Personal Documentation and Information System for Engineers[@]

M S Sridhar[*]

Abstract: Enumerates the information explosion environment and the findings of past research on information gathering patterns of engineers, presents basic concepts of documentation, information and personal documentation and information system, stresses the need for systematic management of personal documentation and information system, explains briefly the various principles, procedures and processes involved in developing and managing personal documentation and information system and concludes by calling for little investment of time and energy on the part of engineers to develop and maintain scientific personal documentation and information systems to derive enormous benefits to the organisations and individuals.

Keywords: Personal documentation system; personal information system; engineers; information gathering pattern

1. Introduction

A typical R&D engineer or scientist, quite often, suffers from not being able to get specific information required at right time and right place as well as by having 'access to excess information' or what is called 'information overload'. If you get some critical information which you should have had yesterday (late detection of information) and it is too late to use it today (as your design might have frozen), you 'repent' on it and you may ignore it. If you get some information which may not be useful to you now and not definite about its utility later and you do not have proper plan and system for processing and storing information for future use, you may intentionally or unintentionally ignore it. If you have exceptional memory you may recall perfectly or partly either the information itself or the place where information is stored. But many of us are not blessed with exceptional memory.

Increasingly more and more information, both sought and unsought, would be pouring on your table and mind requiring to make a systematic selection of it for present as well as future use. Lucky (1993, p223) gives a very interesting account of how a typical engineer or scientist bathe

[@] *CLIS Observer*, 6 (1&2) January - June 1989 : 9-13.

daily in tubs of bits (information) and afterward pull the plug and watch the excess information drain away. Most of their time is spent on either accessing information or pretending to access information, the later activity helps to preserve the illusion of the saving information-age citizen. An average engineer or scientist might get about 300 pages or 1 MB of material every day apart from many technical journals they have to read to keep professionally competent. An average person has an input/output capacity of about 150 bits per second and hence 300 pages require all the office hours (of say 8 hours) to be spent in input mode only. Even if they spend 8 hours a day to absorb all that arise in a day there is very limited memory (say 100 K) each one has to store them. As per pigeon-hole theory of the brain, the excess bits will fallout after the input cycle is over. Further ROM section of the memory may be still smaller to permanently store and recall as and when required. Under the circumstance, formal institutional systems like libraries might be of some help you at a broad level to lay your hands on sources of information, provided you have time and energy to explore them. But tailoring such information to your personal need could consume considerable time and energy. This possibility of shift from Library to personal information system is like shift that occured from public to private transport (DeBuse, 1988).

An R&D engineer, who often works away in the field with definite time targets for his projects, requires or seeks information from a variety of formal, informal and interpersonal sources of information needs to optimise his ways of gathering information to have least of 'repentence' and conserve his productive time and energy by planning for a personal documentation and information system in addition to departmental and organizational information systems.

I plan in this presentation to tell you about what, why and how of personal documentation and information system. Before that, I wish to set the ground by presenting some features of information gathering patterns of engineers in general as discovered in past studies.

2. Information Gathering Patterns of Engineers

Let us first explore the information gathering patterns of engineers based on past research, particularly the nature of information required, accidental acquisition of information, delegation of information gathering task, use and dependence on information sources - formal, documentary, informal, intra-and interpersonal sources of information and late detection of information.

Information-needs of engineers differ widely from those of scientists and other R&D workers. The differences in needs and demands for information are more strongly related to the kind of employment and type of organisation than the discipline.

The accidental or chance or unplanned acquisition of pertinent information occurring in the case of some engineers might have been caused by the process of ‘peripheration' as a preparation for inter-personal contact, specialising in two or more quite different specialities, having open information-system as against perfect and closed system, increase in the size of the team and identification of a central core area and a largely ill-defined peripheral area of interest by engineers. In other words apparently irrelevant or non-specific information also plays an important role in the process of problem-solving by engineers.

Generally, engineers were found to delegate their information gathering task to others more than scientists.

The factors which decide the choice of a source of information by an engineer apart from task and purpose of seeking information, are physical proximity, accessibility, perceived quality and utility, `ease of use' and previous experience about the source or acquaintance with the source. These factors are very much inter-related. It is found that accessibility and `ease of use' are stronger factors than perceived quality and the amount of information expected to yield by a source. Anirrationality (i.e., a curious filtering process) is that engineers use channels in proportion to accessibility and `ease of use', but they accept ideas from those channels in proportion to technical quality (Gerstberger and Allen, 1968; Rosenberg, 1970).

Both formal and informal sources of information are used to meet the information-needs of engineers. Certain needs are associated with certain sources, usage of which varies with job function/nature of work, discipline, professional focus, organisational affiliation/work environment, education, nationality, user evaluation of the channel, stages of the task, etc. For example, mechanical engineers engaged in research and teaching made greater overall use of literature, and those engaged in practical aspects of engineering like design, testing and maintenance used data sheets, hand books, BSI documents and trade literature (Wood and Hamilton, 1967). On the other hand industrial personnel with professional focus tended to seek external personnel communication and journals as compared to internal corporate sources sought by personnel with operational focus (Rosenbloom and Wolek, 1970, p 91-92).

