

An Attempt to Map Information Literacy Skills via Citation Analysis of Final Year Project Reports

N.N. Edzan

Library and Information Science Unit
Faculty of Computer Science and Information Technology
University of Malaya, 50603 Kuala Lumpur
Email: edzan@um.edu.my

Abstract

Throughout the years, librarians have been using citation analysis to determine the usage of their collection for their collection development. Others have used citation analysis to look at undergraduate information behaviour. Various attempts are now being made to relate citation analysis of bibliographies to information literacy competencies by mapping them to the performance indicators of established information literacy standards. Within this context, the main aim of this paper is to analyse the bibliography of the final year project reports produced by the undergraduates from the Faculty of Computer Science and Information Technology, University of Malaya. A total of 73 reports were analysed using a pre-designed scoring sheet and results presented include the number of pages, number of citations, types of sources used, usage of Web resources, currency of sources and citation style. The contents analysis of the bibliographies indicates the following: (a) the least number of citations per report is 6 and the most is 165 with the most number of citations within the range of 11 to 20 cites; (b) there are more Web citations than citations to books, journal articles, undergraduate reports, Masters' dissertations and conference papers; (c) there are more citation to .com than to .org, .edu, .net and other URL extensions; (d) most citations are not dated and most of those dated are from within the last three years with the most current being 2005 and the oldest dated citation is 1935; and (e) most references have their print citations cited correctly but the Web citations cited incorrectly. When mapped to the performance indicators of the ALA/ACRL/STS 2005 Information Literacy Standards for Science and Engineering/Technology, only a handful of indicators could be matched to the information literacy performance indicators. Therefore, bearing this in mind, librarians and academics must make a more conscious effort to play a more important role in imparting information literacy skills to their students' since it has been identified as a necessary lifelong skill.

Keywords: Citation analysis; Information literacy; Final year project reports; Performance indicators.

1. Introduction

Leiding (2005) defines a citation as "a bibliographical entry in a footnote, reference list, or bibliography of a document that contains enough information (for example, author, title, publisher, or journal title) to verify the original item". Librarians have been using citation analysis to determine usage of their collection for collection development purposes. (Leiding, 2005). Citation analysis is a subdivision of citation studies, which was defined by Mosher (1984 cited in Leiding 2005) as being "any specific methodologies that use source citations or references drawn from the scholarly apparatus of articles and books as the basis for manipulation, research, and study." Researchers like Magrill and St Clair (1990), Davis and Cohen (2001), Davis (2002, 2003), and Heller-Ross (2002) have used citation analysis to look at undergraduate information behaviour. Attempts are also being made to relate citation analysis of bibliographies to information literacy competencies.

2. Methodology

Methodologies used in thirteen similar studies were reviewed and fifteen studies were analysed to determine the categorisation of citations used. The *Information Literacy Standards for Science and Engineering/Technology*, which was proposed by the

ALA/ACRL/STS Task Force on Information Literacy for Science and Technology in 2004, was used as the standard to which the information literacy skills will be mapped to. As for the project reports, all the reports produced by the final year undergraduate students from the Faculty are kept in the Faculty's Library. Using reports from the final year undergraduates, 293 reports were identified as the population and every fifth report on the shelf was taken for analysis with a total of 73 reports analysed. There is a tendency for the students to use the word "Reference" to refer to the list of items that they refer to in their report. On the other hand, the word "Bibliography" is used to refer to items that they refer to but are not used within the report. Within this context, only the reference and/or bibliography (if available) were analysed using a pre-designed scoring sheet.

