

# Interactive Visualization Notebooks: A Use Case in Space Weather Exploration

Joshua Reibert  
Institute of Data Science  
German Aerospace Center (DLR)  
Jena, Germany  
joshua.reibert@dlr.de

Matthias Pohl  
Institute of Data Science  
German Aerospace Center (DLR)  
Jena, Germany  
matthias.pohl@dlr.de

## Abstract

The increasing volume and complexity of data in many domains necessitate innovative approaches to data analysis and visualization. This paper presents an interactive visualization notebook method that integrates data management, processing, and visual analytics to facilitate knowledge generation and decision-making. We demonstrate the effectiveness of this approach through a use case in space weather exploration, where interactive visualizations are used to analyze the impact of solar activity on global navigation satellite systems. Our demonstrator shows that interactive visualization notebooks can streamline workflows, enhance transparency, and support reproducibility in data science and visual analytics pipelines. We discuss the limitations and potential extensions of this approach, highlighting the need for standardized data interfaces, customizable components, and high-performance computing infrastructure to support large-scale data analysis. This research contributes to the development of more efficient and effective data-driven decision-making processes in various domains.

## CCS Concepts

• **Human-centered computing** → **Visual analytics; Information visualization.**

## Keywords

Data Science Projects, Interactive Visualization, Space Weather Models, Information Visualization, Visual Analytics

## ACM Reference Format:

Joshua Reibert and Matthias Pohl. 2025. Interactive Visualization Notebooks: A Use Case in Space Weather Exploration. In *Proceedings of the 18th International Symposium on Visual Information Communication and Interaction (VINCI 2025)*, December 01–03, 2025, Linz, Austria. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3769534.3769581>

## 1 Introduction

The rapid increase in data generation [23] has significantly amplified the importance of data science (DS) as organizations recognize its potential to enhance performance and decision-making [3, 12, 14, 15, 22]. DS focuses on extracting actionable insights from complex datasets using advanced analytical techniques [17]. Unlike traditional software engineering projects, DS projects prioritize data

exploration, requiring a more adaptive approach [4]. Methodologies like cross-industry standard process for data mining (CRISP-DM) and the knowledge discovery in databases (KDD) process have been developed to aid data scientists in navigating data processing and model development [5, 18]. While these methodologies include visualization as a means to understand and communicate results, visual analytics (VA) puts more emphasis on the human component in sensemaking [9]. Visualization makes data and models accessible to users through perception, who in turn interact with the system to fine-tune parameters and eventually gain insights by discovering patterns and testing hypotheses [16].

Data scientists value notebooks for their flexibility and ease of use, which facilitates dynamic data exploration and visualization, leading to compelling insights for stakeholders [2]. However, transitioning these exploratory notebooks to production settings requires specialized skills beyond typical responsibilities, such as software deployment and system integration [6]. VA systems have especially high requirements for visualization and interaction in order to support analysts best in sensemaking. In larger organizations, dedicated engineering teams often assist in this process, providing custom solutions that ensure best practices and efficient model performance [6, 15]. Conversely, data scientists in smaller organizations may need to take on these additional tasks, which can overwhelm them and affect deployment quality [1, 2]. Yet, recent improvements to notebooks and visualization ecosystems are making them more viable to take on the full knowledge generation pipeline from data management to interactive data visualization.

This paper investigates the applicability of interactive visualization notebooks for data science and visual analytics pipelines. The approach combines data processing, dynamic visualization, and real-time interactive analyses to streamline workflows and ensure a transparent and reproducible data handling process. We illustrate this through a use case for the interactive exploration and analysis of space weather, which examines model data based on solar activity datasets to uncover patterns and trends.

