TRANSFORMING BIBLIOGRAPHIC DATA BASES INTO BIBLIOGRAPHICS

V.L. KALYANE
Turning Point Centre, A-6/07,
Kendrya Vihar, Kharghar,
Navi Mumbai - 410 210
Maharashtra.

ABSTRACT
Bibliometric data obtained from the MEDLINE bibliographic database was processed into various bibliographies for easy perception and quick understanding of the information generation and dissemination phenomenon. This methodology may be used as a basis for librarians to perform effectively their prescribing and subscribing roles; for researchers to finalise R & D communication dissemination strategy; for human resource managers to evaluate and forecast productivity trends of researchers by channelising their energies in right directions; and for policy makers to derive technical intelligence from the already published literature resulting into maximum utilisation of information generated at global level. A new profession of bibliographic database processors is now emerging.

1 Introduction
The term "Bibliographics" is used here to mean the various graphic representations that can be made out of any bibliometric study. Bibliometrics/ Scientometrics/ Informetrics as a discipline uses quantitative methods to map the status and development of scientific knowledge and technical innovation communication productivity through research output parameters such as number of publications or patents. Due to the availability of fairly comprehensive databases on publications and patents, it has now become possible to arrive at fairly rapid estimates of the nature and extent of scientific or technological activity at the national, regional and global level. This makes scientometrics and informetrics (fig. 1) an important tool for managers, funding agencies and institutions. Information on the nature of collaborative activity, multivariate comparisons between countries and between states and regions within a country in terms of their research priorities, relative contributions made by major scientific organisations, universities and industrial houses, can also provide useful input to policy planners and managers. New text and language processing techniques such as co-word or co-authorship analyses have enabled the mapping of cognitive linkages, with a potential of indicating emergent areas of activity, with implications for policy.

Research information is published through scientific journals, reports, books and patents. The number of publications is growing at an ever increasing pace. To facilitate the handling of the multitude of information, refer-
ence journals are published in which the title, a short summary and bibliographic references to the article, are listed. These references, themselves representing an immense amount of information, are made available in database. The largest STN databases (by number of records in millions as on December 1997) are: WORLDCAT (35), INPADOC (26.8), REGISTRY (16.9), SCISEARCH (15.7), CA/CAPLUS (14), BIOSIS (11), MEDLINE (9.3), WPINDEX (8.3), BEILSTEIN (7.1), EMBASE (7), PROMT (6.1), INSPEC (5.7), JAPIO (5.5), INVESTEXT (4.4), AND COMPENDEX (4).

The clarion call [1] for using bibliographic databases to identify and harness information sources and human resources remained silent till reprinting of the same paper in the International Information Communication and Education journal in 1994. Awareness has now grown further after the MEDLINE study [2].

2. The Medline

MEDLINE database contains information on every area of medicine. The MEDLINE corresponds to the Index Medicus, the Index to Dental Literature, and the International Nursing Index. Bibliographic information, indexing terms, abstracts, chemical names, and CAS Registry Numbers are all searchable. Online thesauri are available for the Medical Subject Heading (/MN), Controlled Terms (/CT), Controlled Term 95 (/CT95), and Chemical Name (/CN) fields. A learning file LMEDLINE is available. MEDLINE (Producer: US National Library of Medicine, USA, Coverage: 1966 to date; File Size: 9.3 million records; Updates: Weekly; Content: Medicine; Language: English;) is a comprehensive bibliographic database indexing more than 3,700 biomedical journals.

MEDLINE guide topic fields are: AB (Abstract), AD (Address of Author), AI (Abstract Indicator), AN (MEDLINE Accession Number), AU (Author/s), CM (Comments) CN (Contract or Grant Numbers), CP (Country of Publication), GS (Gene Symbol), ISSN (International Standard Serial Number), LA (Language of Article), MESH (Medical Subject Headings), MIME (Minor MeSH Headings), MJME (Major MeSH Headings), NM (Name of Substance), PS (Personal Name as Subject), PT (Publication Type), PY (Publication Year), RN (CAS Registry Number of EC Number), SB (Subset), SI (Secondary Source Indicator), SO (Source / Bibliographic Citation), TG (Checktags), TI (Title), TO (Original Title), UD (Update Code), and CITN (Citation).

