

# **Improving the visualisation of statistics: The use of SDMX as input for dynamic charts on the ECB website**

Xavier Sosnovsky, Gérard Salou

European Central Bank

## **Abstract**

The ECB has introduced a number of dynamic graphs on its web site. Those are based on Rich Internet Applications (RIAs) technologies. This paper describes the development process of a generic visualisation framework, based on the SDMX standard. The framework can be used by other visualisation tools, thereby removing the need to hard code data structures within these tools, significantly cutting down development time and making it possible to visualise and to navigate any SDMX coded data file on the Internet.

**Keywords:** visualisation, SDMX, rich internet applications

## **1. Introduction: The need for SDMX dynamic graphs**

In December 2006, the Statistical Information Services Division of the European Central Bank (ECB) started a project to investigate ways to improve the communication of statistics towards the public. More specifically, the project aimed at:

- Improving the visual display of statistical data by following best practices in the field of data visualisation.
- Improving the accessibility of the published statistical data.
- Enabling basic data analysis from the visual display thereby improving the understanding of the statistics published.
- Leveraging the usage of the SDMX data framework already used along the whole statistical process at the ECB.

On a technical level, the project would investigate the potential use of Rich Internet Applications (RIAs) technologies together with SDMX concepts to meet these targets. As RIAs claim to combine the best features of both web applications and traditional desktop applications, they appear very

promising. Like traditional desktop applications, RIAs offer a rich user experience and an improved responsiveness, and like web applications, they also offer a broad reach, platform independence and a low deployment cost.

This paper describes the various visualisation tools based on RIAs technologies which have been developed successively by the ECB, within the scope of the project mentioned above. The project was split into 3 (partially overlapping) phases:

1. Phase 1 (2006-12 – 2007-07): Research and initial developments
2. Phase 2 (2007-05 – still on-going): Consolidation and generalisation
3. Phase 3 (2008-03 – still on-going): The development of ECB dashboards

## **2. Phase 1: Research and initial developments**

In this phase, which lasted from December 2006 to July 2007, some preliminary studies were made to identify the most suitable technologies for the task at hand<sup>1</sup>. During this phase, two applications, currently in use on the ECB website, were developed.

### **2.1 The euro foreign exchange reference rates**

Around mid-April 2007, new graphs were introduced for the [euro foreign exchange reference rates](#). This section is, by far, the most popular on the ECB website with more than 700 million visits since between January 2006 and January 2009. Daily rates are available in various formats, including HTML, SDMX-ML, CSV and PDF.

The new dynamic graphs were developed having in mind the most frequent requests of ECB website users concerning exchange rates and how to best meet them using RIA technologies. As such, the dynamic graphs offer:

- An easy access to frequently requested figures such as the more recent exchange rate and the change compared to the previous observation, the change over the selected period of interest (the last year by default), as well as the minimum and maximum value over the selected period, those three measures being calculated on-the-fly as the observation period is adjusted.
- When hovering over the graph for a selected currency, visitors will see the value of each data point together with the change over the last period. As direct access to historical data was one of the most frequent requests concerning exchange rates on the ECB statistics hotline, the situation has been significantly improved with the introduction of the new graphs. This functionality would not have been possible using traditional charting technologies, and could only be achieved because of the use of RIAs technologies.

---

<sup>1</sup> Out of the various technologies available at that time, Adobe Flex has been selected for the project.

## UNCLASSIFIED

- Different periods of interest (last month, last 3 months, last year, etc) can be selected. Users needing a finer level of control for selecting the period of interest can use the slider at the bottom of the graph.
- Another popular although simple request, the possibility to swap the base currency (for instance the EUR vs. the USD instead of the USD vs. the EUR), has also been implemented.

As all actions performed by visitors (such as changing the selected period for instance) happen on the visitor's computer, the application responds immediately, which enhances the overall user satisfaction and allows increasing the quantity of information extracted from the data.

So far, the application has been very well received by users of the website. Nevertheless, additional functionalities are planned and will include, among others:

- The possibility to change the time series frequency.
- The possibility to compare currency developments on the same graph (for instance the JPY and the USD against the EUR).
- The possibility to display data as a table.
- The possibility to download the data from the graph and import them into a spreadsheet application.