## 2 Background

Data science (DS) provides methodologies for a systematic approach to data analysis that describe and formalize the process of extracting knowledge from data. The cross-industry standard process for data mining (CRISP-DM) framework encapsulates the common practices of data mining professionals across various industries [18]. It is composed of six phases: *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, and *deployment*. While CRISP-DM is primarily used in industrial contexts, the knowledge



This work is licensed under a Creative Commons Attribution 4.0 International License. *VINCI 2025, Linz, Austria*

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1845-8/25/12

<https://doi.org/10.1145/3769534.3769581>

discovery in databases (KDD) framework is more prominent in academic settings, focusing on pattern extraction from large datasets. The KDD process is characterized by its iterative nature and the potential for feedback loops between the various stages [5]. Based on the requirements of the domain and stakeholders, the first step is the *selection* of datasets or specific subsets that are relevant to the discovery objectives. Next, *preprocessing* procedures enhance data quality, collect additional information, and handle missing data. In the *transformation* stage, salient features are identified and selected. This may involve the application of dimensionality reduction techniques or transformation methods designed to minimize the number of variables, thereby reducing the amount of data without sacrificing critical information. The *data mining* phase applies algorithms to discover meaningful patterns within the data. Such patterns can manifest in various forms, including classification rules, decision trees, regression models, and clusters. *Interpretation* turns the discovered patterns into knowledge. This may require revisiting earlier steps for refinement or re-evaluation. Visualization techniques are crucial at this stage to communicate extracted patterns and their implications. Finally, newly uncovered knowledge is integrated into existing systems for practical application or documented for stakeholders. Through this systematic approach, the KDD process facilitates the extraction of valuable insights from complex datasets, thereby enabling informed decision-making and enhanced comprehension within the targeted application domain.

Visual analytics (VA) is “the science of analytical reasoning facilitated by interactive visual interfaces” [20]. It combines multiple research areas such as visualization, data mining, and data management to enable users to derive knowledge and support decisions based on massive datasets, thereby augmenting human reasoning with the computational power of machines [9]. As the name suggests, VA integrates more fundamentally with visualization to communicate intermediate results of semi-automated analyses, which users in turn steer through interactivity. Sacha et al. developed a knowledge generation model that ties together established DS process models and integrates them in a high-level description of VA systems [16]. Their model is split into a computer and a human part, which are linked through interaction and perception. On the computer side, data serves as a foundation for models that are created using processes such as the KDD and visualization of data and models. The human side consists of exploration, verification, and knowledge generation, which are modeled as loops to signal that these processes are often happening in parallel and are spontaneous in nature. The *exploration loop* describes the users’ interaction with the VA system to generate new visualizations and models. They test or reject hypotheses in the *verification loop* in order to eventually gain insights in the *knowledge generation loop*, which manifests itself as patterns in the data and their quality while also gaining experience about the VA system itself.

### 3 Interactive Visualization in Notebooks

Computational notebooks are a popular choice to combine code, data, visualization, and other interactive elements in a single document [11]. Multiple platforms exist, but the most prominent are Jupyter Notebooks [10]. Data science (DS) especially benefits from this combined approach for iteratively building data workflows,

tracking processing steps and lineage, enriching code with documentation, and presenting results to gain insights and support decision-making [11]. As a consequence, the whole DS pipeline can be combined into a single notebook, thereby eliminating potential friction points between different tools.

However, as more work is shifted into notebooks, their limitations also become more apparent [2]. Most notably, notebooks are not reproducible as results depend on the execution order of cells. The management of dependencies poses significant challenges in computational environments. The dissemination of results is often hindered by the difficulty of sharing findings across different platforms. Furthermore, the reusability of code cells is limited, and achieving interactivity within these frameworks is not a straightforward endeavor. Jupyter stores cells as well as their output in JSON, and hence, small changes cause big diffs, thereby making it harder to track changes between versions, e.g., when using git. Cells are not reactive and must be re-run manually whenever the variables they depend on change. Finally, the results cannot be easily shared as a web or command-line application and, therefore, are often only used for prototyping.

Marimo<sup>1</sup> is an open-source notebook solution for Python that aims to overcome these limitations. Being Python-first, marimo notebooks are valid Python code and therefore suitable for versioning and furthermore integrate well with common tooling such as package managers, linters, and code formatters. Cells are reactive and automatically run when their input values change. Consequently, marimo notebooks are executed in deterministic order and therefore reproducible.

