ACM SIGIR 2011

July 24-28, 2011, Beijing, China

The 34th Annual International ACM SIGIR Conference

Program Booklet

SIGIR

Beijing

2011

Table of Contents

Organization 1
Preface from the General Chairs 3
Preface from the Program Committee Chairs 5
Area Chairs 7
Program Committee 9
SIGIR 2011 Sponsors and Supporters 15
Layout of Hotel 17
Social Program 18
Program at a Glance 19
Doctoral Consortium 23
Tutorials 24
Keynote Addresses 25
Detailed Scientific Program 27
Industry Track Sessions 47
Industrial Keynote Presenters 49
Posters 53
Demonstrations 63
Layout of Poster and Demo Session 65
Workshops 67
Reception at Microsoft Research Asia 69
Industry Lab Tour 70
Near-by Restaurants 71

Organization

General Co-Chairs:
Wei-Ying Ma, Microsoft Research Asia, China
Jian-Yun Nie, University of Montreal, Canada

Program Co-Chairs:
Ricardo Baeza-Yates, Yahoo! Research, Spain
Tat-Seng Chua, National University of Singapore
W. Bruce Croft, University of Massachusetts, Amherst, USA

Posters and Demos Co-Chairs:
Kevyn Collins-Thompson, Microsoft Research, USA
Gary Geunbae Lee, POSTECH, Korea
Massimo Melucci, University of Padua, Italy

Tutorial Chair:
Yi Zhang, University of California, Santa Cruz, USA

Workshop Co-Chairs:
Noriko Kando, National Institute of Informatics, Japan
Luo Si, Purdue University, USA

Mentor Chair:
Gareth Jones, Dublin City University, Ireland

Outstanding Paper Award Chair:
Susan Dumais, Microsoft Research, USA

Doctoral Consortium Co-Chairs:
Marteen de Rijke, University of Amsterdam, the Netherlands
Paul Thomas, CSIRO, Australia

Industrial Track Co-Chairs:
David Carmel, IBM Research, Haifa, Israel
Haifeng Wang, Baidu, China

Sponsorship Chair:
Jimmy Huang, York University, Canada

Publicity Chair:
Kam-Fai Wong, Chinese University of Hong Kong, China

Local Organization Chair:
Xiaoming Li, Peking University, China

Finance Chair:
Hang Li, Microsoft Research Asia, China

Chair of Coordination with Local Societies:
Sheng Li, Harbin Institute of Technology, China

Proceedings Chair:
Zaiqing Nie, Microsoft Research Asia, China

Conference Arrangements:
Racemind Racemind Event Marketing, China

Preface from the General Chairs

Welcome to the 34th ACM SIGIR conference on research and development in information retrieval, and welcome to Beijing!

The whole Chinese IR community is delighted to host this year's SIGIR conference for the first time in China in the historical and vibrant city of Beijing. The last decade has witnessed the great surge of China's Internet and search engine industry as one of the largest and most active in the world. The SIGIR conferences have also seen increased Chinese participation: the number of full papers from China has reached a record high of 18 this year, while more than half of them have at least one Chinese name as author. This year's SIGIR conference is a significant opportunity to forge stronger ties between the Chinese IR community and the international SIGIR community. We hope that you will enjoy the great technical program as well as the supporting events in connection with the search engine industry in China.

The conference promises to be a success, with a record number of full papers to be presented. The program committee, chaired by Ricardo Baeza-Yates, Tat-Seng Chua and Bruce Croft and helped by 43 area chairs and many reviewers, did a great job in selecting high quality papers that reflect the latest developments in IR. We are equally grateful to all the other committee members: the Posters and Demos chairs KevynCollins-Thompson, Gary Geunbae Lee, and Massimo Melucci; the Tutorial chair Yi Zhang; the Workshops chairs Noriko Kando and Luo Si; the Mentoring Program chair Gareth Jones; the Outstanding Paper Award chair Susan Dumais; the Doctoral Consortium chairs Marteen de Rijke and Paul Thomas; the Industrial Track chairs David Carmel and Haifeng Wang; and the Publication chair Zaiqing Nie. We are especially grateful to Jimmy Huang for his active attraction of corporate sponsorship. The conference also benefits from the great help of the Local Organization Committee, chaired by Xiaoming Li. Special thanks go to Hang Li and Ling Huang for their involvement throughout the preparation of this conference.

SIGIR'11 is sponsored by ACM and has received generous financial support from a large number of companies and organizations. To date, Microsoft Research supports at the Diamond level; Baidu, EMC, Google, Tencent and Yahoo! provide Gold support; eBay is a Silver supporter; HP, IBM Research and Yandex take part at the Bronze level; while

Springer, Now Publishers and Morgan & Claypool exhibit their products. SIGIR'11 is organized in association with the Chinese Information Processing Society. Finally, we are thankful to the University of Montreal for allowing Jian-Yun Nie to devote himself entirely to this conference during the last 6 months, and we thank ACM and the SIGIR Executive Committee for their constant support.

We hope you will enjoy the conference and your visit to Beijing and China.

Wei-Ying Ma
Microsoft Research Asia

Jian-Yun Nie
University of Montreal

Preface from the Program Committee Chairs

Welcome to the 34th ACM SIGIR International Conference on Research and Development in Information Retrieval. The record number of papers in this year's conference represents both the breadth and depth of the research being done in this vibrant field, both in academia and industry. We have done our best to ensure that these papers meet high standards of quality in terms of presentation, citations, and experimental methodology. At the same time, we have tried to be flexible in the application of these criteria in order to accept papers describing novel and innovative work that may be somewhat unconventional.

The conference received 543 full paper submissions this year, with 240 (44%) coming from Asia and Pacific region, 185 (34%) from the Americas, and 112 (21%) from Europe (the rest were "unknown"). Of these papers, 108 (19.9%) were accepted, up from the acceptance rate of 16.7% in last year's conference. The top five countries in terms of accepted papers were the U.S.A. (52), China (18), Germany (7), and thenthe U.K. and Spain (both 5). In addition, 274 short papers were submitted to the poster track, of which 89 (32.5%) were accepted. In the other categories, there were 15 (42.8%) demonstrations, 8 workshops, and 11 half-day tutorials accepted. In terms of the technical areas that the accepted papers cover, using the primary keyword assigned by the authors, the top five areas are document representation and content analysis (20%), retrieval models and ranking (17%), users and interactive IR (13%), queries and query analysis (11%), and filtering and recommendation (11%). Perhaps the only surprise there is the increase in the number of papers in filtering and recommendation. We believe that the papers at this year's conference provide an excellent cross-section of what is going on in our field. We hope that you find that reading them and listening to the presenters to be a rewarding experience.

SIGIR uses a two-tier double blind review system. For the full papers, the first step is that at least three first-tier reviewers read every paper and provide ratings and comments. Then two additional reviewers, referred to as the primary or secondary area chairs, study those reviews, and introduce their own opinions and summaries where appropriate by making additional comments. In some cases, the area chairs initiate the discussion among the first-tier reviewers to work out any controversial issues or significant differences of opinion. A new step

introduced this year was to request author feedback for specific issues in some papers. Another change this year was that final decisions for nearly all papers were made by the two area chairs together with the reviewers. At the program committee meeting in Barcelona, the program chairs and some area chairs went over the reviews, obtained additional input, and made decisions in the few cases where the area chairs had requested more discussion.

The success of the conference depends entirely on the large number of volunteers that every year devotes large amounts of their time to reviewing and organizing. We are grateful to the more than 200 first-tier reviewers (or general program committee), who performed their tasks very well. We also acknowledge the tremendous job done by the Posters Chairs in handling the short papers, and the great work done by the other Chairs with their tasks. We are particularly indebted to the 43 Area Chairs for their willingness to try a new approach that required more effort from them, and for their responsiveness throughout the reviewing process. We also acknowledge the help of Thomas Preuss with support for the ConfMaster reviewing system. Finally, we extend our best wishes to next year's Program Chairs!

Ricardo Baeza-Yates
Yahoo! Research Barcelona
Africa and Europe

Tat-Seng Chua
National University of Singapore
Asia and Pacific

W. Bruce Croft
University of Massachusetts
Americas

Area Chairs

Ricardo Baeza-Yates *(Yahoo! Research, Spain)* - Chair
Tat-Seng Chua *(National University of Singapore)* - Chair
W. Bruce Croft *(University of Massachusetts, Amherst, USA)* – Chair

1. Document Representation and Content Analysis
James Allan *(University of Massachusetts Amherst, USA)*
Jamie Callan *(CMU, USA)*
Fernando Diaz *(Yahoo! Research, USA)*
Wessel Kraaij *(Radboud University Nijmegen, The Netherlands)*
Douglas Oard *(University of Maryland, USA)*
Maarten de Rijke *(Universiteit van Amsterdam, The Netherlands)*

2. Queries and Query Analysis
Eugene Agichtein *(Emory University, USA)*
Nick Craswell *(Microsoft Research, USA)*
Evgeniy Gabrilovich *(Yahoo! Research, USA)*
Rosie Jones *(Akamai, USA)*

3. Users and Interactive IR
Nick Belkin *(Rutgers, The State University of New Jersey, USA)*
Susan Dumais *(Microsoft Research, USA)*
Diane Kelly *(University of North Carolina, Chapel Hill, USA)*
Ryen White *(Microsoft Research, USA)*

4. Retrieval Models and Ranking
Norbert Fuhr *(Universitat Duisburg-Essen, Germany)*
Djoerd Hiemstra *(University of Twente, The Netherlands)*
Oren Kurland *(Technion, Israel)*
Hang Li *(Microsoft Research Asia, China)*
Tie-Yan Liu *(Microsoft Research Asia, China)*
Donald Metzler *(USC/ISI, USA)*
Luo Si *(Purdue University, USA)*
ChengXiang Zhai *(University of Illinois at Urbana-Champaign, USA)*

5. Search Engine Architectures and Scalability
Alistair Moffat *(University of Melbourne, Australia)*
Fabrizio Silvestri *(ISTI-CNR, Italy)*

6. Filtering and Recommendion
David Carmel *(IBM Research, Israel)*
John Riedl *(University of Minnesota, USA)*
Yi Zhang *(UC Santa Cruz, USA)*

7. Evaluation
Ben Carterette *(University of Delaware, USA)*
Mark Sanderson *(RMIT, Australia)*
Ian Soboroff *(NIST, USA)*
Justin Zobel *(University of Melbourne, Australia)*

8. Web IR and Social Media Search
Soumen Chakrabarti *(Indian Institute of Technology, Bombay, India)*
Yoelle Maarek *(Yahoo! Research, Israel)*
Marc Najork *(Microsoft Research, USA)*
Iadh Ounis *(University of Glasgow, UK)*
Berthier Ribeiro-Neto *(Federal University of Minas Gerais, Brazil)*

9. IR and Structured Data
Mounia Lalmas *(Yahoo! Research, Spain)*
Andrew Trotman *(University of Otago, New Zealand)*
Arjen de Vries *(Centrum Viskunde & Informatica and Delft University of Technology, The Netherlands)*

10. Multimedia IR
Xian-Sheng Hua *(Microsoft Research Asia, China)*
R. Manmatha *(University of Massachusetts Amherst, USA)*

11. Other Applications
Charlie Clarke *(University of Waterloo, Canada)*
Edward Fox *(Virginia Tech, USA)*

Program Committee

Maristella Agosti *(University of Padua, Italy)*
Omar Alonso *(Microsoft, USA)*
Ismail Altingovde *(Bilkent University, Turkey)*
Gianni Amati *(Fondazione Ugo Bordoni, Italy)*
Avi Arampatzis *(Democritus University of Thrace, Greece)*
Paavo Arvola *(University of Tampere Finland, Finland)*
Javed Aslam *(Northeastern University, USA)*
Leif Azzopardi *(University of Glasgow, United Kingdom)*
Ricardo Baeza-Yates *(Yahoo! Research, Spain)*
Peter Bailey *(Microsoft, USA)*
Niranjan Balasubramanian *(University of Massachusetts Amherst, USA)*
Timothy Baldwin *(University of Melbourne, Australia)*
Krisztian Balog *(Norwegian University of Science and Tech., Norway)*
Roberto Basili *(University of Roma, Tor Vergata, Italy)*
Steve Beitzel *(Telcordia Technologies, USA)*
Michael Bendersky *(University of Massachusetts Amherst, USA)*
Paul Bennett *(Microsoft, USA)*
Mikhail Bilenko *(Microsoft Research, USA)*
Gloria Bordogna *(National Research Council of Italy,Italy)*
Mohand Boughanem *(IRIT, University Paul Sabatier, Toulouse, France)*
Giorgio Brajnik *(University of Udine, Italy)*
Martin Braschler *(Zurich University of Applied Sciences, Switzerland)*
Eric Brown *(IBM Research, USA)*
Peter Bruza *(Queensland University of Technology, Australia)*
Chris Buckley *(Sabir Research, USA)*
Stefan Buettcher *(Google, USA)*
Fazli Can *(Bilkent University, Turkey)*
Guihong Cao *(Microsoft, USA)*
Robert Capra *(University of North Carolina at Chapel Hill, USA)*
Ben Carterette *(University of Delaware, USA)*
Vitor Carvalho *(Microsoft, USA)*
Carlos Castillo *(Yahoo! Research, Spain)*
Raman Chandrasekar *(Evri, USA)*
Yi Chang *(Yahoo! Labs, USA)*
Kevin Chang *(University of Illinois at Urbana-Champaign, USA)*
Francine Chen *(FX Palo Alto Laboratory, USA)*
Pu-Jen Cheng *(National Taiwan University, Taiwan)*
Jean-Pierre Chevallet *(Université Grenoble, France)*
Abdur Chowdhury *(Twitter, USA)*
Paul Clough *(University of Sheffield, United Kingdom)*
Jack Conrad *(Thomson Reuters, USA)*
Gordon V. Cormack *(University of Waterloo, Canada)*
Hang Cui *(Google, USA)*

