Augmenting Public Deliberations through Stream Argument Analytics and Visualisations

Brian Plüss*, Fabian Sperrle†, Valentin Gold‡, Mennatallah El-Assady†,
Annette Hautli-Janisz§, Katarzyna Budzynska*,¶ and Chris Reed*

*Centre for Argument Technology, University of Dundee, Dundee, UK
E-mail: b.pluss@dundee.ac.uk, k.budzynska@dundee.ac.uk, c.a.reed@dundee.ac.uk

†Computer and Information Science, University of Konstanz, Konstanz, Germany
E-mail: fabian.sperrle@uni-konstanz.de, menna.el-assady@uni-konstanz.de

‡Center of Methods in Social Sciences, University of Goettingen, Goettingen, Germany
E-mail: valentin.gold@sowi.uni-goettingen.de

§Department of Linguistics, University of Konstanz, Konstanz, Germany
E-mail: annette.hautli@uni-konstanz.de

¶Institute of Philosophy and Sociology, Polish Academy of Sciences, Warsaw, Poland

***Abstract*—Public deliberations are organised by governments and other large institutions to take the views of citizens around controversial issues. Increasing public demand and the associated burden on public funding make the quality of public deliberation events and their outcomes critical to modern democracies. This paper focuses on technology developed around streams of computational argument data intended to inform and improve deliberative communication in real time. Combining state-of-the-art speech recognition, argument mining, and analytics, we produce dynamic, interactive visualisations intended for non-experts, deployed incrementally in real time to deliberation participants via large screens, hand-held and personal computing devices. The goal is to bridge the gap between theoretical criteria on deliberation quality from the political sciences and objective analytics calculated automatically from computable argument data in actual public deliberations, presented as a set of visualisations which work on stream data and are simple, yet informative enough to make a positive impact on deliberative outcomes.**

***Index Terms*—argument analytics, public deliberations, deliberative democracy, stream visualisations.**

I. Introduction

In this paper, we describe first results towards building a visual interventions system to improve public deliberations. The main focus of this paper is on the stream processing technique to augment deliberations in real-time. We introduce the software architecture and describe deliberation analytics and public-oriented visual augmentations aimed at improving the quality of the deliberation. This effort contributes to the "digitalisation" of society by developing a visual discourse architecture to track, analyse, and facilitate face-to-face participatory processes in real-time (see the conceptual scenario in Figure 1).

The inclusion of citizens into political decision-making processes has gained attention in recent years, mainly due to an increased demand by citizens to *have a say, get heard* in political decisions. Citizens lack of understanding of or agreement with decisions taken by their formally elected representatives have resulted in dissatisfaction with regular parliamentary political decision-making procedures (cf. [1], [2]). In some instances, disagreement became obvious, resulting in large protests involving hundreds and thousands of people[1]. Recent electoral success of populist parties may arguable be linked to the persistent alienation of citizens from politics [3]. Overall, as decision-making becomes more problematic for governments, citizens, politicians, and non-governmental organisations have suggested new methods and designs to put citizens and public stakeholders in (partial) decision-making authority. Examples include referenda, plebiscites, participatory budgeting, town

Fig. 1: An augmented public deliberation.

[1] In 2010, around 50K people protested against the reconstruction of the main railway station in Stuttgart, Germany, even though the decision had passed all relevant political councils.

LEVIA'18: Leipzig Symposium on Visualization in Applications 2018

hall meetings, citizen hearings, and public consultations. They all share an increased burden on organisational costs and public spending[2].

Contrary to previous efforts in building deliberation support systems (for an overview see [7]), we propose an architecture that works on stream data and is applied to face-to-face deliberative debates. Existing automated systems are only applied to online, computer-mediated discussions [7]. Research into such support systems has found that problem-solving abilities increase when using argumentation mapping software [8]; online mapping tools lead to more evidence-based discussions [9]; decisions are taken more quickly when individual argumentation maps become visible to the other participants [10]; and these software solutions can help to reach consensus [11]. From these results, it follows that support systems would contribute to face-to-face deliberations. However, there is currently a lack of visual systems mapping arguments and other important dimensions of deliberative communication in face-to-face deliberations.

Research comparing online and offline deliberations has led to inconsistent conclusions on which collaboration mode should be preferred. Some studies conclude that both modes result in rather similar positive changes on the attitudes of citizens and no significant differences can be seen (e.g. [12], [13]). Yet, offline deliberation appears to be more effective in influencing citizens' willingness to participate in politics [14]. Similarly, offline deliberations are characterised by higher quality of deliberation [15], and show stronger effects on attitudes towards policy [16].

Designing a visualisation architecture for face-to-face deliberations is significantly more challenging than for online deliberations. Such discussions are spontaneous in nature: when people communicate, immediate responses are given without long consideration. Consequently, the visual support must be instantaneous, relevant, and as simple as possible, while effectively supporting participants in their deliberation tasks. The real-time nature of this processing makes it inherently complex. For an efficient analysis, strict runtime limitations apply. If these cannot be fulfilled as the discussion moves on, any processed data lose relevance. This becomes increasingly more challenging over the duration of a debate as the amount of data increases. As many analytic models (e.g. for arguments, topics, and named entities) were created without streaming data in mind, significant efforts are needed to make them compatible. This also applies to the visualisations, as they need to be updated over time. However, newly incoming data cannot simply be added to visualisations as they arrive. Instead, to avoid cluttering, information that becomes irrelevant needs to be hidden or removed. This can be achieved by reducing the level of detail, aggregating data, collapsing tree branches, or using sliding-windows.

[2] Costs in Germany are estimated at €300K for each participatory process [4], although the now infamous 9-day public arbitration on the aforementioned reconstruction of the Stuttgart railway station – later dubbed "a failure of deliberative democracy" [5] – involved 230 participants, was broadcast live on TV, and greatly exceeded the €300K estimate. In the UK, citizen participation costs each local government around €2.5 million per year, with a total annual financial burden of around €1.3 billion on the taxpayer [6].

