On behalf of the organizing committee, it is my great pleasure to welcome you to the 36th European Conference on Information Retrieval (ECIR 2014) in Amsterdam. I hope you will find the conference inspiring, useful and interesting.

The annual European Conference on Information Retrieval is the main European forum for the presentation of new research results in the field of information retrieval. The 36th edition of the conference is organized by the Intelligent Systems Lab Amsterdam at the University of Amsterdam.

This year, the conference will see 33 full paper presentations, 50 short paper presentations, 15 demonstrations as well as 8 presentations in the Industry sessions.

Putting together the program for ECIR 2014 represents a tremendous effort by a large team of people. I would like to thank the local team, led by Tom Kenter, who have worked hard to get every detail just right. A big thank you to the program committee led by Arjen P. de Vries and ChengXiang Zhai. I am grateful to Franciska de Jong and Kira Radinsky for doing a fantastic job in selecting the short papers. I thank Leif Azzopardi for putting together a great selection of workshops and Edgar Meij for organizing an inspiring tutorial program. I am very grateful to Katja Hofmann who selected the demonstrations. Finally, I am also very grateful to David Carmel and Thijs Westerveld for composing a stimulating industry session.

John Tait chaired the best paper award committee and Ayse Göker chaired the KSJ Award committee. Thank you both!

ECIR has traditionally had a high level of student participation. I am glad that we are continuing this tradition this year. Thanks to generous support from a large number of sponsors, we have been able to keep the registration rates affordable and the number of fee waivers and travel grants high. The City of Amsterdam, the European Science Foundation (ESF), the ELIAS Research Network Programme, Google, Microsoft Research, the Netherlands Organization for Scientific Research (NWO), Textkernel, Yandex,
the Netherlands Institute for Sound and Vision, CRC Press, NOW Publishers, Springer: your sponsorship made a tremendous difference. And the valuable support by the BCS-IRSG, the Intelligent Systems Lab Amsterdam, Girl Geek Diner, and the Friends of SIGIR program is gratefully acknowledged.

Yahoo! Labs sponsored the best paper award, the best short paper award, the best demo award, and the best reviewer award. Thank you!

Finally, I would like to thank you, the authors, presenters and participants for being part of ECIR 2014.

Enjoy Amsterdam. Enjoy ECIR 2014!

Maarten de Rijke
ECIR 2014 General Chair

---

**ORGANIZATION**

<table>
<thead>
<tr>
<th>Chairs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General chair</strong></td>
</tr>
</tbody>
</table>
| **Program co-chairs** | Arjen P. de Vries (CWI Amsterdam)  
ChengXiang Zhai (University of Illinois at Urbana-Champaign) |
| **Local chair** | Tom Kenter (University of Amsterdam) |
| **Short paper co-chairs** | Kira Radinsky (SalesPredict)  
Franciska de Jong (University of Twente) |
| **Workshop chair** | Leif Azzopardi (University of Glasgow) |
| **Industry day co-chairs** | David Carmel (Yahoo! Research)  
Thijs Westerveld (WizeNoze) |
| **Demo chair** | Katja Hofmann (Microsoft Research) |
| **Tutorial chair** | Edgar Meij (Yahoo! Research) |

<table>
<thead>
<tr>
<th>Local Organization</th>
</tr>
</thead>
</table>
| **Finance** | Anne Schuth (University of Amsterdam)  
Caroline van Impelen (University of Amsterdam) |
| **Location** | Daan Odijk (University of Amsterdam) |
| **Social Events** | Richard Berendsen (University of Amsterdam)  
Wouter Weerkamp (904Labs) |
<p>| <strong>Volunteer</strong> | Ridho Reinanda (Royal Netherlands Academy of Arts and Sciences and University of Amsterdam) |
| <strong>Sponsoring</strong> | Maria-Hendrieka Peetz (University of Amsterdam) |
| <strong>Publicity</strong> | David Graus (University of Amsterdam) |
| <strong>Proceedings</strong> | Zhaochun Ren (University of Amsterdam) |</p>
<table>
<thead>
<tr>
<th>Time</th>
<th>UvA1</th>
<th>UvA2</th>
<th>UvA3</th>
<th>UvA4</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:00</td>
<td>09:00 — 12:15 Text Quantification</td>
<td>09:00 — 12:30 Gamification for Information Retrieval</td>
<td>09:00 — 12:30 Bibliometric-Enhanced IR</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
</tr>
<tr>
<td>09:00</td>
<td>09:00 — 12:30 Gamification for Information Retrieval</td>
<td>Half-day Workshop</td>
<td>Half-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>10:00</td>
<td>09:00 — 12:30 Bibliometric-Enhanced IR</td>
<td>Half-day Workshop</td>
<td>Half-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>11:00</td>
<td>09:00 — 12:30 Gamification for Information Retrieval</td>
<td>Half-day Workshop</td>
<td>Half-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>12:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>13:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>13:30</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>14:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>15:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>16:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>17:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>18:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>19:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>20:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
<tr>
<td>21:00</td>
<td>09:00 — 17:15 Context Aware Retrieval and Recommendation</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
<td>Full-day Workshop</td>
</tr>
</tbody>
</table>