Brian Davison *(Lehigh University, USA)*
Debora Donato *(Yahoo! Labs, USA)*
Antoine Doucet *(University of Caen, France)*
Georges Dupret *(Yahoo! Labs, USA)*
Koji Eguchi *(Kobe University, Japan)*
David Eichmann *(U. Iowa, USA)*
Jonathan Elsas *(CMU, USA)*
David Elsweiler *(Friedrich-Alexander Universitaet Erlangen, Germany)*
Jianping Fan *(UNC-Charlotte, USA)*
Hui Fang *(U. of Delaware, USA)*
Henry Feild *(University of Massachusetts Amherst, USA)*
Shaolei Feng *(Siemens Corporate Research Labs, USA)*
Radu Florian *(IBM Research, USA)*
Martin Franz *(IBM, USA)*
Jim French *(Corporation for National Research Initiatives, USA)*
Luanne Freund *(University of British Columbia, Canada)*
Evgeniy Gabrilovich *(Yahoo, USA)*
Patrick Gallinari *(Université Pierre et Marie Curie - Paris 6, France)*
Bin Gao *(Microsoft Research Asia, China)*
Eric Gaussier *(Université Joseph Fourier, France)*
Fredric Gey *(University of California, Berkeley, USA)*
C. Lee Giles *(Pennsylvania State University, USA)*
Ayse Goker *(City University London, United Kingdom)*
Marcos Goncalves *(UFMG, Brazil)*
Gregory Grefenstette *(Exalead, France)*
Jacek Gwizdka *(Rutgers University, USA)*
Sanda Harabagiu *(USA)*
Donna Harman *(NIST, USA)*
David Harper *(Google Switzerland, Switzerland)*
Ahmed Hassan *(University of Michigan Ann Arbor, USA)*
Claudia Hauff *(TU Delft, Netherlands)*
Daqing He *(University of Pittsburgh, USA)*
William Hersh *(Oregon Health and Science University, USA)*
Timo Honkela *(Helsinki University of Technology, Finland)*
Eric Horvitz *(Microsoft Research, USA)*
Xuanjing Huang *(Fudan University, China)*
Jimmy Huang *(York University, Canada)*
Panos Ipeirotis *(New York University, USA)*
CV Jawahar *(IIIT Hyderabad, India)*
Jiwoon Jeon *(Google, USA)*
Daxin Jiang *(Microsoft Research Asia, China)*
Hideo Joho *(University of Tsukuba, Japan)*
Gareth Jones *(Dublin City University, Ireland)*

Joemon Jose *(University of Glasgow, United Kingdom)*
Vanja Josifovski *(Yahoo! Research, USA)*
Marko Junkkari *(University of Tampere Finland, Finland)*
Jaap Kamps *(University of Amsterdam, Netherlands)*
Noriko Kando *(National Institute of Informatics, Japan)*
Evangelos Kanoulas *(University of Sheffield, United Kingdom)*
David Karger *(MIT, USA)*
Jussi Karlgren *(Gavagai, Sweden)*
Gabriella Kazai *(Microsoft Research Cambridge, United Kingdom)*
Weimao Ke *(Drexel University, USA)*
Qifa Ke *(Microsoft Research, USA)*
Marijn Koolen *(University of Amsterdam, Netherlands)*
Yehuda Koren *(Yahoo! Research Israel, Israel)*
Ravi Kumar *(Yahoo! Research, USA)*
Giridhar Kumaran *(Microsoft Corporation, USA)*
Alberto Laender *(Federal University of Minas Gerais, Brazil)*
Wai Lam *(The Chinese University of Hong Kong, Hong Kong)*
Monica Landoni *(USI, Switzerland)*
Guy Lapalme *(Université de Montréal, Canada)*
Birger Larsen *(Royal School of Library and Information Science Denmark, Denmark)*
Ray Larson *(University of California, Berkeley, USA)*
Victor Lavrenko *(University of Edinburgh, United Kingdom)*
Matthew Lease *(University of Texas at Austin, USA)*
Miro Lehtonen *(University of Helsinki, Finland)*
Jochen Leidner *(Thomson Reuters Global Services, Switzerland)*
Anton Leuski *(University of Southern California, USA)*
David Lewis *(David D. Lewis Consulting, USA)*
Xiao Li *(Microsoft Research, USA)*
Zhisheng Li *(National University of Singapore, Singapore)*
Ee-Peng Lim *(Singapore Management University, Singapore)*
Chin-Yew Lin *(Microsoft Research Asia, China)*
Yi Liu *(Google, USA)*
Ting Liu *(Harbin Institute of Technology, China)*
David Losada *(University of Santiago de Compostela, Spain)*
Jie Lu *(IBM Research, USA)*
Robert Luk *(The Hong Kong Polytechnic University, Hong Kong)*
Yuanhua Lv *(UIUC, USA)*
Yoelle Maarek *(Yahoo! Research, Israel)*
Craig Macdonald *(U. Glasgow, United Kingdom)*
Andrew MacFarlane (City University London, United Kingdom)
Prasenjit Majumder *(DAIICT, India)*
Thomas Mandl *(U. Hildesheim, Germany)*
Stephane Marchand-Maillet *(University of Geneva - CS Dept - Viper group, Switzerland)*
James Mayfield *(JHU HLT/COE, USA)*

Paul McNamee *(Johns Hopkins University, USA)*
Tao Mei *(Microsoft Research Asia, China)*
Massimo Melucci *(University of Padua, Italy)*
Natasa Milic-Frayling *(Microsoft Research Ltd, United Kingdom)*
Zhao-yan Ming *(National University of Singapore, Singapore)*
Mandar Mitra *(Indian Statistical Institute, India)*
Vibhu Mittal *(Root-1 Research, USA)*
Marie-Francine Moens *(Katholieke Universiteit Leuven, Belgium)*
Alistair Moffat *(University of Melbourne, Australia)*
Alessandro Moschitti *(University of Trento, Italy)*
Josiane Mothe *(IRIT, Université de Toulouse, France)*
Isabelle Moulinier *(Thomson Reuters, USA)*
Philippe Mulhem *(Laboratoire d'Informatique de Grenoble, France)*
Vanessa Murdock *(Yahoo! Research, Spain)*
Sung HYon Myaeng *(Korea Advanced Institute of Science and Tech, South Korea)*
Gonzalo Navarro *(University of Chile, Chile)*
Michael Nelson *(Old Dominion University, USA)*
Hwee Tou Ng *(National University of Singapore, Singapore)*
Charles Nicholas *(USA)*
Alexandros Ntoulas *(Microsoft Research, USA)*
Giorgio Maria Di Nunzio *(University of Padua, Italy)*
Paul Ogilvie *(LinkedIn, USA)*
Marius Pasca *(Google Inc., USA)*
Gabriella Pasi *(University of Milano-Bicocca, Italy)*
Virgil Pavlu *(Northeastern University, USA)*
Jan Pedersen *(Microsoft Corp., USA)*
Fuchun Peng *(Microsoft, USA)*
Jeremy Pickens *(Catalyst Repository Systems, USA)*
Ari Pirkola *(University of Tampere Finland, Finland)*
Benjamin Piwowarski *(University of Glasgow, United Kingdom)*
John Prager *(IBM Research, USA)*
Tao Qin *(Microsoft Research Asia, China)*
Filip Radlinski *(Microsoft, Canada)*
Hema Raghavan *(IBM Research, USA)*
Vijay Raghavan *(USA)*
Edie Rasmussen *(University of British Columbia, Canada)*
Andreas Rauber *(Vienna University of Technology, Austria)*
Matthew Richardson *(Microsoft Research, USA)*
Soo Young Rieh *(University of Michigan, USA)*
Hae-Chang Rim *(Korea University, South Korea)*
Thomas Roelleke *(Queen Mary University of London, United Kingdom)*
Dmitri Roussinov *(University of Strathclyde, United Kingdom)*
Stefan Rueger *(United Kingdom)*
Tetsuya Sakai *(Microsoft Research Asia, China)*
Ralf Schenkel *(Saarland University, Germany)*

Hinrich Schuetze *(University of Stuttgart, Germany)*
Jangwon Seo *(University of Massachusetts Amherst, USA)*
Chirag Shah *(Rutgers University, USA)*
Rao Shen *(Yahoo!, USA)*
Jialie Shen *(Singapore Management University, Singapore)*
Dou Shen *(Buzz Labs, USA)*
Milad Shokouhi *(Microsoft, United Kingdom)*
David Smith *(University of Massachusetts Amherst, USA)*
Mark Smucker *(University of Waterloo, Canada)*
Aya Soffer *(IBM, Israel)*
Dawei Song *(The Robert Gordon University, United Kingdom)*
Ruihua Song *(Microsoft Research Asia, China)*
Eero Sormunen *(Finland)*
Torsten Suel *(Polytechnic Institute of NYU, USA)*
Krysta Svore *(Microsoft Research, USA)*
John Tait *(Information Retrieval Facility, Austria)*
Simone Teufel *(University of Cambridge, United Kingdom)*
Martin Theobald *(Max-Planck Institute Informatics, Germany)*
James Thom *(RMIT, Australia)*
Paul Thomas *(CSIRO, Australia)*
Paul Thompson *(Dartmouth College, USA)*
Anastasios Tombros *(Queen Mary University of London, United Kingdom)*
Elaine Toms *(Dalhousie University, Canada)*
Andrew Turpin *(University of Melbourne, Australia)*
Howard Turtle *(Syracuse University, USA)*
Antti Ukkonen *(Yahoo! Research Barcelona, Spain)*
Pertti Vakkari *(University of Tampere, Finland)*
Olga Vechtomova *(University of Waterloo, Canada)*
Nina Wacholder *(Rutgers University, USA)*
Jingdong Wang *(Microsoft Research Asia, China)*
Meng Wang *(National University of Singapore, Singapore)*
Kai Wang *(Institute for Infocom Research (I2R), Singapore)*
Qiuyue Wang *(Renmin Univ. of China, China)*
William Webber *(University of Melbourne, Australia)*
Ingmar Weber *(Yahoo! Research Barcelona, Spain)*
Xing Wei *(Microsoft Corp., USA)*
Gerhard Weikum *(Max-Planck Institute for Informatics, Germany)*
Thijs Westerveld *(Teezir Search Solutions, Netherlands)*
Gu Xu *(Microsoft Research Asia, China)*
Xiaobing Xue *(University of Massachusetts, USA)*
Yiming Yang *(Carnegie Mellon University, USA)*
Linjun Yang *(Microsoft Research Asia, China)*
Xing Yi *(University of Massachusetts Amherst, USA)*
Emine Yilmaz *(Microsoft Research Cambridge, United Kingdom)*
Elad Yom-Tov *(Yahoo! Research, USA)*
Masaharu Yoshioka *(Hokkaido University, Japan)*

Markus Zanker *(Alpen-Adria-Universitaet Klagenfurt, Austria)*
Lei Zhang *(Microsoft Research Asia, China)*
Shenghuo Zhu *(NEC Labs America, USA)*
Nivio Ziviani *(Federal University of Minas Gerais, Brazil)*

SIGIR 2011 Sponsors and Supporters

Sponsors

SIGIR Special Interest Group on Information Retrieval

Association for Computing Machinery

Advancing Computing as a Science & Profession

Corporate Supporters

Diamond

Microsoft

Gold

Baidu 百度 EMC² Google

SOSO 搜搜 YAHOO!

