

Okapi's TREC success

At the recent Text Retrieval Conference (TREC) held just outside Washington DC, an experimental text retrieval system from the UK made an excellent showing in comparison with many experimental and commercial systems from around the world.

TREC is a large-scale experiment involving a number of research groups working on text retrieval. Each participating group takes on the same two gigabytes of text data and the same set of fifty search requests, and runs them through its own system. Output is sent to the US National Institute of Standards and Technology for assessment, and a number of performance measures are calculated. These measures are not independent, but focus on the ability of the system to find items which the assessors regard as relevant. Thus they attempt to measure effectiveness in a user-oriented sense, rather than efficiency in a time or cost sense.

Okapi is the experimental system of the Centre for Interactive Systems Research at City University, London. Its basic approach is statistical, deriving from a probabilistic model of the search process. It ranks the items retrieved in order of probable relevance, and invites the user to indicate which items he or she really likes. Okapi can then refine the search to find other similar items.

In this, the third annual round of TREC, there were two main categories of entries. In the routing section, where the search is assumed to be run regularly against incoming items, one Okapi run gave the best performance of all the results returned, on most performance measures. In the adhoc section, where the search is taken to represent a new requirement on an existing database, the best Okapi run was second on most measures. The top run in this section, on the INQUERY system from the University of Massachusetts, included some manual

intervention, while the Okapi run was automatic.

TREC has stimulated a revival of research interest in information retrieval theory and system design, particularly in the context of large (by experimental standards) text datasets. The format encourages friendly rivalry while promoting a real exchange of ideas and techniques. Some groups collaborate in various ways (for example, in an earlier round, a group at the University of California Los Angeles used the Okapi software for its entry). The Okapi team at City, which has been supported in its efforts by grants from the British Library, certainly learnt from the other participants in the first two rounds, and hopes to continue to take part in future TRECs.

The TREC conferences are reported (results of the tests, papers by the participants, system descriptions etc) in the proceedings, edited by Donna Harman and published by NIST, Gaithersburg, MD, USA. Proceedings of TREC—3 will be published in early 1995. A special issue of Information Processing and Management, also early 1995, will contain papers based on the first two TRECs, including one about Okapi.

Further information about TREC can be obtained from Donna Harman at NIST, and about Okapi from Stephen Robertson at City University, London.

- Steve Robertson (ser@is.city.ac.uk)

Editorial

So here we are with issue number two of the Informer and not very long after number one. It's not that I decided to do it early, it's because the last one was late. About four months late to be precise, since there was a gap between when Mark Sanderson's editorship finished and when I realised mine had started. Then another gap between when I started collecting

SIGIR funding, but hurry!

The European Information Retrieval Specialist Group of the Council for European Professional Informatics Societies (CEPIS/EIRSG) invites applications from young researchers for funding support to attend the 18th International ACM-SIGIR Conference in Seattle, USA in July 1995. Funding is restricted to European or European-based researchers aged 35 or younger on the date of the conference, who work in information retrieval or one of its related areas. Further details of the application procedure may be obtained by email from eirsg-fund@compapp.dcu.ie. The closing date for applications is **24th March 1995**.

"Internet users know what sex is" shock

Emeryville, Calif. - Feb. 22, 1995 - Ziff-Davis Press announced the February 1995 release of the book, *net.sex* featuring the myth-shattering results of the first comprehensive methodologically sound sexual knowledge test completed by Internet users. This comprehensive examination, performed entirely online, reveals that Internet users passed the test, scoring on average a grade of "C," two full grade points higher than the failing "F" grade received by the "offline" population polled by The Kinsey Institute.

Info: computerlife@delphi.com

copy and when I actually produced the finished product. Sorry about that.

Not much to say in the way of editorial really so I guess I'll see you all at the colloquium in April. You'd be best to avoid me or else you can expect to be asked to send me some news, write an article, pen a regular column, take over as editor, etc.


- Mark Magennis (Ed.)

Editor:

Mark Magennis
Department of Computing Science
University of Glasgow
Glasgow G12 8QQ

email: mark3@dcs.glasgow.ac.uk
tel: 0141 339 8855 x8452
fax: 0141 330 4913

The BCS IRSG is sponsored by

 Health Libraries & Information
Network

The Wonderful World of the Internet

In this issue, the WWoI reports on a search tool, a mailing list, some ominous developments and a couple of information resource infrastructure type projects (both destined to "revolutionise the way we ..."). In true News at Ten style we finish with a heart-warming story about everyone's favourite burger clown, Ronald McDonald.

WWWorm

The WWWorm is an example of an Internet worm, a program that wriggles around the web from document to document looking for things that match some search statement. The WWWorm basically does string matching on any objects it finds that are one of the following of four types:

1. URLs of WWW documents
2. Titles of WWW documents (every document has one, made up by the author to describe its contents)
3. The text of links
4. URLs that the links point to (links can point to WWW documents or other WWW objects such as movies, pictures, sounds, gopher objects)

WWWorm uses the UNIX `egrep` command to perform database searches. If you are not familiar with `egrep`, it is advisable to restrict yourself to single keyword searches. Also, because the database being searched is huge, expect to wait a while.

<http://www.cs.colorado.edu/home/mcbryan/WWW.html>

Document Search and Retrieval mailing list

The Document Search and Retrieval Mailing List is a forum for the discussion of topics relating to organisation and retrieval of electronic documents. This is an opportunity to relay information about new products, developments, and methods for finding, organising and using the ever growing amount of information stored on your computer.

To subscribe send mail to `Majordomo@imsworld.com` with the following command in the body of your email message:

`subscribe search <username>`

Internet Public Library

The University of Michigan School of Information and Library Studies proudly announces the advent of the Internet Public Library (IPL), an innovative, online, 24 hour public library designed to revolutionize the way the world thinks about library services. Chaired by an assemblage of internationally renowned librarians and leading information industry professionals, the Internet Public Library will offer an exciting vision of the library of tomorrow as envisioned by many of the brightest talents in the field today.

With a stated mission to "provide services and information which enhance the value of the Internet to its ever expanding and varied community of users," IPL is prepared to provide essential library services to a target audience estimated to number 1/4 of the entire American population by the end of the century. Among the first services to appear will be an online reference division; a youth services division; a user education division; and professional services for librarians.

