Within the European System of Central Banks (ESCB) statistics (data and associated metadata) are exchanged in a fully automated manner. The data exchange system was set up to serve the ESCB’s data exchange requirements in a timely, flexible, robust and secure manner. It is based on platform-independent standards and supports advanced validation mechanisms. It allows partner institutions to build tools which facilitate accessibility and metadata management (using their local infrastructure) and to maximise benefits, especially in the context of multilateral data exchange. Overall, this data exchange framework contributes to improved timeliness, punctuality, accessibility, clarity and comparability of statistics.

1. Introduction

One year ago, the European System of Central Banks (ESCB)\(^1\) celebrated the fifth anniversary of automated statistical data exchanges. Back in March 1998 the Banco de España had created and sent to the European Monetary Institute (EMI)\(^2\) the first GESMES/TS message\(^3\) with live balance of payments data. Almost immediately, the balance of payments areas of Eurostat and the IMF started supporting this message for the exchange of live data. Soon after, the Bank for International Settlements (BIS) also joined the user community, together with its central banking partners from all over the world. The core element of the ESCB’s statistical data exchange (SDE) system is the GESMES/TS message. The SDE system ensures full automation, data integrity of the exchanged information, syntactic and semantic validation and zero conversion costs for all partners involved in the dozens of ESCB statistical interchanges\(^4\) that take place every week and which, nowadays, cover all domains of economic statistics. The message itself was designed jointly by the EMI, the BIS, Eurostat, the IMF and several experts from other institutions (especially central banks). It has also greatly benefited from valuable lessons learnt from previous pilot schemes (e.g. BOPSTA, GESMES/DSIS) and the expertise acquired by the institutions participating in its development from similar projects (i.e. BIS). GESMES/TS enjoys international recognition, its uses are continually expanding and it serves more and more economic domains and user communities. Nowadays, it is maintained under the auspices of

\(^1\) The European Central Bank (ECB) and the 25 EU national central banks (NCBs).
\(^2\) The EMI was the forerunner of the ECB.
\(^3\) Called GESMES/CB at that time.
\(^4\) These interchanges usually imply two-way transmissions, reporting (to the hub institution) and dissemination (from the hub institution).
SDMX, an initiative of seven international and European organisations (BIS, ECB, Eurostat, IMF, OECD, UN and World Bank). Even after five years, the GESMES/TS message is still characterised by the freshness and dynamism it had when first launched and its powerful data model is capable of serving new requirements and taking advantage of new technologies. It is also the core information model of the new emerging SDMX standards. More information on GESMES/TS can be found on the following websites:
- http://www.sdmx.org/

The contribution of GESMES/TS and of the SDE system to the quality of ESCB statistics stems from their positive role in the following quality dimensions:
- Timeliness and punctuality;
- Accessibility and clarity;
- Comparability of statistics.

2. The initial vision and assumptions

The vision of the ESCB data exchange system was to develop a robust, secure, paperless and fully automated data exchange facility which would be free of conversion costs and would incorporate advanced validation mechanisms, while also taking into account some restrictions in a pragmatic context:
- Not all statistical requirements, as they would evolve in future, were precisely known from the outset.
- We live in a statistical world in which complete statistical harmonisation is not always easily achievable; different concept definitions and inconsistent classifications might still need to be used for some years in the context of the different data exchanges administered by the various supranational organisations. Of course, European and international harmonisation should be strongly encouraged and pursued, but its absence should not be used as an excuse to postpone the implementation of automated processes.
- Different institutions use different platforms and the technical standardisation basis of the new facility should ensure interoperability, but without imposing any specific technology or proprietary formats. The telecommunication components may need to impose some common and technologically compatible approaches, but even these restrictions should be kept to a minimum.
- The ECB’s partner institutions in statistics (EU central banks and several EU statistical offices) also exchange data with the BIS (EU central banks) and Eurostat (EU statistical offices and central banks). In order to minimise the overall costs, the ESCB facility should also be compatible with the future directions and strategies followed by the BIS and Eurostat.

Looking back at the initial vision and the assumptions made, even after six years of rapidly evolving technology, the choice of the GESMES/TS message still looks to be the ideal means for satisfying the requirements and the constraints discussed above in an optimal manner.

