

CONFERENCE OF EUROPEAN STATISTICIANS

Joint UNECE/EUROSTAT Work Session on Electronic Data Reporting
(Geneva, Switzerland, 13-15 February 2002)

Topic (iii): Metadata, conceptual models and standards

**ELECTRONIC DATA REPORTING (EDR), METADATA, STANDARDS
AND THE EUROPEAN STATISTICAL SYSTEM (ESS)¹**

Submitted by Eurostat²

Invited paper

ABSTRACT

The questions addressed in this paper include: What is EDR? What is metadata for EDR? Which metadata standards for EDR are available or coming up? And how is the ESS concerned or involved? The current situation of EDI standardisation is examined in more detail: the transition from EDIFACT to XML and the adoption of ebXML for the ESS.

Keywords: ESS, EDR, Metadata, EDIFACT, XML, ebXML

I. INTRODUCTION

1. Article 1 of the Council Regulation on Community Statistics [1] states: "The national authorities at national level and the Community authority at Community level shall be responsible for the production of Community statistics in compliance with the principle of subsidiarity. To guarantee comparability of results, Community statistics shall be produced on the basis of uniform standards and, in specific, duly justified cases, of harmonised methods."

2. The Community is the European Union (EU), national authorities are the National Statistical Institutes (NSIs) and other bodies (e.g. central banks, customs, ...) responsible in each member state for producing Community statistics. The Community authority is Eurostat, the Statistical Office of the European Communities in Luxembourg. Together, these institutions form the Community Statistical System (CSS). The term European Statistical System (ESS) is used to additionally include partner countries in the European Economic Area.

3. The principle of subsidiarity is intended to ensure that decisions are taken as closely as possible to the citizen. Specifically, it is the principle whereby the Union does not take action unless it is more effective than action taken at national, regional or local level. This means for example that statistical raw data collection is normally undertaken in the member states by the Competent National Authorities (CNAs).

4. The Statistical Programme Committee (SPC) was established in 1989 [2] to assist the European Commission in the general coordination of the multi-annual statistical programmes for the ESS [3]. The

¹ The views expressed in this document are those of the author and do not necessarily reflect policies of the European Commission.

² Prepared by Uwe Kunzler, Eurostat unit A-2 – Information and Communication Technologies for the Community Statistical System.

SPC is composed of representatives of the National Statistical Institutes (the Directors-General or Presidents) and chaired by the Director-General of Eurostat. The most important work of the SPC is collaboration to continually improve the operation of the ESS. This objective strongly relies on the success of ESS-wide IT systems.

5. The SPC has set up an IT Steering Committee (ITSC) to make recommendations on IT-related issues. In September 2001, the ITSC presented to the SPC a vision of IT at the service of the ESS [4]. This document stated: "In the ideal world, the ESS would have harmonised or closely coordinated statistical processes [...]. These would cover the full statistical life cycle from survey design and collection of data, through production to dissemination."

6. The document identified eight strategic objectives; the first objective was 'Collaboration on data and metadata standardisation and data collection.' The text explained: "Data collection is an important function of the ESS and it will become even more important with enlargement of the EU. A lot of progress has been made in automation to make this more efficient. But this is still far from complete implementation. Standardisation is an important aspect, as without standardisation automation is very much more difficult and inefficient. It will also be necessary to meet the challenges of the new technologies and update the standards we use to XML versions [...]. Standardisation activities will need to include not only technical standardisation (e.g. the development of standard XML message formats), but also the standardisation of metadata.

7. "A more active cooperation with NSIs and with the Statistics Departments of partner organisations is needed for coordinating our respective developments on metadata issues. The work should progress on two parallel activities: the content harmonisation of statistical metadata and the development of more advanced systems for exchanging and sharing the information. The ESS needs to deal with both standardisation activities ('technical' and 'content-oriented') at the same time.

