar la tecnología para replicar, básicamente, el modelo de aula tradicional basado en la clase magistral. Y, por tanto, se puede decir que las *nuevas tecnologías* a menudo se han utilizado para reproducir los sistemas actuales educativos, sin tener mucho en cuenta

cómo el potencial de la tecnología puede ayudar a diseñar un sistema mejor. ¿Qué causas subyacen al fenómeno descrito?, ¿es que el modelo basado en conferencia tradicional funciona tan bien que no hay necesidad de buscar algo mejor?, y ¿se está mejor servido mediante el uso de la tecnología solo para replicar y aumentar este sistema que ha estado en vigor desde hace cientos de años?

El proceso de la educación superior ha funcionado bien en el contexto para el que fue construido cuando se podría asumir que, seguramente, se estaba enseñando a grupos pequeños de estudiantes con una base de conocimientos, habilidades relevantes y objetivos homogéneos.

### **I.2.1 Limitaciones del proceso de enseñanza/aprendizaje**

A continuación se enumeran los problemas más comunes de la educación y se argumenta cómo han ido disminuyendo por medio de la integración de las ICT en el proceso de enseñanza/aprendizaje hasta dejar expuesta la problemática del presente trabajo.

#### **I.2.1.1 Los instructores enseñan a un porcentaje de la clase**

En la enseñanza tradicional con el aumento de la diversidad de la población estudiantil, los instructores se ven obligados a tomar decisiones cada vez más difíciles sobre el nivel al que se dirige su enseñanza. Algunos instructores enseñan a los mejores estudiantes de la clase, mientras el resto de los estudiantes de clase, con frecuencia la mayoría, se encuentran en la circunstancia de mantenerse al día o quedarse rezagado. Otros se esfuerzan para asegurar que se satisfagan las necesidades de todos los estudiantes, que desgraciadamente, puede resultar en una experiencia aburrida, incluso molesta para muchos de los estudiantes. Otros buscan alcanzar un término medio, lo que puede dar lugar a que la mitad de la clase se pierda y la mitad de la clase pierda la atención requerida.

Ahora con la implementación de las ICT y la educación en línea los estudiantes son capaces de moverse a través del material a su propio ritmo. La tecnología también hace que sea fácil proporcionar múltiples niveles de instrucción y diversos caminos a través del mismo material.

Los estudiantes que entran en el curso con conocimientos previos o que aprenden rápido son capaces de lograr un objetivo de aprendizaje determinado y seguir adelante. Los estudiantes que necesitan más tiempo con el material pueden trabajar a un ritmo más lento y son capaces de acceder a explicaciones adicionales de contenidos alternativos, de esta forma se podría seguir dando al estudiante más instrucciones sobre un concepto hasta que este ha alcanzado el objetivo de aprendizaje, y luego permitir que se traslade al siguiente tema. Sin embargo, se cree que no se estaría ayudando a desarrollar las habilidades meta cognitivas necesarias para guiar su propio aprendizaje (Baer & Campbell, 2012; Bailey, Henry, McBride, & Puckett, 2011).

#### **I.2.1.2 Ineficiencia en creación de recursos educativos**

En general, los profesores crean sus propios materiales didácticos para sus clases. Los conocimientos adquiridos por un instructor con experiencia acerca de la mejor manera de enseñar el material en la mayoría de los casos

se pierden en el momento en que el instructor se jubila o se traslada y un instructor nuevo debe comenzar de cero (Baer & Campbell, 2012; Bailey et al., 2011). La tecnología puede ayudar a resolver este problema y maximizar la productividad al permitir crear cursos que satisfagan las necesidades de los estudiantes y profesores a través de muchas instituciones. Esto permite eliminar la redundancia actual en la creación de la instrucción. En lugar de gastar recursos en miles de diferentes cursos de introducción a la Estadística, se cree que tiene más sentido que expertos en el tema a partir de una amplia gama de colegios y universidades creen un pequeño número de los cursos de Estadística que se adapten a las necesidades de los estudiantes en todos los colegios y universidades. En este sentido surgen en escena los denominados Cursos *Online* Masivos Abiertos (del inglés, *Massive Online Open Courses*, MOOCs). Los MOOC son un evento en el que una comunidad masiva de personas, movidas por el interés en aprender sobre un tema determinado e interactuando de manera proactiva, acceden a información relevante, buscan e intercambian nueva información de maneras diversas y, en el proceso, crean conexiones y relaciones entre sí. Sin embargo, cómo determinar o identificar los recursos y/o cursos de calidad con el objetivo de la mejora del proceso de aprendizaje. Un acercamiento es plausible mediante enfoques que emplean una variedad de análisis cualitativos y cuantitativos, incluyendo técnicas de minería de datos, métricas, análisis de contenido, etc. (Clow, 2013; Milligan & Littlejohn, 2013; Rodríguez, 2012).

### **I.2.1.3 Para el instructor el estado del conocimiento previo del estudiante en la materia es una caja negra**

Los instructores pueden tener pequeños indicios de esta caja negra a través de las tareas y exámenes, pero de nuevo, esta información suele llegar demasiado tarde para ser de utilidad, el punto en el que sería más valiosa la retroalimentación es mientras que la instrucción se está dando.

Dependiendo de la habilidad del instructor para interpretar las expresiones faciales de los estudiantes, el docente puede tener poco o ningún conocimiento del progreso de los estudiantes para comprender los conceptos que se presentan. En grandes salas de conferencia, el profesor se ve privado de los beneficios de interpretar las expresiones faciales de los estudiantes, dependiendo de la capacidad del profesor, sobre el progreso del estudiante más allá de las primeras filas.

Actualmente, en algunos VLE se utiliza el modelo de estado de los conocimientos del estudiante representa por medio de una visualización del estilo de cuadro de mandos (*dashboard*) para el instructor. Esta pantalla proporciona una descripción general de cómo los estudiantes de una clase están realizando los objetivos de aprendizaje para cada módulo en el curso. Esto les da a los instructores una vista de lo que siempre ha sido una caja negra. Antes de entrar en clase, los instructores pueden ver rápidamente los conceptos que los estudiantes están asimilando bien y los conceptos con los que tienen problemas. Esto permite a los instructores pasar su tiempo con los estudiantes en una manera optimizada y haciendo efectivo uso de sus conocimientos. En lugar de gastar tiempo valioso de clase repasando conceptos que los estudiantes fueron capaces de aprender fuera de clase, pueden abordar problemas que los estudiantes están teniendo.

También se centrará en los aspectos más ricos de la materia que puede ser que no han tenido tiempo para cubrir en un modelo de enseñanza tradicional.

#### **I.2.1.4 Los estudiantes, con frecuencia, no reciben retroalimentación inmediata**

Con el fin de aprender la materia, la mayoría de los estudiantes escuchan las lecciones en el aula y consultan y estudian los libros de texto; para demostrar su nivel de competencia, realizan comúnmente tareas y exámenes. En el momento en que reciben retroalimentación sobre su trabajo, la oportunidad de corregir cualquier malentendido o reforzar las respuestas correctas, a menudo ha pasado.

La retroalimentación es crucial para el proceso de aprendizaje, la razón principal de que los tutores humanos pueden ser tan eficaces en el trabajo con los estudiantes es que son capaces de proporcionar información específica e inmediata en los puntos adecuados en el proceso de aprendizaje (Bloom, 1984).

Uno de los objetivos en la creación de un curso en línea es que las interacciones que el estudiante tiene con el curso deberían, en la medida de lo posible, ser el modelo de los tipos de interacciones que este tendría con un tutor humano ayudándolo, a trabajar a través de los recursos. En lugar de presentar al estudiante una gran cantidad de contenido para escuchar o leer, como lo haría con una conferencia o con libros de texto tradicionales, se le presenta un concepto con una cantidad limitada de contenido expositivo. Es decir, "aprender haciendo" en las actividades, no se espera a que el estudiante haya dominado el material, de hecho, se espera que cometa errores (Murnane & Phillips, 1981; Thompson, 2010; von Hippel & Tyre, 1995).

Durante el proceso de aprendizaje y creación de actividades, se trabaja para captar los errores más comunes que cometen los estudiantes en el aprendizaje de un concepto dado. Esto se puede hacer a través de conversaciones con los profesores que tienen amplia experiencia en la enseñanza del curso, o mediante el análisis de los artefactos de trabajo de los estudiantes, por ejemplo, tareas, recursos y exámenes; a continuación, se les escribe comentarios dirigidos hacia la corrección de esos errores comunes. Cuando un estudiante elige una respuesta incorrecta, la plataforma explica por qué y corrige esa idea errónea al igual que un tutor humano lo haría si estuviera ayudando al trabajo de los estudiantes a través del material. Cuando se hace correctamente, esto puede ser muy potente.

#### **I.2.2 Retos y necesidades del proceso de enseñanza/aprendizaje**

En todos estos problemas mencionados los educadores han tenido que cambiar las formas tradicionales de obtener "el pulso de la clase" o la realimentación de sus estudiantes (secciones I.2.1.3 y I.2.1.4), que en una educación presencial tenían a mano fácilmente. Además, han aparecido numerosas formas nuevas de aprendizaje y colaboración entre estudiantes y profesores, debido a las grandes posibilidades que abre la Web 2.0 (O'Reilly, 2007). Estos problemas han generado un renovado sentido de responsabilidad de mejorar de manera urgente la educación, la

transparencia y el rendimiento de esta. La incorporación de la tecnología en la educación sigue estando a la vanguardia de la investigación, centrándose en mejorar el éxito del estudiante (NGLC Insights, 2013). Las instituciones están tratando de combinar e integrar tecnologías que apoyan con éxito el aprendizaje del estudiante (Lochner, 2011; I. f. S. UNESCO, 2010). Dentro de las cuestiones en las que se requiere atención se incluyen el aumento de obtención del título y la disminución de la brecha de rendimiento, así como cambiar el enfoque del "acceso al éxito" a cambio del de "las competencias adquiridas". Al igual que con todos los aspectos del aprendizaje, estos desafíos no se pueden resolver con simples cambios; las instituciones deben esforzarse por desarrollar análisis o "inteligencia procesable" en todas las áreas y niveles institucionales, particularmente en el aprendizaje (Van Barneveld et al., 2012).

Por esta razón, los profesores se han visto obligados a buscar formas nuevas de evaluación basadas en el análisis de la información educativa, que sea capaz de tener en cuenta cuánto trabajo han realizado los estudiantes y qué habilidades han adquirido. De la misma manera que han aparecido nuevas formas de comunicación, también deben crearse métodos que analicen y posibiliten en mayor medida la comprensión y el análisis de esa comunicación para continuar mejorando el proceso educativo. Todo esto conlleva que la evaluación de los estudiantes ya no puede basarse tan solo en los resultados de los exámenes finales, sino en un proceso de aprendizaje en el que, además de las competencias académicas, se valoran otros aspectos.

Aunado a estos retos la educación superior tiene acceso a más datos que nunca, el *software* y recursos tecnológicos son el fortalecimiento de la capacidad institucional para acceder a los datos y con estos mejorar la toma de decisiones. Estas herramientas inteligentes se están llevando al aprendizaje adaptativo y oportunidades personalizadas (A. Berlanga & F. J. García Peñalvo, 2005; A. Berlanga & J. F. García Peñalvo, 2005, 2008; Peter Brusilovsky, 2001; Ferraro, Álvarez, & Peñalvo, 2004). De hecho, el análisis de los datos institucionales es clave para transformar la retención de los estudiantes, la graduación y el éxito (LMTF, 2013). Esto ha generado una problemática nueva en donde se debe tomar en cuenta las diferentes perspectivas de análisis de acuerdo a los profesores o estudiantes o a los gestores. Estos diferentes enfoques se comportan como diferentes perfiles de intereses y objetivos, por lo que la evaluación del proceso de enseñanza/aprendizaje se realiza dependiendo desde qué perfil se realice; el tipo de análisis que se lleve a cabo puede ser para evaluar a los profesores, para la evaluación de los estudiantes o para la evaluación de los recursos, de la estructura, la programación y distribución de los cursos. Actualmente, se cuenta con el acceso a datos de análisis y el uso de esos datos para impulsar múltiples bucles de retroalimentación:

- **Evaluación de los estudiantes:** proporciona al estudiante el apoyo puntual y específico en todo el proceso de aprendizaje. Este apoyo es en forma de autorreflexión (Govaerts, Verbert, Duval, & Pardo, 2012) que se adapta al rendimiento actual de la persona y que promueven la revisión y refinamiento. También esta evaluación se presenta al profesor con el objetivo de identificar "estudiantes en riesgo". La riqueza de los datos que se recogen sobre el uso de los

estudiantes y el aprendizaje ofrece una oportunidad sin precedentes para dar a los instructores una imagen clara del estado de conocimiento actual del estudiante.

- **Evaluación de los instructores:** de la misma forma que se puede evaluar a un estudiante por su participación y compromiso con el proceso de aprendizaje, los profesores son capaces, por ejemplo, de auto evaluarse y decidir si planifica más tiempo de interacción con los estudiantes; de manera que se optimice el tiempo y los conocimientos para orientarles hacia los intereses y las necesidades del estudiante. El análisis de los datos a nivel de interacción permite observar cómo los estudiantes utilizan el material del curso y evaluar el impacto de sus patrones de uso en los resultados del aprendizaje. Entonces se es capaz de tomar ventaja de este análisis para refinar de forma interactiva y mejorar el curso para el próximo grupo de estudiantes.
- **Evaluación del proceso y/o la institución:** consecuentemente, con la información que actualmente se puede obtener, se está habilitado el análisis de la institución y del proceso de enseñanza, con métricas como la participación, cantidad de tiempo ocupado, número de recursos multimedia, adscritos al curso, etc. Los gestores pueden tomar decisiones institucionales o referentes al proceso de aprendizaje, como la asignación de recursos tecnológicos o la desaparición del mismo.

Las mediciones de rendimiento en el aprendizaje del estudiante, la progresión y finalización son cada vez más frecuentes, con el ejemplo de las llamadas nacionales, estatales y locales para una mayor compromiso viene la necesidad de instituciones para desarrollar más capacidad de datos y optimizar la retención y finalización de los estudiantes. No en vano se ha identificado el potencial de la tecnología en la educación (Bailey et al. (2011):

*Estamos en los albores de una era en la que los educadores tienen la posibilidad de aprovechar la tecnología para producir un cambio radical en el rendimiento estudiantil. Aunque algunos visionarios han estado prometiendo durante años que la tecnología podría transformar la educación y primaria y secundaria -a pesar de los miles de millones de dólares gastados en la creación de redes de escuelas y equiparlas con los ordenadores y otros dispositivos- las repercusiones concretas sobre los resultados de los estudiantes hasta la fecha ha sido decepcionantes. Sin embargo, cuando la tecnología se introduce estratégicamente en cada paso importante del proceso pedagógico, es, de hecho, cuando tiene el potencial de mejorar todos los aspectos de la enseñanza y el aprendizaje, (del inglés, "We are at the dawn of an era in which educators have the potential to harness technology to produce a step change in student achievement. Although visionaries have been promising for years that technology would transform primary and secondary education—and despite the billions of dollars spent on networking schools and equipping them with computers and other devices—the actual impact on student outcomes to date has been disappointing. Yet when technology is strategically introduced into every step of the educational value chain, it does, in fact, have the potential to enhance every aspect of instruction and learning").*

Con el fin de mejorar dramáticamente los resultados del estudiante, la tecnología debe estar totalmente alineada con los objetivos educativos, normas, planes, evaluaciones, intervenciones y profesionales desarrollo (Bailey et al., 2011).

En los últimos años se ha constatado un tendencia en lo referente a la realización de diversos análisis y, en algunos casos, estos se acompañan con representaciones visuales y el uso de técnicas de analítica visual, un ejemplo de estas son los trabajos de (Butcher, Bhushan, & Sumner, 2006; Butcher & Kintsch, 2004; Castaño Garrido, 2004; Coll & Monereo, 2008; Davies & Graff, 2005; De Laat, Lally, Lipponen, & Simons, 2007; Dicheva & Dichev, 2006; Dicheva, Dichev, & Wang, 2005; Donath, 2002; Ewy, 2003; Ganoe et al., 2003; Gibbs, Olexa, & Bernas, 2006; Gómez Aguilar, Suárez Guerrero, Therón Sánchez, & García Peñalvo, 2010; Govaerts et al., 2012; Govaerts, Verbert, Klerkx, & Duval, 2010; Granda, Uria, García, Suarez, & Gonzalez, 2008; Guo & Chen, 2006; Hardless & Nulden, 1999; Hwang & Arbaugh, 2009; Kim, 2013; Lauer, 2006; Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003; Macfadyen & Dawson, 2010; Mackinlay, 1986; Mazza & Dimitrova, 2005; Mazza & Milani, 2004; Michinov, Brunot, Le Bohec, Juhel, & Delaval, 2011; Morales Morgado et al., 2009; Q. V. Nguyen, Huang, & I, 2004; Quang Vinh, Yu, MaoLin, & JiaWan, 2013; Rossling et al., 2006; Silva & Figueira, 2012; Tat & Carpendale, 2006; Tuckman, 1991; Weisskirch & Milburn, 2003; Willging, 2008; Williams & Conlan, 2007). Estos trabajos serán descritos y analizados en detalle en el Capítulo V. Una representación visual tiene como principal finalidad permitir al usuario adquirir un cierto conocimiento a partir de los símbolos, colores y formas que percibe y este hecho debería ser directo e intuitivo.

En el caso específico de la implementación de los modelos para la representación del conocimiento, visualización analítica, analítica del aprendizaje y analítica académica, las herramientas actuales únicamente se limitan a ofrecer un conjunto de visualizaciones o interfaces de usuario que no son adecuadas para las metodologías existentes o en otros casos se han quedado rezagadas, un ejemplo de estas son los trabajos de (Dicheva & Dichev, 2006; Dicheva et al., 2005; Donath, 2002; Ganoe et al., 2003; Gibbs et al., 2006; Govaerts et al., 2010; Gretarsson et al., 2012; Hardless & Nulden, 1999; Lauer, 2006; Mackinlay, 1986; Mazza & Dimitrova, 2005; Mazza & Milani, 2004; Ming, Calvo, & Pardo, 2013; Q. V. Nguyen et al., 2004; Quang Vinh et al., 2013; Rossling et al., 2006; Silva & Figueira, 2012; Tat & Carpendale, 2006; Weisskirch & Milburn, 2003; Williams & Conlan, 2007), los cuales se describirán y detallarán sus ventajas y defectos en el Capítulo V. A groso modo estas presentan un flujo de trabajo no claramente definido, ni no son herramientas robustas de soporte a los procesos de análisis e interacción con los LMS, por lo que pueden llevarse a cabo de forma ineficiente y se vuelven extremadamente complejos para el analista, ya sea un estudiante, un profesor o un gestor.

En conclusión, además de haberse expresado la importancia del análisis de la información educativa para la mejora del proceso de enseñanza/aprendizaje, y el actual reto de tener que enfrentarse a una vasta cantidad de datos, la problemática de análisis y visualización de datos en *eLearning* en el presente trabajo de tesis se compone de dos vertientes.

Por un lado, la diversidad de propuestas metodológicas, sin que exista una que resulte claramente dominante, que abarque el proceso, y permita realizar una analítica para ambos niveles jerárquicos: Analítica Educativa y Analítica Académica (Van Harmelen & Workman, 2012). Esta situación incide directamente en los diseños de los recursos de un LMS así como del propio curso, su programación y distribución temporal (Baepler & Murdoch, 2010).

Esta difícil situación está directamente relacionada con el usuario, que es el que se encarga de analizar los datos académicos y/o de aprendizaje y, como resultado, se generan, actualmente, diseños de visualización analítica aún con deficiencias. Y, por tanto, no se logra entender el proceso de aprendizaje y menos aún se logra llegar a la predicción de su resultado.

La segunda es la falta de correspondencia entre las metodologías de analítica y las herramientas de analítica visual, así como el hecho de que las herramientas sean poco robustas, lo que dificulta en demasía el desarrollo de los procesos de analítica visual aplicada al *eLearning*. Esta problemática es atribuible directamente a las herramientas existentes, ya que, como se describirá en el Capítulo V más adelante, se limitan al uso de técnicas simples que son fácilmente llevadas al límite de sus capacidades o rebasadas, y tampoco implementan mecanismos de interacción eficientes.

Las herramientas actuales aplican en su mayoría técnicas de analítica, las cuales se encuentran dentro de los procesos de análisis confirmatorio (definido en la sección IV.4). El presente trabajo proporciona un modelo, descrito en la sección VI.1.2, que además de la aplicación de este tipo de técnicas, también posibilita la realización de un análisis exploratorio.

### I.3 Objetivos y Preguntas de Investigación

La mayoría de las limitaciones mencionadas en la sección anterior persisten incluso con el uso de las nuevas tecnologías. Estas son las razones que impulsan este trabajo de tesis con el objetivo de hacer frente a estos desafíos: en lugar de simplemente mover objetos del modelo de aula tradicional al modelo basado en la Web, ¿cómo se puede utilizar la tecnología para hacer formación en un entorno en línea?, y ¿qué beneficios pueden derivarse de estos usos de la tecnología en el contexto educativo?

Las herramientas actuales para analizar tanto el contenido de los ambientes de *eLearning*, así como la evolución temporal, semántica y social del mismo, no han explotado al máximo las técnicas de análisis visual (Scholtz, 2006) o centran su atención en pocas técnicas de visualización, además de estar solo enfocadas al contenido de conocimiento del ambiente o administración del propio material educativo. En muchos de los casos están enfocadas únicamente para el uso del profesorado y como parte de un análisis confirmatorio de los datos.

El trabajo de esta tesis doctoral básicamente propone el planteamiento de soluciones a la problemática de la representación visual y del análisis de los datos de un LMS. Estas soluciones deben ser planteadas desde una perspectiva diferente a las actuales propuestas, con el objetivo de hacer eficientes los diversos procesos involucrados en el análisis visual de cursos en línea. Más específicamente, se requiere la propuesta de un flujo de trabajo y actividades para la analítica visual en el *eLearning*.

Esta tesis doctoral parte de una serie de cuestiones de investigación, entre las que destacan las siguientes.

### I.3.1 Preguntas de investigación

- 1) ¿Es posible definir un modelo de visualización analítica en *eLearning* capaz de entender y mejorar el proceso de enseñanza y aprendizaje?
- 2) ¿Es posible desarrollar un modelo que pueda aplicarse con relativa independencia de los paradigmas teóricos, enfoques metodológicos, e incluso de las soluciones tecnológicas adoptadas?
- 3) ¿Es posible que sirva el modelo para evaluar y mejorar el aprovechamiento de los recursos ICT utilizados para los cursos en línea y a los involucrados en este?; ¿que sirva de guía tanto en el desarrollo de estos sistemas, como para tomar decisiones apropiadas para adecuar el ritmo de las clases y del aprendizaje de los estudiantes?
- 4) ¿Puede el modelo responder a los exigentes requisitos actuales?
- 5) ¿Es posible desarrollar un modelo que, más allá del escenario de origen, sea exportable también a otros contextos institucionales o corporativos, facilitando así la transferencia de conocimiento y la recopilación sistemática y estructurada de buenas prácticas que contribuyan a la mejora continua de los procesos formativos y que, a su vez, proporcione los datos y la funcionalidad necesaria para desarrollar aún más estos sistemas en el futuro?

Como respuesta a las cuestiones recién formuladas se propone el objetivo principal de este trabajo de investigación, a saber.

#### Objetivo principal

*Definir un modelo capaz de esquematizar los pasos del análisis de la información de forma visual generada por el proceso de enseñanza y aprendizaje, que tome en cuenta los diseños y metodologías más actuales acerca de este proceso de analítica en la educación y el proceso de analítica visual de la información.*

Para plantear soluciones a este objetivo la investigación se parte del enfoque que la Analítica Visual define para la visualización, exploración y análisis de datos, ya que este enfoque basado en explotar la capacidad cognitiva humana, junto a los potentes métodos de procesamiento de datos, ha permitido el desarrollo de diversas herramientas y soluciones exitosas en otros contextos. Este objetivo principal se concreta en los siguientes objetivos específicos.

### I.3.2 Objetivos específicos

- 1) Analizar las distintas aproximaciones para llevar a cabo el análisis de la información en la Web y más concretamente en las plataformas de *eLearning*. Estudiar cuáles son los recursos para el análisis de la información, los elementos necesarios para desarrollarlo y el proceso de minería del uso sobre el que basar el análisis.

- 2) Evaluar y desarrollar en base a los recursos para el análisis de la información una aplicación de analítica visual adecuada para el profesorado, estudiante o gestor académico. Evaluar sus ventajas y desventajas y corroborar su adaptación al modelo propuesto.
- 3) Diseñar un modelo que tome en cuenta las teorías y restricciones actuales del análisis de información educativa y de las técnicas de analítica visual, así como los diferentes perfiles del proceso de aprendizaje que permita:
  - a) Proporcionar retroalimentación para la mejora de los recursos y la mecánica de la enseñanza.
  - b) Suministrar a los docentes la retroalimentación necesaria para fomentar su autoevaluación y autorreflexión con el fin de mejorar su aprendizaje.
- 4) Desarrollar un sistema de analítica visual de la información en base al conocimiento adquirido y que satisfaga el modelo propuesto.
- 5) Analizar la usabilidad y la utilidad percibida por parte de los usuarios de la herramienta desarrollada e introducir ciclos de realimentación y mejora del producto teniendo en cuenta los resultados obtenidos.
- 6) Proporcionar nuevo conocimiento sobre el ambiente de aprendizaje virtual de tal manera que fortalezca y soporte la toma de decisiones educativas.

Por tanto, este trabajo se centra en la creación de un modelo de analítica visual en *eLearning* donde se desarrollen y/o implementen nuevas técnicas que resalten los datos menos obvios por medio del análisis exploratorio y confirmatorio de los datos de un CMS/LMS, que, además, propicie las condiciones necesarias para que tanto los estudiantes como los profesores y los gestores académicos puedan hacer autoevaluación de su participación en los procesos de enseñanza/aprendizaje. Por ejemplo, los estudiantes podrían llegar a conocer cuándo necesitan más práctica y cuándo están listos para avanzar en el temario, en esencia, un objetivo secundario de este trabajo es el de propiciar el autoaprendizaje no solo para aprender las materias concretas, sino también para aprender a aprender.

Además, en el caso de profesores, se facilita tener herramientas para evaluar mejor o de forma diferente a los estudiantes y, en el caso de gestores y profesores, se introducen medios para evaluar el proceso de enseñanza y aprendizaje, además de la evaluación de la plataforma, evolución y estándares.

#### **I.4 Metodología de investigación**

Para lograr los objetivos propuestos se propone el siguiente plan de acción. De manera global la investigación comprende dos grandes enfoques metodológicos: el primero corresponde a una revisión sistemática de la literatura, para poder establecer el estado del arte al comienzo del desarrollo de la tesis; el segundo se basa en una serie de ciclos investigación-acción (Ferrance, 2000), a medida que se va desarrollando el trabajo de tesis, con el objetivo de solucionar problemas específicos.

Una revisión sistemática de la literatura representa generalmente la fase inicial en cualquier proyecto de investigación y desarrollo. Es la forma en la que los investigadores pueden obtener el conocimiento existente sobre una cierta línea de investigación. El primer paso corresponde a la búsqueda en la literatura de artículos relevantes para poder detectar soluciones a problemas similares o para poder analizar los procedimientos llevados a cabo por otros investigadores.

Para el caso específico de este trabajo de tesis, la revisión sistemática comprende las siguientes áreas:

- Los sistemas de representación de información, relaciones sociales y patrones temporales relacionados con el análisis y comprensión del proceso de enseñanza y aprendizaje. Específicamente, las metodologías para el desarrollo de análisis y las métricas de este proceso, así como las herramientas para su mejora.
- Procesos de análisis y soluciones basadas en VA, LA y AA a diversos problemas similares.
- Soluciones basadas en el área de Visualización de Información (*Infovis*) a problemas similares.

La metodología de investigación a seguir tendrá, de manera específica cuatro fases:

1. En primer lugar, y de forma intensiva en la fase inicial, se lleva a cabo una búsqueda bibliográfica exhaustiva del estado del arte en el transcurso de los últimos diez años en torno a los temas mencionados anteriormente.

Además, periódicamente se realiza una revisión de los temas que atañen al proyecto, como son las técnicas de visualización aplicadas al *eLearning*, la modificación de estándares o cambios en la estructura de los sistemas administradores de cursos en línea y otros temas relacionados que puedan resultar de ayuda, como son las técnicas de visualización aplicadas a otros campos, métodos distintos de clasificación de patrones aplicados a LMS y LCMS. La asistencia a congresos y el intercambio de ideas con expertos en el tema a través de seminarios, con el objeto de mantener actualizadas los temas implicados en el proyecto.

2. En segundo lugar, se realiza un estudio y análisis de los CMS, LCMS y posible desarrollo de las técnicas de AA, LA y VA más interesantes de la bibliografía recopilada. El grado de interés se determina según los resultados obtenidos tras el análisis de la expresividad y complejidad de los diversos tipos de visualizaciones elegidas.

Una vez conocido el estado del arte, se definen las pautas de un modelo para implementar algoritmos con la finalidad de elaborar una herramienta eficiente de analítica visual, valorando su rendimiento computacional con datos reales. De todo este trabajo se obtienen resultados que ayudan a determinar las necesidades reales de visualización de ambientes *eLearning*, con lo que se logra centrar la investigación en los campos más importantes.

3. En tercer lugar, una vez obtenido un alto grado de profundización en todos los temas relacionados con los ambientes de *eLearning* y analizadas las demandas computacionales y de técnicas de analítica visual de estos últimos, se diseña un modelo y se implementa un conjunto de técnicas analíticas que mejoran y facilitan el análisis y la obtención de nuevo conocimiento para la mejora, así como el entendimiento, de la interacción en el proceso de aprendizaje. Se adoptan técnicas de visualización existentes en otros campos para analizar los datos definidos como relevantes para el proyecto, buscando optimizar los tiempos de cómputo, la expresividad y eficiencia del análisis. Además, y aún más importante, se desarrolla un conjunto de nuevas técnicas y métricas diseñadas exclusivamente para el análisis de los CMS/LCMS. Durante la fase de implementación de las técnicas visuales, estos tres pasos anteriores se solapan y repiten cíclicamente conforme se incrementa el conocimiento sobre la materia, se generan nuevos resultados y se publican artículos. Estas fases se desarrollan de manera paralela en una serie de ciclos investigación-acción, muy específicos sobre el proceso y las soluciones más adecuadas para su desarrollo. Estos ciclos complementan la revisión sistemática y representan una revisión específica y muy puntual de la literatura sobre ciertos aspectos del estado actual del desarrollo.
4. Finalmente, una vez desarrollado el modelo e implementadas las técnicas visuales y de analítica, la interacción con los usuarios y la finalización del prototipo, se elabora un conjunto de casos de estudio que permiten evaluar, por un lado, la usabilidad de las actividades y flujos de trabajo del análisis y, por otro, la funcionalidad de la herramienta.

Para la revisión bibliográfica, se utilizó principalmente el Servicio de Archivos y Bibliotecas de la Universidad de Salamanca, siendo vital el soporte que brinda para revistas electrónicas. Además, se programan reuniones con los directores y otras personas implicadas en la investigación de modo que se pueda mantener una comunicación directa sobre los avances, ideas e impresiones del grupo, ayudando así a criticar, validar y planificar las actividades futuras.

## I.5 Líneas de investigación de la tesis

La presente tesis ha sido desarrollada dentro del marco de las líneas de investigación de la Ingeniería del Conocimiento, Visualización de Información, Analítica Visual, Analítica del aprendizaje, Analítica académica y Minería de datos educativa dentro de dos grupos de investigación de la Universidad de Salamanca, por un lado, el Grupo de Investigación en Interacción y *eLearning* (GRIAL)<sup>2</sup>, y por otro lado el Grupo de Visualización de la Universidad de Salamanca (VisUSAL)<sup>3</sup>.

---

<sup>2</sup> <http://grial.usal.es/>

<sup>3</sup> <http://vis.usal.es/~visusal/grupo/>

Este trabajo de tesis se relaciona con otros proyectos del grupo de investigación GRIAL, este es un Grupo de Investigación Reconocido de la Universidad de Salamanca y Grupo de Excelencia de la Junta de Castilla y León. Está compuesto por un nutrido grupo de investigadores de diferentes ámbitos de conocimiento, en el que predominan los perfiles técnicos y pedagógicos, pero cuenta también con expertos en gestión de proyectos de *eLearning* procedentes del ámbito de las Humanidades, Ciencias Experimentales, etc.

Las principales líneas de investigación del grupo se refieren a aspectos como: Sistemas interactivos para el aprendizaje, Tecnologías para el aprendizaje, Ingeniería Web y Arquitectura del *Software*, Metodología del aprendizaje *online*, Calidad y evaluación en educación, Teoría de la comunicación y Gestión estratégica del conocimiento y de la tecnología.

Respecto a los proyectos de investigación financiados obtenidos de la cooperación de ambos grupos se tienen:

- oiPLE: entorno abierto, integrado y personalizado para el aprendizaje. Hacia una nueva concepción de los procesos de aprendizaje basados en tecnología - TIN2010-21695-C02. Proyecto coordinado de Investigación Fundamental No Orientada, financiado por el Ministerio de Ciencia e Innovación. Se centra concretamente en la posibilidad de exportar funcionalidades del LMS a otros contextos, así como facilitar la integración de otras aplicaciones dentro de él para personalizar el aprendizaje, dicho proyecto se puede considerar como uno de los pilares fundamentales en que se asienta la presente tesis.
- Layers4Moodle – TSI-020302-2009-35. Financiado por el Ministerio de Industria, Turismo y Comercio dentro del subprograma I+D. Pretende el desarrollo de un entorno de ejecución de aplicaciones, así como estas aplicaciones, para permitir crear capas de información generadas por los usuarios y superponerlas sobre los contenidos generados por un sistema *software* de gestión del aprendizaje como Moodle. El proyecto del Gobierno de Castilla y León GR47 y los proyectos del Ministerio de Ciencia e Innovación FI2010-16234 y TIN2010-21695-C02-01.

Desde el grupo GRIAL se han llevado a cabo varias líneas de trabajo relacionadas con *eLearning*:

- En el área de gestión del conocimiento se desarrolla una propuesta de investigación encaminada a evaluar la calidad de los Objetos de Aprendizaje (OA) desde una perspectiva pedagógica y tecnológica. Para ello propone instrumentos que permiten la evaluación de los objetos de aprendizaje al recopilar información cuantitativa y cualitativa de la calidad de los objetos. Dentro de estos instrumentos tienen especial relevancia la herramienta HEODAR (Herramienta para la Evaluación de Objetos Didácticos de Aprendizaje Reutilizables) como herramienta para medir la calidad de los objetos de aprendizaje integrada en Moodle (Morales-Morgado, Muñoz, Conde, & García-Peñalvo, 2009). Estos mecanismos facilitan la selección de objetos

relacionados con unos objetivos educativos concretos que pueden reutilizarse (Morales, 2008).

- Propuesta basada en herramientas para la definición y empaquetamiento de pruebas de evaluación que siguen especificaciones como IMS QTI (IMS-GLC, 2006a) e IMS CP (IMS-GLC, 2007a). De esta forma se establecen mecanismos de evaluación para Sistemas Adaptativos e Hipermediales de Aprendizaje (Barbosa León & García Peñalvo, 2005; Barbosa León, García Peñalvo, Rodríguez Conde, Morales Morgado, & Ordoñez de Pablos, 2012).
- Propuesta para la definición de diseños instruccionales adaptativos, soportados por una herramienta de autor para la realización de Diseños Instructivos Adaptativos denominada Hyco-LD y que se basa en la especificación IMS-LD (Berlanga & García Peñalvo, 2004; A. J. Berlanga & F. J. García Peñalvo, 2008).
- METHADIS (Metodología para diseñar un sistema Hipermedia Adaptativo para el aprendizaje basados en estilos de aprendizaje y estilos cognitivos) (Prieto, 2006). Con el fin de sistematizar la tarea de diseño de sistemas hipermedia adaptativos se define una metodología, que en función de un objetivo de aprendizaje y mediante una secuencia de etapas, determina las estrategias instruccionales más adecuadas al contexto de aprendizaje y a las particularidades de los diferentes usuarios.
- AHKME (*Adaptive Hypermedia Knowledge Management ELearning System*), sistema de gestión del conocimiento basado en web que combina conceptos de la Web 2.0 y Web 3.0 relativos a la colaboración avanzada, redes sociales, interoperabilidad, estandarización y adaptación. El objetivo de este sistema es preparar a los usuarios para el nuevo paradigma relativo a la Web 3.0 y dotarles de herramientas para la reutilización y para flexibilizar y personalizar el aprendizaje (Rego, 2012; Rego, Moreira, & García-Peñalvo, 2011).
- Definición de una metodología de tutoría *online* en el ámbito de la formación, que recogería las competencias, métodos y elementos cualitativos relativos a la formación, necesarios para garantizar cursos de aprendizaje de calidad (Seoane, García, Bosom, Fernández, & Hernández, 2007; Seoane-Pardo & García-Peñalvo, 2006)(Seoane Pardo, 2014).
- Investigación en el ámbito de la Ingeniería de Ontologías y la Visualización de la Información para facilitar la conceptualización de ontologías de gran tamaño. Para llevar a cabo esta labor se define la herramienta OWL-VisMod (J. García, Peñalvo, Therón, & de Pablos, 2011; García Navarro, 2012; García-Peñalvo, García, & Therón, 2011).
- Uso de la Analítica Visual para observar la evolución de *software* con posibles aplicaciones como el mantenimiento del mismo (Therón, González, & García, 2008; Therón, González, García, & Santos, 2007). En concreto se propone la herramienta Maleku para llevar a cabo tal cometido (González-Torres, Therón, García-Peñalvo, Wermelinger, & Yu, 2011).

- Investigación relativa a entornos personalizados de aprendizaje, especificaciones de interoperabilidad, servicios web, arquitecturas orientadas a servicios y *mobile learning* (Alier Forment, Casany Guerrero, Conde González, & García Peñalvo, 2010; Alier Forment et al., 2013; Casany Guerrero, Alier, Conde González, & García-Peñalvo, 2009; Casany Guerrero, Conde González, Alier, & García-Peñalvo, 2009; Conde González, 2007, 2012; Conde González, Casany Guerrero, Alier, & García Pealvo, 2009; Conde González, García Pealvo, Casany Guerrero, & Alier, 2011; Conde González, García Peñalvo, Guerrero, & Forment, 2009; Conde González, García-Peñalvo, Casany, & Forment, 2010; Conde González, García-Peñalvo, Casany, & Forment, 2013; Conde González, Gómez Aguilar, del Pozo de Dios, Therón Sánchez, & garcía Peñalvo, 2010; Conde González, González, & Moren, 2006).

## I.6 Contenido

El siguiente capítulo de este trabajo, El aprendizaje y la tecnología, describe a groso modo el proceso de la introducción de las ICT en la educación, presenta algunas definiciones y tecnologías ocupadas; cómo ha sido su evolución e influencia en el proceso educativo y cómo ha surgido la aparición de nuevas metodologías y plataformas.

Ulteriormente, se presentan algunas definiciones de *eLearning* y para finalizar el capítulo, la forma en la que actualmente se presenta a los estudiantes, esto es, las herramientas de aprendizaje en línea.

El capítulo Analítica en la educación presenta una definición global de lo que la analítica es y cómo se ha ido adaptado desde el área de negocios hasta la formación educativa, pasando por la formación empresarial y la inteligencia de negocios; además expone las diversas definiciones en el área de la analítica en la educación y los grupos existentes. Dentro de estas líneas de investigación se describen más detalladamente la de la analítica académica y la de la analítica del aprendizaje, para exponer el modelo y el proceso de la analítica educativa. Finalmente, y no por eso menos importantes, dos líneas de investigación que han dirigido su objeto de análisis al área educativa, la minería de datos educativa y el análisis social educativo, y como estas abordan las grandes cantidades de información que actualmente se generan, los *big data*, que también se definen.

El capítulo cuarto, Visualización de la Información y Analítica Visual, presenta la forma en que las representaciones visuales afectan a los procesos educativos desde la psicología educativa, para posteriormente definir estas áreas de investigación y cuál es su aporte al presente trabajo; aunado a esto define la visualización y el proceso de la analítica visual, después, se presenta el proceso de la visualización analítica, el mantra de la búsqueda de información, elementos importantes en el presente trabajo y, por último, se explica el término alfabetización de la información o de los datos, o alfabetización visual y su relación con la visualización analítica.

En el siguiente capítulo, Capítulo V, se presentan los aportes al estado de la cuestión actual de esta tesis doctoral, principalmente en cuanto al análisis de información educativa para la mejora del proceso de enseñanza

y aprendizaje, donde se exponen los trabajos más relacionados con la presente investigación, ya sea por su semejanza en materia visual, teórica, técnica o procedimental y cómo estos, por medio del análisis que se describe y de la presentación de los problemas o deficiencias encontrados y las soluciones posibles, dan lugar la propuesta.

El capítulo sexto, Modelo de Analítica Visual en *eLearning*, expone la propuesta del modelo, sus desafíos y necesidades. A continuación, se describe el proceso de diseño del prototipo y de desarrollo del mismo y se describen las técnicas utilizadas para su construcción.

Posteriormente, en el capítulo dedicado a los Casos de estudio, se describen, para cada uno de estos, los datos a analizar, las preguntas a responder y los resultados obtenidos y, finalmente, la evaluación de la usabilidad realizada.

Por último, en el capítulo octavo se exponen los resultados obtenidos, las conclusiones generales y las publicaciones realizadas. Para cerrar este capítulo se describen las líneas de trabajo futuro.



## Capítulo II El aprendizaje y la tecnología

*"Bringing together students from all walks of life, across cultures and nationalities—that is the ultimate learning experience."*

*—Shai Reshef, UoPeople founder and president*



La globalización tiene una de sus manifestaciones más relevantes en las denominadas ICT, las cuales han permitido llevar la globalidad al mundo de la comunicación, lo que facilita la interconexión entre las personas e instituciones a nivel mundial, y elimina barreras espaciales y temporales. Se denominan ICT al conjunto de tecnologías que permiten la adquisición, producción, almacenamiento, tratamiento, comunicación, registro y presentación de informaciones, en forma de voz, imágenes y datos contenidos en señales de naturaleza acústica, óptica o electromagnética (Kaffash, Kargiban, Kargiban, & Ramezani, 2010). La aplicación de esas tecnologías no garantiza por sí misma una mejora en los procesos de aprendizaje, pero es necesario tenerlas en cuenta de cara a proporcionar las soluciones de aprendizaje más adaptadas las necesidades y circunstancias de los estudiantes. Es decir, se requiere conocer las diferentes modalidades de formación que aparecen en función del soporte tecnológico utilizado, como puede ser el uso de Internet y las tecnologías web (en lo que se denomina *eLearning*), los dispositivos móviles (*mLearning*), los juegos (*gLearning*), la utilización de sensores distribuidos en el entorno del usuario para que pueda aprender del contexto que le rodea (*uLearning*), las tecnologías derivadas de los conceptos web 2.0 (*eLearning 2.0*) y un largo etcétera. Todas estas modalidades tienen, a pesar de sus diferencias, un objetivo común, facilitar el aprendizaje del estudiante, aunque para ello utilicen distintos soportes que suponen una adaptación de la forma en que se llevan a cabo las actividades.

En el presente capítulo se desarrolla brevemente la historia del aprendizaje electrónico y cómo las ICT se encuentran al servicio de la educación, como es el caso de los LMS, CMS y VLE, que son introducidos y definidos. Además se describe cómo estos dan lugar a diversas metodologías, plataformas, procesos de administración, difusión y cómo se han ido adaptando a las necesidades de la educación actual y la interoperabilidad que esto conlleva.

## II.1 Inicios del Aprendizaje Electrónico

El aprendizaje electrónico (en inglés, *electronic learning*, *eLearning*) se basa en el uso de tecnologías de la información e Internet como herramientas de enseñanza y aprendizaje. Tiene sus orígenes en los años 60 cuando se tienen los primeros usos de los ordenadores con fines educativos. En general las primeras generaciones del *eLearning* se basan en introducir las ICT para fomentar nuevas actividades y posibles concepciones del aprendizaje y se comienza a utilizar el ordenador como una herramienta formativa, sin tener que, obligatoriamente, requerir una conexión a la Red, ni contar con enlaces a material didáctico externo; siendo lo más frecuente basarse en soportes como CD ROM, DVD o Discos Flexibles, esto es, consiste en un conjunto de *hardware* y *software* de alto coste económico que ayuda al profesor y al estudiante en el proceso de enseñanza y aprendizaje. Este aprendizaje se basó en un conjunto de técnicas de estructuración de contenidos mediante los que se hace seguir a cada usuario un proceso individualizado y temporalmente predeterminado, que puede ser lineal o ramificado (Azcorra Saloña, Bernardos Cano, Gallego Gómez, & Soto Campos, 2001; RUIPÉREZ, Castrillo Larreta, & García Cabrero, 2006).

En España esta modalidad de aprendizaje surge en los años 80, su denominación más común fue la de Enseñanza Asistida por Ordenador (EAO, en inglés *Computer-Aided Instruction* o *Courseware* o *Computer-Assisted Instruction*, CAI o *Computer-Based Training* CBT); debido a la ausencia de redes en esta etapa, los esfuerzos se centraban en el desarrollo de paquetes autónomos de contenidos educativos, se comienza lo que se conoce como la revolución electrónica y a principios de los 80 aparecen los primeros ordenadores personales que abren las puertas al aprendizaje individualizado por ordenador. No obstante, desde los años 50-60 del siglo XX se acuña el concepto de educación a distancia, esta no se caracteriza, en esta etapa, por avances significativos desde un punto de vista tecnológico, pero sí supone un cambio en la consideración del aprendizaje como una tarea fundamentalmente autónoma por parte del estudiante. Por otra parte, la educación a distancia rompe la vinculación natural del aprendizaje formal con un espacio físico de referencia, el aula, como lugar que ocupa el centro del proceso enseñanza-aprendizaje (Nicholson, 2007).

Posteriormente, partir de los años 90, en la primera mitad se empiezan a contemplar posibilidades multimedia avanzadas, la combinación de diferentes medios, sonidos, textos, imágenes fijas y en movimiento en un mismo diseño y en el mismo soporte. Aumenta la estandarización, el campo de ordenadores aumenta con rapidez y se desarrollan soportes de alta capacidad de almacenamiento que permitieron incorporar elementos multimedia, sobre todo vídeo, de forma que se enriquecen de forma notable los contenidos y se dota de mayor interactividad a las aplicaciones. Además, la inclusión de juegos educativos en estos nuevos medios empieza a proliferar. Con su aparición nacen en España varios proyectos (LETRA, ATENEA, etc.) y empresas (Anaya Interactiva, Telefónica I+D, Chadwyck-Healey, BSI Multimedia, etc.). Este tipo de formación es de tipo *offline*, en la que usuario interactúa meramente con el ordenador cuando él lo desea.

Con la llegada de Internet, a partir de mediados de los años 90, se favoreció, además de la propia concepción de la Red, la facilidad para hacer contenidos distribuibles con un sencillo editor, lo que repercutió de forma positiva en la creación de contenidos, pero a la vez, ha tenido algunos efectos negativos, y todo ello ha implicado un cambio de estrategias y paradigmas. Esta etapa puede considerarse como una fase de transición en la que el usuario realiza los cursos *offline*, pero dispone de ciertos campus virtuales en los que se recrea el marco típico de relaciones de la formación presencial, lo que conduce al aprendizaje *online* y la popularización del Internet se produce gracias al establecimiento de un nuevo servicio, el *World Wide Web* (Alesso & Smith, 2008; Berners-Lee, Cailliau, Groff, & Pollermann, 1992; Shadbolt, Berners-Lee, & Hall, 2006).

Posteriormente, las grandes empresas tecnológicas parecen apostar decididamente por el aprendizaje *online*, se incorporan los cursos a sus intranets o redes corporativas, como complemento o alternativa a la tradicional formación continua presencial. Dado a este cambio, el siguiente paso evolutivo de los sistemas de aprendizaje basados en computador fue el denominado aprendizaje basado en la Red (en inglés, *Internet Based Learning*, IBL), También se conoce como instrucción basada en la Red (en inglés, *Internet Based Training*, IBT) en el cual los contenidos llegan a sus destinatarios a través de Internet o de una intranet (Boneu, 2007). Esto supone el paso de un sistema de aprendizaje que se producía en un entorno *offline* como era el CBT a otro con mayores posibilidades pedagógicas, el IBT. El IBT es un proceso en el que el estudiante tiene un acceso global más de lo que puede tener a través de un *software* local, y esta extensión genera disponibilidad de cantidad potencialmente masiva del *software* educativo y una interactividad directa al profesor particular.

Durante el periodo del 2000 al 2003 puede considerarse que se produce una crisis del *eLearning*, ya que las empresas realizan un importante desembolso en plataformas y prima la premura en la impartición de contenidos que permita rentabilizar la plataforma, en lugar de la calidad de los contenidos en sí o de la tutoría. Teniendo como resultado un efecto de enorme expectación inicial y elevadas dosis de fascinación (acompañadas de una fuerte inversión tecnológica), no parecía que estas costosas innovaciones aportaran nada realmente significativo desde el punto de vista didáctico. En esta época comienza a considerarse la posibilidad del aprendizaje mixto (*blended learning*), que permite compensar las carencias de la formación en línea sin renunciar a sus ventajas. En este período se produce un impulso general de inversiones tecnológicas que favorecieron la distribución de Internet y la creación de una cultura común de la tecnología, sobre todo en los países desarrollados, produciendo, como era de esperarse, una brecha digital en otros países con respecto a estos (Al-Abbasi, 2012; Rey, 2012).

Esta conciencia crítica, sumada a una mayor cultura tecnológica y al aumento masivo de los destinatarios potenciales de formación en línea, facilita una revolución en las dinámicas de interacción *online* e incluso una cierta racionalización de la inversión tecnológica necesaria para la formación en línea, dando como resultado, en los años 2004 al 2006, un cierto movimiento de reivindicación del protagonismo de los roles humanos del aprendizaje. Se buscan entornos tecnológicos mucho más

completos que no solo faciliten la distribución de contenidos sino que, a su vez, debieran proporcionar un elevado grado de interacción entre los actores implicados en este proceso. Es entonces cuando surgen las auténticas “herramientas de *eLearning*”, bien *integradas* en las plataformas o disponibles en una Web cada vez más interactiva, dinámica, viva.

La diversidad de dispositivos de acceso también favorece el uso de estas nuevas herramientas. La difusión de ordenadores portátiles y, posteriormente, de otros aparatos incluso más pequeños con grandes posibilidades de conectividad (*tablets, smartphones, etc.*), impulsan la aparición de nuevas modalidades (o quizá sería mejor decir “adaptaciones”) del *eLearning* como el *mobile learning* o *mLearning*. Gracias a la popularidad de estos dispositivos, así como al desarrollo de herramientas que facilitan la interacción o la creación de contenidos por parte de cualquier usuario con destrezas tecnológicas mínimas y al aumento exponencial de usuarios de la red, a partir de 2004 se produce la revolución de la llamada Web 2.0 o Web Social: a mayor compartición, mayor conocimiento y, de nuevo, mayor posibilidad de compartir (O’Reilly, 2007; Seoane Pardo, 2014).

Un fenómeno que, según algunos expertos, ya no constituye una revolución tecnológica, sino una revolución social; surge así el *eLearning 2.0* (Downes, 2005). Desde entonces y hasta ahora, el aprendizaje se centra en el estudiante y se adapta a sus necesidades, algo que se materializa en el concepto de los entornos personalizados de aprendizaje. Entornos que reciben nombres diversos, entre los cuales los más conocidos son plataformas, Entornos Virtuales de Aprendizaje, *Learning Management Systems* (LMS), Sistemas de Gestión del Aprendizaje (SGA) o, de una forma más amplia y genérica, Campus virtuales. Aunque en muchos contextos estos términos se emplean como sinónimos, este trabajo se decanta únicamente por la denominación de LMS para referirse a este tipo de herramientas *software* y que se definirá en sección II.3.

Actualmente, cada vez más, proliferan las iniciativas de Conocimiento en Abierto (F. J. García Peñalvo et al., 2010), por ejemplo en las universidades (se destaca la propuesta *Open Courseware Consortium*<sup>4</sup>, al que se encuentran adheridas más de 280 instituciones de 40 países, compartiendo cursos universitarios en abierto), y hasta las convocatorias públicas que financian investigación e I+D exigen, con una tendencia creciente, que el conocimiento resultante se haga público y en abierto (declaraciones internacionales sobre el acceso abierto en las instituciones públicas: Budapest del 2002<sup>5</sup>, Bethesda del 2003<sup>6</sup> y la Declaración de Berlín, de 2003<sup>7</sup>).

---

<sup>4</sup> *OpenCourseWare Consortium*, [www.ocwconsortium.org](http://www.ocwconsortium.org)

<sup>5</sup> *Budapest Open Access Initiative*, [www.budapestopenaccessinitiative.org](http://www.budapestopenaccessinitiative.org)

<sup>6</sup> Bethesda Statement on Open Access Publishing, <http://legacy.earlham.edu/~peters/fos/bethesda.htm>

<sup>7</sup> Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, <http://openaccess.mpg.de/286432/Berlin-Declaration>

Quizá uno de los casos más paradigmáticos, por su relevancia, magnitud y simbolismo, sea el Proyecto Genoma Humano, que durante 13 años y con una inversión de 90.000 millones de dólares ha producido resultados científicos que se hacían públicos y se ponían a disposición de cualquier usuario en Internet (accesibles en la página del *Human Genome Project*<sup>8</sup>).

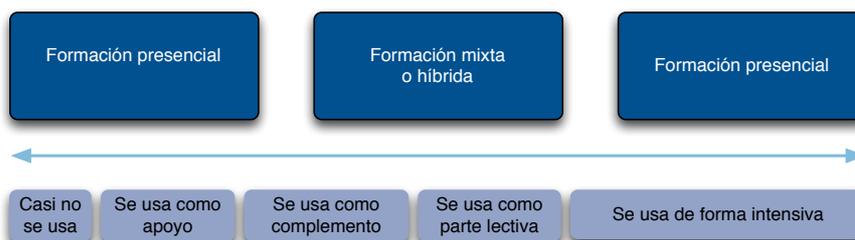


Fig. 1. Espectro de los usos de Internet como herramienta docente en el contexto universitario. Fuente: adaptado de (Duart Montoliu, Gil, Pujol, & Castaño, 2008)

Con la llegada de estas interacciones, esto es, la incorporación de las ICT a la formación, los métodos y características de la diferenciación entre la formación presencial y la formación a distancia se volvieron un poco imperceptibles. La cantidad de tiempo dedicado a un tipo de métodos u otros decide la forma principal de la denominación del programa o proyecto, se puede decir “formación presencial con apoyo” de determinados recursos tecnológicos, si la mayor parte del tiempo del proceso se desarrolla en el aula, o se combina con materiales de tipo auto-educativo, como tutoriales o simuladores. Si la mayor parte de la formación se desarrolla a través de sistemas basados en tecnologías se debería hablar de teleformación.

Existen diferentes clasificaciones de *eLearning*, algunas referidas al modo en que se lleva a cabo, al dispositivo utilizado, etc. La evolución del *eLearning* se describe en una de estas clasificaciones, en concreto, la que atiende al grado de presencialidad del aprendizaje. En función de dicho factor, el *eLearning* se clasifica en (Duart Montoliu et al., 2008): formación presencial, formación mixta y formación virtual (ver Fig. 1).

### II.1.1 Formación Presencial

En la formación presencial el uso de Internet sirve como apoyo o complemento opcional para la formación de los discentes. Este tipo de formación supone un uso ocasional de Internet, la responsabilidad recae sobre el formador, quien coordina y distribuye tareas y tiempos, la transmisión de contenidos se realiza en un centro de formación. Esta, la formación presencial, va modificando su metodología para aprovechar las potencialidades que brinda Internet y las tecnologías web, en general para mejorar sus procesos de enseñanza-aprendizaje; de forma que en muchas acciones de formación empieza a ser poco distinguible la barrera entre la modalidad presencial y la modalidad a distancia (ver Fig. 1).

<sup>8</sup> *Human Genome Project*, [www.ornl.gov/sci/techresources/Human\\_Genome/home.shtml](http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml)

### II.1.2 Formación mixta

En el caso de la formación mixta, supone un uso de Internet como parte lectiva y, por tanto, se considera como parte de la formación no solo las actividades presenciales, sino también otras planteadas en modalidad en línea. Es decir, combina escenarios de formación síncronos y asíncronos. El aprendizaje mixto (del inglés, *blended learning*) también llamado semipresencial, no solamente se usa en entornos de formación empresarial, sino también en instrucciones escolares convencionales.

En los últimos años, muchos investigadores han centrado sus estudios de investigación sobre aprendizaje mixto (D. R. Garrison & Kanuka, 2004). Sin embargo, en el pasado, el aprendizaje mixto se refiere solamente a dos modelos de *blended learning*: físicamente cara a cara, también llamada formación presencial, con formación electrónica asíncrona y formación electrónica cara a cara combinada con formación electrónica asíncrona.

Aunque el término "aprendizaje mixto" alcanzó un uso común recientemente, la idea se ha usado durante años. Diferentes investigadores han definido la formación mixta de forma diferente, por lo que la definición de aprendizaje mixto aún no ha llegado a una conclusión. C. Graham (2005) ha resumido los tres tipos de definiciones de la literatura sobre el aprendizaje mixto, para llegar a la conclusión de que se trata de una combinación de:

- a) Los diferentes métodos de enseñanza.
- b) Las diferentes modalidades o medios de entrega.
- c) Instrucción en línea y cara a cara.

Tradicionalmente se habla de un tipo de formación u otro en función del tiempo que tenga que dedicar el estudiante en su domicilio (particular o puesto de trabajo) o en el centro de formación. Se trataría de un eje continuo en el que cabe situar acciones de carácter mixto, son aquellas que disponen de ambos tipos de formación en mayor o menor grado, son conocidas como *blended learning*.

### II.1.3 Formación virtual

Finalmente, en la formación virtual la formación se realiza de forma totalmente asíncrona, es decir, sin requerir que profesores y estudiantes coincidan en el mismo momento y lugar. Este caso se requiere de alguna herramienta que facilite la interacción entre los actores del proceso formativo, facilite la gestión de recursos y actividades, es decir, de una plataforma de aprendizaje *online*.

Una vez contextualizado el momento actual en que se encuentra la formación en línea en relación con la evolución o "fases de desarrollo" de la tecnología asociada al *eLearning*, puede decirse que, en la segunda década del siglo XXI, existen herramientas prácticamente para cualquier necesidad relacionada con la formación *online*. Sin pretender jugar a visionarios ni afirmar, ni mucho menos, que no quede nada por inventar o descubrir, sí parece existir cierta sensación, al menos entre quienes están iniciados en lo que se refiere a innovaciones tecnológicas, de que buena parte de las necesidades en relación con la formación *online* empiezan a estar cubiertas.

La ausencia de presencia física, elemento que se achaca como un inconveniente al *eLearning* (y lo es, qué duda cabe) es algo que dista todavía mucho de poder solventarse, incluso con las más novedosas soluciones de realidad aumentada y simulaciones virtuales. A continuación se presentan algunas definiciones del aprendizaje *online*.

## II.2 Definición de *eLearning*

En la conceptualización del aprendizaje *online* se habla de tecnologías de la formación cuando se incorpora un elemento digital a procesos de enseñanza y aprendizaje.

En general cuando hay una utilización de tecnologías de la información con algún objetivo didáctico previamente formulado del ordenador. Con la incorporación de las redes (Internet) y por tanto de las telecomunicaciones sería más preciso utilizar la denominación tecnologías de la información y las comunicaciones en formación, ya que define con más exactitud todas las tecnologías aplicadas. En cualquier caso la confusión no parece existir, ambas tecnologías, de la información y de las comunicaciones, coinciden y ofrecen servicios combinados y transparentes para los usuarios.

El término compuesto por las palabras telecomunicación y formación, *teleformación*, es un sistema que aporta formación a distancia, apoyado en las ICT (tecnologías, redes de telecomunicación, videoconferencias, TV digital, materiales multimedia, etc.), que combina distintos elementos pedagógicos: autoestudio, prácticas, los contactos en tiempo real (presenciales, videoconferencias o *chats*) y los contactos diferidos (tutorías, foros de debate, correo electrónico). Es decir, la formación a distancia se convierte en teleformación cuando se lleva a cabo mediante las ICT. La idea esencial está en la distribución abierta, rápida y eficaz de información y contenido didáctico, entre todos los actores implicados en el proceso: profesores-tutores y estudiantes.

El término actualmente en uso es *eLearning*. Teleformación es equivalente a *eLearning*. Su traducción en castellano sería "aprendizaje electrónico". Está asumido que la "e" representa Internet, por ejemplo *eGovernment*, *eBusiness*. En general se llama "*eLearning*" al conjunto de tecnologías, aplicaciones y servicios orientados a facilitar la enseñanza y el aprendizaje a través de Internet/Intranet, que facilitan el acceso a la información y la comunicación con otros participantes. El término formación en línea (en inglés, *online*), donde "en línea" es estar conectado, se usa para indicar la conexión a Internet o a una red.

Hay muchos casos en los que se puede descargar materiales o programas que pueden ser trabajados *offline*, a pesar de ello, si para esa descarga ha sido necesario algún tipo de conexión a una plataforma o página web, o es necesaria o posible la conexión, en ciertos momentos, como por ejemplo para comunicarse con el tutor u otros compañeros, se denomina formación en línea (en inglés, *online learning*). Si se pone el énfasis principal en el estudiante se hablaría de Aprendizaje en Red.

También se puede encontrar otras acepciones que hacen referencia al mismo concepto, como pueden ser formación basada en la Web (en inglés, *Web Based Training*, WBT), este término sí que hace referencia explícita a la tecnología web y, por tanto, a Internet.

WBT consiste en la provisión de contenido educativo a través de un navegador web ya sea en Internet, en una intranet privada o una extranet. Suele hacer referencia a la difusión de contenidos formativos con carácter auto-educativo, de forma similar a la CAI o CBT, pero con difusión y presentación a través de Internet.

Todas estas acepciones tratan de incorporar seudónimos de Internet para indicar que se trata de modalidades de formación que permite utilizar las potencialidades de la Red para acercar la formación a sus posibles usuarios.

Desde el punto de vista conceptual *eLearning* es un término susceptible de diferentes definiciones y a menudo intercambiable por otros: formación *online*, cursos *online*, formación virtual, teleformación, formación a distancia, campus virtual, etc. Estos términos se utilizan habitualmente como sinónimos, aunque se pueden diferenciar según el sentido expresado aquí. Si se tiene en cuenta que el Internet va impregnando todas las acciones y variantes técnicas y metodológicas, pocos elementos quedan en la teleformación claramente distinguibles del *eLearning*, paquetes de autoestudio, tipo tutoriales y simuladores, que en buena medida se están realizando o distribuyendo con tecnologías web.

Simplificando, la teleformación consiste en la utilización de tecnología para incrementar la eficacia y accesibilidad del aprendizaje. Esto incluye aplicaciones multimedia en soportes como el CDROM, utilización de un foro de debate para continuar una discusión fuera del aula, o un curso desarrollado completamente en línea. Todos los demás términos dependen de este. La teleformación consiste en el uso de cualquier tipo de tecnología, desde el nivel más básico, uso del *e-mail* por ejemplo, al más avanzado, un curso completo (o una simulación, o realidad virtual, etc.).

Esencialmente, *eLearning*, a diferencia de la enseñanza tradicional, es otra manera de enseñar y aprender. El aprendizaje electrónico (*eLearning*) es uno de los ambientes de aprendizaje más populares de la era de la información. Por tanto, los esfuerzos de aprendizaje y experimentos actualmente reciben una gran atención en todo el mundo.

Esta complejidad conceptual, así como la diversidad de términos, conceptos y vocablos que se utilizan de forma indiscriminada para referirse a un sistema de formación singular, llevan a matizar lo que se entiende por *eLearning*. Una de las primeras definiciones se debe a la Dirección General de Telecomunicaciones en 1996. Esta definición, a pesar de que no hace una referencia explícita al término *eLearning*, sino que utiliza el vocablo tele-educación, señala cómo se puede catalogar bajo esta etiqueta el desarrollo del proceso de formación a distancia (reglada o no reglada), que basado en el usos de las tecnologías de la información y las telemáticas posibilitan el desarrollo de aprendizajes interactivos, flexibles y accesibles a cualquier posible receptor (MOPTMA- Ministerio de Obras Públicas, 1996). La traducción literal al español llevaría a una concepción de "aprendizaje electrónico". Para este estudio se acepta como punto de partida algunas de las diversas definiciones existentes para el concepto de *eLearning*.

El planteamiento de estas definiciones depende mucho de la perspectiva de los diferentes autores, aunque en todas ellas se plantea el uso de la tecnología y se incide más o menos en las actividades de aprendizaje a través de ellas. Martín Hernández (2006) señala la concepción compleja del *eLearning* que:

*engloba aquellas aplicaciones y servicios que, tomando como base las ICT, se orientan a facilitar el proceso de enseñanza-aprendizaje.*

Por otro lado B. Khan (2000) define que:

*el eLearning abarca aprendizaje basado en la Web (en inglés, Web-Based Learning, WBL), la formación basada en Internet (en inglés, Internet-Based Training, IBT), el aprendizaje distribuido avanzado (en inglés, Advanced Distributing Learning, ADL), y el aprendizaje en línea (en inglés, Online Learning, OL).*

En este sentido pero en otra definición,

*eLearning incluye la instrucción a través de todos los medios electrónicos como Internet, intranets, extranets, y documentos de hipertexto/hipermedia (Govindasamy, 2001).*

Una definición muy relacionada con la tecnología que únicamente considera Internet como canalizador del *eLearning* es la establecida por Rosenberg (2001):

*El uso de tecnologías Internet para la entrega de un amplio rango de soluciones que mejoran el conocimiento y el rendimiento. Está basado en tres criterios fundamentales:*

- 1. El eLearning trabaja en red, lo que lo hace capaz de ser instantáneamente actualizado, almacenado, recuperado, distribuido y permite compartir instrucción o información.*
- 2. Es entregado al usuario final a través del uso de ordenadores utilizando tecnología estándar de Internet.*
- 3. Se enfoca en la visión más amplia del aprendizaje que va más allá de los paradigmas tradicionales de capacitación (Rosenberg, 2001).*

Otra definición es la de Sandars and Kieran (2004), que denota un cambio en la forma en la que el rol del educador cambia de un proveedor de información a un moderador que promueve la idea de que el estudiante es altamente activo y sintetiza una gran cantidad de información desde recursos diversos y dispersos:

*El eLearning se refiere al uso de la tecnología de la información para ofrecer una gama de oportunidades de aprendizaje. El rápido desarrollo de las TI ha permitido contenidos de aprendizaje que se ofrecen en diferentes formatos y son entregados a través de internet, intranet u otros medios electrónicos.*

Además plantea unas características interactivas de la forma en la que el discente aprende:

*Utilizar entornos que son ricos en imágenes multimedia, especialmente visual y audio, se prefieren a los que son en su mayoría compuesto por texto; existe una preferencia para participar activamente en las tareas en lugar de leer sobre los eventos o escribiendo acerca de ellos.*

*La motivación para aprender viene de participar activamente en el aprendizaje y por los intentos de responder a las preguntas que surgen durante una tarea (Sandars & Kieran, 2004).*

Otras definiciones están enfocadas desde la perspectiva de qué ofrece la experiencia en el desarrollo y explotación de plataformas *eLearning*:

*Capacitación no presencial que, a través de plataformas tecnológicas, posibilita y flexibiliza el acceso y el tiempo en el proceso de enseñanza-aprendizaje, adecuándolos a las habilidades, necesidades y disponibilidades de cada discente, además de garantizar ambientes de aprendizaje colaborativos mediante el uso de herramientas de comunicación síncrona y asíncrona, potenciando en suma el proceso de gestión basado en competencias (F. J. García-Peñalvo, 2005).*

En (Stark, Schmidt, Shafer, & Crawford, 2002) se define *eLearning* como:

*La experiencia de obtener conocimiento y habilidades a través de la entrega electrónica de educación, entrenamiento o desarrollo profesional. Esto abarca educación a distancia y aprendizaje asíncrono y puede ser entregada en un formato personalizado para un estudiante en particular.*

Pero antes de adoptar a este trabajo una definición, es interesante revisar la propuesta de (Azcorra Saloña et al., 2001), que incide en el concepto de la interacción y, por tanto, del factor humano, al definir al *eLearning* como:

*Una enseñanza a distancia, abierta, flexible e interactiva basada en el uso de las nuevas Tecnologías de la Información y de las Comunicaciones, y sobre todo aprovechando los medios que ofrece la red Internet.*

A partir de todas estas definiciones pueden observarse multitud de perspectivas del mismo concepto, con ciertos factores comunes como la tecnología, los servicios o la formación, aunque después de lo expuesto, aquí se va a defender una definición de *eLearning* enunciada desde una perspectiva de la calidad del objetivo buscado, el aprendizaje:

*Proceso de enseñanza-aprendizaje, orientado a la adquisición de una serie de competencias y destrezas por parte del estudiante, caracterizado por el uso de las tecnologías basadas en web, la secuenciación de unos contenidos estructurados según estrategias preestablecidas a la vez que flexibles, la interacción con la red de estudiantes y tutores y unos mecanismos adecuados de evaluación, tanto del aprendizaje resultante como de la intervención formativa en su conjunto, en un ambiente de trabajo colaborativo de*

*presencialidad diferida en espacio y tiempo, y enriquecido por un conjunto de servicios de valor añadido que la tecnología puede aportar para lograr la máxima interacción, garantizando así la más alta calidad en el proceso de enseñanza-aprendizaje (Francisco José García-Peñalvo, 2008).*

En esta definición, el concepto de calidad en *eLearning*, en cuanto que hace referencia a un contexto formativo complejo, depende de estos cinco factores: tecnología, servicios, evaluación/acreditación, contenidos y factor humano (tutoría). Siguiendo esta idea García Peñalvo and Seoane Pardo (2015) presentan una definición más actual:

*El proceso formativo, de naturaleza intencional o no intencional, orientado a la adquisición de una serie de competencias y destrezas en un contexto social, que se desarrolla en un ecosistema tecnológico en el que interactúan diferentes perfiles de usuarios que comparten contenidos, actividades y experiencias y que, en situaciones de aprendizaje formal, debe ser tutelado por actores docentes cuya actividad contribuya a garantizar la calidad de todos los factores involucrados (García Peñalvo & Seoane Pardo, 2015).*

Por último una definición igual de actual que la anterior, pero que no toma en cuenta el factor ni humano, ni el de la evaluación es la de Khan y Ally:

*El eLearning puede ser visto como un enfoque innovador para la entrega de material educativo bien diseñado, centrado en el alumno, interactivo y facilitando ambientes de aprendizaje a cualquier persona, en cualquier lugar, en cualquier momento mediante la utilización de los atributos y recursos de diversas tecnologías digitales, junto con otras formas de materiales de aprendizaje adecuados para el entorno de aprendizaje abierto y distribuido (del inglés ELearning can be viewed as an innovative approach for delivering well-designed, learner-centered, interactive, and facilitated learning environments to anyone, anyplace, anytime by utilizing the attributes and resources of various digital technologies along with other forms of learning materials suited for the open and distributed learning environment) (B. H. Khan & Ally, 2015).*

Por tanto, a partir de las definiciones de *eLearning*, reconoce los retos que plantea la diversidad de estudiantes y profesores. De hecho, el *eLearning* se extiende los paradigmas tradicionales de aprendizaje en nuevos modelos de aprendizaje dinámico a través de las tecnologías informáticas y de Internet. Sin embargo, solo una pequeña cantidad de literatura *eLearning* evalúa tanto en los instructores y en los estudiantes actitudes hacia el uso de *eLearning* como la enseñanza y el aprendizaje de las herramientas.

De igual forma se puede concluir de estas definiciones que los tres grandes componentes de los procesos de *eLearning* son:

- Las tecnologías de soporte a los procesos de aprendizaje (redes, *hardware*, *software* y herramientas en forma general).
- Los contenidos o elementos contenedores de información.
- Las personas que interactúan en el proceso de aprendizaje y de soporte al aprendizaje.

En la práctica, para llevar a cabo un programa de formación basado en *eLearning*, se hace uso de plataformas o sistemas de *software* que permiten la comunicación e interacción entre profesores, estudiantes y contenidos. Se tienen principalmente dos tipos de plataformas: las que se utilizan para impartir y dar seguimiento administrativo en los LMS y, por otro lado, las que se utilizan para la gestión de los contenidos digitales, los LCMS.

### II.3 Sistemas de administración de aprendizaje

Los Sistemas de Administración de Aprendizaje son las herramientas desarrolladas para dar soporte al aprendizaje electrónico. Se trata de programas de *software* orientados tanto a la administración como a la docencia. La parte administrativa de un LMS está orientada, entre otras cosas, a gestionar usuarios, recursos como materiales, vinculaciones, proyectos, tareas, administrar accesos, generar informes y gestionar servicios de comunicación como foros de discusión, videoconferencias. La parte docente incluye listados de los cursos ofrecidos, material para los cursos, ejercicios, apuntes, entregables y evaluaciones.

En esencia, los LMS son sistemas de administración/gestión de contenidos o CMS. Un CMS consiste en una interfaz que controla una o varias bases de datos donde se aloja el contenido del sitio. El sistema permite manejar de manera independiente el contenido y el diseño. Así, es posible manejar el contenido y darle en cualquier momento un diseño distinto al sitio sin tener que darle formato al contenido de nuevo, además de permitir la publicación fácil y controlada en el sitio a varios editores. Así que un LMS es un CMS con una interfaz de usuario orientada a la docencia.

Estas plataformas ofrecen ambientes de aprendizaje ya diseñados e integrados donde los diferentes actores acceden a él y se autentican mediante una clave personal, dando lugar a espacios privados, dotados de las herramientas necesarias para aprender (comunicación, documentación, contenidos, interacción, etc.). Estas herramientas permiten que, a grandes rasgos:

- Los profesores coloquen a disposición de los estudiantes la especificación de los cursos.
- Los tutores supervisen el desarrollo del curso y el avance de cada estudiante.
- Los estudiantes accedan a los contenidos, realicen las actividades previstas, se comuniquen entre sí y con el tutor para resolver dudas y realizar trabajos en grupo.
- Los administradores obtengan información en línea del progreso del curso y de las acciones administrativas relacionadas, tales como inscripción de estudiantes, historial de cursos, etc.

Para poder cumplir con su propósito todo LMS posee un conjunto mínimo de herramientas, entre las que cabe destacar las herramientas de distribución de contenidos, las herramientas de comunicación y colaboración (con especial mención a los foros como herramienta central de un sistema *eLearning* para la comunicación y la colaboración asíncronas),

herramientas de seguimiento y evaluación, y herramientas de administración y asignación de permisos. De la experiencia del uso de las plataformas se derivan ciertas carencias destacables. La mayor limitación de las plataformas tecnológicas es su propia definición. Son sistemas de gestión del aprendizaje sin más, que en ocasiones no incorporan otros aspectos interesantes como una adecuada gestión de los contenidos y, en general, funcionalidades atribuidas a los CMS.

A efecto de solventar esta separación surgen los denominados LCMS. Se trata de sistemas de Gestión de Contenidos de Aprendizaje. Una vez que los contenidos están en este sistema ya pueden ser combinados, asignados a distintos cursos, etc., de forma que la reutilización de contenidos educativos se convierte en su principal cometido. Los LCMS se basan en un modelo de objetos de contenido u objetos de aprendizaje. Un objeto de aprendizaje se puede definir de forma genérica como "cualquier recurso digital que puede ser reutilizado para apoyar el aprendizaje" (Wiley, 2002) o "como una entidad, digital o no digital, que puede ser utilizada, reutilizada y referenciada durante el aprendizaje apoyado con tecnología" (LTSC, 2009).

Como el Dr. Charles Severance (fundador de la *Open Source LMS Sakai*) establece que los LMS "están lo suficiente maduros como para que la mayoría de las facultades y estudiantes que los hayan usado, a pesar del sistema que han elegido, hayan quedado satisfechos" (Alier Forment et al., 2010).

Los LMS, como se ha comentado con anterioridad, son algunas de las herramientas más representativas del aprendizaje *online* y han sido ampliamente adoptadas por diferentes instituciones, universidades y empresas. Sin embargo, a pesar de esa aceptación, los LMS no han conseguido las mejoras esperadas en los procesos de aprendizaje, debido a:

- 1) No se aprovechan convenientemente las herramientas provistas y se utilizan como meros espacios para la publicación de los cursos (Alier Forment, Casany Guerrero, Conde González, García Peñalvo, & Severance, 2009; Conole & Alevizou, 2010; Cuban, 2003).
- 2) Los LMS restringen las posibilidades del estudiante para colaborar en el aprendizaje y fomentar un constructivismo social no limitado a un período específico de tiempo (por ejemplo, curso académico) (J. S. Brown & Adler, 2008; Wesch, 2009).
- 3) Se centran en el curso y en la institución, no en el estudiante en sí (Downes, 2005).

Además de esto, debe considerarse que el aprendizaje *online* no acaba en los LMS, sino que hay multitud de herramientas en línea para completarlo y mejorarlo. Herramientas que actúan como fuentes de información, como medios de comunicación, como puntos de intercambio de experiencias, etc. Debe, por tanto, hacerse sitio a las aplicaciones de búsqueda, de noticias, de localización, los repositorios documentales, los foros, los *blogs*, calendarios, juegos, mundos virtuales, etc. Es decir, a las nuevas iniciativas derivadas de la Web 2.0 (Oreilly, 2007).

Queda claro ante esta situación que es necesario que los LMS evolucionen, pero, ¿hacia dónde deben hacerlo? La evolución de los LMS tiene que hacerse hacia el usuario y, por tanto, hacia sistemas personalizados, estos sistemas pueden ser los PLE pero, para incorporar la nueva generación de aplicaciones de apoyo al aprendizaje dentro de los LMS hay una necesidad de interoperabilidad entre sistemas para poder utilizarlos, es necesario tener en cuenta en primer lugar cómo van a abrirse los LMS, y esto se puede hacer mediante el uso de las Arquitecturas Orientadas a Servicios (SOA, *Service Oriented Architecture*) (Al-Ajlan & Zedan, 2008; Casany, Alier, Conde, & García, 2009; Conde González, Gómez Aguilar, et al., 2011; Dagger et al., 2007; Pätzold et al., 2008), que son descritos a continuación.

Este es un enfoque de Ingeniería de *Software* que proporciona una separación entre la interfaz de un servicio y su implementación subyacente. Para aplicaciones de consumo de los servicios, no importa cómo se implementan los servicios, cómo se almacenan los contenidos y la forma en que están estructurados. En las aplicaciones de consumo enfoque SOA puede inter-operar a través del más amplio conjunto de proveedores de servicios (implementaciones), y los proveedores pueden ser fácilmente intercambiados en la marcha sin modificación de código de la aplicación.

Actualmente hay una tendencia hacia la modularización de los sistemas informáticos. Esto es debido a las ventajas que incorpora, como pueden ser la independencia de desarrollo y evolución, aumento de la seguridad, desarrollo escalable, ayudando así a la reutilización, adaptación y escalabilidad de los sistemas, etc. Por otra parte, se está trabajando en la producción de *software* de servicios independientes de la implementación subyacente.

Uniendo ambas concepciones surgen las arquitecturas SOA. Entre los elementos que fomentaron el desarrollo de las arquitecturas SOA se encuentran el desarrollo de diversos tipos de aplicaciones, las redes de ordenadores, las arquitecturas cliente-servidor, etc. (Ramaratnam, 2007). Las arquitecturas SOA suponen un paso más en los avances en arquitecturas de sistemas de información. En su nivel más básico, las arquitecturas orientadas a los servicios son una colección de servicios que se comunican entre ellos (Payne & Baroody, 2006).

SOA es un paradigma con el cual muchos sistemas se podrían beneficiar de las ventajas que conlleva. No obstante estos sistemas legados pueden ser adaptados a la arquitectura SOA a través de un proceso de *refactoring*. Los sistemas de *software* ya existentes deben ser constantemente mantenidos y sometidos a procesos de refactorización para preservar su valor de negocio, más aun teniendo en cuenta la rápida evolución de las y tecnologías *software* (Canfora, Fasolino, Frattolillo, & Tramontana, 2008; Di Lucca, Gold, & Visaggio, 2008)

El programador, diseñador, desarrollador principal y fundador de Moodle, Martin Dougiamas, asignó a principios de 2008 al equipo de la UPC, la tarea de desarrollar una nueva API (del inglés, *Application Programming Interface*) para acceder a los servicios del sistema Moodle, con independencia de su aplicación, que actualmente permanece estable en las versiones de Moodle. Esta tarea se describe en el gestor de Moodle (Moodle, 2010b) y en Moodle Docs (Moodle, 2010a).

Este API consiste en un conjunto de servicios web que encapsulan la mayoría de los servicios que un cliente o módulo externo (e incluso interno) a la aplicación necesitan de un servidor Moodle. En octubre de 2008, esta capa de servicios web se integró en la distribución estándar de Moodle 1.9.3 y es el estándar de interoperabilidad del subsistema hasta la fecha y de las futuras versiones de Moodle.

A modo de resumen, se ha mostrado una perspectiva histórica de cómo el aprendizaje ha ido evolucionando con el uso de las ICTs, posteriormente se ha definido el *eLearning* para describir como llega por medio de plataformas de administración las ICTs al estudiante, sus características y por último su medio de interconexión para aportar flexibilidad y apertura a estos por medio de los servicios web a nuevas herramientas que soportan el proceso de enseñanza y aprendizaje. Sin embargo, es preciso mencionar que todas estas plataformas y herramientas generan un cúmulo de información que se puede reutilizar en beneficio de la propia educación. En el siguiente capítulo se describirá el área de la analítica y como esta llega al área educativa con el objetivo de la mejora de esta.



## Capítulo III Analítica en la educación

“...If our aim is to understand people's behaviour rather than simple record it, we want to know about primary groups, neighbourhoods, organizations, social circles, and communities; about interaction, communication, role expectation, and social contract...”

— Allen Barton, 1968



En este capítulo se expone la definición de la analítica de manera general para continuar con la descripción de cómo se llega a integrar en la educación, con especial atención a los diferentes niveles y definiciones que se han dado en este contexto y cómo se encuentra actualmente definida como Analítica del Aprendizaje o Analítica Educativa (LA) y Analítica Académica (AA). Además, se presentan las definiciones de Minería de Datos para la Educación (del inglés, *Educational Data Mining*, EDM) y de Analítica Social para la Educación (del inglés, *Social Learning Analytics*, SLA), ya que mantienen una estrecha relación con el surgimiento de la AA y LA. Algunos autores sostienen que EDM y SLA son áreas separadas de LA y AA (Siemens & Baker, 2012), mientras que otros las consideran como técnicas o instrumentos utilizadas en las analíticas (Baepler & Murdoch, 2010; Elias, 2011; Greller & Drachsler, 2012) y, finalmente, otros las consideran como sinónimos (Bach, 2010). Parte importante del proceso de la analítica es el modelo de referencia de LA y la estructura básica de su proceso, el cual se considera en este capítulo. Por último, el hecho de que tanto las personas que diseñan y construyen la analítica, como las personas que utilizan dichas herramientas o procesos de analítica, tengan las competencias necesarias para entender la importancia y explotación de los datos, es un factor importante en el cumplimiento del objetivo de la analítica misma. Este término se conoce como alfabetización de los datos y su definición se incluye en este capítulo.

### III.1 Analítica

John P. Campbell, DeBlois, and Oblinger (2007) exponen que la analítica está profundamente relacionada con grandes conjuntos de datos, técnicas estadísticas y modelos de predicción, de forma que podría pensarse como la práctica de la minería de datos institucionales para producir "inteligencia procesable".

En la actualidad el término "analítica" se aplica de forma frecuente a los esfuerzos en la explotación de datos de diversas fuentes para ayudar a las organizaciones a ser más eficaces. Así, un número cada vez mayor de diferentes industrias están desarrollando esta estrategia para lograrlo (Thomas H. Davenport, Harris, & Morison, 2010). Bichsel (2012b), además de colocar el énfasis en la predicción y siguiendo esta idea, define la analítica como:

*El uso de los datos, el análisis estadístico y los modelos explicativos y predictivos para obtener información y actuar sobre temas complejos (Bichsel, 2012b).*

Los añadidos sugeridos con mayor frecuencia a su definición fueron que la analítica es "estratégica" y participa en la "toma de decisiones". Bichsel (2012b) menciona que algunos sugirieren que se deben mencionar los tipos de datos e informes que participan con la analítica, y sugiere que la "visualización de datos" debe ser parte de su definición. Se puede hablar de la analítica como un proceso que contiene un número de pasos a seguir; de esta forma Bichsel (2012b) describe ese proceso como:

- 1) Comenzar con una cuestión/pregunta estratégica.
- 2) Encontrar o recolectar los datos adecuados para responder a dicha pregunta.
- 3) Analizar los datos con miras a la predicción y al descubrimiento de nuevo conocimiento.
- 4) Representar los resultados de manera que sean comprensibles y aplicables.
- 5) Realimentar el proceso para abordar la pregunta y cuestiones estratégicas y para crear otras nuevas.

La analítica puede ser fuente de información para la toma de decisiones humanas o puede soportar la toma de decisiones de forma automatizada. La analítica es subconjunto de la inteligencia de negocios (del inglés, *Business Intelligence*, BI), un conjunto de tecnologías y procesos que utilizan datos para comprender y analizar el rendimiento del negocio (Thomas H. Davenport, Harris, & Shapiro, 2010).

La necesidad de utilizar analítica ha surgido en el momento en que las empresas han ofrecido productos similares y han usado tecnología comparable, de manera que los procesos de negocio de alto rendimiento se convierten en algunos de los últimos puntos en los que se puede conseguir una diferenciación. Muchas de las bases previas para la competencia ya no están disponibles, como la ventaja geográfica única que ya no importa en la competencia mundial, y la regulación proactiva ha desaparecido en gran medida.

Las tecnologías patentadas se copian rápidamente, y romper la innovación en productos o servicios parece cada vez más difícil de lograr. Lo que queda como una base para la competencia consiste en ejecutar su negocio con la máxima eficiencia y eficacia, así como tomar las decisiones de negocio lo más inteligente posible. Entonces la analítica en la competencia se ejerce al exprimir hasta la última gota de valor a partir de los procesos de negocio y las decisiones clave y se define como una organización que utiliza la analítica amplia y sistemáticamente para ejercer la competencia (Thomas H. Davenport et al., 2010).

Desde este punto de vista empresarial, Thomas H. Davenport (2006) menciona que las compañías y organizaciones están sacando partido a sus bases de datos. Tras acumular dichos datos durante años y darse cuenta de que necesitaban mejorar su capacidad transaccional, muchas organizaciones han llegado a un punto en el que han pensado (Thomas H. Davenport, 2006):

*De acuerdo, no tenemos más excusas; tenemos que empezar a gestionar nuestro negocio de una manera diferente, basándonos en nuestros datos. La segunda conclusión es que toda compañía puede mejorar sus capacidades analíticas. El análisis avanzado puede llegar a ser una parte integral de la estrategia de negocio. El tercer elemento que se necesita además de datos y tecnología para llegar a competir en análisis avanzado, y que son los factores diferenciales, tienden a ser humanos, como por ejemplo: el liderazgo, la pasión, las habilidades y las relaciones; aspectos que, creo, han sido desatendidos en las investigaciones sobre inteligencia de negocios, también llamada analítica, de los últimos años. Las compañías que cuentan con una alta sofisticación analítica suelen contar con empleados altamente cualificados trabajando para ellas. Esto no creo que sea casual (Thomas H. Davenport, 2006).*

El potencial de la analítica de acuerdo con estas definiciones es que ayuda al evaluar las acciones pasadas para estimar el potencial de las acciones futuras, con las cuales tomar mejores decisiones y adoptar estrategias más eficaces, ya sea a nivel organizacional o individual.

Un factor menos popularizado a la hora de realizar la explotación efectiva de la analítica es la rica variedad y madurez de las técnicas de análisis de datos. Un analista experto tiene ahora muchas disciplinas en las que inspirarse y muchas herramientas al alcance para usarlas.

La diversidad y la flexibilidad de algunas de las técnicas que se alinearon bajo la bandera analítica se evidencia por las numerosas aplicaciones: los mercados financieros, análisis de los deportes, econometría, precios de los productos y la maximización del rendimiento, de fraude, de detección de la delincuencia, filtros de correo electrónico de (*spam*), comercialización, segmentación de clientes, la eficiencia organizacional. Detrás de estas aplicaciones se pueden encontrar las raíces de la analítica en el nacimiento de las estadísticas en el siglo XVIII, pero desde entonces las diferentes aplicaciones de la estadística y de las ICT han dado lugar a diferentes comunidades de práctica que ahora parecen estar fusionándose (Cooper, 2012).

	Pasado	Presente	Futuro
Información	¿Qué sucedió? <b>Reporte y Descripción</b>	¿Qué está sucediendo ahora? <b>Alerta</b>	¿Qué sucederá? <b>Extrapolación</b>
Conocimiento	¿Cómo y por qué sucedió? <b>Modelos y Explicación</b>	¿Cuál es la mejor acción a seguir? <b>Recomendaciones</b>	¿Qué es lo mejor o peor que puede suceder? <b>Predicción</b>

Tabla 1. Preguntas cuyas respuestas son fuentes de información y conocimiento, mediante un análisis direccional (Thomas H. Davenport et al., 2010).

Para organizar los actuales y diversos enfoques que existen de la analítica, a pesar de que las preguntas clave que la analítica puede abordar son innumerables, se han definido una serie de preguntas clave para esquematizar los enfoques existentes. Estas preguntas esquematizan los diferentes enfoques de acuerdo a la dimensión temporal de las preguntas, así como su nivel de aportación de conocimiento. Estos aspectos y las preguntas claves se resumen en la matriz propuesta en (Thomas H. Davenport et al., 2010), que se recoge en la Tabla 1, que distingue entre enfoques más basados en la realidad (fila de la *información*, del inglés "*information*") y enfoques que están más orientados hacia la comprensión más profunda (la fila *conocimiento*, del inglés "*insight*") y los segmentos de estos de acuerdo a un marco temporal.

Se puede notar que los pioneros de analítica de la Web (del inglés, *Web Analytics*) (Burby & Atchison, 2007; Kaushik, 2007) están explotando los datos de la "Web Social" (Gruber, 2008) mediante el uso de análisis de redes sociales (Margaix Arnal, 2008; Wasserman & Faust, 1994) y apoyados con representaciones interactivas (Carlis & Konstan, 1998) y exploratorias (D. A. Keim, Mansmann, Schneidewind, & Ziegler, 2006; Kumar & Garland, 2006) de análisis (Stuart K. Card, Mackinlay, & Shneiderman, 1999; D. A. Keim, Mansmann, & Thomas, 2010) que forman parte de las técnicas de visualización de información (Ghanbari, 2007; B. Shneiderman, 1996; Spence, 2000; Ware, 2004) y no solo se trata de los gráficos en los informes de gestión.

Los campos que algunos ven anticuados, como por ejemplo la investigación operativa, y otros que a menudo se perciben tan futuristas como inteligencia artificial, están haciendo contribuciones de formas sorprendentes. Mientras tanto, la comunidad educativa ha hecho sus propias contribuciones; el análisis de redes sociales e inteligencia artificial, ambos han surgido de la investigación académica y ahora se está empezando a ver las variantes específicas del sector de la analítica que están siendo utilizados bajo la forma de la EDM, analítica del aprendizaje y bibliometría.

En las siguientes secciones se desarrollan los términos relacionados con la analítica en la educación pero antes, en el trabajo de Cooper (2012), se describen algunas de las comunidades de la analítica más importantes.

Estas comunidades, como son la estadística, inteligencia de negocios, investigación operativa, inteligencia artificial, analítica de la Web, etc., cada uno con sus propios orígenes, las técnicas, las zonas de limitación y preguntas claves (antes mencionadas) y sugiere cómo están contribuyendo al futuro, con especial referencia al contexto de la educación.

Además de las definiciones establecidas en el trabajo anteriormente mencionado, donde la mayoría de ellas tienen que ver con el análisis, en lugar de análisis académico o educativo, cabe mencionar, que la literatura se ocupa de una amplia variedad de definiciones de analítica; documentos enteros se han escrito sobre la terminología, véase por ejemplo Van Barneveld et al. (2012). En este trabajo un amplio y variado conjunto de definiciones funcionales de la analítica del aprendizaje y de los términos relacionados se observan y presentan el uso de la analítica, de la analítica de los negocios, analítica académica, analítica del aprendizaje usada en la academia, analítica del aprendizaje usada en los negocios, analítica predictiva y la analítica de la acción (ver la Tabla 2).

Término	Definiciones	Nivel
<b>Analítica</b>	“Son los procesos de evaluación y análisis de datos que nos permiten medir, mejorar y comparar el desempeño de individuos, programas, instituciones o empresas, grupos de organizaciones y/o industrias enteras” (D. M. Norris, Baer, & Offerman, 2009a)	Institución Departamento Instructor Estudiante
	“Es el proceso de toma de decisiones basada en datos, utilizado para comunicar decisiones a todos los niveles de la empresa” (Ravishanker, 2011)	Institución Departamento Instructor Estudiante
<b>Analítica de negocios</b>	“La categoría entera se refiere simplemente a utilizar datos y análisis para entender y administrar los negocios efectivamente, esto, en contraposición a la simple acción de capturar las direcciones de los consumidores o mantener un seguimiento de los balances vacacionales de los empleados o ese tipo de transacciones” (Hopkins, 2010)	Institución (Empresa)
<b>Analítica de negocios</b>	“La analítica de negocios ( <i>Business Analytics</i> , BA) es la práctica de la exploración metódica e iterativa de los datos de una organización haciendo énfasis en el análisis estadístico. BA se utiliza en compañías comprometidas con la toma de decisiones basadas en datos. BA se utiliza para ganar conocimiento a partir de información relacionada con decisiones de negocios y puede ser utilizada para automatizar y optimizar procesos de negocios. Las compañías orientadas al manejo de información tratan sus datos como un activo común y lo potencian como una ventaja competitiva” (Rouse, 2010)	Institución (Empresa)
<b>Analítica académica</b>	“Las iniciativas tempranas de analítica académica buscan predecir e identificar las dificultades académicas en los estudiantes, permitiendo a los profesores y tutores actualizar o reformular los métodos de aprendizaje-enseñanza o proveer instrucciones dirigidas a necesidades específicas de aprendizaje” (John P. Campbell et al., 2007)	Estudiante
	“Se enfoca en aspectos académicos, principalmente el acceso a los estudiantes, asequibilidad y éxito” (D. M. Norris et al., 2009a)	Estudiante
	“Es el término imperfecto equivalente a Inteligencia de Negocios ( <i>Business Intelligence</i> , BI), el cual esencialmente describe el uso de tecnología de la información para apoyar la toma de decisiones operacional o financiera”( Philip J. Goldstein & Richard N. Katz, 2005)	Institución

Término	Definiciones	Nivel
<b>Analítica académica</b>	“Consiste en combinar datos con técnicas estadísticas y modelos predictivos para apoyar a los profesores y tutores en el proceso de determinar que estudiantes enfrentan dificultades académicas, permitiéndoles intervenir para ayudarlos a alcanzar resultados exitosos” (John P. Campbell & Oblinger, 2007)	Instructor Estudiante
	“Es la minería de datos en sistemas que apoyan la enseñanza y el aprendizaje para aportar actualización, tutorial o intervención dentro del ambiente de aprendizaje” (John P. Campbell, 2007)	Instructor Estudiante
	“El proceso de proveer a las instituciones de educación superior con los datos necesarios para responder a los desafíos de reporte y toma de decisiones que enfrentan las universidades contemporáneas” (Dawson, McWilliam, & Pei-Ling Tan, 2008)	Institución Departamento
	“Se puede referir de manera amplia a la toma de decisiones basadas en datos a nivel universitario o de instituto con propósitos operacionales, pero también puede ser aplicado a aspectos de enseñanza y aprendizaje en estudiantes” (Baepler & Murdoch, 2010)	Institución Instructor Estudiante
	“Puede identificar e incluso predecir cuáles estudiantes se encuentran en riesgo en un curso particular”( Fritz, 2009)	Instructor Estudiante
<b>Analítica de aprendizaje— Academia</b>	“Es la interpretación de un amplio rango de datos producidos y recopilados sobre un grupo de estudiantes con el propósito de evaluar el progreso académico, predecir su desempeño futuro y detectar diferentes aspectos potenciales”(Johnson, Adams, & Cummins, 2012)	Institución Estudiante
	“Es el uso de modelado predictivo y otras técnicas de analítica avanzada para mejorar los recursos instruccionales, curriculares y de apoyo con el fin de alcanzar objetivos específicos de aprendizaje.” (Bach, 2010) (n.b. Bach presenta la analítica de aprendizaje y la analítica académica como sinónimos)	Departamento Estudiante
	“Ofrece la posibilidad a los estudiantes y escuelas de adaptar sus oportunidades educacionales al nivel de cada estudiante en función de su habilidad y necesidades.” (Johnson et al., 2012)	Estudiante
	“Puede ser utilizado para evaluar currículo, programas e instituciones” (Johnson et al., 2012)	Institución Departamento
<b>Analítica de aprendizaje— Academia</b>	“Es el uso de modelos y datos para predecir el progreso y desempeño de los estudiantes y la habilidad para tomar acciones basadas en esa información” (Analytics, 2013)	Estudiante
	“Es la recopilación y análisis de datos de uso asociados con el aprendizaje de estudiantes para observar y entender conductas y comportamientos de aprendizaje con el propósito de facilitar una intervención adecuada” (M. Brown, 2011)	Estudiante

Término	Definiciones	Nivel
<b>Analítica de Aprendizaje— Industria</b>	“El estudio del impacto de la enseñanza y aprendizaje en quienes lo reciben” (Moore, 2005)	Estudiante
	“Recopilación de entradas de múltiples bases de datos, cuando en conjunto con las interrogantes adecuadas pueden, extraer datos y crear una representación en tiempo real de las métricas de entrenamiento de una organización” (Hartley, 2004)	Institución Departamento
	“Un conjunto de actividades que una organización realiza para apoyar el entendimiento de las mejores estrategias para entrenar y desarrollar empleados y consumidores” (Berk, 2004)	Institución Departamento
<b>Analítica predictiva</b>	“Conjunto de [inteligencia empresarial] tecnologías que descubre las relaciones y patrones dentro de grandes volúmenes de datos que pueden ser utilizados para predecir el comportamiento y eventos [...] de análisis predictivo tiene visión de futuro, mediante el uso de eventos pasados para anticipar el futuro” (Eckersen, 2007)	Institución Departamento Instructor Estudiante
	“La analítica predictiva conecta datos para actuar efectivamente estableciendo conclusiones valiosas sobre condiciones actuales y eventos futuros” (IBM, 2010)	Institución Departamento Instructor
	“La analítica predictiva [...] es a la vez un proceso de negocios y un conjunto de tecnologías interrelacionadas. La analítica predictiva potencia el conocimiento de negocios de una organización aplicando técnicas de análisis sofisticadas y datos empresariales. El conocimiento resultante puede llevar a acciones que comprobablemente demuestran cómo las personas se comportar como consumidores, empleados, pacientes, estudiantes y estudiantes” (IBM, 2009)	Institución Departamento Instructor Estudiante
<b>Analítica de Acciones (AcA)</b>	“Más comprehensiva que la analítica académica, abarcando procesos académicos y administrativos, y reconociendo la necesidad de una reinversión [...] enfocada más ampliamente en la productividad y desempeño académico y administrativo” (D. M. Norris, Baer, & Offerman, 2009b)	Institución
	“El término analítica de acciones se refiere a la analítica de habilidades y prácticas que son poderosas, inmediatas y útiles para una gran variedad de actores interesados. Pero, más importante aún, la analítica de acciones solo se lleva a cabo en empresas y ambientes que están genuinamente comprometidos con la medida y mejora de aspectos clave de productividad, innovación y desempeño. Estas organizaciones deben construir activamente sus capacidades y cultivar los comportamientos para alcanzarlos. Lograr una analítica de acciones es algo más orientado al liderazgo y produce cambios más significativos en la cultura organizacional y el comportamiento que la tecnología” (D. M. Norris et al., 2009b)	Institución
	“Es una fusión de nuevas herramientas analíticas con las expectativas crecientes para una mayor rendición de cuentas educativa..” (Schoenecker & Baer, 2010)	Institución

Tabla 2. Definiciones conceptuales y funcionales de diversos tipos de analítica adaptado de (Van Barneveld et al., 2012).

Por otro lado, Thomas H. Davenport et al. (2010) afirman que no en todas las organizaciones se va a utilizar el análisis como medio de diferenciación competitiva, pero cada organización puede beneficiarse de este al mejorar la forma de:

- Utilizar los datos para obtener una visión más profunda.
- Tomar decisiones más inteligentes.
- Ejecutar las decisiones más consistentes y obtener mejores resultados.

