Special Publication 82-4
ISSN 9275-0929

SECOND INTERNATIONAL CONFERENCE ON GEOLOGICAL INFORMATION

Proceedings: Volume 1

May 23-27, 1982
Colorado School of Mines
Golden, Colorado

Edited and compiled by

Claren M. Kidd

Sponsored by:

Geoscience Information Society
Geological Information Group
of the Geological Society of London
International Union of Geological Sciences
Association of Chief Librarians
of National Geological Surveys
Association of Geoscientists
for International Development

Published by the Oklahoma Geological Survey
Charles J. Mankin, Director
October 1982

Special Publication Series

The Oklahoma Geological Survey's Special Publication series is designed to bring new geologic information to the public in a manner efficient in both time and cost. The material undergoes a minimum of editing and is published for the most part as a final, author-prepared report.

Each publication is numbered according to the year in which it was published and the order of its publication within that year. Gaps in the series occur when a publication has gone out of print or when no applicable publications were issued in that year.

This publication is issued and printed by the Oklahoma Geological Survey as authorized by Title 70, Oklahoma Statutes, 1971, Section 3310, and Title 74, Oklahoma Statutes, 1971, Sections 231-238. 600 copies have been prepared for distribution at a cost to the taxpayers of the State of Oklahoma of $8,200.

Local Organizing Committee:

Robert Bier, Chair
Neely Bostick
Diane Brown
Ralph Farrar
Cheryl Kirk
Mary Larsgaard
Joanne V. Lerud
Barbara MacDonald
Hartley K. Phinney, Jr., Chair
Caryl Shields

Conference Organizing Committee:

Julia H. Bichteler
Claren M. Kidd
Daniel F. Merriam
Hartley K. Phinney
Mary W. Scott
Richard D. Walker
Dederick C. Ward, Chair

Editorial Note--

The papers have been reproduced from camera-ready copy provided by the authors. Only typographical errors were corrected. The views of the authors were not changed or modified and remain those of the author, not the publisher.

Mention of specific companies or of their products or brand names does not imply endorsement or recommendation on the part of the Oklahoma Geological Survey or the sponsoring organizations.

C O N T E N T S

Volume I

Contents--cont.

PROGRAM

SUNDAY 23 MAY

Arthur Lakes Library

18:00-21:00 Registration and icebreaker

MONDAY 24 MAY

Green Center

8:30 Registration

Metals Hall

9:05 D. C. Ward, Call to order, announcements

9:15 Guy T. McBride, Jr., Welcome and opening remarks

SESSION 1: GEOSCIENCE INFORMATION: ISSUES AND PROSPECTS I
Presiding: G. Lea

9:30 T. C. Bearman, International exchange of scientific and technical information--policy issues for the 1980's

10:00 A. P. Harvey, Geological information--past, present, and future

10:30 Coffee

11:00 J. Gravesteijn, International cooperation in geological documentation: why, when, ways, and whereto

11:30 N. J. Pruett, Data on the rocks--a cross section of user needs

12:00 G. Lea, Concluding remarks and general discussion

12:30 Lunch

14:00-15:30 SESSION 2: SYMPOSIUM - GEOSCIENCE INFORMATION FOR DEVELOPING COUNTRIES: PROBLEMS AND PROSPECTS
Presiding: J. W. Green

A. R. Berger, Some recent developments in geoscience information in the Third World

G. N. Rassam, Impact of current developments in information management on developing countries

N. Benveniste, Geoscience Information in Mexico; state-of-the-art

J. V. Hepworth, Information needs in developing countries

15:30 Tea

16:00-17:00 Exhibits (*Green Center*)

Monday evening

Optional: Evening at Coors Brewery (limited to 120 persons)

TUESDAY 25 MAY

Metals Hall

9:00- 10:30 SESSION 3: SYMPOSIUM - ONLINE BIBLIOGRAPHIC DATABASES
Presiding: R. K. Farrar

U. H. Rowell, In house information online: creating small databases

R. K. Farrar and J. LeRud, Online searching using geographical coordinates

A. P. Gotto, Economic geology on GeoArchive

J. G. Mulvihill, Impact of international data exchange on GeoRef

R. D. Walker, The future cost of online retrieval

10:30 Coffee

11:00- 12:30 SESSION 4: SYMPOSIUM - MAPS, MAPPING, AND MANAGEMENT
Presiding: S. Klimley

J. Coombs, The reality of online retrieval of geoscience maps

R. Markham, Current Issues in microforms

A. K. Turner, Digital cartography--a geological perspective

12:30 Lunch

14:00- 15:30 SESSION 5: SYMPOSIUM - THE USE OF DIGITAL DATA IN THE GEOSCIENCES
Presiding: J. F. Lander

F. Chayes, The need for retrospective data bases in the naturalistic sciences, and some general problems associated with them

R. Walton, Digital data processing in the U.S. Geological Survey (USGS)

A. L. Clark, National and international data exchange: problems and opportunities

15:30 Tea

16:00- 17:00 Symposium continued

L. G. Batten (speaker), E. D. Greenlee, D. A. Hastings, S. K. Jenson, and C. M. Trautwein, An overview of the geological implications of spatial data bases

Mike F. Horder, Do's and Don't's of Geological Data Management

18:00 Happy Hour

19:00 Banquet

Presiding: R. A. Brown
Speaker: Robert L. Bates

WEDNESDAY 26 MAY

Green Center

9:00–12:00 SESSION 6: TECHNICAL PAPERS

Room 268

Presiding: J. Bailey

9:00 N. Champigny (speaker), C. I. Godwin, P. R. Tate Blanchet, T. Chen and H. J. Mah, Application of the computer-based Geolog System to exploration and development data

9:30 W. L. Chenoweth, Information availability of the National Uranium Resource Evaluation Program

10:00 A. M. Hittelman, Database design and inventory system for trackline data

Presiding: P. Sheahan

11:00 A. R. Masterson, The role of the information specialist in the (U.S.) state geological surveys

11:30 E. P. Shelley, The information services and systems of the Australian Bureau of Mineral Resource

Metals Hall

Presiding: A. Bearer

9:00 T. Barrington, Humanizing science communication--innovation at SEG

9:30 D. S. Crowe (speaker) and D. A. Tellis, The Australian experience with information transfer in the geosciences

10:00 M. D. du Plessis, Geoscience information in the Geological Survey of South Africa

Presiding: M. J. Farmer

10:30 G. A. Pawloski, Continental Scientific Drilling Program data base

11:00 R. L. S. Weaver, Geological Activities of the World Data Center-A: Glaciology (Snow and Ice)

11:30 L. J. Pieri, Data and its retrieval from the Regional Planetary Image Facilities

12:00 unscheduled

15:00 M. F. Horder, A Summary of Issues and Prospects

15:30 Mary W. Scott, Closing of the technical sessions

WEDNESDAY EVENING

Visits to local technical information centers

SECOND INTERNATIONAL CONFERENCE ON GEOLOGICAL INFORMATION

Attendees

Albert, T. M.
USGS
115 National Center
Reston, VA 22092

Andrews, Martha
Dept. Geol. Science--CU-Boulder
Box 250, Univ. Colorado
Boulder, CO 80309

Anthes, Susan H.
Univ. of Colorado - Boulder
Science Library - Norlin
Boulder, CO 80309

Baclawski, Diane
Geology Library
Michigan State University
E. Lansing, MI 48824

Bailey, Judith
French American Metals Corp.
9580 W. 14th Ave.
Lakewood, CO 80215

Baker, Kay M.
USGS Library
Stop 914, Box 25046, Denver, CO 80225

Baldwin, Joan
Pennzoil Expl. & Prod. Co.
Technical Information Services
P.O. Box 2967
Houston, TX 77252-8200

Barrington, Theodore
Soc. of Exploration Geophysicists
Box 3098
Tulsa, OK 74101

Bartlett, Leah A.
Wichita State University
Library - Box 68
Wichita, KS 67208

Bates, Robert L.
Geology & Mineralogy
Ohio State University
125 South Oval Mall
Columbus, OH 43210

Batten, Lawrence G.
Technicolor Government Service
EROS Data Center
Sioux Falls, SD 57198

Bearer, Annabeth
Phillips Petroleum Co.
439 FPB
Bartlesville, OK 74004

Bearman, Toni Carbo
National Commission of Libraries and Information Science
1717 K. Street, N.W., Suite 601
Washington, D. C. 20036

Benveniste, Nelly
Instituto Mexicano Del Petroleo
EJE Central Lazaro Cardenas 152
Mexico D.F. 03100, Mexico

Berger, Dr. Antony R.
Editor, EPISODES
Room 177, Geological Survey of Canada
601 Booth St.
Ottawa K1A OE8 Canada

Bier, Robert A., Jr.
U.S. Geological Survey Library
Mail Stop 914, Box 25046, Federal Center
Denver, CO 80225

Bonn, Jane
U.S. Geological Survey
P.O. Box 25047 MS 914
Denver, CO 80225

Bording, E. D.
Library
Canadian Superior Oil Ltd.
3 Calgary Place
355-4 Ave. S.W.
Calgary Alberta, Canada T2L 0J6

Bostick, N. H., Mail St. 940
Oil & Gas Resources Branch
U.S. Geological Survey
Federal Center, Box 25046
Denver, CO 80225

Bourgeon, Annette
Geological Survey of Canada
601 Booth St.
Ottawa Ontario K1H 0E8

Brown, Diane
Arthur Lakes Library
Colorado School of Mines
Golden, CO 80401

Brown, Gavin H.
Brown's Geological
160 N. Gower St.
London UK NW1 2ND

Brown, Regina
The Ohio State Univ.
Orton Memorial Library of Geology
155 South Oval Dr.
Columbus, OH 43210

Brown, Sharon
Conoco Inc.
555 17th St.
Denver, CO 80202

Casteel, Kyran
Mining Journal Research Services
15 Wilson St.
London UK EC2M2TR England

Castillo, Estella
Molycorp, Inc.
P.O. Box 54945
Los Angeles, CA 90054

Champigny, Normand, M.A.sc.
Geological Engineer
304 1200 W. Pender
Vancouver, B.C. Canada V6E 2S9

Chayes, Felix
Geophysical Laboratory
2901 Upton St.
Washington, D.C. 20008

Chender, Michael
Metals Economics Group
1722 14th St.
Boulder, CO 80302

Chenoweth, William L.
U.S. Dept. of Energy
P.O. Box 2567
Grand Junction, CO 81502

Clark, Allen L.
Inter. Inst. Res. Dev.
#3 Belvedergasse A-1040
Vienna, Austria

Clark, David
NOAA
Code D623, 325 Broadway
Boulder, CO 80303

Clayton, Keith M.
Geo Abstracts
Regency House, 34, Duke St.
Norwich, NR3 3AP, U.K.

Cockburn, George H.
Energie et Ressources-Quebec
1620 Boul. de l'Entente
Quebec, Que, Canada GOS 3LO

Coombs, Jim
Southwest Missouri State Univ.
SMSU Library, Map Collection
Springfield, MO 65802

Conner, Maura
U.S. Geological Survey
Woods Hole, MA 02543

Courain, Margaret E.
Environmental Data & Inform. Service
NOAA
3300 Whitehaven St., SE.
Washington, D.C. 20235

Crowe, D. S.
Australian Mineral Foundation
P.O. Box 97
Glenside, South Australia 5065

Cunningham, Ann Marie
Institute for Scientific Information
3501 Market Street
Philadelphia, PA 19104

Davidson, Veronica E.
Conoco, Inc
2220 Buckthorne, Suite 201
P.O. Box 4800
The Woodlands, TX 77380

Davie, Robert F.
Saskatchewan Geological Survey
1914 Hamilton St., 12th Floor
Regina, Saskatchewan, Canada S4P 4V4

Deger, Beth
URS Co.
3955 E. Exposition
Denver, CO 80209

Dennis, Kathryn
Watts Griffis and McOuat
159 Bay St., Ste. 911
Toronto, Canada M5J 1J7

Derksen, Charlotte
Branner Earth Sciences Library
Stanford University
Stanford, CA 94305

Dirlam, Dona Mary
Gemological Institute of America
1660 Steward St.
Santa Monica, CA 90404

Du Plessis, M. D.
Geological Survey of South Africa
c/o South African Consulate General
425 Park Avenue
New York, NY 10022

Durotoye, Bisi
atural History Museum
University of IFE, ILE-IFE
Nigeria

Duttenhefer, Annie
Petroleum Information
400 E. Dry Creek Rd.
Littleton, CO 80122

Eaglesfield, Jean T.
Lindgren Lib. M.I.T.
54-200 Mass. Inst. of Technology
Cambridge, MA. 02139

Edwards, Mai
NOAA/NESDC
325 Broadway
Boulder, CO 80303

Ellis, Judith A.
Billiton Exploration, USA, Inc
1625 Broadway #1550
Denver, CO 80202

Ellwood, Elsie E. (Miss)
Librarian
Petroleum Corporation of Jamaica
Box 579
Kingston 10, Jamaica

Engle, Monica
Penrose Library
503 Hilltop Drive
Colorado Springs, CO 80906

Erdmann, Charlotte
South Dakota School of Mines & Tech.
Devereaux Library
Rapid City, SD 57701

Ervin, Sharon C.
634 Carlson
Sycamore, IL 60178

Eva, Margaret
Geology Library
University of Queensland
St. Lucia QLD, Australia 4067

Faber, M. L.
Conoco, Denver West Exploration
P.O. Box 1219
Golden, CO 80402

Fagerstrom, David M.
University of Colorado
University Libraries, Box 184
Boulder, CO 80309

Fair, Rosemary R.
Champlin Petroleum Co.
P.O. Box 1257
Englewood, CO 80150

Farmer, Michael
Shell Oil Co.
P.O. Box 4423
Houston, TX 77210

Farrar, Ralph K.
Marathon Oil Co.
P.O. Box 269
Littleton, CO 80160

Felt, Mary Alice
Anschutz Corp.
555 17th St.
Denver, CO 80202

Fernald, Gail
Information Connections
2355 Balsam Dr.
Boulder, CO 80302

Fraser, Gail L.
Home Oil Co. Ltd. (Library)
325- 8th Ave., SW.
Calgary, Alberta, Canada T2P 2Z5

Fraser, Robert B.
Minerals Management Service
National Center - MS 630
Reston, VA 22091

Freeman, C. M.
Bendix FEC
Box 1569
Grand Junction, CO 81502

Gauss, Nancy V.
Colorado Presevation Office
1300 Broadway
Denver, CO 80027

Gese, Diann D.
U.S. Bureau of Mines - IFOC
Denver Federal Center - Bldg 20
Denver , CO 80225

Goodwin, Jr., George H.
U.S. Geological Survey Library
National Center, Mail Stop 950
Reston, VA 22092

Gotto, Anthea
Geosystems
P.O. Box 573
Cambridge, MA 02139

Gouwens, Jodi
Resource Assoc. Of Alaska, Inc
3230 Airport Way
Fairbanks, AK 99701-5596

Granger, Christian
SNEA (P)-Elf Aquitaine, Tour Franklin
Bureau 2349 - Cedex 11
Paris La Defense, France 92081

Grant, Bruce
NOAA/EDIS
325 Broadway
Boulder, CO 80303

Gravesteijn, Jacobus
Bureau De Recherches Geologiques & Minieres (B.R.G.M.)
B.P. 6009
Orleans, France 45060

Green, Julian
Harvard University
Geol. Sci. Library
24 Oxford St.
Cambridge, MA 02138

Groen, Henk A.
Geoscience Data Centre
Ontario Geological Survey
Rm. 804 - 77 Grenville St.
Toronto Ontario, Canada M5S 1B3

Hafner, Jill
Exploration Library
Superior Oil
P.O. Box 1521
Houston, TX 77001

Hale, Kathryn L.
E. K. Lehmann & Assoc.
1409 Willow St., Suite 300
Minneapolis, MN 55403

Harvey, A.P.
Head, Dept. of Library Services
British Museum (Natural History)
Cromwell Road, London SW7 5BD

Heiser, Lois
Indiana University
Geology Library
Bloomington, IN 47405

Henderson, W. G.
Institute of Geological Sciences
Murchison House, West Mains Rd.
Edinburgh, UK EH9 3LA

Hensinger, Jim
Bibliographical Center for Research
2265 Hanover
Aurora, CO 80010

Hepworth, John V.
ESCAP/RMRDC (United Nations)
JL. Jend. Sudirman 623
Bandung, Indonesia

Hinks, Yvonne
Gallagher Library of Geology
University of Calgary
Calgary, Alberta, Canada T2N 1N4

Hittelman, Allen M.
NOAA
325 Broadway (D-62)
Boulder, CO 80303

Holloway, D.
Charter Consolidated Services Ltd.
Charter House, Park St.
Ashford, Kent England TN24 8EQ

Horder, Dr. M. F.
Selection Trust Limited
Selection Trust Bldg.
Mason's Avenue
London EC2V 5BU

How, Sarah E.
Head, Geology-Geophysics Library
Univ. of California, Los Angeles
Los Angeles, CA 90024

Jones, Ruth M.
Chevron Resources Co.
P.O. Box 7147
San Francisco, CA 94120-7147

Kelly, William J.
Phillips Petroleum Co.
230 Frank Phillips Bldg.
Bartlesville, OK 74004

Kemper, Lyle L.
USGS National Cartographic Inform. Ctr.
Box 25046, M.S. 405, Federal Center
Denver, CO 80225

Kidd, Claren M.
Univ. of Oklahoma
830 Van Vleet Oval, Rm. 103
Norman, OK 73019

Kirk, Cheryl L.
Amoco Production Co.
Amoco Bldg.
17th & Broadway
Denver, CO 80202

Kiswanto, Johanes Baptis
Chief, Documentation Section
Directorate of Environmental Geology
DirGen of Mines/Dept. of Mines
JL Diponegoro 57, Bandung, Indonesia

Klamer, Paul
Metals Economics Group
1722 14th St.
Boulder, CO 80302

Klimley, Susan
Lamont-Doherty Geological Observatory
Palisades, NY 10964

Kruelski, Elaine
Getty Oil Co.
P.O. Box 7900
Salt Lake City, UT 84107

Lacot, Mrs. Claude
Elf Aquitaine (Production)
Tour Generale - Cedex 22
F-92088 Paris La Defense 9, France

Lander, James F.
NOAA/EDIS
3763 Spring Valley Rd.
Boulder, CO 80302

Larkin, Joey (Ms.)
DEPCO, Inc.
1000 Petroleum Bldg.
Denver, CO 80202

Larsgaard, Mary L.
Map Librarian
Arthur Lakes Library
Colorado School of Mines
Golden, CO 80401

Lea, Graham
Geosystems
P.O. Box 1024
Westminster, London SW1, UK

Lerud, Joanne V.
Marathon Oil Co.
P.O. Box 269
Littleton, CO 80160

Lukens, Beatrice
Earth Sciences Library
Univ. of California - Berkeley
Berkeley, CA 94720

Lyttle, Norman A.
Nova Scotia Dept of Mines and Energy
1690 Hollis St.
Halifax, N.S., Canada B3J 2X1

MacDonald, Barbara
Maxima Corp.
200 Union Blvd.
Denver, CO 80202

MacGilliuray, Joe
Alberta Geological Survey
3rd Floor, Terrace Plaza
4445 Calgary Trail South
Edmonton, Alberta, Canada T6H 5R7

Markham, Robert
James Mitchner Library
Univ. Of Northern Colorado
Greeley, CO 80639

McGarr, Michael
Head, Library Information Services
Institution of Mining & Metallurgy
44 Portland Place
London, England W1N 4BR

McGarry, Dorothy
Univ. of California- Los Angeles
P.O. Box 5803
Sherman Oaks, CA 91413

Mamoulides, Aphrodite
Shell Development Co.
P.O. Box 481
Houston, TX 77001

Mansour, A. O.
Geological Survey Egypt
3 Salah Salem St.
Cairo, Egypt Abbassiya - P.O.

Pagain, Marina
AGIP S.p.A.
S. Donato Milanese
Milano, Italy 20097

Marshburn, Peg
Standard Oil Co. of California
20 Westfall Ave.
Oakland, CA 94611

Masterson, Amanda R.
Bureau of Economic Geology
The University of Texas at Austin
Box X, University Station
Austin, TX 78712

Mathews, Karen
MS 586
Los Alamos National Laboratory
P.O. Box 1663
Los Alamos, NM 87544

Mohn, Dr. Joachim
Ideal Basic Industries
Expl. Dept., P.O. Box 1949
Fort Collins, CO 80522

Morningstar, Melody M.
Mobil Producing TX & NM, Inc
9 Greenway Plaza Suite 2700
Houston, TX 77046

Moore, Jeane K.
Ernest K. Lehmann & Assoc. Inc
1409 Willow St. #300
Minneapolis, MN 55403

Morrissett, Elizabeth
Montana College of Mineral Sci. & Technology Library
Butte, Montana 59701

Pawloski, Gayle
Lawrence Livermore Nat'l Laboratory
P.O.Box 808 L-222
Livermore, CA 94550

Pearson, Barbara C.
EXXON Prod. Research Co.
P.O. Box 2180
Houston, TX 77001

Phinney, Jr. Hartley K.
Director of the Library
Arthur Lakes Library
Colorado School of Mines
Golden, CO 80401

Pieri, Leslie J.
Jet Propulsion Laboratory
4800 Oak Grove , MS 264-115
Pasadena, CA 91214

Pinkston, Terry G.
EXXON Production Research Co.
P.O. Box 2189
Houston, TX 77001

Powell, Alan
Elf Aquitaine Oil & Gas
P.O. Box 19810
Houston, TX 77024

Powell, Camille M.
Fluor Engineering and Const.
P.O. Box 35000
Houston, TX 77035

Price, Patricia
University of Hawaii
Hawaii Institute of Geophysics
2525 Correa Rd., HIG 252
Honolulu, Hawaii 96822

Price, Patricia A.
Gulf Oil Expl. & Prod. Co.
P.O. Box 36506
Houston, TX 77036

Pruett, Nancy J.
Technical Library
Sandia National Laboratories
Albuquerque, NM 87185

Ramberg, Ellen Marie
Univ. of Oslo
Dept. of Geology, Box 1047
Blindern, Oslo 3, Norway

Ramsden, Anne
Inst. of Geological Sciences
Exhibition Rd.
London UK SW7 2DE

Rassam, G. N.
GeoRef.
Falls Church, VA 22040

Ritzma, Barbara R.
University of Utah
Marriott Library
Salt Lake City, UT 84112

Rowell, Unni Havem
Exmin Corporation
P.O. Box 2655
Bloomington, IN 47402

Schimming, Sue
Petroventures, Inc
Suite 600
4704 Harlan St.
Denver, CO 80212

Scott, Mary W.
Univ. of North Dakota
Chester Fritz Library
Grand Forks, ND 58202

Sharp, Alice L.
Colorado Historical Society
1300 Broadway
Denver, CO 80203

Sheahan, Patricia
SELCO, Inc.
55 University Ave.
Toronto, Ontario Canada MSJ 2H7

Sheaves, Miriam L.
Geology Library
Mitchell Hall 029A
University of North Carolina
Chapel Hill, NC 27514

Shelley, E. Paul
Bureau of Mineral Resources
P.O. Box 378
Canberra City, ACT Australia 2601

Shields, Caryl L.
Library
USGS, MS914
Box 25046
Denver, CO 80225

Shilts, Dorothy D.
Wichita Scouting Co-op. Inc
Oil Information Library
813 Hamilton Bldg.
Wichita Falls, TX 76301

Sinnott, Gertrude M.
U.S. Geological Survey Library
National Center, MS 950
Reston, VA 22092

Sorensen, Janice
Kansas Geological Survey
1930 Avenue "A", Campus West
University of Kansas
Lawrence, KS 66044-3896

Spencer F.L.A., K. J.
Chief Librarian
Inst. of Geological Sciences
Exhibition Rd.
London England SW7 2DE

Spilhaus, Fred
American Geophysical Union
2000 Florida Ave., NW.
Washington D.C. 20009

Stair, Fred G.
Cities Service Co.
P.O. Box 3908
Tulsa, OK 74102

Stark, Marilyn
Reference Librarian
Arthur Lakes Library
Colorado School of Mines
Golden, CO 80401

Stevenson, John J.
U.S. Bureau of Mines - IFOC
Bldg. 20, Denver Federal Center
Denver, CO 80225

Struble, Dorothy
Phillips Petroleum Co.
7800 E. Dorado Place
Englewood, CO 80111

Summers, Kenneth W.
Superior Oil Co., Minerals Div.
P.O. Box 13628
Tucson, AZ 85732

Svetlik, Susan
Library, Bedford Inst. of Oceanography
P.O. Box 1006
Dartmouth, N.S. Canada B2Y 4A2

Thompson, Shirley
Shell Oil Co.
P.O. Box 4423
Houston, TX 77210

Tompkins, Dorothy L.
Billiton Exploration USA, Inc.
1625 Broadway, Suite 1550
Denver, CO 80202

Turner, A. Keith
Geology Department
Colorado School of Mines
Golden, CO 80401

Valenciano, Dominador B.
Bureau of Energy Development
Merritt Rd., Fort Bonifacio
Metro Manila, Philippines

Van Roggen, Dr. J. F. Graadt
Geology Dept., Anglo American
P.O. Box 16587
Marshalltown, 2107 (Johannesburg) R.S.A.

Venkataraman, Dr. S.
Science Division/Main Library
Univ. of British Columbia
1956 Main Mall
Vancouver, B.C. Canada V6T 1Y3

Wagner, Barbara L.
10185 West 25th Ave., #23
Denver, CO 80215

Walcott, Rosalind
Earth & Space Sciences Library
S.U.N.Y. at Stony Brook
Stony Brook, NY 11794

Walker, Richard
Univ. of Wisconsin
600 N. Park St.
Madison, WI 53706

Walton, Ron
USGS
P.O. Box 25046
Denver, CO 80225

Waranius, Fran
Lunar & Planetary Inst.
3303 NASA Road One
Houston, TX 77058

Ward, Dederick
Geology Library, Univ. of Illinois
223 Natural History Bldg.
1301 West Green St.
Urbana, IL 61801

Weaver, Ron
World Data Center A
CIRES, Univ. of Colorado
Box 449
Boulder, CO 80309

Welsh, Harry E.
Devereaux Library
South Dakota School of Mines & Tech.
Rapid City, SD 57701

Wesley, Jannette Sharp
USGS Oil & Gas Branch
MS 971, Box 25046
Denver, CO 80225

Wilkins, Eleanore E.
U.S. Geological Survey Library
345 Middlefield Rd., MS 55
Menlo Park, CA 94025

Zaccaria, Maria
AGIP S.p.A.
S. Donato Milanese
(Milano) Italy 20097

Zar, Kathleen
University of Chicago Library
1100 East 57th St.
Chicago, IL 60637

Zipp, Louise S.
Geology Library
University of Iowa
Iowa City, IA 52242

Supplement 1

(Nonattendees)

Co-authors

Susan Gallanter
TechSearch
9989 West 60th Avenue
Arvada, CO 80004

William F. Graser, Chief
Technical Information, NCIC
USGS Box 25046
Federal Center, MS 504
Denver, CO 80225

Dr. Frank T. Manheim
U.S. Geological Survey
Woods Hole, Mass. 02543

Members of GIS Organizing Committee

Julie H. Bichteler
Graduate School of Library Science
The University of Texas at Austin
Box 7576, University Station
Austin, TX 78712

Daniel F. Merriam, Chairman
Department of Geology
Wichita State University
Wichita, KS 67208

Canadian Liaison

C. F. Burk, Jr.
Office of the Auditor General
240 Sparks Street
Ottawa, Ontario Canada K1A 0G6

FOREWORD

Persons responsible for providing information or the accessibility to geoscience information whether they are called information specialists (usually in industry) or librarians (usually in public libraries or in academia) are scattered around the world. The services and products that they provide may be verbal, online, or in hard copy. To make services and products more attractive and accessible, the producers of information must consult with intermediaries or end-users who are the information specialists and librarians.

Following a successful first meeting held in 1978 in London, conference organizers decided to hold a second meeting in North America. The Colorado School of Mines campus at Golden, Colorado, was selected as our meeting site because of its geologically interesting setting in the foothills of the Rocky Mountains, willing volunteers in the energy rich Denver metropolitan area, and the inexpensive University housing and services.

Approximately 170 delegates from 16 countries listened to 40 papers on the diverse aspects of international geological information, such as information needs in developing countries, digital mapping, online information systems, the role of information specialists/librarians in regional geological surveys, selection of geoscience journals, and the concept of information brokerage.

International working groups on library cooperation, national resources inventory, and editing/publishing/translations were organized at the conference as the result of discussion held among delegates. Collegiality was everywhere evident and paramount to the success of the meeting. The participants left the Colorado foothills rich in new friendships and with the invitation to a future conference in Adelaide, Australia in 1986.

Claren M. Kidd

Claren M. Kidd

INTERNATIONAL EXCHANGE OF SCIENTIFIC AND TECHNICAL INFORMATION

POLICY ISSUES FOR THE 1980'S

Toni Carbo Bearman

National Commission on Libraries and Information Science
1717 K Street, N.W., Suite 601, Washington, D.C. 20036

Abstract--Three developments over the past 10-15 years have complicated the already complex international scientific and technical information issues: (1) technological advances, (2) protection of proprietary and privacy rights, and (3) the growing awareness of scientific and technical information as a valuable resource. Regarding the third point, the debate continues as to whether this information is a unique and valuable resource that is not consumed but possibly enhanced by use, or whether it is a commodity to be traded, taxed, and sold. The questions of whether the flow of information should be regulated or restricted and, if so, by whom, relate directly to this disagreement.

Extensive issues related to scientific and technical information policy include privacy, restraint on scientific and technical publication, transborder data flows, public/private sector interactions, international reciprocity, proprietary rights questions (such as copyright), national security concerns, and cultural and social issues. These issues logically fall into two major groups: (1) economic/technological, and (2) social, political, and cultural.

There is a constant struggle to achieve a reasonable balance between the need to protect national security and proprietary rights, on the one hand, and the principle of unrestricted dissemination of scientific and technical information, on the other. We in the scientific, technical, and information communities need to work closely with our public officials and elected representatives to achieve this balance.

Introduction

I am very pleased to be back in Colorado and honored to have been invited to speak at the Second International Conference on Geological Information at this excellent school. I am delighted to see participants from so many different countries, and I would like to extend a special thanks to our Canadian colleagues who have given up their Victoria Day holiday to be with us today.

Because I have learned that many of you are not familiar with the National Commission on Libraries and Information Science, I would like to take a moment to describe NCLIS to you.

(1) NCLIS serves as resident expert for all branches of the government in the library/information field. With information activities comprising almost half of the Gross National Product, this is a crucial area of expertise.

(2) As an independent agency, it serves as an honest broker bringing together individuals from all branches of government with those at the state and local levels and with representatives from the private sector.

(3) Through its extensive network of contacts with individuals and associations in the library/information field, it is able to provide a forum for the entire community; and

(4) It serves as a catalyst identifying problems, suggesting solutions, and making things happen.

As part of its ongoing process to carry out these roles, the Commission established priorities for 1982. The major priorities set are:

- The development of specifications for revised library and information services legislation;

- Improved dissemination of Federal information; and

- Resource sharing and application of technology.

In carrying out another of its mandated responsibilities, the Commission is studying the informational needs of rural Americans by working closely with the National Agricultural Library and also by giving support and guidance to the program of the Intermountain Community Learning Information Center Project. This is a grassroots activity, essentially dependent on the state and local governments and the state extension services in Colorado, Montana, Utah, and Wyoming. The Commission also is exploring several other areas, including the identification of the library and information needs of senior citizens and persons living in remote areas, and the development of programs to meet those needs efficiently and effectively; and the examination of the changing role of public libraries and the development of ways to assist them in harnessing the new technology, increasing cooperation with the private sector, and improving services by increasing the participation of volunteers for special programs, such as combatting illiteracy.

The Commission staff is now concentrating on eight issue papers to be presented at the June meeting of the Commission. The topics for the issue papers are: (1) Changing Role of Public Libraries; (2) Urban Libraries; (3) Literacy; (4) Rural America; (5) Aging/Senior Citizens; (6) Impact of Information Technology; (7) Information and Productivity; and (8) International Activities. After reviewing these papers and discussing other issues, the Commission will establish its priorities for fiscal year 1983 and 1984.

Let me turn now to the policy issues confronting us in the 1980's. Although I have worked in the field of Scientific and Technical Information for more than 20 years, I am not a geologist. Also, because the international policy issues for geological information are the same as those for STI more broadly my remarks today--which are my own and do not necessarily represent those of the National Commission or of the U.S. Government--will relate to broad scientific and technical information policy issues for the 1980's.

International scientific and technological and information issues are not new. Concerns such as intellectual freedom and access to documents and information have been around for a long time. Over the past 10-15 years, three developments--(1) technological advances; (2) protection of proprietary and privacy rights; and (3) the growing awareness of scientific and technological information as a valuable resource--have further complicated the already complex and interrelated international information issues.

Advances in computer and telecommunications technologies have given rise to large-scale information networks that routinely transmit computerized data on virtually all subjects across national borders. Second, the protection of proprietary and privacy rights has become increasingly difficult with the ease of copying data and documents and with the growing amount of data about individuals and organizations stored in computers.

Related to these developments is a growing awareness of information as a valuable resource, or, some would argue, as a commodity. The debate concerning whether information is a resource or a commodity continues, with one side arguing that information is a unique and valuable resource that is not consumed (and may, in fact, be enhanced) by use. As a resource, it should be nurtured, preserved, managed, and made available--usually for a price. The other side argues that information, like gold or hog bellies, is a commodity to be traded, taxed, and sold. Its cost and price may vary significantly. The questions of whether the flow of information should be regulated or restricted and, if so, by whom, relate directly to this disagreement.

Characteristics of the Information Environment

Developments in technology have led us from an agricultural-based society through an industrialized era to what is now popularly called the Information Age. A brief review of the characteristics of the information environment in which the U.S. is a key player provides a framework for the consideration of international scientific and technical information issues. The list I have compiled is based on the one in the Rockefeller report on National Information Policy published by NCLIS in 1976 (U.S. Domestic Council Committee on the Right of Privacy). These characteristics are:

(1) An expanded increase in the volume of information flow;

(2) A shrinkage of time and distance constraints upon communication;

(3) Greater nationwide dependence upon information and communication services;

(4) An increase in the interdependence of previously autonomous institutions and services;

(5) Conceptual changes in economic, social, and political processes induced by increased information and communications. (To cite just two examples: The projected impact of a "checkless/cashless society" as a result of electronic fund transfer, and possible impacts on institutions of interactive home delivery of information systems, such as Qube);

(6) A decrease in the "time cushion" between social and technical changes and their impact and consequences; and

(7) Global shrinkage and its consequent pressures on increased international information exchange. (One example is the Swedish law regulating personal information, which impacts on multinational corporations with personal data held in computer files in Sweden.)

Overview of the Issues

Among the extensive list of issues related to scientific and technological information policy are the following: privacy, restraint on scientific and technical publication, transborder data flows, public/private sector interactions, international reciprocity, proprietary rights questions (such as copyright), national security concerns, and cultural and social issues.

These issues logically fall into two major groups: (1) economic/technological; and (2) social, political and cultural. In most discussions of economic/technological issues, the terms "cost," "price," and "value" are used without clear definitions or distinctions among them. It is important to distinguish carefully among these three concepts. The determination of costs is a basic economic decision needed for effective management. The pricing of products and services, although clearly related to economic concerns, may also be a political decision, especially in the public sector. Determining the value of a product or service may extend far beyond economic questions to philosophical considerations of quality, or ethical issues, such as the notion of "public good."

Economic Technological Issues

Economic/technological issues include: (1) the imposition of taxes or tariffs on the transfer of information internationally, with the likely monitoring of that information by the taxing authority; (2) pricing of telecommunications services by government monopolies at far above cost; (3) the imposition of inconsistent or narrowly interpreted technical standards upon both providers and users of telecommunications and data processing services; (4) requirements that "domestic" information and transactions be processed in the home country; (5) the denial of entry for a foreign enterprise into a nation's market; and (6) restrictions on the transmission of information across national boundaries because its content reflects a wide variety of social, cultural, and political concerns which relate to economic considerations. These issues represent barriers to the transfer of information and data across national borders, frequently referred to as "transborder data flow."

These barriers create disruptions in operations, reduce competitiveness, and may even totally eliminate the ability to compete in some markets. Another obvious impact of these barriers is on products and services. For users in some countries, it costs much less to search a database on that country's system than it does on a foreign system, or to search a database developed in that country instead of a foreign database. Clearly, these tariffs have an impact on the revenues for services.

Political, Social and Cultural Issues

Political, social and cultural concerns include privacy, the protection of proprietary rights, the preservation of national sovereignty, the protection of national security, and the desire to assure access to information held in databases outside a nation's boundaries but vital to its national development.

Concern for the preservation of national sovereignty can result in a broad range of restrictions on the content of data transmission. It may cause a government to determine that it should monitor the content of information flow through telecommunication channels or prohibit, for economic or cultural reasons, the dissemination of information originating in a foreign country. For example, the Clyne report (Consultative Committee on the Implications of Telecommunications for Canadian Sovereignty, 1979) from Canada recommended that "the government should act to regulate transborder data flows to ensure that we do not lose control of information vital to the maintenance of national sovereignty." Canada is understandably sensitive to this issue because of its relatively small population (25 million, less than the state of California), and the dominance of its information and other industries bv the U.S. National sovereignty concerns of the less-developed countries relates to their dilemma of trying to acquire the modern technological tools to assist in the development process while seeking to break away from their dependence on the industrialized countries. As Ricardo Saur (1979) points out, they fear they will be unable to participate as equals in the world political and economic community unless they can "develop a <u>local</u> capacity to avoid dependency on outside sources for computer power . . . to confine data flow <u>within</u> borders, and . . . to provide protection for those local development efforts." The movement to provide a "New World Information Order" relates directly to national sovereignty concerns. This movement results from a concern by some countries that the U.S. dominance of mass media and international communications can be viewed as a form of "cultural imperialism." For example, developing countries have recently sought to acquire greater allocations of the radio frequency spectrum as a means of countering the predominance of the U.S. in the telecommunications capabilities.

The protection of national security goes beyond the well-publicized concern for military security and relates closely to preservation of national sovereignty. Many barriers to information flow have been erected because of fears of "vulnerability." The storage of data outside the country, the need for ready access to critical information from a foreign location, the flooding of a country with inaccurate or misleading information are three preceived threats to a nation's ability to protect its culture, society and political structure. All of these issues have existed for a number of years;

however, the development of telecommunications and computer technologies has significantly increased the threat of foreign manipulation.

Developed countries share these concerns. For example, the Canadians view the problems of international information flow primarily in terms of the impact that the storage and processing of information in foreign computers--particularly those owned and operated in the U.S.--may have on their development and political authority. To them the major problem is:

> that data processing and communications business may be lost to Canadians as a result of the foreign flow; that [Canadian] data in the U.S. databanks may be pre-emptorily [sic] withheld [from return to Canada], that the U.S. laws may change and leave Canadians less protected; and that, as a sovereign state, Canada feels some national embarrassment and resentment over increasing quantities of sensitive data about Canadians being stored in a foreign country. (Turn, 1979.)

An increasing number of countries worry that the loss of control over information about individuals, companies, and internal functions could jeopardize their sovereignty and national security, leaving them open to possible disruptions.

Assuming that information is a resource, and considering the increasing restraints on natural resources worldwide, some nations are beginning to question whether stronger controls should be placed on the transfer of technology and technical information. The U.S. in particular, is concerned that it is giving away more than it is getting. More than six bills have been introduced in the U.S. Congress related to reciprocity--seeking to increase the return the U.S. gets for sharing its information resources. Some critics contend that we give away valuable information to nations which copy our techniques while avoiding the costs of research and development, then market similar products at lower prices, thereby undercutting U.S. products.

Such sentiments, however, overlook the international character of modern technological development. Many high-technology firms have research and development labs in Europe, North America, and Japan. These facilities are dependent on the open and unrestricted flow of information to carry out their tasks. A recent Congressional study contends:

> The cost of barriers to the international flow of otherwise available technical information does not have a figure. We cannot measure the effect of impediments to innovation, to the growth of productivity, or to the increase of knowledge.
>
> On the other hand, legitimate national defense arguments for limited restrictions can be made. The flow of high technology to the U.S.S.R. has been an area of debate for many years. Recent cancellation of export licenses as a result of Soviet aggression reflects the judgement that providing technical information, services, and equipment to certain foreign governments threatens national defense. (U.S. House of Representatives Committee on Government Operations, 1980.)

Erosion of indigenous cultures has been caused by the introduction of foreign databases and the extensive broadcasting of news, entertainment programs, and advertising which originates in other countries. Foreign information services can bring valuable scientific and technical information to remote areas, but they can also intrude on local social and cultural values. These concerns are not limited to the Third World--for example, Canadians also feel threatened by the implications of the information age for their own national identity and culture.

Even though these social, cultural and political concerns are more difficult to define and quantify, they are equally important elements of national policies related to international library/information issues.

In the past scientific and technical information has been treated differently. In general, few restrictions have been placed on it. In a recent article in Science, Dr. Anna D. Harrison (1982) expressed it well:

> It is generally accepted that the free exchange of scientific know-how and scientific knowledge is a basic tenet of the scientific community. At the same time, it is well known that an individual may create barriers to retard the flow of information within an institution, that an institution may create barriers to inhibit the flow of information to other institutions, and that a nation may regulate the exchange of information with another nation. Herein lie very thorny issues, the resolution of which has a great deal to do with the environment in which the scientific community proceeds. Actions to restrict flow of information reflect the perceived short-term advantage of proprietary scientific know-how and knowledge but discount our mutual dependency and the long-term benefits of the free exchange of scientific knowledge. The long-term costs of mistrust and stagnation are significant indeed. Three types of situations that justify careful consideration are the relation of national security considerations to the free exchange of scientific information with Soviet scientists, the relation of the proprietary interests of industry to the free flow of information within and from academic laboratories, and the impact of the practices of individual scientists upon their colleagues and students.

Here in the U.S. we are seeing an increasing number of attempts to restrict the flow of STI.

> Item: The Department of Energy issued an order requiring government clearance of any communication between its contractors and Soviet scientists.

> Item: The Commerce Department forced the American Vacuum Society to withdraw its invitation to Soviet bloc scientists to attend an international conference on magnetic bubble devices.

> Item: The State Department sent letters to university science departments asking them to restrict the movements of Chinese students and visitors (Nelkin, 1982).

Item: The Reagan Administration has rejected continued government funding of U.S. membership in the International Institute for Applied Systems Analysis (IIASA). A major factor in the decision, according to a recent Science article, is "the belief that the benefits of IIASA membership to the U.S. are much less than those to the Soviet Union and also that U.S. security interests could be compromised." (Science, 1982.)

Item: A major U.S. for-profit database producer was told it could not get an export license to send computer tapes of published bibliography information to its subscribers because the tapes, themselves, are "high technology."

I cite these examples from the U.S. to give an indication of some of the concerns being raised by the U.S. government and to indicate some actions being taken. Understanding the complexities of the issues facing us and the difficulties of developing and implementing appropriate policies are challenging to say the least. There is a constant struggle to achieve a reasonable balance between the needs to protect national security and proprietary rights, on the one hand, and the principle of unrestricted dissemination of scientific and technical information, on the other. It is essential that individuals knowledgeable about STI participate in making information policies. We in the scientific, technical, and information communities need to work closely together with our public officials and elected representatives to achieve this balance, to ensure that the issues are understood, and that appropriate policies are developed. We must not allow the principle of scientific freedom to be forgotten in discussions of our national policies.

Thank you.

References Cited

Canada. Consultative Committee on the Implications of Telecommunications for Canadian Sovereignty [the Clyne report], 1979, Telecommunications and Canada: Ottawa, the Committee.

Harrison, A. J., 1982, Reflections on current issues in science and technology: Science, v. 215 (Feb. 26), p. 1061.

Nelkin, Dorothy, 1982, Intellectual property: The control of scientific information: Science, v. 216 (May 14), p. 706.

Saur, Ricardo, A. C., 1979, Protection without protectionism: Journal of Communication, v. 29, p. 140.

Science, 1982, Lack of reciprocity prompts IIASA cutoff: Science, v. 216 (Apr. 2), p. 35.

Turn, Rein, editor, 1979, Transborder data flows: Concerns in privacy protection and free flow of information: Arlington, Virginia, Report of AFIPS Panel on Transborder Data Flow, v. 1, p. 5.

U.S. Domestic Council Committee on the Right of Privacy, 1976, National information policy: Report to the President of the United States: Washington, D.C., NCLIS, p. 182-183.

U.S. House of Representatives, Committee on Government Operations, 1980, International information flow: Forging a new framework: Washington, D.C., U.S. Government Printing Office, House Report 96-1535, p. 23.

GEOLOGICAL INFORMATION

PAST, PRESENT AND FUTURE

Anthony P. Harvey

Department of Library Services
British Museum (Natural History)
London SW7 5BD

Abstract--Information, whether geological or not, is of no value unless it can be communicated. The channels of communication in geology, as in all science, form a complicated web - interwoven and interactive. The major medium of communication for geology has been ink on paper; initially as the written word but since the beginning of the nineteenth century as the printed word -whether in book or journal form. There is at present little evidence that this channel of communication is about to lose its premier position. The secondary channels of communication, into which may be placed for example publishing, indexing and abstracting services and libraries, do show considerable differences over time especially in the last twenty years. It is these areas which appear to be likely to lead the way in the "electronic information age" we are now in. Although the technology exists (or soon will) for dramatic changes in the organization of geological information and its communication, it will be professional and economic forces which will determine the shape of the future system.

This paper attempts to isolate some threads from the tapestry of science communication, with special reference to geology, and present them with the aim of providing discussion points which can be developed or discarded during the course of the Second International Conference on Geological Information. They may also provide some focal points for the conference as a whole. It is not the intention to offer solutions.

The title includes the word 'information', a word which means 'all things to all men'. However, information, whatever definition is used is of no value unless it is in a form which can be communicated. Information exists to be used. If the topic of information is hard to define then to examine in detail the nature of science communication is even harder. It is therefore, the intention to concentrate on a few specific areas:

i) the scientific, and in particular the geological community

ii) the dissemination of information by whatever means, but paying particular attention to the role of publishing and publishers

iii) the retrieval of information especially through the medium of the secondary services (i.e. those devoted to abstracting and indexing the literature) and their on-line databases; and the role of libraries.

As a line from a very famous lyric bids 'let's begin at the very beginning'. In the mid-eighteenth century geology, like science itself, was in its infancy. There were no professionals, relatively few publications and the main channel of communication was the hand written document.

This era came to an end with the dawning of the nineteenth century, although geology was still in the hands of amateurs, gifted though many of them undoubtedly were. The end of this first phase came with the establishment of specialized societies which had as one of their several aims the publication of the results of members researches. Such societies provided vital focal points for the development of geology. The corpus of published information began to develop in spite of the difficulties often encountered in actually undertaking the investigations, and sometimes the even greater difficulty of getting the results published. For example, many authors paid for their own works to be printed or invited subscribers.

However, one third of the way through the nineteenth century the pace of science development began to quicken. The population of the western world as a whole was also increasing thus providing a larger base of potential personnel. University education was improving, governments were becoming increasingly involved with geology. All of these developments were to cause geology to move away from an amateur based science to a professional one. In turn this was to lead to a fairly rapid increase in all types of publication. Libraries too had begun to move from the private to the public domain.

Although a German secondary service (<u>Taschenbuch fur die Gesamte Mineralogie</u>) had begun in the early years of the nineteenth century, the great increase in literature in the middle decades led to the Geological Society of London establishing an indexing service based on accessions to its library. Initially this was printed in the <u>Proceedings</u> (1851-1894) and later separately as <u>Geological literature added to the Geological Society's Library</u>. Indeed there is a direct link between this and the current <u>Bibliography and Index of Geology</u>.

It should be noted that all these developments were going on increasingly on a global, as opposed to just a European, scale. The foundations laid down for the scientific communication system in the nineteenth century were basically adequate to last until the second half of the twentieth century. Certainly new societies were formed (some older ones even died), new periodicals came into being and the scientific community continued to grow as did the availability of professional employment. But there was relatively little further fragmentation of geology from the major divisions established in Victorian times.

To turn now to the post war period. One palaeontologist has said of the history of his particular department in a national museum that it may be 'characterized as 200 years of relative stability, and 20 years of profound change' (Ball, 1979). These last twenty years correspond to the beginning of the third phase in the evolution of scientific communication; that is the beginning of the electronic information era.

The 1960s witnessed a rapid increase in the number of individuals involved in science and to this trend geology was no exception. Educational establishments flourished, new literature gushed forth but no matter library budgets seemed to keep pace.

The growth of the geological profession, coupled with the seemingly crazy publish or perish syndrome (perhaps those trying to cope with the increasing information flow would say publish and perish) could only lead in one direction, to a growth in the literature. Perhaps more importantly in a literature which was becoming increasingly fragmented as the science itself developed new sub-fields. The professional societies long dominant as publishers found that their financial resources were generally inadequate to capitalize on all these new areas. They also lacked the forceful promotion machinery to ensure sales (Harvey, 1978).

Hence the rise of the commercial giants and, perhaps a little unfairly, the publishing of scientific information in the period 1960-1980 would, in the estimation of many, be adequately summarized by a paragraph from a recent article by William Kay (1982) on Robert Maxwell: 'Pergamon is based on Maxwell's brilliant idea of creating monopolies in abstruse academic journals which universities find they have to buy to keep up with the subject. Pergamon can charge over £2000 a year ...(for some series)... while its contributors tend to be professors who are content to accept little payment for the honour of having their articles published before their deadly rival half-way round the world'.

One interesting development has been the substantial rise in the number of geological (and in particular palaeontological) newsletters. These often represent the visible portion of an invisible college but perhaps also show, through their informality, a certain frustration with the more formal world of publishing, both primary and secondary. Certainly for those groups which have access to computers, this type of publication would be ideally suited to the electronic medium.

Perhaps the greatest changes in the post-war period have been in the abstracting and indexing services. The rise in the amount of literature conveniently occured at the same time as the rapid development of computer technology. This was fortunate for science otherwise the community would now not just be awash in a sea of information but sunk beneath the waves. Of course the use of computers in the printing industry also moved forward but the economies there were not as high as had been hoped for initially, but the base created may form the foundation on which to build in the future.

