MSc Research Skills

Topic: Finding & evaluating information

D G Rossiter

University of Twente.
Faculty of Geo-information Science & Earth Observation (ITC)

April 5, 2013

Copyright © 2007-2013 University of Twente, Faculty ITC.

All rights reserved. Reproduction and dissemination of the work as a whole (not parts) freely permitted if this original copyright notice is included. Sale or placement on a web site where payment must be made to access this document is strictly prohibited. To adapt or translate please contact the author (http://www.itc.nl/personal/rossiter).

Topics for this lecture

- 1. Information and information overload
- 2. Literature review
- 3. Types of sources and their reliability
- 4. How to search

Topic: Information and information overload

- Information
- · Information overload
- Information literacy
- Information skills
- Personal information infrastructure

Information

"Knowledge is power. The more one knows, the more one will be able to control events."

- Francis Bacon, *Meditationes Sacrae* (1597)

Information has become a vital resource for world economies, and drives research. Information is highly synergistic: the more information one has, the more information one can generate.

It is certainly the basic component of any research project.

Information overload

Information is available from many sources and in many formats, such as:

- printed text (newspapers, magazines, journals, reports . . .);
- television (increasingly international);
- videos;
- library databases; and
- · web sites.

Much of the printed material is also on the Internet.

- Question: How to avoid information overload?
- Answer: With a clear information strategy

Information skills

Information skills can be grouped in five categories:

- 1. awareness (that we need reliable information);
- 2. access (how to find it);
- 3. evaluation (how to decide what is most relevant and reliable);
- 4. **use** (how to use it in our work);
- 5. **generation** (how to produce new information).

Information skills for the MSc thesis proposal

- find relevant and accurate information on the research topic;
- · relate these in a literature review

Personal information infrastructure

Slowly you can build up:

- A list of your favourite reference materials such as handbooks, user's guides and texts;
- A list of your favourite journals and book series;
- A list of the best searchable databases for your purpose;
- · A set of your saved searches in the databases, also as e-mail or RSS alerts.

This is in fact an **infrastructure** for *your* research career.

Topic: Literature review in a research proposal

When proposing a research project:

- 1. to establish its originality
 - prove that the proposed work has not been done before;
- 2. to place the proposed research in **context** (related work);
 - how it fits in the "big picture"
 - related work done by others that influence the choices made for this research
- 3. to compare and justify the choice of research methods
 - methods used in similar studies
 - why some methods are preferred or appropriate in this research

Topic: Types of sources and their reliability

"Don't believe everything you read!"

Especially on the internet!

Not all sources are equally reliable, nor have they all had the same quality control.

Recommended: ITC library's **Information Literacy Course**¹ module 2 "Selecting the right information sources"

¹ http://www.itc.nl/Pub/Home/library/Library-Guides/LiteracyCourse

Journal Article

This is an original contribution that appears in a published scientific journal.

These contributions have been peer-reviewed to ensure quality control

Note: not all peer-review is equally effective.

In general, the more influential the journal (i.e. the more its work is cited and considered of top quality), the more likely that peer review has been rigorous.

Peer-review for quality control

- 1. The authors submit a draft of the article to a journal editor.
- 2. The editor checks the relevance for the journal and the format of the paper.
- 3. The **editor** sends the draft to several other scientists familiar with the subject matter; these are the **peer-reviewers**.
- 4. The reviewers read the draft and advise the editor to either:
 - (a) accept;
 - (b) accept with minor revisions;
 - (c) Reconsider if major changes are made; or
 - (d) reject outright.

Where is the quality control?

- 1. Recommendation to revise and possibly re-submit because:
 - Incorrect analysis, not suitable to the data
 - Unjustified conclusions, poor reasoning;
 - Work does not properly consider related work;
 - Poor writing.

2. Outright **rejection** because:

- Fraud, e.g. plagiarism of others' work
- Incorrect data collection or processing methods (therefore the data are not reliable);
- Work repeats what has already been done, nothing new is added to the existing literature;
- Work is too narrow ("light") to justify publishing, but could be incorporated into a bigger study.

Types of journal articles

Research Article Describes an original investigation, method, or procedure. Specific and limited

Dobos, E.; Micheli, E.; Baumgardner, M. F.; Biehl, L.; & Helt, T. 2000. *Use of combined digital elevation model and satellite radiometric data for regional soil mapping. Geoderma* **97**(3-4):367–391

Review Article Summarises a set of research articles; surveying the state-of-art in a particular field. The title may include words like "review", "summary", or "overview".

McBratney, A. B.; Odeh, I. O. A.; Bishop, T. F. A.; Dunbar, M. S.; & Shatar, T. M. 2000. *An overview of pedometric techniques for use in soil survey. Geoderma* **97**(3-4):293-327

Opinion A scientific **editorial**, either by the journal editor or an invited contributor.

Basher, L. R. 1997. *Is pedology dead and buried? Australian Journal of Soil Research* **35**:979-94

Which journals are most reliable?

There are a great many journals, and all may have useful and reliable information.

