



Article

AI-Supported EUD for Data Visualization: An Exploratory Case Study

Sara Beschi ¹, Daniela Fogli ^{2,*}, Luigi Gargioni ² and Angela Locoro ¹

¹ Department of Economics and Management, University of Brescia, 25121 Brescia, Italy; sara.beschi@unibs.it (S.B.); angela.locoro@unibs.it (A.L.)

² Department of Information Engineering, University of Brescia, 25123 Brescia, Italy; luigi.gargioni@unibs.it

* Correspondence: daniela.fogli@unibs.it

Abstract

Data visualization is a key activity in data-driven decision making and is gaining momentum in many organizational contexts. However, the role and contribution of both end-user development (EUD) and artificial intelligence (AI) technologies for data visualization and analytics are still not clear or systematically studied. This work investigates how effectively AI-supported EUD tools may assist visual analytics tasks in organizations. An exploratory case study with eight interviews with key informants allowed a deep understanding of data analysis and visualization practices in a large Italian company. It aimed at identifying the various professional roles and competencies necessary in the business context, understanding the data sources and data formats exploited in daily activities, and formulating suitable hypotheses to guide the design of AI-supported EUD tools for data analysis and visualization. In particular, the results of interviews with key informants yielded the development of a prototype of an LLM-based EUD environment, which was then used with selected target users to collect their opinions and expectations about this type of intervention in their work practice and organization. All the data collected during the exploratory case study finally led to defining a set of design guidelines for AI-supported EUD for data visualization.

Keywords: end-user development; data visualization routines; human–AI interaction; case study research



Academic Editors: Rui Yu, Sooyeon Lee, Syed Masum Billah, John M. Carroll and Gianluigi Ferrari

Received: 27 June 2025

Revised: 20 July 2025

Accepted: 28 July 2025

Published: 1 August 2025

Citation: Beschi, S.; Fogli, D.; Gargioni, L.; Locoro, A. AI-Supported EUD for Data Visualization: An Exploratory Case Study. *Future Internet* **2025**, *17*, 349. <https://doi.org/10.3390/fi17080349>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Organizations encompass disparate work practices, ranging from specialized micro-tasks for operational processes (e.g., monitoring of a production device) to long-term strategic plans conceived as part of managerial competencies. All these work practices need a certain level of knowledge (expertise) and information (data) to allow for good decision making at each step of the organizational continuum. Data-driven activities characterize some of these transversal tasks. They represent knowledge-intensive practices based on skills that each person inside an organization should possess and through which they communicate. These skills should receive adequate support from interaction designers. As recently noted in [1], knowledge-intensive work practices can be supported by better end-user development (EUD) tools, but this is not sufficient. What is often overlooked is the need to analyze the domain of practice and to assess the individual computational skills and data literacy in order to determine what is the level of support that EUD tools can effectively bring in organizational activities. This personalization would yield a benefit by improving the quality of work practices [2].

Visual analytics is massively exploited as a key activity related to data-driven decision making in many professional roles and transversal tasks inside organizations. Recent research [3] has identified in data visualizations (data viz from now on) those kinds of “boundary objects” capable of becoming common informational capital for knowledge sharing inside and across organizational routines, sectors, and goals. Despite the paramount importance of data analysis and visualization in enterprises, there is a lack of systematization in the analysis of data viz practices, individual literacy, and adequate support from the related technologies [4].

Artificial intelligence (AI), and large language models (LLMs) in particular, bring the promise of supporting users with a more personalized and informal interaction style, often equipped with the ability to hide the complexity of low-level coding and database query formulation. As stated in [5], AI systems are characterized by the paradigm of being *adaptive*, autonomously adapting to users and contexts, whereas EUD systems are characterized by their being *adaptable*, i.e., adjustable from the users who decide their degree of adaptation to the problem at hand. The integration of these two paradigms depends on how they are applied to concrete work practices and customized for individual literacy. Work practices should be attentively investigated and, arguably, those that bring a higher business value and a wider benefit to the company at large should be those that are more carefully considered. We argue that data-driven work practices are gaining momentum as key innovation factors, and visual analytics and data viz are increasingly growing among those practices. Visual analytics is the data viz-based process of using data and processing them properly to investigate business phenomena. Data visualizations are the artifacts yielded for the specific process of visual analytics.

This paper aims to investigate how to integrate the adaptive (AI) and adaptable (EUD) paradigms to support visual analytics and data viz design practices, starting from the analysis of those practices in a large Italian company. Therefore, it presents an exploratory case study that aims to answer the following overarching question: *Can we support data analysis and visualization practices in business contexts with EUD enhanced by LLMs?*

We adopted the exploratory case study research method described in [6] to fill a gap in the research on causal studies of data viz-based practices in organizations. In this way, we aim to pose the initial hypotheses in order to “define the necessary questions and hypotheses for developing consecutive studies” (ibidem) in this field. This research fosters an in-depth analysis of an organizational domain and of the potential consequences of technological interventions. In particular, our exploratory case study research allowed us to answer the above overarching question by digging into two more specific research questions (RQs) and work hypotheses:

RQ1: Who are the workers (role, responsibility, competency) performing data analysis and visualization tasks in an organization, and how are their tasks characterized (technologies, literacy, business strategies)?

RQ2: Which behaviors and expectations would emerge if visual analytics and data viz design practices were enhanced with an AI-supported EUD tool?

The first phase of the case study research aimed at answering RQ1. It consisted of a qualitative investigation carried out by means of interviews with key informants, whose role is to modify data visualizations or to create new ones. This phase led to identifying key user profiles that require different levels of support and personalization to carry out their work with the use of data viz.

Based on the results of this phase, a second phase of the case study research aimed at answering RQ2. It consisted of developing an LLM-enabled EUD environment that allowed users to create customized data visualizations that meet their different needs, skills, and work practices, and of the direct observation of a selection of key informants interacting with this prototype. This LLM-enabled environment can be defined as a design probe, rather than a classical prototype [7]. The difference between the two is crucial for our purpose. A full prototype is subject to users' evaluations regarding usability, with the goal of developing and deploying it in the long run. A design probe is purposely designed as an underspecified artifact whose main goal is to be used in everyday practices to challenge, engage, and solicit reactions and responses from the users. For our intent, this probe should bring hypotheses about which design tradeoff between the adaptability and adaptivity of AI-supported EUD tools can be formulated. In this second phase, qualitative data were collected through direct observation and a questionnaire, where selected users provided their opinions about the effectiveness and trustworthiness of LLM-powered interactions. The exploratory case study research finally led us to hypothesize design guidelines for AI-supported EUD tools for visual analytics and data viz design practices.

The paper is structured as follows: Section 2 gives an account of previous works in the domain of data viz in organizational contexts and EUD approaches adopted in this domain; Section 3 describes the framework for our exploratory case study analysis in the data viz domain, and presents the analysis and the outcomes of the qualitative interviews; Section 4 presents our LLM-based EUD environment and the results of the experimental interaction with key users; Section 5 reports the findings, highlighting the potential and limitations of our approach; Section 6 concludes the paper and outlines possible future work.

2. Background and Related Work

In this section, we first provide a synthesis of the motivations underlying the research in the field of data analysis and visualization, highlighting the most important issues still to be addressed. Then, we review the most recent approaches that support end users in modifying or creating their data visualizations.

2.1. Data Analysis and Visualization in Organizational Contexts

The problematization of the topic reported in RQ1 regards the gap between systems' capabilities and users' expectations. The structured review in [8] synthesizes three decades of research in data management and visualization. The primary aim of this review is to highlight the gaps in visualization design, where the usability of a system is often compromised by the mismatch between its design and users' wishes. The review categorizes and discusses several key database optimization techniques that have been shown to benefit interactive analysis systems. At the end, considerations about the need for more robust and scalable solutions to support the increasing complexity and size of datasets in visualization systems are discussed, together with an integrated approach, where such optimization techniques are also available to the designers of visualization systems.

Another work contributing to moving the problem of data analytics and visualization outside the lab was that in [9]. In the vein of analyzing several organizational domains, such as healthcare, finance, and marketing, semi-structured interviews were conducted, and three professional archetypes were identified in the end: hackers, scripters, and application users, each one representing a different approach, skill set, and interaction modality with data analysis tools. The study proposed five high-level tasks that key users perform with data: discovery, wrangling, profiling, modeling, and reporting, as well as three categories of tool functionalities: database, scripting, and modeling. These professional roles, activities, and tools partly overlap with the ones identified in the present study. For ex-

ample, the same considerations emerged pertaining to what were the more tedious tasks in data analysis. Some of these aspects regard bad data quality, still unresolved when not amplified in AI-assisted tasks [10]. Although this study touched on themes and outcomes that are very similar to ours, it was limited to the framing of organizational roles and practices, without any intent of exploring the impact of different technologies on users' behaviors and expectations.