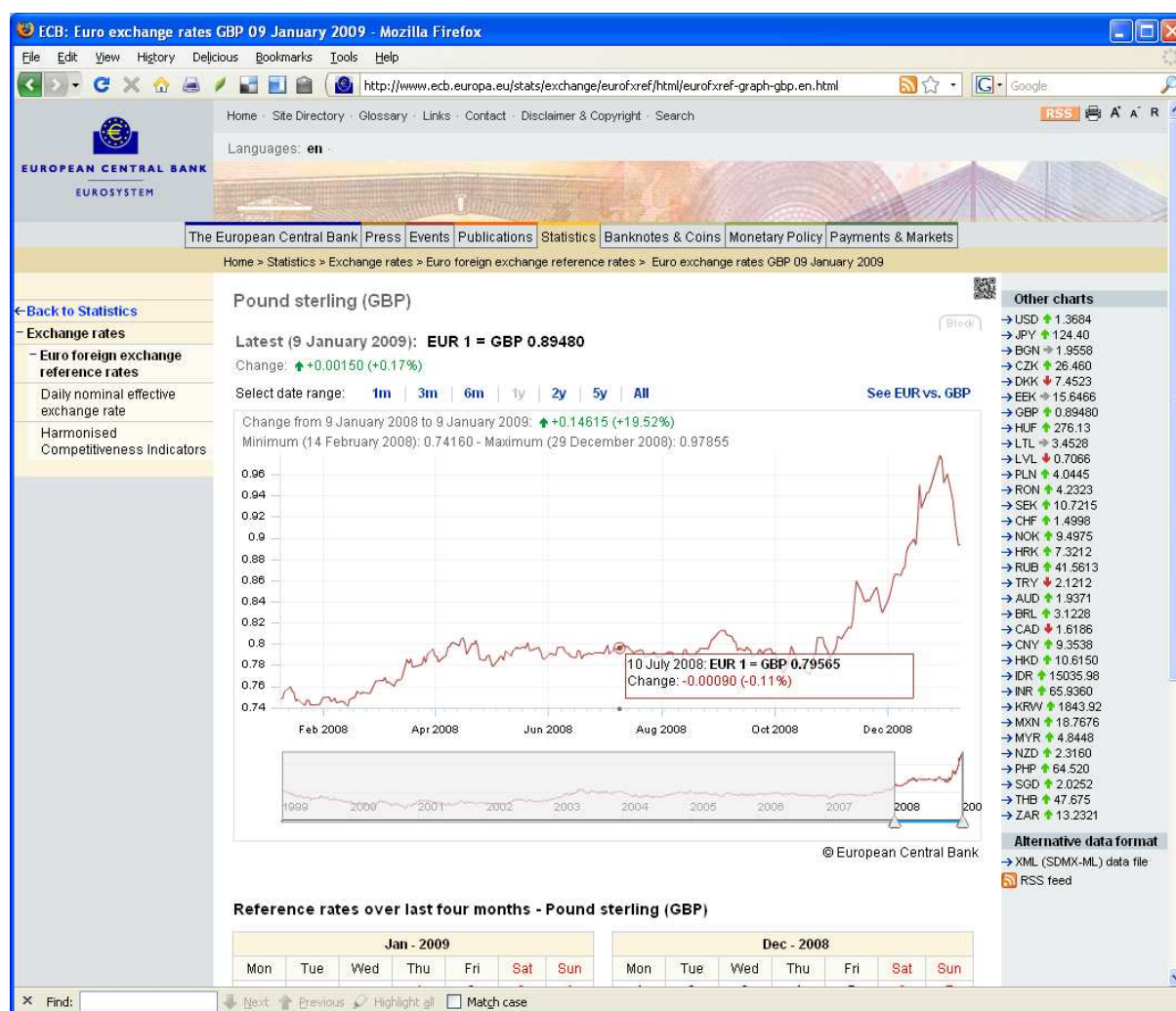


Figure 1 - The euro foreign exchange reference rates

## 2.2 The euro area yield curves

In July 2007, a new interactive application was introduced on the ECB website, to display the [euro area yield curves](#) on AAA-rated euro area central government bonds.

The application allows users to visualise different types of yield curves (spot rate, instantaneous forward, par yield) and their parameters, at various dates. It also supports comparisons of curves at different points in time or with all euro area central government bonds. Evolution of a selected yield curve can easily be grasped by playing an animation of the selected curve over time. Like for the exchange rates, it is also possible to focus on certain periods of interest and to dynamically perform calculations, such as the spread between curves.

Since its introduction on the ECB website, the euro area yield curve application has regularly shown very high usage and ranks among the ten most popular areas of the Statistics section of the ECB website.

