Visualization of the Topic Space of Argument Search Results in args.me

Yamen Ajjour *   **Henning Wachsmuth** **   **Dora Kiesel** *   **Patrick Riehmann** *
Fan Fan *   **Giuliano Castiglia** *   **Rosemary Adejoh** *   **Bernd Fröhlich** *   **Benno Stein** *
* Bauhaus-Universität Weimar, Weimar, Germany, `<first>.<last>@uni-weimar.de`
** Paderborn University, Paderborn, Germany, `henningw@upb.de`

Abstract

In times of fake news and alternative facts, pro and con arguments on controversial topics are of increasing importance. Recently, we presented *args.me* as the first search engine for arguments on the web. In its initial version, args.me ranked arguments solely by their relevance to a topic queried for, making it hard to learn about the diverse topical aspects covered by the search results. To tackle this shortcoming, we integrated a visualization interface for result exploration in args.me that provides an instant overview of the main aspects in a barycentric coordinate system. This *topic space* is generated ad-hoc from controversial issues on Wikipedia and argument-specific LDA models. In two case studies, we demonstrate how individual arguments can be found easily through interactions with the visualization, such as highlighting and filtering.

1 Introduction

For many controversial topics in life and politics, people disagree on what is the right stance towards them, be it the need for feminism, the influence of religion, or the assassination of dictators. Stance is affected by the subjective assessment and weighting of pro and con arguments on the diverse aspects of a topic (Kock, 2007). Building stance in a self-determined manner is getting harder and harder in times of fake news and alternative facts, due to the unclear reliability of many sources and their bias in stance and covered aspects. This was our societal motivation for the development of the first dedicated argument search engine, *args.me*.[1]

args.me allows querying for arbitrary controversial topics. As search results, it opposes pro and con arguments from the web, ranked by their computed relevance to the topic. args.me is non-commercial and aims to avoid bias towards either stance. In

(Wachsmuth et al., 2017b), we introduced the software framework and the initial argument index underlying args.me as well as its basic user interface.

In its first version, args.me presented arguments in textual form with linked sources, similar to the web page snippets of conventional search engines, but with color-encoded stance. Examples are given below in Figures 1 and 2. This is adequate for comprehending those arguments deemed most relevant. Unlike for many general information needs (Croft et al., 2009), however, reading the top results is not enough for building an informed stance. Rather, diverse aspects of a controversial topic need to be explored. In our recent study with 97 international users, *aspect coverage* was seen as the second most important ranking criterion — after *source reliability* but before *recency*, *user ratings*, and others.[2] A simple relevance ranking of possibly thousands of arguments provides weak support in that regard.

This paper shows a novel way of presenting argument search results, which we designed and integrated into args.me to support a rapid exploration of the aspects of a topic. In particular, we visualize this *topic space* in a barycentric coordinate system (Riehmann et al., 2018), representing the distribution of pro and con arguments over the main covered aspects (see Figure 3). Possible aspects were derived offline from the Wikipedia list of controversial issues[3] as well as from LDA topic models built based on the 291k arguments in our index, whereas the aspects actually visualized are derived ad-hoc from the search results. Through interactions with the visualization, a user can easily highlight and filter arguments on the aspects of interest. In two case studies, we demonstrate how the visualization speeds up argument search notably.

---

[1] Also known as just *args*, found at: `https://args.me`

[2] The user study, also including other questions related to argument search, is going to be published in another venue.

[3] Issue list: `https://en.wikipedia.org/wiki/Wikipedia:List_of_controversial_issues`
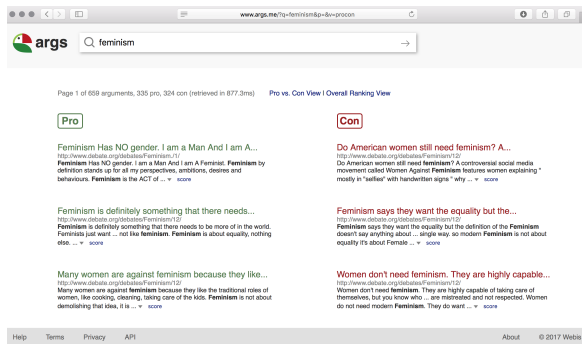
Figure 1: The *pro vs. con view* of args.me, showing the argument search results for the query "feminism".

