

Seminar Innovations in Official Statistics

*Side event of the 40th Session of the Statistical Commission – UNSD, New York -20
February, 2009*

Innovation in Data Dissemination at the Brazilian Institute of Geography and Statistics

1. Introduction

The increase of democracy in the world, the better organization of the several segments of civil society to claim for their rights, globalization and international commercial treaties, the impact of technological advances on the economy and on society, the new priorities of the government and the changes in social characteristics result in new needs concerning information and new challenges for National Institutes of Statistics.

In this context, several challenges are faced by the National Institutes of Statistics, as presented below:

- to produce a bigger and more varied number of consistent and relevant statistics based on principles and measures with continuous evaluation of production, in order to release more updated data, with regular frequency, more varied topics and also spatially detailed
- to invest in the application of new estimation methods, for better use of the information collected and also to increase the use of administrative registers, so as to improve the efficiency of production and the reduction of the information collection burden on respondents
- to invest in increased integration among the social, economic, geographic and environmental milieus
- to define a working program which can balance strategic priorities, new users' demands and the available resources, considering the commitment with the needs, the information collection burden on respondents, and the several quality dimensions
- to develop strategies and dissemination tools which will be able to supply information to different media and in diverse spatial approaches in order to provide the appropriate data for the several existing public segments

IBGE has been trying to develop a better and more complete analysis of the country's reality, investing in the continuous improvement of its ongoing surveys, as well as in the increase of the representativeness of its data collections, through an expansion of the geographic coverage and the adoption of new survey topics as well as new techniques for estimation in small areas.

Nowadays, when productive processes are becoming more intense in the search for methodological advances and technological effectiveness, a team's versatility and skill,

greatly determine the success of any activity, for example, the production and dissemination of information.

2. The Internet and the dissemination of statistical information

Regarding the dissemination of statistical information, we are living in a period of great change, considering that presently the main dissemination channel is the Internet, of which technological and service evolution is hard to predict.

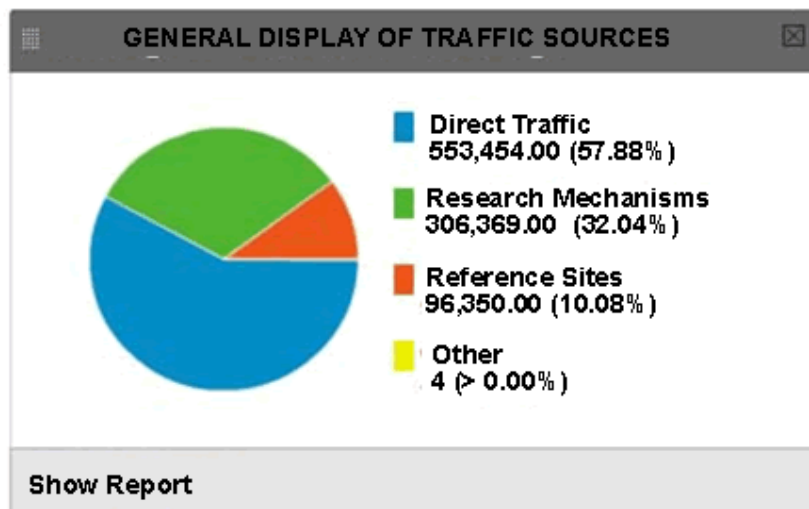
The Internet placed initially the Official Institutes of Statistics facing the challenge of providing information to its users almost in an instantaneous way. This goal, in a general way, was attained, with the release of information in a traditional way, with tables, graphs and thematic maps.

Initially, users were very much satisfied with the gain of time and accessibility. Also, the number of users started to increase exponentially, as a consequence of the growing number of Internet users as a whole, and of the improvement implemented on the web sites of the Institutes, which attract new users.

However, the expectations increased, demanding:

- Databases with online aggregated tables
- Thematic maps constructed “on the fly”
- Download of complete publications
- Interactive databases producing “à la carte” tabulations
- Dynamic graphs which allow the addition of time as a variable

Another phenomenon is that the growth of the number of persons who access the site is encouraged by search sites (Google, Yahoo etc., as seen in the graph below) and this type of occasional user needs immediate answers, which might fulfill their needs, and if possible, overcome expectations and turn this person into a constant user. Hence, mechanisms that will answer directly the requests must be developed, or, in other words, the refinement of topic indexing and of search tools in the site, must quickly address the matter demanded by the user, following the concept: “do not make me think”.



Source: Google Analytics, result of the access ways to the main page of IBGE.

