Facilitating access to electronic resources: matching bibliographic record content with client preferences

Lynne C. Howarth and Joseph P. Cox
Faculty of Information Studies
University of Toronto
howarth@fis.utoronto.c
cox@fis.utoronto.ca

This paper presents findings from a study in progress that addresses the paucity of research concerning which bibliographic, and particularly descriptive, elements should be included in a record so as to facilitate identification and retrieval of information. Data from both checklist and focus-group methodologies were compared to provide a client-oriented framework for assessing content in bibliographic records, and to highlight gaps between user preferences and the availability and presentation of elements in some existing bibliographic systems.

1. Introduction and background to the study

For nearly three decades the *Anglo-American Cataloguing Rules* have been the basis for the description of bibliographic resources in library catalogues. Over that time a great deal of effort and expense has gone into creating catalogues that adhere to the standards set out in the *Rules*. However, very little research has been done to evaluate the usefulness of the specific bibliographic elements contained in bibliographic records. That is, having found a record in a online public access catalogue (OPAC), what individual descriptive elements are used by the client to determine the suitability of the retrieved record, whether to refine or continue the search or to locate and use the material described by the record.

This paper presents findings from a study in progress that address the paucity of research concerning which bibliographic, and particularly descriptive, elements should be displayed, and in what manner so as to facilitate identification and retrieval of information. By means of a second-level descriptive record framework from the *Anglo-American Cataloguing Rules* (an international standard), a checklist of bibliographic elements was developed. Records for ten items in Internet-accessible OPACs of ten public libraries across Canada were evaluated for their content relative to the prescribed bibliographic standard. Concurrently with checklist data collection, focus group interviews were held with a stratified sample of public library users who were asked to rank the

relative importance and usefulness of elements of bibliographic records, and subsequently to "prototype" their ideal record. The data from both the checklist and focus group-methodologies provide a client-oriented framework for assessing content in bibliographic records, and highlight gaps between user preferences and the availability and presentation of elements in some existing bibliographic systems.

2. Literature review

Many authors have claimed that much of what appears in many bibliographic displays is not wanted or used by most clients, and that other information may be more valuable to them. Crawford (1992) refers to the "arcana" of data displayed in bibliographic records. Displaying unnecessary or unwanted elements may confuse the clients and make it more difficult for them to scan the display (several authors cited in Wallace 1984). The cost of recording these pieces of information for the bibliographic record is also a concern (Wallace 1984; IFLA 1992).

Palmer's (1972) study found that 5,000 users at the University of Michigan used, on average, 4.5 elements of the 20 elements on a display; he concluded that many of the elements could be deleted. For example, he found that clients used the following elements less than 10% of the time: authors' birth/death dates, pagination, size, series note, translator, illustrations, bibliography note, and Library of Congress card number. Previous studies (cited in Palmer 1967) had found that names of joint authors, editors, joint editor, compiler, joint compiler, serial editor, added entries for government headings, and personal, corporate, and series added entries were also used less than 10% of the time. Yet many systems still display these elements today.

Most studies in this area have been conducted in university environments. One of the few studies in public libraries was a survey of OPAC users, which showed that only 51% of the respondents were satisfied with the amount of bibliographic data supplied (Muncer 1990). Seal (1983) conducted an experiment with Computer Output Microform (COM) catalogues, a full-entry catalogue and a short-entry catalogue, to assess client needs for bibliographic data elements. Seal concluded that much of the information normally included in the catalogue entry is very rarely used by clients and its inclusion makes catalogues difficult to use. He also notes that if such data were eliminated, we could consider providing other, more valuable information. His study did not identify which information might be more valuable.

3. Methodology

3.1. Designing the checklist of bibliographic elements

In order to evaluate present bibliographic records, a research procedure was developed that would examine the content of ten bibliographic records in ten different public library sites. It was decided to begin with the development of a checklist of those elements currently used in records catalogued according to present cataloguing standards. Therefore, the process of defining a checklist of bibliographic elements was based on the *Anglo-American Cataloguing Rules*, 2nd ed., 1988 rev. (*AACR2R*), *Amendments 1993*, second-level description. The second level of bibliographic description is the one most widely used in national and local catalogues and is believed to address the needs of most catalogue creators and information seekers. While the essential description of an item is contained within the criteria of second level, it was also recognized that additional elements form a critical part of the bibliographic records in catalogues. These additional elements include the access or indexed points of the names of personal or corporate names, and so these were added to the descriptive elements checklist. The *AACR2R* second-level description was further adapted by the removal of the "material specific designation" element, as it is not applicable to monographic materials.