The combination of code and graphical output is a key reason for the surge of computational notebooks [11]. Several Python libraries exist to visualize information, which are especially useful in the context of notebooks. The matplotlib library is the most popular, but mainly produces static charts with limited and cumbersome interactivity [13]. Other libraries, such as Plotly and Bokeh, are based on web technologies and therefore inherently interactive; however, this comes at the cost of limited customization options. In contrast, Altair is a declarative statistical visualization library for Python based on the Vega and Vega-Lite data visualization grammars [21]. Hence, it is very flexible and can produce a wide range of visualizations while supporting high levels of customization. Basic interactions are also supported as part of the specification and therefore relatively easy to add. Marimo notebooks have additional mechanisms to add interactivity to Altair charts automatically, and can access filtered datasets from interactive charts reactively.

As notebooks have evolved, they have become better suited to take on the wide range of tasks required for data science and visual analytics. Python supports data management well, mostly driven by *DataFrame* libraries such as pandas<sup>2</sup> and polars<sup>3</sup>, but also analytical database systems such as DuckDB<sup>4</sup>. With the addition of scientific computing libraries such as NumPy<sup>5</sup> and SciPy<sup>6</sup> and libraries for machine learning, the full KDD process can be implemented in a

<sup>1</sup><https://marimo.io/>

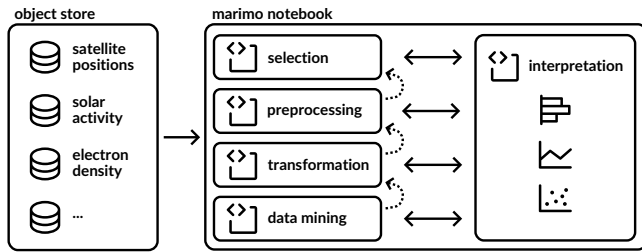
<sup>2</sup><https://pandas.pydata.org/>

<sup>3</sup><https://pola.rs/>

<sup>4</sup><https://duckdb.org/>

<sup>5</sup><https://numpy.org/>

<sup>6</sup><https://scipy.org/>



**Figure 1: The combination of an object store to hold data and a marimo notebook for data science processes such as the steps of the knowledge discovery in databases (KDD) process provides a comprehensive approach for visual exploration and analysis of space weather model data for sensemaking.**

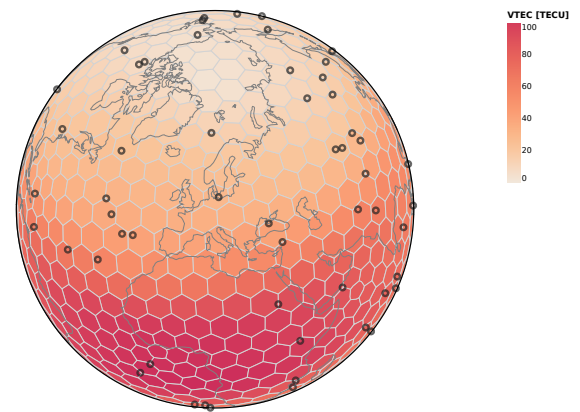
Python notebook. In contrast, business intelligence (BI) focused software for data visualization like Tableau<sup>7</sup> is limited to data readily available in specific file formats or databases. In combination with the aforementioned visualization libraries for interactive data visualization, notebooks can also provide the interfaces to the human sensemaking loop to implement the full VA pipeline in an interactive notebook.

#### 4 Space Weather Exploration Use Case

We demonstrate the applicability of interactive computation and visualization notebooks for the visual analysis and exploration of space weather. Space weather describes the conditions in Earth’s ionosphere, which is mostly driven by solar activity. It affects the quality of satellite communications, particularly during major events such as solar storms. This, in turn, can impact global navigation satellite system (GNSS) such as Galileo and GPS [19]. Hence, the space weather is closely observed using ground- and space-based sensors, which produce large quantities of raw data. Physicists model space weather using these data to analyze it and to correct its effects on GNSS [8]. This demonstrator aims to visualize and analyze the impact of space weather on GNSS satellite links for a given location on Earth.

