IAMSLIC at a Crossroads
(eds) Burkhart & Burkhart, IAMSLIC, 1989

How Good Is Compact Cambridge ASFA? (Aquatic Science Bibliographic Data Base on CD-ROM)

M. P. Tapaswi
National Institute of Oceanography
(Council of Scientific & Industrial Research)
Dona Paula, Goa, India

ABSTRACT

Overviews on the *Aquatic Sciences and Fisheries Abstracts* (ASFA) data base and its latest CD-ROM product, Compact Cambridge ASFA, are presented, in addition to outlining the systems handling bibliographic information at the international level. The difficulties encountered in the use of Compact Cambridge ASFA have been indicated, while extending some suggestions for implementation to increase the utility, reliability and comprehensiveness of this facility.

INTRODUCTION

International efforts: The compilation of a catalogue of scientific papers at the Institut de Bibliographie, Brussels, at the end of the 19th century marked the beginning of the establishment of a data base at the international level (Guha, 1983). The efforts put in, subsequently, during the 20th century towards expansion probably experienced interruption due to the World Wars. Later, at the intervention of and financial support from Unesco and other UN agencies, the scope of the data base as an information system was widened to encompass various disciplines of science, engineering and technology. Such systems in nuclear science (INIS), agriculture (AGRIS), science development and technology (DEVSIS), aquatic and fishery sciences (ASFIS), and science and technology policy information (SPINES) have been set up at the international level for better dissemination, understanding and sharing of knowledge.

These systems emphasize 1) decentralized input, 2) centralized processing and 3) decentralized utilization of the products. The former lays stress on the international collaboration and acceptance of responsibility for collection and consolidation of bibliographic information as an input at the national/regional level. The second aspect underlines and takes into consideration the need for international understanding to evolve standards for use in the preparation of an input record. The third feature indicates considerable change in the nature of documentation services. This decentralized system does not concentrate its efforts for creation of local data bases for international coverage, but limits its monitoring to only local/regional data generated for inclusion in such international data bases. This, in a way, amounts to use of international data bases for local services, like selective dissemination of information (SDI) and/or retrospective literature search.

With the advancement and introduction of computer technology into this field, these data bases have become accessible for on-line searches over and above the information available on magnetic tapes during the last few years. The latest trends in the advances in computer technology could be seen in the introduction of the Compact Disk–Read Only Memory (CD-ROM) system. This disk is simply a small plate coated with plastic which acts as a storage device for information up to 600 megabytes.

ASFA on CD-ROM: The *Aquatic Sciences and Fisheries Abstracts* (ASFA) data base deals with bibliographic information on the science, technology and management of the aquatic environment (marine, brackishwater, and freshwater), including the socio-economic and legal aspects. This is an output of ASFIS, sponsored by the Food & Agricultural Organization of the United Nations (FAO), Intergovernmental Oceanographic Commission (IOC) of Unesco and the Ocean Economics and Technology branch of the United Nations Development of International Economic and Social Affairs (UN/DETB).

The first compact disk named Compact Cambridge ASFA contains records from January 1982 to December 1986 and the second from 1987 to June 1988. This will be updated in the near future. The Compact Cambridge ASFA, with an annual input rate of about 25,000 records, now has more than 1,600,000 records pooled from various input centres world over.

The Compact Cambridge search software provides options – menu driven and dot level – to access the records available on the disk. The search strategy can begin by writing a word (uniterm) or a combination of two or more words (phrase) that can be searched only in one field at a time. Single words can be truncated and searched anywhere or in any selected field in a record. Boolean operators can be used to combine more than one word or sets up to a limit of ten previous (last) searches. A common facility to display/print/download the search results exists. Also available is the 'EXPAND' command, which is used to view the dictionary of terms and number of postings for each of those in a specific data element. This can also be used to decide the root for truncating a word before a search.

The record contains various searchable data elements in fields such as author; author affiliation; corporate author; title; source; text and abstract language; subject; biological, geographic and other descriptors; abstract; classification numbers; environmental regime; conference title; original title; ISSN; ISBN; patent number; other identifying numbers; and classification index.

BIBLIOGRAPHIC SERVICE AT NATIONAL INSTITUTE OF OCEANOGRAPHY (NIO)

This institute caters to the requirements of researchers from academic and R & D institutions in South-East Asia. As a part of the ongoing project "Bibliographic Information System," in collaboration with Unesco, this institute procured the CD-ROM -latest device in data bases- during the second half of 1988. Prior to this, ASFA hard copy and the on-line search facility were available at this institute, and the users were familiar with this data base.

The advantages of Compact Cambridge ASFA over the hard copy and/or on-line searches are

1) fast retrieval,
2) comprehensiveness,
3) easy accessibility,
4) preferred form of output, and
5) economic viability.

These are expected of any latest facility created to meet the demands of the user community.

While using Compact Cambridge ASFA, the author experienced a few difficulties, such as these:

1) failure in downloading search results of complex search strategies,
2) limitation of MACRO utility while reusing search strategy containing phrase with word or phrase,
3) failure in display of first level authors of analytico-monographic documents in any kind of output,
4) non-availability of author affiliation in many references, and
5) suppression of relevant laboratory and theoretical studies while searching on specific environmental regime.

