Hey Dashboard!: Supporting Voice, Text, and Pointing Modalities in Dashboard Onboarding

Vaishali Dhanoa [1,2] **Gabriela Molina León**[1] **Eve Hoggan**[1] **Eduard Gröller**[2]

Marc Streit[3] **Niklas Elmqvist** [1]

[1] Aarhus University [2] TU Wien [3] Johannes Kepler University Linz

Abstract

Visualization dashboards are regularly used for data exploration and analysis, but their complex interactions and interlinked views often require time-consuming onboarding sessions from dashboard authors. Preparing these onboarding materials is labor-intensive and requires manual updates when dashboards change. Recent advances in multimodal interaction powered by large language models (LLMs) provide ways to support self-guided onboarding. We present DIANA (Dashboard Interactive Assistant for Navigation and Analysis), a multimodal dashboard assistant that helps users for navigation and guided analysis through chat, audio, and mouse-based interactions. Users can choose any interaction modality or a combination of them to onboard themselves on the dashboard. Each modality highlights relevant dashboard features to support user orientation. Unlike typical LLM systems that rely solely on text-based chat, DIANA combines multiple modalities to provide explanations directly in the dashboard interface. We conducted a qualitative user study to understand the use of different modalities for different types of onboarding tasks and their complexities.

1 Introduction

Visualization dashboards have become essential tools for data exploration and decision-making across organizations [33]. Modern dashboards integrate multiple linked views, complex filtering mechanisms, and sophisticated interaction techniques to support analytical workflows [4]. However, this complexity creates a significant barrier: users often struggle to understand how to navigate and effectively use these dashboards without guidance [44, 7, 10]. Dashboard authors accordingly invest substantial time creating *onboarding* materials and conducting training sessions [12, 40], yet these resources quickly become outdated as dashboards evolve [13]. Existing automated onboarding solutions often fail to match the effectiveness of human experts who can provide contextual explanations, adapt to user questions, and demonstrate features using the actual data and visualizations present in the dashboard.

We present Diana (Dashboard Interactive Assistant for Navigation and Analysis), a multimodal dashboard assistant that replicates the guidance provided by human onboarding experts. Unlike text-only chatbots or onboarding tutorials (static or dynamic [13]), Diana combines multiple interaction modalities—voice input and output, mouse, and keyboard-based interaction—to provide *situated explanations* [9] directly within the dashboard interface. The system employs a contextual radial menu that organizes dashboard features into logical structures, supports pointing gestures and lasso selection for highlighting specific charts or data points, includes a text-based natural language

*Correspondence: dhanoa@cs.au.dk.

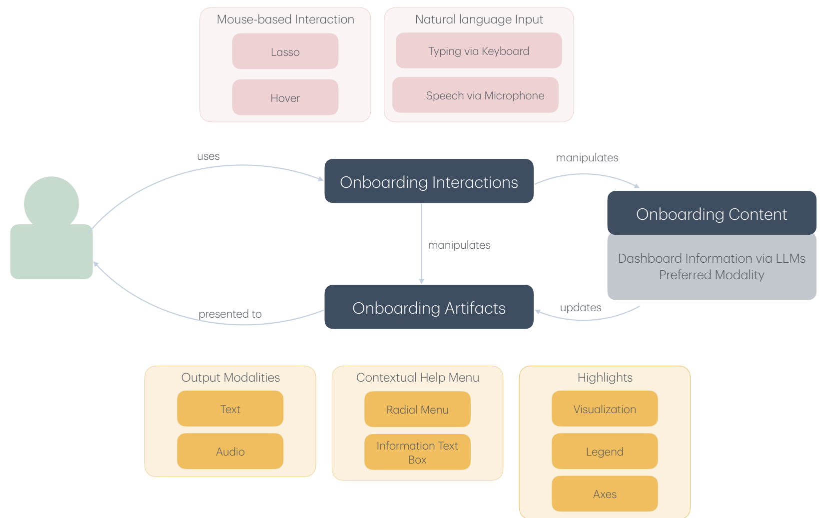

Figure 1: **Multimodal interaction for dashboard onboarding using LLMs.** Using the process model for dashboard onboarding proposed by Dhanoa et al. [12], we extend it with multimodal interactions and features, powered by LLMs for onboarding a user.

interface, and integrates speech recognition and synthesis for natural conversation. Diana delivers contextual help using the actual tasks, visualizations, and datasets from each specific dashboard, enabling users to learn through direct manipulation and exploration of real content.

We implemented Diana as a web-based system using a GPT-powered agent and integrated with Microsoft Power BI. The system supports any Power BI dashboard with appropriate metadata annotations. It uses OpenAI's Realtime API [2] for speech processing and provides real-time visual feedback through dashboard highlighting and overlays. Using the prototype, we conducted an in-depth qualitative user study with novices and expert dashboard users across hour-long sessions involving detailed tasks on varied, unfamiliar datasets and dashboards. The study design incorporated insights from three pilot studies to refine the experimental protocol. Results demonstrate that participants rapidly grasped Diana's capabilities and consistently preferred multimodal interactions over single-modality approaches for onboarding tasks. Visual highlighting combined with voice-based questions and responses emerged as the dominant interaction pattern. All experts expressed strong interest in using Diana for their professional dashboard work.

The contributions of this work are threefold: (1) the concept of a multimodal dashboard onboarding assistant that combines natural language, visual highlighting, and gestural interaction to replicate human-expert guidance; (2) the implementation of Diana, demonstrating how large language models can be integrated with dashboard environments to provide contextual, situated assistance; and (3) empirical findings from a qualitative user study revealing how novice and expert users employ different modalities for various onboarding tasks and their preferences for multimodal over unimodal interaction patterns.

2 Related Work

Here we review prior work on visualization dashboards, onboarding, multimodal interaction for data visualizations, and automatic approaches for dashboard generation.

2.1 Visualization Dashboards

Visualization dashboards have become ubiquitous tools for data exploration and decision-making across diverse domains, yet their complexity often creates significant barriers for users. Few's book on visualization dashboards [18] outlined the many fine points and varied purposes of dashboard design in industry. Sarikaya et al. [34] provided a comprehensive survey of what the scientific visualization community means by "dashboards," revealing their diverse applications and the varied expectations users bring to these systems. Their work highlighted the fundamental challenge: dashboards must balance powerful functionality with usability.

Multiple studies have documented the struggles users face when working with dashboard systems. Tory et al. [43] examined how users attempt to find their "data voice" when working with dashboards, documenting persistent difficulties in navigation, interpretation, and effective use of interactive features. Their findings reveal that even well-designed dashboards can overwhelm users who lack sufficient training. Similarly, Walchshofer et al. [44] observed the socio-technical challenges that arise when organizations transition to new dashboarding systems, noting that successful adoption depends heavily on both technical design and organizational support strategies.

The visualization community has responded to these challenges by developing design frameworks and best practices. Bach et al. [4] identified recurring dashboard design patterns, providing a vocabulary for discussing effective dashboard layouts and interaction paradigms. Lin et al. [23] developed DMiner, which provides design recommendations for dashboards based on analysis of a large dataset of existing dashboards. Their system helps authors create more effective dashboards by suggesting layouts, chart types, and design patterns that have proven successful in similar contexts, representing a shift toward data-driven dashboard design that could reduce user confusion from the outset.

Despite these advances in design guidance, the challenge remains: complex dashboards with multiple coordinated views and rich interactions require significant user investment to master. This persistent gap between dashboard capabilities and user proficiency motivates the need for more sophisticated onboarding and assistance mechanisms that can help users navigate these powerful but complex systems.