Figure 2: The *overall ranking view* of the initial version of args.me, showing results for the query "feminism".

2 Related Work

Different systems to visually create and analyze arguments have been introduced in the past. Some of their visualizations serve as a mind map to support ongoing discussions, such as *gIBIS* (Conklin and Begeman, 1988), *Belvedere* (Suthers et al., 1995), *ArgVis* (Karamanou et al., 2011), *Dicode* (Tzagarakis and Karacapilidis, 2013), and *Debate-Graph* (Baldwin and Price, 2018). Others allow for evaluating the structure and schemes of arguments, such as *Araucaria* (Reed and Rowe, 2004), *Rationale* (van Gelder, 2007), *ArgueApply* (Pührer, 2017), *Argunet* (Betz et al., 2018), and *Truthmapping* (Truthmapping, 2018).

To support achieving consensus in a discussion, *SEAS* (Lowrance et al., 2000), *VUE* (Baroni et al., 2015), and *Dialectic Map* (Niu, 2016) provide a combination of automatic argument analysis and visual argument summaries. With similar intentions, *Lexical Episode Plots* (Gold et al., 2015), *ConToVi* (El-Assady et al., 2016), *NEREx* (El-Assady et al., 2017), and Jentner et al. (2017) visualize specific aspects of transcribed discussions.

All these works focus on single arguments or the set of arguments within a single debate or text. In contrast, we present a visualization that summarizes arguments from many different texts. Unlike in (Wachsmuth et al., 2017a), where we illustrated *structural* argumentation patterns in the texts of a corpus, here we target the *content* of arguments. As the above-mentioned system ConToVi, we visualize the topic space covered by a set of arguments. While ConToVi provides insights into the flow of aspects during the discussion of a controversial topic, our visualization aims to make arguments on specific aspects easily findable. Moreover, we allow arguments to cover a weighted distribution of multiple aspects rather than only a single aspect.
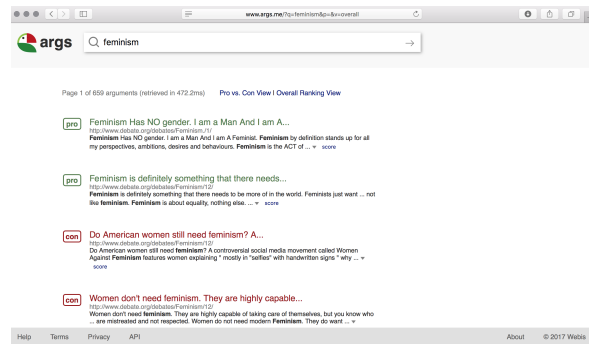
3 Argument Search with args.me

As presented in (Wachsmuth et al., 2017b), the initial version of args.me follows approved concepts of conventional search engines (such as *Google* or *Bing*), but it adapts them to the specific goals of argument search. Via the interface of args.me, users can enter free text queries on controversial topics, such as "feminism" and "assassination of dictators". While conventional search engines return links to web pages along with short textual excerpts as results, args.me directly returns all arguments found to be relevant, linked to their source web pages.

Originally, args.me provided two views that displayed the found arguments in a textual result list with color-encoded stance, as shown in Figures 1 and 2: a *pro vs. con view* that opposes the most relevant pros and cons, and an *overall ranking view* that ranks all arguments by their relevance irrespective of stance. A fundamental question in this regard is what arguments are actually deemed *most relevant*? Argument search implies specific ranking criteria, such as recency, perceived quality, aspect coverage, and source reliability. Assessing these criteria is all but trivial and partly unsolved. In its current state (mid 2018), args.me therefore still relies on standard information retrieval measures (Robertson and Zaragoza, 2009), i.e., it ranks arguments higher the more they match the words in the query.