**SUNDAY APRIL 13**

10:30 — 11:00 Coffee Break

12:30 — 13:45 Lunch at Erasmus

15:15 — 15:45 Coffee Break

18:00 — 20:00 Welcome reception: Riva, Amstelboulevard 1, Amsterdam
In recent years it has been pointed out that, in a number of applications involving (text) classification, the final goal is not determining which class (or classes) individual unlabelled data items belong to, but determining the prevalence (or “relative frequency”) of each class in the unlabelled data. The latter task is known as quantification. In recent years the research community has shown a growing interest in tackling quantification as a task in its own right. One of the reasons is that, since the goal of quantification is different than that of classification, quantification requires evaluation measures different from those used for classification. A second, related reason is that, as it has been shown, using a method optimized for classification accuracy is suboptimal when quantification accuracy is the real goal. A third reason is the growing awareness that quantification is going to be more and more important; with the advent of big data, more and more application contexts are going to spring up in which we will simply be happy with analyzing data at the aggregate (rather than at the individual) level. The goal of this tutorial is to introduce the audience to the problem of quantification, to the techniques that have been proposed for solving it, to the metrics used to evaluate them, and to the problems that are still open in the area.
Gamification for Information Retrieval (GamifIR’ 2014)
Frank Hopfgartner (Technische Universität Berlin), Gabriella Kazai (Microsoft Research), Udo Kruschwitz (University of Essex), Michael Meder (TU Berlin)

Gamification is the application of game mechanics such as leaderboards, badges or achievement points in non-gaming environments. The goal is to achieve more accurate work, better retention rates, and a more cost-effective solution by relating motivations for participating as more intrinsic than conventional methods. Gamification has recently emerged as a major research area, however its adoption in Information Retrieval (IR) is still in its infancy despite the fact that there are many IR tasks that could potentially benefit from gamification techniques including the manual annotation of documents in IR evaluation, the participation in user studies to study interactive IR challenges, and the move from single-user search to social search.

This workshop intends to narrow down and focus on the challenges and opportunities that gamification can present for the IR community. We aim to bring together researchers and practitioners from a wide range of areas including information retrieval, human-computer interaction, computer games, and natural language processing.

Information Access in Smart Cities (i-ASC)
M-Dyaa Albakour (University of Glasgow), Craig Macdonald (University of Glasgow), Iadh Ounis (University of Glasgow), Charles L. A. Clarke (University of Waterloo), Veli Bicer (IBM Research)

Modern cities are becoming smart where a digital knowledge infrastructure is deployed by local authorities (e.g. City councils and municipalities) to better serve the information needs of their citizens, and to ensure sustainability and efficient use of power and resources. This knowledge infrastructure consists of a wide range of systems from low-level physical sensors to advanced sensing devices through social sensors. This proposed workshop will be a venue for research on digesting the city’s data streams and knowledge databases in order to serve the information needs of citizens and support decision making for local authorities. Possible use cases include helping tourists to find interesting places to go or activities to do while visiting a city, or assisting journalists in reporting local incidents. Indeed, this workshop will foster the development of new information access and retrieval models that can harness effectively and efficiently the large number of heterogeneous big data streams in a city to provide a new generation of information services.
Full Day Workshop

The 4th Workshop on Context-awareness in Retrieval and Recommendation (CaRR)
Alan Said (CWI), Ernesto De Luca (FH Potsdam), Matthias Böhmer (DFKI), Daniele Quercia, (Yahoo! Research)

Context-aware information is widely available in various ways such as interaction patterns, devices, annotations, query suggestions and user profiles and is becoming more important for enhancing retrieval performance. At the moment, the main issue to cope with is not only retrieving the most relevant items and content, but defining them ad hoc. Further relevant issues are personalizing and adapting the information and the way it is displayed to the users current situation (device, location, social surrounding) and interests. In the 4th edition of the workshop we want to focus on integrating social context into retrieval and recommendation.

13:45 – 17:15 | Room: UvA1
Tutorial
Designing Search Usability
Tony Russell-Rose (UXLabs)

Search is not just a box and ten blue links. Search is a journey: an exploration where what we encounter along the way changes what we seek. But in order to guide people along this journey, we must understand both the art and science of user experience design.

