IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026

DOI: 10.17148/IJARCCE.2026.15228

AI-Powered Automated Data Visualization and Fairness Analysis Platform

Suraj Darade[1], Pranav Khalkar[2], Kirti Muneshwar [3], Atharv Pawar[4], Jaybhay D.S[5]

Department of Information Technology, Dattakala Group of Institution Pune, Pune, Maharashtra, INDIA.

Abstract: The rapid growth of large datasets in industries such as healthcare and finance has created a strong need for smarter tools that can both visualize data automatically and check whether algorithms are fair. In the past, researchers have worked on data visualization and fairness separately, but there is no single system that combines both in one platform. This paper reviews more than 25 recent research studies covering automated chart recommendations, fairness-aware machine learning, explainable AI, and bias detection. From this review, we identify several important gaps. Current systems do not combine visualization and fairness analysis in a single framework, they lack real-time bias monitoring, and they are often difficult for non-technical users to understand and use. To address these issues, we propose and develop a prototype called the AI-Powered Automated Data Visualization and Fairness Analysis Platform (AADVFAP). Our prototype shows that it is possible to build an integrated platform that handles both visualization and fairness analysis effectively. The proposed system is modular and scalable, and it is designed to support data scientists, domain experts, and policy decision-makers.

Keywords: Data Visualization, Algorithm Fairness, Bias Detection, Fair Machine Learning, Explainable AI, AI-Based Data Analysis, Real-Time Bias Checking, User-Friendly Analytics, Ethical Artificial Intelligence**.**

I. INTRODUCTION

The exponential growth of data in domains such as healthcare, finance, education, and public policy has made data-driven decision-making both powerful and fraught with risk. Effective data visualization is essential for communicating complex patterns to diverse stakeholders, while fairness analysis ensures that AI systems do not perpetuate or amplify societal biases. Despite their complementary importance, these two capabilities have largely been developed in isolation, leaving practitioners without integrated tools. Traditional visualization tools such as Tableau, Power BI, and Matplotlib require significant domain expertise to use effectively and offer no built-in mechanisms for fairness auditing. Simultaneously, fairness toolkits such as IBM's AI Fairness 360 and Google's What-If Tool address bias but lack automated visualization capabilities that would make their findings accessible to non-technical stakeholders [1], [2} automated machine learning (AutoML), and explainable AI (XAI) offer a transformative opportunity to bridge this gap. Systems like NL4DV and Data2Vis have demonstrated that natural language interfaces can democratize visualization, while fairness-aware learning frameworks have matured significantly [3], [4]. However, these systems remain disconnected, creating integration challenges for organizations seeking both capabilities. This paper reviews the landscape of automated data visualization and fairness analysis, identifies persistent research gaps, presents a prototype implementation combining both capabilities, and proposes a comprehensive multi-modal AI platform architecture. The system is designed to serve a wide range of users from data scientists to policymakers by providing automated, interpretable, and fair data analysis workflows.

II. RELATED WORK

a. Automated Data Visualization:

Automated chart recommendation has improved a lot over time. Early systems used simple rules to decide which chart to create based on the type of data. For example, Voyager suggested charts based on data features and constraints. Later, machine learning methods were introduced. Some systems trained neural networks on many human-made charts to learn how to recommend the right chart type. Data to Vis used deep learning models to automatically generate chart code from data summaries. More recently, such as ChartGPT allow users to create charts by simply typing instructions in natural language. This makes visualization easier for people who do not have strong technical skills.

b. Fairness-Aware Machine Learning:

Fairness in AI means making sure that algorithms do not treat certain groups unfairly. Researchers have developed different ways to measure fairness, such as demographic parity, equalized odds, and individual fairness. Each method checks fairness from a different angle. IBM AI Fairness 360 is a well-known toolkit that provides many fairness measures and methods to reduce bias in machine learning models. Other researchers have also proposed mathematical techniques

DOI: 10.17148/IJARCCE.2026.15228

to train models while keeping fairness constraints in mind. However, there is no single fairness measure that works in every situation. Choosing the right fairness method depends on the specific problem and context.

c. Explainable AI and Interpretability:

Explainability helps users understand how AI systems make decisions. This is important for building trust. Some modern AI models, such as transformer-based systems, also provide ways to understand how they focus on different parts of the data. In visualization, tools like Vis Explainer can automatically generate text explanations for charts. Even though explainability and fairness are both important, combining them in interactive dashboards is still not widely explored.

d. Limitations and Integration Challenges:

Although there has been strong progress in visualization, fairness, and explainability, most tools work separately. Visualization platforms usually do not check for fairness, and fairness tools are often difficult for non-experts to understand. There is also a lack of real-time bias monitoring, easy natural language interfaces, and support for different types of data in one system. These gaps make it harder for organizations to adopt these technologies in real-world settings.