3. Informetrics on accidents and trauma

Presentation of the paper 'Informetrics on accidents and trauma' [3] generated active interest among the participants of the 49th FID conference and congress at New Delhi. The conference concluded with emphasis of use of information for human welfare. The ever alert and preplanned mechanism of springing up into immediate action during the 'golden hour' (first hour after the accident) may result in the maximum salvage possible. We consider it a prime priority to effectively put to use information facilities and all up to date information technologies to save the life of persons affected by accidents and trauma and rescue their precious life. Encouraged by good response, further, it was decided to transform the bibliometric database into bibliographies. (fig. 2)
4. Methodology

'Accidents and Trauma' papers retrieved from the MEDLINE database by publication year were 140 papers during 1995, 182 papers during 1996, and 48 papers up to September 1997, when this work was initiated. These 370 papers were studied for various fields. Frequency counts resulted in bibliometric database. Herewith, the bibliometric database is transformed into the bibliographic formats.
5. Bibliographic results

The bibliographic results presented in fig 3 to 13 are self explanatory.

6. Discussion

(a) Researchers often find it difficult to cope up with the wide scattering of publications in their domain of interest. And, no library can claim to subscribe to all journals publishing research papers in a particular domain, however microtopic it may be. Hence, it becomes necessary to find out and standardise a methodology to identify core journals for rationalised subscription.

(b) Bibliographic presentations to the Library Committee can facilitate decision making, for example, while subscribing to journals in a domain. Decision can be taken to subscribe to only core journals forming the Nuclear Zone of high prolific productivity journals in the domain under consideration.

(c) Same methodology can hold good for the identification of expertise at global level, which may facilitate close and timely interactions among specialists in this era of mega-authorships. Causes being: Multidisciplinary research activity, Multidisciplinary application of research results, Resource sharing, and
Advanced information technology. This facilitates team approach at global level.

(d) While submitting research project proposals, project leaders should review all up-to-date bibliographic database in the domain and transform it into quantitative formats and bibliographics. This would help them to defend their projects and convince their requirements forcefully to the funding agencies.

(c) Quinquennial review teams or referees can follow all above mentioned parameters in addition to local specific criteria for the domain, if any. The judgment of qualitative elements depends much on the intuition and knowledge of the experienced referee. An analytical synopsis, of the scoring rate of the different relevant aspects to the dossier in the form of a graph promotes the objectivity through Research Proposal Profile. Evaluation of research proposals and future projects will depend on the subjective/qualitative criteria such as originality and feasibility, and the quantitative analysis will serve as a background. An equilibrium between information of past research and on prospects is neces-
Figure 9: Bradford-Zipf bibliograph and inset cumulative publications in core journals having research papers on 'Accidents and Trauma' retrieved from MEDLINE database on CD-ROM (1995-97)

Figure 10: Number of pagewise frequency of research papers on 'Accidents and Trauma' retrieved from MEDLINE database on CD-ROM (1995-97)

Bibliometric data provides a "monitoring device" for science policy and research management. They can as be used to explore large sale collaboration and team work, in specific domains.

Kelpro Bulletin, 5 No (1) June 2001
have already become a standard method of documenting the process of information generation and dissemination.

(b) Researchers should be encouraged to publish their research in the high 'impact factor' and high 'immediacy index' journals, or at least in the journals having high 'prolific productivity' of publications in their own research domain. Prerequisite to do so is to identify such journals and bring it to the notice of the researchers concerned. Hence, bibliographic databases will be very useful in such an endeavour of prime importance to finalise communication dissemination strategy.

(i) Bibliographic database can be used to evaluate and project productivity of an individual scientist also [12-31]. Further work is needed in this field so as to enable forecasting of the productivity of young scientists so as to tap their talent in real time.

(j) Research Impact Assessment (RIA) by Comparative Intelligence (CI) professional who has interest in these fields has many options for proceeding further from the map, depending on this person's specific interests. For example.

- reading of the most highly cited papers,
- overview of the current literature,
- perusing the journals which contain the highest frequency of recent publications,
- contacting experts in a particular thrust area or technique and
- contacting the institutions identified with given techniques in the theme.

The key conclusion is that, starting from the raw data, the analyst can generate any cross-cutting relationships desired to proceed further in specific directions of personal interest.

