CLM-LM 4/72

Preliminary Notes on Visits to a Number of
On-Line Information Retrieval Systems in the
United States and Canada

J L Hall
6 June 1972

Preliminary Notes on Visits

to a Number of On-Line Information Retrieval

Systems in the United States and Canada

by

J L Hall

INTRODUCTION

During a visit to the USA and Canada in May 1972, on-line information retrieval was discussed with 12 system users and developers. "Hands-on" use was made of 11 systems. A full visit report will be completed in due course.

The purpose of this preliminary document is to make available quickly (before the NATO On-Line Institute, 24 July - 5 August 1972, Lyngby, Denmark) some notes on the systems visited.

Also included are details of recent literature available. All the documents cited are held at Culham. Since these total over 300 pages they are not reproduced as attachments to this brief Memorandum; however, copies can be obtained on loan from the Culham collection. (Citation is limited only to recent documents which may not yet (June 1972) be known to some readers of this Note; in many cases, of course, a number of earlier, often published, documents exist.)

UKAEA Research Group
Culham Laboratory,
Abingdon, Berks.

6 June 1972

BRIEF SUMMARY

Entry	Organization	Subject Area	Size* of On-line File	Page
1.	Biosciences Information Service, Philadelphia	Biological sciences	Small (< 3,000 records)	3
(-)	Business Computer Systems Ltd., Ottawa (not visited but SPIRS system used; see Entry No.2)			
2.	Communications Canada, Ottawa	Communications	Small (< 100 records)	3
3.	Council for Exceptional Children, Arlington	Handicapped or gifted children	Small (8,800 records)	4
(-)	Data Corporation, Dayton (not visited but DATA CENTRAL system used from Bethesda; see Entry No.8)			
(-)	IBM, Gaithersburg (not visited but IBM system designed for New York Times seen in New York; see Entry No.11).			
4.	Kingston, Queen's University (not visited but QUIC/LAW used from Ottawa)	Statutes & Court decisions	Large (10^8 characters)	5
5.	Lehigh University, Bethlehem (not visited; used from Philadelphia)	Science & Engineering	Large ($\sim 5 \times 10^5$ records)	5
6.	Library of Congress, Washington	Bill Status	Medium (21,500 records)	6
(-)	Library of Parliament, Ottawa (uses Kingston, Queen's University system; see Entry No. 4)			
(-)	Lockheed Missiles & Space Company (not visited, but DIALOG system used; see Entry No. 3)			

Entry	Organization	Subject Area	Size* of On-line File	Page
7.	Massachusetts Institute of Technology, Cambridge	Materials science	Medium (20,000 records)	7
8.	National Institute of Neurological Diseases and Stroke, Bethesda	Epilepsy	Medium (12,000 records)	8
9.	National Library of Medicine, Bethesda	Medical sciences	Large (425,000 records)	8
10.	National Science Library, Ottawa	Pollution	Medium (46,000 records)	9
11.	New York Times, New York	News and editorial matter	Medium (50,000 records)	10
12.	System Development Corporation, Falls Church. (N.B. SDC programs are also used by the National Library of Medicine, Entry No. 9)	Education	Medium (87,000 records)	10

. . . .

*   <u>Note</u>

Size of on-line file: small, < 10,000 records   :   3 organizations
                        medium, 10,000-100,000   :   6 organizations
                                records
                        large, > 100,000 records   :   3 organizations.

MORE DETAILED ENTRIES

1. Biosciences Information Service of Biological Abstracts (BIOSIS)
    2100 Arch Street, Philadelphia, Pennsylvania 19103, U.S.A.

    Subject area: biological sciences.

    Brief project aim: to develop a tool for query strategy construc-
        tion, queries to be run later in batch mode search of the
        2×10^6 references in the "Biological Abstracts" data base of
        literature from 1959 onwards.

    Project acronym: -

    Data base on-line: a part of the thesaurus is held on-line
        (May 1972).

    On-line retrieval system: being developed in-house.

    User terminal equipment: teletypewriter.

    Computer used: IBM 360/175, at the University of Pennsylvania;
        eventually an IBM 370/145 will be used.

    Status of project: experimental; also used for tutorial sessions
        for new staff joining BIOSIS.