Defense personnel also turned most of the time to either colleagues or departmental files or personal files for information (Auerbach, 1966, p 106-107; 1965, p 1-12).

The general picture is that scientists are extrovert and depend more on formal sources and engineers are introvert and depend more on informal sources. Within inter-personal sources scientists contact colleagues outside-the-laboratory more than the engineers do. Such a difference is attributed to the nature of work that engineers are concerned with `making things work' (Wolek, 1969), the psychological traits that predispose an engineer to solve problems by himself or with the help of colleagues rather than by finding answer in the literature (Anthony, et.al, 1969), use of relatively (three decade) old basic science inputs for technological innovation (Crane, 1971, p 29-30; Price, 1965, p 553-568) and training and habituations of engineers in the use of formal information-system (Paisley, 1968, p 10-11). Wilkin (1981) refuted the general findings (or belief!) that engineers read less than other professionals stating that the complex relationship between engineers and their sources of information is still not well understood and findings were often interpreted out of contexts.

It is interesting that the engineers do not always turn to information-sources which reward them most. They try to minimise loss than maximising the gain in turning to a particular source, exhibiting a sort of conservative attitude probably due to their objective of doing `better things' than `best things' (Gerstberger and Allen, 1968, p 271).

In a problem-solving, and decision-making situation, an engineer, first, turns most probably to intra-personal reserve supplies such as personal files, one's own head, memory and knowledge (Shuchman, 1981, p 35; 1982, p5; Raitt, 1984, p 213) (may be because of factors like accessibility, approximity and ease of use), failing which or finding it in-sufficient, next turns to informal channels such as contacting a colleague or a `technological gatekeeper', or delegating it to a colleague, internally-generated documents like reports, he then (unless withdraws from the problem or accepts the unsatisfactory situation) proceeds to formal information system like a library or information centre, with or without modifying the nature of the problem.

The ranking of formal sources of information as needed by engineers has varied widely among different studies. Scientific, technical, professional and trade journals are the single most widely used formal source of information especially for keeping up-to-date by engineers. Though abstracting and indexing journals are valued high, they are relatively less

used. Technical reports were found to be next only to technical journals in importance. High-dependence of engineers on suppliers' information or manufacturers' catalogues or trade literature and heavy use of data books and text books are two important aspects of information gathering by engineers.

It is very difficult to say that reading is a superior way of transferring or acquiring information than hearing or other ways of transferring or acquiring information. Informal, oral and inter-personal sources within the organisation are very important sources of information for engineers whether they are direct or indirect (two step/multistage flow), vertical or horizontal in the organisation structure. Great body of information is getting around by a mechanism that can only be termed gossip.

Several reasons were put forth to emphasise the relatively greater role played by informal sources of information in case of for engineers. The practitioner-engineers, who normally work under the conditions of uncertainty and/or anxiety in diverse set of research areas naturally turn first to their colleagues to compare the results with other similar results (Havelock et.al., 1969 p 4-12), to get a tailor made solution synthesised to support a finding (Ackoff et.al., 1976, p 148), to have a source of confidence and reassurance, to get details concerning procedures or experiments, to cut short the lag in publication-time (Hall, 1972, p 14), to have expert assistance in locating diverse set of published material (Crane, 1971, p 30), to communicate inter-personally the complex messages (Wolek, 1970) and to fill the gap between supply of and demand for information (Kunz, et.al., 1977, p 9).

Wolek (1972) found a rich period of preparatory activity preceding the actual exchange of information in informal discussions consisting of: (i) `piggy backing', i.e., storing the information-need with the hope of an accidental encounter with the required information, (ii) `friendly consultations' with the intentions of making friends known about one's interests and (iii) `professional peripheration' to get sufficient background and understanding to enable him to approach a competent person.

Hence, the primary source of engineering information is largely what the engineer keeps in his head (intra-personal source) or possibly knows where to find in books or catalogs in his office" (Shuchman, 1982, p 5). But the effects of knowledge and experience are largely unconscious, and it is felt difficult to make estimate on the nature, size and value of such intra-personal sources of information (Wilson, 1977, p 61).

As far as personal collection or personal file of engineers and technologists are concerned, a large proportion of it consists of unsolicited literature (Langrish, 1972) in the form of trade literature/product catalogues, manuals, technical reports, drawings, trade journals, standard specifications, hand books, reprints, preprints, Xerox copies of papers, drafts, quotations, calculations and notes. These `extracted literature' or `excerpts from books' or `dismantled books' significantly overlaps with formal sources. The proportion of reports, reprints, preprints and trade literature in personal collection differs significantly with the attributes of an engineer and his area of work. Increasingly, The size, quality, arrangement and indexing of an individuals private library will influence the nature of volume of his skirmishes with any external library or information-service" (Rowley and Turner 1978, p 103- 104).

