Supporting Handoff in Asynchronous Collaborative Sensemaking Using Knowledge-Transfer Graphs

Jian Zhao, Michael Glueck, Petra Isenberg, Fanny Chevalier, Azam Khan

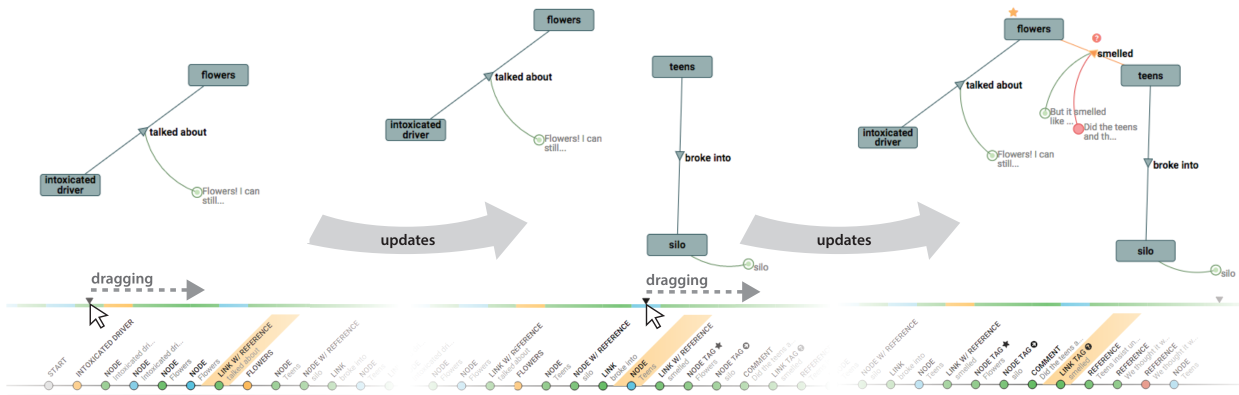

Fig. 1. The Knowledge-Transfer Graph (KTGraph) supports the externalization of the analytic process through dedicated graph elements. Nodes and links to describe discovered concepts and relationships; integrated links to source material (green), comments (red), and tags (e. g., golden star and red question mark badges), as well as an interactive playback of graph creation help support handoff of partial findings to subsequent investigators.

Abstract—During asynchronous collaborative analysis, handoff of partial findings is challenging because externalizations produced by analysts may not adequately communicate their investigative process. To address this challenge, we developed techniques to automatically capture and help encode tacit aspects of the investigative process based on an analyst's interactions, and streamline explicit authoring of handoff annotations. We designed our techniques to mediate awareness of analysis coverage, support explicit communication of progress and uncertainty with annotation, and implicit communication through playback of investigation histories. To evaluate our techniques, we developed an interactive visual analysis system, KTGraph, that supports an asynchronous investigative document analysis task. We conducted a two-phase user study to characterize a set of handoff strategies and to compare investigative performance with and without our techniques. The results suggest that our techniques promote the use of more effective handoff strategies, help increase an awareness of prior investigative process and insights, as well as improve final investigative outcomes.

Index Terms—Collaboration, sensemaking, handoff, handover, structured externalizations, interactive visual analysis.

1 Introduction

Visual representations of knowledge and insights can help analysts track, organize, and synthesize their observations and questions during data analysis [16, 46, 53]. As investigators collaborate, these visualizations of knowledge can evolve into shared thinking and sensemaking spaces in which investigators synthesize and build on each other's findings [26, 37]. Our work is situated in the domain of collaborative sensemaking with visual analytics tools. Sensemaking is a process that helps to bridge gaps in understanding and find meaning in information [34]. In collaboration, sensemaking is required to create a shared understanding of information through communication and/or the creation and manipulation of shared data representations [33].

Interactive systems have been shown to be effective for collaborative sensemaking in synchronous co-located collaborations [20, 29] and for supporting asynchronous remote collaboration over the internet [15, 47, 52]. However, the explicit transfer of knowledge between collaborators remains a key visual analytics challenge in complex sensemaking scenarios [3, 20, 29, 35, 43].

This transfer of knowledge is known as handoff (or handover) in the shiftwork literature. Successful handoff is difficult even when investigators are trained to use standardized workflows and protocols (e. g., in intelligence analysis [35]). In general, non-routine and ill-defined tasks often exacerbate the difficulty of handoff due to the different strategies employed by individual investigators and insufficient communication between collaborators [43]. In addition to a lack of personal communication, incomplete externalizations (external representations of a person's internal thoughts) created by analysts generally do not themselves adequately communicate strategies or thought processes to the subsequent collaborator [43].

Our work focuses on supporting handoff of *partial findings* (i.e., developing, incomplete analysis results) from one asynchronously working investigator to the next. These types of handoff occur when investigators cannot meet in person due to geolocation, time constraints, time zones, privacy concerns, or personal preference. In this context, researchers have previously identified the often required transfer of partial findings as a valuable but rarely supported feature of analysis systems and, thus, an important research challenge to address [3, 43, 54].

In this work, we contribute a set of interactive visualization techniques to mediate handoff by encoding awareness of data coverage through references to source materials, explicitly communicating analysis progress and uncertainty using comments and tags, and sup-

- *Jian Zhao is with FX Palo Alto Laboratory. E-mail: zhao@fxpal.com*
- *Michael Glueck and Azam Khan are with Autodesk Research. E-mail: {michael.glueck|azam.khan}@autodesk.com*
- *Petra Isenberg and Fanny Chevalier are with Inria. E-mail: {petra.isenberg|fanny.chevalier}@inria.fr*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

porting implicit communication of thought process through playback of investigation histories. We implemented our techniques in the Knowledge-Transfer Graph (KTGraph), a novel collaborative visual analysis tool that supports handoff between investigators.

In addition to the tool itself, we contribute design considerations and rationale for our approach to support the handoff of partial findings, based on a review of work studying collaboration and sensemaking. Further, we conducted a two-phase user study to evaluate our techniques in a document analysis scenario. Results provide insights into handoff strategies and suggest that playback of investigation histories is an effective approach to support the handoff of partial findings.

2 Background and Related Work

Our work primarily relates to prior work in three areas: collaborative sensemaking/handoff, collaborative visual analytics, and the support of externalizations of analysis progress.

2.1 Collaborative Sensemaking and Handoff

Sensemaking is a concept that has been studied in many disciplines [39, 50] and in HCI is generally understood as a process of understanding information [34]. Weick [50] suggests sensemaking activities benefit from the social aspects of collaboration. It is through communication with others that multiple perspectives are integrated and debate and negotiations are triggered. These interactions can further elicit externalization of latent insights. In the field of HCI, collaborative sensemaking with shared externalizations has been particularly studied for information retrieval tasks, such as collaborative search [31, 32]. Here, shared workspaces help collaborators to manage, organize, and discuss found information.

In our work, we particularly focus on the problem of handoff of partial findings in data analysis scenarios. Handoff in collaborative sensemaking has been investigated through studies of information help desks [43], product recommendations [44], and intelligence analysis [3]. However, while the sensemaking work most closely related to ours comes from the field of visual analytics [19, 29], we are not aware of work in this domain that shares our focus on handoff and effectiveness of transferring previous investigators' knowledge.

2.2 Collaborative Visual Analytics

Supporting collaborative data analysis is a key research challenge in the domain of visual analytics [18, 46]. Researchers have proposed both novel system designs [19, 29, 40, 49, 51], discussed and implemented software infrastructures [2, 30, 42], studied collaborative data analysis behavior [20, 21, 37, 48], and highlighted design considerations and open research problems [13, 17, 18, 54] to address this challenge.

We were particularly inspired by work on the support of awareness, coordination, and synthesis in collaborative analysis activities. In particular, we drew on the principle of information scent [34] in both co-located (e.g., [19, 29]) as well as distributed collaboration tools (e. g., [15, 47, 49, 52]). These features use cues to seen or unseen information in order to help broaden exploration coverage.

The most closely related systems to ours are Cambiera [19] and CLIP [29] which highlight shared searches and discoveries in documents. These systems, however, focused on co-located collaboration and do not specifically address the handoff of partial findings.

Coordination and synthesis strategies between asynchronous data analysts were also previously studied [37, 40, 41], sharing our interest in understanding and consolidating work of previous investigators. Robinson [37] focused on the co-located synthesis of findings after an asynchronous distributed analysis phase. Sarvghad et al. [40, 41] found visualizing data dimension coverage of a previous analyst's exploration can promote awareness and understanding, as well as the question formation process, in tabular data analysis.

Similarly, KTGraph supports coordination by highlighting previously investigated data, but adds user-created externalizations under the assumption that they provide more explicit representation of thought processes. In addition, the handoff techniques of KTGraph do not assume any specific underlying data structures (e. g., data tables) and is, thus, potentially generalizable for broader applications including both structured and unstructured data.

