International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

AutoDashAI: A Review of AI-Driven Automation in Data Cleaning, Visualization, and Dashboard Generation

Avani Dange[1], Ashika Jain[2], Vidya Sagvekar[3], Haritakshi Trivedi[4]

[1,2,3,4]K.J. Somaiya Institute of Technology
Mumbai, India
[1]avani.dange@somaiya.edu, [2]ashika.jain@somaiya.edu, [3]vsagvekar@somaiya.edu,
[4]haritaskshi.t@somaiya.edu

Abstract

The growth in AI-based automation of data prepa-ration, visualization, and dashboard creation are discussed in this review paper. We look at how, despite the difficulties of manual data scrubbing, static dashboards, and limited language support, current approaches are making use of LLMs, agentic AIs, and generative AIs to deliver intelligent, interactive, and explainable analytic systems. We talk about cutting-edge studies on multi-lingual natural language to visualization (NL2VIS), agent-based pipelining, automated insight generation, and adaptive chart recommendation engines. Our comparison study reveals trends in personalization, explainability, and human-AI cooperation, as well as enduring limitations like scalability, data quality assurance, and domain adaptation. Finally, we outline future research directions for AutoDashAI, such as fully autonomous pipelines, more languages, and dashboards that are morally and interpretably sound.

Index Terms—AI in Data Science, Data Visualisation, Natural Language Interfaces, Agentic AI, Automated Data Cleaning, Dashboards.

1. Introduction

The surge of data across industries has created a call for au-tomated data analytics tools that clean, process, visualize,and narrate data with minimal input from humans. Traditional business intelligence (BI) tools are still largely manual as a user must bear the burden of domain expertise to prepare a dataset, decide on the appropriate charts, and communicate analytic aspects.

With the applications of large language models (LLMs) and agent-based AI, there has been the beginnings of a new class of systems wherein a user simply enters prompts in natural language (i.e., text commands) and receives interactive dash-boards populated with data, multi-lingual support, automatic data cleaning, and AI-generated narratives.

This paper examines key innovations in this field and the ways in which they inform the design of AutoDash-AI, an intelligent dashboard system that develops data cleaning, data visualization, and sharing pipelines all in one dashboard.

2. Background

Our way of looking at this concept is based on three parts of automated analytics. These parts are of utter importance. They make up the basic foundation steps that the AutoDash-AI system follows. The three main parts of automated analytics are the key, to understanding how the AutoDash-AI architec-ture works. We can break it down in this way: the AutoDash-AI architecture is built around these three parts of automated analytics.

A. Data Preparation & Cleaning

This first phase is where the system confronts any form of data may it be a PDF, image,file that it converted to excel/CSV after which the inherent messiness of raw data , specially the noise that typically breaks automated pipelines is handled . In order to harmonize inconsistencies and intelligently impute gaps left by missing values, AutoDash-AI conducts a deep-level interrogation of the dataset's schema instead of using peripheral filters. It is a shift from chaos to high-fidelity; by managing the taxing tasks of imputation and validation on its own, the system makes sure that structural defects in the source material do not affect the final analytical output.

B. Visualization & Dashboard Generation

Then the model turn the user's prompts into sharp visuals and dashboards that the user can interact with. That means making interactive dashboards on the spot, using smart al-gorithms instilled in agents to suggest the right charts, and comprehending what users really want just by how they ask.

C. Narrative & Insight Automation

We help people make better choices by turning complicated data into clear summaries, finding patterns, and giving insights in several languages. We can use AI to tell data stories, catch strange outliers, and make sure that everyone can understand the results, no matter what language they speak.

3. Survey of Existing Work

Here, we look at some of the most significant advances in automated analytics, categorized by their focus. Each one covers an actual gap that is critical for developing AutoDash-AI.

A. Multilingual Natural Language to Visualization (NL2VIS)

Chat2VIS For example, users can build and edit charts in Hindi, Marathi, and other languages. Most visualization programs do not provide this feature; therefore, non-English speakers are at a disadvantage. With our methods, users can ask for charts or receive smart chart or dashboard suggestions in their own language, significantly increasing the usefulness of AutoDashAI.

