Detecting Changes in User Behavior to Understand Interaction Provenance during Visual Data Analysis

Alyssa Peña
alyspena@gmail.com
Texas A&M University

Ehsanul Haque Nirjhar
nirjhar71@tamu.edu
Texas A&M University

Andrew Pachuilo
Texas A&M University

Theodora Chaspari
chaspari@tamu.edu
Texas A&M University

Eric D. Ragan
eragan@ufl.edu
University of Florida

ABSTRACT

Analysts can make better informed decisions with "Explainable AI". Visualizations are often used to understand, diagnose, and refine AI models. Yet, it is unclear what type of interactions are appropriate for a given model and how the visualizations are perceived. Furthering research into sensemaking for visual analytics may be useful in understanding how user's are interacting with visualizations for AI and in developing a naturalistic model of explanation. Conventional approaches consist of human experts applying theoretical sensemaking models to identify changes in information processing or utilizing recorded rationale provided by the users. However, these approaches can be inefficient and inaccurate since they heavily rely on subjective human reports. In this research, we aim to understand how data-driven techniques can automatically identify changes in user behavior (inflection points) based on user interaction logs collected from eye tracking and mouse interactions. We relay the results of a supervised classification system using Hidden Markov Models to predict changes in a visual data analysis of a cyber security scenario. Preliminary results indicate a 70% accuracy in identifying inflection points. These preliminary results suggest the feasibility of data-driven approaches furthering our understanding of sensemaking processes and interaction provenance.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing theory, concepts and paradigms**; **Visual analytics**; • **Computing methodologies** → *Knowledge representation and reasoning*; *Machine learning approaches*.

KEYWORDS

Visual Analytics, Modeling User Behavior, Explainable AI, Sensemaking

ACM Reference Format:
Alyssa Peña, Ehsanul Haque Nirjhar, Andrew Pachuilo, Theodora Chaspari, and Eric D. Ragan. 2019. Detecting Changes in User Behavior to Understand Interaction Provenance during Visual Data Analysis. In *Joint Proceedings of the ACM IUI 2019 Workshops, Los Angeles, USA, March 20, 2019.* , 6 pages.

IUI Workshops'19, March 20, 2019, Los Angeles, USA

© 2019 Copyright ©2019 for the individual papers by the papers' authors. Copying permitted for private and academic purposes. This volume is published and copyrighted by its editors.

1 INTRODUCTION

The effectiveness of AI applications is limited by the disconnect between experts knowledgeable in AI and subject matter experts. Using "Explainable AI" (XAI), analysts can make better informed decisions [17]. In particular, visualizations are used to understand, diagnose, and refine AI models. However, it is unclear what type of interactions are appropriate for a given model and how the visualization is perceived [23]. Sensemaking can aid in the development of a naturalistic model of explanation and visualization tools [17, 19]. In visual analytics, understanding the user's sensemaking process comprises an important milestone for improving tools, supporting collaboration among analysts, and training new workers [30]. Furthering existing research in sensemaking for visual analytics may in turn further research in XAI.

Due to the complexity and ad-hoc nature of sensemaking during exploratory data analysis, it is difficult to comprehensively characterize sensemaking using solely theoretical models without expertise and assumptions to map analyst behavior to the model [20, 29, 32]. Capturing the provenance of sensemaking often requires explicit feedback or annotations from the analyst. Researchers often investigate "thinkaloud protocols" [11] to study analytic processes by asking analysts to verbally describe what they are doing or thinking over time (e.g., [31]). Unfortunately, collecting explicit user feedback—whether from annotations or verbal comments—can be distracting or burdensome for the analyst. However, research has demonstrated the benefits of using interaction behavior to learn about the analysis and sensemaking processes (e.g., [5, 8, 26]). In this way, information about sensemaking can be inferred automatically without requiring explicit user feedback.

