the basements (which many houses in Rajasthan have) are twice the ground floor values.

In summary, we can note that in our country, due to the tropical climate and our living habits, we have plenty of ventilation in our houses, which is one of the factors that help in reducing indoor radon levels. Whether one considers the average level over the country, or even the maximum values recorded in the high background areas, the levels are below the ICRP recommended action level of 200 Bqm^{-3}, and hence indoor radon is *not* a significant health risk in our country.

1. *Ann. ICRP*, ICRP Publication 50, 1987, **17**, 1–60.
2. *Sources, Effects and Risks of Ionizing Radiation*, UNSCEAR, New York, 1988, pp. 62–81.
3. Baxter, M. S., *IAEA Bull.*, 1993, **35**, 35–38.
4. *Ann. ICRP*, ICRP Publication 32, 1981, **6**, 1–24.
5. Methods for the measurement of radon, thoron and their progeny in dwellings. Report of a Task Group, Atomic Energy Regulatory Board, AERB Technical Manual No. TM/RM-1, AERB, Bombay, 1993.
6. World Health Statistics: Annual Vital Statistics and Causes of Death, WHO's Report Geneva, 1983
7. *Ann. ICRP*, ICRP Publication 60, 1991, **21**, 137–141
8. Health Risks of Radon and Other Internally Deposited Alpha Emitters, BEIR IV Report. Washington, DC. 1988.
9. Hui, T. E., Poston, J. W. and Fisher, D. R., *Rad. Prot. Dosim.*, 1990, **31**, 405–411.
10. *Ann. ICRP*, ICRP Publication 39, 1984, **14**, 1–8.
11. Sankaran, A V., Jayaswal, B., Nambi, K S. V. and Sunta, C. M., *U, Th and K distributions inferred from regional geology and the terrestrial radiation profiles in India.* BARC, Bombay, 1988, pp 1–104.
12. Nambi, K. S. V., Bapat, V. N., David, M., Sundaram, V. K., Sunta, C. M. and Soman, S. D. *Natural Background Radiation and Population Dose Distribution in India*, BARC, Bombay, 1986, pp 1–30.
13. Raghavayya, M., *A study of the distribution of radon in uranium mines*, Technical Report BARC/1–452, 1977.
14. Subba Ramu, M. C., Shaikh, A. N., Muraleedharan, T. S., Ramachandran, T. V. and Nambi, K. S. V., in *Proc. Seventh National Conf. on Particle Tracks in Solids* (eds Kumar and Reddy), Defence Laboratory, Jodhpur, 1993, pp. 11–25.
15. Narayanan, K. K, Krishnan, D. and Subba Ramu, M. C., Exposure from natural and man-made sources of ionizing radiation, Indian Soc. Rad. Phys Report 1991. ISRP(K)-BR-3.
16. Subba Ramu, M. C., Shaikh, A. N, Muraleedharan, T S. and Ramachandran, T. V. *Sci. Total Environ*, 1990, **90**, 49–52.
17. Nair, N B., Eapen, C. D. and Rangarajan, C., *Radiat. Prot. Dosim.*, 1988, **11**, 193.
18 Kumar, S., Gopalani, D, Jodha, A. S, Ramaseshu, P. and Nagaratnam, A., *Bull Rad. Prot.*, 1992, **15**, 78–80.

ACKNOWLEDGEMENTS. I thank Dr Subba Ramu and Dr Sudershan Kumar for many stimulating discussions.

Received 5 October 1993; accepted 7 October 1993

Electronic databases, networks and information support for scientific research

T. B. Rajashekar and A. Sreenivasa Ravi

Electronic databases and computer networks are bringing forth significant changes in both formal and informal information transfer mechanisms in science. More than 6,000 electronic databases are estimated to be available today, which include a large number of bibliographic and scientific hard databases. While a majority of these can be accessed 'online' using computer-communication networks, they can also be acquired on tapes, diskettes and CD-ROM discs and searched locally. Electronic databases are used predominantly to meet the current and retrospective information needs of scientists. While electronic databases facilitate a researcher in gaining rapid access to published or about-to-be published information, a variety of 'networked information services' have evolved on academic and research networks like BITNET, NSFNET, Internet and JANET. These include electronic discussion forums, data archives, electronic journals, library catalogues and databases. Network information services deliver information to the end-user right at his terminal and have the advantages of convenience, speed and informality. Although researchers in a few institutions in India are beginning to take advantage of these developments, there is need for more widespread awareness of these. This article is an attempt in this direction.

It is now widely acknowledged that the ability to access, transmit, share and disseminate information will make the difference between success and failure in the 21st century. If we have to improve productivity and innovation of our R&D activities and give competitive edge to our technological products, then convenient, economic

The authors are in National Center for Science Information, Indian Institute of Science, Bangalore 560 012, India

and quick access to results of worldwide scientific and technological research becomes very critical. Given the rapid growth of science, reflected by the growing number of researchers, scientific establishments, scientific fields and research publications, conventional information transfer mechanisms—both formal and informal—are proving to be highly inadequate in achieving this objective.

Solution to this problem has come mainly as the result of a major shift in handling information in its analog and physical form (voice, print, audio visual) to electronic form. The moment information is created in electronic form, it becomes mobile, rather than static. It can then be processed, transmitted, stored, retrieved and presented at enormous speeds and convenience. We are witnessing significant changes in both formal and informal modes of scientific communication due to electronic information handling.

Electronic information, together with computer networks, are making inter-personal communication among researchers more instant, global and interactive. Academic and research networks like BITNET, Internet and JANET, which link universities, laboratories, and research agencies around the world, are enabling more and more information to be distributed in advance of 'normal' publication outlets through electronic conferences and distribution lists (bulletin boards, list servers and news services), info servers, electronic journals and newsletters. We discuss these developments on academic and research networks in a later section of the paper.

But this phenomenon of networked inter-personal communication is a recent one. It is electronic databases which emerged first in the late 1960's offering control over and improved access to the rapidly increasing number of print publications like the journal, conference, patent, report, etc. Information explosion, coupled with literature scatter (dispersion of papers in a field over a large number of journals, including a few 'key' and many peripheral journals) and increasing cost of science publications have rapidly degraded the ability of a scientist to keep well informed of recent developments in one's field.