Table 1 is the comprehensive list derived from *AACR2R*. There are 24 elements included. Also listed are the ten items A–J (see section below, "Selection of Items"). The checklist also shows with an asterisk those items in which particular elements are present. A space indicates that the element was not applicable to that item. As can be seen from the list, all elements, except subseries ISSN, are found at least once in the items. There are 143 occurrences of the elements in the ten items chosen for examination.

3.2. Selection of items

The selection of items used to evaluate the bibliographic elements involved meeting a number of needs. The items had to be English-language monographs and meet the dual criteria of (1) being owned by all the public library sites across the country and (2) illustrating different bibliographic features in order to provide a sampling of how OPACs handle all the bibliographic elements as listed in the checklist in Table 1. Two research assistants identified many potential items for inclusion in the list. All items in this preliminary list were examined to determine if they exhibited the necessary bibliographic elements so that every element in the checklist, if possible, was found at least once in the records for the items. The preliminary list of items was also pre-searched in the library OPACs to determine that they were indeed held by the libraries whose

Table 1. Checklist of bibliographic elements

Bibliographic elements	A	B	C	D	E	F	G	H	I	J	Total
Author(s)	*	*	*	*	*	*	*	*	*	*	10
Title proper	*	*	*	*	*	*	*	*	*	*	10
General material designation	*	*	*	*	*	*	*	*	*	*	10
Parallel title	*								*		2
Other title information			*		*				*		3
1st statement of responsibility	*	*	*	*	*	*	*	*	*	*	10
Subsequent statement of responsibility					*				*		2
Edition statement	*		*		*		*	*			5
Statement of responsibility for edition							*				1
1st place of publication	*	*	*	*	*	*	*	*	*	*	10
1st publisher	*	*	*	*	*	*	*	*	*	*	10
Date of publication	*	*	*	*	*	*	*	*	*	*	10
Extent of item	*	*	*	*	*	*	*	*	*	*	10
Other physical details	*	*	*		*	*			*	*	7
Dimensions	*	*	*	*	*	*	*	*	*	*	10
Accompanying material	*										1
Series title proper		*	*	*	*				*	*	6
Series statement of responsibility				*					*		2
Series ISSN		*									1
Series numbering				*	*				*	*	4
Subseries title										*	1
Subseries ISSN											0
Notes	*	*		*	*	*		*	*	*	8
Standard number	*	*	*	*	*	*	*	*	*	*	10
Total	15	14	14	14	17	12	12	12	18	15	143

bibliographic records would be examined. The following is the final list of items:

- *Canada with love=Canada avec amour*. Patriation ed. (1982). Monk, Lorraine.
- *Canada's Department of External Affairs* (1990). Hilliker, John.
- *Canadian business guide to environmental law : protect yourself! protect us all!* 1st ed. (1988). Rovet, Ernest.
- *Euthanasia: aiding suicide and cessation of treatment* (1982). Law Reform Commission of Canada.
- *Field guide to the birds: a completely new guide to all the birds of eastern and central North America*. 4th ed. (1980). Peterson, Roger Tory.
- *Historical atlas of Canada (1987–1993)*. Matthews, Geoffrey J.
- *Oxford companion to German literature*. 2nd ed. (1986). Garland, Henry.

- *The Oxford English dictionary.* 2nd ed. (1989). Simpson, J.A.
- *A rattle of pebbles: the First World War diaries of two Canadian airmen* (1987). Greenhous, Brereton.
- *Robertson Davies* (1986). Peterman, Michael.