Research in the area of late detection of information and tolerable delay in supply of information, reveal that, generally, engineers and technologists have tended to ignore information found late i.e., after their designs were `frozen'. The late detection can either indicate conscientiousness or laziness. Engineers who are most likely to detect late information are those who are conscientious in their literature-searching and infrequent users in their haphazard search find more of accidental late discoveries (Skelton, 1973, p 145). Late discoveries of information might be due to harmless or harmful-costly ignorance. Latter often leads to repentance after late detection.

3. Personal Documentation and Information System

The word `Document' is often mistaken for historical document. It can be defined as a graphic record which provides data or information or phenomenon or idea in pictures or words or sketches.

Unfortunately, the word `documentation' is also very frequently mistaken for `reprography', `document production', `translation', etc. Rightly or wrongly `documentation' is also associated with `standardisation' and activities of `design and drawing offices'. Documentation, as a technique can exist in a wide spectrum of situations like tax offices, insurance companies, banks, libraries, etc., in addition to standardisation, design and drawing offices. Documentation involves a variety of tools and techniques to systematically plan, collect, store, organise `documents' or even `information' and provide efficient retrieval and

useful services to those who need information. Often the phrases `documentation service' and `information services' are interchangeably used.

To digress a little, information is an amorphous concept, less susceptible to a precise definition. Information is considered as fifth need of man ranking after air, water, food and shelter (Kemp, 1976, p101). Information-collection, transfer and use are all-pervasive and universal activities. Knowingly or unknowingly, intentionally or unintentionally, all of us most of time of our life and work (including the present moment) are concerned with information its generation, recording, processing, repackaging, transfer, receiving, use and application.

We all know, data is `raw or unprocessed' information and `knowledge' is passive and an item of knowledge becomes an item of information only when it undergoes the active process of communication. The potential of information is realised and the purpose of communication is served only when the information is put into intended use.

Research in the past, has repeatedly confirmed the positive relationship of efficient and effective communication and information supply with performance, productivity, innovation and even creativity.

'Information-cycle' broadly has three phases *viz,* (1) information-production, (2) information-transmission including dissemination & distribution, and (3) information-consumption, use and application. These phases are not easily discernable as an information producer simultaneously acts as consumer. Yet personal documentation & information system is based mainly on the role of individual as a user or consumer of information.

It is possible to identify levels of information systems corresponding to Maslow's hierarchy of needs: (I) Survival-level information consists of basic financial and resource data used to keep the organization alive. *(ii)* Safety-level information which supports management functions. *(iii)* Information which moves rapidly to keep all systems healthy and working in concert with all other parts of the institution. *(iv)* Information that relates work at office to activities at home and in the external world correspond to ego and self-esteem need. The information systems for the first three (Lower) levels are institutionally developed and controlled whereas fourth level represents the entry of personal information system. The first three levels are supported by formal systems like MIS and libraries which cater to the educational, recreational and/or research needs of the patrons. The fourth level, an individual-oriented unique system, necessarily has to

be designed and maintained by individual concerned. *(v)* Lastly, information concerned with integration of institutional information systems with personal systems so that individual needs and institutional needs are both met in a personally unique and highly autonomous way correspond to the self-actualization or self-realization need (Dow, 1987).

Most of the scholars, academicians, researchers, scientists and engineers create and maintain level four personal collection of information materials. These collections contain a wide variety of materials and often it is difficult to distinguish personal collection from a departmental or group's collection. They range from personal notes on desk diary, piece of paper, card or note book to technical report or a trade catalogue or a dismantled book, in variety, and few dozens to few thousands in number.

The area of intellectual work is extremely broad and it covers a large number of quite diversified activities like planning and control of work, technique of reading, studying, discussing, presenting, communicating, writing, drawing, designing, etc. This array of intellectual work should essentially include management of personal documentation and information system (Stibic, 1980, p 3-4).

4. Need for Personal Documentation and Information System

A survey (Northup et al 1983) by the American Council of Learned Societies revealed that academics in all fields except classics regard their personal libraries as more important than the institutional libraries they use. Research in information gathering, patterns of engineers has shown that the most significant source of information for an engineer is his own head and personal collection, followed by nearby colleagues and departmental collection. Reasons for such a high reliance of engineers on intra- and inter-personal sources of information and informal personal collection are many.

When an engineer seeks information he tries to match his cognitive, possibly, certain and clear need, for information with a source of information yet it cannot be ruled out that his choices normally slant towards sources and systems which conform to his habits, styles and ideo-syncracies than those sources which need him to adapt himself to them. He could also be biased, subjective, conservative and habituated to his personal information system and unintentionally develop an apathy towards new systems. Even if one has initiative to use an external information system, suitable system must be available and accessible in an easy to use way. On these scores also the personal information system being one's own property stands out. The perceived utility

of personal information system also plays a role in using it.