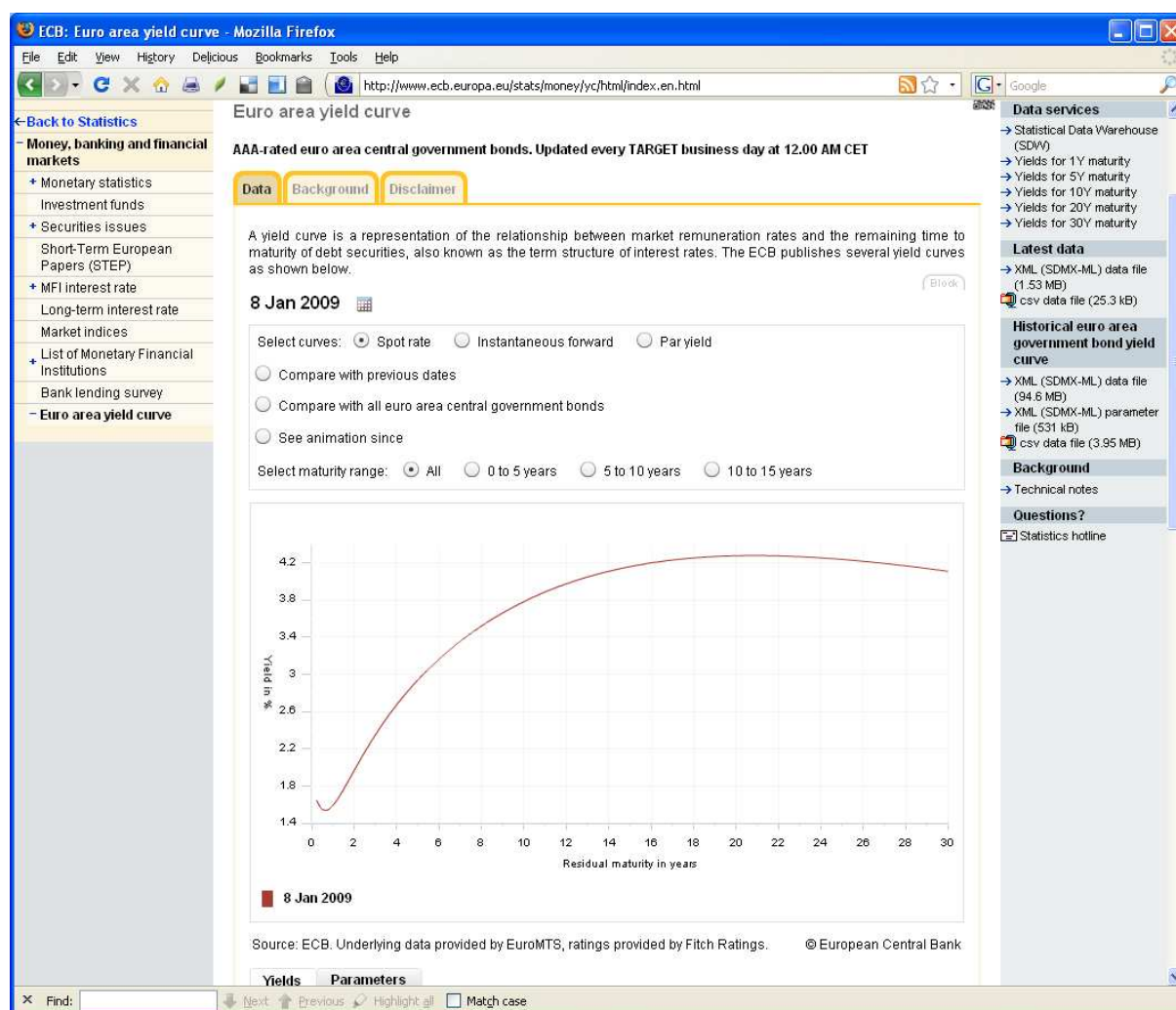


Figure 2 - The euro area yield curve

## 2.3 Conclusions of Phase 1

Rich Internet Applications technologies can meet the goal set initially to improve communication towards the public. Indeed dynamic graphs improve the visualisation and the accessibility of statistics and help making the analysis more efficient, productive and successful, thereby improving the understanding of the data. The applications are very popular and show a rather high usage. Demonstrations to interested parties including to the Press are very well received.

However, the two applications developed in Phase 1 are "ad hoc" applications, where the knowledge of the structure of the data being displayed (for instance, the notion of currency and base currency for the exchange rates application) is "hard coded" into the application itself. This method is inefficient and does not scale well in case many similar visualisation tools need to be built. In order to speed up the development of such applications, there is a need for a generic visualisation tool, which supports the visual display of statistical data and metadata without the need to "hard code" in the application any knowledge about the structure of the data being displayed. This was one of the main objectives of Phase 2.