The following were observed in the study:

- a) Level of analysis: bibliographies of project reports.
- b) Number of concepts to code: citations were coded based on a combination of typologies used by researchers such as Hovde (2000), Davis and Cohen (2000, 2001, 2003) and Leiding (2005). The typology used are Books, Journals, Magazines, Newspapers, Undergraduate project reports, Postgraduate thesis and dissertation, Conference proceedings, Web and Unidentifiable.
- c) Decision on coding: coding were done on the citations to determine the number and currency of each type of source
- d) Distinguishing concepts need to be established so as to avoid ambiguity: a clear distinction of the different categories of sources (Table 1) were made using the criteria of categorisation used by Davis and Cohen (2001), Smith (2003) and Mohler (2005). Print items were either coded as such or coded as "Web" depending on whether the students had stated how they had accessed the source. No effort was made to check for accuracy or persistence of the Internet citations. Each Internet citation was taken as true and it was assumed that the URL given will lead directly to the cited document.
- e) Coding rules will have to be established: scoring for the number of citations, variety of sources and the number of citations per source is as listed in Table 2.
- f) Irrelevant information: any incomplete or irrelevant information were categorised as "unidentifiable".
- g) Coding the bibliography involved the use of a coding sheet for each project report.
- h) Analysis of the results was done on completion of coding of at most, the bibliographies of at least 20% of the 2004/2005 final year project reports.

Table 1: Categorisation for Resources

Category	Scholarly / Non-scholarly	Criteria for Categorisation
Books	Scholarly	
Journals	Scholarly	Scholarly periodical that contains a report of primary research.
Magazines	Non-scholarly	Non-scholarly periodical that reports news, industry information and events
Newspapers	Non-scholarly	
Project reports	Scholarly	
Dissertation	Scholarly	
Conference papers	Scholarly	
Websites		Official, professional and educational resources whose domain names end in .edu, & .gov
Unidentifiable		Resources with insufficient information to fit into any other category

Table 2: Scoring for Bibliography

Attributes	Scoring		
Total number of citations	1-5 = 2 16 - 20 = 8	6-10 = 4	11-15 = 6 > 20 = 10
Number of citations per source			
Book	1 - 3 = 3	4 - 6 = 6	> 6 = 10
Journal	1 - 3 = 3	4 - 6 = 6	> 6 = 10
Magazine	1 - 3 = 10	4 - 6 = 6	> 6 = 3
Newspaper	1 - 3 = 10	4 - 6 = 6	> 6 = 3
Undergraduate project reports	1 - 3 = 3	4 - 6 = 6	> 6 = 10
Masters dissertations	1 - 3 = 3	4 - 6 = 6	> 6 = 10
Conference papers	1 - 3 = 3	4 - 6 = 6	> 6 = 10
Web - scholarly	1 - 3 = 3	4 - 6 = 6	> 6 = 10
Unidentifiable	0		
Time frame	< 3 years = 5 10 - 12 years = 2 4 - 6 years = 4 > 12 years = 1 7 - 9 years = 3		
Citation style	Consistent = 5 Inconsistent = 0		
Total	120 points		

3. Findings and Discussion

(a) Number of Citations

The total number of citations is 2,184 with the least being six cites and the most, 165 cites. The average number of citation per report is 29.9 cites (Table 3).

Table 3: Number of Citations (n=93)

Range of Citations	Frequency	%
1-10	7	9.6
11-20	26	35.6
21-30	16	21.9
26-30	9	12.3
31-40	7	9.6
41-50	2	2.7
51-100	4	5.5
>100	2	2.7
Total	73	100.0

Performance indicator 1.1. specifies the need to define and articulate the need for information. The presence of a reference list at the back of every project report shows that the students do have a need for information and this need is satisfied through the use of various information sources which finally appear as citations in their reports. Even though the number of citations may be as little as six or as many as 165, its presence denotes the need for information. As long as the information need is fulfilled by a certain number of information sources, then the need is deemed as being satisfied.

(b) Types of Sources

The identification and tallying of citations was a straightforward process since all the citations are easily identified. Table 4 shows the breakdown in numbers and percentages of the format of works cited in the students' reference list. Web citations are present in all reports with a minimum of two cites appearing in a report and a maximum of 148 cites appearing in another.

The dependence on the Web reinforces findings from the previous surveys of students doing the final year project and lectures supervising them. There are several possible explanations which would require verification through focus interviews with the respondents. It is possible that they use these reports as "report writing guides" only and do not use the information contained in them. It is also possible that they do use the information contained in the reports verbatim and avoided citing them for fear of being caught plagiarising.