As visualisations are presented to non-experts, they must be comprehensible without extensive explanation. Simplicity is a requirement for the visual design to ensure that participants can instantly absorb the content of every visual element and their interpretations in context with low cognitive load. Another important challenge lays in the dynamic design of the visual elements. As said, the system must carefully incorporate recent actions in the debate, incrementally processing data as new content becomes available. The algorithms for analysing deliberation must also ensure they do not omit important aspects of the debate and must – at all times – objectively reflect the current state of the deliberation. Consequently, there is a need to balance between visual updates, coherence, objectivity, and simplicity.

In general, we aim for an architecture that adds significant augmentation as content layers which support participants during public deliberations. Ultimately, we aim to build an intervention system to achieve high quality deliberative communication. This paper presents the first steps in this effort. We draw support from the recent advances in computational and visual analytics and present disruptive technology to the rapidly expanding domain of deliberative democracy.

II. Background

A. Deliberative Communication

Deliberation promotes a very specific type of communication that emphasises normative rationality and public reasoning. In its extended form, deliberation aims for a rationally motivated consensus instead of taking decisions based on a majority vote [17]. It expects participants to communicate in an inclusive and respectful manner [18]. Moreover, decision-making should be based on extensive justification [19] and, following Habermas, it is the power of the better argument that should lead to agreement and decisions [17]. While the theory originally postulates normative maxims for deliberation, the demand for applications and computational tools have led to an increased empirical understanding of deliberation decision-making.

Research in the field of deliberative communication has concentrated on three aspects of deliberative decision-making: antecedents, dynamics, and consequences [20]. While a multitude of work focuses on the institutional and individual antecedents and their effects on deliberation and decision-making (see e.g. [21], [22]), fewer studies analyse the actual dynamics of deliberative communication. The current framework of analysis is based on manual reading and highlighting of interesting deliberative sections within a discourse (e.g. [23]). Sequential effects, however, do matter and presenting visual feedback to the participants is expected to activate potentials for better deliberation.

B. Argument Analytics

Argument technologies rely on computational models and tools to process natural language argument. The area has flourished in recent years, taking a discipline traditionally linked to philosophy and the law, to the foreground of AI and the news [24], [25], [26]. The Argument Web [27], for instance, is an ecosystem for analysing, storing, and processing argument structures represented as a network of propositions (i.e. claims and evidence) connected by relations such as support and attack. The Argument Interchange Format (AIF) [28] is a domain-independent, extensible formalism to represent argument structures. Arguments are stored in AIFdb [29], [30], a public repository of argument data with more than 30 thousand argument maps in several languages. Among other extensions to AIF, AIF^+ now supports the extraction of arguments from dialogues [31], following a theory in which argument structures are anchored to speaker utterances via illocutionary intentions [32]. This allows for the systematic identification of argument structures in transcripts of public deliberation, manually and, increasingly so, automatically via argument mining techniques [33], [34].

Once arguments are in a format amenable of computational treatment, it is possible to compute properties of the argument network to shed light on the nature of the deliberation. Such argument analytics [35], [36], [37] can then be turned into visualisations and presented to stakeholders. Recent efforts have aimed to make visualisation of argument analytics suitable for the general public. The Election Debate Visualisation project built Democratic Replay [38], a platform to replay videos of election debates enhanced with interactive hypervideo visualisations. The Centre for Argument Technology at the University of Dundee has recently partnered with the BBC to annotate a special programme of BBC Radio 4's The Moral Maze about the 60th anniversary of the UK's Abortion Act, creating visual analytics that were made available to the show listeners [39]. The visualisations work on six increments, and navigation controls allow users to see how they evolved throughout the show.

C. Visualisation Technologies

As mentioned in the introduction, several software frameworks have been proposed for visualising deliberation. Most visual discourse architectures however, focus on a limited subset of deliberation analytics [40], [41], and present argumentation as a map, hoping to increase the social deliberation skills of the participants (see e.g. [42]). As the brief outline in the introduction suggests, research has shown that such software tools indeed enhance deliberation and might also increase the propensity of reaching consensus.

The VisArgue project [43], [44] developed a fully-automated framework for analysing and visualising deliberative communication. It is currently the most coherent framework for deliberation analytics and their visual representation. Although it can only be applied to the debates ex-post, i.e. after a deliberative event, efforts are underway to include support for stream data. The framework automatically identifies the relative amount of deliberation along four dimensions: participation, respect, justification, and accommodation. For each of the dimensions, visualisations were developed to allow thorough visual inspection and interpretation of a debate's deliberative content (for further details, see e.g. [45], [46], [47], [48]). The architecture proposed in this paper and described in the next section builds on the VisArgue pipeline and the Argument Web providing fully-automated visual analytics of public deliberations in real time as the debates take place. The proposed visualisation framework is aligned on concepts for visualising text for close- and distant-reading, as surveyed by Jänicke et al. [49], as well as tasks for visual text analytics, as surveyed by Liu et al [50].

III. System Architecture

Figure 2 shows the pipeline for deliberation analytics, which combines state-of-the art speech recognition, argument mining and data analytics to produce stream visualisations of public deliberations, which are deployed in real time to participants via large analytic displays, hand-held devices or personal computers.

A. Stream Speech Processing

The Input module takes the audio signal of spoken deliberations and produces live transcripts, either using automatic speech-to-text services (e.g. Google's Cloud Speech-to-Text or Amazon's Alexa Voice Service) or a human stenographer. Recent advances in automatic speech recognition allow to distinguish different speakers and work extremely well in many languages, even with strong accents, local dialects, and background noise. This makes them suitable for research system prototypes and demonstrations. For deployment in real public deliberations, the required increase in transcript quality and speaker identification is guaranteed by using a stenographer. In both cases, the result will be a stream of utterances, with a timestamp, the speaker ID, and the locution.

B. Stream Argument Technologies

The stream of text is processed by the Argument Mining Framework (AMF) [51], [52] producing incremental argument structures in the Argument Interchange Framework [28], [53]. Although existing argument technologies (see Section II-B) were conceived to work with static, complete event-related argumentat data, it is straightforward to use them for processing a stream of data. By building a series of increasingly larger argument datasets as new transcribed speech arrives, the entire ecosystem of tools can be used on the successive corpora. Previous experiences with such an approach include the real-time manual annotation of a live radio show using the Argument Analysis Wall [54]. A group of 10 researchers took a stenographic feed of the broadcast, chunked it, and analysed it collaboratively using a very large touchscreen. The entire radio show was analysed by the end of the transmission, and incremental analytics were available throughout the show as the analysis progressed. The approach in the pipeline for deliberation analytics is analogous, except that the team of researchers is replaced by state-of-the-art automatic argument mining algorithms.

