

SIGIR Sheffield 2004

Final Program

The Twenty-Seventh Annual
International ACM SIGIR Conference
on Research and Development
in Information Retrieval

July 25 - 29, 2004

University of Sheffield

Travel Information

Sheffield Station to Conference venues

1. Take a tram at the station (Malin Bridge - Blue line)
2. Get off the tram at University of Sheffield
3. Walk to Octagon Centre or Arts Tower (5 min)

Sheffield Station to Hotels

Those who stay at Hotel Bristol, Travel Inn, IBIS Hotel, or Novotel:

Walk to hotels (5 to 15 min)

Those who stay at the rest of hotels:

Take a taxi from the station

Hotels to Conference venues

Those who stay at Hotel Bristol, Travel Inn, IBIS Hotel, or Novotel:

1. Take a tram at Castle Square (Either Middlewood - Yellow line or Malin Bridge - Blue line)
2. Get off the tram at University of Sheffield
3. Walk to Octagon Centre or Arts Tower (5 min)

Those who stay at Halifax Hall or Sheffield Marriot Hotel:

A coach service will be available (Timetable is given in a separate sheet)

Those who stay at Rutland Hotel:

1. Walk to Octagon Centre or Arts Tower (20 min)
2. Or take a bus (Route 60) and get off at University Students Union (Glossop Road)
3. Walk to Octagon Centre or Arts Tower (2 min)

Tram Fees

1. £1.20 for Station to University (One way)
2. £2.40 for One day ticket (called Dayrider)
3. £7.50 for Week ticket (called Megarider)

Tickets can be purchased from a conductor in the tram.

Sponsorships

SIGIR 2004 is sponsored by the following organizations.

ACM

BCS IRSG

Marie Curie Actions - EU

Microsoft Research

Canon Research Centre Europe

Mirago

IBM Research

Yahoo! Research Labs

Yahoo! Research Labs investigates fundamental computational problems and seeks innovative solutions as part of Yahoo!'s larger mission of being the most essential Internet service.

Ask Jeeves, Inc.

Ask Jeeves, Inc. provides consumers and advertisers with information retrieval products across a diverse portfolio of Web sites, portals and desktop search applications.

Google Inc.

Google's innovative search technologies connect millions of people around the world with information every day. Learn about engineering positions in the U.S., Europe or Asia at www.google.com/jobs/eng.html

ERCIM

Thomson

MITRE

Kluwer Academics

SHARP Laboratory of Europe

Queen Mary College, University of London

Keynote Address

Challenges in Using Lifetime Personal Information Stores

Dr. Gordon Bell (Microsoft Bay Area Research Center)

"Within five years, our personal computers with terabyte disk drives will be able to store everything we read, write, hear, and many of the images we see including video. Vannevar Bush outlined such a system in his famous 1945 Memex article.

For the last four years we have worked on MyLifeBits www.MyLifeBits.com, a system to digitally store everything from one's life, including books, articles, personal financial records, memorabilia, email, written correspondence, photos (time, location taken), telephone calls, video, television programs, and web pages visited. We recently added content from personal devices that automatically record photos and audio."

(Snippet from Abstract)

Dr. Gordon Bell's keynote address will start at 09.30 on Monday 26, July.

Chemoinformatics: an Application Domain for Information Retrieval Techniques

Prof. Peter Willett (University of Sheffield)

"Chemoinformatics is the generic name for the techniques used to represent, store and process information about the two-dimensional (2D) and three-dimensional (3D) structures of chemical molecules. Chemoinformatics has attracted much recent prominence as a result of developments in the methods that are used to synthesize new molecules and then to test them for biological activity. These developments have resulted in a massive increase in the amounts of structural and biological information that is available to support discovery programmes in the pharmaceutical and agrochemical industries."

(Snippet from Abstract)

Prof. Peter Willett's keynote address will start at 11.30 on Wednesday 28, July.

Full Conference Program

Sunday 25 July

Tutorials

Venue: Arts Tower

08.00 - 17.00 **Registration Opens**

08.00 - 10.00 **XML Information Retrieval**

Ricardo Baeza-Yates (University of Chile)
Norbert Fuhr (University of Duisburg Essen)

Machine Learning for Text Classification Applications

David Lewis (Independent consultant)

Multilingual Information Retrieval

Fred Gey (University of California at Berkeley)

User Interfaces and Visualisation for Information Access

Marti Hearst (UC Berkeley)

10.00 - 10.30 **Coffee Break**

10.30 - 12.15 **XML Information Retrieval**

Ricardo Baeza-Yates (University of Chile)
Norbert Fuhr (University of Duisburg Essen)

Machine Learning for Text Classification Applications

David Lewis (Independent consultant)

Multilingual Information Retrieval

Fred Gey (University of California at Berkeley)

User Interfaces and Visualisation for Information Access

Marti Hearst (UC Berkeley)

12.15 - 13.15 **Lunch**

13.15 - 15.00 **Internet search**

Jan Pedersen (Yahoo!, Inc.)
Knut Magne Risvik (Yahoo!, Inc.)

