AnnoLens: Exploration and Annotation through Lens-Based Guidance

Franziska Becker* Steffen Koch† Tanja Blascheck‡

University of Stuttgart

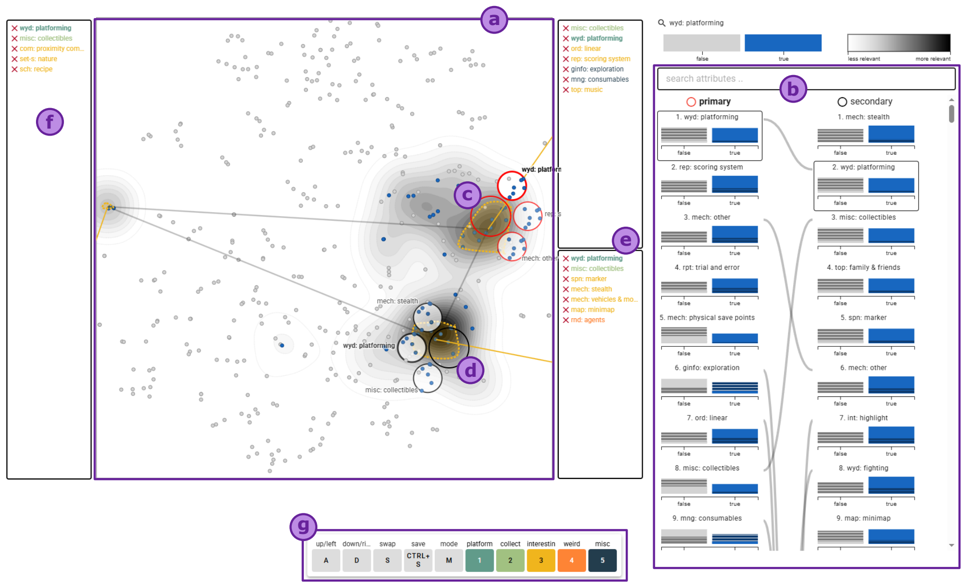

Figure 1: The ANNOLENS system, recreating the annotations made by participant P1 in our pilot study, showing: (a) a scatter plot and relevance score contours, (b) bar charts depicting summary data about attributes for each lens and linking identical attributes, (c) primary lens, (d) secondary lens, (e+f) annotations, and (g) the hotbar with multiple buttons for annotation and navigation.

ABSTRACT

Annotation is often a time-consuming but fruitful activity in data analysis contexts. The manual labor required to create useful annotations is a barrier that keeps users from documenting their analysis, especially intermediate results. To address the needs of exploration and annotation alike, we propose integrating annotation with lens-based interactions, combining both with guidance. We investigate the exploration-annotation requirement space, identifying challenges and extracting five design requirements for annotation in exploration contexts. Based on this investigation, we designed ANNOLENS—a concrete instantiation of such a system that lets users explore and annotate dimensionality-reduced multivariate data. It employs a dual-lens approach for contrastive exploration, using guidance to steer users toward interesting data subsets and attributes. Annotation is directly integrated into the lenses, letting users quickly annotate hunches and discoveries. Automated merging and linking serve to simplify annotation management and reduce disruptions. In a pilot study, we conducted a preliminary evaluation of our approach, which indicated that users find it easy to annotate data and were able to incorporate their knowledge and unique perspective into the process. A free copy of this paper and all supplemental materials are available at `https://osf.io/zpu6c/`.

Index Terms: Annotation, exploration, guidance, visual analytics.

*e-mail: franziska.becker@vis.uni-stuttgart.de
†e-mail: steffen.koch@vis.uni-stuttgart.de
‡e-mail: tanja.blascheck@vis.uni-stuttgart.de