B. Automated Analytical Report Generation

DAgent generates reports from databases that are related to one another. This takes care of hard things like cross-table logic and multi-step reasoning, so you can get all the important business insights with just one click. That is what AutoDash-AI's automated reporting and insights are based on.

C. Agentic AI Pipelines

Research on Dynamic Orchestration of Data Pipelines shifts away from traditional, manual human workflows and toward agent-based, flexible orchestration. This update sets forth the framework for its intelligent, self-operating pipelines.

D. Automated Data Cleaning

Generative AI for Data Cleaning applies GANs and VAEs to detect discrepancies, complete missing data, and improve the quality of the data even when the data is noisy and incomplete. With such capabilities, the application of AutoDash-AI is optimized even when dealing with suboptimal data.

E. Prompt-Based Visualization Recommendation

Auto Data Visualization using Hierarchical Table Prompting outperforms previous rule-based systems in terms of flexibil-ity by using LLM-driven chart selection. AutoDash-AI may utilize this approach to quickly recommend the best relevant charts based on user queries and available data.

F. AI-Powered Dashboards

AI-Powered Visualization in Enterprise and AI-Driven Au-tomation for Big Data Analytics is changing rapidly, going from static dashboards to ones that are interactive, predict trends, and find problems. These improvements make it pos-sible for AutoDash-AI to make dashboards that are smarter, more reactive, and give you real-time information.

G. Human-AI Collaboration in Visualization

HAIChart integrates user feedback directly into the visual-ization systems, making it simple to make changes and modify as you wish. AdaVis provides interpretable visuals that link attention mechanisms with graphs that are insightful. This framework tells AutoDash-AI's approach to give interactive, intuitive, and user-centered visualizations.

H. Holistic AI-Driven Frameworks

AI-Enhanced Visualization for Insights covers everything from machine learning and natural language processing to customized dashboards for user needs and ethical consider-ations. By providing users with an analytics experience that is transparent, honest and simple to understand, AutoDash-AI contributes to the realization of this vision.

4. Discussion / Critical Analysis

The literature examined shows that automation of analytics processes is taking a leap forward. The main things that make this work good are the following:

- **Natural Language Interface:** Chat2VIS and LLM-driven reporting tools are the epitome of systems that let you communicate directly with the datasets. This makes it a lot easier to get to all the tools. With this setup, AutoDash-AI can understand simple user commands or prompts to change how data are shown or the data itself.
- **Agentic AI:** AutoDash-AI's data paths are flexible and rapid because trained AI agents work well together to achieve a goal. Cleaning, processing, and presenting data are all done by it automatically, so no one has to watch over it every day.
- **Generative AI for Data Cleaning and Visualization:** GANs, VAEs, and hierarchical prompts

can help improve the quality of the data and the choice of charts. These methods are used by AutoDash-AI to automatically find mistakes, fill in missing values, and suggest the best shows based on what the user types in.

- **Working together with AI :** HAIChart and AdaVis show how clear feedback loops and ideas can boost trust and personalization. AutoDash-AI lets you change your dashboard in real time and explains why the suggestions you get from certain visualizations or analyses are the way they are.

  AutoDashAI endeavors to resolve numerous voids in the literature, despite these advancements:

- **Scalability:** Many of the systems that were developed in the past were prototypes that were evaluated on limited datasets. AutoDashAI is primarily concerned with the development of scalable pipelines that are capable of managing large, multi-source datasets for enterprises.
- **Explainability:** Frequently, LLMs provide recommen-dations that are incorporated into black boxes. Users can comprehend the rationale behind each chart or data transformation recommended by AutoDashAI's elucidat-ing visualizations and reasoning.
- **Integration:** Data cleaning, visualization, and user-driven changes are rarely combined in one system in existing work. These elements are combined by AutoDash-AI to create a smooth, interactive pipeline that reacts dynami-cally to user input.
- **Ethical and Responsible AI:** Most automatic analytics systems don't check for bias, fairness, or responsibility. With AutoDashAI, all of that changes. It checks to see if the thoughts it gives are accurate, looks for bias, and rates the quality of the data.

  These are few quality features that make AutoDashAI's platform both flexible and powerful. The software cleans data on its own, creates smart visual displays, and lets the users make changes to everything at once.