3.3. Selection of public library sites

Guiding the selection of the public library sites to be used in this research were a number of criteria. Selection of the libraries was narrowed to members of the Canadian Association of Large Urban Public Libraries (CALUPL) with OPACs accessible through the Internet. To provide a more balanced picture, the geographic location and the OPAC system used in each library were also considered. Collection size was not thought to be an issue for the purposes of this research project; however, predominantly French-language institutions were eliminated owing to the low level of overlap with the collections of English-language institutions. In March 1995, two research assistants determined which member institutions of CALUPL adhered to these criteria. There are no representative libraries from Quebec and Eastern Canada, as some members of CALUPL at this time did not fulfil the criteria of Internet-accessible catalogues or the ownership of the selected items. The following is the final list of public library sites chosen:

Burlington Public Library
Edmonton Public Library
Metro Toronto Reference Library
Mississauga Public Library
Ottawa Public Library
Regina Public Library
Saskatoon Public Library
Scarborough Public Library
Vancouver Public Library
Winnipeg Public Library

3.4. Expert and copy cataloguing

During this process, it became apparent that there would be value in developing "ideal" records for items to facilitate comparing how the various institutions handle different bibliographic elements. Therefore, all items were catalogued by a group of Cataloguing Experts to the precise standards of a second-level descriptive record according to *AACR2R.* The value of retaining Library of Congress and/or National Library of Canada cataloguing copy for the items to be incorporated into the comparisons and evaluations was also realized. This additional information would provide insight into the dissemination of bibliographic information as well as trends in cataloguing practices. A cataloguing chart created for each individual item showed how every bibliographic element of each item was handled by the Experts, the National Library of Canada, and the Library of Congress.

3.5. Data collection

Once the items and sites had been selected, and during the period March 26 to April 1, 1995, two research assistants independently collected and evaluated the bibliographic records from each of the sites using a data collection form. In cases where more than one record for the same item was available on a system, the first record served as the basis for evaluation. The OPAC screens were captured and retained for future reference. Additionally, the research assistants made notes of any obvious errors, omissions, or anomalies in the records. Subsequently, one of the researchers compared and checked for agreement/disagreement between the two sets of data and resolved any differences, referring to the captured screen images where necessary.

3.6. Data tabulation

The data collected were initially tabulated and summarized in three ways, namely, through site summaries, item summaries, and element summaries. The site summaries showed how *each site* had treated the specific bibliographic elements for *all of the items*. The item summaries illustrated how *all sites* had dealt with the bibliographic elements for *each item*. The element summaries demonstrated how *all sites* had dealt with individual bibliographic elements for *all items*.

Further analysis of the data was carried out with the intention of tabulating the level of congruence/incongruence between the bibliographic elements used by the Cataloguing Experts in their "to standard" cataloguing of the items and the cataloguing provided by the ten sites and also by the National Library of Canada and the Library of Congress. A scoring method was devised that allowed for very detailed analysis of how elements differed across the sites. A chart was created that contained the Cataloguing Experts' list of bibliographic elements for each item. With the summaries previously created and the screen printouts, as well as national agency cataloguing, every bibliographic element used by the sites was compared to the bibliographic elements in the "to standards" cataloguing record created by the Cataloguing Experts. A scoring method was applied to every element in every item in every site such that every element was rated from 1 to 6, indicating how it differed from the Experts' version (Table 2).

Table 2. Scoring chart for congruence/incongruence between Experts and sites

T1. Typos, punctuation, diacritics, spacing	T4. Partially complete content
T2. Different wording, same meaning	T5. Completely incorrect content
T3. Differences in rule interpretation	T6. Missing data entirely

4. Data analysis and findings

Data summaries suggested that all sites (n = 10 or 100%) consistently provided information for the following bibliographic elements, whenever such information was readily available and/or applicable according to the bibliographic standard *AACR2R*:

- Titles (title proper and other title information)
- 1st place of publication
- 1st publisher name
- Date of publication
- Extent of the item (pagination)
- Notes about the item

Several sites never supplied the following information, even when it was readily available and/or applicable according to the *AACR2R* guidelines:

- Parallel title (the same title in a second or other language)
- General material designation (indicating material format, e.g., "text" for printed monographs)
- Dimensions (size of the item)
- International Standard Book Number (ISBN = a unique identifier for each monograph title)
- International Standard Serial Number for series (ISSN = a unique identifier for serial titles)

Some variability was demonstrated across Canadian public libraries as to their inclusion of the following bibliographic elements:

- Statement of responsibility (first statement of authorship relative to a title, an edition, or a series title)
- Edition statement
- Other physical details (e.g., illustrations)
- Series titles or subtitles
- Series numbering (the number or volume in the series)

Overall, it was observed that the bibliographic records for English-language monographs held in Canadian public libraries were created fully to the *AACR2R* standard of the second level of description in those areas related to the provision of title proper, publication information, and notes about the nature or content of the item—all areas for which data are usually readily available from the monograph itself. Information that might be considered more "obscure" or difficult to determine from the item would include, in particular, series titles or subtitles and series numbering. The inclusion in the bibliographic record of material format (general material designation or "gmd") is optional according to *AACR2R*, and might account for the majority of sites' omitting it from the bibliographic display. The exclusion of the ISBN (and where applicable the

ISSN) is likely not a reflection of bibliographic practice so much as of OPAC system design. While the ISBN (or ISSN) may, in fact, be supplied in the original biliographic record—as indeed it was in several of the sites' machine-readable cataloguing (MARC) records—the system itself may support it as a unique numeric access key while also suppressing its display. In some cases it was neither displayed nor activated as an access key even though it existed in the original MARC record for the item.

While the initial data summaries provided an overview of treatment of relevant bibliographic elements on an item-by-item and site-by-site basis, the researchers were interested in obtaining a more detailed understanding of the type and extent of bibliographic variations within the ten monographs. With both a general and specific bibliographic perspective, it would be possible to speculate on the nature, degree, and extent of variability existing to challenge the searching effectiveness of OPAC users.

With reference to the rating scale described in Table 2 above, data were summarized from the scoring forms (see Table 3 as an abridged example) used to record on an item-by-item basis, each time a type of error pertained to a particular bibliographic element (i.e., frequency of occurrence) along with the number of sites committing the error (i.e., number of all occurrences of the error).

Table 3: Scoring form for error types (abridged example)

Elements	Experts		LC	NLC	P1	P2	P3	P4	P5	P6	P7
Pagination	ca. 125 p. :	1									
		2									
		3									
		4									
		5									
		6									
Illustrations	50 col. Ill ;	1									
		2									
		3									
		4									
		5									
		6									
Dimensions	30 cm. +	1									

As Table 4 illustrates, the type of error that occurred with greatest frequency within individual bibliographic elements, across the ten monographs held at ten Canadian public libraries, was that of missing data. That is, applicable data were not supplied in 69 instances for a total occurrence of 282 missing data errors

Table 4. Summary of errors

Bibliographic elements	T1	T2	T3	T4	T5	T6
Author (as main entry)	1/2	5/8	4/16			
Authors (as added entry)	3/8	12/30	7/19			7/32
Title proper	10/35	1/1	1/1			
General material designation [GMD]			10/11 6			
Parallel title	1/9					1/1
Other title information	1/5					
1st statement of responsibility	7/18	1/7	2/14			7/7
Subsequent statement of responsibility	1/2					2/2
Edition statement	4/7		2/4			2/11
Statement of responsibility for edition	1/2					1/6
1st place of publication		5/16	1/3			
1st publisher	7/7	1/7	3/16	1/2		
Date of publication	1/1	2/7	4/12	1/4		
Extent of item (pagination)	2/7	3/6	5/9	3/6	2/8	1/1
Other physical details (illustrations)	3/4	4/11	1/7	5/14		4/12
Dimensions (size of item)	5/18				3/7	9/34
Accompanying materials		1/1				1/9
Series title proper	7/62	6/14		2/6		5/9
Series statement of responsibility		1/7				1/1
Series ISSN			1/3			
Series numbering	1/1	2/3	1/1			3/7
Subseries title		1/1				1/3
Notes (summarized)	1/3	16/75		8/26		12/51
ISBN #1	7/10	7/12		1/4		10/78
ISBN #2						1/9
ISBN #3						1/9
Total frequency of occurrences by error type	63/	68/	42/	21/	5/	69/
Total number of all occurrences of error	/201	/206	/221	/62	/15	/282