    Literature available: (1.1) "Sample Run of BIOSIS Conversational
        Strategy Building System". 13 April 1972. (8 pages).

    Contact Officer: Louise Schultz.

2. Communications Canada, 100 Metcalfe Street, Ottawa, Ontario KIA OC8,
    Canada

    Subject area: information dealing with the broad field of
        "communications".

    Brief project aim: to provide rapid access to the literature and
        eventually to automate a central microfiche file for display
        via closed-circuit television.

    Project acronym: SPIRS-BIB (Special Purpose Information Retrieval
        System).

    Data base on-line: testing (May 1972) is being carried out on a
        small number of references only; about 1,000 references
        may be available on-line later in 1972 and thereafter about
        5,000 documents may be available within a reasonable time.

On-line retrieval system: SPIRS (currently being developed by
   Business Computer Systems Ltd., Ottawa).

User terminal equipment: teletypewriter-type terminal operating
   at either 10 characters per second or at 30 characters per
   second.

Computer used: Sigma 7, on a leased basis.

Status of project: experimental.

Literature available: (2.1) "An Automated On-Line, Interactive,
   Co-ordinate Index to Information in the Field of Communications".
   (Brief description of system, as at March 1972.) By R.M.
   McMullen. (3 pages).

Contact Officer: Dr R M McMullen, Director, Information Retrieval
   Services Branch.

3.   Council for Exceptional Children, Suite 900, Jefferson Plaza, Building 1,
     1411 South Jefferson Davis Highway, Arlington, Virginia, U.S.A.

   Subject area: exceptional children, either handicapped or gifted.

   Brief project aim: to develop a useful on-line data bank of
      literature relating to exceptional children.

   Project acronym: -

   Data base on-line: 8,800 records (early May 1972).

   On-line retrieval system: based on Lockheed DIALOG system.

   User terminal equipment: visual display unit (CCI 300) with
      linked keyboard (CCI 303).

   Computer used: time is leased on a near-by IBM 370.

   Status of project: operational.

   Literature available: (3.1) "User's Manual. ERIC/DIALOG
      Online Retrieval System". August 1971. (42 pages).

   Contact Officer: Carl F. Oldsen.

4. Kingston, Queen's University, Ontario, Canada.

> Subject area: Canadian Statutes and Supreme Court decisions.
>
> Brief project aim: to construct an on-line interactive information
> retrieval system for searching and displaying the full text
> of documents in a matter of seconds, particularly documents
> in the legal field.
>
> Project acronym: QUIC/LAW.
>
> Data bases on-line: (1) full text of Supreme Court of Canada
> decisions from 1923 to 1970 (87,500,000
> characters),
>
> (2) full text of the 1970 Revised Statutes
> of Canada (11,200,000 characters),
>
> (3) for other data bases using QUIC/LAW
> see National Science Library, Ottawa,
> Entry No. 10.
>
> On-line retrieval system: developed in-house.
>
> User terminal equipment: video display terminals (2000 Data
> Display terminal, Atlantic Technical Corporation, at 2400
> baud) or teletypewriter terminals (at conventional speed).
> Terminals are located from coast-to-coast; one terminal
> is located in the Library of Parliament, Ottawa.
>
> Computer used: IBM 360/65.
>
> Status of project: operational.
>
> Literature available: (4.1) "QUIC/LAW; User's Manual
> (Preliminary Manual)". 1971. (45 pages.)
>
> Contact Officer: Professor H.J. Lawford, Faculty of Law.

5. Lehigh University, Bethlehem, Pennsylvania, U.S.A.

> Subject area: interdisciplinary, particularly science and
> engineering information.
>
> Brief project aim: to provide a user-oriented multi-option on-
> line information retrieval service.
>
> Project acronym: LEADERMART (Lehigh Automatic Device for
> Efficient Retrieval for the MART library).

Data base on-line: titles, authors, bibliographic details,
abstracts, etc. of about half a million articles drawn from
recent COMPENDEX, CAS Condensates, and various other source
tapes.

On-line retrieval system: developed in-house.

User terminal equipment: visual display units (e.g. 2400 baud
Computer Terminal Corporation Data Point 3000; 22 user
groups use CRT equipment); teletypewriters (3 user groups
use teletypewriters).

Computer used: CDC 6400 with a CDC 821 disc store of 10^9
characters.