So far, however, for many queries the top-ranked arguments neither appear to be the strongest on the given topic, nor do they cover the whole diversity of the topic. Having only a textual result list can then make it hard to find the best arguments or specific arguments of interest. This is why we aimed for new ways to support an efficient search space exploration — one main goal of information visualization (Munzner, 2014). While our recent user study suggested that source reliability
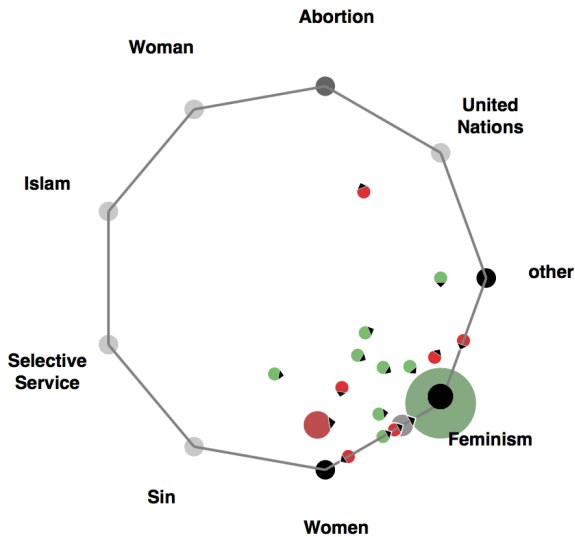
Figure 3: Topic space visualization for the query "feminism", positioning the retrieved arguments according to the eight main covered aspects and *other*.

Figure 4: Positioning an argument glyph in the topic space: the black arrow shows the linear combination of weighted vertices (*Aspect 3* 0.2, *Aspect 5* 0.3, *Aspect 1* 0.5). The glyph itself points to the main covered aspect.

is most important (see Section 1), the 291,440 arguments currently indexed by args.me anyway come from five selected sources only (Wachsmuth et al., 2017b). Instead, we thus focus on the second most important ranking criterion: aspect coverage.

4   Visualization of the Topic Space

For building an informed stance on a controversial topic, obtaining insights into the variety of aspects touched by the topic is crucial. We aid this process by accompanying the textual result list in args.me with a topic space visualization, which puts the aspects covered by the listed arguments into the focus, highlights the main aspect of each argument, and groups arguments covering similar aspects.

4.1   Determining Topical Aspects

The first step to develop the visualization was to build a topic model that can represent the aspects of each argument in the result list. We compared two alternative approaches for this purpose:

First, we computed the relative distribution of all the over 1000 terms from the Wikipedia list of controversial issues in each indexed argument. For instance, if "Women" occurs ten times, "Woman" six times, "Feminism" four times, and no other term, then we have *(Women 0.5, Woman 0.3, Feminism 0.2)* with implicit zeros for all others. Second, we performed LDA topic modeling (Blei et al., 2003) based on the words in all arguments from our index. With an interval size of 10, we tested all numbers of topics from 10 to 100 and chose the
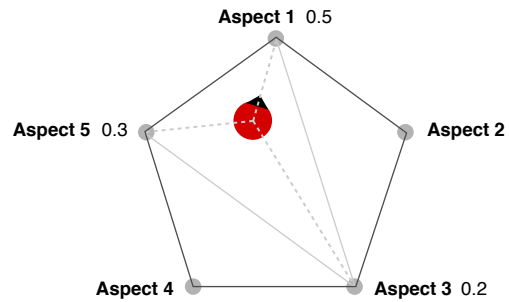
number that minimized perplexity: 40. Each aspect is then represented by all words of one LDA topic, and the relative aspect distribution is calculated by counting the occurrence of all associated words in each argument. We found the Wikipedia-based topic model to be more convincing, which is why it is set as the default in args.me.[4]

4.2   Visualizing Topical Aspects

To visualize the aspect-based topic space, we opted for *generalized barycentric coordinates* (Meyer et al., 2002), as they naturally fit our purpose: We represent an argument as a linear combination of weights for all aspects, while barycentric coordinates represent a point as a linear combination of the vertices of a polygon (both adding up to 1.0). Thus, the topic model can be used as input for the visualization without recalculation. Figure 3 shows the visualization of the results for the query "feminism", consisting of two main elements: the *topic space* and the *argument glyphs* within this space.