A library without windows, walls, or even books, IPL will still provide a user friendly spot to turn to for questions on how to plan a family budget; learn more about the world of Internet based resources; or even turn to for a story hour for children. Bringing the best features of the community library forward into a new technological environment, IPL seeks to challenge our thinking about new "communities" that will arise in the future and a broader range of services the library of tomorrow might provide.

The Internet Public Library is scheduled to go online March 17, 1995.

To join a mailing list for information, mail `Majordomo@sil.sil.umich.edu` with the following command in the body of your email message:

`subscribe ipl-update <username>`

ISWorld Net

The blurb:

"We will provide information management scholars and practitioners with a single entry point to resources related to information systems technology and promote the development of an international information infrastructure that will dramatically improve the world's ability to use information systems for creating, disseminating, and applying knowledge".

This is another one of those "We

don't know what we're creating or what it's going to be used for but we're sure it will revolutionise the way we work, think, interact, eat, sleep, drink, etc.", but the inspiration for this project seems to have come from baseball, new-born children and slices of salami, so it must be interesting. But joking aside, it's this kind of thing that may well produce some really useful networking.

<http://www.cox.smu.edu/mis/iswnet/home.html>

U.S. Network Policing Proposed

Exon Bill seeks to make system operators liable for messaging content.

Senator Jim Exon has proposed the Communications Decency Act, 1995, in the United States Senate. In an effort to stamp out digital pornography, it makes all telecommunications providers doing business in the United States (from the telephone companies all the way down to offices that use LANs) liable for the content of anything sent over their networks. To avoid the possibility of tens of thousands of dollars in fines and up to two years in jail, business owners would be forced to police their networks and monitor in advance all messages sent over them.

This bill is substantially the same as the one he put forward last year. He will offer it as an amendment to the pending telecommunications deregulation legislation in the U.S. Senate, which is expected to be enacted by July. Last year, his amendment was adopted even though many thought it hastily drafted and poorly thought out. Fortunately, the telecommunications deregulation legislation died. This year, a more conservative U.S. Congress may be even more reluctant to challenge a "morality" amendment; and its legislative vehicle, the telecommunications deregulation legislation, stands a much better chance of passage this year.

- Electronic Messaging Association (info@ema.org)

Microsoft Internet Assistant risk

At the end of January middle of February this year Microsoft will be introducing Internet Assistant, an HTML creator and

World of Internet cont...

WEB browser for Word for Windows 6.0. The WEB browser will read Word 6.0 documents directly. Word documents can come with programming that will activate on opening. While this has always been a problem document distribution hasn't generally been widespread until soon from now.

Three types of things I can see happening.

1. Viral type documents. These are documents that will change your normal.dot and copy itself from document to document.

2. Trojan Horse type 1 documents. These are documents that do something on opening, like delete files etc....And possible even harmless things.

3. Trojan Horse type 2 documents. Really scary documents that communicate BACK to the web-server without your knowing it and sending additional information gleaned from your machine and or network. There are some truly scary things that could be done with a creative VBA/CGI programmer.

It is unfortunate that these risks exist, because otherwise the ability to have "programmable" documents on the web is a really cool concept. But nonetheless risks like these have to be dealt with.

- author unknown

Similarly to 3. (above) there has been at least one instance of a company automatically downloading information about users' machine configurations (processor type, etc.) when they access a Web page. After a number of complaints the company stopped doing it. - Ed.

Ronald McDonald's name already taken

Poor McDonalds. Not only are they fighting to save their good name in the McLibel case (<http://anthfirst.san.ed.ac.uk/McLibelTopPage.html>), but when they save it they won't even be able to use it on the Internet. Joshua Quittner, a WIRED magazine contributor, has registered himself as ronald@mcdonalds.com. McDonalds Corporation complained to the InterNIC, who promptly withdrew the domain registration from Quittner, only to reinstate it a week later. This has now caused a rush of people registering names of famous organisations, mostly in the hope that the organisation will want to buy it off them at some later date for a large sum of cash. There are, however, reports of companies registering the names of their rivals to prevent them from using them. Expect a rash of court cases in which corporate giants crush all opposition.

Aslib Telematics Grant Awareness

Aslib, The Association for Information Management is holding an Awareness Day for those interested in cooperation with the European Commission for developing products, software applications or services in the field of Telematics. Financial and Technical Managers and Directors will benefit from attending this seminar.

Representatives from the European Commission in Luxembourg will be on hand to give you expert guidance and practical help in finding funds, partners and making a case for assistance in developing your technology and products.

Calls for proposals in the Information Engineering and Libraries sectors will be released in March 1995. The Department of Trade and Industry has sponsored Aslib to inform British industry and

academia about the opportunities opened up to them by this 4th Framework programme, worth over 6 billion pounds sterling. Special facilities are being created for small businesses within the programme, and Aslib will be announcing a timely new Internet service to support this sector.

A videoconference will take place during the seminar, between London and Luxembourg, allowing delegates to discuss their projects with EC representatives.

24 March 1995, 10.00 - 15.00
Institution of Electrical Engineers, Savoy Place, London WC2
£35 (Aslib does not charge VAT on training events).
Lunch and refreshments will be provided.

For details contact:
Barbara Hobbs
Tel: +(44) 171 253 4488

e-mail: training@aslib.demon.co.uk

Are you Being Formed?

by John Lindsay

For most of the 1970s I wrote a series of articles in Libraries and Social Change, on that topic. Now I've been asked to return to the theme, but this time in the more general context of information retrieval.

People who have heard me sounding off at meetings will know that I am not in sympathy with the dominant paradigm of IR, where information is regarded as the property of an object, a document or a collection; but is rather a construction: a relationship between a user and a search space. Developments in the Internet has been the major significant change in search space, and therefore in the capacity of knowing subjects to be informed. Therefore changes in the political, economic and social construction of that space are important for users to know what is happening.

Participants at the last BCS IRSG colloquium might remember my summary, which Ruben chose not to include in the proceedings. For those who were not there, or who don't have perfectly retentive memories, here is the URL: <http://infosys.kingston.ac.uk/kingston/isschool/jlpapers/Thatcher.html>. But that was nearly a year ago and before open.gov.uk.

Open government in Britain must be a contradiction? That is only the beginning for the confusion they (it?) is likely to get into with the impact of the Internet on the relationship between the citizen and the state.