3. The ESCB’s statistical data exchange system

Main components

For more than five years the SDE system has served all ESCB statistical data exchange requirements in a robust, reliable, secure, automated and timely manner. It is based on a “hub”
data exchange model, with the ECB being in the centre (hub); NCBs\(^5\) prepare data files (GESMES/TS messages) and send them to the ECB (“reporting”)\(^6\). The data exchanges between the ECB and NCBs take place through a closed network (ESCB-net). The ECB performs aggregations and other calculations (e.g. construction of additional indicators) and these are sent to the NCBs (“dissemination”). The dissemination via the ECB’s website and the production of publications are also highly automated, using as a source the ECB Statistical Data Bank (ESDB).

The SDE system is also capable of performing (and is actually used for) fully automated data exchanges with other international organisations; the Directorate General Statistics (DG-S) of the ECB takes care of this dissemination and the data become simultaneously available to internal users and to the NCBs. In addition to the data file-based dissemination, the ECB disseminates data via its website and, of course, prepares publications.

The ECB also exchanges statistical data with international and European organisations such as the IMF, the BIS and, of course, Eurostat. The graph below shows the data flows through the ECB’s SDE system:

Nowadays, the ECB administers fully automated data exchanges (GESMES/TS-based) through three different telecommunication channels:
- With the 15 NCBs that were members of the ESCB prior to 1 May 2004: through a closed network (“ESCB-Net”).

\(^5\) Apart from the NCBs, quite a few national statistical institutes (NSIs) are also involved in data exchanges with the ECB: data are sent to the ECB either directly (via internet and PGP/GPG encryption and digital signature) or via the corresponding NCB.

\(^6\) The SDE system is currently only “data file”-based; of course, when interactive access to statistics becomes available, broader aspects of functionality could be served.
- With the new EU Member States’ central banks and EU NSIs: via e-mail; the GESMES/TS messages are encrypted and digitally signed (with PGP/GPG) files that are attached to the e-mails.

- For the connection between the ECB and the other supranational institutions (i.e. Eurostat, BIS, IMF): special hardware (and software) for ensuring reliable and secure telecommunication.

In all three cases the exchanged files are processed automatically and eventually used to update the production or central databases of the receiving institution.

As stated above, in all data flows shown in the chart, only GESMES/TS messages are used.

The SDE system makes it possible to fully automate data exchanges by virtually “connecting” the databases of partner institutions through a telecommunication infrastructure and the exchange of standardised statistical messages.

Extensive flexibility is also provided: some partners may prefer to use their central or reference statistical databases (from which the “message extraction” takes place) for administering the data exchanges, while other partners may prefer to directly connect their local, decentralised database systems or applications (if these can be easily equipped with an automated GESMES/TS filter). However, other NCBs may prefer a mixed solution (e.g. using the decentralised production systems for reporting data to the ECB and using the central or reference statistical databases for loading the data disseminated by the ECB). Due to the “standardised” interchange means provided by GESMES/TS, the precise local organisational approach would not affect the system’s overall functionality.

Acknowledgements and error handling

GESMES/TS offers advanced error handling possibilities. Free software and source code (written in Java) are available which make it possible to “read”, “check” (syntactically and semantically) and “convert” GESMES/TS messages.

Most partners institutions use locally a special “GESMES/TS syntax checker” (written in Java) to syntactically and semantically check (against key family definitions, concept definitions and

---

7 It is expected that the new Member States’ central banks will also use the ESCB-Net for statistical data exchange by the end of 2004.

8 Currently, the ECB exchanges data with seven NSIs. However, although GESMES/TS is exclusively used as the message format in the data exchanges, full automation has not yet been achieved with all NSIs, due to the pending implementation of the PGP/GPG encryption and digital signatures in some of these NSIs. For security reasons, the ECB would not “trust” and perform automatic processing on incoming messages without digital signatures (i.e. without being able to verify the sender).

9 With the exception of some data file exchanges between the ECB and Eurostat (FAME databases). However, these have also been planned for conversion to GESMES/TS: the exchange of GESMES/TS files provides more advantages than the exchange of files of a proprietary format, even if two institutions use the same database system or applications. The advantage of GESMES/TS is that it also ensures full consistency between the sent and the expected semantics (data model consistency). For example, if two institutions exchange XLS files, although the files could be understood by humans when accessed, only an extremely precise agreement on naming conventions, tables and the location of each item in the cells would allow immediate processing by machines. GESMES/TS does exactly this (a text format) with an even more flexible syntax and via a much shorter and unambiguous language for describing the exact formatting agreements (for the exchange of data and rich metadata) than would be needed to exchange and validate Excel or CSV files comprising data and metadata and eventually make them appropriate for full automation.

