Deb8: A Tool for Collaborative Analysis of Video

Guilherme Carneiro
SACHI, University of St Andrews
St Andrews, Fife, UK
gsmc@st-andrews.ac.uk

Miguel A. Nacenta
SACHI, University of St Andrews
St Andrews, Fife, UK
mans@st-andrews.ac.uk

Alice Toniolo
SACHI, University of St Andrews
St Andrews, Fife, UK
a.toniolo@st-andrews.ac.uk

Gonzalo Gabriel Méndez
Escuela Superior Politécnica del Litoral, Guayaquil, Ecuador, and University of Calgary
Calgary, Canada
gonzalo.mendez@ucalgary.ca

Aaron J. Quigley
SACHI, University of St Andrews
St Andrews, Fife, UK, and
Keio-NUS CUTE Center,
National University of Singapore
aquigley@st-andrews.ac.uk

ABSTRACT

Public, parliamentary and television debates are commonplace in modern democracies. However, developing an understanding and communicating with others is often limited to passive viewing or, at best, textual discussion on social media. To address this, we present the design and implementation of Deb8, a tool that allows collaborative analysis of video-based TV debates. The tool provides a novel UI designed to enable and capture rich synchronous collaborative discussion of videos based on argumentation graphs that link quotes of the video, opinions, questions, and external evidence. Deb8 supports the creation of rich idea structures based on argumentation theory as well as collaborative tagging of the relevance, support and trustworthiness of the different elements. We report an evaluation of the tool design and a reflection on the challenges involved.

CCS CONCEPTS

• **Human-centered computing → Synchronous editors**.

KEYWORDS

Video analysis, evidence linking, collaborative debate analysis, argumentation

TVX '19, June 5–7, 2019, Salford (Manchester), United Kingdom

ACM ISBN 978-1-4503-6017-3/19/06…$15.00
https://doi.org/10.1145/3317697.3323358

ACM Reference Format:
Guilherme Carneiro, Miguel A. Nacenta, Alice Toniolo, Gonzalo Gabriel Méndez, and Aaron J. Quigley. 2019. Deb8: A Tool for Collaborative Analysis of Video. In *ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19), June 5–7, 2019, Salford (Manchester), United Kingdom.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3317697.3323358

1 INTRODUCTION

Broadcast debates are common in modern democracies as a powerful forum to help shape the public's image of those debating and their arguments. Multiple speakers can present arguments and counter-arguments, opposing views and connected arguments each with varying degrees of relevance and trustworthiness. For example, in 2016, over a quarter of the population of the USA watched the final candidate debate in the presidential election while millions more watched live streams and recordings of the debate [36].

However, as debates involve people making arguments, they might be flawed. Untruths, lies or superficial statements litter debates. Simple statements of fact can be open to interpretation and hence support divergent viewpoints in the same debate. For example, "the unemployment rate is at a 10 year low" might support the argument "the economy is doing well" but if people have stopped looking for work then the state of the economy might actually be poor.

During a debate, statements can be fact-checked but if the arguments are based on opinions or the statements require more nuance, then it can be difficult to definitively label something as untrue. Moreover, debates are complex multifaceted events which make it difficult to form an overall picture to make a decision about the opposing arguments.

In the face of this, audiences and the general public have turned to a range of technologies to support discussion (e.g., social media). However, many systems are not adequate for discussing a debate with many arguments, branching to sub-arguments, a range of degrees of relevance in the evidence presented, or requiring opinion deconstruction. The effort to discover and weave together relevant information, on even a single debate, might be too high for a single person.

To overcome the problems inherent in discussing or "debating a debate", and the weaknesses in the existing approaches we propose *Deb8*, a visual language underpinned by a syntactic graph structure [8]. Deb8 is an online system which allows multiple people to collaboratively analyze videos of recorded debates. It supports a structured approach to the collection and linking of information within and around different parts of the video. Deb8 uses an ontology derived from the argumentation literature, and allows data to be linked from the Internet. Starting with quotes or snapshots from the video, Deb8 supports the creation of chains of questions, opinions and evidence where the links among these elements can be weighted by each person based on the degree to which they feel it supports or is relevant. Deb8 supports temporal filtering of argument elements along with the ability to only view those with particular degrees of support and relevance.

We make the following contributions: 1) the design and implementation of a collaborative platform for the group-based analysis of debate videos; 2) a set of argument theory driven features for supporting quotes, questions, opinions, relevance and support in evidence-based analysis, and; 3) a preliminary study of three groups of users suggesting that Deb8 can address existing challenges of video-based debates.

2 EXAMPLE SCENARIO

A group of journalists have been tasked by the editor of their newspaper with analyzing a televised debate that took place recently on national TV. The newspaper, which prides itself on its balanced views and respect for facts, wants to offer analyses of each candidate's discourse. The analysis should weigh supporting and contradicting views for arguments and statements. Journalists will comment on quotes by the candidates, find evidence that corroborates or contradicts statements and collectively evaluate their relevance. Once the analysis is completed, they want to reach some kind of team consensus regarding the strength of the different candidates and how trustworthy their proposals are.

Later, the editors decide that they want to open up the analysis to the general public. Interested readers will have different sources of evidence (missed by the journalists, or from non-reputable sources) and will have different points of view regarding what is relevant, important, and believable. Note that, at this time, we are not considering scenarios involving real-time analysis of debates.

3 RELATED WORK

The growth of online discussion fora has given rise to numerous ways in which existing systems can be used to support group discourse around a debate. Examples include collaborative web-based learning environments (e.g. Blackboard, MOOC), newsgroups, question and answer systems (e.g. Quora), commenting systems (e.g. newspapers, blogs), dedicated discussion channels (e.g. IRC, Slack), online news and discussion communities (e.g. Reddit, Slashdot), and social networks (e.g. Facebook). Such systems are often largely textual with simple reply mechanisms, limited threading for discussions and little support for the identification of arguments and counter-arguments or rebuttals, topics, concepts, emergent questions, evidence, or for providing structure in how these aspects interrelate in the debate.

The simplicity of text-based systems has allowed them to be used in flexible ways supporting a breadth of discourse types ranging from discussion around a debate event, to an ongoing discussion around a controversial social topic [10].

While general purpose systems can support many forms of discussion, the need to support more structured discourse, such as in formal debates, has given rise to systems that work around opposing arguments, often ending with a vote. These afford creating a topic or question (e.g. ConvinceMe.net, Debate.org, EDeb8.com, DebateIsland.com, debatewise.org or artikulate.in), adding polls (e.g. Debate.org), opinions (e.g. Debate.org), arguments for and against a position (e.g. ConvinceMe.net, quibl.com, createdebate.com and debatewise.org) and voting (e.g. Debat [21], Debate.org, EDeb8.com, ConvinceMe.net, netivist.org, quibl.com, createdebate.com, or debatewise.org). Alternatively, dedicated strands within an established system, such as "change my view" in Reddit.com [22], take an existing platform and overlay new rules to afford new forms of interaction for debate.

