DOCUMENT RESUME

ED 171 239 IR 007 147

AUTHOR Veith, Richard H.
TITLE The Internationality of Information Banks--Online Bibliographic Databases.
PUB DATE 78
NOTE 34p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Bibliographies; *Data Bases; *Foreign Countries; Global Approach; Information Networks; Information Systems; *Information Utilization; *On Line Systems
IDENTIFIERS *Access (Information); Scientific and Technical Information

ABSTRACT

The international character of online bibliographic databases is assessed by addressing three major aspects: literature production, article selection, and database use, as well as a number of side issues such as distribution and language. Requests for information were sent to producers and suppliers, most of whom responded, with no attempt being made to be comprehensive in this initial survey. Selected databases were expected to vary widely on several dimensions. While 90 to 95 percent of the world's scientific and technical journal and report literature are estimated to be accessible by online terminals, it appears that the literature itself is not as accessible, the databases of citations are not accessible in a feasible manner in most countries, the citations are not in a language universally understood, and the quality of the indexing and abstracting in relation to the information content of the original articles often goes unexamined. (JEG)

THE INTERNATIONALITY OF INFORMATION BANKS -- ONLINE BIBLIOGRAPHIC DATABASES

by

Richard H. Veith

THE INTERNATIONALITY OF INFORMATION BANKS --ONLINE BIBLIOGRAPHIC DATABASES

Richard H. Veith
School of Information Studies
Syracuse University

ternational information sharing has, during the past dozen years, come to encompass the use of instantly accessible, globally available public databases of facts, figures, literature citation and informational items of sundry kinds. Of total information transfer activity, the use of such databases is assuredly still a very small part. However, the recourse to databases is substantially growing, as computer files are dramatically expanding and as telecommunications facilities specifically for data traffic are steadily expanding. Therefore, an attempt was made to assess the international character of online bibliographic databases.

The question of the internationality of the databases is considerable, ranging from authorship of the original material through the computerization of selected items to electronic access. A number of directories have been published (21, 22, 25, 33) and articles written giving some statistics (6, 7, 29, 34) but these do not cover the breadth of the question, partly because of the newness of database activity and partly because commercial organizations involved are reluctant to surrender internal statistics.

The question is also part of a larger one involving international data traffic in general. There is presently debate in many countries over the facts and fears of transborder data flows, which includes all forms of

computer-readable data moving across national boundaries. Both national and international committees and congresses have recently met or will meet to establish broad policies. Some spokesmen are even calling for enforceable international law as well. The Council of Europe, the European Economic Community, the European Parliament, the Nordic Council, the Intergovernmental Bureau for Informatics, UNESCO, the World Administrative Radio Conference in 1979, a United Nations Conference on Science and Technology, the International Commission to Study the Problems of Communication, the U.N. General Assembly, conferences sponsored by the Organisation for Economic Co-Operation and Development, and meetings of the nonaligned nations are all scheduled to produce (if they have not already done so) transborder data flow policies within the next year or two.

National laws affecting international data transfer have been enacted in Sweden, West Germany and France, with a law pending in Austria and restrictions on transborder data flows in effect in Australia, Canada and Japan.

Much of the concern stems from questions of personal privacy, economic strength, national sovereignty and, in the broadest form, cultural domination. Although the international traffic in bibliographic databases is minute compared to other types of data flows, it is directly concerned with the general question of cultural domination. Is it possible to determine what effects, if any, computer networks of databases are having on existing patterns of international information sharing? Do online databases merely reflect the literature, for example, in terms of authorship, publication or availability? Or, perhaps unintended, have bibliographic databases become the repositories of citations biased toward the country doing the selecting and indexing? Are we able to say now (or ever) whether the advent of informational databases internationally available marks a recognizable improvement in international

sharing?

In order to at least begin to answer these questions, a preliminary assessment was conducted of a small number of databases. Requests for information were sent to database producers and suppliers, virtually all of whom responded. (See "References" for a list of the organizations who supplied data.) No attempt was made to be comprehensive in this initial survey, but an attempt was made to choose databases that could be expected to vary widely on several dimensions. With few exceptions, databases in the U.S. were chosen both because of the position of the U.S. in the database business and because some of the statistics would be more readily obtainable.

The assessment addressed three major aspects, namely, literature production, article selection, and database use, and a number of side issues such as means of distribution and language.

Literature Production

In 1977, some 181 countries were publishing 90,000 serial titles (26). Of that number, approximately 9500 were published in the U.S. (1). Another source, covering only 80 countries and excluding the U.S., puts the number of periodicals worldwide at about 98,000 (31). On a very gross level, then, one can say that the U.S. produces one-tenth of the world's serial publications. Europe, on the other hand, as indicated in Fig. 1, produces over half of the total periodical publication. The statistics for book production are similar, with the share for the U.S. a little higher and the share for Europe a little lower. (See Fig. 1 for comparison of book and periodical production in the seven major regions of the world.)