These behavioral and procedural aspects are better analyzed in the user study in [11], where 22 participants were observed during interaction with an AI-assisted prototype, based on an LLM, while performing different tasks based on the ARCADE benchmark [12]. The study aimed to observe how data analysts understood and verified the correctness of AI-generated analyses. This was not a case study, but it was based on generic data analysis tasks and on crafting a prompt, and it did not focus on visual analytics. Thus, the focus of the study was not on the behaviors and expectations of users regarding organizational decision-making practices.

Regarding the gap to be filled in terms of data visualization capabilities and desiderata from decision makers in enterprises, the study by Franconeri et al. [13] anticipated many of the issues still observable and unresolved in the organizational domain. The study used an online questionnaire asking about organizational data, technological equipment, and visualization usage. Many interesting outcomes emerged from the survey, in particular, the need for qualitative explanations of quantitative data analysis and visualizations; the need to have some background information on the data provenance and preparation; and the need to contextualize data-driven analysis into a wider picture before making sense of data. All of these aspects converged into the hypothesis that AI-based data visualization tools may have a role in them. However, this study did not specifically investigate the use of AI-based tools in enterprises. No design probe was put into place to hypothesize research directions that take into consideration this kind of technology.

2.2. End-User Development for Data Visualization

EUD for data visualization has been, so far, an under-explored topic. In 2013, Pantazos et al. [14] compared different research and commercial visualization development tools to investigate their suitability for end-user developers. The authors underlined how visualization development is traditionally carried out by professional programmers in collaboration with domain experts, but that such collaboration often leads to misunderstandings and longer development time. The examined tools offered graphical user interfaces with direct manipulation of visual objects, but most of them required professional programming skills or presented themselves as 'black boxes', preventing end users from creating visualizations different from predefined templates. Furthermore, it was observed that evaluation with users was usually ignored [14]. Thus, EUD is advocated as a solution to this problem; that is, providing end users (domain experts) with proper tools they can use to directly create or customize the desired visualizations for their data, as also already suggested in [15]. To this purpose, ref. [16] proposes uViz, which presents a very rich but complex graphical user interface, leading users to make errors during formula definition and data binding; indeed, participants in the user study claimed that users need some information technology skills and database knowledge to use it effectively. Starting from the analysis of users' practices, knowledge, and skills in a real organization, this paper aims to create a design probe that allows for the investigation of how to ensure a better fit between users and technology.

Natural language-based interfaces allow one to overcome the difficulties encountered by end users with traditional query languages and manual plotting of visualizations. In their survey, Zhang et al. [17] investigated the evolution of methods and tools implement-

ing the translation of natural language into SQL (Text-to-SQL) and of natural language into visualization specifications (Text-to-Vis). Early approaches to text-to-SQL and Text-to-Vis were rule-based and template-based; then, neural networks and deep learning methods allowed a significant advancement in system performance; recently, pretrained language models (PLMs) and, above all, LLMs have offered new opportunities to develop easy-to-use interfaces for data query and visualization. As for the system architectures, end-to-end systems have emerged as the most suitable for end users; these systems process input questions (text or voice-based) and directly generate the desired output. Examples include Photon [18], VoiceQuerySystem [19], Sevi [20], and DeepTrack [21], but none of them are based on LLMs. To experiment with this novel technique and fill the gap related to the availability of end-to-end systems in the real world [17], this paper presents an LLM-based EUD environment, whose design was informed by the outcomes of the first phase of our case study research.

The survey by Hong et al. [22] analyzes papers proposing LLM-based systems for Text-to-SQL only, underlining that this has the potential to democratize access to data for those users who are not knowledgeable in SQL programming [23]. The LLM-based solution for Text-to-SQL has so far been more investigated than Text-to-Vis (e.g., [24,25]) since it brings superior generation capabilities. However, most of the papers surveyed did not involve users in system design and evaluation, but benchmark datasets and automatic tests were used. Robustness in real-world scenarios is considered a challenge for future work [22]; in particular, it is important to fill the semantic gap between the user question and the database schema and cope with the relatively small size of real databases with respect to research-oriented benchmarks. Scoping our research in a real scenario with the participation of real users is meant to address this challenge. Data privacy is an additional issue underlined in [22]; our approach does not require passing the organization's data to an external service; it is dealt with locally by the developed application.

The work described in [26,27] focuses on the use of LLMs for data visualization, the most recent approach used for Text-to-Vis. Wu et al. [26] compared fine-tuned models and inference-only models with state-of-the-art methods, demonstrating the higher performance of LLMs and analyzing where they failed. The comparison was performed on a benchmark dataset, while a user study with six participants majoring in computer science was used to assess the success rates of user querying for data viz; therefore, a usability study with real end users has not been carried out. Sah et al. [27] present an LLM-based system that generates a structured JSON object representing data attributes, analytic tasks, and relevant data viz; evaluation is performed in this case using an available dataset with human-generated utterance sets, while no real users are involved in the interaction with the system. Our work adopts an exploratory case study approach; thus, real users are involved throughout the research activities.

3. Phase 1: Answering RQ1

This section explores the first phase of the case study, consisting of the preparation of the interview, conducting the interviews, and a thematic analysis of the results of the interviews.

3.1. A Framework for an Exploratory Case Study for Data Analysis and Visualization Practices

As introduced in Section 1, the research method adopted in this study is an exploratory case study research with a single case. This choice, which is allowed by the methodology, was made due to the nature of RQ1: a single case study allowed for a deep exploration of visual analytics practices in different organizational departments; hence, it provided a wider glance at different kinds of key activities supported by visualization artifacts,

different decision-making granularities, and a higher chance to discover variability in expertise in the management of these tools.

Figure 1 depicts the context-dependent workflow of our framework, with its three perspectives: *Technology flexibility and adaptability*, *visual information literacy*, and *business strategy*. Each perspective is instantiated into an entity of the framework: *Routine_VizTask*, *Human_Intepretation*, and *Contextual_Rule*, respectively.

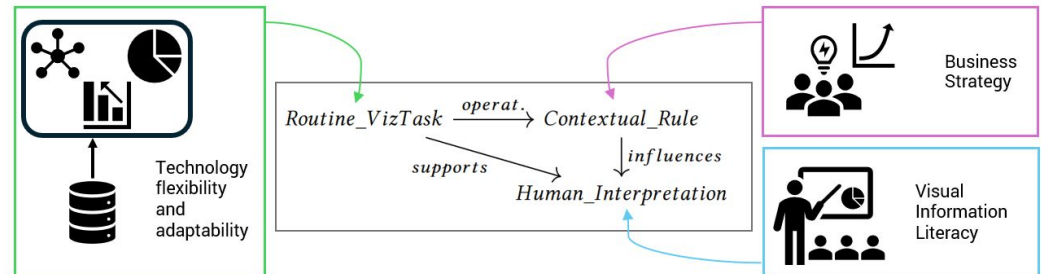


Figure 1. The context-dependent human–viz interaction framework [4] and its relationships with the three perspectives.

The tripartite set of dimensions in Figure 2 was explored in the case study, and is derived from [4], where the above framework for characterizing context-dependent perspectives in human–data viz interaction was designed.

Visual analytics practices are instantiations of the *Routine_VizTask* entity, which concerns the *technology* perspective. The adopted technology should be flexible and adaptable enough for performing visual analytics as often as the user’s role may require. Characterizing the *Routine_VizTask* entity means providing evidence of the users’ difficulties, in terms of challenges when using data viz, integrating data sources together, and choosing the performance measurements (e.g., a key performance indicator—KPI).

Visual analytics practices should boost the *Human_Intepretation* entity in order to answer questions about the organization’s conduct [28]. This step regards the second perspective of the framework, the *visual information literacy* of the user. Visual information literacy has been characterized as “the ability to properly process information related to data graphics, i.e., encoding information into data graphics and decoding information from data graphics” [2]. For example, more complex charts may be automatically provided based on the user’s domain expertise and ability to interact more intensively with data viz.

Visual analytics practices also operationalize the perspective of *business strategy*. This perspective can be regarded as an instantiation of the *Contextual_Rule* entity, and can be easily mapped to the concept of contextual knowledge sharing that a visual artifact promises to deploy. This third perspective may be characterized as the capability of the data viz to reflect the users’ and organization’s strategies. In terms of interactions, these strategies concern two aspects: the information quantity contained in the data and the users’ behavior during interactions with and communication through data viz [29,30]. The former requires a knowledge of how data visualizations respond to the query information from the point of view of the complexity of the data that need to be shown. Complexity depends on the number of entities represented in the data, the number of properties depicted for each entity, the type of data of each represented property (being nominal, ordinal, numeric discrete, and numeric continuous), and the level of detail (e.g., aggregated or not aggregated) [31]. Studying the systematic users’ behavior can help identify the areas of interest that each user focuses on. This behavior depends on their confidence with the visualization at hand, their current strategy, and their literacy level. Also, engagement and collaboration with colleagues may be dimensions related to all the three perspectives of the framework. Assessing the information complexity and the users’ behavior may bring

data visualization tools able to formulate on-the-fly query answering and personalized data viz interactions that could improve both the *business strategy* perspective and the *Human Interpretation* instantiations.