However, certain journals have been identified as the most **important** and **prestigious** journals within a subject field.

The MSc researcher should prefer articles out of these journals; it would be very unusual not to refer to any of these.

These are called **ISI** journals.

- · ISI: Institute for Scientific Information (Thompson Reuters)
- about 1/3 of all journals are ISI (about 3 700 in the sciences)

Criteria for selection as ISI journal

- the journal's publishing standards:
 - * selectivity;
 - * peer-review process.
- its editorial content (including reputation of its editorial board);
- the international diversity of its authorship;
- its citation record in other journals (importance of its papers).

Is a journal ISI?

Search the Journal Citation Reports (JCR).

Example, there are 15 ISI journals² in the "Remote Sensing" category (in alphabetic order):

- 1. Canadian Journal of Remote Sensing (Canadian Aeronautics & Space Institute)
- 2. GIScience & Remote Sensing (Bellwether Publishing, USA)
- 3. GPS Solutions (Springer)
- 4. IEEE Geoscience and Remote Sensing Letters (IEEE)
- 5. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (continued . . .)

²as of 12-April-2010

- 6. IEEE Transactions on Geoscience and Remote Sensing (IEEE)
- 7. International Journal of Applied Earth Observation and Geoinformation (Elsevier Science)
- 8. International Journal of Remote Sensing (Taylor & Francis)
- 9. ISPRS Journal of Photogrammetry and Remote Sensing (Elsevier Science for ISPRS)
- 10. Journal of Geodesy (Springer)
- 11. Photogrammetric Engineering and Remote Sensing (American Society of Photogrammetry and Remote Sensing)
- 12. Photogrammetric Record (Wiley-Blackwell)

(continued ...)

- 13. Photonirvachak (Indian Society of Remote Sensing)
- 14. Radio Science (American Geophysical Union)
- 15. Remote Sensing of Environment (Elsevier Science)
- 16. Survey Review (Commonwealth Association of Surveying and Land Economy)

Relevance to MSc project

ISI journals have been **expertly evaluated** as:

- reliable (especially their quality control and editorial policies);
- important (i.e. their papers are cited by other researchers).

so the MSc student (beginning researcher) has an easier job of **evaluating** papers - these have been **screened** (filtered, pre-selected).

Open Access

Open Access (OA) publications are those made freely available online to libraries and readers, anywhere, with no charges imposed for access.

- The "Green" Road to OA: repositories of papers published commercially but then made available (often after a time lag);
 - * Directory of Open Access Repositories (OpenDOAR):
 http://www.opendoar.org/
- The "Golden" Road to OA: Open Access journals: producer pays costs, papers are immediately OA
 - * Directory of Open Access Journals (DAOJ); http://www.doaj.org/
 - * Example: Hydrology and Earth System Sciences; http://www.hydrol-earth-syst-sci.net/

Conference Paper

- An original contribution that was presented at a scientific meeting;
- Usually not or minimally peer-reviewed, because . . .
- · ... conferences are mainly to discuss work in progress;
- Intensity of peer-review depends on quality of the conference;
- In most fields, papers presented at a conference should appear as journal papers theses within a few years of the conference;
- Some high-prestige conferences use peer review; the proceedings are published to the same standards as an edited book or a journal article

Conference papers in a journal

Conference papers may be collected for a **special issue** of a **peer-reviewed** journal; in that case they are peer-reviewed and cited as a journal paper, even though the work was first presented in a conference:

Conference paper:

Rossiter, D. G. 2006. Classification of urban and industrial soils in the World Reference Base for Soil Resources. In 18th World Congress of Soil Science. Philadelphia, PA (USA): IUSS

Journal article:

Rossiter, D. G. 2007. Classification of urban and industrial soils in the World Reference Base for Soil Resources. Journal of Soils and Sediments 7(2):96-100

There is no reason to cite the conference paper once the journal article is published.

Book chapter

This is an **original contribution** that is collected into an **edited book** on a specific topic.

These are typically invited by the book editor and may undergo some peer review; certainly they are edited. Often they are review articles.

Skidmore, A. K. 1999. *Accuracy assessment of spatial information*. In Stein, A.; Meer, F. v. d.; & Gorte, B. G. F. (eds.), *Spatial statistics for remote sensing*, pp. 197–209. Dordrecht: Kluwer Academic.

Textbook

This is a published book meant to introduce a subject for classroom teaching or self-study.

It can treat a topic at any **level** (i.e. pre-requisites for understanding it), but given that level, it is intended as the **first contact** with the subject.

Not peer-reviewed as such, but are typically extensively **edited** and sent by the publisher to people who might use the text in teaching, to see if they find the book accurate and useful.

