Confirmation Bias Awareness in Visual Information Retrieval via LLM-Guided Interaction-Trace Monitoring

Mariia Tytarenko*
Institute of Visual Computing
Graz University of Technology

Daniel Atzberger†
Hasso Plattner Institute
Digital Engineering Faculty
University of Potsdam

Michael Bedek‡
Department of Psychology
University of Graz

Stefan Lengauer§
Institute of Visual Computing
Graz University of Technology

Tobias Schreck¶
Institute of Visual Computing
Graz University of Technology

Abstract

Large Language Models (LLMs) are becoming a key part of interactive visual analytics systems, helping users engage in natural language interaction and explore data more effectively. However, this kind of support can sometimes unintentionally reinforce confirmation bias by consistently aligning responses with what the user has already expected and focused on. In this work, we introduce an event-driven workflow to raise awareness of confirmation bias, leveraging naturally occurring interaction traces (e.g., navigation patterns, content choices, time spent on items, and query formulation) to detect signs of confirmation bias during exploration. We apply our proposed method in the realm of Consumer Health Information Systems (CHIS), where mitigating biased information seeking and processing is particularly critical. Our work lays the groundwork for future research on cognitive bias-aware LLM agents, extending beyond confirmation bias to other forms of cognitive biases in visual analytics systems.

Index Terms: Visual Analytics; Large Language Models; Natural Language Interfaces; Confirmation Bias.

1 Introduction

Interactive Visual Analytics (VA) brings together engaging visual interfaces and automated data analysis to help users make sense of complex datasets [17]. Users can interact with the data through various means, e.g., filters, selections, and queries, while the system updates visualizations and analysis results in real time. This collaborative approach allows analysts to handle complexity, but it can also lead to a significant cognitive load. To ease this burden, recent efforts have begun incorporating Large Language Models (LLMs) into VA interfaces. These LLMs can understand the current analysis context and provide insights or suggestions. For instance, Zhao et al. presented LEVA – a system that leverages an LLM to recommend insights based on the current state of the visualization [34]. Similarly, the ProactiveVA framework uses an LLM-driven interface agent to recognize when users might need assistance and suggest analytical actions [33]. These initiatives highlight the potential of AI-driven support in VA. However, they are mainly focused on task-related assistance, while identifying and mitigating cognitive biases through user interactions remains a challenge.

*e-mail: mariia.tytarenko@tugraz.at
†e-mail: daniel.atzberger@hpi.uni-potsdam.de
‡e-mail: michael.bedek@uni-graz.at
§e-mail: s.lengauer@tugraz.at
¶e-mail: tobias.schreck@tugraz.at

Cognitive biases play a crucial role in how we perceive, evaluate, interpret, remember, and act on information [23]. Following Soprano et al., cognitive biases are systematic errors in human cognition that can impair the process of information evaluation and lead to deviations from purely rational information processing [21]. One of the main challenges for interactive systems is identifying and addressing these biases without adding extra strain on users. While the detection of some cognitive biases can be performed through questionnaires or tests informed by cognitive psychology, or by subsequently observing deviations in human judgments and decisions from purely rational outcomes, these approaches require explicit user intervention, and may disrupt the exploration process [14]. In contrast, implicit assessment based on interaction traces is attractive because it can operate unobtrusively, i.e., user interactions can provide behavioral evidence of potential cognitive bias while being naturally available to the system [25, 35].

While there are many cognitive biases discussed in cognitive psychology, in this work we focus on two distinct cognitive bias "clusters" (clusters because in both cases there are several overlapping sub-cognitive biases and related concepts): i) the confirmation bias, and ii) the biases when dealing with probabilities. The confirmation bias can be defined as the tendency to confirm an initial belief or hypothesis by proactively seeking information that is consistent with this belief, to consider such pieces of information as more credible and relevant, and to remember them better than information that contradicts the hypothesis [16]. It overlaps closely with the concepts of selective exposure and search for evidence [9], selective perception and biased assimilation [5], as well as positive hypothesis testing; i.e., the tendency to confirm a hypothesis or initial assumption rather than to falsify it. Ideally, such confirmation bias, or "confirmatory search tendency", is implicitly assessed, as it may be unconscious to the user.

As an example, in the context of health-related information seeking, a user searching online or asking an LLM might find themselves returning to content that highlights supposedly simple but ineffective home remedies, while barely touching on important viewpoints such as physician supervision, regular medication, or additional possible behavioural changes. By monitoring information-seeking behaviors, our system can recognize when users might be restricting their understanding by focusing too much on particular perspectives. Identifying these patterns is the initial step in grasping the subtleties of user behavior and preferences.

