Visual Analytics Challenges and Trends in the Age of AI: The BigVis Community Perspective*

Nikos Bikakis†
Hellenic Mediterranean University &
Archimedes/Athena RC
Greece

Panos K. Chrysanthis
University of Pittsburgh
USA

Guoliang Li
Tsinghua University
China

George Papastefanatos
Athena RC
Greece

Lingyun Yu
Xi'an Jiaotong–Liverpool Univ.
China

1 Introduction

This report provides insights into the challenges, emerging topics, and opportunities related to human–data interaction and visual analytics in the AI era.

The BigVis 2024[1] organizing committee conducted a survey among experts in the field. They invite the Program Committee members and the authors of accepted papers to share their views. *Thirty-two scientists* from diverse research communities, including Databases, Information Visualization, and Human–Computer Interaction, participated in the study. These scientists, representing both industry and academia, provided valuable insights into the current and future landscape of the field.

In this report, we analyze the survey responses and compare them to the findings of a similar study conducted four years ago [2]. The results reveal some interesting insights. First, *many of the critical challenges identified in the previous survey remain highly relevant today*, despite being unrelated to AI. Meanwhile, the field's landscape has significantly evolved, with *most of today's vital challenges not even being mentioned in the earlier survey*, underscoring the profound impact of AI-related advancements.

By summarizing the perspectives of the research community, this report aims to shed light on the key challenges, emerging trends, and potential research directions in human–data interaction and visual analytics in the AI era.

2 Survey Overview

The survey is divided into two parts. The first is related to challenges (Sect. 3), and the second focuses on emerging research topics (Sect. 4).

The participants were requested to answer six questions, either by filling out free-text fields or selecting from the options provided. The survey was anonymous, since the questions related to personal information are optional, e.g., name, county, affiliation. The survey required, on average, about 3 to 5 minutes to be completed.

Participants Demographics. We intended to find scientists from different research communities (e.g., Databases, Information Visualization, HCI), and from industry and academia. To this end, the survey was disseminated to the BigVis 2024 Program Committee members (58 members) and to the authors of accepted BigVis 2024 papers (34 authors). At the end, 32 *of the scientists invited completed the survey*.

The following *characteristics of the participants are collected* (Fig. 1):

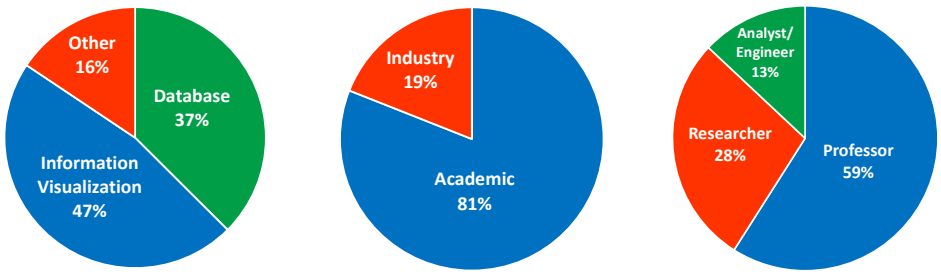
- **Scientific Field** (Fig. 1a): The options were: (a) *Database*; (b) *Information Visualization*; (c) *Data Minning*; (d) *Human–Computer Interaction*; (e) *Computer Graphics*; and (f) *Other*. Most of the participants belong to Information Visualization (47%) and Database (37%) communities, while 16% belong to others research fields.
- **Career** (Fig. 1b): The options were: *Academic* (81%) and *Industry* (19%).
- **Position** (Fig. 1c): The options were: *Professor* (59%); *Researcher* (28%); and *Analyst/Scientist/ Engineer* (13%).

3 Research Challenges

In this section we outline the survey's results regarding research challenges related to data visualization and visual analytics.

