Erschienen in: Computer Graphics Forum ; 35 (2016), 3. - S. 431-440
https://dx.doi.org/10.1111/cgf.12919

ConToVi: Multi-Party Conversation Exploration using Topic-Space Views

Mennatallah El-Assady[1,2], Valentin Gold[1], Carmela Acevedo[1], Christopher Collins[2], and Daniel Keim[1]

[1]University of Konstanz, Germany
[2]University of Ontario Institute of Technology, Canada

Abstract

We introduce a novel visual analytics approach to analyze speaker behavior patterns in multi-party conversations. We propose Topic-Space Views to track the movement of speakers across the thematic landscape of a conversation. Our tool is designed to assist political science scholars in exploring the dynamics of a conversation over time to generate and prove hypotheses about speaker interactions and behavior patterns. Moreover, we introduce a glyph-based representation for each speaker turn based on linguistic and statistical cues to abstract relevant text features. We present animated views for exploring the general behavior and interactions of speakers over time and interactive steady visualizations for the detailed analysis of a selection of speakers. Using a visual sedimentation metaphor we enable the analysts to track subtle changes in the flow of a conversation over time while keeping an overview of all past speaker turns. We evaluate our approach on real-world datasets and the results have been insightful to our domain experts.

1. Introduction

Traditional text analysis techniques have been mostly developed to handle highly-edited written text data since these texts usually have a higher writing-quality with less noise. In contrast, transcribed spoken text is usually based on social interactions between different speakers. This type of text data is typically less informative and more argumentative in its nature [Ast01]. In a broader definition, written online discussions on social media can also be categorized as transcribed conversations. This type of text data is usually created fast and in large quantities. Therefore, depending on qualities of the writing or the speech and their transcription, these texts may contain multiple non-standard lexical items and syntactic patterns. Due to these characteristics, analyzing such corpora using existing automatic analysis methods usually does not achieve the same level of accuracy as with highly-edited written text. Furthermore, as a consequence of the social interactions in conversations, these corpora typically evolve in a serial fashion over time, making the content and form of the texts a direct result of the actions and reactions of speakers. To achieve a better understanding of conversational texts, it is, therefore, necessary to analyze their evolution over time and the underlying dynamics within the conversation.

Across all subfields of political science, scholars are interested in analyzing deliberative democracy. Deliberative democracy is a theoretical construct emphasizing the use of proper modes of reason-giving to achieve mutual consent in a debate [Cha03]. So far, most studies either focus on the antecedents or the consequences of deliberation, but neglect the heart of the discussion which is the dynamics within the debate. An analysis of the dynamics of deliberation requires an understanding of the role of participants in a conversation and their speech and argumentation patterns across raised topics. We conducted a preliminary observational study of three political science researchers, specializing in deliberative democracy, and four computational linguistic scholars, which resulted in identifying two main tasks for multi-party conversation analysis: analyzing the roles of participants in a conversation and their attitude towards a given topic. Our observational study also revealed that the current framework of political science analysis is based on a manual reading of the debate to identify interesting segments and reconstruct the content-based interaction between participants. Hence, for deliberation analysis, it is not only important to know that a given speaker addresses a specific topic, but also their stance towards the topic and how it evolves over time.

Additionally, recent research on deliberation has integrated empirical results to the analysis [SBSS03], relying on a combination of manual coding and subsequent statistical analysis. Both of these manual processes of analysis are very time-consuming and can only be performed on a limited set of corpora [BBGSG10]. The coding and analysis results are often subjective and not reproducible, making them often subject to critical judgments of other researchers.

In this paper, we introduce ConToVi (Conversation-Topic-Visualization), a novel visual analytics system to explore multi-party conversations based on Topic-Space views. ConToVi builds a framework for automatically extracting relevant linguistic and statistical features from a conversation and visualizes them in various

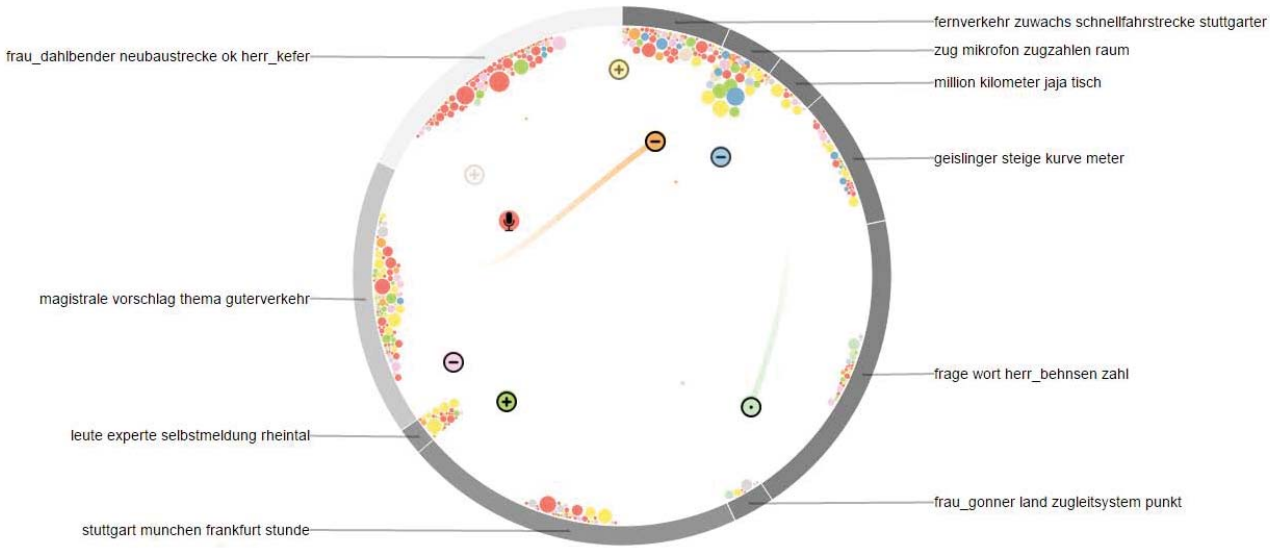


Figure 1: Frame from the animation of the Utterance-Sedimentation View of the complete Stuttgart 21 arbitration.

contexts and at different levels of detail. Our observational study results indicated that speaker behavior patterns and utterance feature analysis have to be regarded in context of the thematic structure of the underlying text. This is because the same speaker might behave differently when speaking about different topics. Additionally, the content of the conversation is important to differentiate and compare the interactions of participants throughout a conversation. We observed that the political science scholars map each utterance in a conversation to a manually created topic structure in order to perform further analysis on a more abstract grouping of the speaker-turns. To group utterances into more abstract thematic clusters, we use a hierarchical topic modeling algorithm [EA15] that was developed to handle linearly structured text corpora and generates, therefore, comprehensive topic structures for conversational text corpora. Our topic modeling algorithm addresses the challenges of automatically processing transcribed spoken text data and ensures a solid foundation for the further analysis, as the computed topic structure is one of the main pillars of the visual analytics approach.

