What Do LLMs Prioritise When Adapting Visualizations to User Personas?

Amir Hosseini*, Jo Wood, Aidan Slingsby, Radu Jianu, Mai Elshehaly

giCentre, City St George's, University of London

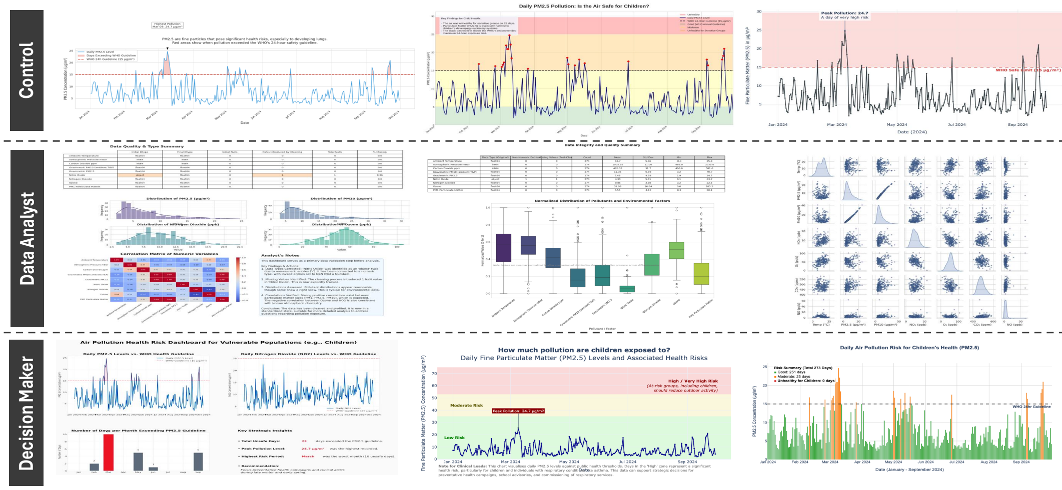

Fig. 1: Example designs generated by the LLM to adapt visualizations to a decision-maker persona (bottom), an analyst persona (middle), and a control (top), in response to the task *"I want to know how much pollution children are exposed to"*.

Abstract—
Large Language Models (LLMs) are increasingly used for generating and adapting visualizations for different user groups. While recent efforts have focused on adapting visualizations to users' cognitive and perceptual abilities, how LLMs cater to the distinct interests and *subjective* priorities of various stakeholder groups remains largely unexplored. Specifically, LLMs utilise rhetorical elements to prioritise data stories, which can shape user interpretation. We present a systematic approach to assessing how LLMs adapt their visualization rhetoric to match the priorities of different user personas in a healthcare context. Based on qualitative interviews with population health stakeholders, we demonstrate LLMs' capabilities for *(i)* understanding user tasks and priorities from interview data, *(ii)* adapting visualizations to these priorities, and *(iii)* justifying design choices for the adaptations. Population health data presents an excellent space for experimentation, given: *(a)* the diversity of stakeholders (e.g., commissioners, population health experts, data analysts, and the public); and *(b)* the varied purposes and key messages for which visualizations are designed. We reflect on patterns in LLM reasoning about persona-specific design choices—*in light of an established analytical framework for rhetorical visualization*—and propose open questions to promote safer, more responsible practices in LLM-assisted visualization.

Index Terms—Generative AI, Visualization, User Tasks

1 Introduction

User-adaptive visualization aims to amplify the human ability to perceive, experience, and reason about visualization by taking into account the needs, cognitive traits, and tacit knowledge of individual users. The ability to personalise visualizations in ways that adapt to the specific needs and preferences of individual user groups relies on a representation of user models, which can be collected explicitly, simulated by Artificial Intelligence (AI) [18], or captured through user interaction [16]. Traditionally, these user models have focused adaptations to long-term user characteristics, such as demographics and personality traits [17]. When considering user priorities, adaptation efforts have focused on low-level tasks, such as mouse and keyboard interactions. However, in real-world visualization authoring, higher-level analytical activities and goals of the users play a crucial role in identifying priorities for visualization design [11].

The proliferation of LLMs is creating new opportunities for user-adaptive visualization to consider these higher-level priorities, while incurring minimal overhead for visualization au-

*amir.hosseini.2@citystgeorges.ac.uk

Manuscript received 05 Jul. 2025; accepted 01 Aug. 2025. Date of Publication 02 Nov. 2025; date of current version xx xxx. 2025. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

thors. Namely, LLMs can be fine-tuned to adapt visualizations to users' fast-evolving priorities and data analytics or decision-making needs, without requiring users to explicitly articulate those priorities. However, this level of LLM automation can introduce several challenges for adaptation, especially in applications that concern population health, where there is a risk that LLM-authored adaptations may promote biased messages and lead to ill-informed decision-making. Therefore, a careful and systematic toolkit to assess *how* LLMs construct visual narratives from data, and *what* subjective and rhetorical elements they incorporate in their design choices, is critical to ensure the responsible generation of user-adaptive visualizations.

