Interdisciplinary Advances in Adaptive and Intelligent Assistant Systems:

Concepts, Techniques, Applications, and Use

Gunther Kreuzberger, Ilmenau Technical University, Germany

Aran Lunzer, Hokkaido University, Japan

Roland Kaschek, Gymnasium Gerresheim, Germany

Information Science REFERENCE

INFORMATION SCIENCE REFERENCE

Hershey · New York

Director of Editorial Content: Kristin Klinger
Director of Book Publications: Julia Mosemann
Acquisitions Editor: Lindsay Johnston
Development Editor: Joel Gamon
Publishing Assistant: Casey Conapitski
Typesetter: Casey Conapitski
Production Editor: Jamie Snavely
Cover Design: Lisa Tosheff

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue,
Hershey PA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

Interdisciplinary advances in adaptive and intelligent assistant systems : concepts, techniques, applications and use / Gunther Kreuzberger, Aran Lunzer and Roland Kaschek, editors.
p. cm.
Includes bibliographical references and index.
Summary: “This book discusses how collaboration between humans and intelligent systems can potentially improve access to and processing of complex information, contributing to the synthesis, storage and integration of knowledge”--Provided by publisher.

ISBN 978-1-61520-851-7 (hardcover) -- ISBN 978-1-61520-852-4 (ebook) 1. Intelligent control systems. 2. Expert systems (Computer science) I. Kreuzberger, Gunther, 1968- II. Lunzer, Aran, 1965- III. Kaschek, Roland H., 1955-
TJ217.5.I5453 2010
006.3--dc22

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 9
Your Personal, Virtual Librarian

Alexander Krumpholz
CSIRO ICT Centre & Australian National University, Australia

David Hawking
Funnelback, Australia

Tom Gedeon
Australian National University, Australia

ABSTRACT

Searching scientific literature is a common and critical activity for research scientists, students, and professionals such as medical clinicians. These search tasks can be time consuming and repetitive, but literature search and management tools are already making the job much easier. This chapter analyses the literature retrieval process, reviews some currently available tools and elaborates on potential future support for the knowledge worker by an intelligent automated assistant. A special focus of this chapter is the automatic retrieval of medical literature and the exploration of the answer space.

INTRODUCTION

Before the Internet revolution, researchers physically went to the library to find the publications they needed for their studies. They spent hours trying to fish the right index cards out of little drawers in numerous rows of cabinets to identify books and papers that might guide their work. Librarians then fetched the needed materials and passed them out to the reader. The turnaround time could be hours, days, or

DOI: 10.4018/978-1-61520-851-7.ch009

even weeks if the publication had to be ordered from overseas. Later, librarians got access to computers and translated readers' needs into queries to retrieve the meta-information about the requested literature. Beyond their retrieval tasks, experienced librarians also helped students to identify popular books or drew their scientists' attention to new publications in their field.

Today, almost every researcher has a computer on their desk and is able to send queries to search engines all over the world—including literature retrieval engines—and retrieve references, even full papers, within minutes. However, do the new systems provide you with all the benefits you got from your librarian? Are they the perfect solution for your literature retrieval needs, or is there much room for improvement?

We will describe existing systems currently available commercially or free, and present current research efforts to improve literature retrieval. The need for literature retrieval depends on the role and the field of the researcher as well as the task they wish to achieve. This chapter focuses on literature retrieval for computer scientists, and on medical literature retrieval for clinicians or medical researchers.

Information Needs

In what circumstances do scientists search the literature?

We will discuss six typical scenarios:

- Finding a specific publication
- Finding related publications
- Staying up to date in the field
- Entering a new field
- Special domain: Medical research.

Finding a Specific Publication

Almost all scientific publications provide a list of references to related material, similar to the way that URLs are used to refer to Web pages on the Internet. While the hyperlink mechanism embedded in the World Wide Web allows us to follow a URL by clicking on the anchor text, bibliographic references once had to be resolved by visiting the library, accessing the referenced journal issue and finding the cited article within. A scientist may need to access cited references to fully understand the paper they are reading or to learn about related work. References to specific publications may also be made by reviewers when the scientist submits his own work for publication. Furthermore, specific references are often made in casual conversations: "I really liked your presentation, but have you come across the paper by Smith and Jones from the University of...?".

Nowadays, various Internet portals can be used to locate abstracts quickly, and even to download the full text or printable version of the publication. Tools to manage retrieved bibliographical records for automatic use by text processing systems are essential and used by students or scientists.

Finding Related Publications

Traditionally, related publications have been found by searching for publications in the same classification as the primary article, or by following journals or conference proceedings in relevant areas. Through its own reference list, each new article might lead to other specific publications and, possibly, other venues to be monitored. Once research groups or individuals working on similar topics have been identified, their progress can be followed and collaboration becomes a possibility. Today, more automated means of identifying related publications have been developed and are available on some Web portals.

Staying up to Date in the Field

Once, staying on top of the field involved skimming through key journals and conference proceedings as they were published. Through conferences, discussions and exchanges with other researchers, scientists could be alerted to new projects, fertilising new ideas and collaborations. In addition to these traditional ways, modern Web-based technologies have provided notification services to provide keyword-based alerts to the publication of relevant new literature. Such services have been implemented economically and are often offered free of charge.

Entering a New Field

Entering a new field involves getting introduced to new terminology, new concepts and relationships between them. Scientific papers are usually too narrow and detailed to quickly obtain a bird's-eye view of a subject, although review articles are available for many topics. Some work has been done to automatically generate such reviews. An alternative approach is to use books and course material to provide good introductions, but there can be significant delays between the journal publication of cutting-edge research and its appearance in longer texts.

Special Domain Medical Research

Retrieval of medical literature is an activity that stands to have a profound effect on our health and wellbeing. Awareness of current medical knowledge can help medical practitioners to make better decisions and so lead to better health outcomes, or avoid adverse events. It is no surprise that vast resources have been invested in organising medical knowledge through ontologies such as Snomed (Stearns et al., 2001) and thesauri like MeSH ("Medical Subject Headings", n.d.). Since medical publications tend to be rich in jargon, advanced knowledge extraction techniques achieve good results. This has enabled the development of specialized visualization prototypes (Plake et al., 2006) and even expert systems.

With the development of *evidence-based medicine*—a term formally defined by the Evidence-Based Medicine Working Group in the 1990s (Evidence-Based Medicine Working Group et al., 1992)(Claridge & Fabian, 2005)—the need to find scientific literature in the medical field and to judge its value has become even more important. Observations reported in case studies trigger randomized controlled trials. Systematic reviews take all medical literature relevant to a given research question, summarize them, weigh the level of evidence they represent and provide a critical discussion. Clinical Practice Guidelines are step-by-step descriptions on how to diagnose and treat a patient with certain symptoms and co-morbidities (co-occurrences of additional diseases) and are generally based on systematic reviews. Due to the needs of evidence-based medicine and the specialization in the medical field, clinical literature retrieval is well developed and rich in features.

The prevalence of computer support in clinical settings has led to electronic health records increasingly being used to store clinical data such as symptoms, diagnoses, test results and treatments for each patient. While privacy issues still cause many concerns about how to share and exploit information, local solutions are implemented to help clinicians to make better-grounded decisions faster. Controlled sharing of electronic patient records between hospitals, clinics and practitioners has the potential to improve the quality of patient treatment and to reduce costs in the health care sector.

Interestingly, Rockliff et al. (2005) show that the services of real librarians remain of high value. They describe 'Chasing the Sun', a project between the South Australian Health Services Libraries' Consortium (SAHSLC) in Australia and the South West Information for Clinical Effectiveness (SWICE) network in the United Kingdom. This out-of-hours emergency virtual reference service allows night shift clinicians in either country to get help related to patient care using a skilled librarian as an intermediary.

Retrieval Process

Ad hoc Web search is a process well known to Internet users. Starting with a favourite search engine, a user enters a query containing some search terms and gets a list of matching documents—mostly Web pages—as a result. Most users only skim the first few results and may alter the query and search again without even leaving the first result page. These steps are repeated until the searcher is satisfied or gives up. The main types of change to the query include modifying the query terms (e.g., when better expressions have been seen in the first results), and adding/removing terms to reduce/increase the number of results. While interaction with literature retrieval systems is quite similar to Web retrieval, specific metadata is often exploited to allow additional means of interaction such as links to other publications by the same author.

Common steps in a search session are:

- Identification of the information need
- Selection of a corpus to query
- Selection of a search tool
- Formulation of a query
- Retrieving the search result
- Identification of the relevant items in the result set
- Identification of the relevant section of each result

Each of those steps has to be conducted manually by the searcher.

Identification of the Information Need

Initially, the user has to become aware of the need to find publications. Depending on the user's current task, it can be one of the needs described above. Often the user needs to find related publications or is trying to gain understanding of a certain topic.

Selection of a Corpus to Query

Then the user has to select the corpus to use. The corpus in this context is the document collection to be indexed by a retrieval system. Corpora differ in their comprehensiveness, currency and correctness. Some collections have full text versions of the documents available; others only store the bibliographic metadata or abstracts. The user's choice of corpus may depend upon the age of the publications they are seeking.