Silver

ebay Research Labs

Bronze

HP Labs IBM Research Yandex

Associated Organization

CIPS

Also Supported by

Université de Montréal

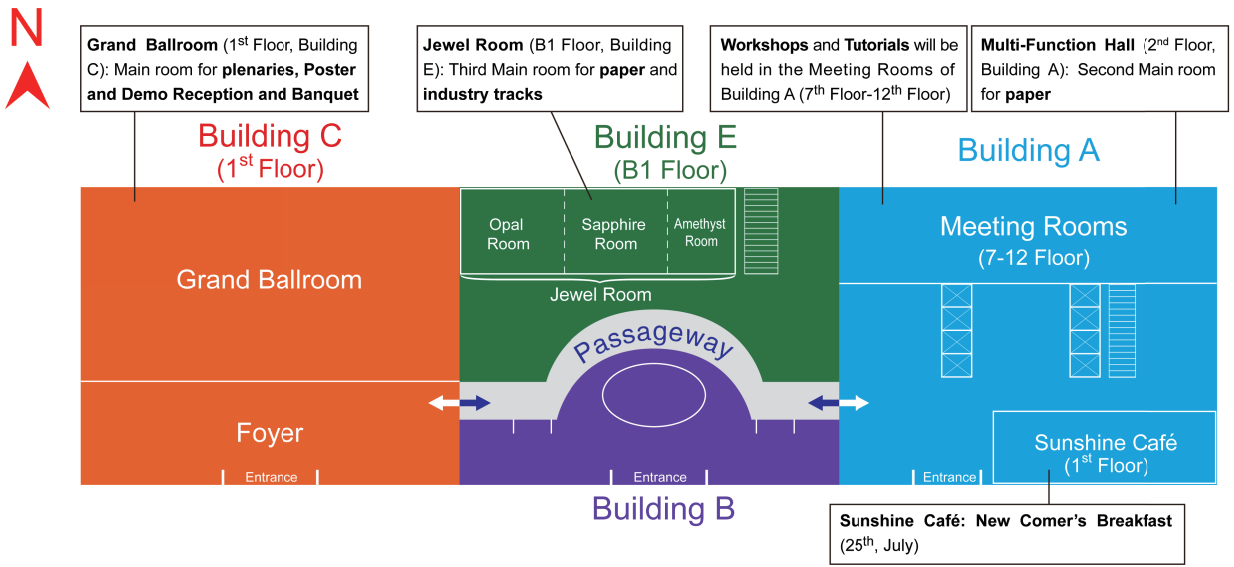
Exhibitors

MORGAN & CLAYPOOL PUBLISHERS

now the essence of knowledge

Springer

Layout for Hotel



Social Program

Welcome Reception
July 24th, Sunday, 18:00-22:00
The Forbidden Terrace, 19th Floor, Building A

Newcomer's Breakfast
July 25th, Monday, 07:00-08:30
Sunshine Cafe

Poster and Demo Reception
July 25th, Monday, 19:00-21:00
Grand Ballroom, 1st Floor, Building C

Banquet
July 26th, Tuesday, 19:00-23:00
Grand Ballroom, 1st Floor, Building C

Microsoft Research Asia Open House and Reception
July 27th, Wednesday, 19:00-21:30, MSR Asia Office
Free shuttle bus will be provided between Beijing Hotel and MSR Asia Office

July 24, 2011 Sunday			
	Tutorials		***Doctoral Consortium***
Rooms	**Meeting Rooms, Building A**		**Meeting Room, Building A**
08:30-10:00	7024, 7026, 8024, 8026		1024
10:00-10:30	Coffee Break		
10:30-12:15	7024, 7026, 8024, 8026		1024
12:15-13:30	Lunch on your own		
13:30-15:15	7024, 7026, 8024, 8026		1024
15:15-15:45	Coffee Break		
15:45-17:15	7024, 7026, 8024, 8026		1024
18:00-20:00	**Welcome reception** Forbidden Terrace, 19th Floor, Building A		
July 25, 2011 Monday			
	Scientific Papers		
Rooms	**Grand Ballroom** **1st Floor, Building C**	**Multi-Function Hall** **2nd Floor, Building A**	**Jewel Room** **B1th Floor, Building E**

07:00-08:30	New Comer's breakfast		
08:30-10:00	**Opening ceremony** **Keynote speech - Qi Lu (Microsoft)** Grand Ballroom, 1st Floor, Building C		
10:00-10:30	Coffee break		
10:30-12:10	Users I	Query Analysis I	Learning to Rank
12:10-14:00	Lunch on your own		
14:00-15:15	Personalization	Retrieval models I	Social Media
15:15-15:45	Coffee break		
15:45-17:25	Content Analysis	Web IR	Collaborative filtering I
18:30-22:00	**Poster-Demo reception** Grand Ballroom, 1st Floor, Building C		
July 26, 2011 Tuesday			
	Scientific Papers		
Rooms	**Grand Ballroom** **1st Floor, Building C**	**Multi-Function Hall** **2nd Floor, Building A**	**Jewel Room** **B1th Floor, Building E**

08:45-10:00	**Keynote speech - ChengXiang Zhai (UIUC)** Grand Ballroom, 1st Floor, Building C			
10:00-10:30	Coffee break			
10:30-12:10	Users II	Query Analysis II	Communities	
12:10-14:00	Lunch on your own			
14:00-15:15	Classification	Retrieval models II	Image Search	
15:15-15:45	Coffee break			
15:45-17:25	Indexing	Web Queries	Collaborative filtering II	
19:00-22:00	**Banquet** Grand Ballroom, 1st Floor, Building C			
July 27, 2011 Wednesday				
	Scientific Papers			***Industrial track***
Rooms	**Grand Ballroom 1st Floor, Building C**	**Multi-Function Hall 2nd Floor, Building A**	**Opal Room B1 Floor, Building E**	**Sapphire & Amethyst Room B1 Floor, Building E**
08:30-09:45	Latent Semantic Analysis	Multimedia IR	Summarization	Opening and Keynote

09:45-10:00	Coffee break			
10:00-11:40	Vertical & Entity Search	Query suggestions	Linguistic Analysis	Keynote and Session
11:45-14:00	SIGIR Business Meeting (Lunch box provided)			
14:00-15:15	Clustering	Effectiveness	Multilingual IR	Keynote and Session
15:15-15:30	Coffee break			
15:30-17:10	Efficiency	Recommender systems	Test collections	Keynote and Session
17:10-17:30	Closing session			
17:30-21:00	Reception at Microsoft Research Asia (Free shuttle buses will leave from the conference hotel at 17:30-18:00, and the reception will start at 18:30.)			
July 28, 2011 Thursday				
	Workshops			
Rooms	**Meeting Rooms, Building A**			
08:30-17:30	Workshops			
09:00-17:30	Industry Lab Tour			

Doctoral Consortium

Sunday, July 24th, Meeting Room, Building A

- Multidimensional Search Result Diversification: Diverse Search Results for Diverse Users
  Sumit Bahta (Pennsylvania State University)

- Using k-top Retrieved Web Snippets to Date Temporal Implicit Queries based on Web Content Analysis
  Ricardo Nuno Taborda Campos (University of Porto)

- Improving Query and Result List Adaptation in Personalized Multilingual Information Retrieval
  M. Rami Ghorab (Trinity College Dublin)

- Search Engines that Learn Online
  Katja Hofmann (University of Amsterdam)

- Efficient and Effective Solutions for Search Engines
  Xiang-Fei Jia (University of Otago)

- Query Expansion Based on a Semantic Graph Model
  Xue Jiang (Chinese Academy of Sciences)

- Domain-Specific Information Retrieval Using Recommenders
  Wei Li (Dublin City University)

- Modeling Document Scores for Distributed Information Retrieval
  Ilya Markov (University of Lugano)

- Descriptive Modelling of Text Classification and its Integration with other IR Tasks
  Miguel Martinez-Alvarez (Queen Mary, University of London)

- Persistence in the Ephemeral: Utilizing Repeat Behaviors for Multi-Session
  Personalized Search Sarah K Tyler (University of California, Santa Cruz)

- Understanding and Using Contextual Information in Recommender Systems
  Licai Wang (Beijing University of Posts and Telecommunications)

Tutorials

Machine Learning for IR: Core Learning Technologies for Information Retrieval
Half-day – Morning Room 7024
Luo Si, Rong Jin

Enhancing Web Search by Mining Search and Browse Logs
Half-day – Morning Room 8024
Daxin Jiang, Jian Pei, Hang Li

A New Look at Old Tricks: The Fertile Roots of Current Research
Half-day – Morning Room 7026
Paul Kantor

Crowdsourcing for Information Retrieval: Principles, Methods, and Applications
Half-day – Morning Room 8026
Omar Alonso, Matthew Lease

Machine Learning for IR: Emerging Learning Technologies for Information Retrieval
Half-day – Afternoon Room 7024
Luo Si, Rong Jin

Practical Online Retrieval Evaluation
Half-day – Afternoon Room 8024
Filip Radlinski, Yisong Yue

Web Retrieval: The Role of Users
Half-day – Afternoon Room 7026
Ricardo Baeza-Yates, Yoelle Maarek

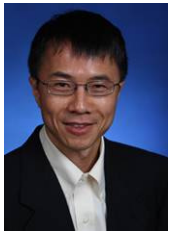
Information Organization and Retrieval with Collaboratively Generated Content
Half-day – Afternoon Room 8026
Eugene Agichtein, Evgeniy Gabrilovich

Keynote Addresses

July 25th, Monday, 08:30-10:00
Grand Ballroom, 1st Floor, Building C

Future of the Web and Search
Qi Lu, President of Microsoft's Online Services Division

No one doubts that we have only scratched the surface of what is possible with the Web. The day is coming fast when the Web will become almost a virtual mind reader. Your intent, interests, and needs will be instantly perceived and the information you want will be promptly delivered -- whether you ask for it directly or not -- based on a deep understanding of the meaning of words in your query, knowledge of your preferences and patterns, what others have done before you, your location, and more. In this talk, I will share some of my thoughts about where the Web is heading and how search will be transformed to align to this new Web, laying out some specifics behind Microsoft's vision to empower people with knowledge.



As president of Microsoft's Online Services Division (OSD), Dr. Qi Lu leads the company's search and online advertising efforts. Dr. Lu oversees the OSD Research & Development team which has responsibility for the evolution of Microsoft's search, portal and advertising services; the Online Audience Business Group; and the Advertiser and Publisher Solutions Business Group. Dr. Lu reports to Microsoft chief executive officer Steve Ballmer.

Prior to joining Microsoft, Dr. Lu spent 10 years as a Yahoo! senior executive. His roles included serving as the executive vice president of engineering for the company's Search and Advertising Technology Group where he oversaw the development of Yahoo!'s Web search and monetization platforms and vice president of engineering responsible for the technology development of Yahoo!'s search, e-commerce and local listings of businesses and products.

Before joining Yahoo!, Dr. Lu worked as a research staff member at IBM's Almaden Research Center and Carnegie Mellon University and was a faculty member at Fudan University in China. He received his bachelor of science and master of science in computer science from Fudan University and his Ph.D. in computer science from Carnegie Mellon University. Dr. Lu holds 20 U.S. patents.

July 26th, Tuesday, 08:30-10:00
Grand Ballroom, 1st Floor, Building C

Beyond Search: Statistical Topic Models for Text Analysis
ChengXiang Zhai, Department of Computer ScienceUniversity of Illinois at Urbana-Champaign, USA.

Search is generally a means to the end of finishing a task. While the current search engines are useful to users for finding relevant information, they offer little help to users for further digesting and analyzing the overwhelming found information needed for finishing a complex task. In this talk, I will discuss how statistical topic models can be used to help users analyze and digest the found relevant information and turn search results into actionable knowledge needed to complete a task. I will present several general statistical topic models for extracting and analyzing topics and their patterns in text, and show sample applications of such models in tasks such as opinion integration, comparative summarization, contextual topic trend analysis, and event impact analysis. The talk will conclude with a discussion of novel challenges raised in extending a search engine to an analysis engine that can go beyond search to provide more complete support for users to finish their tasks.