8. "The goal is to define common e-standards (such as the ones based on the XML syntax) and a series of high-priority standardisation activities in the field of metadata production, exchange and dissemination, that could allow us to gain efficiency and avoid duplication of efforts. This would also allow us to reduce the reporting burden on Member States through the implementation of a common platform for producing and sharing statistical information."

9. The same document stated under the heading 'Opportunities: New technologies, Web technology': "The chief opportunity presenting itself is the explosion in the Internet and Internet related technologies (including XML, middleware, etc.). Internet technologies go a long way towards solving the problems of interfacing different platforms. According to a Gartner Group analysis, we are now entering the third phase of Internet development. The first phase enabled 'advertising' on the Internet in which the important thing was to 'have a presence' and be able to reach one's potential customers. This has been termed 'brochure ware'. Next came the ability to buy and sell, or e-commerce. The important thing was to be able to interact with one's customers. The phase we are entering is the use of the Internet to facilitate interaction with one's collaborators. Applications are appearing which can create deeper relations between partners, which give better ways of dealing with the supply chain and provide better ways of getting at knowledge and using it (whether it comes from information bases or from people). This has been termed 'collaboration ware'. Examples include desktop video-conferencing, GroupWare and knowledge management systems."

10. This paper examines how Internet technologies and standards can be utilised for EDR in the sense of 'collaboration ware'. The following chapters cover the topics 'Electronic Data Reporting', 'Metadata' and 'Standards'. This document is partly based on previously presented papers ([5], [6]).

II. EDR – ELECTRONIC DATA REPORTING

11. What is EDR? There are many definitions, partly overlapping, rarely describing exactly the same phenomenon, sometimes even widely differing (see for example [7]). I will try to narrow the field a bit.

12. Data collection in the ESS comprises two phases: raw data collection – which is mainly undertaken in the member states by national authorities – and transmission of (usually aggregated) data to Eurostat. EDR falls under the first category, raw data collection.

13. The collection of raw data includes direct (or primary) and indirect (or secondary) data collection. Direct data collection is based on surveys with data providers (or their agents). Indirect data collection makes use of existing data, for example from taxation or customs databases. EDR is a method of direct data collection.