1 INTRODUCTION

Data exploration is an integral part of many visual analytics solutions. Documenting insights and rationales can help users recall their analysis process, present their findings, and collaborate with others [23, 28, 37]. While exploration aims to help users uncover insight, externalizing these insights can be costly work that is often performed post-hoc. How low-cost interactions that let users externalize quickly can be integrated into visual analytics solutions and how they affect the sensemaking process is an active area of research [2, 22, 30, 24]. In this article, we take another step towards integrating annotation into data exploration with the help of guidance, making three contributions: First, we discuss the exploration-annotation requirement space, looking at challenges and diverging needs based on prior work. Second, we present ANNOLENS, a concrete system that combines guidance with interactive lenses to facilitate rapid exploration and annotation for dimensionality-reduced multivariate data (cf. Figure 1). In contrast to highly automated or completely manual approaches, ANNOLENS guides users towards interesting features or data subsets, but leaves the decision to annotate up to them. It prioritizes quick interactions, allowing users to create unique annotation maps that reflect their individual viewpoint as it evolves during exploration. Third, we reflect on observations regarding workflow differences and annotation outcomes from a pilot study with three participants which can inform future work.

2 RELATED WORK

Annotations in visualization exist in the form of highly structured data [11], free-form sketches [30], or storytelling devices [29]. They can serve to document insights [11, 14, 27, 31], communicate a narrative [29, 33], or label data for machine learning [6]. Annotations can be seen as a part of provenance, although provenance often exceeds the goals that annotation pursues. Provenance can include recording and analyzing interaction logs, data transformations, or visualization states for user modeling, replication, steering models or evaluating systems [28, 37]. In fewer cases, annotations directly aim to inform the sensemaking process—both functionally and visually—through tight integration with other system components. Click2Annotate [11] lets users annotate findings using different templates that describe different data patterns, but disrupts the data exploration and is less suitable for intermediate results and ideas that are still subject to change. Zhao et al. [38] designed ANNOTATIONGRAPHS, an interactive system that supports the documentation of insights through user-authored annotations, emphasizing extensive and deliberate annotation. The system combines data analysis and annotation, arranging annotations as a graph laid out through a mixed-initiative approach, though experts felt the process of creating annotations was generally time-consuming. Kandogan [19] introduced "just-in-time descriptive analytics," providing automated labels for patterns in scatter plots based on user interactions. They help users to find and understand patterns in a data-driven manner, but do not allow users to create their own annotations.

Exploration, in a visualization context, is "a hypothesis-generation process" [20] that aims to build "rich mental models of the data" [5] and involves "identifying questions of interest, inspecting visualized data, and iteratively refining one's questions and hypotheses" [3]. It is a process that often starts without clear tasks, during which questions and hypotheses arise by finding *interesting* bits and pieces in the data. There are many techniques that support exploration in interactive visual systems, including magic lenses. Lenses let users interact with *data neighborhoods*, supporting *rapid exploration* of that data. Tominski et al. [35] define the conceptual lens model as a pipeline that is attached to the standard visualization pipeline via selection and transforms the data to generate a lens effect. A lens can apply different functions that transform or enrich the underlying data: showing details about the data [7] or changing the layout [34]. By employing multiple lenses [9, 21], users can directly compare specific subsets of the data. Looking at the overlap between lenses and annotation, Bettio et al. [8] use annotations to guide navigation of interactive lenses, focusing on annotations that produce visual overlays. Similarly, Ahsan et al. [1] use audio-visual annotation graphs to guide exploration in a cultural heritage context. These works showcase how annotation data can serve as a basis for guidance when combined with lens-based exploration, but these annotations are static and provided by domain experts beforehand. In contrast, we consider annotations created by users *during* exploration that enable insight externalization, producing annotation maps that reflect users' current mental model of the data, based on their unique knowledge and perspective.

3 REQUIREMENTS & DESIGN

Annotation can be a critical activity for analysis and reporting, but it is often time-consuming. In this section, we introduce the exploration-annotation-requirement space, identifying design challenges and opportunities. Based on this requirement space, we detail our design decisions for ANNOLENS.