ChengXiang Zhai is an Associate Professor of Computer Science at the University of Illinois at Urbana-Champaign, where he also holds a joint appointment at the Graduate School of Library and Information Science, Institute for Genomic Biology, and Department of Statistics. He received a Ph.D. in Computer Science from Nanjing University in 1990, and a Ph.D. in Language and Information Technologies from Carnegie Mellon University in 2002. He worked at Clairvoyance Corp. as a Research Scientist and a Senior Research Scientist from 1997 to 2000. His research interests include information retrieval, text mining, natural language processing, machine learning, and bioinformatics. He is an Associate Editor of ACM Transactions on Information Systems, and Information Processing and Management, and serves on the editorial board of Information Retrieval Journal. He is a program co-chair of ACM CIKM 2004, NAACL HLT 2007, and ACM SIGIR 2009. He is an ACM Distinguished Scientist and the recipient of an Alfred P. Sloan Research Fellowship, IBM Faculty Award, ACM SIGIR 2004 Best Paper Award, and a Presidential Early Career Award for Scientists and Engineers (PECASE).

Detailed Scientific Program

July 25th, Monday

08:30-10:00 Keynote Address Grand Ballroom, 1st Floor, Building C
Future of the Web and Search
Qi Lu (Microsoft)

10:00-10:30 Coffee Break

10:30-12:10
Session 1A: Users I
Grand Ballroom, 1st Floor, Building C
Chair: Nicholas Belkin

- **Modeling and Analysis of Cross-Session Search Tasks.**
  Alexander Kotov (University of Illinois at Urbana-Champaign), Paul N. Bennett, Ryen W. White, Susan T. Dumais, Jaime Teevan (Microsoft Research)

- **The Economics in Interactive Information Retrieval.**
  Leif Azzopardi (University of Glasgow)

- **Seeding Simulated Queries with User-study Data for Personal Search Evaluation.**
  David Elsweiler (University of Erlangen), David E. Losada, José C. Toucedo, Ronald T. Fernandez (Universidad de Santiago de Compostela)

- **Understanding Re-finding Behaviour in Naturalistic Email Interaction Logs.**
  David Elsweiler (University of Erlangen), Morgan Harvey (University of Strathclyde), Martin Hacker (University of rlangen)

Session 1B: Query Analysis I
Multi-Function Hall, 2nd Floor, Building A
Chair: Ben Carterette

- **People Searching for People: Analysis of a People Search Engine Log.**
  Wouter Weerkamp, Richard Berendsen, Bogomil Kovachev, Edgar Meij (University of Amsterdam), Krisztian Balog (NTNU Trondheim), Maarten de Rijke (University of Amsterdam)

- **Learning Search Tasks in Queries and Web Pages via Graph Regularization.**
  Ming Ji (University of Illinois at Urbana-Champaign), Jun Yan (Microsoft Research Asia), Siyu Gu (Beijing Institute of Technology), Jiawei Han (University of Illinois at Urbana-Champaign), Xiaofei He (Zhejiang University), Wei Vivian Zhang (Microsoft Corporation), Zheng Chen (Microsoft Research Asia)

- **Intentions and Attention in Exploratory Health Search.**
  Marc-Allen Cartright (University of Massachusetts, Amherst), Ryen W. White, Eric Horvitz (Microsoft Research)

- **User Behavior in Zero-Recall eCommerce Queries.**
  Gyanit Singh, Nish Parikh, Neel Sundaresan (eBay Research Labs)

Session 1C: Learning to Rank
Jewel Room, B1 Floor, Building E
Chair: Tie-Yan Liu

- **Bagging Gradient-Boosted Trees for High Precision, Low Variance Ranking Models.**
  Yasser Ganjisaffar (University of California, Irvine), Rich Caruana (Microsoft Research), Cristina Videira Lopes (University of California, Irvine)

- **Learning to Rank for Freshness and Relevance.**
  Na Dai (Lehigh University), Milad Shokouhi (Microsoft Research), Brian D. Davison (Lehigh University)

- **A Cascade Ranking Model for Efficient Ranked Retrieval.**
  Lidan Wang, Jimmy Lin (University of Maryland), Donald Metzler (University of Southern California)

- **Relevant Knowledge Helps in Choosing Right Teacher: Active Query Selection for Ranking Adaptation.**
  Peng Cai (East China Normal University), Wei Gao (The Chinese University of Hong Kong), Aoying Zhou (East China Normal University), Kam-Fai Wong (The Chinese University of Hong Kong & Ministry of Education)

12:10-14:00 **Lunch on your own**

14:00-15:15
Session 2A: Personalization
Grand Ballroom, 1st Floor, Building C
Chair: Maarten de Rijke

- **SCENE : A Scalable Two-Stage Personalized News Recommendation System.**
  Lei Li, Dingding Wang, Tao Li (Florida International University), Daniel Knox, Balaji Padmanabhan (DailyMe, Inc.)

- **Inferring and Using Location Metadata to Personalize Web Search.**
  Paul N. Bennett (Microsoft Research), Filip Radlinski (Microsoft), Ryen W. White (Microsoft), Emine Yilmaz (Microsoft)

- **Active Learning to Maximize Accuracy vs. Effort in Interactive Information Retrieval.**
  Aibo Tian, Matthew Lease (The University of Texas at Austin)

Session 2B: Retrieval Models I
Multi-Function Hall, 2nd Floor, Building A
Chair: Oren Kurland

- **CRTER: Using Cross Terms to Enhance Probabilistic Information Retrieval.**
  Jiashu Zhao, Jimmy Xiangji Huang, Ben He (York University)

- **A Boosting Approach to Improving Pseudo-Relevance Feedback.**
  Yuanhua Lv, ChengXiang Zhai (University of Illinois at Urbana-Champaign), Wan Chen (Wolfram Research, Inc.)

- **Enhancing Ad-hoc Relevance Weighting Using Probability Density Estimation.**
  Xiaofeng Zhou, Jimmy Xiangji Huang, Ben He (York University)

Session 2C: Social Media
Jewel Room, B1 Floor, Building E
Chair: Mounia Lalmas

- **Who Should Share What? Item-level Social Influence Prediction for Users and Posts Ranking.**
  Peng Cui, Fei Wang (IBM T.J. Watson Research Center), Shaowei Liu, Mingdong Ou, Shiqiang Yang, Lifeng Sun (Tsinghua University)

- **Mining Tags Using Social Endorsement Networks.**
  Theodoros Lappas (University of California, Riverside), Kunal Punera, Tamas Sarlos (Yahoo! Research)

- **Crowdsourcing for Book Search Evaluation: Impact of HIT Design on Comparative System Ranking.**
  Gabriella Kazai (Microsoft Research), Jaap Kamps, Marijn Koolen (University of Amsterdam), Natasa Milic-Frayling (Microsoft Research)

15:15-15:45 **Coffee Break**

15:45-17:25
Session 3A: Content Analysis
Grand Ballroom, 1st Floor, Building C
Chair: Jamie Callan

- **A Site Oriented Method For Segmenting Web Pages.**
  David Fernandes, Edleno S. de Moura, Altigran S. da Silva, Berthier Ribeiro-Neto, Edisson Braga (Federal University of Amazonas)

- **Composite Hashing with Multiple Information Sources.**
  Dan Zhang (Purdue University), Fei Wang (IBM T.J. Watson Research Lab), Luo Si (Purdue University)

- **Detecting Outlier Sections in US Congressional Legislation.**
  Elif Aktolga (University of Massachusetts, Amherst), Irene Ros, Yannick Assogba (IBM T.J. Watson Research Center)

- **DOM Based Content Extraction via Text Density.**
  Fei Sun, Dandan Song, Lejian Liao (Beijing Institute of Technology)

Session 3B: Web IR
Multi-Function Hall, 2nd Floor, Building A
Chair: Ricardo Baeza-Yates

- **Social Context Summarization.**
  Zi Yang (Tsinghua University), Keke Cai (IBM, China Research Lab), Jie Tang (Tsinghua University), Li Zhang, Zhong Su (IBM, China Research Lab), Juanzi Li (Tsinghua University)

- **Probabilistic Factor Models for Web Site Recommendation.**
  Hao Ma, Chao Liu (Microsoft Research), Irwin King (AT&T Labs Research & The Chinese University of Hong Kong), Michael R. Lyu (The Chinese University of Hong Kong)

- **Efficiently Collecting Relevance Information from Clickthroughs for Web Retrieval System Evaluation.**
  Jing He, Wayne Xin Zhao, Baihan Shu (Peking University), Xiaoming Li (Beihang University), Hongfei Yan (Peking University)

- **Unsupervised Query Segmentation Using Clickthrough for Information Retrieval.**
  Yanen Li (University of Illinois at Urbana-Champaign), Bo-Jun Paul Hsu, ChengXiang Zhai, Kuansan Wang (Microsoft Research)

Session 3C: Collaborative Filtering I
Jewel Room, B1 Floor, Building E
Chair: David Carmel

- **Collaborative Competitive Filtering: Learning Recommender Using Context of User Choices.**
  Shuang-Hong Yang, Bo Long, Alexander J. Smola (Yahoo! Research), Hongyuan Zha (Georgia Institute of Technology), Zhaohui Zheng (Yahoo! Labs Beijing)

- **CLR: A Collaborative Location Recommendation Framework based on Co-Clustering.**

Kenneth Wai-Ting Leung, Dik Lun Lee (Hong Kong University of Science and Technology), Wang-Chien Lee (The Pennsylvania State University)

- **Functional Matrix Factorizations for Cold-Start Recommendation.**
  Ke Zhou, Shuang-Hong Yang, Hongyuan Zha (Georgia Institute of Technology)

- **Exploiting Geographical Influence for Collaborative Point-of-Interest Recommendation.**
  Mao Ye, Peifeng Yin, Wang-Chien Lee (HKUST & The Pennsylvania State University), Dik-Lun Lee (HKUST)

18:30-22:00 Poster and Demo Session
Grand Ballroom, 1st Floor, Building C

July 26th, Tuesday

08:45-10:00 Keynote Address
Grand Ballroom, 1st Floor, Building C
Beyond Search: Statistical Topic Models for Text Analysis
ChengXiang Zhai (University of Illinois at Urbana-Champaign)

10:00-10:30 Coffee Break

10:30-12:10
Session 4A: Users II
Grand Ballroom, 1st Floor, Building C
Chair: Susan Dumais

- **Why Searchers Switch: Understanding and Predicting Engine Switching Rationales.**
  Qi Guo (Emory University), Ryen W. White, Yunqiao Zhang, Blake Anderson, Susan T. Dumais (Microsoft Corporation)

- **Find It If You Can: A Game for Modeling Different Types of Web Search Success Using Interaction Data.**
  Mikhail Ageev (Moscow State University), Qi Guo, Dmitry Lagun, Eugene Agichtein (Emory University)

- **Measuring Improvement in User Search Performance Resulting From Optimal Search Tips.**
  Neema Moraveji (Stanford University), Daniel Russell (Google), Jacob Bien (Stanford University), David Mease (Google)

- **ViewSer: Enabling Large-Scale Remote User Studies of Web Search Examination and Interaction.**
  Dmitry Lagun, Eugene Agichtein (Emory University)

Session 4B: Query Analysis II
Multi-Function Hall, 2nd Floor, Building A
Chair: Bruce Croft

- **CrowdLogging: Distributed, Private, and Anonymous Search Logging.**
  Henry Field, James Allan, Joshua Glatt (University of Massachusetts, Amherst)

- **Out of Sight, Not Out of Mind.**
  Elad Yom-Tov, Fernando Diaz (Yahoo Research)

- **Scalable Multi-Dimensional User Intent Identification using Tree Structured Distributions.**
  Vinay Jethava (Chalmers University), Liliana Calderon-Benavides (UNAB), Ricardo Baeza-Yates (Yahoo! Research), Chiranjib Bhattacharyya (Indian Institute of Science), Devdatt Dubhashi (Chalmers University)

- **Social Annotation in Query Expansion: A Machine Learning Approach.**
  Yuan Lin, Hongfei Lin, Song Jin, Zheng Ye (Dalian University of Technolog)

Session 4C: Communities
Jewel Room, B1 Floor, Building E
Chair: Iadh Ounis

- **Predicting Web Searcher Satisfaction with Existing Community-based Answers.**
  Qiaoling Liu, Eugene Agichtein (Emory University), Gideon Dror, Evgeniy Gabrilovich, Yoelle Maarek, Dan Pelleg, Idan Szpektor (Yahoo! Research)

- **Competition-based User Expertise Score Estimation.**
  Jing Liu (Harbin Institute of Technology), Young-In Song, Chin-Yew Lin (Microsoft Research Asia)

- **Learning Online Discussion Structures by Conditional Random Fields.**
  Hongning Wang, Chi Wang, ChengXiang Zhai, Jiawei Han (University of Illinois at Urbana-Champaign)

- **Mining Topics on Participations for Community Discovery.**
  Guoqing Zheng, inwen Guo, Lichun Yang, Shengliang Xu (Shanghai Jiao Tong University), Shenghua Bao, Zhong Su (IBM China Research Laboratory), Dingyi Han, Yong Yu (Shanghai Jiao Tong University)

12:10-14:00 **Lunch on your own**

14:00-15:15
Session 5A: Classification
Grand Ballroom, 1st Floor, Building C
Chair: Arjen P. de Vries

- **Authorship Classification: A Discriminative Syntactic Tree Mining Approach.**
  Sangkyum Kim, Hyungsul Kim, Tim Weninger, Jiawei Han, Hyun Duk Kim (University of Illinois at Urbana-Champaign)

- **On Theme Location Discovery for Travelogue Services.**
  Mao Ye (Microsoft Research Asia & The Pennsylvania State University), Rong Xiao (Microsoft Research Asia), Wang-Chien Lee (The Pennsylvania State University), Xing Xie (Microsoft Research Asia)

- **Effective Sentiment Stream Analysis with Self-Augmenting Training and Demand-Driven Projection.**
  Ismael S. Silva, Janaína Gomide, Adriano Veloso, Wagner Meira Jr., Renato Ferreira (Federal University of Minas Gerais)

Session 5B: Retrieval Models II
Multi-Function Hall, 2nd Floor, Building A
Chair: Donald Metzler

- **Hypergeometric Language Models for Republished Article Finding.**
  Manos Tsagkias, Maarten de Rijke, Wouter Weerkamp (University of Amsterdam)

- **Estimation Methods for Ranking Recent Information.**
  Miles Efron (University of Illinois), Gene Golovchinsky (FX Palo Alto Laboratory, Inc.)

