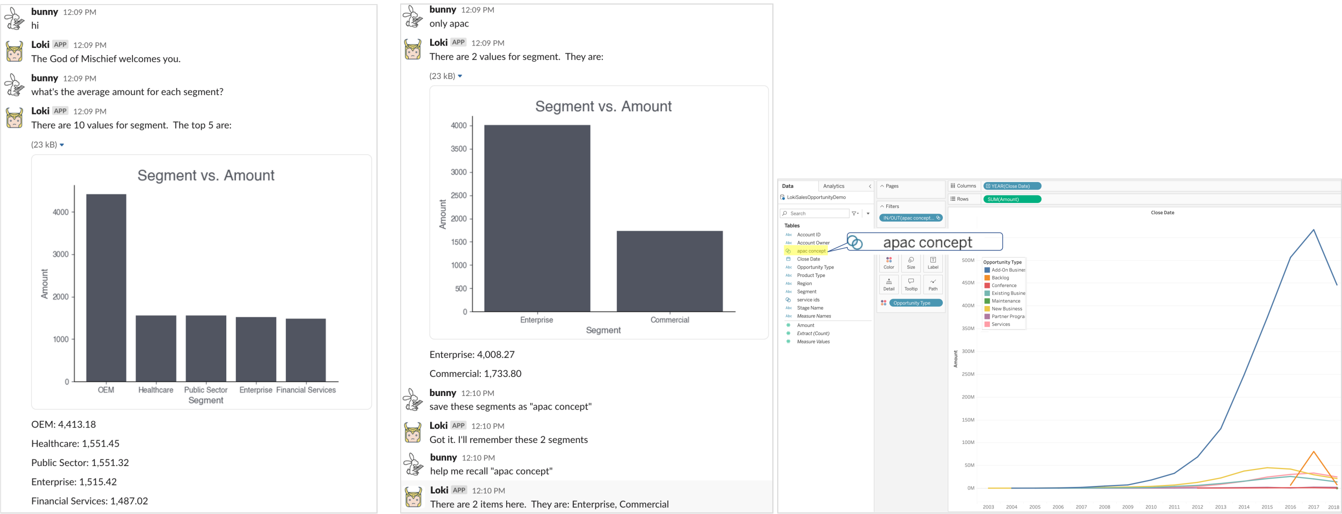
LOKI: Reusing Custom Concepts in Interactive Analytic Workflows

Vidya Setlur and Andrew Beers

Tableau Research, USA



(a) Initial interaction with LOKI (b) Saving custom concept (c) Using concept (highlighted in yellow) in a Tableau workbook

Figure 1: Conversational interaction between a user, 'bunny' and LOKI using a sales transactions data source. (a) LOKI responds for the initial query, "what is the average amount for each segment?". (b) The user continues to interact by filtering to the 'apac' segment and saving the query as a reusable concept, *apac concept*. (c) The reusable concept, *apac concept*, is used as a filter on 'apac' segments in the line chart.

Abstract

Natural language (NL) interaction enables users to be expressive with their queries when exploring data. Users often specify complex NL queries that involve a combination of grouping, aggregations, and conditionals of data attributes and values. Such queries are often reused several times by users during their analytical workflows. Existing systems offer limited support to save these bespoke queries as concepts that can be referenced in subsequent NL queries, leading to users having to respecify these queries repeatedly. To address this issue, we describe a system, LOKI *that allows users to save complex and bespoke queries as reusable concepts and use these concepts in other NL queries and analytics tools. For example, users can save an NL query, "show me the opportunity amount by customer for open opportunities" in a sales dataset, as a concept* 'followup customers' *and reference this custom concept in a query such as "show me the total opportunity amount for followup customers." A qualitative evaluation of* LOKI *indicates the usefulness of supporting the reuse of custom concepts across various analytical workflows. We identify future research directions around in-situ semantic enrichment and dynamic concept maps for data exploration.*

CCS Concepts

• ***Human-centered computing*** → ***Interactive systems and tools; Visualization application domains;***