We present a systematic approach for the capture and evaluation of design patterns in LLM-generated user-adaptive visualizations. Users in our experiments are modelled through personas that are obtained from real interview data with two broad categories of stakeholders of population health data: data analysts and decision-makers. We report on design choices made by the LLM for each persona category and reflect on its use of the four layers of rhetoric identified by Hullman and Diakopoulos in [7]. This deductive approach enables us to identify common patterns and areas of concern where the LLM's choices might divert users' attention or obscure key data insights.

2 Background and Related Work

Adapting visualizations to specific users has been extensively explored in the literature. This section highlights the dimensions most relevant to the use of personas in adaptation and advocates the need for a role-based, activity-centred framework for adaptation. For a more thorough review of considerations in adaptive visualization systems, we refer the reader to [17].

2.1 The Role of Personas in User-Centred Design

Personas are widely recognised as a cornerstone of User-Centred Design (UCD) and Human-Computer Interaction (HCI). They are typically defined as *"fictitious characters representing archetypal users of a system, product, or service"* [1], created from user research to embody the needs, goals, and behaviours of key user segments. The primary purpose of personas is to help designers target their decisions throughout the design process, cater to user needs, and create empathetic representations that guide their choices. While the principles of activity-centred and role-based designs have been established as methodologies to guide human designers in creating effective user interfaces, their application in the domain of generative AI remains largely unexplored. Recent efforts in several domains have recognised the role that LLMs can play in simulating user personas in applications like adaptive narrative systems [3], automated UX design [6], and high-level human-AI reasoning and co-creation [9].

2.2 Persona Characterisation for Visualization

Personas play a crucial role for visualization designers. They help ensure that visualizations are not designed in a vacuum but are instead tailored to the intended users, considering their specific tasks, goals, and the context in which they operate. By representing target users, personas guide designers in selecting appropriate chart types, interaction techniques, and levels of detail, aiming for visualizations that are more effective, engaging, and fit for purpose [18]. The literature reveals several dimensions through which personas could be characterised to tailor data visualizations, which we outline in this section.

2.2.1 Cognitive and Perceptual Characteristics

A prominent approach to persona characterisation, particularly in the context of adaptive systems, involves defining users based on a spectrum of cognitive and psychological attributes. The work by Yanez and Nobre [18] provides a recent example, exploring the use of GPT-4 to tailor visualizations based on personas characterised by such attributes, which they categorise in several key dimensions: *(i)* short-term transient cognitive states like level of attention, feeling of confusion and cognitive load; *(ii)* long-term personality traits such as locus of control, need for cognition and dispositions like neuroticism and extraversion; *(iii)* long-term cognitive abilities, which encompass inherent capacities like perceptual speed, visual working memory, and spatial ability; *(iv)* user's experience and potential biases, such as their visualization literacy and their data expertise in the relevant subject matter; and *(v)* subjective visual preferences, like preferred colour schemes, desired visual complexity, and basic demographic information, including age, colour perception and language. These traits are posited to have an impact on how users perceive, process, and interact with visualizations.

A key trait for developing user personas is visualization literacy, which refers to an individual's ability to effectively find, interpret, evaluate, use, and create data visualizations [2]. This concept is also encompassed within the "Experience/Bias" category in the framework by [17]. The influence of such traits is further explored in [13], where statistical expertise, visual literacy, perceptual speed, and visual working memory have been shown to affect visualization perception.

2.2.2 Activity-Centred Characteristics

Another approach to persona characterisation is to define personas primarily by the tasks and activities they need to perform. This "activity-centred" approach prioritises understanding user activities and tasks [11]. As articulated by Marai [11] in the context of problem-driven scientific visualization, this framework provides a structured methodology for domain characterisation by focusing on user tasks and workflows. In this model, a distinction is made between high-level **activities** and lower-level **tasks**. An activity is a high-level structure, such as understanding genomic relationships, while a task is a more granular component of the activity, such as loading a dataset or locating a specific gene cluster. The model's focus on detailed, concrete tasks helps to reveal the underlying requirements of the domain precisely. The visualization literature acknowledges the importance of tasks in evaluation studies. However, the process of selecting and describing tasks can often be ad-hoc, highlighting the need for more systematic approaches to task characterisation [12].