However, place of publication does not necessarily mean that all of

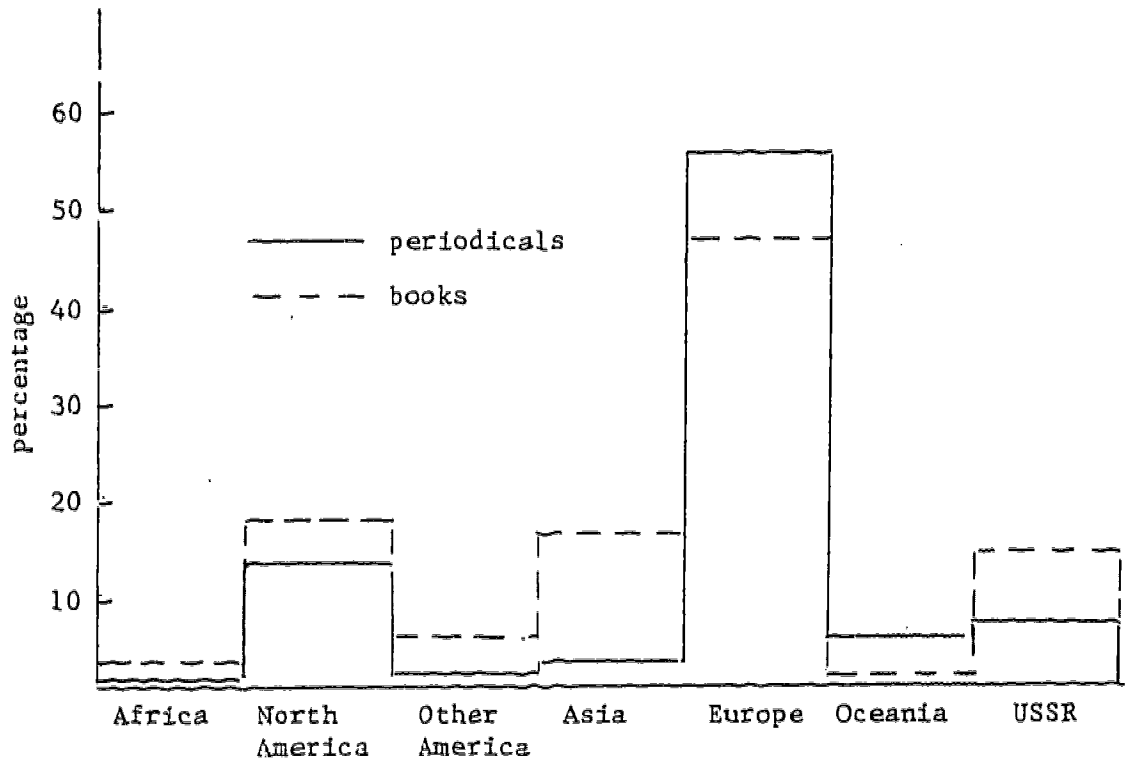

Figure 1: World periodical and book production. Totals for periodicals are based on statistics from years 1972 through 1977; totals for books are based on 1975.

Sources: Ayer (1), Bowker (26), U.N. Yearbook (31).

the articles in the journals or magazines were authored by that country's citizens. In the sciences, for example, there is a tradition of international cooperation. Nearly 30 percent of the science articles published in the U.S. in 1976 were by foreign authors (19). Similarly, 20 percent of U.S. authors in the sciences published outside the U.S. in 1976. The percentages for some of the other countries of publication for U.S. authors are listed in Table 1. Looking at the total world science literature, it is estimated that the U.S. produces some 38 percent--and this figure may be too high since the journals canvassed were those listed in a U.S.-produced database (19). The extent of the international availability of science literature is perhaps underscored by the fact that almost half--42 percent--of the references listed by U.S. authors were to foreign publications.

One might expect more international activity in the sciences than in other serial literature. Therefore, since almost three-fourths of the science articles published in the U.S. are by U.S. authors, the figure is probably higher for serial literature in general. And one might expect that the U.S. share of world science literature is greater than for other types of literature. Thus the U.S. share of all journal literature ranges from about 10 percent to 38 percent at most. It seems, then, that databases heavily filled with citations to U.S. publications are certainly not representative of the available literature.

Article Selection

Although certain databases are advertised as covering the world's literature, selection takes place not only on a journal basis but also on an article-by-article basis. According to one set of estimates, as little as 3 percent of the relevant articles in a field have been cited in a data-

Table 1: U.S. Share of Science Literature

	percent of total, by authors from:	percent of articles in U.S. journals from:	percent of U.S. authors published in:
U.S.	38%	73%	80%
U.K.	-	4	8
West Germany	-	2	2
France	-	2	<0.5
U.S.S.R.	-	1	<0.5
Japan	-	3	<0.5
Canada	-	5	1
all others	62	12	9

Based on journal literature during 1976 from the following fields: clinical medicine; biomedical research; biology; chemistry; physics; earth and space sciences; engineering and technology; psychology; mathematics.

Source: Science Indicators 1976 (19).

base (22). At the other end of the scale, as shown in Table 2, the estimates range up to over 600 percent, which suggests that there may be something wrong with the assumptions or the estimations. For one thing, Pratt (22) only estimated journal articles, while Chemical Abstracts, for example, includes citations to books, conference proceedings, dissertations and technical reports as well. In fact, journal articles accounted for only 69 percent of the Chemical Abstracts files in 1975 (2). It is also possible that Pratt's estimate of the number of articles per year per journal was too low, or his estimate of the relevant journals was too restrictive.