Electronic databases

Electronic databases (EDB), stated simply, are electronic equivalents of their print counterparts, but as is the case most often, are produced first and then used for producing the print version. They can be categorized into two types—reference and source databases. Reference databases refer or point a user to another source, often a document, for more details. Reference databases can be further sub-categorized into bibliographic (containing primarily citations from published information like journal articles, reports, patents, dissertations, conference proceedings and books) and directory (for e.g. listing of companies, associations or people). Source databases contain complete data or the full text of the original source information. These are categorized further as hard, soft-hard or full-text. Hard databases contain original and/or statistically manipulated numeric representations of data; soft-hard databases contain a mixture of numeric data and related textual information (comments about the data and/or bibliographic references to the original data); full-text databases contain records of the complete text. The number of information items in an EDB may vary from a few thousand to several million items (see the box *Some Key Electronic Databases*).

The information flow in database production is shown in Figure 1. Beginning with the National Library of Medicine, USA, with its MEDLARS database in 1960, the number and size of databases has grown dramatically, from about 20 in 1965 to more than 4200 in 1989. Current estimates indicate more than 6000 publicly available electronic databases in a variety of forms: on storage media like magnetic tapes, CD-ROM (Compact Disc-Read Only Memory) and floppy diskettes, and through online/offline database hosts. While a database producer converts information into electronic form and brings it out on a storage medium, a database host (also called a database vendor) company mounts one or more databases on a computer and offers a variety of services in online or offline mode. While there are a few thousand database producers, there are a few hundred database hosts operating around the world today, resulting in the emergence of an entire industry, called the 'database industry'. DIALOG, BRS, STN Intl, ORBIT, Pergamon Infoline, Questel, ESA/IRS are some of the leading online database hosts today.

While online access has been the most popular mode of database searching so far (about 40 million online searches in 1991!), it is CD-ROM that has captured the current imagination of both database producers and users and appears to be the chief medium of database distribution in the 90's. Thanks to its enormous storage capacity (about 600 megabytes on one five and a quarter inch disc—equivalent of 250,000 pages of printed text!) and its compatibility with microcomputers, it is now possible to access large databases locally using a CD-ROM, without establishing online connection to a big computer system thousands of miles away. CD-ROM databases thus come closest to the concept of *Desk top global information systems*, since the entire set-up, including a PC, a CD-ROM drive, CD-ROM discs and associated software and documentation, can be housed on a single table top!

When compared to library card catalogues and printed sources like directories, indexes and abstracting journals, electronic databases are far more easily amenable for exhaustive and rapid processing of complex queries which are a combination of multiple search terms (for

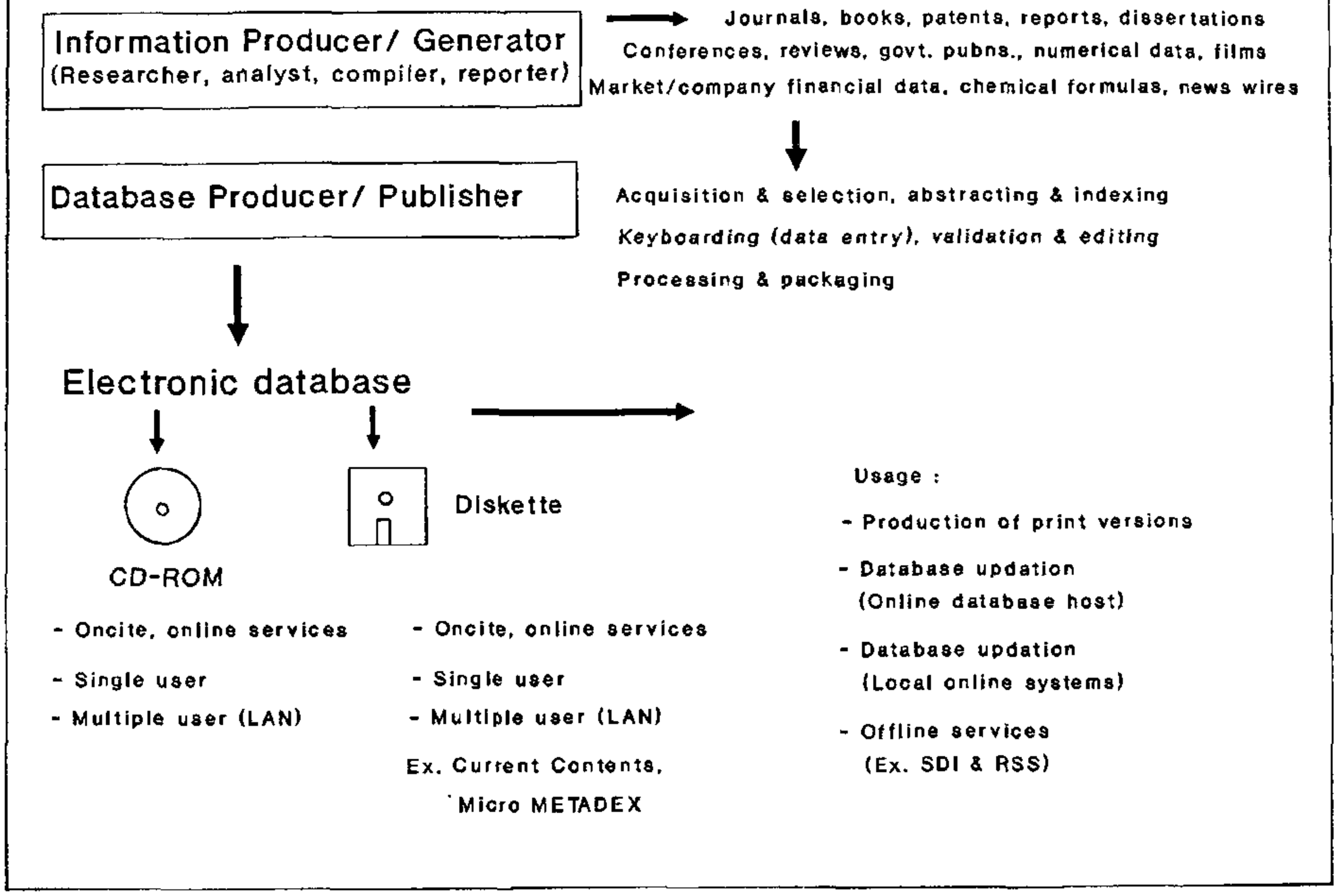

Figure 1. Electronic information path

ex. papers by FC CODD or CJ DATE on RELATIONAL DATABASES or QUERY LANGUAGES published during 1975–1980), retrieval and display of information in a variety of formats. Electronic databases have thus enabled condensing the time element involved in the creation, storing, retrieval and information dissemination process, thereby making more current information available to users than what comes through printed publications.