The aim of this tutorial is to deliver a learning experience grounded in good scholarship, integrating the latest research findings with insights derived from the practical experience of designing and optimizing an extensive range of commercial search applications. It focuses on the development of transferable, practical skills that can be learnt and practised within a half-day session.

13:45 – 17:15 | Room: UvA2
Tutorial
The Cluster Hypothesis in Information Retrieval
Oren Kurland (Technion – Israel Institute of Technology)

The cluster hypothesis (van Rijsbergen ’79) states that “closely associated documents tend to be relevant to the same requests”. This is one of the most fundamental and influential hypotheses in the information retrieval field. We will survey the different lines of work that the hypothesis has given rise to (e.g., cluster-based retrieval, using topic modeling for retrieval). The survey will be accompanied by an in-depth analysis of the retrieval techniques that are inspired by the cluster hypothesis and which are used for various tasks including ad hoc retrieval, meta-search, microblog (e.g., Twitter) retrieval, query-performance prediction, search-results diversification.
Chair: Ayse Göker

A long standing challenge in Web search is how to accurately determine the intention behind a searcher’s query, which is needed to rank, organize, and present results most effectively. The difficulty is that users often do not (or cannot) provide sufficient information about their goals. As this talk will show, it is nevertheless possible to read their intentions through clues revealed by behavior, such as the amount of attention paid to a document or a text fragment. I will overview the approaches that have emerged for acquiring and mining behavioral data for inferring search intent, ranging from robust models of click data in the aggregate, to modeling fine-grained user interactions such as mouse cursor movements in the searcher’s browser. The latter can also be used to measure the searcher’s attention “in the wild”, with granularity approaching that of using eye tracking equipment in the laboratory. The resulting techniques and models have already shown noteworthy improvements for search tasks such as ranking, relevance estimation, and result summary generation, and have applications to other domains, such as psychology, neurology, and online education.
11:00 — 12:15 | Room: UvA1-2
Session 1A: Evaluation I
Chair: Pavel Serdyukov

11:00
Reducing Reliance on Relevance Judgments for System Comparison by Using Expectation-Maximization
Ning Gao (University of Maryland, College Park), William Webber (William Webber Consulting), Douglas Oard (University of Maryland, College Park)

11:00 — 12:15 | Room: UvA3-4
Session 1B: Recommendation
Chair: Leif Azzopardi

11:00
Content + Attributes: a Latent Factor Model for Recommending Scientific Papers in Heterogeneous Academic Networks
Chenyi Zhang (Zhejiang University), Xueyi Zhao (Zhejiang University), Ke Wang (Simon Fraser University), Jianling Sun (Zhejiang University)

11:25
Best and Fairest: an Empirical Analysis of Retrieval System Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow)

11:25
Real-Time News Recommendation using Context-aware Ensembles
Andreas Lommatzsch (TU Berlin)

11:50
Tackling Biased Baselines in the Risk-Sensitive Evaluation of Retrieval Systems
Bekir Taner Dincer (University of Mugla), Iadh Ounis (University of Glasgow), Craig Macdonald (University of Glasgow)

11:50
A Personalised Recommendation System for Context-Aware Suggestions
Andrey Rikitianskiy, Morgan Harvey, Fabio Crestani (University of Lugano)

13:45 — 15:15 | Room: UvA1-4
Poster Session 1

1. Video Clip Retrieval by Graph Matching
Manal Al Ghamdi, Yoshinoko Gotoh (University of Sheffield, UK)

2. EDIUM: Improving Entity Disambiguation via User Modeling
Romil Bansal, Sandeep Panem, Manish Gupta, Vasudeva Varma (IIIT, Hyderabad)

3. A Comparison of Approaches for Measuring Cross-Lingual Similarity of Wikipedia Articles
Alberto Barrón-Cedeño (TaLP Research Center, Universitat Politècnica de Catalunya, Spain), Monica Lestani Paramita (University of Sheffield, UK), Paul Clough (University of Sheffield, UK), Paolo Rosso (ELiRF-DSIC, Universitat Politècnica de València)

4. Challenges on Combining Open Web and Dataset Evaluation Results: The Case of the Contextual Suggestion Track
Alejandro Bellogín, Thaer Samar, Arjen P. de Vries, Alan Said (CWI, Amsterdam)

5. Integrating Multiple Resources for Diversified Query Expansion
Arbi Bouchoucha, Xiaohua Liu, Jian-Yun Nie (University of Montreal, Quebec, Canada)

6. Facet-based User Modeling in Social Media for Personalized Ranking
Chen Chen (Nankai University, China)

7. Spot the Ball: Detecting Sports Events on Twitter
David Corney, Carlos Martin, Ayse Göker (Robert Gordon University, UK)

