
Accessing Indian Numeric and Statistical Data: a critical study of the Suprastructure and Infrastructure in India

Abstract

The use of numeric and statistical data for macro and micro level decision-making, development planning, and socio-economic research has always been critical for governments, international organisations and society at large. While the developments in information and communication technologies (ICTs) have paved way for timely access to validated numeric data on the one hand, these have also posed many challenges before library and information professionals to exploit the opportunities made available by the ICTs to manage and disseminate the numeric and statistical data efficiently and effectively.

In India, numeric data have been published regularly, mainly by the Central and State Governments. Numeric data published by the government ministries, departments, and other agencies is largely print-based, basically brought out in the form of reports, as well as ad-hoc and regular publications. Although the technology for digital storage and dissemination of numeric data had been available for a long time, its importance seems to be realized only recently by the Government. Although the NICNET (a national network for dissemination of the government data and information) has been operational since the 1980s, a comprehensive National Policy on Dissemination of Statistical Data (NPDSP) was announced only in May 1999 by the Government of India (GoI). This paper critically evaluates the provisions of this policy and also looks at the infrastructure being made available in the form of NICNET.

Though a few efforts have been made in India (e.g. by the Reserve Bank of India, and Registrar General Office) to digitise the existing numeric and statistical data, access to the digitised data is not adequate and reliable. In fact, the conduit is available, but the content is far from satisfactory. There is a strong need for assessing user needs, enhancing their awareness, and consolidating the efforts of various ministries, departments and other source agencies to make the collection, validation, organisation and dissemination of numeric and statistical data efficient and effective. A beginning has already been made in this direction by making the Department of Statistics, Government of India responsible to serve as a nodal

*by Dr. Jagtar Singh & H. P. S. Kalra**

agency in this regard. An effort has been in this paper to raise a few issues and put forward a few suggestions to ameliorate the situation and also to enhance global community's awareness regarding the state-of-the-art in India.

Introduction

The importance of reliable and accurate numeric and statistical data has been duly recognised by scientists, social scientists, planners, and decision-makers, entrepreneurs, and governments. Now-a-days, a lot of time and manpower is invested in collecting, analysing, and disseminating information in quantitative form. As the number and variety of data sources increase, be it an academic, industrial or government setting, the process of providing access to data gets complicated. The use of traditional methods of data collection are tedious and cumbersome. All sorts of clarifications and explanations regarding the data need to be mentioned to all data source agencies/individuals time and again. Dissemination of the data analysed (in some cases unanalysed also) to all concerned poses problems for the library and information professionals. The problems are compounded with the increasing demand of users for numeric data customised according to their needs, or in ready to use formats. Users from academic organisations, bureaucracy, government, business, industry, and non-government organisations rely heavily on numeric and statistical data for their work. Though the use of computers and storage media on the one hand has solved the problems of storing and analysing large quantities of such data, it has given rise to operational problems on the other.

With the convergence of computer and communication technologies and the emergence of networks, information handling processes have undergone a profound change. Developments in information and communication technologies (ICTs), particularly in the 1990s have changed very significantly the way we manage information, including statistical information, right from its generation to use. Now it is possible to access numeric and statistical data via the networks, particularly the Internet. In the networked environment, access to, validation, security, and updating of statistical data are some of the challenges facing the library and information professionals.

State-of-the-Art Report

In the context of developments in ICTs at the global level, the situation in India regarding the availability of and access to numeric and statistical data is not so encouraging. Numeric data are collected, analysed, validated and disseminated by ministries, departments, and agencies of the central and the state governments. These data are published as reports, ad-hoc and regular publications mainly in the printed form (e.g., Publications of Central Statistical Organisation, Human Development Reports for various years produced by the state governments of Karnataka, and Madhya Pradesh, and the Statistical Abstract of Punjab published by the Economic and Statistical Organisation, Punjab).

