Vázquez-Ingelmo, A., García-Peñalvo, F. J., & Therón, R. (2020). Generating Dashboards Using Fine-Grained Components: A Case Study for a PhD Programmes. In P. Zaphiris & A. Ioannou (Eds.), Learning and Collaboration Technologies. Design, Experiences. 7th International Conference, LCT 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part I (pp. 303-314). Springer Nature. https://doi.org/10.1007/978-3-030-50513-4_23

Generating Dashboards Using Fine-Grained Components: A Case Study for a PhD Programme

Andrea Vázquez-Ingelmo[1(✉)], Francisco J. García-Peñalvo[1], and Roberto Therón[1,2]

[1] GRIAL Research Group, Computer Sciences Department, Research Institute for Educational Sciences, University of Salamanca, Salamanca, Spain
{andreavazquez, fgarcia, theron}@usal.es
[2] VisUSAL Research Group, University of Salamanca, Salamanca, Spain

Abstract. Developing dashboards is a complex domain, especially when several stakeholders are involved; while some users could demand certain indicators, other users could demand specific visualizations or design features. Creating individual dashboards for each potential need would consume several resources and time, being an unfeasible approach. Also, user requirements must be thoroughly analyzed to understand their goals regarding the data to be explored, and other characteristics that could affect their user experience. All these necessities ask for a paradigm to foster reusability not only at development level but also at knowledge level. Some methodologies, like the Software Product Line paradigm, leverage domain knowledge and apply it to create a series of assets that can be composed, parameterized, or combined to obtain fully functional systems. This work presents an application of the SPL paradigm to the domain of information dashboards, with the goal of reducing their development time and increasing their effectiveness and user experience. Different dashboard configurations have been suggested to test the proposed approach in the context of the Education in the Knowledge Society PhD programme of the University of Salamanca.

Keywords: Domain engineering · SPL · Information dashboards · Information systems · Educational dashboards

1 Introduction

Data visualization is gaining relevance as a method to understand and generate knowledge [1] from datasets. However, the exponential growth in data generation due to the widespread of data-driven technologies [2] asks for new methodologies to support informed decision-making even if relying on complex and large datasets.

Visual analytics [3, 4] is a popular methodology to foster the comprehension of large quantities of data. This field focuses on analytical reasoning and how interactive tools and visual interfaces can support it.

Visualizing data and generating knowledge from them through information visualizations or information dashboards [5–7] require proper representations, that is, appropriate encodings and visual metaphors to convey the information contained (and mainly hidden) within large datasets.

However, although practitioners can follow general guidelines to design information visualizations [8], the process of building these tools is costly, because it is important not only to take into account but also to deeply understand the context in which these tools will be employed.

The audience and the context of application are essential to build an effective dashboard. One of the challenges in this field is the adaptation of visualizations and dashboards based on their context [9]. But the adaptation of these tools is complex; it requires several resources to design and implement an adapted version of the same tool.

That is why some approaches that leverage reusability are extremely useful in these domains [10]. Especially, approaches like the Software Product Line (SPL) [11, 12] paradigm and the Model-Driven Development (MDD) [13] foster the reusability of components and knowledge by abstracting common features of a concrete domain.

A dashboard meta-model has been previously developed [14–16] to account for not only for technical features of these tools but also to account for their users. Involving the end-users' characteristics and goals regarding information is crucial to obtain a dashboard adapted to the context and audience.

This work focuses on the application of the developed dashboard meta-model to generate information dashboards in the context of a PhD programme. The goal of this application is to test the suitability of a generative pipeline to tailor information dashboards in a real-world scenario.

A requirement elicitation process has been carried to identify the information requirements and goals of each involved user profile to achieve the mentioned goal. With this information, it has been possible to map the goals into concrete dashboard features.

The rest of this paper is organized as follows. The next section contextualizes the PhD Programme on Education in the Knowledge Society. Section 3 explains the different materials and methods employed to carry out this work. Section 4 presents the results of the requirement elicitation process, as well as the results regarding the generation of customized dashboards based on the previously identified requirements. Finally, Sect. 5 discusses the obtained results, concluding with Sect. 6, in which the conclusions derived from this work are described.