10 That is, similarly to the SWIFT messages exchanged for payment transfers between banks.

11 This is the case in most central banks: their statistical databases are connected to the ESCB-network, the GESMES/TS messages are generated automatically and the full SDE system functionality (e.g. acknowledgements, logging, etc) can be used.

12 Of course, the role of a proper database infrastructure is crucial. The benefits of GESMES/TS would be considerably reduced if extensive manual work were needed for generating the messages.

13 Conversions to formats CSV, XML, HTML and FAME are possible with the existing software.
code lists) their messages before sending to the ECB. Nevertheless, if a message reaches the ECB with such errors, the ECB’s system automatically responds to the sending institution with a message describing the errors found.

**Benefits in a multilateral data exchange environment**

When a partner institution interacts via GESMES/TS with several institutions in a data exchange environment, the benefits enjoyed grow enormously (especially if there are properly organised database systems from which the automated production of GESMES/TS messages takes place). The graph below shows such a situation:

This picture looks complex. However, when the production and exchange of GESMES/TS messages occurs in a highly automated manner, it actually makes it possible to satisfy several organisational requirements, peculiarities and obligations of the current complex world at very low costs. Such costs would mainly result from the “mapping” work for the codes used, this being the consequence of the lack of statistical harmonisation; such costs would anyway be incurred if different data structure and classification schemes needed to be used for reporting to the various institutions (as compared with the internally used classification and coding schemes).

**The prerequisites for using a fully automated data file exchange system**

It should be noted that the use of the ESCB’s SDE system (and of any fully automated data file or data sharing system) requires two agreements by the user institutions; these agreements simply ensure that full automation is feasible at the local level and in a secure manner:

- **Security: an agreement on the telecommunication means**
  - For example, the ECB with its partners uses a variety of means to ensure maximum flexibility (e.g. a closed network with EU central banks, a special infrastructure with the BIS, Eurostat and the IMF, internet and GPG/PGP for encryption/digital signatures with all other institutions). For security reasons, secure communication is a prerequisite for full automation. In all these cases, GESMES/TS files are used almost exclusively (see bullet point below).

- **An agreement on the syntax and the semantics of the exchanged information**
  - GESMES/TS is the agreed message format in the ESCB.
  - Moreover, in order to start the actual data exchange, an agreement is also needed on the “semantics” that will be used in the exchanged statistical messages (i.e. definitions of

---

14 For example, different reporting requirements for the various central institutions, different release calendars for different versions of data, several legal acts determining possibly different reporting obligations, different data release arrangements by the various national and supranational institutions towards their partner institutions, various levels of confidentiality for different datasets or for versions of datasets under alternative definitions and classifications, the need for data cross-checking between central institutions, time lags between the release of national data for internal use by institutions compiling EU and euro area aggregates and the release to the general public (e.g. national and EU or euro area aggregated data may need to be released simultaneously), etc.
concepts, definitions of data structures (“key families”), underlying code lists, lists of time series to be exchanged). These semantics, in GESMES/TS terms, are called “structural definitions”, and they are disseminated by the ECB to its partners. When changes are needed (e.g. when new data need to be exchanged), the ECB disseminates (in advance) the updated structural definitions.

These prerequisites are very easy to meet for any other institution with which the ECB would like to exchange data because:

(i) GPG is a very flexible and broadly accepted solution\textsuperscript{15} for securing communication via the internet, such as the exchange of e-mail attachments; and

(ii) GESMES/TS is a simple text format (and free software for format conversions is already available).