Complementing the activity-centred view, another key dimension for persona characterisation is the user's professional role. This role-based approach is crucial because individuals in similar roles tend to perform similar activities. The importance of understanding user roles is emphasised in the design study methodology literature, where Sedlmair et al. [14], for instance, identify several distinct collaborator roles that are critical to the success of a visualization project. Different roles imply different goals, responsibilities, and levels of engagement with data. Therefore, defining personas based on roles helps ground the design process in the organisational and professional context in which the visualization will be used.

The framework proposed in this paper adopts a structured, role-based, and activity-centred method for persona development. The novelty here lies in the specific synthesis of "role" and "activity" for the explicit purpose of guiding LLM-driven visualization personalisation in a professional domain like Public Health Management, where understanding the practical, day-to-day activities of stakeholders is crucial for creating effective and relevant data visualizations, that are tailored to support the specific workflows and goals of different user roles.

3 User Persona Characterisation

We extracted user tasks and roles from nine interview transcripts. The interviews were part of a wider study exploring the interests of five district-level decision-makers and four data analysts in routinely collected health data. Interviews aimed at exploring how routine data (such as electronic health records) could supplement their existing workflows and lead to better decision-making. Participants were asked questions about their current data practices and how they envision visualization of routine data to supplement and enhance these practices. This study has institutional ethics approval; interview transcripts were anonymised, and all identifying information was removed prior to analysis.

We employed an activity inquiry approach to extract relevant persona characteristics from the interview data [11]. That is, for each participant, we extracted the tasks, data, and flow information. The tasks were extracted at two levels of granularity, following the definition in [11]: *(i) activities:* are high-level structures, such as *"A1.1: manage a patient tracking list"*; *(ii) tasks:* are lower-level components of the activity, such as *"T1.1: identify potentially missing patient entries"*. We matched each task with one of two user archetypes: **data analyst** and **decision-maker**. This broad categorisation follows a similar definition to that given by Dimara et al. [5], in which decision-makers are characterised by a non-technical background and less familiarity with the details of data capture, storage, and processing. Table 1 shows a sample of three data roles (*P1-P3*) and one decision-making role (*P4*), along with their anonymised activities.

We augmented user tasks for each persona with more holistic professional profiles. For each persona, we constructed a rich character description that included a fictional name, role title, and a narrative summary covering their professional background, key responsibilities, and educational history. Finally, we synthesised and abstracted the relevant role-based and activity-centred information from the character descriptions. This process resulted in comprehensive characteristics for each persona, which constituted the core input for the adaptation task.

4 Visualization Adaptation Experiment

We fed the personas described in Section 3 to a Gemini 2.5 Pro LLM and prompted it to adapt visualizations to match the interests of each persona. At the time of conducting our experiments, this model was recognised for its advanced reasoning and instruction-following capabilities across numerous benchmark tasks, making it a suitable choice for complex visualization adaptation tasks. It also benefits from having a large context window, which makes it suitable for processing larger datasets. We present here our systematic investigation of how different user personas and model temperature settings influence the nature and justification of the LLM-generated visualizations.

4.1 Experimental Design

Our experiment was structured to evaluate the impact of persona-focused prompts on the output of an LLM. To achieve this, we established two experimental groups and a control group:

- **Analyst Persona Group:** The LLM was conditioned with a persona representing a data analyst. This persona emphasised technical proficiency, a focus on data quality, statistical rigour, and complex relationships within the data.
- **Decision-Maker Persona Group:** The LLM was conditioned with a persona representing a strategic decision-maker, specifically a non-technical stakeholder with a clinical background who prioritises high-level, actionable insights and Key Performance Indicators (KPIs) over detailed data examination.
- **Control Group:** The LLM received a generic prompt without any persona information.

This design allows for a comparative analysis between the two professional roles and a non-personalised baseline. Experimental conditions were created by matching these groups with different temperature settings for the LLM. While previous studies set the temperature to 0 to ensure less creativity and increase determinism of the generated output [4, 10, 15], we were interested in investigating the influence of the model's inherent stochasticity on personalisation. Temperature is a parameter that controls the randomness of a model's predictions—at 0, the model consistently selects the most probable next token, while higher values (up to 1.0) increase the likelihood of choosing less probable tokens, introducing greater variability and creativity. In our experiments, we used three temperature settings: 0 to minimise randomness and produce predictable, conservative outputs; 1.0 to maximise randomness and encourage novel, diverse solutions; and 0.5 as a middle ground to explore the trade-off between deterministic and creative personalisation strategies.

Each experimental condition (3 persona groups × 3 temperature levels) was replicated five times to ensure the reliability and consistency of the results, yielding a total of 45 trials per dataset.