The demonstrator uses two key infrastructure components: On one side, an object storage solution (e.g., MinIO<sup>8</sup>) serves as a centralized repository for all relevant data for further analysis. On the other side, a Python-based marimo notebook enables in-depth data processing as well interactive visualization and analysis. Importantly, the notebook environment includes one or more code cells that offer an array of visualization options (e.g., Altair[21], mosaic[7]). These visualizations play a crucial role in displaying the selected and transformed datasets, enabling clear understanding and interpretation of the final patterns revealed during the knowledge generation process.

Data exploration is modeled along the steps of the KDD process (Figure 1). The NEDM space weather model serves as a dataset. It provides global, three-dimensional electron density estimates based on the F10.7 index for the solar activity and along satellite links [8]. Information on satellites and their orbits, as orbit mean-elements



**Figure 2: After preprocessing, electron density is aggregated as the vertical total electron content and visualized as averages over the spatial H3 grid. The GNSS positions in view of the selected receiver position are overlaid.**

messages (OMM), is provided by CelesTrak<sup>9</sup>. In the preprocessing phase, GNSS data from satellite identifiers are derived, focusing on relevant GNSS satellites. A rolling average of the F10.7 solar flux index is computed, essential for assessing the impact of solar activity on the Earth’s ionosphere. During the transformation phase, the data for the data mining phase is selected. GNSS satellite positions are calculated using the simplified general perturbations (SGP4). In a second iteration, the electron density is aggregated on H3<sup>10</sup> grid cells for improved spatial distribution (Figure 2). Relative positions, such as elevation angle and distance, are computed to facilitate interactive map selection. Processed data is stored using DuckDB for integration into the final visualization. In the data mining phase, physicists use the processed F10.7 data to create models of space weather, which also predict ionospheric electron density. A visualization of visible GNSS satellites and their parameters allows the interactive selection of a subset of them based on a user-selected receiver position (Figure 3). The vertical total electron content (VTEC) and slant vertical total electron content (STEC) for accurate evaluations are defined and calculated. The finalized visualizations aid interpretation and showcase satellite positions alongside STEC values. Users can interactively re-calibrate data based on results and re-evaluate models, fostering an iterative process for deriving insights.

By default, the notebook operates in application mode, which hides the underlying code blocks. Users can easily initiate re-runs of the preprocessing, transformation, and data mining phases through a user-friendly dashboard interface, enhancing engagement and workflow efficiency.

#### 5 Discussion

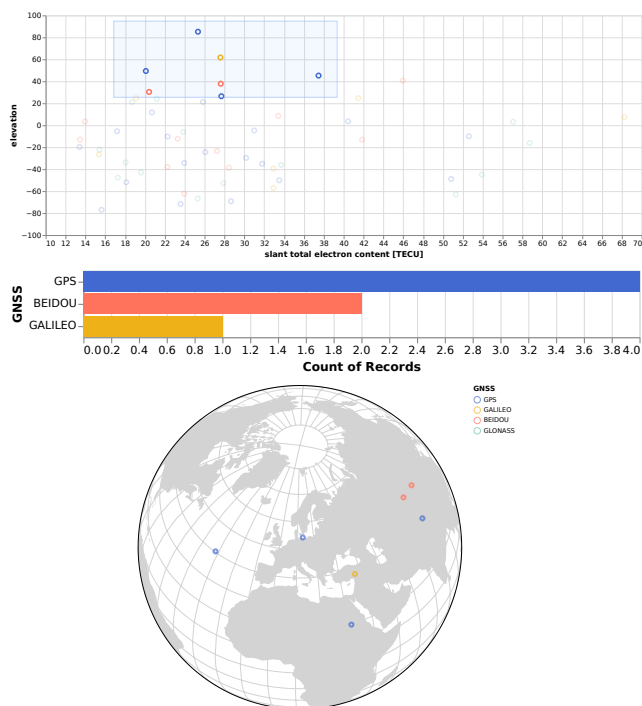
A use case was presented that illustrates the effectiveness of interactive visualization techniques within a data storage and programming notebook environment. This approach enables customized data processing and visualization that can be tailored to specific