Example Textbooks

Some say "text":

- Dupriez, H. & de Leener, P. 1998. *Trees and multistorey agriculture in Africa: a textbook for agro forestry*. Nivelles; Wageningen: Terres et Vie; Technical Centre for Agricultural and Rural Cooperation (CTA)
- Kutilek, M. & Nielsen, D. 1994. Soil hydrology: textbook for students of soil science, agriculture, forestry, geoecology, hydrology, geomorphology or other related disciplines.
 Geoecology paperback. Reiskirchen (D): Catena Verlag

Others don't:

- Lillesand, T. M. & Kiefer, R. W. 1994. *Remote sensing and image interpretation*. New York: John Wiley & Sons, 3rd edition
- Bishop, Y.; Fienberg, S.; & Holland, P. 1975. *Discrete multivariate analysis: theory and practice*. Cambridge, MA: MIT Press

Technical Report

These are publications from an institution or project, and often contain **primary** data and maps which do not appear elsewhere.

They are often difficult for others to obtain, but if they are the only source of information (e.g. of primary data), they should be cited.

They are not peer-reviewed; the quality control was only as good as the project.

- Center for Advanced Spatial Technologies (CAST). 1998. AR-GAP final report: State-wide biodiversity mapping for Arkansas. Report, Center For Advanced Spatial Technologies (CAST), Fayetteville, AR.
- Anonymous. 1985. *Soils and soil conditions, Kali Konto upper watershed, East Java*. Project Report ATA 206, Universitas Brawijaya (Malang), Agricultural University (Wageningen)

On-line access

More and more journals (also books and reports) are only available on-line, and have no printed equivalent.

Some of these are fully **peer-reviewed** and of good quality; they are used like printed journal articles.

Example:

Bourne, P. 2005. *Ten simple rules for getting published. PLoS Computational Biology* 1(5):e57. http://dx.doi.org/10.1371/journal.pcbi.0010057

Note the use of the **Digital Object Identifier (DOI)** system to help the reader locate the article on-line.

All peer-reviewed journal articles which are available on-lion have a permanent DOI.

Web access to printed sources

- · e.g. on-line versions of printed journals, technical reports, books.
- the web source is a copy or a differently-formatted version of a printed source.
- The web alternative is easier to access
- The printed source is cited, optionally with the web address
- Center for Advanced Spatial Technologies (CAST). 1998. AR-GAP final report: State-wide biodiversity mapping for Arkansas. Report, Center For Advanced Spatial Technologies (CAST), Fayetteville, AR. URL: http://www.cast.uark.edu/gap/
- Dobos, E.; Micheli, E.; Baumgardner, M. F.; Biehl, L.; & Helt, T. 2000. *Use of combined digital elevation model and satellite radiometric data for regional soil mapping. Geoderma* **97**(3-4):367-391 DOI: http://dx.doi.org/10.1016/S0016-7061(00)00046-X

Web pages

Some information is only available via the web.

Tomorrow's version may be different from today's, it may move to another cyber-address, or it may even disappear.

- not peer-reviewed;
- not permanent;
- not 'published' in the traditional sense, or even the digital equivalent

Example of a web page

"In the USA, soil conservation is aggressively promoted through attractive web sites [e.g. 1]."

Entry in the list of references:

[1] USDA Natural Resources Conservation Service. No date. *North Carolina NRCS*. On-line document; URL:

http://www.nc.nrcs.usda.gov/. Access date: 05-April-2013

The access date must be given, because digital sources may change, move or disappear.

Judging the reliability of web pages

Accuracy:

- * As far as you can evaluate, is the factual information correct?
- * Are there grammatical, spelling and/or typographical errors?

Authority:

- * Who sponsors the page?
- * Who wrote the material? What are the author's qualifications?
- Who is the copyright holder?

If you can't even tell who sponsors the page or who wrote the text, that is a bad sign: no accountability implies no authority.

```
(continued ...)
```

Objectivity:

- * Is the text an advertisement?
- Does the text seem to only present one side of a known controversy?
- Is the sponsor known to be objective?
- * Is there a reason the sponsor would distort or exaggerate?

Currency:

- When was the page written?
- * Is the information kept current?
- * When is the information from any graphs and charts gathered?

Coverage:

- Is it completed or under construction?
- * Is there a print equivalent (or at least some printed sources to back it up)?
- Is it an entire work or just a fragment?

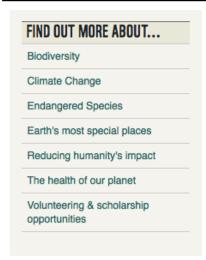
Reliable web pages

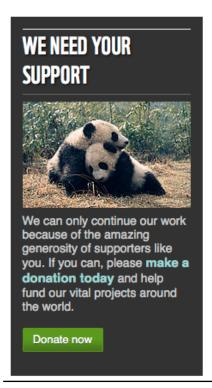
- Technical pages of responsible organizations, e.g. NASA http://www.nasa.gov/
- Official pages of organizations, when presenting facts within their competence,
 e.g. FAO http://www.fao.org/
- Collaborative sites (many small contributors) with community peer-review, e.g.
 CC-CEDICT (Chinese-English dictionary) http://cc-cedict.org/wiki/