The second cognitive bias cluster revolves around the difficulty many people have in correctly interpreting and evaluating probabilities, and in neglecting base rates in the context of conditional probabilities (such as single-event probabilities, conditional probabilities, and relative risks). It has been suggested that such probability statements should be reformulated as (natural) frequency statements to foster better insights and understanding [4]. While this paper focuses on confirmation bias, probability-related biases form a second important cluster of cognitive biases that is planned to be addressed in future work via reframing strategies, e.g., natural frequencies.

In this paper, we introduce an *Agentic workflow* that provides easy-to-implement support for bias recognition and mitigation during exploratory analysis. This workflow includes (i) structured logging of interaction history (e.g., clicks, dwell time, and queries), (ii) a bias monitoring step driven by an LLM using carefully crafted prompts, and (iii) recommendations grounded in knowledge structures that encourage broader exploration. Our work in progress is not a full-featured system or empirical study.

2 Related Work

Natural Language Interfaces (NLI) for VA systems have grown rapidly in recent years, allowing users to express their needs through text or speech and receive back charts, summaries, or analytic actions. Zhang et al. showcased how a fine-tuned LLM can convert clinicians' free-text inquiries into interactive visualizations of electronic health records [32]. More broadly, Hong et al. describe how conversational GenAI agents facilitate ongoing analytic discussions during exploratory visual analysis, as well as related systems that make LLM-driven analysis easier to monitor and guide through visual representations of the model's intermediate steps [8, 29]. Early NLI systems, such as FlowSense, were designed to convert plain English commands into data-flow editing operations, enabling novice users to edit complex visualizations via a semantic parser [30]. Similarly, Talk2Data decomposed complex questions into simple ones and provided annotated charts in response [6].

A second line of research highlights that the traces left by user interactions can provide valuable insights into users' analytical behaviors, and have also been used to derive bias-related signals during analysis [25]. Nguyen et al. showcased how analyzing action sequences and visualizing user logs can uncover common workflows and help reconstruct typical behavioral patterns [15]. As AI becomes more integrated in these workflows, Holter et al. explored how decision-making is distributed in human-AI collaborations and emphasized the need to consider agency when creating AI support systems explicitly [7]. Similarly, Wang et al. investigated the changing dynamics of trust in human-AI communication through VizTrust - a visual analytics device designed to capture how trust develops during interactions [26]. Altogether, these findings point towards solutions that leverage interaction history while also ensuring user control, transparency, and appropriate system behavior.

Recent research has shown how AI-driven access to information can actually reinforce our existing beliefs and limit our exposure to a variety of viewpoints. Lopez-Lopez et al. explained how generative AI can exacerbate confirmation bias when people search for health information [12]. Abul-Fottouh et al. dive into the biases found in vaccine-related video recommendations [1], while Shen et al. raise important safety concerns about deep recommender systems, including the risks of amplification and feedback loops [20]. These insights motivate the need to develop interaction-time mechanisms that support reflection and a diversity of perspectives.

The consumer health information domain further underscores the need for adaptive, user-centered approaches. Schreck et al. argue that traditional CHIS often present static health information and outline research opportunities for adaptive visualization grounded in cognitive psychology [17]. Systems such as HealthLens show the feasibility of combining NLP and visualization to effectively support users in this field [32].

However, many existing CHIS and VA interfaces fall short by not including clear ways to identify when exploration is becoming too narrow or aligned with user preferences and to suggest interventions that encourage users to consider alternative or complementary information. In our work, we take these ideas further by introducing an agentic workflow that aims to raise bias awareness in VA.

```
Prompt:
You are an advanced analytics assistant tasked with
identifying and addressing confirmation bias in user
interactions.  The { interaction data } includes
detailed records of user interactions, and your task
is to process and interpret this data to identify
patterns indicative of confirmation bias.

Data Provided:
- Click Patterns:  Information about the tags user
clicked on.
- Time Spent:  Records of the time users spent
engaging with specific tag.
- Search Queries:  A log of user queries.

Task:
1) Analyze the provided user interaction data for
patterns that suggest confirmation bias.  For example:
- Repeated clicks on same tags and long time (over a
minute) spend on a specific tag.
- Queries that predominantly focus on supporting a
single viewpoint (Look for one-sided focus:  Does the
query exclude opposing evidence?
Check for emotionally charged words:  Words like
"proof", "best", "unfair", or "harmful" often indicate
a bias.  Analyze repeated patterns:  Are similar
queries reinforcing a specific belief without seeking
alternatives?)
2) Summarize your findings by highlighting:
Behavioral patterns that contribute to biased
understanding.
3) Provide actionable recommendations to counteract
the detected bias.
For example:
Generate prompts to encourage users to explore another
tags.

Output Format:
If bias is identified, output:  "Possible bias
identified" and add a recommendation for exploring
other sections.

Constraints:
Ensure all outputs are derived strictly from the
provided data.
Maintain a professional, respectful tone to avoid
alienating the user while nudging them toward unbiased
exploration.
```

Figure 1: Instruction prompt for detecting confirmation bias in user interactions using click patterns, time spent, and query phrasing. The assistant provides respectful suggestions for broadening user exploration.

