TalkTraces: Real-Time Capture and Visualization of Verbal Content in Meetings

Senthil Chandrasegaran
Department of Computer Science,
University of California, Davis
schandrasegaran@ucdavis.edu

Chris Bryan
CIDSE
Arizona State University
cbryan16@asu.edu

Hidekazu Shidara
Department of Computer Science,
University of California, Davis
hshidara@ucdavis.edu

Tung-Yen Chuang
Department of Computer Science,
University of California, Davis
tydchuang@ucdavis.edu

Kwan-Liu Ma
Department of Computer Science,
University of California, Davis
ma@cs.ucdavis.edu

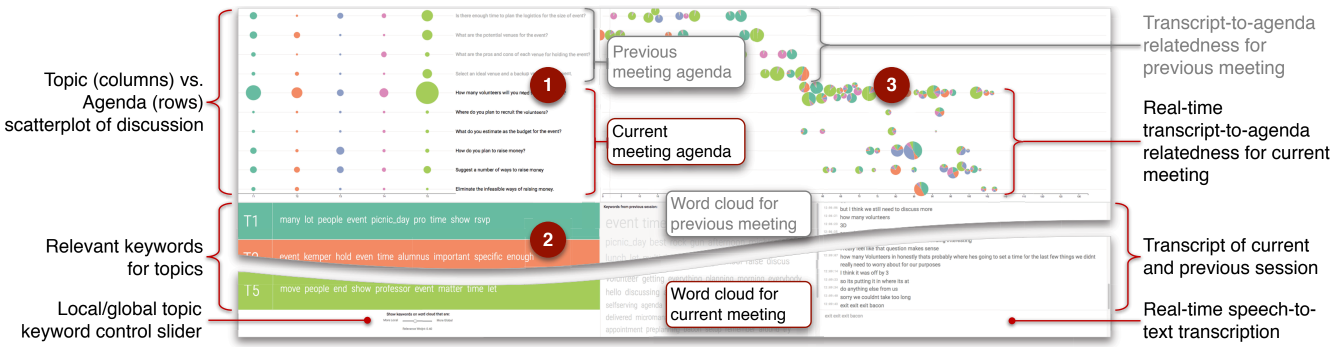

Figure 1: The final iteration of TalkTraces shows (1) agenda items from the previous and current meeting, (2) topics computed from the previous meeting transcript, and (3) transcript lines rendered as pie charts reflecting topic distribution, aligned with the most related agenda item. Streaming audio transcribed from the current meeting updates the visualizations in real time.

ABSTRACT

Group Support Systems provide ways to review and edit shared content during meetings, but typically require participants to explicitly generate the content. Recent advances in speech-to-text conversion and language processing now make it possible to automatically record and review spoken information. We present the iterative design and evaluation of TalkTraces, a real-time visualization that helps teams identify themes in their discussions and obtain a sense of agenda items covered. We use topic modeling to identify themes within the discussions and word embeddings to compute the discussion "relatedness" to items in the meeting agenda. We evaluate TalkTraces iteratively: we first conduct a comparative between-groups study between two teams using TalkTraces and two teams using traditional notes, over four sessions. We translate the findings into changes in the interface, further evaluated by one team over four sessions. Based on our findings, we discuss design implications for real-time displays of discussion content.

CHI 2019, May 4–9, 2019, Glasgow, Scotland UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300807

CCS CONCEPTS

• **Human-centered computing** → **Information visualization**; *Computer supported cooperative work.*

KEYWORDS

Collaboration, Real-Time Visualization, Streaming Data

ACM Reference Format:
Senthil Chandrasegaran, Chris Bryan, Hidekazu Shidara, Tung-Yen Chuang, and Kwan-Liu Ma. 2019. TalkTraces: Real-Time Capture and Visualization of Verbal Content in Meetings. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland UK*. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3290605.3300807

1 INTRODUCTION

The notion of "smart meeting spaces"—interactive software systems that capture, analyze, and represent the information discussed during meetings—has captured the imaginations of multimedia researchers for over two decades. With its origins in Vannevar Bush's *Memex* [9], the idea of a smart system for capturing and retrieving personal and/or group information has evolved over the years. These systems seek to address common pitfalls in group meetings such as non-equitable participation, lack of focus on an agenda, and inordinate focus on irrelevant matters [40]. Examples of such systems include special-purpose meeting rooms with workstations, cameras, and multimedia devices [12] and collaborative tabletop systems for multi-user shared content [25]. Typical smart meeting environments capture and retrieve information that is intentionally created by users. However, capture and representation of verbal utterances in real time can help users keep track of their discussion with little extraneous effort.

In addition to maintaining the real-time representation of the discussion, there is often a need to determine whether the participants of itemized meetings are staying on agenda. Alternately, when the meeting is an open-ended brainstorming session, it would be useful to determine if ideas are being positively or negatively received. Such information can help moderators intervene and steer discussions towards a more desirable outcome. Recent advances in speech recognition and natural language processing (NLP) have made the streaming capture of spoken information feasible. However, designing visualizations that help users glean information at a glance—while also allowing more complex exploration of the data—is a challenge that has not yet been addressed.

In order to address these needs, we present TalkTraces, a multi-view, real-time visualization of conversations that presents a thematic overview of previous and ongoing meetings. We posit that using real-time information displays will aid awareness of the discussion, help participants reflect on their ongoing and previous discussions, and serve as a source of reminders and inspiration for discussion points. In this paper we describe two iterations of TalkTraces:

Iteration 1: A real-time visualization that focuses on discussion *topics* revealed through topic modeling of previous and current discussion content.

Iteration 2: A real-time visualization that indicates the ongoing discussion's relevance to *agenda items*. We compute this relevance using cosine similarity between vector representations of agenda items and speech instances.

Our primary goal is to help users maintain an overview of current and prior discussion content. By overview, we mean in the context of planned topics for the discussion (i.e. agenda), as well as topics that emerge from a discussion upon review. TalkTraces allows participants to answer questions such as *"What topics recur across meetings?"* and *"When did we discuss this agenda item?"*

TalkTraces provides additional views with increased detail, including an automated word cloud of the discussion and a speech transcript for reviewing the actual discussion. Users can proactively engage the display, filtering discussion content to answer questions such as *"What part(s) of the discussion does a particular topic represent?"* and *"in what context was (a given word/phrase) discussed in a particular meeting?"* These views can help identify patterns in the way meeting participants work together.

We evaluated iteration 1 of TalkTraces through a qualitative, longitudinal user study comparing two groups that use TalkTraces and two baseline groups that use standard pen-and-paper. Both groups were studied over four meeting sessions. We collected participant feedback on the meeting and on the interface and performed qualitative coding of their video-recorded meetings to identify actions such as referring to notes and viewing the display. For iteration 1, participants referred more to concrete information on the display, such as the transcript and word cloud, and found the topic-based views too abstract. To address the abstractness issue, we redesigned the interface to center around an agenda-focused visualization, with the topic-based visualization providing auxiliary information. A follow-up study of iteration 2 with one group of participants over four sessions indicated that while topic-based views were primarily retrospective for participants, the new agenda-based views helped them quickly identify unaddressed discussion points and helped them think of new topics to discuss.

Based on our experience in creating and evaluating the two TalkTraces iterations, we propose a set of generalized approaches for designing visualizations that aid real-time awareness and recall of group discussion content. In summary, the contributions of this paper are: (a) identification of requirements for maintaining participants' awareness of discussion content, (b) the design and implementation of TalkTraces, an interactive, real-time information display that uses topic modeling and word embeddings to represent contextualized discussion content, and (c) implications for creating meeting-awareness visualizations, based on the iterative development and evaluation of TalkTraces.

2 BACKGROUND AND RELATED WORK

TalkTraces integrates speech recognition, natural language processing, topic modeling, and word embedding in its data processing pipeline. Though it shares features from smart meeting rooms, it is designed to exist as a peripheral meeting component. The visualization design shares characteristics with high-throughput textual displays based on Pousman and Stasko's taxonomy [43], in that it steadily conveys a high volume of information.

Smart Meeting Environments

Smart meeting rooms (SMRs) are physically and/or digitally shared spaces that support group meetings and/or collaborative activities [67]. The concept of SMRs originated with Xerox PARC's "media spaces" in the 1980s [6]. Given that a significant percentage of meetings fail to achieve their goals [19], there is continued interest in SMRs even today. SMRs often integrate sensors such as audio and video recorders to store content for later playback. Other devices, including digital whiteboards, shared desktop software applications, and sensory tracking equipment, provide additional streams of information and meta-information content. Smart meeting software such as meeting browsers [65] aid analysis, review, and summarization of the meeting.