2.2 Onboarding for Dashboards

Given the popularity of visualization dashboards, *visualization onboarding* has recently emerged as a research area, though solutions remain fragmented across different contexts and applications. Stoiber et al. [40] first characterized the onboarding space for individual visualizations, establishing foundational concepts and identifying the need for systematic approaches to help users learn visualization systems. Their work highlighted various onboarding methods ranging from simple tooltips to more sophisticated guided experiences.

Building on these foundational concepts, Dhanoa et al. [12] proposed a systematic process model for dashboard onboarding, identifying key stages that users go through when learning new dashboard systems. This work provided a framework for understanding onboarding as a structured process rather than ad-hoc training.

Recent advances have focused on automating onboarding creation using tours. Dhanoa et al. [13] introduced D-Tour, a system for semi-automatically generating interactive guided tours for visualization dashboards. Their approach reduces the manual effort required from dashboard authors while maintaining the ability to create engaging onboarding experiences. Similarly, Hoque and Sultanum [19] developed DashGuide, which enables the creation of interactive dashboard tours with minimal author input while allowing for manual refinement when needed.

The comparative evaluation of different onboarding approaches has revealed important insights about their effectiveness. Stoiber et al. [41] systematically compared various onboarding methods, finding that interactive approaches generally outperform static documentation, though the optimal approach depends on user characteristics and dashboard complexity.

Despite these advances, most existing onboarding solutions remain primarily visual and static, relying on annotations [15], tooltips, or guided tours that users follow passively. Chundury et al. [9] represent a notable exception with their contextual, in-situ help features for visual data interfaces, though their approach focuses on help rather than comprehensive onboarding. This gap between the interactive nature of modern dashboards and the static nature of most onboarding solutions motivates the need for more dynamic, responsive onboarding systems that can adapt to user behavior and provide assistance through multiple interaction modalities. In particular, current automated onboarding methods are still far inferior to the gold standard of onboarding: using a human instructor [].

2.3 Multimodal Interaction with Data Visualizations

The visualization community has long recognized the limitations of traditional mouse and keyboard interaction for complex data exploration tasks. Elmqvist et al. [16] articulated the concept of *fluid interaction* for information visualization, arguing that effective visualization systems should support seamless transitions between different interaction modes and reduce cognitive barriers between user intent and system response. This foundational work established fluidity as a key principle and identified the utility of multiple modalities for designing more natural visualization interfaces.

Lee et al. [22] provided a comprehensive call to action, encouraging the visualization community to expand beyond mouse and keyboard to support more natural interactions. Their survey of alternative interaction modalities highlighted the potential for touch, speech, gesture, and other modalities to make visualization more accessible and expressive. This work catalyzed subsequent research into multimodal visualization interfaces.

Early explorations of touch-based visualization interaction demonstrated the potential of direct manipulation using modalities beyond the mouse and keyboard. Baur et al. [6] introduced TouchWave, showing how kinetic multi-touch gestures could enable fluid manipulation of hierarchical stacked graphs. Their work revealed how touch affordances could support more intuitive interactions with complex visualization structures. Building on these foundations, Nielsen et al. [29] studied the use of touch and pen scribbling to support multidimensional visualization queries. Badam et al. [5] analyzed the affordances of different input modalities—touch, speech, proxemics, gestures, gaze, and wearables—for visual data exploration, providing guidelines for combining modalities to leverage their complementary strengths. Thompson et al. [42] developed Tangraphe, which employed single-hand, multi-touch gestures for network visualization exploration, demonstrating how naturalistic gestures could replace conventional pointer-based interactions.

The integration of speech with visualization interaction has proven particularly promising for complex analytical tasks. Srinivasan and Stasko [37] examined what users naturally want to express through natural language when analyzing data, providing insights into the linguistic patterns that visualization systems should support. They later developed Orko [38], one of the first systems to combine voice and direct manipulation for network analysis, showing how speech could complement visual interaction for complex exploration tasks. These systems demonstrated that users could express analytical intents through conversational interfaces, though challenges remained in handling ambiguous queries and maintaining context across extended analysis sessions. Building on these foundations, Srinivasan and Setlur [36] developed BOLT, a natural language interface specifically designed for dashboard authoring, showing how conversational interaction could extend beyond data exploration to visualization creation itself.

Empirical studies have provided crucial insights into user preferences and performance with different modality combinations. Saktheeswaran et al. [32] found that participants preferred combining multimodal over unimodal input when exploring network visualizations on large vertical displays, with touch and speech proving to be a particularly effective combination. Their findings challenged assumptions that users would prefer single-modality interaction for simplicity. In large display contexts, León et al. [28] found that participants preferred speech commands for 10 of 15 exploration tasks over touch, pen, and mid-air gestures, suggesting that speech excels particularly in scenarios where physical interaction is challenging.

Recent work has focused on designing consistent multimodal experiences across different devices. Jakobsen et al. [20] explored how user position and movement relative to large displays could inform interaction design, showing how proxemic cues could enhance information visualization experiences through implicit spatial input. Srinivasan et al. [39] developed InChorus, which provides consistent multimodal interactions across tablet devices, demonstrating how design principles can maintain usability while adapting to different form factors.

2.4 Automatic Generation of Dashboards

The complexity of creating effective dashboards has motivated significant research into automated generation approaches. Early systems such as VizDeck [21] introduced the concept of self-organizing dashboards that automatically recommend visualizations based on statistical properties of data, using a card game metaphor to help users organize visualizations into interactive dashboards with minimal programming effort.

Deep learning approaches have shown promise for dashboard recommendation and generation. Wu et al. [46] developed MultiVision, which uses deep learning models to assist in analytical dashboard design by learning from provenance data and authoring logs. Their mixed-initiative system allows users to provide optional input while the model recommends data column selections and multiple chart combinations. Similarly, Deng et al. [11] proposed DashBot, which uses deep reinforcement learning to generate insight-driven dashboards by constructing training environments based on visualization knowledge.

Intent-based recommendation systems have emerged to support dashboard composition workflows. Elshehaly et al. [17] developed QualDash for healthcare quality improvement, introducing a metric card metaphor that serves as a building block for generating adaptable dashboards across different hospital units. Pandey et al. [31] introduced MEDLEY, which recommends dashboard collections based on analytical intents such as measure analysis, change analysis, and distribution analysis. Users can specify intents explicitly or implicitly through data attribute selection, with the system providing collections of logically grouped views and filtering widgets.

The proliferation of large language models has enabled new approaches to dashboard creation through natural language interaction. Dibia [14] developed LIDA, which uses LLMs for automatic generation of grammar-agnostic visualizations and infographics, demonstrating how conversational interfaces can simplify visualization creation. Building on this foundation, Lisnic et al. [24] developed a LLM-powered system specifically for creating text content within dashboards, showing how artificial intelligence (AI) can assist with narrative elements that provide context and explanation.

Recent work has integrated LLMs into broader visual analytics workflows. Chen et al. [8] developed InterChat, which combines direct manipulation with natural language interaction for visual analytics, including LLM-based prompting capabilities. Zhao et al. [48] proposed LightVA, a framework that uses LLM agents for task planning and execution in visual analytics, employing a recursive process with planner, executor, and controller components. Wen et al. [45] explored multimodal prompting with LLMs for visualization authoring, investigating how different input modalities can enhance the creation process beyond traditional text-based interaction.

While these automated generation approaches have made dashboard creation more accessible, most focus on the authoring process rather than helping users understand and navigate existing dashboards. The challenge of onboarding users to complex, pre-existing dashboard systems—particularly through multimodal interaction—remains largely unaddressed by current automatic generation methods.