Table 4: Distribution of Citations by Category (n=3184)

Format	Minimum	Maximum	Mean	Frequency	%
Web	2	148	20.37	1487	68.1
Book	0	21	6.40	467	21.4
Journal article	0	22	1.75	128	5.9
Conference paper	0	25	.89	65	2.9
Undergraduate report	0	4	.42	31	1.4
Masters' dissertation	0	2	.08	6	0.3
Total				2184	100.0

Table 5: Range of Citations by Category (n=3184)

Range	Book Freq (%)	JA Freq (%)	UR Freq (%)	MD Freq (%)	CP Freq (%)	Web Freq (%)
0	4 (5.5)	50 (68.5)	59 (80.8)	70 (95.9)	61 (83.6)	0 (0.0)
1-5	32 (43.8)	13 (17.8)	14 (19.2)	3 (4.1)	9 (12.3)	10 (13.7)
6-10	28 (38.4)	7 (9.6)	0 (0.0)	0 (0.0)	1 (1.4)	19 (26.0)
11-15	6 (8.2)	2 (2.7)	0 (0.0)	0 (0.0)	1 (1.4)	11 (15.1)
16-20	2 (2.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	9 (12.3)
21-25	1 (1.4)	1 (1.4)	0 (0.0)	0 (0.0)	1 (1.4)	8 (11.0)
26-30	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	5 (6.8)
31-35	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	4 (5.5)
>35	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	7 (9.6)
Total	73 (100.0)	73 (100.0)	73 (100.0)	73 (100.0)	73 (100.0)	73 (100.0)

JA = journal article UR = undergraduate report MD = Masters' dissertation
CP = conference paper

Table 5 shows the numbers and percentages of the format of works within a range of numbers. Most citations to books (43.8%), journal articles (17.8%), undergraduate reports (19.2%), Masters' dissertation (4.1%) and conference papers (12.3%) are in the range of 1 to 5. Most reports do not include citations to journal articles (68.5%), undergraduate reports (80.8%), Masters' dissertation (95.9%), conference papers (83.6%) but only 5.5% of the reports do not have citations to books. All project reports have citations to the Web

with the most being within the range of 6 to 10 (26.0%). A total of 94.5% of the reports have citations to books, 63.0% to journal articles, 19.2% to undergraduate reports, 16.4% to conference papers and 4.1% to Masters' dissertations. Although sources from the Web are most frequently listed in the reference list, the students who wrote the reports do cite books, journal articles, undergraduate reports and conference proceedings. The numbers may be small but the mere presence of these citations conforms to performance indicator 1.2. (identifies a variety of types and formats of potential sources for information).

(c) Web Resources

Citations to the Web can be further decomposed by type using its URL extension and .org, .com, .edu, and .net were used in this study. Any URL that does not fall into any one of this category is identified as "others". Table 6 shows that the .com sites are most favoured over the other sites and this constitutes 65.5% of all the Web citations. The rest of the sites are lagging behind and their totals are less than 15.0% each.

Table 6: Distribution of Web Citations by Category (n=1487)

Type	Frequency	%
.com	959	64.5
.edu	214	14.4
.org	169	11.4
.net	54	3.6
Others	91	6.1
Total	1487	100.0

When looking at the number of reports having Web citations within a certain range, it can be seen that not all the reports have citations to .com sites (Table 7). Only 95.8% of the reports include a .com site. Interestingly, although the numbers are small in terms of frequency of occurrence, 61.1% of the reports have citations to .org sites, 69.9% to .edu sites, 42.4% to .net sites and 49.3% to "others"(42.5%).