3.1 Requirements for Exploration-Annotation

Exploration and annotation are somewhat orthogonal activities during sensemaking processes [25]. While exploration primarily takes place during information foraging, annotating and recording findings are important for supporting later stages when users create schemas, build hypotheses, and report their findings. Data exploration is the process of building understanding about a data space by traversing it in some fashion. It starts out without clear tasks with users wanting to find patterns and understand relationships in the data [5, 17, 26]. Over the course of exploration, users can form questions and hypotheses which typically change their information needs [15]. Given this framing, facilitating exploration includes making this data space traversal more effective. Effective in this context means quickly finding interesting subsets or dimensions that produce insight or point analysis into a more concrete direction. Because finding these interesting subsets is not easy, approaches for exploratory analysis often help users to query and view the data from different angles. Related interactions must consequently enable users to efficiently formulate queries that shed more light on data characteristics. When viewing annotation as a documentation activity, it is more deliberate and slow, as users are trying to mold their found insights into a more structured representation.

If we consider annotation as an activity that supports *sensemaking*—where annotations represent not final insights but instead insights as they emerge and change—then deliberate and slow interactions are at odds with exploration requirements. For such a scenario, annotation interactions should be **quick and easy** (R_1), to not disrupt the flow of exploration. Similarly, **seamless** (R_2) transitions between modes of exploration and annotation prevent users from excessive configuration and mode-switching, which can increase cognitive load and lead to errors [32]. To document from which data the insights were derived, annotations should **contextually link** (R_3) to their data source [22, 28, 30]. Additionally, because they should mirror the user's changing mental model of the data, they must be easy to **add, delete, and edit** (R_4). As annotations accumulate, it becomes more challenging to **maintain clarity** (R_5), particularly when the visualizations themselves are already visually complex. From these requirements derive a multitude of interconnected design decision concerning the representation and function of annotations in systems for exploratory data analysis: When should which annotations be visible? Is it important that annotations are placed in proximity to their context, the annotated data? How should annotations influence each other; how do they influence other system features? To answer these questions, designers must investigate how users want to use annotations and which information they need at which point in the workflow.

Subsequently, we present a system that combines lens-based interaction, guidance, and annotation as complementing parts that let users externalize and integrate insights into the exploration process. Our approach can be described using both the knowledge-assisted visual analytics conceptual model [12] and the model for guidance in visual analytics [10]. In the former, our guidance is part of the *analysis* that produces *visualizations* and alters the *specification*, whereas annotations externalize *tacit knowledge* for consumption by the system and user. For the guidance model, our approach mediates guidance both through the lens interactions and annotations.

3.2 AnnoLens Design

With ANNOLENS, our goal was to create a system that lets users simultaneously explore and annotate data, thereby reflecting their unique perspective and treating annotations not as an addendum or chore, but as an activity that supports the sensemaking process. We focus on exploration of dimensionality-reduced multivariate data, which can be found in many domains for different types of data. Based on the challenges and requirements identified previously, we describe our system design and how it connects exploration and annotation activities with lenses and guidance. The code for ANNOLENS is available online at https://github.com/ArielMant0/anno-lens, with a demo hosted at https://arielmant0.github.io/anno-lens/.

3.2.1 Supporting Exploration

To support data exploration we use lenses and guidance to tackle three related questions: (i) which data attributes are relevant, (ii) where similar data can be found, and (iii) how these data subsets compare to each other. Lenses allow for quick, continuous traversal of a data space by simply moving the cursor. They are a powerful and intuitive approach to supporting exploration for meaningfully-arranged data, like in scatter plots [16], geographic maps [21], and graphs [7, 34]. Simultaneously, lenses let users easily select neighborhoods that are assumed to have some meaningful reason for their spatial proximity—which we can use to identify the data that annotations should be associated with. A scatter plot (cf. Figure 1 a) combined with a lens is our primary means around which exploration is structured. ANNOLENS employs a dual-lens concept, where the primary lens is connected to guidance in multiple ways: First, the data under the primary lens is used to calculate *relevance scores* for each data attribute. Second, the highest-ranked attribute is used to color the dots in the plot according to their value. Third, the secondary lens is placed at another plot location with the most similar relevance score for the same attribute, if one exists. The relevance score calculation depends on the attribute type, but always tries to model the discrepancy between the attribute's global distribution and the lens data's distribution. Attributes that are prevalent under the lens but rare globally get a higher score than attributes prevalent in both. Concrete calculations and more examples are available in the supplemental material. Background contours in the plot visualize the relevance scores for pre-computed lens positions, giving further visual guidance. After being placed at its initial position, users can move the secondary lens to another location on demand. Next to each lens, smaller *lens copies* (cf. Figure 1 c & d) depict the same data colored according to other highly-ranked attributes. They give users immediate visual feedback on the connection between different attributes, enabling direct comparison. Next to the scatter plot, small histograms of all data attributes are shown for each lens, sorted by their relevance scores (cf. Figure 1 b). Histograms for the same attribute are connected via links, indicating how relevance scores compare between lenses.

