

Citation of Sources

CSE5230 Data Mining
Semester 2, 2004
CSSE Monash University

1 Introduction

This document presents an overview of the principles which should be observed when citing sources, published or otherwise, in both your literature review and your research paper. Additional and supporting information is available in the Faculty of Information Technology *Guide to Writing Assignments* (Faculty of Information Technology 1998) (hereafter referred to as the Faculty Guide). You should read the Faculty Guide as well as this document. It provides a great deal of additional useful information and guidance.

2 Citations: What and Why?

The Faculty Guide says of citations:

Citations are the references to others' work which you have used. They are placed in the body of your paper. Their purpose is to provide the reader with an intellectual 'audit trail' which can be followed back from your efforts to the sources you used. For this reason, citations and a matching list of references are necessary where a particular published or unpublished source has provided ideas you have used, *even if you have not used a direct quotation.* [emphasis added]

Correctly citing the sources that you have used is a *vital* part of academic writing. Just as important—if not more so—is the correct use of quotation marks “ ”, for short quotations, or indentation, for long quoted passages (as below), to indicate when text has been taken directly from one of the sources used. As the Faculty Guide says:

If you use *more than five words* taken directly and word for word from a published source, and do not indicate that the material is a quotation, you are guilty of **plagiarism**. This is the case even when you cite the source in the body of your text, or in your references, or both. Plagiarism of any extent is severely frowned upon, and where intent to gain an unfair advantage is evident, becomes **cheating**. Examiners are within their rights to award 0 for any work that entails plagiarism, and *cases of cheating can lead to a fail in the subject, or even exclusion.* [emphasis added]

These are not idle threats. The Faculty and the University regard plagiarism and cheating as very serious offences. We want to help you to get it right, but you must take responsibility for your work, and make use of the resources provided and to which you are directed.

2.1 Why Are Citations Required?

All academic writing includes and requires citations. Citations show readers where ideas have come from, and make it easy for them to find the original source if they want to know more about it. They

are also an indication of the author's scholarship: the appearance of many citations can indicate that the author has read widely and is familiar with the state-of-the-art in his/her field of research.

All research builds on earlier work. Sir Isaac Newton captured this notion beautifully in the sentence "If I have seen further, it is by standing on [the] shoulders of giants" (Newton 1676). Citing earlier research makes the context of new research clear. You will find that all the academic papers you have located for your literature review cite earlier work for this reason, even if they introduce completely new ideas. Non-academic sources, such as industry 'white papers' often do not cite their sources, which means that they are of very limited value.

In a literature review, such as the one that you are writing for this unit, new research is not expected. You are expected to read a variety of publications relevant to your topic, to describe their contents, and then to discuss them, comparing the various approaches to problems and discussing their advantages and disadvantages. Since no original research is expected, it is reasonable to expect that almost every paragraph will contain one or more citations. The only likely exceptions to this are the introduction and/or the conclusion, which may consist of general discussion of the problem.

Some students seem to believe that placing many citations in their work is bad, since (they think that) it makes it look as if they have not done anything original themselves. This is *not* the case! The presence of many citations is *good*, since it indicates that you have read widely, and that you understand the importance of acknowledging sources in academic writing.

2.2 Paraphrasing and Quotation

Although the presence of many citations is good, it is *extremely* important that you write about the articles which you have read *in your own words*; this is called *paraphrasing* and is normal academic practice (see section 5 of the Faculty Guide). Whenever you take words *directly* from a source—*i.e.* without rewriting in your own words—you are *quoting* that source. Some quotation in a research paper is fine: sometimes an author seems to have found exactly the right words (as with Newton (1676) above). Quotations *must* be either enclosed in quotation marks or shown as an indented passage. Remember, however, that a paper consisting entirely or largely of quotations will fail: your lecturer or tutor is interested in *your* work, not simply the words of others pasted together.

3 Information Required for Citations

Each citation in the text of your paper must correspond to an entry in the 'References' section at the end. Remember that one of the primary functions of a citation is to allow the reader to find the source which you used. Consequently the entries in the 'References' section must contain enough information for this to be possible. The main questions you must answer are:

- **Who** wrote this?
- **What** is it called?
- **Where** was it published?¹

¹Not where did you find it!

- **When** was it published?

Below is a list, taken from Lamport (1994), of possible `BIBTEX` entry types and the required and optional information that must be supplied when referring to a source of that type. Even if you do not use `LATEX` and `BIBTEX` this is an excellent guide to the information that should be provided when citing various kinds of publications.

article An article from a journal or magazine.

Required information: author, title, journal, year

Optional information: volume, number, pages, month, note, key, url

book A book with an explicit publisher.

Required information: author or editor, title, publisher, year

Optional information: volume, series, address, edition, month, note, key, url

booklet A work that is printed and bound, but without a named publisher or sponsoring institution.