High Performance Indexing and Query Evaluation for Information Retrieval

Justin Zobel (Royal Melbourne Institute of Technology)

Bioinformatics and Genomics for Information Retrieval

Bill Hersh (Oregon Health and Science University)

Text summarization

Dragomir Radev (University of Michigan)

15.00 - 15.30 **Coffee Break**

15.30 - 17.00 **Internet search**

Jan Pedersen (Yahoo!, Inc.)
Knut Magne Risvik (Yahoo!, Inc.)

High Performance Indexing and Query Evaluation for Information Retrieval

Justin Zobel (Royal Melbourne Institute of Technology)

Bioinformatics and Genomics for Information Retrieval

Bill Hersh (Oregon Health and Science University)

Text summarization

Dragomir Radev (University of Michigan)

Full Conference Program

Sunday 25 July

Doctoral Consortium

Venue: Halifax Hall of Residence

09.00 - 09.15	Introduction
09.00 - 10.45	Presentations/Meetings I
10.45 - 11.15	Coffee Break
11.15 - 12.45	Presentations/Meetings II
12.45 - 14.00	Lunch
14.00 - 16.30	Presentations/Meetings III
16.30 - 16.45	Feedback and Wrap-up

Coach Trip

Venue: Chatsworth House

12.00 - 12.30	Coach pick-up (exact time of pick-up depends on the hotels)
13.30 - 15.00	Guided tour
15.00 - 16.00	Free time
16.00 - 17.00	Depart from Chatsworth and return to hotels

Reception

Venue: Millennium Galleries

18.45 - 19.15	Coach pick-up (exact time of pick-up depends on the hotels)
19.30 - 21.00	Reception

Full Conference Program

Monday 26 July

07.30 - 09.00 **Newcomer's Breakfast** (University House)

08.00 - 18.00 **Registration Opens** (Octagon Centre)

Conference open

Venue: Octagon Centre

Chair: Mark Sanderson

09.00 - 09.30 **Opening speech**

09.30 - 10.30 **Keynote Address: Challenges in Using Lifetime Personal Information Stores**

Gordon Bell (Microsoft Bay Area Research Center)

10.30 - 11.00 **Coffee Break** (University House)

Session I: Opening Session

Venue: Octagon Centre

Chair: David Harper

11.00 - 12.30 **Evaluating High Accuracy Retrieval Techniques**

Chirag Shah (University of Massachusetts) and W. Bruce Croft (University of Massachusetts)

Scaling IR-System Evaluation using Term Relevance Sets

Einat Amitay (IBM Research Lab), David Carmel (IBM Research Lab), Ronny Lempel (IBM Research Lab), and Aya Soffer (IBM Research Lab)

Using Temporal Profiles of Queries for Precision Prediction

Fernando Diaz (University of Massachusetts) and Rosie Jones (Yahoo! Inc.)

12.30 - 14.00 **Lunch** (University House)

Session IIa: Test collections

Venue: Octagon Centre

Chair: Kalervo Järvelin

14.00 - 15.30 **Retrieval Evaluation with Incomplete Information**

Ellen M. Voorhees (National Institute of Standards and Technology) and Chris Buckley (Sabir Research, Inc.)

Test collections with no system pooling

Mark Sanderson (University of Sheffield) and Hideo Joho (University of Sheffield)

Building an Information Retrieval Test Collection for Spontaneous Conversational Speech

Douglas W Oard (University of Maryland), David Doermann (University of Maryland), Xiaoli Huang (University of Maryland), G. Craig Murray (University of Maryland), Jianqiang Wang (University of Maryland), Bhuvana Ramabhadran (IBM), Martin Franz (IBM), Samuel Gustman (Survivors of the Shoah Visual History Foundation), James Mayfield (Johns Hopkins University), Liliya Kharevych (California Institute of Technology), and Stephanie Strassel (Linguistic Data Consortium)

15.30 - 16.00 **Coffee Break** (University House)

Full Conference Program

Monday 26 July

Session IIb: Formal models - 1

Venue: Students' Auditorium
Chair: Fabrizio Sebastiani

- 14.00 - 15.30 **A Formal Study of Information Retrieval Heuristics**
Hui Fang (University of Illinois), Tao Tao (University of Illinois),
and ChengXiang Zhai (University of Illinois)
- Probabilistic Model for Contextual Retrieval**
Ji-Rong Wen (Microsoft Research Asia) and Ni Lao (Tsinghua
University)
- Discriminative Models for Information Retrieval**
Ramesh M Nallapati (University of Massachusetts)