2 PhD Programme on Education in the Knowledge Society

The PhD Program on Education in the Knowledge Society emerges from the Research Institute for Educational Sciences (IUCE – http://iuce.usal.es) at the University of Salamanca (Spain), following the Spanish Royal Decree 99/2011. The primary motivation behind this PhD Program is to feature the teaching-learning forms as main impetus of the Knowledge Society, so as to examine and create new information about the learning as a key component of the Knowledge Society, including both the Social Sciences considers and the new mechanical advances yet inside a synergic and harmonious methodology [17, 18].

The focus of this program is completely interdisciplinary, mainly upheld by the Recognized Research Groups at the University of Salamanca: GRIAL (http://grial.usal.es) [19], GITE (http://gite213.usal.es), OCA (http://campus.usal.es/∼oca/), VISUALMED (http://visualmed.usal.es), Robotics and Society Group (http://gro.usal.es) y E-LECTRA (http://electra.usal.es).

The PhD Program gives a context where knowledge generation and its visibility and dissemination are primary objectives. In order to reach them, the scientific knowledge management of the Programme is supported by a technological ecosystem that fusion methodology and technology to provide tools for both PhD candidates and researchers. The fundamental parts of the ecosystem are the PhD portal (http://knowledgesociety.usal.es) and a set of social instruments, such as SlideShare to share presentations (http://www.slideshare.net/knowedgesociety) or a YouTube channel to share seminars and meetings (http://youtube.com/knowledgesocietyphd).

3 Materials and Methods

3.1 Dashboard Meta-model

Meta-models are the starting point of the Model-Driven Development paradigm [13]. These artifacts allow the high-level definition of entities and relationships, without focusing on the technical details of the domain.

In this case, a dashboard meta-model has been developed to capture the abstract characteristics that define dashboards. As can be seen in Fig. 1, the user is also part of the meta-model because they are the drivers of the design process and the final consumers of the information displayed through these tools, so their characteristics must be accounted for to deliver a properly designed dashboard.

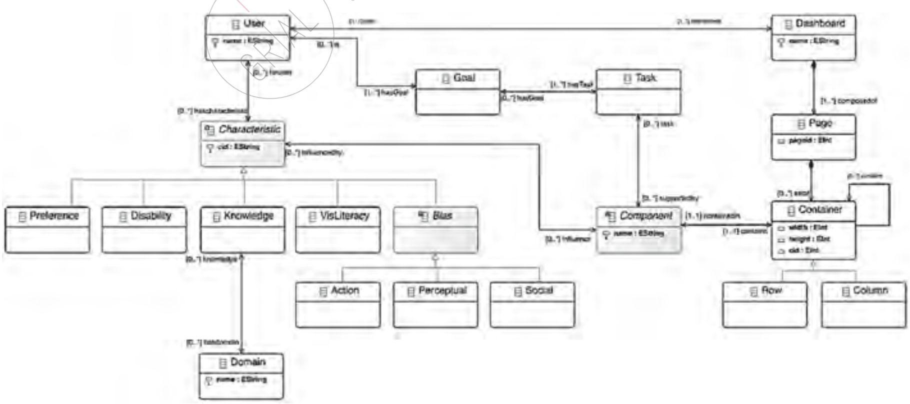

Fig. 1. Overview of the dashboard meta-model. The component definition has been omitted for legibility reasons, but it can be consulted at [20].

The entire meta-model design process is out of the scope of this paper, but it can be consulted in [15, 16].

3.2 Feature Model

The goal of software product lines (SPLs) is to derive final products from different core assets and software components [11, 12], and to allow the adaptation of products to match specific requirements without consuming significant time and resources.

This paradigm has two main phases: the domain engineering phase and the application engineering phase. The mentioned core assets are developed during the first phase by identifying commonalities and variability points within the products' domain. This phase supports the implementation of base components that hold the common logic of the domain as well as the variability points.

Variability points are sections of the logic that can be modified, parameterized, or configured to change the functionality of the products during the second phase (application engineering phase). In SPLs, products' functionalities are seen as features. Stakeholders can select different features for their products to be injected in the base logic, thus obtaining personalized systems that fit their specific requirements. These personalized products are built by reusing and assembling the core assets developed during the domain engineering phase, which reduces the time-to-market of tailored software systems as well as their development efforts.

