

# A proposal to measure the understanding of data visualization elements in visual analytics applications

Andrea Vázquez-Ingelmo<sup>1</sup>, Francisco José García-Peñalvo<sup>1</sup>, Roberto Therón<sup>1,2</sup>, Vetrica Byrd<sup>3</sup> and Jorge D. Camba<sup>3</sup>

<sup>1</sup> GRIAL Research Group, University of Salamanca, Paseo de Canalejas 169, Salamanca, Spain

<sup>2</sup> VisUSAL, Computer Science Department, Universidad de Salamanca, Salamanca, Spain

<sup>3</sup> Computer Graphics Technology, Purdue University, West Lafayette, IN, USA

## Abstract

Data visualizations and information dashboards are useful but complex tools. They must be fully understood to draw proper insights and to avoid misleading conclusions. However, several elements and factors are involved in this domain, which makes it difficult to learn. In previous works, we proposed a meta-model to capture the primitive elements that compose visualizations and dashboards. This meta-model has served as a framework for conducting data visualization research, but also to develop a graphical tool for generating data visualizations and dashboards. This tool (namely MetaViz) enables users to create data visualizations through fine-grained components based on the entities represented in the meta-model. The main goal of the system is to provide a learning experience in which users can freely add and configure elements to understand how they influence the final display. This work describes work in progress to validate the pedagogical value of MetaViz in terms of the understanding of data visualization concepts.

## Keywords

Data visualization, Information Dashboards, Understandability, Instrument.

## 1. Introduction

Conveying information effectively has become a challenge over the years; the increase of data in terms of quantity and complexity has hampered their analysis and communication to different audiences. However, along with the increase in quantity and complexity, there is also an increase in relevance of data within informed decision-making processes [1]. Using a data-driven philosophy to make decisions allows better communication, measurability, accountability and more objective approaches to tackle new problems [2].

Visual tools such as data visualizations and information dashboards are being widely used to address the extraction of knowledge from raw data [3]. Although these mechanisms provide means to convey data easier [4-7], audiences must understand the employed visual metaphors [8]. Many factors are involved at this point: potential biases [9], beliefs [10-12], and even -purposely or not- misleading designs [13, 14].

Regarding the latter, several platforms have been developed to facilitate the process of designing and implementing data visualizations. Systems like Tableau (<https://www.tableau.com/>), Microsoft Excel (<https://www.microsoft.com/microsoft-365/excel>), Power BI (<https://powerbi.microsoft.com/>), etc., provide graphical interfaces that allow users with no experience in programming to create data visualizations and even assist them in the design process to choose the best encodings. However, it is crucial to understand and account for every element involved in understanding data visualizations to deliver effective and clear displays of information [5, 6, 15, 16].

---

Learning Analytics Summer Institute Spain (LASI Spain) 2022, June 20–21, 2022, Salamanca, Spain

EMAIL: [andrevazquez@usal.es](mailto:andrevazquez@usal.es) (A. 1); [fgarcia@usal.es](mailto:fgarcia@usal.es) (A. 2); [theron@usal.es](mailto:theron@usal.es) (A. 3); [vbyrd@purdue.edu](mailto:vbyrd@purdue.edu) (A. 4); [jdorribo@purdue.edu](mailto:jdorribo@purdue.edu) (A. 5)

© 2022 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

In this context, we propose MetaViz, a new system designed and implemented through a model-driven approach. A dashboard meta-model was developed in previous works [17-21] to provide a decomposition of the dashboards and data visualizations domain and to fuel the model-driven development.