15.30 - 16.00 **Coffee Break** (University House)

Session IIIa: XML retrieval

Venue: Octagon Centre
Chair: Norbert Fuhr

- 16.00 - 17.30 **The overlap problem in content-oriented XML retrieval
evaluation**
Gabriella Kazai (Queen Mary University of London), Mounia Lal-
mas (Queen Mary University of London), and Arjen P. de Vries
(CWI)
- Length Normalization in XML Retrieval**
Jaap Kamps (University of Amsterdam), Maarten de Rijke
(University of Amsterdam), and Bkür Sigurbjörnsson (University of
Amsterdam)
- Configurable Indexing and Ranking for XML Information
Retrieval**
Shaorong Liu (UCLA), Qinghua Zou (UCLA), and Wesley W. Chu
(UCLA)

Session IIIb: Dimensionality reduction

Venue: Students' Auditorium
Chair: Stephen Robertson

- 16.00 - 17.30 **Locality Preserving Indexing for Document Representation**
Xiaofei He (University of Chicago), Deng Cai (Tsinghua Univer-
sity), Haifeng Liu (University of Toronto), and Wei-Ying Ma
(Microsoft Research Asia)
- Polynomial Filtering in Latent Semantic Indexing for Informa-
tion Retrieval**
Effrosyni Kokiopoulou (University of Minnesota) and Yousef Saad
(University of Minnesota)
- On Scaling Latent Semantic Indexing for Large Peer-to-Peer
Systems**
Chunqiang Tang (University of Rochester), Sandhya Dwarkadas
(University of Rochester), and Zhichen Xu (Yahoo! Inc.)

Poster and Demo

Venue: Octagon Centre

18.00 - 21.00 **Poster and Demo Session with buffet**

Full Conference Program

Tuesday 27 July

08.00 - 18.00 **Registration Opens** (Octagon Centre)

Session IVa: Formal models - 2

Venue: Octagon Centre

Chair: Mounia Lalmas

08.30 - 10.00 **GaP: A Factor Model for Discrete Data**

John F Canny (University of California, Berkeley)

Belief Revision for Adaptive Information Retrieval

Raymond Y K Lau (Queensland University of Technology), Peter D Bruza (The University of Queensland), and Dawei Song (The University of Queensland)

Tuning Before Feedback: Combining Ranking Discovery and Blind feedback for Robust Retrieval

Weiguo Fan (Virginia Tech), Ming Luo (Virginia Tech), Li Wang (University of Michigan, Ann Arbor), Wensi Xi (Virginia Tech), and Edward A. Fox (Virginia Tech)

Session IVb: Cross-language IR

Venue: Students' Auditorium

Chair: Bruce Croft

08.30 - 10.00 **Translating Unknown Queries with Web Corpora for Cross-Language Information Retrieval**

Pu-Jen Cheng (Academia Sinica, Taiwan), Jei-Wen Teng (Academia Sinica, Taiwan), Ruei-Cheng Chen (Academia Sinica, Taiwan), Jeng-Haur Wang (Academia Sinica, Taiwan), Wen-Hsiang Lu (National Cheng Kung University, Taiwan), and Lee-Feng Chien (National Taiwan University, Taiwan)

Resource Selection for Domain-Specific Cross-Lingual IR

Monica Rogati (Carnegie Mellon University) and Yiming Yang (Carnegie Mellon University)

Using the Web for Automated Translation Extraction in Cross-Language Information Retrieval

Y Zhang (RMIT University) and P Vines (RMIT University)

10.30 - 11.00 **Coffee Break** (University House)

Session Va: Language models

Venue: Octagon Centre

Chair: Gareth Jones

10.30 - 12.30 **Dependence Language Model for Information Retrieval**

Jianfeng Gao (Microsoft Research Asia), Jian-Yun Nie (Université de Montréal), Guangyuan Wu (Tianjin University), and Guihong Cao (Tianjin University)

Parsimonious Language Models for Information Retrieval

D. Hiemstra (University of Twente), S. E. Robertson (Microsoft Research), and H. Zaragoza (Microsoft Research)

Cluster-based retrieval using language models

Xiaoyong Liu (University of Massachusetts) and W. Bruce Croft (University of Massachusetts)

Corpus structure, language models, and ad-hoc information retrieval

Oren Kurland (Cornell University) and Lillian Lee (Cornell University)