3.2.2 Supporting Annotations

Annotating in ANNOLENS is tightly connected to the lenses and consequently facilitated by the guidance that is combined with these lenses. We define annotations as structured tuples that consists of data ids, attribute labels, and optionally an attribute value. This allows us to implement quick, single-click annotations (R_1, R_2, R_4) that are directly connected to the lens-based exploration (R_3). Users can choose from a set of single-click interactions, which creates an annotation for the currently active combination of lens and attribute. We chose to provide alternative ways for triggering the same action to accommodate different user preferences, like for interaction pacing [13]. A hotbar placed at the bottom of the window provides five colored buttons that create an annotation in the respective color (cf. Figure 1 g). The button can be clicked with the mouse, but it also displays its associated key that can be pressed to annotate. Clicking on an attribute label next to a lens copy also annotates that attribute for the related lens data. The same principle is applied to switching the active attribute, for which we provide buttons, hotkeys, and interactions at visualization elements. By giving users a set of different colors with customizable labels to annotate with, they can create categories to which they can assign their own meaning. Thereby, we aim to alleviate the limited expressive freedom users have for annotations, which are restricted to attribute labels and values. During the design phase, we examined various annotation representations and levels of detail. Initially, we considered showing labels inside the plot, but abandoned this approach due to the potential for clutter (R_5), especially given our lens design that already adds multiple visual elements to the plot. Implementing different levels of detail can reduce clutter, but requires the system to decide which information each level shows. Ultimately, we chose to place annotations on the left and right sides of the scatter plot (cf. Figure 1 e+f), dividing the available space into non-overlapping areas based on the number of annotations. This trades less visual clutter (R_5) for reduced scalability, but allows us to constantly depict all annotations, providing instant access (R_3). Annotations are integrated into the system in several ways to let them act as more than notes. Whenever the currently active attribute is contained in an annotation, the respective data points and label are visually highlighted (R_3). These highlights act as guidance cues [10] that remind users of annotations they already made and which data that concerns. To show connections between annotations, ANNOLENS automatically links annotations based on their overlap in labels. These links are visualized in the scatter plot when the lens rests, encoding the amount of overlap via line thickness (cf. Figure 1). We chose to show these links only on demand to reduce clutter (R_5) and because we suspect that users care about these links when they want to reason about their annotations—not when moving the lens to explore. Lastly, ANNOLENS merges annotations with an overlap in data points and the same labels, thereby reducing unnecessary manual labor (R_1, R_2).

4 PILOT STUDY

In a pilot study with three participants (2 male, 1 female, all aged 25 to 30), we conducted an initial examination of ANNOLENS. Specifically, we wanted to get an impression of annotation workflows, resulting annotations, and usability hurdles. We recruited participants from our own and other institutes based on their experience with the dataset's domain and visualization. Participant P1 is familiar with the study dataset, P2 works in visualization and knows the domain but not the dataset, and P3 is not familiar with the dataset, domain, or visualization.

4.1 Study Datasets

The study used two different datasets, one for the training round and one for the study task. For each dataset, we calculated a 2D embedding using t-SNE [36] that is visualized in a scatter plot. During the system demonstration and training phase, participants worked with a dataset that consists of 80 different types of cereals, described through 16 attributes [18]. This dataset was chosen because it contains different data types and its attributes require little expertise to understand. The study task used a dataset of video games collaboratively labeled by three coders [4]. It has 388 data points and 256 attributes, which describe characteristics of a game in a binary `yes` or `no` fashion. We chose this dataset for four distinct reasons: First, it contains complex patterns that the authors are already familiar with. Second, it reflects characteristics of real-world datasets such as highly correlated attributes and semantic inconsistencies. Third, this data has attributes that are subjective by design, leaving room for interpretation that participants can fill with their pre-existing knowledge. Fourth, it is likely that we can find participants with different levels of knowledge for the data domain, which likely impacts the annotations they will create.