Brainstorming and Crowds

Group discussion in a debate can be seen as a form of brainstorming, while the involvement of ever larger groups suggests that the "wisdom of the crowd" may allow for richer discussions and hence agreement, or at least evidenced disagreement. Brainstorming, as an identified concept, dates back over half a century [27], while the notion of the "wisdom of the crowd" is at least a century old [9]. Both concepts have had detractors, yet the development of digital platforms relying on both concepts are now commonplace. Consensus building or the formation of opposing views through brainstorming are common uses of online discussion tools. From Facebook to email, we can see how such systems allow for spontaneous group discussion, harnessing the wisdom of a group to produce new ideas or solve old problems.

Manifestations of such brainstorming systems for collaborative analysis of information appears in numerous forms of related work. Platforms such as SlideShare or Prezi allow for limited markup and discussion of the content while

IdeaMâché [20] affords information composition with concept maps through presentation, discussion and ideation. Alternative approaches for collaborative sense-making and intelligence analysis offer a sense-making canvas to reduce the cognitive effort of analysts processing multi-source information, including The Sandbox [41] and others (e.g., [15, 39]). In addition, interactive visual languages, such as iVoLVER [26], can be employed to facilitate analysis of visual data.

However, many of these approaches are only one step beyond the basic textual discussion system. Systems are emerging which support more structured discourse by allowing graphical structuring of questions and answers and their significance (e.g., debatemap.live) or ideas (e.g. debategraph.org). Mind-maps, collaborative mind-maps [32] and systems grounded in the mindmapping paradigm [20] rely on the ability to create textual or graphical nodes, introduce child nodes, manipulate siblings, and grow the often tree-structured information space by interactively linking elements. Graphical data-flow languages from audio composition [6] to visualization [26], health data-flow [35] or for end-user programming (e.g. LabView), all relate to the visual linking of elements employed in Deb8. We refer the reader to [17] for a survey of dataflow languages and [8] for visual languages in general. The flexibility introduced by such graph components enable the representation of complex argument structures such as argument reinstatements, that is, arguments that support a claim by challenging its undermining arguments (i.e., [31]). This last group of systems are most closely related to our work but do not incorporate the key collaborative debate analysis aspects we introduce.

Much debating systems research focuses on understanding the content of the debate. Prior work has explored fact checking [12, 28], stance identification stance [1], highlight identification [37], argument analysis [19], sentiment analysis and segmentation [23], second-screen experiences [2, 11], live debate collective assessment (e.g. "the worm" [4]), real-time feedback [14], and debate visualisation [29]. Such research can enhance any debate-support system but does not address the inherent problems of collaborative analysis.

Alongside applications that focus on understanding debates, argument-mapping tools help analyze and structure arguments, including Rationale & bCisive (reasoninglab.com), Compendium [34], AGORA-net [13], and OVA+ [16]. These provide graph patterns to guide the construction of well-formed arguments underpinned by a specific *argument ontology* which indicates the meaning of nodes and relationships [3]. Tools such as bCisive and Compendium are based on the Issue-Based Information System (IBIS) representation [18] where *positions* respond to an *issue*, and are expanded with pro and con *arguments*. IBIS is used in many applications [34], including dialogue analysis [5] to build collaborative understanding of an issue. Other systems provide an ontology that focuses on identifying the structure of inferences and conflicts, for example OVA+, based on the Argument Interchange Format [3]. Deb8 builds on these ontologies by specifying elements necessary to represent meaningful structures for collaborative discussion and analysis of debates.

This review has identified a number of aspects of related work which overlap with Deb8. Some tools, such as Compendium [34] and AGORA-net [13], are collaborative while other tools, such as debatemap.live, debategraph.org, Rationale & bCisive (reasoninglab.com), Compendium [34], AGORA-net [13], and OVA+ [16], afford users a graphic canvas. In addition, Compendium [34] and debategraph.org allow users to link external evidence to arguments. By contrast, Deb8 is unique in that it combines these features and original ones (e.g., connecting arguments and external evidence to temporal media, the ability to watch video in parallel with a summary of its analysis) specifically for video-based debates. Further, Deb8 exemplifies the four stages of Collect, Relate, Create, and Donate [33] for sparking creativity in the identification of quotes, development of questions and opinions, provision of an evidence base while allowing for collaborative analysis and identification of relevance, support and belief in one's own, and others' points of view.

4 DESIGN GOALS, SCOPE AND PRINCIPLES

Our overarching goal is to create a new type of web-based media that enhances understanding and communication of people about video-based debates. More specifically, we designed Deb8 to: G1) support deep and close analysis of video debates; G2) facilitate direct linking with existing knowledge and opinions; G3) enable collaboration between people with different opinions, and; G4) allow people to manage complexity. Although we realize the importance of simplicity in the design of the interface and strive to make the UI as accessible as possible, we prioritized the goals above over the creation of an interface for "walk up and use" because we believe that deep analysis requires some training. In addition, we consider arguments in a broad sense and we do not intend to map argument structures, such as premises and conclusions.

Some of our goals above are partly shared with existing tools (e.g., Compendium [34], AGORA-net [13], or debategraph.org), but they all draw from our belief that deeper collaborative engagement with evidence and political ideas by a broader range of the population is a good way forward towards better functioning democracies.

In order to address our design goals, we established the following design principles as a guide: DP1) provide a small set of reusable elements that can be interconnected in rich ways; DP2) map interface elements to the constructs in argument theory to structure discussion; DP3) allow each user to judge relevance, valence and trustworthiness of discussion elements at a fine level of granularity; DP4) support multiple views and filtering of data and; DP5) provide flexible UI navigation and linking to support complex workflows. We strived to apply these principles during iterative development sprints in which we tested the prototype and our assumptions through design critique, sometimes exposing the system to colleagues. In the following sections, we describe the interface using goal and design principle codes above to describe how specific features address them.

5 DEB8

In this section, we use the scenario above to describe the Deb8 system and its underlying argument structure in relation with goals (G1-4) and design principles (DP1-4) of Section 4.

Interface Structure

Deb8 is conceived to run in a large landscape monitor or interactive surface, preferably high-resolution (UHDTV). Besides a thin horizontal bar at the top with the title and log-in buttons, the main structure divides the screen into three vertical panels: the Video and Caption Panel (Fig. 1.A), the Argument Canvas (Fig. 1.B), and the Web Browser (Fig. 1.C). The three panels are interlinked in multiple ways to support G2. We anticipate that the activities supported by the three panels (video watching, argument chain creation, and web browsing) take place in a highly interleaved way, hence all of them are present continuously on the screen (DP5). To accommodate varying importance of the activities at different points in time, the amount of screen real estate that each panel takes can be adjusted by dragging their boundaries.

Video and Caption Panel

The leftmost panel contains a video player to display the debate video, with the usual web-video facilities (play, stop, pause and a timeline). Underneath the video, a scrollable panel contains the video's captions. As the video plays, the corresponding caption is highlighted. Clicking on a specific caption moves the video to the corresponding time. This is an example of cross-element UI linking (DP5). Caption "text snippets" or video frame "snap-shots" are dragged from the caption panel into the central argument canvas to create textual or video quotes respectively, which can hence be manipulated and interconnected within the canvas.