8. Exploiting Result Diversification Methods for Feature Selection in Learning to Rank
Kaweh Djaafari Naini (L3S Research Center, Leibniz University of Hannover), Ismail Sengor Altingovde (Middle East Technical University)

9. Boilerplate Detection and Recoding
Matthias Gallé, Jean-Michel Renders (Xerox Research Centre Europe, France)

10. A Two-level Approach for Subtitle Alignment
Jia Huang (Drexel University, USA), Hao Ding (New York University, USA), Xiaohua Hu (Drexel University, USA), Yong Liu (New York University, USA)

11. Analyzing Discussions on Twitter: Case Study on HPV Vaccinations
Rianne Kaptein, Erik Boertjes, David Langley (TNO, The Netherlands)

12. Automatically Retrieving Explanatory Analogies from Webpages
Varun Kumar, Savita Bhat, Niranjan Pedanekar (Tata Consultancy Services, India)

13. Geo-Spatial Domain Expertise in Microblogs
Wen Li (Delft University of Technology, Netherlands), Carsten Eickhoff (Delft University of Technology, Netherlands), Arjen P. de Vries (CWI, Amsterdam, Netherlands)
Shangsong Liang, Zhaochun Ren, Maarten de Rijke (University of Amsterdam, The Netherlands)

15. Towards Digital Sensitivity Review
Graham McDonald, Craig Macdonald, Iadh Ounis, Tim Gollins (University Of Glasgow, UK)

16. Unsupervised Approach for Identifying Users’ Political Orientations
Youssef Meguebli (SUPELEC Systems Sciences E3S, France), Mouna kacimi (Free University of Bozen-Bolzano, Italy), Bich-Liên Doan (SUPELEC Systems Sciences E3S, France)

17. User Perception of Information Credibility of News on Twitter
Shafiza Mohd Shariff, Xiuzhen Zhang, Mark Sanderson (RMIT University, Australia)

18. Sentiment Analysis and the Impact of Employee Satisfaction on Firm Earnings
Andy Moniz (APG Asset Management, The Netherlands), Franciska de Jong (Erasmus University, The Netherlands)

19. Entity Tracking in Real-Time using Sub-Topic Detection on Twitter
Sandeep Panem, Romil Bansal, Manish Gupta, Vasudeva Varma (IIIT, Hyderabad, India)

20. Time-Aware Focused Web Crawling
Pedro Pereira (University of Minho, Portugal), Olga Craveiro (University of Coimbra-Portugal), Joaquim Macedo (University of Minho, Portugal), Henrique Madeira (University of Coimbra-Portugal)

21. Temporal Expertise Profiling
Jan Rybak (Norwegian University of Science and Technology, Norway), Kristzian Balog (University of Stavanger, Norway), Kjetil Nervåg (Norwegian University of Science and Technology, Norway)

22. Assessing Quality of Unsupervised Topics in Song Lyrics
Lucas Sterckx, Thomas Demeester, Johannes Deleu, Laurent Mertens, Chris Develder (Ghent University – iMinds)

23. An Information Retrieval-Based Approach to Determining Contextual Opinion Polarity of Words
Olga Vechtomova, Kaheer Suleman, Jack Thomas (University of Waterloo, Canada)

24. Query Term Suggestion in Academic Search
Suzan Verberne (Radboud University Nijmegen, The Netherlands), Maya Sappelli (Radboud University Nijmegen, The Netherlands), Wessel Kraaij (Radboud University Nijmegen, The Netherlands, TNO, The Netherlands), Egon L. Van den Broek (Department of Information and Computing Sciences, Utrecht University)

25. The Impact of Future Term Statistics in Real-Time Tweet Search
Yulu Wang, Jimmy Lin (University of Maryland, USA)
15:45 — 17:00 | Room: UvA1-2
Session 2A: Optimization and prediction
Chair: Eric Gaussier

15:45
Optimizing Base Rankers Using Clicks: A Case Study using BM25
Anne Schuth, Floor Sietsma, Shimon Whiteson, Maarten de Rijke (University of Amsterdam)

16:10
Predicting Search Task Difficulty
Jaime Arguello (University of North Carolina at Chapel Hill)

16:35
Crawling Policies Based on Web Page Popularity Prediction
Liudmila Ostroumova, Ivan Bogatyy, Arseniy Chelnokov, Alexey Tikhonov, Gleb Gusev (Yandex)

15:45 — 17:00 | Room: UvA3-4
Session 2B: Semantics and annotation
Chair: Edgar Meij

15:45
Measuring the Effectiveness of Gamesourcing Expert Oil Painting Annotations
Myriam Traub, Jacco van Ossenbruggen, Jiying He, Lynda Hardman (CWI)