Full Conference Program

Tuesday 27 July

Session Vb: Clustering

Venue: Students' Auditorium

Chair: Fabio Crestani

- 10.30 - 12.30 **Document Clustering by Concept Factorization**
Wei Xu (NEC Laboratories America, Inc) and Yihong Gong (NEC Laboratories America, Inc)
- Learning To Cluster Search Results**
Hua-Jun Zeng (Microsoft Research, Asia), Qi-Cai He (Peking University), Zheng Chen (Microsoft Research, Asia), Wei-Ying Ma (Microsoft Research, Asia), and Jinwen Ma (Peking University)
- Document Clustering via Adaptive Subspace Iteration**
Tao Li (University of Rochester), Sheng Ma (IBM Research), and Mitsunori Ogihara (University of Rochester)
- Restrictive Clustering and Metaclustering for self-organizing Document Collections**
Sergej Sizov (Max-Planck-Institut fuer Informatik) and Stefan Siersdorfer (Max-Planck-Institut fuer Informatik)
- 12.30 - 14.00 **Lunch** (University House)

Session VIa: Text classification

Venue: Octagon Centre

Chair: John Tait

- 14.30 - 15.30 **Feature Selection using Linear Classifier Weights: Interaction with Classification Models**
Dunja Mladenic (Jozef Stefan Institute), Janez Brank (Jozef Stefan Institute), Marko Grobelnik (Jozef Stefan Institute), and Natasa Milic-Frayling (Microsoft Research Ltd.)
- Web-page Classification through Summarization**
Dou Shen (Tsinghua University), Zheng Chen (Microsoft Research Asia), Qiang Yang (Hong Kong University of Science and Technology), Hua-Jun Zeng (Microsoft Research Asia), Benyu Zhang (Microsoft Research Asia), Yuchang Lu (Tsinghua University), and Wei-Ying Ma (Microsoft Research Asia)
- Parameterized generation of labeled datasets for text categorization based on hierarchical directory**
Dmitry Davidov (Technion - Israel Institute of Technology), Evgeniy Gabilovich (Technion - Israel Institute of Technology), and Shaul Markovitch (Technion - Israel Institute of Technology)

Session VIb: Disambiguation

Venue: Students' Auditorium

Chair: Doug Oard

- 14.30 - 15.30 **Information Retrieval using Word Sense: Root Sense Tagging Approach**
Sang-Bum Kim (Korea University), Hee-Cheol Seo (Korea University), and Hae-Chang Rim (Korea University)
- An Effective Approach to Document Retrieval via Utilizing WordNet and Recognizing Phrases**
Shuang Liu (University of Illinois at Chicago), Fang Liu (University of Illinois at Chicago), Clement Yu (University of Illinois at Chicago), and Weiyi Meng (Binghamton University)
- Web-a-where: Geotagging Web Content**
Einat Amitay (IBM), Nadav Y. Har'El (IBM), Nadav Y. Har'El (IBM), Ron Sivan (IBM), and Aya Soffer (IBM)

- 15.30 - 16.00 **Coffee Break** (University House)

Full Conference Program

Tuesday 27 July

Session VIIa: Recognising and using named entities

Venue: Octagon Centre

Chair: Wei-Ying Ma

16.00 - 17.30 **Focused Named Entity Recognition Using Machine Learning**

Li Zhang (IBM China Research Laboratory), Yue Pan(IBM China Research Laboratory), and Tong Zhang (IBM T.J. Watson Research Center)

Learning Phonetic Similarity for Matching Named Entity Translations and Mining New Translations

Wai Lam (The Chinese University of Hong Kong), Ruizhang Huang (The Chinese University of Hong Kong), and Pik-Shan Cheung(The Chinese University of Hong Kong)

Text Classification and Named Entities for New Event Detection

Giridhar Kumaran (University of Massachusetts Amherst) and James Allan(University of Massachusetts Amherst)

Session VIIb: Efficiency and scaling

Venue: Students' Auditorium

Chair: Justin Zobel

16.00 - 17.30 **Assigning Identifiers to Documents to Enhance the Clustering Property of Fulltext Indexes**

Fabrizio(F) Silvestri (Università di Pisa), Salvatore(S) Orlando (Università di Venezia - Mestre), and Raffaele(R) Perego (Consiglio Nazionale delle Ricerche)

Filtering Algorithms for Information Retrieval Models with Named Attributes and Proximity Operators

Christos Tryfonopoulos (Technical University of Crete), Manolis Koubarakis (Technical University of Crete), and Yannis Drougas (University of California Riverside)

Hourly Analysis of a Very Large Topically Categorized Web Query Log

Steven M Beitzel (Illinois Institute of Technology), Eric C Jensen (Illinois Institute of Technology), Abdur Chowdhury (Illinois Institute of Technology), David Grossman (Illinois Institute of Technology), and Ophir Frieder (Illinois Institute of Technology)