Libraries and librarians have not so far received much attention in this paper but this should not be taken to indicate a lack of their importance within the communication system. The point is that relatively they have changed very little. They have become larger, use up a greater percentage of the parent institutions funds and space than ever before, but basically they remain much as they were in the nineteenth century. Computer technology when introduced has largely been hidden from users by being concerned with 'house-keeping' routines. However, the two areas of greatest impact have been in on-line searching and in the development of network cataloguing on either local, national or international scales.

To summarize the present scientific communication system is Victorian in basis. It is now being put under considerable strain and its inefficiencies and inadequacies are becoming increasingly apparent.

What of the future? Most prophets do not have a successful record but it is certain that there will be changes within the science communication system. If this is doubted then consider the following:

i) the population of the developed nations is set not only to decline but also to age thus reducing the potential size of the base of practitioners

ii) world economy is still basically in recession; for example examine the current financial cuts universities in the United Kingdom are experiencing. Nearly all libraries report having major funding cuts in real terms

iii) technology continues to develop at what some might regard as a frightening place.

Just as the secondary services changed in the 1960s and 1970s so will the primary publishing system change in the 1980s. It will do so for one simple reason, publishers are in business to make money and they do not care in what format they work. But changes will come because of the relative inefficiency of the present system and the fact that the cost of its component parts (e.g. paper and mailing) are all increasing whilst in the alternative electronic based system costs are, relatively speaking, falling. Publishers are becoming increasingly alarmed at the extent of photocopying and inter-library loans. In other words they want to safeguard their property and make maximum gain out of their investment. What courses are open to them?

With the increasing availability of full text on magnetic tape the trend is likely to follow the path taken by the secondary services and the one major system being developed is certainly heading in that direction.

ADONIS (Articles Delivery Over Network Information Service) is a document delivery system which is due to become operational in 1984. It is based on images of scientific, technical and medical journal article pages, including illustrations, which are stored on optical reading discs. The consortium of publishers involved include Elsevier, Blackwell Scientific, Pergamon, Springer Verlag, John Wiley and Academic Press. This system is open to others to join and it is estimated that by the time it becomes available more than 3000 journals will be represented on the discs.

Two other developments which offer pointers to the future are: the active work being undertaken by the United Kingdom Printing Industry Research Association in evaluating, through a group of publishers, the opportunities which may be beneficial to the publishing industry from advances in both technology and telecommunications systems; and the placing on a data base of the full text of 31 McGraw-Hill technical and business journals.

However, the journal will be with us for many years yet, if indeed it does ever entirely disappear, as will be the printed book. A fear, increasingly expressed, however, is that with escalating costs, publishers will become unwilling (or unable) to publish in convential form new very specialized periodicals and books. What effect would this trend have on the developments of geology as a whole?

One area where change will be rapid will be in the reference book field, where for example, Gale Research have their Directory of Associations available on-line. Many other publishers are now holding the information contained in directories, bibliographies and dictionaries in a form which makes them ideal for on-line searching. There is no doubt that electronic publishing is highly desirable and effective in these areas.

Within the publishing system itself, editors are usually regarded as a conservative group but even they are already using on-line bibliographic searching to identify possible duplicate submissions, or to discover the track record of a potential author. Certainly authors themselves will need to come to terms with word processors.

It now seems highly unlikely that there will be a microform publishing revolution or that in geology the synopsis journal will be well received and supported by the community. It is only necessary to give the example of the Bulletin of the Geological Society of America. Because of the backlog of papers it commenced in 1980 to be issued in synopsis form with full text back-up on microfiche.

Such was the reaction of potential authors that the flow of papers almost ceased and in 1982 the synopsis idea was abandoned. This is strange because an analysis of reading habits by scientists would seem to indicate the synopsis medium as an ideal one. Other forces must be at work and it may be suspected that the root cause of dislike is that it does not help the publish and be promoted syndrome.

A final point in this section relates to discussions with publishers. There is little evidence that major publishers are expecting to go out of business indeed they seem to be quietly flourishing. Certainly commissioning editors in major houses still have targets of between 30 and 40 books in a subject area per annum in addition to perhaps several new journal titles.

Secondary services will continue to develop and there will be an increasing trend away from printed copy sales to on-line searching. This in turn will mean higher on-line costs. The reasons for this trend are not hard to perceive: libraries gain access to a wide variety of data bases many of which they could never justify or afford in the hard copy versions; payment is by use; and increasingly the data bases contain more information than is available in the published service. Further, individual geoscientists were often reluctant to use the printed abstracting and indexing services whereas they seem to much appreciate the flexibility and ease of on-line searching.

Data banks will also continue to develop with a greater availability by direct access than ever before. An excellent review of geological source and reference data bases has been provided by Burk (1982). There will be a continuing fragmentation of the retrieval system as smaller units seek to capitalize on any unique information stores they hold by selling them to the community at large. Further libraries and individual users will be accessing these systems using not dumb terminals but micro-computers; thus allowing a more user friendly approach to the various systems and the ability to store and reformat the information received. These will be important changes leading in many instances to a greater use of such data bases by individual geologists rather than just through libraries.

What will be interesting to observe is what effect the development of publishers data bases will have on the secondary services. There is no doubt that the present system of on-line bibliographical data base searching can be most frustrating. The references may take fifteen minutes to retrieve and the actual papers six weeks or more through conventional libraries and inter-library loan systems. In discussion at the first International Conference on Geological Information there was a feeling that the secondary services should become increasingly able to retrieve facts as well as citations and that the direction would be to full text retrieval. Perhaps this particular line of development will now be taken up by the publishers themselves rather than by the secondary services or is there likely to be a merging between these publisher data bases and bibliographic data bases? Does Robert Maxwell with Pergamon-Infoline show the way? This is a UK host system owned by a very larger publisher and certainly Maxwell (1982) himself believes that in the future emphasis will be on full text retrieval.

What of libraries in the future? (cf Lancaster, Drasgow and Marks, 1980; Drake, 1980). Librarians will be forced to change at least some of their philosophy. They will need to come to realize that information is going to cost more than it has in the past. Librarians will still need, at least those in charge of large collections of archival material, to conserve and make available the older books and journals, for it should be remembered that many facets of geological research still rely heavily on access to this older literature. Conservation will be of increasing concern, especially because of the financial implications, and may well result in the transfer of information from paper to a medium such as the videodisc. There will be a change of emphasis from a holdings policy (except perhaps in major national and subject libraries) to an access policy. Cooperation and coordination between libraries will be increasingly evident. International cooperation will increase through such systems as the Online Computer Library Center (OCLC). Librarians will need to be conscious, more than ever before, of the fact that they are part of a service industry and if they fail to accurately estimate what service is required and its appropriate level of funding, they will be heading for serious problems. Because of the fragmenting of communication channels librarians will retain their important position as intermediaries between information creators and users and the information itself.

There will be interesting developments in the non-academic libraries. Indeed, in many industrial libraries, holdings of actual books and journals are already down to a minimum level. Librarians in these areas could do much to increase their status (and the desirability of their library to the parent organization) if they become the keepers of the unique store of company information to create an in-house data base. If this in-house data base can then be capitalized on further by selling the information outside the parent company then the library is even more secure. Even within academic establishments, librarians should try to become the centre of their own organizations information network and act as a focal point for information requirements.

Three factors will be of importance in determining the speed of the main areas of change which have been outlined above - they are professional considerations, economics and technology.

There is little doubt that a great deal of today's publishing is to satisfy the system which has developed to support academic promotions. However, in accepting, or fighting the new systems, the geological profession will have to decide what it wants and what it is prepared to accept. For example, it certainly seems unwilling to accept, at present, synopsis publishing. It will need to carefully determine which parts of the highly fragmented communication system are appropriate for information in different facets of geology (cf Line, 1980). It will be necessary to re-examine the needs of the creators and users of information and to understand the differences between them for there is an area of potential conflict in accepting the electronic medium. This is because those who demand instant information are less likely to publish and it is those who do publish who are likely to put up the greatest resistance to change. Certainly there is now an urgent need to evaluate, over the whole geological spectrum, the nature of the material, the likely audience and the use to be made of the information before the present communication system is irretrievably destroyed.

What all of us must strive to ensure is that changes in the primary communication system do not lead to a deterioration of the quality of information. This threat of quality deterioration was one continually voiced by European editors in the earth and life sciences at their recent conference entitled 'The scientific editor in the electronic age'. With the introduction of the new technologies is also the fear for the integrity of the archival copy, because of the ease with which it can be altered. Such considerations are particularly important where matters of priority are concerned.

Information will cost more and there will be a shift of costs so that the purveyor of information invests less in the initial product and sells as needed. Both creators and users will pay more. Publishers and data base managers will be looking to earn as much revenue as possible. Publishers will certainly be doing all they can to see a tightening in the laws regarding photocopying and devising systems of rapid document delivery to replace inter-library loans. Librarians will become more cost conscious and all of their systems will be examined in detail leading to cut backs in acquisitions, binding and their allotment of building space (Kent, 1978). Already in many cases book budgets have been savagely attacked to enable journals to be purchased. All other library processes will be carefully analyzed and costed. Libraries will also seek to capitalize on any unique information they hold.

Cut backs in institutional funding will influence the introduction of new technology. Where speed of access to information is important the investment will be made, where speed is not important, such investment will not be forthcoming. This will inevitably cause problems within the communication system.

If a technologist and particularly one in the computer field, was summing up the chances are that information creators, users and providers would be promised utopia. Advances will indeed be great during the next few years and what is more the next generation of geologists will have been brought up on home computers, teletext, etc. (Mathews, 1978). Perhaps mention of three recent developments will serve to show a little of what is in store: magnetic bubble memory will greatly enhance the storage of vast quantities of data; there is already available a portable micro-computer which is capable of storing 15-20 pages of text and displaying it line at a time; and the development of teletext offers the opportunity for standards to be developed.

Maybe the introduction of such technological wonders will not be as rapid as is often thought largely due to different standards and the need to wait for telecommunications systems to catch up. Perhaps it will be no bad thing if there is an odd touch on the brakes otherwise there is a real danger that the communication system will become dehumanized. However, it is likely that technology will press on rapidly simply because market forces will prevail.

Geology is both an historic and an international science which has many different facets all of which have different information requirements. For many decades yet a great deal of its published literature will remain valid and used. Not only that in many parts of the world it is still cheaper to hand-set type; whilst in many others there is no use having an electronic device because there is just no power system to plug into.

Care must be taken not to injure the whole communication system by by widening the gap between the information 'haves' and the information 'have nots'. It was A.K. Kent (1980) who after a fairly gloomy review of the future wrote: 'Perhaps the only possibility for ensuring that my scenario turns out to suffer the fate of all prophecy is for the major creators, processors and users of information to set together to understand each other's problems and develop together a common strategy for development and advance. We are all in the same boat together. If technology and the opportunity it offers for the future causes us to run before we can walk then none of us may pass the finishing line'.

Could there be any better reason for an international conference which will, it is to be hoped lay down some aims for geological communication and help develop the strategy and the tactics for realizing them.

Acknowledgements

I am indebted to many of the delegates and speakers to the recent conference, entitled 'The scientific editor in the electronic age' (Fifth European Life Science Editors' Association / European Association of Earth Science Editors General Assembly, Pau, France, 11-14 May 1982) for valuable discussions.

The proceedings of the meeting are to be published by The European Space Agency. My many friends in the publishing industry I thank for sharing with me their thoughts on the future.

I also acknowledge the help given to me by the Aslib library staff. In any such review many sources are used, the major ones are cited below, but to those interested in primary communication, I would recommend a new bibliography edited by Margaret Mann and published by the Primary Communications Research Centre, Leicester University and entitled Primary communications: a review of literature.

Also a most useful periodical for keeping up-to-date is Outlook on Research Libraries published by Elsevier Sequoia.

References Cited

Ball, H. W., 1979, The evolution of a national collection: Special Papers in Palaeontology 22, p. 49-56.

Burk, C. F., 1982, A worldwide list of source and reference databases in the geosciences: Database 5, p. 11-21.

Drake, M. A., 1980, The environment for special libraries in the 1980s: Special Libraries 71, p. 509-518.

Harvey, A. P., 1978, A history of geological serial publication in the United Kingdom: Earth & Life Science Editing 6, p. 7-12.

Kay, W., 1982, The resurrection of Robert Maxwell: The Sunday Times 4 April, p. 57.

Kent, A., 1978, Crystal gazing into the future: Journal of Library Automation 11, p. 329-337.

Kent, A. K., 1980, Scientific and technical publishing in the 1980s, in Hills, P. (editor), The future of the printed word: London, Pinter, p. 163-169.

Lancaster, F. W., Drasgow, L. S. and Marks, E. B., 1980, The role of the library in an electronic society, in Lancaster, F. W (editor), Proceedings of the 1979 clinic on library applications of data processing: Urbana-Champaign, University of Illinois, Graduate School of Library Science, p. 162-191.

Line, M. B., 1980, Some questions concerning the unprinted word, in Hills, P. (editor), The future of the printed word: London, Pinter, p. 27-35.

Mathews, W. D., 1978, Advances in electronic technologies: Journal of Library Automation 11, p. 299-307.

Maxwell, R., 1982, Data retrieval systems and international publishing: Aslib Proceedings 34, p. 38-46.

INTERNATIONAL COOPERATION IN GEOLOGICAL DOCUMENTATION : WHY, WHEN, WAYS AND WHERETO

J. Gravesteijn
Bureau de Recherches Géologiques et Minières
B.P. 6009
45060 Orléans Cedex (France)

G.N. Rassam
GeoRef/American Geological Institute
5205 Leesburg Pike
Falls Church, VA. 22041 (USA)

Abstract - Several cooperation projects have been undertaken in the last decades under the auspices of the International Union of Geological Sciences through its Documentation Commission or Advisory Board for Publication. For example : the preparation of review articles, the creation of Episodes, the setting up of the joint IUGS/ICSU AB Working Group on a Multilingual Thesaurus in Geology.

However, facing a steadily growing volume of published information, increasing cost and largely unsatisfied demand of the scientific community, international cooperation is a challenge which will be met only when conditions for a successful operation are set.

The technological revolution in terms of computer storage capacity and telecommunications makes it easier to consider and implement worldwide cooperation in building bibliographic data bases and information tools ready to meet the needs of research and industry.

The strength of this type of cooperation is the fact that individual national centers act under the auspices of or on behalf of an international body thus combining the scientific platform of the international institution and the operational capacity and detailed knowledge of the participating organizations.

The cooperative agreement between the American Geological Institute and the European network for geological documentation connects the GeoRef data base and the PASCAL-GEODE data base and is an important step into the direction of a better integration of national and international efforts in geological information handling.

GEOINFORM the East European project for cooperation on geological information between CMEA countries forms another example of practical cooperation.

In the next decade the development of cooperation programs in the field of geological documentation will be marked or influenced by several factors :

- National and international copyright legislation related to computer applications
- Integration of primary and secondary information
- Integration of factual and bibliographic information handling
- Technological development
- Balance between international data bases and national data bases
- Cooperation between the developing countries and the industrial information world.

But above all, the efficiency of international cooperation will depend upon mutual understanding, a well-defined common information philosophy, much good will and confidence between the partners involved.

Introduction

During the last two decades, international scientific cooperation has been increasing in volume, gaining in efficiency and contributing to a better understanding of scientific and technological problems.

The scale of cooperation varies from individual exchange of knowledge to implementation of internationally conducted projects.

The International Union of Geological Sciences (IUGS) played in many cases a catalytic or operational role together with other international organizations, such as UNESCO.

The International Geological Correlation Program (IGCP) is a good example of international cooperation dealing with fundamental problems and the "Lithosphere Project" will have a significant impact on a better understanding of the geological processes and the building of the continents.

IUGS activities

In the field of geological documentation - restricting this area to bibliographic information handling - international cooperation and exchange have existed since journals, books and bibliographies have been printed.

Some examples of cooperative projects implemented by the Advisory Board for Publications (ADP) and the Commission for Geological Documentation (COGEODOC) may illustrate the IUGS policy and practical approach in bibliographic information handling.

Through its Advisory Board for Publications, the IUGS participated actively in the creation of regionally organized associations of earth sciences editors : North and South America, Europe and Asia have their own associations. The activities vary from one continent to another depending essentially on the cultural, economic and technical environment of earth science editing in the area covered by the regional association.

The North American organization is concentrating its activity on editorial practice, copyright and marketing problems. The South American focuses more on linguistic terminological problems and in Europe, Editerra is especially involved in standardization, terminology and nomenclature.

Acting as editor itself, through ADP, IUGS offers its series of "IUGS publications" for dissemination of research results conducted within the Union's domain of responsibility.

"Episodes" and its predecessor, the "IUGS Newsletter", were managed in the beginning by the IUGS Secretariat, disseminating news from the Union and its commissions.

For several years - actually since the IUGS Secretariat moved to Canada - Episodes has played an important role as an international communication medium for gathering news, review articles, congress proceedings, etc.. and offering a platform for an exchange of ideas and scientific results.

The Commission for Storage, Processing and Retrieval of Geological Data (COGEODATA) deal with standardization and coordination of factual geological data handling and exchange but this activity is out of the scope of this paper.

The Commission for Geological Documentation (COGEODOC) covers the broad field of bibliographic information handling. Its objective is to promote internationally better access to published literature. Its activities include improvement of documentation tools, cooperation between existing documentation centers and production of review articles.

- Set up in 1964, the Commission tackled first the problem of review articles. There was a strong need for this type of literature and practically none was being produced. Until the early seventies, international review articles were published at the occasion of IUGS symposia or colloquia.

This activity slowed down and has now practically disappeared for several reasons :

* the need to concentrate more energy on the computer processing of bibliographic material. The dispersion of efforts in this field could be dangerous and a coordination could save time and stimulate cooperation projects,

* the appearance of several national or international journals specializing in review articles and it is not the aim of an IUGS Commission to enter into competition with existing publications.

- In 1972, the Commission decided together with the International Council of Scientific Unions - Abstracting Board (ICSU AB) to sponsor a joint Working Group on a Multilingual Thesaurus in Geology. Ten languages are represented : English (USA and Canada), French, German, Russian and Spanish form the core of the Multilingual Thesaurus but other language groups are participating namely, Czech, Finnish, Hungarian, Italian and Swedish specialists prepared their versions of the Multilingual Thesaurus. Negotiations are underway with Turkish and Arabic groups to include also these languages.

The objective of the joint IUGS - ISCU AB Working Group is to provide a switching mecanism between different systems operating in different languages. The aim is not to build a classic multilingual thesaurus showing one overall hierarchical thesaurus structure for all languages but to create a link between thesauri used in various monolingual systems.

Such a multilingual tool is of course important in itself and essential for the interconnection of bibliographic data bases operating in different linguistic environments. But even more important is the opportunity to discuss during the working sessions of the international group, general indexing problems and experiences of individual documentation centers. This led to a better mutual understanding and a progressive assimilation of indexing and processing problems occuring in the various centers. Two examples illustrate the impact of the dialog between the documentation centers on practical international cooperation :

* The European network in geological documentation, grouping geological surveys in eight countries : France, Germany, Spain, Czecoslovakia, Finland, Hungary, Poland and Romania, contributes to the PASCAL-GEODE data base and its printed bibliography "Bulletin signalétique - Bibliographie des sciences de la Terre". The group came to an agreement with the American Geological Institute, responsible for the GeoRef data base and "Bibliography and Index of Geology" to implement a joint international data base continuing both GeoRef and PASCAL-GEODE. This cooperation project eliminates duplication, improves coverage of both separate data bases and will later facilitate searching for the end user.

  This project was not the direct concern of COGEODOC but could only be implemented because the information managers of both data bases had the opportunity to discuss the practical indexing problems related to the Multilingual Thesaurus project. From these discussions it was learned that the gaps between the two indexing systems were more formal than fundamental. Bilateral discussions made it then possible to build a common indexing system and classification scheme. Exchange of American and European material is now operational showing a practical outcome of what was started as a methodological exercice.

* The second example is the methodological approach chosen by the participants in the GEOINFORM project to build a common indexing system.

  GEOINFORM was created in 1978 as a subsystem of the International System of Scientific and Technical Information (ISSTI), a cooperation project set up by the member countries of the Commission for Mutual Economic Aid (CMEA). Bulgaria, Czechoslovakia, German Democratic Republic, Hungary, Mongolia, Poland, USSR and Yugoslavia participate in the GEOINFORM program.

Managed by national committees, an executive board and a technical and scientific council, GEOINFORM has the following objectives :

- Eliminating duplication of efforts
- Developing national geological information systems in the members countries
- Developing automated information system both for bibliographic and factual data.

Activities are focusing on :

- Preparation of bibliographies and review articles in the field of geology
- Development of a common thesaurus
- Standardization of bibliographic description and format
- Creation of a joint data base.

Experimental sharing of input (indexing, processing) is planned for 1982 and the operational stage of a common system and a joint data base is to be implemented in 1985.

The multilingual thesaurus of geosciences, which will serve as the common GEOINFORM indexing tool, is based upon the same principles as the IUGS/ ICSU AB multilingual thesaurus : distinction of subfields in the earth sciences, flexibility in the structure of the national thesauri and no rigid hierarchical structure for the international thesaurus which is meant to be a switching mechanism between the national indexing systems using various languages.

Links between both multilingual indexing tools are not yet operational but these attempts to cooperate beyond a national level carry the germs for a further integration of international cooperation in data base construction.

International cooperation : Why

In the past most of the international projects were related to the preparation and publication of specialized bibliographies. Unfortunately these operations were mostly characterized by an unacceptable time lag between the publication of the bibliographic tool and the referenced documents. So, it is not surprising that the major international geoscience bibliographies were produced by national organizations. In France, Germany, UK, USSR and USA the documentation centers covering the earth science literature started all their bibliographic activities during the last 50 years. It was quite logical that the implementation of computerized bibliographic data bases was initiated by the same organizations.

However, in an expending information world, international cooperation is an excellent card to play in order to overcome the problems of comprehensive coverage of the literature and elimination of overlap between data bases.

All geological data bases cover theoretically the same population of documents. Actually, overlap in coverage can be estimated at a rate of 60 % - 80 %. It is also obvious that an American center has a better coverage of the American literature and that an English, a French or a German center is likely to receive more literature published in Western Europe. In the same way a Russian center deals more comprehensively with East European literature.

The effects of both factors, - overlap and better coverage by a regional center, - in addition to the steadily growing volume of published and unpublished material serve to induce international cooperation. Increasing costs and unsatisfied demand of the international scientific community transform the management of such a cooperative system or network into an economic and technical challenge for the existing geological information centers.

International cooperation : When

For information handling, the introduction of computer technology was a revolution, then became a constraint and is now developing into a producer -and user-friendly direction.

As the efficiency of information handling is directly related to computer storage capacity and access facilities, the scientific community had to wait until the last decade before it could fully use the possibilities of the electronic revolution.

It is noteworthy that the major international cooperation programs on information systems as well as exchange between data base producers started in the Seventies. Most experiments could only be implemented through computer applications resolving format and capacity problems and creating transfer conditions adapted to operational needs.

As in most cases, things happen when the appropriate time has come. Try to have your six month old baby walk and you will discover that the baby will not do its first steps before he or she is ready physically to do so. In international cooperation, try to force the decision and you will be disappointed and miss the objective. When economic and technical conditions are fulfilled you reach more or less easily an agreement and the project can be successful. In the Sixties, there was a need for review articles and this work could by that time only be accomplished by an international body.

The technological revolution in terms of computer storage capacity and telecommunications makes it easier to consider and implement worldwide cooperation in building bibliographic data bases and information tools ready to meet the needs of research and industry.

International cooperation : Ways

As stated earlier, cooperation in geological documentation may vary from off-prints exchange between individual scientists, through institutional exchange between libraries or documentation centers, to large scale projects involving the production of data bases. In all cases efficiency relies upon willingness and user-oriented operational effectiveness.

The strength of the type of cooperation calling for much operational manpower, for example the building of a multilingual thesaurus or in the case of publishing activities, lies in the fact that individual national centers act under the auspices or on behalf of an international body, thus combining the scientific platform of the international institution and the operational capacities and detailed knowledge of the participating national organizations.

However, care must be given to the balance between national and international input as to avoid political, economic or technical problems.

International cooperation : Whereto

In the next decade, the development of international cooperation programs in the field of geological documentation will be marked or influenced by several factors.

National and international copyright legislation related to computer storage

At present the legal situation is confusing.

No country provides a clear copyright law in terms of input and output of information stored in a computer device. In many countries, national committees or commissions are dealing with the problem and UNESCO is preparing guidelines for national legislators. But the legal base for input of references, indexing terms and abstracts is still under discussion and several cases have been brought to court.

On the output side, the protection of data base producers against piracy becomes more urgent as high-speed transmission of bibliographic data is technically possible and economically cost-effective.

An international consensus has to be reached and a balance to be found between protection and free flow of information, taking into account the interest of authors of primary publications, data base producers and end users. A divergent approach and pressure by the various professional groups or a floating development in national legislations will certainly have a negative impact on dissemination of information in general and on international cooperation in particular.

Integration of primary and secondary information

Among the bottlenecks in information flow, document delivery is still a major issue.

Computer and telecommunication technology resolved a great deal of the problems related to indexing and data base organization. The use of data bases is still increasing and, in the industrialized world, access to information on existing documents has been brought much closer to the user.

Libraries and documentation centers still play a prominent role in supplying primary documents on demand.

Copyright on reproduction, missing journals, distance between the user and the appropriate library, cost and delay in time, are problems which are restraining an efficient use of available information, especially in the case of international networks.

However, with respect to document delivery, electronic technology will also help to resolve the problems. Experiments with Digital Optical Recording (DOR) are encouraging and the study undertaken by six of the major publishing houses in the world promises a technical solution within three years. The huge storage capacity of this type of disk (50,000 compacted pages on a disk) and the combined efforts of British Library Lending Division (BLLD) and the publishers could show the way for an easy retrieval of primary documents, resolving in the same time the issue of the copyright fees due for reproduction.

Of course, it is not a romance and the ideal solution is not for tomorrow. But no doubt this technological development will impact heavily on library activities, data base production and publishing. Some of the existing conflicts will be resolved but others will appear. Only a good contact between the partners in the information chain on the international scene will avoid disturbance in the information flow.

Integration of factual and bibliographic information handling

Geological bibliographic data base are well established, factual data base are being implemented but used only on a small scale. Management of factual data bases is expensive and cooperation is more difficult because of data confidentiality. But the demand exists : a user of an information system wants not only references but also drilling sections, water levels, geotechnical measurements or data on mineral resources.

International public exchange of data will probably be limited as the demand will essentially concern local, regional or national data.

Extrapolating earlier experiments these integrated information systems grouping factual and bibliographic data should be heavily decentralized. In many cases regional centers operating mini-or micro-computers will provide the information to local and regional users. A national organization for processing all collected data is certainly desirable but probably unrealistic in the long run.

Technological development

The importance of development of storage devices have already been mentioned. High-speed transmission networks will further facilitate transfer of large information masses and foster international cooperation. Micro-computer equipment make further decentralization of document analysis and input easier and will probably contribute also to the development of smaller specialized or regional data bases extracted from the larger field-oriented data bases offering also a possibility for repackaging information originating from related fields.

Balance between international and national data bases

The international data bases are here defined as current files containing available information on documents published throughout the world gathered by an individual center or through a network. National data bases can be considered as the information compiled from all publicly-available documents related to the geology of a country. These national data bases contain mostly more information concerning gray literature and cover a longer period of time.

It is in this area that international cooperation cannot play fully its role unless the national system is compatible with the indexing philosophy used by the international data base. This was the way the partners in the European network for geological documentation have chosen. For example, the French national data base goes back to 1755 whereas the international PASCAL-GEODE data base contains bibliographic information on the world literature since 1968.

Special attention will be given in this conference to the developing countries but at this point we can already state that there is a need in these countries for national and international information. More than half of the available information on the geology of these countries is missing in the international data base.

Reasons are often restricted distribution of reports, ignorance of existing journals and absence of publicity. Creation of national centers for geological documentation could serve to complete the widely-available international information.

User habits

The past, present and future were described from a producer's point of view. By no means should we forget the user. The tools and the technological development are there. But are they adequate and are they being used ? The answer is probably "Yes" in the sense of increasing usage by young geoscientists (at least in the industrial world). And "No" in the sense that too few geoscientists are using the available mass of information. Too many individuals rely still on off-prints and on the invisible college of the scientific conferences. This type of information transfer is certainly important but much more is available and at the present state of technological and economic conditions access is relatively easy. A constant dialog between producers and users, especially the potentiel users, will certainly help to improve the information flow.

Conclusion

Many factors mark and influence international cooperation. Efforts should be encouraged when they attempt to improve the various aspects of geological information.

But above all efficiency of international cooperation will depend upon mutual understanding, a well-defined common information philosophy, much goodwill and confidence between the partners involved.

We are entering in an information society. Information becomes the raw material of the eighties. It will be the task of the information specialists to move from a stage of providing information for geoscientists into a situation where informed geoscientists demand better information.

References Cited

Rassam, G.N., Gravesteijn, J., and Lagarde E., 1980-1981, A joint bibliographic data base in geology : an experiment whose time has come: Bull. B.R.G.M., section IV, n° 3, pp. 215-221.

Glashoff H., 1980-1981, The multilingual Thesaurus of Geology, a switching tool between bibliographic data bases developped by IUGS and ICSU AB: Bull. B.R.G.M., section IV, pp. 201-206.

DATA ON THE ROCKS

A CROSS-SECTION OF USER NEEDS

Nancy Jones Pruett

Sandia National Laboratories

Technical Library, Organization 3144, Albuquerque, NM 87185

Abstract--Geoscientists show communication patterns of both scientists and engineers. However, the nature of the geoscience field means that some specific information needs are different from those of other scientists and these needs have implications for information services. Geoscientists often apply basic physics and chemistry to geological problems and need to have access to a broader range of journals. Geological information stays relevant for longer periods of time than in other sciences. Much geoscience data is space-dependent, time-dependent or both, and is thus different from the simpler numerical data of physics or chemistry. So handling "data on the rocks" requires different techniques and display formats.

Introduction

In any discussion of information services it is important to touch base with the needs of the users. Published information about the information needs of geoscientists specifically is scarce. And much of what is published is descriptive and doesn't give us much help in predicting or understanding behavior. So it is helpful to look at some of the literature of scientific communication behavior and the information behavior of engineers. Then we will consider what is unique about geoscientists. Finally, we'll discuss the implications of this behavior for information services.

Although the material about communication behavior of scientists and engineers is probably internationally applicable, the implications I draw for information services are based on my own experience which is largely in library information services in the United States.

Scientist's Information Behavior

I have observed that scientists constantly point out that the information libraries can supply isn't current enough. But when we propose ways of speeding up the publication process by doing away with the review stage or putting everything online, they object. Why is this? And why is it that library users rarely include the scientists who are well established in their fields?

W. D. Garvey's work cumulated in Communication: The Essence of Science (1979), shows that it isn't anything our information services are doing wrong, and that the scientist's behavior is consistent from his or her point of view.

In order to understand the scientist's view, let's look at a time scale for a typical publication process (Figure 1).

Note that it is about 28 months from the initiation of a work to its publication in a journal. The communication during that 2-1/2 years is in the informal domain. Once the work is published in a journal, it enters the formal domain. It is public, unambiguously retrievable and carries with it the stamp of approval of the scientist's peers. "Scientists behave quite differently and apply different sets of standards when communicating in one or the other domain." (Garvey, 1979, p.22) Let's look at the two domains and the function of the scientific journal in more detail.

The Informal Domain

There is a much larger volume of communication before publication than there is after publication. Prepublication communication includes personal correspondence, conversations and presentations at professional meetings, drafts that are circulated for comments, etc. The communication is an exchange and at this phase it is under the control of the scientist. He or she will get feedback on the same work many times, so communication is also redundant. It is also unstable, in that the work may look different in various communications. The informal domain includes the communication which actually enters into the creation of scientific information, and is thus of the most interest to the scientist. Information which he or she seeks in this phase will only be accepted if it is personally relevant.

The informal domain is indeed relevant to geoscientists. Paleontologist Richard Bambach in discussing how he keeps current listed "keeping up with colleagues" as the most important of four approaches.

> "C. P. Snow's concept of the invisible university is a meaningful one. There is a network of people that are active in forefront research in almost every field. They seem to keep in touch with each other by correspondence, phone and personal visits. These people are doing much of the ground-breaking in their disciplines and, because of their need for information that is not yet common published knowledge, they seek and supply knowledge within their active research cadre.
>
> "I try to keep in touch with the active scholars in my fields of interest. In our exchanges I 'pick' their brains and I trust they 'pick' mine. People in other environments than your own do different things and pay attention to different things than you do at home. Every campus or laboratory visit, speaking engagement, hosting of a visiting scholar, or scientific meeting produces new ideas and new directions to search. The remark 'did you know so-and-so just discovered such-and-such' is about as current as you can get unless you do it yourself." (Bambach, 1981, p. 17)

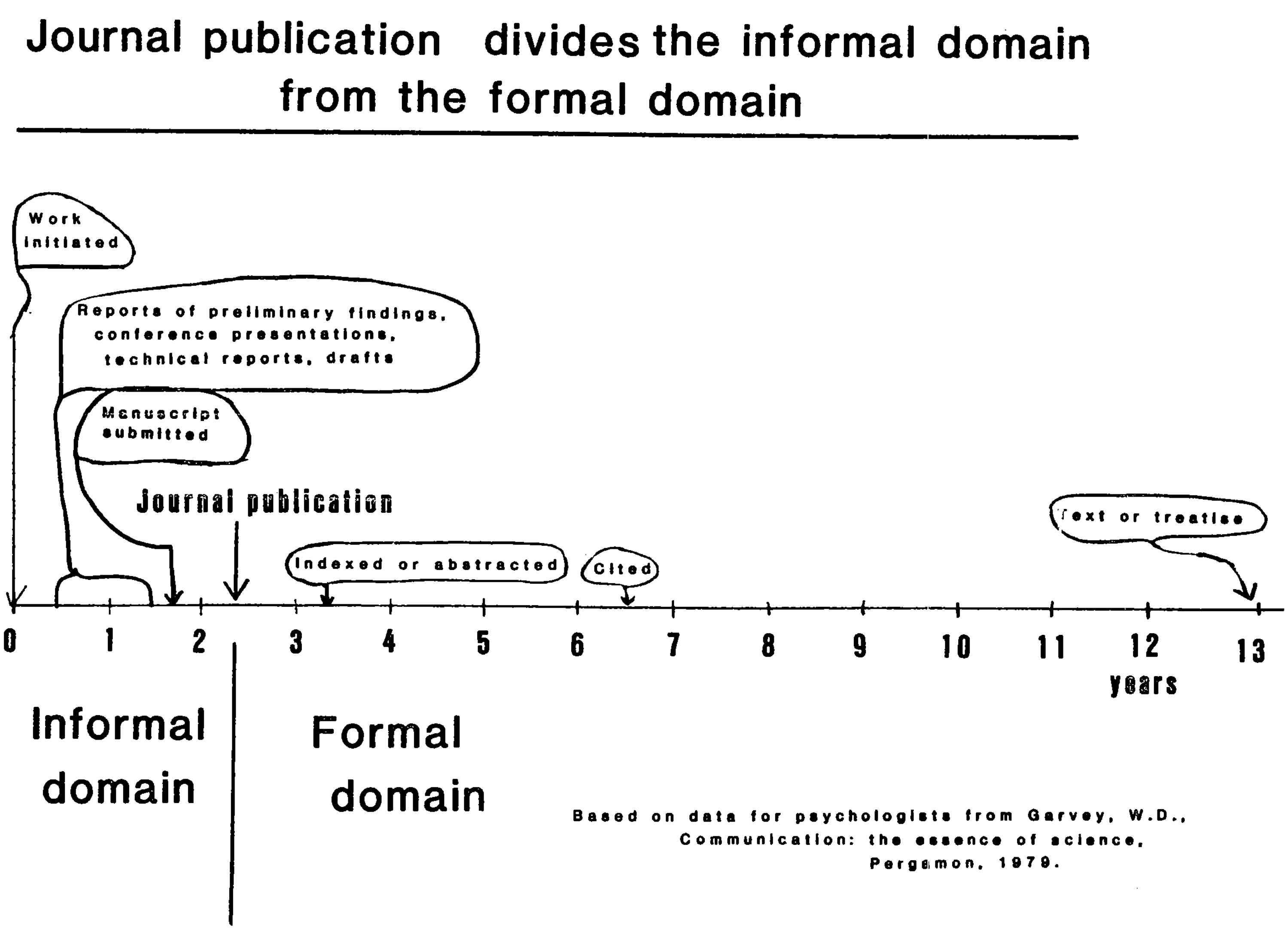


Figure 1: A time scale for the publication process.

The types of information included in the informal domain are: conference presentations, preprints, technical reports, drafts, theses and dissertations, in-house publications (reports, bulletins or memos), and conference proceedings. Essentially any publication which has not been reviewed is informal. Conference proceedings are an interesting case of an informal publication appearing formal and they are troublesome for that reason.

The Scientific Journal

The scientific journal marks the publication boundary between the informal and formal domain.

> "The scientific article is, and will remain for some time, vital to the scientific community. It is the basic unit of the scientific journal process which provides a system for formal, public, and orderly communication among scientists (see Appendix D). Journals are formal in the sense that article manuscripts have been reviewed, revised to near perfection, and then allowed to pass into the formal domain where they may be explicitly cited and unambiguously retrieved. They are public both in the sense that anyone can submit a manuscript for publication in them and that they are available to anyone in libraries or by subscription. The orderliness of journals is founded on their articles being selected on the basis of scientific merit, which means that (a) the research reported is flawlessly conducted and (b) its results are relevant to scientific progress in the sense that they have explicit continuity with previous work and foreshow the future course of work on the research front." (Garvey, 1979, p. 69)

The prime purpose of publishing in a scientific journal is <u>not</u> communication, it is the attainment of professional goals. The scientist publishes in the journal to gain and maintain visibility among his or her peers. The journal also is the socially accepted medium for establishing priority of discovery. It is who publishes first which decides who gets credit in the annals of science for a discovery. The second use for the scientific journal is for the establishment of a public body of knowledge.

Publication is the prime product of science. Thus the entering of a work into the formal domain is essentially the completion of the research.

The Formal Domain

As we've discussed, the characteristics of the formal domain are that it is public, orderly and stable. The types of publication which are considered in the formal domain are the journal, the treatise and the text. You'll note from Figure 1 that it takes an average of 13 years for work to appear in a text or treatise.

One part of the formal domain of particular interest to information specialists is post-publication processing. This includes indexing, abstracting and citation analysis. Post publication processing, in contrast with the informal domain and the publication and review process, has been primarily in the hands of librarians and information specialists. In the last ten years we have made great strides in automating indexes and cutting down the time it takes to index an article. There is no question that these improvements were necessary, but it is clear from Garvey's work that even if articles were indexed the moment they were published, the index would be too slow for the scientist. There's no sense in our being frustrated about it: we know now that the reason is that the scientist's prime communication goes on in the informal domain. Nothing we do after publication is going to satisfy him.

The Scientist's Use of Information

Garvey's book also includes research on information use by scientists. He considered ten kinds of information needed, eleven stages of scientific work, and seven sources of information. By combining the various categories, he could determine, not only that "journals" and "local colleagues and students" were the most often successfully used sources, but also that journals were usually used to satisfy two of the ten kinds of needs:

> "Journals are most useful in providing information needed to place a scientist's work in proper context with similar work already completed and to integrate his findings into current scientific knowledge; these two are the kinds of information least frequently obtained from local colleagues and students. On the other hand, local colleagues and students are superior providers of information needed to select a design for data collection, to design equipment or apparatus or to choose a data-gathering technique... Books seem to play a dual role in providing needed information. On the one hand they are effective in providing general information needed to formulate a scientific solution and on the other, specific information needed to choose a data-analysis technique. Reprints and meeting presentations are most effective in providing information needed to relate a scientist's work to ongoing work in his area. Technical reports displayed a rather flat spectrum of kinds of information useful to the scientists. Relative to the other kinds of information they provide, they appear to function best in yielding information needed to design equipment and to select data-gathering techniques." (Garvey, 1979, p. 265)

Garvey found variation in information needs depending on the stage of research. He also noted that most scientists were engaged simultaneously on several "different" research projects at once and were generally in different stages in each. The only information need that was constant through all stages was the need "to place the work in context."

Other factors that Garvey noted include variations in use between:

1) Physical and social scientists. (Meeting presentations and technical reports are more useful to the physical scientists. Books and local colleagues and students are more useful to the social scientists).

2) Those working in a new subject area and those working in the same area. ("Journals seem especially useful in satisfying information needs of scientists working in a new area.")

3) Less experienced and more experienced scientists ("In most instances the information needs of the least experienced scientists are greater than those of the most experienced scientists." Local colleagues and students and journals seemed especially useful to the least experienced.)

4) Basic scientists and applied scientists. (Applied scientists use all media less than the basic scientists, except that they seem to find technical reports more useful than the basic scientists.) (Garvey, 1979, p. 273)

Information Needs of Engineers

Garvey's research showing a difference in information use between applied scientists and engineers is supported by numerous other use studies.

Rosenbloom and Wolek (1965) found that "The operationally oriented engineer (a) used information sources internal to his organization more than a scientist (68% as against 21%), (b) used oral sources more (69% as against 41%) and (c) used professional literature less (9% as against 46%)...

"Slater and Fisher showed that in the U.K. engineers used libraries markedly less than did scientists." (Gralewska-Vickery, 1975, p. 3)

Gralewska-Vickery's own study of earth science engineers reported that:

> "Despite the publisher's efforts, the formal written media are not extensively read by many engineers. A mining engineer said: 'I don't think that the mining engineer makes sufficient use of published literature, for two reasons: first of all there is not much valuable information published in mining, and secondly, he is not very familiar with how to go about finding the relevant information.' Another more senior mining engineer was doubtful about the contents of journals: 'Some of the journals (and here in particular I can mention the rock mechanics literature) are too theoretical, too much mathematics, too little practical applications, almost no case studies or descriptions of simple methods and equipment'." (Gralewska-Vickery, 1975, p. 274)

Engineers seem to want digested information, and if they are given too much unevaluated information they will select based on authors they know.

> "We were told: Sometimes when I get information from you, I look at the list, but there is frankly so much of it, I look at the titles and authors and unless I recognise the author, and know that he is a good man, I don't choose to read the paper. It's good to get the lists of references from you, but it is advisable to take short-cuts on the basis of people whom you know and whose work in the field is appreciated." (Gralewska-Vickery, 1975, p. 277)

Engineers overall read less and they use literature and libraries less. Even information services which are directly oriented to the engineer are little used except by academics.

We also know that the literature of technology is not cumulative, as science is. "Citations to previous papers or patents are fewer and are more often to the author's own work." (Allen, 1977, p. 39)

Allen, in Managing the Flow of Technology (1977), insists that much of the confusion in information use studies comes as a result of not differentiating between scientists and engineers. It's important for us to consider the difference, because in many cases geologists are applied scientists and show the information patterns of engineers rather then scientists.

Allen quotes research that shows that engineers differ from scientists in their basic values. For instance,

> "Engineers and scientists, despite surface similarities, are so fundamentally different in their natures that one could hardly expect similarity in communication behavior. Not only are the two groups socialized into entirely different subcultures but their educational processes are vastly different, and there is a considerable amount of evidence to show that they differ in personality characteristics and family backgrounds as well. Krulee and Nadler (1960) contrast the values and career orientation of science and engineering undergraduates in the following ways:
>
> '(Students) choosing science have additional objectives that distinguish them from those preparing for careers in engineering and management. The science students place a higher value on independence and on learning for its own sake, while, by way of contrast, more students in the other curricula are concerned with success and professional preparation. Many students in engineering and management expect their families to be more important than their careers as major sources of satisfactions, but the reverse pattern is more typ for science students. Moreover, there is a sense in which the science students tend to value education as an end in itself, while the others value it as a means to an end'." (Allen, 1977, 37)

Engineers also differ from scientists in their levels of education and their work goals. Ritti (1941) found a marked contrast between the work goals of scientists and engineers after graduation. Allen (p. 38-39) summarized Ritti's findings as follows:

> "While publication of results and professional autonomy are clearly valued goals of Ph.D. scientists, they are just as clearly the least valued goals of the baccalaureate engineer... Furthermore, both groups desire career development or advancement, but for the engineer advancement is tied to activities within the company, while for the scientist advancement is dependent upon the reputation established outside the company.
>
> "The type of person who is attracted to a career in engineering is fundamentally quite different from the type who pursues a scientific career. On top of all of this lies the most important difference: level of education. Engineers are generally educated to the

baccalaureate level; some go on to a Master of Science degree; some have no college degree at all. The scientist is almost always assumed to have a doctorate. The long, complex process of academic socialization that is involved in reaching this stage is bound to result in a person who differs considerably in his lifeview. These differences in values and attitudes toward work will almost certainly be reflected in the behavior of the individuals. To treat both professions as one and then to search for consistencies in behavior and outlook is almost certain to produce error and confusion of results."

Most of this research on the differences between engineers and scientists was done with U.S. engineers, and may not hold true for European engineers.

In addition to differences between engineers and scientists as people there is a difference in the nature of science and technology.

In science, as we've seen, the literature is the repository for all scientific knowledge and it is permanently recorded. To a technologist, the document is not so important. The creation stands for itself.

"The names Wilbur and Orville Wright are not remembered because they published papers... the technologist's principal legacy to posterity is encoded in physical, not verbal, structure. Consequently, the technologist publishes less and devotes less time to reading than do scientists." (Allen, p. 40)

The communication structure of technology is also different.

"Information is transferred in technology primarily through personal contact. Even in this, however, the technologist differs markedly from the scientist. Scientists working at the frontier of a particular specialty know each other and associate together in what Derek Price has called "invisible colleges." They keep track of one another's work through visits, seminars, and small invitational conferences, supplemented by an informal exchange of written material long before it reaches archival publication. Technologists, on the other hand, keep abreast of their field by close association with co-workers in their own organization. They are limited in forming invisible colleges by the imposition of organizational barriers." (Allen, p. 40)

Engineers work mostly for large bureaucratic organizations. Much of the information they have is proprietary and sharing between companies is not as easy as sharing between universities. In Allen's study of engineering information use, he found that the source for the piece of information which resulted in a better solution to a design problem was most often an internal consultant rather than literature or an external source. Those internal sources often turned out to be "gatekeepers," the exceptions to the rule, people who were high communicators, read the literature and had extensive outside contacts. These people translate the "outside" information into terms that are applicable "inside."

Allen also pointed out that because of the importance of personal communication, turnover plays an essential part in furthering technology.

When an engineer changes jobs he carries practical information with him. Allen noted that the only really successful technology transfer has occurred when someone involved in the scientific or technological development left the research organization and went into business.

Geoscientists

So, now that we have all this background, we can talk about geoscientists.

Are geoscientists more like scientists or more like engineers?

Engineering, mining and oil company geologists seem more like engineers. They're usually looking for how to do something, they're covered by proprietary information rules, they want targeted, selected, evaluated, digested information. But many geoscience researchers show the patterns of scientists. They have invisible colleges, concentrate on the informal domain for communication, and place a high value on the scientific journal and the review process. These are primarily academics, but some scientists in more bureaucratic organizations also fit here.

So how do geoscientists' literature needs differ from those of other scientists and engineers?

Geoscientists' Literature Needs

1) Geoscientists need access to a wider range of subjects because the field applies other sciences to earth problems, and because geology is increasingly used in social science applications (e.g., city planning) (Craig, 1969; Thurston, 1980, p. 20; Bambach, 1981, p. 20)

2) Geoscientists need access to older material. (Craig, 1969; Burton and Kebler, 1960) This is particularly true in paleontology and regional geology.

> "Apparent obsolescence rates vary from physically oriented subjects (relatively brief half-lives) to those biologically oriented (relatively long half-lives). Journals that publish papers in general geology, or across a broad spectrum of geoscience subdisciplines, fall between the extremes. Most characteristic of these is the Bulletin of the Geological Society of America. For traditional disciplines such as paleontology and field geology, obsolescence rates seem to have varied little over a twenty year span. For fast changing disciplines such as solid earth geophysics, obsolescence rates have varied widely within this interval." (Kohut, 1974, p. 250)

3) Geoscientists need access to a wider range of formats. In particular, their use of maps, guidebooks, and theses is significantly higher than other sciences. For instance, Eugene Garfield discovered that theses ranked number 20 in a list of highly cited journals in geology and geophysics. "In few of the lists we have compiled have theses figured so prominently." (Garfield, 1974, p. 105)

4) Geoscientists use informal published media more than other "scientists."

5) Geoscientists use abstracts and indexes less than other "scientists." Bichteler quoted King (1976, 2:312) in her paper on user needs for the First International Conference: "According to King (1976, 2:312), environmental scientists depend less on abstracts and indexes, either printed or machine readable and more on references from colleagues than other scientists." (1979, p. 265) Gralewska-Vickery found the same pattern.

> "The use of abstracts journals is not very widespread. They are used mostly by senior engineers employed in governmental institutions which have a library, which in turn disseminates information to the senior staff; by some consultants, especially when they approach a new problem; and by engineers of any discipline and stage of their career, if they have been educated to use them and have a library in their work which subscribes to them. It is very difficult to generalise on this topic. One suspects that there will be always quite a considerable number of engineers who have never used and never will use abstracts journals. They will either rely completely on their own experience and knowledge of authors whose work may be of value to them, or will depend on oral information given them by others who act either as gatekeepers, as friendly colleagues or as expert consultants.
>
> "A considerable number of the interviewed mining engineers had never seen an abstracts journal, despite the fact that the Institution of Mining and Metallurgy produces one and distributes it without charge to its members." (Gralewska-Vickery, 1975, p. 275)

6) Geoscientists, because of the nature of their work, need access to literature by geographic area.

They need to be able to get all of the information relevant to their work pinpointed by a small area and also all the information that's relevant in the region around it.

7) Geoscientists, because they do field work, have a need for hard copy publications.

One can't plug in a microfiche reader or a terminal in the middle of the Andes.

8) Because geology does not conform to political boundaries, geoscientists need access to foreign language materials and translations of them. And the need is not just for common foreign languages, but also for the obscure ones.

Geoscientist's Data Needs

What's different about geoscientist's data needs?

Shapley and Tomlinson (1980), writing in Data Handling for Science and Technology, pointed out these differences:

1) Geoscience data are space-dependent, time-dependent, or both. In contrast, physics and chemistry data are fundamental, unvarying values.

2) In geoscience there is a huge quantity of data.