The topic space is depicted as a regular polygon with one vertex for each represented aspect. Both given topic model alternatives comprise too many aspects to depict them all. To reduce visual clutter in favor of a lean visualization, we limit the maximum number of visualized aspects, such that readability is not diminished. In particular, we keep only those eight aspects that are the most frequent in the argument search results. All other aspects are merged into a ninth aspect "other". The labels for the aspects are short terms in case of the Wikipedia-based topic model or visualized as word clouds in case of the LDA topic model.

Each argument glyph represents one argument in the form of a colored circle (green for pro, red

---

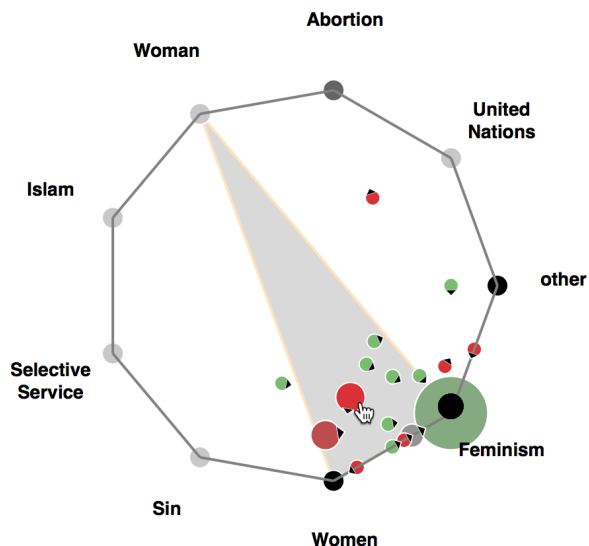[4]The LDA alternative can be activated in args.me by changing the value of the *v*-parameter in the URL field to "lda".

Figure 5: Hovering over an argument reveals the aspects it covers (main aspect marked by a small arrow).

Figure 6: Selecting arguments in the topic space visualization filters them in the textual result list of args.me.

for con) with a small arrow pointing to the main covered aspect. The glyphs are positioned based on their aspect distribution: the stronger one aspect, the stronger a glyph is "pulled" in that direction, as sketched in Figure 4. Accordingly, similar arguments are placed spatially near to each other. To ensure the visibility of all glyphs and to avoid overplotting, arguments placed on top of each other are aggregated into a single glyph. The glyph size depends on a logarithmic mapping of the number of represented arguments. Since arguments with both stances may be grouped, the color of an aggregate glyph represents the majority stance of all arguments contained, from green (all pro), over gray (balanced pro/con), to red (all con).

4.3 Interacting with Topical Aspects

The integration of our visualization into args.me is shown below in Figures 7 and 8. This new *topic space view* replaces the old overall ranking view: it includes the textual argument ranking and adds the visualization to the right. At first, the visualization shows only the information outlined above, but it provides further details upon interaction.