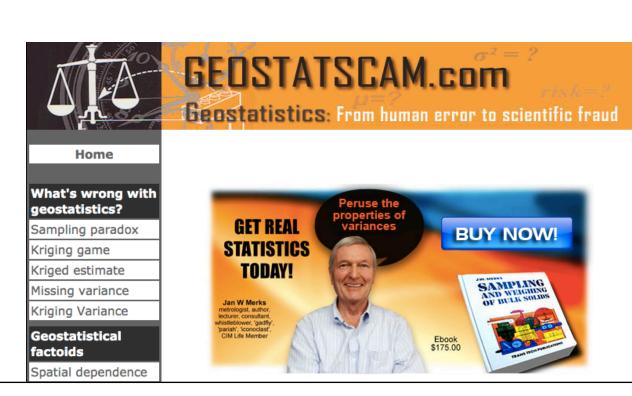
Wikipedia is a special case of a collaborative site: generally reliable because it requires a neutral point of view but this is difficult to enforce; there is no final authority.

Problematic web pages

- Commercial sites
 - obviously, promoting their product but they can be reliable sources for technical information about the product
- · Special-interest and advocacy groups, promoting a programme or viewpoint
 - May have valid information but often is one-sided ("slanted")
 - * May have purposely misleading or exaggerated information to stimulate activism or financial support
- Lecture notes or tutorials: check the credentials of the author (their peer-reviewed journal articles)
- "Hobby" sites: an amateur with enthusiasm for a subject but not necessarily sufficient knowledge







Topic: How to Search

Finding relevant material, and especially the most important for your purpose, is not easy.

It requires patience, detective skills, and continued hard work.

Fortunately, in the **digital age** it is possible to make much more rapid progress than previously. But, you should have a sound **search strategy**.

And, you have help from the information specialists in the ITC library.

Recommended: ITC library's **Information Literacy Course**³ module 3 "Searching the literature"

³http://www.itc.nl/Pub/Home/library/Library-Guides/LiteracyCourse

ITC library search resources

Recommended: ITC library's Information Literacy Course⁴ module 5 "Databases"

- Search and view **full-text** (if ITC has a subscription):
 - 1. Web of Science
 - 2. Science Direct / Elsevier
 - 3. SpringerLink
 - 4. JSTOR
- · Search abstracts, some links to full-text
 - 1. Scopus
 - 2. Google Scholar
- Subject-oriented bibliographic databases
 - 1. Geobase
 - 2. CAB Abstracts

⁴http://www.itc.nl/Pub/Home/library/Library-Guides/LiteracyCourse

Starting points for search

- Keyword searches in electronic resources (e.g. ScienceDirect)
- Reference lists in lecture notes;
- Reference lists in earlier theses;
- Reference lists in textbooks;
- Review papers
 - * Each reference is placed in **context**, with an indication of **importance**;
- · Reference lists in **reseach papers** in recent issues of relevant **journals**.

Search strategy

- build concept groups from your research topic;
- develop a set of terms for each concept group;
- find synonyms;
- decide which Boolean logic is needed, place brackets and combine concepts;
- decide whether to use truncation or not;
- · choose the databases and use proper search commands for each.

Example: Concept groups

From the ITC tutorial:

"These concepts are the main topics which best describe the information you seek. Analyze the topic you need information about by writing out a few detailed sentences about this topic. Underline the main words in these sentences."

Example: "Kampala has a severe problem with flooding after heavy rains, which is caused by rapid runoff from imperemable areas (paved roads, compacted soils) and insufficient infiltration in vegetated areas. I want to assess this spatially with a distributed hydrological model, for urban flooding which needs a gridded map of soil parameters, especially infiltration rate, over a large area. Has anyone made a soil properties map of Kampala? Has anyone developed a methodology for rapidly identifying compacted soils from imagery? What methods have been used to interpolate from point samples of soil infiltration to a grid?"

Example: Terms from concept groups

- 1. Kampala (specific area), African urban/city/citites (similar areas)
- 2. distributed hydrological/runoff/infiltration model for urban flooding
- 3. soil properties/hydrology map
- 4. identify/map compacted/impermeable/impervious soil from imagery
- 5. interpolate point to grid/raster

These may have synonymns, to be joined with OR (see next)

Boolean (logical) operators

• And operator:

- * Retrieves records that include both terms
- * narrows your search
- * Used for terms or concepts that are not related

Or operator:

- * Retrieves records that include either of the terms
- * Widens your search
- Used for related terms or concept

Not operator:

- Retrieves records that include one term but not another term
- Eliminates all the records containing the second term
- * Narrows your search
- * May eliminate relevant records