4.2 Study Procedure

At the start, the experimenter explained the study and demonstrated system capabilities using the Cereal dataset. Afterwards, participants were asked to perform a number of interactions and answer basic comprehension questions as training. Then the main study task began. We asked participants to imagine working in a game development context where the experimenter is the superior, providing the data and task. Participants were instructed to extract information about clusters of platform games, a subset of skill-based movement games, to find differences and similarities between them, which they should present afterwards. This task design aimed to imitate more realistic data exploration scenarios, in which users have some level of pre-existing knowledge and a high-level goal. Entirely free exploration may overwhelm users or lead them to ignore system functionalities. Participants had 20 minutes to complete the task and 5 minutes to present their results. We asked participants to think aloud during the task and ask questions about system features if they were unsure about them. At the end, we asked six questions, described in the supplemental material, regarding the use of annotations and general system utility.

4.3 Results

This section describes results from our pilot study, starting with task performance followed by observations concerning differences in user workflows and resulting annotations.

4.3.1 Task Results

All participants solved the study task to a satisfactory degree, finding the three target clusters and describing at least two important characteristics and differences. P1 was the most thorough, finding more details than the other two, but taking longer to find the initial attribute that correlates with platform games. In the reporting phase, all participants found the same high-level characteristics, but diverged for the specifics, informed by their unique perspective.

4.3.2 Workflow Differences

Looking at the workflows of participants, we observed both similarities and differences between them. For similarities, we saw that all of them were a little reluctant to annotate directly, preferring to explore for some time before starting to annotate. This may be because they wanted to get an overview of the dataset before creating annotations, even if these are easily deleted. Another reason may be that the system complexity was too overwhelming at the beginning, requiring some initial warm-up. When asked how easy or difficult it was to create annotations, all participants said they felt it was easy as soon as they understood the concept and had an idea of what to annotate. Although all participants started the task in a similar manner—getting to know the dataset and finding the platforming attribute—their approaches diverged after that initial phase. P1 did not move the lenses too much, instead preferring to get an in-depth understanding of all the related attributes for each group. To inspect attributes, P1 often used the small lenses and hotkeys, whereas P2 and P3 preferred looking at the histograms, only changing the active attributes a few times to inspect distributions in the scatter plot. Participants only deleted annotations when they made a mistake, preferring to add annotations once they felt reasonably sure that their understanding was correct. All participants noted the relevance contours as a positive component that helped them in getting a quick overview and finding other interesting spots easily. Similarly, everyone positively mentioned having two lenses to make direct comparisons between data subsets.

4.3.3 Annotation Differences

P1 treated annotation colors as different categories that signify how the respective labels relate to each other and the dataset as a whole. For example, labels that clusters have in common were marked as dark green, whereas labels unique to a cluster were yellow. This usage of the colors conforms with how we conceptualized it. In contrast, P2 assigned a color to each cluster, giving them a name he thought fitting, while P3 always used the same color. Looking at the number of annotations, P1 annotated 19 labels distributed over the 3 clusters, P2 annotated only 3 labels, and P3 annotated 10 labels for the three clusters. P1 rigorously inspected and compared all attributes for all clusters, thereby discovering and annotating more details about cluster differences than other participants. In the discussion, P2 said that he felt that his previous knowledge allowed him to only annotate a few labels as a reminder. Participant P3, constrained her annotations to the highest ranking attributes due to her limited knowledge about the domain.