In the first part (Sect. 3.1), the participants were asked to vote on today's importance of the challenges stated four years ago in a 2020 report, titled "*Big Data Visualization and Analytics: Future Research*

*To appear in ACM SIGMOD Record, 2025
†Corresponding author.
[1] 7th *Intl. Workshop on Big Data Visual Exploration & Analytics*, in conjunction with the 50th *Intl. Conf. on Very Large Databases* (VLDB 2024), Guangzhou, China. More details about the BigVis workshops can be found in [7].



(a) Scientific Field (b) Career (c) Position

Figure 1: Participants Demographics

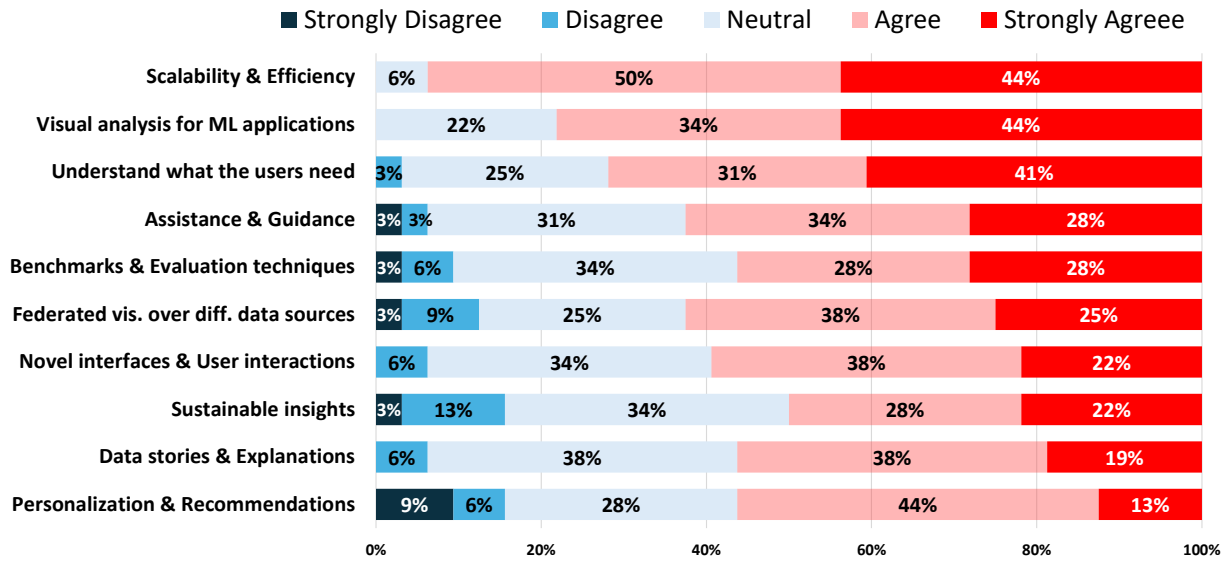


Figure 2: The Importance of the 2020 Challenges Today ["*Is this challenge important today*?"]

Challenges and Emerging Applications" [2]. In the next part (Sect. 3.2), participants suggested a challenge they consider the most important today, regardless of whether it was included in 2020 challenges.

3.1 The Challenges Stated in the 2020 Report

This section presents the results of the survey regarding today's importance of the ten challenges identified four years ago in the 2020 report. Particularly, in the context of the 3rd International Workshop on Big Data Visual Exploration and Analytics (BigVis 2020), the organizing committee invited 14 *distinguished scientists*, from different communities to provide their insights regarding the *challenges and the applications they find more interesting in coming years, related to the areas of Big Data visualization and analytics*.

The challenges indicated in the 2020 report were: (a) *Support scalability & efficiency*; (b) *Enable visual analysis for ML applications*; (c) *Understand what the users need*; (d) *Build novel interfaces & user interactions*; (e) *Assistance & guidance*; (f) *Generate data stories & explanations*; (g) *Enable federated visualization over different data sources*; (h) *Develop benchmarks & evaluation techniques*; (i) *Provide sustainable insights*; and (k) *Enable personalization & recommendations*.

