

**UNITED NATIONS STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS**

**EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES
(EUROSTAT)**

**ORGANISATION FOR ECONOMIC
COOPERATION AND DEVELOPMENT
(OECD)
STATISTICS DIRECTORATE**

Joint UNECE/Eurostat/OECD work session on statistical metadata (METIS)
(Luxembourg, 9-11 April 2008)

OUTCOMES OF PRE-WORK SESSION SURVEY OF METIS PARTICIPANTS

Submitted by the UNECE Secretariat¹

I. INTRODUCTION

A request to complete an on-line questionnaire was sent to organizations attending METIS 2008 and past meetings. The purpose of the questionnaire was to gather information about existing metadata practices and issues. The results provide an overview of the current situation and challenges being experienced in national and international statistical offices. It provides a starting point for discussion and will help guide the METIS steering group in addressing the needs of the participants at the meeting.

This follows on from a similar survey conducted in 2007, which focused on Part C of the Common Metadata Framework (CMF), the results of which are published at <http://www.unece.org/stats/documents/ece/ces/ge.40/2007/wp.9.e.pdf>.

The questionnaire was completed by 35 national and international / intergovernmental statistical offices. They were:

National Statistical Offices

Albania	Moldova
Australia	Netherlands
Austria	New Zealand
Bulgaria	Norway
Canada	Poland
Croatia	Portugal
Cyprus	Romania
Czech Republic	Slovakia
Estonia	Slovenia
France	Spain
Germany	Sweden
Greece	Switzerland
Ireland	Turkey
Lithuania	U.S. Bureau of Labor Statistics (BLS)

International / Intergovernmental Statistical Offices

Eastern Caribbean Central Bank
Eurostat
European Central Bank (ECB)
International Monetary Fund (IMF)
Organization for Economic Cooperation and
Development (OECD)
United Nations Economic Commission for
Europe (UNECE)
United Nations Industrial Development
Organization (UNIDO)

¹ Prepared by Jessica Gardner (jessica.gardner@unece.org).

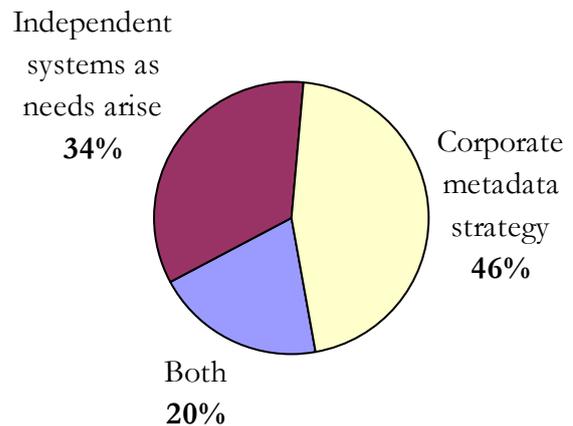
II. QUESTIONS AND RESULTS

Question 3²: Does your organization have a corporate metadata strategy and/or are they building independent systems to manage metadata as needs arise?

Respondents were asked to indicate if they used either or both approaches to manage metadata. Twenty percent of respondents have both a corporate metadata strategy and independent systems, and the remainder have specified either one or the other.

The comments made were:

- (a) Local systems still try sometimes to argue the case for local metadata solutions.
- (b) A corporate strategy requires executive staff support. This hasn't occurred yet.
- (c) Independent systems are still in existence and are being replaced by a centralised system.
- (d) We are presently developing an enterprise strategy.
- (e) Most existing metadata systems are integrated but do not cover all statistical survey life cycle.
- (f) Independent systems are still in place for continuity purposes; the corporate strategy is aimed at integrating the different modules into a system handling data and metadata across the statistical life cycle.
- (g) The independent but loosely-coupled systems are part of the strategy.
- (h) STC has a corporate metadata strategy that first focussed on providing metadata for disseminated products but this role has expanded to survey design data analysis and data archiving. For the other phases of the survey life cycle (i.e. from build to processing) there are currently independent metainformation systems in the organization.
- (i) Now Statistics Lithuania is implementing new system: Integrated statistical information system (ISIS)
- (j) We have start the work to build a METADATA system in INSTAT
- (k) Data Warehouse Project to be implemented soon.
- (l) Collection management system manages statistical surveys through collection storage and data production (besides description of new derived variables) does not cover dissemination to the end users via web services. There are metadata systems for specific needs – for production of statistics from administrative data. Dissemination is managed by own metadata systems. Collection management system and other metadata systems are not fully harmonised as concerns variables and code lists.
- (m) IS Planning & Design of Statistical Survey; IS Statistical Classifications
- (n) Developed in the context of the migration of the complete system from Mainframe to Client/Server platform
- (o) Beginning process of developing a corporate metadata strategy
- (p) A corporate metadata strategy is developing now
- (q) Under design and development
- (r) Dispositif de documentation structurée - logiciel qui permet de stocker toute forme de méta-données
- (s) There are built successive subsystems of the metadata system (depending on the current needs)



² Questions 1 and 2 relate to the contact details of each respondent.