Searching electronic databases

How does one search an electronic database? There are primarily two ways of searching a database—online or offline. In online searching a computer terminal is used to access a database residing on a computer situated nearby or in a remote location and the searcher interacts with the database by typing in the search statements (queries) and viewing the results on the terminal screen. Such interactive searching facilitates speed and convenience in modifying the search statements to retrieve all useful information. In offline searching, the user submits his/her query to the database service agency in written form which is converted into a set of search statements and processed on an appropriate database to retrieve all matching records which are then supplied to the user. Offline searching therefore takes longer time and more importantly it does not give scope for taking quick corrective actions. Irrespective of the type of media on which the database is stored and the search mode employed (offline or online), the basic techniques of database searching are the same.

Searching bibliographic databases

Databases on a computer are organized in terms of records. Let us consider searching a bibliographic database like INSPEC, which is the world's leading database covering physics, electrical and electronic engg., computers, control and information technology. A search in a bibliographic database of this type does not result in

SOME KEY S&T ELECTRONIC DATABASES

Following is a brief description of some key electronic databases in science and technology.

A. Bibliographic Databases :

1. CA-SEARCH (Chemistry) : References of more than 9 million chemistry publications, dating from 1967. Produced by Chemical Abstracts Service, USA. Equivalent of the printed Chemical Abstracts.

2. INSPEC (Physics, Electrical & Electronic Engg., Computers, Control and Information Technology) : About 4 million bibliographic records, dating from 1969. Producer : Institution of Electrical Engineers, U.K. Also available in CD-ROM since 1991.

3. BIOSIS Previews (Life Sciences) : More than 7 million bibliographic records, since 1969. Producer : Bioscience Information Services, U.S.A. Also on CD-ROM from 1988.

4. COMPENDEX Plus (Engineering) : Machine readable version of 'Engineering Index'. About 3 million bibliographic records, since 1970. Producer : Engineering Information Inc., U.S.A. Also on CD-ROM from 1986.

5. MATHSCI (Mathematics) : Contains evaluative reviews and abstracts of mathematics literature, since 1959. File size : More than a million records. Producer : American Mathematical Society, U.S.A.

B. Full Text and Image Databases :

I. Full Text Databases : Examples include full text of S&T journals like 'Science', 'Electronic Design', 'Electronics', 'Byte', etc., handbooks and encyclopedias like 'Kirk-Othmer Encyclopedia of Chemical Technology' and 'Beilstein Handbook' (organic chemistry), full text legal databases like LEXIS and news databases like NEXIS, Financial Times, etc. All these include only textual information, without any graphics like diagrams, charts, etc.

II. Image Databases : These databases contain complete images of print publications, captured using optical scanning. Examples include ADONIS CD-ROM discs, produced by ADONIS B.V., Amsterdam, covering about 400 leading biomedical journals and IEEE/IEE Publications Ondisc (IPO), a joint venture product of IEEE, IEE and University Microfilms Intl., providing access to complete page images of about 80 journals, 500 standards and 360 conference proceedings published by IEEE (USA) and IEE (UK).

C. Directory Databases :

Examples include 'Research Centres and Services Directory', by Gale Research Inc., providing detailed information on over 26,000 research organisations worldwide; 'Computer Readabale Databases', by Gale Research Inc., providing description of more than 6,000 publicly available databases; 'D&B Dun's Market Identifiers', by Dun's Marketing Services, providing detailed information on business establishments in USA, Europe, Africa, Asia and other countries.

D. Scientific Numeric Databases :

Examples include 1. CSD - Cambridge Crystallographic Database, produced by University Chemical Laboratory, Cambridge, UK. CSD contains crystallographic, chemical and bibliographic information for organic solids and includes unit cell data, atomic coordinates, chemical connectivity, and reference information; 2. Metals Datafile, produced by Institute of Metals, UK. Includes mechanical properties like crack propagation, elongation, fatigue life, etc. and physical properties like density, electrical resistivity, etc. of metals and alloys; 3. GENBANK, produced by Intelligenetics Inc., USA. Contains information on nucleotide, DNA and amino acid sequences; 4. Hybridoma Data Bank, produced by American Type-Culture Collections, USA. Contains information on hybridomas and monoclonal bodies.

information which will directly solve the R&D problem a user has, instead it provides the enquirer with a list of publications which are likely to contain the desired information. Each record in INSPEC describes a document like a paper in a journal or conference proceedings, a patent, a technical report, etc. using fields like author, title of the publication, etc. A sample INSPEC record is shown in Figure 2. As can be seen from this sample record, each document record is also characterized by fields like keywords and classification codes, which facilitate more accurate search and retrieval by subject content.

One searches a database with the objective of obtaining all useful information pertaining to the search topic or query. Retrieving maximum number of useful document records depends on how well the query is posed to the database in question. A search query is formulated in terms of 'search terms' like author names, keywords (subject words taken from document title and abstract or assigned by an indexer), classification codes and journal title. The context or relationship among search terms is indicated using Boolean logic operators–AND, OR and NOT. Boolean operator AND is used to relate two or more search terms when the presence of all these search terms is essential for any record to be retrieved, for instance, as in,

AU = Martin, James AND JL = Data Communications

which will retrieve all papers of 'James Martin' in the journal 'Data Communications', excluding all other publications of this author.

The OR operator is used to relate two or more search terms when the presence of any one of these search terms is sufficient for any record to be retrieved, for instance, as in,

KW = X-ray Diffraction OR KW = Crystal Structure

which will retrieve all references having either or both of these terms. Finally use of NOT operator, as in,

KW = Computer Networks NOT AU = Martin, James

would retrieve all references on 'computer networks' except those which are authored by 'James Martin'.