From the above description of the three perspectives and the related entities comprising the research framework, several dimensions may be identified and operationalized. Figure 2 depicts those that were investigated in the current case study. Each group of dimensions (12 in total) is related to one perspective and can be seen as one of the several aspects of each instantiated entity.

In this study, the choice of these 12 dimensions was functional to acquire knowledge from key informants to characterize users’ profiles. For example, investigating the *technology flexibility and adaptability* perspective may bring to light both behaviors and expectations. Observing what the most frequent interactions with data and business intelligence (BI) tools are, whether and to what extent AI is exploited, and what are the artifacts designed during the visual inspection of data in visual analytics routines, may shed light on their intrinsic and perceived quality. Investigating the *visual information literacy* perspective may be useful to scrutinize the level of expectation and confidence of individuals with visual analytics tools, the self-perception of expertise and skills, their educational background, and the degree of engagement with tools to manipulate data and related visualizations. Finally, investigating the dimensions related to the *business strategy* may help enter into behavioral stances, such as the visual analysis routines of interviewees, their level of collaboration with other colleagues, the final outcomes of this collaboration in terms of effective design, and their understanding of data manipulation and data viz design. The adoption of data-driven measures and KPI in the context at hand is also assessed through the dimension of routine and task exploration.

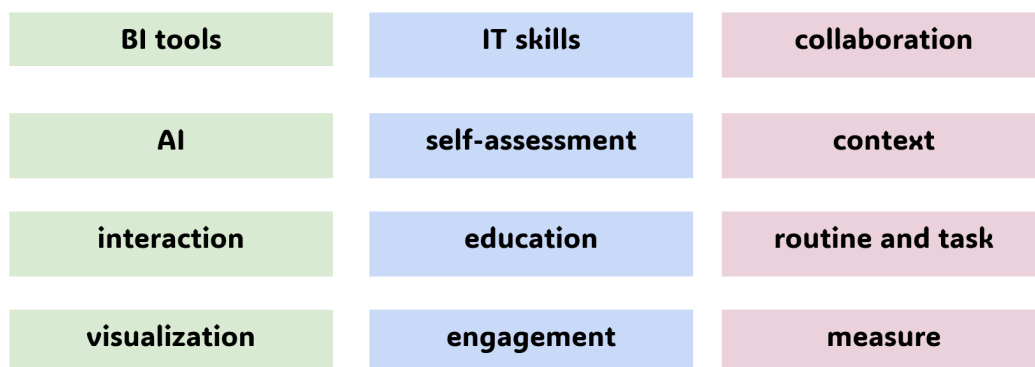


Figure 2. The 12 dimensions of the 3 perspectives adopted in the case study to drive the interviews. Colors identify the framework perspectives: Technology (green), visual information literacy (blue), and business strategy (pink).

3.2. The Semi-Structured Interviews

The organization investigated in this study is a multinational company in the manufacturing domain, and it is considered a large organization.

One of the authors conducted an internal audit to identify key informants based on the above framework and dimensions. A canvas of the proposed interview was made available to the potential key informants to facilitate their consent to take part in the study. After two of the authors had drawn up the interview canvas, a small pilot was carried out by the same authors with the organizational referents about the opportunity of posing those questions to allow visual analytics routines, data-driven practices, and decision-making processes to emerge. After the adjustments that followed from this informal discussion and approval, the two authors reviewed the canvas and designed its final version. The final canvas of the interview is reported in Table 1, where, for each question, a label associates

it with each of the 12 dimensions characterizing the three entities of the framework (the latent constructs of our exploratory analysis), which are reported in Figure 2.

Table 1. Interview canvas, including an association with the dimensions characterizing the entities seen as latent constructs of the framework depicted in Figure 2.

Interview Questions	Dimension
1. What types of data viz do you most frequently create for your analyses?	Visualization
2. Which BI tools or software do you use to generate these data viz?	BI tools
3. What are the main challenges you encounter when creating data viz?	Interaction
4. How do you verify the accuracy of the visualizations you create?	Education
5. Have you ever used advanced analytics features integrated into your BI tools?	IT skills
6. Have you ever used ChatGPT or other LLM-based systems to analyse data?	AI
7. Have you ever used ChatGPT or other LLM-based systems to analyse a data viz?	AI
8. How would you assess your skills in using tools and programming languages for data analysis and data viz creation?	Self-assessment
9. How would you describe your ability to read, interpret, and design effective data viz?	Self-assessment
10. Have you received or independently followed specific training on using these tools?	Education
11. Do you use predefined templates or create completely new data viz each time?	Engagement
12. Have you ever collaborated with colleagues to improve your skills in data viz creation?	Collaboration
13. Has using data viz changed your work?	Context
14. Which measures/statistics or KPIs do you find most useful in your analyses?	Measure
15. At what point in your work routine are you required to use visual analytics?	Routine and task

The aim of the interviews was to collect insights about data-driven practices inside the company and to allow the profiling of informants to identify target users and administer the experimental phase to them. The interviews were recorded, and two of the authors analyzed the transcripts. The methodology followed to analyze the interview material was affinity clustering on digital copies of transcripts. A hybrid approach was adopted for codification, combining inductive analysis, to yield preliminary results, with deductive thematic analysis [32]. Snippets of the transcripts were codified under the dimensions related to the framework entities and to each question asked during the interview. Disagreements between coders were managed during a post-coding phase, where the authors discussed until a final concordance was reached.

3.3. Results

After having carried out the interviews, the role of each of the eight key informants was identified. A summary is reported in Table 2, together with the classification of key informants based on their self-assessment skills in visual analytics and data viz design, as emerged from questions no. 8 and 9 of the interview.

Table 2. Key informants and their role, with labels codified for the answers to the self-assessment of data viz skills.

Key Informant	Role	Self-Assessment of Data Viz Skills
1	IT Engineer	Very high
2	Customer Care Manager	Low
3	Digital Marketing Manager	Medium
4	Production and Manufacturing Manager	Medium
5	Logistics and Demand Planning Expert	Low
6	Production Process Engineering Manager	High
7	Data Science Manager	Very high
8	Logistics, Planning, and Supply Chain Manager	High

From the analysis of the interviews, the following three key profiles emerged. All three profiles primarily use data viz for business decision making and internal reports. Each profile description is structured around the three perspectives of the theoretical framework.

3.3.1. Basic User

Technology: Regarding data viz usage, the basic user exploits pre-configured dashboards and Excel with simple line, bar, and pie charts.

Literacy: This user rarely modifies visualizations, and does not experiment or use ChatGPT, advanced analytics, or more powerful BI tools. Regarding self-perceived skills, this user wants to learn new tools and improve skills in the creation of appropriate data viz (even in Excel) to become more effective in data communication. Communication abilities with data viz are crucial for knowledge sharing, and, for this reason, this user consistently uses the same charts. Sometimes, the user experiments with more advanced data viz, but time is a challenging factor for self-improvement.

Business strategy: Regarding the frequency of interaction, data visualizations are produced on a monthly basis, while tables are used daily.

3.3.2. Intermediate User

Technology: Regarding data viz usage, the intermediate user uses multiple software tools offering pre-built data viz and dashboards, performing independent analyses from exported data. This user does not apparently need advanced analytics or ChatGPT-like functionalities, but is open to trying advanced tools, even by self-training.

Literacy: Regarding self-perceived skills and communication abilities with data viz, the intermediate user focuses on data communication for different stakeholders, using advanced data viz such as maps and bar charts.

Business strategy: The only limit in experimenting is the target audience. Regarding the frequency of interaction, this user performs daily analyses.

3.3.3. Advanced User

Technology: Regarding data viz usage, the advanced user creates customized visualizations with advanced languages or tools such as Python, R, Power BI, Power Query, and Trevor. This user has experimented with advanced analytics or ChatGPT-like tools.

Literacy: Regarding self-perceived skills and communication abilities with data viz, this user serves many other colleagues in preparing data viz. This user uses advanced data viz, such as stacked bar charts, gauges, scatter plots, heatmaps, and box plots.

Business strategy: In collaboration with colleagues and the target audience, this user provides a visual analytics tool to support their learning and promote their autonomy in data-driven

interpretation and decision making. Regarding the frequency of interaction, this user massively engages in data and visual analytics as part of their daily routine.

3.4. Challenges Common to the Three Profiles

A summary of the challenges common to the three profiles is reported in Table 3. Given the generalization of these challenges in spanning all three profiles, each challenge was more deeply related to the dimensions of the theoretical framework, rather than to the three framework perspectives only.

Table 3. Common challenges for the three key user profiles, anchored to dimensions of the theoretical framework.