4. Why GESMES/TS has proven so successful for the ESCB

GESMES/TS has proven extremely beneficial for exchanging statistical data and metadata, at least within the ESCB. Partner institutions have managed over the last six years to increase spectacularly the number of statistical domains served and the number of time series exchanged. Everything that could theoretically be automated has been automated, already at an early stage. Over the last two years, the central banks of the then 12 accession countries have also successfully joined the GESMES/TS exchange community, participating in the two-way data exchanges with the ECB (and with Eurostat). The reasons why GESMES/TS has proven so successful can be summarised as follows:

- **GESMES/TS is platform-independent**: institutions can use any internal database already familiar to staff. The messages are extracted from the database system used, and when GESMES/TS messages are sent by another institution, the receiving institution can (automatically) load the contents onto the internal database system and offer the data to end-users (or for further processing) via the databases, browsers and tools already familiar to the internal users.

- **GESMES/TS makes it possible to build completely automated systems for data and metadata exchange**, including the exchange of data structure descriptions, code lists and definitions of statistical concepts. It can also be combined with appropriate monitoring of the incoming and outgoing messages, error handling, logging and archiving of related information.

- **It supports the exchange of “structural definition” messages**: moreover, these definitions do not interfere with the core syntax of the message and they are exchanged separately from the messages carrying the actual data/metadata content. In order to cover new domains, it is easy to quickly develop the definition of additional data structures and to describe them in a GESMES/TS message or (for facilitating reading by humans) just on one page, without complex technologies.

- **Free software, code and tools** are already available, developed mainly by Eurostat and the ECB, but also by other partners and institutions. Commercial software is not needed, as a large stock of computer code is already available for free.

- **The message can also be used easily in two-way data exchange**: institutions report data to the ECB in GESMES/TS, but the ECB also sends GESMES/TS data files to its partners, allowing the same information to be shared easily within the ESCB. The message, and the

\textsuperscript{15} The software is free and can be downloaded from the internet.
automated processing it supports in a platform-independent manner, acts as a catalyst allowing also smaller institutions with limited human resources to access and load large quantities of statistical information.

- **A very rapidly growing user community with international support** (see also next section on critical success factors). The GESMES/TS data model and user guide is maintained under the auspices of the SDMX initiative.

- **A powerful data model:**
  - GESMES/TS allows easy coverage of any domain of statistics (or any domain or phenomenon where time would be a relevant dimension).
  - GESMES/TS is characterised by syntactic stability: the message structure has remained virtually fixed over the past five years. This makes it possible to minimise development and maintenance costs, basically limiting them to the initial investment.
  - GESMES/TS can also be used as the data model for internal storage of the data and the associated metadata.

- **GESMES/TS encourages statistical harmonisation** (in data structures, concepts, code lists) across central organisations using GESMES/TS (it makes the benefits of reducing the “mapping” costs more transparent). This is a very important indirect benefit: as GESMES/TS needs structured information, it encourages institutions to harmonise concepts and to converge towards common definitions. Although harmonisation is not a prerequisite for using GESMES/TS (allowing, therefore, full automation even in today’s complex statistical world), the benefits of GESMES/TS for partner institutions are even more apparent in the context of European or international harmonisation.

5. **Critical success factors**

It has been found that, under certain conditions, GESMES/TS can help user institutions to maximise their benefits on an even larger scale. So although the factors described below are not preconditions for using GESMES/TS, if these conditions are satisfied, further optimisation would be possible:

- **Well structured internal database systems**: unfortunately, if internal data are not properly managed and stored, GESMES/TS cannot help to optimise processes. GESMES/TS shows its strength and makes benefits visible mainly when combined with good database systems (possibly through centralised or reference databases) from which the production of GESMES/TS can be planned as a regular generation of text files under certain conditions (e.g. when data to be reported are due or when a new dataset has been used to update the database or, on request, by pressing a button).

- **Metadata to be managed in close relation to the data**: this ensures that the metadata aspects will not be disregarded or given low priority. Managing metadata either in the same system as the data or in different systems but with strong links is helpful for the production process, but also for the manner data and manadata are delivered and accessed by users. For example, if important functions related to managing changes in metadata or to confidentiality flags are managed through completely separated (and unlinked from the data) systems, the underlying processes tend not to be properly maintained over time.