7. Database tomography

Database Tomography can be used to derive technical intelligence from the published literature. Database Tomography is a patented system for analysing large amounts of computerised textual material. It includes algorithms for extracting multi-word phrase frequencies and performing phrase proximity analyses. Phrase frequency analysis provides the pervasive themes of a database, and the phrase proximity analysis provides the relationships among the pervasive themes and between the pervasive themes and sub-themes. One potential application of Database Tomography is to obtain the thrusts and interrelationships of a technical field from papers published in the literature within that field [32].
8. Production of papers

As per the Research Output Database [33] total number of biomedical papers have increased by one-third between 1988 and 1995; London, Cambridge, Oxford and Edinburgh produce the most number of papers. Belfast and Leicester are the areas where the output is increasing rapidly; UK is relatively strong in tropical medicine and arthritis research and its shares of papers in cardiology, genetics, nursing and ophthalmology have increased significantly in the eight years; both national and international collaboration have increased, with increase in the average number of authors, addresses and funding sources acknowledged on each paper. UK collaboration with Portugal and Spain is increasing; the government and private non-profit agencies are involved in funding equal number of papers.
India should encourage studies in the areas of patent bibliometrics and the economics of R & D [34]. India should have a full-fledged observatory for science and technology, similar to the one in France headed by Remi Barre. This may facilitate evidence-based decisions on funding.

9. Human resources development

Information needs of endusers specific to a particular domain should be satisfied. Even it may enable positioning of the trustworthy information professionals to take up the information prescription providing role for study in a specific domain. Human resources development in every specific field should be given priority. Potential users include doctors, nurses, laboratory scientists, pharmacists, occupational therapists, speech therapists, ambulance drivers, general administrators, computer scientists, architects, medical records officers, and the patents themselves. The provision of information to clinicians at the time and place of need can make a visible and measurable (“life and death”) difference in the outcome of patient care.

10. Literatherapy and infotherapy

In psychosocial approaches, the therapist - psychiatrist, psychologist, psychiatric social worker, or psychiatric nurse - establishes a therapeutic or professional relationship with the patient. It may involve reading, writing, story telling, speaking, nonverbal communication, music therapy, literacy materials, photographs, audiocassettes, audio-visuals, etc. Mental health centres, de-addiction centres, outpatient services, nursing homes, counselling centres, schools counsellors, AIDS rehabilitation centres, juvenile delinquents, etc. are the target groups.

Preventive medical care can help by providing information to the patient, not only for illnesses, but for prevention and wellness. Indeed, information professionals are emerging as bibliotherapists. Literatherapy has evolved from Bibliotherapy into Multimedia Communication Therapy or Infotherapy. "Information therapy" or "Infotherapy" is a domain for supplying patients with health information, enabling them to make informed decision about their health and care, participate in their own well being, and thus, decrease the utilisation of health care resources. The basis of this enduser revolution lies in bibliotherapy, patient education, consumer health trends, patients rights, right to information, and the freedom of information.

The electronic age with its information superhighways is expected to provide a unique delivery system for Infotherapy instructions for self - treatment, whether printed, presented via computer or by audio visual menus, are effective in the management of phobias, panic disorders, other anxieties, depression, bulimia nervosa, alcohol problems, nicotine abuse, myocardial infarction, AIDS, compliance problems, and the counseling to relatives of patients.

11. The human touch

Today medical care revolves around the hinges of constant uninterrupted flow of information and hi-tech advancement. ”But let us remember that there is one technology that will never fade out to change. That technology has never failed for centuries, and will perhaps never do. This technology has everything to do with and deals entirely with, the art of human touch. The touch not necessarily physical but a touch of compassion, a touch of dignity and the touch of humanity and its relationships [35]". The recommended ABCD rule is : ABILITY ; BEING THERE ; COMMUNICATION ; and DOCUMENTATION.

12. Conclusion

Bibliographic databases can be processed into bibliometrics and bibliographics which has its use as a standard methodology to visualise and facilitate quick decision making process. It can also help information professionals to effectively exercise their role as information managers and consultants for prescribing study material and for subscribing core journals.
Informetrics has its utility in the decision making process. It can identify strengths and weaknesses in a research domain. Impact of Information Technology in reducing the barriers of distance by shrinking the globe into a globule is evident. Trespassing all of the unnatural barriers made by human beings (due to language, religion, geographical, political, knowledge, etc.) is the highest achievement of Homo sapiens of the 20th century evolving themselves into the generations of *Homo scientificus* or *Homo technoficus* or *Homo infotechnoficus*.

**References**


