

Common Round: Application of Language Technologies to Large-Scale Web Debates

Hans Uszkoreit, Aleksandra Gabryszak, Leonhard Hennig, Jörg Steffen,
Renlong Ai, Stephan Busemann, Jon Dehdari, Josef van Genabith,
Georg Heigold, Nils Rethmeier, Raphael Rubino, Sven Schmeier,
Philippe Thomas, He Wang, Feiyu Xu
Deutsches Forschungszentrum für Künstliche Intelligenz, Germany
firstname.lastname@dfki.de

Abstract

Web debates play an important role in enabling broad participation of constituencies in social, political and economic decision-taking. However, it is challenging to organize, structure, and navigate a vast number of diverse argumentations and comments collected from many participants over a long time period. In this paper we demonstrate *Common Round*, a next generation platform for large-scale web debates, which provides functions for eliciting the semantic content and structures from the contributions of participants. In particular, *Common Round* applies language technologies for the extraction of semantic essence from textual input, aggregation of the formulated opinions and arguments. The platform also provides a cross-lingual access to debates using machine translation.

1 Introduction

Debating is a very useful approach to individual and collective decision making; it helps the formation of ideas and policies in democracies. Web-based debates allow for reaching a wider audience, therefore bringing more arguments and diverse perspectives on a debate topic compared to face-to-face discussions. The asynchronous mode of web debates also enables participants to explore debate content more thoroughly (Salter et al., 2016). However, these advantages often get lost in a large-scale debate since its content can become unmanageable. Moreover, missing background knowledge on debate topics or language barriers for international topics can prohibit users from a proper understanding of debate content.

The majority of the existing web discussion

platforms offer the following participation model: users can formulate a discussion topic, share their opinions on that topic and respond to others in the form of unstructured posts. However, they do not offer effective functionalities for 1) easy access to the argumentative structure of debate content, and 2) quick overviews of the various semantic facets, the polarity and the relevance of the arguments. Some platforms¹ allow users to label posts as *pro* or *con* arguments, to cite external sources, to assess debate content or to create structured debates across the web, but do not offer any deeper automatic language technology-based analyses. Argumentation mining research, which could help in automatically structuring debates, has only recently been applied to web debate corpora (Boltuzic and Snajder, 2015; Petasis and Karkaletsis, 2016; Egan et al., 2016).

Our goal is to address these issues by developing a debate platform that:

- Supports debate participants in making substantial and clear contributions
- Facilitates an overview of debate contents
- Associates relevant information available on the web with debate topics
- Connects regional discussions to global deliberations
- Supports advanced participation in deliberations, without sacrificing transparency and usability.

In this paper, we present the *Common Round* platform, which implements various functions towards these goals, with following contributions: (a) we describe the architecture of the platform (Section 2), including a model for argumentative dialogues (Section 3) and (b) we present our implementation of several language technologies for supporting collective deliberation (Sections 4–6).

¹Examples: ProCon.org, OpenPetition.de, Arvina and ArguBlogging (Bex et al., 2013)

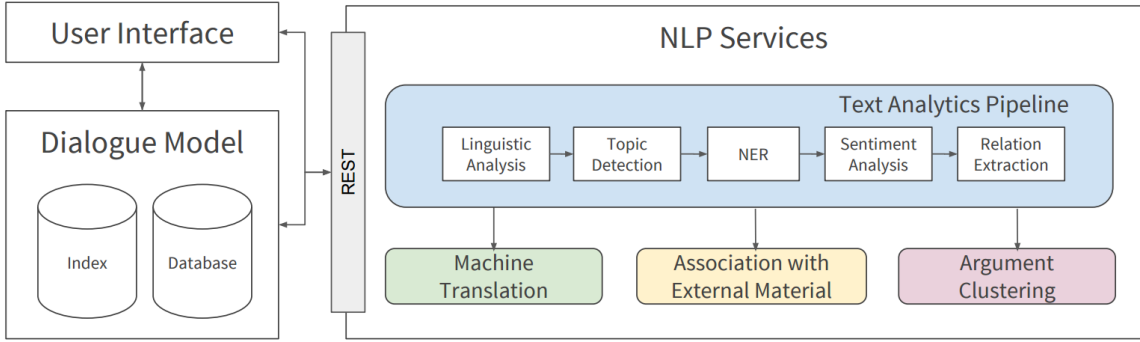


Figure 1: Architecture of Common Round.