- **Query by Document via a Decomposition-Based Two-Level Retrieval Approach.**
  Linkai Weng (Tsinghua University & Microsoft Research Asia), Zhiwei Li, Rui Cai (Microsoft Research Asia), Yaoxue Zhang, Yuezhi Zhou (Tsinghua University), Laurence T. Yang (St. Francis Xavier University), Lei Zhang (Microsoft Research Asia)

Session 5C: Image Search
Jewel Room, B1 Floor, Building E
Chair: Lei Zhang

- **Integrating Hierarchical Feature Selection and Classifier Training for Multi-Label Image Annotation.**
  Cheng Jin (Fudan University), Chunlei Yang (UNC-Charlotte)

- **Efficient Manifold Ranking for Image Retrieval.**
  Bin Xu, Jiajun Bu, Chun Chen, Deng Cai, Xiaofei He (Zhejiang University), Wei Liu (Columbia University), Jiebo Luo (Eastman Kodak Company)

- **Mining Weakly Labeled Web Facial Images for Search-based Face Annotation.**
  Dayong Wang, Steven C. H. Hoi, Ying He (Nanyang Technological University)

15:15-15:45 **Coffee Break**

15:45-17:25
Session 6A: Indexing
Grand Ballroom, 1st Floor, Building C
Chair: Alistair Moffat

- **Temporal Index Sharding for Space-Time Efficiency in Archive Search.**
  Avishek Anand (Max-Planck Institute for Informatics), Srikanta Bedathur (IIIT - Delhi), Klaus Berberich (Max-Planck Institute for Informatics), Ralf Schenkel (Saarland University)

- **Inverted Indexes for Phrases and Strings.**
  Manish Patil, Sharma V. Thankachan, Rahul Shah (Louisiana State University), Wing-Kai Hon (National Tsing Hua University), Jeffrey Scott Vitter (The University of Kansas), Sabrina Chandrasekaran (Louisiana State University)

- **Faster Temporal Range Queries over Versioned Text.**
  Jinru He, Torsten Suel (Polytechnic Institute of New York University)

- **Indexing Strategies for Graceful Degradation of Search Quality.**
  Shuai Ding (Polytechnic Institute of New York University), Sreenivas Gollapudi, Samuel Ieong, Krishnaram Kenthapadi, Alexandros Ntoulas (Microsoft Research)

Session 6B: Web Queries
Multi-Function Hall, 2nd Floor, Building A
Chair: Nick Craswell

- **Incremental Diversification for Very Large Sets: a Streaming-based Approach.**
  Enrico Minack, Wolf Siberski, Wolfgang Nejdl (Leibniz Universität Hannover)

- **Intent-Aware Search Result Diversification.**
  Rodrygo L. T. Santos, Craig Macdonald, Iadh Ounis (University of Glasgow)

- **Parameterized Concept Weighting in Verbose Queries.**
  Michael Bendersky (University of Massachusetts), Donald Metzler (University of Southern California), W. Bruce Croft (University of Massachusetts)

- **UPS: Efficient Privacy Protection in Personalized Web Search.**
  Gang Chen, He Bai, Lidan Shou, Ke Chen, Yunjun Gao (Zhejiang University)

Session 6C: Collaborative Filtering II
Jewel Room, B1 Floor, Building E
Chair: Yi Zhang

- **Handling Data Sparsity in Collaborative Filtering using Emotion and Semantic Based Features.**
  Yashar Moshfeghi, Benjamin Piwowarski, Joemon M. Jose (University of Glasgow)

- **Fast Context-aware Recommendations with Factorization Machines.**
  Steffen Rendle (University of Konstanz), Zeno Gantner, Christoph Freudenthaler, Lars Schmidt-Thieme (University of Hildesheim)

- **Filtering Semi-Structured Documents Based on Faceted Feedback.**
  Lanbo Zhang, Yi Zhang (University of California, Santa Cruz), Qianli Xing (Tsinghua University)

- **Learning Relevance from Heterogeneous Social Network and Its Application in Online Targeting.**
  Chi Wang (University of Illinois at Urbana-Champaign), Rajat Raina (Facebook), David Fong (Stanford University), Ding Zhou (Facebook), Jiawei Han (University of Illinois at Urbana-Champaign), Greg Badros (Facebook)

19:00-22:00 Banquet
Grand Ballroom, 1st Floor, Building C

July 27th, Wednesday

08:30-09:45
Session 7A: Latent Semantic Analysis
Grand Ballroom, 1st Floor, Building C
Chair: Fernando Diaz

- **ILDA: Interdependent LDA Model for Learning Latent Aspects and their Ratings from Online Product Reviews.**
  Samaneh Moghaddam, Martin Ester (Simon Fraser University)

- **Clickthrough-Based Latent Semantic Models for Web Search.**
  Jianfeng Gao, Kristina Toutanova, Wen-tau Yih (Microsoft Research)

- **Regularized Latent Semantic Indexing.**
  Quan Wang (Peking University), Jun Xu, Hang Li (Microsoft Research Asia), Nick Craswell (Microsoft)

Session 7B: Multimedia IR
Multi-Function Hall, 2nd Floor, Building A
Chair: Tat-Seng Chua

- **Multimedia Answering: Enriching Text QA with Media Information.**
  Liqiang Nie, Meng Wang, Zheng-Jun Zha, Guangda Li, Tat-Seng Chua (National University of Singapore)

- **Enhancing Multi-Label Music Genre Classification Through Ensemble Techniques.**
  Chris Sanden, John Z. Zhang (University of Lethbridge)

- **PICASSO - To Sing, You Must Close Your Eyes and Draw.**
  Aleksandar Stupar, Sebastian Michel (Saarland University)

Session 7C: Summarization
Opal Room, B1 Floor, Building E
Chair: Chin-Yew Lin

- **Enhanced Results for Web Search.**
  Kevin Haas (Microsoft), Peter Mika (Yahoo! Research), Paul Tarjan (Facebook), Roi Blanco (Yahoo! Research)

- **Summarizing the Differences in Multilingual News.**
  Xiaojun Wan, Houping Jia, Shanshan Huang, Jianguo Xiao (Peking University)

- **Evolutionary Timeline Summarization: A Balanced Optimization Framework via Iterative Substitution.**
  Rui Yan, Xiaojun Wan (Peking University), Jahna Otterbacher (Illinois Institute of Technology), Liang Kong, Xiaoming Li, Yan Zhang (Peking University)

09:45-10:00 **Coffee Break**

10:00-11:40
Session 8A: Vertical & Entity Search
Grand Ballroom, 1st Floor, Building C
Chair: Charles Clarke

- **Ranking Related News Predictions.**
  Nattiya Kanhabua (Norwegian University of Science and Technology), Roi Blanco, Michael Matthews (Yahoo! Research)

- **Collective Entity Linking in Web Text: A Graph-Based Method.**
  Xianpei Han, Le Sun (Chinese Academy of Sciences), Jun Zhao (National Laboratory of Pattern Recognition)

- **From One Tree to a Forest: a Unified Solution for Structured Web Data Extraction.**
  Qiang Hao (Microsoft Research Asia & Tianjin University), Rui Cai (Microsoft Research Asia), Yanwei Pang (Tianjin University), Lei Zhang (Microsoft Research Asia)

- **Improving Local Search Ranking through External Logs.**
  Klaus Berberich (Max-Planck Institute for Informatics), Arnd Christian König, Dimitrios Lymberopoulos (Microsoft Research), Peixiang Zhao (University of Illinois at Urbana-Champaign)

Session 8B: Query Suggestions
Multi-Function Hall, 2nd Floor, Building A
Chair: Yoelle Maarek

- **Query Suggestions in the Absence of Query Logs.**
  Sumit Bhatia (The Pennsylvania State University), Debapriyo Majumdar (IBM Research India), Prasenjit Mitra (The Pennsylvania State University)

- **Synthesizing High Utility Suggestions for Rare Web Search Queries.**
  Alpa Jain, Umut Ozertem, Emre Velipasaoglu (Yahoo! Labs)

- **Post-Ranking Query Suggestion by Diversifying Search Results.**
  Yang Song, Dengyong Zhou, Li-wei He (Microsoft Research)

- **Automatic Boolean Query Suggestion for Professional Search.**
  Youngho Kim, Jangwon Seo, W. Bruce Croft (University of Massachusetts, Amherst)

Session 8C: Linguistic Analysis
Opal Room, B1 Floor, Building E
Chair: Noriko Kando

- **Improved Video Categorization from Text Metadata and User Comments.**
  Katja Filippova, Keith B. Hall (Google Inc.)

- **Multifaceted Toponym Recognition for Streaming News.**
  Michael D. Lieberman, Hanan Samet (University of Maryland)

- **Enriching Document Representation via Translation for Improved Monolingual Information Retrieval.**
  Seung-Hoon Na, Hwee Tou Ng (National University of Singapore)

- **A Novel Corpus-Based Stemming Algorithm using Co-occurrence Statistics.**
  Jiaul H. Paik, Dipasree Pal, Swapan K. Parui (Indian Statistical Institute)

11:45-14:00 SIGIR Business Meeting (Lunch box provided)

14:00-15:15
Session 9A: Clustering
Grand Ballroom, 1st Floor, Building C
Chair: Andrew Trotman

- **Document Clustering with Universum.**
  Dan Zhang (Purdue University), Jingdong Wang (Microsoft Research Asia), Luo Si (Purdue University)

- **Identifying Points of Interest by Self-Tuning Clustering.**
  Yiyang Yang, Zhiguo Gong, Leong Hou U (University of Macau)

- **Cluster-Based Fusion of Retrieved Lists.**
  Anna Khudyak Kozorovitsky, Oren Kurland (Technion – Israel Institute of Technology)

Session 9B: Effectiveness
Multi-Function Hall, 2nd Floor, Building A
Chair: Mark Sanderson

- **System Effectiveness, User Models, and User Utility: A Conceptual Framework for Investigation.**
  Ben Carterette (University of Delaware)

- **Evaluating the Synergic Effect of Collaboration in Information Seeking.**
  Chirag Shah, Roberto González-Ibáñez (Rutgers University)

- **Repeatable and Reliable Search System Evaluation using Crowdsourcing.**
  Roi Blanco (Yahoo! Research), Harry Halpin (University of Edinburgh), Daniel M. Herzig (Karlsruhe Institute of Technology), Peter Mika (Yahoo! Research), Jeffrey Pound (University of Waterloo), Henry S. Thompson (University of Edinburgh), Thanh Tran Duc (Karlsruhe Institute of Technology)

Session 9C: Multilingual IR
Opal Room, B1 Floor, Building E
Chair: Doug Oard

- **Cross-Language Web Page Classification via Dual Knowledge Transfer Using Nonnegative Matrix Tri-Factorization.**
  Hua Wang, Heng Huang, Feiping Nie, Chris Ding (University of Texas at Arlington)