C. Stream Argument Analytics and Visualisations

The series of argument datasets are then incrementally processed by a combination of analytics (Analytics Service): special-purpose variations of developments from the Argument Analytics [36], [36] and VisArgue [44] projects. A subset of these will be described in detail in the next section.

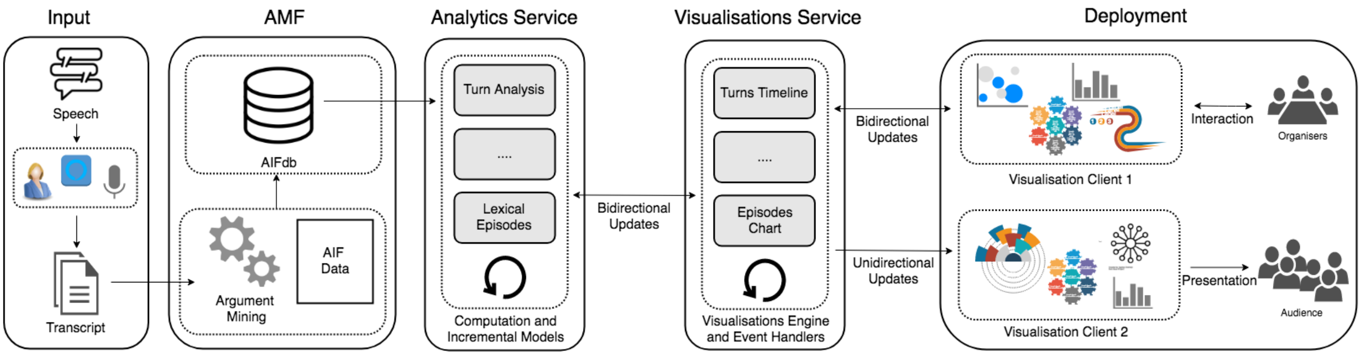


Fig. 2: Deliberation Analytics Pipeline: Input takes a speech signal and turns it into a transcript utterance-by-utterance; these utterances (timestamp, speaker ID and locution) are processed by AMF, the Argument Mining Framework into incremental argument structures in the argument Interchange Format (AIF); the Analytics and Visualisations Services work in tandem computing properties on the stream of argument data and producing dynamic visualisations for Deployment to deliberation participants via diverse visualisation clients.

The information stream resulting from the analytics is then turned into dynamic visualisations (Visualisations Service), tailored to the deliberation stakeholders, which update as new data becomes available. Two aspects of the visualisations are crucially relevant: their behaviour throughout the event, i.e. how they start, develop and end as data from the analytics services flows in; and their clarity in conveying information to untrained users. Examples of visualisations intended for the general public are available in [39], and for experts in [55].

Finally, the visualisations are presented live to participants on a large screen, or made available interactively to organisers and moderators via personal computing devices (Deployment). Individual visualisations are combined and arranged on canvases depending on the specific client: a few simpler, non-interactive visualisations are presented to the audience, for instance; while a larger set of more detailed, customisable options can be made available to moderators and organisers. The four visualisations presented below are designed to be presented to all users, but this is likely to change as more visualisations are added to the system. In general, however, visualisations that require user interaction, e.g. to expose specific parts of the information, will necessarily be shown on handheld devices. The Named Entity Graph and the Argumentation Map are examples of such visualisations. Similarly, visualisations with potential for misinterpretation or that require extensive explanations would only be shown to organisers and moderators.

IV. Deliberation Analytics

The VisArgue project (introduced above) developed a number of visualisation approaches to analyse discourse data [56]. For example, to analyse the contributions of speakers in a debate and their dynamics, the system uses *ConToVi* [46], a radial visualisation showing the so-called Topic-Space of a debate. Using staged animations, the flow of the discussion can be retraced, highlighting speaker-behaviour and topic-shifts. Similarly, the visual argument analytics in the EDV and ARG-tech BBC projects mentioned above [38], [39] provide tools to navigate and understand election debates non-linearly around sense-making features and to inspect argumentation properties of controversial deliberations in radio shows.

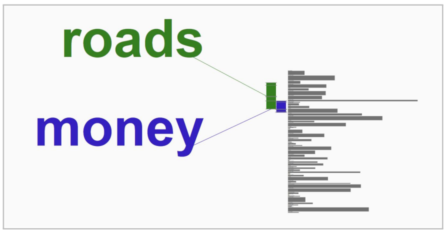
These visualisations are tailored to enable different perspectives on the data, building a large portfolio for analysing deliberative communication. All of them, however, were developed targeting a scenario in which users review past debates to get deeper insights into what has been said, enabling a detailed and focused analysis.

In the scenario of real-time face-to-face discussion support, there is a need for simplification to reduce the cognitive load on users, enabling them to focus on the essential. We propose a layered analysis concept allowing for the reduction of complexity: the faster the stream of new utterances, the more visualisations are simplified to ensure that users can focus on the most important information. If the stream slows down, layers can be added back on, allowing for more in-depth analysis. Users could, for example, view details of visualisations in individual "sandboxes", e.g., on mobile devices. These sandboxes could elaborate on specific aspects of the debate, and could be configured individually and independently for each user.

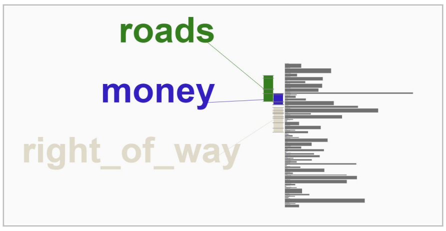
A. Content Analytics

The contents participants contribute to the debate are directly relevant to the overall quality of a public deliberation. We focus here on two types of content analysis: topics, i.e. *what* participants talk about, and argumentation, i.e. *how* participants talk about those topics.