Thomas H. Davenport et al. (2010) declararon que aunque la mayoría de las empresas recopilan y almacenan enormes cantidades de datos, no la utilizan de forma efectiva. Algo muy similar puede decirse de la educación superior. Aunque se recogen y almacenan grandes cantidades de datos, los datos no se utilizan eficazmente para hacer predicciones o desencadenar respuestas proactivas efectivas.

En la década de los noventa, cuando los sistemas administrativos eran autónomos, se inició la evolución de la empresa y de los sistemas de enseñanza, esto dio como resultado la aparición de la analítica en la educación. Con el tiempo, los avances tecnológicos permitieron sistemas totalmente integrados, lo cual consintió una mayor eficiencia transaccional, integración de la información, la mejora de la retención de informes y análisis de negocio, inteligencia de negocios, así como el reclutamiento (D. M. Norris & Baer, 2009; D. M. Norris & Lefrere, 2010).

A partir de 2004, se produce la revolución de la llamada Web 2.0 o Web Social (Oreilly, 2007). Un fenómeno que, según algunos expertos, ya no constituye una revolución tecnológica, sino una revolución social, con el surgimiento del llamado *eLearning 2.0* (Downes, 2005). En este punto, la infraestructura tecnológica fusionada en red, *software* integrado (por ejemplo, los sistemas de planificación de recursos empresariales del inglés, *Enterprise Resource Planning*, ERP), la seguridad y los vastos recursos digitales fomentaron la creación de aplicaciones con la tecnología creciendo de forma sostenida, centrada en la seguridad, servicios web y servicios de red.

Los sistemas de enseñanza eran más lentos en desarrollarse, pero con el tiempo se desarrollaron sistemas de cursos de información, sistemas de gestión de cursos y los sistemas de gestión de aprendizaje. Esto permitió la convergencia entre los sistemas administrativos y gestores en el ámbito de la empresa. Las instituciones comenzaron a adoptar los sistemas de toda la empresa, incluyendo LMS y sistemas de información de los estudiantes. Portales empresariales siguieron la evolución de la capacidad de acceder e integrar los LMS, los sistemas de planificación de recursos empresariales y los activos de conocimiento, creando una base de autoservicio para estudiantes, profesores y el personal (Downes, 2005).

En la siguiente sección se describe el acercamiento de la analítica a la educación y se define la analítica académica y, posteriormente, la analítica del aprendizaje.

### III.2 Analítica Académica

Durante los últimos quince años ha habido un énfasis de las ICT en el desarrollo de la infraestructura tecnológica. Hoy el énfasis va más allá de la infraestructura e incluye los procesos de negocio y los criterios estratégicos de forma y modo para llevarlos a cabo. Existe una gran continuidad dentro de las actividades en el ambiente de la analítica, que exige un mayor rendimiento y el demostrar resultados se ha trasladado al desarrollo de los *software* código abierto (del inglés, *open-source*)/arquitectura abierta (del inglés, *open-architecture*) y la capacidad de aprovechar la nube.

El aprendizaje en línea ha llegado progresivamente a ser considerado como equivalente o incluso superior en algunos aspectos al tradicional, el aprendizaje cara a cara. Versiones transformadas de aprendizaje en línea, mixto e *eLearning* tienen el potencial de ser elementos esenciales para reimaginar la educación superior y hacerla sostenible, teniendo como resultado aumentar el rendimiento y la eficiencia institucional (D. M. Norris & Baer, 2009), reducir el costo marginal de la matrícula de manera significativa.

La demanda de plazas se estimula aún más si estas transformaciones abordan la facilidad de aprendizaje, reducen el tiempo para la competencia y fortalecer los vínculos entre las experiencias de aprendizaje y competencias auténticas para el lugar de trabajo. La capacidad de hacer esto se ha mejorado con el aprendizaje en línea y herramientas de la Web 2.0 y los patrones de interactividad. Esto ha tenido como resultado la primera generación de carteras de estudiantes/instituciones, cuadros de mando ejecutivos y sistemas de gestión de evaluación/rendimiento en un ambiente de arquitectura de código abierto (D. M. Norris & Lefrere, 2010).

Además, en las versiones más recientes de LMS 2.0 (Baer & Campbell, 2012) se han consolidado opciones LMS de campus. Incluso, se han desarrollado infraestructuras de base de datos incluyendo los almacenes de datos (del inglés, *Data Warehouse, DW*). Los esfuerzos anteriores de las entidades en la captura de datos, proporcionando disponibilidad en los DW, y los esfuerzos iniciales de minería de datos (del inglés, *Data Mining, DM*) son fundamentales para la generación siguiente de actividades de analítica.

La educación superior se puede beneficiar de los avances del *Business Intelligence* (BI) en el mundo corporativo para desarrollar nuevas soluciones integradas dentro del ámbito del aprendizaje, como son las plataformas LMS, SIS, etc. Se estima que la analítica será un importante apoyo para la educación superior del futuro (Bichsel, 2012b; Börner, 2012; M. Brown, 2012; Diaz & Brown, 2012; J. Foley, Card, Ebert, MacEachren, & Ribarsky, 2006; Johnson et al., 2012; LMTF, 2013; NGLC Insights, 2013; Van Harmelen & Workman, 2012).

La aparición de la analítica en la educación proviene de una evolución de sistemas empresariales y de la educación en conjunto. Un ejemplo de los primeros proyectos que analizan el impacto en el éxito del estudiante fue el realizado por (John P. Campbell et al., 2007), estos autores citan que la respuesta de las instituciones "a las presiones internas y externas para la entrega de resultados en la educación superior, especialmente en las áreas

de mejora de los resultados del aprendizaje y el éxito de los estudiantes, requerirá que los líderes de las ICT den un paso adelante y se conviertan en socios importantes en los asuntos académicos y estudiantiles”. Las ICT pueden ayudar a responder a la necesidad de ofrecer resultados y evidencias educativas a través de la analítica académica (AA) (Johnson, Smith, Willis, Levine, & Haywood, 2011), que se está convirtiendo en un componente fundamental de los entornos de aprendizaje de próxima generación.

En el estudio realizado por Bichsel (Bichsel, 2012a), donde se realizó una encuesta a una muestra de las instituciones miembros de EDUCAUSE, así como miembros de la Asociación para la Investigación Institucional (del inglés, *Association for Institutional Research*, AIR) y además, siete grupos constituidos por ambos profesionales, los de la tecnología de la información (del inglés, *Information Technology*, IT) y los investigadores de las instituciones (del inglés, *Institutional Research*, IR), fueron convocadas para tratar los temas de: definición de la analítica, identificación de retos, y proposición de soluciones.

Después de haber eliminado las encuestas incompletas o repetidas de los grupos seleccionados quedaron 339 encuestados. En este trabajo, gran parte de los encuestados identificaron las áreas de destino en la que sus instituciones están participando o intentan participar en la analítica; sin embargo, parece que la mayoría de las instituciones están en la etapa de recopilación de datos o control de datos en la mayoría de sus áreas y que aún no han madurado hasta el punto de llegar a utilizarlos para la predicción o la toma de decisiones. El hecho de que la recogida de datos se produzca ya en áreas de instituciones es un buen comienzo, pero existen barreras antes de llegar a la utilización de los datos de forma proactiva para tomar decisiones, cosa que es crucial para avanzar en la adopción de la analítica.

Las iniciativas actuales de la analítica pueden ayudar a unificar la institución, centrándose en estratégicas clave y la centralización de los datos; la integración de estos desde múltiples fuentes hace que sean más consistentes y aumenta su accesibilidad, visibilidad y utilidad.

Los IR y los profesionales de la IT tienen una gran cantidad de conocimientos sobre los posibles objetivos de la analítica, los beneficios de la analítica, y las barreras que existen en sus instituciones, ya sea para iniciar o ampliar los programas de analítica. Siempre que se cuente con los recursos adecuados, la formación y una cultura de la toma de decisiones basada en los datos que incluya el apoyo ejecutivo, podrán sobreponerse a estos obstáculos.

A medida que el interés por el análisis académico en la educación superior ha crecido, también lo han hecho las demandas de demostración de resultados que están exigiendo la medición del desempeño y la mejora de las intervenciones en los procesos educativos, para poder ir más allá de la evidencia o informar, y así poder ir a la decisión/acción. Sin embargo, el reto de tomar esta decisión/acción es complejo. La medición del rendimiento educativo incluye tanto el rendimiento operativo (sistemas administrativos y de apoyo) como el rendimiento académico (diseño y ejecución de estrategias académicas para lograr experiencias de aprendizaje, los resultados y las competencias de la vida real).

La mejora de forma generalizada del desempeño requiere la medición coordinada, la intervención y la acción a través de toda la educación y fuerza de trabajo en todo su espectro, es decir, esta mejora del rendimiento requiere de articulaciones y transiciones más eficaces entre las empresas de aprendizaje y entre el aprendizaje y el trabajo (D. Norris, Baer, Leonard, Pugliese, & Lefrere, 2008).

Hay una gran continuidad de las actividades en el ambiente de la analítica académica, Long and Siemens (2011) mencionan:

*La analítica abarca todo el ámbito y alcance de la actividad de la enseñanza superior, lo que afecta la administración, la investigación, la enseñanza y el aprendizaje, y los recursos de apoyo. El colegio / universidad por tanto debe convertirse en una organización más intencional, inteligente, con datos, pruebas y análisis que juegan un papel central en esta transición.*

John P. Campbell and Oblinger (2007) describen el proceso de la analítica académica como "un motor para tomar decisiones o acciones guiadas" (del inglés, "engine to make decisions or guide actions") que consiste en 5 pasos: capturar, informar, predecir, actuar y refinar.

La analítica académica se basa en la extracción de datos de una o más fuentes, tales como un CMS o un sistema de información del estudiante. Los datos, que pueden ser almacenados en un almacén de datos para su uso en un curso, programa educativo o campus universitario, se analizan utilizando *software* estadístico, y se genera un modelo matemático. Sobre la base del modelo y los valores predeterminados, se puede activar una acción particular, como, por ejemplo, el envío al estudiante de una notificación electrónica el inicio de una intervención personal del equipo de la institución educativa.

Dentro de la educación hay usos del análisis de datos que se encuentran fuera de la definición de aprendizaje centrado en el análisis del aprendizaje. Por ejemplo, el uso de análisis para seleccionar buenos candidatos de un grupo de solicitantes que deseen estudiar en un programa en particular. Esto es, se utiliza la analítica académica de forma que se pueda llegar a comprender la actividad, esta analítica que no es estrictamente analítica del aprendizaje, mientras se ayuda a que se cumplan los objetivos de las instituciones educativas (Van Harmelen & Workman, 2012).

En (Van Harmelen & Workman, 2012) se destaca también que la analítica existe como una parte de un sistema socio-tecnológico en donde la toma de decisiones y las consecuentes acciones son de igual importancia que los componentes técnicos de esta en las soluciones obtenidas de la analítica. Además, existen factores para la adopción de la analítica en la educación que son más de naturaleza organizacional y humana que técnica. El liderazgo y la organización cultural y sus habilidades son de gran importancia. También se incide en la importancia de mantener la perspectiva y no perderse en los datos, la tecnología, o en salidas de análisis sin sentido, para centrar los esfuerzos en la consecución de los objetivos sociales, pedagógicos o empresariales. Más allá de los datos, la tecnología y los requisitos estadísticos, los proyectos de analítica académica requieren, coincidiendo con lo ya mencionado sobre BI, habilidad y liderazgo: las tres características propias de los proyectos más exitosos de

analítica académica. Así, involucran (Philip J. Goldstein & Richard N. Katz, 2005):

**Gestores académicos, equipo de gobierno o cualquier persona con la capacidad de tomar decisiones educativas, es decir, los líderes en la toma de decisiones basada en la evidencia de los datos.**

Cualquier esfuerzo de AA comienza con unos líderes comprometidos en tomar las decisiones sobre la base de los datos institucionales adquiridos. La analítica puede utilizarse para examinar las cuestiones institucionales clave, tales como la inscripción o la retención, que por su naturaleza son complejas y, a menudo, sensibles, pero la decisión de seguir adelante con la analítica depende del conocimiento entre los administradores de alto nivel en una institución educativa.

**El personal experto en el análisis de datos.**

El segundo componente fundamental para la construcción de una iniciativa de AA es la dotación de personal. Los miembros del personal que participan en los esfuerzos de la analítica incluyen a menudo a administradores de bases de datos, investigadores institucionales, investigadores educativos, programadores y especialistas de dominio (por ejemplo, servicios estudiantiles, retención, desarrollo/promoción). Se puede necesitar personal informático y académico para recopilar información de diversos sistemas académicos como el CMS. El equipo debe tener la habilidad para construir modelos predictivos orientados por la investigación educativa basados en datos institucionales. Otros miembros del personal pueden necesitarse para centrarse en el desarrollo de políticas y aclarar quién tiene acceso a los datos, cómo se pueden utilizar los datos y qué modelos de datos de seguridad se requieren. Dado que la AA requiere el análisis de datos, las instituciones tendrán que invertir en la formación eficaz para obtener personal analítico cualificado. La obtención o el desarrollo de personal cualificado puede representar la mayor barrera y el mayor costo para cualquier iniciativa de AA. Si este tipo de personal se suma a las unidades institucionales de investigación existentes o se instruyen en la organización, departamento de asuntos estudiantiles o unidades académicas dependerá de la cultura organizacional y el lugar de los recursos.

**Una plataforma tecnológica flexible que esté habilitada para recoger, realizar minería y analizar los datos.**

El tercer elemento en cualquier proyecto de AA es la tecnología. Un almacén de datos es el componente clave de la infraestructura tecnológica para recopilar la información procedente de una variedad de fuentes en una estructura común que permita el análisis de datos. Para llenar este almacén, la institución tendrá que construir un "puente" entre la aplicación y el almacén.

El papel que puede desempeñar la analítica dentro del entorno de aprendizaje dependerá en gran medida de la visión de la institución. Esta visión puede ser la "inteligencia práctica", donde las herramientas y los datos reducen el riesgo de fracaso escolar y maximizan las probabilidades de éxito de los estudiantes. La ola inicial de las herramientas de AA busca mejorar la comprensión de las maneras en que los estudiantes, profesores y asesores pueden mejorar la retención y el éxito. Por mucho que los distintos sistemas de gestión de cursos surgieron en la década de los noventa, la educación superior está viendo la primera generación de modelos de predicción, herramientas de aprendizaje adaptativas, herramientas de alerta temprana, y las nuevas herramientas de visualización de datos para que los tomadores de decisiones puedan acceder y utilizar los datos de manera oportuna (Baer & Campbell, 2012).

Los líderes, gestores académicos o equipo de gobierno, deben crear una cultura institucional para utilizar las herramientas de análisis con el objetivo de maximizar el potencial de mejora de acceso de los estudiantes, el aprendizaje, la progresión y el éxito. Una institución debe considerar varios pasos clave para la adopción de la AA (Baer & Campbell, 2012):

- Identificar líderes de opinión para el uso de los datos para resolver los retos educativos.
- Desarrollar la capacidad de modelado predictivo existente y expandirse a través de los programas.
- Identificar qué datos son importantes y los indicadores utilizados para medirlos.
- Identificar las mejores herramientas en su clase de análisis y modelado predictivo, aplicaciones y procesos.
- Realizar cambios en la analítica de los procesos institucionales.
- Desarrollar agresivamente la capacidad de organización para el uso de análisis.
- Crear un plan de comunicación.

Los líderes deben tener en cuenta los métodos de ir más allá de informes y encontrar nuevas maneras de ayudar a los estudiantes de manera proactiva. Además de un enfoque programático, la institución debe identificar mecanismos en los que se dispone de datos para todas las áreas (Baer & Campbell, 2012). Aunque la necesidad de las habilidades técnicas de minería de datos y análisis estadístico es obvia, habilidades tales como el análisis de procesos, evaluación y diseño de instrucción son también esenciales para el proceso. La meta para la analítica debe seguir siendo el ir más allá de los datos y estadísticas y centrarse en cómo se debe utilizar la información. Para crear un plan de comunicación con el fin de sostener y ampliar el desarrollo de un programa de análisis, es imperativo que los líderes se comuniquen regularmente con las partes interesadas sobre el proceso y los resultados. La concentración y atención en cómo las estrategias analíticas afectan al rendimiento, la productividad y el valor serán la base para un modelo de aprendizaje de última generación.

La organización utiliza las herramientas de visualización para maximizar el mensaje al personalizar informes para los interesados y por los medios adecuados. La comunicación regular aumentará la confianza y en general el uso de los análisis, que puede construir la cultura de la evidencia y la investigación necesaria para mantener los esfuerzos.

### III.3 Analítica del Aprendizaje

La investigación en analítica de aprendizaje usa el análisis de datos para orientar decisiones tomadas en cada etapa del sistema educativo. Mientras que los analistas en el ámbito empresarial usan los datos de consumo para dirigirse a los consumidores potenciales y personalizar la publicidad, las analíticas de aprendizaje hacen uso de los datos de los estudiantes para construir mejores aproximaciones pedagógicas, centrarse en las poblaciones de estudiantes en riesgo, y evaluar qué tipos de programas han sido más efectivos y cuáles deberían mantenerse orientados a mejorar los porcentajes de retención del estudiante (Long & Siemens, 2011).

Van Harmelen and Workman (2012) mencionan que tal como se aplica la analítica en el ámbito educativo está proporcionando avances mensurables para el aprendizaje y la enseñanza, y ofrece la esperanza de la más conveniente toma de decisiones basadas en evidencia, la acción y la personalización en diversas áreas de la educación. En esto, los beneficiarios son los estudiantes, profesores, departamentos, instituciones educativas y los interesados a niveles regionales y nacionales. Además definen a la analítica del aprendizaje como:

*La aplicación de técnicas analíticas para el análisis de datos educativos, incluyendo datos sobre actividades para el estudiante y el maestro, para identificar patrones de comportamiento y proporcionar información útil para mejorar el aprendizaje y las actividades relacionadas con el aprendizaje (Van Harmelen & Workman, 2012).*

Por ejemplo, los datos extraídos de un sistema de información del estudiante proporcionan datos demográficos, el rendimiento académico y la información de aptitud. El CMS ofrece una visión de los esfuerzos relacionados con el curso del estudiante, proporcionando información de la interacción en tiempo real que permite la comparación con sus compañeros (autoevaluación, conocida en inglés como *self-assessment*) (Bromham & Oprandi, 2006; Govaerts et al., 2012; Mazzola & Mazza, 2012).

Las dos fuentes de datos se combinan para predecir la probabilidad de éxito de los estudiantes. Usando esta probabilidad, la institución puede decidir si tomar ciertas acciones, como invitar a un estudiante a una sesión de ayuda a través del correo electrónico o proponer a un estudiante una reunión con un asesor. La autorregulación del aprendizaje (del inglés, *Self-Regulated Learning*, SRL) no es una habilidad mental o una habilidad de rendimiento académico; sino que es el proceso de auto-directiva por la que los estudiantes transforman sus capacidades mentales en habilidades académicas.

Zimmerman enumera tres fases de SRL:

1. Fase de previsión, en esta se realiza el análisis y la automotivación, por tanto, es donde se crean los objetivos y estrategias de planificación.
2. Fase de control de rendimiento, la cual comprende el autocontrol y auto observación, y es donde se aplican estrategias cognitivas y de aprendizaje.
3. Auto-reflexión, comprende el auto juicio, la auto observación y la autoevaluación con el objetivo de realizar una auto reacción, en esta fase es donde se asigna la causa y el efecto de las acciones.

Zimmerman (2002) cree que todos pasan por estas fases a medida que se identifican los intereses personales, se administran los esfuerzos cognitivos para aprender, y se identifican las causas de las acciones.

El aprendizaje se percibe como una actividad que los estudiantes hacen por ellos mismos de una manera proactiva, más que como un evento encubierto que les ocurre en reacción a la enseñanza. La autorregulación se refiere a pensamientos auto-generados, sentimientos y comportamientos que están orientadas a alcanzar los objetivos (Farajollahi & Moenikia, 2011; Zimmerman, 2002).

Estos estudiantes son proactivos en sus esfuerzos por aprender, ya que son conscientes de sus fortalezas y limitaciones, y se guían por objetivos establecidos personalmente y estrategias relacionadas con la tarea. Estos estudiantes hacen seguimiento de su comportamiento en términos de sus objetivos y hacen auto-reflexión sobre su eficacia. Esto aumenta su autosatisfacción y motivación para seguir mejorando sus métodos de aprendizaje. A causa de su motivación superior y los métodos de aprendizaje adaptativo, los estudiantes autorregulados no solo son más propensos a tener éxito académico, sino a ver su futuro con más optimismo (Farajollahi & Moenikia, 2011; Zimmerman, 2002).

Chatti, Dyckhoff, Schroeder, and Thüs (2012) consideran que la analítica del aprendizaje (LA) se basa en áreas de investigación como la minería de datos para la educación (del inglés, *Educational Data Mining*, EDM) (Ryan S. J. d. Baker, 2011; Bhardwaj & Pal, 2012; Bienkowski, Feng, & Means, 2012; C. Romero & Ventura, 2010), sistemas de recomendación (del inglés, *Recommended Systems*), sistemas inteligentes de tutoría (del inglés, *Intelligent Tutoring System*, ITS), análisis de redes sociales (del inglés, *Social Network Analysis*, SNA) (Wasserman & Faust, 1994), entre otras. Además, considera LA como un término global genérico para describir un área de investigación del aprendizaje potenciado por la tecnología (del inglés, *Technology Enhanced Learning*, TEL) que se centra en el desarrollo de métodos de análisis y detección de patrones en los datos recogidos de los centros educativos y aprovecha los métodos para apoyar la experiencia de aprendizaje. En su trabajo propone la siguiente descripción del proceso de analítica del aprendizaje.

### III.3.1 Proceso de la Analítica del Aprendizaje

En el proceso general de analítica, el procedimiento se inicia con la captura de datos sin sentido, los cuales después se transforman como información para permitir predicciones basadas en el conocimiento y, finalmente, tomar una decisión adecuada. Los datos son la base de cualquier análisis y, por tanto, para el caso de LA los datos educativos son muy importantes.

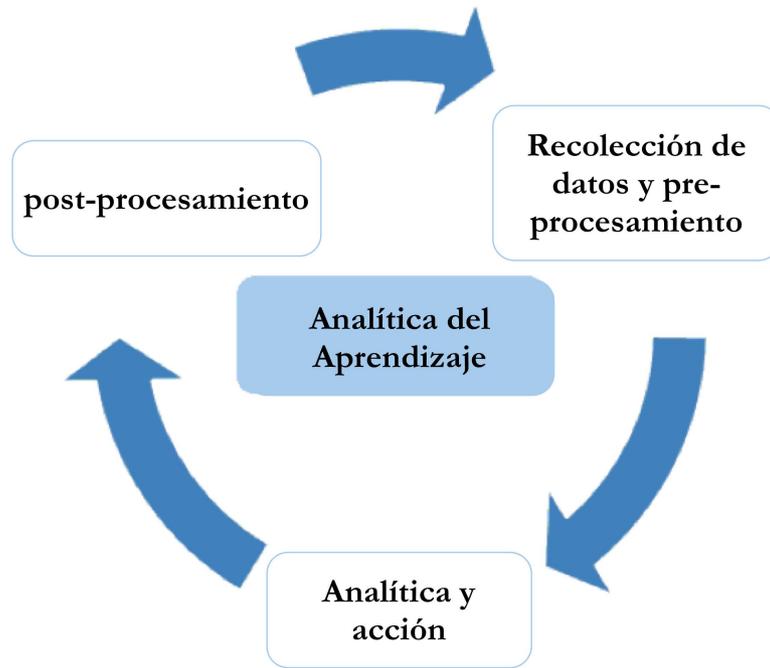


Fig. 2. Proceso de la analítica del aprendizaje, modificado de (Chatti et al., 2012).

Como se ilustra en la Fig. 2, el proceso general de LA es un proceso iterativo que generalmente consta de tres pasos, que se pueden encontrar también en procesos de EDM (E. García et al., 2010; Cristóbal Romero, Ventura, & García, 2008), VA (D. Keim, J. r. Kohlhammer, G. Ellis, & F. Mansmann, 2010), *Data Mining* (Liu, 2006), VA (D. Keim et al., 2010). Estos pasos, que se describen a continuación, son la recolección de datos y pre-procesado, la analítica y la acción y, por último, el post-procesado de los datos.

#### Recolección de datos y pre-procesamiento.

El primer paso de LA es la recolección de datos de diferentes entornos y sistemas educativos. Estos datos pueden ser demasiado grandes y/o contener muchos atributos irrelevantes para el estudio y, por tanto, se requiere de un pre-procesamiento de los datos (también llamado como preparación de los datos). Este proceso permite la transformación de los datos en un formato y estructura adecuados para ser usados como datos de entrada en determinados métodos de LA.

Durante este proceso se pueden realizar varios y diversos pasos, como la limpieza, integración de datos, reducción de datos, modelado de datos, identificación de usuarios y sesiones, entre otros (Han & Kamber, 2006; Liu, 2006; C. Romero & Ventura, 2007). La adición de la etapa final de refinamiento se reconoce como un proyecto de auto mejora en el que el seguimiento del impacto de las decisiones tomadas es un esfuerzo continuo y los modelos estadísticos deben actualizarse continuamente (Eliás, 2011).

### **Analítica y acción.**

Sobre la base de los datos pre-procesados y en el objetivo del ejercicio de análisis, se pueden aplicar diferentes técnicas de LA para explorar los datos con la intención de descubrir patrones ocultos que pueden proveer nuevas pautas para tener una experiencia de aprendizaje más efectiva. Dentro de las técnicas de la fase de analítica, no solo se incluye la visualización y análisis de la información, sino también las acciones que se pueden tomar con respecto a esta información. La adopción de medidas es el objetivo principal de todo proceso de análisis. Dentro de las acciones se incluye, el seguimiento, el análisis, la predicción, la intervención, la evaluación, la adaptación, la personalización, la recomendación y la reflexión (Bienkowski et al., 2012).

### **Post-procesamiento.**

Para lograr la mejora continua del ejercicio de analítica, el post-procesamiento es una pieza fundamental. Este además de refinar el conjunto de datos de salida del proceso puede implicar nuevos datos obtenidos de fuentes externas adicionales, la determinación de nuevos atributos necesarios para la siguiente interacción del proceso y la determinación de nuevos indicadores/métricas. Además de la posible modificación y refinamiento de las variables actuales del análisis o la elección de un nuevo método y/o técnica de analítica (Chatti et al., 2012).

El proceso descrito anteriormente se centra únicamente en el proceso de analítica dentro del complejo procedimiento del LA y, por tanto, está ausente el concepto de comportamiento pedagógico. En el trabajo de Greller and Drachsler (2012) Greller and Drachsler (2012), se describe la relación entre estos (ver Fig. 3). El LA puede trabajar en apoyo a multitud de estrategias pedagógicas y actividades de aprendizaje gracias a la representación y manifestación del conocimiento oculto tras los datos disponibles. Esto significa que solo se puede ver la parte pedagógica a través de los datos. Debido a esto, no se incluye como parte del proceso de análisis (Fig. 2), pero implícitamente ya está contenida en los conjuntos de datos de entrada que encapsulan el comportamiento pedagógico de los usuarios. Como es sabido, este comportamiento depende en gran medida de la plataforma y la visión pedagógica de los desarrolladores (Anderson & Dron, 2011).

Además, la pedagogía se puede abordar de manera explícita en las metas y objetivos que establece el diseñador de la estrategia de LA (dimensión "objetivos/metás").

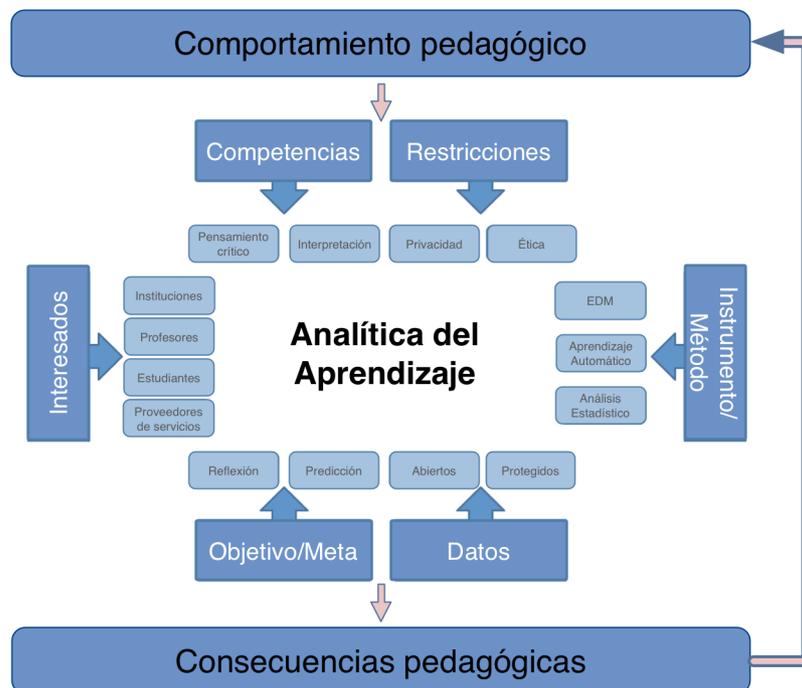


Fig. 3. Analítica del aprendizaje y la pedagogía; modificado de (Greller & Drachsler, 2012).

El método de LA (dimensión "instrumentos/métodos") determinará el resultado del análisis y, junto con la interpretación aplicada, puede dar lugar a una gran variedad de opciones para las consecuencias y las intervenciones. Si se hubieran aplicado dichas intervenciones pedagógicas, que daría lugar a nuevos comportamientos que, de nuevo, podrían ser analizados a través de los datos disponibles (Fig. 3). La conducta pedagógica aparece como un comportamiento estudiante/profesor que está motivado por los diseños didácticos (diseños de aprendizaje). Las consecuencias pedagógicas, de manera similar, son los ajustes a la estrategia didáctica o el diseño de aprendizaje basado en los resultados del proceso de LA. En la Fig. 3 la parte nuclear del proceso corresponde a el modelo de referencia de las dimensiones de LA, con la diferencia de que los elementos son más específicos, con el objetivo de ejemplificar el elemento, por ejemplo EDM en los instrumento o pensamiento crítico en las competencias.

Una definición de base para el análisis de aprendizaje es la que proporciona *Learning Analytics and Knowledge (LAK)*, la conferencia anual en el ámbito de la analítica del aprendizaje:

*El análisis del aprendizaje es la medición, recopilación, análisis y presentación de los datos sobre los estudiantes y sus contextos, con el objetivo de entender y optimizar el aprendizaje y los entornos en el que se produce (LAK, 2011).*

Diaz and Brown (2012) por otra parte, definen la analítica del aprendizaje en términos de sus objetivos:

1. Analizar datos múltiples y variados que están relacionados con y producidos por los estudiantes.
2. Tratar de monitorizar la actividad de aprendizaje y desarrollo y predecir los resultados del aprendizaje.
3. Permitir tanto a los estudiantes como a los tutores la intervención y toma de decisiones sobre el aprendizaje.

Una vista basada en una fusión de las actividades y procesos, propósito y uso final de la analítica la presenta M. Brown (2011) lo que resulta bastante útil al desgranar el concepto de LA:

- **Recopilación de datos:** esto supone el uso de programas, *scripts* y otros métodos para recopilar datos. Puede implicar datos de una sola fuente o de una variedad de fuentes, además puede conllevar grandes cantidades de datos. Los datos pueden ser estructurados (por ejemplo, los registros del servidor) o no estructurados (por ejemplo, mensajes en los foros de discusión). El diseño específico de la actividad de recolección debe ser congruente con los objetivos del proceso de LA.
- **Análisis:** a los datos no estructurados generalmente se les da algún tipo de estructura antes del análisis. Los datos se someten a una combinación adecuada de análisis cualitativo y cuantitativo. Los resultados del análisis se presentan mediante una combinación de visualizaciones, tablas, gráficos y otros tipos de visualización de la información.
- **Aprendizaje estudiantil:** este objetivo principal es el que distingue a la analítica del aprendizaje de otros tipos de analítica. Los procesos de LA se centran en el aprendizaje de los estudiantes: lo que los estudiantes están haciendo, dónde están invirtiendo su tiempo, a qué contenidos están accediendo, la naturaleza de su discurso, cómo están progresando, y así sucesivamente a nivel individual o de un grupo de estudiantes que trabajan juntos.
- **Audiencia:** la información que un proceso de LA aporta puede ser utilizada para:
  1. Informar a los instructores.
  2. Informar a los estudiantes.
  3. Informar a los equipos de gestión.

Los tres grupos comparten los informes que permiten las intervenciones apropiadas. Por lo general (1) y (2) posibilitan las intervenciones a nivel de curso, mientras que (3) informa a las intervenciones a nivel departamental, de centro e institucional. Los tipos de datos y análisis utilizados dependen de la audiencia objetivo.

- **Intervenciones:** la razón de realizar procesos de LA es permitir las intervenciones adecuadas a nivel individual, departamental o institucional. La analítica del aprendizaje puede hacer más que simplemente identificar a los estudiantes en riesgo. Mediante el análisis de las "migajas de pan" digitales que generan los estudiantes al participar en todos los aspectos de un curso, es posible observar el progreso del estudiante en etapas específicas y en las actividades de un curso. El potencial de los procesos de LA radica en ser capaz de indicar lo que está y lo que no está funcionando en un nivel mucho más fino de granularidad que nunca, incluso cuando un curso está transcurriendo (M. Brown, 2011).

Una vez definida y desgranados la definición y elementos de un proceso de LA es importante exponer el modelo de referencia, donde se destacan los participantes, los interesados y las restricciones y responsabilidades para la construcción de LA. Greller and Drachsler (2012) proponen un modelo de referencia. Con este marco se da por supuesto que los diseñadores responsables de los procesos analíticos no solo implementarán lo que es técnicamente posible hacer y está legalmente permitido (o por lo menos que no esté prohibido), pero teniendo en cuenta de manera integral los resultados de los grupos de interés y, aún más importante, las consecuencias para los interesados, es decir, la gente que suministra los datos (véase la Fig. 4). A diferencia de la Fig. 3 en esta se presentan los elementos de forma genérica, dado que es un modelo de referencia y tiene el objetivo de ser más abstracto.



Fig. 4. Dimensiones críticas de la analítica del aprendizaje (Greller & Drachsler, 2012).

### III.3.2 Modelo de referencia para la Analítica del Aprendizaje

El marco de referencia tiene la intención de ser una guía y un descriptor de las zonas problemáticas. Por tanto, puede referenciarse como un "marco de diseño" que puede y debe utilizarse para el diseño de servicios de LA desde una perspectiva integradora.

En este mismo sentido Chatti et al. (2012) describen un modelo de referencia para LA basado en cuatro dimensiones que identifica diversos retos y oportunidades de investigación en el área de LA en relación con cada dimensión. Como se representa en la Fig. 5, las cuatro dimensiones del modelo de referencia propuesto para LA son:

- ¿Qué?: ¿Qué tipo de datos reúne el sistema para su análisis?
- ¿Quién?: ¿Quién participa en el proceso análisis?
- ¿Por qué?: ¿Por qué se requiere el sistema de análisis de los datos recogidos?
- ¿Cómo?: ¿Cómo funciona el sistema para realizar el análisis de los datos recogidos?

Algunos usos potenciales de las aplicaciones de LA proporcionan apoyo a:

- Identificar a los estudiantes en riesgo a fin de proporcionar intervenciones positivas diseñadas para mejorar la retención.
- Brindar recomendaciones a los estudiantes en relación con el material de lectura y actividades de aprendizaje.
- Detectar la necesidad de mejoras pedagógicas y medir los resultados de estas.
- Adaptar la oferta de cursos.
- Identificar los profesores que tienen un buen desempeño y los profesores que necesitan ayuda con los métodos de enseñanza.
- Ayudar en el proceso de reclutamiento de estudiantes.

El informe *Horizon* del *New Media Consortium* (Johnson et al., 2011) ofrece una definición más completa:

*Analítica del aprendizaje se refiere a la interpretación de una amplia gama de datos producidos por y recopilados a favor de los estudiantes con el fin de evaluar su progreso académico, predecir el rendimiento futuro y los posibles problemas puntuales. Los datos se obtuvieron de las acciones explícitas estudiantiles, tales como completar las tareas y tomar los exámenes, de las acciones tácitas, que incluyen las interacciones sociales, actividades extracurriculares, mensajes en foros de discusión y otras actividades que no son evaluadas directamente como parte del progreso educativo del estudiante" (Johnson et al., 2011).*

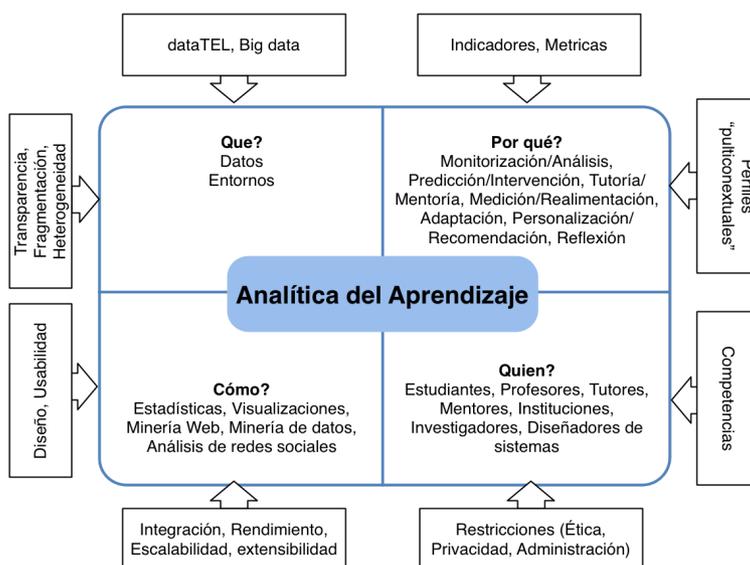


Fig. 5. Modelo de referencia para la analítica del aprendizaje, modificado de (Chatti et al., 2012).

El objetivo de la analítica del aprendizaje es permitir que los profesores y los centros educativos adapten las oportunidades educativas al nivel de las necesidades y capacidades de cada estudiante de la manera más rápida posible. La promesa de la analítica del aprendizaje es la de aprovechar el poder de los avances en la minería de datos, la interpretación y modelado para mejorar la comprensión de la enseñanza y el aprendizaje, y para adaptar la educación individualmente a los estudiantes con una mayor eficacia.

Todavía en sus primeras etapas, el análisis del aprendizaje responde a las peticiones de responsabilidad en los campus y su objetivo es aprovechar la gran cantidad de datos producidos por los estudiantes en las actividades académicas. Estos análisis se basan de forma intensa en la teoría del aprendizaje y prestan mucha atención a los elementos de la enseñanza en el momento de la participación *online* (Ferguson & Buckingham Shum, 2012b).

Los enfoques de la analítica que pueden ser clasificados de esta manera incluyen intrínsecamente formas de analítica social: analítica de las redes sociales y analíticas del discurso (De Liddo, Buckingham Shum, Quinto, Bachler, & Cannavacciuolo, 2011; Ferguson & Buckingham Shum, 2011). La analítica del discurso es un añadido al conjunto de herramientas de la analítica del aprendizaje relativamente reciente, sin embargo, se basa en un amplio trabajo previo en áreas como diálogo exploratorio (del inglés, *Exploratory Dialog*) (Ferguson & Buckingham Shum, 2011; Littleton & Whitelock, 2005; Mercer, 2000; Mercer & Wegerif, 1999) el análisis semántico latente (del inglés, *Latent Semantic Analysis*) (Thomas K. Landauer, Peter W. Foltz, & Darrell Laham, 1998; T. K. Landauer, P. W. Foltz, & D. Laham, 1998) y la argumentación asistida por ordenador (del inglés, *Computer Supported Argumentation*) (Rider & Thomason, 2008).

Los modelos se crean en algún punto en el tiempo para realizar analíticas o sistemas de predicción y apoyo al proceso de enseñanza y aprendizaje; pero, dado que las maneras en que la transmisión de la información en la educación se han estado manteniendo en cambios constantes tanto en género como en número y, por tanto, son cambiantes los factores que influyen en los modelos de predicción y por ende, es importante a la hora de la toma de decisiones que un estudiante/profesor o administrador puede y/o debe decidir que el modelo es inexacto en el momento presente y tomar una acción no justificada por los modelos analíticos que se utilizan. Esta circunstancia se debe permitir, pero también se deben registrar dichas circunstancias para tenerse en cuenta en los modelos nuevos o para mejorarlos.

Para M. Brown (2012) la recolección, la estandarización y el modelado de datos son solo una parte de cualquier sistema de producción, para el análisis y los datos derivados de una analítica se necesitan presentar a los usuarios de una manera eficaz y viable. Visualización y presentación de informes son los elementos que hacen la inteligencia del LA verdaderamente viables, ya que hacen visibles los patrones en los datos. Esto permite a instructores, asesores y estudiantes tomar las medidas adecuadas. Algunas de las disciplinas que contribuyen a la visualización son el diseño gráfico y la visualización de la información, la cartografía, la informática, la psicología y la estadística (M. Brown, 2012).

Para los educadores y los investigadores, las analíticas del aprendizaje han sido cruciales para obtener una mayor evidencia acerca de la interacción de los estudiantes con textos *online* y materiales de enseñanza informatizados. Los estudiantes están empezando a experimentar los beneficios de las analíticas del aprendizaje en la medida que se involucran en plataformas que utilizan datos específicos suyos para desarrollar sistemas alineados con sus necesidades de aprendizaje.

Una vez descrita la analítica del aprendizaje, así como la analítica académica es importante mencionar, dado lo expuesto en la Tabla 2, que esta multiplicidad de definiciones no debe ser tomada como recurso o alternativa para indicar un campo de confusión, sino más bien se propone con la intención de indicar la riqueza del dominio y la multiplicidad de enfoques a la misma. Ferguson proporciona una guía muy legible para el desarrollo del campo y su terminología. El propósito es útil también en la agrupación y discusión de la analítica. Ferguson (2012) propone la siguiente agrupación de la analítica dependiendo del propósito de esta:

- LA se centra principalmente en el reto de la educación: ¿cómo optimizar las oportunidades de aprendizaje en línea?
- AA y AcA se centran en el desafío político/económico: ¿cómo se puede mejorar sustancialmente las oportunidades de aprendizaje y los resultados educativos a nivel nacional o internacional?
- EDM se centra principalmente en el desafío técnico: ¿cómo se puede extraer valor de estos grandes conjuntos de datos relacionados con el aprendizaje?

A pesar de esta diversidad en la nomenclatura, Van Barneveld et al. (2012), define un marco de trabajo (del inglés, *framework*):

*We acknowledge that, functionally, different analytics are intended to work as a cohesive and integrated whole that serves the needs of the academy at a variety of levels.*

Existe una frontera difusa de las definiciones entre estos grupos, pero ha habido varios intentos para eliminar la ambigüedad de los campos. Long and Siemens (2011) se centraron en los significados actuales y futuros, distinguiendo entre la analítica del aprendizaje -que beneficia a los estudiantes y profesores, y se centran en el nivel de los cursos y del departamento- y la analítica académica -que benefician a los donantes/financiadores, los administradores y la comercialización a nivel institucional; financiadores y administradores a nivel regional, y los gobiernos y las autoridades educativas de internacional- (ver Tabla 3).

En este mismo sentido Baepler and Murdoch (2010) examinaron las diferencias entre la minería de datos, el análisis académico y la auditoría de los sistemas institucionales. Educause, que ha desarrollado sus definiciones de analítica durante varios años, ha tenido una visión más a largo plazo, para establecer un panorama más amplio de la terminología y destacar la variedad de definiciones que han surgido en la última década (Van Barneveld et al., 2012) (ver la Tabla 2).

En la Tabla 3 se muestra una discriminación entre la analítica del aprendizaje y la analítica académica, así como sus similitudes, y se presenta, incluso, sus principales beneficiarios. A diferencia de la AA, que utiliza los datos recogidos para tomar decisiones a nivel institucional, el objetivo de la LA es el uso eficaz de estos datos con el fin de tener un impacto directo en los estudiantes.

Tipo de analítica	Nivel u objeto de análisis	Beneficiario
Analítica del Aprendizaje	<b>Nivel curso:</b> análisis de las redes sociales, el desarrollo conceptual, y del discurso, "currículo inteligente"	Estudiantes, facultades
	<b>Departamental:</b> modelado predictivo, patrones de éxito / fracaso	Estudiantes, facultades
Analítica Académica	<b>Institucional:</b> Perfiles de estudiantes, rendimiento de las academias, flujo del conocimiento	Administradores, financiadores y marketing
	<b>Regional (estado/provincia):</b> comparación entre sistemas	financiadores y administradores
	<b>Nacional e internacional</b>	Gobiernos nacionales y autoridades educativas

Tabla 3. Analítica del aprendizaje y analítica académica - Enfoque de las actividades analíticas y de los beneficiarios. Modificado de (Long & Siemens, 2011).

A nivel individual, la LA puede apoyar la reflexión de los procesos de aprendizaje y ofrecer información personalizada sobre los progresos del estudiante (Govaerts et al., 2012; Govaerts et al., 2010). En el plano institucional, la LA puede mejorar los procesos de seguimiento y sugerir intervenciones o actividades para los estudiantes en particular. Sin embargo, se debe tener cuidado de no confundir los objetivos y las partes interesadas en el diseño de un proceso de la LA y no dejar que, por ejemplo, las consideraciones económicas y de la eficiencia en el plano institucional dicten estrategias pedagógicas, ya que esto posiblemente conduzca a la industrialización en lugar de personalización.

Con respecto al tercer elemento de la agrupación de Ferguson (2012), la minería de datos para la educación, este la coloca al mismo nivel de LA y de AA y se describe en la siguiente sección.

### III.4 Minería de datos para la educación

Al igual que con los primeros esfuerzos para entender los comportamientos en línea, los primeros intentos de minería de datos de datos educativos involucraban la minería de los web *log* (Ryan S.J.d. Baker & Yacef, 2009), pero ahora más integrados, instrumentados y sofisticados, los sistemas de aprendizaje en línea ofrecen más tipos de datos. Una característica importante y única de los datos sobre la educación es que son jerárquicos. Los datos a nivel de pulsación (del inglés, *click*), nivel de respuesta, nivel de sesión, nivel de los estudiantes, nivel de la clase y el nivel de los profesores y los centros educativos están incluidos y relacionados unos dentro de otros (Ryan S. J. d. Baker, 2011; C. Romero & Ventura, 2010).

Otras características importantes son el tiempo, la secuencia y el contexto. El tiempo es importante para capturar datos, tales como la duración de las sesiones de práctica o tiempo para aprender. La secuencia representa cómo los conceptos se construyen unos detrás de otros y cómo las prácticas y tutorías deben ser ordenadas. El contexto es importante para explicar los resultados y saber que un modelo puede o no puede funcionar. Los métodos de minería de datos jerárquica y longitudinal de modelado de datos han sido avances importantes en la EDM (Ryan S. J. d. Baker, 2011).

La minería de datos para la Educación en general hace hincapié en la reducción del aprendizaje en pequeños componentes que pueden ser analizados y luego influidos por el *software* que se adapta al estudiante (Siemens & Baker, 2012). Los datos del aprendizaje de los estudiantes recogidos por los sistemas de aprendizaje en línea se están explorando para desarrollar modelos predictivos mediante la aplicación de métodos de minería de datos educativos que clasifican los datos o encuentran relaciones. Estos modelos tienen un papel clave en la construcción de sistemas de aprendizaje adaptativos, en los que las adaptaciones o las intervenciones basadas en las predicciones del modelo se puede utilizar para cambiar lo que los estudiantes experimentan o incluso recomendar experiencias fuera de los servicios académicos para apoyar su aprendizaje (E. García et al., 2010).

La EDM se está convirtiendo en un área de investigación con un conjunto de métodos computacionales y psicológicos además de diversos planteamientos de investigación para entender cómo los estudiantes aprenden. Nuevos métodos interactivos de aprendizaje basados en ordenador y herramientas (sistemas inteligentes de tutoría, simulaciones, juegos) han abierto oportunidades para recopilar y analizar datos de los estudiantes, para descubrir patrones y tendencias en los datos, y para hacer nuevos descubrimientos y probar hipótesis sobre cómo los estudiantes aprenden. Los datos recogidos de los sistemas de aprendizaje en línea se pueden agregar más de un gran número de estudiantes y pueden contener muchas variables que los algoritmos de minería de datos pueden explorar para construir modelos.

Un significado, más restrictivo para la minería de datos educativa, que se refiere al uso de técnicas de minería de datos es el de C. Romero and Ventura (2010): para analizar datos con el fin de resolver los problemas de investigación educativa. La EDM se refiere al desarrollo de métodos para explorar los tipos de datos de los centros educativos y, con estos métodos, comprender mejor a los estudiantes y las características del entorno en el que aprenden (C. Romero & Ventura, 2007). La EDM explota la estadística, las máquinas de aprendizaje y los algoritmos de minería de datos sobre los distintos tipos de datos educativos. El Departamento de la Tecnología Educativa de EE.UU. ofrece una discusión de minería de datos educativos y, en general, sobre la analítica del aprendizaje (Bienkowski et al., 2012):

*EDM desarrolla métodos y aplica las técnicas de la estadística, aprendizaje automático y minería de datos para analizar los datos recogidos durante la enseñanza y el aprendizaje. EDM pone a prueba las teorías de aprendizaje e informa a la práctica educativa.*

*LA aplica técnicas de la informática, la sociología, la psicología, estadística, aprendizaje automático y minería de datos para analizar los datos recopilados durante los servicios educativos y la administración de la educación, la docencia y el aprendizaje. LA crea aplicaciones que influyen directamente en la práctica educativa.*

El desarrollo de la analítica del aprendizaje se ha visto impulsado al incluirlo en el informe Horizon del 2011 del NMC (Johnson et al., 2011). Este informe es parte de una serie y el NMC es una empresa establecida en el 2002 dedicada a identificar y describir las tecnologías emergentes y su potencial impacto en los próximos cinco años en una variedad de sectores en todo el mundo. Específicamente, el reporte Horizon del 2011 examina las nuevas tecnologías aplicadas a la educación, el aprendizaje y en la investigación creativa y su potencial impacto en estas áreas. Este identifica al LA como una tecnología para la observación. El informe Horizon del 2012 (Johnson et al., 2012) coloca a un subconjunto del área de la LA, el análisis social del aprendizaje (del inglés, *Social Learning Analytics*, SLA), en el plan de adopción de dos a tres años como tecnología emergente para el apoyo a la pedagogía (Buckingham Shum & Ferguson, 2012). La siguiente sección presenta la SLA con más grado de detalle.

### III.5 Analítica social del aprendizaje

El término analítica social del aprendizaje abarca también los enfoques "socializados", los cuales se pueden aplicar de manera fácil en entornos sociales. Estos incluyen el análisis de contenido, sistemas de recomendación y métodos automatizados de examinación, indexación y filtrado de recursos multimedia en línea con el fin de guiar al discente a través de la gran cantidad de recursos (Drachsler et al., 2010; Verbert et al., 2011). Esta analítica toma un aspecto social basándose en los datos obtenidos de las etiquetas, clasificaciones y metadatos suministrados por los estudiantes (véase, por ejemplo, (Clow & Makriyannis, 2011)).

El desarrollo de la SLA representa un alejamiento de la investigación basada en los datos hacia una investigación más fuertemente basada en las ciencias del aprendizaje y, cada vez más, se trata la complejidad de la formación continua que se lleva a cabo en una variedad de contextos. Una sub-rama socio-cultural de la investigación educativa demuestra cómo el lenguaje es en sí mismo es uno de los instrumentos principales a través del cual los aprendices construyen el conocimiento, y su uso se ve influenciado por los objetivos, los sentimientos y las relaciones de sus usuarios, todo lo cual cambia según el contexto (Wells & Claxton, 2008).

Otra línea de investigación hace hincapié en que el aprendizaje no puede entenderse por centrarse únicamente en la cognición, el desarrollo o el comportamiento de los estudiantes individuales, ni puede entenderse sin referencia a su naturaleza situada (J. Gee, 1997; Wertsch, Baker-Sennett, Rogoff, & Bell, 1992).

El pensamiento y el sentido de la construcción también se abordan desde el campo socio/cultural de los teóricos. Desde esta perspectiva, la construcción de significado está ligada a contextos y propósitos específicos. Se desarrollan formas compartidas de responder a los patrones y las características de los contextos particulares que Gee (J. Gee, 1997) llama a los discursos.

*Los discursos son coordinaciones socio históricas de personas, objetos, formas de hablar, actuar, interactuar, pensar, valorar y (a veces) la escritura y la lectura que permite la visualización y el reconocimiento de las identidades socialmente significativas.*

Dentro de los discursos, la gente puede intercambiar ideas y explicaciones. Explicaciones o teorías sobre la práctica vinculada a grupos socioculturales se llaman modelos culturales (J. P. Gee, 1992). Los modelos culturales no solo se llevan a cabo por los participantes individuales, sino que también residen en las prácticas en las que el grupo se involucra, las herramientas que utilizan, así como la configuración contextual.

Actualmente, se ve como una posibilidad real de la construcción de una visión integral del progreso del estudiante y tener en cuenta el sentimiento con el fin de permitir a los "sistemas basados en computadoras para interactuar con los estudiantes en forma de apoyo emocional" (Blikstein, 2011).

Nuevas herramientas, tales como el Servicio de Infraestructura Visualización GRAPPLE (GVIS) no se refieren a un solo VLE, sino que se pueden extraer datos de diferentes partes del Entorno Personal de Aprendizaje de aprendizaje (PLE) y utilizar estos datos para apoyar las habilidades meta-cognitivas como la auto-reflexión (Mazzola & Mazza, 2012).

El análisis de redes sociales, también denominado análisis estructural, investiga los lazos, las relaciones, los roles y las formaciones de la red, y un análisis de la red de aprendizaje social tiene que ver con cómo estas se desarrollan y se mantiene para apoyar el aprendizaje. Una red es un conjunto de elementos, que se denominan vértices o algunas veces nodos, con conexiones entre ellos, llamados bordes. Sistemas que toman la forma de redes (también llamados "grafos" en gran parte de la literatura matemática) abundan en el mundo. Algunos ejemplos son Internet, la *World Wide Web*, las redes sociales de conocimiento u otras conexiones entre individuos, redes organizacionales y redes de relaciones comerciales entre empresas, redes neuronales, redes metabólicas, redes de distribución, como los vasos sanguíneos o rutas postales de entrega, y muchos otros. Debido a su enfoque en el desarrollo de las relaciones, y su opinión de que la tecnología forma parte de este proceso, este tipo de análisis ofrece la posibilidad de identificar intervenciones que puedan aumentar el potencial de la red para apoyar el aprendizaje de sus actores. Por ejemplo, se ha demostrado que tener una amplia red es crucial para el desarrollo personal y profesional, además, ambos lazos débiles, que se sostuvieron con los conocidos, y los fuertes lazos de amistades duraderas junto con la pertenencia en una comunidad son importantes para el aprendizaje (Granovetter, 1973; D. Z. Levin, Cross, & Abrams, 2004).

El análisis de redes sociales puede abordarse desde la perspectiva de un individuo o de toda la red. Un enfoque egocéntrico puede identificar a las personas que apoyan el aprendizaje de un individuo, el origen de los conflictos en la comprensión, y algunos de los factores contextuales que influyen en el aprendizaje. Por otra parte, una visión general de la red proporciona información sobre los intereses y las prácticas de un conjunto de personas, la identificación de los elementos que sostienen la red juntos; Las visualizaciones de redes sociales pueden mostrar si las interacciones se producen entre todos los miembros de un grupo o si algunos miembros del grupo se comunican más (o menos) con otras personas concretas (Haythornthwaite & de Laat, 2010).

No obstante, si bien es cierto que este tipo de análisis ofrece una representación visual de la interacción social en Internet, el análisis de redes comienza prestando atención especial al estudio de las estructuras sociales insistiendo, por tanto, menos en por qué la gente hace lo que hace y más en la comprensión de los condicionantes estructurales de sus acciones. La asunción básica del análisis de redes es que la explicación de los fenómenos sociales mejoraría analizando las relaciones entre actores. El análisis de redes sociales generalmente estudia la conducta de los individuos a nivel micro, los patrones de relaciones (la estructura de la red) a nivel macro, y las interacciones entre los dos niveles (Sanz Menéndez, 2003).

A menudo, tanto la EDM como la SLA se refieren a la utilización de técnicas de "*big data*" a grandes conjuntos de datos para descubrir información útil con fines educativos, como se discute en (Ferguson, 2012). Este tipo de minería de datos y análisis de redes sociales educativas son parte de una serie de técnicas que se pueden aplicar en la analítica del aprendizaje y académica.

Como se ha mencionado con anterioridad, en estos campos, AA y LA, además de en las técnicas de EDM y SLA se requieren de varios elementos para la realización del estudio y la obtención de nueva información y/o conocimiento para hacer uso de esta. Entre estos elementos se encuentran, en el inicio de proceso, los datos, dentro de los cuales se mencionó el concepto que se definirá en la siguiente sección, *big data*.

Además de los datos, en la parte final del proceso, se encuentra la representación de los datos que se ve influenciada por la experiencia y conocimiento que tenga la persona analista, líder de la institución o gestor académico para poder comprender el comportamiento de los datos y/o la representación de la información durante el proceso de analítica. Este concepto se conoce como alfabetización de los datos (del inglés, *Data Literacy*, DL) y por ende, a la hora de la interpretación de las visualizaciones de datos y diversas técnicas de representación y análisis de datos, la alfabetización visual (Gershon, Eick, & Card, 1998).

Por ejemplo, Stephenson and Schifter Caravello (2007), presentan un estudio en el cual realizan una integración de lo que se define como alfabetización de los datos en un curso de Ciencias Sociales de la Universidad de Los Ángeles California (UCLA). Como conclusión de este ejercicio, se expone que en muchas de las Ciencias Sociales, y particularmente en la Sociología, la comprensión del expediente académico y el desarrollo del pensamiento crítico se basan en parte en la comprensión de la información numérica y sus representaciones. Por tanto, la alfabetización de los datos es un componente integral de la alfabetización informacional para estas disciplinas.

Por otro lado, en la Universidad de Wollongong se realizó un uso distinto de la analítica, donde por medio de un cubo en función de la biblioteca (en este caso específico cubo es una función de base de datos y presentación de informes a medida que se une a los datos de uso de la biblioteca con los datos de los estudiantes, incluyendo información demográfica y de rendimiento académico) se intentó predecir las notas de los estudiantes dependiendo del uso que le daban a la biblioteca. El análisis de los datos reveló una fuerte correlación entre las notas y el uso de los recursos de información que la biblioteca ofrece a los estudiantes (Cox & Jantti, 2012).

La instrucción de análisis fuera de la educación superior implica a los instructores directamente en el proceso de analítica. El proyecto DATA de Oregon entrena profesores de K-12 para aplicar la analítica en la selección de las decisiones educativas con base en las necesidades individuales de los estudiantes, lo que resulta en la adopción generalizada de toma de decisiones basada en los datos a nivel de aula (M. Garrison & Monson, 2012). Sobre la base de esta idea y teniendo en cuenta que la toma de decisiones basada en los datos (del inglés, *Data Driven Decision Making*, DDDM) es el proceso por el cual los administradores y profesores recogen y analizan datos para tomar una serie de decisiones educativas

(Ikemoto y Marsh, 2007); además de que el uso de datos se ha convertido en una característica predominante de las políticas de los numerosos países y agendas de reforma, en particular en los Países Bajos (Schildkamp & Teddlie, 2008), Inglaterra (Ehren & Visscher, 2008) y Canadá (Earl & Katz, 2006; Raby, 2009). DDDM es también un proceso importante de la recuperación y reinversión de la Ley de 2009 y la carrera a la mejor competencia patrocinado por el Departamento de Educación de EE.UU (Hamilton et al., 2009; J. A. Levin & Datnow, 2012).

En definitiva, la alfabetización de los datos es un componente importante en la analítica y, por tanto, también en la analítica del aprendizaje y académica (Greller & Drachsler, 2012).

### III.6 Big data

Durante las últimas décadas han tenido lugar varios cambios en la forma en que las organizaciones y los individuos generan y procesan la información y la participación. Por un lado, la aparición de nuevos dispositivos, como teléfonos inteligentes y tabletas, y la mejora en las capacidades de procesamiento y almacenamiento de las computadoras y servidores, han aumentado las capacidades de los individuos y organizaciones para crear formatos de contenido de datos nuevos y diferentes. Mientras que, por otro lado, los datos generados podrían ser compartidos en tiempo real a altas velocidades, gracias a las nuevas tecnologías de la comunicación.

Por tanto, los datos son más variados y con formatos más complejos que ahora se crean, se recogen y se transmiten de numerosas fuentes, incluyendo redes de sensores, GPS, etiquetas RFID, redes inalámbricas, satélites, flujos multimedia y *software* utilizados por los dispositivos móviles (Krishnan, 2013). En consecuencia, los individuos y las organizaciones tienen acceso a más datos de lo que son capaces de procesar y transformar en conocimiento, lo que está provocando la sobrecarga de información en todos los ámbitos del quehacer humano. Siguiendo esta idea, *big data* es un término que se usa para describir el nuevo contexto de la abundancia, definiéndose en relación con las capacidades actuales. El Instituto Global McKinsey define *big data* como "conjuntos de datos cuyo tamaño está o va más allá de la capacidad de las herramientas de *software* de bases de datos típicos para capturar, almacenar, gestionar y analizar" (Manyika et al., 2011). Este término se utiliza con frecuencia para hacer referencia a tal volumen, la variedad, la velocidad y la complejidad de los datos que se producen diariamente. Aunque las características de los datos han cambiado con el tiempo, muchos retos asociados con los grandes datos no son tan nuevos y se han discutido en el ámbito de la investigación durante años, incluso décadas.

Estos desafíos están relacionados con la recuperación y la integración de fuentes de datos heterogéneas en grandes volúmenes de datos, su procesamiento y almacenamiento y la escalabilidad de las herramientas que realizan análisis de datos automática. Sin embargo, no se trata solo de tamaño, también se refiere a la velocidad, el volumen y la capacidad de cálculo y de análisis requerida para gestionar los datos y obtener conocimiento (Hemerly, 2013).

El objetivo principal de la investigación relacionada con los *big data* y las tecnologías es administrar y transformar datos disponibles en tiempo real y datos históricos en conocimiento para informar las decisiones de acuerdo a los requerimientos y necesidades organizacionales.

En respuesta a las limitaciones de las técnicas de gestión de los datos existentes, se ha desarrollado una nueva generación de tecnologías (por ejemplo, Hadoop<sup>9</sup>), bases de datos y técnicas (por ejemplo, la minería de datos o descubrimiento de conocimiento en bases de datos)

Como consecuencia de ello, los teóricos han postulado que algo fundamental ha cambiado en los datos en sí, creando un mundo en el que casi todas las interacciones de datos, incluyendo la investigación científica, se ven afectados:

*Este es un mundo donde grandes cantidades de datos y matemáticas aplicadas reemplazan cualquier otro instrumento que pueda usarse. ¡Fuera todas las teorías del comportamiento humano, de la lingüística a la sociología. Olvídense de la taxonomía, ontología y la psicología. Quién sabe por qué las personas hacen lo que hacen? El punto es que lo hacen, y podemos seguir y medir con fidelidad sin precedentes. Con suficientes datos, los números hablan por sí mismos. (Long & Siemens, 2011).*

Las analíticas en la educación constituyen la aproximación educativa a los *big data*, una ciencia originalmente utilizada en el ámbito empresarial para analizar las actividades del consumidor, identificar tendencias de consumo y predecir el comportamiento de los consumidores. La irrupción de Internet ha impulsado la investigación en torno a los *big data* y hacia métricas de todo tipo, así como la proliferación de las herramientas de rastreo web, que permiten a las empresas construir grandes reservas de información que eventualmente podrían poner en valor en sus campañas de *marketing*.

Los *big data* capturados de los comportamientos en línea de los usuarios permite a los algoritmos inferir el conocimiento de los usuarios, las intenciones e intereses y para crear modelos con el objetivo de predecir el comportamiento y el interés futuro de estos. La educación se ha embarcado en un proceso de búsqueda de datos similar, con objeto de incrementar la retención de los estudiantes y proporcionarles una experiencia de aprendizaje personalizada y de alta calidad. El énfasis fundamental en grandes volúmenes de datos es que los datos en sí son un punto o un camino hacia la generación de valor en las organizaciones.

Los datos no son simplemente el subproducto de interacciones y actividades de una organización. *Big data* es una capa de valor crítico para los gobiernos, las empresas y las instituciones de educación superior<sup>10</sup>. Esta capa comprende una estructura de cantidades grandes de datos, la cual debe ser construida y esquematizada con tripletas de información introspectiva a medida de cada institución (Berman, 2013;

<sup>9</sup> <http://hadoop.apache.org/>.

<sup>10</sup> <http://www.whitehouse.gov/blog/2012/03/29/big-data-big-deal>

Haythornthwaite, de Laat, & Dawson, 2013). Como se ha mencionado anteriormente, *big data* es una parte del proceso de cualquier analítica, ya sea académica o del aprendizaje, y la consecuencia de este análisis se debe post-procesar para ser presentado a los tomadores de decisiones, estudiantes, analistas, equipo de gobierno o académicos. Esta presentación de los datos puede ser de diferentes formas y estas se ven influidas por las capacidades y habilidades del receptor.



## Capítulo IV Visualización de la Información y Analítica Visual

*“By visualizing information, we turn it into a landscape that you can explore with your eyes, a sort of information map. And when you’re lost in information, an information map is kind of useful.”*

*-David McCandless, data journalist, and*

*information designer*



Desde el comienzo de la ciencia, los seres humanos han experimentado momentos de introspección en el que una cierta idea o solución se vuelve tan clara como nunca lo había sido antes. A esto se le llama mapa mental, modelo mental o imagen mental, que comúnmente es fruto de la interacción de los sentidos con el mundo o con representaciones de este.

Estas técnicas para la representación de los datos no son nuevas, este tipo de técnicas han evolucionado durante unos 500 años de diseño de la información. Desde el Renacimiento italiano del siglo XV, cuando los arquitectos florentinos perfeccionaron la geometría y el dibujo en perspectiva enriqueciendo las representaciones de los objetos físicos, varios de estos métodos emprendedores han evolucionado, de forma que a menudo se encuentran en los esquemas cotidianos de quienes se ven enfrentados a una cantidad abrumadora de datos. Algunas de estas técnicas están bien documentadas, como por ejemplo, la estructuración de la tabla periódica de los elementos químicos (E. Tufte, 1990).

El auge de las tecnologías de la información y la comunicación ha creado nuevas condiciones para la aparición de sociedades del conocimiento. Como se mencionó anteriormente, la Sociedad del Conocimiento se deriva de la Sociedad de la Comunicación y de la Información (Levis, 2004). En la actualidad la sociedad se enfrenta a una cantidad cada vez mayor de datos, en la última década la mejora constante de los dispositivos de almacenamiento de datos y los medios para crear y recopilar datos han influido en la forma en que se trata la información. La mayor parte del tiempo, los datos se almacenan sin filtrado y refinamiento para su uso posterior. Hoy en día prácticamente todas las ramas de la industria o negocio, y cualquier actividad política, empresarial o personal, generan grandes cantidades de datos. Estas son circunstancias que caracterizan a la Sociedad de la Información.

Esta llamada Sociedad de la Información es un concepto que se construyó durante las últimas décadas del siglo pasado, obteniendo su validación internacional en la Cumbre Mundial de la Sociedad de la Información (CMSI). En esta cumbre se firmó una declaración de principios cuyo primer punto del documento afirma el "deseo y compromiso comunes de construir una Sociedad de la Información centrada en la persona, incluyente y orientada al desarrollo, en el que todos puedan crear, consultar, utilizar y compartir la información y el conocimiento" (CMSI, 2004).

La Sociedad de la Información en gestación solo cobrará su verdadero sentido si se convierte en un medio al servicio de un fin más elevado y deseable: la construcción a nivel mundial de sociedades del conocimiento que sean fuentes de desarrollo para todos, y sobre todo para los países menos adelantados. Para lograrlo, dos desafíos planteados por la revolución de la información revisten una importancia particular: el acceso a la información para todos y el futuro de la libertad de expresión (Martín-Barbero, 2012).

Un elemento central de las sociedades del conocimiento es la "capacidad para identificar, producir, tratar, transformar, difundir y utilizar la información con vistas a crear y aplicar los conocimientos necesarios para el desarrollo humano. Estas sociedades se basan en una visión de la sociedad que propicia la *autonomía* y engloba las nociones de pluralidad, integración, solidaridad y participación". Tal como la UNESCO puso de relieve en la primera parte de la Cumbre Mundial sobre la Sociedad de la Información (CMSI), la noción de sociedades del conocimiento es más enriquecedora y promueve más la autonomía que los conceptos de tecnología y capacidad de conexión que a menudo constituyen un elemento central en los debates sobre la sociedad de la información. Las cuestiones relativas a la tecnología y la capacidad de conexión hacen hincapié en las infraestructuras y la gobernanza del universo de las redes. Aunque revisten una importancia fundamental evidente, no deberían considerarse como un fin en sí mismas. En otras palabras, la sociedad mundial de la información solo cobra sentido si propicia el desarrollo de sociedades del conocimiento y se asigna como finalidad.

Actualmente, además de estas diferencias en infraestructura y comunicación y la vasta y diversa cantidad de información que está en constante flujo, agravando la situación, las posibilidades para recoger y almacenar datos van en incremento a una velocidad mayor que la capacidad de utilizarla para la toma de decisiones. Aunado a esto, en la mayoría de las aplicaciones, los datos en bruto no tienen valor en sí mismos, sino que se requiere extraer el conocimiento contenido en estos (Bindé, 2005). En efecto, cabe preguntarse si la desigualdad de acceso a las fuentes, contenidos e infraestructuras de la información no pone en tela de juicio el carácter mundial de la Sociedad de la Información. Estas actuales circunstancias presentan, entre otros, el problema de la sobrecarga de información, la cual, se refiere al peligro de perderse en los datos: al tener datos irrelevantes para la tarea deseada, que los datos estén procesados de manera inapropiada o que se presenten en una forma no adecuada. Se puede considerar dos formas generales de presentar información: como texto o como alguna forma de representación pictórica abstracta.

El problema clave con la representación de texto es que a pesar de permitir un alto nivel de definición, requiere una gran cantidad de esfuerzo cognitivo y no genera una imagen en la capacidad inherente de las personas para el reconocimiento de patrones y análisis. Por el contrario, la representación gráfica se adopta, ya que permite al usuario reducir su carga cognitiva, y se basa en sus habilidades de reconocimiento de patrones y análisis visual. Estas capacidades de cada individuo dependen del nivel de alfabetización de los datos o de información que se tenga, y más específicamente de la alfabetización visual, ya que en el campo de las representaciones visuales el sentido de la visión tiene el potencial de estar en el centro de la comprensión y el reconocimiento de patrones.

Este capítulo consta de cinco secciones en las cuales se describe, como primera sección, qué es la visualización y las definiciones actuales con sus diferentes enfoques, seguido por los fundamentos de la psicología educativa, que describen el lugar que la visualización ocupa y su importancia en la educación. Posteriormente se define qué es la visualización de la información en la sección tres y en la cuarta sección la analítica visual y su proceso, en la quinta sección se describen algunas técnicas y estrategias de visualización de la información (del inglés, *Information Visualization, infoVis*) y visualización analítica y, finalmente, dada la anteriormente mencionada estrecha relación con la visualización, se expone el concepto de alfabetización de los datos.

## IV.1 Visualización

La visualización (del inglés, *Visualization, Vis*) es un término antiguo que ha recibido una gran cantidad de interés recientemente en la comunidad de Ciencias de la Computación. Según el diccionario Merriam-Webster Online (2014) la visualización es la "formación de imágenes visuales, el acto o proceso de interpretar en términos visuales o de puesta en forma visual" Merriam-Webster Online (2014). La cual coincide con la definición de la RAE de *visualización*: acción y efecto de *visualizar*, donde visualizar es "Representar mediante imágenes ópticas fenómenos de otro carácter; por ejemplo, el curso de la fiebre o los cambios de condiciones meteorológicas mediante gráficas, los cambios de corriente eléctrica o las oscilaciones sonoras con el oscilógrafo, etc.". Por su parte, ya con el uso del ordenador, McCormick, DeFanti, and Brown (1987) ha añadido: "visualización es una herramienta o método tanto para la interpretación de los datos de imágenes introducidos en un ordenador, como para la generación de imágenes de conjuntos de datos multidimensionales complejos".

Spence (2000) ha observado que existe una diversidad en el uso del término "visualización", algunos autores usan visualización para referirse a dos ideas diferentes, una es a la representación gráfica y la otra al proceso cognitivo de entender una imagen. Ware (2004) ha definido la visualización como una imagen mental o una representación visual de un objeto, de una escena, persona o de la abstracción que es similar a la percepción visual. Estas definiciones reflejan la existencia de una visualización externa y una interna.

El término de visualización, como representaciones internas, se utiliza para describir construcciones mentales internas, es decir, modelos mentales, que se crean en el ojo de la mente y se utilizan en las imágenes mentales y para resolver problemas en los que las personas leen su modelo mental (Johnson-Laird, 1983). Las visualizaciones, como representaciones externas, se refieren a representaciones usadas típicamente para el aprendizaje y utilizadas en la ciencia, como por ejemplo los gráficos, diagramas, modelos, simulaciones, etc., y son, por definición, semánticamente ricos.

La visualización, como el lenguaje y otras herramientas cognitivas y comunicativas, varía en efectividad. La visualización es una representación gráfica, que es la mejor en la transmisión de las ideas complicadas en una forma clara, precisa y eficiente (Stuart K. Card et al., 1999; Edward R. Tufte, 1997). La visualización es la transformación de la representación simbólica a la representación geométrica. Esta se genera a partir del contenido de los datos (numéricos, caracteres, lógicos o abstractos), y cuando se ofrecen a un observador son procesados a través de la vista (el sentido que mejor permite captar una gran cantidad de información).

En general, el ancho de banda de presentación de la información es potencialmente mayor en el dominio visual que para los medios de comunicación que llegan a cualquiera de los otros sentidos. Los seres humanos tienen capacidades perceptivas notables, que son muy infrutilizados en los diseños actuales. Los usuarios pueden escanear, reconocer y recordar las imágenes rápidamente, también pueden detectar cambios en el tamaño, el color, la forma, el movimiento o la textura. Así mismo, pueden apuntar a un solo píxel, incluso en una pantalla de megapíxeles, y pueden arrastrar un objeto a otro para realizar una acción. La capacidad humana de reconocer patrones subconscientemente (Miemis, 2010) es un factor determinante, y todo ello permite al observador acceder a conocimiento nuevo en solo un vistazo, a través de un modelo mental que se ha generado en la mente, y que es algo que no se puede imprimir en papel o desplegar en una pantalla de ordenador.

Considerando esto, se puede decir que la visualización es una actividad cognitiva, facilitada por las representaciones gráficas externas, a partir de las cuales la gente construye una representación mental del mundo (Ware, 2004). En otras palabras, la visualización es el proceso de transformación de los datos, la información y el conocimiento en forma visual, explotando las capacidades del cerebro humano para procesar significativamente mayor cantidad de información visual (Gershon et al., 1998).

Aun cuando las representaciones visuales se utilizan para diversos fines (analíticos, de negocios, gubernamentales, educativos), la finalidad de la visualización es la transmisión de una idea o imagen mental al receptor, lo que conlleva el uso de las capacidades perceptivas y cognitivas tanto del receptor como del emisor. Estas son parte del proceso de transmisión de conocimiento y, por ende, del proceso de enseñanza y aprendizaje, por lo que en la siguiente sección se describirán los fundamentos de la psicología de la educación.

## IV.2 Fundamentos de la psicología de la educación

Una parte significativa de la información necesaria para la vida se obtiene a través del uso del sistema visual. Si se compara la percepción visual con los otros sistemas sensoriales, esta es el canal sensorial primario con el que la mente construye sus representaciones.

Esto es cierto porque "se extiende al ser humano más allá de su propio cuerpo, es el mediador de otras impresiones sensoriales, y actúa como un estabilizador entre la persona y el mundo externo" (Villalba Simón et al., 2000).

Entre todos los estímulos visuales, los que agregan el mayor aporte para el conocimiento, y en el período más corto de tiempo, son las representaciones gráficas visuales. Esta es la razón de su gran importancia en el proceso de aprendizaje. La función de los sistemas de análisis visuales en la educación debe abarcar no solo la exposición de la información, sino también, y sobre todo, la transformación de la información en conocimiento.

Los estudios sobre representaciones textuales y gráficas como estrategias orientadas a mejorar el proceso de construcción de representaciones mentales, confirman "que la estructura de los gráficos afecta a la estructura del modelo mental...", por lo que los "estudiantes necesitan oportunidades para participar en actividades de procesamiento en las que las exigencias impuestas por la presentación de la información se corresponden con las capacidades del individuo para hacer frente a esa información" (Ploetzner & Lowe, 2004). Con este enfoque, los sistemas analíticos visuales no solo se ocupan de la promoción del aprendizaje visual del contenido de aprendizaje, sino también con la visión de cómo el componente visual puede mejorar las condiciones de aprendizaje en y a través de los entornos virtuales de aprendizaje, donde los materiales, sin duda, juegan un papel mediador decisivo. Por tanto, la comprensión de la función docente de las representaciones visuales en el aprendizaje puede entenderse desde perspectivas socio-culturales como la mediación instrumental.

Según Vygotsky (1989), toda la actividad depende del material con el que opera. Desde el punto de vista sociocultural, la presencia de la actividad instrumental -la interacción con los diversos tipos de productos físicos o simbólicos de la cultura- es fundamental en el aprendizaje. No trata con una comprensión accesoria de la función de los instrumentos en la construcción de la mente, sino más bien, con esta concepción de los procesos psicológicos superiores, el desarrollo no existe sin la presencia de la actividad instrumental; el aprendizaje lleva la forma particular del instrumento con el que la actividad se lleva a cabo.

De la misma manera que una calculadora media la actividad de un estudiante en la solución de un problema de matemáticas, por ejemplo, mediante la modificación de su forma de representar y actuar sobre el problema que hay que resolver -ya que el artefacto, como sistema simbólico que es, opera como un amplificador de la acción- un VLE, como un instrumento y medio ambiente más complejo, según sugiere un sistema simbólico más diverso para el aprendizaje y, por tanto, con una atención técnica y pedagógica más elaborada.

Un factor importante es saber que el VLE y los diferentes recursos simbólicos utilizados para aprender tienen una función mediadora que transforma, modifica o altera las condiciones en las que el estudiante aprende en el aula virtual. No es una cuestión solo de saber que se puede estudiar con herramientas de aprendizaje visual, sino de la comprensión de cómo estas herramientas pueden mejorar la comprensión de la actividad en el aprendizaje virtual. Este es un punto que justifica la teoría socio-cultural de la utilidad de los VLE como instrumento de mediación socio-cultural. Como Bruner (2000) afirma de una manera sintética:

*La cultura da forma a la mente y nos proporciona la caja de herramientas a través del cual se construye no solo nuestros mundos, sino también nuestras concepciones de nosotros mismos y nuestras facultades. (Bruner, 2000).*

Además, Bruner (2000) identificó tres principios que sirven de guía para el desarrollo de la instrucción:

- (1) La instrucción debe estar relacionada con las experiencias y los contextos que hacen que el estudiante esté deseoso y sea capaz de aprender (disposición).
- (2) La instrucción debe estar estructurada de modo que el estudiante pueda aprehenderla fácilmente (organización espiral).
- (3) La instrucción debe estar diseñada para facilitar la extrapolación y/o para completar las brechas de conocimiento (llegando más allá de la información dada)

Por tanto, hablar de una imagen es también hablar de la oportunidad de la representación, de la mediación, de la actividad de aprendizaje en el VLE. Un recurso visual no es una entidad neutra, sino que afecta al sujeto en dos formas: como una acción externa, permitiendo que el sujeto haga algo (o no) con el apoyo de este recurso, y también por afectar a la representación interna de la asignatura: se aprende a pensar con esas formas de hacer algo. Esta doble orientación es válida cuando se habla de materiales de enseñanza y aprendizaje en la formación virtual, cuando el canal de la acción educativa es un medio como un VLE.

Además, existe la necesidad de entender cómo estas relaciones se ajustan entre la representación interna, el aprendizaje, y la representación externa, recursos visuales, qué medios se deben utilizar y qué tipo de ayuda se debe proporcionar a los estudiantes cuando aprenden a través de la representación visual. El objetivo principal es crear modelos útiles que sirvan para proponer mejores condiciones visuales y no visuales para el aprendizaje.

Los diseñadores han descubriendo cómo utilizar las pantallas de color y de alta resolución para presentar grandes cantidades de información de manera rápida, ordenada y controlada por el usuario. Psicólogos perceptuales, estadísticos y diseñadores gráficos ofrecen una valiosa orientación sobre la presentación de información estática, pero la oportunidad de utilizar representaciones dinámicas ofrece a los diseñadores de las interfaces de usuario mucho más que la sabiduría de los años 90 (B. Shneiderman, 1996; E. Tufte, 1990; Edward R. Tufte, 1997), dando como resultado métodos y técnicas de visualización de la información, que se desarrollan en la siguiente sección.

### IV.3 Visualización de la información

Actualmente, los investigadores del lenguaje visual y los diseñadores de interfaces de usuario han desarrollado potentes métodos de visualización de la información, al tiempo que se van integrando con los avances tecnológicos y la finalidad del objeto de estudio. La aparente simplicidad de la visión oculta los complejos mecanismos para lograr el éxito del diseño de sistemas de visualización de la información, que depende de la caracterización adecuada de la tarea, el sistema visual humano, las representaciones, así como la interacción dinámica entre todos estos elementos (Ware, 2004). Según Stuart K. Card et al. (1999) la visualización de la información es:

*El uso de representaciones visuales interactivas con el apoyo de los ordenadores, que tienen como objetivo amplificar la cognición.*

En otras palabras, esta trata de aprovechar el medio dinámico e interactivo y de bajo costo de los ordenadores para diseñar nuevos apoyos gráficos que mejoren las capacidades cognitivas. Esto significa que con el uso de los ordenadores, se realiza la selección de la información esencial, la eliminación de la información irrelevante y la estructuración de la información para que pueda ser precisa y fácilmente comprendida de forma eficaz. Después se despliega la información a través de representaciones espaciales o gráficas para facilitar la comparación, el reconocimiento de patrones, la detección de cambios y otras habilidades cognitivas, haciendo uso del sistema visual humano (Spence, 2000).

Estas representaciones son simplificaciones elegidas para ayudar a la creación de percepciones visuales (modelos mentales) de lo que sucede en un nivel macro del objeto de estudio o de análisis. Este proceso de simplificación y la representación dentro del alcance de los sentidos humanos con la ayuda de modelos se vuelve de mayor importancia en cada secuencia de preguntas y explicaciones de los fenómenos estudiados en un nivel micro de la información.

Los modelos y las representaciones o visualizaciones (imagen o modelo mental) de estos y de las entidades, las relaciones, las causas, los efectos del fenómeno o de la información representada se convierten en vitales, ya que son una parte crucial en la producción de nuevo conocimiento. Una definición de visualización muy mencionada en la literatura, es la de Stuart K. Card et al. (1999) mencionada anteriormente. Donde cognición se entiende como el poder de la percepción humana, o en palabras más sencillas, la adquisición o utilización de conocimientos (Stuart K. Card et al., 1999; Teyseyre & Campo, 2009). Por tanto, el objetivo de la visualización de la información es analizar, explorar, descubrir y comunicar información en una forma muy comprensible. La visualización de la información se usa para presentar gran cantidad de información de forma coherente y compacta, desde diferentes puntos de vista, para proporcionar así diferentes niveles de detalle (Kowalski & Maybury, 2000). El número y complejidad de los datos obligan a administrar eficientemente los grandes volúmenes de información producidos, lo que incrementa su dificultad de procesamiento y análisis.

La importancia de la visualización gráfica de la información radica en que permite a las personas analizar de mejor manera las diferentes relaciones de los datos. El diseño de dicha visualización se enfoca básicamente en dos aspectos, los contenidos y el diseño gráfico, para los cuales se deben tener en cuenta las técnicas de visualización adecuadas que se han de utilizar para generar la presentación de la información a los diversos usuarios. Uno de los aspectos más importantes a tener en cuenta en la visualización de la información es escoger adecuadamente la técnica de visualización relevante que se aplique de mejor manera al problema a tratar. Por consiguiente, las clasificaciones o taxonomías relacionadas con la visualización de la información proporcionan elementos básicos para ayudar a resolver el problema de implementar eficientemente dicha visualización (Amar, Eagan, & Stasko, 2005; Arens, Hovy, & Vossers, 1992; Stuart K Card & Mackinlay, 1997; Chung, Chen, Chaboya, O'Toole, & Atabakhsh, 2005; Lee, Plaisant, Parr, Fekete, & Henry, 2006b; H.-J. Schulz & Schumann, 2006; B. Shneiderman, 1996; Wehrend & Lewis, 1990; Michelle X Zhou & Feiner, 1996). Además, facilitan el entendimiento de cómo aplicar determinada visualización de acuerdo con la naturaleza de la información y su campo de acción. Y para ello se debe tener en cuenta el tipo de información a representar y su naturaleza. A continuación se relacionan los tipos de información a representar.

#### IV.3.1 Tipos de información

- **Lineal:** Tablas, código fuente del programa, listas alfabéticas, artículos ordenados cronológicamente, etc.
- **Jerárquica:** Estructuras de árbol.
- **Redes:** Estructuras generales de grafos, tales como hipermedia, gráficos de nodos y enlaces (cíclicos, con nodo origen y sin él, dirigidos y no dirigidos), redes semánticas, redes, etc.
- **Multidimensional:** Atributos de metadatos, como el tipo, el tamaño, el autor, la fecha de modificación, etc. Los elementos con n-atributos se convierten en puntos en el espacio n-dimensional.
- **Espacios de características:** En recuperación de información un vector de expresión representa cada objeto de una colección. Para las colecciones de documentos de texto, las frecuencias de las palabras se utilizan para construir vectores de expresión. El espacio de características se proyecta en dos o tres dimensiones de la pantalla.
- **Espacios de consulta:** Artículos distribuidos en función de su afinidad con (combinaciones de) los términos de una consulta en particular.
- **Espacial:** Datos intrínsecamente 2D o 3D, tales como planos, mapas, modelos CAD, etc. Dado que la información espacial tiene una obvia y natural representación en lugar de una abstracta, no se considera que sea visualización de información en un sentido estricto, sin embargo existen representaciones abstractas de este tipo.

Las representaciones de la información se hacen aún más atractivas al aportar orientación o contexto, para permitir la selección de las regiones y para dar información dinámica con la finalidad de la identificación de los cambios. Este tipo de actividades realizadas de manera automática o manualmente por el usuario se conocen como interacciones.

### IV.3.2 Interacción

Como el análisis visual se ocupa de la relación entre representaciones visuales y la cognición humana, más que del desarrollo de nuevas metáforas visuales, rara vez es suficiente para desencadenar esta generación de conocimiento (donde el conocimiento puede ser un nuevo descubrimiento o la confirmación o la negación de una creencia previa). Estas representaciones visuales deben estar incrustadas en un marco interactivo, donde los andamios del proceso de construcción del conocimiento humano, en conjunto con las herramientas y los métodos adecuados, están para apoyar la acumulación de pruebas y observaciones sobre las teorías y creencias del analista.

La "ciencia de la interacción" se ocupa del estudio de los métodos por los cuales los seres humanos crean conocimiento a través de la manipulación de una interfaz. Como toda ciencia, implica el desarrollo y prueba de las teorías sobre los medios más eficaces para apoyar la investigación.

Interacción, sin embargo, es un término sobrecargado. En un nivel, la interacción típicamente se refiere al conjunto de los controles previstos para el usuario con el objetivo de manipular una interfaz y las relaciones entre el usuario y la interfaz. A un nivel más abstracto, es la interacción entre el usuario y el espacio del problema.

Esta interacción de alto nivel es un acto cognitivo que está habilitado por herramientas computacionales, pero no tiene lugar exclusivamente dentro de ellas (ni, para el caso, a través del uso de cualquier herramienta única). La ciencia de la interacción, por tanto, es un concepto mucho más amplio que simplemente los principios para la creación de *widgets* de interfaz (Pike, Stasko, Chang, & O'Connell, 2009).

La comunidad de visualización de la información ha comenzado a distinguir entre las interacciones de bajo nivel (aquellas entre el usuario y la interfaz de *software*) y las interacciones de alto nivel (aquellas entre el usuario y el espacio de la información).

En la Fig. 6 se resume la relación de entre alto y bajo nivel de interacción. Los controles interactivos proporcionados por la representación visual individual proporcionan acceso a un conjunto de técnicas de representación y de interacción de bajo nivel que apoyan las intenciones de nivel superior. En esta misma figura el discurso analítico es la relación entre estas técnicas y los objetivos y las tareas del usuario, que implican decisiones de bajo nivel sobre la manipulación de controles interactivos y metas de más alto nivel que rodean el problema que se investiga. El discurso se forma alrededor de la retroalimentación entre los objetivos del usuario y el resultado del análisis de una representación en particular o cambiarla a través de la interacción.

Las técnicas de representación -el desarrollo de lo que es normalmente del campo de la visualización de la información- son artefactos de bajo nivel que apoyan la intención de un usuario para representar la información de tal manera que logre una determinación particular, es decir una representación de los datos o una interacción específica con estos que apoye la intención del analista. El reconocimiento de las diferencias entre las técnicas y las intenciones es fundamental para el desarrollo de herramientas de apoyo al análisis. La técnica no debe ser considerada como un fin en sí mismo, sino un medio para apoyar la comprensión de la información por parte del usuario.

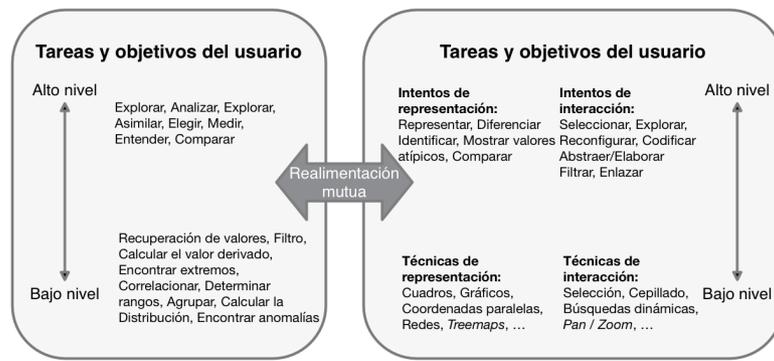


Fig. 6. Discurso analítico implica la retroalimentación mutua entre los objetivos y las tareas del usuario con las posibles comunicaciones y usos de las visualizaciones interactivas, adaptado de (Pike et al., 2009).

Becker, Cleveland y Wilks definen de manera compacta a la interacción como la manipulación, dirección y cambio instantáneo (Becker, Cleveland, & Wilks, 1987); asimismo posteriormente Dix (2009) sintetiza la interacción como "la comunicación entre el usuario y el sistema". Dado que la interacción puede ocurrir incluso con una imagen estática, como describe Spence (2007), la interacción no es ciertamente un concepto tangible. Sin embargo, las técnicas de interacción son menos difíciles de definir y son conceptos más tangibles que la interacción en sí. Una imagen estática no tiene una técnica de interacción asociada a pesar de que los usuarios pueden interactuar con ella. James Foley, van Dam, Feiner, and Hughes (1995) definen una técnica de interacción como una manera de utilizar un dispositivo de entrada/salida física para realizar una tarea genérica en un diálogo persona-ordenador.

En la visualización de la información las técnicas de interacción están más diseñadas para cambiar y ajustar la representación visual que para la introducción de datos en los sistemas, que es claramente un aspecto importante de la interacción (James Foley et al., 1995). Además, se considera a las técnicas de interacción como las características que proporcionan a los usuarios la capacidad de manipular e interpretar representaciones directa o indirectamente. De acuerdo con este punto de vista, una imagen estática o una representación autónoma de animación no tienen asociadas técnicas de interacción. Sin embargo, en una interfaz, un menú para cambiar de un gráfico de dispersión a una vista de coordenadas paralela es una técnica de interacción, ya que permite a los usuarios manipular una representación, a pesar de que puede ser menos interactivo o directo, y sin embargo es capaz, mediante una representación de los

mismos datos pero de diferente manera, de propiciar la creación de hipótesis o de nuevo conocimiento acerca de estos (Yi, Kang, Stasko, & Jacko, 2007). Las taxonomías de las técnicas de interacción son útiles para lograr una mejor comprensión del espacio de diseño de interacción. Algunas tratan de categorizar las técnicas de interacción de bajo nivel (por ejemplo, (Buja, Cook, & Swayne, 1996; Chuah & Roth, 1996; Dix & Ellis, 1998; D. A. Keim, 2002; B. Shneiderman, 1996; Wilkinson, 2005)); algunas proporcionan dimensiones para describir las técnicas de interacción (por ejemplo, (Spence, 2007; Tweedie, 1997)); otras se mueven más allá de las técnicas de interacción de bajo nivel para ofrecer una visión más amplia de la interacción que incluya nociones tales como espacios de interacción y los parámetros (por ejemplo, (M. Ward & Yang, 2004)); mientras que otras se centran más en las tareas de los usuarios (por ejemplo, (Amar et al., 2005; Michelle X. Zhou & Feiner, 1998)).

Esta divergencia sugiere que puede haber múltiples formas o granularidades para describir las técnicas de interacción, que también están en consonancia con el ciclo de la acción de Norman (2002) que describe la interacción entre un usuario y el mundo a través de múltiples pasos (es decir, la creación de la meta, el diseño de la intención, la especificación de una acción, la ejecución de la acción, la percepción de la situación del mundo, la interpretación del estado del mundo, y la evaluación del resultado).

Ware (2004) propone una taxonomía simple, pero expresiva en un nivel abstracto, que incluye las principales técnicas clásicas de interacción para la exploración y la navegación de datos. Esta no distingue claramente entre estos dos tipos de acciones, sin embargo, está más centrada en las acciones del usuario, no en las intenciones (es decir, se trata de una taxonomía de bajo nivel). Esta clasificación, sin embargo, está en línea con el modelo de referencia de la visualización de información de Stuart K. Card et al. (1999), donde las acciones de interacción siguen los fundamentos de visualización de la información.

Yi et al. (2007) proponen una nueva taxonomía basada en las intenciones del usuario. Esta incluye siete tipos de intenciones y sus técnicas asociadas. Tatiana Von Landesberger, Fiebig, Bremm, Kuijper, and Fellner (2014) hacen coincidir esta taxonomía de Yi et al. (2007) con la taxonomía de Ware (2004). Posteriormente, Gotz and Zhou (2008) crean una taxonomía de las actividades del usuario en el proceso analítico con tres categorías principales: acciones de exploración, las acciones de obtención de conocimiento, y acciones de metadatos. Se puede notar que las acciones de exploración corresponden aproximadamente a las acciones de interacción discutidas anteriormente por Ware (2004) y Yi et al. (2007) con respecto a la visualización de la información. Asimismo, con respecto a la actividad analítica, Gotz and Zhou (2008) mejoran los tipos de interacciones con la clasificación de las acciones de obtención de conocimiento y acciones de metadatos, que en su mayoría fueron olvidadas antes de que fuesen referenciadas por estos y después categorizadas en (Tatiana Von Landesberger et al., 2014) como la parte de las interacciones de razonamiento de la analítica visual. No obstante, tanto Yi et al. (2007) como Gotz and Zhou (2008) utilizan las intenciones del usuario para caracterizar interacciones.

Por otro lado Heer and Shneiderman (2012) proponen una taxonomía de las dinámicas interactivas del análisis visual, que se compone de tres categorías principales: especificaciones de datos y visuales, vista de manipulación y proceso y procedencia. Las dos primeras categorías se refieren a los cambios en la representación visual de los datos (visualización de datos, filtrar, ordenar, obtener datos, navegar, seleccionar, enlace y organizar visitas). La tercera categoría se refiere a acciones de razonamiento (historias de análisis de registro, anotaciones de patrones encontrados, compartir puntos de vista y la asistencia a los usuarios a través del proceso de análisis). Un aspecto interesante de esta es la inclusión de los aspectos de colaboración de proceso analítico (compartir puntos de vista y guiar a los usuarios), sin embargo, en la parte racional es similar a la taxonomía de (Gotz & Zhou, 2008).

La Tabla 4 resume varios estudios en visualización de la información que proponen taxonomías de importancia para el estudio de las técnicas de interacción. Esta ha sido modificada del estudio de Yi et al. (2007), que presenta un amplio panorama de las taxonomías de interacción, las tareas y las dimensiones para la visualización de información. A pesar de que muchos de los estudios comparten unidades comunes, las taxonomías tienen niveles significativamente diferentes de granularidad.

Publicaciones	Unidades taxonómicas
<i>Las taxonomías de técnicas de interacción de bajo nivel</i>	
<b>B. Shneiderman (1996)</b>	Información general, <i>zoom</i> , filtro, detalles a pedido; relacionar, crear una historia, y extraer.
<b>Buja et al. (1996)</b>	Centrándose (elección de [proyección, relación de aspecto, <i>zoom</i> , <i>pan</i> ], la elección de [la variable, orden, escala, relación de aspecto de escala-, animación, y la rotación 3-D]), la vinculación (cepillado como acondicionado /seccionamiento /consulta de base de datos), y la organización de visitas (matriz gráfica de dispersión y graficado condicional).
<b>Chuah and Roth (1996)</b>	Operaciones de la Interacción básica de la visualización (BVI): gráficas (codificar datos, ajustar el valor de la gráfica, manipular objetos), de conjuntos (crear conjuntos, borrar conjuntos, resumir conjuntos, otros), y las operaciones de datos (añadir, eliminar, derivar atributos, otros).
<b>Dix and Ellis (1998)</b>	Resaltando y centrarse, Acceder a información adicional: deslizándose hacia abajo y los hipervínculos, Información general y contexto: <i>zoom</i> y ojo de pez, Misma representación/cambio de parámetros, Mismos datos/cambio de representación, Vinculación de la representación-uni6n temporal.
<b>D. A. Keim (2002)</b>	Proyecciones dinámicas, filtrado interactivo, <i>zoom</i> interactivo, distorsión interactiva, cepillado y

	vinculación interactiva.
<b>Amar et al. (2005)</b>	Recuperar de valores, filtro, calcular el valor derivado, encontrar valor extremo, clasificar, determinar rango, caracterizar la distribución, y encontrar anomalías, agrupaciones, y correlacionar.
<b>Wilkinson (2005)</b>	Filtrado (categórico, continuo, múltiple, filtrado rápido), navegación ( <i>zoom</i> , panorámica, lente), la manipulación (arrastrando el nodo, reordenamiento categorizado), el cepillado y el enlace (cepillo de formas, lógica cepillo, cepillado rápido), animación (animación por fotograma), rotación, transformación (por especificación, montaje, exhibición, grifo, grifos 2, 3 tomas).
<b><i>Dimensiones taxonómicas de técnicas de interacción</i></b>	
<b>Tweedie (1997)</b>	Tipos de interacción (manual, mecanizada, enseñable, dirigibles y automáticas). La franqueza (directa e indirecta manipulación).
<b>Ware (2004)</b>	Clases de bucles de retroalimentación de menor a mayor nivel: <b>Bucle de minería de datos:</b> se seleccionan los objetos y mover usando la coordinación ojo-mano. <b>Exploración y navegación bucle:</b> la búsqueda de un camino en el espacio de datos y construyendo así un modelo mental de los datos. <b>Resolución de problemas de bucle:</b> analista forma hipótesis sobre los datos y los refina a través de un proceso de visualización aumentada.
<b>Lee, Plaisant, Parr, Fekete, and Henry (2006a)</b>	Tarea Gráficas: tareas basadas en topología, las tareas basadas en atributos, las tareas de visualización, y la tarea de resumen.
<b>Spence (2007)</b>	Modos de interacción: Continuo, intensivo, pasivo, y la interacción compuesto.
<b>Bertini and Lalanne (2009)</b>	Se miran si solo la acción del usuario cambia el estado actual (afinación, cambio de parámetros) o la cambia por completo (cambio del esquema): <b>La manipulación y puesta a punto:</b> cambiar los parámetros en el contexto de un esquema dado (En la visualización: cambio de parámetros de representación –por ejemplo <i>zoom</i> , etc.-; y en la minería de datos: cambian los parámetros del modelo –por ejemplo, el cambio de funcionalidad distancia). <b>Cambiar el esquema:</b> el cambio del modelo de datos o representación (en visualización: cambiar el mapeo visual o representación, y en minería de datos: cambio del modelo de datos – por ejemplo, el cambio de generación de reglas para encontrar clústeres-).

***Taxonomía de las operaciones de interacción***

<b>M. Ward and Yang (2004)</b>	Interacción: Operadores de interacción (navegación, selección, distorsión); Espacios de interacción (espacio de pantalla, espacios del valor de los datos, espacio de la estructura de los datos, espacio de los atributos, espacio de los objeto, y la visualización del espacio de la estructura); Parámetros de interacción (enfoque, extensión, transformación, y mezclador).
--------------------------------	---

***Las taxonomías de tareas del usuario***

<b>Michelle X. Zhou and Feiner (1998)</b>	Tareas visuales relacionales (asociado, trasfondo, categorizar, agrupar, comparar correlacionar, distinguir, enfatizan, generalizar, identificar, localizar, crear rangos, revelar, conmutar).
---	--

Tareas de organización visual directa y de codificación (ENCODE).