3 Agentic Workflow

In our work, we present an agentic workflow that aims to identify and reflect on confirmation bias during exploratory analysis. It combines interaction logging, prompt-guided monitoring of LLMs, and knowledge-based guidance. Central to our approach is the use of *tags*, which indicate topics or themes the user interacts with. While in this work, we consider interactive document exploration, tags also refer to visual elements or meaningful patterns in a data visualization. Hence, our workflow is tailored for VA environments where users navigate a semantically organized information space, e.g., within a CHIS, that categorizes topics with tags, descriptive texts, and related visual representations. The goal is not to diagnose or label users as biased, but rather to identify interaction patterns indicative of a narrowing of exploration and to foster reflective sensemaking by providing contextually grounded suggestions.

Our minimal workflow follows these steps: Interaction capture → Trigger Policy → Bias Monitoring (LLM prompt) → Reflective Guidance (knowledge-grounded).

1. Interaction capture. During user exploration of the VA interface, the system records user behavior:

- Click Patterns – keeping track of which articles or data points users click on the most, e.g., a selected tag such as "Complications", "Nutrition", or "Treatment", or interactions with related visualizations that highlight these concepts.

- Search Queries – noting the details in user queries can reveal if they are searching for information that aligns with their existing beliefs [10].

- Time Spent - Looking at how long users engage with a specific chapter or viewpoints, e.g., repeatedly looking at sections of the text indicating the risks of insulin therapy while not reading those sections reporting its benefits. It may indicate a biased tendency towards negative outcomes.

These are stored in a structured (e.g., JSON) form that contains an ordered record of the user's engagement over time.

2. Triggering. Bias monitoring is executed when a trigger policy is activated. In our proof-of-concept prototype, we trigger bias monitoring on demand (when explicitly invoked). This approach lets us showcase the entire workflow in a controlled environment. Looking ahead, this straightforward trigger might be improved with automatic heuristics that will run monitoring after a set period, after a certain number of interactions, or when specific patterns emerge.

3. Bias monitoring through prompt-guided LLM analysis. On activation, the system takes a snapshot of a recent interaction history and sends it to the LLM via an instruction prompt (see Fig 1).

4. Knowledge-grounded reflective guidance. If a potential bias pattern is detected, the system uses the knowledge assets (tag relations, texts, and associated visualizations/views) to recommend complementary topics and generates a brief, non-judgmental suggestion that encourages broader exploration.

To tackle identified biases and promote a broader range of exploration, the system might utilize a prompt similar to the one in the figure 1 (which is based on preference-inconsistent recommendations, e.g., [18], as mitigation strategy, a form of "considering the opposite" [13]). In this paper, we present a proof-of-concept that uses a single prompt to monitor cognitive biases via interaction tracking, focusing on the confirmation bias as an illustrative example. This prompt not only signals when a "possible bias" is detected but also offers an initial suggestion to broaden the exploration. The design is such that the bias signal remains hidden from the user; instead, it initiates a specific recommendation step. Moving forward, we plan to develop actionable knowledge-grounded recommendations that leverage tag relations, descriptive texts, and associated visualization views.

Since prompts are essential to our system's successful integration of NLI into the VA environment, prompt engineering is a key element in ensuring seamless user-LLM communication [27]. However, creating effective prompts can be challenging, particularly for those without prior experience in prompt optimization, design, and formulation. Papers on guidance for creating prompts [31, 22] might be helpful in this case. In our design, prompts are defined mainly by the system administrator, who designs them to control LLM behavior for tasks, such as summarization or bias detection. At the same time, user input via the chat interface serves as implicit prompts—free-text input that initiates an LLM response. This arrangement combines expert-defined structure and user-defined flexibility.

4 APPLICATION EXAMPLE

To illustrate the agentic workflow in context, we use an existing CHIS system, called A⁺CHIS [19], that supports the structured exploration of a health topic – *type-2 diabetes mellitus* (T2DM) – through a document-derived knowledge structure. In our case, the system uses a 124-page document on Diabetes provided by the German public health insurance AOK, titled "Diabetes under Control" ("Den Diabetes im Griff" in German) [2]. A⁺CHIS offers a unique way to navigate documents by combining various coordinated views. This setup allows users to see the structure of topics alongside the text and related visuals, making it easier to dive into the material at different levels of detail [24]. Figure 2 presents some of the visualisation components implemented in our system.