Further in the process of adaptation to different information systems, an engineer optimises his cost in terms of time and efforts and prefers sources which adapt to his needs than those require him to adapt himself. It is common experience that unless a person who wants information is fairly sure of getting it without much trouble, he is apt to do without it, if it is not essential, conforming to law of least effort',' why bother theory of information usage' (Cooper, 1978) and Mooers' law which states that users will utilize an information-service only when doing so costs them less than not using it (Mooers, 1960 p (II)). In such circumstances, relying on memory, skirting around the issue or making do with incomplete or vague information from easily available sources are usual behaviour. So personal collections are frequently and extensively used by engineers as they are always available to them whenever they want without any barriers or restrictions whatsoever.

The location of personal collection, time required to access it, the flexibility in using it and permanence nature of information received from it and arrangement of personal collection are in favour of the taste and idiosyncrasies of engineers.

The most vexing problem comes to engineers when their personal collections are disorganized or not so well organized and they are unable to find and retrieve quickly what they want. Efforts to accumulate huge personal collections becomes waste in such circumstances. A need for an organized or improved system is quite often felt by engineers. Some may also feel that they cannot afford more cost and time on personal documentation and information system and it may not even be worth. The fear that cost and time needed to improve a personal documentation system is more than what it would save is unfounded as long as an engineer restricts his input data to the minimum, does not replace existing and available services and systems and chooses an effective technique.

"Personal documentation ensures *re-usage of information* that would be otherwise lost. It saves not only *time wasted in searching* for documents, but also *time wasted in repetitive work,* in duplicate computations, programming, analysis, drafts, design or experiments" (Stibic, 1980, p 2). In addition to saving time and cost, by making access to information easier and faster, well organized personal information system facilitates and stimulates better and more intensive utilization of accumulated information that in turn stimulates creative thinking and that improved style of intellectual work of an engineer.

In nutshell, a personal documentation and information system *(i)* prevents information once gained from being lost, *(ii)* intensifies the use of available information resources, *(iii)* improves organization of knowledge, *(iv)* provides opportunity for creative use of information, *(v)* allows for linking of facts and ideas and *(vi)* helps to discover hither to unseen relations, associations and conclusions (Stibic, 1980, p 3).

5. Management of Personal Documentation and Information System

Though organizing a personal documentation and information system is different from organizing an institutional system, the underlying procedures are the same. Here, the system manager, system operator and system user is the same person. The main purpose of personal documentation of diverse collection of documents is to make 'rapid recall of relevant' information easy. Information collected by individual is useless if it is not easily and quickly accessible. Some of the basic principles of personal documentation (Stibic, 1980, p35-38) are:

(i) Personal documentation must help and not hinder: Time, energy and cost spent on creation and updation of a personal documentation system should be offset by improved quality and speed of retrieval as well as overall usage of the system.

(ii) Simple devices can provide a good service: Unlike large documentation and information centres and bibliographic databases, personal system is relatively small in size and hence simple devices like a peek-a-boo card technique or a sequential search in a PC would do.

(iii) Deviations from documentation standards are permitted: Since saving time without loss of clarity and usability is the supreme rule, deviations from standards are permitted in personal documentation. Probable exception may be in noting bibliographic data and providing citations and references in reports, papers or theses.

(iv) Subjectivity of personal documentation: The individual point of view should be respected in classification, indexing, recording and abstracting, amending titles for KWIC index preparation, etc. to give expression to individual's evaluation or criticism and remarks on usefulness, originality$_4$ credibility, reliability etc. of documents.

In brief, personal documentation is NOT professional documentation on a smaller scale; it has its own rules and laws, deviating substantially from established rules.

6. Conclusion

Engineers are sandwitched between 'information overload' and highly personalized need for information in their endeavour. They have resorted to their own special ways of seeking and gathering information required for their work, career and life. They are capable of developing and managing highly sophisticated documentation and information systems. But it appears that they are not fully convinced of the need and benefits of such as personal system, partly due to lack of formal training, orientation and exposure to such systems and partly due to habitual working on 'crises management' mode. Little thought, planning and efforts can greatly rationalize their information gathering habits and developing and organizing pertinent personal collection of documents.

As a logical extension of institutional documentation and information systems, engineers need to design, develop, maintain and manage their own personal documentation and information systems. Personal systems are based on the procedures akin to institutional systems but they are more flexible, individual oriented and hence not miniature institutional systems.

Lastly, an engineer who has an efficient personal documentation and information system not only derivesen ormous benefits by investing some time and energy, but also contributes positively to infor*mation* transfer and communication within the organization as well as to institutional documentation and information system.

REFERENCES:

1. Ackoff, Russell L. et al. Designing a National Scientific and Technological Communication System (The SCATT Report). Pennsylvania: University of Pennsylvania Press, 1976.

2. Anthony, L. J. et al. "Growth of the literature of physics." Reports of Progress in Physics 32 1969: 709-767.

3. Auerbach Corporation. DoD User Needs Study, Phase I. Final Technical Report, 2 vols. DoD 1151-TR-3 ; AD 615 501,2. Philadelphia, pa.: Auerbach Corporation, 1965.

