

SEMINAIRE

E+; 3=! C

SEMINAR

STATISTICAL COMMISSION AND
 ECONOMIC COMMISSION FOR EUROPE



Distr.
 GENERAL

CONFERENCE OF EUROPEAN
 STATISTICIANS

CES/SEM.43/20
 17 February 2000

ENGLISH ONLY

Seminar on integrated statistical information
 systems and related matters (ISIS 2000)

(Riga, Latvia, 29-31 May 2000)

Topic III: Innovations in data collection and exchange

EUROSTAT'S INNOVATIONS IN DATA EXCHANGE

Invited paper

Submitted by Eurostat¹

I. INTRODUCTION

1. Eurostat is currently collecting over 250 dataflows from Member States with varied volume and periodicity. This represents about 10,000 files that are received each year in Eurostat. To assist statisticians either at Eurostat or in the National Administration in their data transmission tasks, several projects have been launched in order to solve specific problems:

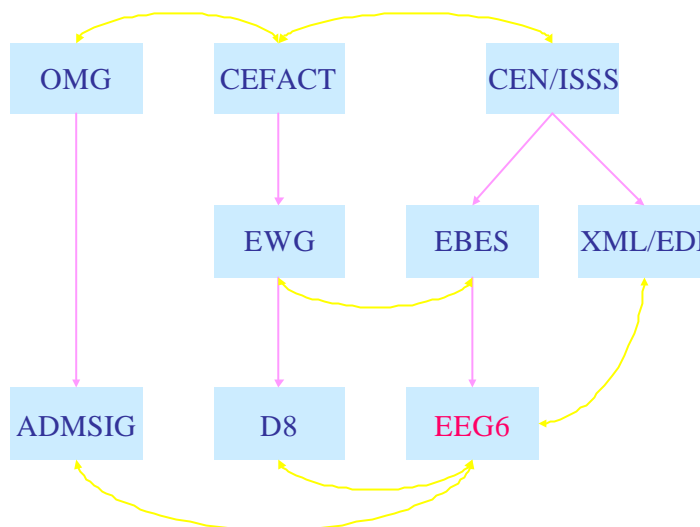
- from the message standardisation (**GESMES**),
- the data collection monitoring (**STADIUM**) and its inventory (**EDIFLOW**),
- up to the transmission modules hiding the telecommunication layer (**STATEL**).

II. GESMES

2. GESMES (Generic Statistical Message) is an EDI message developed by EBES Expert Group 6, as an official international standard of the UN EDIFACT Working Group.

3. GESMES has been developed for the exchange of aggregated statistical data or, in other words, the exchange of multidimensional statistical tables and time series. The principal organisations involved in the exchange of this data are the national statistical offices, national central banks, and those organisations playing a role at the international level, such as the UN/Economic Commission for Europe (UN/ECE), OECD, Eurostat, the European Central Bank and the Bank for International Settlements.

1 Prepared by Wolfgang Knueppel.



Example of a GESMES message for Maritime Transport Statistics

4. The message is built up by three-letter segments containing administrative data (sender, receiver, name of dataflow, date etc), and the data itself (ARR - array segment):

```

UNA:+,? '
UNB+UNOC:3+TEST-GESMES+EUROSTAT+990528:1447+ES97Q3A1++GESMES++++1'
UNH+ES97Q3A1+GESMES:2:1:E6'
BGM+74'
NAD+MS+TEST-GESMES'
NAD+MR+EUROSTAT'
DSI+TEST-GESMES-TRANS-Q'
STS+3:142+9:142'
DTM+Z02:19973:608'
IDE+5+SMRTM-A1-Q'
GIS+AR1'
ARR++1997:3:ESHUV:2:ESBIO:0111:1:.....:5023'
ARR++1997:3:ESHUV:2:ESCAS:0112:1:.....:656'
ARR++1997:3:ESHUV:1:ESSCT:0112:1:.....:523'
ARR++1997:3:ESVIL:2:GBAVO:0061:9:.....:5123'
UNT+14+ES97Q3A1'
UNZ+1+ES97Q3A1'
  
```

5. GESMES is able to cope with the many and varied types of array data to be exchanged and supports the exchange of statistical figures together with their associated metadata (e.g. footnotes, methodological notes, data set structures). GESMES works in synergy with CLASET, the message for the exchange of classification.

