Systematically and Efficiently Searching for Scientific Information

Edited by

Ir. A.A.K. Boxem and Ir. W.C. Oosterling

Content

Foreword

The University Library of the University of Twente has a long and innovation oriented tradition of training in information literacy. This also includes the development of training tools which can be tuned to differences in level of education, scientific discipline, specific training demands, available time, etc.

This syllabus will be used in the course 'Systematically Searching for Scientific Information'. This course is developed primarily for scientific researchers of the University of Twente. The main goal is to teach methods and techniques for systematically searching and registering scientific information.

This syllabus has recently been updated. It originates from a Dutch language syllabus accompanying the digital interactive training programme Meewiz. Although an English version of this programme is not available, this syllabus can help you to expand your knowledge of systematically searching for information. If necessary information specialists of the University Library are available to support you.

Of course we welcome any comments on the content of this syllabus.

Drs. P.G.G.M. Daalmans,

University Librarian

October 2010

0. Systematically searching for scientific information: An overview

Searching for information is time-consuming. If it is done methodically, it will save time and has the highest output. The SSI course and this syllabus, which belongs to the course, offers you the possibility to learn how to search for information efficiently and methodically.

You will reap the benefits of the SSI course if you carefully read this syllabus beforehand. Some parts of the syllabus will get further attention in contact hours during the course.

Chapter 0 is an overview. It briefly describes the four phases of an "extended search" when searching for scientific information. Whereas, Chapters 1 to 6 will treat these aspects much more thoroughly.

Phases search process   In the process of methodical searching for information, four phases are to be distinguished:
1. defining the problem statement and getting orientation on the subject;
2. developing strategy and planning,
3. performing a systematic search and organising the results;
4. evaluating the process and output; and if necessary, restarting earlier phases.

0.1 Defining the problem statement and getting orientation on the subject

In order to search for (literature) information efficiently, a good problem statement is necessary. The problem statement defines the subject that you are going to search for, as well as the objective you want to achieve. On the basis of the problem statement, and the side conditions formulated in that framework, you must decide which referring sources you are going to search for, how you are going to search in these sources, and in what order you are going to do this.

Provisional problem   Draw up a provisional problem statement and determine what demands
statement   have to comply with the compilation report, for example. On the basis of this, you start general orientation. This must lead to a more detailed demarcation of your subject, the manner of searching and the means that are used in this. For that purpose you trace relevant reference literature, by means of the library catalogue, such as (specialized) encyclopaedias, and handbooks. You can also use the Internet in an orientating way.

Demarcation	Further demarcation occurs according to:

- The period within which the searched literature has been published.
- The language in which that information may be phrased.
- The kinds of documents (e.g. journal articles, dissertations, Internet sites).

Now, you also determine with what search terms you are going to search for literature. With the help of a thesaurus, lists of keywords, multilingual specialized dictionaries and the like, these search terms can be extended; synonyms and translations can be found, more specific terms or just more general ones. If you also have to search for information via classification systems you can, , quote the relevant classification codes at this stage.

Sources

You trace the relevant (referring) sources for your subject and then skim through them. They give you an indication about the amount of (literature) information that has been published on this subject. You may make notes on this in your personal documentation system for example. Finally, you should search for persons and institutions that can/could be of use when searching for information (e.g. via referring sources). During orientation, it may appear that the subject must be adjusted, restricted or extended.

0.2 Developing strategy and planning

In the strategic phase you must finalize the preparation of your systematic search process. By doing this well, you can save a lot of precious time. You should consider drawing up a schedule within which the searching must take place, and add any relevant data and agreements. You can record this in your personal documentation system.

Planning

You should draw up a strategic planning for the search process, i.e. you could determine the manner in which and the sequence in which the selected referring sources will be searched. This means for example, that of each source that is written down, what years will be searched through, and what source specific search terms are going to be used in the search (e.g. in your personal documentation system).

The previously determined search terms must therefore be "translated" into the search terms that are specific for the sources to be searched.

0.3   Realizing a systematic search process and managing the results

Tracing information

After undertaking all of the described preliminary work, including the development of a personal documentation system, the actual trace of the information takes place with the help of the selected referring sources. You should record data on 'found titles' in your own personal documentation system (possibly via downloading) and you try to obtain the documents by downloading or borrowing them/applying for them in libraries.

Your personal documentation system also minimizes the risk of constantly re-obtaining the same information over and over again. You do this by constantly taking down notes whether the information is already in your possession, or has been ordered, what arrangements have been made with specialists, if available information has already been processed, what the possible benefit will be, where it is stored etc.

Processing information

Now you also need to start to study all of the obtained information. In this way you can determine whether the found literature indeed deals with the subject to be studied. If this is not the (sufficient) case, then the search terms will have to be adjusted. All relevant data should be recorded in the personal documentation system.

Retrieve more

Particularly when you have already found key articles or reviews, you can, departing from those articles, search for more recent literature with the help of the "Citation Indexes" as discussed in chapter 3.4. Older literature can be found via the references in the articles that you have already obtained (the so-called snow ball method). If necessary, you could contact a number of institutions or authors who are working on the same subject.

Finally, you should take steps to keep informed of new publications via alerts or SDI's.

Lack of time

During the search process, if you notice that you are running out of time, you can, after determination of the cause, take the following measures:
- consult your subject librarian / information specialist;
- claim more time;
- restrict your subject, i.e.: focus your problem statement;
- restrict your search area, e.g.: only publications of after 1990, only English-language literature, no report literature or only publications in full text databases or present in your own library;
- only search for referring source(s) of which you expect the highest yield;
- contract out part of the searching.

All measures you take must be recorded clearly and expressed with reasons. After all, you must be able to justify your method of searching. Moreover, you will be able to fill in any possible gaps at a later stage.

0.4 The evaluation of process and results

Responsibility

Through the search process you constantly make choices and establish priorities. This happens when developing the problem statement and when taking decisions about the sources to be consulted. Choosing and deciding are professional activities, which you will have to be able to justify. Therefore it is important that this happens explicitly and is recorded in your personal documentation system. Of course this also true for any necessary modifications that occur during the process.

Scope search process

Passing through all phases and steps of the search process seems very time-consuming. However, this is not necessary with good planning. However, you should remember that the whole search process does not always have to be carried out down to the minutest detail.

It is not advisable to start more and more endless refinements again and again, but you should not hesitate to return to an earlier point in the cycle.

1. Formulating a problem statement

For an efficient search of literature, a good problem statement is necessary. The problem statement indicates the subject on which you are going to search and also the target you want to achieve. On the basis of the problem statement and the side conditions you decide in which referring sources you are going to search, how you are going to search in those sources and in which order you are going to do that.

Formulating a good problem statement is not simple. If your problem statement is insufficiently clear or too elaborate, then there is a great chance you will find too much information, which will divert you from your problem. In this course some general characteristics of good problem statements will be discussed.

What? why?

In the first place, it must become clear what you want know and why you want to know this. Formulating this part of the problem statement in the form of a question often works.

Kind of information

In the second place, it must become clear in the problem statement what kind of information is demanded for the problem solution, e.g.:
- facts or opinions;
- very new information and/or older information;
- information of a high theoretical calibre or practical information;
- written in which languages;
- what kind of documents (articles, folders, Internet sites, theses, advices, policy documents ...)

Terminology

In the third place, terminology is of the utmost importance. Which concepts are at the centre in your problem statement, are there equivalents for these concepts, are these concepts clear, are there relations between the concepts taken etc.

Side conditions

Finally, it is important that there is clarity within which side conditions you are going to search for information. Think of questions such as: how much time and money is available, where can you search (libraries, Internet, experts), when can you search / when do you want to search (day time, evening, at the weekend), how do you want to process the found information and how much time do you have for this.

2. Knowledge and choice of referring sources

2.1 What are referring sources ?

When tracing scientific information, referring sources are indispensable because they provide you with titles and sites of the information you are searching for. There are different kinds of referring sources. Some traditionally distinguished sources are: bibliographies, catalogues and citation indexes. In the last few years, electronic sources (databases) have been added. Databases traditionally use referring sources. On the Internet, search engines and web indexes play an important role. Internet can serve different purposes: as referring and as informing source. Searching on the Internet is dealt with in sub module 3.3. Some knowledge about kinds of referring sources is essential. Next to this, it is important to know how you can choose from the multitude and where to find them. All these questions are dealt with here.

2.1.1 Bibliographies

A bibliography is an orderly list with references to documents, and is meant to signal the existence of those documents, irrespective of the place where they are. A bibliography can occur as an independent publication (book, journal, database or cd-rom), as a journal article or as a supplement to a book or article.

Kinds

There are many kinds of bibliographies. These can be divided up into:
- Publication areas (national bibliographies)
- Subject area (general or special bibliographies)
- Publication period of documents (retrospective and current bibliographies)
- Publication form of documents (bibliographies of books, dissertations, congress papers, journal articles, etc.)

There are also bibliographies of bibliographies.

Abstract journals

A special form of a special, current bibliography is the abstract journal. This is a periodical publication in which brief summaries of recent literature on a certain discipline are published (also mentioned as abstracts or excerpts). It mostly concerns articles from journals and/or proceedings. Abstract journals are often electronically accessible.

Current awareness service lists

Another category worth mentioning is the current awareness service lists. These lists are continuously kept informed about any recently published material in a particular discipline.
Acquisition lists of libraries are to be considered as current awareness service lists.
Current Contents are journals that offer current awareness service in the form of lists of contents of certain journals.

Individual current awareness service

In many literature databases it is possible to set up a quick form of individual current awareness service, called 'Alert' or 'Selective Dissemination of Information' (SDI). On the basis of a so-called 'interest profile' you can periodically review the latest titles that are recorded in a database. In the Picarta database, you can personally select which journals you want to receive lists of contents. Then you will receive, by e-mail, the lists of contents of new issues as soon as they have been added to the database. Also a current awareness service from Picarta is possible on the basis of a personal interest profile.

2.1.2   Catalogues

Catalogues form an essential tool when tracing and obtaining information. A catalogue is a, mostly electronic, database containing title descriptions of documents such as books, reports, journals and/or cd-roms which form a part of a collection in a specific place. Characteristic for library catalogues is the indication for the location of each document: the shelf number. In a catalogue, like the one of the University of Twente's library, you can, among other things, search for names of authors or editors, title words, year of publication and keywords or on a combination of these so-called search keys.

In the past, catalogues were found in university libraries mainly in the form of card trays. In libraries abroad and in smaller libraries you will still find them. Also the Universiteit Library Twente started with a card tray system. However, in 1986, the card system was replaced by a computer database system. The catalogue of the University of Twente's library is now one of the many PICA-databases. PICA is an organization that has been established by a number of university libraries that wanted to automate their card tray system.

2.1.3   Citation indexes

Within the referring sources the citation indexes have an important and special place. In such indexes you can, originating with a publication of which the bibliographic data are known, find later references to that publication. So it is possible to find more recent publications about your subject via a citation index. More information about working with citation indexes see paragraph 3.4.

2.1.4   Databases

There are many different kinds of databases that serve as referring sources. However, they have one characteristic in common: they are to be consulted electronically. It can be traditional catalogues, bibliographies in electronic form but also combinations of referring sources, with sometimes possibilities to click direct the online available full text informing sources as, e.g. in Picarta.

National and international organizations, trade unions and ministries are also computerizing their systems, and hence providing searchable online databases. You may search these databases via the University of Twente's library webpage.

You should realize that most databases are strongly aimed at literature in the English language.

2.1.5   The Internet

The databases mentioned in section 2.1.4 are mostly all accessible via the Internet – with some requiring subscription. In order to use the open Internet as referring source, there are search engines and web directories at your disposal. Search engines often pretend to index the "whole" Internet. You can also use them to search the parts selected by you (e.g. certain countries or discussion/news groups). Web directories work with subsequent subdivided rubrics from which you can choose. See for more information on searching on the Internet in paragraph 3.3.