Truncation

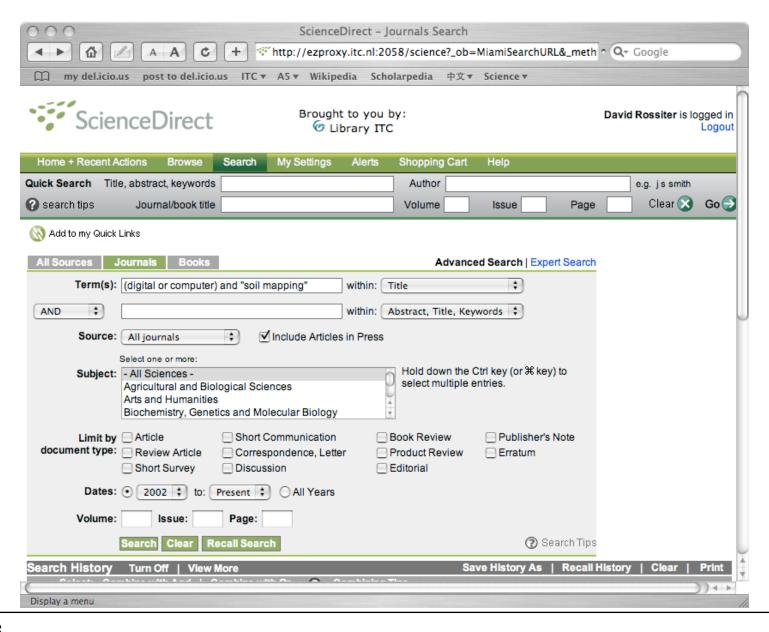
Replace part of a word with a symbol, the search engine will match anything for the missing part.

Truncation symbols can be used either inside the word or at the end of it.

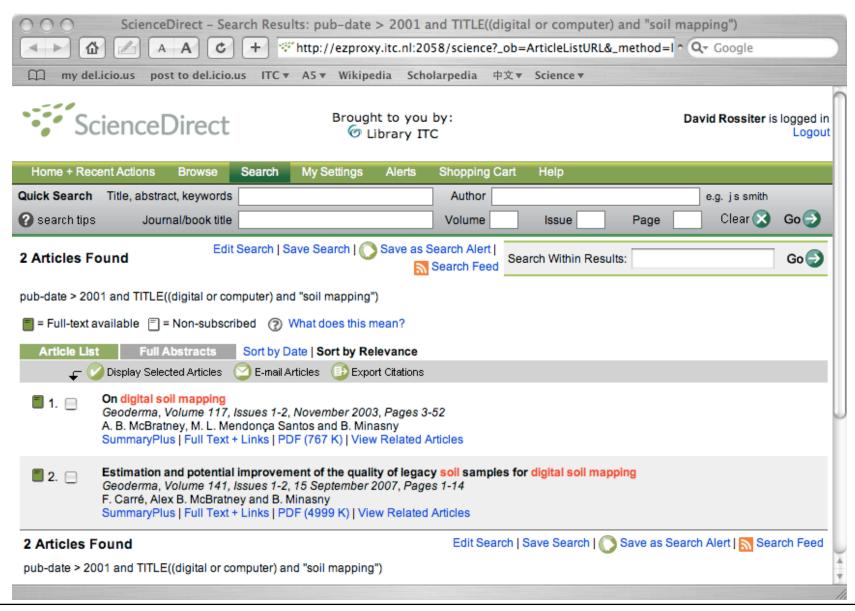
Each database uses its own truncation symbols; the asterisk * is most common.

- · 'wom*n' finds 'woman' or 'women'
- · 'diet*' finds 'diets', 'dietician', 'dieting', 'dietary', but also 'diethylstilbestrol' (a synthetic estrogen)
- · 'cartograph*' finds 'cartographic', 'cartography', 'cartographical'
- · 'plan*' finds 'plan', 'plans', 'planning' but also 'planet', 'planetary'
- 'system*' finds 'system', 'systems'
- 'science*' finds 'science', 'sciences'