<b>Gotz and Zhou (2008)</b>	Tres categorías principales de las actividades del usuario en el proceso analítico:
-----------------------------	---

**Acciones de exploración:** estos se dividen en la exploración de datos (Filtro, Inspección, Consulta, restauración) y en la exploración visual (Cepillado, Cambio de Metáfora, cambio de rango, *Zoom*, *Pan*, fusionar, clasificar, *Split*);

**Acciones de introspección:** en lo visual (anotar, marcar) o basada en el conocimiento (eliminar, modificar, crear).

**Meta-Acciones:** cuenta con los siguientes tipos: rehacer, deshacer, volver, eliminar.

<b>Yi et al. (2007)</b>	Seleccione (marcas de algo tan interesante).
-------------------------	--

Explora (panorámica y paseo directa).

Vuelva a configurar, Codificar, Resumir/Elaborar (detalles en la demanda, enfoque en los datos, información sobre herramientas, y *zoom*).

Filtrar (consultas dinámicas).

Conectar (resaltar asociaciones y mostrar elementos de datos ocultos que son relevantes para un artículo específico).

<b>Heer and Shneiderman (2012)</b>	Datos y Especificaciones Ver, Ver Manipulación y Proceso y procedencia. Las dos primeras categorías se refieren a cambios en la representación visual de los datos (visualizar datos, filtrar, ordenar, derivar datos, navegar, seleccionar, enlazar y organizar visitas). La tercera categoría con acciones de razonamiento (historias de análisis de registros, anotación de patrones encontrados, compartir puntos de vista y guiar a los usuarios a través del proceso de análisis).
------------------------------------	--

<p><b>Tatiana Von Landesberger et al. (2014)</b></p>	<p>Visualización:</p> <p><b>Cambios en los datos:</b> selección: incluye filtrado (consultas dinámicas), detalles sobre la demanda; edición: edición de datos a través de la manipulación directa.</p> <p><b>Cambios de visualización:</b> (cambios en la representación): cambios de esquema (incluyen el cambio entre otras cosas, del tipo de visualización o de mapeo visual); Los cambios de parámetros (incluyen cambios de vista (p.ej., <i>zoom</i>, panorámica) o el cambio de la combinación de colores utilizada).</p> <p>Razonamiento:</p> <p><b>Cambios de datos:</b> (seguimiento del proceso analítico, Edición: anotación).</p> <p><b>Cambios de razonamiento:</b> Cambios en representación- (cambios del Programa: el cambio de tipo de análisis, los cambios de parámetros: deshacer / rehacer, etc.).</p> <p>Procesamiento de datos:</p> <p><b>Cambios de datos:</b> (Selección: filtrado de los datos utilizados para el procesamiento de datos; Edición: creación de nuevos datos a través de procesamiento de datos (por ejemplo, la transformación de datos)).</p> <p><b>Cambios de procesamiento:</b> Cambios en la representación- (cambios Esquema: el cambio de tipo de algoritmo de procesamiento; parámetro cambia: el cambio de parámetros algorítmicos).</p>
--	--

Tabla 4. Modificado de las taxonomías de importancia sobre las técnicas de interacción (Yi et al., 2007).

Por otra parte, además de la interacción en el área de la visualización de la información, la evaluación de la visualización y la búsqueda de estándares tienen una relevante importancia y se utilizan para determinar cómo de cerca está la visualización de su objetivo. En la evaluación se toman en cuenta muchas dimensiones a observar, que incluyen la evaluación de la usabilidad, cómo de eficientes son las técnicas de visualización, cómo de precisas, robustas y fáciles de usar son.

La visualización es una herramienta potente que puede utilizarse en diferentes procesos cognitivos: exploratorios, analíticos y descriptivos (Butler, Almond, Bergeron, Brodlie, & Haber, 1993). Existen muchas pautas para el diseño visual en la visualización de la información, pero el principio básico se podría resumir en el mantra de la búsqueda de información:

*Vista general primero, después filtrado y acercamientos, a continuación detalles bajo demanda (del inglés, “Overview first, zoom and filter, then details-on-demand” (B. Shneiderman, 1996).*

Aunado a esto, E. Tufte (2001) introdujo los principios de diseño para una visualización de la información exitosa, que debería:

- 1) Mostrarlos datos.
- 2) Evitar la distorsión de lo que los datos tienen que decir.
- 3) Presentar los datos condensados.
- 4) Hacer los grandes conjuntos de datos coherentes.
- 5) Fomentar los procesos de inferencia.
- 6) Proveer diferentes perspectivas sobre los datos, desde lo global a lo particular y viceversa.
- 7) Tener un objetivo claro.

Posteriormente, J. J. Thomas and Cook (2006) añaden que las visualizaciones y técnicas de interacción tienen como objetivo:

- Facilitar la comprensión del enorme y en continuo crecimiento conjunto de colecciones de datos de múltiples tipos.
- Ofrecer marcos de desarrollo para el análisis de datos espaciales y temporales.
- Apoyar la comprensión de la información incierta, incompleta y a menudo engañosa.
- Proporcionar al usuario representaciones adaptables que permitan la plena conciencia de la situación, mientras que apoyan a la toma de decisiones y acciones detalladas.
- Apoyar a múltiples niveles de datos y abstracción de la información, incluida la integración de los diferentes tipos de información en una sola representación.

El término visualización de la "información", en contraposición a "visualización" en general, se utiliza generalmente para contrastarla con la visualización "científica". En la ciencia hay muchos fenómenos que tienen una conexión directa con el mundo físico, pero de algún modo invisible, por ejemplo, el flujo de aire alrededor de un ala de avión.

Estos datos científicos tienen a menudo la forma de campos de números de los vectores definidos continuamente durante un espacio 2D o 3D. Por el contrario, la visualización de información a menudo se refiere a veces a datos más complejo estructuralmente, pero casi siempre discretos: jerarquías, tablas, datos de puntos. Por otra parte, los datos de visualización de la información a menudo incluyen los datos categóricos (por ejemplo, el género masculino/femenino), así como los datos continuos (por ejemplo, altura). Los dos no son totalmente distintos, por ejemplo, los sistemas de información geográfica se refieren a datos sobre mapas 2D. Los métodos utilizados para mostrar el consumo regional de gasolina para fines de *marketing* no serán tan diferentes de los que muestran los patrones de temperatura promedio para la modelización del clima y, por supuesto, se podría querer utilizar estos dos conjuntos de datos para comprender los patrones de calentamiento global.

Con las mejoras en las interfaces gráficas de usuario y en los dispositivos de interacción, una comunidad de investigación dedicó sus esfuerzos a la información de visualización (Stuart K. Card et al., 1999; C. Chen, 2006; Spence, 2007; Ware, 2004). En algún momento, esta comunidad reconoció el potencial de la integración del usuario en el proceso de descubrimiento de conocimientos, las capacidades de interacción y transferencia de conocimientos y en la minería de datos a través de técnicas de visualización eficaces y eficientes. Esto llevó a la exploración visual de datos y minería de datos visual (D. A. Keim, 2001). Esta integración se amplió considerablemente en el ámbito tanto de la visualización de la información como de los campos de minería de datos, lo que resulta en nuevas técnicas y muchas oportunidades de investigación interesantes e importantes.

Dos de los primeros usos del término analítica visual fueron Pak Chung and Thomas (2004) y un año más tarde en la agenda de investigación y desarrollo para la Analítica Visual (James J. Thomas & Cook, 2005). Posteriormente, el término se utilizó en un contexto más amplio, que describe un nuevo campo multidisciplinar que combina diferentes áreas de investigación, incluyendo la visualización, la interacción persona-ordenador, el análisis de datos, la gestión de datos, el procesamiento de datos geo-espacial y temporal, y el apoyo a la toma las decisiones (Andrienko et al., 2007; D. Keim, Mansmann, Schneidewind, Thomas, & Ziegler, 2008). A continuación se describe el campo multidisciplinar de la Analítica Visual.

#### IV.4 Analítica Visual

Como se ha mencionado al inicio de este capítulo, actualmente, las capacidades de almacenamiento han estado facilitando el acopio de grandes cantidades de datos, los cuales, a menudo, se adquieren y se almacenan sin ningún filtro o procesamiento; lo que convierte a estas circunstancias en ventajas para la Analítica Visual (del inglés, *Visual Analytics*, VA), motivo por el cual se ha popularizado.



Fig. 7. La analítica visual integra visualización como núcleo en conjunto con disciplinas adyacentes y depende de la disponibilidad de instalaciones de infraestructura y evaluación adecuados (D. Keim et al., 2010).

Durante la definición de esta área, la ciencia se ha basado en teorías perceptivas y cognitivas que abarcan la integración entre cognición, percepción y acción: por otro lado, tiene en cuenta aspectos cognitivos fundamentales para la interacción como la atención y la memoria.

Una importante síntesis de aspectos de la percepción, la memoria y los sesgos cognitivos que afectan el análisis es la contribución de Heuer (2001). Su trabajo se centra en los límites fundamentales que restringen el proceso de análisis y proporciona métodos analíticos para la compensación de estas restricciones.

La VA se desarrolló con la finalidad de apoyar el proceso de razonamiento analítico en sí. La integración de la interacción a un nivel básico en la teoría de la percepción y la cognición potencia las visualizaciones interactivas (ver Fig. 7), que son una pieza esencial en las herramientas de VA, y que son diferentes a las visualizaciones estáticas (James J. Thomas & Cook, 2005).

Estas visualizaciones interactivas de la analítica visual integran disciplinas científicas para mejorar la división del trabajo entre el hombre y la máquina; combina el razonamiento analítico (que permite a los usuarios obtener una visión profunda de los datos, para soportar directamente la evaluación, la planificación y la toma de decisiones) con visualizaciones interactivas.

Estas visualizaciones interactivas son representaciones visuales acompañadas de técnicas de interacción con el objetivo de explotar el ancho de banda del ojo humano para ampliar la capacidad mental y permitir a los usuarios ver, explorar y entender grandes cantidades de información de forma simultánea. Además, se consideran técnicas de apoyo a la producción, la presentación y la difusión de los resultados analíticos que comuniquen la información en el contexto adecuado, que son el uso principal para una variedad de audiencias. Estos sistemas están sujetos a las fortalezas y debilidades de las capacidades perceptivas y cognitivas humanas (ver Fig. 8).

Un aspecto importante en la ciencia del razonamiento analítico es crear modos que representen datos en formas que ofrezcan interacción y permitan procesos de pensamiento para convertir datos en información, información en significado y significado en comprensión. Todas estas fases se explican en detalle en (James J. Thomas & Cook, 2005).

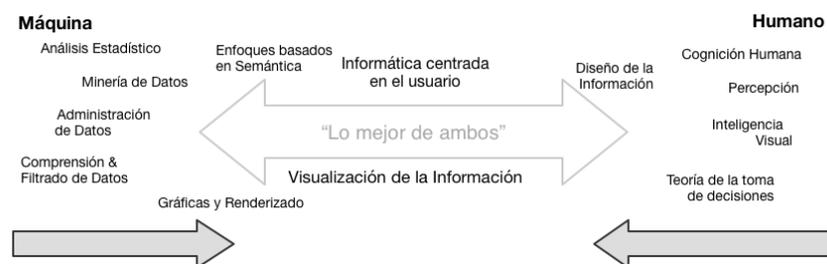


Fig. 8. La analítica visual integra disciplinas científicas para mejorar la división del trabajo entre el hombre y la máquina, modificado de (D. Keim, Andrienko, et al., 2008).

En este documento se sientan las bases de la ciencia Analítica Visual que se define como:

*La analítica visual es la ciencia del razonamiento analítico facilitado por interfaces visuales interactivas (del inglés, "Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces") (James J. Thomas & Cook, 2005).*

La VA combina un proceso analítico semiautomático con técnicas de visualización de la información, por tanto, se utilizan las fortalezas del procesamiento de datos electrónicos y las cognitivas del ser humano (ver Fig. 8). Es decir, la VA integra las capacidades analíticas de la computadora y las habilidades de las personas, lo que permite nuevos descubrimientos y la potestad de las personas para tomar el control del proceso analítico. La VA propicia el esclarecimiento de conocimiento oculto e inesperado que puede conducir a una innovación beneficiosa y rentable. Anteriormente se ha mencionado el mantra de la búsqueda de información, el cual se describe como: Vista general primero, después filtrado y acercamientos, a continuación detalles bajo demanda de B. Shneiderman (1996), por su parte, D. Keim, Andrienko, et al. (2008) amplían esta e idea y definen el mantra de VA como:

*Analizar Primero - Mostrar lo Importante - Acercamientos, filtros y analizar más - detalles sobre la demanda (del inglés, "Analyse First - Show the Important - Zoom, Filter and Analyse Further - Details on Demand") (Keim et al., 2008).*

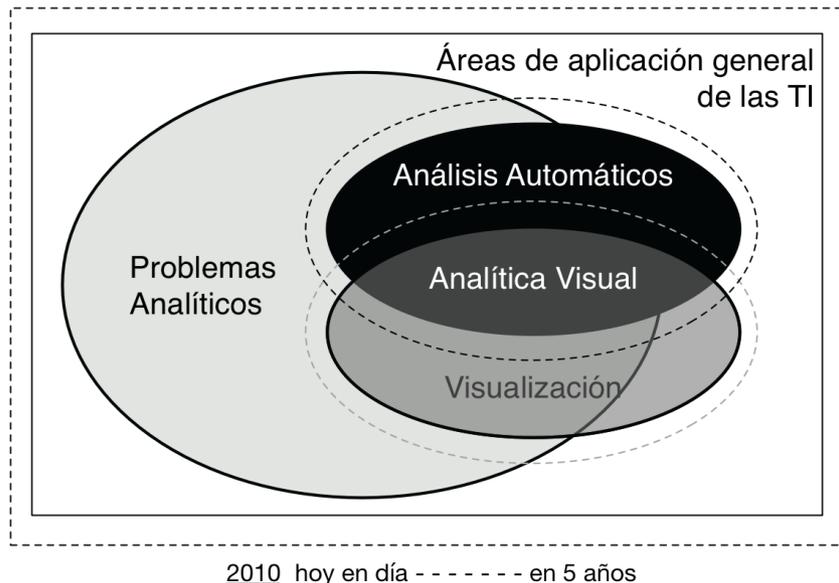


Fig. 9. Las dos clases de problemas. Se pueden observar los problemas generales de aplicación del área de TI (blancos) y los problemas analíticos (gris claro); ambos se pueden resolver mediante el Análisis Automático (negro), técnicas de Visualización (gris) y Analítica Visual (gris oscuro). Debe tenerse en cuenta que no todos los problemas de un análisis automático o análisis visual se pueden resolver por medio de la analítica visual, si existen otras maneras efectivas y eficientes de resolver el problema. Modificado de (Daniel A. Keim et al., 2010).

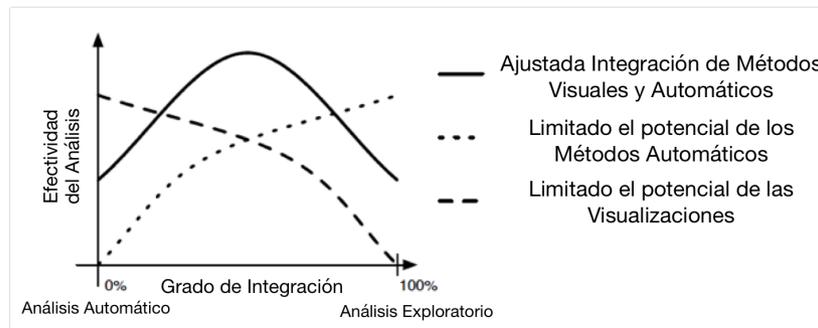


Fig. 10. Potencial de la analítica visual. Modificado de (Daniel A. Keim et al., 2010).

Más adelante, Daniel A. Keim, junto a otros autores, proporcionan una definición más actual y específica de lo que es la VA:

*La analítica visual combina técnicas de análisis automatizados con visualizaciones interactivas para crear un proceso de comprensión, el razonamiento y la toma de decisiones eficaces, basados en conjuntos de datos muy grandes y complejos (del inglés, "Visual analytics combines automated analysis techniques with interactive visualizations to create an effective understanding, reasoning and decision making process, based on very large and complex data sets") (D. A. Keim, J. Koblhammer, G. Ellis, & F. Mansmann, 2010).*

Las técnicas de análisis automático, como la Estadística y la Minería de datos, se han desarrollado independientemente de las técnicas de visualización e interacción. Sin embargo, algunos conceptos claves cambiaron el alcance bastante limitado de estas disciplinas en analítica visual. Uno de los pasos más importantes en esta dirección fue la necesidad de pasar del análisis confirmatorio al análisis exploratorio de datos (Ver Fig. 9 y Fig. 10).

Para que una herramienta de VA sea efectiva debe basarse en una comprensión profunda de cómo las personas perciben, razonan y crean nuevo conocimiento. Es importante para comprender mejor la abstracción y cómo la gente crea, evalúa y compara los "modelos mentales" que construyen en el proceso de VA (Fig. 11) en un primer sentido y luego tomar medidas basadas en estos modelos.

Entender la abstracción no solo apoya claramente el diseño de herramientas para crear (o ayudar a los usuarios a crear) abstracciones, sino también la capacidad de capturar el proceso de razonamiento y sus mecanismos. Esto implica y confirma el importante papel de los humanos en el proceso de análisis. Como se mencionó anteriormente, los seres humanos están capacitados para interpretar las impresiones visuales, pero cada persona posee una forma de interpretar representaciones, la cual muchas veces es errónea, y más aún cuando estas son inadecuadas. La frase "Uso de imágenes para pensar" (del inglés, "Using pictures to think") resume una idea fundamental de la VA, pero el análisis es un proceso que debe incluir acciones, sus interacciones, y sus etapas (Fig. 11).

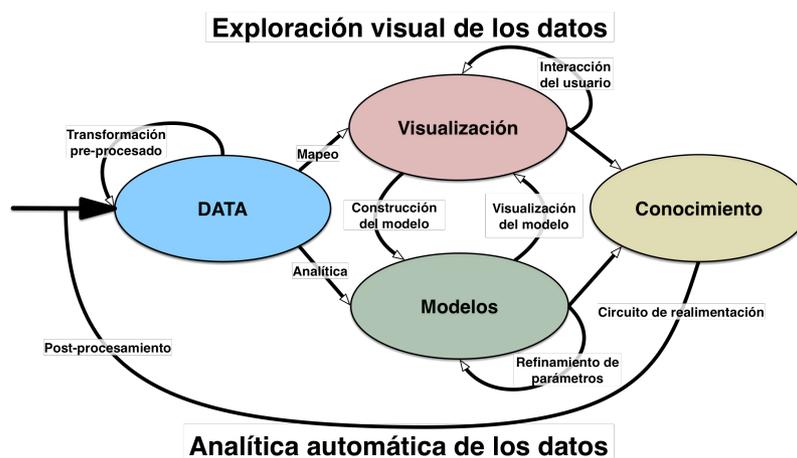


Fig. 11. Proceso de VA propuesto por Keim et al. Presentado y modificado de (D. Keim, Andrienko, et al., 2008).

Actualmente las fuentes de datos que se utilizan para la obtención de nuevo conocimiento son heterogéneas y distribuidas, por lo que el primer paso en cualquier proceso de VA es el procesamiento y transformación de los datos, que puede incluir una estandarización, normalización, agrupación y limpieza de los mismos (Fig. 11). Después de esto, el analista puede elegir entre realizar un proceso automático o visual (ver Fig. 10), para, mediante la alternancia entre métodos visuales y automáticos, lograr el mejor procedimiento que conduce a un refinamiento continuo y verificación de los resultados preliminares.

Es necesaria la interacción del usuario con la visualización para revelar información detallada, por ejemplo, hacer *zoom* sobre distintas áreas de datos, o considerar diferentes puntos de vista sobre los datos visuales. Los hallazgos en las visualizaciones se pueden utilizar para dirigir la construcción de modelos en el análisis automático.

En resumen, en el proceso de análisis visual, el conocimiento se puede obtener a partir de la visualización y el análisis automático, así como mediante las interacciones entre las visualizaciones, los modelos y los analistas humanos. El objetivo de la analítica visual es crear sistemas de *software* que apoyen el proceso visual de razonamiento analítico. Siguiendo este objetivo se utilizan técnicas y estrategias de visualización de la información (*infoVis*) y analítica visual, algunas de las cuales se describirán brevemente a continuación.

#### IV.5 Técnicas y estrategias de *infoVis*/VA

La visualización de la información y la analítica visual han generado numerosas técnicas de visualización, donde la visualización de estructuras jerárquicas ha generado numerosas técnicas (en el trabajo de H. Schulz, Hadlak, and Schumann (2011) se puede ver una revisión de las técnicas más relevantes). A continuación se describirán las que destacan por ser muy populares y/o las que están directamente relacionadas con el presente trabajo y se requieran para una mejor comprensión de este. En primer lugar se comenzará por las técnicas de diseño de visualización, posteriormente las de navegación y por último las de interacción.

### IV.5.1 Estrategias de diseño de visualización

El proceso de representación visual ofrece un punto de partida inicial para el diseño de visualización. Hoy en día son necesarios, a medida que aumenta la complejidad y el tamaño de los datos, métodos más avanzados para la representación de estos.

La identificación de las estructuras subyacentes dentro de la información a analizar, ayuda a orientar aún más el proceso de diseño. Estas estructuras proporcionan organización de alto nivel para un conjunto de datos y a menudo, proporcionan una guía para el diseño de visualizaciones apropiadas. Desde estas estructuras tienden a construirse los modelos mentales en los analistas. Estos atributos de los datos por lo general, son representados en una disposición espacial y conforman el diseño principal de la visualización. A continuación se definen algunas de estas estrategias de diseño de representación.

#### IV.5.1.1 Árbol jerárquico

Los árboles pertenecen a las estructuras con un enfoque de enlace, el cual utiliza diagramas de nodo de enlace. Los elementos se asignan a los nodos visuales, y las conexiones se asignan a los vínculos visuales entre los nodos. Así de forma más común se usan figuras geométricas para los nodos y líneas para los enlaces que los unen. Los diseños espaciales de árboles de nodo enlace pueden ser de tipo radial, hiperbólico o Cono en 3D, entre otros. Estos sistemas son útiles en la representación de estructuras solo con un atributo de datos expresado en una etiqueta de texto en los nodos. Este tipo de representaciones tienden a consumir mucho espacio debido a la cantidad de espacio blanco necesario dentro de cada uno de estos diseños espaciales (Modjeska, 1997).

#### IV.5.1.2 TreeMap

Los *TreeMap* son representaciones que solucionan el problema del espacio a la hora de representar grandes cantidades de datos con estructuras jerárquicas, llenando el espacio para maximizar el uso de todos los píxeles disponibles, y logrando así representar cantidades grandes de elementos. Los atributos de los datos se asignan a las propiedades de la forma, tamaño y color de los nodos, que comúnmente son rectángulos. Por tanto, los *TreeMap* son muy útiles en la visualización de los atributos no textuales, ya que en los casos de grandes cantidades, la representación se vuelve densa, limitando el espacio libre para las etiquetas de un nodo textual. Los nodos suelen estar dispuestos dentro de su nodo padre de acuerdo a una variedad de algoritmos existentes (Ben Shneiderman & Wattenberg, 2001; Vliegen, van Wijk, & van der Linden, 2006), formando generalmente en un rectángulo que contiene a sus nodos hijo y este a su vez a los suyos, generando una granularidad en la representación que refleja la profundidad.

#### IV.5.1.3 Coordenadas paralelas

Es un método de visualización de datos multivariados que se utiliza para los datos de n-dimensiones, en los que hay n ejes verticales (u horizontales) igualmente espaciados, y cada punto de datos se representa como una poli línea que cruza cada eje en una posición proporcional a su

su valor para esa dimensión, teniendo así una representación óptima para destacar patrones en los datos de muchas dimensiones en un espacio reducido (Ghanbari, 2007). A cada valor de datos le corresponde una posición a lo largo de la línea (dependiendo del valor del dato, se ubicaría entre el mínimo en la parte inferior y el máximo en la parte superior); una colección pura de puntos en estas posiciones de los datos no sería extremadamente útil, por lo que los puntos que pertenecen a un mismo registro (fila) se conectan con líneas, lo que ayuda a la comparación entre registros y el descubrimiento de patrones (Inselberg, 1985).

#### IV.5.1.4 Gráfico en espiral

Estos gráficos representan una simbiosis de dos técnicas y toman las ventajas de ambos; por un lado las gráficas circulares han demostrado ser útiles para exponer y comparar los comportamientos periódicos en pequeños conjuntos de datos de series de tiempo, y por otro lado, los gráficos clásicos de puntos, de barras y gráficos de líneas que ya han demostrado ser muy eficaces para el análisis de datos en serie. La estructura circular de espirales permite una fácil detección de los ciclos y la comparación de conjuntos de datos periódicos. Además, la continuidad de los datos se expresa mediante el uso de una espiral en lugar de un círculo (Weber, Alexa, & Müller, 2001).

#### IV.5.1.5 Matriz de adyacencia

Este tipo de visualización utiliza los atributos explícitos y relaciones entre los elementos de una colección. La visualización se basa en una matriz utilizada por Becker, Eick, and Wilks (1995), donde los ejes X e Y corresponden a una lista ordenada de los elementos en un subconjunto de la colección y las relaciones que se producen en este subconjunto. Es decir, si un elemento "i" está relacionado con un elemento "j", una pequeña forma corresponderá con la celda "i, j" de la matriz. Existen algunos trabajos donde extienden estas dos dimensiones a una tercera en el eje Z que corresponde a los diversos tipos de relaciones, así una dimensión z existe cuando el elemento "i" tiene múltiples relaciones para el elemento "j" y, por tanto, habrá varias pequeñas formas - tal vez una pila de cuentas - en varias posiciones Z (Hetzler, Harris, Havre, & Whitney, 1998).

#### IV.5.1.6 Gráfico de radar

El gráfico radial también se conoce como gráfico de radar, tabla *web*, gráfica de araña, mapa de las estrellas o gráfico de telaraña. Los primeros en utilizarlo fueron Smith, Hixon, and Horan (2001) y Becker et al. (1995), quienes mostraron la ubicación mediante rectángulos en un espacio de trabajo de dos dimensiones. Un gráfico de radar es un método de visualización de datos multivariados en la forma de un gráfico bidimensional de tres o más variables cuantitativas representadas en ejes a partir del mismo punto central (también llamados radios). Cada uno de los aspectos de los elementos de los datos forman ejes individuales que han sido dispuestos radialmente alrededor de un punto central. El valor de cada atributo de cada elemento se representa con un nodo (punto) en el radio (eje), la distancia desde el centro depende del valor de este atributo y la longitud de datos de un radio es proporcional a la magnitud del atributo

para el nodo siendo relativo a la magnitud máxima de los atributos a través de todos los valores de los datos. Se traza una línea que conecta los valores de datos para cada nodo generando el gráfico radial que muestra todos los valores de los atributos de un elemento en una sola línea (Gutwin & Greenberg, 2002).

#### IV.5.1.7 Nubes de palabras

Las nubes de etiquetas conceptualmente se asemejan a los clásicos histogramas, mientras los histogramas se utilizan normalmente para representar las frecuencias de una docena de elementos, las nubes de etiquetas pueden representar las frecuencias de un centenar de elementos, un ejemplo claro es el del trabajo de Ya-Xi, Santamaría, Butz, and Therón Sánchez (2009). Además al mostrar etiquetas como hipervínculos, se obtiene la posibilidad de interacción de la cual carecen los histogramas. En el etiquetado la fuente de la etiqueta representa la frecuencia de esta en la colección de elementos a analizar y la coloración y la posición puede ocuparse para atributos como el tema o la temporalidad de cada elemento, modificando la estructura y apariencia de la nube de palabras (Kaser & Lemire, 2007; Millen, Feinberg, & Kerr, 2005).

#### IV.5.1.8 Mini gráfico (*Sparkline*)

Un mini gráfico es una pequeña intensa gráfica, simple, del tamaño de una palabra con la resolución tipográfica. Sparkline significa que los gráficos son un mini gráfico puede estar en todas partes, en una palabra o un número, puede ser incrustado en una oración, titular, mapa, hoja de cálculo, gráfico, etc.(Bongshin, Riche, Karlson, & Carpendale, 2010; Edward R Tufte, 2006).

#### IV.5.1.9 Metáforas

Una metáfora permite a un concepto construir parcialmente otro, mientras que al mismo tiempo oculta aspectos del concepto creado. Las metáforas no establecen asignaciones completas entre conceptos; sino que mejoran algunos aspectos y suprimen otros, estas se cree que surgen directamente de las experiencias físicas, sociales y culturales. Así, cada experiencia es una estructura coherente que tiene varias dimensiones: participantes, partes, etapas, secuencias lineales, su causalidad y propósito, y estas son a menudo descritas por metáforas emergentes, formas básicas que pueden estructurar los conceptos más difíciles, por ejemplo, las emociones. Conceptos complejos a menudo requieren dos o más metáforas, para estructurar parcialmente el concepto. Es decir, cada metáfora subordinada representa un aspecto del concepto complejo, y estas metáforas cohesionan juntas. Un ejemplo sencillo es del de las metáforas orientacionales, que se utilizan posiciones como arriba-abajo, dentro-fuera, delante-detrás, *on-off*, de profundidad superficial, y centro-periferia (estas metáforas se utilizan en la navegación en línea) (Modjeska, 1997).

#### IV.5.1.10 Múltiples vistas enlazadas

La técnica de vistas enlazadas nace, por un lado, de la necesidad de los usuarios de ver los datos complejos e intrincados; intentando explorar y descubrir algunos hechos que no son fáciles de encontrar.

Estas investigaciones complejas requieren que el usuario considere muchos escenarios, para comparar las visualizaciones generadas a partir de varios conjuntos de datos diferentes, para agregar y minar los datos, o tal vez los datos de múltiples y diversos conjuntos de datos para generar nueva información, y ser capaz de revertir fácilmente a una iteración o filtro anterior. Por otra parte, los usuarios pueden estar muy familiarizados con las técnicas del uso y, por tanto, pueden estar perdiendo la riqueza de datos subyacentes. Así, mediante la utilización de un entorno de diseño de visualización que permite al usuario examinar las diferentes representaciones y también gestiona sus interacciones y automáticamente enlazar las operaciones entre las vistas, se pueden percibir las relaciones y hechos nuevos e interesantes de sus datos. Esto quiere decir que cualquier tipo de manipulación de selección, exclusión o realzado de un subconjunto de los datos en una de las representaciones visuales, se verá reflejado en todas las representaciones que la herramienta tiene a disposición con esta técnica implementada, permitiendo al usuario ver los datos en diversas formas, para manipular la presentación visual de diferentes maneras, y también interactuar y coordinar la interacción entre los diferentes puntos de vista (Roberts, 2007).

Después de emplear una estrategia de información general para proporcionar una amplia visión de una gran cantidad de información, la siguiente preocupación es el diseño de la navegación a través del espacio de información.

#### **IV.5.2 Estrategias de diseño de navegación**

Se necesitan métodos interactivos para apoyar la navegación entre una visión general del conjunto de los datos y de los detalles de toda la información. Para apoyar esta necesidad existen las siguientes estrategias de diseño de navegación primaria:

##### **IV.5.2.1 Planos visuales y contextos (*layering and separation*)**

La confusión y el desorden son fallas de diseño y no se encuentran en los atributos de la información, como dice Tufte, por lo que es importante encontrar las estrategias de diseño óptimo que revelen la complejidad de los datos y sus detalles. Una de las técnicas más poderosas para reducir el ruido y enriquecer la expresividad de la representación es la creación de planos visuales y contextos. La creación de planos eficiente basa su utilidad en la idea de la unión de más de uno estos planos, capa de atributo de los datos, interactuando en un terreno plano que genera patrones de información y texturas de los atributos de los datos, dotando al usuario mediante el intercambio e interacción de estos de diferentes aspectos y patrones del conjunto de datos (E. Tufte, 1990).

##### **IV.5.2.2 Vista panorámica y *zoom* (*zoom + pan*)**

Con el objetivo de una buena manipulación de conjuntos de datos de gran tamaño y complejidad, las visualizaciones ampliables comienzan con la información general, una vista panorámica, y luego permiten a los usuarios ampliar de forma dinámica en el espacio de información para llegar lo con datos de interés. Los usuarios pueden salir del *zoom* para volver a la vista

general, y de nuevo acercar a otra parte. Los usuarios también pueden desplazarse por el espacio sin reducir el *zoom*. El *zoom* puede ser una navegación suave y continua a través del espacio de la vista panorámica, o puede ser utilizado para avanzar a través de niveles discretos de escala como en *TreeMap*. Esta técnica proporciona una interacción a través de los niveles de *zoom*, el principal problema que queda es la desorientación al hacer *zoom*.

#### IV.5.2.3 Vista general más detalle (*overview + detail*)

Esta estrategia utiliza múltiples puntos de vista para mostrar simultáneamente una visión general y una vista de detalle. Un indicador de campo de visión en la vista general, es decir un elemento visible, indica la ubicación de la vista de detalle dentro del espacio de información. Las vistas están enlazadas de tal manera que la interacción en la visualización de la vista general provoca un cambio en la vista de detalle para navegar en consecuencia. Asimismo, cuando los usuarios navegan directamente en la vista de detalle, las actualizaciones de la vista general se reflejan para proporcionar información de la ubicación (Spence, 2000).

#### IV.5.2.4 Foco + contexto (*foco + context*)

Durante la exploración e interacción de representaciones visuales uno de los principales problemas es cuánta información mostrar en la vista general y en la vista en detalle. Para contribuir en la solución de este problema la técnica de foco más contexto se basa en tres premisas: en primer lugar, el usuario necesita tanto información general (contexto) y la información de detalle (foco) de forma simultánea. En segundo lugar, la información necesaria en el resumen puede ser diferente de la que se necesita en detalle. Finalmente, estos dos tipos de información se pueden combinar dentro de una sola pantalla (dinámicamente), igual que en la visión humana (Stuart K. Card et al., 1999).

Además de las técnicas de diseño de las representaciones visuales y las estrategias de navegación como apoyo para la visualización y análisis de los datos existen mecanismos que se proporcionan al usuario para interactuar con las herramientas

### IV.5.3 Técnicas y estrategias de interacción

#### IV.5.3.1 Vista de ojo de pez

Los usuarios generalmente requieren información relativa al contenido al interactuar con las estructuras de información, que a menudo se pierden y/o les falta información contextual interpretativa, lo cual supone un problema central en los sistemas de información debido a la falta de espacio de exhibición para mostrar toda la información disponible. Furnas (1986) propuso generalizar vistas de ojo de pez como una solución a este problema. Esta técnica imita la estructura de percepción del ojo humano: muestra los detalles locales y el contexto global simultáneamente. Una vista de ojo de pez es un gráfico que muestra un área de interés bastante grande en detalle, el foco, y muestra el resto de la gráfica, el contexto, sucesivamente, más pequeño y con menos detalle.

Esto logra una integración sin problemas de los detalles locales y el contexto global mediante el reposicionamiento de elementos de la gráfica el cambio de tamaño, donde el punto central del foco es el punto de enfoque; comúnmente este punto de enfoque se relaciona con el puntero del ratón y a medida que se arrastra cambia el foco y el contexto suavemente en la pantalla actualizándose en tiempo real. El tamaño y el detalle de un vértice en la vista de ojo de pez dependen de la distancia del vértice del foco, comúnmente los valores de algunos parámetros son controlados y pre asignados por el usuario.

Una lente de ojo de pez parece tener todas las ventajas de los otros enfoques para la visualización y consulta gráfica, pero sin sufrir ninguno de los inconvenientes. En un experimento de usabilidad los puntos de vista de ojo de pez generalizados superan significativamente vistas jerárquicas planas tradicionales (Furnas, 1986; Sarkar & Brown, 1994).

Las deformaciones en la mayoría de las representaciones de ojo de pez se realizan por medio de deformaciones en los objetos visuales, estas distorsiones se realizan por medio de dos funciones de transformación (*bifocal display* o *multifocal display*) y pueden involucrar una o dos dimensiones (Leung & Apperley, 1994).

#### IV.5.3.2 *Zoom* y *zoom* semántico

Otra técnica gráfica para equilibrar el detalle y el contexto es conocida como interfaces de *zoom* semántico o multiescala. Un *zoom* físico, por una parte, cambia el tamaño y el detalle visible de los objetos. El *zoom* semántico, por otra parte, cambia el tipo y el significado de la información mostrada por el objeto. El *zoom* semántico evita las distorsiones físicas de puntos de vista de ojo de pez, mediante el uso de una transición semántica entre vistas detalladas y generales de información (Modjeska, 1997).

#### IV.5.3.3 Cepillado (*brushing*)

Proceso en el que un usuario puede resaltar, seleccionar o eliminar un subconjunto de elementos que se muestra gráficamente por medio del señalamiento de los elementos con un ratón u otro dispositivo adecuado de entrada. En situaciones donde se muestran simultáneamente múltiples puntos de vista de los datos (por ejemplo, diagramas de dispersión), el cepillado es a menudo asociado a un proceso conocido como vistas enlazadas, en el que los elementos de cepillado en una sola vista afectan a los mismos datos en todos los otros puntos de vista. Una de las técnicas de cepillado fue aplicada por primera vez a los diagramas de dispersión de alta dimensional permitiendo a los usuarios obtener información sobre las relaciones espaciales de N dimensiones destacando los datos que caen dentro un sub espacio especificado por el usuario (Becker & Cleveland, 1987; M. O. Ward, 1994)

### IV.6 Alfabetización de datos e información

Una vez definidas las técnicas y estrategias de *infoVis* y VA es importante mencionar que en el capítulo anterior se han descrito la AA y LA, donde se ha hablado de los líderes. Estos líderes son parte del proceso de analítica, ya sea académica o de aprendizaje, líderes educativos y "a

menudo no tienen idea de qué significan los datos o cómo usarlos" (Earl & Katz, 2006), y no hay que sorprenderse de que se sienten inseguros acerca de sus esfuerzos en DDDM (Wu, 2009). Dado que en la AA y LA se realizan procesos estadísticos y representaciones gráficas para la presentación de resultados y en la visualización de la información y en la VA se utilizan técnicas de visualización e interacción para la representación de los datos, estos procesos se ven influidos por las capacidades y habilidades de los usuarios. Los desarrolladores, a su vez, incluyen consideraciones humanas, la creación de sistemas de visualización más usables. El desarrollo de los principios científicos y de ingeniería para la generación de visualizaciones (para los usuarios con diversas necesidades y capacidades) y una metodología para la solución de problemas con la visualización de información son muy necesarios (Gershon et al., 1998).

Además, como el uso de la Web se populariza y se vuelve más sofisticado, repercutiendo en un gran impacto en la manera en que se presenta la información. Con el creciente uso de los medios visuales, la alfabetización visual de la población aumentará, y como resultado los usuarios estarán más cómodos interactuando con los objetos visuales y, por tanto, obtendrán más información de ellos. Estas capacidades de los usuarios se corresponden con las habilidades del ser humano, su percepción visual y la destreza en el reconocimiento de patrones y estas se pueden desarrollar a partir de la llamada alfabetización de los datos o de la información, que a continuación se detallan.

La alfabetización de los datos también se conoce como alfabetización estadística, alfabetización cuantitativa o aritmética, o alfabetización de la información. Esta alfabetización es la capacidad de entender las tablas, los gráficos, los datos y los conceptos principales y terminologías de la estadística. Ellen B. Mandinach and Gummer (2013) definen la alfabetización de los datos como la habilidad de entender y usar los datos de manera efectiva para tomar decisiones informadas. Esta alfabetización de los datos está compuesta por un conjunto base de habilidades y conocimientos que habilitan para transformar los datos en información y finalmente en conocimiento para llevar a cabo una acción. Este conjunto de habilidades comprende el conocimiento de cómo identificar, coleccionar, organizar, analizar, contabilizar y priorizar los datos. Además incluye la capacidad de realizar hipótesis, identificar problemas, interpretar los datos y determinar, planear e implementar acciones. Una persona alfabetizada en lo que se refiere a los datos, según J. A. Levin and Datnow (2012), se define como alguien que es capaz de:

- 1) Pensar en los efectos de los datos.
- 2) Reconocer datos poco sólidos y los sólidos.
- 3) Tener conocimientos acerca de los conceptos estadísticos y de medición.
- 4) Hacer una interpretación primordial (Earl & Katz, 2006).
- 5) Prestar atención a la presentación de informes y a las audiencias.

El problema, sin embargo, es que muchos gestores académicos o directivos no han tenido una formación adecuada en la comprensión, el análisis y la interpretación de datos (Ellen Beth Mandinach & Honey, 2008; Wu, 2009).

El personal de la Universidad Abierta del reino unido durante muchos años se ha venido ocupando de la preparación de los nuevos estudiantes que comienzan sus estudios en la Universidad y han proporcionado una serie de materiales que pueden ayudar con la instrucción y apoyo continuo a medida que estudian. Estos datos han sido base en el proyecto del Inventario de Aprendizaje Permanente Efectivo (ELLI del inglés, *Effective Lifelong Learning Inventory*), un instrumento de investigación validado que produce un perfil individual de aprendizaje y que ofrece un marco para la reflexión y el debate sobre los enfoques de aprendizaje, los contextos de aprendizaje y las limitaciones en el aprendizaje de los estudiantes y los educadores. Esto permite la identificación de las siete dimensiones del aprendizaje y escalas fiables para evaluarlas. Estas dimensiones parecen ser capaces de diferenciar entre los estudiantes eficaces, comprometidos y activos, de los discentes pasivos, dependientes y frágiles (Buckingham Shum & Deakin Crick, 2012; Crick, Broadfoot, & Claxton, 2004; Edwards, 2010).

En esta investigación la base de la comunicación y de la extracción del conocimiento se realiza mediante la autorreflexión y autodiagnóstico, por lo que es de suma importancia las capacidades y la alfabetización visual. Además de los resultados ya mencionados, se argumenta, tomando en cuenta las bases de la AA y LA del capítulo anterior y de la VA de este capítulo, que el análisis del aprendizaje visual debe dar un paso más: cuando este análisis se oriente a potenciar a los estudiantes, de los que se debe entender sus capacidades pedagógicas, el conocimiento se dará tanto para los educadores y los educandos; por un lado aprendiendo a cómo realizar una eficiente reutilización de los datos de las experiencias previas y por otro a instruirse sobre la manera más eficiente de aprender. Por otra parte, (Buckingham Shum & Deakin Crick, 2012) mencionan que lo que ha surgido del programa de investigación que ha acompañado el modelo de cambio que describieron en su trabajo, con el *warehouse* de aprendizaje como parte nuclear, es la evidencia inseparable de que la alfabetización visual es una parte crucial del desarrollo personal y social y que este cambio profundo se puede lograr por medio de prácticas pedagógicas.

El origen de la analítica educativa y académica, como se ha mencionado anteriormente, se transfirió desde el área de los negocios, de igual manera y como parte de este desarrollo se encuentra la idea de la gran importancia que tienen los datos en la toma de decisiones. Una organización basada en datos adquiere, procesa y aprovecha los datos en el momento oportuno para crear eficiencias, repetir y desarrollar nuevos productos, y navegar por el panorama de la competencia. La adquisición y procesamiento de los datos son primeros pasos en el trabajo estos. Las mejores organizaciones basadas en datos se centran implacablemente en mantener sus datos limpios. Los datos deben ser organizados, bien documentados, con formato consistente y libre de errores.

Las organizaciones exitosas invierten mucho en las herramientas, procesos y auditorías regulares de los datos. Estos han desarrollado una cultura que entiende la importancia de la calidad de los datos. El análisis de los datos no solo se realiza reuniendo estadísticas básicas, también se requiere desarrollar una intuición para cualquier defecto en los datos y poder ser capaz de explicar cualquier cosa inesperada de estos.

No es una cuestión de comprobar las estadísticas de una lista, sino de construir un modelo mental de lo que dicen los datos sobre el mundo (Shron, 2014). Para hacer el cambio hacia una cultura basada en datos también se requiere una infraestructura fundamental que incluya las herramientas tecnológicas de apoyo basadas en datos para la toma de decisiones, así como una alfabetización de los datos entre el personal para incluir estructuras de apoyo, procesos sistemáticos y recursos que faciliten implementar y mantener una cultura de datos. Además de esto, existen varios factores que se unen en el momento de fomentar el interés en hacer un mayor uso de la analítica y de otorgar la importancia que tienen los datos. Uno de ellos es el aumento de la disponibilidad, el detalle, el volumen y la variedad de datos sobre el uso casi omnipresente de las ICT a través de casi todas las facetas de la vida diaria. Este aspecto tiende a ser el foco de los medios de comunicación, pero los datos por sí solos no son suficientes para darse cuenta de los beneficios de la analítica.

Se puede finalizar este capítulo con la siguiente definición: la alfabetización de los datos para los profesores o alfabetización de los datos pedagógicos o alfabetización de los datos para el conocimiento de la enseñanza es la habilidad de transformar la información en conocimiento e instrucciones prácticas por medio de la recolección, el análisis y la interpretación de diversos tipos de datos con la finalidad de determinar directrices educativas.



El primero, constituido por representaciones visuales, está relacionado directamente con el *eLearning*, ya sea para la comprensión del proceso educativo, para el análisis de este o como representación de contenidos y/o información educativa.

Por otra parte, el segundo grupo de trabajos recopilados se centra en las técnicas y los procedimientos que influyeron en el desarrollo de la investigación del presente trabajo y que están relacionados directamente con *eLearning*, pero sin una representación visual.

Por último, el tercer grupo de representaciones visuales, está relacionado con otras áreas de investigación, pero cuyas metodologías y/o técnicas utilizadas se han tomado en cuenta o se han implantado para su experimentación con el objetivo de evaluar su efectividad y posible incorporación al ámbito de este trabajo.

A continuación se presenta la primera sección del capítulo en el que se describen herramientas visuales y sus técnicas, y, posteriormente, los problemas encontrados y las soluciones propuestas.

## V.1 Las representaciones visuales: Técnicas de interacción

La visualización de la información forma parte de la interfaz directa entre el usuario y la máquina. Como se mencionó en la sección IV.1, la visualización abarca tanto la comprensión de la imagen como la síntesis de la imagen, es decir, la visualización es una técnica o método para la interpretación de los datos de la imagen introducidos en un ordenador y para la generación de imágenes de los conjuntos de datos multidimensionales complejos (McCormick et al., 1987).

En la visualización de datos se representa gráficamente la información mediante formas más comprensibles de los datos y se revela información que se encuentra oculta en los contenidos de, por ejemplo, las plataformas de aprendizaje. El objetivo de la visualización es entender los datos que pueden recuperarse y analizarse, de forma que al representarse visualmente se proporciona una visión cuantitativa y cualitativa de los conjuntos de datos grandes y complejos, con lo que se consigue resumir los datos y ayudar a la identificación de regiones de interés con los parámetros adecuados para un análisis más enfocado o específico. Una visualización de éxito puede reducir el tiempo que se necesita para entender los datos subyacentes, encontrar relaciones u obtener la información requerida para la toma de decisiones.

Como se comentó en la sección IV.4, la VA es un campo multidisciplinar que incluye las siguientes áreas de enfoque: técnicas de razonamiento analítico; representaciones visuales; técnicas de interacción; representaciones de datos; transformaciones de datos; y técnicas de apoyo a la producción, presentación, y difusión de los resultados analíticos, para lograr que comuniquen la información en el contexto adecuado para una variedad de audiencias (James J. Thomas & Cook, 2005).

Por otro lado, el *eLearning* habitualmente se soporta técnicamente por LCMS o LMS, como se expresó en la sección II.3. Tales entornos de formación basados en la Web se caracterizan por el uso de grandes cantidades de datos, con una fuerte interactividad y sin restricciones en el

espacio y el tiempo. Estas plataformas de aprendizaje almacenan un registro de las actividades del curso de los estudiantes en una base de datos. Por lo general, los LMS, proporcionan servicios para la supervisión del estudiante, lo que permite al instructor ver determinados datos estadísticos, tales como el número de accesos realizados por el estudiante a cada recurso, un registro de las páginas visitadas, el número de entradas para cada día, etc. Los profesores pueden utilizar esta información para seguir las actividades de los estudiantes e identificar posibles problemas.

Sin embargo, esta información se proporciona generalmente en una forma que plantea varios problemas. Por ejemplo, la información de este tipo tiene una organización lógica muy simple que consiste en una enorme lista de números y los identificadores de cada elemento que interviene; en algunos casos la forma de presentarlos está muy centrada en el detalle, es decir, muestra los identificadores de curso, módulo y actividad, etc. y comúnmente esta información se muestra como un archivo de texto o por medio de gráficos muy básicos, lo cual separa al analista del contexto del curso y/o de toda la plataforma de aprendizaje electrónico de la institución, lo que complica el reconocimiento de patrones y la obtención de una visión global de lo que sucede. Además, estas plataformas tienden a ofrecer una interactividad limitada y como resultado el usuario se frustra con estas estadísticas tan sencillas, los datos de seguimiento de actividad en bruto son difíciles de interpretar, casi no se puede percibir lo que está sucediendo y, en muchos casos, los datos resultan ser prácticamente incomprensibles.

Aunado a esto, el sistema no ofrece la adquisición de nuevo conocimiento, ni expresa los patrones de uso, ni divulga casos específicos que requieran una atención especial. Hoy en día, como se expuso en el Capítulo II: El aprendizaje y la tecnología, la creciente utilización de las tecnologías para apoyar el aprendizaje ha fomentado la creación de herramientas que ayudan a extraer información que no está disponible a primera vista. Esto es esencial para la mejora del proceso de aprendizaje desde el punto de vista de quienes tienen la responsabilidad de tomar decisiones a nivel institucional, de los proveedores de contenidos educativos, de los profesores y de los estudiantes, los cuales se deberían beneficiar del uso de herramientas analíticas eficaces para las plataformas de *eLearning* actuales.

El *eLearning* y la VA han tenido un crecimiento y han obtenido excelentes resultados como campos separados de la investigación, sin embargo, se han hecho pocos esfuerzos para reunir a las dos disciplinas en un ambiente de cooperación con el fin de darse cuenta de su enorme potencial (Börner, 2012; Mottus, Kinshuk, Graf, & Nian-Shing, 2013; Spector, 2013).

Por otra parte, como se expresa en el Capítulo III: Analítica en la educación, en la literatura sobre la investigación educativa se ha establecido que el seguimiento del aprendizaje del estudiante es un componente crucial de la educación de alta calidad (Agudo-Peregrina, Iglesias-Pradas, Conde-González, & Hernández-García, 2014; Armstrong & Anthes, 2001; Dyckhoff, Lukarov, Muslim, Chatti, & Schroeder, 2013; Hardless & Nulden, 1999; Naps et al., 2002; Tempelaar, Heck, Cuypers, van der Kooij, & van de Vrie, 2013; Wolff, Zdrahal, Nikolov, & Pantucek, 2013).

Una vez que los instructores y los estudiantes no están en el ambiente tradicional presencial, se deben explorar nuevas formas de seguimiento de la actividad o del proceso de aprendizaje del estudiante. El uso efectivo de un LMS requiere que los instructores puedan proporcionar los medios adecuados para el diagnóstico de problemas, y así poder tomar medidas inmediatas con el objetivo de prevenir o superar las dificultades (Mottus et al., 2013). Siguiendo el enfoque de la VA (mencionado en la sección IV.4), el camino es el de proponer y desarrollar soluciones visuales interactivas e innovadoras con el fin de ayudar a los usuarios, con sus diferentes perfiles específicos en las plataformas de *eLearning*, a extraer conocimientos específicos relacionados con el complejo proceso de la enseñanza y el aprendizaje.

A pesar de la promesa sobre visualizaciones para el aprendizaje, se sabe relativamente poco sobre cómo las personas piensan o aprenden con visualizaciones y multimedia. Cuanto más se utilizan las visualizaciones complejas y más accesibles sean, más investigación será posible realizar sobre el efecto y la aceptación por parte de sus usuarios. Siguiendo este punto de vista constructivista, y como se mencionó en el Capítulo III, actualmente se está reconociendo cada vez más que el buen material de aprendizaje, incluyase contenido educativo o herramienta de apoyo y seguimiento, en la educación virtual es aquel que no solo permite a los estudiantes alcanzar los objetivos propuestos, sino que también facilita la oportunidad de aprender cómo aprender (Guardia, Mas, & Girona, 2005).

Pero para que las visualizaciones se diseñen y apliquen en situaciones de aprendizaje de forma óptima, se debe saber más acerca de los procesos cognitivos implicados en el aprendizaje con visualizaciones. También hay que estudiar los aspectos específicos de si (y cómo) las visualizaciones tienen impacto en los procesos de comprensión. Un modelo interno de un ser humano -mapa cognitivo-, se genera en base a lo que observa, como anteriormente se explicó en las secciones IV.1 y IV.2. La representación de la realidad mediante la tecnología es, en la mayoría de los casos, determinante en la aceptación y la comprensión de la solución a los problemas complejos. Además, como se ha explicado anteriormente, la visualización amplifica las capacidades cognitivas humanas (James J. Thomas & Cook, 2005).

Desde 1990 ha habido una serie de estudios que examinan los efectos de los diagramas frente al texto en materia del rendimiento de los usuarios de computadoras. También ha habido estudios sobre los efectos de las metáforas sobre el aprendizaje y la búsqueda de información (Butcher et al., 2006; Butcher & Kintsch, 2004; Hsu, 2007), en estos se compararon los efectos de las metáforas visuales frente a las verbales en cuanto a facilitar a principiantes y expertos la comprensión y construcción de modelos mentales, obteniéndose diferencias significativas entre la variable metáfora visual/verbal o la interacción entre la metáfora y el experto. El estudio de Butcher and Kintsch (2004) indica que los diagramas tienen efectos significativos en los estudiantes mediante la formación de modelos mentales, y que los sujetos con un alto grado de conocimiento forman mejores modelos.

En comparación con los diagramas detallados, diagramas simplificados pueden apoyar un mejor aprendizaje de hechos. En el estudio de Butcher et al. (2006) se demostró que los diagramas simplificados apoyan mejor la integración de la información, aunque ambos tipos de diagramas pueden aumentar la precisión de la comprensión y facilitar la generación de inferencias. En su investigación, Butcher and Kintsch (2004) estudiaron los procesos cognitivos y los modelos mentales de los estudiantes que utilizan diagramas visuales, además de texto.

Los participantes estudiaron con diagramas moderadamente detallados o solo con texto. Los diagramas más detallados enfatizaron información estructural correcta sobre el dominio y los diagramas menos detallados se simplificaron para enfatizar las relaciones funcionales. En ambos estudios, (Butcher et al., 2006; Butcher & Kintsch, 2004), la presencia de diagramas aumentó significativamente el número de inferencias producidas por los participantes, pero no influyó en cualquier otro proceso de aprendizaje (medido), seguimiento, interpretación o elaboración. Además, los participantes que usaron diagramas demostraron mayor desarrollo del modelo mental, pero el efecto fue más pronunciado para los estudiantes que utilizaron los diagramas simplificados. La interpretación de estos resultados conduce a que los diagramas con el diseño adecuado pueden promover con éxito el aprendizaje porque orientan correctamente al estudiante a participar en los procesos cognitivos esenciales para la comprensión.

El diseño de sistemas interactivos de aprendizaje electrónico multimedia se describe en (Granda et al., 2008). En ese trabajo, la investigación se centra principalmente en los aspectos de diseño de la pizarra compartida para permitir una visualización remota escalable. Entre estos materiales, los medios de comunicación notables incluyen: texto, ilustraciones, animaciones, audio, vídeo, hipertexto o simulaciones que ofrecen diferentes grados de abstracción y concreción. En el caso de las representaciones visuales, se propone una clasificación de acuerdo con su función educativa en la educación virtual.

Las representaciones visuales se pueden clasificar de diferentes maneras: por la forma en que se crea la representación visual (metafórica, no metafórica) (Stuart K. Card et al., 1999; Wilkinson, 2005); por el número de datos (una, dos o tres dimensiones, datos temporales y multidimensionales, datos de red y multidimensionales; para más información sobre tipos de datos en profundidad se remite al lector a (B. Shneiderman, 1996); por el tipo de dato representado (temporal, social y jerárquica, etc.) (Mackinlay, 1986); por el espacio de diseño de acuerdo a las etapas del proceso de visualización de (Chi, 2000; Daassi, Nigay, & Fauvet, 2006); o por los datos dependientes del tiempo (Aigner, Miksch, Muller, Schumann, & Tominski, 2008). Una clasificación de los diferentes tipos de representación visual, de acuerdo con su función educativa en el entorno de aprendizaje puede tener tres tipos de uso: como soporte para representar el contenido de aprendizaje, como una técnica de aprendizaje y como ayuda para representar el proceso de aprendizaje (Gómez Aguilar, Suárez Guerrero, et al., 2010).

En esta tesis doctoral se han tenido en cuenta solo los trabajos que entran en la segunda y tercera categoría, es decir, aquellos que hacen aportaciones en la comprensión o en la mejora del proceso de aprendizaje. Además, se incluyen las investigaciones que, ya sea por su representación o su metodología, están relacionadas con la presente tesis doctoral. Estos últimos se clasificarán de acuerdo a su metodología o su tipo de representación implantada dependiendo de cuál sea el caso.

A continuación se desarrollan tres subsecciones donde se describirán inicialmente las contribuciones visuales que apoyan al *eLearning*, después las contribuciones no visuales que apoyan al *eLearning* y, finalmente, los trabajos visuales que no apoyan al *eLearning*.

### V.1.1 Contribuciones visuales que apoyan al *eLearning*

Las representaciones visuales que se utilizan como técnicas de aprendizaje son conocidas como gráficos didácticos o imágenes didácticas (Prendes Espinosa, 1994) y también como organizadores gráficos (del inglés, *Graphic Organizers*) (Ewy, 2003). Estas herramientas visuales tienen como objetivo permitir a los propios estudiantes organizar el contenido de forma gráfica, lo que hace más eficiente el aprendizaje, así como el aumento de los factores de motivación de los estudiantes. Entre sus ventajas, se pueden destacar los siguientes elementos: su valor en la representación de información compleja, el procesamiento de información, la identificación de conexiones entre conceptos, la identificación de errores de conceptualización, la organización de las relaciones de jerarquía o de inclusión, y la mejora personal de la comprensión del aprendizaje (Carr & Mazur-Stewart, 1988; Daniel H Robinson & Kiewra, 1995).

Sin embargo, para que estos organizadores gráficos logren tener un efecto positivo en el aprendizaje, no es suficiente con simplemente usarlos; los profesores, así como los estudiantes, deben comprender sus principales características y, además, ser conscientes del objetivo para el que está destinado su uso (Gallavan & Kottler, 2007). Por tanto, se necesita un profesor con experiencia en esta área para sugerirlos, dado que cada organizador gráfico propone la representación de la información en una forma diferente, y se requiere de una cierta competencia de los estudiantes sobre el proceso visual y la organización de la información con el fin de lograr el mejor resultado.

Existe una gran variedad de organizadores gráficos, como mapas conceptuales (útiles para organizar y representar las relaciones de la jerarquía conceptual visualmente), mapa de ideas (útiles para promover la lluvia de ideas, elaborar planes de acción y analizar los problemas), diagramas de causa-efecto (útiles para el reconocimiento de las relaciones causales entre dos o más fenómenos, hechos o conceptos), líneas de tiempo (útiles para la creación de secuencias de acontecimientos o hitos de manera temporal), organigramas (gráficos útiles en la organización de las relaciones funcionales en una jerarquía), cíclicas (útiles para representar una secuencia de eventos que se repiten periódicamente), diagramas de flujo (útil para representar esquemáticamente la secuencia de una instrucción algoritmo o los pasos de un proceso), diagramas de Venn (útiles para establecer diferentes formas en que los conjuntos se pueden

relacionar), entre otros (Manoli & Papadopoulou, 2012; Daniel H. Robinson, 1997). Todo esto va dirigido a conseguir representar la información de forma gráfica para los estudiantes, lo que facilita su aprendizaje. Dada su utilidad, existe una línea de trabajo e investigación en educación que apoya el uso de estos organizadores gráficos en diferentes experiencias educativas y que se ha extendido gracias a los avances informáticos en la educación virtual (Inspiration Software, 2003). Hoy en día, se está reforzando el uso de herramientas web para crear estos organizadores gráficos y ayudar a los estudiantes con la construcción de sus modelos de conocimiento (Castaño Garrido, 2004).

Consecuentemente, existen propuestas en los campos de teoría educativa, estándares y ontologías en apoyo a la creación, organización y clasificación de los objetos de aprendizaje y para la estructuración y clasificación de contenidos educativos en los LCMS y/o LMS (Guo & Chen, 2006; Karunananda, 2006).

La mayor fortaleza de este conjunto de herramientas visuales es que obligan al individuo a hacer explícito lo que es aparentemente información incomprendible (Coll & Monereo, 2008). En algunas ocasiones el representar conceptos o ideas y objetos de aprendizaje plantea la circunstancia de expresar el mismo tema de diferentes maneras, esto significa que los conceptos semánticamente idénticos pueden expresarse por diferentes términos del vocabulario de dominio.

En su estudio, Guo and Chen (2006) integran un léxico de dominio en una ontología y de este modo definen asignaciones de términos del vocabulario de dominio a su significado tal y como se define por los conceptos de la ontología, por tanto, presentan el material de aprendizaje en diferentes contextos de aprendizaje o de presentación, lo que permite el acceso a determinado material de aprendizaje (ver la Fig. 12).

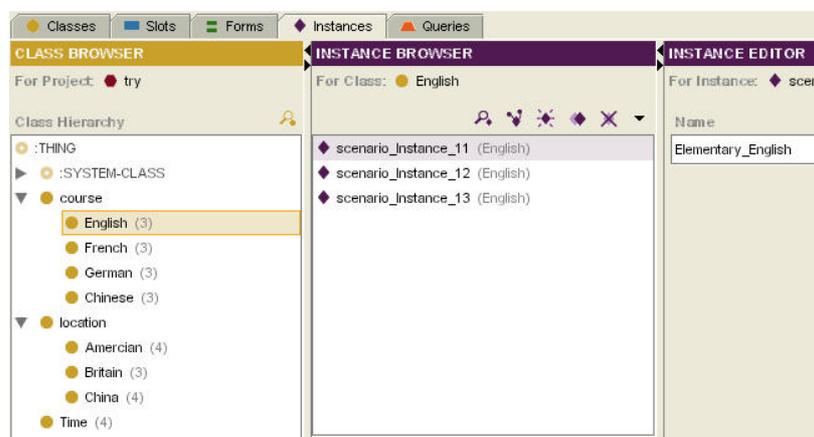


Fig. 12. Diseño y representación ontológica del escenario de *eLearning* de Guo and Chen (2006). Diseño y representación ontológica del escenario de *eLearning* de Guo and Chen (2006).

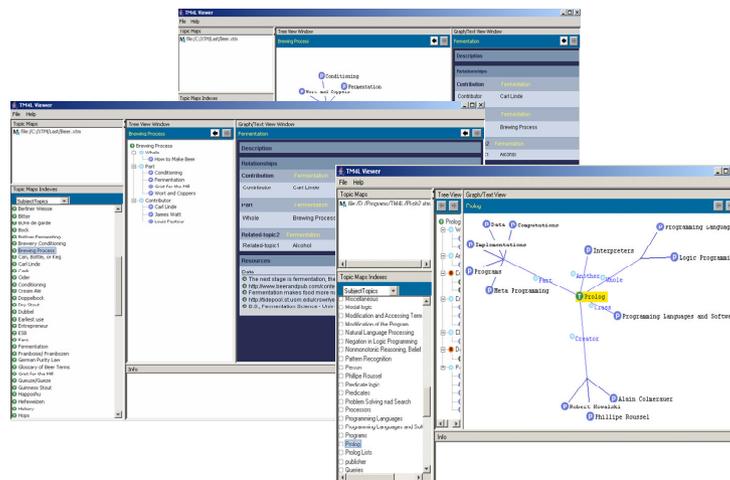


Fig. 13. Captura de pantalla del visor de TM4L (Dicheva & Dichev, 2006; Dicheva et al., 2005).

Otro grupo de trabajos abordan el uso de la visualización, pero no para análisis de la información, sino como parte de un objeto de aprendizaje o como recurso de un curso (Lauer, 2006; Rossling et al., 2006). Siguiendo con los trabajos que representan conceptos, Dicheva et al. (2005) desarrollaron una herramienta para crear y visualizar mapas conceptuales para el *eLearning*, (*Topic Maps For e-Learning*, TM4L), que hace uso de ontologías y proponen la visualización de mapas temáticos de colecciones de objetos de aprendizaje con el apoyo de la información semántica, además de su administración interactiva.

Esto con el fin de complementar los editores de mapas conceptuales (del inglés, *Topic Maps*, TM) existentes y herramientas de visualización para el área de *eLearning*. Este se basa en el proyecto de código abierto TM4J<sup>11</sup> y combina el editor y el visor. El objetivo del visor del *TM4L* es proporcionar una interfaz gráfica de usuario intuitiva para la navegación de contenido de aprendizaje basado en el TM. Dicheva et al. (2005) consideran que la navegación como una práctica de exploración, es un proceso de búsqueda de información muy relevante en la tarea de aprendizaje real del usuario.

Para habilitar la exploración el visor de usos múltiples de *TM4L* es compatible con múltiples vistas que ofrecen diferentes perspectivas sobre el objeto seleccionado. El visor ofrece tres vistas: vista de gráficos, vista de texto y vista de árbol. Estos puntos de vista están destinados a facilitar la navegación en los "puntos calientes", es decir, no está diseñado para transmitir la riqueza de un repositorio basado en TM, sino para mostrar qué temas están presentes y cómo se relacionan. Con el objetivo de reducir la sobrecarga de información, se ha elegido cada paso de navegación para mostrar solo los temas más relacionados con el objeto seleccionado. La vista de gráfico muestra un gráfico semánticamente expresivo (ver Fig. 13).

<sup>11</sup> <http://tm4j.org/>

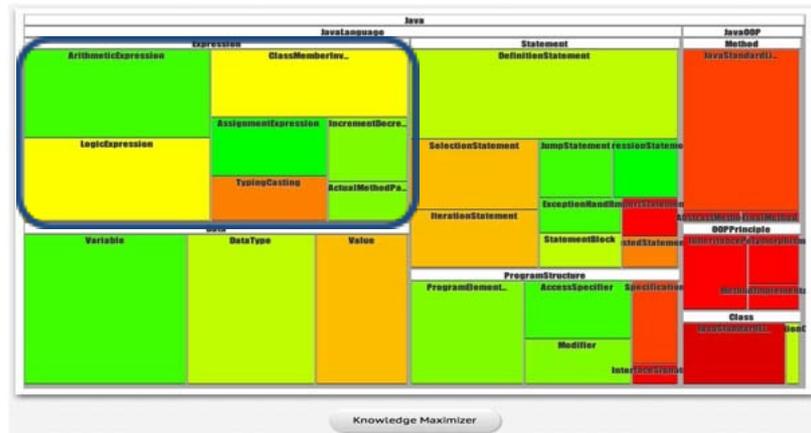


Fig. 14. Mapa del explorador de conocimiento (enfoque abierto de modelado del estudiante) y un botón para lanzar el maximizador de conocimiento (enfoque de secuenciación de problemas) (P. Brusilovsky, Baishya, Hosseini, Guerra, & Miner, 2013).

Otra utilidad de los mapas nace de la circunstancia por la que, dentro de un corto período de tiempo (normalmente una semana o menos), un estudiante tiene que revisar el contenido que fue estudiado durante todo el semestre, identificar las posibles lagunas de conocimiento y conceptos erróneos, y llenar estos vacíos. Una herramienta de aprendizaje personalizado basada en un modelo de estudiante a largo plazo podría ser de gran ayuda en este proceso. Al reflejar el progreso del estudiante durante todo el semestre, un modelo de estudiante puede distinguir los temas que se aprendieron y necesitan solo una actualización rápida de los temas que se pasaron por alto y puede ser necesario un examen a fondo.

Utilizando este modelo, P. Brusilovsky et al. (2013) presentan una herramienta personalizada de preparación para el examen que puede guiar de forma individual a cada estudiante a través del proceso de estudio con el uso de la técnica de *TreeMap*. El objetivo fue explorar los dos enfoques más populares de modelos de orientación personalizada al estudiante: modelado abierto del estudiante y la secuenciación adaptativa del problema.

La idea de modelado abierto del estudiante es para mostrar el estado de un modelo de estudiante a estudiante con el fin de ayudar a reflexionar sobre sus conocimientos, identificar las brechas, y centrarse en subsanar estas deficiencias. La idea de la secuenciación adaptativa del problema es generar una secuencia personalizada de los problemas que ayudarán al estudiante a la práctica de manera eficiente su conocimiento faltante. *KnowledgeZoom* (KZ) es una herramienta de preparación de exámenes desarrollada por P. Brusilovsky et al. (2013) que combina ambos enfoques en una única interfaz. La interfaz de KZ (ver Fig. 14 y Fig. 15) ofrece acceso directo al primer enfoque mencionado y un botón para iniciar el segundo enfoque.



Fig. 15. Esta figura es resultado de hacer *zoom* sobre el nodo *Expressions* (esquina superior izquierda como se marca en la Fig. 14). Revela el siguiente nivel de la jerarquía de dicho concepto. Ahora el usuario puede, ver el todo, que consta de varios conceptos desconocidos (coloreados en rojo) y varios bien aprendidos (coloreados en verde), que tiene conocimiento intermedio del concepto *Logic Expression* nodo que se muestra en color amarillo (P. Brusilovsky et al., 2013).

Por otra parte, bajo la influencia del paradigma constructivista, donde el intercambio de información y su retroalimentación son la base para la creación de nuevos conocimientos, resulta de utilidad dar la oportunidad a los profesores de explorar la interacción que se produce en una plataforma virtual. Romani and da Rocha (2000) desarrollaron la herramienta *InterMap*, que presenta este tipo de información cuantitativa de una manera sencilla mediante una pantalla en la que se le presentan al usuario las opciones para la formulación de una búsqueda. El usuario puede definir el tipo de presentación (gráfico o tabla); el período (día, semana o mes) y la categoría (todos los participantes, solo los estudiantes o solo los profesores).

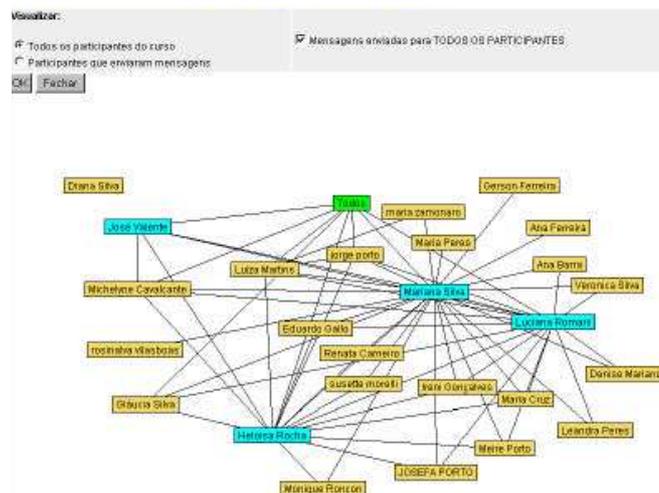


Fig. 16. Visualización de la interacción. Se presenta un gráfico que representa la interacción entre los participantes de un curso de ejemplo. Los nodos en color cian representan los maestros, los nodos en naranja indican los estudiantes y un nodo en color verde representa a todos los participantes en el curso. Por tanto, ya que es posible enviar mensajes a todos los participantes, en esta visualización se pone en evidencia las personas que han enviado mensajes a todos los participantes. Utilizando la técnica de consultas dinámicas, los usuarios pueden modificar dinámicamente la visualización a través de controles en la parte superior de la ventana (Romani & da Rocha, 2000).

En *InterMap* se usa un código de color para distinguir los profesores y los estudiantes en la herramienta. En toda la herramienta, los docentes se representan en cian y los estudiantes en naranja. Además de diagramas de barras, en *InterMap* es posible obtener información de las búsquedas realizadas mediante gráficos de dispersión que representan los datos mensuales, semanales o diariamente dependiendo de la búsqueda realizada. Finalmente, *InterMap* posibilita la representación de la interacción a través de una red. En esta los vértices (nodos) representan a las personas -es decir, los participantes en el curso- y las aristas (líneas) representan el intercambio de mensajes entre ellos. Entonces, si el participante A envía un mensaje a los participantes B, se dibuja una línea que une los dos nodos. En esta representación, los nodos sin aristas indican que estos participantes estuvieron inactivos (ver la Fig. 16).

Cuando se selecciona un nodo de la gráfica con un *click* de ratón, este nodo y todas las líneas que enlazan con este cambian de color, lo que facilita la visualización. Además, es posible modificar la presentación del gráfico seleccionando un nodo y arrastrándolo con el ratón.

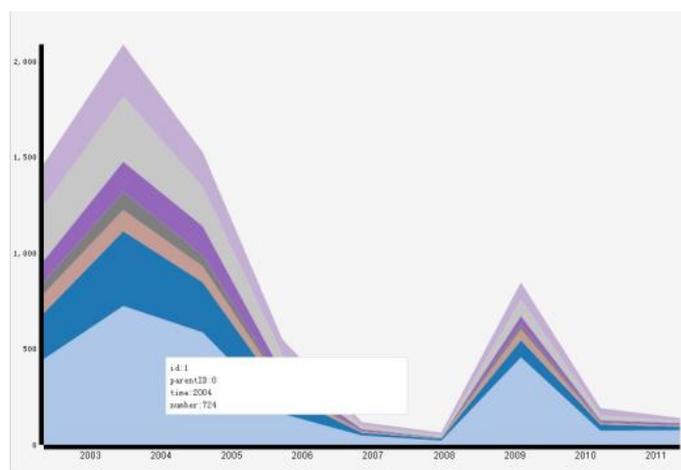


Fig. 17. Flujo de temas (Feipeng et al., 2013).

Existen herramientas de visualización de las tendencias de los temas, estos se centran en demostrar la evolución del tema con un modelo de tema apropiado o un enfoque atractivo para ilustrar la tendencia de su evolución (Chuang, Manning, & Heer, 2012; Gretarsson et al., 2012). Sin embargo, así como el modelado de la evolución temática y el efecto visual, la relación jerárquica de temas en conjuntos grandes de textos y la interacción de la interfaz de visualización es también muy importante para la visualización de la evolución de los temas. Sin embargo, la mayoría de los trabajos se ocupan de la evolución temática y del problema de interacción por separado, sin tener en cuenta las características distintivas de la relación jerárquica y la interacción en las grandes colecciones de textos. Feipeng et al. (2013) presentan una herramienta que, por medio de minería de la relación jerárquica de temas y tendencia evolutiva, ofrece tres vistas junto con funcionalidad interactiva que permite a los usuarios comprender la evolución de los temas de las discusiones de un LMS de una manera flexible y fácil (ver Fig. 17).



Fig. 18. Representación de HEODAR (Morales Morgado et al., 2009). Se divide en tres partes: 1) la representación de la media de los resultados de todas las evaluaciones realizadas de los LO (esta es la evaluación del LO); 2) el valor multiplicado por el porcentaje de profesores que han evaluado el LO con respecto al número total de profesores, lo que resulta en la evaluación ponderada y, por último, 3) el porcentaje de profesores que han evaluado el LO con respecto al número total de docentes.

Siguiendo con el uso y el análisis de la interacción en los LMS para la mejora del aprendizaje, se expone la forma de presentación de los resultados de la evaluación de la herramienta *HEODAR* (Morales Morgado, Aguilar, & Peñalvo, 2008; Morales Morgado et al., 2009; Munoz, Conde, & García Peñalvo, 2010), un mecanismo para aportar nueva información sobre la evaluación de objetos de aprendizaje (del inglés, *Learning Object*, *LO*). Esta suministra información relevante tanto para los profesores como para los estudiantes. La forma visual de la evaluación del LO, como se muestra en la Fig. 18, utiliza una calificación representada por 5 estrellas siendo la nota de 1 a 5, rellenado según el caso, ya sea por los valores o valores ponderados.

Dentro de las estrategias para la mejora del proceso de enseñanza y aprendizaje, la representación de la evaluación de los contenidos de aprendizaje no es única. Williams and Conlan (2007) presentan un prototipo de un analizador de narrativa en su trabajo, que muestra la correlación entre la actividad de un estudiante y un curso de *eLearning* adaptativo. Muestran las estructuras narrativas y el estilo de aprendizaje de los estudiantes en los sistemas de *eLearning* y también el uso de una sencilla línea de tiempo para la selección de las estructuras narrativas.

La visualización del analizador de narrativa, ver Fig. 19, se divide en cuatro áreas visualmente separadas. Estas son: el menú de configuración de los parámetros de visualización (en la parte superior), la visualización de la narrativa (en el centro a la derecha), la vista estilo de aprendizaje (en el centro a la izquierda) y el área de visualización de la línea de tiempo (en la parte de abajo).

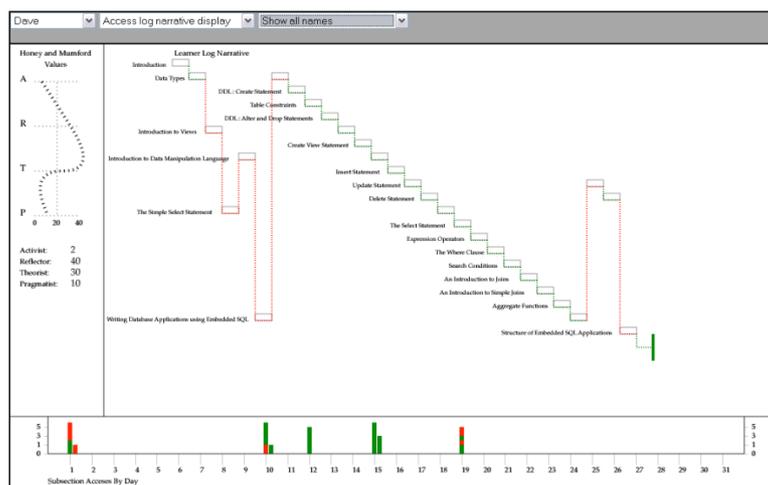


Fig. 19. Prototipo del analizador de narrativa (Williams & Conlan, 2007).

Cada área tiene propiedades únicas cuando se seleccionan diferentes opciones de narrativas o estudiante. El área de ajuste de parámetros de visualización se compone de una serie de menús desplegables que permiten la selección de los dicentes y los tipos de visualización narrativa, así como opciones para la visualización de la subsección de los nombres al mostrar el *log* de las narrativas. En el área de visualización de la narrativa se muestran las distintas permutaciones de la asignatura y del *log* de estructuras narrativas. Las narraciones se han dibujado desde la esquina superior izquierda del área de visualización y se extienden por la anchura del bloque de la subsección de la narrativa a la derecha para cada evento de narrativa del curso o del *log*. Cuando la historia de acceso lógico de un estudiante se combina en esta pantalla la escalera de desviaciones de la narración sugiere que muestra el desplazamiento y la dirección del salto respecto a la secuencia nominal de la narración. La dirección de salto puede interpretarse como la exploración o la revisión para adelante o para atrás respectivamente. Cuando se combina con la pantalla de estilo de aprendizaje los autores de cursos pueden vigilar o validar los aspectos relevantes de la evolución de los cursos. Esta característica permite una rápida asimilación de las principales características del estudiante que, cuando se combina con las pantallas narrativas pertinentes, tiene como objetivo apoyar a los estudiantes en el desarrollo de la meta cognición mediante el contraste de su estilo de aprendizaje auto declarado con la realidad de sus interacciones con el sistema de *eLearning* adaptativo.

Además de la representación visual de los contenidos de aprendizaje con objetos de aprendizaje, el estilo de narrativas de los estudiantes y el uso de técnicas para maximizar el aprendizaje a través de una serie de organizadores gráficos, no hay que olvidar que un estudiante en este tipo de educación está actuando en un entorno virtual, el cual que exige tanto la posibilidad como la necesidad de representar la dinámica de interacción del proceso de aprendizaje, que son apenas visibles en la forma orgánica de aprendizaje.

A diferencia de la educación presencial, todas las acciones o intervenciones que realizan los individuos pueden ser reutilizadas y representadas en la misma plataforma, a la vez que se registran en la base de datos. Esto es necesario (debido a que el aprendizaje en estos entornos virtuales toma prestado de otros tipos de ayuda visual, de las cuales a lo largo de este capítulo se describirán ejemplos -además de los de los objetos de aprendizaje y organizadores gráficos) para comprender las relaciones establecidas entre los individuos involucrados, tales como el reconocimiento de las acciones respectivas en la plataforma y la mejora de las estrategias de enseñanza y aprendizaje.

En otras palabras, como se mencionó en la sección III.3, existe una posibilidad técnica (todas las acciones o intervenciones que realizan los individuos pueden ser reutilizadas y representadas en la misma plataforma, a la vez que se registran en la base de datos) con una finalidad pedagógica (la eficaz transmisión de conocimiento y comprobación de este), que ofrece una serie de ayudas analíticas por medios visuales reutilizando todo lo que está sucediendo, evaluándose y aprendiéndose en un LMS para representar, comprender y analizar la forma en que se están desarrollando los procesos de aprendizaje y al reutilizar estos datos mejorar el proceso.

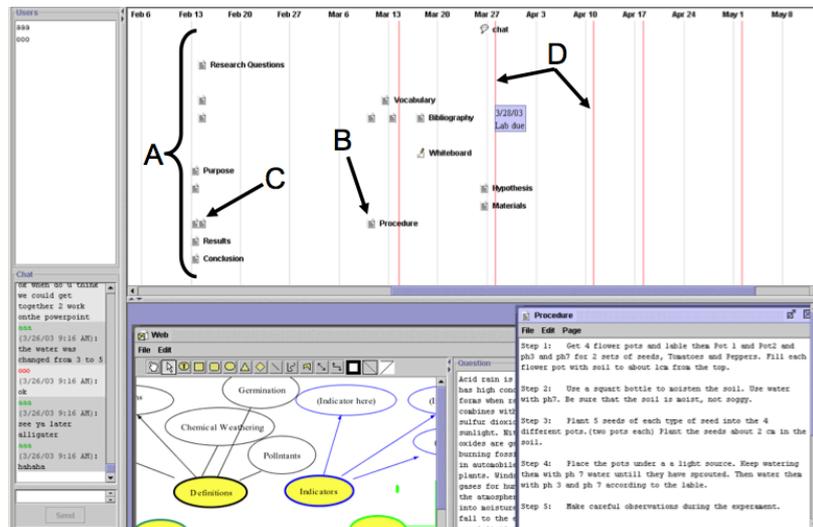


Fig. 20. Cliente de escritorio. El tiempo se representa de izquierda a derecha, desplazable desde el principio hasta el final del proyecto (Ganoë et al., 2003).

En pocas palabras, una representación visual del proceso, como un instrumento de mediación, puede mejorar la capacidad de aprendizaje de los estudiantes ya que les permite visualizar la interacción que tiene lugar durante el proceso de la educación virtual.

El trabajo aula *BRIDGE* (Ganoë et al., 2003) muestra, como se ve en la Fig. 20, proyectos distribuidos de grupos entre las aulas escolares que pueden colaborar como parte de un equipo más grande, incluyendo grupos similares de otras clases. Por medio de una sesión estándar de *chat*. Los mensajes de *chat* se muestran en una lista deslizable con el tiempo y el remitente por encima de los mensajes. La sesión de *chat* es persistente, de forma que los miembros del grupo que se pierdan una discusión puedan revisar lo que se dijo. Esta sesión es un objeto compartido al igual que los demás documentos creados y compartidos por el grupo. A medida que ingresan mensajes de *chat*, la última fecha de modificación de la sesión se actualiza en la línea de tiempo. Los iconos en la Fig. 20 representan los artefactos de trabajo; cada fila contiene la historia de un elemento individual (A), el extremo derecho es por tanto la versión actual (B), con versiones archivadas a la izquierda (C), y las líneas verticales más oscuras representan los plazos (D). Las líneas de tiempo presentan una pantalla integrada de información de la actividad. La información que se muestra se genera a través de tareas de control de versiones y de calendario automático.

A la izquierda, una barra lateral, en la parte superior, que se puede cerrar contiene una lista de usuarios, y una sesión de chat persistente en la parte inferior de esta. En el aula *BRIDGE* la lista de usuarios proporciona información acerca de la presencia de otros miembros del grupo, mostrando los nombres de usuario de cada miembro del grupo que está utilizando actualmente el cliente. Una técnica útil de este trabajo es presentar las actividades de cada grupo de forma simultánea. Esto permite una rápida comparación entre los grupos para conseguir, con ayuda de iconos que representan las diversas actividades, un sentido del progreso de las metas del proyecto.

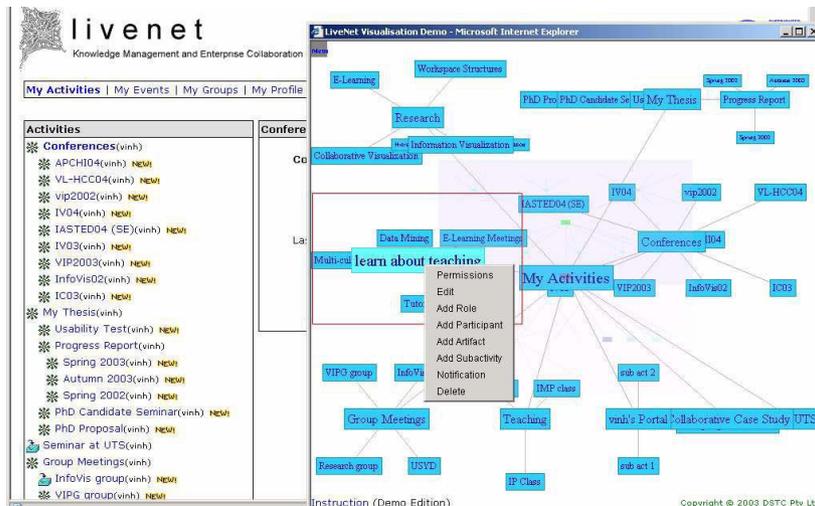


Fig. 21. Un ejemplo de un menú de operación que aparece en un nodo particular que muestra todas las acciones, propiedades y atributos asociados del objeto correspondiente. En ENCCON los nodos representan los objetos (es decir, actividades, artefactos, grupos, eventos, agentes, perfiles, etc.), mientras que los líneas se utilizan para presentar relaciones entre los objetos, en este caso se limita a las jerarquías relacionales (Q. V. Nguyen et al., 2004).

Quang Vinh et al. (2013) presentan una herramienta que reconoce automáticamente el tipo de dato multidimensional ingresado y permite la tabulación de estos en diferentes graficas de dispersión de puntos, ya sea por el eje X o el eje Y; se permite mostrar la etiqueta del dato en el diagrama de dispersión y permite personalizar la variable del eje utilizado. Además, habilita al usuario el uso de la representación de atributos por medio del color y el tamaño de los puntos.

Q. V. Nguyen et al. (2004) utilizan técnicas de visualización en apoyo al *eLearning* en ENCCON, representado en la Fig. 21. Utilizan rectángulos para la división de zonas. ENCCON se utiliza para representar la estructura del espacio de trabajo colaborativo y las relaciones lógicas entre los objetos (por ejemplo, elementos de aprendizaje). En su última versión (Q. Nguyen & Huang, 2010) presentan mejoras que consisten en una colección de múltiples algoritmos de disposición de los nodos en la red, tales como ENCCON (Q. V. Nguyen & Huang, 2005; Q. V. Nguyen et al., 2004), *Spring Embedded* (Eades, 1984), dibujo circular, y la combinación de múltiples algoritmos en un diseño de disposición de los nodos en la red (Huang & Nguyen, 2008).

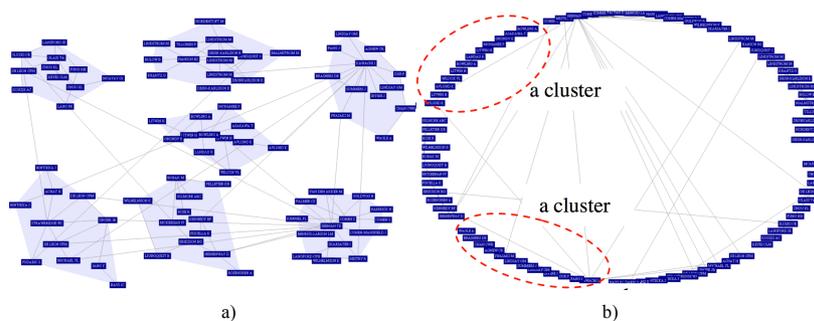


Fig. 22. Un ejemplo de dos diseños de gráficos de la misma red: a) mediante un *Spring Embedded* con particiones de ENCCON, y b) utilizando un algoritmo de diseño circular (Q. Nguyen & Huang, 2010).

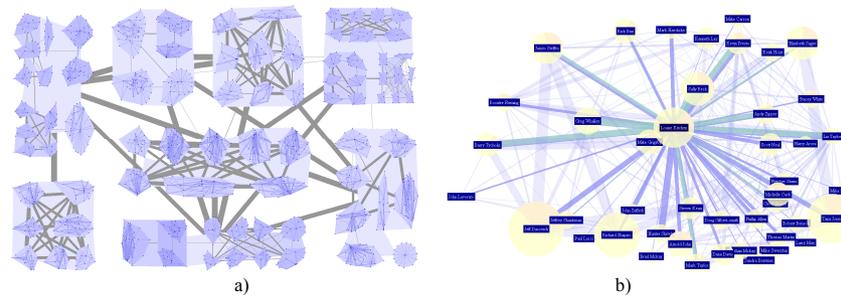
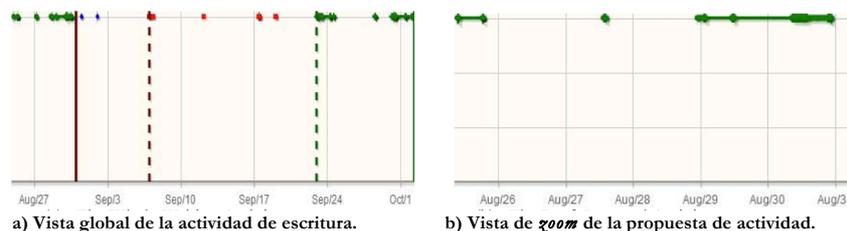


Fig. 23. Un ejemplo de las diferentes visualizaciones de redes: a) visión abstracta de un gráfico jerárquico y b) vista de detalle asociada a un elemento (Q. Nguyen & Huang, 2010).

Estos algoritmos se emplean de forma automática o se seleccionan manualmente por el usuario para proporcionar el diseño óptimo para la gran red. Los procesos de refinación, incluyendo el ajuste de maquetación y eliminación de superposición de los nodos, también se aplican sobre la pantalla para mejorar la claridad, la amabilidad estética y propiedades del Agrupaciones de nodos. En cualquier momento, el usuario puede cambiar entre diferentes diseños y viceversa para obtener las visualizaciones óptimas y/o preferibles. La Fig. 22 muestra un diseño multi-escala y una visualización utilizando los algoritmos *Enccon* y *Spring Embedded*. La Fig. 23 muestra el diseño de la misma red utilizando diferentes algoritmos en el mismo sistema.

Los entornos virtuales de aprendizaje, como espacios educativos, como se comentó en las secciones II.1 y II.3, ofrecen una serie de posibilidades de interacción social, profesor-estudiante o entre los estudiantes, a través de diferentes canales síncronos o asíncronos. También ofrecen posibilidades de interacción con una serie de diferentes materiales educativos, pero no dan una imagen o representación general de lo que sucede en estos espacios, como bien se ha dicho en la sección I.2.2. Para cualquier persona que opera en estos espacios educativos, es de crucial importancia ubicar su acción en el ámbito educativo y reconocer todas las interacciones que tienen lugar en la formación virtual. Siguiendo esta idea Ming et al. (2013) presentan un sistema de LA llamado *Tracer* que recopila eventos desde una herramienta de escritura.



a) Vista global de la actividad de escritura.

b) Vista de zoom de la propuesta de actividad.

Fig. 24. Visualización del modelo basado en la línea. Cada fila representa la actividad de escritura de un usuario. Los puntos están conectados por líneas y el grosor de una línea indica la intensidad de la conducta del usuario durante un período de tiempo. Esta figura muestra un ejemplo de gráficos generados usando esta visualización: a) muestra toda la actividad, donde los puntos están conectados a las líneas en cada sub-actividad (se definió cada línea como una serie, que tiene diferente color y grosor); b) es una figura complementaria (zoom sobre la figura 2a), lo que da más detalles acerca de la conducta del estudiante sobre la actividad propuesta del proyecto de escritura (Ming et al., 2013).

Mientras que los estudiantes trabajan en su tarea de escritura ofrece visualizaciones para mostrar el nivel de compromiso del usuario con relación a sus tareas, además de incluir medidores de tiempo de compromiso. El sistema tiene dos componentes principales. El primero de ellos es el módulo de recopilación de datos, que en la versión estudiada se basa en la API de Google Doc<sup>12</sup> y la herramienta iWrite de escritura (Calvo, O'Rourke, Jones, Yacef, & Reimann, 2011). iWrite es un sistema de gestión de asignaciones basada en web, que proporciona una plataforma para que los estudiantes escriban y presenten su trabajo escrito en un documento de Google Docs. *Tracer* utiliza la API de Google Doc para acceder a la historia de las revisiones de los documentos gestionados por iWrite. Estas dos herramientas se combinan en *Tracer* para registrar la actividad y proporcionar información al autor acerca de la asignación. Su contribución se centra en cómo obtener una medida del nivel de compromiso del escritor durante la asignación y mostrar esta información a profesores y estudiantes. En este trabajo se incluyen dos modelos de visualización derivados de los datos recogidos: Modelo Visualización basada en Línea (LbVM) y Modelo de Visualización basado en Puntos (PbVM). Ambos modelos muestran una puntuación de compromiso basada en el tiempo que el usuario pasó en la actividad (ver Fig. 24). Las interacciones durante el curso se consideran un nivel de compromiso, por ejemplo los debates o discusiones, y en estas cada mensaje tiene un remitente, la fecha de envío y tema. Dentro de un debate, se llama hilo de discusión a un comentario inicial y todas sus respuestas, un conjunto de hilos sobre un tema forman una discusión. A la persona que envía el mensaje inicial en un hilo se la denomina "iniciador".

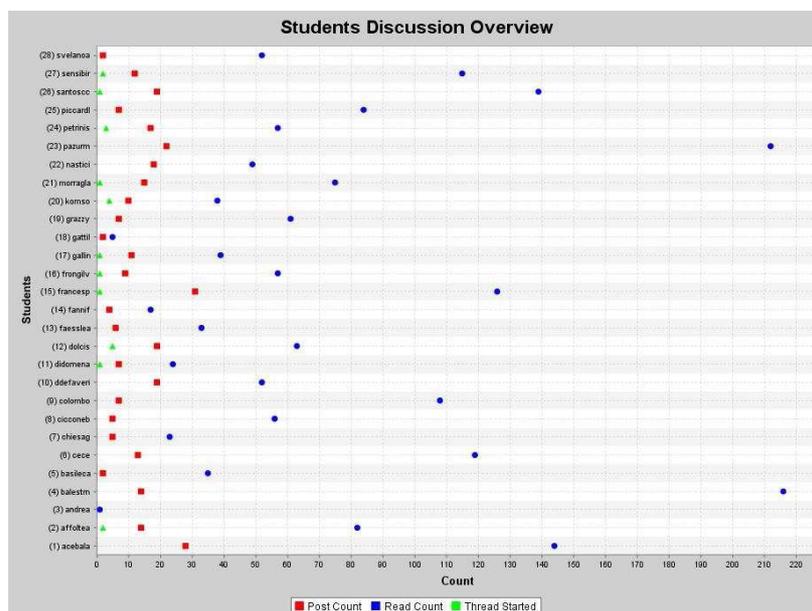


Fig. 25. Representación de las discusiones de la herramienta de (Mazza & Milani, 2004).

<sup>12</sup> <https://developers.google.com/google-apps/documents-list>

Mazza and Milani (2004) muestran el instante en el que los usuarios entran en la plataforma, los accesos de los estudiantes a un recurso particular y una representación de la frecuencia de la lectura y la escritura en los foros a lo largo del curso, tanto de los usuarios así como del iniciador (ver Fig. 25). Aquí los profesores tienen una visión general de todas las discusiones llevadas a cabo por los estudiantes. Para cada estudiante del curso que da indicios del número de mensajes enviados (con un cuadrado), el número de mensajes de lectura (con un círculo) y, finalmente, el número de las discusiones iniciadas por el estudiante en las discusiones (con un triángulo).

En (Donath, 2002) se analizan tres proyectos: la *Coterie*, *PeopleGarden*, y el *Proyecto Loom*. El primero de ellos destaca dos elementos clave: la actividad de los participantes y la estructura de la conversación. En segundo lugar, *PeopleGarden* utiliza la metáfora del jardín de flores para visualizar la participación en un foro. Por último, el objetivo del proyecto *Loom* es crear visualizaciones basadas semánticamente en los grupos de noticias de Usenet (Boyd, Hyun-Yeul, Ramage, & Donath, 2002). Existe un amplio conjunto de estadísticas que se pueden derivar de una base de datos de publicaciones de Usenet.

La representación de estas estadísticas en clasificaciones socialmente significativas es la esencia de la parte analítica de la visualización semántica. Parte del proceso inicial es identificar las categorías de interés y definir las en términos de características reconocibles. Por ejemplo, es posible que se desee representar a los líderes de un grupo o la identificación de personas que publican con frecuencia y a menudo obtienen respuestas de forma positiva. Esto distingue a estos usuarios de otros reseñistas frecuentes como *trolls* (alborotadores deliberados), fuentes de noticias automáticas, etc., y los distingue también de los comentarios frecuentes que a menudo se han contestado negativamente -estos se pueden etiquetar inicialmente como "provocadores".

Siguiendo esta idea de análisis y seguimiento de los mensajes y discusiones, en un LMS se almacenan las visitas y mensajes a través del tiempo para cada persona. Sin embargo, dada la naturaleza asíncrona de las actividades de aprendizaje en los entornos electrónicos, en la mayoría de los casos, no existe la posibilidad de proporcionar a las personas involucradas una visión completa de las actividades, el progreso y los patrones de uso de la plataforma. Hardless and Nulden (1999) tratan de ayudar a los tutores en estas circunstancias mediante la presentación de las visitas y los comentarios en los foros en una línea de tiempo.

Una herramienta con una interfaz muy semejante a la mencionada de Hardless and Nulden (1999), llamada *CourseVis*, se presenta en (Mazza & Dimitrova, 2005) genera varias representaciones gráficas en 3D y 2D, que se clasifican de acuerdo a los aspectos sociales, cognitivos y conductuales. Su objetivo es ayudar a los instructores para formar un modelo mental de sus clases con el fin de poder ofrecer a los estudiantes la ayuda adecuada; esto lo hace a través de una representación basada en un gráfico de dispersión de los debates en línea y una matriz para visualizar el desempeño de los estudiantes en las pruebas relacionado con los conceptos de dominio.

## Summary of student's behaviours

from 2002-01-15 to 2002-04-11  
Student: "francesco"

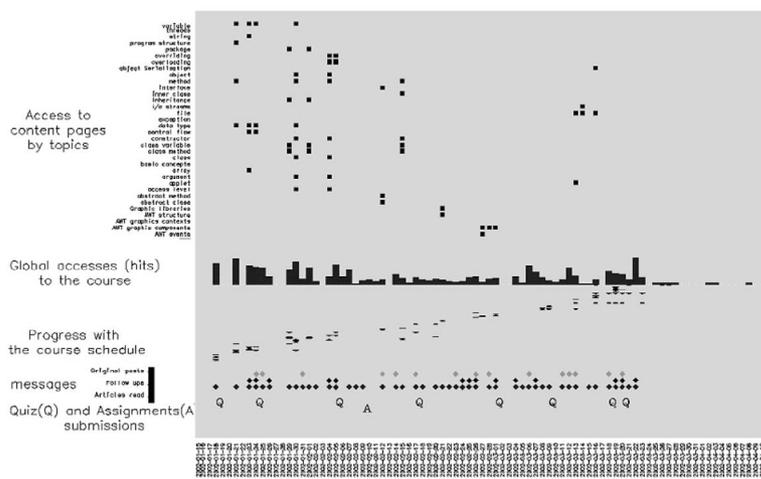


Fig. 26. El gráfico de Comportamiento representa el comportamiento del estudiante (Mazza & Dimitrova, 2005).

Ilustrado en la Fig. 26, el estudio de Mazza and Dimitrova (2005) representa la información respecto a un estudiante en particular. Se toma ventaja del método de composición de un solo eje (Mackinlay, 1986) para presentar un gran número de variables en un espacio métrico 2D. Con un eje X común, se cartografían las fechas del curso y se presenta una serie de variables: los accesos a páginas de contenido por temas, los accesos totales al curso, el progreso con el programa del curso, los mensajes y los envíos de exámenes y tareas.