In our research, we explore summarizing analysis activity from collected visual analytics interaction logs. Our post-hoc approach focuses on techniques to summarize different periods of analysis time based on changes in the analyst's goals and behaviors, referred to as inflection points. Inflection points can be indicative of changes in a analyst's thought process, but qualitatively identifying these changes can be a laborious process. Finding alternative ways to detect inflection points could help quantitatively model such behaviors and provide additional insights to analytic provenance. The problem is formulated as a supervised classification task, according to which the goal is to determine the presence or absence of inflection points from a time-sequence of eye-tracking and mouse interaction logs over a pre-determined length of time. We capture interaction logs and think-aloud comments from human participants during a data analysis activity with a sample visual analytics

CEUR-WS.org/Vol-2327/IUI19WS-UIBK-3.pdf

application. We employ a simple heuristic rationale to label inflection points during the analysis session and extract features from the interaction logs. Two supervised Hidden Markov Models were trained with the features and labels to predict inflection points.

We present this research as preliminary work from an analysis of sample records collected from a seven-participant user study. We share our results on identifying changes in analyst behavior along with insights learned regarding the challenges of using features from different interaction logs to understand sequence segments. The results of our study can be used in future research aimed towards explaining the training/model building process of AI.

2 RELATED LITERATURE

Research of analytic provenance in visual analytics broadly consists of integrated workflow management, post processing of captured interactions, and visualization of captured tool states and/or interactions using a combination of analytics and sensemaking theory [30]. Some visual analytics tools utilize node-based workflows by allowing users access to a view showing previous states and operations used to manipulate the data flow (e.g., [6, 9]). Alternatively, the overall visual analysis task can be composed of subtasks derived from user interactions [4, 15]. Dou et al. [8] compared human-coded interpretations of interaction logs to their ground truth of think-a-loud transcriptions based on analysis findings, strategies, and methods. While these results were based on human interpretation of the reasoning process, they showed that there was a correlation between interactions and cognitive reasoning. Research has also been made in summarizing the analysis task through meta data visualization using topic modelling for breaks in time [22, 25]. However, there is debate in determining parameters for segmentation and deciding which interactions are most meaningful to indicate changes in the analysis process.

Given feature rich information from interaction logs, it is possible to use machine learning techniques to gain insights on the user's sensemaking process. Work by Harrison et al. [18] employed interaction features such as mouse clicks in conjunction with Hidden Markov Models (HMMs) to predict transitions in user frustration. Research also shows that mouse and interface interactions can be used to predict task performance and infer user personality traits [5]. Kodagoda et al. [21] compared machine learning techniques to infer reasoning provenance. However, they relied on using experts to code user's interactions as processes based on Klein's data frame theory of sensemaking [20] and were not successful in identifying all processes. Similarly, Aboufoul et. al explore using HMMs to determine the cognitive states of users by mapping hidden states to sensemaking processes defined by Piroli et. al. to validate their models [1].

Recent techniques introduce objective signal-based indices of eye-tracking and mouse log data that provide an alternative view of the interaction between the user and the visualization. The work by Coltekin et al. [7] provide a deeper understanding of how people make inferences and decisions when using visualizations tools and complex displays by comparing a hypothetical sequence of strategies to recorded sequences of participants. Eye tracking-derived metrics can also be highly beneficial for understanding visualization usage and analysis behavior. Eye tracking metrics often highlight the importance of *areas of interests* (AOIs) by including factors such as number of fixations within areas of the display and sequential transitions among areas [2, 3, 13]. AOIs are specified regions of an interface such as a particular view, button, or piece of content.

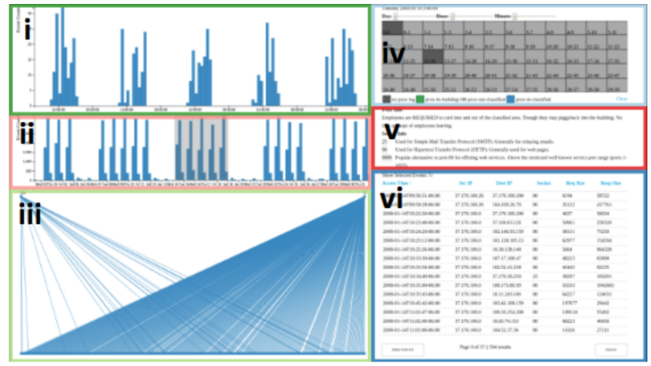

Figure 1: The visual analytics tool with six areas of interest: Detailed Histogram (i), Overview Histogram (ii), Network Graph (iii), Offices (iv), Info (v), Table (vi).

