

Generación automática de interfaces software para el soporte a la toma de decisiones. Aplicación de ingeniería de dominio y machine learning.

Tesis doctoral

Programa de doctorado en Ingeniería Informática

Departamento de Informática y Automática

Universidad de Salamanca (<https://ror.org/02f40zc51>)

Salamanca, España

Doctoranda

Andrea Vázquez Ingelmo

Directores

Dr. Francisco José García-Peñalvo

Dr. Roberto Therón Sánchez



Resumen extendido

Palabras clave: *Visualización de datos, Visualización de información, Dashboards, Desarrollo dirigido por modelos, Arquitectura dirigida por modelos, Líneas de productos de software, Metamodelado, Generación de conocimiento, Interfaces gráficas de usuario, Interacción persona-ordenador.*

El análisis de datos es un proceso clave para fomentar la generación de conocimiento en determinados dominios o campos de estudio. Con una sólida base informativa derivada del análisis de datos, los responsables de la toma de decisiones pueden diseñar estrategias con el objetivo de obtener beneficios en sus ámbitos de actuación específicos. Sin embargo, dado el constante crecimiento de los volúmenes de datos, el análisis de éstos debe contar con herramientas potentes que permitan la extracción de conocimientos.

Pese a que conceptos técnicos como "dirigido por datos" (en inglés, *data-driven*) y "toma de decisiones" pueden considerarse jerga empresarial, es necesario ser conscientes de que estos conceptos influyen en todos los aspectos de nuestro día a día. Llevamos a cabo procesos de toma de decisiones cuando organizamos nuestro tiempo para coger el autobús puntualmente o cuando miramos las previsiones meteorológicas para decidir si debemos llevar un paraguas en un día nublado. Estamos, conscientemente o no, tomando decisiones basadas en datos constantemente.

Pero, aunque nos enfrentemos a estas decisiones cada día, la comprensión de los datos no es una tarea simple ni un proceso sencillo que pueda darse por sentado. Hay varios mecanismos que se activan [1] para ayudarnos a generar nuevos conocimientos a partir de pequeñas porciones de datos [2-4]. Sin embargo, estos mecanismos no son infalibles. De hecho, vivir en una sociedad altamente conectada, donde se generan

continuamente flujos de datos, puede dificultar estos procesos debido a la sobrecarga de información [5, 6].

En este contexto, el apoyo tecnológico es crucial para facilitar la generación de conocimiento en entornos complejos y con importantes cantidades de datos. Los *dashboards* o cuadros de mando y las visualizaciones de datos ofrecen una solución de software para analizar grandes volúmenes de datos de forma visual con el fin de identificar patrones y relaciones, así como tomar decisiones en función de la información presentada.

Las visualizaciones de datos y los *dashboards* están compuestos de un conjunto de elementos visuales colocados y configurados en función de los datos de entrada. Pero estos elementos se influyen mutuamente, y hay conceptos que ni siquiera aparecen en la pantalla pero que son cruciales, como las características del usuario. Debido a esta complejidad, es importante apoyarse en conocimiento experto a la hora de desarrollar visualizaciones de datos y *dashboards*.

Sin embargo, los usuarios quieran crear una visualización no siempre tendrán a su disposición a un experto en la materia para guiar el proceso y aplicar los principios de diseño adecuados.

Varias herramientas han intentado abordar esta cuestión asistiendo al usuario y automatizando la generación de visualizaciones de datos y *dashboards* mediante la implementación de procesos generativos que capturan y aplican el conocimiento de los expertos adaptando los elementos visuales en función de los datos y el contexto. Este es el caso de herramientas comerciales como Tableau, Microsoft Excel, Google Charts, etc.

Aunque estas plataformas son muy potentes, sigue existiendo un problema relacionado con la transferencia del conocimiento experto a los profesionales, y con la expresividad de las visualizaciones obtenidas.

Por otro lado, las librerías de programación declarativa e imperativa pueden mejorar la expresividad y funcionalidad de las visualizaciones de datos y *dashboards* desarrollados, pero, en este caso, suelen venir acompañadas de una pronunciada curva de aprendizaje que dificulta el proceso de implementación.

Teniendo en cuenta lo anterior, esta Tesis Doctoral aborda el reto de mejorar el proceso de desarrollo de *dashboards* y visualizaciones de datos, así como su calidad y prestaciones en términos de personalización, usabilidad y flexibilidad, entre otros.

La hipótesis principal de este trabajo se plantea como sigue: “*La automatización del desarrollo de interfaces de usuario a medida para apoyar los procesos de toma de decisiones incrementa sus beneficios en términos de características funcionales y no funcionales.*”

En otras palabras, el objetivo principal de la investigación es diseñar e implementar un marco generativo para el desarrollo automático, sistemático y personalizado de *dashboards*, además de discutir los conocimientos alcanzados al automatizar la generación de estas herramientas. Mediante este enfoque, se trata de fomentar la individualización, la usabilidad y la flexibilidad para maximizar los beneficios derivados de las herramientas generadas.

Debido a la naturaleza mixta de los artefactos y escenarios propuestos, esta tesis se ha llevado a cabo siguiendo un proceso iterativo en el que el conocimiento adquirido a través de las experiencias pasadas y los resultados de los diferentes ciclos es crucial para las siguientes etapas. Para llevar a cabo este proceso se ha seguido el marco metodológico de la Investigación-Acción [7].

Kemmis planteó la Investigación-Acción [8] como un método de indagación llevado a cabo por los participantes en situaciones sociales con el objetivo de mejorar y comprender sus propias prácticas sociales y sus contextos.

Posteriormente, McTaggart y Kemmis describieron las características de esta metodología. La metodología de Investigación-Acción se basa en una espiral cíclica de investigación y acciones compuesta por una serie de fases y secuencias [9]. Por lo tanto, la Investigación-Acción es un proceso iterativo en el que cada ciclo proporciona un resultado que será la entrada para el siguiente ciclo.

El marco elegido para el desarrollo de *software* es un enfoque ágil basado en SCRUM [10]. Este marco proporciona los procesos, reglas, prácticas, roles y artefactos necesarios para aumentar la productividad de los equipos de desarrollo mediante un ciclo de desarrollo de software iterativo e incremental.

Para evaluar los artefactos derivados se ha utilizado un enfoque de investigación. La investigación se ha llevado a cabo utilizando métodos cuantitativos y cualitativos [11], aprovechando las dos perspectivas para obtener una visión más amplia de los resultados para afrontar los siguientes ciclos de investigación-acción.

Sin embargo, es necesario formalizar el problema a tratar antes de poder iniciar dichos ciclos. Para ello, se ha realizado una revisión sistemática de la literatura (SLR, por sus siglas en inglés) bajo las directrices propuestas por Kitchenham [12, 13] y García-Peñalvo [14], con el fin de estudiar problemas similares y soluciones desarrolladas previamente para comprender el contexto y el estado actual del campo.