5 Discussion

We based our design requirements on the needs of exploration processes, prioritizing quick interactions for annotation such that users may easily annotate without interrupting their exploration needlessly. The pilot study showed that all participants could find relevant information and annotate it in a way that suited their preference—using different interactions to annotate with different levels of detail. The study task was more focused than completely free exploration, to give users some guidance as to what they could be looking for and to reflect realistic scenarios where data exploration is subject to high-level goals. However, this also means that annotation behaviors might change for other task scenarios that are more open or more focused. Participants in the study agreed that AnnoLens made it easy to annotate, although they still started the task by exploring for some time before actually annotating anything. This may indicate that even low-cost annotation is only engaged with after an initial familiarization phase that lets users form an initial *schema* of the data. Integrating guidance, via attribute sorting and relevance contours, was unanimously appreciated by all study participants. Participants generally agreed with the guidance on which attributes are relevant, but also annotated according to their individual perspective, resulting in different annotation maps. The annotations users create offer an opportunity to expand guidance further, such as suggesting attribute labels based on their occurrence in already annotated data. All participants had some initial difficulty in understanding the different system features. They struggled a bit with navigating between attributes and understanding how they could make use of the different annotation colors. AnnoLens trades expressive power of annotations for speed; users can only annotate attribute labels and values in different colors. While the colors give users some way to encode meaning, two participants mentioned wanting to also take free-form notes, which clashes with the need for quick interactions. Whether that poses an undesirable disruption of the exploration process or is instead an opportunity for users to structure their thoughts is an open question for future work. Our work is a small step towards expanding our understanding of annotation in visual analytics system. However, larger studies are needed to test whether approaches like ours can be used in a generalized way and to understand how they can best support users.

6 Conclusion

AnnoLens combines lens-based guidance with annotation to support externalization without interrupting the exploration process. By integrating lens-based interactions directly with annotation interactions, it makes annotating quick and less disruptive. Our approach constitutes a step towards introducing fast and semi-automated annotations for interactive exploration. In a small pilot study, participants created unique annotations that reflect their knowledge and unique approach. Every participant found it easy to annotate, though usability issues and system complexity occasionally posed problems. Guidance was viewed positively by all participants and could be extended further by incorporating existing annotations for suggestions or visual cues. Extending the annotation mechanism with foundation models could be an interesting option to make annotation more powerful and personalized. Large-language models (LLMs) could take over the burden of creating more detailed reports from annotations, letting users focus on exploration. They could also try to describe what the user is looking for based on the created annotations, thereby acting as a mirror that could help users reflect on their results so far. AnnoLens is a first venture into the exploration-annotation space, opening up avenues for improvement and extension in future work.

Acknowledgments

Tanja Blascheck is funded by the European Social Fund and the Ministry of Science, Research and Arts Baden-Württemberg. We want to thank René Warnking for his continued feedback and support during this project.

REFERENCES

[1] M. Ahsan, F. Marton, R. Pintus, and E. Gobbetti. Audio–visual annotation graphs for guiding lens-based scene exploration. *Computers & Graphics*, 105:131–145, 2022. doi: 10.1016/j.cag.2022.05.003 2

[2] S. K. Badam, S. Chandrasegaran, and N. Elmqvist. Integrating annotations into multidimensional visual dashboards. *Information Visualization*, 21(3):270–284, 2022. doi: 10.1177/14738716221079591 1

[3] L. Battle and J. Heer. Characterizing Exploratory Visual Analysis: A Literature Review and Evaluation of Analytic Provenance in Tableau. *CGF*, 38(3):145–159, 2019. doi: 10.1111/cgf.13678 2

[4] F. Becker, R. P. Warnking, H. Brückler, and T. Blascheck. Beyond Entertainment: An Investigation of Externalization Design in Video Games. *CGF*, 2025. 3

[5] J. T. Behrens. Principles and procedures of exploratory data analysis. *Psychological Methods*, 2(2):131–160, 1997. doi: 10.1037/1082-989X.2.2.131 2

[6] J. Bernard, M. Zeppelzauer, M. Sedlmair, and W. Aigner. VIAL: A Unified Process for Visual Interactive Labeling. *The Visual Computer*, 34(9):1189–1207, 2018. doi: 10.1007/s00371-018-1500-3 1

[7] E. Bertini, M. Rigamonti, and D. Lalanne. Extended excentric labeling. *CGF*, 28(3):927–934, 2009. doi: 10.1111/j.1467-8659.2009.01456.x 2, 3

[8] F. Bettio, M. Ahsan, F. Marton, and E. Gobbetti. A novel approach for exploring annotated data with interactive lenses. *CGF*, 40(3):387–398, 2021. doi: 10.1111/cgf.14315 2