Common Challenges to the Three Profiles	Technology	Literacy	Business
Heavily rely on manual work with Excel	BI Tools	IT skills	Routine and Task
Identify data viz for audience with different expertise is complex	Visualization	Education	Context
Tendence to simplify data viz and visual analytics	Visualization	Education	Context
BI tools are too rigid, thus Excel is preferred	BI Tools	IT skills	Routine and Task
Manual checking for data consistency is due to avoid poor decisions	BI tools	Engagement	Routine and Task
More dynamic solutions are complex and time is a barrier	BI Tools	IT skills	Routine and Task

4. Phase 2: Answering RQ2

An EUD environment fostering natural language interaction has been developed to allow the target users of the case study to easily create their data visualizations. The final goal was to observe their behaviors during the interaction with the prototype and collect their opinions and expectations to explore the role that such an EUD environment could play in a real organization where decision making is often supported by data analysis and visualization.

4.1. The AI-Supported Data Visualization Prototype

The design probe is a prototype web application that follows the client-server paradigm. Specifically, PostgreSQL is employed as the database management system. To leverage its LLM capabilities, the API provided by OpenAI is used, exploiting one of the most advanced models currently available, i.e., *GPT-4o*. The LLM is responsible for (i) generating the extraction SQL query based on the user’s intent, and (ii) recommending the most suitable visualization for the requested data and user profile.

To instantiate the model, a dynamic prompt engineering approach was adopted to ensure adaptability to individual users. The prompt structure consists of a fixed part that provides general system instructions about the task (i.e., “generate a SQL query and provide a suggested data visualization type”). At the same time, user-specific information is dynamically injected at runtime based on the logged user. These adaptive elements include user preferences, such as commonly used visualization types and interaction patterns, as well as the schema of the database relevant to the user’s daily activity.

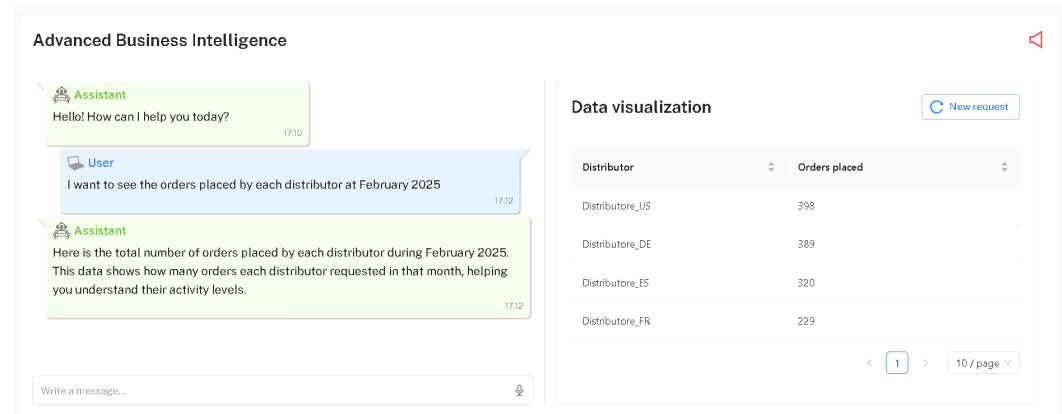
After the SQL query generation, the system undergoes a formal syntax validation to confirm adherence to SQL standards. Additionally, this validation ensures that the query is limited to data retrieval operations (*SELECT* statements) and does not include modification or deletion commands (*UPDATE* or *DELETE* statements), thereby mitigating potential security risks. To ensure response consistency, minimizing the model temperature parameter, which controls response variability, is crucial. A lower temperature setting enhances determinism, ensuring that identical inputs yield consistent outputs. For this case study, a temperature equal to 0.2 has been chosen.

The proposed solution enhances users' data management by leveraging an end-to-end workflow that integrates query generation (Text-to-SQL) and visualization suggestions (Text-to-Vis). The main objective is to simplify user interaction, ensuring that individuals without programming expertise can easily retrieve and analyze data related to their work practices.

In particular, the user initiates an interaction by submitting a natural language request. The user request, the model initialization instructions, the history of the previous messages in the same conversation, the user profile, and the database structure are processed and forwarded to the LLM. Having been instructed with contextual information about the database structure and application domain, the model generates the SQL query to extract the requested data and a suggested visualization format. After its validation, the SQL query is executed and the relevant data is extracted. It is important to note that it is not the LLM that executes the query but the backend of our web-based prototype, thus avoiding privacy and security issues regarding data dissemination. Without a specified user preference for data representation, the model autonomously selects the most appropriate visualization based on the requested data and the user profile. If the user needs to refine the request further, adjust the visualization format, or explore additional insights, they can, enabling an iterative and dynamic data analysis experience.

The user profiles described in the previous section have been employed to personalize the interaction with the application. The primary factors considered in this personalization include the type of chart generated, any additional explanations provided in the chat, and potential suggestions for alternative visualizations suitable for the requested data and the user profile.

For the *basic* profile (see Section 3.3.1), the preference for simple chart types is respected, and explanatory messages in the chat can be particularly useful. An example of interaction with the system is presented in Figure 3, where the simplest visualization type is displayed (i.e., a table), along with an additional explanation of it in the chat. No suggestions for more advanced visualizations are offered, as this profile targets users with limited data visualization experience.



The screenshot displays the 'Advanced Business Intelligence' interface. On the left, a chat window shows the following interaction:

- Assistant:** Hello! How can I help you today? (17:10)
- User:** I want to see the orders placed by each distributor at February 2025 (17:12)
- Assistant:** Here is the total number of orders placed by each distributor during February 2025. This data shows how many orders each distributor requested in that month, helping you understand their activity levels. (17:12)

On the right, a 'Data visualization' panel shows a table with the following data:

Distributor	Orders placed
Distributore_US	398
Distributore_DE	389
Distributore_ES	320
Distributore_FR	229

The table includes a 'New request' button and pagination controls showing '1' of '10 / page'.

Figure 3. An example of interaction with the prototype for the basic profile.

In the case of the *intermediate* profile (see Section 3.3.2), Figure 4 illustrates a slightly more complex visualization type (i.e., a bar chart), along with an explanatory message. Additionally, the system presented the user with a suggestion to view the data through an alternative representation (i.e., a pie chart).

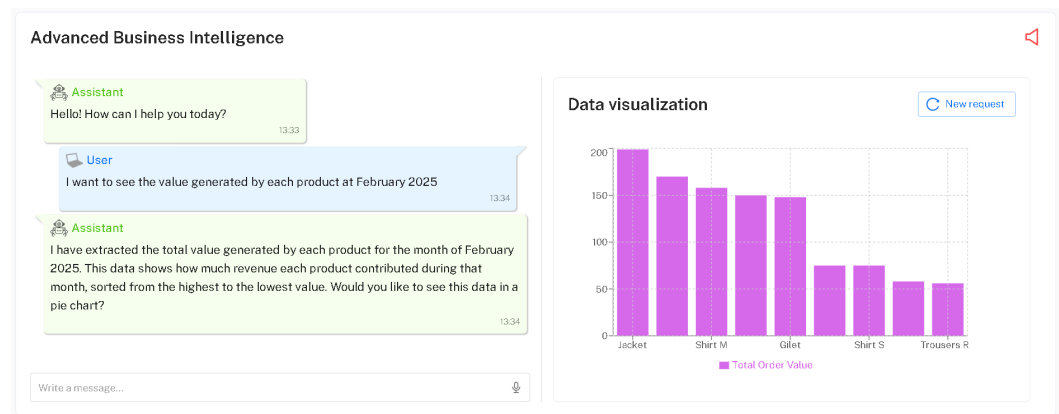


Figure 4. An example of interaction with the prototype for the intermediate profile.

Lastly, the *advanced* profile (see Section 3.3.3) is provided with more complex visualizations and no or limited explanations in the chat. For example, Figure 5 shows a heatmap visualization of the data requested by the user. In this case, no explanations are provided in the chat, reflecting the assumption that advanced users are already familiar with the visual representation. Similarly, no alternative suggestions are offered, as these users are expected to know what type of chart they require and may request the one they consider most appropriate for their needs.

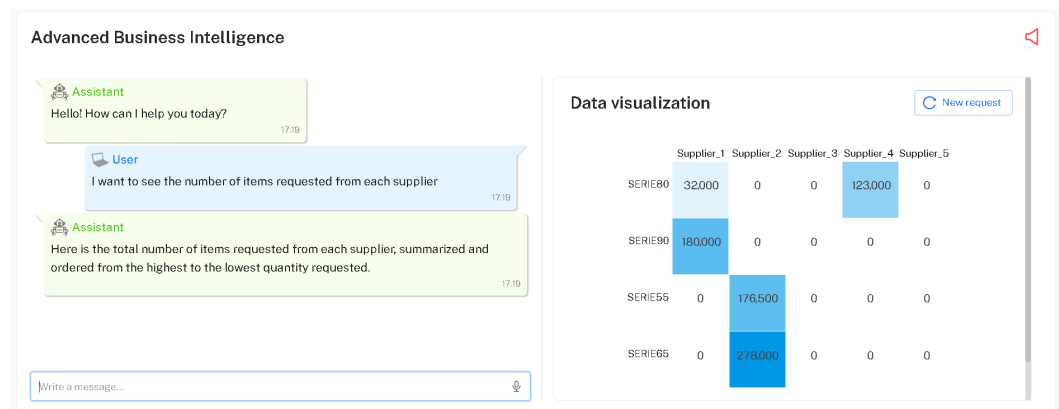


Figure 5. An example of interaction with the prototype for the advanced profile.