Selection of a Search Tool

When a collection is indexed by different parties, the user must also decide which search system to use. Freely available corpora are used by researchers to explore new ways of pre-processing the data in order to provide additional support to the searcher. Some prototypes are available free of charge for everyone on the Internet to explore, evaluate and send feedback. Such prototypes however are often unknown to researchers who are focusing on their work rather than tracking such developments.

Formulation of a Query

Finding the ideal search terms is often not a trivial task. Information retrieval systems have to match the query to the indexed corpus to find the documents best matching the users query. This matching makes it necessary for the user to know the jargon and writing style used in the corpus. The information retrieval community has developed mechanisms like *stop word lists* to remove words not relevant for the matching process, *stemming* to include documents containing identical words with different inflections, and *query expansion* to add synonyms of query terms to the query. Searchers must try to find a query that achieves a good balance between *precision* and *recall*, i.e., a query that retrieves everything of importance yet nothing of irrelevance. Recall expresses the proportion of relevant documents in the corpus that have been retrieved. Precision indicates the proportion of the retrieved documents that are relevant. Usually a query that achieves high precision will suffer from poor recall and vice versa. It's important to recognize that the user's need might not be well expressed in the user's query.

But why should we aim for the ideal search terms in times when users can make use of modern interfaces that allow them to iteratively reformulate queries interactively and submit hundreds of queries in a single session? On one hand because it is a frustrating task to miss out on the bulk of results just because the correct term has not been used. A PubMed search for "*high blood pressure*" for example currently returns 8,163 results if it is directly used as a query. PubMed however automatically expands queries before executing them to include the appropriate Medical Subject Headings (MeSH). Once the query is expanded to "*hypertension*" *[MeSH Terms] OR* "*hypertension*" *[All Fields] OR (*"*high*" *[All Fields] AND* "*blood*" *[All Fields] AND* "*pressure*" *[All Fields]) OR* "*high blood pressure*" *[All Fields]*, PubMed is able to allocate 376,007 matches. On the other hand, this chapter introduces an intelligent assistant that should at least support the user in the query formulation process by automatically generating an initial query, which can be modified by the user.

Retrieving the Search Result

The result of a Web search is often a ranked list of matching web pages or documents. In literature retrieval, the result set is usually a list of publications, represented by bibliographical records containing metadata such as the article title, the journal title, authors, publisher and the publication date.

Identifying the Relevant Items in the Result Set

Results returned by the system generally match the user's query in some way, but it is very unusual for all of them to be actually useful. The user must skim the results to find the ones most likely to help them with their task. Polysemies—words with multiple meanings—are a source of results matching the query but not the information need. The query term *apple* may return documents related to fruits as well as those related to computing equipment. Established methods to help the process of disambiguation include clustering and relevance feedback. Web search engines have been developed that build clusters according to the content of the result documents and present them in a way that allows the searcher to narrow the results to a particular cluster. Relevance feedback systems trace the user's interaction with the result set to identify results that are of interest to the user and from them to derive and submit a more complex query more closely related to the user's actual information need.

Identification of the Relevant Section of Each Result

Finally, the searcher has to find the part of each result document matching the query to find the section of interest to decide if the whole document has to be read and potentially cited. The size of publications available electronically for indexing is increasing and the work on information retrieval systems that deliver only parts of documents has commenced long ago (Salton et al., 1993) (Wilkinson, 1994). There are also tools for highlighting relevant sections, or individual query term occurrences, within long documents.

Techniques for Matching and Ranking Documents

Web search engines typically identify a set of candidate documents using a very narrow matching criterion such as "contains all of the query words in either singular or plural forms". They then rank the candidate set by combining degree-of-query-match with a number of query-independent factors such as popularity (estimated from user click frequencies), authoritativeness (estimated from patterns of hyperlinking) and inverse spam score.

The literature retrieval process is analogous but differs in many important ways. Popularity and authoritativeness may be estimated from analogous indicators such as number of downloads and time-weighted citation frequency. Note that for certain research tasks all documents satisfying the inclusion criteria must be considered, regardless of popularity or authority. (Spam is hopefully never an issue!) The InexBib system (Krumpholz & Hawking, 2006) shows that the anchor text exploited by Web search engines also has an analogue in literature retrieval and can be used to extend recall.

Choosing a candidate set in literature search is much more of a challenge, because of the diversity of language typically used to discuss a single scientific concept or issue. This might sound naïve given that each field has well defined terminology, but overlapping fields do not necessarily use the same terms for the same concept and sometimes the correct terminology is just not being used. At the time of writing this chapter, the PubMed search "*high blood pressure*" *[All Fields] NOT hypertension[All Fields]* returns 2125 results, and slight variations like "*higher blood pressure*" *[All Fields] NOT hypertension[All Fields]* or "*increased blood pressure*" *[All Fields] NOT hypertension[All Fields]* yield result sets of hundreds of documents.

On the Web, when a layperson submits a simple query like "bird flu" they can confidently expect to find a set of pages giving good advice and information on that subject. However, a great deal of scientific literature relevant to that topic will not match those terms. The query must be expanded to include obvious terms such as influenza, H5N1 and many highly specialized terms relating to viruses, proteins, and the human immune system. Which query is ideal is highly context-dependent: What is the searcher's field of interest? Why are they searching on this particular occasion? What do they already know?

By convention, scientific articles are structured into sections such as Abstract, Related Work, Method, Results and Conclusions. They typically include other structural elements such as figures, tables and equations. There has been some recent research interest in whether document structure can be exploited in order to improve the retrieval performance, the matching process and the units of retrieval. Query matches in certain elements of a document (such as its title or conclusions) may be weighted more heavily than those in the general text of the document.

The rest of this section focuses on how to exploit structure and search context.

Structured Document Retrieval

The Initiative for the Evaluation of XML retrieval (INEX) has been active since 2002 as a platform for XML retrieval experiments (Fuhr et al., 2002). Its participants normally build an information retrieval collection and use the relevance values attached to each result for each given query to compare their search engines and improve performance (and derive better evaluation metrics). In an annual process, the participants download the corpus and suggest meaningful topics, which are collected by the organisers. After the elimination of redundant and meaningless queries, the remaining ones are distributed to the participants, who use their search engine to produce a list of potential hits for each topic. Those hits are collected and judged by the community. Once the relevance of each result for each query has been assessed, the collection can be used to calculate a set of quality measures characterizing a specific result set for a given query. Those metrics are used to compare different search engines or different tuning parameters on a given search engine. Results are not full documents as in traditional document retrieval, but parts of documents specified with XPath (Berglund et al., 2002) expressions, allowing specification of individual XML elements or regions within XML documents.

Scientific publications are written once, but may be read hundreds or thousands of times. The structure of scientific publications is defined to support the reader's need to access relevant literature. The title is used to get a first glimpse of relevance, the abstract is read for a better understanding and the conclusions are often consumed next, while figures or images can quickly give an idea of what the paper is reporting. Based on the information gathered, the reader assesses the likely value of reading the whole paper thoroughly. The importance of images in communicating the essence of a scientific paper has also been identified by (Xu et al., 2008), who show their PubMed results as thumbnail images of the containing graphs. This allows searchers to quickly identify publications relevant to their research in some cases and to identify previously seen papers in others.

Scientific articles in the medical field have a much higher degree of structure than publications in other research areas like computer science. The abstract is usually already partitioned, clearly describing the type of study undertaken, the demographics of the subjects, the results and so on. The Instructions for Authors by the Journal of the American Medical Association for example specify the requirements for different paper types in fine-grained detail. (http://jama.ama-assn.org/misc/ifora.dtl):

Abstracts for Reports of Original Data: Reports of original data should include an abstract of no more than 300 words using the following headings: Context, Objective, Design, Setting, Patients (or Participants), Interventions (include only if there are any), Main Outcome Measure(s), Results, and Conclusions. For brevity, parts of the abstract may be written as phrases rather than complete sentences.

This paragraph is followed by a description for each of the defined sections.

But is this enough? An intelligent system should be able to rely on concepts and their relations described in a publication without the need of interpretation and the inherent risk of misinterpretation by natural language processing (NLP) systems, as they try to map natural language into knowledge representation data structures. Since a researcher creates only a few publications per year, but is supposed to consume the essence of thousands of them at the same time, it only makes sense to shift the effort away from the reader to the writer.

One step in this direction is the creation of metadata by numerous indexers at the National Library of Medicine in the USA. Every publication in PubMed is tagged with a set of relevant medical subject heading (MeSH) terms, allowing a searcher to compose their query from a controlled vocabulary.

Context

It seems obvious that a literature retrieval system should be able to perform better if it can exploit the contextual information in addition to the explicitly specified query. Indeed there are a number of studies supporting this expectation. For example (Teevan et al., 2005) show that Web search results can be improved if an initial raw result set is reranked locally using information derived from the user's previous interactions and the documents they have on their own computer. (Kelly & Teevan, 2003) survey the use of human behaviours in deriving implicit measures for use in ranking or re-ranking.