Table 7: Range of Web Citations by Category (n=1487)

Range	.org Freq (%)	.com Freq (%)	.edu Freq (%)	.net Freq (%)	Others Freq (%)
0	28 (38.4)	3 (4.1)	22 (30.1)	42 (57.5)	37 (50.7)
1-5	38 (52.1)	23 (31.5)	39 (53.4)	31 (42.5)	34 (46.6)
6-10	4 (5.5)	16 (21.9)	8 (11.0)	0 (0.0)	2 (2.7)
11-15	1 (1.4)	10 (13.7)	1 (1.4)	0 (0.0)	0 (0.0)
16-20	2 (2.7)	8 (11.0)	1 (1.4)	0 (0.0)	0 (0.0)
21-25	0 (0.0)	6 (8.2)	2 (2.7)	0 (0.0)	0 (0.0)
26-30	0 (0.0)	2 (2.7)	0 (0.0)	0 (0.0)	0 (0.0)
31-35	0 (0.0)	1 (1.4)	0 (0.0)	0 (0.0)	0 (0.0)
>35	0 (0.0)	4 (5.5)	0 (0.0)	0 (0.0)	0 (0.0)
Total	73 (100.0)	73 (100.0)	73 (100.0)	73 (100.0)	73 (100.0)

(d) Currency of Sources

It is regrettable that half of the citations (51.0%) do not have a date and when looking through the reports again, the absence of the dates is more often than not associated with Web citations (Table 8). Of the citations that do have a date, most of them are within the last three years (29.9%). However, 2.1% of the citations refer to publications that were

published more than 12 years ago with the earliest year being 1935 followed by 1966, 1967, 1976, 1977, 1978, 1987, 1988, 1989 and 1990. The absence of the date of publication in a citation could indicate that the writer is unaware of the proper way to cite or could not locate the date within the information source itself.

Table 9: Date of Publication of Citations (n=2184)

Date of Publication	Frequency	%
No date	1113	51.0
< 3 years	653	29.9
4 – 6 years	208	9.5
7 – 9 years	125	5.7
10 – 12 years	39	1.8
> 12 years	46	2.1
Total	2184	100.0

The former would indicate an inability to conform to performance indicators 4.1. (understands many of the ethical, legal and socio-economic issues surrounding information and information technology), 4.2. (follows laws, regulations, institutional policies, and etiquette related to the access and use of information resources), and 4.3. (acknowledges the use of information sources in communicating the product or performance). The latter would indicate that the writer is unable to fulfill performance indicator 2.5. (extracts, records, transfers, and manages the information and its sources).

The use of current information sources is commendable since it is reflective of performance indicator 3.2. (selects information by articulating and applying criteria for evaluating both the information and its sources) since currency is one of the criteria for evaluation of information sources.

(e) Citation Style

There were glaringly obvious discrepancies in citing print and Web resources. Therefore, the print and Web citations had to be analysed as two separate entities and categorised as follows:

- i. Print citations correct but Web citations incorrect
- ii. Print citations incorrect but Web citations correct
- iii. Print and Web citations correct
- iv. Print and Web citations incorrect

Only 12.3% of the reports had both print and Web citations written out in the proper format, but 31.5% had both of them wrong (Table 9). Another 41.0% had only the print citations in the correct format.

There is cause for concern here since presumably the owners of these reports would have attended the compulsory Information Skills Course in their first year where they were taught how to interpret a bibliographic record as well as compile a bibliography using the APA style. Apart from that, these students would have also completed their Industrial Training Programme which requires them to write a comprehensive report inclusive of a reference list. The format for citing sources are also clearly displayed and accessible via the Industrial Training Programme website. The absence of citations, which conforms to an agreed style, indicates the inability to conform to performance indicators 4.1. (understands many of the ethical, legal and socio-economic issues surrounding information and information technology), 4.2. (follows laws, regulations, institutional policies, and etiquette related to the access and use of information resources), and 4.3.

(acknowledges the use of information sources in communicating the product or performance).

