ACL 2016

The 4th BioASQ Workshop
A challenge on large-scale biomedical semantic indexing and question answering

Proceedings of the Workshop

August 13, 2016
Berlin, Germany

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-945626-22-7
ACL anthology W16-31xx

Preface

The fourth BioASQ workshop on biomedical semantic indexing and question answering took place in Berlin, Germany on August 13th, 2016 and was hosted by the Humboldt University. The workshop was supported by the BioASQ project[1], which organizes the corresponding annual challenge. The goals of the workshop were to present the results of the fourth BioASQ challenge and further the interaction with the wider community of biomedical semantic indexing and question answering. The presenters represented research teams from different parts of the globe and with different viewpoints to the problem. This contributed to a very lively and interesting discussion among the participants of the workshop. Six papers were presented during the workshop. All were selected by peer review for presentation. This volume includes 7 papers and one abstract: The first paper gives an overview of the challenge, including especially the datasets that were used throughout the challenges and the overall results achieved by the participants.
The remaining six papers are those presented at the workshop. The first of these papers is a new extension of the MTI system. In particular, Learning to Rank methodology is used as a boosting component of the MTI system. The second paper present a system which includes several multi-label classifiers (MLC) that are combined in ensembles. Elastic-Search for indexing is the object of discourse of the third workshop paper. The fourth paper presents a system which uses TmTool in addition to MetaMap, to identify possible biomedical named entities, especially out-of-vocabulary concepts. They also introduced a unified classification interface for judging the relevance of each retrieved concept, document, and snippet, which can combine the relevant scores evidenced by various sources. The system presented in the fifth paper relies on the Hana Database for text processing. It uses the Stanford CoreNLP package for tokenizing the questions. Each of the tokens is then sent to the BioPortal and to the Hana database for concept retrieval. The concepts retrieved from the two stores are finally merged to a single list that is used to retrieve relevant text passages from the documents at hand. The last paper focuses on the retrieval of relevant documents and snippets. The proposed system uses a cluster-based language model. Then, it reranks the retrieved top-n sentences using five independent similarity models based on shallow semantic analysis.

Finally, the proceedings also include the abstract of one paper that was presented in the poster session only, which describes an approach for extending the web services in order to retrieve the relevant documents, concepts, snippets and triples for the question-answering task.

We wish to thank all who participated to the success of this workshop, especially the authors, reviewers, speakers and participants.

Ioannis A. Kakadiaris, George Paliouras and Anastasia Krithara
August 2016

[1] http://www.bioasq.org

Organizers:

Ioannis A. Kakadiaris, University of Houston, USA
George Paliouras, NCSR "Demokritos", Greece and University of Houston, USA
Anastasia Krithara, NCSR "Demokritos", Greece

Program Committee:

Ion Androutsopoulos, Athens University of Economics and Business, Greece
Nicolas Baskiotis, Université Pierre et Marie Curie, France
Dimitris Galanis, National Technical University of Athens, Greece
Brigitte Grau, LIMSI/CNRS, France
Aris Kosmopoulos, NCSR “Demokritos”, Greece
Zhiyong Lu, National Library of Medicine, USA
Prodromos Malakasiotis, Athens University of Economics and Business, Greece
Jim Mork, National Library of Medicine, USA
Diego Molla, Macquarie University, Australia
Henning Müller, University of Applied Sciences, Switzerland
Claire Nedellec, INRA, France
Mariana Neves, University of Potsdam, Germany
Harris Papageorgiou, ILSP, Greece
Ioannis Partalas, Viseo group, France
John Prager, Thomas J. Watson Research Center, IBM, USA
Francisco J. Ribadas-Pena, University of Vigo, Spain
Hagit Shatkay, University of Delaware, USA
Grigoris Tsoumakas, Aristotle University of Thessaloniki, Greece
Christina Unger, Bielefeld University, Germany
Ellen Voorhees, National Institute of Standards and Technology, USA

Invited Speaker:

Sherri Matis-Mitchell, independent consultant for Text, Data and Social Media Analytics at Data Star Insights

Table of Contents

Conference Program

9:00-9:15	**Welcome**
9:15-10:15	**Invited speaker: Sherri Matis-Mitchell** "Solving Problems and Supporting Decisions in Pharma R& D using Text Analytics: A Recent History"
10:15-10:30	*Results of the 4th edition of BioASQ Challenge* Anastasia Krithara, Anastasios Nentidis, Georgios Paliouras and Ioannis Kakadiaris
10:30-11:00	**Coffee break**
11:00-12:30	**BioASQ participant session**
11:00-11:15	*Using Learning-To-Rank to Enhance NLM Medical Text Indexer Results* Ilya Zavorin, James Mork and Dina Demner-Fushman
11:15-11:30	*LABDA at the 2016 BioASQ challenge task 4a: Semantic Indexing by using ElasticSearch* Isabel Segura-Bedmar, Adrián Carruana and Paloma Martínez
11:30-11:45	*Learning to Answer Biomedical Questions: OAQA at BioASQ 4B* Zi Yang, Yue Zhou and Eric Nyberg
11:45-12:00	*HPI Question Answering System in BioASQ 2016* Frederik Schulze, Ricarda Schüler, Tim Draeger, Daniel Dummer, Alexander Ernst, Pedro Flemming, Cindy Perscheid and Mariana Neves
12:00-12:15	*KSAnswer: Question-answering System of Kangwon National University and Sogang University in the 2016 BioASQ Challenge* Hyeon-gu Lee, Minkyoung Kim, Harksoo Kim, Juae Kim, Sunjae Kwon, Jungyun Seo, Yi-Reun Kim and Jung-Kyu Choi
12:15-12:30	*Large-Scale Semantic Indexing and Question Answering in Biomedicine* Eirini Papagiannopoulou, Yiannis Papanikolaou, Dimitris Dimitriadis, Sakis Lagopoulos, Grigorios Tsoumakas, Manos Laliotis, Nikos Markantonatos and Ioannis Vlahavas
Poster	*WS4A: a Biomedical Question and Answering System based on public web services and ontologies* Miguel J. Rodriques, Miguel Fale and Francisco M. Couto