Contextualized methods have the potential to assist with disambiguating ambiguous terms and with generating more effective queries, however there are many open questions about whether automatically derived profiles are best applied at the user's computer, where full context is available without privacy concerns, or at a remote retrieval system, where there is unlimited potential to affect matching and ranking.

EXISTING LITERATURE RETRIEVAL TOOLS

This section introduces some illustrative literature retrieval tools and maps them to the previously described researchers' needs.

Finding a Specific Publication

Databases

With the availability of computers and the Internet, publishers and scientific associations started to offer online access to their full text articles for their members through online portals.

The Association for Computing Machinery (ACM) (http://www.acm.org/), as an example, built the ACM Digital Library and offers access through the ACM Portal (http://portal.acm.org/), while the Institute of Electrical and Electronics Engineers (IEEE) (http://www.ieee.org/) created the IEEE Xplore digital library providing a search interface called IEEE Xplore (http://ieeexplore.ieee.org/).

Nowadays, most publishers of scientific material allow searching their articles, eBooks and so on through their portals like Springer's (http://www.springer.com/) SpringerLink (http://www.springerlink.com/) and Elsevier's (http://www.elsevier.com/) ScienceDirect (http://www.sciencedirect.com/)

However, those portals are usually limited to their own publications and allow only members to get access to the full text or PDF version of the publications or sell each article individually. The advantage of publishers' portals is in the high quality of the meta-data attached.

Manually Compiled Collections

The DBLP Computer Science Bibliography (Ley, 1997) (Ley, 2002) (http://www.informatik.uni-trier.de/~ley/db/) based at Universität Trier is a database of bibliographic references. Development on DBLP commenced in 1993 as a collection of computer science related publications. The high quality of manually maintained bibliographical references (Ley & Reuther, 2006) allowed the development of tools implementing bibliometrics developed long ago by (Garfield, 1963) and (Small, 1973). Despite the development of far larger indices, DBLP is still used frequently.

Other researchers improved the DBLP database by adding an enhanced search component called CompleteSearch DBLP (http://dblp.mpi-inf.mpg.de/dblp-mirror/index.php) (Bast & Weber, 2007) as well as faceted search capabilities (Tunkelang, 2009) via an interface named FacetedDBLP (Diederich et al., 2007) (http://dblp.l3s.de/dblp++.php).

Automatic Indices

Citeseer (http://citeseer.ist.psu.edu/) was developed by NEC in 1997 (Lawrence et al., 1999) and moved to Pennsylvania State University in 2003. It started as an attempt to bring the bibliographic information previously published distributed on websites of publishers, research organizations and individual authors into one searchable index and to automate indexing citations and citation linking. Citeseer also acts as a research platform and the improved system, CiteseerX (http://citeseerx.ist.psu.edu/), incorporates additional features like citation statistics, related document identification and query-sensitive summaries. CiteseerX now indexes over 1.4 million documents and over 27 million citations.

ScientificCommons (http://www.scientificcommons.org/) is an index built by the Institute for Media and Communications Management at the University of St.Gallen in Switzerland. (Kirchhoff et al., 2008). According to their website they indexed 13 million publications by January 2007.

Google Scholar (http://scholar.google.com) is Google's take on literature retrieval. Compared to other systems, Google started relatively late to offer a specialized portal for scientific publications, but has grown to probably the largest index. While numbers are not openly published on the Google Scholar website, simple searches claim hundreds of millions results. Google approaches publishers to streamline access to all their publication material. Google also analyses citations to build citation graphs and allow searching for publications that link to the current one.

Since the metadata is extracted automatically, the metadata provided by such systems tends to contain more errors.

Finding Related Publications

Finding related publications is based on a given context. This context can be based on other publications, the current work of the author, the author's long-term research interests and many other features.

Some bibliographic search systems like CiteseerX shows publications related to a given one. Such relatedness can for example be identified by similarity of the terms contained in both texts, by the number of shared references cited in both papers or the number of publications citing both texts.

The Mac OS X application Papers helps the researcher to build a library of documents, by supporting queries to most repositories and the easy incorporation of search results including the meta data into the user's local collection. In cases where a URL to a PDF-version of a publication is provided, the user can double-click the entry and the file will be automatically downloaded and stored with the metadata. The paper can then be printed or read and annotated on screen within the same application.

In terms of finding related publications, the application creates an editable list containing the authors of the publications of the collection, and the author view as shown in Figure 1 can be configured to automatically check for recent publications when an author is selected. This helps to find follow-up publications and become aware of the author's recent work. In an analogous way, recent articles for known journals are retrieved on the selection of a journal.

Staying up to Date in the Field

Traditionally, researchers tend to stay on top of their field via subscriptions to newsletters and journals, by visiting symposiums and conferences, and by exchanging references with colleagues. The latter seems to be used in increasing rates since

Web 2.0 based tools to share bookmarks and bibliographical metadata collections are more commonly available. An additional feature of modern bibliographic search sites is the availability of email or RSS notification.

Notification Services

The search interface of PubMed as shown in Figure 2 allows the user to specify RSS feed as a result output option. This allows researchers to specify their information need in a very precise form. The RSS feeds can be added to so-called feed aggregators that collect the incoming RSS items and present them for example in a list similar to an email inbox. (http://www.nlm.nih.gov/pubs/techbull/mj05/mj05_rss.html)

Sente (http://www.thirdstreetsoftware.com/) is a Mac OS X application that allows the user to manage references and their PDF files. Additionally, the user can define hierarchical collections by defining queries. The queries associated with those collections will be executed in the background and new publications will be indicated similar to incoming mail.

BioMed Central (BMC) Bioinformatics allows researchers to specify their research interests and send email notifications once new publications match the searcher's profile as illustrated in Figure 3. The service covers the broader interest, but is not as selective as PubMed's RSS implementation. (http://www.biomedcentral.com/)

Figure 1. Screenshot showing Paper app's author view (©2010 mekentosj.com Used with permission)

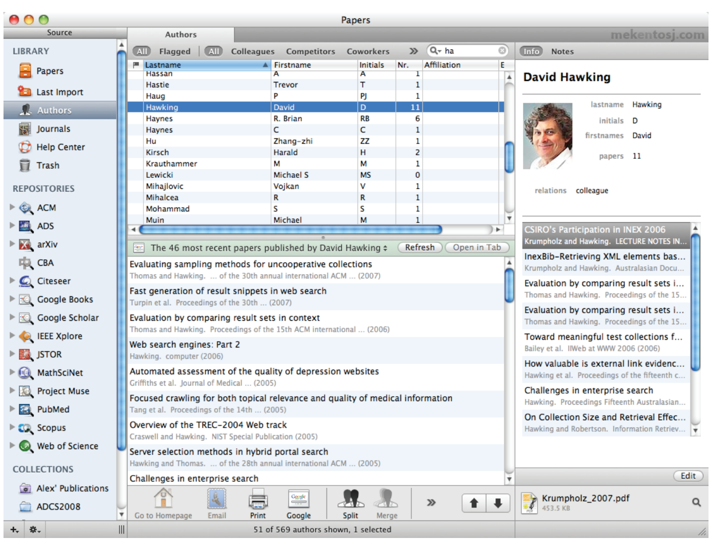
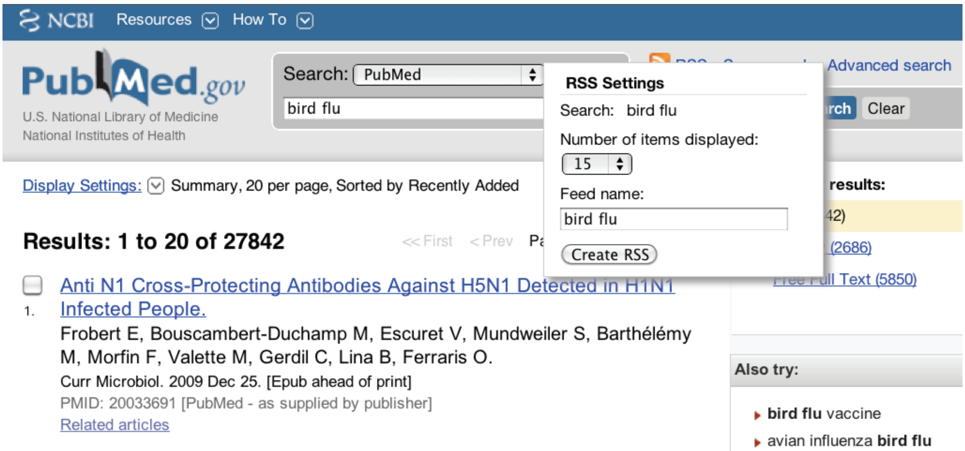


Figure 2. Screenshot showing PubMed's RSS Feed creation (Image courtesy of the U.S. National Library of Medicine)



