

Research Skills

Advanced MSc and PhD Students

<http://www.cs.bham.ac.uk/~jxb/rs.html>

Lecture 2

Literature Searching & Review

Andrea Soltoggio

<http://www.cs.bham.ac.uk/~ags>

Literature searching & review

The fundamental **starting point** of a research project

Scientific literature is the **source of knowledge** we deploy when working on a research project

Excellent **knowledge** in a field is a requirement when starting a research project

Literature searching & review

Knowledge

1. Gives authority to one's opinion/ideas/findings
2. Prevents from wasting time investigating something already known
3. Increases the probability of successful research by generating:
 - A good grasp/overview of a field
 - Foreseeing/guessing capabilities
 - Mastery in language and terminology

Searching : what can be found

- All relevant literature that covers the field of interest
- Taxonomy : how the field is organised
- Timeline of progress :
 - how quick an area has developed
 - how recent some concepts are
- Departments, laboratories, institutions, people, projects

Terms

Scatter

Bradford's law of documentation suggests:

- 90% of the important concepts of a subject in 10% of publications;
- the other 10% of the concepts of a subject is scattered in the remaining 90% of publications.

Relevance

How relevant / important is the content of a source with respect to the topic that is being reviewed

Forms of literature - books

- Often provide an understandable, well organised overview of the field
- Books usually written for money or promotion rather than *only* communication.

Currency

- Often quite delays in publishing - eg a year or more - *with no revisions* in between submitting manuscript and publication.

Forms of literature - journals

- Refereed journals - usually high status
- Unrefereed journals - usually up-market newsletters
- Indices
 - **Acceptance rate**
 - **Impact Factor (IF)**
example: IEEE Transaction on Evolutionary Computation, **impact factor** in 2006 : 3.77
(in 2002 was 1.48)

Currency

- Often long delays in publishing

Forms of literature - conferences

- Limited space, concise description of work
- Quicker reviewing process than journals, if accepted no compulsory changes are required
- Quality of papers depends on acceptance rate, authors

Currency

- Can be very up-to-date.
- The spoken version can be more up-to-date than the pre-printed version

Forms of literature - reports

- Anyone can produce a report and “publish” it by putting it on the WWW.
- Quality of reports depends entirely on authors

Currency

- Can be very up-to-date - but could just be an old and often rejected journal article

Forms of literature - patents

- Can be detailed, but not explanatory
- The purpose is not scientific communication, rather to protect the rights on an invention
- Often industry research produces more patents than university research

Currency

- Can contain rather recent research (6 months/1 year)

Forms of literature - theses

- Can contain a *survey* of a field
- Many factors (student's skills, supervisor, department) affect the quality

Currency

- At least 6 months old - may be 3½ years old or more

Forms of literature - WWW pages

- Anyone can publish anything.
- The source / web site usually tells the reliability and quality of the information
- Difficult to cite web pages as the content and location can change any time