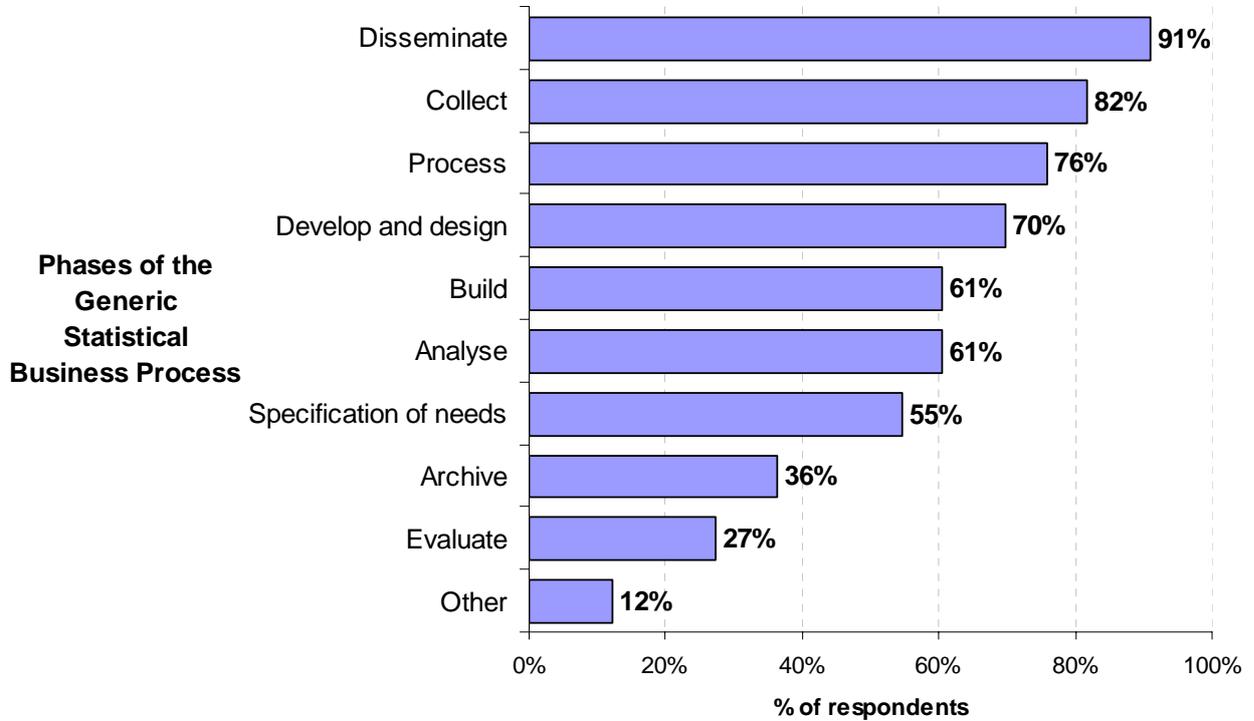
Question 4: Please identify the existing or planned metadata system(s) in your organization and indicate the current stage(s) of implementation for each.

The table below shows which stage each responding agency is at with their metadata systems. This is to help respondents identify other organizations at the same stage(s) of implementation.