In our own research, instead of employing elaborate coding schemes that might be time-consuming and hard to replicate, we relied solely on inflection points observed when a user changed tactics in tool usage as labels. We included eye tracking data in addition to mouse interaction features and tested for which features produced the highest model precision. The extracted features required less dependency on the structure of the visual interface compared to previous approaches that relied on a variety of task-dependent indices (e.g., features related to keyword highlighting, new search). Our automated system was designed to be implemented on a variety of visual analytics tasks, including those towards XAI (e.g., feature analysis [17]). Having an understanding of when inflections occur can later aid in semantic interactions [10] or developing adaptive UIs [28].

3 METHODOLOGY

In our approach we attempt to identify inflection points using mouse and eye tracking interaction logs as features to a supervised classification model. We discuss a prior study that collected interaction logs from a visual analysis scenario and the associated visualization tool. A human experimenter labelled perceived changes in behavior during analysis to create a "ground truth" set of labels. Processed features and labels were used to train two HMM models.

3.1 Analysis Scenario and Interaction Logs

Our research makes use of prior data collected from a visual analysis scenario, available on a public repository of analytic provenance records and interaction logs [24]. The analysis scenario implemented used the Badge and Network Traffic dataset from the VAST 2009 mini-challenge [16]. The challenge had two high-level tasks regarding an embassy employee suspected of sending data to a malicious corporation within the network. The tasks were finding which computer was used to send confidential data and characterizing patterns of suspicious computer use. Information such as employee identification numbers, proximity log card uses, and network traffic logs were provided for a period of one month.

As the basis for data collection, the visual analytics tool was designed to be similar to common designs of visualization tools with multiple coordinated views used for network analysis (e.g., [14, 27]). Using this tool, test participants conducted multi-dimensional data analysis. The tool included six views in total that supported interactive exploration of the cyber security data. Throughout the paper, we refer to these views as *areas of interest* (AOI). Five AOIs provided interactive tabular and visual presentations of the network data, and one AOI provided static information relevant to network traffic (see Figure 1). They are summarized as follows:

- Detail Histogram: An interactive histogram for filtering the period of time being analyzed by the user.
- Overview Histogram: An interactive histogram view of the data at a more macro scale.
- Network Graph: A network graph visualization showing network traffic between IP addresses. Mousing over a node dimmed unconnected nodes to facilitate visual inspection.
- Offices: A graphical layout showing employee desks and whether or not employees were in the office during the time selected via the histograms.
- Info: A static textual view with basic information about network traffic and common usage of socket numbers.
- Table: An interactive table containing network traffic between IP addresses. Users could click to sort the table by attribute, select particular exchanges for further inquiry, or filter to a specific time.

3.2 User Study for Data Collection

Seven participants engaged in an analysis session using our tool for a 90 minute duration. All participants majored in computer science and had taken at least one university level class in computer networking. The median age was 23 years. Participants were briefed on the visual analysis tool and tasked with finding the computer being used to send confidential data to malicious corporations. They were also asked to "think out loud" by verbalizing their thoughts throughout the process.

The tool was a standalone web application viewed on a 27-inch monitor with mouse input. Mouse interactions such as the position of the cursor, click events, hovering, and filtering were recorded via the application. The screen was also recorded using screen capture software. Eye tracking data was collected using a Tobii EyeX (70 Hz) eye tracking equipment attached to the bottom of the monitor. A standard microphone was employed to record audio of the think-aloud data.

3.3 Extracting Features from User Interactions

We categorized the interactions from the raw eye tracking x-y coordinates and mouse events to lay out a foundation for feature extraction. Following Brehmer et al. [4], we identified abstract analysis tasks by methods of interaction: *select*, *navigate*, *arrange*, *change*, *filter*, and *aggregate*. For our study, navigation solely relied on the x-y coordinates of the eye tracking data as a means of navigating

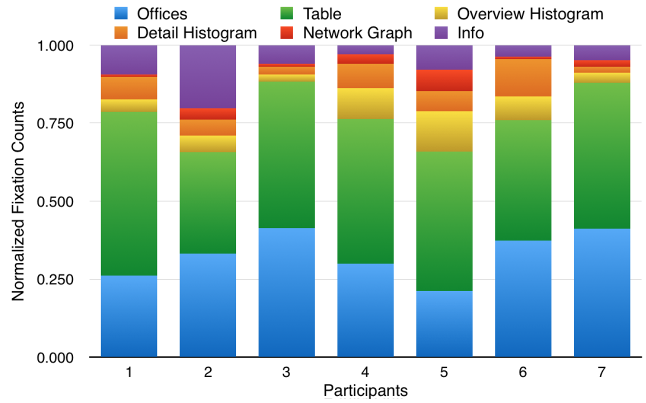

Figure 2: Ratio of eye fixations for each AOI per participant.

