The world of engineering information – 10 points to survive

Thomas Hapke
www.tu-harburg.de/b/hapke/
hapke@tu-harburg.de
University Library of the Hamburg University of Technology
October 26, 2004

Point 1: Be aware of your information behavior.

Information literacy is a crucial key skill for self-directed learning in scholarly and professional everyday life. In addition to efficient retrieval and navigation strategies information literacy includes - above all - the creativity to organize and shape your own information process in a conscious and demand-oriented way. For the searcher it is not longer questionable to find some information, but to filter reliable information out of a lot of similar offers. Every subject has its own special information media and also particular retrieval strategies to meet the subject-related information needs. The so-called "invisible web" or "deep web" contains information sources which are not collected by most search engines like Google. For example it includes the content of special databases e.g. for patents, websites secured by password access or only available in an intranet as well as script-based websites, which offer dynamic content.

There is a whole range of reasons for reading and informing for research: To give you ideas and enhance your creativity, to understand and be able to effectively criticize what other researchers have done in your subject, to broaden your perspective and view your work in context to others (direct personal experience is never enough), to legitimate your arguments, to avoid double efforts in research, to learn more about research methods and their application in practice, and to find new areas for research (Blaxter et al., 2001). Before beginning to search information reflect on your topic and specific information need, gather background information and focus your research.

Point 2: Use tutorials, subject gateways and literature guides to inform yourself about searching information

There are a lot of tutorials on the net to improve your information skills. DISCUS (Developing Information Skills & Competence for University Students) is an example of a web-based bilingual (German, English) learning tutorial for information literacy in engineering which can be used independent of time and space. DISCUS was developed at the University Library of the Hamburg University of Science and Technology and is offered at http://discus.tu-harburg.de. An American example is the Texas Information Literacy Tutorial (TILT) at http://tilt.lib.utsystem.edu.
So-called subject gateways containing collections of subject-specific links are good starting points for relevant web sites. Two examples are “EEVL (Enhanced and Evaluated Virtual Library)- the Internet Guide to Engineering, Mathematics and Computing” at http://www.eevl.ac.uk and the Engineering Subject Gateway of the TIB/UB, the German National Library of Science and Technology at http://vifatec.tib.uni-hannover.de/index.php3?L=e. Special link collections you find in web catalogs like Yahoo (http://dir.yahoo.com/) or the Open Directory Projekt (http://www.dmoz.org/) as well as in further subject gateways for narrower subjects like the “U.S. National Center for Biotechnology Information” (http://www.ncbi.nlm.nih.gov).

Literature guides give a comprehensive overview about all forms of primary and secondary literature of the treated subject. The most actual guide to information sources in chemistry, which is also of value for the process and materials engineer, is written by Maizell (1998). The whole field of electronic information in chemistry in its broadest sense – including bioinformatics - is covered by the 4 volumes of the “Handbook of chemoinformatics” (Gasteiger, 2003), which exists in a ‘short version’ also as a textbook. The most important general engineering guide is “Information sources in engineering” (MacLeod & Corlett, 2004). Guides for the life sciences come from Schmidt (2002) and Wyatt (1997), for business information from Moss (2004).

**Point 3: Use your local research library and consult a librarian or information specialist.**

Your local academic research library, e.g. a university library, can be your gateway through the jungle of information. Browsing the open shelves can give you new ideas and stimulation. The library catalog with its local holdings is an important database to search for books as well as to locate scholarly journals.

Even if the library does not possess the item you are interested in, library union catalogs offer a wide range of library materials which can be ordered through interlibrary loan or document delivery. In many countries special libraries functions as a National Library for Science and Technology, e.g. in Germany the TIB/UB in Hannover (http://www.tib.uni-hannover.de). Databases available in the local intranet give reference to further information (e.g. journal articles) not necessarily housed by the library itself. Last but not least contact the local information specialist or subject librarian who is usually keen and interested to help and to give information consulting.