Argument Canvas

The central and core panel of the application is an infinite zoomable argument canvas that allows the journalists to build a shared graph of argument chains (G2) by creating arbitrarily complex (G4) combinations of simple elements (DP1) derived from argument theory (DP2). This model and interface are inspired by existing interfaces used for brainstorming [20], argument mapping [16, 34, 39], and visual programming [6, 26]. The types of elements that can be connected are described below, and were chosen to represent the smallest atomic argument components to analyze debates (DP1) and to connect to each other in as many meaningful ways as possible (see the Argument Ontology sub-section).

Argument elements. Deb8 offers four key atomic argument analysis element classes (quotes, opinions, questions, and evidence). Arguments are meant to be built from left to right expanding the analysis of previous points, therefore elements connect to other elements from their left connectors and are connected to other objects on their right connectors.

Quote widgets display primary content (i.e., caption snippets and snapshots of the video) on the argument canvas. When the journalist drags a selection of text from the captions that she finds arguable, or a telling frame of the video, onto the canvas, this creates a quote element that contains the caption text or the snapshot of the video. Quotes are designed to be the roots of the argumentation threads in a graph, to which all other elements can connect. This is an explicit design decision to help the discussion stay focused on what is in the primary content itself (the video–G1) rather than allowing any opinion or evidence to exist independently, without connecting to any content. We believe that this supports collaboration despite people having different opinions (G3), as material drawn from the video is a primary source.

The quote widget (Figure 2) has an outbound connection port on its left (Fig. 2.A–because quotes can also be used as a type of evidence on a deeper part of the graph structure) and an inbound connector port on the right, to receive connections. In all widgets, inbound and outbound ports allow multiple connections (G2, G4, DP1, DP5). The main body of the widget displays the quote itself (Fig. 2.A) or the snapshot (Fig. 2.B). The small icon on the top left sets the video play location to the quote. The "G" icon on the top right of the caption quote widget launches a search on the browser with the current selection of words within the quote (G2, DP5).

Question. A journalist can create a question widget by dragging the question icon from the icon bar (Fig. 1.D) and typing the question text. Any member of the team can use the three color-coded buttons marked with a plus sign to create positive answers (e.g., in our example, Yes, as shown in Fig. 3), negative answers, or neutral answers to the question.

This widget is the only one which contains internal "sub-widgets": the answers. This breaches DP1 somewhat because we could have made answers separate widgets that connect to a question. However, here we decided to limit the flexibility of answering a question for the sake of simplification (G4) and, importantly, to help the team maintain an understandable common structure of their arguments (DP2). Answer sub-widgets are designed to group a variety of elements that can connect to them. I.e., we prefer to encourage many

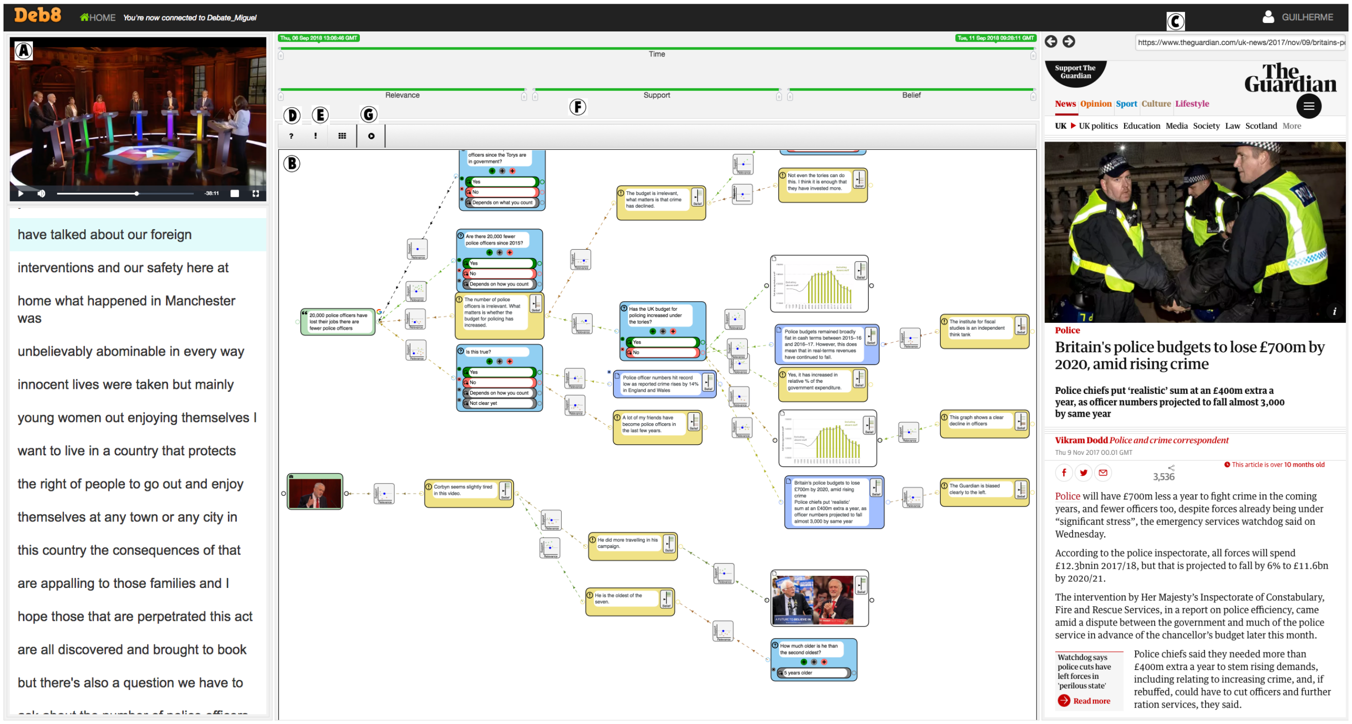

Figure 1: Main structure of the Deb8 interface.

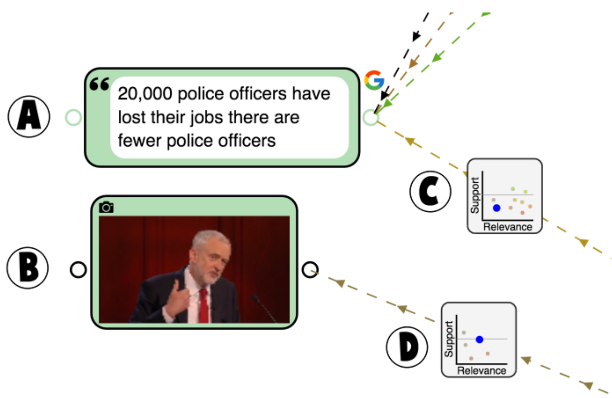

Figure 2: Caption (A) and snapshot quote widgets (B).

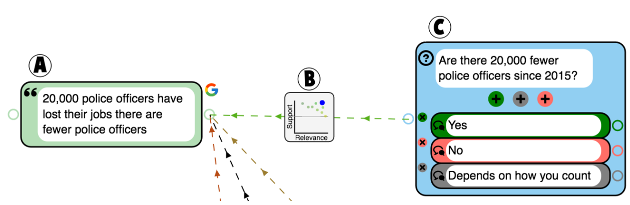

Figure 3: Question widget (right) connected to a quote (left).

opinions and bits of evidence (of different types) connected to a few answers (Yes, No, Maybe) rather than long lists of answers potentially similar to each other.