Banquet

Venue: Cutler's Hall

19.00 - 23.00 **Banquet**

Full Conference Program

Wednesday 28 July

08.30 - 17.30 **Registration Opens** (Octagon Centre)

Session VIIIa: Content-based filtering & collaborative filtering

Venue: Octagon Centre

Chair: James Allan

09.00 - 11.00 **A Collaborative Filtering Algorithm and Evaluation Metric that Accurately Model the User Experience**

Matthew R. McLaughlin (Oregon State University) and Jonathan L. Herlocker (Oregon State University)

An Automatic Weighting Scheme for Collaborative Filtering
Rong Jin (Michigan State University), Joyce Y. Chai (Michigan State University), and Luo Si (Carnegie Mellon University)

Using Bayesian Priors to Combine Classifiers for Adaptive Filtering

Yi Zhang (Carnegie Mellon University)

A Nonparametric Hierarchical Bayesian Framework for Information Filtering

Kai Yu (Siemens AG), Volker Tresp (Siemens AG), and Shipeng Yu (University of Munich)

Session VIIIb: Image retrieval, users and usability

Venue: Students' Auditorium

Chair: Ian Ruthven

09.00 - 11.00 **Automatic Image Annotation by Using Concept-Sensitive Salient Objects for Image Content Represent**

Jianping Fan (University of North Carolina), Yuli Gao (University of North Carolina), Hangzai Luo (University of North Carolina) and

A Search Engine for Historical Manuscript Images

Toni M Rath (University of Massachusetts), R Manmatha (University of Massachusetts), and Victor Lavrenko (University of Massachusetts)

Display Time as Implicit Feedback: Understanding Task Effects

Diane Kelly (University of North Carolina) and Nicholas J. Belkin (Rutgers University)

Human versus Machine in the Topic Distillation Task

Mingfang Wu (CSIRO), Gheorghe Muresan (Rutgers University), Alistair McLean (CSIRO), Muh-Chyun Morris Tang (Rutgers University), Ross Wilkinson (CSIRO), Yuelin Li (Rutgers University), Hyuk-Jin Lee (Rutgers University), and Nicholas J. Belkin (Rutgers University)

11.00 - 11.30 **Coffee Break** (University House)

11.30 - 12.30 **Keynote Address: Chemoinformatics: an Application Domain for Information Retrieval Techniques**

Peter Willett (University of Sheffield)

12.30 - 14.00 **Lunch/ACM Business Meeting** (University House)

Full Conference Program

Wednesday 28 July

Session IXa: Machine learning for IR

Venue: Octagon Centre
Chair: Alistair Moffat

14.00 - 15.30 **Learning Effective Ranking Functions for Newsgroup Search**
Wensi Xi (Virginia Polytechnic Institute and State University),
Jesper Lind (Microsoft Research), and Eric Brill (Microsoft Research)

Language-specific Models in Multilingual Topic Tracking
L S Larkey (University of Massachusetts), F Feng (University of Massachusetts), M Connell (University of Massachusetts), and V Lavrenko (University of Massachusetts)

Web Taxonomy Integration through Co-Bootstrapping
Dell Zhang (National University of Singapore) and Wee Sun Lee (National University of Singapore)

Session IXb: Natural language processing

Venue: Students' Auditorium
Chair: Nick Belkin

14.00 - 15.30 **Evaluation of an Extraction-based Approach to Answering Definitional Questions**
J Xu (BBN Technologies), R Weischedel (BBN Technologies), and A Licuanan (BBN Technologies)

Query Based Event Extraction along a Timeline
Hai Leong Chieu (DSO National Laboratories) and Yoong Keok Lee (DSO National Laboratories)

Sentence Completion
Korinna Grabski (Otto-von-Guericke-Universitaet Magdeburg) and Tobias Scheffer (Humboldt-Universitaet zu Berlin)

15.30 - 16.00 **Coffee Break** (University House)

Session X: Web structure

Venue: Octagon Centre
Chair: R Manmatha

16.00 - 17.30 **Block-Level Link Analysis**
Deng Cai (Tsinghua University), Xiaofei He (The University of Chicago), Ji-Rong Wen (Microsoft Research Asia) and Wei-Ying Ma (Microsoft Research Asia)

Usefulness of Hyperlink Structure for Query-Biased Topic Distillation
V Plachouras (University of Glasgow) and I Ounis (University of Glasgow)

Block-based Web Search
Deng Cai (Tsinghua University), Shipeng Yu (University of Munich), Ji-Rong Wen (Microsoft Research Asia) and Wei-Ying Ma (Microsoft Research Asia)