We visualize a conversation's thematic floor by spanning the topics as dimensions of a circular plot, called *Topic-Space* (see Figure 1). In this space we place speakers and their utterances to track their movement and interactions over time. Inside the Topic-Space, utterances are placed according to their relation to the outer topic dimensions. This relation is derived using the relative membership degree of each utterance to each topic. These relative membership degrees build the weights for the force-directed layout that determines the positions of all objects inside the Topic-Space. Similar to the Dust and Magnet metaphor [YMSJ05], topics can be regarded as magnets, whereas speakers and utterances could be seen as dust particles in the space. Through this placement, the position of every speaker in the topic space at a given time point is determined from her corresponding speaker-turn at that time. Furthermore, the movement of a speaker inside the space is computed by interpolating the positions of all her utterances throughout the conversation. Using this design principle, we generate four different views to address our users' analysis requirements–the analysis of deliberative dynamics. The metaphor of a conversation floor depicted by the Topic-Space creates a comprehensive mental model for the use of our visualization that has been favored by the domain experts.

The requirements for creating this visualization address the above-mentioned gaps in the research of deliberative democracy, namely the analysis of the dynamics of a conversation and the exploration of speaker interactions across the thematic landscape of a debate [ASW12]. As a goal of deliberative analysis is to identify the use of proper modes of reason-giving, we extract the most frequent linguistic cues for speech and argumentation to facilitate the analysis of speaker behavior patterns. The requirements can be categorized as follows: **Speaker Movement Analysis**: **[R1]** Retracing a sequence of speakers and topics over time. **[R2]** Identifying the influence of a speaker on the conversation. **[R3]** Analyzing how speakers push a specific topic. **Behavior Pattern Analysis**: **[R4]** Exploring speech and argumentation patterns for each speaker and topic. **[R5]** Analyzing the speaker stances at a given point in time towards a given topic. **[R6]** Comparing the pattern similarities of different speakers. To address these requirements, we created the following animated and interactive visualizations: **[V1]** Discussion Overview with all speakers. **[V2]** Speaker-Trail-View for a selection of speakers. **[V3]** Utterance-Sedimentation View for a selection of speakers. **[V4]** Speaker-Path View for one speaker. All views allow the abstraction of utterance features into glyphs tailored to highlight the speech and argumentation patterns of a speaker-turn. The Topic-Space features an on-demand glyph overlay and displays more details about an utterance, as well as the underlying text, upon interaction with the glyphs. We present two types of glyphs: a *topic glyph* and an *argumentation glyph*.

In general, ConToVi is designed to be independent of the specific algorithms used for feature computation and topic modeling, thus allowing the analysis of numerous types of transcribed spoken texts with conversational character. The main contributions of our approach are: **[C1]** A characterization of the challenges and requirements for automatic analysis of conversational text data. **[C2]**

Introducing *Topic-Space Views* as an alternative to existing ThemeRiver-based visualizations for topic-based text exploration. **[C3]** Proposing design solutions for overcoming the known drawbacks of radial force-directed layouts. **[C4]** Integrating motion chart trails into the visualization to overcome motion-blindness and allow the animation to be more traceable. **[C5]** Using a visual-sedimentation-based metaphor for setting the animation into the context of the prior speaker-turns. **[C6]** A characterization and abstraction of speech and argumentation features into tailored glyphs.

2. Related Work

Topic Visualizations — In recent years, various interactive visualization approaches have been developed for the exploration of topical structures of document collections, mostly based on Latent Dirichlet Allocation [BNJ03] (LDA). Some approaches, such as FacetAtlas [CSL*10] and TopicPanorama [LWC*14], visualize topics by meaningfully grouping their keywords. Both of these visualizations rely on a graph structure of the keywords and on graph layout algorithms to position them.

The majority of the topic analysis approaches, on the other hand, adapt ThemeRiver-based visualizations to show the temporal trends of topics. The ThemeRiver metaphor was first used for text data by Havre et al. [HHWN02], while TIARA [WLS*10] is one of the first systems that applied LDA in a ThemeRiver-based visualization, showing the evolution of topics over time. TextFlow [CLT*11] and RoseRiver [CLWW14] incorporate topic birth, merge, and death in order to analyze critical events by detecting emerging or diminishing topics. Dou et al. presented Paralleltopics [DWCR11] as an approach for understanding document characteristics. This approach does not focus on temporal changes in the topics, but rather on linking different visualizations to explore the underlying text content features. In contrast to ThemeRiver-based visualizations, which are typically used for visualizing the evolution of content over time, our approach regards the thematic structure of a debate as closed space and focuses on emphasizing the dynamics within a conversation and the interactions of speakers with respect to the content of the discussion.

Other tools focus on visualizing the topic structure of hierarchical topic modeling results and go beyond the usage of LDA for the topic generation, for example, Hiérarchie [SHM14]. Hierarchicaltopics [DYW*13] is an extension of the Paralleltopics [DWCR11] approach and uses an adaption of the Bayesian Tree Rose Algorithm [LWSG] to build topic hierarchies that can be modified by the user. In contrast, UTOPIAN [CLRP13] uses Nonnegative Matrix Factorization to create topic clusters, which can be interactively adjusted. In contrast tools such as Termite [CMH12] and Serendip [AKV*14] present more abstract model-driven visualization techniques. To avoid unstable spatial layouts, ConToVi is based on a topic modeling algorithm [EA15] which produces deterministic results, making multiple runs of the visualization on the same data comparable.

A particularly relevant approach for ConToVi is ArgVis [KLT11], a tool for interactively visualizing topic dynamics in multi-party conversations. This approach is based on the Speaker Identity for Topic Segmentation model [NBGR12], which identifies topic shifts in conversations. Along with the approaches such as Flash forums [DWM04] for navigating online discussion or ConVis [HC14] for the interactive exploration of asynchronous conversations, the ArgVis tool is one of the few approaches to visualize conversational text corpora. However, these dynamic visualization approaches do not have the same analytic focus as our framework.

From the visualization point of view, SolarMap [CGS*11] by Cao et al. is similar to the Topic-Space view introduced in this paper. However, this approach significantly differs in the usage of the mapping dimensions. In the outer part of their radial visualization, Cao et al. set different topic keywords, using the inner space to cluster similar documents and show topical themes of the analyzed corpus. In contrast, ConToVi uses the pre-computed hierarchical topic structure as outer dimensions to create a topic-space that is utilized to place conversation speakers and utterances.