4.2 Datasets

Our aim was to assess the LLM's ability to abstract user tasks and priorities from personas, without constraining its design choices to the specifics of an individual dataset. Therefore, we selected data that were not explicitly mentioned in either the interviews or the extracted domain-specific tasks. Instead, we utilised an air quality dataset from the city of Bradford, UK, compiled over a nine-month period from January to September 2024. This dataset comprises daily recordings from a network of sensors, capturing nine variables (Carbon Dioxide, PM10, PM2.5, Nitric Oxide, Nitrogen Dioxide, Ozone, PM1, Ambient Temperature, Atmospheric Pressure). The dataset was selected for its relevance to population health and its inherent complexity, which allows for analytical exploration.

In Experiment 1, our analysis was restricted to the data stream from a single sensor location. For Experiment 2, we included data from twelve distinct locations within the city. The data included geospatial information such as location name, latitude, and longitude for each sensor.

4.3 Prompt Engineering

We employed a zero-shot prompting approach and designed two primary prompt structures: a *Personalised Prompt* for the persona-based groups and a *Control Prompt* for the baseline group. All prompts instructed the LLM to assume the role of a visual analytics expert tasked with generating appropriate visualizations in response to a unified user task: *"I want to know how much pollution children are exposed to"*. The prompts included general information about the dataset and specific instructions for generating executable Python code, as well as providing clear justifications for the LLM's personalisation and design choices.

The critical distinction lay in the instruction of the personalisation task. The Personalised Prompt included an explicit directive for the LLM to tailor the visualization and its underlying data transformations to the specific needs, goals, and context of the persona it had been given. The Control Prompt omitted this personalisation directive.

4.4 Data Collection and Analysis

For each trial, the LLM generated two primary outputs as instructed: *(a)* a Python code snippet for producing a data visualization; and *(b)* a textual justification of its design choices.

Participant	Role	Key activities
P1	Data role	1. Managing, cleaning, and analysing patient lists for [disease] assessments. 2. Reporting on waiting list statistics, service performance, and referral trends. 3. Responding to data-specific queries, particularly for outsourced assessments and service delays.
P2	Data role	1. Managing, processing, and ingesting incoming data feeds for a secure data environment. 2. Ensuring data quality, standardisation, pseudonymisation, and linkage. 3. Managing user access, local and cloud-based data platform infrastructure, and documentation.
P3	Data role	1. Developing and maintaining aspects of the data infrastructure. 2. Conducting data analysis on specific datasets (e.g., primary care) for various projects. 3. Critically assessing data quality, provenance, and fitness for research.
P4	Decision role	1. Identifying, and stratifying vulnerable patient groups within a primary care practice network (PCPN). 2. Inform and prioritise local health improvement initiatives at the PCPN level. 3. Balance population-level targeted health interventions with nuances of individual patient care.

Table 1: Example high-level activities identified for interview participants. Data roles tend to involve more technical activities.

To structure our analysis of the LLM's personalisation strategy, we instructed the model to categorise its justifications according to three types of adaptive visualization interventions identified by Yanez et al. [17]:

- **Pre-visualization:** Data-level adaptations made before rendering (e.g., filtering, aggregation).
- **Within-visualization:** Adaptations to the visual encoding in a chart (e.g., colour or annotations).
- **Between-visualizations:** Adaptations that involve suggesting or generating entirely new or alternative visualizations.

The generated Python code was rendered to visually inspect the output, while the categorised justifications provided a structured dataset for analysing the model's output generation process. Initial inspection of the LLM-generated visualizations and justifications revealed argumentative themes that align with elements of the visualization rhetoric framework in [7]. Therefore, we use this framework as a basis for human validation of the LLM-generated design patterns and justifications, as we detail next.

5 Human Validation

The results produced by the LLM were first organised into Excel sheets for initial inspection, where each row captured the run ID, persona category, temperature, run index, justification text, and produced visualization(s). The authors discussed the results in a series of meetings and noted down high-level patterns that were clear in the LLM's design choices. This discussion led to a consensus that the observed design patterns could be further categorised and analysed according to the Visualization Rhetoric design framework by Hullman and Diakopoulos [7]. Therefore, a deductive approach was taken to group design choices and their justification under the four layers of rhetoric: data adaptations, visual representation, interaction, and annotation (both graphical and textual). The design dimensions within each layer were listed and further refined inductively as more themes emerged from the LLM-generated text and charts. Table 2 summarises the final set of design dimensions included in our analysis.

5.1 Layer I: Data Adaptations

The LLM dedicated more pre-visualization data adaptations to the data analyst persona. Adaptations primarily focused on improving data quality, including data imputation, interpolation, and type conversion. Aggregations were applied to focus insights on value distribution and correlation, allowing the analyst to inspect data integrity. This meant that the temporal patterns in the data were consistently obscured in response to this persona, which contrasts with the LLM's response for the other experimental conditions, especially when considering data from a single data source (see Figure 1).