1. Introduction

Curated data with semantically meaningful attributes and concepts is an important aspect of the visual analysis workflow [ask21]. With the prevalence of natural language interfaces (NLIs) for data exploration [GDA*15, SBT*16, SS18, ibm21, pow21, tho21], curated concepts are a requisite aspect of supporting expressiveness when users craft natural language (NL) utterances. However, the process of information seeking is often dynamic and complex, where pre-defined curated static concepts are rather limiting as users express their utterances during the flow of their analysis. These utterances often contain complex analytical concepts involving *mul-*

tiple attributes that also specify conditions like *filter expressions* as users explore different aspects and subdomains of the underlying data [STD19]. Further, users tend to reuse these complex utterances in subsequent interactions or to other datasets. In other words, there is a need for *in-situ semantic curation* of analytical concepts that could be reused across different analytical workflows.

To realize custom analytical concepts during data exploration, NL utterances need to be represented as concrete descriptions that match the structure or content of the underlying data. Specifically, subexpressions within these utterances have to be converted into a set of analytical sub-expressions that resolve into structured queries containing attributes and values from the underlying data source. Concept-based or conceptual query interfaces focus on this problem by enabling users to directly name and save constructs when querying databases [BS94, CWS93, OWE00].

To explore how analytical concepts can be dynamically created and subsequently reused in analytical workflows, we introduce a system, LOKI.[†] We implemented LOKI on the Slack platform, a popular NL messaging and analytical chatbot environment [ST22] that provides APIs to programmatically implement parser logic for interpreting NL utterances and customize system responses to include text and visualization image responses (Figures 1a and 1b). A parser determines intents and extracts entities from the utterances based on a set of ML-trained analytical conversation patterns, returning text and chart responses to the user. During the interaction with the interface, users can save their conditional utterances as reusable concepts in the datasource that they are exploring and use these concepts in other queries and visual analytics tools such as Tableau. An evaluation of LOKI provides useful insights for supporting in-situ semantic enrichment through the creation of reusable analytical concepts.

2. Related Work

2.1. Semantic enrichment of data

Semantic enrichment encapsulates rich semantic knowledge into a structured dataset of concepts to provide a better understanding of the underlying data and, consequently, the ability to reason about the data. Knowledge graphs are often employed in semantic enrichment using well-structured, encyclopedic knowledge about a wide spectrum of entities [ABK*07, VK14, PTWS20]. Other research has explored the use of domain-specific concepts ontologies to provide entity relationships between data attributes [BKOH17, LdKLC13]. Systems have also explored improving the expressiveness of database queries by augmenting ontologies with synonymous concepts [LEGÖ20]. While upfront semantic enrichment can support a wider repertoire of analytical inquiry, none of this prior work has explored ways to support *in-situ* semantic enrichment of data in the flow of visual analysis. Our work explores this gap by enabling semantic enrichment in the context of NLIs for visual analysis.

[†] The name LOKI is inspired by the fictional Marvel Comic character who possesses a brilliant intellect with some knowledge of technology, capable of influencing humans' actions [LLI*14].

2.2. NLIs for visual analysis

Augmenting data with additional concepts has enabled a variety of NLIs for data exploration. Ontology-based NLIs augment linguistic patterns with domain knowledge and allow reasoning at the level of real-world entities for interpreting the user's utterance and converting it into a structured query [LÖQ*18]. However, these patterns are identified ahead of time and do not account for user-defined concepts during the query process.

NLIs for visual analysis [GDA*15, DMN*17, SS21, tho21, ibm21, pow21, LTL*21, LHJY21] focus on interpreting users' utterances by providing useful visualization responses and recommendations. The methods of interpreting intent typically rely on explicitly named data attributes and values in the user's input queries. Systems like Eviza [SBT*16] and Ask Data [STD19] handle impreciseness around vague numerical concepts such as 'cheap' and 'high' by dynamically inferring a range based on the statistical properties of the data. Users can adjust the range and save those preferences for subsequent interaction. While the concepts and their interpretation were dynamically computed during the user session, they were simple range filters with only the bounds being modifiable by the user. Our paper explores more complex filter concepts that can be saved and reused across data sources and visual analysis workflows.