Multi-Party Conversation Visualizations — More related to the analysis task of ConToVi are works that attempt to visualize some aspects of multi-party conversation, such as social interactions. Erickson and Kellogg [EK00] give a more general framework on important aspects to consider while designing what they call "socially translucent systems". The Topic-Space metaphor used in ConToVi can be considered an adaptation of the *social proxy* introduced in their paper. Other tools model the social interactions in chat systems to allow a more focused interaction on the participant dynamics, e.g. Chat Circles [DV02] and GroupMeter [LPH*09]. Approaches like Conversation Clusters [BK09] and Trains of Thought [SGH12] group the content of conversations dynamically for a better understanding of the thematic structure of conversations.

Radial Visualizations and Ring Plots — Radial methods for information visualization have been widely used for the visual analysis and exploration of different datasets. In their recent survey, Draper et al. [DLR09] categorize different kinds of radial visualization techniques. Under this categorization, ConToVi can be considered a ring plot, extending basic concepts of RadViz [HGM*97] and similar to Dimensional Anchors [HGP99]. Additionally, through the animated views, ConToVi is also categorized as a dynamic motion chart or temporal data visualization.

Most radial visualizations have to deal with major limitations regarding their scalability towards the number of outer dimensions in order to create meaningful views [DBB10]. The order of these dimensions significantly influences object placement in the inner space. According to Rubio-Sanchez et al. [RSRDS16], when using a force-directed layout technique for object placement, the force of opposite dimensions might be averaged out, creating a large central clutter. Moreover, as the number of dimensions increase, the force of each dimension is reduced, making the impact of each dimension on the placement of the objects harder to detect. To overcome these known limitations, ConToVi introduces a set of visual design solutions, discussed in the next section.

3. Topic-Space View

Adopting the metaphor of a discussion floor, Topic-Space views span all topics of a conversation as dimensions of a radial plot. A position in the 2D-Space can thus be used to encode given relations of an object to all topics instead of a combination of topics and time. This allows the focus of the analysis to be on the dynamics of a conversation, enabling the comparison of speaker patterns and sequence in addressing the topics. The Topic-Space is defined through aligning the most general topics in a radial fashion on the outside of the circular space. Each segment on the ring represents an upper-level topic and corresponds in its length to the size of the represented topic. The order of the ring segments is initially determined by the order of appearance of the topic in the conversation. This can, however, be changed by the user through selecting a topic of interest. In the center of the created space, a force-directed layout determines the position of each speaker according to the topic membership degrees of her statements. Each speaker is represented by a colored dot in the space which corresponds in its size to the amount of text associated with this speaker. Using these characteristics, a conversation can be explored using different views on the Topic-Space which are described in this section.

While such a radial visualization enables us to encode (for our task) more relevant information into the position of objects, it also introduces new challenges for the interpretation of a point in the Topic-Space. As described by Rubio-Sanchez et al. in their recent study [RSRDS16], RadViz plots [HGM*97] suffer from two main drawbacks, namely [1] the interpretation of positions inside the plot and [2] the scalability to a large number of data points, since these visualizations tend to generate a high central clutter. To address these challenges, we propose a number of design solutions. In order to disambiguate the topics contributing to the position of objects in the Topic-Space, we introduce topic glyphs and reorder the outer topic-dimensions according to their similarity to a selected reference topic.

Topic glyphs surround an object on-demand, displaying its normalized similarity to all outer dimensions of the space. These glyphs enable the users to understand why a certain object is positioned in this part of the Topic-Space. These glyphs are designed to correspond to the outer edge of the topic space, encoding the topics using angle and their similarity using color (with darker segments being more similar). To make the glyphs comparable across objects, the angles are not rotated to face the topics but stay fixed. This, however, requires a higher mental effort for interpreting the glyphs.

We interactively reorder the outer topic-dimensions according to their similarity to a selected reference topic. This topic is interactively selected by the user, depending on the focus of her analysis and all other topics are ordered clockwise with decreasing similarity to this reference topic (with the most similar topic as a direct successor of the reference topic). Using a black-to-white colormap, these similarities are additionally indicated. This step has the underlying assumption that utterances are most likely to be mentioning similar topics together rather than opposing topics. Thus, by ordering highly correlated dimensions beside each other on the outside of the plot, points in the space become more interpretable and the central clutter is reduced. In addition, through reordering the topics, both the Topic-Space and the topic glyphs become more scalable, since the influence of smaller topics is reduced by placing them next to a larger, similar topic.

Other solutions we implemented to reduce the central clutter of objects in the Topic-Space views are introducing a repelling weighted force from the center of the space, showing only a subset of the data at a time, and reducing the number of analyzed utterances through filtering and time-range selection.

The first measure directly influences object positioning in the Topic-Space, as we define repelling forces from the center of the topic space that contrast the attracting forces of the topic-dimensions on the edge of the space. By weighting these forces, we can reduce the central clutter in the space while maintaining the relative position of objects and thus the encoded information about the object's relation to the topics. The second step limits the number of objects shown in the space at a time. Depending on the analysis task, the users might be interested in the general movement of speakers in the space, the interaction of a selection of speakers, or the detailed analysis of one speaker. Using different views for the Topic-Space, we show only the relevant data objects for the analysis. Additionally, by using animation in some of these views, the amount of objects displayed at a given time point can be reduced considerably. Thus, the animation allows a focused analysis of specific time frames and reduces the overall visual clutter that would result from displaying the dynamics of the complete conversation in a single view.

All Topic-Space views implement a number of filtering methods to enhance data exploration. In addition to the speaker selection, a filtering of the utterances according to their attributes is supported. Using a range query, utterances are filtered according to numerical features. Logical queries filter them according to their speech and argumentation patterns, which are derived using statistical and linguistic features. Moreover, as only a small time-frame of the discussion might be of interest to an analyst, time-range filters enable the selection of an interesting analysis window.

Our approach is designed to analyze conversations in 4 different views, each corresponding to a general task. We derive the following 4 tasks from our requirement analysis: **[T1]** Exploration of the conversation and overview generation (addressing R1, R2, R3). **[T2]** Analyzing and understanding the speaker movement through topic space for selected speakers (addressing R1, R2, R4). **[T3]** Analyzing the contribution of each speaker to each topic over time (addressing R2, R3, R6). **[T4]** Analyzing the movement and patterns of one speaker in detail (addressing R4, R5, R6). In the following sections, each view is described in more detail.