16:10
Relevance-Ranked Domain-Specific Synonym Discovery
Andrew Yates, Nazli Goharian, Ophir Frieder (Georgetown University)

16:35
Towards Generating Text Summaries for Entity Chains
Shruti Chhabra (IIIT-Delhi)
The relation between academic and industrial research in information retrieval has been a frequent topic of discussion. In discussions about this relation, the use of proprietary data in IR research is one issue that gives rise to a range of positions. For instance, it was recently decided to turn ICTIR into a more ‘foundational’ conference, with an emphasis on experiments on data that is openly available to the research community. This move probably reflects the position that for scientific results to be reproducible, the data being used should be open. The inclusion of industry days or industry sessions at recent editions of SIGIR and ECIR represents a different position, namely that valuable insights can be gained from proprietary data.

But there’s more. As the community’s research agenda develops, a strong emphasis on online algorithms and online evaluation is emerging. Is this shift a source of divergence between academic and industrial research in IR? Work on online evaluation may require access to live systems. And it already seems that expertise on online algorithms and time-aware ranking methods can more easily be found in industrial research environments than in academia.

The ECIR 2014 organizers organize a panel to discuss the IR research ecosystem and the roles of academic and industrial research in this ecosystem. Our goal is to stimulate discussion in the IR community that may inform future strategy development.
11:00 — 12:15 | Room: UvA1-2
Session 3A: Evaluation II
Chair: Jaime Arguello

11:00
Tolerance of Effectiveness Measures to Crowdsourced Relevance Judging Errors
Le Li, Mark D. Smucker (University of Waterloo)

11:25
Evaluation of IR Applications with Constrained Result Panel
Yuanhua Lv, Ariel Fuxman, Ashok Chandra (Microsoft Research)

11:50
Dissimilarity based Query Selection for Efficient Preference based IR Evaluation
Gabriella Kazai (Microsoft UK + USA), Homer Sung (Microsoft USA)

11:00 — 12:15 | Room: UvA3-4
Session 3B: Aggregation
Chair: Fabio Crestani

11:00
Blending Vertical and Web results
Damien Lefortier (Yandex & University of Amsterdam), Pavel Serdyukov (Yandex), Fedor Romanenko (Yandex), Maarten de Rijke (University of Amsterdam)

11:25
Personalizing Aggregated Search
Stanislav Makeev, Andrey Plakhov, Pavel Serdyukov (Yandex)

11:50
HetPathMine: A Novel Transductive Classification Algorithm on Heterogeneous Information Networks
Chen Luo (Jilin University, China), Renchu Guan (Jilin University, China), Zhe Wang (Jilin University, China), Chenghua Lin (University of Aberdeen, Scotland)

13:45 — 15:15 | Room: UvA1–4
Poster Session 2

Fawaz Alarfaj, Udo Kruschwitz, Chris Fox (University of Essex, UK)

27. Score Normalization using Logistic Regression with Expected Parameters
Robin Aly (University of Twente, The Netherlands)

28. More Informative Open Information Extraction via Simple Inference
Hannah Bast, Elmar Haussmann (University of Freiburg, Germany)

29. Learning a Theory of Marriage (and other relations) from a Web Corpus
Sandro Bauer (University of Cambridge, UK), Stephen Clark (University of Cambridge, UK), Laura Rimell (University of Cambridge, UK), Thore Graepel (Microsoft Research Cambridge, UK)

30. Learning from User Interactions for Recommending Content in Social Media
Mathias Breuss, Manos Tsagkias (University of Amsterdam, The Netherlands)

31. Towards an Entity-based Automatic Event Validation
Andrea Ceroni, Marco Fisichella (L3S Research Center, Leibniz Universität Hannover)

32. Bringing Information Retrieval into Crowdsourcing
Qinmin Hu, Jimmy Huang (York University, Canada)

33. Query Expansion with Temporal Segmented Texts
Olga Craveiro (Polytechnic Institute of Leiria, Portugal, University of Coimbra, Portugal), Joaquim Macedo (University of Minho, Portugal), Henrique Madeira (GISUC, University of Coimbra, Portugal)

34. On Clustering and Polyrerepresentation
Ingó Frommholz, Muhammad Kamran Abbasi (University of Bedfordshire, UK)

35. Effects of Position Bias on Click-Based Recommender Evaluation
Katja Hofmann (Microsoft Research, UK), Anne Schuth (ISLA, University of Amsterdam), Alejandro Bellogin (CWI), Maarten de Rijke (ISLA, University of Amsterdam)