"One satellite image of terrestrial cloud cover contains 10^8 data points; and images are taken every few minutes--mounting up to 10^{12} bits per year per satellite. Monitoring the earth's magnetic field, the ionosphere, or earth motions by seismographs by some 200-600 stations throughout the world accumulates only slightly less data. Modern seismic soundings of the ocean bottom lead to the accumulation of data at the rate of some 10^4 per kilometer, and there are over 10^6 cruise kilometers made each year. Thus it is readily apparent that, if a large part of these data is considered at one time, the largest computers in the world must be used for modeling the dynamics of the geosciences areas. In addition, the largest mass storage capacity must also be available. The techniques for handling these kinds of space-varying and time-varying data are quite different from those for most kinds of laboratory data." (p. 41)

3) Selective retrieval is required.

"Usually the data disseminated are selections from the total file. Most seismologists study seismograms only for earthquakes of special interest to them individually, and do not want seismograms for days without earthquakes. Thus there are problems of selective retrieval from a large homogeneous data set." (p. 43)

4) The data are not repeatable under identical or completely controlled conditions. (p. 42)

5) The data often require "data massaging" to change the form of data so that they may be stored or may be more amenable to analysis. Data massaging includes:

"smoothing of lines and boundaries, the elimination of distortion in graphic records, the rotation and translation of coordinate systems, the matching of edges between adjoining sheets of graphic data, and the rectification of topological errors." (p. 43)

6) The data are not disseminated in the literature.

"The scientific literature is not the principal medium for disseminating many large bodies of geoscience data; the quantities are too great, and the needs for dissemination can, in such cases, be better met through the use of alternate (e.g., computer-compatible) formats. Accordingly some categories of geoscience data are disseminated primarily through scientist-to-scientist exchange or through data centers." (p. 43)

Implications for Information Services

What does our knowledge about scientists, engineers and geoscientists imply for information services?

W. D. Garvey's Research

Garvey's description of the importance of the informal domain explains why the most well-established faculty are not heavy library users and why they're never satisfied with the currency we can provide: their interest is focussed in the prepublication period. Even an index produced at the moment a journal came out would be too slow. Is there anything we can do to help the scientist in the informal domain? Garvey suggests helping track the conference literature.

> "Any assistance which can be given in locating important meetings and conferences and acquiring the relevant information presented there will be of immense assistance to the scientists, and any effective information service to a scientific community should include it." (p. 57)

Conference Papers Index has recently made a step in the right direction. They announced that they'll be including announcement of conferences and not just the published papers from them. Also GeoRef has begun putting some abstracts into the database again after stopping due to lack of funds. We need to do more with the conference literature.

Garvey also told us that the most important reason scientists publish is for prestige. This may explain why the GSA Bulletin's move to microfiche failed: nobody wanted his or her contribution published in microfiche. And his description of the interrelatedness of the publishing system should make us understand why changing it is so hard. ("In the present-day system of scientific communication, the roles of user, producer, and disseminator are so dynamically intertwined that they cannot be treated separately when attempting to provide information services to scientists." (p. 26) Proposals to speed up the delivery of information by putting everything online and cutting out the review process are unacceptable because the formality is necessary. Electronic journals will only succeed if there comes a time that they are more prestigious than hard copy journals.

Garvey's research also pointed up the importance of journals as sources of information. We must find ways to continue to buy them, even though the increasing costs are frightening.

Engineer's Behavior

What are the implications from our knowledge of engineer's behavior? For the short-term, we should seek out the gatekeepers. Make sure they know about available information services such as computerized literature searching, SDI, displays of new books, etc. And talk to them as sources of information about what their colleagues are doing. In the long term, we should be working toward digested, evaluated information. Our aim should be some kind of information system which retrieves the way an internal consultant does, applying broad knowledge to a specific problem.

Geoscientist's Unique Needs

Because of the geoscientist's needs, libraries need to provide access to a wide range of subjects. Online searching makes available indexes for many subjects on one terminal. As Bambach said, "The importance of on-line searching for the reference desk cannot be overstated. All geoscience librarians must take all the necessary steps to ensure they understand fully and can manipulate effectively these bibliographic data bases. Librarians must become as familiar with their 'ins and outs' as they are with those of the standard, and not so standard reference books. The advent of the on-line source creates the possibility of using a whole range of services which would never be purchased in hard copy just for a geology library. It is rather like having readily available the secondary services off the shelves of engineering, biology and chemistry libraries. Such multidisciplinary use of subject services can only benefit the user." (1981, p. 88)

Also because of the geoscientist's unique needs, libraries need to keep the older material around, collect maps, guidebooks and other informal publications, and fight to keep access to theses available (and not just in microfilm).

Because geology knows no boundaries, our databases must become more comprehensive. This may require networks of local databases. They must also accommodate all languages.

We also need to work to improve the specificity of retrieval from databases. As Garvey said,

"The problem for contemporary scientists is that even if they had perfect retrieval systems they would be presented with so many items that they could not assimilate and process them." (p. 107)

The most pressing need is to find a better way to respond to questions such as "I'm working at the Nevada Test Site, and I need all the information on the geology of the area." The current geographic index[illegible] and retrieval on the bibliographic databases is not adequate, (Pruett, 19[illegible], pp. 1-18) and too much of the literature (particularly local and informal literature) is not included. Retrieval of data by geographic area on numeric databases such as the Earthquake Data File of NOAA's World Data Center in Boulder, is much more sophisticated and user friendly than anything the bibliographic databases have available. (See Hittleman, 1982 for a technical description of this system.)

On the other hand, how does one find out what databases exist? The bibliographic services might be able to help the numeric databases with "bibliographic control" of their proliferating databases.

Summary

Not enough is known about the information needs of geoscientists. We have studied what is known about the communication behavior of scientists and the information behavior of engineers and looked at the ways geoscientists fit those patterns. We have also looked at the ways geoscientists differ from other engineers and scientists and listed their unique literature needs and data handling problems.

Although our knowledge of our users is incomplete, the more those providing information services understand the more likely they are to provide services that meet the user's needs.

References Cited

Allen, T. J., 1977, Managing the flow of technology: technology transfer and the dissemination of technological information within the R&D organization, MIT Press, 320 p.

Bambach, R. K., 1981, Keeping current in an interdisciplinary field: paleontology, Geoscience Information Society Proceedings, v. 11, p. 11-22.

Bichteler, J., 1979, Geoscience information sources and services from the user's viewpoint, in Harvey, A. P. and Diment, Judith A. (editors), Geoscience Information: A state-of-the art review, Broad Oak Press, p. 263-270.

Burton, R. E., and Kebler, R. W., 1960, The half-life of some scientific and technical literature, American Documentation, v. 11, p. 18-22

Craig, J. E. G., Jr., 1969, Characteristics of use of geology literature, College and Research Libraries, v. 30, p. 230-236.

Garfield, E., 1974, Journal citation studies X: geology and geophysics, Current Contents, no. 30, p. 102-106.

Garvey, W. D., 1979, Communication: the essence of science, Pergamon, 279 p.

Gralewska-Vickery, A., and Roscoe, H., 1975, Earth science engineers: communication and information needs, Final Report, Imperial College Research Report, no. 32, 365 p.

Gralewska-Vickery, A., 1978, Communication and information needs of earth science engineers, Information Processing and Management, v. 12, p. 251-282.

Hittleman, A. M. and Metzger, D. R., 1982, "Marine Geophysics: database management and supportive graphics", In press (scheduled to be published in Computers and Geosciences, v. 8, no. 4)

King, D. W., 1976, Statistical indicators of scientific and technical communication 1960-1980, King Research, 4 vols.

Kohut, J. J., 1974, A comparative analysis of obsolescence patterns of the U.S. geoscience literature, Journal of the American Society for Information Science, p. 242-251.

Krulee, G. K. and Nadler, E. B., 1960, Studies of education for science and engineering: student values and curriculum choice, IEEE Transactions on Engineering Management, v. 7, p. 146-158.

Pruett, N. J., 1981, Searching for maps in GeoRef, Western Association of Map Libraries Information Bulletin, November, pp. 1-18.

Ritti, R. R., 1971, The engineer in the industrial corporation, Columbia University Press.

Shapley, A. H., and Tomlinson, R., 1980, Treatment of observational data in the geosciences: estimation and approximation of data, in Rossmassler, S. A. and Watson, D. G., (eds.), Data Handling for Science and Technology, North-Holland, p. 41-53.

Slater, M. and Fisher, P., 1969, Use made of technical libraries, Aslib Occasional Publication, no. 2, 86 p.

Thurston, N., 1980, Collection development in geoscience libraries: serials, Geoscience Information Society Proceedings, v. 10, p. 17-34.

SOME RECENT DEVELOPMENTS IN GEOSCIENCE INFORMATION IN THE THIRD WORLD

A.R. Berger
Geological Survey of Canada
601 Booth., Ottawa, K1A OE8

Abstract--- Geoscientists in developing countries place high priority on the need for improvements in publishing local journals and newsletters, in establishing information centers and data bases, and in gaining access to foreign scientific publications. This paper reviews some aspects of the bleak information "conditions" in developing countries and suggests a number of ways of offering practical assistance. In the long-run, however, Southern scientists themselves will have to resolve these problems.

Introduction

The 2nd International Conference on Geological Information addresses an important frontier in science - that of information management in all its aspects from data collection, manipulation, storage and dissemination to publishing, information services and policy matters. Naturally this frontier is marked by the continual appearance of new techniques, hardware, software, organizations, and the rate of change is not less than for other areas of the modern technological world. But there is another frontier - or more aptly barrier - that exists in science, one that is generally ignored by the scientists in the richer and more industrialized countries. That frontier is on the far side of a growing gap between science in the industrial nations of the North and that in the developing countries of the South.

The new generation of scientists and engineers in industrialized countries may well be brought up on microcomputers and videoscreens, but this will certainly not be the case in most developing countries. While we in the North race ahead in the search for new techniques of mineral exploration, new ideas on the origin of life, and a better understanding of petrological and tectonic processes, many of our counterparts in the South struggle hard to reach levels of scientific capability that we achieved a decade ago. For them the lack of access to Northern science through conferences, seminars, short courses, journals, text books and equipment leads inevitably to professional isolation and demoralization. There is a growing body of literature that attests to this (e.g. Berger, 1975, Moravcsik, 1975, Carman, 1979).

I believe that the rational development and management of the world's natural resources for which geoscientists are the main experts - mineral, fossil fuels, and groundwater - and the effective control of many aspects of our physical environment - volcanism, earthquakes, floods, and accelerated erosion - require cooperative efforts from scientists in all parts of the world. Those who live in Africa, Asia, and Latin America together with most of the world's population live under conditions which in far too many cases prevent them from playing the important role that they should in this process. Scientists are not the main cause of these conditions and neither can they hope themselves to close the gap between rich and poor. There are, however, a number of practical measures that scientists both Northern and Southern can take. This paper deals with some of them that concern information.

Geological Information Management in the South

At almost any gathering of scientists in developing countries the discussion inevitably turns to the need for more, better, cheaper information services. There can be little doubt therefore of the desire to improve their access to and management of scientific information.

It is easy to generalize about conditions regarding geological information management in developing countries but difficult to substantiate these general statements, for there are many stages of development to be found in such diverse countries as, for example, Saudi Arabia, Ghana, Sri Lanka, Brazil, Algeria, Chile, Tanzania, and Indonesia. Nevertheless despite the variations and with many exceptions, a number of common features exist, as anyone can verify for herself by a brief visit to many a university geology department or a geological survey.

Many libraries are cramped, poorly organized, dirty, and obviously little-used, as a glance at the borrower's card file - if one exists - will establish. Many local geoscientists will admit readily that no one knows exactly what information exists about local geology, or where to find that that does exist. As many visitors to the South have discovered to their bitter frustration, one of the hardest situations to cope with is this lack of knowledge about what scientific work is going on now or which has in the past. In addition there is apt to be an appalling lack of commmunication within a local scientific community. A special problem is what appears to the outsider to be the excessive confidentiality with which many scientific reports are treated, though one suspects that in some cases documents are labelled confidential simply to avoid the physical problem of making them available to the scientific community.

On an individual level, too, many scientists take the attitude that reading scientific works is only for students and researchers. There appears thus little desire to keep up with current trends, and new developments in their own fields. No doubt such attutudes result from cultural attributes as well as from rigid and excessively authoritative models of education. There is in addition a common reluctance to commit ideas and research results to paper, despite the fact that new work has been done. Language barriers are important factors here especially where opportunities to publish in local language are lacking. Unfamiliarity and unease with the main international scientific languages - English, French or Spanish - are powerful restraints to publication in the major journals.

There are other important resons too for the poor management of information. One is the dire lack of facilities in many areas. Duplicating machines, microfiche readers, computer facilities, reliable postal services, steady electricity supplies, are all too often not obtainable or attainable. Added to these are the pressures of hot climates and a voracious insect population (leading to deterioration of documents) and, more serious, a severe lack of dedicated people, especially at the levels of technicians and library personnel. Obviously finances are in short supply, and it is difficult to persuade politicians and policy makers in these regions to divert scarce monetary resources to bolster such scientific "infrastructure".

A ready way to gauge the lack of finances is to scan the foreign literature which comes into a Southern geoscience library. If libraries (and scientists) in Europe and North America find it hard to keep up with the increasing flow of information, for many developing countries the information explosion is already a relic of the past. Foreign exchange restrictions meant that international journals may have been stopped as much as 10-15 years ago. Even those institutions which have managed to acquire limited foreign exchange have great difficulty in monitoring new and recent publications in the North. Again isolation is the result.

This is a indeed bleak picture. But there are bright spots and there are a number of cooperative approaches involving geoscientists and information specialists both Northern and Southern which deserve encouragement and support.

Publications from the South

Newsletters

The growth in the number of geological newsletters produced in developing regions is, I believe, a healthy sign. Not only do they provide a practical means of forging communication links (in local languages) within a country, a discipline, or a region, but they also provide valuable experience in publishing for their editors. The quality is variable and their frequency of appearance irregular, but they are a start and their proliferation is a measure of the need for communication which they help to fill.

To quote a few examples, among these newsletters are those produced by national, regional or international geoscience societies:

- CCOP Newsletter (ESCAP Committee for Coordination of Joint Prospecting for Mineral Resources in Asian Offshore Areas)
- SEGELA (from the Council of Directors of Latin American Geological Services -in Spanish)
- AGID News (Association of Geoscientists for International Development)
- African Geoscience Bulletin (Geological Society of Africa and AGID - in English and French)
- RMRDC Newsletter (ESCAP Regional Mineral Resources Development Center -for Asian region)
- Geosciences (SE Asia Regional Network of Geosciences and AGID)

There are also Southern newsletters which focus on a particular discipline:

- Aguas Subterraneas (AGID - on groundwater, in Spanish)
- AGE News (from Asian Information Center for Geotechnical Engineering)
- ASEQUA Bulletin de Liaison (regional newsletter of Senegalese Association for African Quaternary - in French and English)
- IGCP Project 129 Newsletter (on laterisation processes, from India)

And there are, of course, those newsletters produced in the North largely or primarily about or for developing countries, such as

- CGLO Newsletter (Commonwealth Geological Liaison Office, London)
- African Geology (Association of African Geological Surveys, Paris)
- Natural Resources and Energy Newsletter (UN, New York)

This is not to mention those international newsletters and newsmagazines produced in the North and which are read widely in developing countries (e.g. from IUGS, IGCP, AGI...).

Despite their repetitive and often incestuous nature, the advantage of these newsletters is that they can be produced rapidly and cheaply, and in local languages where this is appropriate (Hindi, Thai, Malay, Arabic, Korean, Chinese, Portuguese...). Quality control does not have to be strict and they offer their sponsors a chance to acquire foreign publications on an exchange basis. Indeed scientific societies in the North wishing to bridge the communication gap could make a start by encouraging and entering into such exchange agreements.

Journals

The pace of growth of more formal journals published in most developing regions has been much slower. Latin America has perhaps the longest tradition with scientific periodicals like:

- Ameghiniana: Revista de la Asociacion Palaeontologica Argentina
- Revista Geofisica (Mexico)
- Noticia Geomorfologica (Brasil)

And elsewhere there are respected, if little known, journals like:

- Bulletin of the Geological Society of Malaysia
- Bulletin of the Faculty of Earth Science, King Abdulaziz University, (Saudi Arabia)
- Bulletin of the Geophysical Observatory of Ethiopia
- Journal of the Geological Society of India

To judge by comments from developing country geoscientists there seems to be a need for more such journals to provide an outlet for scientific papers of too local interest to be acceptable to the international journals. Yet it is costly to start new periodicals and much experience is needed - money and capabilities which are in very short supply in many parts of the South.

Government Documents

The most important source of geological documentation in most developing countries are government documents, especially bulletins and memoirs of geological surveys and mines departments. In Africa and Asia many of these publications, pre-date independence and were produced during the former, colonial times. With the change-over to national management came for many countries a slowing down or halting of government publishing. Indeed a frequent criticism of geological survey departments in developing countries is that they do not appear to be doing anything useful geologically. Thus it is often very hard to detect, especially from a distance or when visiting a country briefly, any appreciable progress in regional mapping in the last 10 or 20 years. On closer examination substantial advances can sometimes be seen, but the fact that this progress is barely visible to those outside the Survey itself underlines a continuing problem that faces many developing countries. To illustrate this situation an example from Ghana follows.

Ghana has one of Africa's older geological surveys, an establishment which in 1979 had some 60 Ghanaian professionals, excluding those on study leave. From Independence in 1957 to the end of 1978, 14 map-sheets (mainly at 1" = 1 mile) and accompnying memoirs were published, some of these being the result of work done earlier when the Survey was under colonial status. Yet this represents only 22% of the total map-sheets completed during this period.

The delays in publication have been caused first by the lack of staff with editorial experience, a need which becomes obvious when one realizes that much of this work, particularly in the 1960's, was the result of aid schemes involving geologists from the US, Bulgaria, Holland, Poland, India, and no less than 59 from the USSR. Fortunately most of the necessary translating, rewriting and editing is nearing completion, thanks to the efforts of editors sent on technical assistance missions from North America and Eastern Europe, but the second hurdle remains - the difficulty of getting the material printed and distributed. Not only are there problems of cost, but huge delays still occur at the Government printers, due to shortages of paper, breakdowns of equipment, backlog of work, etc...

The reports unpublished in 1979 included 57 describing map areas, giving a total map coverage (mostly 1" = 1 mile) of nearly half the country. And if one omits the area underlain by the Voltaian sedimentary cover which has little mineral potential and has thus been the subject of less geological work, the coverage of the rest of Ghana where the mines and mineral occurrences are increases to over 70%, a significant achievement. But most of the 63 reports unpublished in 1979, some having lain unpublished for up to 20 years, have still to appear in public.

The Ghana example may not be typical, but it illustrates graphically the need for editorial assistance which a number of developing countries have. Such situations have unfortunate side-effects too in that they add to the frustrations of scientists who are unable to see their work made available to others. As many Third World geoscientists are well aware, policy-makers and planners who are responsible for budget allocations are generally reluctant to provide funds for editing and publication, except through the normal (clogged) Government channels. And geoscientists in turn do not seem to have been very successful in explaining the importance of their work seeing the "light of day".

Comment

The growth in geoscience publications issued in developing countries is encouraging, especially where informal literature like newsletters is concerned. In the case of the more formal journals and government documents, however, practical support mechanisms are needed in terms of finances, editing and advice on procedures and policies. For example, Northern geoscience organizations with in-house publications might consider internships of several months which would allow prospective editors from developing countries to gain practical experience in editing, layouts, preparing copy, printing and even distribution. Professional societies like GIS, Editerra (now EASE) and AESE could also play a role by extending their membership, on "concessionary" terms, to organizing special sessions (workshops, short courses) for developing countries.

Information Services

Biblographic services, data-bases, and information centers for the earth sciences are rare in developing countries, as are even the simplest aids that we take for granted in the North such as printed bibliographies, indexes, reprographic services and so forth.

The only established, automated regional bibliographic data base (BDB) in geoscience that I know of in developing countries is AGE - the Asian Center for Geotechnical Information at the Asian Institute of Technology in Bangkok (Valls and Brenner, 1980 Burk, 1982). Though this was started and is still run mainly by civil engineers, it is an excellent model, as long as one has a ready source of funds and a willing host institution with resources to spare.

There are plans to develop other BDB services in the South. J.V. Hepworth's paper to this Conference describes one now being planned for coal in Asia and the Pacific. And there are other organizations with long-range plans to develop information bases (numeric as well as bibliographic) - the UN Economic Commission for Africa (mineral resources), the UK-based Intermediate Technology Indust rial Services (techniques for small-scale mining, mainly in developing countries), the Comité Interafricain d'Etudes Hydrauliques (water resources in Sahelian Africa), and AGID (groundwater in Latin America). However, in view of the rapidity with which new data bases and information services develop in the North, one may also why there is not more urgency in developing countries.

An obvious reason is the lack of support mechanisms, and guidelines, especially as regards manual methods of data handling. How does one identify key-words and operate manual files? How should the development of an information center or data base proceed? What are the necessary inputs? What equipment is needed? How many people, with what kind of training? How long does it take to set one up? How does one encourage the use of new centers? How much will it cost? There is surely a need for case-histories here, such as the one presented by U.H. Rowell at this Conference. COGEODATA has produced a wealth of information on automated methods, but these aim too high for the present circumstances of many developing countries. Observations and suggestions on geoscience information systems can be found in Hepworth's paper.

In a sense the need for these kind of information centers can be measured by the degree of coverage of Southern geoscience literature by Northern bibliographic data bases such as GEOREF, PASCAL/GEODE, GEOARCHIVES etc. After all, if these BDB's are doing the job well then it should be a simple matter of developing ways to allow access to them by developing country scientists, who would at least be able to know what literature they should acquire. A little reflection on the difficulties of monitoring scientific publications in Chinese, Thai, Malay, Korean and other rather local languages will suffice to show that serious gaps must occur.

Appendix I presents a detailed study undertaken in 1980 of the coverage by the major BDB's of the English-language geological literature of two very different countries: Ghana and Thailand. This study, originally presented in abstract form at the 26th International Geological Congress in 1980, showed that the BDB's were able to pick up some literature largely unknown in but relevant to these countries, but they missed 35% to 90% of the literature known to local geoscientists. The proportions vary depending on the country and the BDB, but despite the ridiculously small sample used in this study the conclusion is clear that the Northern BDB's in their present stage of development are not able to monitor Southern geoscience literature effectively enough. This must be at least partly because too few in the South know themselves what local information exists. I believe that the BDB's can help to change this situation by encouraging and supporting Southern groups aiming to compile bibliographies themselves and to improve their capabilities of monitoring local information.

Access to Foreign Literature

Many of the problems in building science in developing countries stem from the great difficulty their scientists have in gaining access to the international scientific literature. The main barrier is one of finances and currency restrictions. Libraries (and individuals) can rarely afford to subscribe to foreign journals or to purchase books, and even when they can do so there are severe restrictions on the availability of foreign currency. Where these restrictions exist it may be as difficult to acquire the convertible equivalent of $5 as $50. The provision of discounts and reduced subscriptions for developing countries by Geological Magazine, AGID, and Pergamon Press (for AGID members in the South), for example, are commendable. It remains, however, to see whether these reductions make any substantial difference.

Attempts to get around the currency restrictions barriers have been made in various ways. AGID accepts payment "in-kind" for membership dues for developing countries. However no more than about 25 out of some 700 members in the South use this method, sending in geological maps, survey memoirs or postage stamps of a value equivalent to the dues. Unesco Book Coupons were designed to circumvent currency restrictions and can work for those countries participating in the program. However, they are not easy to acquire and notoriously difficult to redeem. International Postal Coupons can also be used, but again are difficult to cash.

A different approach was begun by AGID nearly eight years ago when it began to "recycle" books and journals donated by libraries and individuals to institutional libraries in developing countries. This scheme is still going under the direction of John Moore of the Geology Department of Carleton University in Ottawa. He receives and sends out an average of 1000 to 2000 separate items each year. Libraries requesting donations are sent a list of available titles from which they select useful items and return to Carleton. Publications in stock are then sent out on a first come first serve basis, with due regard to need and to geographic balance. This service is possible because of the presence in Ottawa of the Overseas Book Center, a charitable agency which packs and ships the books and journals at a modest charge, which is in turn covered by the Canadian International Development Agency.

Obviously this service depends upon volunteer contributions and is thus rather hit-and-miss. No attempt has been made to build up collections in any systematic way or to provide on-going support for designated libraries. However, it might be possible to identify scientists who subscribe to but do not keep back issues of journals and to persuade them to donate on a steady basis. There is clearly an opportunity here for groups and individuals in the North to make a practical contribution to international development, by gathering donations and starting other clearing stations.

Final Comment

This paper suggests a number of ways in which Northern-based societies, groups individuals can assist in bridging the information gap to the developing countries. Most of these approaches are stop gaps or band-aids - temporary measures as best. In the long run their problems must be resolved by Southern scientists themselves, but the struggle is a hard one and there is a need for constant encouragement, support, and "solidarity" from the international geo-science community at large.

Appendix: The Ability of Computerized Bibliographic Data Bases to Monitor the Geoscience Literature of Developing Countries: Two Test Cases

Introduction and Procedure:

In recent years I have asked many geologists in developing countries about their use of bibliographic data bases (BDBs), increasingly essential to scientists in the North. Most had either never heard or never used them, and a few claimed that "We can do much better ourselves for the geoscience literature of our country". So I decided to test the usefulness of BDBs in a few selected countries. Studies begun on the Caribbean region and on Tanzania were not completed, but for Ghana and Thailand analyses were possible thanks to a major contribution by D.R. Workman now of the University of Hong Kong. The sample size is ridiculously small, but the results appear to be symptomatic of a larger problem.

A basic requirement for the study was the recent appearance of geoscience bibliographies compiled by local geologists who are in the best position to judge the relevance of the literature and to identify all literature produced locally. Their bibliographies were compiled manually, without reference to the BDBs. Print-outs were then obtained (in the second half of 1979) from major BDBs for these countries. The most "hits" were scored on GEOREF, GEOARCHIVE and PASCAL/GEODE, but also searched were NTIS, AGRIS, CAC, DEL, COMPENDEX, and SWR, none of which produced more than a few citations. The print-outs were then compared carefully with the local bibliographies (LBs) for publications dated in the same time period. The results are presented graphically in Figures 1 and 2.

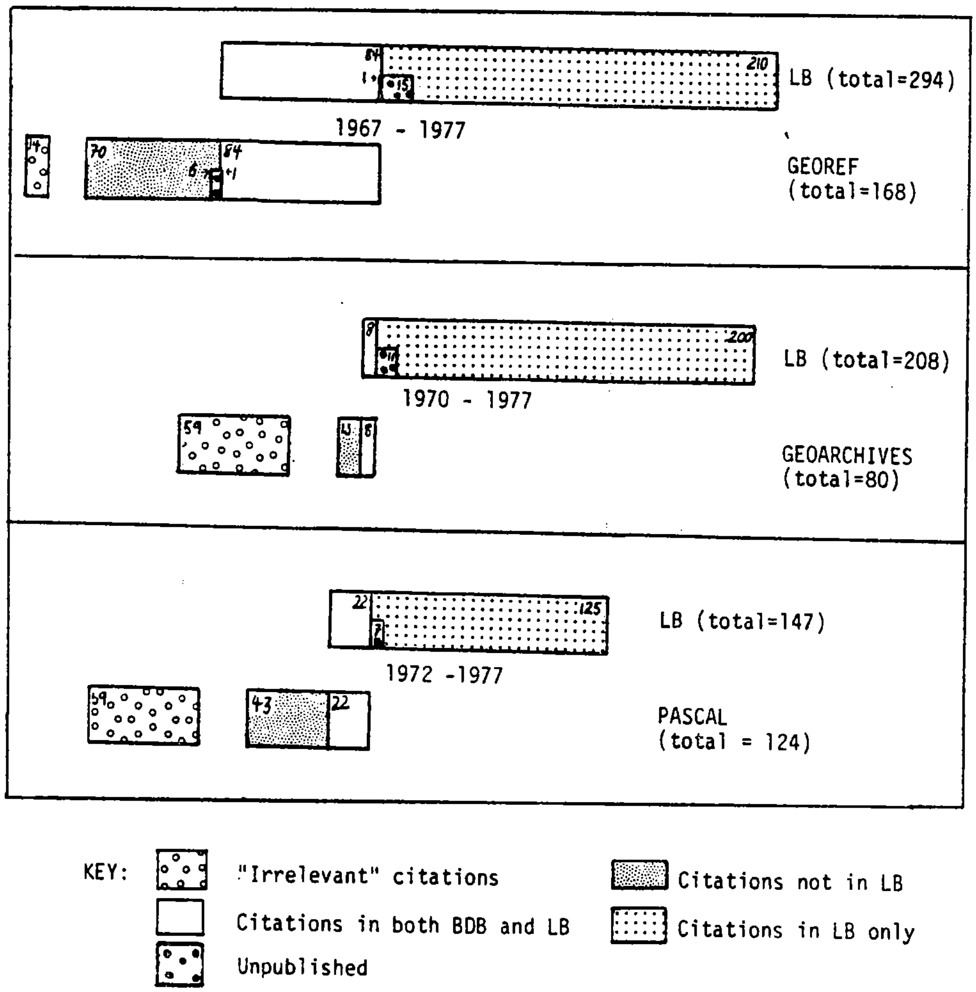


Figure 1: GHANA - Graphical illustration of the degree of overlap in the number of citations (given in figures) contained in the local bibliography (LB) by Jones and Kesse (1978) and that in various bibliographic data bases (BDBs).

Ghana

Omitting certain citations in the BDBs judged to be irrelevant within the context of the coverage of the LB (Jones and Kesse, 1978) including duplicated citations, citations wrongly included, e.g. on a place called Ghana, New Mexico, the situation is as follows:

GEOREF: Overlap with LB = 23% of the total relevant literature
Cited by BDB but not by LB = 19%
Cited by LB but not by BDB = 58%

GEOARCHIVES: Overlap = 4%
Cited by BDB but not LB = 6%
Cited by LB but not BDB = 90%

PASCAL: Overlap = 12%
Cited by BDB but not LB = 23%
Cited by LB but not BDB = 65%

Some comments need to be made. First, this analysis omits reference in the LB to work on soils, archaeology, newspaper articles and on adjacent regions. The unpublished works from the LB referred to in Figure 1 do not include the many undated Ghana Governement reports cited in this bibliography. Second, the coverage of PASCAL and GEOARCHIVES is rather different from that of GEOREF; they include many citations from the mining press, for example, which were excluded from the LB. Hence there were a large number of "irrelevant" citations in these two BDBs. Citations in GEOREF missed by the LB included 15 papers in French, German, Russian and Czech, 11 on engineering geology and limnology, 5 overseas theses, 7 abstracts, and 15 regional works with passing mention of Ghana. Those in the LB but missed by GEOREF have not been analysed thoroughly, but about half are citations of local government and university publications, and 10-20% of West African journals and conference reports. A special search for citations on groundwater in Ghana on several other BDBs (including SWR and AGRICOLA) turned up no more than 10-20% of the titles included in a recent LB produced by the Water Resources Reesearch Unit in Accra; nearly all the latter were published locally.

Thailand

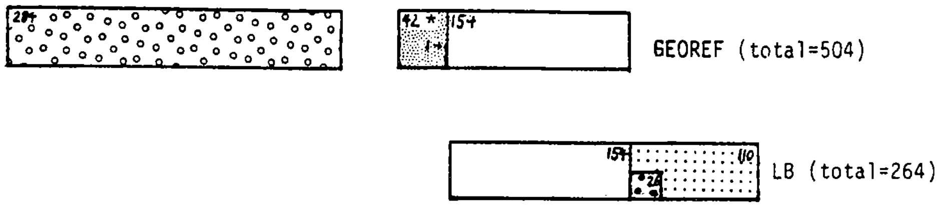
Here, GEOREF again turned up the largest number of citations, and was the only BDB analysed in against the LB (Workman, 1978). Figure 2 show the details.

Overlap with LB = 50% of the total relevant literature
Cited by BDB but not LB = 14%
Cited by LB but not by BDB = 36%

The 284 "irrelevant" citations in the BDB include:

- -41 published in but not about Thailand
- -31 groundwater maps of NE Thailand reduced to a single citation in the LB
- -23 citations with only passing mention of Thailand
- -21 citations to newspapers, newsmagazines, and yearbooks
- -5 citations in languages other than French, English, and German
- -5 citations duplicated
- -95 citations in engineering geology, soils and hydrology
- -22 citations on mining, mineral dressing and petroleum
- -41 citations on tektites and other miscellaneous topics

The latter three categories were all excluded from the LB. Most of the other citations were known to the compiler of the LB, but omitted intentionally



* Of these 16 were known but omitted from LB through oversight
5 are published abstracts of unpublished papers
3 were published in Japan, 1 in India, 1 in Hungary, 1 in Italy, and 1 in Sweden
3 are from gemmological journals
3 are in German
2 are on seismicity
1 is an M.Sc. thesis
3 are on miscellaneous topics
42

Figure 2: THAILAND - Graphical illustration of the degree of overlap in the number of citations contained in the local bibliography (LB) by Workman (1978) and that in GEOREF (the BDB). Key as in Figure 1.

Of the 110 citations in the LB but not in GEOREF:

-46 were published in Thailand (university, government, and geological society
-23 were UN reports
-26 were unpublished (10 consultant reports, 8 university theses and reports, 7 Thai government and 1 foreign government reports)
-7 were published in Malaysia and Singapore
-3 were published in USA
-2 were published in Germany
-1 each published in France, South Africa, and UK

It should also be mentioned that neither Georef or the LB included citations to 392 geological papers written in the Thai language and cited in Nutalaya et al. (1978). Probably no non-Thai geologist can read these.

Conclusions

A maximum overlap of 50% is probably not unreasonable for many other LDCs too. For example, a preliminary analysis of the water resources literature on the Caribbean region indicates that the 5 major BDBs searched (GEOREF, GEOARCHIVES, PASCAL, NTIS, and ENVIRONLINE) identified less than 50% of the literature compiled manually for the same time period.

The BDBs certainly cite some important literature unknown to local geoscientists, but this is numerically much smaller than that identified by LBs and missed by the BDBs. It seems therefore that a geologist with access to good local and international libraries and a thorough familiarity with the local geology can probably do better than the BDBs, given time and money.

Yet the BDBs are performing a very useful service, one which should be much more widely advertised in developing countries, and which would undoubtedly be used more there if (a) the coverage of LDCs were increased, and more importantly (b) efforts were made to bypass the inevitable barriers of foreign exchange and inertia on the part of LDC users.

This preliminary study says nothing of the primary problem of gaining access to the actual literature cited, a virtual impossibility for far too many LDC geoscientists. It also says little about the problem of the immense volume of unpublished literature "stored" in LDCs, and for the most part un-monitored and inaccessible.

References Cited

Berger, A.R. (ed.), 1975, Geoscientists and the Third World: a collective critique of existing aid programs: Geological Survey of Canada Paper 74-57, 32p.

Burk, C.F.Jr., 1982, A worldwide list of source and reference databases in the geosciences: Database, June, 1982, p. 11-21.

Carman, J.S., 1979, Obstacles to mineral development: a pragmatic view: Pergamon Press, New York, 175p.

Jones, E. and Kesse, G.O., 1978, Bibliography of geology of Ghana and related topics: Geological Survey of Ghana Bulletin 48, 66p.

Moravcsik, M.J., 1975, Science development: the building of science in less developed countries: PASITAM, International Development Research Center, Bloomington, Indiana, 262p.

Nutalaya, P., Petchvarun, V. and Saengsuwan, C., 1978, Bibliography of geology and mineral resources of Thailand (in Thai language), in Nutalaya, P. (editor), Geology and mineral resources of Southeast Asia: Asian Institute of Technology, Bangkok, p. 823-843.

Valls, J. and Brenner, R.P., 1980, The Asian Information Center for Geotechnical Engineering - Operation, Services and Experiences: Workshop on Road Research and Information for Southeast Asia, Melbourne, Australia, 14p.

Workman, D.R., 1978, Bibliography of geology and mineral resources of Thailand, 1951-1978, in Nutalaya, P. (editor) Geology and mineral resources of Southeast Asia: Asian Institute of Technology, Bangkok, p. 805-821.

IMPACT OF CURRENT DEVELOPMENTS IN INFORMATION MANAGEMENT ON DEVELOPING COUNTRIES

G. N. Rassam and J. Gravesteijn

GeoRef, American Geological Institute
One Skyline Place-5205 Leesburg Pike, Falls Church, VA 22041

Abstract--Recent trends and techniques of information management in the bibliographic field are sketched briefly and the possible and actual utilization of such techniques by Third World countries is explored. It is demonstrated that the Third World can be divided into (1) developed, (2) developing, and (3) underdeveloped, in terms of utilization of information technology of the Industrial World.

It is argued that being "underdeveloped" in this sense may give the country concerned a chance to move faster than the "developed" country. Several proposals are made to achieve better cooperation between the Third World and the Industrial World in the field of information transfer.

Introduction

Information has been defined in different ways by different specialists, the definition mostly reflecting the particular concerns and needs of the author concerned. Many have called information a "resource" or an economic product whose value can be quantified, although attempts at such quantification have been scarce and not very successful. Most agree that information, approached from a socio-economic and socio-political point of view, forms an identifiable body of products or goods which provide the essential underpinning of modern society in the "North" or the "Industrial World". Information in that sense is the "raw material" of the "knowledge" industry which is an integral part of this Industrial World. But before we try to make the contrast with the status of information in the rest of the world, let us dispose of the terminology issue.

The terms "industrial world", "third world", "developing countries", "socialist countries", "underdeveloped nations", "countries on the way of development", "communist countries", "East-West confrontation", "North-South confrontation", all have been used to describe, more or less, certain geographic areas and politico-economic issues that have characterized our world in the past few decades and that continue to take shape this day. Be that as it may, we will use the terms Industrial World (IW), basically denoting the U. S., Europe, Japan, and the USSR, and Third World (TW) denoting essentially the rest of the world. According to Price (1981), these two worlds can be called "information-rich" and "information-poor". That, of course, is a typical IW view that tells us, in effect, that someone is rich simply because they possess certain "resources" while another is "poor" because they do not possess these very same "resources", all the time neglecting to tell us that these "resources" were arbitrarily defined to be of a certain kind, and only of that kind.

For if we look at the status of information in the Third World, we will see that although it does not measure up to the IW standards (for a hilariously condescending article on this, somewhat typical of the genre, see Bortnick, 1981), it is still quite often very rich and varied and does act in all cases as the base of all knowledge just as it does in the Industrial World.

In other words, there are cases of information-poverty and information-wealth in both IW and TW and while the efficiency of the use (and misuse) of this information varies widely all over the globe, it is not necessarily true that it is always more efficiently used in the IW. There is no evidence, for example, that the city government of Richmond, Virginia knows how to use the information available to it (or even that it has more useful information available to it) than the city government of Fez, Morocco. The two city governments simply have two different kinds of information sets available which are used differently.

Although this argument is important and deserves a large measure of documentation and analysis, it will not be further developed here except to emphasize the particularities of both the problems and the solutions facing the Third World in this regard, and to try to change the focus from quantity of information (wealth-poverty arguments) to the quality of information (kinds and options), and to mention in passing that much of the TW present ruling elite itself has been thoroughly indoctrinated in the notion of the IW information-superiority, even when this same elite sometimes rails against "cultural imperialism" and the like.

To examine the issues confronting TW geoscience information decision-making, this paper will briefly focus on the status of geologists, information management, geoscience information (especially the bibliographic part), and computer and communication technology. It will then discuss particular TW problems, transfer problems, models, and finally make some concrete suggestions for mutual help between the Third World and Industrial World.

The emphasis throughout will be on this dynamic mutuality which is the *sine qua non* of all information transfer and meaningful progress.

Geoscientists and Information, a State-of-the-Art Capsule

A recent survey (Barroul et al., 1980), indicates the presence of some 550,000 geoscientists in the world (less than 2% in Africa, more than 38% in Asia, and close to 43% in Europe) with the type of economy having no apparent effect on the number of geoscientists, but with the gross national product (GNP) showing a not surprising close relationship with the total number.

From the point of view of "theory" vs "practice" these geoscientists are split 1 to 3 which is not surprising either.

The geoscientific work that these scientists and technologists do varies from teaching to drill-site geology and, depending on specific country characteristics, they tend to congregate into scientific societies and professional organizations closely parallel to their specialties and interests. In the August 1981 issue of Geotimes, some 500 different such societies are listed. In the Geologists' Year Book of 1977 (Dolphin Press, London), 250 societies are listed. Considering that most Russian and Chinese societies are not listed, and the lack of reporting of some other countries, a number close to 1000 seems to be not unrealistic. Many of these societies publish journals and books, i.e. produce primary information. Perhaps as many as 3500 geoscientific journals exist worldwide today and 8000 if we include those journals dealing with subjects related to geology.

Many other organizations such as museums, oil companies, survey organizations, universities and research institutes produce their own printed material (little on microfiche or tape or film material) in the form of monographs, special papers and the like. The total number of items published each year would be (depending on the assumed inclusion and exclusion factors such as including abstracts, grey literature, etc.) 70,000-100,000 documents increasing at a rate of 3-6 percent a year. Journals are sprouting all over and, in many cases, with no clear functional value to the user.

This is so much so and the effect in cost to libraries and individual scientists is such that a crisis situation is fast approaching. Heath (1982, p. 684) goes as far as suggesting that scholarly institutions should impose a fee for the services of editors of proprietary journals and that new journals should be reviewed for quality and demonstrated need.

As libraries face high cost, they tend to drop some journals, relying on bibliographies while many more are dropping printed bibliographies in favor of online services. No statistical evidence exists for such trends in TW countries, if any exists, and in the case of GeoRef's Bibliography and Index of Geology, the overall trend is for a constant subscription rate despite increased subscription cost.

The field of geoscience itself has broadened its scope to include such studies as correlation of soil type with disease incidence, geoarchaeology, paleontological anthropology and other multidisciplinary approaches. Many factors enter into these changes affecting the careers of young geoscientists and the sociology of these changes is not quite clear but certainly the economic factor looms, as it always does, as the most significant. For the past three or four years for example, geologists in the United States were actively pursued, recruited, even wooed by oil companies, a situation that quickly (if one is to believe the stories) became farcical with oil executives scrambling to offer young BS graduates such amenities as new cars and salaries often higher than their professors', affecting quickly the delicate balance of education and industry in the United States. But now, with a very short-duration slump in oil prices, the competition suddenly grows less fierce.

How does all this affect geoscience information? and where are the TW geologists in this boom-bust cycle?

To begin with, younger geologists in the IW are less inhibited in using computer technology and therefore their active involvement in early careers has definitely contributed to the rapid increase in demand for online services. The scramble for geologists has created a parallel scramble for information of all kinds.

In the Third World the situation was and is less hectic: most of the organizations able financially to perform exploration work are state- or foreign-owned. In either case, there is a large measure of government control and "planning", making the freedom of movement of young geologists (and information available to them) much more restricted than their counterparts in the Industrial World. This situation, among others to be tackled later, creates less of a real demand for information than should be expected simply by the size of geoscience population.

Recent Developments in Geoscience Information Management

During the seventies, the information industry in the IW came of age, and very rapidly, inroads were made in all the scientific disciplines known to man, geoscience being not excluded. The existence of the present meeting and its predecessor in 1978 is an indicator of these inroads and the direction of things to come. But these meetings, though "international" in intent, are really IW in construct.

Not only are they held in IW places (London and Golden) and sponsored by IW people, but are attended by mainly IW geoscience-information people (not more than twenty participants out of a total of 140 at last count are from outside North America!).

Looking at the programs of the two meetings, we notice the following: For the London meeting: only three presentations out of 31 total were from TW countries (India and Brazil). Four of the total were about TW topics.

In the Golden meeting only 2 out of some 50 are by TW people, and 5 of 50 are on TW topics.

Aside from these meetings, there are several other venues and services of geoscience information available both online or batch, in print or in microfiche. The status and trends of all of this is delineated in several papers presented at this meeting and others in the past. For more details, see the references by Gravesteijn (1982).

The Computer Revolution

The major vehicle for modern information is the computer and the telecommunications network that has evolved quickly over the past few years.

The rapid developments in this field are fortunate for they facilitate the process of transfer of information, making it both cheaper and faster. But they present further complications in dealing with TW countries which are for the most part unequipped (both in machines and skilled manpower) to tackle the problems involved in machine processing.

It is also difficult for these countries, many of which just emerged from centuries of colonial past, to formulate policies dealing with information transfer. The political rigidity and centralized government apparatuses of many of these countries exacerbate a tendency to hinder the flow of information even at its most innocuous.

Brazil, for example, requires government approval of any computer processing done by foreign agencies. The official aim is to encourage a budding domestic processing industry and to protect local manpower and industry. The overall effect however, is higher cost of processing relative to U. S. vendors and less international interaction. Such a policy, it must be said, is by no means atypical and in fact may be less restrictive than many.

From a purely economic technical viewpoint, the electronic revolution will provide many poorer TW countries opportunities for possessing very sophisticated systems at a relatively low cost -- all in this decade.

Abelson (1982) summarizes the cost and efficiency trends: The price of small computers has been dropping at a rate of 25 percent per year while the annual increase in available computer power is about 40 percent. Miniaturization resulting from placing more circuits on a silicon chip has improved reliability per circuit 10,000 times better than it was 25 years ago.

Much effort is going into the improvement of the human-machine interface and of computer graphics -- again of significance to TW countries where machines are still viewed with great suspicion and where any attempt to "humanize" the machine therefore becomes of crucial value.

Software programming is not advancing as fast, being much more complicated in its logic than hardware. But here again, this could be an opportunity for TW countries, where mathematical skills are much more abundant relative to technological skills.

Furthermore, it is the local manpower which is much more aware of its particular needs and linguistic intricacies, thus it is better able to analyze problems and construct useful computer programs that are adapted to the particular issues at hand.

In the area of information processing and transfer, major advances are being made in the field of computer-controlled switching: glass fiber transmission lines able to transmit data at a rate of 10 billion bits per second and major improvements in satellite technology are foreseen.

The remaining TW problems of infrastructure deficiencies are major obstacles in the path of taking advantage of this computer revolution. Electricity shortages and outages, lack of trained personnel and lack of maintenance play a much more important role in the TW than in the IW. But such problems are not insurmountable even in the technologically least-developed countries. As management skills improve and the general economic infrastructure is modernized, quick adaptation to computer technology is possible.

In Kuwait online searches are routinely being made by the Kuwait Scientific Research Institute. In Iraq billing for electricity, water and other services is handled by the computer. Seismic data for Nigerian oil exploration is processed and stored by computers.

The examples abound. What is important to remember here is that the general lack of infrastructure may not be universal within a given TW country, i.e., a highly-advanced institution, information-rich and technically superior, such as a national oil company replete with highly-trained personnel may coexist with a society and institutions that are essentially pre-industrial in their outlook.

From personal experience, comes the case of a national oil company sending a geologist to investigate a report of an oil seep in a city near the capital of the country. After meeting the reporter of the seep, the geologist -- accompanied and cheered by a host of curious kids -- finds that the "oil seep" comes from a ruptured fuel tank in a nearby house. Such a straight-forward explanation however, would not have been sufficient for the local population already dreaming of sudden riches, so a hole is dug in the street, some of the oil is collected in a soft-drink bottle, and a solemn pronouncement is made to the effect that thorough analysis will be made, all to the murmured approval of the gathering throng. Science is the new magic, and the people know it. (The analysis, properly made, confirms the fuel tank source, but the appearances are correctly observed.)

Transfer Mechanisms

Information exchange involving published material is a process of transfer of information from the primary product (document) to the ultimate user involving, usually, the agency of a third party, the secondary information organizations or abstracting and indexing agencies.

Primary Product

TW countries still produce a very small proportion of the total published scientific output. As an indication, it is estimated that in 1976, Arabic book titles did not exceed 8,000 titles, roughly 1.5% of the total for the world. In geoscience, the proportion is even smaller since most books published in the TW relate to the arts and humanities.

Thus in the exchange process the total TW published material is small but it can be of major significance to the IW if it deals (as it does quite often) with basic descriptions of the geology of a region. Detailed discussions of stratigraphy, tectonics, and mapping projects are often more useful in the producing country than a theoretical definition of the mechanism driving continental drift.

With respect to periodical literature, TW countries vary widely in the number and quality of the journals they produce. In the appendix are listed journal titles covered by the GeoRef Information System for the three model countries of this report: Brazil, India and Iraq. From these lists, it appears that India produces some 150 earth-science related scientific journals (many are general science journals), Brazil produces 110, and Iraq 5.

These numbers are bound to increase in the future as more and more geologists join the workforce and as educational opportunities increase.

Indexing, Abstracting and Processing

While geoscience documentation began very early in India (3rd quarter of the 19th century according to Ghosh, 1978), most of the effort has been in the nature of printed bibliography compilations and no centralized computerized data base seems to have emerged.

The same state of affairs appears to be the fate of geoscience documentation in both Iraq and Brazil.

TABLE 1.
GEOSCIENCE BIBLIOGRAPHIC WORKS IN INDIA, BRAZIL AND IRAQ

Major bibliographic centers		Major bibliographies	Commencement
India:	Geological Survey of India	Earth science abstracts	1975
	Indian Geological Index	Indian geological index	1971
Brazil:	Instituto Brasileiro de Bibliografia e Documentacoa	Bibliotecas especializades brasileiras	?
	Instituto de Pesquisas Tecnologicas, S/A Divisao de Minas e Geologia	Bibliografia e indice da geologia do Brasil	?
Iraq:	University of Baghdad	Bibliography of Iraqi geology	?

Burk, C. F. (Australian Mineral Foundation, 1981), identified 82 source (or numeric) databases in the geosciences, none of which belonged to a TW country. Of 46 public reference databases, 5 were international in scope and only one, that of Asian Geotechnical Engineering based in Thailand, belongs to a TW country.

In Iraq a National Bibliographical Agency working with the National Library Office has been set up and is preparing the national bibliography and index of Iraqi periodicals as well as an index of Iraqi theses for 1970-1980.

The Language Barrier

One of the major difficulties in the flow of information has always been the language barrier and although it is widely assumed nowadays that English has become the lingua franca of technology, problems still exist in accepting this concept for various national, psychological and technical reasons.

Linguistic imperialism as perceived by countries of the TW, takes the shape of English, French or Russian domination. The "experts" who come to help don't speak the local language of the host country. Add to that a commonly-perceived disastrous decline in the facility of American scientists, for example, in other languages, and you can see some of the problems involved (Mulholland; Fedmann, 1981).

