



## UvA-DARE (Digital Academic Repository)

### Foreword

Alonso, O.; Kamps, J.; Karlgren, J.

**Publication date**

2014

**Document Version**

Final published version

**Published in**

ESAIR'14

[Link to publication](#)

**Citation for published version (APA):**

Alonso, O., Kamps, J., & Karlgren, J. (2014). Foreword. In *ESAIR'14: proceedings of the 7th International Workshop on Exploiting Semantic Annotations in Information Retrieval : November 7, 2014, Shanghai, China* (pp. iii). ACM. <http://dl.acm.org/citation.cfm?id=2663712>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.



## UvA-DARE (Digital Academic Repository)

### ESAIR'14: proceedings of the 7th International Workshop on Exploiting Semantic Annotations in Information Retrieval: November 7, 2014, Shanghai, China

Alonso, O.; Kamps, J.; Karlgren, J.

#### Publication date

2014

#### Document Version

Final published version

[Link to publication](#)

#### Citation for published version (APA):

Alonso, O., Kamps, J., & Karlgren, J. (2014). *ESAIR'14: proceedings of the 7th International Workshop on Exploiting Semantic Annotations in Information Retrieval: November 7, 2014, Shanghai, China*. Association for Computing Machinery. <http://dl.acm.org/citation.cfm?id=2663712>

#### General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

November 7, 2014  
Shanghai, China



Association for  
Computing Machinery

*Advancing Computing as a Science & Profession*



# ESAIR'14

Proceedings of the 7th International Workshop on  
**Exploiting Semantic Annotations  
in Information Retrieval**

*Sponsored by:*

**ACM SIGIR and ACM SIGWEB**

*In-cooperation with:*

**CCF**

*Supported by:*

**SAP, Baibu, Google, Alibaba Group, EMC<sup>2</sup>, Yahoo! Labs, Yandex, &  
Fudan University**

*co-located with:*

**CIKM 2014**



**Association for  
Computing Machinery**

*Advancing Computing as a Science & Profession*

**The Association for Computing Machinery  
2 Penn Plaza, Suite 701  
New York, New York 10121-0701**

Copyright © 2014 by the Association for Computing Machinery, Inc. (ACM). Permission to make digital or hard copies of portions of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permission to republish from: [permissions@acm.org](mailto:permissions@acm.org) or Fax +1 (212) 869-0481.

For other copying of articles that carry a code at the bottom of the first or last page, copying is permitted provided that the per-copy fee indicated in the code is paid through [www.copyright.com](http://www.copyright.com).

**Notice to Past Authors of ACM-Published Articles**

ACM intends to create a complete electronic archive of all articles and/or other material previously published by ACM. If you have written a work that has been previously published by ACM in any journal or conference proceedings prior to 1978, or any SIG Newsletter at any time, and you do NOT want this work to appear in the ACM Digital Library, please inform [permissions@acm.org](mailto:permissions@acm.org), stating the title of the work, the author(s), and where and when published.

**ISBN: 978-1-4503-1365-0**

Additional copies may be ordered prepaid from:

**ACM Order Department**

PO Box 30777  
New York, NY 10087-0777, USA

Phone: 1-800-342-6626 (USA and Canada)  
+1-212-626-0500 (Global)  
Fax: +1-212-944-1318  
E-mail: [acmhelp@acm.org](mailto:acmhelp@acm.org)  
Hours of Operation: 8:30 am – 4:30 pm ET

Printed in the USA

# Foreword

These proceedings contain the contributed papers of the *Seventh Workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR 2014)*, held at CIKM 2014 in Shanghai, on November 7, 2014. After successful workshops at ECIR'08 in Glasgow, WSDM'09 in Barcelona, CIKM'10 in Toronto, CIKM'11 in Glasgow, CIKM'12 at Maui, and CIKM'13 in San Francisco, this year's workshop will remain to advance the general research agenda on this core problem, with an explicit focus on one of the most challenging aspects to address in the coming years. The main remaining challenge is on the user's side—the potential of rich document annotations can only be realized if matched by more articulate queries exploiting these powerful retrieval cues—and a more dynamic approach is emerging by exploiting new forms of query autosuggest. How can the query suggestion paradigm be used to encourage searcher to articulate longer queries, with concepts and relations linking their statement of request to existing semantic models? How do entity results and social network data in “graph search” change the classic division between searchers and information and lead to extreme personalization—are you the query? How to leverage transaction logs and recommendation, and how adaptive should we make the system? What are the privacy ramifications and the UX aspects—how to not creep out users?

These and other related questions will be discussed at this open format workshop -- the aim is to provide paths for further research to change the way we understand information access today!

The workshop will consist of three main parts:

1. Two keynotes to help us frame the problem, and create a common understanding of the challenges: Peter Mika (Yahoo Labs) and Silviu-Petru Cucerzan (Microsoft Research).
2. A boaster and poster session with 11 papers selected by the program committee from 15 submissions (a 73% acceptance rate). Each paper was reviewed by at least two members of the program committee.
3. Breakout groups on different aspects of exploiting semantic annotations, with reports being discussed in the final session.

These papers represent the ideas and opinions of the authors who are trying to stimulate debate. It is the combination of these papers and the debate that will make the workshop a success! We thank Google for sponsoring the workshop, which helped us build a program with great keynote speakers and allowed us to continue the tradition of organizing social events after the workshop. We thank the ACM and CIKM for hosting this workshop, and Huan Liu, Xiaofeng Meng, and Lisa Tolles for their outstanding support in the organization. Thanks also go to the program committee, the authors of the papers, and all participants without whom there would be no workshop.