(or 28.5% of all occurrences of error, $n = 987$). A "Type 2 " error was assigned where different wording was used to convey the same meaning (for example, a separate note each for a bibliography and index, as opposed to a combined note for both). This error type ranked second in frequency of occurrence ($n = 68$), for a total of 206 errors (or 20.9% of the overall total, $n = 987$). Errors in typography, spelling, diacritics, spacing, or punctuation ("Type 1" errors) ranked third with 63 unique occurrences, while accounting for 206 errors across all sites (or 20.9% of the overall total, $n = 987$). The fourth most frequently occurring error was that of a difference in *AACR2R* rule interpretation between the Experts and the individual sites. Examples of such errors would include the

specification of illustration (e.g., "50 col. ill.") versus a generic term (e.g., "ill."). A total of 42 separate instances of Type 3 errors were recorded, accounting for 221 total occurrences (or 22.4% of the overall total, $n = 987$). It is interesting to note that, while there were more unique occurrences of Types 1 ($n = 63$) and 2 ($n = 68$) errors, the cumulative total of sites repeating those errors was less than that for Type 3 (total occurrences $n = 221$). Fewer sites were responsible for more Type 3 errors, while more sites replicated fewer errors of Types 1 (total occurrences $n = 201$) and 2 (total occurrences $n = 206$), respectively.

Most commendable from the perspective of quality of bibliographic record was the low frequency and overall number of errors recorded for partially complete content (Type 4) or for completely incorrect content (Type 5). The latter were considered to be errors of a larger magnitude because of their potential to mislead or impede client identification of, or access to, items. Only five occurrences of Type 5 errors were recorded (total occurrence = 15, or 1.5% of the overall number of errors occurring, $n = 987$), while Type 4 errors were uniquely registered in 21 instances but repeated 62 times (for 6.3% of the overall number of errors, $n = 987$).

In summary, then, data were more likely to be missing from records than to be included but entirely incorrect, or only partially complete in content. More libraries erred in their interpretations of *AACR2R* relative to "to standard" records created by Cataloguing Experts, but more individual instances of errors were attributable to fewer sites as pertained either to differences in typography, spelling, diacritics, spacing, or punctuation (Type 1), or to using different words or phrases to convey the same meaning (Type 2).

Having determined the nature, type, and degree of categories of errors identified in bibliographic records accessible via public library OPACS, it then remained to relate those findings to information-seeking clients, the intended "consumers" of bibliographic records. To draw this relationship, the researchers referred to findings from a focus-group study (Luk 1996) of two linguistic "sets" of public library users—one group for whom English was the first language, and the second for whom Cantonese was first. Stratified by age, three groups of each linguistic cohort were asked to rank (from most to least) the relative importance of various bibliographic elements (based on *AACR2R*, and with terminology based on second-level description), and subsequently to design, as part of an unstructured exercise, an "ideal" display both in terms of format/presentation and bibliographic content. Findings suggested that the most *used* bibliographic elements were title, author, and subject, while the most infrequently *used* were International Standard Book Number (ISBN) and Library of Congress Control Number (LCCN). For English-language par-

ticipants, the *most important* bibliographic elements were title, author, and summary (note), while for the Cantonese-language participants, title, author, and call number ranked as *most important.*

Expanding the list to the ten top-ranked *descriptive* bibliographic elements (out of a total of 37 possible elements, excluding subject and call numbers), those designated as *most important* were, in descending order by participant group:

English as first language
1. Title
2. Author
3. Summary [note]
4. Date of publication
5. Type of material
6. Reading level [note]
7. Publisher
8. Other author(s)
9. Extent of item (pagination)
10. Edition

Cantonese as first language
1. Title
2. Author
3. Publisher
4. Summary [note]
5. Type of material
6. Date of publication
7. ISBN
8. Reading level [note]
9. Other author(s)
10. Edition

Each of the focus groups was conducted separately, by age group and also by first language. It is somewhat remarkable to observe the relatively high level of congruence and overlap in the selection of *most important* descriptive elements on the part of the six (total) groups. If one relates the above findings back to a consideration of the nature, degree, and type of "errors" characteristic of bibliographic records accessible via the Internet from public library OPACs, one notes that two of the top-ranked elements, while "permitted" by *AACR2R* and possible candidates for any of the ten monographs selected for the study, never appear in any of the records. "Summary notes" and "Audience level" or level of reading notes are accommodated in the bibliographic standard, but not included by any of the sites. Similarly, "Type of material", optionally assignable in *AACR2R* as a general material designation (e.g., "text"; "sound recording"; "computer file") is an element appearing in only two of ten sites. If one considers the number of different types of errors found in one bibliographic element to be an expression of high variability within that element, then four of the top-ranked *most important* elements would harbour from one to six types of errors. For the English-language group, the extent of the item ranks ninth; Types 1–6 can be found in this element. Date of publication and (name of) publisher record instances of Types 1, 2, 3, and 4 errors, with "Other author(s)" substituting Type 6 or "Missing data" for the less serious recording of a Type