Search in ScienceDirect



Search results

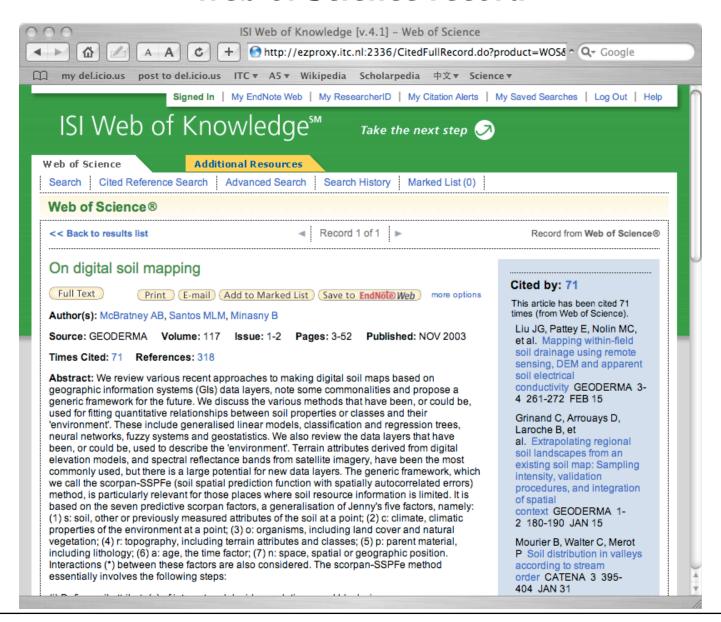


The "spider" approach

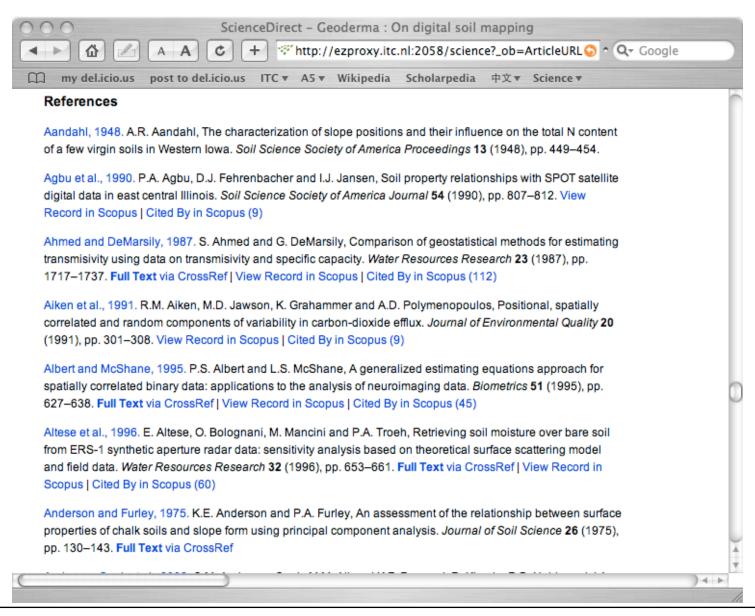
Once you have found some relevant literature, look for:

- Works that are cited in the papers you have found ("backwards spider"); note these have been put into context for you by the paper's authors
- Works that cite the papers you have found ("forwards spider")
 - * use the forward search of an electronic resource such as Web of Science
- Works by the same author(s)
- Related articles links in the search results ("sideways spider")
- Papers in the same journal

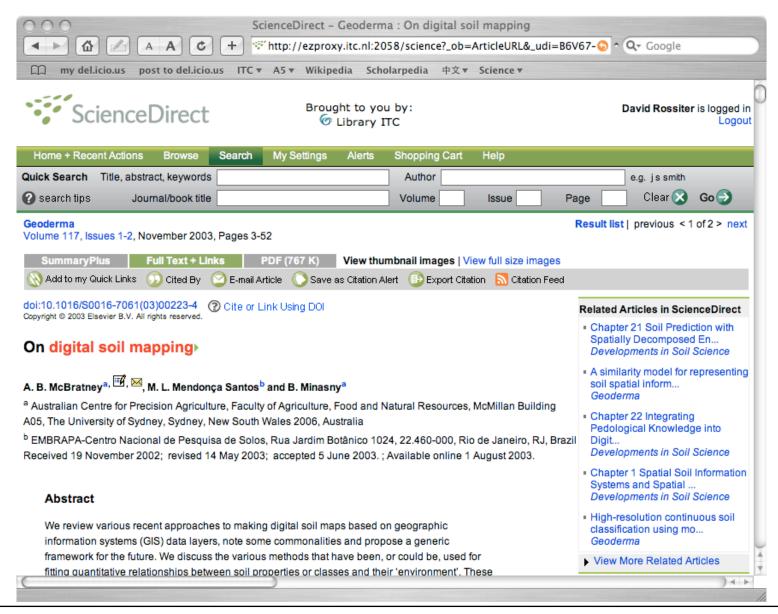
Web of Science record



References in the paper ("backwards spider")



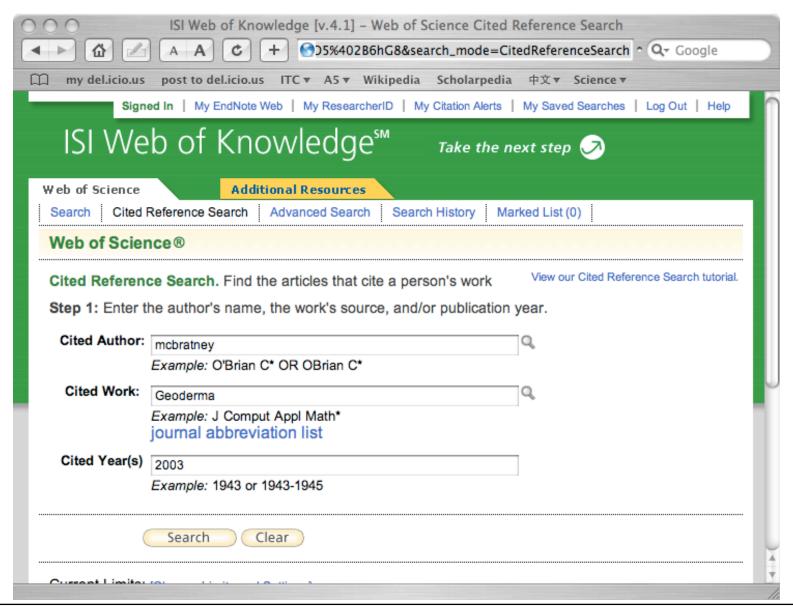
"Related articles" links ("sideways spider")