6. Various profiles of GESMES have been developed under which the time series oriented GESMES/CB has found widest acceptance in the field of central and international banking, within and outside the European System of Central Banks.

7. GESMES can also be used as a request message for data. This message can be used to trigger the sending of specified data from a partner. It is, however, not often used because complex systems are required to handle it.

8. The reason for GESMES implementation is to improve the quality of the information transmitted, reduce the costs for national administrations and Eurostat,

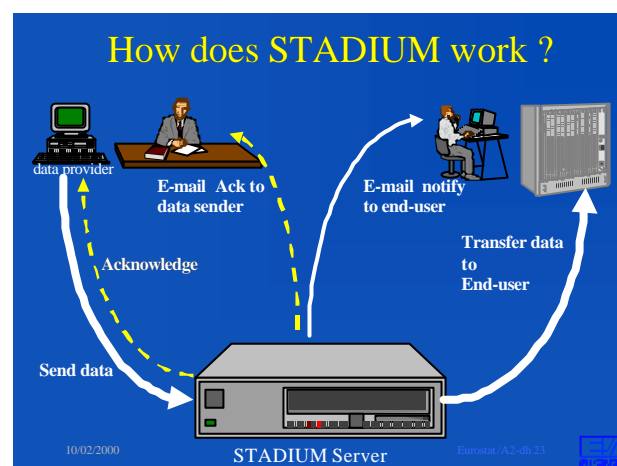
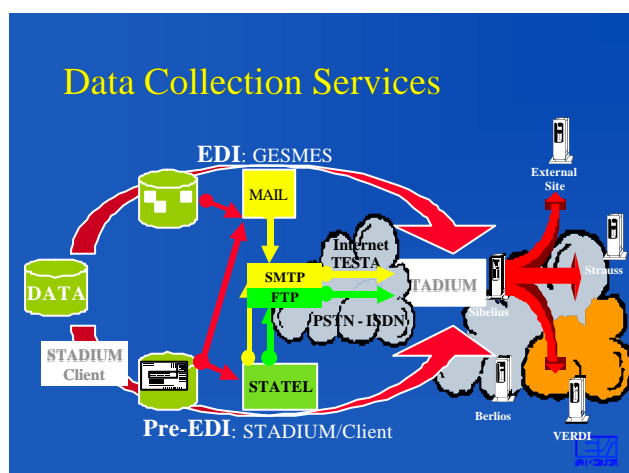
and to improve the timeliness of production of statistical data at the international level.

9. In order to increase the acceptance of GESMES, XML and OO representations are under development, in addition to the existing EDIFACT message. The world-wide standardisation of these representations is planned under the umbrella of XML-EDI (in cooperation with the newly formed ebXML initiative, bringing together the WWW Committee, OMG, UN/CEFACT, and others) and the OMG, respectively.

10. For several years, countries outside of the European Union and EFTA have been actively participating in EEG6 and its working groups. Training sessions in GESMES, especially for central European countries, have been held, further seminars will be available on request and budgetary availability, and an EDI road show for member states, accession countries, and other parties interested in our standards is in preparation. The newly formed EDI implementation task force shall be open to accession countries.

III. STADIUM

11. Eurostat has developed a solution for electronic data collection and related monitoring software called STADIUM. There are currently about 80 STADIUM-client workstations installed in Member States, each sending statistical data to Eurostat, from a few Kb to several Mb on a daily or an annual basis. In addition, the STADIUM server can also receive EDI files (GESMES/EDIFACT) generated by our partners.



12. STADIUM was conceived as a means of rationalising the process of data collection, by creating a centralised collection service. It is, therefore, a general service for collecting statistical data submitted to Eurostat, and distributing them to the Eurostat Units in charge of processing them. It is a portable, open, wide-area application-to-application statistical data transfer management service. It is a managed and controlled environment: the various applications, and the people in charge, are registered in the STADIUM database as senders or receivers of specific information. STADIUM performs the routing and monitoring operations, but is not responsible for the content or format of transfers. It is presently installed at Eurostat, but could be installed at any statistical data collection centre.