The features of software product lines can be specified by different means. One of the most popular methods to identify and arrange SPL features are feature models [21]. Feature models are useful for documentation purposes, but also essential artifacts for guiding the development process of the product line. These models provide a skeleton for designing the core assets and for materializing the variability points at code level.

In this specific domain (i.e., the dashboards domain), the feature model captures the dashboards' visualizations' low-level characteristics, corresponding with visual encodings, visual marks, etc. [8]. Figure 2 shows an excerpt from the feature model employed to design the core assets for the dashboards product line.

The hierarchical structure of the feature model enables the definition of high-level characteristics and their refinement to reach finer-grained features [22].

3.3 Requirements Elicitation

The requirements elicitation phase is an essential phase in any software development process. In this case, it is crucial to understand the context of application (the PhD Programme on Education in the Knowledge Society) to generate adapted information dashboards.

Firstly, there is a variety of profiles to take into account involved in the programme: from PhD candidates to their advisors and managers. These users will be the drivers of the dashboards' design and generation processes because the outcome must match their information needs and functional requirements.

A questionnaire is proposed to gather the information needs of each involved profile. However, to ensure the proper design of this questionnaire, an initial interview was conducted. The goal of this interview was to collect information regarding the business processes of the PhD programme, as well as the available data and their structure.

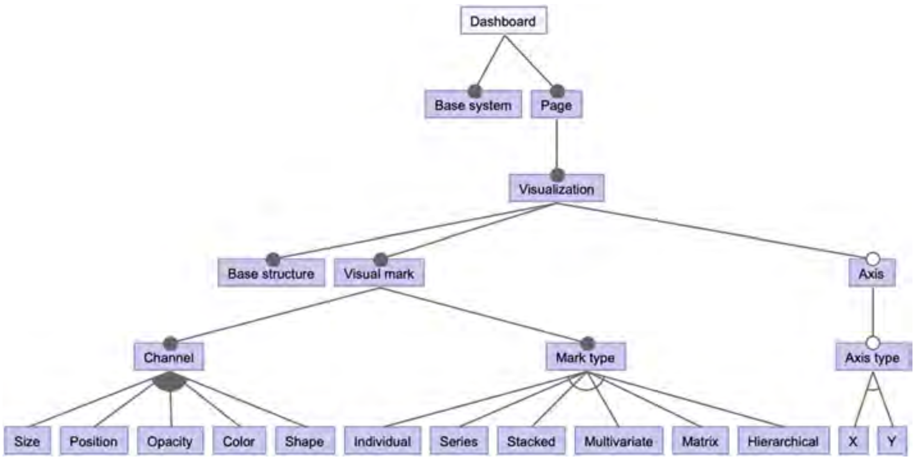

Fig. 2. Simplified feature model of a dashboards product line.

A few questions were asked to a member of the PhD programme's quality committee to understand which data could be displayed in a potential information dashboard. The PhD portal gathers information about the milestones achieved by the members of the PhD programme. For example, data about publications, research visits, conferences, reports, awards, seminars, patents, etc., are collected in a semi-structured way.

On the other hand, the PhD programme on Education in the Knowledge Society has a series of minimum requirements to be able to defend the thesis or to obtain some recognitions like an international PhD mention. These requirements are covered in [23].

In addition, other aspects of the PhD candidates are stored, such as their academic year, associated research lines, PhD advisors, scholarships, contracts, PhD modality (full- or part-time), etc.

Finally, the users of the portal can have different roles: students (including current PhD candidates or post-doctoral students), PhD advisors, and managers.

This interview supported not only the development of the questionnaire, but also the comprehension of the variables available to be displayed in potential PhD information dashboards.

3.4 Instrumentation

Once the interview was carried out, an instrument to collect information regarding the users' information requirements was designed. First, some demographic variables are collected to contextualize the sample: age, gender, and birthplace.

The next section focuses on the collection of the user situation within the PhD programme: role, research lines, PhD modality, and academic year (in the case of PhD students) and the number of PhD thesis being directed (in the case of PhD advisors). Questions regarding the usage of the PhD portal were also included in this section to understand how users employ this platform.

