

Unclassified

STD/SIMS(2004)9



Organisation de Coopération et de Développement Economiques
Organisation for Economic Co-operation and Development

29-Mar-2004

English - Or. English

**STATISTICS DIRECTORATE
STATISTICAL INFORMATION MANAGEMENT AND SUPPORT DIVISION**

**STD/SIMS(2004)9
Unclassified**

The Use of Gesmes/TS in the Statistical Data Exchanges of the European System of Central Banks

EXPERT GROUP ON STATISTICAL DATA AND METADATA EXCHANGE

**1-2 April
Château de la Muette, Paris**

JT00161021

Document complet disponible sur OLIS dans son format d'origine
Complete document available on OLIS in its original format

English - Or. English

**A CASE OF FULLY AUTOMATED INTERCHANGE OF STATISTICS:
THE USE OF GESMES/TS IN THE
STATISTICAL DATA EXCHANGES OF THE
EUROPEAN SYSTEM OF CENTRAL BANKS**

Introduction

One year ago, the European System of Central Banks (ESCB)¹ celebrated the fifth anniversary of automated statistical data exchanges. It was back in March 1998 that the Banco de España created and sent to the European Monetary Institute² the first GESMES/TS message³ with live balance of payments data. Almost immediately, the balance of payments areas of Eurostat and the IMF started supporting the message for the exchange of live data. And, 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 GESMES/TS message is the core element of ESCB's statistical data exchange system (SDE), ensuring full automation, data integrity, syntactic and semantic validation and zero conversion costs for all partners involved in the dozen of ESCB's statistical interchanges⁴ that take place every week and which cover nowadays 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. 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 SDMX, an initiative of six international and European organisations (the BIS, the ECB, Eurostat, the IMF, the OECD and the UN). Even after five years, the message is still characterised by the freshness and dynamism of its first appearance and its powerful data model is ready to serve new requirements and take advantage of new technologies. More information on GESMES/TS can be found on the following web pages:

- <http://www.sdmx.org/>
- <http://www.ecb.int/> (“Statistics”, “GESMES/TS”)

¹ The European Central Bank (ECB) and the 15 EU national central banks (NCBs).

² the European Monetary Institute (EMI) was the forerunner of the European Central Bank

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).

The Initial Vision and Assumptions

The vision about the ESCB data exchange system was to develop a robust, secure, free of conversion costs, paperless and fully automated data exchange facility, incorporating advanced validation mechanisms, however taking also 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 the complete statistical harmonisation is not always easily achievable; different concept definitions and inconsistent classifications might need to be used still for some years in the context of the different data exchanges as administered by the various supranational organisations; of course, European and international harmonisation should be strongly encouraged and pursued, however, 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 to be used; only the telecommunication components could be the ones that might need to impose some common and technologically compatible approaches, but even these restrictions should be kept at a minimum.
- ECB's partner institutions in statistics (EU central banks and several EU statistical offices) exchange data also 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 into the initial vision and the assumptions made, even after six of rapidly evolving technologies, the choice of the GESMES/TS message looks still to be the ideal means for satisfying the requirements and the constraints discussed above in an optimal manner.

ESCB's Statistical Data Exchange System (SDE)

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): NCBs⁵ 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 through the ECB web site 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 (and it is actually used) for fully automated data exchanges with other international and The Directorate General Statistics of the ECB (DG-S) takes