The Statistics Wing (SW) of the Ministry of Statistics and Programme Implementation (MoSPI), Government of India, (earlier the Department of Statistics) is the apex body for official statistical system in India. Two organisations under it, namely the National Sample Survey Organisation (NSSO) and the Central Statistical Organisation (CSO) are responsible for carrying out socio-economic surveys, field work for surveys, training, dissemination and publication, and coordination of statistical activities. Since government agencies are largely responsible for collection and dissemination of numeric data, there is generally a big time lag in the publication of numeric data. Appropriate technology was available in India for quite a long time, but has not been used optimally for quick analysis and timely dissemination of such data.

Other agencies, such as the Reserve Bank of India (RBI) and the Registrar General's Office, are also engaged in providing financial and census information respectively in statistical form. Statistical publications by various government departments and agencies are in the broad areas of national income, industry, banking and finance, census, trade, agriculture, labour, and education. In the last few years however, the Computer Centre of MoSPI has also been instrumental in making available some of the publications of NSSO in magnetic tapes, e.g., the Annual Survey of Industries 1995-96, and the Report on Energy Statistics, 1998-99. As far as the bibliographic control of publications containing statistical information is concerned, there is no single source in printed form, though efforts have been made, e.g., the Statistical System in India, 1989; the Catalogue of Government of India's Civil Publications; and the announcements in newspapers entitled 'List of new Arrivals' by the Controller of Publications, Government of India.

National Policy on Dissemination of Statistical Data

The Government of India (GoI) seems to have realized only recently the importance of disseminating the numeric information in digital form. In May 1999, The Government of India announced a comprehensive National Policy on Dissemination of Statistical Data (NPDS) and specific

guidelines for release of data. The provisions of the policy are reproduced as Annex 1 given at the end of this paper. The policy of the GoI is a welcome step in the direction of dissemination of numeric data in digital form, but certain lacunae in the policy are worth discussion. Provisions of NPDS have been examined critically below.

Under clause (i) of the policy, it is written that the data would be available to users in the form of hard copies and magnetic media. As the policy was announced in May 1999, developments in data storage technology at that time were ahead of magnetic media. Infrastructure for data storage in optical media (CD-ROMs, DVDs) is also available with government departments and agencies. A comprehensive term to include optical media would have been better. Similarly, provisions for availability of validated data via networks, particularly the Internet, could also have been incorporated in the clause.

Clause (v) of the policy will act as a hindrance in quick and timely release of data. Under this clause, it is said the data users shall have to wait for three years after the completion of field work to get the data, in case the reports based on survey data work can not be released by the concerned government agencies earlier. Moreover, the access mechanism for data collection in such a situation has not been specified. Other clauses in the policy such as clause (vi) and clause (viii), deal with non-commercial use of data, and the Department of Statistics (DoS), GoI, acting as the nodal agency for dissemination of statistical data, respectively.

Statistics Wing (SW) in the MoSPI (earlier the DoS) has been entrusted with the responsibility of data collection from source agencies; the organisation of data and ensuring its quality; conducting studies regarding data collection and validation for each type of data source; and the dissemination of official statistics under clause (viii) of the policy, and clauses (i), (ii), (iii), and (iv) of the guidelines for release of data. Regarding secondary publications, e.g., bibliographies, indexes, and directories in electronic form, provision has been made in the guidelines for release of data (clause viii), but not much information is available on the web site of SW in MoSPI (<http://www.nic.in/stat>)

Though dissemination of statistical data is the focus, mechanisms have been spelt out only for release of data, and not for its dissemination. SW can act as disseminator of statistical information, to one and all only if it has a network of branch offices. It does not have any such network, and under the present provisions of the policy and guidelines, therefore, either the dissemination activity, largely print-based, will be centralised or will have to rely on some other government department/agency. Public libraries offer such a network in almost all the states and union territories in the country. Traditionally, public

libraries have been playing the role of information disseminators. Many states in India have working public library systems today. Ten out of 28 states have public library legislation. The policy could have incorporated the role of public libraries in collaboration with SW for dissemination of statistical data.