Del mismo modo, *CrystalChat* (Tat & Carpendale, 2006), mediante una representación visual en 3D, conmutable a 2D, revela los patrones de comunicación de un individuo con las personas que forman parte de su historial de *chat* personal. En esta herramienta se pueden descubrir patrones de frecuencia de mensajes, temporalidad de estos, frecuencia de los participantes, así como quien inicia la conversación y patrones en el uso de los emoticonos. Este programa no solo permite a los usuarios descubrir sus propios patrones personales, sino que también les permite reflexionar sobre sí mismos para cambiar el propio comportamiento social en el futuro. Como se muestra en la Fig. 27, cada faceta se compone de una serie de conversaciones que se construye como sigue: un disco de color representa un solo mensaje y cada color indica una persona diferente, una lectura de izquierda a derecha en forma secuencial se utiliza para entender el orden de estos mensajes (es decir, los mensajes están dispuestos de modo que el primer mensaje en la conversación comienza a la izquierda), una conversación se define, como en los datos, por la apertura y cierre de la ventana de chat, donde una fila de discos de mensaje representa una conversación.

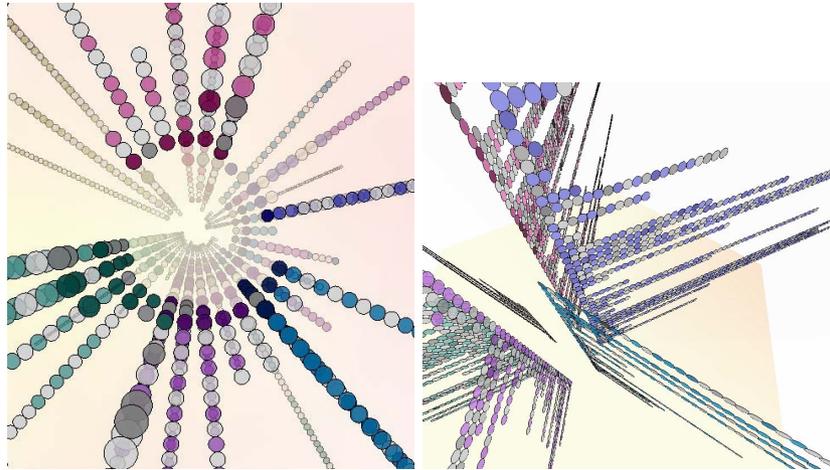


Fig. 27. A la izquierda la vista desde arriba de *CrystalChat* y a la derecha, una vista mirando hacia abajo en la estructura de las conversaciones en *CrystalChat* (Tat & Carpendale, 2006).

En este mismo sentido, Gibbs et al. (2006) presentan el *software* de mapeo de las relaciones temporales de los debates (del inglés, *Mapping Temporal Relations of Discussions Software*, MTRDS) destinado a ayudar a analizar los aspectos temporales de las discusiones del curso en línea, de forma que se resalta la información importante relacionada con los patrones de interacción. MTRDS es un prototipo web que representa dinámicamente los aspectos temporales de discusiones.

Cualquier persona que utilice la plataforma *WebCT* puede codificar los mensajes de los foros de discusión en un archivo de texto y luego cargar el archivo en el MTRDS utilizando un navegador web. Un mensaje que origina un hilo de conversación (es decir, el mensaje iniciador) es aquel con el que se crea una cadena de respuestas a este mensaje iniciador; cuando se responde, esta respuesta tiene también uno o más mensajes como hijos (respuestas a este último).

Una línea con una única flecha que apunta hacia atrás representa el vínculo entre un nodo de respuesta (hijo) y su mensaje de origen (el padre). A partir de los mapas se puede determinar de forma sencilla la ascendencia de un nodo de respuesta particular (nodo hijo). MTRDS dibuja flechas que indican que se ha respondido el mensaje o se sigue por un mensaje específico, por tanto, las flechas apuntan hacia atrás. Es importante señalar que este tipo de mapa ilustra el patrón de respuesta. Por otro lado, si se está interesado en el flujo de la discusión del tema, puede tener más sentido que las flechas señalen hacia delante. Sin embargo, esto depende de la finalidad del instructor/investigador para utilizar el mapa. Además, MTRDS genera una representación visual de los debates basados en hora y fecha (ver la Fig. 28). MTRDS colorea los mensajes por el tema, porque las discusiones a menudo contienen varios temas/hilos de discusión. El código de colores también revela cadenas de mensajes aislados y aquellos que son dominantes. Cuando pasa el usuario el ratón sobre un nodo de mensajes, los aspectos más destacados de nodo y el tiempo de *post* aparece en la columna de la izquierda. Si el usuario hace *click* en un nodo de mensajes, la vista se expande y presenta un cuadro de desplazamiento que contiene el mensaje, el nombre del autor y el tiempo.

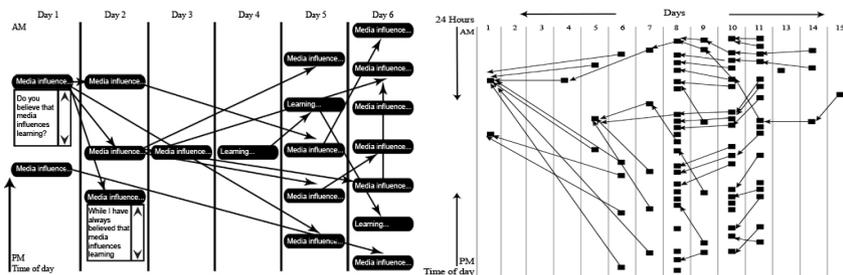


Fig. 28. Presentación del mapa de discusiones de MTRDS, los ejes X e Y denotan el día de la fecha y hora, respectivamente. Cada mensaje de discusión (*post*) está representado por un círculo con un código de color (nodo del mensaje) y una línea (enlace) que conecta un nodo respuesta al mensaje originario (Gibbs et al., 2006).

En el trabajo anterior se crean estructuras que relacionan mensajes formando redes de mensajes entre participantes, se logra así estudiar los patrones en las discusiones. El objetivo es utilizar estos métodos para estudiar la naturaleza de los patrones de interacción dentro de una comunidad de aprendizaje en red y la forma en que sus miembros comparten y construyen conocimiento.

Este objetivo ha dado lugar a una tendencia a usar representaciones visuales de todo lo que sucede en los entornos virtuales de aprendizaje, que es el mismo proceso, como una herramienta tanto para cualquier persona que quiera entender y actuar dentro de este espacio (ya sea cuando están aprendiendo de forma individual o cooperativamente), y para cualquier persona que quiera entender y/o ayudar como un tutor, así como para aquellos que quieren mejorar operativamente y enriquecer su enseñanza en un LMS. Dentro de esta tendencia se puede destacar, como se mencionó en la sección III.5, el Análisis de Redes Sociales (del inglés, Social Network Analysis, SNA); se trata de una alternativa útil para el análisis de las estructuras sociales que surgen de las relaciones entre los diferentes actores sociales, que se utiliza a menudo en la psicología social, la antropología, la comunicación y la política (Ferguson & Buckingham Shum, 2012a).

El SNA, como un conjunto de técnicas de análisis para el estudio formal de las relaciones entre los actores y para el análisis de las estructuras sociales que surgen de estas relaciones, se basa en la teoría de grafos para representar el comportamiento de las redes. Los nodos de las redes pueden representar personas, organizaciones, eventos o naciones, mientras que los enlaces se corresponderían con las direcciones, las frecuencias y puntos fuertes de las relaciones entre los nodos (Scott & Carrington, 2011). Estas técnicas de visualización también se están comenzando a validar en los cursos en línea y hay una tendencia que muestra que las métricas SNA y las visualizaciones de las interacciones son herramientas potencialmente efectivas para analizar patrones de interacción en entornos virtuales, por ejemplo en foros virtuales asincrónicos en la educación virtual (Willing, 2008). Es decir, el SNA se centra en el reconocimiento de los patrones y relaciones de la estructura social que se tejen en una relación social.

Dentro de esta línea de investigación, De Laat et al. (2007) representan cómo de equilibrada es la participación dentro de la comunidad de aprendizaje. Estos autores muestran (Fig. 29) y discuten el patrón participativo que cambia con el tiempo.

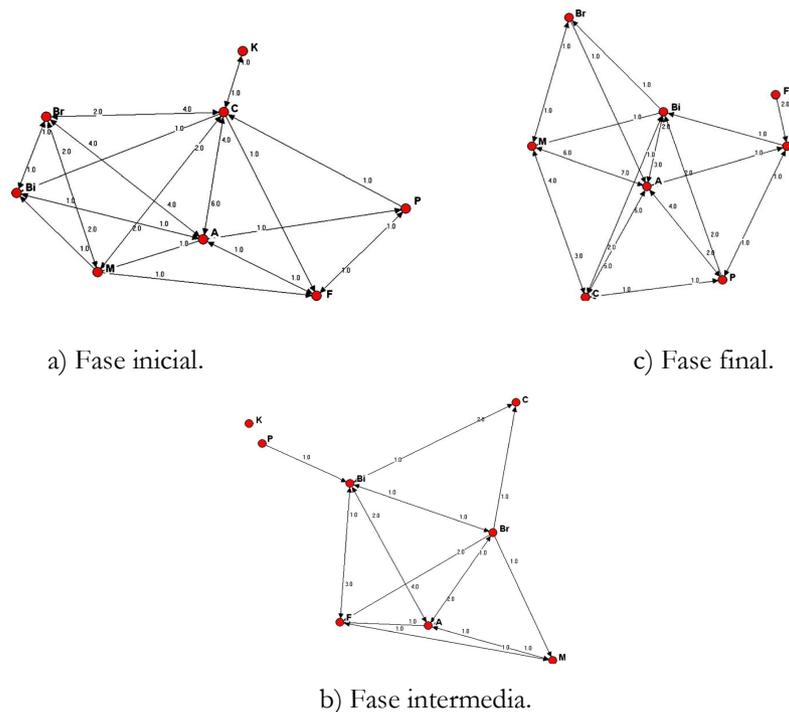


Fig. 29. Patrones de interacción entre los ocho participantes en las diferentes fases de una tarea de aprendizaje (De Laat et al., 2007).

Para comprender la compleja dinámica de la Red de Aprendizaje/Aprendizaje Colaborativo Apoyado por Ordenador (del inglés, *Networked Learning/Computer-Supported Collaborative Learning, NL/CSCL*), es importante la utilización de métodos que ayudan a ver los patrones de la interacción entre los participantes y sus dinámicas temporales. De Laat et al. (2007) representan los patrones de interacción en su caso de estudio de una NL/CSCL. Esta acción formativa fue dividida en tres fases (a, b y c en la Fig. 29) de diez semanas de duración con el objetivo de mostrar los patrones sociales y su evolución, se dividió el período de 10 semanas en tres secciones: inicio, medio y fin. A partir de cada período, se tomó una muestra de mensajes de un plazo de 10 días para formar el conjunto de datos. En este estudio se encontró que la densidad de los grupos era bastante estable, esto significa que los niveles de conectividad y de participación en esta comunidad estaban relativamente repartidos con igualdad. Estos son los resultados positivos en cuanto a la cohesión del grupo.

Sin embargo, cuando se combinaron estos resultados con los resultados del análisis de contenido, se hizo evidente que, aunque la posición de los participantes en la red sigue siendo la misma, la naturaleza o el enfoque de sus contribuciones cambiaron con el tiempo. Esto sugiere que los participantes desarrollan diferentes roles o intereses durante su trabajo de colaboración. Por tanto, es importante cuando se estudia NL/CSCL centrarse no solo en los patrones globales de participación, colaboración y construcción del conocimiento, sino también tener en cuenta la evolución de estos.

Otro aspecto a tomar en cuenta dentro del estudio de SLA es el de Schreurs, Teplovs, Ferguson, de Laat, and Buckingham Shum (2013), donde proponen que una SLA basada en perspectivas que diseñan modelos y visualizaciones capaces de mostrar las métricas habituales de SNA (como la densidad, granularidad, grados de entrada de los nodos, grados de salida de los nodos, etc.), extiendan sus capacidades a expresar los tipos de vínculo sociales entre actores (es decir relación laboral, parental, afectiva, marital, etc.) y las subredes que se crean con estos y con temas específicos que se desarrollan (subred laboral, subred familiar, etc.).

Se describe una implementación técnica que demuestra este enfoque, que extiende la herramienta *Network Awareness Tool* (NAT) por medio de la repoblación automática con los datos de una plataforma de aprendizaje social *SocialLearn*. El resultado es la capacidad de visualizar y filtrar visualmente las redes mediante la identificación de las relaciones entre las personas que interactúan en torno a los mismos temas (ver Fig. 30). Además, con el fin de representar y diseñar las interfaces de las estructuras sociales en gráficos de red, hay que tener en cuenta que debe ser ligero y del lado del cliente para evitar que el servidor se sobrecargue. Por otra parte, esta tecnología debe poder integrarse fácilmente en otras plataformas, como por ejemplo Moodle. Ya existen muchas tecnologías capaces de representar gráficos de red (*Graphviz* - *Graph Visualization Software*<sup>13</sup>). Sin embargo, la mayoría de ellos se basan en formato DOT<sup>14</sup> para la definición de la gráfica y, por tanto, requerirían de una conversión previa de los datos recogidos de un foro en un archivo de este tipo, como por ejemplo *Canviz*<sup>15</sup> y *Grappa*<sup>16</sup>.

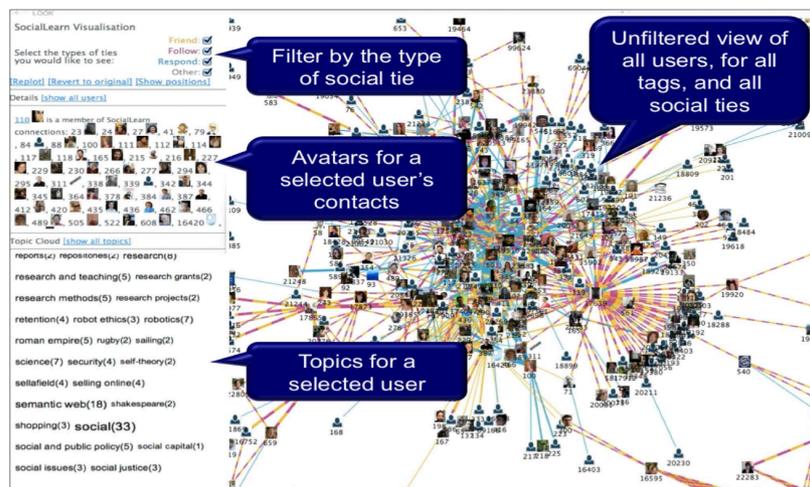


Fig. 30. El uso de NAT para visualizar y filtrar los lazos sociales de la persona, el tipo de lazo, y tema. Debido a limitaciones de espacio, se trata de una imagen compuesta que muestra toda la red sin filtrar, pero cuando se selecciona un usuario de esta, se filtra el panel izquierdo, como se muestra, para mostrar solo su ego-red y temas (Schreurs et al., 2013).

13 <http://www.graphviz.org>

14 <http://www.phil.uu.nl/~js/graphviz/dotguide.pdf>

15 <https://code.google.com/p/canviz/>

16 <http://www2.research.att.com/~john/Grappa/>

Este tipo de herramientas solo muestran el archivo en una imagen, de forma que la interactividad es nula. Otra tecnología disponible es *ZGRViewer*<sup>17</sup>, que es capaz de ofrecer una representación interactiva de un archivo DOT. Esta tecnología, sin embargo, no es capaz de arrastrar los nodos (por tanto, el reposicionamiento no es posible) y se convierte en muy confuso para utilizar en los gráficos de tamaño medio. Para el LMS Moodle, el sistema IGRAPH (Silva & Figueira, 2012) está listo para utilizarse fuera del LMS de donde se obtienen los datos y se acopla con un simple análisis de contenido para rechazar los mensajes irrelevantes y recopilar métricas de red social. En este sistema se utiliza como base la librería *Dracula*<sup>18</sup> para el módulo de dibujo, y fue alterada de acuerdo a las necesidades del sistema, además de ser capaz de analizar la red en tiempo real. Silva and Figueira (2012) representan las interacciones como se ilustra en la Fig. 31. En esta se puede observar tres tipos diferentes de representación, una, el inciso a), que ilustra una red centralizada, otra, el inciso b), una red creada a partir del análisis de los *click* realizados en la plataforma y, por último, inciso c), una red que representa las interacciones entre personas que utilizaron el foro de la plataforma.

Otros trabajos que estudian la participación a través del tiempo y toman en cuenta el contenido son los debidos a Govaerts *et al.* (2010) (2012) con su herramienta Medidor de Actividad del Estudiante. Además es importante destacar que nunca con anterioridad se habían utilizado técnicas de coordenadas paralelas en los campos de LA y AA.

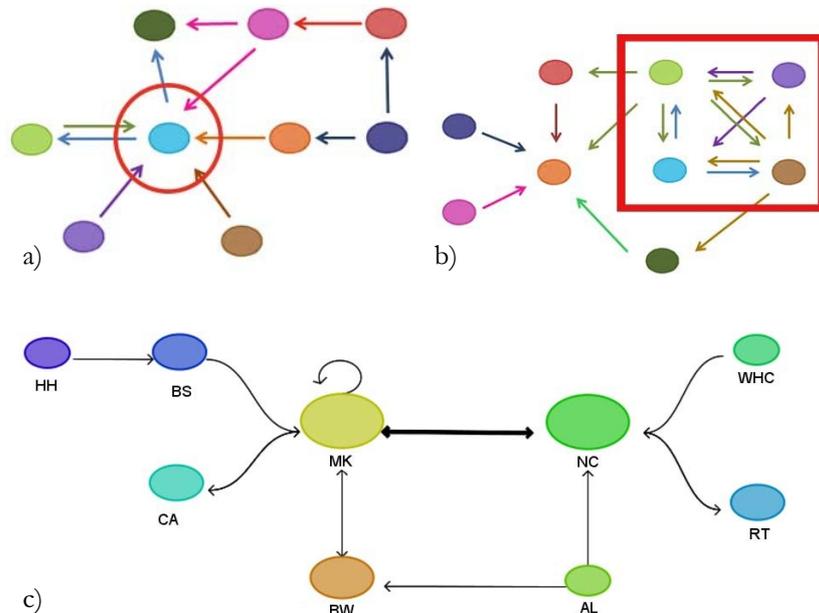


Fig. 31. Ejemplo de la representación de (Silva & Figueira, 2012). a) Una red centralizada; b) Detección de *click*; c) Gráfico del foro “Moodle creates success”.

<sup>17</sup> <http://zvtm.sourceforge.net/zgrviewer.html>.

<sup>18</sup> <http://www.graphdracula.net>.

En esta herramienta se visualizan las acciones de los estudiantes. Se presentan cuatro iteraciones de diseño y los resultados de los dos estudios de evaluación, cuantitativo y cualitativo, en entorno a datos reales que evalúan la usabilidad, el uso y la utilidad de las diferentes visualizaciones.

En la Fig. 32 el gráfico de líneas (etiqueta 4) permite la búsqueda de tendencias, cada línea representa un estudiante. Las líneas se conforman por medio de la conexión de todas las marcas de tiempo en el cual el estudiante estaba trabajando, más que conectándose, realizando alguna acción dentro de la plataforma, además de la cantidad acumulada de tiempo invertido en esta. La inclinación de la línea muestra el esfuerzo del estudiante: una línea escarpada significa un período de trabajo intensivo; una línea plana muestra inactividad. Por ejemplo, el estudiante seleccionado (en rojo) trabajó intensamente al inicio, a continuación, se estancó un poco y dejó de interactuar a finales de septiembre. Debido a que la información en pantalla puede llegar a ser muy densa, se puede ampliar por fecha (etiqueta 7), *zoom* vertical (8), búsqueda (9) y filtrado (10). El gráfico permite comparar estudiantes y saber en qué actividades o recursos han trabajado a través del tiempo.

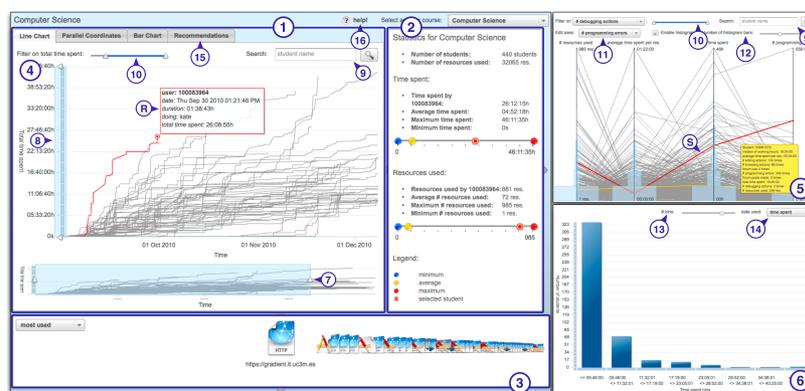


Fig. 32. La interfaz de usuario con tres visualizaciones (Govaerts et al., 2012).

Las coordenadas paralelas (etiqueta 5) son una forma común de visualizar datos de elevada dimensionalidad en el espacio limitado de una pantalla y permitir la exploración de las tendencias. Los ejes verticales representan diferentes métricas. Las métricas difieren sobre la base de los datos disponibles. El gráfico de barras (etiqueta 6) muestra la distribución de los estudiantes en un histograma para el tiempo total gastado y los recursos utilizados. La métrica representada se puede cambiar (14), mostrando una impresión general visual de la métrica y de los estudiantes. El panel de recomendación (etiqueta 3) permite la navegación a través de los recursos utilizados, se proporcionan diferentes iconos en función de tipo de recurso.

La pestaña "Recomendaciones" (etiqueta 15) contiene una nube de etiquetas animada de las recomendaciones. La mayor contribución del trabajo de (Govaerts et al., 2012) radica en la presentación de ideas sobre cómo y por qué los profesores y los estudiantes utilizan visualizaciones para mejorar la obtención del conocimiento y la auto-reflexión en los escenarios de la vida real en el proceso de aprendizaje.

También permite el descubrimiento de los promedios y tendencias de las actividades, además de los indicadores de uso de los recursos del curso, que proporcionan información sobre el material utilizado por otros estudiantes a la hora de aprender.

Una vez descritos los trabajos que apoyan al *eLearning*, además existen trabajos que no necesariamente son visuales y que apoyan a la educación, aportando métricas o nuevos procesos para el análisis de la información educativa. En la siguiente sección se presentan algunos trabajos que apoyan el aprendizaje electrónico y que no utilizan una representación visual para mostrar sus resultados.

### **V.1.2 Contribuciones sin representación visual que apoyan al *eLearning***

Como se mencionó en las secciones III.2 y III.3, existen algoritmos, métricas y procesos que contribuyen al análisis de la información procedente de los LMS y que son de utilidad para el entendimiento y/o mejoramiento del aprendizaje. Por ejemplo, la participación de los estudiantes en términos de cantidad de mensajes publicados en un foro parece estar influenciada por la calidad de las infraestructuras puestas a disposición de los estudiantes (Cacciamani & Luciani, 2007; Cesareni & Martini, 2005). Sin embargo, un elevado número de contribuciones de los estudiantes no siempre se correlaciona con su capacidad de aprendizaje. De hecho este es el caso que describen Hwang and Arbaugh (2009), que encontraron que los estudiantes que escriben menos mensajes pero participan de la discusión con respecto a una serie más amplia de temas, logran mejores resultados en la prueba de conocimiento final, en comparación con los que escriben más, pero están limitados a un solo tema de discusión.

Con el fin de evaluar el grado de participación en los cursos en línea, se han adoptado diversos enfoques metodológicos. En cuanto a los métodos cuantitativos, el recuento de las visitas (Davies & Graff, 2005) y del número de publicaciones de mensajes (Lipponen et al., 2003; Weisskirch & Milburn, 2003) se han utilizado a menudo (Kim, 2013).

Otro trabajo es el presentado por Macfadyen and Dawson (2010), en el que se realiza un estudio exploratorio en el cual se extrajeron y se analizaron variables relevantes de seguimiento de los estudiantes procedentes de tres trimestres (cinco clases) de un curso de biología de pregrado, impartido totalmente en línea, ofrecido en la Universidad de Columbia Británica en 2008. Sus hallazgos indicaron que un modelo de regresión de éxito de los estudiantes, desarrollado con variables de seguimiento relevantes a las intenciones de los instructores y de diseño de sitios web de cursos en línea (variables relevantes como por ejemplo: herramientas implementadas para permitir la entrega de contenido, y/o participación de los estudiantes, y/o la evaluación y calificación, y/o administración, etc.), combinada con medidas de tiempo en cada tarea (variables que indican el número de inicios de sesión y el tiempo pasado en línea) explica más del 30% de la variación en la calificación final del estudiante.

Para el curso utilizado en el modelo de regresión se encontró que las tres medidas de la función de actividad en línea del estudiante (número de mensajes en los foros, mensajes de correo electrónico enviados y evaluaciones completadas) como variables significativamente predictivas de la calificación final del estudiante.

Davies and Graff (2005) en su análisis reunieron y utilizaron el acceso de los estudiantes a la zona de grupo y su acceso a las áreas de comunicación, para representar el grado de participación en la discusión en línea. El acceso a la zona principal de contenidos y el acceso a la zona de estudiantes también se combinaron y se utilizaron para representar la actividad “*blackboard*” sin participar en discusiones de grupo. Estos investigadores concluyen que los efectos beneficiosos reportados de la participación y la interacción en línea no se traducen necesariamente en notas más altas al final del año. Sin embargo, los estudiantes que suspendieron en uno o más módulos, interactuaron con menos frecuencia que los estudiantes que lograron calificaciones de superación de la asignatura.

En este estudio, la participación en foros de discusión en línea tiene un doble propósito: mejorar el aprendizaje y proporcionar apoyo. Se puede, por tanto, según Davies and Graff (2005), dar el caso de que los factores tales como la frecuencia de las interacciones tienden a ser más importante en el suministro de apoyo, mientras que la calidad y la dinámica de las interacciones son los factores que influyen más en el aprendizaje y el rendimiento.

Algunos estudios como el de Romero-Zaldivar, Pardo, Burgos, and Delgado Kloos (2012) utilizan el número de operaciones de actualización de eventos y el número de operaciones de actualización de archivos para investigar el rendimiento de los estudiantes. Esto es, cada operación de actualización está en realidad compuesta por dos sub-operaciones: actualizar los archivos solicitados por el usuario, y subir los archivos que contienen los eventos registrados. Este trabajo obtuvo como resultado que de los siete tipos de eventos registrados, tres de ellos tenían una correlación significativa con la calificación final del curso. Se realizó un análisis de regresión múltiple y el modelo resultante considera solo dos de estos factores.

El modelo obtenido representa más del 20% de la variabilidad de la nota final. Por tanto, como segunda conclusión obtuvieron que la observación detallada de las actividades de los estudiantes en el espacio de trabajo del curso ofrece un marco fiable para predecir su rendimiento académico. Por su parte, Che-Cheng and Chiung-Hui (2013) tienen como propósito en su estudio identificar qué variables de seguimiento de curso se correlacionan significativamente con el rendimiento académico en los cursos asíncronos mixtos en línea a través de un análisis empírico de datos de un LMS. En este estudio, las variables de seguimiento se refieren al número de sesiones en línea, número de comentarios originales creados, número de mensajes de respuesta creados, número de páginas de contenido visualizadas y el número de comentarios leídos.

En este trabajo el rendimiento académico se define con la nota/calificación final del estudiante. Utilizaron un análisis de regresión múltiple para evaluar el funcionamiento de las variables de seguimiento de curso en la predicción del rendimiento académico. Los resultados indicaron que aproximadamente el 16,4% de la variabilidad en el rendimiento académico se contabilizó para las variables de seguimiento de los estudiantes del curso, y tres (sesiones en línea, mensajes de respuesta creados, y comentarios leídos) de las cinco variables fueron estadísticamente significativas.

Medeiros, Gomes, Amorim, and Medeiros (2013) investigaron y desarrollaron una arquitectura para el seguimiento de las interacciones sociales llamada Amadeus-SIMA, integrada en el LMS de Amadeus. Este trabajo validó dos hipótesis a través de un experimento empírico: mejorar el conocimiento de los maestros de la presencia social del estudiante y predecir el rendimiento de los estudiantes sobre la base de los comportamientos sociales generados por Amadeus-SIMA. Estos comportamientos de análisis relacionados con el grupo son la cohesión y la heterogeneidad, y en relación con cada uno de los estudiantes, se tomaron el grado de interactividad, el grado de aislamiento, el grado de intermediación, el prestigio y el grado de compromiso. En este trabajo se encontró que el grado de interactividad y compromiso resultaron ser las variables predictivas.

Por su parte, Giovannella, Scaccia, and Popescu (2013) investigan dos posibles indicadores de resultados de aprendizaje: la participación activa de los estudiantes con las herramientas 2.0 y estilos de aprendizaje de los estudiantes (del inglés *Learning Styles*, LS). Mediante el uso de análisis de componentes principales (del inglés *Principal Component Analysis*, PCA) han sido capaces de identificar un subconjunto de actividades de los estudiantes que son relevantes para construir un espacio ortogonal de representación y las actividades que pueden apoyar la predicción de éxito del estudiante (medido como calificaciones finales). Para investigar la relación entre el rendimiento de los estudiantes, su implicación con las herramientas web 2.0 y sus LS, se recogieron tres tipos de datos: i) calificaciones finales de los estudiantes; ii) la actividad de los estudiantes en el *blog*, *wiki*, *Delicious* y *Twitter*, y iii) LS de los estudiantes. Como resultado obtuvieron que la técnica de PCA es muy eficaz en la identificación de los componentes principales para el análisis educativo, dado a los resultados obtenidos y que debe vigilarse para predecir el éxito de los estudiantes.

Las acciones que, en este trabajo, contribuyen más son aquellas que involucran la producción de contenidos (es decir, acciones que implican un proceso de apropiación anterior) y su organización (es decir, acciones que también implican la conciencia sobre el contenido). Proporcionar información a los contenidos producidos por los compañeros resultó ser una variable totalmente independiente para determinar los valores tomados por la segunda componente principal, pero su correlación con la calificación final es incierta. Por otra parte, las dimensiones obtenidas del modelo de LS no parecen ser predictores fiables del rendimiento de los estudiantes.

Agudo-Peregrina et al. (2014) definen tres clasificaciones independientes del sistema de interacciones (basadas en el modo de participación, en la frecuencia de uso, y en el agente) y evalúan la relación de sus componentes con el rendimiento académico a través de dos modalidades diferentes de aprendizaje: ambiente de aprendizaje virtual apoyando el tradicional cara a cara y el aprendizaje en línea. Para ello, se realizó un estudio empírico con datos de seis cursos en línea y dos cursos F2F apoyados por entornos virtuales de aprendizaje. La extracción y análisis de los datos requiere el desarrollo de una herramienta *ad hoc* basada en la clasificación de interacción propuesta. Los resultados de este análisis muestran tendencias similares en el comportamiento de uso de los entornos virtuales de aprendizaje por los estudiantes a través de los diferentes cursos, y han ayudado a identificar qué interacciones pueden tener influencia real en el rendimiento académico de los estudiantes en estos entornos virtuales que, para cada clasificación, hay una relación entre un cierto tipo de interacciones y el rendimiento académico en los cursos en línea; por otro lado, esta relación no es significativa en el caso de los cursos F2F apoyados por los entornos virtuales. Estos resultados, que deben ser confirmados por otros estudios, proporcionan una primera base teórica para la selección de los datos pertinentes en el aprendizaje de los procesos de análisis.

Naveh, Tubin, and Pliskin (2010) examinan el uso de los estudiantes del LMS, su nivel de satisfacción, y cómo se correlacionan estas variables dependientes con las variables organizacionales en una universidad. Datos sobre 1212 sitios web de cursos se reunieron en el 2007 a partir de los *log* de los LMS y de cuestionarios de satisfacción. Los resultados indican un uso variado del LMS, un alto nivel de satisfacción y una baja correlación significativa entre el uso y la satisfacción. En cuanto a las variables organizativas, se encontró que el contenido del curso se correlaciona significativamente con el uso y la satisfacción; el tamaño del curso, el estado del instructor y la existencia de foros mostraron correlación significativa con el uso del LMS; finalmente, la disciplina del curso tuvo una baja correlación con la satisfacción.

Michinov et al. (2011) realizaron un análisis de los componentes principales sobre los ítems de la escala de procrastinación de Tuckman (1991). El objetivo de este estudio fue comprobar si la procrastinación influiría en el éxito de los estudiantes en el aprendizaje en línea, y que esto podría explicarse por el nivel de participación de los estudiantes en los foros de discusión. Siguiendo numerosos estudios sobre la procrastinación académica (por ejemplo, (Akinsola, Tella, & Tella, 2007; Beck, Koons, & Milgrim, 2000)), Michinov et al. (2011) extienden los resultados anteriores hacia entornos de aprendizaje en línea, lo que demuestra que los estudiantes que tienen más tendencia a posponer las cosas son los que presentan el rendimiento más pobre. Más importante aún, este estudio encontró que el nivel de participación en foros de discusión interviene en la relación entre la dilación y el rendimiento. En este estudio, los muy procrastinadores tienden a escribir algunos mensajes en los foros de discusión y, por tanto, realizan menos comunicación con sus compañeros (y su tutor/instructor) lo que tiene un impacto perjudicial sobre el rendimiento. Mediante la conexión tardía, pierden el hilo de las discusiones y luego temen aparecer como "recién llegados".

En consecuencia, se mantienen aislados hasta el final del curso. Por el contrario, los menos procrastinadores tienden a escribir un gran número de mensajes en los foros de discusión, y esta estrategia, que implica un alto nivel de participación e interacción con los otros, tiene un impacto positivo en su rendimiento; estos suelen mejorar su comprensión de la asignatura con la ayuda de otros estudiantes y el tutor/instructor.

Tales resultados sugieren que la simple oferta de foros de discusión asincrónica en los cursos en línea no genera necesariamente más interacciones de los estudiantes, esta depende en parte de las tendencias de prórroga de los estudiantes. Los autores del estudio concluyeron que, para medir la participación del estudiante, la investigación no debe basarse únicamente en el número de mensajes enviados, sino que también debe considerarse la posibilidad de que los estudiantes también aprenden a través de la participación pasiva en los foros con la lectura de las contribuciones de los demás estudiantes. Es posible que las personas con altas tendencias a la procrastinación aprendan a través de la observación, mientras que aquellos con baja tendencia a la procrastinación prefieran aprender mediante la participación con otros en los foros de discusión. Sin embargo, esto no permite concluir que las tendencias de alta procrastinación se relacionan con bajo rendimiento.

En los anteriores trabajos se ilustran algunas iniciativas para encontrar las actividades que pueden tomarse como indicadores para la predicción, reconocimiento de estilos de aprendizaje, o simplemente, semáforos de atención para mejorar el aprendizaje. Además, en la sección V.1.1 de este capítulo se mencionaron algunas herramientas visuales que comparten este objetivo, pero, lógicamente, lo abordan con el apoyo de técnicas de visualización e interacción.

Sin embargo, existen muchas formas de representar la información y procesarla con el objetivo de resaltar los patrones y/o tendencias que contenga o información importante para la toma de decisiones. Cada una de ellas presenta ventajas e inconvenientes en función del tipo, volumen y fuentes de los datos y siempre es preciso explorar diferentes técnicas de interacción y visualización. A continuación se presentan algunos trabajos que utilizan la visualización de la información y la analítica visual como medio de expresión y transmisión de nuevo conocimiento o exploración del espacio de información representado, y que de alguna manera u otra están relacionados con el presente trabajo de tesis.

### **V.1.3 Contribuciones visuales independientes al *eLearning***

Es importante mencionar que los intentos crecientes para adaptar o apoyar la educación tradicional con *eLearning* han traído como consecuencia un esfuerzo continuo para almacenar y organizar el contenido de los LMS. En este sentido, es importante no perder de vista la creación y el desarrollo de normas y herramientas visuales intuitivas que permitan el análisis y la gestión de este contenido en función de diversos puntos de vista.

Como se mencionó en las secciones IV.3 y IV.4, existen múltiples y diversas formas de representación de la información dependiendo del tipo y estructura de datos a analizar, explorar y/o representar. En los LMS, como se expresó en el modelo de referencia de Chatti et al. (2012) y las

diferentes dimensiones críticas de la analítica del aprendizaje Greller and Drachsler (2012) en la sección III.3.2, la estructura y las dimensiones de análisis de la información educativa son complejas.

Estas dimensiones comprenden información temporal, dado que los entornos educativos tienen un *log* que registra cada actividad y el instante de tiempo en que tiene lugar; información semántica, ya que el flujo de información tanto en los recursos como en las discusiones es analizable; social, dado que la interacción entre los diferentes participantes genera una red de comunicación y, por tanto, una red social; y, finalmente, jerárquica, puesto que cada universidad, facultad y profesor organiza su curso o conjunto de cursos de diferente forma.

A continuación se presentan algunos trabajos que no necesariamente están enfocados hacia el apoyo al *eLearning*, pero que por su metodología o propuesta se han tomado en cuenta en la realización de esta tesis. Para una mejor comprensión de estos se han dividido dependiendo del tipo de datos a representar y, por ende, del tipo de visualización. Se comenzará con las representaciones de espacios temporales –para representar la información temporal almacenada, después la representación de contenido semántico –utilizadas para la representación de textos, discusiones, bibliotecas, ontologías, etc. y, finalmente, la representación de la estructura de los datos –técnicas utilizada para la representación de datos (habitualmente multidimensionales), como por ejemplo los datos jerárquicos, las conexiones que se realizan entre las personas de un entorno, etc.

### V.1.3.1 Representación de espacios temporales

Existen muchas formas para representar la información y diferentes técnicas para interactuar con ella. El investigador y/o desarrollador tiene que concentrarse en las técnicas de visualización con el fin de obtener los máximos beneficios que ofrecen, por tal motivo, estas técnicas de visualización e interacción deben comprobarse por su utilidad y eficacia en el aprendizaje electrónico. Otro aspecto importante es especificar claramente qué tipo de datos tiene que representar y con qué método de visualización, de tal forma que el usuario interprete y obtenga la máxima comprensión de los datos (M. Khan & Shah Khan, 2011).

El caso más común es el uso de una disposición lineal (usando coordenadas cartesianas) con el tiempo representado en el eje X, de ahí que comúnmente se le denomina línea de tiempo.

El diseño de representaciones temporales no es tarea fácil ya que conlleva, en la mayoría de los casos, otra variable o variables a tomar en cuenta; Daassi et al. (2006) pasaron revista a las diferentes técnicas de visualización temporal disponible y las clasificaron, también Aigner, Miksch, Müller, Schumann, and Tominski (2007) han revisado las diferentes formas de mostrar datos temporales de acuerdo a las características de la tarea de análisis. Por otra parte, aunque visto con mucha menor frecuencia, el tiempo también se puede representar en forma de espiral.

Carlis and Konstan (1998) introducen una técnica de visualización en espiral, que muestra los datos a lo largo de una espiral para resaltar los atributos de serie de datos a lo largo del eje de espiral y los periodos por medio de los radios.

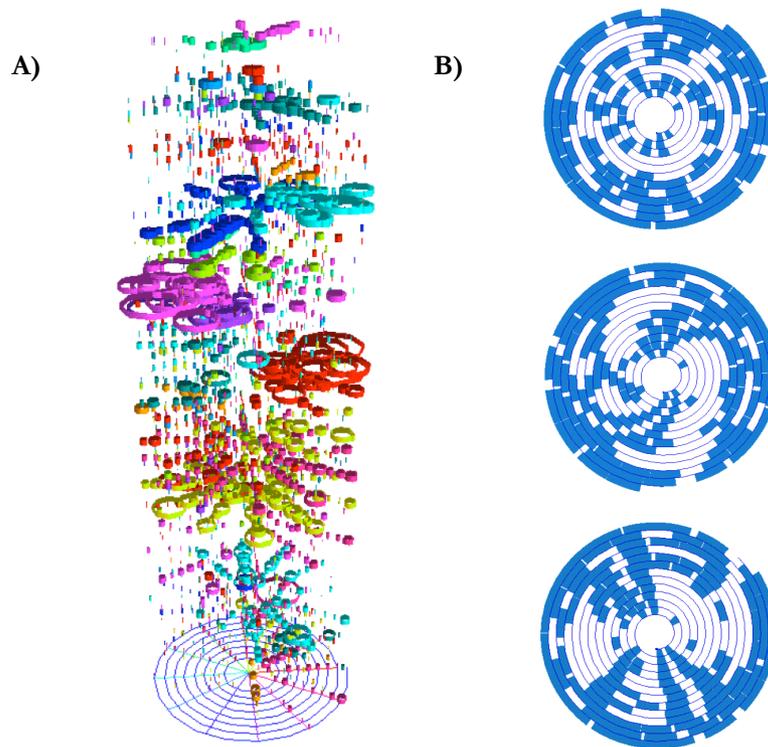


Fig. 33. A) Una pantalla espiral de porcentajes de consumo mensuales para los 112 alimentos durante el período 1980-1988. Cada alimento se ha asignado uno de los doce colores. B) El ajuste de una vista espiral de datos de sonido de cinco instrumentos. De arriba hacia abajo la estructura del sonido se devela por sí sola. Esta muestra tres puntos de vista diferentes sobre la base de tiempo diferente por los valores de vuelta. Los datos provienen de una grabación de un disco compacto de cinco instrumentos (Carlis & Konstan, 1998).

En este trabajo de Carlis and Konstan (1998) se explora esta manera de representación tanto en 2D como en 3D (ver Fig. 33), y un trabajo similar se puede encontrar en (Weber et al., 2001) (ver Fig. 34), pero en este caso la propia espiral sirve como un filtro para los datos de 3D. En (Bergstrom & Karahalios, 2007) este concepto se aplica a un conjunto de datos mucho más restringido, como es el caso de una sola conversación. Batty (2006) presenta una representación gráfica denominada el "reloj de rango" para examinar la dinámica de tres distribuciones: el tamaño de las ciudades de EE.UU. desde 1790, en el Reino Unido desde 1901 y en el mundo desde el año 430 a C.

Esta visualización es una combinación entre una representación en espiral y una representación tipo radar. En esta las órdenes de rango se trazan para cada ciudad en sentido de las agujas del reloj con el rango más alto en el centro y el más bajo en la circunferencia.

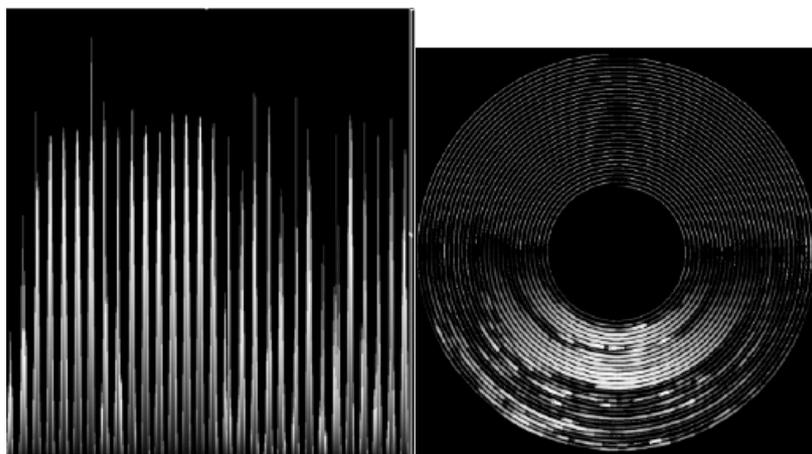


Fig. 34. Dos visualizaciones de la intensidad del sol utilizando aproximadamente el mismo espacio de la pantalla y el mismo esquema de codificación por color, en este caso en escala de grises. En la visualización de espiral es mucho más fácil comparar los días para detectar períodos nublados, o para ver eventos como el amanecer y la puesta del sol (Weber et al., 2001).

Otro trabajo con una representación en espiral es el de Karrie G. Karahalios and Bergstrom (2009), los cuales presentan un espejo social, que es un tipo de visualización social, con tres cualidades particulares. En primer lugar, según Karrie G Karahalios and Viégas (2006), las visualizaciones sociales son las visualizaciones sobre las personas, para las personas; son una visualización en las que la gente puede ver información acerca de sí mismos dentro de la información acerca de otros.

En cierto sentido, puede ser percibida la información de "nosotros y ellos", dado que cada participante ve la misma visualización. Además, los cambios perceptibles sutiles de cualquier participante aparecen en la visualización a medida que ocurren en tiempo casi real. Esto puede incluir la captura y visualización de los comportamientos sociales sutiles, como reír o toser. Debido a esta característica en tiempo casi real, la gente puede alterar rápidamente su comportamiento y, por tanto, la visualización para que los demás perciban cómo les gustaría ser percibidos en este espejo público. Además, los espejos sociales permiten explorar patrones y comportamientos de grupo a través de la experimentación en tiempo real, reproducción, anotación y reconfiguración (Karrie G Karahalios & Viégas, 2006).

Karrie G. Karahalios and Bergstrom (2009) presentan tres tipos de espejos sociales que se utilizan para demostrar cómo pueden servir como indicadores sociales: en primer lugar, el reloj de conversación ilustra la visualización del comportamiento de los participantes; representa los roles sociales, las normas y patrones de toma de turnos (ver Fig. 35). En segundo lugar, los votos de la conversación ilustran el aumento de la visualización de los participantes con la votación anónima (ver Fig. 36). La participación individual, así como las señales anónimas aportadas por los participantes, influye en esta visualización, que hace hincapié en la identidad social, los roles y la cohesión del grupo y la evolución de estos durante el tiempo de la interacción.

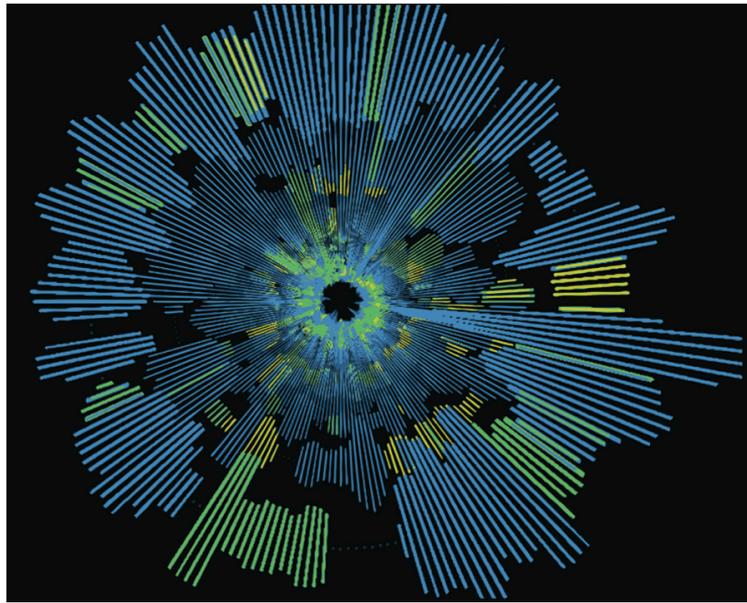


Fig. 35. El reloj de la conversación representa cada participante en un color diferente. La longitud de cada rectángulo es proporcional al volumen capturado en ese punto en el tiempo. Los puntos a lo largo de la circunferencia de cada anillo implican que la conversación está activa pero nadie está hablando. Cada anillo representa un minuto de conversación (Karrie G. Karahalios & Bergstrom, 2009).

Por último, las agrupaciones de la conversación ilustran la visualización de temas de contenido de cada participante en el tiempo (ver Fig. 37). Estos espejos sociales trazan la similitud de contenido, la formación de contenido con el tiempo y la contribución de los participantes. Dependiendo del tipo de tarea que se quiera apoyar, esta representación temporal en espiral deber ser complementada con una interacción adecuada. Así, por ejemplo, Therón, Gonzalez, Garcia, and Santos (2007) han desarrollado un árbol de revisión para visualizar las aportaciones de los miembros del equipo a través de varias revisiones, referencias y largos períodos de tiempo, en el mismo elemento o documento dentro de un proyecto de *software*.

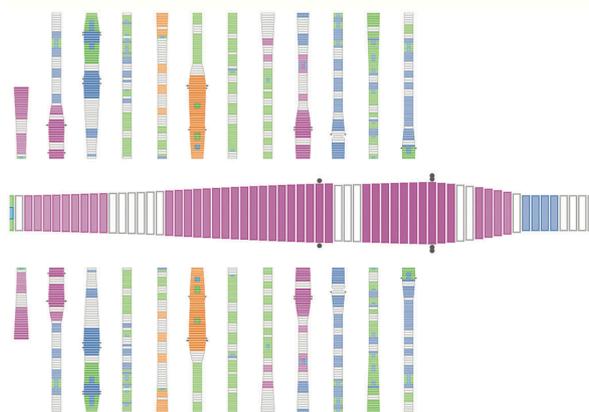


Fig. 36. La representación de los votos de la conversación es compatible con un botón de votación para contribuir de manera significativa, y este voto está representado por medio de los puntos en los bordes estrechos de las barras rectangulares (Karrie G. Karahalios & Bergstrom, 2009).

Su propuesta consiste en la representación de los grandes árboles de revisiones donde las líneas de base tienen varias ramas y cada rama múltiples revisiones de elementos; este tipo de navegación a través del árbol de versiones ofrece una de las técnicas importantes de la visualización de la información, el foco + contexto, en apoyo a la interactividad para permitir la inspección de más de una línea de base a la vez, exhibiéndose la colaboración de los desarrolladores para cada línea de base y correlacionar toda la información con la línea de tiempo.

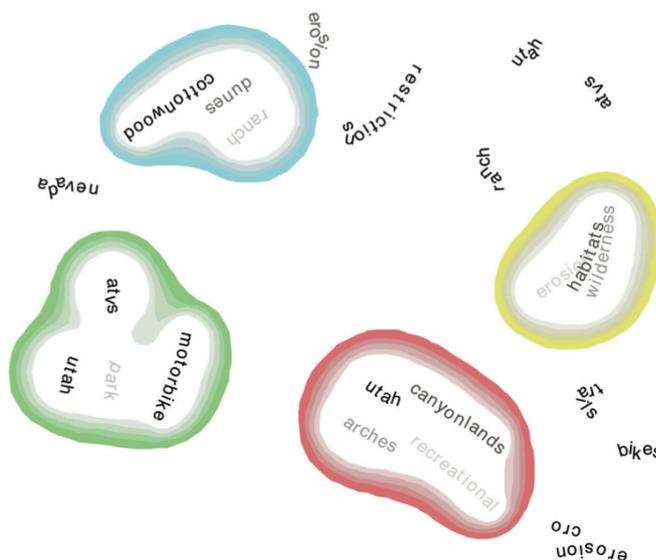


Fig. 37. Una vista de la visualización de agrupaciones de la conversación. Las palabras encerradas están a una distancia común dada su similitud. Las palabras fuera de las áreas cerradas están por encima del umbral de similitud de ese clúster tema. Las palabras aparecen en tiempo real conforme se habla (Karrie G. Karahalios & Bergstrom, 2009).

En la visualización exploratoria de información, como se describió en la sección IV.4, el objetivo es presentar la información de manera que los usuarios pueden fácilmente discernir patrones; los patrones revelan tendencias, relaciones, anomalías y estructura de los datos, y pueden ayudar a los usuarios a confirmar conocimiento o hipótesis. Quizás lo más importante es que también plantean preguntas inesperadas de los usuarios que finalizan en nuevos conocimientos. El reto es crear visualizaciones que permiten a los usuarios encontrar patrones rápida y fácilmente. Una técnica muy eficiente en el caso de datos temporales la que S. Havre, Hetzler, and Nowell (2000); (2002) proponen, que parten de la metáfora de un río (de ahí el nombre, *ThemeRiver*, de la técnica), la cual proporciona a los usuarios una visión macro de los cambios temáticos en un corpus de documentos de más de una dimensión. Está diseñada para facilitar la identificación de tendencias, patrones, y la aparición inesperada o no de temas o tópicos.

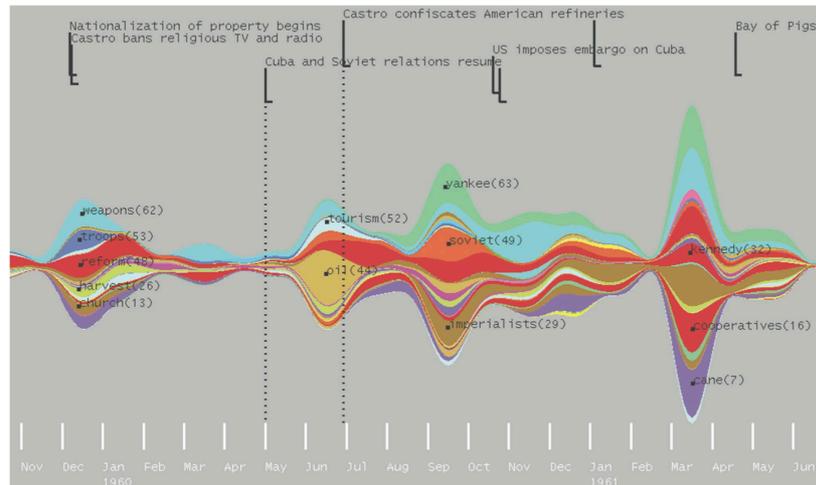


Fig. 38. *ThemeRiver* utiliza una metáfora del río para representar los temas de una colección de discursos de Fidel Castro, entrevistas y artículos desde finales de 1959 hasta mediados de 1961 (S. Havre et al., 2000; Susan Havre et al., 2002).

En este prototipo se usa el tiempo como la dimensión de serie, se proporciona información contextual a través de una línea de tiempo y los marcadores de interés de los eventos que ocurren. En la Fig. 38 del prototipo del sistema, que trata de evidenciar variaciones temáticas en el tiempo dentro de una gran colección de documentos, el "río" fluye de izquierda a derecha a través del tiempo, el cambio del ancho del mismo representa cambios en la fuerza temática de los documentos asociados temporalmente, por lo que se crean "corrientes" de colores que fluyen en el río estrechándolo o ensanchándolo para indicar disminuciones o aumentos en la fuerza de un tema individual o de un grupo de temas en los documentos asociados.

Sin embargo, el éxito interactivo de la visualización depende no solo de las codificaciones visuales, sino también en los mecanismos para navegar por el espacio de la información visualizada, como ya se mencionó en el apartado IV.3. Estos mecanismos de interacción pueden tomar muchas formas, incluyendo la vista panorámica y *zoom* (Dachselt & Weiland, 2006), o consultas de texto y *widjets* de consultas dinámicas (Willett, Heer, & Agrawala, 2007). Además, depende también de las señales apropiadas de navegación visual, estas pueden ayudar a los usuarios por el recorrido y el perfeccionamiento de su exploración.

En un intento de resolver el problema que representan los amplios períodos temporales, Dachselt and Weiland (2006) propusieron una deformación bifocal en su propuesta *TimeZoom*, que permite la definición de vistas simultáneas con nivel de *zoom* (globales o detalladas) de regiones de tiempo independiente de la línea de tiempo global (ver la Fig. 39.1). Las regiones pueden ser definidas y seleccionadas, indicadas por barras de selección coloreadas (compárense las zonas a, b y c en la Fig. 39.2). Las regiones seleccionadas pueden ser modificadas en intervalo de tiempo y en nivel de detalle (compárense la Fig. 39.3).

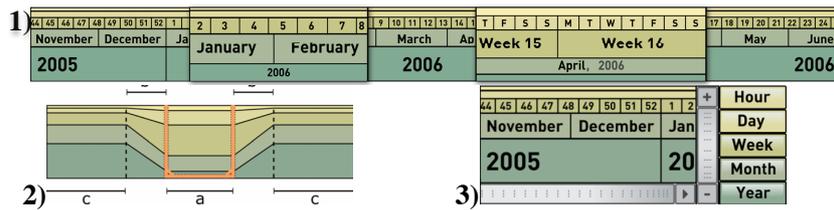


Fig. 39. 1) *TimeZoom* con dos regiones de atención adicionales que muestran diferentes niveles de detalle. 2) Definición de las diferentes zonas de las regiones de enfoque. 3) Las partes de la interfaz de usuario, (Dachsel & Weiland, 2006).

En este mismo sentido, André et al. (2007) con el fin de representar los datos temporales, con la necesidad de poder visualizar la información temporal (a) dentro de las jerarquías dinámicas, (b) a través del concepto relaciones/asociaciones, y (c) en una visión general a gran escala con una vista del detalle, sugieren un filtro visual en la vista general de la línea de tiempo con el objetivo de cambiar dinámicamente los datos del detalle que se representan en su propuesta (ver la Fig. 40). El cronograma propuesto hace varias contribuciones al trabajo de visualización de línea de tiempo: permite representar visualizaciones anidadas de las jerarquías asociadas temporalmente; apoya firmemente la toma de decisiones por parte del usuario al permitir al usuario controlar el detalle de cada faceta individual, el usuario puede llegar a lo que quiere saber; ayuda a proporcionar información útil, incluso cuando una vista llega al máximo detalle, permite elegir, mediante los controles de usuario, qué atributos se desean ver destacados en la visión general.

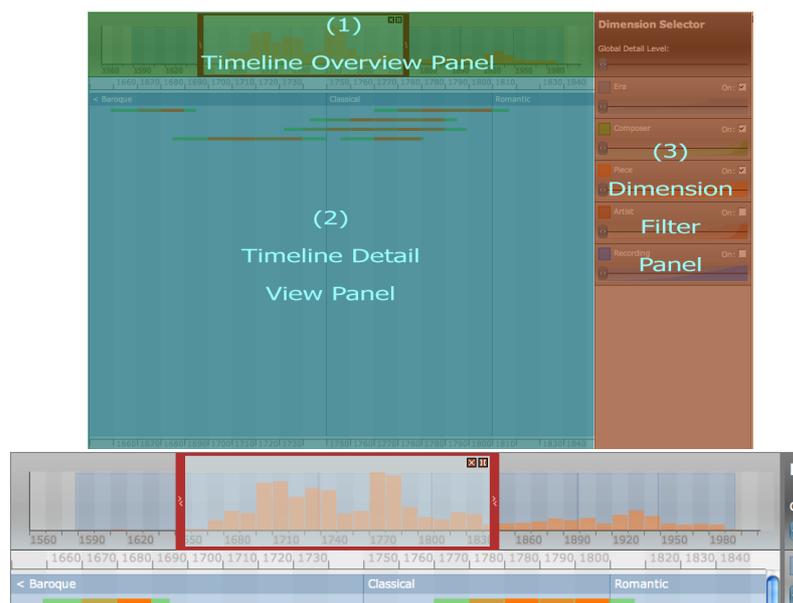


Fig. 40. Arriba, 1) panel de la visión general línea de tiempo, 2) panel de la vista cronológica de detalles y 3) el panel del filtro de dimensión. Los tres paneles tienen atributos que han sido diseñados para hacer frente a los problemas de escala, la jerarquía y las relaciones entre estos y el valor temporal. Abajo, el panel de información general continuo que muestra los datos temporales como un histograma, para ampliar grandes conjuntos de datos sin perder el contexto -visión detallada del panel (1)- (André et al., 2007).

Como se mencionó anteriormente en los entornos educativos *online*, además de la información temporal, se encuentra almacenada y continuamente compartida información del contenido de los cursos, comentarios, dudas, discusiones, consejos, etc. A este contenido se le conoce como contenido semántico y es esta interacción en sí la que forma parte de la transferencia de conocimiento, por tanto, es objetivo importante del análisis. A continuación se describen algunos trabajos que representan visualmente contenido semántico fuera del ámbito del *eLearning*.

### V.1.3.2 Representación de contenido semántico

La forma más común y directa de representar una idea, concepto u objeto es mediante la utilización de la palabra o palabras que lo definen directamente o a sus principales características, es decir, el contenido semántico principal que lo define (Viégas & Wattenberg, 2008).

El término *folksonomy*, o etiquetado social, es una forma relativamente nueva y de gran alcance para organizar la información. Se han tenido etiquetas o palabras clave por algún tiempo, pero solo con la naturaleza distribuida de Internet, los ordenadores, y la participación de gente que suministra un rico vocabulario para una variedad de recursos, se ha logrado ver la opinión colectiva y observarse en maneras interesantes (Mathes, 2004; Russell, 2007). Las nubes de palabras (del inglés, *Tag Clouds*) o nube de etiquetas por lo general tienen un fin determinado, por ejemplo presentar una descripción visual de una colección de texto (Donath, 2002) por medio de algoritmos automáticos de organización (Salton, Allan, & Buckley, 1994) y/o de creación de las etiquetas (Porter, 1980), también se utilizan para representar el etiquetado social (Bielenberg & Zacher, 2005; Hammond, Hannay, Lund, & Scott, 2005; Mathes, 2004).

En este sentido, quizás el primer ejemplo de etiquetado social puede ser el resultado de un experimento llevado a cabo por el psicólogo social Stanley Milgram and Jodelet (1976). Milgram y Jodelet pidieron a múltiples personas que nombraran los monumentos históricos de París, para posteriormente crear un "mapa mental" colectivo de la ciudad mediante la propiedad tamaño de la fuente de las palabras para mostrar con qué frecuencia se menciona cada lugar.

Hoy en día, los sitios web de alto perfil promueven un considerable esfuerzo de investigación para mejorar los diseños de nubes de etiquetas: *Flicker*<sup>19</sup>, *Delicio.us*<sup>20</sup>, *TagCrowd*<sup>21</sup>, *TagLine Generator*<sup>22</sup>, *Tag Cloud Generator*<sup>23</sup>, *Wordle*<sup>24</sup>, *Manyeyes*<sup>25</sup>, etc. Por ejemplo, *TagLine Generator* permite generar una secuencia de nubes de etiquetas, que se asocia con el tiempo, a partir de una colección de documentos.

---

19 <http://www.flickr.com/photos/tags>

20 <http://delicious.com/tag>

21 <http://www.tagcrowd.com>

22 <http://chir.ag/projects/tagline>

23 <http://www.tagcloudgenerator.com>

24 <http://www.wordle.net>

25 [http://maneyes.alphaworks.ibm.com/maneyes/page/Tag\\_Cloud.html](http://maneyes.alphaworks.ibm.com/maneyes/page/Tag_Cloud.html)

Además implementa un control deslizante dinámico que se utiliza para navegar por los puntos de tiempo, pero solo se muestra una nube de etiquetas en un momento determinado de tiempo. Stefaner (2007) presenta en su herramienta *elastic tag map* una nube de etiquetas donde todas ellas se ajustan en función de su frecuencia (ver Fig. 41). Las etiquetas tienden a seguir una distribución de ley de potencias o ley potencial, una escala proporcional no es aconsejable, ya que esto daría lugar a algunas etiquetas enormes y un gran número de las más pequeñas. En consecuencia se ha aplicado una escala logarítmica, lo que lleva a una disminución lineal en tamaño con respecto al rango de la etiqueta. Lo primero que se puede observar es que las etiquetas más frecuentes también definen los extremos del espacio distribuido y en consecuencia se colocan en las zonas exteriores de la nube de etiquetas. En segundo lugar, para la mayoría de los usuarios, algunos de los códigos que se utilizan con frecuencia están también muy fuertemente relacionados (véase por ejemplo, la filosofía, la religión, la cultura), mientras que otros definen su propia "esquina" del conjunto de datos en ausencia de otras etiquetas de uso frecuente. Además, un histograma muestra las frecuencias globales de las etiquetas para exponer la naturaleza del histograma de las colecciones de etiquetas y subcolecciones especialmente.

Es interesante ver que todas las colecciones examinadas muestran una clara distribución de histograma largo, sin embargo, con diferentes pendientes (ver las esquinas inferiores izquierdas en cada imagen de la Fig. 41). Una segunda idea interesante es la composición del histograma al seleccionar una de las etiquetas de bajo rango, casi siempre producidas en una de las etiquetas de rango superior, junto con sus etiquetas de rango medio asociados.

Sobre las nubes de etiquetas se han desarrollado mejoras adicionales que incluyen el uso de algoritmos espaciales (para organizar y colocar las palabras en una nube de etiquetas en un área más pequeña) y los algoritmos de agrupación (procedimiento en el que las etiquetas que se utilizan juntas o tienen significados similares se colocan cerca las unas de las otras (Y. Chen, Yang, & Ribarsky, 2009)). Por ejemplo, Kaser and Lemire (2007) en su trabajo organiza las etiquetas en las tablas anidadas para sitios basados en HTML mediante un algoritmo de empaquetado por medio de una automatización de diseño electrónico (del inglés, *Electronic Design Automation*, EDA).

En este sentido, Torniai, Jovanovic, Gasevic, Bateman, and Hatala (2008) proponen que el tamaño de la etiqueta empleada en la nube de etiquetas y el color sirvan para transmitir a los profesores la información que describe la popularidad de las etiquetas y la relevancia respectivamente. Por otro lado, Bateman, Gutwin, and Nacenta (2008); Rivadeneira, Gruen, Muller, and Millen (2007) comparan nueve propiedades visuales de las nubes de etiquetas por sus efectos sobre la búsqueda visual de las etiquetas. Sus resultados muestran que el tamaño de la fuente y el peso de la fuente tienen efectos más fuertes que otros, tales como la intensidad del color, el número de caracteres o el área de la etiqueta.

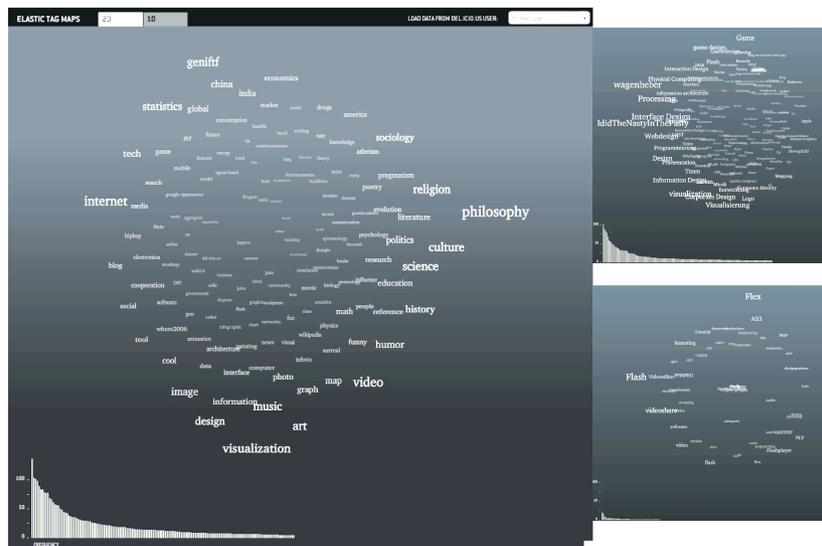


Fig. 41. Aplicación *elastic tag maps* (Stefaner, 2007). [http://well-formed-data.net/experiments/tag\\_maps\\_v5/](http://well-formed-data.net/experiments/tag_maps_v5/).

El orden de las palabras en la nube de etiquetas es también un punto a tomar en cuenta, Halvey and Keane (2007) comparan las nubes de etiquetas con las listas tradicionales (horizontales y verticales), en cada caso comparan un orden alfabético contra un orden normal.

Por medio de entrevistas y encuestas sobre la facilidad de encontrar una etiqueta específica Halvey and Keane (2007) llegaron a la conclusión de que las listas se comportan mejor que las nubes de etiquetas y que el orden alfabético acelera aún más la velocidad de búsqueda en las nubes de palabras; el tamaño de la fuente es muy importante para la rapidez y la facilidad con las que los usuarios encuentran la información; la posición de las etiquetas es también muy importante y mejora el escaneo de listas y nubes en lugar de leerlas.

Otra manera encontrada en la literatura de optimizar las nubes de etiquetas es por medio de la utilización de varias técnicas de visualización en una misma representación gráfica de los datos. Un ejemplo de esto es el *software* de nubes de palabras paralelas (del inglés, *Parallel Tag Clouds*, PTC) que mezcla el uso de dos representaciones de información, como son las coordenadas paralelas y las nubes de etiquetas. Esta aplicación está diseñada para proporcionar una visión general de una colección de documentos mediante la incorporación de elementos gráficos de coordenadas paralelas, para lo que se codifica el tamaño de texto de nubes de etiquetas tradicionales (Collins, Viegas, & Wattenberg, 2009). En la Fig. 42 se puede observar cómo PTC puede exponer un conjunto de nubes de palabras verticalmente y luego relacionarlas mediante líneas dependiendo de lo que se esté buscando. Aunque PTC muestra múltiples nubes simultáneamente, no representan explícitamente las tendencias y, por tanto, la comparación de múltiples nubes de etiquetas.

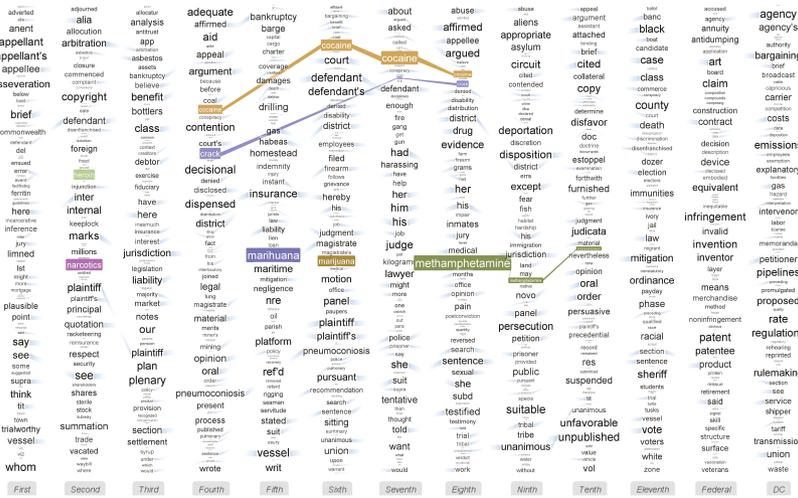


Fig. 42. Un PTC que revela las diferencias en la prevalencia de drogas entre las decisiones escritas de los Tribunales de Estados Unidos del Circuito de Apelaciones. Los Tribunales de Circuito de Apelación se componen de las 12 divisiones judiciales regionales (numeradas de primero a undécimo, más el Circuito DC) (Collins et al., 2009).

En un esfuerzo de expresar las tendencias en las nubes de etiquetas, Dubinko et al. (2007) presentan una visualización de nubes de etiquetas distribuidas a través del tiempo, con la intención de representar la evolución de las etiquetas en tiempo real. Con el mismo objetivo Bongshin et al. (2010) integran mini gráficos (del inglés *sparklines*), que son un tipo de representación de información gráfica que se caracteriza por su pequeño tamaño y alta densidad de datos (ver Fig. 43). En este trabajo se implantó en una nube de etiquetas para transmitir las tendencias entre múltiples nubes. Además proporcionaron un estudio controlado para explorar la eficacia de la representación temporal incrustada en la nube, que se ha denominado *SparkClouds*.

En este mismo sentido de expresar las tendencias, Weiwei et al. (2010) introdujeron un método de visualización en el que se expresa la evolución de las etiquetas sobre el tiempo en un cuadro de tendencias independiente pero enlazado con la nube, lo que permite mediante la selección de un período de tiempo, visualizar las nubes de palabras con el objetivo de expresar las evoluciones temporales de contenido en un conjunto de documentos.



Fig. 43. Alternativas de diseño para unificar las nubes de etiquetas y los mini gráficos (Bongshin et al., 2010).

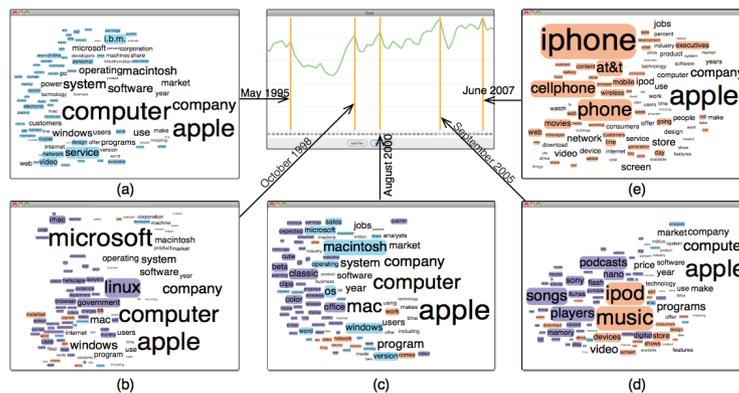


Fig. 44. En el centro de la parte superior de la figura se presenta un visor gráfico de tendencias que muestra una curva de importancia extraído de una colección de documentos con diferentes puntos de tiempo (Weiwei et al., 2010).

En la Fig. 44 el eje X codifica el tiempo y el eje Y codifica la significancia de la nube de palabras. La curva verde en el gráfico representa la importancia de medición de las nubes de palabras en diferentes pasos de tiempo. Cinco nubes de palabras ((a) - (e) en la figura) se crean utilizando el algoritmo para cinco puntos de tiempo seleccionados en los que se observaron altos valores de significación.

Uno de los primeros intentos de utilizar la vista nube de etiquetas como parte de un sistema con múltiples vistas coordinadas (del inglés, *Coordinated Multiple Views*, CMV, Roberts (2007)) fue presentado en el trabajo de (Therón, Santamaria, García, Gómez Aguilar, & Paz-Madrid, 2007). El cual es una herramienta de analítica visual, que fue desarrollada para el análisis de datos cinematográficos para el Concurso de *InfoVis*, el cual consistió en analizar de forma visual un conjunto de 20.204 películas y 149.397 personas, recogidas de <http://www.imdb.com>; en su trabajo ampliaron el conjunto de datos original para incluir nueva información importante, resultando un incremento a un total de 21.247 películas, 235.442 personas en el elenco.



Fig. 45. En la figura se muestran las actrices que han ganado premios como actriz principal o secundaria (en la parte central y abajo). Algunas de ellas son Julia Roberts, Nicole Kidman, Jennifer Connely y Renee Zellweger. Se puede ver el perfil de sus películas en las coordenadas paralelas (arriba a la derecha) y el género más relevante de sus películas en la nube de palabras (arriba a la izquierda) (Therón, Santamaria, et al., 2007).

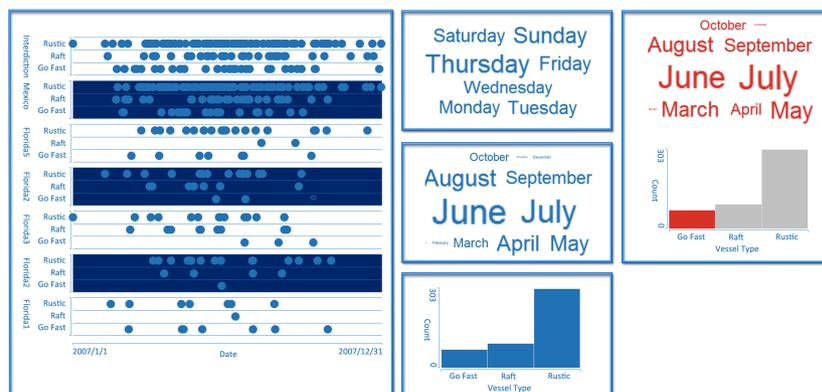


Fig. 46. La figura representa los eventos en el año 2007 de tres diferentes barcos utilizados para inmigración. Es muy obvio que no hay tendencias y/o patrones de uso del tipo de barco, y también que el número de barcos de algún tipo no son lo mismo (Matkovic, Ammer, Gracanin, Purgathofer, & Lez, 2010).

Esta herramienta utiliza las técnicas de visualización de las coordenadas paralelas, *overlaper*, nube de palabras (ver Fig. 45). Estas últimas se utilizan para representar el género de las películas siendo esta una forma de filtrar los datos y cambiar las representaciones que están enlazadas.

Posteriormente Matkovic et al. (2010) en su trabajo utiliza la etiqueta para identificar los meses en que los inmigrantes tienen más éxito (la mayoría de los desembarques), de forma que modifica en cada interacción las otras visualizaciones enlazadas (ver la Fig. 46). Mientras los tipos *Go Fast* y *Raft* son muy similares en número, los barcos del tipo *Rustic* se utilizan en un número mucho mayor. Algo muy importante es que existe mucha menos actividad inmigrante durante el otoño y el invierno, que es probablemente debido al mal tiempo. Esto también se confirma en la vista de nube de etiquetas con los meses en que enero, febrero, noviembre y diciembre (meses de invierno) son apenas visibles.

La vista nube de etiquetas con los días de la semana muestra que el jueves y el domingo son los días más activos y hay una tendencia cada vez mayor hacia el final de la semana. El histograma confirma qué tipo de embarcación *Rustic* es la más utilizada. A la derecha se puede ver que los barcos de tipo *Go Fast* rara vez se utilizan durante el invierno (si se selecciona la barra de *Go Fast* se puede observar en el vista nube de etiquetas que los meses de invierno son apenas visibles o no visibles en absoluto).

Además, en esta propuesta también han integrado la información temporal en la nube de etiquetas. En la Fig. 47 se puede observar cuándo los inmigrantes tienen más desembarcos exitosos y podría identificarse por meses. También, se puede visualizar en dos puntos de vista, una con y otra sin contexto, la parte roja representa desembarcos exitosos. Las inmigraciones son más exitosas durante marzo, mayo y junio. Es importante resaltar la técnica de cepillado para mostrar el contexto de la información en gris, ya que potencia la expresividad de la herramienta.



Fig. 47. El éxito de los inmigrantes. El punto de vista de la izquierda muestra fracaso (contexto) como gris y éxito (cepillado) en rojo. Marzo, mayo y junio fueron meses de mayor éxito para los inmigrantes. La vista derecha muestra lo mismo pero con una única opción seleccionada (por lo que no se muestra el contexto) (Matkovic et al., 2010).

Las nubes de palabras no solo se representan en listas o párrafos formados por las palabras o etiquetas de la nube, Gambette and Véronis (2010) presentan una herramienta para el análisis de textos que construye un árbol. El tamaño de las palabras clave simplemente refleja la frecuencia de las palabras, como es generalmente el caso en las nubes de etiquetas, o puede utilizarse de manera diferente, por ejemplo, para reflejar la significación estadística de las distintas palabras con respecto a un corpus de referencia.

Por ejemplo, los árboles de palabras que representan los discursos de Barack Obama y George Bush podrían ser contrastadas de esa manera: las palabras más grandes serían aquellas que son las más destacadas de cada uno de ellos. En la Fig. 48 los colores tienen un uso obvio, como es la categorización de acuerdo a los temas (por ejemplo, deportes, política, negocios, etc. en un sitio web de noticias); el brillo se utiliza para demostrar si la palabra aparece en un mismo lugar en el texto o en muchos lugares (según algunos coeficiente de dispersión).

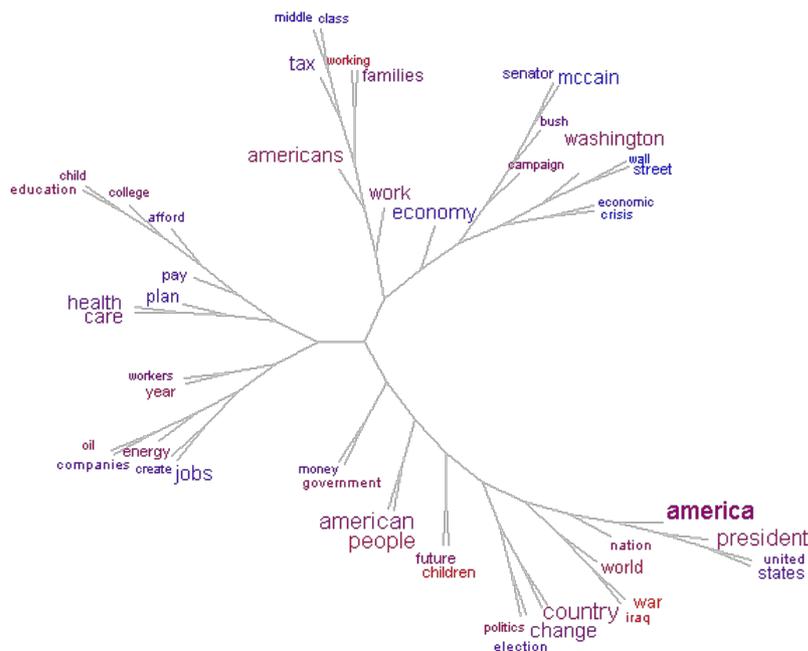


Fig. 48. Árbol de palabras de los discursos de campaña presidencial de Obama, con la distancia de Jaccard, y la coloración cronológica. El rojo corresponde al comienzo de la campaña ("niños", "Irak", "guerra", "mundo"), mientras que azul corresponde a la final ("McCain", "Wall Street", "crisis", "impuestos") (Gambette & Véronis, 2010).

Si el corpus se asocia con fechas, las palabras más recientes se pueden visualizar con la mayor intensidad o con un color diferente, en el caso de la Fig. 48 se utiliza para expresar la antigüedad. La información puede ser transmitida también por el espesor de borde, longitud y color. Sin embargo, queda por ver cuánta información se puede superponer en el mismo árbol sin molestar a su legibilidad global. Un buen truco para mejorar el aspecto general del árbol de la nube es forzar la longitud de la arista a la unidad. Esto evita el problema de rama larga que se produce con la mayoría de las distancias semánticas: las ramas principales a las hojas son muy largas y la estructura del árbol se oculta en el centro. La visualización obtenida refleja la distancia semántica menos fielmente, pero la topología del subárbol aparece más claramente.

Los árboles o gráficos de red se utilizan más frecuentemente en la representación de redes sociales o de estructuras jerárquicas. Estos son una forma de representación muy intuitiva y útil para la representación de estructuras de datos. Como se ha comentado con anterioridad en el proceso de aprendizaje se cuenta con estructuras jerárquicas que conforman el organigrama del contenido de un curso y además la red de conexiones que se forma a través de la interacción de los usuarios del LMS. A continuación se detallan algunos trabajos de representación de estructuras de datos.

### V.1.3.3 Representación de la estructura y dimensionalidad de los datos

Como se mencionó en el Capítulo III, la importancia de proveer al profesorado con datos en tiempo real relacionados con la creación de estructuras sociales y relaciones es fundamental para informar sobre cualquier modificación necesaria en las actividades con el objetivo de mejorar el aprendizaje. Por tanto, el nivel de supervisión de los medios de comunicación se ha convertido recientemente en un elemento importante. Existe una fuerte evidencia para sugerir que dentro de las actividades bien estructuradas, los procesos de construcción del conocimiento alcanzan niveles superiores de pensamiento crítico y de que los estudiantes son capaces de establecer y sostener grupos cohesivos. Por otro lado, algunos profesores cuentan con la participación de los foros de discusión en línea como una pequeña parte de la nota final. Así que la participación en foros de discusión por los estudiantes también puede estar influida por este tipo de supervisión. Esta prueba sirve para justificar la necesidad de análisis de redes sociales automatizada y de herramientas que sean capaces de ofrecer para los tutores y profesores una representación en "tiempo real" del análisis (Bakharia & Dawson, 2011; Srivastava & Gupta, 2014).

El estudio de las redes, en base a la teoría matemática de grafos, es uno de los pilares fundamentales de la matemática discreta. Euler en 1735 presentó la solución al problema de los puentes de Königsberg y esta se cita a menudo como la primera verdadera prueba de la teoría de redes, y durante el siglo XX la teoría de grafos se ha convertido en un importante cuerpo de conocimientos (Sachs, Stiebitz, & Wilson, 1988). Las redes también han sido ampliamente estudiadas en las ciencias sociales.

Ya en la década de 1930, los sociólogos se dieron cuenta de la importancia de los patrones de conexión entre las personas en la comprensión del funcionamiento de la sociedad humana. Estudios de redes típicas en sociología implican la circulación de cuestionarios, preguntando a los encuestados sobre los detalles de sus interacciones con los demás. Uno puede entonces utilizar las respuestas para reconstruir una red en la que los vértices representan individuos y los bordes de las interacciones entre ellos. Estudios típicos de redes sociales abordan cuestiones de centralidad (que individuos están mejor conectados a los demás o que tienen más influencia) y la conectividad (cómo las personas se conectan entre sí a través de la red) (Newman, 2003).

De acuerdo con Barry and Fulmer (2004), el grado en que otras personas pueden recuperar y leer los mensajes de una persona puede afectar su uso; es decir, el hecho de que todo el mundo tenga acceso a los mensajes en el foro de discusión en línea, tanto los profesores como los estudiantes, en cualquier momento podría afectar el uso del LMS. Por ejemplo, un comentario no adecuado, una pregunta que resulte muy obvia para otros, o incluso una falta de ortografía de un estudiante podría cambiar la opinión de los compañeros sobre él/ella y, por tanto, su intención y ánimo de seguir participando.

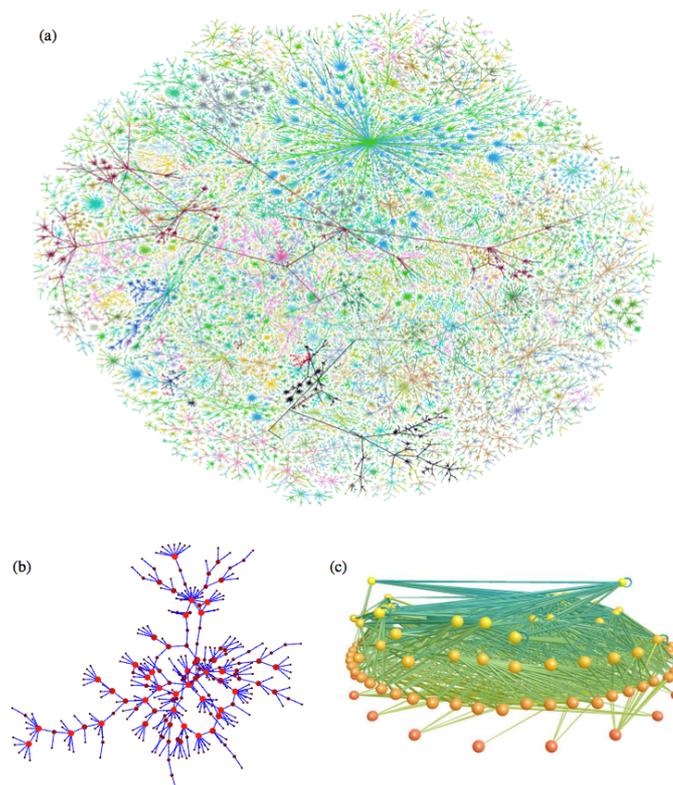


Fig. 49. Tres ejemplos de los tipos de redes que son el tema de esta revisión. (a) Una visualización de la estructura de la red de Internet a nivel de "sistemas autónomos" - grupos locales de computadoras cada uno en representación de cientos o miles de máquinas. (b) Una red social, en este caso de los contactos sexuales. (c) Una red trófica de las interacciones depredador-presa entre las especies en un lago de agua dulce (Newman, 2003).

Newman (2003) revisa recientes (y algunos no tan recientes) trabajos sobre la estructura y función de los sistemas en red de este tipo. En este artículo se revisaron estudios empíricos centrados en una serie de propiedades estadísticas de las redes que han recibido una atención especial, incluyendo longitudes de trayectoria, grados de distribución, agrupación, y la capacidad de recuperación. La observación más importante es que las redes están en general muy lejos de ser creadas con aleatoriedad, por el contrario estas tienen firmas estadísticas altamente distintivas, algunas de las cuales como altos coeficientes de agrupamiento y grados de distribución altamente sesgados son comunes a las redes de una amplia variedad de tipos (ver Fig. 49).

Los grafos aparecen en numerosas aplicaciones como la navegación web, los diagramas de estado de transición o las estructuras de datos. La capacidad de visualizar y navegar por estas potencialmente grandes y abstractas estructuras, es a menudo una parte crucial de una aplicación.

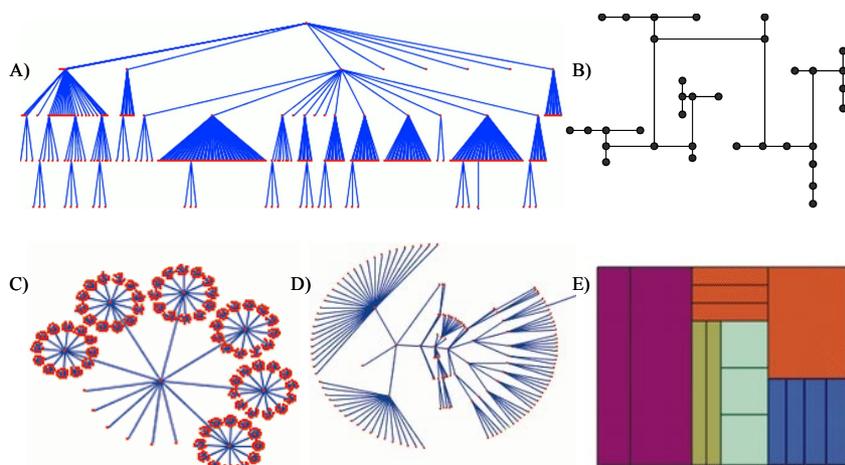


Fig. 50. A) Un diseño de árbol para un modelo de datos que genera un gráfico grande. B) Disposición H-árbol. C) Vista Globo. D) Vista Radial. E) *TreeMap*: rectángulos con color pertenecen al mismo nivel de la (árbol) jerarquía ((Herman, Melancon, & Marshall, 2000).

En este sentido, como se mencionó en el Capítulo IV, la visualización de información tiene requisitos específicos, lo que significa que se toma en cuenta los resultados del dibujo tradicional de un grafo desde una perspectiva diferente. Herman et al. (2000) exponen las diferentes formas de representar un grafo en 2D (ver Fig. 50) y 3D (ver Fig. 51), de las cuales se pueden resaltar aspectos útiles, por ejemplo si se requiere expresar una estructura de relaciones muy grande, se requiere notar la profundidad a simple vista y en poco espacio el *TreeMap* (ver Fig. 50.E) o el árbol radial (ver Fig. 50.D) son una opción para el caso de 2D y para el de 3D un árbol hiperbólico (ver Fig. 51.E) o un cubo de información (ver Fig. 51.B) son técnicas aplicables, sin embargo, si no se tiene restricción de espacio un diseño de árbol estilo dendograma (ver Fig. 50.A) para 2D y árbol radial 3D (ver Fig. 51.A) para el uso de tres dimensiones; por otro lado, en caso que se requiera solo explorar el árbol de la estructura de los datos, una técnica por la que se podría optar es un H-árbol (ver Fig. 50.B) o un árbol en vista globo (ver Fig. 50.C) haciendo uso de dos dimensiones y para 3D un árbol de conos (ver Fig. 51.C).

Además de se deben tomar en cuenta las diferentes formas de interacción y navegación (consultar la sección IV.3.2) con el objetivo de mejorar la expresividad y potencia en el reconocimiento de patrones y creación de nuevas hipótesis o confirmación de estas. Un ejemplo de la utilidad de un grafo es *Storylines*, presentado por Zhu and Chen (2007), el cual proporciona un marco que permite análisis visuales para estudiar un conjunto de textos no estructurados sin conocimiento previo de su estructura temática. Por medio de la creación de una representación de red, el sistema integra de manera innovadora la indexación semántica, el procesamiento del lenguaje natural y el análisis de redes sociales.

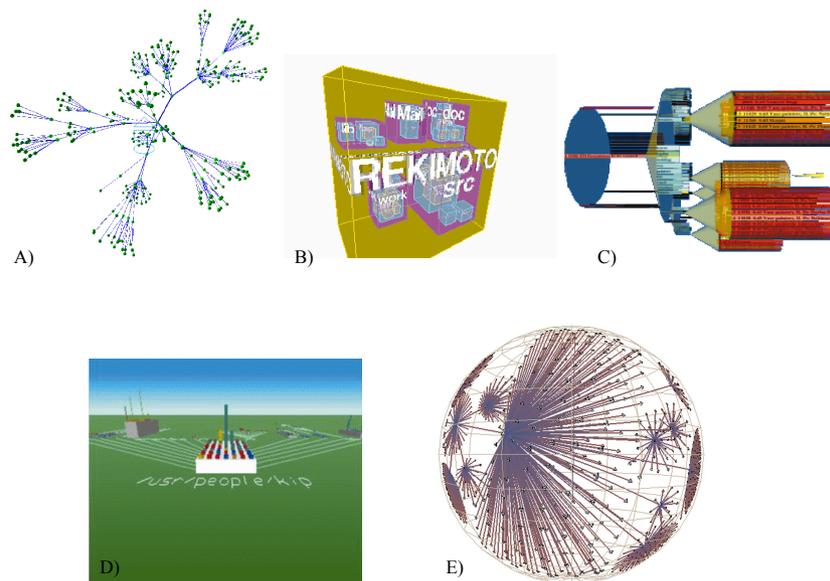


Fig. 51. A) Versión en 3D de un algoritmo radial. B) Información Cube. C) Un árbol de conos. D) Navegador de archivos de sistema en SGI. E) Vista hiperbólica de un árbol en 3D (Herman et al., 2000).

Este trabajo, genera grupos de documentos en función de su concordancia de palabras clave en el espacio de documentos. Tales grupos de documentos forman una historia (ver Fig. 52). *Storylines* comienza con una visión general de toda la colección de documentos para que los analistas puedan explorar la distribución de documentos a través del tiempo mediante búsquedas de palabras clave. Estas palabras clave se han agrupado temporalmente, permitiendo así la navegación a través del tiempo por parte de los analistas.

Al igual que la naturaleza jerárquica o de red que puede encontrarse en muchos conjuntos de datos (ciertamente el caso de datos en contextos educativos), a menudo los datos son multivariantes/multidimensionales. Una de las técnicas de visualización más exitosas para este tipo de datos se denomina Coordenadas Paralelas (del inglés, *Parallel coordinates*, PC), desde su introducción temprana de Inselberg (Inselberg & Dimsdale, 1990), (Inselberg & Dimsdale, 1990), han sido ampliamente utilizadas (Tory, Potts, & Moller, 2005).

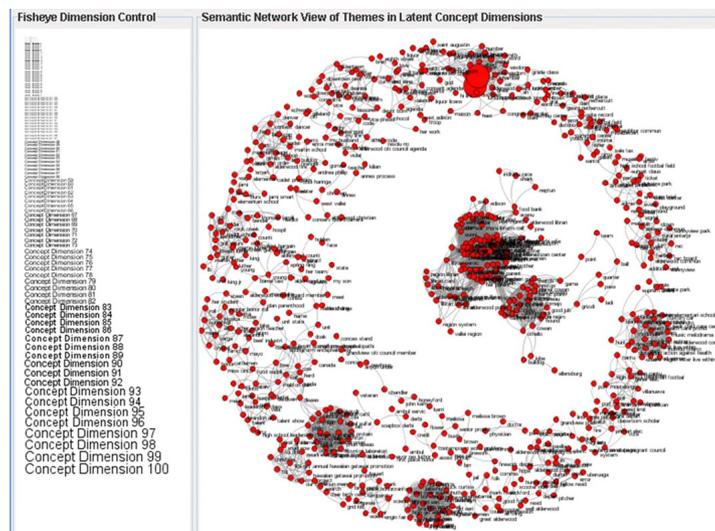


Fig. 52. Una visión general de una red semántica de términos sobre la base de las 100 dimensiones semánticas latentes en el conjunto de datos VAST más significativas. La lista de dimensiones conceptuales se filtra a través de una vista de ojo de pez para una fácil selección. Cada término está representado por un nodo rojo. El tamaño de los nodos refleja la contribución global del término a todo el espacio concepto (Zhu & Chen, 2007).

Las coordenadas paralelas son una de las técnicas de visualización más famosas, además de ser uno de los temas más comunes en los artículos académicos sobre visualización (Almeida, de M. Lourenco, Meiguins, & Meiguins, 2009; Stuart K. Card et al., 1999; Collins et al., 2009; Govaerts et al., 2012; M. Graham & Kennedy, 2003; Hauser, Ledermann, & Doleisch, 2002; Heer & Shneiderman, 2012; D. A. Keim, 2001, 2002; M. Khan & Shah Khan, 2011; Tory et al., 2005; M. O. Ward, 1994; Wegman, 1990; H. Zhou, Yuan, Qu, Cui, & Chen, 2008). Esta técnica aunque en un principio puede ser confusa, ya que no se está acostumbrado al uso de este tipo de representaciones multidimensionales, es una herramienta muy potente para la comprensión de los conjuntos de datos de más de dos dimensiones. Las coordenadas paralelas han sido ampliamente utilizadas para el análisis de conjuntos, dibujando dimensiones como ejes paralelos y datos del hiperespacio con poli líneas que conectan sus valores escalares en los ejes. Las coordenadas paralelas pueden representar los datos de N-dimensiones en un espacio de 2 dimensiones (ver Fig. 53 y Fig. 54).

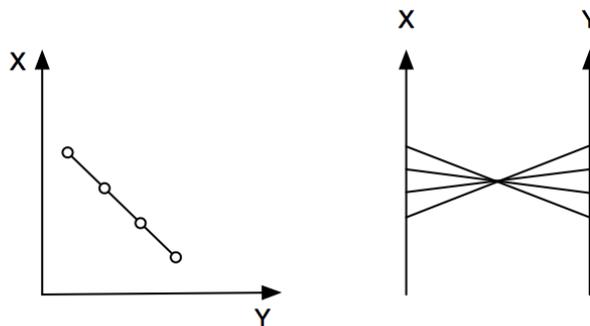


Fig. 53. La idea de las coordenadas paralelas es el uso de ejes desplegados paralelamente en lugar de ortogonalmente. Los cuatro puntos de la izquierda están representados por las cuatro líneas de la derecha (Siirtola, Laivo, Heimonen, & Raiha, 2009).

Sin embargo, cuando los tamaños de los conjuntos de datos se vuelven más grandes (millones de registros en una tabla, por ejemplo), llenan a la pantalla de poli líneas, que hacen que puedan llegar a ser demasiado densas para ser interpretadas. La forma habitual de describir las coordenadas paralelas sería hablar de espacios de alta dimensión representados con la técnica de establecer los ejes de coordenadas en paralelo en lugar de ortogonales entre sí (ver Fig. 53). El lector puede imaginarse que a cada valor de datos le corresponde una posición a lo largo de la línea (dependiendo del valor del dato, se ubicaría entre el mínimo en la parte inferior y el máximo en la parte superior); una colección pura de puntos en estas posiciones de los datos no sería extremadamente útil, por lo que los puntos que pertenecen a un mismo registro (fila) se conectan con líneas (lo que ayuda a la comparación entre registros y el descubrimiento de patrones).

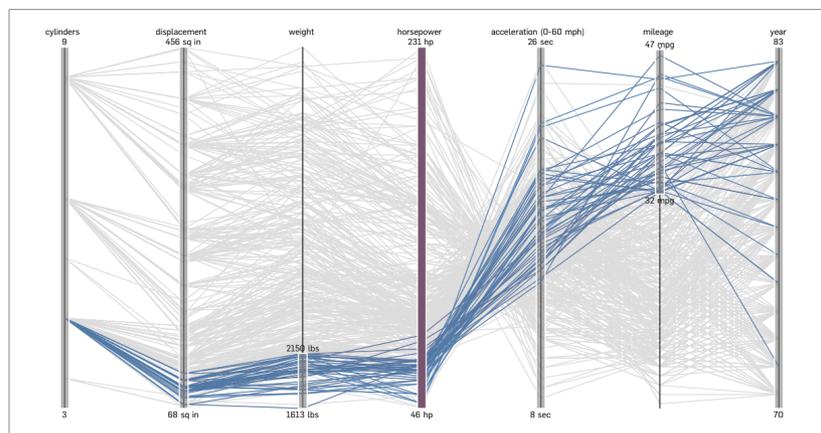


Fig. 54. Consultas de selección en coordenadas paralelas (Heer & Shneiderman, 2012).

Para dar un poco de sentido a esto, la forma más fácil es olvidar la parte de los registros que abarcan todo el ancho y mirar el espacio entre cada par de ejes. De esta forma se pueden encontrar los patrones en cada uno de los espacios entre los ejes. Además de un poco de experiencia en la lectura de coordenadas paralelas, la mejor manera de comprender un conjunto de datos es mediante el uso de las técnicas de interacción que proporcione la visualización (consultar sección IV.3.2).

La técnica principal de interacción en las coordenadas paralelas se denomina cepillado (del inglés, *brushing*) en coordenadas paralelas. Esta es una técnica muy eficaz para el proceso de selección (cepillado) de objetos en una pantalla para resaltar (u ocultar) los datos correspondientes de los otros puntos de vista enlazados, es decir especificar un enfoque explícito de la información visualizada (Becker & Cleveland, 1987; Heer & Shneiderman, 2012; M. O. Ward, 1994). En esta técnica el usuario marca activamente subconjuntos del conjunto de datos mediante el uso del ratón y la realización de selecciones de áreas de la visualización, independientemente de cuál sea la metáfora de la visualización o el diseño de esta, se dibujan, en muchos casos, como cuadros sombreados delimitadores o rangos en los ejes depende la representación.

En la Fig. 54 se puede notar en los ejes unos cuadros sombreados que delimitan los valores seleccionados por el usuario, esta técnica si se utiliza en combinación con múltiples vistas enlazadas, permite a los usuarios entender las correlaciones a través de múltiples dimensiones (Becker & Cleveland, 1987; Gresh, Rogowitz, Winslow, Scollan, & Yung, 2000; Hauser et al., 2002).