A search statement may consist of only one search term (along with an indication of search term type, e.g.

TITLE	: Externally excited semi-infinite one-dimensional models
AUTHORS	: Cannas, S.A., Prato, D.
AUTHOR AFFILIATION	: Fac. de Matematica, Astron. y Fisica, Univ. Nacional de Cordoba, Argentina
JOURNAL TITLE	: Am. J. Phys. (USA)
ISSN	: 0002-9505
VOL. AND ISSUE	: Vol. 59, No.10
PAGE NO.	: 915-20
PUBLICATION DATE	: Oct. 1991
CLASSIFICATION CODES	: A0150,A7320,A0365,A7125C,A0365G,A6310
THESAURUS TERMS	: Kronig-Penney model,Lattice dynamics,Potential scattering,Quantum theory,Surface electron states,Teaching
TITLE/ABSTR. TERMS	: Quantum particle scattering,Lattice vibrations,Extended solutions,Electronic states,Externally-excited 1D models,Semi-infinite harmonic chain,Forced periodic motion, Semi-infinite Kronig-Penney chain, Localized solutions, Band edges
ABSTRACT	: The problem of a semi-infinite harmonic chain with the first mass subject to a forced periodic motion is solved exactly. The scattering of a quantum particle by a semi-infinite Kronig-Penney chain is analyzed analytically. The behavior of extended and localized solutions is analyzed in both problems and many analogies are found between them. Special attention is paid to the asymptotic behavior of localized solutions near the band edges.

Figure 2. A sample INSPEC database record

'AU' for author, 'JL' for journal, etc.) or several search terms connected using appropriate boolean operators. For example, search statement for 'studies on lung cancer in smokers' may look as follows:

(Cancer OR Carcinoma OR Malignant OR Neoplasm) AND
(Lung OR Pulmunory OR Respiratory System) AND
(Smokers OR Smoking)

While query formulation and searching by fields like author and journal title is quite straightforward, it is the subject search, i.e., searching using keywords, that poses the biggest problem. Which keywords to use for searching depends on whether the search is being conducted to ensure exhaustivity (search recall) 2—retrieval of all related information, or specificity (search precision)—retrieval of information directly relevant and specific to the search topic. Exhaustivity can be achieved by using keywords which not only represent all major and minor concepts of the search topic but are also at slightly broader level (e.g. using the term 'cancer' in place of more specific terms like 'blood cancer' OR 'lung cancer'). Retrieval precision can be achieved by using keywords which are at the same level as concepts in the search topic (e.g. using 'parallel sorting algorithms' instead of the broader 'parallel algorithms').

It can also be seen that while OR operator increases the search recall, AND and NOT operators increase search precision as they tend to narrow down or increase the specificity of the search topic.

Effective database searching requires good knowledge of the database(s) being searched, subject knowledge and the search system being used. It is for this reason that most often the end user does not carry out the search himself, but delegates it to a search intermediary (search specialist).

Scientific hard databases

Classed as source databases in the hard or soft-hard category, scientific hard databases (also referred to as numeric databases) are an ordered collection of numbers, the values of which (a) correspond to various properties, parameters or attributes of elements, substances or systems and (b) are critically evaluated by experts prior to being included in the database. Production of these databases involves more rigorous functions of review, evaluation and correction, which are not often found in reference databases. This serves to make the data in scientific hard databases more reliable and useful than those found in the open literature because of the rationalization of factors like uncertainty statements and units of measurement. Given the specialized nature of the data available in these databases and the background required to formulate queries, as well as to interpret and to manipulate the results, the end-user of these databases tends to be a practising scientist, engineer or technician than an information scientist. In Figure 3 we show a sample amino acid sequence entry taken from the SWISS-PROT protein sequence database. Since the past few years there is rapid increase in the number and size of scientific hard databases, particularly chemical structure and property databanks (e.g. 'Beilstein current facts in chemistry' on CD-ROM published by Springer–Verlag containing structures, data and literature references for about 300,000 organic compounds) and sequence databanks in molecular biology (e.g. 'EMBL' produced by IntelliGenetics Inc., which contains descriptions of DNA and RNA sequences).

Research information needs and database services

What services can a researcher obtain from electronic databases? There are two major database services—Current Awareness Service (CAS) and Retrospective Database Search Service (RSS). CAS fulfils the current information needs of a researcher during the 'ongoing' stage of research by helping him/her to keep in touch with current research work carried out elsewhere in the world. RSS fulfils the retrospective information needs during the initial stages of research, by providing a researcher with state-of-the-art of the concerned research area helping a researcher in the identification, selection and definition of research problem and in avoiding duplication of research.

For a particular research area, both CAS and RSS can be provided using the same database(s), the major difference being the portion of the database that is used for searching and the frequency with which these searches are carried out, for providing these services. RSS is usually a one-time, exhaustive, search on the entire database or a major portion of it (say the past 5 or 10 years), with the objective of identifying and extracting all relevant information, published during this period. Since search for RSS is done on a large portion of a database, the output size (number of records retrieved) may be very large. RSS systems provide facilities for controlling this size by using limiting factors like language, year(s), document type, etc., for instance, as in,

KW = Raman Spectroscopy/LA = Eng, DT = Art, YR = 1985–1992

which will retrieve only journal articles in English published during 1985 to 1992 in the field of Raman spectroscopy.