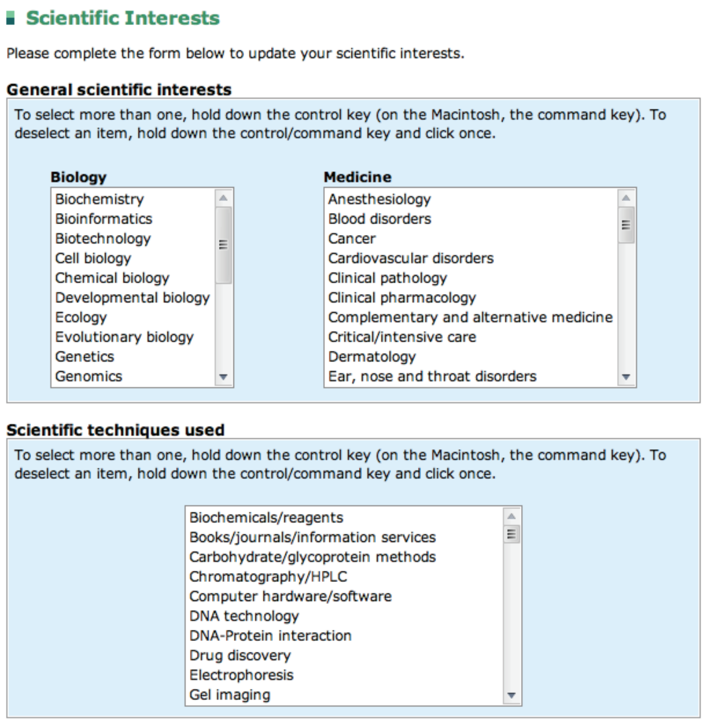
Social Networks

The rise of the Web 2.0 technologies enabled the development of Web applications that were as feature rich as binary applications for desktop computers. Their server-based nature also allowed those applications to share data between users and user groups. One class of such social networks is based on the idea of sharing bibliographic references and web links.

The publisher Elsevier (http://www.elsevier.com/) offers the free service called 2collab (http://www.2collab.com) that allows researchers and students to manage and share bookmarks and references to scientific literature. Similar services are provided by Bibsonomy (http://www.bibsonomy.org/) and CiteULike (http://www.citeulike.org/).

Such bibliography-related social networking sites allow researchers to store their bibliographical references on the server, tag them with key words and upload a PDF version. Based on the information the community entered, the collection of references can be browsed by tag, 'hot papers' are indicated and researchers with similar research interests are identified based on the overlap in both users' bibliographies. The references of others can be viewed, as well as who else has a certain reference in their collection. CiteULike even allows notification on updates of other user's collections via RSS feeds. Other users of the systems can usually be contacted or linked in as colleagues to allow additional collaboration features. The bibliographies can be exported into various formats like BibTeX or RIS.

Figure 3. Screenshot showing BioMed Central's email notification service setup (© BioMed Central, 2010. Used with permission)



Entering a New Field

Someone who wants to enter a new science area needs to understand the field's terminology, concepts and their relationships and which questions have been solved or are still open problems. Lecturers at universities often spend vast amounts of time preparing course material or even books that are specifically designed to introduce students to new areas.

Researchers who have to do a thorough literature review sometimes summarize the current state of literature in surveys. Such surveys cover the publications of selected topics or concepts well and are a valuable source of information for students and other researchers. Journals like the ACM computing surveys publishing nothing but such surveys.

Due to the relatively high quality of community-authored websites like Wikipedia (http://www.wikipedia.org/) and knol (http://knol.google.com/) they are also becoming a viable entry point for getting an overview of a field or concept. While they cannot be taken as source of scientific knowledge like reviews, they are often well written and provide a quick overview including links to more specific or generic concepts and scientific publications.

Many of the tools discussed in other sections are helpful for researchers entering a new field as well, such as the extraction of concepts and their relationships (Plake et al., 2006) described in section *Knowledge representation* or the automatic generation of surveys (Mohammad et al., 2009) (see section *Summaries*).

Special Domain: Medical Research

PubMed (http://www.pubmed.gov/) is an index of Medline provided by the United States' National Library of Medicine (NLM) and the National Institutes of Health. With about 19 million citations, it is the largest corpus of medical literature and therefore widely used by researchers.

PubMed already supports the user with rigorous query expansion facilities. In order to make use of PubMed's MeSH term tagging, entered query terms are mapped into MeSH term combination, sometimes resulting in a complex Boolean query. As an example if the user enters the search terms *colorectal cancer* the query will be rewritten by PubMed to: "*colorectal neoplasms*" *[MeSH Terms] OR (*"*colorectal*" *[All Fields] AND* "*neoplasms*" *[All Fields]) OR* "*colorectal neoplasms*" *[All Fields] OR (*"*colorectal*" *[All Fields] AND* "*cancer*" *[All Fields]) OR* "*colorectal cancer*" *[All Fields]*

Since PubMed also provides an API to allow their search component to be used in complex scenarios, a number of research prototypes incorporate PubMed's search features to create new retrieval tools. This section describes tools that are based on PubMed such as AliBaba, a knowledge extraction and visualization tool, and GoPubMed, a service providing a comprehensive interface to semantically navigate through the PubMed results to a given query.

Knowledge Representation

Ali Baba is a research prototype built by the Knowledge Management in Bioinformatics research group at the Humboldt University in Berlin (Plake et al., 2006). It performs a PubMed query, analyses the results and visually presents an interactive graph showing the semantic relationship between biological concepts, which they call entities, contained in the result set. One of the problems with PubMed and similar resources is the plethora of documents they contain and the vast amount of matches they produce to common query terms. AliBaba helps the user to visualize the concepts described as well as relationships between them in an interactive graph. Extracting concepts mentioned together within one sentence, for example, identifies relationships between such concepts. Selected GUI components link back to the publications from which that knowledge was gained. This helps the searcher to obtain an overview quickly.

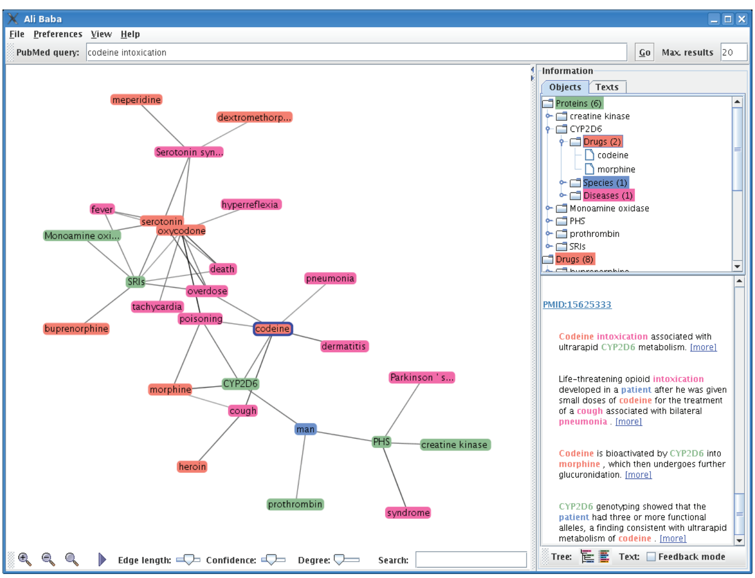
AliBaba performs entity as well as relationship extraction by learning language patterns and deriving consensus patterns representing clusters of identified initial patterns for performance reasons. The entities extracted belong to classes like cells, diseases, drugs, proteins, species and tissues. The techniques used in AliBaba have been described in various papers including (Plake et al., 2006) and (Palaga et al., 2009).

The screenshot in Figure 4 is an example from the AliBaba Website (Knowledge Management in Bioinformatics Group, Department of Computer Science, Humbold-Universität zu Berlin). It is described there as follows:

A Patient with cough becomes unresponsive after normal dosage of codeine—what is going on? The query entered in Ali Baba was "codeine intoxication".

Ali Baba shows the relationship between codeine (marked in the graph with blue frame), cough, morphine, and poisioning. Poisioning is also connected to morphine and CYP2D6. The solution thus is that codeine is bioactivated by CYP2D6 into morphine, certain patients show an ultrarapid form of this metabolism, which leads to a life-threatening intoxication (see (Gasche et al., 2004)). The connection codeine->CYP2D6->morphine is directly visible in Ali Baba. ("AliBaba Screenshots", n.d.).

Figure 4. Screenshot showing AliBaba's user interface (© 2006-2010 Ulf Leser. Used with permission)



Faceted Interfaces

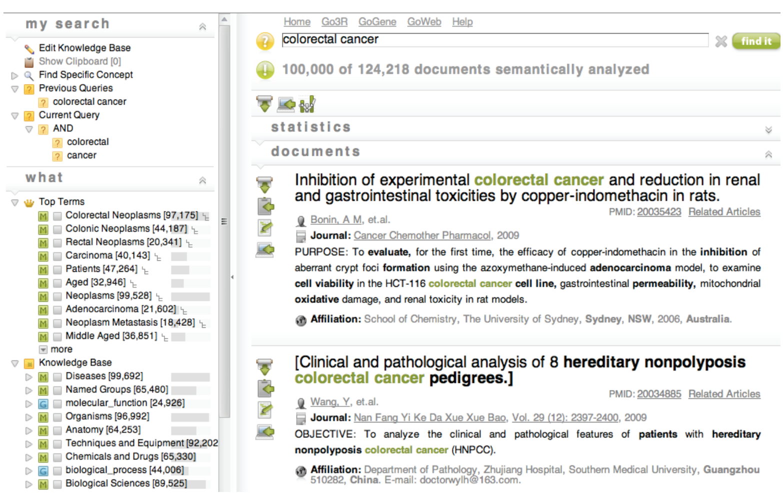
The German company Transinsight provides another interface to PubMed: GoPubMed (Delfs et al., 2004) as illustrated in Figure 5. It shows matching sentences of the abstract, with search terms highlighted, and provides ontology-based ways to navigate through the search results, concepts discovered within and metadata, such as authors, publishers and year of publications. It shows statistics for the metadata as well as timelines indicating when publications including the search terms have been published. These features of GoPubMed help searchers to explore and understand new fields.

