International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 15 Issue 3, March-2026

AI Data Assistant

Prof. Baliram B. Deshmukh[1], Taware Tejaswi[2], Yadav Shruti[3], Tanaya Bandal[4], Ankita Hake[5]

[1,2,3,4,5]Department of Artificial Intelligence and machine learning, Navsahyadri Education Society's Group Of Institutes, Pune,

ABSTRACT

This paper presents a web-based automated data visualization system designed to simplify the process of data analysis. The proposed platform enables users to upload structured datasets and automatically prepares the data for visualization by performing essential preprocessing operations. It supports multiple chart types and allows users to generate dashboards through simple column selection without requiring programming knowledge or complex configurations. The system focuses on usability and automation to make data exploration faster and more accessible. Experimental observations indicate that the platform reduces manual effort and improves the efficiency of dashboard creation compared to traditional visualization tools.

Keywords— Automated Data Visualization, Dashboard Generation, Data Preprocessing Automation, Web-Based Analytics, Non-Technical Users.

INTRODUCTION

The continuous growth of digital data has increased the importance of effective data analysis and visualization in decision-making processes. Visual representations such as charts and dashboards help users understand patterns, trends, and relationships more clearly than raw data. Because of this, business intelligence tools are widely used across industries.

However, many existing visualization platforms require technical expertise. Users often need to perform manual data cleaning, manage data types, and apply formulas or query-based operations before generating meaningful dashboards. For non-technical users such as business professionals, students, or small organizations, these steps can be difficult and time-consuming. As a result, they frequently depend on technical experts to analyze data, which reduces efficiency.

To overcome these challenges, this research proposes a web-based automated data visualization system designed for non-technical users. The system simplifies the analytics workflow by automatically detecting data types, handling missing or blank values, removing unnecessary spaces, and preparing the dataset for visualization. After preprocessing, users can generate different types of charts and create dashboards by simply selecting the required columns, without writing code or using complex formulas.

The objective of this work is to reduce the complexity of traditional business intelligence tools and provide an accessible, automation-driven platform that enables efficient and independent data analysis.

Problem Statement

Although many data visualization and business intelligence tools are available, they often require technical knowledge for effective use. Users must manually clean data, manage data types, and apply formulas or queries before generating dashboards. For non-technical users, these steps are complex and time-consuming. As a result, many individuals depend on technical experts to perform data analysis, which reduces efficiency and delays decision-making. There is a clear need for a system that can automatically handle data preprocessing and enable simple dashboard generation without requiring programming skills or advanced analytical knowledge.

Objectives of the Study

The main aim of this project is to develop a user-friendly data visualization platform that reduces the technical effort required in traditional analytics tools. The system is designed with a focus on simplicity and automation so that even users without a technical background can analyze data independently.

The objectives of this study are to create a web-based application that allows easy dataset upload and automatic data preparation. Another important objective is to automate common preprocessing steps such as identifying data types, handling missing or blank values, and cleaning unwanted spaces in the dataset. The system also aims to provide multiple visualization options that can be generated by simple column selection, without requiring coding or formula-based operations.

Additionally, the project focuses on simplifying dashboard creation and reducing the overall time required to convert raw data into meaningful visual insights.

PROPOSED SYSTEM ARCHITECTURE

The proposed system follows a modular architecture designed to simplify the data visualization workflow while maintaining flexibility. The overall process begins with data input and ends with automated dashboard generation.

The first component of the system is the Data Input Module, which allows users to upload structured datasets in supported formats such as CSV or Excel. Once the dataset is uploaded, it is passed to the preprocessing layer.

The Data Preprocessing Module automatically analyzes the dataset to detect data types such as numerical, categorical, and date values. It also identifies missing or blank entries and removes unnecessary white spaces. This step ensures that the dataset is cleaned and properly structured before visualization.

After preprocessing, the data is forwarded to the Visualization Engine. This module enables the generation of different chart types based on the selected columns. The system dynamically prepares appropriate visual representations without requiring manual configuration.