17.30 - 17.40 **Conference close** 16

Full Conference Program

Thursday 29 July

Workshops

Venue: Arts Tower

09.00 - 17.00 **New Directions For IR Evaluation**
Anton Leuski (USC Institute for Creative Technologies) and Douglas W. Oard (University of Maryland)

RIA and "Where can IR go from here?"
Donna Harman (National Institute of Standards and Technology)

Peer-to-Peer IR
Jamie Callan (Carnegie Mellon University), Norbert Fuhr (University of Duisburg-Essen), and Wolfgang Nejdl (University of Hannover)

Information Retrieval in Context
Peter Ingwersen (Royal School of LIS), Keith van Rijsbergen (University of Glasgow), and Nick Belkin (Rutgers University)

Search and Discovery in Bioinformatics
Javed Mostafa (Indiana University, Bloomington) and Padmini Srinivasan (University of Iowa)

Integration of Information Retrieval and Databases (IR + DB)
Thomas Roelleke (Queen Mary University London) and Arjen P. de Vries (CWI)

XML and Information Retrieval
Ricardo Baeza-Yates (University of Chile) and Yoelle S. Maarek (IBM Haifa Research Lab)

Geographical IR
Ross Purves (University of Zurich) and Chris Jones (Cardiff University)

Mathematical Formal Methods
Sandor Dominich (University of Veszprém) and C.J. "Keith" van Rijsbergen (University of Glasgow)

Semantic Web
Ying Ding (University of Innsbruck), Keith van Rijsbergen (University of Glasgow), Iadh Ounis (University of Glasgow), and Joemon Jose (University of Glasgow)

Information Retrieval For Question Answering
Rob Gaizauskas (University of Sheffield), Mark Hepple (University of Sheffield), and Mark Greenwood (University of Sheffield)

10.30 - 11.00 **Coffee break**

12.30 - 13.30 **Lunch**

15.00 - 15.30 **Coffee break**

Posters

Image Based Gisting in CLIR

Mark Sanderson (University of Sheffield) and Robert Pasley (University of Sheffield)

Classifying Racist Texts Using A Support Vector Machine

Edel Greevy (Dublin City University) and Alan F. Smeaton (Dublin City University)

Discovery of Aggregate Usage Profiles based on Clustering Information Needs

Azreen Azman (University of Glasgow) and Iadh Ounis (University of Glasgow)

Merging Retrieval Results in Hierarchical Peer-to-Peer Networks

Jie Lu (Carnegie Mellon University) and Jamie Callan (Carnegie Mellon University)

The Effect of Back-Formulating Questions in Question Answering Evaluation

Tetsuya Sakai (Toshiba), Yoshimi Saito (Toshiba), Yumi Ichimura (Toshiba), Tomoharu Kokubu (Toshiba), and Makoto Koyama (Toshiba)

Effect of Varying Number of Documents in Blind Feedback

J Montgomery (Clairvoyance Corporation), Luo Si (Carnegie Mellon University), Jamie Callan (Carnegie Mellon University), and David A Evans (Clairvoyance Corporation)

Eye-Tracking Analysis of User Behavior in WWW Search

Laura A Granka (Cornell University), Thorsten Joachims (Cornell University), and Geri Gay (Cornell University)

Subwebs for Specialized Search

Raman Chandrasekar (Microsoft Corporation), Harr Chen (Microsoft Corporation), Simon Corston-Oliver (Microsoft Corporation), and Eric Brill (Microsoft Corporation)

Comparison of Using Passages and Documents for Blind Relevance Feedback in Information Retrieval

Zhenmei Gu (University of Waterloo) and Ming Luo (Virginia Tech)

Measuring Pseudo Relevance Feedback & CLIR

Mark Sanderson (University of Sheffield) and Paul Clough (University of Sheffield)

A Two-stage Mixture Model for Pseudo Feedback

Tao Tao (University of Illinois at Urbana Champaign) and ChengXiang Zhai (University of Illinois at Urbana Champaign)

Natural Language Processing for Browse Help

Eric Crestan (Sinequa) and Claude de Loupy (Sinequa)

Triangulation without Translation

James Mayfield (The Johns Hopkins University) and Paul McNamee (The Johns Hopkins University)

A Session-Based Search Engine

Smitha Sriram (University of Illinois at Urbana-Champaign), Xuehua Shen (University of Illinois at Urbana-Champaign), and Chengxiang Zhai (University of Illinois at Urbana-Champaign)

Evaluation of Filtering Current News Search Results

Steven M Beitzel (Illinois Institute of Technology), Eric C Jensen (Illinois Institute of Technology), Abdur Chowdhury (Illinois Institute of Technology), David Grossman (Illinois Institute of Technology), and Ophir Frieder (Illinois Institute of Technology)