Además de obtener un panorama de soluciones, el principal resultado de la revisión bibliográfica es un análisis crítico de las metodologías y arquitecturas encontradas en los trabajos seleccionados. Este tipo de análisis ofrece un buen punto de partida para diseñar e implementar la primera propuesta de un sistema para la generación automática de *dashboards*.

Las preguntas de investigación abarcaron aspectos relevantes para tener en cuenta a la hora de abordar los flujos generativos de *dashboards*. Con la información recopilada, fue posible seleccionar la mejor estrategia para implementar enfoques que abordan la generación automática de estas herramientas.

En virtud de los resultados del SLR, la decisión fue seguir un enfoque de metamodelado para conceptualizar el marco generativo y el paradigma de la línea de productos de software (SPL, por sus siglas en inglés) para materializar y transformar las características abstractas en código fuente. El análisis de los artículos recuperados ha demostrado que estos dos enfoques son viables en este ámbito; casi un tercio de los trabajos seleccionados (8 de 30) emplean uno de estos dos paradigmas.

Sin embargo, la viabilidad de las soluciones no fue el único objeto de estudio. Otros atributos, como la flexibilidad y las capacidades evolutivas, la posibilidad de transferir las soluciones a cualquier dominio de datos, la trazabilidad de los requisitos y el potencial de integrar algoritmos de inteligencia artificial (IA) para adaptar las características del *dashboard* a los cambios del entorno también fueron objeto de esta revisión.

Por ejemplo, algunos trabajos mencionan el uso de modelos de aprendizaje automático (ML, por sus siglas en inglés) para liberar a los usuarios de tareas complejas como la configuración del diseño del *dashboard*. Sin embargo, estas aplicaciones no están detalladas o se encuentran en sus primeras fases de desarrollo.

Esta línea de investigación es muy prometedora, ya que los enfoques de IA podrían aportar varios beneficios para ayudar a los usuarios y proporcionarles pautas útiles para aprender y comprender cómo diseñar *dashboards* y visualizaciones eficaces. En este sentido, la elección de un enfoque de metamodelado también es adecuada para aplicar métodos de ML, ya que estos modelos requieren datos estructurados de los que aprender. Las instancias del metamodelo pueden proporcionarse como entradas para identificar los patrones que hacen que determinadas configuraciones sean útiles, eficientes, eficaces, utilizables, etc.

En definitiva, la realización de este análisis ha proporcionado evidencias claras de que las soluciones que siguen los paradigmas de meta-modelado y/o SPL cumplen con estas propiedades, concluyendo que la versatilidad de estas metodologías proporciona buenos cimientos para implementar un sistema generativo basado en un metamodelo.

Concretamente, los metamodelos son artefactos cruciales en los paradigmas de desarrollo dirigido por modelos (MDD, por sus siglas en inglés) [15-17], ya que permiten mapear entidades de niveles muy abstractos a entidades más detalladas e incluso al código fuente mediante transformaciones.

El *Object Management Group* (OMG) propone la arquitectura dirigida por modelos (MDA, por sus siglas en inglés) como guía para implementar el desarrollo dirigido por modelos. Esta arquitectura proporciona un marco para el desarrollo de software que emplea modelos para describir y definir el sistema objetivo [18]. La principal diferencia entre MDD y MDA es que la propuesta del OMG utiliza un conjunto de estándares: meta-objeto (MOF), lenguaje de modelado unificado (UML), intercambio de metadatos XML (*Extensible Markup Language*), y consulta/vista/transformación (QVT).

El marco MDA está compuesto por cuatro capas arquitectónicas. Cada capa representa un nivel de abstracción de las entidades representadas. La capa más abstracta (nivel M3) se conoce como el nivel meta-metamodelo. Esta capa define estructuras y conceptos básicos para representar capas menos abstractas, así como a sí misma, y puede implementarse con el mencionado estándar MOF.

El nivel M2, es decir, el nivel de metamodelo, cumple con el metamodelo y representa entidades y relaciones abstractas. Los metamodelos pueden verse como lenguajes específicos de dominio (DSL, por sus siglas en inglés) que expresan características comunes y genéricas del dominio objetivo.

El nivel M1, define modelos que instancian y especifican las características abstractas contenidas en el metamodelo, y su sintaxis debe cumplir con el nivel M2.

Por último, el nivel M0 representa aplicaciones del mundo real basadas en un modelo M1 previamente definido.

Siguiendo este paradigma y la metodología de investigación-acción, el metamodelo desarrollado ha sido sometido a diferentes iteraciones antes de obtener la versión final. Estos incrementos no sólo han permitido centrarse en entidades de dominio específicas en cada iteración, sino también revisar y resolver posibles problemas que pudieran surgir en iteraciones posteriores.

Las primeras iteraciones se centraron en los elementos tangibles de los *dashboards* (diseño, recursos, componentes visuales, etc.), seguidas de un par de iteraciones para abordar la caracterización de la audiencia (el usuario).

Los dos últimos incrementos abordaron conceptos de más alto nivel, como los patrones de interacción (comportamientos dinámicos que no son explícitamente tangibles) y las características del dominio de los datos (el conocimiento implícito sobre el contexto en el que se enmarcan los conjuntos de datos del mundo real).

El desarrollo del metamodelo no solo proporcionó un artefacto para construir un marco generativo para los *dashboards*, sino que también incrementó el conocimiento del autor respecto a este dominio, que es otro beneficio bien conocido del uso de la ingeniería de dominio: capturar, generar y reutilizar el conocimiento.

Este enfoque también impulsó la decisión de utilizar plantillas de código como método para materializar los puntos de variabilidad de la SPL. Las plantillas de código se seleccionaron dada su idoneidad en este contexto y su parecido con el meta-modelado en la filosofía de encapsular y descomponer entidades complejas en elementos primitivos.

El emparejamiento de las entidades del metamodelo con sus respectivos fragmentos de código permitió el desarrollo de un generador de *dashboards*. Este hito marcó el inicio de las fases de validación y aplicación de los artefactos desarrollados.

En este sentido, se realizó un estudio de validación del contenido del metamodelo para comprobar la coherencia, la relevancia y la claridad de las diferentes secciones de este artefacto. Al validar el metamodelo, es posible identificar las posibles limitaciones e inconvenientes de la representación del dominio de los *dashboards* y abordarlos antes de utilizar este artefacto para instanciar estas herramientas en el mundo real.

En primer lugar, se aplicó el marco de calidad del metamodelo propuesto en [19] para comprobar la calidad de la versión Ecore del metamodelo de *dashboards* antes y después de introducir las modificaciones. El metamodelo cumplía con las treinta características del marco relacionadas con el diseño, las mejores prácticas, las convenciones de nomenclatura y las métricas, lo que demuestra su calidad.

En segundo lugar, se llevó a cabo una validación por expertos de la versión final del metamodelo de *dashboards*. El objetivo de esta validación es comprobar si las secciones del metamodelo son claras, coherentes y pertinentes. En particular, se aplicó el juicio de expertos [20].