Barycentric coordinates are ambiguous and may place arguments with different aspects at similar locations. For disambiguation, users can hover over a glyph to reveal all covered aspects, as exemplified in Figure 5. The represented arguments are also highlighted in the textual list, given that they appear on the current result page. Vice versa, hovering over a textual argument highlights the respective glyph with a wide green or red border.
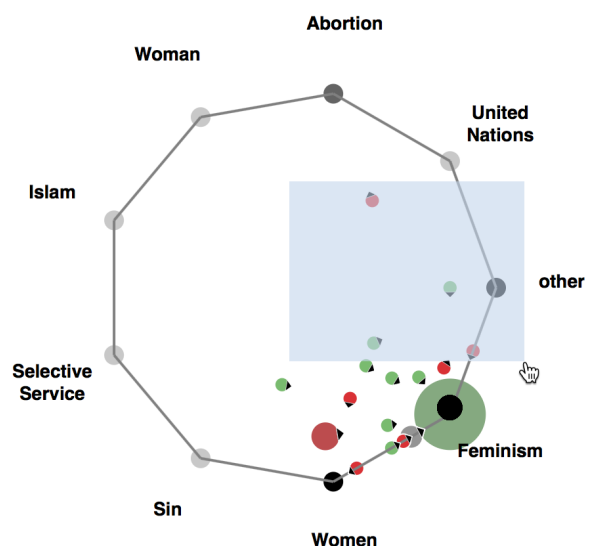
In addition, the visualization enables a filtering of the textual results: A user can select one or more arguments by clicking or brushing (see Figure 6), in order to narrow down the list to the aspects of interest. All other arguments are grayed out.

5 Case Studies

To verify the benefit of our visualization, we finally explore two typical use cases of argument search: *topic space exploration* and *search refinement*.

5.1 Topic Space Exploration

First, we consider a query for "feminism". 659 arguments are returned by args.me for this topic, as shown in Figure 7. While the top-ranked arguments seem highly relevant in general, our visualization reveals that also some rather specific aspects are covered by the search results, such as "Abortion" and "United Nations". Interacting with the visualization helps explore the entire topic space.

In particular, hovering over the argument glyphs clarifies what aspects they exactly cover, such as "Woman", "Women", and "Feminism" itself for the highlighted argument in Figure 7. After a first exploration via hovering, a result subset of interest can be filtered through brushing, say, the four topmost glyphs (see Figure 6 above). The selected arguments are then shown at the top of the textual result list (all below are grayed out). From the selected arguments, we learn that Emma Watson has made the need for feminism a point at the United Nations, whereas the claimed necessity of abortion is used as an analogy to justify the necessity of
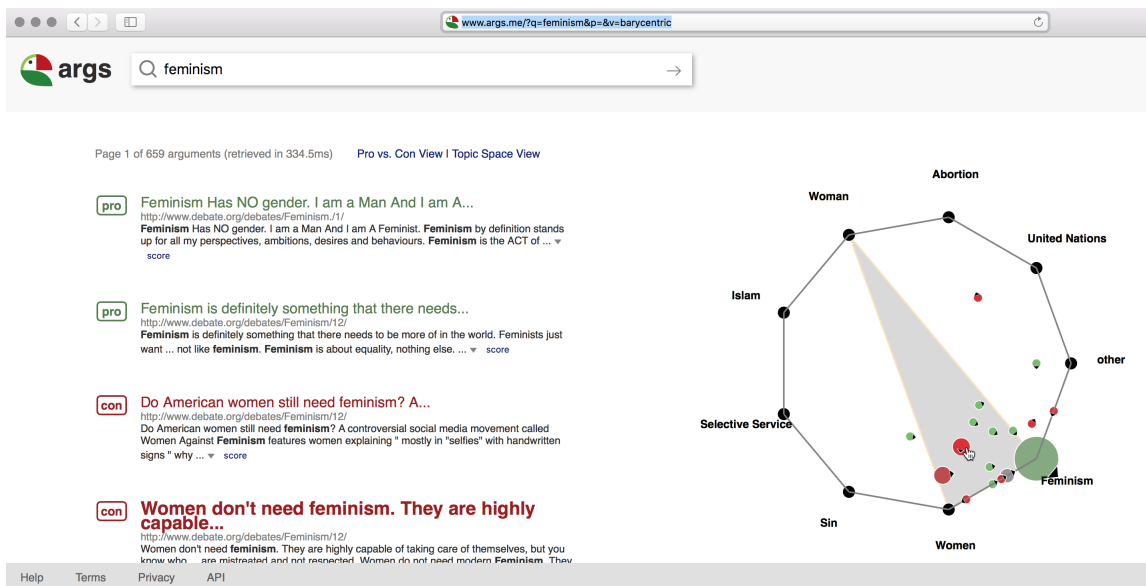
Figure 7: The args.me search results for the query "feminism", along with the integrated topic space visualization. The argument hovered over in the visualization is highlighted in the textual result list ("Women don't need...").
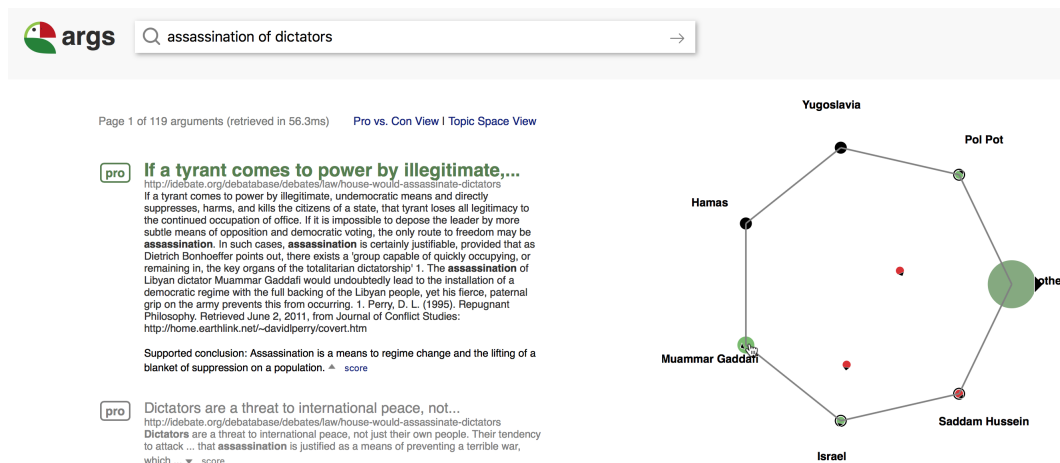