National Statistical Commission

The National Statistical Commission (NSC) was setup by the GoI in January 2000 to critically examine the shortcomings and deficiencies of the present statistical system with a view to recommending measures for a systematic revamping of the system. NSC has 12 members under the chairmanship of Dr. C. Rangarajan, the Governor, Andhra Pradesh. Terms of reference of NSC include timeliness, reliability, and adequacy of statistical system in India; dissemination of these statistics; and the coordinating mechanism using statistical information for policy making and planning. Although the commission was expected to submit its report to the Government within a period of twelve months from the date of its establishment, it has not submitted the report. (The text version of the information on the NSC downloaded from the MoSPI web site is enclosed as Annex 2.)

Infrastructure for Dissemination of Digital Data

The infrastructure for storing and disseminating the data in digital form was set up in 1977 by the Government of India under the Department of Electronics. Later, it was entrusted to the direct control of Planning Commission. Recently, recognising the growing importance of ICTs, a separate Ministry of Information Technology (MIT) was created by the GoI, with NIC and its infrastructure under the MIT. NIC has a large distributed network infrastructure, known as NICNET with its nodes in all parts of the country, including many remote areas. The conduit, i.e. NICNET, is available, but the availability of and access to the content, i.e. the statistical data, via NICNET is not easy. Effort of the NIC to provide statistical, numeric and other factual information through General Information Service Terminals of NIC (GISTNIC) is not successful as these are placed in the district commissioners' offices. Even the web site of GISTNIC does not provide users with much information (<http://gist.ap.nic.in>).

Although the NPDS in India has been announced only in 1999, and the NSC commissioned in 2000, efforts by the concerned government agencies to provide statistical and numeric data in the digital form and via networks started earlier. Examples of these are given below. The economic and monetary data and information are available via the Reserve Bank of India (RBI) web site (<http://www.rbi.org.in>). RBI is the central bank of India. In addition to its main web site, it has also created special URLs for frequently accessed documents, a list of which appears as Annex 3. The documents available via the RBI

web sites provide textual as well as numerical and statistical information. The Weekly Statistical Supplement, available via its web site (<http://www.wss.rbi.org.in>) provides economic information in numeric form under many headings (The text version of a downloaded document is enclosed as Annex 4.) Census data in the floppy discs is also available at select institutions, but its format is not user-friendly for search purposes. Brief information on census is also available from Registrar General's Office (RGO) web site (<http://www.censusindia.net>).

In spite of the excellent web site maintained by the RGO, not much data are available via it. Web sites of nearly all ministries and departments of GoI and their agencies exist today. A list of these sites has been compiled by Varun. A cursory look at the URLs of the web sites of various ministries, departments, and agencies reveals that NIC has created web sites for many government ministries, departments, and agencies, but adequate information is not available via the government web sites, and is not updated in some cases. In some cases, there is no email address for contacting the concerned department. Thus it becomes clear that while, with the help of the elite NIC the conduit for information dissemination has become available to data source agencies, the content (data and relevant information) available via the conduit is far from satisfactory.

Looking Towards the Future

Reforms in the telecom sector in India have been very rapid in the last few years, and with the increasing competition, prices of computers, telecommunication equipment, and services have come down heavily. More and more bodies in India are now using the computers and communication facilities and services. These are likely to increase manifold in the coming years. With such a market scenario, need for information (including numeric information) available via the computer and communication facilities, both at workplace and at home would increase considerably. Therefore, assessing user needs for statistical and numeric information and enhancing users' awareness of the existing resources and services through which they can access such information is the need of the hour.

The initiatives by the GoI, such as the announcement of NPDS in May 1999 and setting up of NSC in January 2000 are in the right direction, but are in reverse order. The Government should have set up NSC earlier, as one of the terms of its reference is with regard to collection and dissemination of timely, reliable and adequate statistics. While the focus of NPDS is on centralisation of statistical information system, one of the terms of reference of NSC deals with decentralisation of statistical information system. In the light of report and recommendations of NSC (whenever submitted), the

NPDS will have to be amended or a new policy will have to be announced.