Additional features such as ontology building tools are provided by the commercial extension, called GoPubMed Pro. (http://www.gopubmed.org/)

Novo|seek's interface (http://www.novoseek.org/) is visually similar to GoPubMed, allows users to filter by medical concepts and bibliographic properties, including authors and journals (see (Allende, 2009)). In addition to the PubMed corpus, it also indexes full text articles and around 500,000 research grants from Canada and the US provided by SciSight (http://www.scisight.com/).

The point of faceted interfaces is to use hierarchies of facets based on properties of the collection's structure or artificially created ones to provide the user means to navigate through the collection. Often the interface is enhanced with the estimated number of results to give the user advance information about the potential result set of each facet.

Figure 5. Screenshot showing GoPubMed's faceted user interface (© 2005-2010 Transinsigh, Germany. Used with permission)



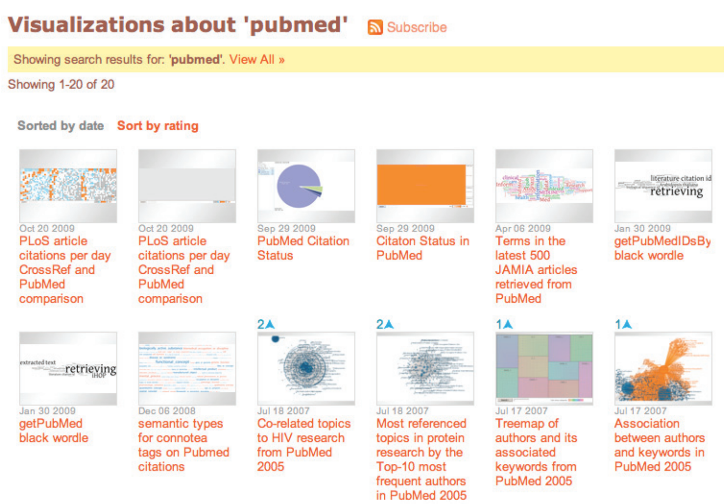
Result Visualisation

ManyEyes (http://manyeyes.alphaworks.ibm.com/manyeyes) is a site by IBM alphaWorks that allows Internet users to create, explore and discuss interactive visualisations of their datasets (Viegas et al., 2007). Users pick or upload a dataset, select a predefined visualisation type and configure it according to their visualisation ideas. Various PubMed related datasets are already available on the server and 20 visualisations about PubMed have been created and discussed by the community. Figure 6 shows an overview of the first ten visualisations matching the search term *pubmed*.

The delivery of the search results in a graphical alternative to a ranked list might take some burden off the user. Instead of parsing each result, understanding its key points and building a mental image of the answer space, the user sees a graphical representation of an aspect of the matching documents and their relations. This can potentially be of great help in getting an overview of the field.

Other work on the visualisation of PubMed results can be seen on the website of the graph visualisation software AiSee (http://www.aisee.com/graph_of_the_month/pubmed.htm) by the German company AbsInt Angewandte Informatik GmbH. It shows visualisations based on data processed using tools like botXminer (http://www.aisee.com/graph_of_the_month/botxm.htm), which has been described in (Mudunuri et al., 2006) or the publication network graph utility PubNet (http://www.aisee.com/graph_of_the_month/pubmed.htm) by Douglas et al. (2005).

Figure 6. Many Eyes shows a variety of example graphs representing PubMed search results (© 2010, IBM Corporation. Used with permission)



Tailored Delivery

Tailoring the delivery is the process of adapting a search engine's output specifically for the current user or user group. In order to allow for a system to deliver user-specific output, it has to represent preferences of the user. Such information is stored in a user model and can be predefined to match the needs of known classes of user like clinical practitioners versus nurses, or learned for specific users.

MiSearch (http://misearch.ncibi.org/) offers adaptive PubMed search (States et al., 2009). Using relevance feedback techniques, the system identifies preferred MeSH terms, substance names and author names, and uses this information to rank result sets of future queries according to the searcher's interests. Relevance feedback has been explored in the information retrieval field since the early 1970s (Rocchio, 1971).

AN INTELLIGENT LITERATURE RETRIEVAL SYSTEM

The section *Retrieval process* described common steps in a search session. Using an intelligent system of the sort we envisage, the user experience will not feel like *searching*. Rather, all the steps in the retrieval process, such as query formulation, retrieval, and result delivery, will integrate seamlessly in a behind-the-scenes workflow whose behaviour can be modified by the user either implicitly through monitoring interactions or explicitly. In order to achieve this, an intelligent literature retrieval assistant should have a user model, task model and knowledge structure model in order to effectively support the searcher.

Some aspects of an intelligent assistant for literature retrieval will be described here. The user should be supported in

- Query generation
- Result comprehension

Query Generation

The first important component of the literature retrieval assistant is the automation of the retrieval steps preceding the search engine interaction, namely the identification of the information need, the selection of a corpus to query and of a search tool, and the formulation of a query.

The system needs to be built into the tools used by the researcher in the phase where literature retrieval is necessary. This can be a program, such as a word processor in the case of an author, or the electronic patient record, in the case of clinical decision support. As part of the user model, the assistant should know about the user's education, for example if the information is needed by an undergraduate student, a post-doc or a professor with years of experience in the field.

In the medical field, the user's current role or job position could be used to decide which type of publication is likely to be most useful: for example, whether a randomized clinical trial, a clinical guideline, or even a consumer information sheet would be more helpful.

Utilizing the user's previous publications as part of the search context, the intelligent search assistant can help to disambiguate query terms in favour of the user's research interests.

Three main points according to (Budzik et al., 2001) are

- Relevance of active goals
- Word-sense ambiguity
- Audience appropriateness.

One way to derive the user's context is to attach the literature retrieval assistant to applications used by the searcher. Some effort has been made to explore the context of the user and offer matching search results, including working prototypes.

(Budzik et al., 2001) designed information management assistants and implemented a system called Watson that ran on a user's computer, observing the user's current context and task. This information was used to automatically retrieve documents for immediate display to the user. Alternatively graphical user interface components to trigger a search were offered to the user.

The information management assistant extracted the information need, based on the user's action and task models, as well as the content representation, based on the currently manipulated document and content models. Using information source descriptors, queries were sent to external information sources and the results were collected and presented to the user.

Watson incorporated adapters to access documents from applications like Microsoft Word and Web browsers and could be configured to retrieve similar or related documents. Based on text analysis, information entities such as addresses could be extracted and augmented by a link to a map service showing the location of the address. When the user created an empty image and a caption, the system used the caption text as query terms and displayed matching images from the Web. Since the rest of the document formed the context of that query, the results generally matched well and did not need to be clustered for disambiguation.

Watson was available as a free desktop tool for the Windows platform, but has been developed into a server-based platform for Web publishers and is now commercialized as Perfect Market. (http://www.perfectmarket.com/) by Idea Labs (http://www.idealab.com/).

A less generic type of application onto which to attach a literature retrieval assistant would be authoring software. Just as text editors evolved into word processors, special purpose authoring software is being developed for screenplay writers, novel writers and scientific writers.

One such development is the digital scholar's workbench (Barnes, 2007). It helps students and scientists to focus on writing their publications, while taking care of the publication workflow. The workbench supports the conversion of produced word processor documents into DocBook XML (ref) for storage in a revision control system as well as long-term preservation. It automatically generates XHTML versions of the student's website and PDF for printing.

Other systems, such as Scrivener (http://www.literatureandlatte.com/scrivener.html) for Mac OS X, support the research phase by providing a storage space for background material including web pages, images, word processing documents, notes, and so on. It then aids the author in writing long documents by allowing the creation of text snippets of arbitrary size that can easily be arranged and rearranged within the structure of the document.

In systems like the ones described above, the currently written publication could be taken to automatically allocate publications that are close to the currently written one and should be cited similar to what Watson (Budzik et al., 2001) is capable of doing. However in addition to only taking the current document into account, the context should also make use of the current location in the document for a more tactical view, or the authors' previously published articles, and their bibliographies to identify long-term interests.

This could even go down to a finer grain, for example by highlighting paragraphs as coming from another publication, reminding the author to quote text correctly in order to avoid unintentional plagiarism.

In the clinical context, (Price et al., 2002) describe their prototype called SmartQuery, which was integrated into an electronic health record system. They used relevant terms out of the patient record, converted them into MeSH terms and used them to query PubMed and other data sources. They also built a model of pathological test results and the interpretations of those results. This allows, for example, the mapping of the lab report entry "Meas ICA, Wh B" to the search term "Calcium", "Hypercalcemia" or "Hypocalcemia".