Figure 8: The single filtered args.me search result on the aspect *Muammar Gaddafi* for the query "assassination of dictators". The filtering is the result of clicking in the respective argument glyph in the topic space visualization.

feminism. Without the visualization, these insights would have been hard to gain; the two respective arguments were ranked at position #43 and #46.

5.2 Search Refinement

As a second example, we assume that a user looks for new arguments on the "assassination of dictators", for which args.me provides 119 results. If the user wants to refine a search to restrict it to a specific aspect of the topic only (e.g., to arguments covering *Muammar Gaddafi*), a simple click on the respective argument glyph in the topic space visualization suffices, as illustrated in Figure 8. The associated arguments are filtered and placed at the top of the result list (only one argument in the illustrated case). With the existing interaction methods

of args.me, the argument text can be extended and its source web page shows up after clicking on it. In the old overall ranking view, the argument would have been ranked at position #34.

6 Conclusion

This paper has presented a new visual interface for our argument search engine, args.me. The visualization supports users in learning about the topical aspects covered by the arguments returned in response to a query. In two case studies, we have demonstrated how this topic space visualization enables an efficient exploration and refinement of argument search results. Future research on modeling the aspects of an argument can further enhance the usability of the visualization.

References

Peter Baldwin and David Price. 2018. Debategraph. https://debategraph.org.

Pietro Baroni, Marco Romano, Francesca Toni, Marco Aurisicchio, and Giorgio Bertanza. 2015. Automatic evaluation of design alternatives with quantitative argumentation. *Argument & Computation*, 6(1):24–49.

Gregor Betz, Sebastian Cacean, and Christian Voigt. 2018. Argunet. http://www.argunet.org.

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022.

Jeff Conklin and Michael L. Begeman. 1988. gIBIS: A hypertext tool for exploratory policy discussion. *ACM Transactions on Information Systems*, 6(4):303–331.

Bruce Croft, Donald Metzler, and Trevor Strohman. 2009. *Search Engines: Information Retrieval in Practice*, 1st edition. Addison-Wesley Publishing Company, USA.

Mennatallah El-Assady, Valentin Gold, Carmela Acevedo, Christopher Collins, and Daniel Keim. 2016. ConToVi: Multi-party conversation exploration using topic-space views. *Computer Graphics Forum*.