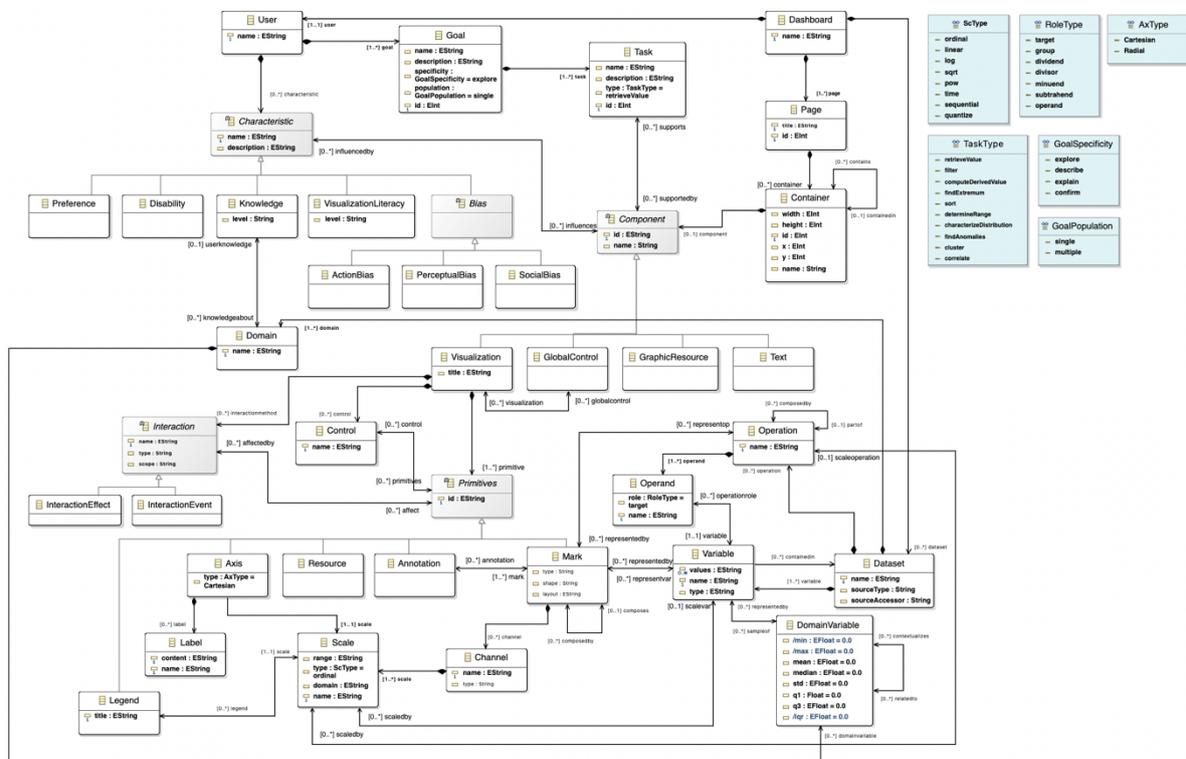
The meta-model represents the most primitive elements of data visualizations (tangible and conceptual entities) and how they influence one another. In this sense, MetaViz takes advantage of this artifact and provides a graphical interface to generate data visualizations and dashboards through the composition of basic shapes and configurations. The aims of this research are two-fold: 1) to generate data visualizations through a usable interface and 2) to provide a learning experience in which users can learn how different configurations and visual elements influence how we see data and our reached insights.

This paper describes a work-in-progress for a pilot study to measure the understanding of the elements involved in the design of data visualizations and information dashboards. To do so, we designed a questionnaire focused on determining the level of awareness of users of Tableau and MetaViz in terms of the visual metaphors and configurations they employed to create a data visualization.

The rest of this work is organized as follows. Section 2 outlines the dashboard meta-model and the MetaViz platform. Section 3 details the study procedure and the developed questionnaire. Finally, section 4 discusses the impressions of the pilot study.

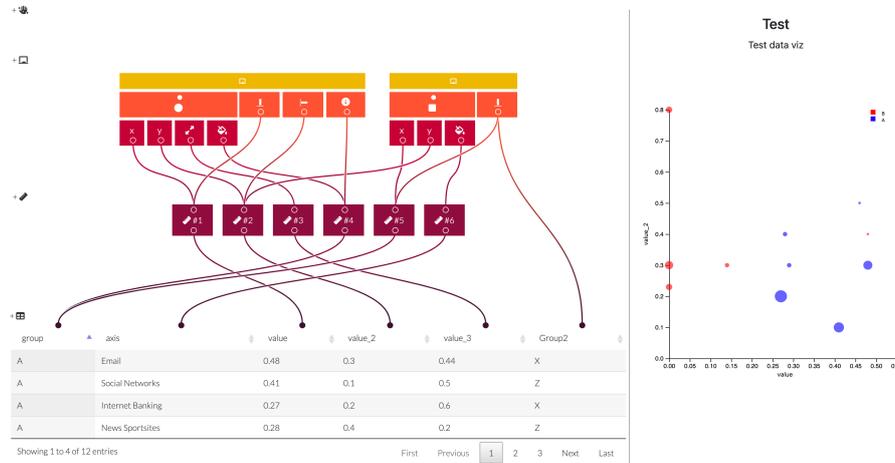
## 2. Background

The foundation of this work is a meta-model which has been subject to several improvements and modifications through domain engineering [22, 23] to capture the most relevant elements and factors in the domain of data visualizations and information dashboards. The current version of the dashboard meta-model is shown in Fig 1. The user [21], the layout and visual components [20] are represented along with concepts related to data, such as its domain and context [17].