The LLM applied source selection in response to the full dataset, which includes all sensors. Assuming the analyst is only concerned with data relating to children's exposure, the LLM selected only three sensors as data sources which were deemed to be in close proximity to schools, in 93.3% of the visualizations it generated for the analyst persona. This pattern dropped to 66.6% in response to the decision-maker persona.

While the LLM's tendency to eliminate data sources was more prevalent for the analyst persona, the opposite was true for variable selection, where it included more variables, focusing its messages on only one or up to two pollutants in the decision-maker and control groups. Aggregation and binning were applied consistently in 100% of the visualizations for the data analyst across both single- and multi-sensor datasets. They were, however, rarely applied in the decision-maker category for single-sensor data, but were prevalent for multi-sensor data. All aggregations and data adaptations were supplemented with consistent justification text dedicated to provenance information that was deemed relevant to each persona. Additional data provenance tables were included only for the data analyst (see Figure 1—*middle*).

Key takeaway: The LLM made several persona-specific assumptions about which data sources and variables to include or exclude in the visualization. Therefore, there is a need to supplement the textual justifications provided by the LLM with provenance visualization, which would enable users to detect assumptions made and override them if necessary.

5.2 Layer II: Visual Representation

Both within- and between-visualization adaptations were observed in the LLM-generated responses.

Within-visualization adaptations. When not presented with a user persona (**control** runs), the LLM consistently resorted to a line chart (Figure 1—*top*), while making minor within-visualization adaptations to present additional information, such as the World Health Organization (WHO)'s 24-hour PM2.5 guideline (15 µg/m³). Values that exceeded the benchmark were *emphasised* using colour bands, shading, and additional visual marks (such as Xs). This pattern of adaptation remained consistent in control runs even with higher temperature levels (0.5 and 1.0), which were meant to allow the LLM a higher level of creativity when composing the visualization.

When adapting visualizations for the data analyst persona, no within-visualization adaptations were detected. Instead, this persona exhibited a high level of between-visualization adaptations, which we detail below. In contrast, some within-visualization adaptations were applied in response to the decision-maker persona. An example is shown in Figure 1—*bottom row*, where

<table>
<tr><th colspan="9">Layer I: Data Adaptations</th></tr>
<tr><th colspan="6">Omission</th><th>Provenance</th><th colspan="2">Transformation</th></tr>
<tr><td>Data source selection</td><td>Variable selection</td><td>Outliers</td><td>Aggregation</td><td>Categorization</td><td>Axis thresholding</td><td>Provenance</td><td>Filtering</td><td>Scaling</td></tr>
<tr><th colspan="9">Layer II: Visual Representation / Mapping Rhetoric</th></tr>
<tr><th colspan="3">Uncertainty</th><th colspan="6">Obscuring</th></tr>
<tr><td>Statistical Graphics</td><td>Distributions</td><td>Correlations</td><td>Visual noise</td><td>Third dimension</td><td>Oversizing</td><td>Emphasis</td><td>Position on axis</td><td>Double axes</td></tr>
<tr><th colspan="2">Metaphor</th><th>Classification</th><th colspan="6">Layer III: Interaction</th></tr>
<tr><td>Implicit mappings</td><td>Contrasts</td><td>Grouping</td><td>Selection</td><td>Brushing & linking</td><td>Rescaling</td><td>Zooming</td><td>Panning</td><td>Annotating</td></tr>
<tr><th colspan="9">Layer IV: Annotation (graphical or textual)</th></tr>
<tr><th colspan="4">Uncertainty</th><th colspan="2">Emphasis</th><th colspan="3">Classification</th></tr>
<tr><td>Leap of faith forecast</td><td>Explicit labelling</td><td>Expression of doubt</td><td>Outlier annotation</td><td>Explicit labelling</td><td>Font bolding or italicizing</td><td>Explicit boundaries</td><td>Typographic features</td><td>Legends</td></tr>
<tr><th colspan="6">Provenance</th><th colspan="3">Linguistic-based rhetoric</th></tr>
<tr><td>Data sources</td><td>Additional references</td><td>Methodological choices</td><td>Exceptions</td><td colspan="2">Corrections</td><td colspan="3">Quotations, over/understatements, etc.</td></tr>
</table>

Table 2: Taxonomy of design choices made for adaptation, based on [7] and our own observations.

the colour bands for levels of risk to human health were adapted from global WHO levels specified in the control case (see Figure 1—*top*) to the UK's Daily Air Quality Index (DAQI).