36. Quality-based Automatic Classification for Presentation Slides
Seongchang Kim, Wonchul Jung, Kjeejun Han, Jae-Gil Lee, Mun Y. Yi (KAIST)

37. Context of Seasonality in Web Search
Tomáš Kramár, Mária Bieliková (Slovak University of Technology, Slovak Republic)

38. A Language Modeling Approach to Personalized Search based on Users’ Microblog Behavior
Arjumand Younus (National University of Ireland, Ireland, University of Milano-Bicocca, Italy), Colm O’Riordan (National University of Ireland, Ireland) Gabriella Pasi (University of Milano-Bicocca, Italy)
39. On the Effect of Locality in Compressing Social Networks
Panagiotis Liakos (University of Athens, Greece), Katia Papakonstantinopoulou (University of Athens, Greece), Michael Sioutis (Université Lille-Nord de France, France)

40. Cross-Domain Collaborative Filtering with Factorization Machines
Babak Loni, Yue Shi, Martha Larson, Alan Hanjalic (Delft University of Technology, The Netherlands)

41. Improvements to Suffix Tree Clustering
Richard Elling Moe (University of Bergen, Norway)

42. Deep Learning for Character-based Information Extraction
Yanjun Qi (University of Virginia, USA), Sujatha G Das (Penn State University), Ronan Collobert (IDIAP, Switzerland), Jason Weston (Google Research)

43. Text-Image Topic Discovery for Web News Data
Mingjie Qian (University of Illinois at Urbana-Champaign, USA)

44. Detecting Event Visits in Urban Areas via Smartphone GPS Data
Richard Schaller (University of Erlangen-Nuremberg, Germany), Morgan Harvey (University of Lugano, Switzerland), David Elsweiler (University of Regensburg, Germany)

45. A Case for Hubness Removal in High-Dimensional Multimedia Retrieval
Dominik Schnitzer (OFAI, Austria), Arthur Flexer (OFAI, Austria), Nenad Tomasev (Jožef Stefan Institute, Slovenia)

46. Multi-evidence User Group Discovery in Professional Image Search
Theodora Tsikrika (CERTH, Thessaloniki, Greece), Christos Diou (Aristotle University of Thessaloniki, Greece)

47. Analyzing Tweets to aid Situational Awareness
Tim L.M. van Kasteren, Birte Ulrich, Vignesh Srinivasan, Maria E. Niesen (AGT International, Germany)

48. Query-dependent Contextualization of Streaming Data
Nikos Voskarides, Daan Odijk, Manos Tsagkias, Wouter Weerkamp, Maarten de Rijke (University of Amsterdam)

49. An Exploration of Tie-Breaking for Microblog Retrieval
Yue Wang, Hao Wu, Hui Fang (University of Delaware, USA)

50. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

51. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

52. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

53. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

54. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

55. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

56. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

57. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

58. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

59. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)

60. Efficiently Estimating Retrievability Bias
Colin Wilkie, Leif Azzopardi (University of Glasgow, UK)
15:45 — 17:00 | Room: UvA1-2
Session 4A: Queries
Chair: Jaap Kamps

15:45
Leveraging Dynamic Query Subtopics for Time-aware Search Result Diversification
Tu Ngoc Nguyen, Nattiya Kanhabua (L3S Research Center)

16:10
A Study of Query Term Deletion using Large-scale E-commerce Search Logs
Bishan Yang (Cornell University), Nish Parikh (eBay Research Labs), Gyanit Singh (eBay Research Labs), Neel Sundaresan (eBay Research Labs),

16:35
Detecting Missing Content Queries in an SMS-Based HIV/AIDS FAQ Retrieval System
Edwin Thuma (University of Glasgow & University of Botswana), Iadh Ounis (University of Glasgow), Simon Rogers (University of Glasgow)

15:45 — 17:00 | Room: UvA3-4
Session 4B: Mining Social Media
Chair: Iadh Ounis

15:45
Cross-Language Pseudo-Relevance Feedback Techniques for Informal Text
Chia-Jung Lee, Bruce W. Croft (University of Massachusetts Amherst)

16:10
Hierarchical Multi-label Conditional Random Fields for Aspect-Oriented Opinion Mining
Diego Marcheggiani (ISTI-CNR), Oscar Täckström (SICS), Andrea Esuli (ISTI-CNR), Fabrizio Sebastiani (ISTI-CNR)