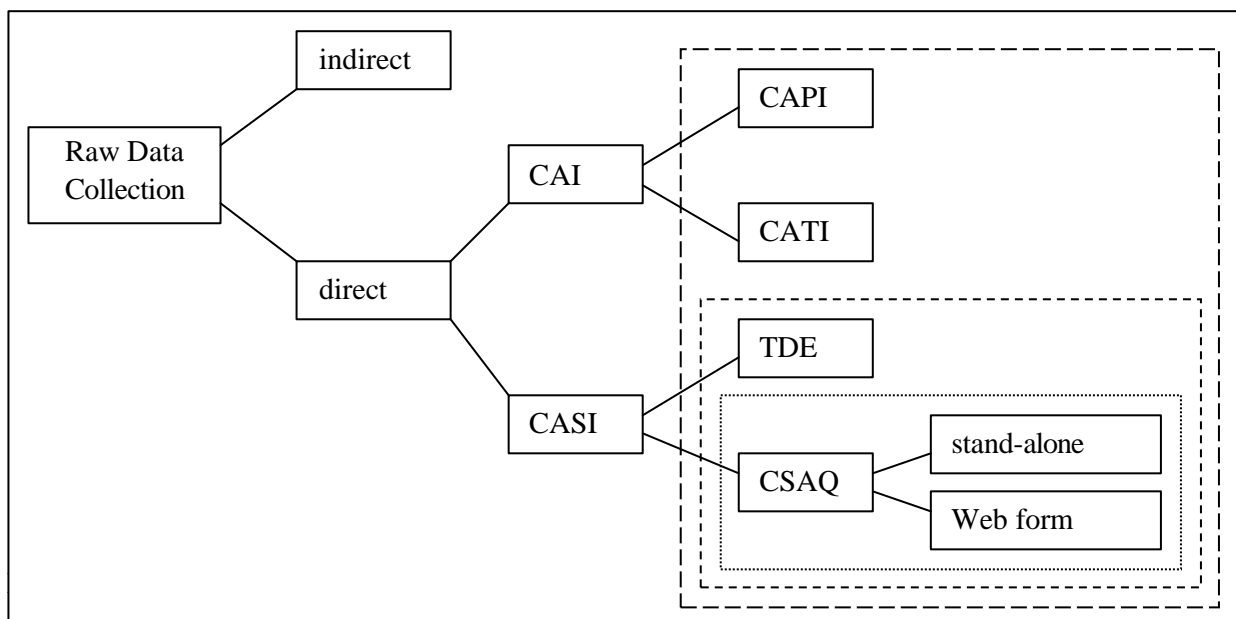


Figure 1: What is EDR?

14. Regarding direct data collection, we can distinguish between interview and self-interview techniques. Interviews may be conducted face-to-face or by telephone. Both interview methods may be computer supported, which is known as CAI – Computer Assisted Interviewing (including CAPI and CATI – Computer Assisted Personal or Telephone Interviewing).

15. Self-interviews are based on questionnaires handed out to the respondent. Traditionally, this has been a paper questionnaire sent out by mail, completed by the respondent and then returned to the data collector. Today, Information and Communication Technology (ICT) is used to automate this process as much as possible; this is known as CASI – Computer Assisted Self-Interviewing. CASI includes electronic questionnaires for self-interviewing and TDE – Touch-tone (or Telephone) Data Entry. CASI e-questionnaires are often called CSAQ – Computerised Self-Administered Questionnaire.

16. This context is illustrated in figure 1. I will use the term EDR in the sense of direct raw data collection based on electronic questionnaires for self-interviewing (CSAQ). A wider understanding of EDR would include TDE and probably also CAI.

17. There are two basic electronic questionnaire technologies: in the first case, software has to be installed on the respondent's workstation, the second solution only requires the availability of an Internet

connection and a normal Web browser. In this paper I will call the first type of system ‘stand-alone CSAQ system’, the second ‘Web form’.

18. A stand-alone CSAQ system can be used without an Internet connection, using for example magnetic media for data transmission. Such CSAQ systems were already in existence before the Internet revolution. They allow statistical data to be captured and packed in a format accepted by the data collector. Modern stand-alone CSAQ systems make more and more use of the Internet: data is sent over the Internet; program updates can be downloaded from the data collector’s Web site; long code lists may be consulted on the collector’s web server; and so on.

19. A successful example of a stand-alone CSAQ is IDEP/CN8 ([8], [9]), an electronic questionnaire for Intrastat declaration; Intrastat is the statistical system relating to the trading of goods between EU member states. IDEP/CN8 was developed and is maintained by Eurostat. It is available in most EU member states in their respective languages. The competent national administrations take care of distribution and user support. IDEP/CN8 is currently used by over 50,000 enterprises. The existence and success of IDEP/CN8 is the result of a close and fruitful co-operation between the relevant partners within the ESS.

20. While traditional stand-alone CSAQ systems like IDEP/CN8 cover only one statistical survey, multi-questionnaire CSAQ systems can be used for several surveys in parallel. New questionnaires can be added whenever needed, for example through download via the Internet. EDISENT, one of the first multi-questionnaire systems, was developed in the TELER research project (1996-1999) [10]. A more recent development is now being used in Austria (e-Quest, opened to the public in April 2001) [11]: e-Quest will be presented by Statistics Austria at this work session.

21. Web forms are electronic questionnaires that do not require software installation on the provider’s computer. A Web form is available on the data collector’s Internet server and can be accessed and completed with a usual Web browser. This kind of solution is offered more and more by statistical offices in most EU member states [12]. As examples you can take the UK Intrastat Web form (available since 1998) [13] or the German w3stat system (available since 1999) [14].

22. Web forms based on XML are also being developed in the IQML research project (2000-2003) [15] which will be presented separately in this work session.