An earlier study of abstracting services themselves found that anyone attempting full coverage of a subject by using abstracts journals will certainly fail (17). In that study, nearly 30 percent of the desirable references for given topics were not contained in the abstracts journals, while 46 percent of the references were covered more than once. It is probably still true that in most fields journals of citations and databases of citations are considerably less than 100 percent representative of the field's literature.

It might be useful to know, then, even if the bulk of the world's literature is in fact monitored or scanned, whether the process of scanning and selecting is performed in one country or in a number of countries and whether or not the location of the activity has an effect on what is selected.

For the databases in this review, the scanning and selecting of articles to cite or abstract is primarily performed in one country, the U.S. (see Table 3). There are, however, two notable exceptions on the list which merit further description: INIS (International Nuclear Information System), and Index Medicus (or MEDLARS).

The production of the INIS database is truly an international effort,

Table 2: Articles in a Database as a Percentage of Total Articles in the Field

database	percentage
ISI (public administration)	3%
International Labour Doc. (labor)	4
ISI (education)	7
ISI (geography)	8
ISI (economics)	10
ISI (demography)	10
ISI (law)	11
STIS (textiles)	13
ISI (meteorology)	14
Economic Titles (economics)	14
ISI (social science)	27
Textiles Tech. Digest (textiles)	30
ISI (mathematics)	36
F&S Index (business)	38
Petroleum Abstracts (petroleum)	43
Mathematics Titles Service (mathematics)	52
APILIT (petroleum)	70
TITUS (textiles)	70
Food Science and Technology Abstracts	70
MEDLARS (medicine)	74
CAIN (agriculture)	74
METADEX (metals)	79
Excerpta Medica (medicine)	91
GEO-REF (earth sciences)	95
Psychological Abstracts (psychology)	109
SPIN (physics)	125
Toxicon (environment)	179
BIOSIS (biology)	256
Mechanical Engineering (mechanical engineering)	444
Nuclear Science Abstracts (nucleonics)	453
Chemical Abstracts Condensates (chemistry)	656

Based on Ulrich's subject categories, and assuming sixty articles per year per journal. Databases selected from those available in Europe, 1976.

Source: Information Economics (22).

Table 3: Number of Countries Selecting Articles, and Number of Countries Represented

database	number of countries where scanning/indexing performed	number of countries whose publications are scanned
INIS	49	?
Index Medicus	12	76
Chemical Abstracts Service	5*	150
AGRICOLA	2**	101
BIOSIS	1 (U.S.)	116
INSPEC	1 (U.K.)	72
Language and Language Behavior Abstracts	1 (U.S.)	61
ISI	1 (U.S.)	60
Sociological Abstracts	1 (U.S.)	60?
COMPENDEX	1 (U.S.)	34
Oceanic Abstracts	1 (U.S.)	24
Pollution Abstracts	1 (U.S.)	18
ABI/Inform	1 (U.S.)	4
Pharmaceutical News Index	1 (U.S.)	1

Sources: CAS Catalog (3), INIS Today (15), National Library of Medicine (18), INSPEC Catalogue (14), National Agricultural Library (32), and correspondence.

*The U.S. is the primary country; 5 percent is done by the United Kingdom Chemical Information Service Input Center, and a lesser amount by volunteer groups in Japan, Poland and Czechoslovakia.

**The U.S. is the primary country; Agriculture Canada produces input for a small subset covering Canadian cooperatives.

coordinated and operated by the International Atomic Energy Agency, Vienna. The decentralized approach to producing input was chosen not only to spread the cost of data gathering and to help improve national information infrastructures, but also to provide "the most comprehensive coverage of the nuclear literature" (15). A map of the member countries of INIS shows some two-thirds of the land mass of the world (49 countries) as part of the INIS system, with the major exceptions being the People's Republic of China, some Central American countries, and much of the African continent. There is more to the international production of INIS, though, than the mere fact that 49 countries provide input. Organizations in some countries serve as centers for nearby countries, or provide technical services for other centers. The INSPEC organization in Great Britain, for example, produces the machine-readable version of "something like 10 percent" of the total INIS database by receiving input from Scandinavian, Dutch, Austrian and Swiss centers scanning journal literature (13). For the non-conventional literature, though, some of those centers prepare the machine-readable versions themselves.

The success of the decentralized input process for INIS has encouraged the U.N. Food and Agriculture Organization to use the same procedures for its Agricultural Science and Technology Information System (AGRIS). Another system on the INIS model, under consideration, is the Development Sciences Information System (DEVSIS). (Preparations for online access to INIS were only begun in 1977.)

The Index Medicus, which is available online as MEDLINE and a number of related databases, is produced by the U.S. National Library of Medicine using input from the U.S. and eleven other countries, the latter under quid pro quo agreements. Great Britain, Germany and Sweden, for example, provide their own inhouse indexing. France and Brazil both use inhouse and

contract indexing services. And Australia, Iran, Mexico and South Africa all use contract indexing exclusively as their form of quid pro quo payment. The National Library of Medicine system, which includes fifteen separate databases online and several more available offline, is the largest single system in the U.S. and is considered to be the most heavily used online database in existence. Nearly 900,000 searches were performed on the medical databases as a group during fiscal year 1977, both online and offline (18).