Finally, the Dashboard Generation Module allows users to create dashboards by selecting required columns and assigning a dashboard name. The generated dashboard organizes the selected visualizations in a structured format, making it easy to interpret insights.

This modular design ensures smooth data flow between components and supports automation at each stage of the analytics process.

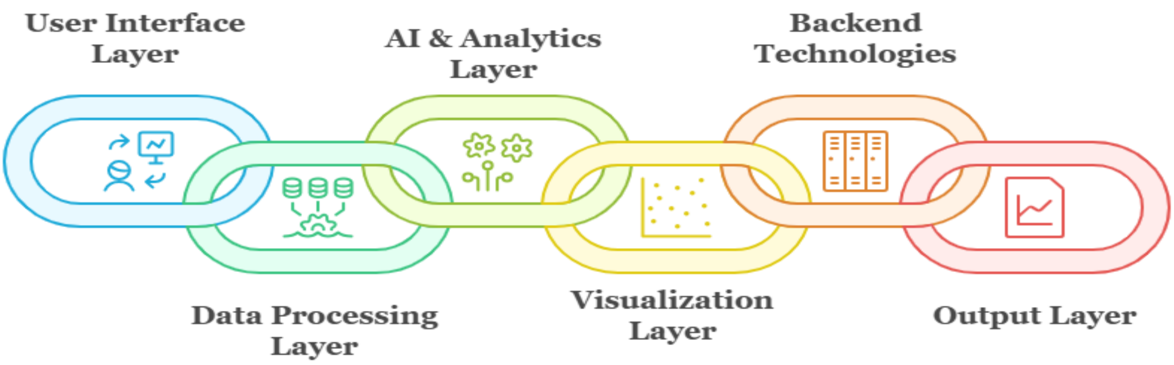


Fig. System Architecture

METHODOLOGY

The methodology of the proposed system is designed to automate the complete workflow from raw data upload to dashboard generation. The process follows a structured sequence of steps to ensure accuracy and efficiency.

Initially, the user uploads a structured dataset through the web interface. Once uploaded, the system reads the dataset and analyzes its schema to identify column names and data structure. In the next step, the preprocessing mechanism automatically detects the data type of each column, such as numerical, categorical, or date-based attributes.

After data type identification, the system performs cleaning operations. Missing or blank values are detected and handled appropriately. Unnecessary white spaces are removed, and inconsistent formatting is corrected wherever possible. These automated transformations ensure that the dataset is ready for visualization without requiring manual intervention. Following preprocessing, the visualization logic is applied. When the user selects a specific column, the system evaluates its data type and generates suitable visualizations accordingly. For example, categorical data may be represented using pie or bar charts, while numerical relationships can be displayed using scatter or area charts. This dynamic mapping between data type and visualization type enables efficient chart generation.

Finally, the selected visualizations are organized into a dashboard structure. The user can assign a dashboard name, and the system automatically arranges the chosen charts in a clear and structured layout. This step-by-step methodology ensures minimal user effort while maintaining functional flexibility and automation throughout the analytics process.

DATASET DESCRIPTION

To evaluate the performance and usability of the proposed system, structured tabular datasets were used. The system supports commonly used formats such as CSV and Excel files, where rows represent records and columns represent attributes.

For experimental analysis, datasets containing mixed data types were selected, including numerical, categorical, and date-based attributes. The datasets also contained common real-world issues such as missing values, blank entries, and inconsistent formatting. These characteristics were intentionally considered to test the effectiveness of the automated preprocessing module.

The size of the datasets varied to examine system performance under different conditions. Small and medium-sized datasets were used to evaluate dashboard generation speed and preprocessing accuracy. This approach helped assess how efficiently the system handles data cleaning and visualization without manual intervention.

The selected datasets were suitable for generating multiple chart types and validating the dynamic visualization logic of the system. By using structured but imperfect datasets, the robustness of the automation process was effectively examined.