III. METADATA

23. What is metadata? More specifically: what is metadata in our statistical context? This question was dealt with in many papers and conferences over the last 30 years. Many different definitions of statistical metadata models have been presented by various bodies and institutions without achieving international agreement. Recently six important international organisations – BIS (Bank for International Settlements), ECB (European Central Bank), Eurostat, IMF (International Monetary Fund), OECD (Organisation for Economic Co-operation and Development) and UN (United Nations) – decided to join forces and to “explore common e-standards and ongoing standardisation activities that could allow us to avoid duplication of effort in our own work and possibly for the work of others in the field of statistical information” [16]. On 6-7 September 2001, the SDMX (Statistical Data and Metadata eXchange) group met in Washington, D.C. to exchange information and to launch common work on statistical metadata. This work has just started; it will be important to ensure that the SDMX initiative also covers EDR.

24. There are different types of metadata. Metadata may help to search data (information retrieval metadata); metadata may describe the characteristics of statistical data (statistical metadata); metadata may be used by IT systems to manage, process and exchange data (IT infrastructure metadata). There are more types of metadata, and each type may be subdivided in more specific types (for more information see for example [17] and [18]).

25. But our question here is: what is metadata for EDR? In order to answer this question, I will examine another question: how does EDR work? As explained above, this paper is about electronic questionnaires, so-called CSAQ systems. There are stand-alone CSAQ systems and Web forms. Stand-alone CSAQ systems can be single-questionnaire (like IDEP/CN8) or multi-questionnaire systems (like EDISENT or e-Quest). In this paper, I will look at IT related metadata only; I will not consider statistical metadata, like for example data about survey populations, samples, methods, etc.

26. An EDR based survey starts with the questionnaire design. The electronic questionnaire that will be made available to the respondent is built up of a number of different questionnaire metadata:

- ~~///~~ metadata describing the actual questionnaire (you could name them structural questionnaire metadata): questions, answer types, screen texts, routing information, etc.;
- ~~///~~ questionnaire layout metadata: fonts, positions, colours, logos, backgrounds, etc.;
- ~~///~~ metadata supporting the respondent in completing the questionnaire; these may be:
 - help texts, error messages, warnings;
 - nomenclatures, classifications, code lists;
 - validation rules;
 - autofill rules;
 - rules to generate output (the data to be returned);
 - pre-filled data (but this may be seen as data, not metadata).

27. This list is not exhaustive, nor is it meant to be a model definition; it is just an example to show the context of our discussion.

28. In the case of a single-questionnaire stand-alone CSAQ system, the questionnaire metadata is normally built in to the CSAQ software. This software also includes the functionality to present the questionnaire on the screen. In the other cases, the respondent will view the questionnaire using questionnaire presentation software (multi-questionnaire stand-alone CSAQ) or a standard browser (Web form); in these cases, presentation software and questionnaire metadata are separated. Some of the benefits in the latter case are: only one software is needed for many different questionnaires; shorter development cycles and thus lower costs; the questionnaire files are much smaller than a complex software package i.e. distribution is easier and cheaper.

29. In the next step of the e-questionnaire's life cycle, there are three options: (a) the respondent completes it manually, or (b) files generated by specific software are imported, or (c) the questionnaire directly accesses existing databases for data retrieval. In all these cases, validation rules (metadata!) are applied automatically. Options (b) and particularly (c) require more metadata support: the import data format has to be defined, and in case (c) the IT systems need to know how to talk to each other in order to identify the wanted data.

30. Once the questionnaire is completed, data is submitted to the collector. Data submission is undertaken via HTTP, e-mail or other available file transfer options. There are two options for the data format: the complete original questionnaire description including the data filled in by the respondent is submitted, or a specific (smaller) return data file is generated and sent. In the latter case, metadata is required to specify the return data format.