Incentive came from the Gore proposition for a national information infrastructure in the US and the Bangemann report on information society in Europe. The result was a position paper by the CCTA on information superhighways in Britain. The influence of the US shows:

- encourage private investment
- provide and protect competition
- provide open access to the network
- avoid creating information "have/have nots"
- encourage flexible and responsive government

so the CCTA suggest that for Britain policy should revolve around:

- improve public access to information
- improve public access to the democratic process

Are you being formed? cont...

- improve efficiency, effectiveness and responsiveness to the public
- promote the use of email for inter-departmental communication
- reduce the amount of paper or paper-based transactions used in government.

In order to facilitate the process a series of consultative groups on open government (cogs) have been set up, including *policy@gate.ccta.gov.uk*, from which these points are drawn.

What is any of this going to mean for us? At this stage it is too early to say. With the commercialisation of the Internet, all the issues of pricing, costing, charging, payment and funding are likely to become more complicated and within Higher Education in Britain the role of the Joint Information Services Council (JISC) will probably become more important in negotiating deals for complex licensing arrangements. Meanwhile individual universities will find it more complicated to decide how to provide services and navigating around search spaces will become harder. For commercial organisations, building the support environment for information retrieval appears not to have even started.

- John Lindsay (*lindsay@kingston.ac.uk*)

ESF scientific network

European Science Foundation (ESF) 'scientific network' on "Converging Computing Methodologies in Astronomy" makes its debut. A new ESF scientific network provides funding for work visits and workshops over a 3-year period. Among central topics are:

From vision models to image information retrieval. Methods such as wavelets and multiresolution approaches, mathematical morphology, and fuzzy methods have proven their worth in the framework of accessing appropriate information from large image databases. Such methods must be moulded together to allow semantically driven access to data.

The data life-cycle - methodological aspects. The astronomical data life-cycle is highly digital: data capture is increasingly on CCD electronic detectors, data are subject to image processing and statistical treatment, and the final major stage in this process involves data archiving, and publication. Not surprisingly, the

Announcing IDOM-Web Server

The IDOM-Web Server for Databases, Information Retrieval & Multimedia

IDOM-Web is a World-Wide-Web server set up by the IDOMENEUS European ESPRIT Network of Excellence (No. 6606). IDOM-Web serves academic and industrial users by providing a consolidated and focused view on recent R&D activities in the convergence areas of databases, information retrieval and multimedia in Europe.

As of today, IDOM-Web gives you online access to the following information resources maintained by the IDOMENEUS administrative servers.

- A database of up-to-date research site descriptions in Europe. Currently, this list covers about 40 leading groups in the field of DB, IR and MM. This database gives extensive site information and it lists key publications, research projects and technical achievements. Some sites also give electronic access to abstracts and full paper versions of their latest technical reports.

- An Online Database of Acronyms and Standards (OASIS) relevant to databases, information retrieval and multi-media. Where possible, the acronym descriptions

are enriched with WWW pointers to related Internet information resources.

- An online (searchable) bulletin-board and calendar of conference, workshop and summer-school announcements containing the latest description of the event (e.g. submission dates, schedule, registration details, ...) as distributed, for example, by international mailing lists like "DB-World".

- References to the EU Fourth Framework Programme (FP4), its calls, deadlines and background; there is also an interactive service intended to help formation of consortia and writing of IDOM-related FP4 proposals.

- Information on the IDOM Network itself, such as a project descriptions and reports from the IDOM working groups.

- The IDOM Front Desk explains how you can contribute actively to IDOM-Web

- announcing information via IDOM Web and the IDOM mailing list

- joining IDOM as an Associate Member or Technical Contributor to be included in the IDOM-Web site descriptions and the IDOM consolidated bibliography.

The IDOM-Web server has been operational since June 1993. Some of its services are supported by ESPRIT Basic Research. It can be found at <http://idom-www.informatik.uni-hamburg.de>

- Joachim W. Schmidt, NoE Coordinator (*J_Schmidt@dbis1.informatik.uni-hamburg.de*)

issues of electronic publishing and of digital libraries are increasingly central.

From data integration to information integration. Particular data integration (data fusion) problems, such as integration of data associated with different wavelength ranges, are of great relevance in the context of large space- and ground-based observing projects (e.g. co-addition in image restoration; image restoration and filtering approaches which incorporate semantic information on the cosmic objects of interest; close, complementary use of multi-million object astronomical catalogs; classification of terabyte data collections). Long-term access to stored data - what should be the "future of 'society's' memory?" Beyond data, astronomy is all about information. Compression is central - in a broad sense, compression is summarization, and therefore is part of the overall process of scientific analysis.

Coordinating committee of the scientific network is A. Bijaoui (Nice), V. Di Gesu (Palermo), A. Heck (Strasbourg), M.J. Kurtz (Harvard), P. Linde (Lund), M.C. Maccarone (Palermo) - Chair, R.

McMahon (Cambridge), R. Molina (Grenada), F. Murtagh (Munich) - Secretary, E. Raimond (Dwingeloo).

Further information available from <http://www.hq.eso.org/conv-comp.html>

Joke

There's a Doctor, an Information Scientist, and a Computer Scientist debating which is the oldest profession. The Doctor says "ah, that's easy - in the beginning God created man, Adam and Eve, biology, anatomy,.... The Medical profession has to be the oldest". "Ah ha", says the Information Scientist. "No, in the beginning there was chaos and when God created the earth he imposed order on the scheme of things, that is our job. The Information profession has to be the oldest". "Well, says the Computer Scientist. Who do you think created the chaos in the first place?"

- Frances Johnson

Documentation Abstracts Inc. Research Grant

The board of directors of Documentation Abstracts Inc. (DAI) announces it will award annually the Information Science Abstracts (ISA) research grant of \$1,000 to one or more information professionals to conduct a research project oriented toward the study of the primary or secondary literature of information science. European researchers may apply for this grant.

Documentation Abstracts Inc. is owner/sponsor of the monthly abstracting and indexing publication, Information Science Abstracts (ISA), published for DAI by Plenum Publishing Corporation. International in scope, ISA provides references to and abstracts of the useful literature in the fields of information science and library science dating back to 1966. Its coverage extends to journal articles from over 300 journals, to conference proceedings, books, reports and patents. The entire file is available for searching via Dialog File 202 and on Silver Platter CD-ROM.