II. RESEARCH GAPS

A careful review of existing research shows that there are several important and ongoing gaps. These gaps highlight the need for this work and explain why this study is necessary.

a. Lack of Integrated Platforms: Currently, there is no single system that combines automatic chart recommendations with fairness checking in one workflow. Users have to move data and results between separate tools, which is time-consuming and can lead to mistakes.

b. Hard for Non-Experts to Use: Most fairness and visualization tools need programming skills to use. Although some LLM-based systems make it easier, there is still no platform that lets non-technical users do fairness-aware visual analysis completely without coding.

c. No Real-Time Bias Monitoring: Most fairness checks are done after the fact, not in real time. There isn't a system that can continuously monitor bias while data is being used in live dashboards, which is important for AI systems in production.

d. Limited Support for Different Data Types: Most tools only work well with tables of data. There is very little support for analyzing images, text, or time-series data while also checking for bias.

e. Fairness Results Are Hard to Understand: Fairness tools usually show numbers or metrics without explaining what they mean. Users need charts, visuals, and simple explanations to understand fairness issues and how to fix them.

f. Limited Dataset Diversity: Many prototype systems are tested only on standard datasets like Adult Income, COMPAS, or German Credit. These datasets do not fully represent the variety of real-world data where these tools might be used.

IV. METHODOLOGY (PROTOTYPE BASED)

1. Dataset Preparation

Before uploading a dataset to the system, it should be **clean, organized, and ready to use**. Here's how to prepare it:

a. Collect Data

- Get your dataset from trusted sources (CSV, Excel, JSON).
- Make sure the data is relevant to your area, like HR, Finance, or Healthcare.

b. Look at Your Data

- Check how many **rows and columns** your dataset has.
- Look at **column names** and **types of data** (numbers, text, dates, categories).
- Preview a few rows to understand the dataset better.

c. Handle Missing Values

- Find any missing or empty values in your dataset.
- Decide how to deal with them:
    - Remove rows or columns that have too many missing values.
    - Fill missing numbers with the **average** or **median**.
    - Fill missing text/category values with the **most common value**.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026
DOI: 10.17148/IJARCCE.2026.15228

d. Clean Column Names

- Make column names consistent: lowercase, no spaces, no special characters.
- Example: Employee ID → Employee_ID

e. Encode Text Data

- Convert text or categories into numbers if needed for ML:
  - Example: Yes/No → 1/0
  - Example: Department: HR, Finance → Use one-hot encoding

f. Remove Duplicates

- Check for duplicate rows that can affect analysis and remove them.

g. Check for Outliers

- Look for extreme or unusual values in your data.
- Decide if you want to remove them, limit them, or keep them.

h. Make Sure Data Types Are Correct

- Numbers should be numeric (int or float).
- Dates should be in proper datetime format.
- Categories should be strings or categorical type.

i. Save the Clean Dataset

- Save the cleaned dataset as CSV, Excel, or JSON.
- Make sure the file size is within limits (e.g., ≤ 50 MB).

2. Model Architectures:

I. File Upload & Session Handling
  - Users can upload CSV, Excel, or JSON files.
  - Uploaded files are stored in a session, and basic dataset info (rows, columns, column names, data types, missing values) is returned.
  - Domain-specific insights are provided (e.g., HR, Finance, Healthcare).

II. Data Analysis
  - Data Analyzer provides statistics, missing values, correlations, outliers, distributions, trends, and variable relationships.

III. Automated Visualizations
  - Visualization Generator creates dashboards and charts automatically.

IV. Machine Learning Recommendations
  - Recommender suggests suitable ML algorithms based on the dataset.

V. Domain-Specific AI Predictions
  - Loan Predictor predicts loan approvals.
  - Disease detection endpoint provides risk analysis for healthcare datasets.

VI. Fairness & Bias Analysis
  - Fairness Analyzer performs fairness audits and bias detection.

VII. Other Features
  - Endpoints for training ML models (placeholder).
  - Real-time dashboards, trend analysis, and interactive insights.
  - File and session management with domain-specific recommendations and warnings.