<sup>7</sup><https://www.tableau.com/>

<sup>8</sup><https://min.io/>

<sup>9</sup><https://celestrak.org/>

<sup>10</sup><https://h3geo.org>



**Figure 3: The interactive visualizations show GNSS satellites relative to a ground position and derived values. The scatter plot visualizes their elevation angle in relation to the slant vertical total electron content (STEC) based on the NEDM model [8]. Finally, the map depicts the position of the selected satellites.**

use cases, allowing for insightful analysis and interpretation of complex datasets. However, this method has certain limitations in terms of generalizability, making it less adaptable for a wide range of applications across different domains. Authors of such notebooks require knowledge of programming, visualization, and data management, as well as domain-specific expertise. In contrast, other low- or no-code environments allow them to focus more on the domain and analysis, albeit being less customizable. The implementation of standardized data interfaces and configurable components would mitigate this, yielding a more robust and adaptable framework.

The selection of libraries used in the demonstrator imposes constraints on the interactivity, thereby limiting user engagement with the system. For instance, Altair and Mosaic offer interactive selections within their systems, but these are not always accessible programmatically for use in other code cells. The implementation of custom visualization widgets could circumvent this, but at the cost of additional effort and reduced integration with these libraries.

Large datasets pose a challenge for these visualization libraries, especially for Altair, which includes the data used as JSON in its specification. As a result, the chart is self-describing, but in turn, performance suffers heavily when visualizing larger datasets. In contrast, the Mosaic architecture decouples data processing from visualization specification and thereby boasts high performance [7]. However, there is currently no Python API for Mosaic, and thus

the specifications must be provided directly as JSON, making its use more error-prone and debugging errors more difficult.

When implementing a solution in a production environment, it is vital to select a stable toolset and technology stack that aligns with the project’s specific requirements. The current implementation of the use case is more suited as a proof of concept than for full-scale deployment, as it does not yet possess the robustness required for production. For example, interactions can easily trigger re-evaluations of multiple code cells, which in turn reduces reactivity of the notebook and negatively impacts user experience. Integration of high-performance computing (HPC) infrastructure, for example, with the widely used dask<sup>11</sup> alongside scalable, highly available data storage solutions, would improve reliability and facilitate more efficient data analysis, enabling near real-time processing capabilities. This combination not only enhances computational efficiency but also supports the management of large datasets, ensuring that the system can adequately meet the demands of complex data workflows in a production setting.

Finally, there is a need to confirm the utility of this notebook-based approach for VA using quantitative and qualitative user studies. This includes applicability to different domains, as well as usability and the effects on the quality of the insights found. Additionally, scalability and responsiveness should be tested using real-world scenarios as benchmarks to ensure that the user experience is not negatively impacted by high workloads.

## 6 Conclusion

This paper presents a notebook-based approach to data science and visual analytics that includes the full pipeline from data management and transformation to interactive, visual exploration and analysis. The interactive environment is suited for real-time data processing and dynamic visualizations, allowing users to engage with and adjust their analyses instantly. This streamlined workflow offers a comprehensive approach to the knowledge generation process, ensuring transparency and reproducibility. We discuss current limitations of this approach, yet also demonstrate its utility for a space weather analysis use case that empowers users to conduct thorough data analysis and produce impactful visual outputs that effectively communicate their findings.