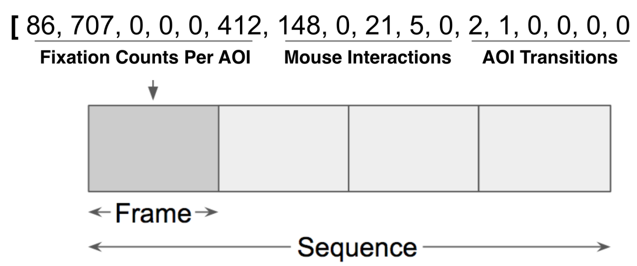

Figure 3: Schematic representation of the feature design method. Sequences were segmented by frames that include counts for: fixations per AOI, categorized mouse interactions, and eye tracking transitions originating from each AOI.

between AOIs. The remaining methods were used to categorize the mouse interactions.

To summarize the navigation data, we mapped the x-y coordinates of the eye tracking data to the six AOIs. By examining the ratio of fixations per participant across the AOIs (see Figure 2), we observed that the *offices* and *table* views had the highest occurrence of fixations, while the *network graph* and *detail histogram* had the least. The interactions were heavily skewed towards *select* interactions (Table 1), wherein a participant either hovered over an item for more information or selected an item to single it out visually. This might have been a result of it being the easiest and most natural way of viewing information while using the tool.

We derived three feature sets from the described interaction methods and AOIs. Due to the high sampling rate of the eye-tracking data in comparison with mouse interactions, we binned interactions into analysis frames of pre-determined length (Figure 3). Within each analysis frame we computed the number of fixations of the eye-tracking data per AOI, the number of mouse interactions methods, as well as the total number of eye tracking transitions that originated from each AOI. The maximum number of features per frame was seventeen (six for the fixation counts in each AOI, six for transitions between AOIs, and five for mouse interaction methods)

Table 1: Normalized mouse data categorized by method of interaction for each participant.

ID	Total Mouse Interactions	Select	Arrange	Change	Filter	Aggregate
1	29253	0.951	0.003	0.026	0.016	0.004
2	10923	0.866	0.003	0.012	0.009	0.110
3	25959	0.890	0.005	0.027	0.014	0.065
4	32077	0.952	0.001	0.009	0.013	0.025
5	9219	0.898	0.003	0.032	0.025	0.042
6	23426	0.863	0.004	0.036	0.018	0.079
7	17539	0.775	0.001	0.029	0.020	0.174

3.4 Hidden Markov Model Classification

Our system was designed to solve a binary classification problem of whether or not a sequence of interactions contained an inflection point. The inflection points were manually labeled by the experimenter based on the participant's think out louds and corresponding interactions with the visualization tool as the study was conducted. The experimenter marked an inflection when he observed the participant changing tactics in tool usage. While the use of one experimenter is not ideal, the experimenter took notes of the entire session, was well versed in the tool's usage, and listened for verbal cues. An example of an inflection would be a participant shifting from a period of time using the *detail histogram* and scanning *offices* for changes to a new strategy of looking through the *table* of network traffic. There was a mean of 15 inflections per participant, with counts ranging from 9 to 22.

We implemented a supervised learning algorithm based on two Hidden Markov Models (HMM): one modeling interaction sequences that included inflections and one with sequences that did not [33]. Approximate inflection points were classified by using the forward algorithm to determine the log-likelihood score of a sequence occurring from each model and selecting the model with the higher score. Specifically, we marked the time at the beginning of a sequence as an inflection point if the inflection model had a higher score. Sequences were formed using a sliding window of frames on the 90 minute session. The duration of the frames was always under one minute and the number of features ranged between 5 and 17. Figure 3 provides an example of the input sequence, that consists of successive analysis frames. For training, if an inflection occurred within the first half of the sequence length, the sequence would be used to estimate the model parameters of the inflection model. Remaining sequences were used on the non-inflection model. The number of hidden states for each model was either 4 or 5.