Required information: title

Optional information: author, howpublished, address, month, year, note, key, url

inproceedings (or **conference**) An article in a conference proceedings.

Required information: author, title, booktitle, year

Optional information: editor, pages, organization, publisher, address, month, note, key, url

inbook A part of a book, which may be a chapter (or section or whatever) and/or a range of pages.

Required information: author or editor, title, chapter and/or pages, publisher, year

Optional information: volume, series, address, edition, month, note, key, url

incollection A part of a book having its own title.

Required information: author, title, booktitle, year

Optional information: editor, pages, organization, publisher, address, month, note, key, url

manual Technical documentation.

Required information: title

Optional information: author, organization, address, edition, month, year, note, key, url

phdthesis (or **mastersthesis**) A Ph.D. or Master's thesis.

Required information: author, title, school, year

Optional information: address, month, note, key, url

proceedings The proceedings of a conference.

Required information: title, year

Optional information: editor, publisher, organization, address, month, note, key, url

techreport A report published by a school or other institution, usually numbered within a series.

Required information: author, title, institution, year

Optional information: type, number, address, month, note, key, url

unpublished A document having an author and title, but not formally published.

Required information: author, title, note

Optional information: month, year, key, url

misc Use this type when nothing else fits.

Required information: howpublished (note really required, but strongly recommended)

Optional information: author, title, month, year, note, key, url

Note that the ‘required information’ above indicates the *minimum* required. You should supply as much information as you can. The ‘References’ section at the end of this document gives examples of how an journal article (Borges & Levene 2000), a conference article (Agrawal & Srikant 1995), a book (Westphal & Blaxton 1998) and an on-line industry or government white paper should be cited (Houle, Cadigan, Henry, Pinnamaneni & Lundahl 2000).

3.1 Citing On-line Sources

Many students find sources for their theses on-line. This is fine, but it is important to be careful when citing sources found on-line, for two reasons:

- often the on-line version is not the primary medium in which the article was published (especially for journal and conference articles). You must find out where the article was actually published and cite those details. Publication in a conference or journal generally implies peer review and thus quality control. Anyone, however, can put up a web page;
- URLs are subject to frequent change. Often documents disappear from the web, or are moved to a new URL.

The second point above means that it is **not acceptable** simply to give a URL for an on-line document. First you must try to find out if it has been published elsewhere (almost all academic articles have been), and if you cannot find any such information, you must give at least the author(s) and title of the document, as well as the date on which it was downloaded. See the Faculty Guide for directions on citing articles with no apparent author.

4 Tools to Help Manage Citations

Whilst it is reasonably easy to manage citations and references for a small essay or paper manually, it becomes much more difficult when writing large papers, theses or books. Authors of such works use tools to help them with this task.

The document preparation system used by many scientists, particularly in computer science, physics and mathematics, is \LaTeX . It has a bibliography management system called $\text{BIB}\TeX$, and is also excellent for managing large documents and typesetting mathematical formulae. \LaTeX is available on most, if not all, UNIX systems (including Linux). It is also available for MS Windows. There is further information available on-line at:

<http://www.csse.monash.edu.au/software/latex/>.

If you use MS Windows software, there is a tool for MS Word and Word Perfect called *EndNote*. It is available free to Monash staff and students. Details on how to obtain and use EndNote are available on-line at:

<http://www.lib.monash.edu.au/v1/endnote/endncon.htm>

References

- Agrawal, R. & Srikant, R. (1995). Mining sequential patterns, in P. S. Yu & A. L. P. Chen (eds), *Proceedings of the 11th International Conference on Data Engineering*, Taipei, Taiwan, pp. 3–14.
- Borges, J. & Levene, M. (2000). A fine grained heuristic to capture web navigation patterns, *ACM SIGKDD Explorations* 2(1): 40–50.
- Faculty of Information Technology (1998). Guide to writing assignments, Faculty of Information Technology, Monash University, Australia; On-line resource.
<http://www.csse.monash.edu.au/~ajh/adt/studentguide/studentguide.html>
- Houle, J. L., Cadigan, W., Henry, S., Pinnamaneni, A. & Lundahl, S. (2000). Database mining in the human genome initiative, Bio-databases.com white paper, Amita Corporation, 1420 Blair Place, Suite 500, Ottawa, Ontario, Canada, K1J 9L8.
<http://www.biodatabases.com/whitepaper01.html>
- Lamport, L. (1994). *L^AT_EX: a document preparation system*, second edn, Addison-Wesley Publishing Company, Inc., Reading, MA, USA.
- Newton, I. (1676). Letter to Hooke.
- Westphal, C. R. & Blaxton, T. A. (1998). *Data Mining Solutions: Methods and Tools for Solving Real-World Problems*, John Wiley & Sons, New York, USA.