Consolidation of the efforts of various ministries, departments and other source agencies of the GoI is also needed to make the collection, validation, organisation and dissemination of numeric and statistical data efficient and effective. The SW has been entrusted with the job of coordination with other government ministries, departments, and data source agencies, and to act as a nodal agency for dissemination of statistical data (clause v of the guidelines). But nothing concrete has come out even after one and a half years of the adoption of the policy. Some guidelines and clauses of the policy act as hindrance in accessing data. In such a situation, the role of library and information professionals would be to convince the government to provide validated quality numeric information to society at large. The GoI through the SW can provide quality information by strengthening the existing infrastructure of the NIC and ensuring access to the information by making available the NICNET terminals available to people via public libraries.

Further Reading

Chidambaram, S. Siva. (1999) Access and availability of statistical information. IASLIC Bulletin, 44(3), Sep., 133-141.

Goswami, P.R. (1998) Access to socio-economic data with particular reference to India. DESIDOC Bulletin of Information Technology, 18(4), Jul, 29-38.

Goswami, P.R. (2000) Official statistical information: Indian scenario. Information Today and Tomorrow, 19(1) Jan-Mar, 11-16, 21.

India. Ministry of Statistics and Programme Implementation. (2000) Annual Report 1999-2000. [New Delhi: MoSPI]

India. National Statistical Commission. <http://www.nic.in/stat> (visited 10-1-2001)

Kathuria, Rajat (2000) Telecom policy reforms in India. Global Business Review, 1(2) Jul-Dec, 301-326.

National Policy on Dissemination of Statistical Data. (1999) Information Today and Tomorrow, 18(3), Jul-Sep, 16-17.

Saha, A. and Thulasi, K. (1998) Techno-commercial information on the Internet. Information Today and Tomorrow, 17(1) Jan-Mar, 6-18.

Satish Chander (1998) Access to legal information in India. DESIDOC Bulletin of Information Technology, 18(4), Jul, 21-28.

Seshagiri, N. and Reddy, C.L.M. (1997) Evolution of ethical aspects of digital information in India. International Information and Library Review, 29(2), June, 227-235.

Varun, V.K. (1998) RU on Internet? Information Today & Tomorrow, 17(4) Oct-Dec, 19-20.

Annex 1

National Policy on Dissemination of Statistical Data

i. Dissemination of official statistics in the form of reports, ad-hoc and regular publication etc. by the Central Ministries / Departments / Agencies as at present shall continue. Validated data, though published, including unit/household/establishment level data after deleting their identification particulars to maintain confidentiality should also be made available to the national and international data users in the form of hard copies and on magnetic media on payment basis;

ii. No data, which are considered by the concerned official in data source agency to be of sensitive nature and the supply of, which may be prejudicial to the interest, integrity, and security of the nation, should be supplied. The Central Government, or a state Government or the concerned Government agency, as the case may be, shall exercise its overriding prerogative to decide the degree of sensitivity of the official statistics produced by it. The data source agency will reserve the right to withhold its release altogether or to release selectively.

iii. Price of data to be supplied under (i) above should include the cost of stationery, computer consumables and computer time for sorting information. However

iv. price may be fixed in Indian currency as well as in Sterling Pound and American Dollar. Foreign currency prices may be determined using relevant official multiplier fixed from time to time for printed government publications;

v. Survey results/data should be made available to the data users in India and abroad simultaneously after the expiry of three years from the completion of the field work or after the reports based on survey data are released, whichever is earlier;

vi. Data users will give an undertaking in the prescribed form to the effect, inter alia, that the official statistics obtained by him for his own declared use will not be passed on with or without profit to any other data user or disseminator of data with or without commercial purpose;

vii. Data users will have to acknowledge the data sources in their research work based on official statistics. One copy of research study along with short summary of conclusions, if required by the concerned data source agency, should be supplied in the form of hard copy or on electronic media, free of cost; and

viii. The Department of Statistics will be the nodal agency for dissemination of official statistics provided by Central Government Ministries and Departments. However the concerned subject matter Ministries and Departments of the Central Government will be the final authority on issues arising out of this policy with a view to resolving any dispute between a data user and a data source agency.