All applicants for the ISA Research Grant must be information professionals and hold a graduate degree in library science or information science. No individuals who are associated with ISA are eligible. This includes members of the Board of Directors of DAI, employ-

ees of Plenum Publishing Corporation, Silver Platter, and Dialog.

Applicants must submit a completed application package by June 30, 1995 outlining the scope and nature of the proposed project, providing evidence of an established methodology and a viable research design. Examples of possible topics for research include the use of information resources, comparison of tools, quality of the literature, and bibliometric analysis. The goal of the research must be to produce a publishable paper.

Half the amount of the \$1,000 grant will be paid upon announcement of the award and the balance will be paid upon successful completion of the research project. In addition to the cash award, access to ISA CD-ROM will be provided for the length of the grant.

For further information and for application instructions:

Taissa Kusma, Vice-Chair of Documentation Abstracts Inc.,
Director, Electronic Publishing, American Institute of Physics,
500 Sunnyside Boulevard, Woodbury, NY 11797,
516/576-2260,

Taissa@aip.org FAX 516/576-2499

MSc Advanced Information Systems

University of Glasgow, Computing Science Department

The MSc in Advanced Information Systems at the University of Glasgow is a new advanced masters programme designed to give students the intellectual and practical skills required to address the explosive growth in large-scale, networked, information systems. With world-class research groups in databases, information retrieval and interactive systems, the Computing Science Department at Glasgow is uniquely placed to provide a course to develop these skills. This one year advanced masters is composed of five components:

- database and persistent systems
- information retrieval

- interactive multimedia systems
- an introductory foundations fortnight
- a large project resulting in a thesis

These components are designed to address the challenges of expanding access to large collections of information, for example over the internet. As well as technical issues of efficient and accurate searching of traditional data many other issues need to be addressed, such as accessing non-structured text and multimedia information, providing access to computer naive users, social implications of large information systems, browsing based access and general user interface issues. The graduates of this course will be well qualified for research or development work in academia, industry, commerce, or the public sector.

The foundations fortnight gives an overview of the three areas and covers introductory material which is common to the other components. This is followed by three specific courses on the main topics of the programme - all of these modules are compulsory, as the course is

already specialised, and contain a high literature reading and presentation element. Finally, a research project must be carried out, under weekly supervision, to MSc standards. The course is assessed primarily on examinations of the main topics and submission of the research project.

Although all successful students will graduate with an advanced MSc after one year, many are expected to carry on into PhD programmes (either in Glasgow or elsewhere). A review point of 8 months into the course will allow limited streaming so that potential PhD candidates can have their project orientated towards their continuing studies.

The Department of Computing Science at Glasgow is a large department which has achieved the highest assessment in both research and teaching - having received a 5 rating in the last UK research assessment and a rating of excellent in the 1994 SHEFC teaching assessment. The department currently has about 120 students on an MSc IT degree and around 40 on our PhD programme. The University has an overall postgraduate student population of approximately 2 500 and has extensive sports and social facilities (including a dedicated postgraduate research club).

For further information on the MSc in Advanced Information Systems either contact us for a paper booklet (and application form) or look at the information on our web pages.

MSc AIS Information
Computing Science Department
University of Glasgow
Glasgow G12 8QQ
Scotland
UK

E-mail: *ais@dcs.gla.ac.uk*
Phone: +44 141 330 6041
Fax: +44 141 330 4913
Web: *http://www.dcs.gla.ac.uk/courses/MScPhD_AIS/*

- Mark Dunlop, course director
(*mark@dcs.gla.ac.uk*)

Informer archive

Many thanks to Malcolm Kendall for unearthing almost the entire back-catalogue of IRSG newsletters all the way from June 1983. However, we're still missing issues 1, 2, 5-8, 18, 19, 40, 42 and 43, so if you've got some old issues gathering dust in a corner of your office then we'd be pleased to have them.

Conferences, Workshops and Seminars

Calls for papers

IEE Developments in Personal Systems

IEE Computing and Control Division
Professional Group C10 (Consumer and Domestic Systems)
Colloquium on Developments in Personal Systems

23rd June, 1995
Savoy Place, London

The Colloquium will explore the development of computing technology towards truly personal systems which provide users with freedom of mobility, data and functionality. Contributions are invited that discuss:

- developments in enabling technologies for personal and mobile computing systems (e.g. wireless and IR and RF networks, flatpanel displays, digital audio/video, PC Cards);
- the development of Personal Communications Networks (PCNs) and other communications services which increase an individual consumer's access to the global digital medium;
- the impact of 'ubiquitous' computing technology;
- 'quality of working life' issues created by the development of personal systems and other advanced communications technologies;
- innovative case studies of the conception, design, development and use of personal systems and devices (e.g. personal communicators, personal digital assistants, the 'digital wallet' or computer-telephony integration applications).

The emphasis should be on applications which break new ground and on should provide information which is accessible to wide audience of research scientists, application developers and decision-makers. Case studies of particular personal systems in both academic or commercial contexts are welcomed. Research studies which evaluate technology developments, as well as contributions which explore current trends and future possibilities, are also invited.

Prospective authors are invited to submit a summary of approximately 500 words before 1 April 1995 to:

Professor Peter J. Thomas
Centre for Personal Information Management
Faculty of Computer Studies and Mathematics
University of the West of England
Coldharbour Lane
Bristol BS16 1QY.
Tel: 0117 9762529
Fax: 0117 9763973

Email: Peter.Thomas@csm.uwe.ac.uk

ASIS SIG/CR Workshop

6th ASIS SIG/CR Classification Research Workshop: An interdisciplinary meeting

October 8th, 1995
Chicago

The American Society for Information Science Special Interest Group on Classification Research (ASIS SIG/CR) invites submissions for the 6th ASIS Classification Research Workshop, to be held at the 58th Annual Meeting of ASIS in Chicago, IL. The workshop will take place Sunday, October 8th, 1995, 8:30 a.m. - 5:00 p.m. ASIS '95 continues through Thursday, October 12th.

