

The challenge of disseminating European statistics: the Eurostat experience

Pedro Díaz Muñoz¹

Introduction

This paper first describes the role of Eurostat in the dissemination of European data – a role at the end of a production chain that begins with the statistical organisations of the Member States. The paper then explains the legal mandate establishing this dissemination role, as well as recent enhancements to it through the European Statistics Code of Practice and the revision (currently in progress) of the basic legal framework for EU statistics. Finally, the paper reflects on future challenges, such as the need for statistics to be immediately available, the growing need for global access to global information, the demand for more explanations of data, competition from private statistical data providers, the issue of access to very large databases and the varied needs of different users. These are the challenges that frame the issues relating to the dissemination of public statistics as this process evolves.

Background

As the Statistical Office of the European Union, Eurostat is responsible for designing, collecting, producing and disseminating European statistics. One important aspect of this responsibility is making European data available to all users, European institutions, national administrations, businesses, analysts, researchers and the general public. The statistical production process begins with the collection and production of data by national statistical organisations, pursuant to standards developed at the European Union level and usually spelled out in binding legal instruments. The data produced by Member States are subsequently transmitted electronically in standardised formats to Eurostat, where they are validated and processed before the European totals are computed. These processed national data plus the European aggregates constitute the official European statistics. As a final step in the production chain, the data are disseminated by Eurostat and the national statistical organisations.

The statistical dissemination process is embedded in the basic legal instrument dealing with the production of European Statistics, and is currently being enhanced in two different ways:

- Revision of the legal framework, which is currently being drafted. It underlines the importance of statistical dissemination and stresses the need for effective cooperation with the European Statistical System, in order to optimise user access to data.
- Two of the principles – principles 6 and 15 – set forth in the European Statistics Code of Practice, which was adopted by the Statistical Programme Committee in February 2005, are devoted to dissemination, emphasising matters such as impartiality, equal access, clarity and accessibility.

¹ Eurostat, Statistical Methods & Tools; Dissemination, 5 rue Alphonse Weicker, L-2721 Luxembourg-Kirchberg, Luxembourg. E-mail: Pedro.diaz@ec.europa.eu.

Eurostat dissemination policy

Eurostat's dissemination policy was substantially redefined at the beginning of 2004. An internal Reflection Group led by Michel Glaude, Eurostat's Director of Social Statistics, put forth a fundamental policy document. Its basic principles are that:

1. Information should be available to all users at the same time and under the same conditions. Eurostat is currently drafting an impartiality protocol that clearly spells out the rules governing implementation of this principle. The rules – already, in fact, in place – are based on the requirement that, except in duly justified and publicly explained situations, data shall be made available to everyone at the same time. The only exception to this is the transmission of data, under embargo, to selected institutions in very specific cases.
2. All statistical data shall be disseminated free of charge. At the beginning of 2004, Eurostat decided to make all information on its website accessible to everyone without cost. The website functions as the front end for reference databases that store all of the data to be disseminated, including the electronic versions of printed publications.
3. The main vehicle for statistical dissemination shall be the internet. The Eurostat website, providing full access to all Eurostat data, was made available to the public on 1 October 2004. Since then, online access has been the principal means of access for all users of Eurostat data.
4. Printed publications shall continue to be a vehicle for dissemination. Eurostat believes that there is a need for printed publications. Consequently, a publication programme is being maintained, though the number of titles (currently about 40 per year) has been reduced, with a focus on compendium and thematic publications covering different topics. A large number of analytical information sheets (Statistics in Focus) and rapid data sheets (Data in Focus) are also produced (approximately 250 per year). As mentioned, all of this information is uploaded to the website.
5. User feedback is a cornerstone of statistical dissemination. Eurostat monitors use of the site in various ways (number of visits, page views, number of PDF, data and table downloads, etc). Eurostat also conducts periodic user satisfaction surveys on the web to solicit feedback. Lastly, user support activity carried out jointly with the NSIs is also an excellent source of feedback on user satisfaction and future needs.
6. Cooperation within the European Statistical System should enhance outreach. Eurostat has an ambitious programme of cooperative dissemination activities. Projects include coordinated user support activity, sharing of tools, dynamic links to different websites, exchange of news via RSS feeds and exchange of data in XML formats to customise display on national websites. To track and develop these coordination activities, a Dissemination Working Group with representatives of Member States meets twice a year, while several Task Forces are also active.

Future challenges in disseminating European statistics

As the information age develops and accessibility to information evolves, there are important challenges for statistical dissemination in terms of keeping users accurately informed and providing sufficient information for decision-making by stakeholders. The era when official statistics were the sole source of information is behind us. Official statistical organisations today confront a new paradigm, in which other information brokers also play a role. Let us examine the main challenges to be dealt with.