Upgrading existing system	<ol style="list-style-type: none"> 1. Canada 2. France 3. Germany 4. Norway 5. Poland 	<ol style="list-style-type: none"> 6. Portugal 7. Eastern Caribbean Central Bank 8. Eurostat 9. UNECE 10. UNIDO 	
26% of respondents			
Advocating / writing business case	<ol style="list-style-type: none"> 1. Australia 2. Austria 3. Cyprus 4. Czech Republic 5. Germany 	<ol style="list-style-type: none"> 6. Moldova 7. New Zealand 8. Portugal 9. USA 10. Eastern Caribbean Central Bank 	
29% of respondents			
Approved and now planning	<ol style="list-style-type: none"> 1. Czech Republic 2. Germany 3. Greece 4. Ireland 	<ol style="list-style-type: none"> 5. Switzerland 6. IMF 7. UNIDO 	
21% of respondents			
Under development	<ol style="list-style-type: none"> 1. Albania 2. Austria 3. Canada 4. Czech Republic 5. Lithuania 	<ol style="list-style-type: none"> 6. New Zealand 7. Norway 8. Portugal 9. Romania 10. Slovenia 	<ol style="list-style-type: none"> 11. Spain 12. Turkey 13. Eurostat 14. IMF 15. UNECE 16. UNIDO
44% of respondents			
Testing	<ol style="list-style-type: none"> 1. Albania 2. Australia 3. Canada 4. Croatia 5. Czech Republic 6. Netherlands 	<ol style="list-style-type: none"> 7. New Zealand 8. Norway 9. USA 10. Eurostat 11. IMF 12. UNIDO 	
35% of respondents			
Implementation complete / maintaining system	<ol style="list-style-type: none"> 1. Albania 2. Australia 3. Austria 4. Bulgaria 5. Canada 6. Estonia 7. France 8. Germany 	<ol style="list-style-type: none"> 9. Ireland 10. Lithuania 11. New Zealand 12. Norway 13. Poland 14. Portugal 15. Slovakia 16. Slovenia 17. Spain 	<ol style="list-style-type: none"> 18. Sweden 19. Switzerland 20. European Central Bank 21. Eurostat 22. OECD 23. UNECE 24. UNIDO 25. IMF
71% of respondents			

Question 5: Which phase(s) of the generic statistical survey lifecycle model are covered by your metadata system(s)? (Check as many as apply)

Responses show that most metadata systems concentrate on the 'Disseminate', 'Collect', 'Process' and 'Develop and design' phases. The least covered phases are 'Evaluate' and 'Archive'.



Comments relating to those who chose 'other' were:

- (a) Tool for corporate planning - allocation of budget by program activities; monitoring applications and software use
- (b) None are covered as well as they should be!
- (c) Definitional metadata could be apply into phase of design and disseminate

Question 6 & 7: Please indicate your organization's use of the following standards, languages, etc. (check as many as apply).

		Currently in use	Can output data in this format	Consider-ing	No current use
Electronic Data Interchange (EDI) or UN/EDIFACT	12	Austria, Czech Republic, France, Netherlands, Norway, Portugal, Sweden, Turkey, US BLS, ECB, Eurostat, IMF	3	0	13
Statistical Data and Metadata Exchange (SDMX) ISO 17369	8	Lithuania, Norway, Turkey, US BLS, ECB, Eurostat, IMF, OECD	7	16	5
Neuchatel classification model	7	Canada, Croatia, Netherlands, Norway, Portugal, Sweden, Switzerland	5	9	12
Dublin Core	6	Australia, Canada, Croatia, Czech Republic, Norway, IMF	3	5	14
Metadata Registry Standard ISO/IEC 11179	6	Australia, Canada, Norway, Portugal, Sweden, IMF	3	7	15
Extensible Business Reporting Language (XBRL)	3	Norway, Netherlands, France	0	12	12
Data Documentation Initiative (DDI)	4	Canada, Croatia, New Zealand, US BLS	2	7	14
Neuchatel variable model	4	Netherlands, Norway, Sweden, Switzerland	1	10	14
GIS Metadata Standard ISO 19115	3	Canada, Portugal, Eurostat	2	7	14
Common Warehouse Metamodel (CWM)	1	Canada	1	5	20
Metadata encoding and transmission standard (METS)	1	US BLS	0	3	21
Research Description Framework (RDF) or Web Ontology Language (OWL)	0	N/A	0	4	22

The comments and other standards mentioned were:

- (a) SDDS-IMF (4 responses mentioning this)
- (b) METANET Reference Model - partly contains indicated standards (2 responses mentioning this)
- (c) XML (CLASET) GEMMES-TS
- (d) ABS broadly uses the OECD "Quality Dimensions" (with local subdivision of one dimension and relabelling of another) for providing metadata relevant to assessing the quality (fitness for purpose) of metadata we disseminate.
- (e) Various in-house xml-standards for data and metadata transmission as part of production systems
- (f) SDMX Cross-Domain Concepts
- (g) FGDC - Federal Geographic Data Committee (sw ArcGis)
- (h) SCB Metamodel v. 2.1
- (i) ISO/IEC 11179 and Neuchâtel variable model are only partly in use

Question 8 & 9: What challenges does your organization face in implementing metadata systems and frameworks?