Para esta tarea, se creó un cuestionario en línea con seis secciones diferentes del metamodelo (diseño, características del usuario, objetivos y tareas, relaciones entre el usuario y el *dashboard*, componentes primitivos de las visualizaciones de datos, y dominio y operaciones de datos), además del metamodelo completo. Cada sección se valoró en función de las dimensiones mencionadas utilizando puntuaciones de 1 a 4,

donde 1 implica que la sección no cumple el criterio, y 4 que lo cumple en gran medida.

Los resultados arrojaron buenas puntuaciones (puntuaciones altas -4- y medias -3-) y sólo cinco puntuaciones bajas -2- relacionadas con la relevancia y coherencia de la sección de diseño del *dashboard* (debido a su simplicidad) y la relevancia de las primitivas de visualización y el dominio de los datos (debido al alto nivel de detalle que podría ser "a demasiado bajo nivel para los usuarios finales").

En lo que se refiere a aplicaciones, se han llevado a cabo varias investigaciones para probar el funcionamiento del metamodelo y del generador de *dashboards* en diferentes escenarios, tanto teóricos como prácticos.

En cuanto a la dimensión teórica de las investigaciones, el metamodelo se ha integrado con éxito con otros metamodelos para apoyar la generación de conocimiento en ecosistemas de aprendizaje. Aunque la existencia de un metamodelo de ecosistemas de aprendizaje resuelve la mayor parte de los problemas asociados a la definición y desarrollo de estas soluciones tecnológicas, existen algunas cuestiones relacionadas con el análisis de los flujos de información y el apoyo a los procesos de toma de decisiones que deberían mejorarse.

La conexión de ambos metamodelos dio lugar a un metamodelo holístico que se puso a prueba mediante la instanciación de un modelo centrado en la provisión de un ecosistema de salud para apoyar a cuidadores [21]. Se diseñó un *dashboard* para lograr diferentes objetivos de información relacionados con la gestión de los cuidadores.

El metamodelo también se aplicó como marco para conceptualizar e instanciar *dashboards* en diferentes dominios. Los escenarios de esta categoría incluyen la instanciación manual del metamodelo en modelos que representan visualizaciones de datos del mundo real, la identificación de sesgos, la automatización del proceso de instanciación a través de herramientas externas y el concepto de proporcionar *dashboards* como servicio a través de interfaces de programación de aplicaciones (API, por sus siglas en inglés).

En cuanto a las aplicaciones prácticas, se ha hecho hincapié en cómo transformar el metamodelo en una instancia adaptada a un contexto específico, y cómo

transformar finalmente este modelo posterior en código, es decir, en el producto final y funcional. En este contexto, se probó la generación automática de *dashboards* en un programa de doctorado.

En primer lugar, se llevó a cabo un proceso de elicitación de requisitos para modelar los *dashboards* que se generarían. El proceso de obtención de requisitos incluyó una entrevista con un miembro del comité de calidad del Programa de Doctorado para entender qué datos podrían mostrarse en un posible *dashboard*.

Estas propuestas fueron finalmente implementadas en el Programa de Doctorado de Educación en la Sociedad del Conocimiento de la Universidad de Salamanca (España) [22]. Se realizó un estudio de usabilidad para puntuar la integración de la visualización de datos en el portal de doctorado utilizando el cuestionario *System Usability Scale* (SUS) [23]. 35 participantes (entre los que se encontraban estudiantes de doctorado, asesores y gestores) respondieron al cuestionario SUS. La media de usabilidad percibida de las visualizaciones del portal de doctorado fue de 75,36, lo que puede considerarse una buena puntuación en el contexto del cuestionario SUS.

Esta aplicación sirvió para comprobar la viabilidad de la generación automática de código en entornos reales. Sin embargo, la transformación del metamodelo en modelos efectivos de *dashboards* y visualizaciones fue llevada a cabo manualmente. En este sentido, se exploró una combinación del pipeline generativo con la IA. La IA se encargaría de aprender las buenas prácticas y directrices en el diseño de la visualización de datos y de seleccionar la mejor configuración dadas las necesidades específicas del contexto.

En este sentido, se realizó un estudio para probar la idea de utilizar el generador de *dashboards* para entrenar modelos de ML e identificar configuraciones potencialmente perjudiciales en las visualizaciones de datos [24]. Para ello, se generó un conjunto de datos de entrenamiento de diferentes visualizaciones de información mediante el generador desarrollado. Además, al utilizar este enfoque, las configuraciones de las visualizaciones generadas ya estaban estructuradas y

preparadas para su procesamiento y uso como entrada para algoritmos de ML, lo que también ahorra tiempo y permitía centrarse más a fondo en el proceso de etiquetado. Este proceso de etiquetado consistió en evaluar las visualizaciones generadas como “útiles” o “no útiles”, teniendo en cuenta si los datos mostrados con distintas configuraciones podrían llevar a engaño o a conclusiones erróneas.

Una vez generado el conjunto de datos, se entrenaron diferentes modelos de ML: Naïve Bayes, Support Vector Machine (SVM), Ada Boost y Random Forest (RF). El clasificador RF obtuvo los mejores resultados en términos de exactitud, precisión, recuperación y puntuación F1. Dados estos resultados, se eligió el clasificador RF para evaluar los resultados de las predicciones individuales introduciendo manualmente los valores de otras visualizaciones.

Una de las razones de la alta precisión es que el modelo imita los criterios definidos previamente para clasificar cada visualización. Sin embargo, esto también indica que el modelo aprendió satisfactoriamente las características más importantes del proceso de clasificación. Las características más importantes fueron los valores máximos y mínimos del dominio de las escalas del eje X, del eje Y, y del tamaño de los elementos visuales. Este resultado se alinea con investigaciones anteriores encontradas en la literatura, en las que la definición de los rangos de las escalas de una visualización es determinante para su correcta comprensión [25].

Por último, se desarrolló una plataforma que permite a los usuarios instanciar modelos a partir del metamodelo de *dashboards*. Esta plataforma, denominada MetaViz, permitió convertir el metamodelo en un recurso más accesible para las aplicaciones prácticas, así como implementar los mecanismos para abordar transformaciones complejas del modelo de forma visual.

La arquitectura de MetaViz se basa en los conceptos del paradigma MDA descritos anteriormente, y las capas MDA se materializan en el sistema a través de diferentes módulos de software.

No sólo la arquitectura del software se basa en las capas MDA, sino también la interfaz de usuario. Una barra de herramientas permite a los usuarios añadir nuevos elementos básicos a su espacio de trabajo. Los parámetros de configuración de cada

elemento también se basan en los atributos y relaciones del metamodelo. El espacio de trabajo permite la instanciación de nuevos *dashboards* conectando y configurando los diferentes "meta-elementos". Por último, un lienzo muestra el *dashboard* o las visualizaciones generadas en tiempo real.

Finalmente, se ha llevado a cabo un conjunto de casos de estudios en diferentes dominios. El metamodelo y el sistema generativo se probaron en el ámbito del empleo y la empleabilidad. Después de esta integración, el metamodelo se mejoró y la nueva versión, junto con el enfoque de ecosistema conceptualizado en [26], se integraron en diferentes plataformas de salud basadas en datos. La dimensión de aprendizaje del metamodelo y el instanciador gráfico -MetaViz- también se ha evaluado en el contexto educativo.