Two recent survey papers [20, 67] provide overviews of the current research, technologies, and trends for SMRs. This includes the design of architectures and systems that provide meeting capture, meeting recognition, semantic processing, and evaluation methods. Meeting recognition and semantic processing (respectively, forms of low- and high-level processing) rely on techniques from image and speech processing, computer vision, human-computer interaction, and ubiquitous computing to successfully analyze, synthesize, index, annotate, and display the captured meeting content [20].

Today, SMRs are connected to the broader ideas of Ambient Intelligence and the Internet of Things, where interconnected digital computing devices and sensors in our physical world interact with us intelligently and unobtrusively [44, 66]. Commercial, cloud-connected SMR solutions are available today via technical organizations such as Cisco [13], Intel [27], and Microsoft [38]. Notably, while the devices themselves collect data automatically and in real-time, SMRs typically require active user engagement and interaction.

Topic Modeling Approaches for Text Data

Meeting recognition is responsible for low-level analysis of captured data, such as speaker recognition and detection of structural features. In particular, topic modeling is a (generally unsupervised) machine-learning approach for discovering abstract themes ("topics") within a collection of one or more text documents. This is done as a probabilistic approach, whereby classification is accomplished based on the maximal likelihood that an element bins to a group or cluster [3]. Common algorithms for topic modeling include latent Dirichlet allocation (LDA) [5], Pachinko allocation [34], and hidden Markov models [22]. For these approaches, the order of documents does not matter, an assumption we rely on for TalkTraces as topics in meeting contexts may be revisited at later times throughout the meeting or in subsequent discussions. Recently, interactive human-in-the-loop topic modeling has gained traction, both with and without visualization support. Wang et al. [62] propose an evolutionary Bayesian rose tree algorithm that provides a multi-branch hierarchical representation of topics and their evolution over time. TopicStream [35] is a visual analytics system that generates a set of hierarchical nodes based on user-identified topics and visualizes temporal changes in these topics. More recently, interactive topic modeling approaches [51] refine topics based on user intervention, allowing for better interpretation. Since our interest is in real-time awareness of topics with minimal interaction, we use LDA for topic modeling in a streaming data context (see Sec. 3).

Another way of categorizing meeting discussions is based on relatedness to agenda items. A common approach is to represent words, phrases, or sentences as vectors; vector operations such as cosine similarity are used to determine relatedness [14]. Using word embeddings is a popular technique to determine this relatedness, with pre-trained models such as Word2Vec [42] and FastText [30]. For iteration 2 of TalkTraces, we use ConceptNet Numberbatch [53], a pre-trained word embedding that incorporates data from other word embeddings and open knowledge databases.

Visualizing Text and Meeting Data

Text visualization is an increasingly important subfield within information visualization [33]. When text data is time-varying, text labels, topic models, and word embeddings can be spatially plotted [15, 16, 54, 61, 64]. In these cases, the temporal evolution(s) of the data points are plotted along an axis to show the dynamics. Alternatively, text and topic data can be plotted via ordination or network-based plots [1, 10, 18, 21, 31, 58]. To accommodate streaming data, the visualization must update using animation or a display refresh.

For data collected in meetings, visualization-driven meeting browsers summarize discussions between participants. Similar to text visualization approaches, these systems generally highlight relevant topical or thematic content for later analysis [2, 11, 17, 26, 47, 56]. For instance, the CALO Meeting Assistant system [56] provides features similar to TalkTraces for post-meeting transcript review and topical analysis. In contrast, TalkTraces is active *during* the meeting, though it may also be used for off-line analysis.

Collaborative and Real-Time Visualization

TalkTraces is designed as a collaborative, real-time visualization. In collaborative visualization scenarios [28], data is accessed, viewed, interpreted, and interacted with by multiple persons—sometimes simultaneously and on the same screen. Real-time visualizations, specifically when they provide non-critical information, fall under the umbrella of ambient or peripheral displays, characterized as "portraying non-critical information on the periphery of a user's attention" [37, p. 169]. Such displays are normally designed for public deployment [50, 59, 60], though they can also be placed in home, corporate, and team environments [24, 29, 41, 48]. Vogel and Balakrishnan identify several design principles for public ambient visualizations, including calm aesthetics, comprehension, immediate usability, and shared use [60].

3 DESIGN

Our motivation and design guidelines stem from Nunamaker et al. [40]—specifically their list of group process gains and losses. We focus on group aspects that can be helped with timely information updates, such as *cognitive inertia*: a discussion getting fixated because team members do not want to speak unless their comments are related to the current discussion. *Concentration blocking*, *attention blocking*, and *failure to remember* are also group process losses that can be helped by providing timely information on the discussion. Nunamaker et al. [39] provide a thorough classification of group work along the orthogonal dimensions of productivity processes (thought, communication, and information access) and levels of work (individual, coordination, and group dynamics). We use this classification in our design to address information access at a group dynamics level. Our goal was to design and evaluate a technique to (a) capture the spoken content of a meeting, (b) provide an interpretable, conceptual representation of this content for the meeting participants, and (c) allow participants to view their discussion in real time, and interactively review it after the meeting.

While (a) can mostly be achieved through a standard speech-to-text conversion engine such as Google Speech API, (b) and (c) require an interface. For design guidelines, we again draw from Nunamaker et al. [39] to "keep the user learning curve short; use simple interfaces" (p. 178). Given that the primary activities of meeting attendees are talking, listening, recording (e.g. taking notes), and thinking, a complex interface requiring substantial interaction (or one that calls attention to itself) may be distracting. Information should be presented in a subtle and non-intrusive manner.

Rationale

While the kinds of computational and representational support by group support systems can be manifold, we restrict our scope to *spoken content* only. We do this for two reasons: first, we are interested in real-time information capture and representation in group meetings. Speech between group members occurs in most meetings regardless of the domain expertise or purpose of the group; it is the primary form of discourse. Second, while there are additional multimodal components in which group members engage, such as note-taking, sketching, or searching for information, properly integrating these modes requires consideration on public vs. private note-taking practices, which is not currently our focus. With these criteria in mind, we identify the following requirements for an interface that allows participants of a meeting to obtain a high-level perspective of the meeting's history and current state.

R1. Historical Recall: The interface should provide a way to recall the content of prior meetings without requiring a close examination of detailed content, such as a transcript.

R2. Conceptual Overview: The interface should provide a conceptual overview of the discussion at a glance without interrupting the meeting. Meetings are inter-personal activities, and a common interface for the group should follow a *"periphery-passive"* mode [55], i.e. provide relevant information without distracting the users.

R3. Situational Awareness: During the meeting, the interface should provide a view of "recent content" to reorient a participant if their attention drifts. The interface should keep users aware of changes in information (i.e. meeting content) through subtle changes in the visualization [43].

R4. Conversation Impetus: In order to overcome challenges such as cognitive inertia, we draw from *persuasive systems*: systems that can potentially change user behavior by displaying relevant information. Thus, when participants get fixated on topics, the display can provide information on which they can act, for example, move on to other topics that need attention.

The above-identified requirements inform the design of the visualization and interactive components of TalkTraces. In the following sections, we will outline the design and rationale behind both iterations of TalkTraces. Designs for iteration 2 were motivated by our findings in the user studies of iteration 1, so these findings are mentioned briefly in this section, and in detail in the user study and results.

Linguistic Processing

The data processing and visualization pipelines for both iterations of TalkTraces are shown in Fig. 2. We individually discuss each below.

Iteration 1: Identifying Discussion Topics: A thematic overview of meeting discussions is automated through topic modeling, since these methods can be used to identify themes present in a given text corpus of text. In particular, we chose Latent Dirichlet allocation (LDA) [5] as our statistical method. While LDA is meant primarily for collections of documents, and while there exist other topic modeling techniques such as Dynamic Topic Models [4] and Topics Over Time [63] for time-variant topics, we chose LDA for two reasons. First, most topic modeling techniques designed to track changes over time require certain assumptions that we cannot make for meetings, for instance the time window over which topics are expected to change. Second, we represent each topic with a set of extracted keywords using a technique created specifically for LDA [49]. This provides users the control to balance keyword ranking between frequency and relevance.