The number of countries providing input, or even the number of countries whose journals are scanned, does not tell the full story, however. Even though twelve countries provide input into the National Library of Medicine system, for example, the foreign centers combined provide only slightly more than half the total number of citations. Table 4 shows that the U.S. share of the citations in various databases may range from about 20 percent to 90 percent or more for databases supposedly concerned with more than U.S. literature. For BIOSIS, the U.S. share is 22 percent, while the shares for the other 115 countries represented in BIOSIS range from about 8 percent for the U.S.S.R. to a low of 0.01 percent for eighteen countries including El Salvador, the Central African Republic, Chad, Guam, Hong Kong, Liberia, Liechtenstein, Nepal, Sierra Leone and Vatican City. For Chemical Abstracts databases, the U.S. share (26 percent) is almost matched by the Russian share of 23 percent. Inform, on the other hand, a business literature database, is about 90 percent U.S. references.

Database Use

The production and use of databases is definitely not a world-encompassing activity. The compilation of databases by Williams and Rouse (33) covers only databases produced in the U.S., Canada and Western Europe (with one exception in Japan). The compilation by Pratt (21) covers only databases

Table 4: Percentage of Database by Country of Publications Cited

	BIOSIS	Chemical Abstracts	Index Medicus	COMPENDEX	INIS	Pollution Abstracts	ABI/Inform
U.S.	22.14%	25.9%	48%	53% (1974)	53% (1974)	60%	90%
all others	9.11	8.2	52	47	47	40	10
U.S.S.R.	8.25	23.2					
U.K.	6.94	6.2					
Japan	6.54	8.2					
West Germany	4.95	6.0					
France	4.47	4.2					
Italy	3.32	1.9					
India	2.96	3.0					
Switzerland	2.78	1.0					
Brazil	2.54	-					
Netherlands	2.40	1.2					
Australia	2.18	1.2					
Canada	2.16	2.7					
Poland	1.97	2.1					
Czechoslovakia	1.71	1.4					
Spain	1.47	-					
Sweden	1.31	1.1					
East Germany	1.23	1.2					
Denmark	1.19	-					
Belgium	1.12	-					
South Africa	1.05	-					
Yugoslavia	.97	-					
Argentina	.94	-					
Romania	.93	.5					
Finland	.78	-					
Hungary	.77	.8					
New Zealand	.71	-					
Mexico	.68	-					
Austria	.66	-					
Norway	.66	-					
Portugal	.61	-					
Venezuela	.50	-					

Sources: Tomberg (28), and correspondence.

available in Europe. But both directories imply that there is not much else to cover. (Both directories list a little over 300 individual bibliographic databases, but these are not necessarily the same 300; some 142 of the databases listed in Pratt were not listed in Williams--or the databases were listed so differently as to appear to be different. Table 5 gives a breakdown of the 142 databases by country.)

Part of the picture of the availability of the databases can be seen from the viewpoint of where the databases are produced. Nearly two-thirds of all databases are produced in the U.S., containing 85 percent of all bibliographic records available online, according to one source (34). Using somewhat older figures, Table 6 shows 56 percent of all database producers as U.S. companies while, for databases available in Europe, only one-fourth of the originators are U.S. companies. The next two largest countries of origin are the U.K. and France, or the U.K. and West Germany, depending upon the source.

A more recent account puts the "world's output" at just over 400 databases of bibliographic information and 250 data banks of non-bibliographic material, almost exclusively in North America and Western Europe (28).

Looking at database use in terms of searches, most searches of bibliographic databases take place in the United States and Canada by a factor of twelve-to-one (see Table 7). For 1976, 68 percent of all online searches were in the U.S., 19 percent were in Canada and 13 percent were in Europe (28). Use, then, is overwhelmingly North American.

To give a more specific look at a few of the users, Table 8 presents some percentages of use by U.S. and non-U.S. customers. The National Library of Medicine states that 11 percent of their total number of access codes are used by foreign online centers, which, of course, indicates users, not actual use. BIOSIS, at the other end, estimates that 50 percent of the online use is

Table 5: Number of Databases Listed in European Directory But Not Listed in U.S. Directory, by Country

country	number of databases
Germany	41
U.K.	38
France	31
U.S.	18
Sweden	3
Netherlands	3
Denmark	1
Switzerland	1
Belgium	1
Austria	1
Czechoslovakia	1
Canada	1
Japan	1
Norway	1
total	142

Compiled from Pratt (21), Williams and Rouse (33).

Table 6: Database Producers and Originators, Numbers and Percentages by Country

	Williams & Rouse (producers)		Pratt (originators/operators)	
	number	%	number	%
U.S.	158	56%	77	27%
U.K.	43	15	60	21
France	38	13	33	12
Norway	9	3	3	1
West Germany	6	2	53	19
Canada	7	2	3	1
Netherlands	5	2	14	5
Sweden	4	2	7	2
Switzerland	3	1	5	2
Denmark	2	1	6	2
Italy	2	1	5	2
Belgium	1		9	3
Austria	1		3	1
Czechoslovakia	-		2	1
Poland	-		2	1
Luxembourg	1		1	
Australia	1		-	
Japan	1		1	
Spain	-		1	
Finland	-		1	
totals:	282	98+%	286	100%

Compiled from Pratt (21), Williams and Rouse (33).

Table 7: Number of Searches of Online Files

	1972	1973	1974	1975	1976	1977
			(in thousands)			
Europe--						
bibliographic	10	15	45	70	100	-
North America--						
bibliographic	-	-	700	1000	1200	2000
Europe--						
data banks*	50	220	550	1600	2200	-

*Data banks are non-bibliographic files. Although the online use in Europe of data banks appears to rise dramatically, most of the use can be attributed to three operators (28).