Question 1

The participants were asked to vote on the *10 challenges stated in the 2020 report, based on their importance/emerge*. Particularly, the participants rated each challenge using a five-level Likert scale (i.e., Strongly Disagree, Disagree, Neutral, Agree, and Strongly Agree) on the question "*Is this challenge important today*?".

Question 1 Responses. The results are presented in Figure 2 via a percent stacked bar chart. Participants voted "*Scalability & efficiency*" as the most important challenge", with 94% of participants indicating they agree or strongly agree, and 0% disagree or strongly disagree. The second most significant is the "*Visual analysis for ML applications*" challenge, with 78% (resp. 0%) of the participants state that agree or strongly agree (resp. disagree or strongly disagree). "*Sustainable insights*" is the challenge in which most of the participants disagree (13%) or strongly disagree (3%). Finally, "*Personalization & recommendations*" is voted as the least important challenge.

It is worth mentioning that for *all challenges, at least* 50% *of the participants strongly agree or agree on today's importance of the challenge. Similar results can be observed when considering importance scores*.

3.2 The Challenges Stated by the Survey Participants

This section presents the challenges stated by the participants as the most important, regardless of whether they were included in 2020 challenges.

Question 2

The participants were asked to provide in a free text the *challenge they consider most important for the coming years*, along with a brief description.

Question 2 Responses. The participants *indicate* 16 *challenges*. The challenges are presented in Table 1; the number in the parentheses that appears in some challenges *indicates the number of participants that mention this challenge*. Furthermore, *red font highlights the challenges that are mentioned for the first time in this survey*, i.e., challenges that were not mentioned in the 2020 survey.

The most commonly suggested challenge is the "*Use of LLMs in visualization and analytics*" (voted by 16% of the participants), whereas "*Fairness and Trustworthiness*", as well as "*Visualization for non-expert users*" are the next most common (each voted by 13%). Note that none of the terms LLMs, fairness or trustworthiness is mentioned in the four years ago challenges. Also note that explainability, which emerged as one of the most frequently mentioned challenges, is also not mentioned in the 2020 report.

Other common challenges (mentioned by at least two participants) are related to: "*User assistance & guidance*"; "*Understanding what the users need*"; "*High dimensional & stream data*"; "*Progressive data analysis & visualization*; and "*Immersive visualization*". Among these challenges, "*High dimensional & stream data*", "*Progressive analysis*", and "*Immersive visualization*" appeared for the first time.

Table 1: Survey Challenges ⋆

Exploit LLMs$^{(5)}$, Ensure fairness & trustworthiness$^{(4)}$, Enable visualization for non-expert users$^{(4)}$, Offer assistance & guidance$^{(2)}$, Generate explanations$^{(2)}$, Understand what the users need$^{(2)}$, Handle high dimensional & stream data$^{(2)}$, Enable progressive data analysis & visualization$^{(2)}$, Develop immersive visualization systems$^{(2)}$, Provide sustainable insights, Support data abstraction, Implement novel scalable interfaces, Design context-specific visualizations, Formulate fundamental visualization problems, Use surrogate modeling, Develop energy consumption-based solutions

⋆ c$^{(x)}$: x indicates the number of participants that mention the challenge c in the survey. Red font: Challenges mentioned for the first time in the current survey (considering only those indicated by at least two participants).

4 Emerging Topics

In this section we present the participants' responses regarding the most emerging research topics in Big data visualization and analytics field. The candidate topics list consists of the topics of interest included in the BigVis call-for-papers.

Question 3

The participants were asked to *select (vote) from a list of candidate topics, up to three topics that they consider the most emerged.*

Question 3 Responses. Figure 3 shows the percentage of the participants' vote for each topic. The topics with the most votes are "*Human–in–the–loop processing*"; "*Interactive & human–centered ML*"; and "*Progressive analytics*", where 31% of the participants select. On the other hand, the topics with the less votes are: "*Scientific visualization*"; *"Setting-oriented visualization*"; and *"Distributed & parallel techniques*", which are voted by 3% of the participants.