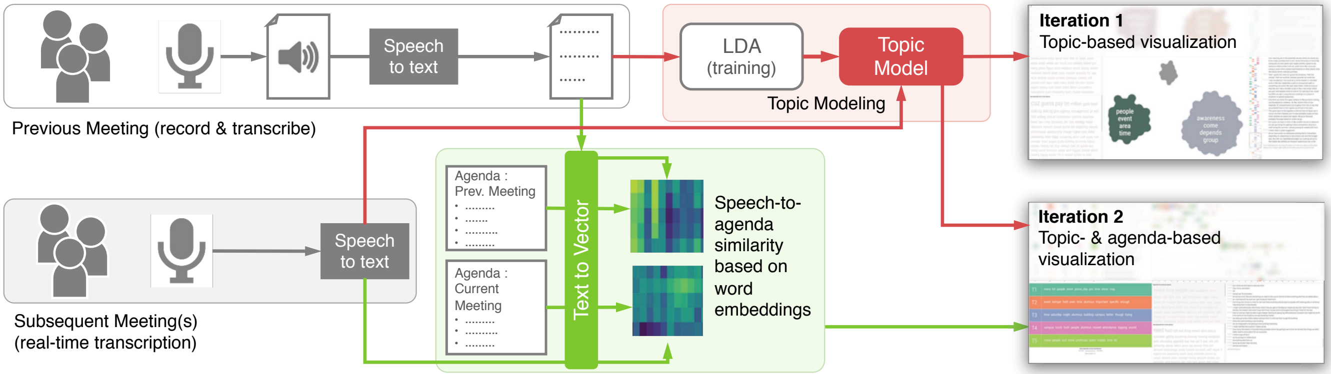

Figure 2: The data pipeline for each iteration. In iteration 1, the initial meeting is transcribed using speech-to-text algorithms, and with LDA, used to train a topic model (red boxes & arrows). The model is used to categorize utterances from subsequent meetings. Iteration 2 additionally uses word embeddings to compute speech-to-agenda-item similarities (green boxes & arrows). In both iterations, the visualizations update in real time to represent the results of these computations.

The initial meeting is recorded and transcribed; the transcript lines are grouped into sets of "documents." To reduce noise, we perform stop word removal, including filler words used in speech (e.g. *"like," "yeah,"* and *"right,"*). We then lemmatize words to remove differentiation between inflected forms. Dominant n-grams are also identified to preserve context (e.g. *"picnic_day"* in Fig. 1 and *"healthy_eating"* in Fig. 3). A document-term matrix is constructed with the resulting text and used to train an LDA model. The model categorizes the ongoing discussion in real time into the computed topics. This process is represented by the red arrows in Fig. 2.

A drawback of topic models is that new incoming words may not be clearly assignable to one of the predefined topics. We mitigate this by defining a probability threshold $p_T = min(p_{max1}, p_{max2}, \cdots p_{maxN})$ for N documents (lines in the transcript). p_{max} is calculated for each document as $max(p_1, p_2, \cdots p_k)$, where $p_1 \cdots p_k$ are probabilities that the document belongs to each of the K topics. If a new spoken line or set of lines is assigned a maximum probability lower than p_T, the line is assigned to a new, "unknown" topic. Thus users can determine in in real time the underlying topics/themes of their current discussion, and track digressions or potentially new topics (requirement **R3**).

Iteration 2: Agenda-Discussion Relatedness: Our study with iteration 1 of TalkTraces revealed that topic-based representations were found to be too abstract and difficult to follow. The "new/emergent topic" computation was not deemed useful since it did not help participants distinguish new topics from digressions. We also observed that participants preferred using a predefined meeting agenda (made available to them) to keep track of the discussion.

With iteration 2, in addition to topic modeling, we provide a way for users to categorize and keep track of their discussion with reference to the meeting agenda. We use word embeddings [57]—vector representations of words in a linguistic space—to determine the similarity between any spoken line and the agenda items for that meeting. For our computation, we use a pre-trained word embedding called ConceptNet Numberbatch [53]. We look up the vector representation for each word in an agenda item and average them to compute a single "agenda vector" representation for that item. Using a similar method, we obtain "sentence vectors" for each line in the transcript. We then compute a pairwise cosine similarity between every "agenda vector" and every "sentence vector", to find the agenda item most similar to each utterance (see green highlights in Fig. 2).

In comparison to the topic modeling approach, new words introduced in the discussion *can* be used to compute agenda similarities as long as they are part of the word embedding's vocabulary. As an additional advantage, the metric obtained is more tangible for the user as it directly references the meeting's predefined agenda items (requirements **R2** & **R3**).

Interface Design

TalkTraces is designed to display important information that keeps the user aware of the state of their discussion through subtle changes. Fig. 3 shows iteration 1 of our interface, while iteration 2 is shown in Fig. 1, Central to each interface is a high-level representation of the discussion (requirement **R2**), visualized as topics (iteration 1) and agenda items (iteration 2). Additional overviews are presented by interactive word clouds of the previous and current meetings. The actual transcript provides a "ground truth" (requirement **R1**).

Visualizations and Animations: We describe in the section on linguistic processing our topic modeling approaches used for iteration 1, and the additional agenda-to-discussion similarity computation using word embeddings for iteration 2. Here we discuss how the results are visualized.

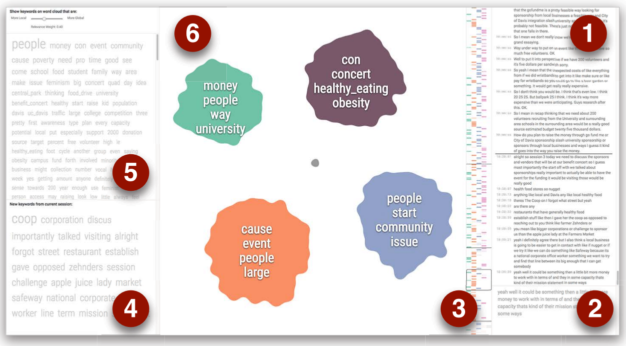

Figure 3: Iteration 1 of the interface features views seen in the final interface (Fig. 1) such as the transcript (1), real-time transcription (2), and word clouds of current (4) and previous (5) meetings. The main differences are a topics view (6) with relevant keywords overlaid on each topic "blob", and a heatmap (3) showing topic probability distributions for each transcript line. New utterances merge into each topic blob, or an "unknown" topic blob in the center.

Iteration 1: Topic-Focused Visualization: For a model with k topics in a discussion transcript, each group of spoken line(s) is assigned k probabilities, one for each of the topics. Interpreting this as a k-dimensional vector, we use multi-dimensional scaling (MDS) to represent each utterance as a point/node on a 2D plane. The nodes are clustered according to their most dominant topic, and each topic assigned a color from a perceptually-uniform color palette. The cluster size represents the prevalence of a topic across the discussion(s). In order to help users glean information at a glance, the cluster is then converted to a topic "blob" using SVG filters [8]. The user thus sees the nodes as an amorphous topic group, rather than a cluster of discrete utterances. Finally, to help the user assign meaning to each topic, we overlay the top 3 (for more topics) to 5 (for fewer topics) keywords from each topic over its corresponding blob. The resulting visualization gives a thematic overview of the entire discussion over multiple sessions (requirement **R2**).

New utterances from an ongoing meeting are assigned either to one of the existing topics or to a catch-all "unknown" topic as described in the linguistic processing section. As the group discussion progresses, each spoken utterance by a participant appears on the display as a node in the center of the screen, which then moves toward its topic centroid and "merges" with its topic blob. Changes in chroma and luminance across the CIELCh color space [46] highlight recently-addressed topics from previously-addressed and unaddressed topics, respectively (see *Iteration 1* in Fig. 4). At any point of the discussion, participants are provided with situational information (requirement **R3**): they can tell which topics have not been addressed at all, which topics were discussed earlier in the discussion, and which ones have been addressed most recently. Utterances that cannot be categorized in one of the predefined topics are assigned to the "unknown" topic blob representing emerging topics.

Iteration 2: Agenda-Focused Visualization: In evaluating iteration 1 (see Sec. 5), the topic bubbles were considered useful for determining topic dominance. Unfortunately, the animated transitions (of new utterances into topic bubbles) were often too fleeting or subtle to be noticed. While the animations were designed to be minimal (i.e. not attention-seeking), it resulted in a loss of situational awareness for many participants. Iteration 2 thus focuses on a more persistent encoding of each spoken line.

Like iteration 1, speech instances are encoded to nodes. However, we now use the *position* of the node to indicate both its relatedness to a specific agenda item and its recency (requirement **R3**). The topic blobs of iteration 1 are replaced with two views on either side of a list of agenda items (see Fig. 1). On the right of the agenda list, each sentence spoken in the previous and current meeting is aligned in a beeswarm plot against each corresponding agenda item. The density of the beeswarm gives a sense of agenda coverage. The utterances are arranged left-to-right in the order in which they are spoken. This provides both a thematic and a temporal sense of the discussion, as a user can see at a glance *what* agenda items have been discussed and *when* this happened (requirement **R1**). Each node in the beeswarm is overlaid with a pie chart indicating the topic distribution of that sentence in the transcript. The topic distribution for all the sentences related to an agenda item is aggregated and displayed (requirement **R2**) on the left of the agenda list as a topic vs. agenda scatterplot (Fig. 5).