Table 9: Citation Style (n=73)

Citation Style	Frequency	%
Print citations correct but Web citations incorrect	41	56.2
Print citations incorrect but Web citations correct	0	0.0
Print and Web citations correct	9	12.3
Print and Web citations incorrect	23	31.5
Total	73	100.0

4. Conclusion

In terms of information literacy competence and in relation to the *ALA/ACRL/STS 2005 Information Literacy Standards for Science and Technology*, the citation of the project reports complied to the various performance indicators of Standards 1 and 3 only (Tables 11 and 12). The information literacy of the authors of these reports who are the final year undergraduates are such that they:

- i. Listed a number of different information sources in their reference lists and this complies to Standard 1 which states that “the information literate student determines the nature and extent of the information needed”.
- ii. Use current information sources in their project reports reflecting their ability to evaluate information sources. This complies with Standard 3 where “the information literate student critically evaluates the procured information and its sources, and as a result, decides whether or not to modify the initial query and/or seek additional sources and whether to develop a new research process”.

Standards 2, 4 and 5 are not traceable through the project report. Almost all of the performance indicators in Standard 2 can only be traced through observation of an actual work in progress and not through a completed project report. The skills associated with Standard 2 cannot be traced through the literature review and bibliography since it is in reference to the student’s ability to obtain the information they require for their project report. Skills associated with Standard 2 can be best gauged through actual observations of the students interacting with the different sources of information or through the students’ journal entries of their information seeking process

As for Standard 4, this warrants the analysis of the project report as well as a presentation. Performance indicators for Standard 5, can only be gauged via a presentation by the authors of the reports or an interview with them. Standard 5 is a higher level skill and refers to the students’ ability to keep abreast with current developments in the discipline of computer science and information technology as well as understanding that information literacy is an ongoing process and realising that it is an important component of lifelong learning. The ability to be kept informed about current developments in their field can be ascertained from the currency of the information sources that the students used, most of which are within the last three years. The other components of Standard 5 are more suitably determined via interviews or presentations

Through this study, the researcher was able to ascertain compliance to Standards 1 and 3 only. However, whether the students have internalised the searching and evaluation process as a useful skill, which they can use cannot be determined from analysing the bibliography. The analysis only provided the study with an in-situ picture of the

information literacy competencies of the final year undergraduate students as seen through their bibliographies. Further works should explore other avenues for assessing the skills.

While most of the indicators do not match the standards, the findings do have practical implications for educators. The educators should seriously look into this matter in order to define and identify the role of educators and other academic fields in defining acceptable types of resources for papers and citation formats. At the same time, librarians need to seriously look into the provision of an information literacy course for their undergraduates. These students come into the system with different levels of skills which has to be identified and subsequently will result in the design and development of viable and time tested information literacy courses. Through these courses, the students should have obtained the necessary skills which comply to a given standard.

References

- ALA/ACRL/STS Task Force on Information Literacy for Science and Technology. 2005. *Information Literacy Standards for Science and Engineering/Technology*. Chicago: ALA. Accessed on 16 December 2006 from: <http://www.ala.org/ala/acrl/acrlstandards/infolitscitech.htm>
- American Library Association. Association of College and Research Libraries. 2000. *Information literacy competency standards for higher education*. Chicago: ALA
- Australian and New Zealand Institute for Information Literacy. 2004. *Australian and New Zealand information literacy framework: principles, standards, and practice*. 2nd ed. Adelaide: ANZIL
- Davis, Philip M. 2002. The effect of the Web on undergraduate citation behavior: a 2000 update. *College & Research Libraries*, Vol.63, no.1:53-60.
- Davis, Philip M. 2003. Effect of the Web on undergraduate citation behavior: guiding students scholarship in a networked age. *portal: Libraries and the Academy*, Vol.3, no.1: 47.
- Davis, Philip M and Suzanne A. Cohen. 2001. The effect of the Web on undergraduate citation behavior 1996-1999. *Journal of the American Society for Information Science and Technology*, Vol.52, no.4:309-314.
- Heller-Ross, H. 2002. Nursing research assignments and citation analysis of student bibliographies. *The Reference Librarian*, Vol.77:121-139.
- Hovde, Karen. 2000. Check the citation: library instruction and student paper bibliographies. *Research Strategies*, Vol.17, no.1:3-9.
- Leiding, Reba. 2005. Using citation checking of undergraduate honors theses bibliographies to evaluate library collections. *College & Research Libraries*, Vol.66, no.5:417-429.
- Magrill, R.M. and G. St. Clair. 1990. Undergraduate Term Paper Citation Patterns by Disciplines and Level of Course. *Collection Management*, Vol.12, no.3/4:25-56
- Mohler, Beth A. 2005. Citation analysis as an assessment tool. *Science & Technology Libraries*, Vol.25, no.4:57-64.
- Society of College, Universities and National Libraries. 1999. *Information skills in higher education: a SCONUL position paper*. London: SCONUL.