Finally, the last section included questions regarding users' past experiences with information visualization and regarding the users' information requirements for a hypothetical PhD programme dashboard.

The instrument to collect these requirements was implemented employing a customized version of LimeSurvey (https://www.limesurvey.org), an Open Source online statistical survey web application. The instrument was applied in Spanish because the PhD portal users were Spanish speakers.

4 Results

4.1 Requirements Elicitation Results

The questionnaire was answered by 24 participants. The distribution of the participants in terms of their role is summarized in Fig. 2. All the participants stated that they employ the PhD portal, and the majority had experience using dashboards or visualization tools (70.83%).

Regarding how the PhD students save their milestones, an interesting pattern of answers were found. The information regarding their milestones is scattered among different research profiles (Mendeley, Scopus, WoS, Google Scholar, etc.) and/or the PhD programme portal, in personal Microsoft Word files created by themselves and even some students state that they don't save their milestones at all. These aspects are summarized in Fig. 3.

Finally, the collected requirements for a hypothetical personal dashboard within the PhD programme portal suggest that users have different information goals. On the one hand, PhD advisors were very interested in obtaining information regarding the publications (count, typology, and the number of citations), conference attendance/participation, and research profiles of their PhD students (Fig. 4).

In relation to managers or academic committee members, in addition to the information related to the progress of all the PhD students and their deadlines, some pointed out as an information requirement the number of PhD students associated to each research group, as well as metrics regarding the interactions between their PhD advisors. Another participant asked for information about publications by research line and the number of publications during specific periods of time.

As will be discussed, the most diverse requirements were found within the PhD students' answers. The majority share the necessity of displaying information related to their PhD progress and their achieved milestones (publications, conferences, seminars, etc.), but other requirements were also mentioned:

- Remaining required activities.
- Recommended activities vs. required activities.
- Comparison with other PhD students.
- Distribution of activities/milestones (publications, seminars, etc.) by type.
- Percentage of progress based on the requirements of the PhD programme.
- Status of each research milestone uploaded to the portal.
- Deadlines and enrollment dates.

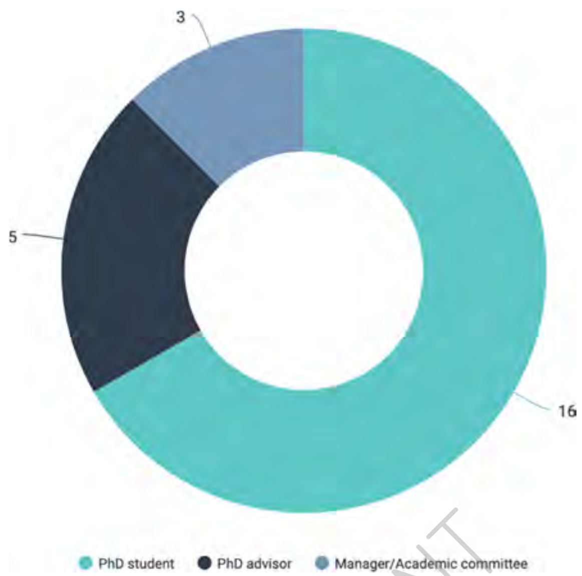

Fig. 3. Distribution of participants regarding their role in the PhD programme (n = 24).

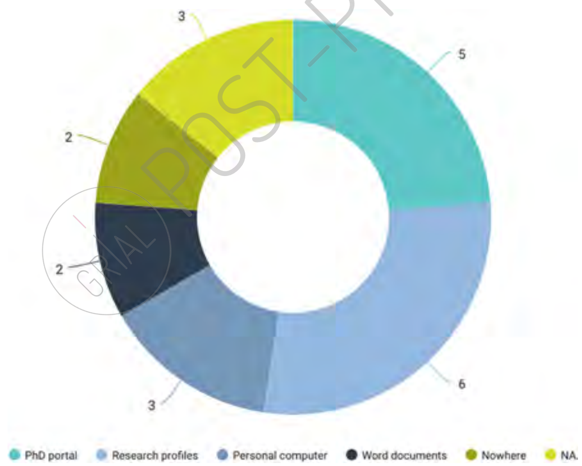

Fig. 4. Distribution of the methods employed by PhD students to save their progress/milestones obtained within the PhD programme (n = 16; some participants pointed out more than one method to save their milestones).