New utterances in this view are already aligned vertically with the corresponding agenda item, but "fly in" (animate) from the right to occupy the target x-position (see *Iteration 2* in Fig. 4). This gives more time for the user to process the information, as throughout the animation it is clear to which agenda item the node belongs. Even after the animation, its position as the right-most node in the beeswarm plot allows users who miss the animation event to still maintain situational awareness.

Topic-Relevant Keyword View & Control: Dominant keywords for each topic are overlaid on each topic blob in iteration 1 of the interface, helping users interpret each topic. We use a list of words to represent topics as it has been shown to aid quick identification of topic themes [52]. Dominant keywords are ordered according to Sievert & Shirley's *relevance* metric [49] and displayed for each topic. The metric uses a weight parameter $\{\lambda\colon 0 \leq \lambda \leq 1\}$ to determine the relevance of a word to a topic. Selecting a lower λ emphasizes words that are more dominant within that topic, while higher λ emphasizes words that are more frequent across the transcript. We allow the user to fine-tune λ with a slider

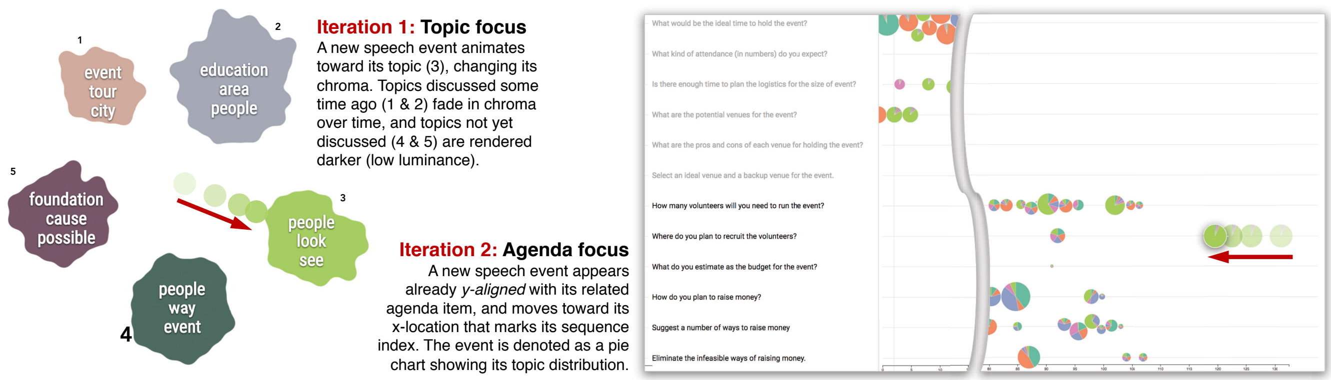

Figure 4: The difference in the visualizations and animations between iterations 1 and 2. Iteration 1's topic-focused view has new utterances move towards the topic blob to which the topic model has assigned it. The color changes of topic blobs in the CIELCh color space indicate topics addressed in the discussion and their recency. Iteration 2's agenda-focused view uses y-position to encode agenda categorization, and x-position to encode recency. The animation starts and ends in the same y-position, allowing more time for users to process the relevance of the latest utterance to the agenda.

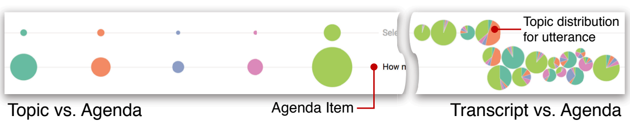

Figure 5: A detail of the agenda-centric display (iteration 2) shows speech instances as a beeswarm plot aligned to their most-relevant agenda items. Each speech instance is shown as a pie chart of probabilities that the instance belongs to each topic. The size of the pie is proportional to the utterance's similarity to the agenda item. A topic vs. agenda scatterplot on the left shows the topic distribution for each agenda, aggregated over the utterances. The colors are categorical, indicating topics, while size indicates the affinity of the agenda to that topic.

control until they are satisfied with the topic representation. This keyword display is maintained in iteration 2, but in the form of a simple topic-relevant keyword list (see Fig. 1)

Word Clouds: While LDA and word embeddings provide thematic overviews of the discussion content, simple word cloud displays provide a lower-level overview of discussion content (requirement **R2**). As seen in Fig. 3 for iteration 1 and Fig. 1 for iteration 2, we use two word clouds: one that represents the previous meeting(s), and another that represents the current meeting. More specifically, the second word cloud shows words that are unique to the current discussion, or new discussion items (requirement **R3**). Both word clouds are independent of topics, update dynamically based on new utterances, and use both word size and word order to indicate the frequency of occurrence.

Transcript View: Both iterations of TalkTraces feature a view of the discussion transcribed to text. The transcript updates in real time using streaming speech-to-text conversion. The transcript and real-time speech-to-text views provide a "ground truth" of the discussion, to help recall specific details (requirement **R1**). It also helps users interpret the thematic overviews based on the accuracy of the speech-to-text conversion.

Interactions: For an effective review of discussion content during or at the end of a meeting (or the beginning of a meeting to review the previous meeting's content), we incorporate linked views and interaction. The view linking is conceptually similar in both iterations, but adapts to the specific designs of each version. For both iterations, almost all the views are linked based on topics. Fig. 6 shows an instance of this interaction performed at the beginning of a meeting when there is no new content. Here, previous meeting content can be reviewed. Topic 2 is selected from the topic wordlist (or topic blobs in the case of iteration 1 of TalkTraces), highlighting corresponding portions of the pie charts in the beeswarm plots, showing that the topic is most prevalent in agenda item 1, and to a lesser extent, item 4. The topic-agenda scatterplot confirms this. The word cloud is filtered to highlight the most frequent words in this topic, as well as lines in the transcript that are allocated to this topic. Similar interactions can be performed to view keyword occurrences in the transcript or identify the related agenda item from a line in the transcript.

Implementation Details

The data processing pipeline is built using Python, including the Natural Language Toolkit [36] for text processing, the Gensim library [45] for topic modeling (LDA), and a python implementation of LDAVis [49] for the interactive topic keyword display. Transcript-to-agenda similarity is computed using the Conceptnet Numberbatch [53] word embeddings. We use the Google Speech API for real-time

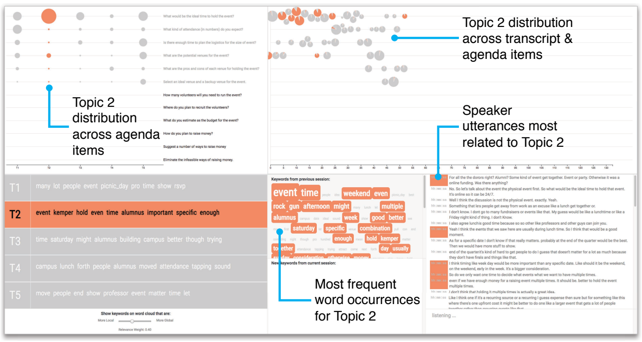

Figure 6: The interface can be used to review meeting content at the beginning of a meeting (previous content), or the end of the meeting. The interactive linked views show—using the topic colors—how selected topics are linked to agenda items and to lines in the transcript.

speech-to-text conversion. The interfaces are implemented in HTML5/JavaScript, with D3 [7] for the visualizations.

4 USER STUDY

We sought to examine whether a real-time visualization with the added functionality of skimming and filtering captured text content would aid collaborative work before, during, and at the end of a group discussion. Specifically, we wanted to study any effects TalkTraces might have on group awareness during ongoing discussions, and whether its real-time visualization provided reminders, cues, and/or new ideas for the participants to continue a thread of discussion or take it in new directions. We also wanted to examine whether TalkTraces could supplement traditional means of recording meetings—i.e., manual note-taking. To this end, we conducted a between-subjects, qualitative study comparing teams that used TalkTraces with teams that used a traditional meeting setting (pen and paper) over multiple meeting sessions. The study was conducted in two stages: a longitudinal, comparative study with 4 teams over 4 sessions using iteration 1 of the interface, and a follow-up study of iteration 2 with 1 team over 4 sessions.

Participants

We recruited 15 paid participants (4 female, 11 male) aged 18–44 years, 2 of whom were university employees in research labs, 2 postdoctoral researchers, 10 graduate students, and one undergraduate student. Twelve participants specialized in computer science and had minor to significant background in visualization. Participants were grouped into 5 teams of 3. Pilot studies revealed that the speech-to-text engine was error-prone with multiple accents in a group. Thus, of the 9 participants who used any one of the TalkTraces interfaces (iteration 1 or 2), 7 were native English speakers.