EXPERIMENTAL EVALUATION

The proposed system was experimentally evaluated to measure its performance, automation efficiency, and usability. The evaluation focused on three main aspects: preprocessing time, dashboard generation time, and user interaction effort.

To analyze performance, structured datasets of small and medium sizes were used. The time required to clean and prepare the dataset was recorded and compared with the estimated time taken using traditional business intelligence tools where preprocessing is performed manually. The results indicated that the automated preprocessing module reduced data preparation time significantly, as common tasks such as data type detection, blank value handling, and space removal were completed automatically.

Dashboard generation efficiency was also observed. In conventional tools, users are required to configure charts manually and perform multiple setup steps. In the proposed system, visualizations were generated through simple column selection, reducing the number of interactions required. This demonstrated improved workflow simplicity and faster dashboard creation.

Additionally, usability was evaluated by analyzing the number of steps required to generate insights. The proposed system minimized manual operations and eliminated the need for formula-based configurations, making the analytics process more accessible to non-technical users.

RESULTS AND DISCUSSION

The performance of the proposed system was evaluated by measuring data preprocessing time, dashboard generation time, and user interaction effort. The results were compared with the estimated workflow followed in traditional data visualization tools.

For a structured dataset containing approximately 5,000 records, the automated preprocessing module completed data cleaning tasks such as data type detection, blank value handling, and removal of unnecessary spaces in an average time of 3–5 seconds. In contrast, manual preprocessing in traditional tools required approximately 2–4 minutes, depending on user expertise. This demonstrates a significant reduction in preparation time due to automation.

Similarly, dashboard generation efficiency was measured. In conventional tools, users typically perform multiple configuration steps, including chart selection, field mapping, and parameter adjustments. This process required approximately 3–5 minutes for basic dashboard creation. In the proposed system, users were able to generate

International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 15 Issue 3, March-2026

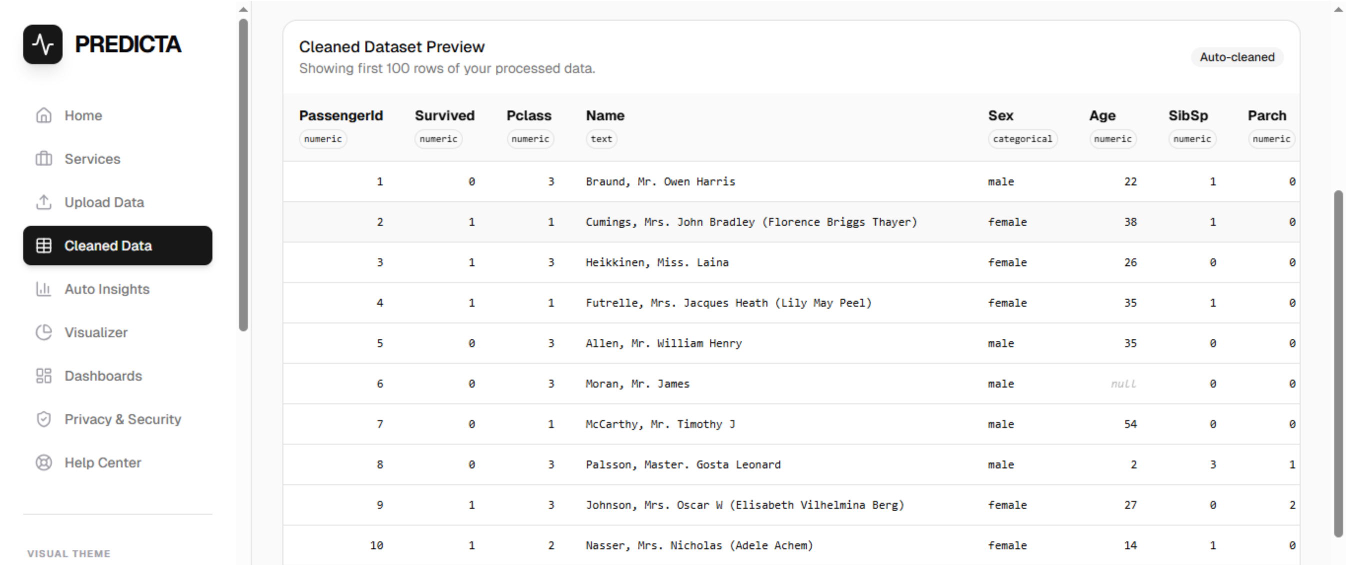
dashboards within 15–30 seconds by selecting the required columns and assigning a dashboard name. The reduction in steps and configuration effort highlights the simplicity of the proposed approach.