A modo de resumen, en esta primera sección se han descrito un conjunto de herramientas y/o prototipos visuales y posteriormente no visuales que soportan la mejora de la enseñanza, consecutivamente se han expuesto una variedad de técnicas de representación de información agrupadas en función del tipo de información a representar. Dado que en la investigación de campo de la visualización de información y la visualización analítica continuamente se sugieren enfoques alternativos para presentar información cuantitativa de manera más sencilla y eficiente, estas agrupaciones anteriormente expuestas se han generado tomando en cuenta las características de los datos a los que se tiene acceso en un LMS. Dado a la naturaleza no solo de los atributos de los datos sino a la complejidad del proceso de aprendizaje, además de las formas de producción de nuevas perspectivas sobre los atributos de los datos de actividad existentes en un LMS, deben investigarse aspectos más cualitativos de análisis (cognitivas y meta cognitivas).

En la siguiente sección se presentan los problemas encontrados de manera general en las propuestas y técnicas descritas en las secciones anteriores y, de manera particular, en los más cercanos al presente trabajo. Posteriormente, se expondrán las soluciones propuestas y cómo las diferentes alternativas descritas en el apartado anterior han influido en el presente trabajo de tesis o se han implementado.

## V.2 Análisis, problemas encontrados y posibles soluciones

Los entornos virtuales de aprendizaje, como espacios educativos, ofrecen una serie de posibilidades de interacción social, profesor-estudiante o entre los estudiantes, a través de diferentes canales síncronos o asíncronos. Dichos entornos también ofrecen posibilidades de interacción con diferentes materiales educativos, pero no dan una imagen o representación general de lo que sucede en estos espacios, y mucho menos una interacción con esta imagen para obtener de ella un análisis o nuevo conocimiento útil para la mejora del proceso de aprendizaje y la explotación total del entorno utilizado (o al menos una comprensión de este).

La importancia de la evaluación de la calidad de los mensajes publicados en las discusiones en línea sigue siendo un área de investigación activa en el contexto educativo. Sin embargo, aún quedan por resolver problemas asociados al análisis de esta gran cantidad de datos: la evolución temporal del contenido del curso, la interacción estudiante-tutor, las redes sociales formadas por la relación material-estudiante-tutor, las métricas de evaluación o la predicción del rendimiento de los estudiantes, entre otras cosas.

Con un proceso y un modelo de análisis eficientes, combinados con un buen pre-procesamiento de la información y técnicas de *infoVis/VA* (Ghanbari, 2007) como por ejemplo: el uso del color (Borland & Taylor, 2007), uso de diferentes perspectivas de los atributos de los datos (Pretorius & Van Wijkk, 2006), *temas* (Vliegen et al., 2006), coordenadas paralelas (H. Zhou et al., 2008), etc., proporcionarían nuevo conocimiento para retroalimentar el proceso de enseñanza y aprendizaje con el objetivo de la mejora continua. Este trabajo comparte la premisa de Bratitsis and Dimitracopoulou (2007), que considera que la investigación debe centrarse en la búsqueda de nuevas formas de análisis y visualización de datos en bruto, con el fin de proporcionar un examen más profundo de la colaboración, a favor de los usuarios involucrados y la actividad general.

Así, esta sección se centra en plantear las problemáticas encontradas en las herramientas de analítica visual y visualización de información en las representaciones visuales y técnicas de interacción (ver sección V.1). Dentro de las técnicas y visualizaciones mencionadas en la sección anterior (Las representaciones visuales: Técnicas de interacción) del presente capítulo, del Capítulo V, existen dos grupos; el primero de ellos dedicado a la representación y exploración de datos educativos o contenidos de la plataformas de aprendizaje y un segundo grupo dedicado a su análisis.

Conviene recordar en este momento que el objetivo principal de este trabajo es:

*"Definir un modelo capaz de esquematizar los pasos del análisis de la información de forma visual generada por el proceso de enseñanza y aprendizaje, que tome en cuenta los diseños y metodologías más actuales a cerca de este proceso de analítica en la educación y el proceso de analítica visual de la información."*

Se describen a continuación algunos de los problemas encontrados en las herramientas analizadas, puesto que encontrar soluciones a estos proporcionaría garantías de éxito en el cumplimiento del objetivo planteado. Primero se realiza un recorrido, desde el punto de vista de la analítica visual y la analítica del aprendizaje, por las herramientas encontradas exponiendo las limitantes y aportaciones de estas. Posteriormente, se analiza de forma general la problemática actual del análisis de los datos educativos y se presentan los problemas encontrados y algunas propuestas que los resuelven.

### V.2.1 Análisis

Se comienza este análisis con la propuesta de (Guo & Chen, 2006), que se puede observar en la Fig. 12, dado que no se trata de una herramienta de análisis como el resto de las propuestas (y el presente trabajo), sino de una representación semántica del contenido y recursos en *eLearning*. En este caso el trabajo se desarrolló a partir de *Protégés*<sup>26</sup>, una herramienta para edición y visualización de ontologías, que muestra la jerarquía de clases, las relaciones entre las clases y las propiedades de los objetos y de los datos.

---

<sup>26</sup> The Protégé homepage, <http://protege.stanford.edu/>

Dicha representación toma la forma de un grafo dirigido, el cual no incorpora mecanismos de interacción visual y en algunos casos la cantidad de datos es tan alta que presenta un grave problema de la saturación de información. Un caso análogo, como herramienta que visualiza y administra información ontológica, se puede encontrar en TM4L (Dicheva & Dichev, 2006; Dicheva et al., 2005).

Esta herramienta permite representar mapas conceptuales mediante tres vistas (*Graph View*, *Tree View* y *Text View*). Este trabajo de visualización mejora, con respecto al de (Guo & Chen, 2006), que agrega interacción visual con el usuario dotándole de manipulación de los objetos, selección de estos, posibilidad de ver información contextual y realzado de algunos elementos mediante la selección (diferenciando el objeto seleccionado con el uso del color y cepillado) y *zoom* (ver Fig. 13). Sin embargo, cuando el nivel de los datos aumenta puede dar como resultado una saturación de la pantalla.

Por último, dentro de este tipo de herramientas o visualizaciones que no analizan la información de un entorno educativo, sino que representan y permiten explorar información de este, se encuentra la visualización de P. Brusilovsky et al. (2013), que presentan una herramienta soportada por la representación de *TreeMap*, la cual permite además de la selección y de *zoom* de los objetos temáticos para el apoyo al estudio, la utilización del *zoom* semántico. Además del uso de la identificación del objeto seleccionado con el uso del color, también codifica por colores, de verde a rojo, los temas de estudio dependiendo de si el estudiante los tiene dominados o hace falta trabajar sobre ellos (ver Fig. 14 y Fig. 15). Sin embargo, al realizar el *zoom* semántico pierde el contexto de la información.

Por otro lado, dentro de las herramientas que tienen como objetivo el análisis de información educativa, se encuentra la de Romani and da Rocha (2000), que por su parte, presentan una herramienta que permite la selección del modo de representación de los atributos de los datos, dependiendo de lo que se quiera analizar. Esto es una característica importante ya que permite ver los datos desde diferentes puntos de vista; sin embargo, estas vistas no son simultáneas ni enlazadas. Las opciones de visualización de los resultados de la búsqueda son en forma de gráfico de barras o de representación en forma de red.

Estas representaciones incluyen las características mencionadas anteriormente: selección, *zoom*, *zoom* semántico y codificación por color; logrando una mayor expresividad en la herramienta con la representación de los nodos inactivos en la red y de un nodo específico que representa a todos, permitiendo identificar quiénes han enviado mensajes a todos los participantes (ver la Fig. 16). El inconveniente es que al realizar estas búsquedas se pierde la vista global de los datos (el contexto).

Así como el estudio de la interacción de los estudiantes es importante, los contextos educativos contienen mucha más información que podría representarse, como por ejemplo la relación jerárquica de temas en conjuntos grandes de textos para la visualización de la evolución de estos. Feipeng et al. (2013) presentan una herramienta que, por medio de minería de la relación jerárquica de temas y tendencia evolutiva, ofrece tres visualizaciones que, junto con funcionalidad interactiva (codificación por color, *zoom*, *zoom* semántico, cepillado, etc.), permite a los usuarios

comprender la evolución de los temas de las discusiones (ver Fig. 17). Pero esta propuesta no presenta estas vistas enlazadas, por lo que no fomenta la vista de las características de los datos desde diversos puntos de vista y por tanto se pierde su fácil comparación y análisis, además de la exploración de los datos en estas diferentes vistas simultáneamente.

Con otro objetivo, y aunque su representación carece de interacción con el usuario, Morales Morgado et al. (2009) proponen una de las pocas visualizaciones que permiten una representación sencilla e intuitiva que represente la evaluación por parte de estudiantes y profesores de los objetos de aprendizaje, propiciando la mejora continua de estos (ver Fig. 18). Esta forma de visualizar la información tiene la característica de fomentar la participación y retroalimentación sobre los objetos de aprendizaje en la plataforma por medio de la evaluación de los mismos, ayudando así a su perfeccionamiento.

Esta característica de evaluación y realimentación para la mejora también puede verse en el prototipo de análisis de narrativa de (Williams & Conlan, 2007) (ver Fig. 19), en la que ya se implementan vistas simultáneas, uso del color, mayor interacción con el usuario, una representación temporal de los patrones y una comparación con una teoría de tipos de narrativas existentes obtenidas de las encuestas de evaluación a los estudiantes. Pero carece de la interconexión de vistas simultáneas, y por tal, de la rápida comparación y exploración de las características de los datos de manera simultánea y desde diferentes puntos de vista.

El trabajo aula BRIDGE (Ganoë et al., 2003) muestra, como se ve en la Fig. 20, una visualización para la organización de proyectos por medio de sesiones de *chat* entre grupos de diversas aulas escolares. En este trabajo se utilizan las vistas simultáneas, el *zoom*, la codificación por color, la representación temporal en una línea de tiempo y una red social de los temas que se tratan; permitiendo con estas técnicas la comparación entre grupos de trabajo. En este trabajo el aumento de la cantidad de datos puede resultar en el amontonamiento de las figuras y, por tanto, en la imposibilidad de interpretación de los mismos.

Otra representación red que implementa otras técnicas es la propuesta por Q. V. Nguyen et al. (2004), que emplean técnicas de visualización para el apoyo al *eLearning* en ENCCON, representado en la Fig. 21. Utilizan rectángulos para la división de zonas, realizando agrupaciones. En su última versión (Q. Nguyen & Huang, 2010) presentan mejoras que consisten en una colección de múltiples algoritmos de diseño modificados, tales como *Enccon*, *Spring Embedder* y *Circular Drawing*, con lo que permite la combinación de múltiples algoritmos en un diseño (Huang & Nguyen, 2008). Estos algoritmos se emplean de forma automática o se seleccionan manualmente por el usuario para proporcionar el diseño óptimo para la red. En las Fig. 22 y Fig. 23 se muestra el diseño de la misma red utilizando diferentes algoritmos.

La aportación principal de este trabajo, además de la utilización de las técnicas anteriormente mencionadas, es la representación de las agrupaciones por medio de figuras geométricas pero también con las aristas dependiendo de los datos; y el inconveniente es no poder ver simultáneamente y enlazados los diferentes diseños que los algoritmos crean en la representación de la red.

Una herramienta que utiliza el color para categorizar, el *zoom*, el *zoom* semántico, la representación en una línea de tiempo y una métrica de LA (tiempo dedicado) es la propuesta de Ming et al. (2013). El aporte principal con respecto a las técnicas utilizadas es el uso del grosor en las aristas que unen los nodos en la representación temporal y la representación de la métrica del tiempo dedicado en la plataforma *online*, ya que en los trabajos anteriores esta técnica no se había utilizado para representar una característica de los datos. Por su parte, los inconvenientes, como en los anteriores trabajos, se centran en el hecho de perder el contexto al realizar un *zoom* semántico y en que no se pueda ver simultáneamente otro tipo de dato a analizar (ver Fig. 24).

Con una representación tradicional de diagrama de dispersión y con el uso del color y la forma para la caracterización de las actividades de los estudiantes, Mazza and Milani (2004) aportan como puntos fuertes de su trabajo que muestran métricas de LA, como, por ejemplo, el instante en el que los usuarios entran en la plataforma, los accesos de los estudiantes a un recurso particular y una representación de la frecuencia de la lectura y la escritura en los foros a lo largo del curso, tanto de los usuarios así como del que inicia el hilo de discusión (ver Fig. 25). Además se permite la comparación entre los estudiantes. Como principal punto negativo se tiene la imposibilidad de analizar las conexiones que se crean entre las personas involucradas en la interacción que se analiza, dejando una oportunidad importante del análisis de la red que se construye a partir de las interacciones.

Con un parecido muy alto en el método de representación (dispersión de puntos) implementado en el trabajo de Mazza and Milani (2004) y empleando, además, el uso del color en las figuras y diferentes figuras para representar las actividades en la visualización (ver Fig. 26), el estudio de Mazza and Dimitrova (2005) integra técnicas de visualización que usan la tercera dimensión y el tamaño de la figura para la representación de los atributos de los datos. Además, incorpora una representación en matriz de adyacencia de los puntos fuertes y debilidades de los estudiantes en los temas del curso. Sin embargo, comparte el mismo problema con el trabajo de Mazza and Milani (2004). Del mismo modo, y compartiendo la misma carencia de representación de otros aspectos de los datos de la interacción en los LMS, CrystalChat (Tat & Carpendale, 2006) permite la traducción de los patrones de comunicación de un individuo con las personas que forman parte de su historial de *chat* personal (ver Fig. 27). Utiliza el color y el tamaño como en los anteriores trabajos, pero también la posición para la representación de los atributos de los datos. Además incorpora la posibilidad de la auto-medición y autorreflexión por parte de los estudiantes al permitirles el acceso para observar sus patrones en las discusiones.

Como última herramienta en la que se representa la interacción de los involucrados en el proceso de aprendizaje pero no de las conexiones sociales que entre estos se forman, se menciona la de Gibbs et al. (2006), que presentan el *software* de mapeo de las relaciones temporales de los debates mediante una representación de red con nodos y aristas.

Utilizando las técnicas de cepillado, codificación por color, *zoom* semántico, etc. (ver la Fig. 28). Sin embargo, como ya se ha mencionado previamente,

estas propuestas solo analizan la interacción en los foros y en otros casos la estructura de los temas del contenido o de las discusiones, pero no el aspecto social de las conexiones que se crean en esta interacción.

Como alternativa a esta carencia surge la propuesta de De Laat et al. (2007), en la que representan cómo de equilibrada es la participación dentro de la comunidad de aprendizaje. Estos autores muestran (ver Fig. 29) y discuten los patrones participativos que cambian con el tiempo para comprender la compleja dinámica de las conexiones en el proceso de enseñanza /aprendizaje.

La propuesta de Schreurs et al. (2013), además de mostrar las conexiones entre las personas en el proceso de aprendizaje, uso del *zoom*, *zoom* semántico, codificación por colores, presenta también los temas de discusión en una misma herramienta con vistas múltiples y enlazadas, donde proponen no solo las métricas habituales de SNA, sino también los tipos de vínculo sociales entre actores y subredes sobre temas específicos de discusión (ver Fig. 30).

Menos potente que la propuesta de (Schreurs et al. (2013) es la del sistema IGRAPH (Silva & Figueira, 2012) que está basada en una representación en red, utiliza el color para codificar los participantes, el grosor de las líneas para destacar la frecuencia y el sentido de las aristas para expresar el sentido de la discusión (ver Fig. 31). Sin embargo, no representa los temas de discusión y al aumentar el volumen de los datos, dado que en su trabajo no se mencionan los métodos de ocultamiento, filtrado, *zoom* o selección, la representación dificultaría la interpretación de estos.

Finalmente, la propuesta más potente, creada por Govaerts et al. (2010) y mejorada en (2012), que presentan una herramienta denominada Medidor de Actividad del Estudiante (ver Fig. 32). Es importante mencionar que además de utilizar diferentes representaciones de los datos (representación lineal, de barras, lista horizontal con ojo de pez, nubes de palabras y coordenadas paralelas), además de estar enlazadas, uso del color para la codificación, *zoom*, el cepillado etc. Además, esta herramienta utiliza métricas de auto-reflexión, como la frecuencia en los recursos utilizados por los estudiantes y el tiempo invertido y permite la fácil comparación entre los estudiantes. Sin embargo, no se representan las conexiones sociales en las interacciones que la herramienta permite explorar.

A modo de resumen, se puede notar una diversidad en los tipos de datos a representar, la forma de representarlos y de analizarlos, por lo cual, para una mejor comprensión y comparación entre las diferentes propuestas encontradas y con el objetivo de resumir los avances que se han ido presentado en LA, AA y VA como soporte al *eLearning*, a continuación se presenta una visión global desde diferentes perspectivas, para lo cual se han realizado tablas resumen para su mejor apreciación (se ha incluido un trabajo que constituye el trabajo de tesis en las comparativas con el nombre de “prototipo de VeLA” el cual se describirá en la sección VI.2.2).

Estos criterios son: el tipo de dato a representar, las técnicas de LA utilizadas, el tipo de representación visual utilizada y las técnicas de VA utilizadas.

### V.2.1.1 Tipo de dato a representar

Reflejado en los trabajos encontrados, el tipo de dato a representar y a analizar en un ambiente de aprendizaje mixto o completamente *online* es extenso y variado. Sin embargo, no todos los datos que se encuentran almacenados en los *logs* se analizan. En este apartado se detallan cuáles se han analizado con la finalidad de encontrar alguna tendencia, carencias posibles y, por consiguiente, oportunidades de investigación. Como se puede ver en la Tabla 5, existen propuestas que analizan de forma eficiente uno o varios tipos de datos del proceso de aprendizaje *online* o mixto, como son la representación de temas en los textos extraídos (P. Brusilovsky et al., 2013; Q. Nguyen & Huang, 2010), las interacciones en las discusiones de los foros (De Laat et al., 2007; Gibbs et al., 2006; Romani & da Rocha, 2000; Silva & Figueira, 2012), el análisis de métricas de auto-medición y reflexión (Govaerts et al., 2012); en algunos casos alguna propuesta une en su análisis diferentes tipos de datos, como por ejemplo el trabajo de (Schreurs et al., 2013), que analiza los temas de los debates así como las conexiones sociales que se crean en estos.

Como se mencionó en la sección IV.4, la posibilidad de explorar la información desde diferentes puntos de vista y su representación en distintas formas por medio de técnicas de interacción es de suma importancia para su comprensión, la creación de un modelo mental y la extracción de nuevo conocimiento. Sin embargo, aunque desde hace más de una década se llevan realizando análisis de los datos, estos en su mayoría han abordado el análisis usando de forma aislada diferentes perspectivas de los datos, pero no varias simultáneamente.

Los ejemplos más comunes que se pueden encontrar en los trabajos mencionados que estudian diferentes tipos de datos, o por lo menos dos de forma simultánea, analizan la actividad temporal y las relaciones sociales (Ferguson & Buckingham Shum, 2012a; Mazza & Milani, 2004; Romani & da Rocha, 2000) o, por otro lado, la actividad temporal de los estudiantes y el análisis de texto (Boyd et al., 2002; Donath, 2002; Feipeng et al., 2013; Ganoë et al., 2003; Gibbs et al., 2006; Govaerts et al., 2012; Mazza & Dimitrova, 2005; Quang Vinh et al., 2013).

Además, en la Tabla 5 se revela la tendencia del aumento en el alcance de análisis de las herramientas, es decir, la exploración de la información desde diferentes perspectivas de los datos simultáneamente. De manera clara, ha ido aumentando el número de herramientas con la posibilidad de análisis y/o exploración de combinaciones de: análisis semántico y/o frecuencia temporal de actividades y/o análisis de texto y/o de relaciones sociales, etc. Ejemplo de esto son los trabajos de (P. Brusilovsky et al., 2013; Gibbs et al., 2006; Gretarsson et al., 2012; Guo & Chen, 2006; Q. Nguyen & Huang, 2010; Quang Vinh et al., 2013; Schreurs et al., 2013).

Referencias	Jerárquica	Actividad temporal	Ontologías	Relaciones sociales	Contenido educativo	Contenido semántico	# de tipos de dato usados
Hardless and Nulden (1999)		✓					1
Romani and da Rocha (2000)		✓		✓			2
Donath (2002)				✓		✓	2
Boyd et al. (2002)				✓		✓	2
Ganoe et al. (2003)		✓			✓		2
Q. V. Nguyen et al. (2004)					✓		2
Mazza and Milani (2004)		✓		✓			2
Dicheva et al. (2005)			✓		✓		3
Mazza and Dimitrova (2005)		✓			✓		2
Gibbs et al. (2006)		✓		✓		✓	3
Guo and Chen (2006)			✓		✓		3
Lauer (2006)					✓		1
Rossling et al. (2006)					✓		1
Tat and Carpendale (2006)		✓					1
De Laat et al. (2007)		✓		✓			2
Williams and Conlan (2007)		✓					1
Morales Morgado et al. (2009)					✓		1
Q. Nguyen and Huang (2010)				✓	✓		3
Govaerts et al. (2012)		✓			✓		2
Gretarsson et al. (2012)					✓	✓	3
Chuang et al. (2012)					✓	✓	2
Ferguson and Buckingham Shum (2012a)		✓		✓			2
Silva and Figueira (2012)				✓		✓	2
<b>Prototipo de VeLA</b>	✓	✓		✓	✓	✓	5
P. Brusilovsky et al. (2013)					✓	✓	3
Feipeng et al. (2013)		✓				✓	2
Quang Vinh et al. (2013)		✓			✓		3
Ming et al. (2013)		✓					1
Schreurs et al. (2013)				✓		✓	3
<b>Frecuencia de los tipos de dato</b>	<b>9</b>	<b>15</b>	<b>2</b>	<b>11</b>	<b>15</b>	<b>10</b>	

Tabla 5. Tabla comparativa de los tipos de datos educativos que se utilizan en las herramientas visuales que soportan la mejora y/o la comprensión del proceso de aprendizaje.

De forma general se puede concluir que las nuevas herramientas deben y están en el camino de incluir diferentes puntos de análisis de diversos tipos de datos simultáneamente para proporcionar un análisis visual exploratorio más robusto. Esto conlleva, como se mencionó en párrafos anteriores, el uso de técnicas de interacción y de métricas de análisis para la consecución de este objetivo. Así, en el siguiente apartado se detallan las diferentes técnicas de LA que se han encontrado en las herramientas estudiadas y descritas hasta ahora, con el objetivo de identificar posibles tendencias, carencias y, por tanto, oportunidades de desarrollo de posibles soluciones.

### V.2.1.2 Técnicas de LA utilizadas

Como se mencionó en el Capítulo III, la analítica en la educación es:

*La aplicación de técnicas analíticas para el análisis de datos educativos, incluyendo datos sobre actividades para el estudiante y el profesor, para identificar patrones de comportamiento y proporcionar información útil para mejorar el aprendizaje y las actividades relacionadas con el aprendizaje (Van Harmelen and Workman (2012)).*

Dado que existen diversas áreas (EDM y SLA), como se mencionó en las secciones III.3, III.4, III.5 y III.6, a las que se les considera sub-áreas dentro de esta, las técnicas existentes de analítica del aprendizaje son muy diversas y en constante evolución.

La forma más concreta de compararlas es ir reconociéndolas en los trabajos relacionados descritos anteriormente y que mencionan cómo se aplica la analítica en el ámbito educativo, proporcionando avances mensurables para el aprendizaje y la enseñanza, y ofreciendo la esperanza de la más conveniente toma de decisiones basadas en la evidencia, la acción y la personalización en diversas áreas de la educación.

En la tabla comparativa de las técnicas de analítica educativa que se utilizan en las herramientas visuales que soportan la mejora y/o la comprensión del proceso de aprendizaje (ver Tabla 6), se resalta que desde el inicio ha existido el interés por analizar las actividades de los estudiantes, utilizando diversas técnicas de análisis, aunque en un principio no reconocidas como técnicas de analítica del aprendizaje. En esta tesis doctoral, dada su metodología e interés, se han categorizado como tales; dentro de estas algunas se han utilizado desde hace ya más de una década con relativa constancia, como por ejemplo la búsqueda de patrones temporales, el análisis de la frecuencia de escritura en las discusiones, la exploración semántica y el análisis del rol social que ejerce cada participante en el proceso de aprendizaje/enseñanza.

Sin embargo, el número de técnicas de LA, al igual que el tipo de dato a analizar por parte de las herramientas actuales, han ido en aumento. Esto no es algo fortuito, dado que a mayor diversidad de datos, lo óptimo es que exista mayor diversidad de técnicas usadas para el análisis de los atributos de los mismos. Si se revisa la Tabla 6 con detenimiento, se puede notar un aumento en el uso simultáneo de diversas técnicas de LA, proporcionando así mayor robustez.

Reference	Temporal Pattern	Semantic Explore	Self-Reflection, Self-monitoring, Self-assessment					Social pattern			Type of # técnicas					
			Login : Score	Starter discussion	Time Spent	Read content	Write Post	Grouping	Any Activity	Centrality		Degree	Density			
Hardless and Nulden (1999)			✓				✓									2
Romani and da Rocha (2000)			✓				✓									6
Donath (2002)							✓									4
Boyd et al. (2002)				✓			✓									4
Graoe et al. (2003)		✓														2
Q. V. Nguyen et al. (2004)		✓														1
Mazza and Milani (2004)																1
Dicheva et al. (2005)		✓														1
Mazza and Dimitrova (2005)				✓												3
Gibbs et al. (2006)								✓								4
Guo and Chen (2006)		✓														1
Lauer (2006)		✓														1
Rosling et al. (2006)		✓														1
Tat and Carpendale (2006)								✓								1
De Laat et al. (2007)																3
Williams and Conlan (2007)								✓								4
Morales Morgado et al. (2009)		✓														5
Q. Nguyen and Huang (2010)		✓														2
Govaerts et al. (2012)																6
Gretarsson et al. (2012)		✓														6
Chuang et al. (2012)		✓														4
Ferguson and Buckingham Shum (2012a)																2
Silva and Figueira (2012)		✓														7
Prototipo de VeLA	✓	✓														6
P. Brusilovsky et al. (2013)		✓														11
Feipeng et al. (2013)		✓														2
Quang Vinh et al. (2013)		✓														4
Ming et al. (2013)		✓														2
Schreurs et al. (2013)		✓														4
<b>Frecuencia de las técnicas</b>	<b>15</b>	<b>16</b>	<b>5</b>	<b>4</b>	<b>5</b>	<b>3</b>	<b>6</b>	<b>15</b>	<b>4</b>	<b>1</b>	<b>11</b>	<b>9</b>	<b>10</b>	<b>1</b>		<b>6</b>

Tabla 6. Tabla comparativa de las técnicas de analítica educativa que se utilizan en las herramientas visuales que soportan la mejora y/o la comprensión del proceso de aprendizaje.

Por otro lado, existen técnicas que también se desarrollan sin una representación visual en las cuales comúnmente se utilizan regresiones estadísticas (Che-Cheng & Chiung-Hui, 2013; Macfadyen & Dawson, 2010; Romero-Zaldivar et al., 2012), análisis de componentes principales, análisis estadístico cuantitativo (Davies & Graff, 2005; Giovannella et al., 2013; Kim, 2013), análisis de redes sociales (Lipponen et al., 2003; Medeiros et al., 2013) y/o minería de datos (Weisskirch & Milburn, 2003). Estos trabajos se han descrito en la sección V.1.2 del presente capítulo. Sin embargo, para una mayor comprensión y comparación se presenta la Tabla 7. En la Tabla 7 se puede observar que, desde hace más de una década, se vienen realizando estudios con la intención de predecir y mejorar el aprendizaje en línea (sin una representación visual o procesamiento automático) mediante el análisis de los datos almacenados en una plataforma de aprendizaje *online*. Además se resalta el hecho que el acceso de lectura y escritura a las discusiones son las métricas de LA más utilizadas, seguidas en menor grado por la distinción del iniciador del hilo de la discusión y los grados de participación de recepción y envío en una red de conexiones sociales.

Asimismo, existen comprobaciones de la existencia de una correlación entre estas métricas con el rendimiento de los estudiantes (marcadas en la Tabla 7 con un ★). Es importante resaltar que la métrica más utilizada, la escritura de mensajes en los foros de discusión, coincide con la más corroborada como indicador de predicción de rendimiento, seguida por la lectura de las discusiones y el análisis del contenido de los mensajes.

Se debe aclarar que en la Tabla 7 existe una columna en la que se indica una categorización dinámica de las actividades, pero no todos los trabajos categorizan todas las actividades (marcados con un \* al inicio de la fila), sino solo las actividades estudiadas por el trabajo (representado por cada fila) y que aparecen en la tabla (como otra columna marcada con ✓); otros que categorizan un número mayor de 30 actividades (marcados con \*\* al inicio de la fila) que no aparecen en la tabla por razones de espacio, pero sí aparecen aquellas que resultaron importantes para el estudio realizado.

Para continuar con el análisis es importante recordar lo que se mencionó en la sección IV.4, la analítica visual integra disciplinas científicas para mejorar la división del trabajo entre el hombre y la máquina; combina el razonamiento analítico (que permite a los usuarios obtener una visión profunda de los datos, para soportar directamente la evaluación, la planificación y la toma de decisiones) con visualización interactivas. Por tanto, en el aparato V.2.1.3 se analizan las diferentes visualizaciones implementadas y/o diseñadas en los trabajos mencionados.

### V.2.1.3 Tipo de representación visual

Por un lado, los tipos de datos más analizados, como se mencionó en el apartado V.2.1.1 (ver Tabla 5), son la frecuencia temporal de las actividades, el análisis de textos extraídos (de contenido *eLearning* y de discusiones den los LSM) y las relaciones sociales que se crean con las interacciones en las plataformas de aprendizaje *online*.

Referencias	Diversidad de escritura en discusiones	Acceso de lectura a discusiones	Acceso de lectura a recursos	Acceso de escritura a discusiones	Iniciador de la discusión	Centralización	densidad	grados	Análisis del contenido de los <i>posts</i>	Tareas completadas	Inicio de sesión	Tiempo en línea total	Categorías de actividades
(Lipponen et al., 2003)	✓			✓	✓	✓	✓	✓					
* (Weisskirch & Milburn, 2003)				✓★					✓★				✓★
* (Davies & Graff, 2005)	✓			✓				✓	✓				✓
(Hwang & Arbaugh, 2009)	★			✓									
(Macfadyen & Dawson, 2010)	✓	✓	✓★			✓	✓	✓		✓★		✓	
(Romero-Zaldivar et al., 2012)	✓	✓	✓★									✓	
** Prototipo de VeLA	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
* (Kim, 2013)	✓			✓					✓★				✓
(Che-Cheng & Chiung-Hui, 2013)	✓★	✓	✓★	✓							✓★		
(Medeiros et al., 2013)						✓	✓	✓★					
@ (Giovannella et al., 2013)	✓★		✓★	✓★									
** (Agudo-Peregrina et al., 2014)	✓	✓	✓	✓	✓					✓	✓		✓★

Tabla 7. Tabla comparativa de las técnicas de LA que se encontraron en los trabajos no visuales indicadas con ✓ e indicando con doble símbolo (✓★) las que se comprobó que existe una relación con el rendimiento del estudiante. Los trabajos marcados con un asterisco \* al inicio de la fila, se refieren a que han categorizado las actividades pero solo aquellas, indicadas en la tabla, que han estudiado. En el trabajo de (Giovannella et al., 2013), marcado al inicio de la fila con @, es de relevancia mencionar que aunque se marcan estas técnicas de búsqueda de correlación, en este trabajo se analizan no solo las del LMS sino también las de otras herramientas sociales como *twitter* y *Facebook*. La caracterización de las actividades se realiza en un número mayor a 30 actividades categorizando les desde diferentes enfoques en el trabajo de (Agudo-Peregrina et al., 2014), marcado con \*\* en el inicio de la fila.

Por otro lado, también se conocen las métricas más utilizadas como técnicas de LA (ver la Tabla 6 y la Tabla 7) son, en relación lógica con los tipos de datos analizados más comúnmente, las de búsqueda de patrones temporales de la actividad y escritura en los foros de discusión, seguidas por el análisis de contenidos semánticos, la centralidad y el grado de participación en las estructuras sociales creadas por la participación en los LSM. Es de esperarse, por tanto, que las representaciones visuales más utilizadas estén relacionadas con visualizaciones temporales, de redes sociales y de análisis de textos.

La Tabla 8 muestra los trabajos que se han analizado en la bibliografía (y se han descrito anteriormente), pero esta vez se comparan con respecto al tipo de representación visual que utilizan. De esta manera se puede resaltar el hecho, que ya se intuía, de que las representaciones de línea de tiempo y de redes sociales han sido las más empleadas; y, en segundo lugar según frecuencia de uso, están las nubes de palabras, las gráficas lineales (gráficas que utilizan las coordenadas cartesianas y líneas) y los diagramas de dispersión de puntos.

Sin embargo, el patrón mostrado hasta ahora en los anteriores criterios (secciones V.2.1.1 y V.2.1.2), a saber, el aumento de tipos de dato a analizar y el aumento de técnicas utilizadas simultáneamente, no se manifiesta de forma clara para este criterio, es decir, no se pudo concluir que exista un aumento en el uso de múltiples representaciones visuales en las herramientas. Esto es a razón de que los trabajos de (Ferguson & Buckingham Shum, 2012a; Govaerts et al., 2012; Mazza & Dimitrova, 2005; Williams & Conlan, 2007), que son los que utilizan una mayor cantidad de técnicas, están distribuidos en el período de tiempo completo (del más antiguo en el 2005 al más nuevo en el 2012), y además, el resto de trabajos han utilizado dos representaciones diferentes en sus herramientas y como máximo tres durante todo el período de tiempo comprendido. Es decir, desde el principio se han utilizado en muchos casos dos diferentes formas de visualizar los datos y como máximo cuatro.

Otro aspecto importante que se puede extraer de la Tabla 8 es que el uso de representaciones abstractas o complejas, como las coordenadas paralelas, la representación en espiral, la gráfica de radar, las deformaciones y las metáforas en las representaciones y las representaciones en tercera dimensión de los datos no se han utilizado con frecuencia en el área de análisis de datos educativos, Esto se debe posiblemente a la falta de alfabetización de la información y/o alfabetización visual mencionada en la sección IV.6. Sin embargo, se las debe tomar en cuenta para futuros diseños de herramientas, dada su gran expresividad y potencial para explorar y revelar patrones en los datos.

Las visualizaciones interactivas recogidas en la Tabla 8 son representaciones visuales que, junto a las técnicas de interacción, de visualización analítica y de visualización, explotan el ancho de banda del ojo humano para ampliar la cognición y permitir a los usuarios ver, explorar y entender grandes cantidades de información de forma simultánea. Por tal motivo, sin estas técnicas de visualización y de interacción de VA no se podría lograr tal objetivo. Teniendo esto en cuenta, en la siguiente sección se analiza el uso de estas técnicas en los trabajos revisados en las anteriores secciones.

Referencias	Word Cloud	Lifelines, timelines	Bar Graph	Tree Map	Other	Network Graph	Line Graph	Tree	Scatterplot	Radar Chart	Spiral	Parallel Coordinates	# de técnicas usadas
Hardless and Nulden (1999)	✓												1
Romani and da Rocha (2000)	✓					✓			✓				3
Donath (2002)					People Garden								1
Boyd et al. (2002)					Text & shapes	✓			✓				4
Ganoe et al. (2003)	✓					✓							2
Q. V. Nguyen et al. (2004)						✓							1
Mazza and Milani (2004)	✓		✓			✓							3
Dicheva et al. (2005)	✓					✓							1
Mazza and Dimitrova (2005)	✓		✓		Matrix	✓			3D				4
Gibbs et al. (2006)	✓					✓							2
Guo and Chen (2006)						✓							1
Lauer (2006)								✓					1
Rosling et al. (2006)						✓							1
Tat and Carpendale (2006)	✓		✓										3
De Laat et al. (2007)					CrystalChat	✓							1
Williams and Conlan (2007)	✓					✓							3
Morales Morgado et al. (2009)					Score Stars								1
Q. Nguyen and Huang (2010)						✓							1
Govaerts et al. (2012)	✓		✓			✓						✓	4
Gretarsson et al. (2012)						✓							2
Chuang et al. (2012)													1
Ferguson and Buckingham Shum (2012a)						✓				✓			4
Silva and Figueira (2012)	✓					✓							1
<b>Prototipo de VeLA</b>	✓	✓	✓		<b>Deformation</b>	✓					✓	✓	6
P. Brusilovsky et al. (2013)				✓									1
Feipeng et al. (2013)	✓					✓							2
Quang Vinh et al. (2013)									✓				1
Ming et al. (2013)	✓					✓							2
Schreurs et al. (2013)						✓							2
<b>Frecuencia de las técnicas</b>	4	13	5	1	6	15	6	1	6	1	1	2	2

Tabla 8. Tabla comparativa de las representaciones que se utilizan en las herramientas visuales que soportan la mejora y/o la comprensión del proceso de aprendizaje.

#### V.2.1.4 Técnicas de VA utilizadas

Además de examinar la representación de los diferentes tipos de datos, las formas de representarlos, las métricas y técnicas de LA, los modos de interacción con el usuario (como parte de las técnicas de analítica visual), también han sido objeto de estudio. Dos ejemplos sencillos de este tipo de enfoque son el trabajo de Koike, Sugiura, and Koseki (1997), en el que propusieron un control deslizante eficiente y una ventana temporal que sirve como un filtro, y el trabajo de Willett et al. (2007), en el que presentaron unos *widgets*, controles de la interfaz gráfica de usuario, que al usarlos de forma integrada en las visualizaciones, se consigue mejorar y facilitar la navegación en los espacios informativos.

Así, en la Tabla 9 se presenta una comparativa de las técnicas de analítica visual que se emplean en las herramientas que soportan la mejora y/o la comprensión del proceso de aprendizaje mediante representaciones visuales. En esta tabla se puede distinguir un aumento en el uso simultáneo de las técnicas de VA, o sea, al igual que en el uso de las técnicas de LA (ver Tabla 7), los diferentes trabajos comparados ahora en la tabla han utilizado y/o implementado cada vez un mayor número de diversas técnicas de visualización analítica a través del tiempo. Sin embargo existen muchas técnicas de VA de enorme potencial aún sin utilizarse.

Otro aspecto importante que se puede resaltar en la Tabla 9 es el hecho de que las técnicas más utilizadas hasta el momento son: la posibilidad de obtener desde una vista general los detalles de los elementos representados (*overview + detail*), el *zoom* gráfico y el uso de la codificación por colores para la representación de los atributos de los datos. Asimismo, pero en menor medida, la utilización de metáforas para la representación de los atributos de los datos, las búsquedas dinámicas (*dynamic queries*), la posibilidad de observar en la misma ventana de la herramienta los elementos seleccionados (*foco*) y el resto de los datos (*contexto*) en el mismo momento y con la posibilidad de interactuar con ellos (*foco + contexto*) y la creación de agrupaciones visuales de los datos por sus características intrínsecas.

Las agrupaciones en los datos pueden estar ya dadas por el tipo de dato en sí (como por ejemplo los datos que contienen jerarquías), pero también existe la posibilidad de crear agrupaciones por medio de algoritmos específicos. Este es el caso de las agrupaciones por medio de algoritmos de minería de datos o representaciones de métricas de análisis, los cuales buscan entre los atributos de los datos características en común o conjuntos de valores de las métricas, agrupando los datos en consecuencia. Este tipo de agrupaciones se expresa en la columna de agrupación (*cluster*). Esta técnica junto con las vistas enlazadas (*interlinked views*) son las menos utilizadas en los trabajos de análisis de información educativa, a pesar de ser de las más recomendadas por el área de VA.

Referencias	Zoom	Vista general + detalle	Agrupación	Agrupación visual	Consultas dinámicas	Foco+ contexto	Vistas Enlazadas	Zoom semántico	Metáforas	Codificación por color	# de técnicas usadas
Hardless and Nulden (1999)		✓				✓				✓	3
Romani and da Rocha (2000)		✓			✓	✓				✓	5
Donath (2002)				✓					✓	✓	4
Boyd et al. (2002)		✓		✓		✓			✓	✓	6
Ganoe et al. (2003)		✓				✓					3
Q. V. Nguyen et al. (2004)		✓			✓	✓					4
Mazza and Milani (2004)		✓								✓	3
Dicheva et al. (2005)		✓			✓					✓	4
Mazza and Dimitrova (2005)		✓			✓				✓	✓	4
Gibbs et al. (2006)		✓		✓	✓	✓				✓	6
Guo and Chen (2006)								✓			1
Lauer (2006)		✓							✓	✓	3
Rossling et al. (2006)		✓				✓			✓	✓	5
Tat and Carpendale (2006)		✓									2
De Laat et al. (2007)							✓				2
Williams and Conlan (2007)		✓			✓	✓					3
Morales Morgado et al. (2009)									✓	✓	2
Q. Nguyen and Huang (2010)			✓	✓	✓			✓	✓	✓	7
Govaerts et al. (2012)		✓		✓	✓	✓	✓	✓	✓	✓	9
Gretarsson et al. (2012)		✓	✓	✓	✓			✓		✓	7
Chuang et al. (2012)		✓								✓	3
Ferguson and Buckingham Shum (2012a)		✓									2
Silva and Figueira (2012)				✓					✓	✓	3
<b>Prototipo de VeLA</b>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	10
P. Brusilovsky et al. (2013)		✓			✓	✓		✓		✓	6
Feipeng et al. (2013)		✓			✓	✓		✓		✓	6
Quang Vinh et al. (2013)		✓		✓	✓	✓	✓		✓	✓	8
Ming et al. (2013)		✓		✓	✓	✓		✓	✓	✓	7
Schreurs et al. (2013)		✓		✓	✓	✓	✓	✓	✓	✓	9
<b>Frecuencia de la técnica:</b>	<b>22</b>	<b>23</b>	<b>3</b>	<b>11</b>	<b>15</b>	<b>14</b>	<b>5</b>	<b>9</b>	<b>13</b>	<b>22</b>	

Tabla 9. Tabla comparativa de las técnicas de analítica visual que se utilizan en las herramientas que soportan la mejora y/o la comprensión del proceso de aprendizaje por medio de representaciones visuales.

Finalmente, se pueden concluir ciertas normas que cumplen tanto la teoría del Capítulo III y del Capítulo IV y lo encontrado en los trabajos analizados en esta tesis doctoral (ver Tabla 10):

1. Si se tienen escasos tipos de datos a analizar se puede permitir tener pocas representaciones visuales y, por tanto, utilizar escasas técnicas de LA y VA.
2. Sin embargo, aunque sea poca la variedad, en el caso de datos complejos, como las ontologías, se recomienda el uso ya sea de múltiples representaciones o un alto despliegue de técnicas de LA y/o VA.

3. Por el contrario, si se tiene una variedad en el tipo de dato a analizar, comúnmente se utilizan diversas formas visuales para representar los datos simultáneos y por ende, múltiples técnicas de LA y VA.
4. Pero si se tienen múltiples tipos de datos a analizar pero se tienen pocas representaciones visuales, se deben utilizar variadas técnicas ya sea de LA como de VA o en el mejor de los casos de ambos.

Después de haber descrito y analizado de manera global los trabajos del estado actual de la investigación del área de la visualización analítica en el *eLearning*, y ya habiendo encontrado algunas pautas y tendencias, en la siguiente sección se describen los problemas encontrados y las soluciones posibles.

Referencias	Tipo de conexión a los datos	Tipos de dato	Visualizaciones	Técnicas de LA	Técnicas de VA
Hardless and Nulden (1999)		1	1	2	3
Romani and da Rocha (2000)	Database	2	3	6	5
Donath (2002)		2	1	4	4
Boyd et al. (2002)		2	3	4	6
Ganoe et al. (2003)	Database	2	2	2	3
Q. V. Nguyen et al. (2004)	Database	2	1	1	4
Mazza and Milani (2004)	Database	2	3	1	3
Dicheva et al. (2005)	TM4J	3	1	1	4
Mazza and Dimitrova (2005)	XML	2	3	3	4
Gibbs et al. (2006)	XML	3	2	4	6
Guo and Chen (2006)	OWL/RDF	3	1	1	1
Lauer (2006)	XML	1	1	1	3
Rossling et al. (2006)	XML	1	1	1	5
Tat and Carpendale (2006)	XML	1	2	3	2
De Laat et al. (2007)	WebCT log-files	2	1	4	2
Williams and Conlan (2007)	Database	1	3	5	3
Morales Morgado et al. (2009)	Database	1	1	2	2
Q. Nguyen and Huang (2010)	XML	3	1	6	7
Govaerts et al. (2012)	attention-XML	2	4	6	9
Gretarsson et al. (2012)	ND/TopicNets	3	2	4	7
Chuang et al. (2012)		2	1	2	3
Ferguson and Buckingham Shum (2012a)	WS- SocialLearn	2	4	7	2
Silva and Figueira (2012)	Database	2	1	6	3
<b>Prototipo de VeLA</b>	<b>WS-Moodle</b>	<b>5</b>	<b>6</b>	<b>11</b>	<b>10</b>
P. Brusilovsky et al. (2013)	OWL/database	3	1	2	6
Feipeng et al. (2013)	Text-web crawler	2	2	4	6
Quang Vinh et al. (2013)	FCM/TXT file	3	1	2	8
Ming et al. (2013)	Google Doc API	1	2	4	7
Schreurs et al. (2013)	NAT plugin-SocialLearn	3	2	6	9

Tabla 10. Tabla comparativa de la forma de obtención de los datos en las herramientas y frecuencia simultánea de: tipo de datos, representación visual, número de técnicas de analítica del aprendizaje implementadas y por último, el número de técnicas de visualización analítica implementadas en las herramientas. Coloreadas desde el color rojo para un valor menor pasando por el amarillo hasta el verde para un valor mayor con el objetivo de ayudar a su lectura.

## V.2.2 Problemas y soluciones posibles

En las herramientas que usan las técnicas de representación visual, analítica académica y visualización analítica para el análisis de datos educativos que se han comparado anteriormente se ha encontrado una serie de problemas, ya sea por la ausencia o escasa presencia de algunas características deseables: desde una perspectiva más de herramienta en general y no desde el diseño de interfaces de VA, LA o AA, la independencia de las herramientas actuales con respecto a la conexión con las fuentes de datos y las plataformas de aprendizaje es una carencia notable, ya que supone un contexto de uso excesivamente rígido para profesores y gestores académicos. Además, existe un controvertido problema comentado en (Hardless & Nulden, 1999) y en el modelo de referencia para la analítica del aprendizaje de Chatti (Chatti et al., 2012), que no es otro que la violación de la integridad personal: aunque en la mayoría de los trabajos de investigación se obtiene un permiso por parte de los involucrados, esto no significa que en la vida real sea siempre posible obtener dicho permiso, por lo que es un punto a tomar en cuenta para el diseño de herramientas de análisis de datos educativos. Esta circunstancia, aunque ajena a la analítica de los datos educativos, influye a la eficiencia del análisis.

Por otra parte, y más enfocado a las técnicas de LA, AA, y VA, en *eLearning* se ha constatado poca expresividad visual, escasa personalización de la exploración, saturación de elementos visuales y, por último, falta de mecanismos de interacción. A continuación se explican cada uno de estos factores y se presentan algunas soluciones posibles.

### V.2.2.1 Flexibilidad e independencia de conexión con las fuentes de datos y respeto a la privacidad e integridad de los datos personales

Como se mencionó con anterioridad en el modelo de referencia de las analíticas para la educación en la sección III.3.2, la ética, la privacidad de los datos y la violación de la integridad personal (Hardless & Nulden, 1999), son algunas de las restricciones que deben tenerse en cuenta a la hora de realizar sistemas de análisis de datos. Por otro lado, en su mayoría, las herramientas descritas y comparadas en las tablas de la sección V.2.1 están vinculadas y/o embebidas en las plataformas, por tanto: son dependientes de estas, afectando a su portabilidad además de no permitir la obtención de datos o el intercambio de estos con otros entornos de aprendizaje, funcionar únicamente *online* y/o no permitir la obtención de datos de varias plataformas simultáneamente para su comparación.

En los trabajos analizados se encontraron diversas formas de obtención de los datos. Si se observa la segunda columna de la Tabla 10, partiendo desde arriba hacia abajo, de forma global se puede descubrir una tendencia: desde una dependencia directa del sistema o plataforma al conectarse a través de bases de datos (Ganoë et al., 2003; Mazza & Milani, 2004; Morales Morgado et al., 2009; Q. V. Nguyen et al., 2004; Romani & da Rocha, 2000; Silva & Figueira, 2012; Williams & Conlan, 2007), continuando con la utilización de estándares propios, como son, *attention-XML* por parte de Govaerts et al. (2012), *TopicNets* de Gretarsson et al. (2012), archivos FCM en Quang Vinh et al. (2013), Google Doc API de

Ming et al. (2013), entre otros (Quang Vinh et al., 2013; Schreurs et al., 2013) o estándares más genéricos como XML (Gibbs et al., 2006; Lauer, 2006; Mazza & Dimitrova, 2005; Q. Nguyen & Huang, 2010; Rosling et al., 2006; Tat & Carpendale, 2006), estándares de ontologías (P. Brusilovsky et al., 2013; Guo & Chen, 2006), de redes (Dicheva et al., 2005; Gretarsson et al., 2012) y finalizando con la utilización de servicios web (Ferguson & Buckingham Shum, 2012a) o el mecanismo más genérico que es la obtención de textos directamente por medio de un *web crawler* (Feipeng et al., 2013). Sin embargo, en ningún trabajo se encontró explicación o indicación de haber tomado en cuenta la restricción de la privacidad.

Por otra parte, el estar ligado a un estándar personalizado, es decir, un estándar creado específicamente para el desarrollo de la herramienta y por tanto también específico de una versión de esta o de la plataforma de donde se obtengan los datos, o a una conexión directa a la base de datos presentan la problemática de que las herramientas son dependientes de estas plataformas o estructuras de base de datos, no permiten la obtención de datos o el intercambio de estos con otros entornos de aprendizaje, trabajan únicamente *online* y no permiten la obtención de datos de varias plataformas simultáneamente para su comparación.

No obstante y como alternativa a esta situación, trabajos como (Ferguson & Buckingham Shum, 2012a; Rosling et al., 2006) usan técnicas de recuperación de la información basadas en servicios web. En estos hay una dependencia de las capacidades del perfil del usuario que la plataforma de origen asigne, es decir, la cantidad y el tipo de dato que se obtendrá depende directamente de los permisos asignados al usuario con el que se utiliza el servicio web, siendo estos permisos otorgados por el administrador académico o persona responsable de la plataforma educativa y en caso de no estar implementado el servicio web necesario para la obtención de algún dato o la infraestructura para ello, su realización es más sencilla, flexible y adaptable que en los otros casos. El problema de la dependencia de la estructura de la base de datos, estándar de archivo exportado o plataforma, como se mencionó anteriormente en los anteriores trabajos, no puede existir en el para estos modelos, ya que el usuario podrá analizar solamente aquello que sus privilegios como usuario de la plataforma permitan y su implementación es independiente de la herramienta de análisis y, por ende, su adaptación en las siguientes versiones es más rápida de realizar sin afectar el funcionamiento del sistema.

#### V.2.2.2 Poca expresividad visual

La expresividad visual se define como el número de variables visuales utilizadas para hacer más eficiente una visualización. Como se ha descrito en la sección V.2.1, existe una tendencia en la utilización de más tipos de datos en las herramientas, pero aún quedan espacios de investigación por cubrir. Además, una de estas variables es la distribución espacial de los elementos en el plano, un aspecto fundamental en una visualización. La distribución física de los elementos se utiliza también para proporcionar información visual.

El propósito más común de una distribución espacial de elementos es establecer una relación de orden entre estos, como se ha ilustrado en la Fig. 29, donde la distribución de los elementos se ve afectada a través del tiempo, expresando la evolución de las conexiones creadas por la interacción de los participantes. Sin embargo, una distribución espacial eficiente puede proveer de mucha más información, como por ejemplo, relaciones de propiedad, de dependencia o similitudes entre elementos. Es justo en este aspecto en el que el uso de una distribución espacial de los elementos puede mejorar considerablemente las visualizaciones.

Una solución a esto es la implementación de métricas que se expresen a través de la distribución espacial de los elementos algún atributo de los datos u otro tipo de dato de la plataforma educativa a analizar (por ejemplo las figuras Fig. 22 y Fig. 23, donde se realizan agrupaciones en los nodos dependiendo de la semejanza entre estos). Además, el uso del color, el tamaño, la forma o la implementación de alguna otra representación visual da solución a esta problemática.

### V.2.2.3 Manejo adecuado de los planos visuales y los contextos

Otro problema con las herramientas visuales es la falta de un manejo adecuado de los planos visuales y los contextos. Este problema se ilustra en la Fig. 17 con el prototipo de (Feipeng et al., 2013), en ella aparece un recuadro que indica datos (*id: 1, time: 2014, number: 724*) de algún elemento seleccionado, sin embargo, la representación de las relaciones de este con otros elementos no resulta lo suficientemente clara y sencilla, ya que no se destaca qué elemento o parte del elemento está seleccionado o resaltado, ya sea la columna de 2014, o la forma en color celeste, azul o algún otra. En este caso se podría resolver identificando con un sombreado o aclarado del color del elemento seleccionado, o en el caso de ser parte de un elemento seleccionado se podría dibujar una línea, con puntos o círculos sombreados resaltando los promedios y máximos y mínimos de los datos de la representación.

En el caso de las redes, las relaciones se muestran como aristas que unen los propios elementos relacionados. Para el caso de la selección de un elemento, las aristas involucradas se vuelven más anchas para resaltarlas del resto. Sin embargo, esta solución no es suficiente para que realmente el usuario pueda percibir de forma efectiva cuál es el elemento seleccionado, las relaciones de este elemento y los otros elementos relacionados. Esto se debe principalmente al hecho de que no hay un buen manejo de los focos, los contextos y los planos visuales, ver las figuras: Fig. 13, Fig. 16, Fig. 22 y Fig. 23 de los trabajos de (Dicheva et al., 2005; Q. V. Nguyen et al., 2004; Romani & da Rocha, 2000).

Una solución a este problema es el uso de otras técnicas visuales como son el uso de varios niveles de opacidad y/o la técnica de cepillado: mediante el difuminado de un conjunto de elementos se puede mantener el contexto global del problema, permitiendo así identificar elementos, tanto seleccionados como los que no lo están, y resaltar mucho más los elementos que se pretenden mostrar en un primer plano (el elemento seleccionado, sus relaciones y los elementos relacionados con este). Esta solución está relacionada con la técnica de Foco + Contexto, una de las técnicas de interacción definidas en el área de la Analítica Visual.

Otra solución a este mismo problema de la selección de los elementos visuales, es el uso de la técnica de *zoom* semántico. Esta técnica se utiliza justamente en los casos en los que se quiere seleccionar alguno de los elementos y destacar la información relevante a este, como pueden ser sus relaciones y los elementos relacionados. Sin embargo, esta opción puede provocar la pérdida del contexto de navegación, problema descrito a continuación.

#### V.2.2.4 Pérdida del contexto de navegación

Este es un problema que comparten la gran mayoría de las herramientas que tratan de representar taxonomías de conceptos en grandes cantidades de datos, como es el caso del contenido educativo en las plataformas de aprendizaje o su estructura. Esto se debe a que suelen usar una solución basada en listas, en redes o en árboles (ver las figuras: Fig. 20 (Ganoe et al., 2003), Fig. 21 (Q. V. Nguyen et al., 2004), Fig. 24 (Ming et al., 2013), Fig. 27 (Tat & Carpendale, 2006), Fig. 28 (Gibbs et al., 2006) y Fig. 30 (Schreurs et al., 2013)), siendo esta una solución poco eficiente desde el punto de vista de las visualizaciones. Básicamente, la principal solución radica en el uso de técnicas de visualización especializadas en la representación de grandes taxonomías de conceptos, como pueden ser los *TreeMap* y los árboles de jerarquías.

Entre las ventajas en el uso de estas técnicas se puede destacar que aprovechan de una manera eficiente el espacio visual disponible y que la distribución espacial (*layout*) de los elementos resulta bastante intuitiva. Sin embargo, incluso estas visualizaciones pueden llegar a resultar insuficientes, por lo que es necesario proporcionar mecanismos de interacción robustos (Averbukh, Konovalov, & Vorzopov, 1997; Gallagher, Hatch, & Munro, 2008; Scholtz, 2006; Youn-ah, Gorg, & Stasko, 2009). Uno de los mecanismos de interacción que han implementado los árboles jerárquicos es la posibilidad de contraer y expandir ramas completas de la taxonomía o de deformarlos. Este mecanismo evita la saturación de la vista con elementos visuales que están fuera del alcance del análisis del usuario, ya que este no puede analizar simultáneamente todas las ramas de la taxonomía.

Como en el epígrafe anterior, otra de las soluciones para la pérdida del contexto es el uso de la técnica de *zoom* semántico, esta técnica se utiliza para desplegar información específica sobre un elemento seleccionado. Todas las visualizaciones en esta propuesta utilizan el *zoom* semántico en combinación con la técnica de Foco + Contexto, en donde el foco del elemento seleccionado se realiza a través del *zoom* semántico, y el contexto de la vista permanece visible. Esta solución se genera como resultado de la interacción con una vista general, a partir de la cual el usuario está interesado en analizar detalles específicos de alguno de los elementos, por lo que se combinan en una sola vista tanto el contexto de la visualización general como los detalles específicos del elemento seleccionado.

#### V.2.2.5 Saturación de elementos visuales

Como se mencionó en la sección III.6 la tendencia continua de almacenamiento y producción de cantidades inmensas de información procesable provoca el problema de la saturación de información en las

visualizaciones, que genera a su vez una saturación de los elementos visuales que se representan (ver Fig. 23 (Q. Nguyen & Huang, 2010), Fig. 27 CrystalChat (Tat & Carpendale, 2006) y Fig. 30 de (Schreurs et al., 2013)). Una solución a este problema radica en cambiar el enfoque de la visualización del modelo, en lugar de presentar todos los elementos en la misma visualización (cursos, foros, discusiones, estudiantes, profesores, académicos, comentarios, elementos de aprendizaje), como hacen la mayoría de las herramientas, sería conveniente definir nuevos procesos y flujos de trabajo, en los que el usuario pueda analizar el modelo en varias fases y, en cada una de estas, analizar un aspecto específico.

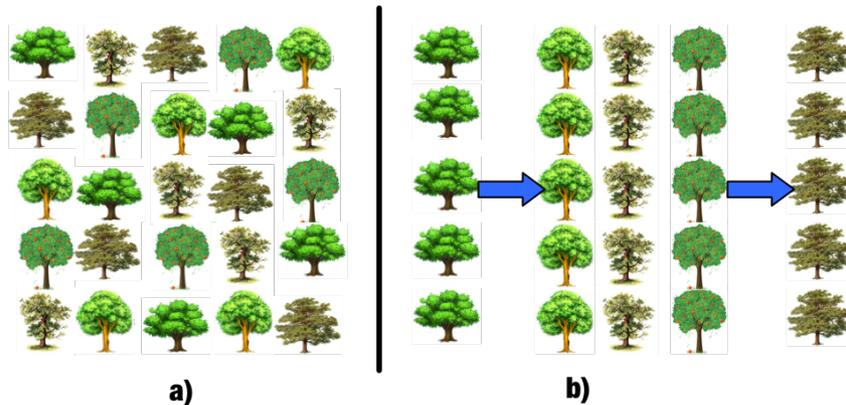


Fig. 55. (a) Muestra un esquema en el que toda la información se muestra al mismo tiempo, lo que complica en demasía el proceso de abstracción. (b) Muestra un enfoque diferente, en el que el usuario sigue un flujo de trabajo ordenado para ir conociendo parte del sistema modelado.

Esta solución, primero, evita el tener que desplegar demasiada información que el usuario es incapaz de analizar y, segundo, permite definir un flujo de actividades específico para cada fase del análisis, de acuerdo al interés del usuario. Las Fig. 55 (a) y (b) ilustran una analogía de una solución. Por un lado, la Fig. 55 (a) muestra un esquema en donde toda la información se muestra al usuario al mismo tiempo, en este caso, se trata de un bosque lleno de árboles. En este esquema resulta difícil, primero, identificar toda la información y, segundo, establecer claramente los tipos de elementos (árboles en el bosque), el número de elementos de cada tipo, etc. Por otro lado, la Fig. 55 (b) muestra un enfoque diferente para mostrar la información, en donde se define un flujo de actividades que permiten al usuario analizar la información de forma parcial, de acuerdo a sus intereses.

En conclusión las aportaciones actuales de las herramientas analizadas en el presente capítulo han solucionado problemas puntuales, analizando la información educativa desde una específica perspectiva, lo que propicia las problemáticas descritas. Todo ello ha servido como base y oportunidad de desarrollo de soluciones a través de la implementación de técnicas de VA, LA y mecanismos para la navegación e interacción en un prototipo de visualización analítica en *eLearning* que se describe en detalle el siguiente capítulo, y que constituye la principal aportación de este trabajo de tesis doctoral.



## VI.1 Diseño del modelo de analítica visual en *eLearning*

¿Por qué una herramienta visual en vez de un análisis puramente numérico? Por un lado, por medio de un análisis numérico se tiene una pregunta específica a responder siendo este un análisis poco dinámico, mientras que en el caso de las herramientas visuales no se tiene una pregunta específica sino que son la exploración y la satisfacción de la curiosidad los factores principales que llevan a la obtención de nuevo conocimiento y nuevas preguntas a responder. Para este segundo caso, las herramientas basadas en técnicas de analítica visual explotan las capacidades visuales del ser humano logrando la rápida comprensión de las correlaciones existentes y encontradas en los datos, gracias a la comparación de los elementos de datos y la exploración.

Es importante también tomar en cuenta los aspectos de evaluación y para el desarrollo de las herramientas visuales, recordando las secciones III.2 y III.3 donde se describen las consideraciones para llevar a cabo una analítica del aprendizaje y analítica académica respectivamente, quedando pendiente las consideraciones para los procesos analíticos visuales. En (Scholtz, 2006) se detallan cinco áreas o aspectos de los entornos analíticos visuales que deben considerarse como indicadores y métricas para las evaluaciones que se desarrollarán (como por ejemplo utilidad, interacción, organización espacial, entre otras).

Los aspectos de evaluación deben incluir la facilidad de uso, pero es necesario ir más allá de la facilidad de uso básica. En resumen, el objetivo de los investigadores de visualización es producir representaciones visuales de relaciones complejas que ayuden a los usuarios a centrar su atención de manera eficiente.

En el caso específico de aprendizaje electrónico, esto se puede hacer mediante el análisis interactivo visual de:

- Las actividades registradas de un usuario, almacenadas en las plataformas de *eLearning*, que pueden proporcionar un medio para comprender los procesos de aprendizaje y el comportamiento de los estudiantes durante todo su período de aprendizaje.
- El contenido semántico subido en las plataformas de *eLearning* (los comentarios aportados por los participantes y fuentes externas de aprendizaje depositadas en las conversaciones o materiales y recursos almacenados en la plataforma), así como el texto y los objetos de aprendizaje que se ponen a disposición en el curso, lo que puede llevar a cabo la visión general de un curso, tema o toda la plataforma, y también puede ser una especie de acceso directo al contenido específico.
- Las relaciones sociales que comúnmente están ocultas, pero pueden ser de gran utilidad para comprender la agrupación de los usuarios en el proceso de aprendizaje.

Recordando los objetivos de este trabajo (ver secciones 0 y I.3.2), además de suministrar información que retroalimente el proceso de aprendizaje en el *eLearning*, el resultado de aplicar la analítica visual en los procesos de interacción en los LMS debe permitir:

- Analizar de forma visual del flujo de la actividad y la distribución de los participantes (estudiantes, profesores, administradores, clientes, etc.) en la plataforma virtual.
- Contar con elementos de juicio empírico constante para mejorar el diseño de la organización de interacción, así como el diseño y la gestión de la plataforma virtual.
- Facilitar a los profesores la identificación precisa de las necesidades de aprendizaje de los diferentes estudiantes y, por tanto, la oportunidad de intervenir.
- Reutilizar la información para ajustar las estrategias de aprendizaje de los estudiantes o el enfoque de enseñanza.

Con el propósito de completar estos objetivos y utilizar la mayor diversidad posible de la información actualmente disponible en las plataformas *eLearning*, a continuación se describe cómo la analítica visual junto a la analítica educativa y analítica académica confluyen en un modelo de analítica visual aplicada al *eLearning*.

### **VI.1.1 Analítica Visual aplicada al *eLearning* (*Visual eLearning Analytics*)**

La informática y las tecnologías ubicuas han provocado una evolución en la educación hacia nuevos paradigmas de enseñanza/aprendizaje, concediendo a los usuarios acceso a aulas virtuales y a bases de datos de conocimiento desde cualquier lugar y en cualquier momento. Como consecuencia, los procesos de aprendizaje están en continua evolución, apareciendo nuevos enfoques en la universidad hacia la docencia basada en medios electrónicos y sociales (Berlanga et al., 2010; Francisco José García Peñalvo, 2008; García Peñalvo, 2011; Mateo, 2006).

Como se mencionó en la sección III.2, el término "Analítica Académica" fue acuñado por Goldstein y Katz, definiéndolo como la aplicación de herramientas de inteligencia de negocios para el área de aprendizaje (Philip J. Goldstein & Richard N. Katz, 2005). A diferencia de la AA, donde se utilizan los datos capturados para tomar decisiones a nivel de institución, el objetivo de la Analítica del Aprendizaje es el uso de estos datos, y de cualquier otra observación adicional de la que se pueda obtener información, para tener un impacto directo sobre los estudiantes, los profesores y el proceso de aprendizaje (Long & Siemens, 2011).

Siguiendo con el objetivo de mejora del aprendizaje, la Minería de Datos Educativos, como se menciona en la sección III.4, es un campo que persigue un objetivo similar al de LA. EDM propone el uso de diversas técnicas, como el análisis estadístico, aprendizaje automático, Minería de Datos, etc., para resolver los problemas de investigación educativa y entender el entorno en el que los estudiantes aprenden (Ryan S.J.d. Baker & Yacef, 2009; C. Romero & Ventura, 2010). Dentro de estas, las regresiones (Agudo-Peregrina et al., 2014) son uno de los métodos más utilizados para este propósito.

En la misma línea, la socialización y la colaboración tienen un papel clave en los procesos de aprendizaje (J. S. Brown & Adler, 2008). Aunada a esto, existe una fuerte evidencia para sugerir que dentro de las actividades bien estructuradas, los procesos de construcción del conocimiento alcanzan niveles superiores de pensamiento crítico, donde los estudiantes son capaces de establecer y sostener grupos cohesivos.

Esto sirve como justificación de la necesidad de Análisis de Redes Sociales y de herramientas que son capaces de ofrecer en tiempo real análisis para los tutores y profesores (Bakharia & Dawson, 2011). Por otra parte, la Analítica Visual, explicada en la sección IV.4, es un área emergente de investigación y su práctica tiene como objetivo apoyar el razonamiento analítico a través de interfaces visuales interactivas (J. J. Thomas & Cook, 2006). La visualización se integra como núcleo y, en conjunto con otras disciplinas científicas, mejora la división del trabajo entre el hombre y la máquina, al combinar el razonamiento analítico con la visualización interactiva. Es decir, representaciones visuales y técnicas de interacción (que explotan las capacidades perceptivas y el gran ancho de banda del ojo humano para permitir a los usuarios ver, explorar y entender grandes cantidades de información de forma simultánea), así como técnicas de apoyo a la producción, presentación y difusión de resultados analíticos que comuniquen información en el contexto adecuado para una variedad de audiencias (D. Keim et al., 2010). Las herramientas de VA hacen posible obtener un modelo mental de los datos complejos y nuevo conocimiento (D. Keim, Andrienko, et al., 2008; James J. Thomas & Cook, 2005). El objetivo es ayudar a los usuarios a descubrir las anomalías inesperadas y sorprendentes, cambios de patrones y relaciones que luego son evaluadas para desarrollar nuevas ideas (Cook, Earnshaw, & Stasko, 2007).

En la presente tesis se pretende mejorar la eficiencia del proceso de análisis de aprendizaje. El objetivo es proponer y validar un modelo que tome en cuenta técnicas de analítica visual y la información existente en un LMS, con el fin de mejorar el proceso de LA y AA, sin olvidar asegurar el éxito y desarrollo de la formación educativa. Para ello, a manera de comprobación del modelo, se ha implantado un sistema y aplicado sobre un LMS para validar su expresividad.

### **VI.1.2 Definición del modelo de analítica visual aplicada al *eLearning***

Al igual que con la mayoría de las áreas de investigación, la terminología que identifica y diferencia EDM, AA y LA no es tan homogénea como sería deseable (Van Barneveld et al., 2012), sin embargo, el aspecto que las tres áreas tienen en común, además del objetivo de mejorar y comprender el proceso de aprendizaje, es la necesidad de una gran cantidad de observaciones de dicho proceso. Por otro lado, en el trabajo (Siemens & Baker, 2012; Siemens et al., 2011) se plantea que las comunidades de LA y EDM se da un solapamiento considerable (tanto en términos de investigación como en sus investigadores), y que ambas comunidades creen firmemente en la realización de investigaciones y aplicaciones que beneficien a los estudiantes, así como la información y la mejora de las ciencias del aprendizaje; además, plantea algunas diferencias entre estas comunidades.

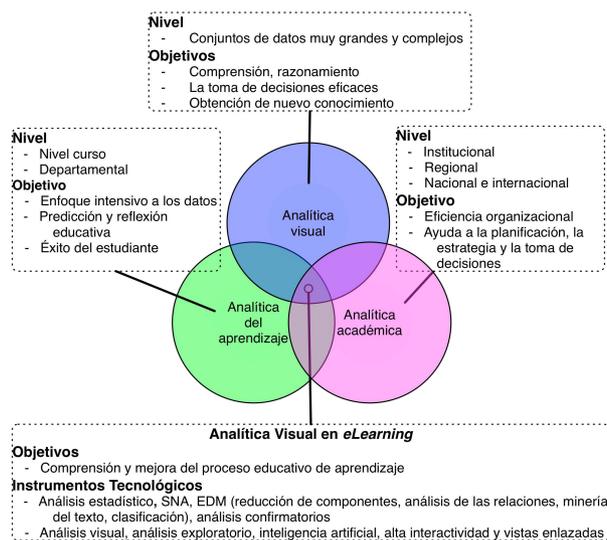


Fig. 56. Modelo teórico de la Visualización Analítica en *eLearning* (Gómez Aguilar, García Peñalvo, & Therón, 2014).

Las técnicas, métodos y objetivos que diferencian e identifican a la LA del EDM se encuentran mencionados y aplicados en diversas áreas de la ciencia dentro del alcance de la VA, por tanto, realizar una aplicación de VA en el área de la educación, más que excluir a la AA y a LA, las complementa. La notable intersección de estas áreas y sus diferencias da lugar al presente modelo de Visualización Analítica en *eLearning* (del inglés, *Visual eLearning Analytics*, VeLA).

El presente trabajo considera al EDM y el SNA como parte de los métodos, técnicas e instrumentos tecnológicos que se utilizan en VeLA. En la Fig. 56, se muestra cómo las áreas de LA y AA comparten un conjunto de objetivos e instrumentos tecnológicos y cómo el área de VA en el modelo VeLA complementaría y enriquecería al proceso de AA y LA (Chaomei, Ibekwe-SanJuan, SanJuan, & Weaver, 2006; D. Keim et al., 2010; Daniel A. Keim et al., 2010; T. von Landesberger, Gerner, & Schreck, 2009), lo que constituye una contribución del presente trabajo de investigación. La unión de las métricas y técnicas utilizadas en LA, AA y los procedimientos propios de la EDM (reducción de componentes, análisis de las relaciones, minería del texto, clasificación) y de la SNA, sus instrumentos tecnológicos como son el análisis estadístico, análisis confirmatorios, vistas enlazadas, *zoom*, foco más contexto, *zoom* semántico, etc., con el objetivo de la mejora de la educación en sus diferentes niveles, se ve fortalecida y optimizada con las técnicas y las métricas que se toman como parte de la visualización analítica y se complementan con las técnicas de interacción y de visualización de la misma. Lo que da como resultado unos objetivos en común:

- a) Enfoque intensivo en los datos;
- b) Éxito del estudiante;
- c) Ayuda a la planificación, la estrategia y la toma de decisiones; y
- d) Contribución al campo base de investigación.

El modelo de la Fig. 57 (VeLA) se basa en el modelo de referencia de LA de Chatti et al. (2012), la teoría del ciclo de LA de Clow (2012), el proceso de la exploración visual de datos de D. Keim et al. (2010) y, además, toma en cuenta las restricciones y elementos del modelo de dimensiones críticas de la LA de Greller and Drachsler (2012).

Por tanto, la aportación teórica principal que surge de la utilización de la VA en el área de la educación es que el proceso del modelo Vela combina métodos de analíticas automáticas con técnicas visuales exploratorias que, con una alta interacción disponible para el usuario, tienen la finalidad de obtener nuevo conocimiento de los datos educativos.

Ben B. Shneiderman (1996) propuso un paradigma para la búsqueda de la información que fue extendido por Daniel D. Keim, Andrienko, et al. (2008) para dar más capacidad de análisis en el bucle de antes y después de la generación de representaciones visuales interactivas. Con estas el usuario explora los datos para extraer modelos abstractos a partir de conjuntos de datos que son demasiado grandes o demasiado complejos para ser analizados de una manera directa, de forma que las preguntas del usuario se respondan y se creen durante el ciclo de exploración.

Asimismo, en este trabajo, al unirse las áreas de LA, AA y VA, el paradigma de Keim aplicado a la analítica educativa, se extiende en un paso final, la intervención, que proporciona al ciclo analítico una retroalimentación, para quedar de la siguiente forma:

*“Analizar primero, mostrar lo importante, hacer zoom, filtrar y analizar más profundamente, detalles bajo demanda, intervención”;*

del inglés,

*“Analyze First; Show the Important; Zoom, Filter and Analyze Further; Details on Demand, Intervention”.*

Este ciclo está representado en la Fig. 57 en la parte central del proceso. Este esquema se debe interpretar como un proceso de análisis visual exploratorio, a la vez que se tiene en cuenta que no es suficiente simplemente recuperar y mostrar datos mediante el uso de metáforas visuales, sino que es necesario analizar los datos de acuerdo a su valor de interés, que muestran los aspectos más importantes de los datos y, al mismo tiempo, proporcionan formas de interacción que habilitan al usuario para obtener detalles de los datos dentro de los ciclos exploratorios.

En la Fig. 57 se muestra una visión general abstracta de las diferentes etapas (representadas a través de círculos) y sus transiciones (flechas) en el proceso de VeLA. Dentro de cada etapa se encuentran representadas las dos áreas de la analítica en la educación, la AA y la LA, y su respectiva descripción en cada etapa. Como se aprecia en el diagrama, el modelo contempla distintos perfiles de usuario, la diferencia entre estos es el objetivo y, por tanto, el tipo de información de su interés.

# Proceso de la analítica visual en eLearning

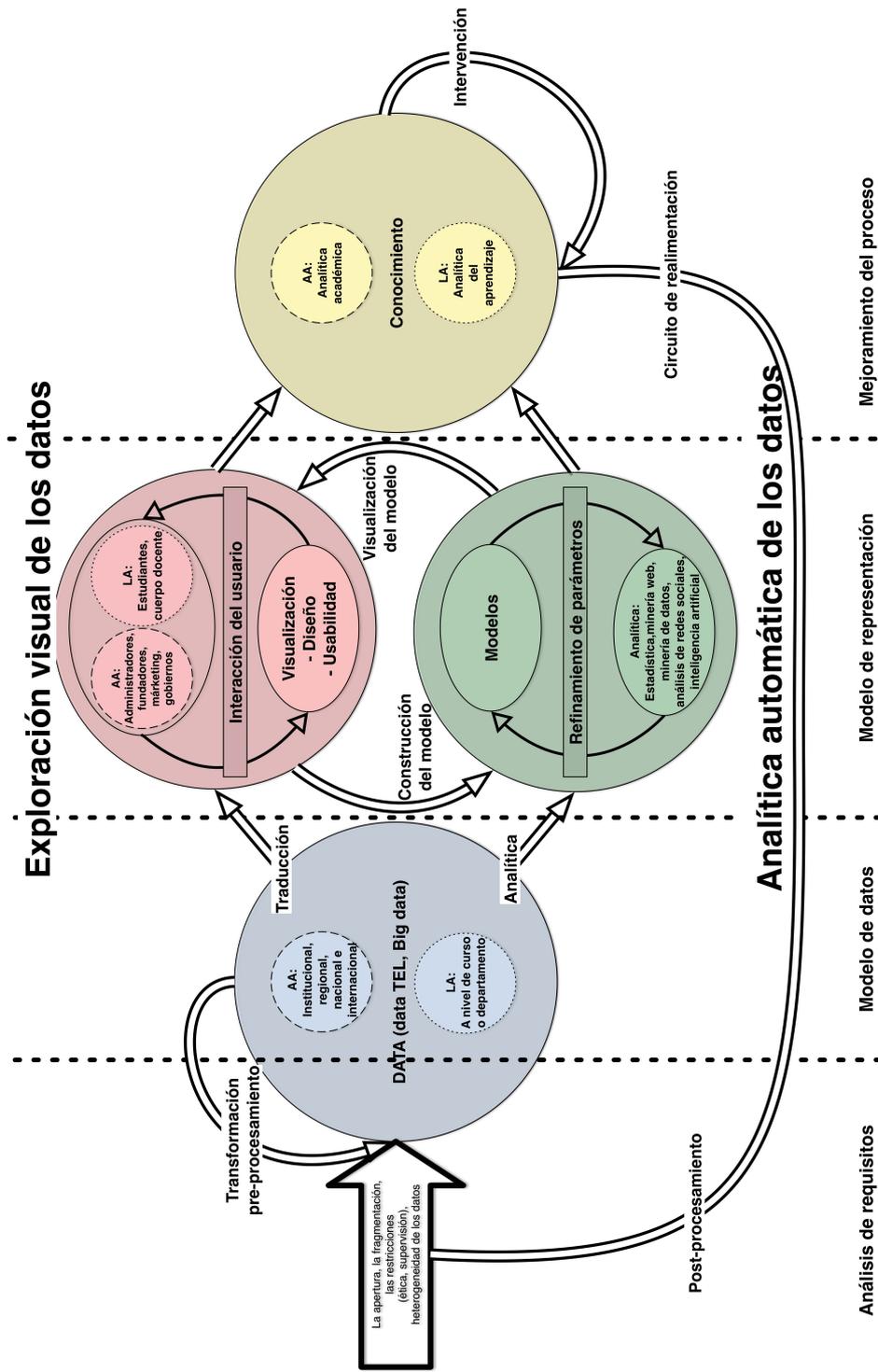


Fig. 57. Proceso de Analítica visual en eLearning (Gómez Aguilar, García Peñalvo, & Therón, 2014).

En la parte inferior del modelo se presentan unas etapas en el proceso de creación de análisis:

- **Primera fase del modelo.**  
Consiste en realizar un análisis de los requisitos generales del sistema que se deben cubrir para el análisis de las interacciones. Esta fase da respuesta a la pregunta inicial: “¿Qué necesitan saber los diferentes perfiles sobre el análisis de los datos educativos?” El producto final de esta fase es la especificación conceptual del problema en estudio y los requisitos de usuario que deben satisfacer sistemas de este tipo.
- **Segunda fase del modelo.**  
Permite definir un Modelo de Datos Conceptual sobre los datos necesarios para definir una base de datos con la información necesaria para un sistema de análisis analítica visual en *eLearning*. El producto final de esta fase es el Modelo de Datos representado como un modelo de clases en UML (Object Management Group, 2014), que contiene las entidades y las relaciones entre ellas. También se estudian las posibles consultas que producirán una serie de entradas a la siguiente fase.
- **Tercera fase del modelo.**  
Sirve de ayuda en el proceso de representar eficazmente los datos provenientes de las consultas a la base de datos. Para ello, se ha llevado a cabo el proceso que identifica la manera de representar estos datos, mediante técnicas y métodos de visualización analítica, visualizaciones interactivas y estadísticas que permitan una mayor variabilidad en los datos y faciliten la interpretación de los datos a través de las interacciones. En esta fase se intentará conseguir identificar la mejor manera de mostrar eficazmente las consultas de los usuarios. Esta fase del proceso coincide con el ciclo exploratorio mencionado anteriormente, que comprende el descubrimiento y el análisis de los patrones de comportamiento de los usuarios, que, a su vez, permitirá al profesor conocer cómo trabajan los estudiantes. Este nuevo conocimiento es el resultado final del modelo de representación que categoriza las múltiples vistas enlazadas del sistema del prototipo de VeLA.
- **Cuarta fase del modelo.**  
Tiene el objetivo general de mejorar el proceso de enseñanza y aprendizaje. El producto final de esta fase del Modelo del Proceso de Analítica Visual para el *eLearning* es la intervención en el proceso de aprendizaje, es decir, la realimentación del proceso, que propondrá una solución para cada problema, falta o disfunción en el aprendizaje o la creación de nuevas hipótesis y preguntas a responder.

Dado que el tipo de dato que se maneja para el análisis tiene una temporalidad variable, una socialización y una alta cantidad de actividades sin restricción de tiempo o espacio, este trabajo, con el fin de mejorar el reconocimiento de patrones, aborda específicamente las siguientes necesidades/desafíos en el proceso de VeLA.

### VI.1.2.1 Necesidades/desafíos en el proceso de VeLA:

- **Temporalidad.**
  - Analizar el contenido del curso de *eLearning* a través del tiempo y no solo en un curso, sino también un campus o toda una institución educativa.
  - Definir una representación temporal compacta, a fin de encontrar patrones (en meses, horas o semanas) a través de interacción.
  - Personalizar el análisis por agrupación de actividades, tiempo, persona específica y perfil.
  
- **Análisis del contenido semántico.**
  - Determinar un diseño nube de etiquetas que represente su evolución en el tiempo, en múltiples niveles de detalle, para equilibrar así la coherencia semántica, de contenido y la estabilidad espacial de la visualización.
  - Plantear un diseño compacto que no impacte negativamente en la legibilidad de la nube de etiquetas, de forma que la importancia o la frecuencia y la evolución en el tiempo de una etiqueta se codifiquen directamente en espacio del tamaño de la palabra.
  
- **Análisis de redes sociales.**
  - Realizar la comparación de la estructura jerárquica de la plataforma y las relaciones que se crean entre las personas (foros y los contribuyentes compartidos entre foros).
  - Definir una herramienta gráfica social que brinde la posibilidad de encontrar la relación directamente proporcional del desempeño de los estudiantes con la frecuencia de lectura en foros y/o recursos.
  - Ofrecer la posibilidad de ocultar elementos de la red social a fin de ver más claramente la estructura y la relación de los elementos que interesen.
  
- **Métricas estadísticas**
  - Definir un sistema que permita la múltiple selección de subconjuntos y/o el uso de diferentes categorías personalizadas de actividades, para hacer posible su diferenciación por su color, con el objetivo de la identificación de patrones. Además, se permite generar un archivo con ellas, lo que posibilita la importación y exportación.
  - Soportar métodos de interacción, como el filtro de búsqueda, o la ocultación y la reordenación de las diferentes métricas.

Con el fin de comprobar la utilidad y expresividad del modelo y su aplicabilidad se ha implementado un sistema de *software* del modelo VeLA como resultado de cuatro ciclos de investigación-acción (Ferrance, 2000), los primeros dos en julio 2010 y 2012, posteriormente otros dos en febrero del 2013 y 2014, basados en los trabajos mencionados y en el modelo anteriormente descrito.

A continuación se explica la implementación y diseño del prototipo y posteriormente se describe el proceso de evaluación.

## VI.2 Proceso del diseño y evaluación del prototipo de VeLA

A manera de resumen el proceso de diseño y evaluación del prototipo del modelo VeLA sigue la metodología de investigación-acción y la metodología de investigación basada en el diseño, con el objetivo de diseñar y desarrollar de forma rápida de prototipos para evaluar las ideas en ciclos. De esta forma, y siguiendo las metodologías antes mencionadas, el sistema del prototipo del modelo VeLA se desarrolló durante cuatro iteraciones. Antes de cada iteración se estableció una versión del prototipo y un caso de estudio, resultando tres versiones del prototipo de VeLA. A continuación se describe la versión final del prototipo, su implementación y acceso a los datos a analizar.

### VI.2.1 Implementación y acceso a las fuentes de datos

Se toma como base la plataforma MOODLE Moodle (2014), ya que en numerosas fuentes bibliográficas la mencionan como una de las plataformas con mayor adaptabilidad entre otras características (Graf & List, 2005; Ohta et al., 2005). En cuanto a la implementación y/o modificación de nuevas técnicas de visualización, se utilizó Java como lenguaje de programación. En este caso, la herramienta se ha adaptado para trabajar con Moodle y hacer uso de varios años de datos almacenados en la base de datos de la plataforma. Sin embargo, los principios de diseño son lo suficientemente generales, de forma que el prototipo del modelo VeLA podría adaptarse, sin demasiado esfuerzo, para trabajar con cualquier otro CMS o cualquier otro tipo de datos.

La herramienta se ha desarrollado con una arquitectura Modelo-Vista-Controlador (del inglés, *Model-View-Controller*, MVC) (Buschmann et al., 1996), y la conexión entre la herramienta y Moodle se ha realizado mediante una capa de servicios web, para lo que se ha seguido el siguiente proceso: (1) es necesario definir si el usuario puede usar un servicio web de autenticación de usuarios; (2) el cliente puede invocar un servicio para recuperar la información requerida; (3) en función de las capacidades del usuario se devolverá distinta información; (4) una vez que se recupera la información, esta se pre-procesa y se le da formato para mostrarse adecuadamente a través de la herramienta de visualización analítica.

La herramienta tiene capacidad de exploración y comparación para lo que emplea cuatro técnicas de visualización basadas en líneas de tiempo en espiral (STL) (Gómez Aguilar, García Peñalvo, & Therón, 2013b); Gómez Aguilar, Therón, and García-Peñalvo (2009), nubes de etiquetas (WC) (Gómez Aguilar, Conde-Gonzalez, García-Peñalvo, & Therón, 2011), coordenadas paralelas (PC) (Gómez Aguilar, García Peñalvo, & Therón, 2013a) y gráficos de redes sociales (SG) (Gómez Aguilar, Therón, & García-Peñalvo, 2013).

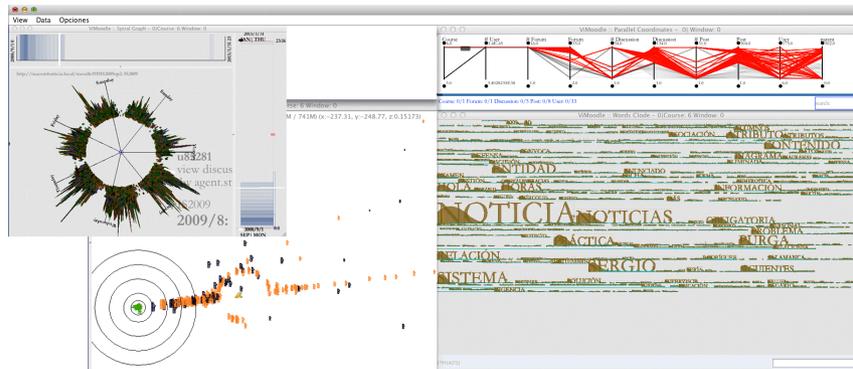


Fig. 58. Vista general del prototipo de VeLA.

Estas se utilizan con diferentes propósitos dependiendo de sus características de expresividad, distribución visual y potencia. Una vista global de las técnicas de visualización se puede observar en la Fig. 58. Debido a que la herramienta permite perfiles con diferentes privilegios de información, y estos varían desde tener acceso a un curso hasta toda una institución educativa, las diferentes técnicas mencionadas se han implementado siguiendo un esquema de flujo de análisis adecuado (ver Fig. 59). Este se inicia con la conexión a un LMS, posteriormente dadas sus características óptimas en la representación de jerarquías, se muestra la estructura de la información a la que se tiene acceso, la cual depende de los privilegios del usuario, con las coordenadas paralelas.

Esta representación de coordenadas paralelas sirve como una vista global de la información del LMS consultado y se describirá en la siguiente sección VI.2.1.1. Cabe mencionar que el prototipo permite múltiples conexiones y, por tanto, se desplegarían múltiples representaciones de coordenadas paralelas en diferente color para cada una de las fuentes del LMS. Desde esta vista global, usando coordenadas paralelas, se puede dirigir el analista hacia uno o varios cursos internos de la fuente de datos del LMS para su análisis y comparación por medio de cualquiera de las representaciones visuales antes mencionadas (línea de tiempo en espiral, nube de palabras, red social e incluso la representación de coordenadas paralelas, técnicas que se describen a continuación).

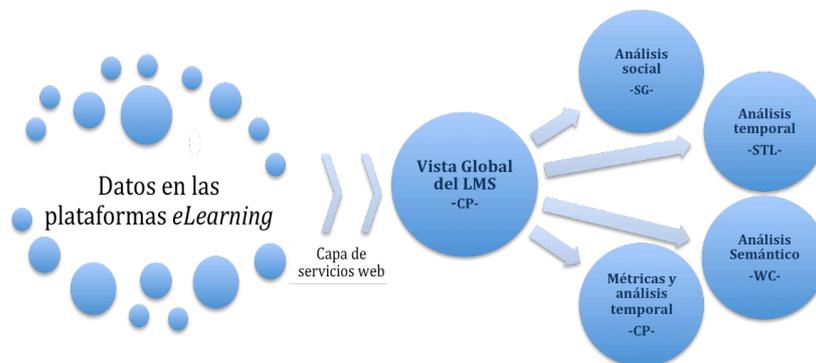


Fig. 59. Esquema de flujo de trabajo del prototipo de VeLA.

En primer lugar se describirán las coordenadas paralelas (PC), ya que es la primera representación que se despliega después de realizarse la conexión y el pre procesamiento de los datos. Posteriormente se describirá la línea de tiempo en espiral (STL), a continuación de esta se detallará la representación de nube de palabras (WC) y, finalmente, la red social (SN).

Una característica importante es que los cuatro puntos de vista (nubes de palabras, red social, línea de tiempo y coordenadas paralelas) están interrelacionados. Esto significa que, dado que todos los puntos de vista son diferentes maneras de transmitir la información, cuando el usuario interactúa con uno de ellos y cambia la representación, los otros puntos de vista también cambian dependiendo de la acción realizada originalmente. Con esto se da solución a unos de los problemas explicados en las secciones V.2.2.3, sobre el manejo adecuado de los planos y los contextos, y la sección V.2.2.4, sobre la pérdida del contexto en la navegación.

### VI.2.1.1 Representación de las coordenadas paralelas

Las coordenadas paralelas son una forma común de visualizar datos de alta dimensión en un espacio limitado, a la vez que permiten la exploración de las tendencias. Desde la introducción temprana de Inselberg (Inselberg, 1985; Inselberg & Dimsdale, 1990), las coordenadas paralelas han sido ampliamente utilizadas para visualizar datos multivariados (Tory et al., 2005), estas son una de las técnicas de visualización más conocidas y, aunque en un principio al analista le puede resultar confuso, es una herramienta muy potente para la comprensión de conjuntos de datos complejos multidimensionales. Además de un poco de experiencia en la lectura de coordenadas paralelas, la mejor manera de conocer un conjunto de datos es mediante las técnicas de la interacción que proporcione la visualización.

r: 0.58fps - 31919 items - fonts(4) colors(17) mem(146M / 2039M) (x--706.45, y--369.80, z:0.00565)

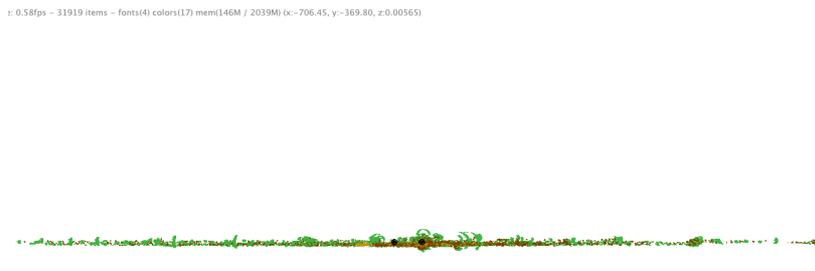


Fig. 60. Vista global del campus completo de una universidad representado con una red social, en esta se muestran los cursos en color naranja, sus foros en color marrón y sus discusiones en color verde únicamente.

Para entender mejor esta técnica refiérase a la sección V.1.3.3. Este potencial del uso eficiente del espacio al representar datos multidimensionales es la razón principal por la cual se ha elegido, como antes se ha mencionado, para desplegar la vista global de la estructura de datos. Un ejemplo de esto se puede observar, con el objetivo de mostrar el potencial de esta técnica en comparación con otras herramientas y representaciones, con la siguiente prueba de expresividad y usabilidad antes de ser elegida esta como óptima para la investigación, y mostrándola como ejemplo de la posible visión global de exploración.

El ejercicio consiste en representar la misma cantidad de información con dos de las técnicas de representación más comúnmente usadas, las redes sociales (ver Fig. 60 y Fig. 61) y las coordenadas paralelas (ver Fig. 62). Con el uso de estas técnicas, se exponen en estas la misma gran cantidad de datos educativos, tal como se necesita en AA, LA y que es materia de estudio de VA. En las imágenes de las Fig. 60 y Fig. 61 se representaron los datos de un campus completo. Como se puede observar en la Fig. 60, donde se visualizan los cursos, sus foros y sus discusiones, la cantidad de datos obtenidos para ser representados como vista global en una red social, resulta poco eficiente ya que se convierte en una red de tal inmensidad que resulta complicado identificar alguna estructura o encontrar información relevante para la creación de hipótesis o nuevas preguntas de análisis.

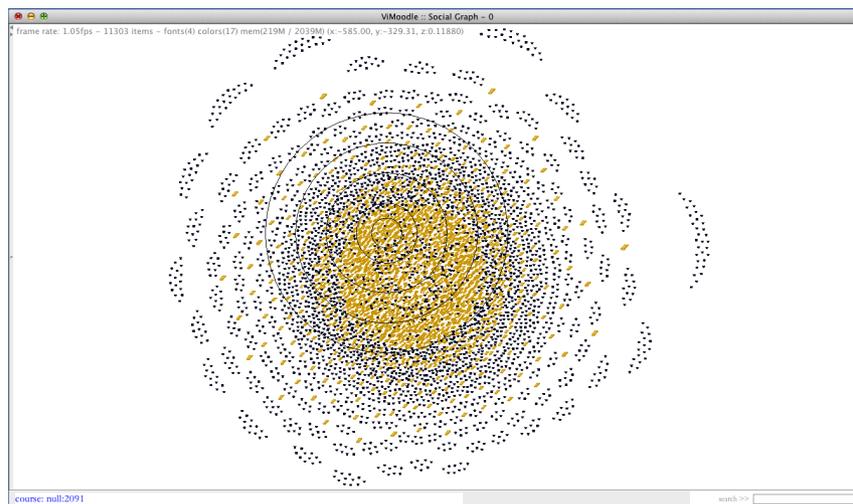


Fig. 61. Vista global del campus completo de una universidad representado con una red social, en esta se muestran los cursos en color naranja y sus foros en color marrón únicamente.

Antes de continuar es preciso comentar que estos datos, tomados de un campus universitario completo, comprende diez meses (de junio del 2008 a abril del 2009) con un total de más de 9 millones de registros, más de 2300 cursos, más de 4800 foros, cerca de 12000 discusiones, casi 35 mil publicaciones en foros y cerca de 20 mil estudiantes. Por este motivo se decidió eliminar las discusiones de la visualización con el objetivo de corroborar si era posible realizar una vista global con esta representación de red social (ver Fig. 61) desde una perspectiva más general que elimina una jerarquía en la estructura de la institución. El resultado, aunque mejor que el anterior, continua siendo poco eficiente con en uso del espacio y la expresividad, impidiendo encontrar algún patrón en la estructura o nuevas preguntas de investigación. Finalmente, como en Govaerts et al. (2012) y así como en el presente prototipo, se prosiguió con la representación de los mismos datos, en los ejes verticales de las coordenadas paralelas (ver Fig. 62). Estos ejes representan diferentes métricas, como por ejemplo, el promedio de *posts* de los estudiantes, el promedio de notas de estos, el número de foros, discusiones y *post*, etc., y depende de la cantidad de información a la que tiene acceso la herramienta.

La forma habitual de describir las coordenadas paralelas sería hablar de espacios de alta dimensión representados con la técnica de establecer los ejes de coordenadas en paralelo en lugar de ortogonales entre sí. Para que sea más fácil de interpretar y entender, se puede imaginar una tabla de datos en donde cada una de las columnas están asignadas a un eje vertical en la imagen de la Fig. 62. Cada valor de datos le corresponde una posición a lo largo de la línea, dependiendo del valor del dato terminaría entre el mínimo en la parte inferior y el máximo en la parte superior. Una colección pura de puntos en estas posiciones de los datos no sería extremadamente útil, por lo que los puntos que pertenecen a un mismo registro (fila) se conectan con líneas. Para dar un poco de sentido a esto, la forma más fácil es olvidar la parte de los registros que abarcan todo el ancho y mirar el espacio entre cada par de ejes. De esta forma se pueden encontrar los patrones en cada uno de los espacios entre los ejes.

Sin embargo, cuando los tamaños de los conjuntos de datos se vuelven tan grandes, como millones de líneas que llenan la pantalla, las coordenadas paralelas pueden llegar a ser demasiado densas para interpretarse. La imagen de la Fig. 62 muestra la estructura de todos estos cursos y, para este ejemplo, se han ocultado algunos ejes de métricas e información privada. Sin embargo, siendo las columnas mostradas: curso, foro, discusión, *post*, usuario y usuario que contesta a una publicación determinada, además de la capacidad de representar más características del campus universitario que en la técnica anterior de red social, ya a simple vista, con la ayuda de la transparencia en las líneas y las tonalidades, y sin haber realizado ninguna interacción, se perciben las zonas dónde se acumula mayor frecuencia.

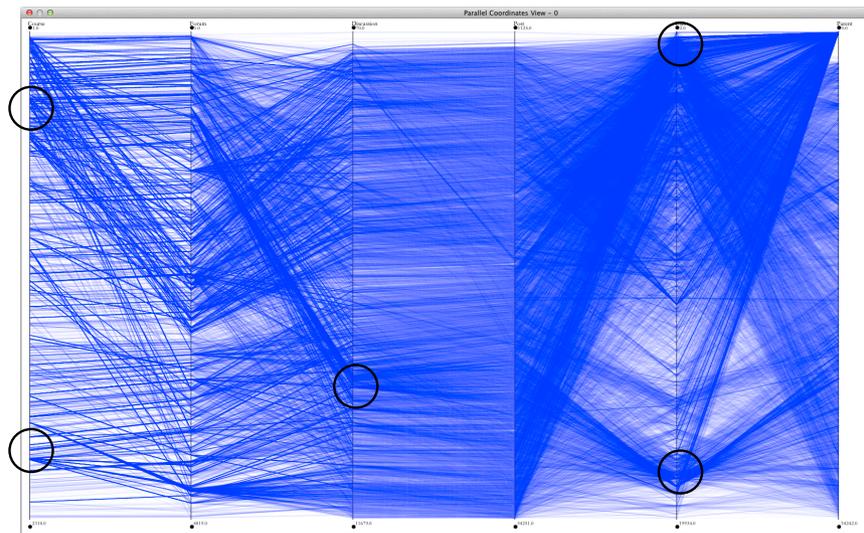


Fig. 62. Vista Global del campus completo de una universidad en coordenadas paralelas. En las columnas, de izquierda a derecha, se expresan los cursos, sus foros, sus discusiones, sus *posts*, quién escribió el *post*, y quién contestó. <http://analiticavisual.es/video-gallery/item/12-vela-vimoodle-sep-2013>.

En dichas zonas, indicadas con círculos negros sobre la imagen de la Fig. 62, ya sea de foros de un curso (segunda columna) o de *posts* de un usuario específico (quinta columna) o de discusiones de un foro (tercera columna), se pueden observar algunos detalles que llaman la atención. Por ejemplo, entre el segundo y tercer eje (ejes de foro y discusión), se notan unas zonas más oscuras que otras (señaladas con los círculos sobre la Fig. 62), estas revelaciones serían un buen punto de partida para la exploración. Además, en la columna de usuario y usuario que contesta a una publicación de foro determinado (último y penúltimo eje) también se notan unas zonas más oscurecidas que otras (marcadas con los círculos sobre la imagen de la Fig. 62), estos son usuarios que participan mucho más que el resto, e inmediatamente surgen nuevas preguntas como: ¿estos usuarios participan de tal forma que sobresalen del resto en uno o varios cursos?, ¿los cursos son de un área en específico?, y ¿los usuarios que contestan significativamente más a las publicaciones en foros, son profesores o estudiantes?, ¿coinciden con los que publican significativamente más que el resto de usuarios?, etc. Si bien este es un ejemplo muy simple, muestra las estructuras típicas que se puede encontrar en la mayoría de los conjuntos de datos de este tipo, y resulta razón suficiente para ser elegida como representación para la vista global del prototipo presentado.