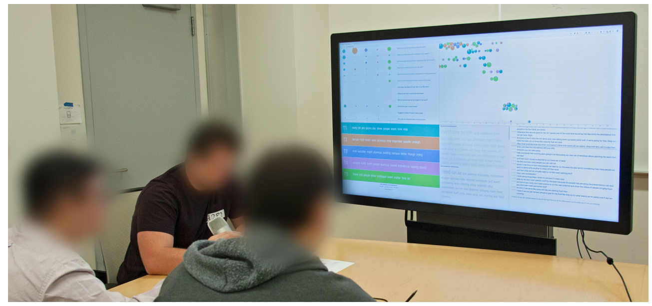

Figure 7: The setup of the user study showing participants using the final iteration of TalkTraces. A similar setup was used with the corresponding interface (see Fig. 3) for iteration 1. In the case of baseline teams, no display was used.

Study Setup

Every meeting session was performed in a dedicated meeting room (Fig. 7). For iteration 1, two teams (labeled V1 and V2) were assigned to use TalkTraces; the two other teams (B1 and B2) were assigned to a baseline “no-screens” condition. Pilot testing revealed that the Google Speech-to-Text service worked best with no overlapping conversation. To minimize cross-talk, we provided each team with a single shared microphone and asked participants to speak only when in possession of it. To mitigate confounding effects, we imposed the same setup for teams B1 and B2 as well. While this setup might potentially impact team dynamics, the constraint is due to technological limitations outside the scope of our work. Additionally, the goal of our study was not to study team dynamics; it was to see how teams used and interpreted the visualizations. For the follow-up study using iteration 2, the one team involved (team F) used the same setup as teams V1 and V2.

The interface for both iterations was displayed using a Chrome browser on a 55-inch LCD display (3840 × 2160 pixels), connected to a Laptop with 16 GB RAM, a 2.8 GHz processor, and a 4 GB GPU, running Mac OS 10.13. All teams regardless of condition were given printed copies of the agenda and additional sheets of paper for note-taking.

Procedure

All teams were given the same overall goal: identify a cause about which they felt strongly and plan an event to raise funds for this cause. This goal was split up into 4 agendas, one for each meeting. The first meeting was 30 minutes long, while subsequent meetings were 15 minutes each. Each group selected a team leader who had the additional responsibility of keeping track of agenda items and verbally summarizing each meeting session. For teams that used TalkTraces, the leader was tasked with identifying which of five topic models (trained assuming 3, 4, . . . 7 topics) best represented their prior discussion. They could do so using the linked views (Fig. 6) and using the relevance slider to choose a value for λ. Since the dataset—the transcript of prior meeting(s)—was small and familiar to the team leader, we relied on their intuition and knowledge of the previous discussion(s) to compensate for a lack of expertise in topic modeling. For teams V1, V2, and F, the first meeting (session 0) did not use the TalkTraces interface as there was no text data to model and visualize. Subsequent meetings displayed all prior meeting content.

Data Collection & Analysis

Apart from demographic and technical background data collected from each participant at the start of the study, participants were asked to evaluate their meeting experience in terms of satisfaction, adherence to a topic, fixation etc. on a 5-point Likert scale. Participants reported whether or not they covered all items on the given agenda and if they discussed any topics *not* on the agenda. They also reported on the usefulness and relevance of the visualizations and on how distracted they felt during the study. For iteration 2, we also interviewed the team after every session.

We were also interested in the effect of the visualization on participant behavior. The videos of each session (total 20 videos including team F for iteration 2, ~75 minutes for each team) were coded by two of the authors of this paper, with an overlap of four videos totaling 75 minutes (~20%) between them. Each coder marked the following behavior: for all participants, they identified instances of speaking and of looking at their notes/agenda. For teams using TalkTraces, they additionally coded instances of looking at the visualization. We posited that the time spent by participants looking at the visualization could indicate whether it could serve as an alternative or an extension of notes and agenda items to which meeting participants typically refer. Speaking time was used as a baseline in case we needed to look closer at instances of non-equitable participation. Both coders independently coded the same video to compare codes. Using a two-second window to compare codes, they reached an agreement Cohen's kappa (κ) of 0.66 ($p < 0.001$).

5 RESULTS

We first report the results of the study with iteration 1 of TalkTraces (teams V1, V2, B1, B2), followed by the results of the follow-up study that used iteration 2 (team F). The results are based on participant feedback of their meeting experience and—for the teams that used TalkTraces—feedback on the relevance and usefulness of the interface, along with the level of distraction they felt during the meeting. Fig. 8 shows the overall participant feedback on their meeting experience across sessions for baseline and vis teams. Coded data of time spent by participants speaking, looking at notes/agenda, and looking at visualizations is shown in Table 1.

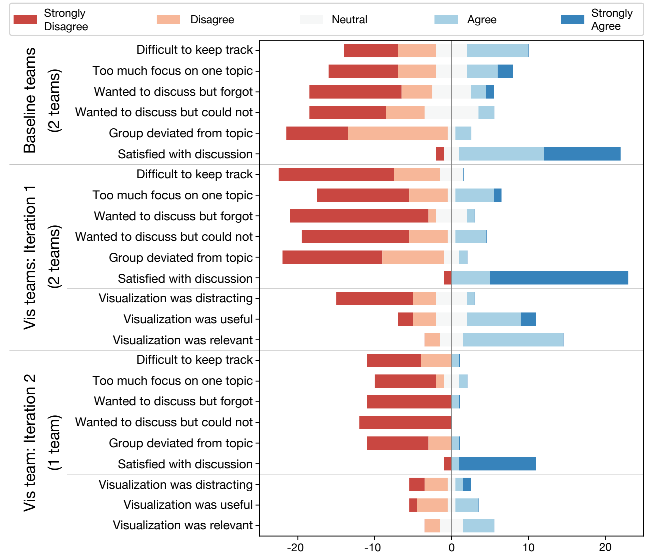

Figure 8: Participant responses on a Likert scale to questions on the meeting for each meeting condition, represented as diverging stacked bars [23]. Iteration 2 has fewer participants than the other conditions, and thus shorter bars.

Iteration 1 Study: Topic-Focused Display

Participants who used iteration 1 of TalkTraces reported that they mostly did not find the visualization distracting. They found the visualizations largely relevant to their discussion but were divided on whether they were useful. While the number of participants is too low for statistical analysis, this data can be used as a background for participant comments which we categorize under relevance, utility, and distraction.

Relevance: Participant perception of the relevance of the information shown on TalkTraces depended on factors such as the novelty of the visualization, the relation of the current discussion to prior sessions, and occasionally the accuracy of the speech-to-text transcription. One participant (team V1) found the visualization relevant in sessions 1 and 2, but not in session 3. Her comments demonstrate a decline in enthusiasm: *"it was interesting to see which category the speaker's sentence was being added to and see which category got bigger"* (session 1), *"there was one topic we did not talk about at all"* (session 2), and *"I didn't see how the main part of the vis was benefiting me. But I liked (having) the transcript to read back"* (session 3). By session 3, she chose to simply refer to the transcript to keep track of the discussion, a sentiment a teammate of hers mirrored. In fact, her team did not take any notes after session 0, completely relying on the printed agenda and TalkTraces for recall and discussion. Another participant's comments highlighted one of the weaknesses of using topic modeling in this setting: *"today I think the ...projected budget was so dissociated from the topics in the visualization that I did not use the visual* (sic) *actively."*

Table 1: Mean time per team per session for each action.

Team	Session	Time Spent on Actions (min:sec)					
		Speak		Look (notes)		Look (vis)	
	# (mins)	M	S.D.	M	S.D.	M	S.D.
Iteration 1							
V1	0 (30:00)	7:29	2:59	6:20	3:49	–	–
	1 (15:00)	3:56	1:40	1:35	1:18	1:42	1:23
	2 (15:00)	4:21	1:53	1:59	0:32	0:47	0:29
	3 (15:00)	4:06	1:29	1:41	0:10	1:33	1:17
V2	0 (30:00)	9:00	5:58	7:32	4:43	–	–
	1 (15:00)	4:03	2:44	2:49	2:22	1:15	1:12
	2 (15:00)	2:59	4:03	2:39	2:06	2:13	1:35
	3 (15:00)	4:06	2:54	1:59	1:46	1:16	1:42
B1	0 (30:00)	9:27	3:05	23:02	4:54	–	–
	1 (15:00)	4:43	3:33	10:00	3:01	–	–
	2 (15:00)	4:46	1:45	9:29	2:17	–	–
	3 (15:00)	4:44	1:28	11:06	2:03	–	–
B2	0 (30:00)	8:04	2:53	6:32	4:22	–	–
	1 (15:00)	4:13	2:20	2:48	0:27	–	–
	2 (15:00)	2:40	1:00	1:33	1:30	–	–
	3 (15:00)	3:52	1:23	2:01	0:43	–	–
Iteration 2							
F	0 (30:00)	6:26	2:12	9:32	3:10	–	–
	1 (15:00)	2:56	1:50	3:39	1:10	1:00	0:25
	2 (15:00)	1:56	1:29	3:04	1:27	1:47	1:07
	3 (15:00)	2:20	1:31	4:41	2:51	0:59	0:48