In terms of user interaction, the system minimized technical dependency by eliminating the need for formula writing or complex configuration settings. Non-technical users were able to create visualizations without prior knowledge of analytical tools, indicating improved usability.

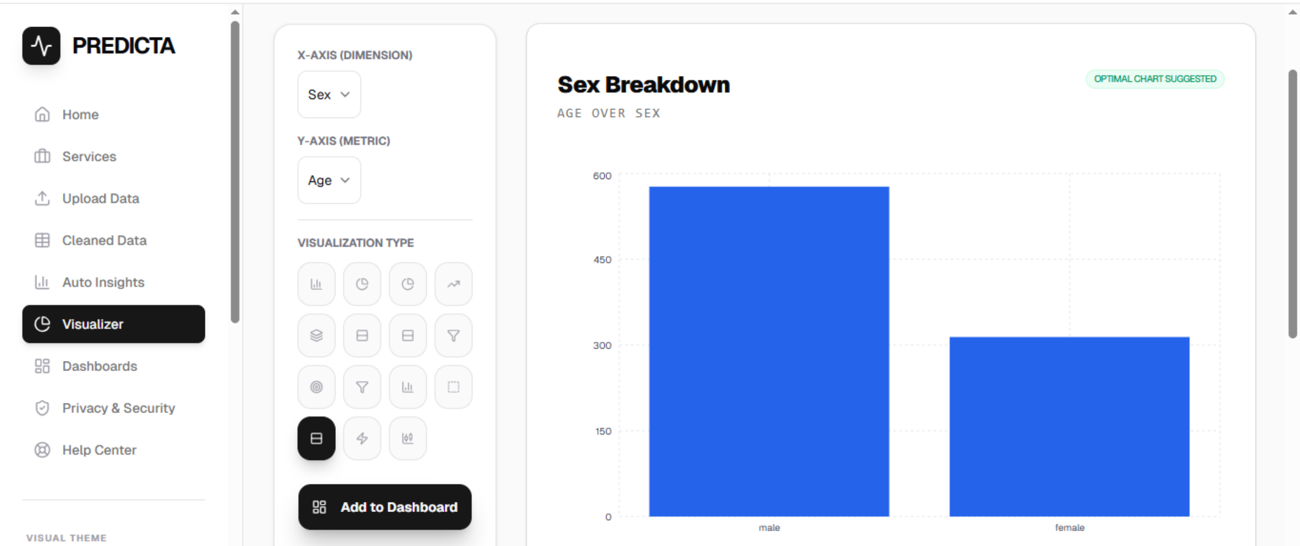
The results confirm that integrating automated preprocessing and visualization logic within a single workflow significantly enhances efficiency. However, performance may vary for very large datasets, and optimization for high-volume data processing remains an area for future improvement.

Overall, the proposed system successfully reduces time consumption, simplifies workflow complexity, and improves accessibility in data visualization tasks.

After uploading the dataset, the system automatically performs data cleaning and preprocessing operations. It identifies missing values, removes inconsistencies, and standardizes the data format. The cleaned dataset is then displayed in a tabular preview for user verification. This process reduces manual effort and prepares the data for visualization and analysis.