First, there is a need for data to be immediately available. This is very acute in the case of European statistics, since in many if not all cases the corresponding national data have already been published by the originating country. Since users are aware of this, they are unwilling to accept delays in the publication of the relevant European data. Furthermore, delays may encourage information brokers to compile their own aggregates, which may end up being the framework for economic decisions before the official European statistics are available. The European Statistical System must therefore develop ways to make data available to the entire ESS network as soon as national data become available.

Second, there is a need for global information to be globally accessible. Information needs today transcend borders, and users from any part of the world may be interested in data from any other. The structuring of links must be enhanced to allow users to navigate effectively through the vast amount of information offered by different statistical organisations, and statistics must be presented in a way that is coherent and understandable.

Third, and following from the foregoing point, measures must be taken to facilitate users' access to large amounts of information. Tools must be provided that present information synthetically and dynamically, giving users a rapid overview of large databases.

Fourth, since users are no longer required to go through information brokers to access data, the data must have clear and adequate explanations of their context, purpose and limitations. Therefore, metadata will play an essential role in future dissemination schemes, and navigation and search capabilities must be enhanced so that users will be able to find, on their own, the information they need.

Fifth, just as not all users are the same, neither are their needs. As the information market becomes more global, a wider variety of users become interested in statistical data. Organisations responsible for statistical dissemination today and in the future must therefore understand, and take account of, users' varying needs, motivations and ways of viewing data.

Sixth, researchers' access to data must be optimised. Currently, the richness of the data collected from respondents is not being fully exploited, due to concerns about the confidentiality of data. Current protective measures restrict access too much. The research community could use this information securely if adequate techniques and procedures were put in place to evaluate the risk of disclosure when granting access.

One further challenge facing official statistics organisations is competition from private providers of statistical data. In many cases, statistical data on economic and social phenomena can be obtained very easily through non-official statistical sources, and can be made easily available online. One of the future roles that official statistics organisations must play – and a subject they must carefully study – is helping users to judge the quality, relevance and continuity of such data.

Dissemination in the ESS. Where are we now?

The European Statistical System is already examining several ways of meeting these challenges. For example, there are numerous ongoing projects involving cooperation between Eurostat and the statistical offices of EU Member States. Some of them, such as RSS news feeds and the exchange of XML data for customised presentation, have been mentioned above. There is also an ESS intranet initiative, known as INSITE, which was launched in early 2007, and to which most of the Member States have already subscribed. Great efforts have also been made in creating and upgrading an ESS virtual network.

Another strand of international cooperation projects is designed to address the issue of standardising data and metadata models. This is part of the SDMX (Statistical Data and Metadata Exchange) initiative sponsored by several organisations (IMF, OECD, BIS, UN,

ECB, Eurostat and World Bank). SDMX permits data and metadata to be exchanged and automatically managed by recipients. For this purpose, SDMX employs standard data structures that can be shared by all partners, so that whenever a data set arrives, it can be displayed according to the shared standard structure. SDMX also provides a set of tools that will make it possible to share information placed on the websites of different organisations, so that a user could potentially select data from different data providers and produce a table combining the information – and do so in real time – through a registry system that sends out a notification each time new data are uploaded by one of the partners. Eurostat is currently working on a project called SDMX Open Data Interchange (SODI), which applies this concept, using SDMX for the simultaneous dissemination of short-term statistics.

At present, Eurostat is developing tools for improved access to, and display of, data. By the time this conference takes place, a new generation of maps, graphs and country profiles will be operational on the Eurostat website, and it will be possible to graphically display data from selected parts of the databases. An enhanced user interface for online access to all Eurostat databases will also be operational before the summer of 2007. All of these tools should be seen as the first phase of a new generation of utilities for displaying information stored in large databases, which will gradually be implemented in the next couple of years. Many of these tools are being developed jointly with Member States.

The SDMX initiative also provides an opportunity to develop standards for the structuring and presentation of metadata – another issue Eurostat is working on – using the above-mentioned SODI project as a testing ground. Eurostat has recently reorganised its activities so as to have a specific unit whose primary mission is to improve the content of information.

The new XML publishing initiative is designed to produce publications much faster, and to produce hardcopy and PDF publications simultaneously with their web (HTML) counterparts. Eurostat has launched a task force to study relevant previous experience and to develop a strategy that can then be implemented by both Eurostat and national statistical offices.

Finally, information search utilities must be improved, in order to keep up with increasing demand for faster data search capabilities. To this end, Eurostat is working on improving search engines and navigation tools.

Conclusion

The dissemination of European Statistics has been substantially transformed in the last few years to adapt to new technologies and changing needs. New challenges lie ahead, however, as information becomes more accessible to all users. The ESS must collaborate on developing new tools and policies, in order to evolve in tandem with user expectations and technical progress.