3 Diana: Dashboard Interactive Assistant for Navigation and Analysis

Existing onboarding methods for dashboards, such as documentation, training, videos, and meetings, often fail to support users in situ. They are static, linear, and cannot adapt to individual preferences or questions that may arise during the onboarding process. Additionally, the user needs to constantly switch programs, manually test interactions while cross-referencing instructions, which can lead to frustration with onboarding help. Self-designed tutorials, such as step-by-step interfaces, also face the challenge of being ignored by many users. To address the balance between autonomy and agency, we designed Diana (Dashboard Interactive Assistant for Navigation and Analysis), a multimodal onboarding assistant powered by Large Language Models.

Diana is designed not to replace domain expertise, but to bridge the gap between static onboarding and in-situ analysis. Its role is to help users navigate dashboards, understand visualizations, and learn how to perform analysis through guided, multimodal interaction. Users can interact with Diana via three interaction modalities: keyboard (chat), mouse (lasso selection), and voice (speech).

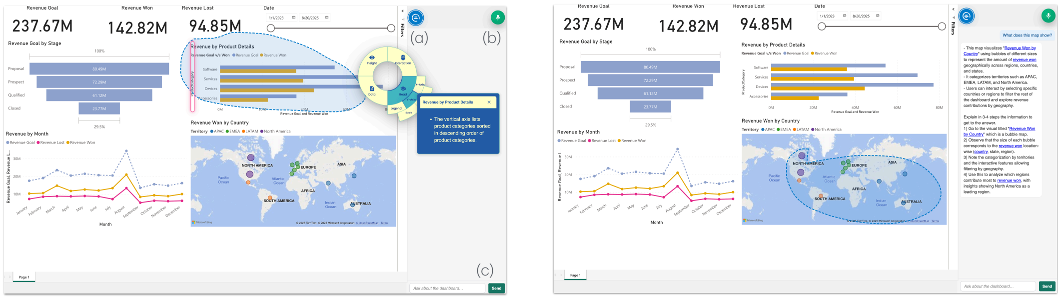

Figure 2: **DIANA User Interface.** The left side shows a visualization dashboard with Diana activated: (a) represents the lasso selection button, (b) push-to-talk for speech-based interaction, and (c) chat-based interaction. Currently, the lasso selection is activated and the y-axis of the visualization is highlighted, along with the radial menu which is opened for the y-axis information. The right side shows multimodal interaction with mouse-based lasso selection and chat-based interaction that infers the context of the conversation based on the selected visual.

Based on the chosen modality, Diana provides feedback through text, audio, a contextual menu, and visual highlights to orient users within the dashboard.

Our design goals (DG) combine lessons from the state of the art in digital assistants (e.g., Microsoft Co-Pilot), common modalities already available in dashboards, and practices of human onboarders:

DG1 Build on familiar environments: create an onboarding system that can be used on a familiar dashboard platform (such as Microsoft Power BI) to increase the likelihood of adoption.

DG2 Support multimodal interaction: allow users to interact via keyboard, mouse, and voice, individually or in combination, to accommodate different preferences and contexts.

DG3 Provide human-like onboarding support: simulate the role of an expert by answering questions, guiding navigation, and scaffolding analytical steps.

DG4 Facilitate orientation in complex dashboards: employ features such as lasso selection, contextual help menu, and visual highlights to direct user attention and clarify relationships across views.

DG5 Balance guidance and autonomy: give users the freedom to request help when needed, rather than forcing a step-by-step walkthrough.

We designed Diana according to these goals that support the user's orientation and onboarding within the dashboard. It is implemented using a Microsoft Power BI dashboard embedded in a custom web application, thereby satisfying DG1. We distinguish between modalities, which define how users interact with Diana (keyboard, mouse, voice), and features, which define how the system provides feedback and orientation (contextual help menu, visual highlighting). We explain them below.

3.1 Mouse-based Interaction

The mouse is the primary input device in most dashboard environments, and users are generally familiar with interacting with visualizations through clicking. To extend this familiar mode of interaction for onboarding (DG2), we included two complementary features: lasso selection (or free-form selection) and hover functionality. Lasso selection enables users to mark a region of interest within the dashboard, mimicking the pointing gestures a human expert might use when directing attention to a particular element. Once a region is selected, the system captures both the highlighted area and the corresponding visualization. At the same time, the dashboard remains fully interactive for further exploration.

Hover functionality provides additional detail in two ways: (i) through contextual help options in the radial menu (see Section 3.4), and (ii) through embedded hyperlinks returned in responses to text-based queries (see Section 3.3). This way lasso serves as a way of anchoring questions to a region, while hover enriches that region with further information. In a typical onboarding scenario, users also point and ask questions, therefore, through mouse interaction, we provide human-like onboarding support (DG3) and also facilitate orientation by anchoring onboarding interactions to specific areas of the visualization (DG4).

3.2 Voice-based Interaction

With the growing support for multimodal interaction in LLM-powered assistants (e.g., Microsoft Co-Pilot), we included a voice modality to mimic the experience of conversing with a human onboarder (DG2 and DG3). Speech input allows users to issue queries in natural language and receive real-time guidance without relying on typing.

Users can enable this modality via a push-to-talk model, where the users can press and hold the mouse button while speaking, and upon release, Diana processes the input using the Realtime API [2] and immediately returns a spoken response (DG5). This design ensures that users maintain full control over when the system listens, avoiding interruptions during regular conversation. We initially experimented with a wake-word approach (*"Hey dashboard!"*), but found it prone to false activations, as the system occasionally responded to background discussions that were not intended for the assistant. We therefore selected the push-to-talk model as a more reliable and less intrusive interaction style. In addition to the audio output, Diana also highlights specific regions of the dashboard to direct user's attention (DG4).

3.3 Keyboard-based Interaction

The keyboard-based interaction complements the voice-based interaction by offering a written chat interface that supports deliberate, text-based interaction. Unlike spoken dialogue, written exchanges create a persistent trace of the conversation, allowing users to revisit instructions and follow them step by step (DG2 and DG3).

The chat component uses the OpenAI APIs to process user queries and provides responses in textual form. In addition to plain text, these responses may also include embedded hyperlinks and hovering over them highlights the corresponding regions on the dashboard (e.g., the visualization, its legends, or its axes) (DG4 and DG5).

3.4 Contextual Help Menu

To support onboarding in a deterministic and structured way, we designed a contextual radial menu that complements the natural language modalities (chat and speech), which rely heavily on LLMs for generating responses. While conversational interaction adapts to the user's phrasing and provides guidance step by step, the menu makes the full range of onboarding information visible, even if the user does not know what to ask.

We extended the component graph for dashboards described by Dhanoa et al. [12] for the design of our contextual help menu. As a visualization can contain many elements, such as data, legends, axes, and interaction options, that may require explanation, we adopted a bottom-up approach to assemble this information hierarchically. To present this information visually, we used a circular radial layout that progressively reveals detail across levels. It captures the complexity of the component graph while presenting it in a structured and navigable form.

The menu is organized into four top-level categories:

- **Read**, describing how to interpret the visual encoding (e.g., axes, labels, scales, legend order).
- **Data**, explaining underlying sources such as tables, columns, keys, and measures.
- **Interact**, making explicit both self-interactions (e.g., highlighting or comparing categories) and cross-visual interactions (e.g., filtering other charts).
- **Insight**, outlining descriptive and smart insights such as trends or potential drivers.

The menu appears automatically when a region of a chart is selected with the lasso (DG4). It opens at the relevant category and displays a pop-up information text box with a description of the selected region. Users can also hover over menu items, prompting the text box to display the corresponding explanation.

