Eighth Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR'15)

CIKM 2015 Workshop

Krisztian Balog
University of Stavanger
Norway
krisztian.balog@uis.no

Jeffrey Dalton
Google Research
jeffdalton@google.com

Antoine Doucet
University of La Rochelle
L3i Laboratory
France
antoine.doucet@univ-lr.fr

Yusra Ibrahim
Max Planck Institute for
Informatics
Saarbrücken, Germany
yibrahim@mpi-inf.mpg.de

ABSTRACT

The amount of structured content published on the Web has been growing rapidly, making it possible to address increasingly complex information access tasks. Recent years have witnessed the emergence of large scale human-curated knowledge bases as well as a growing array of techniques that identify or extract information automatically from unstructured and semi-structured sources. The ESAIR workshop series aims to advance the general research agenda on the problem of creating and exploiting semantic annotations. The eighth edition of ESAIR sets its focus on applications. We dedicate a special "annotations in action" track to demonstrations that showcase innovative prototype systems, in addition to the regular research and position paper contributions. The workshop also features invited talks from leaders in the field. The desired outcome of ESAIR'15 is a roadmap and research agenda that guides academic efforts and aligns them with industrial directions and developments.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: H.3.1 Content Analysis and Indexing; H.3.3 Information Search and Retrieval; H.3.4 Systems and Software

Keywords

Semantic annotations; knowledge bases; complex search tasks

1. OVERVIEW

Over the recent years, major search engines have redesigned their search results to accommodate semantic information, including rich search results and direct question answering. Recently, web search engines all developed large-scale entity databases, including Google's Knowledge Graph, Yahoo!'s Web Of Objects, and Bing's Satori. At the heart of these efforts is large-scale semantic annotation of both queries and documents.

The goal of the ESAIR workshop series is to create a forum for researchers interested in the application of semantic annotations for information access tasks. By semantic annotations we refer to linguistic annotations (such as named entities, semantic classes or roles, etc.) as well as user annotations (such as micro-formats, RDF, tags, etc.). ESAIR'15 continues, with a renewed group of organizers, on the path set by the previous edition(s) of the workshop: clarifying the exact role of semantic annotations in supporting complex search tasks.

Annotations come in a variety of flavors (named entities, temporal information, geo-positional markers, semantic roles, sentiment, etc.) and there is a growing repertoire of tools and techniques available for extracting these annotations automatically from text. The question then presents itself: what, if anything, is missing? We seek to answer this question by focusing on applications that are rooted in specific, real-world use cases. To further sharpen this application-oriented focus, the workshop features a dedicated "annotations in action" track for presenting innovative prototype systems. A Best Demonstration Award, sponsored by Google, is presented to the participant(s) with the most outstanding demo at the workshop.

2. PROGRAMME AND ORGANIZATION

The workshop called for regular papers (4+ pages), position papers (2+1 pages), and demo papers (4 papers), which were each reviewed by at least 3 members of the program committee. The committee was keen on selecting papers that introduce novel ideas and will stimulate discussion in

the workshop. The accepted contributions cover a wide spectrum of topics related to semantic annotations and information retrieval, including: healthcare, automatic translation, semantic search, question answering, knowledge profiling, topic modeling, contextual enrichment, named entity disambiguation, temporal tagging, and semantic annotation of video content. Out of the 19 submissions, the committee accepted a total of 10 papers (5 regular, 2 position, and 3 demo papers) to be presented at the workshop. In addition to the regular paper presentations and demonstration session, the programme includes invited talks by experts in the field, which will focus on the uprising challenges in semantic information access and retrieval. The workshop is intended to be highly interactive where everyone contributes and is planned to inspire the attendees to think "outside the box." Thus, the workshop will include focused discussion sessions at which researchers with similar interests can showcase their ideas and concerns. A final wrap-up session will conclude the event with the objective to formulate critical insights into the potential of semantic annotations, the barriers to success, and specific steps to take this research forward. We continue with the tradition of earlier ESAIR editions and organize a social event for further, more informal discussions among workshop participants and other CIKM attendees.

The list of accepted contributions, corresponding resources (slides, links, etc.), and all other outcomes of the workshop are available at http://esair.org.

Acknowledgments