The way that a question is formulated and the negative (-) or positive (+) polarity of each answer is important. This is because a comment or evidence that supports a positive answer should also support whatever the question is connected to (G3, G4, DP2, DP3). This is further described in the Section on "Linking and Rating" below.

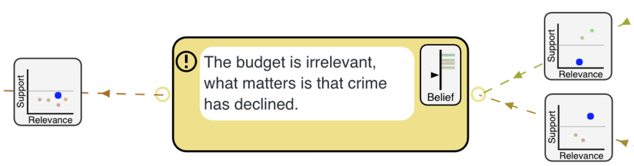

Figure 4: An opinion widget.

Opinion widgets contain text to explain or introduce an argument or idea, without providing direct evidence or stating a question (Fig. 4). Opinions are created by dragging the exclamation mark (Fig. 1.E) from the icon bar and have the usual ports as well as an additional widget that allows anyone connected to this debate to rate the degree of belief that they have on this particular opinion. The system records a data point for each analyst who moves the belief slider. The belief that other people have expressed on a particular opinion is visible in the widget as a shadow, forming a kind of histogram (see Fig. 4–supports G3, DP3, DP4). In the second part of the scenario, one could imagine how a large number of readers could weigh in to provide a well-sampled crowdsourced poll of the believability of this item.

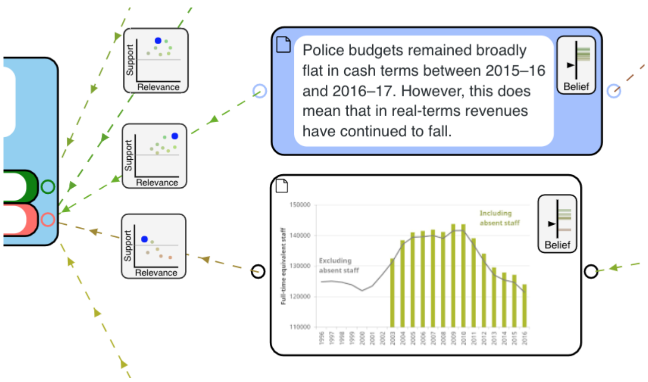

Figure 5: Two evidence widgets: text-based (top), and image-based (bottom).

Evidence is represented through a widget that can be text- or image-based (Fig. 5). A journalist can create evidence widgets by selecting and dragging text or an image from the browser of the rightmost panel of the application (Fig. 1.C). Evidence widgets store the link of the original source document. Any journalist in the team can later retrieve the web document to their right-panel web-browser by clicking on the document icon on the top left of the widget (Fig. 5). Evidence widgets have belief sliders identical to those of opinions. This widget allows the journalists to store their fact checking efforts within the debate analysis itself and also provides an appropriate entry point for those who want to follow up on the sources that they have linked.

Argument Linking and Rating System. The *belief rating slider* of the opinion and evidence widgets described above is one part of the rating system, and supports G3, enacting DP3. In addition, all other links between widgets are tagged in two dimensions with a square widget as shown in figures 2 to 4, with a close-up in Figure 6. This *link rating widget* allows each analyst to rate the connection between the origin and the destination widgets in two dimensions: *relevance* and *support*. Relevance denotes whether the origin of the link is on topic with respect to its destination. Moving the blue dot to the right means increased relevance. For example, a question such as "has the government forbidden alcohol advertisement" might be judged to have very low relevance to a quote stating that "20,000 police officers have lost their jobs" (i.e., most in the team will move the dot to the left). An opinion stating that "The number of police officers is irrelevant, what matters is crime stats trends" might be considered of middle relevance by some, for the same quote; evidence stating that "the number of police officers has declined in the last four years by 14%" might be considered very relevant, and most raters will move the blue dot to the right.

The other dimension of the square link rating widget is *support*. Support indicates the polarity of the relationship between the origin and the destination of the link. For example, an opinion that reads "police officers have not lost their jobs, they have retired" would be judged to be on the negative side of the support dimension (i.e., it contradicts the quote, which shows the blue dot lower in the widget), whereas a piece of evidence from a web article indicating that "15,000 police officers have been made redundant in 2018" would be considered highly supportive of the quote. In general, positive support means that the higher the importance and belief of the origin widget, the higher the importance or belief in the destination. Negative support inverts this relationship.

Although questions do not intrinsically support or contradict a quote, opinion or evidence, they have a polarity according to how they are formulated. The rule is that a link from a question should be rated as supportive if a *yes* or affirmative answer (indicated with green at the time of answer creation), is supportive of the element in the destination. This allows flexibility in the formulation of the question. For example, a question on the same quote about police officers above that is formulated as "Are there currently fewer police officers?" would be considered supportive of the quote, since a *yes* answer works in the same direction as the quote. If the question was "Are there currently more police officers?" the question should be rated non-supportive.

What each journalist sees on their own canvas are their own blue dots (one per link), along with a group-aware view of where all the other people's views of the relevance lay across the 2D chart (Figure 6). As before, a large number of ratings from the general readership of the newspaper could offer a quick overview of whether that connection is controversial (broad distribution of dots) or not (sharp distribution), and along which dimension.

Notice that even fairly straightforward relationships between widgets can be somewhat controversial. In the example above, an extremely meticulous journalist might consider whether a question about there being fewer police officers is relevant, since this might depend on the meaning of losing a job, and police force attrition could also happen through retirement. For this reason (i.e., to support different opinions–G3, at low granularity–DP3) each journalist can make their own judgment on relevance, support and belief. Finally, the color of the graphical link and of the dots behind double-encode support and relevance visually: average positive and negative support change the hue from green to red respectively, and relevance increases saturation (low relevance is close to gray, and high relevance makes the colors vivid).

Argument Ontology. The widgets described above are based on an argumentation ontology that we derived from the argumentation research field (reviewed in Section 3) to make the

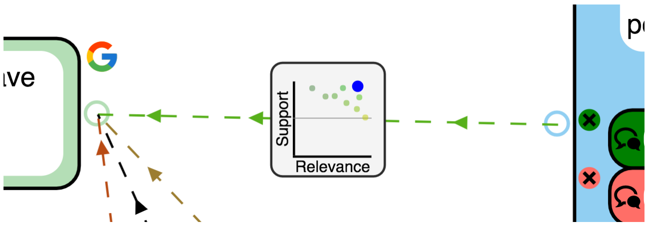

Figure 6: Close-up of the link rating widget connecting a quote with a question.

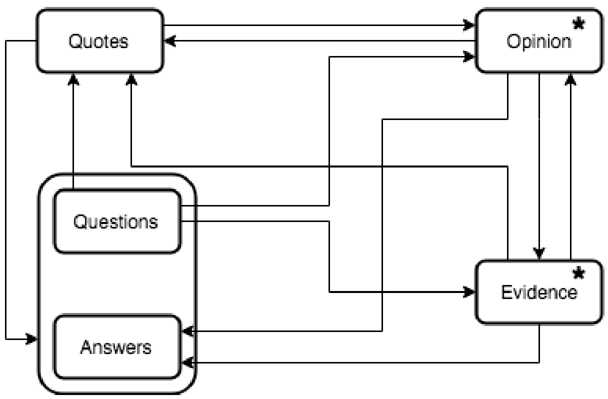

Figure 7: The ontology of argumentation widgets and their connections. Elements marked with an '*' have a belief rating slider. All links have support & relevance rating widgets.

interface theoretically sound (DP2) and simultaneously practical to use (DP5). Figure 7 summarizes the existing ontology. Note that *questions* can only be connected to via *answers* and that most of other links can go in either direction. All relationships can be many-to-many. The ontology that we derived is most closely related to the IBIS representation [18, 34] extended for collaborative analysis of debate. IBIS's 'Issues' would be posed as *questions* in Deb8, and Deb8's *quotes* and *answers* represent 'positions' in IBIS jargon. IBIS's 'Arguments' are instantiated instead through *opinion* and *evidence* widgets, of which plausibility and polarity can be rated via their relevance, support, and belief widgets.