Although the schema is deterministic, we used an LLM to polish the narrative text associated with each menu entry. For instance, instead of a bare label such as *Currency Scale*, the menu might display *The X-axis represents values in USD on a continuous scale.* In this way, the LLM enriches the text with more natural explanations, while the structure and content remain grounded in the underlying visualization metadata. In addition to plain text, icons are used to enhance readability.

The menu dynamically adapts to the visualization type by pruning irrelevant categories. For instance, for a Key Performance Indicator (KPI) card, interaction options are hidden, as direct highlighting or filtering is not possible. Similarly, for a filter visual, the Insight category is suppressed, since filters do not directly encode data-driven findings.

Unlike the chat or speech modalities, which reveal information only when prompted, the menu lays out the full structure of what can be learned, encouraging discoverability and providing a complete description of the visualization. It packages information by category but requires the user to assemble the pieces into an understanding of the visualization.

3.5 Visual Highlighting

We designed the visual highlighting feature to complement all three modalities by directing user attention to specific dashboard elements, similar to how a human onboarder might point to a chart or annotate it (DG3). When triggered through a query, whether via lasso selection, voice command, or a written input, the system highlights the relevant visualization, its axes, or its legend. These highlights appear as rectangular pulsating bars that serve as visual cues, helping users orient themselves and understand where to focus (DG4).

During a lasso selection, for example, if the system detects a region such as a legend, the highlight is shown around the legend area. For voice-based interaction, highlights are synchronized with the spoken guidance, though slight rendering delays can occur. In keyboard-based interaction, highlights are triggered through embedded hyperlinks and appear when hovered.

3.6 Combining Modalities and Features

In real onboarding scenarios, users rarely rely on a single modality when seeking onboarding support. Therefore, we designed Diana to allow users to combine modalities and features, in parallel or sequentially, to reproduce the onboarding experience closer to that provided by a human onboarder (DG3). Our system enables users to lasso a region (mouse-based interaction) and simply use the voice-based interaction to ask a question for the selected area. The user can also view more detailed information in the contextual menu, while receiving audio onboarding. The use of visual highlights can prompt further exploration, prompting the user to open the contextual menu on that specific visualization or simply selecting that visualization and asking a chat-based query (keyboard-interaction).

These combinations allow Diana to have different kinds of support: the mouse-based anchoring to a selected region, the contextual menu for structured information, natural language support in the form of voice and keyboard-based interactions, and visual highlighting for orientation and directed attention. Altogether, they can create an adaptive onboarding experience in which onboarding can be tailored to both the task at hand and the user's preferred way of interacting.

Unlike typical LLM-based assistants that aim to deliver direct answers, Diana focuses on orientation and learning. The system is prompted with information about the structure of the dashboard (such as available visuals, axes, maps, and overview descriptions) but not with the underlying data values. As a result, Diana cannot return analytical results itself, except for basic descriptive queries (e.g., *what is the y-axis?*). This design choice ensures that the assistant focuses on orientation and guidance rather than replacing the user's own analytical work. We explain the working of the system in the next section.

4 Implementation

We implemented Diana using the Microsoft Power BI Embedded report [27], to integrate our solution within existing commercial dashboard tool, such as Microsoft Power BI. We obtain information about all the visualizations in the dashboard using Power BI's REST API [35]. We also apply simple heuristics to deduce the information not present within these APIs. Additionally, we use OpenAI chat completion models [1] and Realtime API [2] for chat and voice-based interactions, respectively.

The system itself is built using the Vue framework [3] for the front-end and Flask API for the Python backend. All calls to the Open AI APIs are routed through the backend. We also created prompts for chat and voice-based interactions; they are provided in the supplementary materials. These prompts are designed such that the system does not output answers, but rather define steps for getting to an answer. The reason for this is that the purpose of Diana is not to create yet another assistant that solves user queries, but that provides onboarding and navigation support so that users can remain responsible for the analytical tasks.

We implemented Diana using Microsoft Power BI Embedded [27] to integrate within a widely used commercial dashboard platform without building a custom dashboard stack. We obtain visualization information, such as titles, axes, data, and layout bounds via the Power BI Embed API for Typescript [25], and use the Power BI REST API [35] for embedding logistics (e.g., report loading and issuing logging in token). When certain attributes are unavailable, we apply lightweight heuristics to infer them from the rendered view.

For the natural language interaction, we use the OpenAI APIs: Chat Completions with gpt-4.1-mini for text [1] and the Realtime API [2] with gpt-4o-realtime-preview for voice I/O (with automatic speech transcription). All OpenAI calls are proxied through the Flask backend [30], which mints ephemeral Realtime session tokens so the browser never sees the primary API key.

Diana cannot query the underlying dataset nor retrieve off-screen values. Therefore, it cannot provide data-derived answers or numeric values. It only uses the metadata and content available from the embedded dashboard (e.g., visual titles, layout bounds, and author-provided descriptions). Diana has no authoring control over the dashboard, and therefore, it cannot apply filters, or interact with visuals. When a user request requires information that is not available from the visible dashboard context, the system emits a structured "NOT IN DATA" outcome and redirects the user to the appropriate navigation or exploration steps.

5 Study Design and Methodology

Our goal was to examine how users engage with DIANA (DG2 and DG3: multimodal interaction and human-like onboarding) and how it affects the balance between guidance and user autonomy (DG5). We therefore adopted a qualitative, task-based study design combining think-aloud protocols with semi-structured interviews. We conducted the study in two phases: (1) a formative study (n=4), to refine the task design and the set of supported modalities, and (2) main interviews (n=6) with participants spanning novice to experienced dashboard users.

During the main sessions, participants completed a series of dashboard-oriented tasks of varying difficulty using Diana. Participants were free to choose the interaction modality (or modalities) per task and could ask follow-up questions on tasks in more detail when needed. Comparing real-world onboarding scenarios, we also provided a digital standard user guide for the dashboard as an optional reference but participants were not required to use it.

5.1 Participants

We first did a formative study with four participants, out of which one was (A1) dashboard author who regularly onboards several colleagues and the other three (A2-A4) were domain experts and dashboard users with different level of visualization and dashboard expertise, as shown in Figure 3. We used the insights from the formative study to inform the task phrasing and overall improvement of the prototype for the main session. We do not include the formative study data in the main analysis.

Both the formative study and main session participants were recruited from different university departments, such as Business Intelligence, Economics, Finance & Controlling, and Computer Science. For the main session, the average age of the participants was 36.7 years (σ = 11.1) and their experience in their field of expertise ranged from 2 to 25 years. Out of the six main study participants, three were well-versed with dashboards (P2, P3, and P4), with P2 providing regular onboarding to many of his colleagues. The other three participants (P1, P5 and P6) were novices and had

Figure 3: **Participant demographics.** Demographics of the formative study participants (A1-A4) and main study participants (P1-P6). Gender is encoded as male (M) and female (F). Visualization and dashboard experience is measured from novice to expert. AI interface experiences is noted for chat and voice interfaces.

little to no experience with visualization dashboards. Each interview was conducted in person and individually.

We also asked all these participants about their familiarity of AI tools and their experience with different modalities. These participants were chosen for two main reasons: (i) to understand how varied level of expertise (novice-expert) leads to differences in self-onboarding practices, and (ii) how does their experience with AI and multimodalities inform their preferences during the onboarding process.