(Cimino & Del Fiol, 2007) describes *infobuttons*, a more recent prototype exploring ways to link from clinical systems to online information resources.

Result Delivery

Once the search results are retrieved, the user needs to understand the content of the retrieved information and the relations between publications. The concept of showing a list of ranked result documents as offered by traditional web search tools is often not the most intuitive way to present results. The user needs to find an answer to a question quickly and should not have to wade through long lists of results. A lot of effort is therefore going into the research on better user interfaces, with results being displayed for example as interactive graphs or even geographic maps, trying to generate direct benefit for the user.

Another key question is how to alert the user of the availability of search results. The user should not feel interrupted by a paperclip-like annoyance (Swartz, 2003). Instead the results should be shown on demand or probably via a non-interruptive notification in the graphical user interface. As example of a less interruptive interface, the Watson prototype (Budzik et al., 2001) was able to search for images to embed in a document based on the caption text. Since the user already indicated the need for a new image, the results are likely not to be seen as in interruption.

It should be noted that various research areas address the result delivery task; the examples given here are very selective.

Result Visualisation

Visualizations showing citation graph data or relationships in medical subjects have been increasingly explored recently. Some of those visualizations are interactive and allow the searcher to browse the answer space. (Whitelaw, 2009) explains his work on the Visible Archive project (http://visiblearchive.blogspot.com/) that aims to make the vast amount of documents in the Australian National Archive more accessible. Visualizations like the one in Figure 7 show the number of publications along timelines, indicate the size and the series, and the amount available digitally as well as related material.

A literature retrieval assistant in the future should have the ability to explore relations beyond those explicitly available, such as citations. If and when technologies to extract semantics from general text become sufficiently advanced, systems linking concepts like those available in the medical field today will become generally available.

Figure 7. One of the interactive interfaces created by the VisualArchive project. It shows series from the Australian National Archive as squares. The size of each square indicates the physical size of the collection in storage, the inner square indicates the percentage of the series available digitally. Selecting a series shows a summary and links to all series that are related. (© 2009 Mitchell Whitelaw. Used with permission)

Summaries

Once automatically generated summaries are of sufficient quality, they will help to describe a field to get a quick overview and might also be useful for creating an update for a researcher returning from an absence.

(Mohammad et al., 2009) compare ways of automatically generating technical surveys. They use multi-document summarizing techniques based on the citation texts.

With the increasing rate of research papers and articles being published, researchers have an increased need to quickly estimate the relevance of a text for their research interests. Apart from the title, the abstract is used to get a first impression about the content of a publication. The abstract, however, is written by the author. A more objective description of a publication might be gained by looking how authors citing that publication describe it.

A recent development in information retrieval is an increased focus on the delivery aspect of the search result. The traditional way of returning a ranked list has transformed into ways of aggregating different result types like images, movies and geo-referenced documents. This research area is now referred to as aggregated search. The workshop on aggregated search at the SIGIR 2008 conference focused on this topic. One of the more advanced approaches is the automated generation of documents based on the needs of the searcher. (Wan et al., 2008) describe a system that plans the structure of a result document, triggers searches in heterogeneous information sources, analyses the results, and summarises them to compile the result document.

Responsiveness to the User

Fellow human beings respond to a number of signals we present unconsciously about our responses to information. There is increasing work on computer understanding of such signals. It is accepted that one of the most significant signals is the direction of eye gaze. The pattern of eye gaze has been used for determining user task (Iqbal & Bailey, 2004) and for tuning enterprise search (Hawking et al., 2009). There are two constructive models of eye gaze motion for reading based on the properties of the oculomotor system, and effects of word recognition and can explain many of the experimental phenomena faced in reading (the E-Z Reader model—(Pollatsek et al., 2006) and SWIFT—(Engbert et al., 2005)). When fused with a predictive eye gaze model (e.g. (Gedeon et al., 2008)), they could be used to make some predictions as to the degree of comprehension or even appropriateness of the information presented. The latter is possible because the eye gaze path is under cognitive but not conscious control (e.g., for faces see (Palermo & Rhodes, 2007)), and so the content of view affects the way the eyes move.

The next step would be to use this information in the query formulation phase or at least to use the predicted user response to pre-fetch further information and so on.

Future Research Directions

Why do big organizations not have a pool of secretaries helping all managers, but rather one personal secretary for each manager, usually called a 'personal assistant'? It is because secretaries can play a much more valuable role if they go beyond just typing letters or mechanically completing administrative tasks, and instead grow to be the extended arm of their manager. A good secretary knows all the preferences of their manager, their preferred travel arrangements, their goals and how they like to achieve them, how to handle their incoming calls and email, how they organize their daily routines and even how they like their coffee. It usually takes a while for a new secretary to become a manager's most important companion.

Similarly, an automated literature retrieval assistant can be much more effective if it is contextualized—i.e., if it knows about its user, their tasks and information needs. It has to gather this information and learn about the user's preferences in order to build a user model, a task model and a knowledge structure model.

The literature retrieval assistant could make assumptions about a researcher's interests by taking into account their recent publications and drafts, the bibliography file they used, the references they accessed on online bibliographical sites like the ones discussed above, and the research areas of colleagues connected to them via work-related social networks such as LinkedIn (http://www.linkedin.com/) or Xing (http://www.xing.com/). Analysing the publications read by the user, as well as their type and publisher, the system can create a model of the searcher's information need. In the medical context, does the user prefer to read clinical guidelines, systematic reviews, randomized clinical trials or case studies?

But a literature retrieval assistant can be particularly helpful when a researcher's context is changing. It could analyse the publications in a field to help researchers entering a new research area at the start of a new project or when changing jobs. How did the field develop? What have been the major terms and streams, did they merge with other areas or divert into separate directions? What are the key publications in the area? Who are the key players, researchers, organisations, companies? The area could be even be presented in a way that would help newcomers to quickly gain an understanding of where ideas originated and what are the open issues. Interactive timelines and influence graphs coupled with automatically generated summaries could potentially be highly effective in this regard.

Research in the bio-informatics area has progressed well in recent years and currently available tools can take a lot of the burden off the medical researchers by extracting and pre-processing knowledge from scientific literature and aggregating it into comprehensible graphs. Modern tools allow the researcher to interactively explore the answer space by filtering the publications using medical terms from thesauri or even focus on genomes, tissues, and other medical entities and their relations. Future breakthroughs in technologies related to the automated processing of documents written in natural language may one day significantly improve bio-medical retrieval by allowing searches of relationships between biological entities such as genes and proteins.

A future intelligent literature assistant may assist medical practitioners and specialists in their daily routines. Systems that make use of task or working context, for example using data of the current patient's electronic health record to provide decision support about diagnosis or care, have been explored but are not generally available to practitioners.

EVALUATING PERFORMANCE

It is all very well to use one's intuition to design an intelligent literature search assistant and to propose all sorts of exciting features, but eventually we need to ask: Is the assistant useful as a whole? How useful? Is it more useful than other assistants? What about its specific features – do they make the system more useful, or do they make it worse?

Search for scientific literature is a sub-area of the well established field of *Information Retrieval*, which has usually based its measures on *precision:* the proportion of a retrieved set of documents which are judged useful; and *recall*: the proportion of the total number of useful (or relevant) documents which have been retrieved. In fact, information retrieval has been described as a signal detection problem (Swets, 1963) where useful documents are seen as signal and irrelevant ones as noise. Swets proposed that a retrieval system, like a missile-detecting radar, could be characterised with an ROC (Receiver Operating Characteristic) curve. In the retrieval case recall is plotted against *fallout:* the proportion of irrelevant documents in the collection, which have been retrieved. Definitions of precision, recall and fallout assume a binary definition of usefulness, but this is over-simplistic -- humans are often capable of distinguishing multiple levels of utility.

Obviously, judgments of the usefulness of a document retrieved are subjective and highly dependent upon the information need that prompted the search. Returning to the information need scenarios in beginning of the chapter, it would be natural to expect that the different scenarios would lead to very different judgments. For example, an expert searching for new material in their field would judge as useless many documents that would be very useful to a novice entering the field.

Long-Ago Evaluations from Which we can Learn

We would now like to discuss two literature search evaluations conducted many decades ago, which illustrate important principles that will be of use to us in evaluating modern systems.

The first is the study conducted by Cleverdon and others at the Cranfield Aeronautical Laboratories (Cleverdon, 1967). These studies are acclaimed by many as providing the foundation for most modern evaluations of information retrieval systems, but interestingly, utility judgments in the Cranfield collection were explicitly related to a specific task. Researchers at Cranfield who deposited a technical report in the library were asked to judge all the other technical reports (over a thousand of them) on a multi-point scale of usefulness. Essentially: (1) does the paper completely obviate the need for the new report? or (2) was the new author able to avoid experiments or derivations by citing the older report? or (3) does the older report provide background information worth citing in the newer one, or is the older report irrelevant or otherwise not useful?

