Salut, our leetle Breeteesh chooms and bienvenu to zee latest sexy edition of zee Informer! In zees issue, we preezent details of Le Colloquium ‘98 <http://irsg.eu.org/ colloq/>. Why not join us een our mountain resort where zee croissants are warm and zee vin ees cold. We can watch zee sun set over zee Vercors, hold hands in zee moonlight, or we can leesten to zee sceentilating programme of talks preezented zees year.

Also, in zees edition, we have a winner for l’Inform competition - Stefano Mizzaro - and une profile on Mark Stairmand. Ooh la la!!!
Who's Who
IRSG Committee Contact
List 1997 - 1998

Wondering who you should contact about what? Well, here's the current list of contacts.

Chair
Mark Dunlop
Department of Computing Science
University of Glasgow
Tel: 0141 330 6035 Fax: 0141 330 4913
Email: mark@dcs.gla.ac.uk (or chair.irsg@bcs.org.uk)

Vice-Chair
John Lindsay
Kingston University
Tel: 0181 547 2000 ext 2699 Fax: 0181 547 7887
Email: lindsay@kingston.ac.uk

Secretary
Jane Reid
Department of Computing Science
University of Glasgow
Tel: 0141 330 5006 Fax: 0141 330 4913
Email: jane@dcs.gla.ac.uk

Treasurer
Jonathan Furner
School of Information and Media
The Robert Gordon University
Tel: 01224 283835 Fax: 01224 492608
Email: j.furner@rgu.ac.uk

Colloquium 98
Yves Chiaramella
CLIPS-IMAG Universite Joseph Fourier 38041 Grenoble cedex 9 Grenoble France
Email: chiara@imag.imag.fr

One-day events
Mounia Lalmas
Department of Computing Science
University of Glasgow
Tel: 0141 330 5323 Fax: 0141 330 4913
Email: munia@dcs.gla.ac.uk

David Harper
School of Computer and Mathematical Sciences
The Robert Gordon University
Tel: 01224 262701 Fax: 01224 262727
Email: d.harper@scms.rgu.ac.uk

Clare Harvey
School of Computing and Information Systems University of Sunderland
Tel: 0191 515 3508 Fax: 0191 515 2781
Email: clare.harvey@sunderland.ac.uk

Ruairi O'Donnell
Department of Information Science
Strathclyde University
Tel: 0141 552 4400 ext 3700/3096 Fax: 0141 553 1393
Email: rodonnel@dis.strath.ac.uk

Margaret Graham
Department of Information & Library Management University of Northumbria at Newcastle
Newcastle-Upon-Tyne, NE1 8ST
Email: margaret.graham@unn.ac.uk

Jan J IJdens
Sharp Laboratories of Europe
Oxford Science Park Oxford OX4 4GA Tel: 01865 747711 Fax: 01865 714170
Email: jan@sharp.co.uk

Informer Team
John Davies Information Access Research BT Laboratories
Tel: 01473 605536 Fax: 01473 642459
Email: john.davies@bt-sys.bt.co.uk

Frances Johnson
Department of Library and Information Studies Manchester Metropolitan University
Tel: 0161 247 6156 Fax: 0161 247 6351
Email: F.Johnson@mmu.ac.uk

Jon Ritchie (co-editor)
Department of Computing Science
University of Glasgow
Tel: 0141 330 6854 Fax: 0141 330 4913
Email: jon@dcs.gla.ac.uk

Ian Ruthven (co-editor)
Department of Computing Science
University of Glasgow
Tel: 0141 330 6292 Fax: 0141 330 4913
Email: igr@dcs.gla.ac.uk

Research Position
in Agent-Based Information Filtering
Computing Science Department,
University of Dortmund, Germany

For the EU-funded Telematics project EUROgatherer, the Information Retrieval group at the University of Dortmund is looking for a research staff member, starting 1 April 1998.

EUROgatherer involves three commercial European Web search engines and four research institutes from six countries. The project aims at the development and pilot application of agent-based methods for information gathering from Web-based information sources. Starting from user-defined interest profiles, the system will monitor relevant Web sites as well as Web databases and forward new, relevant documents to the user as soon as they become available. The major responsibility of our group will be the development of the filtering component.

