SIGIR 2016 Workshop WebQA II: Web Question Answering Beyond Factoids

Alessandro Moschitti, Lluís Màrquez, Preslav Nakov
Qatar Computing Research Institute, HBKU, Qatar
{amoschitti, lmarquez, pnakov}@qf.org.qa

Eugene Agichtein
Emory University, Atlanta, USA
eugene.agichtein@emory.edu

Charles Clarke
University of Waterloo, Canada
claclark@gmail.com

Idan Szpektor
Yahoo Research, Haifa, Israel
idan@yahoo-inc.com

1. OVERVIEW

Web search engines have made great progress in answering factoid queries, such as

How many people live in Australia?

They can provide a succinct answer, up to a few words in length, and can sometimes offer additional information such as related facts or entities. However, Web search engines are currently not well-tailored for managing more complex questions, especially when they require explanation or description, e.g.,

Can I get a Qatar residence permit for my wife while she is currently in Doha with a tourist visa?

Given a question like this, currently, search engines resort to returning a link to a detailed Web document, which does not make sure the user can find an answer. Alternatively, such a question might be posted on a Community Question Answering (CQA) site, e.g., Qatar Living,[1] hoping to get a human-authored and detailed response. Other questions submitted on the Web can be short and ambiguous (such as Web queries to a search engine). These issues make the WebQA task more challenging than traditional question answering, and finding the most effective approaches for it remains an open question.

This workshop is a second edition of the successful WebQA workshop, which was held at SIGIR'2015 [1]. The new edition continues the exploration of the boundaries of Web question answering for better understanding the spectrum of approaches and possible responses that are more detailed than a short fact, yet are more useful than a full document. In particular, we also focus on methods that can handle complex questions involving the interdependencies between different entities and facts.

[1]http://www.qatarliving.com/forum

SIGIR '16 July 17-21, 2016, Pisa, Italy

ACM ISBN 978-1-4503-4069-4/16/07…$15.00

DOI: http://dx.doi.org/10.1145/2911451.2917767

2. THEME AND GOAL

This workshop aims to explore diverse approaches to answering questions on the Web. Unlike the more formal format of conferences, the aim of this workshop is to bring together researchers in diverse areas working on this problem, including those from natural language processing, information retrieval, social media and recommender systems communities. Yet, it is designed to be of interest for the SIGIR audience. However, due to its format, its goal, as compared to the main conference, is to conduct a more focused and open discussion. Both academic and industrial stakeholders were welcome to the workshop, including keynotes and invited speakers. In particular, we encouraged the discussion of ongoing research and late–breaking, preliminary results on the following topics:

- Social and user generated question analysis
- Identifying question intent in Web queries
- Answer aggregation from various sources
- Answer summarization from various sources
- Using collaboratively generated content for QA
- Inferring answers for Web questions using knowledge bases or graphs
- Answering opinion questions, including sentiment analysis
- Answering complex, multi-sentence questions
- Evaluation of question answering systems (e.g., via crowd-sourcing)

3. RELEVANCE TO SIGIR

Web question answering is a central topic in information retrieval. However, answering Web queries poses challenges harder than in traditional Web search, as the expectation of the user is much higher when the system promises an answer, not just 10 blue links to click.

While an active area of research exists, this workshop provides a more focused forum addressing specific challenges such as query understanding, answer extraction, and source selection, that are more critical to Web QA compared to traditional web search.

We believe that a continuation of the first WebQA workshop is important as many relevant publications have continued to be disseminated across conferences. See for example, [6, 12, 13, 17, 18, 19, 20, 21, 23], also given the advent of neural networks, e.g., [5, 8, 9, 22, 24]. WebQA offers a forum to researchers and practitioners to discuss and possibly collaborate, thus helping advance the state of the art.

This workshop also coincides and complements the LiveQA track at TREC 2015 and 2016 [2]. LiveQA is a revival of the TREC Question Answering track. The track provides a challenge and data for answering *real user questions*, posted live to the Yahoo! Answers site. The TREC 2016 LiveQA challenge evaluation was held around the time of the workshop, which would provide both the active and the potential participants a way to discuss ideas and approaches.

Another relevant activity is the challenge on CQA organized in 2015 and 2016 at SemEval, i.e., Task 3 [14, 15], which focused on answering new questions using a CQA forum (Qatar Living). In particular, participants were asked to rerank the results returned by a search engine, and in addition to select the good answers from a community forum (see for example the systems developed in [3, 10, 16]). Additionally, the challenge proposed a question-question similarity and an answer selection subtasks (see e.g., [4, 7, 11]).

The WebQA II workshop was held shortly after SemEval 2016, and thus it allowed the participants to discuss its outcome as well as further ideas in more detail. It is valuable for two more reasons: (*i*) there is still a lot of disagreement regarding the goals and the nature of Web question answering, mostly related to the *question intent* (what kind of queries benefit from question answering compared to other methods); and (*ii*) leading search engines are eager to provide question answering services, especially for mobile devices.