In addition to the profile-based personalization just described, users always have the option to further refine their requests in a dynamic and interactive way. Any user may request a refinement of the type of visualization; for instance, by asking to view the requested data in an alternative chart. They may also request a refinement of the data itself, such as specifying a particular time interval or applying specific filters. Furthermore, users can request a refinement of the chat response, asking for more detailed explanations of the extracted data, including details on how to better understand the provided visualization. The objective of this refinement process is to allow users to tailor the interaction based on their individual goals and familiarity with the data. By adjusting the type of visualization, selecting specific subsets of data, or requesting further clarification through the chat, users can gradually adapt both the content and the form of the output to better suit their needs.

4.2. The Interaction Experiment

The experimental interaction with the design probe consisted of an interactive session with the above prototype application. Three out of the eight key informants participated in the experiment as representatives of the three key profiles. It is important to underline that the interaction experiment did not aim at evaluating the usability of the application, but at

collecting further information to answer RQ2, and subsequently the overarching question of this research activity. Therefore, the experiment exploited qualitative methods to gain insights about the behaviors and expectations of different types of knowledge workers accomplishing their decision-making goals through a visualization tool enhanced with EUD and LLMs.

Based on the key profiles identified, the experiment was divided into a common part and a profile-based part. The common part consisted of an explanation to the participants of the types of visualization that the AI-based prototype was capable of providing: numeric or percentage KPIs, tables, line graphs, bar charts, area charts, pie charts, heatmaps, and scatter plots. Participants were encouraged to use natural language to request data visualizations and to receive explanations of the generated data visualizations, fostering an environment where they could talk aloud and execute tasks that were similar to their routine tasks. Participants should have to interact with the system in written natural language: on the right panel (see, for example, Figure 3), they were able to see the generated data viz, while on the left panel, the chat was at their disposal to continuously interact.

In the profile-based part, each experimental task script diverged based on the users' profile domain of expertise, data viz literacy, and on the complexity of interactions that form their visual analytics routine. For example, a task could be as easy as generating a KPI and a basic data viz, e.g., requesting a table of distributor orders, identifying top distributors by order quantity in a specific month, and observing order variations over time. A task could be more complex, such as requesting a comparative analysis of temporal changes in the data, e.g., compare product demands by customers, visualize sales volume changes over time, and assess market dynamics through pie charts or a breakdown view of active countries in e-commerce. More advanced tasks might include an in-depth analysis of supply chain data and vendors' performances, e.g., to evaluate supply chain efficiencies, understand product line order trends, and explore the impact of a supplier's performance on delivery delays through a series of cascaded visual analyses.

The four tasks assigned to the participants and customized to the three user profiles are reported in Tables 4–6.

Table 4. Basic user experiment.

Basic User Experiment
<p>Task 1: Ask to see the table of distributors' orders. You also want to know which distributor ordered the most products in January 2025 and what that number is.</p>
<p>Task 2: You want to see the variation in the number of orders over time. You want to see the service level for the distributors.</p>
<p>Task 3: You want to see a bar data viz to compare the orders made by each distributor in February 2025. Now you want to compare January and February. You want to see a pie data viz with the product categories in 2025. Now you want to visualize the breakdown by product.</p>
<p>Task 4: Why is our service level to distributors low in some cases? What actions can we take to ensure a more efficient processing of orders?</p>

Table 5. Intermediate user experiment.

Intermediate User Experiment
<p>Task 1: Ask to see the table of customers' orders. You also want to know which product was the most requested in January 2025 and the quantity.</p> <p>Task 2: You want to see the variation in sales volume over time. You want to compare the value generated by the products.</p> <p>Task 3: You want to see a bar data viz to compare the value generated by each type of product in February 2025. Now you want to compare January and February. You want to see a pie data viz of the most active countries in e-commerce. Now you want to visualize the best-selling products in Italy.</p> <p>Task 4: You need to create a report on e-commerce sales for January 2025. The goal is to show the value of orders and the quantity purchased by analyzing the differences between product categories, countries, and payment methods.</p>

Table 6. Advanced user experiment.

Advanced User Experiment
<p>Task 1: Ask to see the table of orders to suppliers. You also want to know which product line has the most orders and the respective quantity.</p> <p>Task 2: You want to see the value of orders for each product line. You want to see the number of items per product line.</p> <p>Task 3: You want to see a bar data viz to compare the ordered quantity and the requirement for each product line. Now you want to compare the ordered and received quantities. You want to see a pie data viz showing the value of orders for each supplier. Now you want to visualize the number of items requested to each supplier.</p> <p>Task 4: You want to understand how to optimize the procurement process. Therefore, you ask which suppliers have the most significant impact on delivery delays and which items are most involved.</p>

A final questionnaire was administered to the users, where each answer was an open-ended one. As each user profile had different tasks to accomplish, there was no standard evaluation test that could be uniform across different users' profiles. Hence, we preferred to investigate the qualitative opinions of users after using the system. The questionnaire items are reported in Table 7 (some of its questions are inspired by the Bot Usability Scale presented in [33]).

The experiment aimed at testing what behaviors and expectations emerge from key users based on the ability of the system to drive them through (i) the exploration of a dataset based on their specific needs and expertise; and (ii) the flow of interaction with a data viz generated by the system, which they contribute to designing for visual analytics purposes.

Table 7. Final questionnaire.

No.	Question
1	Was it easy for you to express your needs to the system?
2	Was the system response clear and understandable?
3	Was the system response satisfactory and accurate, or would you have preferred another response?
4	Did the system seem to remember previous interactions or your preferences?
5	Do you think this system could improve your work in creating data viz?
6	Would you trust using this system to perform your work?

During the experiment, users could receive some suggestions on the tasks that could be executed with the system, but were essentially left free to express themselves to require a specific data viz, to refine their request, or to obtain an explanation of the results. For privacy reasons, the system did not have any knowledge of the database, but only of its schema. Based on the dataset schema, an SQL query to the dataset was executed by the system, and a data viz based on the user's request was created and proposed in the right panel of the prototype.

Each individual session was carried out remotely through Microsoft Teams. The entire session, including audio, video, and automatic transcription, was recorded for later analysis. One researcher played the role of facilitator, introducing the project, explaining the nature of the experimental activity, and sharing a document with the assigned four tasks. The facilitator also took notes of the user's behaviors and comments and at the end proposed the questions of the final questionnaire, annotating the received responses. Another researcher managed the technical aspects of the experiment, selecting the correct system version and database for that specific user. Each user session lasted about 30 min.

4.3. Results

The assigned tasks were completed successfully by all three users, who obtained the desired visualization for their data. Only the basic user encountered some difficulties during the interaction with the prototype, especially at the beginning of the experimental session; this required the intervention of the facilitator, who provided suggestions about the possible natural language requests accepted by the system. This problem and other significant aspects are discussed in the following subsections.

Recordings, direct observation, and note-taking were adopted to collect qualitative data during the interaction with the system by the key users. Transcripts and notes were then integrated with the answers to the open-ended questions of the post-questionnaire. A thematic analysis was then performed by two researchers on the collected material, following a deductive and semantic approach. The extracted codes were discussed in a meeting, which also served to identify the emerging themes. These themes are reported in the following, distinguishing them based on the user profile and reporting participants' quotes when considered useful.

4.3.1. Findings for the Basic User

The following main themes emerged for the basic user profile.

Prompting requires training. During interaction with the system, the basic user often asked us for confirmation before proceeding with the task at hand; this happened in particular when task 1 started, as there were some hesitations in formulating the natural language requests to the prototype. Crafting prompts may be difficult for users who rarely use LLMs, leading to complex queries that should be refined at each new iteration. Despite these initial difficulties, the user found that interacting was beneficial, mainly when a new

data viz was exploited for the first time, i.e., a heatmap. In fact, the user observed: *“Unlike searching on the Internet, here I was working with data I’m already familiar with, so asking for an example based on that made learning much faster. As soon as I saw the table, I understood how it worked and thought that this was nice and useful, I could actually use it in the future”*.

Problem-solving based on trial and error. Task execution was characterized by a trial-and-error style. For example, when the system generated a line chart instead of the intended bar chart for analyzing suppliers and their service level, the user successfully identified the issue and solved it. Likewise, some attempts to incorporate multiple categories in a bar chart were unsuccessful, requiring continuous adjustments and many prompts. In synthesis, the prompting activity often required iterative refinement to yield the desired result. Nonetheless, the user seemed to acknowledge and appreciate the tool capability of providing data viz and data manipulation, recognizing some advantages over human error for some tasks (*“Yes, it definitely makes fewer mistakes than I do when I manually copy formulas over and over. Once the data and databases are loaded into the system, there’s significantly less room for human error”*).