4. ----------. "What's wrong with information retrieval?" Machine Design 38 (16) 7 July 1966: 106-109.

5. Bennett,Martin. "Application of library techniques outside the library service", Library Association Record 77(1)Jan 1975, p8.

6. Cooper, William S. "Why bother theory of information usage. "Journal of Informatics 2(l)

April 1978: 2-5.

7. Crane, Diana. "Information needs and uses". In Carlos A.Cuadra ed. Annual Review of Information Science and Technology. Chicago: Encyclopedia Britanica Inc., 1971. Vol. 6: 3-39.

8. De Buse, Raymond "So that's a book... advancing technology and the library" information technology and Library, 7 No.1 March 1988: 7-18.

9. Dow, Elizabeth. “Personal information STIBIC (V) “Tools of the mind: techniques and systems: the library role”, Library Journal 1 Nov 1987, 29-32.

10. Gerstberger, Peter G. and Thomas J. Allen. "Criteria used by research and development engineers in the selection of an information source." Journal of Applied Psychology 52 (4) August 1968: 272-279.

11. Hall, Robert W. Personal Effectiveness in Industrial Research. PhD Thesis. Indiana: Indiana University, Graduate School of Business, 1972.

12. Havelock, Ronald G. et al. "Planning for Innovation Through Dissemination and Utilization of Knowledge” Michigan: The University of Michigan,The Institute for Social Research,CRUSK, 1969.

13. Kemp,D.A. "Nature of knowledge: an introduction for libraries", London: Clive Bingley 1976.

14. Kunz, Werner et al. "Methods of Analysis and Evaluation of Information Needs: A Critical Review"Munchen: Verlag Dokumentation, 1977.

15. Langrish, L. et al. "Wealth from Knowledge"London: Macmillan, 1972.

16. Lucky, Robert W. “Lucky strikes... again” New York:IEEE Press, 1993.

17. Lundeon,Gerald. "Software for managing personal files", Database 12(3) June 1989, 36-48.

18. Mooers, Calvin N. “Mooer’s law or why some retrieval systems are used and others are not.” American Documentation 11(3), July1960: ii.

19. Northup, Dcana E, *et al.* “Characteristics of clinical information searching: investigation using critical incident technique”, Journal of Medical Education Nov 1983, p873-881.

20. Paisley, William J. "Information needs and uses". In Carlos A. Cuadra ed. Annual Review of Information Science and Technology. Chicago: Encyclopedia Brittanica Inc., 1968. Vol. 3: 1-30.

21. Price, Derek J. De Solla. "Is technology historically independent of science?: a study of statistical historiography." Technology and Culture 6 (4) Fall 1965: 553-568.

22. Raitt, David Iain. Communication and Information Seeking and Use Habits of Scientists and Engineers in International Organisations and National Aerospace Research Establishments in Europe. PhD Thesis. Loughborough: Loughborough University of Technology, 1984.

23. Rosenberg, Victor. "Factors affecting the preferences of industrial personnel for information gathering methods ". In Tefko Saracevic comp.& ed. Introduction to Information Science. New York: R.R.Bowker, 1970, 95-100.

24. Rosenbloom, Richard S. and Francis W. Wolek. Technology and Information Transfer : A Survey of Practice in Industrial Organisations. Boston: Harvard University, Graduate School of Business Administration, 1970.

25. Rowley, J. E. and C. M. D. Turner. "Dissemination of Information"London: Andre Deutsch/A Grafton Book, 1978.

26. Shuchman, Hedvah L. Information Transfer in Engineering. Rep-461-46-27. Connecticut: Futures Group Inc., 1981.

27. Skelton, Barbara. "Scientists and social scientists as information users: a comparison of results of science user studies with the investigation into information requirements of the social sciences." Journal of Librarianship 5 (2) April 1973: 138-156.

28. Soper,Mary Ellen. "Characteristics and use of personal collections", Library Quarterly Oct 1976, 397-415.

29. Stibic,V. "Personal documentation for professionals : means and methods", Amsterdam: North Holland 1980.

30. Stibic, V. "Tools of the mind: techniques and methods for intellectual work", Amsterdam: North Holland 1982.

31. Wilkin, Anne. Information Needs of Practitioners : A Review of the Literature. BLR&DD-5611. London: The London University, 1981.

32. Wilson, Patrick. "Public Knowledge and Private Ignorance : Toward a Library and Information Plolicy"London: Greenwood Press, 1977.

33. Wolek, Francis W. "Engineer : his work and needs for information". In J.B.North ed. Proceedings of the 32nd Annual ASIS Meeting. New York: Greenwood, 1969.

34. ----------. "Complexity of messages in science and engineering : an influence on patterns of communication ". In Carnot E.Nelson and Donald K. Pollock eds. Communication among Scientists and Engineers. Massachusetts: D.C.Heath Lexington, 1970. 233-265.