2.3 Externalization Approaches

Graph-based visualizations have been used extensively to externalize investigations in visual analysis systems, including investigating document collections (e. g., [23, 28, 45]), intelligence analysis (e. g., [3, 7, 29, 35], and visual analytics (e.g., [6, 55]). KTGraph builds on the general visual design of our previous work, Annotation Graphs [55], that offered specific exploration capabilities for graphs based on added data annotations. KTGraph, while visually similar, has a very different goal. Annotation Graphs focused on providing meaningful layouts for data exploration based on user annotations, but KTGraph focuses on providing capabilities for building graphs as externalizations of analysis results and handing them over to subsequent analysts.

CommentSpace [52] is related to our work in that it effectively used tags and links to structure comments, which improved analytics results. Although we also support user generated comments and tags to explicitly communicate intent, uncertainty, and progress of the investigation, we apply them in a very different context: an interactive and evolving graph visualization. In addition, we integrate features, such as references to source data (e. g., [29]) and document visitations (e. g., [1]) directly into the interface to improve awareness of prior investigators work through analytic provenance [54].

Timelines are often used in visual analysis tools to represent temporal relationships within the data being investigated (e. g., [4, 5, 7, 8, 28]). Heer & Agarawala suggest that asynchronous collaborations can benefit from timelines to communicate the temporal progression of investigations [13]. The value of communicating the temporal progression and development of an analysis has been demonstrated using static visualizations for individual investigators (e. g., [14, 23]) and in collaborative settings (e. g., [6]), as well as for interactive tutorials of software workflows (e. g., [11]). KTGraph extends this concept to an interactive playback of externalization creation and management to implicitly communicate temporal aspects of investigations, such as externalization updates, analyst insights, and analysis rationales.

3 Design Considerations

Prior work has significantly shaped and inspired our design of handoff features. Here, we describe the design considerations we derived from the literature and briefly discuss which related work they are based on.

G1. Support interactive externalizations. The representation should be sufficiently abstract to provide flexibility of expression and to support a variety of analysis styles [24, 25]. A tight coupling between data exploration and externalization can help mediate switching between deductive and inductive modes of reasoning [3, 13, 38]. Above all, the representation should be interactive, enabling analysts to continuously build, re-evaluate, re-structure, expand, and refine the externalization as their understanding evolves [13, 25, 35, 38].

G2. Encode analytic provenance. Understanding where in the data insights came from is a critical aspect of communicating the current state of an analysis [54]. The system should support quick lookups to understand the source of insights in the data [35]. This enables analysts to double-check the logic of the prior collaborator and build confidence in the existing state of the analysis. It also helps analysts to deduce and revisit the rationale behind prior analyses [25].

G3. Facilitate common understanding. Sharma & Furnas [44] underscore the importance of a consistent method of externalization to support common understanding between analysts. Drawing on the principle of information scent [34], the system should, thus, enable analysts to explicitly encode their progress and process [35], for example, through notification mechanisms (i. e., action flags or commentary [13, 25, 52]). The system should summarize which data has already been considered and explored, such as revealing most and least frequently visited regions of the data collection [13].

G4. Provide interaction and analytic provenance. Given a lack of direct communication between collaborators, the system should provide implicit methods of transferring an awareness of the progression and development of the analysis [6, 13, 54]. These analysis histories can

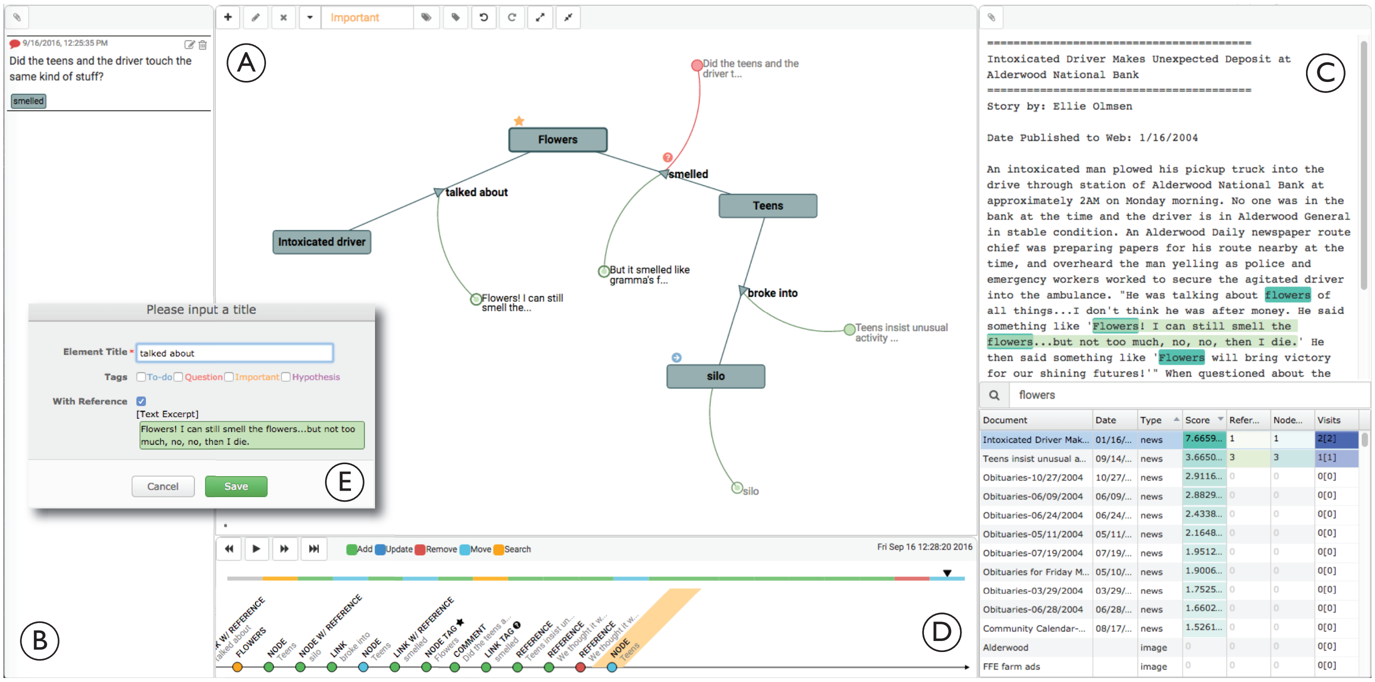

Fig. 2. The KTGraph interface. An investigator uses the Graph Panel (A) to externalize their investigation. They can review comments related to the investigation in the Comments Panel (B). The Dataset Panel (C) displays the dataset under investigation. The Timeline Panel (D) enables investigators to playback the investigative history. An input dialog (E) is popped-up when creating a new node or link.

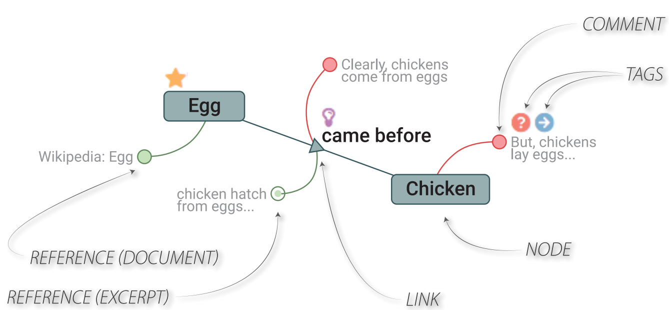

Fig. 3. Example externalization illustrating the visual language of the interactive graph visualization: nodes, links, tags, comments, and references. Comments and references can be attached to nodes and links; tags can be attached to nodes, links, references, and comments.

support rapid uptake of prior collaborators' work [35] and help infer the rationale underlying their insights [36].

4 Knowledge-Transfer Graph

Taking the aforementioned considerations together, we designed KT-Graph (Fig. 2) to offer the following features to facilitate handoff in asynchronous collaborative data analysis:

• KTGraph extends the traditional concept graphs through *embedded references, comments, and tags*, with which investigators can explicitly communicate their progress and uncertainty using a coherent visual language (Fig. 3). This design offers a thorough picture of the externalized knowledge promoting externalization during the early stages of sensemaking, when an investigator's understanding may be unresolved. For example, a reference connected to many nodes might be an important evidence to look at for the following investigator.

• KTGraph provides an *interactive timeline* that records every historical state of the investigation. An interactive playback function allows to review how prior investigators encoded and built up the visualization, implicitly communicating intent and strategy (Fig. 1).

• KTGraph supports *tagging* of any element of the graph. Tags can help an analyst convey the meta-information of their thoughts, e. g., pointing out promising directions to dig in (Fig. 3) or to-do items.

Thus, KTGraph facilitates handoff through specific knowledge transfer features that convey what has been done, why it was done, and what remains to be done. In the remainder of this section, we briefly describe the KTGraph interface and its four coordinated panels: Graph, Comments, Document, and Timeline (Fig. 2).

4.1 Graph Panel

In the *Graph Panel* an analyst can build a graph visualization (G1) (Fig. 1, Fig. 2A) to externalize investigative progress. Possible graph components are nodes, links, tags, comments, and references (Fig. 3).

Nodes Egg and *links* came before can be created and labeled to encode abstract concepts or entities (e. g., people, places, or events) and their relations to each other. The nodes can be placed freely at any position on the 2D canvas. In addition to using the toolbar + ✎ ×, the graph can be edited using direct manipulation, for example, dragging from one node to another creates a link and click-and-drag displaces nodes on the canvas.