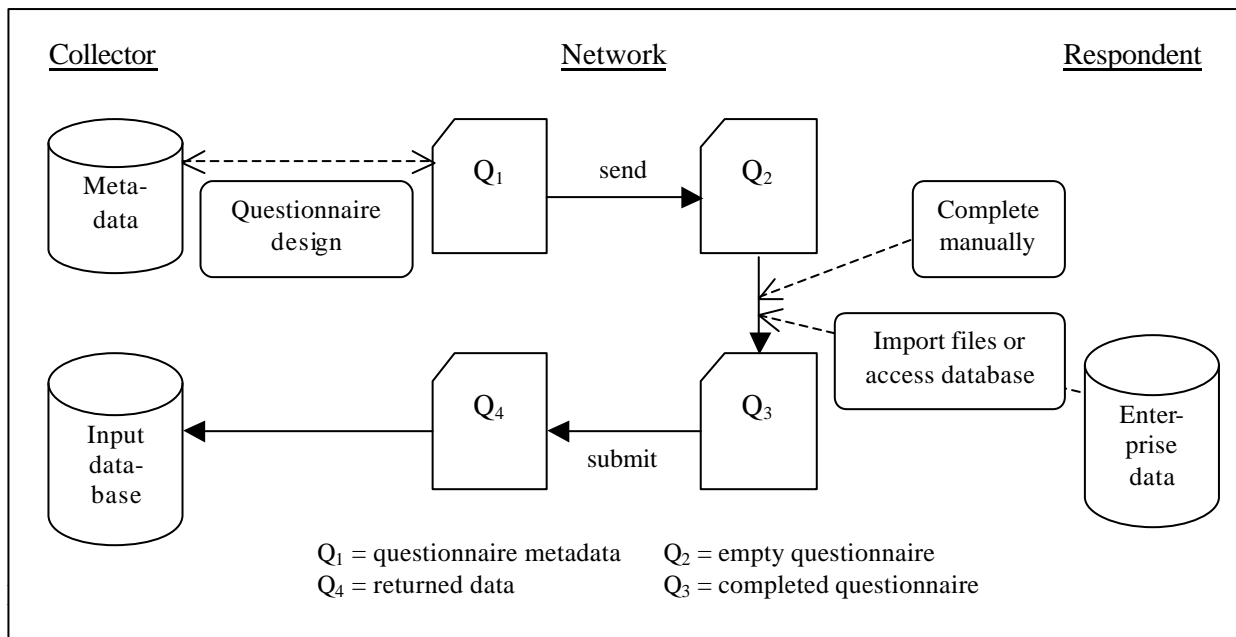


Figure 2: How does EDR work?

31. This scenario is illustrated in figure 2. It comprises different types of EDR metadata, including:
- ☞ questionnaire metadata sent to the respondent, including structural metadata, layout metadata and supporting metadata like validation rules, classifications, etc;
 - ☞ data retrieval metadata to access databases in an automated way;
 - ☞ data exchange i.e. EDI metadata.

32. EDI metadata is the best known type of metadata in the EDR context. EDIFACT and ANSI X.12 are in use for quite a while. XML is entering the scene dynamically. We will see more about standards in the next section.

IV. STANDARDS

33. I would like to start with a look at standards for EDI metadata. In general terms EDI (Electronic Data Interchange) means the exchange of information by two computer applications in a structured way without human interaction. This is mainly needed to automate business-to-business (B2B) communication, e.g. to send supply orders, bills and so on in an automated way.

34. EDI can also be applied in the statistical world. Two statistical institutions may exchange aggregated data, or a provider of raw data (e.g. an enterprise) may send data to the collector using EDI technology.

35. EDIFACT (and in America ANSI X.12) is the traditional standard for EDI. But XML is increasingly taking the stage. What does this mean for the collection of statistical data in the ESS?

36. EDIFACT was the pre-Web standard for EDI. The typical communication media were dedicated lines or Value Added Networks (VANs). The explosion of the Internet does not automatically mean the end of EDIFACT as EDIFACT messages can equally well be sent over the Internet, as e-mail attachments for example, or via file transfer.

37. EDIFACT became a UN standard in 1987. From the beginning, the ESS participated actively in the definition of EDIFACT messages. The relevant EDIFACT body is EBES, the European Board for EDI Standardisation. EBES is organised in different expert groups according to the different domains

specifying EDIFACT messages. EBES Expert Group 6 (EEG6) is the expert group for statistical purposes.

