

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS**

UNECE Work Session on the Communication of Statistics
(30 June – 2 July 2010, Paris, France)

- (v) Internal communication and training

**STATISTICS EXPLAINED –
INTRODUCING A NEW APPROACH TO DISSEMINATION**

Submitted by Eurostat¹

I. INTRODUCTION

1. In order to reach larger user groups, and inspired by the success of Wikipedia, Eurostat started a project on disseminating statistical information through a wiki-type system. The main objective is to create an easy to use and frequently updated set of articles about European statistics. As a starting point, the content from existing publications as well as from other sources was uploaded into the wiki. The articles are interlinked with Eurostat data bases and publications. Also, links to related information on other parts of the Eurostat website and on other websites are included. The content is being gradually improved and completed by all Eurostat staff.
2. The system has been open to the public since September 2009. It is still considered to be a beta version since the content structure has still to be stabilised and various technical developments are ongoing. Content is continuously added, both on the initiative of the statistical units and of the dissemination unit. Whereas statistical units concentrate on their own topic and introduce a certain creativity and diversity, the dissemination unit tries to keep a harmonised appearance and a balance between the statistical topics.
3. This article describes the process of introducing the system into the organisation and discusses the reasons for success and the difficulties encountered.

II. THE START

4. Eurostat publishes annually 5-15 printed statistical books, 15 pocketbooks, 120 Statistics in Focus and various methodological publications. Large amounts of content have been created and are disseminated in paper and PDF format. Similar content is often created repetitively in different forms for different purposes.
5. Especially for large publications such as the Eurostat yearbook, the production process is quite cumbersome. Drafts are circulated and discussed several times with the statistical units before a final version is reached. As a consequence, data at the moment of publishing is already 6-9 months older than what can be found on the website.
6. In order to improve the situation, the Eurostat management committee asked the dissemination unit in January 2008 to examine possibilities to disseminate the content of publications more effectively.

¹ Prepared by Ulrich Wieland, ulrich.wieland@ec.europa.eu.

III. DEVELOPING THE CONCEPT

7. In the initial discussions, the remarkable success of Wikipedia was highlighted. Wikipedia became the main global reference for encyclopaedic information only a few years after its creation in 2001. The success of Wikipedia lies in the fact that everybody who feels expert in a field can contribute to its content. The editing language is simple and easy to learn, and well documented. Producing an attractive layout does not require high technical skills. For the end user, the system is easy to use.

8. However, it took several months to clarify the ideas about how to apply such a system to Eurostat. Discussions in the Eurostat Dissemination Committee (Codif) took place in May 2008 on the basis of a first Mediawiki prototype. The Codif contains both members from statistical units and members from the dissemination units. Discussion points concerned the content (individual indicators, predefined tables, replacing the panorama publications or including all value-added content of the website), the relation to other parts of Eurostat dissemination, the problems of responsibility for the content and the additional workload.

9. On the basis of these discussions, the concept was further refined and open points clarified. The Eurostat Management Committee approved the concept in November 2008.

10. Up to this point, a trainee with technical background made some first developments and tests. Rapid progress was facilitated by the fact that the required software was freely available on the internet and could be implemented on a PC. Testing and experimenting could start without any time-consuming software acquisition and, in the first phase, without the agreement of the central informatics units.

11. After the positive decision of the Management Committee, the system had to be put on a stable technical basis with well-defined support by the information management department. The technical choices were further discussed since Mediawiki is not an approved software of the European Commission. The discussion delayed the project by 2 months but eventually the software was accepted. A technical support was organised on a contractual basis.

12. In this phase, statistical units were asked for the first time to work on the system and to check and review the existing articles. This was a crucial step in the process, and several units showed a critical attitude, especially towards the additional workload and also the conceptual issues and overlaps with the existing website.

13. However, Statistics Explained received very high-level support. Even the Commission's communication on the '...production method of EU statistics: a vision for the next decade', (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2009:0404:FIN:EN:PDF>) which was adopted on 10 August 2009 but debated much earlier, refers explicitly to Statistics Explained as a modern tool for communicating with users. The clear support by the top management motivated also reluctant units to invest in the system.

14. In September 2009 the system was opened to the public. At that time, the content of the Eurostat yearbook and the regional yearbook as well as selected articles from other publications and Statistics in Focus were included. Statistics Explained was opened to the public on 28 September 2009 as a beta version. Currently, the system contains 137 statistical articles and 24 background articles, plus about 500 definitions.

IV. MAIN DISCUSSION POINTS

15. This section summarises the main discussions which took place during the early phase of the project. It was important that concerns of experienced users were discussed in an early stage of the project. This helped to shape the system and also to take account of concerns of the users.

