Quality Control for a Bibliographic Database

Vinh-The Lam
Manager, Database and Information Resource Development
Canadian Centre for Occupational Health and Safety

(Published in ***Canadian journal of information science***, v. 17, no. 1, 1992, p. 24-31)

Abstract.- *The Canadian Centre for Occupational Health And Safety (CCOHS) was created in 1978 to promote the right of Canadians to a safer work-place. CCOHS has developed a computerized information systen called CCINFO. As a biblio-graphic database component of CCINFO, CANADIANA provides references to occupational Health and safety literature published in Canada, about Canada, or written by Canadian authors. Quality control efforts have been made in database design and record creation procedures for optimal retrieval of accurate and reliable bibliographic information.*

Introduction

The Canadian Centre for Occupational Health and Safety (CCOHS) was created by an Act of Parliament in 1978 *"to promote the fundamental right of Canadians to a healthy and safe working environment"* (Canadian Centre for Occupational Health and Safety Act 1978). CCOHS has created a computerized information system called CCINFO. This service is now available both online through connections to CCINFOline and on CD-ROM through subscriptions to the quarterly CCINFOdisc. CCINFO provides a variety of databases (bibliographic, direct information, full text) as well as colourful videotex information packages. Among the bibliographic databases, CANADIANA is the one that provides references to occupational health and safety (OHS) literature published in Canada, about Canada, or written by Canadian authors. It rounds out the OHS information base of CCINFO, which also provides access to the international literature, through CISILO, produced by CIS (Centre international d'information de sécurité et d'hygiène du travail, an agency of the International Labour Office in Geneva, Switzerland), and to the American literature, through NIOSHTIC, produced by NIOSH (National Institute of Occupational Safety and Health in Cincinnati, Ohio). This paper describes the structure and content of CANADIANA and discusses the quality control issues in its production.

Database Description

CCINFO is mounted on an HP3000 minicomputer and run by MINISIS software developed by the International Development Research Centre (IDRC) in Ottawa. As a bibliographic database, CANADIANA provides fields for data necessary to identify documents, such as author's name, title, publisher, ISBN, ISSN, report number(s), and organization name. Subject retrieval is provided through fields for descriptors (taken from the CIS Thesaurus), CAS Registry Numbers (for chemicals), contents notes, and abstracts or extracts (abstract-substitutes created by extracting useful information from the item being described). Most of these fields are searchable. Inversion types vary according to the nature of data entered or according to retrieval needs and patterns. Table 1 lists the searchable fields.

Table 1

Field Name	Mnemonic	Inversion Type
Author	AU	Word
Title	TI	Word
Language of publication	LA	Term
CAS Registry number	RN	Term
Descriptors	DE	Word, Term
Date of publication	PD	Full field
Country of publication	CY	Full field
Series number	SERNO	Full field
Report number	RE	Full field
ISBN	BN	Full field
ISSN	SN	Full field
Organization name	ON	Word
Source of document (Monograph)	LC	Word
Source of document (Periodical)	JN	Word, Full field
Series title	SE	Word, Full field

The following fields may also appear when records are browsed:

- ISN (Internal Sequence Number, the identification number of records)
- Reason for CANADIANA designation
- Conference name
- Publisher
- Edition statement

- Physical description
- Notes
- ISNs of related records
- Contents
- Abstract

Three print formats are provided for the users:

- short format: to provide just a basic bibliographic citation with descriptive information
- medium format: to provide basic bibliographic citation plus subject content information
- long format: the full record

With the CCINFOdisc, the users can create a print format of their own choice.

Database Content

CANADIANA gives references to the useful, relevant, and updated information available. Subjects of current importance in OHS such as WHMIS (Workplace Hazardous Materials Information System), back problems, drug screening, and passive smoking are given high priority. To achieve this, a number of core Canadian journals are analysed on a regular basis. Another important feature of this database is its ability to provide references to the "fugitive" documents such as position papers by workplace groups for important OHS issues as well as unpublished and limited-distribution reports by consulting firms. OHS conference proceedings are analyzed and individual papers fulfilling CANADIANA criteria are entered. Because one of the criteria for CANADIANA is Canadian authorship, efforts have been made to include articles written by Canadians published in non-Canadian journals.

CANADIAN can answer the folloing types of questions:

1. We understand that the Ontario Ministry of Labour has proposed important changes in OHS legislation and it has put out in 1989 a background paper on these changes. Can we identify that paper through a search in CANADIANA ?

   The response is shown in Figure 1.