3.1. Discussion Overview

The first step scholars usually take before analyzing discussion transcripts is watching videos or listening to audio recordings of the conversations. This gives them a first indication of the content and the social dynamics of a conversation. Moreover, the first exploration of a discussion on such an overview-level allows the analyst to identify portions of particular interest for their analysis tasks. In addition, interesting speech and argumentation patterns that are used by some speakers while addressing the different topics are first revealed through this exploration. To enable users of ConToVi to perform these tasks, we introduce the first view as a discussion overview visualization. This view is a simple animated visualization in the Topic-Space showing the movement and involvement of all participants of a conversation over time. It serves as a starting point for the analysis, allowing the users to mentally reconstruct the discussion dynamics.

The speaker movement through the Topic-Space is based on their turns in the conversation. In keeping the metaphor of a discussion floor in the Topic-Space, participants enter and leave the floor according to their level of activity. A point representing a speaker first appears in the space at the time of the first utterance of the speaker. The positioning of each speaker, i.e. the representation at a certain time point, corresponds to the position of her last turn. Each speaker point is assigned a different color and overlaid with an icon that indicates the speaker's position in the conversation. The color intensity of a dot at each time point corresponds to the amount of time that has passed between the speaker's last utterance and the current time of the animation. Therefore, dots representing speakers who have not talked recently begin to fade out. At the end of the animation, all speakers are shown with respect to their position according to the average placement of all their corresponding utterances. To facilitate the tracking of the speakers' movements, speaker-trails can be blended in this view, as described in the next section. For instance, Figure 2 demonstrates the movement of the three speakers of a presidential debate.

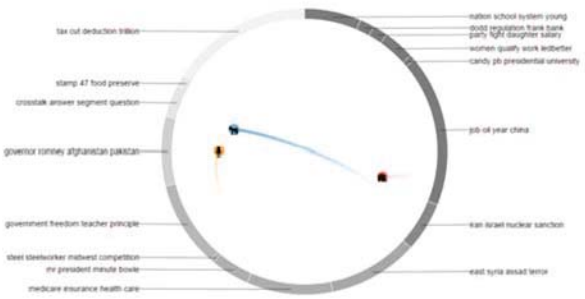


Figure 2: Speaker-Trails [V1] showing the movement of Barack Obama in the Topic-Space created for the Obama vs. Romney Presidential Debates.

3.2. Speaker-Trail-View

ConToVi displays speaker movement in the Topic-Space via animation. Due to motion blindness, a focused analysis of animated objects might prove to be more challenging than still objects. To overcome this issue, we introduce speaker-trails and topic-glyphs. In a trade-off between visual clutter in the Topic-space and the interpretability of the speaker movements, we designed the visual features to assist detection and comprehension of user movement changes over time.

Figure 2 shows an example of speaker-trails. These short path segments follow a speaker point during the animation and indicate the direction and length of a speaker's movement. Trail trajectories at each time point are defined by the topic membership of the underlying utterances and an interpolation between consecutive speaker utterances. It is calculated using b-spline interpolations with control point duplication on the utterance positions. This is the same interpolation that is used for drawing the complete speaker path in V4. The length of each trail indicates the distance covered in the topic space between two consecutive utterances and the color gradient reveals the direction of movement. As a default setting, the color of each trail corresponds to its speaker. However, in addition to showing the speakers' paths in the Topic-Space, a selection of linguistic or statistical features can be mapped to the thickness and color of the trails, opening a view on the speakers' behaviors towards a topic at a given time point. These features include the politeness and sentiment scores of the utterances, the certainty of the speaker, and her eloquence. Using such mappings, an analyst can explore attributes of speakers throughout the conversation and track their change between different topics. Moreover, she can discover similarities in speaker behaviors.

To facilitate the comprehension of the movements in the Topic-Space and its underlying utterances, the Speaker-Trail View allows the users to integrate the topic glyphs into the animation. Similar to the other objects in the space, these glyphs fade out after some user-defined time in order to avoid visual clutter. In addition, for a more focused analysis, users can inspect a single glyph by itself by hovering over it. The topic glyphs visualize the utterances as a mixture of topics and thus make the positioning of the utterances and their respective speakers more understandable. Figure 3 shows a screenshot of the animation of topic glyphs and speaker trails in the Topic-Space.

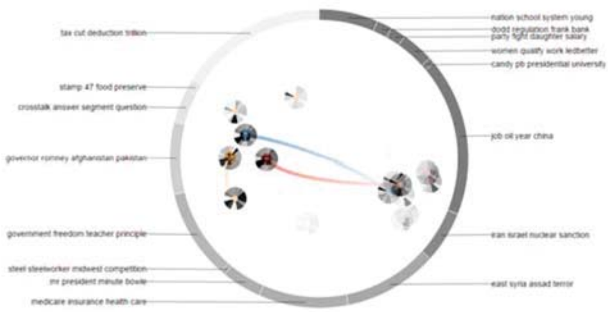
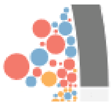


Figure 3: Topic Glyph overlays on top of the Speaker-Trail view [V2] showing the position of the last spoken utterance for the time-frame of the animation.

3.3. Utterance-Sedimentation View

To analyze the contributions of speakers to the topics over time, we propose the Utterance-Sedimentation View. This animated visualization allows an analysis not only on the speaker-level but also on the utterance-level. To display all previous utterances to the animation time, we simplify the membership information of utterances to topics and regard the top topic for each utterance as its main topic. In the metaphoric sense of the Topic-Space, this corresponds to the main topic of each utterance having the largest pulling force that overshadows all other topics. In order to achieve a coherent visualization, we apply a visual sedimentation metaphor to the Topic-Space. By using a physics engine, we, therefore, define a gravitational force for each topic-dimension that attracts its corresponding utterances. As the speakers move through the topic space, they "place" their utterances in the space which then gets pulled out by the topic gravity. These utterances are then piled up at the edge of the space. Each utterance in this view is represented by a colored dot with the size corresponding to the length of its underlying text. As the sizes of utterances are defined to all fit into their corresponding segment in the Topic-Space, we do not aggregate the utterances as is usual for visual sedimentation-based visualizations. This choice has



the main advantage that we can roughly retrace the sequence of utterances in a topic-segment by their distance to the edge of the space. Utterances close to the edge are spoken at the beginning of the conversation and utterances close to the center of the space occurred later in the discussion, as shown in Figure 4.

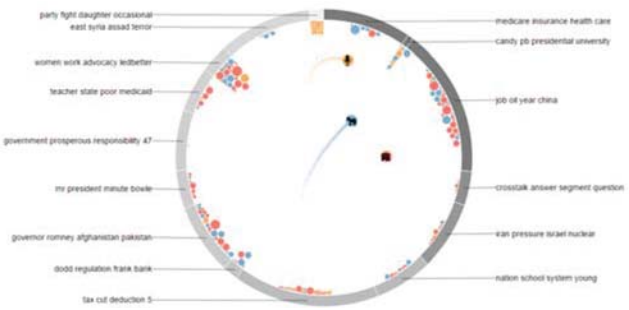


Figure 4: Utterance-Sedimentation view [V3] for the presidential debate of Obama vs. Romney.