2 Common Round platform - Overview

Figure 1 depicts the platform architecture. Common Round provides a top-level dialogue model that supports users to make decisions and enter labeled contributions at appropriate places in the model. This keeps debates organized and transparent, and at the same time allows users to participate in a free and unrestricted discussion. Furthermore a predefined coarse-grained debate structure facilitates the development of automatic text analysis algorithms for deliberation support, since it enables the system to integrate domain and context knowledge for better interpretation of user intentions. Within our platform, textual debate content is automatically further structured and enhanced employing argument mining, text analytics, retrieval of external material and machine translation. The results of these analyses are fed back to the front-end to give users a good overview of discussion content, and to enable a structured, semantic search for additional evidences and background material. We will exemplify the features of our platform with four selected topics *Should Cannabis be legalized?*, *Are German cars better than those from Japan?*, *Should we use wind power?*, *How to address online fake news?*

3 Dialogue model and content assessments

The Common Round platform provides a structure that aims to cover the most essential aspects of argumentative dialogues (Figure 2).

The top level defines the semantic categories of the debate questions. Given the categories, the questions themselves can be posted as the next level. There are two types of debate questions: (a) yes-no questions, and (b) multi-proposal questions (Table 1). A question reflects the major claim

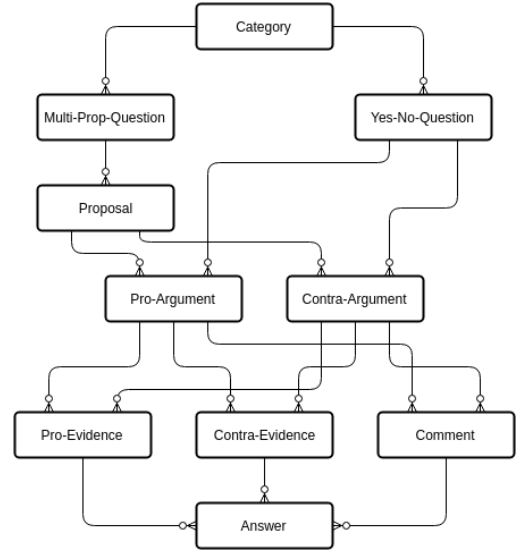


Figure 2: Dialogue model of Common Round.

for the debate. The subsequent level is reserved for

question type	example
Yes-No	Should Cannabis be legalized?
Multi-proposal	How to address online fake news?
Proposal 1	Impose sanctions on fake news authors.
Proposal 2	Impose sanctions on website providers.

Table 1: Yes-no and multi-proposal questions.

arguments. An argument can either be a *pro*- or a *con*-argument. For the yes-no questions, an argument directly refers to a *yes* answer to the question, while for multi-proposal questions an argument refers to a single proposal. At the next level, a user can add evidence in favor of or against an argument. Such evidence could be a textual contribution or a link to external sources. Finally, a user can freely answer to either an evidence or a comment or just refer to a previously posted answer. The length of the contributions is restricted in order to enforce separated, focused posts instead of

lengthy and noisy information.

Besides the dialogue structure, assessing postings is an important aspect in the Common Round forum. The assessment reflects the community view of a main debate claim, an argument or evidence. The degrees of consent a user can express are based on the common 5-level Likert-type scale (Likert, 1932). Based on the ratings a first time viewer of a question can get a quick overview of the state of the discussion.