We provide four basic *tags* that are represented by icons in the graph: *To-do*, *Question*, *Important*, and *Hypothesis*, similar to the fixed set of tags used in CommentSpace [52] (G3). Further, investigators can also define custom tags to extend this basic vocabulary using Important. Tags can be used to highlight important findings, track progress, and denote uncertainty. Tags can be added to any element in the graph: node, link, comment, or reference.

4.2 Comments Panel

The text of *comments* is displayed in the *Comments Panel* (Fig. 2B) and represented by a small icon in the graph (G3). Selecting a comment in the Comments Panel highlights the associated graph elements in the Graph Panel, and vice versa. Comments are free-form text and can be used to encode thought process (e. g., document findings, explain reasoning, propose hypotheses, or pose questions). Comments can be added to any node or link in the graph using the button, and they are sorted in a reverse chronological order in the Comments Panel.

4.3 Dataset Panel

The *Dataset Panel* displays the data being investigated (Fig. 2C). Our current implementation supports investigative document analysis, and thus, this panel displays a keyword search interface to explore a set of documents with a detail view for one document. Documents are ranked by a search result score and matching keywords in the selected document are highlighted. For each document, meta-information such as the number of visits and references is also visualized (G3).

References associate nodes and links to the source evidence in the dataset, which can be whole documents or excerpts (G2). Text excerpts within documents are highlighted with a green background. Selecting a reference loads the corresponding document in the Dataset Panel, and based on the type of reference, the text excerpt or the whole document is visually emphasized. References can be added to any node or link by first selecting the objects and clicking the button. To streamline authoring tags and references, KTGraph auto-populates creation dialogs with titles and references based on selected text in the Dataset Panel (Fig. 2E) (G1).

While in this paper we focus on document analysis tasks, KTGraph is generalizable by replacing the Dataset Panel with other visualizations for specific applications, such as the time-series data tables in Annotation Graphs [55]. Instead of representing text and text excerpts, references could then link to data chunks selected by an analyst.

4.4 Timeline Panel

To explore the evolution of the graph over time, an analyst can interactively playback all historical states of the graph using the *Timeline Panel* via (Fig. 1, Fig. 2D) (G4). The timeline captures and shows all graph operations, including creating, deleting, and editing of nodes and links; layout modifications; attaching comments, tags, and references; as well as keyword search terms. Each operation type is color-coded and the overall history is summarized using principles of overview+detail. The entire history is displayed as a thin line on the top, and the detailed operations close to the current focusing timestep are represented as circles below. To save space, multiple similar layout modifications are aggregated visually.

KTGraph creates a session for each investigator. The timeline shows the start and end of each session with a small visual break (whitespace). An analyst can use the timeline to perform undo or redo operations *only* in their investigation session. Most importantly, the timeline provides implicit awareness of what previous investigators have done through the animated playback. An analyst can pause this animation at any time to explore the elements of the historical state of the graph without modifying them. The final state of the graph serves as the starting point for the subsequent analyst and can, however, be freely modified.

5 Sensemaking Task

The current implementation of KTGraph supports investigative document analysis, a task which has previously been shown to benefit from collaboration [20,29]. We used the VAST 2006 Challenge, Stegosaurus, [10] during development and testing of KTGraph. Stegosaurus is recognized as a benchmark visual analytics task solvable by a general population, and it has been used in previous studies of sensemaking [1, 12] and collaborative analysis [20, 29]. The goal of the scenario is to investigate suspicious activities in a small town. To successfully solve the challenge, analysts have to filter out irrelevant documents, read and discover facts in the news articles, make logical connections, and generate hypotheses. The task starts with a vague clue about a man crashing into a bank; from there, an analyst needs to work through the dataset to reveal additional clues to solve the mystery. To mimic real-world scenarios, there are distractors that could lead an analyst in the wrong direction. Our version of the synthetic dataset contained 246 documents (with about 3,000 entities), a majority of which are fictional news articles, as well as additional image resources and several longer information documents. Only 11 documents contain pertinent clues to the investigation. The task is expected to take a single investigator 2–6 hours to solve using a basic keyword search interface [29].

5.1 Example Handoff Scenario

Ryan and Emma want to collaborate to analyze the Stegosaurus data. Each can contribute various expertise related to the small town under investigation. However, since Ryan and Emma do not share the same timezone it is difficult for them to work synchronously or communicate directly. They decide Ryan will start the investigation using KTGraph and Emma will pick up where he stops. Please refer to the accompanying video for a demonstration of the scenario.

5.1.1 First Session: Initial Investigation.

Ryan externalizes his mental model piece-by-piece by creating a graph representation. By abstracting salient evidence and connections as he discovers them, he consolidates his understanding and develops hypotheses. He creates a breadcrumb trail to the source articles by embedding references. Using comments and tags, he is able to express important discoveries, uncertainties, and flag nodes and links requiring further investigation. This not only helps him track his own progress, but also helps encode his thought process for subsequent investigators. Ryan stops his investigation because his expertise is requested for another case. The result is an interactive graph visualization of his partial findings that Emma can build upon.

5.1.2 Second Session: Transferring Knowledge.

To continue the investigation that Ryan started, Emma opens the investigation in KTGraph, which starts a new session. At first glance, Ryan's graph is complicated and difficult to fully grasp. Although the tags are good indicators of the key concepts, (e. g., `Flowers`), a clear starting point is not evident. Emma shifts her attention to the Timeline Panel to review the operations Ryan performed during his session.

Emma rewinds the overview timeline to the beginning and hits the *Play* button to watch a smooth animation of the graph construction history. Since all the panels are synchronized with the timeline, Emma can explore the graph at any moment of Ryan's investigation and relate the state of the graph to the material he was consulting.

Emma pauses the playback because she notices many references were attached to `Silo`, which was tagged `To-do` at the time. She explores the graph by hovering over graph elements to reveal connected elements and associated items in other panels. She clicks on the references to view the associated articles which highlights the referenced text. This enables Emma to quickly skim articles to understand how they relate to the representations in the graph.

Next, Emma sorts the Dataset Panel by the number of visits and references to see what other documents Ryan read. Emma spots articles that are related to the `Silo` but have not been visited or referenced. She decides this is a good place to pick-up the investigation.

This access to history implicitly communicates when content in the graph changed, helping Emma infer why these changes were made. Emma is thus able to quickly catch-up on the investigation, discovering that Ryan had started an investigation of articles about the `Silo`, but not finished reading all the matching articles. Like Ryan, Emma continues augmenting the graph to document her thoughts, thereby leaving traces for potential subsequent investigations.

6 User Study of Handoff

To better understand how our tool would be used in settings inspired by our fictional handoff scenario, we conducted a user study in two phases. We evaluated handoff performance and characterized handoff strategies using the Stegosaurus document analysis challenge [10].

6.1 Research Questions

Very little is known about how interface features impact handoff in visual analytics. Based on our literature analysis, we hypothesized that in particular the availability of the Timeline Panel as well as the tagging features would positively impact handoff between investigators. We thus collected both performance measures as well as subjective preferences in two study phases: Phase 1 in which we studied the activities of the follow-up analyst exclusively and Phase 2 which captured data from both a starting and follow-up analyst. As prior work [20] found that work styles in collaborative data analysis are extremely varied and can impact the usefulness of specific interface features, we chose to observe and analyze handoff strategies as well. Knowledge about different handoff strategies allowed us to assess whether different strategies similarly impact the usefulness of interface features and the success of the joint data analysis.

6.2 Participant Recruitment and Apparatus

We recruited participants from a pool of university graduate students and professional researchers. All participants had a background in computer science or engineering, were proficient in English, and had no prior experience with the Stegosaurus challenge. Participants were only eligible to participate in one of the two phases. Each session lasted ~1.5 hours and participants received a $25 gift card in compensation.

The study was conducted using a 24-inch desktop monitor with a mouse and keyboard. The visual interface was presented in full-screen mode. Screen- and audio-recordings of the sessions were captured, as well as participant interaction logs and the graph creation histories.

In both phases, we used a Baseline that emulates systems in previous synchronous collaboration studies (e. g., CLIP [29]). To keep the consistency of our interface design in both conditions, we built the Baseline from the codebase of KTGraph, which includes identical Graph, Comments, Datasets Panels, and user interactions, but does not comprise the novel features (i.e. no Timeline Panel, and no tagging).

6.3 Phase 1—Handoff Performance

The Phase 1 focused on the evaluation of handoff performance. We simulated an asynchronous collaboration scenario by preparing a starting (handoff) graph. The same graph was used by all participants, enabling us to analyze repetitions of the same handoff instance.

6.3.1 Design and Methods

We used a between-subjects design with 20 participants (6 females); between 21-52 years of age (mean 32). Participants were randomly assigned to one of the two conditions (KTGraph or Baseline), resulting in 10 handoff instances in each condition.

We designed the handoff graph based on observations from several pilot studies. We referenced evidence from 2/11 key documents that included information for solving the mystery. In addition, we added comments and tags that documented uncertainty of facts and highlighted correct directions for investigation. We did not include misleading information. The handoff graph included 17 nodes, 20 links, 22 references, 3 comments, and 15 tags; covering 2 key documents.