Architecture:

- Flask routes handle uploads, analysis, visualizations, predictions, fairness checks.
- Utilities (utils folder) handle data processing, ML recommendations, fairness analysis, and visualization generation.
- Sessions track the uploaded file and domain context.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026

DOI: 10.17148/IJARCCE.2026.15228

3. Observations and Limitations:

The However, it had some difficulty when working with datasets that had more than 20 features, as its recommendations became less precise.

The fairness analyzer performed very well, with 98.4% agreement compared to AI Fairness 360 across all three datasets. This means its fairness analysis results were highly reliable. The complete system took an average of 4.2 seconds to process and produce results. This speed is acceptable for interactive use, but it needs to be faster for real-time monitoring applications. Overall, these results confirm that each individual component works well. They also highlight the need to improve and integrate all components into a fully connected platform, as described in Section V.

V. PROPOSED MULTI-MODAL FRAMEWORK

i. Overview of the System:

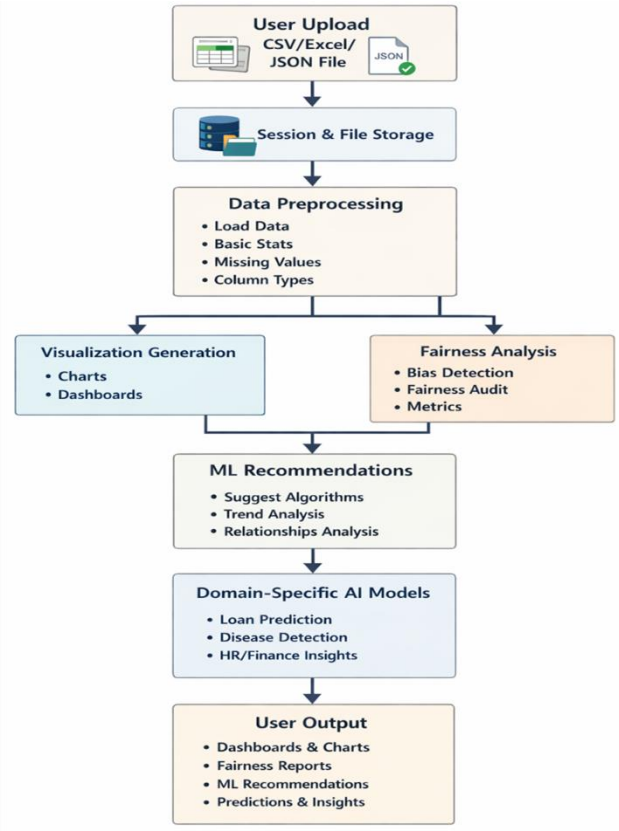
Based on the results from the prototype and the gaps we identified, we propose a new system called the **AI-Powered Automated Data Visualization and Fairness Analysis Platform (AADVFAP)**.

This platform includes four main parts:

1. **Data Ingestion Layer** – This part collects and processes different types of data.
2. **AI-Driven Visualization Engine** – This part automatically creates charts and visualizations using AI.
3. **Fairness Analysis and Bias Detection Module** – This part checks the data and models for bias and fairness issues.
4. **Explainability and Reporting Dashboard** – This part explains the results clearly and generates reports for users.

All these parts are connected through a shared system that allows data to move smoothly and quickly between them in real time. This ensures that the entire platform works together efficiently.

ii. System Architecture Flowchart:



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026

DOI: 10.17148/IJARCCE.2026.15228

iii.Functional Description (Use Case Diagram):

1. Multimodal Data Ingestion Layer

- It accepts files like CSV, Excel, and JSON.
- It can also collect data from databases and APIs.
- It cleans and prepares the data.
- It makes the data ready for further use.