4 error. Only title and author elements demonstrate a relatively high level of consistency and completeness, though the existence, in many records, of Type 1 errors (typography, spelling, spacing, diacritics, punctuation) could have deleterious consequences for the success of a search. Spelling or input errors can relegate titles, in particular, to the netherworld of the "misplaced" and irretrievable access point.

To recap, the content and integrity of the bibliographic records being created currently may be insufficient and inadequate to appropriately address even a "short list" of bibliographic "wants and needs" on the part of potential public library OPAC users. Both the focus-group study (Luk 1996) and the present study of bibliographic elements in selected records for monographs held in selected sites across Canada lack sample sizes fully representative of, or generalizable to, the universe of OPAC users or bibliographic records. Nonetheless, the message that the apparent "gap" between user preferences and the structure, content, and current manifestation of OPAC records brings is a warning about—or perhaps a window of opportunity for—better adapting our bibliographic delivery mechanisms to suit the needs of an increasingly literate and demanding clientele in an expanding electronic environment.

If the bibliographic structures we have created to facilitate user identification of, and access to, materials have proved insufficient, incomplete, or even useless for clients in the OPAC environment, one wonders at the further alienation of bibliographic "product" from "consumer" in the increasingly visually oriented and graphically flexible world of electronic resources. These questions become even more critical when addressed to the growing availability of electronic resources accessible via local networks, the Internet, or intranets. For example, bibliographic records that are created to identify and describe "virtual" textual or multimedia materials add front-end value to the relative "information anarchy" characteristic of the Internet. Essential to developing and maintaining virtual library collections, bibliographic records for electronic resources must be sufficiently and appropriately descriptive to facilitate client access to materials that are not literally tangible, and therefore not readily verifiable as to their relevance to user needs. Bibliographic records continue to have the potential to provide that relatively interpretable, "thumbnail" sketch of the nature and content of items in a variety of formats. Moreover, the structures do exist currently for accommodating records with enhanced content and richer detail (with little extra effort at the creation stage). As this paper has attempted to demonstrate, albeit on a small scale, there appears to be a gap between user preference and the content and integrity of bibliographic records in their present form. To continue to deny or ignore that gap may be at our own peril as facilitators of information access and exchange.

Acknowledgements

The authors gratefully acknowledge the financial assistance provided by the Social Sciences and Humanities Research Council of Canada in funding this research project. They also wish to thank Trina Richard, Catalogue Librarian at Ryerson Polytechnic University Library, Toronto, for her many and invaluable contributions to this project, including the scoring of errors in bibliographic records and for creating the tables included here.

References

Crawford, Walt. 1992. Starting over: current issues in online catalog user interface design. *Information Technology and Libraries* (March): 62–76.

IFLA Standing Committee for Section on Cataloging. 1992. *Terms of reference for a study group on functional requirements for bibliographic records.* September 4.

Luk, Annie T. 1996. *Evaluating bibliographic displays from the users' point of view: a focus group study.* Master of Information Studies Research Project Report. Toronto: Faculty of Information Studies, University of Toronto.

Muncer, Joy, Carol Hill, and Martin Dudley. 1990. Raised eyebrows: a survey of public reactions to an OPAC. *Public Library Journal* 5 (3): 71–72.

Palmer, Richard P. 1972. *Computerizing the card catalogue in the university library: a survey of user requirements.* Littleton, CO: Libraries Unlimited.

Seal, Alan. 1983. Experiments with full and short entry catalogues: a study of library needs. *Library Resources and Technical Services* 27: 144–55.

Wallace, Danny P. 1984. *The user friendliness of the library catalog.* University of Illinois, Graduate School of Library and Information Science. Occasional papers, No. 163.