Status of project: operational.

Literature available: (5.1) "The LEADERMART System and Service".
By A.J. Kasarda and D.J. Hillman. (Preprint of paper to
be presented later in 1972.) (22 pages.)

Contact Officer: Professor D.J. Hillman.

6. Library of Congress, Washington D.C., U.S.A.

Subject area: legislative information.

Brief project aim: an initial aim is the provision of instant
information on Bill status.

Project acronym: -

Data base on-line: 21,469 records (8 May 1972).

On-line retrieval system: developed in-house.

User terminal equipment: IBM 2260 display terminal with a linked
IBM 1053 printer.

Computer used: IBM 360/50.

Status of project: experimental.

Literature available: (6.1) "Computer Support of Congressional
Operations - Selected References: Revised", Document
TK 6565 C, 72-73 SP. By R.L. Chartrand. 15 March 1972.
(7 pages.)

Contact Officer: Robert L. Chartrand.

7. Massachusetts Institute of Technology, Cambridge, Massachusetts, U.S.A.

Subject area: (initially) materials sciences.

Brief project aim: a programme of research and experiment intended to provide a foundation for the design of future information systems - on-line retrieval is seen as likely a necessary component of these future systems.

Project acronym: INTREX (Information Transfer Experiment)

Data base on-line: 20,000 references on-line in computer store. 15,000 references available in microfiche form for display, at terminal station, during a search.

On-line retrieval system: being developed in-house.

User terminal equipment: ARDS (Advanced Remote Display Station) terminal developed in-house. This allows query formulation and also display on the same video screen of full text of article(s) drawn from text held in microfiche form in two CARD units (combined maximum capacity = full text of 15,000 articles so oldest 5,000 microfiche are not held in CARD store and are no longer available for immediate scan at terminal).

Computer used: IBM 7094.

Status of Project: experimental/operational.

Literature available: (7.1) "Project INTREX: a Brief Description". By C.F.J. Overhage. M.I.T. 1972. (Unnumbered booklet, 37 pages).

(7.2) "Project INTREX: Samples of Catalog Interactions". M.I.T. 1971. (Unnumbered document, 28 pages).

Contact Officer: Professor J.F. Reintjes, Electronic Systems Laboratory.

8.   National Institute of Neurological Diseases and Stroke, National
      Institutes of Health, 8600 Rockville Pike, Bethesda, Maryland, U.S.A.

>   Subject area: epilepsy.
>
>   Brief project aim: to make available to interested workers the
>   literature of epilepsy for on-line, free text searching.
>
>   Project acronym: EARS (Epilepsy Abstracts Retrieval System).
>
>   Data base on-line: 12,000 abstracts from 3,000 source journals,
>   essentially the literature contained in "Epilepsy Abstracts"
>   from 1947 to December 1971.
>
>   On-line retrieval system: Data Corporation's DATA CENTRAL.
>
>   User terminal equipment: Computer Communications CC 300 colour
>   TV terminal with a CC 303 keyboard, and a linked CC 305
>   video printer.
>
>   Status of project: operational.
>
>   Literature available: (8.1) "Epilepsy Abstracts Retrieval System
>   (EARS): a Critical Evaluation of a Computer-based Medical
>   Search and Retrieval System". By R.L. Rapport, F.W. Lancaster
>   and J.K. Penry.   (In press; "Postgraduate Medicine", 1972).
>   (Preprint, 15 pages.)
>
>   Contact Officer: Dr J.K. Penry.

9.   National Library of Medicine, Lister Hill National Center for Biomedical
      Communications, 8600 Rockville Pike, Bethesda, Maryland, U.S.A.

>   Subject area: Medical sciences.
>
>   Brief project aim: to provide an on-line bibliographic searching
>   capability for libraries at medical schools, hospitals, and
>   research institutions throughout the USA.
>
>   Project acronym: MEDLINE (MEDLARS On-LINE).
>
>   Data base on-line: 425,000 references (early May 1972).
>
>   On-line retrieval system: SDC's ORBIT II as adapted to the
>   requirements of the Lister Hill Center where the program
>   suite is known as ELHILL II.
>
>   User terminal equipment: normally teletype or typewriter-type
>   terminals operating at 10, 15 or 30 characters per second.
>
>   Computer used: IBM 370/155 at the National Library of Medicine.