In some literature search scenarios, such as carrying out systematic reviews in medicine, achieving very high recall is important. Unfortunately, in such a scenario, recall should ideally be computed relative to "all published documents". In other words, useful/relevant documents should not be ignored even if they are not included within the scope of the search system being used. In these cases, the *coverage* of the collection is a vital factor. Fascinatingly, a study of the effectiveness of MEDLARS (pre-cursor of PubMed) published by Lancaster (1969) took this into account.

The MEDLARS evaluation sampled real searches conducted over a period and sampled the documents retrieved by the corresponding query. Sampled documents were printed and sent to the submitter along with the original query for judgment on a four-point scale. Uniform sampling is critical to ensuring that conclusions drawn from the experiments may be extrapolated to real-world usage. To estimate coverage, and therefore true recall, subject experts at the U.S National Library of Medicine identified candidate documents not included in MEDLARS that they believed could be relevant. These additional documents were also printed and included in the set of documents sent to the participants for evaluation.

Between them, these long-ago studies illustrate principles that should be applied to the evaluation of a virtual librarian:

- The evaluation of a retrieval system should occur in the context of a real task.
- Evaluations should uniformly sample the population of real searches applicable to the scenario to be evaluated, so that conclusions drawn may be meaningfully generalised.
- It is appropriate to take account of degrees of usefulness. More recent work by Järvelin and Kekäläinen (2000) has proposed a measure called Normalised Discounted Cumulative Gain (NDCG) which takes into account multiple levels of utility and which may be a good choice for a single number score on which to compare alternative systems.

Other Types of Evaluation Methodology

Where the scenario being evaluated is oriented to precision rather than recall, different types of evaluation can be considered. First, we can develop an effectiveness score based on the user's interactions with the system. How many documents did they open, print or email? How much time did they spend looking at a document? Second, we can present two alternative sets of retrieved documents side-by-side to randomly selected users on a large screen and ask them to compare the sets as a whole and make a preference judgment (Krumpholz & Hawking, 2006) (Thomas &

Hawking, 2006). This second method would be a very appropriate choice for comparing two alternative virtual librarians, or for comparing a base system against the same system with a particular enhancement. Evaluating result sets in their entirety by this method has many advantages, as outlined by Thomas and Hawking, but the method is not always applicable.

Other Evaluation Factors

Since, as we have seen, there are many different functions that a Virtual Librarian might perform, evaluation and comparison of systems obviously cannot be restricted to comparing the sets of documents retrieved. Some added-value enhancements such as automatic query generation and enhancement, automatic deduction of a user's task context, and sophisticated linguistic analysis of the content of documents, may indeed be measurable by their effect on the quality of results retrieved. However others, such as tools for exploring, summarising, visualising and navigating within the results set must be measured in other ways: perhaps the side-by-side preference approach is applicable here, or perhaps we must resort to the pattern of more traditional psychological experiments with Latin Square designs, controls for individual differences, rating scales and pre and post questionnaires.

A fully-fledged Virtual Librarian is a complex system with many components, hopefully performing tasks of which previously only humans were capable. It is clear that the methodology for any attempt to evaluate or compare such systems must be very carefully thought out and tailored not only to the features of the systems being evaluated, but to the scenarios being modelled.

CONCLUSION

The modern scientist can gain access to most current, and many past, journal articles and research papers without even leaving their desk. For many of us, still remembering dusty index cards in libraries, this is a major achievement of the library science and information retrieval community and of invaluable help for researches in all fields. It must be close to what Vannevar Bush envisaged, in his highly prescient description of the hypothetical *Memex* system (Bush, 1945)

However, an intelligent literature retrieval assistant of the future could take even more pressure off the researcher by automatically retrieving literature in the absence of a specific search and by adding value to retrieved literature through pointers to related work, automatic summaries of single and even multiple documents, highlighting of relationships between concepts and alerts to new and emerging ideas.

The field is still in very active development and we can expect to see a change of paradigm away from explicit search (pull paradigm) toward automated monitoring, delivery and knowledge extraction (push paradigm). Finding appropriate methodologies for evaluating and comparing the sophisticated intelligent assistants of future will be an additional challenge for researchers.

REFERENCES[1]

Barnes, I. (2007). The digital scholar's workbench. In Chan, L. & Martens, B., editors, ELPUB2007. *Openness in Digital Publishing: Awareness, Discovery and Access - Proceedings of the 11th International Conference on Electronic Publishing*, (pp. 285–296), Vienna, Austria.

Bast, H., & Weber, I. (2007). The completesearch engine: Interactive, efficient, and towards ir&db integration. In *CIDR, 7*, pp. 88–95. Citeseer.

Berglund, A., Boag, S., Chamberlin, D., Fernandez, M., Kay, M., Robie, J., & Simeon, J. (2002). *Xml path language (xpath) 2.0. W3C working draft*.

Berner, E., & Ball, M. (1998). *Clinical Decision Support Systems: Theory and Practice*. New York: Springer Verlag.

Budzik, J., Hammond, K., & Birnbaum, L. (2001). Information access in context. *Knowledge-Based Systems*, *14*(1-2), 37–53. doi:10.1016/S0950-7051(00)00105-2

Bush, V. (1945). As we may think. *Atlantic Monthly*, *176*, 101–108.

Cimino, J. J., & Del Fiol, G. (2007). *Clinical Decision Support: The Road Ahead, chapter Infobuttons and point of care access to knowledge*. New York: Academic Press.

Claridge, J. A., & Fabian, T. C. (2005). History and development of evidence- based medicine. *World Journal of Surgery*, *29*(5), 547–553. doi:10.1007/s00268-005-7910-1

Cleverdon, C. (1967). The Cranfield tests on index language devices. *Aslib Proceedings*, 19,173–192. (republished In Sparck Jones, Karen and Peter Willett, eds, Readings in Information Retrieval.San Francisco: Morgan Kauffman. Collen, M. F. (1987). Health care information systems: a personal historic review. In Bluhm, (Ed.) *Proceedings of ACM conference on History of medical informatics*, pages 123–136, New York:ACM.

Delfs, R., Doms, A., Kozlenkov, A., & Schroeder, M. (2004). Gopubmed: ontology-based literature search applied to gene ontology and pubmed. In German Conference on Bioinformatics, pages 169–178. Citeseer.

Diederich, J., Balke, W., & Thaden, U. (2007). Demonstrating the semantic growbag: automatically creating topic facets for FacetedDBLP. In *Proceedings of the 7th ACM/IEEE-CS joint conference on Digital libraries*, page 505. ACM.

Douglas, S., Montelione, G., & Gerstein, M. (2005). Pubnet: a flexible system for visualizing literature derived networks. *Genome Biology*, *6*(9). doi:10.1186/gb-2005-6-9-r80

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). Swift: A dynamical model of saccade generation during reading. *Psychological Review*, *112*, 777–813. doi:10.1037/0033-295X.112.4.777

Evidence-Based Medicine Working Group. Guyatt, G., Cairns, J., Churchill, D., Cook, D., Haynes, B., Hirsh, J., Irvine, J., Levine, M., Levine, M., Nishikawa, J., Sackett, D., Brill-Edwards, P., Gerstein, H., Gibson, J., Jaeschke, R., Kerigan, A., Neville, A., Panju, A., Detsky, A., Enkin, M., Frid, P., Gerrity, M., Laupacis, A., Lawrence, V., Menard, J., Moyer, V., Mulrow, C., Links, P., Oxman, A., Sinclair, J., and Tugwell, P. (1992). Evidence-based medicine: A new approach to teaching the practice of medicine. *Journal of the American Medical Association*, *268*(17), 2420–2425. doi:10.1001/jama.268.17.2420

Fuhr, N., Gövert, N., Kazai, G., & Lalmas, M. (2002). INEX: INitiative for the Evaluation of XML retrieval. In *Proceedings of the SIGIR 2002 Workshop on XML and Information Retrieval*.

Garfield, E. (1963). Citation indexes in sociological and historical research. *American documentation, 14*(4),289–291.

Gasche, Y., Daali, Y., Fathi, M., Chiappe, A., Cottini, S., Dayer, P., & Desmeules, J. (2004). Codeine intoxication associated with ultrarapid CYP2D6 metabolism. *The New England Journal of Medicine*, *351*(27), 2827–2831. doi:10.1056/NEJMoa041888

Gedeon, T., Zhu, D., & Mendis, B. S. U. (2008). Eye gaze assistance for a game-like interactive task. *International Journal of Computer Games Technology*, 2008.

Hawking, D., Thomas, P., Gedeon, T., Jones, T., & Rowlands, T. (2009). New methods for creating testfiles: Tuning enterprise search with c-test. In *SIGIR 2009 Workshop on the Future of IR Evaluation*, page 5.

Iqbal, S., & Bailey, B. (2004). Using eye gaze patterns to identify user tasks. In *The Grace Hopper Celebration of Women in Computing*. Citeseer.

Järvelin, K., & Kekäläinen, J. (2000). IR methods for retrieving highly relevant documents. In [Athens, Greece.]. *Proceedings of SIGIR*, *00*, 41–48.