The LLM deemed that the UK-specific definition of risk to population health is more contextually relevant to the decision-maker persona than the global WHO ranges. It is clear that these adaptations of the same visualization convey two different messages about the level of risk to human health implied by the data. The decision of which benchmark is more appropriate for a given decision-making context requires domain expert feedback.

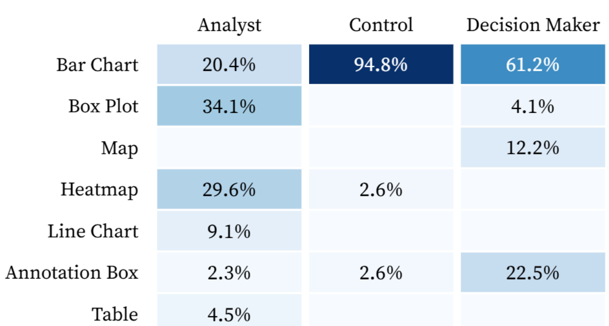

Fig. 2: Distribution of chart types by persona group.

Between-visualization Adaptations. The LLM devoted more between-visualization adaptations to the data analyst persona than to the decision-maker or control groups (Figure 2). While the latter two exhibited relatively consistent visualization patterns, the analyst persona prompted a notably diverse range—from complex multi-view analytical dashboards to statistical plots such as box plots and small multiples like scatterplot matrices (Figure 1). This diversity suggests that the LLM interprets the analyst role as requiring flexible, exploratory approaches tailored to varied analytical questions.

Box plots and heatmaps dominated the analyst outputs, together accounting for 63% of all visualizations. Box plots were the most frequently used, reflecting the LLM's emphasis on tasks central to data quality assessment—understanding statistical distributions, identifying outliers, and comparing summaries across variables. Heatmaps appeared consistently, indicating an intention to reveal correlations in data missingness between sensor locations and pollutant variables. Bar charts were used for the analyst persona only 20% of the time—less than half as often as for decision-makers (61%) and control groups (95%)—suggesting their selective use for simple categorical comparisons rather than as a default choice.

In contrast, the output of the control group was characterised by a significant lack of variety, highlighting the LLM's baseline behaviour. An overwhelming 95% of all the charts created for the control group were bar charts. In some cases, a high number of bar charts were produced in a single output. The visuals for this group lacked complexity, with a single heatmap and annotation box being notable outliers.

The decision-maker persona's visualizations focused on high-level, interpretable and actionable insights, often with geographical or narrative elements. This was the only persona that generated maps, which appeared in 32% of its runs. It also used text labels more than any other, demonstrating a priority not just for showing data but also for explaining them to support a decision-making process. While bar charts were frequently used, they were often complemented by maps or text labels to create a more comprehensive picture.

Key takeaway: The LLM considered human expertise and data-related activities for each persona when making both within- and between-visualization adaptations. More research is needed to validate these choices for each of the intended audiences.

5.3 Layer III: Interaction

Our experiment revealed a stark contrast in the use of interactive elements across personas, linking back to their implied needs. The decision-maker persona had the highest rate of interactive outputs, with over half of its runs (53%) producing an interactive visual. This strongly suggests that the LLM interprets this role as one that requires active engagement with the data, utilising interactive dashboards to filter, drill down, and explore various facets to make informed decisions.

In contrast, the analyst persona produced no interactive charts. This finding implies that the LLM views the analyst's role as one who performs an analysis and then presents a definitive, static report of their findings, where the output is the conclusion, not a tool for further exploration. The control persona was in the middle, with 18% of its outputs being interactive. This suggests that while interactivity is within the LLM's general capabilities, it is not a default feature and is used sparingly without persona-specific prompting.

The tendency to generate multi-chart layouts, or dashboards, was another clear distinction. Both the analyst (94%) and decision-maker (89%) personas showed a strong preference for multi-view visuals, aligning with their respective roles of providing either a deep statistical analysis or a strategic overview. Conversely, the control group was far less likely to create a multi-view layout, doing so in less than half of its runs (41%). It typically defaulted to a single, direct visualization, reinforcing the idea that creating a dashboard is a deliberate, persona-driven choice for the LLM.

Finally, the decision to create a multi-view layout appears to be independent of temperature, with consistent rates across all levels. This indicates that the persona, not model creativity, is the driving factor behind dashboard creation. However, temperature did have a significant impact on interactivity. A temperature of 0, which makes the model more deterministic, resulted in far fewer interactive visuals (7%). Once the temperature was increased to 0.5 (35%) or 1.0 (29%), the model had the creative freedom to employ interactive elements more frequently.

Key takeaway: The LLM focused on engagement and interactivity for the decision-maker persona. This contrasts with the expectation of a data analyst who may want to conduct exploratory data analysis, for which interaction is key. Expectations for interactivity need to be more clearly articulated to the LLM.