5 Discussion

First, the survey *highlights the broad acceptance of the importance of all the challenges identified four years ago* (Question 1), with at least half of the participants strongly agreeing or agreeing on their today's importance.

The results regarding the current challenges (Question 2) revealed the importance of AI–related problems. Notably, *the most frequently mentioned challenges today were entirely absent four years ago*.

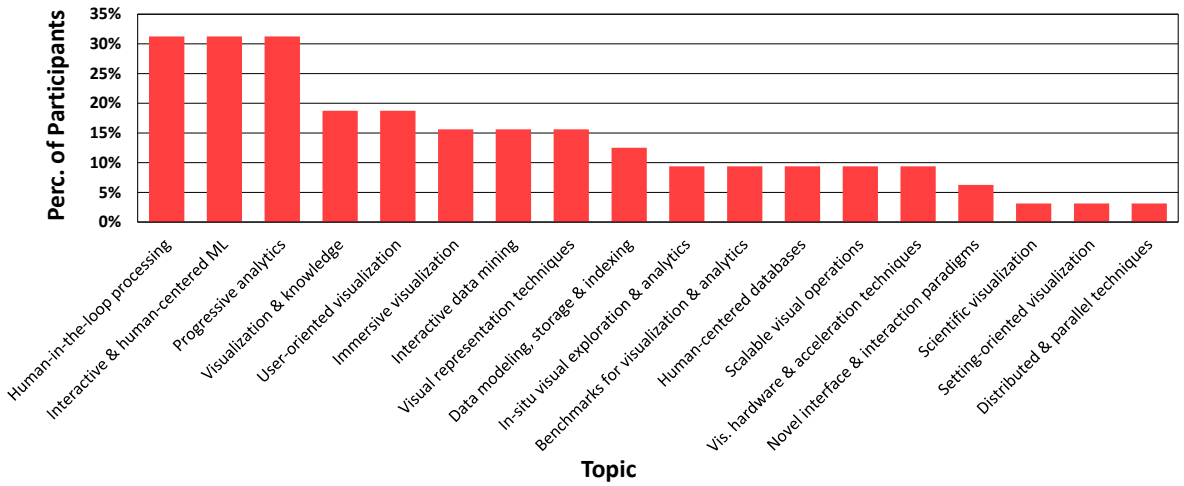


Figure 3: Emerging Topics: The Percentage of Participants that Voted for Each Topic

For example, problems related to LLMs, fairness & trustworthiness, and explanations are some of the newcomers. Furthermore, challenges such as "*Non-expert users*", "*High dimensional & stream data*", "*Progressive analysis*", and "*Immersive visualization*" appeared also for the first time.

Further comparison of responses reveals that "*Scalability & efficiency*", the most important challenge in 2020 (Question 1), was not mentioned in Question 2. One possible explanation is that nearly 95% of participants had already rated it in Question 1 as (very) important, reducing the need to highlight it again. Similarly, "*Visual analysis over ML applications*", the second most important challenge in 2020, was absent from Question 2 responses, despite being recognized as one of the most emerging topics (Question 3).

Additional discussion of the current challenges and state-of-the-art approaches can be found in [1, 3–6, 8–15].

Acknowledgments. This work has been partially supported by project MIS 5154714 of the National Recovery and Resilience Plan Greece 2.0 funded by the European Union under the NextGenerationEU Program.

References

[1] Sihem Amer-Yahia, Leilani Battle, Yifan Hu, Dominik Moritz, Aditya Parameswaran, Nikos Bikakis, Panos K. Chrysanthis, Guoliang Li, George Papastefanatos, and Lingyun Yu. 2025. Data Exploration and Visual Analytics Challenges in AI Era. ACM SIGMOD Blog, https://wp.sigmod.org/?p=3820.