Many of the English-speaking scientists today in fact share the attitude that "English has become not only truly international . . . but also irreplacable, since many modern scientific terms are in it" (Arditti, 1981).

But despite such views there is no question that language problems can and do hamper the free flow of information and, in fact, they will increasingly add to flow restrictions as more and more countries and cultures participate in the further development of science and technology and as political awareness increases in relation to increased economic power in some TW countries.

Tools to break down the language barriers have been developed. These include the creation of multilingual thesauri, artificial international scientific languages, and computer translation.

Machine-aided human translation remains perhaps the most effective means of circumventing some of the problems and there seems to be no real reason for the fear of an impending catastrophic clogging of information flow through the use of a cacophony of languages. In most TW cases, the colonial past determines the language of exchange and the assumption has been made a priori that one or more European languages are "necessary" for "progress" on a national level. Thus students in the Arab countries of North Africa, for example, learn French or English or Italian or some combination thereof, depending on what European countries colonized them in the past. Such patterns are hard to break. Algeria, for example, has been trying to "Arabize" its curricula for decades now and while the results are indeed impressive, the official but informal language of communication remains the French.

Technology Transfer

That transfer of technology is not a new process, albeit its current fashionability as a phrase, is recognized by many thoughtful writers. De Bettignies (1978, p. 321), for example, emphasizes the transfer of such important technologies as printing and the smallpox vaccine from the developed country of China to underdeveloped Europe of the time.

The developed Arabs, with the "help" of the Crusaders managed the transfer of sugar and silk technologies, as well as the introduction of the zero concept in mathematics leading ineluctably to the development of all modern computer technology. Other examples of such transfer abound through the ages.

Thus the idea of transfer has been with us for a long time, changing orientation and significance constantly. But in the past two decades, the phrase has acquired a special meaning: the transfer of technology from the industrial world to the pre-industrial world (from IW to TW) as an integral part of development. More recently, particularly after the Report of the Club of Rome and the investigations of the Brandt Commission, such stress on "development" as a goal in itself is being questioned, or rather, qualified. What is being forgotten in the debate, is that technology transfer has always been to some degree, a two-way street. Information transfer is even more fundamentally a reciprocal process, a "process of negotiation". Information, particularly in geoscience, is of fundamental strategic consequence to nations.

Nau (1976) identifies three major political motivations of technology transfer for the United States: military context, foreign aid and private industry, all of which are part of the national effort to maintain superiority and security.

All countries share this concern and all make an effort at acquiring information relevant to that concern especially in the area of raw materials.

Technology and information transfer is a dynamic process that creates new conditions in itself: that is, it changes the premises from which one starts and, given the preponderance of "hard" technology (machines) in the IW and therefore the predominance in quantity at least of the data produced in the IW over that produced by the TW, the latter is, as a recipient of technology and information, cast into the role of a dependent. A dependent who needs his routine fix which can be cut off at the pleasure of the supplier.

The relationship furthermore, as pointed out earlier, is basically different: it is a mutual flow. The IW needs geoscience information from and about the TW more than the other way around: most of the major reserves of mineral and fuel resources lie in TW areas, a situation that suggests to some the possibility of TW countries charging for online information about their countries regardless of its source, the money to go into the development of local information resources.

While this idea may be at present far-fetched and impractical (no collecting mechanism or agency), it would not be surprising, given the available online technology, to see some version of it coming into being soon.

More fundamentally, the information coming from the IW is often tailored to the needs of IW users and contains information of irrelevant or minimal use to TW scientists. For example, references to articles of local geological descriptions in the U. S. or USSR is of very little use to TW geologists. Deep Sea Drilling Project reports are of use to Chinese and Indian geoscientists perhaps but to few other TW scientists. Lunar and planetary exploration reports have no significance for a geologist in Malaysia, and the geology of Pakistan is hardly of interest (except theoretically) to a petroleum geologist in Argentina.

Present Status of Geoscience Information in the TW

1. Acquisition

Primary information availability varies according to source and location of information system.

An Indian information system would necessarily have more access to Indian information than does, say, a British system. But this is not strictly true either. Because of historical reasons (the colonial past) and geographic reasons (India is a huge country) Indian Bengali information may travel faster to London than it does to Delhi (Berger, 1974).

There is no comprehensive geoscience information database existing today and there may never be (Rassam, 1982). Nor is there perhaps a real necessity for an absolutely comprehensive system. What is needed is an acquisition program in the international databases that is aggressive and conscientious in trying to contain all possible published material. To do that, cooperative arrangements must be set up between the literature-producing countries and the international system.

One example of such cooperation occurred when GeoRef was commissioned by UNESCO to compile a catalog for IGCP projects. The first such catalog was published in 1980 and a second edition will come out in 1983.

For the compilation of this catalog and index, GeoRef had to work in close tandem with both the UNESCO Secretariat in Paris which tries to be a repository of all IGCP reports and with individual IGCP project leaders throughout the world. This was particularly important since a good portion of the material was published in ephemeral-type publications such as workshop proceedings and project reports, available only through the person organizing the workshop or taskforce.

Still, the experiment worked to a substantial degree through the high motivation of all concerned.

2. Processing

Information processing has become in many cases computer-oriented. But geoscience information processing in many countries is not. This is not necessarily a drawback. What is always needed is a clear assessment of actual requirements of a particular country. In many cases, a mix of manual and automatic systems will be quite satisfactory. In some cases, bypassing the manual stage to a completely computerized system is possible and even desirable.

3. Dissemination

Geoscience information dissemination remains to a large degree in the traditional forms of publications (mainly of journals, rarely of books), conferences, and word-of-mouth. The rigidity of the local bureaucracy in many TW countries inhibits the flow of information from one organization to another within the country. In fact, there may be more flow between a given organization and outside the country than with another sister organization. In a sense, this is due to the realization by the managers of these organizations of the importance of information (and of withholding it) to their careers and authority. Information is power.

Present Status of Information Utilization in the TW

It becomes clear from the previous discussion that within the Third World itself, there is a wide variation in terms of the assimilation and utilization of modern information systems produced by the Industrial World, just as there is a wide variation between the amount of information available in a small town in Tennessee and in Washington, D.C.

This parallel, however cannot be pushed beyond a certain limit as the lack of information in the small Tennessee town is primarily a function of money and need or lack of them whereas in the TW it is often a function of the overall infrastructure of society that tends to short-circuit, so to speak, advanced-technology systems.

These TW countries are not uniform. Some like Brazil and India have already used modern information systems from computers to satellites and in fact, are beginning to produce systems of their own. In both cases, strange to add, the geosciences seem to have not been accorded a priority in the area of documentation and information control. Let us call these countries Model A.

Other countries such as Iraq, Algeria, and Pakistan belong to Model B in this sense: they are developing their information systems more or less rapidly and more or less haphazardly and to varying degrees of control over final product. Saudi Arabia, for example, which belongs to this group of developing countries, is trying to create overnight a complete information system modeled on the latest American methods (the best money can buy syndrome), but without having the personnel (or even potential personnel some would say) to man these projects, thus relying on foreigners. This kind of development is neither stable nor will it, in the long run as some people assume, automatically become national in character. The situation (dependence on foreigners) may perpetuate itself simply because of convenience.

This model is marked by a situation in which many, but not all, elements of the infrastructure are in place and what is most needed is a sense of direction -- a statement of national policy.

Still other TW countries, usually the poorer ones, such as Mali, Chad, and Surinam, are Model C -- underdeveloped in this sense.

Our argument here is that it is not such a bad thing to be underdeveloped, because as was shown before, the whole field of information is in a state of ferment and despite all the futurology, only the vaguest outline of the shape of things to come is discernible. Still, the rate of change is slowing down which means that long-range planning is becoming meaningful.

Several reasons give an advantage in terms of time and cost to Model C and Model B countries:

1. While it is estimated that only 10 percent of the population of a TW country is reached by existing information dissemination programs, the situation is rapidly changing via television; the McLuhanesque global village is becoming more of a reality than ever.

Furthermore, geoscientists form always, by definition, an "information elite" within their countries as they possess access to often confidential information of politico-economic nature (oil drilling statistics, mineral reserves, etc.), and they can command usually some access to computer technology.

2. The cost of computers is going down drastically, making it possible for many of the poorer countries to construct central information systems based on them.

3. The rapid advances in availability of online information systems and the telecommunications systems necessary for their dissemination, means that the democratization of information is moving faster than previously envisioned. A geoscientist in Dacca can practically instantaneously recall information stored in Palo Alto. The problems here are increasing cost of searching, the front cost of setting up facilities, complexity of searching (exaggerated by vendors) and lack of communication infrastructure.

A more serious problem is often the inaccessibility of the primary information once a reference is found. This is especially true for a TW geoscientist because of money-exchange restrictions. It is essential therefore that all efforts at providing TW geoscience libraries with primary information undertaken by many individuals and agencies such as AGID be continued and aided.

Future Needs and Proposal

To achieve the best results, a TW country could set up a national automated information network. Such a network will reduce duplication of work between various agencies and organizations in the country, free scarce technical manpower from essentially redundant work to concentrate on other tasks, and can provide a link with outside (IW) information networks through online vendors, for example, in the West and through the COMECON network in the East.

It is simpler to start with a computerized network than to try to "reinvent the wheel" or as a recent comment by a hydrologist put it "reinvent the wood in the wheel". The case of AESIS in Australia (Tellis, 1981) can serve as a good model here. The network can exist only if a conscious effort at standardization is made. For bibliographic material, the UNISIST Reference Manual, the second edition of which just appeared (Dierickx and Hopkinson, 1981) can serve as a model. Already several documentation centers are using it, or systems close to it in conception: Excerpta Medica, GeoRef, PASCAL, Zoological Record, DEVSIS (International Information System for Development Sciences), and DOCPAL (the Latin-American Population Documentation System).

Such a national network, or the relevant part of it, then can proceed to areas of regional and international cooperation.

In the latter, use can be made of the existing geoscience systems and, at the same time, there could be actual participation in the system to the mutual benefit of both.

One idea that has been discussed, is to have the national centers be responsible for the indexing of the country's literature, convert it to an international format and feed it to the international center. In the meantime, national literature of restricted accessibility can be integrated in the network but not made available to the international center.

A prototype already exists in the case of the cooperation of several European geological surveys with the PASCAL-GEODE data base (Gravesteijn, 1982) and the cooperation between PASCAL-GEODE on the one hand and GeoRef on the other (Rassam, 1981).

This contribution by the national center will entitle it to a proportional decision-making authority in the international center as to content and direction. An existing model would be that of the AGRIS agriculture data base.

In addition, the national center could receive items of interest to it from the total international package in proportion to its contribution to that package.

The technical details of such a proposal remain to be worked out and while we don't want to minimize them, they are essentially technical and solvable given the desire and the will.

Finally, a multilingual thesaurus in geology (rather, a switching mechanism for national thesauri -- a device for these thesauri to communicate with each other) already exists (Glashoff, 1981, pp. 201-206; Dvorkina, 1979). Ten languages are included with plans for adding Arabic, Turkish and Portuguese being discussed. It would be very useful for national or regional agencies in the Third World (language groupings would work here) to support the addition of their languages to this thesaurus and to start using it in transferring information.

Specifics of Geoscience Database Construction in the TW

Geoscientists the world over (in the IW and TW alike) need a certain amount of information and data in order to perform their work fruitfully. They need: a) detailed knowledge of their area: mapping and concomittant programs; b) general advances in theory; and c) advances in practice (machines, techniques).

All of these requirements need some kind(s) of information system to sustain them, but the first requirement's documentation needs are met best by some locally designed system combining the most advanced features of existing numeric databases and some access to a bibliographic system. Other documentation needs can be met by using an existing global bibliographic system in conjunction with such additional local material (report literature and the like) that may not have made its way into the global system (Berger, 1974).

To create such a "super bibliographic file" is not difficult and completely within the range of present technology and possibilities. The advantages for the TW country or region is that it has complete control on the file (which satisfies political and psychological considerations) while in fact sharing into a global network of information.

The cost for such a system should not exceed $100,000 initially and some $10,000 a month thereafter. To be practical, such a system need not be designed only for the geosciences but within a broader, national information system. Such modularity is again quite feasible.

What is very important however, is that the designers of any such system should opt for standardization at the very outset. Should they do that (in input forms, data elements, system design and output), they will not only reduce future cost, but also achieve an advantage over major IW systems which were designed and developed in a lurching, stop-and-go fashion. The new system will be much more adaptable and long-lived plus being easier to use.

The cooperative aspects with the IW could take the shape of bilateral agreements, international cooperation or regional agreement. Thus use could be made of relevant UN agencies such as UNISIST, international unions (such as IUGS), trade agreements and the like.

The most essential thing is that the design of the system be done with as full a participation of the TW country or region as possible. The geoscientists and information scientists of the TW must have the final say on the shape of the system to be, not some self-appointed consultant or group coming from the IW with their prejudices and preconceptions.

An example of where things could go wrong so easily occurred recently: the Industrial Documentation Center of Baghdad, Iraq is being established as a central information repository for all functions of the Ministry of Industry and Minerals. It will handle both externally and internally generated documents. The Center contracted with Computas, a Norwegian company, to provide for the training of two Iraqi information persons.

From interviews with these two people, it was found that within a period of 6-8 months, they had to attend courses in information systems at Case Western Reserve University as non-degree students, then have a whirlwind tour of professional groups, online services and others in both the U. S. and Europe. In the U. S. alone, they visited and talked to people from twenty organizations within less than three weeks, all, fortunately for them, on the East Coast.

They stated that they would rather have stayed in one place and learned how things are actually done than attend all those theoretical courses and then listen to all those sales pitches! Of course, none of this attitude came out officially except perhaps in the final report to their Center.

Concluding Observations

Development in the popular sense used by most can be defined by technical feasibility, popular participation and the socio-economic cost/benefit balance.

There is, in fact, no absolute development: it is always a function of time, place and material ingredients.

In all cases development is philosophy- or ideology-dependent. In the IW, the ideology prevalent in almost all institutions but especially in the scientific manifestations of society is that of the dominance of man over nature. History, "progress", art and technology are explained in that context. Only recently has there been some questioning of this fundamental ideology and then only by the fringes of society. In the TW, where development in this sense lags behind and where technology is mostly "primitive", the underlying ideology is that of man-in-balance-with-nature. That kind of philosophy, true for an Indian in Bombay, a bedouin in the Arabian desert and a Brazilian-native in the Amazon, has been essentially dependent on a low level of technology which requires much less expenditure of energy and material resources.

Can development then take place without changing the ideology? This is perhaps the most difficult and most important question for the rest of the century, not only for the TW but for an IW reaching the "limits of growth".

As with all such questions, no simple answer is possible or meaningful. The answer must be in the attitude which defines the nature and orientation of progress. Many legitimate complaints have been made (from both sides of the debate): the avarice of the West, the dogmatism of the East, the backwardness of the South, etc.

Clearly however, a balance between the two sides, a negotiation, must be struck in an interdependent world.

Many TW countries cannot feed their people. Some squander their meager resources on wars and weapons. Most have not reached the minimum political maturity needed for useful cooperative schemes with their neighbors. Most feel dependent and resentful. Najm Aboud Najm (1982) writes that technological development for the IW is mostly for finding the limits for the future while an increasing gap is forming with the TW. He charges that the IW tries to perpetuate the gap, even increase it. The numbers suggest that he may be right in that the economic gap is growing: the 1973 World Bank statistics give the gross national product per capita (in U. S. dollars) for North America and Japan as $5340 while for Africa and Asia (excluding Japan) it was $215 -- a ratio of 24.8. In 1978, after years of technology transfer the two corresponding numbers were $9029 and $328, and the ratio was 27.5. Something is clearly wrong here.

In the IW however, in many areas unchecked growth has meant a quick approach toward an ecological disaster. The economic cost of environmental concerns has meant a slow-down in the implementation of many programs designed to avert this disaster and the present trend in the U. S., vocally supported by most geoscientists, is indeed for more "mastery of nature", more exploration, more drilling, less regulation, less environmental curbs. The trend as applied by IW scientists in TW countries has been and is even worse. The concern there is for megabucks and miniresults. And this goes for individual consultants, companies and government agencies such as the U. S. Corps of Engineers. The underlying attitude one suspects is "who cares"?

To approach the stage of negotiated advance in the mutual transfer of information there must be first a breakdown of old attitudes and a buildup of new ones that are at once more realistic and more fruitful.

TABLE 2

GeoRef Serial List for Model Countries

INDIA

Akashvani
Birbal Sahni Inst. Palaeobot., Birbal Sahni Mem. Lect.
Cochin, Univ., Dep. Mar. Sci., Bull.
Geoviews
India, Meteorol. Dep., Mem.
Indian Bur. Mines, Miner. Econ. Div., Mark. Surv. Ser.
Recent Res. Geol.
Bombay Geogr. Mag.
Deccan Coll. Postgrad. Res. Inst., Publ.
India, Cent. Water Power Res. Stn., Annu. Res. Mem.
India, Geol. Surv., Libr. Bull.
Bioviqvanam
Indian Colloq. Micropalaeontol. Stratigr., Proc.
Photonirvachak
Mausam
Monogr. Palaeontol. Soc. India
Problems of the Arctic and the Antarctic, Collection of Articles
Hydrology Review
Geoscience Journal
Journal of Engineering Geology
Journal of Association of Exploration Geophysicists
Mineral Resources Series - Indian Bureau of Mines, Mineral Economics Division
Records of the Geological Survey of India, Part 5: Southern Region
Annual Report - National Institute of Oceanography
Mineral Information Series - Indian Bureau of Mines
Indian Sci. Congr. Assoc. [53rd Sess.], Proc.
Indian Sci. Congr. Assoc. [54th Sess.], Proc.
Indian Sci. Congr. Ass. [55th Sess.], Proc.
Indian Sci. Congr. Assoc. [56th Sess.], Proc.
Indian Sci. Congr. Ass. [57th Sess.], Proc.
Indian Sci. Congr. Assoc., 58th Sess., Proc.
Indian Sci. Cong. Assoc., 60th Sess., Proc.
Aligarh Muslim Univ., Geol. Dep., Ann.
Natl. Geophys. Res. Inst. (Hyderabad), Annu. Rep.
Geol., Min. Metall. Soc. India, Bull.
Indian Bur. Mines, Bull.
Indian Natl. Sci. Acad., Bull.
Indian Geol. Assoc., Bull.
Indian Soc. Soil Sci., Bull.
Mysore Geol. Assoc., Bull.
Natl. Inst. Sci. India, Bull.
Chayanica Geol.
Andhra Pradesh, Ground Water Dep., Dist. Ser.
Explorer (Orissa, Dir. Mines)
Geogr. Rev. India
Geol. Soc. India, Bull.

Table 2 (cont'd.)
INDIA

India, Geol. Surv., Misc. Publ.
Geol. Soc. India, Mem.
Himalayan Geol.
India, Cent. Board Irrig. Power, Publ.
Indian Geohydrol.
Indian Geol. Index
Indian Geotech. J.
Indian Geophys. Union, Bull.
Indian J. Earth Sci.
Indian J. Meteorol. Geophys.
Indian J. Mar. Sci.
Indian J. Pure Appl. Phys.
Indian J. Power River Val. Dev.
Indian J. Phys.
Indian Miner. Yearb.
India, Oil Nat. Gas Comm., Dir. Geol., Spec. Pap.
India, Oil Nat. Gas Comm., Bull.
Inst. Eng. (India), J.
Indian Acad. Geosci., J.
Indian Geophys. Union, J.
Karnatak Univ., J., Sci.
Mar. Biol. Assoc. India, J.
Madras Univ., J.
Nagpur, Univ., Geol. Soc., J.
Mysore, Dep. Mines Geol., Bull.
Mysore, Dep. Mines Geol., Geol. Stud.
Mahasagar
Miner. Res. (Nagpur)
Mysore, Dep. Mines Geol., Groundwater Stud.
Mines Miner. (Indian Bur. Mines)
Geol. Assoc. Tamil Nadu, Newsl.
Natl. Geophys. Res. Inst. (Hyderabad), Bull.
Natl. Miner. Dev. Corp., Newsl.
India, ONGC, Rep.
Bhu-Vidya
Prog. Geophys. (Hyderabad)
Indian Natl. Sci. Acad., Proc., Part B
Natl. Inst. Sci. India, Proc., Part A
Natl. Inst. Sci. India, Proc., Part B
India, Geol. Surv., Mem., Palaeontol. Indica, New Ser.
Palynol. Bull. (Lucknow)
Pramana
Panjab Univ. (Chandigarh), Cent. Adv. Stu. Geol., Publ.
Andhra Pradesh, Ground Water Dep., Res. Ser.
Poona, Cent. Water Power Res. Sta., Annu. Res. Mem.
Indian Roads Congr., Roads Res. Bull.
Roorkee, Univ., Res. J.
Birbal Sahni Inst. Palaeobot., Sir Albert Charles Seward Mem. Lect.
Birbal Sahni Inst. Palaeobot., Spec. Publ.
Birbal Sahni Inst. Palaeobot., Silver Jubilee Lect.
Andhra Pradesh, Ground Water Dep., State Ser.

Table 2 (cont'd.)
INDIA

Min., Geol. Metall. Inst. India, Trans.
Uttar Pradesh, Dir. Geol. Min., Monogr.
Vasundhara (Saugar, Univ., Geol. Soc., J.)
Vikram Univ., Inst. Geol., J.
Agra Univ. J. Res., Sci.
Bombay Nat. Hist. Soc., J.
Curr. Sci.
Geol., Min. Metall. Soc. India, Q. J.
Geol. Soc. India, J.
Indian J. Hist. Sci.
Indian J. Technol.
Indian Mineral.
Indian Miner.
Bulletin of the Indian Society of Earthquake Technology
Indian Soc. Soil Sci., J.
J. Mines, Met. Fuels (Calcutta)
J. Palynol. (Palynol. Soc. India)
J. Sci. Ind. Res. (New Delhi)
J. Soil Water Conserv. India
Met. Miner. Rev. (Calcutta)
Miner. Wealth (Gujarat, Dir. Geol. Min.)
Natl. Geogr. J. India
Palaeobotanist
Sci. Cult. (New Delhi)
Zool. Soc. India, J.
Inst. Fr. Pondicherry, Sect. Sci. Tech., Trav.
Proc. Annu. Conv. - Indones. Pet. Assoc.
Osmania Univ., Cent. Explor. Geophys., C.E.G. Bull.
Irrig. Power (New Delhi)
India, Geol. Surv., Mem.
Indian Natl. Sci. Acad., Proc., Part A
Indian Acad. Sci., Proc., Sect. A: Earth Planet. Sci.
Indian Acad. Sci., Proc., Sect . B
India, Geol. Surv., Rec.
Indian Ceram. Soc., Trans.
Indian Sci. Congr. Assoc., Proc.
Records of the Zoological Survey of India
Highway Research Bulletin
Indian J. Meteorol., Hydrol. Geophys.
Adv. Pollen-Spore Res.
Geophytology
Vignana Bharathi
India, Geol. Surv., News
Geophys. Res. Bull. (Hyderabad)
India, Geol. Surv., Bull., Ser. B
India, Geol. Surv., Bull., Ser. A
Poona, Univ., J., Sci. Technol.
Palaeontol. Soc. India, J.
Panjab Univ. (Chandigarh), Res. Bull.
Res. Bull. Panjab Univ., Sci.
Symposium on Earthquake Engineering

Table 2 (cont'd.)

IRAQ

Iraqi Chem. Soc., J.
Baghdad, Univ., Coll. Sci., Bull.
Sumer
Iraqi J. Sci.
Geol. Soc. Iraq, J.

BRAZIL

Acta Geol. Leopold.
Braz., Com. Exec. Plano Lavoura Cacaueira, Cent. Pesqui. Caucau, Bol. Tec.
Braz., Dep. Nac. Prod. Miner., Projeto RADAM, Levantamento Recursos Nat.
Naturalia (Brazil)
Parana, Univ. Fed., Cent. Pesqui. Estudos Geod., Bol., Geod.
Bol. - Sao Paulo, Inst. Geol.
Braz., Com. Exec. Plano Recuperacao Econ.-Rural Lavoura Cacaueira, Cent. Pesqui. Cacau, Inf. Tec.
Bibliogr. - Braz., Dep. Nac. Prod. Miner.
Projeto REMAC (Reconhecimento Global Margem Cont. Bras.)
Rev. Bras. Cartogr.
Bol. Tec. - Sao Paulo, Dir. Planej. Controle
Biblio-Geo - Univ. Sao Paulo, Inst. Geogr.
Atlantica
Boletim Fluviometrico
Anais da Semana Paulista de Geologia Aplicada
Anuario do Instituto de Geociencias
Boletim do Departamento de Geologia, Universidade Federal de Ouro Preto, Publicacao Especial
Arquivos do Museu de Historia Natural
Anuario Mineral Brasileiro
Monografias - Instituto de Pesquisas Tecnologicas do Estado de Sao Palo
Simposio Regional de Geologia
Semanas de Estudos Geologicos
Congr. Brasil. Geol., [22d, Belo Horizonte], Resumo Comun.
Congr. Brasil. Geol., [22d, Belo Horizonte], Roteiros Excurso#7.es
Acta Biol. Parana.
A#2.qua Subterra#4.nea
Braz., Dep. Nac. Prod. Miner., Anu. Miner. Bras.
Mus. Nac. (Rio de J.), Arq.
Braz., Com. Nac. Energ. Nucl., Dir. Exec. Area Miner., Bol.
Braz., Lab. Prod. Miner., Bol.
Bibl. Geogr. Brasil., Se#2.r. A, Publ.
Braz., Div. Geol. Mineral., Avulso
Braz., Div. Geol. Mineral., Notas Prelim. Estud.
Braz., Dir. Hidrogr. Navegacao, (Relat.)
Sao Paulo, Univ., Inst. Geocienc., Bol.
Bol. Paulista Geogr. (Sao Paulo)
Braz., Div. Fom. Prod. Miner., Bol.
Braz., Div. Fom. Prod. Miner., Avulso
Braz., Dep. Nac. Prod. Miner., Publ. Espec.
Mus. Nac. (Rio de Janeiro), Bol., New Ser., Geol.
Soc. Brasil. Geol., Bol.

Table 2 (cont'd.)
BRAZIL

Braz., Supt. Desenvolvimento Nordeste, Div. Geol., Bol. Estud.
Mus. Para. Emilio Goeldi, Bol., Antropol.
Mus. Para. Emilio Goeldi, Bol., Bot.
Bol. Parana. Geocie#4.nc.
Mus. Para. Emilio Goeldi, Bol., Zool.
Parana#2., Univ. Fed., Inst. Geol., Bol., Geol.
Braz., Div. Geol. Mineral., Bol.
Braz., Lab. Prod. Miner., Avulso
Braz., Supt. Desenvolvimento Nordeste, Div. Geol., Ser. Geol. Econ.
Braz., Supt. Desenvolvimento Nordeste, Div. Hidrogeol., Ser. Hidrogeol.
Braz., Supt. Desenvolvimento Nordeste, Div. Geol., Ser. Espec.
Braz., Supt. Desenvolvimento Nordeste, Div. Geol., Ser. Geol. Reg.
Congr. Bras. Geol., An.
Catastrophist Geol.
Sao Paulo, Univ., Inst. Geogr., Geogr. Planejamento
Gemol. (Sao Paulo)
Gondwana Newsl.
Sao Paulo, Inst. Geogr. Geol., Bol.
Iheringia, Ser. Geol.
Iheringia, Ser. Zool.
Braz., Inst. Pesq. Mar., Publ.
Inst. Oswaldo Cruz, Mem.
Mus. Para. Emilio Goeldi, Bol. Geol.
Pernambuco, Univ. Fed., Inst. Geocie#4.nc., Se#2.r. B, Estud. Pesqui.
Pernambuco, Univ. Fed., Inst. Geol., [Publ.], Geol.
Pernambuco, Univ. Fed., Inst. Geocie#4.nc., Se#2.r. C, Notas Pre#2.vias
Sao Paulo, Univ., Inst. Geogr., Paleoclimas.
Mus. Nac. (Rio de J.), Publ. Avulsas
Rev. Bras. Biol.
Rev. Cult. Para
Rio Grande do Sul, Univ. Fed., Inst. Geocien., Mapa Geol.
Rio Grande do Sul, Univ., Esc. Geol., Notas Estud.
Inst. Geogr. Geol. (Brazil), Rev.
Sao Paulo, Inst. Geogr. Geol., Relat.
Pesquisas
Rio de J., Univ. Fed., Inst. Geocie#4.nc., Geol., Bol.
Soc. Interca#4.mbio Cult. Estud. Geol., Sem. Estud.
Sao Paulo, Univ., Inst. Geogr., Ser. Teses Monogr.
Sudene
Rio Grande do Sul, Univ., Esc. Geol., Bol.
Rio Grande do Sul, Univ., Esc. Geol., Publ. Espec.
Sa#7.o Paulo, Univ., Fac. Fil., Cie#4.nc. Letras, Bol.
Acad. Bras. Cie#4.nc., An.
Bol. Geogr.
Bol. Te#2.c. Petrobra#2.s
Eng., Min. Metall.
Jornal de Mineralogia
Min., Metal.
Notic. Geomorfol.
Rev. Brasil. Geogr.
Acta Amazonica.
Mus. Para. Emilio Goeldi, Publ. Avulsas
Cad. Cienc. Terra - Univ. Sao Paulo, Inst. Geogr.

Table 2 (cont'd.)
BRAZIL

Cienc., Tec., Pet., Sec. Explor. Pet.
Sa#7.o Paulo, Univ., Inst. Geogr., Geomorfol.
Sao Paulo, Univ., Inst. Geogr. Sedimentol. Pedol.
Biogeogr. - Univ. Sao Paulo, Inst. Geogr.
Braz., Dep. Nac. Prod. Miner., Bol.
Revista do Instituto Geologico
Geogr. Urbana - Univ. Sao Paulo, Inst. Geogr.
Geol. Metal. (Sao Paulo)
Metodos Questao - Univ. Sao Paulo, Inst. Geogr.
Mineracao Metalurgia (1968)
Serie Recursos de Solos
Bol. Mineral.
REM (Rev. Esc. Minas)
Sao Paulo, Univ., Inst. Oceanogr., Bol.
Rev. Bras. Geocienc.
Sao Paulo, Univ., Inst. Geogr., Cartogr.
Soc. Bras. Geol., Nucl. Norte, Bol.

References Cited

Abelson, Philip H., 1982, The revolution in computers and electronics: Science, v. 215, no. 4534, p. 751-874.

Agris, 1978, Introduction: Food & Agriculture Organization of the UN, 38 p.

Alleyne, D. H. N., 1978, Role of state petroleum enterprise as transfer of technology: Natural Resources Forum, no. 3, p. 5-17.

Arditti, Joseph, 1981, The language problems, "letters": Science, v. 214, p. 1194-1195.

Australian Mineral Foundation, 1981, Seminar on "Geoscience numeric and bibliographic data": Adelaide, 30th March - 1st April, variously paginated.

Baker, Dale B., 1974, USSR/USA scientific and technical information in perspective, Miles Conrad Memorial Lecture: National Federation of Abstracting & Indexing Services, Philadelphia, Pa., Report, no. 8, 24 p.

Barroul, Jacques, Bodelle, Jacques, and Rolet, Philippe, 1980, Distribution of geologists around the world: Episodes, v. 1980, no. 3, p. 15-18.

Baxter, G. G., and Horder, M. F., 1981, Databanks and databases in geology: Journal of the Geological Society of London, v. 138, pt. 5, p. 573-630.

Berger, A. R. (editor), 1974, International workshop on earth science aid to developing countries, May 17-19, 1974, Memorial University, St. John's, Newfoundland, Canada, variously paginated.

Bortnick, Jane, 1981, Notes on Nigeria: Bulletin of American Society for Information Science, v. 8, no. 2, p. 27-29.

Cooper, Marianne, 1982, Secondary information services in science and technology; a wide-angle view: Journal of the American Society for Information Science, v. 33, no. 3, p. 152-156.

de Bettignies, H.-C., 1978, The management of technology transfer; can it be learned?: Impact of Science on Society, v. 28, no. 4, p. 321-327.

Delbos, Leon, 1981, Cooperation internationale dans le domaine de la documentation "Sciences de la Terre", plus specialement avec les pays en developpement: Bulletin du B.R.G.M. (deuxieme serie), sect. IV, no. 3, p. 235-240.

di Castri, F., and Hadley, M., 1979, A typology of scientific bottlenecks to natural resources development: Geojournal, v. 3, no. 6, p. 513-522.

Dierickx, H., and Hopkinson, A., 1978, Toward a common bibliographic exchange format?: International symposium on bibliographic exchange formats, Taormina, Sicily, 27-29 April 1978, UNIBID, London, 214 p.

_______ (editors), 1981, Reference manual for machine-readable bibliographic descriptions, 2nd revised edition, PGI/81/WS/22: UNESCO, Paris.

Drake, Charles L. and Schmitt, Lois E. (editors), 1981, Science special issue on solid earth: Science, v. 213, no. 4503, 138 p.

Dvorkina, B. D., Moskalenko, Z. D., Nilova, L., and others, 1979, Mnogoyazchnij tezaurus po geologii: Leningrad, VSEGEI, 16 p., French translation, Bureau de Recherches Geologiques et Minieres, no. 5550.

El Fassi, Idriss, 1980, I'dad al-bia al-tabeeia fi al-manatiq al-Sudaniya; tanawul manhaji li-tajribat Casmance (Senegal) [The preparation of the natural environment in the Sudanese regions; a systematic approach to the Casmance (Senegal) experiment]: Revue de Geographie du Maroc, no. 4, p. 31-36.

Eres, Beth Krevitt, 1981, Transfer of information technology to less developed countries; a systems approach: Journal of the American Society for Information Science, v. 32, no. 2, p. 97-102.

ESCAP (Economic & Social Commission for Asia and Pacific) and RMRDC (Regional Mineral Resources Development Centre), Sept. 1981, Newsletter, no. 4, p. 4-4.

Exxon Corporation, 1980, World energy outlook: Exxon Background Series, Exxon Corporation, 40 p.

Feldmann, Albert A., 1981, The language problems, "letters": Science, v. 214, p. 1194-1195.

Fesharaki, Fereidun, 1981, World oil availability; the role of OPEC policies: Annual Review of Energy, v. 6, p. 267-308.

Glashoff, H., 1980-81, The multilingual thesaurus of geology, a switching tool between bibliographic data bases developped by IUGS and ICSU AB: Bulletin B.R.G.M., sect. IV, pp. 201-206.

Granick, Lois W., 1982, The emergence and role of common interest groups in secondary information: Journal of the American Society for Information Science, v. 33, no. 3, p. 175-182.

Gravesteijn, J., and Rassam, G. N., 1982, International cooperation in geological documentation; why, when, ways and whereto [abstr.]: Proceedings, Second International Conference on Geological Information, May 23-27, 1982, edited by Claren M. Kidd, Golden, CO.

Green, Julian W., 1981, Geoscience information for developing countries; a forgotten priority: Geoscience Information Society, Proceedings, v. 11, p. 189-200.

Heath, James E., 1982, Cost of new journals: Science, v. 216, no. 4547, p. 684.

Horton, Forest Woody, 1982, A software crisis is indicative of a return to organizational basics: Information & Records Management, v. 16, no. 3, p. 42-43.

International Atomic Energy Agency (INIS), 1979, An introduction to INIS: INIS Today, Austria, 51 p.

Istituto per i Beni Artistici Culturali e Naturali, Regione Emilia-Romagna, 1980, Biblioteche Sistemi informativi e documentazione, materiali di laroro del corso regionale di aggiornamento per operatori di biblioteca 1978: Documenti 14, editor Paola Casagrande, Italy, 190 p.

Kent, A. K., 1981, The information industry in the year 2000: Information Services & Use, v. 1, no. 1, p. 11-15.

Lancaster, F. W. and Neway, Julie M., 1982, The future of indexing and abstracting services: Journal of the American Society for Information Science, v. 33, no. 3, p. 183-189.

Leigh B., 1981, Online searching of bibliographic geological databases and their use at the science reference library: Journal of the Geological Society of London, v. 138, pt. 5, p. 589-597.

Levitan, Karen B., 1982, Information resources as "goods" in the life cycle of information production: Journal of the American Society for Information Science, Jan. 1982, v. 33, no. 1.

Library of Congress, Congressional Research Service, Science Policy Research Division, 1981, Information and telecommunications; an overview of issues, technologies, and applications: Report for the Subcommittee on Science, Research and Technology of the Committee on Science and Technology, U. S. House of Representatives, Ninety-seventh Congress, first session, Serial J, U. S. Govt. Printing Office, Washington, D.C.

Longo, Rose Mary Juliano, and Machado, Ubaldino Dantas, 1981, Characterization of databases in the agricultural sciences: Journal of the American Society for Information Science, v. 32, no. 2, p. 83-91.

McKelvey, V. E., 1972, Mineral resource estimates and public policy: American Scientist, v. 60, p. 32-40.

Maxwell, Robert, 1968, Presentation speech, information storage and retrieval, v. 4, Pergamon Press, p. 87-90.

Moses, Lincoln E., and Glessey, C. Roger, 1981, Energy information: Annual Review of Energy, v. 6, p. 445-482.

Mulholland, J. Derral, 1981, The language problems, "letters": Science, v. 214, p. 1194-1195.

National Federation of Abstracting and Indexing Services (NFAIS), 1976, A study of coverage overlap between major science and technology abstracting and indexing services, report: National Federation of Abstracting and Indexing Services, Philadelphia, PA, 80 p.

Nau, H., 1976, Technology transfer and U. S. foreign policy, Praeger, New York.

Pachersky, T., 1982, Problems of information services with respect to integration of the sciences: Journal of the American Society for Information Science, v. 33, no. 3, p. 115-123.

Price, William H., 1981, The information future and its challenge to information managers: Information Servuces & Use, v. 1, no. 1, p. 3-9.

Rassam, G. N., 1982, Questions and answers: GeoRef Newsletter, v. 5, no. 1, p. 1-2.

Ricaldi, Victor (editor), 1980, Directory of geoscience departments in universities in developing countries 1979-1980: Association of Geoscientists for International Development, 28 p.

Shahad, B., 1981, Iranian Library of Medicine local computer network: Online Review, v. 5, no. 2, p. 133-137.

Tellis, D. A., Conroy, C. M., Matthews, J. E., 1981, AESIS; Analysis and comment, March 1981: Australian Mineral Foundation, Project no. 154/81, Seminar on "Geoscience Numeric and Bibliographic Data", Papers and Recommendations, Adelaide, 30th March - 1st April 1981.

UNESCO, UNISIST, 1974, Using and improving national information systems for development; some practical implications for policy-makers (draft): Paris, France, 45 p.

von Ledebur, Ernst-Joachim Freiherr, 1981, The German experiment in Third World countries; a report on progress after five years in East Africa: Bulletin of the American Society for Information Science, v. 8, no. 2, p. 24-26.

Votaw, Gregory B., 1981, Prospects for Third World development; some thoughts on the 1980s: Dames & Moore Engineering Bulletin, no. 57, p. 5-18.

Werdel, Judith A. and Steale, Richard A., 1978, The information age; worldwide data warfare?: Bulletin of the American Society for Information Science, v. 4, no. 6, p. 10, 12.

Withington, Frederic G., 1981, Computer technology; the state of the art: Journal of the American Society for Information Science, v. 32, no. 2, p. 124-130.

GEOSCIENCE INFORMATION IN MEXICO: STATE-OF-THE ART

Nelly Benveniste

Instituto Mexicano del Petróleo
Av. Eje Central Lázaro Cárdenas 152, Col. San Bartolo
Atepehuacán, Deleg. Gustavo A. Madero, 07730-México,D.F.
México

Abstract--Within the historical, economic and social framework of Mexico, a developing country, the different kinds of Geoscience Information are analyzed.

The management of electronically handled information is discussed, as well as the programs to design data bases at institutional and domestic levels, and to establish a network that will allow not only the retrieval of information from foreign data bases, but also their feed-back, in order to spread the results of technical projects and research in the field of the earth sciences, that are currently published.

Introduction

Whenever we speak of information in developing countries, we must keep in mind their interrelations with other nations of the world, especially the industrialized countries, as approximately two thirds of the world's population is struggling to overcome their poverty while developing at the same time.

The Third World is composed of countries with their own historical, cultural, social, economic and political values, which determine their individual patterns of development. These countries are excercising a collective political pressure upon the industrialized nations to obtain more profitable industrial and commercial policies, as well as a parallel insistence for the restructuration of the international communication media, as part of their struggle to control their economic processes. Most developing countries have attained their political independence, but this has not been followed by the control of their economies, and consequently, they do not have a source of indigenous information.

Within the context of the great impact of the technological advance of industrialized nations upon developing countries, Mexico has strived to establish mechanisms to allow its development, and to adjust to the world's economical crisis, which has affected scientific activities, including the Geosciences.

In the time alloted, I will try to give you a general view of how the geological information is managed in Mexico, and which are the trends for the future.

Publications

Publications can be classified as conventional literature, such as books, periodicals, indexes, etc., that can be acquired commercially or by exchange, and which are readily available. The non-conventional literature includes technical reports, preprints, reprints, patents, thesis, etc., that usually have a restricted circulation.

For the purpose of this paper, I will refer only to conventional literature, but when evaluating the degree of productivity in the field of the earth sciences in mexico, one must take into account that many of the results of research or operational work are included in the non-conventional literature.

An important source of geosciences publications should be the societies and associations of professionals that work in this field, which issue their own publications, and have an elected editor who is responsible for them. Unfortunately, they are frequently published irregularly mostly because of insufficient material, or lack of funds. The editor's dilemma is whether to chose excellence, based upon enough high quality publications that have been reviewed, and among which he can have selecting options, or to make the best of the available material and keep the publication updated, while still striving to attain excellence in the future.

Another kind of publications is institutional, and according to their activities, they are either highly specialized or multidisciplinary.

In general terms, there is not enough dissemination of our publications, and one of its main causes is that they are written in spanish. Sometimes they have abstracts in english which can be retrieved to be included in foreign data bases, and in other cases, the authors prefer to publish their papers in foreign periodicals, especially in the United States, that have more prestige, and a much larger edition and wider dissemination.

To promote indigenous publications, including Periodica, a reference index compiled in Mexico by the Centro de Información Científica y Humanística (Center of Scientific and Humanistic Information, CICH) at the Universidad Nacional Autónoma de México (National University of Mexico, UNAM) which analyzes the more important latinamerican periodicals on science and technology, professionals should be encouraged and motivated to contribute more papers to domestic publications, and associations to promote a closer relationship between domestic and international earth sciences societies.

As to the collections of books, periodicals, indexes and other bibliographic materials in our information units, we depend 95% on foreign publications. The majority are acquired in the United States, and others come from England, France, Germany and other industrialized countries.

Government and private agencies prepare and print maps of different kinds: geological, topographic, geographic, hydrological, of metallic mineral deposits, metallogenetic, non-metallic, etc. in different scales; and aerial photographs taken at different altitudes, in black and white or in color, for photogrametric and photointerpretation studies.

General Services

Information units offer the following services:

- Consultation (catalogs, indexes, microforms, data bases, etc.).
- Translation of earth sciences literature.
- Retrospective bibliographic information (manually, using indexes, or electronically, using data bases).
- Question - answer (by telephone or telex).
- Loaning (personal or interlibrary).
- Photocopies of literature.
- Acquisition of bibliographic material.
- Recent acquisitions information bulletin.
- Selective dissemination of information.
- Catalogs of information units' periodicals.
- Information and distribution of government publications.
- Microform reading and printing equipment.
- Document's control.
- Exchange.
- Services to regional information units related to the geosciences.

In relation to the last point, it would be advisable to establish a larger number of such information units, in order to decentralize the information services.

Computerized Services

The Consejo Nacional de Ciencia y Tecnología (National Council of Science and Technology, CONACYT), offers the Servicios de Consulta a Bancos de Informacíón (Consulting Services to Data Bases, SECOBI), which acts as a liason to multidisciplinary foreign data bases, and to which all the institutions with terminals are connected, thus saving considerable amounts of time and money.

With regard to the geosciences, by using the ORBIT system we have access to the Tulsa data base and GeoRef; through DIALOG to GeoRef and Geoarchives, in the United States, and through QUESTEL, in France, to Geode. This allows for retrospective bibliographic research, retrieval of references, preparation of selective dissemination of information profiles, etc.

The service of selective dissemination of information is offered to institutions, universities, government agencies, libraries, etc. that do not have terminals, and which deal with the geosciences.

There are instances in which institutions that do research and develop technology, give advice, service and manage the libraries on earth sciences of institutions in the same field that do the operational work; for instance, the Instituto de Investigaciones Eléctricas (Institute of Electric Research, IIE) with regard to the Comisión Federal de Electricidad (Federal Comission of Electricity, CFE); the Instituto Nacional de Investigaciones Nucleares (National Institute of Nuclear Research, ININ), with respect to Uranio Mexicano (Mexican Uranium, URAMEX) and the Instituto Mexicano del Petróleo (Mexican Petroleum Institute, IMP) with regard to Petróleos Mexicanos (the national petroleum company, PEMEX).

In the case of Mexican Uranium, this institution imported the hardware, and designed its software according to their needs. These programs include all the information concerning the exploration, development, processing and commercialization of radioactive materials. They have six terminals that service its Delegations, which group several States, thus having an in-house network.

The Institute of Electric Research also has an in-house network, with seven terminals which service the centers where the Federal Comission of Electricity technicians do operational work, mainly on geology, geothermy and power plants.

The Mexican Petroleum Institute has a geophysical data processing center of field information supplied by PEMEX (the national petroleum company). The researchers in this center are designing programs that in a near future will substitute gradually the imported software.

The Mexican Petroleum Institute and the National University of Mexico have computerized their libraries' book collections, to speed-up the bibliographic retrieval. This information can be consulted on paper or through terminals connected to the library.

The Mexican Petroleum Institute also has a statistical data base of the world's oil industry, based on the analysis of both foreign and domestic publications, which can also be consulted on paper or through terminals.

Interinstitutional Cooperation Programs

Besides these services, there are interinstitutional cooperation programs, such as the Collective Catalog of Periodicals of Mexico, which is multidisciplinary.

The Instituto de Geología (Geological Institute) of the National University of Mexico and the Consejo de Recursos Minerales (Council of Mineral Resources) are compiling the mineral and geological information of Mexico.

The Geological Institute has already computerized this information, with its own software programs. The Council of Mineral Resources is still at the stage of compiling the information included in unpublished technical reports, and codifying it. The ultimate purpose is to establish eventually an electronic information center of geological and mineral data.

Therefore, there is a trend to integrate all the geological data of Mexico, so that eventually there will be a domestic computerized information network.

Training

A very important matter related to the management of the electronically handled geological information, is the adequate training of the persons who manage it.

Generally speaking, the purveyors explain how to operate the equipment, but this is only the first step. It is essential that the operators learn how to develop the full potential of the equipment and its use and, furthermore, that they learn how to meet the requirements of the users. Therefore, it is very important to program training courses by specialists, and to apply this knowledge to design adequate communication policies for developing countries, to be able to exchange technical information on the earth sciences at local and regional levels, and to advise less developed countries, in order to apply rationally the technology transfer from industrialized nations.

Conclusions

1.- In Mexico there is a high level of productivity in the geosciences that is not well known due to lack of publishing material, the language barrier, insufficient promotion of publications, and because a relatively large percentage of it is contained in non-conventional publications.

2.- The services of information units are adequate, but it would be advisable to establish more regional earth sciences libraries to decentralize these services.

3.- With regard to geosciences publications and data bases' services, we depend totally on foreign information sources, but there are ongoing projects to design software programs adequate to the specific needs of projects developed by researchers and technicians in this field.

4.- Several institutions have in-house data base networks, but the ultimate objective is to have a domestic information network on earth sciences.

5.- It is important to train technical personnel to operate the data bases' equipment to its full potential, and give adequate service to the users, according to their needs.

INFORMATION NEEDS IN DEVELOPING COUNTRIES

John V. Hepworth

ESCAP Regional Mineral Resources Development Centre (United Nations)
Jalan Jenderal Sudirman No. 623, Bandung, Indonesia

Abstract--Observations have been made on geoscience information systems in a number of developing countries in the ESCAP region, such as Indonesia, Fiji, Papua New Guinea, Philippines, Thailand and Solomon Islands.

The ways in which information is handled, including observational data and departmental records, is outlined. An attempt is made to assess the effectiveness of the systems and to identify the greatest needs.

The basic needs of good management, standard procedures and simple mechanisation are often lacking. Much improvement could be achieved by the development and use of small, appropriate, computerized systems especially in listing and maintaining departmental records, in text handling and report production.

Elaborate, total information systems, although technically feasible (at least in 'the north') are largely unrealistic in the present context. Nevertheless a start is to be made on a computerized Coal Data Base for the Region.

Introduction

The RMRDC is a small group of United Nations consultants who provide advice in the earth sciences to any of the 36 regional member countries in Asia and the Pacific, the ESCAP region. As Information Specialist in RMRDC the writer has been able to make an assessment of the information systems used, and the information needs, in Geological Surveys and Bureaus of Mines (hereafter referred to 'GSDs' to denote the major, national earth science institution in a country which may or may not, in fact, be called 'The Geological Survey', and may or may not include the 'Department of Mines') in some of these countries. The countries of the ESCAP region include such giants as China and India, and such small members as Bhutan and Tonga. There are no 'typical' members, but those on which most observations were made are of the medium size.

The observations have been made during visits to 16 developing countries in Asia and Africa, on matters concerning Technical Collaboration and development in the geosciences, upon specific enquiry into the information systems in 6 developing and one industrialized country in Asia and the Pacific (ESCAP region: Fiji, Indonesia, Papua New Guinea, Philippines, Solomon

Islands, Thailand and Australia), and from a general knowledge of a number of institutions in 'developed' countries.

Although it has not been possible to make a thorough analysis (which might take 3 weeks) of any one organisation's information system, it is observed that these consist of similar recognizable functional elements arranged in a great diversity of ways, that they attempt to carry out a broadly similar set of functions usually with a degree of success considerably below expectations, and that they commonly share similar inadequacies and deficiencies.

The promises of advanced technology and information handling techniques as solutions to their common problems appear, at first sight, to be very attractive.

'Information' is used in this paper in the broad sense to include the range of activities from the observation and recording of scientific data, through processing and reporting, to dissemination, with emphasis towards the latter end of the spectrum.

While similar functional elements can be recognized in the organizations observed, they differ greatly in scope and detail. Also the ways in which information passes around the particular organisation, and the difference between this and the theoretical channels of communication implied by the departmental 'organisation chart' is noteworthy.