Respondents were asked to select from a list of ten common challenges, with the option of specifying other challenges, and then rank their selection from most to least problematic.

% of respondents who face this challenge	Change compared to 2007 survey ³	
77%	Not asked in 2007	Lack of resources
74%	↓ 5%	Getting subject-matter statisticians support/buy-in
69%	↓ 15%	Implementing standards
69%	↓ 10%	Employing and retaining people with the right IT skills
66%	Not asked in 2007	Identifying standard metadata requirements across different areas of the organization
63%	↓ 11%	Adopting or developing an end-to-end process model
57%	Not asked in 2007	Insufficient integration of standard metadata system with data production systems
57%	↓ 17%	Metadata quality
49%	↓ 19%	Getting top management support/buy-in
49%	Not asked in 2007	Difficulties embedding the end-to-end process within actual workflows

Three respondents (9%) specified other challenges they face:

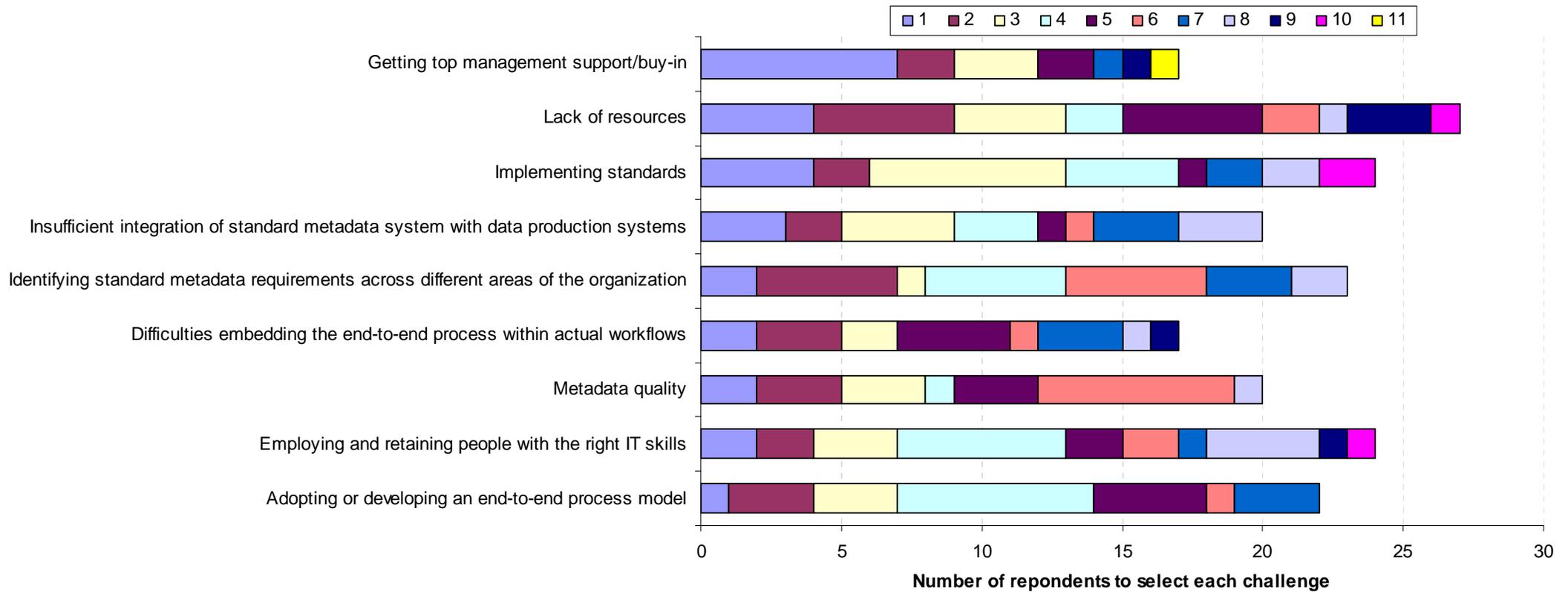
- Identifying a global strategy
- Multi-language support (English Spanish and French)
- Standard metadata requirements across different areas are one thing but there are cases of "legitimate" specialised requirements as well. The trick is taking very heterogeneous current practices in each area and working out how to move the "standard" part over to standard systems how to address "unnecessary" difference and how to support "legitimate" difference while marrying up those specialised requirements with use of the "standard" systems and processes to the greatest extent possible.