We are looking for applicants owning a degree in computer science who have knowledge in at least one of the following areas: - Web-based information systems, - Information Retrieval, - Agent systems.

Due to the short duration of the project, the initial contract will be limited to 15 months. However, there may be possibilities for working in other projects afterwards. Our research group is involved in a number of other national and European research projects on digital libraries, Web-based information systems and multimedia information systems. Further details are available at http://ls6-www.informatik.uni-dortmund.de/ir/.

Payment will be according to German BAT IIa (approx. 70000 DM p.a.).

Please send your resume to: Prof. Dr. Norbert Fuhr, Informatik VI, University of Dortmund, D-44221 Dortmund, Germany
Wednesday

Session 1 (1400 - 1545)

Introduction
Chiaramella, CLIPS IMAG

NHS: A tool for the automatic construction of news hypertext
Dalamagas, University of Athens

Agents in cyberspace - towards a framework for multi-agent systems in information discovery
Wondergem, Van Brommel, Huibers and Van der Weide, University of Nijmegen

An architecture for efficient document clustering and retrieval on a dynamic collection of newspaper texts
Smeaton, Burnett, Crimmins and Quinn, Dublin City University

Session 2 (1615 - 1745)

A Dempster-Shafer model for document retrieval using noun phrases
Theophylactou and Lalmas, University of Glasgow

Modeling information retrieval with probabilistic augmentation systems
Picard and Haenni, University of Neuchatel

Efficient query processing using conceptual graphs
Ounis and Pasca, CLIPS-IMAG

Evening: Conference Dinner

Thursday

Session 3 (0900 - 0945)

Keynote: The logical approach to IR: models, applications and knowledge representation issues.
Amati, Fondazione Ugo Bordoni.

Session 4 (0945 - 1115)

Content based annotation and retrieval in RAIDER
Fountain and Tan, University of Reading

Text-image interaction for image retrieval and semi-automatic indexing
Duffing, CRIN-CNRS

German indexing and retrieval test database: some results of the pre-test
Kluck, Social Sciences Information Centre of Bonn

Session 5 (1145 - 1315)

User-chosen phrases in interactive query formulation for information retrieval
Smeaton and Kelly, Dublin City Uni. & Broadcom Eireann Research

Evaluation of interfaces for IRS: modelling end-user searching behaviour
Manglano, Beaulieu and Robertson, City University

Social and cognitive factors in the design and evaluation multimedia systems
Gabrielli, University of Padova

Evening: Conference Social Afternoon

Friday

Session 6 (0930 - 1130)

Combining information retrieval with information extraction for efficient retrieval of calls for papers
Lazarinis, University of Glasgow

The relationship between IR and multimedia databases

De Vries and Blanken, University of Twente

Independence of contributing retrieval strategies in data fusion for effective information retrieval
Smeaton, Dublin City University.

Coupled hierarchical IR and stochastic models for surface information extraction
Zaragoza and Gallinari, University Pierre and Marie Currie

Session 7 (1200 - 1300)

Keynote: IRSG Best Student Paper in IR competition winner

How many relevances in information retrieval?
Mizzaro, University of Udine.

Close
Dunlop, University of Glasgow

Winner

As you can see from the IRSG programme, the winner of this year’s Springer-Informer prize for best student paper in IR has been won by Stefano Mizzaro from the University of Udine. The judges were impressed by all the entries to the competition this year but were particularly keen on Stefano’s paper “How many relevances in information retrieval?”. Thanks to all who entered the competition and especially to all the judges!
What my research is about

Information Retrieval

Cohesion with Applications in

A Computational Analysis of Lexical

Title of thesis

1QD

Black, Department of Language

Ex-supervisor is Bill

Supervisor:

http://www.pac.soton.ac.uk

760834 Email: mas@pac.soton.ac.uk

Southampton SO16 7NP Tel: 01703

Centre, 2, V enture Road, Chilworth,

Contact details

Name: Mark Stairmand

Profile

Implementation of a generic textual

analysis based on lexical cohesion

analysis to identify topical themes and

three IR-related applications: a context-

based IR system, a process for

recognising topic boundaries in text, and

automatic production of a domain-

dependent lexical resource.