La versión preliminar del metamodelo y la línea de productos de software se utilizaron para abordar la personalización del análisis visual [27, 28] en el dominio del empleo y la empleabilidad. En concreto, este enfoque se aplicó en el Observatorio Español de Empleo y Empleabilidad Universitaria (OEEU) [29, 30]. Esta red de investigadores y técnicos realiza estudios sobre estos campos en el contexto académico [30-32], a través de un enfoque basado en datos para recoger, analizar, visualizar y difundir datos de empleo y empleabilidad de los egresados de las universidades españolas.

Para los *dashboards* del Observatorio, se definieron tres componentes visuales configurables: un diagrama de dispersión, un diagrama de cuerdas y un mapa de calor. Estas visualizaciones responden a los requisitos de los datos del Observatorio, pero pueden reutilizarse para otros dominios de datos. Se diseñó un DSL para lograr la conexión entre el metamodelo y el código fuente, y se basó en las características del dominio identificado del SPL.

Los resultados muestran el éxito de la implementación de la SPL en este contexto. La versión anterior del sistema del OEEU proporcionaba *dashboards* estáticos con los mismos indicadores para cada usuario. Con este enfoque, los *dashboards*

pueden personalizarse mediante la instanciación de los componentes y funcionalidades deseados.

En lo que respecta al ámbito de la salud, se han llevado a cabo diferentes trabajos de investigación relacionados con plataformas dirigidas por datos y centrados en proporcionar interfaces que ayuden tanto a los procesos de toma de decisiones como a las aplicaciones de IA. En primer lugar, el desarrollo de una plataforma (CARTIER-IA) que da soporte a todas las tareas relacionadas con la gestión de datos (incluida la recopilación de datos estructurados e imágenes médicas) y que también permite tanto a los profesionales sanitarios como a los científicos de datos aplicar modelos de IA a las imágenes almacenadas.

El diseño de la plataforma ha hecho posible la aplicación de algoritmos de IA en los datos de imagen sin requerir conocimientos de programación. Los algoritmos de IA están disponibles para que cualquier usuario los ejecute y recupere sus resultados. Esta arquitectura también permite la integración de fuentes de datos externas (como REDCap [33]) y otros módulos, como el generador de *dashboards* desarrollado en esta tesis.

En segundo lugar, se ha implementado otra plataforma gráfica (KoopamL) para ofrecer interfaces intuitivas y educativas con el fin de construir y ejecutar enfoques de ML en el contexto médico. La arquitectura de KoopaML se basa en diferentes módulos conectados por flujos de información. Uno de los principales propósitos de este diseño es proporcionar flujos flexibles con componentes y algoritmos reutilizables.

Para dar soporte al análisis exploratorio de datos [34], que es una parte crucial de las aplicaciones de ML, se integró en KoopaML el generador de *dashboards*. Esta integración permitió la generación dinámica de resúmenes de los datos de entrada, así como informes de los resultados derivados del proceso de entrenamiento adaptados a las métricas requeridas por el usuario.

En tercer lugar, el metamodelo y el generador también se integraron en un servicio web para visualizar los resultados del estudio SALMANTICOR [35], un estudio descriptivo transversal de base poblacional sobre la prevalencia de la cardiopatía estructural y sus factores de riesgo con un total de 2.400 individuos. La

plataforma no sólo se centró en proporcionar visualizaciones básicas para resumir los resultados, sino también en ofrecer una buena experiencia de usuario para alcanzar conocimientos sobre el estudio [36]. Por estas razones, la arquitectura se compone de módulos que supervisan la recuperación, la visualización y el análisis de cada sección del estudio. La división de la arquitectura en módulos individuales pero relacionados entre sí permite la flexibilidad para ampliar la plataforma con más funcionalidades, análisis y visualizaciones.

El *front-end* proporciona una interfaz utilizable y visualizaciones de datos para navegar por los resultados del estudio, y el *back-end* ofrece funcionalidades de cálculo, almacenamiento y recuperación de datos a través de llamadas a la API. Además, las visualizaciones de datos también pueden generarse a demanda mediante interacciones simples utilizando el generador de *dashboards*.

Por último, en lo que respecta al contexto educativo, se diseñó un estudio para probar el rol del metamodelo y el instanciador gráfico (MetaViz) como recursos de aprendizaje.

Existen varias plataformas para facilitar el proceso de diseño e implementación de visualizaciones de datos. Sistemas como Tableau (<https://www.tableau.com/>), Microsoft Excel (<https://www.microsoft.com/microsoft-365/excel>), Power BI (<https://powerbi.microsoft.com/>), etc., proporcionan interfaces gráficas que permiten a los usuarios sin experiencia en programación crear visualizaciones de datos e incluso les ayudan en el proceso de diseño para elegir las mejores configuraciones. Sin embargo, es fundamental comprender y tener en cuenta todos los elementos que intervienen en las visualizaciones de datos para ofrecer una visualización de la información eficaz y que no sea confusa o engañosa [37-40].

Para explorar el potencial papel educativo del metamodelo, se desarrolló un estudio piloto para medir la comprensibilidad de los elementos que intervienen en el diseño de visualizaciones de datos y *dashboards*. El objetivo de este estudio es comprobar si MetaViz ofrece una experiencia más educativa y proporciona a los

usuarios una mayor comprensión de los conceptos del campo de la visualización de datos que otras herramientas comerciales.

Siguiendo los resultados y aplicaciones descritas, se ha podido observar que el desarrollo de un metamodelo de *dashboards* ha abierto varias vías de investigación, tanto a nivel teórico como práctico.

La versatilidad de este artefacto y del sistema generativo ha sido validada a través de los casos de estudio descritos previamente. En cuanto a la validación de expertos, los resultados demostraron que las entidades y relaciones identificadas eran pertinentes, coherentes y comprensibles. Estas características son cruciales, especialmente en este contexto. Debido a la complejidad inherente al dominio de la visualización de datos y los cuadros de mando, es necesario transmitir el conocimiento y las entidades relacionadas con este contexto de forma comprensible.

La realización de estudios de casos en diferentes dominios también ha mejorado el metamodelo y el sistema generativo, ya que de las integraciones surgieron nuevas entidades y relaciones. La evolución de estos recursos puede observarse a través de los estudios presentados, partiendo de una versión muy básica y de grano grueso [27, 28] y terminando en un marco completo que ha servido para entrenar modelos de ML [24], conceptualizar *dashboards* a alto nivel [41-45], apoyar los procesos de toma de decisiones en plataformas de salud [36, 46-49], e incluso conducir la arquitectura e interfaz de una plataforma completa (MetaViz).

Todos estos estudios han formado parte de los ciclos de investigación-acción definidos, los cuales han permitido identificar nuevas relaciones, conceptos y matices en el ámbito de la visualización de datos, dando lugar a un metamodelo potente y versátil para definir *dashboards* y visualizaciones de datos.