38. For specific statistical domains subsets of the relevant EDIFACT messages developed by other competent bodies were adopted. Foreign trade statistics, for example, use subsets of customs messages specified by the Customs Expert Group of EBES (EEG3). However, the statistical community, i.e. EEG6, also developed their own generic EDIFACT messages: GESMES for multi-dimensional data or chronological series, CLASET for the exchange of classifications and RDRMES for raw data reporting i.e. EDR. Each of these messages is maintained by a specific working group within EEG6.

39. EDIFACT was widely accepted in some areas. The European System of Central Banks (ESCB) adopted GESMES as their data transmission format (in fact a specific dialect of GESMES named GESMES/CB). Data exchange between Eurostat and its partners (NSIs, ECB and others) is also based on GESMES. The EDIFACT message for Intrastat reporting (CUSDEC/INSTAT) is used by over 50,000 enterprises; several million copies of this message are sent every year.

40. In other areas, acceptance of EDIFACT has been poorer. This is particularly the case when small and medium sized enterprises (SMEs) are involved. They are reluctant to introduce EDIFACT solutions for several reasons: high complexity, high costs, little flexibility, to name only the most important. These problems may be overcome by a new approach, especially tailored to the needs of the Web: XML (eXtensible Mark-up Language).

41. There are three languages that are often mixed up: SGML, HTML and XML. XML is the youngest in this row. Why was XML developed? Why wasn't SGML taken as the Web language? And why not stick to HTML?

42. SGML (Standard Generalised Mark-up Language) is a general document description language. It was conceived before the Web came up (adopted by ISO in 1986). SGML allows the definition of different document types, in other words, of different languages describing specific types of documents. In this sense SGML is a meta-language. And SGML is quite complex – too complex for the Web (making implementation expensive and applications slow).

43. HTML (Hypertext Mark-up Language) was developed as the document description language of the Web (around 1990). HTML is one specific SGML application i.e. it defines one fixed type of document. The problem with HTML is: it is not extensible, there will always be applications that cannot be based on HTML.

44. XML was conceived to overcome the problems linked to SGML (overly complex) and HTML (inflexible). XML is a subset of SGML, designed to enable the use of SGML functionality on the Web. Like SGML (and unlike HTML) it is a meta-language, allowing you to specify your own languages for your specific purposes.

45. XML 1.0 has been adopted by the World Wide Web Consortium (W3C) in 1998. XML is well suited for WebEDI i.e. for EDI over the Web. XML is supported by the major players in the software industry (Microsoft, IBM, Sun, Oracle, ...) and by international administrations (OECD, IMF, ECB, EU, ...). A further advantage of XML is: it makes business-to-consumer (B2C) communication possible; this was not the case with EDIFACT – which individual or household would have installed an EDIFACT solution at home, expensive and complex as it is (while it is not offered for free)? XML however is accessible using a normal browser. XML forms, made available by businesses on their Web site, can be used by everybody – no extra costs arise for the consumer.

46. So XML is the ideal candidate for the future of EDI. But XML is only the basis. For WebEDI you need a bit more. XML is only the syntax. Semantics are defined separately. This is where ebXML comes in.

ebXML (electronic business XML) is a joint venture of OASIS i.e. the software industry (including Microsoft, IBM, SUN, SAP, ...) and UN/CEFACT, the international business standards body responsible for EDIFACT. ebXML has defined an XML framework (core components, business processes, registry services, ...) for electronic business. On 14 May 2001, approval of the ebXML specifications was announced to the press [19]:

47. “Geneva, Switzerland and Boston, MA, USA; 14 May 2001 – UN/CEFACT and OASIS today announced that participants from around the world approved ebXML specifications at a meeting in Vienna, Austria on 11 May 2001. ebXML, which began as an 18-month initiative sponsored by UN/CEFACT and OASIS, is a modular suite of specifications that enables enterprises of any size and in any geographical location to conduct business over the Internet. Using ebXML, companies now have a standard method to exchange business messages, conduct trading relationships, communicate data in common terms and define and register business processes.”