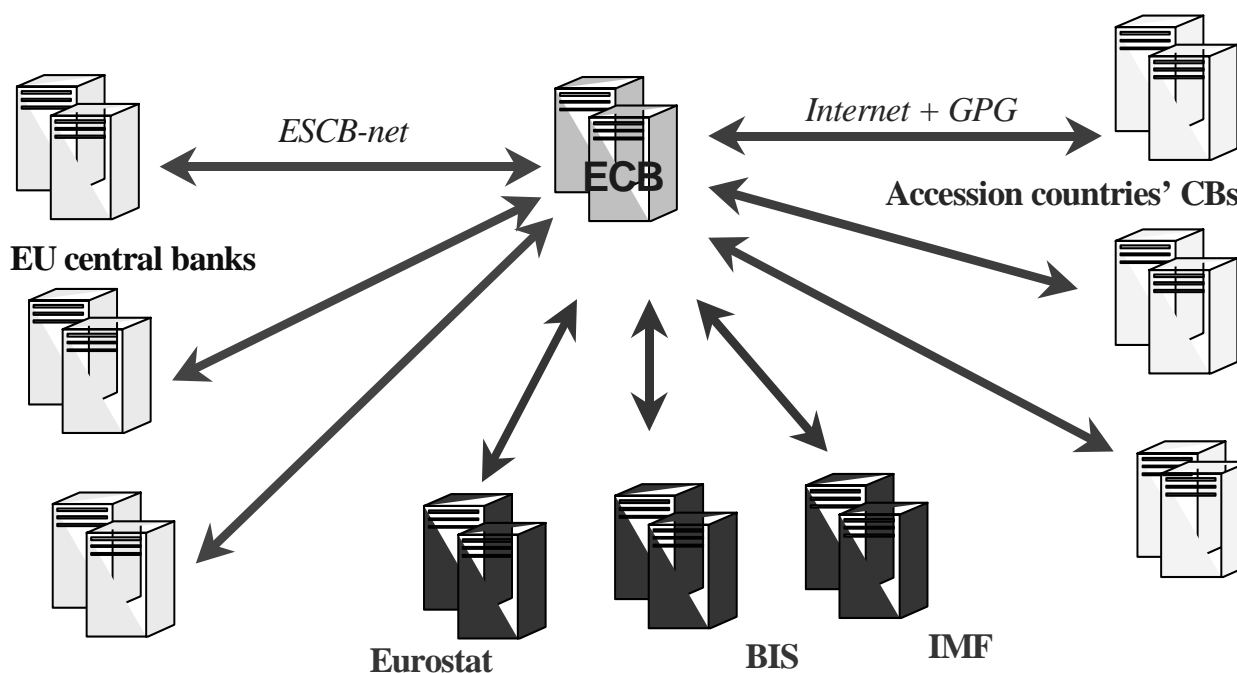
⁵ Apart from the national central banks (NCBs), also quite a few national statistical institutes (NSIs) are also involved in data exchanges with the ECB: data are sent to the ECB either directly (via internet and PGP/GPG encryption and digital signature) or via the corresponding NCB.

⁶ The SDE system is currently only "data file" based; of course, when in future "interactive access to statistics" will also become available, broader aspects of functionality could be served.

care of this dissemination and the data become simultaneously available to its internal users and to the NCBS. Apart from the data file based dissemination, the ECB disseminates data through its web site 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 ECB's SDE system:

Nowadays, the ECB administers fully automated data exchanged (Gesmes/TS based) through three



different telecommunication channels:

- With the EU central banks: through a closed network (“ESCB-Net”);
- With Accession countries' central banks and EU national statistical institutes (NSIs)⁷: through internet mail; the GESMES/TS messages are encrypted and digitally signed (with PGP/GPG) files that are attached in 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 they are eventually used to update the production or central databases of the receiving institution.

⁷ Currently, the ECB exchanges data with seven NSIs. However, and though GESMES/TS is exclusively used as the message format in the data exchange, full automation has not yet been achieved with all, 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 (thus, without being able to verify the sender).

As said, in all data flows shown in the chart, exclusively Gesmes/TS messages are used⁸.

The SDE system allows to fully automate the data exchanges by virtually “connecting” the databases of the partner institutions through a telecommunication infrastructure and the exchange of standardised statistical messages⁹.

Extensive flexibility is also provided: some partners may prefer to use their central or reference statistical databases (out of which the “message extraction” will take place) for administering the data exchanges¹⁰, while other partners may prefer to directly connect their local - decentralised database systems or applications (of course, if these can be easily equipped with an automated Gesmes/TS filter). Yet, 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, a decision on the precise local organisational approach would not affect system's overall functionality.¹¹

Acknowledgements and error handling

Gesmes/TS offers advanced error handling possibilities. Free software and source code (written in Java) is available: it allows to “read”, “check” (syntactically and semantically) and “convert” Gesmes/TS messages¹².

Most partners institutions use the “Java syntax checker” to syntactically and semantically check (against key family definitions, concept definitions and code lists) their messages before sending to the ECB. Nevertheless, if a message reaches the ECB with such errors, ECB's system automatically responds back to the sending institution with a message describing the errors found.