- **Central institutions need to pay special attention to ensuring the proper coordination and management of their “structural definitions”**: the function of “structural definition” coordination is very important in a central institution for maximising partners’ benefits from
the use of GESMES/TS. This coordination function needs to ensure interaction in a coordinated manner with the subject matter areas that manage the requirements at the statistical level. These requirements then need to be properly formulated in technical terms (structural definitions) and commonalities to be maximised (e.g. in concept definitions or code lists used). Another important aspect is the interaction needed with other central institutions devising structural definitions, in order to ensure convergence to a more advanced level of international harmonisation. These aspects are very important not only for the statisticians, but also for reducing “mapping” costs (e.g. for codes and keys used) and making access to end-users more homogeneous across the various domains of statistics.

- **Easy access to combined statistical and technical expertise** (e.g. on database issues): it seems that statistics departments with poor access to dedicated statistical IT or database support tend to use GESMES/TS in a sub-optimal manner, e.g. by introducing additional manual work through spreadsheets or through risky (in terms of possible errors) and labour-intensive operations with intermediate formats such as CSV. The ECB and Eurostat frequently organise GESMES/TS courses that help to develop GESMES/TS expertise.

- **The possibility of exchanging data with more than one institution**: this is simply a factor that makes the benefits even more significant. For example, when an institution has automated the production of GESMES/TS messages and uses them to report to the ECB, Eurostat and other institutions (and/or to receive such messages from them) the investment pays off even faster. Also, with a minimum investment in a “loading” filter, the possibility of receiving GESMES/TS messages from various central organisations allows the receiving institutions to process massive amounts of incoming data from various sources at a low cost.

- **Good and secure telecommunication means** allow for the complete integration and automation of the data exchange. For this to be possible, at least authentication of the sender and encryption of the transmitted data (at least when data are confidential) are necessary.

- **Commitment through proper consultations and/or appropriate legal framework**: commitment, communication and consultations are very important, especially in the first phases of a data exchange project in which GESMES/TS is used for the first time. In this phase interaction with other experienced institutions also proves very useful.

- **Independence from commercial software and providers**: free and open source created and exchanged between partners facilitates the use of the agreed standards and, moreover, contributes to a spectacular improvement of the code/software used in production by the partner institutions.

- **Openness to future enhancements and developments**: this reflects the dynamic aspects and the capability to respond to new (not initially foreseen) requirements in an efficient, coordinated and timely manner (e.g. dissemination via websites in a more harmonised manner across countries through, for example, compliance with the emerging SDMX standards).

6. **How the automated data exchange system has been contributing to the quality of ESCB statistics**

The quality of the statistics it produces is a major concern for the ESCB. Evaluating the experience of automated data exchange in the ESCB for more than six years, a positive contribution to quality can be attributed through the following quality dimensions:
• Timeliness\textsuperscript{16} and punctuality\textsuperscript{17}: automated systems contribute to faster processing and reporting. Moreover, the ESCB’s system includes automated error alert mechanisms and these considerably speed up the problem-solving process.

• Accessibility: the discipline imposed on the metadata structure (and their syntax) and the clear “prescription” of the exchanged data (through the predefined structural definitions of the GESMES/TS data model) ensure that the data and metadata used by the browsers to access the data/metadata are optimised with respect to the “technical” factors (e.g. ability to configure browsers at a low maintenance cost through the GESMES/TS structural definitions).

• Clarity: the exchange of metadata is foreseen through the GESMES/TS data model and, for each dataset, the actual metadata to be exchanged (attributes) are defined through the agreed corresponding key family.

• Comparability: the GESMES/TS data model imposes clear descriptions for each dimension value. In particular:
  • \textit{Comparability over time}: this is ensured by the continuing effort to avoid “restructuring” actions in the GESMES/TS key families.
  • \textit{Comparability between geographical areas}: for a given dataset the same GESMES/TS “key family” is used across the reference regions (EU countries); it is expected that in future, with the growing use of GESMES/TS (and the growing effort towards harmonisation) around the globe, there will be greater motivation to share the use of harmonised key families and improve the comparability also with regions beyond Europe.
  • \textit{Comparability between (non-geographical) domains}: the use of GESMES/TS makes very transparent the inconsistencies in concept definitions and code lists between domains (e.g. national accounts, monetary statistics, financial accounts, government finance statistics). This helps to intensify the effort towards convergence to common semantics (when/where feasible).

\textsuperscript{16} Time elapsed between the reference period (to which the data refer) and the actual delivery date.

\textsuperscript{17} Time lag between actual delivery date of data and target delivery date (when data should be delivered).