Expectations about how LLMs work. Managing dates within the system proved to be another dull activity, and the expected handling was not always clear (*“I cannot really read the dates... Oh, right. You can hover over the line to see specific values. But I asked for data from December, and it only shows February. Is it just me not reading it correctly?”*). The user also showed a strong wish for accuracy by frequently checking interpretations and emphasizing proper wording, having little confidence and trust in the LLM.

Difficulties in feature discovery. Some functionalities, like sorting data, did not initially come to mind but were successfully used after our suggestions. In the same vein, there was a lack of familiarity with many types of available data viz, but once reminded of available options, the user actively explored new data visualizations (such as the heatmap mentioned above) and extended this understanding to other tasks. The basic user often expressed challenges in interpreting data visualizations, sometimes unaware of interactive features like the mouseover functionality. Another remark was on the importance of precisely entering dataset column names (*“Oh, I think it’s because I didn’t write it correctly... I was supposed to use a capital letter. I made a mistake”*). A key request was the possibility to place different data visualizations side by side for easier comparison.

Ease of natural language interaction. The basic user found it easy to interact with the system for expressing their needs and making requests (*“In our work, we usually start from a dataset, then it’s all about pivoting, filtering, searching, ... That’s why I think this could be extremely convenient: you just use a single sentence instead”*). It was especially appreciated how the system accumulates knowledge about the interaction and the fact that there was no need to formulate the previous requests again.

Quality of the output. Mostly, the system responded in an understandable way, especially with some basic visualization, i.e., table, and advanced data viz requests, i.e., heatmaps. The responses of the system were considered satisfactory and accurate (*“The output matched my expectations, it showed exactly what I was looking for”*).

Usefulness and effectiveness of the tool. Regarding the extent to which the system could improve daily work, the opinion was positive. It was found convenient to use a tool of this kind, rather than searching for help online (*“It’s definitely more convenient than searching online; since the system already works with data I’m familiar with, learning is much faster”*). The platform already had information available and had already made examples on available data. Compared to a human being, the system was perceived as making fewer mistakes.

4.3.2. Findings for the Intermediate User

Different themes emerged for the intermediate user.

Prompting encourages exploration of the system's capabilities. The intermediate user was quicker in grasping how the system worked, readily trying different types of requests to better understand what the output was. Since the beginning, this user intuitively knew what to write and how to format the inputs to enter the desired measures. Engagement in trial and error was the main attitude to assess the capabilities of the system and to experiment with various request formats. Data viz choices included bar charts, line charts, pie charts, tables, and scatter plots. However, when attempting to design a scatter plot, the initial result was not the one intended, and the user soon realized that two measures were required (*"I asked for a scatter plot to show the correlation between sales and number of orders, but then I realized I needed to define two separate measures, one for each axis"*). After this attempt, the user proceeded with an alternative data viz.

Proactive management of the system's limitations. Some frustration arose when the system failed to display expected data, such as a KPI. However, the user seemed to stick to the given tasks, trying to improve data viz independently, e.g., by modifying the granularity of the data, transforming monthly orders into weekly orders, and the like. This user proactively refined the analyses by introducing new measures and modifying requests as some insights developed during the interaction; for instance, after thinking aloud *"Hmm, okay, but what if I wanted to ask, for example, to include the percentage variation?"*, the user successfully obtained the output update.

Dealing with complex data visualization requests. There were moments when the system did not immediately respond as expected. For example, when trying to design a heatmap but a bar chart was returned instead, and upon specifically asking, the user obtained the correct data viz. A second example was when a complex request was also attempted, highlighting both the difficulty in making the system become responsive in this respect and realizing how challenging it could become to articulate complex analytical needs. In general, when dealing with complex data visualizations, difficulties in fully expressing analytical intentions were expressed (*"There's also a human limitation, obviously, when it comes to explaining things. I'm one of those people who knows exactly what I want in my head, but explaining it..."*). Furthermore, there was uncertainty regarding the system's capabilities and limitations, leading to a continuous process of request exploration and subsequent refinement.

Misunderstandings addressed easily. On the one hand, when faced with misunderstanding from the system, the user independently refined or rewrote prompts, adjusting the approach accordingly. On the other hand, when occasional input errors occurred, the user noticed that the system accurately interpreted their intent (*"Oh, but it even corrected the name I mistyped. Nice, I got it"*). In general, the response of the system was found to be clear, also because there was feedback even when the system was unable to perform the request.

Trust in the system. The system was found to be adequate from the satisfaction and accuracy perspectives, but further refinements to the prototype were perceived as necessary. Therefore, trust in the system could be gained provided that further improvements in the interaction modalities are made (*"I imagine these systems still need to be refined and trained, but the beginning looks promising. After a proper testing phase, yes, I could trust it"*). Despite this, the tool was evaluated as substantially improving daily work.

4.3.3. Findings for the Advanced User

Finally, additional themes associated with the advanced user were identified.

Effectiveness and efficiency of human-AI interaction. During the interaction with the system, the advanced user always yielded precise and detailed results, often testing the

system capability to interpret datasets accurately and demanding greater transparency in data representation. The focus of the interaction soon extended beyond basic interactions, aiming to optimize a workflow through advanced data viz requests for correlation analysis. When requesting correlations between measures, the system sometimes generated an incorrect scatter plot, leading the user to make request formulations more intricate to try to yield the expected result. The system efficiency was appreciated in its capacity for reducing the time spent on manual Excel processing, recognizing its potential for streamlining data analysis and decision making (*"I'd normally spend half an hour doing this in Excel, but here I can see it instantly"*). Using the system was found to be a notable improvement in daily work, especially in saving time (*"It saves me a lot of time, especially when I need to perform monthly comparisons. This also spares me from building pivot tables every single time"*).

Retaining human control over the AI-based system. With good skills in prompt engineering, this user emphasized that the system's effectiveness was heavily influenced by how well requests were formulated (*"I need to be careful with how I phrase the request, otherwise it gives me something different"*). Prompts tended to be lengthy and complex, reflecting a desire for full control over data viz and system outputs. There was also a need to go beyond data viz and understand how trustworthy the system was in terms of checking the data at the source from the table.

Strengths and weaknesses of LLM-based interaction. The system was found to be able to express needs and answer requests adequately, as seamlessly as interacting with a common chatbot (*"The experience is like chatting: I talk to it as if it were a person. There's no need to use strange commands"*). The clarity and understandability were judged quite well, despite the fact that the system did not understand when it had made a mistake, and it was not easy to try to communicate the mistake. Expectations went beyond the system's capabilities, showing a wish to exploit this kind of system for complex analytical needs, which are, however, still difficult to support with this technology.

Trustworthiness and transparency of the system. After thorough testing, the user expressed trust in using the system for work-related tasks, provided that a proper verification of the dataset was conducted before a full reliance on the system could be admitted. While SQL query visibility was initially valued, this user acknowledged that, once trust is established, such transparency may become unnecessary (*"As long as the results are consistent, I do not feel the need to verify the code every time"*).

5. Discussion

5.1. Empirical Analysis

The exploratory case study research described in this paper allows us to respond positively to the overarching research question reported in Section 1. The AI-supported EUD environment developed for the case study is aimed at democratizing data access while prioritizing user-defined preferences and computational efficiency. The natural conversation capability of the model ensures that different users (e.g., supply chain managers, customer care managers, and data analysts) can access insights without requiring technical expertise in querying a database or designing a complex data viz. This flexibility may support more rapid decision-making processes in a dynamic environment.

All users, notwithstanding their expertise in data viz, considered the system useful and able to speed up their daily activities.

Different behaviors and expectations emerged during the interaction experiment depending on the user profile, providing interesting insights for reflection on the characteristics that a valuable LLM-based EUD environment should offer in real contexts.

Design guidelines emerged from the analysis of the users' profiles and the interaction experiment themes. They are outlined and discussed in the following sections.

5.2. Design Guidelines for AI-Supported EUD for Data Visualization

Design guidelines emerged from the empirical analyses of Sections 3.3 and 4.3 and are reported in Table 8, where they are prioritized depending on the target users of the system. In the next paragraphs, they are analyzed from the theoretical point of view, based on the initial framework and the tradeoff between the adaptivity and adaptability of AI, and from the practical point of view, based on concrete actions for the context and practices at hand.

Table 8. Guideline priorities for the three key user profiles (1 = high priority, 5 = lower priority).

Design Guideline	Basic	Inter.	Adv.
Integrate multimodality in the AI interface	1	5	5
Support advanced and iterative request refinements	3	1	3
Provide support for complex data viz requests	5	2	1
Customize systems	2	3	4
Leave control to the user	4	4	2

5.2.1. Theoretical Implications

Integrate multimodality into the AI interface. The empirical analysis has revealed the gap between common BI tools functionalities and the need for flexibility and usability, which may be ensured by leveraging multimodality. These findings suggest we should investigate the technology perspective and all of its dimensions more, in the direction of both *adaptivity* and *adaptability*, depending on the user profile.