Usefulness: Most of the participants found the visualization useful for recalling meetings, but less so for real-time awareness of the ongoing discussion. One participant emphasized consistently how he used the visualization for recall at the start of the meeting, *"but didn't look at it moving forward"* partly because he found it difficult to interpret the topic blobs. Two participants found that the word cloud showing words unique to the current discussion helped them think of more ideas to discuss. One of them said, *"I did not pay enough attention to the visualization during talking turns but used it for support in the next information to present."*

Distraction: Most participants did not often refer to the visualization when talking to each other, one of whom noted how it was *"easy to ignore when you're focused on talking."* In Table 1, the total time teams V1 and V2 spend looking at their notes/agenda was consistently higher (by a small margin) than the total time they spent looking at the visualization. However, it is difficult to say from observation alone if looking at the visualizations indicated distraction or intent. Prior work on measuring distraction in group settings has used a combination of wearable sensors and self-reporting [32]. While we also use self-reporting, a more focused study on distraction may be needed for a more accurate assessment. The mean time that teams V1 and V2 spent looking at notes/agenda is close to team B2, which suggests that the visualizations were supplementary to the printed agenda and notes, not a replacement. Team B1 spent the most time looking at their notes/agenda, but we noticed that they seldom made eye contact with each other throughout the meeting, preferring to look at their notes/agenda while speaking or listening, which could indicate introversion.

Inaccuracies in transcription sometimes drew untoward attention: *"sometimes when the system didn't pick up what we were saying accurately, I got a little distracted."* Participants also found the topic visualization somewhat abstract and difficult to interpret. While they seemed intrigued if they found unaddressed topics during a session, they did not find it as useful as the (written) agenda list and the transcript.

Overall, the study revealed that the topic-focused visualization central to iteration 1 was helpful in reviewing previous meetings, but was less helpful in keeping track of the session in real-time. The main issue seemed to be that, when the discussion progressed in a direction not covered earlier, newer keywords could not be allocated clearly to any one of the topics. They were assigned to the "unknown topic" blob, which was thought to have limited utility. The keyword and transcript views were moderately useful as they helped participants think of, or remember items to discuss.

Iteration 2 Study: Agenda-Focused Display

Iteration 2 of TalkTraces sought to address the weaknesses in the previous iteration through a redesign of the real-time visualization, as discussed in the design section. Participant feedback showed that the agenda-focused visualization was more useful than the topic-focused version, and the potential benefits would be revealed in greater relief if employed in longer and more complex discussions. Feedback on the transcript and word clouds were similar to those in iteration 1.

The visualizations here were more complex, with pie charts showing the topic distribution for each sentence. This initially made it difficult for participants to keep track of the visualization. In the interview after the first session, the team leader reported difficulty with keeping track of topics: *"trying to dissect what the charts do versus the coloring code because the topics don't make concrete sense to me."* Participants mostly felt confused by the pies (topic distributions) in the beeswarm plot but were able to understand the agenda-focused visualization. The same participant said, when reviewing the visualization: *"We can say that we did talk very little about that (pointing to an agenda item) because we had very little to say about that, but we can see that from the plot. But when you add those three (topic) colors, we immediately need to associate the colors with that (pointing to the list of topic keywords), and we're pulled away from it."*

In subsequent sessions, they became used to the visualizations, and by session 3 the same participant found himself referring to the list of agenda items on the display and trying to address them in real time. *"I did see the visualization once to see what (agenda items) we didn't cover, which was vendors, originally (pointing using the mouse), which was about here…(points to specific nodes on the visualization)."* This observation is a significant validation of the position-based encoding because the participant seemed to subconsciously not only relate a discussion item to an agenda item, but also the *time* in the meeting at which it was addressed.

Finally, participants could not easily interpret some of the topics and agenda items and felt that they should have more control over it. This caused some disagreements within the participants about what agenda items a particular discussion should be assigned under. This was partly due to the nature of the study, where agenda items were given to participants for consistency, and partly because the topic model relied on prior meetings instead of a representative corpus of text.

6 DISCUSSION & IMPLICATIONS

TalkTraces is designed to address four main requirements: (R1) to aid historical recall, (R2) to provide a conceptual overview at a glance, (R3) to provide situational awareness of the current state of the meeting, and (R4) to provide an impetus for conversation. Study results with the two iterations of TalkTraces indicate that all of these requirements are addressed to some extent, but there remain a number of ways in which real-time awareness needs to be balanced with the ability to review discussions.

Choice of Conscious vs. Preattentive Processing

Participants found it difficult to keep track of topic-focused visualizations in both iterations as they had to continuously re-interpret the topics based on the updating list of words, requiring conscious processing. In contrast, the agenda aspects visualized in iteration 2 facilitated much better discussion, partly because they enabled preattentive processing, as position encoded both the specific agenda items as well as their times. We thus recommend using design principles that enable preattentive processing (such as position and size), as they minimize user effort required to interpret the visualization and are more effective for real-time meeting awareness. One participant from the study of iteration 2 suggested removing topic-based representation *during* meetings and allowing them to be overlaid as an option during meeting review. This allows thematic exploration when needed.

Providing Thematic Representation

In both iterations, participants found it difficult to interpret the topics, partly because the dataset was sparse and partly because of the topic modeling (LDA) and topic representation (keyword list). Our choice was motivated by prior research that showed how novice users preferred using keywords to interpret topics [52], and because LDAVis [49] allowed a simple criterion and control to order the keywords. However, newer, interactive topic modeling approaches [51] now allow users fine-grained control over topic refinement. These might potentially address the difficulty in interpretation of meeting discussions datasets at the cost of a steeper learning curve; they are worth exploring. It would also help to supplement meeting transcripts with additional data, such as a corpus of text from team communication (emails, messages, documents, etc.) or an ontology that represents domain-specific knowledge. Such project-relevant communication and references can thus be mined to model the relevant knowledge for providing a suitable thematic overview.

Providing Conversation Impetus

While participants from iteration 1 did not find the visualization distracting, they also failed to notice when they fixated on one topic or switched between topics. This information, encoded by changes in the topic blob's chroma and luminance values, was too subtle to perceive. Those who reported that TalkTraces helped them think of discussion points either used the word cloud during the meeting (mainly in iteration 1) or the agenda-transcript beeswarm plot (in iteration 2). Some participants remarked on the need to look up outside information such as locations or numbers (e.g. the capacity of the campus stadium), but this was prohibited in the study setup as participants were not allowed access to other computers. Speech-inspired search and display representations are already being explored [48] and could prove useful in providing new directions for the discussion.

7 LIMITATIONS

TalkTraces has been designed as a real-time visualization for helping meeting participants' situational awareness. In addition to the strengths and limitations of the interface itself that have been uncovered in the study, there are aspects of the system dependencies and study setup that need to be considered when drawing conclusions from the results.

One limitation that has been discussed in the previous section is the choice of topic modeling approach and its representation. New, interactive topic modeling approaches and representations allow users to interactively analyze topic emergence, splitting and merging [15]. They also help track hierarchical and temporal changes and alignments of topics over text streams [35]. Other techniques help users refine topic modeling results [51]. More recent systems are available to improve the interpretability of our topic modeling approach, perhaps at the cost of a steeper learning curve.

Another limitation is the agenda and duration of the meetings in our study. With the exception of agile scrum meetings, discussions typically last longer than 15 minutes, and participants have some agency over determining agenda items. We chose a generic agenda of event planning over potential technical discussions to minimize the occurrence of jargon, which may not be accurately transcribed by the used speech-to-text API. This limitation is fast disappearing with improvements in speech recognition. Keeping these considerations in mind, TalkTraces should be interpreted as a supplement to existing SMR technologies that would benefit from its integration.

8 CONCLUSION

In this paper, we present TalkTraces, a multi-view, real-time representation of conversations that presents a thematic view of discussion content. We present two iterations of TalkTraces, both of which use transcribed audio to create a topic model of the discussion using LDA. Iteration 1 features a topic-focused visualization that updates in real time as the discussion is transcribed and processed using the topic model. A user evaluation revealed the difficulty of participants in quick, at-a-glance interpretation/recall of topics. Participants also found color-based animations too subtle to follow updates to the visualization. In response to the feedback, Iteration 2 additionally uses word embeddings to compute agenda-to-discussion similarity in real time and displays the result as a beeswarm plot to help participants keep track of agenda items covered. Participants found the agenda-based visualization useful and intuitive, though the topic representations were still difficult to interpret. We discuss the implications of these observations and provide guidelines for real-time visualizations of discussion content.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their feedback and suggestions that helped improve the quality of this paper. This research is sponsored in part by U.S. National Science Foundation through grants IIS-1320229 and IIS-1528203.