**Point 4: Use encyclopedias and other reference works for first orientation**

For first-time orientation encyclopedias, handbooks and other reference works are often the first step when beginning research in a new subject. Encyclopedias contain a detailed view of evaluated knowledge and additionally also references for further reading. Libraries offer a selected range of such reference works in printed form in their reading rooms. Maybe, electronic versions are available in your local intranet. There exists voluminous encyclopedias in many specific subjects.

Examples:
“Encyclopedia of catalysis” (ed. I. T. Horváth. 2003, 6 volumes)
“The optics encyclopedia : basic foundations and practical applications” (2004. 5 volumes)
“The Sage encyclopedia of social science research methods” (2004. 3 volumes)
“Encyclopedia of information systems” (2003. 4 volumes)
Point 5: Play with search terms when exploring database features (Boolean logic, wildcard symbols, neighborhood operators, search fields, ...)

When searching a database do you often feel that you cannot take advantage of all the possibilities of the system? Changing search interfaces of varying databases or search engines force the searcher to explore the databases by themselves. He has to play with databases and to learn more about their functions: How to use logical, so-called Boolean operators to link search terms? Which special search fields are available e.g. to search only for descriptors or authors? Which neighborhood operators - to search terms in the same field or as a phrase - are available in the specific database you use? Which wildcard (joker) symbols? The latter like ‘?’ or ‘*’ or ‘$’ - which one depends on the search interface - simplify the Boolean combination with "or" of terms beginning with the same character string; e.g. searching with ‘biodegr?’ retrieves documents containing ‘biodegradation or biodegradable or biodegraded or biodegradability or ...’.

It is also necessary to "play" with search terms which describe the topic as clear as possible. The selection of search terms is the most important step during a search. Which synonyms, quasi-synonyms, or related terms should be included? Does there exist so-called controlled vocabularies like classifications, subject headings and descriptors or other database specific vocabulary e.g. like abbreviations? Too general key words lead to too many hits from which often only a fraction is useful; when using too specific key words, important information might not be found. Begin to search with one or two search terms and explore all possibilities of the database interface with multifaceted trying and quick & dirty searches! This is the only way to be sure not to overlook too much.

A search term worksheet can help to structure your query. For it the topic must be divided into components. To formulate the topic, you should think of an article or book title which exactly meets your information need. Write down terms for each component into the columns of the worksheet.

<table>
<thead>
<tr>
<th>Component 1</th>
<th>Component 2</th>
<th>Component 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>microbi? degrad?</td>
<td>aromat?</td>
<td>soil?</td>
</tr>
<tr>
<td>biodegrad?</td>
<td>polyaromat?</td>
<td>clay?</td>
</tr>
<tr>
<td>bioremed?</td>
<td>Benzene</td>
<td>compost?</td>
</tr>
<tr>
<td>microbi? decompos?</td>
<td>PAH</td>
<td>sediment?</td>
</tr>
</tbody>
</table>

For Boolean search you have to combine the terms in each of the worksheet's columns with the operator "OR" and then combine the resulting sets with the operator "AND". If possible brackets should be used because the ranking order of the operators can be different in different systems: Component 1 AND Component 2 AND Component 3, where e.g. Component 1 = (microbi? degrad?) OR biodegrad? OR bioremed? OR...

The "Boolean Restaurant" (Figure 1) is a good analogy to learn more about the difference between logical operators and the everyday usage of the conjunctions "and" and "or". In this "restaurant" the dishes are the records of the "database" and the retrieval interface corresponds to the menu.
Point 6: Search for journal articles in subject-specific databases

Actual research results you cannot find in reference works or in books. In this case you have to search in further subject-specific secondary information sources to find the primary sources like journal articles, conference papers, reports, patents, theses, ...