In order to have fuller utilization of this facility, the following are some of the suggestions for implementation at the appropriate level:

1) introduce artificial intelligence/expert systems for searching,
2) provide facility to search for the type of documents/literary style,
3) enable output in any preferred, sorted order, and
4) disclose input centres along with (3) above on output to strengthen document delivery system.

To increase the reliability and comprehensiveness of the data base, the author also recommends evolving a dual monitoring system to check the coverage of cover-to-cover scanned journals and a well defined policy for coverage of type/kind of material published in journals.

The following paragraphs describe the above points in some detail:

1. **Failure in downloading**: If the query has more parameters with many phrases and Boolean operators, the system does not download the searched set. It gives an error message saying that the file already exists, though it did not exist previously.

2. **Reusing saved search strategy**: A search strategy can be saved in MACRO file which can be recalled for search at other time(s) or even for the other disk(s) of ASFA. However, this facility fails when the search strategy has a phrase combined with any other word or phrase, for the simple reason that only one field can be searched at a time while searching a phrase without Boolean operators. Experiences show that at least one phrase exists for most of the search strategies. In such cases, this facility appears to be of no use because of the limitation of searching only in one field at a time without Boolean operators while searching for a phrase.

3. **Author field**: The citations of Analytico-Monographs –AM, AMS, AMC– (viz. paper from conference proceedings) have two levels of title and authors. Although Compact Cambridge ASFA has the capability to search on authors at both the levels, the display in the author field is only of the authors of title level two (e.g., editors of a conference proceedings in the above example). The author of title level one (contributor to the article) is suppressed, thereby providing an incomplete citation. This error needs to be rectified.

4. **Author affiliation:** This field is most useful in getting a reprint from an author for a user's personal collection. Also, this field can be used to find out what kind of studies are being carried out in a given country or laboratory. Recently, this field was also put to use for an editor of a journal, who was interested in finding scientific personnel working on a specific subject, to whom he could eventually write for refereeing articles to be submitted to the concerned journal.

At times, the message "Address not stated" flashes on the screen in this field. In such cases, it is understood that the input agencies involved are helpless. But, on many occasions, the author observed that this field is totally absent from the display. The reasons for such an absence are not known.

5. **Environmental regime**: This is one of the important header fields in ASFA, which notes whether the document is relevant to the regimes of marine and/or brackish-water and/or fresh water. This is particularly useful when the user is interested only in one or two of these regimes. However, if the user is interested in only a laboratory study or a theoretical model, narrowing the search results by looking for documents of any one of the above three regimes, the user is unable to retrieve it as this field is likely to be blank. If the user insists on retrieving documents on the latter, he is forced to accept documents from other regimes, which are of no interest. For this reason, it is felt that 'laboratory' and 'theory' may be considered as 'regimes,' along with the conventional ones.

6. **Introduction of artificial intelligence/expert system**: Most of the computer search strategies compare words or phrases to a given word or phrase. Attempts have been successful with other bibliographic software like CDS/ISIS to interface the thesaurus with the database to give a chance to select hierarchial or related terms for a given term before searching. This has added to an artificial intelligence of the system, and the results appear to be much better than just searching for a word or phrase match. It will be very useful if this facility is added to the software of Compact Cambridge ASFA. Fortunately, ASFIS already has a very useful thesaurus; what is now required is its interfacing.

7. **Search on type of document/literary style:** Though grey literature is gaining more and more importance in the referring system, it has not yet attained a fully reputable status. Experience shows that many users are still interested in journal article citations only. The "Type of document" and "Literary style" fields in header group in ASFIS input format can be made searchable to allow for such queries. Besides, this facility would enable one to take up a study of distribution of a literature in different kinds of documents.

8. **Orderliness in output:** A display/print/download facility exists for the searched results without any definite sequence for the output. If the sort facility is made available for sorting the searched sets in required order (alphabetical, classified or yearwise), the output will have a definite sequence and better appearance. This will facilitate the user's search for the record/reference of interest in the output.

9. **Index/display/search on input centres:** Of the header fields, "Environmental regime" is currently available as a searchable field (Anon, 1986). Incidentally, the other fields, although equally important, have not been categorized under searchable/display fields. For example, a display of Input Centre Code field would help in document procurement. Though the foreward of ASFA hard copy mentions that the full text copies of documents cited in ASFA are generally available through agencies like the British Library Document Supply Centre (BLDSC), Information on Demand (IOD), and National Technical Information Service (NTIS), they have been quite expensive. Such a display might encourage the inter-input centres to exchange documents required by the users. Though there is no binding from ASFIS on input centres to act as document supply centres, some centres might come forward for such an exchange or service with affordable charges.