2.2   Choosing referring sources

Problem statement

Choosing referring sources is not easy. There is a great offer spread over many different places. Most of them are to be searched and consulted via libraries and/or the Internet. Here your problem statement plays an important role. It determines in first instance what kinds of referring sources you will need. Problem statements can be restricted to one discipline or on the contrary be related to different disciplines. Likewise, there are referring sources that are restricted to a certain discipline whereas others are multidisciplinary. Also, in relation to the problem statement, the size and the nature of the information to be searched for contributes to the choice of the referring sources to be consulted. It makes a difference, whether you are searching for information for a lecture or for a thesis. Note that the latter search is less demanding than the first.

Assessment

Once you have traced some referring sources, it is necessary to assess them globally before you determine your choice. The first assessment mostly concerns the title and other formal characteristic of a source. A further assessment is to be formed on the basis of the prefaces or explanations with sources. In these there is often more to be found concerning the significance of the covered area, but also about the character of the information to be found and the size of the database. This makes it possible to estimate what meaning the referring source can have for you.

2.3   Finding referring sources

Search systems

Searching for referring sources will often be done in libraries and/or via the Internet. On the Web site of a scientific library, surveys of the referring sources are presented summarizing what is available.
Referring sources that are important to the University of Twente are listed on the University Library's website www.utwente.nl/ub. You can search using 'Searching information'. Most of the sources you find there can only be accessed from the University of Twente network or externally through a VPN connection.

You can also carry out an Internet search using a variety of search engines. On www.scholar.google.com, which is freely accessible, you can search for information in scientific sources.

3.   Search techniques

When you are going to search for literature about a certain subject in referring sources (such as abstract journals, catalogues or databases) you will have to 'translate' your subject into the search terms used in those sources. These can be for example, keywords or classification codes. Sources mostly give information themselves about the search possibilities. Printed sources do this in the introduction, and electronic ones in the help function. The three most common ways of subject searching are with the help of classifications, keywords or free text.

3.1   Subject searching

3.1.1.   Classifications

Classification codes

A classification is a scheme in which subjects or concepts are arranged in classes or headings. In the library, world classifications are used to arrange literature. The headings are often indicated by alpha-numeric notation codes: the so-called classification codes or systematic codes. Thus, these codes refer to subjects. E.g. see a part of the International Patent Classification in 4.3.2.
So, if a classification is used in a database, all of the publication descriptions will include a field in which the relevant classification codes are stated. These codes are allocated by experts.
There are major advantages to using a classification, as it provides a standard way of referring to certain topics. This means that you no longer have to take account of variations such as synonyms, alternative spelling or singular/plural when you carry out a search. Classifications are also an effective solution to the disadvantages of titles that do not accurately reflect the content (as is often the case with patents).
When tracing literature you should realize that a subject can belong to different headings. Moreover, in one document different subjects can be dealt with next to each other. Therefore it is advisable to search in all appropriate headings.
Before searching, it also useful to look at the complete classification, i.e. the list with available codes and their descriptions, so that you can select the desired codes. Such a list is usually available in print and/or electronically. Should the list not be available, you can still work with desired codes as long as you have found some good hits using another method. By looking at the headings of the hits that have been stored and which codes belong to these, you will discover codes that are usable, in order to continue searching.

Another type of classification is the categorization of publications in a limited number of publication types such as 'articles', 'reviews', 'conference papers', 'reports' or 'books'. This provides a useful way of refining your search, also in the second instance. (See e.g. the Scopus database.)

3.1.2   Controlled terms

A controlled term is a word or a combination of words by which a subject occurring in a document is stipulated as accurately as possible. Specialists allocate controlled terms. Controlled terms are also called keywords, descriptors or subject headings.
Using keywords gives the same advantages as classifications, in the sense that they also provide a standard way of referring to certain topics.

However, certain words are fashionable, so that in the course of time different terms have been used for one and the same subject. When searching with controlled terms in a referring source it is advisable to consult an alphabetical list of these terms used in that source.

Thesaurus   An extended form of a list of controlled terms is a thesaurus. A thesaurus is a kind of dictionary that systematically arranges the mutual relations between terms used within a discipline. It is in fact a list consisting of preferred terms (also called "descriptors") chosen from terminology. Via the descriptors can be searched in the referring source. A thesaurus often consists of non-preferred terms as well as synonyms, for example, accompanied by references to preferred terms (such as from the extra entrance "bike" to "bicycle"; see the example overleaf).

To illustrate this, look at the example below from a thesaurus in the field of traffic and transportation.

Bicycle

Extra entrance	Bike
Broader term	Means of transportation
Narrower term	Carrier cycle
	City bike
	Mountain bike
	Racing bike
Related term	Moped

Explanation of the terminology:

From both synonyms "bike" and "bicycle" the latter has been chosen as preferred term (in bold letters).

- *Extra entrance* indicates in this case that "bike" is not the preferred term in this thesaurus; along with the term "bike" see "bicycle" is mentioned.
- *Broader term* gives a reference to more general keywords, for there is a possibility that in a publication about "means of transportation" something is mentioned about a bicycle.
- *Narrower term* indicates a reference to more specific keywords; you now see that "racing bike" is also a preferred term. So if you are searching for something in that field you had better directly try this more specific, narrower term.
- *Related term* gives a reference to keywords that are not narrower or broader but do have a relation with the term used: "moped" and "bicycle" are very much traffic regulation related.

3.1.3   Free text

Other search keys

Next to classification codes or keywords other search keys are also to be used when searching in referring sources, e.g. "title words" or "abstract words". When searching in this free text you do not need to consider the lists of controlled terms when 'translating' your search question, but you do have to take into account that authors use their own terminology in their publications. That terminology comes back in titles and summaries. Then it is important to observe different styles of writing (e.g. fibre and fiber, in British and American English respectively), synonyms, singular/plural, different languages and so on. In electronic databases free text searching can often be done with search keys as a 'basic index' or 'subject'. Note that search keys allow you to simultaneously search within all abstract words, title words and controlled terms. Searching with such search keys can provide more hits but less specific results than only using title words or controlled terms.

N.B. When writing a publication, use words in the title or in the summary that someone else is likely to choose when searching for literature about that subject (e.g. terms from a frequently used thesaurus). If you have thought of a funny or playful title, add a subtitle that describes the subject clearly.

3.2 Searching databases

If you want to do a search as completely as possible it counts that a good 'translation' of the subject in search terms is important for the result. Such a translation will seldom provide one search term. Therefore it is important to know how in a search system terms can be combined to a search phrase. Unfortunately, there is no question of standardization: each search system has its own (im)possibilities.

3.2.1 Boolean operators

To combine search terms in search systems you often make use of so-called Boolean operators: AND, OR and NOT. The methodology of how to do this is described in some detail in chapter 5 of Sieverts and De Jong-Hofman (1996). You will find an abstract of this text in the appendix at the back of this syllabus.

3.2.2 Combining search terms

Analyzing in aspects

Regarding a basis set-up, first the search question is analyzed in terms of independent aspects. E.g. the demand for literature about '*corrosion of steel constructions under water*' has three aspects. Then the separate aspects are 'translated' into suitable search terms. The different possibilities are combined with the OR-operator. In the case of a search in the 'basic-index' of a database this could be as follows:

"Translation" in search terms

(aspect)	(search terms)
Corrosion	→ rust OR corrosion
Steel constructions	→ steel
Water	→ water OR sea OR maritime

N.B.:
- On behalf of the survey ability not all possibilities are mentioned in this example.
- Notice that the word 'constructions' is not included in the search terms. This word is not strictly necessary, so using it will decrease the amount of hits. Even very good hits can be excluded, e.g. about corrosion of steel test plates in water.

Combination

Finally the different aspects are connected with the AND-operator so that the search can be written as:

(rust OR corrosion) AND steel AND (water OR sea OR maritime)

If this search sentence is entered into a search system, hits will be found that include all the three aspects. Each hit contains a combination of at least one of the mentioned words per aspect.

'Simple search' modes

Most search engines on Internet do not accept extensive search sentences with a lot of nesting. In that case the search sentence must be split up in a way according to the system.

In many 'simple search' modes it is not obvious how well the system performs when more then one search term is entered; if so, use the Help-function. If Help is absent: first try one search term and then add another term. In the AND-function the number of hits decreases ('restrict set') whereas in the OR-function the number of hits increases ('enlarge set').

If you want to search systematically, it is therefore essential that you familiarize yourself with the system beforehand so you have a clear idea of how the AND, OR and possibly the NOT search operators work.

Truncation

A frequently occurring extra-ordinary form of the OR-function is the so-called truncation ('wildcard'). If, for example, you type the truncated 'polym?' into the catalogue of the University of Twente's library, the system searches all terms that begin with 'polym', such as polymer, polymers, polymerization.

Sometimes truncation at the beginning or in the middle of a word is possible, e.g. colo?r for colour/color.

Truncation saves time and often results into unexpected search terms thus more hits.

It is generally advisable to use at least five letters when truncating the end of the word. Otherwise the search is likely to include words which bear no relation to the search term. The characters used to indicate truncation vary depending on the source. A list can be found in the Help function.

Proximity operators

In some search systems it is also possible to indicate, with the help of proximity operators, the relative position of two search terms. When two search terms in a text are closer together, it may be concluded that they have a stronger relation to each other and thus the relevancy of the search result is greater.

The most used proximity operator is NEAR/n. The search phrase 'airport NEAR/5 construction' indicates that it is allowed to have a maximum of 5 words between the search terms 'airport' and 'construction'. If you do not enter the n-value then a default value is chosen in many search systems. Note that some search systems do not allow the definition of an n-value, in which case only a simple NEAR operator is available.

Phrase searching

Most search systems are able to search on two or more search terms in a specific order. Type (double) quotes at the beginning and the end of the phrase, e.g. "harley davidson" or 'harley davidson', see the Help function. Phrase searching (also called Sentence Search) and proximity-operators are an extra-ordinary form of the AND-function.

3.2.3   Search tips

Say 'and' but use 'OR'   When splitting up your search question in aspects you must pay attention to the fact that in common speech the word 'and' in enumerations means 'and/or', so the Boolean OR must be used.
E.g. when searching for books on the subject 'traffic **and** transportation', you really mean: traffic **and/or** transportation. The search sentence must be 'traffic **OR** transportation'.

Selecting aspects   Begin by selecting aspects which you definitely want to appear in a hit. Don't start with general search terms such as 'properties', 'cause', 'effect', 'correlation', or 'relation'. Scientific articles often deal with such aspects without including them explicitly in the title or summary. If you search on them, you run the risk of missing good hits. You can always refine your search at a later stage, for example if your initial searches generate too many hits.
It is also important to avoid searching on trivial aspects. For example if your topic is "the number of road accidents involving a motorcycle in the Netherlands", you would be well advised to leave the aspect "road" out of your search.

Finding search terms   Don't forget singular and plural forms (use truncation), English and American spelling (e.g. fibre/fiber). Use classification codes (e.g. patents) wherever possible. Further it is important to use synonyms and broader or narrower terms when appropriate. E.g. if you search for articles on the economic importance of the Dutch large rivers, may search for the latter as ('large rivers' OR rhine OR waal OR meuse).
More and more it is possible to use so called 'auto stemming' or 'stemming'. In that case large lists of synonyms, plurals/singulars and spelling variations are incorporated in the search system. If desired, these terms will be used automatically when the user enters a search term. Be aware of the simplicity of these stemming systems. The example previous regarding the narrower terms of the 'large rivers' in the Netherlands is not possible with auto stemming yet!

Capitals   Do not use capitals in search terms unless you only want hits written with these capitals in particular. Search systems find both, Holland or holland, when the search term 'holland' has been entered.
On the contrary, always use capitals when typing the search operators AND, OR, NOT, NEAR. This leads to better surveyability of the search sentence, and moreover some search systems demand capitals when you want to use Boolean operators.

Stop words   Avoid simple words like 'a', 'the', 'in', 'with', 'if', as a search term, because often they are suppressed in the search index of databases. If so, there is a great chance that the system will give no hits at all when this kind of words are entered as a search term.
Stop words however are allowed within phrase searching, so the search phrase "a tale of two cities" will suffice.

Too many hits	If you get too many hits on your subject to handle properly, it is possible to restrict the set by making more demands, such as

- Add more search aspects (more AND …….)
- Ask for young results, hoping older ones are included
- Important words must occur in the title instead of the abstract
- Search for 'review' or 'portal' or 'overview'
- specifying that the hits come from prestigious journals or proceedings or that they have been frequently cited (a possibility offered by the Scopus database for example);
- by opting for books that have been assessed as "good", as can be seen on websites such as Amazon.com.