**Figure 1:** The current version of the dashboard meta-model. A high-resolution version of the image can be consulted at <https://zenodo.org/record/7037624>.

Starting from this meta-model and following the model-driven development [24, 25] and software product lines [26, 27] approaches, we implemented a graphical meta-model instantiator, i.e., a platform that allows users to create models following the meta-model abstract entities. MetaViz also enables the generation of functional, real-world visualizations following the instantiated models' specifications. MetaViz's interface is shown in Fig 2. The platform is available at <https://metaviz.grial.eu/>.



**Figure 2:** MetaViz's interface. A high-resolution version of the image can be viewed at <https://bit.ly/3G3JUPK>

### 3. Study procedure

MetaViz is focused on developing factual, conceptual, and procedural visualization knowledge [28]. The study focuses on the learning experience that the MetaViz system offers. MetaViz's strengths are flexibility and the fine-grained configuration of every element. In this sense, we want to test if the system's interface and architecture improve attention during the design process of data visualizations.

While powerful tools like Tableau offer assisted and straightforward implementation of data visualizations, MetaViz forces the user to be aware of the elements and underlying mappings they define when creating a visualization.

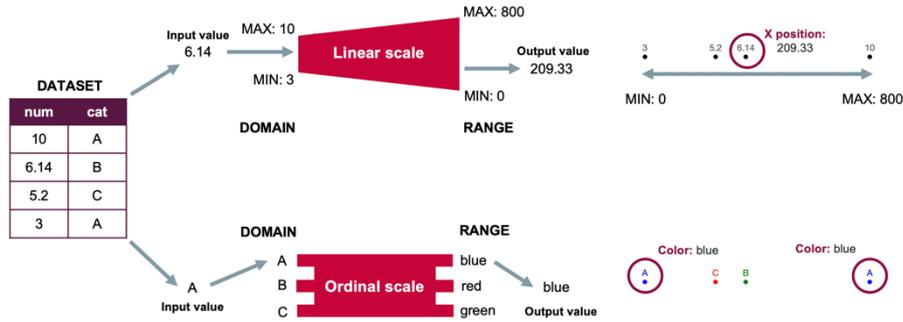
For this matter, we designed a procedure and a questionnaire that aims at measuring two cognitive dimensions following Bloom's taxonomy of educational objectives [28-30]:

- Remember - Identification of the elements that compose a data visualization at first sight, i.e., to test if the student can recognize and recall basic elements of data visualizations.
- Understand – Understandability of the data visualization design process, i.e., to test if the student is aware of the dimensions and elements involved in the display.

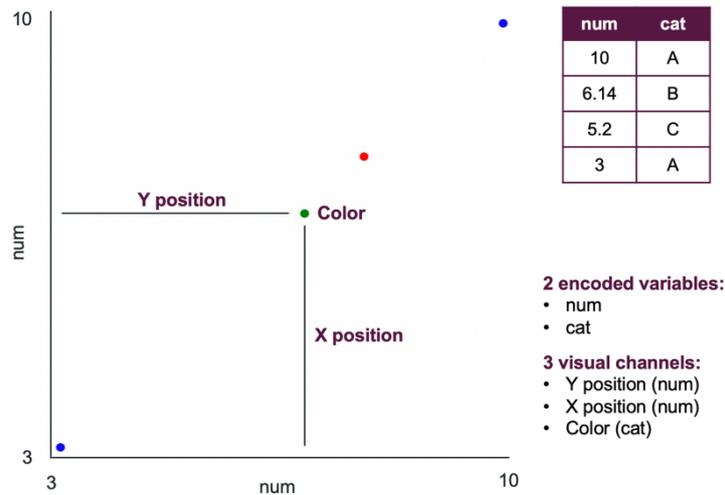
The study procedure consists of two similar parts. The first part involves a widely used data visualization tool, in this case, Tableau. Users are asked to download a test dataset and create a scatter plot that shows data values from two numerical variables and one categorical variable with Tableau. When finished, users close Tableau and answer the following questions about the scatter plot they just created:

1. Which variable was represented in the X-axis?
2. What was the maximum value of the Y-axis scale?
3. How many visual encodings were employed in your visualization?
4. How many scales were involved in your visualization?