REFERENCES

[1] Daniel Angus, Andrew Smith, and Janet Wiles. 2012. Conceptual recurrence plots: Revealing patterns in human discourse. *IEEE Transactions on Visualization and Computer Graphics* 18, 6 (2012), 988–997. https://doi.org/10.1109/TVCG.2011.100

[2] Tony Bergstrom and Karrie Karahalios. 2009. Conversation clusters: grouping conversation topics through human-computer dialog. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*. 2349–2352. https://doi.org/10.1145/1518701.1519060

[3] David M. Blei. 2012. Probabilistic Topic Models. *Commun. ACM* 55, 4 (2012), 77–84. https://doi.org/10.1145/2133806.2133826

[4] David M Blei and John D Lafferty. 2006. Dynamic topic models. In *Proceedings of the International Conference on Machine Learning*. 113–120. https://doi.org/10.1145/1143844.1143859

[5] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, 4/5 (2003), 993–1022.

[6] Sara A Bly, Steve R Harrison, and Susan Irwin. 1993. Media spaces: bringing people together in a video, audio, and computing environment. *Commun. ACM* 36, 1 (1993), 28–46. https://doi.org/10.1145/151233.151235

[7] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D^3: Data-Driven Documents. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2301–2309. https://doi.org/10.1109/TVCG.2011.185

[8] Nadieh Bremer. 2015. Create a “gooey effect” during transitions in d3.js. https://www.visualcinnamon.com/2015/05/gooey-effect.

[9] Vannevar Bush. 1945. As we may think. *The atlantic monthly* 176, 1 (1945), 101–108. https://doi.org/10.1145/227181.227186

[10] Nan Cao, Jimeng Sun, Yu-Ru Lin, David Gotz, Shixia Liu, and Huamin Qu. 2010. Facetatlas: Multifaceted visualization for rich text corpora. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (2010), 1172–1181. https://doi.org/10.1109/TVCG.2010.154

[11] Sandro Castronovo, Jochen Frey, and Peter Poller. 2008. A generic layout-tool for summaries of meetings in a constraint-based approach. In *International Workshop on Machine Learning for Multimodal Interaction*. 248–259. https://doi.org/10.1007/978-3-540-85853-9_23

[12] Patrick Chiu, Ashutosh Kapuskar, Sarah Reitmeier, and Lynn Wilcox. 2000. Room with a rear view. Meeting capture in a multimedia conference room. *IEEE MultiMedia* 7, 4 (2000), 48–54. https://doi.org/10.1109/93.895154

[13] Cisco Systems Inc. 2018. Cisco Collaboration Meeting Rooms. https://www.cisco.com/c/en/us/solutions/collaboration/pervasive-conferencing/index.html. [Online; accessed 03-April-2018].

[14] Stephen Clark. 2015. Vector space models of lexical meaning. *The Handbook of Contemporary semantic theory* (2015), 493–522. https://doi.org/10.1002/9781118882139.ch16

[15] Weiwei Cui, Shixia Liu, Li Tan, Conglei Shi, Yangqiu Song, Zekai Gao, Huamin Qu, and Xin Tong. 2011. Textflow: Towards better understanding of evolving topics in text. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2412–2421. https://doi.org/10.1109/TVCG.2011.239

[16] Wenwen Dou, Xiaoyu Wang, Drew Skau, William Ribarsky, and Michelle X Zhou. 2012. Leadline: Interactive visual analysis of text data through event identification and exploration. In *IEEE Conference on Visual Analytics Science and Technology*. 93–102. https://doi.org/10.1109/VAST.2012.6400485

[17] Patrick Ehlen, Matthew Purver, John Niekrasz, Kari Lee, and Stanley Peters. 2008. Meeting adjourned: off-line learning interfaces for automatic meeting understanding. In *Proceedings of the ACM international conference on Intelligent user interfaces*. 276–284. https://doi.org/10.1145/1378773.1378810

[18] Mennatallah El-Assady, Rita Sevastjanova, Bela Gipp, Daniel Keim, and Christopher Collins. 2017. NEREx: Named-Entity Relationship Exploration in Multi-Party Conversations. In *Computer Graphics Forum*, Vol. 36. 213–225. https://doi.org/10.1111/cgf.13181

[19] Sayed M Elsayed-Elkhouly, Harold Lazarus, and Volville Forsythe. 1997. Why is a third of your time wasted in meetings? *Journal of Management Development* 16, 9 (1997), 672–676. https://doi.org/10.1108/02621719710190185

[20] Carlos Filipe Freitas, Joao Barroso, and Carlos Ramos. 2015. A Survey on Smart Meeting Rooms and Open Issues. *International Journal of Smart Home* 9, 9 (2015), 13–20.

[21] Brynjar Gretarsson, John O'donovan, Svetlin Bostandjiev, Tobias Höllerer, Arthur Asuncion, David Newman, and Padhraic Smyth. 2012. Topicnets: Visual analysis of large text corpora with topic modeling. *ACM Transactions on Intelligent Systems and Technology* 3, 2 (2012), 23. https://doi.org/10.1145/2089094.2089099

[22] Amit Gruber, Yair Weiss, and Michal Rosen-Zvi. 2007. Hidden topic markov models. In *Proceedings of the International Conference on Artificial intelligence and statistics*. 163–170.

[23] Richard M Heiberger, Naomi B Robbins, et al. 2014. Design of diverging stacked bar charts for Likert scales and other applications. *Journal of Statistical Software* 57, 5 (2014), 1–32. https://doi.org/10.18637/jss.v057.i05

[24] Steven Houben, Connie Golsteijn, Sarah Gallacher, Rose Johnson, Saskia Bakker, Nicolai Marquardt, Licia Capra, and Yvonne Rogers. 2016. Physikit: Data engagement through physical ambient visualizations in the home. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*. 1608–1619. https://doi.org/10.1145/2858036.2858059

[25] Seth Hunter, Pattie Maes, Stacey Scott, and Henry Kaufman. 2011. MemTable: an integrated system for capture and recall of shared histories in group workspaces. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*. 3305–3314. https://doi.org/10.1145/1978942.1979432

[26] Seth Hunter, Pattie Maes, Stacey Scott, and Henry Kaufman. 2011. MemTable: an integrated system for capture and recall of shared histories in group workspaces. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*. 3305–3314. https://doi.org/10.1145/1978942.1979432

[27] Intel Inc. 2018. Intel Unite Solution. https://www.intel.com/content/www/us/en/architecture-and-technology/unite/intel-unite-overview.html. [Online; accessed 03-April-2018].

[28] Petra Isenberg, Niklas Elmqvist, Jean Scholtz, Daniel Cernea, Kwan-Liu Ma, and Hans Hagen. 2011. Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization* 10, 4 (2011), 310–326. https://doi.org/10.1177/1473871611412817

[29] Nassim Jafarinaimi, Jodi Forlizzi, Amy Hurst, and John Zimmerman. 2005. Breakaway: an ambient display designed to change human behavior. In *ACM CHI extended abstracts on Human factors in computing systems*. 1945–1948. https://doi.org/10.1145/1056808.1057063

[30] Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. 2017. Bag of Tricks for Efficient Text Classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Association for Computational Linguistics, 427–431.

[31] KyungTae Kim, Sungahn Ko, Niklas Elmqvist, and David S Ebert. 2011. WordBridge: Using composite tag clouds in node-link diagrams for visualizing content and relations in text corpora. In *IEEE Hawaii International Conference on System Sciences*. 1–8. https://doi.org/10.1109/HICSS.2011.499

[32] Taemie Kim, Agnes Chang, Lindsey Holland, and Alex Sandy Pentland. 2008. Meeting mediator: enhancing group collaborationusing sociometric feedback. In *Proceedings of the ACM conference on Computer supported cooperative work*. 457–466. https://doi.org/10.1145/1460563.1460636

[33] Kostiantyn Kucher and Andreas Kerren. 2015. Text visualization techniques: Taxonomy, visual survey, and community insights. In *IEEE Pacific Visualization Symposium*. 117–121. https://doi.org/10.1109/PACIFICVIS.2015.7156366

[34] Wei Li and Andrew McCallum. 2006. Pachinko allocation: DAG-structured mixture models of topic correlations. In *Proceedings of the ACM international conference on Machine learning*. 577–584. https://doi.org/10.1145/1143844.1143917

[35] Shixia Liu, Jialun Yin, Xiting Wang, Weiwei Cui, Kelei Cao, and Jian Pei. 2016. Online visual analytics of text streams. *IEEE Transactions on Visualization and Computer Graphics* 22, 11 (2016), 2451–2466. https://doi.org/10.1109/TVCG.2015.2509990

[36] Edward Loper and Steven Bird. 2002. NLTK: the natural language toolkit. In *Proceedings of the ACL Workshop on Effective tools and methodologies for teaching natural language processing and computational linguistics - Volume 1*. 63–70. https://doi.org/10.3115/1118108.1118117

[37] Jennifer Mankoff, Anind K Dey, Gary Hsieh, Julie Kientz, Scott Lederer, and Morgan Ames. 2003. Heuristic evaluation of ambient displays. In *Proceedings of the ACM CHI conference on Human factors in computing systems*. 169–176. https://doi.org/10.1145/642611.642642

[38] Microsoft Inc. 2018. IoT Smart Buldings | Microsoft. https://www.microsoft.com/en-us/internet-of-things/smart-building. [Online; accessed 03-April-2018].