Support advanced and iterative request refinements. This is in line with human–AI interaction guidelines proposed in [34], where the initial settings of AI should clearly provide the potential and limitations of the system actions. When misunderstandings occur, clear motivations are advocated and, over time, learning from and adapting to the users’ behavior may improve the sense of engagement, alignment, and satisfaction that improve task automation, adoption rates, and design practices. These findings suggest we should further investigate the literacy aspects and the contribution of AI systems in supporting self-assessment, IT skills, education, and engagement dimensions. Furthermore, design prioritization based on user profile may help tune personalizations of AI systems, thus its *adaptability*.

Provide support for complex data viz requests. Within current LLMs, there are still limitations regarding the management of complex requests from users, and having a coherent and consistent flow of reasoning to fix LLM errors, misunderstandings, and bad behaviors. These findings may suggest investigating the business strategy and literacy of users further, in order to systematically identify where the key issues are, and provide remedies to them through the AI. This implies strengthening the *adaptivity* aspects.

Customize systems. Many suggestions may come under the broader scope of system customization. They are related to the property of these systems to be “user-aware” in terms of their skills, literacy, and expertise. Training the model with users’ interactions and feedback could represent a direction to investigate to obtain a true customization of the system. The perspectives to be more investigated in this respect are the literacy and business strategy, in all of their dimensions, through the *adaptivity* aspects of AI.

Leave the control to the user. The theme of control emerged often in connection with trust and transparency. For these reasons, more clear and explicit displays of dataset measures, of the logic of the columns selected, and better handling of ambiguous selection are important recommendations for a system dealing with the full lifecycle of data retrieval and visualization. These findings suggest we should further investigate the technology

and business strategy perspectives. One may consider the possibility of an AI immersed in the context, as a collaborator throughout the data lifecycle, which helps users in their routine tasks and consistency checking of queries and results. This suggests strengthening the *adaptability* aspect of AI.

5.2.2. Practical Implications

Integrate multimodality in the AI interface. Regarding the system's capability, potential, and limitations, natural language must be complemented with a suitable graphical user interface that guides users in the system usage and allows them to deepen data analysis without additional prompts that could be difficult to formulate. This integration should also include a more guided interface with clearer suggestions on how to formulate requests and more detailed explanations of available data viz types and their best use cases. This integration would also improve the usability of the system. More transparent system behaviors (also exploiting visual feedback) must be studied to avoid users having to check for output correctness through other tools.

Support advanced and iterative request refinements. As emerged in the experiment regarding the trial-and-error behaviors of all three user profiles, and the necessity to address misunderstandings, an LLM-based EUD environment for data viz should provide advanced capabilities for iterative refinement of requests, supported by informative feedback about what the system "knows" and can do.

Provide support for complex data viz requests. On the one hand, there is a need to support intermediate and advanced users in the creation and refinement of complex data viz to support complex analytical tasks effectively. This complexity may be supported, e.g., by providing a suitable display of dataset attributes within the interface, giving AI-based suggestions about the queries that are feasible with certain data or the data viz types that could be obtained with them. On the other hand, the correct interpretation of complex queries formulated by the user should be guaranteed by the system.

Customize systems. Customization features may include, for example, interactive tutorials for first-time users; enhanced text interpretation, potentially through auto-completion; refinement feedback, ensuring readability and comprehension of visualizations; and deeper numerical insights, with incremental details upon the user's demand, showing trend-based insights, numerical evidence, and clear indication of interactivity of charts, whenever available, through explicit indications.

Leave the control to the user. Readability remains a priority, as it allows users to exercise full control of visual attributes, the order of displayed elements, and of 2D vs. 3D visualizations. Transparency in all the phases of the data lifecycle is a strong expectation. In particular, the more complex the generated SQL query, the higher the need to provide a design allowing full control of the further refinements, processing, and display of the data.

5.3. Limitations of the Work

Some limitations affect this study. The use of one case study vs. more case studies may be seen as an obstacle to the generalizability of our results. On the other hand, this gave us the opportunity to gain an extensive knowledge about a company, its specific business practices and visual analytics routines, which allowed us to control many variables that are usually difficult to control in an *in vivo* environment; to distill the most important practices, behaviors, and expectations from all of the key informants in order to focus on visual analytics practices; to complement and triangulate the information given by each key informant when collaborating with another key informant in daily practices, so as to gain knowledge from studying their interactions; and to gain a deep and more situated knowledge of how to hypothesize design guidelines to be exploited in that specific context.

Another limitation may reside in the number of participants involved in the experimental part. Even though we carefully selected the three key users to reflect the full spectrum of behaviors, expectations, interaction styles, and communication ability that emerged with the interviews in phase 1, we recognize that further information might have been obtained by involving a higher number of participants. However, due to the nature of the exploratory case study, we claim that limiting this experimental part to a few participants, and to an in-depth qualitative analysis, allowed us to better grasp the nuances of interaction with an LLM-based EUD system for a specific range of data analysis and visualization tasks. The obtained results will guide the design and development of a full prototype, to be quantitatively evaluated with an adequate number of participants working in the same company or in other organizations, with the purpose of demonstrating the generalizability of the approach.

A further limitation may regard the use of a design probe instead of a full user study. Exploratory case study research requires having hypotheses to formulate in the end, and a user study can be seen as a subsequent step, i.e., the verification of a hypothesis about the usability of a system. In this study, we aimed at posing the correct hypotheses about the motivations and relevance of exploiting an LLM-based support for EUD in data visualization. Thus, we were a step behind the design of a full prototype. As the probe term suggests, we designed an underspecified artifact that could be immersed in everyday practices to challenge, engage, and solicit reactions and responses from the users [7].

A final limitation may concern the consistency validation between items and dimensions, due to the small number of participants. However, the framework exploited to describe the perspectives to be investigated provided a robust methodological and theoretical anchor to the results and findings.

6. Conclusions

The contribution of this study is both practical and theoretical. The practical contribution is the empirical analysis of the visual analytics process in the mix of business practices, whenever “doing so by other means is difficult or impossible” [6] (p. 1), and local in-depth transparency becomes a priority over generalizability of results. The theoretical contribution is based on the application of a methodology whose “intuitive approach” allowed the emergence of the “hitherto unknown [...] when phenomena are studied that are as yet unrecognized” [6] (p. 1). Put simply, design guidelines were drawn from the distillation of the interviews, practice observation, and experiments carried out.

Going back to our overarching question, the emergence of key user profiles, the observation of the users’ behaviors and interaction patterns, and the collection of feedback from the users’ experience with the system have paved the way for a thorough exploration of possible ways to obtain EUD-supported data visualization practices with the aid of LLMs. In particular, we have shown how we could support data analysis and visualization practices in business contexts with EUD enhanced by LLMs, and derived clear indications for the development of this type of application, which emerged from a thorough analysis of routines, strategies, and micro-tasks of key users, with different skills, roles, and interaction characteristics. AI-supported EUD tools for data visualization and decision making should provide multimodality, customization, trustworthiness, and control to let humans be at ease with them, in a more symbiotic way than current systems. Having studied the issue in its original context may result in valuable aspects that may, however, be difficult to quantify. On the other hand, the depth and flexibility of this approach may benefit other, mostly overlooked, qualitative aspects of the scientific challenge posed by underestimated issues, such as those related to everyday practices, subjects interactions, self-perception, and the like.

Author Contributions: Conceptualization, all; methodology, D.F. and A.L.; software, L.G.; validation, all; formal analysis, all; investigation, all; resources, S.B.; data curation, S.B.; writing—original draft preparation, A.L., D.F. and L.G.; writing—review and editing, A.L. and D.F.; supervision, D.F. and A.L.; project administration, A.L.; funding acquisition, A.L. All authors have read and agreed to the published version of the manuscript.