16:35
Generating Pseudo-ground Truth for Predicting New Concepts in Social Streams
David Graus, Manos Tsagkias, Lars Buitinck, Maarten de Rijke (University of Amsterdam)
Twitter’s search engine faces some of the most unique challenges in information retrieval and distributed systems today. On the scaling front, it is a relatively young system with a massive user base, billions of queries daily, and many billions of indexed documents – with thousands being added every second. On the ranking side, the combination of realtime and social requires new solutions to relevance estimation for newly-created documents, blending different types of live content, evaluation in the absence of direct user feedback, and more. On top of this, the dynamic nature of a nascent company and product leads to a multitude of operational challenges and opportunities. In this talk, I will cover some of these challenges and how we approach them at Twitter.
11:00 — 12:15 | Room: UvA1-2
Session 5: Digital Libraries
Chair: Nicola Ferro

11:00
Thumbnail Summarization Techniques for Web Archives
Ahmed AlSum, Michael Nelson (Old Dominion University)

11:25
CiteSeerX: A Scholarly Big Dataset
Cornelia Caragea (University of North Texas), Jian Wi (Pennsylvania State University), Alina Ciobanu (University of Bucharest), Kyle Williams (Pennsylvania State University), Juan Fernandez-Ramirez (University of the Andes), Hung-Hsuan Chen (Pennsylvania State University), Zhaohui Wu (Pennsylvania State University), C. Lee Giles (Pennsylvania State University)

11:50
"User Reviews in the Search Index? That'll Never Work!"
Marijn Koolen (University of Amsterdam)

13:45 — 15:00 | Room: UvA1-2
Session 6: Efficiency
Chair: Andrew Trotman

13:45
A Scalable Gibbs Sampler for Probabilistic Entity Linking
Neil Houlsby (University of Cambridge), Massimiliano Ciaramita (Google Research Zurich)

14:10
Effective Kernelized Online Learning in Language Processing Tasks
Simone Filice, Giuseppe Castellucci, Danilo Croce, Roberto Basili (University of Rome Tor Vergata)

14:35
On Inverted Index Compression for Search Engine Efficiency
Matteo Catena (Gran Sasso Science Institute), Craig Macdonald (University of Glasgow), Iadh Ounis (University of Glasgow)
15:30 — 16:45 | Room: UvA1-2
Session 7: Information Retrieval Theory
Chair: Norbert Fuhr

15:30
Exploring the Space of IR Functions
Parantapa Goswami (Joseph Fourier University, France), Simon Moura (Joseph Fourier University, France), Eric Gaussier (Joseph Fourier University, France), Massih-Reza Amini (Joseph Fourier University, France), Francis Maes (University of Leuven)

15:55
Metric Spaces for Temporal Information Retrieval
Matteo Brucato (University of Massachusetts, Amherst), Danilo Montesi (University of Bologna)

16:20
Local Linear Matrix Factorization for Document Modeling
Lu Bai, Jiafeng Guo, Yanyan Lan, Xueqi Cheng (Institute of Computing Technology, Chinese Academy of Sciences)

15:30 — 16:45 | Room: UvA3-4
Industry session 3: eCommerce and Product Search
Chair: Thijs Westerveld

15:30
Scalable Recommender Systems and Its Similarity with Advertising Systems
Joaquin Delgado (Intel media, USA)

15:55
eCommerce Search Platform Based on Lucene and Solr
Mikhail Khudnev (GridDynamics, USA)

16:20
Addressing Volatility and Catering to the Long-Tail: Unique Challenges in Product Search
Nish Parikh (eBay, US)
SOCIAL EVENTS

Sunday 18:00 — 20:00 Welcoming Reception
After the workshops and tutorials on Sunday we’ll take a short walk from Hotel Casa 400 to Riva to enjoy some drinks, snacks and a marvelous view on the Amstel river.

Where: Riva, Amstelboulevard 1 (see map on p. 44)
How to get there: 10 min walk

Monday 17:15 — 18:00 Women Networking Event
At Erasmus there is a women networking event sponsored by Girl Geek Dinners Amsterdam, who will provide ice cream. This meeting intends to connect women in IR from all stages of their career, to foster a strong network of women in IR.

Monday 20:30 — late (Discretionary) Downtown Drinks
For those who are interested, on Monday evening there’s a small ECIR get-together at De Bekeerde Suster, near the “Wallen” area in Downtown Amsterdam. There will be opportunity to enjoy some beautiful Amsterdam-brewed beers. By its discretionary nature, we’ll have a cash bar at De Bekeerde Suster.

Where: De Bekeerde Suster, Kloveniersburgwal 6
How to get there: Take Metro 51, 53 or 54 from Amstel Station, to metrostop Nieuwmarkt

Tuesday 17:30 Boats ‘n’ Banquet
The ECIR banquet dinner will take place at Jamie Oliver’s Fifteen. We leave from Hotel Casa 400 to board four canal boats, around the corner at the Amstel river (see map on p. 44). The canal boats will take us down the Amstel river giving us an excellent view of Amsterdam’s waterside and canal district. The canal boats take us up to the IJ lake, where we will disembark to enter Fifteen.