Very few good hits	If you wonder, "is this all there is ?" (always do!) perform another search with:

- More/better search terms, e.g. that you found in your hits or asked an expert (more OR ….)
- Less demands: drop aspect(s)
- More languages or publication dates (older)
- More databases/search engines

Too many 'bad' hits	If there is a lot of 'noise': hits that don't have any relation with your subject: search again with adjustments, such as

- Replace truncation by full terms, e.g. car*→ (car OR cars) to avoid the words: careful, cara etc.
- If search terms occur far from each other (no relation), use NEAR or Phrase Searching
- Homonyms: exclude unwanted meanings, e.g. jaguar NOT car , if you want the animal
- searching on title words/keywords instead of words from the summary
- specifying that hits come from a particular field (a possibility offered by the Scopus database for example).

3.3 Searching via the Internet

Since the beginning of September 1973 the Internet has grown steadily. In the meantime it is, especially thanks to the introduction of the World Wide Web (WWW) in 1993, an indispensable source of information.
 Recently, scientific information on the Internet has become a valuable source, although highly qualified information is generally not freely available. Large online sources of electronic scientific information (databases, e-book/e-journal collections) are accessible via WWW, but they are not free of charge.
 An important development in this respect is the growing full text availability of scientific output of academic institutions by means of their so-called institutional repositories. In the Netherlands for instance, all universities and some other academic organizations have their own repository full of text scientific publications offering free access to these via their own website and via other services on the Internet, like Google. The repository of the University of Twente can be searched via www.utwente.nl/ub (UT publications). The repositories of all the Dutch universities and academic institutions can be searched via www.darenet.nl.

Other sources of free scientific information are web sites with:
- catalogues of (scientific) libraries
- contents of journals (via publishers)
- free electronic journals, e.g. for a list see www.doaj.org
- preprints of journal articles

 www.scholar.google.com has free full text, but also paid full text via commercial article providers

Reliability and quality of information

The reliability and quality of information on the Internet is often not, as by the traditional ways of publishing in books or journals, controlled by publishers, editors or peers. Especially with scientific information it is important to get an idea of the quality and reliability, this can be done by
- knowing who is the owner of the site and whether the author is qualified.
- verifying why the information is published on Internet (commercial?)
- checking to see if the information is published on a personal site, for which the organization named in the URL is not responsible. (Then you see in the URL the word 'users' or 'members' or the sign ~).
- verifying if the information has also been published in other media is reliable.

Search Tools	A lot of electronic information is available via the Internet. In order to make this information much more accessible, different tools and services are available. The most common way of classifying web search tools is to make difference between 'search engines' and 'web directories'. A search engine is a database of automated collected and indexed websites and offers the possibility to search in that database with the help of search terms. A web-directory (or web index) is a database of websites that have (partly) been selected by hand and have been classified according to subjects and made suitable for searching. In some cases (e.g. Google, Yahoo) both ways of searching are provided.

For both, search engines and web directories, it is possible to register your own web site in order to make it 'findable'. After registration, it is advisable to check on a regular basis whether the website appears in the database's engine/index.

3.3.1   Search engines

Search engines are the most popular tools for searching information on the web. Awarded examples (2005) are Google (www.google.com), Yahoo (www.yahoo.com) and MSN Search (http://search.msn.com). Generally the search engines have larger databases than web-directories. Moreover, the variety of the entered websites is greater. Yet the databases of the largest search engines contain no more than about 50 percent of the total free accessible web (2005). The 'deep web' information in catalogues and other databases (that is only accessible after searching in that specific database) is generally not easy to find with normal search engines. Some of them do find PDF-documents, pictures or messages from discussion or news groups (Usenet).
For possibilities, scope and other information of individual search engines, see www.searchenginewatch.com.

Parts search engine	A search engine consists of three parts:

1.   a program which collects web-sites by tracing from link to link, also indicated with spider, crawler or robot,
2.   a database (index) with copies of collected websites and
3.   The actual search-engine, software with which the database can be searched and the hits are ranked on relevance.

Indexation	Collecting and indexing of websites occurs for the most part automatically. The spider crawls the Web on a regular basis and selects websites. Changes in pre-selected and indexed websites will automatically be adjusted in the database.

How the collected websites have been indexed differs per search engine. These differences are expressed in the search possibilities that a search engine offers, and the accuracy in which a search order is carried out.

So, if you enter a search into a search engine, it does not search the entire web then, but a pre-existing index of previously visited websites.

Search possibilities	Searching for information with the help of a search engine is comparable to searching literature databases (see par. 3.2). Together with entering search terms there is a possibility to,

- make use of Boolean operators (AND, OR, NOT, sometimes AND NOT, nesting);
- demand or exclude certain words in already found results
- search for literal text ("phrase searching").

Searching with truncation and proximity is in general not possible.

Many search engines offer the possibility, besides in the complete text, to search in specific areas, such as in the title or the address (URL). Searching in a certain area provides a more accurate result than searching the complete text.

Relevance

Most search engines present the search result in a sequence of relevance. Although it is not completely understood how the search engine determines this relevance. In general, the relevance of a 'found web page' is assessed by (among others):

- the number of times that search terms occur in that web page
- the distance between 2 or more search terms
- the place of the search terms (in the URL, title, bold text or normal text)
- number of links in other popular web pages to that page
- number of links in that page to other popular web pages

Tips

Tips how to operate search engines:

- Try a number of different search engines. There is often surprisingly little overlap among the hits generated! On http://ranking.thumbshots.com, you can perform a search using two major search engines simultaneously (e.g. Google, Yahoo, MSN or AltaVista). The website generates a list of the unique hits and the overlap in the top 100 hits: the latter is usually only a few per cent!
- Use the Help-function in order to know which search facilities are offered and how to apply these.
- Use the possibility of advanced searching and the application of Boolean operators (AND, OR, NOT)
- Enter search terms as specifically as possible, use 'phrase searching'.
- Enter as many search terms as possible. Make smart use of phrase searching (e.g. "the European Union was founded in" or "Putin has * children").
- Mind a possibly double meaning of search words (Jaguar for example, is both an animal and a name of a car).
- Think of synonyms, plurals and language.
- Do not use capitals, unless specifically required.
- Do not trust the sequence in relevance of the search result and also observe hits mentioned further on in the list.

| *Meta-search engines* | You can also search using different search engines at the same time with the so-called meta-search engines. A meta-search engine has its own system of relevance ranking. Awarded examples are Dogpile (www.dogpile.com), Vivisimo (http://vivisimo.com) and Kartoo (www.kartoo.com, if you like seeing your search results visually, these meta search engines show the results interconnected by keywords). |

However, meta-search engines have two important disadvantages:
- Only simple search orders are possible, since more extensive search orders requires negotiation with other search engines with a different set of rules.
- Meta-search engines only search a random part of the databases of other search engines.

3.3.2 Web-directories or Web-indexes

| ***Difference with search engines*** | Whereas search engines index websites automatically, web-directories effect this partly by hand. With some web-directories the websites are both selected and classified by editors. Changes to a website are not automatically recorded. |

The most remarkable difference with search engines is the way in which can be searched in the database. Whereas in search engines it is found by using search terms in combination with all kinds of extra facilities. However, in web-directories searching is done by means of browsing through the headings and sub headings.
The gap between search engines and web-directories is narrowing, almost all search engines have classified a number of popular websites, so these can also be searched via a list of headings and web-directories offer the possibility to use search terms.

| ***Examples of Web-directories*** | Yahoo (www.yahoo.com) is a general web-directory. Suggested and collected web-sites are automatically added to the database. Virtual Library (http://vlib.org) and Infomine (www.infomine.ucr.edu) are web-directories with selected scientific information, just like Librarians Index to the Internet (www.lii.org). |

3.3.3 Choice of means

When do you choose a search engine and when do you choose a web-directory, when searching for information on the Internet? This choice is determined by the kind of question you have or the kind of information you are searching for. The following table can be used as a guideline.

Kind of question/information	Means in searching
The general question, survey of information about a certain subject, mostly in the orientating phase, such as which technical journals are available on the Internet.	Web-directories, both general (Yahoo) and more specific and selective (Argus Clearinghouse, MedWeb, Librarians Index, Virtual Library).
Specific, but frequent questions such as flight schedules of a well-known airline company.	General web-directories (Yahoo). Search engines (Google, MSN Search, AltaVista, etc.).
Specific, but less frequent questions such as information about the construction of an airport in the sea.	Search engines
Questions about a topical, much-discussed theme, specifically or generally, such as a book that has had publicity lately.	Discussion groups or news groups (Usenet, available through Google, Yahoo, AltaVista, etc.).

3.3.4 Searching for specific information on the Internet

Scientific information

Many catalogues of scientific libraries and other (subscribed) databases with scientific information are accessible via the home-pages of Universities and/or their libraries (e.g. www.utwente.nl click on 'University Library'). There are Universities that, independently or in mutual co-operation, offer selecting, indexing and search facilities.
For instance, a network of British universities maintains a large database called INTUTE, with information on academic websites in all kinds of areas (www.intute.ac.uk).
Search engines that specialize in searching websites that contain scientific information include www.scholar.google.com and www.scirus.com, which is run by Elsevier.
Web directories that specialize in accessing scientific information on the Internet include Infomine (http://infomine.ucr.edu, mentioned above) and Academic Info (www.academicinfo.net).

Via the internet-address http://xxx.lanl.gov you can search for pre-publications (pre-prints) of articles in the field of physics, mathematics and informatics.

Geographic determined web-sites

Websites originating from a certain country can be selected with the help of their www-address. Google and AltaVista e.g. ('advanced search') offers the possibility to search in the field 'domain'. This field contains the country-codes that have been entered in many URLs, such as 'nl', for The Netherlands, or 'de' for Germany.

Discussion groups	If you are searching for discussion groups on a certain subject, then you can consult http://lists.topica.com or http://groups.google.com.

3.4 Searching citation indexes

Citation indexes play an important role in many literature searches. The essence of 'indexes' is that you are able to find more recent articles, departing from previous bibliographic data that is already known. Therefore, if you have an important article or book, you can find more relevant recent articles including any colleagues or institutions that are actively working in that area. Citation data are often used to obtain an indication of the quality of journals or scientific output.

Web of Science	Until around 2005, citation data were only collected by the publisher of Web of Science. This database (formerly a printed reference work) consists of three sections: • Science Citation Index (SCI), natural sciences/technology; • Social Sciences Citation Index (SSCI); • Arts & Humanities Citation Index (A&HCI). These indexes only include articles from international scientific journals which meet certain quality criteria.
More options	Nowadays there are more options available when searching citation data. For example the Scopus database and www.scholar.google.com both state for every hit whether the document has been cited by other documents included in their database. They also provide the option of clicking directly to the publications containing the citations.
Advantage Web of Science	Web of Science also enables its users to look up a source article and click through to the articles in which it is cited. However, Web of Science also features and extra search facility: it enables you to search directly in the index of documents cited: the Cited Reference Index. This has the advantage of enabling users to trace articles that cite the source article in a reference that contains minor errors or spelling variations. It is widely known that many references contain errors such as incorrect or missing initials in authors' names or mix-ups between volume numbers and page numbers. Such errors also find their way into the information contained in citation databases. This means that a garbled version of the title is often provided when books, proceedings or dissertations are cited. Authors' names with prefixes such as de and van der are also particularly likely to generate errors. If a source article in Web of Science is first traced by means of the cited authors index using a comprehensive search on the author's name, all of the various descriptions of the source publication can be clicked on in the Cited Reference Index shown (including those with minor errors), thereby allowing all of the articles containing citations to be shown at once.

Example Web of Science

Imagine that you are looking for citations of a thesis by J.P. de Jongh, who obtained his doctorate at Utrecht University in 1971.

Start by using Cited Reference Search to search for (* is truncation character):

Cited Author: dejong*   Cited work: thesis*

Then you will find (albeit not in a row) the following 7 hits with a total of 32 citations.