In describing the information systems in GSDs of those Asian and Pacific countries we shall consider: Records and Reports; Publications; Libraries. All are integral parts of the information system in a developing country GSD. It is considered that divisions into 'administrative' and 'scientific' information, although often (but not always) made is artificial, and is also inappropriate to modern methodology.

The Departmental Records

Most of the basic information or data bank of a GSD is generated as written reports, whether describing a laboratory determination of a mineral, the investigation of a prospect, or an account of a mapped region.

Comprehensive lists, much less bibliographic reference systems, of these records (reports) rarely exist in developing countries. Surprisingly, even lists of published records may not exist. A notable exception is Papua New Guinea in which the Geological Survey recently carried out a complete listing of every report, published and unpublished, in the 'Catalogue of Data Files, 1980'. This list, unique in the developing countries known to the writer, is inevitably rather cumbersome, being reproduced on A4 paper and about 3 cm thick, and cries out to be transformed into a computerized data base.

About 2500 to 3000 record titles are listed, divided into map sheets.

There is no index, and no key-wording, and retrieval must be through geographic location, by a limited number of place names, by a limited number of subjects, or by looking through the whole thing.

This document is of great interest for several reasons, first that it exists and is a unique attempt at total record storage in a GSD, second that the concept obviously owes its origin to computerized data handling, and third because it is related to microfiche data storage. Thus the main reason for making the list arose from a decision to microfiche all records because of lack of space for the files, and this enforced the sorting and listing of all existing files. The working record is therefore the microfiche: unfortunately the Department had only a small microfiche reader, no printer and no copier (1981), readily identifiable 'needs' in this developing country.

The Papua New Guinea system by its advanced nature, high-lights its own 'needs': the need for key-wording, the need for indexing, and for methods of data retrieval. Clearly the 'Catalogue of Data Files' is suitable for computerization, at an initial level. With additional work it could be given more elaborate bibliographic status and search capability.

Although the Papua New Guinea 'Catalogue of Data Files' system stands out in the extent and orderliness of its presentation of record titles, some approximations to it exist elsewhere. For example, the Indonesian Department of Mineral Resources lists many of its reports on the Library card index and uses a modified UDC classification. However, (and this is the reason for taking this particular example) the old Geological Survey of Indonesia has recently split into four directorates each with a budding library and information system of its own, and there is no common indexing or reference system.

At the other end of the spectrum in some countries known there is no list of departmental records or reports in existence, and no way of 'retrieving' other than shuffling through the files, if they can be found.

In the middle of the spectrum there are countries which are more or less well-furnished with lists of reports, published and unpublished. In Fiji for example there are the 'Bibliography of the Geology of Fiji' (Duberal and Rodda, 1969) and 'Mineral Deposits of Fiji' (Colley, 1976); and in Solomon Islands 'Mineral Occurrences of the Solomon Islands' (Arthurs, 1979) and 'List of Company Reports held on Open File', and 'List of Publications of the Geological Survey' is awaiting completion. In Philippines a 'List of Publications of the Bureau of Mines and Mineral Resources' is published periodically by the Library Section of the Mineral Economics and Information Division. However, delays in compiling, and publishing such lists means that even in those developing countries where publication lists and bibliographies are produced these are often several years out of date. The advantages of computer-produced cumulative lists are obvious.

In Solomon Islands a 'universal' filing system has been in use for a number of years and this in itself does improve the record retrieval capacity in as much as more information tends to be entered into the system than in the divided system. However, it unlikely that any record will appear in more than one file so the options for retrieving are limited.

Some estimates of the magnitude of records held in GSDs can be made and are given below (Table 1.). However, experience has shown that this apparently simple aim of establishing how many records are stored in a department is fraught with problems, first by confusion as to what constitutes a record; whether published and unpublished information should be treated together or separately and how these terms are defined; the plain lack of cumulative lists; and lack of time and opportunity to walk around other people's offices classifying and counting their accumulated papers. Within these limitations the following figures are given. Almost certainly there is a large hidden and unlisted mass of records in each case, especially with 'administrative affinities'.

Table 1.--Records held in some GSDs

	Fiji	Indonesia (CGRD)	Papua New Guinea	Philippines	Solomon Islands	Thailand	Australia
Published	78	94	2500 to 3000	346	163	43	470
Unpublished	242				436		5000
Computerized data system	No	In part	No	In part	No	No	Yes

The records refer to single title documents concerning the geology/ mineral resources of the country.

The above brief description of the record system in some developing countries leads to a fairly clear recognition of common information needs, which may appear quite obvious:

a. all records should bear explicit titles;
b. titles should be briefly informative and include up to about 3 key-words;
c. all records should carry brief bibliographic data such as author, date, location, primary subject classification, and serial number;
d. every record (report), from whatever division of a GSD, should carry the above brief bibliographic details and be entered on to the central record system, or data base, whether thus is manual or computerized. This would have to be recognized as an obligatory action and supervised.

Obvious as the above 'needs' might appear for the systematic designation of records, they are the minimum requirement to establish a coordinated information system, and it is certain that this is very rarely achieved in developing country GSDs. Thus out of 7 countries studied only 3 have anything approaching it.

The number of data elements proposed is very small: this is deliberate in the hope of encouraging adoption as routine (see Fig. 1.).

Evidently, at this stage we have arrived at, or very close to the need for a computerized data base. Few of the developing countries under

PROPOSED 'BRIEF BIBLIOGRAPHIC RECORD'

DESIGNED TO BE PRINTED ON COVER OF EVERY DOCUMENT TO ENSURE MINIMUM REFERENCE DATA

BRIEF BIBLIOGRAPHIC RECORD: EXPLANATION
Country name: *printed on all files.*
Department name: *printed on all files.*
Document number: *year, sequence.*
Date: *As appears on document.*
Document description: *Report, Bulletin, Memoir, Conference Document, Thesis, Map, Monograph.*
Status: *Published, Unpublished, Open File, Confidential.*
Title: *As on title page.*
Author(s): *Surname(s) and initials.*
Locality: *1. Name* *2. Name of district* *3. By reference system e.g. national, geographic, coordinate, grid, alphanumeric etc.*
Subject description: *Descriptive terms, key words, perhaps thesaurus.*
Summary: *Short indicative Summary.*
Classification: *From some chosen system e.g. UDC, DDS or other.*

BRIEF BIBLIOGRAPHIC RECORD: THE PRINTED FORM
Country name: **REPUBLIC OF UTOPIA**
Department name: **GEOLOGICAL SURVEY DEPT.**
Document number:
Date:
Document description:
Status:
Title:
Author(s):
Locality:
Subject description:
Summary:
Classification:

Fig. 1.

consideration have begun to computerize either their 'administrative' or 'scientific' information systems (see Table 1.). In Indonesia, however, the Mining and Minerals Information System aims to cover the entire range of mining and minerals resources information within the Directorate General of Mines. Also Pertamina, the state oil corporation, has an advanced system. Papua New Guinea has probably gone furthest in organising its data in a form virtually suitable for computer storage but has yet to acquire a computer. In Philippines it is intended to install a total computerized information system to cover the entire range of records of survey and mining activities covered by the Bureau of Mines and Geosciences. The other countries studied do not have computers, at least for geoscience information. However, the on-going Indonesian — Australian regional mapping, 10 years reconnaissance of Irian Jaya has developed a computerized system for all field and analytical data based on Hewlett-Packard 1000 (45) system (see RMRDC, 1981). The Indonesian — UK Kalimantan Coal Exploration Project devised a system to store a large amount of borehole and detailed mapping data (Thomas et. al., 1980). The Indonesian — UK North Sumatra Project used a computerized system for a reconnaissance geological and geochemical survey and this is fully operative (Page and Young, 1981).

It is not yet clear to what extent the departments that have attempted to computerize their geoscience information have benefitted. On the other hand it is abundantly clear that all would benefit from a more systematic and orderly treatment of their information. Perhaps a main, indirect benefit of computerization is (as in the example of Papua New Guinea) in introducing this sense of order. It is then illogical to argue for 'computerization without computers': it is more a matter of applying the concept of 'appropriate technology' to information and of identifying hardware and software within the price range (say US $ 50,000) and within the organisational capacity of the GSDs. Clearly there is a need to do this, and possibly to develop new systems that are appropriate. Thus systems such as CDS/ISIS and MINISIS may not be appropriate (in cost and training required and in upkeep and maintenance) to perform the comparatively simple functions that appear to be necessary for basic information handling (filing, sorting, indexing and limited search capacity) in the developing countries.

Publications

From 'reports' it is a natural step to 'publications' and it is clear that a certain amount of information has to be published by a GSD. There are, however, some countries in the ESCAP region that have never published any geological document, some that may publish one document a year, while one country (China) publishes over 500 serials in a year! In a number of countries it is not possible to find out, except by personal enquiry on the spot, what if anything is in fact published. At the same time countries with as much as ten years back-log of material awaiting publication are known — and this in a region where small printing businesses are plentiful, skilful and inexpensive. One of the clearest information needs in the developing countries is to find ways of removing the obstacles to publication.

Dramatic changes in rate of departmental publication usually take place as a result of T.C. projects, although 'northern' donors and planners have been painfully slow to learn that production of reports and publications is a slow, skilful and expensive business. For example, one T.C. project which made a reconnaissance of 190,000 km^2 finished with its maps and reports drafted but an additional two years project was required to print and publish them. This kind of underestimation of information requirements is not at all unusual and reflects the 'poor relation' attitude towards dissemination.

A comparable, although smaller, example of improved publication was provided in one ESCAP country that had published no map (including national geological map) during some 30 years since independence. Carefully selected technical assistance (field checking, map drafting, printing knowhow, provision of screens and locally unobtainable equipment) resulted in a flow of locally published maps which continues and nor produces about 5 sheets per year. The need for this kind of 'technology transfer' by intensive training and example in specialized information methodology such as the production of geological maps (and perhaps reports) cannot be over-emphasized.

Apart from the initial difficulties suffered by many geoscientists in writing accounts of what they have done, the ensuing processes of typing, correcting and editing (sometimes in a second language, English) are often overpowering. These contribute to the delay in publishing so commonly observed.

It is understood that in USGS it is normal practice for reports to be typed on word processors. No such machines (and few memory typewriters) are known to exist in the GSDs under discussion.

It may be supposed (but by no means proven) that the facilities of the word processor in helping to produce clean well-edited text are as desirable in the developing countries as in the developed ones. Thus the possibility of producing, relatively easily and quickly, camera-ready copy suitable for litho-offset reproduction, would seem potentially to be a way of cutting through many a publication log jam.

The constraints to the use of word processors in developing countries are, first, financial — a basic IBM Displaywriter costing US \$ 15,000 in Indonesia; second the lack of choice of makes and models which are inevitably aligned towards business, and third the problems of inadequate or non-existent maintainance.

Nevertheless the word processor, with its editing (and spelling correcting?) facilities demands careful consideration as a desirable technical need, or perhaps a study assessment is needed of alternative ways of providing such equipment to developing country GSDs. No T.C. project is known to the writer which has included a word processor to assist in report production, although electronic equipment ('data-generators' rather than 'data-processors') comparable in size and cost such as atomic absorption photospectrometers for example, have long been acceptable hardware for T.C. planners and donors.

There is also some reluctance on the part of information people to accept the word processor (which is a standalone microcomputer) as 'serious' and an inclination towards the supposedly more powerful terminal to a mini- or mainframe. It seems that there is a real need to examine the capability and suitability of existing word processors as practical means of facilitating publication (and in elementary data handling) in developing countries. Also perhaps to design and manufacture machines appropriate to the needs of GSDs and similar organisations.

Libraries

Library accomodation in the developing countries varies from large well-lit rooms with adequate reading and administrative space to others in which the bare necessities exist. Most wear an air of fading respectability and suffer from low incomes and inflated prices. In one library (Solomon Islands) a scatter of geoscience literature shares room with other disciplines. In another (Philippines) miraculous feats of double flooring and close packing contain (but not for much longer) the huge stock. In yet another (Fiji) the space problem — including the pressure of increasing amounts of marine survey data — in addition to library and records, has produced plans, still unrealized, for an 'information building' including a library.

Table 2.--Some examples of library holdings, from verbal information

<table>
<tr><th></th><th>Fiji</th><th>Indonesia (CGRD)</th><th>Papua New Guinea</th><th>Philippines</th><th>Solomon Islands</th><th>Thailand</th><th>Australia (BMR)</th></tr>
<tr><td>Total Monographs</td><td>1264</td><td>7067</td><td rowspan="2">520</td><td>20000</td><td></td><td>40000</td><td>18000</td></tr>
<tr><td>Text books (where distinguished)</td><td></td><td></td><td>3000</td><td>320</td><td>1400</td><td></td></tr>
<tr><td>Serials</td><td></td><td>580</td><td></td><td></td><td>2000</td><td></td><td></td></tr>
<tr><td>Reprints</td><td></td><td>5441</td><td></td><td></td><td>4000</td><td></td><td></td></tr>
<tr><td>Serials received</td><td></td><td></td><td></td><td></td><td></td><td>55</td><td>2500</td></tr>
<tr><td>System</td><td>DDS</td><td>UDC</td><td></td><td>L of C</td><td></td><td></td><td>UDC</td></tr>
</table>

The figures are obviously inadequate but are included to give an indication of orders of magnitude. An eyeball assessment of the library shelves usually reveals a majority of old well-worn books and a deficiency of new dust covers.

Needs are easy to recognize: more money for book purchase, better facilities for library work, and probably some change of emphasis in the role of the library from a mausoleum for dead books towards a service for supplying information.

More than one library is known where no complete set, or list of departmental publications exists, and neither departmental publications nor reports are indexed. On the other hand in another institution (Asian Institute of Technology, Bangkok) a computerized library system has been installed (Michael Sherwood, pers. comm.) which extends from book acquisition, to printing-out library catalogue cards. But this is extremely unusual and probably not an example to be widely followed. Only one other computerized geoscience library service is known in the region, the BMR, Australia where a computerized index is being built-up, together with a restricted on-line enquiry service through the reference librarian.

A number of T.C. projects have been performed to improve library services in the region, and the professional level of librarianship is high. Looking for 'needs' that are capable of being realized it may suggested that

a. continuation of such projects should be positively encouraged;
b. bodies such as I.U.G.S. and ASLIB might look more actively into ways in which they can cooperate in library work, including exchange visits;
c. that studies should be made of actual deficiencies in books and periodicals caused by lack of funds in geoscience libraries in the developing countries, so that real figures can be used and brought to the attention of the international aid donors such as ADB, to UNESCO, EEC and so on. This is necessary because, while lip-service is widely paid to the need for information, few donors consider books (and libraries) worthy of more than low order financial support.

Editing and Language

Although only a minority of the ESCAP region's population has English as a first tongue, their geoscientists are normally expected to be able to write in English suitable for publication. This is, in fact, often not the case, and considering the difficulties that even native English speakers have in writing English at a publication standard it is not surprising that 'the editing problem' constitutes one of the most obvious impediments to the dissemination of information.

The promise of automatic translation, although already taking place to a small extent (e.g. Bulletin Signalétique) belongs more to the final section of this paper than to immediate practicalities.

One need is to recognize that there is, in any organisation, a requirement for editing, that this is a slow, time-consuming job and cannot be fitted into to the run of administrative duties. The only organisation in the region known to the writer that does recognize and make provision for this is the BMR, Australia with six editor posts. In all other regional GSDs editing is assumed to an 'invisible asset' and there are no editorial posts. There is, however, the familiar back-log of unpublished, or never-to-be-published work.

One slightly oblique solution of the editing problem is adopted by one regional country (Philippines) in issuing most of its reports as 'open-file, unpublished' reports, with the statement that they are not stringently edited. In fact they closely resemble the 'published reports' which are issued in much more limited numbers. It may be considered that to lower the standard of editing is a sensible partial solution to the problem. Outrageous as this will seem to purists it is perfectly possible to convey meaning clearly in bad or ungrammatical language. Nevertheless it would require a programme of investigation to establish what kind of 'bad' English was, in fact, a positive and acceptable improvement. Perhaps editors should be chosen who are not familiar with the more subtle points of the language?

There is probably no lack of sufficiently qualified and experienced people in the Region, and in the industrialized countries, who could undertake the editing of currently written reports and papers. It is more a matter of identifying a suitable agency (and funds) which could draw up lists of suitable persons and then act as intermediary in apportioning the work. Some of the suitable persons would have retired from active geology, and others would be people with a limited amount of time available. RMRDC might undertake this role, and AGID might like to consider the possibility of creating a panel of voluntary editors who would be prepared to undertake editing on request.

It is allowable to think of memory typewriters and word processors being used in such editorial exchanges, disks being sent by mail between client and editor: or even (final section) by telecommunication link with discussion of problems at convenient terminals.

In the realm of reality, however, a simple and resolvable 'need' for every GSD is to have at least one copy of the new "AGI Glossary" in the department.

The International Geoscience Data Bases

A powerful component of present day information work is the international data base such as GEOREF, GEOSYSTEMS and GEODE-PASCAL. To the extent that these are accepted as a 'need' of the 'northern' geoscientist they may also be presumed to be necessary to his 'southern' equivalent in the developing countries. The validity of this assumption is open to discussion, but the constraints that have prevented the establishment of a regional data base while primarily financial, also include lack of an organisation to undertake it. An alternative to establishing a regional geoscience data base would be to use the 'northern' data bases. But as far as the writer knows no GSD subscribes to an international data base and the T.C. project which includes, say, a 5 year subscription to one of the international data bases has yet to be conceived.

One step has, however, been taken in establishing a regional data base. As a part of the UN Regional Energy Development Programme it has been agreed that ESCAP/RMRDC shall set up a Coal Information Service, including a bibliographic Coal Data Base for the countries of the region. At the same time the opportunities for exchange of regional information for world-wide, or 'northern' coal information will not be neglected. The Coal Information Service will comprise three parts, to include data acquisition; data storage, processing and retrieval; and a facility for current awareness in the form of a newsletter or bulletin. When the working model has been established it will be handed over to an ESCAP regional institution to run.

It can be argued that if a coal data base is needed for the region then equally a minerals and mining data base is needed. Also Berger (1980) has suggested that the 'northern' data bases only capture about half the 'southern' data. Perhaps one solution is improved acquisition by the international data bases (which really means better feeding by 'southern' organisations) and subsidized dissemination of information, including the provision of telecommunication links and terminals to the developing countries. It is by no means certain that such a solution is acceptable to either 'north' or 'south'.

The writer does not know the answers to the questions raised but suspects that to establish data bases on comparatively restricted subjects (for example the Coal Data Base, which might be followed by perhaps a Clay Minerals Data Base or Metallic Minerals) rather than to attempt total coverage may be the right solution.

As an example the Asian Institute of Technology's (Bangkok) geotechnical data base AGE may be cited. This appears to be successful within its field which is clearly defined and only embraces part of the geoscience literature.

Conclusions

Observations have been made on information activities in some of the developing countries of the ESCAP region.

It has been concluded that significant improvements could be, and in most developing countries need to be made, in the handling of information at a managerial level. This could be facilitated by the provision and use of some of the simplest equipment which is often surprisingly absent and unavailable locally. Antiquated methods inherited from the past are still common and modern ways of classifying, indexing, filing, recording and circulating information still await introduction into many countries.

Thus it is stilloften impossible, or very difficult, in many GSDs to make a retrospective search for information about, say, a mineral such as mica, the only route being through the brain of some old and faithful resident of the library, or an individual who has worked on the topic. Indexes, card catalogues, or files are often not maintained, and lists of neither report titles nor the periodicals produced are kept. There are, of course some countries where the basic information handling is done moderately well and, rarely, with exemplary efficiency. More often, antiquated methods inherited from the past, or 'ad hoc' systems thought up by administrators, or none at all, exist. The need for the adoption of orderly, systematic and disciplined management of a department's files is obvious: perhaps the greatest contribution the computer revolution in information is in the obligatory sense of order in data entry that it brings. Given, then, that this first 'need' is satisfied, by a rigorous analysis of the existing information system and a good shaking up into the form suggested by a clear identification of objectives, it is possible to proceed further and ask whether, and at what stage computerization should be introduced.

If we identify the primary information functions of a GSD as being to produce recorded data in the form of reports (i.e. narrative records) and to communicate effectively through correspondence and publication, it is apparent that modern methods have much to offer. It is arguable as to whether a piecemeal approach is to be preferred — office calculators and accounting equipment plus memory typewriters plus word processor, plus mini-computer for observational data such as geophysics and geochemistry, versus a computer configuration capable of handling all these aspects, including accounting, text handling and numeric data processing. The answer may be predetermined by finance, in which the hardware and software to accomplish the extended, or total, system must cost in the region of US $ 250,000 to US $ 500,000 which is outside the range of probabilities for most GSDs of the developing countries. Therefore the piecemeal approach is probably unavoidable, in which case the writer would place at a high priority the acquisition of a standalone word processor, with the best indexing, classifying, sorting, retrieval, and arithmetical programs that are available for the particular make and model. Thereafter the qualitative step-up to equipment capable of running, for example CDS/ISIS or MINISIS, plus a management system, means a financial increase of an order of magnitude (US $ 20,000 to US $ 200,000).

At last, we may dream a little: supposing that each GSD is a national centre for earth resource information, on which ministries depend for decision material; and which is respected and supported by a public which believes that the geologist is the revealer of true national wealth. Communication between departments and ministries would be televisual, and by 'electronic mail'. Filing of incoming information, both administrative and scientific, would largely be automatically entered, and retrievable at the interactive terminals in every office.

Typing of all reports would be on word processor, or by the text-handling facility of the mainframe computer, and authors would correct first drafts on hard copy printed out for them, and subsequently they and the editor would check final drafts on the v.d.u. Spelling would, of course, be checked automatically, together with lay-out, spacing, hyphenation and tabulation. Printing of camera-ready copy would be rapid and of high quality so that the department's publications would appear quickly in elegant, printed form.

<u>All</u> departmental records would be stored and processed in a data base allowing search and retrieval by key-words, concepts, report numbers, location, dates and authors.

Several terminals in the Library would allow on-line reference to one or more of the international data bases, the regional data bases such as the Coal Information Service, and to sister surveys' records systems. Dedicated transmission links would be available.

Microfiches could be demanded on the communication system, transmitted and printed as hard copy as desired, and would actually be legible. The Library catalogue would be on cards or computer, the Librarian would handle purchase and acquisition, classification and dataloguing and stock control from her/his terminal.

A modest expenditure of between US $ 0.5 to 1 million would secure for any GSD of a developing country such a dream facility.

The object of this paper has been to identify some information needs in the geoscience organisations of developing countries: that most of them will appear fairly obvious does not detract from the necessity of recognizing them. Nor is this the place to spell out the methodology, but in conclusion a list of steps to be taken, each of which virtually amounts to a project objective, is given below:

i. Improve and upgrade office management, using simple well-tried aids, especially to handle correspondence and observational data.

ii. Devise acceptable and appropriate data entry formats; and encourage the use of keywords and thesauri to create brief bibliographic data for all records.

iii. Encourage and assist these aims by T.C. programmes and by support from international organisations such as I.U.G.S. and its constituents, e.g. AGID.

iv. Promote and encourage widely the construction of departmental record lists and indexes.

v. Investigate and encourage provision of small computer systems to support steps i. to iv., and encourage integration of 'scientific' and 'administrative' data.

vi. Establish computerized data bases on specific subjects in the geosciences.

vii. Investigate ways of facilitating editing of reports, for internal use and publication, in two ways
   - formation of English speaking editorial panels under T.C. programmes or voluntary arrangements,
   - promote use of word processors for text-handling, editing and production of camera ready copy.

viii. Promote 'technology transfer' in map production.

ix. Support Library improvement, including grants for text-books and periodicals.

x. Provide subscriptions to the northern international data-bases, and barter for their information.

xi. If, however, the scale of T.C. funds available should rise to the order of US $ 0.5 to 1 million for improving information services in any one developing country's geological survey/ bureau of mines it would be possible to make a quantitative leap ahead to an 'ideal' information system with extensive e.d.p. facilities.

References

_____ 1980, Catalogue of Data Files. Geol. Surv. Papua New Guinea, Dept. of Minerals and Energy.

_____ 1981, List of publications, Nov. 1981. Library Section, Mineral Economics and Information Division, Bureau of Mines and Geosciences, Rep. of the Philippines.

_____ 1981, Bibliography on Philippines geology, mining and mineral resources 1972-1975. Pt I Author and title index, Pt II Subject Index. Library Section, Bureau of Mines and Geosciences, Rep. of the Philippines.

Arthurs, J.W., 1979, Mineral Occurrences of the Solomon Islands, Bull. Geol. Surv. 13.

Berger, A.R., 1980, The Usefulness of computerized bibliographic data bases for monitoring the geoscience literature on developing countries, Int. Geol. Congr. 26 (12) Abstracts. p. 843.

Colley, H., 1976, Mineral deposits of Fiji (Metallic Deposits). Mem. Min. Res. Div. Fiji 1.

Duberal, R.F. and Rodda, P., 1969 (?). Bibliography of the geology of Fiji. Dept. Geol. Surveys, Min. of Nat. Resources, Fiji.

Greenbaum, D., 1979, A catalogue of non-confidential mineral exploration reports (M.R.D. Exploration Report Series). Rep. Geol. Surv. Fiji 12. (with supplements).

Manser, W., 1978, Earth Science Abstracts 1974-1976. Mem. Geol. Surv. Papua New Guinea 5.

Page, B.G.N. and Young, R.D., 1980, Anomalous geochemical patterns from northern Sumatra: their assessment in terms of mineral exploration and regional geology. Journ. Geochemical Exploration, 15, p. 325-365.

RMRDC, 1981, Newsletter, No. 4, RMRDC, Bandung, p. 4-20.

Thomas, L.P., Hardjono, Land, D.H., Addison, R., 1980. Coal Exploration in East Kalimantan, Indonesia. Int. Geol. Congr. 26(14), p. 327-334.

IN-HOUSE INFORMATION ONLINE:

CREATING A SMALL DATABASE

Unni Havem Rowell

Exmin Corporation
P.O.Box 2655, Bloomington, IN 47402

Abstract--The most familiar type of data file in the scientific information community is the commercial, bibliographic database made available to the public through large, online service systems. In recent years a huge number of less visible files have been established by federal and state agencies, industry and research institutions, due to the security or special vocabulary or indexing depth required for data collections of a confidential or highly specialized nature. Additional demands on the performance of a file, as minimum retrieval time, system flexibility, consistency of information, and coordination and combinations of data elements have further led to extensive use of interactive systems.

The decisions to create a private file, with the initial analyses and evaluations, with plans for format, functions and development, for its administration and use, and with the choice of systems and software, are complex and need to be the joint responsibility of management, operator, and end user.

The present paper attempts to analyze the functions and characteristics of an in-house bibliographic data file, in contrast to those of a commercial database. The practical steps in the development of a database are treated in some detail and illustrated by the author's experience with and examples from an online database designed for a small exploration and mining company. The emphasis is on choice of system, on database definition, structure and size, on indexing vocabulary and the various kinds of bibliographic data included--with particular concern for performance, ease of use, cost and security.

Introduction

To develop a database is *great* fun--if you are in excellent health, have strong nerves and a loving family, and needless to say: time and money to burn. If you miss out on any of the above and still go ahead, either you are a masochist, or your need for improved data management is serious indeed.

I might as well state my position at this point: I am not an expert; I know next to nothing about computer languages and program analysis. But, as a research geologist in an exploration and mining company, the newly established U.S. arm of an international concern, I was given the task of setting up a data management system for our collection of documents. All kinds of documents: telexed, taped, typed, blue-lined, xeroxed, drawn-in-ink, photographed, printed, or hand-scrawled--they all needed a base. My directives about hardware were none; I could use vertical files or computers, as long as security and privacy were airtight, retrieval time minimal, the system functional, flexible and reliable--and operational within a few months. My paper will be based on my experiences, but rather than present a complete case history, I will give my requirements for private vs. public databases and explore some characteristics of in-house files that I find particularly interesting or intriguing. Finally I will illustrate some practical steps in database development with examples from our own system.

I have chosen to take the term *in-house* literally and will deal only with files containing limited amounts of local data, as for a research project, a special library, or a mining company.

Database Requirements

In-House Files

My requirements for an in-house file--I like to think of it as *Tailored information*--are ranked as follows:

1. Coverage in depth of a limited subject area
2. High security
3. Minimum retrieval time
4. Positive cost/benefit analyses
5. Performance quality
6. Simple procedures for entry and retrieval of data

Commercial Databases

Similarly, my priorities as a user for a commercial database, or *Information off-the rack*, would be:

1. Coverage of a broad area of related subjects
2. Low costs
3. Performance quality
4. Document relevance
5. Document availability
6. Frequent updating

Database Characteristics

Environmental Characteristics of an In-House Database

The reasons for setting up an online database vary greatly; improved information access is the prime objective; system flexibility, file coordination, information sharing and consistency, a special indexing vocabulary, and file protection are other factors. Special environmental features of a private file are: it is designed not only by administrators, but also by operators and users; the data are usually indexed, entered and retrieved by the very same people, some of which are the end users; operators are trained in data management on-the-job within a limited subject area; access to the file is limited, sometimes possible only through power and in-house status, while gold paves the way to the commercial database.

Database Contents

Database scope and coverage. An in-house file is designed to accommodate a special data collection for a special group of people. It holds data from both public and private documents, regardless of format and must be self-contained and cover all aspects of a subject. A commercial database, on the other hand, has data from many related subject areas of potential interest to many types of users, but is limited to information in the public domain. For special aspects the user must rely on a combination of databases.

Indexing vocabulary and depth. The indexing vocabulary of an in-house file does not pretend to give a complete outline of a subject--as should a thesaurus--but is practical and grown out of a particular collection of data, with uneven coverage of different subject areas. A standard thesaurus would hardly do. The number of index terms for each document is limited by computer storage costs, and since it is more important to pick out relevant information than to get a description of level, approach, or format of a document, *identifiers* are chosen in indexing, rather than *descriptors*. Keywords are typically informative and will not merely state that a document deals with absolute age, or that something is Devonian, but give specifics and often also relations or directional information.

Database Format

In a database of your own design the number, types, and combinations of bibliographic elements depend solely on your own data. The format can be simple, and no coordination is needed with sections in a printed index; a feature I often find disturbing in commercial databases.

Practical Use of the Database

Data entry. Most data management systems allow use of word processing equipment or conventional terminals for file loading; some require transfer of the data to magnetic tape, while others have automatic data compiling. Data are conveniently entered directly from the printed text of periodicals and books on a word processor, a mini-computer, or a terminal with a memory and preferably a CRT screen for immediate editing. With a collection ranging from maps through correspondance to slides, and a terminal without editing capabilities--so that file loading takes place directly online--it is useful first to prepare a bibliographic worksheet for each document.

Search strategies. Because of the small, but highly specialized selection of index terms, or a simple format, search strategies in private file tend to be straightforward and therefore rather rigid. Nevertheless the file is expected to yield information of both high precision and high recall.

Practical Steps in Database Development

Initial Analyses and Decisions

Analyses of the status and performance of the existing file system, expectations of type and size of future data collections, and the information needs and satisfaction level among the users will indicate whether there is a real *need* for an interactive database. The cost and time involved in the conversion and in the operation of a new system may, however, be the *deciding* factors.

Choice of System

A choice of system means evaluation of both program packages and equipment. A good system should save us trouble, space and time--maybe even money in the long run? The insiders advise us to try it all out, and at Exmin we had our fair sampling.

We found that a word processor is only a good choice when training for the Olympics in disc throwing. Constantly flipping diskettes in and out of the drives is exhausting. Admittedly there is ample time to recover while waiting for the printouts. We never seriously considered getting our own mini-computer but explored the option of time-sharing on a fair-sized neighbouring computer and purchasing the software. In our experience the library-oriented software was far less expensive than that intended for the business market. The stumbling block was our absolute privacy requirement. We were also interested in the private file services of the commercial online systems, but again the privacy at the stage of data transfer was at stake; a problem which later has been solved.

Finally our choice was made: to rent both information processing programs and computer storage from a large national company, a time-sharing network with additional batch-processing features.

Database Definition

We started the design of our database by listing our essential bibliographic elements, dividing them into nine fields, setting the maximum length of each field, deciding which fields needed to be repeatable, always multiple, or single. For the resulting worksheet, see the Appendix. Five fields are required, and all except title and source fields are searchable. Since many of our various "documents" have no title at all, we chose to rely on keywords instead of searching on words in a constructed title. The fields are combined in four records which can be searched separately without going through the whole database.

Decisions on Indexing Policies

Decisions on indexing rules and policies are time-consuming indeed. Which are our prime subject areas? Which are less important, and which can be disregarded? Do we want a pre-designed thesaurus? A hierarchical vocabulary? Or more down to earth: Do we need both *granite* and *granites*? How should we distinguish consistently between the 17 American Clay Counties? Do we want to split the Precambrian into narrower age terms?

At Exmin we favour the lexical principles of the *GeoRef* thesaurus, but add or change as we please: use very general terms, or none at all, for issues peripheral to us and go into minute detail with multiword terms for our specialties. We have f.i. added substantially to the exploration terminology. The format of the program allows terms to be added as we go along, and we keep an authority file with detailed directives and cross-references for each keyword used.

Loading the File

From the data sheets we can enter the data at our TI65 terminal, either directly or using the bubble memory, creating a file in the workspace, and subsequently saving it in the preliminary mass storage. Editing or listing of the file takes place in the workspace, but no searching. A pre-processor program has been written in an extended FORTRAN-4 version in time-sharing format with line numbers for easy editing. The compiler will check for syntactical errors before compiling the data and loading them onto permanent disc storage.

Testing and Adjusting the Program

The stage of testing is interesting, but quite frustrating. You add and change and change and add. A new system even needs to be changed and added to after the testing period. Maybe the specifications were not specific enough? Maybe the situation and the requirements changed more quickly than anticipated? The final test comes with the use of the database. Does it really meet the needs of the end users? Did the marketing people promise more than the system can give? At Exmin we wrote an operator manual describing the system and the procedures in detail and found that to be a valuable exercise leading to important program adjustments. Unfortunately we also found, quite recently, after three months of testing and development of what now seemed a very workable database, that the operating costs dramatically exceed the estimates given in our contract--and have yet to decide where to go from here!

Afterthoughts

We are constantly and wisely reminded to seek informed advice before deciding on a data processing system. The lucky ones among us are part of a large institution or concern with computer specialists and facilities at hand. Others have to get advice from outside. The field of database planning and management has developed explosively the last few years, and--to use a tectonic metaphor--subject spreading has split our informed *outside* into one library science plate, one EDP plate, and one business application plate that already have drifted far apart. The gap between software knowledge and a real understanding of specific user needs is a serious one. Have you ever tried to get clear statements on software performance and of actual expenses from the marketing representatives? We are told that a particular program is *exactly* what we need, when it obviously has been designed for a completely different purpose and only can serve us in an inefficient and expensive way. We are given costs per storage block, or worse: per SBU, and meet with great reluctance to translate this into bytes, hours, or bibliographic units. The rates for the use the various intelligence levels of a program package can remain an eternal mystery--after all: an analyst writes no invoices, and a computer answers no questions.

Any solution? Advice. A d v i c e . A D V I C E ! At an early stage. From consultants--but do find a knowledgeable one. From experienced users--find one compatible to you.

One last thing: It takes

T I M E

to develop a database: the experts reckon from one to four years, according to your shoe size. Fortunately most of us have an old system that will have to do in the meantime.

A few references to in-house database development literature:

Davis, B., 1975, *Database management systems: User experience in the United States*: Manchester, Engl., National Computing Centre, 81 p.

Database directions; the next steps: Proceedings of the Workshop of the National Bureau of Standards and the Association for Computing Machinery, Fort Lauderdale, FL., 1975. Washington, D.C., U.S. Govt. Print. Off., 1976, 158 p.

Levine, G.R., 1981, Developing databases for online information retrieval: *Online Review,* v.5, # 2, p. 109-120.

Hagberg, B.S., 1978, Deere archival retrieval system: *Agricultural Libraries Information Notes,* v. 4, # 4/5, p. 1-2.

Swanson, R.W., 1980, Probing private files: *Database,* v. 3, # 2, p. 70-76.

Swanson, R.W. and Engler, J.A., 1980, Probing private files: Polaroid Corporation's photoindex: *Database,* v. 3, # 4, p. 57-67.

IN-HOUSE INFORMATION ONLINE: (APPENDIX)
Creating a small database

Handout at the Golden conference, giving an outline of the topic of the paper presented.

A. Introduction

B. Database Requirements

In-House Files *"Tailored information"*	Commercial Databases *"Information off-the-rack"*
1. Coverage in depth of a limited subject area	1. Coverage of a broad area of related subjects
2. High security	2. Low costs
3. Minimum retrieval time	3. Performance quality
4. Acceptable costs	4. Document relevance
5. Performance quality	5. Document availability
6. Simple procedures for entry and retrieval of data	6. Frequent updating

C. Characteristics of an in-house database

1. Environment
   a. Purpose: *Optimal information access*
   b. End users: *A small, specialized group*
   c. User input: *At all stages of the database development*
   d. Typical operators: *Subject specialists, in both data entry & retrieval*
   e. Accessibility: *Limited to very restricted*

2. Database contents
   f. Scope and coverage: *Extensive coverage of limited subject*
   g. Indexing vocabulary: *Specialized & specific, Identifiers, not Descriptors*
   h. Indexing depth: *Very uneven*

3. Database structure
   i. Format and fields: *Rel. simple, based on the data collection*

4. Practical use of database
   j. Data entry: *On terminal, with or without worksheets*
   k. Search strategies: *Simple, rigid, high expectations*
   l. Operator language: *Simple, spoken language*

D. Practical Steps in Database Development

1. Initial analyses and decisions
2. Choice of system
3. Database definition
4. Indexing policies
5. Decisions on database administration
6. Conversion of existing files
7. Testing and adjustment of program
8. Loading the file
9. Training and education of users
10. User feedback and file maintenance

WORKSHEET FOR THE EXMIN DATABASE

(APPENDIX)

Location code:

LOC **1** (Fixed field, 15 characters)

First author:

AUT1 **2** (25 char.)

Second author:

AUT2 **2** (")

Third author:

AUT3 **2** (")

Fourth author:

AUT4 **2** (")

Year:

YR **3** (Fixed field, 4 characters)

Title:

TI **4** (60 char.)

Source:

SO **5** (30 char.)

6 - **7** are special fields for exploration information

1, 3, 4, 5 & 9 are required fields.

All fields except 4 & 5 are searchable.

Fields 1 through 5 are combined in a general record
Fields 1 & 6 form an exploration record
Fields 1 & 8 form a state record
Fields 1 & 9 form a keyword record

State:

ST **8** **8** **8** **8** **8** **8** **8** **8**

A complete document record has a maximum of 454 characters.

(fixed field of 2 characters,
repeatable 8 times, if needed)

Keywords:

KW **9** **9** **9** **9** **9** **9**

(Fixed field of 30 characters, repeatable 6 times, if needed)

ONLINE SEARCHING USING GEOGRAPHIC COORDINATES

Ralph K. Farrar
Joanne V. Lerud

Marathon Oil Company
Denver Research Center
P.O. Box 269, Littleton, CO 80160

Abstract-- There are many problems associated with using geographic terms either in natural language or from a thesaurus to index and retrieve documents. Some areas may not have a name. A geologic province or district may not be easily defined with geographic terms. The name of an area may not be readily known. Some areas may have several names or names with variant spellings. One possible solution to these problems may be to use geographic coordinates to specify an area. Many coordinate systems are in existence, as are many forms of expression for coordinates. The best known system is latitude and longitude expressed in degrees, minutes, and seconds. Two online databases containing bibliographic information, GeoRef, created by the American Geological Institute, and CRIB, created by the United States Geological Survey, use geographic coordinates. To retrieve information from these databases using geographic coordinates, database vendor software mainly employs ranging techniques on an inverted file containing the geographic coordinates. Retrieval results from searches using geographic coordinates versus index terms illustrate the relative effectiveness of using geographic coordinates for searching.

Problems Using Geographic Terms

There are many problems associated with using geographic terms to index and retrieve documents. They are similar to problems of assigning any subject term for indexing and retrieval. The principal concern is to find and use terms that properly express what the document or search request is about. This is not always an easy task when dealing with a location or geographic area.

Areas Not Named with a Geographic Term

Some areas may not have a name. This is most common in large areas without prominent landmarks, such as the deep water portions of oceans or broad expanses of desert. In some cases an area of interest may overlap several adjacent areas. An already named example of such a location would be the Four Corners area located at the boundary intersections of Colorado, Utah, Arizona, and New Mexico. For the most part, though, such overlaps may not have a name.

Areas of the oceans. The Atlantic Ocean covers about a fifth of the world's surface, yet there are not a great many terms to describe this area. The "GeoRef Thesaurus and Guide to Indexing" (Riley, 1981, p. 24) lists 68 narrower terms for the Atlantic Ocean. Geosystems' "Geosaurus" (Geosystems, 1981, p. 139) lists 43 terms. Petroleum Abstracts' "Geographic Thesaurus" (Information Services Division, 1978, p. 220-221) lists 355 terms under Atlantic Ocean in its hierarchical listing. This is about 2% of the approximately 14,500 terms in the "Geographic Thesaurus." A similar lack of vocabulary for describing locations in other oceanic areas is also common, even though much exploration of the deep oceans has occurred in the last 20 years.

Areas that overlap. Geologic trends or areas of interest more often than not cross political or jurisdictional boundaries. Such areas or trends may not have a name and may never have a name. They are usually difficult to index and it is even more difficult to formulate a logical search statement for them that will return highly relevant results. An area of this class could be a large mineralized zone extending over many political jurisdictions, overlapping a couple of geologic provinces, and containing many well recognized mining districts.

Name of Area Not Well Known

Usage varies. Many areas may have names but the usage of these names may not be well known or fixed. As evidence of this are the many "use for" or "UF" designations in thesauri. In the "GeoRef Thesaurus and Guide to Indexing" there are many examples of the sort as "Carlsberg Ridge UF Arabian Ridge [and] Arabian-Indian Midoceanic Ridge [and] Arabian-Indian Ridge." Although an attempt at vocabulary control has been made, instructions are given to also search the terms Arabian-Indian Midoceanic Ridge and Arabian-Indian Ridge along with the term Carlsberg Ridge.

Usage not fixed. In a recent update on the progress of activities to create a geologic and geographic thesaurus for the Committee for Coordination of Joint Prospecting for Mineral Resources in Asian Offshore Areas, it was reported in the "CCOP Newsletter" (1981, p. 11) that, "The members of the WGPD [Working Group on Petroleum Data] fully understand that in many countries the names for geographic areas [provinces, basins, fields, pools, and formations] have not been designated fully yet, so it was agreed to compile information on what was available." This report well illustrates the problem of selecting geographic terms for indexing and retrieval for areas still not fully known.

Variant spellings. Another instance where "use for" designations are common is for variations in spelling. Again from the "GeoRef Thesaurus" there are many examples such as "Majorca UF Mallorca," "Ankara UF Angora," or "Saint-Severin UF St. Severin." In the last example instructions are also given to search the term St. Severin along with the preferred term Saint-Severin. This problem can be compounded when romanization schemes are not well established or changed. The change to Pinyin from the Wade-Giles system of romanization for Chinese is in a class all by itself.

Political name changes. Many areas have changed their names over the years for political reasons. Thus the "GeoRef Thesaurus" lists "Madagascar- A valid term through 1974. After 1974, use Malagasy Republic." Another example is "Namibia- Term introduced in 1981 to replace South-west Africa." Sometimes a new name may not be used, as "Kampuchea use Cambodia" or "Kymer Republic use Cambodia." Sometimes, though, the name remains the same but the area will change, for example, again from the "GeoRef Thesaurus," "Congo- This term is now limited to the present Congo (People's Republic of the Congo). Formerly it also included what is now Zaire."

Geographic Coordinates

One possible solution to the above problems of using geographic terms is to use geographic coordinates to describe an area. Many different coordinate or grid systems for describing a location are in existence today. The Universal Transverse Mercator system is well recognized throughout the world. Individual countries may have their own national grids and states or provinces within countries may have their own coordinate systems, as is the case in the United States. Indeed, a location or area may be represented by several different coordinate systems. All of these coordinate or grid systems, though, are usually designed with some special purpose in mind. The most common system, the one that all the other systems relate to and the one that everyone thinks of when geographic coordinates are mentioned, is latitude and longitude.

Latitude and longitude can be expressed in many different ways. Generally they take the form of degrees, minutes, and seconds based on 360 degrees to a full circle. A few countries for special purposes, though, use a centisimal system, but this is very much an exception to the rule (Maling, 1973, p. 31). For latitude a direction from the equator is given being either north or south. For longitude a direction, being east or west, is given from the prime meridian, the Greenwich meridian, which was agreed upon internationally in 1884. Geographic coordinates in this form are the most well known and recognized as the best way of designating locations on the earth especially for geography.

Database Use of Geographic Coordinates

Two databases that contain bibliographic information and use geographic coordinates are GeoRef and CRIB. GeoRef, created by the American Geological Institute, is a database that currently contains citations from five major geological reference sources and many specialized bibliographies. The five major sources are (1) "Bibliography and Index of North American Geology" (1961-1970), (2) "Bibliography of Theses in Geology" (1965-1966), (3) "Geophysical Abstracts" (1966-1971), (4) "Bibliography and Index of Geology Exclusive of North America" (1967-1968), and (5) "Bibliography and Index of Geology" (1969 to present). CRIB, Computerized Resources Information Bank, is created by the United States Geological Survey. It contains information on the metallic and nonmetallic resources of the United States and other countries, but coverage by commodities and geography is not uniform (Calkins and others, 1978, p. A2). Access to GeoRef is provided by the ORBIT system of the System Development Corporation and the DIALOG system of DIALOG Information Services, Inc. CRIB is accessed through General Electric Information Services Company with software provided by Information Systems Programs of the University of Oklahoma.

Assignment of Geographic Coordinates

GeoRef. Since September 1977, GeoRef has been assigning geographic coordinates to its references. Not all references, though, receive coordinates. For a document to be assigned coordinates it must meet the following requirements.

"a. The paper must deal with geography in a way relevant to the geology of the area. Studies on geologic research in West Germany do not receive coordinates, whereas studies on the geology of West Germany do.

"b. The paper must deal with geographic area of reasonable size, usually a political subdivision (country) or smaller. The United States or Europe do not receive coordinates but Southwestern Kansas does. Parts of seas and oceans receive coordinates. Usually one set of coordinates is supplied but sometimes two (when the study compares samples from two widely separate areas). Studies on samples from scattered regions (such as geochemical studies) may not get coordinates at all, since it is felt that the geography is incidental and not relevant to the results in most cases." (Rassam and Tahirkheli, 1981, p. 84). Of the approximately 243,200 references added to GeoRef from September 1977 to April 1982, 139,023 or 57% have been assigned coordinates.

CRIB. Like GeoRef not all items in CRIB are assigned geographic coordinates. Unlike GeoRef, CRIB may use three other coordinate systems for describing location besides geographic coordinates. The other systems, besides latitude and longitude, are Universal Transverse Mercator coordinates, State Grid coordinates, and the U.S. Public Land Survey System of township, range, and section. Keefer and Calkins (1978, p. B9) state that "Usually just one of these methods is used by a reporter, but in some records, two methods (such as latitude and longitude and also township, range and section) may be used." Of the 52,456 items in CRIB, 37,847 or 72% have been assigned a latitude or longitude. Not all of the items in CRIB have references; of the 37,847 records which have geographic coordinates 30,011 or 79% also have references. Some records have many more than one reference and often cite the older literature.

Description of Areas with Geographic Coordinates

GeoRef. An area may be described in one of two ways, as a rectangle or a point. Latitudes are always given first in the field followed by longitudes. If the location is to be described as a rectangle, the southernmost latitude is given first followed by the northernmost latitude, then the easternmost longitude is given followed by the westernmost longitude. For a point location the latitude and longitude are just repeated, for example, N100000 N100000 W1350000 W1350000. If more than one area is being described for a record, another series of latitudes and longitudes is given. In GeoRef the direction comes before the degrees, minutes, and seconds of latitude or longitude.

CRIB. Locations are described only as points. The latitude is set forth in its own field as is the longitude. The direction designation follows the degrees, minutes, and seconds.

Retrieval Use of Geographic Coordinates

GeoRef and CRIB are databases, which are machine-readable files of information. To retrieve information from these databases, some other system of programs is needed. The American Geological Institute has selected two companies, System Development Corporation with their ORBIT system and DIALOG Information Services, Inc. with their DIALOG system, to provide this service for the GeoRef database. The U.S.G.S. has selected the Information Systems Programs of the University of Oklahoma to provide this service for the CRIB database. The retrieval function of these different systems is similar, but the procedures by which they obtain results may be very different.

GeoRef

Dialog. Two separate indexes or inverted files have been made specifically for all the geographic coordinates in the GeoRef file. One index is for latitudes, the other for longitudes. The latitude index is organized in an ascending sort. First all the north latitudes are listed from smallest to largest and then all the south latitudes are listed in a similar manner. The longitude index is arranged in the same way, east followed by west.

There are four different search strategies that can be applied to the latitude and longitude indexes (DIALOG Information Services, 1981, p. 22-27). First, the use of any single latitude or longitude can be searched by doing a SELECT LT= or LN= and stating the latitude or longitude in degrees, minutes and seconds after the =. Second, the use of any two latitudes or longitudes in a record can be searched by doing a SELECT LT= or LN= and using the (F) operator between the two latitudes or longitudes. The (F) operator requires the latitudes or longitudes to be within the same field, but in any order. Third, truncation can be used, for example, SELECT LN=102? or SELECT LT=27?. Fourth, a range of latitudes or longitudes can be searched by doing a SELECT LT= or LN= and using a colon (:) between the lower and upper bounds of the range. The lower bound should always be stated first. If the search area crosses a hemisphere, either east and west or north and south, then the ranging strategy must be used for the coordinates in one hemisphere and then the other. Once retrievals have been made from the latitude index and longitude index, they are anded together with a COMBINE statement.

ORBIT. A completely different tack has been taken by the ORBIT system to provide for searching by geographic coordinates. The coordinates in a record are matched against a table of pre-set ranges, which are made searchable by listing them in the basic index. The pre-set ranges are in increments of five degrees, for example, N00 to N04 or E000 to E004. To find all records indexed with the same latitude or longitude, the term standing for the pre-set range, which includes the desired latitude or longitude, is entered first. If, for instance, you wanted to search for N030000, you would enter FIND N0-4. Then a full text search must be done on the coordinates field of the records that were retrieved by the search on the pre-set range. So to finish the search for N030000, you would enter SENSEARCH CORD:N030000: (System Development Corporation, 1979, p. 5). The above method may be used repetitively to search for a range of coordinates, but it is cumbersome.