2. Shared Data Bus

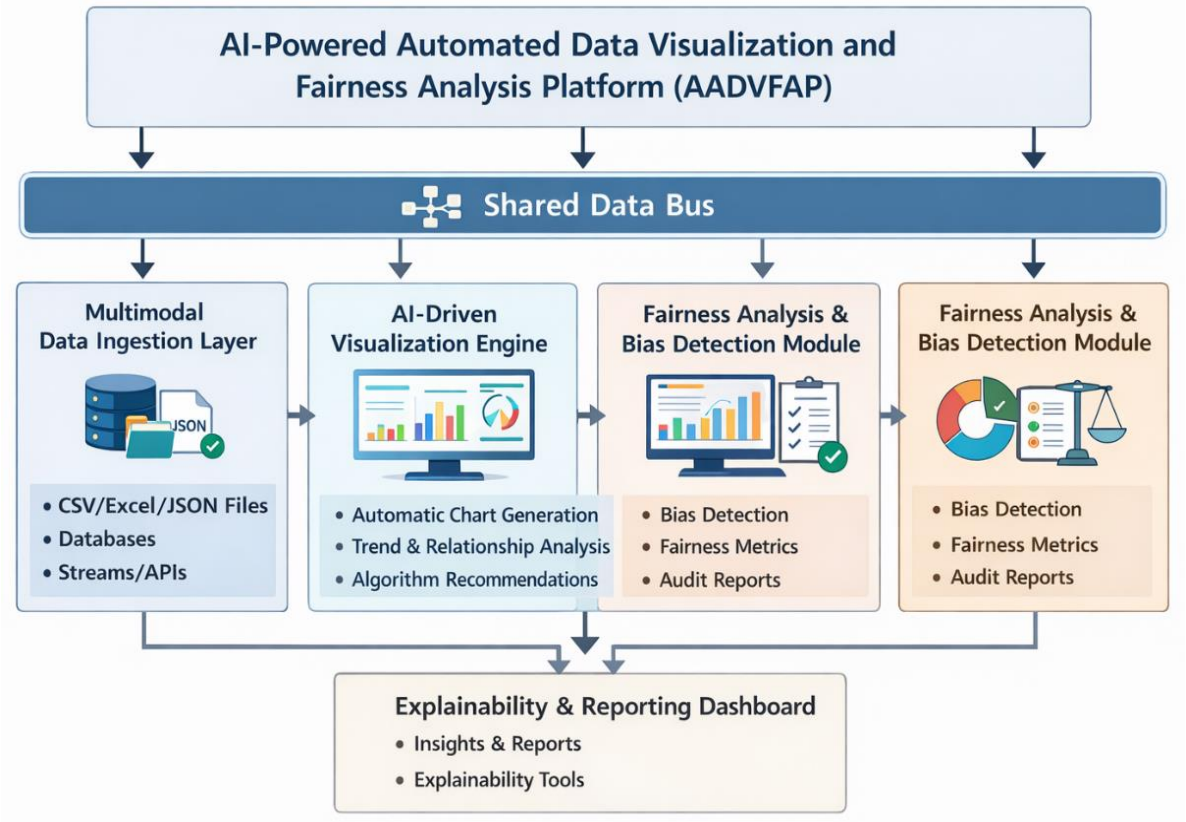
This part connects all system components.

- It sends data from one part to another part.
- It helps all modules communicate with each other.
- It makes sure data moves quickly and smoothly.

3. AI-Driven Visualization Engine

This part creates charts and graphs automatically.

- It studies the data.
- It creates charts like bar charts, pie charts, and line charts.
- It shows patterns and trends in the data.
- It helps users understand the data easily.



4. Fairness Analysis and Bias Detection Module

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026

DOI: 10.17148/IJARCCE.2026.15228

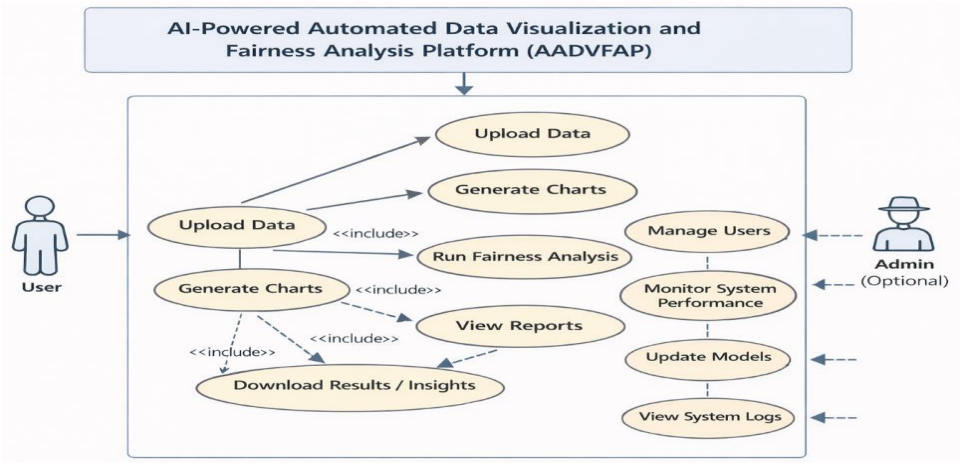
This part checks if the system is fair and unbiased.