- **No Free Lunch: Brute Force vs. Locality-Sensitive Hashing for Cross-lingual Pairwise Similarity.**
  Ferhan Ture (University of Maryland), Tamer Elsayed (King Abdullah University of Science and Technology), Jimmy Lin (University of Maryland)

- **An Event-centric Model for Multilingual Document Similarity.**
  Jannik Strötgen, Michael Gertz, Conny Junghans (Heidelberg University)

15:15-15:30 **Coffee break**

15:30-17:10
Session 10A: Efficiency
Grand Ballroom, 1st Floor, Building C
Chair: Justin Zobel

- **Posting List Intersection on Multicore Architectures.**
  Shirish Tatikonda (IBM Almaden Research), B. Barla Cambazoglu, Flavio P. Junqueira (Yahoo! Research)

- **Timestamp-based Result Cache Invalidation for Web Search Engines.**
  Sadiye Alici (Bilkent University), Ismail Sengor Altingovde (L3S Research Center), Rifat Ozcan (Bilkent University), B. Barla Cambazoglu (Yahoo! Research), Özgür Ulusoy (Bilkent University)

- **Energy-Price-Driven Query Processing in Multi-center Web Search Engines.**

Enver Kayaaslan (Bilkent University), B. Barla Cambazoglu, Roi Blanco, Flavio P. Junqueira (Yahoo! Research), Cevdet Aykanat (Bilkent University)

- **Faster Top-k Document Retrieval Using Block-Max Indexes.**
  Shuai Ding, Torsten Suel (Polytechnic Institute of New York University)

Session 10B: Recommender systems
Multi-Function Hall, 2nd Floor, Building A
Chair: Luo Si

- **Utilizing Marginal Net Utility for Recommendation in E-commerce.**
  Jian Wang, Yi Zhang (University of California, Santa Cruz)

- **Recommending Ephemeral Items at Web Scale.**
  Ye Chen (Microsoft Corporation), John F. Canny (University of California, Berkeley)

- **A Unified Framework for Recommendations Based on Quaternary Semantic Analysis.**
  Wei Chen, Wynne Hsu, Mong Li Lee (National University of Singapore)

- **Associative Tag Recommendation Exploiting Multiple Textual Features.**
  Fabiano Belém, Eder Martins, Tatiana Pontes, Jussara Almeida, Marcos Gonçalves (Universidade Federal de Minas Gerais)

Session 10C: Test collections
Opal Room, B1 Floor, Building E
Chair: Ian Soboroff

- **Evaluating Diversified Search Results Using Per-intent Graded Relevance.**
  Tetsuya Sakai, Ruihua Song (Microsoft Research Asia)

- **Evaluating Multi-Query Sessions.**
  Evangelos Kanoulas (University of Sheffield), Ben Carterette (University of Delaware), Paul D. Clough (University of Sheffield), Mark Sanderson (RMIT)

- **Quantifying Test Collection Quality Based on the Consistency of Relevance Judgements.**
  Falk Scholer (RMIT University), Andrew Turpin (University of Melbourne), Mark Sanderson (RMIT University)

- **Pseudo Test Collections for Learning Web Search Ranking Functions.**
  Nima Asadi (University of Maryland), Donald Metzler (University of Southern California), Tamer Elsayed (King Abdullah University of Science and Technology), Jimmy Lin (University of Maryland)

17:10-17:30 **Closing Session**
Grand Ballroom, 1st Floor, Building C

17:30-21:00 **Reception at Microsoft Research Asia**
Free shuttle buses will leave from the conference hotel at 17:30-18:00, and the reception will start at 18:30

Notes

Industry Track Sessions

Sapphire & Amethyst Room, B1 Floor, Building E

08:45-09:00
Opening Remarks

09:00-09:45
Keynote: Sensor-Aided Mobile Information Management and Retrieval
Edward Y. Chang (Google Research, China)

09:45-10:15 Coffee Break

10:15-11:00
Keynote: Box Computing: The Simplest Way to Use the Internet
Ruoxue Liao (Baidu, China)

11:00-11:30
Review of MSR-Bing Web Scale Speller Challenge
Kuansan Wang and Jan Pedersen (Microsoft)

11:30-12:00
From information needs to action needs: towards contextual app search and recommendation
Evgeniy Gabrilovich (Yahoo! Research)

12:00-14:15 Lunch and SIGIR Business Meeting

14:15-15:00
Keynote: Building Search Systems for the Enterprise
Shivakumar Vaithyanathan (IBM Research)

15:00-15:30
Predicting eBay Listing Conversion
Ted Yuan, Zhaohui Chen (eBay)

15:30-16:00 Coffee Break

16:00-16:30
A large scale machine learning system for recommending heterogeneous content in social Network
Yanxin Shi (Facebook Inc.)

16:30-17:15
Keynote: Task Engine: A User-centric Service
Gordon Sun (Tencent Technology, China)

17:15-17:30
Elsevier App Challenge
Jukka Valimaki, Remko Caprio

Industrial Keynote Presenters

Sensor-Aided Mobile Information Management and Retrieval
Edward Y. Chang, Google Research, China

The number of “smart” mobile devices such as wireless phones and tablet computers has been rapidly growing. These mobile devices are equipped with a variety of sensors such as camera, gyroscope, accelerometer, compass, NFC, WiFi, GPS, etc. These sensors can be used to capture images and voice, detect motion patterns, and predict locations, to name just a few. This keynote depicts techniques in configuration, calibration, computation, and fusion for improving sensor performance and conserving power consumption [1, 2]. Novel information management and retrieval applications that can benefit a great deal from enhanced sensor technologies are also presented.

Edward Chang heads Google Research in China since March 2006. He joined the department of Electrical & Computer Engineering at University of California, Santa Barbara, in 1999 after receiving his PhD from Stanford University. Ed received his tenure in 2003, and was promoted to full professor of Electrical Engineering in 2006. His recent research activities are in the areas of distributed data mining and their applications to rich-media data management and social-network collaborative filtering. His research group (which consists of members from Google, UC, MIT, Tsinghua, PKU, and Zheda) recently parallelized SVMs (NIPS 07), PLSA (KDD 08), Association Mining (ACM RS 08), Spectral Clustering (PAMI 10), and LDA (WWW 09) (see MMDS / CIVR / EMMDS / MM / AAIM / ADMA / CIKM keynotes and tutorials for details) to run on parallel machines for mining large-scale datasets. His team at Google have developed and launched Google Confucius (a Q&A system, VLDB 10) at 68 countries including China, Russia, Thailand, Vietnam, Indonesia, 17 Arabic, and 40 Africa nations. Ed also directs Google Mobile 2014 research focused program, which funds novel mobile research projects world-wide. Ed has served on ACM (SIGMOD, KDD, MM, CIKM), VLDB, IEEE, WWW, and SIAM conference program committees, and co-chaired several conferences including MMM, ACM MM, ICDE, and WWW. He is a recipient of the NSF Career Award.

Box Computing: The Simplest Way to Use the Internet
Ruoxue Liao, Baidu, China
The concept "Box Computing" was first announced in Baidu Technology Innovation Conference 2009, in response to the constantly changing world of Internet and its users. More and more information is generated by users and the contents on the web have become more and more dynamic. Users want not only mark-up webpages, but also multi-media content which can be played immediately and applications which allows more interactions. By "Box Computing", users can simply input what they want. No matter whether it's a webpage, a game, online shopping tool, or virus-scanning software. Ideally, our box will respond to users with exact "answers" - fast, directly, simply, and reliably. Furthermore, our "Box Computing" offers an open platform for people -- individuals and groups -- to share data and applications, and subsequently allows every internet user to benefit from the power of the crowd, which means Box Computing is also an eco-system of collected intellectual and users. This talk will guide you to the world of Baidu's "Box Computing".

Ruoxue Liao is senior architect & chair of technical committee at Baidu, where he focuses on Internet products and technologies including search, recommendation and architecture. He is one of the chief technical leaders of Baidu's "Box Computing".

Building Search Systems for the Enterprise
Shivakumar Vaithyanathan, IBM Research

In contrast to the radical advances in Web search over the last several years, search over enterprise intranets has remained a difficult and largely unsolved problem. Among the reasons is the fact that content creators lack the economic incentive to make the relevance of their pages easily discoverable. Towards a solution, we developed the notion of a search database system, which is designed around two core necessities. First, auxiliary databases compiled from corpus-text analysis are crucial to understanding the intent of a search query, and the degree to which this intent is met by a given document. Second, administrators should have the ability to hook into the ranking logic in order to specify critical

up-to-date knowledge, and thereby ensure high quality even as content, trends, and users change over time. In a search database system, such knowledge is specified by means of a simple yet expressive rule language for query rewriting. I will describe the efforts in developing a robust implementation of a search database system for the IBM intranet, which will go live in early August 2011. I will also discuss our theoretical modeling of search database systems, and some of the research directions worthwhile pursuing.

Task Engine: A User-centric Service
Gordon Sun,Tencent Technology, China

When a user starts a search he or she usually has a task in mind. The normal web search engine is meant to retrieve relevant information for the user, but not for the fulfillment of the user's ultimate task. In this talk, I will propose a new product concept – “task engine”, with its conceptual model and functional characteristics. An effective task engine can act as user's personal online assistant for various tasks and generate greater user loyalty and business revenue in the Internet industry. I start the talk with an analysis of query logs from the SOSO commercial search engine, categorize query sessions into various types of user task. I will describe the research work and product development at Tencent for several tasks, our successes and the technical challenges we had. Although a task engine could be partly implemented using the current web search engine infrastructure, two major limitations of the latter need to be addressed for a more appropriate implementation: “single query box input” and “one-shot search process”. Our experience showed that in addition to search engine, a task engine should also involve decision engine, recommendation engine and various data mining engine.

Dr. Gordon Sun, Chief Scientist, Tencent Technology, China.
He is leading the next generation search technology program at Tencent, serving as the General Manager of Research department of Search Technology. Dr. Sun has been working on algorithmic search technology since 1998 when he joined Inktomi (the leading US search engine company during 1990s) as the senior scientist and architect. He joined Yahoo early 2004

where he was leading the Global Search Relevance team as the director of research until 2009 when he joined Tencent and moved to China. Gordon received his Ph.D. in theoretical physics from University of Iowa, 1984. He worked as the Research faculty in University of Maryland for 9 years before he moved to Silicon Valley, 1993. He has broad work experiences, knowledge and publications in neural networks, pattern recognition, machine learning, data mining, information retrieval, speech recognition, hand-writing recognition and non-linear dynamics.

Posters

01 Parallel Learning to Rank for Information Retrieval
Shuaiqiang Wang (Shandong University of Finance), Byron J. Gao, Ke Wang (Simon Fraser University), Hady W. Lauw (Institute for Infocomm Research)

02 Learning Features Through Feedback for Blog Distillation
Dehong Gao, Renxian Zhang, Wenjie Li (The Hong Kong Polytechnic University), Yiu Keung Lau (City University of Hong Kong), Kam Fai Wong (The Chinese University of Hong Kong)

03 Time-based Relevance Models.
Mostafa Keikha, Shima Gerani, Fabio Crestani (University of Lugano)

04 Improved Query Performance Prediction Using Standard Deviation
Ronan Cummins, Joemon M. Jose (University of Glasgow), Colm O'Riordan (National University of Ireland, Galway)

05 Learning to Rank Using Query-Level Regression
Jiajin Wu, Zhihao Yang, Yuan Lin, Hongfei Lin, Zheng Ye, Kan Xu (Dalian University of Technology)

06 Diversifying Product Search Results
Xiangru Chen, Haofen Wang, Xinruo Sun, Junfeng Pan, Yong Yu (Shanghai Jiao Tong University)

07 Ad Hoc IR – Not Much Room for Improvement
Andrew Trotman, David Keeler (University of Otago)

08 Image Annotation Based on Recommendation Model
Zijia Lin, Guiguang Ding, Jianmin Wang (Tsinghua University)

09 Utilizing Minimal Relevance Feedback for Ad Hoc Retrieval
Eyal Krikon, Oren Kurland (Technion – Israel Institute of Technology)

10 Sense Discrimination for Physics Retrieval
Christina Lioma, Alok Kothari, Hinrich Schütze (Stuttgart University)

11 When Documents Are Very Long, BM25 Fails!
Yuanhua Lv, ChengXiang Zhai (University of Illinois at Urbana-Champaign)

12 Location and Timeliness of Information Sources During News Events
Elad Yom-Tov, Fernando Diaz (Yahoo Research)

13 What Deliberately Degrading Search Quality Tells Us About Discount Functions
Paul Thomas (CSIRO), Timothy Jones (Australian National University), David Hawking (Funnelback Pty Ltd & Australian National University)

14 Collective Topic Modeling for Heterogeneous Networks
Hongbo Deng, Bo Zhao, Jiawei Han (University of Illinois at Urbana-Champaign)

15 Graph-Cut Based Tag Enrichment
Xueming Qian, Xian-Sheng Hua (Microsoft Research Asia)

16 Personalized Social Query Expansion Using Social Bookmarking Systems
Mohamed Reda Bouadjenek (Bell Labs France & Université de Versailles), Hakim Hacid, Mokrane Bouzeghoub (Université de Versailles), Johann Daigremont (Bell Labs France)

17 What Are the Real Differences of Children's and Adults' Web Search?
Tatiana Gossen, Thomas Low, Andreas Nürnberger (University of Magdeburg)

18 Cognitive Coordinating Behaviors in Multitasking Web Search
Jia Tina Du (University of South Australia)

19 Optimizing Multimodal Reranking for Web Image Search
Hao Li (Chinese Academy of Sciences), Meng Wang, Zhisheng Li (National University of Singapore), Zheng-Jun Zha, Jialie Shen (Singapore Management University)

20 Multi-Layer Graph-Based Semi-Supervised Learning for Large-Scale Image Datasets Using MapReduce
Wen-Yu Lee, Liang-Chi Hsieh, Guan-Long Wu, Winston Hsu (National Taiwan University), Ya-Fan Su (Chunghwa Telecom Co., Ltd.)