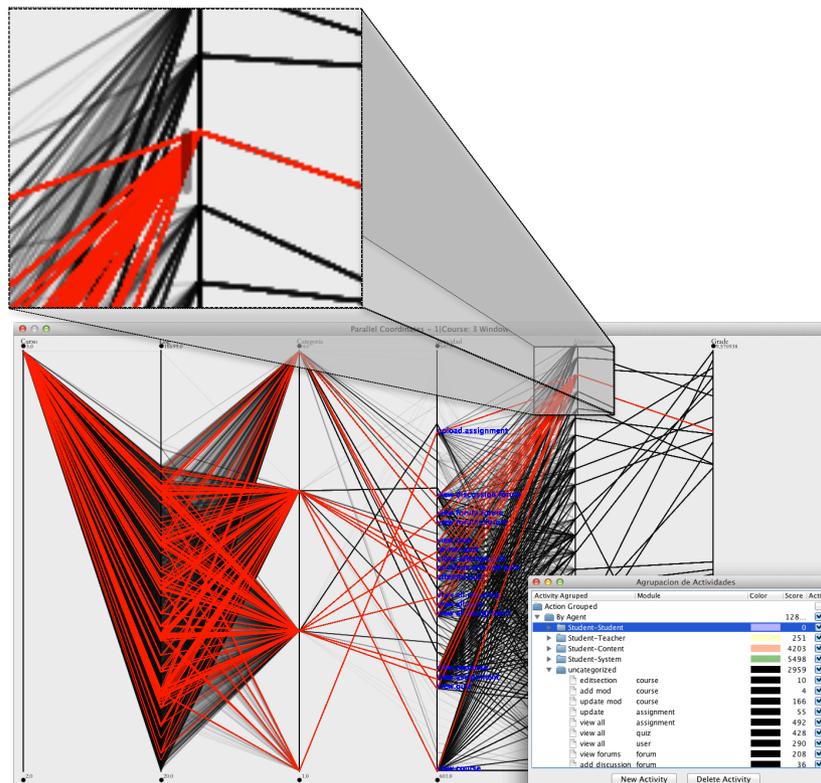


Fig. 63. Vista del subconjunto de las actividades realizadas durante todo el curso de un estudiante con una nota de más de 8.

Consecuentemente, con esta técnica se resuelve una de las necesidades VI.1.2.1 y problemáticas descritas en la sección V.2.2.5. Ahora bien, dada la versatilidad y optimización del espacio de esta herramienta, en la presente propuesta esta técnica se utiliza en dos puntos del flujo de análisis, cambiando el tipo de dato que representa, siendo este un *zoom* semántico.

Primero se elige para utilizarse como vista global y, por tanto, como punto de partida para el análisis y comparación de los cursos que en esta se representen; pero también se utiliza como parte de las técnicas de representación para la comparación y análisis de los cursos contenidos en la fuente de datos junto con la nube de palabras, la red social y la línea de tiempo en espiral.

El paso de un punto a otro se realiza colocándose sobre alguna de las líneas que representan los cursos y con el botón derecho se despliega una lista en un menú contextual que contiene las visualizaciones posibles, incluida también las coordenadas paralelas, y posteriormente al colocarse sobre una visualización se despliega un submenú de todos aquellos cursos que se encuentren bajo la selección. Al elegir la representación de las coordenadas paralelas, o cualquiera de las disponibles, el sistema realiza un pre-procesado de los datos y después despliega la información en la visualización seleccionada, sin embargo para el caso de la selección de las coordenadas paralelas funciona como un *zoom* semántico ya que con la misma técnica de visualización se despliega diferente información, en la vista global se explora y se analiza todo un campus (ver Fig. 62) y en este caso el curso seleccionado (ver Fig. 63, información jerárquicamente contenida en la de la vista global pero mostrada en detalle).

Obsérvese la diferencia entre la Fig. 62 y la Fig. 63. A nivel de los datos que se representan en esta misma técnica pero en diferentes niveles jerárquicos, por una parte, a nivel global, cada línea poligonal formada por la unión de los diferentes puntos de cada eje vertical representa un curso (ver Fig. 62), y, por otro lado, después de interesarse en un curso específico, los datos representados son otros, siendo cada estudiante del curso en cuestión una línea horizontal poligonal de las coordenadas paralelas (ver Fig. 63).

Los datos utilizados en la Fig. 63 se obtuvieron de una plataforma Moodle con el fin de analizar un curso. La duración de este fue de Noviembre 2011 a Mayo 2012, casi el ciclo escolar con 13000 registros de actividades que fueron realizados por un total de 30 estudiantes. Los datos consisten en las filas: "Log", "Curso", "Estudiante", "Rol", "Actividad", "Modulo", "Categoría", "Grade", no obstante es posible cambiar el orden y ocultar o mostrar cualquiera de estas. Para este estudio se ha ocultado la columna de "rol" y "Modulo" además se ordenaron los ejes de la siguiente manera: "Curso", "Log", "Categoría", "Actividad", "Estudiante", "Grade".

En la Fig. 63 se muestra la selección de un estudiante con nota alta con el objetivo de satisfacer la curiosidad de qué actividades o que caracteriza a este grupo de estudiantes que sobresale. Se puede observar, a primera vista, que a mayor frecuencia de actividades (nótese la cantidad de líneas en rojo que cruzan el segundo eje de izquierda a derecha, eje de "Log") es más alta la nota. Los datos para este estudio se recogieron durante todo el curso.

En primer lugar, el número de accesos, la lectura de acceso y los mensajes para cada discusión, se contó el número de accesos de lectura a un recurso para producir datos cuantitativos para el análisis, se obtuvo el módulo de Moodle utilizado y el contexto con el fin de realizar la categorización adecuadamente. Para el caso de este segundo nivel, el uso de categorías tiene importancia para esta propuesta dado que se ha decidido implementar la manera de categorizar dinámicamente las actividades de los estudios mencionados en la Introducción, descritos en la sección V.1.2 y analizados en la sección V.2.1.2 (Agudo-Peregrina et al., 2014; Che-Cheng & Chung-Hui, 2013; Davies & Graff, 2005; Giovannella et al., 2013; Hwang & Arbaugh, 2009; Kim, 2013; Lipponen et al., 2003; Macfadyen & Dawson, 2010; Medeiros et al., 2013; Romero-Zaldivar et al., 2012; Weisskirch & Milburn, 2003).

Además de permitir el uso de categorías dentro de los ejes de las coordenadas paralelas y la codificación del color de las líneas en base a estas, también permite la creación de agrupaciones personalizadas y generar un archivo con estas para su importación y exportación. Con este tipo de implementación de categorizaciones y representación de métricas y los métodos de interacción como: el filtro de búsqueda a través de la caja de texto abajo a la izquierda de la visualización, el ocultamiento y la reordenación de columnas también son de vital importancia para un análisis visual. La principal de estas interacciones en coordenadas paralelas se llama *brushing*. Esta es una técnica muy efectiva para especificar una señalización y enfoque de un conjunto de los datos de forma explícita en la visualización de información (Becker & Cleveland, 1987; M. O. Ward, 1994).

Mediante esta técnica el usuario marca activamente subconjuntos de los datos de ajuste mediante el uso del ratón, seleccionando zonas de intersección con las líneas, dibujadas como rectángulos de selección en muchos casos o rangos en los ejes (mírese el pequeño rectángulo sombreado de selección de subconjuntos de la Fig. 63 en la penúltima columna señalado con el cuadro negro de *zoom* a la derecha de la imagen). Si se utiliza en conjunción con múltiples puntos de vista vinculados, el *brushing* (Fig. 63) puede permitir a los usuarios entender las correlaciones a través de múltiples dimensiones (Becker & Cleveland, 1987; Gresh et al., 2000; Hauser et al., 2002). Cabe mencionar que la visualización implementada permite la multi-selección de subconjuntos en los ejes verticales o fuera de ellos. Otra técnica es el uso de las diferentes agrupaciones de las actividades, para colorear las líneas de las coordenadas paralelas con el objetivo de identificar patrones. La interacción y más detalles de la representación de las coordenadas paralelas se puede observar en el vídeo de la URL de la leyenda de la Fig. 62) y el que se encuentra en <http://analiticavisual.es/video-gallery/item/11-vela-vimoodle-may-2013>.

### VI.2.1.2 Línea de tiempo en espiral

La línea de tiempo en espiral es una representación visual interactiva que se utiliza para analizar el uso de un LMS a través del tiempo. Su objetivo es proporcionar una representación compacta de la actividad global del LMS, lo que aporta una visión general de los entornos de aprendizaje con

posibilidad de ir al detalle que se desee y algunas otras técnicas que se describirán en esta sección. La distribución de la interfaz gráfica (ver la Fig. 64), más allá de la zona central donde se encuentra la espiral, se organiza de la siguiente manera: los otros dos paneles proporcionan vistas adicionales de los datos mostrados en la espiral, en la vista superior se dibuja de forma lineal horizontal, los datos actualmente seleccionados. En el lado derecho se encuentra una vista general (el contexto), representada de forma lineal vertical y que incluye un control deslizante para seleccionar el período de tiempo que se quiere analizar. En estas tres representaciones (temporal lineal vertical, temporal lineal horizontal y temporal en espiral) existen dos elementos de información, representados con objetos visuales, que se deben considerar en el cuerpo de las líneas de tiempo, la vertical, la horizontal y la representada en espiral.

El elemento principal es una barra de colores que representan las actividades (Fig. 64, indicado con líneas punteadas de color naranja); en este caso el color se corresponde con el tipo de las actividades, pero se puede personalizar para que también represente el curso o estudiante que realiza las actividades, alguna agrupación personalizada o el rol entre otras.

El segundo elemento se utiliza para representar las unidades de tiempo de nivel superior (Fig. 64, señalado con líneas punteadas azules) que se pueden personalizar en función de la finalidad, en días, semanas, meses o años, es posible pintar una o varias de estas simultáneamente, aunque no es recomendable, ya que podrían saturar la visión del analista; así, por ejemplo, las actividades (barras) que se produjeron dentro de un día se envuelven con un rectángulo, y el color del rectángulo (unidad de tiempo) puede ser designado de acuerdo con el tiempo (día, semana, mes, año) o la frecuencia.

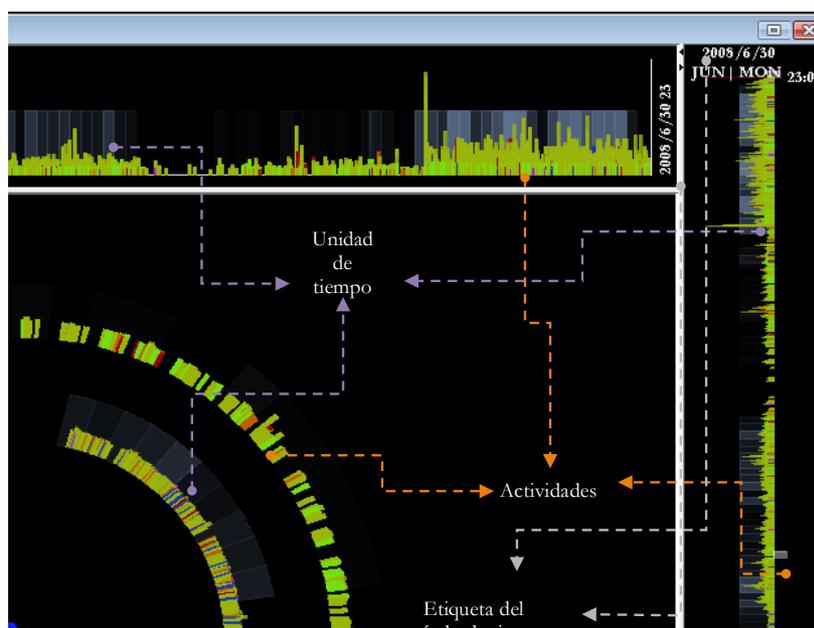


Fig. 64. Elementos representados en la visualización de línea de tiempo en espiral. En esta se aprecian las tres representaciones del tiempo: en la esquina izquierda de abajo la espiral, a la derecha la línea de tiempo vertical del contexto y arriba la línea de tiempo horizontal de los datos seleccionados. Además, se aprecian, indicados con líneas punteadas, los tres elementos gráficos que esta visualización aporta.

En el caso de la Fig. 65 el color depende del número total de las actividades llevadas a cabo durante un día: cuantas más actividades realizadas contenga, más brillante será el relleno del rectángulo (designado por la frecuencia). También estos elementos se pueden dibujar en el mismo color, por ejemplo, cuando el objetivo de la exploración es el patrón temporal en general, en lugar de uno particular, como la persona o el patrón de una actividad específica. En ambos métodos de cálculo de dibujo, el color de las figuras se puede personalizar en todo momento del análisis. La altura de las barras depende de la cantidad de actividades del mismo tipo que se produjeron exactamente en el mismo minuto.

Finalmente, el tercer elemento de la Fig. 64 (indicado por las flechas de color gris) son las etiquetas de los períodos seleccionados. En estas dos representaciones lineales (horizontal arriba y vertical a la derecha en la figura) se encuentran en los extremos, para cada caso; arriba y abajo para la vertical e izquierda y derecha para la horizontal. Estas etiquetas informativas describen el período que se está representando con el siguiente formato:

**0000/00/00:** [año]/[número del mes del año]/[número del día del mes].

**XXX|XXX:** [nombre del mes del año] |[nombre del día del mes].

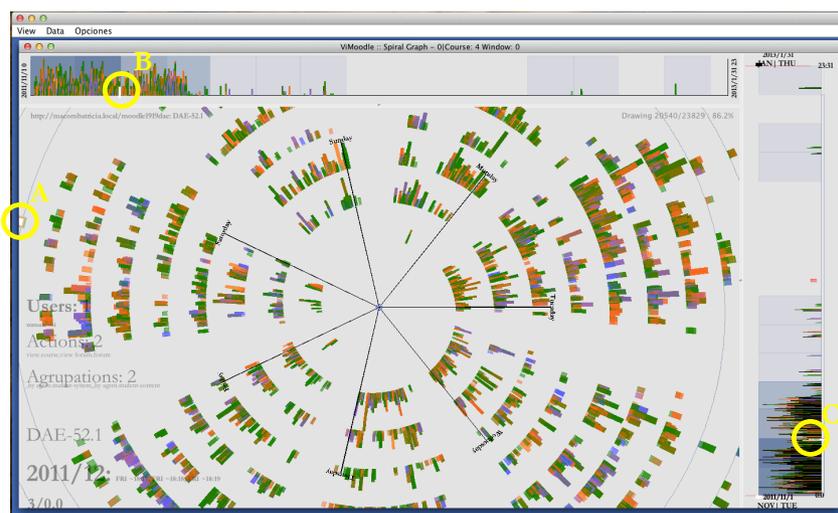


Fig. 65. Visión global de la línea de tiempo en espiral. Esta visualización muestra el aspecto general del patrón temporal del curso expresando el patrón semanal. Cuando se coloca sobre cualquier representación de una actividad, esta se resalta en blanco. Ténganse en cuenta los círculos (A, B y C) colocados sobre la imagen, que se han puesto para indicar cuál es la última actividad que se puede ver en la representación espiral sin deformación. Todos los círculos ponen de relieve el mismo momento de tiempo, pero se puede ver en la representación lineal en la parte superior (ver el círculo B de ese punto) que con la espiral solo puede verse la mitad de los datos, de modo que la deformación de la espiral es una técnica útil.

En la representación central de la visualización, la espiral, se utilizan las esquinas como espacios para mostrar etiquetas informativas (Fig. 65):

- En la esquina superior izquierda del panel de la espiral se muestra (aunque se puede ocultar) la ruta de la fuente de los datos.

- En la esquina superior derecha del panel de la espiral se muestra, en función de los datos seleccionados, la cantidad seleccionada, la cantidad total de los datos y el porcentaje que representa esta cantidad seleccionada de los datos con respecto al total en el siguiente formato:

**0000/00000 00.00%:** [cantidad de actividades mostradas]/[cantidad de actividades totales] [porcentaje proporcional]%;

- En la esquina inferior izquierda se expresa en el siguiente formato la información del objeto sobre el que se posiciona el ratón y puede ser una unidad de tiempo, ya sea un día, un mes o un año; y por otra parte un momento en el tiempo, es decir un minuto:

- Una unidad de tiempo (día, mes, año):

- día:

**0000/000:XXX:** [año]/[número de día del año]:[nombre del día de la semana]

**00.00% (0000 /0000)** [porcentaje proporcional con el total de actividades]% ([cantidad de actividades en el período resaltado] /[actividades totales])

- Mes:

**0000/00** [año]/[número de mes del año]

**00.00% (0000 /0000)** [porcentaje proporcional con el total de actividades]% ([cantidad de actividades en el período resaltado] /[actividades totales])

- Año:

**0000** [año]

**00.00% (0000 /0000)** [porcentaje proporcional con el total de actividades]% ([cantidad de actividades en el período resaltado] /[actividades totales])

- Un momento en el tiempo (un minuto):

- **Users:**

**000** [número de usuarios que realizan actividad]

**aaa..., bbb..., ccc...** [nombre de los usuarios]

- **Actions:**

**000** [número de actividades realizadas]

**aaa..., bbb..., ccc...** [nombre de las actividades]

- **Agrupations:**

**000** [número de agrupaciones de las actividades]

**aaa..., bbb..., ccc...** [nombre de las agrupaciones]

XXXX

[nombre del curso]

0000/00: xxx~00:00, ... [año]/[número de mes del año]:[nombre del día]~[hora]:[minuto]

- En la esquina inferior derecha se expresa la información del objeto (momento o unidad de tiempo) que se ha seleccionado mediante un *click* sobre él y se muestra con el mismo formato que la esquina inferior izquierda.

Este enfoque proporciona dos beneficios: primero, los rectángulos brillantes proporcionan una distinción previa del período más activo de tiempo, y en segundo lugar los diferentes rectángulos de color separan naturalmente períodos de más alto nivel de tiempo.

Cuando se coloca por encima de cualquier elemento, ya sea de una actividad, una unidad de tiempo, se repinta en blanco el objeto gráfico; como se puede observar en los círculos negros colocados sobre la Fig. 65. Estos señalan la última actividad representada, que se puede resaltar al colocarse sobre esta en la representación en espiral, para mostrar así sus correspondientes representaciones en la visualización horizontal lineal en la parte superior, y en la parte derecha en la visualización vertical lineal (ver los círculos negros de la imagen). De esta forma se evita la pérdida del foco y el contexto durante los ciclos de análisis y exploración de los datos temporales, solucionando una de las problemáticas mencionadas en la sección V.2.2.4, la Pérdida del contexto de navegación.

Además, con esto se puede reconocer que, con la espiral, limitados por el espacio, solo puede observar la mitad de los datos, de manera que la deformación es una técnica útil para el reconocimiento de patrones y optimización del espacio de representación. Dando, de esta forma, solución a una de las problemáticas presentadas en la sección V.2.2.2 y proporcionando mayor expresividad. En este mismo sentido, esta visualización cuenta con una técnica visual para equilibrar detalle y contexto en la visualización de datos, que se conoce como función de *zoom* semántico (Modjeska, 1997) o interfaces de múltiples escalas, además del tradicional *zoom* gráfico o físico. Esta técnica se realiza en la visualización pulsando dos veces sobre cualquier unidad de tiempo en cualquiera de las tres diferentes representaciones temporales y tiene como objetivo, seleccionar el período deseado y cambiar la representación actual por el período seleccionado, con la misma técnica y objetos pero diferentes datos temporales.

Las espirales planas más comunes son la espiral de Arquímedes, la espiral hiperbólica y la espiral logarítmica. La línea de los dos últimos se mueve rápidamente de distancia desde el origen. Se utilizó la espiral de Arquímedes porque se mueve lejos desde el origen a una velocidad constante y, por tanto, es más compacta. Para el resto de este prototipo cuando se usa "espiral" se refiere a la espiral de Arquímedes (Carlis & Konstan, 1998). Una espiral de Arquímedes se expresa en una ecuación polar:

$$r = \langle U \text{ donde } \begin{array}{l} x = r \cos U \\ y = r \sin U \end{array} \quad \begin{array}{l} r = \sqrt{x^2 + y^2} \\ U = \tan^{-1}(y/x) \end{array}$$

Se piensa en términos de un "eje de caracol", donde se enrolla en un eje para una variable de serie periódica en una espiral de manera que los períodos se alinean como se quiere que se haga. Si  $T$  es un eje y  $c$  es una constante, entonces el tamaño de  $t$ , un punto en  $T$ , su posición en el plano se puede encontrar usando:

$$x = c * t * \cos(t * 2 * \angle) \quad y = c * t * \sin(t * 2 * \angle);$$

El aspecto global de la espiral sin deformación se puede observar en la Fig. 65. Debido a que se deben representar grandes cantidades de información, se ha agregado un desplazamiento, esto es un valor de sangría ( $s$ ) que es aumentado al espacio en donde se dibuja la actividad ( $t$ ) puede ser suministrado por el usuario. Es decir,  $T = t + s$ . La sangría es la distancia que existe entre la barra más alejada del centro, el valor  $t$  (valor máximo de frecuencias en un mismo momento del tiempo) de un giro, y el giro siguiente de la espiral, esta se puede ver representada en la Fig. 66.

La línea de tiempo en espiral se puede adaptar a las necesidades del usuario, por lo que puede explorar todos los datos temporales disponibles, al pasar de la vista general a la vista del detalle (curso, persona, tiempo o actividad) dentro de un período de tiempo determinado. Esta visualización consta de tres paneles visuales internos que proporcionan diferentes vistas de los datos. La representación principal es la línea de tiempo en espiral, en la parte central, que en su forma más simple, no es más que una secuencia de eventos codificados por colores. Estos eventos se ordenan en sentido horario, en orden cronológico desde el centro, por lo que el evento más antiguo se representa en el centro y el evento más nuevo más alejado al centro.

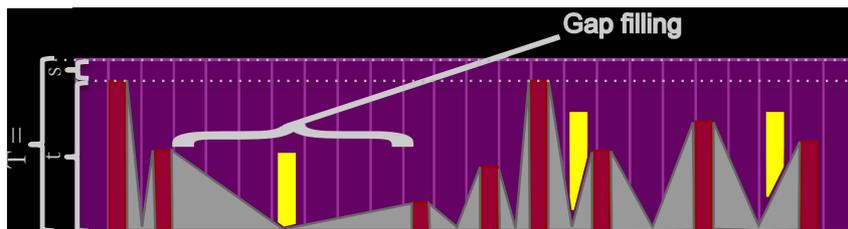


Fig. 66. Distorsión de la espiral, las barras rojas son las actividades, los polígonos grises representan los períodos de inactividad, que forman un polígono envolvente para cada vuelta de la espiral de  $360^\circ$  y las barras amarillas son las actividades del siguiente turno. Por último, la distancia total del espacio donde se despliegan las barras es  $T$ , y este está formado por la distancia de la barra más alta, denominada  $t$  más una sangría  $s$ .

Como se ha mencionado, esta representación temporal explora las actividades realizadas en una plataforma *eLearning*, siendo una visualización excelente para los analistas ya que coloca períodos similares alineados, facilitando su comparación.

En la Fig. 67 están representados los mismos datos de la Fig. 65, es decir, se han utilizado para representar un total de 91 estudiantes y se analizaron casi 55.000 eventos. Los datos para este estudio fueron recogidos durante un período de noviembre de 2011 hasta marzo de 2012. El análisis consiste en la selección, la unificación, la estandarización y el filtro de eventos registrados en un LMS para que se eliminen los mensajes en los foros que no son relevantes (como "sonrisas", "abrazo", "hola", etc.).

Sin embargo, en estas figuras se puede notar claramente la diferencia en la eficiencia del uso del espacio y, por tanto, algunos patrones temporales permanecen ocultos debido a la forma regular de la espiral (ver la Fig. 65). Por lo que se aplica una técnica de distorsión aplicada a la espiral (ver la Fig. 67) que es muy útil para desentrañar los patrones de evolución. La idea general detrás de la distorsión en espiral es la siguiente: en lugar de tener una forma espiral regular, con un aumento constante de su radio, la distancia desde el centro a la siguiente barra para dibujarse depende de la altura de las barras anteriores sobre el mismo ángulo.

La evolución de la espiral es muy similar a la que se observa en los cortes transversales de los árboles, en los que anillos registran la forma en que los árboles crecen bajo condiciones ambientales diferentes (Therón, 2006).

En el presente trabajo, las diferentes anchuras de las líneas concéntricas producen una forma similar a una espiral distorsionada, donde la variación de anchura se produce por la acumulación de diversas actividades por unidad de tiempo seleccionada. Esto se puede apreciar en Fig. 66. Las actividades se representan como barras de histograma dentro de la espiral y los espacios entre las barras de un giro a otro que se aprecian en la Fig. 65 en color gris y Fig. 64 en color negro, en los casos de la Fig. 66 y la Fig. 67 están cubiertas con las actividades de la siguiente vuelta de la espiral, en caso de que existan.

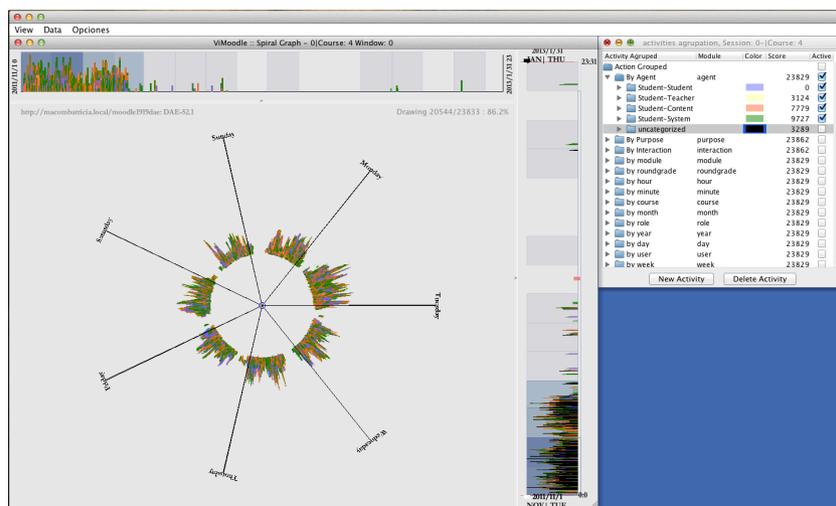


Fig. 67. Visión global de la representación de la línea de tiempo, esta visualización muestra el aspecto general del patrón temporal semanal del curso ya deformado en la espiral. Obsérvese la tabla de la derecha que muestra la agrupación de las actividades seleccionadas, donde se han excluido todas las actividades sin categoría agrupadas bajo la etiqueta *uncategorized*.

Esta deformación pierde la estructura de la línea de tiempo casi por completo, por lo que es difícil reconocer a qué período de tiempo pertenece cada actividad; sin embargo, aumenta el reconocimiento de patrones de tiempo al alinearse los períodos de tiempo similares en un mismo sector de la espiral, optimizando el espacio de la visualización y resaltando los patrones de frecuencia de actividad en el patrón de tiempo seleccionado. El resultado de la aplicación de este método para los mismos datos de la Fig. 65 se muestra en la Fig. 67.

Además, la línea de tiempo en espiral, con la finalidad de permitir una personalización de la exploración y análisis de los datos (Xiaoyu, Wenwen, Butkiewicz, Bier, & Ribarsky, 2011), habilita a los usuarios para crear, exportar e importar diferentes agrupaciones personalizadas de las actividades (*Student-Teacher, Passive, Create-Interaction* (Agudo-Peregrina et al., 2014), o cualquier otra agrupación personalizada que se quiera crear), y algunos de estos grupos se producen de forma automática (grado, rol, año, días, usuario, mes, etc.). Una vez que se han creado grupos de actividades, estos pueden utilizarse para seleccionar y excluir los datos del conjunto de datos analizados.

También puede ser útil cambiar el color de la agrupación con el fin de destacarla entre las demás, con el objetivo de facilitar el reconocimiento de tendencias. Por medio del menú contextual de la espiral o de la ventana de personalización y reagrupación de categorías de actividades (véase en la Fig. 67 la ventana de la derecha) se pueden delinear y rellenar todos y cada uno de los elementos de la representación visual, también se pueden elegir los filtros y selecciones de cada elemento representado y de cada agrupación, es decir, día, mes, año, persona específica, actividad específica, agrupación personalizada, etc.; además, se puede elegir el patrón a mostrar, es decir, que significado tendrá un giro completo de 360° grados, por ejemplo en la Fig. 67 abajo, el patrón elegido es semanal, por lo que el giro está dividido en 7 sectores que representan cada día de la semana, el patrón también puede cambiarse por anual, donde el giro completo de la espiral representará un año y, por tanto, se dividirá en 12 sectores y, por último, también se puede elegir el patrón diario, donde el giro entero representará las 24 horas del día y, por tanto, se dividirán los 360 grados en 24 sectores diferentes.

Esta técnica de interacción puede apreciarse con más detalle en el Capítulo VII en la sección VII.1.2, en el Caso de estudio 1. Sin embargo, todas estas técnicas y características mencionadas sobre esta representación de líneas de tiempo pueden percibirse con mayor claridad al consultar el vídeo de la URL: <http://analiticavisual.es/video-gallery/item/11-vela-vimoodle-may-2013>.

### VI.2.1.3 Representación de nube de palabras

La nube de etiquetas proporciona un medio para que los usuarios se formen una impresión general del conjunto subyacente de los contenidos y de la "esencia" de lo que se está tratando. Son una lista de las etiquetas que, por lo general, se muestran en orden alfabético y ponderadas visualmente en función de su popularidad. Tanto el orden como la ponderación pueden ser manipulados por el analista.

En las nubes de palabras la popularidad de una etiqueta se expresa por el tamaño de la fuente (en relación a las otras etiquetas) y, por tanto, se reconoce fácilmente (Hassan-Montero & Herrero-Solana, 2006); es frecuente que otras propiedades visuales, tales como el color de la fuente, la intensidad, o peso, también se manipulen (Bateman et al., 2008).

Esta propuesta de visualización está formada por un panel que contiene la nube y en la base en la cual se muestra una etiqueta de información contextual (palabra seleccionada/frecuencia de la palabra/total de las palabras) y en la esquina derecha una caja de búsqueda. Una etiqueta en la nube representa un cierto número de las etiquetas que se utilizan con mayor frecuencia en un área definida de la interfaz de usuario (ver la Fig. 68). En esta figura se representan dos diferentes plataformas de *eLearning*, en donde uno de los cursos cuenta con más de 4.500 palabras (a la izquierda) y el otro con más de 4.000 palabras (a la derecha).

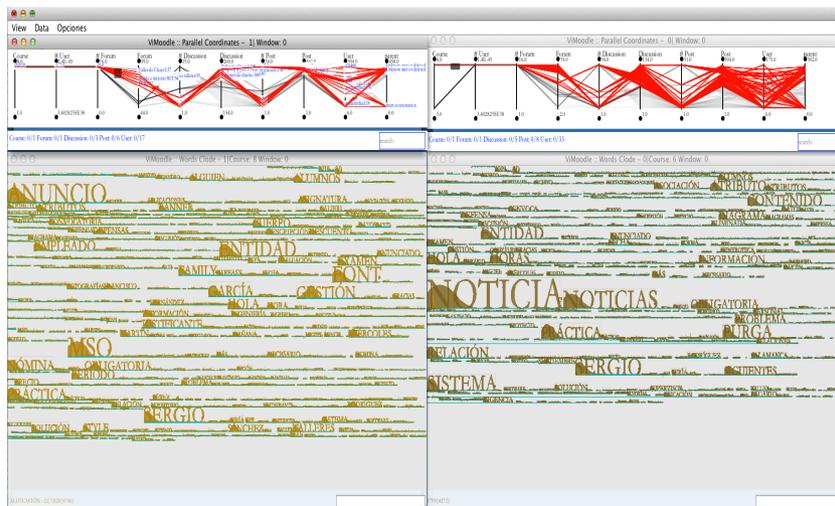


Fig. 68. Representaciones de las coordenadas paralelas y la nube de palabras/etiquetas. En esta imagen se puede observar la comparación de las nubes de palabras de dos cursos diferentes, en la derecha están representadas más de cuatro mil palabras y en la izquierda casi cinco mil palabras.

Las nubes de palabras pueden evolucionar a medida que los datos asociados cambian a través del tiempo. Por tal motivo, surge la necesidad de estudiar las tendencias y comprender cómo el contenido del texto o temas evolucionan con el tiempo; esto ha sido el propósito de otras visualizaciones (Bongshin et al., 2010; Weiwei et al., 2010).

Sin embargo estas propuestas, como se mencionó en las secciones V.2.2.2 y V.2.2.5, resultan deficientes en el uso del espacio, por tanto, para transmitir mejor la evidencia del cambio a través del tiempo de múltiples nubes de etiquetas, en esta propuesta se desarrolla un nuevo tipo de nube de etiquetas que integra en la etiqueta misma la información de evolución temporal de sí misma (ver Fig. 69). El diseño de esta propuesta se puede observar en la Fig. 69, esta es una visión general del diseño de una etiqueta, en donde se aprecian sus cinco componentes principales:

- El gráfico de barras.
- La gráfica de onda.
- El tiempo.
- La etiqueta.
- La información contextual.

La representación de la gráfica de onda (obsérvese la Fig. 69, las curvas azules y naranjas) y gráfico de barras (obsérvese la Fig. 69, bajo las curvas azules y naranjas) se resaltan líneas más oscuras que nacen en la base de la etiqueta y crecen hacia la parte superior) representan la misma información de significancia pero en diferente forma visual (barras o en onda para cada caso). El siguiente elemento, la etiqueta, representa mediante el tamaño de la fuente de la palabra la importancia general de esta en los foros y contenido compartido en el LMS.

Posteriormente, la información contextual (obsérvese la Fig. 69, el número en vertical al final de la palabra), expresa numéricamente la frecuencia con la que se encontró la palabra en cuestión, ayudando en el caso de dos palabras que resulten semejantes en su tamaño a distinguir con mayor facilidad cual es más representativa en un conjunto de datos analizados. Por último, el eje X a lo largo de la palabra codifica el tiempo total seleccionado y el eje Y codifica la significancia de las etiquetas en la nube de palabras. Todos estos elementos pueden ser visibles o no según las necesidades del usuario, es decir, se pueden dibujar u ocultar todos los elementos de la representación en la visualización por medio de los menús contextuales.

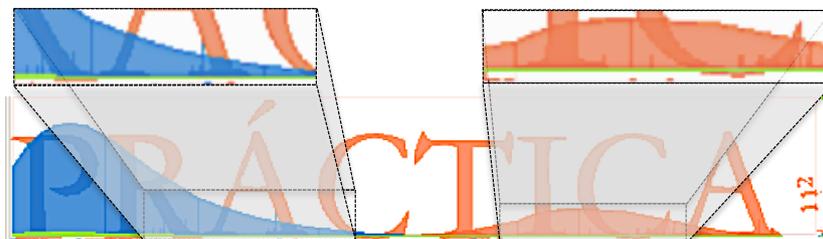


Fig. 69. En esta imagen se puede notar que la etiqueta “práctica” fue escrita al inicio del curso (expresado por la onda en azul sobre la palabra) y, sin embargo, fue leída hacia el final del curso (expresado por la onda en naranja).

Las curvas y las barras que se muestran en la Fig. 69 representan la importancia del contenido del documento, representado por una palabra, durante el tiempo, es decir, representan cómo de importante, dada su frecuencia, es una determinada etiqueta (que se representa con una palabra en el fondo de la Fig. 69) durante el período de tiempo seleccionado. Esta información se extrae de la información obtenida del contenido del curso y de las discusiones de mensajes en los foros de la plataforma, posteriormente se realiza un pre-procesamiento para eliminar palabras vacías (comúnmente llamadas en inglés, *stop words*).

Es importante mencionar, que este diseño, además de representar la evolución con mayor eficiencia espacial, aumenta la expresividad de la visualización, ya que no solo se centra en un aspecto de la evolución temporal de la etiqueta, sino que representa para mejorar aún más la expresividad y la legibilidad del significado de las curvas y las barras, por medio de la codificación de colores y transparencias, la evolución temporal de la etiqueta con respecto su actividad asignándoles diferentes colores.

Así bien, dependiendo la actividad, el verde se asignó a la actualización (en el caso de los foros el *post*, obsérvese la Fig. 69, la curva verde), el naranja para lectura de mensajes y contenido (obsérvese la Fig. 69, la curva

naranja), y el azul para la creación de mensajes, con respecto a la actividad de los foros, o contenido, con respecto a los recursos educativos del curso (para el caso de los foros el *add-post*, obsérvese la Fig. 69, las curva azul).

Además, habilitan al usuario tanto el *zoom* físico como el semántico. Si el usuario quiere saber más sobre el entorno de un contenido semántico en un período específico de tiempo, entonces puede seleccionar un período específico de tiempo, a través de la interacción manteniendo presionada la tecla Ctrl y moviéndose encima de cualquiera de las letras de una palabra, lo cual despliega unas etiquetas informativas del período que corresponda (ver Fig. 70), de esta forma se contextualiza a la persona en un determinado momento en el tiempo para ayudarle a la toma de decisiones y seguir explorando, ya sea con un *zoom* semántico o cambiando de estrategia. Después de explorar y elegir una palabra o letra (período de tiempo, representado por esta), el usuario puede hacer doble *click* en el período que decida, para obtener un *zoom* semántico.



Fig. 70. En esta figura se pueden observar las etiquetas informativas sobre el período que corresponde únicamente a la letra sobre la que el ratón se encuentra y la tecla Ctrl presionada.

Esta visualización se adapta a las necesidades del usuario, explorando todas las discusiones, foros, cursos, recursos y datos de los usuarios disponibles, y pasando de la vista general al detalle de una determinada persona, curso o discusión dentro de un período de tiempo determinado por la selección realizada en las otras visualizaciones. Cuando los usuarios utilizan el *zoom* semántico se analizan solo los mensajes y el contenido que se haya realizado, ya sea en el período seleccionado o por contener la etiqueta seleccionada, en función del contexto, de tal forma que la herramienta selecciona para su nuevo análisis el foro, comentarios y el contenido educativo relacionado con la palabra seleccionada o el tiempo seleccionado para la reconstrucción de la nube de palabras.

Por último, se puede mencionar que si de las palabras que aparecen en primera instancia, a pesar del pre-procesamiento que se realiza para eliminar las palabras vacías automáticamente, al analista le parece que algunas no son relevantes debido a que no reportan información de su interés (entre ellos se podrían encontrar los sinónimos, los plurales, las preposiciones, los números de identificación del estudiante y los nombres), aunque el prototipo descrito en este trabajo habilita manipular el uso de sinónimos y plurales, uniéndolos en una sola palabra a través del algoritmo de prefijos y sufijos ajustados al español (Porter, 1980), el analista puede eliminarlos manualmente a través de la interacción con los menús contextuales, logrando de esta manera que la representación optimice el espacio de las palabras de un foro o de un curso, y ayude a resaltar la importancia de cada etiqueta. Todas las características mencionadas de esta visualización pueden percibirse mejor en el vídeo: <http://analiticavisual.es/video-gallery/item/10-vela-vimoodle-abril-2011>.

#### VI.2.1.4 Análisis de redes sociales

Aunque los usuarios de los LMS han construido estructuras masivas de conectividad social, las interfaces web típicas de estas construcciones sociales siguen siendo relativamente pobres. Esto causa problemas a sus miembros para explorar su comunidad en línea y evaluar tanto la escala como los individuos a los que está expuesta su información personal de auto-medición y auto-reporte.

La visualización de la red social se utiliza para representar el mapa de relaciones, enlaces y frecuencias de actividades de los estudiantes, profesores y cualquier otro participante. Estas características se representan por medio del tamaño y el color de los iconos y las líneas (enlaces) en la misma gráfica. En este tipo de gráficos de nodos y enlaces se pueden representar la información mediante varias técnicas de analítica visual. Las más comunes de estas son el uso del tamaño de los nodos, la forma física de los nodos, el color de los nodos, la distancia entre un nodo y otro, el grosor de las líneas de enlace entre nodos, la posición de los nodos, etc. A continuación se describirán las características de la propuesta.

La visualización social de este estudio es un grafo dirigido por fuerzas y se interpreta como un gráfico de un sistema físico con fuerzas entre los nodos y que intenta minimizar la energía del sistema, encontrado un equilibrio, para obtener un dibujo estable. Tales algoritmos se utilizan para la elaboración de redes arbitrarias (parciales), como diagramas de flujo, gráficos de planificación de programas, gráficos de llamadas telefónicas, etc. (Kaufmann & Wagner, 2001). También se pueden aplicar a los diseños de agrupaciones (del inglés, *clusters*).

La posición de cada elemento en el gráfico de este prototipo se basa en la simulación de las fuerzas físicas que interactúan entre los nodos, que generan atracciones y repulsiones entre sí. Estas fuerzas dependen del peso de los elementos, del número de relaciones que cada uno de ellos tiene y, a su vez, de las métricas que se representen por medio de los enlaces internos y externos (*internal-spring* y *external-spring* desarrolladas más ampliamente en las secciones 8.5.2 y 8.5.3 de Kaufmann and Wagner (2001) para crear las agrupaciones en los gráficos de redes sociales). En la visualización se implantaron como nodos algunas entidades codificadas, obsérvese la Fig. 71. En esta figura se pueden distinguir cómo se verían los nodos que representan a las personas, los cursos, los foros, los comentarios y las discusiones. Además, el tamaño de estos iconos depende de su grado de entrada y salida y de su profundidad en la jerarquía, que se explican a continuación.

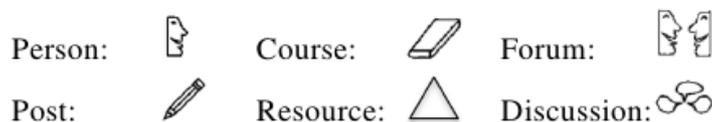


Fig. 71. Codificación gráfica de los nodos que representan los elementos de la red social (Gómez Aguilar, Theron, & García Peñalvo, 2008).

Los grados de entrada se refieren al número de enlaces hacia el nodo en cuestión (concretamente recibir una contestación a un *post* realizado, estar matriculado en un curso o inmerso en una discusión, ser observado en el perfil de usuario, entre otras) y los grados de salida son aquellos enlaces desde el nodo (específicamente sería la actividad realizada por el individuo como, por ejemplo, leer un recurso, escribir un *post*, leer una discusión entre otras).

Cabe mencionar que estos grados están directamente relacionados con la frecuencia de interacción del usuario, de forma que repercute directamente de varias formas en el grafo: una es modificando la distancia de los nodos hacia el punto central (el curso), para quedar más alejados del nodo del curso aquellos nodos menos activos; y la otra es repercutiendo en el tamaño del nodo, resaltándose así visiblemente más grandes los cursos más activos, foros más concurridos, etc.

Por otra parte, en la red social, los diferentes nodos se pueden ocultar o cambiar de color, de forma individual o en grupo. En el caso del nodo de las personas, hay dos tipos como se puede ver en la Fig. 72, por una parte profesores (en azul) y por otra estudiantes (rojo a la izquierda y naranja en la parte derecha de dicha figura). Además, en esta visualización es posible ocultar los enlaces, así se logra mostrar únicamente los aspectos y nodos deseados, lo que permite optimizar el espacio y la claridad en la exploración de los datos, deficiencia mencionada en la sección V.2.2.5. Por ejemplo mostrar únicamente las personas y las líneas que van a los nodos que hayan sido iniciadores de conversación (ver la Fig. 72), o mostrar únicamente las líneas de interacción (con diferentes grosores o distancias en función de la frecuencia de interacción) entre las diferentes personas del curso (ver la Fig. 73), para identificar así fácilmente los elementos centrales y/o grupos sociales que se forman.

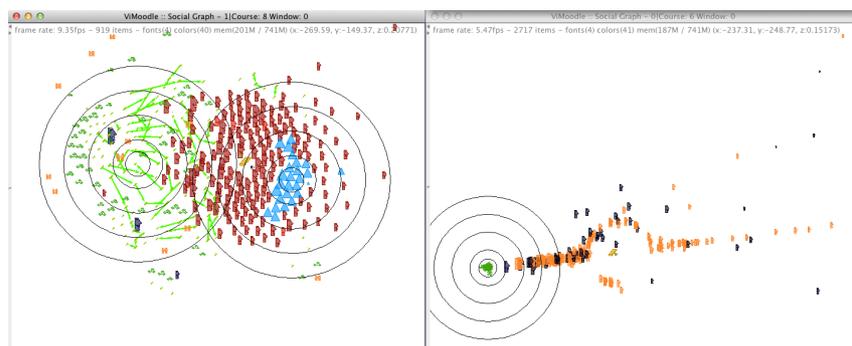


Fig. 72. Representación de la red social. A la izquierda se representa un curso completo y se muestran todos sus elementos sin activar los enlaces internos y externos (procesándose junto con los elementos invisibles más de novecientos elementos entre enlaces y nodos) y a la derecha un curso en el cual se han ocultado todos los elementos con excepción de las discusiones y los participantes, con el fin de resaltar una métrica de LA (cuanta más participación en discusiones, mejor rendimiento) por medio de los enlaces internos y externos (procesándose junto con los elementos invisibles más de dos mil elementos entre enlaces y nodos). Dado que las distancias en representaciones que no son distribuidos los elementos en una línea, sino depende de las fuerzas de atracción y repulsión, se pueden desplegar, o no, los círculos concéntricos que se ven en la figura, siendo estos una ayuda visual que permite la comparación de las distancias entre dos nodos.

Por último, dado que se ha demostrado en otros estudios que hay relación entre la frecuencia de la lectura y la escritura en los foros con el rendimiento de los estudiantes, y entre la frecuencia de los recursos de lectura y el rendimiento de los estudiantes (Agudo-Peregrina, Hernández-García, & Iglesias-Pradas, 2012; Lipponen et al., 2003), solucionando el problema mencionado en la sección V.2.2.2, a través del menú contextual, el usuario puede elegir que se represente por medio de los *springs* (enlaces invisibles que implementan fuerzas de atracción y repulsión entre los nodos, mencionados anteriormente) la distribución de los estudiantes en función de dichas relaciones.

Aunque parece simple este tipo de representaciones de las métricas (ver Fig. 72 a la derecha), y sobre todo las que se han corroborado como significantes en el rendimiento del estudiante, si se observa bien, los nodos de los participantes se dispersan en la figura de la derecha partiendo desde el centro de los círculos concéntricos (agrupación creada de todas las discusiones por medio de *springs*) hacia el exterior de estos. Esta dispersión no es casual ya que los nodos más cercanos son los que más han participado, mientras que los más alejados los que menos lo han hecho, relación que frecuentemente es directamente proporcional a su nota final (en caso de tener el dato) o, como es el caso de la figura, una posible predicción del rendimiento de los participantes.

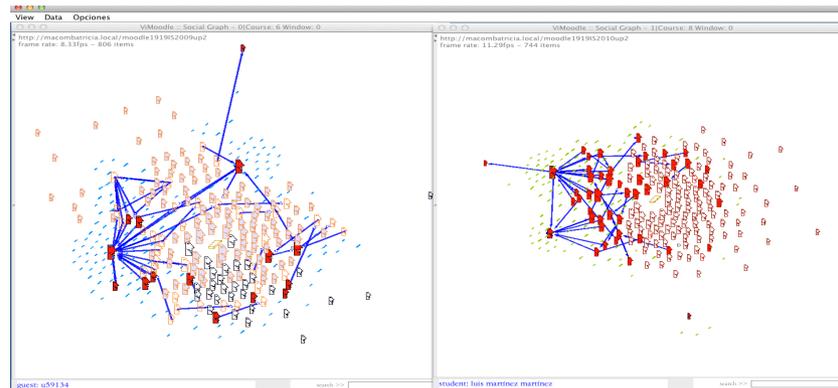


Fig. 73. Representación de la red social, dotada de funcionalidad de búsqueda, un cuadro de texto en la base de la visualización. En esta representación se muestran dos cursos diferentes, ocultando todos los nodos con excepción de los participantes y mostrando, por medio de las líneas, las relaciones de comunicación entre los participantes. La distancia de estos representa la frecuencia de participación entre ellos.

Finalmente, en la Fig. 73 se puede distinguir, gracias a los diferentes métodos de interacción del prototipo, las relaciones de participación de los involucrados en el curso. En ambas representaciones de red social, las líneas azules representan la interacción existente entre dos personas, y su distancia la frecuencia de esta. Por tanto, en el lado derecho se puede percibir la existencia de un participante dominante en color rojo en el cuadrante de abajo y a la izquierda, y, por otra parte, en la red representada a la derecha, la participación está distribuida entre dos personas.

Sin embargo, es importante resaltar que no solo se muestra en el gráfico de red social la participación en los foros, caso de las líneas azules de esta figura, sino también la actividad de lectura, la cual influye en el tamaño de los nodos, por esta razón en ambas representaciones existen nodos de un

tamaño casi igual a los participantes dominantes que se han mencionado, pero que no están conectados con líneas azules. Es importante mencionar que, dado que las distancias en representaciones que no son distribuidos los elementos en una línea, sino depende de las fuerzas de atracción y repulsión, el analista pueden desplegar -o no- los círculos concéntricos que se ven en la figura, siendo estos una ayuda visual que permite la comparación de las distancias entre dos nodos.

Todas estas características e interacciones habilitadas para el analista se pueden encontrar descritas más detalladamente en la sección VII.1.3 y en el vídeo: <http://analiticavisual.es/video-gallery/item/13-vela-vimoodle-jul-2013>.

Esta propuesta está dedicada a la solución del problema de análisis de una gran cantidad de información que se asocia a los procesos de enseñanza y aprendizaje mediados por plataformas tecnológicas. Este enfoque proporciona nuevos conocimientos útiles a través de una herramienta de *software* eficiente que soporta el pre-procesamiento de la información, el análisis, la visualización y la interacción. Además, se utiliza de forma independiente de la versión del LMS, y puede utilizarse *offline*, *online*, en web o como aplicación de escritorio.

Este prototipo de herramienta está dirigido tanto a estudiantes, como profesores y gestores de una institución educativa. Los datos a los que el usuario puede acceder dependen de los permisos que tenga en el LMS, y estos pueden variar desde un curso hasta un campus universitario completo. Una vez descrito el prototipo final del sistema a continuación se describe la metodología empleada en el proceso de evaluación y diseño utilizada durante el desarrollo del sistema del prototipo de VeLA.

## VI.2.2 Proceso de evaluación del diseño del prototipo

Como se mencionó en la introducción, en la presente tesis se ha utilizado la metodología de investigación-acción. Dentro de sus bases, Pring (2004) señala cuatro características significativas de la metodología de investigación-acción: cíclica/recursiva; participativa; cualitativa; y reflexiva. Además, menciona que un rasgo específico de la investigación-acción es la necesidad de integrar la acción, por lo cual, el foco de la investigación será el plan de acción para lograr el cambio o mejora de la práctica o propósito establecido; es decir, «Hacer algo para mejorar una práctica», esto es un rasgo de la investigación-acción que no se da en otras investigaciones.

Este tipo de metodología se adapta y coincide con la metodología de investigación basada en el diseño. Para diseñar y desarrollar el sistema del prototipo de VeLA, se aplicó la metodología de investigación basada en el diseño (Obrenovic, 2011), que está basada en la creación rápida de prototipos para evaluar las ideas en ciclos.

De esta forma, y siguiendo las metodologías antes mencionadas, el sistema del prototipo del modelo VeLA se desarrolló durante cuatro iteraciones, con una duración en conjunto de más de 24 meses. Antes de cada iteración se estableció una versión del prototipo y un caso de estudio, resultando tres versiones del prototipo de VeLA, cuatro iteraciones y cuatro casos de estudio, para un mejor entendimiento observe la Tabla 11.

Nombre   Versión	Objetivo	Técnicas y estrategias de Visualización
prototipo del modelo VeLA   0	Entender las relaciones que se crean en la interacción realizada en Moodle (M. Á. Conde González et al., 2010; Gómez Aguilar, Theron, et al., 2008; Gómez Aguilar, Therón Sánchez, & García Peñalvo, 2008; Gómez Aguilar, Therón Sánchez, & García Peñalvo, 2007a, 2007b)	<ul style="list-style-type: none"> <li>✓ Gráfica de caracol <sup>z, cc</sup></li> <li>✓ Nube de palabras <sup>z, zs</sup></li> <li>✓ Gráfico de red social <sup>cc</sup></li> <li>✓ Codificación por color (CC)</li> <li>✓ Zoom(Z)</li> <li>✓ Zoon semántico (ZS)</li> <li>✓ Vistas enlazadas</li> </ul>
<b>Caso de estudio 0</b>		
<b>Iteración 1: 30 Jul 2010</b>		
prototipo del modelo VeLA   1	Además de la comprensión de la evolución temporal de las actividades realizadas (Conde González, Gómez Aguilar, et al., 2011; García Peñalvo, Conde González, Bravo Martín, Gómez Aguilar, & Therón Sánchez, 2011a, 2011b; Gómez Aguilar, Conde González, Therón, & García Peñalvo, 2010a, 2010b; Diego Alonso Gómez Aguilar, M. Á. Conde González, Roberto Therón Sánchez, & F. J. García Peñalvo, 2011; Diego Alonso Gómez Aguilar, Miguel Ángel Conde González, Roberto Therón Sánchez, & Francisco José García Peñalvo, 2011; Gómez Aguilar, Conde-Gonzalez, et al., 2011; Gómez Aguilar, Suárez Guerrero, et al., 2010; Gómez Aguilar et al., 2009; Gómez Aguilar, Therón Sánchez, & García Peñalvo, 2008; Gomez-Aguilar, Conde-Gonzalez, Theron, & Garcia-Peñalvo, 2011)	<ul style="list-style-type: none"> <li><del>✓ Gráfica de caracol <sup>z, ee</sup></del></li> <li>✓ Gráfica de espiral <sup>z, zs</sup></li> <li>✓ Manejo de sinónimos, plurales, eliminación de etiquetas en las nubes de palabras</li> <li>✓ Datos exportables para análisis <i>offline</i></li> </ul>
<b>Caso de estudio 1</b>		
<b>Iteración 2: 30 Jul 2012</b>		
prototipo del modelo VeLA   2	Permitir la comparación entre dos o varios cursos (Conde González, García Pealvo, Gómez Aguilar, & Therón Sánchez, 2014; Gómez Aguilar, García Peñalvo, et al., 2013a, 2013b; Gómez Aguilar, García Peñalvo, & Therón, 2014; Gómez Aguilar, García Peñalvo, & Therón Sánchez, 2014; Gómez Aguilar, García Peñalvo, & Therón Sánchez, 2014; Gómez Aguilar, Therón, et al., 2013)	<ul style="list-style-type: none"> <li>✓ Gráfica de espiral <sup>z, zs, cc, a</sup></li> <li>✓ Nube de palabras <sup>z, zs, cc, a</sup></li> <li>✓ Gráfico de red social <sup>z, cc, a</sup></li> <li>✓ Agrupaciones (A)</li> <li>✓ Coordenadas paralelas <sup>z, zs, cc, a</sup></li> </ul>

Nombre   Versión	Objetivo	Técnicas y estrategias de Visualización
<b>Caso de estudio 2:</b>		
<b>Iteración 3: 10 Feb 2013</b>		
<b>prototipo del modelo VeLA   3</b>	Aumentar el aumento del alcance no solo a la comparación de cursos individuales en cada conexión y visualización, sino a una universidad entera (Gómez-Aguilar, Hernández-García, García-Peñalvo, & Therón, 2015)	✓ Personalización de las agrupaciones de las actividades.
	<b>Caso de estudio 3:</b>	
<b>Iteración 4: 10 Feb 2014</b>		
<b>Caso de estudio 4:</b>		

Tabla 11 Cuadro general del proceso de diseño del prototipo del modelo VeLA y las principales técnicas implementadas en cada versión. En esta tabla se puede observar el objetivo a lograr en cada mejora del prototipo y el aumento de técnicas de visualización en cada una de las diferentes versiones (léase la última columna de la derecha de arriba hacia abajo, es importante mencionar que solo se ha escrito lo que se ha incrementado o eliminado, el resto de características continúan igual que en la versión anterior) y en cada una de estas se ha indicado el nombre de la o las técnicas que se han implementado específicamente para la visualización por medio de superíndices, por ejemplo *zoom* semántico en la nube de palabras estaría como nube de palabras<sup>28</sup>.

Las figuras presentadas en este trabajo muestran el prototipo final del modelo VeLA (versión 3 de la tabla de diseño, Tabla 11, e iteración 4 en la de iteraciones, Tabla 12). El prototipo desarrollado representa un sistema inminentemente interactivo, es decir, la parte de interfaz tiene un peso especialmente importante y va a ser clave la usabilidad final de la herramienta. La Ingeniería de la Usabilidad (Scholtz, 2004) provee métodos estructurados para alcanzar dicha usabilidad. La usabilidad de un producto de *software* es un aspecto fundamental que se evalúa a través de pruebas que involucran la interacción de los usuarios con el sistema.

Parte de la evaluación de este tipo de pruebas se lleva a cabo generalmente mediante el uso de cuestionarios con una serie de preguntas específicas sobre la experiencia del usuario con el sistema y, a través de las respuestas de los usuarios, los diseñadores del sistema pueden advertir en una etapa temprana de la fase de desarrollo, sobre los posibles problemas y conflictos que pudieran haber pasado inadvertidos.

En el contexto específico de los sistemas de analítica visual de información educativa es fundamental realizar diversas evaluaciones de usabilidad, que aseguren que estos sistemas sean, por un lado, fáciles de aprender, pero también efectivos, funcionales y eficientes desde la perspectiva del usuario. Todos estos aspectos permiten asegurar la usabilidad de los productos de *software* en general y, en particular, también de los sistemas de *Learning Analytics*, *Academic Analytics* y, por tanto, del sistema del prototipo de VeLA. En función del tipo de sistema e interacción con el usuario se han definido diversos métodos de evaluación. La técnica de evaluación heurística es el método más ampliamente utilizado, ya que las evaluaciones centradas en el usuario suelen ser menos costosas en cuanto a tiempo. Más aún, es una excelente forma de obtener una retroalimentación directa de los usuarios. Nielsen (1994) propone el uso de una técnica de evaluación heurística, que requiere un conjunto de especialistas que evalúen y juzguen que una interfaz de usuario cumpla con los principios del diseño de usabilidad.

Scholtz (2004) resume los diversos tipos de evaluaciones, entre los que se encuentran las evaluaciones basadas en expertos y las evaluaciones basadas en modelos. Uno de estos métodos es una evaluación centrada en el usuario (del inglés, *user-centred evaluation*), en la que los usuarios son la principal fuente de retroalimentación. Para la evaluación de la usabilidad se ha tomado el enfoque de una evaluación de usabilidad centrada en el usuario y en expertos de Scholtz (2004), mediante la aplicación de cuestionarios a los participantes de la evaluación para el caso de los usuarios y para el caso de los expertos mediante una entrevista para el intercambio de ideas, mejoras y errores encontrados.

Todas las técnicas se han implementado tomando en consideración la capacidad cognitiva de una persona, por lo que la evaluación de usabilidad también debe considerar a la persona como el centro del proceso. Este tipo de evaluación centrada en el usuario y en expertos es consecuencia de la propia naturaleza de la herramienta, ya que, como se ha descrito previamente, el diseño del sistema del prototipo y modelo de VeLA está completamente orientado al usuario como el centro de los ciclos interactivos del análisis.

#:Fecha	Información demográfica	Metodología	Objetivo	Adaptaciones al diseño	Resultado de evaluación
1: 30 Jul 2010	Profesor de la Universidad de León del Área de Arquitectura y Tecnología de Computadores del Departamento de Ingenierías Mecánica Informática y Aeroespacial	Se le presentó la herramienta y se les entrevistó para evaluarla y discutir sobre los puntos fuertes, innovaciones, problemas y posibles mejoras	Evaluación de usabilidad, alcance del análisis y satisfacción del usuario, evaluación del alcance, prueba de comparaciones entre cursos, satisfacción del usuario y comentarios de mejoras	Primera implementación, conformada de una línea de tiempo en caracol, una nube de palabras y una red social que mostraba, de un curso, la estructura, contenido y actividad de los diferentes perfiles de la plataforma	<b>Negativa:</b> Problemas de usabilidad -Conexión a los datos por DAO posible causa de incompatibilidades con versiones de Moodle superiores, solo social -limitado para comparar entre cursos o ver métricas de cursos, tiempos de respuesta muy lentos, problemas para interpretar el significado de la espiral, problemas para interpretar las líneas temporales en el la nube de palabras <b>Positiva:</b> Rápido aprendizaje, alta expresividad de las visualizaciones y de información útil
2: 30 Jul 2012	Profesor de la Universidad Politécnica de Madrid del Departamento de Ingeniería de Organización, Administración de Empresas y Estadística	Se le presentó la herramienta y se les entrevistó para evaluarla y discutir sobre los puntos fuertes, innovaciones, problemas y posibles mejoras	Evaluación de usabilidad, alcance del análisis y satisfacción del usuario, prueba de comparaciones entre cursos, satisfacción del usuario y comentarios de mejoras	Se optimizaron las estructuras de datos, se mejoraron los tiempos de respuesta, se implementó la manipulación de plurales y la eliminación de etiquetas en las nubes de palabras, así como su importación y exportación desde y a archivo, se cambió la conexión a los datos por servicios web, incrementando la compatibilidad de futuras versiones, se implementó la línea de tiempo en espiral y la clasificación de las actividades por mes, estudiante, rol, nota y período temporal, se implementaron las coordenadas paralelas en dos niveles semánticos, de usuarios y de cursos, para mostrar métricas	<b>Negativa:</b> Ayuda en la comprensión de la representación, habilitar la personalización de las categorías, mejora en las etiquetas de información contextual, poca claridad en la red social y representación muy básica en comparación a las posibilidades en la analítica de redes sociales educativas (SLA) <b>Positiva:</b> Análisis más completo de la plataforma, alta interacción y exploración de los datos, portabilidad y flexibilidad, análisis temporal, semántico, social y métricas útil
3: 10 Feb 2013	15 estudiantes del Máster Universitario en Tecnologías de la Información y de la Comunicación en Educación impartido en la Universidad de Salamanca	Se les impartió una charla relacionada con la analítica del aprendizaje, analítica académica y analítica visual, finalmente, se les mostró la herramienta y se les explicaron algunos detalles del diseño de las visualizaciones, así como una introducción general de la interacción con los usuarios	Evaluación de usabilidad, utilidad y facilidad de aprendizaje de utilización y la posible auto-evaluación	Se implementaron algunas métricas en la representación social sobre las correlaciones entre rendimiento y actividad de los estudiantes, se implementó la personalización de las categorías de actividades, se agregaron etiquetas informativas contextuales	<b>Negativa:</b> Algunos problemas de adaptación y dificultad para interpretar los datos (resueltos con una explicación de la representación) <b>Positiva:</b> Muy útil interesante y adaptable
4: 10 Feb 2014	17 estudiantes del Máster Universitario en Tecnologías de la Información y de la Comunicación en Educación impartido en la Universidad de Salamanca	Se les impartió una charla relacionada con la analítica del aprendizaje, analítica académica y analítica visual, finalmente, se les mostró la herramienta y se les explicaron algunos detalles del diseño de las visualizaciones, así como una introducción general de la interacción con los usuarios	Evaluación de usabilidad, utilidad y facilidad de aprendizaje de utilización y la posible auto-evaluación	Se resolvieron algunos temas pendientes de usabilidad se aumentaron algunas métricas de iniciador de discusión para los estudiantes	<b>Negativa:</b> Algunos problemas con la instalación de la herramienta (falta de JRE en las máquinas) <b>Positiva:</b> una herramienta que cubre una carencia actual en la educación en línea

Tabla 12 Cuadro general de iteraciones de diseño del prototipo del modelo VeLA y los principales resultados de la evaluación.

La diferencia con la evaluación basada en expertos es que este último esquema involucra un reducido número de usuarios con un alto grado de conocimiento y muy específico, mientras que el esquema basado simplemente en los usuarios, involucra a un mayor número de usuarios sin un alto grado de especialización. Estas evaluaciones, ya sea por entrevista en el caso de expertos o por cuestionario en el caso de usuarios, se realizaron al final de cada iteración de la presente investigación. La Tabla 12 presenta una visión general de todas las iteraciones y los principales resultados de la evaluación. Todas las evaluaciones se han realizado en entornos del mundo real y han involucrado cursos reales en una variedad de temas, sus profesores y estudiantes.

Este tipo de evaluaciones se basan en las siguientes actividades: la identificación de los usuarios representativos, la identificación y el desarrollo de las tareas representativas y, finalmente, el desarrollo de un cuestionario que permita al usuario otorgar una evaluación de los diversos aspectos de la herramienta. Esta evaluación de la usabilidad se presentará después de haber descrito las características del prototipo de VeLA, sus diversas representaciones visuales (en la siguiente sección) y de haber detallado los casos de estudio (presentados en el capítulo siguiente), permitiendo así una comprensión más amplia de este. A continuación se describen las actividades realizadas para llevar a cabo el proceso de diseño del prototipo presentado en este trabajo.

La primera actividad consiste en la identificación de aquellos usuarios representativos que participan en la evaluación del prototipo. Para el desarrollo de esta tarea es necesario determinar quiénes son los usuarios representativos del sistema del prototipo de VeLA. Este segmento de usuarios está formado por todos aquellos, principalmente profesores y gestores académicos, interesados en analizar, entender y optimizar el proceso de enseñanza y aprendizaje.

En este contexto, los usuarios de la herramienta son, por un lado, los estudiantes interesados en autoevaluarse y mejorar su aprendizaje y, por otro, aquellos con un cierto grado de especialización, o al menos con cierto conocimiento de analítica del aprendizaje, que en este segundo caso serían algunos gestores académicos y profesores.

Para el desarrollo de esta evaluación se utilizó, por un lado, como segmento de usuarios un total de treinta y dos estudiantes del Máster Universitario en Tecnologías de la Información y de la Comunicación en Educación impartido en la Universidad de Salamanca. A estos estudiantes se les impartió una charla relacionada con la analítica del aprendizaje, la analítica académica y la analítica visual, finalmente, se les mostró la herramienta y se les explicaron algunos detalles del diseño de las visualizaciones, así como una introducción general de la interacción con los usuarios. Finalmente, se les entregó un cuestionario que incluía 30 preguntas (ver el anexo en la sección IX.1) sobre usabilidad, practicidad y eficiencia del sistema, expuesto a los treinta y dos estudiantes participantes.

El grado de conocimiento de los estudiantes sobre LA, AA, EDM o VA no era muy especializado, sin embargo, al ser estudiantes de nivel de máster universitario, todos ellos tenían conocimiento sobre el uso de las nuevas tecnologías en y para la educación.

Por otro lado, como segmento de expertos se tomó a dos profesores: uno de la Universidad de León, del Área de Arquitectura y Tecnología de Computadores del Departamento de Ingenierías Mecánica Informática y Aeroespacial y el segundo de la Universidad Politécnica de Madrid, del Departamento de Ingeniería de Organización, Administración de Empresas y Estadística. Ambos con conocimiento en las áreas de LA y AA a los que se les presentó la herramienta y se les entrevistó para evaluarla y discutir sobre los puntos fuertes, innovaciones, problemas y posibles mejoras.

Una vez definido el segmento de los usuarios, la segunda actividad en el esquema de evaluación corresponde a la identificación de las tareas representativas. Estas tareas deben estar siempre enfocadas a los aspectos específicos que intentan evaluarse; en este caso, los tres aspectos más importantes a evaluar son la efectividad, la eficiencia y la interacción con el usuario. La efectividad es un objetivo muy general y se refiere a lo bueno que es un producto al hacer lo que se supone que tiene que hacer. Por otro lado, la eficiencia se refiere a la forma en la que un producto apoya a los usuarios en el desarrollo de sus tareas (Lazar, Feng, & Hochheiser, 2010; Preece, Rogers, & Sharp, 2001).

Finalmente, la interacción con el usuario se refiere a los mecanismos con los que el usuario puede ejecutar acciones sobre el modelo visual, lo que representa un aspecto fundamental en cualquier sistema de *software* y, especialmente, en el sistema de visualización analítica aplicada al *eLearning*.

Teniendo claro el diseño final del prototipo el proceso de evaluación y desarrollo de este, en el siguiente capítulo se describirán los casos de estudio desarrollados antes y después de cada iteración del proceso anteriormente descrito. Esto es, si son tres iteraciones (ver Tabla 11) y un caso de estudio antes y después de cada una, en el siguiente capítulo se describirán los casos de estudio comenzando con el caso cero anterior a la iteración uno y finalizando con el caso cuatro realizado después de la cuarta iteración.





A continuación, para cada uno de estos casos de estudio del sistema, se detalla el ambiente de análisis, es decir, el conjunto de datos a analizar con la última versión del sistema, posteriormente se expone el problema y las preguntas a responder, se desarrolla la solución planteada y los resultados obtenidos.

## VII.1 Descripción del prototipo a través de casos de estudio

Como se ha mencionado este desarrollo se ha realizado de acuerdo a las iteraciones y las versiones del sistema, ver Tabla 11, por lo que se han desarrollado diversos casos de estudio en diferentes etapas e iteraciones en el orden siguiente: caso de estudio 0 de la versión 0 analizado en la primera iteración, caso de estudio 1 después de la primera iteración, caso de estudio 2 al finalizar la segunda iteración y, por último, un caso de estudio para cada una de las iteraciones con los usuarios, iteraciones 3 y 4, casos de estudio 3 y 4. En cada caso de estudio se detallarán los datos analizados y los resultados obtenidos.

### VII.1.1 Caso de estudio 0

En su versión cero el sistema presentó tres visualizaciones, las cuales fueron un gráfico de red social (SNG), un gráfico de caracol (del inglés, *Snail Graph*, SG) y una nube de palabras (del inglés, *Words Cloud*, WC), las cuales pueden observarse en la Fig. 74. Los datos analizados se corresponden a 45 tipos diferentes de actividades, 243 estudiantes, 23558 actividades, que se realizaron en 2 diferentes cursos en un período de 2 años y medio. Los foros se ven representados en el WC, que muestra las palabras más relevantes de estos. Aquí, el usuario puede elegir entre diferentes vistas dependiendo de los datos que quiera analizar, eligiendo entre las palabras clave: los usuarios, los cursos, los debates y los temas de mensajes en el foro, obtenidos a partir del análisis de todo el contenido semántico y la información intercambiada en la plataforma.

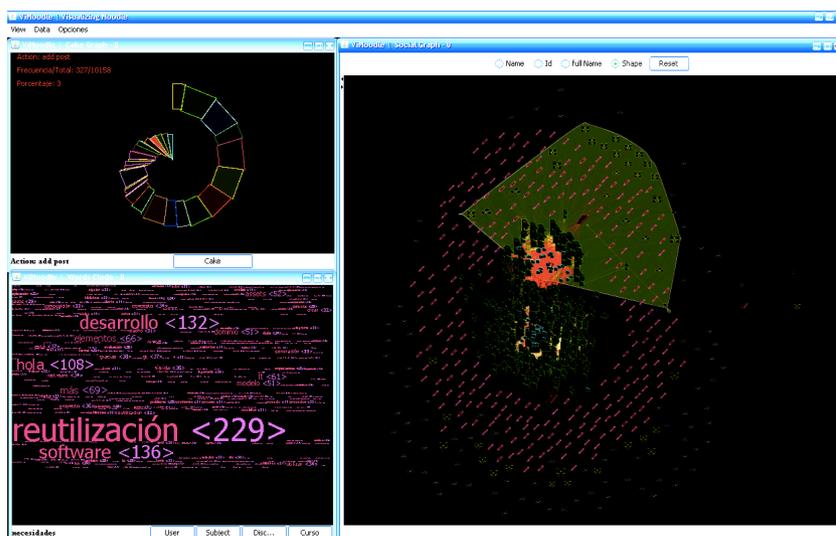


Fig. 74. Visualización global de las relaciones entre los usuarios.

En la Fig. 74 a la derecha se encuentra la SNG que representa un mapa de relaciones y la frecuencia de las actividades de los estudiantes y profesores. Las líneas entre los iconos y los tamaños relativos de los iconos reflejan estas relaciones y sus frecuencias. Otra vista del sistema en la misma figura es el SG, que crea un mapa general de la frecuencia de las acciones realizadas por los usuarios en la plataforma organizadas por orden alfabético.

Pensando como un estudiante nuevo, se intentó identificar a una persona con una alta frecuencia de actividad, posiblemente la persona más accesible en los foros, y que sabe acerca de la palabra "Dominio". Con la ayuda del filtro de conectividad: "grado de separación" (del inglés, *degree of separation*, DOS), es decir, el número de pasos que separan a dos personas en una red.

En la Fig. 74, donde se representan las relaciones entre las personas no solo en la creación de conversaciones sino también en las visitas a los perfiles y lecturas de los *posts*, la información de la SNG es de tal magnitud que es imposible llegar a la meta propuesta. Por tanto, es necesario aplicar un primer filtro de los datos: en el SG se hace *click* en un sector (blanco Fig. 75, izquierda) el filtrado de todas las actividades, excepto "agregar post". El resultado se muestra a la derecha de la Fig. 75 (observe se el cambio de los números y los tamaños de fuente de las palabras en comparación con los de la Fig. 74) filtrando y mostrando los elementos más activos. Ahora, se hace *click* en la palabra "Dominio" (obsérvese el color blanco de la palabra) con el WC (Fig. 75, derecha). Como resultado se muestra la SNG solo para aquellos usuarios que han utilizado esa palabra (Fig. 76).

El conglomerado verde que se puede ver en la SNG (Fig. 76) se utiliza para mostrar el curso que tiene el mayor volumen de datos (es decir, número de enlaces en Moodle).

En la red que se muestra en la Fig. 76, se observa a una persona con la mayor actividad, representado por el icono de una "cara" de tamaño mayor al resto. Una información en la descripción emergente sobre la herramienta muestra que es Francisco. La SNG sigue siendo compleja y se puede explorar aún más. Ahora se tiene interés en el DOS de la red y en la comparación de este en la SNG entre las personas.



Fig. 75. Filtro de Actividad (izquierda) y el filtro de la Palabra (derecha).

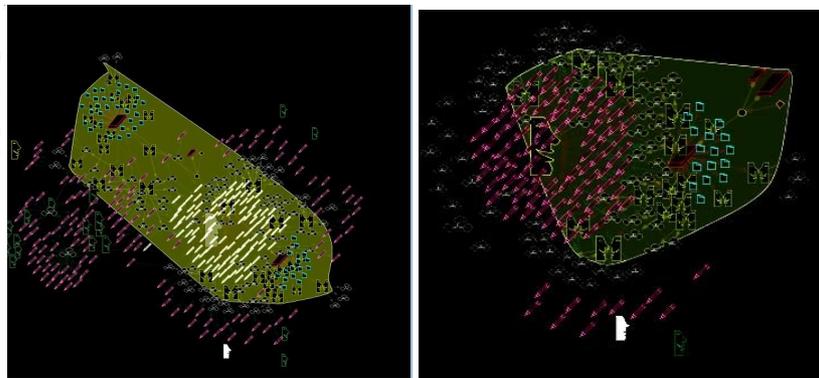


Fig. 76. Francisco (izquierda) y Patricia (derecha) redes sociales con filtro de conectividad "6".

Por tanto, la atención se centra en dos personas (Fig. 76), en este caso Francisco (que ya hemos mencionado) y Patricia (que hemos buscado por su nombre, de ahí su color blanco). Se hace uso de la personalización de la SNG para reflejar el cambio de DOS entre la persona y los demás seleccionados (en el caso actual, 6). Finalmente se comparan las dos pantallas resultantes.

En la Fig. 76 se muestra que Francisco (izquierda) es la persona con la mayor conectividad en el CMS, la actividad en los foros, y que ha utilizado la palabra "Dominio" en nuestra plataforma de aprendizaje en línea.

Se debe destacar que los múltiples puntos de vista en las versiones del prototipo del modelo VeLA están vinculados, de modo que cada vez que cualquier tipo de manipulación se lleva a cabo en uno de ellos, por ejemplo, el filtrado de datos, el resto de las vistas se actualizan dinámicamente según corresponda.

### VII.1.2 Caso de estudio 1

El análisis de este caso de estudio, utilizando la versión uno del prototipo, se dirige a un LMS basado en Moodle, los datos son los utilizados en el caso de estudio anterior, que contiene un total de veintidós (22) cursos, cuarenta y cinco (45) diferentes tipos de actividades, doscientos cuarenta y tres (243) estudiantes y cincuenta y seis mil novecientos treinta y dos (56.932) actividades en diversos cursos, que abarcan un período de dos años y medio.

Este caso de estudio es complementario al análisis que se hizo de las interacciones sociales que tienen lugar en una plataforma de *eLearning*. Como es sabido y mencionado en la sección de Retos y necesidades del proceso de enseñanza/aprendizaje y el capítulo de las Aportaciones al estado del conocimiento actual de Analítica Educativa, Analítica Académica y Analítica Visual en *eLearning*, cualquier intento de sacar conclusiones acerca de un gran número de datos educativos mediante las herramientas proporcionadas por el propio Moodle es inviable. Por lo que se desarrolló la versión uno y se siguió con la investigación.

En el primer caso de estudio el foco de interés estuvo en la visualización de las estructuras sociales que fueron construidas dentro de una plataforma Moodle, representándolas mediante un grafo dirigido por fuerzas, en el que las personas se representaban mediante nodos y las relaciones entre estos se codificaron como las aristas de la red. Asimismo, se facilitó el análisis visual en base a la frecuencia de las actividades en un gráfico circular de caracol, y, finalmente, a través de una pantalla de nubes de palabras, que tiene como objetivo visualizar el contenido semántico de la plataforma.

Al igual que en el presente prototipo, y como ya se mencionó con anterioridad, las diferentes visualizaciones desde su versión cero están enlazadas, por lo que de una manera rápida y fácil es posible obtener información nueva e indispensable para la creación de nuevo conocimiento. Sin embargo, en el caso de estudio que se describe a continuación no se parte de cero, sino que se toman las conclusiones extraídas en el primer caso de estudio como punto de partida.

Así, se parte de la enorme base de datos que se ha comentado anteriormente, pero teniendo en cuenta los nuevos conocimientos proporcionados por el trabajo previo. Este caso de estudio requirió no solo de la implementación de técnicas de visualización y representación de los datos, sino que además se necesitó una adaptación y rediseño de las mismas del prototipo anterior (véase el Caso de estudio 0, la Tabla 11 y Tabla 12, donde en materia temporal la visualización de caracol limita al analista a una cantidad pequeña de datos, teniendo como resultado una versión nueva del prototipo) y por lo que, lógicamente, comparando esta con su anterior, el análisis se mejora.

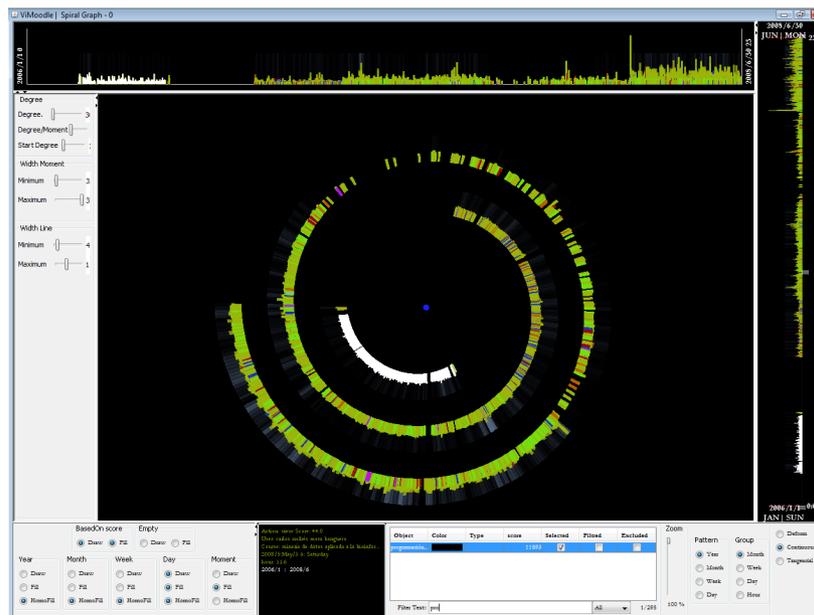


Fig. 77. Esta figura muestra el subconjunto de los datos de un curso resaltado por medio de la interacción de la caja de texto de búsqueda y de selección, localizados en el panel central inferior.

El objetivo concreto de este caso de estudio es la búsqueda de los patrones temporales en los datos y completar los conocimientos previamente adquiridos sobre el(los) curso(s) identificado(s) y la(s) persona(s), así como para descubrir nuevos conocimientos relacionados con las actividades específicas de los usuarios en la plataforma. También se trata de encontrar maneras de estudiar mejor las herramientas de educación virtual en los próximos años y descubrir los problemas potenciales que causan malos usos de los CMS.

Finalmente, también se pretende ser capaces de empezar a ver algunas soluciones, mejorando así el uso de la plataforma y la mejora de los recursos que ofrece esta tecnología aplicada a la educación y por ende a nuestra exploración. En el anterior caso de estudio se encontró que, incluso con la ayuda de herramientas de visualización de redes sociales, los datos correspondientes a un solo curso fueron suficientes para formar una red social con una conectividad alta y, por tanto, tan poblada que no se podía llegar a conclusiones interesantes. La forma más adecuada de avanzar en el análisis fue filtrando los datos por tipo de actividad. Esto condujo a la identificación de la persona(s) más conectada(s).

Como resultado de ello, el primer paso es centrarse en el curso de interés. Por tanto, se va a proceder al resaltar todas las actividades del curso deseado; esto se hace mediante la realización de una búsqueda por nombre en el panel adecuado. El resultado de esta acción se puede ver en la Fig. 77. Téngase en cuenta que una vuelta completa en la espiral corresponde a un año, con enero situado en ángulo de 90 grados y el tiempo avanzando hacia la derecha. También se puede ver que el curso (resaltado en blanco) solo ocupa un cuadrante de la espiral, lo que significa que estuvo activo desde marzo a junio de 2006 (detalles adicionales como las fechas exactas de los eventos se proporcionan como etiquetas de texto contextuales), lo que corresponde al 34% del conjunto de datos total almacenado en la plataforma.

A continuación, se filtra cualquier dato relacionado con otros períodos de tiempo. Esto se realiza arrastrando los controles deslizantes en el resumen de la línea de tiempo. Como se puede ver en la Fig. 78 (panel derecho) el contexto del período entero se mantiene en el fondo (zonas borrosas de los períodos de tiempo no seleccionados en la visión general de la derecha). La atención se centra ahora en el período seleccionado, y, como resultado tanto de la espiral y la línea de tiempo lineal (panel superior de la representación visual) solo mostrará cuatro meses de actividades.

El siguiente paso es activar la deformación de la espiral, específicamente dando *click* sobre la opción en el panel de abajo a la izquierda, panel de patrón de la espiral, que nos ayudará a visualizar mejor los patrones ajustando la representación en espiral de manera que una vuelta completa corresponde a un mes, lo que permite hacer un mejor uso del espacio disponible en la pantalla y que se desvelen patrones. El resultado de estas acciones se puede ver en la Fig. 78, que representa a un curso único, con 57 estudiantes y 11.893 actividades diferentes (codificando su tipo por colores) dentro de un período de 4 meses.

Como acaba de mencionar, un giro de la espiral representa un mes, así que el primer día de cada mes se inicia con un ángulo de 0 grados. Además, la espiral se divide en 4 cuadrantes correspondientes a las 4 semanas del mes. De este modo, en la Fig. 78 se puede observar que, para cada mes del curso, las semanas en la que menos se ha trabajado (la espiral es más estrecha en un ángulo específico), son por lo general el primero y el último. Por otro lado, durante la segunda y tercera semanas más actividades en la plataforma de aprendizaje se han llevado a cabo.

Anteriormente se mencionó la codificación de color, así es cómo funciona: cada color contenido en la espiral representa un tipo diferente de actividad (cada actividad se produce dentro de un minuto), pero también se puede observar la línea de tiempo lineal (panel superior, Fig. 78), los rectángulos (semana de este caso) son de color homogéneo gris, con la particularidad de que el brillo/claridad representa la mayor cantidad de actividades acumuladas, por tanto, la segunda y tercera semana del curso son aquellas en los que la plataforma de *eLearning* ha sido más utilizada.

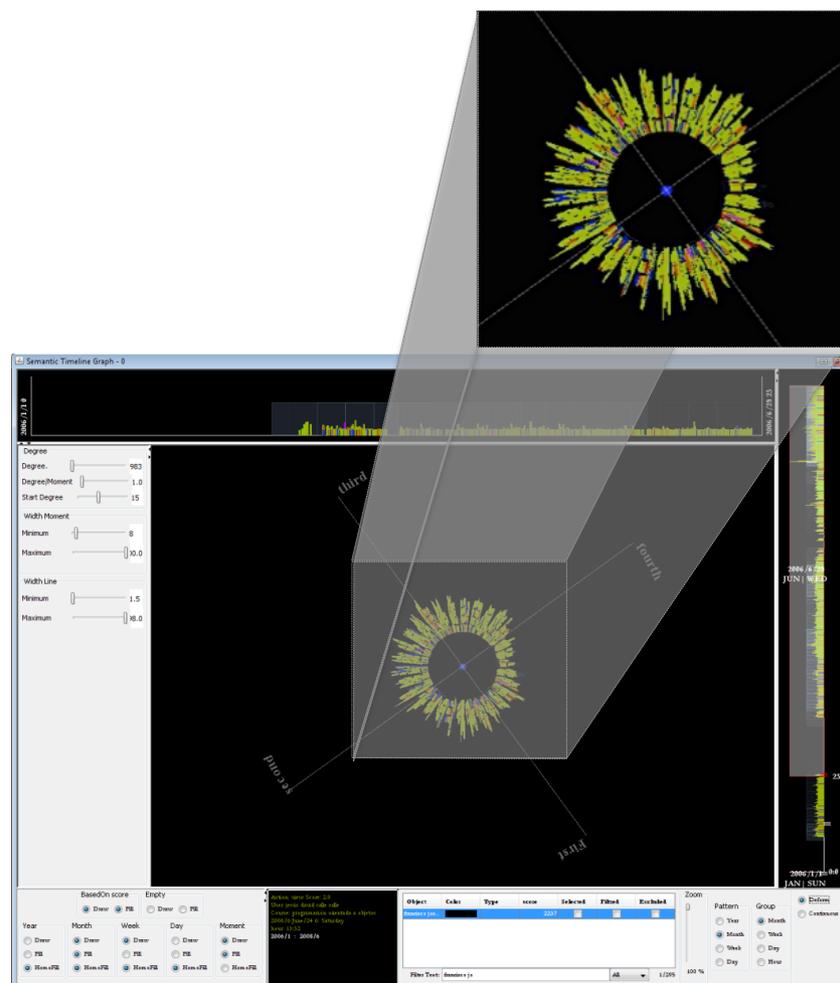


Fig. 78. Se muestra solo el periodo de tiempo seleccionado a la derecha de la visualización de SST.

Por tanto, las primeras cuatro semanas del período completo tuvieron la mayor acumulación de actividades por semana, y, dentro de esas cuatro, la segunda y tercera, como se ha indicado (ver los recuadros en tonos de gris en la parte superior de la Fig. 78); este hecho no es suficiente, sin embargo, para hacer una diferencia notoria en el período de una de ellas, en sentido global, que significa hacer que, por ejemplo, siendo la semana 3 y 4 del primer mes las más activas del curso, no significa que el período comprendido por el acumulado de cada la tercer o cuarta semana de todos los meses durante el período completo del curso logre ser la más prolífica. Dado que en este caso es la primera semana (como se puede observar en la parte central de la Fig. 78 mírese la distancia más alejada desde el centro del espiral, siendo la más alejada la del acumulado de todos los meses del periodo completo de cada primer semana de cada mes, que corresponde a la parte de debajo de la espiral).

Gracias a las ventajas de la visualización espiral distorsionada, se continua cambiando el significado de una vuelta completa ( $360^\circ$ ) a una semana (con el panel de patrón; obsérvese cómo cambia el patrón espiral de la Fig. 78 al de la Fig. 79). El resultado es una espiral diferente ahora con siete sectores que representan cada uno de los días de la semana. Este cambio pone de relieve el hecho de que tres días (jueves, viernes y sábado) contienen la mayor acumulación de actividades, siendo el jueves el día con la mayor concentración de actividad.

También es bastante obvio (ya se podía observar en la Fig. 77) que la actividad más popular es la representada en color verde ("vista", es decir, mirar las fotos de los demás participantes en la plataforma, visitar para solo leer las discusiones, los foros y los recursos). Un paso más allá en el análisis sería explorar las actividades de una persona determinada. De esta manera se puede evaluar el uso que un individuo está haciendo de la plataforma educativa.

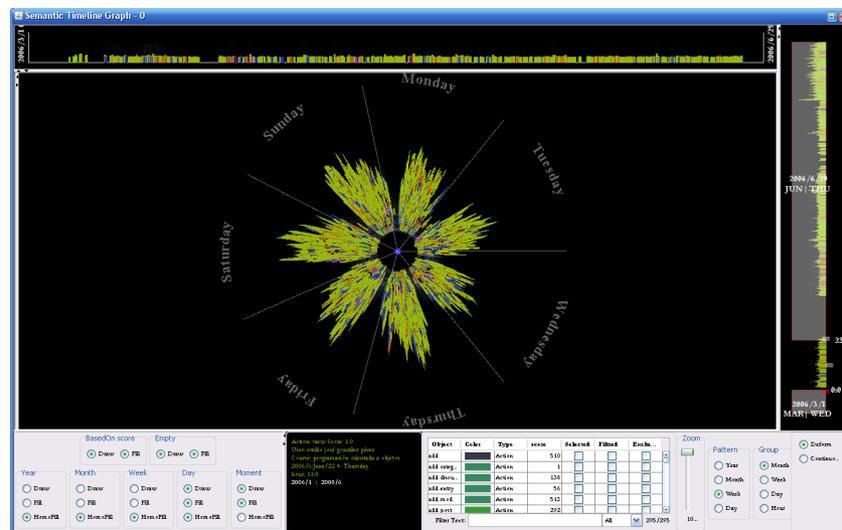


Fig. 79. Patrón temporal de las actividades semanales de un curso.

Por tanto, a continuación se filtran algunos datos, mediante la búsqueda del nombre de la persona que se quiere analizar en el panel de abajo y a la derecha de la Fig. 80 (una de las personas identificadas en el análisis de la interacción social en el Caso de estudio 0). También se puede hacer uso del *zoom* gráfico para ajustar la imagen al tamaño adecuado para el análisis.

El resultado de estas operaciones es que las actividades realizadas por el usuario seleccionado se resaltan en blanco (véase la Fig. 80). En este punto, es importante señalar que la visualización implementa la técnica de foco + contexto, difuminando el contexto en el fondo (nótese el halo gris). Esto evita perder el patrón completo y de esta manera se puede comparar con el patrón del foco elegido (datos resaltados y/o de interés en la exploración). De forma que los detalles de la persona enfocada dicen que ha hecho un total de 2.308 actividades durante los cuatro meses del curso.

Se recuerda que un giro en espiral corresponde a una semana, que se divide en siete sectores, uno para cada día de la semana. Es curioso descubrir que, a pesar del hecho de que las plataformas en línea permiten la libre programación de las actividades de aprendizaje, la persona analizada hace un uso muy metódico de la misma. Téngase en cuenta que esta persona utilizó la plataforma los siete días de la semana (incluyendo sábados y domingos) y cuando lo hizo, la mayoría de las actividades se llevaron a cabo dentro de las segundas 6 horas de cada día (es decir, en las primeras horas de la mañana).

Además, los domingos y los lunes son los días que más uso presentan de la plataforma, para luego disminuir los miércoles y los jueves y volver a incrementar la actividad de nuevo hasta el nivel de uso del domingo. Se puede añadir que esta es la persona con la mayor proporción de la actividad, ya que contribuyó hasta un 4% de las actividades totales de la plataforma.

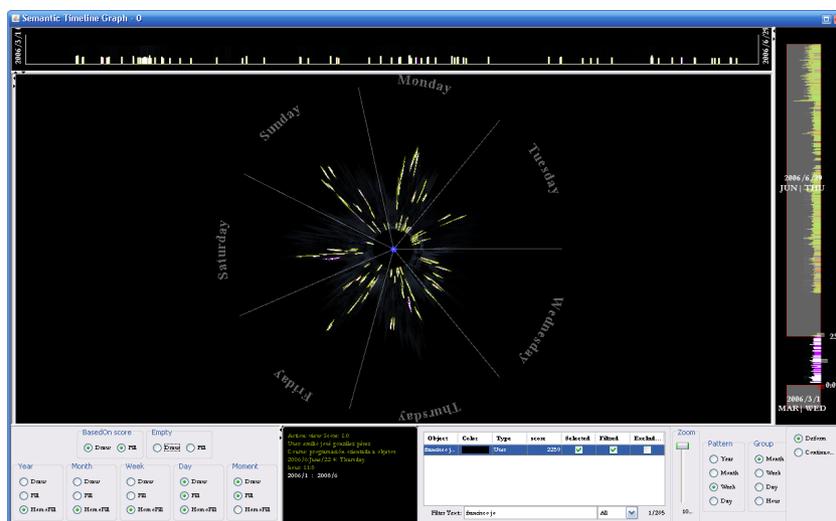


Fig. 80. La figura muestra el patrón de la actividad semanal de una persona.

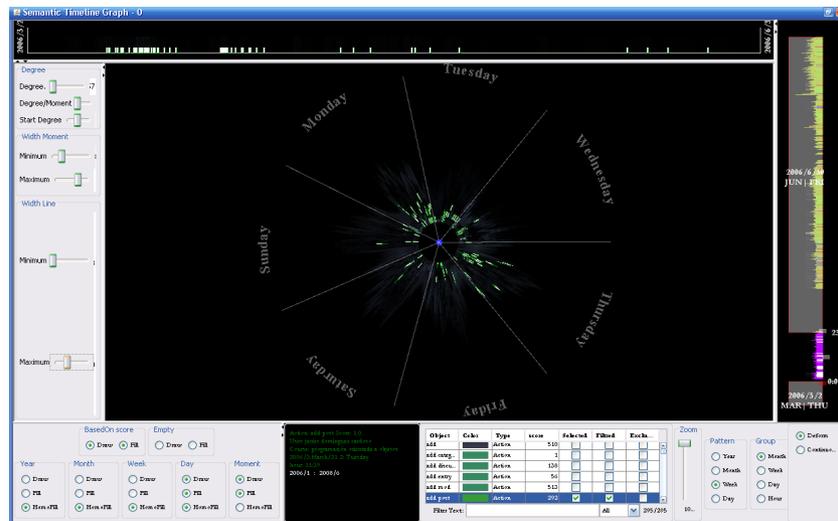


Fig. 81. Se muestra el patrón de actividad semanal de la acción "agregar post" en la plataforma.

Por último, también se tenía interés en el análisis de las actividades, por tanto y de forma consecuente, se procedió de una manera similar a la etapa anterior, buscando el nombre de la actividad que se quería explorar. Dado que las actividades que resultan de mayor interés para el análisis son las que involucran la creación de conocimiento o participación activa del mismo, se seleccionó la actividad "agregar post" y el resultado se muestra en la Fig. 81.

La visualización de la Fig. 81 representa un total de 181 actividades del tipo "agregar post", que se realizan por la herramienta, manteniendo el resto de las actividades en tonos oscuros para mantener el contexto del análisis. Esto es una cantidad muy pequeña de las actividades con respecto a las actividades totales de la curso.

Una vez más, una vuelta completa de la espiral representa una semana, por lo que el patrón indica que este tipo de actividad se ha llevado a cabo con mucha más frecuencia al principio del curso (las actividades aparecen sobre todo en los giros más internos Fig. 81). Sin embargo, los lunes y jueves en la última parte del curso se produjo un aumento de este tipo de actividad. Además, los jueves y los lunes se utilizó la plataforma para agregar mensajes, y se experimentó una notable disminución los viernes, sábados y domingos. Uno de los hechos más difíciles de entender, que se puede ver fácilmente con esta visualización en espiral, es que la actividad explorada muestra un patrón que rompe con el patrón general: téngase en cuenta que no existe una forma borrosa de los lunes, es decir, los lunes no se encuentran entre los días con mayor acumulación de actividades en general, en oposición a la acumulación "agregar mensaje" (del inglés, *add post*).

Otro hecho interesante (Fig. 82) es que la acumulación general y la acumulación de las unidades de tiempo no siempre coinciden, como en este caso ocurre con las actividades por mes. Si se tiene en cuenta la suma de todas las actividades por mes de manera individual durante todo el período, junio 2008 es el más activo (el rectángulo brillante en la línea de tiempo lineal y el pico de la forma más externa en junio en la Fig. 82).

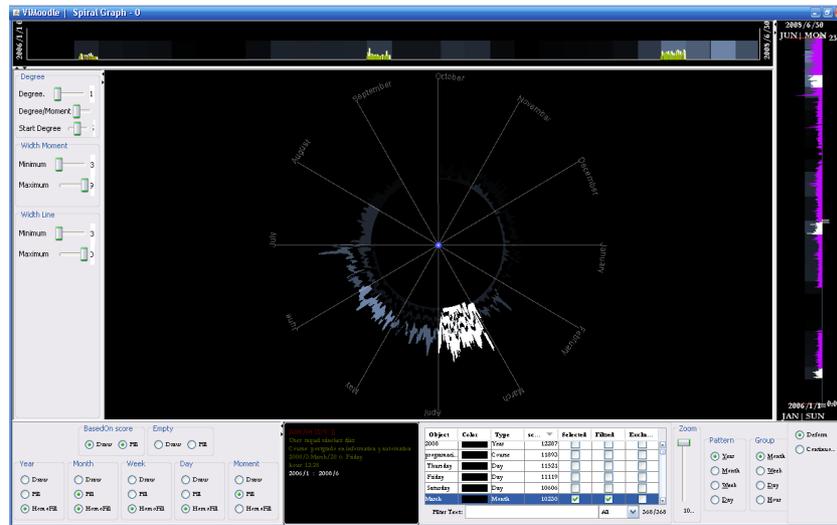


Fig. 82. Un ejemplo de la utilidad del uso de ambas representaciones de línea de tiempo, tanto la espiral como la lineal.

Por otra parte, la acumulación de actividades para todo el período muestra que marzo es el mes con mayor actividad (la forma blanca en la visualización de espiral central), es decir, que marzo ha promediado más actividades durante el total de años, que junio, que solo un año tuvo un aumento de actividad atípico.

Además, con la ayuda de esta herramienta se ha encontrado que hay una concentración de actividades que representan un uso poco activo y participativo de la plataforma de aprendizaje. Entre estos usos se pueden destacar: la visualización de los perfiles de los usuarios, la lectura de las discusiones sin aportar ningún comentario, explorar la lista de los contenidos del curso sin acceder a ellos, y así sucesivamente.

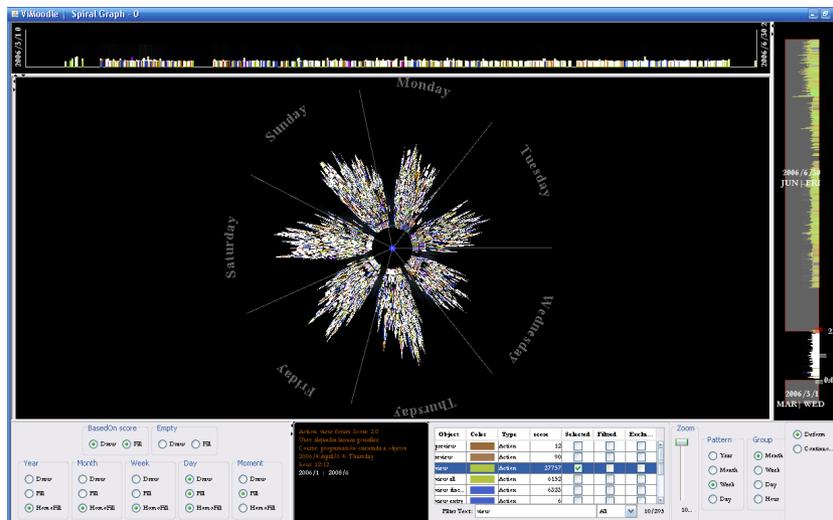


Fig. 83. La figura representa el patrón temporal semanal de la actividad “view”.

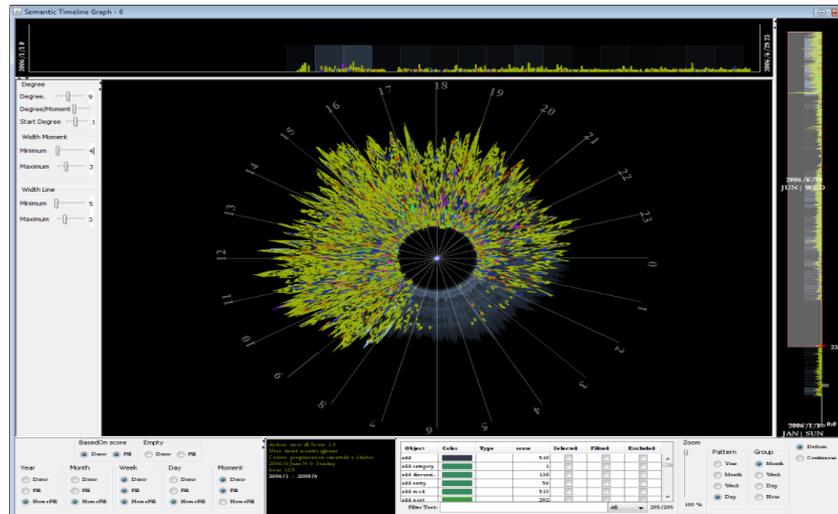


Fig. 84. Se representa el patrón temporal diario de las actividades de un curso específico.