## References

- [1] Titus Barik, Robert DeLine, Steven Mark Drucker, and Danyel Fisher. 2016. The bones of the system: a case study of logging and telemetry at Microsoft. In *Proceedings of the 38th International Conference on Software Engineering, ICSE 2016, Austin, TX, USA, May 14–22, 2016 - Companion Volume*, Laura K. Dillon, Willem Visser, and Laurie A. Williams (Eds.). ACM, 92–101. doi:10.1145/2889160.2889231
- [2] Souti Chattopadhyay, Ishita Prasad, Austin Z. Henley, Anita Sarma, and Titus Barik. 2020. What’s Wrong with Computational Notebooks? Pain Points, Needs, and Design Opportunities. In *CHI ’20: CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, April 25–30, 2020*, Regina Bernhaupt, Florian ‘Floyd’ Mueller, David Verweij, Josh Andres, Joanna McGrenere, Andy Cockburn, Ignacio Avellino, Alix Gogney, Pernille Bjøn, Shengdong Zhao, Briane Paul Samson, and Rafal Kocielnik (Eds.). ACM, 1–12. doi:10.1145/3313831.3376729
- [3] Hsinchun Chen, Roger H. L. Chiang, and Veda C. Storey. 2012. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Q.* 36, 4 (2012), 1165–1188. [http://www.misq.org/skin/frontend/default/misq/pdf/V36I4/SI\\_ChenIntroduction.pdf](http://www.misq.org/skin/frontend/default/misq/pdf/V36I4/SI_ChenIntroduction.pdf)
- [4] Manirupa Das, Renhao Cui, David R. Campbell, Gagan Agrawal, and Rajiv Ramnath. 2015. Towards methods for systematic research on big data. In *2015 IEEE International Conference on Big Data (IEEE BigData 2015), Santa Clara*,