5.2 Dataset and Task

We used an example dataset from Microsoft Power BI [26] to create a custom visualization dashboard. The dashboard consisted of three key performance indicators (KPIs), one date filter, a funnel chart, a line chart, a clustered bar chart, and a map visualization. All visualizations, except the KPIs, were interactive. In addition to highlighting and filtering, a few visuals supported drill-down functionality, allowing users to explore information in greater depth.

We leveraged this interactivity when designing the task set, which spanned three dimensions: task type, cardinality, and level of difficulty.

- **Task type:** Tasks were categorized into three groups: *information lookup*, *exploratory*, and *interpretive*. Information lookup tasks ranged from simple retrievals (e.g., reading a KPI value) to slightly more involved actions (e.g., hovering over a visual to reveal details). Exploratory tasks required participants to interact with one or more visuals to uncover the relevant information. Interpretive tasks went further, requiring participants to make sense of patterns or relationships that were not explicitly presented.

- **Cardinality:** This dimension captured the number of views a task required. As the dashboard consisted of multiple coordinated views, some tasks could only be solved by interacting with more than one visualization.
- **Level of difficulty:** We assigned difficulty levels (easy, medium, hard) based on our prior experience with dashboard onboarding and the complexity of the interactions required to solve the task, also informed by the formative study results.

For instance, an easy look-up task involving only a single view could be *How do I figure out the scaling of the x-axis of the bar chart?*. Meanwhile, a difficult task involving multiple views could be *How can I figure out the revenue goal for Australia in the Services subcategory for the Proposal stage?*. We designed ten such tasks based on our own onboarding experience and in discussion with the formative study participant A1. Figure 4 shows the task types and their difficulty levels. The task questions are provided in the supplementary material.

5.3 Study Procedure

Each session began with an introduction to the study goals and a request for informed consent regarding audio and screen recording. After receiving their consent, we asked participants to fill out a pre-questionnaire, containing a set of preliminary questions about their prior experience with visualizations, dashboards, and AI interfaces (summarized in Figure 3).

Following this, participants completed a short training session that lasted approximately five minutes on average. During the training, we introduced Diana and asked participants to try each interaction modality at least once, as well as to experiment with combining modalities. This was intended to ensure that participants felt confident in using Diana before starting the tasks with the dashboard. In addition to the modalities provided by the system, a dashboard user guide (as a digital document) was available as an optional resource throughout the study.

The main task phase of the study lasted around 25 minutes. Participants were free to choose whichever modalities they preferred for each task and could switch between them at any time. They could ask the system as many questions as they wanted until they were satisfied with an answer, or solve the tasks on their own without assistance. If a task could not be solved, participants were allowed to skip it and continue to the next one. After completing all tasks, participants were interviewed about their onboarding experience, comparisons to their usual onboarding practices, and their preferred modalities.

5.4 Data Collection and Analysis

To capture participants' backgrounds and experiences, we collected data using the pre-questionnaire as well as follow-up interviews after the tasks focusing on onboarding practices, modality preferences, and overall reflections on the experience.

For each session, we also recorded audio and screen activity, capturing both the dashboard interactions and the participants' think-aloud comments. Responses to the pre-questionnaire and the follow-up interview were also documented. The recordings were later analyzed qualitatively to identify task-solving strategies and modality choices.

For each task we recorded whether it was solved, the modality and any switches between modalities, the Diana features that participants engaged with, and their rationales for switching. In their post task interview, we collected qualitative data on their preferred modality, their perceived guidance and autonomy and their experience compared to their previous onboarding experiences. The time taken to complete a task was not taken into account as it depends heavily on how participants understood the task phrasing and their familiarity with various modalities.

6 Results

In the following, we present the findings from the four formative study sessions and the six main study sessions.

Figure 4: **Modalities and features used during onboarding.** Filled boxes show the use of a modality or a feature. Empty boxes indicate that it was not used. Boxes are colored by participants ID: P1, P2, P3, P4, P5, and P6.

6.1 Formative Study Interviews

We did a formative study with four participants (A1-A4) to refine both the design and implementation of Diana and the experimental procedure. The list of the participants included a Power BI specialist (A1), a regular Power BI user (A2), a dashboard author and occasional user (A3), and a daily Power BI user (A4). In the following, we describe their experience with dashboards, onboarding, and their impressions about the system.

A1 regularly created onboarding materials for dashboards, typically using videos and bookmarks in Power BI. Their approach involved taking screenshots of dashboards, adding markers and pointers, and displaying them as a translucent overlay on the first page, which users can dismiss after viewing. They noted that this process is time-consuming and requires frequent updates as dashboards evolve. A1 also emphasized that dashboard features often contain subtle behaviors that are difficult to onboard effectively. For example, drill-down interactions usually confused users because the view changes immediately, and users may not recognize the update. During task performance, A1 suggested that visuals should be highlighted directly and that the AI should provide more specific guidance, while also conveying the purpose of the dashboard.

A2 relied heavily on dashboards in Microsoft Fabric and typically used official documentation for onboarding. If documentation was insufficient, they turned to community forums, and only as a last resort, to AI systems. Having used voice interfaces for more than seven years, A2 expressed disappointment that these technologies have not met their early promises, preferring instead what they described as "ground truth" onboarding methods. They also noted that seeking help was particularly difficult for novices who may not yet know what to ask. A2 valued contextual menus that provided embedded information about a visualization. Based on this feedback, we included an additional option for onboarding: a user guide containing annotated images and instructions for how to read, interact with, and interpret the dashboard.

A3 used Power BI dashboards infrequently. When joining her institution, she received a single 45-minute onboarding session covering relevant dashboards. For support, she typically experimented on her own or asked a mentor. During the tasks, A3 expected contextual menus to provide richer explanations. Although she did not normally use voice interfaces, in the study, she quickly shifted to voice to ask clarification questions—about abbreviations, interaction methods, and task instructions. She appreciated the multimodality of the onboarding, as she valued having options. While she preferred chatting via text over talking out loud, she emphasized the importance of having a choice. She also combined mouse-based selection with writing to find answers.

A4 used dashboards daily and, like A3, had received only a single onboarding session. When encountering issues, she usually consults her mentor. Although she used Microsoft Co-Pilot for other tasks, she had not used it with dashboards. During the tasks, she favored chat for its step-by-step guidance, but also experimented with speech interaction. She appreciated that the Open AI's Realtime API recognized her accent and changed to a different language accordingly, but suggested improvements such as synchronized animations, slower speech, and a transcription option. The user guide, by contrast, felt burdensome. A4 preferred using text and speech together: the chat allowed her to get written instructions while the voice output guided and confirmed her actions in real-time. She saw this combination as a way to save time for both herself and her mentor, reducing the need to ask experts for help.

Our formative study highlighted four changes: (i) Visual highlighting is essential to direct attention and orient users within the dashboard, (ii) prompts must be instructional so that they can teach a user how to perfrom concrete actions (e.g., drill-down on a visualization), (iii) the information text box that appears with the contextual menu should have narrative style information with icons to enhance readability, and (iv) a ground-truth fallback is needed, such a user guide so that the users can use it, if they prefer documentations over Diana's support.

6.2 Main Study

In the main study, we had six participants whose performance varied systematically with the task type, number of views required, and the difficulty level. They frequently switched modalities when tasks became complex, combining different modalities along with features such as lasso selections, radial menus, and visual highlights.

6.2.1 Easy Tasks

Most participants solved easy lookup or hover tasks independently, relying little on Diana. Onboarding was mainly used for confirmation rather than discovery. For instance, P2 and P3 used speech interaction and employed the voice and visual highlights feedback to double-check their answers, while P5 used the chat interface to clarify how to solve one task. As P1 explained, she was *"used to figuring things out on [her] own,"* describing the multimodal help as *"brand new... but pretty cool to have these options."* The learning effect was evident by the ninth task, which every participant solved without any onboarding, regardless of prior expertise.