4 EVALUATION

To further understand the impact of interactions in determining the occurrence of inflections, experiments were performed using a leave-one-user-out cross-validation. One participant was held out in the test set for each fold, while the remaining data was used in the training set. The number of folds was equivalent to the number of users. Evaluation metrics were averaged over all folds. We experimented with the number of hidden states N for the inflection and non-inflection models ($N \in \{4, 5\}$), frame length K ($K \in \{0.1, 0.5, 1.0\}$minutes), sequence length L ($L \in \{4, 8, 12\}$frames), and all possible combinations of the feature sets.

Evaluation metrics include the recall for the inflection and non inflection classes. In addition, unweighted accuracy is computed as the average of the separate recalls per class. Unweighted measures are usually employed to yield unbiased estimators of the system accuracy respective of the number of samples per class. They are common measures in applications with unbalanced data, since they give the same weight to each class, regardless of how many samples of the dataset it contains [12]. Table 2 shows the results from the highest scoring models (overall and for each feature set).

Results indicate that sequences of 4 one-minute frames generally had the best scores. Intuitively, this can be justified by the fact that reducing granularity in time provides more room for error, resulting in improved accuracy. Counts of eye tracking fixations per AOI and mouse interactions were the most helpful feature sets. These were the only sets used in the highest scoring system and individually scored better than the AOI transition feature set. The highest unweighted accuracy of our system was 70.8% with an average of 23 classified inflection points. While the achieved accuracy is greater than chance (50%), additional experimentation with a greater amount of participants is required to more-reliably classify inflection points.

5 DISCUSSION AND CONCLUSION

Visualizations used to understand, diagnose, and refine models are a subset of "Explainable AI" (XAI). However, it is unclear what types of interactions are appropriate for a given model and how the visualizations are perceived. By developing methods to understand the user's sensemaking process in visual analytics we may be able to gain insight on how visualizations for AI are being used and in development of a naturalistic model of explanation. The study of sensemaking processes in visual analytics typically relies on having elaborate coding schemes, where human experts with domain specific knowledge provide annotations. However, this requires time-consuming reporting from analysts who are burdened with explaining their rationale. Our research explores the use of supervised classification to detect inflection points and its features derived solely from interaction logs. The proposed classification system of two HMM models reached an unweighted accuracy of 70% when detecting inflection points, suggesting feasibility of inflection-point classification.

Best results were obtained using a reduced time granularity of 4-minute time precision within the total 90-minute user session. The count of AOIs from eye tracking fixations and mouse interactions proved to be slightly more useful feature sets in comparison to transitions between AOIs. As part of our future work, we will experiment with time trajectory approaches that will not bin eye-tracking and mouse interaction features over an analysis frame, as this might remove valuable sequential information. We will also consider more detailed feature extraction techniques such as saccade information, raw x-y coordinates, or specific parts of the AOI to increase granularity.

Table 2: Recall rates for classifying inflection and non-inflection points using eye tracking and mouse interaction features with Hidden Markov Models. Results are from the highest scoring model across all experiments and per feature set.

Number of Hidden States		Time Sequence Specifications		Features			Performance Detecting Inflections		
Inflection Model	Non-Inflection Model	Frame Length (minutes)	Sequence Length (frames)	Eye Tracking	Mouse Interactions	Eye Transitions	Non-Inflection Recall	Inflection Recall	Unweighted Accuracy
5	5	1	4	6	5	0	93.9%	47.6%	70.8%
5	4	1	4	0	5	0	90.9%	46.3%	68.6%
4	5	1	8	0	0	6	87.1%	34.3%	60.7%
4	4	1	4	6	0	0	92.1%	43.8%	68.0%

Going forward, we seek to increase the number of participants for more generalizable models and to group participants based on interaction sequences. The analysis activity per participant consisted of highly diverse combinations of tasks and subtasks due to the nature of complex task and unstructured data exploration. Cleaner analysis tasks with more clearly defined subtasks may help our understanding of participants' inflection points and guide insights on individual sensemaking. While we do not expect to be able to perfectly capture the human analytic process through interaction logs alone, the results of this work are promising for further analysis such as comparing sequences between inflection points amongst users or segmenting the analysis session to augment visualizations of meta data.