CAS (also called 'Selective Dissemination of Information' or SDI), on the other hand, is an alerting service, periodically informing a researcher of current

```
ID   ALR1_ECOLI     STANDARD;      PRT;    64 AA.
AC   P29743;
DT   01-APR-1993 (REL. 25, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
DE   ALANINE RACEMASE, BIOSYNTHETIC (EC 5.1.1.1) (FRAGMENTS).
                              -----> Protein description
GN   ALR.
OS   ESCHERICHIA COLI.        -----> Organism
OC   PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC   ENTEROBACTERIACEAE.
RN   [1]
RP   SEQUENCE OF 1-24 FROM N.A.
RM   84161990
RA   NAKAYAMA N., ARAI N., BOND M.W., KAZIRO Y., ARAI K.;  ---> Reference
RL   J. BIOL. CHEM. 259:97-101(1984).
RN   [2]
RP   SEQUENCE OF 25-64 FROM N.A.
RC   STRAIN=K12;
RM   86076977
RA   KURAMITSU S., INOUE K., OGAWA T., OGAWA H., KAGAMIYAMA H.;
RL   BIOCHEM. BIOPHYS. RES. COMMUN. 133:134-139(1985).
CC   -!- FUNCTION: CELL WALL FORMATION.
CC   -!- CATALYTIC ACTIVITY: L-ALANINE = D-ALANINE.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC   -!- SUBUNIT: MONOMER.
CC   -!- PATHWAY: ALONG WITH D-ALANINE-D-ALANINE LIGASE, IT MAKES UP THE
CC       D-ALANINE BRANCH OF THE PEPTIDOGLYCAN BIOSYNTHETIC ROUTE.
CC   -!- SIMILARITY: TO OTHER ALANINE RACEMASES.
DR   EMBL; K01174; ECDNAB.
DR   EMBL; M12047; ECTYRBA.
DR   ECOGENE; EG10001; ALR.
DR   ECOGENE; EG10038; ALR.
KW   ISOMERASE; PYRIDOXAL PHOSPHATE; CELL WALL; PEPTIDOGLYCAN SYNTHESIS.
FT   NON_CONS      24      25
SQ   SEQUENCE    64 AA;  7392 MW;  22707 CN;      ---> Amino Acid Sequence
   MQAATVVINR RALRHNLQRL RELVDPVILW GEGLPVERIA EMTKVSAYEL ITRLTSRVAM
     KYVD
//
```

Figure 3. Entry from SWISS-PORT protein sequence database.

developments in his/her field of specialization, by extracting relevant items of information from the latest 'updation' of the database, i.e., all the records added during the previous week, month, or quarter. CAS is thus often subscribed for a period of one or more years. The major benefit of CAS is reduction in the effort and time, otherwise a researcher has to incur in manual scanning for relevant literature from the ever-expanding volume of published literature, which can be neither exhaustive, current or accurate.

CAS is profile-based, a profile of a researcher consisting of search statements constructed using search terms and Boolean operators, expressing the current research information needs of the researcher or research group. Every time a database update is received in the form of a magnetic tape, CD-ROM or floppy diskette, a search program automatically scans this database update, extracting all references and abstracts which have matched the profile. The matched abstracts are then printed out and mailed to the concerned user.

A critical component of CAS is tuning the profile to arrive at the appropriate search terms and their combinations which will maximize retrieval of relevant abstracts. Profile tuning is carried out in response to the feedback supplied by the user for each batch of abstracts mailed to him/her.

A sample CAS profile of a researcher in the field of computer science and portion of the CAS output for this profile are shown in Figures 4 and 5 respectively.

Computer networks and informal communication in science

Despite a well-established formal information flow mechanism in the science world, scientists depend heavily on informal communication to collect the latest or 'nascent' information in their fields of research. A scientist working in the cutting edge of a research area invariably exchanges information informally with his peers in the field, thus participating in an 'invisible college.' The significance of this informal domain cannot be over-emphasized as today by the time an article appears in print the research is usually old news.

Current Awareness Service in Physics
National Centre for Science Information
Indian Institute of Science, Bangalore - 12

USER PROFILE

User : XXXXXXXXXX
ASST. PROF.
ECE DEPT.,
INDIAN INSTITUTE OF TECHNOLOGY
XXXXXXXXXX

Profile No. : P0918

Search terms :

Code	Type	Search Term
A	C	C5620 ------> (Computer Networks)
B	K	COMPUTER COMMUNICAT?
C	K	MODEL?
D	K	STOCHASTIC?
E	K	QUEUEING?
F	K	PERFORMANCE?
G	K	OPTIMISATION?
H	K	CONTROL?

Search Expressions :

1. (A+B)*(C+D+E+F+G+H)# (Boolean Search Expression)

(Note : '+' = OR, '*' = AND and '^' NOT)

Figure 4. A sample CAS profile of a scientist

Major impact on the informal domain of science communication has come from academic and research computer networks. Internet, Usenet and BITNET are the most prominent ones among these networks with a very large reach and high activity (see box *'Academic and Research Networks'*). These networks offer several communication services, including electronic mail, file transfer and remote log-in. Electronic mail, or e-mail, is a store-and-forward messaging facility, available on all these networks. E-mail is probably the most widespread of the three network services because it is often the only way to exchange information between these networks. Users on these networks can also invoke file transfer commands like FTP (File Transfer Protocol) on Internet, to copy computer files containing information such as software, documentation, maps, etc. Remote log-in is the most sophisticated service provided on these networks through programs like TELNET on Internet which allows a user at one site to work on a computer at another site.

Network information services

These basic communication facilities have been used to evolve a variety of *'network information services'*—information services delivered over these networks to the

1. Evaluating the field bus data link layer by a Petri net- based simulation. Di Stefano, A.; Mirabella, O. (Istituto di Inf. e Telecomun., Catania Univ., Italy). IEEE Transactions on Industrial Electronics(USA), vol.38, no.4, p 288- 97 (Aug. 1991).

   The performance of the data link layer of the field bus, an emerging LAN architecture for control systems presently under standardization, is discussed and evaluated. Some aspects of the communication mechanism are presented, referring to the use of the two kinds of tokens of the protocol and pointing out their main features and the advantages and disadvantages offered by each one. The protocol was specified by extended timed Petri- nets and was evaluated by a suitable simulation tool. The results provide some criteria according to which the designer can choose, among the various mechanisms made available by the standard proposal, those which provide the most efficient solution for the scenario being considered (6 refs). (Journal paper) (NCSI No. - P20/4014982)

2. Performance evaluation of layered communication protocol. Inai, H.; Yokohira, T.; Murata, M.; Miyahara, H. (Inf. Process. Center, Kobe Univ., Japan). Electronics and Communications in Japan, Part 1 (Communications)(USA), vol.74, no.4, p 13- 25 (April 1991).