Sneak Pique [SHKC20] examined how textual and visual variants of autocompletion with data previews provide users guidance within the context of NLIs for visual analysis. GeoSneakPique extends the concept of data-driven scaffolds from Sneak Pique by exploring how vague definitions for places can be expressed as user-defined concrete specifications of cognitive regions [SBW21]. Our work, LOKI, further generalizes the notion of user-defined concepts beyond places to any attribute in a tabular dataset, where complex filter expressions can be saved as named concepts and used in subsequent NL queries and in data analysis tools, such as Tableau.

3. System

3.1. Overview

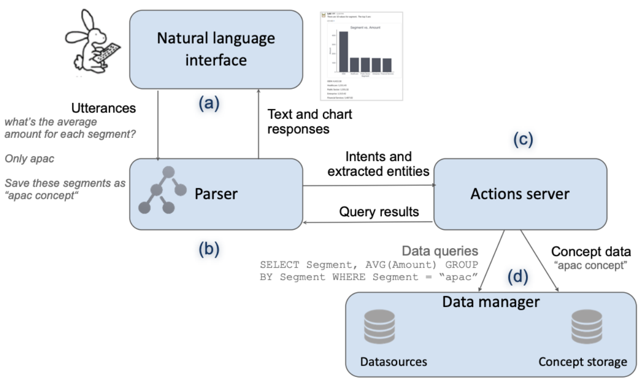


Figure 2: System overview of LOKI. (a) The NLI allows users to interact in Slack, (b) the parser interprets the intent of the NL utterances, (c) the actions server triggers actions based on the intent and executes queries to the data sources, and (d) the data manager manages query execution and concept persistent in the data sources.

LOKI is an analytical chatbot (i.e., a software application that engages in an NL dialogue with the user about data [ST22]) with a user interface exposed via Slack. Figure 1 shows a sample user session in LOKI. We employ a parser trained on a set of analytical intents using an open-source conversational AI library [ras21]. The intents and data entities extracted from the input utterance are passed to an action server. The action server runs analytical queries against a set of configured data sources, returning the results to the user. The system can create named concepts for later use in NL utterances, as well as persist these concepts in the data sources for use by other systems such as Tableau. The data sources could be any tabular CSV file, but for the purpose of this prototype, we consider data sources across topics such as sales [sup23, Sal23], world events [cov23], and entertainment [Ban21]. Figure 2 shows the system overview. A video of the system is included as supplementary material. We now describe the various components of the system.

3.2. Natural language interface (NLI)

LOKI supports both informational and analytical utterances. *Informational utterances* allow the user to understand the data that is available. An utterance, "what data is available?" responds with a list of the data sources that were configured during training, and "tell me about sales" responds with metadata about the data source named *Sales*, listing out the names of available columns that can be used in queries. LOKI also displays the data source's metadata in the event that it is unable to interpret the user's utterance.

We draw precedence from previous NLI research [STD19, TS19] to support a range of *analytical utterances* that include aggregation (e.g., "what is the average amount?"), group (e.g., "show me the average amount for each segment"), and filter (e.g., "filter to central") operations. LOKI produces a text response for a single result (e.g., a single average value for "what is the average amount?") and visualizations for results that return multiple rows (e.g., a bar chart for "show me the average amount for each segment").

3.3. Parser Module

The parser is trained on a set of analytical inquiries using templates of crowdsourced analytical utterances [STD19, ask21, SS18]. Templates include slots for attributes such as "what is the total (measure)?" where 'measure' is the name of the attribute the parser would need to recognize from the NL utterance. As part of training, a set of data sources are specified over which LOKI operates. Sample utterances and entities are dynamically generated as input to the training using the templates as a guide. For example, consider the data source with attributes, `sales`, `profit`, and `segment`. If there exist training templates - "show me (measure) by (dimension)" and "only for (value)", LOKI will classify `sales` and `profit` as measures because they are numeric, and `segment` as a dimension. The parser module handles imprecision such as misspellings and incomplete input, by applying fuzzy matching and lemmatization on the tokens in the utterances with attributes in the data sources [MRS08].