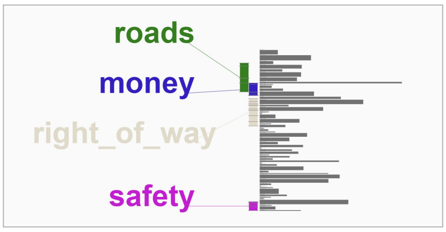
1) Topics: Lexical episodes plots are used to show which concepts are important during the deliberation. They are intended to give an overview of those concepts and highlight in which parts of the discussion they played an important role. The plots incrementally visualise analysed utterances as grey bars, mimicking the formatting of the underlying text as it would be seen when zooming out far enough. The "episodes" are defined as sequences in the text where a given keyword appears with a density higher than expected. These episodes are coloured for easier differentiation and shown on the left-hand side. The expected density of a keyword is calculated as the number of



(a) 52 utterances



(b) 53 utterances



(c) 54 utterances

Fig. 3: Lexical Episodes Plot: individual utterances can lead to the creation of new concepts if they mention a significant keyword.

times the keyword occurs in a word sequence, divided by the total number of words in the sequence. For a keyword that appears four times in a text with 100 keywords, we would expect the repetitions of the keyword to appear with gaps of 24 other words, assuming an equidistant density distribution. If the gap between two occurrences of a keyword in a group of at least three occurrences is smaller than its expected distance, a lexical episode is created. This episode then spans the region of the text from the first occurrence in the group to the last.

Figure 3 shows the development of the episodes over time as new utterances are processed. Depending on their topics and any mentioned keywords, new episodes can be created with every new utterance. Generally, adding new utterances can lead to an increase or decrease in the expected density for each keyword, in turn leading to the creation or removal of episodes. Figure 4 shows the visualisation by the end of the deliberation. For example, the episode "right of way" that was added in Figure 3b, was later removed as other, more significant episodes were detected after processing further utterances. The significance level required for an episode to be shown is automatically adjusted over time to avoid cluttering. Less significant episodes are automatically removed to make room for newer, more relevant ones. As a consequence, users might not see the same past episodes when looking at snapshots of the visualisation at different times in the deliberation. As episodes are created

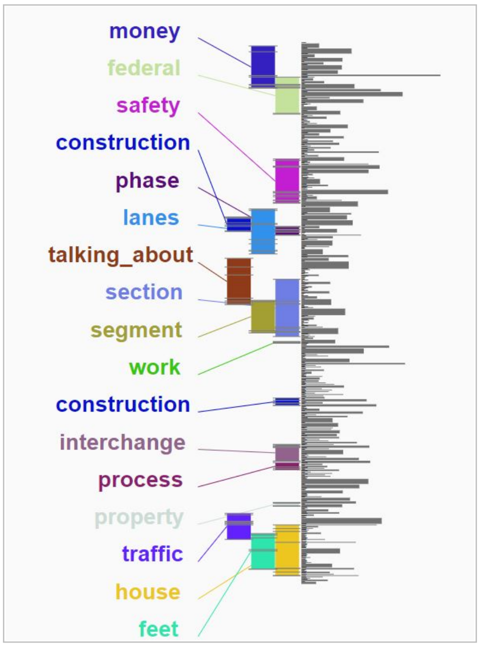


Fig. 4: Lexical Episodes Plot at the end of the discussion.

every time words are used more frequently than expected, users can conclude that removed episode keywords have become less significant as the discussion moved forward. The overview of the discussion provided in Figure 4 shows that participants were concerned about road safety, the number of lanes, the location of interchanges, the traffic, their houses, etc.

2) Argumentation: During deliberations, participants will speak in support or against issues related to the main question(s) that motivate the meeting. These relations between participants' utterances are captured by the analysis of the argument structure described above. By counting the number of propositions in favour and against a given position, it is possible to establish which issues are most divisive. These structures are presented as a collapsible tree, showing the issues in connection to the deliberation question to which they relate. As new arguments and issues are identified, older ones are collapsed, to make room and allow focus on the most recent information. This visualisation is aimed at making explicit what are the main issues in the discussion, how the different positions relate to these issues and to each other, and which of the participants hold those positions. The expectation is that this information will allow participants, and especially the moderator, to better manage the deliberation.

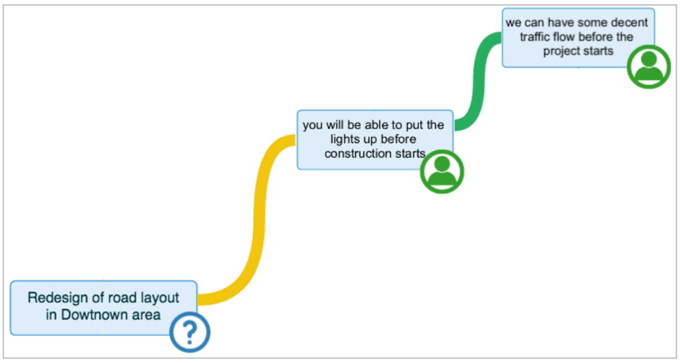
Figure 5 shows an example of how the Dynamic Argument Map evolves during the deliberation. The root node represents the question motivating the public deliberation. One or more questions are generally available before the event, published by the organisers to motivate attendance and structure the discussion. The links between issue and questions are made by automatic analysis of topic similarity. Issues are connected to questions via yellow lines and are decorated with an avatar representing the participant who put it forward. Similarly, reasons for and against issues and other propositions are also decorated by the speaker, but connected with green and red lines, respectively. The thickness and opacity of the lines are used to indicate the path between the root of the tree and the most recent addition to the map. Older branches, such as when there is a change of issue, are collapsed to avoid cluttering and ease focusing attention to the most relevant section of the map. This is illustrated in Figure 5c, where the top branch constructed in Figures 5a and 5b is partially collapsed showing two levels with thinner lines and 50% opacity. The thicker lines leading to the bottom right proposition indicate that this is the most current, and therefore relevant, addition to the structure.

B. Participation Analytics

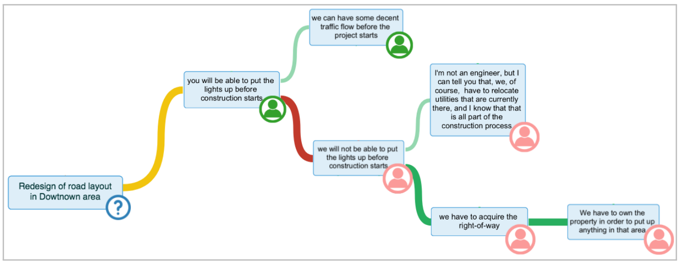
Opportunities for participants to contribute to the debate are also directly relevant to the quality of public deliberations. Below, we describe two types of participation analysis: entities, i.e. *who* takes part, and diachrony, i.e. *when* and *how much* participants contribute throughout the debate.