**Omar Alonso**

*ESAIR 2014 Co-chair  
Microsoft*

**Jaap Kamps**

*ESAIR 2014 Co-chair  
University of  
Amsterdam*

**Jussi Karlgren**

*ESAIR 2014 Co-chair  
Gavagai & KTH  
Stockholm*

# Table of Contents

<b>ESAIR 2014 Workshop Organization</b> .....	vii
---	-----

## Keynote Address I

Session Chair: Jaap Kamps (*University of Amsterdam*)

• <b>Semantic Search at Yahoo!</b> .....	1
Peter Mika ( <i>Yahoo Labs</i> )	

## Keynote Address II

Session Chair: Omar Alonso (*Microsoft*)

• <b>Linking to Web Knowledge Bases and Applications to Web Search</b> .....	3
Silviu Petru Cucerzan ( <i>Microsoft</i> )	

## Boaster Session

Session Chair: Jaap Kamps (*University of Amsterdam*)

• <b>Documents Search Using Semantics Criteria</b> .....	5
Santiago Cotelo, Alejandro Makowski, Luis Chiruzzo, Dina Wonsever ( <i>Universidad de la República, Montevideo, Uruguay</i> )	
• <b>Towards Named-Entity-based Similarity Measures: Challenges and Opportunities</b> .....	9
Tom De Nies ( <i>Ghent University-iMinds, Belgium</i> ), Christian Beecks ( <i>RWTH Aachen University, Germany</i> ), Wesley De Neve ( <i>Ghent University &amp; KAIST</i> ), Thomas Seidl ( <i>RWTH Aachen University</i> ), Erik Mannens, Rik Van de Walle ( <i>Ghent University</i> )	
• <b>Can Corpus Similarity-Based Self-Annotation Assist Information Retrieval?</b> .....	13
Vinay Deolalikar ( <i>Hewlett-Packard Research</i> )	
• <b>AIDA-Social: Entity Linking on the Social Stream</b> .....	17
Yusra Ibrahim, Mohamed Amir Yosef, Gerhard Weikum ( <i>Max Planck Institute for Informatics, Germany</i> )	
• <b>A Probabilistic Concept Annotation for IT Service Desk Tickets</b> .....	21
Ea-Ee Jan, Kuan-Yu Chen, Tsuyoshi Ide ( <i>IBM T.J. Watson</i> )	
• <b>Semantic Annotation with RescoredESA: Rescoring Concept Features Generated from Explicit Semantic Analysis</b> .....	25
Zhuoren Jiang ( <i>Dalian Maritime University, China</i> ), Miao Chen, Xiaozhong Liu ( <i>Indiana University Bloomington</i> )	
• <b>Using Semantic Role Labeling to Predict Answer Types</b> .....	29
Zuyao Li ( <i>University of Southern California</i> ), Peter Exner, Pierre Nugues ( <i>Lund University, Sweden</i> )	
• <b>Leverage the Associations between Documents, Subject Headings and Terms to Enhance Retrieval</b> .....	33
Jin Mao ( <i>Wuhan University, China</i> ), Kun Lu ( <i>University of Oklahoma</i> )	
• <b>Bringing Head Closer to the Tail with Entity Linking</b> .....	37
Manisha Verma ( <i>University College London</i> ), Diego Ceccarelli ( <i>ISTI CNR, Pisa</i> )	
• <b>A Fragment-Based Similarity Measure for Concept Hierarchies and Ontologies</b> .....	41
Hui Yang ( <i>Georgetown University</i> )	
• <b>Exploiting Inference from Semantic Annotations for Information Retrieval: Reflections from Medical IR</b> .....	43
Guido Zuccon ( <i>Queensland University of Technology</i> ), Bevan Koopman ( <i>CSIRO, Australia</i> ), Peter Bruza ( <i>Queensland University of Technology</i> )	

<b>Author Index</b> .....	46
---------------------------	----

# ESAIR 2014 Workshop Organization

**General Chairs:** Omar Alonso (*Microsoft, Mountain View CA, USA*)  
Jaap Kamps (*University of Amsterdam, The Netherlands*)  
Jussi Karlgren (*Gavagai Stockholm, Sweden*)

**Program Committee:** Omar Alonso (*Microsoft Bing, USA*)  
Krisztian Balog (*University of Stavanger, Norway*)  
Shlomo Geva (*QUT, Australia*)  
Djoerd Hiemstra (*University of Twente, The Netherlands*)  
Jaap Kamps (*University of Amsterdam, The Netherlands*)  
Noriko Kando (*National Institute of Informatics, Japan*)  
Jussi Karlgren (*UC Berkeley, USA*)  
Ray Larson (*Microsoft Research, USA*)  
Liz Liddy (*Syracuse University, USA*)  
Maarten Marx (*University of Amsterdam, The Netherlands*)  
Peter Mika (*Yahoo! Research Barcelona, Spain*)  
Patrick Pantel (*Yahoo! Research, USA*)  
Livia Polanyi (*Stanford University, USA*)  
Ralf Schenkel (*Universität Passau, Germany*)  
Andrew Trotman (*University of Otago, New Zealand*)  
Roman Yangarber (*University of Helsinki, Finland*)

**Sponsors:**

**SIGIR**  
Special Interest Group  
on Information Retrieval



**In cooperation with:**