[39] Jay F. Nunamaker, Robert O. Briggs, and Daniel D. Mittleman. 1996. Lessons from a decade of group support systems research. In *Proceedings of the Hawaii International Conference on System Sciences Volume 3: Collaboration Systems and Technology*. 418–427.

[40] Jay F. Nunamaker, Alan R Dennis, Joseph S Valacich, Douglas Vogel, and Joey F George. 1991. Electronic meeting systems. *Commun. ACM* 34, 7 (1991), 40–61. https://doi.org/10.1145/105783.105793

[41] Chris Parnin and Carsten Gorg. 2007. Design guidelines for ambient software visualization in the workplace. In *IEEE International Workshop on Visualizing Software for Understanding and Analysis*. 18–25. https://doi.org/10.1109/VISSOF.2007.4290695

[42] Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the conference on empirical methods in natural language processing*. 1532–1543.

[43] Zachary Pousman and John Stasko. 2006. A taxonomy of ambient information systems: four patterns of design. In *Proceedings of the ACM conference on Advanced visual interfaces*. 67–74. https://doi.org/10.1145/1133265.1133277

[44] Carlos Ramos, Juan Carlos Augusto, and Daniel Shapiro. 2008. Ambient intelligence—the next step for artificial intelligence. *IEEE Intelligent Systems* 23, 2 (2008), 15–18. https://doi.org/10.1109/MIS.2008.19

[45] Radim Řehůřek and Petr Sojka. 2010. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC Workshop on New Challenges for NLP Frameworks*. 45–50.

[46] Alan R Robertson. 1977. The CIE 1976 Color-Difference Formulae. *Color Research & Application* 2, 1 (1977), 7–11.

[47] Yang Shi, Chris Bryan, Sridatt Bhamidipati, Ying Zhao, Yaoxue Zhang, and Kwan-Liu Ma. 2018. MeetingVis: Visual Narratives to Assist in Recalling Meeting Context and Content. *IEEE Transactions on Visualization and Computer Graphics* (2018). https://doi.org/10.1109/TVCG.2018.2816203

[48] Yang Shi, Yang Wang, Ye Qi, John Chen, Xiaoyao Xu, and Kwan-Liu Ma. 2017. IdeaWall: improving creative collaboration through combinatorial visual stimuli. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing*. 594–603. https://doi.org/10.1145/2998181.2998208

[49] Carson Sievert and Kenneth Shirley. 2014. LDAvis: A method for visualizing and interpreting topics. In *Proceedings of the workshop on interactive language learning, visualization, and interfaces*. 63–70.

[50] Tobias Skog. 2004. Activity wallpaper: ambient visualization of activity information. In *Proceedings of the ACM Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*. 325–328. https://doi.org/10.1145/1013115.1013171

[51] Alison Smith, Varun Kumar, Jordan Boyd-Graber, Kevin Seppi, and Leah Findlater. 2018. Closing the Loop: User-Centered Design and Evaluation of a Human-in-the-Loop Topic Modeling System. In *Proceedgs of the ACM Conference on Intelligent User Interfaces*. 293–304. https://doi.org/10.1145/3172944.3172965

[52] Alison Smith, Tak Yeon Lee, Forough Poursabzi-Sangdeh, Jordan Boyd-Graber, Niklas Elmqvist, and Leah Findlater. 2017. Evaluating visual

representations for topic understanding and their effects on manually generated labels. *Transactions of the Association of Computational Linguistics* 5, 1 (2017), 1–15.

[53] Robyn Speer, Joshua Chin, and Catherine Havasi. 2017. ConceptNet 5.5: An Open Multilingual Graph of General Knowledge. In *AAAI Conference on Artificial Intelligence*. 4444–4451.

[54] Yun Sun, Yubo Tao, Geng Yang, and Hai Lin. 2013. Visitpedia: Wiki article visit log visualization for event exploration. In *International Conference on Computer-Aided Design and Computer Graphics*. 282–289. https://doi.org/10.1109/CADGraphics.2013.44

[55] Yuzuru Tanahashi and Kwan-Liu Ma. 2015. Stock lamp: An engagement-versatile visualization design. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems*. 595–604. https://doi.org/10.1145/2702123.2702323

[56] Gokhan Tur, Andreas Stolcke, Lynn Voss, Stanley Peters, Dilek Hakkani-Tur, John Dowding, Benoit Favre, Raquel Fernández, Matthew Frampton, Mike Frandsen, Clint Frederickson, Martin Graciarena, Donald Kintzing, Kyle Leveque, Shane Mason, John Niekrasz, Matthew Purver, Korbinian Riedhammer, Elizabeth Shriberg, Jing Tien, Dimitra Vergyri, and Fan Yang. 2010. The CALO meeting assistant system. *IEEE Transactions on Audio, Speech, and Language Processing* 18, 6 (2010), 1601–1611. https://doi.org/10.1109/TASL.2009.2038810

[57] Joseph Turian, Lev Ratinov, and Yoshua Bengio. 2010. Word represen-tations: a simple and general method for semi-supervised learning. In *Proceedings of the 48th annual meeting of the association for computa-tional linguistics*. 384–394.

[58] Frank Van Ham, Martin Wattenberg, and Fernanda B Viégas. 2009. Mapping text with phrase nets. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009). https://doi.org/10.1109/TVCG.2009.165

[59] Sergio A Velastin, Boghos A Boghossian, Benny PL Lo, Jie Sun, and Maria Alicia Vicencio-Silva. 2005. PRISMATICA: toward ambient intelligence in public transport environments. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 35, 1 (2005), 164–182. https://doi.org/10.1109/TSMCA.2004.838461

[60] Daniel Vogel and Ravin Balakrishnan. 2004. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the ACM symposium on User Interface Software and Technology*. 137–146. https://doi.org/10.1145/1029632.1029656

[61] Xiting Wang, Shixia Liu, Yang Chen, Tai-Quan Peng, Jing Su, Jing Yang, and Baining Guo. 2016. How ideas flow across multiple social groups. *IEEE Transactions on Visualization and Computer Graphics*, 51–60. https://doi.org/10.1109/VAST.2016.7883511

[62] Xiting Wang, Shixia Liu, Yangqiu Song, and Baining Guo. 2013. Mining evolutionary multi-branch trees from text streams. In *Proceedings of the ACM SIGKDD international conference on Knowledge Discovery and Data mining*. 722–730. https://doi.org/10.1145/2487575.2487603

[63] Xuerui Wang and Andrew McCallum. 2006. Topics over time: a non-Markov continuous-time model of topical trends. In *Proceedings of the ACM SIGKDD international conference on Knowledge Discovery and Data mining*. 424–433. https://doi.org/10.1145/1150402.1150450

[64] Furu Wei, Shixia Liu, Yangqiu Song, Shimei Pan, Michelle X Zhou, Weihong Qian, Lei Shi, Li Tan, and Qiang Zhang. 2010. Tiara: a visual exploratory text analytic system. In *Proceedings of the ACM SIGKDD international conference on Knowledge discovery and data mining*. 153–162. https://doi.org/10.1145/1835804.1835827

[65] Steve Whittaker, Simon Tucker, and Denis Lalanne. 2012. Meeting browsers and meeting assistants. In *Multimodal Signal Processing: Human Interactions in Meetings*, Steve Renals, Hervé Bourlard, Jean Carletta, and Andrei Popescu-Belis (Eds.). Cambridge University Press, Chapter 12, 204–217.

[66] Felix Wortmann and Kristina Flüchter. 2015. Internet of things. *Business & Information Systems Engineering* 57, 3 (2015), 221–224. https://doi.org/10.1007/s12599-015-0383-3

[67] Zhiwen Yu and Yuichi Nakamura. 2010. Smart meeting systems: A survey of state-of-the-art and open issues. *ACM Computing Surveys (CSUR)* 42, 2 (2010), 8. https://doi.org/10.1145/1667062.1667065