**An User-Customized Data Visualization System Based on Chat2VIS**

**Group Number**: 13

Dewei Tan 2093688 ([d.tan@student.fdu.edu](mailto:d.tan@student.fdu.edu))

Xiangjun Li 2092016 ([x.li3@student.fdu.edu](mailto:x.li3@student.fdu.edu))

Zihuan Zhu 2086903 ([z.zhu2@student.fdu.edu](mailto:z.zhu2@student.fdu.edu))

Lina Gu 2088991 ([l.gu@student.fdu.edu](mailto:l.gu@student.fdu.edu))

Jing Li 2091325 ([j.li8@student.fdu.edu](mailto:j.li8@student.fdu.edu))

**Original Paper Overview:** The original paper "Chat2VIS: Generating Data Visualizations via Natural Language Using ChatGPT, Codex, and GPT-3 Large Language Models" generates data visualizations from natural language descriptions. By using OpenAI LLMs, the system essentially translates user natural language instructions into python code and displays analysis charts and graphs.

**Problem Statement:** Chat2VIS's analysis graph output is primitive, fixed, far from satisfactory for direct research use. Furthermore, the current interface could be improved with new level user interactivity. Notably, this article used GPT-3 model ‘‘text-davinci-003’’ and the Codex model ‘‘code-davinci-002’’, as well as ChatGPT 3.5, but these models are shut down by OpenAI on January, 2024.

**Proposed Improvement:** Our project aims to rebuild the Chat2VIS website with new intuitive user interface using our own code. Apart from new looks, we will introduce powerful data visualization library realizing advanced visualization settings for users to customize which include color schemes, shapes, fonts, labels, legends and more. During this process, we would replace the retired models with GPT-4o mini and GPT-4o as well.

**Impact of Improvement:** By enhancing Chat2VIS with additional customization features, we will significantly improve user experience and make it a professional platform for data analysis and presentation in daily use and academic application. However, one unknown side effect of LLM model update would be the difference of prompt output results from the paper, but the logic and ideas behind remain the same.

**Original Paper Reference**

P. Maddigan and T. Susnjak, "Chat2VIS: Generating Data Visualizations via Natural Language Using ChatGPT, Codex, and GPT-3 Large Language Models," 2023. DOI: 10.1109/ACCESS.2023.3274199. https://ieeexplore.ieee.org/document/10121440

**Project Scope:**

**1.Goals and Improvements:**

1.1 **New user interface**: offer reshaped intuitive, user-friendly and efficient interface.

1.2 **New features**: provide advanced customization including color changing, shape and font options, label and legend styling and more.

1.3 **New API support**: abandon deprecated LLM APIs and replace with whole new systems.

1.4 **New prompt formats**: redesign prompt formats for LLM models and visualization libraries achieving flexible data representation.

**Novel Aspect**

Our project extends the Chat2VIS framework by a full suite of customization tools with brand-new UI/UX. Unlike the original system, which only produces an analysis graph prototype, we introduce flexible and advanced user visualization settings delivering sufficiently high standard analysis graphs.

**Technology Stack:**

Languages: Python, JavaScript

Libraries: openai, pandas, numpy, matplotlib, streamlit, pyecharts

APIs: OpenAI LLM APIs for natural language processing

Frameworks: Streamlit.