[9] S. Butscher, K. Hornbæk, and H. Reiterer. Spacefold and physiclenses: simultaneous multifocus navigation on touch surfaces. In *AVI '14: Proc. 2014 Int. Work. Conf. Adv. Vis. Interf.* ACM, 2014. doi: 10.1145/2598153.2598177 2

[10] D. Ceneda, T. Gschwandtner, T. May, S. Miksch, H.-J. Schulz, M. Streit, and C. Tominski. Characterizing guidance in visual analytics. *IEEE Trans. Vis. Comp. Graph.*, 23(1):111–120, 2017. doi: 10.1109/TVCG.2016.2598468 2, 3

[11] Y. Chen, S. Barlowe, and J. Yang. Click2annotate: Automated insight externalization with rich semantics. In A. M. MacEachren, ed., *Symp. Vis. Analyt. Sci. Tech (VAST)*, pp. 155–162. IEEE, 2010. doi: 10.1109/VAST.2010.5652885 1, 2

[12] P. Federico, M. Wagner, A. Rind, A. Amor-Amoros, S. Miksch, and W. Aigner. The role of explicit knowledge: A conceptual model of knowledge-assisted visual analytics. In B. Fisher, S. Liu, and T. Schreck, eds., *Conf. Vis. Analyt. Sci. Tech. (VAST)*, pp. 92–103. IEEE, 2017. doi: 10.1109/VAST.2017.8585498 2

[13] A. Goguey, C. Gutwin, Z. Chen, P. Suwanaposee, and A. Cockburn. Interaction pace and user preferences. In *Proc. 2021 CHI Conf. Hum. Fact. Comp. Sys.*, pp. 1–14. ACM, 2021. doi: 10.1145/3411764.3445772 3

[14] D. P. Groth and K. Streefkerk. Provenance and annotation for visual exploration systems. *IEEE Trans. Vis. Comp. Graph.*, 12(6):1500–1510, 2006. doi: 10.1109/TVCG.2006.101 1

[15] M. Hearst. *Search User Interfaces*. Cambridge University Press, 2009. 2

[16] F. Heimerl, M. John, Qi Han, S. Koch, and T. Ertl. DocuCompass: Effective Exploration of Document Landscapes. In *Conf. Vis. Analyt. Sci. Tech. (VAST*, pp. 11–20. IEEE, 2016. doi: 10.1109/VAST.2016.7883507 3

[17] S. Idreos, O. Papaemmanouil, and S. Chaudhuri. Overview of data exploration techniques. In T. Sellis, S. B. Davidson, and Z. Ives, eds., *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*, pp. 277–281. ACM, New York, NY, USA, 2015. doi: 10.1145/2723372.2731084 2

[18] P. Isenberg, P. Dragicevic, and Y. Jansen. 80 cereals, 2018. https://www.kaggle.com/datasets/crawford/80-cereals (last accessed on 29. April 2025). 3

[19] E. Kandogan. Just-in-time annotation of clusters, outliers, and trends in point-based data visualizations. In G. Santucci, ed., *Conf. Vis. Analyt. Sci. Tech. (VAST)*, pp. 73–82. IEEE, 2012. doi: 10.1109/VAST.2012.6400487 2

[20] D. A. Keim. Visual exploration of large data sets. *Communications of the ACM*, 44(8):38–44, 2001. doi: 10.1145/381641.381656 2

[21] R. Krüger, D. Thom, M. Wörner, H. Bosch, and T. Ertl. Trajectorylenses—a set-based filtering and exploration technique for long-term trajectory data. *CGF*, 32(3pt4):451–460, 2013. doi: 10.1111/cgf.12132 2, 3

[22] H. Lin, D. Akbaba, M. Meyer, and A. Lex. Data hunches: Incorporating personal knowledge into visualizations. *IEEE Trans. Vis. Comp. Graph.*, 29(1):504–514, 2023. doi: 10.1109/TVCG.2022.3209451 1, 2