3.4. Actions Server

In LOKI, actions represent the analytical operations that can satisfy a user's intent. Each action is realized with a corresponding function that parameterizes the recognized entities from the NL utterances and the current conversational state of the system. The actions then trigger queries to the data sources for retrieving results for the utterances as well as persisting named concepts. This conversational state encompasses the current data source, the most recent query posed, the result for that query, any filters that are in play, and any saved concepts created by the users.

3.5. Reusable Concepts

A user can create a *reusable concept* from the result by saving the underlying data as a named concept or by querying the result and saving the resulting data as a named concept. In the former case, an utterance, "let's call these segments *c1*", would save the data in the `segment` attribute as a concept *c1*. In the latter case, for an utterance such as "let's call these accounts *c2*" the referenced attribute `account` is not part of the current query results, and so a new query is issued that adds data from `account` as the concept *c2*. LOKI also supports nested custom concepts with utterances such as, "show me profit by *c1*" and "save this concept as *c3*," where a nested custom concept *c3* is created from *c1*, grouped by the attribute, `profit`.

These concepts can subsequently be used in analytical utterances and in other related data sources with shared attributes. For example, "what is the average amount for *c1*" would return the average of the attribute, `amount` for those records that matched the concept *c1*. Concept *c1* can be subsequently used in other data sources that have an attribute, `segment`. Figure 1 shows a concept, *apac concept* based on a data source containing sales transactions by salespeople to particular customer accounts. A concept created on the first data source based on the account column could be used in queries against another data source containing marketing campaign activity associated with customer accounts.

Reusing concepts in other visual analysis workflows. Reusable concepts are useful in other contexts besides the LOKI chatbot application. LOKI supports the persistence of these concepts in the data source from which they were created by the user. A user can make a concept available in a data source via an utterance such as "publish apac concept", which will persist the concept, *apac concept* in the data source from which it was created. This concept can then be used in other visual analysis tools that import the data source for analysis. Figure 1c shows how the concept, *apac concept*, is used as a custom field to filter values in a line chart.

4. Preliminary User Study

We conducted a preliminary user study of LOKI to (1) collect feedback on how people express and save user-defined concepts and (2) identify system limitations and opportunities for how reusable concepts can be used to further data exploration.

4.1. Participants

We recruited 12 volunteers (P1-P12, six males, six females) from a local town mailing list. The participants had a variety of backgrounds - user researcher, sales consultant, engineer, product manager, real estate broker, and marketing manager. Based on self-reporting by the participants, all were fluent in English and regu-

Proceedings published by Eurographics - The European Association for Computer Graphics.

larly used some type of NL interface, such as Google. Seven participants regularly used a visualization tool, and the rest had limited proficiency.

4.2. Method

Participants could interact with two sales datasets [Sal23] - (1) 'Sales Opportunities' with attributes describing product sales such as, `Account ID`, `Account Owner`, `Amount`, `Region`, `Segment`, `Opportunity Type`, `Product Type`, and `Close Date`; (2) 'Online Activity' with attributes describing activity in an online service such as, `Sales Account Id`, `Activity Count`, and `Activity Name`.

We commenced with a short introduction of how to use LOKI and the types of intents it recognizes. Participants were instructed to phrase their queries in whatever way that felt most natural and to tell us whenever the system did something unexpected. We discussed reactions to system behavior throughout the session and concluded with an interview. Each session took approximately 45 minutes. We employed a mixed-methods approach involving qualitative and quantitative analysis but considered the quantitative analysis as a complement to our qualitative findings.

4.3. Results

Overall, participants were positive about the premise of saving and reusing custom concepts in their analytical workflows. They appreciated the convenience of being able to save complex and compound queries that involve grouping, aggregation, and filter expressions as named concepts. "It's like bookmarking web pages as you are browsing. I can save them and then come back to it later [$P6$]" The total number of queries that participants typed ranged from 7 to 28 ($\mu = 12.2$). The number of times participants saved an utterance as a concept in each study session ranged from 3 to 12 ($\mu = 6.5$). Participants reused these saved concepts 3 to 10 ($\mu = 5.1$) times in subsequent utterances in their user sessions. The most common analytical queries involved group and filter expressions (48%) such as "show me account owners by segment filter region to apac." The second most common concept involved aggregation, group, and filter expressions (45%) such as "average amount by product type, only services." The remaining concepts involved group utterances such as "account owner by segment."