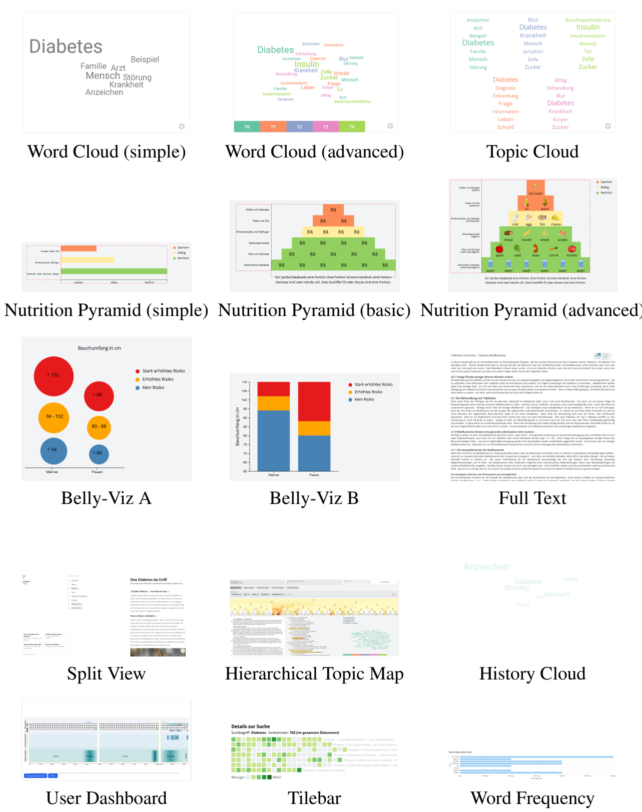

Figure 2: Gallery of some of the A⁺CHIS components. The system offers various document visualisations and infographics, some of which are available at multiple levels of detail (e.g., the Nutrition Pyramid). Together, these views could provide interaction signals for bias tracking.

In such systems, cognitive biases may emerge as users selectively seek information that aligns with an evolving concern or prior beliefs, especially when the information is highly personal [16, 25]. Detecting relevant cognitive biases is essential in such systems because unaddressed biases can skew decision-making and undermine the reliability of insights.

Within this broader system, we have created a specific view that enables the workflow in our proof-of-concept scenario (see Fig. 3). This view combines essential components: (i) an NLI in the form of a chatbot, (ii) a structured information space, and (iii) synchronized views of the document and visualizations. On the left side (panel ⓐ), it shows a *User History* panel that displays the user's recent exploration history in a chronological format (e.g., clicked topics/tags, recent conversational queries, text highlights, etc.). This