The handoff graph represented a critical point in the investigation. It presented verified initial suspicions and some hypotheses, but raised more questions than it answered, leaving many opportunities for follow-up investigations. The handoff graph enabled us to measure the effectiveness of handoff by asking participants what they learned from the graph and then observing how well they continued the investigation. The handoff graphs used in both conditions were identical but tags were not supported in the Baseline condition. However, we ensured that they provided the same level of information.

6.3.2 Performance Metrics

Inspired by the literature, we employed the following three metrics to evaluate handoff performance.

The *handoff score* quantifies the critical facts a participant explained to the researcher after reviewing the handoff graph, similar to the scoring used in previous studies [20,22,29]. Participants could mention 6 critical facts and 3 directions for investigation, for a total score out of 9. The handoff score is, thus, a measure of the amount of information participants successfully understood from the handoff graph.

The *debriefing score* quantifies task performance based on a previously used scoring schemes [20, 22, 29]. Participants had to describe their hypotheses in detail and identify key insights in the scenario. Participants received up to 11 positive points for correct insights and negative points for incorrect hypotheses and statements.

The *key documents score* counts references to key documents in the graph [20], excluding the two documents from the handoff graph. Participants received a positive point for every key document attached to elements of the graph. Merely visiting a key document did not affect the score. This was a score out of 9 (2/11 key documents excluded).

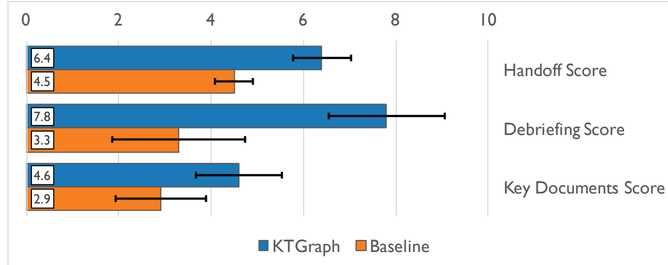

Fig. 4. Different measurements of the task performance in Phase 1. Mean values; error bars indicate 95% confidence intervals.

6.3.3 Procedure

Each study session began with a 15-minute training period about the technique (KTGraph or Baseline). We asked participants to try all the interface features and resolved any questions they had. In the tutorial we used an example dataset of five news articles from the New York Times. The investigation period followed directly after training.

At the start of the investigation period, each participant received the Stegosaurus introduction document on paper that described the scenario and provided the first clue. Participants were told that they were working in teams to solve the mystery, but that they would not be able to meet or communicate with each other. We advised them to utilize the interface to document their progress and thought process to help subsequent investigators continue the investigation.

We assessed handoff at two times during the investigation, first after an initial exploration of the handoff graph, and then again after the end of the investigation period. This methodology is similar to the one used by Sarvghad et al. [40] to study comprehension of externalizations.

Participants were given a total of 40 minutes to work on the investigation. When starting, participants were instructed to explore the handoff graph and prepare to report what had been done in the investigation so far. They were given up to 10 minutes to freely explore the handoff graph, but could choose to report sooner to gain time for their own investigation. During this initial exploration, they were not able to make any modifications to the graph. When ready, we asked them to explain the prior investigation to the researcher, during which the timer was paused. To assess knowledge-transfer after this initial exposure, the researcher conducted a short informal interview. The researcher was careful to guide the conversation towards salient points, without directly asking questions providing hints. At the end of the study, the researcher analyzed interaction logs, recorded videos, and interview notes to obtain the performance metrics described earlier.

6.3.4 Results and Discussions—Handoff Performance

The results of Phase 1 demonstrate that KTGraph was more effective at supporting handoff than Baseline, improved understanding, and led to a better overall performance in the Stegosaurus task (Fig. 4).

KTGraph's mean handoff score was 71% (6.4/9, CI [5.8, 7.0]) compared to 50% (4.5/9, CI [4.1, 4.9]) in Baseline. Regardless of condition, all participants spent similar amounts of time reading the handoff graph, 6.8 minutes ($\sigma = 2.1$) in KTGraph and 6.7 minutes ($\sigma = 1.9$) in Baseline. This suggests that facts were more effectively understood when timeline and tagging features were available.

The mean debriefing score for KTGraph was 71% (7.8/11, CI [6.5, 9.1]) compared to 33% (3.3/11, CI [1.9, 4.7]) in Baseline, indicating that participants were able to advance the investigation much more successfully with the additional tagging and timeline features.

We also found evidence that participants discovered more key documents with KTGraph, 51% (4.6/9, CI [3.7, 5.5]), than Baseline, 32% (2.9/9, CI [1.9, 3.9]) [27]. We further analyzed the number of visits and references for each key document (Fig. 5). In both conditions, we saw fairly similar patterns of document visitation and referencing. However, in KTGraph no participant added references to either of the two key documents included in the handoff graph, while in Baseline some participants did (Fig. 5: D1, D2). Moreover, participants using Baseline did not create any references to two later documents, even though they visited them (Fig. 5: D9, D10).

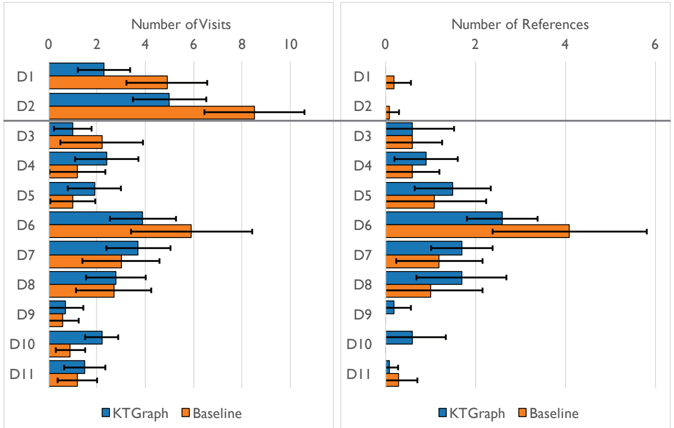

Fig. 5. Number of additional visits (viewed in Dataset Panel) and references (included in graphs) to the 11 key documents by participants in Phase 1. D1 and D2 are the initial two key documents included in the handoff graph. Error bars indicate 95% confidence intervals.

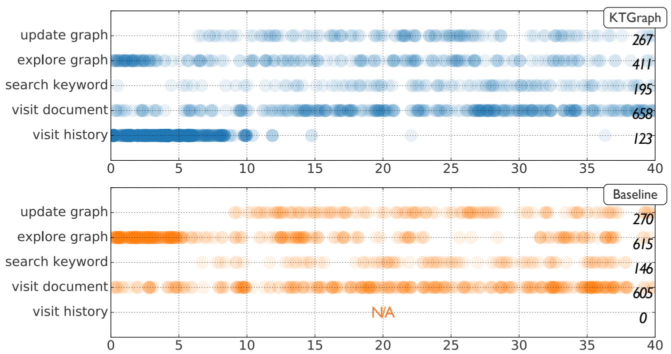

Fig. 6. Plots of participant investigative interaction patterns over the investigative session in Phase 1. The x-axis represents time in minutes. Numbers on the right indicate total counts of the actions.

This finding is interesting when considered relative to the handoff and debriefing scores. Despite having searched for and having found evidence in similar documents, the participants using KTGraph were better able to identify and integrate evidence from the documents that were pertinent to the investigation. Based on the higher handoff scores, we speculate this is because they had developed a better understanding of the intent of the first investigator, which helped them more efficiently focus their subsequent investigative work.

In short, KTGraph outperformed Baseline on all three performance metrics. The 95% confidence intervals indicate that the differences were significant for handoff and debriefing scores (Fig. 4) [27].

6.3.5 Results and Discussions—Interaction Logs

We further analyzed the interaction logs of participants captured during the experiments (Fig. 6). In general, the sequence of actions was similar between the KTGraph and Baseline conditions. One noticeable difference, however, was that "explore graph" operations (e. g., clicking a node or link) appeared more evenly distributed in the KTGraph condition, while in Baseline, there were more gaps and clusters. There were also fewer "explore graph" actions in KTGraph, which may be due to the higher demand of exploring the handoff graphs in Baseline (because the timeline was not available). There were greater numbers of document views and searches with KTGraph (658 and 195 respectively) compared to Baseline (615 and 146 respectively), which might indicate that participants consumed more information. In KTGraph, the vast majority of operations performed on the Timeline Panel occurred at the start of the investigation session, although two participants revisited the history at later points during their investigation.