How I got into this:

The language engineering

department at UMIST is involved with

a wide range of application areas, and

my supervisor had a specific interest

in Language Processing techniques

applied to information retrieval. Whilst

following an MSc. in Cognitive

Science I had come across a paper in

the "Computational Linguistics" journal describing a computational

means of identifying topics in text

based on an analysis of lexical cohesion relations. At the time the

paper was written a suitable on-line lexical resource was not available to

the authors, and there hypothesis was based on a manual thesaurus-based

analysis. At the time of reading the

article I knew two on-line resources

were available, namely the GNU Roget

Thesaurus and Princeton University's

WordNet, yet I could find no reference
to work in more recent literature which
followed up the method described in
the paper. This sounded like a good
project to me, and it seemed that any
system which purports to identify topical themes should be put to the test
in the notoriously unforgiving IR domain!

My contribution to IR:

The thesis describes the design and

implementation of a system I chose to
call QUESCOT (after too long spent

acronym searching). This stands for
"Quantification and Encapsulation of
Semantic Context in Text". The system
takes raw ascii text as input and from
this produces a set of weighted topics
which it thinks are most pertinent in the
text. Several resources are required in
this process: a fast part of speech
tagging program, a morphological
analyzer, a means of recognizing
collocational terms, and a lexical
resource which describes semantic
relations between concepts. This latter
resource came in the shape of
WordNet, the lexical resource freely
available from Princeton University.

Although influenced by Morris and
Hirst's original paper published in "The
Journal of Computational Linguistics"
the algorithm used to identify topics
was different to that reported in the
paper. Lacking a suitable on-line
Morris and Hirst did not have the
benefit of applying their proposed
analysis to a large number of texts and
addressing any problems this revealed.

QUESCOT incorporated a variety of
heuristics which addressed problems
revealed through testing the process.

The motivation behind producing such a tool was that automatic topic
identification which accounts for the
context in which key terms appear had
the potential to substantially improve
IR performance. The next phase of the
PhD work involved producing an IR
system which would accept natural
language queries, pick out and weight
concepts in queries, and match these
queries against a pre-specified
document set. Prior to this, the entire
document set must have been analysed
with the QUESCOT to produce
'semantic indexes' which detailed the
pertinent concepts in a document, and
weighted these according to their
significance in the document in which
they occur. Having processed the entire
document set (90 000 documents
which required over 1 week of
computing time), it was possible to
produce an 'idf of topics', that is to
produce inverse document frequency
weights for topics rather than simply
for keywords. These figures could then
be used when determining the
significance of query items. The
hypothesis was that accounting for
text context of occurrence when indexing
documents would improve IR
precision performance.

The resulting system, named
COATER (Context-activated Text
Retrieval) was evaluated against the
SMART system. Although this is not
representative of a state-of-the-art IR
system, it is freely available, flexible,
and referred to in many experiments
so is useful for comparative purposes.
Evaluating the system with a set of 12
queries, based on relevance
judgements obtained from a group of
12 people for each query, indicated that
COATER retrieved documents more
relevant to the set of queries. However,
these results had to be tempered with
the fact that the queries chosen were
very simple and typically could not
contain proper names since these
concepts were not represented in
WordNet. The set of WordNet concepts
should be viewed as the system's
controlled vocabulary.

Other applications based on the set
of indices QUESCOT produced for the
document collection included an
automatic text segmentation process.
This is useful in IR since indexing
concepts with respect to segments,
rather than entire documents, has been
shown to result in improved retrieval
performance. However, evaluation
indicated that QUESCOT only
performed on a par with a much less
computationally expensive alternative
method for the same task. (cont on
page 7)
Informer Vol 8 Spring 1998

Conferences

DEXA '98


About the workshop

Today's technology has made it possible for databases to include different types of media objects such as text, image, audio and video. In particular, visual information systems are getting more popular than text-based information systems. The growing use of multimedia data has introduced new challenges in storage, retrieval and communication of these diverse multimedia objects. This has increased research in the area of image retrieval and multimedia databases. Some of the issues addressed and yet to be solved completely are:

* Appropriate and efficient representation methods for images need to be established.

* As these images and other multimedia objects are very large in size, we need better models and techniques for compressing these objects.

* One of the most important aspects of image communications is the quality of the image. The basic arbiter of image quality is the human visual system. Due to the complexity of the human visual system, it is hard to find an automated measure which accurately reflects human judgement. Therefore, we need to define some measures which can be used to evaluate the quality of an image more precisely.