4.2 Dashboard Generation Results

Based on some of the requirements identified through the questionnaire, three dashboard prototypes for each major PhD role have been generated by using the software product line paradigm. These proposals can be seen in Figs. 5, 6, and 7.

For example, based on the collected requirements, a student dashboard might be composed of views that show their achievements classified by type and reference marks to show if they have reached the minimum requirements of the PhD programme. On the other hand, these achievements can also be displayed through time, allowing students to inspect their productivity (Fig. 5).

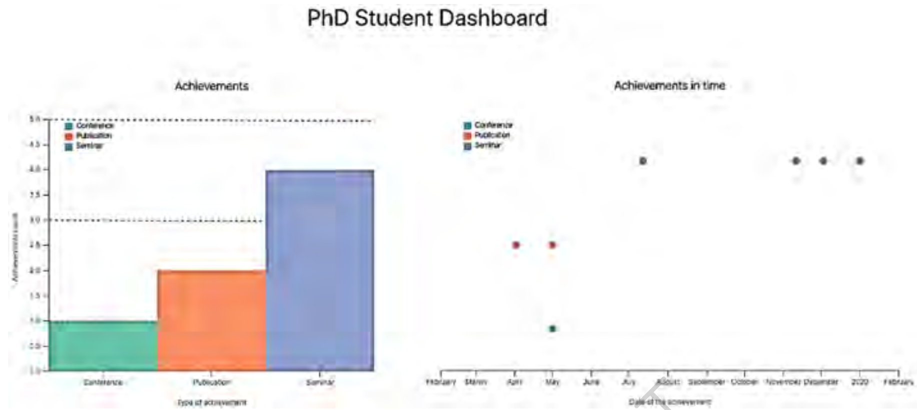

Fig. 5. PhD student dashboard proposal.

On the other hand, the PhD advisor dashboard (Fig. 6) could show each student individual progress as well as a general overview through the time of the advisors' PhD candidates. The flexibility of this approach allow the adaptation of this view to the specific number of students of each PhD advisor.

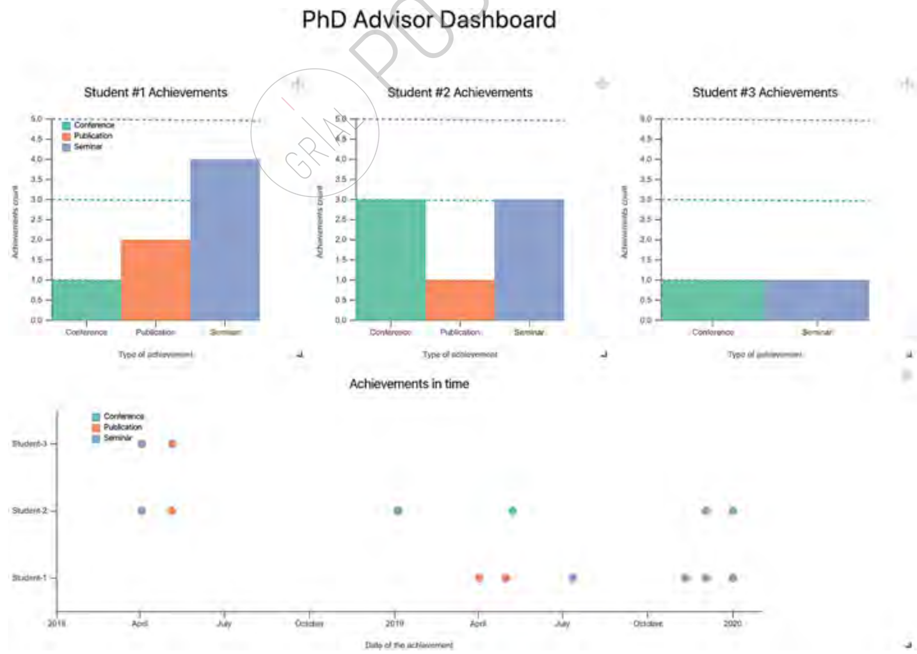

Fig. 6. PhD advisor dashboard proposal.