<sup>11</sup><https://www.dask.org/>

- CA, USA, October 29 - November 1, 2015. IEEE Computer Society, 2072–2081. doi:10.1109/BIGDATA.2015.7363989
- [5] Usama M. Fayyad, Gregory Piatetsky-Shapiro, and Padhraic Smyth. 1996. Knowledge Discovery and Data Mining: Towards a Unifying Framework. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), Portland, Oregon, USA*, Evangelos Simoudis, Jiawei Han, and Usama M. Fayyad (Eds.). AAAI Press, 82–88. <http://www.aaai.org/Library/KDD/1996/kdd96-014.php>
- [6] Christian Haertel, Daniel Staegemann, Christian Daase, Matthias Pohl, Abdulrahman Nahhas, and Klaus Turowski. 2023. MLOps in Data Science Projects: A Review. In *2023 IEEE International Conference on Big Data (BigData)* (Sorrento, Italy, 2023-12-15). IEEE, 2396–2404. doi:10.1109/BigData59044.2023.10386139
- [7] Jeffrey Heer and Dominik Moritz. 2024. Mosaic: An Architecture for Scalable & Interoperable Data Views. *IEEE Trans. Vis. Comput. Graph.* 30, 1 (2024), 436–446. doi:10.1109/TVCG.2023.3327189
- [8] Mohammed Mainul Hoque, Norbert Jakowski, and Fabricio S. Prol. 2022. A new climatological electron density model for supporting space weather services. *12 (2022)*, 1. doi:10.1051/swsc/2021044
- [9] Daniel A. Keim, Gennady L. Andrienko, Jean-Daniel Fekete, Carsten Görg, Jörn Kohlhammer, and Guy Melançon. 2008. Visual Analytics: Definition, Process, and Challenges. In *Information Visualization - Human-Centered Issues and Perspectives*, Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North (Eds.). Lecture Notes in Computer Science, Vol. 4950. Springer, 154–175. doi:10.1007/978-3-540-70956-5\_7
- [10] Thomas Kluyver, Benjamin Ragan-Kelley, Fernando Pérez, Brian E. Granger, Matthias Bussonnier, Jonathan Frederic, Kyle Kelley, Jessica B. Hamrick, Jason Grout, Sylvain Corlay, Paul Ivanov, Damián Avila, Safia Abdalla, Carol Willing, and Jupyter Development Team. 2016. Jupyter Notebooks - a publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas, 20th International Conference on Electronic Publishing, Göttingen, Germany, June 7-9, 2016*, Fernando Loizides and Birgit Schmidt (Eds.). IOS Press, 87–90. doi:10.3233/978-1-61499-649-1-87
- [11] Robert Kosara. 2023. Notebooks for Data Analysis and Visualization: Moving Beyond the Data. *IEEE Computer Graphics and Applications* 43, 1 (2023), 91–96. doi:10.1109/MCG.2022.3222024
- [12] Oliver Müller, Maria Fay, and Jan vom Brocke. 2018. The Effect of Big Data and Analytics on Firm Performance: An Econometric Analysis Considering Industry Characteristics. *J. Manag. Inf. Syst.* 35, 2 (2018), 488–509. doi:10.1080/07421222.2018.1451955
- [13] Jorge Piazzentin Ono, Juliana Freire, Cláudio T. Silva, João Comba, and Kelly P. Gaither. 2021. Interactive Data Visualization in Jupyter Notebooks. *Comput. Sci. Eng.* 23, 2 (2021), 99–106. doi:10.1109/MCSE.2021.3052619
- [14] Matthias Pohl, Christian Haertel, and Klaus Turowski. 2023. Value Creation from Data Science Applications - A Literature Review. In *Perspectives in Business Informatics Research - 22nd International Conference on Business Informatics Research, BIR 2023, Ascoli Piceno, Italy, September 13-15, 2023, Proceedings (Lecture Notes in Business Information Processing, Vol. 493)*, Knut Hinkelmann, Francisco J. López-Pellicer, and Andrea Polini (Eds.). Springer, 327–338. doi:10.1007/978-3-031-43126-5\_23
- [15] Matthias Pohl, Daniel Gunnar Staegemann, and Klaus Turowski. 2022. The Performance Benefit of Data Analytics Applications. In *The 13th International Conference on Ambient Systems, Networks and Technologies (ANT 2022) / The 5th International Conference on Emerging Data and Industry 4.0 (EDI40 2022) / Affiliated Workshops, March 22-25, 2022, Porto, Portugal (Procedia Computer Science, Vol. 201)*, Elhadi M. Shakshuki and Ansar-Ul-Haque Yasar (Eds.). Elsevier, 679–683. doi:10.1016/j.PROCS.2022.03.090
- [16] Dominik Sacha, Andreas Stoffel, Florian Stoffel, Bum Chul Kwon, Geoffrey P. Ellis, and Daniel A. Keim. 2014. Knowledge Generation Model for Visual Analytics. *IEEE Trans. Vis. Comput. Graph.* 20, 12 (2014), 1604–1613. doi:10.1109/TVCG.2014.2346481
- [17] Michael Schulz, Uwe Neuhaus, Jens Kaufmann, Daniel Badura, Stephan Kühnel, Wolfgang Badewitz, David Dann, Simon Kloker, Emal M. Alekozai, and Carsten Lanquillon. 2020. Introducing DASC-PM: A Data Science Process Model. In *Australasian Conference on Information Systems, ACIS 2020, Wellington, New Zealand, December 1-4, 2020*. 45. <https://aisel.aisnet.org/acis2020/45>
- [18] Colin Shearer. 2000. The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of data warehousing* 5, 4 (2000), 13–22.
- [19] Ke Su, Shuanggen Jin, and Mohammed Mainul Hoque. 2019. Evaluation of Ionospheric Delay Effects on Multi-GNSS Positioning Performance. *Remote. Sens.* 11, 2 (2019), 171. doi:10.3390/RS11020171
- [20] James J. Thomas and Kristin A. Cook. 2006. A Visual Analytics Agenda. *IEEE Computer Graphics and Applications* 26, 1 (2006), 10–13. doi:10.1109/MCG.2006.5
- [21] Jacob VanderPlas, Brian E. Granger, Jeffrey Heer, Dominik Moritz, Kanit Wongsuphasawat, Arvind Satyanarayan, Eitan Lees, Ilia Timofeev, Ben Welsh, and Scott Sievert. 2018. Altair: Interactive Statistical Visualizations for Python. *J. Open Source Softw.* 3, 32 (2018), 1057. doi:10.21105/JOSS.01057
- [22] Samuel Fosso Wamba, Angappa Gunasekaran, Shahriar Akter, Steven Ji fan Ren, Rameshwar Dubey, and Stephen J. Childe. 2017. Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research* 70 (2017), 356–365. doi:10.1016/j.jbusres.2016.08.009
- [23] Shen Yin and Okyay Kaynak. 2015. Big Data for Modern Industry: Challenges and Trends [Point of View]. *Proc. IEEE* 103, 2 (2015), 143–146. doi:10.1109/JPROC.2015.2388958