The CR Workshop is designed to be an exchange of ideas among active researchers with interests in the creation, development, management, representation, display, comparison, compatibility, theory, and application of classification schemes. Emphasis will be on semantic classification, in contrast to statistically based schemes. Topics include, but are not limited to:

- Warrant for concepts in classification schemes.
- Concept acquisition.
- Basis for semantic classes.
- Automated techniques to assist in creating classification schemes.
- Statistical techniques used for developing explicit semantic classes.
- Relations and their properties.
- Inheritance and subsumption.
- Knowledge representation schemes.
- Classification algorithms.
- Procedural knowledge in classification schemes.
- Reasoning with classification schemes.
- Software for management of classifica-

tion schemes.

- Interfaces for displaying classification schemes.- Data structures and programming languages for classification schemes.
- Image classification.
- Comparison and compatibility between classification schemes.
- Applications such as subject analysis, natural language understanding, information retrieval, expert systems.
- Representation and access on the Internet

The workshop's early registration fee is \$35.00 for SIG/CR members and/or participants; \$45.00 for ASIS members; \$60.00 for non-members. (The workshop is separate from the ASIS Annual, an additional registration is required for the Annual Conference).

Previous proceedings are titled "Advances in Classification Research: proceedings of the ... ASIS SIG/CR Classification Workshop" and are published by Learned Information, Inc., Medford, NJ.

Submissions should be made by email, or diskette accompanied by paper copy, or paper copy only (fax or postal), to arrive by April 15, 1995, to:

Ray Schwartz, 530 Jefferson St., #13,
Hoboken, New Jersey 07030, USA
Work Phone: 212-305-3294; Fax: 212-305-6193;

For additional information:
email rps4@columbia.edu
URL: <http://www.columbia.edu/~rps4/sigcr.html>

IEE Information Overload

IEE Computing and Control Division
Professional Group C5 (Human-Computer Interaction)
and C10 (Consumer and Domestic Systems)
Colloquium on Information Overload

29th November, 1995
Savoy Place, London

The phenomenal growth in the use of information and communication technologies has had the result that the volume of information available to individuals, at least in the Western World, has dramati-

c.f.p.'s...

cally increased. Even those individuals who have limited access to the 'global information superhighway' - made possible by developments in the Internet, for example - are experiencing what has been termed 'information overload'. This phenomenon, widely noted in research into the social dimensions of information technology, is acknowledged to be a problem which has to be addressed in the design of future information and communication technologies. As a result, there is an increasing interest in the development of information and communication technologies which assist users in managing information. These new generations of interactive systems have a requirement not only for usability, but also for actively aiding users in managing information.

A colloquium is being organised by Professional Groups C5 (Human-Computer Interaction) and C10 (Consumer and Domestic Systems) to consider the latest developments in information management in interactive systems both from the user's and system developer's perspective. The Colloquium will provide an opportunity to learn about the phenomenon of 'information overload', new developments in information management, and provide an opportunity for attendees to update their knowledge on the state-of-the-art interactive information systems.

Topics that are intended to be addressed in the Colloquium include, but are not limited to:

- New Developments in Information and Communication Technologies (e.g. the 'information superhighway', Internet, World-Wide Web, personal communications networks and interactive multimedia);
- Interface Design and Information Management (e.g. techniques and guidelines for the design of the user interfaces to complex, information-rich systems in order to combat information overload);
- Systems Development Issues (e.g. security, access, global information systems);
- Social Aspects of New Technologies (e.g. the dangers of information piracy, and electronic copyright and theft issues).

The Colloquium will be held at the IEE, Savoy Place, London on Thursday 30th November, 1995.

Prospective authors are invited to submit a summary of approximately 500 words by 1st May 1995 to either:

Dr. Jan Noyes
University of Bristol
Department of Psychology
8 Woodland Road
Bristol BS8 1TN
Tel. 0117-9288560
Fax. 0117-9288588
EMail *J.Noyes@uk.ac.bristol*

Professor Peter J. Thomas
Centre for Personal Information Management
Faculty of Computer Studies and Mathematics
University of the West of England
Coldharbour Lane
Bristol BS16 1QY
Tel. 0117-9762529
Fax. 0117-9763973
EMail *Peter.Thomas@csm.uwe.ac.uk*
Compuserve: 100255,2506.

The emphasis for papers should be on concepts and developments which 'break new ground' and should provide information which is accessible to a wide audience of research scientists, application developers and decision-makers. Case studies of particular information management systems in both academic or commercial contexts are welcomed. Research studies which evaluate technology developments, as well as contributions which explore current trends and future possibilities, are also invited.

Eastern European Librarianship Workshops

Workshops on Librarianship and Information Science to be presented in Eastern Europe and the Former Soviet Union between June and December, 1995

Please respond directly to:
Open Society Institute
Regional Library Program
Arvacska 6
H-1022 Budapest, Hungary
PHONE : 36-1-2122029
FAX : 36-1-2122092

The Regional Library Program, Open Society Institute, invites proposals to conduct three to five day workshops in Eastern Europe and the former Soviet Union between June 1 and December 15, 1995. We anticipate offering a total of 8 workshops in the following countries: Belarus, Croatia, the Czech Republic, Hungary, Poland, Kazakhstan/Kyrgyzstan, Lithua-

nia and Slovenia during this period. These workshops will be a part of a continuing series that began in 1994; other workshops may be planned for 1996 as well.

The mission of the Open Society Institute (OSI) is to develop, plan and implement regional programs and policy initiatives in the areas of education, economic, legal, and social reform in Central and Eastern Europe. As an umbrella organization, OSI provides administrative and financial support to various Soros organizations in the region.

Funding for the training workshops in librarianship and information science is immediately available.

The Regional Library Program has targeted five topics:

- Project Management
- Managing Change in Libraries and Information Services
- Budget Management
- Personnel Management
- Collection Development

These workshops should enable the participants to: accomplish the following:

- Implement a new project within a library and deal effectively with any stress brought on by this new project
- Negotiate effectively with a supervisor or group
- Devise new job descriptions
- Implement new marketing strategies for libraries
- Deal effectively with all levels of library personnel
- Analyze costs within a library on a variety of subjects such as book collections, automation, personnel, buildings, etc.

Qualifications and experience in the topic will be given the highest priority in evaluation. You are invited to send a program description for each of the workshops and how they would be conducted.

The successful prospectus will be advertised by the local Soros Foundations in each of the eight countries. It is up to the local library coordinator and the library board to select a topic from the five targeted subjects.