³ This question was asked in the survey conducted in 2007 (refer to page 1) but with fewer common challenges to select from. Additional options and a larger respondent sample in 2008 may account for some of the differences between the results.

(Question 8 & 9 continued) What challenges does your organization face in implementing metadata systems and frameworks?

The graph below aims to illustrate how respondents ranked each of the challenges selected, that is which of them they consider to be most problematic.

How respondents ranked each challenge
 (from most problematic (1) to least problematic (11))



Question 10: Please list the categories in use in your organization to differentiate metadata into types (e.g., structural, definitional, methodological, etc).

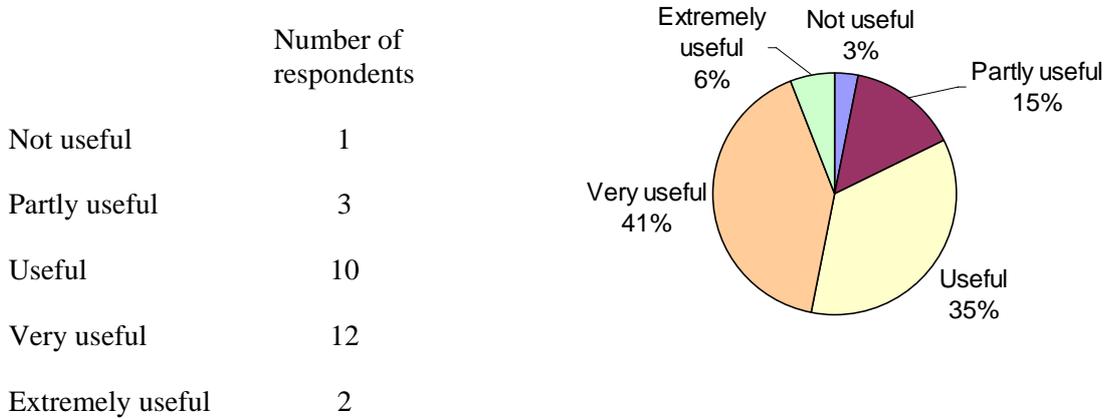
Respondents were given a free text field to describe the categories used to differentiate between types of metadata. The responses indicate a variety of terms are used, with the most common categories being 'structural', 'definitional', 'methodological' and 'quality'.

Structural	13
Definitional	11
Methodological	9
Quality	6
Referential	5
System	4
Technical	4
Administrative	4
Classifications	3
Process	3
Conceptual	3
Operational	3
Survey	2
Procedural	2
Validation	1
Descriptive	1
Variables	1
Exploratory	1
Discovery	1
Physical	1
Semantic	1
Contextual	1

Comments included:

- (a) We think of changing to the six categories mentioned in the emerging METIS Corporate Metadata Framework (CMF) once they have been formulated in a more detailed way.
- (b) Currently we are not using a classification for organizing metadata; and await the proposal of the CMF for direction.
- (c) Currently we do not have any such grouping nor have we needed it ...
- (d) We do not have such a classification. In practice we do of course use different types of metadata but the experience is that this type of distinction tends to cause confusion.

Question 11: How useful is the Common Metadata Framework for your organization?



Question 12 related to presentation of demonstrations at the METIS meeting and the results are not relevant to this report.

Question 13: Finally, are there any other comments or suggestions about METIS, the Common Metadata Framework (CMF) or related issues?

Three respondents gave additional comments or suggestions:

- (a) METIS generally and CMF are seen as very valuable collaborative and information sharing initiatives in regard to metadata management between National (and International) Statistical Offices at the conceptual strategic and practical level. The initiatives are yet to achieve their full potential but continued and expanded participation by members will allow that.
- (b) 1. In question 4 it might be helpful to have more response categories between "Testing" of a metadata system and "Implementation complete". When is the implementation of a metadata system seen as complete? When it is 100% filled? 50%? 90%? Or is it sufficient to have a metadata system but no metadata in it? 2. METIS and the research papers provided by the participating institutions have helped us a great deal in working ourselves into the matter.
- (c) There exist a number of standards; they are more general. That is why such works as Neuchâtel terminologies which could be mapped to ISO standards are useful. They are more familiar to statisticians and their needs.