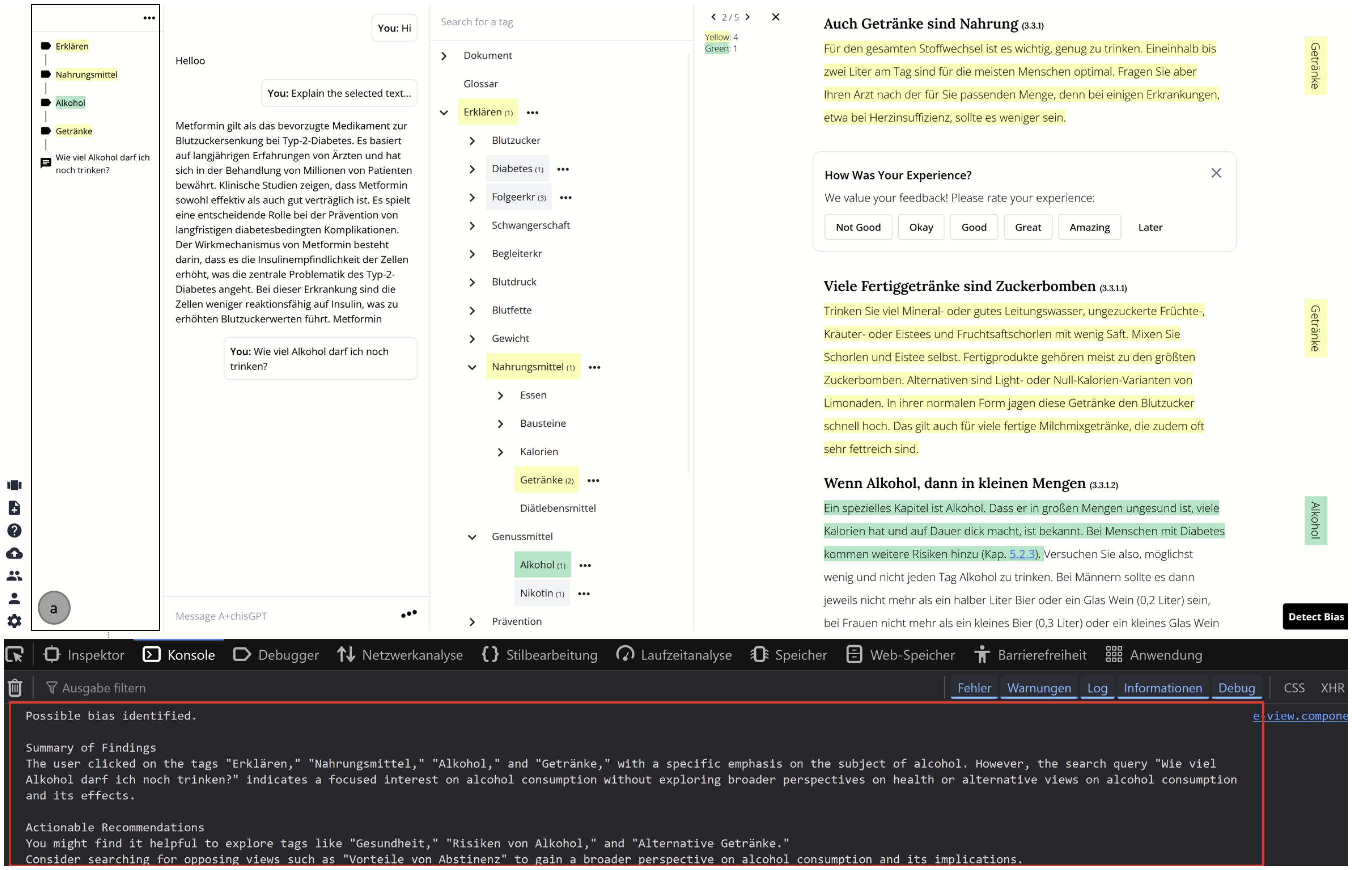

Figure 3: The system demonstrates its ability to detect bias by analyzing user interactions, particularly those related to the query “Alkohol” (alcohol). Through zero-shot prompting, it uncovers potential biases and recommends alternative tags, such as “Gesundheit” (health) and “Vorteile von Abstinenz” (benefits of abstinence), encouraging a more balanced exploration of the topic. Notably, the identified biases are only visible to the system (as indicated in the console), providing subtle guidance that aims to support reflective exploration, enhance decision-making, and encourage thorough user engagement.

panel serves as a concise representation of the behavioral evidence we use for the bias monitoring step.

One of our main design goals is to present the monitoring results in a structured format that the system can easily read [11]. In our current prototype, the monitoring step is triggered on demand and examines the recorded interaction trace for signs of a *possible* confirmation bias. It is important to note that the resulting bias flag is considered *system-internal*, i.e., it is used solely as a control signal for subsequent recommendation actions and is not intended to be visible to users.

For example, in a recorded session (see Figure 3), the user mainly engaged with the categories “Nahrungsmittel” (food), “Getränke” (beverages), and specifically “Alkohol” (alcohol). The user query, “Wie viel Alkohol darf ich noch trinken?” (How much alcohol can I still drink?) indicates a focused interest in alcohol consumption, which may overlook broader dietary or health-related aspects of managing T2DM. When the bias identification was triggered, the system flagged this interaction as a potential bias and provided actionable recommendations. In this case, the evidence comes from the frequent revisits to a small cluster of topics, the limited exploration of other diet-related areas, and the way the user phrased their question, focusing mainly on maintaining an existing behavior rather than considering alternatives or broader context. Following the internal signal, the system issues preliminary reflective guidance, consistent with our current prompt design (that further can be shown to the user): it suggests concrete next exploration steps that broaden topical coverage by leveraging the document-derived structure (e.g., moving from “Alkohol” to related dietary topics, recommendations, or risk factors) and by pointing to relevant document sections and associated visual views. In this example, these suggestions serve as an initial demonstration of feasibility rather than a finalized recommendation strategy. Designing stronger, knowledge-grounded recommendation policies remains an important direction for future work.

This example illustrates how an interaction trace can activate a system-internal “possible bias identified” signal, leading to light, thoughtful next-step suggestions based on the document’s structure. While we showcase this in a visual CHIS environment, the same approach can be adapted for other interactive systems that track user behavior.

5 Discussion & Future Work

In the context of our current example, detecting a possible confirmation bias pattern from interaction traces should be understood as only the first step toward mitigating the bias, rather than a complete solution. However, there is work to be done to transform this demonstration into a generalizable method.