CRIB. The system of programs, originated by Information Systems Programs of the University of Oklahoma to perform retrieval functions, is GIPSY (General Information Processing System). Besides retrieval, GIPSY has many capabilities, such as mathematical manipulation of numerical data elements, that are not usually found with other systems. Like the DIALOG and ORBIT systems, the information from only certain fields is placed in an inverted file or indexed. The latitude and longitude fields do not fall in that category.

Several different search strategies are available to search the latitude and longitude fields (Information Systems Programs, 1981, p. 103-104). First, existence of any specific coordinate can be checked by entering LAT or LONG and the specific coordinate in between less than and greater than symbols, which are used as delimiters, for example LAT <21-30-15N> or LONG <113-45-20W>. Second, left and right truncation is allowed but it must be used with caution. Since the direction for a coordinate is placed at the end of a field, a set of items must be created first by searching with truncation on the direction, then the numerical part of the coordinate field of the retrieved items can be searched again with another truncation strategy. Third, a ranging strategy is available. This is done with the THRU operator, for example, LAT <03-30-00N> THRU <10-00-00N>. To avoid problems with this strategy a truncation search should first be performed on the direction also.

Retrieval Results

False Drops

From our experience all the retrieval systems operated as they were designed to operate, when doing searches using geographic coordinates. They retrieved according to the search statements that were entered. Unfortunately, when doing area searches using ranging techniques on the DIALOG and ORBIT systems as described in their searching manuals, up to 80% of the items retrieved were not within the areas being searched. This occurs because areas directly bordering the search area to the east, west, north, and south are also retrieved. If the left of a search area or rectangle is longitude E095, all items in the database that have that longitude will be retrieved, but false drops will occur because E095 will be common to the left edge of the search area and the right edge of the false drop. Another reason for this occurrence is the prevalent use by both indexers and searchers to describe areas with whole degrees, thereby increasing the chances for common edges.

This problem does not occur when searching CRIB with GIPSY, because the geographic coordinates in CRIB are always entered as points, unlike GeoRef where most locations are described as rectangles. If areas were described as rectangles in CRIB, GIPSY would function like DIALOG and ORBIT, because the overall retrieval logic is the same in all of them.

Coordinate Versus Terms

GeoRef. If only one kind of search is going to be done, it is usually more effective to search with geographic terms than with coordinates due to the above search logic problems and since only a small percentage, 19%, of the total database is indexed with geographic coordinates.

Coordinate searching, though, should not be ignored. In doing searches on areas that are not named or where the name of an area is not well known, we have found coordinate searching to be valuable. This was especially true for an area in the North Atlantic for which the only suitable term from the "GeoRef Thesaurus" appeared to be Northwest Atlantic. Too many items to be of use, 645, were retrieved when that geographic term was used. The geographic area in question was then searched using coordinates, and 22 items were retrieved of which several were useful. Another example where coordinates were useful was searching for items on the Khorat Plateau, Thailand. The "GeoRef Thesaurus" term for this area is Khorat Plateau. Searching with this geographic term retrieved 16 items. A coordinate search of this area was then performed which retrieved 43 items, 7 of which were useful and had not been found by using the geographic term.

CRIB. Since 72% of the items in CRIB have a latitude and longitude associated with them and false drops are not a problem using coordinate searching, searching with geographic coordinates is usually more effective than searching with geographic terms when a geographic area is being searched. If a specific mine is being searched for and the name of the mine is well known, it would be more appropriate to search using the mine name.

References Cited

Calkins, J. A., and others, 1978, Description of CRIB, the GIPSY retrieval mechanism, and the interface to the General Electric Mark III Service:- U.S. Geological Survey Circular 755-A, 49 p.

CCOP Newsletter, 1981, CCOP/ASCOPE co-operation: v. 8, no. 3, p. 8-11.

DIALOG Information Services, Inc., 1981, GeoRef; DIALOG information retrieval service, 89: Palo Alto, California, 41 p.

Geosystems, 1981, Geosaurus; Geosystems thesaurus of geoscience (4th ed.): London, 212 p.

Information Systems Programs, University of Tulsa, 1978, Geographic thesaurus (3rd ed.): Tulsa, Oklahoma, 245 p.

Information Systems Programs, University of Oklahoma, 1981, The minerals data system of the U.S. Geological Survey: Norman, Oklahoma, 207 p.

Keefer, E. K., and Calkins, J. A., 1978, Description of data items and codes in CRIB: U.S. Geological Survey Circular 755-B, 32 p.

Maling, D. H., 1973, Coordinate systems and map projections: London, George Philip and Son, 255 p.

Rassam, G. N., and Tahirkheli, S., 1981, GeoRef online workshop training manual: Falls Church, Virginia, American Geological Institute, 151 p.

Riley, S. J., ed., 1981, GEOREF thesaurus and guide to indexing (3rd ed.): Falls church, Virginia, American Geological Institute, 431 p.

System Development Corporation, 1979, GeoRef user manual: Santa Monica, California, 42 p.

ECONOMIC GEOLOGY ON GEOARCHIVE

Anthea Gotto

Geosystems
PO Box 573, Cambridge, MA 02139

Abstract--Over the last few years economic geology has become an increasingly important component of GeoArchive, Geosystems' international, bibliographic, geoscience database. At the end of 1981, commercially-related geoscience subjects constituted approximately 33% of the total database available through Dialog Information Services. The amount of economically-relevant information will increase significantly when data from over 100,000 maps are integrated into the publicly-available file.

In response to the needs of users, the economic geology content of GeoArchive is being increased, with no detriment to the coverage of the core geoscience subjects. In addition, more data, for example, mineral reserves figures are being incorporated as extensions to the reference titles, thereby reducing the necessity of obtaining the source documents. It is envisaged that GeoArchive becomes less of a strictly bibliographic database and more of a value-added file, which complements source databases containing numeric, textual and graphic data.

GeoArchive is an international, bibliographic database which covers all aspects of the geosciences -- from academic, core geology subjects to more applied areas such as mining and petroleum production. Table 1 is a summary of the main subject fields covered. More background information about GeoArchive can be found in Lea (1972, 1975, 1978) and in Walker (1978).

The sources, which are all publicly-available, are summarised briefly by form in Table 2. Further details are provided by Shearer (1979) and Charles (1981).

Since the database was started in 1969, there has been an ever-increasing demand from existing and prospective users, for coverage of economically-oriented geoscience literature. Consequently, Geosystems have been expanding steadily, the coverage of industry-related publications, and at the end of 1981, at least 33% of the references on GeoArchive were

Table 1.--GeoArchive subject scope

Regional geology	Sedimentology
Mineral exploration	Paleontology
Mining & petroleum production	Geochemistry
Mining & petroleum management	Geophysics
Environmental geology	Geomathematics
Engineering geology	Tectonics
Oceanology	Structural geology
Energy sources	Geomorphology
Mineral deposits	Geochronology
Hydrology & hydrogeology	Stratigraphy
Mineralogy	Methodology & equipment
Petrology	Geoscience information

Table 2.--GeoArchive sources

Journals	Maps
Newspapers	Field trip guides
Magazines	Conference proceedings
Reports	Conference abstracts
Press releases	Conference reports
Government publications	Collected papers
Theses	Dictionaries
Books	Directories
Book reviews	Audio-visual materials

concerned primarily with exploration and production from various geological, technical and management aspects.

At the same time, there has been refinement of the classification, and hence indexing, of those subjects related to oil, gas, and mining operations. Such developments are reflected by the Geosystems' thesaurus, Geosaurus (Charles, 1981). This, apart from being a thesaurus, is a classification where the subject areas, including stratigraphic and geographic terms, are hierarchically arranged with a coding system which facilitates the broadening and narrowing of searches. Figure 1 is an extract from Geosaurus that illustrates the main features, including the introduction dates of new terms.

346000 Mining engineering

See also: 332700 Coal exploration; 333000 Mineral exploration; 347000 Mineral processing; 358500 Mining hazards; 410000 Coal; 420000 Metallic deposits; 450000 Non-metallic deposits; 858000 Mining equipment; 860000 Mineral processing equipment; 965000 Mining organisations; 978000 Materials policy

346050 Mine evaluation & planning

Includes: Mine feasibility studies

See also: 250000 Resource maps & surveys; 352000 Reserve & resource estimation; 353000 Production managenment

346070 Mine surveying

Introduced: August 1981

See also: 255000 Mine & quarry plans & surveys; 858110 Surveying equipment

346080 Mine layout design

Introduced: August 1981

Includes: Openpit design

See also: 255000 Mine & quarry plans & surveys

346090 Mine sampling

Introduced: August 1981

Includes: Bulk sampling; Channel sampling; Chip sampling; Scout sampling

See also: 338200 Formation testing & sampling; 352000 Reserve & resource estimation; 691400 Regionalised variables (geostatistics); 851500 Field equipment

Figure 1. Extract from Geosaurus

The economic geology content of GeoArchive is soon to be greatly enhanced by the addition of information on over 100,000 maps, which are from the collections of the Institute of Geological Sciences (IGS) in the United Kingdom. These maps have been indexed in detail using the Geosaurus classification codes, together with title, edition, source, scale, latitude and longitude, form (including type of material, printing process, colour), publication number (for series which are numbered using two sequences), IGS library accession and location numbers, and finally, the accession number of related material, such as the text describing the map.

There has also been a specific effort to provide more data from an article, by using free-form title extensions. These are particularly useful for short news items that contain only a few important facts. Figure 2 shows an example where the GeoArchive editor has extended the title by adding more detail in parentheses.

Teutonic Bore Mine opens (Western Australia; copper-zinc-silver reserves 2,500,000 tonnes)

Mining Journal (London) 297/7611 P10 1981

Figure 2. GeoArchive citation with a title extension

This value-added approach has been taken further with the mining and minerals database, MinSys, which attempts to fulfil the user's need for easily-retrievable, condensed information. MinSys is not a full-text database. Many full-text databases have no controlled indexing, so that only natural language searching is possible, which means that important (relevant) concepts often may not be retrieved. Also, of course, a great volume of material may be retrieved and it all has to be read through for relevance. The information that goes into MinSys is indexed using controlled terms, and the retrieved product is a full extract of the source item. The extract can contain text, data such as production figures or operating costs, and graphics. (There is also the capability of collecting and manipulating data from different sources in the database and then specifying the type of graphic display output.) Figure 3 illustrates some of the types of data sets obtainable from MinSys.

In order to handle all these types of information, it was necessary to develop a new type of indexing system, which is called Omnisaurus. This is explained more fully in Gotto (in press) and is summarised in Table 3. It should be pointed out that material in GeoArchive is indexed at the first level of detail, with the hierarchically arranged descriptors of Geosaurus being used for indexing concepts. Source material for MinSys is analysed at any or all of the five levels.

Table 3.--Omnisaurus indexing system

LEVEL	OMNISAURUS	APPLICATION	FILE STRUCTURE
1	Geosaurus	Concepts	Hierarchical
2	Minsaurus	Names	Index sequential
3	Datasaurus	Data	Relational
4	Dictosaurus	Text	Inverted
5	Graphosaurus	Diagrams	Matrix

RESOURCE: Cobalt
RESOURCE: Copper
RESOURCE: Nickel
RESOURCE: Precious metals
RESOURCE: Palladium
RESOURCE: Platinum group metals
RESOURCE: Platinum
LOCATION: Western Platinum Mine, Rustenburg District, Transvaal
ORGANISATION: Western Platinum Ltd
MILL PRODUCTION: 1 October 1980 to 30 September 1981. Western Platinum Mine, mill feed 1,316,000 t
OPERATING COSTS: 1 October 1980 to 30 September 1981. Western Platinum Mine, R18.13 /t
SALES: 1 October 1980 to 30 September 1981. Net after commission, R50,385,000
OPERATING (WORKING) PROFIT: 1 October 1980 to 30 September 1981. R22,255,000
CAPITAL COSTS: 1 October 1980 to 30 September 1981. Net, R18,402,000
UNDERGOUND SAMPLING RESULTS: 1 October 1980 to 30 September 1981. Western Platinum Mine, sampled 13,757 m, channel width 95 cm, Platinum group metals, 4.86 g/t, 462 cmg/t
MINE DEVELOPMENT: 1 October 1980 to 30 September 1981. Good progress has been made with the expansion programme announced last year incorporating the mining, milling and treatment of ore from the UG 2 reef. Test milling should commence on schedule in March 1982
AVERAGE MILL RECOVERY: 1 October 1979 to 30 September 1980. Western Platinum Mine, Cobalt 16 t, Copper 1057 t, Nickel 1706 t
AVERAGE MILL RECOVERY: 1 October 1979 to 30 September 1980. Western Platinum Mine, Precious metals 706 kg, Palladium 1198 kg, Platinum 2774 kg
SOURCE: 20 October 1981. Press Release, Western Platinum Ltd
SOURCE TITLE: 1 October 1980 to 30 September 1981. Western Platinum Ltd, Preliminary profit statement for the quarter ending 30 September 1981.

Figure 3. Example of MinSys output

Thus, by using the Omnisaurus framework, all the information is captured and easily retrieved with a high degree of relevance. Market research and user-feedback have confirmed that the MinSys type of output is extremely useful for those people who require economically-related geological information. Already, more descriptors and data, in the MinSys style, are being added to the economic material being indexed for GeoArchive, so that it is becoming more of a 'value-added' bibliographic database. Certainly, with the user's increasing impatience when seeking information, more databases are being developed in a similar way, where the distinction between bibliographic or reference and non-bibliographic or source databases is no longer clear, and instead, becomes a continuum.

References cited

Charles, R., 1981, Geosaurus: Geosystems' thesaurus of geoscience: Geosystems, London, 212 p.

Gotto, A., (in press) Omnisaurus: a multilevel indexing system for natural resources data: Proceedings of the Geoscience Information Society, 13.

Lea, G., 1972, GeoArchive: an information retrieval system for geoscience: Proceedings of 24th International Geological Congress, Montreal, Section 16, p. 204-211.

_____1975, GeoArchive: Geosystems' information network: Geoscience Documentation 7, p. 61-63.

_____1978, GeoArchive: Geosystems' indexing policy: Proceedings of the Geoscience Information Society, 8, p. 42-56.

Shearer, J., 1979, Geosources: Geosystems' list of geoscience serial titles and non-serial publishers: Geosystems, London, 103 p.

Walker, R. D., 1978, GeoArchive: a brief review: Online 2, no. 3, p. 40-43.

IMPACT OF INTERNATIONAL DATA EXCHANGE ON GEOREF

J. G. Mulvihill

American Geological Institute
One Skyline Place - 5205 Leesburg Pike, Falls Church, VA 22041

Abstract--This year marks the beginning of an exchange of bibliographic references in machine-readable form between the GeoRef and Pascal-Geode data bases. To initiate this exchange many differences between the data bases had to be reconciled. In general, features unique to one data base were extended to both, and conflicting practices were resolved by discussion and compromise.

In preparation for the exchange, programs were written to convert records in the internal format of each data base to the UNISIST Reference Manual exchange format. The serial lists of the two data bases were merged and redivided. Major differences in indexing were identified and resolved. Lesser differences were and will continue to be looked at as practice brings them to light. An English-French translation file was developed for all standard index terms. Conventions were agreed to for capitalization, place names, transliteration, illustrations, coordinates, etc. These steps are described and related to their impact on GeoRef and the Bibliography and Index of Geology.

Introduction

On July 15, 1980 AGI issued the following news release:

JOINT GEOLOGICAL DATA BASE

The American Geological Institute (AGI) and the Bureau de Recherches Geologiques et Minieres (BRGM), and the Centre National de la Recherche Scientifique (CNRS) have signed an Agreement to begin production of a joint geological data base in 1981. This cooperative venture reflects a major effort to improve worldwide coverage and bibliographic control of the geologic literature. The existing data bases, AGI's GeoRef, BRGM's GEODE, and the geology sections of CNRS's PASCAL, will continue, but each will contain all references in the joint data base. French and English will be the base languages.

Cooperation will begin slowly in 1981 and accelerate as the year progresses. During 1981, documents with a 1981 imprint will be subject to the Agreement. Pre-1981 documents will continue to be covered by both data bases. This procedure has been adopted to provide a clear break between the duplication of the past and the shared input under the Agreement.

The printed bibliographies now produced from the individual data bases will be produced from the joint data base. AGI will continue to publish and sell the Bibliography and Index of Geology, and BRGM-CNRS will continue to publish and sell the Bulletin Signaletique-Bibliographie des Sciences de la Terre.

Coverage will include all materials previously indexed by the individual data bases. To eliminate duplication, but assure full coverage, the parties have produced a master list of serials and have worked out the handling of books, reports, theses, and maps. AGI will provide references for North American publications and BRGM-CNRS for European publications. Other areas will be divided between the parties to assure worldwide coverage. West Germany, Spain, Romania, Poland, Czechoslovakia, Hungary, and Finland will furnish references from their countries. The Agreement also provides for other countries to participate.

* * * * * * * * *

. . . to improve worldwide coverage . . .
. . . to eliminate duplication . . .

These are worthwhile goals, but we did not fully appreciate on July 15, 1980 the amount of preparatory work needed to achieve them. Our intent was to begin the exchange in January 1981. Already, by July 1980 much work had been done. An Agreement had been negotiated and signed and a multilingual geology thesaurus was nearing completion. This thesaurus, a project sponsored jointly by IUGS and ICSU/AB had been years in the making. Two of the people who had been involved in it, Jacques Gravesteijn of BRGM and Ghassan Rassam of AGI were already engaged in discussions on indexing for the joint data base in July 1980. Their previous cooperation on the thesaurus formed a basis for these discussions. But the exchange was not to begin in 1981. Not until April 1982 were the first "live" tapes exchanged.

Before the exchange could begin, differences in indexing, subject coverage, data elements, document coverage, and tape format had to be reconciled. These have been worked out. The effect of these decisions on GeoRef is the subject of this paper. I have not attempted to detail the changes to PASCAL-GEODE which also occurred, but obviously there were some. It was a matter of give-and-take in which we weighed the merits of various options and attempted to choose the best.

Since the actual exchange began just recently, we expect to see additional effects on GeoRef, particularly changes in coverage. These will not become evident for months. We will be monitoring these carefully. Had the exchange begun on schedule we could cover these here, but as it is they will have to wait for another paper.

General Principles and Assumptions

We should exchange data in computer-readable form -- This was decided early on. We had tried and abandoned an exchange of indexer work sheets as not cost-effective. One reason the work sheets didn't work was that they had to be reindexed. They might have worked had there been fully compatible data elements, vocabulary, etc. Computer-readable exchange forced some of this compatibility on us. For it the AGI computer had to be able to put the data from BRGM-CNRS into the AGI format. This meant not only that each data element used in PASCAL-GEODE should fit into a data element in GeoRef. It also meant that the computer should be able to use the actual data as provided or be able to convert that data to the required form. To give one example, if PASCAL-GEODE recorded the language of the source document, then the AGI computer should be able to put that information somewhere in GeoRef. Moreover, there must be agreement on whether the language will be a symbol or spelled out, be in upper or lower case letters, be repeatable or not, etc. If the language comes as a symbol and AGI wishes to convert it to another symbol or to full form, then the computer needs a complete list of the symbols used. Etc., etc. Such are the details the computer forces upon one. This is good but requires some considerable work on specifications and programming. To minimize this effort, we decided to establish an exchange format. AGI might have any sort of internal format it wished but for the exchange, the data must be put into the exchange format. Similarly any data coming to AGI had to be in the exchange format. For the exchange format we decided upon an existing standard, the UNISIST-ICSU/AB Reference Manual for Machine-Readable Bibliographic Descriptions. The Reference Manual was intended to be used as an exchange format. Adoption of this exchange format gave our programmer detailed, ready-made specifications to which and from which to convert data. Even so, each data element had to be discussed and specifics filled in tailored to our needs. This application of the Reference Manual format was written out and dubbed the Atlas format. Extensive programming was required, but the common exchange format at least gave the programmers detailed specifications to work from.

We should minimize the changes to exchanged data -- One of the reasons the exchange of indexer work sheets failed was that so many changes were required to use the exchanged sheets it was actually more efficient to do the citations from scratch. Ideally AGI should not have to look at the references received from BRGM-CNRS before adding them to GeoRef. For this to be done there would have to be automatic checking in of references in serial and monograph files in the library, no possibility of duplication of references, confidence in the completeness and accuracy of the data, machine translation of index terms, titles and abstracts, and incorporation of all data into the GeoRef format by computer only, without human intervention.

Unassisted addition of records has not been achieved and probably never will be. However this principle guided us toward using the computer for all feasible tasks and attempting to eliminate differences between us. For example, if we could decide on a common set of category codes and apply them consistently, then the AGI indexer reviewing a reference from BRGM-CNRS would not need to assign a category code to the reference. Past practices and the decision by both parties not to change their printed bibliographies worked against a common set of categories, but a compromise was achieved (see Categories below).

We should establish a common indexing vocabulary -- This was a guiding principle from the beginning. By far, most of the weeks spent working out problems were taken up with indexing terminology. This was so important because if an exchanged document had to be completely reindexed, much of the savings of the exchange would be lost. Further, since AGI might not have the source document from which BRGM-CNRS had prepared the references, at the minimum there should be enough indexing provided in that reference for the AGI indexer to work from. So the indexing of areas, commodities, elements, rocks, minerals, fossils, etc. was discussed. For each of these key topics, indexing practice and terminology were compared and what each party expected from the other was established. A second phase involved a term-by-term comparison of the two vocabularies, complicated by the involvement of two languages. The questions what do you mean by such-and-such a term and how do you use this term were asked over and over.

We tried to keep changes to a minimum, and where a change was made, to make it in a way that would be compatible with our previous indexing. For example, we agreed to distinguish between aluminum as a commodity and as an element by term form while retaining our old index terms and adding a set of new ones (see Commodities and Elements below).

Effects on GeoRef

From the discussions preparatory to the exchange, we agreed to certain changes in GeoRef. These are specified below, grouped under the headings (1) general effects, (2) effects on indexing, and (3) effects on data elements.

1. General Effects

Base language -- English was the base language of GeoRef. As of the exchange GeoRef will have two base languages, English and French. Exchanged references will have titles, index terms, and abstracts (if present) in both English and French. BRGM-CNRS provides the French translations, AGI the English. Controlled index terms are translated by computer.

Diacriticals -- These were not used in GeoRef after 1971. Exchanged references will have them for French, German and some other languages.

Data elements -- GeoRef had used those in the UNISIST-ICSU/AB Reference Manual since October 1975. The data elements of the Atlas exchange format, an adaptation of the Reference Manual, are very similar to those in GeoRef. There are very few changes, mostly additions, e.g. the indication of the language of titles, abstracts and index terms, addition of special fields for translation source, and indication of controlled and uncontrolled terms.

Processing -- A translation step has been added to the processing of a reference from BRGM-CNRS. In this step, English translations of French titles, abstracts and uncontrolled index terms are added to the references from BRGM-CNRS by the GeoRef staff and these translations are also sent to BRGM-CNRS. Similarly French translations of English titles, abstracts and uncontrolled index terms in AGI references are prepared by BRGM-CNRS and sent to AGI to be added to those references.

Bibliography and Index of Geology -- It was decided that no change be made in the structure of the Subject Index. It will continue as it has been, composed of three level sets with cross-references. Changes in individual index terms will, of course, affect the Subject Index. These are described in the GeoRef Thesaurus (1981) and in the User Guide to the Bibliography and Index of Geology to be published in 1982. The changes enumerated in this paper relating to subject coverage and document coverage will of course be reflected in the Bibliography. The sets are not provided by BRGM-CNRS in the exchange references. The GeoRef indexers add these sets to the exchanged references, specifically for use in the Bibliography and in other indexes prepared from GeoRef.

GeoRef Thesaurus -- The third edition (1981) of the Thesaurus includes the changes introduced for the exchange. The more important of these are discussed below under indexing. With these changes and the normal changes and additions between editions, the number of terms in the Thesaurus increased from about 12,500 in the second edition to about 13,500 in the third edition.

Coverage -- AGI is responsible for North American publications. BRGM-CNRS is responsible for European publications. Publications from other areas are divided between them. Before the exchange, GeoRef had worldwide coverage with emphasis on North America. With the exchange the coverage of Europe should improve. This is because the European references are originating in Europe. At the same time the coverage of North America should be as good as before. To assure no serials fall between the cracks, we compared serials lists with BRGM-CNRS and retained serials on either list. For monographs, we are notifying BRGM/CNRS of monographs we see in their geographic area of coverage which we would have covered and they are to do likewise. There will be much better coverage of European reports, meeting abstracts and theses.

1981 publications -- References to 1981 and later publications are subject to the exchange. We are continuing to prepare references to pre-1981 publications, worldwide, as these come to our attention, but these are not exchanged.

2. Effects on Indexing

Commodities and elements -- In GeoRef the same terms were used for chemical elements and for commodities. For example, "aluminum" was used for both the element aluminum and for aluminum ores. We depended upon coordination with other terms online and set structure in the Bibliography to distinguish between these meanings. Now we have agreed to add a separate set of terms for the commodities, e.g. "aluminum ores", and to restrict the unmodified term to use as a chemical element.

Fossils -- We had used the same systematic fossil terms for paleontological papers as for stratigraphy papers. Now a small number of common names for fossils has been added, restricted to use in stratigraphy papers.

Minerals -- We had indexed specific mineral names as well as more generic names. This is not changed.

Seismology -- We have added a number of terms in seismology.

Areas -- We had indexed specific areas plus broader terms from our Geographical Terms List (List O). We are continuing this but we are also making sure to include a term from a certain minimum hierarchical level keyed to a set of index maps (see below). E.g. we would add the new term "Southern California" to a paper also indexed with the terms "San Diego County" and "California".

Stratigraphy -- We will continue to index the specific and general stratigraphic terms as in the past, but will add a stratigraphic term from a list of required minimum hierarchical level terms.

Fossils -- We will continue to index the specific and general fossil terms as in the past, but will add a fossil term from a list of required minimum hierarchical level terms.

Index maps -- We will use a set of index maps covering the world, from which at least one term for a minimum hierarchical level will be selected for each paper on a geographical area.

3. Effects on Data Elements

Affiliation -- GeoRef is adding more specific address information to the organization in this field.

Geographical coordinates -- These will continue to be included.

Categories -- We will continue to use the same 29 categories as in the past, but will include subdivisions for some of them, giving a total of 43.

Serial titles -- We will continue our gradual transition to ISSN-Key Title-Imprint.

Language abbreviations -- No change.

Map types -- We have adopted a standard list of map types, to facilitate translation from French to English.

Numbers of tables, plates and maps -- We had simply noted the presence of tables, plates or maps. For the exchange, we will add the number of these in each reference.

Abstracts -- We will be receiving abstracts in some of the BRGM-CNRS references. These we plan to translate and include in GeoRef and the Bibliography. We had already begun to add a few abstracts to GeoRef and will continue to do so.

Language of title, of abstract, and of index terms -- This is being added since the data base now has dual base languages.

Translation source -- A number of fields are being added for recording information on the source from which a translation was made. Previously in GeoRef this information was put in an unformatted note.

THE FUTURE COST OF ONLINE RETRIEVAL

Richard D. Walker

University of Wisconsin, Madison, Wisconsin. Library School
600 N. Park Street, Madison, Wisconsin 53706

Abstract--An accelerated rate of migration from subscriptions of print products (also known as paper or hard copy products) to online database products (also known as tape or computer-readable products) is predicted for the next few years in all fields. There is some evidence that new or expanding libraries tend to choose online services over subscriptions to their paper alternatives.

Expanded use of online searching of data bases of all types will continue but the impact on print products is not yet clear. In the geosciences there does not yet appear to be a significant decrease in subscriptions of the print products even with an ever increasing level of use of the online services.

In a recent issue of the Library Association Record, the new President of the Library Association, Ken Stockman, (1982, p.3) quoted Edmund Burke, the 18th century politician, who wrote, "you can never plan the future on the past," and also "to lament the past, to conceive extravagant hopes of the future, are the common dispositions of the greatest part of mankind."

Not to heed the wisdom of Burke would be folly. But, whatever may develop in the years ahead will somehow be an extension of the present and we can at least know the present; at the sametime we will not "lament the past" nor predict "extravagant hopes of the future," but be prepared as best we can to deal with the future as it presents itself as the extension of the present.

Ever since the question was just publicly aired at the 1974 meeting of ICSU-AB (International Council of Scientific Unions-Abstracting Board) in Berlin (Stern, 1974), the problem of the impact of increasing use of online searching of bibliographic databases on subscriptions of their paper equivalents has been discussed and studied. (In some cases "equivalent" may not be strictly precise, as the machine readable databases are not always exactly equivalent to the printed counterpart.) The most recent study of this impact was reported in Online Review by Lancaster and Goldhor (1981, p.301-311). They predict a great accelerated level of "migration" will occur in the next five to ten years. (Migration is defined as switching subscriptions from the paper equivalent to the use of online searching of its database.) Also out of the University of Illinois at Urbana-Champaign comes a recent thirteen year analysis of the finances of a major database producer, which also publishes abstracting and indexing (A&I) products. This analysis, reported by Williams (1981, 263-276), considers the data, although from only one database producer, to be representative of numerous A&I services. Her findings suggest there is "no alternative to increasing prices for online users."

The primary problem is that of maintaining a financial balance in view of increased income from the database revenue (online royalties), decreased income from their paper equivalent A&I services, and increased general operation costs. The two major geoscience database producers have reported recently in personal communiques that migration is not a major problem for their organizations (Mulvihill, 1982 & Lea, 1982). The American Geological Institute's subscriptions for the Bibliography and Index of Geology have not decreased significantly in recent years. From 1975 through 1981 there was only a 3.6 percent decrease even with substantial increases in subscription prices (due more to general increasing production costs than any other factor). Between 1975 and 1981 the subscription price of the Bibliography and Index of Geology increased from $650 p.a. to $825 p.a. Earlier figures are not complete but there was only a slight drop in the number of subscriptions when the price was increased from $150 p.a. (1969-71) to $250 p.a. (1972-74) and to $650 p.a. in 1975 (Mulvihill, 1982). Geosystems "have not experienced any drop off in Geotitles Weekly subscriptions." While there is a steady increase in the number of online users, "it seems to be that new users go online" (Lea, 1982).

The primary focus here is what the costs of information in general will be in the next five to ten years and specifically what the costs of information in the geosciences will be. I have no special expertise in predicting the future, but I do have some data and some knowledge about the geosciences that may give me a bit of an edge in making an intelligent guess about this important aspect of geological information.

Database construction is expensive. It is expensive if it is undertaken as production for an A&I service and it is expensive if it is a separate enterprise in its own right. Any consideration of future costs must include both the financial and the intellectual input needed for either production. Database producers have recovered the cost of production of printed publications, for the most part A&I services, from the revenues received from subscriptions to these print publications. New technological advances and increased use of online searching of these A&I services have caused producers to consider changes in the methods employed to recover these costs.

A study financed by EEC (European Common Market) in the form of a grant to ICSU-AB and conducted by Barwise of the London Business School reported that approximately 10% of the producer's creation costs were recovered from online searchers and other users of database services and 90% from printed A&I services (Barwise, 1979). Additionally, it was pointed out that 80% of the revenue from printed product subscriptions is available to the producer to recover the costs of production but only about 20% of the income from online searching is returned to the producer because 80% of the charges for such use are to cover telecommunication costs and computer operations offered by the vendors of the service.

From this ICSU-AB study was developed a modelling strategy to look into the future. Applied to a fictitious database producer using eight scenarios to project the several outcomes, conclusions were drawn. The main conclusion was that online services were subsidized by printed products and that this practice was less and less appropriate as the online services became more highly developed and a larger part of the producers overall service offerings. This strongly suggests a gradual increase in online charges in relation to print services by a factor of 2 or 3 times.

In September of 1978 it was announced in Library Journal that Psychological Abstracts would increase fees from $20 to $30 per connect hour for online searches beginning January 1979, so as to have the database start paying its appropriate share of the costs of production. The announcement stated, "it is unfair to have the print version... absorb so much of the costs...." ("Psych Abstracts," 1981, p.1681).

In the analysis by Williams referred to earlier, it is stated that the existence or non-existence of the so-called "migration phenomenon" is immaterial as it seeks to lay blame on the technology and the concomitant service it brings. Furthermore, Williams suggests competition between products is healthy and desirable, albeit, with many problems (1981, p.265).

Lancaster and Goldhor, in their study of 200 academic and special libraries to determine the extent to which these libraries had discontinued subscriptions to printed A&I services as a result of the accessibility of equivalent databases online, reported significant levels of cancellation but their data "suggest that online services have so far had rather little effect on subscriptions to printed services." (1981, p.319). Roger Summit of Dialog Information Services, Inc., in a letter in Online Review wrote that this study, "was conducted with an apparent bias toward showing a positive relationship between online use and subscription cancellations" and suggests questions that might have mitigated the bias, had they been asked (1981, p.496). He further suggests, to the extent that database suppliers conclude online usage and subscriptions to print services are competitive as opposed to complementary and increase royalties or forced tie-in arrangements with subscriptions, both of which can lead to decreased overall use, the results will be unfavorable to all concerned.

Summit's critique ends with:

> Short term increased revenue can come from growth in usage at the same price as well as from increased prices with no (or even declining) growth. Sustaining long term increases from revenue must come from growth.
>
> It is extremely important to all concerned--database producers, online services, and users--that the phenomenon of online usage be objectively and pragmatically understood, and the tendency toward oversimplification of cause and effect be resisted. With the growth rates reported, we should adjust our focus on the reality of the expanding marketplace. Such a

> focus should suggest strategies to maintain and increase growth, rather than pricing tactics which, although they may increase short term revenue, most surely stifle growth) (1981, p.496).

Summit's criticism notwithstanding, it is worthwhile to take note of the prediction by Lancaster and Goldhor concerning the transition from print to database use. They suggest a process through three phases, with the third phase being established libraries discontinuing printed products subscriptions in favor of online access within the next ten years. The process of conversion from print to online use will be accelerated by (1) simplification of search procedures, (2) increased availability of online services, and (3) "changing attitudes among librarians regarding the proportion of the budget to be allocated to subsidizing print on paper access versus the proportion to be allocated to subsidizing online access" (1981, p.310).

There is little reason to suspect the geosciences are different than other disciplines in this respect, and I would predict the trends developing generally have application in our field. The two established, widely available systems have responded to my request for information and I shall speak specifically to their situations.

The American Geological Institute's royalty charges have increased twice in recent years. From 1973 to June of 1981 royalties charged for use of GeoRef were $15 per connect hour and 10 cents per citation offline and no charge for online printing of citations. From July 1981 through February 1982 the charges were $16.50 per connect hour, 10 cents for an offline citation, and 5 cents for an online citation. Since March 1982 the connect charge has been $24 per hour and both offline and online printing of a citation are at 10 cents each. John Mulvihill of AGI has written: "We don't know how many migrations from B.I.G. to online have occurred.... But aside from migration, if it were not for online I feel fairly certain that there would have been an increase in B.I.G. subscriptions during that period [1975-1981], given the increased attention to geology which has occurred. Each year a number of subscribers drop, but others take their place. And some who would have subscribed decided on online instead" (Mulvihill 1982).

Revenue at AGI from online use of GeoRef has steadily risen from 3% of the total in 1975 to 14% in 1981.

From Graham Lea I learned: "Each month we seem to get more online users of GeoArchive.... We have not experienced any drop off in Geotitles Weekly subscriptions--it seems to be that new users go online."

Revenue at Geosystems from online use is increasing by 1% to 5% per year and represents between 5% and 20% of the total revenue.

In conclusion, all indications are that there will be increases in costs of both printed and online services in the geosciences and that as long as data producer's revenue keeps increasing, the ratio between online use and printed service subscriptions does not matter. There has been steady growth in revenue from GeoArchive and its print products since 1968 and AGI has also experienced healthy growth during the last decade.

Costs will rise, a result of inflation; while there will be lower computer costs. A different mix of online and print product revenue will develop, perhaps different for each database producer.

So, while we can not plan the future based on the past, we can at least observe the present trends in this specific area. And, while I have not predicted "extravagant hopes of the future" I think we can look forward to continued affordable bibliographic services in the years immediately ahead.

References Cited

Barlow, D. H., Kent, A. K., and Stern, B. T., 1979, Information costs in the future: International Forum on Information and Documentation, v. 4(1), p. 29-31.

Barwise, T. P., 1979, Online Searching: The Impact on user charges of the extended use of online information services: International Council of Scientific Unions, Abstracting Board, Paris.

______1979, The Cost of literature search 1985: Journal of Information Science, v. 1, p. 195-201.

Clayton, A., 1981, Factors affecting future online services: Online Review, v. 5, p. 287-300.

Dammers, H. F., 1975, The Economics of computer based information systems: A Review: Journal of Documentation, v. 31, p. 38-45. [For comments by W. A. Martin see Journal of Documentation, v. 31, p. 218.]

Lancaster, F. W. and Goldhor, H., 1981, The Impact of online services on subscriptions to printed publications: Online Review, v. 5, p. 301-311.

Lancaster, F. W., and Neway, J. M., 1982, The Future of Indexing and Abstracting Services: Journal of the American Society for Information Service, v. 33, p. 183-9.

Lea, Graham, 1982, Personal letter. 28 April.

Mulvihill, John, 1982, Personal letter. 23 April.

Pfaffenberger, A., and Echt, S., 1980, Substitution of SciSearch and Social SciSearch for their Print Versions in an Academic Library: Database, v. 3, p. 63-71.

"Psych Abstracts" Price boost stirs protest: 1981, Library Journal, v. 103, p. 1681-1682.

Stanley, W. G., 1979, Changing Revenue Patterns from Online Use. Paper presented at the National Information Conference and Exposition, NICE III, Washington, D.C., April, 1979.

Stern, B. T., Annual Board Meeting: ICSU-AB, Berlin.

Stockman, K., 1982, The Supreme opportunity for service: Library Association Record, v. 84(1), January, p. 3.

Summit, R. K., 1981, Online and Print: Online Review, v. 5, p. 496.

Timour, J. A., 1979, Use of Selected Abstracting and Indexing Journals in Biomedical Resource Libraries: Bulletin of the Medical Library Association, v. 67, p. 330-5.

Trubkin, L., 1980, Migration from print to online use: Online Review, v. 4, p. 5-12.

Williams, M. E., 1977, Online problems: Research today, solutions tomorrow: Bulletin of the American Society for Information Science, v. 3(4), p. 14-16.

______1981, Relative impact of print and database products on database producer expenses and income: Information Processing and Management, v. 17, p. 263-276.

______and Brandhorst, T., 1979, Future trends in A&I data-base publication: Bulletin of the American Society for Information Science, v. 5(3), p. 278.

THE REALITY OF ONLINE RETRIEVAL OF GEOSCIENCE MAPS

James A. Coombs

Map Collection
Southwest Missouri State University
Springfield, Missouri
65804

Introduction

The need for conferences such as this is evidence of the growing demand for geoscience information. In particular, the importance of geoscience information to environmental planning and legislation is "... contributing to increased demand for maps, information about maps, and improved bibliographic control of maps" (Schroeder, 1981,p.435).

Bibliographic control of maps has been a long-standing problem for book-oriented libraries, primarily because the arrangement of bibliographic information on maps is frequently inconsistant, and the written description of cataloging records frequently cannot adequately describe the graphic information displayed on maps. These characteristics are primary reasons why the map format has received inadequate and unequal levels of bibliographic control in both general and specialized library environments.

Recent Developments

This situation has been improving, however. In the past few years, several developments have helped librarians achieve a level of bibliographic control for maps equivalent to that provided for books. The first was the development and implementation of the Machine Readable Cataloging (MARC) map format at the Library of Congress Geography and Map Division, and the publication of David Carrington and Elizabeth Mangan's "Data Preparation Manual for the Conversion of Map Cataloging Records to Machine Readable Form" in 1971. The MARC format provides a standardized method of identifying the various types of bibliographic information in a cataloging record by numbered fields, and, through the use of indicators, provides information necessary to make the record easily manipulated by computers (Hudson, 1981, p.117). MARC has since become the internationally accepted standard supported by major government subsidized programs in the United States, Canada, Great Britian, France, and Germany (Parr, 1975, p.43, cited in Schroeder, 1981, p.423).

The next development was the implementation of a MARC format-compatible, online map cataloging workform by the Online Computer Library Center (OCLC). OCLC's relationship with maps began in 1975, when it asked the United States

Geologicl Survey (USGS) to recommend six professional map librarians to serve on a Task Force for Cataloging Maps. The Task Force met with OCLC staff several times during 1976 to assist the Center with: (1) the design of the workform for input cataloging of maps, (2) the fields that should be indexed for record retrieval, (3) the design of the format for printed catalog cards, and (4) to recommend a set of minimum standards for inputting maps. (See Cobb, 1977, p.30; Minton, 1978, p.27). The workform became available to member libraries in September of that year, and online map cataloging became reality. Since then, the Research Libraries Information Network (RLIN) and Washington Library Network (WLN) have also developed online map cataloging capability.

The third event was the creation of the International Standard Bibliographic Description for Cartographic Materials [ISBD(CM)] by the International Federation of Library Associations and Institutions Joint Working Group on ISBD(CM), and the publication of "International Standard Bibliographic Description for Cartographic Materials" by the Group in 1977. The idea of an ISBD for maps was first suggested by the IFLA Sub-section of Geography and Map Libraries in 1973, and the Joint Working Group was formed, consisting of representatives from the national map libraries of Hong Kong, Finland, France, the Netherlands, West Germany, the United States, and Canada. The Group compiled four drafts; the second of which was circulated to about 100 members of the map library community for review. (See Karrow, 1981, p.2-3). The final version of ISBD(CM) thus represents a milestone for international cooperation among map libraries, as well as for bibliographic control of maps.

The most recent event occurred in October 1979, when the Anglo-American Cataloguing Committee for Cartographic Materials (AACCCM), with representatives from the Library of Congress, the Public Archives of Canada, the British Library, the National Library of Australia and the map library associations of respective countries, met in Ottawa, Canada to discuss the shortcomings of the Anglo-American Cataloging Rules, Second Edition (AACR2) as it relates to cartographic materials. (See Farrell, 1979, p. 19-32; Stevens, 1980, p. 117-126). Although the concept, arrangement, and general principles of AACR2 are more logical than that of AACR1, apparently there was minimal peer review of the cartographic chapter while it was in the planning stage. (See Karrow, 1981, p.2-11). The AACCCM decided at this first meeting that one of its primary tasks would be the compilation of a map cataloging manual (Farrell, 1979, p.20). After two years of hard work, the Committee's "Cartographic Materials: a Manual of Interpretation for AACR2," will be published by the American Libraries Association (ALA) in August 1982.

The Pros and Cons of Online Map Cataloging

The advantages and benefits of online cooperative cataloging have been widely reported in library journals, as have the problems. Cobb (1977, p.30-33), Minton (1979, p.25-40), and Schroeder (1981, p.419-423) have described these as they apply to cartographic materials. Standardization of map records, shared cataloging, catalog cards tailored to individual libraries, and access to a large data base containing bibliographic records of cartographic materials in the Library of Congress, the USGS Library, and other member libraries are the major benefits for geoscience map collections. The major drawbacks are lack of retrospective coverage, including the Library of Congress (pre-1968) and USGS (pre-1979), slow terminal response time, and lack of quality control.

In general, the key to the overall success of any online cataloging system is that all records meet the standards set by AACR and the MARC formats, and that the cataloging and formatting of the records be free of errors (Hudson, 1981, p.116-117). Yet a major problem with the use of contributed cataloging has been the amount of revision needed to bring records up to these standards. Hudson's analysis (1981, p.118-119) of OCLC monographic cataloging records used by the State University of New York at Albany from November 1979 to July 1980 revealed that 58% (589 out of 1017) member-contributed records required revision. Of these, 40% contained formatting errors, 29% contained incomplete or incorrect fixed fields, and 22% had access point errors or omissions. The situation regarding duplicate records in OCLC is just as serious. In Johnson and Josel's sample (1981, p.44) of 3428 monographs, scores, sound recordings, and audiovisual material records, 94 titles (9.4%) had duplicate records. Two records were found for 84 titles, 3 records for 8 titles, 4 records for 1 title, and 5 records for 1 title, making a total of 201 records for 94 titles.

The studies, while being generally relevant, did not involve the map format. Therefore, a similar survey was conducted in May 1982 of 497 geoscience maps in the Southwest Missouri State University Map Collection. These included state geological survey maps, USGS thematic maps, and various geologic, hydrologic and physical maps of foreign countries. Of these, 283 (57%) were found on OCLC, with 155 member-contributed records and 128 MARC records. Of the member records, 119 (77%) needed revision, 54 (35%) had formatting errors, 22 (14%) had fixed field omissions, 17 (11%) had access point omissions or errors, and 104 (87%) were input prior to January 1981, and needed revision to comply with AACR2 standards. Concerning duplicate records, 104 records were found for 49 titles, with 2 records for 43 titles and 3 records for 6 titles. Library of Congress MARC records were involved with 43 titles, and these duplicates were created when the MARC tapes made from 1968 to 1979 were loaded into the Online Union Catalog in July 1979 (SLA, 1979, p.33). Of the 214 titles not found on OCLC, most were either pre-1968 (98 titles), locally produced (30 titles), or were received within the last month (35 titles).

Improving Database Quality

Changes to input requirements, changes in cataloging rules, differences in interpretation of AACR, differences in local cataloging needs, and human error all contribute to inaccurate and incomplete biliographic records. In the OCLC Cataloging Sub-system, users are responsible for the accuracy of data entered and for complying with AACR2, the Library of Congress's interpretation of AACR2 (for map catalogers, that would be the AACCCM manual), as well as with OCLC's standards

Users can, of course, correct and/or modify records before producing cards or updating holdings, but only OCLC or the inputting library can correct a Master Record. The current procedure for correcting other institutions' errors is to fill out a "Change Request" form according to the guidelines in Section 7, "Quality Assurance," of OCLC's "Online Systems Cataloging: User Manual." This time consuming and expensive process, as well as the slowness with which OCLC takes action on reported errors, may easily discourage members from filing reports. Also, for maps, duplicate records are no longer even reported (Johnson and Josel, 1981, p.46).

Dissatisfaction with the level of bibliographic errors in the OCLC database, as well as delay of enhancements such as subject search capability, has caused some member libraries to investigate other online cooperative cataloging systems. Comparisons of the currently available systems have been published in the library journals (see Matthews, 1979, p.665-738; Webster, 1980, p.1-32; University of Oregon, 1981, p.215-230), but as these systems occasionally offer new services or experience problems, the relevance of these comparisons diminishes. RLIN's recent improvements (see Library Journal, 1981, p.2067) and crisis (see Savage, 1982, p.1027-1030) are good examples.

OCLC is aware of user dissatisfaction and is attempting to improve quality control with its new ENHANCE program, to be installed in 1983. ENHANCE will have four levels of corrections ability and will provide some online error correction capability for member libraries. "Master Mode" will be restricted to OCLC use in deleting records and making type code corrections, as well as making additions, changes and corrections. The level one step below Master Mode is for the Library of Congress and other national libraries, who will be able to make any additions, changes and corrections; but not to delete records or make type code changes. The next lower level will allow a few OCLC-selected libraries to correct any errors found in member-contributed records. The lowest level will allow any member library to expand a record, but not to change or correct one. (See Johnson and Josel, 1981, p.46).

Map catalogers have made their own efforts to improve quality control. In June 1980, the Map Online Users Group (MOUG) was formed by a merger of the OCLC Map Users Group and map catalogers using the RLIN system. MOUG sponsors map workshops (the most recent was at OCLC April 15-17), and is active in lobbying OCLC to accelerate implementation of specific technical capabilities which would be more responsive to online map catalogers and other users, such as geographic coordinate and geographic area searching. Also, the MOUG Cooperative Cataloging Coordinating Committee (C^4) was created to initiate cooperative cataloging agreements (MOUG, June 1980, p.1). So far, the U.S. Geological Survey Library has volunteered to be responsible for cataloging all USGS maps except topographic quadrangles (MOUG, October 1980, p. 2-3).

The Special Libraries Association Geography and Map Division Cataloging Committee is also active in establishing cooperative cataloging in the map library community. In 1981, the Committee mailed a questionnaire to several members of the G & M Division in order to obtain information about each library's area of expertise or specialization. The Committee will have an open meeting at the SLA Annual Conference in June 1982 at which representatives of the cataloging committees of the Committee for Southeastern Map Libraries, ALA Map & Geography Round Table, MOUG, North American Cartographic Information Society, and Western Association of Map Libraries will discuss cooperative cataloging and means of achieving such cooperation (McGarry, written commun., 1982).

Information on the new developments in online map cataloging can be found in MOUG's Newsletter; Dorothy McGarry's regular feature "On the Cataloging/Cataloguing Front," in the ALA Map & Geography Round Table newsletter, "base line;" and Myrna Fleming's "Cataloging Cartographic Materials," in the Western Association of Map Libraries "Information Bulletin." The Library of Congress "Cataloging Service Bulletin" includes AACR rule interpretations and other general cataloging news.

The use of cataloging worksheets in conjunction with online map cataloging has also been recommended for map catalogers who do not have their own CRT and have to rely on staff members in technical services for searching and inputting records. (See Minton, 1978, p.28, 34-36; OCLC Map Users Group, February 1980, p.2; Coombs, 1980, p.5-6). Filling out worksheets is a good way of insuring that proper indicators, subfield codes, and bibliographic information is used and, therefore, is a useful means of contributing error-free records. Appendix B is a sample of the worksheet used for OCLC map cataloging at Southwest Misssouri State University.

Summary

The growing demand for maps, as well as the related increased production of maps, has placed pressure on geoscience and map librarians to provide better bibliographic control of maps in their collections. To date, access to online cooperative cataloging systems offers the best means of achieving better control. However, to assure the success of online data bases, they must be utilized with the utmost efficiencey. Improving the quality of cataloging input is the only way that libraries will be able to take full advantage of the data base services. While member libraries are only partially responsible for substandard records that exist in the data bases, only member libraries working cooperatively can raise the level of quality in bibliographic records. Many geoscience and map libraries are active toward this end. The presence of new faces at presentations such as this is evidence that these efforts are successful.

References Cited

Cobb, D.A., 1977, OCLC: its impact on map librarianship: SLA Geography and Map Division Bulletin no.108, p.30-33.