35. ----------. "Preparation for interpersonal communication." Journal of the American Society for Information Science 23 (1) January-February 1972: 3-10.

36. Wood, D. N. and D. R. L. Hamilton. Information Requirements of Mechanical Engineers : Report of a Recent Survey. London: The Libary Association, 1967.

* About the Author

Dr. M.S.Sridhar is a post graduate in mathematics and business management and a doctorate in library and information science. He is in the profession for last 35 years. Since 1978 he is heading the Library and Documentation Division of ISRO Satellite Centre, Bangalore. Earlier he has worked in the libraries of National Aeronautical Laboratory (Bangalore), Indian Institute of Management (Bangalore) and University of Mysore. Dr. Sridhar has published four books ('User research: a review of information-behaviour studies in science and technology', 'Problems of collection development in special libraries', 'Information behaviour of scientists and engineers' and 'Use and user research with twenty case studies') and 74 research papers, written 19 course material for BLIS and MLIS, presented over 22 papers in conferences and seminars, and contributed 5 chapters to books. **E-mail:** sridharmirle@yahoo.com, mirlesridhar@gmail.com, sridhar@isac.gov.in **; Phone:** 91-80-25084451; **Fax:** 91-80-25084475.

PS: Further notes used in lectures (Stibic, 1980):

5.1 Rational Information Gathering Strategy

We have already seen the general information gathering habits of engineers including those which are not always rational. Having clarity and answering the questions such as what type of documents and information are relevant, what is the area of interest, what forms and what channels and sources to be consulted, etc. are necessary before start gathering documents and information.

5.1.1 Clarity on Primary area of interest

Except probably managers and teachers, all other engineers are specialists having their own core area of interest within the profession. Occasionally, some peripheral areas do attract their attention in addition to a more dynamic temporary `zone of interest' within core area of interest and work. At times, an engineer may reach a stage of `over-specialisation' which may substantially reduce his chances of `lucky accidents' of discovering insightful information unexpectedly. On the other hand interdisciplinary and multidisciplinary character of many areas of work really make it difficult to define a clear boundary of core area of interest of engineers. Balancing with `over-specialisation' and `jack of all trade' trends and drawing a fairly clear `radius of interest' is needed.

5.1.2 Know the Information Sources to be Used

As mentioned earlier there is a plethora of formal and documentary sources of information to be used by engineers. In addition, inter-personal and informal sources also provide information. Irrespective of the type of source, some information come to you unsought and others need to be deliberately searched and gathered.

Formal and documentary sources of information are available at three levels:

(i) Primary sources or documents like books, journals, standards, reports, minutes of meetings, etc., which contain original information.

(ii) Secondary sources or documents like library catalogues, abstracting and indexing journals, online information system (majority of databases housed in them), index to correspondence, etc., which `provide information about information'.

(iii) Tertiary sources or documents like bibliography of bibliographies, directory of libraries, etc., which help us to know about secondary sources.

Large volume of information is received via informal channels like personal and telephone contacts, at meetings, conferences, lectures, etc., internal documents like letters, circulars, company regulations, memo's, notes, minutes, contracts, project documentation and finally self created notes, drafts, remarks, drawings, calculations, laboatory diaries and logs, excerpts from books, etc.

5.1.3 Optimise Ways and Means of Gathering Information

There are a number of ways of gathering information. A suitable combination of them has to be adopted depending on need, available time, resources and facilities. Knowing high ways and byeways of secondary sources of information and approching through secondary sources saves a lot of time and efforts. If the purpose is to have a very comprehensive collection of information , it is worth going for online search or systematic search through abstracting and indexing journals. On the other hand, if one wants some basic information on a specific subject, it may be better to directly browse the shelves of a library without even going to card catalogues.

5.2 Information Collection and Recording

Out of the information encountered, one may ignore a large majority of irrelevant, retain `documents' which can be owned and study `documents' accessed in public systems like libraries. Methods employed to record information studied in documents are: underlining

or highlighting significant words or passages, attaching personal marginal notes or flags, making a short abstract and/or compiling excerpt from the document being read.

Flexible conventions of `shorthand' marking and writing system may be employed in recording information. However, strict adoption of abstracting rules may not be necessary. The most important of all these methods of extracting information is recording complete bibliographical data in an abbreviated way about original documents including its whereabouts like in which library it was consulted and what is its classification or accession number.

It is always advisable to make such notes, abstracts or excerpts on standardised preferably preprinted cards (say 5"x8") on the spot. Recording in the final way with fixed form, sequence, layout and punction and avoiding notes making on bus tickets, canteen coupons, etc. in an adhoc way are necessary.

Following standards for recording bibliographical data about documents becomes imperative, not only to avoid pitfalls in later retrieval, but also to conform to any of the standards or style manuals expected of you in your paper or thesis. For example procedure to identify and enumerate author's name (particularly surname), title of document, imprint (place of publication, name of publisher and year of publication) for books and monographs, standard abbreviations for journals, locating volume number, issue number and date of publication of journals, etc. are to be learnt. The individual researcher or engineer or scientist is not expected to have mastery on rules for recording bibliographical data, but a fair knowledge is necessary to maintain consistency.