Funding: Research funded by the European Union—Next-Generation, Mission 4 Component 1 CUP D53D23008690006, project name: “Characterizing and Measuring Visual Information Literacy” ID 2022JJ3PA5, CUP D53D23008690006.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: No further data than those available in the paper is provided (e.g., interviews transcripts), due to privacy reasons related to a non-disclosure agreement between the authors and the Company.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Barricelli, B.R.; Fogli, D.; Locoro, A. EUDability: A new construct at the intersection of End-User Development and Computational Thinking. *J. Syst. Softw.* **2023**, *195*, 111516. [\[CrossRef\]](#)
2. Locoro, A.; Fisher, W.P.; Mari, L. Visual information literacy: Definition, construct modeling and assessment. *IEEE Access* **2021**, *9*, 71053–71071. [\[CrossRef\]](#)
3. Otto, J.T.; Davidoff, S. Visualization Artifacts are Boundary Objects. In Proceedings of the 2024 IEEE Evaluation and Beyond-Methodological Approaches for Visualization (BELIV), St. Pete Beach, FL, USA, 14 October 2024; pp. 81–88. [\[CrossRef\]](#)
4. Buono, P.; Locoro, A. Modelling Data Visualization Interactions: From Semiotics to Pragmatics and Back to Humans. In Proceedings of the AVI '20: 2020 International Conference on Advanced Visual Interfaces, Salerno, Italy, 28 September–2 October 2020. [\[CrossRef\]](#)
5. Fischer, G. Adaptive and Adaptable Systems: Differentiating and Integrating AI and EUD. In *End-User Development*; Spano, L.D., Schmidt, A., Santoro, C., Stumpf, S., Eds.; Springer: Cham, Switzerland, 2023; pp. 3–18.
6. Mills, A.J.; Durepos, G.; Wiebe, E. *Encyclopedia of Case Study Research*; Sage Publications: Los Angeles, CA, USA, 2010.
7. Boehner, K.; Vertesi, J.; Sengers, P.; Dourish, P. How HCI interprets the probes. In Proceedings of the CHI '07: SIGCHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 28 April–3 May 2007; pp. 1077–1086. [\[CrossRef\]](#)
8. Battle, L.; Scheidegger, C. A Structured Review of Data Management Technology for Interactive Visualization and Analysis. *IEEE Trans. Vis. Comput. Graph.* **2021**, *27*, 1128–1136. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Kandel, S.; Paepcke, A.; Hellerstein, J.M.; Heer, J. Enterprise data analysis and visualization: An interview study. *IEEE Trans. Vis. Comput. Graph.* **2012**, *18*, 2917–2926. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Sambasivan, N.; Kapania, S.; Highfill, H.; Akrong, D.; Paritosh, P.; Aroyo, L.M. “Everyone wants to do the model work, not the data work”: Data Cascades in High-Stakes AI. In Proceedings of the CHI '21: 2021 CHI Conference on Human Factors in Computing Systems, Yokohama, Japan, 8–13 May 2021. [\[CrossRef\]](#)
11. Gu, K.; Shang, R.; Althoff, T.; Wang, C.; Drucker, S.M. How Do Analysts Understand and Verify AI-Assisted Data Analyses? In Proceedings of the CHI '24: 2024 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 11–16 May 2024. [\[CrossRef\]](#)
12. Yin, P.; Li, W.D.; Xiao, K.; Rao, A.; Wen, Y.; Shi, K.; Howland, J.; Bailey, P.; Catasta, M.; Michalewski, H.; et al. Natural language to code generation in interactive data science notebooks. *arXiv* **2022**, arXiv:2212.09248. [\[CrossRef\]](#)
13. Dimara, E.; Zhang, H.; Tory, M.; Franconeri, S. The Unmet Data Visualization Needs of Decision Makers Within Organizations. *IEEE Trans. Vis. Comput. Graph.* **2022**, *28*, 4101–4112. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Pantazos, K.; Lauesen, S.; Vatrapu, R. End-User Development of Information Visualization. In *End-User Development*; Dittrich, Y., Burnett, M., Mørch, A., Redmiles, D., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 104–119.
15. Heer, J.; van Ham, F.; Carpendale, S.; Weaver, C.; Isenberg, P. Creation and Collaboration: Engaging New Audiences for Information Visualization. In *Information Visualization: Human-Centered Issues and Perspectives*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 92–133. [\[CrossRef\]](#)
16. Pantazos, K.; Lauesen, S. End-User Development of Visualizations. *Electron. Imaging* **2016**, *28*, art00007. [\[CrossRef\]](#)
17. Zhang, W.; Wang, Y.; Song, Y.; Wei, V.J.; Tian, Y.; Qi, Y.; Chan, J.H.; Wong, R.C.W.; Yang, H. Natural Language Interfaces for Tabular Data Querying and Visualization: A Survey. *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 6699–6718. [\[CrossRef\]](#)

18. Zeng, J.; Lin, X.V.; Hoi, S.C.; Socher, R.; Xiong, C.; Lyu, M.; King, I. Photon: A Robust Cross-Domain Text-to-SQL System. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Online, 5–10 July 2020; pp. 204–214. [[CrossRef](#)]
19. Song, Y.; Wong, R.C.W.; Zhao, X.; Jiang, D. VoiceQuerySystem: A Voice-driven Database Querying System Using Natural Language Questions. In Proceedings of the SIGMOD '22: 2022 International Conference on Management of Data, Philadelphia, PA, USA, 12–17 June 2022; pp. 2385–2388. [[CrossRef](#)]
20. Tang, J.; Luo, Y.; Ouzzani, M.; Li, G.; Chen, H. Sevi: Speech-to-Visualization through Neural Machine Translation. In Proceedings of the SIGMOD '22: 2022 International Conference on Management of Data, Philadelphia, PA, USA, 12–17 June 2022; pp. 2353–2356. [[CrossRef](#)]
21. Luo, Y.; Li, W.; Zhao, T.; Yu, X.; Zhang, L.; Li, G.; Tang, N. DeepTrack: Monitoring and exploring spatio-temporal data: A case of tracking COVID-19. *Proc. VLDB Endow.* **2020**, *13*, 2841–2844. [[CrossRef](#)]
22. Hong, Z.; Yuan, Z.; Zhang, Q.; Chen, H.; Dong, J.; Huang, F.; Huang, X. Next-Generation Database Interfaces: A Survey of LLM-based Text-to-SQL. *arXiv* **2024**, arXiv:2406.08426.
23. Ma, P.; Wang, S. MT-teql: Evaluating and augmenting neural NLIDB on real-world linguistic and schema variations. *Proc. VLDB Endow.* **2021**, *15*, 569–582. [[CrossRef](#)]
24. Gao, D.; Wang, H.; Li, Y.; Sun, X.; Qian, Y.; Ding, B.; Zhou, J. Text-to-SQL Empowered by Large Language Models: A Benchmark Evaluation. *Proc. VLDB Endow.* **2024**, *17*, 1132–1145. [[CrossRef](#)]
25. Li, H.; Zhang, J.; Liu, H.; Fan, J.; Zhang, X.; Zhu, J.; Wei, R.; Pan, H.; Li, C.; Chen, H. CodeS: Towards Building Open-source Language Models for Text-to-SQL. *Proc. ACM Manag. Data* **2024**, *2*, 127. [[CrossRef](#)]
26. Wu, Y.; Wan, Y.; Zhang, H.; Sui, Y.; Wei, W.; Zhao, W.; Xu, G.; Jin, H. Automated Data Visualization from Natural Language via Large Language Models: An Exploratory Study. *Proc. ACM Manag. Data* **2024**, *2*, 115. [[CrossRef](#)]
27. Sah, S.; Mitra, R.; Narechania, A.; Endert, A.; Stasko, J.; Dou, W. Generating Analytic Specifications for Data Visualization from Natural Language Queries using Large Language Models. *arXiv* **2024**, arXiv:2408.13391. [[CrossRef](#)]
28. Hong, S.R.; Hullman, J.; Bertini, E. Human Factors in Model Interpretability: Industry Practices, Challenges, and Needs. *Proc. ACM Hum.-Comput. Interact.* **2020**, *4*, 68. [[CrossRef](#)]
29. Kahneman, D.; Sibony, O.; Sunstein, C.R. *Noise: A Flaw in Human Judgment*; Harper Collins Publishers: Dublin, Ireland, 2021.
30. Dimara, E.; Franconeri, S.; Plaisant, C.; Bezerianos, A.; Dragicevic, P. A task-based taxonomy of cognitive biases for information visualization. *IEEE Trans. Vis. Comput. Graph.* **2018**, *26*, 1413–1432. [[CrossRef](#)] [[PubMed](#)]
31. Holliman, N.S.; Coltekin, A.; Fernstad, S.J.; McLaughlin, L.; Simpson, M.D.; Woods, A.J. Visual entropy and the visualization of uncertainty. *arXiv* **2019**, arXiv:1907.12879.
32. Fereday, J.; Muir-Cochrane, E. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *Int. J. Qual. Methods* **2006**, *5*, 80–92. [[CrossRef](#)]
33. Borsci, S.; Schmettow, M.; Malizia, A.; Chamberlain, A.; van der Velde, F. A confirmatory factorial analysis of the Chatbot Usability Scale: A multilanguage validation. *Pers. Ubiquitous Comput.* **2022**, *27*, 317–330. [[CrossRef](#)]
34. Amershi, S.; Weld, D.; Vorvoreanu, M.; Founrey, A.; Nushi, B.; Collisson, P.; Suh, J.; Iqbal, S.; Bennett, P.N.; Inkpen, K.; et al. Guidelines for Human-AI Interaction. In Proceedings of the CHI '19: 2019 CHI Conference on Human Factors in Computing Systems, Glasgow Scotland, UK, 4–9 May 2019 ; pp. 1–13. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.