   Proposes a queueing network model for the layered communication protocol and its approximate analysis method. The performance of the layered communication protocol is evaluated by using the proposed model and analysis. First, the queueing network model is constructed for the connection- oriented and connectionless data transmissions between peer layers. In each queueing network model, the transmission delay in the lower layer protocol is represented by an infinite- server queue. The performance model for the layered communication protocol is constructed by piling up each model. Then an approximate analysis is presented for each queueing network model. Based on the relation to be satisfied by the packet transmission delay in each layer and the throughput between adjacent layers, an approximate analysis of the whole model can be executed iteratively. In numerical examples, the validity of the approximate analysis for the model is shown. The effects of window size, the time- out period, and receiver buffer size on the transmission delay are discussed (11 refs). (Journal paper) (NCSI No. - P20/4018608)

Figure 5. Portion of the CAS output from the INSPEC database.

end user right at his/her terminal. Two major network information services, both of which use e-mail as their main communication medium, are electronic discussion groups and infoservers. An electronic discussion group (also called variously as electronic bulletin board, electronic conference, etc.) uses e-mail to set up an informal discussion by interconnecting people of specific interest, over a network. Members of such lists can exchange messages with others irrespective of the network they are on and of course their geographic location. More importantly, these messages get archived in electronic form and can be retrieved for later use. The software used for setting up an electronic discussion group is known as mail server. Mail servers are electronic mail delivery systems, which when they receive a message can resend it to a group of users/subscribers whose e-mail addresses are maintained as a mailing list. Each subscriber sees all the mail forwarded by the server and if he/she wants to add his/her comments on the issue, sends in a message to the server. This in effect creates a discussion group over the network on a particular topic. An info-server (also called as data archive) enables a network user to search an information archive or a database, such as an archive of an electronic discussion list, a bibliographic or library database, a directory of network users, etc., resident on a remote computer system on the network, by sending his/her query as an e-mail message and receive the search results by e mail.

Academic and Research Networks

Most significant impact on informal communication among researchers has come from academic and research networks like NSFNET, BITNET, JANET, USENET and Internet. The capabilities and reach offered by these computer-to-computer networks has enabled them to be viewed as 'electronic high-ways' or 'data carriers' for development of a multitude of information oriented activities and applications.

Internet - A network of networks: The Internet is a worldwide collection of thousands of interconnected computer networks which is used by over a milion people daily. The ancestry of the Internet is deeply rooted in the ARPANET, the first ever network developed by the Advanced Research Projects Agency of United States' Department of Defense to aid sharing of resources and information among researchers. Since its creation in 1983, the Internet has grown exponentially in terms of number of networks connected to it. According to one estimate there are more than 5,000 announced networks in over 50 countries connected to Internet as of October, 1992 with an estimated user base of 3 million people. The services available on this net are e-mail, file transfer and remote login to other systems.

Resources available on Internet include computing centres, electronic discussion forums (e.g. IR-L@UCCVMA.UCOP.EDU for information retrieval), electronic journals and news letters (e.g. TeXMag for Tex typesetting system), library catalogues (e.g. MELVYL at Univ. of California), info servers or data archives (e.g. netlib@ornl.gov which returns mathematical software by e-mail), directory databases that contain basic contact information about network users. There are also network navigation tools like WAIS, Gopher and Archie which facilitate identification and access to these resources through user friendly interfaces.

Usenet - USENET began in 1979 and has about 265,000 users and 9,700 hosts in five continents according to a recent estimate. It supports only one service called *news* - an electronic conferencing service. USENET news is read widely in the world and gets contributions from all over the world. The total data traffic everyday is in several mega bytes. The newsgroups are divided into some broad categories:

comp: Computer science and software source
sci: Technical discussions on Physical, Chemical, Biological sciences etc.
misc: Miscellaneous topics
soc: Social issues and socialising
talk: Debate oriented subjects
news: The news network and the software
rec: Recreational activities and hobbies

There are hundreds of newsgroups that exist today on almost every conceivable topic starting from the group that discuss as technical a topic as nuclear fusion in *sci.physics.fusion* to a group named *rec.food.cooking*. And groups exist for most areas of computer science including those for distribution free software source codes. Each USENET host collects the news from its neighbouring host, called *newsfeed*, using the software *news*. Unlike the BITNET lists the USENET news is distributed at host level. Of course a host need not subscribe to all the newsgroups available at the newsfeed.

A number of hosts archive the postings of important newsgroups and provide them on archive servers as mentioned above. There are some systems which provide gateways and re-distribute the newsgroups to BITNET so that they are available as BITNET lists. An example of a group that is available in this way is *bionet* for biological sciences.

In this context, mention must be made of *UUNET* which is a non-profit subscription-based network relay service for traffic on UUCP network and others like USENET and Internet etc. Physically *uunet* is a single system located in Arlington, Virginia. Subscribers are charged hourly connection costs for mail exchanges and a monthly flat fee. There are hundreds of subscribers on *uunet* now including backbone machines of national networks from India, Australia, Europe etc.

BITNET - BITNET (Because It's Time NETwork) is a cooperative network started for universities in USA. Presently it is serving over 2300 host and several hundred sites in 32 countries and is basically limited to academic institutions. The underlying protocol is NJE (Network Job Entry). The major communication facilities it supports are e-mail and a restricted kind of file transfer. BITNET has a very sophisticated mailing list maintainer called LISTSERV. This software helps to set up discussions on the network using e-mail. When a message is sent to LISTSERV, it resends the message to a group of users whose e-mail addresses are maintained in its mailing list. Each user can send his response to LISTSERV which will be distributed by the system to all the users again. This in effect creates an electronic group discussion on the network. To join a list a user needs to send a *subscribe* command as an e-mail message to LISTSERV. Similarly there are commands to leave a list like *sign off* or *unsubscribe*. LISTSERV also has some archive server functions. It maintains archives of the postings of a list and users can retrieve required postings by sending queries via e-mail.