5. Discussion and Future Work

The study also reveals several shortcomings and provides opportunities for better supporting and leveraging custom concepts:

Automatic construction of concept maps While participants appreciated the ability to name and persist concepts, they wanted additional system capabilities for constructing concept maps that take into account hierarchies and relationships between concepts and the data attributes. For example, $P4$ expressed, "I first saved a concept for customers in Europe, and I created another one where I filtered by Close Date. I would have liked a way to see my concepts with the hierarchy between them." An extension of LOKI would be to explore concept map construction based on common attributes and determining relational dependencies in a concept hierarchy.

Repair and refinement of concepts: LOKI enables users to refer back to the concepts they have saved. However, there were consistent requests to have the ability to repair the concepts, especially modifying the filter criteria saved in the concept. $P8$ remarked, "I created a concept, 'customer watch' based on the region, but then I wanted to change it and couldn't find an easy way to do that other than creating a new one." However, this feedback opens an interesting topic on permissions and role for updating concepts as $P11$ observed - "I don't want someone to mess with my concept. Maybe they can copy and personalize it if they want." In contrast, $P2$ wanted a more collaborative approach to the refinement of concepts if the data source was being shared within their team - "I can see my team wanting to update the same concept rather than duplicating efforts so that we are all on the same page." These contrasting viewpoints indicate a need for a thoughtful approach to managing updates and refinements to concepts based on scope and organization.

Recommendations: A prevalent challenge with NLIs is the need to provide relevant scaffolding and guidance to support users as they frame their NL utterances while exploring data [SHKC20,SBW21]. With LOKI we observed that users expected recommendations to suggest pertinent concepts that they could create and explore with the data sources. $P10$ expressed - "I wish there was some onboarding feature that could show me, hey, these are some concepts that others have saved. I could then use them or tweak them to what I wanted to do." Future work should explore techniques to support guided discovery learning through concept recommendations.

Leveraging large language models (LLMs): Due to their ease of use and their fluent text-generative capabilities, LLMs are garnering attention for both text generation and conversational capabilities [MEL*22]. Custom-trained GPT models could potentially be trained on data repositories and user queries to automatically identify and save custom concepts. In addition, future research should explore whether custom LLMs could provide additional semantics and organize the concepts as part of a concept map hierarchy.

Additional interaction modalities and interfaces: In this paper, we explore how NLIs, specifically in a Slack chatbot context, can be used to explore the creation of custom concepts for analytical reuse. Future research directions could consider additional visualization response types as well as modalities such as gestures, voice, and alternative chatbot platforms to assess how people express and use custom-created concepts in these new environments.

6. Conclusion

This paper presents LOKI, an analytical chatbot system for enabling users to save bespoke named concepts during their analytical workflows. We demonstrate how these concepts can be reused in subsequent utterances and concepts. Further, LOKI supports the persistence of these named concepts in the data source that they originated from and can be used in another visual analysis tool that imports the updated data source. We conducted a qualitative evaluation of LOKI that indicates the usefulness of supporting the reuse of custom concepts across a variety of visual analytics workflows. Feedback from interacting with LOKI identifies research opportunities concerning in-situ and dynamic curation of custom concepts during visual analytics workflows.