3.4. Speaker-Path View

The last view features the complete speaker path for a selected speaker. This view shows a static path that indicates the complete movement of a speaker through the conversation. As with the speaker-trails, these paths are calculated using b-spline interpolations. In contrast to linear interpolation, the calculated splines avoid occlusion of path segments. Depending on the activity and participation of a speaker in the conversation these paths might get visually cluttered and thus harder to understand. Therefore, we include the feature of retracing the paths using the animated speaker dots in order to reconstruct the movement patterns. Using the various selection and filtering methods embedded in ConToVi , the overplotting of long paths can be reduced. The main task of this view is to enable a focused analysis of one speaker and to allow the comparison of multiple speaker paths. Comparing paths can be achieved through overlay in the same Topic-space or by a side-by-side comparison. This feature is important to address tasks related to speaker similarity. Figure 5 shows a comparative view of three participants of the Stuttgart 21 arbitration. Hereby, the topics addressed by each speaker become clear. The moderator of the arbitration Dr. Heiner Geißler moves back and forth between all topics over the course of the discussion, while the other speakers address topics more specifically.

In addition to the animated speaker dots, this view enables the overlay of the different design elements, such as topic or argumentation glyphs. Moreover, the width and color saturation of the path can be used to encode information about the underlying utterances, such as length or sentiment value. Interacting with these overlays enables the users to get more details about the utterances and speakers on demand and facilitates the detailed analysis of the linguistic and statistical features of the text. Such a detailed analysis of single utterances is key to proving the hypotheses of the analysts and allows the generation of real empirical evidence through going down to the annotated text level. For analyzing speech and argumentation patterns, users, therefore, rely on the interaction with the argumentation glyphs and the speaker paths. These glyphs represent abstractions of important features for their analysis and are the doorway to the concrete texts. The design of the argumentation glyphs is discussed in the next section.

4. Speech and Argumentation Patterns

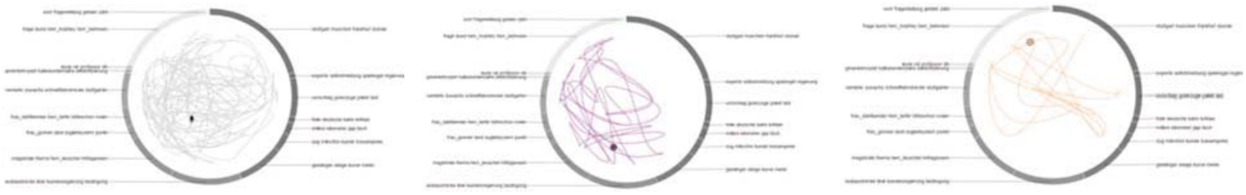
When analyzing conversations, analysts are usually trying to find measurable evidence to support their hypotheses about the dynamics of a conversation. Depending on the analysis task at hand, different statistical and linguistic features may form a solution. In contrast to the task-independent algorithms deployed for generating the Topic-Space views, measures for detecting task-relevant patterns are usually depending on the language usage and general text quality. In order to satisfy the requirements of the analysis, we need to tailor the choice and visual design of the measures to the concrete analysis. Hence, in the context of political discourse analysis, we extract features that highlight the dynamics of a political debate.

For assessing the deliberative quality of debates, speech and argumentation patterns of speakers are particularly interesting. Together with our domain expert we derived a set of features that are inferred from the theoretical dimensions of deliberation based on Steenbergen et al. [SBSS03]. For detecting speakers' speech patterns, we use a rule-based annotation system to extract speaker stances. In general, the approach is based on a lexical-grounded procedure [BHJSB14], annotating the occurrence of deliberative relations within utterances. These relations are defined through a combination of nominal attributes that are extracted using a set of rules from the text. In essence, the rules disambiguate the usage of functional words and connectors in the text that indicate speaker stances towards their utterances and define the parts of the text that indicate these stances. For our tool, we use the most common 8 speaker stances to characterize the utterances. Since the usage of other stances is generally very infrequent they are not shown in our visualizations.

To model the argumentation patterns of speakers we extract the scope of arguments within an utterance with respect to justification, i.e. reason and conclusion. Using the lexical-grounded procedure, we extract not only the occurrence of justification in an argument but also its scope within an utterance. In general, the level of uncertainty in the identification of these markers is low.

4.1. Argumentation Glyphs

The argumentation glyphs are abstract representations of some important features of speech and argumentation. For the analysis of deliberation, we designed the argumentation glyph to encode the above-mentioned features, in addition to



(a) Moderator: Dr. Heiner Geißler (b) Project supporter: Dr. Volker Kefer (c) Project opponent: Michael Holzhey

Figure 5: Speaker-Path view [V4] for different speakers of the Stuttgart 21 arbitration.

the length of an utterance. These 11 dimensions are carefully mapped to visual elements in the glyphs to give a quick understanding of the mapping and a fast recognition. All ten speech and argumentation attributes have a double visual encoding to enhance a pre-attentive cognition of the information, in compliance with Borgo et al. [BKC*13] design guidelines (DG5). The size of the glyph corresponds to the length of a given utterance. For the argumentation patterns, the degree of justification of a speaker is measured using two indicators: (A) Reason, i.e. explanation and justification of the argument. (B) Conclusion, i.e. results of the argument. Both indicators are measured in relation to the utterance length and are visually represented by the outer segments of the glyph. These segments encode the relative amount of both indicators by their size and color (white/light-gray for 0% to black for 100%). The eight nominal dimensions of speaker stances are mapped to equally-sized circle-segments in the inside of the glyph. These features are mapped to both color and angle (position). The choice of the colors ensured that dimensions that highly-correlate in practice are assigned different colors.

The argumentation glyphs can be mapped on to the different views of the Topic-Space to represent an abstraction of the utterances. They are used interchangeably with the topic glyphs, depending on the analysis task. In Figure 6, a subset of the glyphs of three speakers from the Stuttgart 21 arbitration is shown. With the help of this table view, the evolution of the argumentative structure of a speaker and her behavior can be analyzed. While analyzing the patterns for different speakers, users tend to immediately detect typical glyphs for each speaker and thus are able to spot outliers in the conversation. When an outlier is found, it is important for analysts to be able to analyze the underlying text of its corresponding utterance manually to disclose automatic processing mistakes and to verify their hypotheses. For such a detailed inspection, we provide a text-level analysis.