* Better data structures have to be defined for storing and accessing images efficiently.

* Proper query languages have to be designed and implemented for retrieving images.

* Automatic object recognition techniques are needed to help in the process of feature based image extraction.

* Images need to be classified into separate groups in order to allow for better query performance.

The workshop seeks to bring together researchers from the image processing and image retrieval communities to discuss new approaches in their fields. The workshop encourages researchers to introduce new ideas and present experimental results from their work in progress. Participants are encouraged to contribute papers on any of the above topics or related areas.

Submission details

All papers will be refereed by at least 3 members of the DEXA'98 program committee. All accepted papers will be published by IEEE Computer Society Press in a volume of the DEXA'98 workshops.

Authors are invited to submit research contributions or practical experience reports in English. Papers should be double spaced and no longer than 5000 words. Authors should submit papers either:

* by email as uuencoded PostScript to dexa98@pluto.ct.monash.edu.au, or

* by posting 4 copies of their paper to the Workshop Chairs:

  Prof Bala Srinivasan, DEXA’98 Workshop, Department of Computer Technology, Monash University, Caulfield East, VIC - 3145, Australia.

  Prof A M Tjoa, DEXA’98 Workshop, Department of Computer Science, Technical University of Vienna, Austria

Submission deadline: March 20, 1998
Notification of acceptance: May 8, 1998
Camera ready copies: June 5, 1998

Second European Conference on Research and Advanced Technology for Digital Libraries
19 - 23 September, 1998
Heraklion, Crete, Greece
http://www.csi.forth.gr/2EuroDL
e dl@cc.uch.gr

Objectives

This conference is the second of a series of European conferences on research and technology for digital libraries funded by the European Commission's TMR Programme. Its objectives are: to bring together researchers from multiple disciplines whose science relates to the development of digital libraries; to provide an opportunity for these scientists to form a research community in Europe specific to digital library development and to enable them to discuss issues and strategies specific to the European context; to assist young researchers in establishing relationships with senior scientists in their areas of interest; to enable review and discussion of research under way in Europe, the US, Japan and other countries on digital libraries; to stimulate researchers, especially young scientists, to explore new areas of interest in digital library development; to establish a forum for discussion of issues specific to Europe such as interoperability, multilinguality, intellectual property policy, and information commerce; to provide an opportunity for researchers in the
relevant enabling technologies and information sciences, to discuss issues related to interoperability between world wide distributed digital libraries.

From a technical point of view, the European Conferences series aims to contribute to the definition of those digital library parameters which especially influence issues of access, retrieval, and interaction with information; to identify key problems which must be solved to make digital library services an effective reality; to identify a general structure or framework for integrating research and solutions; and to propose and encourage specific, high priority research directions within such a framework.

Topics
The conference organisers solicit papers on topics related to digital libraries, including but not limited to the following list:

- Digital Library Models, Frameworks, and System Requirements
- Metadata
- System Integration and Architecture Issues
- Interoperability, Scalability
- Networked Information Discovery, Agent Technologies
- Information Retrieval, Organisation, Navigation - Tools and Paradigms
- Multilinguality
- Role of Knowledge Representation Systems in Digital Library Interactions
- Collecting, Capturing, Filtering, Cataloging, Indexing,
- Preserving
- Intellectual Property Rights, Terms and Conditions, Rights Management
- Authoring, Electronic Publishing, Electronic Commerce and Information Economies
- Economic and Social Implications and Issues
- User Interfaces
- Handling of Graphics, GIS, Medical Data, Multimedia Information, Experimental Data and
- Scientific Models

Conference Programme
The conference will be held in Heraklion, Crete, Greece. Tutorials will be organised on the 19th and 20th of September 1998. The opening session will take place at 9.00a.m. on Monday the 21st of September 1998 and the final session will take place on Wednesday afternoon, the 23th of September 1998. Full details on the scientific programme of the Conference will be published on our Web site by the 1st of July 1998.