6.2.2 Medium Tasks

Medium-level tasks typically required interacting with two to three views and triggered frequent back-and-forth interactions with Diana and switching heavily between modalities. Participants P1 and P5 used the keyboard to chat with Diana. The text-based response, coupled with embedded visual highlighting of the relevant region, helped participants understand where to look at and execute the onboarding steps to solve the tasks (DG4). P1 remarked that *"it is pretty clever that it doesn't tell you the answer, because sometimes you write [or type] and you get the answer, and you don't know how you actually got there! That's pretty cool".*

P2, P3, P4, and P6 used speech interaction, sometimes in combination with the mouse (DG2). The audio and visual feedback helped them orient themselves on the dashboard and find the right steps to get to their answers. P3 and P6 asked several follow-up questions to Diana, in order to interpret the question, engaging with at least two modalities at a time (DG2 and DG3). P2 also switched modalities, from keyboard to speech, to rephrase the question in a more natural-sounding way. P1 and P2 also used the mouse to interact with the radial menu to gain more information on the onboarding content. This illustrates how participants used Diana not just to retrieve answers, but also to scaffold their own orientation and reasoning in the dashboard.

6.2.3 Hard Tasks

Hard exploratory and interpretive tasks were the most demanding, often involving multiple steps and rephrasing. Participants engaged in multi-modality sequences (e.g., mouse → voice → keyboard). P2, P3, P4, and P6 used the mouse and voice modality and rephrased the questions in their own way to interact with Diana. P2, P3, and P6 asked several follow-up questions using the voice modality, indicating that as tasks became more complex, the need for human-like onboarding became more prevelant. P1 and P5 switched between keyboard and other modalities. Despite the use of all the modalities, P1 could not solve one of the hard tasks, indicating that the instruction prompt to Diana could have been improved. This was the only unsolvable task during the entire user study session across all participants. P2 also used the mouse modality, utilizing the radial menu feature that helped him confirm his answer to one of the tasks.

6.2.4 Preferred Modalities and Features

As the participants performed the user study, their choice of preferred modality became clear. Despite varying levels of prior experience, all participants experimented with speech interaction. Even those with no prior exposure (P1, P2, P6) adopted it during the study, with P2 initially hesitant but later describing it as *"comfortable, almost not robotish."* Voice was valued for its speed and natural phrasing, while chat was preferred for step-by-step guidance (P1 and P5). Mouse interactions, such as lasso selection, were consistently used to define a focus before turning to another modality (P3 and P4). Across modalities, the visual highlight feature was universally appreciated to orient user attention.

6.2.5 Underused Modalities and Features

The radial menu was the least used feature. While P2 found it helpful for confirmation, most others described it as unintuitive, overly detailed, or unfamiliar: *"I didn't know what specifically to use it for [...] the surface things were good enough"* (P4). The user guide was largely ignored, with P1 and P3 admitting that they forgot it existed. P6 said she would have preferred to read it if time allowed. This indicates that static documentation can be easily overshadowed by interactive support.

6.3 Reflection on the System

In the post-task interview, participants reflected on three main themes: their choice of modality, their experience with the onboarding, and the comparison to previous onboarding experiences.

6.3.1 The Value of Multimodality.

Several participants emphasized that having multiple options was useful, as it allowed them to adapt depending on time, task, or personal style. P1 summarized this flexibility: *"If you are in a hurry, you can just talk to it and it goes faster. If you have time to sit down and think, then you can use the chat and the radial menu... But I would probably use writing at most, because I'm old-fashioned."*

Similarly, P2 appreciated having "fallbacks" between modalities. P3 and P4 also found the use of multiple modalities, especially mouse and voice quite helpful, with P4 saying that she *"definitely liked the combination of selecting [lasso] and speech."*

6.3.2 Strengths of Voice-based Interaction and Visual Highlights Feature.

Speech interaction was well received, even among those without prior experience. P2 remarked that *"it was very comfortable voice, I could hear that it was mechanical, but it was not that mechanical."* He also appreciated the real-time response, and when paired with the visual highlighting feedback, it helped him see what Diana was saying in response to his question (DG4). As P2 was an expert, he could verify the correctness of the answer provided by Diana. As the answer was phrased naturally and accurately, he said *"and when it did that [gave him the right steps to the answer] it had my trust, that I could actually ask the voice interaction a question."* P3 described the voice modality and the response from Diana informative and that he *"really [had] fun interacting with her,"* while

P4 appreciated that speech *"doesn't stop when you are doing something else on the dashboard."* The mouse modality, in the form of lasso selection, was also valued when combined with voice, as the combination helped participants anchor their queries to specific visuals.

6.3.3 Limitations of Chat and Radial Menu.

Although useful for step-by-step guidance, the chat via keyboard was sometimes described as slow or effortful (*"chatting took too long to write compared to just speaking,"* P2). The radial menu was the least favored feature: while P2 found it helpful for confirming a difficult task, others found it overwhelming or unfamiliar. P4 explained, *"It had more deep details that I normally wouldn't look for... Unfamiliarity with it made me use it less."*

6.3.4 Comparison to Prior Onboarding.

Compared to step-by-step onboarding methods, Diana was described as more flexible and empowering. P4 compared it to video game tutorials where you are forced to proceed linearly: *"You have to explore one step at a time... Here you have the freedom to onboard yourself the way you want to."* P1 echoed this, preferring to ask for help when needed rather than *"click all over"* through a fixed tutorial (DG5). For P2, who regularly onboards colleagues via videos and one-on-one sessions, Diana represented a potential time saver. Although he values personal interaction, he saw the system as a *"virtual project finance controller"* (DG3) that could free up some of his workload.

7 Discussion

We now discuss the implications of our findings in the context of dashboard onboarding and the use of large language models for onboarding.

7.1 Speech Interaction as the Human Onboarder Assistant

In the main study, participants frequently preferred speech over mouse or keyboard-based interactions for onboarding help and follow-up questions. This result matches the findings of León et al. [28] on modality preferences for exploration tasks. This is also in line with the findings from Zhan et al. [47] who studied trust in healthcare voice assistants and suggested that perceived usefulness of a system is a significant factor that could explain the trust a user has in the system.

Rather than replacing human onboarders, Diana assists and augments their capabilities by handling routine questions such as *how do I...?* that are time-consuming to document and maintain. In the context of onboarding, the result also aligns with one of our initial goals of providing human-like support (**DG3**). The findings suggest that participants appreciated real-time responses and being less dependent on asking a human expert for help.

In our formative study, we also came across a surprising result: the Realtime API [2] detected that the participant was not a native English speaker and switched to another language that the person understood. While it was an unexpected behavior, the person was positively surprised about the detection. Furthermore, other participants wished for natural language interaction in their native language. While this is an option with the current large language models, it may go beyond the capabilities of a human onboarder in most cases.

7.2 Multimodal Interaction for Onboarding

The different preferences and interaction patterns we found suggest that multimodality is a suitable approach for onboarding. While some participants had a clear preference for one modality, others, such as keyboard-based interaction for chat or voice-based interaction, others preferred combining two or more. This was visible when the tasks were difficult.

The less use of the contextual menu could also suggest that sophisticated visual encodings for help are not necessary for onboarding, as they were perceived as advanced or quite detailed. But, they would be extremely helpful at later stages after users have been onboarded, to allow them to access more fine-grained information, as echoed by P1.