The Document as an Ergodic Markov Chain

Eduard C Hoenkamp (University of Nijmegen) and Dawei Song (Queensland University)

Expertise Community Detection

Raymond J D'Amore (The MITRE Corporation)

Posters

Learning Patterns to Answer Open Domain Questions on the Web

Dmitri Roussinov (Arizona State University) and Jose Robles (Arizona State University)

Email is a Stage: Discovering People Roles from Email Archives

Anton Leuski (University of Southern California)

Searching Databases for Semantically-related Schemas

Gauri Shah (IBM Almaden Research Center) and Tanveer Syeda-Mahmood (IBM Almaden Research Center)

Topic Prediction Based on Comparative Retrieval Rankings

Chris Buckley (Sabir Research)

Context-Based Question-Answering Evaluation

Elizabeth D. Liddy (Syracuse University), Anne R. Diekema (Syracuse University), and Ozgur Yilmazel (Syracuse University)

Design of an E-Book User Interface and Visualizations to support Reading for Comprehension

Y Sun (The Robert Gordon University), D J Harper (The Robert Gordon University), and S N K Watt (The Robert Gordon University)

Toward Better Weighting of Anchors

David Hawking (CSIRO), Trystan Upstill (Australian National University), and Nick Craswell (CSIRO)

Aggregated Feature Retrieval for MPEG-7 via Clustering

Jiamin Ye (Dublin City University) and Alan F. Smeaton (Dublin City University)

Answer Models for Question Answering Passage Retrieval

Andres Corrada-Emmanuel (University of Mass. at Amherst) and W. Bruce Croft (University of Mass. at Amherst)

A Study of Topic Similarity Measures

R. W. White (University of Glasgow) and J. M. Jose (University of Glasgow)

Effectiveness of Web Page Classification on Finding List Answers

Hui Yang (National University of Singapore) and Tat-Seng Chua (National University of Singapore)

Detection and Translation of OOV Terms Prior to Query Time

Y Zhang (RMIT University) and P Vines (RMIT University)

Evaluation of the real and perceived value of automatic and interactive query expansion

Y. Nemeth (Ben-Gurion University), B. Shapira (Ben-Gurion University), and M Taeib-Maimon (Ben-Gurion University)

Information Extraction Using Two-Phase Pattern Discovery

L Ma (University of New South Wales) and J Shepherd (University of New South Wales)

Context Sensitive Vocabulary And its Application in Protein Secondary Structure Prediction

Y Liu (Carnegie Mellon University), J Carbonell (Carnegie Mellon University), J Klein-Seetharaman (Carnegie Mellon University), and V Gopalakrishnan (University of Pittsburgh)

Formal Multiple-Bernoulli Models for Language Modeling

D Metzler (University of Massachusetts), V Lavrenko (University of Massachusetts), and W B Croft (University of Massachusetts)

User Biased Document Language Modelling

L Azzopardi (University of Paisley), M Girolami (University of Glasgow), and C J van Rijsbergen (University of Glasgow)

A Unified Model of Literal Mining and Link Analysis for Ranking Web Pages

Xu Yinghui (Toyohashi University of Technology) and Kyoji Umemura (Toyohashi University of Technology)

Posters

A Joint Framework for Collaborative and Content Filtering

Justin Basilico (Brown University) and Thomas Hofmann (Brown University)

Refining Term Weights of Documents Using Term Dependencies

Hee-soo Kim (Ajou University), Ikkyu Choi (Ajou University), and Minkoo Kim (Ajou University)

Multiple Sources of Evidence for XML Retrieval

Börkur Sigurbjörnsson (University of Amsterdam), Jaap Kamps (University of Amsterdam), and Maarten de Rijke (University of Amsterdam)

Query-Related Data Extraction of Hidden Web Documents

YL Hedley (Coventry University), M Younas (Coventry University), A James (Coventry University), and M Sanderson (University of Sheffield)

The Patent Retrieval Task in the Fourth NTCIR Workshop

Atsushi Fujii (University of Tsukuba), Makoto Iwayama (Hitachi), and Noriko Kando (National Institute of Informatics)

Information Retrieval Using Hierarchical Dirichlet Processes

Philip John Cowans (University Of Cambridge)

A Study of Methods for Normalizing User Ratings in Collaborative Filtering

Rong Jin (Michigan State University) and Luo Si (Carnegie Mellon University)

A Review of Relevance Feedback Experiments at the 2003 Reliable Information Access (RIA) Workshop.