Currency

- Could be very up-to-date indeed - although it is often difficult to tell

Search strategies

No rule / often subjective approaches prevail

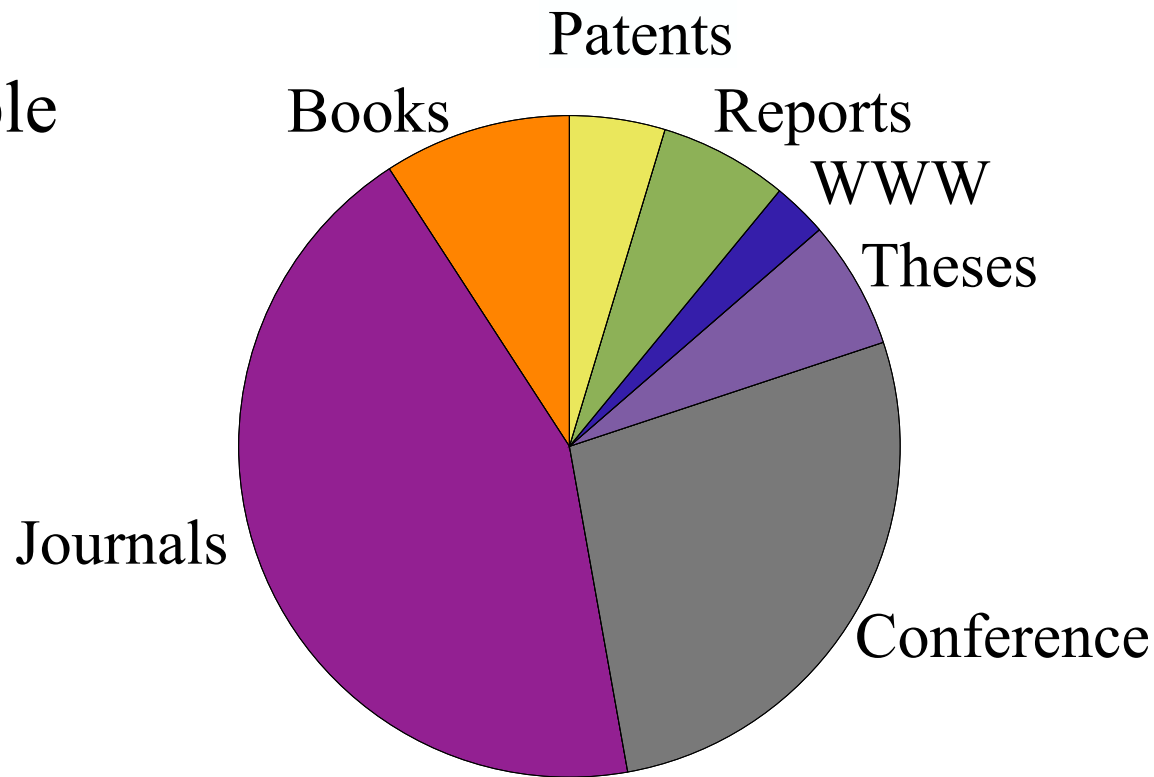
A number of **factors** determines the type and outcome of a literature search

- Sources of information considered
- Geographical coverage
- Languages
- Databases employed for search
- Use of keywords and incremental search
- Purpose of the search

Search strategies - sources

A good literature review includes all or nearly all sources

Example



Search strategies - geographical coverage

A good literature review has papers from many geographical areas and institutions

Problems with research in restricted geographical areas

- American / European / Japanese approaches and methodologies in some field can differ. Some search engines are biased
- Research in single institutes and laboratories can also be biased by a specific methodology
- Terminology can be also specific and misunderstood elsewhere

Search strategies - Languages

Most research papers nowadays are written in English

However, some important and useful exceptions exist

- Textbooks written by lectures are often in a national language other than English
- Research statements, technical reports and thesis outside English speaking countries are often written in the national language

Search strategies - Databases

There are a number of databases, catalogues, electronic search engines that can be very useful in search.

There is **NOT** a **universal or best** search engine

Often a systematic literature search is carried out using many search engines.

Examples are Citation Indexes, IEEE Xplore, Google scholar, CiteSeer, Index to theses

Search strategies - Keywords - 1

Keywords are meta-data used to retrieve information that is relevant to our interests

Identifications of keywords and search.

Problems:

- Researches use **different keywords** to name the **same or similar concepts**
- Keywords or terms name or are linked to **different concepts** for different researchers

Search strategies - Keywords - 2

Keywords are learnt during literature search and are effectively part of the outcome of the search

Solution: incremental approach to literature search.

Search queries are repeated in time considering the newly learnt keywords and terms (the process might take months)

Search strategies - purpose - 1

1. **Basic coverage** and overview of a field
 - Consider books, journal reviews, wikipedia pages
 - Focus on taxonomy, successful applications

2. **Deeper understanding** of specific topics or latest techniques
 - Also consider up-to-date journal and conference papers
 - Be prepared to deal with maths or difficult terminology

Search strategies - purpose - 2

3. Research focussed

- Make sure you carried out previous points 1 and 2
- Read future works in papers
- Link concepts from different fields (interdisciplinary approach)
- Identify research questions, areas poorly covered
- Identify precise key concepts from which to build your research

Summary

- Importance of literature search & review
- Main forms and characteristics of research dissemination
- Suggestions for a “good and systematic search strategy”, shortly:

Consider all sources, search on wide geographical areas and in more languages, use several search engines, update continuously keywords and terms, remember the purpose of the search

Literature searching & review

Next two lectures on

- Use of search engines: Science Citation Index, Google, CiteSeer...
- Example of using keywords, expanding search using different terminology
- Bibliography software and systematic search