7.3 Using Large Language Models for Onboarding

While the natural interaction via voice and text was successful, it required a substantial effort at the technical level to minimize the number of hallucinations. The information sent to the LLMs was carefully designed based on the available REST APIs information and heuristics that ensure that information from the REST APIs is not wildly extrapolated. The prompts were fine tuned several times during the formative study process. We were able to significantly reduce the hallucinations related to both chat and voice based interactiions, but participants still encountered issues when they highlighted a specific section of the dashboard and then asked for guidance about it via natural language.

Additionally, as the Diana has no access to the data, it could be used in any domain with ease because it only has information about the high level aspects of the visualization. It can infer trends and drivers, however, it cannot give detailed data-level information that is usually critical in business-context.

7.4 Limitations

While we included a user guide in our study to compare with standard onboarding materials, we did not compare our approach with other onboarding methods, such as in-person sessions. Future work should compare our LLM-powered approach with in-person onboarding to identify the effects on performance and user experience.

8 Conclusion

We designed and developed Diana, a multimodal dashboard assistant that combines voice, text, and in-situ visual highlight and contextual help, to help users in dashboard onboarding. Through a qualitative user study, we found that participants leveraged all interaction modalities and preferred voice-based interaction over keyboard-based and mouse-based interaction. Our findings suggest that LLMs are useful for onboarding tasks when they are paired with visual affordances. It also encouraged users to ask for help without it getting into the way of their autonomy. We suggest that onboarding systems should be multimodal by design and offer a low-friction path to help users onboard themselves on dashboards.

Acknowledgments

This work was supported partly by Villum Investigator grant VL-54492 by Villum Fonden. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the funding agency.

References

[1] . Gpt-4.1 mini. https://platform.openai.com/docs/models/gpt-4.1-mini, 2023. Last accessed on 2025/09/09.

[2] . Realtime api. https://platform.openai.com/docs/guides/realtime, 2023. Last accessed on 2025/09/09.

[3] . Vue js. https://vuejs.org/, 2023. Last accessed on 2025/09/09.

[4] Benjamin Bach, Euan Freeman, Alfie Abdul-Rahman, Cagatay Turkay, Saiful Khan, Yulei Fan, and Min Chen. Dashboard design patterns. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):342–352, 2023. doi: 10.1109/TVCG.2022.3209448.

[5] Sriram Karthik Badam, Arjun Srinivasan, Niklas Elmqvist, and John Stasko. Affordances of input modalities for visual data exploration in immersive environments. In *Proceedings of the IEEE VIS Immersive Analytics Workshop*, Los Alamitos, CA, USA, 2017. IEEE Computer Society. URL https://karthikbadam.github.io/files/mixing-modalities.pdf.

[6] Dominikus Baur, Bongshin Lee, and Sheelagh Carpendale. TouchWave: kinetic multi-touch manipulation for hierarchical stacked graphs. In *Proceedings of the ACM Conference on Interactive Tabletops and Surfaces*, pages 255–264, New York, NY, USA, 2012. ACM. doi: 10.1145/2396636.2396675.

[7] Davide Ceneda, Theresia Gschwandtner, Thorsten May, Silvia Miksch, Hans-Jorg Schulz, Marc Streit, and Christian Tominski. Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):111–120, January 2017. doi: 10.1109/TVCG.2016.2598468.

[8] Juntong Chen, Jiang Wu, Jiajing Guo, Vikram Mohanty, Xueming Li, Jorge Piazentin Ono, Wenbin He, Liu Ren, and Dongyu Liu. InterChat: Enhancing generative visual analytics using multimodal interactions. *Computer Graphics Forum*, 44(3), 2025. doi: 10.1111/CGF.70112.

[9] Pramod Chundury, Mehmet Adil Yalçin, Jonathan Crabtree, Anup Mahurkar, Lisa M. Shulman, and Niklas Elmqvist. Contextual in situ help for visual data interfaces. *Information Visualization*, 22(1):69–84, 2023. doi: 10.1177/14738716221120064.

[10] Christopher Collins, Natalia Andrienko, Tobias Schreck, Jing Yang, Jaegul Choo, Ulrich Engelke, Amit Jena, and Tim Dwyer. Guidance in the human–machine analytics process. *Visual Informatics*, 2(3):166–180, September 2018. doi: 10.1016/j.visinf.2018.09.003.

[11] Dazhen Deng, Aoyu Wu, Huamin Qu, and Yingcai Wu. DashBot: Insight-driven dashboard generation based on deep reinforcement learning. *IEEE Transactions on Visualization and Computer Graphics*, 29(1):690–700, 2023. doi: 10.1109/TVCG.2022.3209468.

[12] Vaishali Dhanoa, C. Walchshofer, Andreas Hinterreiter, Holger Stitz, Eduard Groeller, and Marc Streit. A process model for dashboard onboarding. *Computer Graphics Forum*, 41: 501–513, 2022. doi: 10.1111/cgf.14558.

[13] Vaishali Dhanoa, Andreas P. Hinterreiter, Vanessa Fediuk, Niklas Elmqvist, Eduard Gröller, and Marc Streit. D-Tour: Semi-automatic generation of interactive guided tours for visualization dashboard onboarding. *IEEE Transactions on Visualization and Computer Graphics*, 31 (1):721–731, 2025. doi: 10.1109/TVCG.2024.3456347.

[14] Victor Dibia. LIDA: A tool for automatic generation of grammar-agnostic visualizations and infographics using large language models. *CoRR*, abs/2303.02927, 2023. doi: 10.48550/ARXIV.2303.02927.

[15] Micheline Elias and Anastasia Bezerianos. Annotating BI visualization dashboards: needs and challenges. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pages 1641–1650, New York, NY, USA, 2012. ACM. doi: 10.1145/2207676.2208288.

[16] Niklas Elmqvist, Andrew Vande Moere, Hans-Christian Jetter, Daniel Cernea, Harald Reiterer, and T. J. Jankun-Kelly. Fluid interaction for information visualization. *Information Visualization*, 10(4):327–340, 2011. doi: 10.1177/1473871611413180.

[17] Mai Elshehaly, Rebecca Randell, Matthew Brehmer, Lynn McVey, Natasha Alvarado, Chris P. Gale, and Roy A. Ruddle. QualDash: Adaptable generation of visualisation dashboards for healthcare quality improvement. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):689–699, 2021. doi: 10.1109/TVCG.2020.3030424.

[18] Stephen Few. *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly Media, Sebastopol, CA, USA, 2006.

[19] Md. Naimul Hoque and Nicole Sultanum. DashGuide: Authoring interactive dashboard tours for guiding dashboard users. *Computer Graphics Forum*, 44(3), 2025. doi: 10.1111/CGF.70107.

[20] Mikkel R. Jakobsen, Yonas Sahlemariam Haile, Søren Knudsen, and Kasper Hornbæk. Information visualization and proxemics: Design opportunities and empirical findings. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2386–2395, 2013. doi: 10.1109/TVCG.2013.166.

[21] Alicia Key, Bill Howe, Daniel Perry, and Cecilia R. Aragon. Vizdeck: self-organizing dashboards for visual analytics. In *Proceedings of the ACM Conference on Management of Data*, pages 681–684, New York, NY, USA, 2012. ACM. doi: 10.1145/2213836.2213931.

[22] Bongshin Lee, Petra Isenberg, Nathalie Henry Riche, and Sheelagh Carpendale. Beyond mouse and keyboard: Expanding design considerations for information visualization interactions. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2689–2698, 2012. doi: 10.1109/TVCG.2012.204.