Table 11: Matching Features of Project Reports to Performance Indicators

Features	Performance Indicators	Compliance
Number of citations	1.1. Defines and articulates the need for information.	Yes
Types of sources	1.2. Identifies a variety of types and formats of potential sources for information	Yes
Currency of sources	2.5. Extracts, records, transfers, and manages the information and its sources.	No
	3.2. Selects information by articulating and applying criteria for evaluating both the information and its sources.	Yes
	4.1. Understands many of the ethical, legal and socio-economic issues surrounding information and information technology.	No
	4.2. Follows laws, regulations, institutional policies, and etiquette related to the access and use of information resources.	No
	4.3. Acknowledges the use of information sources in communicating the product or performance.	No
Citation style	4.1. Understands many of the ethical, legal and socio-economic issues surrounding information and information technology.	No
	4.2. Follows laws, regulations, institutional policies, and etiquette related to the access and use of information resources.	No
	4.3. Acknowledges the use of information sources in communicating the product or performance.	No

Table 12: Matching Features of Project Reports to Standards

Standards	Performance Indicators	Features in Project Reports
1. The information literate student determines the nature and extent of the information needed and constructs a course of action for obtaining the information.	1.1. Defines and articulates the need for information 1.2. Identifies a variety of types and formats of potential sources for information 1.3. Has a working knowledge of the literature of the field and how it is produced. 1.3. Considers the costs and benefits of acquiring the needed information.	Number of citations Types of sources used

<p>2. The information literate student procures needed information effectively and efficiently</p>	<p>2.1. Selects the most appropriate investigative methods or information retrieval systems for accessing the needed information 2.2. Constructs and implements effectively designed search strategies. 2.3. Retrieves information using a variety of methods 2.4. Refines the search strategy if necessary. 2.5. Extracts, records, transfers, and manages the information and its sources..</p>	
<p>3. The information literate student critically evaluates the procured information and its sources, and as a result, decides whether or not to modify the initial query and/or seek additional sources.</p>	<p>3.1. Summarizes the main ideas to be extracted from the information gathered. 3.2. Selects information by articulating and applying criteria for evaluating both the information and its sources. 3.3. Synthesizes main ideas to construct new concepts 3.4. Compares new knowledge with prior knowledge to determine the value added, contradictions, or other unique characteristics of the information. 3.5. Validates understanding and interpretation of the information through discourse with other individuals, small groups or teams, subject-area experts, and/or practitioners. 3.6. Determines whether the initial query should be revised. 3.7. Evaluates the procured information and the entire process.</p>	<p>Currency of sources</p>
<p>4. The information literate student understands and respects the economic, ethical, legal, and social issues surrounding the use of information and its technologies and either as an individual or as a member of a group, uses information effectively to accomplish a specific purpose</p>	<p>4.1. Understands many of the ethical, legal and socio-economic issues surrounding information and information technology. 4.2. Follows laws, regulations, institutional policies, and etiquette related to the access and use of information resources 4.3. Acknowledges the use of information sources in communicating the product or performance. 4.4. Applies creativity in use of the information for a particular product or performance. 4.5. Evaluates the final product or performance and revises the development process used as necessary. 4.6. Communicates the product or performance effectively to others.</p>	
<p>5. The information literate student recognizes the need to keep current regarding new developments in his or her field and understands that information literacy is an ongoing process and an important component of lifelong learning.</p>	<p>5.1. Recognizes the value of ongoing assimilation and preservation of knowledge in the field. 5.2. Uses a variety of methods and emerging technologies for keeping current in the field.</p>	