Kelly, D., & Teevan, J. (2003). Implicit feedback for inferring user preference: a bibliography. In *ACM SIGIR Forum, 37*, 18–28. New York: ACM

Kirchhoff, L., Stanoevska-Slabeva, K., Nicolai, T.& Fleck, M. (2008). *Using social network analysis to enhance information retrieval systems*.

Krumpholz, A. & Hawking, D. (2006). *InexBib - retrieving XML elements based on external evidence*.

Lancaster, F. (1969). Medlars: Report on the evaluation of its operating efficiency. [republished In Sparck Jones, Karen and Peter Willett, eds, Readings in Information Retrieval, Morgan Kauffman, San Francisco, 1997]. *American Documentation*, *20*, 119–142. doi:10.1002/asi.4630200204

Lawrence, S., Bollacker, K., & Giles, C. L. (1999). Indexing and retrieval of scientific literature. In *Proceedings of the Eighth International Conference on Information and Knowledge Management,* pp. 139–146. New York: ACM.Ley, M. (1997). Die trierer informatik-bibliographie dblp. *GI Jahrestagung*, pages 257–266.

Ley, M. (2002). The dblp computer science bibliography: Evolution, research issues, perspectives. *Lecture Notes in Computer Science*, 1–10. doi:10.1007/3-540-45735-6_1

Ley, M., & Reuther, P. (2006). Maintaining an online bibliographical database: The problem of data quality. Actes des Sixi`emes Journ´ees Extraction et Gestion des Connaissances, Lille, France, pages 5–10.

Medical Subject Headings. (n.d.). Retrieved December 23, 2009 from National Library of Medicine Web site: http://www.nlm.nih.gov/mesh/

Mohammad, S., Dorr, B., Egan, M., Hassan, A., Muthukrishan, P., Qazvinian, V., et al. (2009). Using citations to generate surveys of scientific paradigms. In *Proceedings of Human Language Technologies: The Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2007)*, pages 584–592. Association for Computational Linguistics.

Mudunuri, U., Stephens, R., Bruining, D., Liu, D., & Lebeda, F. (2006). botxminer: mining biomedical literature with a new web-based application. *Nucleic Acids Research, 34*(Web Server issue):W748–W752.

Palaga, P., Nguyen, L., Leser, U., & Hakenberg, J. (2009). High-performance information extraction with AliBaba. In *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology*, pp.1140–1143. New York: ACM.

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? a review of how face perception and attention interact. *Neuropsychologia*, *45*(1), 75–92. doi:10.1016/j.neuropsychologia.2006.04.025

Plake, C., Schiemann, T., Pankalla, M., Hakenberg, J., & Leser, U. (2006). Alibaba: Pubmed as a graph. *Bioinformatics (Oxford, England)*, *22*(19), 2444. doi:10.1093/bioinformatics/btl408

Pollatsek, A., Reichle, E., & Rayner, K. (2006). Tests of the ez reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*(1), 1–56. doi:10.1016/j.cogpsych.2005.06.001

Price, S., Hersh, W., Olson, D., & Embi, P. (2002). Smartquery: context-sensitive links to medical knowledge sources from the electronic patient record. In *Proceedings of the AMIA Symposium*, page 627. American Medical Informatics Association.

Rocchio, J. J. (1971). Relevance feedback in information retrieval. In G. Salton,(Ed.) *The SMART Retrieval System—Experiments in Automatic Document Processing (pp.* 313–323).Upper Saddle River, NJ: Prentice Hall Inc.

Rockliff, S., Peterson, M., Martin, K., & Curtis, D. (2005). Chasing the sun: a virtual reference service between sahslc (sa) and swice (uk). *Health Information and Libraries Journal*, *22*(2), 117–123. doi:10.1111/j.1471-1842.2005.00569.x

Salton, G., Allan, J., & Buckley, C. (1993). Approaches to passage retrieval in full text information systems. In *Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval,* page 58. New York: ACM.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science American Society for Information Science*, *24*(4), 265–269. doi:10.1002/asi.4630240406

States, D. J., Ade, A. S., Wright, Z. C., Bookvich, A. V., & Athey, B. D. (2009). MiSearch adaptive PubMed search tool. *Bioinformatics (Oxford, England)*, *25*(7), 974–976. doi:10.1093/bioinformatics/btn033

Stearns, M., Price, C., Spackman, K., & Wang, A. (2001). SNOMED clinical terms: overview of the development process and project status. In *Proceedings of the AMIA Symposium,* (p.662). American Medical Informatics Association.

Swartz, L. (2003). *Why people hate the paperclip: Labels, appearance, behavior and social responces to user interface agents*. bachelor thesis.

Swets, J. A. (1963). Information retrieval systems. *Science*, *141*(3577), 245–250. doi:10.1126/science.141.3577.245

Teevan, J., Dumais, S., & Horvitz, E. (2005). Personalizing search via automated analysis of interests and activities. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, (p. 456). New York: ACM.

The Journal of the American Medical Association (2009). *JAMA Instructions for Authors*.

Thomas, P., & Hawking, D. (2006). Evaluation by comparing result sets in context. In [New York: ACM.]. *Proceedings of CIKM*, *2006*, 94–101.

Tunkelang, D. (2009). *Faceted Search (Synthesis Lectures on Information Concepts, Retrieval, and Services)*. Morgan and Claypool Publishers.

Viegas, F., Wattenberg, M., Van Ham, F., Kriss, J., & McKeon, M. (2007). Many Eyes: A site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, *13*(6), 1121. doi:10.1109/TVCG.2007.70577

Wan, S., Paris, C., & Krumpholz, A. (2008). From aggravated to aggregated search: Improving utility through coherent organisations of an answer space. In Lalmas, M. & Murdock, V.,(Eds.) *Proceedings of the SIGIR 2008 Workshop on Aggregated Search*, Singapore.

Whitelaw, M. (2009). Exploring archival collections with interactive visualisation. In To appear in*: Proceedings of eResearch Australasia2009*: No boundaries.

Wilkinson, R. (1994). Effective retrieval of structured documents. In *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, (pp.311–317). New York: Springer-Verlag

Xu, S., McCusker, J., & Krauthammer, M. (2008). Yale image finder (yif): a new search engine for retrieving biomedical images. *Bioinformatics (Oxford, England)*, *24*(17), 1968. doi:10.1093/bioinformatics/btn340

ADDITIONAL READING[2]

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Addison Wesley.

Blum, B. I. (1987). *Proceedings of ACM conference on History of Medical Informatics*, New York: ACM.

Greenes, R. A. (2006). *Clinical Decision Support: The Road Ahead*. New York: Academic Press.

Ingwersen, P., & Jarvelin, K. (2005). *The turn*. New York: Springer.

Kopliku, A. (2009). *Aggregated search: potential, issues and evaluation. Technical report*. France: Institut de Recherche en Informatique de Toulouse.

Lehmann, H. P., Abbott, P. A., Roderer, N. K., Rothschild, A., Mandell, S., Ferrer, J. A., Miller, R. E., & Ball, M. J. (2006). Aspects of Electronic Health Record Systems. *Health Informatics Series*. Springer Science+Business Media, Inc.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge, UK: Cambridge University Press.

Murdock, V., & Lalmas, M. (2008). Workshop on aggregated search. *SIGIR Forum, 42*(2),80–83.

Shortliffe, E. H., & Cimino, J. J. (2006). *Biomedical Informatics: Computer Applications in Health Care and Biomedicine. Springer Science+Business Media* (3rd ed.). LLC.

Tunkelang, D. (2009). *Faceted Search (Synthesis Lectures on Information Concepts, Retrieval, and Services)*. San Rafael, CA: Morgan and Claypool Publishers.

ENDNOTES

1. Please notice that the Bibliography provided is very selective. Numerous publications exist in this field, some dating back half a century. Allende, R. A. (2009). Accelerating searches of research grants and scientific literature with novo|seek. *Nature Methods*, 6.
2. Introductions on Information retrieval are *Introduction to Information Retrieval* (Manning et al., 2008) and *Modern Information Retrieval* (Baeza-Yates & Ribeiro-Neto, 1999). *The turn – Integration of Information Seeking and Retrieval in Context* (Ingwersen & Jarvelin, 2005) is a book focusing on the

context aspect of information retrieval. The recently published book *Faceted Search* (Tunkelang, 2009) is a brief introduction to the topic. Aggregated search is addressed in *Workshop on Aggregated Search* (Murdock & Lalmas, 2008) and thoroughly discussed in the technical report *Aggregated search: potential, issues and evaluation* (Kopliku, 2009).

Readers interested in aspects of clinical literature retrieval and clinical decision-making might find *Clinical Decision Support – The Road Ahead* (Greenes, 2006) very helpful. Furthermore, books like *Biomedical Informatics: Computer Applications in Health Care and Biomedicine* (Shortliffe & Cimino, 2006) and *Aspects of Electronic Health Record Systems* (Lehmann et al., 2006) are helpful to get a better understanding of the clinical aspects of information science. The papers collected in the proceedings of the *ACM conference on History of medical informatics* provide an interesting insight into early experiences in the area (Blum, 1987).