16. One of the biggest concerns was the responsibility for the individual articles. Following the discussions, a double validation was implemented, requiring the approval of a responsible person in the statistical unit, and the approval of a person in the dissemination unit. This solution gives a reasonable guarantee of control without introducing an overly bureaucratic procedure. Moreover, controls are less strict than for publications which are read several times during their production process.
17. Some people were also uneasy with the idea that other people could change their articles. However, since all changes are traced, it is unlikely that changes are introduced without sufficient care. In addition, inadequate contributions can be quickly removed and previous versions restored.
18. Several colleagues mentioned the need to write differently for the web than for traditional publications. In fact, web readers jump much quicker between pages and expect to understand the main messages of a page immediately. On the other hand, you find also many articles on Wikipedia going over many pages and requiring sequential reading over several pages. Eurostat concluded that a short introduction should be added to each article to support the reader with the main results, but to leave existing published material nearly unchanged.
19. A major discussion point of Statistics Explained was the updating of data. As with PDF publications, there is no automatic link between Statistics Explained and the data base, and Statistics Explained data does not change when the data in the data base changes. Two measures have been taken to tackle this problem. All pages containing data start with information about when the data has been extracted. In addition, there is always a deep hyperlink to the statistical database where the most recent data can be found.
20. However, care must be taken that the articles and tables/graphs do not become too outdated. For example, in the current situation of the unexpected economic decline, it might be advisable to update outdated texts talking about continuous growth during the past years. The need for updating is very dependent on the subject covered. Statistical units are asked to check their pages at least on an annual basis. The dissemination unit regularly checks the system for outdated pages.
21. Early discussions concerned the role of Statistics Explained in relation to Eurostat publications. It became clear that several existing publications become obsolete with the introduction of Statistics Explained. However, a few major publications such as the Eurostat yearbook will remain. In order to assure coherence of the yearbook with Statistics Explained, future versions of the yearbook will be based on Statistics Explained articles. To support this approach and to maintain coherence, all articles suitable for the yearbook will be earmarked and their updates closely monitored.
22. Statistics Explained is an integral part of the Eurostat website. Technically, however, it is a separate system, and decisions had to be taken on how to integrate Statistics Explained with the Oracle Portal, which is the underlying software of the Eurostat website. The integration is mainly done through intensive hyperlinking in both directions.
23. Availability of language versions is a regular topic of discussion in Eurostat. As for publications, multiple language versions are desirable but the resources for translation are scarce. For Statistics Explained, a flexible approach has been decided, with English as the master language. Further language versions (especially German and French) will be offered when translations are available.
24. A major concern was the additional workload and the need to learn yet another new system. This concern could not be entirely eliminated in the short run, although Statistics Explained will replace several publications and potentially other dissemination tools such as Coded. Also the yearbook production, which requires an active collaboration by all units, will be simplified by Statistics Explained.

V. CURRENT SITUATION AND FURTHER EVOLUTION

25. Participation of Eurostat staff in the content development varies. Whereas some units show a high motivation to extend or update the content, others see it still as an additional burden to an already overloaded staff. It will be the task of the Statistics Explained project team to demonstrate that Statistics Explained in the end simplifies and rationalises dissemination. Currently, about 135 people (out of a total Eurostat staff of 835) have logged into the system, of which an estimated 50 make regular updates.

26. A major step of integrating the system into the organisation will happen during the coming months, when the content of the 2011 yearbook will be prepared in the Statistics Explained system. Rather than exchanging Word and Excel files with draft text and tables, units will be asked to update the 2010 content directly in Statistics Explained or to comment on existing content using the discussion function in Mediawiki. All statistical units will then have to actively work with the system.

27. Some statistical units started to include the task of updating Statistics Explained into public tenders. One of the biggest Eurostat publications, the European Business Statistics publication, will be integrated in Statistics Explained before the end of this year, and the paper and PDF publication will be discontinued. Other statistical units follow this example.

28. A further essential step is the integration of the Eurostat terminology data base Coded. Apart from the fact that definitions become much more accessible when they can be reached through hyperlinks from existing texts, it also allows the elimination of the specific Coded software system. The integration of other methodological information such as the ESMS files is in discussion.

VI. CONCLUSION

29. Compared to other major dissemination systems, Statistics Explained was introduced in a very short time period. In fact, it took only 18 months from the initial mandate and the first reflections to the opening of the system to the public. The main reasons for success were as follows:

- The strong support of the management which helped to overcome the initial reluctance;
- The availability of suitable content in the yearbook and in other publications. An important stock of valuable information could be introduced into Statistics Explained before statistical units had to work with the system;
- The use of open source software by avoiding technical developments whenever possible. Although some technical support is required for installation, experimenting, testing, troubleshooting and a few other specific developments, the total informatics cost did not exceed 1.5 man-years.

30. Although most people are now convinced about the usefulness and the advantages of the new system, it has not yet fully penetrated the organisation. Such a process takes time but there are no doubts that this will happen during the coming years.