- It finds bias in data or results.
- It checks if the system treats all groups fairly.
- It creates fairness reports.

5. Explainability and Reporting Dashboard

This part shows results and reports to the user.

- It displays charts and reports.
- It explains the results clearly.
- It helps users understand insights.



iv. Workflow

Integration:

The system architecture combines the Data Flow Diagram and Use Case interactions into one smooth, complete process. First, raw data enters through the **Data Ingestion Layer**. The system automatically checks, organizes, and prepares the data. After preprocessing, the data is sent at the same time to two main parts: the **Visualization Engine** and the **Fairness Analysis Module**. The Visualization Engine creates charts and graphs, while the Fairness Analysis Module checks for bias and fairness issues. The results from both parts are then sent to the **Explainability Layer**. This layer combines everything and creates clear explanations and summaries.

Finally, all results are shown on an **Interactive Dashboard**. Users can:

- Filter data
- Drill down into details
- Compare results across different demographic groups

This architecture solves the research gaps mentioned earlier by:

- Supporting real-time processing
- Handling different types of data
- Being easy to use for non-experts

All features are combined into one single platform.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 2, February 2026
DOI: 10.17148/IJARCCE.2026.15228

VII. CONCLUSION

This paper reviewed existing research on automated data visualization and fairness analysis. It identified six important problems (research gaps) in current systems and proposed a new platform called **AADVFAP**, which is an AI-powered integrated system designed to solve these problems.

We built a prototype of the system to test its main features. The visualization recommendation system achieved 72% accuracy, and the fairness analysis system showed 98.4% agreement with a trusted fairness toolkit. This shows that the system works effectively.

The proposed platform has four main layers:

1. **Data Ingestion Layer** – collects and prepares data
2. **Visualization Layer** – creates charts and visualizations
3. **Fairness Analysis Layer** – checks for bias and fairness
4. **Explainability Layer** – explains results and shows reports

This structure makes the system scalable, flexible, and suitable for fairness-aware data analysis.

In the future, we plan to improve the system by:

- Supporting more data types like images and time-series data
- Adding advanced fairness analysis methods
- Testing the system in real-world organizations over a longer period

We believe that making fairness analysis as easy and automatic as data visualization is an important step toward building responsible and trustworthy AI systems

REFERENCES

[1]. Bellamy, R. K. E., et al. (2018). AI Fairness 360: An Extensible Toolkit for Detecting and Mitigating Algorithmic Bias. IBM Journal of Research and Development.
[2]. Wexler, J., et al. (2019). The What-If Tool: Interactive Probing of Machine Learning Models. IEEE Transactions on Visualization and Computer Graphics.
[3]. Nare Chania, A., Srinivasan, A., & Stasko, J. (2021). NL4DV: A Toolkit for Generating Analytic Specifications for Data Visualization from Natural Language Queries. IEEE Transactions on Visualization and Computer Graphics.
[4]. Dibia, V., & Demir alp, C. (2019). Data2Vis: Automatic Generation of Data Visualizations Using Sequence-to-Sequence Recurrent Neural Networks. IEEE Computer Graphics and Applications.
[5]. Wongsuphasawat, K., et al. (2016). Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations. IEEE Transactions on Visualization and Computer Graphics.
[6]. Hu, K., et al. (2019). Vimy: A Machine Learning Approach to Visualization Recommendation. CHI Conference on Human Factors in Computing Systems.
[7]. Dibia, V. (2023). LIDA: A Tool for Automatic Generation of Grammar-Agnostic Visualizations and Infographics using Large Language Models. ACL.
[9]. Hardt, M., Price, E., & Srebro, N. (2016). Equality of Opportunity in Supervised Learning. NeurIPS.
[10]. Dwork, C., et al. (2012). Fairness Through Awareness. ITCS.
[11]. Agarwal, A., et al. (2018). A Reductions Approach to Fair Classification. ICML.
[12]. Mehrabi, N., et al. (2021). A Survey on Bias and Fairness in Machine Learning. ACM Computing Surveys.
[13]. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). Why Should I Trust you? Explaining the Predictions of Any Classifier. KDD.
[14]. Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. NeurIPS.
[15]. Vaswani, A., et al. (2017). Attention Is All You Need. NeurIPS.
[16]. Liu, Y., et al. (2023). Vis Explainer: Automatic Explanation Generation for Data Visualizations. IEEE VIS.