Important Dates
15 March 1998
Proposals for tutorials, panels and demos due to the Programme Chair

15 April 1998
Notification of tutorial, panel and demo acceptance

15 May 1998
Papers and proposals for posters due to the Programme Chair

25 June 1998
Notification of paper and poster acceptance

1 July 1998
Scientific Programme on the Web

25 July 1998 Final papers due

19-20 September 1998 Tutorials

21-23 September 1998 Conference

Panels
Suggestions for the organisation of panel sessions on one of the proposed topics or on related topics are welcomed. Proposals should include a short CV and position paper for each panelist.

Posters
During the conference a space will be reserved for poster sessions. Research projects of any scale are invited to illustrate innovative concepts and prototype systems.

Poster proposals should include title, names of presenters and outline (max. 500 words).

Tutorials
Tutoral days will be held before the Conference, on Saturday the 19th and Sunday the 20th of September 1998. Proposals for tutorials are solicited. Tutorials would be either half day (3 hours) or full day (6 hours). Each proposal should include a title, a summary (intentions, objectives, etc.), duration and a short CV of the instructor(s).

Demos
Result demonstrations of on-going projects are strongly encouraged. Those interested should submit a description of the intended Demo to the Programme Chair.

Papers
Papers (max 20 pages, double spaced) should be submitted electronically in HTML format, either by e-mail to the Conference Secretariat, ecdl@cc.uch.gr, or to our ftp site, ftp://ftp.ics.forth.gr/2EuroDL.

Proceedings
The Proceedings will be published by Springer as a volume in their Lecture Notes in Computer Science series and will be distributed at the Conference. Fellowship for Young Researchers

A limited number of fellowships for the Conference and also for Tutorials are available for young researchers who are citizens of European Union countries or Liechtenstein, Norway and Iceland. The fellowship offers free registration for the participants and, in special cases where necessary and appropriately justified, may pay for or reimburse travel and lodging expenses.
Programme Chair

Christos Nikolaou, University of Crete & IC8-FORTH, Leoforos Knossou, GR-71110 Heraklion, Crete, Greece Tel: +30 81 393199, Fax: +30 81 210106 E-mail: nikolau@cc.uch.gr

Contact Info

For more information regarding this conference contact the conference secretariat,

Rena Kalaitzaki and Maria Stavrakaki, University of Crete, Computer Science Department, Tel: +30 81 393504, Fax: +30 81 393501 E-mail: ecdl@cc.uch.gr

CoLIS3

Third international conference on concepts in library and information science

Inter-University Centre Dubrovnik (IUC) Dubrovnik, Croatia, 23 - 26 May 1999
http://www.ffzg.hr/infoz/colis3 http://www.scils.rutgers.edu/colis3 colis3@ffzg.hr

The general aim of CoLIS conferences is to provide a broad forum for critically exploring and analyzing library and information science as a discipline and as a field of research from historical, theoretical, and empirical perspectives.

Theme for CoLIS3

The goal is to critically explore the evolving concepts, research, and development related to digital libraries from the perspective of a number of approaches and disciplines. The conference aims to focus on fundamental and integrating issues and problems that reflect efforts and thinking from a number of disciplines and countries. Contributions are invited that address the topics listed below. They are further elaborated at conference web sites: Fundamental and theoretical aspects and questions underlying digital libraries. Nature of digital library collections. Organization of digital libraries. Access to digital libraries. Use of digital libraries. Technical infrastructure. Evaluation of digital libraries. Social issues. Economics of digital libraries. Integration of various kinds of information resources. Relations to traditional libraries - roots, comparisons, complements, integration.

Research papers. Full length papers (4 copies) of up to 12 pages covering any of the topics should be submitted to the Program Chair responsible for the geographic region. A Best Paper Award and a Best Student Paper Award are planned.

Poster presentations. Presentation of a work in progress that is best communicated in an interactive way. A synopsis (4 copies) of up to 3 pages should be submitted to the Poster Chair.

Panels. Proposals for panel sessions on issues and problems of general interest that will stimulate a lively debate should be sent to the Panels Chair.

Tutorials. Proposals for half or full day tutorials held before the opening of the conference should be submitted to the Tutorial Chair.

Workshops. Proposals for one day workshops held at the conclusion of the conference should be submitted to the Workshops Chair. Both, tutorials and workshops are designed to be self-sustaining economically to cover expenses of presenters.