1) Entities: To analyse the participation of speakers in a debate and explore their influence, we developed a named entity relationship graph. The resulting visualisation gives an overview of relevant named entities, which are extracted using a combination of supervised and unsupervised machine learning algorithms. These entities are put in relation with the participants of the debate, showing which themes each participant has talked about. We use the Stanford Named Entity Recognition system [57] to extract entities in seven categories: Location, Person, Organisation, Money, Percent, Date, and Time. We also use information provided by topic models and lexical chaining algorithms to extract important context-related keywords; and heuristics such as word lists, to detect date and time keywords or those expressing positive or negative emotion, among others. Relations between the extracted entities are identified using a distance-restricted model to find entity pairs. We consider all entities that appear within a distance of at most five words from each other to be in a potential relationship, and create an entity pair. This definition ensures fast and efficient computation of entity pairs.

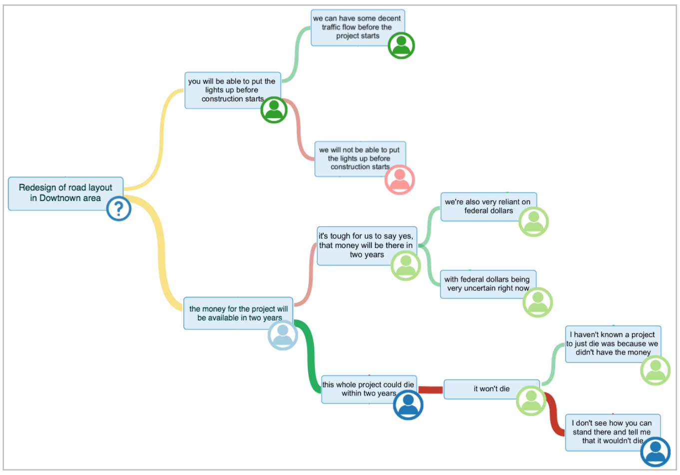
The extracted entities are shown in a directed node-link graph and positioned using a force-directed layout algorithm. The nodes are scaled corresponding to the frequency of the respective entity in the text. Their colour is determined based on the membership of the entities in ten categories like *person*, *location*, *emotion words* or *context words*. Entities that appear in pairs are connected with an edge. The edge length is scaled to the average distance between the entities in the text, bringing entities closer in the text closer together in the visualisation. The brightness and width of the edges is scaled proportionally to the frequency of the entity pair. To select the degree of detail for the visualisation, users can adjust the minimum entity-pair frequency that is still shown. As new utterances are added in the data stream, extracted named entities are added to the entity graph as is shown in Figure 6. Small cards represent the current speakers in the discussion and show their positions within the entity space. We again use a force-directed layout algorithm to position the speaker cards. Each card is attracted



(a) 2 utterances



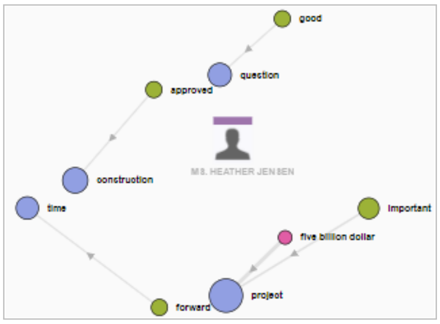
(b) 6 utterances



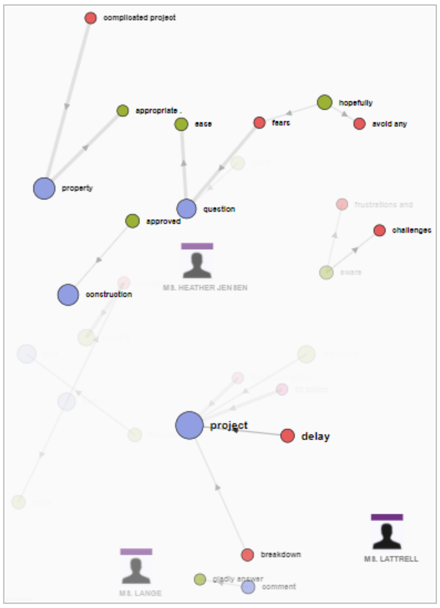
(c) 14 utterances

Fig. 5: Dynamic Argument Map: yellow lines connect divisive issues with deliberation questions, while green and red show support and attack relations; older branches of the map collapse as new ones are added by processing the most recent utterances in the stream.

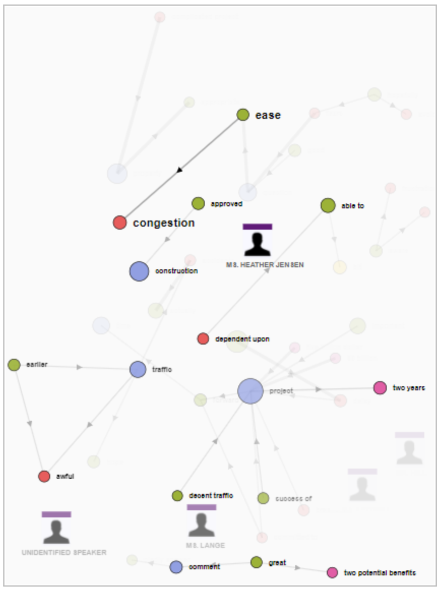
by the entities the speaker has talked about. All attraction forces are proportional to the amount that a participant has talked about an entity. Consequently, participants talking about many entities move towards the centre of the graph. To avoid overplotting issues, an entity or speaker card can be anchored to a given position. This causes the remaining entities to re-orientate around the fixed point based on the original attraction forces. Throughout the debate, we fade out both entities and speakers that have not been updated in a while, enabling the users to focus on who is participating right now, and in which



(a) 5 utterances



(b) 15 utterances



(c) 50 utterances

Fig. 6: Named Entity Graph: As the discussion progresses, new entities and speakers are added to the graph and older ones are faded out to avoid cluttering. The speakers' position in the entity graph shows what they have been talking about.

context. The three snapshots shown in Figure 6 show that a budget of five billion US dollars is available and that the project has been approved, that *Ms. Lande* and *Mr. Lattrell* voiced concerns about potential delays, and that *Ms. Jensen* elaborated on potential improvements to traffic congestion.