Aunque existen otros metamodelos de *dashboards* en la literatura [50-52], no tienen en cuenta las características de grano fino, que son cruciales en este dominio, ya que una ligera modificación en el diseño de un *dashboard* o de una visualización de datos podría llevar a conclusiones significativamente diferentes [25, 53, 54].

El desarrollo del metamodelo no ha sido un proceso lineal. Se han realizado varias mejoras y modificaciones antes de obtener la versión actual. Estas

modificaciones son el resultado de los estudios que se han derivado de la aplicación de la ingeniería de dominio en este contexto.

La primera versión del metamodelo carecía de detalles tanto en lo que respecta a los componentes visuales como al usuario final. La inclusión del usuario como elemento extremadamente significativo dentro del ámbito de los cuadros de mando y la visualización de datos es crucial. Los procesos de desarrollo de visualizaciones y *dashboards* comienzan con el usuario (obtención de requisitos) y terminan con el usuario (refinamiento del producto) [37, 55, 56], por lo que no sólo se deben tener en cuenta las características técnicas de un cuadro de mando a la hora de modelar estas herramientas, ya que estas características surgen de los requisitos de los usuarios y están influenciadas por ellos [57].

De hecho, los resultados de la instanciación en [41] y [43] han demostrado que los objetivos de información de los usuarios son necesarios para diseñar *dashboards* que soporten diferentes roles y necesidades.

Lo mismo ocurre con los trabajos derivados de la integración del sistema generativo en las plataformas sanitarias descritas en [36, 46-49]. La heterogeneidad de las fuentes de datos y la variedad de roles involucrados en el ámbito de la salud exigen interfaces altamente personalizables. En este sentido, la provisión de *dashboards* como servicio [26] permitió la posibilidad de generar estas herramientas de forma transparente y construir bajo demanda en diferentes etapas del análisis de datos estructurados, imágenes DICOM y resultados de ML.

Además, ha sido necesario incluir en el metamodelo entidades y relaciones relacionadas con el dominio datos. El experimento realizado en [24] permitió etiquetar las visualizaciones de datos utilizando el conocimiento de los expertos sobre el contexto de los datos. Sin embargo, en el segundo experimento [58], en el que los datos se generaron de forma aleatoria (y, por tanto, no pertenecían a ningún contexto específico), fue casi imposible decidir si una visualización era potencialmente engañosa, porque se necesitaba más información para realizar el etiquetado. Gracias

a esta experiencia, se identificó la falta de la noción de contexto de los datos y de las variables de dominio en el metamodelo, lo cual se corrigió posteriormente [58].

En cuanto a los detalles técnicos, el generador de *dashboards* se ha probado con diferentes lenguajes: VegaLite [59] en [60], React (<https://es.reactjs.org/>) en [61] y, finalmente, D3.js [62] en [24, 26, 36, 43, 44, 48, 49] y en el desarrollo de MetaViz.

La decisión de utilizar D3.js fue impulsada por la expresividad proporcionada por este lenguaje, que se alinea con la estructura de grano fino del metamodelo. Sin embargo, el uso de plantillas de código para materializar las características de la SPL en código [63] proporciona la posibilidad de desarrollar los activos *software* a través de diferentes tecnologías.

El metamodelo también fue un recurso esencial para explorar la aplicación de la IA al ámbito de la visualización de datos [24]. En primer lugar, siguiendo el desarrollo dirigido modelos y la ingeniería de la línea de productos de software, fue posible generar automáticamente las visualizaciones que posteriormente se etiquetaron para construir un conjunto de entrenamiento. Además, en segundo lugar, el metamodelo proporcionó las características y relaciones para estructurar el esquema del conjunto de datos de entrenamiento, lo que supuso un paso crucial antes de aplicar cualquier algoritmo de ML.

Las métricas de precisión y exactitud muestran que el modelo de ML resultante aprendió de los conocimientos implícitos y las heurísticas que se utilizaron para etiquetar manualmente el conjunto de entrenamiento. Los criterios seguidos para entrenar los modelos podrían considerarse obvios o muy básicos (como no exagerar los valores de la escala o utilizar determinadas codificaciones) [24], pero este estudio estaba motivado en ofrecer un método para que los novatos o los usuarios no expertos sean conscientes de las configuraciones engañosas que pueden estar introduciendo inconscientemente en sus diseños.

Además, el metamodelo ha proporcionado la columna vertebral para el desarrollo de una plataforma de generación de *dashboards* a través de una interfaz gráfica. La plataforma MetaViz (<https://metaviz.grial.eu/>) representa la unificación de todo el conocimiento derivado de la investigación realizada durante esta tesis. La

implementación de MetaViz ha servido para materializar el metamodelo en un recurso funcional que puede ser explorado interactivamente, instanciado y transformado de forma transparente en código real a través del sistema generativo.

El número de aplicaciones del metamodelo en diferentes dimensiones (teóricas y prácticas) y dominios (empleo y empleabilidad, salud y educación) es también un resultado en sí mismo. Todos los resultados asociados a esta tesis están impulsados por el metamodelo de *dashboards*, lo que demuestra también su versatilidad y flexibilidad a la hora de conceptualizar, generar y capturar conocimiento relacionado con los *dashboards* y las visualizaciones de datos.

En este sentido, es evidente concluir que los enfoques MDD y SPL son beneficiosos para mejorar los procesos de desarrollo de cuadros de mando de información y visualizaciones de datos a medida.

Esta tesis también ha servido no sólo para analizar la automatización del diseño e implementación de estas herramientas, sino para explorar cómo concienciar sobre buenas prácticas mientras se desarrollan.

La combinación de los resultados y las conclusiones anteriores han permitido afirmar que el objetivo principal y los subobjetivos derivados de esta tesis se han alcanzado, y tras sus resultados, que la hipótesis planteada al inicio de este proyecto de investigación es válida.

En lo que respecta al trabajo futuro, el desarrollo del metamodelo y el sistema generativo ha desbloqueado varias oportunidades de investigación relacionadas con diferentes áreas.

En primer lugar, el metamodelo ha demostrado ser un artefacto útil para impulsar la investigación relacionada con la visualización de datos. Varios conceptos capturados en el metamodelo pueden ser explorados más a fondo, específicamente los relacionados con las características y objetivos del usuario. Aunque modelar y detectar los sesgos y creencias de los usuarios es un reto, los beneficios derivados de la recolección de esta información serían muy valiosos para luchar contra las noticias falsas, la polarización y las visualizaciones de datos engañosas.

En el ámbito de la ingeniería de software, el metamodelo puede mejorarse con reglas más específicas [64] que recojan buenas prácticas y directrices relacionadas con el ámbito de la visualización de datos. Estas mejoras pueden sentar las bases de un sistema de recomendación basado en reglas o una base de conocimiento para el diseño de visualizaciones de datos.

Por otro lado, en el ámbito de la Interacción Persona-Ordenador, una línea de investigación futura es continuar con la validación del usuario en términos de rendimiento, usabilidad o satisfacción, entre otras métricas, de los *dashboards* generados, así como de la plataforma instanciadora gráfica.