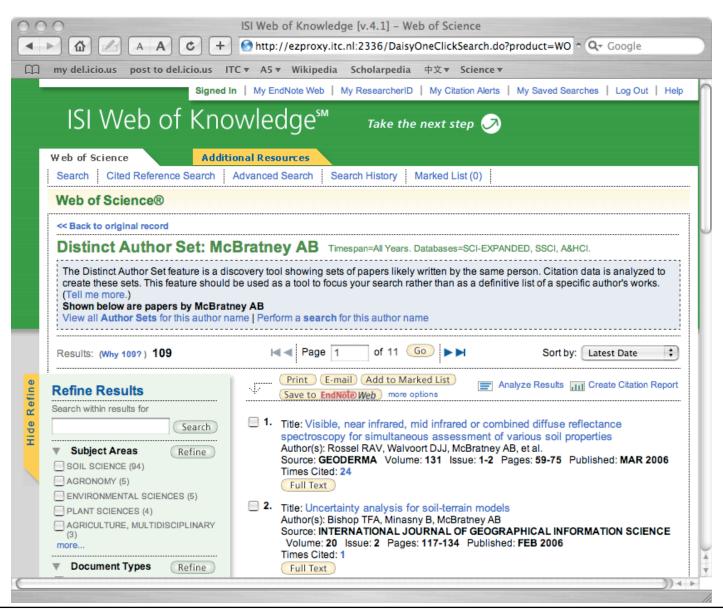
Web of Science "cited reference search"

- Purpose: discover how a known idea or innovation has been confirmed, applied, improved, extended, or corrected
- A feature of Web of Science (WoS)
- · Search for articles that have cited a **previously published** work
- · So, you don't re-do something already known

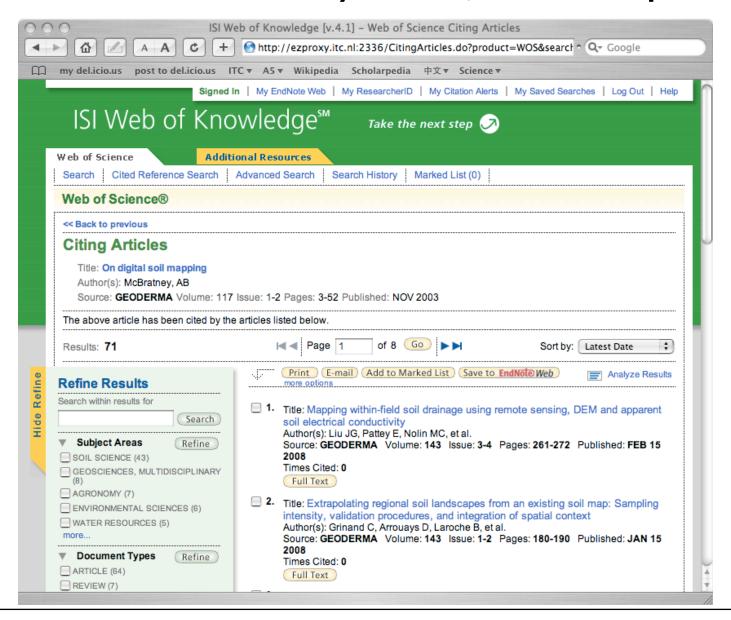
Finding works that cite a known reference



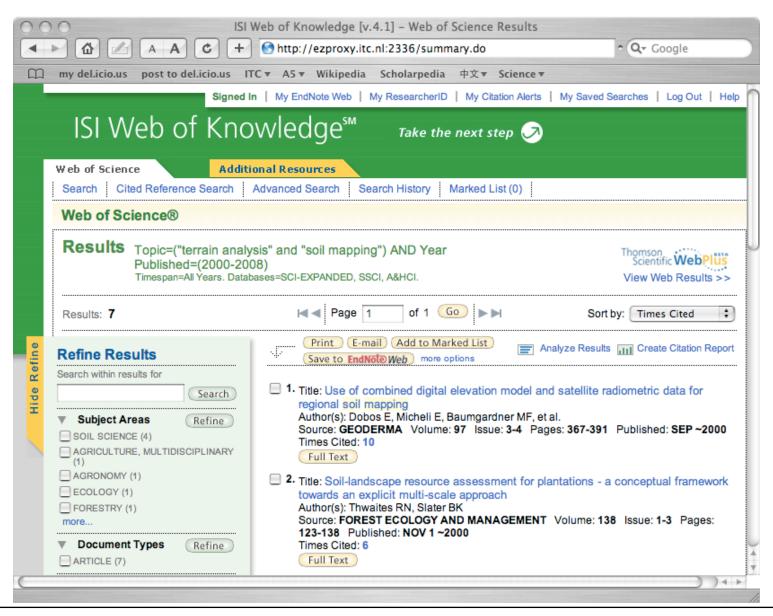
Web of Science: recent articles by the same author



Web of Science "Cited by" links ("forwards spider")



Web of Science: sort results by number of times cited



Evaluating the search results

Recommended: ITC library's **Information Literacy Course**⁵ module 6 "Evaluating the search results"

Don't just take the first search results you get! Ask: are there:

- Not enough references?: broaden the search, or use a spider approach to find more from the few good references you have
- Too many references?: narrow the search by adding search terms with 'and'
- Too specific references?: use some general terms, or use a spider approach to look for general references cited in these specific references
- Too general references?: use more specific terms, or look at the references cited in the general reference; these may be specific studies supporting the general ones.