5.4 Layer IV: Annotation

The application of annotations varied across the three experimental conditions, revealing a persona-driven approach by the LLM. This was evident not only in the content of the annotations but also in their frequency. Annotations were present in 84% of visualizations for the decision-maker and 82% for the data analyst, but this figure dropped to just 29% for the control group.

For the data analyst persona, annotations centred on enhancing precision, transparency, and data provenance. The LLM frequently added detailed labels to visual elements, such as the exact numerical values on bar charts and heatmap cells. This was justified as a way to provide the expert user with the precise figures required for validation and reporting, removing reliance on visual approximation alone. Furthermore, annotations were used to document data processing steps, specify units of measurement, and state data limitations, for instance by noting how missing data were handled. In several instances, dedicated "Data Quality Report" sections were added as textual annotations.

In contrast, annotations for the decision-maker persona were geared towards providing strategic context and highlighting actionable insights. A dominant pattern was the inclusion of external benchmarks, like the WHO guidelines, which were added as reference lines. The LLM justified this as a method to transform raw data into a meaningful assessment against a health standard, providing an evidence base for interventions. The LLM also used graphical annotations to direct the user's attention, such as using arrows, callout boxes, or distinct colours to highlight the most critical data points. This was done to surface key insights without requiring a deep exploration. Finally, the LLM often supplemented the visuals with textual summaries that translated the data into a synthesised narrative to support decision-making.

The low frequency of annotations (29%) in the control group indicates that, without a guiding persona, the LLM does not make many assumptions about insights or provenance information to include. The annotations that were present primarily focused on general readability and clarity, such as adding standard data labels for precision and ensuring titles and axes were clear. The significant drop in usage demonstrates that sophisticated annotation is a deliberate, context-driven choice for the LLM, employed as a rhetorical tool to tailor the visualization's message to the specific goals of the intended user.

Key takeaway: The LLM uses annotations as a primary tool to meet the specific and complex needs of a given persona, rather than as a default feature. Annotations were employed not only for clarification but as a rhetorical device to frame the data in a manner consistent with the perceived needs of each persona.

6 Conclusion, Limitations, and Future Work

We presented a systematic investigation into how LLMs adapt data visualizations to the needs of different professional roles and personas. We prompted Gemini 2.5 Pro with personas for a data analyst and a decision-maker—derived from real interviews in a public health context—and analysed the outputs through the lens of visualization rhetoric. Our findings demonstrate that LLM outputs show sophisticated persona-driven rhetorical framing that extends beyond simple chart generation. For the data analyst persona, the model prioritised data quality, statistical rigour, and provenance, generating complex, multi-view static layouts. In contrast, for the decision-maker persona, it focused on producing actionable and high-level insights using more accessible charts, interactive elements, and contextually relevant annotations. These distinct patterns highlight the LLM's ability to interpret the implicit goals and priorities embedded in user personas and translate them into specific rhetorical choices across data, visual representation, interaction, and annotation layers.

As these models become increasingly integrated into data analysis and communication workflows, understanding the rhetorical stances they adopt is critical to ensuring responsible, effective use of LLM-assisted visualization authoring, particularly in high-stakes domains such as public health. While our study offers useful insights, several limitations remain for future work.

First, our experiments were conducted exclusively with a single LLM (Gemini 2.5 Pro). The specific patterns of rhetorical adaptation and justification we observed are, therefore, a product of its unique architecture, training data, and fine-tuning. Although pilot findings suggest similar behaviour among other LLMs (e.g., models from the GPT families), they were not included in our full experiment. Therefore, our findings may not be **generalisable** to other LLMs. Our future work will involve comparative studies across multiple LLMs to build a robust understanding of how different models approach adaptation.

Second, while the LLM was able to articulate justifications for its design choices, we must be critical of these explanations, as they could include **hallucinations** or **post-hoc rationalisations** constructed from patterns in the model's training data, rather than a genuine reflection of an internal design process. Future work should contrast these with human-articulated design justifications, such as those in [8].

Importantly, while our analysis applies a rhetorical framework to interpret the LLM's design rationale, the LLM is not a deliberate agent taking a calculated rhetorical stance. Its outputs reflect statistical patterns learnt from training data rather than conscious design decisions or true comprehension of user needs. A promising direction for future research is to prompt the model to construct competing narratives. For instance, one that emphasises the urgency of a public health crisis versus one that suggests otherwise. Finally, the personas used, while grounded in real interview data, remain synthesised archetypes. Future studies should focus on evaluating LLM-driven adaptation with real users in real-time, interactive scenarios.

Acknowledgements

This work is co-funded by the Bradford Institute for Health Research and EPSRC Doctoral Training Partnership at City St George's, University of London. We thank the VAHC reviewers for their valuable suggestions, which helped improve this paper.