Mennatallah El-Assady, Rita Sevastjanova, Bela Gipp, Daniel Keim, and Christopher Collins. 2017. NEREx: Named-entity relationship exploration in multi-party conversations. *Computer Graphics Forum*, 36(3):213–225.

Tim van Gelder. 2007. The rationale for rationale. *Law, Probability & Risk*, 6(1-4):23–42.

Valentin Gold, Christian Rohrdantz, and Mennatallah El-Assady. 2015. Exploratory text analysis using lexical episode plots. In *Eurographics Conference on Visualization (EuroVis) - Short Papers*. The Eurographics Association.

Wolfgang Jentner, Mennatallah El-Assady, Bela Gipp, and Daniel A. Keim. 2017. Feature alignment for the analysis of verbatim text transcripts. *EuroVis Workshop on Visual Analytics (EuroVA)*.

Areti Karamanou, Nikolaos Loutas, and Konstantinos Tarabanis. 2011. ArgVis: Structuring political deliberations using innovative visualisation technologies. *Electronic Participation*, pages 87–98.

Christian Kock. 2007. Dialectical obligations in political debate. *Informal Logic*, 27(3):233–247.

John D. Lowrance, Ian W. Harrison, and Andres C. Rodriguez. 2000. Structured argumentation for analysis. In *Proceedings of the 12th International Conference on Systems Research, Informatics, and Cybernetics: Focus Symposia on Advances in Computer-Based and Web-Based Collaborative Systems*, pages 47–57.

Mark Meyer, Alan Barr, Haeyoung Lee, and Mathieu Desbrun. 2002. Generalized barycentric coordinates on irregular polygons. *Journal of Graphics Tools*, 7(1):13–22.

Tamara Munzner. 2014. *Visualization Analysis and Design*. AK Peters Visualization Series. CRC Press.

Hui Niu. 2016. *Pedagogical Efficacy of Argument Visualization Tools*. Ph.D. thesis, Education: Faculty of Education.

Jörg Pührer. 2017. ArgueApply: A mobile app for argumentation. In *International Conference on Logic Programming and Nonmonotonic Reasoning*, pages 250–262. Springer.

Chris Reed and Glenn Rowe. 2004. Araucaria: Software for argument analysis, diagramming and representation. *International Journal on Artificial Intelligence Tools*, 13(04):961–979.

Patrick Riehmann, Dora Kiesel, Martin Kohlhaas, and Bend Fröhlich. 2018. Visualizing a thinker's life. *IEEE Transactions on Visualization and Computer Graphics*.

Stephen Robertson and Hugo Zaragoza. 2009. The probabilistic relevance framework: BM25 and beyond. *Foundations and Trends in Information Retrieval*, 3(4):333–389.

Daniel Suthers, Arlene Weiner, John Connelly, and Massimo Paolucci. 1995. Belvedere: Engaging students in critical discussion of science and public policy issues. In *Proceedings of the 7th World Conference on Artificial Intelligence in Education*, pages 266–273.

Truthmapping. 2018. Truthmapping. https://www.truthmapping.com.

Manolis Tzagarakis and Nikos Karacapilidis. 2013. On the exploitation of semantic types in the visualization of complex argumentative discourses. In *Proceedings of the 2nd International Workshop on Intelligent Exploration of Semantic Data*, page 3. ACM.

Henning Wachsmuth, Giovanni Da San Martino, Dora Kiesel, and Benno Stein. 2017a. The impact of modeling overall argumentation with tree kernels. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2379–2389. Association for Computational Linguistics.

Henning Wachsmuth, Martin Potthast, Khalid Al Khatib, Yamen Ajjour, Jana Puschmann, Jiani Qu, Jonas Dorsch, Viorel Morari, Janek Bevendorff, and Benno Stein. 2017b. Building an argument search engine for the web. In *Proceedings of the 4th Workshop on Argument Mining*, pages 49–59. Association for Computational Linguistics.