48. ebXML is supported by EBES. In order to reflect this fact and the future focus on XML, EBES decided to change its name. The new name is e-business Board for European Standardisation (eBES). Other standardisation bodies are undergoing similar changes, also at structural level. The former EWG (EDIFACT Working Group of UN/CEFACT) will become ebWG (e-business Working Group) and will be responsible for the non-technical aspects of ebXML: core components, business processes and message development.

49. It is declared Eurostat strategy to develop and introduce ebXML compliant versions of the statistical messages used in the ESS. The basis for these developments will be UML (Unified Modelling Language) data models. Eurostat’s data transmission and dissemination tools (like STADIUM or NewCronos) will support XML. However, the existing EDIFACT solutions will be maintained as long as required.

50. ebXML is not the only XML-based semantic standard. Essentially each field of activity can have its own specific set of standards built on XML. Some of those standards could be used for statistical purposes: compliant databases could be interrogated automatically by statistical questionnaires, for example.

51. The eXtensible Business Reporting Language (XBRL) is an XML-based specification for the preparation and exchange of financial reports and data [20]. XBRL is driven by the accounting profession, in particular the American Institute of Certified Public Accountants. XBRL uses accepted financial reporting standards and practices to exchange financial statements across all software and technologies, including the Internet. In 2001, Eurostat set up a Task Force on Accounting. This will assist ESS activities in different areas, including collection and interpretation of business data from firms. In this context, the Task Force will certainly have to examine the XBRL achievements.

52. The IQML project is building upon the Common Warehouse Metamodel (CWM) [21]. CWM is the result of a request for proposal issued by OMG in 1998. Partners in the project include IBM, Unisys, Oracle and others. “The purpose of OMG’s Common Warehouse Metadata Initiative (CWMI) is to enable easy interchange of metadata between data warehousing tools and metadata repositories in distributed heterogeneous environments.”

53. The XForms project was launched by the World Wide Web Consortium (W3C) [22]. They claim that XForms is the XML language for the next generation of Web forms. “XForms is W3C’s name for a specification of Web forms that can be used with a wide variety of platforms including desktop computers, handheld devices, information appliances, and even paper.”

54. There are many more XML applications and industry initiatives (see for example [23]). While ebXML has been chosen as the semantic XML standard for the ESS, it is important not to lose sight of

the other standards. Under certain circumstances they may be the better solution, and you should at least know what they do better.

55. We have seen that there is a lot of standardisation activity in the area of EDI metadata. In respect of questionnaire metadata including questions, answer types, validation rules, classifications and so on, XForms is one standardisation initiative. There are also questionnaire metadata projects within the ESS.

56. During the questionnaire design phase, questionnaire metadata can be stored for later re-use in a metadata repository, or taken from this metadata repository to build a questionnaire. This option is being examined in the IQML project [24] (another approach is presented in [25]). The TADEQ research project has analysed a different problem: the automatic generation of electronic questionnaire documentation [26].

57. Initiatives like XBRL or CWM push in the direction of automated data retrieval based on standardised metadata. Automated data retrieval is also the basis of the ‘Semantic Web’. This expression was coined by Tim Berners-Lee, the inventor of the World Wide Web. “The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in co-operation. The first steps in weaving the Semantic Web into the structure of the existing Web are already under way.” The vehicle of the Semantic Web is metadata, expressed in XML, RDF (Resource Description Framework) and ontologies [27]. These ideas and activities may deserve more attention by the ESS in the next future.

V. CONCLUSION

58. The Web revolution is changing the way we communicate, including in the area of EDR for statistical purposes. XML is taking the lead as the document and data exchange standard for the Web. XML-based framework standards like ebXML and XBRL will have a heavy impact on the automation of data collection.