4.2. Text-Level Analysis

Across all components of ConToVi , the analyst can explore a discussion on different levels of detail, starting with the most general overview animations and digging down to interesting argumentation glyphs. To verify and comprehend these abstract glyphs, a detail panel with additional information about the underlying glyph

(a) Moderator: Dr. Heiner Geißler

(b) Project supporter: Dr. Volker Kefer

(c) Project opponent: Michael Holzhey

Figure 6: Glyph-based representation for three participants over the course of the debate. Each glyph represents one utterance. For the glyph in the rectangle, a detailed view is provided in Figure 7.

is shown on demand. Figure 7 shows the detail panel for a selected glyph of the moderator of the Stuttgart 21 arbitration. In this view, the position of the utterance in relation to the complete discussion is highlighted in the upper horizon-graph. This graph represents the conversation timeline, with the upper two values corresponding to the amount of *Reason* and *Conclusion* in a conversation and the lower values corresponding to the eight speaker stances. On the left-hand side of the detail panel, the complete text of the selected utterance is shown, highlighting the occurrence of speech or argumentation features.

5. Expert Case Study

To evaluate our framework, we conducted an expert case study with two political scientists with a research interest in discourse analysis and expertise in deliberative communication analysis. Both experts were involved in the initial requirements analysis study and were regularly consulted during the development of the framework. However, only one of the two **[E1]** was part of the design team and the other **[E2]** was completely unfamiliar with the visual design of ConToVi . Hence, we could observe the time it takes E2 to become familiar with the system. Both experts were given two real-world political debates to analyze over a period of one hour. Due to familiarity with the general contexts—but not the exact content—of the debates, we asked them to state their assumptions for the conversation dynamics. After deriving a set of hypotheses, the experts used our system to examine them while analyzing the debates.

Figure 7: Panel providing a detailed view on the glyph dimensions and its content. The top chart illustrates the complete conversation and indicates the position of the glyph in the discourse. The left panel presents the content of the utterance; on the right, the glyph and its dimensions are given.

5.1. Datasets

The first conversation is the publicly available (transcribed) dataset of the Stuttgart 21 (S21) arbitration on a controversial railway and urban development project in Germany. The complete arbitration builds a corpus of around 6000 utterances, involving 60 speakers.

The second data source we used for the expert case study are the transcriptions of the televised US presidential election debates. The main objective of these debates is to persuade yet undecided voters to vote for the party of the candidate. These debates typically contain fewer off-topic fragments and cross-talk segments than the arbitration. On average, each of the presidential debates has 160 to 180 utterances.

5.2. Verifying Deliberative Hypotheses

ConToVi was designed based on a set of general requirements [R1-R6] for the analysis. These are the basis to answer more specific questions. On the level of **Speaker Movement Analysis**, questions include: **[Q1]** Who is leading/following/most influential in the conversation? **[Q2]** Which speakers introduce new topics or switch between topics? **[Q3]** Which topics are most addressed by whom? **[Q4]** What is the influence of the moderator on the conversation? As for **Behavior Pattern Analysis**, some questions are: **[Q5]** Which linguistic features are characteristic for a speaker/group? **[Q6]** What is the "attitude" of a speaker towards a given topic? **[Q7]** How similar are movement, speech, and argumentation patterns of selected speakers? These and other questions related to the dynamics of conversations inspired the two domain experts to come-up with relevant hypotheses for each of the use cases. In the following paragraphs a selection of the hypothesis of both experts are presented, grouped according to their respective dataset.

Stuttgart 21 With respect to this arbitration the first hypothesis of E1 concerned the moderation style of the arbitrator Dr. Heiner Geißler. Based on his previous knowledge about previous moderations, E1 expected that **[H1]** Dr. Geißler is steering the conversation through an active moderation style and expected that he will be often participating in non-moderational topics and thus contribute to the content of the conversation. To verify his hypothesis, E1 first watched the animation of V1 to track the movement of all speakers in the Topic-Space. By focusing on the movement of the moderator he could find different points in time where Dr. Geißler changed the topic, followed by all active speakers. To verify his involvement in the topics, E1 then switched to V3. E1 found several large utterances associated with Dr. Geißler in non-moderational topics, in addition to the small moderational utterances that were mainly clustered in a moderation topic. Through this analysis E1 concluded that Dr. Geißler indeed follows an active moderator role and thus verified his hypothesis. With respect to the three-fold categorization of mediation styles used in political science research, E1 could identify the arbitrator relying on the manipulative style—a mediation style that is characterized by a high influence of the negotiation process, the content, and even the outcome. E1 appreciated the clarity of V1 which enabled him to focus on the conversation dynamics without distractions and enjoyed watching accumulation of utterances at the Topic-Space edges in V3, showing the history of the conversation while it evolved. E1 said that through identifying large topic segments that did not have many utterances attached to them, he could, in addition to knowing the past, predict the future of the conversation, since these segments are designed to contain a large number of utterances by the end of the animation.

On the other hand, E2 was more interested in analyzing the collective behavior of all participants of the arbitration. Through her knowledge about the structure of this discussion, she expected that **[H2]** the debate starts with small focused discussions and presentation with a smaller number of speakers and then the discussion floor is opened for a broader debate with all participants. Using V3, she was able to reconstruct the sequence of speakers and the discussed topics. While following along the animation, E2 could confirm her hypothesis, since the debate involved more speakers and a faster switching between topics towards the end of the discussion. Figure 1 displays one time-slice of this animation, where two speakers are moving in sequence.

In addition to the behavior of single speakers, E2 was also interested in the behavior of speaker groups, i.e. speakers that have the same position in the argumentation. She expected **[H3]** that each group of participants would want to push their own topics. This hypothesis contradicts one of the main assumptions of deliberation: to respect the positions and reasons of others and to allow an inclusive and intense debate. To analyze these dynamics, E2 colored all participants of a group in the same color and watched the accumulation of their utterances over time in V3. She concluded that while this claim could be partly true, she could not confirm it on this broader group level. However, by analyzing the individual speaker paths in V4, she was able to identify similarities and differences between speakers of the same or opposing parties, respectively. Figure 5 shows the paths of three speakers representing different positions in the arbitration: As the individual paths reveal, the arbitrator Dr. Geißler addresses most topics, while both other speakers only address a subset of the topics. E2 was surprised to find Dr. Kefer addressing only so few topics, since she expected him to be more deliberatively involved in the debate.

With respect to deliberation and in particular the attitude of speakers, E1 and E2 could confirm their hypotheses that were derived from a qualitative inspection of the conversation. In general, E1 and E2 made extensive use of the sequence of speakers combining the overview with a detailed inspection of the content. For instance, they identified two strategies of the moderator Dr. Geißler **[H4]** (see Figure 6): His utterances are characterized by linguistic particles establishing a common ground between the counterparts. In each of his utterances, he indicates that the facts and propositions are already known to the opponents and expects the opponents to implicitly agree upon. The moderator also consistently tries to achieve mutual consent by using particles demonstrating consensus willing. Both experts were especially interested to see the reactions to the statements by the moderator that involved such particles. Consequently, they have found that this demonstrates the importance of an active moderator within the debate. For the other participants in the arbitration, similar patterns could be identified.