Coombs, J.A., 1980, Experiences with OCLC cataloging procedures: OCLC Map Users Group Newsletter no.3, p.4-6.

Farrell, Barbara, 1979, Minutes of the first meeting of the Anglo-American Cataloging Committee for Cartographic Materials, Ottawa, Canada, October 1-5, 1979: SLA Geography and Map Division Bulletin no.118, p.19-32.

Hudson, Judith, 1981, Revisions to contributed cataloging in a cooperative cataloging data base: Journal of Library Automation, v.14, no.2, p.116-120.

Johnson, J.J., and Josel, C.S., 1981, Quality control and the OCLC data base: a report on error reporting: Library Resources & Technical Services, v.25, no. 1, p.40-47.

Karrow, R.W., Jr., 1981, Innocent pleasures: ISBD(CM), AACR2, and map cataloging: SLA Geography and Map Division Bulletin no.126, p.2-12.

Library Journal, 1981, Research Libraries Group reconfigures RLIN data base: v.106, November 1, 1981, p.2067-2068.

McGarry, Dorothy, 1982, written communication dated April 2, 1982: on file at the Physical Sciences and Technology Libraries, UCLC, Los Angeles, California.

Map Online Users Group Newsletter, June 1980, OCLC centers of responsibility: map cataloging: no.4, p.1.

------October 1980, Cooperative cataloging: no.5, p.2-3.

Mattews, J.R., 1979, The four online bibliographic utilities: a comparison:

Library Technology Reports, v.16, no.6, p.665-838.
Minton, J.O., 1978, OCLC map cataloging: an emerging standard practice: SLA Geography and Map Division Bulletin no.113, p.25-40.
OCLC Map Users Group Newsletter, February 1980, Sample workform: no.2, p.2.
Parr, Thomas, 1975, Automation of cartobibliography; review of MARC for map library information retrieval & cartographic bibliography: SLA Geography and Map Division Bulletin no. 100, p.38-63.
Savage, Noelle, 1982, RLIN plagued by downtime, staff, and $$ shortages: Library Journal, v.107, no.11, p.1027-1030.
Schroeder, J.R., 1981, Perspectives on map cataloging and classification, in Larsgaard, M.L., ed., Map librarianship and and map collections: Library Trends, v.29, no.3, p.419-438.
SLA Geography and Map Division, 1979, Map Librarian's Professional Concerns Committee on cooperation with the Library of Congress Geography and Map Division Bulletin no.118, p.33-41.
Stevens, S.D., 1980, Anglo-American Cataloguing Committee for Cartographic Materials: Western Association of Map Libraries Information Bulletin, v.11, no.2, p.117-126.
University of Oregon, 1981, A comparison of OCLC, RLG/RLIN, and WLN: Journal of Library Automation, v.14, no.3, p.215-230.
Webster, J.K., 1980, The bibliographic utilities: a guide for the special librarian: New York, Special Libraries Association, 32p.

Appendix A

Acronyms Used

AACCCM	- Anglo-American Cataloguing Committee for Cartographic Materials
AACR2	- Anglo-American Cataloging Rules, Second Edition
ALA	- American Libraries Association
INFLA	- International Federation of Library Associations and Institutions
ISBD(CM)	- International Standard Bibliographic Description for Cartographic Materials
LC	- Library of Congress
MAGERT	- Map & Geography Round Table
MARC	- Machine Readable Cataloging
MOUG	- Map Online Users Group
MOUG C^4	- MOUG Cooperative Cataloging Coordinating Committee
OCLC	- Online Computer Library Center
RLIN	- Research Libraries Information Network
SLA	- Special Libraries Association
SUNYA	- State University of New York at Albany
USGS	- United States Geological Survey
WAML	- Western Association of Map Libraries
WLN	- Washington Library Network

Appendix B

OCLC MAP CATALOGING WORKSHEET rev 2/81

OCLC: ________ Rec stat: n Entrd: ________
Type: e Bib lvl: m Lang: ____ Source: d Form: Relief: ____
RecG: _ Enc lvl: I Ctry: ____ Dat tp: _ Govt pub: _ Indx: __
Desc: a Mod rec: Base: ____ Dates: _____,_____

010 ___-_________ (LC card number)

040 ≠c ^XX (cataloging source)

020 ______________ (ISBN)

034 _, a ≠b ________ (scale den.) ≠d ________ (west) ≠e ________ (east) ≠f ________ (north)

≠g ________ (south) (coordinates)

041 ______________ (languages)

043 ______________ (geographic area codes)

052 ______ ≠b ____ (LC geog. class. & subarea code)

086 _, ______________ (government doc. number)

090 G_____.________ _____ ≠b ._____ (LC call number, local)

049 XXXX (local holdings)

1__ _, ____ (main entry)

245 _, ____ (title statement)

250 _______ ed. (edition statement)

255 Scale 1:__________ ≠b ______________________ proj.
≠c (__ ___°__'__"--__ ___°__'__"/__ ___°__'__"--__ ___°__'__")

260 _, ______________ : ≠b ______________________,
≠c _____. (imprint)

265 ____ (source)

300 ___ map(s) : ≠b col. ; ≠c ____ x ____ cm. (collation)

4__ _, ____ (series statement)

5__ ____ (notes)

6__ _, ____ (subject headings)

7__ _, ____ (add. ent.)

8__ _, ____ (series add. ent.)

CURRENT ISSUES IN MICROCARTOGRAPHY

Robert P. Markham
University of Northern Colorado
Librarian, Coordinator of Archives and Micrographics, the James A. Michener Library, Greeley, CO 80639

Abstract--Map storage and retrieval presents special problems because of the size of maps, the degree of detail included, and the use of colors and color coding. Several techniques are currently being used for microform storage and retrieval systems, including roll microfilm, aperture cards and jackets, microfiche, and color microforms.

Projects using these systems are described, including some of the values and limitations of each of these systems, with special attention given to user access and updatability. Observations are made about computer output microform cartography.

Current Issues

Since the days of the Great Library at Alexandria, scholars, librarians and cartographers have been bothered by the recurring problem of information storage and retrieval on maps. Maps do present a series of problems because of their great diversity in size and content, their wealth of detail in many cases, and the use of color to delineate necessary features. For some of these problems, microforms contribute toward a solution and provide the user with access not possible with any other current format.

"Microcartography" is a term which seems to have been coined by Larry Cruse of the University of California. It is used to refer to the art and science of transferring maps and cartographic information from paper to a microformat, especially microfilm or microfiche. This paper will describe some of the projects currently underway using microformat to convey a variety of cartographic information to a variety of map users. It is by no means an attempt to provide an exhaustive survey; it will note some major projects and some of their unique features.

1. Roll Microfilm Applications

As is well known to users, roll microfilm, especially 35mm, has several distinct advantages and has been used for several decades. The chief advantage is that roll film will retain exactly the file order in which the documents, charts, drawings, or illustrations were assembled. In some cases, differing reduction ratios can be used to accommodate the information to the image size available on the film. It is these qualities which probably led to the use of roll microfilm by the United States Geological Survey when it prepared its excellent collection of topographic maps, "The Historical Reference File." The Survey collected all of the published maps, including a great many which were out of print or superseded and assembled them by state. Most of the maps were quadrangle maps in the 7½' and 15' series, 1:62,500 and 1:24,000 in scale.

The maps were filmed in alphabetical order by map name, and within that sequence they were filmed in chronological order beginning with the oldest printed map located among those published by the Survey. The maps are filmed in cine mode at what appears to be a uniform reduction ratio of 20:1. The first 184 rolls form a superb historical collection. Few libraries if any except the Survey Library can claim so fine an historical collection. Users can obtain this historical collection from the Survey for a relatively small cost (1).

Soon an additional 60 rolls of microfilm updates were added to the collection. Other supplements and additions were made to the file, including the 15' Corps of Engineer maps for some areas. The fact that it was necessary to add these supplementary rolls indicates some of the values and the limitations of roll microformat: roll film does indeed maintain the exact sequence in which the materials are filmed but it cannot be easily updated. For the intended historical purpose, that was acceptable; for the user who may not know a specific publication date or who wants the latest edition, roll format becomes increasingly difficult to use. One can rarely be sure which roll has the specific map one wants; furthermore, eyelegible title targets on the box or roll rarely contain enough specific information to satisfy a variety of users. Theoretically one could splice new images into the roll, but that is rarely possible or desirable. If, for example, one is studying the terrain of the Dinosaur National Monument, one needs maps from at least two states, Colorado and Utah. One also needs such alphabetically disparate maps as Jones Hole, Zenobia Peak, Canyon of Lodore South, Hells Canyon, Stuntz Reservoir, and Split Mountain. Roll microfilm has the data; they are not easily used for studying adjacent maps, and finding updated maps is even more difficult in many cases.

2. Aperture Card and Jacket Applications

Tab Cards with an aperture provided for the insertion of a microfilm master negative 35mm film "chip" have been in use for some 20 years (2). Aperture cards provide a unit-document system in most cases. Each card normally has only one film chip. Identification can be easily provided by typing or keypunching the necessary information on the card thereby markedly improving storage and retrieval. One writer involved in design and implementation for space satellites notes one of the advantages of microformat:

> It didn't make much sense to have to support our... small, 50 pound (22 kilogram), satellite at the launch site with over 100 times its weight in paper prints. (3)

Aperture card format has been widely used by engineers, draftsmen, architects, and many others. The Colorado Department of Highways is collecting a series of its historical maps on aperture card so that they can be easily updated, added to a collection, and provided with adequate title information on the card in eyelegible letters.

Furthermore, the micrographics industry has a range of equipment to handle filming, card insertion, duplication, and filing. It is a practice in wide use and will not soon be outdated.

All of this may well be true, but the system is not as simple as one might think. Aperture cards were designed for machine sorting and filing but obviously the space occupied by the film chip cannot be used. I personally have never seen or known a system in which such machine filing was in fact used to any significant degree. Some of us have seen disaster when ordinary tab cards go through a sorter; we would not want costly film to suffer the fate of some cards. Some of my colleagues tell me that in most cases, the computer staff have not even seen aperture cards before when they are first brought to the computer facility. Other colleagues report that equipment specially modified to work with aperture cards is no longer manufactured and may only be obtained second-hand. It would seem, therefore that while machine sorting of aperture cards is technically feasible is not always practical. Is there perhaps a better alternative?

In the last decade or less, a variety of alternative microforms have been introduced as microfilm jackets with a number of different size channels into which film may be inserted. They are commonly used for inserting 16mm roll film but they can also be used for inserting 35mm film. The result is a microfiche form of approximately 4" x 6" or more correctly, 105mm x 148mm. These jackets have ample space for a printed or typewritten title or header in order to provide title or filing information for full identification.

This jacket format makes full use of the entire fiche; space is not wasted as it is with aperture cards. Maps and drawings can be filmed and inserted in the jacket and some jackets are so manufactured that film of different widths can be used, both letter-size pages containing relevant information, and the full map or drawing. The resulting microfiche can be read in normal office microfiche readers. The fiche can be duplicated easily and filed along with other fiche of the same size. It is especially convenient to file these jackets with other types of reports or computer output microfiche data.

Jackets have other advantages. The film is protected from scratching. Fiche can be updated rather easily by inserting new film into the channel. The fiche are very light weight and can easily be mailed to other locations.

But there are disadvantages too. Jackets normally have unsealed ends and film can slip out or be dislodged. Perhaps even more critical is the fact that though the mylar jackets are very thin, they do place a film layer between the film emulsion and the duplicating film. One can expect to loose 5% to 15% resolution quality with the image when using a jacket. When using typewritten materials, this loss may not be evident; when duplicating maps with extremely fine detail, the loss is very noticable.

Furthermore, mylar is very slippery and flexible. It does not file easily; it can slip from one's grasp with little effort and will not stay upright in a file drawer.

Both microfilm and aperture cards/jackets have good and bad features. Both fit some aplications and are less suitable for others. That is as true with office and information files as well as it is for maps and drawings.

3. Microfiche Applications

The last decade has seen several projects which have employed standard microfiche for maps and drawings. Standard microfiche is normally now defined as having 98 pages on a fiche, seven rows and 14 columns. This is the size used for 8½" x 11" office files. Legal size sheets, fold out pages, and other irregular sizes do not fit the format at the standard reduction ratio of 24:1. But the microfiche size and convenience makes it suitable for a variety of applications. It can carry one image, or many; it can carry printed pages or charts, drawings, and maps. A project may require many fiche, or just a few. A single fiche can be added to the file for updating without updating an entire roll.

Some current projects use this format. Nautical Charts on Microfiche have been available for a number of years. The current president of the company described the collection several years ago (4). Each chart is filmed on one microfiche, though only about one-half of the fiche is actually used. There is no title information beyond the eyelegible title on the chart itself. Maps are apparently filmed at about 5x and the charts can be interfiled and updated or replaced. The low reduction ratio means that the maps are easily read. Users report that they are in a number of seagoing vessels and are especially convenient because of the compact size.

Microfiche was also chosen by Updata Publications for its new collection of Water Supply Papers 1896-1970, but the fiche was a quite different format. Among the numerous papers were a large number of maps in a variety of different sizes. All of the maps were first copied on standard office machine copiers. No one type of equipment proved able to copy all of the maps; different machines copied different features better than others. Maps were copied on 8½" x 11" sheets allowing a slight overlap on each edge of the sheet so that no information would be lost. A given map might have from four to 15 or more sheets to cover the entire map. The image becomes a mosaic made up of the frames on the fiche in grid format. This particular camera could not use a page size larger than the standard 8½" by 11" page. The result is that the image which one finally obtains on the microfiche is a 4th generation: original, photocopy, film negative, and finally the user's duplicate fiche (5).

Those of us who have worked with copies of copies know that each successive generation is inferior to the previous and the fewer generations allowed, the better the quality. Nevertheless, very great care was taken to assure the highest possible quality.

A quite different format was chosen for use by the Micrographics Laboratory of the University of Northern Colorado for its collections of Topographic Maps on Microfiche. The paramount concern was to produce maps of the highest resolution possible because of the extremely fine detail on the maps themselves. The goal of the project initially was to introduce microforms to users in innovative ways, to introduce maps on microfiche as one inexpensive format for users who might not be able to have or maintain a large map collection, and to develop a collection systematically arranged and accessed.

The project staff are currently filming the United States Geological Survey World Topographic Maps (IMW) for the United States, the National Topographic Maps, and the state Topographic Quadrangle Maps, some 40,000 maps in all. After much experimentation, the following features emerged through the years:

a. Map Arrangement. Topographic maps are filmed in geographical arrangement so that adjacent maps may be consulted easily and accurately. Alphabetic arrangement does not allow this.

b. Microfiche Format. Topographic maps are arranged systematically by latitude and longitude coordinates so that maps can be selected and refiled easily. Even if only the coordinates are known, as might be the case when data are obtained from the U.S. National Atlas, the specific map can be easily located which has the desired location.

c. Microfiche Image placement. World Topographic Maps are filmed with two rows of two maps, normally four maps per fiche. The National Maps are filmed with two rows of three maps each. Topographic quadrangle maps are filmed with two rows of four maps. With this format, the collection can easily be updated, or expanded as new information becomes available. In addition, the 15' maps (scale 1:62,500) were actually filmed in enlarged quadrants so that on the microfiche they are seen at the same scale as the 7½' maps.

d. Microfiche Titles. Because of the filming procedures used, there are actually seven rows of title information available. All of the maps are identified by the series to which they belong, World, Nation, or Quadrangle. The maps are assigned a number based on the geographic area in which they are located, e.g., N or S, referring to the Northern or Southern Hemisphere, and A to R, referring to the standard latitude map enumeration. In addition, they are given the numeric designation of the specific National Topographic Area in which they occur, such as NJ 12-12, and the location of that map in the area. Finally, to this number is appended a letter from A to P indicating the sequence of the microfiche within the national area map. This allows the user to know immediately not only the quadrangle name but the national and world maps of which it is a part.

   Titles also include the latitude and longitude, state(s) which are included on the map, map publication date, complete map name, and the date of the microfilm edition and reduction ratio of the map image.

e. Filing Arrangement. Microfiche are filed in microfiche panels to simplify access and refiling. Facing each panel is a printed map guide which shows the layout of the microfiche and all of the quadrangles in that national area. The guide also includes information about special features such as national parks, forests, monuments, wilderness areas, and historic sites.

f. Map Indexes. Each set of maps is provided with indexes appropriate to those maps. The national maps have a special National Lands Index on microfiche, as well as the complete index of names from the U.S. National Atlas on microfiche. The state maps are accompanied where possible by the new Geographic Names Information System indexing all features occuring on topographic maps in a given state under some 60 different feature categories.

g. Large Scale Maps. State sets of topographic maps also include the relevant national and world maps so that one actually has the state maps in the collection at three different scales, 1:1,000,000, 1:250,000, and 1:24,000.

h. Flexible Arrangement. The particular arrangement of the maps is so planned that, although the fiche are duplicated individually by the staff, they may be assembled in various ways. One of the national areas, for example, consists of 128 quadrangles on 16 microfiche; these fit in one microfiche panel, or they may be collected in a handy pocket fichepac. Or, the maps can be assembled to form a collection of the 14,000' peaks, the Continental Divide, or other specific feature. Hikers, bikers, field parties, and travellers can have a collection unique to their needs (6).

The microfiche are designed to be read on an office/desk type viewer or a full-screen PEPCO viewer in which any of the maps can be blown back to full size, whether oriented vertically or horizontally.

The collection is inexpensive and can be obtained for about 1/10 of the cost of the paper maps. Black Diazo negatives are recommended because they cause less eyestrain and are easier to read than positive images. In addition, the negative image is less affected by dust which my fall on the film, lens, or glass flats in the reader.

The last example of microfiche maps is the unusual collection produced by the EROS Data Center of the United States Geological Survey by satellite or by aerial photography. These maps are systematically arranged much like the collection of Topographic Maps on Microfiche. Because they are aerial photographs they must be seen with somewhat different equipment and they give the type of information unique to that type of photograph. This type of image, however, has the great merit over any other form in that it is updated sometimes monthly. For current information and study of the type of detail, such as vegetation or community growth, the maps are unsurpassed. Because these projects will be discussed elsewhere in this conference, we do not wish to detract from that presentation. Microfiche was the format chosen to fit best the particular needs of a rapidly updatable collection of maps.

4. Color Microforms

For many cartographers the use of color is so essential that adequate microcartography or print cartography cannot be accepted without the use of color. One of the best examples is, of course, the various state geologic maps produced by the U. S. Geological Survey. No degree or combination of cross hatch, shading, or other techniques can be quite as effective.

It is to be emphasized at the outset that the best use of color microforms is when there are few distinctive colors and fairly large type. For visual and training purposes, diagrams, and similar applications, color microfiche works well especially when 100s of duplicates are needed. The unit price per fiche becomes very low.

When the same efforts are directed toward color microfiche maps, the results are mixed. Kodak color microfilm does have quite good color fidelity when all of the steps are performed properly. Users have reported effective results (9). In addition, the Surveys and Mapping Branch of the Canadian Department of Energy, Mines and Resources has produced quite satisfactory color microfiche maps on 35mm microfilm. They fill orders upon demand and insert the 35mm film strip duplicates in a microfiche jacket with a title typed on the header. At least the 1:250,000 series of map areas in Canada are available in that format; the samples seen by this writer were well done and moderately priced. The chief limitation of that system was that they did not have thorough or adequate title information. The EROS Data Center has completed a similar project but this writer has not seen the actual samples of work completed. No reports have been given about the plans to update either collection.

Telephone conversations with the color laboratory at Eastman Kodak in Rochester, NY, have turned up a few other samples of maps but these were only isolated projects. A very brief report was made recently about plans to produce nautical charts on microfiche but the report was too limited to elicit further comment.

Color, then, does have advantages, but there are also built in disadvantages. In our own work with color microfiche, we can expect at least $100.00 per single master microfiche and as much as $1.00 to $3.00 for each duplicate, depending upon the quantity. Furthermore color microfilm has poor film resolution quality. In our own tests we could do no better than 40 to 60 lines per millimeter; for black and white film, the minimum should be 100. With our topographic map collection, we exceed 140 lines per millimeter. Given this poor resolution quality, to the trained eye the colors can be true but the image never appears sharply in focus. We well know the constant struggle Hollywood technicians face in trying to preserve the color film masters of such classics as "Gone with the wind." Color experts assure us that color film looses quality even if stored under the optimum conditions. Users can rarely afford such storage in their local situation.

Even though it is highly desirable, color quality and stability do not seem to be likely in the forseeable future unless there is an enormous demand. Meanwhile, silver halide products continue to have a proven archival life greatly surpassing 100 years.

6. Computer-Produced Color Microforms

Only limited reports have been made about color microforms produced by means of computer output (11). There can be little doubt that optical scanning and digital information transmission can result in significant new advances. Another participant in this conference will report on those aspects. All of us have seen enough color enhanced digitized photographs from satellite views from planets not to recognize the great potential for mapping. At present, the process is limited by the quality of the film and the high cost of the hardware needed to produce such images.

Conclusion

Projects such as those cited, and others produced by a variety of sources, support our conviction that there surely will be microfiche in your cartographic future. There is no likelihood that they will replace paper; they can achieve wide acceptance where quantity, size, volume, cost, portability, and convenience are factors.

Those projects will best succeed in which there is a planned information system which integrates the microformat, the user, and the equipment into a cohesive system. When that is accomplished microcartography can achieve its best results and achieve its greatest impact.

ENDNOTES

1. For further information on their historical collection of maps on roll microfilm, contact the United States Geological Survey, National Cartographic Information Center [NCIC], 507 National Center, Reston, VA, 22092.

2. Kinney, R. H., 1982, Engineering microfilm systems: IMC Journal, volume 18, Number 2, pp 13-19.

3. Kinney, 1982, p. 14.

4. Barkee, Charles J., 1981, Maps and Charts on Microfilm? in Cruse, Larry, (editor), Microcartography: applications for archives and libraries: Western Association of Map Libraries, Occasional Paper Number 6, pp. 9-15.

5. Water Supply Papers 1896-1976 & Ongoing, 1982: Updata Publications, Inc.

6. Markham, Robert P., 1978, Topographic Maps on Microfiche: The Journal of Micrographics, May/June, pp. 315-318.

7. For further information on the EROS products, consult their Lansat Microcatalog and their Aerial Photography Micrographic Index. EROS Data Center, Sioux Falls, SD, 57198.

8. Goulard, Claude, 1981, Color microfiche: myth or reality? IMC Journal, volume 17, number 3, pp. 11-13.

9. Energy, Mines and Resources Canada, Surveys and Mapping Branch, 615 Booth Street, Ottawa, Canada, K1A OE9.

10. Grimaldi, John E, 1981, Computer-generated color microfiche today, IMC Journal, Volume 17, number 4, pp. 11-13.

11. Exelbert, Rodd S., 1982, Engineering Documentation: the Big Camera, IMC Journal, volume 18, number 2, pp. 9-11. See also Nelson, Carl E., 1965, Microfilm Technology: Engineering and Related Fields: McGraw-Hill Book Company.

DIGITAL CARTOGRAPHY-A GEOLOGICAL PERSPECTIVE

A. Keith Turner

Department of Geology
Colorado School of Mines
Golden, Colorado 80401

Abstract — Geological applications of digital cartography are developing more slowly than for many other applications. Some examples of digital geological maps and hardware and software development trends are reviewed to identify the constraints affecting the implementation of digital cartographic methods to the production of geological maps.

The unique characteristics of traditional geological maps, the conservatism of many geologists and geological organizations, and the lengthy field investigation times which yield low map production rates appear to be the most important factors preventing the existing digital techniques from being cost-effective. Conversely, current and anticipated digital mapping technologies can be used effectively in support of geological, geochemical, and geophysical exploration, property evaluation, and development of resources.

Therefore it appears that digital cartography will become a widely accepted tool for the exploration geologists, but will retain a much more limited role in the production of the traditional, smaller-scale, reconnaissance geology maps produced by government agencies.

Introduction

Maps are fundamental to geology; from the earliest days a geological map was an important product, sometimes the only product, of a geological investigation. Map production is slow and expensive. A review of historical records of geological organizations reveals frequent complaints about the lack of skilled draftsmen, the delays in getting lithographic plates produced, and the costs.

It has been at least 20 years since computer graphics techniques were first applied to cartography. The introduction of digital techniques to cartographic processes offered the potential for solving the production problems of traditional methods. Great strides have been made, yet digital cartography has not yet replaced, to any significant extent, the traditional methods for producing geological maps.

A geological map is a very specialized document. Its formulation has been developed to suit the specific needs of a geologist. Most regional geological maps at 1:250000 scale or larger are printed on topographic base maps. Traditionally, rock units are differentiated by colors and symbols, although cost saving efforts have led to the increasing production of uncolored geologic maps. A number of specialized symbols have been developed to indicate the interrelationships among the geologic units.

It is this last aspect of geologic mapping that differentiates it so profoundly from other thematic mapping such as land cover maps, soils maps, forestry maps, etc. Geologic mapping attempts to portray on the two-dimensional paper the three dimensional interrelationships among geologic units. The symbology to do this is necessarily complex. It is also standardized and largely internationally accepted, although some differences occur among the various national surveys. The symbology has evolved very slowly over the last century, so that maps of 50 or even 100 years ago are readily readable by the modern geologist.

It is widely believed, without much supporting evidence, that the existing geologic mapping standards are the best possible system for transferring information from the geologist producing the map to the map readers - the users. Bickmore (1975, p. 349) reports some disturbing results from measurements of the legibility of maps. These experiments showed the levels of map reading accuracy on geochemical maps were generally low. For example, there was a less than 50 percent success rate in counting the number of symbols, about 20 percent of the answers were not even in the correct contour interval, and, of those within the correct contour interval, a third of the answers were only accurate to within $\pm$ 25 percent of the true value. A further interesting result was the discovery that age or experience did not improve map reading accuracies, only speed! In the study, 11-year old school children performed as accurately as professional geologists.

The lack of extensive publicity about such tests, and of more experimental testing to confirm or disprove such data, has led to a general satisfaction with the status quo. The traditional geological mapping symbols have posed a real challenge to the use of computer graphics. Standard computer equipment of nominal cost either cannot reproduce these symbols at all, or can do so only in an unconvincing manner. Digitally produced geologic maps have been produced, and some are superior in cartographic quality to traditional maps, but only with the use of comparatively expensive hardware.

Digital Cartographic Procedures

System Components

Digital cartographic systems must have four components:

1) Data entry - a method of converting graphical products into a numerical format suitable for entry into a computer, usually through a process called "digitizing".
2) Data storage and retrieval - the numerical cartographic data constitutes a data base, whose organization must be carefully formulated to allow the user to efficiently extract desired combinations of information.
3) Data editing - the user must be able to examine the data base and to make changes and corrections, preferably in an interactive fashion.

4) Data display - the user must be able to produce suitable map products having the information content, scale, and map projection as desired.

Manual versus Digital Procedures

There is a fundamental difference between manual and digital cartography. With manual techniques the map is the primary product; while with digital techniques the numerical data base is the primary product, and maps are secondary products. Digital systems are rarely if ever cost-effective if only a single map product is to be produced; rather they become attractive because of their selective editing capabilities. Maps with a variety of information contents, or scales, or projections can be produced from a single data base. The data base can be selectively updated and new products produced.

The above comments suggest both the advantages and disadvantages of digital cartography. Digital cartography is cost effective where multiple products must be produced, or where the user requests are variable, or where the map data are changing with time (such as land cover data), or where the map is but one product and the data base can be used for other purposes, such as physical modeling for instance. Conversely, manual techniques are usually cheaper where a single map is to be produced and it will remain unchanged for a long period of time, or if scale changes would invalidate the suitability of the data being mapped.

Of course cost is not always the only concern, although it is generally an important one. Digital systems, once fully operational, give the users the option of speeding up the production process. Most digital drafting devices can produce a drawing much more rapidly than a human, and they can work multiple shifts, and do not take coffee breaks. Digital systems allow the users the options of producing different products, ones which cannot be produced by manual techniques. This additional capability is hard to quantify but is a real benefit in many situations.

Digital Data Base Design

Of the four components in a digital cartographic system, the data base design is probably the most critical to the overall success of the system. A poorly designed data base will just not be able to function and produce the products the user desires. Data entry, edit, and display functions can be changed or upgraded over time as user demands change, but the data base structure represents a large and growing investment and it must be suitable.

One problem with cartographic data base design is the shear size of these data bases. Rhind and Adams (1981) give a very instructive review of the problem. They quote Doyle (1978) who pointed out that a gridded elevation model of the Earth's land surface with one point per meter would require 1.6 million magnetic tapes to hold the data and take three years of computer time to read! While such a data base is clearly unworkable, we are producing vast amounts of remote sensor data by LANDSAT for example. The World Data Base II, a numerical outline map of the world's coastlines and related data developed for the CIA, has over 5 1/2 million points located by latitude-longitude. The organization of such data bases is not a trivial task.

Three Competing Data Base Formats

Almost from the beginning of digital cartography there has been debate concerning the most appropriate format for cartographic data bases. Three approaches have been championed, derided, resurrected, and have endured. They are:

1) Gridded, or cellular, or raster data
2) Vector,or are, or polygonal data
3) Triangular mesh data

The detailed discussions of the merits of these quite different approaches are beyond the limits of this paper. It should be pointed out, however, that much of this discussion revolves around the use of these techniques for topographic mapping.

The gridded data method is a popular technique, especially for contouring. It has a number of computational advantages, at a cost of increased data storage. This last concern is diminishing as the cost of memory hardware drops. Gridded or raster data are desired by many of the newer display devices, including refresh screen CRT's, electrostatic plotters, and ink-jet plotters. As the resolution of these devices improves and competes with the more traditional pen plotter devices, raster data structures appear more desirable.

For many years, the best quality computer graphics display devices were vector oriented units - pen plotters of the drum or flatbed variety, and storage tubes. These devices coincided with the still expensive computer memory restrictions of the early 1970's to boost the interest in vector data. The procedure was analogous to the classic map making procedures.

Triangular mesh techniques have always been the "third-force". The early IBM contouring programs used this method, and it continued to be the subject of research throughout the early 1970's. Several commercial applications are based on this method (Males and Gates, 1980), including one for handling geological subsurface data (Gold, 1980).

Review of Pertinent Developments in Digital Cartography

Geologists are interested, to varying degrees, in a variety of digital cartography developments. While geological maps are definitely thematic maps, the geologist is often dependent on topographic maps to form his map base. Frequently a map base at suitable scale is unavailable, in which case photographic enlargement or reduction of one or several maps is undertaken to provide such a base. Such procedures often violate map accuracy standards, yet the geologist rarely concerns himself with such matters, perhaps because the positional reliability of many of his geologic contacts is much less than any topographic features. Geologists also extensively use aerial photographs and remote sensing imagery.

Digital Topographic Mapping

By the early 1970's a number of digital topographic mapping systems were in place and operational at civil engineering organizations. One of the earliest was at the Ontario Ministry of Transportation and Communications which became operational in 1970 (Turner, 1981). Other systems have been installed in such states as California, Texas, Georgia and Michigan. All of these sytems were concerned with the production of large scale (1:500 to 1:2500) construction plans. A logical extension of these technologies has led to the production of utilities mapping by utility firms or municipalities, such as Houston. Boyle (1981, p. 32) points out that there is a difference between such mapping, which closely emulates engineering drawings and other CAD/CAM applications, and national topographic mapping programs. Boyle suggests that the size and complexity of the spatial relationships among map features is so much greater for standard map quadrangles that the earlier digital map drafting procedures will not be satisfactory.

Some experimentation by national mapping concerns, chiefly in Canada (Zarzycki, 1978) and Britain (Rhind and Adams, 1981) led to the implementation of a number of experimental installations. The Canadians have installed a system designed to map at 1:50000 scale in the Arctic and produced maps beginning in 1978 (Zarzycki, 1978).

These systems utilized high resolution flatbed plotters to develop the map plates for printing. Such maps meet or exceed the drafting quality of traditional methods. The development of such systems became increasingly affected by the development of analytical plotters which allowed essentially automated map production. These were introduced in the 1960's but they were, and have remained, extremely expensive, so that by the late 1970's only a few were operational. Their presence is profoundly affecting the longer range planning of national mapping agencies (Zarzycki, 1978; Case, 1981; Southard, 1980; Williams, 1980).

Digital Thematic Mapping

Thematic mapping covers an enormous variety of applications and there are a variety of sytems designed to meet some portion of them. Dudycha (1981) gives a most readable overview of these various types of systems, and the reader is referred to the various volumes of the Harvard Library of Computer Graphics for many examples.

On the one extreme are low cost, low resolution cellular systems producing graphical products on the line printer; while on the other extreme are sytems designed to produce full color maps rivaling the best a human draftsman can produce.

Many thematic mapping systems involve the combining or compsiting of several map products to produce new themes. Polygon intersection algorithms do exist, but most commercially available ones tend to operate inefficiently when very large numbers of polygons are encountered. Raster, or cellular, operations are often more efficient but this requires vector to raster and raster to vector conversions which may be expensive (Peuquet, 1981a; 1981b).

Tomlinson and Boyle (1981) describe a review of most major commercial thematic mapping systems undertaken for the Saskatchewan government. This review concluded that none of the existing systems met the desired operational criteria and recommended that no system be purchased by Saskatchewan at this time.

In spite of the undoubted problems documented by Tomlinson and Boyle, research into thematic mapping continues. The Experimental Cartography Unit (ECU), established in London England in 1967, has developed a number of thematic maps; including several British geologic maps (such as the Abingdon map produced in 1970) and a folio of maps jointly with the Kansas Geological Survey (Campbell, et al; 1979).

In the United States much work has been done by the U.S. Forest Service, the U.S. Fish and Wildlife Service, the U.S. Geological Survey for land use and land cover mapping (Guptill; 1981), and the Soil Conservation Service (Johnson; 1980). All these products use vector oriented data bases, with the exception of some of the Forest Service products. Digital cartography has been used primarily to speed up the production of the map products. The maps are still looked on as a primary product and the digital data base appears secondary. While these data bases are publically available, it is not clear to the author at this time whether they are being used to any extent by second or third party users. The lack of software, standardized data formats, and the incompleteness of coverage to date have tended to limit the utility of these data bases.

Software and Hardware

It is not the intent of this paper to review software or hardware availability. Calkins and Marble (1980) have recently issued a three-volume index of software; while Petrie (1981) has reviewed many important hardware trends.

Obviously the availibility of suitable software and hardware is a necessary condition for the successful use of digital cartography. Most geologists and geological agencies have neither the expertise nor the desire to invent such systems.

A number of trends in technology development promise to encourage the greater use of digital cartography in the near future. We have an enormous amount of existing traditional map data. How can such data be rapidly digitized? Manual line following digitizing appears much too inefficient. Fortunately, a number of scanning systems are now available which can accomplish this task efficiently (Montuori, 1980; Leberl and Olson, 1982).

Optical disk technology is now advanced to the point where home video playback systems are on the market. Such systems fare poorly on the home video market because recording requires very expensive equipment and "clean room" conditions. However, they have the ability of storing vast amounts of map data very economically, without discernable deterioration over time, and capable of very rapid access to any desired scene. These are predicted to be the master storage medium for much topographic data (Southard, 1980).

Assessment

Current Digital Cartography in Geology

At present, the use of digital cartography to produce traditional geologic maps is very limited. Greater use of appropriate technology by the Soil Conservation Service and the Geography Program of the U.S. Geological

Survey demonstrate the technical feasibility, as does the Kansas Geological Survey experimental quadrangle (Campbell et al, 1979). Why then are the geological organizations not using this technology more? The answer appears to lie partly in the conservatism of many geologists and geological organizations, and partly in the economics.

Geology maps represent the results of lengthy field investigations and their production may take years. Users are not as readily apparent or clamoring for the data, as are the agricultural interests for the soils surveys. Geology maps do not change with age and the survey is usually done with a precision commensurate with the expected map scale, so the ability to enlarge or reduce the map scale is minimal value.

These factors appear to minimize the desire for geological organizations to invest in the very expensive hardware needed to produce traditional geologic map equivalents using digital methods.

On the other hand, geologists involved in exploration have often departed from traditional map displays. Depending on the type of exploration being undertaken, a variety of geological, geochemical and geophysical data may be acquired. Such data may be plotted, contoured, or stratified in many ways. Maps, cross sections and isometric views may be employed to aid the geologist make his interpretations. As exploration continues and drilling is undertaken a much more detailed three-dimensional model of subsurface conditions may be developed; one in which the depth coordinate may be significant with respect to the north-south and east-west extent of the study.

Such considerations have led geologists to evaluate the possible benefits of data manipulation by computers. Whether the exploration be for oil and gas, coal, base metals, or precious metals, the use of computers is increasing. Oil and gas exploration is the most advanced because the thousands of well logs were early candidates for computer storage. Feineman (1980) and Gold (1980) give two examples of the methods being used.

The Future for Digital Cartography in Geology

I anticipate an expansion of the use of digital cartography methods by geologists. More and more geological data are being processed by computer as a result of routine data management procedures. Chemical results for a series of rock samples are stored in a computer as part of a "word processing" sequence, or for convenience in transmitting data from laboratory, to office, to field. Once entered, the data need only the addition of coordinates to be plotted. Geologists still have a graphical, or cartographical, bias and the urge to see the results in correct spatial relationsnips is overwhelming.

The production of "traditional" maps may be more slow in coming, but I expect to see increased digital activity in this area also. The availability of digital terrain map products is going to increase quickly in the developed countries and once it is available, the need to merge geological and other thematic data with the topographic data will become obviously desirable. The availability of methods of rapidly digitizing existing older maps, via scanner techniques, and the ability to create large browse capabilities using optical disk technology with digital data base creation on demand, will encourage the trend toward digital techniques.

Such developments are likely to be constrained in the near term by a lack of financial resources to purchase the expensive hardware required, and by the lack of suitable software (Petrie, 1981; Zarzycki, 1979). There appears to be less of a restriction concerning available cartographic data bases (Burkart, 1981; Rhind and Adams, 1981).

References Cited

Bickmore, D.P., 1975; The relevence of cartography: Display and Analysis of Spatial Data (J.D. Davis and M.J. McCullough, eds.), J. Wiley and Sons, p. 328-351.

Boyle, A.R., 1981, Concerns about the present applications of computer-assisted cartography: Cartographica vol. 18, no. 1, p.31-33.

Burkhart, B.R., 1981; Cartographic base files at Lawrence Berkeley Laboratory Harvard Library of Computer Graphics, vol. 13, p. 11-46.

Calkins, H. and Marble, D.F., 1980; Computer software for spatial data handling: Intl. Geographical Union, Commission on Geographical Data Sensing and Processing, Ottawa Canada (3 vols.).

Case, J.F., 1981; Automation in Photogrammetry: Photogrammetric Engineering and Remote Sensing, vol. XLVII, no. 3, p. 355-364.

Doyle, F.J., 1978; Digital terrain models: an overview: Photogrammetric Engineering and Remote Sensing, vol. XLIV, no. 12, p. 1481-1485.

Dudycha, D.J., 1981; The impact of computer cartography: Cartographica, vol. 18, no. 2, p. 116-150.

Feineman, D.R., 1980; A case study in computer mapping in petroleum exploration : Harvard Library of Computer Graphics, vol. 10, p. 41-51.

Gold, C.M., 1980; Drill-hole data validation for subsurface stratigraphic modeling: Harvard Library Computer Graphics, vol. 10, p. 52-58.

Guptill, S.C., 1981; Thematic map production from digital spatial data: Harvard Library of Computer Graphics, vol. 15, p. 121-124.

Johnson, C.G., 1980; National Soils data base from the Soil Conservation Service: Harvard Library of Computer Graphics, vol. 8, p. 21-30.

Leberl, F.W. and Olson, D., 1982; Raster scanning for operational digitizing of graphical data: Photogrammetric Engineering and Remote Sensing, vol. XLVIII, no. 4, p. 615-628.

Males, R.M. and Gates, W.E., 1980; ADAPT: A digital terrain model-based geographic information system: Harvard Library of Computer Graphics, vol. 8, p. 182-196.

Montuori, J.S., 1980; Image Scanner Technology: Photogrammetric Engineering and Remote Sensing, vol. XLVI, no. 1, p. 49-64.

Peuquet, D.J., 1981a; An examination of techniques for reformatting digital cartographic data. Part 1: the raster to vector process: Cartographica, vol. 18, no. 1, p. 34-48.

Peuquet, D. J., 1981b; An examination of techniques for reformatting digital cartographic data. Part 2. The vector to raster process: Cartographica, vol. 18, no. 3, p. 21-33.

Petrie, G., 1981; Hardware aspects of digital mapping: Photogrammetric Engineering and Remote Sensing, vol. XLVII, no. 3, p. 307-322.

Rhind, D. and Adams, T., 1981; Coordinate data bases: availability and characteristics: Harvard Library of Computer Graphics, vol. 13, p. 53-60

Southard, R.B., 1980; The changing scene in survey and mapping: Photogrammetric Engineering and Remote Sensing, vol. XLVI, no. 11, p. 1415-1420.

Tomlinson, R.F. and Boyle, A.R., 1981; The state of development of systems for handling natural resources inventory data: Cartographica, vol. 18, no. 4, p. 65-95.

Turner, J.B., 1981; Computer aided photogrammetry at M.T.C.: Workshop for Automated Photogrammetry and Cartography; Amer. Society Photogrammetry, Falls Church, VA., p. 73-80.

Williams, O.W., 1980; Outlook on future mapping, charting, and geodesy systems: Photogrammetric Engineering and Remote Sensing, vol. XLVI, no. 4, p. 487-494.

Zarzycki, J.M., 1978; An integrated digital mapping system: Canadian Surveyor, vol. 32, no. 4, p. 443-452.

Zarzycki, J.M., 1979; Opportunities and prospectives in digital mapping and automated cartography: Proceedings USGS Centennial Symposium, Reston (in press).

RETROSPECTIVE DIGITIZATION OF GEOSCIENCE DATA:

THE EXPERIENCE OF IGCP PROJECT 163

Felix Chayes

Geophysical Laboratory, Carnegie Institution of Washington
2801 Upton St., N.W., Washington, DC 20008

Abstract--The stock of published information about igneous petrology has for some time been expanding at such a rate that its partition by further specialization is exceedingly undesirable and perhaps no longer even feasible. Probably this point already has been or soon will be reached in other naturalistic disciplines.

Although no longer manageable by traditional devices available to the individual scholar, the public corpus of petrology is not particularly large by electronic standards. Experience with International Geological Correlation Project 163, charged with designing and stimulating the development of a world data base for igneous petrology, has brought into focus a number of problems--some dominantly scientific, others essentially social--that will probably recur in attempts to computerize the information bases of other naturalistic disciplines.

The dominantly scientific problems are largely concerned with the nature of the data. Analytic data are numeric and thus easily handled, but much information of interest to petrologists is qualitative; storage methods that will permit efficient machine sorting based on stratigraphic age, mineral assemblages, and various petrographic properties must be developed.

Information available per specimen varies widely in amount and kind. Numerical information is usually in tabular form, but qualitative data are embedded in text. Heterogeneity of data and the difficulty of extracting nonnumeric information from text lead to the first dominantly social problems--the design of coding forms and the organization and administration of the data-capture operation.

Subsequent movement of data from hard copy to the first machine-readable form is a mixed scientific and social problem. The final disposition of the base and its associated information system is again dominantly social. The base must be readily available--either verbatim, in subsets, or in reduced form--to the concerned scientific public, yet some means must be found to preserve its integrity and provide regular updating.

The Need for Retrospective Bases

Until public notice of this meeting was circulated I was not aware that the term "retrospective base" had attained any broad currency, and after some inquiry I still do not know whether it yet has a broadly accepted denotation. What I have in mind by an ideal retrospective base is the electronic equivalent of a specialized public library. The user of such a base would have at his disposal about the same kind and quality of information he would expect to find in a first-class reference library that catered to his needs. And by having this material "at his disposal," I do not mean merely that he could retrieve all or any part of it *verbatim* in humanly legible form, but that at his option he could also either reduce selected data in any way he saw fit before transforming them from bits and bytes into diagrams or tabular summaries, or partially reduce and prepare them for further electronic processing. The electronic library I have in mind would provide its users with all the information of the conventional library and all the data-processing potential of the computation center.

In my view, the retrospective base would contain data only, or data accompanied by minimal amounts of text or paraphrase. It is most unlikely that bases of this sort would ever replace the conventional library. Editors who worry on this score are worrying unnecessarily and, to put it very mildly, unrealistically. But if the data it is to contain are already in print and the conventional libraries that house this printed matter will continue to be indispensable, how do we justify the rather large effort required to build a good retrospective base?

The answer to this question is scarcely any longer novel. For even within closely constrained naturalistic specialties, there is now far more scholarly work to be done than can be done by conventional scholarly means. This is so well known as hardly to require extensive documentation, but my own field of interest is so rich in examples that I cannot resist citing one.

If in 1920 you had wished to become an expert on the chemical composition of Hawaiian lavas, you would have had to familiarize yourself with just 26 analyses; that's all there were in print. The petrologist of that day could be expected to have firm and ready command of what was known about his subject, and if his memory was less than photographic or his recall less than instant, there were admirably organized and remarkably exhaustive encyclopedic tabulations to help him out.

If you had attempted this Hawaiian exercise after World War II instead of World War I, in, say, 1955 instead of 1920, you would have needed thorough familiarity with or ready access to perhaps 500 published analyses; by 1976 the number had risen to well over 1000. And, if your memory had flagged, you would have found there was no longer a reasonably current, printed encyclopedic tabulation you could turn to for assistance. Nor is there likely to be one again.

This has been the situation for some time in igneous petrology, and I suspect it also is or soon will be the situation in other fields of geological specialization in which either random-sampling variation or the secular distribution of identical or similar objects in time or space are major concerns. It seems to me that the penalty for not learning to computerize the information corpus of any such field of inquiry will be either a precipitous decline in scholarly standards or a fragmentation of scholarly interest and responsibility. For whenever we are faced with more information about a subject than we can control we must either tacitly agree to know less about the subject as a whole or explicitly agree to concern ourselves only with particular parts of it. (Usually, I suspect, we end up doing both these things.)

In these days of already excessive specialization, and of declining academic standards in so many other respects, it seems to me that neither alternative is acceptable. I believe, further, that both can be avoided. In petrology, at least, although the amount of published information is far larger than can be managed efficiently by conventional scholarly methods, it is not particularly large from the viewpoint of machine management.

The problems we would--and, I believe, will--encounter in shifting the retrieval and reduction activity of any naturalistic discipline from conventional scholarly data processing to the electronic variety are many and varied. Further, they are for the most part political, social, or human problems, so they do not have clean, neat, unique solutions. But experience with IGCP Project 163, whose mission is to design and stimulate the development of a retrospective world data base for igneous petrology, persuades me that practical solutions could be found for most of them.

The rest of my talk will be based on the work of Project 163, and I know very few of you will have any interest in a detailed account of our history and activity. What you do and should want is not detail but generality. Although I'm persuaded there is, or soon will be, a general need for retrospective data bases in many naturalistic disciplines, I'm not entirely sure of the conceptual generality of the problems to be faced by our different disciplines in constructing them. Using the term "general" in its common rather than mathematical sense, it is on the apparently more general problems--or, at any rate, those that seem more general from the point of view of a specialist in petrology--that I shall focus the remainder of my talk.

Variations in the Amount and Kinds of Information per Sample Unit or Item

The basic "atom" or unit concerning which information is to be collected will of course differ from discipline to discipline. Experience with IGCP Project 163's base and in the construction of its better known predecessors leaves little doubt that in igneous petrology the "natural" sample unit or item is the chemically analyzed specimen. The three largest predecessors--my own (Chayes, 1971; Chayes *et al.*, 1976), Roger LeMaitre's CLAIR (LeMaitre, 1973), and Felix Mutschler's PETROS (Mutschler *et al.*, 1976)--were all begun with the notion of providing encyclopedic collections of what are called, in petrology, "major-element" or "essential-oxide" analyses. Each makes provision for inclusion of other information--CLAIR is in fact a data-base management system that can be used for data of any type--but I think it is fair to say that in the now rather impressive retrospective data gathering activities associated with all three there has been no systematic attempt to extend information coverage beyond essential oxide analysis, rock name, and specimen location.

The information coverage being attempted by IGCP Project 163 is of an entirely different order. In addition to essential oxide, name, and location information, as in existing bases, the IGCP base is to include, for each specimen, published information about trace-element content, stratigraphic and radiochemical age, mode of occurrence, texture, structure, mineral assemblage, and mode and type of alteration. Finally, an "additional information" section of the specimen description is to contain tagged and framed but otherwise unstructured information about field occurrence and association, modal analysis, refractive index, specific gravity, and other properties either uncodable or not yet coded.

For most specimens, most of this information will of course be lacking. Our glossary, for instance, contains more than 150 petrographic descriptors, no more than 15 of which have ever been cited in a single specimen description; to date the average number per specimen is less than 5. The situation with regard to mineral descriptors is very similar. Again, any element in the periodic table may be included in the trace-element vector of a specimen description but most specimen descriptions contain no trace-element information at all, and in those that do the average number of element determinations recorded is less than 10. The potential size of a specimen description and the incompleteness of most actual specimen descriptions bring into question the once conventional concept of an item description consisting of a fixed number of lists of fixed length. Indeed, the most succinct way to characterize an item description of the sort we must process is to say that it consists of a number of lists most of which are optional and of variable length. My own view of this situation--not fully shared within the project--is that it creates no severe technical or programmatic difficulties. It powerfully affects every social aspect of the effort, however, and it is to these social effects that I now turn.

Data Capture

Numerical analytical data, whether of "major" or "trace" components, are usually posted in tables, from which they may be extracted efficiently by bright, well supervised, and well motivated student labor. The nonnumerical information, on the other hand, is nearly always imbedded in text--indeed it _is_ the text--and proper extraction of it requires close, discriminating reading by or under the immediate supervision of skilled professional labor. The amount of it awaiting extraction is so vast, grows so rapidly, and appears in so many journals and in so many languages, that there seems no practical prospect of funding its extraction as an in-house operation, whatever the house. Data capture will have to be a cooperative venture and essentially a volunteer one. I suppose that whether or not such an effort can be mounted and sustained is a pretty good test of whether a subject needs and merits a comprehensive retrospective base.