Por tanto, resulta de gran interés analizar el uso real de la plataforma. La Fig. 83 se refiere al mismo ejemplo anterior con la actividad "add post" que se ha estado analizando, pero si se selecciona la actividad "vista", se resaltan estas actividades. Si se comprueban los números, se va a obtener como resultado que 57 estudiantes realizaron 8.266 actividades de "vista".

Es muy llamativo si se comparan la Fig. 81 (actividad con un fuerte contenido educativo activo, "agregar post") y la Fig. 83 (actividad con los contenidos educativos participativos mínimos, "vista"). Es importante mencionar que la herramienta ofrece la posibilidad de filtrar cualquier tipo de actividades, si se quiere analizar solo actividades con mayor carga activa para evitar el "ruido" de que esas actividades más pasivas y que, a priori, podrían tener un menor peso en el proceso de aprendizaje.

Teniendo en cuenta que en la representación por semanas de la espiral se muestra una ligera diferencia en el huso horario, para continuar el análisis ahora se cambia el significado de un giro de la visualización a un día, dividiendo la espiral en 24 horas. De esta manera se tiene la oportunidad de ver la distribución horaria de las actividades en la plataforma de *eLearning* (Fig. 84). Como puede verse fácilmente, se puede concluir que las horas en que hay una mayor cantidad de actividad en la plataforma son justo antes y justo después de la comida (téngase en cuenta que en España la hora del almuerzo es alrededor de las 14:00), es decir, ente las 11:00 y las 12:00 de la mañana y las 15:00 horas. También se pueden notar algunos sectores vacíos, que corresponden a las horas de la madrugada, que por supuesto, es que normalmente, cuando los usuarios de la plataforma están dormidos. Sin embargo, no es un caso excepcional que haya personas que utilizan la plataforma varias veces entre las 02:00 y las 07:00.

Para completar el análisis, se seleccionó la misma persona que en el caso de la Fig. 81, pero esta vez eligiendo una espiral distorsionada con el patrón por día; el resultado se muestra en la Fig. 85. El usuario seleccionado tiene una mayor cantidad de actividad en las segundas seis horas del día.

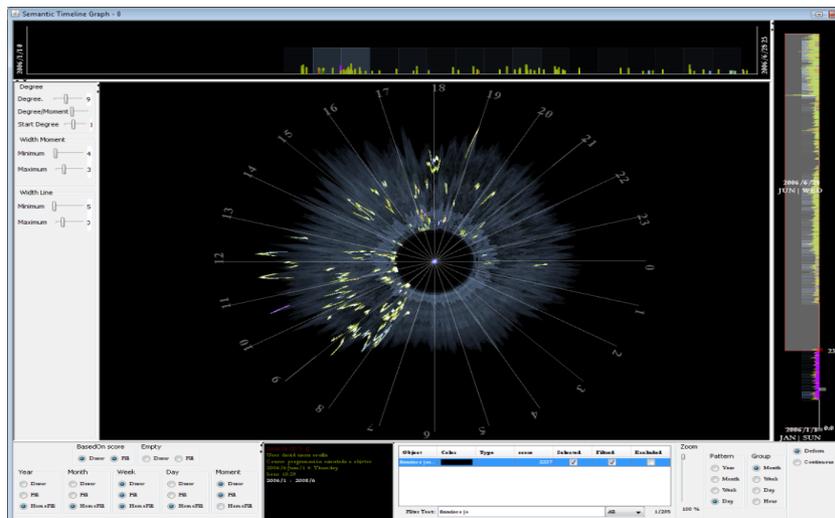


Fig. 85. Se representa el patrón temporal diario de las actividades de una persona específica.

Es curioso ver que existe una concentración de la actividad en torno a las 18:00 horas, además de una concentración de la actividad entre las 8:00 y las 9:00 horas. Finalmente, también es interesante descubrir que, en contra de lo que cabría esperar en una plataforma educativa no regular, todas estas actividades se llevaron a cabo justo antes de ir a trabajar o después del trabajo, como la jornada de trabajo más común es de 9:00 a 14:00 y de 15:00 a 18:00, se puede inferir que esta persona también está trabajando además de utilizar la plataforma educativa o participa en otras actividades y es por eso que normalmente utiliza la plataforma antes y después de un tiempo específico.

### VII.1.3 Caso de estudio 2

Previamente se ha abordado el análisis, por medio de servicios web en un entorno virtual de aprendizaje (VLE) (Gómez Aguilar, Conde González, et al., 2010b), para analizar de forma visual un curso basado en Moodle. En este caso de estudio se utilizó el sistema con el fin de analizar (utilizando la visualización de redes sociales utilizada en el Caso de estudio 0) diferentes datos con respecto a los casos de estudio anteriores. En este caso se tomó la asignatura de Ingeniería del *Software* (García Peñalvo, Bravo Martín, & Conde González, 2008) del tercer curso de Ingeniería Técnica en Informática de Sistemas de la Universidad de Salamanca durante el año académico de Septiembre de 2009- a Septiembre 2010. Se analizaron casi 46.500 eventos para un total de 115 estudiantes.

Se desarrolló una optimización e implementación de métricas encontradas a lo largo de la investigación, tomando como punto de partida el trabajo anterior (Caso de estudio 0 (Gómez-Aguilar, Therón Sánchez, & García Peñalvo, 2008)), con el fin de mejorarlo y llevar a cabo este estudio. Tomando ventaja de sus beneficios, es decir, que se puede analizar temporalmente el curso, que es independiente de la versión de la plataforma y se puede utilizar fuera de línea.

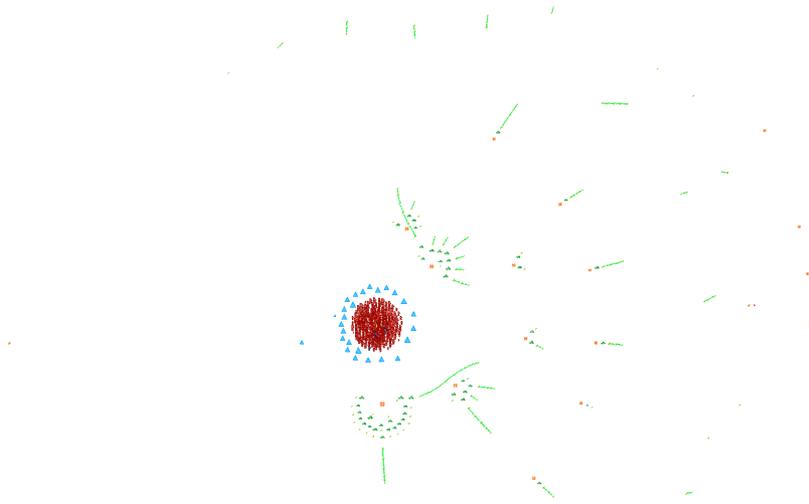


Fig. 86. Vista global de la red social. Esta visualización muestra el aspecto general de la estructura del curso. Obsérvese las líneas de conversación, formadas por un grupo de *post* en un foro, también algunas de ellas están relacionadas con sus discusiones, las cuales a su vez están relacionadas con su foro. Además, nótese el grupo de triángulos azules que son recursos de la asignatura, que están alrededor de la gente en el curso.

El desarrollo realizado amplía las posibilidades de la herramienta de analítica visual con el fin de revelar otro tipo de patrones. La visualización gráfica social de este estudio es un grafo de fuerza dirigido. Se interpreta un gráfico como un sistema físico con fuerzas entre los nodos y, a continuación este trata de minimizar la energía del sistema para obtener un equilibrio, resultando un dibujo agradable. También se pueden aplicar a los diseños en clúster (Kaufmann & Wagner, 2001), para el propósito de este caso de estudio se aplicaron resortes internos y externos (del inglés, *internal-spring* y *external-spring*) para crear los grupos.

Los datos para este caso de estudio fueron recogidos durante el curso, por ejemplo: el número de visitas, la lectura y el acceso a los mensajes de cada debate, el número de accesos de lectura a un recurso se contabilizaron para producir datos estadísticos y cuantitativos para el análisis. El análisis de la parte semántica consiste en la selección, la unificación, la normalización y el filtro de los eventos del registro de la plataforma. Este filtro se realizó para eliminar las palabras vacías del contenido y de los mensajes del foro que no son relevantes (como "sonrisas", "abrazos", "hola", etc.) (Cristóbal Romero et al., 2008).

Como contribución complementaria a la herramienta de analítica visual en la cual se pueden utilizar otras medidas existentes (por ejemplo, el grado de separación, la centralidad, etc. (Bakharia & Dawson, 2011; Gómez-Aguilar et al., 2008)); este estudio muestra la correlación existente entre dos y tres variables que se han encontrado durante la investigación, que son directamente proporcionales a la nota, como es la participación activa del estudiante. Todo esto representado en la misma visualización interactiva.

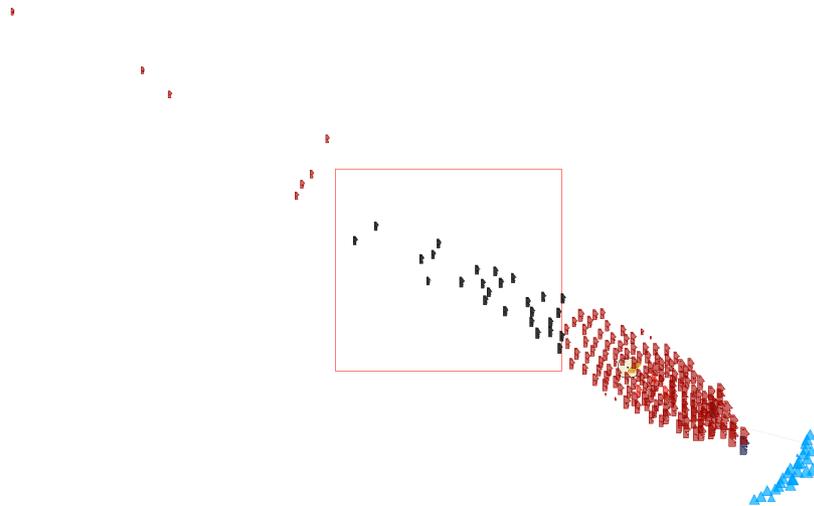


Fig. 87. Vista de la red social filtrada previamente con el fin de mostrar solo la relación entre la lectura de los recursos, el número de personas con quien se relaciona y gente que lo hizo. Es importante mencionar que el recuadro rojo es un modo de interacción para la selección de un grupo de personas a analizar.

Se ha creado un gráfico que muestra las actividades (agregar comentario, leer comentario, leer recurso, agregar recurso, ver perfil, ver discusión, ver foro, etc.) entre todos los estudiantes. Esta gráfica es bastante densa (Fig. 86), pero se ve la información general de la estructura. Las formas de color rojo y gris que forman un círculo son las personas, los triángulos azules son los recursos, los objetos verdes son discusiones, los elementos naranja son los foros y las líneas verdes y grises son conversaciones. Además, la opacidad y anchura de las aristas hace posible la ilustración de la intensidad de la relación entre los estudiantes y demás participantes con el acceso de lectura de un recurso o debate y la frecuencia de los participantes en los foros. El debilitamiento de la intensidad de la relación reflejado en la arista que une dos nodos se ve representado también en el distanciamiento secuencial de los nodos.

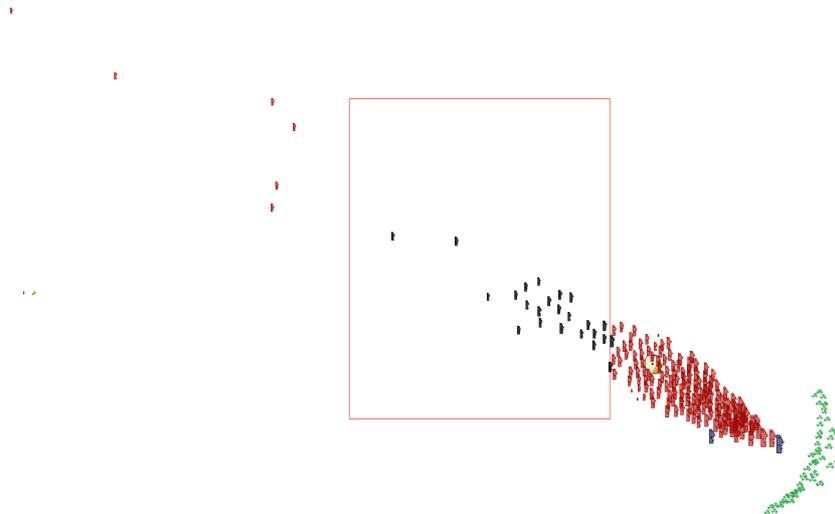


Fig. 88. Vista de la red social filtrada previamente con el fin de mostrar solo la relación entre la lectura de los debates, los mensajes del foro, el número de personas con quien se relaciona y gente que lo hizo. Obsérvese también el marco rojo, que es un modo de interacción para la selección de un grupo de personas para su posterior análisis.

Por esta razón, el comportamiento de la gráfica se representa de manera diferente dependiendo de los elementos visibles (ver Fig. 87, Fig. 88, Fig. 89). Esto significa que cuanto mayor sea la frecuencia de actividades entre los estudiantes, más cerca del grupo de debates o recursos o ambos, dependiendo del caso y, por tanto, también más cerca que se colocan entre ellos. Siguiendo la idea, cuanto más lejos están los estudiantes del conjunto de las discusiones o de recursos, según el caso de grupos, la frecuencia de la actividad es menor. Además, cuanto menor es la diferencia entre sus frecuencia de actividad entre los estudiantes más cercanos se ubicarán entre ellos. Esto es debido a que el tamaño de las líneas de las aristas se calcula de modo que represente una dispersión de los estudiantes en relación a la correlación entre las variables cuantitativas visualizadas. Téngase en cuenta que este tipo de cálculo a través de las aristas en el grafo de fuerzas dirigido no se ha realizado con anterioridad, por lo que es uno de los aportes de esta iteración de la investigación. En este caso de estudio se pretende responder si es posible encontrar una relación entre el rendimiento de los estudiantes y la participación, por medio de cambios visibles en nuestra herramienta de análisis visual (Gómez-Aguilar et al., 2008). Para confirmar esta hipótesis, se analizaron tres casos:

**R:** La relación entre la frecuencia de acceso a la lectura de los recursos y el número de personas con las que se relaciona cada estudiante y su rendimiento escolar, es decir, la nota del estudiante Fig. 87.

**D:** La relación entre la frecuencia de acceso de escritura a los foros, la frecuencia de acceso de lectura a las discusiones, el número de personas con quien se relaciona el estudiante y la calificación del estudiante Fig. 88.

**DR:** La relación entre la frecuencia de acceso de escritura a los foros, la frecuencia de acceso de lectura a los recursos, la frecuencia de acceso de lectura a los debates, el número de personas con quien se relaciona el estudiante y las calificaciones del estudiante Fig. 89.

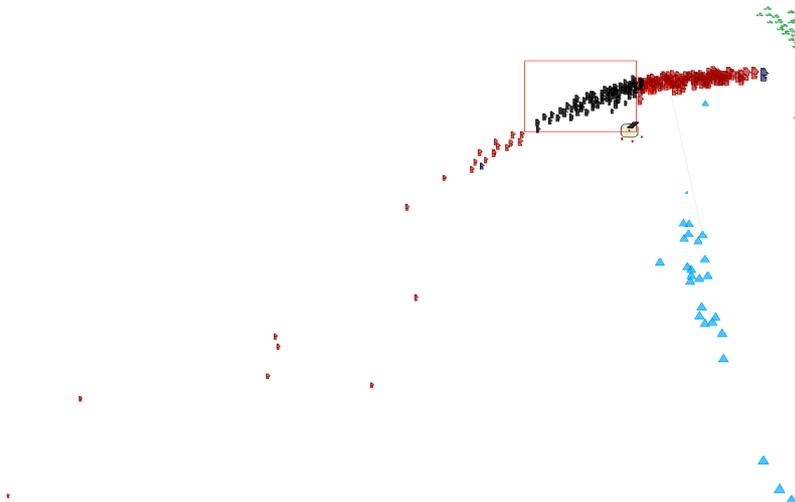


Fig. 89. Vista de la red social que se filtró previamente con el fin de mostrar la relación entre la lectura de los recursos, la lectura de las discusiones, los mensajes en los foros, el número de personas con las que se relaciona cada participante y las personas que lo hicieron. Obsérvese también que existe un cuadro rojo, este es un modo de interacción para la selección de un grupo de personas para analizar posteriormente.

Por tanto, se ocultaron los elementos de la visualización que no se quería que se correlacionaran (en este caso, los mensajes, los invitados y los foros) y se dejaron visibles los iconos de las discusiones (los verdes sobre la Fig. 86, Fig. 88 y Fig. 89) o de los recursos (los azules en las Fig. 86, Fig. 87 y Fig. 89) o ambos, dependiendo del caso.

Además, para cumplir con el propósito de este caso de estudio se dividió el grupo de personas en tres partes para cada caso anterior. Se seleccionaron 3 grupos a través de la visualización, como de costumbre, utilizando el evento interacción con el ratón (manteniendo pulsado el botón izquierdo del ratón y arrastrando).

Cuando se selecciona un grupo de personas a través de la interacción de visualización, se crea un cuadro rojo (como se puede ver en la Fig. 87, Fig. 88 y Fig. 89). Por tanto, al seleccionar tres veces, se tienen tres grupos de personas con una relación específica analítica con diferentes variables correlacionadas como métrica: asignando una **D**, una **R** o **DR**, dependiendo del caso analizar. Finalmente se asignaron las calificaciones de los estudiantes para cada persona en los grupos (ver Fig. 90) y la media.

<b>D: 1</b>	→ 5.7	<b>D: 2</b>	→ 5.8	<b>D: 3</b>	→ 5.0
<b>R: 1</b>	→ 5.7	<b>R: 2</b>	→ 6.0	<b>R: 3</b>	→ 4.5
<b>DR: 1</b>	→ 6.0	<b>DR: 2</b>	→ 5.6	<b>DR: 3</b>	→ 5.4

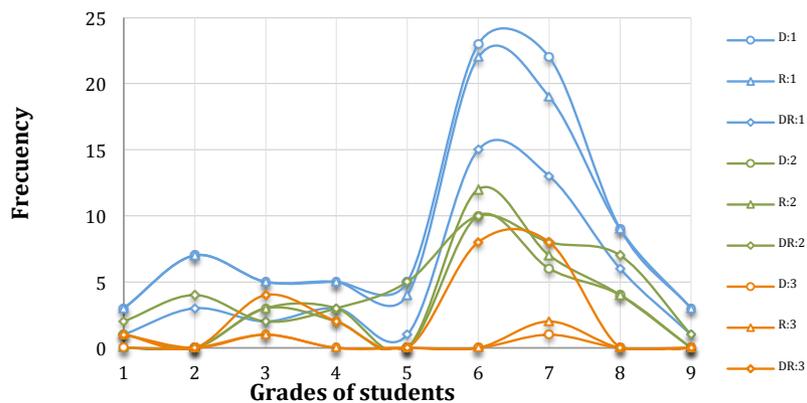


Fig. 90. Distribución de las agrupaciones D:1, D:2, D:3; R:1, R:2, R:3 y DR:1, DR:2, DR:3, correspondientes a los tres grupos de personas seleccionados de los tres casos a analizar con la finalidad de comprobar la hipótesis.

En la Fig. 90 las líneas representan el grupo formado a través de la selección visual de las personas (**1**, **2**, **3**) en la visualización del gráfico social de cada una relaciones a analizar en este caso de estudio (**D**, **R**, **DR**). El eje X representa las calificaciones del estudiante y el eje Y el número de personas. Por tanto, **1**, **2**, **3** de cada relación **D**, **R** y **DR** son los grupos de personas de una relación específica de análisis visuales de variables.

El **1** que representa el grupo más participativo (línea azul), el **2** que es el del medio (línea verde) y el **3** es el grupo participativo más bajo (línea naranja). Las letras **D** y **R** significan la relación de cada persona con el número de acceso de lectura a una discusión (**D**) o a un recurso (**R**). **DR** es la combinación de ambos.

El grupo de personas cercanas al conglomerado de las discusiones o los recursos, según el caso, son las personas más activas del curso (por ejemplo, **D:1** es el grupo de personas que accedieron a la plataforma y leyeron discusiones de forma más frecuente e hicieron más comentarios en los foros); las personas más distantes en el grupo de discusiones o de los recursos son las personas de actividad más baja (grupo **3**, en el caso del mismo ejemplo mencionado anteriormente **D:3**, línea naranja en Fig. 90).

En general, y teniendo en cuenta los datos de la Fig. 90, se puede ver la mejora en las calificaciones de los estudiantes en los tres casos (**1**, **2**, **3** independientemente de la frecuencia de lectura o escritura en foros o recursos, es decir, independientemente del grupo **D**, **R** o **DR**). Esto es, las líneas de color naranja son los estudiantes con las notas más bajas y las líneas azules tienen las más altas.

Sin embargo, si se centra la atención en el promedio de los grupos de las **D**'s o grupos de las **R**'s, el patrón anterior parece permanecer oculto. Esto es, para el caso de los grupos de **R** (Fig. 87, el caso **R** recuerde que es la correlación entre la frecuencia de acceso de lectura a los recursos, el número de personas con quien se relaciona y la calificación del este), donde el grupo **1** (en teoría los estudiantes con el más alto resultado) debería ser el mejor promedio, pero es el grupo **R:2**, que corresponde a los estudiantes con un rendimiento medio. Los resultados de los grupos de **D**'s son similares.

Esto probablemente parece una mala noticia, pero nótese que la diferencia entre la **D:1** (caso que deben ser los máximos) y el grupo **D:2** (tienen 0,1 de diferencia) es más baja que la misma relación en los grupos de **R** (0,3); ahora se observa la misma diferencia, pero esta vez entre los espacios entre los estudiantes (Fig. 87, Fig. 88). Obsérvese que los estudiantes están más cerca entre ellos en la Fig. 88 en comparación con la Fig. 87. Como conclusión, esta visualización ayuda a revelar una tendencia con facilidad y de una manera interactiva, así, rápidamente, proporcionando información oculta.

### VII.1.4 Caso de estudio 3

Esta tercera versión del sistema de analítica visual para el *eLearning* aborda, como anteriormente (Gómez Aguilar, Conde González, et al., 2010b; Gómez Aguilar et al., 2009; Gómez Aguilar, Therón Sánchez, & García Peñalvo, 2013), el análisis de los datos por medio del uso de servicios web en un entorno virtual de aprendizaje (del inglés *Virtual Learning Environment*, VLE). Por este motivo, Los datos a los que el usuario puede acceder dependen de los permisos que tenga en la plataforma, estos pueden variar desde un curso hasta un campus universitario completo; teniendo la capacidad de exploración y comparación de estos.

Los datos utilizados se obtuvieron de una plataforma Moodle con el propósito de analizar un curso. La duración de este mes de Noviembre 2011 a Mayo 2012, casi el ciclo escolar, con 13000 registros de actividades que fueron realizadas por un total de 30 estudiantes.

El objetivo de este caso de estudio es responder si es posible encontrar una relación entre el rendimiento de los estudiantes y la participación a través de los cambios visibles en una herramienta de análisis visual. Para confirmar esta hipótesis se analizaron los siguientes casos:

- a) La relación entre la frecuencia de acceso a las publicación de los foros y la frecuencia de acceso a la lectura de las discusiones con la nota del estudiante Fig. 92.
- b) La relación entre la frecuencia de acceso a recursos de lectura y las calificaciones del estudiante Fig. 96, Fig. 97 y Fig. 98.

En este caso de estudio que se presenta se ha utilizado el sistema de análisis visual para el *eLearning* y para analizar los casos mencionados se ha implementado una visualización interactiva multidimensional usando la técnica de las coordenadas paralelas.

El uso de categorías tiene importancia para este estudio dado que se requirió implementar la manera de categorizar dinámicamente las actividades de los estudios. Además de permitir el uso de categorías dentro de los ejes de las coordenadas paralelas y la coloración de las líneas en base a estas, también permite la creación de agrupaciones personalizadas y generar un archivo con estas para su importación y exportación.

Finalmente, los métodos de interacción como: el filtro de búsqueda a través de la caja de texto abajo a la izquierda de la visualización, el ocultamiento y reordenación de columnas también son de vital importancia para un análisis visual.

Desde la introducción temprana de Inselberg (Inselberg & Dimsdale, 1990), las coordenadas paralelas han sido ampliamente utilizadas para visualizar datos multivariados (Tory et al., 2005). Las coordenadas paralelas son una de las más famosas técnicas de visualización, y entre los temas más comunes de trabajos académicos en la visualización. Aunque en un principio confuso, es una herramienta muy poderosa para la comprensión de conjuntos de datos numéricos multidimensionales.

Las coordenadas paralelas (Inselberg & Dimsdale, 1990) han sido ampliamente utilizadas para el análisis de conjuntos dibujando dimensiones como ejes paralelos y datos del hiperespacio con poli-líneas que conectan sus valores escalares en los ejes, las coordenadas paralelas pueden representar los datos de N-dimensiones en un espacio de 2 dimensiones. Sin embargo, cuando los tamaños de los conjuntos de datos se vuelven muy grandes, millones de líneas pueden llenar la pantalla, pudiendo llegar a ser demasiado densas para ser interpretadas.

Basándose en los anteriores casos de estudio, este desarrollo amplía las posibilidades de la herramienta de analítica visual con el fin de revelar otro tipo de patrones y complementar el conjunto de visualizaciones, mejorando así el proceso de análisis. Este sistema está dirigido tanto a estudiantes, como profesores y gestores de una institución educativa.

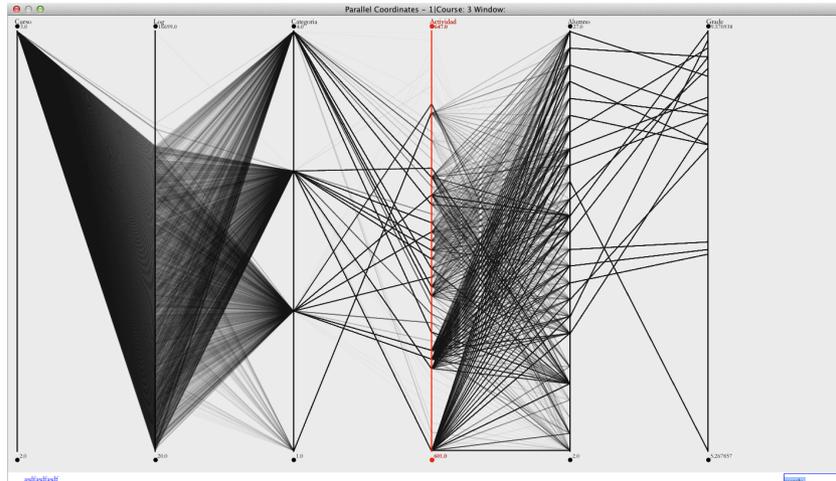


Fig. 91. Vista global de las coordenadas paralelas, donde se percibe por medio de las diferentes oscuridades de las figuras formadas por las líneas que se entrecruzan y correlacionan de eje en eje, información a primera vista. Por ejemplo se puede observar a la derecha del eje en rojo en la parte central de la imagen, cuatro bifurcaciones más oscuras que el resto, estas se corresponden a las actividades más realizadas ("view resource", "view assignment", "view quiz", "view course").

Los datos para este estudio se recogieron durante todo el curso. En primer lugar, el número de accesos a la lectura y mensajes para cada discusión, se contó el número de accesos de lectura a un recurso para producir datos cuantitativos para el análisis, se obtuvo el módulo de Moodle utilizado y el contexto con el fin de realizar la categorización adecuadamente.

Los datos consisten en las filas: "Log", "Curso", "Estudiante", "Rol", "Actividad", "Módulo", "Categoría", "Grade" como puede verse en la Fig. 91. No obstante es posible cambiar el orden y ocultar o mostrar cualquiera de estas. Para este estudio se ha ocultado las columnas "rol" y "Módulo", además se ordenaron los ejes de la siguiente manera: "Curso", "Log", "Categoría", "Actividad", "Estudiante" y "Grade" (para mejor claridad véase la Fig. 91, la cual no tiene seleccionado ningún subconjunto de los datos). A primera vista en las imágenes se puede reconocer fácilmente unos triángulos formados por las líneas entre un eje vertical y otro, que también pueden apreciarse como cilindros en la Fig. 91. El triángulo más grande es el que se reconoce entre el primer y segundo eje (eje de cursos y el eje de *log*) esto es dado a que solo se presenta un curso (primer eje).

Entre el eje del *log*, que representa todas las actividades realizadas en orden cronológico de abajo hacia arriba, y el tercer eje, el número de figuras triangulares formadas solo puede ser un número entero, ya que es el de las categorías y no hay más de cuatro, por lo que todas las líneas tienen que pasar por un pequeño número de puntos.

Además, en el espacio entre *Log* y Categoría, se puede decir que la categoría que tiene menos actividades realizadas es la 1, la categoría 3 es la más recurrida seguida por la 2 y después la 4. Esto se percibe como resultado de observar cuál es la más oscura de las figuras formadas entre los ejes, y resultado de usar la transparencia de las líneas, siendo la figura formada más oscura la de mayor frecuencia. Para algunos casos solo hay que seguir las líneas y ver cómo se cruzan: un conjunto grande de líneas que se cruzan son un indicio de una relación inversa.

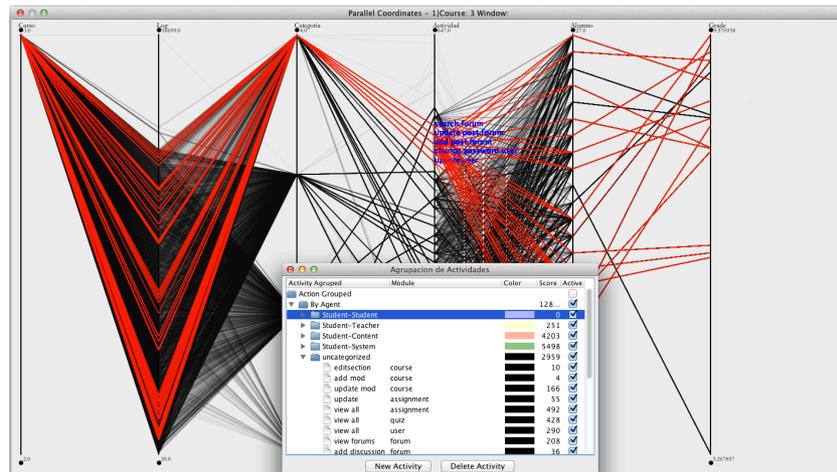


Fig. 92. Vista del subconjunto de las actividades comunes con la interacción en foros como: *search forum*, *update post forum*, *add post forum*, etc.

La correlación es mucho más directa en algunos conjuntos de datos, pero para este caso de estudio y con el objetivo de corroborar las correlaciones de los estudios antes mencionados, como por ejemplo:

- Cuanta mayor interacción mejor será nota.
- Cuanto mayor sea el número de entradas de lectura a los foros y publicación en estos mejor será la nota.
- Cuanto mayor es la frecuencia de lectura de recursos y la participación en los foros mayor será la nota.

Para comprobar estas hipótesis se exploraron los datos seleccionando diferentes subconjunto de datos, y se han elegido las imágenes más representativas. Si se observan las figuras Fig. 93, Fig. 94 y Fig. 95, en ellas se seleccionó un estudiante con nota baja, media y alta respectivamente con el objetivo de comprobar la hipótesis a).

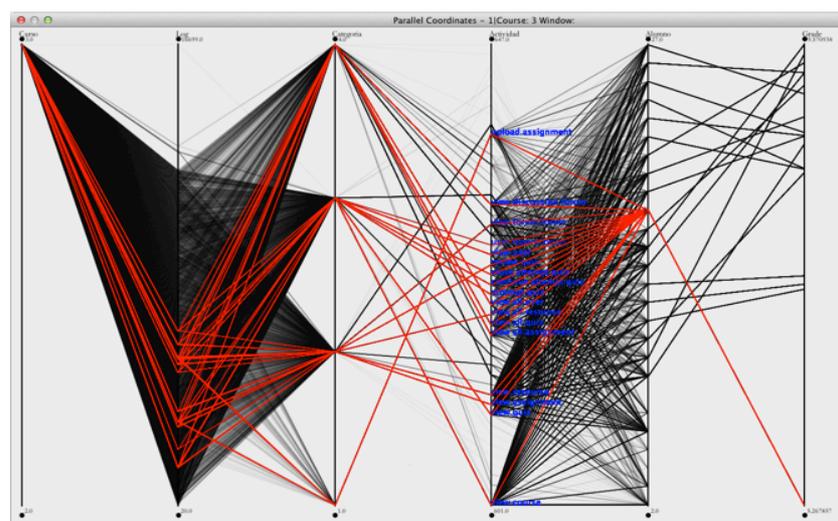


Fig. 93. Vista del subconjunto de las actividades realizadas durante todo el curso de un estudiante con una nota de 5.2.

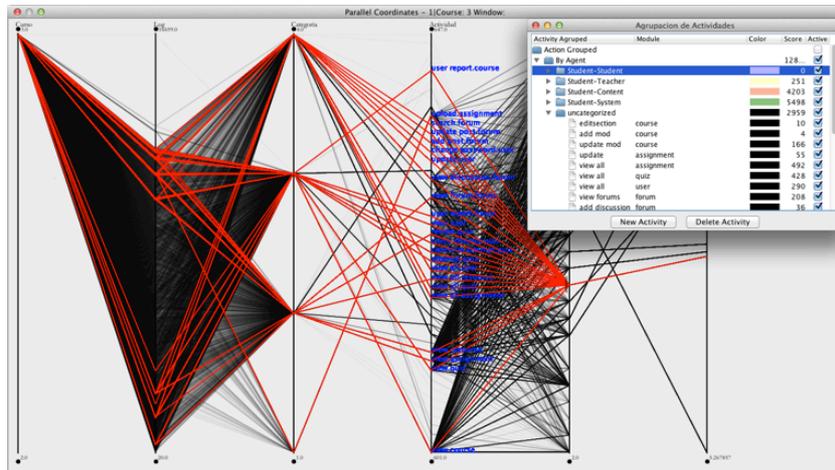


Fig. 94. Vista del subconjunto de las actividades realizadas durante todo el curso de un estudiante con una nota de 6.5.

Se puede observar, a primera vista, que a mayor frecuencia de actividades (nótese la cantidad de líneas en rojo que cruzan el segundo eje de izquierda a derecha, eje de *Log*) es más alta la nota. Sin embargo, entre el estudiante de nota media (Fig. 94) y nota baja (Fig. 93), la diferencia es muy pequeña lo que hace resaltar el patrón de periodicidad. Es decir, que no solo la cantidad de actividades que realice el estudiante es representativa para deducir que obtendrá buena nota, sino también la periodicidad con la que se realizan estas.

Para el caso del estudiante de baja nota (Fig. 93), solo realizó actividades al inicio del curso (mírense las líneas rojas que cruzan el eje de *Log* que solo se colocan en la parte inferior); en comparación con el estudiante de nota media que en la Fig. 94 sus actividades fueron realizadas tanto al inicio del curso como al final.

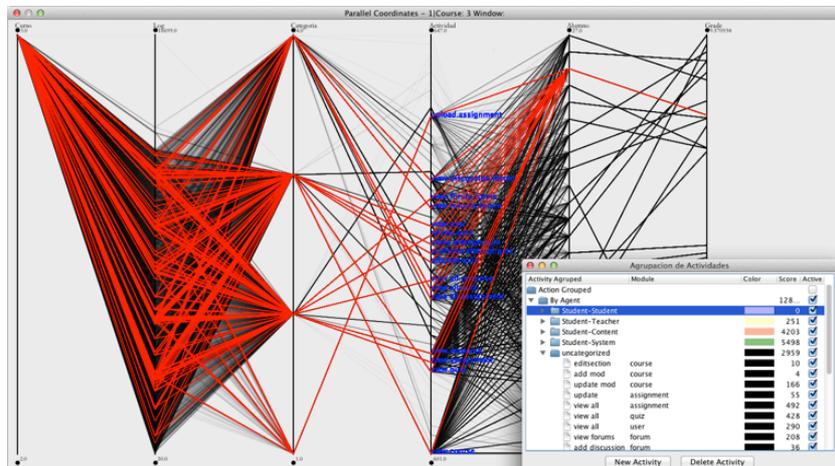


Fig. 95. Vista del subconjunto de las actividades realizadas durante todo el curso de un estudiante con una nota de más de 8.

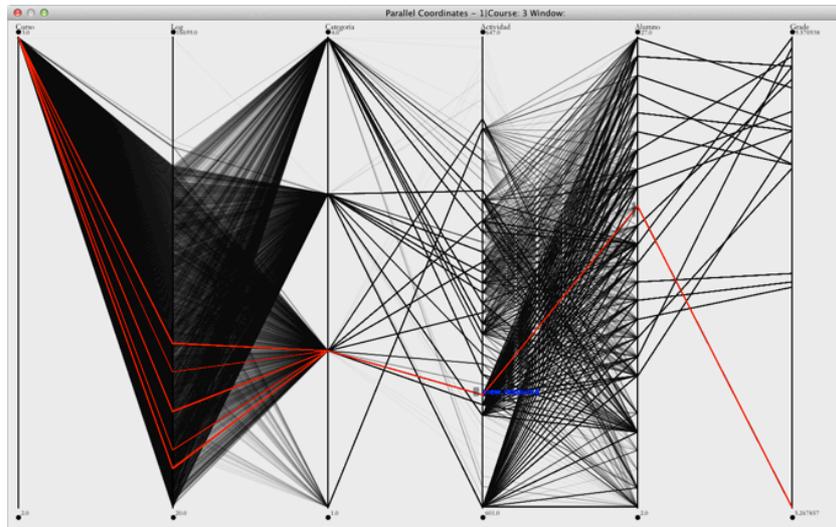


Fig. 96. Vista del subconjunto de las actividades de "view resource" realizadas durante todo el curso de un estudiante con una nota de 5.2.

Además, se puede observar que no son las mismas actividades las que se realizan, razón por la cual se ha llevado a cabo el siguiente experimento: se seleccionó el conjunto de categorías de actividades y los subconjuntos de las actividades de estas que fueron estudiadas anteriormente en el trabajo mencionado (Agudo-Peregrina et al., 2012) con el fin de comprobar esa relación en este conjunto de datos y desvelar nuevos patrones encontrando los siguientes resultados.

En la categoría 4 mencionada en el estudio como representativa se denotó (ver Fig. 92) que las actividades "search forum, update post forum, add post forum" fueron realizadas únicamente por los estudiantes con buena nota y alta. Con esto se puede comprobar la relación de que cuanto más sea la frecuencia de las actividades de esta categoría mayor será la nota y en específico de estas actividades.

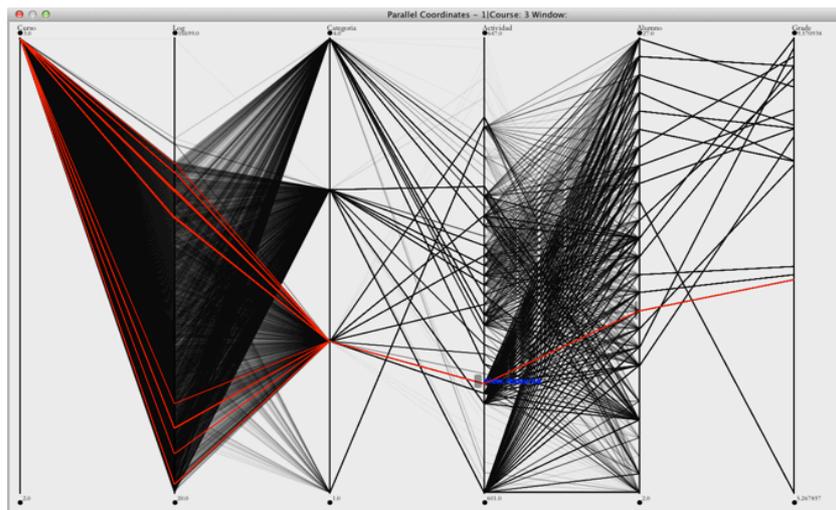


Fig. 97. Vista del subconjunto de las actividades de "view resource" realizadas durante todo el curso de un estudiante con una nota de 6.5.

Posteriormente, con el objetivo de comprobar la hipótesis b) en este conjunto de datos y observar cómo se manifiesta la periodicidad, las Fig. 96, Fig. 97 y Fig. 98 muestran la selección del subconjunto de la actividad "view resource" de un estudiante con nota baja, media y alta respectivamente. En estas imágenes se observa claramente que se repite el patrón, esto es, que cuanto más frecuencia de actividad mejor será la nota.

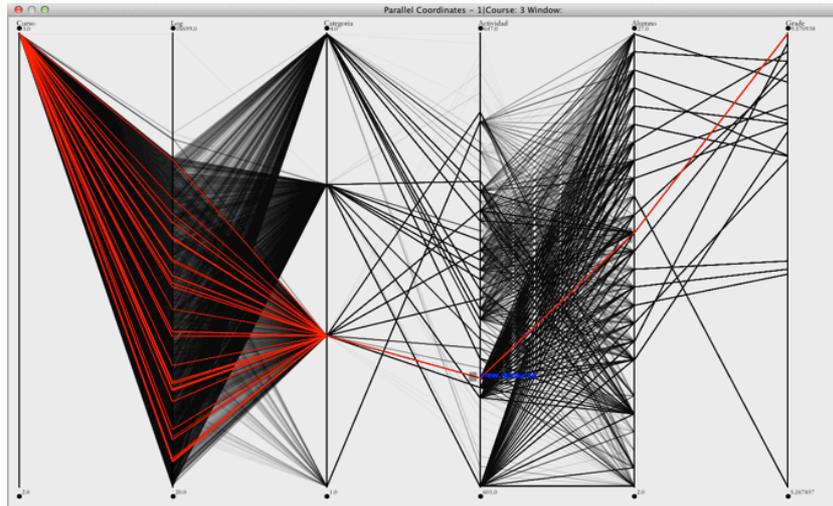


Fig. 98. Vista del subconjunto de las actividades de "view resource" realizadas durante todo el curso de un estudiante con una nota de 9.6.

Sin embargo, no siempre significa mejor nota cuanto mayor frecuencia de actividad, pero la tendencia general es claramente esa. Cabe mencionar que también el patrón de la periodicidad se ve en este ejemplo, ya que en las imágenes Fig. 96 y Fig. 97 se puede ver que la diferencia de frecuencias no es muy grande pero sin embargo la periodicidad de la Fig. 97 está mejor distribuida (entre el inicio y el final del curso) en comparación con la Fig. 98, teniendo un efecto positivo en la nota, que varía de 5.2 a 8.

#### VII.1.5 Caso de estudio 4

Para este estudio, se han utilizado tres grupos de datos de distinta naturaleza (curso 1, curso 2, y del curso 3). Cabe señalar que en el caso de estos datos, los cursos son presenciales y a pesar de la documentación impresa se distribuye a los estudiantes, también se encuentra disponible en la plataforma de aprendizaje configurada para este efecto. Por otra parte, hay tareas y pruebas en los tres cursos que se contestaron a través de la plataforma *eLearning*.

La plataforma de elección para la realización de los cursos fue Moodle. A todos los estudiantes y profesores se les dio un manual de usuario antes de comenzar el curso, y se han establecido cauces para la resolución de las cuestiones técnicas a través de correo electrónico o foros de la plataforma y los mecanismos de recogida de estadísticas de la interacción proporcionada por la plataforma.

El objetivo de este caso de estudio fue responder a la siguiente pregunta:

*¿Es posible encontrar una relación entre el rendimiento y la participación de los estudiantes mediante el sistema de analítica visual y la caracterización de las actividades?*

Para confirmar esta hipótesis se analizaron los siguientes casos:

- Relación temporal entre las categorizaciones de actividades personalizadas y el rendimiento de los estudiantes Fig. 99.
- Búsqueda de patrones en las frecuencias de las categorizaciones de las actividades y/o el rendimiento de los estudiantes Fig. 100.

En este caso se ha utilizado el sistema de analítica visual para analizar un total de 91 estudiantes y casi 55.000 eventos. Los datos para este estudio fueron recogidos durante diferentes períodos: en el caso del Curso 1, mayo 2011 a agosto 2011, y para los casos del Curso 2 y Curso 3 desde noviembre 2011 hasta marzo 2012. El análisis automático, como en los anteriores casos de estudio, consiste en la selección, la unificación, normalización y filtro del registro de eventos de la plataforma.

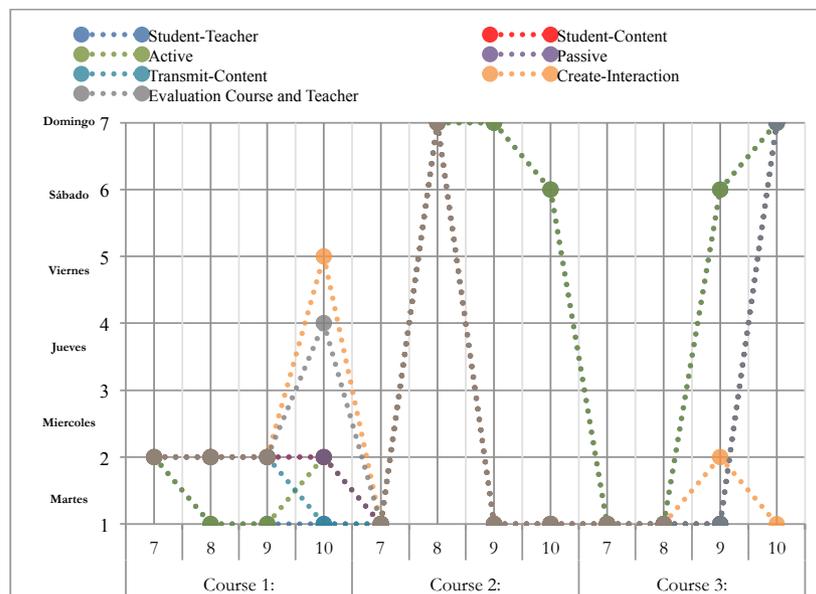


Fig. 99. Se representan los días de la semana codificados por número (1 es el lunes 2 es martes, y así sucesivamente) en el eje vertical, y en el eje horizontal, son los diferentes grados de desempeño de los estudiantes en los cursos.

En este estudio se utilizó para analizar las actividades de los estudiantes a través de tres agrupaciones diferentes de actividades que se ha mencionado anteriormente y se puede ver en la tabla de la Fig. 100.

- 1) La agrupación de actividades por agente:
  - a) Estudiante-Profesor.
  - b) Estudiante-Contenido.
- 2) La agrupación de actividades por finalidad:
  - a) Activo.
  - b) Pasivo.

- 3) La agrupación de actividades por la interacción:
  - a) Transmisión de contenido.
  - b) Creación de interacción.
  - c) Evaluación del curso y del profesor.

Para responder a la primera pregunta de la planteada en este estudio, el análisis consistió en seleccionar para cada curso solo aquellas actividades que fueron realizadas por los estudiantes, que luego se filtraron por cada grado de desempeño de los estudiantes y, por último, por cada una de las agrupaciones de actividades.

Una vez que se obtuvo el conjunto de datos, se exploraron las diferentes representaciones del patrón temporal de la herramienta (año, semana, día), con el fin de identificar algunos de los patrones de comportamiento existentes. Como resultado, no se ha encontrado un patrón diario, sin embargo, se encontraron algunos detalles curiosos en los patrones semanales. Dado que el análisis es exploratorio y el descubrimiento de algún patrón se crea en el usuario por medio de la observación de diferentes representaciones del sistema, por consiguiente este comportamiento encontrado se obtiene a partir de un gran número de imágenes.

Para expresar los resultados de forma concisa, se ha incluido un resumen de los resultados en la Fig. 99. Aquí lo interesante no es el día de mayor actividad específica de los datos seleccionados, es decir, el caso del grupo de *Create-Interaction* de las actividades (como se indica por la línea de color naranja en Fig. 99), los días con mayor actividad son lunes, martes y en un segundo nivel los viernes y sábados. La relación más importante encontrada es que cada uno de los cursos cambia el mayor día de actividades en el mismo momento que cambia la calificación del desempeño del estudiante. Es decir, independientemente de cada curso y de cada agrupación de actividades, casi en todos los casos cambian su principal día las actividades y el grado de rendimiento también, esto es que un grupo de estudiantes que tienen determinadas calificaciones en un curso, no comparten un día de mayor actividad con el otro grupo de estudiantes. Y más concretamente, el grupo de estudiantes con las mejores calificaciones no compartía sus días de mayor actividad con los otros grupos de estudiantes. Esto puede ser una consecuencia de:

1. Factores sociológicos que conducen a los mejores estudiantes a interrelacionar se y trabajar juntos.
2. Factores aleatorios que llevan a los mejores estudiantes a trabajar en el mismo día.
3. Factores relacionados con la organización del curso que obliga a los grupos de estudiantes a trabajar juntos.

Otra cosa a destacar en la Fig. 99 es que cuanto mayor es la nota de los estudiantes, también el número de días de mayor actividad, es decir, si se toma como ejemplo el curso 1, los estudiantes con calificaciones de 7 a 9 estuvieron realizando con mucha más frecuencia actividad únicamente los lunes y martes, con independencia del tipo de actividad, pero estudiantes los con la nota 10 de ellos tuvieron lunes, martes, jueves y viernes.

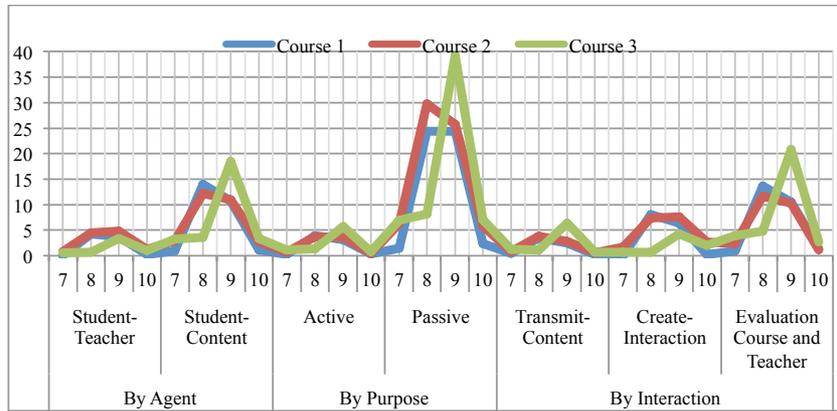


Fig. 100. Vista del conjunto de 3 cursos, que se filtró previamente con el fin de mostrar la relación entre la agrupación de las actividades y el desempeño de los estudiantes de los datos. El eje vertical es el porcentaje de actividades que fue hecho en total y en el horizontal es la agrupación de las actividades para cada grado de desempeño de los estudiantes para cada curso.

Para una mejor comprensión de cómo la Fig. 99 fue diseñada, es necesario hacer referencia a la Fig. 101, en esta última, previamente se han filtrado solo las actividades de los estudiantes que se ven, que corresponden al curso 1 y los estudiantes de nota 9, y, donde se puede ver que el martes es el día de mayor actividad.

Por tanto, tomando en cuenta el día de mayor actividad del ejemplo anterior y para cada conjunto de datos, en la Fig. 99, los días se codifican en forma consecutiva, el lunes será 1, el martes será el número 2 y así sucesivamente. Al hacer esto con cada uno de las agrupaciones de actividades y cada uno de las notas del curso, para cada curso, la Fig. 99 y la Fig. 100 fueron construidas.

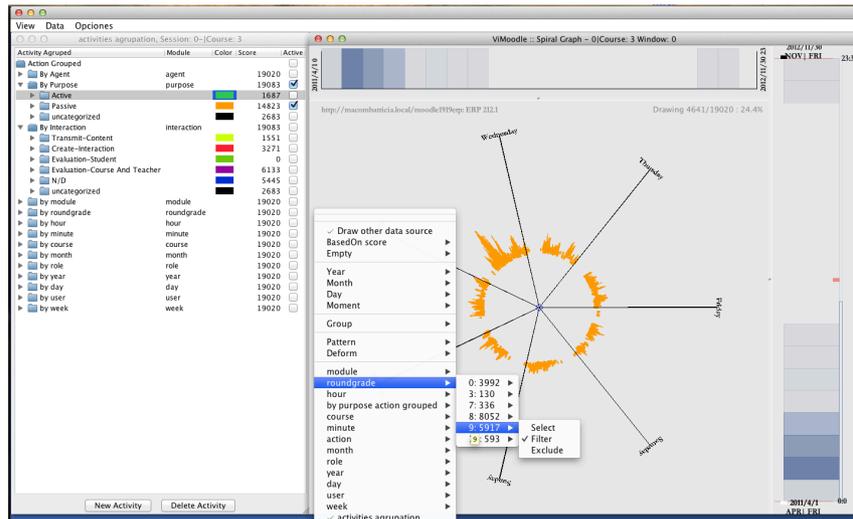


Fig. 101. Vista de la SST que se filtró previamente con el fin de mostrar solo las actividades de los estudiantes con nota 9 y sus actividades pasivas de la agrupación por propósito. Obsérvese también que la zona de color naranja en una de las secciones de la representación deformada (que representa martes) es el más grande, por tanto, el día más activo. Los datos de esta imagen son del Curso 1. (<http://analiticavisual.es/images/dialogo/TEEM2013/fig4-ERM-9-Tuesday-Passive.png>).

Si se pone atención otra vez en la Fig. 101, en la esquina superior derecha del panel central puede notarse que está señalado el porcentaje de las actividades de los datos filtrados con respecto al total del conjunto. Teniendo en cuenta este porcentaje se construyó la Fig. 100. Por otra parte, esta misma es el resultado obtenido después de analizar los datos con el fin de encontrar patrones en las frecuencias de las categorizaciones de las actividades y/o el rendimiento de los estudiantes.

Es muy interesante observar que, independientemente de la época del año en que se realizaron los cursos, del tipo de los cursos y los estudiantes, todos, los tres cursos, parecen seguir el mismo comportamiento.

Estos resultados indican que es necesario continuar la investigación y un nivel más profundo, con el fin de encontrar la razón o alguna idea más concreta del motivo del patrón. Este tipo de patrón ayuda al profesor a identificar estudiantes en riesgo o identificar las actividades que los estudiantes con mejores notas llevan a cabo y, a continuación, realizar los cambios en el programa del curso o se pueden tomar medidas de fomento a las actividades que realizan los mejores estudiantes. Como mínimo, estos resultados pueden ayudar a tomar decisiones, las cuales mejorarán las calificaciones de los estudiantes.

## VII.2 Evaluación de usabilidad

Esta evaluación de usabilidad se ha enfocado en la efectividad en el uso de las técnicas visuales y de interacción que soportan la tarea de análisis y retro-alimentación del proceso de aprendizaje, ya que estas fases son las que requieren el uso de la herramienta. También se ha considerado en la evaluación, la satisfacción del usuario en general, así como sus comentarios sobre diversos aspectos a mejorar.

La evaluación, al igual que las iteraciones del proceso de investigación, fue dividida de acuerdo a los perfiles que el modelo Vela tiene en cuenta, es decir, tanto los gestores académicos y profesores, como los estudiantes. De esta forma la evaluación se realizó en cuatro ciclos en total divididos en dos partes:

- La primera parte consistió en el análisis de una plataforma *online* desde el perfil de un profesor (experto), con la finalidad de que el usuario se familiarizara con la herramienta y también probará el rendimiento de las visualizaciones con una gran cantidad de datos.
- La segunda parte correspondió al análisis y autorregulación por parte de los estudiantes (usuarios). Ambos basados en el modelo descrito en la sección VI.1.2.

En la segunda parte de estos ciclos, como se mencionó anteriormente, tuvieron como realimentación el cuestionario realizado a los estudiantes del Máster Universitario en Las TICs en Educación: Análisis y Diseño de Procesos, Recursos y Prácticas Formativas de la Universidad de Salamanca (ver resultado de las encuestas Fig. 102). El análisis de los resultados se basa en el cálculo del valor de la media aritmética de todas las evaluaciones de los usuarios a cada una de las preguntas. Este tipo de evaluación, aunque sencilla, es bastante robusta ya que permite reconocer las preguntas peor evaluadas por parte de la mayoría de los usuarios.

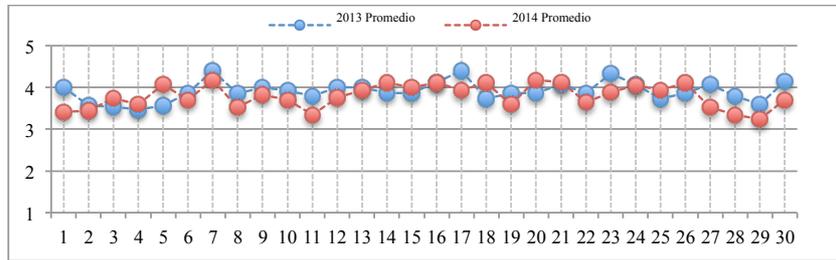


Fig. 102. Promedio de las Encuestas realizadas de los ciclos de investigación acción. Los valores van de 1 al 5, donde el 1 una respuesta "muy mal" y el 5 como "excelente".

Todas las preguntas detectadas con un valor de media relativamente bajo, son las que en mayor medida hubo que analizar para poder corregir estos aspectos específicos. En términos generales los resultados de la evaluación son aceptables, los valores de las medias son en general altos, lo que indica que los objetivos generales de las visualizaciones se han cumplido, sobre todo el hecho de ofrecer una herramienta intuitiva, fácil de usar y robusta para el análisis de datos educativos.

De acuerdo a las respuestas de los usuarios en las preguntas abiertas, se pudo obtener en conclusión que algunos de ellos consideraban que el proceso de analítica visual no era un asunto intuitivo a primera instancia. Es decir, que los usuarios requerían una breve explicación antes de comenzar a utilizar la herramienta. Este aspecto es importante considerarlo, aunque el hecho de que los usuarios no habían utilizado ninguna herramienta de analítica visual previamente a su uso con el prototipo del modelo VeLA es un factor que repercute en sus respuestas.



## Capítulo VIII Conclusiones y trabajos futuros

*"Not everything that counts can be counted, and not everything that can be counted counts"*

*Albert Einstein (Sign hanging in Einstein's office*

*at Princeton)*



Esta sección expone brevemente el proceso de investigación realizado, resaltando las problemáticas encontradas que dan lugar a los objetivos formulados, y de qué forma se lograron completar, o no, resultando en las principales conclusiones que se extraen del desarrollo y puesta en marcha de este trabajo.

Además se exponen algunos de los principales hitos de investigación que se esperan desarrollar para dar continuación a las líneas de investigación abiertas con esta tesis doctoral.

## VIII.1 Conclusiones

Inicialmente se abordó el tema de la historia y definición del *eLearning*, se definió qué es la analítica y se presentó cómo esta a través de la inteligencia de negocios ha llegado al ámbito educativo. Asimismo, se afrontó cómo la analítica en la educación (LA, AA, EDM, SLA) se concreta en un conjunto de herramientas de gestión de elementos de aprendizaje, pero también de análisis y mejora del aprendizaje, ya sean visuales o procedimentales, que se interconectan (por medio de servicios web o algún otro mecanismo). Al mismo tiempo, campos disciplinares como la visualización analítica y las técnicas de *infoVis* se han estado acercando a otras múltiples áreas de aplicación. Consecuentemente, se hizo una revisión exhaustiva de aplicaciones prácticas en las áreas relacionadas con Analítica del Aprendizaje, Analítica Académica, Analítica Visual, Minería de Datos para la Educación y Análisis de Redes Sociales Educativas.

Durante este proceso prospectivo de estudio se llega a percibir que la tendencia futura se encaminaba hacia una mayor adopción del aprendizaje en línea, por lo que habría una mayor oportunidad para instructores y administradores de supervisar la actividad de los estudiantes y la interacción con el contenido del curso y sus compañeros. También se detectaron un conjunto de deficiencias y problemáticas en las herramientas actuales, que se tomaron en cuenta en este trabajo. Además, todo esto condujo a la delimitación de una zona común donde se entrelazaban las áreas antes mencionadas (AA, LA, EDM, SLA, VA, e *infoVis*), estando poco explotadas las áreas de VA e *infoVis* en el ámbito educativo. Estos hechos constituyen las bases de la adopción de las técnicas de analítica visual en la educación, y que se han contemplado a partir de las propuestas más relevantes en el ámbito de la analítica académica y la analítica del aprendizaje y bajo el esquema de la analítica visual.

Así, bajo el paradigma “Analizar primero, mostrar lo importante, hacer un *zoom*, filtrar y analizar más profundamente, dar detalles bajo demanda, intervención” convergen en su adopción y adaptación, por un lado, el análisis de datos educativos, que es el proceso por el cual se crea nuevo conocimiento a partir de un conjunto de datos educativos diversos (LA, AA, EDM y SLA) y, por otro, el paradigma de la búsqueda de información propuesto (técnicas de *infoVis* y VA). Siendo este el punto de partida de la investigación sobre todas las restricciones, técnicas, métricas y especificaciones, conduciendo, como resultado del análisis de esta circunstancia teórica, a la creación de un modelo que toma en cuenta tanto el área educativa como la analítica y las técnicas de representación visual y la analítica visual, cumpliendo el objetivo principal de esta tesis.

Por tanto, la definición de un modelo de analítica visual para el *eLearning* (VeLA) es el principal resultado teórico de este trabajo. Este modelo comprende un ciclo interno de análisis y búsqueda de nueva información, tareas y flujos de trabajo para los principales procesos involucrados en la analítica visual de datos educativos. Estos flujos de trabajo se basan en las principales actividades definidas en aproximaciones existentes, como LA, AA, VA, propuestas que se han tomado como base en este trabajo, además de las restricciones y técnicas del área de *infoVis* y VA.

El trabajo de investigación resumido en los párrafos anteriores ha llevado a las siguientes conclusiones:

- El proceso de analítica en eLearning tiene cuatro componentes importantes: las restricciones de acceso a la información desde los diferentes perfiles (profesores estudiantes y gestores académicos), los datos obtenidos, su procesamiento y el análisis en sí mismo. Debido a esto, en este trabajo de tesis se han tomado en cuenta las técnicas de infoVis y VA, ya que permiten una representación visual de los datos, lo que facilita la abstracción y el entendimiento de su estructura, en comparación con el uso de técnicas no visuales; además, la interacción sobre las visualizaciones permite un mejor entendimiento de los datos representados por parte del usuario.
- El análisis de datos educativos es un proceso que resulta complejo en demasía, especialmente para fuentes de datos (estructuras de cursos y prácticas) creados por terceras partes. Así, el modelo propuesto no está vinculado a un contexto específico ni presupone una determinada concepción del aprendizaje o visión de la formación virtual. Tampoco es aplicable solo a un número reducido de escenarios ni depende de condiciones particularmente definidas
- Es posible la definición de los procesos, restricciones, técnicas y métricas básicas necesarias para el desarrollo de una analítica visual de datos educativos, mediante la creación de un modelo de visualización analítica en eLearning (VeLA). Dicho modelo parte de la teoría del ciclo, el modelo de referencia, y tiene en cuenta las restricciones y los elementos del modelo de dimensiones críticas de la analítica del aprendizaje; todo ello bajo un esquema de VA centrado en el usuario, considerándose el proceso de la exploración visual de datos de infoVis.
- El modelo propuesto supone un marco teórico sobre el que desplegar un proceso de análisis visual al eLearning. Así, con el objetivo de comprobar su validez y viabilidad, el modelo propuesto se materializó en un prototipo, que integra diversos mecanismos de visualización e interacción, propuestos específicamente para resolver la problemática del modelo VeLA. La implementación de dicho prototipo constituye un importante resultado que supone la consecución del segundo objetivo del presente trabajo de tesis. La fase de implementación incluyó la participación del usuario como el actor principal, pero requirió el soporte de diversos mecanismos y herramientas que permitieran su desarrollo de forma eficiente, correspondiendo parte de su desarrollo al usuario. Siguiendo esta idea y con la finalidad de obtener retroalimentación de los usuarios y de estimar el rendimiento y funcionalidad de la herramienta, se desarrollaron dos tipos de evaluaciones: una primera evaluación de usabilidad y una segunda evaluación de funcionalidad. La validación positiva del prototipo ha sido posible gracias a un enfoque de desarrollo centrado en el usuario y guiado por casos de estudio. Esta retroalimentación representa un aspecto crucial, ha sido fundamental para corregir y mejorar diversos aspectos, permitiendo la detección de algunos errores, que posteriormente fueron resueltos, así como la detección de aspectos que se debían mejorar, y otros no contemplados durante la fase del diseño original.

- La agrupación de las actividades para el apoyo en el análisis ha tenido resultados positivos. Así, el prototipo del modelo VeLA permite la personalización de las agrupaciones de las actividades, la creación y la codificación por color, el filtrado y la exclusión de actividades de forma dinámica. Además, se ha encontrado que en el análisis de textos la utilización de lematización aporta una disminución en el número de palabras y, por tanto, un uso más eficaz del espacio visual y un análisis más eficiente. En consecuencia, el prototipo del modelo VeLA permite principalmente, la personalización por medio de la eliminación y adición de nuevas agrupaciones de sinónimos y plurales de forma dinámica, lo que involucra un proceso creativo que permite incorporar conocimiento experto dentro del proceso de análisis, mejorando de esta forma el proceso de análisis de datos educativos.
- Ha sido especialmente relevante la incorporación de mecanismos de reutilización, que involucran una primera fase de análisis de los modelos y métricas de análisis actuales, con el objeto de tomar una decisión acerca de la posible reutilización de dichos modelos existentes. De este modo, el prototipo desarrollado aporta la posibilidad de exportación e importación tanto de las agrupaciones de las actividades, como de los sinónimos y plurales manipulados, mejorando de esta forma el proceso de análisis de datos educativos y la reutilización de durante el proceso.
- Las herramientas analizadas presentaban deficiencias tales como la saturación de información, la poca expresividad visual, así como la falta de mecanismos de interacción robustos. Los problemas identificados se atajaron a través de la implementación de cuatro representaciones para la información recuperada de los LMS. En concreto, las herramientas aplicadas están orientadas a analizar y descubrir las diferentes dimensiones de la información: temporal, social y semántica. Desde la perspectiva del contexto global de la herramienta, la navegación entre las diversas visualizaciones se desarrolla de acuerdo a los flujos de trabajo propuestos a través de niveles jerárquicos de los datos dependiendo del perfil y los privilegios, lo que supone dotar a la herramienta desarrollada de una excelente flexibilidad que se traduce en una capacidad de adaptación para responder a las preguntas que van surgiendo en el proceso de análisis.
- Las herramientas analizadas presentan dependencias a plataformas específicas, y en algunos casos, a una versión específica de la plataforma a analizar. Por tal motivo, las estrategias en el diseño del prototipo del modelo VeLA han sido, por un lado, la obtención de los datos de forma independiente de la plataforma y la separación de la funcionalidad analítica entre las diversas técnicas implementadas; y, por otro lado, el diseño de ciertas técnicas específicas para el análisis de datos educativos. Estas estrategias han permitido establecer una secuencia definida en cuanto a los flujos de análisis, y en consecuencia, evitar la saturación de información que se muestra al usuario, que suponía una enorme barrera en el caso de otras herramientas. Así, la herramienta define claramente la secuencia de actividades en los procesos, a través de la navegación de las técnicas visuales. El resultado es que los procesos se realizan de forma más clara e intuitiva por el usuario, lo que facilita el desarrollo de estas actividades.

- El análisis de conjuntos de datos educativos está directamente relacionado con la participación del estudiante, así como la evaluación de las actividades de aprendizaje implementadas. Las preguntas relacionadas con cómo y cuándo los estudiantes están conectados y activos, y qué actividades promueven la participación del estudiante, pueden responderse a través de la observación y el análisis del comportamiento en línea de los estudiantes. En este sentido, el estudio proporciona una plataforma para investigar nuevas suites de herramientas de diagnóstico, técnicas y metodologías que pueden, a su vez, proporcionar nuevas oportunidades y nuevos conjuntos de datos para informar al instructor y fomentar la reflexión para la mejora continua y progresiva de la práctica pedagógica.

A modo de colofón, cabe reflexionar sobre el hecho de que el prototipo del modelo VeLA consiste en un conjunto de técnicas muy especializadas e intuitivas, que utilizan la capacidad de procesamiento de los ordenadores y que permiten al usuario utilizar su capacidad cognitiva visual para la realización de un análisis visual de los datos educativos extraídos de una plataforma de aprendizaje en línea. Por otro lado, se consideró una interacción robusta, con un conjunto de funcionalidades que no se limitan simplemente a visualizar la información, sino que permiten al usuario, entre otras cosas, tener múltiples vistas y perspectivas de los datos educativos, así como navegar a través de los diferentes niveles de abstracción de los datos. En suma, se logra la mejora significativa de los procesos de análisis de datos educativos, con respecto a otras propuestas. En términos generales, los usuarios han evaluado la herramienta satisfactoriamente en cuanto a las visualizaciones y formas de interacción. Esta implementación ofrece mejores prestaciones con respecto a las propuestas analizadas actuales, en cuanto a visualización e interacción y, por ende, a su capacidad para desarrollar el análisis. Todo lo dicho da pie a enunciar la siguiente conclusión general:

*Este trabajo de investigación ha permitido: en primer lugar, detectar la problemática principal de las actuales herramientas; en segundo lugar, en base a los problemas detectados, plantear soluciones muy específicas al dominio de la Analítica Académica y Analítica del Aprendizaje por medio de las técnicas visuales y de interacción propuestas; y, en tercer lugar, por medio de la propuesta del modelo VeLA enfocado a las necesidades de las actividades de análisis de los diferentes tipos de datos y perfiles educativos requeridos y mediante la exitosa aplicación de la Analítica Visual en el análisis y comprensión del proceso de enseñanza y aprendizaje. Es posible, por tanto, explicitar más claramente los patrones de uso de las plataformas de eLearning, lo que ha permitido demostrar el enorme potencial del enfoque aplicado y abrir una prometedora línea de investigación.*

## VIII.2 Trabajos futuros

Este trabajo de tesis se ha basado en los principales procesos involucrados en la analítica visual aplicada al *eLearning*, sin embargo, existen algunos otros que han quedado fuera de esta investigación por sobrepasar los objetivos establecidos. Las líneas de investigación futuras se relacionan principalmente con estos diversos procesos que no han podido abordarse hasta el momento.

Entre estas líneas, se puede mencionar primeramente un aspecto que ha sido discutido en varias fuentes bibliográficas y que se ha quedado para su incorporación en proyectos posteriores: el uso de una estructura ontológica base para el análisis semántico del contenido y la actividad en los entornos virtuales educativos. Esta posibilidad fortalecería la implementación de estrategias de lematización propuestas en esta tesis.

Otra línea de investigación posible es la utilización de aplicaciones como servicios por medio de la nube (*cloud computing*), que mediante de la cual se facilitaría (aún más) la reutilización en el prototipo de VeLA. Especialmente interesante sería explorar el desarrollo de herramientas en la nube mediante técnicas y mecanismos de visualización, siguiendo el enfoque de la Analítica Visual en *eLearning*.

Un tercer trabajo futuro abordaría el desarrollo y optimización del análisis con el objetivo de trasladarlo a los dispositivos móviles, permitiendo el intercambio de información, especialmente en ambientes distribuidos que incorpore algunos logros significativos en relación con la Web Semántica. También el perfeccionamiento de las técnicas y/o implementación de nuevos algoritmos de análisis y agrupación de los datos del prototipo VeLA queda abierto como proyecto futuro.

La tecnología actual permite el almacenamiento de la información suficiente para hacer frente al tipo de estudios realizados en esta tesis. Sin embargo, el tamaño de los conjuntos de datos requiere de herramientas diseñadas *ad hoc*, por lo que se abre una cuarta línea de investigación que se dirige hacia la implementación de técnicas de *BigData*.

Se pretende que las líneas de investigación descritas sean enfocadas desde la perspectiva del modelo VeLA, de la misma forma en que se han propuesto las soluciones en este trabajo de tesis.

## VIII.3 Publicaciones relacionadas con la tesis

Entre los resultados obtenidos se puede enumerar diversas publicaciones en congresos y revistas internacionales, que validan el trabajo de investigación de esta tesis al haber sido sometido al escrutinio de diversos investigadores en el área. Esta lista se detalla a continuación:

### VIII.3.1 Revistas

- **Gómez-Aguilar, D. A., Hernández-García, Á., García-Peñalvo, F. J., & Therón, R.** (2015). "Tap into visual analysis of customization of grouping of activities in eLearning". *Computers in Human Behavior*, 47(0), 60-67. doi: 10.1016/j.chb.2014.11.001. **Q1 - IF: 2.273.**

- **Gómez-Aguilar, Diego-Alonso, García-Peñalvo, Francisco-José y Therón, Roberto**, (2014). "Analítica visual en e-learning". *El Profesional de la Información*, 23, 3, 236-245. **Q3 - IF: 0.402**
- **García Penalvo, Francisco José; Conde González, Miguel Ángel; Bravo Martín, Sergio; Gómez Aguilar, Diego Alonso; y Therón Sanches, Roberto**, (2011). "Visual Analysis of a Moodle-based Object Oriented Programming Course". *IJCA Proceedings on Design and Evaluation of Digital Content for Education (DEDCE)*, 1, 0, 8-14.
- **Conde, Miguel A.; Gomez, Diego A.; Pozo, Alberto Del; y Garcia, Francisco J.**, (2011). "Web services layer for Moodle 2.0: a new area of possibilities in web based learning". *Int. J. Technol. Enbanc. Learn.*, 3, 3, 308-321. DOI= [10.1504/ijtel.2011.040227](https://doi.org/10.1504/ijtel.2011.040227).
- **Garcia, Juan; Aguilar, Diego A. Gomez; Gonzalez, Antonio; Garcia, Francisco J.; y Theron, Roberto**, (2010). "A middleware framework to create data structures for a visual analytics object oriented approach". *International Journal of Knowledge and Learning*, 01/01/, 6, 2, 256-267. DOI= [10.1504/IJKL.2010.034757](https://doi.org/10.1504/IJKL.2010.034757).
- **Morales Morgado, Erla; Gómez Aguilar, Diego Alonso; García Peñalvo, Francisco José; y Therón Sánchez, Roberto**, (2009). "Supporting the Quality of Learning Objects Through Their Ranking Visualization". *International Journal of Emerging Technologies in Learning (iJET)*, 4, pp. 24-29.
- **Gómez Aguilar, Diego Alonso, Theron, Roberto y García Peñalvo, Francisco José**, (2009). "Semantic spiral timelines used as support for e-learning". *Journal of Universal Computer Science (j-jucs)*, April, 15, 7, 1526-1545. **Q4 - IF: 0.669**

### VIII.3.2 Congresos

- **Conde González, Miguel Ángel; García Pealvo, José Francisco; Gómez Aguilar, Diego Alonso; y Therón Sanches, Roberto**, (2014). Visual learning analytics techniques applied in software engineering subjects. En: 2014 IEEE Frontiers in Education Conference Proceedings (Madrid, Spain, 22-25 Oct. Year), IEEE, 3009-3017.
- **Gómez Aguilar, Diego Alonso, García Peñalvo, Francisco José y Therón Sánchez, Roberto**, (2014). Modelo de Analítica Visual para Datos Educativos. En: En Sistemas y Tecnologías de Información. Actas de la 9ª Conferencia Ibérica de Sistemas y Tecnologías de Información – CISTI 2014 (Barcelona, España, 18-21 Jun Year), AISTI (Asociación Ibérica de Sistemas y Tecnologías de Información), 1132-1137.

- **Gómez Aguilar, Diego Alonso, Therón, Roberto y García-Peñalvo, Francisco José**, (2013). Reveal the Relationships among Students Participation and Their Outcomes on E-Learning Environments: Case Study. En: *Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on* (15-18 July 2013 2013c), 443-447. DOI= [10.1109/ICALT.2013.136](https://doi.org/10.1109/ICALT.2013.136).
- **Gómez Aguilar, Diego Alonso, García Peñalvo, Francisco José y Therón, Roberto**, (2013). Tap into visual analysis of the customization of grouping of activities in eLearning. En: *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality* (Salamanca, Spain2013), ACM, 2536575, 253-259. DOI= [10.1145/2536536.2536575](https://doi.org/10.1145/2536536.2536575).
- **Gómez Aguilar, Diego Alonso, García-Peñalvo, Francisco José y Therón, Roberto**, (2013). Evaluación visual de las relaciones entre participación de los estudiantes y sus resultados en entornos de e-learning. En: *XV Simposio Internacional de Tecnologías de la Información y las Comunicaciones en la Educación Conjuntamente con el Congreso Español de Informática (CEDI 2013)* (2013).
- **Gómez Aguilar, Diego Alonso; Conde-Gonzalez, Miguel Ángel; García-Peñalvo, Francisco José; y Therón, Roberto**, (2011). Reaveling the Evolution of Semantic Content through Visual Analysis. En: *Advanced Learning Technologies (ICALT), 2011 11th IEEE International Conference on* (6-8 July 2011 2011), 450-454. DOI= [10.1109/ICALT.2011.141](https://doi.org/10.1109/ICALT.2011.141).
- **Gomez-Aguilar, Diego; Conde-Gonzalez, Miguel; Theron, Roberto; y Garcia-Peñalvo, Francisco**, (2011). "Supporting Moodle-Based Lesson through Visual Analysis". En *Human-Computer Interaction – INTERACT 2011*, P. CAMPOS, N. GRAHAM, J. JORGE, N. NUNES, P. PALANQUE and M. WINCKLER Eds. Springer Berlin Heidelberg, 604-607. DOI= [10.1007/978-3-642-23768-3\\_93](https://doi.org/10.1007/978-3-642-23768-3_93).
- **Gómez Aguilar, Diego Alonso; Conde González, M. Á.; Therón Sánchez, Roberto; y García Peñalvo, F. J.**, (2011). Improving Moodle-based eLearning through visual analysis, a case study. En: *The 2011 CAL conference, Learning Futures: Education, Technology, Sustainability*. (Manchester, UK 13-15 April 2011 2011), Computers & Education.
- **Gómez Aguilar, Diego Alonso; Conde González, Miguel Ángel; Therón Sánchez, Roberto; y García Peñalvo, Francisco José**, (2011). Supporting Moodle-based lesson of Software Engineering through visual analysis. En: *XII Congreso Internacional de Interacción Persona-Ordenador – Interacción 2011* (Lisboa, Portugal, 2-6 Sep. 2011), Ibergarceta Publicaciones S.L., Madrid, 319-328.

- **Gómez Aguilar, Diego Alonso; Conde González, Miguel Ángel; Therón, Roberto; y García Peñalvo, Francisco José,** (2010). "Retrieval information model for Moodle data visualization". *Advanced Learning Technologies 2010. ICALT 2010. 10th IEEE International Conference on*, July 5-7, Product Number E4055, 526-527. DOI= [10.1109/ICALT.2010.150](https://doi.org/10.1109/ICALT.2010.150).
- **Conde González, M. Á.; Gómez Aguilar, Diego Alonso; del Pozo de Dios, Alberto; Therón Sánchez, Roberto; y garcía Peñalvo, F. J.,** (2010). Moodle 2.0 Web Services Layer and Its New Application Contexts. En: *1st International Conference, TECH-EDUCATION 2010*. (Athens, Greece.2010), pp. 110-116.
- **Gómez Aguilar, Diego Alonso; Conde González, Miguel Ángel; Therón, Roberto; y García Peñalvo, Francisco José,** (2010). Modelo basado en Servicios Web para la Visualización de Moodle. En: *En las Actas del XI Congreso Internacional de Interacción Persona-Ordenador – Interacción 2010. Celebrado dentro del III Congreso Español de Informática* (Universidad Politécnica de Valencia, España, 7-10 de septiembre de 2010), Madrid: Ibergarceta Publicaciones S.L., 445-454.
- **García, Juan; González Torres, Antonio; Gómez Aguilar, Diego Alonso; Therón, Roberto; y García Peñalvo, Francisco José,** (2009). "A Visual Analytics Tool for Software Project Structure and Relationships among Classes". En *Smart Graphics*, A. BUTZ, B. FISHER, M. CHRISTIE, A. KRÜGER, P. OLIVIER and R. THERÓN Eds. Springer Berlin Heidelberg, 203-212. DOI= [10.1007/978-3-642-02115-2\\_17](https://doi.org/10.1007/978-3-642-02115-2_17).
- **García Navarro, Juan Francisco ; Gómez Aguilar, Diego Alonso; Gonzalez Torres, Antonio ; García Peñalvo, Francisco José; y Therón Sánchez, Roberto,** (2009). A Middleware Framework to Create and Manage Data Structures for Visual Analytics. En: *Second World Summit on the Knowledge Society, WSKS 2009* (September 2009 2009), Springer pp. 466-473.
- **Gómez Aguilar, Diego Alonso, Theron, Roberto y García Peñalvo, Francisco José,** (2008). Understanding Educational Relationships in Moodle with ViMoodle. En: *Advanced Learning Technologies, 2008. ICALT '08. Eighth IEEE International Conference on* (1-5 July 2008), 954-956. DOI= [10.1109/ICALT.2008.276](https://doi.org/10.1109/ICALT.2008.276).
- **Gómez Aguilar, Diego Alonso, Therón Sánchez, Roberto y García Peñalvo, Francisco José,** (2008). Semantic Spiral Timeline como apoyo al e-learning. En: *Informática Educativa, 2008. SIIE 2008. El X Simposio Internacional de* (2008).
- **Morales Morgado, Erla, Gómez Aguilar, Diego Alonso y García Peñalvo, Francisco José,** (2008). Herramienta para la Valoración de Objetos de Aprendizaje mediante Ranking en Moodle. En: *Informática Educativa, 2008. SIIE 2008. El X Simposio Internacional de* (Salamanca 2008).

- **Gómez Aguilar, Diego Alonso, Therón Sánchez, Roberto y García Peñalvo, Francisco José**, (2007). "ViMoodle: Visualizando Moodle". En *I Congreso Internacional de Tecnología, Formación y Comunicación (EuniverSALearning'07)*, LOGO Ed. Asociación Española de Estudios sobre Lengua, Pensamiento y Cultura Clásicas, Salamanca.
- **Gómez Aguilar, Diego Alonso, Therón Sánchez, Roberto y García Peñalvo, Francisco José**, (2007). Visual analytics, for virtual education platforms. En: *I Congreso Internacional de Tecnología, Formación y Comunicación* (2007), LOGO, CEUR-WS.
- **Therón, Roberto; Santamaría, Rodrigo; García, Juan; Gómez Aguilar, Diego Alonso; y Paz-Madrid, Vadim**, (2007). "Overlapper: movie analyzer". *Infovis Conference Compendium*, 140-141.

### VIII.3.3 Capítulos de libro y Postproceedings

- **Gómez Aguilar, Diego Alonso, García Peñalvo, Francisco José y Therón Sánchez, Roberto**, 2014. "Visual analytical model for educational data". En *9th Iberian Conference on Information Systems and Technologies (CISTI)* Institute of Electrical and Electronics Engineers. IEEE, 1-6. DOI= <http://dx.doi.org/10.1109/CISTI.2014.6877098>.
- **García Penalvo, Francisco José; Conde González, Miguel Ángel; Bravo Martín, Sergio; Gómez Aguilar, Diego Alonso; y Therón Sanches, Roberto**, 2011. "Visual Analysis of a Moodle-based Object Oriented Programming Course". En *International Journal of Computers Applications Proceedings on Design and Evaluation of Digital Content for Education (DEDCE)* Foundation of Computer Science, USA, 13-19. ISBN: 978-93-80746-64-6
- **Gómez Aguilar, Diego Alonso; Suárez Guerrero, Cristobal; Therón Sánchez, Roberto; y García Peñalvo, Francisco José**, 2010. "Visual Analytics to Support e-Learning". En *Advances in Learning Processes*, M.B. ROSSON Ed. In-Tech, Croatia, 207-228.
- **Gómez Aguilar, Diego Alonso, Therón Sanches, Roberto y García Peñalvo, Francisco José**, 2008. "Visual Analytics for Virtual Education Platforms". En *Post-proceedings of the International Conference on Technology, Training and Communication eUniverSALearning 2007*, A. LÓPEZ EIRE, F.J. GARCÍA PEÑALVO, A.M. SEOANE PARDO and E.M. MORALES MORGADO Eds. CEUR Workshop Proceedings, Salamanca, Spain. ISBN: 1613-0073

## Capítulo IX Anexos

### IX.1 Encuesta<sup>27</sup>

Cuestionario de efectividad del uso de vMoodle para el apoyo en los ambientes de aprendizaje en línea.

#	Valore los siguientes apartados relacionados con la herramienta.	1	2	3	4	5
1	Ayuda a encontrar debilidades aprendizaje con su aplicación en situaciones reales					
2	Ayuda a salvar la distancia entre teoría y practica					
3	Proporciona una realimentación adecuada de la asignatura					
4	Implica a los participantes en su propio aprendizaje					
5	Crea una actitud de participación activa.					
6	Las metas generales del programa de formación están comprometidas con la mejora de la calidad del aprendizaje del alumnado					
7	El programa asume un modelo de profesor como profesional reflexivo, crítico e innovador, y del eLearning como entorno de formación					
8	El programa responde a necesidades formativas fundamentadas en estudios y análisis evaluativos que están relacionadas directamente con la calidad del aprendizaje del alumnado.					
9	Los objetivos del programa son relevantes y están vinculados con las necesidades de los profesores y estudiantes					
10	Los objetivos muestran con claridad la incidencia del programa en los resultados de aprendizaje de los estudiantes					
11	El programa integra las diferentes fases del proceso de enseñanza/aprendizaje					
12	El diseño del programa toma en consideración los diferentes roles que el profesorado puede desempeñar en su tarea profesional: docente, orientador, tutor, director, formador					
13	El programa presenta a los profesores múltiples perspectivas de análisis y se fundamentan en los hallazgos de la investigación educativa					
14	El programa facilita que los profesores profundicen según sus intereses y colaboren para construir conocimiento					
15	El programa ayuda a prevenir al profesorado acerca de las posibles dificultades de aprendizaje que los estudiantes pueden encontrarse					
16	El programa promueve la utilización de estrategias de aprendizaje basadas en la reflexión e indagación, mediante las cuales el profesorado analiza sobre su enseñanza para diseñar procesos de mejora docente					

<sup>27</sup> <http://analiticavisual.com.es/images/dialgoag/Tesis/Encuesta.pdf>

<b>-----AYUDA A MEJORAR COMPETENCIAS DEL TIPO:-----</b>						
17	Organización del tiempo					
18	Resolución de problemas					
19	Toma de decisiones					
20	Planificación					
21	Búsqueda de información					
22	Creatividad					
23	Gestión por objetivos					
24	Gestión de proyectos					
25	Estimulación intelectual					
28	Comunicación interpersonal					
29	Trabajo en equipo					
30	Tratamiento de conflictos					
31	Negociación					
32	Liderazgo					

## IX.2 Acrónimos

### A

AA: Analítica Académica, 44, 45, 67, 77, 79, 80, 90, 91, 92, 96, 128, 129, 154, 186, 198, 205, 206, 207, 208, 215, 237, 238, 270

### B

BI: Inteligencia de negocios, 68, 72, 76, 79

### C

CBT: Educación Basada en el Uso del Ordenador, 32, 52, 53, 58  
 CMS: Sistemas de Administración de Cursos, 26, 31, 44, 45, 51, 62, 63, 78, 79, 81, 212, 242, 244

### D

DDDM: Toma de decisiones basada en los datos, 96, 97, 128

### E

EDM: Minería de Datos para la Educación, 67, 70, 82, 83, 85, 90, 92, 93, 95, 96, 189, 205, 206, 207, 237, 270  
*eLearning*: Aprendizaje electrónico, 1, V, VI, VII, VIII, 23, 28, 32, 33, 34, 41, 42, 43, 45, 46, 47, 49, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 65, 75, 76, 131, 132, 133, 134, 136, 137, 138, 142, 143, 145, 156, 160, 161, 168, 182, 184, 186, 191, 197, 198, 202, 203, 204, 205, 206, 207, 209, 210, 211, 224, 227, 238, 243, 245, 251, 257, 263, 270, 271, 274, 276, 277, 281  
 ERP: Sistemas de planificación de recursos empresariales, 75

### G

GRIAL: Grupo de Investigación en Interacción y eLearning, III, 46, 47

### I

*I + D + i*: Investigación, Desarrollo e Innovación, 24  
 IBL: Aprendizaje basado en la Red, 53  
 IBT: Instrucción basada en la Red, 53, 59

ICT: Tecnologías de la Información y la Comunicación, 24, 28, 29, 33, 35, 42, 49, 51, 52, 55, 57, 59, 69, 76, 77, 130

ILEs: Entornos de Aprendizaje Inteligentes, 32

*infoVis*: Visualización de la información, 103, 121, 128, 182, 270

IR: Investigadores de las instituciones, 77

IT: Tecnología de la información, 77

ITS: Sistemas inteligentes de tutoría, 82

### L

LA: Analítica del Aprendizaje, 29, 44, 45, 67, 82, 83, 84, 85, 86, 87, 88, 90, 91, 92, 93, 96, 128, 129, 146, 154, 185, 186, 188, 189, 191, 192, 195, 196, 197, 198, 202, 205, 206, 207, 208, 215, 231, 237, 238, 270

LCMS: Sistema de administración de contenidos de aprendizaje, 26, 27, 28, 31, 45, 62, 63, 131, 133, 137

LMS: Sistema de administración de cursos en línea, 26, 27, 28, 31, 34, 41, 42, 44, 45, 47, 51, 54, 62, 63, 64, 75, 76, 131, 133, 134, 137, 141, 142, 143, 148, 151, 154, 156, 158, 159, 160, 161, 175, 176, 181, 185, 192, 203, 204, 206, 213, 219, 224, 228, 230, 233, 239, 243, 272

LORs: Repositorios de objetos de aprendizaje, 31

LS: Estilos de aprendizaje de los estudiantes, 158, 159

### M

MOOC: Cursos Online Masivos Abiertos, 36

MVC: Modelo-Vista-Controlador, 212

### O

OA: Objeto de Aprendizaje, 27, 47

### P

PCA: Análisis de componentes principales, 158

PLE: Entornos Personales de Aprendizaje, 33, 34, 64, 95

### S

SGA: Sistemas de Gestión del Aprendizaje, 54

SLA: Análítica Social para la Educación, 67, 93, 94, 95, 96, 153, 189, 236, 270  
 SNA: Análisis de Redes Sociales, 82, 151, 153, 186, 207  
 SOA: Arquitectura orientada a servicios, 34, 64, 65  
 SRL: Autorregulación del aprendizaje, 81, 82

## V

VA: Análítica Visual, 28, 44, 45, 83, 117, 118, 119, 120, 121, 128, 129, 131, 132, 133, 134, 182, 186, 193, 195,

196, 197, 198, 202, 206, 207, 208, 215, 237, 270  
 VeLA: Análítica Visual en eLearning, VI, VIII, 186, 188, 190, 192, 194, 196, 197, 207, 208, 210, 211, 212, 213, 233, 234, 235, 237, 239  
 Vis: La visualización, 103  
 VisUSAL: Grupo de Visualización de la Universidad de Salamanca, 46  
 VLE: Entorno de Aprendizaje Virtual, 26, 31, 37, 51, 94, 105, 106, 252, 257

## W

WBT: Formación basada en la Web, 58

## Referencias

- Adell, J., & Castañeda, L. (2010). Los Entornos Personales de Aprendizaje (PLEs): una nueva manera de entender el aprendizaje. In R. Roig Vila & M. Fiorucci (Eds.), *Claves para la investigación en innovación y calidad educativas. La integración de las Tecnologías de la Información y la Comunicación y la Interculturalidad en las aulas. Strumenti di ricerca per l'innovazione e la qualità in ambito educativo. La Tecnologie dell'Informazione e della Comunicazione e l'interculturalità nella scuola.*: Marfil Roma TRE Università degli Studi.
- Adell, J., & Castañeda, L. (2013). El ecosistema pedagógico de los PLEs. In L. Castañeda & J. Adell (Eds.), *Entornos Personales de Aprendizaje: claves para el ecosistema educativo en red*: **Alcoy: Marfil**.
- Afgan, N. H., & Carvalho, M. G. (2010). The Knowledge Society: A Sustainability Paradigm. *CADMUS: The Personality of a Journal*, 1(1), 27-41.
- Agudo-Peregrina, Á. F., Hernández-García, Á., & Iglesias-Pradas, S. (2012). *Predicting academic performance with learning analytics in virtual learning environments: A comparative study of three interaction classifications*. Paper presented at the Computers in Education (SIIE), 2012 International Symposium on.
- Agudo-Peregrina, Á. F., Iglesias-Pradas, S., Conde-González, M. Á., & Hernández-García, Á. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior*, 31(0), 542-550. doi: <http://dx.doi.org/10.1016/j.chb.2013.05.031>
- Aigner, W., Miksch, S., Muller, W., Schumann, H., & Tominski, C. (2008). Visual Methods for Analyzing Time-Oriented Data. *Visualization and Computer Graphics, IEEE Transactions on*, 14(1), 47-60. doi: 10.1109/TVCG.2007.70415
- Aigner, W., Miksch, S., Müller, W., Schumann, H., & Tominski, C. (2007). Visualizing time-oriented data—A systematic view. *Computers & Graphics*, 31(3), 401-409. doi: 10.1016/j.cag.2007.01.030
- Akinsola, M. K., Tella, A., & Tella, A. (2007). Correlates of academic procrastination and mathematics achievement of university undergraduate students. *Eurasia Journal of Mathematics, Science & Technology Education*, 3(4), 363-370.
- Al-Abbasi, D. (2012). The Effects of Modality and Multimedia Comprehension on the Performance of Students with Varied Multimedia Comprehension Abilities when Exposed to High Complexity, Self-paced Multimedia Instructional Materials. *Journal of Educational Multimedia and Hypermedia*, 21(3), 215-239.
- Al-Ajlan, A., & Zedan, H. (2008). *The extension of web services architecture to meet the technical requirements of virtual learning environments (Moodle)*. Paper presented at the

- Computer Engineering & Systems, 2008. ICCES 2008. International Conference on.
- Alesso, H. P., & Smith, C. F. (2008). *Thinking on the Web: Berners-Lee, Godel and Turing*: Wiley-Interscience.
- Ali, L., Asadi, M., Gašević, D., Jovanović, J., & Hatala, M. (2013). Factors influencing beliefs for adoption of a learning analytics tool: An empirical study. *Computers & Education*, 62(0), 130-148. doi: 10.1016/j.compedu.2012.10.023
- Alier Forment, M., Casany Guerrero, M. J., Conde González, M. Á., & García Peñalvo, F. J. (2010). Interoperability for LMS: the missing piece to become the common place for e-learning innovation. *International Journal of Knowledge and Learning*, 6(2), 130-141.
- Alier Forment, M., Casany Guerrero, M. J., Conde González, M. Á., García Peñalvo, F. J., & Severance, C. (2009). Interoperability for LMS: The Missing Piece to Become the Common Place for Elearning Innovation. In M. Lytras, E. Damiani, J. Carroll, R. Tennyson, D. Avison, A. Naeve, A. Dale, P. Lefrere, F. Tan, J. Sipior, & G. Vossen (Eds.), *Visioning and Engineering the Knowledge Society. A Web Science Perspective* (Vol. 5736, pp. 286-295): Springer Berlin Heidelberg.
- Alier Forment, M., Casany Guerrero, M. J., Poch, J. P., Galanis, N., Mayol, E., Conde González, M. Á., & García Peñalvo, F. J. (2013). Integration of Google Docs as a Collaborative Activity within the LMS Using IMS BasicLTI *Information Systems, E-learning, and Knowledge Management Research* (pp. 677-683): Springer.
- Almeida, L. H., de M. Lourenco, R. A., Meiguins, B. S., & Meiguins, A. S. G. (2009). *WebPrisma: An Interactive Web-Based Tool for Exploratory Visualization Using Multiple Coordinated Views*. Paper presented at the Information Visualisation, 2009 13th International Conference. 10.1109/IV.2009.17
- Amar, R., Eagan, J., & Stasko, J. (2005, 23-25 Oct.). *Low-Level Components of Analytic Activity in Information Visualization*. Paper presented at the Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization.
- Analytics, G. (2013). Learning Analytics. Retrieved Sep. 2014, from <http://www.gilfuseducationgroup.com/learning-analytics>  
<http://www.gilfusacademicanalytics.com/capabilities/learning-analytics>
- Anane, R., Crowther, S., Beadle, J., & Theodoropoulos, G. (2004). eLearning content provision. *Database and Expert Systems Applications, 2004. Proceedings. 15th International Workshop on*, 6, 420-425.
- Anderson, T., & Dron, J. (2011). Three generations of distance education pedagogy. *International Review of Research in Open and Distance Learning*, 12(3), 80-97.
- André, P., Wilson, M. L., Russell, A., Smith, D. A., Owens, A., & m.c, s. (2007). *Continuum: designing timelines for hierarchies, relationships and scale*. New York, NY, USA.
- Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M.-J., MacEachren, A., & Wrobel, S. (2007). Geovisual analytics for spatial decision support: Setting the research agenda.

- International Journal of Geographical Information Science*, 21(8), 839-857. doi: 10.1080/13658810701349011
- Arens, Y., Hovy, E. H., & Vossers, M. (1992). On the knowledge underlying multimedia presentations. 280-306. <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA278690>
- Armstrong, J., & Anthes, K. (2001). How Data Can Help: Putting Information to Work to Raise Student Achievement. *American School Board Journal*. <http://www.ecs.org/html/Document.asp?chouseid=3178>
- Averbukh, V. L., Konovalov, A. V., & Vorzopov, V. V. (1997). *An approach to evaluation of Software Visualization*. Paper presented at the CHI '97 Extended Abstracts on Human Factors in Computing Systems, Atlanta, Georgia.
- Azcorra Saloña, A., Bernardos Cano, C. J., Gallego Gómez, Ó., & Soto Campos, I. (2001). INFORME SOBRE EL ESTADO DE LA TELEEDUCACIÓN EN ESPAÑA. In J. Solá (Ed.), *Estado de la teleeducación en España*. España: Departamento de Tecnologías de las Comunicaciones, Universidad Carlos III de Madrid & Asociación de Usuarios de Internet.
- Bach, C. (2010). *Learning Analytics: Targeting Instruction, Curricula and Student Support*. Paper presented at the Education and Information Systems, Technologies and Applications: EISTA 2010, Orlando, Florida, USA. [http://www.iis.org/CDs2010/CD2010SCI/EISTA\\_2010/PapersPdf/EA655ES.pdf](http://www.iis.org/CDs2010/CD2010SCI/EISTA_2010/PapersPdf/EA655ES.pdf)
- Baepler, P., & Murdoch, C. J. (2010). Academic Analytics and Data Mining in Higher Education. *International Journal for the Scholarship of Teaching and Learning*, 4(2), 9. <http://digitalcommons.georgiasouthern.edu/ij-sotl/vol4/iss2/17>
- Baer, L., & Campbell, J. (2012). From Metrics to Analytics, Reporting to Action: Analytics' Role in Changing the Learning Environment *Game Changers: Education and Information Technologies* (pp. 53-65).
- Bailey, A., Henry, T., McBride, L., & Puckett, J. (2011). Unleashing the Potential of Technology in Education
- Baker, R. S. J. d. (2011). Data Mining for Education. *International Encyclopedia of Education*, 7, 112-118. [http://www.columbia.edu/~rsb2162/Encyclopedia\\_Chapter\\_Draft\\_v10-fw.pdf](http://www.columbia.edu/~rsb2162/Encyclopedia_Chapter_Draft_v10-fw.pdf)
- Baker, R. S. J. d., & Yacef, K. (2009). The State of Educational Data Mining in 2009: A Review and Future Visions. *Journal of Educational Data Mining*, 1(1), 1-3. <http://www.educationaldatamining.org/JEDM/index.php/JEDM/article/view/8/2>
- Bakharia, A., & Dawson, S. (2011). *SNAPP: a bird's-eye view of temporal participant interaction*. Paper presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge, Banff, Alberta, Canada.