US Presidential Debates For this study we presented the experts with three presidential election debates between President Obama and Governor Romney, concatenated. In contrast to S21, these debates have a clearer format with defined question and answer times. Therefore the moderators are usually more passive. However, after watching the animated discussion overview V1, E1 suggested that **[H5]** the moderator might not only be passive, but unheard on several occasions. He found various instances where the moderator attempts to change the topic and the two candidates just continue on talking. He then commented that this is a common phenomenon in presidential debates when it comes to controversial topics.

On the other hand, E2 was interested in the dynamics between participants. She noticed that **[H6]** throughout the debate, Obama takes a defensive role in contrast to Romney, who appears to be more offensive. This finding is based on the topic glyphs in V2, where she noticed that Romney is actively introducing new topics, while Obama addresses these topics afterwards. She first thought this is due to the speaker turns in the debate, but as this became a notifiable trend throughout the discussion, she started a more detailed analysis using V3. By following the sequence of utterances in each topic and differentiating between utterances of different speakers, E2 concluded that Romney indeed has more instances of first-topic-switches than Obama. These subjective findings were confirmed through a text-level analysis of relevant utterances.

Overall Feedback The political science scholars were very eager to explore the discussions and debates using the interactive Topic-Space views of ConToVi. Expert E2, who was unfamiliar with the system, analyzed the debates within the same amount of time than E1. Both experts were interested in exploring the conversation in the multiple levels of detail provided by the tool and appreciated the ability to dig deeper for more information during the analysis. The power of the visual representation and abstraction of the datasets allowed the analysts to get a new perspective of the conversations and, through focusing on the relevant aspects, to find new patterns and outliers, even in the discussions they were already familiar with. Moreover, the overlay of the topic and argumentation glyphs was perceived as particularly helpful to understand the positioning of points in the space and to identify speech and argumentation patterns for different speakers. Both experts agreed that the use of ConToVi has an added value to their research on deliberation and they see potential for future applications.

6. Conclusion and Future Work

We presented ConToVi, a novel visual analytics approach for multiparty conversation exploration that assists political science scholars to identify deliberative patterns debates. We introduce the concept of a Topic-Space, relying on the metaphor of a discussion floor. In an observational study, we derived six important requirements [R1-R6] for the analysis of conversations and addressed these by designing four different views [V1-V4] for the Topic-Space. The main contributions [C1-C6] of our work provide a coherent visualization solution to a radial layout of topic-dimensions and propose measures to overcome the drawbacks that might occur with this layout strategy.

We evaluate our approach using two real-world political debates and through an expert case study, we demonstrated the use of ConToVi to verify the hypotheses [H1-H8] of two political science scholars. Furthermore, ConToVi can be used to analyze any kind of conversational text data; it is not restricted to the use of oral communication, but can also be applied to online forum discussions or chats. With the exception of the specific language features for the speech and argumentation analysis, ConToVi is language independent and can be used to analyze debates in any language. Hence, ConToVi provides a generic framework to for the analysis of dynamics in multi-party conversations.

Our approach has been well received and has been found to be intuitive and understandable. The users of the tool appreciated the consistency in the usage of the visual metaphor, which helped them create a mental model of conversations that it tightly coupled with the visualizations of ConToVi. However, the visualization has also some limitations. The main open task for the visual design is to expand the detail provided in the topic labels and optimize their placement, which is part of our future work. Moreover, the navigation through time will be also supported by directly moving the speaker points in the overview visualization, similar to the direct interactions in DimpVis [KC14]. Lastly, we plan to extend the automatic linguistic feature generation by including the measurement of direct speaker interactions, in order to visualize when speakers are directly talking to each other.

References

[AKV*14] ALEXANDER E., KOHLMANN J., VALENZA R., WITMORE M., GLEICHER M.: Serendip: Topic model-driven visual exploration of text corpora. In *Proc. of IEEE Conf. on Visual Analytics Science and Technology* (Oct 2014), pp. 173–182. 3

[Ast01] ASTON G.: Text categories and corpus users: a response to David Lee. *Language learning and technology 5*, 3 (2001), 73–76. 1

[ASW12] ANGUS D., SMITH A., WILES J.: Conceptual recurrence plots: Revealing patterns in human discourse. *Visualization and Computer Graphics, IEEE Transactions on 18*, 6 (2012), 988–997. 2

[BBGSG10] BLACK L. W., BURKHALTER S., GASTIL J., STROMER-GALLEY J.: Methods for Analyzing and Measuring Group Deliberation. In *Sourcebook of Political Communication Research: Methods, Measures, and Analytic Techniques*, Bucy E. P., Holbert R. L., (Eds.). Routledge, New York, NY, 2010, ch. 17, pp. 323–345. 1

[BHJSB14] BÖGEL T., HAUTLI-JANISZ A., SULGER S., BUTT M.: Automatic Detection of Causal Relations in German Multilogs. In *Proc. of the EACL 2014 Workshop on Computational Approaches to Causality in Language* (Gothenburg, Sweden, 2014), ACL, pp. 20–27. 6

[BK09] BERGSTROM T., KARAHALIOS K.: Conversation clusters: grouping conversation topics through human-computer dialog. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2009), ACM, pp. 2349–2352. 3

[BKC*13] BORGO R., KEHRER J., CHUNG D. H., MAGUIRE E., LARAMEE R. S., HAUSER H., WARD M., CHEN M.: Glyph-based visualization: Foundations, design guidelines, techniques and applications. *Eurographics State of the Art Reports* (2013), 39–63. 7

[BNJ03] BLEI D. M., NG A. Y., JORDAN M. I.: Latent dirichlet allocation. *J. Machine Learning Research 3* (2003), 993–1022. 3

[CGS*11] CAO N., GOTZ D., SUN J., LIN Y.-R., QU H.: SolarMap: Multifaceted Visual Analytics for Topic Exploration. In *Proc. of IEEE Int. Conf. on Data Mining* (Dec 2011), pp. 101–110. 3

[Cha03] CHAMBERS S.: Deliberative Democracy Theory. *Annual Review of Political Science 6*, 1 (2003), 307–326. 1

[CLRP13] CHOO J., LEE C., REDDY C. K., PARK H.: UTOPIAN: User-Driven Topic Modeling Based on Interactive Nonnegative Matrix Factorization. *IEEE Trans. on Visualization and Computer Graphics 19*, 12 (Dec. 2013), 1992–2001. 3