These workshops should be as practical as possible and concentrate on real problems rather than grand theory. Developing professional links with Western European and American librarians and librarianship will be an important benefit of the workshops and should be addressed. The workshops will include prior visits to libraries in the region by workshop leaders so they can understand local con-

c.f.p.'s...

ditions.

The Open Society Institute in consultation with the local Soros Foundations will advertise the workshops, select the delegates (usually 20 per workshop); arrange for translation of workshop materials; pay for accommodation for the trainer(s) and delegates; and provide local expenses for the workshop leaders.

The courses will be conducted in English. Translators will be provided. It is hoped that a significant body of training materials will emerge from these workshops, which could then be further distributed within the region.

Please submit a budget with your proposal which includes the following costs: personnel, development, delivery, materials and overhead. It will be up to the trainers to arrange their air travel (which OSI will cover). OSI will also cover hotel accommodation, per diem (\$25 per day) and in-country travel. Your submission should also consist of a prospectus including the names and credentials of workshop leaders, experience working in CEE and/or FSU and other training experience in non-English speaking countries.

The timeline for the selection process is as follows:

- April 1, 1995: Proposals due
- April 15, 1995: RLP decision on bids given
- June-December, 1995: Workshops given

Please send proposals to:
Open Society Institute
Regional Library Program
Arvacska 6
H-1022 Budapest
HUNGARY
PHONE : 36-1-2122029
FAX : 36-1-2122092

Other conferences

The deadline for submissions to the conferences and workshops below either has passed or will have by the time you receive this.

Making Information Systems Work

IEE Computing and Control Division

Professional Group C10 (Consumer and Domestic Systems)

Lecture on Making Information Systems Work

by Professor Ray J. Paul (Brunel University)

Chairman: Professor Peter J. Thomas (University of the West of England, Bristol)

3 May, 1995, 5.30 pm
Savoy Place, London

The most generous observation which can be made about current information systems developments is that most, if not all, systems disappoint. This disappointment rests mostly with the customer - either user or owner - but also with the analysts and developers of those systems. An example is the sad case of the London Ambulance Service computer system that spectacularly failed and had to be disconnected in its first days of operation. On Monday and Tuesday 26 and 27 October 1992 between ten and twenty people are alleged to have lost their lives because the emergency service failed them.

Clearly in this case the customers were disappointed, but so were the 'owners' of the system: the systems manager is alleged to have said that "the computer system did not fail on that Monday - it was working exactly as it was designed to". It was also suggested that the users of the system - in the control room and the ambulances - were "doing things wrong". That is, they were not doing what the system was designed to expect. Similar examples abound. Taurus, the collapsed system for share dealing in the City of London is a more recent and financially spectacular failure. The question that may be asked is "are there any examples of complete success stories?", and this question often becomes "why don't systems do what people want?".

However, the more important issue, which will be explored in this lecture by Professor Ray J. Paul of Brunel University, is "how can people know what they want" in the development of information systems. Professor Paul will discuss, through examples of the development of information systems, approaches to 'making information systems work'.

CKBS-SIG Intelligent Agents

Intelligent Agents and the next information revolution

A One-Day Meeting of the International

CKBS-SIG

May 9th, 1995
Manchester

It is by now a cliché that the widespread use of distributed information services will radically alter the way in which both organisations and individuals work. There are many indicators of this coming information revolution. The growth of network technology in commercial organisations, the routine use of email within academia, and the astonishing extent of interest in the World Wide Web are three obvious examples. Yet while the enormous potential presented by distributed information services is widely recognised, the software required to fully realise this potential is not yet available. There are several reasons for this, but among the most important is that current software paradigms simply do not lend themselves to developing the kind of applications required. In order to build computer systems that must operate in large, open, distributed, highly heterogeneous environments, we must make use of entirely new software technologies. The concept of an intelligent agent, that can operate autonomously and rationally on behalf of some user in such complex environments, is increasingly promoted as the foundation upon which to construct such a technology. The purpose of this meeting is to bring together researchers and practitioners interested in realising and exploiting this important emerging technology.

The day-long meeting will consist of: an introductory overview of the area and issues; keynote presentations from influential researchers; long presentations describing major applications, projects, and research results; and short presentations describing ongoing work. The emphasis throughout the day will be on informality, discussion, and informed speculation.

If you would like to give a presentation, then email one of the organisers (below) enclosing your full contact details and a short (one paragraph) summary of your intended presentation. Topics of interest include, but are by no means limited to, the following:

Network agents, WWW agents, agent-based information systems, agents in decision support, agents for resource location, distributed information services, software agents/softbots, knowbots, authentication and security issues in cooperative systems, agent communication languages, KQML and KIF, cooperative information retrieval

Coming soon...

and management, shared ontologies, the electronic marketplace, information management and filtering agents, knowledge sharing, interface agents.

Register by simply emailing one of the organisers, enclosing your full contact and affiliation details. All are welcome. No charge will be made for attendance. Please do not turn up without registering.

email: M.Fisher@doc.mmu.ac.uk
tel (+44 1 61) 247 {1531, 1488}
fax (+44 1 61) 247 1483
URL: <http://www.doc.mmu.ac.uk/STAFF/mike/ckbs95.html>

Digital Libraries '95

The Second International Conference on the Theory and Practice of Digital Libraries

June 11-13, 1995
Austin, Texas

As a community, we are just beginning the process of addressing the challenging research problems posed by the advent of the digital library. Addressing these problems requires changes in our research environments—effective solutions require cooperation within widely interdisciplinary groups of scholars and the forging of new links not previously tested. Issues to be addressed include the technological, the representational, the organizational, the social, and the legal. In the founding conference in this series, Digital Libraries '94, we began to develop the common language that permits communication among the diverse communities of interest. We acknowledged strengths and weaknesses of technologies and techniques drawn from traditional areas of study. We recognized that digital libraries must build on the past experiences of traditional libraries but also must not be constrained by the physical limitations of the traditional library.

