UNECE High-level Group for the Modernisation of Official Statistics

Business Case for Metadata Glossary

This business case was prepared by Supporting Standards Group and is submitted to the HLG-MOS for their approval.

Type of Activity			
☐	New project	☐	New activity
☐	Extension of existing project	☒	Extension of existing activity

Purpose

The metadata glossary is intended to be a central and definitive source for the terms and definitions contained in the various statistical metadata standards managed under HLG-MOS (GSIM, GSBPM, GAMSO, and CSPA). The glossary will also serve as the source of disambiguation between the same spellings with (slightly) different meanings contained across the standards (homographs), and will serve as the source of identifying terms of different spellings with similar or the same meanings across the standards (synonyms). Based on work in 2016, 2017, and 2019, the metadata glossary now includes terms for GAMSO, GSBPM phases and part of the GSIM information objects. The purposes of this activity are to complete the glossary work for the rest of GSIM information objects and provide a unified glossary for GSIM, GSBPM, and GAMSO. Terms from CSPA will be added if needed.

Description of the activity

To conduct the work, group participants bring their knowledge of statistical practice and the expertise around the particular standards under consideration. This deep understanding allows group members to analyse definitions, identify nuances between similar ideas, and arrive at consensus wording (among the group) that captures and conveys the intended meaning. When the technical meaning for a term is questioned, the group creates an issue for the respective standards group to consider.

The output of the Metadata Glossary Group will be an alphabetically organized list of terms, their definitions, sources for the definitions, source of the term (in which standard(s) does it appear), explanatory text that further contextualizes the meaning, and references to other similar, broader, and narrower terms. The basic organizational (distinguished from presentational) principles applied are those for a thesaurus.

The task team has completed review of GSIM Business Group and Base Group and plans to complete the rest of the groups by early 2020. For the GSIM terms the task team thinks unclear or need improvement, the team is documenting issues and proposals for the next GSIM revision.

Alternatives considered

An alternative is not to complete the work. The lack of a definitive resource for terms and definitions in the standards will result. Another alternative is to continue the glossary work as part of the effort to build a Core Ontology. Here, the conflation of separate activities is a detriment to both. Glossary work is fundamentally about obtaining and refining definitions of terms. Ontology work is about inter-relating the concepts defined in the Glossary and creating a specially tagged file of the terms, definitions, and relationships that allows for a semantically rich query capability. A decision was made to continue both groups (glossary and core ontology) separately. This appears optimal.

How does it relate to the HLG-MOS vision and other activities under the HLG-MOS?

The metadata glossary will provide a more accurate and richer understanding of the terms and concepts in each of the HLG-MOS standards – GSIM, GSBPM, GAMSO, and CSPA. Time spent understanding how the standards relate to each other is greatly reduced. The ability to build more pervasive metadata driven systems within and across NSOs is enhanced, in large part due to the increased ability to make the standards

work together. As input to the Core Ontology work, the glossary holds a foundational position within the family of HLG-MOS standards.

Proposed start and end dates	
Start: January, 2020	**End:** April, 2020