Latest developments in information technologies have enabled scientists and engineers to down load their online search results from databases kept on host computer or from CD-ROM databases and directly import them as input to their personal file on PCs (Lundeon, 1989).

5.3 Organisation of Personal Collection of Documents

The collection of documents in the possession of an engineer is usually the main component of a personal documentation and information system. In the broadest sense of the word the documents not only include sources of information mentioned earlier but also microforms, audio-visual materials like slides, video and audio tapes and discs, computer outputs on

paper, floppy or magnetic tape and even physical objects like pictures, specimens, sculptures, postage-stamps, coins, etc.

Physical storage and arrangement of personal documents depend on physical properties of documents. For example, books go on shelves, letters and memos in files, microforms and AV materials in their own specific form of storage. Each variety has to be filed and arranged in separate group by its own criteria. Often, an identification code or number is given to each document in order to relate this to a record of its surrogate in the retrieval system. Identification code consists of a document form code represented mnemonically like B for books, M for microforms, S for slides, etc., prefixed to an unique identification code like sequence number. Physical form in one's personal collection may be different from original form of a document.

Individual need to decide which type of documents are to be systematically stored with proper retrieval system developed and others kept informally. Similarly whether or not to combine documents providing profession and work related information with documents providing private and personal information (eg. personal letters) is to be decided in advance. It is also necessary to decide whether a temporary special purpose system like that needed for writing a paper, report or thesis is being designed or a permanent one is being designed.

5.4 Development of Information Retrieval System

If the collection is small search frequency is low and search requirements are simple the physical arrangement itself serves as retrieval tool. Physical arrangement is linear and hence a single criterion like author or subject or chronology could be used. Heterogeneous collections like journal articles, drawings, microforms, slides, etc., cannot be placed in one sequence in one container. Polythematic documents cannot be easily arranged by their subjects. Hence, an auxiliary or secondary or surrogate system has to be created in which each document in personal collection is represented as a record (eg., a catalogue card, an entry in a database in magnetic system). The set of records with additional features of retrieval techniques embedded would constitute a retrieval system. In other words, it is a tool to find a certain document or a group of documents concerning one or more subjects or characteristics and provide information about their location in the physical storage.

Asymmetries between the collection of personal documents and the retrieval system consisting of surrogates are:(i) Some types of documents may not need recording in retrieval

system as they may not need extensive searching and retrieval. (ii) some records in retrieval system like facts, data, remarks, notes, etc., may not have corresponding document in personal collection but surrogate itself may contain the entire document. (iii) Some records like those explained in (ii) above might have already been excerpts from documents external to personal collection (eg. library books, journals, etc.).

5.4.1 Record Structure and Bibliographical Data

Though record structure in a large database and retrieval system is exhaustive and lengthy, personal system has to have a modest record. A record in the retrieval system consists of a number of elements called fields which contain:

(i) Data identification and description like unique record number, date of input, status of the record, etc.

(ii) Basic identification data or bibliographic data like author, title, date of publication, etc. and data about the source of the document like volume number, issue number, title of a journal, etc.

(iii) Location (or a pointer) to the original document

(iv) Physical properties of the document

(v) Subject of the document in terms of classification number and keywords or descriptors

(vi) Abstract or summary, etc.

5.4.2 Subject Description: Classification and Indexing

One of the most frequently asked query to any information retrieval system, whether personal or institutional, is subject query. Due to polythematic nature of documents, labelling each document to represent one subject or theme is inadequate and lead to loss of information or failure to retrieve all the information on a given theme in the system. Two ways of overcoming this is by classifying and indexing the documents.

For personal documentation and information system, a special purpose classification scheme like that used by abstracting and indexing journals (eg. Engineering Index) or a subset of a library oriented general or specific classification scheme be adopted. Though both classification and indexing try to analyse a document and label its contents, the classification with its artificial language of notation helps to sequentially arrange the documents in storage. However, indexing with natural language words (either free text or controlled vocabulary) enhances flexibility of search strategies and retrieval capabilities to

some extent. You may find it more easier to adopt an indexing system with subject thesaurus than a rigid hierarchical classification scheme with notation. The main advantage of classification is its logical structure, dividing a subject field from the generic into the specific, bringing related concepts together, and making orientation easier. The primary advantage of indexing is its flexibility and freedom. One may adopt both notational classification scheme and an indexing method combining advantages of both.

A controlled indexing language consists of a limited number of previously selected, semantically defined and structured collection of indexing terms called descriptors free from synonyms and homonyms (ie., a thesaurus) with rules for its application. Computers have not only enabled handling free text indexing workable but also are helping in automatic indexing. Free text indexing systems are in a way `post-controlled' vocabulary indexing systems. It is also necessary to distinguish between pre-coordinated and post-coordinated indexing and retrieval systems. In a post-coordinated system, any combination (think of Boolean logic) of keywords can be used for searching. With the advent of computers, post-coordination systems have become more easier and popular.