Finally, the PhD manager dashboard (Fig. 7) could hold information regarding the workload (in terms of directed thesis) of each PhD advisor and research group involved in the programme. Also, following the collected requirements, this dashboard can also have a view displaying the achieved publications through time by research line.

Fig. 7. PhD manager dashboard proposal.

5 Discussion

The questionnaire yielded very interesting results and confirmed that dashboards must take into account different user roles and user requirements. There is not a "universal" dashboard that fits for any user; that is why dashboards should be tailored depending on their context.

Three main roles have been identified within the PhD programme through an interview with a member of the quality committee: PhD students, PhD advisors, and the programme managers. The questionnaire asked different questions depending on the role of the participant. While students asked for information about their progress or publications, PhD advisors wanted to gain insights regarding their own students' process. On the other hand, managers mostly asked to gain overall insights regarding all the students enrolled in the programme.

The greatest diversity of requirements was found among PhD students; although these users share the same role, significant differences were encountered regarding their information goals. In this work, a single dashboard proposal was presented for PhD students, but different dashboards could be developed for this role; for example, one user asked for comparisons with other students' progress.

While this could be a beneficial feature for this user, it could be counterproductive for other users (for example, showing their progress compared with other pre-doctoral users could have a negative impact on some students' motivation [24]).

The three dashboards were generated using a DSL based on the feature tree and meta-model presented in Sect. 3. This approach not only improved the development time of the dashboards, as the tools were automatically generated through code templates (the core assets of the SPL, in this case) but also improved the requirements management of the users.

A requirement file using the DSL could be maintained for each user, supporting fine-grained management of requirements and controlling changes in a straightforward manner.

The main effort, however, is still in the requirements' elicitation process. If users' information requirements are taken for granted or are given low relevance, the approach would be useless, because the output would be an inefficient and ineffective dashboard. That is why current research is focused on how to automatize or infer users' requirements from their characteristics or from the data to be displayed [25, 26].

This approach is yet to be integrated and tested within the Education in the Knowledge Society PhD programme context, but the viability results of using this approach to generate information dashboards seem promising.

6 Conclusions

The SPL paradigm has been applied to generate dashboards in an educative context; specifically, in a PhD programme context.

The diversity of information goals and requirements ask for flexible design and development process to build these tools in order to accelerate the delivery time and leverage the information displayed.

The application of the SPL paradigm requires initial efforts to identify the domain's abstract features and to develop the different core assets, but relying on these artifacts have subsequently reduced the required time to develop tailored dashboards.

The presented dashboard generative process has been driven by the PhD programme involved actors and their information goals, which were gathered through an online questionnaire.

Future research lines will involve the improvement of the dashboard product line to automatize the generation process, as well as in-depth user testing to validate the utility of this approach to integrate tailored dashboards in learning contexts.

Acknowledgments. This research work has been supported by the Spanish *Ministry of Education and Vocational Training* under an FPU fellowship (FPU17/03276). This work has been partially funded by the Spanish Government Ministry of Economy and Competitiveness throughout the DEFINES project (Ref. TIN2016-80172-R).

References

1. Zeleny, M.: Management support systems: Towards integrated knowledge management. Hum. Syst. Manag. **7**, 59–70 (1987)
2. Patil, D., Mason, H.: Data Driven. O'Reilly Media Inc., Sebastopol (2015)
3. Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., Melançon, G.: Visual analytics: definition, process, and challenges. In: Kerren, A., Stasko, J.T., Fekete, J.-D., North, C. (eds.) Information Visualization. LNCS, vol. 4950, pp. 154–175. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-70956-5_7