In this sense, IBGE is willing to construct, in terms of web design, a minimalist portal (where *less is more*), in which the user, through a search program, will be served with the lowest number of clicks:

The image shows the top section of the IBGE website. At the top, there is a navigation bar with links: [About IBGE](#), [News](#), [Surveys](#), [Indicators](#), [Geosciences](#), [Highlights](#), [Channels](#), [Topics](#), and [Website Statistics](#). Below this is a banner with the IBGE logo on the left and a quote on the right: "Retratar o Brasil com informações necessárias ao conhecimento da sua realidade e ao exercício da cidadania". Below the banner, there is a news snippet dated 11/14/2008 with the headline "IBGE publishes the Regional Accounts 2003-2006" and a "POPLOCK 190,200,486" indicator. A search bar with a "Search" button and links for "Advanced Search" and "Site Map" is present. At the bottom of the screenshot, there is a list of various links including "Cities@", "Popclock", "Statistical Series", "Countries@", "Brazil in Summary", "States@", "Virtual Shop", "Catalog", "Profile of Brazilian Municipalities", "Library Online", "IBGE 7-12", "BME Electronic Questionnaires", "Publishing Calendar", "Maps", "IBGE Teen", "Statistical Classifications - CONCLA", "Institutional Memory", "Forums", "ENCE", "SIDRA", "Statistics of the Twentieth Century", "ALOS", "Public transparency", "Edicts and tenders", and "Annual Accounts Processes".

3. Adding the spatial factor to the process of data collection, analysis and dissemination

Geographic Information Systems - computerized systems which capture, store, edit, analyze, share and visualize geo-referenced information and answer queries, have had a fast evolution as a consequence of the advances in Information Technology, and their use has been spreading out in an accelerated way.

The traditional role of cartographic information in the production of statistical data consisted in giving support to the data collection, through boundary maps and access ways, with results presented in thematic maps. Today, with the additional capacities of Geographic Information Systems, the Global Positioning System - GPS and satellite images, in spatial resolutions similar or better than the ones of aerial photographs, the possibilities of using these technologies supporting all the collection process, besides the statistical information processing and dissemination are enormous.

The application of digital tools adapted to handheld computers with GPS is used in the collection of field data for the correction of cartographic bases (access ways, limits, etc)

which will be used in the process of information collection. Besides this, the Geographic Information System includes the association of the vectors of streets and access ways to a central address list (CNEFE, in the case of IBGE), which facilitates the work of the data collection team.

The possibility of combining the data collection through an electronic questionnaire with a global positioning system, with the use of an address list, allows the spatial location, for example, of the supply of service goods (school and health units) provided by the public service and the identification of areas where there is lack of these services. As a consequence, decision-takers will have one more tool to help them in the elaboration of public policies.

Regarding sample surveys, the use of satellite images helps in the determination of more representative samples, improving, this way, the reliability of the data. The use of satellite images, with different spectral, spatial and temporal resolutions is also very important to the collection of information about natural resources, since it provides a reduction of the collection period and costs.

Geographic Information Surveys constitute a powerful tool considering the analysis of data, since they provide the integration of statistical data to other type of information (highway-railway systems, hydrographic network, airports and ports systems, etc), besides providing the crossing of statistical information, aggregated by regions, Federative Units, municipalities and census sectors, in a quick and speedily form, thus providing a spatial vision of data, combined with other important components.

The dissemination of results must reach the biggest possible number of persons, of all spheres of knowledge and power, in a quick way with an accessible language. In order to do so, IBGE is trying to release its information on the Internet, with tools specifically designed for the visualization of geographic and statistical data, in synergy with other technologies available in the web.

4. Sharing of Spatial Data

Several countries or groups of countries have been developing what is called Infrastructure of Spatial Data, using the communication network provided by the Internet and recommending the shared use of geographic information in a decentralized way. In this scenario, each producer releases information with metadata using their own tools and services, thus increasing the potential use of information of each one and bringing benefits to all of them.

The fact that IBGE has as its attributions Geoscientific and Statistical coverage, leads to the idea of a natural convergence of both into a governmental level system denominated “SIG Brasil”, that is, an Information System where information would be shared in the federal government sphere.

The project is presently in course, updating the digital cartographic bases and installing hardware and software structure for the release of the 2010 Census results in this platform. This will cover all the census cartography, such as maps, orbit maps, orthophotos, census sectors and address lists which will be associated and geo-referenced, allowing, as a consequence, its use through browsers on the WEB, keeping, evidently, the statistical secrecy and opening up unlimited possibilities of use.

This way, we are creating the embryo of what we denominate National Infrastructure of Spatial and Statistical Data.