⁵http://www.itc.nl/Pub/Home/library/Library-Guides/LiteracyCourse

Using Journal Citation Reports to evaluate journals

- At the level of MSc research, the student is advised to concentrate on the most important journals in the field.
- How to decide which are most important?
- Thomson Reuters (ISI) took the lead invented a journal impact factor: a ratio between the number of citations and recent citable items published in that journal.
 - * There are also an "immediacy index" (measures how quickly articles are cited) and "half-life" (measures how long articles are cited)
- · Used in an annual Journal Citation Report, available via Web of Science⁶

⁶http://www.itc.nl/Pub/Home/library/Search-for-information/Web_of_Science_link.html,

Journal Citation Report

Example: 2011, subject area "Remote sensing", sorted by impact factor.

Rank	Abbreviated	2011	Impact	Immediacy	2011	Cited
	Journal Title	Cites	Factor	Index	Items	Half-life
1	REMOTE SENS ENVIRON	18449	4.574	0.654	312	7.6
2	IEEE T GEOSCI REMOTE	16126	2.895	0.490	420	7.8
3	ISPRS J PHOTOGRAMM	1879	2.885	0.323	93	6.2
4	J GEODESY	1413	2.414	0.662	68	5.1
5	INT J APPL EARTH OBS	866	1.744	0.267	90	4.9
10	INT J REMOTE SENS	10865	1.117	0.164	535	8.9
14	PHOTOGRAMM ENG REM S	4056	1.048	0.098	82	>10.0

IJRS very many papers, long duration, but low impact and immediacy **PE&RS** few papers, low impact and immediacy, but used for a long time

Using citation counts to evaluate papers

- What makes a given paper 'important'? One obvious answer: if the paper is cited by many others, it helped science progress.
 - * For the MSc students, this indicates that the paper probably contains valuable information, and should be used by preference.
- A newly-published paper will not yet have any citations; the citation count isn't meaningful until about three years after a paper is published.
 - * But this may be a very good paper, at the "cutting edge" of the field.
- ScienceDirect: citation count with the "Cited in" link; only finds citations to other Elsevier journal articles
- ISI Web of Science: citation count with each paper's record. Search results can be sorted by citation count.

Example Web of Science result

TI=("remote sensing") AND TI=(soil) NOT TI=(moisture or hydrology or water); executed on 05-April-2013

1 Title: EXPLORING A V-I-S (VEGETATION-IMPERVIOUS SURFACE-SOIL) MODEL FOR URBAN ECOSYSTEM ANALYSIS THROUGH REMOTE-SENSING - COMPARATIVE ANATOMY FOR CITIES

Author(s): RIDD MK

Source: INTERNATIONAL JOURNAL OF REMOTE SENSING Volume: 16 Issue: 12 Pages: 2165-2185 Published: AUG

1995; Times Cited: 273

3 Title: Remote sensing of soil salinity: potentials and constraints

Author(s): Metternicht GI, Zinck JA

Source: REMOTE SENSING OF ENVIRONMENT Volume: 85 Issue: 1 Pages: 1-20

Published: APR 25 2003; Times Cited: 110

52 Title: Mapping within-field soil drainage using remote sensing, DEM and apparent soil electrical conductivity

Author(s): Liu JG, Pattey E, Nolin MC, et al.

Source: GEODERMA Volume: 143 Issue: 3-4 Pages: 261-272

Published: FEB 15 2008; Times Cited: 12

181 Title: Assessing land cover and soil quality by remote sensing and geographical information systems (GIS)

Author(s): Obade, Vincent de Paul; Lal, Rattan

Source: CATENA Volume: 104 Pages: 77-92 DOI: 10.1016/j.catena.2012.10.014

Published: MAY 2013; Times Cited: 0

308 Title: USING REMOTE-SENSING TECHNIQUE TO STUDY SOIL SEDIMENTATION FLOW

Author(s): KOLAWOLE MO

Source: ENVIRONMENTAL MANAGEMENT Volume: 17 Issue: 1 Pages: 73-81

Published: JAN-FEB 1993; Times Cited: 0

Evaluation of results by citations and publication date

- 1 classic paper (18 yr), highly cited → major contribution to science
- 3 newer paper (10 yr), relatively almost as highly cited
- 52 typical recent paper (5 yr) with few citations → modest contribution to science
- 181 new paper, no time to be cited
- 308 old paper (20 yr), never cited \rightarrow has not further contributed to science

Next steps

- 1. Follow library tutorials
- 2. Try search strategy for your own research problem
- 3. Develop personal information infrastructure