4. REFERENCES

[1] E. Agichtein, D. Carmel, C. L. Clarke, P. Paritosh, D. Pelleg, and I. Szpektor. Web question answering: Beyond factoids: SIGIR 2015 workshop. In *SIGIR*, 2015.

[2] E. Agichtein, D. Carmel, D. Harman, D. Pelleg, and Y. Pinter. Overview of the TREC 2015 LiveQA track. In *TREC*, 2015.

[3] A. Barrón-Cedeño, S. Filice, G. Da San Martino, S. Joty, L. Màrquez, P. Nakov, and A. Moschitti. Thread-level information for comment classification in community question answering. In *ACL*, 2015.

[4] A. Barrón-Cedeño, G. D. S. Martino, S. Joty, A. Moschitti, F. A. A. Obaidli, S. Romeo, K. Tymoshenko, and A. Uva. ConvKN at SemEval-2016 Task 3: Answer and question selection for question answering on Arabic and English fora. In *SemEval*, 2016.

[5] A. Bordes, S. Chopra, and J. Weston. Question answering with subgraph embeddings. In *EMNLP*, 2014.

[6] L. Braunstain, O. Kurland, D. Carmel, I. Szpektor, and A. Shtok. Supporting human answers for advice-seeking questions in CQA sites. In *ECIR*, 2016.

[7] S. Filice, D. Croce, A. Moschitti, and R. Basili. KeLP at SemEval-2016 Task 3: Learning semantic relations between questions and answers. In *SemEval*, 2016.

[8] F. Guzmán, L. Màrquez, and P. Nakov. Machine translation evaluation meets community question answering. In *ACL*, 2016.

[9] M. Iyyer, J. Boyd-Graber, L. Claudino, R. Socher, and H. Daumé III. A neural network for factoid question answering over paragraphs. In *EMNLP*, 2014.

[10] S. Joty, A. Barrón-Cedeño, G. Da San Martino, S. Filice, L. Màrquez, A. Moschitti, and P. Nakov. Global thread-level inference for comment classification in community question answering. In *EMNLP*, 2015.

[11] T. Lei, H. Joshi, R. Barzilay, T. S. Jaakkola, K. Tymoshenko, A. Moschitti, and L. Màrquez. Semi-supervised question retrieval with gated convolutions. In *NAACL-HLT*, 2016.

[12] Q. Liu, E. Agichtein, G. Dror, Y. Maarek, and I. Szpektor. When web search fails, searchers become askers: Understanding the transition. In *SIGIR*, 2012.

[13] Q. Liu, T. Jurczyk, J. Choi, and E. Agichtein. Real-time community question answering: Exploring content recommendation and user notification strategies. In *IUI*, 2015.

[14] P. Nakov, L. Màrquez, W. Magdy, A. Moschitti, J. Glass, and B. Randeree. SemEval-2015 task 3: Answer selection in community question answering. In *SemEval*, 2015.

[15] P. Nakov, L. Màrquez, A. Moschitti, W. Magdy, H. Mubarak, A. A. Freihat, J. Glass, and B. Randeree. SemEval-2016 task 3: Community question answering. In *SemEval*, 2016.

[16] M. Nicosia, S. Filice, A. Barrón-Cedeño, I. Saleh, H. Mubarak, W. Gao, P. Nakov, G. Da San Martino, A. Moschitti, K. Darwish, L. Màrquez, S. Joty, and W. Magdy. QCRI: Answer selection for community question answering - experiments for Arabic and English. In *SemEval*, 2015.

[17] A. Omari, D. Carmel, O. Rokhlenko, and I. Szpektor. Novelty based ranking of human answers for community questions. In *TREC*, 2016.

[18] B. Petersil, A. Mejer, I. Szpektor, and K. Crammer. That's not my question: Learning to weight unmatched terms in CQA vertical search. In *SIGIR*, 2016.

[19] A. Severyn and A. Moschitti. Structural relationships for large-scale learning of answer re-ranking. In *SIGIR*, 2012.

[20] A. Severyn and A. Moschitti. Learning to rank short text pairs with convolutional deep neural networks. In *SIGIR*, 2015.

[21] G. Tsur, Y. Pinter, I. Szpektor, and D. Carmel. Identifying web queries with question intent. In *WWW*, 2016.

[22] K. Tymoshenko, D. Bonadiman, and A. Moschitti. Convolutional neural networks vs. convolution kernels: Feature engineering for answer sentence reranking. In *NAACL-HLT*, 2016.

[23] K. Tymoshenko and A. Moschitti. Assessing the impact of syntactic and semantic structures for answer passages reranking. In *CIKM*, 2015.

[24] W.-T. Yih, M.-W. Chang, C. Meek, and A. Pastusiak. Question answering using enhanced lexical semantic models. In *ACL*, 2013.