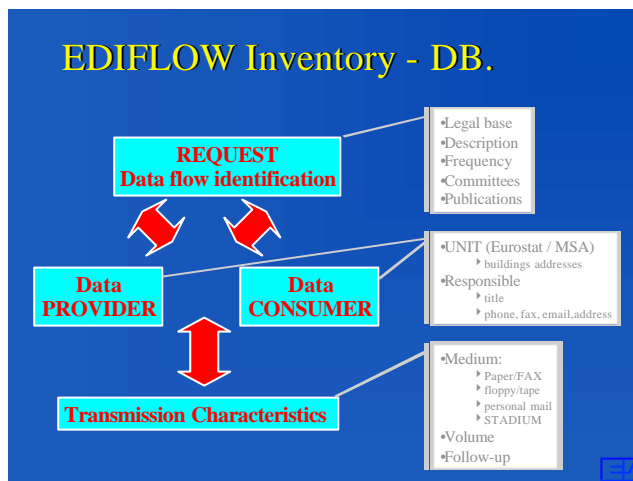
13. In the future, new features will be added to the STADIUM services, mainly in the "workflow" area.

14. Eurostat has extended data exchange beyond the boundaries of the European

Communities. In particular, it has begun to organise data collection via STADIUM from some candidate and western Balkan countries. This concerns currently two important data flows, the Labour Force Survey and External Trade Statistics.

IV. EDIFLOW

15. EDIFLOW, the inventory of all dataflows between Eurostat and the national administrations providing data, serves as the backbone of data transmission. It identifies all dataflows and all partners involved on the level of the individual sender and receiver.



16. In addition, it serves two purposes:

- Promotion of the use of telematic tools (in relation and in coordination with the projects TESTA, STATEL, STADIUM, and GESMES) in the process of data exchange between Eurostat and the Member States.
- Linked to STADIUM it will provide an alert system allowing to identify and to request overdue data.

➤ Sites distribution in EDIFLOW :

	EC	EFTA	CEEC	MEDSTAT	TOTAL
Countries	15	4	12	11	42
Administrations	278	23	36	11	348
People	1274	79	62	25	1440

➤ Distribution of the periodicity for the data flow defined in EDIFLOW :

	D Daily	W Weekly	M Monthly	Q Quarterly	A Annual	X more	S Sequential	Total
EDIFLOW	---	---	29	46	135	27	8	245

17. EDIFLOW may be consulted and updated by all partners intervening in the data exchange. EDIFLOW covers mainly the Member States of the European Communities but Eurostat has already started to collect information on data flows from central European countries.

V. STATEL

18. STATEL provides common telecommunication interfaces for trans-European applications. STATEL is an API (Application Programming Interface) hiding all the telecommunication layers to an application and allowing file transfer between remote partners.

19. The main STATEL characteristics are:

- Unique API with a multi-protocol ability (Kermit, FTP, mail),
- Automatic segmentation for big volume transfer, with automatic retransmission requests,
- Data encryption, protecting files content during transmission.

20. STATEL runs under Win-XX, NT and UNIX. Telecommunication between the STADIUM-Client and the STADIUM-Server is mainly based on file exchange through STATEL services. A STATEL Help desk assists Member States and candidate countries, from the installation time up to the day-to-day follow-up and support.

Acronyms:

CEC	Central European Countries.
CEFACT	Centre for Facilitation of Practises & Procedures for Administrations, Commerce and Transport.
CLASET	an EDI (Electronic Data Interchange) message - for the exchange of classification structures and definition, correspondence tables, code lists and lifecycle.
EBES	European Board for EDI Standardisation
FTP	File Transfer Protocol
GESMES	an EDI message - Generic Statistical Message - for the exchange of any type of multi-dimensional data or chronological series
MEDSTAT	Mediterranean Statistical Co-operation Programme
OMG	Object Management Group
TESTA	Trans-European Services for Telematics between Administrations.