### **3. Phase 2: Consolidation and gradual extension to the ECB website and the ECB Statistical Data Warehouse**

Following the success of the dynamic visualisation tools developed during Phase 1, it was decided to extend the use of such dynamic graphs to other sections of the ECB website and to all graphs available on the ECB Statistical Data Warehouse. However, as mentioned earlier, the visualisation tools developed during Phase 1 have hard coded the internal knowledge of the data structure, which varies widely among different sets of data. Adopting a similar approach for all graphs published on the ECB website would significantly slow down the development of such visualisation tools, would hamper code reusability and would be vastly inefficient. As such a generic visualisation framework, based on the SDMX standard, has been developed during Phase 2. The framework can be used by the other visualisation tools, thereby removing the need to hard code data structures within these tools, significantly cutting down the development time.

#### **3.1 Consolidation: The development of an SDMX framework**

This phase started in May 2007 and finished in February 2008 and aimed at developing a visualisation framework based on the SDMX standard. SDMX was selected for various reasons:

- SDMX is already in use in all layers of the ECB Statistical environment (data exchange, data model, data dissemination, etc).
- It is an ISO technical specification (ISO/TS 17369), sponsored by 7 international institutions (including the ECB).
- The [SDMX information model](#) describes all artefacts needed to represent the statistical data and metadata being displayed.
- The SDMX information model can be represented into various XML formats ([SDMX-ML](#)), which is ideal for feeding data to the visualisation tools.

The framework is organised in three packages. It contains approximately 22.000 lines of code and is covered by more than 700 unit tests.

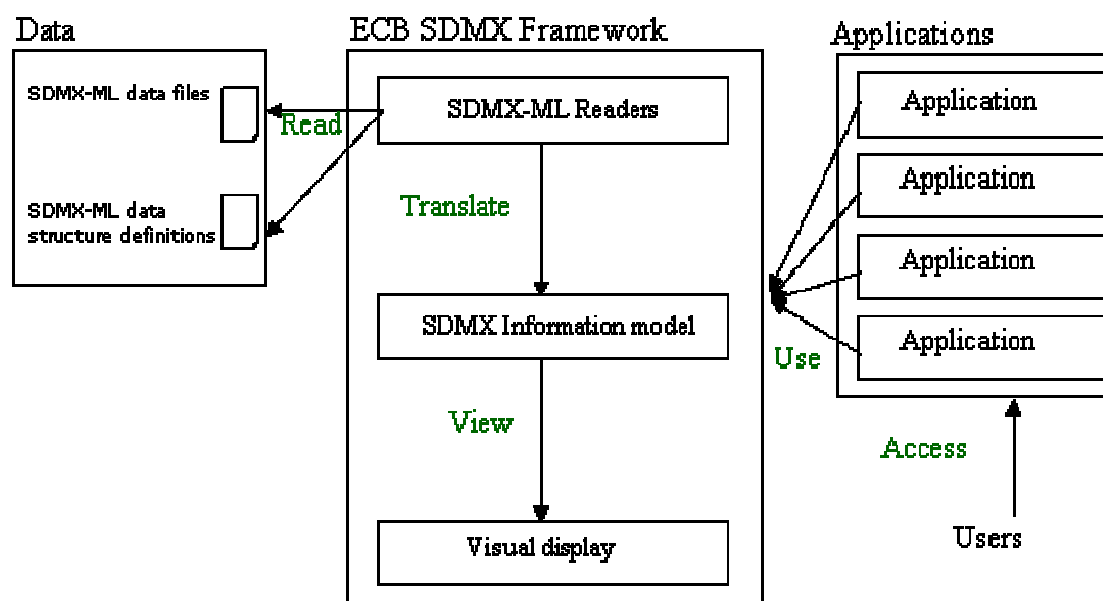


Figure 3 - The ECB SDMX Framework

### 3.1.1 SDMX-ML Readers Package

The SDMX-ML Readers package reads the data structure definitions and the data files, in the SDMX-ML Structure and Compact formats respectively.

The readers will parse SDMX-ML data structure definitions and extracts code lists, concept schemes, organisation schemes, data flows and key families. They will also parse SDMX-ML data files and extracts data sets, groups, series and observations. They will use the information available in the data structure definition (the dimensions, the attributes, the measures, etc) to interpret the statistical data.