Sources: Pratt (22), Tomberg (28), Williams (34).

from outside the U.S. Sociological Abstracts, and Language and Language Behavior Abstracts, both were used about 25 percent of the time by non-U.S. customers, on the basis of actual figures for two months in 1978. Of the foreign users of Sociological Abstracts online, 11.5 percent were Canadian; of the foreign users of Language Abstracts online, 10 percent were Canadian. As the first three databases of Table 8 indicate, the non-U.S. use was spread over a number of countries.

Although it is apparent that U.S. users use U.S. databases, do Europeans largely use European databases? Perhaps not. Tomberg suggests that "the large part of what growth there has been in recent years (in European use) has come from the increased use of U.S. services rather than European services" (28).

Japan

The focus on North America and Western Europe does not mean that there is nothing available in other parts of the world. The few examples on Table 8 indicate databases used in Australia, Brazil, Iran, Japan, Mexico, the Philippines and South Africa. Japan is probably a good example of an active database country often overlooked. Within Japan, there are a number of public and semi-public databases internally produced, a few examples of which are listed in Table 9. Japan also provides input for Chemical Abstracts, MEDLINE, INPADOC and INIS as well as using those databases and others including MARC, Science Citation Index and RINGDOC. Japanese centers also use databases such as the Protein Structure Database (Brookhaven National Institute) and the Astronomical Information Database (Strasbourg Astronomical Information Center) as well as being the major source of international databases such as one on active volcanoes.

Table 8: Percentage of Database Usage, U.S. and non-U.S.

(Online and Tape)

	National Library of Medicine*	Sociological Abstracts**	Language and Language Behavior Abstracts**	Pollution Abstracts, Oceanic Abstracts, and ABI/Inform***	BIOSIS***
U.S.	89%	76%	76.5%	75%	50%
non-U.S.	11	24	23.5	25	50
countries of use:					
Australia	x	x		not available	not available
Belgium		x			
Brazil	x				
Canada	x	x	x		
Denmark		x	x		
Finland		x			
France	x	x	x		
Iran	x				
Italy			x		
Japan	x				
Mexico	x	x	x		
Netherlands		x	x		
Norway		x			
Philippines		x			
South Africa	x				
Spain		x	x		
Sweden	x	x	x		
Switzerland		x	x		
U.K.	x	x	x		
West Germany	x	x	x		

*Percentage of foreign use based on access codes for online use.

**Based on figures for January-February 1978.

***Estimates of general use.

Sources: correspondence.

Table 9: Examples of Japanese Databases

Science and Engineering Document File (Japan Information Center of Science and Technology)
Developing Nations Investment Law Text Data (Institute of Developing Economies)
Overseas Statistics Catalog--STAT Data (Institute of Developing Economies)
Environmental Information Data (National Institute for Environmental Studies)
Japan Comprehensive Economics File (Nihon Keizai Shimbun, Ltd.)
IBJ Data (Industrial Bank of Japan)
Corporate Credit Information--COSMOS (Teikoku Koshinsho Ltd.)
Japan Patent Retrieval System (Japan Patent Information Center)
NEEDS-IR (news articles) (Nihon Keizai Shimbun Ltd.)

Source: Japan Information Processing Development Center (23).

Databases of the Nomura Research Institute, the Industrial Bank of Japan, the Teikoku Kohinsho, Ltd., and Tokyo Shoko Research, Ltd., are available internationally on the General Electric Mark III network. Nihon Keizai Shimbun exports NEEDS-IR, a news file, to the U.S., and the Teikoku Koshinsho, Ltd., is preparing to market COSMOS, corporate credit information, in the U.S.

As another measure of online activity in Japan, at least fifty international computer communications vendors operate time-sharing services in Japan using 170 leased circuits (27).

Countries other than Japan, and outside North America and Western Europe, have equally been neglected. Without a doubt, more information is needed on databases and their use in, for instance, Eastern Europe, Australia and Mexico.

Distribution

There are a variety of means of having online files available internationally other than dialing an international telephone call. A Japanese company, Maruzen, will accept search requests at their Japanese offices for any files on Lockheed Information System's DIALOG; the Japanese office then telexes the requests to Maruzen's New York office where the online searches are performed.

A much simpler arrangement, though, is one in which tapes are transported so that the database may be located and loaded locally.

The massive database of Scientific and Technical Aerospace Reports (STAR) produced by the National Aeronautics and Space Adminstration, for example, is not available online outside the U.S. (to the database in the

U.S., that is). It is, of course, very much available online in Europe because the bibliographic tapes are provided to the ESRIN Space Documentation Service of the European Space Agency. At the end of 1977, about 160 organizations in fourteen European countries were using the NASA file online, for a total of 1800 hours of use. The negotiations required to establish this arrangement do not seem to foster complete worldwide use, however, since no such agreements exist with governments or organizations in the Far East; thus neither Australia nor Japan have online access to STAR. (Interestingly, NASA is one of the chief generators of the data flow between the U.S. and Australia, along with the two governments and the major transnational companies operating in Australia.)