This set of questions are focused on testing if users were aware of the design process of their own data visualizations and if they remembered basic features of their charts. Textual and graphical indications about the meaning of the data visualization terms involved in the questions are provided to avoid confusion. Explanations of the meaning of "scale" and "visual channel/encoding" are shown in Fig. 3 and Fig. 4, respectively.



**Figure 3:** Indications regarding the scale concept in the data visualization domain.



**Figure 4:** Indications regarding the channels/encodings concept in the data visualization domain.

The remaining questions in this part test the ability of users to identify the elements that compose an already implemented data visualization. Different screenshots of data visualizations created in Tableau are displayed, following the next questions:

1. How many variables are involved in the (screenshot's) visualization?
2. How many scales are involved in the (screenshot's) visualization?
3. How many visual channels or encodings are involved in the (screenshot's) visualization?

Figure 5 (a) shows an example of a question from this block. Once finished, the user is asked to create an additional scatter plot with MetaViz with the same dataset. A tutorial is also provided due to the novelty of this system. When the task has been completed, the previous questionnaire is presented in the context of the MetaViz system, that is, questions 1 to 4 referring to the scatter plot created by the user in MetaViz and questions 5 to 7 with screenshots of data visualizations created in this same system, as shown in Fig. 5 (b).

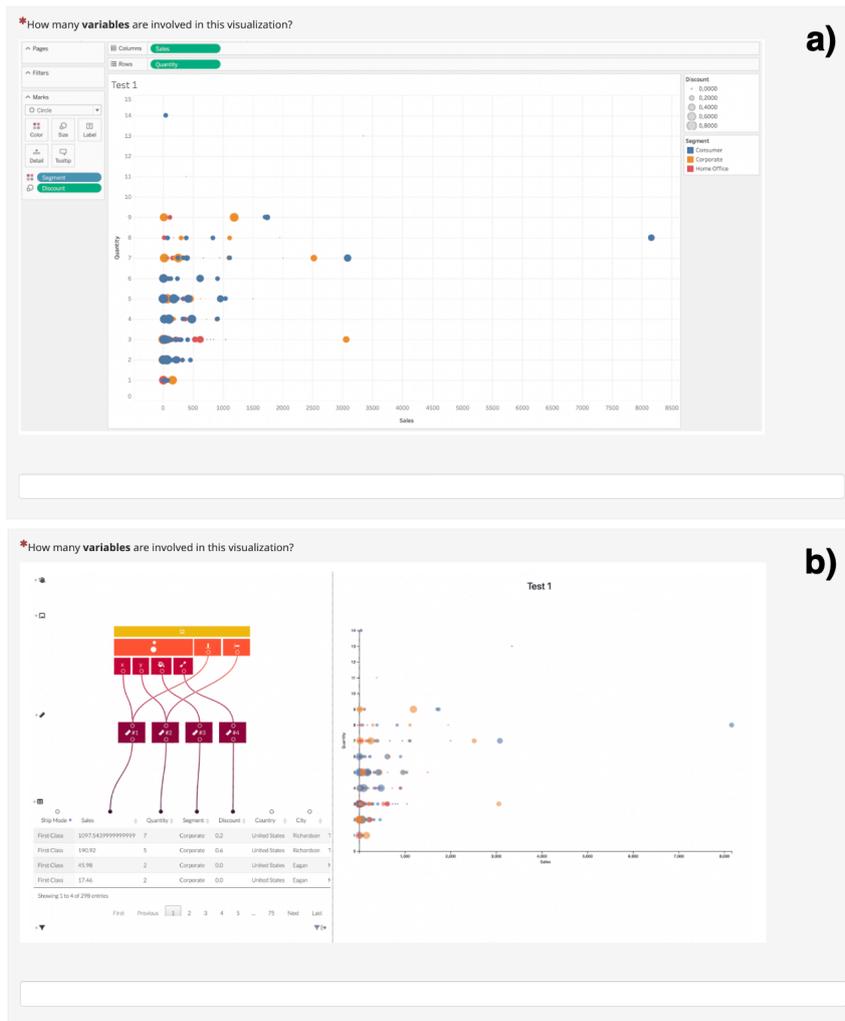


Figure 5: Questions regarding visualizations created in Tableau (a) and MetaViz (b). A high-resolution version of this image can be viewed at: <https://bit.ly/3Mx4ryu>.