The system allows users to create visualizations by selecting the required columns for the X-axis and Y-axis. Based on the selected attributes, an appropriate chart is generated automatically. Users can choose different visualization types to analyze relationships and patterns within the dataset. This interactive approach simplifies data exploration and enables effective graphical representation of data without requiring technical expertise.



After creating visualizations, the system automatically combines selected charts into an interactive dashboard. The dashboard provides a consolidated view of key insights and helps users analyze data more effectively. Users can customize the dashboard according to their requirements and easily monitor multiple visualizations in a single interface. Additionally, the generated dashboard can be downloaded or shared, enabling easy access and collaboration among users.

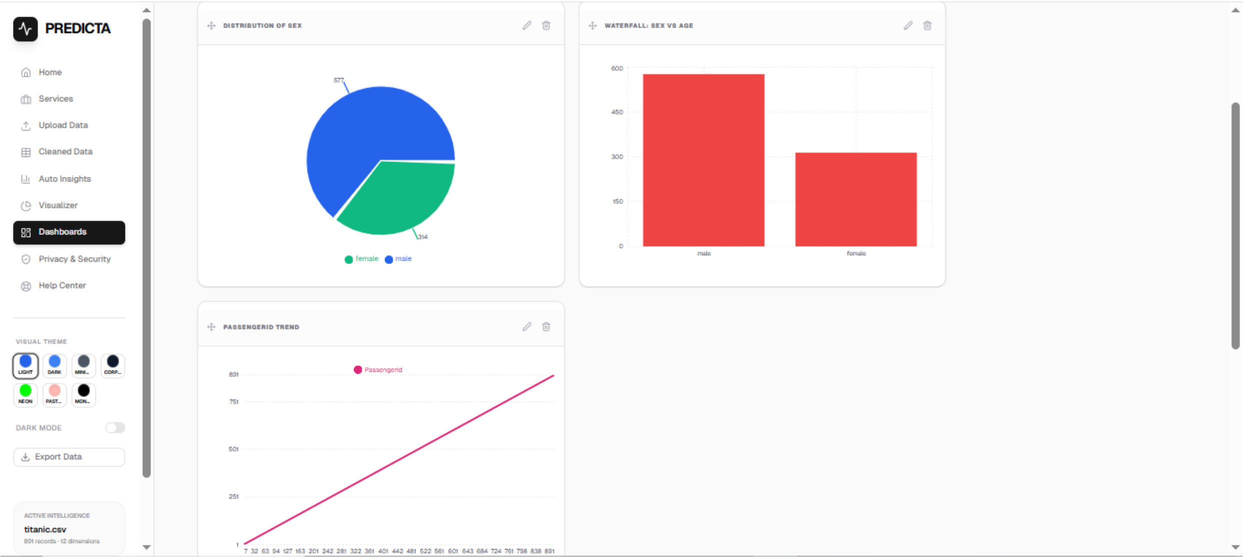


Table 1: Performance Comparison Between Traditional Tools and Proposed System:

Parameter	Traditional Tools	Proposed System
Data Preprocessing Time (5,000 records)	2–4 minutes (manual cleaning)	3–5 seconds (automated)
Dashboard Generation Time	3–5 minutes	15–30 seconds
Data Type Detection	Manual configuration required	Automatic detection
Missing Value Handling	Manual handling	Automatic handling
Number of Steps for Basic Dashboard	6–8 steps	2–3 steps
Technical Knowledge Requirement	Moderate to High	Minimal

Table 1 highlights the comparative performance between traditional data visualization tools and the proposed automated system. The results indicate a significant reduction in preprocessing and dashboard generation time. Furthermore, automation reduces user effort and technical dependency, making the system more suitable for non-technical users.

CONCLUSION AND FUTURE WORK

This paper presented a web-based automated data visualization system designed to simplify data analysis for non-technical users. The system integrates automated data preprocessing and dashboard generation within a single workflow, thereby reducing manual effort and technical complexity.