The Coding Form

Both the heterogeneous nature of the data and the complicated volunteer organization almost certainly required for their harvesting must strongly influence the design of the coding form. We are currently using a single-sheet form that spares the contributor as much writing as possible. It contains the complete glossary of mineral and petrographic descriptors, each keyed by a two-character symbol; the contributor merely circles those symbols applicable to a particular specimen description. He may also have a couple of short external "table look-ups" to do--for mineral-habit symbols and geologic age terms--but we have kept these as undemanding as possible. We had hoped that our rather unconventional form would involve only slight modification of the reading habits of our volunteer contributors and would at the same time be usable as copy by commercial key-punch personnel. Most contributors nevertheless report having to read _much_ more closely and carefully than they usually do, and many of us have discovered that this is, by turn, invigorating, exasperating, and remarkably enlightening.

Our coding form is almost a personal affront to anyone firmly committed to the traditional, column-per-character, line-per-card coding form. And the project does have a few traditionalists whose contribution to date consists largely of objecting to the form rather than filling it out or proposing alternatives that would be at the same time viable and more conventional. The hardly novel lesson to be learned here is that nothing much is to be expected in the way of data capture from people who don't like the data forms they are asked to work with. Perhaps there should be alternative data sheets and alternative procedures for moving data from them to the first machine-readable form. This is an administrative accommodation we have so far been unable to afford. Currently, we are trying to compensate in the opposite direction, preparing a translator that will permit contributors who exercise their privilege of drawing from the base in its prenascent stage to obtain output in just about any format they are willing to specify in sufficient detail.

Data Transfer

In connection with the critical transfer of data from hard copy to the first machine-readable form, the heterogeneity of the data and the requirement that as much of it as possible be available for machine sorting almost proved our undoing. It was not that the grammatical rules were difficult, but just that there were so many of them that the rate of card generation was painfully slow.

For the same reason, I suppose, commercial key punchers were remarkably uninterested in our work. As many of you no doubt recall, in the U.S. this was both a labor-intensive industry and one in which much of the labor was casual. A good shop could do wonders working from a straightforward column-per-character, line-per-card data form and wasn't interested in doing anything else. But that wasn't what we had to offer.

About 2 years ago, in an attempt to reduce a backlog rapidly approaching critical dimensions, we developed a conversational program that prompts copiously, accepts character strings as input, and generates system grammar and syntax, down to the subfield level, internally. This has proved highly successful in our central office, and a number of the U.S. group are beginning to use it in preparing their own initial submissions.

For a marginally financed project like ours--and I suspect marginal support is one property we shall have in common with many base-building ventures in the naturalistic sciences--decentralization of the data-transfer work is highly desirable. Financial considerations aside, it is neither an unmixed blessing nor an unmixed curse. Strongly in its favor, I think, is the sense of community involvement and commitment it creates. Although we aren't really past the planning stage in decentralizing our data transfer, for everyone involved it has been a stimulating, even--if one may use a rather preppy expression--an exciting experience. But it does create problems. Or perhaps it only directs attention to problems not so far satisfactorily solved. One of these is verification.

Verification

If a single office is responsible for all data transfer, responsibility for verification clearly rests upon the transfer operator. At current support levels we cannot afford the duplicate typing required by conventional verification, and this disability would probably persist even if support were somewhat increased.

A conversational program designed to detect and report the commoner species of transcription error and general data infelicity is being developed by one of our members. Although not strictly a verifier, it will promptly focus attention on file entries that are either suspect or clearly wrong; providing it does not deflect attention from suspect areas it fails to detect, it should greatly facilitate visual verification. To be used first as an ex post facto inspection device, this module may ultimately be attached to the data-transfer program, so that its major function will be preventive rather than corrective. But all this is in the future; currently, we rely on visual comparison of computer listing with hard copy.

When we are satisfied that transcription errors have been eliminated, a listing, or "galley," of each file is mailed to the concerned contributor for "proofing." At the present time each contributor reaches his own decision about what constitutes satisfactory proofing; our only lever is a reminder that, in recognition of his time and labor, the contributor's name will be included in the base as part of the data he submits. The resulting working definitions of "proofing" vary widely. Some contributors simply respond with a thank-you note; some conscientiously compare the galley with the original source references rather than merely with the coding forms.

Lack of standardization of the proofing operation to one side, the overall verification procedure is now fairly straightforward. But what becomes of it if we are successful in decentralizing the data-transfer operation? I suspect we shall then have to turn our current procedure upside down, assume that the contributor has proofed his file before he submits it, and hold the central office responsible for a final verification of file against coding forms.

Keeping the Base Current

In the event the base is brought to full currency, it will not remain there for long without systematic updating, and unless it is kept current it will soon be obsolete. Practical arrangements for organizing and managing periodic updating will no doubt differ from discipline to discipline. In petrology it seems to me that something similar to the IGCP project, but more stably organized and funded, would be appropriate. The continuing cooperation of national groups like those being developed by Project 163 would certainly be advantageous and may be indispensable; since such bodies tend to have governmental or quasi-governmental affiliation, the central office ought to have some international scientific affiliation. Such an office would serve as liaison between the concerned scientific public and the organization that housed and serviced the base. Establishing appropriate organizational structure--the in word for this sort of thing, I think, is infrastructure--will be critical to the success of almost any such enterprise, and I wish I had sound advice to offer on the subject. Alas, I do not. We have not yet made satisfactory arrangements, and it is not impossible that we shall cease to exist as an IGCP project before they are completed. I devoutly hope this will not be so, for it could well be lethal. Perhaps the most important lesson to be drawn from our experience is that a long lead time must be provided for planning and arranging final disposition of the product.

Data Integrity

In computer jargon "data integrity" refers only to the preservation of information during its electronic manipulation. This is a very important concept, and of course we are deeply concerned with it, but largely as spectators. There is a broader sense of the term, however, that raises issues not peripheral to our competence, issues that will have to be faced in every similar base-building activity. Several such issues surface in the early stages of data capture, and a very serious one may be expected to arise in connection with dissemination of information from the operating base.

In the context of a retrospective base, the very term "data capture" creates the presumption that clearly specifiable, perfectly objective information is to be moved from the printed page to the coding form, and this is certainly the case for most of the data. Contributors are urged to regard themselves not as editors but as reporters, and most of the time that is what they are. But all of us closely involved with the project have come to realize that we cannot altogether avoid editorial responsibility, that in fact there may sometimes be strong editorial overtones in capture operations that seem at first glance purely reportorial. Time permits introduction of only a few examples.

There is intermittent difficulty with the names of rocks and minerals, even though the system glossary contains over 400 of the former and 170 of the latter. Obsidian and rhyolite for instance, are both system-recognized rock names, but which do we use for the IGBA name of a rock referred to in the source reference as a rhyolitic obsidian? Again, in processing numerical data we conscientiously try to distinguish between major oxide analytical results reported to tenths and hundredths of a percent, but how do we classify an analysis in which the second digit after the decimal point is always or nearly always a zero? Our procedure for recording "tenths' percent only" data is simply to leave the hundredths' place blank; a system program encountering a blank here rather than a digit will "know" it is dealing with tenths' percent data. If the author has put a string of zeroes in this place, are we violating the integrity of the data by deleting them?

In the computational sense of the term, one is not concerned with maintaining data integrity until, as in the preceding examples, it has been decided that the data are indeed to be included in the base. In the common sense of the term, however, it is sometimes intricately involved in reaching that decision, forcing us to reconsider the concepts and definition of the subject matter specialty the base is designed to serve. The mission of Project 163, for instance, is to build a base containing information about terrestrial igneous rocks. Do we then include data about rocks given "igneous" names in a source reference whose author is persuaded they are of anatectic or metamorphic origin? Again, if someone infers that the metabentonite seams found in Ordovician limestones of the Midwest are volcanic ash, and considers the interpretation sound enough to warrant use of the name "ash" as a column caption in a table of analyses, are the data to be included in our base?

Similar "within-" and "pre-base" questions bearing on data integrity will surely arise in attempts to computerize the corpus of information of any natural-historical discipline whose field of interest is more than razor thin. The experience of Project 163 to date suggests that

(a) practical solutions of "within-base" integrity problems can usually be found;

(b) solutions of "pre-base" questions are nearly always social or political rather than scientific and may even be impractical though unavoidable;

(c) neither type of question is likely to have more than casual interest for most participants in the base-building operation;

(d) interesting or dull, profound or trivial, questions of this kind arise only in connection with a very small part of the information to be stored in the base, and should not be allowed to interfere with or inordinately delay its development.

These generalizations are of course subject to modification by further experience; clearly, however, if either (a) or (d) becomes untenable it is time to redesign or abandon the base-building activity.

Although we have so far had no formal discussion of it, I believe a much more difficult problem of data integrity is likely to arise at the other end of the line, when the base begins to be widely used and, presumably, acknowledged. Created by broadly based cooperative activity, our base, and others like it, will be a kind of public property. Yet with regard to data display, applications of these bases will be increasingly private.

The mere fact that the data are publicly available will make it unnecessary to reproduce them in full in any specific application; a brief acknowledgment that data summarized in certain figures or tables have been drawn verbatim or selected in some systematic fashion from a particular version of the base will probably be considered sufficient in most cases. Providing the way in which it has been used is adequately described and readers can be assured of the integrity of the base, there seems nothing to fear in this respect. Assurance of continuing postrelease integrity, however, may be extremely difficult to guarantee.

To accommodate the concerned scientific public it will probably be necessary to distribute copies of the base to a considerable number of national or regional offices. These offices will all have other fish to fry, and most of them will justify domestic support of their activity in handling the base on the grounds that it is so like their other duties. These duties, the ones for which governments are willing to pay, will usually involve maintaining large local data bases, and it may often be politically and fiscally advantageous to pool the domestic and international bases.

It is also possible that commercial information vendors will add naturalistic bases to their inventories. This possibility is remote enough for a base concerned with igneous rocks sensu strictu that I don't feel obliged to worry much about it in my present capacity. But the probability of commercial interest in a base closely concerned with rock alteration and mineralization, for example, would be much higher, and one may anticipate almost immediate vendor interest in bases having anything to do with fossils or stratigraphy.

Finally, with continued improvement and expansion of all computation facilities, increasing numbers of individual petrologists will wish to maintain complete copies of the base in their own institutions, whether educational or industrial.

With very rare exceptions, probably mostly in the last category, none of these contingencies creates a serious threat to data integrity in the computational sense of the term. All of them, however, may sometimes put data integrity in the common sense of the term in serious hazard. No public library deletes information from or adds it to the reference works in its collection, and most public libraries try to penalize users who do so. Although individuals who own reference works often make marginal notations in them, in technical applications no one considers such a notation part of the reference in which it occurs. But it will be easy to fall into the computational equivalents of such malpractices.

Pooling and editing of data is a constant temptation in the manipulation of electronically stored information, for the ability to merge and edit files is surely one of the major advantages of electronic over conventional data processing. Inadequate description and comparison of sources in publications reporting results based on mixed or edited files readily converts what is intended as a scientific argument into an argumentum ad hominem. When the pooling is of public and private files, the former will be cited in the text and included in the list of references, but even the careful reader will have no way of appraising the relation between the public data and the announced conclusions unless the text contains a fairly detailed comparison of data from the two sources. Because of their ready availability, it will often happen that the specifically public data and their bearing on the argument simply disappear from view. Yet the immediate source of these data, the public data base, will probably be generally considered to support the conclusion, and will almost certainly be so cited in secondary publications. As far as its audience is concerned, what looks like--and may well be--a scientific discussion then degenerates into little more than a statement of opinion.

It is incumbent on all of us participating in the construction of public bases to see to it that there is always a ready escape from this dilemma. The bibliography of a discussion that utilizes previously published data always provides entry to the source references. An analogous search facility must be available to readers attempting to appraise, modify, extend, or apply results based on electronically stored public data. For the creation and maintenance of this facility some kind of reasonably stable oversight committee, acting under the sponsorship of a professional society or group of such societies, will probably be indispensable. The major substantive responsibility of this committee, as already suggested, would be the organization and management of periodic updates of the base; its principal social obligation would be to maintain and, whenever possible, improve procedures for providing access to the base, so that the base is public not only in principle but in fact.

References Cited

Chayes, F., 1971, Electronic storage, retrieval and reduction of data about the chemical composition of igneous rocks: Carnegie Institution of Washington Year Book 70, p. 197-201.

Chayes, F., McCammon, K., Trochimczyk, J., and Velde, D., 1976, Rock Information System RKNFSYS: Carnegie Institution of Washington Year Book 76, p. 635.

LeMaitre, R., 1973, Experiences with CLAIR; a computerized library of analyzed igneous rocks: Chemical Geology, v. 12, p. 301-308.

Mutschler, F., Rougon, D. J., and Lavin, O. P., 1976, PETROS--a data bank of major-element chemical analyses of igneous rocks for research and teaching: Computers and Geosciences, v. 2, p. 51-57.

APPLICATION AND MANAGEMENT OF DIGITAL DATA COLLECTION IN THE GEOSCIENCES

Ronald J. Walton

U.S. Geological Survey
P.O. Box 25046, MS 911-Denver, CO 80225

Abstract--Some of the systems utilized within the U.S. Geological Survey, the largest repository of earth science data, include the Rock Analysis Storage System (RASS), the Petroleum Data System (PDS), the Computerized Resource Information Bank (CRIB), the National Coal Resources Data System (NCRDS), the Earth Resources Observation System (EROS), and the Water Data Storage and Retrieval System (WATSTORE).

One of the problems faced in data management is the lack of policy for funding the creation, maintenance, and storage/retrieval of data. At our EROS Center this problem has been addressed, however, and a resulting computerized cost recovery system has been developed that works very well. Another problem is how to convey the knowledge that the research scientist has of the stored data to the secondary data users. Perhaps numeric data base users could tie their data more closely to bibliographic files for conveying this information.

I would like to briefly describe some of the current distributed data processing systems in USGS and some of the large numeric data bases in USGS which we are creating and analyzing on these computers. I would like to describe some of the background for our earth science data collection programs. When I discuss data bases, I will start with the RASS data base, which serves as a prototype approach to meeting multiple levels of users. And finally I will try to describe some of the typical uses of RASS and in more detail some of the problems related to making all of these data bases available to the public (secondary users beyond our own project scientists) in a more accurate and timely fashion.

Most computer applications in the earth sciences field involve the collection and analysis of digital data, as Toni Bearman mentioned yesterday. Government is the largest producer and user of data. USGS is a good example. The USGS has numerous digital data collection and analysis programs in support of its resource evaluation programs. USGS is the largest repository of earth science data. The analysis and interpretation of these data and publishing of the results of its interpretation of these data is its primary mission. The gathering, storing, and retrieving of these data in the Survey are done on a highly decentralized basis within groups of what we call special purpose computing facilities. In the Geologic Division alone there are over 30 groupings of mini and micro computer facilities supporting similar programs. Here in Denver alone we have about 20 of these facilities. The largest ones are the Earthquake facility in Golden, the Uranium/Thorium Branch facility in Golden, the five special purpose Geophysics facilities at Denver West, which include these three systems--VAX 11/780 Geophysical modeling, P/E 3320 Image processing lab, and DEC PDP 11/34--and we also have an Exploration Research facility at McIntyre, Golden, the Analytical Lab facility on Ward Road, and the Oil and Gas Seismic Data Processing and Energy Research Information Center and Central Mineral Resources facilities in the Federal Center in Lakewood. All of these facilities include three DEC VAX 11/780's, one VAX 11/750, one VAX 730, and several PDP 11 LSI-11 minis. Also included are many micros, such as H/P, Superbrains, Apples, and IBM Micros. If you get the impression computers are overwhelming us with data, that's partly true.

Other work going on at these facilities includes common computer applications which involve purely mathematical computations and modeling, but these are not as numerous as those applications that collect and analyze data. However, with the current computerized procedures for collecting digital data, the growing volumes of digital data have created a substantial data management requirement within all of these facilities. Data management is essential for effective utilization of all the geoscience data we're collecting. We are working in a highly distributed data processing environment, which complicates the data base management problem. Proliferation of computer tools leads to lack of compatibility. As we get more and more minis and micros, it gets more difficult to coordinate and avoid "re-inventing data management wheels."

Large volumes of data are being collected by USGS in support of research projects. These data most frequently are collected for different purposes by many individuals over a long period of time. There are few, if any, standards and guidelines which address the storage of these data so that others may share the data. These data are collected primarily for a specific purpose or in support of a specific investigation to solve a particular geologic or geophysical problem. There is little guidance in procuring hardware and software and services to ensure compatibility in data management. Once these data are analyzed and research findings published, then these data are made available in accordance with the Freedom of Information Act to the public for secondary users. This process takes generally a minimum of 2 years or more. And there is lots of data that still doesn't get published and sits in scientists' filing cabinets and offices. An active policy on data dissemination is being developed currently in USGS to assist in the data management effort. This is part of a larger information system policy and reorganization effort in USGS--a new Assistant Director for Information Systems, a new Information Systems Division, and a new division also to address standards, compatibility, and to give advice/assistance on procurement.

The first data base I would like to describe is RASS because it is a prototype of how we plan to meet multiple user retrieval access needs. Rock Analysis Storage System (RASS) is a large geochemical data base of the

Geologic Division's Analytical Laboratories and Exploration Research Branches. This data base contains descriptive information on the location, sample type, who took the sample, etc., and the chemical analyses of these samples. The A-Labs portion of the file is a more detailed research chemical file of 200,000 samples as compared to the Exploration file which is a more brief chemical analysis of about 400,000 exploratory samples. Due to the increasing size of the RASS file to support in-house requirements, efforts have been under way this past year to make these data available to users through a commercial computer network which can be accessed nationally by a local telephone call. The A-Labs file is on-line at the General Electric Information System Network. Users can get access simply by going to their local GE computer office and setting up retrieval services. At the same time, for users who want to get magnetic tapes of the RASS file, plans are to make these tapes available to those users through the NOAA/EDIS NGSDC in Boulder, and data for Colorado is on tape and available at NTIS. For casual users who perhaps can not afford either the on-line retrieval service or the complete set of tapes, listings by geographic areas are to be made available for browsing in the USGS library. Here again, however, is one of our problems--over 300 oil and mineral companies in Denver use the library. This prototype approach to Rass provides access at three different levels of use--on-line, batch (tapes), and casual (library).

The Petroleum Data System (PDS) file in the Oil & Gas Branch of the Geologic Division includes information on petroleum reserves, analysis, and production in the U.S. The PDS file is a joint government/industry file developed by the University of Oklahoma Information Systems Programs of the Energy Resources Center and administered through a joint advisory council. USGS underwrites initial development. USGS funding of PDS has been substantially reduced in recent years with stringent budgetary constraints, and in fact as of 1 Oct 82 we are just one of many subscribers. PDS is now being supported through user subscription service charges. PDS is available through the General Electric Computer Network on a service contract basis. This is an arrangement that OU and GE have had for some time. In general in USGS, we're encouraging consideration of other commercial nationwide networks like Tymnet, Computer Science Corp. (CSC), Boeing, Cybernet (CDC), etc. for numeric data bases. Some of these networks have specific software that are useful that others don't.

The Computerized Resource Information Bank (CRIB) file in the Branch of Resource Analysis, Geologic Division, is a data file on metallic and non-metallic mineral resources of the world. CRIB contains information on name and location of mineral deposits, commodity information, production, estimated reserves, and potential resources. CRIB is available also through the GE network for the public. CRIB is also available to USGS in-house users on the Amdahl computer in Reston, VA, at the national headquarters. Efforts are currently under way to revise the CRIB system to incorporate more quantitative commodity data and to conform to other Department of Interior mineral data systems to serve the nucleus of a Mineral Policy Assessment System at the Department level. A preliminary report by the CRIB Revision Committee is to be completed in 1982. This new approach may enhance some of the rapidly developing distributed processing database management/networking technologies. And it may be a protype for future Division systems.

The National Coal Resources Data System (NCRDS) file provides information on location, quantity, and physical/chemical characteristics of coal and coal-related resources. NCRDS is divided into two regional geographic files--Eastern (E Coal), Western (W Coal)--and two active files of the Bureau of Mines--BMALYT, analyses of heat content data and USChem, chemical analyses of coal beds. An NCRDS Task Force in USGS recently conducted a study of data management needs for NCRDS. A decision to place the management of NCRDS in the new Information Systems Division was recently made by the Director. However, the implementation of NCRDS has been delayed by the more recent establishment of the Mineral Management Service by taking the Conservation Division out of USGS. Currently NCRDS is available for retrieval in-house on the Multics computer in Reston, and batch tapes are available from NGSDC. NCRDS as a major new data-base undertaking in the new Information System Division, if funding is available, is expected to be a future prototype for USGS-wide data bases.

There are several large geophysical data bases in USGS. These include land gravity data, world seismic data, volcanic activity data, earthquake data, aeromagnetic data, magnetic anomaly data, and marine geology data. These files are grouped by common characteristics and collection frequencies. I discussed the VAX geophysical modeling, PDP 11/34 and Perkin/Elmer 3320 Image processing system earlier.

Other geologic data files include the Radiometric Age Data Base (RADB) file of radiometric ages published for the U.S. The RADB file is searchable in-house on the Amdahl computer, and batch tapes are available from NGSDC.

Other large data bases outside of the Geologic Division in USGS which I will mention include the water data base in the Water Resources Division, Landsat remote sensing and imagery data in the Earth Resources Observation System (EROS), and map data bases in the National Mapping Division. Water Data Storage and Retrieval System (WATSTORE) is a large data base of stream, river, reservoir, and ground water information. Water sample analyses data are included in this file. WATSTORE is currently searchable in-house on the Amdahl computer in Reston. However, just as the Geologic Division has done, Water Resources is now in the process of getting minis for WATSTORE. The EROS Data Center in Sioux Falls, South Dakota, of the USGS provides access primarily to all NASA's Landsat imagery, aerial photography, and remote sensing data. Recently digital map data have become available through the National Cartographic Information Center (NCIC). You will be hearing more about map data in tomorrow's session. At any rate, map quadrangles are being digitized on a prioritized basis nationwide, and eventually digital map data will be available for all the U.S. In relation to all of these data, one of the major problems in USGS is that there is really no one place one can go to find out about all of these data.

At any rate, I would like to describe some of the examples of RASS data. The applications of RASS geochemical data to resource evaluation and related geological studies are so numerous as to make it impossible to fully describe. We can give some examples representative of these applications. In-house users are, of course, the largest group of users. Other examples of users are as follows: a state geologic survey requests all chemical analyses of rocks from their state for a compilation on chemistry, mineralogy, and physical properties of rocks; a large steel corporation requests multi-elemental rock-geochemical analyses for the states of Colorado and New Mexico for resource development and mineral resources exploration; a consulting geologist requests data on the molybdenum content of all Colorado soils, stream sediment, and rock samples for research purposes to support work for private industry clients; an exploration and mining company requests uranium data for igneous rocks and uranium data for metamorphic rocks for seven specific geographic areas of the U.S. by increasing and decreasing uranium abundance; and in a final example, the Department of Energy requests all geochemical data for all sedimentary rock samples analyzed for Oklahoma. In all of these examples, substantial cost in terms of labor and computer costs would have to be absorbed by USGS internally if RASS were not available to the public through a commercial network, NGSDC, or in the USGS library. In some cases, these needs could be met best by on-line interactive retrieval on the commercial network, in some cases by ordering the data tapes from NGSDC, and in others simply by going to the library, for example, for browsing through the geographic listings.

One of the major problems we face in data management is the lack of policy for funding the creation, maintenance, and storage/retrieval of these earth science data. USGS is not funded for data management and earth science repository functions such as maintenance of secondary files nor for handling earth science data requests. Often it is more difficult to attempt to recover costs than it's worth. Each research project is funded on its own scientific merits and often suffers from lack of funds for data management. At the same time, if public requests are answered in-house, most of the time it is at the expense of crucial in-house research project funds. Accordingly, there is a dilemma: how to have good data base administration and conform with the Freedom of Information Act while, at the same time, how to accomplish cost/beneficial research results. Perhaps the greatest difficulty is to convince USGS administrators to recognize that quality data services are just as important as survey research objectives.

There are organizations in the federal sector such as our own EROS Center that have been given specific delegation of authority to recover costs and fund their data base activity in serving the public. In fact, a computerized cost recovery system has been developed at EROS that works extremely well. In some cases, federal bureaus have even been able to partially fund their activity related to their data base creation and maintenance functions. The most successful marketable files have been primarily in the bibliographic data base area, but I believe large numeric data base organizations could likewise benefit from this approach as they become more marketable. If funds could be recovered sufficiently to help sustain not only in-house secondary data management costs but actual creation and maintenance of primary files, this would allow the federal establishment to do a much better job of numeric data management.

One of the other problems associated with making earth science data available to secondary users is related to the problem of how to convey the information that the research project scientist has, who knows the data best. How does this knowledge get transferred to the secondary users? This is the gatekeeper discussed earlier. It is difficult to "pick the brains" of the individual research scientist and carry that knowledge effectively in the data base. The quality of the data and the interpretive knowledge of the primary user are extremely important to a specific data base. Qualitative descriptive information on the data can be conveyed in the data base, but the interpretive knowledge which can only be gathered by interrogating the collector and interpreter is difficult, if not impossible, to carry in the data base. This is where the referral concept comes in handy. By referring the secondary user to the original collector, additional data can be individually gathered. Perhaps the future artificial intelligence and computer-aided instruction (CAI) technology will assist in addressing this problem. Also, perhaps numeric data base users could more closely tie their data to bibliographic files for conveying this information. In fact, I think in future we will see close ties between numeric and text-oriented data base services.

In summary, there are substantial amounts of earth science data being collected in digital form with the advanced computer data acquisition systems today. This has led to the significant concerns for better numeric data management of our earth science data bases. Policies are needed which assist in the solution of these data management problems. Federal funding is barely adequate to assist primary research project scientists in meeting their data management problems. Secondary data base users are at the mercy of outdated existing information system policies in the federal sector.

In all of these primary and secondary user needs, the subsidizing of creation and maintenance of these data bases for users through recovery of costs is one major policy consideration. Another policy consideration is regarding the estalishment of standards and guidelines for procurement and development of "information manipulating" tools such as hardware, software, and related services. These tools need to be made more compatible with one another and conducive to information sharing. It would help if we could just get some standardized sizes of holes in boots to match our boot laces.

A large geochemical numeric data base such as RASS can serve as a prototype for meeting different levels of users at their level of need and at their own level of available resources. If a casual user can't afford computer access or can't afford to buy the RASS tapes and run them on a computer, then that user can browse through the listings in the library. Conversely, if a user needs quick access and can afford on-line interactive computer use, then that user can go to the public commercial computer network and set up his or her own retrievals on-line. The third type of user, who may have in his or her organization a large computer system, can buy the RASS tapes and "spin" them on his or her own internal computer system.

Other problems faced by workers in numeric geological data exchange include how to convey descriptive and interpretive information clarifying the data. One way of improving this could be to tie numeric data more closely to the descriptive interpretive analysis of these data in their published bibliographic files. Having worked in both fields, I think there is a lot to learn from each. The combined group of geological data and information users must address these problems and push for more effective data management and information policies and procedures to improve data and information exchange to the mutual benefit of all in the earth science community.

As someone mentioned yesterday, we all have similar data management and information problems. As a suggested recommendation from these meetings, I would like to recommend that a task force on data management be formed to address these problems.

If I can be of any assistance to you on any of these data bases, I am here in Denver. Feel free to see me later, and, as time permits, I'll be glad to take those of you interested on a tour of the earthquake special purpose computing facility next door.

NATIONAL AND INTERNATIONAL DATA EXCHANGE:

PROBLEMS AND OPPORTUNITIES

Dr. Allen L. Clark

IIRD - International Institute for Resource Development, Belvederegasse 2, A-1040 Vienna, Austria

Abstract--National and international data exchange programs are rapidly expanding while at the same time facing a wide range of constraints which reduce the effectiveness of such programs. The primary constraints are those related to information access i.e. problems of economics, technology, and inadequate data, structural problems i.e. problems relating to the exchange partners and technical problems i.e. inherent in the data themselves.

Major areas for increased data exchange exist with respect to mineral and energy data bases, digital base maps, cooperative resource assessment programs and the development of new and standardized definitions for resources and reserves on a commodity by commodity basis. Future data exchange programs will benefit from the resolution of the many constraints and from the implementation of truly national and international data exchange programs which are multi-disciplinary in nature.

Introduction

The international geoscience community has a rapidly expanding need for both more quantitative and more disaggregated data for use as inputs to resource assessments, land use evaluation and classification, econometric models, national resource policy formulation and a large number of specialized scientific investigations. This need for data has resulted in a parallel activity, virtually a worldwide preoccupation, for the development of data bases and information systems to handle the data and support the required analysis. Indeed the latter activity is continuing to expand as a result of several factors including an every increasing improvement in the ability to generate data, extension of data requirements into multiple disciplines, data inputs from newly emerging nations and organizations and technological advances in data capture, data input, and telecommunications for data exchange.

There are probably few, if any, general statements that are applicable to the complete field of national and international data exchange. However it is worthwhile to review some present and past data exchange programs, particularly in terms of what and how much data were exchanged and how the constraints of the program were overcome.

Data Exchange Programs

A primary concern in most, if not all, ongoing and proposed data exchange programs is the fundamental and largely unanswered question as to the value of these programs to the respective participants. Any assessment of value is highly subjective, even when all of the parameters for evaluation are stated primarily because different individuals envision different uses for the exchanged data. In most cases programs which are related specifically to the exchange of basic scientific data are successful and have considerable potential for expansion without serious constraints. This success of scientific exchange programs can be attributed primarily to a simple matter of economics i.e., most scientific data has little or no economic value in the sense that its possession by one organization or individual doesn't provide any real competitive advantage. Such is not the case with respect to data, particularly pertaining to reserves, resources, technology and costs, which have economic value and are much less freely exchanged although they are a major element of many

proposed present day data exchange programs. It is to these areas of data exchange that I should like to address the following discussion; particularly to highlight the attributes of some successful data exchange programs which have overcome the above problems.

Petroleum Data System (PDS) - The petroleum data exchange program which led to the development of the PDS is highly illustrative of the procedure and data required to make such a program operational. The PDS was an outgrowth of a joint program between the U.S. Geological Survey and the University of Oklahoma to compile all available data on the oil and gas fields of the United States and Canada. Initially the PDS program functioned without the active support of industry; receiving and compiling its data from state oil and gas agencies, the American Association of Petroleum Geologists, U.S. Bureau of Mines and State and local geological agencies - Surveys. The development of the data file was successful for two main reasons: First, there was an adequate amount of publically available data which could be used in the initial development of the data file i.e., to reach a point where there was a "critical mass" of data making the file useful to a wider user community. Secondly, initial and continuing support was available through the U.S. Geological Survey for a sufficient period of time to insure that the data file could be built, maintained and used by the user community.

However, the PDS became even more successful with the addition in 1977 of the Petroleum Data System Advisory Committee composed partially of representatives of U.S. and Canadian petroleum companies. These corporate representatives through their companies have provided new industrial data and verification of existing data to improve the data file. Additionally the PDS now contains a data file of digital outlines for all U.S. oil and gas pools which was a direct outgrowth of the users needs expressed by this group.

The development and success of the PDS provides both a useful insight into the development of a data file by data exchange and also some basic aspects of the requirements relative to the maintenance of the data file. Some of the more important features are summarized as follows:

1. In the area of non-renewable resources there must be a large amount of publically available information outside the industrial sector upon which a file can be initiated.

2. Initial and continuing support are required to insure the data file will survive a period when its utility is limited by inadequate data.

3. Once a critical mass of data are available the file should be released to the public in order to develop a diverse and supportive user commodity.

4. Industry should be brought in as soon as possible to provide corporate data and verification.

5. File usage must be enhanced primarily through providing easy access, analytical software and graphics representation.

Even with the tremendous success of the PDS system it should be recognized that there are serious data deficiencies in respect to economic and engineering data which will continue because of the proprietary nature of the data. Regardless of these deficiencies the data file is rapidly developing and is unique in the amount of data contained with respect to reserves and resources; many of the figures provided and verified by industry.

International Strategic Minerals Inventory - Unlike the activities within the energy field there have been virtually no international programs for the exchange of mineral deposit data. The general lack of data exchange is primarily attributable to the following:

1. There is a general lack of specific data, within the published literature on many of the major mines of the world. Therefore any compilation necessarily needs to rely heavily upon industry whose participation to date has been minimal.

2. The mineral industry is much more fragmented than in the energy sector therefore fewer summary studies and statistics are available.

3. Virtually no data are available in the public domain pertaining to cost and other economic data related to a specific mine.

4. There has been an overall lack of emphasis on mineral deposit studies because of a general perceived sufficient supply of most commodities. As a result there is both less interest and less money expended with respect to mineral deposit data.

Given these factors it is much more difficult to construct a mineral deposits data file, certainly at the level of detail found in energy related files, and as a result the exchange of data will be inhibited: unless undertaken by a consortium of major international geoscience agencies.

In recognition of this fact and the ever increasing importance of mineral resources in the world economy the Geological Surveys of Canada, Federal Republic of Germany, United States and the U.S. Bureau of Mines have begun an "International Strategic Minerals Inventory". This program, which is to result in a publically available file, will compile data on the major producing mines, major developing mines and major but undeveloped deposits, in descending priority (Greenwood, W.R. 1982, personal communication). A prototype phase of the Inventory will concentrate on the commodities Ni, Cr, Mn and phosphate rock to test and develop procedures for data exchange and compilation. The proposed Inventory will emphasize data relating to detailed geology, reserves, resources, commodity information, ownership and production statistics.

Ideally the proposed Inventory will serve both as a basis and a catalyst for the development of a truly international mineral deposit data file.

The two examples I have discussed are intended to demonstrate some of the more general problems and also the potential associated with developing both national and international energy and mineral resource data bases. Clearly the situation is different for each sector and continues to vary down to the level of deposit data. Regardless of the individual problems there does appear to be some common approaches, as evidenced by the PDS and International Strategic Minerals Inventory, which can be applied and are successful. There are however several other problems of equal or greater significance associated with the overall field of data exchange.

Information Access: As evidenced by the attempts to establish a New International Information Order, the development of the OPECNA News Service for the developing nations, recent attempts at the Summit meeting in Cancun, Mexico to manipulate the press and the more general national and international programs to compile and exchange geoscience data, information access is rapidly evolving as one of the most sensitive and difficult problems of the 1980's. In attempting to address this problem however several political, technological, economic and basic data problems have emerged, which must be resolved before any significant progress can be made relative to information access. A detailed analysis of any one of these areas or problems is beyond the scope of this paper and indeed seems beyond the immediate scope of any number of special commissions and conferences that are presently being convened or that are working. There are however some primary constraints on information access and exchange between the developing and developed nations which are specific to geoscience data exchange.

Information Classification - Within both the developing and developed nations large volumes of data, indeed all data in some instances, are considered to be inappropriate for outside access or exchange. In the majority of cases this classification is a national political decision made in terms of an assessment of national self interest. Too often, within the scientific community, these decisions are viewed as being illogical, inappropriate and an impediment to the advancement of science. In many instances there may be some validity to these assertions and indeed specific cases exist where such a classification of data was not in the best interest of a nation. However, in any discussion of data access or exchange, the political realities of national policy should and must be taken as a basis for any proposed program; even though such considerations inhibit or preclude the proposed exchange. Indeed the challenge of the future is to reduce such restrictions by demonstrating the mutual advantage of data exchange.

Economic Constraints - The basis economic constraints on data access and exchange can be grossly divided into how much it costs to capture data that is publically available and what it costs to acquire data that is not publically available but which can be purchased. As stated earlier, in many instances this is the difference between scientific data and industrial data.

Scientific, or readily available data, has in most cases been acquired through the expenditure of national, international or organizational funds and is available at some small proportion of cost. Such data will normally have little or no commercial value without subsequent processing or analysis. The majority of available national and international data is of this type. Even though readily available the costs of acquiring such data may in many cases be beyond the means of many scientific investigators or developing nations. Indeed the rising costs of data collection, storage, retrieval and distribution is becoming a primary constraint on access even to scientific data.

Even more restricted economically is access to industrial data which unlike scientific data has a high corporate cost that normally must be recouped. Additionally such data may have a commercial value well in excess of the cost of collection and as a result is even more expensive. Access and exchange of such data may be either impossible, becuase it represents a competitive advantage by a corporation, or very costly when acquisition costs must be amortized and marginal utility paid for by the user. Therefore, the majority of such data are not available to the national or international user community: another reality which must be accommodated in any proposed data access and exchange program.

Technological Constraints - Ironically the very technology which is supposed to make geoscience data accessible and exchange possible in many cases serves as a major barrier, particularly in the case of developing nations. One is constantly reminded that computer technology, telecommunications and data bases put geoscience information within reach of anyone who has a telephone and a terminal. What many people fail to understand is that such access is not possible, for many reasons, in a very large number of developing nations nor is it "just around the corner". Technological exclusion is becoming one of the major problems in international data access and exchange and in many cases is widening, not reducing, the information gap between the developing and developed nations.

Data Insufficience - The 1960's were a time when large all inclusive data files were going to be created which would support the entire range of potential uses. This concept gave way in the 1970's to specialized data files, which would later be combined into a data base, constructed to resolve special problems. The hopes and activities of the 1960's and 1970's have given way to a realization in the 1980's that in general the data available in existing files and data bases is insufficient: in large part as a result of problems related to access. Therefore if the geoscience community is ever to undertake the analysis and assessments that it desires it must first solve the problem of data insufficiency by improving data access and exchange. This problem is inexorably linked to the constraints previously discussed and without the resolution of these constraints the problems of data exchange will continue and increase in complexity and intensity while the potential uses of such data are inhibited if not precluded.

Structural and Technical Problems: In addition to the access problems associated with data exchange there are a number of additional factors that materially influence whether a data exchange program will be entered into, or if it is even possible, on a national or international basis. These problems are divided into structural (dealing primarily with problems relating to the exchange partners) and technical, (dealing primarily with problems relating to the data itself).

The majority of structural problems are related to the variable requirements for data that an individual organization or nation has relative to its own needs. It should always be recognized that individual nations or organizations will participate in an exchange program when they foresee some tangible benefit to be received from participation. This is a factor that is too often overlooked when scientists discuss data exchange and too strongly emphasized when administrators or policy personnel discuss data exchange. Regardless it is clear that nations or well established organizations that have (a) a well established resource policy, (b) that have a high technology base for the collection and analysis of data, (c) and who have a major emphasis on mineral and energy imports and exports have both a higher interest and ability to participate in resource data exchanges. Conversely many nations and/or organizations have only a limited interest in data exchange programs for exactly the opposite reasons from those given above and therefore will not normally be active participants except in very restricted areas. Examples are Bolivia (Tin), Botswana (Diamonds), Jamaica (Bauxite), Turkey (Chromite) and Sri Lanka (Graphite) all of which rely very heavily on the export of their resources as a basis for their economy.

A second major structural problem in data exchange programs is that participation is impeded because of a lack of either a responsible geoscience organization or the lack of technical personnel to participate in the program. In large part this problem is restricted to the developing countries but also occurs in some developed nations which lack a designated responsible organization to handle such activities. In response to this problem many organizations are attempting, on a regional basis, to develop libraries of relevant resource data; the regional mineral resource development centers of Bandung, Indonesia and Dodoma, Tanzania are excellent examples for regional national cooperation. The designation of responsible organizations is a critical first step in developing the capability for many nations or regions to participate in larger international data exchange programs. Regrettably however in most cases these modest efforts are poorly funded and lack adequate technical support.

The technical problems of data exchange are perhaps best kwown and are those most discussed with respect to resource data exchanges. However, although the problems are well known there is still little or no national and international progress made toward resolving many of the more pressing problems including:

International Data Standards - The major technical problem in the exchange of existing resource data is the lack of standardized data definitions and standard measurement and reporting procedures. These problems are particularly acute in the energy and mineral sectors of the geosciences because of the variable occurrence modes of the individual resources, the broad geographic distribution of the resources and the all too often subjective or poorly quantified nature of the resource data. For these reasons international data standards have long been discussed but little progress has been made in developing such standards for even the broadest usage. Particularly illustrative of the other extreme of the problem is the definition of reserves and resources which was conceptually presented by McKelvey (1972, p. 32-40). This effort led to a modified classification by the Canadian Geological Survey in 1974 and an even more modified classification by the UN in 1978. As a result there exists not one but a number of "International Standards" for defining reserves and resources - the two most critical elements of most resource data bases and major items in resource data exchange programs.

The UNESCO sponsored IGCP program #98 "Computer Applications to Resource Studies" evaluated the problem of data standards for resource data files and was forced to conclude that available standards were too diverse to be generally applicable to the broad spectrum of data elements; data compatibility would then have to be provided by subsequent conversion to a local standard upon receipt of the file. Such a solution is by and large unacceptable but is clearly exemplary of the present status of data standards relative to much of present spatial data.

Incompatible Data Formats - Equally troublesome as the diversity or lack of international data standards is the even more diverse data formats by which data are collected and stored. In virutally every organization resource data (geologic, geochemical, geophysical etc.) are collected and stored utilizing a unique format which makes data exchange almost impossible in the worst cases and exceedingly costly and time consuming in even the best cases. Clearly there are many geochemical and geophysical applications in which a unique format is called for in the context of data use, requirements of a data base management system or to facilitate direct data reduction. There are however, many more applications for which standardized data formats would not only be possible but highly desirable, as in the case of field data collection, regional geochemical surveys and deposit evaluations. Unfortunately, both the need for standardization and the methodology for accomplishing it are only now being recognized in many geoscience organizations. Although these are fields which should be pursued actively by the geoscience community there is at present no mechanism by which this can be accomplished internationally, therefore, the problem persists and becomes greater each year.

Validity of Existing Data - Separate and apart from the problem of insufficient data, which may reflect either lack of collection or simply that the data are not available for a variety of other reasons, is the problem of validity of the data that does exist. Few if any problems relative to data exchange are as complex or more difficult to resolve than ascertaining the validity of the data within an exchange program. Too often data exchanges take place on the basis that "regardless of the validity of the data that is all there is and we must use it as it exists". Somehow that is supposed to resolve the problem of validity. Obviously this approach is not endorsed by many, but only some scientists. It does however present a major problem in the exchange of data in that all too often data are used by someone unaware of the validity problem and this can and does lead to major problems.

Many suggestions and practices have been made and/or implemented relative to providing a statement of the validity of the data contained in a file; ranging from the common "these data have not been reviewed or approved by ..." to a detailed probability statement of accuracy for each data item or, as is becoming more common, a listing of the individual to whom the data are attributable. All of these methods provide some general information relative to the validity of the data but fall short of the necessary standards and validation procedures required to assure the geoscience community of overall validity of the data.

Multi-disciplinary Data Requirements: A rapidly emerging need for data exchange is in the field of national and/or international resource assessments where the earlier purely geological assessments have given way to complex assessment methodologies requiring data on not only geology but on exploration activities, reserve accumulation, engineering, economics, environment, and policy decisions relative to resource development. A simplified example of such an assessment program is the Clark-Drew Conceptual Model of a Petroleum Supply System developed and implemented by the Author and Dr.L.J. Drew of the U.S. Geological Survey's Office of Resource Analysis to assess the long term resource supply of oil and gas for the United States (Clark, A.L., 1977, p. 221-233).

To implement a model such as the Clark/Drew model requires data inputs from at least 10 different disciplines associated with the petroleum industry. Such complex data demands pose several major problems with respect to data exchange; a detailed discussion is beyond the scope of this paper, however, several are worth summary comment as they represent problems which will become more multi-disciplinary. Among the most significant are:

1. The diverse data types rarely exist on the same machine or system therefore there is the initial problem of simply identifying the sources of the required data for exchange.

2. Requisite data must be brought onto a single system, or network, and placed under a common data base management system which is both costly and time consuming.

3. Sub-files must be created and made compatible with the analytical software and output devices of the host system requiring a high level of technical support and expertise.

4. Because of the diversity of data and sources considerable time and effort must be spent in data validation and updating.

5. Maintenance of the data system becomes increasingly expensive monetarily and in terms of machine resources necessitating constant monitoring to insure cost effective operations.

In the future providing either the required data or access to the data for multi-disciplinary studies will be a major challenge that must be met and resolved to insure that geoscience information continues to be a fundamental input into national and international resource development projects, program and policy.

Cooperative Opportunities in Data Exchange: Activities in data base development and data exchange, on both the national and international level, are experiencing the growing pains of rapid expansion for which there are few guidelines and even fewer successful programs that might be emulated. Regardless of the many problems, most of which will be solved with time and hard work, the entire field of data exchange can materially benefit the geoscience activities of most if not all nations. Among the most immediate cooperative opportunities for data exchange are the following:

Basic Mineral and Energy Data: large volumes of both energy and mineral data are either publically available or can be acquired on a bi-lateral basis. To facilitate this exchange an ongoing effort to catalogue available geoscience data should be undertaken. As a preliminary step the work of Burke (1981) in assessing international geoscience data bases should be expanded and continued as a major international effort. Additionally available resource data files should be placed on a common system to provide uniform access to the widest possible user community. Finally all such data should be made available at minimal cost, on microfiche or in hard copy, for those without access to the computer system.

Digital base maps - A common and largely unresolved program in the exchange of international or national digital data is the lack of standardized, readily available, low-cost digital base maps. Although a large number of available bases exist there are serious problems associated with their exchange and utilization. First, the majority of such packages are too comprehensive for the needs of most organizations. Secondly, most programs require a large and sophisticated computer system in order to handle both the file and the operating system. Third the majority of such programs are not easily adapted for use with existing data files of plotting equipment.

To resolve this problem a system of partitioned data files should be developed on a national or regional basis, which are compatible with plotting software programs usable on mini to micro computer systems. A central repository where such programs, and required technical expertise, would be available to an international user community would be a major step in resolving the needs of the developing nation, and many others, for digital base maps.

Cooperative Resource Data Programs: Although many constraints exist on international and national data exchange specific areas for international cooperation do exist and programs should be initiated, or existing ones extended, to undertake such cooperative efforts. In particular international data files on mineral deposits, regional geochemistry and geophysics, coal, geothermal, oil shale, phosphate and ocean resources are primary areas for cooperation. Such programs would be best undertaken through cooperative programs between the existing geoscience organizations of the developed and developing nations and made available through an impartial international agency.

Reserve/Resource Definitions - Considerable emphasis has been placed, and continues to be placed, on the definition of internationally acceptable reserve and resource definitions. What is lacking however are the definitions, relative to specific mineral and energy commodities, which are applicable to data exchange programs. Individual data exchange programs should undertake, as a primary objective of their program, the establishment of standardized reserve and resource definitions. Such definitions should be an integral part of the data collection and evaluation procedure and should accompany any file definition. To the extent possible such definitions should include, or conform where possible, to existing proposed definitions and not fall into the universal trap of developing a totally new set of definitions unless absolutely necessary.

References Cited

Burke, C.F., 1981, International Review of Geoscience Source and Reference Databases; Australian Mineral Foundation Seminar in Geoscience Numeric and Bibliographic Data, Adelaide, Australia 20 p.

Clark, A.L., 1977, Guidelines for International Data Collection and Resource Assessments and Analysis, Journal of the International Association for Mathematical Geology, Vol. 9, No. 3, p. 221-233.

McKelvey, V.E., 1972, Mineral Resource Estimates and Public Policy: American Scientist, V. 60, p. 32-40.

Paper Unavailable--

AN OVERVIEW OF THE GEOLOGICAL APPLICATIONS OF DIGITAL DATA BASES

BATTEN,[1] Lawrence G., JENSON,[1] Susan K., HASTINGS,[1] D. A., GREENLEE,[2] D. D., and TRAUTWEIN,[1] C. M.

The U.S. Geological Survey's Earth Resources Observation Systems (EROS) Data Center has been actively engaged since 1979 in research on and development of spatial data handling techniques applied to geoscience data. Currently, digital geologic data bases are being evaluated for their applicability to mineral, and hydrogeologic resource assessment requirements of the U.S. Department of the Interior.

Initial investigations to determine the feasibility and utility of the geologic data base concept for mineral resource assessment involved the design, implementation, manipulation, and management of data bases, and the development of geologic models. A case study in the Interior Porphyry Copper Belt of Alaska incorporated Landsat multispectral scanner data, digital topographic data, and regional geochemical, geophysical, and geological data from the Alaska Minerals Resource Appraisal Program (AMRAP) folio of the Nabesna Quadrangle in Alaska. Geologic modeling of porphyry copper potential was accomplished through a series of interactive analyses that included both arithmetic and statistical integration of multivariate data within the data base. Of the 16 original data types and 18 derived data sets within the data base, 10 parameters were identified as being descriptors of porphyry copper potential within the area. These parameters were quantified with respect to the range of model values, weighted according to their regional significance, and tested within the data base. The grid-cell modeling technique identified three porphyry copper prospects, two of which are known occurrences.

Applications of digital data bases to hydrologic resource assessment are directed at automated procedures for defining drainage basin morphology and hydrogeology. A case study in the Black Hills of South Dakota and Wyoming has incorporated Landsat data, digital topographic data, geologic data, and hydrologic data for basin analysis. Basin morphology has been described in terms of surface area, slope, and topographic aspect using Digital Elevation Model (DEM) data. Basin hydrogeology is modeled through the interactive analysis of lineament density, length and slope of drainageways, land cover conditions, geology, and morphologic parameters (i.e., basin shape, relief and surface area). Correlations between adjacent basins will be used to establish subsurface relationships that influence the hydrologic regime of the Black Hills.

Geological applications of digital data bases require that both image processing and statistical analysis techniques be applied interactively to multivariate geoscience data. The development of a geoscience model is a decisionmaking process in which the scientist converts separate, quantitative data sets into spatially defined, integrated results. The opportunities available to the scientist working with digital data bases are numerous, particularly in efficiently utilizing large amounts of regional geologic, geophysical, geochemical, topographic, remotely sensed, and hydrologic data to establish resource criteria and potential.

---

[1]Work performed under U.S. Geological Survey Contract Number 14-08-0001-20129, Technicolor Graphic Services, Inc., EROS Data Center, Sioux Falls, South Dakota.

[2]USGS/EROS Data Center, Sioux Falls, South Dakota.

The Geoscience Applications Section of the Earth Resources Observation Systems (EROS) Data Section is involved in research on the uses of remotely sensed data for geologic, hydrologic and mineral resource applications. It is currently investigating the utilization of comprehensive spatial data bases in the geo-sciences. Developed techniques are presented via short courses offered at EROS and abroad and through cooperative projects with various governmental agencies. In these investigations, specialties include hydrogeology (Batten), geophysics (Hastings), geochemistry (Trautwein) and digital analysis techniques (Jenson and Greenlee[1]).

[1] Greenlee is a member of the Data Analysis Section of the EROS Data Center.