References

[1] D. Amin, J. Salminen, F. Ahmed, S. Tervola, S. Sethi, and J. Jansen. How is generative ai used for persona development?: A systematic review of 52 research articles, 04 2025. doi: 10.48550/arXiv.2504.04927 2

[2] S. Beschi, D. Falessi, S. Golia, and A. Locoro. Characterizing data visualization literacy for standardization: A systematic literature review. *IEEE Access*, PP:1–1, 01 2025. doi: 10.1109/ACCESS.2025.3559298 2

[3] F. M. Calisto, J. a. Fernandes, M. Morais, C. Santiago, J. a. M. Abrantes, N. Nunes, and J. C. Nascimento. Assertiveness-based agent communication for a personalized medicine on medical imaging diagnosis. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, article no. 13, 20 pages. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3580682 2

[4] V. Dibia. LIDA: A tool for automatic generation of grammar-agnostic visualizations and infographics using large language models. In D. Bollegala, R. Huang, and A. Ritter, eds., *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, pp. 113–126. Association for Computational Linguistics, Toronto, Canada, July 2023. doi: 10.18653/v1/2023.acl-demo.11 3

[5] E. Dimara, H. Zhang, M. Tory, and S. Franconeri. The unmet data visualization needs of decision makers within organizations. *IEEE Transactions on Visualization and Computer Graphics*, 28(12):4101–4112, 2022. doi: 10.1109/TVCG.2021.3074023 3

[6] Y. Huang. Generating user experience based on personas with ai assistants. In *Proceedings of the 2024 IEEE/ACM 46th International Conference on Software Engineering: Companion Proceedings*, ICSE-Companion '24, 3 pages, p. 181–183. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3639478.3639810 2

[7] J. Hullman and N. Diakopoulos. Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2231–2240, 2011. doi: 10.1109/TVCG.2011.255 2, 4, 5

[8] M. Hutchinson, R. Jianu, A. Slingsby, J. Wood, and P. Madhyastha. Capturing visualization design rationale. *arXiv preprint arXiv:2506.16571*, 2025. 6

[9] Y. Liu, P. Sharma, M. Oswal, H. Xia, and Y. Huang. Personaflow: Designing llm-simulated expert perspectives for enhanced research ideation. In *Proceedings of the 2025 ACM Designing Interactive Systems Conference*, pp. 506–534, 2025. 2

[10] P. Maddigan and T. Susnjak. Chat2vis: Generating data visualizations via natural language using chatgpt, codex and gpt-3 large language models. *IEEE Access*, 11:45181–45193, 2023. doi: 10.1109/ACCESS.2023.3274199 3

[11] G. E. Marai. Activity-centered domain characterization for problem-driven scientific visualization. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):913–922, 2018. doi: 10.1109/TVCG.2017.2744459 1, 2, 3

[12] A. Pandey, U. H. Syeda, and M. A. Borkin. Towards identification and mitigation of task-based challenges in comparative visualization studies. In *2020 IEEE Workshop on Evaluation and Beyond - Methodological Approaches to Visualization (BELIV)*, pp. 55–64, 2020. doi: 10.1109/BELIV51497.2020.00014 2

[13] T. Poetzsch, P. Germanakos, and L. Huestegge. Toward a taxonomy for adaptive data visualization in analytics applications. *Front Artif Intell*, 3:9, Mar. 2020. 2

[14] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012. doi: 10.1109/TVCG.2012.213 2

[15] L. Wang, S. Zhang, Y. Wang, E.-P. Lim, and Y. Wang. LLM4Vis: Explainable visualization recommendation using ChatGPT. In M. Wang and I. Zitouni, eds., *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pp. 675–692. Association for Computational Linguistics, Singapore, Dec. 2023. doi: 10.18653/v1/2023.emnlp-industry.64 3

[16] K. Xu, A. Ottley, C. Walchshofer, M. Streit, R. Chang, and J. Wenskovitch. Survey on the analysis of user interactions and visualization provenance. In *Computer Graphics Forum*, vol. 39, pp. 757–783. Wiley Online Library, 2020. 1

[17] F. Yanez, C. Conati, A. Ottley, and C. Nobre. The state of the art in user-adaptive visualizations. *Computer Graphics Forum*, 44(1):e15271, 2025. doi: 10.1111/cgf.15271 1, 2, 4

[18] F. Yanez and C. Nobre. User-Adaptive Visualizations: An Exploration with GPT-4. In D. Archambault, I. Nabney, and J. Peltonen, eds., *Machine Learning Methods in Visualisation for Big Data*. The Eurographics Association, 2024. doi: 10.2312/mlvis.20241126 1, 2