The guidelines for the release of the data are:

i. A data warehouse in the Department of Statistics will be created to enable the data users and general public to have easy access to the published as well as unpublished validated data from one source.

ii. The data warehouse will collect data from various source agencies, integrate the data into logical subject areas, store the data in a manner that is accessible and understandable to non-technical decision-makers and deliver data/information to decision makers through report writing and query tools.

iii. As data source agencies are generating data at various levels, the responsibility of data supplied and receipt will be shared between the respective Central Ministries/Departments/ Agencies and the Department of Statistics by establishing and maintaining close collaboration.

iv. For each data type and source, detailed studies will be undertaken by the Department of Statistics in cooperation with the concerned data source agency on

- (a) the concepts, definitions, classifications and methods used in data collection and processing including validation,
- (b) formats of data collection,
- (c) media on which data will be supplied,
- (d) frequency of supply of data and
- (e) procedures and modalities for preservation, updation and dissemination of data.

v. The volume of data flowing from each source agency into the data warehouse will be assessed by the Department of Statistics in order to formulate the various parameters required for designing, establishing and maintaining a data warehouse.

vi. Each data source agency will be required to adopt for itself a calendar for preparation and release of data data which it will share with the Department of Statistics. As part of its nodal responsibility of dissemination of data from the source the Department of Statistics will keep track of the data release calendar of each source agency.

vii. The data source agency will be required to supply on computer compatible media, validated data, published or unpublished free of cost to the data warehouse.

viii. The Department of Statistics will prepare Directories of all available data in the data warehouse and update the same at frequent intervals. A web site will be created for the data warehouse and Directories will be available on the web site.

ix. From the data warehouse, data/information will be made available free of cost to the data source agencies for official use and also the approved research institutes and universities for research purposes.

x. The price of data to be supplied to other users will depend upon system hardware and software used for data storage, retrieval etc. and also on medium of supply of data.

Annex 2

National Statistical Commission

The Government of India has setup a National Statistical Commission to critically examine the deficiencies of the present statistical system with a view to recommending measures for a systematic revamping of the system (Gazette of India, Extraordinary Part 1, No.10, Resolution No.M-13011/3/99-Admn.IV dated 19.01.2000). The Commission consists of Dr. C.Rangarajan, Governor, Andhra Pradesh, as its part-time Chairman and the following 11 eminent experts as part-time members:

1. Mr. V.R.Rao, ex-Director, Central Statistical Organisation and UN Advisor
2. Mr. S.M.Vidhwans, ex-Director (Economics & Statistics), Govt. of Maharashtra and UN expert
3. Prof. J.Roy, Professor Emeritus, Indian Statistical Institute
4. Dr. Prem Narain, Emeritus Scientist, IARI and ex-Director, Indian Agricultural Statistics Research Institute

5. Dr. Rakesh Mohan, Director-General, National Council of Applied Economic Research (NCAER)

6. Dr. V.R.Panchmukhi, Director-General, Research and Information System for the Non-Aligned and other Developing Countries

7. Dr. Y.Venugopal Reddy, Deputy Governor, Reserve Bank of India

8. Dr. K.Srinivasan, Executive Director, Population Foundation of India and ex-Director of International Institute of Population Studies

9. Prof S.Tendulkar, Delhi School of Economics and Vice-Chairman, N.A.B.S.

10. Dr. A.B.L.Srivastava, Chief Consultant, Educational Consultants India Ltd. & ex-Professor, National Council for Educational Research and Training (NCERT)

11. Dr. Fredie Ardeshir Mehta, Eminent private sector economist and Director, M/s Tata Sons Ltd.

The terms of reference of the National Statistical Commission are as follows:

1. To examine critically the deficiencies of the present statistical system in terms of timeliness, reliability and adequacy

2. To recommend measures to correct the deficiencies and revamp the statistical system to generate timely and reliable statistics for the purpose of policy and planning in Government at different levels of administrative structure

3. To recommend permanent and effective coordinating mechanism for ensuring integrated development of the decentralised statistical system in the country