R H Warren (University of Waterloo) and T Liu (ILS University at Albany)

Supporting Federated Information Sharing Communities

B. Liu (The Robert Gordon University), D. J. Harper (The Robert Gordon University), and S. N.K. Watt (The Robert Gordon University)

The Effect of Document Retrieval Quality on Factoid Question Answering Performance

Kevyn Collins-Thompson (Carnegie Mellon University), Egidio L Terra (University of Waterloo), Jamie Callan (Carnegie Mellon University), and Charles L.A. Clarke (University of Waterloo)

Exploiting hyperlink recommendation evidence in navigational Web search

Trystan G Upstill (Australian National University) and Stephen Robertson (Microsoft Research Labs)

Context-based methods for text categorisation

William John Teahan (University of Wales) and David Samuel Hunnisett (CQRData)

eMailSift: Mining-based Approaches To Email Classification

Manu Aery (University of Texas at Arlington) and Sharma Chakravarthy (University of Texas at Arlington)

Constructing a Text Corpus for Inexact Duplicate Detection

Jack G. Conrad (Thomson Legal & Regulatory) and Cindy P. Schriber (Thomson-West)

Why Current IR Engines fail

Chris Buckley (Sabir Research)

Automatic Sense Disambiguation for Acronyms

Manuel Zahariev (Amware Enterprises Ltd.)

Filtering for Personal Web Information Agents

Gabriel L. Somlo (Colorado State University) and Adele E. Howe (Colorado State University)

Evaluating Content-Based Filters for Image and Video Retrieval

Michael G Christel (Carnegie Mellon University), Neema Moraveji (Carnegie Mellon University), and Chang Huang (Carnegie Mellon University)

Semantic Video Classification and Feature Subset Selection under Context and Concept Uncertainty

Jianping Fan (University of North Carolina at Charlotte) and Hangzai Luo

Demos

Implicit Queries (IQ) for Contextualized Search

Susan Dumais (Microsoft), Edward Cutrell (Microsoft), Raman Sarin (Microsoft), and Eric Horvitz (Microsoft)

Geotemporal Querying of Multilingual Documents

Fredric C Gey (University of California, Berkeley), Aitao Chen (University of California, Berkeley), Ray R Larson (University of California, Berkeley), and Kim Carl (University of California, Berkeley)

ACES: A Contextual Engine for Search

Xuehua Shen (University of Illinois at Urbana-Champaign), Smitha Sriram (University of Illinois at Urbana-Champaign), and Chengxiang Zhai (University of Illinois at Urbana-Champaign)

An Implicit System for Predicting Interests

R. W. White (University of Glasgow) and J. M. Jose (University of Glasgow)

Armadillo: Harvesting Information for the Semantic Web

Sam Chapman (University of Sheffield), Alexiei Dingli (University of Sheffield), and Fabio Ciravegna (University of Sheffield)

UKSearch - Search with Automatically Acquired Domain Knowledge

Udo Kruschwitz (University of Essex) and Hala Al-Bakour (University of Essex)

Geographic Information Retrieval (GIR): Searching Where and What

Ray R. Larson (University of California, Berkeley) and Patricia Frontiera (University of California, Berkeley)

Poster and Demo session with buffet will start at 18.00pm on Monday 26, July.

SIGIR'04 Conference Organization

Organizing Committee

General Chair:	Mark Sanderson (University of Sheffield)
Program Co-Chairs:	Kalervo Järvelin (University of Tampere) James Allan (University of Massachusetts) Peter Bruza (Distributed Systems Technology Centre)
Posters Chair:	Gareth Jones (Dublin City University)
Demonstrations Chair:	Ian Ruthven (University of Strathclyde)
Tutorials Chair:	Fabio Crestani (University of Strathclyde)
Workshops Chair:	Mounia Lalmas (Queen Mary University of London)
Mentoring Chair:	Douglas W. Oard (University of Maryland)
Best Papers Chair:	John Tait (University of Sunderland)
Doctoral Consortium Chair:	David Harper (Robert Gordon University)
Treasurer:	Micheline Beaulieu (University of Sheffield)
Assistant Treasurer:	Kathryn Cobley (University of Sheffield)
Local Arrangements:	Peter Willett (University of Sheffield)
Registration:	Daniela Petrelli (University of Sheffield)
Student Volunteers:	Yazdan Mansourian (University of Sheffield) Mark Stevenson (University of Sheffield)
Publicity & Technology Arrangements:	Hideo Joho (University of Sheffield)
Web Designer:	Stephen Levin (University of Sheffield)
Travel Fund:	Stefan Rürger (Imperial College London) Justin Zobel (Royal Melbourne Institute of Technology)
PC Meeting Organization:	Mounia Lalmas (Queen Mary University of London)