Email: DL95@bush.cs.tamu.edu
Telephone: (409)-862-3217
FAX: (409)-847-8578
WWW: <http://bush.cs.tamu.edu/dl95/README.html>

ACL SIGDAT Very Large Corpora

ACL SIGDAT 3rd workshop on very large

corpora

June 30, 1995 - following ACL-95
MIT, Cambridge, Massachusetts, USA

The workshop will again offer a general forum for new research in corpus-based and statistical natural language processing. Areas of interest include (but are not limited to): sense disambiguation, part-of-speech tagging, robust parsing, term and name identification, alignment of parallel text, machine translation, lexicography, spelling correction, morphological analysis and anaphora resolution.

This year, the workshop will be organized around the theme of supervised training vs. self-organizing methods

Is annotation worth the effort? Historically, annotated corpora have made a significant contribution. The tagged Brown Corpus, for example, led to important improvements in part-of-speech tagging. But annotated corpora are expensive. Very little annotated data is currently available, especially for languages other than English. Self-organizing methods offer the hope that annotated corpora might not be necessary. Do these methods really work? Do we have to choose between annotated corpora and unannotated corpora? Can we use both?

The workshop will encourage contributions of innovative research along this spectrum. In particular, it will seek work in languages other than English and in applications where appropriately tagged training corpora do not exist. It will also explore what new kinds of corpus annotations (such as discourse structure, co-reference and sense tagging) would be useful to the community, and will encourage papers on their development and use in experimental projects.

The theme will provide an organizing structure to the workshop, and offer a focus for debate. However, we expect and will welcome a diverse set of submissions in all areas of statistical and corpus-based NLP.

Registration:
Ken Church
Room 2B-421
AT&T Bell Laboratories
600 Mountain Ave.
Murray Hill, NJ 07974 USA
e-mail: kwc@research.att.com

David Yarowsky
Dept. of Computer and Info. Science
University of Pennsylvania
200 S. 33rd St.
Philadelphia, PA 19104-6389 USA
email: yarowsky@unagi.cis.upenn.edu

Multimedia DBMSs

1st international workshop on multimedia database management systems
August 28-30, 1995
New York

The rapid advances in many facets of Computer Technology have intensified Research and Development interests in Multi-Media Information Processing (MMIP). Despite the current existence of some forms of Multi-Media Systems, the true realization of MMIP rests on the ability to access and manipulate the multimedia objects. That can only be adequately addressed and achieved via Multi-Media Data Base Management Systems (MMDBMSs). For conventional text-based information, data access and manipulation have advanced considerably; however, for multi-media systems that incorporate continuous media objects, e.g., audio and video, and other media types like images, voice, and graphics, they pose novel problems.

This workshop, which is intended to become an annual event, is a follow-up to the highly successful multi-media workshop held in conjunction with the ACM Annual Multi-Media Conference in San Francisco, California, in October 1994. The workshop will also feature two internationally reputable guest speakers.

The workshop will emphasize on the issues relevant to designing, developing, and utilizing MMDBMSs. Especially multi-media database models, decomposition/organization strategies, storage structures and techniques, system, schema, and functional architectures, query language designs and specifications, retrieval/search techniques/strategies/algorithms, metadata management, query optimization, application/user interfaces, access methods and indexing techniques, distributed multi-media data management, multi-media data synchronization, QOS delivery/presentation issues, performance evaluation, data compression, access security issues, object-oriented paradigms, image processing (storage, retrieval, etc), multi-media network protocols, conventional database compatibility issues, operating systems implications.

For registration contact:

Emileen Butler
CASE Center
Suite 2-212, CST,
Syracuse Univ.
Syracuse, NY 13244.
Phone: (315)443-1062

Calls for Journal Papers

Artificial Intelligence Review

Special issue on the use of AI in information retrieval

Artificial Intelligence and Information Retrieval have long shared a common interest in the representation of knowledge for the purposes of classification and retrieval. AI researchers have found that the IR task provides a practical testbed for the application of knowledge-based techniques, while IR researchers have looked to developments in AI for ways to augment methods based on statistics and information science. The combination of knowledge-based and statistical paradigms is likely to play an ever increasing role in enhancing on-line information search and navigation.

This special issue will highlight some of the current efforts in applying AI techniques, such as case-based reasoning, machine learning, neural nets, natural language processing, knowledge acquisition, and intelligent interfaces, to the broad range of problems encountered in on-line information retrieval, including but not limited to:

- Query formulation and query expansion
- Document/image representation and indexing
- Information discovery and filtering
- Document/query matching algorithms; ranking of search results
- User interfaces and user modelling
- Development and use of thesauri and other search tools
- Managing and accessing distributed information sources

In addition to the call for full-length papers, we request that any researchers doing work in this area submit abstracts and/or pointers to recently published works for the purpose of compiling a comprehensive survey of the current state of the art.

Manuscripts due: Mon May 1, 1995

Acceptance notification: Wed July 5, 1995

Final manuscript due: Mon October 2, 1995

Publication date of issue: January, 1996

Guest editor:
Peter Anick

Peter Anick
Digital Equipment Corp.
OGO1-1/E17
40 Old Bolton Road
Stow, MA. 01775
(508) 496-9173 (voice)
(508) 496-9929 (fax)
E-mail: anick@aiag.enet.dec.com

Standpoints: The Electronic Journal of Information Contexts

The theme of the 1994 ALISE (Association of Library and Information Science Education) conference in Los Angeles, "Intellectual Diversity: Cross Disciplinary Connections and Perspectives," reflects the growing interest library educators have in acknowledging the urgent need for broader and more diversified approaches to library and information studies. Among the topics discussed at the conference were various perspectives, such as multiculturalism and feminist standpoint theory, and alternative research methodologies, such as ethnography. While these perspectives and approaches are not new to the field, the conference amplified voices which were previously only whispers buried in the chatter.

With the publication of *Standpoints: The Electronic Journal of Information Contexts*, we hope to address at least one of the barriers that prevent researchers from generating such scholarship - lack of incentive. Positivist models still dominate the research paradigms of the field. Researchers are reluctant to engage in projects that employ alternative perspectives due to perceived difficulties in obtaining funding and scholarly recognition. Qualitative studies may not be valued by positivist reviewers and therefore be rejected for publication. Tenure committees may not look favorably on less well-established methodologies and may be suspicious of interdisciplinary efforts.