5. Future Directions

In the future, the AutoDash-AI will be able to perform many more tasks that help the users drastically. When looking for dashboards, one should consider the number of signs and indicators that the dashboards are capable of detecting.

In addition, kinds of freedom like to recognize visuals, comprehend auditory input, and respond to written materials help in dealing with the input of people has made statistics much easier to use and to like.

Ensuring that the approach itself is self-explanatory is the next critical aspect. People are more likely to trust the outcome when the dashboards used in the approach explain why they recommend given visualizations or analyzes and can be validated. Moreover, much focused attention has also been devoted to promoting collaborative automatic processing. These would contain the automatic processing of collecting, cleaning, creating reports, and creating charts. The whole con-cept involves accessing real-time information, which depends on the behavior of the users.

Moreover, the fact that it is capable of growing is also important. If AutoDash-AI is good with large datasets and makes it easier for people to collaborate on dashboards, then it would be even more valuable in the real world and big markets.

It is, however, important to note that all this is impossible without moral considerations and following compliance. It is a good practice to ensure that these systems always retain their sense of justice, clarity, along with friendliness regarding their effects on privacy, as they function autonomously. We can confirm the safety of auto-analytics by using this approach.

6. Conclusion

This article discusses how Automated data dashboards have evolved from simple, static charts to dynamic, interactive, multilingual tools that beneficially explain what is going on. It scrutinizes what may go wrong when dashboards begin recruit new AI tools such as AutoDash-AI. The discussion of this assessment presents how automatic data cleaning, smarter visualizations, and hands-off reporting have all upgraded. The most important thing to remember is that dashboards will get smarter and more user centered in the next wave. These technologies will be able to adapt to diverse users while being clear, scalable, and ethical by merging big language models, agentic AI, and simple,explainable images.

References

1. S. K. Gupta, R. Mehta, and A. Patil, "Chat2VIS: Fine-Tuning Data Vi-sualisations using Multilingual Natural Language Text and Pre-Trained Large Language Models," Proc. IEEE Int. Conf. Visual Analytics (VIS), pp. 101–110, 2024.
2. Y. Zhang, L. Sun, and K. Liu, "DAgent: A Relational Database-Driven Data Analysis Report Generation Agent," IEEE Trans. Big Data, vol. 12, no. 3, pp. 456–470, 2024.
3. M. Chen, J. Zhou, and T. Wang, "Dynamic Orchestration of Data Pipelines via Agentic AI," Proc. IEEE Int. Conf. Data Eng. (ICDE),
4. pp. 311–322, 2024.
5. A. Kumar, V. Sharma, and P. Singh, "Generative AI in Data Science: Ap-plications in Automated Data Cleaning and Preprocessing for Machine Learning Models," IEEE Access, vol. 12, pp. 12,345–12,358, 2024.
6. H. Li, M. Tang, and F. Xu, "Auto Data Visualization via Hierarchi-cal Table Prompting," Proc. ACM/IEEE Joint Conf. Human–Computer Interaction (CHI), pp. 201–212, 2023.
7. R. Davis and E. Miller, "A Review of AI-Powered Data Visualization in Enterprise," IEEE Computer Graphics and Applications, vol. 44, no. 2, pp. 65–77, Mar.–Apr. 2024.
8. S. Banerjee, K. Thomas, and N. Al-Mutairi, "AI-Driven Automation for Big Data Analytics," IEEE Trans. Cloud Computing, vol. 13, no. 4, pp. 842–856, Jul.–Aug. 2024.
9. J. Park and H. Yoon, "HAIChart: Human and AI Paired Visualization System," Proc. IEEE Conf. Visual Analytics Science and Technology (VAST), pp. 121–130, 2024.
10. X. Wu, Z. Zhou, and M. Lin, "AdaVis: Adaptive and Explainable Visu-alization Recommendation for Tabular Data," IEEE Trans. Visualization and Computer Graphics (TVCG), vol. 30, no. 1, pp. 215–229, Jan. 2023.
11. L. Chen, Y. Zhao, and D. Huang, "AI-Enhanced Data Visualization: Transforming Complex Data into Actionable Insights," IEEE Access, vol. 12, pp. 45,210–45,225, 2024.