As another example of transporting tapes, Chemical Abstracts licenses approximately 40 organizations to use their computer tapes. Seventeen of these processors are in the U.S.; the remaining 23 processors are in 20 countries including Argentina, Australia, Israel, Japan, Korea, Mexico and South Africa. Of the 40 organizations, however, only six provide online access. AGRICOLA, produced by the National Agricultural Library, similarly provides tapes to 20 organizations including centers in Sweden, Australia and Brazil.

In some instances, the international arrangements for providing files in computer-readable form include the sharing of production costs and the developing of ancillary products from the database tapes. The Internationale Dokumentationsgesellschaft für Chemie (IDC) in West Germany contributes a share of the production cost of Chemical Abstracts equal to the percentage of documents originating in West Germany--about 7 percent. The IDC then incorporates Chemical Abstracts data into their own computer system. In

Japan, the Japan Association for International Chemical Information not only provides input for and markets Chemical Abstracts files but also develops their own publications and services based on the Chemical Abstracts databases.

Yet the transportation of tapes does not in fact seem to have been the basis for increased use of databases online. Instead, a more direct contributor to growth has been the expansion of data networks and particularly packet switched networks.

Packet networks, while not the only way to transmit data, are reportedly the most efficient and cost effective means over a wide range of traffic loads (20). The packet networks divide data streams into short groups or packets, add error-checking information, and use node computers to route the packets over leased wideband trunks to other node computers where error checking is performed, faulty transmissions requested again, and packets reassembled. Thus, the conventional error-contributing telephone switching systems are avoided, and wideband channels are used where noise problems are relatively minor.

Telenet Communications Corporation, the first U.S. packet switched network, operates about 190 nodes in North America, Western Europe, Puerto Rico and Hawaii, and interfaces with Trans-Canada Telephone's Datapac and Mexico's Teleinformatica de Mexico. Tymnet, Inc., another U.S. packet network, provides similar service including a link with Italcable, Italy's international record carrier. Both packet switched and circuit switched networks are established or being established across Europe. EURONET, for example, is a transnational packet network, while the Nordic Data Network and the West German data network are both circuit switched. Japan's international packet network, when fully operational, may be one of the most comprehensive data networks anywhere. And reportedly the International Institute for Applied

Systems Analysis, Vienna, is coordinating a packet network into Eastern Europe with nodes in Budapest, Vienna, Pisa, Bratislava, Moscow, Kratowice, Kiev, Prague, Sofia and Berlin.

The rise of data networks are presumed to encourage greater international use of databases at lower cost. The formal establishment of the connection between Telenet and Canada's Datapac, since February 1978, is expected to provide at least one user with an 85 percent reduction in communications costs, and to perhaps draw some U.S. customers to Canadian systems and databases (5). The reasoning behind the latter thought is that some files, available both in the U.S. and Canada, would be less expensive to use on the Canadian system even for U.S. users, and other files in Canada simply are not on U.S. systems. Similarly, a connection between Tymnet and the U.K.'s Datel 200 service is expected to produce a marked savings for U.K. users in comparison to previous arrangements, and to encourage U.S. researchers to use British databases (30).

Although not exclusively so, the establishment of data networks seems to be a major precondition of increased use of online databases.

Language

Finally, there is the persistent problem of language barriers. Databases in English, no matter how representative of world literature and no matter how retrievable worldwide, are still unintelligible to millions of people. A 1976 survey of 337 bibliographic databases found that 76 percent were in English, 10 percent had only German as a carrier language, 10 percent had only French as a carrier language, and 4 percent were in Russian, Japanese, Czech or Swedish (12).

There are a number of databases that are multilingual, according to the

same survey. The TITUS-textile database (Institut textile de France) is in English, French, German and Spanish. The IRRD database (International Road Research Documentation) is in French, English and German. SABIR-Cancernet (Institut Gustave Roussy, France) is in French, English, German, Czech, Polish and Spanish. And CREDOC (Centrum voor Rechtsdocumentatie, Belgium) is in Dutch and French. Some databases are in effect translations of others; for example, there is a German language version of the MESH Vocabulary, a National Library of Medicine database. Other databases are only partially in several languages. PASCAL, produced by France's Centre National de la Recherche Scientifique, contains parts in English as well as French; half of all queries to the PASCAL file, though, are from France, and another 20 percent are from Morocco.

Until automatic translation is perfected, or economic reasons for providing translations arise, there is little to alter the fact that "databases made in France and Germany are almost invariably made in French and German. Databases made elsewhere are almost invariably made in English" (21).

Discussion

Although it is a little sobering to be told that probably half of all international data flow is the mundane matter of airline reservations and related messages (9), there is still reason to ponder the international characteristics of information databases.

Even though Summit estimates that 90 to 95 percent of the world's scientific and technical journal and report literature is accessible by online terminals (6), the literature itself is not so accessible, the databases of citations are not accessible in a feasible manner in the majority of the world's countries, the citations are not in a language universally

understood, and there is reason to believe that his estimates are too generous. Moreover, the matter of the quality of the indexing and abstracting in relation to the information content of the original articles often goes unexamined.

Even when it is possible to retrieve citations internationally, is it desirable, when the citations may have in fact originated much closer to home? As one commentator writes:

> The French supply data in machine readable form on medical, nuclear, space, agricultural, oil and other topics. This gives the U.S. a monopoly of storage of such data. But on top of this, the French have to pay for access, they are dependent on U.S. firms for keeping the "tap turned on" and they expose themselves to having their questions analyzed . . . (10)

One would reasonably expect some sort of quid pro quo arrangement (and there is, in the case of medical data to the National Library of Medicine at least) for such international cooperation in input.