Fig. 7 shows the total number of graph elements that participants created during the experiment. In general, the results of KTGraph and Baseline are similar, although confidence intervals for Baseline are larger (except for comments). On average, participants created

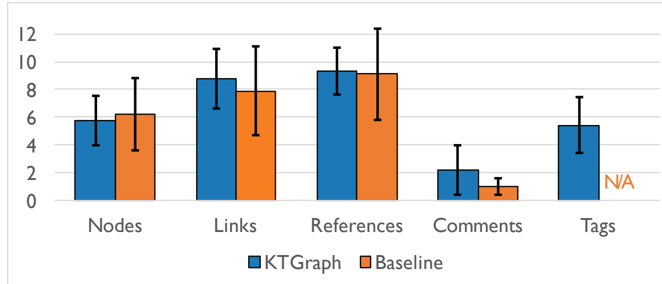

Fig. 7. Number of nodes, links, references, comments, and tags created by participants. Error bars indicate 95% confidence intervals.

more comments with KTGraph, however, the effect is not significant. Moreover, although participants in both conditions generated similar number of references, the references in Baseline have less coverage of key documents (Fig. 5) – most references were to document D6. Combining the results of Fig. 4, Fig. 5, and Fig. 7 indicates that participants were more effective with KTGraph while using roughly similar numbers of graph elements to encode information.

6.4 Phase 2—Externalization Diversity and Preferences

Phase 2 focused on a more externally valid end-to-end collaborative scenario during which participants, who did not communicate face-to-face, worked in asynchronous groups to first start and then follow-up the investigation. This enabled us to investigate various handoff graphs and a larger variance of handoff strategies between investigators at different stages in the investigation.

6.4.1 Design and Methods

We again used a between-subjects design, and recruited another 18 participants (7 females). They were between 22-34 years of age (mean 26). Since we wanted to evaluate different stages of the handoff process, we divided participants into groups of three. Participants were randomly assigned to one of the two conditions (Baseline and KTGraph) and randomly assigned to a group of three, resulting in three investigative sessions for each condition.

Each group worked serially on a single investigation, resulting in one initial session and two handoff sessions. For each team, participants came in at different times and did not have any contact with each other.

6.4.2 Procedure

The procedure for the training period and the investigation period were identical to Phase 1. Participants were given a total of 45 minutes to work on the investigation, which included reviewing the work of prior investigators (if any) and conducting their own investigation. We did not interrupt the investigative session so as not to interfere with the progression; we provided no help or guidance to participants.

After the investigative session, participants completed a questionnaire, followed by a semi-structured interview. The questionnaire used Likert-scale ratings to query the perceived usefulness of each interface feature in the context of performing current investigations and also in understanding prior investigations (1 – strongly disagree, 7 – strongly agree). The first participant in each group worked without a handoff graph and, thus, skipped questions regarding handoff. The questionnaire also asked participants to assess their behavior during handoff and their confidence in their understanding of the investigation.

Semi-structured interview questions focused on the interface features, participants' investigative strategies, and the results of their investigation. Particularly, we asked participants how they conducted their investigation and what factors played a role. Finally, we also asked participants to explain their hypotheses, describe the plot, and identify the key players. During the interview, participants were allowed to use their created externalizations during explanations. The same researcher who coded data in Phase 1, analyzed the experimental data to obtain the performance metrics for this phase.

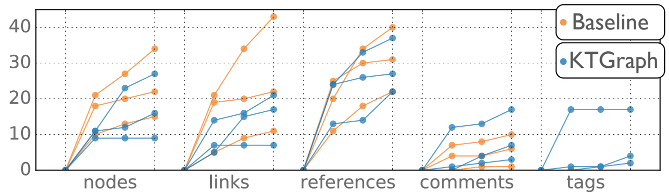

Fig. 8. Plots of numbers of graph elements generated in Phase 2. Each connected line indicates a complete investigation (including one initial and two handoff sessions).

6.4.3 Results and Discussion—Externalization Generation

Fig. 8 gives an overview of the diversity of the graph externalizations generated by each group. Each polyline corresponds to one team and each dot to an individual investigator. As can be seen, graphs were generally extended by subsequent investigators but there was a spread of externalization styles.

Comparing the numbers of links in both conditions, participants with Baseline generated larger and denser graphs with more nodes and links. However, the numbers of references were similar, which may indicate that KTGraph helped participants more effectively embed information from the dataset. The second and third dots correspond to the handoff graphs that subsequent investigators encountered during the study. We can observe that the variability of graphs is relatively high. Comparing the slopes of line segments in each session reveals that, in general the graphs, in terms of node, link, and reference numbers, were extended less by the third investigator compared to the second. However, several exceptions appeared for comments and tags. This may indicate that investigators shifted from objective discovery to more subjective deduction as the investigation progressed. It is worth noting that future studies are warranted to see more generalizable patterns of handoff graphs as there were not many data points collected.

6.4.4 Results and Discussion—Ratings of Usefulness

In the questionnaire, KTGraph and Baseline received similar ratings on usefulness of interface features (Fig. 9). Both systems were thought to be easy to use and learn. The ability to externalize one's thinking with an interactive graph with references and comments embedded was rated useful (G1, G2). However, we also found several larger differences. First, participants rated reviewing information (e. g., visit counts) on the Dataset Panel (Fig. 2C) more useful in Baseline than in KTGraph (C11, F11). This may be because Baseline is not equipped with the Timeline Panel so that this meta-information became a critical means of awareness. Second, comments were surprisingly rated less useful in Baseline. We examined the interaction logs and found that fewer comments were created in the Baseline condition (C4, F4), which may be due to participants' personal working styles or not wanting to break the flow of the investigation. Although we instructed participants to use comments to communicate their thought process for future investigators, in practice most became engrossed in the task and did not document their progress. Third, the handoff graphs were thought less useful in Baseline for understanding prior discoveries (F1, F7), which is possibly because participants found them difficult to interpret.

Of the features specific to KTGraph, tags rated well, but were infrequently used. On average, only 2.6 tags were created per session and three participants did not use tags at all. Two of them commented that they regretted not creating more tags during their investigations. Although the interactive history exploration was not viewed as useful for the initial investigators, these features were found very important by subsequent investigators. A participant mentioned that *"The timeline definitely helps because it shows where [the investigation] started and what the thought process was and how [the graph] was developed."* In the six handoff sessions, four participants consistently used the Timeline, one used it sporadically, and one not at all.

For knowledge transfer question (S1-S7), the most interesting difference was that participants found Baseline unhelpful for understanding prior investigation, while the opposite was true for KTGraph (S3). Also, participants in Baseline had less awareness of what their partners had done (S1) and tended to double-check previous discoveries (S5).

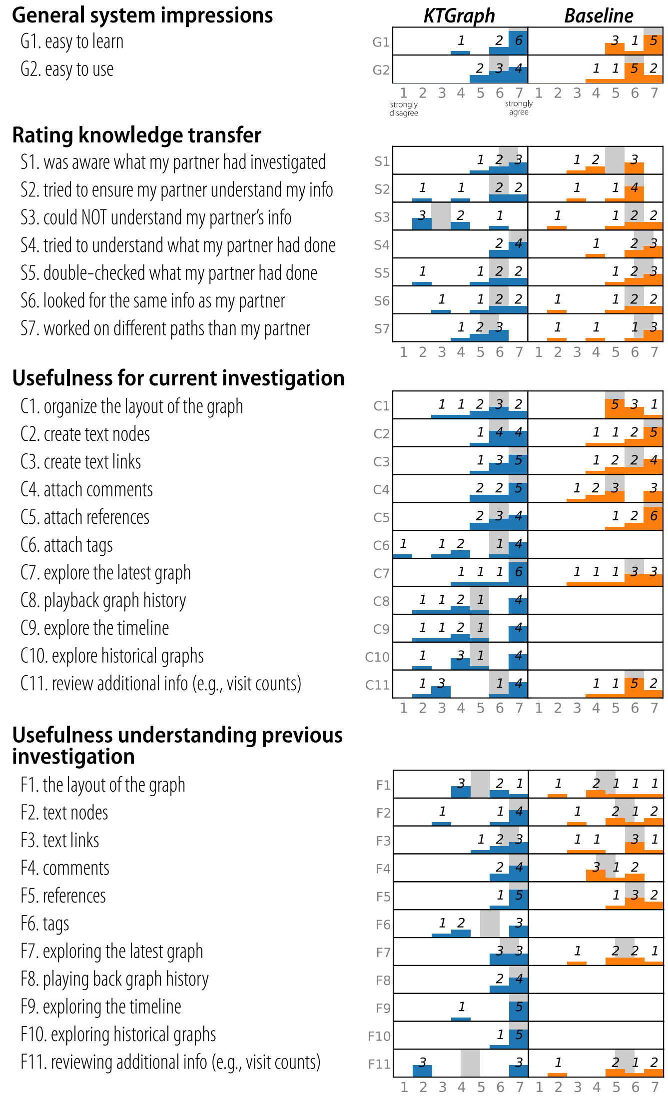

Fig. 9. Participants' ratings in the questionnaire in Phase 2. Median ratings are indicated in grey.

Participants also provided some general feedback about interface usability. For example, one participant requested arc links to avoid overlap, and two requested *"being able to mark and filter documents in the list [on Dataset Panel]."* Several participants also wanted to attach references to comments in the graph and vice versa because *"you will know which comment addresses which reference"*, which is an interesting feature to implement in the future.

6.5 Analysis of Handoff Strategies

Combining the results from Phase 1+2, we were interested in understanding how handoff strategies affected teams and their use of the tools. We derived an initial set of strategies by combining interview notes and a first video-coding. We then re-analyzed and consolidated the found strategies in a second video-coding pass. From this two-pass analysis, we identified five major strategies that participants used to assimilate knowledge from partial findings.