Experimental results show a noticeable reduction in preprocessing and dashboard creation time compared to traditional visualization tools. The platform enables users to generate insights through simple column selection without requiring advanced analytical knowledge. Future work will focus on improving scalability for larger datasets, enhancing performance optimization, and integrating intelligent chart recommendation mechanisms. Additionally, voice-based interaction can be incorporated to allow users to generate visualizations through voice commands, further improving accessibility and user experience.

REFERENCES

1. T. Liu, S. Han, H. Wang, X. Luo, P. Lu, B. Zhu, Y. Wang, K. Li, J. Chen, R. Qu, Y. Liu, X. Cui, A. Yaish, Y. Chen, M. Hao, C. Li, K. Li, A. Cohan, H. Xu, M. Gerstein, J. Zou, and H. Zhao, "Advancing AI Research Assistants with Expert-Involved Learning," 2024.

2. A. Kumar, J. Aurisano, B. Di Eugenio, and A. Johnson, “Intelligent Assistant for Exploring Data Visualizations,” University of Illinois at Chicago, 2024.
3. M. Schaekermann, G. Beaton, E. Sanoubari, A. Lim, K. Larson, and E. Law, “Ambiguity-aware AI Assistants for Medical Data Analysis,” University of Waterloo and University of Toronto, Canada, 2024.
4. S. R., M. A. V. P., L. A. K., P. Sumathi, M. R. P., and G. C. H., “AutoStream: Your All-in-One Data Science Assistant Using LLM,” Int. J. Sci. Res. Eng. Manage., vol. 9, no. 4, pp. –, Apr. 2025.
5. S. H. Berry, N. Burger, H. Dogo, K. B. Kumar, A. Malchiodi, J. Martini, T. Mengistu, H. J. Shatz, A. C. Smith, A. Usanov, and J. K. Yoong, Designing a System for Collecting Policy-Relevant Data for the Kurdistan Region—Iraq, RAND Corporation. [Online]. Available: https://www.jstor.org/stable/10.7249/j.ctt6wq9bd.13
6. D. P. Tegarden, “Business Information Visualization,” Communications of the Association for Information Systems, vol. 1, Jan. 1999.
7. R. Strauch, Data Analysis Techniques: Special Processing, NOAA/ERt/Wave Propagation Laboratory, Boulder, CO, USA, 2024.
8. G. C. Nair, “Data Visualization Tools: A Comparative Analysis,” College of Applied Science Perissery, 2024.
9. T. Petricek, G. J. J. van den Burg, A. Nazábal, T. Ceritli, E. Jiménez-Ruiz, and C. K. I. Williams, “AI Assistants: A Framework for Semi-Automated Data Wrangling,” IEEE Trans. Knowl. Data Eng., 2022.
10. A. Mandamadiotis, S. Eleftherakis, A. Glenis, D. Skoutas, Y. Stavrakas, and G. Koutrika, “DatAgent: The Imminent Age of Intelligent Data Assistants,” Athena Research Center, Athens, Greece, 2024.
11. C. D’Ignazio and R. Bhargava, “Data visualization literacy: A feminist starting point,” in Data Visualization in Society, M. Engebretsen and H. Kennedy, Eds., Amsterdam University Press, 2024. [Online]. Available: https://www.jstor.org/stable/j.ctvzgb8c7.19
12. M. Lagisz, A. Mizuno, K. Morrison, P. Pollo, L. Ricolfi, Y. Yang, and S. Nakagawa, “Using Elicit AI research assistant for data extraction in systematic reviews: a feasibility study across environmental and life sciences,” 2024.
13. A. Kumar, J. Aurisano, B. Di Eugenio, and A. Johnson, “Intelligent Assistant for Exploring Data Visualizations,” University of Illinois at Chicago, 2024.
14. T. Hiippala, “A multimodal perspective on data visualization,” in Data Visualization in Society, M. Engebretsen and H. Kennedy, Eds., Amsterdam University Press, 2024. [Online]. Available: https://www.jstor.org/stable/j.ctvzgb8c7.23