The implication in this case is that France is doing the work of input and the U.S. is the major beneficiary, since it is much easier and less expensive for U.S. customers to access the data than for French customers. It is not clear, though, that U.S. databases really contain large amounts of data from foreign sources. In fact, most indexing and abstracting done to produce U.S. databases is probably performed in the the U.S., to a great degree.

However, this quantity of database production in the U.S. presents several other dangers. If the citations are predominately to U.S. publications, the online data bases may be deceiving U.S. users as to world output. At the very least, the U.S. databases, when overstocked with U.S. material, may be doing little or nothing to encourage international information transfer. And if non-U.S. users begin to rely on U.S. databases more

and more, such users may find a bias against their own publications. (See Fig. 2 for a comparison of the production of periodical literature and the production of databases; North America produces 10 to 15 percent of the world's periodical literature and over half of the world's bibliographic databases, containing up to 85 percent of all online records.)

In the event that a bias does exist, there may well be reason to worry about cultural domination. A proposal being studied by the U.S. Department of State would subsidize terminals in less developed countries, would subsidize high volume, low cost telecommunications links to the U.S., and would maintain catalogs of publicly available databases in the U.S., all so that the less developed countries may access the information in U.S. files. (This proposal is being studied in preparation for the U.N. Conference on Science and Technology for Development to be held in Vienna in late 1979.) Such a project may be opposed as "new colonialism for the information age," aside from being opposed because of the possibility of "locking" less developed countries into U.S. technology (11). (There is also cause for international disagreement over the content of some the databases, since some contain data derived from satellite reconnaissance.) As part of the opposition, France is supposedly preparing a similar plan as a "counterproposal" (16).

There are good reasons to believe that the U.S., as a people, has no desire to impose cultural domination, nor to be the world's database supplier. But what the U.S. does want, says one theory, is access (24). This was the thrust of the Open Door policy of the nineteenth century and the basis for whatever 'empire' the U.S. gained following World War II. It is also the basis for the general transnational concept--supported in the U.S.--which is that significant roles are played by nongovernmental organizations

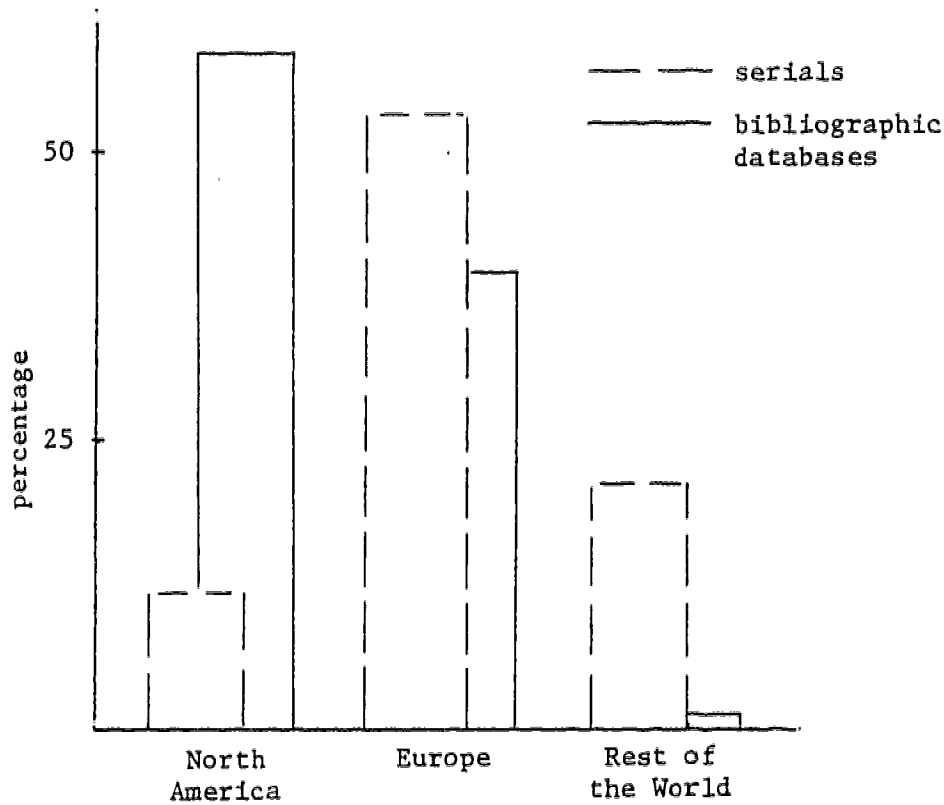

Figure 2: Serials production compared to bibliographic database production.

Compiled from U.N. Yearbook (31), Ayer (1), Williams (34).

in world politics, organizations engaged in international activity aimed not at securing territory but at having access to it.

One might hope, if indeed that is what is happening, that international telecommunications would be a good thing for all countries concerned, an inducement to peaceful realtions and a boon to the separate economies. But as Colin Cherry points out, international telecommunications can reasonably be expected, at best, to merely provide the mechanisms for some organizations to operate (4). How the mechanisms are used is another matter entirely.