[2] Gennady L. Andrienko, Natalia V. Andrienko, Steven Mark Drucker, Jean-Daniel Fekete, Danyel Fisher, Stratos Idreos, Tim Kraska, Guoliang Li, Kwan-Liu Ma, Jock D. Mackinlay, Antti Oulasvirta, Tobias Schreck, Heidrun Schumann, Michael Stonebraker, David Auber, Nikos Bikakis, Panos K. Chrysanthis, George Papastefanatos, and Mohamed A. Sharaf. 2020. Big Data Visualization and Analytics: Future Research Challenges and Emerging Applications. In *Workshop on Big Data Visual Exploration & Analytics (BigVis 2020)*.

[3] Natalia V. Andrienko, Gennady L. Andrienko, Linara Adilova, Stefan Wrobel, and Theresa-Marie Rhyne. 2022. Visual Analytics for Human-Centered Machine Learning. *CG&A* (2022).

[4] Rahul C. Basole and Timothy Major. 2024. Generative AI for Visualization: Opportunities and Challenges. *TVCG* (2024).

[5] Leilani Battle and Carlos Scheidegger. 2020. A structured review of data management technology for interactive visualization and analysis. *IEEE TVCG* 27, 2 (2020).

[6] Nikos Bikakis. 2022. Big Data Visualization Tools. In *Encyclopedia of Big Data Technologies, 2nd Ed.* Springer.

[7] Nikos Bikakis, George Papastefanatos, Panos K. Chrysanthis, Olga Papemmanuil, David Auber, Steffen Frey, Issei Fujishiro, Hanna Hauptmann, Shixia Liu, Kwan-Liu Ma, Tobias Schreck, Michael Sedlmair, and Mohamed A. Sharaf. 2024. Visualizing, Exploring and Analyzing Big Data: A 6-Year Story. *ACM SIGMOD Record* 53, 2 (2024).

[8] Muhammad Raees, Inge Meijerink, Ioanna Lykourentzou, Vassilis-Javed Khan, and Konstantinos Papangelis. 2024. From Explainable to Interactive AI: A Literature Review on Current Trends in Human-AI Interaction. *J. Hum. Comput. Stud.* (2024).

[9] Junpeng Wang, Shixia Liu, and Wei Zhang. 2024. Visual Analytics for Machine Learning: A Data Perspective Survey. *IEEE TVCG* 30, 12 (2024).

[10] Qianwen Wang, Zhutian Chen, Yong Wang, and Huamin Qu. 2022. A Survey on ML4VIS: Applying Machine Learning Advances to Data Visualization. *IEEE TVCG* 28, 12 (2022).

[11] Aoyu Wu, Dazhen Deng, Min Chen, Shixia Liu, Daniel A. Keim, Ross Maciejewski, Silvia Miksch, Hendrik Strobelt, Fernanda B. Viégas, and Martin Wattenberg. 2023. Grand Challenges in Visual Analytics Applications. *IEEE CG&A* (2023).

[12] Aoyu Wu, Yun Wang, Xinhuan Shu, Dominik Moritz, Weiwei Cui, Haidong Zhang, Dongmei Zhang, and Huamin Qu. 2022. AI4VIS: Survey on Artificial Intelligence Approaches for Data Visualization. *IEEE TVCG* 28, 12 (2022).

[13] Weikai Yang, Mengchen Liu, Zheng Wang, and Shixia Liu. 2024. Foundation models meet visualizations: Challenges and opportunities. *Comput. Vis. Media* 10, 3 (2024).

[14] Yilin Ye, Jianing Hao, Yihan Hou, Zhan Wang, Shishi Xiao, Yuyu Luo, and Wei Zeng. 2024. Generative AI for visualization: State of the art and future directions. *Vis. Informatics* 8, 1 (2024).

[15] Jun Yuan, Changjian Chen, Weikai Yang, Mengchen Liu, Jiazhi Xia, and Shixia Liu. 2021. A survey of Visual Analytics Techniques for Machine Learning. *Comput. Vis. Media* 7, 1 (2021).