- *Starting over.* Participants generally ignored the existing graph and started from the first clue to build their own graphs. Later, they identified duplications or related segments between their own and the prior graphs, and tried to merge them together. However, they mainly focused on their own graphs throughout the process.
- *Random access.* Participants randomly visited graph elements that looked interesting, and then tried to understand the whole picture. In KTGraph, this strategy was largely influenced by the tagging feature, which drew participants' attention first.

Table 1. Number of participants for each handoff strategy (some participants used more than one strategy, thus the scores may due to multiple factors). For Phase 1, there were in total 10 handoffs for each condition (KTGraph or Baseline), reported below as *# of participants (mean debriefing score)*; and for Phase 2, there were 6 handoffs each, reported below as *# of participants (Δmean debriefing score)*.

Strategy	Phase 1		Phase 2	
	KTGraph	Baseline	KTGraph	Baseline
Starting over		1 (4.0)	2 (-0.5)	2 (+0.0)
Random access	1 (7.0)	5 (2.6)		3 (+1.3)
Naive browsing	2 (10.0)	2 (4.0)		
Hubs and bridges	1 (6.0)	5 (4.6)		1 (+0.0)
Tracing from the origin	9 (8.0)	1 (0.0)	5 (+1.0)	2 (+2.0)

- *Naive browsing.* Participants explored the entire graph with a natural order, such as from top–bottom or left–right. Then, they combined and sorted all the information to determine which directions to pursue. Similar to *Random access*, tags were helpful for suggesting specific nodes or links to inspect more carefully.
- *Hubs and bridges.* Participants browsed the parts of the graph that had special topological features, such as the densest parts of the graph (e. g., a node with many references or many links) or bridges of the graph (e. g., a node linking two dense components), and tried to connect them to form a bigger picture. In both conditions, the fact that references, comments, and tags changed the topology of the handoff graph helped participants to identify critical information.
- *Tracing from the origin.* Participants tried to understand the thought process of prior analysts by following the graph structure from a logical starting point. This strategy was conducted using different actions in two conditions. In KTGraph, participants used the timeline to browse the graph following the process from the beginning. In Baseline, participants had to guess the development of graphs and trace it manually (e. g., `intoxicated driver` and `flowers` nodes were usually considered to start with).

The existing graphs were generally treated similarly with all strategies, except for *Starting over*. In general, participants first discovered a few interesting points in the existing handoff graph, and then further extended those parts based on their continued investigation. In *Starting over*, participants practically ignored the existing graph and focused on building their own parallel graph.

The handoff strategies we observed are similar to strategies previously identified by Kang et al. for individual investigators when processing information and deciding what to do next [24]. The identified strategies, "Build from detail", "Find a Clue, Follow the Trail", and "Hit the keyword" closely align to our *Naive browsing*, *Tracing from the origin*, and *Hubs and bridges*, respectively. This suggests that strategies of handoff are closely related to data analysis and sensemaking in more traditional, non-collaborative settings.

Table 1 shows the number of participants who used each strategy in both phases of the study. In the following, we discuss the prevalence of the strategies in each phase in detail.

Phase 1 analysis. In the Baseline condition, we observed that participants used a wider combination of strategies. Some participants appeared to become frustrated with *Random access* and switched to *Starting over*, while others turned to *Random access* after *Naive browsing* and *Hubs and bridges* proved too difficult. Overall, the most common strategies were: *Random access* and *Hubs and bridges* (Table 1). However, these resulted in relatively low debriefing scores. Without access to the Timeline, it seemed that *Hubs and bridges* was the most effective strategy. The five participants who used it achieved higher than average debriefing scores (42%, 4.6/11, CI [2.3, 6.8]), however, still lower than the average score of participants using the KTGraph (71%, 7.8/11, CI [6.5, 9.1]). *Tracing from the origin* was the least effective strategy in the Baseline condition.

We observed an interesting difference in the application of *Tracing from the origin* between the Baseline and KTGraph conditions. In Baseline, this strategy was less effective because participants would have to identify the first clue in the graph to determine a starting point.

In contrast, in the KTGraph condition, the Timeline Panel enabled participants to identify the temporal starting point of the investigation and thus used the topology of the graph within a temporal context. Nine out of ten participants ended up using *Tracing from the origin* via the timeline. Six started with this strategy, while three others began with other strategies and then later switched to using *Tracing from the origin*. This strategy proved the most effective in the KTGraph condition and we observed far less switching of strategies. Two participants achieved the highest debriefing scores (10.0) used *Tracing from the origin* + *Naive browsing*, and one participant with the lowest score (6.0 with *Hubs and bridges*) was the only one who did not use *Tracing from the origin*. These findings mirror the results of Kang et al. [24], who found "Find a Clue, Follow the Trail" generally led to positive outcomes for non-collaborative investigation. However, more studies are needed to confirm these results and further examine how investigators combine strategies in various scenarios.

Phase 2 analysis. As shown in Table 1, we also observed a high adoption rate (5/6 participants) for *Tracing from the origin* via the timeline in KTGraph. Two participants employed *Starting over*, but one got stuck and switched to *Tracing from the origin* with timeline later. In the Baseline condition, a larger variety of strategies were used by participants (which is consistent with the results in Phase 1), of which *Random access* seemed to be the most chosen strategy. This may indicate that participants felt overwhelmed by some of the graphs and thought less logically about how to utilize them. Two of the six participants did try to use *Tracing from the origin*, which is a more logical approach, but encountered difficulty and switched to *Starting over* and *Random access*. This is because the *Tracing from the origin* search process relied on the graph topology alone in Baseline. If the topology was significantly complex, participants tended to switch to other strategies. Further, from Fig. 9, we observe that participants generally dealt with larger handoff graphs with more nodes and links. Participants tended to build graphs for their own investigations rather than utilizing information in the handoff graphs due to less understanding, which generated redundancy in their graphs.

We also evaluated the effectiveness of the strategies in Phase 2 based on the increase of the debriefing scores after each handoff. In KTGraph, participants who used *Tracing from the origin* had a mean debriefing score increase of 1.0, larger than those who used *Starting over*, -0.5 (i.e., actually decreasing). For the Baseline condition, we found no major difference between the performance of participants using different strategies. On average, participants performed worse in Baseline (a mean debriefing score of 1.7) than KTGraph (a mean debriefing score of 5.0). It is important to note that these results are only suggestive, because we observed large variances in Phase 2 due to fewer data points and very different handoff graphs (Fig. 8). Better understanding of the handoff graphs in general does not necessarily lead to higher debriefing scores, because the previous participant could point into the wrong directions of the investigation. Thus, more user studies are needed to further evaluate the effectiveness of strategies in such experimental settings.

Overall, by combining both phases of the study, these findings further support that exposure to the temporality of investigations is a critical factor to effective handoff externalizations. Without temporality, other methods of reducing graph readability may be necessary to support comprehension of externalizations.

7 Discussion

In this section, we discuss the key aspects we learned from our study, including the effects of interface features that support handoff, limitations of the current prototype, and our generalizable observations.

7.1 Effects of KTGraph Features

The results of our study demonstrate the benefits of temporal awareness in supporting handoff during asynchronous collaborative investigations.

The higher handoff and debriefing scores in the KTGraph condition suggest that participants more effectively assimilated the work of previous investigators and leveraged this understanding leading to better investigative outcomes. In addition, three participants interviewed in the Baseline condition requested features similar to tagging and interactive history features: highlighting important nodes, adding timestamps to nodes, and showing how the graph was created. The Timeline Panel also best supported the most effective and logical handoff strategy: *Tracing from the origin*. However, more experiments are required to extend our findings on how the handoff strategies affect graph usage.

The results of our study suggest that although tags were rated useful, they were not frequently used. Although we attempted to make tagging easier by integrating it into the node and link creation process (Fig. 2E), many participants were still reluctant to use the tags. Based on their feedback, the main reasons were that tagging *"breaks the flow of thinking"* (deciding which tag to use) and *"wondering why"* (understanding why a tag was used). Reluctance to use tags may also be attributed to the lack of further features that used the tags, such as tag-based filtering of the graph. We were also surprised that no participants created custom tags. Tags have been shown to be powerful when used consistently [52]. Further research to encourage tag usage is warranted, for example methods that reduce barriers to use by mitigating workflow interruption or through automatic suggestions.

Interestingly, adding comments, which requires much more effort than tagging, was considered more useful, for understanding both previous and current investigations (Fig. 9). This may be because comments are richer and interpreting their meaning is less ambiguous. The flexibility of comments resulted in creative uses. For example, one participant in the Baseline condition created a `general` node and attached a summary report in a comment. This was viewed as useful by subsequent participants and supported the *Tracing from origin* strategy. A generalized feature to explicitly communicate starting points may be beneficial to supporting handoff.

7.2 KTGraph Limitations

Although shown effective in our study, the current KTGraph prototype can be further improved. First, the graph visualization does not scale well as the size of the graph grows. After the last participant of each investigation task in Phase 2, the graph became quite complex (with an average of 20.5 nodes, 20.2 links, 29.8 references, and 7.3 comments; also see Fig. 8), and thus subsequent handoffs would become more and more difficult due to visual clutter. Existing methods, such as searching and filtering graph elements or methods for collapsing nodes and links [6, 55] could address these scalability issues.