15. S. D. Mishra, A. Dhiman, and D. Dhyani, “AIAssistant for Daily Use,” Int. J. Sci. Dev. Res., vol. 9, no. 10, Oct. 2024.
16. H. Kennedy, W. Weber, and M. Engebretsen, “Data visualization and transparency in the news,” in Data Visualization in Society, Amsterdam University Press, 2024. [Online].Available: https://www.jstor.org/stable/j.ctvzgb8c7.17
17. T. Uberg Nærland, “The political significance of data visualization: Four key perspectives,” in Data Visualization in Society, Amsterdam University Press, 2024.[Online].Available: https://www.jstor.org/stable/j.ctvzgb8c7.10
18. [18] Y. Yu, L. Shen, F. Long, H. Qu, and H. Chen, “PyGWalker: On-the-fly Assistant for Exploratory Visual Data Analysis,” IEEE Visualization Conference, 2024. [Online]. Available: arXiv:2406.11637v1 [cs.HC]
19. Y. Wu, J. Yang, F. Jorge, and B. Shneiderman, “Visualizing Online Conversations: A Multilevel Approach,” IEEE Transactions on Visualization and Computer Graphics, vol. 23, no. 1, pp. 151–160, Jan. 2017.
20. B. Bilalli, A. Abelló, T. Aluja-Banet, R. F. Munir, and R. Wrembel, “PRESISTANT: Data Pre-processing Assistant,” Universitat Politècnica de Catalunya & Poznan University of Technology, 2024.
21. J. Schwabish, D. Harvey, M. Langness, V. Pancini, A. Rogin, and G. Velasco, Communication and Data Visualization, Urban Institute, 2023. [Online]. Available: https://www.jstor.org/stable/resrep69054.8
22. A. L. Lansky, A Data Analysis Assistant, Recom/NASA Ames Research Center, 2024.
23. S. K. P. Rajagopal, “Analysis of US Population using Data Analytics and Data Science Tools,” Int. J. Sci. Res., vol. 9, no. 10, 2024.
24. H. Taherdoost, “Different Types of Data Analysis; Data Analysis Methods and Techniques in Research Projects,” Int. J. Adv. Res. Manag., vol. 9, no. 1, pp. 1–9, 2020.
25. J. van Mil and J. P. Quintais, “A Matter of (Joint) control? Virtual assistants and the general data protection regulation,” Institute for Information Law (IViR), University of Amsterdam, 2024.
26. T. Luo, C. Huang, L. Shen, B. Li, S. Shen, W. Zeng, N. Tang, and Y. Luo, “nvBench 2.0: Resolving Ambiguity in Text-to-Visualization through Stepwise Reasoning,” The Hong Kong University of Science and Technology, 2024.
27. J. E. Thatcher and C. M. Dalton, What Are Our Data, and What Are They Worth?, Pluto Press, 2022. [Online]. Available: https://www.jstor.org/stable/j.ctv249sg9w.9
28. W. Weber, Exploring narrativity in data visualization in journalism, in Data Visualization in Society, Amsterdam University Press, 2024. [Online]. Available: http://www.jstor.com/stable/j.ctvzgb8c7.24
29. M. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer, “Vega-Lite: A Grammar of Interactive Graphics,” IEEE Trans. Visualization & Comp. Graphics, vol. 23, no. 1, pp. 341–350, 2017.

30. J. Heer, F. B. Viégas, and M. Wattenberg, "Voyagers and Interactive Data Analytics," IEEE Trans. Visualization & Comp. Graphics, vol. 19, no. 12, pp. 2436–2445, 2013.
31. K. Shneiderman, "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations," in Proc. IEEE Symp. Visual Languages, 1996, pp. 336–343.
32. M. 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, Jan. 2017.
33. D. Pomarède, H. M. Courtois, Y. Hoffman, and R. B. Tully, "Cosmography and Data Visualization," Publ. Astron. Soc. Pac., vol. 129, 058002, May 2017.