[23] N. Mahyar, A. Sarvghad, M. Tory, and T. Weeres. Observations of record-keeping in co-located collaborative analysis. In R. H. Sprague, ed., *46th Hawaii International Conference on System Sciences (HICSS), 2013*, pp. 460–469. IEEE, Piscataway, NJ, 2013. doi: 10.1109/HICSS.2013.420 1

[24] A. Mathisen, T. Horak, C. N. Klokmose, K. Grønbæk, and N. Elmqvist. InsideInsights: Integrating Data–Driven Reporting in Collaborative Visual Analytics. *Computer Graphics Forum*, 38(3):649–661, 2019. doi: 10.1111/cgf.13717 1

[25] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proc. Int. Conf. Intel. Analy.*, vol. 5, pp. 2–4, 2005. 2

[26] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of International Conference on Intelligence Analysis*, vol. 5, pp. 2–4, 2005. 2

[27] P. G. Poličar and B. Zupan. VERA: Generating Visual Explanations of Two-Dimensional Embeddings via Region Annotation. 1

[28] E. D. Ragan, A. Endert, J. Sanyal, and J. Chen. Characterizing provenance in visualization and data analysis: An organizational framework of provenance types and purposes. *IEEE Trans. Vis. Comp. Graph.*, 22(1):31–40, 2016. doi: 10.1109/TVCG.2015.2467551 1, 2

[29] D. Ren, M. Brehmer, B. Lee, T. Hollerer, and E. K. Choe. Chartaccent: Annotation for data-driven storytelling. In D. Weiskopf, Y. Wu, and T. Dwyer, eds., *2017 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 230–239. IEEE, 2017. doi: 10.1109/PACIFICVIS.2017.8031599 1

[30] H. Romat, N. Henry Riche, K. Hinckley, B. Lee, C. Appert, E. Pietriga, and C. Collins. ActiveInk: (Th)Inking with Data. In *Proc. 2019 CHI Conf. Hum. Fact. Comp. Sys.*, CHI '19, pp. 1–13. ACM, 2019. doi: 10.1145/3290605.3300272 1, 2

[31] A. Ruangrotsakun, D. Oh, T.-V. Nguyen, K. Lee, M. Ser, A. Hiew, R. Ngo, Z. Shureih, R. Khanna, and M. Kahng. Viva: Visual exploration and analysis of videos with interactive annotation. In *Comp. Proc. 28th Int.l Conf. Intel. User Interf.*, ACM Digital Library, pp. 162–165. Association for Computing Machinery, New York,NY,United States, 2023. doi: 10.1145/3581754.3584160 1

[32] N. B. Sarter and D. D. Woods. How in the World Did We Ever Get into That Mode? Mode Error and Awareness in Supervisory Control. *Human Factors*, 37(1):5–19, 1995. doi: 10.1518/001872095779049516 2

[33] D. Shi, A. Oulasvirta, T. Weinkauf, and N. Cao. Understanding and automating graphical annotations on animated scatterplots. In J. Nonaka, ed., *2024 IEEE 17th Pacific Visualization Conference*, pp. 212–221. IEEE, 2024. doi: 10.1109/PacificVis60374.2024.00031 1

[34] C. Tominski, J. Abello, and H. Schumann. Cgv—an interactive graph visualization system. *Computers & Graphics*, 33(6):660–678, 2009. doi: 10.1016/j.cag.2009.06.002 2, 3

[35] C. Tominski, S. Gladisch, U. Kister, R. Dachselt, and H. Schumann. Interactive lenses for visualization: An extended survey. *CGF*, 36(6):173–200, 2017. doi: 10.1111/cgf.12871 2

[36] L. Van der Maaten and G. Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(11), 2008. 3

[37] K. Xu, A. Ottley, C. Walchshofer, M. Streit, R. Chang, and J. Wenskovitch. Survey on the analysis of user interactions and visualization provenance. *CGF*, 39(3):757–783, 2020. doi: 10.1111/cgf.14035 1, 2

[38] J. Zhao, M. Glueck, S. Breslav, F. Chevalier, and A. Khan. Annotation graphs: A graph-based visualization for meta-analysis of data based on user-authored annotations. *IEEE Trans. Vis. Comp. Graph.*, 23(1):261–270, 2017. doi: 10.1109/TVCG.2016.2598543 2