Moreover, there is a need for better support of workspace organization, especially when the graph is large. From our observations in the studies, participants did not tend to manually organize the graph *"[...] unless I was forced to, because it takes too much effort."* Thus, automatic or semi-automatic graph layout methods based on certain similarity measures among nodes, links, comments, and references (as in e. g., AnnotationGraphs [55]) are worth further consideration.

7.3 Generalization

In our study we focused on investigative document analysis to evaluate KTGraph. Yet, the features that support this task could also be applied to a variety of other exploratory data analysis scenarios where investigators generate hypotheses. For example, our tool can support workflows of collaborative visual analysis tools that involve selection and annotation of data points, such as for time-series data [55] or charts [13, 52]. With an extension of LCW support like CLIP [29], KTGraph may be used in both synchronous and asynchronous collaborations. As the general approach of externalizations is common, we also expect that KTGraph is also suitable to other application domains.

Similar to the collaboration models proposed by Mahyar et al. [29] and Gava et al. [9] for synchronous scenarios, we also observed that awareness of past actions played a critical role in improving the effectiveness of asynchronous collaborations. In particular, awareness facilitated communication and coordination which are two key aspects in collaboration. As in Mahyar et al.'s model [29], we also acknowledge that users' *externalizations* of their findings and thought processes increased the *awareness* of their work for other team members, in our case, facilitating handoffs and ultimately generating better results. Yet, our work also cannot entirely be described by the previous models for synchronous collaboration: First, in our studies, there was no direct *communication* or *discussion* between different team members. The graph was the only means for communication, through its elements (i. e., nodes, links, tags, comments, and references), topological structure, and creation process (via the timeline). Thus, understanding the graph from previous investigations became essential for handoff. This infers that, in collaborative settings such as ours, the externalization itself and the process of generating it are of a greater importance for capturing thought processes and serve as a means for explicit or implicit communication. Second, *coordination* across participants was achieved sequentially in our studies, different from co-located scenarios where coordination is achieved in real-time and in parallel. This indicates that incorrect directions or misunderstandings from previous sessions misled following investigators, which could be amplified and resulted in worse outcomes in the end. This observation reveals the real-world challenges of handoffs and the importance of offering better support to transferring knowledge through the process.

8 Study Limitations

While our study has provided several considerations for the design of externalizations in handoff scenarios, our results should be read in light of the study's limitations. Certainly, more participants could strengthen our results of Phase 2. As the handoff graphs that participants designed varied significantly it is possible that other handoff strategies may surface and that the description of the existing strategies can be further refined. This could particularly be the case with varied types of groups, where investigators know each other well or share common work styles or habits of thinking and recording.

As in related previous studies [20, 29], our participants, while familiar with research practices, were not trained investigators. Strategies of professional analysts may differ from the ones we observed, especially if specific work and documentation requirements are imposed, or if tacit domain knowledge influences information sharing. Conducting further studies with expert users in a specific application domain is an excellent opportunity for future work. It will also be interesting to follow-up with work on hybrid scenarios in which externalizations such as KTGraph are used within brief periods of face-to-face handoff. Certainly, the research space for collaborative handoff is still large. We provide starting considerations for the design of handoff externalizations that will hopefully be extended by future work in this direction.

9 Conclusion and Future Work

We described KTGraph, a graph-based tool to externalize findings in analysis scenarios. KTGraph was designed with a number of techniques to facilitate the handoff of partial findings to subsequent investigators in asynchronous collaborative sensemaking. Specifically, our techniques support handoff through explicit user-annotations and implicit playback of the analytic process. Results of a user study in investigative document analysis suggest that temporal awareness is critical and effectively supports handoff. The study also identifies participants' strategies used in handoff and indicates that the most successful strategy is supported by temporal awareness.

In the future, we aim to extend the analytical capabilities of KTGraph to general data analysis scenarios, including charts and numerical data. We also plan to conduct more user studies to further evaluate the usefulness of KTGraph in handoff, for example, with more participants, deeper factorial design, or other collaborative analysis scenarios. In addition, we want to further enhance the tool by more effectively supporting scalability and workspace organization.

Acknowledgments

We thank John Hancock, Michelle Annett, and Jonathan Deber for their help and valuable input. This work was partially supported by Mitacs Accelerate Program and CPER Nord-Pas de Calais / FEDER DATA Advanced Data Science & Technologies 2015-2020.

References

[1] J. Alexander, A. Cockburn, S. Fitchett, C. Gutwin, and S. Greenberg. Revisiting read wear: analysis, design, and evaluation of a footprints scrollbar. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pp. 1665–1674. ACM, 2009. doi: 10.1145/1518701.1518957

[2] S. K. Badam and N. Elmqvist. Polychrome: A cross-device framework for collaborative web visualization. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS)*, pp. 109–118. ACM, 2014. doi: 10.1145/2669485.2669518

[3] E. A. Bier, S. K. Card, and J. W. Bodnar. Principles and tools for collaborative entity-based intelligence analysis. *IEEE Transactions on Visualization and Computer Graphics*, 16(2):178–191, 2010. doi: 10.1109/TVCG.2009.104

[4] S. P. Callahan, J. Freire, E. Santos, C. E. Scheidegger, C. T. Silva, and H. T. Vo. Vistrails: Visualization meets data management. In *Proceedings of the 2006 ACM SIGMOD International Conference on Management of Data*, pp. 745–747, 2006. doi: 10.1145/1142473.1142574

[5] R. Chang, M. Ghoniem, R. Kosara, W. Ribarsky, J. Yang, E. Suma, C. Ziemkiewicz, D. Kern, and A. Sudjianto. Wirevis: Visualization of categorical, time-varying data from financial transactions. In *IEEE Symposium on Visual Analytics Science and Technology*, pp. 155–162, 2007. doi: 10.1109/VAST.2007.4389009

[6] Y. Chen, J. Alsakran, S. Barlowe, J. Yang, and Y. Zhao. Supporting effective common ground construction in asynchronous collaborative visual analytics. In *Proceedings of the Conference on Visual Analytics Science and Technology (VAST)*, pp. 101–110. IEEE, 2011. doi: 10.1109/VAST.2011.6102447

[7] H. Chung, S. Yang, N. Massjouni, C. Andrews, R. Kanna, and C. North. Vizcept: Supporting synchronous collaboration for constructing visualizations in intelligence analysis. In *Proceedings of the Symposium on Visual Analytics Science and Technology (VAST)*, pp. 107–114. IEEE, 2010. doi: 10.1109/VAST.2010.5652932

[8] W. Dou, D. H. Jeong, F. Stukes, W. Ribarsky, H. R. Lipford, and R. Chang. Recovering reasoning processes from user interactions. *IEEE Computer Graphics and Applications*, 29(3):52–61, 2009. doi: 10.1109/MCG.2009.49

[9] V. L. Gava, M. D. M. Spinola, A. C. Tonini, and J. C. Medina. The 3c cooperation model applied to the classical requirement analysis. *Journal of Information Systems and Technology Management*, 9(2):235–264, Aug. 2012. doi: 10.4301/s1807-17752012000200003

[10] G. Grinstein, T. O'Connell, S. Laskowski, C. Plaisant, J. Scholtz, and M. Whiting. Vast 2006 contest – a tale of alderwood. In *Proceedings of the Symposium On Visual Analytics Science And Technology*, pp. 215–216. IEEE, 2006. doi: 10.1109/VAST.2006.261420

[11] T. Grossman, J. Matejka, and G. Fitzmaurice. Chronicle: Capture, exploration, and playback of document workflow histories. In *Proceedings of the Symposium on User Interface Software and Technology (UIST)*, pp. 143–152. ACM, 2010. doi: 10.1145/1866029.1866054

[12] P. Hamilton and D. J. Wigdor. Conductor: Enabling and understanding cross-device interaction. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pp. 2773–2782. ACM, 2014. doi: 10.1145/2556288.2557170

[13] J. Heer and M. Agrawala. Design considerations for collaborative visual analytics. *Information Visualization*, 7(1):49–62, Mar. 2008. doi: 10.1057/palgrave.ivs.9500167

[14] J. Heer, J. Mackinlay, C. Stolte, and M. Agrawala. Graphical histories for visualization: Supporting analysis, communication, and evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1189–1196, Nov. 2008. doi: 10.1109/TVCG.2008.137

[15] J. Heer, F. B. Viégas, and M. Wattenberg. Voyagers and voyeurs: Supporting asynchronous collaborative information visualization. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pp. 1029–1038. ACM, 2007. doi: 10.1145/1240624.1240781

[16] R. J. Heuer, Jr. *Psychology of Intelligence Analysis*. Center for the Study of Intelligence, Central Intelligence Agency, 1999.