2) Diachrony: The evolution of when and to what extent participants contributed to the deliberation over time is computed from the argument data stream by noting the timestamp and identity associated with each locution. It is worth noting that the information is not based simply on speech actions, but on actual contributions to the argument structure, i.e. only utterances that have actual argumentative force are taken into consideration. This prevents the inclusion of turn-taking and session management contributions (e.g. a moderator explaining the dynamics of the event) from skewing the results. In addition to the timeline of contributions, we also aggregate participation as a pie chart, to show explicitly how much each participant contributed so far to the entire discussion. This explicit information is intended to help to moderate the discussion, e.g. by avoiding imbalances in how time is distributed across participants, presumably leading to fairer, better deliberation.

Figure 7 shows the Turns Timeline and Participation Chart. These are colour-coded, in the same way as the avatars in the Dynamic Argument Map. The timeline presents one line per speaker, and new speakers and contributions are added as new argument data flows in. To help users focus on most recent section of the timeline, we use a sliding window which shows the latest 10 minutes for the discussion. Data from earlier in the discussion are placed in the bottom, zoomed out visualisation which shows every turn since the beginning of the event. In the interactive version, users can scroll back to visualise the details of earlier portions of the debate. The participation chart is simple, yet powerful, allowing all participants to see who has contributed to the discussion so far and whether anyone has dominated the discussion.

V. Conclusions

In this paper, we presented the first steps towards developing a system to augment public deliberations with analytics and visualisations, which work in real time on stream argument data and are suitable for non-expert users. Building on theoretical criteria from the political sciences, we identified an initial set of visualisations which we expect will contribute towards improving the quality of deliberative communication in such events. These visualisations and the underlying technologies were adapted to work on stream, incremental data and pipelined to provide an end-to-end system which turns a speech signal into a collection of visual interventions.

We have presented challenges for stream analytics and visualisations in the introduction, and have addressed these in the design of our visualisation components. Yet, one additional challenge remains. As new data are added, users should be guided to the regions of the visualisations that are changing. This would reduce their cognitive load and allow users to focus on the relevant information, instead of having to identify what has changed between updates before they can make sense of it.

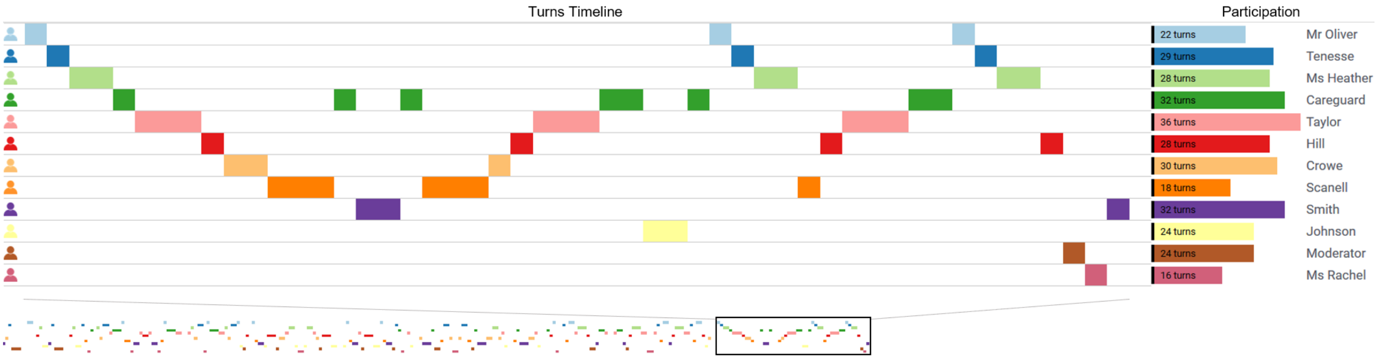


Fig. 7: Turns Timeline and Participation Chart.

The work reported here continues a series of efforts that led to the development of argument technologies [27], argument analytics [36], [37], and interactive visualisations for experts [43] and the general public [38]. Challenges as we move forward include participatory design workshops in which citizens help revising our visualisations towards simpler, yet equally effective versions that we can present to audiences without extensive explanation. We will also work towards developing the current prototype implementation of the pipeline into a fully-fledged, deployable framework. This framework will then be tested in A/B laboratory studies to identify and extract differences between the process of deliberations with and without the visual support system. For instance, we expect to see an increase in deliberation quality when participants in a debate are provided with a visual augmentation of the debate. In addition, each of the visualisations in the system will be evaluated in isolation by a panel of intended final users (e.g. audience members, organisers, etc.), in order to determine to what extent they succeed in effectively augmenting deliberations. This will be done by a combination of user experience and task-specific questionnaires [58]. Finally, we will apply the visual interventions *in the wild* in actual public deliberations in Scotland and Germany, in order to assess their impact on deliberative democracy, and the suitability of the visualisations to be catered to the masses.

Acknowledgments

We would like to acknowledge that the work reported in this paper has been supported in part by VW Foundation (VolkswagenStiftung) under grant 92 182, in part by the DFG-FOR2111 Research Unit at the University of Konstanz, and in part by the DFG under grant 376714276 within the SPP-1999.

References

[1] Pippa Norris. *Democratic Deficit: Critical Citizens Revisited*. Cambridge University Press, Cambridge, MA, 2011.

[2] Roberto Stefan Foa and Yascha Mounk. The Danger of Deconsolidation: The Democratic Disconnect. *Journal of Democracy*, 27(3):5–17, 2016.

[3] Gerry Stoker. *Why politics matters: Making democracy work*. Palave, London, 2nd edition, 2017.

[4] Stadtverwaltung Potsdam. Kosten für Bürgerbeteiligung / Bürgerhaushalt 2013-2015. https://fragdenstaat.de/a/8883, 2015.

[5] The Guardian. Stuttgart 21 is a failure of deliberative democracy. https://www.theguardian.com/commentisfree/2010/oct/24/stuttgart-21-failure-deliberative-democracy, 2010.