What, then, can we expect from the international side of online databases? Probably, in the near future, nothing that can be attributed to the international aspects. What we should try to detect, nonetheless, are the subtle effects (if any) the international use of online informational databases is having on both real and perceived aspects of world literature use and worldwide information sharing.

References

1. Ayer Directory of Publications '78. Philadelphia, Pa.: Ayer Press, 1978.

2. Baker, Dale B. "Recent Trends in Growth of Chemical Literature." Chemical and Engineering News 54, May 10, 1976, pp. 23-27.

3. Chemical Abstracts Service Information Tools--1978 Catalog. Columbus, Ohio: Chemical Abstracts Service, 1977.

4. Cherry, Colin. World Communication: Threat or Promise? London: Wiley-Interscience, 1971.

5. Clinton, Marshall. "International Packet Switching: Its Impact on On-Line Bibliographic Searching." Canadian Library Journal 35(3), 1978, pp. 185-193.

6. "Computerized Information Retrieval Systems." Chemical and Engineering News 55, February 28, 1977, p. 20.

7. Cuadra, Carlos A. "Commercially Funded On-Line Retrieval Services--Past, Present and Future." ASLIB Proceedings 30(1), 1978, pp. 2-15.

8. "The Debate on Transborder Data Flows." EDP Analyzer 16(4), 1978, pp. 1-14.

9. Emmett, Ralph. "Strangulation of World Economies." Datamation 24(3), 1978, pp. 201-203.

10. "Foreign Data Flow Patterns Studied." Transnational Data Report 1(1), 1978, p. 8.

11. "Global Net May Serve Third World." Computerworld, June 26, 1978, pp. 1, 6.

12. Iljon, Ariane. "Scientific and Technical Data Bases in a Multilingual Society." On-Line Review 1(2), 1977, pp. 133-136.

13. INSPEC. An Overview of Inspec. Unpublished, undated paper.

14. INSPEC Catalogue 1978. London: The Institute of Electrical Engineers, 1978.

15. International Atomic Energy Agency. INIS Today. Vienna: International Atomic Energy Agency, 1977.

16. Malik, Rex. "French Pondering Offer to LDCs." Computerworld, July 31, 1978, pp. 1, 6.

17. Martyn, John, and Margaret Slater. "Tests on Abstracts Journals." Journal of Documentation 20(4), 1964, pp. 212-235.

18. National Library of Medicine Programs and Services. Fiscal Year 1977 (Annual Report). Bethesda, Md.: U.S. Department of Health, Education and Welfare, 1978.

19. National Science Board. Science Indicators 1976. Washington, D.C.: U.S. Government Printing Office, 1977.

20. "Planning to Use Public Packet Networks." EDP Analyzer 16(7), 1978, pp. 1-13.

21. Pratt, Gordon (Ed.). Data Bases in Europe. European User Series 1. London: ASLIB, 1975.

22. Pratt, Gordon (Ed.). Information Economics. European User Series 2. London: ASLIB, 1976.

23. The Present Status of the Japanese Data Base Service. Paper provided by the Japan Information Processing Development Center, 1978.

24. Read, William H. "Foreign Policy: The High and Low Politics of Telecommunications." In Anthony G. Oettinger, Paul J. Berman and William H. Read High and Low Politics: Information Resources for the 80s. Cambridge, Mass.: Ballinger Publishing Company, 1977.

25. Sessions, Vivian S. (Ed.) Directory of Data Bases in the Social Sciences. New York: Science Associates/International Inc., 1974.

26. Sources of Serials. 1st ed. New York: R. R. Bowker Company, 1977.

27. Symposium on Transborder Data Flows and Protection of Privacy--Impacts and Trends, Summary Report. Vienna, September 1977. Paris: Organisation for Economic Co-Operation and Development, 1977.

28. Tomberg, Alex. "On-Line Services in Europe." On-Line Review 1(3), 1977, pp. 177-193.

29. "Toward a Modern Secondary Information System for Chemistry and Chemical Engineering." Chemical and Engineering News, June 16, 1975, pp. 30-38.

30. "Transatlantic Data Hookup." Computer News 21(2), 1977, pp. 1-2.

31. UNESCO. Statistical Yearbook 1976. Paris: United Nations Educational, Scientific and Cultural Organization, 1977.

32. U.S. Department of Agriculture. List of Journals Indexed by the National Agricultural Library 1974-76. Vol. II. Beltsville, Md.: National Agricultural Library, 1977.

33. Williams, Martha E. and Sandra H. Rouse (Eds.) Computer-Readable Bibliographic Data Bases. A Directory and Data Sourcebook. Washington, D.C.: American Society for Information Science, 1976.

34. Williams, Martha. "1977 Data Bases and On-Line Statistics." Bulletin of the American Society for Information Science 4(2), 1977, pp. 21-23.

Correspondence

Statistics and information concerning their products and systems were supplied by the following: Biosciences Information Service; Centre National de la Recherche Scientifique; Chemical Abstracts Service; Data Courier, Inc.; Engineering Index, Inc.; Informatics, Inc.; The Information Bank (New York Times) INSPEC; Institute for Scientific Information; International Atomic Energy Agency; Japan Information Processing Development Center; National Aeronautics

and Space Administration; National Agricultural Library; National Library of Medicine; Overseas Telecommunications Commission (Australia); Sociological Abstracts.