Las ideas presentadas en esta tesis también pueden aplicarse en contextos educativos. El metamodelo de *dashboards* captura el conocimiento del dominio, proporcionando un recurso de aprendizaje para enseñar los elementos básicos de las visualizaciones de datos. Otros trabajos podrían consistir en medir el rendimiento del instanciador gráfico en cuanto a su capacidad para proporcionar una experiencia de aprendizaje a sus usuarios.

Finalmente, también hay oportunidades de investigación en el campo de la IA. La noción de identificar visualizaciones engañosas mediante el entrenamiento de algoritmos de IA a través de las entidades capturadas en el metamodelo puede seguirse explorando para mejorar el modelo, e incluso desarrollar un detector automatizado de visualizaciones de datos engañosas. Además, los modelos de IA también podrían aplicarse para ayudar al proceso de instanciación a través de configuraciones recomendadas dado el contexto del *dashboard* o visualización y las características del usuario.

Para concluir, a lo largo del desarrollo de esta tesis doctoral se han realizado una serie de publicaciones científicas para validar la propuesta. El proceso de publicación en diferentes medios ha permitido obtener la retroalimentación de expertos en la materia. En concreto, se han publicado 11 artículos en revistas indexadas y 23 ponencias en congresos internacionales, así como 1 capítulo de libro.

Cabe destacar la realización de dos estancias de investigación. La primera, una estancia de investigación virtual desde el 1 de julio de 2021 hasta el 10 de octubre de

2021, en Østfold University College, Departamento de Informática (Halden, Noruega). Esta estancia de investigación se centró en la validación del metamodelo.

La segunda estancia de investigación se realizó del 10 de enero de 2022 al 14 de abril de 2022 en el Departamento de Tecnología de Gráficos por Ordenador de Purdue University (West Lafayette, Indiana, Estados Unidos de América). La investigación estuvo relacionada con aplicaciones de visualización de datos y su dimensión educativa.

Por último, esta tesis doctoral ha recibido financiación del Ministerio de Educación y Formación Profesional de España con una beca FPU (FPU17/03276).

References

- [1] A. Abran, J. W. Moore, P. Bourque, and R. Dupuis, Eds. *Guide to the Software Engineering Body of Knowledge. 2004 Version*. Los Alamitos, CA, USA: IEEE Computer Society, 2004.
- [2] P. Cooper, "Data, information, knowledge and wisdom," *Anaesthesia & Intensive Care Medicine*, vol. 18, no. 1, pp. 55-56, 2017/01/01/ 2017, doi: <https://doi.org/10.1016/j.mpaic.2016.10.006>.
- [3] J. Rowley, "The wisdom hierarchy: representations of the DIKW hierarchy," *Journal of Information Science*, vol. 33, no. 2, pp. 163-180, 2007/04/01 2007, doi: 10.1177/0165551506070706.
- [4] C. Zins, "Conceptual approaches for defining data, information, and knowledge," *Journal of the American Society for Information Science and Technology*, vol. 58, no. 4, pp. 479-493, 2007, doi: <https://doi.org/10.1002/asi.20508>.
- [5] D. Bawden and L. Robinson, "Information overload: An overview," *Oxford encyclopedia of political decision making*, 2020, doi: <https://doi.org/10.17613/n4b2-1w48>.
- [6] D. Bawden, C. Holtham, and N. Courtney, "Perspectives on information overload," in *Aslib proceedings*, 1999: MCB UP Ltd.
- [7] K. Lewin, "Action research and minority problems," *Journal of social issues*, vol. 2, no. 4, pp. 34-46, 1946.
- [8] S. Kemmis, "Point-by-point guide to action research," *Victoria: Deakin University*, 1984.
- [9] R. McTaggart and S. Kemmis, *The action research planner*. Deakin university, 1988.
- [10] K. Schwaber, "Scrum development process," in *Business object design and implementation*: Springer, 1997, pp. 117-134.
- [11] J. W. Creswell and J. D. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications, 2017.
- [12] B. Kitchenham, "Procedures for performing systematic reviews," Keele University Technical Report TR/SE-0401 and NICTA Technical Report 0400011T.1, 2004.
- [13] B. Kitchenham and S. Charters, "Guidelines for performing Systematic Literature Reviews in Software Engineering. Version 2.3," School of Computer Science and Mathematics, Keele University and Department of Computer Science, University of Durham, Technical Report EBSE-2007-01, 2007. [Online]. Available: <https://goo.gl/L1VHcw>
- [14] F. J. García-Peñalvo, "Developing robust state-of-the-art reports: Systematic Literature Reviews," *Education in the Knowledge Society*, vol. 23, 2022, doi: 10.14201/eks.28600.

- [15] N. Anquetil *et al.*, "Traceability for model driven, software product line engineering," in *ECMDA Traceability Workshop Proceedings*, 2008, vol. 12: SINTEF, pp. 77-86.
- [16] A. G. Kleppe, J. Warmer, and W. Bast, "MDA Explained. The Model Driven Architecture: Practice and Promise," ed: Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 2003.
- [17] S. Trujillo, D. Batory, and O. Diaz, "Feature oriented model driven development: A case study for portlets," in *Proceedings of the 29th international conference on Software Engineering*, 2007: IEEE Computer Society, pp. 44-53.
- [18] S. J. Mellor, K. Scott, A. Uhl, and D. Weise, "Model-Driven Architecture," in *Advances in Object-Oriented Information Systems: OOIS 2002 Workshops Montpellier, France, September 2, 2002 Proceedings*, J.-M. Bruel and Z. Bellahsene Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 290-297.
- [19] J. J. López-Fernández, E. Guerra, and J. De Lara, "Assessing the Quality of Meta-models," in *MoDeVVa 2014. Model-Driven Engineering, Verification and Validation. Proceedings of the 11th Workshop on Model-Driven Engineering, Verification and Validation co-located with 17th International Conference on Model Driven Engineering Languages and Systems (MODELS 2014) (Valencia, Spain, September 30, 2014)*, F. Boulanger, M. Famelis, and D. Ratiu, Eds., no. 1235, Aachen, Germany: CEUR-WS.org, 2014, pp. 3-12.
- [20] R. Skjong and B. H. Wentworth, "Expert judgment and risk perception," in *Proceedings of the Eleventh (2001) International Offshore and Polar Engineering Conference (Stavanger, Norway, June 17-22, 2001)*: International Society of Offshore and Polar Engineers, 2001.
- [21] A. Vázquez-Ingelmo, A. García-Holgado, F. J. García-Peñalvo, and R. Therón, "A Meta-Model Integration for Supporting Knowledge Discovery in Specific Domains: A Case Study in Healthcare," *Sensors*, vol. 20, no. 15, p. 4072, 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/15/4072>.
- [22] F. J. García-Peñalvo, "Education in knowledge society: A new PhD programme approach," in *Proceedings of the First International Conference on Technological Ecosystems for Enhancing Multiculturalism (TEEM'13) (Salamanca, Spain, November 14-15, 2013)*, F. J. García-Peñalvo, Ed., New York, NY, USA: ACM, in ACM International Conference Proceeding Series (ICPS), 2013, pp. 575-577, doi: 10.1145/2536536.2536624.
- [23] J. Brooke, "SUS-A quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, no. 194, pp. 4-7, 1996.
- [24] A. Vázquez-Ingelmo, A. García-Holgado, F. J. García-Peñalvo, and R. Therón, "Proof-of-concept of an information visualization classification approach based on their fine-grained features," *Expert Systems*, e12872, 2022, doi: 10.1111/exsy.12872.
- [25] A. V. Pandey, A. Manivannan, O. Nov, M. Satterthwaite, and E. Bertini, "The persuasive power of data visualization," *IEEE transactions on visualization and computer graphics*, vol. 20, no. 12, pp. 2211-2220, 2014.
- [26] A. Vázquez-Ingelmo, F. J. García-Peñalvo, and R. Therón, "Towards a Technological Ecosystem to Provide Information Dashboards as a Service: A Dynamic Proposal for Supplying Dashboards Adapted to Specific Scenarios," *Applied Sciences*, vol. 11, no. 7, art. 3249, 2021, doi: 10.3390/app11073249.
- [27] A. Vázquez-Ingelmo, F. J. García-Peñalvo, and R. Therón, "Domain engineering for generating dashboards to analyze employment and employability in the academic context," in *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturalism*, Salamanca, Spain, October 24th-26th 2018, New York, NY, USA: Association for Computing Machinery, 2018, pp. 896-901, doi: 10.1145/3284179.3284329. [Online]. Available: <https://doi.org/10.1145/3284179.3284329>
- [28] A. Vázquez-Ingelmo, F. J. García-Peñalvo, and R. Therón, "Taking advantage of the software product line paradigm to generate customized user interfaces for decision-making processes: a case study on university employability," *PeerJ Computer Science*, vol. 5, e203, 2019, doi: 10.7717/peerj-cs.203.