For getting the fulltext of journal articles found in the reference databases mentioned below you have to locate the journal in your local library, to order the article through document delivery or – the most convenient way – to read the pdf file of the article on your screen. There exist some directories to search for electronic journals available on the net, e.g. the “German Electronic Journals Library” at http://www.bibliothek.uni-regensburg.de/ezeit/ or the “DOAJ Directory of Open Access Journals” at http://www.doaj.org/. The use of normally commercial reference databases and the access to the fulltext of a specific journal is usually subject to a fee, but they are often offered within the intranet of universities or companies.

Examples of databases to search journal articles:

<table>
<thead>
<tr>
<th>Database</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical Abstracts Service at <a href="http://www.cas.org">http://www.cas.org</a></td>
<td>For all areas of chemistry, and related sciences like the materials sciences and the environmental sciences (with user interface SciFinder perhaps in your local intranet)</td>
</tr>
<tr>
<td>COMPENDEX (COMPutered ENgineering InDEX) at <a href="http://www.ei.org/eicorp/compendex.html">http://www.ei.org/eicorp/compendex.html</a></td>
<td>Most important and comprehensive database for general engineering</td>
</tr>
<tr>
<td>INSPEC (Information Service in Physics, Electrotechnology, Computer and Control) at <a href="http://www.iee.org/publish/inspec/">http://www.iee.org/publish/inspec/</a></td>
<td>Of importance because information technology plays a considerable role in all areas of engineering today.</td>
</tr>
<tr>
<td>“Science Citation Index” in the “Web of Science” of the Institute of Scientific Information at <a href="http://www.isinet.com/">http://www.isinet.com/</a></td>
<td>In interdisciplinary citation databases you can search with documents as “search terms” and answer questions like: Who have cited a specific document? How much is a document cited?</td>
</tr>
<tr>
<td>Ulidat at <a href="http://doku.uba.de">http://doku.uba.de</a></td>
<td>German database for the environmental sciences.</td>
</tr>
</tbody>
</table>
Point 7: Don’t forget searching for patents and data

Point 7: Don’t forget searching for patents and data

By granting inventors a temporary monopoly to make use of, and exploit, their invention, patents play a key role in industry around the world. Patents are an important source of information because in the research process the patent application is often the first published account of the invention and it also may be the only published. It is said that about 60 % of the information in patents is not published elsewhere. The patent specification must give the complete technical details of the invention containing often many further references to the literature.

The use of patent information is important for research and development to avoid duplication of effort (and possible infringement), to get ideas for new products and applications, to trace the development of a technology, to find solutions to your technical problems by getting detailed technical descriptions of products, and to survey the state-of-the-art. By searching patents you can identify your competitors and monitor their activities (market analysis). Please be aware that comprehensive searches for patents should be made by specialized information brokers and specialists in the complex field of patent information.

DEPATISnet - the German patent information system at http://depatisnet.dmpa.de is the first choice for German patents. Depatis contains free of charge the fulltext of every German and American patent in pdf-format, also patents from further countries. You have to know the exact patent number. Searching in other data fields of the database, e.g. in the title, the patent inventor, or the abstract field, is possible from a distinct year. So you can search for German patents in the title or inventor field from the year 1981.

Every country offers through its patent office possibilities to search its national patents and to get the fulltext, for example for US patents the US Patent and Trademark Office at http://www.uspto.gov/patft/.

Searching for substance properties’ data (Wagner, 2001/2003) is often a really difficult task. Rarely you find exactly the data you are looking for neither in locally available reference sources nor by searching on the net. Substance property data of common substances you find also in some specialized printed reference works like the “CRC handbook of chemistry and physics : a ready-reference book of chemical and physical data” (81. ed., 2000-2001), or the catalog of the Aldrich company. Data about the properties of hazardous substances (materials safety data) as well as about their handling play an important role in science and technology and also on the net.