References

[ABK*07] AUER S., BIZER C., KOBILAROV G., LEHMANN J., CYGANIAK R., IVES Z.: Dbpedia: A nucleus for a web of open data. In *Proceedings of the 6th International The Semantic Web and 2nd Asian Conference on Asian Semantic Web Conference* (Berlin, Heidelberg, 2007), ISWC'07/ASWC'07, Springer-Verlag, p. 722–735. 2

[ask21] Tableau's Ask Data. https://www.tableau.com/products/new-features/ask-data, 2021. 1, 3

[Ban21] BANSAL S.: Netflix movies and tv shows. https://www.kaggle.com/shivamb/netflix-shows, Sep 2021. Accessed: 2023. 3

[BKOH17] BROWNE O., KRDZAVAC N. B., O'REILLY P., HUTCHINSON M. C.: Semantic ontologies and financial reporting: An application of the fibo. In *JOWO* (2017). 2

[BS94] BATRA D., SEIN M. K.: Improving conceptual database design through feedback. *International Journal of Human-Computer Studies 40*, 4 (1994), 653–676. URL: https://www.sciencedirect.com/science/article/pii/S1071581984710305, doi:https://doi.org/10.1006/ijhc.1994.1030. 2

[cov23] Covid-19 Dataset, 2023. CC-BY Dataset: https://covid19.ca.gov. 3

[CWS93] CHAN H. C., WEI K. K., SIAU K. L.: User-database interface: The effect of abstraction levels on query performance*. *MIS Q. 17*, 4 (dec 1993), 441–464. URL: https://doi.org/10.2307/249587, doi:10.2307/249587. 2

[DMN*17] DHAMDHERE K., MCCURLEY K. S., NAHMIAS R., SUNDARARAJAN M., YAN Q.: Analyza: Exploring data with conversation. In *UIST* (2017). 2

[GDA*15] GAO T., DONTCHEVA M., ADAR E., LIU Z., KARAHALIOS K. G.: DataTone: Managing ambiguity in natural language interfaces for data visualization. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)* (2015). 1, 2

[ibm21] IBM Watson Analytics. http://www.ibm.com/analytics/watson-analytics, 2021. 1, 2

[LdKLC13] LEE D., DE KEIZER N., LAU F., CORNET R.: Literature review of SNOMED CT use. *Journal of the American Medical Informatics Association 21*, e1 (07 2013), e11–e19. URL: https://doi.org/10.1136/amiajnl-2013-001636, arXiv:https://academic.oup.com/jamia/article-pdf/21/e1/e11/17375846/21-e1-e11.pdf, doi:10.1136/amiajnl-2013-001636. 2

[LEGÖ20] LEI C., EFTHYMIOU V., GEIS R., ÖZCAN F.: Expanding query answers on medical knowledge bases. In *EDBT* (2020). 2

[LHJY21] LIU C., HAN Y., JIANG R., YUAN X.: Advisor: Automatic visualization answer for natural-language question on tabular data. In *2021 IEEE 14th Pacific Visualization Symposium (PacificVis)* (2021), pp. 11–20. doi:10.1109/PacificVis52677.2021.00010. 2

[LLI*14] LEE S., LIEBER L., IVIE L., KIRBY J., HECK D., AYERS D.: *Avengers Epic Collection: Earth's Mightiest Heroes*. Marvel, 2014. URL: https://books.google.com/books?id=L7C6vwEACAAJ. 2

[LÖQ*18] LEI C., ÖZCAN F., QUAMAR A., MITTAL A. R., SEN J., SAHA D., SANKARANARAYANAN K.: Ontology-based natural language query interfaces for data exploration. *IEEE Data Eng. Bull. 41* (2018), 52–63. 2

[LTL*21] LUO Y., TANG N., LI G., TANG J., CHAI C., QIN X.: Natural language to visualization by neural machine translation. *IEEE Transactions on Visualization and Computer Graphics PP* (2021), 1–1. 2

[MEL*22] MEYER S., ELSWEILER D., LUDWIG B., FERNANDEZ-PICHEL M., LOSADA D. E.: Do we still need human assessors? prompt-based gpt-3 user simulation in conversational ai. In *Proceedings of the 4th Conference on Conversational User Interfaces* (New York, NY, USA, 2022), CUI '22, Association for Computing Machinery. URL: https://doi.org/10.1145/3543829.3544529, doi:10.1145/3543829.3544529. 4

[MRS08] MANNING C. D., RAGHAVAN P., SCHÜTZE H.: *Introduction to Information Retrieval*. Cambridge University Press, USA, 2008. 3