Times cited	Cited Author	Cited work	Year
1	DEJONG	THESIS U UTRECHT	
1	DEJONG JP	THESIS U UTRECHT NET	1971
1	DEJONG H JP	THESIS RIJKSUNIVERSI	1971
24	DEJONG H JP	THESIS U UTRECHT	1971
3	DEJONG H JP	THESIS U UTRECHT NET	1971
1	DEJONG H JP	THESIS U UTRECHT UTR	1971
1	DEJONG H P	THESIS U UTRECHT	1971

It is not entirely clear whether the correct Cited Author is being referred to in the first hit, but for the others this is very likely, given that the correct year is stated in every case.

If you then go on to search on

Cited Author: jong* Cited work: thesis*

you will see another three hits (again not presented one after the other) with a total of 9 citations.

Times cited	Cited Author	Cited work	Year
1	JONGH JP	THESIS U UTRECHT NET	1971
7	JONGH JPD	THESIS U UTRECHT	1971
1	JONGH JPD	THESIS UTRECHT	1971

In these cases, the prefix de was not included or only half included after the

Search tips for citation indexes	• When searching on a person's name, it is sensible to first search on the surname (followed by the first initial at most) followed by the truncation character * so that any variations in initials (or first names) will also be included. You can then select all the relevant names from the resulting list. • If you are searching on an author's name that includes a prefix, search for both possibilities: with and without the prefix or prefixes. • Do not necessarily assume that details such as initials, volume numbers and years of publication will be stated correctly. • Consider whether the surname might give rise to a range of spellings, e.g. "jong" or "jongh". • Differences between an author's official name and the name by which they are known can also make them less traceable, as can a change of surname due to marriage. • If you search for cited journal titles, it is important to make use of a special list of abbreviated titles. • If you are searching for citations of a certain book, only fill in the author's name, as the titles of cited books often appear in the system in abbreviated and/or garbled form.
Journal Citation Reports	ISI, the publisher of the Science Citation Index, annually launches a report with numerical data regarding the citations of journals, which have been entered in the citation indexes. It is prudent to use these Journal Citation Reports if you want to judge the quality of a found article, but also if you want to choose a journal you want to publish in. In the Journal Citation Reports a list is given, per discipline, of journals mentioning the 'Impact Factor' and 'Cited Half Life'. These figures give an indication to the quality of a journal.
Impact Factor	A high Impact Factor means that a journal publishes articles, which are in general cited a lot. The Impact Factor of a journal is the average number of times that an article from that journal is cited in the two calendar years following the publication of the article.
Cited Half Life	The Cited Half Life of e.g. 1998 indicates 50 % of the total number of publication years (calculating back to 1998) that the journal in question has been cited in that year (i.e., 1998). For example, if a journal has a Cited Half Life of 8, this means that half of the articles cited are over 8 years old and the other half are under 8 years old. There are obviously many older articles cited from a journal with a high Cited Half Life. So that journal includes articles that are generally cited long after publication (standard articles).

Eigenfactor Metrics   Since 2007, even more comparative figures have been included in the Journal Citation Reports, such as Eigenfactor Metrics. The Eigenfactor Score takes into account measurements of the average amount of time that scientists devote to reading certain journals and eliminates self-citation from the citation figures.
The Article Influence Score calculates a kind of standardized number which reflects the relative influence of a journal in its field. In addition to being available through a subscription to the Journal Citation Reports, these Eigenfactor Metrics can also be viewed free of charge on www.eigenfactor.org.

Quality of journals   Journal Citation Reports (and also citation figures per author) are used by the management of scientific research institutions in order to determine the quality of scientific output. However, it must be considered that the figures are strongly dependent on the discipline; it is possible that the publication culture and the number of scientists working in this area can vary considerably. Publications or journals that do not have the same target audience, are not allowed to be simply compared with each other. Moreover, the (S)SCI is mainly aimed at the Anglo-Saxon language regions, whereby the emphasis lies on the scientific disciplines considered as important there. It is also worth remembering that far from every journal is included in Web of Science. Nor does it include conference proceedings, books or theses/dissertations.

Hirsch Index: a measure of the quality of *scientists?*

H-index

In 2005 physicist J.E. Hirsch published an article in which he introduced his Hirsch Index or h-index. In simple terms it can be described as follows: a scientist has an h-index of 5 if he/she has written 5 articles all of which have been cited at least 5 times. This also indicates that he/she has not written 6 articles which have all been cited at least 6 times, as then he/she would have an h-index of 6.
Hirsch argued that this index is a better indicator of the quality of individual scientists than the figures that were commonly used at the time, such as the total number of articles written (which says nothing about the quality of the articles) or the total number of citations (which can be strongly influenced by a small number of survey articles).

The h-index therefore incorporates both the number of articles written and the number of citations, summed up simply and directly. The h-index of a scientist cannot diminish. It is a measure of a person's cumulative performance. As long as an author continues to be cited, their h-index can grow. But in its simplicity, the h-index is also a fairly rough method of calculation: large differences in the number of citations need not influence the value of the h-index.
For the original definition and purpose, see the article by Hirsch on http://arxiv.org/abs/physics/0508025.

Comparison between scientists?

As with journals, it is not possible to make a qualitative comparison between scientists from various disciplines on the basis of these types of scores. The publishing and citation culture can often differ considerably, even between adjacent fields of research. Of course, these scores do not take into account factors such as applicability and the actual use to which scientific work is put. They should therefore be treated with caution. Within a single field, this type of score can be used to produce an easy comparison of the differences between scientists in terms of the number of publications and citations, but these are not the only factors that determine the quality of a scientist.

What is more, the h-index also depends on the database from which the citation figures are drawn.

Databases with h-index

Both Web of Science and Scopus contain not only the data on the articles but also data on references included in the articles. In both cases, an individual's h-index is reasonably simple to determine.

First they always produce a collection of the publications by a certain author which are included in a given database. When doing this, it is important to consider any differences in the way in which an author's name might be written (e.g. initials/first name, surname prefixes).

At this stage the 'Create Citation report' function in Web of Science automatically determines an h-index. This value is based on the citation data of articles in quality peer-reviewed journals from 1945 onwards.

In Scopus, an h-index is calculated using the 'Citation Tracker', based on citation data for journal and conference articles from 1996 onwards.

Scopus also provides an 'Author h-graph', which shows at a glance how often a writer's articles have been cited. The number of citations per article are presented in descending order.

The h-index is then presented as the intersection with the y=x diagonal. In the example given, the h-index equals 5.

It is important to remember that, in addition to the author's name, other data on the cited articles are not always stated uniformly in the databases.

This means that an automatically calculated h-index sometimes has to be adjusted manually, as the same citation can be registered as different citations if they have been recorded with slight differences.

In principle, it is possible to calculate an h-index on the basis of combined citation figures from both Scopus and Web of Science. Particularly in cases involving a large number of citations, this is a laborious process.

When people refer to "an author's h-index", they are usually referring to the figure based on the data from Web of Science (which contains articles from quality journals). This only consists of article

citations whereby both the article cited and the article making the citation are included in Web of Science. When introducing his index, Hirsch himself took these data as his basis.

Google Scholar
Sometimes an h-index is also determined on the basis of data from Google's academic database www.scholar.google.com. However, no quality control is applied to this data. For example, contributions to symposia and workshops are included: publications which are normally not assessed on quality and which are not therefore classified as scientific literature. For this reason, we will not describe this method in greater depth in this document. For more information, you can consult www.harzing.com/pop.htm.

Article in Wikipedia
You can also consult the extensive Wikipedia article, which also contains substantial criticism of this type of quality measurement: (http://en.wikipedia.org/wiki/H-index).

4.   Informing sources

Informing sources can be classified in
- General informing sources, aimed at one broad or more disciplines (encyclopaedia, dictionaries, large literature reviews and handbooks)
- Specific informing sources, aimed at a limited field of interest (dissertations, textbooks, reports, journal articles, etc.). Some particular kinds of specific informing sources are statistics, patents, standards, parliamentary papers and European Union documents. We refer to these as 'special informing sources'.

In this chapter the different kinds of informing sources are discussed whereby the special sources will be treated extensively because of their complicated nature. It should be obvious that the importance of the different kinds of informing sources differ per discipline and per situation. You can find informing sources via referring sources. When searching a special kind of informing source, in catalogues often a specific type of descriptors can be used: the form descriptors (e.g. dissertations, encyclopaedias). In most bibliographic databases there is the option to search for special 'document types' (e.g. conference articles) or 'treatment types' (e.g. applications, review). See the Help-function.

4.1   General informing sources

Encyclopaedia

An encyclopaedia can be considered as a descriptive and educational dictionary. Per entry descriptions are presented, mostly written by specialists in the concerning fields. Occasionally some literature references are given along with those descriptions. There are roughly two kinds of encyclopaedia that need to be distinguished: general encyclopaedia (e.g. the Winkler Prins or the Encyclopaedia Britannica) and a subject encyclopaedia, that aims at a specific discipline or theme. Although encyclopaedia are principally used in the first phases of the search process, when the problem statement and the search strategy under discussion, it is often very useful to refer to them at a later stage. Certainly subject encyclopaedias often yield more fundamental information, which can be useful when processing information.

   Most University Libraries offer entrance to one or more large online encyclopaedias. A free online encyclopaedia is www.wikipedia.org. Everyone can contribute content and there is a limited form of quality control. The quality and depth of the items therefore varies considerably.

Dictionaries	A dictionary is an alphabetically or systematically arranged inventory of words or phrases of importance. Dictionaries are principally used in the first phases of the search process, certainly when searching for information concerns a new and/or relatively unfamiliar subject. Apart from that thesauri (see 3.1.2) can play a part in this, and in their own way also function as dictionaries. Online thesauri are findable via the search page of bibliographic databases, a free online linguistic dictionary for all EU languages is offered by the European Union at http://ec.europa.eu/translation/index_en.htm. The Web site of the University library of Twente provides online access to Van Dale dictionaries.
Literature reviews	Literature reviews provide a review about the state of knowledge or insight with relation to a certain subject on the basis of relevant literature. They are also called 'advances', 'reviews' or 'state-of-the-art' publications. They mostly contain substantial lists with references to the original sources and they are written in a compact way. So literature reviews are very informative and can save a lot of time, certainly when information must be collected and incorporated in a short space of time. Often it is not important to search in literature older than the literature review. Literature reviews are presented in different ways: as reports, articles, volumes or textbooks. As mentioned before, most bibliographic databases offer the possibility to search for reviews via search options like 'document types' or 'treatment codes'. See the Help-function.
Handbooks	Handbooks give a systematic summary of a scientific field. They are useful when developing and formulating problem statements and search concepts to be used, but they can also have a useful function as reference work. The
Textbooks	same applies to textbooks.

4.2 Specific informing sources

In specific informing sources (aimed at a specialist audience) new information is presented. Such sources can vary considerably in quality, since most scholarly publications are usually assessed by specialists, editors or publishers. Therefore, if such an assessment is made beforehand, then it is an indication of the higher quality of the publication.

Congress reports	Congress reports contain recent research findings presented on congresses. The separate contributions often give new and reasonably detailed information. The bundles as a whole, present a good overview of the state of affairs with respect to the subject of interest.
Theses	Theses or dissertations are academic writings on the basis of which the author obtains the doctor's degree after successful defence of their thesis. Before the author is admitted into that defence, a committee of specialists assesses the content of his/her thesis. Theses usually furnish the reader with an extensive literature review and new results (obtained from research) on the subject. Therefore, they can be very valuable sources of information.