Beyond it research libraries in science and engineering offer extensive printed substance property collections (“Handbooks”), which perhaps are also accessible in your local intranet in electronic form, e.g. the “Landolt-Börnstein - Numerical data and functional relationships in science and technology”, the “Gmelin Handbook of Inorganic and Organometallic Chemistry” or the “Beilstein Handbook of Organic Chemistry”. For learning how to use these really voluminous works it is helpful to look for tips in special literature guides for chemistry or to ask the information specialist of your local library.

Examples of databases to search chemical property data

| ChemFinder at http://chemfinder.camsoft.com | Meta-search engine for chemical substances information |

Point 8: Evaluate your search results with respect to relevance as well as quality of the document you’ve found and think about processing your information

After searching successfully you have to evaluate your search findings with respect to relevance. How to be sure, that all the potentially important documents are included in your resulting set? How to modify your query to reach this goal? But it is also important to critically evaluate the quality of the documents you found. In case the document is published in a scholarly peer-reviewed journal, the article has been evaluated by independent experts before acceptance or publication. Who is the author and what is his or her background? Why is the document being provided? How current is it?

Further processing of your search results in a form of personal information logbook – e.g. a weblog, a diary-like personal website - can help to record what, where and how you have found your items. If you want to cite your
resource or to get the fulltext you have to keep the exact bibliographical details of your source (e.g. name of the author, title of the journal paper, title of the journal, year and volume of the journal, page numbers).

For summarizing the content of the resources you have found it can help to prepare a concept map (mind map). Even today working with index cards, perhaps in electronic form, is a good way to summarize, to process, and to keep texts and learning matter. When marking and completing an index card you are forced to think about systematics and structuring as well as to identify the key points of your texts and material.

Communicating your results and research findings through an article, a paper or a poster at a conference or in a course, through a thesis, a book or a webpage (Paradis & Zimmerman, 2002) completes the lifecycle of information. The whole chain of processing information should be accompanied by critical reflection on your own information seeking behavior: Can I improve my search? Researching is often an iterative process.

**Point 9: Keep yourself up-to-date by browsing through journal contents, subscribing mailing lists and reading weblogs.**

To keep yourself up-to-date several publishers of journals offer free access at least to the contents of the journal issues or offer free table of contents via email whenever an issue is published. Another way to stay current is to subscribe in subject-specific mailing lists or read subject-specific weblogs like the bioinformatics weblog at http://www.nodalpoint.org.

**Point 10: Reflect about information ethics (intellectual property, copyright and plagiarism) and policy (ownership, privacy) as well as economics (commercial and open access).**

What is publication, what an author, a document, a journal, a collection, or a library? In the electronic world of the Internet all these terms have changed their meaning and use. It is necessary to reflect about information and its communication and use (Feather & Sturges, 2003). Issues in intellectual property and copyright increase in a “cut-and-paste” environment. Why is it important to cite sources of information? What is the right way to cite? Questions of information ethics (plagiarism) as well as information policy (ownership, access, privacy) become important. Does there exist a digital divide? Even think of the preservation of information. What will be happened with electronic records or data in 30 or 50 years?

In spite of information overload only a limited part of information is freely available on the net. Access to commercial information sources for scholarly research like reference databases and the fulltext of a specific journal is usually subject to a license fee and controlled by password. But they are often offered within the intranet of universities or companies. Open access activities – look at http://www.zim.mpg.de/openaccess-berlin/links.html - try to free access to scholarly publications at least for research and educational purposes. An example is the journal PLOS biology at http://www.plosbiology.org.
**Literature:**


**Links:**

Online-Version of this text with integrated links

www.tub.tu-harburg.de/2552.html

University Library Hamburg University of Technology (TUHH)

www.tub.tu-harburg.de

Tips for doing research online

(Susan Gilroy, Lamont Library, Harvard)

hcl.harvard.edu/lamont/resources/guides/onlinereseach.pdf

More on information literacy

www.tub.tu-harburg.de/index.php?id=911