[OWE00] OWEI V.: Natural language querying of databases: an information extraction approach in the conceptual query language. *International Journal of Human-Computer Studies 53*, 4 (2000), 439–492. URL: https://www.sciencedirect.com/science/article/pii/S1071581999903816, doi:https://doi.org/10.1006/ijhc.1999.0381. 2

[pow21] Microsoft Q&A. https://powerbi.microsoft.com/en-us/documentation/powerbi-service-q-and-a/, 2021. 1, 2

[PTWS20] PELLISSIER TANON T., WEIKUM G., SUCHANEK F.: Yago 4: A reason-able knowledge base. In *The Semantic Web* (Cham, 2020), Harth A., Kirrane S., Ngonga Ngomo A.-C., Paulheim H., Rula A., Gentile A. L., Haase P., Cochez M., (Eds.), Springer International Publishing, pp. 583–596. 2

[ras21] Rasa: Open source conversational AI, Nov 2021. URL: https://www.rasa.com/. 3

[Sal23] Salesforce Customer Relationship Management (CRM) data, 2023. CC-BY Dataset: https://www.salesforce.com/resources/guides/crm-analytics-for-sales-service/. 3, 4

[SBT*16] SETLUR V., BATTERSBY S. E., TORY M., GOSSWEILER R., CHANG A. X.: Eviza: A natural language interface for visual analysis. In *ACM UIST* (2016). 1, 2

[SBW21] SETLUR V., BATTERSBY S., WONG T.: Geosneakpique: Visual autocompletion for geospatial queries. In *2021 IEEE Visualization Conference, IEEE VIS 2021 - Short Papers, New Orleans, LA, USA, October 24-29, 2021* (2021), IEEE, pp. 166–170. URL: https://doi.org/10.1109/VIS49827.2021.9623324, doi:10.1109/VIS49827.2021.9623324. 2, 4

[SHKC20] SETLUR V., HOQUE E., KIM D. H., CHANG A. X.: Sneak Pique: Exploring autocompletion as a data discovery scaffold for supporting visual analysis. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2020), Association for Computing Machinery, pp. 966–978. 2, 4

[SS18] SRINIVASAN A., STASKO J.: Orko: Facilitating multimodal interaction for visual exploration and analysis of networks. *IEEE Transactions on Visualization and Computer Graphics 24*, 1 (2018). 1, 3

[SS21] SRINIVASAN A., SETLUR V.: Snowy: Recommending utterances for conversational visual analysis. In *The 34th Annual ACM Symposium on User Interface Software and Technology* (New York, NY, USA, 2021), UIST '21, Association for Computing Machinery, p. 864–880. URL: https://doi.org/10.1145/3472749.3474792, doi:10.1145/3472749.3474792. 2

[ST22] SETLUR V., TORY M.: How do you converse with an analytical chatbot? revisiting Gricean maxims for designing analytical conversational behavior. In *In Proc. of of the SIGCHI Conference on Human Factors in Computing Systems* (2022). 2, 3

[STD19] SETLUR V., TORY M., DJALALI A.: Inferencing underspecified natural language utterances in visual analysis. In *IUI* (New York, NY, USA, 2019), IUI '19, ACM, p. 40–51. doi:10.1145/3301275.3302270. 2, 3

[sup23] Tableau Superstore, 2023. CC-BY Dataset: https://help.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-connect.htm. 3

[tho21] ThoughtSpot. http://www.thoughtspot.com, 2021. 1, 2

[TS19] TORY M., SETLUR V.: Do what I mean, not what I say! design considerations for supporting intent and context in analytical conversation. In *2019 IEEE Conference on Visual Analytics Science and Technology (VAST)* (2019), pp. 93–103. doi:10.1109/VAST47406.2019.8986918. 3

[VK14] VRANDEČIĆ D., KRÖTZSCH M.: Wikidata: A free collaborative knowledgebase. *Commun. ACM 57*, 10 (sep 2014), 78–85. URL: https://doi.org/10.1145/2629489, doi:10.1145/2629489. 2

Proceedings published by Eurographics - The European Association for Computer Graphics.