- [29] A. Vázquez-Ingelmo, J. Cruz-Benito, F. J. García-Peñalvo, and M. Martín-González, "Scaffolding the OEEU's Data-Driven Ecosystem to Analyze the Employability of Spanish Graduates," in *Global Implications of Emerging Technology Trends*, F. J. García-Peñalvo Ed. Hershey, PA, USA: IGI Global, 2018, pp. 236-255.
- [30] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo, J. Cruz-Benito, and A. Vázquez-Ingelmo, *Barómetro de empleabilidad y empleo universitarios. Edición Máster 2017*. Madrid, España: Observatorio de Empleabilidad y Empleo Universitarios, 2018.
- [31] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo, and J. Cruz Benito, "Empleabilidad de los titulados universitarios en España. Proyecto OEEU," *Education in the Knowledge Society*, vol. 19, no. 1, pp. 21-39, 2018, doi: 10.14201/eks20181912139.
- [32] F. Michavila, J. M. Martínez, M. Martín-González, F. J. García-Peñalvo, and J. Cruz-Benito, *Barómetro de empleabilidad y empleo de los universitarios en España, 2015 (Primer informe de resultados)*. Madrid: Observatorio de Empleabilidad y Empleo Universitarios, 2016.
- [33] P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, and J. G. Conde, "Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support," *Journal of Biomedical Informatics*, vol. 42, no. 2, pp. 377-381, 2009/04/01/ 2009, doi: <https://doi.org/10.1016/j.jbi.2008.08.010>.
- [34] J. W. Tukey, *Exploratory data analysis*. Reading, MA, 1977.
- [35] J. I. Melero-Alegria et al., "SALMANTICOR study. Rationale and design of a population-based study to identify structural heart disease abnormalities: a spatial and machine learning analysis," *BMJ Open*, vol. 9, no. 2, p. e024605, 2019, doi: 10.1136/bmjopen-2018-024605.
- [36] A. Vázquez-Ingelmo et al., "A platform to support the visual analysis of the SALMANTICOR study outcomes: conveying cardiological data to lay users," in *Proceedings TEEM'21. Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality (Barcelona, Spain, October 27th – 29th, 2021)*, New York, NY, USA: Association for Computing Machinery, in ACM International Conference Proceedings Series, 2021, pp. 335–341, doi: 10.1145/3486011.3486471. [Online]. Available: <https://doi.org/10.1145/3486011.3486471>
- [37] S. Berinato, *Good charts: The HBR guide to making smarter, more persuasive data visualizations*. Brighton, MA, USA: Harvard Business Review Press, 2016.
- [38] A. Cairo, *The Functional Art: An introduction to information graphics and visualization*. San Francisco, CA, USA: New Riders, 2012.
- [39] A. Cairo, *The truthful art: Data, charts, and maps for communication*. San Francisco, CA, USA: New Riders, 2016.
- [40] C. O. Wilke, *Fundamentals of data visualization: a primer on making informative and compelling figures*. Sebastopol, CA, USA: O'Reilly Media, 2019.
- [41] A. Vázquez-Ingelmo, A. García-Holgado, F. J. García-Peñalvo, and R. Therón, "Dashboard Meta-Model for Knowledge Management in Technological Ecosystem: A Case Study in Healthcare," *Proceedings*, vol. 31, no. 1, p. 44, 2019. [Online]. Available: <https://www.mdpi.com/2504-3900/31/1/44>.
- [42] A. Vázquez Ingelmo, F. J. García-Peñalvo, and R. Therón Sánchez, "Aggregation Bias: A Proposal to Raise Awareness Regarding Inclusion in Visual Analytics," in *Trends and Innovations in Information Systems and Technologies, WorldCIST 2020*, Á. Rocha, H. Adeli, L. P. Reis, S. Costanzo, I. Orovic, and F. Moreira, Eds., 2020, vol. 3, Cham, Switzerland: Springer Nature, in *Advances in Intelligent Systems and Computing Series* Series, no. AISC 1161, doi: https://doi.org/10.1007/978-3-030-45697-9_40. [Online]. Available: <http://hdl.handle.net/10366/148283>
- [43] A. Vázquez-Ingelmo, F. J. García-Peñalvo, and R. Therón, "Generating Dashboards Using Fine-Grained Components: A Case Study for a PhD Programme," in *Learning and Collaboration Technologies. Design, Experiences. 7th International Conference, LCT 2020, Held*