[23] Yanna Lin, Haotian Li, Aoyu Wu, Yong Wang, and Huamin Qu. DMiner: Dashboard design mining and recommendation. *IEEE Transactions on Visualization and Computer Graphics*, 30 (7):4108–4121, 2024. doi: 10.1109/TVCG.2023.3251344.

[24] Maxim Lisnic, Vidya Setlur, and Nicole Sultanum. Plume: Scaffolding text composition in dashboards. In *Proc. CHI*, New York, 2025. ACM. ISBN 9798400713941. doi: 10.1145/3706598.3713580. URL `https://doi.org/10.1145/3706598.3713580`.

[25] Microsoft Corporation. Power bi embedded analytics client apis. `https://learn.microsoft.com/en-us/javascript/api/overview/powerbi/`, 2023. Last accessed on 2025/09/09.

[26] Microsoft Corporation. Revenue opportunities sample for power bi. `https://learn.microsoft.com/en-us/power-bi/create-reports/sample-revenue-opportunities/`, 2023. Last accessed on 2025/09/09.

[27] Microsoft Corporation. Power BI Embedded. `https://azure.microsoft.com/en-us/products/power-bi-embedd` 2023. Last accessed on 2025/09/09.

[28] Gabriela Molina León, Petra Isenberg, and Andreas Breiter. Eliciting multimodal and collaborative interactions for data exploration on large vertical displays. *IEEE Transactions on Visualization and Computer Graphics*, 30(2):1624–1637, 2024. doi: 10.1109/TVCG.2023.3323150.

[29] Matthias Nielsen, Niklas Elmqvist, and Kaj Grønbæk. Scribble query: fluid touch brushing for multivariate data visualization. In *Proceedings of the Australian Conference on Computer-Human Interaction*, pages 381–390, New York, NY, USA, 2016. ACM. doi: 10.1145/3010915.3010951.

[30] Pallets. Flask. `https://flask.palletsprojects.com/en/stable/`, 2023. Last accessed on 2025/09/09.

[31] Aditeya Pandey, Arjun Srinivasan, and Vidya Setlur. MEDLEY: intent-based recommendations to support dashboard composition. *CoRR*, abs/2208.03175, 2022. doi: 10.48550/ARXIV.2208.03175.

[32] Ayshwarya Saktheeswaran, Arjun Srinivasan, and John Stasko. Touch? speech? or touch and speech? Investigating multimodal interaction for visual network exploration and analysis. *IEEE Transactions on Visualization and Computer Graphics*, 26(6):2168–2179, 2020. doi: 10.1109/TVCG.2020.2970512.

[33] Alper Sarikaya, Michael Correll, Lyn Bartram, Melanie Tory, and Danyel Fisher. What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics*, 25(1):682–692, 2019. doi: 10.1109/TVCG.2018.2864903.

[34] Alper Sarikaya, Michael Correll, Lyn Bartram, Melanie Tory, and Danyel Fisher. What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics*, 25(1):682–692, 2019. doi: 10.1109/TVCG.2018.2864903.

[35] Kesem Sharabi, Mahir Diab, laurent-mic, Owen Duncan, Jian Dong, and Sudeep Kumar. Using the Power BI REST APIs. `https://learn.microsoft.com/en-us/rest/api/power-bi/`, 2023. Last accessed on 2025/09/09.

[36] Arjun Srinivasan and Vidya Setlur. BOLT: A natural language interface for dashboard authoring. In *Short Paper Proceedings of the Eurographics Conference on Visualization*, pages 7–11, Eindhoven, The Netherlands, 2023. Eurographics Association. doi: 10.2312/EVS.20231035.

[37] Arjun Srinivasan and John T. Stasko. Natural language interfaces for data analysis with visualization: Considering what has and could be asked. In *Short Paper Proceedings of the Eurographics Conference on Visualization*, pages 55–59, Eindhoven, The Netherlands, 2017. Eurographics Association. doi: 10.2312/EUROVISSHORT.20171133.

[38] Arjun Srinivasan and John T. Stasko. Orko: Facilitating multimodal interaction for visual exploration and analysis of networks. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):511–521, 2018. doi: 10.1109/TVCG.2017.2745219.

[39] Arjun Srinivasan, Bongshin Lee, Nathalie Henry Riche, Steven M. Drucker, and Ken Hinckley. *InChorus: Designing Consistent Multimodal Interactions for Data Visualization on Tablet Devices*, pages 1–13. ACM, New York, NY, USA, 2020. ISBN 9781450367080. doi: 10.1145/3313831.3376782. URL `https://doi.org/10.1145/3313831.3376782`.

[40] Christina Stoiber, Florian Grassinger, Margit Pohl, Holger Stitz, Marc Streit, and Wolfgang Aigner. Visualization onboarding: Learning how to read and use visualizations. In *Proceedings of the Visualization for Communication Workshop at IEEE VIS*, Los Alamitos, CA, USA, 2019. IEEE Computer Society. doi: 10.31219/osf.io/c38ab.

[41] Christina Stoiber, Conny Walchshofer, Margit Pohl, Benjamin Potzmann, Florian Grassinger, Holger Stitz, Marc Streit, and Wolfgang Aigner. Comparative evaluations of visualization onboarding methods. *Visual Informatics*, 6(4):34–50, December 2022. doi: 10.1016/j.visinf.2022.07.001.

[42] John Thompson, Arjun Srinivasan, and John T. Stasko. Tangraphe: interactive exploration of network visualizations using single hand, multi-touch gestures. In *Proceedings of the ACM Conference on Advanced Visual Interfaces*, pages 43:1–43:5, New York, NY, USA, 2018. ACM. doi: 10.1145/3206505.3206519.

[43] Melanie Tory, Lyn Bartram, Brittany Fiore-Gartland, and Anamaria Crisan. Finding Their Data Voice: Practices and Challenges of Dashboard Users. *IEEE Computer Graphics and Applications*, 43(1):22–36, 2021. doi: 10.1109/MCG.2021.3136545.

[44] Conny Walchshofer, Vaishal Dhanoa, Marc Streit, and Miriah Meyer. Transitioning to a commercial dashboarding system: Socio-technical observations and opportunities. *IEEE Transactions on Visualization and Computer Graphics*, 30(1):381–391, 2023. doi: 10.1109/TVCG.2023.3326525.

[45] Zhen Wen, Luoxuan Weng, Yinghao Tang, Runjin Zhang, Yuxin Liu, Bo Pan, Minfeng Zhu, and Wei Chen. Exploring multimodal prompt for visualization authoring with large language models, 2025. URL `https://arxiv.org/abs/2504.13700`.

[46] Aoyu Wu, Yun Wang, Mengyu Zhou, Xinyi He, Haidong Zhang, Huamin Qu, and Dongmei Zhang. MultiVision: Designing analytical dashboards with deep learning based recommendation. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):162–172, 2022. doi: 10.1109/TVCG.2021.3114826.

[47] Xiao Zhan, Noura Abdi, William Seymour, and Jose Such. Healthcare voice ai assistants: Factors influencing trust and intention to use. *Proceedings of the ACM Conference on Human–Computer Interaction*, 8(CSCW1), April 2024. doi: 10.1145/3637339. URL `https://doi.org/10.1145/3637339`.

[48] Yuheng Zhao, Junjie Wang, Linbing Xiang, Xiaowen Zhang, Zifei Guo, Cagatay Turkay, Yu Zhang, and Siming Chen. LightVA: Lightweight visual analytics with LLM agent-based task planning and execution. *IEEE Transactions on Visualization and Computer Graphics*, 31 (9):6162–6177, 2025. doi: 10.1109/TVCG.2024.3496112.