Filtering and Layout. One downside of the atomic/constructive approach (DP1) is the potential complexity of the generated constructions [24, 25]. To manage complexity (G4), we support filtering and layout mechanisms (DP4). The filter bank is on top of the argument canvas (Figure 1.F) and has four filters controlling the visibility of widgets depending on time of creation, relevance, support and belief. Each slider bar has two handles, which allow low-, high- and band-pass filtering of each dimension.

The time filter allows the journalists to collaborate (G3). Moving the left handle to the right makes widgets created before this handle's time fade. By moving this slider's handle the analyst can see the most recent areas of discussion. This can be useful if the journalist has been elsewhere on the canvas for a while. Moving the right handle to the left makes any widgets created after the handle's position disappear. This supports a "manual replay" to understand how the graphs grew (a form of *provenance* support [30]). Creating new widgets brings the slider's handles to full positions, avoiding the complexity of alternative futures (G4, DP5).

The other three filters hide widgets that do not fit within the ranges of relevance, support and belief indicated through the filter sliders. However, unlike with the time slider, a disappearing widget will also make everything else downstream in the argument graphs disappear. This design decision is meant to simplify filtering very large trees (G4) based on the assumption that relevance, valence and trustworthiness propagate throughout the graph (DP3). For example, an element that supports something irrelevant is likely to be irrelevant. Through a combination of settings, the journalists can filter the current canvas to show, for example, only relevant supportive arguments, relevant contradicting arguments, or only those arguments which are considered irrelevant.

Multi-user Collaboration Policies

Deb8 is conceived as a synchronous/asynchronous distributed collaboration tool which can also work in co-located settings. Our scenarios suggest small to large numbers of people, which requires specific UI design decisions. We prioritized two principles here. First, to enable people with different opinions to collaborate (G3), we wanted to avoid edit wars as in Wikipedia [40]. For this, we enable anyone to judge any of the argument constructions through the link rating mechanism described above (DP3) while also locking each construction, including for deletion, once it has been built upon. Therefore widgets, including links, can only be removed and/or edited while nothing connects to them. Otherwise it would be easy for anyone to subvert the meaning of a branch by, e.g., adding the word "not" in a question's text. Objects are not visible to everyone until they are connected, directly or indirectly, to a quote.

Simultaneously, we recognize that people's ways of building arguments and organizing information is personal and constructive (DP1, DP4). Therefore the view in each client is unique: each journalist can choose to rearrange the elements in the 2D canvas however they like. When the number of elements gets large and others' contributions start popping up too fast (G4), it is possible to invoke an automatic force-directed constraint algorithm (based on [7]) to rearrange the elements in the argument canvas, starting from quotes on the left, and avoiding overlaps (DP4).

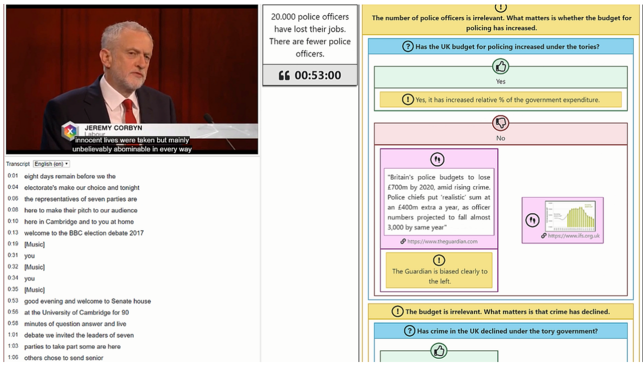

Figure 8: The Analysis Player Window uses containment to play sequentialized discussion graphs alongside the video.

Web browser

Most of the linking functionality between the web browser pane (Fig. 1.C) and the rest of the interface has been already described: journalists can search for evidence on the web through the "G" button in the widgets, and corresponding evidence widgets can be created on the canvas by dragging text or images from the browser. In addition, journalists can simply search and navigate the web through the search bar and by clicking on links, as with a regular web browser. All this functionality is individual; analysts cannot see what others are seeing or searching in their browsers.

Analysis Player

The features and panes described so far provide ways to build almost arbitrarily complex analyses by using the argument canvas to link content from the video and evidence from the Internet. However, they do not provide a simple way to consume the content of the analysis. We wanted to facilitate visualizing specific views of the video analysis in a way that is easier than traversing the argument graphs.

The Analysis Player is an independent feature that appears on a separate screen and sequentializes the structure of the debate to be played alongside the video. For example, a journalist could press the Analysis Player button in the icon bar (Fig. 1.G), to export the elements in the current view (i.e., it will ignore the filtered elements) to an external window (Fig. 8). The new window shows a Video and Caption Panel similar to that of the main interface. As the video plays, the quotes of the canvas appear in a middle panel with their corresponding timestamps. The right panel shows all the argument elements associated with each quote.

The player re-represents the connected elements from the argument canvas using nesting. For example, a question with two answers is shown as an object containing two elements. If one of these answers happens to be associated with a piece of evidence, this will, in turn, contain a visual representation of the evidence (text or image).

As in the main interface, the player's sections are interlinked. That is, interactions with the video's timeline or its captions trigger changes in the displayed quotes and argument elements, and vice versa. This allows the team of journalists to generate summaries of the argument analysis that can be consumed by just playing the video (with the argument elements rolling on the side). Different versions are easy to generate by filtering in different ways prior to invoking the player. For example, journalists might want to have a version with only extremely relevant arguments and another one with only supportive arguments.

Other features

The Deb8 interface tracks different users with a log-in system to allow future personalization of the configuration and saving of different debates and views. There is also a small interface to create new debates out of captioned video.

Implementation

Deb8 works as a web-client implemented in Javascript on the Electron framework[1] using Node-JS[2] and Video.js libraries[3] against a MongoDB database server. The argument canvas is implemented with the publicly available iVoLVER toolkit[4], which is, itself, implemented on top of the Fabric.js library[5].

6 INITIAL EVALUATION

We conducted an initial study of Deb8 to validate our design with respect to the goals discussed in Section 4 and to assess the viability of the approach. We observed three groups of participants performing an analysis of an election debate to gather insights on the use of Deb8 in collaborative analysis of video material (G1). Furthermore, we studied the support that Deb8 offers in facilitating linking evidence and opinions (G2), enabling different forms of collaborative debate analysis (G3) and the complexity of such interactions (G4).