RELIABILITY OF THE DATA BASE

Encouraged by the advantages of the Compact Cambridge ASFA available with this institute, the author carried out a study to evaluate this latest facility for its reliability. According to the Fishery Information Data and Statistics Service of FAO (Needham, 1986), this data base contains/covers 65% of the current accessions derived from journal articles published in some 5,500 journals, which are monitored cover-to-cover. In order to test this, six core journals, two each published from America, Europe and Asia, were taken as a sample. The idea of selecting journals from different geographical areas was done with a view that these are likely to be monitored by different input centres. Only three years' (1982, 1983 and 1984) volumes were considered for testing on a disk having ASFA 1982 through 1986. If the volume of the journal continued in the next year, it was considered as the volume of that year in which the first issue/number was published. The results (Table 1) surprisingly show that the data base did not cite at all (100% loss) the references from one of the Indian journals, *Indian Journal of Marine Sciences* (IJMS), and none of the journals under study are fully cited (cover to cover). One can understand that the non-availability of Indian journals to an input centre might be a factor for not citing the IJMS, but even internationally reputed journals, which are supposed to be available at all input centres, like *Marine Geology* (year 1982: 61% loss) and *Journal of Physical Oceanography* (year 1984: 47% loss), have not been monitored to the quoted extent.

The losses are mainly in the form of all the articles appearing in a specific missing issue(s)/number(s), but not a partial loss from the number(s)/issue(s) from which other articles are cited (Table 1).

The articles lost/not cited by ASFA by way of "partial loss" from any issue/number of a volume of a specific journal are very few. Among these are the articles falling under the category "discussions and replies" or "comments and replies."

Usually, a body of the journal consists of an editorial, preface to a special volume/number, research articles, short communications, notes and correspondence, letters, letters to the editor, comments and replies, book reviews, personalia, obituaries, announcements, addenda/corrigenda, popular articles, and news items. Perusing the available documents about ASFIS, the author could not find whether any policy exists to consider the kind of material appearing in journals for citation in ASFA. However, the General Aspects Category in the document - ASFIS Subject Categories and Scope Descriptions (Anon, 1983) - has given a place to cite a few of the less important kinds of material appearing in the journal.

It is felt that the discussions and replies should be given due consideration for citing at an appropriate subject category in ASFA, as they have very high value for the promotion of scientific knowledge. In the words

Table 1

Number of Articles Cited from Core Journals in ASFA

	J Phys Oceanogr			Limnol Oceanogr			Mar Geol			Estuar, Coast Shelf Sci			Mahasagar			Indian J. Mar Sci		
Description/Year	**1982**	**1983**	**1984**	**1982**	**1983**	**1984**	**1982**	**1983**	**1984**	**1982**	**1983**	**1984**	**1982**	**1983**	**1984**	**1982**	**1983**	**1984**
Total articles Published (A)	124	186	164	126	130	144	119	90	143	101	110	96	38	60	33	80	60	52
Total articles cited in ASFA (B)	103	156	87	91	116	112	46	68	81	75	85	75	36	59	27	–	–	–
Articles not cited (A-B = C)	21	30	77	35	14	32	73	22	62	26	25	21	2	1	6	80	60	52
% loss of information (D)	17	16.1	47	27.7	10.7	22.2	61.3	24.4	43.3	25.7	22.7	20.8	5.2	1.6	18	100	100	100
No. of issues not at all covered[1] (E)	1(16)	2(25)	5(63)	1(23)	–	1(24)	7(73)	1(10)	4(53)	3(25)	2(21)	2(18)	–	–	–	4(80)	4(60)	4(52)

[1]Numbers in brackets indicate articles not covered in those issued.

of Greenfield (1988), "this illustrates the danger that when we rely on computerized sources of information, we are tempted to regard the computer as infallible; yet, when one relies on a computer, one is actually relying on hundreds if not thousands of human beings and every one of them makes mistakes."

ASFA, as the largest and most comprehensive data base in the field of marine and freshwater studies, and supported by international agencies, should be able to strengthen its coverage with the cooperation of the input centres in the form of a dual monitoring system. The strength of the data base, thus, without any ambiguity, lies in its comprehensiveness, reliability and utility.

ACKNOWLEDGEMENTS

The author expresses his deep gratitude to Dr. C. S. Murty, Scientist and Chairman, Library Committee, for encouragement and help in making this manuscript more expressive, and to Dr. B. N. Desai, Director, NIO, for encouragement and facilities provided.

REFERENCES

1. Anon. 1983. *ASFIS Subject Categories and Scope Descriptions.* Rome, Food and Agricultural Organisation of the United Nations. ASFIS Reference Series 2 (Revision 1).

2. Anon. 1986. *Guidelines to Bibliographic Description.* Rome, Food and Agricultural Organisation of the United Nations. ASFIS Reference Series 3 (Revision 1).

3. Greenfield, R. 1988. "Computers and the Information Industry." *NISSAT Newsletter* 7: 3-4.

4. Guha, B. 1983. "Towards International Network." In *Documentation and Information Services, Techniques and Systems.* 339-53.

5. Needham, R. 1986. *ASFIS: Vehicle for International Exchange of Scientific Information.* El Nino Regional Information Seminar, Cartagena, Colombia, 28-30 July.