[6] Involve. The True Costs of Public Participation. https://www.involve.org.uk/sites/default/files/uploads/True-Costs-Full-Report2.pdf, 2005.

[7] Mark Klein. A Critical Review of Crowd-Scale Online Deliberation Technologies. http://ssrn.com/abstract=2652888, 2015.

[8] Kyoo-Lak Cho and David H. Jonassen. The effects of argumentation scaffolds on argumentation and problem solving. *Educational Technology Research and Development*, 50(3):5, Sep 2002.

[9] Daniel D Suthers. Towards a systematic study of representational guidance for collaborative learing discourse. *Journal of Universal Computer Science*, 7(3):254–277, 2001.

[10] Tanja Engelmann and Friedrich W. Hesse. How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *International Journal of Computer-Supported Collaborative Learning*, 5(3):299–319, 2010.

[11] Nikos Karacapilidis and Dimitris Papadias. Computer supported argumentation and collaborative decision making: the HERMES system. *Information Systems*, 26(4):259 – 277, 2001.

[12] Robert C Luskin, James S Fishkin, and Shanto Iyengar. Considered opinions on US foreign policy: Face-to-face versus online deliberative polling. *International Communication Association, New Orleans, LA*, pages 1369–1387, 2004.

[13] Kimmo Grönlund, Staffan Himmelroos, and Staffan Himmelroos. The challenge of deliberative democracy online–A comparison of face-to-face and virtual experiments in citizen deliberation. *Information Polity*, 14(3):187–201, 2009.

[14] Seong-Jae Min. Online vs. Face-to-Face Deliberation: Effects on Civic Engagement. *Journal of Computer-Mediated Communication*, 12(4):1369–1387, 2007.

[15] Cindy Boyles Tucey. Online vs. Face-to-Face Deliberation on the Global Warming and Stem Cell Issues. Western Political Science Association 2010 Annual Meeting Paper, 2010.

[16] Peter Muhlberger. Attitude change in face-to-face and online political deliberation: Conformity, information, or perspective taking? American Political Science Association Annual Meeting, Washington, DC, 2005.

[17] Jürgen Habermas. *The Theory of Communicative Action*. Beacon, Boston, MA, 1984.

[18] John Gastil and Laura W. Black. Public Deliberation as the Organizing Principle of Political Communication Research. *Journal of Public Deliberation*, 4(1):1–47, 2008.

[19] Simone Chambers. Deliberative democracy theory. *Annual Review of Political Science*, 6(1):307–326, 2003.

[20] Peter B. Edwards, Richard Hindmarsh, Holly Mercer, Megham Bond, and Angela Rowland. A Three-Stage Evaluation of a Deliberative Event on Climate Change and Transformation Energy. *Journal of Public Deliberation*, 4(1, Article 6), 2008.

[21] Marco R. Steenbergen, André Bächtiger, Markus Spörndli, and Jürg Steiner. Measuring political deliberation: A discourse quality index. *Comparative European Politics*, 1(1):21–48, 2003.

[22] Christopher Lord and Dionysia Tamvaki. The politics of justification? Applying the 'Discourse Quality Index' to the study of the European Parliament. *European Political Science Review*, 5:27–54, 3 2013.

[23] Léa Roger and Gary Schaal. The Quality of Deliberation in Two Committees of the European Parliament: The Neglected Influence of the

Situational Context and the Policymaking Stage. *Politics and Governance*, 1(2), 2013.

[24] BBC News. The computers being trained to beat you in an argument. https://www.bbc.co.uk/news/technology-41010848, last accessed 2018-07-13, 2017.

[25] The Guardian. Man 1, machine 1: landmark debate between AI and humans ends in draw. https://www.theguardian.com/technology/2018/jun/18/artificial-intelligence-ibm-debate-project-debater, last accessed 2018-07-13, 2018.

[26] The New York Times. IBM Unveils System That Debates With Humans. https://www.theguardian.com/technology/2018/jun/18/artificial-intelligence-ibm-debate-project-debater, last accessed 2018-07-13, 2018.

[27] Chris Reed, Katarzyna Budzynska, Rory Duthie, Mathilde Janier, Barbara Konat, John Lawrence, Alison Pease, and Mark Snaith. The argument web: An online ecosystem of tools, systems and services for argumentation. *Philosophy & Technology*, 30(2):137–160, 2017.

[28] C. Chesnevar, J. McGinnis, S. Modgil, I. Rahwan, C. Reed, G. Simari, M. South, G. Vreeswijk, and S. Willmott. Towards an Argument Interchange Format. *Knowledge Engineering Review*, 21(4):293–316, 2006.

[29] John Lawrence, Floris Bex, Chris Reed, and Mark Snaith. AIFdb: Infrastructure for the argument web. In *Proceedings of the Fourth International Conference on Computational Models of Argument (COMMA 2012)*, pages 515–516, 2012.

[30] John Lawrence and Chris Reed. AIFdb corpora. In S. Parsons, N. Oren, C. Reed, and F. Cerutti, editors, *Proceedings of the Fifth International Conference on Computational Models of Argument (COMMA 2014)*, pages 465–466, Pitlochry, 2014. IOS Press.

[31] Chris Reed, Simon Wells, Joseph Devereux, and Glenn Rowe. Aif+: Dialogue in the argument interchange format. *Frontiers in Artificial Intelligence and Applications*, 172:311, 2008.

[32] Katarzyna Budzynska, Mathilde Janier, Chris Reed, and Patrick Saint-Dizier. Theoretical foundations for illocutionary structure parsing. *Argument and Computation*, 2016.

[33] Raquel M. Palau and Marie-Francine Moens. Argumentation mining: the detection, classification and structure of arguments in text. In *Proceedings of the 12th international conference on artificial intelligence and law*, pages 98–107. ACM, 2009.

[34] Katarzyna Budzynska, Mathilde Janier, Juyeon Kang, Chris Reed, Patrick Saint-Dizier, Manfred Stede, and Olena Yaskorska. Towards argument mining from dialogue. In *Proceedings of the Fifth International Conference on Computational Models of Argument (COMMA 2014)*, pages 185–196. IOS Press, 2014.

[35] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning and logic proamming. In *IJCAI*, volume 93, pages 852–857, 1993.