Boarding starts at 17:30, and the canal boats leave at 18:00 sharp, in order to make sure we arrive at Fifteen around 19:00.

The Banquet Dinner will additionally be the venue for ECIR’s awards ceremony:
• Best Paper Award
• Best Poster Award
• Best Demo Award
• Best Reviewer Award

Note: Make sure to bring your banquet voucher to the banquet dinner!

Getting back
To return to the conference hotel, the easiest way is to go to Central Station by tram (or take a ~20min walk), and take a Metro from Central Station to Amstel Station. From there you can walk to Hotel Casa 400. The tram stop Kattenburgerstraat is at the Piet Heinkade, close to Fifteen.

Metro lines from Central Station that stop at Amstel Station:
• Metro 51 (Direction Westwijk)
• Metro 53 (Direction Gaasperplas)
• Metro 54 (Direction Gein)

Tram 26 from Kattenburgerstraat to Centraal Station
Departs every ~8 minutes (travel time: 5 minutes)

Metro 51, 53 and 54 from Centraal Station to Amstel Station.
Departs every ~3 minutes (travel time: 7 minutes)

Last two public transportion options to Hotel Casa 400

<table>
<thead>
<tr>
<th>Time</th>
<th>Destination</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:00</td>
<td>Tramstop Kattenburgerstraat Tram 26 (direction Centraal Station) to Centraal Station</td>
</tr>
<tr>
<td>00:12</td>
<td>Centraal Station Metro 53 (direction Gaasperplas) to Metrostation Amstel</td>
</tr>
<tr>
<td>00:16</td>
<td>Tramstop Kattenburgerstraat Tram 26 (direction Centraal Station) to Centraal Station</td>
</tr>
<tr>
<td>00:26</td>
<td>Centraal Station Metro 54 (Direction Gein) Metrostation Amstel</td>
</tr>
</tbody>
</table>
Hotel Casa 400 is a 10 minute walk from Amstel Station. At Amstel Station you can take the metro, a tram, bus or train to the city center of Amsterdam. The metro is probably the quickest way to get there (Central Station is about 5 minutes from Amstel Station).

Amsterdam Amstel Station

Café Restaurant Riva
Amstelboulevard 1, 1096 HH Amsterdam

Boat Boarding Point

USEFUL LINKS
Public transport Amsterdam: www.9292.nl/en
General information Amsterdam: www.iamsterdam.com
RESTAURANTS

Walking distance from Hotel Casa 400

Dauphine (French) ± € 35
Prins Bernhardplein 175 (4min)

Cafe Hesp (International) ± € 21
Weesperzijde 130 (7min)

Girassol (Portuguese) ± € 28
Weesperzijde 135 (7min)

VandeMarkt (French) ± € 36
Weesperzijde 144-147 (7min)

La Vallade (French) ± € 34
Ringdijk 23 (9min)

TrouwAmsterdam (French) ± € 25
Wibautstraat 127 (10min)

BAUT (International) ± € 33
Wibautstraat 125 (10min)

Lulu (French/International) ± € 25
Beukenplein 17 (13min)

Thaicoon (Thai) ± € 24
Beukenplein 10 (14min)

Bidou (French) ± € 22
Beukenplein 19-21 (14min)

Riva (International) ± € 20
Amstelboulevard 1 (14min)

Eetcafé Ibis (Ethiopian) ± € 14
Weesperzijde 43 (15min)

Kees. drinken, eten en drinken (Fusion) ± € 27
Weesperzijde 44 (15min)

Sa Seada (Italian) ± € 30
Eerste Oosterparkstraat 3 (15min)

Near metro stop Nieuwmarkt

New King (Chinese) ± € 13
Zeedijk 115

Kam Kee (Chinese) ± € 12
Zeedijk 103

Little Thai Prince (Thai) ± € 20
Zeedijk 33

Café Bern (Swiss) ± € 16
Nieuwmarkt 9

Il Mare (Italian) ± € 20
Zeedijk 56/58

Rodizio.nl (BBQ/Grill) ± € 20
Zeedijk 28

Engelbewaarder (Dutch) ± € 20
Kloveniersburgwal 59

Wau (Malaysian) ± € 27
Zeedijk 35

Ganesha Indian Restaurant
(Indian) ± € 23
Geldersekade 5

Near metro stop Waterlooplein

De Hapjeshoek (Surinamese) ± € 15
Metro stop Waterlooplein 6

OCHA (Thai) ± € 15
Binnen Bantammerstraat 1

Restaurant Greetje (Dutch) ± € 39
Peperstraat 23

Café Stevens (Organic) ± € 17
Geldersekade 123