ACKNOWLEDGMENTS

This material is based on work supported by NSF 1565725 and the DARPA XAI program N66001-17-2-4031.

REFERENCES

[1] Mohamad Aboufoul, Ryan Wesslen, Isaac Cho, Wenwen Dou, and Samira Shaikh. 2018. Using Hidden Markov Models to Determine Cognitive States of Visual Analytic Users. (2018).

[2] Tanja Blascheck, Markus John, Kuno Kurzhals, Steffen Koch, and Thomas Ertl. 2016. VA 2: a visual analytics approach for evaluating visual analytics applications. *IEEE transactions on visualization and computer graphics* 22, 1 (2016), 61–70.

[3] Tanja Blascheck, Kuno Kurzhals, Michael Raschke, Michael Burch, Daniel Weiskopf, and Thomas Ertl. 2014. State-of-the-art of visualization for eye tracking data. In *Proceedings of EuroVis*, Vol. 2014.

[4] Matthew Brehmer and Tamara Munzner. 2013. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2376–2385.

[5] Eli T Brown, Alvitta Ottley, Helen Zhao, Quan Lin, Richard Souvenir, Alex Endert, and Remco Chang. 2014. Finding waldo: Learning about users from their interactions. *IEEE Transactions on visualization and computer graphics* 20, 12 (2014), 1663–1672.

[6] Steven P Callahan, Juliana Freire, Emanuele Santos, Carlos E Scheidegger, Cláudio T Silva, and Huy T Vo. 2006. VisTrails: visualization meets data management. In *Proceedings of the 2006 ACM SIGMOD international conference on Management of data*. ACM, 745–747.

[7] Arzu Çöltekin, Sara Irina Fabrikant, and Martin Lacayo. 2010. Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *International Journal of Geographical Information Science* 24, 10 (2010), 1559–1575.

[8] Wenwen Dou, Dong Hyun Jeong, Felesia Stukes, William Ribarsky, Heather Richter Lipford, and Remco Chang. 2009. Recovering reasoning processes from user interactions. *IEEE Computer Graphics and Applications* 29, 3 (2009).

[9] Cody Dunne, Nathalie Henry Riche, Bongshin Lee, Ronald Metoyer, and George Robertson. 2012. GraphTrail: Analyzing large multivariate, heterogeneous networks while supporting exploration history. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1663–1672.

[10] Alex Endert, Patrick Fiaux, and Chris North. 2012. Semantic interaction for sensemaking: inferring analytical reasoning for model steering. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2879–2888.

[11] K Anders Ericsson and Herbert A Simon. 1998. How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity* 5, 3 (1998), 178–186.

[12] Michael Glodek, Stephan Tschechne, Georg Layher, Martin Schels, Tobias Brosch, Stefan Scherer, Markus Kächele, Miriam Schmidt, Heiko Neumann, Günther Palm, et al. 2011. Multiple classifier systems for the classification of audio-visual emotional states. In *Affective Computing and Intelligent Interaction*. Springer, 359–368.

[13] Joseph H Goldberg and Jonathan I Helfman. 2010. Comparing information graphics: a critical look at eye tracking. In *Proceedings of the 3rd BELIV'10 Workshop: BEyond time and errors: novel evaLuation methods for Information Visualization*. ACM, 71–78.

[14] John R Goodall, Eric D Ragan, Chad A Steed, Joel W Reed, G David Richardson, Kelly MT Huffer, Robert A Bridges, and Jason A Laska. 2018. Situ: Identifying and explaining suspicious behavior in networks. *IEEE transactions on visualization and computer graphics* (2018).

[15] David Gotz and Michelle X Zhou. 2009. Characterizing users' visual analytic activity for insight provenance. *Information Visualization* 8, 1 (2009), 42–55.

[16] Georges G Grinstein, Jean Scholtz, Mark A Whiting, and Catherine Plaisant. 2009. VAST 2009 challenge: An insider threat.. In *IEEE VAST*. 243–244.