Journal articles	Journal articles can have various values in searching for (scientific) information. In the first place, it is important to make a distinction between professional journals and scientific journals. Professional journals present developments and news from the discipline, whereas scientific journals focus more on results of and developments in scientific research. Many scientific journals are only published electronically and are distributed via the Internet (usually not free of charge!). Professional journals do not have a high scientific status, but they can be very informative when the searcher is not yet familiar with the subject. If articles appear in leading scientific journals, they are firstly reviewed by a group of specialists (peer review). They contain a list of references and describe the research method used. It often takes a long time before a submitted article is actually published in a scientific journal: a year or more is by no means exceptional. Waiting lists are usually long, since the number of publications is increasing and because the peer review procedure is very time consuming. The status of the journal is an indication for the quality of an article. In many disciplines the Journal Citation Reports (see 3.4) play a role when determining that status.
Report literature	The quality of report literature varies considerably. In many cases new information is presented in this format. This is sometimes to be regarded as confidential or classified and therefore hard to procure.
Books/Monographs	Books (in scientific language 'monographs') are about a defined subject or part of science. They are marketed via publishers that deal with quality control. Leading scientific publishers have to keep a good reputation and will therefore have to take care that information they publish is reliable and new. Since publishing a book takes a considerable time, the information they contain can also often be found in previous conference reports or articles. A book usually remains on sale for a number of years. Successful books may then be reissued and updated in a new edition.
E-books	There are two kinds of electronic books or e-books. For personal use, people can download entire books or purchase them on CD to be read on a computer screen or e-book reader. Institutions usually purchase access to e-books which are stored on the publisher's server. In this set-up, there is a maximum to the amount of information that can be downloaded. A major advantage of e-books is the ability to search the entire text of the book. But because many e-book providers use their own platforms, it is not yet possible to search all available e-books at the same time.
News or discussion groups	Via news or discussion groups new information is often circulated as well, e.g. in the form of preprints (see also 3.3.4).

4.3 Special informing sources

Statistics, patents and standard specifications, parliamentary papers and documents of the European Union can also be regarded as special informing sources. Usually new information is to be found but because of their complex structure that is not always simple. Therefore, special attention will be given to both the content and the accessibility of these informing sources. The University of Twente's library possesses actual collections of statistics, standards, parliamentary papers and documents of the European Union.

4.3.1 Statistics

In statistic publications mainly numerical data about societal phenomena are presented which have been registered in tables and/or graphics. For The Netherlands the institute Statistics Netherlands (abbrev. in Dutch: CBS) plays an important part in this field. Statistic publications are published both in print and electronically. The StatLine database can be consulted via www.cbs.nl.
International statistic material is made available by the EUROSTAT, the ILO, the UNESCO and UN.
In order to be able to find the demanded statistic material easily, the use of the available registers is recommended.

4.3.2 Patents

A patent is a right of ownership awarded by the government for a certain invention. This means that the invention may only be exploited by others after an arrangement has been made with the inventor, i.e., a licence agreement. This ownership is granted on the condition that the description of the invention in the patent is in the public domain. What kinds of inventions can be patented differs per country. In The Netherlands, a patent is exclusively related to new products or processes whereas new plants fall under the breeder's right. In the United States, a distinction is made,

Utility Patents
Design Patents
Plant Patents
Legal demands

resulting in three different patent categories: utility patents for new products or processes, design patents related to the appearance of something and plant patents for vegetative reproduction of new plants.
Moreover, the legal demands appointed to patents differ per country. In The Netherlands inventions are patented if they are

- new (not published earlier regardless where, when and by whom),
- inventive (sufficient level of invention; not obvious for the expert) and
- industrial applicable.

Period of validity	A patent is a legal document of which the validity is limited to those countries where the patent has been awarded. A patent is awarded for a certain period of time only, this referred to as 'validity'. In The Netherlands, validity is 20 years. In most countries patents are charged annually in order to maintain the patent. Note that a patent loses its validity if the annual subscription is not paid. In 2008, the retrieval service in The Netherlands for 6-year patents was discontinued.
Construction	A Dutch patent publication is constructed from a number of settled parts: • preface (bibliographic data; mostly an extract; sometimes a figure); • content (the familiar standard technique, and what is wrong with it); • description (of the invention and the working/construction); • conclusions (summary of what the invention implies); • possible figures/graphs/schemes (making the invention visual).
Application procedure	The application procedure is just like the demands a patent has to comply with, it is legally arranged and differs per country. In The Netherlands, applications are submitted to the Patent Office. Usage can be made of supranational patent routes, arranged in international treaties, such as via the European Patent Office (EPO), via the Worlds Intellectual Property Organization (WIPO) or in conformity with the Patent Cooperation Treaty (PCT). When a first patent application has been submitted, it is possible to submit an application in or for other countries up to one year after initial submission (the priority date or priority application).
Employees	If an employee invents something in the course of their employment, in the Netherlands the rights to the invention are registered under the name of the employer. However, the employee is named in the patent as the inventor and he/she is usually rewarded by the employer for the invention.
Novelty search	The novelty search, which is carried out by the Patent Office, must give insight in that what is already known and is protected in a patent. The report of the study can be a reason to rewrite or adjust the patent application. You are not obliged to undertake a novelty search. Also the regulations to award a patent differ per country. In The Netherlands, the 1995 Patent Act allows a patent to be awarded without novelty search. The applicant him/herself will have to calculate the possible legal consequences of his/her application.

Searching for Patent publications

Searching for patent publications can be useful for three reasons. Firstly, the information in patents is unique. It is estimated that 70 per cent of the information in patent publications has never been published elsewhere. Secondly, a patent shows whether a certain invention is protected or not. Thirdly, patent publications contain important information about activities of research and development in a certain field or within (competitive) companies.

Patent publications can be traced with the help of a patent databases. These databases contain in general the information that is in the preface. Thus, patents can be found with the familiar search keys, such as name of author or applicant, words from the title of the patent or words from the abstract.

Nevertheless, searching for information in patents is not simple. Because of the demand of publication there is sometimes masking language. Often the patent is limited to sub techniques and secondary parts of an invention so that essential details are not made public. The titles of patents are often rather insignificant. Moreover the text in a patent is not focussed so much on what a certain invention is, but more on what it does or how it works. A chair for example can be indicated as a 'sitting device'. So you don't find such a patent with the search term 'chair'. Therefore, it is advisable to leave the search for patents to a specialist if a thorough search for patents is demanded and specifically when it has to be a novelty search.

Classification systems

An important means when searching for patents is the classification system. At least one classification code is awarded to a patent, but often sub codes are also awarded to sub aspects.

An example of a part of the International Patent Classification (IPC) follows below:

B:	Processing; Transport
B66:	To hoist; To lift; To pull up
B66B:	Lifts; Escalators or moving floors
B66C:	Taps, Load-bearing elements or fixtures for taps, capstans, winches or hoists
B66D:	Capstans; Winches; Hoists, e.g. pulley blocks; Hoisting apparatus
B66F:	Not classified elsewhere hoist, lift, pull up or push, e.g. fixtures that practice a direct lifting power or pushing force on the surface of a load.

The IPC is used for the European patents and is to be found amongst others on a webpage of the World Intellectual Property Organization (WIPO): http://www.wipo.int/classifications/en/ .
Via this webpage you can also find the official keyword list. The American patent office, the US Patent and Trademark Office (USPTO) uses a classification system of its own: http://www.uspto.gov/go/classification/.
Also the European Patent office uses its own classification, see www.espacenet.com.

Patent families   When searching for patent publications it is important to know that patents that have the same priority date or priority application form a patent family. Mostly this concerns the collection of patents that protect one invention internationally. An example of such a patent family from the World Patent Index is shown here:

PATENT INFORMATION:

PATENTNR	KIND	DATE	WEEK	LA	PG	MAIN IPC
DE3805522	A	880908	(8837)*		6	
US4841144	A	890620	(8931)		6	
SE8800561	A	80828	(8941)			
DE3805522	C	920827	(9235)		5	H04B010-02

APPLICATION DETAILS:

PATENTNR	KIND	APPLICATION	DATE
DE3805522	A	DE 88-3805522	880222
US4841144	A	US 88-160769	880226
DE3805522	C	DE 88-3805522	880222

PRIORITY APPLN. INFO: JP 87-U29813   870227

From the last line can be derived that the priority application in Japan (JP) was submitted on January 27, 1987. The 'application details' on top of that show that just within a year two patent applications have been submitted in Denmark (DE) and one in the USA (US). These applications were honoured according to the 'patent information' in DE in respectively 1988 and 1992 and in the US in 1989, whereas also in Sweden (SE) patent was granted. Moreover, it appears from this example that every country has its own number system.

World Patent Index   Via a patent specialists The World Patent Index (Derwent) databases can be consulted. This is the most important patent database if you want to search according to subject. The patents of about 40 countries and institutions are indexed here, among which the European Patent Office (EPO). The World Patent Index goes back to 1963 (for medicines). There is a record present per patent family completed with an abstract and various indexes. Nowadays the most important illustrations from the patent can also be asked for. There is also a Derwent patent database by which, e.g. in the novelty search, can be searched for cited patents: Derwent Citation Index.

	Some other patent databases that you can consult yourself within the University of Twente are:
Espacenet	The European Patent Office provides the Espacenet on http://ep.espacenet.com/, which contains more than 60 million patent descriptions, both from inside and outside of Europe. Many of these patents are available in full text. You can search in the title and abstact words, wich are all loaded in English, so you don't have to search with words in any other languages. No more than ten search terms are allowed in the text boxes. Many patents are available in full text via Espacenet. If a patent is not available in full text on the Internet, it can be obtained through an interlibrary loan like any other document. The University of Twente does not have its own patent collection.
Google Patents	With www.google.com/patents you can search in the full text of the USA patents. It is not clear how complex search strings will be handled. Proximity operators and truncation can not be used.
Chemical Abstracts	The Chemical Abstracts, in which next to other information also the chemical elements in patents of 27 countries, of the EPO and of the WIPO have been indexed as from 1967. Keywords are awarded from the Index Guide and Registry numbers, so that you can search by both subject and material name. Regarding a patent family, one patent is indexed with an abstract and various indexes. The patent family is only available in the printed Patent Index.
More information	More information about patents or retrieving patent information the reader is referred to the following institutions: • NL Patent Office, Rijswijk http://en.octrooicentrum.nl/ • International Patent Research Office (IPRO), The Hague (www.ipro.nl) Furthermore information may be found via the following internet addresses: • www.uspto.gov (United States Patent and Trademark Office) • www.european-patent-office.org (European Patent Office) • www.micropat.com (MicroPatent's Patent Web Services)

4.3.3 Standards

Standards are (inter-)national agreements for any field of technology or company for the advancement of unity and inter-changeability. Standards are determined and published by standardization institutes. Especially in the technology sciences, standards can be very informative. They provide mainly factual information.

In a standard it is indicated, among other things, what applications there are, how the concerning 'product' can be tested, and what the demands are for the test results.

There are many different standards, e.g. DIN (Germany), ISO (International), NEN (The Netherlands) and ASTM (US). An attempt is made in order to 'tune' the national standards more and more to the international standards.

NEN

The NEN (www.nen.nl) acts in The Netherlands as central point where national and international standards are obtainable. The NEN advises on the use of standards and standardization issues.

If there are still no standards available for a new product a producer can report this to the NEN. If a new standard is demanded, then NEN will form a (inter)national committee with representatives from the industry to produce a draft standard. After some time the draft standard, possibly with changes, is transformed into a definite standard. So producers can participate actively in realizing standards.

4.3.4   Parliamentary papers

Parliamentary papers contain reports of meetings and all printed meeting papers of the First Chamber and Second Chamber of the Netherlands.

The OPmaat database (see the University Library home page) is an important source for official publications. It contains full-text information from 1995 onwards, accessed by a search engine.

Full-text parliamentary documents from 1879 to 1995 can be accessed via http://www.statengeneraaldigitaal.nl/.

The University of Twente's library has printed copies of parliamentary documents from before 1879 in its collection.

All parliamentary documents are only available in Dutch.

4.3.5   Documents of the European Union

At different universities in The Netherlands there is a European Documentation Centre (EDC). Such a centre possesses all official documents of the institutions of the European Union (EU, formerly European Communities). The publications are arranged in twenty rubrics, based upon the areas of work inside the EU.

Official Journal of the European Communities

The Official Journal of the European Communities is the official newspaper of the European Union. It provides information on legislation, proposals for legislation, official announcement and information on activities of the institutions of the EU. Over the years, there have been different amendments to the title, format, set up and registers.

Since 1968, the following division has dealt with: Series C, Series L, Supplement (Series S) and Proceedings of the European Parliament.

From 1-1-2000 the Official Journal, Series L and C, are published monthly on a CDROM. The Supplement has been published and has been available on CDROM for several years now.