Some of us who attended the 1994 ALISE conference see the need for creating a new forum that will provide an open environment for different voices representing a variety of points of view on and approaches to a wide range of concerns in the library and information field. This refereed electronic journal will be interdisciplinary, rather than disciplinary, and inclusive, rather than exclusive, in nature. It welcomes any sound research that deals with issues related to information contexts. The preference will be given to those studies taking nontraditional perspectives

and/or approaches; to studies asking questions starting from lives of a wider variety of people; and to studies exhibiting interdisciplinarity. It is the sincere hope of the editors that the publication of this journal will add diversity to our field and contribute to its scholarly development.

Standpoints: The Electronic Journal of Information Contexts is a new forum for different voices representing a variety of points of view on approaches to a wide range of concerns in information contexts. Our vision of the journal is an expansive one. We wish to support an open, interdisciplinary, and inclusive environment devoted to the development of theory and publication of research about information and its contexts. Some examples of the relationships we would like to see this journal support include: interpersonal communication in information contexts; gender studies in telecommunications; research in the communication of information about "taboo" subjects; intercultural communication of information; critical theory as applied to information systems design; information culture and counter-cultures; analyses of contemporary technological responses to the "information gap" and building a conceptual framework for the organization of electronic information. We welcome contributions from all standpoints and methodologies.

email: burnet@zodiac.rutgers.edu

The Information Society (a Quarterly Journal)

"The Information Society" journal, published since 1981, is a forum for thoughtful analysis of the impacts, policies, system concepts, methodologies and cultural change related to these trends. It is a refereed journal that publishes scholarly articles, position papers, short communications and book reviews.

"The Information Society" is a multidisciplinary journal whose audience includes policy- and decision-makers and scientists in government, industry and education; managers concerned with the effects of the information revolution on individuals, organizations and society; and scholars with an interest in issues regarding the Information Society.

Among the topics addressed within the journal are:

- changing National Information Infrastructures, especially as they influence cultural expectations and social prac-

tices

- the politics of change in National Information Infrastructures
- the implications of the coming surge in electronic data interchange (EDI) and electronic commerce among businesses globally
- the ability of companies to “outsource” portions of their information processing to different countries around the world, creating transborder data flow issues for the countries involved and increasing the rapidity with which jobs migrate globally
- meanings and implications of political/economic systems that are based on universal access to baseline information services or fees-for-all-services
- options for, and implications of, various forms of “electronic democracy”
- the rise of “virtual communities” of persons worldwide engaging in “many-to-many” communication among their participants, irrespective of borders or corporate structures
- the role of place and space in visions and practice of digital libraries and electronic forums
- cultural changes in relation to cyberspace — both empirical studies and studies of their representation in

popular culture

- the structure of the information industries, including markets, industrial alliances, the character of work, and management-labor relations

Manuscripts should be sent in triplicate (or electronically by Internet) to the editor-in-chief. For manuscript format details, contact the editor or see the inside back cover of an issue of the journal.

email: klings@ics.uci.edu

URL: <http://www.cs.uci.edu/dir/faculty/CORPS/klings>

Product news

Sun automated tape storage library.

Mountain View, Calif. - Feb 21, 1995 - Sun Microsystems Computer Co. (SMCC) Tuesday introduced the SPARCstorage Library Model 8/140 automated tape storage system designed for cost-effective data management at either the enterprise or department level. Because of its storage capacity of 140 gigabytes and the low cost of the media, SMCC's new library offers economical

online and offline storage for unattended backup-and-restore, archival, data acquisition and hierarchical storage management (HSM) applications.

Gigatek minicartridge drives store over 30Gb

La Costa, Calif. - Feb. 21, 1995 - Gigatek Memory Systems Inc. announced today a technology breakthrough for its quarter-inch cartridge (QIC) product line, advancing storage capacities beyond 30 gigabytes (GB) using existing QIC technology. Gigatek's new EC-Family of QIC cartridges operate in today's minicartridge tape drives and drives currently under development.

PageKeeper 2.0 Intelligent search capabilities (hmmmm - Ed.)

Los Gatos, Calif. - Feb. 21, 1995 - Caere today announced the immediate availability of version 2.0 of its Windows-based PageKeeper document management software for the desktop. PageKeeper 2.0 combines intelligent search capabilities, access to multiple sources of information, optical character recognition (OCR) technology, and integrated application support for easy information access, analysis and retrieval.

The Informer Interview - John Lindsay

This time John Lindsay, IRSG secretary, is asked some rather searching questions and gives back some rather strident answers.

What is your position in the BCS IRSG?
Vertical.

What is your job title?
Reader in information systems design.

What projects or ideas are you currently working on or planning to work on in the future?

Information architecture for the BCS. The role of information in the planning and development of Third World economies. Information audit for architects who claim their building projects are either environmentally sound or improve the environment.

What other areas of IR are you interested in?

The whole rather than the parts. Information, when systematised, changes organisation. Information is a doing word, a relation between a subject and an object, not a thing, or a possession of an object

What areas outside of IR are you active or interested in?

You mean there are any?

What is it with you and leather?

Protects my stomach from knives

What was the greatest moment of your career (or what will it be when it finally happens)?

Rather ideologically loaded wot? Actually, getting Gay Switchboard and its information system up and running was probably the most useful thing I've done.

What are your main interests outside work?

Politics and sex.

Why are there so many Marks from Glasgow on the committee of the IRSG?

God wants to be surrounded by saints?

Predict the future of IR - what do you think will happen in the next 20 years?

Hmm - two options - Bosnia or utopia? I think it is at a paradigm shift boundary. If the shift isn't made it'll disintegrate as the wars of the contending parties bring the whole thing tumbling down - much

wider than IR - this involves telecommunications policy, regulatory environments, competition over markets, standards, and within this client server, controlled vocabularies, tcp/ip, Z.39.50 are merely little parts. What happens actually though depends on what we do and how we organise it. Which is why I think arguments about things like the shape of JANET and who owns intellectual property rights, electronic publishing etc. are in fact important.

What future directions can you imagine for the BCS IRSG?

Actually I suspect I think it rather ought to stay the way it is - small, quirky, aimed at getting new generations up and running more quickly than they otherwise would.

What do you think is the most neglected area of IR?

People - users - their understanding of the world, the codes and structures by which they function more effectively.

How would you improve the BCS IRSG?

A few more active young radical people rather than just the old fogey Marks.

Informer Vol 4 Spring 1995

Yes, the back page is blank again I'm afraid.