The readers will then translate the extracted data into artefacts of the SDMX information model (see next section).

The code being modular, it is possible to add readers for other SDMX-ML formats, which are not currently supported, such as the SDMX-ML Generic format.

### 3.1.2 SDMX Information Model Package

The SDMX Information Model Package represents the artefacts defined in the SDMX information model (code lists, concepts, dimensions, data sets, series, observations, etc) as objects that can be stored in memory. These are the objects that can then be visualised, using the views defined in the third package (see below).

### 3.1.3 ECB Visualisations Package

The ECB Visualisation Package provides views displaying the artefacts of the SDMX information model in various ways (charts, tables, metadata panels, etc). It also supports the possibility to combine various views on a screen.

It follows the Model-View-Controller pattern and implements a few other design patterns (Command, Abstract Factory, Singleton, Observer, etc.).

### **3.1.4 A proof of concept: the SDMX navigator for the SDMX Statistical Web**

In order to prove that the ECB SDMX framework meets the requirements of displaying various types of statistical data in a generic way, an SDMX navigator was developed.

The [SDMX navigator](#) is a simple visualization tool for any SDMX-ML data file, in the Compact format, in the version 2 of the standard. It allows viewing the data as a graph or as a table. It also offers the possibility to look at the metadata. Using the appropriate filters at the top, it is possible to browse the SDMX-ML data file and select different series.

It can be tested with a file containing a reference to the data structure definition, for example, the Harmonised Competitiveness indicators ([http://stats.ecb.europa.eu/stats/df\\_test/hci.xml](http://stats.ecb.europa.eu/stats/df_test/hci.xml)) or the MFI interest rates ([http://stats.ecb.europa.eu/stats/df\\_test/interest\\_rates.xml](http://stats.ecb.europa.eu/stats/df_test/interest_rates.xml)). If the SDMX-ML data file does not contain a reference to the DSD (e.g.: [http://stats.ecb.europa.eu/stats/df\\_test/usd.xml](http://stats.ecb.europa.eu/stats/df_test/usd.xml)), the navigator asks to upload the DSD, in the Structure format (e.g.: [http://stats.ecb.europa.eu/stats/df\\_test/ecb\\_exr1.xml](http://stats.ecb.europa.eu/stats/df_test/ecb_exr1.xml)).

Many improvements would need to be made in order for the SDMX navigator to become an application of production quality, but it nevertheless achieved to prove that the ECB Flex framework meets the requirement to display various statistical data, without the need to hard code the knowledge of the data structure within the application.

### **3.1.5 Open-sourcing the SDMX framework**

In November 2008, the ECB released the SDMX framework as an open source product<sup>2</sup>. The Federal Reserve Bank of New York and Bank of Canada are partners in this initiative. Within 2009, various improvements will be made to the framework by the various contributors, such as adding support for other SDMX-ML Data formats as well as multilingual functionality. The code has been released under a so-called open source permissive licence.

## **3.2 The extension to the ECB website and the ECB Statistical Data Warehouse**

In this phase, new visualisation tools will be introduced on the ECB website and all the charts available in the ECB Statistical Data Warehouse will be replaced by dynamic graphs. This phase has started in March 2008 and produced a first deliverable in the form of the ECB Inflation Dashboard in November 2008.

---

<sup>2</sup> The code is available at the following location: <http://flex-cb.googlecode.com/>

#### **4. Phase 3: “Story-telling” approach. The development of ECB Dashboards**

As demonstrated above, the use of RIA technologies can vastly improve the visual display of statistical data, as well as the accessibility of the underlying data. However, if a “story-telling” approach is adopted, these technologies can also contribute making the data analysis more efficient and help improving the understanding of the data being displayed.

Therefore, the ECB started the development of interactive analytical dashboard with key statistics of interest for the general public. This 3<sup>rd</sup> phase of the project started in March 2008 and produced a first deliverable in the form of the ECB Inflation Dashboard in November 2008, with the following functionalities:

- Possibility to see, on one screen, key information about inflation in the euro area.
- Possibility to see the inflation, broken down by purpose of consumption or following the classification used at the ECB for economic analysis.
- Possibility to drill-down into the components of the two classifications mentioned above.
- Possibility to compare the evolution of components of inflation over time on a single graph.
- Possibility to play a movie showing the evolution of inflation over time.
- Possibility to sort the data following various criteria, such as the inflation rate or the component weight.
- Possibility to compare data for various countries.
- Possibility to view data as graphs or tables.
- Possibility to select the period of interest.