Participants

We recruited 9 participants from a local university (2 females, ages 24-50, 5 native English speakers) in groups of 3. Previous experience in online discussions, reading online news and mind mapping tools varied among participants. Most participants (7) were readers of online news but with little or no active engagement, 7 had contributed to online discussions in social media, and 6 had some mind mapping experience.

[1] http:\electronjs.org
[2] http:\nodejs.org
[3] http:\videojs.com
[4] https:\github.com\ggmendez\iVoLVER
[5] http:\fabricjs.com

Procedure, Tasks and Analysis

Participants took place in sessions lasting approximately one hour and provided written consent according to local ethical procedures. We chose a seven-way 2017 UK general election TV debate for this study. Each session involved three phases: 1) a demo of the system by a facilitator to train the participants in the functionality of Deb8; 2) an individual task where each participant performed analyses of three selected quotes from the video regarding specific claims made by the debaters (T1—15 minutes); and 3) a collaborative task of wider scope, where the group focused on a specific point in the debate regarding priorities for making Britain a safer place (T2—15 minutes). All three participants were co-located in the room, each operating on an individual computer for both tasks, working on an individual canvas of Deb8 for T1, and on a shared canvas for T2.

Each session was video recorded and we took observational notes of the participants' interactions. We captured the screen of each participant's computer in order to see their interaction and the process of how each graph was built. To understand the use of widgets, we collected the argument graph built by each participant as well as the one resulting from the collaborative task. A questionnaire followed each task with open questions requiring participants to comment on features of Deb8. For T2, we collected opinions on whether the tool improves the groups' shared understanding.

The analysis was accomplished by: observing the recorded video of each session; observing the captured screen of each participant's computer; analyzing the final graph generated by each participant for each task and, finally, analyzing the questionnaires. We used these observations to understand the participant's behaviour in terms of interaction, communication and collaboration. We focused on group strategies for building the graph and to search for evidence on the web.

Evaluation observations

The analysis graphs generated during the study have the following characteristics. In T1, we see the use of 9.9 widgets on average divided as: 0.8 questions, 1.9 opinions, 3.0 quotes, and 4.2 evidence nodes. The maximum depth of the analysis was 2 with a branching factor on average of 2.3 (excluding unconnected quotes). On average in T2 participants used 25.3 widgets divided as: 13.7 quotes (of which 3.0 were video snap shots), 4.7 evidence nodes, 3.7 opinions, and 1.7 questions. The maximum depth was 3 with an average branching of 1.5. Figure 9 shows an example of collaborative analysis in T2.

Use of tool. Overall, participants were able to construct argumentation graphs of a reasonable size in a short amount of time. We think that the average graph sizes of over 10 for T1 and 25 for T2 linked widgets are notable for just a 15 minute period. The resulting graphs show plausible analyses and

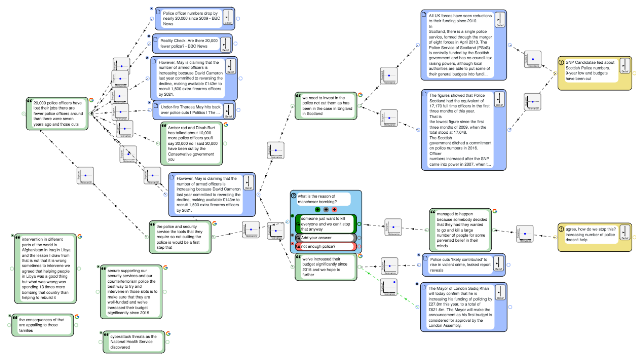

Figure 9: An example of T2 final analysis graph

use of different types of widgets with a mostly coherent semantic. This suggests that Deb8 can support the generation of structured commentary and that its functionality can be learned after a short training.

Participants' responses from the questionnaires highlight that Deb8 was particularly helpful for searching and connecting external evidence in the analysis of quotes. Participants emphasized how this support reduced the effort of switching between different tools or applications. Participants also complimented the ability to extract quotes and replaying the debate from the location of a particular quote.

Collaboration. A collaborative task would usually introduce higher interaction complexity. We observed that, despite the increased complexity, the three graphs from T2 are meaningful and of a reasonable size with respect to the time allocated for this task. In T2 participants created relatively larger constructions than in T1, with a higher number of threads of analysis. Advancing at different speeds, it was common for participants to switch their focus to different sub-graphs of the canvas argument to share opinions and provide evidence.

Because of the diverse social dynamics, often characterized by interleaving of individual and cooperative phases [38], working in collaboration to construct shared understanding is a complex process. Nevertheless, participants were able to cope with these complexities by using Deb8's UI flexibly (DP4, DP5). While participants focused and contributed on the collaborative analysis, they also created 'personal spaces' of analysis within the canvas to focus on individual perspectives before bringing them into view to share them with the other members of the group. While the entire canvas is shared, patterns of territoriality emerged, giving rise to various forms of sharing and personal use.

The questionnaire responses also show that most participants appreciated the ability to add questions and learn from others' opinions. However, they suggested that collaboration could be more effective if Deb8 provided additional support for group tasks. For example, highlighting dynamic updates of canvas elements would allow members of the group to focus on points of analysis being currently considered.

Additionally, we noticed that the Deb8 relevance, support and belief rating system was not fully used in T2. This may be attributed to the limited time available to complete the task. Participants might have constructed their arguments but have not yet had the opportunity to critically reflect on the significance of the claims from other members of the group to converge on a conclusion. Such graph sketching may be followed by periods of reflection, as more people employ the filters to focus their attention on material with a particular level of support and relevance.

Arguments and analysis. We observed varied reasoning patterns and argumentation structures. We noticed instances of debate argument analysis, where quotes are linked to other quotes to map the linear thread of debate to an argument graph, that are similar to existing analytical approaches (e.g., [19]). The process involved many instances of fact checking to establish whether a debater's claim was plausible, and information seeking to find out more about the topic. Both tasks were accomplished by introducing evidence from the web and linking to quotes or answers of questions. These patterns align with requirements identified for audience engagement in televised debates (e.g., [29]). Participants used opinions to state personal conclusions after analyzing existing evidence (as shown in Figure 9), and to provide additional support and share conflicting views with existing claims as is typical in argumentation processes.

While an in-depth analysis of the types of underpinning reasoning processes is out of scope, we noticed that the simple modular design of Deb8 elements (DP1, DP5) enables the formation of rich argumentative structures. Our initial findings suggest that the capabilities of Deb8 provide support for many interlinked reasoning processes such as debate analysis, evidence formation, and collaborative discussion in a coherent format and space for analysis.

7 DISCUSSION AND LIMITATIONS

This paper has introduced the design and implementation of Deb8, a tool for collaborative analysis of video debates that introduces a number of novel features and that is based on a principled design supported by current knowledge in argumentation. The system addresses a complex problem and offers a sophisticated interface that will likely require training and might not be accessible to everyone. This interface borrows elements from the design of graphically-structured argumentation tools such as debatemap.live, debategraph.org, Rationale & bCisive (reasoninglab.com), AGORA-net [13], and OVA+ [16]. Unlike many of these tools, however, Deb8 provides specific features to connect the video/captions source, with the argumentation schema and the evidence. Furthermore, the tool offers collaborative features that extend its use beyond the relatively rare expertise of argument analysts. For example, the ability to integrate the relevance, support and belief ratings across a large number of people collaborating synchronously over the web is distinctive of Deb8, as is the ability of each separate user to create their own filters and layouts to support their own thinking.