Choice of a classification scheme, indexing system and a thesaurus for personal documentation and information system depends on the size of collection, complexity of searches desired, availability of hardware (card index, uniterm indexing system, peek-a-boo card, PC) and softwares (SCIMATE, ASKSAM, Notebook, Pro-cite, Reference Manger, Bib/Search), nature and breadth of area of interest, etc. (Lundeon, 1989).

5.4.3 Technical Gadgets and Means for Personal Documentation

The gadgets and means used to maintain the retrieval part of personal documentation and information system varies from a traditional card index to a sophisticated PC with versatile software for managing personal files. Even though PC based systems have made other personal systems outdated, a selected few techniques are explained below.

5.4.3.1 Card Index

A card index can be developed using standard, preferably preprinted, cards (3"x5" or 4"x6") together with leading cards or guide labels and a card box. Surrogate information of each document is written or typed on standard cards and arranged in a logical/alphabetical sequence of keywords.

5.4.3.2 Terse Documentation

For a small collection of documents or for an adhoc documentation system, the simplest, cheapest and most modest form of personal documentation is to use a system of loose leaves with information recorded in telegraphic way and with pointer to location of original documents. For larger personal collection, maintenance and updation in loose leaf form becomes difficult.

5.4.3.3 Peek-a-boo Cards or Optical Coincidence Cards and Edge Notched Cards

This technique is based on post-coordinate indexing system particularly uniterm indexing system and it is useful where the personal collection is not very large. Each document in the collection is assigned a serial number (document number) and a couple of `uniterms' or keywords describing the subject of document. In case of uniterm indexing system each keyword is written on the top of a standard card which has ten preprinted columns, each of which is meant to write document numbers of documents dealing with the subject represented by the keyword. Document numbers are written in one of the ten columns depending on ending digit of the document number. For example document number 354 will be entered on 4th column. Within each column the nubmers automatically stand in ascending order as numbers are assigned to documents in the chronology of their taking into system.

Retrieval from uniterm index is based on choosing the appropriate keywords/terms like DOCUMENTATION and INDUSTRY and matching for document numbers on both cards to select common numbers which represent documents dealing with DOCUMENTATION and INDUSTRY.

Optical coincidence cards also work very similar to uniterm indexing system. Here a matrix (similar to graph sheet) is provided with as many as 2000 tiny squares, each representing a document number from 1 to 1999. While indexing document number 354 for keywords DOCUMENTATION and INDUSTRY, 354th square on cards having keywords INDUSTRY and DOCUMENTATION are punched so that when we take these two cards and inspect we find light passing through 354th square to indicate that document numbered 354 deal with these two aspects. In the above systems, instead of keyword, a property of an object or a name of a person or any other aspect can also be used.

Another system similar to optical coincidence cards is `edge notched cards' wherein each card has a provision to punch adjacent to a series of holes on the periphery of the card

representing a subject/ keyword. By inserting a needle through the appropriate holes of the cards in a deck, the cards with edges notched off will fall down indicating document numbers dealing with the subjects.

In all the above systems, apart from handling a limited number of documents and keywords, the additional limitations are inability to change or reindex a document, use logical OR and NOT and profound false coordinations like searching for SCHOOLS and LIBRARIES leading to retrieve documents dealing with both school libraries and library schools.

5.4.3.4 Computer (PC) Based Retrieval System

Personal information management is increasingly becoming a computer (PC) based endeavour with off the shelf availability of a wide variety of software packages for personal file management (Lundeon, 1989). PC based personal information systems easily lend themselves for importing information from online information systems, integrating with institutional information systems and extend help in report and paper preparation (i.e., publishing) activities.

In view of such a link between personal system and international systems, personal databses are becoming an indispensable tool for engineers and scientists who wish to gain and maintain adequate and efficient access to the literature. Though serendipity and holism of manual system do not exist, computer based systems make data input and file management easy in addition to unbelievably versatile and fast search features. They allow not only for importing and exporting records form personal files but also allow for updating, breaking and combining files. They provide flexible templates and worksheets to make data input easy. Extensive help facilities and user documentation are provided in these packages. Flexibility in taking output in different formats on both magnetic media and paper is maintained. Above all, if not as extensive as online information systems, they also provide versatile search features like set building, boolean logic, truncation searching, field directed searching, proximity searching, parenthetical grouping of search terms, searching through inverted indexes etc. With these advanced features of computerised personal file/database handling, the earlier uses of computer to generate printed indexes (KWIC,KWOC, etc.) and bibliographies are quite obsolete.

There are already hundreds of international and national databses available to engineers to access, search and obtain information right from their own PC linked through telephone

line. Advances like PAD system of P&T together with advances in computer, communication and network technologies are making online access to these databases simple and cheap. However an orientation training on using online information systems and fair acquaintance with the structure, scope and other features of relevant databases may be necessary.