Demonstrations and exhibits. Proposals for demonstrations of digital library projects or for exhibit of related materials should be submitted to Demonstration & Exhibit Chair.

Important dates and deadlines

Research papers deadline. 15 September 1998:

Deadline for posters, tutorials, workshops, panels, demonstrations and exhibits. 15 October 1998:


1 February 1999: Final camera ready copy.

(Contd from page 4) Finally, a domain-dependent lexical resource was produced based on post-processing of QUESCOT indices. This resource could automatically deduce clusters of related concepts in a corpus, useful for topic identification purposes. The novelty of this approach lay in the fact that the process produced clusters of related concepts, as opposed to potentially ambiguous terms. Early work was presented at BCS-IRSG colloquium held at Manchester in 1996, and a later presentation at SIGIR-97 ("Textual Context Analysis for Information Retrieval") referred to all the work carried out in the course of the ERSC-funded thesis work.

What I want to do next:

Since completing the thesis I have moved to Southampton, where I am employed in the Information Management Section of the Parallel Applications Centre. The work is very interesting, and one of the projects I am working on has a strong IR theme. Having been introduced to the field through this project (the MEMOIR project) my IR interest is now in collaborative information retrieval in large corporations.

Why you should read my thesis (in less than 10 words):

"It'll make your chosen acronyms seem positively inspired in comparison!".

If you are interested in PAC, where Mark now works and specifically to the IR-based Memoir project which he mentions, you can find more details (as well as job opportunities at: PAC home page: www.pac.soton.ac.uk Memoir home page: www.pac.soton.ac.uk/memoir Job offers: www.pac.soton.ac.uk/employ
On the move....

The Department of Information Studies at the University of Sheffield is delighted to announce that Prof. Micheline Beaulieu will be moving to Sheffield in May 1998 and will take over the Headship of the Department in August, at the end of Professor Willet's term as Head.

Prof. Beaulieu started her career as an Assistant Librarian, first at Manchester University and then at City University, and joined the Department of Information Science at City in 1987. She was awarded a Personal Chair for her research into interfaces for OPAC and IR systems in 1993, and has been the Head of Department there since 1996. She has extensive teaching experience and has served on many national and international committees on various aspects of librarianship and information science.

Peter Willet

...and on the up

News reaches us that John Davies has been made manager of the Online Information Research group at BT Labs.

When asked about the post, John said, "The new group pulls together BT's work in the areas of push technology, knowledge management and WWW-based information retrieval. I'm looking forward to what should be a very interesting programme of work. In the coming months, I will be seeking to strengthen our ties with academia."

Be a Star!

The Really Useful Show

Internet users offered 15 minutes of cyberspace fame on the BBC online ‘Web 100’

100 users of the BBC Online Internet site are being offered fame, free games and software, in an interactive initiative for the BBC one daily consumer programme The Really Useful Show, presented by Ruth Langsford and Tony Morris.

The live daytime show already has a consumer panel of 100 testers around the country, conducting consumer tests and reports for the programme. Now, with an established site on the Internet, which has an estimated 50,000 hits a week, The Really Useful Show, is looking for 100 Web users to help them carry out a variety of consumer tests, reports and polls on technology and entertainment - including new software, PC games, and previews of big movie releases.

Online shopping, security, design software, newsgroups and connection speeds are just some of the techno-subjects that the consumer Web 100 panel will be assisting The Really Useful Show, with.

Signing up to the initiative couldn’t be easier: users just send a simple e-mail marked ‘Web 100’, in which they must provide a brief review of their Internet Service Provider and why they chose them.

Internet users are invited to apply to become one of the Web 100 by accessing BBC Online’s The Really Useful Show site at www.bbc.co.uk/useful. Alternatively, users can call The Really Useful Show hotline number: 0500 500334 for further information.

The INFORMER

The Informer is published quarterly by the British Computer Society Information Retrieval Specialist Group.

Editorial Team: Jon Ritchie/Ian Ruthven, Department of Computing Science, University of Glasgow, Glasgow G12 8QQ. Tel: 0141 330 6292/6854 fax: 0141 330 4913. email: informer@dcs.gla.ac.uk

Advertising: Mark Dunlop, BCS IRSG Chair, Department of Computing Science, University of Glasgow, Glasgow G12 8QQ. email mark@dcs.gla.ac.uk