[17] David Gunning. 2017. Explainable artificial intelligence (xai). *Defense Advanced Research Projects Agency (DARPA), nd Web* (2017).

[18] Lane Harrison, Wenwen Dou, Aidong Lu, William Ribarsky, and Xiaoyu Wang. 2011. Analysts aren't machines: Inferring frustration through visualization interaction. In *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on*. IEEE, 279–280.

[19] Robert R Hoffman, Gary Klein, and Shane T Mueller. 2018. Explaining Explanation For "Explainable AI". In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 62. SAGE Publications Sage CA: Los Angeles, CA, 197–201.

[20] Gary Klein, Jennifer K Phillips, Erica L Rall, and Deborah A Peluso. 2007. A data-frame theory of sensemaking. In *Expertise out of context: Proceedings of the sixth international conference on naturalistic decision making*. New York, NY, USA: Lawrence Erlbaum, 113–155.

[21] Neesha Kodagoda, Sheila Pontis, Donal Simmie, Simon Attfield, BL William Wong, Ann Blandford, and Chris Hankin. 2017. Using machine learning to infer reasoning provenance from user interaction log data: based on the data/frame theory of sensemaking. *Journal of Cognitive Engineering and Decision Making* 11, 1 (2017), 23–41.

[22] Rhema Linder, Alyssa M Pena, Sampath Jayarathna, and Eric D Ragan. 2016. Results and challenges in visualizing analytic provenance of text analysis tasks using interaction logs. In *Logging Interactive Visualizations and Visualizing Interaction Logs (LIVVIL) Workshop*.

[23] Shixia Liu, Xiting Wang, Mengchen Liu, and Jun Zhu. 2017. Towards better analysis of machine learning models: A visual analytics perspective. *Visual Informatics* 1, 1 (2017), 48–56.

[24] Sina Mohseni, Andrew Pachuilo, Ehsanul Haque Nirjhar, Rhema Linder, Alyssa Pena, and Eric D Ragan. 2018. Analytic Provenance Datasets: A Data Repository of Human Analysis Activity and Interaction Logs. *arXiv preprint arXiv:1801.05076* (2018).

[25] Sina Mohseni, Alyssa Pena, and Eric D Ragan. 2018. ProvThreads: Analytic Provenance Visualization and Segmentation. *arXiv preprint arXiv:1801.05469* (2018).

[26] Phong H Nguyen, Kai Xu, Ashley Wheat, BL William Wong, Simon Attfield, and Bob Fields. 2016. SensePath: Understanding the sensemaking process through analytic provenance. *IEEE transactions on visualization and computer graphics* 22, 1 (2016), 41–50.

[27] Steven Noel, Michael Jacobs, Pramod Kalapa, and Sushil Jajodia. 2005. Multiple coordinated views for network attack graphs. In *Visualization for Computer Security, 2005.(VizSEC 05). IEEE Workshop on*. IEEE, 99–106.

[28] Andrew Pachuilo, Eric D Ragan, and John R Goodall. 2016. Leveraging Interaction History for Intelligent Configuration of Multiple Coordinated Views in Visualization Tools. *LIVVIL: Logging Interactive Visualizations & Visualizing Interaction Logs* (2016).

[29] Peter Pirolli and Stuart Card. 2005. 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. McLean, VA, USA, 2–4.

[30] Eric D Ragan, Alex Endert, Jibonananda Sanyal, and Jian Chen. 2016. Characterizing provenance in visualization and data analysis: an organizational framework of provenance types and purposes. *IEEE transactions on visualization and computer graphics* 22, 1 (2016), 31–40.

[31] Eric D Ragan and John R Goodall. 2014. Evaluation methodology for comparing memory and communication of analytic processes in visual analytics. In *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*. ACM, 27–34.

[32] Dominik Sacha, Andreas Stoffel, Florian Stoffel, Bum Chul Kwon, Geoffrey Ellis, and Daniel A Keim. 2014. Knowledge generation model for visual analytics. *IEEE transactions on visualization and computer graphics* 20, 12 (2014), 1604–1613.

[33] Dario D Salvucci and Joseph H Goldberg. 2000. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*. ACM, 71–78.