Benefits in a multilateral data exchange environment

When a partner institution is involved in a data exchange environment interacting via Gesmes/TS with several institutions, the benefits enjoyed grow enormously (especially, if there are properly organised

⁸ With the exception of some data file exchanges between the ECB and Eurostat (FAME databases). However, also these have been planned for conversion to GESMES/TS: actually, 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 data base system or applications; the advantage of GESMES/TS is that its use also ensures full consistency in the sent and the expected semantics. For example, let's think of an exchange of XLS files between two institutions: though these files could be understood by humans, only an extremely precise agreement on the table formats and locations of each item in the cells would allow processing by machines: GESMES/TS is doing 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 what one would need invent for exchanging and validating Excel or CSV files comprising data and metadata and making them eventually appropriate for full automation.

⁹ E.g. similarly to the SWIFT messages exchanged for payment transfers between banks.

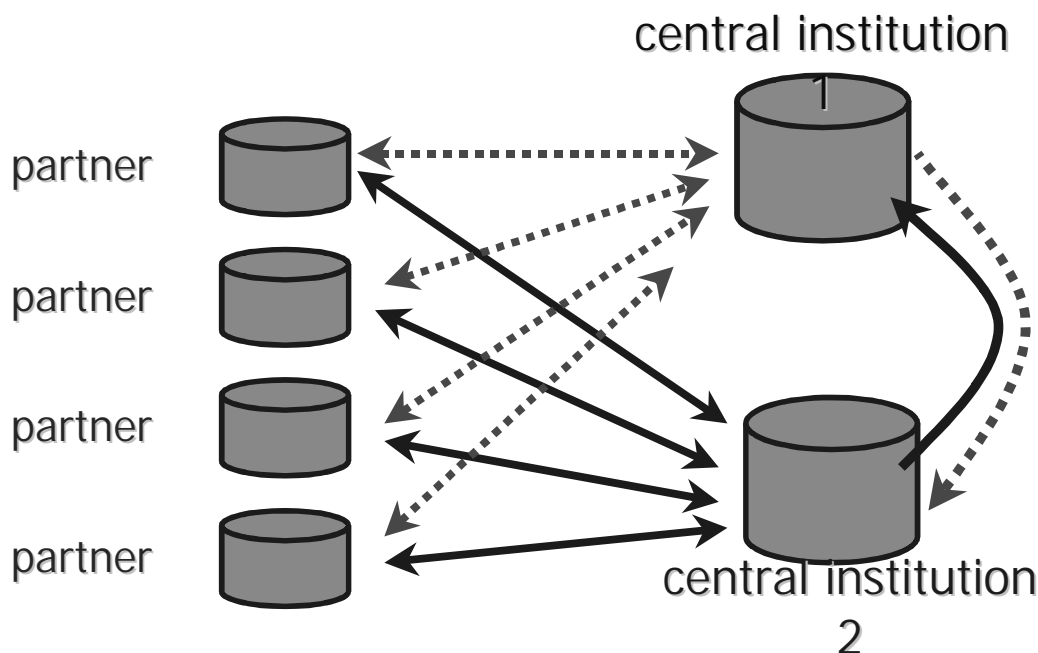
¹⁰ This is the case in most central banks: their statistical databases are connected to the ESCB-network, the GESMES/TS messages are generated automatically and a use of the full Statistical Data Exchange System functionality becomes possible (e.g. acknowledgements, logging, etc).

¹¹ Of course, the role of a proper database infrastructure is crucial. The benefits from the use of GESMES/TS would be considerably reduced if extensive manual work was needed for generating the messages.

¹² Conversions to text, XML, HTML, FAME are possible with the existing software.

database systems out of which the automated production of Gesmes/TS messages takes place). The graph below shows such a situation:

Though this picture looks complex, when the exchange of Gesmes/TS occurs in a highly automated



manner, it actually allows to satisfy several organisational requirements and obligations of the current complex world¹³ at very low costs: these costs should simply cover the “mapping” work for the codes used, this being the consequence of the lack of statistical harmonisation; and these costs would anyway be present, as long as there are different data structure and classification schemes that need to be used for reporting to the various institutions.