- as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, P. Zaphiris and A. Ioannou, Eds., July 19–24 2020, no. 12205, Cham, Switzerland: Springer Nature, in Lecture Notes in Computer Science, 2020, pp. 303-314, doi: 10.1007/978-3-030-50513-4_23.
- [44] A. Vázquez-Ingelmo and R. Therón, "Beneficios de la aplicación del paradigma de líneas de productos software para generar dashboards en contextos educativos," *RIED. Revista Iberoamericana de Educación a Distancia*, vol. 23, no. 2, pp. 169-185, 07/01 2020, doi: 10.5944/ried.23.2.26389.
- [45] A. Vázquez Ingelmo, A. García-Holgado, H. Hernández-Payo, F. J. García-Peñalvo, and R. Therón Sánchez, "Following up the progress of doctoral students and advisors' workload through data visualizations: a case study in a PhD program," *Proceedings of LASI-SPAIN 2021. Learning Analytics Summer Institute Spain 2021: Learning Analytics in times of COVID-19: Opportunity from crisis (Barcelona, Spain, July 7-9, 2021). CEUR Workshop Proceedings Series*, 2021. [Online]. Available: <http://ceur-ws.org/Vol-3029/paper06.pdf>.
- [46] A. Vázquez-Ingelmo *et al.*, "A platform for management and visualization of medical data and medical imaging," in *Proceedings TEEM'20. Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality (Salamanca, Spain, October 21st - 23rd, 2020)*, F. J. García-Peñalvo, Ed., New York, NY, USA: Association for Computing Machinery, in ACM International Conference Proceedings Series, 2020, pp. 518–522, doi: 10.1145/3434780.3436652. [Online]. Available: <https://doi.org/10.1145/3434780.3436652>
- [47] F. J. García-Peñalvo *et al.*, "Application of Artificial Intelligence Algorithms Within the Medical Context for Non-Specialized Users: the CARTIER-IA Platform," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 6, pp. 46-53, 2021, doi: 10.9781/ijimai.2021.05.005.
- [48] A. García-Holgado *et al.*, "User-Centered Design Approach for a Machine Learning Platform for Medical Purpose," in *HCI-COLLAB 2021, Sao Paulo, Brazil, 8-10, September 2021*, Cham, Switzerland: Springer International Publishing, in Human-Computer Interaction, pp. 237-249, doi: 10.1007/978-3-030-92325-9_18.
- [49] A. Vázquez-Ingelmo *et al.*, "Bringing machine learning closer to non-experts: proposal of a user-friendly machine learning tool in the healthcare domain," in *Proceedings TEEM'21. Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality (Barcelona, Spain, October 27th – 29th, 2021)*, New York, NY, USA: Association for Computing Machinery, in ACM International Conference Proceedings Series, 2021, pp. 324–329, doi: 10.1145/3486011.3486469. [Online]. Available: <https://doi.org/10.1145/3486011.3486469>
- [50] R. Morgan, G. Grossmann, M. Schrefl, M. Stumptner, and T. Payne, "VizDSL: A Visual DSL for Interactive Information Visualization," Cham, 2018: Springer International Publishing, in *Advanced Information Systems Engineering*, pp. 440-455.
- [51] I. Logre, S. Mosser, P. Collet, and M. Riveill, "Sensor data visualisation: a composition-based approach to support domain variability," in *European Conference on Modelling Foundations and Applications*, York, United Kingdom, R. J. Cabot J., Ed., 2014, vol. 8569: Springer, in *Modelling Foundations and Applications. ECMFA 2014*, pp. 101-116, doi: https://doi.org/10.1007/978-3-319-09195-2_7.
- [52] T. Palpanas, P. Chowdhary, G. Mihaila, and F. Pinel, "Integrated model-driven dashboard development," *Information Systems Frontiers*, vol. 9, no. 2-3, pp. 195-208, 2007, doi: <https://doi.org/10.1007/s10796-007-9032-9>.
- [53] M. Correll, E. Bertini, and S. Franconeri, "Truncating the y-axis: Threat or menace?," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA, 2020: ACM, pp. 1-12, doi: 10.1145/3313831.3376222.
- [54] S. L. Franconeri, L. M. Padilla, P. Shah, J. M. Zacks, and J. Hullman, "The Science of Visual Data Communication: What Works," *Psychological Science in the Public Interest*, vol. 22, no. 3, pp. 110-161, 2021, doi: 10.1177/15291006211051956.
- [55] S. Few, *Information dashboard design*. Sebastopol, CA, USA: O'Reilly Media, 2006.

- [56] C. Plaisant, "The challenge of information visualization evaluation," in *Proceedings of the working conference on Advanced visual interfaces*, 2004: ACM, pp. 109-116.
- [57] A. Sarikaya, M. Correll, L. Bartram, M. Tory, and D. Fisher, "What Do We Talk About When We Talk About Dashboards?," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 682-692, 2019, doi: 10.1109/TVCG.2018.2864903.
- [58] A. Vázquez Ingelmo, A. García-Holgado, F. J. García-Peñalvo, and R. Therón Sánchez, "A Meta-modeling Approach to Take into Account Data Domain Characteristics and Relationships in Information Visualizations," in *Trends and Innovations in Information Systems and Technologies, WorldCIST 2021*, Azores, Portugal, Á. Rocha, H. Adeli, G. Dzemyda, F. Moreira, and A. M. R. Correia, Eds., 2021, vol. 2, Cham, Switzerland: Springer Nature, in *Advances in Intelligent Systems and Computing Series*, no. 1366, 2021, pp. 570-580, doi: 10.1007/978-3-030-72651-5_54. [Online]. Available: <http://hdl.handle.net/10366/145626>
- [59] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer, "Vega-Lite: A Grammar of Interactive Graphics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 341-350, 2017, doi: 10.1109/TVCG.2016.2599030.
- [60] A. Vázquez-Ingelmo, F. J. García-Peñalvo, R. Therón, D. Amo Filvà, and D. Fonseca Escudero, "Connecting domain-specific features to source code: towards the automatization of dashboard generation," *Cluster Computing*, vol. 23, no. 3, pp. 1803-1816, 2020/09/01 2020, doi: 10.1007/s10586-019-03012-1.
- [61] A. Vázquez-Ingelmo, F. J. García-Peñalvo, R. Therón, and A. García-Holgado, "Specifying information dashboards' interactive features through meta-model instantiation," in *LASI-SPAIN 2020. Learning Analytics Summer Institute Spain 2020: Learning Analytics. Time for Adoption?*, Valladolid, Spain, A. Martínez-Monés, A. Álvarez, M. Caeiro-Rodríguez, and Y. Dimitriadis, Eds., June 15-16 2020, Aachen, Germany: CEUR-WS.org, in *CEUR Workshop Proceedings Series*, no. 2671, pp. 47-59.
- [62] M. Bostock, V. Ogievetsky, and J. Heer, "D³ Data-Driven Documents," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2301-2309, 2011, doi: 10.1109/TVCG.2011.185.
- [63] A. Vázquez-Ingelmo, F. J. García-Peñalvo, and R. Therón, "Addressing Fine-Grained Variability in User-Centered Software Product Lines: A Case Study on Dashboards," in *Knowledge in Information Systems and Technologies*, Á. Rocha, H. Adeli, L. P. Reis, and S. Costanzo, Eds., 2019, vol. 1, Switzerland: Springer Nature, in *Advances in Intelligent Systems and Computing*, no. AISC 930, 2019, pp. 855-864, doi: 10.1007/978-3-030-16181-1_80.
- [64] J. B. Warmer and A. G. Kleppe, *The object constraint language: getting your models ready for MDA*. Boston, MA, USA: Addison-Wesley Professional, 2003.