- Baraniuk, R. G., Burrus, C. S., Hendricks, B. M., Henry, G. L., Hero, A. O., Johnson, D. H., . . . Wilson, W. L. (2002). *Connexions: DSP education for a networked world*. Paper presented at the Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on.
- Barbosa León, H., & García Peñalvo, F. J. (2005, July 7-9). *Importance of Online Assessment in the E-learning Process*. Paper presented at the 6th International Conference on Information Technology Based Higher Education and Training, 2005. ITHET 2005., Santo Domingo, Dominican Republic.
- Barbosa León, H., García Peñalvo, F. J., Rodríguez Conde, M. J., Morales Morgado, E. M., & Ordoñez de Pablos, P. (2012). Adaptive assessments using open specifications. *International Journal of Distance Education Technologies (IJDET)*, 10(4), 56-71. doi: 10.4018/jdet.2012100105
- Barry, B., & Fulmer, I. S. (2004). The Medium and the Message: The Adaptive Use of Communication Media in Dyadic Influence. *The Academy of Management Review*, 29(2), 272-292. doi: 10.2307/20159033
- Bateman, S., Gutwin, C., & Nacenta, M. (2008). *Seeing things in the clouds: the effect of visual features on tag cloud selections*. Paper presented at the Proceedings of the nineteenth ACM conference on Hypertext and hypermedia, Pittsburgh, PA, USA.
- Batty, M. (2006). Rank clocks. *Nature*, 444(7119), 592-596. doi: [http://www.nature.com/nature/journal/v444/n7119/supinfo/nature05302\\_S1.html](http://www.nature.com/nature/journal/v444/n7119/supinfo/nature05302_S1.html)
- Beck, B. L., Koons, S. R., & Milgrim, D. L. (2000). Correlates and Consequences of Behavioral Procrastination: The Effects of Academic Procrastination, Self-Consciousness, Self-Esteem and Self-Handicapping. *Journal of Social Behavior & Personality*, 15(5), 3-13.
- Becker, R. A., & Cleveland, W. S. (1987). Brushing scatterplots. *Technometrics*, 29(2), 127-142. doi: 10.2307/1269768
- Becker, R. A., Cleveland, W. S., & Wilks, A. R. (1987). Dynamic Graphics for Data Analysis. *Statistical Science*, 2(4), 355-383. doi: 10.2307/2245523
- Becker, R. A., Eick, S. G., & Wilks, A. R. (1995). Visualizing network data. *Visualization and Computer Graphics, IEEE Transactions on*, 1(1), 16-28.
- Bergstrom, T., & Karahalios, K. (2007). *Conversation Clock: Visualizing audio patterns in co-located groups*. Paper presented at the Proceedings of the 40th Hawaii International Conference on System Sciences - 2007.
- Berk, J. (2004). The State of Learning Analytics. *T&D*, 58(6), 34-39.
- Berlanga, A., & García Peñalvo, F. J. (2005). Learning technology specifications: semantic objects for adaptive learning environments. *Int. J. Learn. Technol.*, 1(4), 458-472. doi: 10.1504/ijlt.2005.007155
- Berlanga, A., & García Peñalvo, J. F. (2005). IMS LD reusable elements for adaptive learning designs. *Journal of Interactive Media in Education - JIME*, 1(12), 1-16. doi: 10.5334/2005-11

- Berlanga, A., & García Peñalvo, J. F. (2008). Learning Design in Adaptive Educational Hypermedia Systems. *j-jucs*, 14(22), 3627-3647.
- Berlanga, A. J., & García Peñalvo, F. J. (2004). *A Proposal to Define Adaptive Learning Designs*. Paper presented at the Proceedings of the Workshop on Applications of Semantic Web Technologies for Educational Adaptive Hypermedia (SW-EL 2004).
- Berlanga, A. J., & García Peñalvo, F. J. (2008). Learning Design in Adaptive Educational Hypermedia Systems. *Journal of Universal Computer Science, J. UCS*, 14(22), 3627-3647. doi: 10.3217/jucs-014-22-3627
- Berlanga, A. J., Peñalvo, F. G., & Sloep, P. B. (2010). Towards eLearning 2.0 University. *Interactive Learning Environments*, 18(3), 199-201. doi: 10.1080/10494820.2010.500498
- Berman, J. J. (2013). Chapter 4 - Introspection *Principles of Big Data* (pp. 49-61). Boston: Morgan Kaufmann.
- Berners-Lee, T., Cailliau, R., Groff, J.-F., & Pollermann, B. (1992). World-Wide Web: the information universe. *Internet Research*, 2(1), 52-58.
- Bertini, E., & Lalanne, D. (2009). *Surveying the complementary role of automatic data analysis and visualization in knowledge discovery*. Paper presented at the Proceedings of the ACM SIGKDD Workshop on Visual Analytics and Knowledge Discovery: Integrating Automated Analysis with Interactive Exploration, Paris, France.
- Bhardwaj, B. K., & Pal, S. (2012). Data Mining: A prediction for performance improvement using classification. *International Journal of Computer Science and Information Security*, abs/1203.3832(4), 51-56.
- Bichsel, J. (2012a). Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations. (*Research Report*), 31. Retrieved from <http://www.educause.edu/ecar>. website: <http://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf>
- Bichsel, J. (2012b). Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations. 31. Retrieved from <http://www.educause.edu/ecar>. website: <http://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf>
- Bielenberg, K., & Zacher, M. (2005). *Groups in Social Software: Utilizing Tagging to Integrate Individual Contexts for Social Navigation*. (Master of Science in Digital Media Master of Science in Digital Media). Retrieved from [http://groop.us/blog/Thesis\\_KBMZ.pdf](http://groop.us/blog/Thesis_KBMZ.pdf)
- Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief. In O. o. E. Technology (Ed.), <http://www.ed.gov/technology> (U.S. Department of Education ed.): U.S. Department of Education Office of Educational Technology.
- Bindé, J. (2005). Towards knowledge societies: UNESCO world report. In U. r. works (Ed.), (pp. 237). Paris: UNESCO.

- Blikstein, P. (2011). *Using learning analytics to assess students' behavior in open-ended programming tasks*. Paper presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge, Banff, Alberta, Canada.
- Bloom, B. S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13(6), 4-16. doi: 10.3102/0013189x013006004
- Boneu, J. M. (2007). Plataformas abiertas de e-learning para el soporte de contenidos educativos abiertos. *Revista de Universidad y Sociedad del Conocimiento (RUSC)*, 4(1), 36-47.
- Bongshin, L., Riche, N. H., Karlson, A. K., & Carpendale, S. (2010). SparkClouds: Visualizing Trends in Tag Clouds. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6), 1182-1189.
- Borland, D., & Taylor, R. M. (2007). Rainbow Color Map (Still) Considered Harmful. *Computer Graphics and Applications, IEEE*, 27(2), 14-17.
- Börner, K. (2012). *Visual analytics in support of education*. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada.
- Boyd, D., Hyun-Yeul, L., Ramage, D., & Donath, J. (2002). *Developing legible visualizations for online social spaces*. Paper presented at the System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on.
- Bratitsis, T., & Dimitracopoulou, A. (2007). *Collecting and Analyzing Interaction Data in Computer-Based Group Learning Discussions: An overview*. Paper presented at the Workshop at the 11th International Conference on User Modeling, Corfu, Greece.
- Bromham, L., & Oprandi, P. (2006). Evolution online: using a virtual learning environment to develop active learning in undergraduates. *Journal of Biological Education*, 41(1), 21-25.
- Brown, J. S., & Adler, P. R. (2008). Minds on Fire: Open Education, the Long Tail, and Learning 2.0. *EDUCAUSE Review*, 43(1), 16-32.
- Brown, M. (2011). Learning analytics: The Coming Third Wave: ELI Briefs, EDUCAUSE Learning Initiative (ELI).
- Brown, M. (2012). Learning Analytics: Moving from Concept to Practice. *ELI Briefs, EDUCAUSE Learning Initiative (ELI)*. Retrieved from <http://www.educause.edu/eli> website: <http://net.educause.edu/ir/library/pdf/ELIB1203.pdf>
- Bruner, J. (2000). *La educación, puerta de la cultura* (F. Díaz, Trans. 3 ed.). Madrid.
- Brusilovsky, P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction*, 11(1-2), 87-110. doi: 10.1023/a:1011143116306
- Brusilovsky, P., Baishya, D., Hosseini, R., Guerra, J., & Miner, L. (2013). *KnowledgeZoom for Java: A Concept-Based Exam Study Tool with a Zoomable Open Student Model*. Paper

- presented at the Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on.
- Buckingham Shum, S., & Deakin Crick, R. (2012). *Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics*. Paper presented at the 2nd International Conference on Learning Analytics & Knowledge, Vancouver, British Columbia, Canada.
- Buckingham Shum, S., & Ferguson, R. (2012). Social Learning Analytics. *Journal of Educational Technology & Society*, 15(3), 3–26.
- Buja, A., Cook, D., & Swayne, D. F. (1996). Interactive High-Dimensional Data Visualization. *Journal of Computational and Graphical Statistics*, 5(1), 78-99. doi: 10.2307/1390754
- Burby, J., & Atchison, S. (2007). *Actionable Web Analytics: Using Data to Make Smart Business Decisions*: SYBEX Inc.
- Buschmann, F., Meunier, R., Rohnert, H., Sommerlad, P., Stal, M., Sommerlad, P., & Stal, M. (1996). *Pattern-oriented software architecture, volume 1: A system of patterns*: John Wiley and Sons.
- Butcher, K. R., Bhushan, S., & Sumner, T. (2006). Multimedia displays for conceptual discovery: information seeking with strand maps. *Multimedia Systems*, 11(3), 236-248. doi: 10.1007/s00530-005-0004-y
- Butcher, K. R., & Kintsch, W. (2004). Learning with Diagrams: Effects on Inferences and the Integration of Information. In S. B. Heidelberg (Ed.), *Diagrammatic Representation and Inference* (Vol. 2980/2004, pp. 337-340).
- Butler, D. M., Almond, J. C., Bergeron, R. D., Brodlie, K. W., & Haber, R. B. (1993). *Visualization reference models*. Paper presented at the Proceedings of the 4th conference on Visualization '93, San Jose, California.
- Cacciamani, S., & Luciani, M. (2007). *Una knowledge building community nella formazione universitaria on line*: Polimetrica.
- Calvo, R. A., O'Rourke, S. T., Jones, J., Yacef, K., & Reimann, P. (2011). Collaborative Writing Support Tools on the Cloud. *Learning Technologies, IEEE Transactions on*, 4(1), 88-97. doi: 10.1109/TLT.2010.43
- Campbell, J. P. (2007). The Grand Challenge: Using Analytics to Predict Student Success: EDUCAUSE Midwest Regional Conference.
- Campbell, J. P., DeBlois, P. B., & Oblinger, D. (2007). Academic Analytics: A New Tool for a New Era. *EDUCAUSE*, 42(4), 40–57.
- Campbell, J. P., & Oblinger, D. (2007). Academic Analytics. *EDUCAUSE white paper*, 1-20. Retrieved from <http://www.educause.edu/> website: <http://net.educause.edu/ir/library/pdf/PUB6101.pdf>
- Canfora, G., Fasolino, A. R., Frattolillo, G., & Tramontana, P. (2008). A wrapping approach for migrating legacy system interactive functionalities to Service Oriented Architectures. *Journal of*

- Systems and Software*, 81(4), 463-480. doi: <http://dx.doi.org/10.1016/j.jss.2007.06.006>
- Card, S. K., & Mackinlay, J. (1997). *The structure of the information visualization design space*. Paper presented at the Information Visualization, 1997. Proceedings., IEEE Symposium on.
- Card, S. K., Mackinlay, J., & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*.
- Carlis, J. V., & Konstan, J. A. (1998). *Interactive visualization of serial periodic data*, New York, NY, USA.
- Carr, E. M., & Mazur-Stewart, M. (1988). The effects of the vocabulary overview guide on vocabulary comprehension and retention. *Journal of Literacy Research*, 20(1), 43-62.
- Casany Guerrero, M. J., Alier, M. F., Conde González, M. Á., & García-Peñalvo, F. J. (2009). *SOA initiatives for eLearning: A Moodle case*. Paper presented at the 23rd International Conference on Advanced Information Networking and Applications, AINA 2009, Bradford, United Kingdom.
- Casany Guerrero, M. J., Conde González, M. Á., Alier, M. F., & García-Peñalvo, F. J. (2009). *Applications of Service Oriented Architecture for the Integration of LMS and m-Learning Applications*. Paper presented at the WEBIST.
- Casany, M. J., Alier, M., Conde, M. Á., & García, F. J. (2009). *SOA Initiatives for eLearning: A Moodle Case*. Paper presented at the 23rd International Conference on Advanced Information Networking and Applications, AINA, Bradford, United Kingdom, May 26-29, 2009.
- Castaño Garrido, C. (2004). Utilización de mapas conceptuales para la construcción de modelos de conocimiento. *Primeras Noticias. Comunicación y Pedagogía*(No. 194), 75-78.
- Cesareni, D., & Martini, F. (2005). Costruire conoscenza in un forum universitario. *Rassegna di Psicologia*, 22, 89-112.
- Chaomei, C., Ibekwe-SanJuan, F., SanJuan, E., & Weaver, C. (2006). *Visual Analysis of Conflicting Opinions*. Paper presented at the Visual Analytics Science And Technology, 2006 IEEE Symposium On.
- Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *Journal: Int. J. of Technology Enhanced Learning*, 4(5/6), 318 - 331. doi: 10.1504/IJTEL.2012.051815
- 10.1504/12.51815
- Che-Cheng, L., & Chiung-Hui, C. (2013). *Correlation between Course Tracking Variables and Academic Performance in Blended Online Courses*. Paper presented at the Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on.
- Chen, C. (2006). *Information Visualization: Beyond the Horizon*: Springer-Verlag New York, Inc.
- Chen, Y., Yang, J., & Ribarsky, W. (2009). *Toward effective insight management in visual analytics systems*. Paper presented at the Visualization Symposium, 2009. PacificVis' 09. IEEE Pacific.

- Chi, E. H. (2000). *A Taxonomy of Visualization Techniques Using the Data State Reference Model*. Paper presented at the Proceedings of the IEEE Symposium on Information Visualization 2000, Salt Lake City, UT.
- Chuah, M. C., & Roth, S. F. (1996). *On the semantics of interactive visualizations*. Paper presented at the Information Visualization '96, Proceedings IEEE Symposium on.
- Chuang, J., Manning, C. D., & Heer, J. (2012). "Without the clutter of unimportant words": Descriptive keyphrases for text visualization. *ACM Trans. Comput.-Hum. Interact.*, 19(3), 1-29. doi: 10.1145/2362364.2362367
- Chung, W., Chen, H., Chaboya, L. G., O'Toole, C. D., & Atabakhsh, H. (2005). Evaluating event visualization: a usability study of COPLINK spatio-temporal visualizer. *International Journal of Human-Computer Studies*, 62(1), 127 - 157.
- Clow, D. (2012). *The learning analytics cycle: closing the loop effectively*. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada.
- Clow, D. (2013). *MOOCs and the funnel of participation*. Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium.
- Clow, D., & Makriyannis, E. (2011). *iSpot analysed: participatory learning and reputation*. Paper presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge, Banff, Alberta, Canada.
- CMSI, C. M. S. I. S. d. I. I. (2004). Plan de Acción. Túnez: Unión Internacional de Telecomunicaciones (UIT).
- Coll, C., & Monereo, C. (2008). *Psicología de la educación virtual: Morata*.
- Collins, C., Viegas, F. B., & Wattenberg, M. (2009). *Parallel Tag Clouds to explore and analyze faceted text corpora*. Paper presented at the Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on.
- Conde González, M. Á. (2007). *mLearning, de camino hacia el uLearning*. (Master en Sistemas Inteligentes Tesis de Máster), Universidad de Salamanca, Salamanca. Retrieved from <http://hdl.handle.net/10366/21829>
- Conde González, M. Á. (2012). *Personalización del aprendizaje: Framework de servicios para la integración de aplicaciones online en los sistemas de gestión del aprendizaje*. (Tesis Doctoral), Universidad de Salamanca, Salamanca, España.
- Conde González, M. Á., Casany Guerrero, M. J., Alier, M. F., & García Pealvo, J. F. (2009). *Back and forth: From the LMS to the mobile device. A SOA approach*. Paper presented at the IADIS International Conference Mobile Learning, Barcelona, Spain.
- Conde González, M. Á., García Pealvo, J. F., Casany Guerrero, M. J., & Alier, M. F. (2011, 10-12 July). *Merging learning management systems and personal learning environments*. Paper presented at the Proceedings of the The PLE Conference, Southampton, UK.

- Conde González, M. Á., García Pealvo, J. F., Gómez Aguilar, D. A., & Therón Sánchez, R. (2014, 22-25 Oct.). *Visual learning analytics techniques applied in software engineering subjects*. Paper presented at the 2014 IEEE Frontiers in Education Conference Proceedings Madrid, Spain.
- Conde González, M. Á., García Peñalvo, F. J., Casany Guerrero, M. J., & Alier Forment, M. (2009). *Adapting LMS architecture to the SOA an Architectural Approach*. Paper presented at the Internet and Web Applications and Services ICIW 2009, Venice/Mestre, Italy.
- Conde González, M. Á., García Peñalvo, F. J., Guerrero, M. J. C., & Forment, M. A. (2009). *Adapting LMS architecture to the SOA: an Architectural Approach*. Paper presented at the Internet and Web Applications and Services, 2009. ICIW'09. Fourth International Conference on.
- Conde González, M. Á., García-Peñalvo, F. J., Casany, M. J., & Forment, M. A. (2010). Open Integrated Personal Learning Environment: Towards a New Conception of the ICT-Based Learning Processes *Knowledge Management, Information Systems, E-Learning, and Sustainability Research* (pp. 115-124): Springer.
- Conde González, M. Á., García-Peñalvo, F. J., Casany, M. J., & Forment, M. A. (2013). Personal Learning Environments and the Integration with Learning Management Systems *Information Systems, E-learning, and Knowledge Management Research* (pp. 16-21): Springer.
- Conde González, M. Á., Gómez Aguilar, D. A., Del Pozo de Dios, A., & García Pealvo, J. F. (2011). Web services layer for Moodle 2.0: a new area of possibilities in web based learning. *Int. J. Technol. Enhanc. Learn.*, 3(3), 308-321. doi: 10.1504/ijtel.2011.040227
- Conde González, M. Á., Gómez Aguilar, D. A., del Pozo de Dios, A., Therón Sánchez, R., & García Peñalvo, F. J. (2010). *Moodle 2.0 Web Services Layer and Its New Application Contexts*. Paper presented at the 1st International Conference, TECH-EDUCATION 2010., Athens, Greece.
- Conde González, M. Á., González, J. C., & Moren, R. M. M. (2006). Arquitectura para un LMS basada en portlets: Claynet 2.0. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, 7(2), 213-234.
- Conole, G. i., & Alevizou, P. (2010). A literature review of the use of Web 2.0 tools in Higher Education. In H. E. Academy (Ed.), (pp. 111). York, UK.
- Cook, K., Earnshaw, R., & Stasko, J. (2007). Guest Editors' Introduction: Discovering the Unexpected. *Computer Graphics and Applications*, 27(5), 15-19. doi: 10.1109/MCG.2007.126
- Cooper, A. (2012). A Brief History of Analytics. *CETIS Analytics Series*:, 1(9). <http://publications.cetis.ac.uk/wp-content/uploads/2012/12/Analytics-Brief-History-Vol-1-No9.pdf>
- Cox, B., & Jantti, M. (2012). Discovering the Impact of Library Use and Student Performance. *EDUCAUSE Review*.

<http://www.educause.edu/ero/article/discovering-impact-library-use-and-student-performance>

- Crick, R. D., Broadfoot, P., & Claxton, G. (2004). Developing an Effective Lifelong Learning Inventory: the ELLI Project. *Assessment in Education: Principles, Policy & Practice*, 11(3), 247-272. doi: 10.1080/0969594042000304582
- Cuban, L. (2003). *New Technologies in Old Universities Oversold and Underused: Computers in the Classroom* (pp. 99-175). Cambridge: MA: Harvard UP.
- Daassi, C., Nigay, L., & Fauvet, M. C. (2006). A taxonomy of temporal data visualization techniques. *Revue Information-Interaction-Intelligence, Revue en Sciences du Traitement de l'Information (A journal in the Sciences of Information Engineering) Vol. 5, No 2*, 41-63
- Dachselt, R., & Weiland, M. (2006). *TimeZoom: a flexible detail and context timeline*, New York, NY, USA.
- Dagger, D., O'Connor, A., Lawless, S., Walsh, E., & Wade, V. P. (2007). Service-Oriented E-Learning Platforms: From Monolithic Systems to Flexible Services. *Internet Computing, IEEE*, 11(3), 28-35. doi: 10.1109/MIC.2007.70
- Davenport, T. H. (2006). COMPETING ON ANALYTICS. *Harvard Business Review*, 84(1), 98-107.
- Davenport, T. H., Harris, J., & Shapiro, J. (2010). Competing on Talent Analytics. *Harvard Business Review*, 88(10), 52-58.
- Davenport, T. H., Harris, J. G., & Morison, R. (2010). *Analytics at Work: Smarter Decisions, Better Results*: Harvard Business School Press.
- Davies, J., & Graff, M. (2005). Performance in e-learning: Online participation and student grades. *British Journal of Educational Technology*, 36(4), 657-663.
- Davis, H. C., Carr, L., Hey, J. M. N., Howard, Y., Millard, D., Morris, D., & White, S. (2010). Bootstrapping a Culture of Sharing to Facilitate Open Educational Resources. *Learning Technologies, IEEE Transactions on*, 3(2), 96-109. doi: 10.1109/TLT.2009.34
- Dawson, S., McWilliam, E., & Pei-Ling Tan, J. (2008). *Teaching smarter: How mining ICT data can inform and improve learning and teaching practice*. Paper presented at the Proceedings of the Australasian society for computers in learning in tertiary education.
- De Laat, M., Lally, V., Lipponen, L., & Simons, R.-J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for Social Network Analysis. *International Journal of Computer-Supported Collaborative Learning*, 2(1), 87-103. doi: 10.1007/s11412-007-9006-4
- De Liddo, A., Buckingham Shum, S., Quinto, I., Bachler, M., & Cannavacciuolo, L. (2011). *Discourse-centric learning analytics*. Paper presented at the LAK 2011: 1st International Conference on Learning Analytics & Knowledge, Banff, Alberta. <http://oro.open.ac.uk/25829/>

- Di Lucca, G. A., Gold, N., & Visaggio, G. (2008). Guest Editor's Introduction. 10th Conference on Software Maintenance and Reengineering. . *Journal of Systems and Software*, 81, 461-462. doi: <http://dx.doi.org/10.1016/j.jss.2007.06.001>
- Diaz, V., & Brown, M. (2012). Learning Analytics A Report on the ELI Focus Session. *ELI White Papers, EDUCAUSE Learning Initiative (ELI)*, 2. Retrieved from <http://www.educause.edu/eli> website: <http://www.educause.edu/library/resources/learning-analytics-report-eli-focus-session>
- Dicheva, D., & Dichev, C. (2006). TM4L: Creating and Browsing Educational Topic Maps. *British Journal of Educational Technology - BJET*, 37, 391-404.
- Dicheva, D., Dichev, C., & Wang, D. (2005). Visualizing topic maps for e-learning. *Advanced Learning Technologies, 2005. ICALT 2005. Fifth IEEE International Conference on*, 6, 950-951.
- Dix, A. (2009). Human-computer interaction *Encyclopedia of Database Systems* (pp. 1327-1331): Springer US.
- Dix, A., & Ellis, G. (1998). *Starting simple: adding value to static visualisation through simple interaction*. Paper presented at the Proceedings of the working conference on Advanced visual interfaces, L'Aquila, Italy.
- Donath, J. (2002). A semantic approach to visualizing online conversations. *Commun. ACM*, 45(4), 45-49
- Downes, S. (2005). E-learning 2.0. *Elearn magazine*, 2005(10), 1--. doi: 10.1145/1104966.1104968
- Drachsler, H., Bogers, T., Vuorikari, R., Verbert, K., Duval, E., Manouselis, N., . . . Wolpers, M. (2010). Issues and considerations regarding sharable data sets for recommender systems in technology enhanced learning. *Procedia Computer Science*, 1(2), 2849-2858. doi: <http://dx.doi.org/10.1016/j.procs.2010.08.010>
- Dradilova, P., Martinovic, J., Slaninova, K., & Snasel, V. (2008). *Analysis of Relations in eLearning*. Paper presented at the Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT '08. IEEE/WIC/ACM International Conference on.
- Drigas, A. S., & Vrettaros, J. (2006). An Intelligent Search Engine Assessing Learning Material to Improve Learning Procedures. *Information Technology Based Higher Education and Training, 2006. ITHET '06. 7th International Conference on*, 875-883.
- Duart Montoliu, J. M., Gil, M., Pujol, M., & Castaño, J. (2008). *La universidad en la sociedad red : usos de Internet en educación superior* Barcelona, España : Ariel.
- Dubinko, M., Kumar, R., Magnani, J., Novak, J., Raghavan, P., & Tomkins, A. (2007). Visualizing tags over time. *ACM Trans. Web*, 1(2), 7. doi: 10.1145/1255438.1255439
- Dyckhoff, A. L., Lukarov, V., Muslim, A., Chatti, M. A., & Schroeder, U. (2013). *Supporting action research with learning analytics*. Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium.

- Eades, P. (1984). A heuristics for graph drawing. *Congressus numerantium*, 42, 146-160.
- Earl, L. M., & Katz, S. (2006). *Leading schools in a data rich world : harnessing data for school improvement*. Heatherton, Vic: Hawker Brownlow Education.
- Eckersen, W. W. (2007). Predictive Analytics: Extending the Value of Your Data Warehousing Investment *First Quarter: TDWI Best Practices Report*.
- Edwards, C. (2010). *An introduction to the Effective Lifelong Learning Inventory*. Paper presented at the Milton Keynes: Open University.
- Ehren, M. C. M., & Visscher, A. J. (2008). the relationships between school inspections, school characteristics and school improvement. *British Journal of Educational Studies*, 56(2), 205-227. doi: 10.1111/j.1467-8527.2008.00400.x
- Elias, T. (2011). Learning Analytics: Definitions, Processes and Potential.
- Ewy, C. A. (2003). *Teaching with Visual Frameworks: Focused Learning and Achievement Through Instructional Graphics Co-Created by Students and Teachers*: Corwin Press, Inc.
- Farajollahi, M., & Moenikia, M. (2011). *The effect of computer-based learning on distance learners self regulated learning strategies*.
- Feipeng, S., Yanyan, L., & Zhiqiang, Z. (2013). *A Tool for Visualizing Topic Evolution in Large Text Collections*. Paper presented at the Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on.
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 304-317.
- Ferguson, R., & Buckingham Shum, S. (2011). *Learning analytics to identify exploratory dialogue within synchronous text chat*. Paper presented at the LAK 2011: 1st International Conference on Learning Analytics & Knowledge, Banff, Alberta. <http://oro.open.ac.uk/28955/>
- Ferguson, R., & Buckingham Shum, S. (2012a). Social learning analytics. *Journal of Educational Technology and Society*, 15(3). <http://oro.open.ac.uk/34092/>
- Ferguson, R., & Buckingham Shum, S. (2012b). *Social learning analytics: five approaches*. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada.
- Ferrance, E. (2000). Action Research. In B. University (Ed.), *Themes in Education*. NY, USA: LAB: Northeast and Islands Regional Educational.
- Ferraro, M., Álvarez, H., & Peñalvo, F. (2004). Adaptive Educational Hypermedia Proposal Based on Learning Styles and Quality Evaluation. In P. E. De Bra & W. Nejdl (Eds.), *Adaptive Hypermedia and Adaptive Web-Based Systems* (Vol. 3137, pp. 316-319): Springer Berlin Heidelberg.

- Fidalgo Blanco, Á., Sein-Echaluce Lacleta, M. L., Borrás Gené, O., & García Peñalvo, F. J. (2014). Educación en abierto: Integración de un MOOC con una asignatura académica. *Education in the Knowledge Society (EKS)*, 15(3), 233-255.
- Foley, J., Card, S., Ebert, D., MacEachren, A., & Ribarsky, B. (2006). *Visual Analytics Education*. Paper presented at the Visual Analytics Science And Technology, 2006 IEEE Symposium On.
- Foley, J., van Dam, A., Feiner, S., & Hughes, J. (1995). *Computer Graphics: Principles and Practice in C*: Pearson.
- Fournier, H., Kop, R., & Sitlia, H. (2011). *The value of learning analytics to networked learning on a personal learning environment*. Paper presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge, Banff, Alberta, Canada.
- Fritz, J. (2009). *Using Course Activity Data to Raise Awareness of Underperforming College Students*. Paper presented at the World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2009, Vancouver, Canada. <http://www.edutlib.org/p/32850>
- Furnas, G. W. (1986). *Generalized fisheye views* (Vol. 17): ACM.
- Gallagher, K., Hatch, A., & Munro, M. (2008). Software Architecture Visualization: An Evaluation Framework and Its Application. *Software Engineering, IEEE Transactions on*, 34(2), 260-270. doi: 10.1109/TSE.2007.70757
- Gallavan, N. P., & Kottler, E. (2007). Eight Types of Graphic Organizers for Empowering Social Studies Students and Teachers. *The Social Studies*, 98(3), 117-128. doi: 10.3200/TSSS.98.3.117-128
- Gambette, P., & Véronis, J. (2010). Visualising a Text with a Tree Cloud. In H. Locarek-Junge & C. Weihs (Eds.), *Classification as a Tool for Research* (pp. 561-569): Springer Berlin Heidelberg.
- Ganoe, C. H., Somervell, J. P., Neale, D. C., Isenhour, P. L., Carroll, J. M., Rosson, M. B., & McCrickard, D. S. (2003). *Classroom BRIDGE: using collaborative public and desktop timelines to support activity awareness*. New York, NY, USA.
- García, E., Romero, C., Ventura, S., de Castro, C., & Calders, T. (2010). Association Rule Mining in Learning Management Systems *Handbook of Educational Data Mining* (pp. 93-106): CRC Press.
- García, J., Peñalvo, F. J. G., Therón, R., & de Pablos, P. O. (2011). Usability Evaluation of a Visual Modelling Tool for OWL Ontologies. *J. UCS*, 17(9), 1299-1313.
- García Navarro, J. F. (2012). *Análítica visual aplicada a la Ingeniería de Ontologías*. (Doctorado de Informática y Automática), Universidad de Salamanca, Salamanca. Retrieved from <http://hdl.handle.net/10366/116140>
- García Penalvo, F. J., Conde González, M. Á., Bravo Martín, S., Gómez Aguilar, D. A., & Therón Sánchez, R. (2011a). Visual Analysis of a Moodle-based Object Oriented Programming Course. *IJCA Proceedings on Design and Evaluation of Digital Content for Education (DEDCE)*, 1(0), 8-14.

- García Penalvo, F. J., Conde González, M. Á., Bravo Martín, S., Gómez Aguilar, D. A., & Therón Sánchez, R. (2011b). Visual Analysis of a Moodle-based Object Oriented Programming Course *International Journal of Computers Applications Proceedings on Design and Evaluation of Digital Content for Education (DEDCE)* (Vol. 2, pp. 13-19). USA: Foundation of Computer Science.
- García Peñalvo, F. J. (2005). Estado Actual de los Sistemas E-Learning. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, 6(2). <http://www.redalyc.org/articulo.oa>
- García Peñalvo, F. J. (2008). Docencia. In L. M. P. c. Jaime Laviña Orueta (coord.) (Ed.), *Libro Blanco de la Universidad Digital* (pp. 29-62). Madrid, España: Editorial Ariel.
- García Peñalvo, F. J. (2008). *Preface of Advances in E-Learning: Experiences and Methodologies*. Hershey, PA, USA: Information Science Reference.
- García Peñalvo, F. J. (2011). La universidad de la próxima década: la universidad digital.
- García Peñalvo, F. J., Bravo Martín, S., & Conde González, M. Á. (2008). Ingeniería del Software OCW-USAL Web.
- García Peñalvo, F. J., & García Carrasco, J. (2002). Los espacios virtuales educativos en el ámbito de Internet: Un refuerzo a la formación tradicional. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, 3.
- García Peñalvo, F. J., García de Figuerola, C., & Merlo, J. A. (2010). Open knowledge management in higher education. *Online Information Review*, 34(4), 517-519.
- García Peñalvo, F. J., & Seoane Pardo, A. M. (2015). Una revisión actualizada del concepto de eLearning. *Education in the Knowledge Society, Décimo Aniversario*, 16(1).
- García Peñalvo, J. F., García de Figuerola, C., & Merlo, J. A. (2010). Open knowledge. Challenges and facts. *Online Information Review*, 34(4), 520-539. doi: 10.1108/14684521011072963
- García-Peñalvo, F. J. (2005). Estado actual de los sistemas e-learning. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, 6(2). Retrieved from <http://www.redalyc.org/> website: <http://www.redalyc.org/articulo.oa?id=201021055001>
- García-Peñalvo, F. J. (2008). *Advances in E-Learning: Experiences and Methodologies* (Vol. Hershey, PA, USA): IGI Global.
- García-Peñalvo, F. J., García, J., & Therón, R. (2011). Analysis of the owl ontologies: A survey. *Scientific Research and Essays*, 6(20), 4318-4329.
- Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, 7(2), 95-105. doi: <http://dx.doi.org/10.1016/j.iheduc.2004.02.001>
- Garrison, M., & Monson, M. (2012). The Oregon DATA Project: Building a Culture of Data Literacy. *EDUCAUSE Review*. <http://www.educause.edu/ero/article/oregon-data-project-building-culture-data-literacy>

- Gee, J. (1997). Thinking, Learning and Reading: The situated sociocultural mind. In D. Kirshner & J. A. Whitson (Eds.), *Situated Cognition: Social, semiotic, and psychological perspectives*. (pp. 235-260). Hillsdale, NJ:: L. Erlbaum.
- Gee, J. P. (1992). *The social mind: language, ideology, and social practice*. New York: Bergin & Garvey.
- Gershon, N., Eick, S. G., & Card, S. (1998). Information visualization. *interactions*, 5(2), 9-15. doi: 10.1145/274430.274432
- Ghanbari, M. (2007). Visualization Overview. *System Theory, 2007. SSST '07. Thirty-Ninth Southeastern Symposium on*, 6, 115-119.
- Gibbs, W. J., Olexa, V., & Bernas, R. S. (2006). A Visualization Tool for Managing and Studying Online Communications. *Journal of Educational Technology & Society*, 9, 232-243.
- Gil, R., Sancristobal, E., Martín, S., Díaz, G., Colmenar, A., Llamas, M., . . . Castro, M. (2009, June ). *S-Learning: New Web Services in E-Learning Platforms*. Paper presented at the American Society for Engineering Education - ASSE Annual Conference, Austin, TX.
- Giovannella, C., Scaccia, F., & Popescu, E. (2013). *A PCA Study of Student Performance Indicators in a Web 2.0-Based Learning Environment*. Paper presented at the Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on.
- Goldstein, P. J., & Katz, R. N. (2005). Academic Analytics: The Use of Management Information and Technology in Higher Education—Key Findings. *ECAR publication*. Retrieved from <http://www.educause.edu/> website: [http://net.educause.edu/ir/library/pdf/ecar\\_so/ers/ers0508/EKF0508.pdf](http://net.educause.edu/ir/library/pdf/ecar_so/ers/ers0508/EKF0508.pdf)
- Goldstein, P. J., & Katz, R. N. (2005). Academic analytics: The uses of management information and technology in higher education *Tech. Rep.* (Vol. 8): EDUCAUSE Center for Applied Research (2005).
- Gómez Aguilar, D. A., Conde González, M. Á., Therón, R., & García Peñalvo, F. J. (2010a). *Modelo basado en Servicios Web para la Visualización de Moodle*. Paper presented at the En las Actas del XI Congreso Internacional de Interacción Persona-Ordenador – Interacción 2010. Celebrado dentro del III Congreso Español de Informática, Universidad Politécnica de Valencia, España.
- Gómez Aguilar, D. A., Conde González, M. Á., Therón, R., & García Peñalvo, F. J. (2010b). Retrieval information model for Moodle data visualization. *Advanced Learning Technologies 2010. ICALT 2010. 10th IEEE International Conference on, Product Number E4055*, 526-527. doi: 10.1109/ICALT.2010.150
- Gómez Aguilar, D. A., Conde González, M. Á., Therón Sánchez, R., & García Peñalvo, F. J. (2011). *Improving Moodle-based eLearning through visual analysis, a case study*. Paper presented at the The 2011 CAL conference, Learning Futures: Education, Technology, Sustainability. , Manchester, UK

- Gómez Aguilar, D. A., Conde González, M. Á., Therón Sánchez, R., & García Peñalvo, F. J. (2011, 2-6 Sep.). *Supporting Moodle-based lesson of Software Engineering through visual analysis*. Paper presented at the XII Congreso Internacional de Interacción Persona-Ordenador – Interacción 2011, Lisboa, Portugal.
- Gómez Aguilar, D. A., Conde-Gonzalez, M. Á., García-Peñalvo, F. J., & Therón, R. (2011). *Reaveling the Evolution of Semantic Content through Visual Analysis*. Paper presented at the Advanced Learning Technologies (ICALT), 2011 11th IEEE International Conference on.
- Gómez Aguilar, D. A., García Peñalvo, F. J., & Therón, R. (2013a). *Evaluación visual de las relaciones entre participación de los estudiantes y sus resultados en entornos de e-learning*. Paper presented at the XV Simposio Internacional de Tecnologías de la Información y las Comunicaciones en la Educación Conjuntamente con el Congreso Español de Informática (CEDI 2013), Madrid, España.  
<http://repositorio.grial.eu/handle/grial/270>
- Gómez Aguilar, D. A., García Peñalvo, F. J., & Therón, R. (2013b). *Tap into visual analysis of the customization of grouping of activities in eLearning*. Paper presented at the Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality, Salamanca, Spain.
- Gómez Aguilar, D. A., García Peñalvo, F. J., & Therón, R. (2014). *Analítica visual en e-learning*. *El Profesional de la Información*, 23(3), 236-245.
- Gómez Aguilar, D. A., García Peñalvo, F. J., & Therón Sánchez, R. (2014). *Visual analytical model for educational data* 9th Iberian Conference on Information Systems and Technologies (CISTI) (Vol. CFP1403K-POD pp. 1-6): Institute of Electrical and Electronics Engineers. IEEE.
- Gómez Aguilar, D. A., García Peñalvo, F. J., & Therón Sánchez, R. (2014, 18-21 Jun). *Modelo de Analítica Visual para Datos Educativos*. Paper presented at the En Sistemas y Tecnologías de Información. Actas de la 9ª Conferencia Ibérica de Sistemas y Tecnologías de Información – CISTI 2014, Barcelona, España.
- Gómez Aguilar, D. A., Suárez Guerrero, C., Therón Sánchez, R., & García Peñalvo, F. J. (2010). *Visual Analytics to Support e-Learning*. In M. B. Rosson (Ed.), *Advances in Learning Processes* (pp. 207-228). Croatia: In-Tech.
- Gómez Aguilar, D. A., Theron, R., & García Peñalvo, F. J. (2008). *Understanding Educational Relationships in Moodle with ViMoodle*. Paper presented at the Advanced Learning Technologies, 2008. ICALT '08. Eighth IEEE International Conference on.
- Gómez Aguilar, D. A., Therón, R., & García-Peñalvo, F. J. (2009). *Semantic spiral timelines used as support for e-learning*. *Jornal of Universal Computer Science (j-jucs)*, 15(7), 1526-1545.
- Gómez Aguilar, D. A., Therón, R., & García-Peñalvo, F. J. (2013). *Reveal the Relationships among Students Participation and*

- Their Outcomes on E-Learning Environments: Case Study.* Paper presented at the Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on, Beijing.
- Gómez Aguilar, D. A., Therón Sánchez, R., & García Peñalvo, F. J. (2008). Visual Analytics for Virtual Education Platforms. In A. López Eire, F. J. García Peñalvo, A. M. Seoane Pardo, & E. M. Morales Morgado (Eds.), *Post-proceedings of the International Conference on Technology, Training and Communication eUniverSALearning 2007*. . Salamanca, Spain: CEUR Workshop Proceedings.
- Gómez Aguilar, D. A., Therón Sánchez, R., & García Peñalvo, F. J. (2013). *Reveal the relationships among student's participation and its outcomes on eLearning environments: Case study.* Paper presented at the 13th IEEE International Conference on Advanced Learning Technologies, 2013. ICALT 2013. , Beijing.
- Gómez Aguilar, D. A., Therón Sánchez, R., & García Peñalvo, F. J. (2007a). ViMoodle: Visualizando Moodle. In LOGO (Ed.), *I Congreso Internacional de Tecnología, Formación y Comunicación (EuniverSALearning'07)*. . Salamanca: Asociación Española de Estudios sobre Lengua, Pensamiento y Cultura Clásicas.
- Gómez Aguilar, D. A., Therón Sánchez, R., & García Peñalvo, F. J. (2007b). *Visual analytics, for virtual education platforms.* Paper presented at the I Congreso Internacional de Tecnología, Formación y Comunicación.
- Gómez Aguilar, D. A., Therón Sánchez, R., & García Peñalvo, F. J. (2008). *Semantic Spiral Timeline como apoyo al e-learning.* Paper presented at the Informática Educativa, 2008. SIIE 2008. El X Simposio Internacional de.
- Gomez-Aguilar, D., Conde-Gonzalez, M., Theron, R., & Garcia-Peñalvo, F. (2011). Supporting Moodle-Based Lesson through Visual Analysis. In P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, & M. Winckler (Eds.), *Human-Computer Interaction - INTERACT 2011* (Vol. 6949, pp. 604-607): Springer Berlin Heidelberg.
- Gómez-Aguilar, D. A., Hernández-García, Á., García-Peñalvo, F. J., & Therón, R. (2015). Tap into visual analysis of customization of grouping of activities in eLearning. *Computers in Human Behavior*, 47(0), 60-67. doi: 10.1016/j.chb.2014.11.001
- Gómez-Aguilar, D. A., Therón Sánchez, R., & García Peñalvo, F. J. (2008). Understanding educational relationships in Moodle with ViMoodle. *Advanced Learning Technologies, 2008. ICALT 2008. 8th IEEE International Conference on*, 6, 954-956 doi: DOI 10.1109/ICALT.2008.276
- Gotz, D., & Zhou, M. X. (2008). *Characterizing users' visual analytic activity for insight provenance.* Paper presented at the Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on.
- Govaerts, S., Verbert, K., Duval, E., & Pardo, A. (2012). *The student activity meter for awareness and self-reflection.* Paper presented at the CHI '12 Extended Abstracts on Human Factors in Computing Systems, Austin, Texas, USA.

- Govaerts, S., Verbert, K., Klerkx, J., & Duval, E. (2010). Visualizing Activities for Self-reflection and Awareness. In X. Luo, M. Spaniol, L. Wang, Q. Li, W. Nejdl, & W. Zhang (Eds.), *Advances in Web-Based Learning – ICWL 2010* (Vol. 6483, pp. 91-100): Springer Berlin Heidelberg.
- Govindasamy, T. (2001). Successful implementation of e-Learning: Pedagogical considerations. *The Internet and Higher Education*, 4(3-4), 287-299. doi: [http://dx.doi.org/10.1016/S1096-7516\(01\)00071-9](http://dx.doi.org/10.1016/S1096-7516(01)00071-9)
- Graf, S., & List, B. (2005). *An evaluation of open source e-learning platforms stressing adaptation issues*. Paper presented at the Advanced Learning Technologies, 2005. ICALT 2005. Fifth IEEE International Conference on.
- Graham, C. (2005). *Blended Learning Systems: Definition, Current Trends, and Future Directions*. Paper presented at the In. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.123.9357>
- Graham, M., & Kennedy, J. (2003). *Using curves to enhance parallel coordinate visualisations*. Paper presented at the Information Visualization, 2003. IV 2003. Proceedings. Seventh International Conference on.
- Granda, J. C., Uria, C., Garcia, D. F., Suarez, F. J., & Gonzalez, F. (2008). *Design Issues in Remote Visualization of Information in Interactive Multimedia E Learning Systems*.
- Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6), 1360-1380. doi: citeulike-article-id:99857 doi: 10.2307/2776392
- Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Journal of Educational Technology & Society*, 15(3), 42-57.
- Gresh, D. L., Rogowitz, B. E., Winslow, R. L., Scollan, D. F., & Yung, C. K. (2000). *WEAVE: a system for visually linking 3-D and statistical visualizations applied to cardiac simulation and measurement data*. Paper presented at the Visualization 2000. Proceedings.
- Gretarsson, B., O'Donovan, J., Bostandjiev, S., Höllerer, T., Asuncion, A., Newman, D., & Smyth, P. (2012). TopicNets: Visual Analysis of Large Text Corpora with Topic Modeling. *ACM Trans. Intell. Syst. Technol.*, 3(2), 1-26. doi: 10.1145/2089094.2089099
- Gruber, T. (2008). Collective knowledge systems: Where the Social Web meets the Semantic Web. *Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1), 4-13. doi: <http://dx.doi.org/10.1016/j.websem.2007.11.011>
- Guardia, L., Mas, X., & Girona, C. (2005). *Los materiales de aprendizaje en contextos educativos virtuales. Pautas para el diseño tecnopedagógico* (A. Sangrà Ed.). Barcelona: Editorial UOC.
- Guo, W., & Chen, D. (2006). Semantic Approach for e-learning System. *Computer and Computational Sciences, 2006. IMSCCS '06. First International Multi-Symposiums on*, 2, 442-446.

- Gutwin, C., & Greenberg, S. (2002). A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)*, 11(3-4), 411-446.
- Halvey, M. J., & Keane, M. T. (2007). *An assessment of tag presentation techniques*. Paper presented at the Proceedings of the 16th international conference on World Wide Web, Banff, Alberta, Canada.
- Hamilton, L., Halverson, R., Jackson, S. S., Mandinach, E. B., Supovitz, J. A., & Wayman, J. C. (2009). Using student achievement data to support instructional decision making (U. S. D. o. Education., Trans.). Washington, DC.: National Center for Education Evaluation and Regional Assistance, (NCEE 2009-4067).
- Hammond, T., Hannay, T., Lund, B., & Scott, J. (2005). Social Bookmarking Tools (I): A General Review. *D-Lib Magazine*, 11(4). doi: 10.1045/april2005-hammond
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*: Morgan Kaufmann Publishers Inc.
- Hardless, C., & Nulden, U. (1999). *Visualizing Learning Activities to Support Tutors*. Paper presented at the CHI EA '99, Pittsburgh, Pensilvania, USA.  
<http://www.informatik.gu.se/~hardless/publications/chi99.pdf>
- Harnad, S. (2005). The Implementation of the Berlin Declaration on Open Access. *D-Lib Magazine*, 11(3).  
<http://www.dlib.org/dlib/march05/harnad/03harnad.html>  
doi:10.1045/march2005-harnad
- Hartley, D. E. (2004). A Love-Hate Thing. *T&D*, 58(6), 20-22.
- Hassan-Montero, Y., & Herrero-Solana, V. (2006). *Improving Tag-Clouds as Visual Information Retrieval Interfaces*. Paper presented at the I International Conference on Multidisciplinary Information Sciences and Technologies, Mérida, Spain.  
<http://www.nosolousabilidad.com/hassan/visualizious/>
- Hattie, J. C. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London & New York: Routledge, Taylor& Francis Group.
- Hauser, H., Ledermann, F., & Doleisch, H. (2002). *Angular brushing of extended parallel coordinates*. Paper presented at the Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on.
- Havre, S., Hetzler, B., & Nowell, L. (2000). ThemeRiver: visualizing theme changes over time. *Information Visualization, 2000. InfoVis 2000. IEEE Symposium on*, 115-123
- Havre, S., Hetzler, E., Whitney, P., & Nowell, L. (2002). ThemeRiver: Visualizing Thematic Changes in Large Document Collections. *IEEE Transactions on Visualization and Computer Graphics*, 8(1), 9-20
- Haythornthwaite, C., & de Laat, M. (2010). Social Networks and Learning Networks: Using social network perspectives to

- understand social learning. In L. Dirckinck-Holmfeld, V. Hodgson, C. Jones, M. de Laat, & T. Ryberg (Eds.), *Proceedings of the 7th International Conference on Networked Learning* (pp. 183-190).
- Haythornthwaite, C., de Laat, M., & Dawson, S. (2013). Introduction to the Special Issue on Learning Analytics. *American Behavioral Scientist*, *57*(10), 1371-1379. doi: 10.1177/0002764213498850
- Heer, J., & Shneiderman, B. (2012). Interactive dynamics for visual analysis. *Commun. ACM*, *55*(4), 45-54. doi: 10.1145/2133806.2133821
- Hemerly, J. (2013). Public Policy Considerations for Data-Driven Innovation. *Computer*, *46*(6), 25-31. doi: 10.1109/MC.2013.186
- Henze, N., Dolog, P., & Nejd, W. (2004). Reasoning and ontologies for personalized e-learning in the semantic web. *Journal of Educational Technology & Society*, *7*(4), 82-97. doi: citeulike-article-id:6013130
- Herman, I., Melancon, G., & Marshall, M. S. (2000). Graph visualization and navigation in information visualization: A survey. *Visualization and Computer Graphics, IEEE Transactions on*, *6*(1), 24-43.
- Hetzler, E., Harris, W. M., Havre, S., & Whitney, P. (1998). Visualizing the full spectrum of document relationships. *ADVANCES IN KNOWLEDGE ORGANIZATION*, *6*, 167-174.
- Heuer, R. J. (2001). *Psychology of intelligence analysis*. Washington, D.C.: Center for the Study of Intelligence, Central Intelligence Agency.
- Hopkins, M. S. (2010). Are You Ready to Reengineer Your Decision Making? *Magazine: Fall 2010*, *52*(1), 1-7. <http://sloanreview.mit.edu/the-magazine/2010-fall/52102/are-you-ready-to-reengineer-your-decision-making/>
- Hou, H.-T., & Wu, S.-Y. (2011). Analyzing the social knowledge construction behavioral patterns of an online synchronous collaborative discussion instructional activity using an instant messaging tool: A case study. *Computers & Education*, *57*(2), 1459-1468. doi: <http://dx.doi.org/10.1016/j.compedu.2011.02.012>
- Hsu, Y.-c. (2007). The Effects of Visual Versus Verbal Metaphors on Novice and Expert Learners' Performance. In J. Jacko (Ed.), *Human-Computer Interaction. HCI Applications and Services* (Vol. 4553, pp. 264-269): Springer Berlin Heidelberg.
- Huang, M., & Nguyen, Q. (2008). Large graph visualization by hierarchical clustering. *Journal of Software*, *19*(8), 1933-1946.
- Hwang, A., & Arbaugh, J. B. (2009). Seeking feedback in blended learning: competitive versus cooperative student attitudes and their links to learning outcome. *Journal of Computer Assisted Learning*, *25*(3), 280-293. doi: 10.1111/j.1365-2729.2009.00311.x
- IBM. (2009). Predictive Analytics: Defined. Retrieved 01 Sep, 2013, from <http://www.spss.com.hk/corpinfo/predictive.htm>.

- IBM. (2010). Real World Predictive Analytics: Putting Analysis into Action for Visible Results. In I. Software (Ed.), *Business Analytics*.
- Inselberg, A. (1985). The plane with parallel coordinates. *The Visual Computer*, 1(2), 69-91.
- Inselberg, A., & Dimsdale, B. (1990). *Parallel coordinates: a tool for visualizing multi-dimensional geometry*. Paper presented at the Visualization, 1990. Visualization '90., Proceedings of the First IEEE Conference on.
- Inspiration Software, I. (2003, Nov. 2003). Executive Summary Graphic Organizers: A Review of Scientifically Based Research. Retrieved Sep., 2014, from <http://www.inspiration.com/vlearning/research/index.cfm>
- Instituto de Estadística, U. (2010). Trends in Tertiary Education: Sub-Saharan Africa (Vol. 10): UIS Fact Sheet.
- Jeff A. Estefan, Propulsion, J., Ken Laskey, M., Francis G. McCabe, & Danny Thornton, d. (2006). Reference Architecture for Service Oriented Architecture 1.0. Retrieved from <http://www.OASIS-open.org> website: <http://www.oasis-open.org/committees/download.php/19679/soa-rm-cs.pdf>
- Johnson, L., Adams, S., & Cummins, M. (2012). The NMC Horizon Report: 2012 Higher Education Edition. Austin, Texas: NMC, New Media Consortium.
- Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2011). Horizon Report: 2011 Higher Education Edition. In T. N. M. C. (NMC) (Ed.), (Vol. ISBN 978-0-9828290-5-9). Austin, Texas: NMC, New Media Consortium.
- Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*: Harvard University Press.
- Kaffash, H. R., Kargiban, Z. A., Kargiban, S. A., & Ramezani, M. T. (2010). A Close Look in to Role of ICT in Education. *Online Submission*, 3(2), 63-82.
- Karahalios, K. G., & Bergstrom, T. (2009). Social mirrors as social signals: transforming audio into gaphics. *IEEE Comput. Graph. Appl.*, 29(5), 22-33. doi: 10.1109/mcg.2009.85
- Karahalios, K. G., & Viégas, F. B. (2006). *Social visualization: exploring text, audio, and video interaction*. Paper presented at the CHI'06 Extended Abstracts on Human Factors in Computing Systems.
- Karunananda, A. S. (2006). A theoretical-based approach to e-Learning. *Industrial and Information Systems, First International Conference on*, 127-132.
- Kaser, O., & Lemire, D. (2007). Tag-Cloud Drawing: Algorithms for Cloud Visualization. *CoRR Tagging and Metadata for Social Information Organization Workshop*, abs/cs/0703109.
- Kaufmann, M., & Wagner, D. (2001). *Drawing Graphs: Methods and Models (Lecture Notes in Computer Science)*: Springer.
- Kaushik, A. (2007). *Web Analytics: An Hour a Day*: Sybex.
- Ke, F., & Kwak, D. (2013). Online learning across ethnicity and age: A study on learning interaction participation, perception, and

- learning satisfaction. *Computers & Education*, 61(0), 43-51. doi: <http://dx.doi.org/10.1016/j.compedu.2012.09.003>
- Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual Analytics: Definition, Process, and Challenges. In A. Kerren, J. Stasko, J.-D. Fekete, & C. North (Eds.), *Information Visualization* (Vol. 4950, pp. 154-175): Springer Berlin Heidelberg.
- Keim, D., Kohlhammer, J. r., Ellis, G., & Mansmann, F. (2010). *Mastering the Information Age Solving Problems with Visual Analytics*. Goslar, Germany: Eurographics Association.
- Keim, D., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008). Outlook for Visual Analytics Research Funding. In S. Simoff, M. H. Boehlen, & A. Mazeika (Eds.), *Visual Data Mining: Theory, Techniques and Tools for Visual Analytics* (pp. 76-90). Berlin, Heidelberg: Springer-Verlag.
- Keim, D. A. (2001). Visual exploration of large data sets. *Commun. ACM*, 44(8), 38-44. doi: 10.1145/381641.381656
- Keim, D. A. (2002). Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1), 1-8. doi: 10.1109/2945.981847
- Keim, D. A., Kohlhammer, J., Ellis, G., & Mansmann, F. (2010). *Mastering The Information Age-Solving Problems with Visual Analytics*: Florian Mansmann.
- Keim, D. A., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006). *Challenges in Visual Data Analysis*. Paper presented at the Tenth International Conference on Information Visualization. IV' 06., London, England.
- Keim, D. A., Mansmann, F., & Thomas, J. (2010). Visual analytics: how much visualization and how much analytics? *SIGKDD Explor. Newsl.*, 11(2), 5-8. doi: 10.1145/1809400.1809403
- Khan, B. (2000). A framework for web-based learning. *TechTrends*, 44(3), 51-51. doi: 10.1007/BF02778228
- Khan, B. H., & Ally, M. (2015). *International Handbook of E-Learning Volume 1: Theoretical Perspectives and Research*: Routledge.
- Khan, M., & Shah Khan, S. (2011). Data and Information Visualization Methods, and Interactive Mechanisms: A Survey. *International Journal of Computer Applications*, 34(1), 1-14.
- Kim, J. (2013). Influence of group size on students' participation in online discussion forums. *Computers & Education*, 62(0), 123-129. doi: <http://dx.doi.org/10.1016/j.compedu.2012.10.025>
- Koike, Y., Sugiura, A., & Koseki, Y. (1997). *TimeSlider: an interface to specify time point*. New York, NY, USA.
- Koper, R. (2006). Current research in learning design. *Educational Technology & Society*, 9(1), 13-22.
- Kowalski, G., & Maybury, M. T. (2000). *Information Storage and Retrieval Systems: Theory and Implementation*: Kluwer Academic Publishers.
- Krishnan, K. (2013). *Data Warehousing in the Age of Big Data*: Morgan Kaufmann Publishers Inc.

- Kruger-Ross, M. J., & Waters, R. D. (2013). Predicting online learning success: Applying the situational theory of publics to the virtual classroom. *Computers & Education*, *61*(0), 176-184. doi: <http://dx.doi.org/10.1016/j.compedu.2012.09.015>
- Kumar, G., & Garland, M. (2006). Visual Exploration of Complex Time-Varying Graphs. *IEEE Transactions on Visualization and Computer Graphics*, *12*(5), 805--812.
- Laal, M. (2011). Knowledge management in higher education. *Procedia Computer Science*, *3*(0), 544-549. doi: <http://dx.doi.org/10.1016/j.procs.2010.12.090>
- LAK. (2011). *Definition of Learning analytics from the 1st International Conference on Learning Analytics and Knowledge*. Paper presented at the 1st International Conference Learning Analytics & Knowledge, Banff, Alberta, Canada.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2-3), 259-284. doi: 10.1080/01638539809545028
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Latent semantic analysis passes the test: Knowledge representation and multiple-choice testing. *Manuscript in preparation*. <http://citeseerx.ist.psu.edu/showciting;jsessionid=454EFFEE1536A792B9A122C28DC38605?cid=1795709>
- Lauer, T. (2006). *Learner interaction with algorithm visualizations: viewing vs. changing vs. constructing*. Paper presented at the Proceedings of the 11th annual SIGCSE conference on Innovation and technology in computer science education, Bologna, Italy. [http://portal.acm.org/ft\\_gateway.cfm?id=1140179&type=pdf&coll=GUIDE&dl=GUIDE&CFID=72708914&CFTOKEN=91799524](http://portal.acm.org/ft_gateway.cfm?id=1140179&type=pdf&coll=GUIDE&dl=GUIDE&CFID=72708914&CFTOKEN=91799524)
- Lazar, J., Feng, J. H., & Hochheiser, H. (2010). *Research Methods in Human-Computer Interaction*: Wiley Publishing.
- Lee, B., Plaisant, C., Parr, C. S., Fekete, J.-D., & Henry, N. (2006a). *Task taxonomy for graph visualization*. Paper presented at the Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization, Venice, Italy.
- Lee, B., Plaisant, C., Parr, C. S., Fekete, J.-D., & Henry, N. (2006b). *Task taxonomy for graph visualization*. Paper presented at the Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization.
- Leighton Álvarez, H., Prieto Ferraro, M., & García Peñalvo, F. J. (2012). Metodología para determinar atributos y métricas de calidad en sistemas hipermedia adaptativos educativos basados en estilos de aprendizaje. *Revista Educación*, *29*(1), 91-101. doi: 10.15517/revedu.v29i1.2026
- Leung, Y. K., & Apperley, M. D. (1994). A review and taxonomy of distortion-oriented presentation techniques. *ACM Trans. Comput.-Hum. Interact.*, *1*(2), 126-160. doi: 10.1145/180171.180173

- Levin, D. Z., Cross, R., & Abrams, L. C. (2004). The Strength of Weak Ties You Can Trust: The Mediating Role of Trust in Effective Knowledge Transfer. *Manage. Sci.*, *50*(11), 1477-1490. doi: 10.1287/mnsc.1030.0136
- Levin, J. A., & Datnow, A. (2012). The principal role in data-driven decision making: using case-study data to develop multi-mediator models of educational reform. *School Effectiveness and School Improvement*, *23*(2), 179-201. doi: 10.1080/09243453.2011.599394
- Levis, D. (2004). Cumbre Mundial sobre la Sociedad de la Información: modelo para armar. *Signo y Pensamiento*, *23*(44), 28-40.
- Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction*, *13*(5), 487-509. doi: 10.1016/S0959-4752(02)00042-7
- Littleton, K., & Whitelock, D. (2005). The negotiation and co-construction of meaning and understanding within a postgraduate online learning community. *Learning, Media and Technology*, *30*(2), 147-164.
- Liu, B. (2006). *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data (Data-Centric Systems and Applications)*: Springer-Verlag New York, Inc.
- LMTF, L. M. T. F. (2013). *Toward Universal Learning: A Global Framework for Measuring Learning*. Montreal and Washington:: UNESCO Institute for Statistics and Center for Universal Education at the Brookings Institution.
- Lochner, L. (2011). Non-Production Benefits of Education: Crime, Health, and Good Citizenship *National Bureau of Economic Research Working Paper Series* (Vol. 4). Amsterdam: Elsevier
- Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*, *46*(5), 30-32. doi: citeulike-article-id:9958912
- LTSC, I. L. T. S. C. (2009). Standard for Information Technology -- Education and Training Systems -- Learning Objects and Metadata *WG12: Learning Object Metadata*: IEEE.
- Luo, J., Li, W., Cao, J., & Ge, L. (2006). Integrating Heterogeneous E-learning Systems. *Telecommunications, 2006. AICT-ICIW '06. International Conference on Internet and Web Applications and Services/Advanced International Conference on*, 9-9.
- Lynch, M., Engle, J., & L. Cruz, J. (2011). Priced Out: How the Wrong Financial-Aid Policies Hurt Low-Income Students. In T. E. Trust (Ed.).
- Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education*, *54*(2), 588-599. doi: <http://dx.doi.org/10.1016/j.compedu.2009.09.008>
- Mackinlay, J. (1986). Automating the design of graphical presentations of relational information. *ACM Trans. Graph.*, *5*(2), 110-141. doi: 10.1145/22949.22950

- Mallinson, B., & Sewry, D. (2004). eLearning at Rhodes University - a case study. *Advanced Learning Technologies, 2004. Proceedings. IEEE International Conference on*, 708-710.
- Mandinach, E. B., & Gummer, E. S. (2013). A Systemic View of Implementing Data Literacy in Educator Preparation. *Educational Researcher*, 42(1), 30-37. doi: 10.3102/0013189X12459803
- Mandinach, E. B., & Honey, M. (2008). *Data-driven school improvement : linking data and learning*. New York :: Teachers College Press.
- Manoli, P., & Papadopoulou, M. (2012). Graphic Organizers as a Reading Strategy: Research Findings and Issues. *Creative Education*. 3(3), 348-356. doi: 10.4236/ce.2012.33055.
- Mansell, R. (1999). Information and communication technologies for development: assessing the potential and the risks. *Telecommunications Policy*, 23(1), 35-50. doi: [http://dx.doi.org/10.1016/S0308-5961\(98\)00074-3](http://dx.doi.org/10.1016/S0308-5961(98)00074-3)
- Mansell, R., & Wehn, U. (1998). *Knowledge Societies: Information Technology for Sustainable Development*. Oxford: Oxford University Press.
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity.
- Margaix Arnal, D. (2008). Informe Asociación Profesional de Especialistas en Información sobre web social *Informes APEI* (José Antonio Merlo Vega ed., Vol. 1, pp. 63). Gijón, España: APEI.
- Martín Hernández, A. (2006). CONCEPTOS. In A. de Alvarado Yniesta, A. Martín Hernández, L. Lozano Pérez, J. Lozano Moreno, M. García de Castro, & L. V. Martín Martín (Eds.), *La Formación Sin Distancia* (pp. 28). España.
- Martín-Barbero, J. (2012). Los desafíos estratégicos de la sociedad de la información. *Signo y pensamiento: Sociedad de la información: ¿De qué sociedad estamos hablando?*, 23(44), 7-18.
- Mateo, J. L. (2006). Sociedad del conocimiento *ARBOR Ciencia, Pensamiento y Cultura* (Vol. CLXXXII, pp. 145-151).
- Mathes, A. (2004). Folksonomies - Cooperative Classification and Communication Through Shared Metadata. *Computer Mediated Communication*, 47(10), 1-13.
- Matkovic, K., Ammer, A., Gracanin, D., Purgathofer, W., & Lez, A. (2010). Event Line View: Interactive Visual Analysis of Irregular Time-dependent Data. In T. Robyn, B. Pierre, A. Krager, & P. Olivier (Eds.), *10th international conference on Smart graphics. SG'10* (Vol. 6133, pp. 208-219). Banf, Canada: Springer Berlin Heidelberg.
- Mazza, R., & Dimitrova, V. (2005). *Generation of graphical representations of student tracking data in course management systems*. Paper presented at the Information Visualisation, 2005. Proceedings. Ninth International Conference on, London, UK.

- Mazza, R., & Milani, C. (2004). *GISMO: a Graphical Interactive Student Monitoring Tool for Course Management Systems*. Paper presented at the TEL'04 Technology Enhanced Learning'04 International Conference, Milan, Italy. <http://citeseer.ist.psu.edu/mazza04gismo.html>
- Mazzola, L., & Mazza, R. (2012). *Visualizing Learner Models through data aggregation: a test case*. Paper presented at the Red-conference, rethinking education in the knowledge society [ ISBN: 978-88-6101-010-9 ].
- McCormick, B. H., DeFanti, T. A., & Brown, M. D. (1987). Definition of visualization. *SIGGRAPH Comput. Graph.*, 21(6), 3-3
- Medeiros, F., Gomes, A. S., Amorim, R., & Medeiros, G. (2013). *Architecture for Social Interactions Monitoring in Collaborative Learning Environments as a Support for the Teacher's Awareness*. Paper presented at the Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on.
- Mercer, N. (2000). *Words and minds: How we use language to think together*. London and New York: Routledge.
- Mercer, N., & Wegerif, R. (1999). Is 'exploratory talk' productive talk? In K. Littleton & P. Light (Eds.), *Learning with Computers: Analysing Productive Interaction* (pp. 79): Psychology Press.
- Merriam-Webster Online, D. (2014). webster. Retrieved Sep., 2014, from <http://www.webster.com>
- Michinov, N., Brunot, S., Le Bohec, O., Juhel, J., & Delaval, M. (2011). Procrastination, participation, and performance in online learning environments. *Computers & Education*, 56(1), 243-252. doi: doi:10.1016/j.compedu.2010.07.025
- Miemis, V. (2010). Essential Skills for 21st Century Survival: Part I: Pattern Recognition. <http://emergentbydesign.com/2010/04/05/essential-skills-for-21st-century-survival-part-i-pattern-recognition/>
- Milgram, S., & Jodelet, D. (1976). Psychological Maps of Paris. In H. Proshansky, W. Ittelson, & A. Rivlin (Eds.), *Environmental Psychology: People and Their Physical Settings* (pp. 104-124). New York, USA.: Holt, Rinehart and Winston.
- Milicic, N., Krcadinac, U., Jovanovic, J., Brankov, B., & Keca, S. (2013, 9 Apr.). *Paperista: Visual Exploration of Semantically Annotated Research Papers*. Paper presented at the Proceedings of the LAK Data Challenge, Leuven, Belgium.
- Millen, D., Feinberg, J., & Kerr, B. (2005). Social bookmarking in the enterprise. *Queue*, 3 %6(9), 28-35.
- Milligan, C., & Littlejohn, A. (2013). Patterns of Engagement in Connectivist MOOCs. *Journal of Online Learning and Teaching*, 9(2), 149-159.
- Ming, L., Calvo, R. A., & Pardo, A. (2013). *Tracer: A Tool to Measure and Visualize Student Engagement in Writing Activities*. Paper presented at the Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on.

- Modjeska, D. (1997). Navigation in Electronic Worlds: Research Review for Depth Oral Exam David Modjeska Department of Computer Science. [http://www.dgp.toronto.edu/people/modjeska/Pubs/lit\\_rvw.pdf](http://www.dgp.toronto.edu/people/modjeska/Pubs/lit_rvw.pdf)
- Moodle. (2010a). Moodle Docs: Development: Web services. Retrieved February 20, 2014, from [http://docs.moodle.org/en/Development:Web\\_services](http://docs.moodle.org/en/Development:Web_services)
- Moodle. (2010b). Moodle Tracker. Retrieved February 20, 2014, from <http://tracker.moodle.org/browse/MDL-12886>
- Moodle. (2014). <http://moodle.org>. Retrieved Mar, 2015, from <http://moodle.org/>
- Moore, C. (2005). Measuring Effectiveness With Learning Analytics. Retrieved from Chief Learning Officer Solutions for Enterprise Productivity website: <http://www.clomedia.com/articles/measuring-effectiveness-with-learning-analytics>
- MOPTMA- Ministerio de Obras Públicas, T. y. M. A. (1996). *ESTUDIO TECNICO PARA LA ELABORACION DE UN PLAN DE ACCION PARA LA ADMINISTRACION DESTINADO A IMPULSAR LA TELE-EDUCACION*. Madrid: Ernst & Young.
- Morales Morgado, E. M. (2010). *Gestión del conocimiento en sistemas «e-learning», basado en objetos de aprendizaje, cualitativa y pedagógicamente definidos* (Vol. 273): Universidad de Salamanca.
- Morales Morgado, E. M., Aguilar, D. A. G., & Peñalvo, F. J. G. (2008). *HEODAR: herramienta para la evaluación de objetos didácticos de aprendizaje reutilizables*. Paper presented at the X Simposio Internacional de Informática Educativa SIIE 2008.
- Morales Morgado, E. M., García Pealvo, J. F., & Barrón, Á. (2007). Improving LO quality through instructional design based on an ontological model and metadata. *Journal of Universal Computer Science*, 13(7), 970-979.
- Morales Morgado, E. M., Gómez Aguilar, D. A., García Peñalvo, F. J., & Therón Sánchez, R. (2009). Supporting the Quality of Learning Objects Through Their Ranking Visualization. *International Journal of Emerging Technologies in Learning (IJET)*, 4, 24-29.
- Mottus, A., Kinshuk, Graf, S., & Nian-Shing, C. (2013). *Visualization and Interactivity in the Teacher Decision Support System*. Paper presented at the Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on.
- Munoz, C., Conde, M. Á., & García Peñalvo, F. J. (2010). Moodle HEODAR implementation and its implantation in an academic context. *International Journal of Technology Enhanced Learning*, 2(3), 241-255.
- Murnane, R. J., & Phillips, B. R. (1981). Learning by doing, vintage, and selection: Three pieces of the puzzle relating teaching experience and teaching performance. *Economics of Education Review*, 1(4), 453-465. doi: [http://dx.doi.org/10.1016/0272-7757\(81\)90015-7](http://dx.doi.org/10.1016/0272-7757(81)90015-7)

- Naps, T. L., Rössling, G., Almstrum, V., Dann, W., Fleischer, R., Hundhausen, C., . . . Velázquez-Iturbide, J. Á. n. (2002). *Exploring the role of visualization and engagement in computer science education*. Paper presented at the Working group reports from ITiCSE on Innovation and technology in computer science education, Aarhus, Denmark.
- Naveh, G., Tubin, D., & Pliskin, N. (2010). Student LMS use and satisfaction in academic institutions: The organizational perspective. *The Internet and Higher Education*, 13(3), 127-133. doi: <http://dx.doi.org/10.1016/j.iheduc.2010.02.004>
- Newman, M. E. J. (2003). The Structure and Function of Complex Networks. *SIAM Review*, 45(2), 167-256. doi: citeulike-article-id:1033358
- NGLC Insights, N. G. L. C. N. (2013). Building Blocks for College Completion: Learning Analytics. Retrieved from <http://net.educause.edu/> website: <http://www.educause.edu/library/resources/building-blocks-college-completion-learning-analytics>
- Nguyen, Q., & Huang, M. (2010). A New Interactive Platform for Visual Analytics of Social Networks. In M. L. Huang, Q. V. Nguyen, & K. Zhang (Eds.), *Visual Information Communication* (pp. 231-244): Springer US.
- Nguyen, Q. V., & Huang, M. L. (2005). EncCon: An Approach to Constructing Interactive Visualization of Large Hierarchical Data. *Information Visualization*, 4(1), 1-21. doi: 10.1057/palgrave.ivs.9500087
- Nguyen, Q. V., Huang, M. L., & I, H. (2004). *A new visualization approach for supporting knowledge management and collaboration in e-learning*. Paper presented at the Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on. <http://ieeexplore.ieee.org/search/srchabstract.jsp?arnumber=1320217&isnumber=29243&punumber=9225&k2dockey=1320217@ieeecnfs&query=1320217%3Cin%3Earnumber&pos=0>
- Nicholson, P. (2007). A history of e-learning *Computers and education* (pp. 1-11): Springer.
- Nielsen, J. (1994). *Enhancing the explanatory power of usability heuristics*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, Massachusetts, USA.
- Norman, D. A. (2002). *The Design of Everyday Things*: Basic Books, Inc.
- Norris, D., Baer, L., Leonard, J., Pugliese, L., & Lefrere, P. (2008). Action analytics: measuring and improving performance that matters in higher education. *EDUCAUSE Review*, 43(1), 42-67.
- Norris, D. M., & Baer, L. L. (2009). Linking Analytics to Lifting out of Recession. *National Symposium on Action Analytics*, 1-16. Retrieved from <http://www.educause.edu/> website: <http://www.educause.edu/library/resources/academic-analytics-uses-management-information-and-technology-higher-education>

- Norris, D. M., Baer, L. L., & Offerman, M. (2009a). *A National Agenda for Action Analytics*. Paper presented at the National Symposium on Action Analytics. <http://lindabaer.efoliomn.com/Uploads/SettingaNationalAgendaf orActionAnalytics101509.pdf>
- Norris, D. M., Baer, L. L., & Offerman, M. (2009b). Why Action Analytics for Higher Education? *National Symposium on Action Analytics*. <http://www.edu1world.org/PublicForumActionAnalytics/23338>
- Norris, D. M., & Lefrere, P. (2010). *Transformation through expeditionary change using online learning and competence-building technologies* (Vol. 19).
- O'Reilly, T. (2007). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *International Journal of Digital Economics*, 65, 17-37.
- Object Management Group, I. (2014). UML. 2014, from <http://www.uml.org/>
- Obrenovic, Z. (2011). Design-based research: what we learn when we engage in design of interactive systems. *interactions*, 18(5), 56-59. doi: 10.1145/2008176.2008189
- Ohta, Y., Nakano, H., Suzuki, K., Kiyan, T., Shimizu, Y., Noguchi, C., . . . Akiyama, H. (2005). *Practical study of instructional environments for lifelong e-learning*. Paper presented at the 6th International Conference on Information Technology Based Higher Education and Training, 2005. ITHET 2005.
- ONU. (2013). INFORME SOBRE DESARROLLO HUMANO 2013. In O. d. I. s. D. Humano (Ed.), *Programa de las Naciones Unidas para el Desarrollo*.
- Oreilly, T. (2007). What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. *Social Science Research Network Working Paper Series*, 65(1), 17-37. <http://ssrn.com/abstract=1008839>
- Pak Chung, W., & Thomas, J. (2004). Visual Analytics. *Computer Graphics and Applications, IEEE*, 24(5), 20-21. doi: 10.1109/MCG.2004.39
- Pätzold, S., Rathmayer, S., & Graf, S. (2008). *Proposal for the Design and Implementation of a Modern System Architecture and integration infrastructure in context of e-learning and exchange of relevant data* E. I. F. E-Learning (Ed.) *ILearning Forum 2008*. (pp. 82-90).
- Payne, B. R., & Baroody, A. J. (2006). Service-Oriented Architecture: Technology Selection and Strategic IT Management. In M. Khosrow-Pour (Ed.), *Emerging Trends and Challenges in Information Technology Management* (Vol. 1, pp. 3): Idea Group Inc.
- Pike, W. A., Stasko, J., Chang, R., & O'Connell, T. A. (2009). The science of interaction. *Information Visualization*, 8(4), 263-274. doi: 10.1057/ivs.2009.22
- Ploetzner, R., & Lowe, R. (2004). Dynamic visualisations and learning. *Learning and Instruction*, 14(3), 235-240. doi: <http://dx.doi.org/10.1016/j.learninstruc.2004.06.001>

- Porter, M. F. (1980). An algorithm for suffix stripping. *Program: electronic library and information systems*, 14(3), 130-137.
- Preece, J., Rogers, Y., & Sharp, H. (2001). *Beyond Interaction Design: Beyond Human-Computer Interaction*: John Wiley; Sons, Inc.
- Prendes Espinosa, M. P. (1994). *La imagen didáctica: análisis descriptivo y evaluativo*. Universidad de Murcia, Murcia, España. Retrieved from <http://www.tdx.cat/handle/10803/38710>
- Pretorius, A. J., & Van Wijck, J. J. (2006). Visual Analysis of Multivariate State Transition Graphs. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5), 685-692.
- Prieto Ferraro, M. I. (2006). *METHADIS: metodología para el diseño de sistemas hipermedia adaptativos para el aprendizaje, basada en estilos de aprendizaje y estilos cognitivos*. (Phd.), Universidad de Salamanca, Salamanca, España. Retrieved from <http://hdl.handle.net/10366/21856>
- Pring, R. (2004). *Philosophy of educational research* ((2nd ed.). ed.). London: Continuum.
- Quang Vinh, N., Yu, Q., MaoLin, H., & JiaWan, Z. (2013). TabuVis: A tool for visual analytics multidimensional datasets. *SCIENCE CHINA Information Sciences*, 56(5), 1-12. doi: 10.1007/s11432-013-4870-1
- Raby, R. L. (2009). Comparative and International Education: A Bibliography (2008). *Comparative Education Review*, 53(S1), S1-S125. doi: 10.1086/603609
- Ramaratnam, R. (2007). *An analysis of service oriented architectures.*, Massachusetts Institute of Technology, Master of Science in Engineering and Management. Retrieved from <http://dspace.mit.edu/bitstream/handle/1721.1/42372/234382950.pdf?sequence=1>
- Ravishanker, G. (2011). Doing Academic Analytics Right: Intelligent Answers to Simple Questions. *Research Bulletin*. Retrieved from <http://www.educause.edu/ecar> website: <http://www.educause.edu/library/resources/doing-academic-analytics-right-intelligent-answers-simple-questions>
- Reis Figueira, Á., & Bras Laranjeiro, J. (2007). *Interaction visualization in web-based learning using igraps*. Paper presented at the Proceedings of the eighteenth conference on Hypertext and hypermedia, Manchester, UK.
- Rengarajan, R. (2001). LCMS and LMS: taking advantage of tight integration. [http://www.e-learn.cz/soubory/lcms\\_and\\_lms.pdf](http://www.e-learn.cz/soubory/lcms_and_lms.pdf)
- Rey, G. D. (2012). How seductive are decorative elements in learning material? *Journal of Educational Multimedia and Hypermedia*, 21(3), 257-283.
- Rider, Y., & Thomason, N. (2008). Cognitive and Pedagogical Benefits of Argument Mapping: L.A.M.P. Guides the Way to Better Thinking. In T. Sherborne, S. J. Buckingham Shum, & A. Okada (Eds.), *Knowledge Cartography* (pp. 113-130): Springer London.

- Rivadeneira, A. W., Gruen, D. M., Muller, M. J., & Millen, D. R. (2007). *Getting our head in the clouds: toward evaluation studies of tagclouds*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, California, USA.
- Roberts, J. C. (2007). *State of the art: Coordinated & multiple views in exploratory visualization*. Paper presented at the Coordinated and Multiple Views in Exploratory Visualization, 2007. CMV'07. Fifth International Conference on.
- Robinson, D. H. (1997). Graphic organizers as aids to text learning. *Reading Research and Instruction, 37*(2), 85-105. doi: 10.1080/19388079809558257
- Robinson, D. H., & Kiewra, K. A. (1995). Visual argument: Graphic organizers are superior to outlines in improving learning from text. *Journal of Educational Psychology, 87*(3), 455-467. doi: 10.1037/0022-0663.87.3.455
- Rodríguez, C. O. (2012). MOOCs and the AI-Stanford like courses: Two successful and distinct course formats for massive open online courses. *European Journal of Open, Distance and E-Learning, II*(9), 13 p.
- Romani, L. A. S., & da Rocha, H. V. (2000). Interaction Map: Information Visualization Techniques in Web-Based Distance Education Environments: Institute of Computing, University of Campinas.
- Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications, 33*(1), 135-146. doi: <http://dx.doi.org/10.1016/j.eswa.2006.04.005>
- Romero, C., & Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 40*(6), 601-618. doi: 10.1109/TSMCC.2010.2053532
- Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education, 51*(1), 368-384. doi: <http://dx.doi.org/10.1016/j.compedu.2007.05.016>
- Romero-Zaldivar, V.-A., Pardo, A., Burgos, D., & Delgado Kloos, C. (2012). Monitoring student progress using virtual appliances: A case study. *Computers & Education, 58*(4), 1058-1067. doi: <http://dx.doi.org/10.1016/j.compedu.2011.12.003>
- Rosenberg, M. (2001). *E-Learning: Strategies for Delivering Knowledge in the Digital Age*: McGraw-Hill.
- Rossling, G., Naps, T., Hall, M. S., Karavirta, V., Kerren, A., Leska, C., . . . Velázquez-Iturbide, J. Á. (2006). Merging interactive visualizations with hypertextbooks and course management. *SIGCSE Bull., 38*(4), 166-181. doi: 10.1145/1189136.1189184
- Rouse, M. (2010). "Business analytics" definition. Retrieved 01 Sep 2013, from <http://searchbusinessanalytics.techtarget.com/definition/business-analytics-BA>
- Ruipérez, G., Castrillo Larreta, M. D., & García Cabrero, J. C. (2006). SITUACIÓN DEL "E-LEARNING" EN ESPAÑA: EL ÁMBITO

- UNIVERSITARIO. In A. de Alvarado Yniesta, A. Martín Hernández, L. Lozano Pérez, J. Lozano Moreno, M. García de Castro, & L. V. Martín Martín (Eds.), *La Formación Sin Distancia* (pp. 28). España.
- Russell, T. (2007). *Tag Decay: A View Into Aging Folksonomies*. Paper presented at the ASIS&T Annual Meeting. <http://weblog.terrellrussell.com/2007/11/tag-decay-poster-from-asist-is-online/>
- Sachs, H., Stiebitz, M., & Wilson, R. J. (1988). An historical note: Euler's Königsberg letters. *Journal of Graph Theory*, 12(1), 133-139.
- Salinas, J. (2008). *Algunas perspectivas de los entornos personales de aprendizaje*. Paper presented at the III Jornadas Nacionales TIC y Educación, Lorca (MU).
- Salton, G., Allan, J., & Buckley, C. (1994). Automatic structuring and retrieval of large text files. *Commun. ACM*, 37(2), 97-108.
- Sandars, & Kieran, W. (2004). E-learning for general practitioners: lessons from the recent literature. *Work Based Learning in Primary Care*, 2(4), 305-314.
- Sanz Menéndez, L. (2003). Análisis de redes sociales: o cómo representar las estructuras sociales subyacentes. In C. S. d. I. C. (CSIC), U. d. P. C. (UPC), T. Grupo de Investigación sobre Políticas de Innovación, Formación y, & E. (SPRITE) (Eds.), *Unidad de Políticas Comparadas (CSIC)*  
*Documento de Trabajo* (Vol. 21, pp. 20-29).
- Sarkar, M., & Brown, M. H. (1994). Graphical fisheye views. *Communications of the ACM*, 37(12), 73-83.
- Schildkamp, K., & Teddlie, C. (2008). School performance feedback systems in the USA and in The Netherlands: a comparison. *Educational Research and Evaluation*, 14(3), 255-282.
- Schoenecker, C., & Baer, L. L. (2010). *Measuring What Matters: A Dashboard for Success*. Paper presented at the Annual International Conference Chair Academy 19th, Minneapolis, MN. [http://www.hr.mnscu.edu/training\\_and\\_development/documents/Measuring\\_What\\_Matte.pdf](http://www.hr.mnscu.edu/training_and_development/documents/Measuring_What_Matte.pdf).
- Scholtz, J. (2004). Usability evaluation *Technical report*: National Institute of Standards and Technology (NIST).
- Scholtz, J. (2006). *Beyond Usability: Evaluation Aspects of Visual Analytic Environments*. Paper presented at the Visual Analytics Science And Technology, 2006 IEEE Symposium On.
- Schreurs, B., Teplovs, C., Ferguson, R., de Laat, M., & Buckingham Shum, S. (2013). *Visualizing social learning ties by type and topic: rationale and concept demonstrator*. Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium.
- Schulz, H., Hadlak, S., & Schumann, H. (2011). The design space of implicit hierarchy visualization: A survey. *Visualization and Computer Graphics, IEEE Transactions on*, 17(4), 393-411.
- Schulz, H.-J., & Schumann, H. (2006). *Visualizing graphs-a generalized view*. Paper presented at the Information

- Visualization, 2006. IV 2006. Tenth International Conference on.
- Scott, J. P., & Carrington, P. J. (2011). *The SAGE Handbook of Social Network Analysis*: Sage Publications Ltd.
- Seoane Pardo, A. M. (2014). *Formalización de un modelo de formación online basado en el factor humano y la presencia docente mediante un lenguaje de patrón*. (Doctorado de Formación en la Sociedad del Conocimiento), Universidad de Salamanca, Salamanca. Retrieved from <http://hdl.handle.net/10366/123342>
- Seoane Pardo, A. M., & García Peñalvo, F. J. (2014). Pedagogical Patterns and Online Teaching *Online Tutor 2.0: Methodologies and Case Studies for Successful Learning* (pp. 298-316): IGI Global.
- Shadbolt, N., Berners-Lee, T., & Hall, W. (2006). The Semantic Web Revisited. *IEEE Intelligent Systems*, 21(3), 96-101. doi: <http://dx.doi.org/10.1109/MIS.2006.62>
- Shneiderman, B. (1996). *The eyes have it: a task by data type taxonomy for information visualizations*. Paper presented at the Visual Languages, 1996. Proceedings., IEEE Symposium on.
- Shneiderman, B., & Wattenberg, M. (2001). *Ordered treemap layouts*. Paper presented at the Information Visualization, IEEE Symposium on.
- Shron, M. (2014). *Thinking with Data: How to Turn Information Into Insights*: " O'Reilly Media, Inc."
- Siemens, G., & Baker, R. S. J. d. (2012). *Learning analytics and educational data mining: towards communication and collaboration*. Paper presented at the Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, British Columbia, Canada.
- Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Buckingham Shum, S., Ferguson, R., . . . d Baker, R. S. (2011). Open Learning Analytics: an integrated & modularized platform. *Proposal to design, implement and evaluate an open platform to integrate heterogeneous learning analytics techniques*. <http://cmapspublic3.ihmc.us/rid=1KC16KK3Y-1DGTX1Y-H2/KG-OpenLearningAnalytics.pdf>
- Siirtola, H., Laivo, T., Heimonen, T., & Raiha, K. J. (2009). *Visual Perception of Parallel Coordinate Visualizations*. Paper presented at the Information Visualisation, 2009 13th International Conference.
- Silva, A., & Figueira, Á. (2012). *Visual Analysis of Online Interactions through Social Network Patterns*. Paper presented at the Advanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on.
- Smith, R. B., Hixon, R., & Horan, B. (2001). *Supporting flexible roles in a shared space*. Paper presented at the Collaborative Virtual Environments.
- Snodin, N. S. (2013). The effects of blended learning with a CMS on the development of autonomous learning: A case study of different degrees of autonomy achieved by individual learners.

- Computers & Education*, 61(0), 209-216. doi: <http://dx.doi.org/10.1016/j.compedu.2012.10.004>
- Spector, J. M. (2013). Emerging Educational Technologies and Research Directions. *Journal of Educational Technology & Society*, 16(2), 21-30.
- Spence, R. (2000). *Information Visualization* (M. Chalmers Ed. Vol. 1): John Wiley & Sons, Ltd.
- Spence, R. (2007). *Information visualization: design for interaction* (2nd ed ed.): Pearson/Prentice Hall.
- Srivastava, A., & Gupta, D. J. (2014). *Social Network Analysis: Hardly easy*. Paper presented at the Optimization, Reliability, and Information Technology (ICROIT), 2014 International Conference on.
- Stark, C., Schmidt, K. J., Shafer, L., & Crawford, M. (2002). *Creating e-learning programs: a comparison of two programs*. Paper presented at the Frontiers in Education, 2002. FIE 2002. 32nd Annual.
- Stefaner, M. (2007). *Visual tools for the socio--semantic web*. (Master of Interface Design Programme Master of Interface Design Programme), University of Applied Sciences Potsdam. Retrieved from <http://well-formed-data.net/thesis/>
- Stephenson, E., & Schifter Caravello, P. (2007). Incorporating data literacy into undergraduate information literacy programs in the social sciences A pilot project. *Reference Services Review*, 35(4), 525-540. doi: 10.1108/00907320710838354
- Süral, I. (2010). Characteristics of a sustainable Learning and Content Management System (LCMS). *Procedia - Social and Behavioral Sciences*, 9(0), 1145-1152. doi: <http://dx.doi.org/10.1016/j.sbspro.2010.12.298>
- Suriá, R. (2010). Las TIC en las titulaciones universitarias de grado: análisis del conocimiento y uso en el alumnado de la Universidad a Distancia. *Electronic Journal of Research in Educational Psychology*, 8, 1179-1200.
- Tat, A., & Carpendale, S. (2006, 04-07 Jan.). *CrystalChat: Visualizing Personal Chat History*. Paper presented at the Proceedings of the 39th Annual Hawaii International Conference on System Sciences, 2006. HICSS '06. , Hawaii
- Tempelaar, D. T., Heck, A., Cuypers, H., van der Kooij, H., & van de Vrie, E. (2013). *Formative assessment and learning analytics*. Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium.
- Teyseyre, A. R., & Campo, M. R. (2009). An Overview of 3D Software Visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 15(1), 87-105. doi: 10.1109/TVCG.2008.86
- Therón, R. (2006). Hierarchical-Temporal Data Visualization Using a Tree-Ring Metaphor. In A. Butz, B. Fisher, A. Krüger, & P. Olivier (Eds.), *Smart Graphics* (Vol. 4073, pp. 70-81): Springer Berlin Heidelberg.
- Therón, R., Gonzalez, A., Garcia, F. J., & Santos, P. (2007). *The use of information visualization to support software configuration*

- management*. Paper presented at the Proceedings of the 11th IFIP TC 13 international conference on Human-computer interaction - Volume Part II, Rio de Janeiro, Brazil.
- Therón, R., Santamaria, R., Garcia, J., Gómez Aguilar, D. A., & Paz-Madrid, V. (2007). Overlapper: movie analyzer. *Infovis Conference Compendium*, 140-141.
- Thomas, J. J., & Cook, K. A. (2005). Illuminating the Path: The Research and Development Agenda for Visual Analytics *Illuminating the Path: The Research and Development Agenda for Visual Analytics* (pp. 32-68): National Visualization and Analytics Ctr.
- Thomas, J. J., & Cook, K. A. (2006). A visual analytics agenda. *Computer Graphics and Applications, IEEE*, 26(1), 10-13.
- Thompson, P. (2010). Chapter 10 - Learning by Doing. In H. H. Bronwyn & R. Nathan (Eds.), *Handbook of the Economics of Innovation* (Vol. Volume 1, pp. 429-476): North-Holland.
- Torniai, C., Jovanovic, J., Gasevic, D., Bateman, S., & Hatala, M. (2008). *E-Learning meets the Social Semantic Web*. Paper presented at the Eighth IEEE International Conference on Advanced Learning Technologies, 2008. ICALT '08, Santander, Cantabria, Spain. 10.1109/ICALT.2008.20
- Tory, M., Potts, S., & Moller, T. (2005). A parallel coordinates style interface for exploratory volume visualization. *Visualization and Computer Graphics, IEEE Transactions on*, 11(1), 71-80. doi: 10.1109/TVCG.2005.2
- Tuckman, B. W. (1991). The Development and Concurrent Validity of the Procrastination Scale. *Educational and Psychological Measurement*, 51(2), 473-480. doi: 10.1177/0013164491512022
- Tufte, E. (1990). *Envisioning information*. Cheshire, CT, USA: Graphics Press.
- Tufte, E. (2001). *The Visual Display of Quantitative Information*: Graphics Pr.
- Tufte, E. R. (1997). *Visual explanations: images and quantities, evidence and narrative* (Vol. 6). Cheshire, CT, USA: Graphics Press.
- Tufte, E. R. (2006). *Beautiful evidence* (Vol. 23): Graphics Press Cheshire, CT.
- Tweedie, L. (1997). *Characterizing interactive externalizations*. Paper presented at the Proceedings of the ACM SIGCHI Conference on Human factors in computing systems, Atlanta, Georgia, USA.
- UNESCO. (2002). Forum on the Impact of Open Courseware for Higher Education in Developing Countries. In UNESCO (Ed.), *Final Report* (pp. 30 p.). Paris: UNESCO.
- UNESCO, I. f. S. (2010). Global Education Digest 2010: Comparing Education Statistics Across the World: Instituto de Estadística, UNESCO.
- UNESCO, I. M. d. I. (2005). Hacia las sociedades del conocimiento. *Editorial UNESCO*, 244.

- United States, W.-b. E. C., Kerrey, R., & Isakson, J. (2000). The power of the Internet for learning: Moving from promise to practice: report of the Web-based Education Commission. United States.
- Van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in Higher Education: Establishing a Common Language *EDUCAUSE Learning Initiative* (Vol. ELI Paper 1).
- Van Harmelen, M., & Workman, D. (2012). Analytics for Learning and Teaching. *CETIS Analytics Series*, 1(3). Retrieved from <http://publications.cetis.ac.uk/> website: <http://publications.cetis.ac.uk/wp-content/uploads/2012/11/Analytics-for-Learning-and-Teaching-Vol1-No3.pdf>
- Verbert, K., Drachsler, H., Manouselis, N., Wolpers, M., Vuorikari, R., & Duval, E. (2011). *Dataset-driven research for improving recommender systems for learning*. Paper presented at the Proceedings of the 1st International Conference on Learning Analytics and Knowledge, Banff, Alberta, Canada.
- Viégas, F. B., & Wattenberg, M. (2008). Tag clouds and the case for vernacular visualization. *interactions*, 15(4), 49-52. doi: <http://doi.acm.org/10.1145/1374489.1374501>
- Villalba Simón, M. R., Martínez Liébana, I., Checa Benito, F. J., Robles, M. M., Martín Andrade, P., Nuñez Blanco, M. Á., & Vallés Arándiga, A. (2000). *Aspectos evolutivos y educativos de la deficiencia visual* (Vol. 1): Organización Nacional de Ciegos Españoles (ONCE). Dirección de Educación.
- Vliegen, R., van Wijk, J. J., & van der Linden, E. J. (2006). Visualizing Business Data with Generalized Treemaps. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5), 789-796.
- von Hippel, E., & Tyre, M. J. (1995). How learning by doing is done: problem identification in novel process equipment. *Research Policy*, 24(1), 1-12. doi: [http://dx.doi.org/10.1016/0048-7333\(93\)00747-H](http://dx.doi.org/10.1016/0048-7333(93)00747-H)
- Von Landesberger, T., Fiebig, S., Bremm, S., Kuijper, A., & Fellner, D. (2014). Interaction Taxonomy for Tracking of User Actions in Visual Analytics Applications. In W. Huang (Ed.), *Handbook of Human Centric Visualization* (pp. 653-670): Springer New York.
- von Landesberger, T., Gorner, M., & Schreck, T. (2009). *Visual analysis of graphs with multiple connected components*. Paper presented at the Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on.
- Vygotsky, L. (1989). *El desarrollo de los procesos psicológicos superiores* (M. Cole, V. John-Steiner, S. Scribner, & E. Soubberman Eds.). Barcelona: Crítica.
- Ward, M., & Yang, J. (2004). *Interaction spaces in data and information visualization*. Paper presented at the Proceedings of the Sixth Joint Eurographics - IEEE TCVG conference on Visualization, Konstanz, Germany.
- Ward, M. O. (1994). *XmdvTool: integrating multiple methods for visualizing multivariate data*. Paper presented at the

- Visualization, 1994., Visualization '94, Proceedings., IEEE Conference on.
- Ware, C. (2004). *Information Visualization: Perception for Design*: Morgan Kaufmann Publishers Inc.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). New York, NY: Cambridge university press.
- Weber, M., Alexa, M., & Müller, W. (2001). *Visualizing Time-Series on Spirals*. Paper presented at the Proceedings of the IEEE Symposium on Information Visualization 2001 (INFOVIS'01).
- Wegman, E. J. (1990). Hyperdimensional Data Analysis Using Parallel Coordinates. *Journal of the American Statistical Association*, 85(411), 664-675. doi: 10.2307/2290001
- Wehrend, S., & Lewis, C. (1990). *A problem-oriented classification of visualization techniques*. Paper presented at the Proceedings of the 1st Conference on Visualization'90.
- Weisskirch, R. S., & Milburn, S. S. (2003). Virtual discussion: Understanding college students' electronic bulletin board use. *Internet and Higher Education*, 6(3), 215-225.
- Weiwei, C., Yingcai, W., Shixia, L., Furu, W., Zhou, M. X., & Huamin, Q. (2010). *Context preserving dynamic word cloud visualization*. Paper presented at the IEEE Pacific Visualization Symposium - PacificVis 2010, Taipei, Taiwan. 10.1109/PACIFICVIS.2010.5429600
- Wells, G., & Claxton, G. (2008). Introduction: Sociocultural Perspectives on the Future of Education *Learning for Life in the 21st Century* (pp. 1-17): Blackwell Publishing Ltd.
- Wertsch, J. V., Baker-Sennett, J., Rogoff, B., & Bell, N. (1992). Voices of the Mind: A Sociocultural Approach to Mediated Action. *The American Journal of Psychology*, 105, 506-510. <http://www.jstor.org/stable/1423207>
- Wesch, M. (2009). From knowledgable to knowledge-able: Learning in new media environments. *Academic Commons*. Retrieved from <http://www.academiccommons.org> website: <http://www.academiccommons.org/2014/09/09/from-knowledgable-to-knowledge-able-learning-in-new-media-environments/>
- Wiley, D. (2002). Connecting learning objects to instructional design theory: a definition, a metaphor, and a taxonomy. In D. Wiley (Ed.), *The Instructional Use of Learning Objects* (pp. 3-23): The Agency for Instructional Technology.
- Wilkinson, L. (2005). *The Grammar of Graphics (Statistics and Computing)*: Springer-Verlag New York, Inc.
- Willett, W., Heer, J., & Agrawala, M. (2007). Scented Widgets: Improving Navigation Cues with Embedded Visualizations. *Visualization and Computer Graphics, IEEE Transactions on*, 13(6), 1129-1136.
- Willging, P. (2008). *Técnicas para el análisis y visualización de interacciones en ambientes virtuales* (Vol. 14).
- Williams, F. P., & Conlan, O. (2007). Visualizing Narrative Structures and Learning Style Information in Personalized e-Learning

- Systems. *Advanced Learning Technologies, 2007. ICALT 2007. Seventh IEEE International Conference on*, 872-876.
- Wolff, A., Zdrahal, Z., Nikolov, A., & Pantucek, M. (2013). *Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment*. Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge, Leuven, Belgium.
- Wu, P. (2009). *Do we really understand what we are talking about? A study examining the data literacy capacities and needs of school leaders* (Doctor of Education ProQuest Dissertations and Theses), University of Southern California, Los Angeles, California. Retrieved from <http://search.proquest.com/docview/304996387?accountid=17252>
- Xiaoyu, W., Wenwen, D., Butkiewicz, T., Bier, E. A., & Ribarsky, W. (2011). *A two-stage framework for designing visual analytics system in organizational environments*. Paper presented at the Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on.
- Ya-Xi, C., Santamaría, R., Butz, A., & Therón Sánchez, R. (2009). *TagClusters: Semantic Aggregation of Collaborative Tags beyond TagClouds*. Paper presented at the Proceedings of the 10th International Symposium on Smart Graphics, Salamanca, Spain.
- Yi, J. S., Kang, Y. a., Stasko, J., & Jacko, J. (2007). Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6), 1224-1231. doi: 10.1109/tvcg.2007.70515
- Youn-ah, K., Gorg, C., & Stasko, J. (2009). *Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study*. Paper presented at the Visual Analytics Science and Technology, 2009. VAST 2009. IEEE Symposium on.
- Zenha-Rela, M., & Carvalho, R. (2006). Work in Progress: Self Evaluation Through Monitored Peer Review Using the Moodle Platform. *Frontiers in Education Conference, 36th Annual*, 26-27.
- Zhou, H., Yuan, X., Qu, H., Cui, W., & Chen, B. (2008). Visual Clustering in Parallel Coordinates. *Computer Graphics Forum*, 27(3), 1047-1054. doi: 10.1111/j.1467-8659.2008.01241.x
- Zhou, M. X., & Feiner, S. K. (1996, 28-29 Oct). *Data characterization for automatically visualizing heterogeneous information*. Paper presented at the Proceedings IEEE Symposium on Information Visualization'96, , San Francisco, CA.
- Zhou, M. X., & Feiner, S. K. (1998). *Visual task characterization for automated visual discourse synthesis*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Los Angeles, California, USA.
- Zhu, W., & Chen, C. (2007). Storylines: Visual exploration and analysis in latent semantic spaces. *Computers & Graphics*, 31(3), 338-349. doi: <http://dx.doi.org/10.1016/j.cag.2007.01.025>

Zimmerman, B. J. (2002). Becoming a Self-Regulated Learner: An Overview. *Theory Into Practice*, 41(2), 64-70. doi: 10.1207/s15430421tip4102\_2