59. In parallel, metadata management is increasingly seen as a critical success factor, equally for statistical offices. Initiatives like SDMX are promising.

60. Combining Web standards and metadata harmonisation could weave a ‘Semantic Statistical Web’, covering the whole life cycle of statistical data – from collection through production to dissemination. The ESS has set the course.

REFERENCES

- [1] Council Regulation (EC) No 322/97 of 17 February 1997 on Community Statistics. OJ L 52, 22.2.1997
- [2] Council Decision (89/382/EEC, Euratom) of 19 June 1989 establishing a Committee on the Statistical Programmes of the European Communities. OJ L 181, 28.6.1989
- [3] Council Decision (EC) No 1999/126 of 22 December 1998 on the Community statistical programme 1998 to 2002. OJ L 42, 16.2.1999
- [4] CPS 2001/42/14/EN: IT co-operation strategy in the ESS. 42nd meeting of the SPC, Luxembourg, 19 and 20 September 2001
- [5] Electronic Data Reporting (EDR) Strategy for the European Statistical System (ESS). Uwe Kunzler, Eurostat. NTTS/ETK conference, Crete, Greece, June 2001
- [6] Influence of the Internet on data collection and dissemination in the European Statistical System (ESS). Wolfgang Knüppel and Uwe Kunzler, Eurostat. IAOS Satellite Meeting on Statistics for the Information Society, Tokyo, Japan, August 2001
- [7] Methodologies for Internet surveys and other telesurveys. Gad Nathan, Hebrew University, Jerusalem. NTTS/ETK conference, Crete, Greece, June 2001
- [8] IDEP/CN8 – A European Software Package. Eurostat, August 1998

- [9] Electronic Collection of Raw Data (eCoRD) – a European Perspective. Uwe Kunzler, Eurostat. EDR workshop, Hull, Canada, September 2000
- [10] TELER Final Report. Philippe Caille (Cesia), Louis-Aimé de Fouquières (Cesia) and Hans Stol (Statistics Netherlands). May 1999
- [11] e-Quest: a Metadata-based Software for Electronic Raw Data Collection at Statistics Austria. Wolfgang Koller and Günther Zettl, Statistics Austria. NTTS/ETK conference, Crete, Greece, June 2001
- [12] Web-forms for Intrastat. Eurostat, March 2000
- [13] <http://www.hmce.gov.uk>
- [14] <http://www.w3stat.de>
- [15] <http://www.epros.ed.ac.uk/iqml/>
- [16] Task Force on Statistical Data and Metadata Exchange (SDMX) – Common Statement by Participating Institutions. Luxembourg, May 2001
- [17] Sharing best methods and know-how for improving generation and use of metadata. Joanne Lamb, University of Edinburgh. NTTS/ETK conference, Crete, Greece, June 2001
- [18] Developing a common understanding of standard metadata components: a statistical glossary. Denis Ward (OECD) and Marco Pellegrino (Eurostat). Workshop on Statistical Data and Metadata Exchange (SDMX). Washington, D.C., September 2001
- [19] <http://www.ebxml.org>
- [20] <http://www.xbrl.org>
- [21] <http://www.cwmforum.org>
- [22] <http://www.w3.org/MarkUp/Forms>
- [23] <http://xml.coverpages.org>
- [24] Intelligent Use of Metadata in the Questionnaire Design Process. Karen Brannen, University of Edinburgh. NTTS/ETK conference, Crete, Greece, June 2001
- [25] Connection Bits and Pieces: Context Tokens in Survey Design. Stephen Jenkins (Mercator Computer Systems Ltd) and Tony Solomonides (University of the West of England). NTTS/ETK conference, Crete, Greece, June 2001
- [26] TADEQ, a Tool for the Analysis and Documentation of Electronic Questionnaires. Jelke Bethlehem and Anco Hundepool, Statistics Netherlands. NTTS/ETK conference, Crete, Greece, June 2001
- [27] The Semantic Web. Tim Berners-Lee, James Hendler and Ora Lassila. Scientific American, May 2001