This [ECB inflation dashboard](#) has been released on the ECB web site in November 2008.

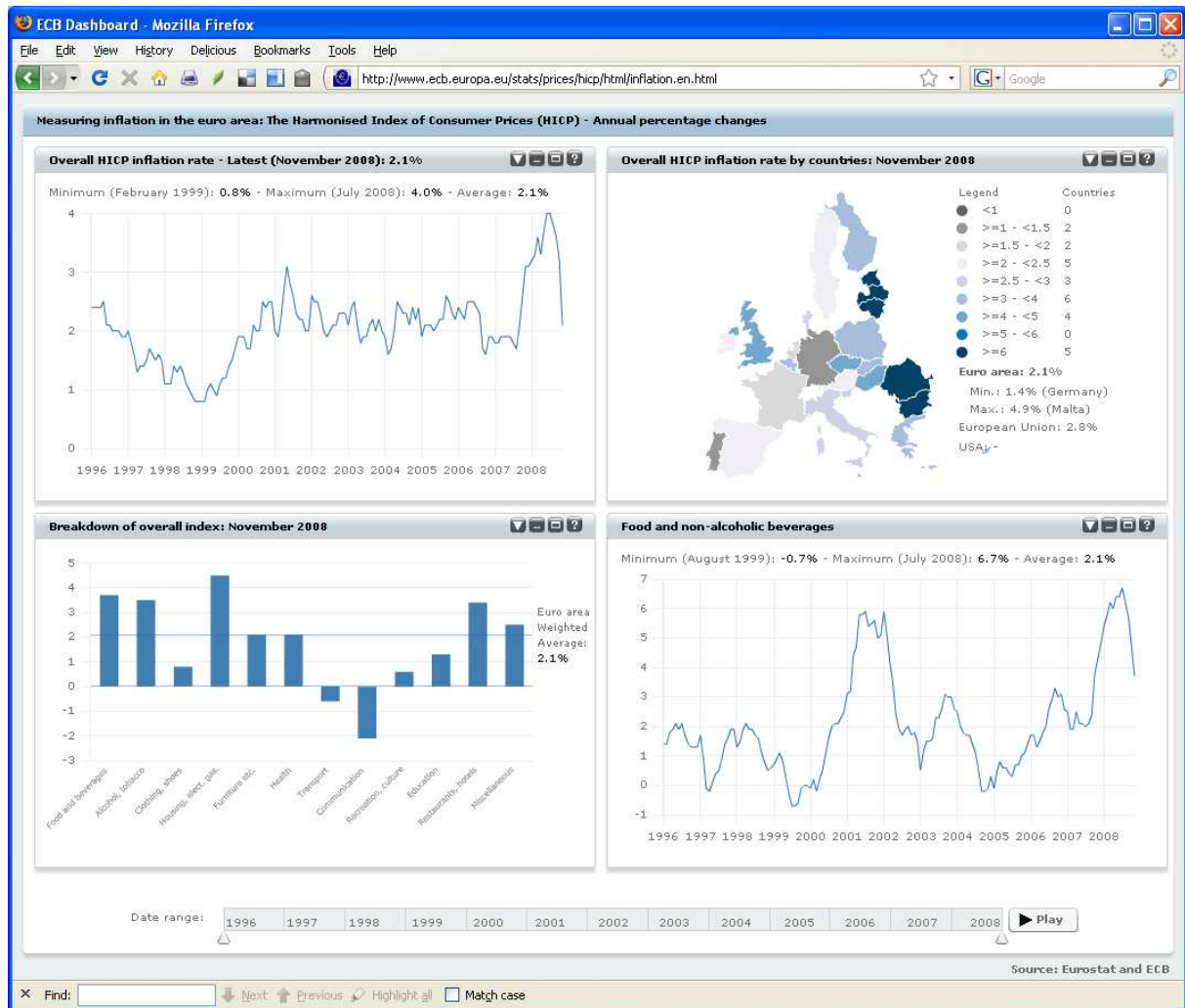


Figure 4 - The ECB inflation dashboard

## 5. Conclusions

Building visualisation tools using RIA technologies definitely improves the visual display and the accessibility of statistical data. Furthermore, adopting a “story-telling approach” (for example via the use of analytical dashboards), helps making data analysis easier and more efficient, thereby improving the understanding of the statistics published. Therefore, dynamic charts can be invaluable tools to improve communication towards the public. Finally, the use of a formal data framework based on a standard such as SDMX makes it possible to rapidly develop a rich and consistent statistical application.