The e-mail address to which the commands should be sent starts with the word *LISTSERV* and the address to which a posting is to be sent starts with the name of the list itself. The number of these BITNET lists runs to a few hundreds covering a range of topics from scholarly discussions on superconductivity on the list *SUPER@FROMP11.BITNET* to small talk by coffee enthusiasts on *COFFEE-L@UBVM.BITNET*. There are about a dozen lists that discuss various topics in library and information profession. We found *PACS-L* (Public Access Computer Systems List) monitored from University of Houston to be very useful. In order to join this list one needs to send an e-mail with a one-line message: *SUBSCRIBE PACS-L <first_name> <last_name>* to the address: *LISTSERV@UHUPVM1.BITNET*. An active list not only serves as a valuable source of information on current trends in the field but also provides answers to specific problems as a large number of professionals participate in the discussion lists.

JANET of UK: JANET (Joint Academic NETwork) was established to provide consolidated network links among universities and research institutions in UK. The communication facilities provided on JANET are remote login, file transfer, remote job entry and e-mail. E-mail can be sent to JANET from any other major network. A remarkable development in the case of JANET is the highly coordinated approach adopted by the library & information community for planning and providing information services over JANET. JUGL - JANET User Group for Libraries was instrumental in establishing several projects for network information services. A group of organisations are set up now for looking into provision of information services on JANET such as UKOLN - UK office of Library Networking - to coordinate network initiatives and promote understanding of library needs among the system suppliers; CHEST - Combined Higher Education Software Team - to negotiate special deals with hardware and software suppliers and purchase of databases necessary for the academic community. Besides these some free agencies operate from the universities for providing specific information services on JANET like BIDS - Bath Information and Data System from University of Bath which provides access to Current Contents and Science Citation Index databases from Institute for Scientific Information, U.S.A., procured by CHEST, charging annual subscription to the network users; and NISP - Networked Information Services Project of University of Newcastle upon Tyne, which developed a mail server software *mailbase* for establishing discussion groups on JANET.

Indian scene

Database services

Compared to many of our Asian neighbours like Japan, the Koreas, Hong Kong, China and Singapore, India has been a late starter in offering international level database services to its S&T research community. Some progress has been made since the mid 1980's by establishing a few database service facilities at national level which can be used by a researcher to obtain database services. The services provided by these agencies vary in terms of subjects and publications covered and also the range of services offered. A list of a few major database service agencies in India is given in the box *Database Service Agencies in India*. A few positive developments are noticeable over the past couple of years in the country—launching of a few metropolitan library networks (e.g. Delhi Library Network DELNET, Calcutta Library Network CALIBNET), improvement in data communications and networking facilities and a few database creation efforts (e.g. the National Union Catalogue of Scientific Serials or NUCSSI, by INSDOC).

Accessing network information services from India

Education and Research Network (ERNET)—a project funded by the Department of Electronics (DOE), Government of India and the United Nations Development Programme—is a commendable effort in popularizing the networks and e-mail among the academic community in our country. It was started by linking computer systems (all running under UNIX environment) at the National Centre for Software Technology (NCST), Bombay; DOE, New Delhi; Indian Institute of Science (IISc), Bangalore; and the five IITs. The data transmission medium was the public telephone network. The node—as each site on the network is called—at NCST has an international gateway to 'uunet' to facilitate exchange of e-mails with people on the other international networks. The number of nodes on ERNET has grown since its inception and now it links not only academic institutions but also other research organizations like CDOT, ITI and some private establishments such as PSI Data Systems and so on. ERNET is more like 'uunet' explained above than any other network. Recently ERNET nodes have changed over to dedicated links to NCST gateway, enabling use of full Internet services, including FTP and TELNET for file transfer and interactive access. Currently, only those user machines having dedicated line connectivity to a ERNET host can avail all these services, limiting dial-up users to only e-mail facility. SIRNET (Scientific and Research NETwork) is another noteworthy attempt in promoting and spreading e-mail links in the research world in our country. This network is set up and managed by the Indian National Scientific Documentation Centre, New Delhi (INSDOC) a constituent centre of CSIR. This can be said to be a sub-network of ERNET with the ERNET node at DOE in New Delhi forming the 'gateway' between the two. It links a number of CSIR laboratories and a few other institutions across the country with the e-mail facility.

Database Service Agencies in India

Following is a non-exhaustive list of agencies known to be offering information services in the country using electronic databases in science and technology

Agency	Subject areas
National Centre for Science Information, Indian Institute of Science, Bangalore (Contact: Chairman)	Science and Technology
Indian National Scientific Documentation Centre (INSDOC), 14, SV Marg, New Delhi (Contact: Director)	Science and Technology
National Access Centres for International Databases (NACID) of NISSAT, DSIR, located at CLRI, Madras; NAL, Bangalore, CDRI, Lucknow; NCL, Pune and INSDOC, New Delhi (Contact: Head, Information Centre of these institutions)	Science and Technology
NIC–ICMR MEDLARS Project, National Informatics Centre, 'A' Block, CGO Complex, Lodhi Road, New Delhi	Medicine
Bio-Technology Information System (BTIS) Centres of Dept. of Biotechnology (DBT), located at IISc., Bangalore; MKU, Madurai; Poona University, Pune; NII, New Delhi; JNU, New Delhi; NIV, Pune (Contact: Coordinator, BTIS Centre in respective institutions)	Biotechnology
Informatics India (Pvt) Ltd, Bangalore	Databases on DIALOG
Central Machine Tools Institute (CMTI), Bangalore	Machine Tools
Indian Agricultural Research Institute (IARI), New Delhi	Agriculture
Central Food Technological Research Institute (CFTRI), Mysore	Food Science & Technology
Information Centre, R&D Division, SAIL, Ranchi	Metallurgy
Library and Information Centre, BARC, Bombay	Nuclear Science
National Information Centre for Crystallography (NICRYS), Madras University, Madras	Crystallography
National Institute of Oceanography (NIO), Goa	Oceanography, Fisheries Science
Technology Information Forecasting and Assessment Council (TIFAC) Centres, DST, New Delhi	Technology
Inter-University Centre for Astronomy and Astrophysics, Pune	Astronomy and Astrophysics data