## 4. Discussion and conclusions

This research presents a work-in-progress to measure the understanding of the elements involved in data visualizations and information dashboards. The study is set to validate the learning dimension of a novel system focused on the creation of data visualizations and dashboards through fine-grained components. This system (MetaViz) is based on a dashboard meta-model developed through domain engineering that captures different concepts and relationships from the data visualization domain.

We developed a study procedure and a questionnaire to determine whether users can understand and recall concepts related to data visualization with MetaViz. The questionnaire aims at measuring two cognitive dimensions –understanding and remembering– [28-30] in the context of education. In this case, MetaViz is compared to a commercial tool (Tableau), but other visual analytics tools can be employed to conduct this research.

The study will be carried out with a sample of students with data visualization skills. Future research will involve replicating this study with other samples, including lay users and people with knowledge from other domains, to compare the outcomes.

## 5. Acknowledgments

This research work has been supported by the Spanish Ministry of Education and Vocational Training under an FPU fellowship (FPU17/03276). This research was partially funded by the Spanish

Government Ministry of Economy and Competitiveness through the DEFINES project grant number (TIN2016-80172-R) and the Ministry of Science and Innovation through the AVisSA project grant number (PID2020-118345RB-I00).

## 6. References

- [1] A. Álvarez-Arana, M. Villamañe-Gironés, and M. Larrañaga-Olagaray, "Improving Assessment Using Visual Learning Analytics," *Education in the Knowledge Society*, vol. 21, no. 9, pp. 1-9, 2020, doi: 10.14201/eks.21554.
- [2] D. Patil and H. Mason, *Data Driven*. Sebastopol, CA, USA: O'Reilly Media, Inc., 2015.
- [3] T. H. Davenport and D. D. D. Prusak, *Working knowledge: How organizations manage what they know*. Brighton, Massachusetts, USA: Harvard Business Press, 1998.
- [4] C. Ware, *Information Visualization: Perception for Design*. Burlington: Elsevier, 2004.
- [5] A. Cairo, *The Functional Art: An introduction to information graphics and visualization*. San Francisco, CA, USA: New Riders, 2012.
- [6] A. Cairo, *The truthful art: Data, charts, and maps for communication*. San Francisco, CA, USA: New Riders, 2016.
- [7] 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 Computer Graphics*, vol. 25, no. 1, pp. 682 - 692, 2018, doi: 10.1109/TVCG.2018.2864903.
- [8] F. García-Sánchez, R. Therón, and J. Gómez-Isla, "Visual literacy in New Media: Systematic Review and Mapping of the Literature," *Education in the Knowledge Society*, vol. 20, p. 44, 2019, doi: [https://doi.org/10.14201/eks2019\\_20\\_a6](https://doi.org/10.14201/eks2019_20_a6).
- [9] E. Dimara, S. Franconeri, C. Plaisant, A. Bezerianos, and P. Dragicevic, "A task-based taxonomy of cognitive biases for information visualization," *IEEE transactions on visualization and computer graphics*, vol. 26, no. 2, pp. 1413-1432, 2018, doi: 10.1109/TVCG.2018.2872577.
- [10] P. Mena, "Reducing misperceptions through news stories with data visualization: The role of readers' prior knowledge and prior beliefs," *Journalism*, vol. 0, no. 0, p. 14648849211028762, 2021, doi: 10.1177/14648849211028762.
- [11] E. M. Peck, S. E. Ayuso, and O. El-Etr, "Data is Personal: Attitudes and Perceptions of Data Visualization in Rural Pennsylvania," presented at the Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland Uk, 2019. [Online]. Available: <https://doi.org/10.1145/3290605.3300474>.
- [12] Y.-S. Kim, K. Reinecke, and J. Hullman, "Data through others' eyes: The impact of visualizing others' expectations on visualization interpretation," *IEEE transactions on visualization and computer graphics*, vol. 24, no. 1, pp. 760-769, 2018, doi: 10.1109/TVCG.2017.2745240.
- [13] A. Cairo, *How charts lie: Getting smarter about visual information*. New York, NY, USA: WW Norton & Company, 2019.
- [14] D. A. Szafir, "The good, the bad, and the biased: Five ways visualizations can mislead (and how to fix them)," *Interactions*, vol. 25, no. 4, pp. 26-33, 2018, doi: 10.1145/3231772.
- [15] S. Berinato, *Good charts: The HBR guide to making smarter, more persuasive data visualizations*. Brighton, MA, USA: Harvard Business Review Press, 2016.
- [16] C. O. Wilke, *Fundamentals of data visualization: a primer on making informative and compelling figures*. Sebastopol, CA, USA: O'Reilly Media, 2019.
- [17] 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 *9th World Conference on Information Systems and Technologies*, Azores, Portugal, Á. Rocha, H. Adeli, G. Dzemyda, F. Moreira, and A. M. R. Correia, Eds., 2021, vol. 2: Springer Nature, in Trends and Innovations in Information Systems and Technologies, WorldCIST 2021, pp. 570-580, doi: 10.1007/978-3-030-72651-5\_54. [Online]. Available: <http://hdl.handle.net/10366/145626>