The prerequisites for using a fully automated data file exchange system

It should be noted that the use of 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 ensuring 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

¹³ e.g. different reporting requirements by the various central institutions, different release calendars for different versions of data, several legal acts determining possibly various reporting obligations, different data release arrangements by the various national and supranational institutions towards their partner institutions, various levels of confidentiality for various data sets or for versions of data sets under alternative definitions and classifications, needs 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 release to the general public (simultaneously, national and EU or euro area aggregated data), etc.

reasons, the secure communication is a prerequisite for allowing full automation. And in all these cases, almost exclusively GESMES/TS files are used (see bullet 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, **also an agreement is needed on the “semantics” that will be used in the exchanged statistical messages** (i.e. definitions of concepts, definition 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.

Actually, these prerequisites are very easy to meet for any other institution which the ECB would like to exchange data with:

- (i) for securing the exchange of e-mail attachments: GPG is a very flexible and broadly accepted solution for securing the communication over the internet¹⁴;
- (ii) GESMES/TS is a simple text format (and free software for format conversions is already available).

Why GESMES/TS Proved So Successful for the ESCB

GESMES/TS has been proven extremely beneficial for exchanging statistical data and metadata, at least in 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 they exchange. And everything that could theoretically be automated has been automated, already at an early stage! Recently (last two years), the central banks of the 12 accession countries also joined the GESMES/TS exchange community very successfully, 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 data base already familiar to the staff; the messages are extracted out of the database system used; and, when GESMES/TS messages are sent by another institution, the receiving institution can (automatically) load the content on 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. Also its use can 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 intervene in the core syntax of the message and they are exchanged separately from the messages carrying the data/metadata actual content; and, 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 needing complex technologies.

¹⁴ The software is free and it can be downloaded from the Internet.

- **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 also the ECB sends GESMES/TS data files to its partners, allowing the same information to be shared easily within the ESCB. The message, and the automated processing it supports in a platform independent manner, act 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 maintenance of the GESMES/TS data model and user guide is 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 has enabled the message with syntactic stability: the message structure has remained virtually fixed over the past five years! This makes it possible to minimise development and maintenance costs, limiting them basically to the initial investment.
 - GESMES/TS can also be used as the data model for internal data 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 from 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 to common definitions. Although harmonisation is not a prerequisite for using GESMES/TS (allowing, therefore, full automation also in today's complex statistical world), the benefits of GESMES/TS for partner institutions are even more apparent in a context of European or international harmonisation.

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 of messages' use 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) out of 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 data set has been used to update the database or, on request, by pressing a button).
- Central institutions need to pay special attention ensuring a proper co-ordination and management of their “structural definitions”: The function of the “structural definition” co-ordination is very important in a central institution for maximising partners' benefits from the use of GESMES/TS. This co-ordination function needs to ensure an interaction in a co-ordinated

manner with the subject matter areas that manage the requirements at the statistical level. These requirements would 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). Also, an important aspect is the interaction needed with other central institutions devising structural definitions, in order to ensure the 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.

- **The possibility to easily have access to some 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. introducing additional manual work through spreadsheets or through risky (for 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, the possibility to receive GESMES/TS messages from various central organisations allows the receiving institutions with a minimum investment in a “loading” filter to massively process at low costs incoming data from various sources.
- **Good and secure telecommunication means** allow to completely integrate and automate the data exchange. For this, 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 the interaction also with other experienced institutions proves very useful.

Does GESMES/TS Have a Future in the ESCB?

Definitely! The benefits that the ECB, the 15 EU national central banks, the EU statistical institutes and the central banks of the 12 accession countries (and all other user institutions in Europe and all over the globe) already enjoy reporting data (or making data available to other institutions) with minimal maintenance costs. These benefits are impressive compared with the past, when paper, fax or unstructured spreadsheets and CSV files were used. Moreover, new technologies and approaches, such as the XML, are expected to further enrich the current range of GESMES/TS applications and tools available (e.g. in the context of the SDMX initiative). The old dream of allowing official statisticians to be devoted to quality and further development of statistics, minimising their distraction and involvement to manual production and paper data exchange has never before come so close to becoming a reality. And the GESMES/TS message has played a major part in making this dream come true, in the ESCB, over the last five years.