21 Tackling Class Imbalance and Data Scarcity in Literature-Based
Gene Function Annotation
Mathieu Blondel, Kazuhiro Seki, Kuniaki Uehara (Kobe University)

22 Bootstrapping Subjectivity Detection
Valentin Jijkoun, Maarten de Rijke (ISLA, University of Amsterdam)

23 The Effects of Choice in Routing Relevance Judgments
Edith Law (Carnegie Mellon University), Paul N. Bennett, Eric Horvitz (Microsoft Research)

24 Statistical Feature Extraction for Cross-Language Web Content Quality Assessment
Guang-Gang Geng, Xiao-Dong Li, Li-Ming Wang, Wei Wang, Shuo Shen (Chinese Academy of Sciences)

25 Exploiting Endorsement Information and Social Influence for Item Recommendation
Cheng-Te Li, Shou-De Lin, Man-Kwan Shan (National Chengchi University)

26 Modeling Subset Distributions for Verbose Queries
Xiaobing Xue, W. Bruce Croft (University of Massechusetts, Amherst)

27 Domain Expert Topic Familiarity and Search Behavior
Sarvnaz Karimi (NICTA and The University of Melbourne), Falk Scholer (RMIT University), Adam Clark (Alfred Health), Sadegh Kharazmi (RMIT University & NICTA)

28 Sample Selection for Dictionary-Based Corpus Compression
Christopher Hoobin, Simon Puglisi (RMIT University), Justin Zobel (University of Melbourne)

29 Evaluating Medical Information Retrieval
Bevan Koopman (CSIRO), Peter Bruza, Laurianne Sitbon (Queensland University of Technology), Michael Lawley (CSIRO)

30 Region-Based Landmark Discovery by Crowdsourcing Geo-Referenced Photos
Yen-Ta Huang, An-Jung Cheng, Liang-Chi Hsieh, Winston Hsu (National Taiwan University), Kuo-Wei Chang (Chunghwa Telecom Co., Ltd.)

31 Towards Effective Short Text Deep Classification
Xinruo Sun, Haofen Wang, Yong Yu (Shanghai Jiao Tong University)

32 Temporal Latent Semantic Analysis for Collaboratively Generated Content: Preliminary Results
Yu Wang, Eugene Agichtein (Emory University)

33 Self-Adjusting Hybrid Recommenders Based on Social Network Analysis
Alejandro Bellogin, Pablo Castells, Iván Cantador (Universidad Autónoma de Madrid)

34 BlogCast Effect on Information Diffusion in a Blogosphere
Sang-Wook Kim, Sang-Wook Kim (Hanyang University), Christos Faloutsos (Carnegie Mellon University), Jiwoon Ha (Hanyang University)

35 Product Comparison Using Comparative Relations
Si Li (Beijing University of Posts and Telecommunications & National University of Singapore), Zheng-Jun Zha, Zhaoyan Ming, Meng Wang, Tat-Seng Chua (National University of Singapore), Jun Guo, Weiran Xu (Beijing University of Posts and Telecommunications)

36 Collaborative Cyberporn Filtering with Collective Intelligence
Lung-Hao Lee, Hsin-Hsi Chen (National Taiwan University)

37 Do IR Models Satisfy the TDC Retrieval Constraint?
Stéphane Clinchant (Xerox Research Center Europe & Université Grenoble I, LIG), Eric Gaussier (Université Grenoble I, LIG)

38 On Diversifying and Personalizing Web Search
David Vallet, Pablo Castells (Universidad Autónoma de Madrid)

39 Semantic Tag Recommendation Using Concept Model
Chenliang Li, Anwitaman Datta, Aixin Sun (Nanyang Technological University)

40 Recommending Interesting Activity-Related Local Entities
Jie Tang (University of California, Berkeley), Ryen W. White (Microsoft Research), Peter Bailey (Microsoft Bing)

41 Cross-Corpus Relevance Projection
Nima Asadi (University of Maryland), Donald Metzler (University of Southern California), Jimmy Lin (University of Maryland)

42 Location Disambiguation for Geo-Tagged Images
Zhu Zhu, Lidan Shou, Kuang Mao, Gang Chen (Zhejiang University)

43 Towards an Indexing Method to Speed-Up Music Retrieval
Benjamin Martin, Pierre Hanna, Matthias Robine, Pascal Ferraro (University of Bordeaux)

44 An Investigation of Decompounding for Cross-Language Patent Search
Johannes Leveling, Walid Magdy, Gareth J. F. Jones (Dublin City University)

45 Detecting Seasonal Queries by Time-Series Analysis
Milad Shokouhi (Microsoft Research)

46 Learning to Rank Under Tight Budget Constraints
Christian Pölitz, Ralf Schenkel (Saarland University)

47 A Novel Hybrid Index Structure for Efficient Text Retrieval
Andreas Broschart, Ralf Schenkel (Universitaet des Saarlandes and Max-Planck-Institut fuer Informatik)

48 A Weighted Curve Fitting Method for Result Merging in Federated Search
Chuan He (Beijing University of Posts and Telecommunications), Dzung Hong, Luo Si (Purdue University)

49 Effect of Different Docid Orderings on Dynamic Pruning Retrieval Strategies
Nicola Tonellotto (National Research Council), Craig Macdonald, Iadh Ounis (University of Glasgow)

50 Time-Based Query Performance Predictors
Nattiya Kanhabua, Kjetil Nørvåg (Norwegian University of Science and Technology)

51 Search Task Difficulty: The Expected vs. The Reflected
Jingjing Liu (Southern Connecticut State University), Nicholas J. Belkin (Rutgers University)

52 On the Suitability of Diversity Metrics for Learning-to-Rank for Diversity
Rodrygo L. T. Santos, Craig Macdonald, Iadh Ounis (University of Glasgow)

53 How Diverse Are Web Search Results?
Rodrygo L. T. Santos, Craig Macdonald, Iadh Ounis (University of Glasgow)

54 Analysis of an Expert Search Query Log
Yi Fang, Naveen Somasundaram (Purdue University), Luo Si, Jeongwoo Ko (Google Inc.), Aditya P. Mathur (Purdue University)

55 A Model for Expert Finding in Social Networks
Elena Smirnova (INRIA Sophia Antipolis)

56 Transductive Learning Over Automatically Detected Themes for Multi-Document Summarization
Massih-Reza Amini (National Research Council of Canada), Nicolas Usunier (Université Pierre et Marie Curie)

57 Rating-Based Collaborative Filtering Combined with Additional Regularization
Shu Wu, Shengrui Wang (University of Sherbrooke)

58 Words-of-Interest Selection Based Temporal Motion Coherence for Video Retrieval on
Lei Wang, Dawei Song, Eyad Elyan (Robert Gordon University)

59 Aggregating Multiple Opinion Evidence in Proximity-Based Opinion Retrieval
Shima Gerani, Mostafa Keikha, Fabio Crestani (University of Lugano)

60 Enhancing Mobile Search Using Web Search Log Data
Yoshiyuki Inagaki, Jiang Bian, Yi Chang, Motoko Maki (Yahoo! Inc.)

61 Award Prediction with Temporal Citation Network Analysis
Zaihan Yang, Dawei Yin, Brian D. Davison (Lehigh University)

62 Rating Prediction Using Feature Words Extracted from Customer Reviews
Masanao Ochi, Makoto Okabe, Rikio Onai (The University of Electro-Communications)

63 Ranking Tags in Resource Collections
Dimitrios Skoutas, Mohammad Alrifai (L3S Research Center)

64 Identifying Similar People in Professional Social Networks with Discriminative Probabilistic Models
Suleyman Cetintas (Purdue University), Monica Rogati (LinkedIn Corp.), Luo Si, Yi Fang (Purdue University)

65 Intent-Oriented Diversity in Recommender Systems
Saúl Vargas, Pablo Castells, David Vallet (Universidad Autónoma de Madrid)

66 Disambiguating Biomedical Acronyms Using EMIM
Nut Limsopatham, Rodrygo L. T. Santos, Craig Macdonald, Iadh Ounis (University of Glasgow)

67 Best Document Selection Based on Approximate Utility Optimization
Hungyu Henry Lin, Yi Zhang, James Davis (University of California, Santa Cruz)

68 Forecasting Counts of User Visits for Online Display Advertising with Probabilistic Latent Class Models
Suleyman Cetintas (Purdue University), Datong Chen (Yahoo! Labs), Luo Si, Bin Shen, Zhanibek Datbayev (Purdue University)

69 Knowledge Effects on Document Selection in Search Results Pages
Michael J. Cole (Rutgers University), Xiangmin Zhang (Wayne State University), Chang Liu, Nicholas J. Belkin, Jacek Gwizdka (Rutgers University)

70 Learning to Rank From a Noisy Crowd
Abhimanu Kumar, Matthew Lease (University of Texas at Austin)

71 How to Count Thumb-Ups and Thumb-Downs?: An Information Retrieval Approach to User-Rating Based Ranking of Items
Dell Zhang (Birkbeck, University of London), Robert Mao, Haitao Li, Joanne Mao (Hughes Network Systems)

72 Predicting Users' Domain Knowledge from Search Behaviors
Xiangmin Zhang (Wayne State University), Michael Cole, Nicholas J. Belkin (Rutgers University)

73 The Interactive PRP for Diversifying Document Rankings
Guido Zuccon, Leif Azzopardi, C. J. Keith van Rijsbergen (University of Glasgow)

74 Detecting Success in Mobile Search from Interaction
Qi Guo, Shuai Yuan, Eugene Agichtein (Emory University)

75 Measuring Assessor Accuracy:A Comparison of NIST Assessors and User Study Participants
Mark D. Smucker, Chandra Prakash Jethani (University of Waterloo)

76 A Bipartite Graph Based Social Network Splicing Method for Person Name Disambiguation
Jintao Tang (Hong Kong Polytechnic University & National University of Defense Technology), Qin Lu, Ting Wang (National University of Defense Technology), Ji Wang (National Laboratory for Parallel and Distributed Processing), Wenjie Li (Hong Kong Polytechnic University)

77 Link Formation Analysis in Microblogs
Dawei Yin, Liangjie Hong, Xiong Xiong, Brian D. Davison (Lehigh University)

78 Evolution of Web Search Results Within Years
Ismail Sengor Altingovde (L3S Research Center), Rifat Ozcan, Özgür Ulusoy (Bilkent University)

79 Decayed DivRank: Capturing Relevance, Diversity and Prestige in Information Networks
Pan Du, Jiafeng Guo, Xue-Qi Cheng (Chinese Academy of Sciences)

80 Multi-Objective Optimization in Learning to Rank
Na Dai (Lehigh University), Milad Shokouhi (Microsoft Research), Brian D. Davison (Lehigh University)

81 A Large-Scale Study of the Effect of Training Set Characteristics Over Learning-to-Rank Algorithms
Evangelos Kanoulas (University of Sheffield), Stefan Savev, Pavel Metrikov, Virgil Pavlu, Javed Aslam (Northeastern University)

82 Exploring Term Temporality for Pseudo-Relevance Feedback
Stewart Whiting, Yashar Moshfeghi, Joemon M. Jose (University of Glasgow)

83 MSSF: A Multi-Document Summarization Framework Based on Submodularity
Jingxuan Li, Lei Li, Tao Li (Florida International University)

