Metadata Exchange in the European Statistical System: Status and Future

January 27, 2021

Luca Gramaglia
Team Leader – Data and metadata exchange standards
Eurostat – European Commission
Metadata and Quality Reporting in the European Statistical System

User-oriented and producer-oriented reports in the European Statistical System:

- **User-oriented**: allows users to correctly understand and interpret the data released
  
  Basis: European Statistical Code of Practice, Principle 15: "European Statistics are available and accessible with supporting metadata and guidance".

- **Producer-oriented**: monitor the quality of the statistics produced
  
  Basis: Regulation 223/2009, Article 12: "Member States shall provide the Commission (Eurostat) with reports on the quality of the data transmitted"
Standards for Metadata and Quality reporting in the ESS

- ESMS structure
- ESS Handbook for quality reports

2009

- ESQRS
- ESS Quality and Performance Indicators

2010

- SIMS, ESMS, ESQRS v.2.0

2015

Eurostat starts implementing ESMS

SIMS approved as the ESS standard for metadata and quality reporting
Standards for Metadata and Quality reporting in the ESS

Objective: Once for all purposes reporting using SIMS

SIMS: set of concepts relevant for metadata/quality reporting

ESMS: Subset for user-oriented report

ESQRS: Subset for producer-oriented reports

Statistical Data and Metadata eXchange
~100% of disseminated Eurostat datasets accompanied by harmonised metadata at European level

~85% of datasets collected from Member States accompanied by harmonised national metadata
SDMX framework for quality and metadata reporting

Reference metadata
- SDMX Cross-Domain Concepts (Glossary)
- Non-SDMX SIMS Concepts
- Domain-specific concepts (if needed)

Metadata Structure Definitions
- Standard SIMS/ESMS/ESQRS
- Extended SIMS/ESMS/ESQRS

Statistical Data and Metadata eXchange
SDMX framework for quality and metadata reporting

Euro-SDMX Registry

ESS Metadata Handler

Users

Production Systems
"Once for all purposes" reference metadata
The dream

SIMS -> Global MSDs

SIMS

-> eurostat

-> User report

-> Producer report
Future work: Exchange of footnotes

- “Footnotes” – textual information that refers to specific subsets of a data cube. Usually needs to travel along with the data and can be attached to different subsets depending on the situation.

- For example, “footnotes” could be attached to a whole country or to a whole time period to indicate specific methodological notes or deviations from standard methodology. Such notes could be useful in the context of data validation.

- SDMX 3.0 will offer more features to help solve this problem (e.g. attachment of attributes to partial keys).
Eurostat's experience shows that Eurostat domain managers often require detailed information in a tabular format (e.g. coefficient of variation by region / variable)

Future work:
Integration of tabular data

<table>
<thead>
<tr>
<th>NUTS-2</th>
<th>Regional Code</th>
<th>Region</th>
<th>Number of employed</th>
<th>Number of part-time employed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Regional Code</td>
<td>Region</td>
<td>Number of employed</td>
<td>Number of part-time employed</td>
</tr>
<tr>
<td>CZ01</td>
<td>Praha</td>
<td></td>
<td>1.21</td>
<td>5.88</td>
</tr>
<tr>
<td>CZ02</td>
<td>Střední Čechy</td>
<td></td>
<td>1.00</td>
<td>6.39</td>
</tr>
<tr>
<td>CZ03</td>
<td>Jihozápad</td>
<td></td>
<td>0.97</td>
<td>5.97</td>
</tr>
<tr>
<td>CZ04</td>
<td>Severozápad</td>
<td></td>
<td>1.27</td>
<td>7.70</td>
</tr>
<tr>
<td>CZ05</td>
<td>Severovýchod</td>
<td></td>
<td>0.98</td>
<td>5.25</td>
</tr>
<tr>
<td>CZ06</td>
<td>Jihozápad</td>
<td></td>
<td>0.97</td>
<td>4.77</td>
</tr>
<tr>
<td>CZ07</td>
<td>Střední Morava</td>
<td></td>
<td>1.20</td>
<td>6.50</td>
</tr>
<tr>
<td>CZ08</td>
<td>Moravskoslezko</td>
<td></td>
<td>1.20</td>
<td>6.58</td>
</tr>
</tbody>
</table>
Future work:
An SDMX solution to the challenge

```
<table>
<thead>
<tr>
<th>MSD</th>
<th>Concepts</th>
<th>Dataflow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Grouped by</td>
<td>Annotation</td>
</tr>
<tr>
<td></td>
<td>Represented by</td>
<td></td>
</tr>
</tbody>
</table>

MSD, Concepts, Dataflow, SDMX dataset
```
Discussion [Q&A]

For follow-up questions: luca.gramaglia@ec.europa.eu