4. Keim, D., Kohlhammer, J., Ellis, G., Mansmann, F.: Mastering the Information Age Solving Problems with Visual Analytics. Eurographics Association, Goslar (2010)
5. Tufte, E., Graves-Morris, P.: The Visual Display of Quantitative Information. Graphics Press, Cheshire (1983, 2014)
6. Ware, C.: Information Visualization: Perception for Design. Elsevier, Amsterdam (2012)
7. Few, S.: Information Dashboard Design. O'Reilly Media Inc., Sebastopol (2006)
8. Munzner, T.: Visualization Analysis and Design. AK Peters/CRC Press, Boca Raton (2014)
9. Sarikaya, A., Correll, M., Bartram, L., Tory, M., Fisher, D.: What do we talk about when we talk about dashboards? IEEE Trans. Vis. Comput. Graph. **25**, 682–692 (2018)
10. Vázquez-Ingelmo, A., García-Peñalvo, F.J., Therón, R.: Information dashboards and tailoring capabilities - a systematic literature review. IEEE Access **7**, 109673–109688 (2019)
11. Clements, P., Northrop, L.: Software Product Lines. Addison-Wesley, Boston (2002)
12. Gomaa, H.: Designing Software Product Lines with UML: From Use Cases to Pattern-Based Software Architectures. Addison Wesley Longman Publishing Co., Inc., Boston (2004)
13. Kleppe, A.G., Warmer, J., Bast, W.: MDA Explained. The Model Driven Architecture: Practice and Promise. Addison-Wesley Longman Publishing Co., Inc., Boston (2003)
14. Vázquez-Ingelmo, A., García-Holgado, A., García-Peñalvo, F.J., Therón, R.: Dashboard meta-model for knowledge management in technological ecosystem: a case study in healthcare. In: UCAmI 2019. MDPI (2019)
15. Vázquez-Ingelmo, A., García-Peñalvo, F.J., Therón, R.: Capturing high-level requirements of information dashboards' components through meta-modeling. In: 7th International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2019), León, Spain (2019)
16. Vázquez-Ingelmo, A., García-Peñalvo, F.J., Therón, R., Conde González, M.Á.: Extending a dashboard meta-model to account for users' characteristics and goals for enhancing personalization. In: Learning Analytics Summer Institute (LASI) Spain 2019, Vigo, Spain (2019)
17. García-Holgado, A., García-Peñalvo, F.J., Rodríguez Conde, M.J.: Definition of a technological ecosystem for scientific knowledge management in a PhD programme. In: Proceedings of the Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM 2015), Porto, Portugal, 7–9 October 2015, pp. 695–700. ACM, New York (2015)
18. García-Peñalvo, F.J.: Engineering contributions to a knowledge society multicultural perspective. IEEE Revista Iberoamericana de Tecnologías del Aprendizaje (IEEE RITA) **10**, 17–18 (2015)
19. García-Peñalvo, F.J., Rodríguez-Conde, M.J., Therón, R., García-Holgado, A., Martínez-Abad, F., Benito-Santos, A.: Grupo GRIAL. IE Comunicaciones. Revista Iberoamericana de Informática Educativa, 33–48 (2019)
20. Vázquez-Ingelmo, A., García-Holgado, A., García-Peñalvo, F.J., Therón, R.: Ecore version of the metamodel for information dashboards (2019). https://doi.org/10.5281/zenodo.3561320
21. Kang, K.C., Cohen, S.G., Hess, J.A., Novak, W.E., Peterson, A.S.: Feature-oriented domain analysis (FODA) feasibility study. Carnegie-Mellon University, Software Engineering Institute (1990)
22. Vázquez-Ingelmo, A., García-Peñalvo, F.J., Therón, R.: Addressing fine-grained variability in user-centered software product lines: a case study on dashboards. In: Rocha, Á., Adeli, H., Reis, L.P., Costanzo, S. (eds.) WorldCIST'19 2019. AISC, vol. 930, pp. 855–864. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-16181-1_80
23. Programa de Doctorado Formación en la Sociedad del Conocimiento. https://www.slideshare.net/knowedgesociety/kickoff-phd-programme-20192020

24. Teasley, S.D.: Student facing dashboards: One size fits all? Technol. Knowl. Learn. **22**, 377–384 (2017). https://doi.org/10.1007/s10758-017-9314-3
25. Hu, K., Bakker, M.A., Li, S., Kraska, T., Hidalgo, C.: VizML: a machine learning approach to visualization recommendation. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, p. 128. ACM (2019)
26. Dibia, V., Demiralp, Ç.: Data2Vis: automatic generation of data visualizations using sequence to sequence recurrent neural networks. IEEE Comput. Graph. Appl. **39**, 33–46 (2019)