   Search:

   Q on ontario ministry of labour

1 : POSTINGS = 261
Q and de legislation
2 : POSTINGS = 19
Q and pd 1989
3 : POSTINGS = 1
Q b (for browse)

* *
* C A N A D I A N A *
* *
* Canadian Centre for Occupational Health and Safety *
* *

ISN: 56597

Title / Statement of responsibility: New directions in workplace health and safety in Ontario : a background paper

Language of publication: English

Reason for CANADIANA designation: Published in Canada

Organization name(s): Ontario. Ministry of Labour

Publisher or Source of Document: [Toronto] : Ontario Ministry of Labour, 1989

Physical description: 1 v. (various pagings)

Date of publication: 1989

Country of publication: Canada

Report number(s): BR-REF-2 (#3-SPU-REP)

CIS Descriptors: Future trends, Workers participation, Responsibilities of employers, Responsibilities of employees, Satey and health training, Legislation, Ontario

Abstract / Extract: The purpose of this paper is to set out a proposed approach to improving the safety and health of Ontario workers. The paper outlines an approach to improved safety and health which places more authority and greater control over workplace risks in the hands of the workplace parties. And it links this greater authority to

responsibility and training. The approach also offers incentives for employers who follow positive health and safety programs which produce a safer workplace.

Figure 1

2. We are implementing WHMIS within our company. We want to do an audit of health hazards in our workplace. Does CANADIANA have any references on how to conduct such an audit ?

   The response is shown in Figure 2.

   Search:

   Q WHMIS
   1 POSTINGS = 94
   Q and audit
   2: POSTINGS = 1
   Q b

   ```
   * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
   *                                                                               *
   *                                C A N A D I A N A                              *
   *      Canadian Centre for Occupational Health and Safety                       *
   * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
   ```

   ISN: 55652

   Title / Statement of responsibility: Implementing WHMIS : specific solutions for your organization / by Tom Robinson

   Language of publication: English

   Reason for CANADIANA designation: Published in Canada

   Author(s): Robinson, T.

   Organization name(s): Centre for Professional Learning , Workplace Hazardous Materials Information System

Publisher or Source of Document: Occupational health & safety : legal and practical advice. [Toronto] : Centre for Professional Learning, [1988]. Vol. 2, ch. XIII. p. U- U14.

Date of publication: 1988

Country of publication: Canada

CIS Descriptors: Harmful substances, Inspection, Safety and health training, Data sheet, Labelling, Legislation, Rights of workers

Abstract / Extract: I have 5 key points that you should take away from this presentation. 1) The real impact of WHMIS; 2) Why an audit of specific hazards in your workplace is essential; 3) The difference between generic and specific WHMIS trainng; 4) Why WHMIS training must be related to your overall hazard control program; 5) The three phases of WHMIS implementation.

Figure 2

3. Does the Canadian Medical Association have an official position on drug screening in the workplace ?

   The response is shown in Figure 3.

   Search:

   Q drug and screening
   1: POSTINGS = 7
   Q and on canadian medical association
   2: POSTINGS = 1
   Q b

   *
   * *
   * C A N A D I A N A *
   * Canadian Centre for Occupational Healt and Safety *
   *

   ISN: 49843

   Title / Statement of responsibility: CMA position paper on pre-employment drug screening

Language of publication: English

Reason for CANADIANA designation: Published in Canada

Organization name(s): Canadian Medical Association

Publisher or Source of Document: Proceedings of the 120th annual meeting : including the transactions of the General Council, Charlottetown, Prince Edward Island. Ottawa : Canadian Medical Association, 1987. p. 236-241.

Date of publication: 1987

Country of publication: Canada

CIS Descriptors: Pre-employment medical examinations, Drugs, Determination in urine, Drug dependence, Rights of workers

Abstract / Extract: The practice of urine testing in the workplace, particularly as a pre-employment screen, is a recent phenomenon. It appears to have originated in the USA generated, in part, by a high profile campaign by President Ronald Reagan to attempt to control drug abuse. The question has been asked: what is the offical CMA policy on the issue ?

Figure 3

Quality Control

Like other CCOHS products, CANADIANA records, before being approved for public access, must go through appropriate checks for quality control. The checks, performed partly by the computer and partly by staff, aim at achieving accuracy and consistency.

1. *Computer-generated control*

Many studies have shown that bibliographic databases are not error-free, especially with respect to typing and spelling errors. Bourne found that the spelling error rates could go from less than 0.5% to over 22% (Bourne, 1977). With CANADIANA, efforts have been made to eliminate potential errors that might be caused by repeated entry of alphanumeric data. Special computer programs have been written by CCOHS systems staff to check validity of ISBNs and ISSNs. CANADIANA contains a field for descriptors taken from the CIS