Series C	In series C the following documentation is published:
	• Statements of the Commission;
	• Opinions, minutes of meetings, resolutions and questions in writing of the European Parliament;
	• Advice from the Economic and Social Commission;
	• Proposal from the Commission to the Board of Ministers for new legislation;
	• Cases brought before the Court of Justice and also a brief summary of the verdict;
	• Vacancies at institutions of the European Union.
Series L	In series L, all community legislation is published. A great part of the legislation published is formed by resolutions with relation to daily management, which have been taken in the framework of the agriculture policy and which in general have a restricted validity. The other legislation and administrative resolutions are of more importance. Two kinds of notations are dealt with, viz.:
	• Notations for regulations, e.g. (EEG) 4283/88 88 is the year and 4283 is a serial number within the concerning year.
	• Notations for all resolutions which are no regulation (ordinances, resolutions, guide lines and recommendations), e.g.: 88/667/EEG 88 is the year and 667 is a serial number.
Searching in the Official Journal of the European Communities	Official Journals of a terminated publication year can be searched by means of the year index. This year index contains an alphabetical and a systematic part. The alphabetical index is a keyword index based upon the so called Eurovoc-thesaurus. In the systematic index, regulations, other resolutions and cases that have been brought before the Court of Justice are recorded by number. Searching in the Official Journals published on CDROM (from 2000) can be done with the help of a search menu available on the same CDROM.
Directory of Community legislation in force	Twice per year the Directory of Community legislation in force is published in addition to the series L of the Official Journal. The Directory gives insight into all applicable legislation on a certain date, and is arranged by subject. The Repertory is published in two parts. Part I forms the main part in which the legislation is classified in 20 chapters according to subject. Part II contains the chronological and alphabetical index for the legislation included in part I.
Supplement	In the Supplement of the Official Journal (series S) invitations for public tenders are published for contracting-out work and delivery of goods and services. The Supplement is published twice a week on CDROM. Although it does not contain any registers, it does have a menu-based search tool.
Proceedings	The Proceedings contains the complete reports of the meetings of the European Parliament, see http://europe.eu.

Bulletin   The monthly Bulletin of the European Union gives an overview of the activities of the Commission and other EU-institutions. It is an important document as starting point for questions on recent developments. References are given to other publications of the EU, such as the Official Journal.

COM-documents   COM-documents are work documents of the Commission. In the COM-documents important initiatives and preparatory legislation on all policy areas of the Commission are published. The commission documents are numbered per year, e.g. COM(89)360. The number between brackets is the year, 360 is the serial number of the concerning year. A monthly register with quarterly and yearly accumulations is published. The register is divided up into work areas of the EU, with index on keywords and on document number. The register does not only pertain to the COM-documents but also the documents of the European Parliament and of the Economic and Social Committee.

                              Some other official publications of the European Union, which are in the EDC, are the law reports of the Court of Justice and annual reports of the EU. Next to information sources in writing, the EDC disposes of electronic sources with relation to the EU, such as the juridical database CELEX and bibliographic database SCAD.

More information   More information about the European Union can be found at http://europe.eu.

5.   Information management

Information management is important for an efficient search process. Thus, Information management may be summarized as:
- organization of the search process; i.e., the planning of the searching operations, obtaining and storing found information and process monitoring,
- processing of information with regards to content, including evaluation of the information, and
- keeping pace with the latest developments in the subject.

Personal documentation system

A separate element that is important with all these aspects is the personal documentation system. In this chapter the 'usual way' of information management is presented, which can be used for efficiently searching for information and processing it.

5.1   The organization of the search process

5.1.1   Planning

A rule of thumb for a good planning is based upon the 'one-third rule'. This means that for the following three activities, one-third of the total available time will be used:
- collect information,
- elaborate information,
- report.

This division seems rather unbalanced, but works very well in practice. However, since the time taken for reporting is often underestimated, an insufficient recording of the findings during the search process, often contributes to panic and stress around completion time.

5.1.2   Obtaining information and process monitoring

Obtaining information

In the first instance, you will be able to search the via referring sources, found by searching for documents in your University Library as long as they are not available via the Internet. Should the required document be lent out, then it is possible to reserve it from your work station via the catalogue. If the titles you are searching for are not available at the University of Twente, then it is possible to order them via the University Library for entry into the collection. If necessary, there is an emergency procedure. For more information, please talk to a library member staff.

Besides, there are the so-called Inter Library Loan facilities (abbreviated in Dutch: IBL). If you have a library card with a zip code you can via Picarta order from any location, as many documents (books and articles) as you like. You can also order documents using the SFX button that appears alongside the hits in database searches. If you have any problems, or need a document from abroad, then contact the University Library helpdesk for more information.

Dutch library and documentation guide

Of course you can also search the other library catalogues via the Dutch library and documentation guide. This guide includes data about more than 1800 libraries and documentation institutions in The Netherlands and provides an overview of all the national organizations.

Process monitoring

The monitoring the search process is important with a view to your planning, and any subsequent adjustments. Of up most importance, certainly with regard to larger projects, is to have continuous insight into the state of the process.
What is already in your possession and/or has been processed. What has been applied for? What about possible agreements with others? Does the evaluation of the process and the found information lead to satisfaction about the process or adjustment of your search activities, planning and the like.
With the assistance of your personal documentation system it is relatively simple to implement such a monitoring system (see 5.4).

5.1.3 Storing information

Storing found information demands special attention. In principle you want to store two cases: a description of the found information (reference) and the actual document (perhaps electronically). You should enter the descriptions or references in your personal documentation system, which have been designed in such a way that it is easy to generate literature lists, references and the like on behalf of the eventual reporting. You can add all kinds of information to these descriptions, i.e. a code for the location in your own filing system (boxes, (electronic) document files, etc.). Furthermore it is important to pay attention to the sequence and codification of the documents and boxes, document files, etc. You should pay attention to the following considerations:
- Storing printed documents according to a specific keyword (and subsequently on publication year) has the advantage that literature on a certain topic is grouped together. A drawback is that you may need to re-organize the documents now and then, since headings that contain too many documents will usually have to be split up.
- Storing documents simply by serial number has the advantage that you don't need to re-order the documents. A drawback is that literature about a subject is not grouped together.

5.2 Elaborating information

Elaborating information consists of both the evaluating and the processing of information. The evaluation and processing of information are closely related with each other.

5.2.1 Evaluating information

The value of found referring sources and informing documents determines the progress of your search process and also the quality of problem-solving. Is it worthwhile to spend more time (and money) on this found information? There are different criteria by which you can answer this question.

Criteria

First of all, there are content related criteria you can use on qualitative judgement of found information:
- is the argumentation good ?
- has the information been reproduced accurately and objectively ?
- is the text understandable ?
- have the facts and opinions been distinguished ?
- have illustrations, graphs and the like been entered ?
- is it based on decent empirical research and/or literature study ?
- has the same information already been found via other sources ?

Also, since some criteria can be used concerning the origin of information, it can be meaningful to work out the name of the author, publisher, journal or authors institute with a Citation Index. However, the danger here is the limited scientific circle in which the same authors continually refer to each other. An author who is not in that particular circle does not necessarily supply unscientific work.

More specifically, content pertaining to the quality criteria can be dealt with: is the information relevant ?
Finally, there are 'use' related criteria such as:
- relevance for the problem statement and for the target group
- availability
- do you have firm understanding of the language in which the information is presented?
- is the year of publication suitable ?

With the help of such considerations, you must decide whether or not certain information is important for your requirements.

5.2.2 Processing information

There are many ways of processing information that are often used subconsciously. It is informative to pay attention to the different methods of processing, and assess which one is the most suitable.

Paraphrasing, *summarizing and citing*	The processing of information consists of activities such as paraphrasing (described in other words), summarizing and citing. The latter is important, since exact references including page numbers are essential for cross-referencing your work.
Analyzing	For further processing, the information can be analyzed: split up in logical parts such as hypotheses, arguments, measurements and conclusions.
Synthesizing	After this, the information can be synthesized or integrated with aspects of your own problem statement. This often leads to the identification of important elements.
Interpreting	Of course, the interpretation (explanation) of information also plays an important role here. Closely related to interpretation is the evaluation of the information (also see 5.2.1).
	It should be clear that when processing a large amount of information the personal documentation system (see 5.4) can serve us well.

5.2.3   Plagiarism and copyright

Plagiarism	In academic circles it is a well-established convention to begin a publication with an extensive overview of previous work in the field. What is the difference between this and plagiarism? Are you allowed to present in your own words what you have read on a website or in a book? The Oxford English Dictionary defines plagiarism as "the action or practice of taking someone else's work, idea, etc., and passing it off as one's own". In scientific writing, authors must always state the basis for their assertions. This makes their work scientifically rigorous and verifiable. It is absolutely out of the question for an author to present another person's work as if it is their own: this is regarded as literary theft and a deliberate attempt to mislead. This principle applies to illustrations as well as to text and ideas. Many academic study programmes now make use of plagiarism scanners which automatically check students' work for plagiarism. Source references therefore play a major part in rules governing plagiarism (see 5.2.4 and 5.2.5).
copyright	However, there are also rules governing copyright. Even a correct source reference does not automatically enable the use of long sections of text or excerpts from films and books produced by others or their public distribution. In such cases, you have to ask for the permission of the copyright owner. This is usually the author or publisher of a book or article, or the "owner" of a website. The concept of "public distribution" is strictly applied. For example, even giving a copy of a music CD to a friend constitutes an infringement. There are exceptions to the general rule stated above, though these may vary from country to country. One universal principle is that no one can claim copyright on matters of common knowledge. In the Netherlands, the copyright owner's permission is not required in the following cases: • If the author has been dead for over 70 years. • If the copy is for personal use only. • Short text citations can be included and figures may be cited in full, with source references. However, citations may only make up a small part of

the work as a whole and must always be relevant to the content. For example, illustrations may not be appropriated in this way if they are purely being used for decorative purposes.

For an extensive account of the general Dutch rules, go to www.auteursrecht.nl. Rules for the application of copyright in Dutch higher education can be found on www.surf.nl/auteursrecht. Special arrangements have been made for academic readers/syllabuses in the Netherlands: see www.cedar.nl/pro/documenten/php for details.

5.2.4   References

References in the text

A description of another person's work and brief citations are therefore permitted, provided that these are clearly indicated by means of a source reference.
For example, Jansen (1983) wrote that ………..
There are a number of standard formats for references, such as the APA rules used in this document.

Short citation

In the literal rendering of text (citation), the section in question is placed between quotation marks ("") or is indented, often in a font different to the main body of the text. The page number is included in the reference: Jansen (1983, p.43) wrote that ……….. "……………".

Long citation

If a long section of text is rendered literally, the copyright owner's permission must first be obtained. This permission must also be stated in the text. For example, Jansen 1983, pp.43-44, reproduced by kind permission of the publisher) wrote as follows: "…………".

List of references

At the end of the document, an alphabetical list of references is included, containing standard descriptions of the sources used (see 5.2.5). Personal documentation systems such as Endnote automatically generate reference lists in all kinds of different styles on the basis of literature information compiled in a database.

5.2.5   Rules for list of references

Bibliographic data

Descriptions of sources play an important role in working with scientific information.
How can you make useful descriptions of sources efficiently? First of all, it is necessary to stipulate essential bibliographic data from the beginning of the search process. By doing this in a uniform and consistent way, it will save you a lot of time in the process of ordering the information and making the report ready for publishing.
For the descriptions of printed and electronic titles international rules have been designed which require very special punctuation. Examples are APA and MLA. In practice, these rules are neglected in many cases. Also, editors of journals often deal with their own rules for literature references.
Special software packages for personal documentation like Endnote or Reference Manager, can compose references according to many

different rules. The program uses data from its general database that has been built up by downloading descriptions from referring sources or typing in (copy/paste) the information by hand.

When you are not obliged to use special rules for references, it is not that important which rules you prefer. But it is important that you apply the rules consistently and that you provide a full description of your sources, such that a third party may easily refer to them. .

ISBN and ISSN

When referring to the official regulations, the so-called International Standard Book Number (ISBN) and International Standard Serials Number (ISSN) receive very little attention. Yet these are valuable pieces of publication data. With the help of these unique numbers, publications are easier to trace than a search by author or title. Books have an ISBN consisting of 10 figures. The first two figures indicate a country code; for example, The Netherlands has the country code 90. Journals and serial works have an ISSN consisting of 8 figures. Occasionally, a publication has both an ISSN and an ISBN. This is also true with works in a series, which can also be considered as an independent publication.