The next step is to build this into a more solid pipeline with clearer component boundaries – distinguishing between monitoring (producing a system-internal bias signal and evidence), triggering (deciding when sufficient evidence has accumulated), and intervention (generating appropriate guidance). Future design could explicitly use agents to maintain user states over time, plan multi-step guidance, and adjust its behavior based on how users interact afterward, while still preserving user control over whether and how guidance should be applied.

A key area of research is to apply effective interventions to a broader range of cognitive biases that come into play during exploratory information seeking. A systematic literature review on a wide range of cognitive biases in fact-checking – an area that is similar to our health-related exploratory (as well as goal-oriented) information seeking – is given in [21]. Many of the cognitive biases described in the literature are conceptually similar or partially overlapping, suggesting that further research is needed to determine which distinct interactions can distinguish among similar biases. This requires identifying which combinations of signals (navigation diversity, revisits, concentrated dwell time, cues from query intent, interactions with documents and visualizations, etc.) provide reliable evidence of bias. It is also crucial to tackle ambiguity and minimize false positives, mainly when goals rather than biases drive focused exploration. However, in some cases, a valid monitoring that is capable of distinguishing similar cognitive biases might not be necessary if the associated interventions have no negative effects anyway (for example, if an incorrect classification in bias detection leads to an intervention that either does not reinforce the actual cognitive bias, or perhaps even reduces it, albeit less effectively and tailored).

As mentioned above, the other cognitive bias cluster considered highly relevant in the context of health-related information seeking and processing revolves around the correct (i.e., rational) interpretation and understanding of probabilities, particularly conditional probabilities. In cases a user is confronted with conditional probabilities, the intervention component could directly recommend a reformulation as natural frequency statement, without relying on the monitoring and triggering component, because measuring the presence of such a cognitive bias on the side of the user is prone to errors due to insufficient or limited interaction data or per se unnecessary: Even experts benefit from such intervention, as studies by Gigerenzer and others with samples of medical doctors show (e.g. [4]).

At the same time, we should enhance our current recommendation process beyond the initial demonstration by developing strategies grounded in knowledge that utilize the document structure (e.g., hierarchies, prerequisites, and related topics), associated texts, and visualization views. Designing stronger, knowledge-grounded recommendation policies remains an important direction for future work, including selecting complementary tags based on hierarchical and dependency relations, balancing diversity and relevance, and incorporating visualization-specific actions (e.g., prompting users to inspect comparison views or overview summaries that contextualize risks and alternatives).

Looking ahead, it is important to refine how we design prompts into stable "contracts" between different components of the system. This means keeping the internal monitoring outputs — e.g. bias signals, separate from the guidance that users see. Finally, the workflow opens opportunities for richer user modeling, the same sources of interaction, such as chatbot interaction patterns, document navigation, and visualization interactions, can provide insights into user intent, visualization literacy [3] or textual literacy [28]. This, in turn, allows for more tailored levels of explanation and guidance. For evaluating these directions, we'll need empirical studies that validate the quality of signals, measure how guidance affects exploration and decision-making quality, and evaluate user perceptions of agency, trust, and usefulness in realistic exploratory contexts.

6 Conclusion

We introduced a workflow designed to provide users bias-aware support during the exploration in VA systems. This approach leverages interaction traces and prompt-guided LLM monitoring to identify signals of confirmation bias and to offer light, domain-, and knowledge-based suggestions. Showcased in a CHIS context, our workflow encourages thoughtful exploration while keeping the bias signal within the system, ensuring users maintain their agency. Our findings pave the way for future research into proactive triggering policies, broader bias coverage, and the practical assessment of bias-aware support during interactive exploration.

Acknowledgements

This work was funded by the Austrian Science Fund (FWF) as part of the project 'Human-Centered Interactive Adaptive Visual Approaches in High-Quality Health Information' (Adaptive and interactive CHIS (A^+CHIS); Grant No. FG 11-B).

References

[1] D. Abul-Fottouh, M. Y. Song, and A. Gruzd. Examining algorithmic biases in YouTube's recommendations of vaccine videos. *International Journal of Medical Informatics*, 140:104175, 2020. doi: 10.1016/j.ijmedinf.2020.104175 2

[2] AOK. Den Diabetes im Griff – Handbuch zu Diabetes mellitus Typ 2, 2023. URL: https://www.aok.de/pk/magazin/cms/fileadmin/pk/pdf/patientenhandbuch-diabetes.pdf, Accessed on 29.12.2024. 3

[3] M. Avgerinou and J. Ericson. A Review of the Concept of Visual Literacy. *British journal of Educational Technology*, 28(4):280–291, 1997. doi: 10.1111/1467-8535.00035 5