The results from the preliminary evaluation are promising and indicate that our participants were adept at creating arguments and finding evidence to back up or disprove arguments, but there are still important outstanding questions. Specifically, what happens when larger numbers of people use the tool? Will the filtering and rating mechanisms enable analysts to cope with the complexity of dealing with large numbers of potentially conflicting opinions and sources of evidence? Do the shared canvas, ontology and widgets have significant weaknesses that allow small numbers of users to hijack the arguments or vandalize the work of many?

Further research is needed to determine to what extent the structured approach that Deb8 enforces supports deep analysis (G1) and enables collaboration of people with different opinions (G3–our preliminary study did not support sessions long enough to see rating behavior). However, even with our small sample, G2 (linking) seems well supported. G4 (managing complexity) is probably the hardest goal, and it is possible that the design will require further collaboration features and policies to become more suitable, especially for larger groups. We also acknowledge that our ontology only focuses at defining analytic elements of debate, and can be further refined with additional widgets capturing different conceptual levels of analysis such as authors or the candidate that made a quote. This kind of extension would enable more powerful filtering and navigation of the arguments.

Comparisons with different approaches such as Rationale & bCisive (reasoninglab.com), Compendium [34], AGORA-net [13], debatemap.live, and OVA+ [16] or other baselines would be useful to understand what features are more and less useful, but appropriate design, presentation and discussion of empirical comparisons with other systems requires significant additional work beyond the scope of this paper.

There are also deeper questions related to the nature of the interface and argumentation itself. Further research, which can be supported by Deb8, can shed light on exactly how graphical layout interfaces can provide better support of argumentation than their mostly linear text-based counterparts, and why. It is also important to further understand how structuring argumentation might be able to help people with opposite opinions to reach certain types of consensus or at least agree to exchange ideas in civilized rule-based ways. Although Deb8 is a simple step in this direction, we believe that it could be instrumental in answering these questions.

8 CONCLUSION

We presented the design, implementation and preliminary evaluation of Deb8, a system that enables deep collaborative analysis of video-based debates. The system enables dense linking of information across multiple types of media to structure arbitrarily complex analyses based on argumentation theory. The combination of features is unique and follows a set of design principles that address the intricate space of political argumentation in the public sphere. Deb8 and other related tools may, directly or indirectly, help encourage evidence-based debate and better political accountability.

9 ACKNOWLEDGMENTS

This research is supported by the Brazilian National Council of Technological and Scientific Development (CNPq) and the National Research Foundation, Prime Minister's Office, Singapore under its International Research Centres in Singapore Funding Initiative.

REFERENCES

[1] Pranav Anand, Marilyn Walker, Rob Abbott, Jean E. Fox Tree, Robeson Bowmani, and Michael Minor. 2011. Cats Rule and Dogs Drool!: Classifying Stance in Online Debate. In *Proceedings of the 2Nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA '11)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 1–9. http://dl.acm.org/citation.cfm?id=2107653.2107654

[2] Olfa Ben Ahmed, Gabriel Sargent, Florian Garnier, Benoit Huet, Vincent Claveau, Laurence Couturier, Raphaël Troncy, Guillaume Gravier, Philémon Bouzy, and Fabrice Leménorel. 2017. NexGenTV: Providing Real-Time Insight During Political Debates in a Second Screen Application. In *Proceedings of the 2017 ACM on Multimedia Conference (MM '17)*. ACM, New York, NY, USA, 1259–1260. https://doi.org/10.1145/3123266.3127929

[3] Carlos Chesnevar, Jarred McGinnis, Sanjay Modgil, Iyad Rahwan, Chris Reed, Guillermo Simari, Matthew South, Gerard Vreeswijk, and Steven Willmott. 2006. Towards an Argument Interchange Format. *The Knowledge Engineering Review* 21, 4 (2006), 293–316. https://doi.org/10.1017/S0269888906001044

[4] Helen Clark. 2000. The Worm that Turned: New Zealand's 1996 General Election and the Televised 'Worm' Debates. In *Televised Election Debates: International Perspectives*, Stephen Coleman (Ed.). Palgrave Macmillan UK, London, 122–129. https://doi.org/10.1057/9780230379602_7

[5] Jeff Conklin. 2003. Dialog Mapping: Reflections on an Industrial Strength Case Study. In *Visualizing Argumentation: Software Tools for Collaborative and Educational Sense-Making*, Paul A. Kirschner, Simon J. Buckingham Shum, and Chad S. Carr (Eds.). Springer London, London, 117–136. https://doi.org/10.1007/978-1-4471-0037-9_6

[6] Cycling '74. 2018. Max Software Tools for Media | Cycling '74. https://cycling74.com/products/max

[7] Tim Dwyer, Yehuda Koren, and Kim Marriott. 2006. IPSep-CoLa: An incremental procedure for separation constraint layout of graphs. *IEEE Transactions on Visualization and Computer Graphics* 12, 5 (2006), 821–828.

[8] Martin Erwig, Karl Smeltzer, and Xiangyu Wang. 2017. What is a visual language? *Journal of Visual Languages & Computing* 38 (2017), 9 – 17. https://doi.org/10.1016/j.jvlc.2016.10.005 SI:In honor of Prof SK Chang.

[9] F. Galton. 1907. Vox populi. *Nature* 75, 1949 (1907), 7. https://doi.org/10.1038/075450a0

[10] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. 2017. The Effect of Collective Attention on Controversial Debates on Social Media. In *Proceedings of the 2017 ACM on Web Science Conference (WebSci '17)*. ACM, New York, NY, USA, 43–52. https://doi.org/10.1145/3091478.3091486

[11] Katerina Gorkovenko and Nick Taylor. 2016. Politics at Home: Second Screen Behaviours and Motivations During TV Debates. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI '16)*. ACM, New York, NY, USA, Article 22, 10 pages. https://doi.org/10.1145/2971485.2971514

[12] Naeemul Hassan, Chengkai Li, and Mark Tremayne. 2015. Detecting Check-worthy Factual Claims in Presidential Debates. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management (CIKM '15)*. ACM, New York, NY, USA, 1835–1838. https://doi.org/10.1145/2806416.2806652

[13] Michael H. G. Hoffmann. 2015. Changing Philosophy Through Technology: Complexity and Computer-Supported Collaborative Argument Mapping. *Philosophy & Technology* 28, 2 (01 Jun 2015), 167–188. https://doi.org/10.1007/s13347-013-0143-6

[14] Bernd Huber, Sarah Tausch, and Heinrich Hußmann. 2014. Supporting Debates with a Real-time Feedback System. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14)*. ACM, New York, NY, USA, 2257–2262. https://doi.org/10.1145/2559206.2581281

[15] IBM. Accessed March 2019. i2 Analyst's Notebook. https://www.ibm.com/us-en/marketplace/analysts-notebook.

[16] Mathilde Janier, John Lawrence, and Chris Reed. 2014. OVA+: An argument analysis interface. In *Proceedings of the Fifth International Conference on Computational Models of Argument*, Simon Parsons, Nir Oren, Chris Reed, and Federico Cerutti (Eds.). Vol. 266. IOS Press, 463–564.