[36] John Lawrence, Rory Duthie, Katarzyna Budzysnka, and Chris Reed. Argument analytics. In P. Baroni, M. Stede, and T. Gordon, editors, *Proceedings of the Sixth International Conference on Computational Models of Argument (COMMA 2016)*, Berlin, 2016. IOS Press.

[37] John Lawrence, Mark Snaith, Barbara Konat, Katarzyna Budzynska, and Chris Reed. Debating technology for dialogical argument: Sensemaking, engagement, and analytics. *ACM Transactions on Internet Technology*, 17(3):24:1–24:23, 2017.

[38] Brian Plüss and Anna De Liddo. Democratic replay: Enhancing tv election debates with interactive visualisations. In *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018.

[39] ARG-tech. The Morality of Abortion. http://bbc.arg.tech/, last accessed 2018-07-13, 2017.

[40] Mark Klein and Luca Iandoli. Supporting Collaborative Deliberation Using a Large-Scale Argumentation System: The MIT Collaboratorium. In *Directions and Implications of Advanced Computing: Conference on Online Deliberation*, University of California, Berkeley, 2008.

[41] Tom Murray, Leah Wing, Beverly Park Woolf, Alexander Wise, Shijun Wu, Lori Clark, Lee Osterweil, and Xiaoxi Xu. A prototype facilitators dashboard: Assessing and visualizing dialogue quality in online deliberations for education and work. In *Proceedings of The 2013 International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE-2013)*, Las Vegas, 2013.

[42] Tom Murray, Lynn Stephens, Beverly Park Woolf, Leah Wing, Xiaoxi Xu, and Natasha Shrikant. Supporting Social Deliberative Skills Online: The Effects of Reflective Scaffolding Tools. In A. Ant Ozok and Panayiotis Zaphiris, editors, *Online Communities and Social Computing: 5th International conference, OCSC 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013. Proceedings*, pages 313–322. Springer, Berlin, 2013.

[43] Mennatallah El-Assady, Valentin Gold, Annette Hautli-Janisz, Wolfgang Jentner, Miriam Butt, Katharina Holzinger, and Daniel Keim. VisArgue - A Visual Text Analytics Framework for the Study of Deliberative Communication. In Daniela Širinić, Jan Šnajder, Zoltán Fazekas, and Shaun Bevan, editors, *Proceedings of The International Conference on the Advances in Computational Analysis of Political Text (PolText2016)*, pages 31–36. University of Zaeb, Zaeb, 2016.

[44] Valentin Gold, Mennatallah El-Assady, Annette Hautli-Janisz, Tina Bögel, Christian Rohrdantz, Miriam Butt, Katharina Holzinger, and Daniel Keim. Visual linguistic analysis of political discussions: Measuring deliberative quality. *Digital Scholarship in the Humanities*, 32(1):141–158, 2017.

[45] Valentin Gold, Christian Rohrdantz, and Mennatallah El-Assady. Exploratory Text Analysis using Lexical Episode Plots. In E. Bertini, J. Kennedy, and E. Puppo, editors, *Euroaphics Conference on Visualization (EuroVis) - Short Papers*. The Euroaphics Association, 2015.

[46] Mennatallah El-Assady, Valentin Gold, Carmela Acevedo, Christopher Collins, and Daniel Keim. Contovi: Multi-party conversation exploration using topic-space views. In *Computer aphics Forum*, volume 35, pages 431–440. Wiley Online Library, 2016.

[47] Annette Hautli-Janisz and Mennatallah El-Assady. Rhetorical strategies in German argumentative dialogs. *Argument & Computation*, 8(2):153–174, 2017.

[48] Mennatallah El-Assady, Rita Sevastjanova, Bela Gipp, Daniel A. Keim, and Christopher Collins. NEREx: Named-Entity Relationship Exploration in Multi-Party Conversations. *Computer aphics Forum*, 36(3):213–225, 2017.

[49] Stefan Jänicke, eta Franzini, Muhammad Faisal Cheema, and Gerik Scheuermann. On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In R. Borgo, F. Ganovelli, and I. Viola, editors, *Euroaphics Conference on Visualization (EuroVis) - STARs*. The Euroaphics Association, 2015.

[50] Shixia Liu, Xiting Wang, Christopher Collins, Wenwen Dou, Fangxin Ouyang, Mennatallah El-Assady, Liu Jiang, and Daniel A. Keim. Bridging Text Visualization and Mining: A Task-Driven Survey. *IEEE Transactions on Visualization and Computer aphics*, pages 1–1, 2018.

[51] John Lawrence and Chris Reed. Argument mining using argumentation scheme structures. In P. Baroni, M. Stede, and T. Gordon, editors, *Proceedings of the Sixth International Conference on Computational Models of Argument (COMMA 2016)*, Berlin, 2016. IOS Press.

[52] Katarzyna Budzynska, Mathilde Janier, Juyeon Kang, Chris Reed, Patrick Saint-Dizier, Manfred Stede, and Olena Yaskorska. Towards argument mining from dialogue. In *Proceedings of the Fifth International Conference on Computational Models of Argument (COMMA 2014)*, pages 185–196. IOS Press, 2014.

[53] Zablith F. Rahwan, I. and C. Reed. Laying the Foundations for a World Wide Argument Web. *Artificial Intelligence*, 171:897 – 921, 2007.

[54] ARG-tech. Argument Analysis Wall. http://www.arg-tech.org/index.php/projects/argument-analysis-wall/, last accessed 2018-07-13, 2012.

[55] VisArgue project. Visual Analysis and Exploration of the 2016 US Presidential Debates. http://presidential-debates.dbvis.de, last accessed 2018-07-13, 2016.

[56] Mennatallah El-Assady, Annette Hautli-Janisz, Valentin Gold, Miriam Butt, Katharina Holzinger, and Daniel Keim. Interactive visual analysis of transcribed multi-party discourse. In *the 55th Annual Meeting of the Association for Computational Linguistics*, pages 49–54, 2017.

[57] Jenny Rose Finkel, Trond enager, and Christopher Manning. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics*, pages 363–370. ACL, 2005.

[58] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. Applying the user experience questionnaire (ueq) in different evaluation scenarios. In Aaron Marcus, editor, *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience*, pages 383–392, Cham, 2014. Springer International Publishing.