[4] G. Gigerenzer and A. Edwards. Simple tools for understanding risks: from innumeracy to insight. *The BMJ*, 327:741–744, 2003. doi: 10.1136/bmj.327.7417.741 1, 5

[5] T. Greitemeyer, P. Fischer, D. Frey, and S. Schulz-Hardt. Biased assimilation: The role of source position. *European Journal of Social Psychology*, 39(1):22–39, 2009. doi: 10.1002/ejsp.497 1

[6] Y. Guo, D. Shi, M. Guo, Y. Wu, N. Cao, and Q. Chen. Talk2data: A natural language interface for exploratory visual analysis via question decomposition. *ACM Transactions on Interactive Intelligent Systems*, 14(2):articleno. 8: 1–24, 2024. doi: 10.1145/3643894 2

[7] S. Holter, C. Moruzzi, and M. El-Assady. Towards Agency in Human-AI Collaboration. *IEEE Computer Graphics and Applications*, 46(1):13–25, 2025. doi: 10.1109/MCG.2025.3623892 2

[8] M.-H. Hong and A. Crisan. Data Has Entered the Chat: How Data Workers Conduct Exploratory Visual Analytic Conversations with GenAI Agents. *ACM Transactions on Interactive Intelligent Systems*, 15(4):articleno. 21: 1–40, 2025. doi: 10.1145/3744750 2

[9] E. Jonas, S. Schulz-Hardt, D. Frey, and N. Thelen. Confirmation bias in sequential information search after preliminary decisions: an expansion of dissonance theoretical research on selective exposure to information. *Journal of Personality and Social Psychology*, 80(4):557–571, 2001. doi: 10.1037/0022-3514.80.4.557 1

[10] S. Li, T. J. Davidson, C. Xiong Bearfield, and E. Wall. Confirmation Bias: The Double-Edged Sword of Data Facts in Visual Data Communication. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, pp. articleno. 1172: 1–16. ACM, 2025. doi: 10.1145/3706598.3713831 3

[11] Y. Liu, D. Li, K. Wang, Z. Xiong, F. Shi, J. Wang, B. Li, and B. Hang. Are LLMs good at structured outputs? A benchmark for evaluating structured output capabilities in LLMs. *Information Processing & Management*, 61:103809, 2024. doi: 10.1016/j.ipm.2024.103809 4

[12] E. Lopez-Lopez, C. M. Abels, D. Holford, S. M. Herzog, and S. Lewandowsky. Generative artificial intelligence–mediated confirmation bias in health information seeking. *Annals of the New York Academy of Sciences*, 1550(1):23–36, 2025. doi: 10.1111/nyas.15413 2

[13] C. G. Lord, M. R. Lepper, and E. Preston. Considering the opposite: A corrective strategy for social judgment. *Journal of Personal-*

ity and Social Psychology, 47(6):1231–1243, 1984. doi: doi.org/10.1037/0022-3514.47.6.1231 3

[14] G. Mark, D. Gudith, and U. Klocke. The cost of interrupted work: more speed and stress. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pp. 107–110. ACM, 2008. doi: 10.1145/1357054.1357072 1

[15] P. Nguyen, C. Turkay, G. Andrienko, N. Andrienko, and O. Thonnard. A visual analytics approach for user behaviour understanding through action sequence analysis. In *EuroVis Workshop on Visual Analytics*. EG, 2017. doi: 10.2312/eurova.20171122 2

[16] R. S. Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2):175–220, 1998. 1, 3

[17] T. Schreck, D. Albert, M. Bedek, K. Horvath, K. Jeitler, B. Kubicek, T. Semlitsch, L. Shao, and A. Siebenhofer. *Adaptive Visualization of Health Information Based on Cognitive Psychology: Scenarios, Concepts, and Research Opportunities*, pp. 165–195. Springer, 2023. doi: 10.1007/978-3-031-34738-2_7 1, 2

[18] C. Schwind and J. Buder. Reducing confirmation bias and evaluation bias: When are preference-inconsistent recommendations effective–and when not? *Computers in Human Behavior*, 28(6):2280–2290, 2012. doi: 10.1016/j.chb.2012.06.035 3

[19] L. Shao, S. Lengauer, H. Miri, M. Bedek, B. Kubicek, C. Kupfer, M. Zangl, B. Dienstbier, K. Jeitler, C. Krenn, T. Semlitsch, C. Zipp, D. Albert, A. Siebenhofer, and T. Schreck. Visual Document Exploration with Adaptive Level of Detail: Design, Implementation and Evaluation in the Health Information Domain. In *Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2023) - IVAPP*, pp. 133–141. SciTePress, 2023. doi: 10.5220/0011621800003417 3