Thesaurus. The latter is a controlled vocabulary devised and maintained by CIS in both printed and electronic form. The electronic form of the CIS Thesaurus is the TERMS database. A special version of TERMS containing only facet codes, English descriptors, and French descriptors have been created by CCOHS systems staff and used in a software package to provide automatic linkage between the facet codes and the descriptors. Staff key in only the codes and the descriptors are automatically displayed in CANADIANA. Another feature of automatic quality control in data entry is the existence of the relational database for organization names. This database plays the role of corporate body authority file for CANADIANA. Headings for corporate bodies are created based on National Library of Canada Canadiana Authority microfiches and/or Anglo-American Cataloging Rules, Second Edition (AACR2). Only the numeric code for an organization is entered and the full form of AACR2-approved heading for it is displayed in CANADIANA. Finally, another linkage mechanism is created to automatically provide each record with "Reason for CANADIANA designation," based on a list of codes:

1: Published in Canada
2: Canadian topic
3: Written by Canadian author(s)

Only codes 2 and 3 are entered by data entry operators. Code 1 is automatically assumed by the computer when the record contains the word Canada in the Country of publication field. With all these measures, not only is quality control achieved but also data entry time and storage space requirements are greatly reduced.

2. *Human / organizational control*

The Quality Assurance (QA) procedures for CANADIANA require that within each step in the record production, searching and verification, creation of bibliographic record, subject analysis, and technical edit / approval, standards be set up and responsibilities clearly defined. The first step aims to avoid duplication of records as well as to establish relationships between the item in hand and others already in the collection. The preliminary record, created at the second step, will contain bibliographic information prescribed by AACR2. It is reviewed / revised and enhanced at the next step. A separate document entitled "Subject Analysis : Policy and Guidelines" and appended to the QA procedures provides guiding principles for the indexing carried out by professionally trained subject analysts. The CIS Thesaurus is used for this vocabulary control exercise. The last step, technical edit and approval, is required to ensure accuracy and consistency of subject analysis work, an indispensable condition for effective retrieval. At the end of each step, a record status code is entered.

3. *Retrieval*

CANADIANA is designed for optimal retrieval. Potential access points expected and prescribed by AACR2 for conventional library catalogues are adequately provided: personal and corporate authors, title, indexing terms, ISBN, ISSN, series title, etc. CCOHS practices have gone far beyond these criteria and fulfilled what Pauline Atherton Cochrane was expecting from online retrieval systems: liberation of bibliographic databases from the "notion of main entry and the limitations of the established rules for added entry" (Cochrane 1985). CANADIANA is designed to serve in a main-entry-free environment. The library tradition of providing access to not more than three names is completely abolished. Fields for personal authors and corporate bodies are made repeatable, and procedures are set up to create as many personal and corporate headings as needed for a document. Other additional access points such as series number, report number(s), country of publication, date of publication, and language of publication are also provided. Relations between documents, for example, separate reports of multi-phase studies, different editions of reference materials, and version in other official language, are also given special treatment. They are made known to searchers through two mechanisms: first, a note concerning the existence of the document in the other offical language, and second, a field for the ISN(s) of the record(s) of the related documents.

For subject retrieval, a strong subject index has been designed to contain keys drawn from the title, contents note, descriptors, and abstract fields and to serve as the default search index. It supplements the advantages of the controlled subject access through CIS descriptors with the strength of free text subject access, something in line with the findings of a recent literature review on online subject retrieval. In this review, the author, though personally recognizing advantages and weaknesses of free text and controlled vocabulary, and thus cautious in his recommendations, still thinks that the option to use both is increasingly favoured (Dubois 1987). Besides, as mentioned earlier, because the inversion type used for the descriptor index is both "word" and "term" (that is, phrase), the fear of negative effects of pre-coordinated terms on retrieval is eliminated (Harter 1988). Finally, unlike the traditional bibliographic databases, CANADIANA also gives references to chapters of monographs (mostly reference materials such as encyclopedias, handbooks, and manuals) and to individual papers in conference proceedings.

Conclusion

CANADIANA is an important source of OHS literature by Canadian content, especially those hard-to-find limited-distribution documents by workplace groups and consulting firms. It is the result of quality control efforts made in database design and record production procedures and thus reflects CCOHS' objective of producing and distributing accurate and reliable information, the first step in achieving change for a better working environment for Canadian working people.

References

An Act to establish the Canadian Centre for Occupational Health and Safety. 1978. Ottawa : Queen's Printer for Canada.

Bourne, Charles P. 1977. "Frequency and impact of spelling errors in bibliographic data bases," *Information processing and management* 13: 1-12.

Cochrane, Pauline Atherton. 1985. "Bibliographic data bases : their effect on user interface design in interactive retrieval system," In *Redesign of catalogs and indexes for improved online subject access : selected papers of Pauline A. Cochrane,* 28-34. Tucson, Ariz. : Oryx Press.

Dubois, C.P.R. 1987. "Free text vs. controlled vocabulary : a reasssessment," *Online review* 11: 243-53.

Harter, Spethen P. 1988. "Detrimental effects of searching with precoordinated terms," *Online review* 12: 205-10