84 SEJoin: An Optimized Algorithm Towards Efficient Approximate String Searches
Junfeng Zhou, Ziyang Chen, Jingrong Zhang (Yanshan University)

85 Bag-of-Visual-Words vs Global Image Descriptors on Two-Stage Multimodal Retrieval
Konstantinos Zagoris, Savvas A. Chatzichristofis, Avi Arampatzis (Democritus University of Thrace)

86 Query Term Ranking Based on Search Results Overlap
Wei Song, Yu Zhang, Yubin Xie, Ting Liu, Sheng Li (Harbin Institute of Technology)

87 Tossing Coins to Trim Long Queries
Sudip Datta, Vasudeva Varma (International Institute of Information Technology)

88 A Comparison of Time-aware Ranking Methods
Nattiya Kanhabua, Kjetil Nørvåg (Norwegian University of Science and Technology)

89 Learning for Graphs with Annotated Edges
Fan Li (Yahoo! Labs)

90 Formulating Effective Questions for Community-based Question Answering
Saori Suzuki, Shin'ichi Nakayama, Hideo Joho (University of Tsukuba)

Demonstrations

01 ClusteringWiki: Personalized and Collaborative Clustering of Search Results
David C. Anastasiu, Byron J. Gao (Texas State University-San Marcos), David Buttler (Lawrence Livermore National Laboratory)

02 OrientSTS: Spatio-Temporal Sequence Searching in Flickr
Chunjie Zhou (Renmin University of China & Ludong University), Dongqi Liu, Xiaofeng Meng (Renmin University of China)

03 A Toolkit for Knowledge Base Population
Zheng Chen, Suzanne Tamang, Adam Lee, Heng Ji (City University of New York)

04 iMecho: A Context-aware Desktop Search System
Jidong Chen (Fudan University & EMC Research China), Hang Guo (EMC Research China), Wentao Wu, Wei Wang (Fudan University)

05 Visualizing and Querying Semantic Social Networks
Aixin Sun, Anwitaman Datta (Nanyang Technological University), Ee-Peng Lim (Singapore Management University), Kuiyu Chang (Nanyang Technological University)

06 What-You-Retrieve-Is-What-You-See: A Preliminary Cyber-Physical Search Engine
Lidan Shou, Ke Chen, Gang Chen, Chao Zhang, Yi Ma, Xian Zhang (Zhejiang University)

07 QuickView: Advanced Search of Tweets
Xiaohua Liu (Microsoft Research Asia; Harbin Institute of Technology), Long Jiang, Furu Wei, Ming Zhou, Microsoft QuickView Team (Microsoft Research Asia)

08 Personalized Video: Leanback Online Video Consumption
Krishnan Ramanathan, Yogesh Sankarasubramaniam, Vidhya Govindaraju (HP Labs India)

09 GreenMeter: A Tool for Assessing the Quality and Recommending Tags for Web 2.0 Applications
Saulo Ricci, Dilson Guimarães, Fabiano Belém, Jussara M. Almeida, Marcos A. Gonçalves, Raquel Prates (Universidade Federal de Minas Gerais)

10 JuSe: A Picture Dictionary Query System for Children
Tamara Polajnar, Richard Glassey, Leif Azzopardi (University of Glasgow)

11 CrowdTracker: Enabling Community-Based Real-Time Web Monitoring
James Caverlee, Zhiyuan Cheng, Brian Eoff, Chiao-Fang Hsu, Krishna Kamath, Jeffrey McGee (Texas A&M University)

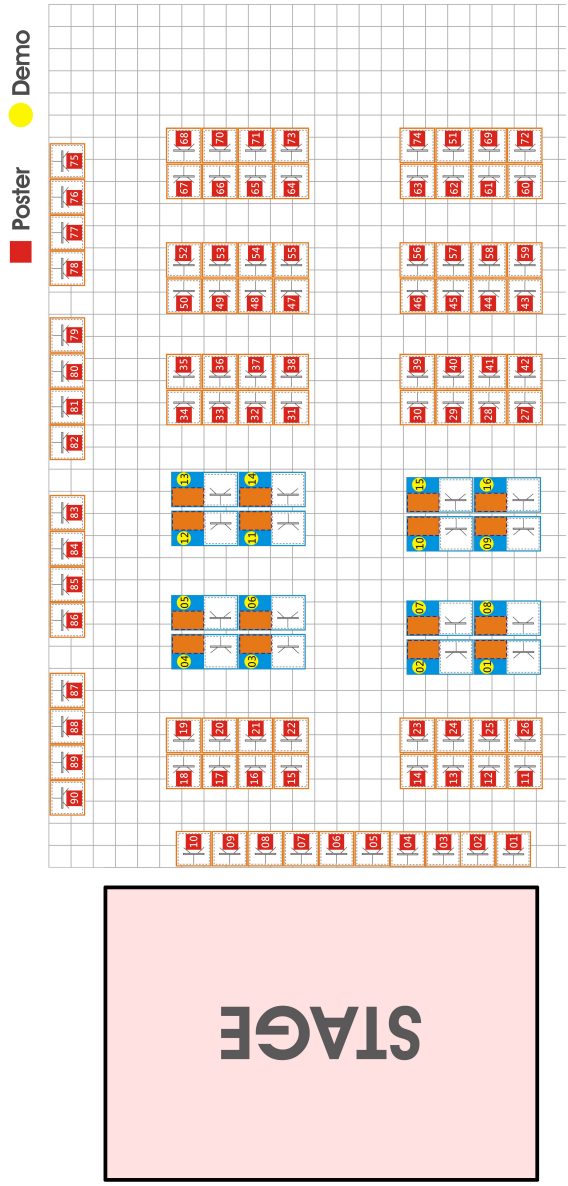
12 The Meta-Dex Suite: Generating and Analyzing Indexes and Meta-Indexes
Michael Huggett, Edie Rasmussen (University of British Columbia)

13 Tulsa: Web Search for Writing Assistance
Duo Ding (Shanghai Jiao Tong University), Xingping Jiang, Matthew R. Scott, Ming Zhou (Microsoft Research Asia), Yong Yu (Shanghai Jiao Tong University)

14 The TREC Files: The (Ground) Truth Is Out There
Savvas A. Chatzichristofis, Konstantinos Zagoris, Avi Arampatzis (Democritus University of Thrace)

15 A Tool for Comparative IR Evaluation on Component Level
Thomas Wilhelm, Jens Kürsten, Maximilian Eibl (Chemnitz University of Technology)

Layout for Poster and Demo Session



Workshops

All workshops will run full-day on **Thursday July 28th, 2011**

Query Representation and Understanding
Hang Li, Gu Xu, W. Bruce Croft, Michael Bendersky
Webpage: http://ciir.cs.umass.edu/sigir2011/qru/
Room 7024

"Entertain me" Supporting Complex Search Tasks
Charlie Clarke, Ning Gao, Jaap Kamps, Jussi Karlgren, Diane Kelly
Webpage: http://staff.science.uva.nl/~kamps/entertainme/
Room 8024

Crowdsourcing for Information Retrieval
Vaughn Hester, Matthew Lease, Alex Sorokin, Emine Yilmaz
Webpage: https://sites.google.com/site/cir2011ws/
Room 9024

Information Retrieval for E-discovery SIRE
Jason R. Baron, Maura R. Grossman, David D. Lewis, Douglas W. Oard
Webpage: http://www.umiacs.umd.edu/~oard/sire11/
Room 1024

Internet Advertising
Tie-Yan Liu, Tao Qin, James G. Shanahan
Webpage: http://research.microsoft.com/en-us/um/beijing/events/ia2011/
Room 7026

Enriching Information Retrieval
Paul N. Bennet, Khalid El-Arini, Thorsten Joachims, Krysta M. Svore
Webpage: http://select.cs.cmu.edu/meetings/enir2011/
Room 8026

Entity-Oriented Search
Krisztian Balog, Arjen P. de Vries, Pavel Serdyukov, Ji-Rong Wen
Webpage: http://research.microsoft.com/en-us/um/beijing/events/eos2011/
Room 1026

Social Web Search and Mining, Analysis under crisis

Fernando Diaz, E. Hovy, Irwin King, Juanzi Li, Donald Metzler, Marie-Francine Moens, Jie Tang, Lei Zhang

Webpage: http://arnetminer.org/SWSM_2011

Room 1226

Reception at Microsoft Research Asia

Wednesday, July 27th, 2011

17:30-18:00	Departure from Beijing Hotel (Gathering at Hotel Lobby, Building A)
18:30-19:30	Arrival at BJW (Microsoft Research Asia Office)
18:30-19:30	Group visit to Visitor Center
19:30-20:00	Welcome Speech
20:00-21:00	Reception and Demonstrations
21:00-22:00	Back to Hotel

MSRA reception Address:
Microsoft Research Asia
Building 2, No. 5 Dan Ling Street, Haidian District, Beijing,
P.R. China, 100080

微软亚洲研究院
北京市海淀区丹棱街 5 号，微软大厦二号楼
邮编：100080

Location map:

Industry Lab Tour

Thursday, July 28th, 2011

09:00-10:00	Departure from Beijing Hotel (Gathering at Hotel Lobby, Building A)
10:00-12:00	Visit Tencent
12:00-14:30	Visit Baidu (Lunch at Baidu)
14:30-16:30	Visit Yahoo! China
16:30-17:30	Back to Hotel

Near-by Restaurants

Tan Jia Cai Restaurant 谭家菜
Address: 7/F, Unit C, Beijing Hotel, 33, Dong Changan Jie, Dongcheng District

Tiandiyijia Restaurant 天地一家
Address:No.140,Nanchizi Avenue,Dongcheng District,Beijing

Donglaishun Hualong Jie Branch 东来顺华龙街店
Address:South courtyard of Guibinlou, E Bldg, Hualongjie, Nanheyan Dajie, Dongcheng District

Beijing Wangfujing Quanjude Roast Duck Restaurant
全聚德烤鸭店（王府井店）
Address: 13, Shuaifuyuan Hutong, Dongcheng District

Golden Jagur Wangfujing Branch (Buffet Restaurant)
金钱豹国际美食百汇
Address:2F Wangfu Century Bldg., No. 55 Dong An Men Street, Dong Cheng District, Beijing

La Cite 紫禁阁西餐厅
Address:115 Nanchizi Dajie,

Wangfujing Snack Street 王府井小吃街
Address:south of Haoyou Department Store, Wangfujing Business Street, Dongcheng District

Donghuamen Snack Night Market 东华门美食坊夜市
Address:North of Donganmen Street, Dongcheng District

Muslim Restaurants

Hongbin Lou Restaurant 鸿宾楼
Address: No.11, Exhibition Hall Road, Xicheng District (北京市西城区展览馆路 11 号)

Hongyunlou Restaurant 鸿运楼
Address: No.7, East Majiabao Road, Fengtai District (丰台区马家堡东路 7 号)

Kaorouji Restaurant 聚德华天烤肉季
Address: No.14, Qianhai East Bank Road, Di'anmen Outer Avenue, Xicheng District (西城地安门外大街前海东沿 17 号)

Baikui Restaurant 白魁老号饭庄
Address: No.158, Jiaodaokou South Avenue, Dongcheng District (东城区交道口南大街 158 号)

Jubaoyuan Restaurant 聚宝源
Address: No.5-2 of 1st Xili Business Building, Niu Street, Xuanwu District (宣武区牛街西里商业 1 号楼 5-2 号)

Vegetarian Restaurants

Baihe Vegetarian Restaurant 百合素食
Address: No.23, Caoyuan Alley, North Street of Dongzhimen, Dongcheng District (东城区东直门北小街草园胡同甲 23 号)

Pure Lotus 静心莲 (长虹桥店)
Address: Tongguang Mansion, No.12, Nongzhanguan Road, Chaoyang District (朝阳区农展馆南路 12 号通广大厦院内)

Natural Vegetarian Lohas 博味堂自然素食
Address: No.14-9, 1st Anhui Li, Chaoyang District

Jingsi Vegetarian Restaurant (静思素食坊)
Address: No.18, East Street, Big Buddhist Temple, North Art Museum, Dongcheng District.

Gong De Lin Vegetarian Restaurant 功德林素菜馆
Address: No.2, Qianmen East Srteet, Congwen District

Notes

SIGIR
Beijing
2011

SIGIR
Beijing
2011

Map of Beijing Hotel

北京饭店
BEIJING HOTEL

Address: No. 33 East Chang An Avenue, Dongcheng District, Beijing
Tel: +86-10-6513-7766
地 址：北京市东城区东长安街33号
电 话：+86-10-6513-7766