Status of project: operational.

Literature available: (9.1) "Operating Guide for On-line
Searchers for MEDLINE (MEDLARS
On-LINE Citation Retrieval Service".
December 1971. (44 pages)

(9.2) "MEDLINE". (statistical and other
details, 20 pages.)

(9.3) Sample printout search. (1 page.)

(9.4) "MEDLINE". NLM Fact Sheet. January
1972 (2 pages).

(9.5) "AIM-TWX". (Brief details, 8 pages.)

Contact Officer: Davis B. McCarn.

N.B. Unfortunately the NLM's AIM-TWX service was not
available on the day of this visit to NLM.

10. National Science Library, National Research Council, 100 Sussex Drive,
Ottawa, Canada.

Subject area: pollution, environmental quality, and water resources
management.

Project acronym: PIP (Pollution Information Project).

Data bases on-line: (1) pollution data base containing 45,984
references (mid-May 1972).

(2) water resources management sample file
containing 2,097 references (mid-May
1972).

On-line retrieval system: QUIC/LAW (for description, see Kingston,
Queen's University; Entry No. 4).

User terminal equipment: video display terminal (2,000 Data Display
terminal, Atlantic Technical Corporation, at 2400 baud).

Computer used: IBM 360/65 at University of Ottawa.

Status of project: (1) pollution data base - operational.

(2) water resources management data base -
experimental.

Literature available: (10.1) "Pollution and Rapid Access to the
Literature". By R.J. Norstrom, C. Quadling, J.R. Marier,
and P. Wolters, Canadian Research & Development, March/April

9

1972 (3 pages).

Contact Officer: Georg Mauerhof.

11. New York Times Information Bank, Times Square, New York, NY 10036, U.S.A.

Subject area: News and editorial matter from the New York Times, plus (eventually) selected material from over 60 other newspapers and periodicals.

Brief project aim: to create a new tool for NYT staff, for outside subscribers, and, in part, to replace the present "morgue" of clippings.

Project acronym: -

Data base on-line: about 50,000 items available on-line (early May 1972).

On-line retrieval system: developed for NYT by IBM, Gaithersburg.

User terminal equipment: IBM 4506 video display unit. The system now under development allows the display (on the same screen as the user dialogue) of the full text of documents stored on microfiche in a Foto-Mem RISAR unit connected to the terminal by co-axial cable.

Computer used: IBM 360/50 with IBM 2314 disc storage and IBM 2321 data cell storage.

Status of project: experimental.

Literature available: (11.1) "NY Times Info Bank to be On-line Late 1972". Information, Pt.1/News, sources, profiles, vol.4, no.2, p.65 (1 page) March-April 1972.

Contact Officer: Dr John Rothman.

12. System Development Corporation, 5827 Columbia Pike, Falls Church, Virginia 22041, U.S.A.

Subject area: education.

Brief project aim: to develop a commercial on-line service in the area of the literature of education.

Project acronym: SDC/ERIC (Educational Resources Information Center).

Data base on-line: 87,168 items dating from 1965 onwards, drawn from ERIC files.

On-line retrieval system: SDC's ORBIT II (On-line Retrieval of Bibliographic Information Time-shared).

User terminal equipment: teletype or typewriter-type terminal operating at 10, 15 or 30 characters per second; will also service CRT terminals.

Computer used: from May 1972 SDC/ERIC will operate on SDC's IBM 370/155 at Santa Monica, California, via the TYMSHARE Corporation's nationwide communications network.

Status of project: operational.

Literature available: (12.1) "Examples of Conversation with SDC/ERIC Search Service". November 1971. (24 pages.)

(12.2) "SDC/ERIC News", Vol.1, Issue 1, February 9, 1972 (2 pages.)

Issue 2, March 17, 1972. (3 pages.)

Issue 3, April 20, 1972. (2 pages.)

Contact Officer: Mary Jane Ruhl, at above address, or Diana DeLanoy, at 2500 Colorado Avenue, Santa Monica, California 90406, U.S.A.

N.B. Search dialogue used in SDC/ERIC is very similar to that used in MEDLINE (for example of this, see National Library of Medicine, Entry No. 9, reference (9.3)).