Digital Object Identifier

A digital object identifier (or DOI) is a unique and permanent means of identifying a file on the World Wide Web. Even if the internet address changes, the file will still be traceable, since this number is stored in a central databank. The system is managed by the International DOI Foundation.

Many articles made available electronically, in scientific journals for example, are given a DOI. For example: 10.1016/j.triboint.2009.12.058 In this string, "10.1016" stands for the databank. After / there is an abbreviation of the journal title, followed by the year and the article number where applicable. For instance, if you have access to the journal in question (e.g. through the University Library), you can gain immediate access to the full text of the article by using the following URL: http://dx.doi.org/10.1016/j.triboint.2009.12.058.

If you are using a source which has a DOI, this must be included in the list of references at the end of the description.

Elements of title descriptions

A title description for literature references is short and enumerates the characteristics of books and other documents according to fixed regulations. For example, the author's name is mentioned without entitling. When the author of a publication is not known this can be indicated by the term Anonymous. When a publication is written by more than two authors, only the name of the first author is mentioned followed by "et al.". In some cases the document has a corporative author, which means that an institute, company or other organization must be regarded as the author.

Furthermore the literature reference should contain both the main title and subtitle. If publisher, year or place of publication are not known this can be indicated by: no publ., no year or no place (or in Latin: s.n., s.a., s.l.) respectively. If a document is published by more than one publisher, or in more than one place, you only mention the first followed by "etc."

The elements of a literature reference differ according to the type of

document: books, contributions or papers in a composite work (i.e. congress proceedings) or journal articles.

Book

In the case of a book, the elements in APA style are:
- author(s)
- year of publication within ( )
- title *in italics*
- series, edition, translator and the like
- publisher, place of publication

For example:
Butterworth, I. (ed). (1998). *The Impact of Electronic Publishing on the Academic Community. An International Workshop organized by the Academia Europea and the Wenner-Gren Foundation.* Wenner-Gren International Series, vol. 73. Portland, London, etc.

Composite work

In these references often the "in"-reference is used. Moreover, for the sake of clarity, the title of the composite work is or in italics. In the case of a contribution or paper in a composite work, the APA elements of a title description are:
- author(s) of the contribution/paper
- year of publication witin ( )
- title of the contribution/paper
- In: editor, title of the composite work *in italics*
- series, edition,
- pages within ( )
- publisher, place of publication

For example:
Mackenzie Owen, J. (1998). Organizing for digital archiving: new distribution models in the scientific information chain. In: *The Impact of Electronic Publishing on the Academic Community.* I. Butterworth (ed). Wenner-Gren International Series, vol. 73. (pp. 128-133). Portland, London, etc..

Journal article

In the case of a journal article, the elements in APA style are:
- author(s)
- year of publication within ( )
- title of the article
- title of the journal *in italics*
- volume number, journal number within ( ), pages.
- DOI if available

Also in these title descriptions, for the sake of clarity, the title of the journal is in italics.

For example:
Speier, C. et al. (1999). Faculty Perceptions of Electronic Journals as

	Scholarly Communication: A Question of Prestige and Legitimacy. *Journal of the American Society for Information Science*, 50 (6), 537-543. DOI 8.0013/JASIS.1999.06.015
Electronic documents	References of electronic documents have basically the same regulations as for printed documents. However, in order to become to usable references, some information is added at the end of the description, such as • the statement 'published on CD ROM' • www-address followed by the date of entry • the organization that is responsible for the Web site • DOI When it is known that an electronic document is also published in print, then, if possible, also mention the necessary data to retrieve it in that form.

5.3 Keeping pace

	Certainly in long-term projects it is important to keep informed of new developments. There are many possibilities for this, which mostly demand a simple 'one-off' effort.
Attention profiles	As for keeping pace of publications, one can think of 'alerts' or 'attention profiles' in electronic databases. When these profiles have been drawn up and implemented, perhaps with the help of an information specialist, newly added items that are important to you will be highlighted. This may take the form of a monthly, electronically forwarded descriptions of publications for example.
RSS	You can simply enter your attention profiles into most referring sources, including scholarly publisher web sites e.g. in RSS way (for more information about RSS see www.wikipedia.org).
Citation Alert	The Science Direct database of Elsevier and Web of Science offer the possibility to install a 'Citation Alert' that draws your attention to new citations of a given article. So Citation Alert presents also a simple and effect method of keeping track of any citations to your own articles. If there is no possibility of entering an automatic alert in a referring source, then you can just store one or more search sentences. You may then frequently repeat your search without much typing work.
Contents	Contents of key journals in your field of interest are mostly very informative too. If a zip-code is coupled to your Univeristy of Twente library card, then you can program Picarta or Online Contents simply to e-mail you the complete contents of the journals that you are interested in.
Current research projects	The Dutch Research Database (abbrev. in Dutch: NOD) presents an overview of current research projects in The Netherlands. You can consult this database with the help of different search keys, such as keywords and names of persons and institutions.

Personal contacts   Naturally, traditional activities such as congress visits and maintaining personal contacts are also excellent to keep pace.
A lot of 'keeping pace' will only provide brief information, but there will also be information with which you want to do more. For this the personal documentation system is very useful.

5.4   Personal documentation system

When searching for information it is always important to enter all relevant search results in such a way that data are easily to be found again. A personal documentation and information system can be of great help. With the help of a flexible system, you can easily arrange the vast amount of found information, remembering that adjustments remain possible all the time. You can expand your system in the course of time and it can serve you well for each new process of the (literature) information search. In short, such a documentation system will make it an efficient process as possible.

An extensive personal documentation and information system contains information of all kinds. For instance you can store the "titles" (and/or literature references and URLs) of informing and referring sources. It also enables you to record the locations of these sources, the subject they refer to, whether or not you have ordered or consulted them already and to whom you have lent them, if relevant. It also enables you to note the names of organizations and individuals of importance to the research (including address details, of course), how the selected reference sources were searched and what the results of the search were.
If you also include extensive comments on the value of each informing source for your research, it is possible to generate a basic literature review from your own system with very little effort.

5.4.1   Set up of a personal documentation system

The set up of a personal documentation system demands a 'one-off' design effort, which eventually pays for itself in time saved. Therefore you will first have to determine a 'program of demands' by which you stipulate how you wish to use your personal documentation system. On the basis of this program of demands, you develop a 'format' that meets your demands.
It is therefore important to think ahead about which details you want to record and which details you want to use for searching/sorting.

There are many different software packages on the market for general database management, especially for personal documentation. On the Internet you can find various demo versions of such packages to download. Examples are Endnote, Reference Manager, Procite and BibTeX/LaTeX. These programs also offer standard formats. The choice of a package strongly depends on your personal demands and circumstances. If you work in a group it can be useful for reasons of interchangeability to work with the same package. Other packages are primarily geared towards use in combination with Word.

You may think that setting up a personal documentation system is time consuming. However, investment in a good system at the beginning of a search process will save you a lot of time in the overall process.

5.4.2   Strategy file

In principle, you need two files for each research project in order to record information: a strategy file and a results file.
The strategy file should include the referring sources in their broadest sense (possibly including people, institutions and the like) and state how you put them to use. For example, it could take the form of a database with the search history (search terms, number of hits and years searched) or the categories you consulted in a library. It can be very useful to have these details to hand for future reference, for instance if you think you have come across new search terms and you are not sure whether you have already used them or not. For researchers taking part in a series of studies, it is very valuable to know exactly how their predecessors went about their search.
In longer term projects, it is advisable to include more details:
- names of institutes where comparable research is taking place, including address, telephone number, e-mail etc.
- the same applies to persons:
  meetings, telephone conversations and recommendations given, complete with dates.

5.4.3   Results file

In addition, you will, of course, save the data on the informing sources you have found (books, articles) in a results file.
Each record in the results file should at least contain fields for all source data, as previously described in 5.2.5: e.g. author(s), title, publisher, year of publication and other data you regard as necessary. With regard to articles from journals or conference reports, you also need an "in" field, while for electronic publications you need fields for the medium, the URL and the date on which the source was consulted. Other relevant fields include:
- location: an indication for example of the library where the publication is located, with the shelf number or a serial number if you have ordered your own collection of publications by number
- abstract: a summary of the source content, if applicable
- classification codes: if you have compiled a wide range of literature, an existing classification can be used (e.g. the one used by your own library). See 3.1.1 for more information on classifications. However, if it concerns literature brought together for a specific purpose, it is advisable to make your own classification
- keywords: the comments on classification codes also apply to keywords
- source type: referring sources, informing sources
- own comments: stating how the source contributes to the research.

Process data

The data of all informing sources, which you find during the search process, must be firstly recorded in the results file. Furthermore, these records can contain a field with the so-called process codification, which indicate at what stage of the search process you are.

For example:

A	= applied for	O =	appeared to be of no importance
V	= processed	C =	cited, but not applied for
N	= not received	S =	important (star-)publication

Arranging results

With the help of the aforementioned data the results of the search process can be arranged according to many characteristics, for example:

- alphabetically according to author
- stage of the search processsource
- keywords and/or systematic codes
- finding place
- chronologically
- on provisional arrangement of the end-product

Data output

All of the systems enable you to produce lists of references in many different styles. If the system is compatible with your word processing program, the references in the text and the associated list of references can be produced with very little effort.

5.5   Checklist

With the help of the following checklist you can structure your search process. It will enable you to search more systematically and efficiently.

Problem statement

Have you formulated the problem statement accurately?

Is there a clear statement of

- purpose
- search terms to be used
- side conditions for the search process: time, money, language, ….

Have you familiarized yourself with the subject, if necessary, with the help of the informing sources?

Is there a clear recording of your sources?

- handbooks
- Internet
- encyclopaedias
- experts

Planning

Have you planned the search process?

Is there a record of

- the available time for searching, obtaining and processing/ reporting
- the sources you are going to consult.

Search

Did you follow the rules for systematic searching (see 3.2)?
Have you recorded during the search:

1. What referring sources have you consulted ?	• catalogues • databases • Internet search engines • citation indexes • bibliographies, abstract journals
2. What search keys have been used ?	• systematic codes • keywords • free text • author names • other search keys ….

3. Which combinations of search terms have been used

4. What the output is per source, how many hits.

5. Which relevant titles you found and where they are available

Information management	***Have you paid attention to the management of searching and processing*** • are you following the time plan ? • are you keeping score of the information acquisition ? • have the title and place descriptions been recorded in a personal documentation system (if necessary, with your own comments) ? • did you take appropriate steps in order to keep informed of any new developments (alerts, current contents, RSS) ?
Quality	***Do you have an opinion about the quality of the used referring and informing sources***

6. Searching for information within the University of Twente

For general information about the University Library Twente and its locations, including information about specific services such as courses, for the access to many catalogues and databases and for the latest news concerning the provision of scientific information in the University of Twente you can consult the Web site via www.utwente.nl, click on 'Library and Archive'.

Collections

The University Library of Twente contains approximately 400,000 books and audiovisual media in the Vrijhof Library. It is subscribed to many journals, most of them in electronic form. Next to this, the library has licenses for the use of many electronic databases by employees and students of the UT. You can find a survey of databases and e-journals on the Web site.

The greater part of the collection in print is accommodated in the reading rooms, the remaining printed publications are in the stack-room of the Vrijhof Library. The Vrijhof Library also contains more general works such as papers (not only in Dutch), magazines, encyclopaedias and atlases. You can also find basic literature for disciplines that are beyond the scope of the University of Twente, such as documents/articles on art and theology. Moreover the Vrijhof Library contains some special collections like DIN-standards and Dutch and European Government documents.

Borrowing, reserving, renewing and ordering

You can search the University of Twente's library catalogue in order to establish what literature is available in the University's library. You will find this catalogue via the homepage. Most books can be borrowed. After you have found the shelf-number in the library catalogue you can take the book from the shelf and have it registered for loan at the loan desk. Moreover, the catalogue offers the employees the possibility to request books for loan and to have them sent to the desired campus location within two days.

You can reserve 'lent out' literature without the need of library staff, and also renew borrowed literature (with the help of your borrower number and the first 3 letters of your surname). If a searched book or journal is not to be found in the catalogue of the University Twente's library there is the possibility to apply for it at other libraries via the SFX button (see below) or via the Picarta database .