4 Argument Mining

The dialogue model enables users to distinguish between argumentative and non-argumentative parts of a debate. Users can explore the components of argumentations, i.e. major claims, arguments and evidences. The categorization in *pro/con* arguments and *pro/con* evidences easily allows finding posts supporting or refuting the major claim. Furthermore, similar arguments are automatically clustered. To aggregate similar arguments we represent posts with Paragraph2Vec (Le and Mikolov, 2014) and then compute clusters from the cosine similarity between post vectors using agglomerative clustering. In order to keep the clustering unsupervised we automatically discover the number of clusters using the Silhouette Coefficient (Rousseeuw, 1987). To help users identify the most relevant content within a cluster, the arguments are sorted by their assessed validity. Reading only a few arguments from each cluster enables users to gain a quick overview of different argument facets. For example, in a debate on cannabis legalization, arguments from a medical point of view can be grouped separately from arguments that focus on economic reasons. This functionality is an important contribution to structuring large-scale debates.

5 Text Analytics and Association with External Material

The Common Round platform enriches the contents of debates and posts by extracting information about topics, sentiment, entities and relations. Topic detection helps to find semantically related debates. Sentiment analysis allows determining the emotion level in discussions. By identifying, for example, instances of the relation *MayTreatDisorder* in a discussion about the legalization of cannabis, we can aggregate references to the same or similar medical facts introduced by different de-

bate participants. Additionally, our platform incorporates a web search to enable finding supporting evidence in external knowledge sources. Table 2 shows an example of the text analytics results and the association with external material.

<p>Cannabis is a good way to reduce pain</p>	
NER and RE	
Sentiment	positive
External Material	www.scientificamerican.com/article/could-medical-cannabis-break-the-painkiller-epidemic

Table 2: Example of the text analytics results.

Information extraction pipeline currently supports English and German texts. We utilize Stanford CoreNLP tools for segmentation, POS-tagging and dependency parsing of English texts. POS-tagging and dependency parsing of German texts are realized with Mate Tools. The topic detection is realized by the supervised document classification module PCL (Schmeier, 2013). PCL is proven to be very robust and accurate especially for short texts and unbalanced training data. For named entity recognition, we employ two complementary approaches. We apply Stanford NER models to recognize standard entity types, such as persons, organizations, locations and date/time expressions. For non-standard concept types, we use SProUT (Drozdzyński et al., 2004), which implements a regular expression-like rule formalism and gazetteers for detecting domain-specific concepts in text. Relation extraction is performed by matching dependency parse trees of sentences to a set of automatically learned dependency patterns (Krause et al., 2012). Relation patterns are learned from corpora manually annotated with event type, argument types, and roles, or using distant supervision (Mintz et al., 2009). For sentiment analysis, we apply a lexicon-based approach that additionally makes use of syntactic information in order to handle negation.

6 Cross-lingual Access

In order to support cross-lingual access to debate posts we developed a new character-based neural machine translation (NMT) engine and tuned on Common Round domains. Translation is available as a real-time service and translation outputs

are marked. Our NMT is based on an encoder-decoder with attention design (Bahdanau et al., 2014; Johnson et al., 2016), using bidirectional LSTM layers for encoding and unidirectional layers for decoding. We employ a character-based approach to better cope with rich morphology and OOVs in Common Round user posts. We train on English-German data from WMT-16² and use transfer learning on 30K crawled noisy in-domain segments for Common Round domain-adaptation. As our NMT system is new, we compare with s-o-t-a PBMT: while both perform in the top-ranks in WMT-16 (Bojar et al., 2016), NMT is better on Common Round data (NMT from 29.1 BLEU on WMT-16 data to 22.9 car and 23.8 wind power, against 30.0 to 13.6 and 17.7 for PBMT) and better adapts using the supplementary crawled data (25.8 car and 28.5 wind against 14.1 and 19.3).³