[17] Wesley M. Johnston, J. R. Paul Hanna, and Richard J. Millar. 2004. Advances in Dataflow Programming Languages. *ACM Comput. Surv.* 36, 1 (March 2004), 1–34. https://doi.org/10.1145/1013208.1013209

[18] Werner Kunz and Horst Rittel. 1970. *Issues as elements of information systems.* Working paper no. 131. Center for Planning and Development Research.

[19] John Lawrence, Mark Snaith, Barbara Konat, Katarzyna Budzynska, and Chris Reed. 2017. Debating Technology for Dialogical Argument: Sensemaking, Engagement, and Analytics. *ACM Trans. Internet Technol.* 17, 3, Article 24 (June 2017), 23 pages. https://doi.org/10.1145/3007210

[20] Rhema Linder, Nic Lupfer, Andruid Kerne, Andrew M. Webb, Cameron Hill, Yin Qu, Kade Keith, Matthew Carrasco, and Elizabeth Kellogg. 2015. Beyond Slideware: How a Free-form Presentation Medium Stimulates Free-form Thinking in the Classroom. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition (C&C '15)*. ACM, New York, NY, USA, 285–294. https://doi.org/10.1145/2757226.2757251

[21] Anders Sundnes Løvlie. 2017. From Playground to Salon: Challenges in Designing a System for Online Public Debate. In *Proceedings of the 21st International Academic Mindtrek Conference (AcademicMindtrek '17)*. ACM, New York, NY, USA, 164–173. https://doi.org/10.1145/3131085.3131111

[22] Mattias Mano, Jean-Michel Dalle, and Joanna Tomasik. 2018. The Consent of the Crowd Detected in an Open Forum. In *Proceedings of the 14th International Symposium on Open Collaboration (OpenSym '18)*. ACM, New York, NY, USA, Article 14, 9 pages. https://doi.org/10.1145/3233391.3233538

[23] Rolando Medellin-Gasque, Chris Reed, and Vicki Hanson. 2014. Older Adults Interaction with Broadcast Debates. In *Proceedings of the 16th International ACM SIGACCESS Conference on Computers & Accessibility*

(ASSETS '14). ACM, New York, NY, USA, 269–270. https://doi.org/10.1145/2661334.2661394

[24] Gonzalo Gabriel Méndez, Uta Hinrichs, and Miguel A. Nacenta. 2017. Bottom-up vs. top-down: trade-offs in efficiency, understanding, freedom and creativity with InfoVis tools. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 13. https://doi.org/10.1145/3025453.3025942

[25] Gonzalo Gabriel Méndez, Miguel A. Nacenta, and Uta Hinrichs. 2018. Considering Agency and Data Granularity in the Design of Visualization Tools. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 13. https://doi.org/10.1145/3173574.3174212

[26] Gonzalo Gabriel Méndez, Miguel A. Nacenta, and Sebastien Vandenheste. 2016. iVoLVER: Interactive Visual Language for Visualization Extraction and Reconstruction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 4073–4085. https://doi.org/10.1145/2858036.2858435

[27] Alex F Osborn. 1953. *Applied imagination, principles and procedures of creative thinking.* Charles Scribner\'s Sons.

[28] Ayush Patwari, Dan Goldwasser, and Saurabh Bagchi. 2017. TATHYA: A Multi-Classifier System for Detecting Check-Worthy Statements in Political Debates. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17)*. ACM, New York, NY, USA, 2259–2262. https://doi.org/10.1145/3132847.3133150

[29] Brian Plüss and Anna De Liddo. 2015. Engaging Citizens with Televised Election Debates Through Online Interactive Replays. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*. ACM, New York, NY, USA, 179–184. https://doi.org/10.1145/2745197.2755521

[30] Eric D. Ragan, Alex Endert, Jibonananda Sanyal, and Jian Chen. 2016. Characterizing Provenance in Visualization and Data Analysis: An Organizational Framework of Provenance Types and Purposes. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan 2016), 31–40. https://doi.org/10.1109/TVCG.2015.2467551

[31] Iyad Rahwan, Mohammed I. Madakkatel, Jean-François Bonnefon, Ruqiyabi N. Awan, and Sherief Abdallah. 2010. Behavioral Experiments for Assessing the Abstract Argumentation Semantics of Reinstatement. *Cognitive Science* 34, 8 (2010), 1483–1502. https://doi.org/10.1111/j.1551-6709.2010.01123.x

[32] Patrick C. Shih, David H. Nguyen, Sen H. Hirano, David F. Redmiles, and Gillian R. Hayes. 2009. GroupMind: Supporting Idea Generation Through a Collaborative Mind-mapping Tool. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work (GROUP '09)*. ACM, New York, NY, USA, 139–148. https://doi.org/10.1145/1531674.1531696

[33] Ben Shneiderman. 2000. Creating Creativity: User Interfaces for Supporting Innovation. *ACM Trans. Comput.-Hum. Interact.* 7, 1 (March 2000), 114–138. https://doi.org/10.1145/344949.345077

[34] Simon J. Buckingham Shum, Albert M. Selvin, Maarten Sierhuis, Jeff Conklin, Charles B. Haley, and Bashar Nuseibeh. 2006. Hypermedia Support for Argumentation-Based Rationale. In *Rationale Management in Software Engineering*, Allen H. Dutoit, Raymond McCall, Ivan Mistrík, and Barbara Paech (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 111–132. https://doi.org/10.1007/978-3-540-30998-7_5

[35] Z. Skrba, B. O'Mullane, B. R. Greene, C. N. Scanaill, C. W. Fan, A. Quigley, and P. Nixon. 2009. Objective real-time assessment of walking and turning in elderly adults. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 807–810. https://doi.org/10.1109/IEMBS.2009.5333934

[36] Brian Stelter. 2016. Debate breaks record as most-watched in U.S. history. Retrieved April 3, 2019 from https://money.cnn.com/2016/09/27/media/debate-ratings-record-viewership/

[37] Chenhao Tan, Hao Peng, and Noah A. Smith. 2018. "You Are No Jack Kennedy": On Media Selection of Highlights from Presidential Debates. In *Proceedings of the 2018 World Wide Web Conference (WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 945–954. https://doi.org/10.1145/3178876.3186142

[38] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. 2006. Collaborative Coupling over Tabletop Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 1181–1190. https://doi.org/10.1145/1124772.1124950

[39] Alice Toniolo, Timothy J. Norman, Anthony Etuk, Federico Cerutti, Robin Wentao Ouyang, Mani Srivastava, Nir Oren, Timothy Dropps, John A. Allen, and Paul Sullivan. 2015. Supporting reasoning with different types of evidence in intelligence analysis. In *Proceedings of the Fourteenth International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 781–789.

[40] Frank van Ham, Fernanda B. Viegas, Jesse Kriss, and Martin Wattenberg. 2007. Talk Before You Type: Coordination in Wikipedia. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*. 78. https://doi.org/10.1109/HICSS.2007.511

[41] William Wright, David Schroh, Pascale Proulx, Alex Skaburskis, and Brian Cort. 2006. The Sandbox for Analysis: Concepts and Methods. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 801–810. https://doi.org/10.1145/1124772.1124890