4. To review the existing legislation for the collection of statistical information and to recommend amendments where necessary, to achieve the objective of collection and dissemination of timely, reliable and adequate statistics

5. To review the existing organisation of the Ministry of Statistics and Programme Implementation (Statistics Wing) and other statistical units of the Government and to make recommendations on their staffing and training requirements to enable them to cope with the increase and development of statistical sources

6. To examine the need for instituting statistical audit

of the range of services provided by the Government and the local bodies and make suitable recommendations thereof and

7. To recommend any other measures for improving the statistical system in the country.

The Commission is expected to submit its report to the Government within a period of twelve months from the date of its establishment.

Annex 3

List of Special URLs for frequently accessed documents of Reserve Bank of India (RBI).

- * Currency Museum <http://www.museum.rbi.org.in>
- * Exchange Control Manual <http://www.ecm.rbi.org.in>
- * Weekly Statistical Supplement <http://www.wss.rbi.org.in>
- * RBI Bulletin <http://www.bulletin.rbi.org.in>
- * Monetary and Credit policy <http://www.cpolicy.rbi.org.in>
- * 9% Government of India Relief Bonds <http://www.goirb.rbi.org.in>
- * RBI Notifications <http://www.notifics.rbi.org.in>
- * RBI Press Releases <http://www.pr.rbi.org.in>
- * RBI Speeches <http://www.speeches.rbi.org.in>
- * RBI Annual Report <http://www.annualreport.rbi.org.in>
- * Credit Information Review <http://www.cir.rbi.org.in>
- * Report on Trend and Progress of Banking in India <http://www.bankreport.rbi.org.in>
- * FAQs <http://www.faqs.rbi.org.in>
- * Committee reports <http://www.reports.rbi.org.in>
- * Y2K <http://www.y2k.rbi.org.in>
- * FII List <http://www.fiilist.rbi.org.in>
- * Electronics Clearing Service <http://www.ecs.rbi.org.in>
- * Facilities for NRIs <http://www.nri.rbi.org.in>
- * SDDS-National Summary data page-India <http://www.nsdprbi.org.in>
- * Foreign Exchange Management Act, 1999 <http://www.fema.rbi.org.in>

Annex 4

Reserve Bank of India. Weekly Statistical Supplement
Dec 02, 2000

1. Reserve Bank of India
2. Foreign Exchange Reserves
3. Scheduled Commercial Banks - Business in India
4. Interest Rates (per cent per annum)
5. Accommodation Provided by Scheduled Commercial Banks to Commercial Sector in the form of Bank Credit and Investments in Shares/Debentures/Bonds/Commercial Paper etc.

6. Foreign Exchange Rates - Spot and Forward Premia
7. Money Stock: Components and Sources
8. Reserve Money: Components and Sources
9. Auctions of 14-Day Government of India Treasury Bills
10. Auctions of 91-Day Government of India Treasury Bills
11. Auctions of 182-Day Government of India Treasury Bills
12. Auctions of 364-Day Government of India Treasury Bills
13. Certificates of Deposit Issued by Scheduled Commercial Banks
14. Commercial Paper Issued by Companies (At face value)
15. Index Numbers of Wholesale Prices (Base: 1993-94 = 100)
16. BSE Sensitive Index and NSE Nifty Index of Ordinary Share Prices - Mumbai
- 17a. Average Daily Turnover in Call Money Market
- 17b. Turnover in Government Securities Market (Face Value)
- 17c. Turnover in Foreign Exchange Market
- 17d. Weekly Traded Volume in Corporate Debt at NSE
18. Bullion Prices (Spot)
19. Government of India: Treasury Bills Outstanding (Face Value)
20. Government of India: Long and Medium Term Borrowings - 2000-2001
21. Secondary Market Transactions in Government Securities (Face Value)

Dr. Jagtar Singh and H. P. S. Kalra, jagtar@pbi.ernet.in
harry@pbi.ernet.in, Department of Library & Information
Science, Punjabi University, Patiala - 147 002. (India)