From uucp Thu Sep 19 15.11 IST 1991
>From shakti!VM1.NoDak.EDU!FUSION Wed Sep 18 08:13:33 1991 remote from vigyan
Received: by vigyan ernet in (smail2 3)
id AA20737; 18 Sep 91 08:13:33 PST (Wed)
Received: by shakti.ncst.ernet.in (5 61/Ultrix3 1-B) for ernet in
id AA25779, Tue, 17 Sep 91 15:45 50 +0530
Received: by sangam ncst.ernet.in (5.61/Ultrix3.0-B) for ernet.in
id AA06301; Tue, 17 Sep 91 11:36:09 +0530
Received. from VM1 NoDak.EDU by relay1.UU NET with SMTP
(5.61/UUNET-internet-primary) id AA09854; Tue, 17 Sep 91 02 04 59 -0400
Message-Id: <9109170604.AA09854@relay1.UU.NET>
Received: from NDSUVM1.BITNET by VM1.NoDak.EDU (IBM VM SMTP R1.2.1MX) with BSMTP id 1287, Tue, 17 Sep 91 01:04:41 CDT
Received: from NDSUVM1.BITNET by NDSUVM1.BITNET (Mailer R2.07) with BSMTP id 8202; Tue, 17 Sep 91 01:04:39 CDT
Date: Mon, 16 Sep 1991 17:52:00 PDT
Reply-To: shakti!VM1.NoDak.EDU!AMES!ZORCH!FUSION.ARC.NASA.GOV
Sender: "Fusion - Redistribution of sci.physics.fusion"
<shakti!NDSUVM1.BITNET!FUSION>
From: shakti!VM1.NoDak.EDU!zorch!fusion.sf-bay org
Subject: Fusion Digest 69
To: Multiple recipients of list FUSION <uunet.UU.NET!NDSUVM1!FUSION>
Status: R

Originally-From: kb@jet.uk (ken blackler)
Newsgroups: sci.physics.fusion
Subject: Re: Current status/News
Date: 16 Sep 91 09 17:57 GMT
Organization: Joint European Torus

THIS IS A RE-POST as I seem to have had distribution problems!

In <1991Sep9.125113 2141@inmos.co.uk> pauls@penguin.inmos.co.uk (Paul Sidnell) writes:

>As an interested bystander without the technical knowledge to grasp much of the detail >in this news group, I would be very interested if someone could post a brief summary of >the current status of fusion research, who's building what, what is the best energy yield >so far etc. Is there enough happening for someone to make this a regular "news" fea->ture ? Thanks in advance.
>Paul

I am not very up on the best results gained by other experiments but here are the best JET (Joint European Torus) results, which are I think currently the best. ;-) JET, aproject funded by 14 European countries is currently the worlds biggest fusion experiment, but it is given much competition from many other sites around the world (who will hopefully also reply !)

Best plasma temperature: 300 Million desgrees C
Best central plasma density: 4*10e20 m-3
Best Energy Confinement Time: 1.18 seconds

The best fusion rate is: 7*10e16 fusions per second
producing: 3.7*10e16 neutrons per second

This is equivalent to 40kW of fusion power.

We are currently using Deuterium and plain Hydrogen as fuel.
If JET were to use a 50:50 mixture of Deuterium:Tritium this would increase to a few Tens of Megawatts of fusion power which is nearly equal to the power currently used to heat the plasma.
Behind all this is the fact that we still put more energy in than is produced in the fusion reaction, the plasma does not 'Ignite' and become self heating, which is what would be required in a 'real reactor'. It is generally accepted that JET (and all current experiments) are all too small to behave as real reactors, since it is easier to contain a plasma in a larger tokomak (basically a doughnut shape) than in a small one (because the bends are less tight).

Thus there is currently discussion going on for a project to follow on from JET and other current work. This new project will either be another European project called NET, or an inter-continentally funded project called ITER. This device is likely to be twice the size of JET, be able to sustain fusion for up to an hour at a time, and to generate 1000MW of energy It is very possible that ITER would not be a single machine but several in different places all looking at different design and physics prolems (and competing which is probably good for the science!).

I hope this is of help to you, and any others who have read it I hope also
... .
... .

Figure 6 A sample message posted on FUSION discussion list

All e-mail-based information sources and services (e.g. electronic discussion forums, info servers, electronic journals) on Internet and associated national and regional networks can be accessed by Indian users who are having e-mail connectivity to ERNET.

Developing network information services in India

We can use the existing network infrastructure in the country and develop several interesting and useful services, utilizing just the e-mail facility. We discuss a few possibilities here. *Electronic discussion forums can* be set up and operated for specific research groups or areas. In addition to improving communication among researchers and research administrators in the country, such forums would also encourage the users to develop more sophisticated network skills and to explore more fully the potentials of electronic communication. Accessibility to existing databases can be increased by creating e-mail based front ends which will receive user queries, formulated in a specific syntax, by e-mail, carry out the search and send the results to the user by e-mail. The database could be a bibliographic database, periodicals contents pages database, etc. Several hundred discussion forums exist today on Internet on a wide variety of topics. The messages posted on these forums are an important source of informal, raw knowledge and often untried new ideas. They can serve as a valuable complement to publications and conferences for professional development. Such forums on important research areas can be subscribed from a node on ERNET and the contents can be distributed periodically to the researchers. Their responses to these discussions may also be channelized through this gateway system. An excerpt from a list called FUSION is shown in Figure 6 to illustrate the kind of discussions on the bulletin boards. This particular bulletin board is actually a redistribution of the Usenet news group 'sci. physics. fusion'. More importantly, there is need for setting up a network information monitoring group to systematically scan and monitor information related activities on academic and research networks and assist our researchers in effective use of network information sources and services.

Conclusion

We have made an attempt in this article to give a broad perspective of both formal and informal information transfer mechanisms, facilitated by electronic databases and computer networks. The possibilities of scientific information services discussed in this article are only illustrative. Seemingly limitless developments are taking place both in the technology of computers, computer networks and their utilization. It remains to be seen how the librarians, information professionals and the scientists in India will respond to this challenge and adopt their services and information gathering habits to the emerging networked environment.

ACKNOWLEDGEMENTS. We thank Professor V. Rajaraman, Chairman, National Centre for Science Information, and Dr A. Ratnakar, Librarian, Raman Research Institute, for their useful suggestions during preparation of this paper. We gratefully acknowledge the support received through a project funded by the Department of Electronics, Govt. of India, in preparing this paper.

Received 20 August 1993; accepted 29 September 1993