7 Conclusion & Future Work

We presented Common Round, a new type of web-based debate platform supporting large-scale decision making. The dialogue model of the platform supports users in making clear and substantial debate contributions by labeling their posts as *pro/con* arguments, *pro/con* evidences, comments and answers. The user input is automatically analyzed and enhanced using language technology such as argument mining, text analytics, retrieval of external material and machine translation. The analysis results are fed back to the user interface as an additional information layer. For future work we suggest a more fine-grained automatic text analysis, for example by detecting claims and premises of arguments as well as types and validity of evidences. Moreover we will conduct a thorough evaluation of the system.

Acknowledgments

This research was supported by the German Federal Ministry of Education and Research (BMBF) through the projects ALL SIDES (01IW14002) and BBDC (01IS14013E), by the German Federal Ministry of Economics and Energy (BMWi) through the projects SDW (01MD15010A) and SD4M (01MD15007B), and by the European Union’s Horizon 2020 research and innovation programme through the project QT21 (645452).

²<http://www.statmt.org/wmt16/translation-task.html>

³All EN→DE, similar for reverse.

References

- D. Bahdanau, K. Cho, and Y. Bengio. 2014. Neural machine translation by jointly learning to align and translate. *CoRR*, abs/1409.0473.
- F. Bex, J. Lawrence, M. Snaith, and C. Reed. 2013. Implementing the argument web. *Commun. ACM*, 56:66–73.
- O. Bojar, R. Chatterjee, C. Federmann, Y. Graham, B. Haddow, M. Huck, A. Jimeno Yepes, P. Koehn, V. Logacheva, C. Monz, M. Negri, A. Neveol, M. Neves, M. Popel, M. Post, R. Rubino, C. Scarton, L. Specia, M. Turchi, K. Verspoor, and M. Zampieri. 2016. Findings of the 2016 Conference on Machine Translation. In *Proc. of MT*.
- F. Boltuzic and J. Snajder. 2015. Identifying prominent arguments in online debates using semantic textual similarity. In *ArgMining@HLT-NAACL*.
- W. Drozdowski, H. Krieger, J. Piskorski, U. Schäfer, and F. Xu. 2004. Shallow processing with unification and typed feature structures — foundations and applications. *Künstliche Intelligenz*, 1:17–23.
- C. Egan, A. Siddharthan, and A. Z. Wyner. 2016. Summarising the points made in online political debates. In *ArgMining@ACL*.
- M. Johnson, M. Schuster, Q. V. Le, M. Krikun, Y. Wu, Z. Chen, N. Thorat, F. Viégas, M. Wattenberg, G. Corrado, M. Hughes, and J. Dean. 2016. Google’s multilingual neural machine translation system: Enabling zero-shot translation. In <https://arxiv.org/abs/1611.04558>.
- S. Krause, H. Li, H. Uszkoreit, and F. Xu. 2012. Large-Scale Learning of Relation-Extraction Rules with Distant Supervision from the Web. In *ISWC*.
- Q. V. Le and T. Mikolov. 2014. Distributed representations of sentences and documents. In *ICML*.
- R. Likert. 1932. A technique for the measurement of attitudes. *Archives of Psychology*, 22(140):1–55.
- M. Mintz, S. Bills, R. Snow, and D. Jurafsky. 2009. Distant Supervision for Relation Extraction Without Labeled Data. In *IJCNLP*.
- G. Petasis and V. Karkaletsis. 2016. Identifying argument components through TextRank. In *ArgMining@ACL*.
- P. J. Rousseeuw. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.*, 20:53–65.
- S. Salter, T. Douglas, and D. Kember. 2016. Comparing face-to-face and asynchronous online communication as mechanisms for critical reflective dialogue. *JEAR*, 0(0):1–16.
- S. Schmeier. 2013. *Exploratory Search on Mobile Devices*. DFKI-LT - Dissertation Series, Saarbruecken, Germany.