- [18] A. Vázquez-Ingelmo, F. J. García-Peñalvo, R. Therón, and M. Á. Conde, "Representing Data Visualization Goals and Tasks through Meta-Modeling to Tailor Information Dashboards," *Applied Sciences*, vol. 10, no. 7, p. 2306, 2020. [Online]. Available: <https://www.mdpi.com/2076-3417/10/7/2306>.
- [19] 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.
- [20] A. Vázquez-Ingelmo, F. J. García-Peñalvo, and R. Therón, "Capturing high-level requirements of information dashboards' components through meta-modeling," presented at the Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality, León, Spain, 2019. [Online]. Available: <https://doi.org/10.1145/3362789.3362837>.
- [21] A. Vázquez Ingelmo, F. J. García-Peñalvo, R. Therón Sánchez, and M. Á. Conde González, "Extending a dashboard meta-model to account for users' characteristics and goals for enhancing personalization," (in eng), *Proceedings of LASI-SPAIN 2019. Learning Analytics Summer Institute Spain 2019: Learning Analytics in Higher Education (Vigo, Spain, June 27-28, 2019)*. CEUR Workshop Proceedings Series, 2019. [Online]. Available: <http://hdl.handle.net/10366/139803>.
- [22] I. Reinhartz-Berger, S. Cohen, J. Bettin, T. Clark, and A. Sturm, *Domain engineering*. Heidelberg, Germany: Springer, 2013.
- [23] D. Bjørner, "Domain Engineering," in *Formal Methods: State of the Art and New Directions*, P. Boca, J. P. Bowen, and J. Siddiqi Eds. London: Springer London, 2010, pp. 1-41.
- [24] A. Pleuss, S. Wollny, and G. Botterweck, "Model-driven development and evolution of customized user interfaces," in *Proceedings of the 5th ACM SIGCHI symposium on Engineering interactive computing systems*, 2013: ACM, pp. 13-22.
- [25] 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.
- [26] P. Clements and L. Northrop, *Software product lines*. Boston, MA, USA: Addison-Wesley, 2002.
- [27] K. Pohl, G. Böckle, and F. J. Van der Linden, *Software Product Line Engineering: Foundations, Principles and Techniques*. New York, NY, USA: Springer-Verlag New York, Inc., 2005.
- [28] M. Keck, E. Stoll, and D. Kammer, "A Didactic Framework for Analyzing Learning Activities to Design InfoVis Courses," *IEEE Computer Graphics and Applications*, vol. 41, no. 6, pp. 80-90, 2021, doi: 10.1109/MCG.2021.3115416.
- [29] D. R. Krathwohl, "A Revision of Bloom's Taxonomy: An Overview," *Theory Into Practice*, vol. 41, no. 4, pp. 212-218, 2002/11/01 2002, doi: 10.1207/s15430421tip4104\_2.
- [30] B. Bloom, M. Engelhart, E. Furst, W. Hill, and D. Krathwohl, *Taxonomy of educational objectives: The classification of educational goals: Handbook I: Cognitive Domain*. New York, NY, USA: D. Mckay, 1956.