[20] H. Shen, Y. Liu, K. Zhang, Q. Cao, and X. Cheng. The rising safety concerns of deep recommender systems. *The Innovation*, 6(10):101038, 2025. doi: 10.1016/j.xinn.2025.101038 2

[21] M. Soprano, K. Roitero, D. La Barbera, D. Ceolin, D. Spina, G. Demartini, and S. Mizzaro. Cognitive Biases in Fact-Checking and Their Countermeasures: A Review. *Information Processing & Management*, 61(3):103672, 2024. doi: 10.1016/j.ipm.2024.103672 1, 5

[22] H. Subramonyam, R. Pea, C. Pondoc, M. Agrawala, and C. Seifert. Bridging the gulf of envisioning: Cognitive challenges in prompt based interactions with LLMs. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, CHI '24, pp. articleno. 1039: 1–19. ACM, 2024. doi: 10.1145/3613904.3642754 3

[23] A. Tversky and D. Kahneman. Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157):1124–1131, 1974. doi: 10.1126/science.185.4157.1124 1

[24] M. Tytarenko, C. W. Burtscher, D. Atzberger, A. Jobst, W. Scheibel, S. Lengauer, and T. Schreck. A user-centric adaption model for document visualizations with different levels of detail within a consumer health information system (short paper). In *IUI Workshops*, 2025. 3

[25] E. Wall, L. M. Blaha, L. Franklin, and A. Endert. Warning, Bias May Occur: A Proposed Approach to Detecting Cognitive Bias in Interactive Visual Analytics. In *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 104–115, 2017. doi: 10.1109/VAST.2017.8585669 1, 2, 3

[26] X. Wang, S. Tulk Jesso, S. Kojaku, D. M. Neyens, and M. S. Kim. VizTrust: A Visual Analytics Tool for Capturing User Trust Dynamics in Human-AI Communication. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '25, pp. articleno. 581:1–10. ACM, 2025. doi: 10.1145/3706599.3719798 2

[27] Y. Wang, Z. Jiang, Z. Chen, F. Yang, Y. Zhou, E. Cho, X. Fan, Y. Lu, X. Huang, and Y. Yang. RecMind: Large Language Model Powered Agent For Recommendation. In *Findings of the Association for Computational Linguistics: NAACL 2024*, pp. 4351–4364. ACL, 2024. doi: 10.18653/v1/2024.findings-naacl.271 3

[28] P. Wright. Textual literacy: An outline sketch of psychological research on reading and writing. In *Processing of Visible Language*, pp. 517–535. Springer, Boston, MA, USA, 1980. doi: 10.1007/978-1-4684-1068-6_39 5

[29] L. Xie, C. Zheng, H. Xia, H. Qu, and C. Zhu-Tian. WaitGPT: Monitoring and Steering Conversational LLM Agent in Data Analysis with On-the-Fly Code Visualization. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*, UIST '24, pp. articleno. 119:1–14. ACM, 2024. doi: 10.1145/3654777.3676374 2

[30] B. Yu and C. T. Silva. FlowSense: A Natural Language Interface for Visual Data Exploration within a Dataflow System. *IEEE Transactions on Visualization & Computer Graphics*, 26(1):1–11, 2020. doi: 10.1109/TVCG.2019.2934668 2

[31] J. Zamfirescu-Pereira, R. Y. Wong, B. Hartmann, and Q. Yang. Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, pp. articleno. 437:1–21. ACM, 2023. doi: 10.1145/3544548.3581388 3

[32] H. Zhang, S. Ning, Q. Zheng, Y. Song, and L.-J. Zhang. HealthLens: A Natural Language Querying System for Interactive Visualization of Electronic Health Records. In *Proceedings of the Thirty-Fourth International Joint Conference on Artificial Intelligence (IJCAI-25)*, pp. 11123–11126, 09 2025. doi: 10.24963/ijcai.2025/1276 2

[33] Y. Zhao, X. Shu, L. Fan, L. Gao, Y. Zhang, and S. Chen. ProactiveVA: Proactive Visual Analytics with LLM-Based UI Agent. *IEEE Transactions on Visualization and Computer Graphics*, 32(1):451–461, 2026. doi: 10.1109/TVCG.2025.3642628 1

[34] Y. Zhao, Y. Zhang, Y. Zhang, X. Zhao, J. Wang, Z. Shao, C. Turkay, and S. Chen. LEVA: Using Large Language Models to Enhance Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 31(3):1830–1847, 2025. doi: 10.1109/TVCG.2024.3368060 1

[35] Z. Zhou, X. Wen, Y. Wang, and D. Gotz. Modeling and Leveraging Analytic Focus During Exploratory Visual Analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, 2021. doi: 10.1145/3411764.3445674 1