Employees can do the latter from their work area via their Inter Library Loan (abbrev. in Dutch: IBL) -account (with the help of their borrower number and the awarded zip code); students can do this via the desk of the Vrijhof Library. You should contact the library desk if you cannot find a title in one of the aforementioned databases.

Finally, you can also make proposals at the loan desk or via an e- mail to your information specialist to purchase literature that is not yet available in the library collection. With the order you can indicate whether you wish to be the first borrower or not. If necessary, an urgent order can also be carried out.

Metalib with Combi-search
Since November 2005 the University Library has been offering a service called Combisearch.
Combisearch uses the Metalib technology, which enables you to find all of the university's available electronic resources and search a number of them simultaneously. If a source cannot be searched with the Combisearch interface, you will be able to use the native interface of that source.
Metalib also offers you the possibility to personalize your search environment by creating your preferred sets of sources, save and re-run your searches, and run alerts on your searches. In order to personalize your search environment you have to log on. Metalib uses the University's Single-Sign-On feature, so once you've been logged on to My Campus, it will use your personal ID to log you on automatically. If you prefer not to log on, you enter Metalib as a Guest user, which means that your personal settings will not be available. At any moment, however, a personal log on is possible.

SFX
Incorporated into Metalib is the SFX-service. It provides you with context-sensitive linking to the university's e-collections and services.
An SFX-button appears along with each retrieved reference from a growing number of resources like Web of Science, Scopus, Compendex, Inspec, Picarta, Google Scholar, etc.
SFX stands for 'Special Effects'. While the delivery of full text to the desktop offers a key service to researchers and students, SFX also enables users to use many additional services. For example, if the full text of an article is not available electronically, via SFX 'advanced' the user can check the library catalogue for print holdings or can perform a search with a web search engine like Google, or can look for articles by the same author in Web of Science. Also interlibrary loan is possible via SFX 'advanced'.

Information specialist
In case of questions during your search for information or the publication of scholarly material, you can ask for the support of an information specialist resident at the University's library. They specialize in searching the UT's databases for all disciplines, and they have a wealth of knowledge regarding journal quality and the repository procedures. For a survey of information specialists, see the Web site: 'Visiting us', 'Information specialists'.

7.   Scientific publication

Thanks to the Internet, research results can be published in different ways, often in the form of a journal article. This chapter presents three publication models (7.1) and provides some information on publishing via the University of Twente repository (7.2).

7.1   Publication models

In addition to publishing in a "traditional" and commercially published scientific journal available to subscribers only, in recent years an increasing number of articles are being made available on the Internet, free of charge. This type of publishing is known as Open Access. What is more, more and more possibilities are emerging which combine Open Access with the traditional subscription journal. This combination is known as the hybrid publication model.

Traditional publication model   In the traditional publication model or publisher's model, the researcher submits an article that he/she has written to the publisher of a scientific journal. After a period in which the publisher has the article peer-reviewed and the author adapts his/her article in accordance with the feedback, the final version of the article is approved and is published in the journal. But this process is not completed until the author signs a copyright agreement in which he/she transfers the rights to the publisher. The journal in which the article appears is made available to the reader by subscription.
Advantages and disadvantages
The traditional publication model has a number of advantages for the author, including:
•   a wide selection of journals in all kinds of areas
•   a journal reputation ("branding") based on thorough quality control (peer review)
The disadvantages are:
•   the fact that access to the article is restricted by price means that fewer people will read it
•   fewer readers means fewer citations
•   transferring copyright means that the author loses control of his/her own work

Open Access model
The Open Access model makes full-text articles open and freely available to everyone. Open Access originated among scientists and libraries:
• as a response to the significant rises in subscription prices which made access to scientific journals more restrictive and
• as a result of the opportunity that the Internet provides to make articles available to a wide audience without having to go through a publisher.

There are two forms of Open Access:

OA journals

Open Access journals: While the traditional scientific journal is based on readers' subscriptions, OA journals are available free of charge and the authors have to pay a certain amount (€3000 per article on average) for publication. See www.doaj.org for a list of OA journals.

Self-archiving

'Self-archiving': Many researchers publish full-text versions of their publications on their own website and/or their employer's website. In the latter case, they usually make use of a special archive: the institutional repository (see 7.2 for more information). Unlike OA journals, self-archiving does not cost the author anything.

Advantages and disadvantages
The Open Access model has the following advantages for the author:
- the article is published more quickly
- unrestricted access means that more people read the article
- more readers increases the likelihood of more citations
- retaining copyright gives the author maximum control over his/her own work

Depending on the form it takes, the Open Access model also has disadvantages for the author:
- In the case of OA journals, it is the author who pays (€2000-€4000)
- At present the selection of OA journals is limited and also restricted to a relatively small number of fields
- To date, few OA journals have a good academic reputation
- 'Rapid' publication of the first version of the article (preprint) via 'self-archiving': no option for citing a peer-reviewed version.

Hybrid publication model
The hybrid publication model combines the advantages of both the models described above.
There are two types of hybrid:

Parallel publishing

Parallel publishing: It is often possible to publish an article both in a journal via a publisher and by means of self-archiving. Many leading scientific publishers have adapted their terms and conditions of publication to the growing desire for Open Access. On the website Sherpa/RoMEO (http://www.sherpa.ac.uk/romeo/) you can find an extensive overview of the opportunities and conditions relating to parallel publication.

Open Choice

Open Choice: A growing number of publishers offer the opportunity - on payment of a certain amount - to have an article included in the journal once it has been accepted as an Open Access publication. The Sherpa/RoMEO website mentioned above also gives a list of the publishers/journals which offer this possibility.

7.2 The University of Twente repository

University of Twente repository
The University of Twente repository is the digital archive of university publications to which University of Twente staff have contributed as authors. It archives the full text of articles, conference contributions, dissertations, research reports, books and other publications and makes them freely available if possible. The University of Twente repository can be found on the University Library website (select 'UT publications' in the left-hand menu). The URL is http://doc.utwente.nl/. In addition to a search function, the website provides the opportunity to upload publications. Once it has been uploaded, the publication is stored in the University of Twente repository, supplied with bibliographical data and, following a check for any publication conditions stipulated by a publisher, access is determined.

Advantages of University of Twente repository
Uploading a publication and making it available through the University of Twente repository has the following advantages, in addition to those relating to self-archiving, and parallel publishing in particular:
• Since the University of Twente repository is part of a national and international network of similar digital archives, its publications are not only traceable through general search engines such as Google, but also using Google Scholar and specific search portals such as Narcis and Scientific Commons. This considerably increases the visibility of these publications.
• In the interests of ensuring lasting archiving and access, the publications from the University of Twente repository are automatically included in the e-Depot of the National Library of the Netherlands.

Table of publication models

	Traditional	Open Access		Hybrid	
		OA journals	Self-archiving	Parallel publishing	Open Choice
Access to article	Journal subscription	Free access	Free access	Journal subscription and free access	Free access
Peer review	yes	yes	no	yes	yes
Visibility (relative)	low	high	high	high	high
Speed of publication (relative)	low	low	high	high	low
Copyright	publisher	author	author	publisher/author*	author
Cost	paid by reader	paid by author (per article)	none	none	paid by author (per article)

* Conditions are set by the publisher and are laid down in the copyright agreement. For a survey of the options per publisher, consult www.**sherpa**.ac.uk/**romeo**.php.

Appendix to section 3.2.1 and 3.2.2 - Boolean operators *

Boolean operators
Search terms can be combined by means of Boolean operators (named after G. Boole, the founder of mathematical logic; it is also mentioned as logic operators) AND, OR and NOT. The application of this is illustrated with the so-called Venn-diagrams (see figure 1 and figure 2).

AND
With the combination A AND B, a search is undertaken that contains both references to the term A and term B.

search term	setnumber	number of found references
POLLUTION	SET 1	2833
NETHERLANDS	SET 2	369
BELGIUM	SET 3	225
POLLUTION AND NETHERLANDS	SET 4	9
POLLUTION AND BELGIUM	SET 5	15

Thus, using the AND-operator the search operation can be specified more precisely and the results are reduced.

Figure 1: Venn-diagrams 2-sets
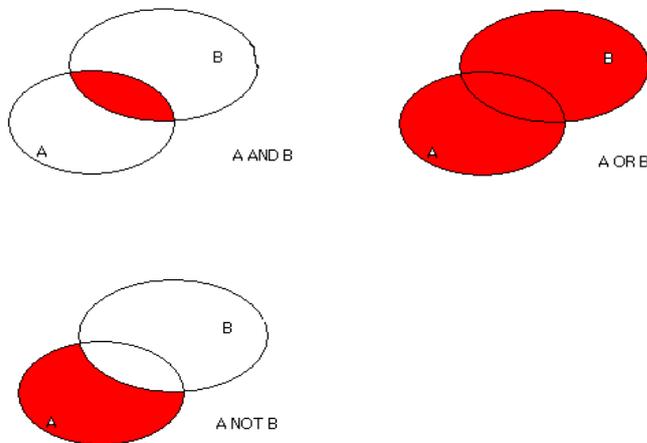
OR

With the combination A OR B, a search is undertaken that contains references to term A and/or B.

NETHERLANDS   SET 2   369
BELGIUM   SET 3   225

NETHERLANDS OR BELGIUM   SET 6   581

So, with the OR-operator a search result is extended. Note that the number of found references is less than the sum of set 2 and 3, because the 'double' references (overlap) are counted singularly.

NOT

With the combination A NOT B a search is undertaken for the references that contain term A but not term B. It should be noted, that when applying the NOT-relation the sequence of the terms is important. This is different to applying the AND- and the OR-relation. A NOT B gives another result than B NOT A.

When searching for the references that contain the word 'pollution', but are not written in the French language, the NOT-relation can be entered:

POLLUTION   SET 1   2833
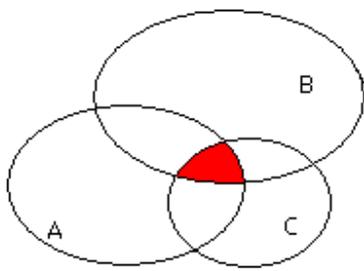POLLUTION NOT language=FRENCH   SET 7   2798

Comparing the number of found references with that of set 1 (2833) it appears that the NOT-operator specifies the search operation more precisely on the basis of negative criteria.

The Boolean operators can be applied with each other in all kinds of combinations. In specific search systems, the sequence in which they are completed is fixed (see the HELP function). Changes to the sequence can be obtained by adding parentheses. The combinations placed in parentheses are completed first. Examples of multiple combinations are (see also figure 2):
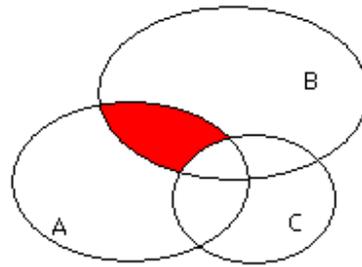
POLLUTION AND WATER AND NETHERLANDS
(NETHERLANDS OR BELGIUM) NOT language=FRENCH
POLLUTION AND (NETHERLANDS OR BELGIUM)

Omitting the parentheses in the last case (in a search system that simply applies the operators from left to right) would be equivalent to the combination NETHERLANDS OR (BELGIUM AND POLLUTION), which would contain next to references about 'pollution in the Netherlands' everything about 'Belgium' without necessarily dealing with pollution (compare the two diagrams at the bottom of Figure 2).
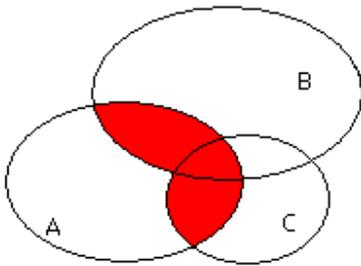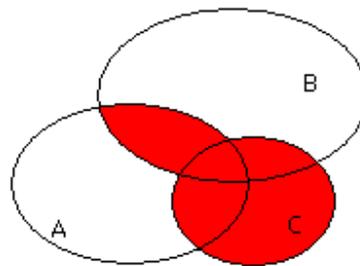
Figure 2: Venn-diagrams 3-sets