[CLT*11] CUI W., LIU S., TAN L., SHI C., SONG Y., GAO Z., QU H., TONG X.: Textflow: Towards better understanding of evolving topics in text. *IEEE Trans. on Visualization and Computer Graphics 17*, 12 (dec. 2011), 2412 –2421. 3

[CLWW14] CUI W., LIU S., WU Z., WEI H.: How hierarchical topics evolve in large text corpora. 3

[CMH12] CHUANG J., MANNING C. D., HEER J.: Termite: Visualization techniques for assessing textual topic models. In *Proc. of Int. Conf. on Advanced Visual Interfaces* (2012), ACM, pp. 74–77. 3

[CSL*10] CAO N., SUN J., LIN Y., GOTZ D., LIU S., QU H.: Facetatlas: Multifaceted visualization for rich text corpora. *IEEE Trans. on Visualization and Computer Graphics 16*, 6 (2010), 1172–1181. 3

[DBB10] DIEHL S., BECK F., BURCH M.: Uncovering strengths and weaknesses of radial visualizations–an empirical approach. *IEEE Trans. on Visualization and Computer Graphics 16*, 6 (2010), 935–942. 3

[DLR09] DRAPER G., LIVNAT Y., RIESENFELD R. F.: A survey of radial methods for information visualization. *IEEE Trans. on Visualization and Computer Graphics 15*, 5 (2009), 759–776. 3

[DV02] DONATH J., VIÉGAS F. B.: The chat circles series: explorations in designing abstract graphical communication interfaces. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques* (2002), ACM, pp. 359–369. 3

[DWCR11] DOU W., WANG X., CHANG R., RIBARSKY W.: Paralleltopics: A probabilistic approach to exploring document collections. In *Proc. of IEEE Conf. on Visual Analytics Science and Technology* (oct. 2011), pp. 231 –240. 3

[DWM04] DAVE K., WATTENBERG M., MULLER M.: Flash forums and forumReader: navigating a new kind of large-scale online discussion. In *Proc. of ACM Conf. on Computer Supported Cooperative Work* (2004), ACM, pp. 232–241. 3

[DYW*13] DOU W., YU L., WANG X., MA Z., RIBARSKY W.: Hierarchicaltopics: Visually exploring large text collections using topic hierarchies. *IEEE Trans. on Visualization and Computer Graphics 19*, 12 (2013), 2002–2011. 3

[EA15] EL-ASSADY M.: *Incremental Hierarchical Topic Modeling for Multi-Party Conversation Analysis*. Master's thesis, University of Konstanz, 2015. 2, 3

[EK00] ERICKSON T., KELLOGG W. A.: Social translucence: an approach to designing systems that support social processes. *ACM transactions on computer-human interaction (TOCHI) 7*, 1 (2000), 59–83. 3

[HC14] HOQUE E., CARENINI G.: Convis: A visual text analytic system for exploring blog conversations. In *Computer Graphics Forum* (2014), vol. 33, Wiley Online Library, pp. 221–230. 3

[HGM*97] HOFFMAN P., GRINSTEIN G., MARX K., GROSSE I., STANLEY E.: DNA visual and analytic data mining. In *Proc. of Visualization '97* (Oct. 1997), pp. 437–441. 3, 4

[HGP99] HOFFMAN P., GRINSTEIN G., PINKNEY D.: Dimensional anchors: a graphic primitive for multidimensional multivariate information visualizations. In *Proc. of ACM Int. Conf. on Information and Knowledge Management* (1999), ACM, pp. 9–16. 3

[HHWN02] HAVRE S., HETZLER E., WHITNEY P., NOWELL L.: Themeriver: Visualizing thematic changes in large document collections. *IEEE Trans. on Visualization and Computer Graphics 8*, 1 (2002), 9–20. 3

[KC14] KONDO B., COLLINS C.: DimpVis: Exploring Time-varying Information Visualizations by Direct Manipulation. *IEEE Trans. on Visualization and Computer Graphics (Proc. of the IEEE Conf. on Information Visualization) 20*, 12 (Dec. 2014). 9

[KLT11] KARAMANOU A., LOUTAS N., TARABANIS K.: Argvis: Structuring political deliberations using innovative visualisation technologies. In *Proc. of the Third IFIP WG 8.5 Int. Conf. on Electronic Participation* (Berlin, Heidelberg, 2011), ePart'11, Springer-Verlag, pp. 87–98. 3

[LPH*09] LESHED G., PEREZ D., HANCOCK J. T., COSLEY D., BIRNHOLTZ J., LEE S., MCLEOD P. L., GAY G.: Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (2009), ACM, pp. 537–546. 3

[LWC*14] LIU S., WANG X., CHEN J., ZHU J., GUO B.: Topicpanorama: A full picture of relevant topics. In *Proc. IEEE Conf. on Visual Analytics Science and Technology* (Oct 2014), pp. 183–192. 3

[LWSG] LIU S., WANG X., SONG Y., GUO B.: Evolutionary Bayesian Rose Trees. 3

[NBGR12] NGUYEN V.-A., BOYD-GRABER J., RESNIK P.: SITS: A Hierarchical Nonparametric Model using Speaker Identity for Topic Segmentation in Multiparty Conversations. In *Association for Computational Linguistics* (2012). 3

[RSRDS16] RUBIO-SANCHEZ M., RAYA L., DIAZ F., SANCHEZ A.: A comparative study between RadViz and Star Coordinates. *IEEE Trans. on Visualization and Computer Graphics 22*, 1 (2016), 619–628. 3, 4

[SBSS03] STEENBERGEN M. R., BÄCHTIGER A., SPÖRNDLI M., STEINER J.: Measuring Political Deliberation: A Discourse Quality Index. *Comparative European Politics 1*, 1 (2003), 21–48. 1, 6

[SGH12] SHAHAF D., GUESTRIN C., HORVITZ E.: Trains of thought: Generating information maps. In *Proceedings of the 21st international conference on World Wide Web* (2012), ACM, pp. 899–908. 3

[SHM14] SMITH A., HAWES T., MYERS M.: Hiérarchie: Interactive Visualization for Hierarchical Topic Models. In *Proc. of the Workshop on Interactive Language Learning, Visualization, and Interfaces* (2014), p. 71. 3

[WLS*10] WEI F., LIU S., SONG Y., PAN S., ZHOU M. X., QIAN W., SHI L., TAN L., ZHANG Q.: TIARA: A visual exploratory text analytic system. In *Proc. of ACM Int. Conf. on KnowledgeDiscovery and Data Mining* (2010), KDD '10, ACM, pp. 153–162. 3

[YMSJ05] YI J. S., MELTON R., STASKO J., JACKO J. A.: Dust & Magnet: Multivariate Information Visualization using a Magnet Metaphor. *Information Visualization 4*, 4 (2005), 239–256. 2