[17] P. Isenberg and S. Carpendale. Interactive tree comparison for co-located collaborative information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1232–1239, Nov. 2007. doi: 10.1109/TVCG.2007.70568

[18] P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen. Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, Oct. 2011. doi: 10.1177/1473871611412817

[19] P. Isenberg and D. Fisher. Collaborative brushing and linking for co-located visual analytics of document collections. In *Computer Graphics Forum*, vol. 28, pp. 1031–1038. Wiley Online Library, June 2009. doi: 10.1111/j.1467-8659.2009.01444.x

[20] P. Isenberg, D. Fisher, S. A. Paul, M. R. Morris, K. Inkpen, and M. Czerwinski. Co-located collaborative visual analytics around a tabletop display. *Transactions on Visualization and Computer Graphics*, 18(5):689–702, May 2012. doi: 10.1109/TVCG.2011.287

[21] P. Isenberg, A. Tang, and S. Carpendale. An exploratory study of visual information analysis. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pp. 1217–1226. ACM, 2008. doi: 10.1145/1357054.1357245

[22] M. R. Jakobsen and K. Hornbæk. Up close and personal: Collaborative work on a high-resolution multitouch wall display. *Transactions on Computer-Human Interaction (TOCHI)*, 21(2):11, Feb. 2014. doi: 10.1145/2576099

[23] N. Kadivar, V. Chen, D. Dunsmuir, E. Lee, C. Qian, J. Dill, C. Shaw, and R. Woodbury. Capturing and supporting the analysis process. In *Proceedings of the Symposium on Visual Analytics Science and Technology (VAST)*, pp. 131–138. IEEE, 2009. doi: 10.1109/VAST.2009.5333020

[24] Y.-a. Kang, C. Gorg, and J. Stasko. Evaluating visual analytics systems for investigative analysis: Deriving design principles from a case study. In *Proceedings of the Symposium on Visual Analytics Science and Technology (VAST)*, pp. 139–146. IEEE, 2009. doi: 10.1109/VAST.2009.5333878

[25] Y.-a. Kang and J. Stasko. Characterizing the intelligence analysis process: Informing visual analytics design through a longitudinal field study. In *Visual Analytics Science and Technology (VAST), 2011 IEEE Conference on*, pp. 21–30. IEEE, 2011. doi: 10.1109/VAST.2011.6102438

[26] P. E. Keel. Collaborative visual analytics: Inferring from the spatialorganization and collaborative use of information. In *Proceedings of the Symposium on Visual Analytics Science and Technology (VAST)*, pp. 137–144. IEEE, 2006. doi: 10.1109/VAST.2006.261415

[27] M. Krzywinski and N. Altman. Points of significance: Error bars. *Nature Methods*, 10(10), Oct. 2013. doi: 10.1038/nmeth.2659

[28] B. C. Kwon, W. Javed, S. Ghani, N. Elmqvist, J. S. Yi, and D. S. Ebert. Evaluating the role of time in investigative analysis of document collections. *IEEE Transactions on Visualization and Computer Graphics*, 18(11):1992–2004, 2012. doi: 10.1109/TVCG.2012.89

[29] N. Mahyar and M. Tory. Supporting communication and coordination in collaborative sensemaking. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1633–1642, Dec. 2014. doi: 10.1109/TVCG.2014.2346573

[30] W. McGrath, B. Bowman, D. McCallum, J. D. Hincapié-Ramos, N. Elmqvist, and P. Irani. Branch-explore-merge: facilitating real-time revision control in collaborative visual exploration. In *Proceedings of the Conference on Interactive Tabletops and Surfaces (ITS)*, pp. 235–244. ACM, 2012. doi: 10.1145/2396636.2396673

[31] M. R. Morris and E. Horvitz. SearchTogether: An interface for collaborative web search. In *Proceedings of Symposium on User Interface Software and Technology (UIST)*, pp. 3–12. ACM, 2007. doi: 10.1145/1294211.1294215

[32] S. A. Paul. *Understanding Together: Sensemaking in Collaborative Information Seeking*. PhD thesis, The Pennsylvania State University, 2010.

[33] S. A. Paul and M. C. Reddy. Understanding together: Sensemaking in collaborative information seeking. In *Proceedings of Conference on Computer Supported Cooperative Work (CSCW)*, pp. 321–330. ACM, 2010. doi: 10.1145/1718918.1718976

[34] P. Pirolli and S. Card. Information foraging. *Psychological Review*, 106(4):643–675, Oct. 1999. doi: 10.1037/0033-295X.106.4.643

[35] B. Prue, M. Jenkins, L. D. Stern, and J. Pfautz. Overcoming limited collaboration channels in distributed intelligence analysis: Visualization tools and design seeds. In *Proceedings of the Intelligence and Security Informatics Conference (JISIC)*, pp. 17–24. IEEE, 2014. doi: 10.1109/JISIC.2014.13

[36] 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. *Transactions on Visualization and Computer Graphics*, 22(1):31–40, Jan. 2016. doi: 10.1109/TVCG.2015.2467551

[37] A. C. Robinson. Collaborative synthesis of visual analytic results. In

Proceedings of the Symposium on Visual Analytics Science and Technology (VAST), pp. 67–74. IEEE, 2008. doi: 10.1109/VAST.2008.4677358

[38] D. M. Russell, R. Jeffries, and L. Irani. Sensemaking for the rest of us. In Proceedings of the ACM CHI Sensemaking Workshop, 2008.

[39] D. M. Russell, M. J. Stefik, P. Pirolli, and S. K. Card. The cost structure of sensemaking. In *Proceedings Conference on Human Factors in Computing Systems (CHI/INTERACT)*, pp. 269–276. ACM, 1993. doi: 10.1145/169059.169209

[40] A. Sarvghad and M. Tory. Exploiting analysis history to support collaborative data analysis. In *Proceedings of the Graphics Interface Conference (GI)*, pp. 123–130. Canadian Information Processing Society, 2015. doi: 10.20380/GI2015.16

[41] A. Sarvghad, M. Tory, and N. Mahyar. Visualizing dimension coverage to support exploratory analysis. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):21–30, Jan. 2017. doi: 10.1109/TVCG.2016.2598466

[42] J. Scholtz and A. Endert. User-centered design guidelines for collaborative software for intelligence analysis. In *International Conference on Collaboration Technologies and Systems (CTS)*, pp. 478–482. IEEE, 2014. doi: 10.1109/CTS.2014.6867610

[43] N. Sharma. Sensemaking handoff: When and how? *Proceedings of the American Society for Information Science and Technology*, 45(1):1–12, June 2008. doi: 10.1002/meet.2008.1450450234

[44] N. Sharma and G. Furnas. Artifact usefulness and usage in sensemaking handoffs. *Proceedings of the Association for Information Science and Technology*, 46(1):1–19, Nov. 2009. doi: 10.1002/meet.2009.1450460219

[45] J. Stasko, C. Görg, and Z. Liu. Jigsaw: Supporting investigative analysis through interactive visualization. *Information Visualization*, 7(2):118–132, June 2008. doi: 10.1057/palgrave.ivs.9500180

[46] J. J. Thomas and K. A. Cook, eds. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society, 2005.

[47] F. B. Viégas, M. Wattenberg, F. Van Ham, J. Kriss, and M. McKeon. ManyEyes: A site for visualization at internet scale. *Transactions on Visualization and Computer Graphics*, 13(6), Nov. 2007. doi: 10.1109/TVCG.2007.70577

[48] K. Vogt, L. Bradel, C. Andrews, C. North, A. Endert, and D. Hutchings. Co-located collaborative sensemaking on a large high-resolution display with multiple input devices. In *IFIP Conference on Human-Computer Interaction (INTERACT)*, pp. 589–604. Springer, 2011. doi: 10.1007/978-3-642-23771-3_44

[49] M. Wattenberg and J. Kriss. Designing for social data analysis. *Transactions on Visualization and Computer Graphics*, 12(4):549–557, July-Aug. 2006. doi: 10.1109/TVCG.2006.65

[50] K. E. Weick. *Sensemaking in Organizations*. Foundations for Organizational Science. SAGE, 1995.

[51] W. Willett, J. Heer, and M. Agrawala. Scented Widgets: Improving navigation cues with embedded visualizations. *Transactions on Visualization and Computer Graphics*, 13(6):1129–1136, novdec 2007. doi: 10.1109/TVCG.2007.70589

[52] W. Willett, J. Heer, J. Hellerstein, and M. Agrawala. CommentSpace: Structured support for collaborative visual analysis. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pp. 3131–3140. ACM, 2011. doi: 10.1145/1978942.1979407

[53] W. Wright, D. Schroh, P. Proulx, A. Skaburskis, and B. Cort. The sandbox for analysis: Concepts and methods. In *Proceedings of the Conference on Human Factors in Computing Systems (CHI)*, pp. 801–810. ACM, 2006. doi: 10.1145/1124772.1124890

[54] K. Xu, S. Attfield, T. J. Jankun-Kelly, A. Wheat, P. H. Nguyen, and N. Selvaraj. Analytic provenance for sensemaking: A research agenda. *IEEE Computer Graphics and Applications*, 35(3):56–64, May 2015. doi: 10.1109/MCG.2015.50

[55] 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 Transactions on Visualization and Computer Graphics*, 23(1):261–270, Jan. 2016. doi: 10.1109/TVCG.2016.2598543