

SDMX CONTENT-ORIENTED GUIDELINES

DRAFT - February 2008

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45 © SDMX 2008

46 <http://www.sdmx.org/>

47

CONTENTS

48	1. INTRODUCTION	4
49	1.1. Background	4
50	1.2. Cross-domain and domain-specific concepts	4
51	1.3. Scope of the Content-Oriented Guidelines	5
52	2. CROSS-DOMAIN CONCEPTS.....	6
53	2.1. Introduction	6
54	2.2. Cross-Domain Concepts in data/metadata exchange	7
55	2.3. Describing Cross-Domain Concepts	9
56	2.4. Cross-Domain Concepts and code lists	10
57	2.5. Representation of Cross-Domain Concepts/code lists	10
58	3. STATISTICAL SUBJECT-MATTER DOMAINS	10
59	3.1. Introduction	10
60	3.2. Classification of International Statistical Activities	11
61	3.3. Using SDMX Statistical Subject-Matter Domains.....	11
62	3.4. Representation of Statistical Subject-Matter Domains.....	12
63	4. METADATA COMMON VOCABULARY	12
64	4.1. Introduction	12
65	4.2. Structure of the MCV	13
66	4.3. Representation of Metadata Common Vocabulary	14
67	5. GOVERNANCE, MAINTENANCE AND OUTREACH	14
68	5.1. SDMX Cross-domain developments	14
69	5.2. Domain-specific developments	14
70	5.3. Outreach	15
71	6. REFERENCES	15
72	7. ANNEXES.....	16

73

1. INTRODUCTION

1.1. *Background*

The Statistical Data and Metadata Exchange (SDMX) initiative (<http://www.sdmx.org>) sets technical standards and content-oriented guidelines to facilitate the exchange of statistical data and metadata using modern information technology, with an emphasis on aggregated data.

Version 1.0 specification of the technical standards has been approved by the International Organization for Standardization (ISO) as a Technical Specification (ISO/TS 17369: 2005 SDMX).

The Version 2.0 specification (November 2005) broadens the framework to support wider coverage of metadata exchange as well as a more fully articulated architecture for data and metadata exchange. Steps are also being taken to bring this work forward within the context of ISO, assuring that SDMX technical standards build on other recognized standards and providing the basis for interoperability with them.

Unlike the technical specifications, the development of the SDMX content-oriented guidelines outlined in this document is being undertaken outside the ISO framework. This should facilitate steps by the SDMX sponsoring institutions to broadly involve content-oriented efforts of a wide range of experts that are already working in many subject-matter domains within the global statistical community.

1.2. *Cross-domain and domain-specific concepts*

The SDMX Content-Oriented Guidelines contained in this document recommend practices for creating interoperable data and metadata sets using the SDMX technical standards. They are envisaged to be applicable generically across statistical subject-matter domains. The Guidelines focus on the harmonization of specific concepts and terminology that are common to a large number of statistical domains. Such harmonisation is necessary to accommodate the efficient exchange of comparable metadata and builds on the experience gained in implementations to date.

In addition to proposing cross-domain content-oriented guidelines, the SDMX initiative also provides a structure for the development of domain-specific content-oriented guidelines. Within that framework, recognized international agencies and bodies involved in setting standards for particular statistical domains will play an important role in developing domain-specific content guidelines and related terminologies.

While the SDMX Technical Standards and the SDMX Content-Oriented Guidelines can be used independently of each other, it is especially conducive to standardisation when they are used together. The evolving work associated with concept harmonization can be supported by a known technical framework, with exchange processes already taking place on a more efficient basis.

When SDMX Technical Standards are used, exchange of data structure definitions and data messages provide a means to take advantage of commonly understood structural metadata among statistical institutions, allowing mapping or translation from and to their own internal statistical data bases and systems. Similarly, when metadata structure

definitions and metadata messages are exchanged, this involves using commonly understood reference metadata that can map from and to internal representations.¹

What is most important is that there be sufficient granularity and terminological consistency in the set of cross-domain concepts to allow for mapping to internal data and metadata structures at institutions. This permits institution-specific definitions to be maintained while using a common "transport" structure for data and metadata exchange between institutions, whether bilaterally (e.g. point-to-point transmissions) or multilaterally (e.g. web disseminations).

To advance on these common structures, SDMX sponsoring institutions are collectively identifying cross-domain concepts that are commonly used in SDMX messages. Domain experts in existing groups and institutions are working to provide common structures for those parts of the messages that are domain-specific (e.g. national accounts, balance of payments, labour statistics, education statistics, millennium development goals indicators, external debt statistics). In addition, SDMX sponsoring institutions will facilitate information-sharing about these domain-specific developments through the SDMX website (e.g. announcement of new activities, maintenance activities, opportunities to provide expert input and links to domain-specific websites).

While all of these common efforts are evolving, the SDMX technical framework can already support bilateral and multilateral exchanges because the structures used in SDMX conformant messages need to contain a clear indication of the data and metadata being transported - whether they come from an institution's own structure or from an agreed set of mappings.

1.3. Scope of the Content-Oriented Guidelines

This is the first version of the SDMX Content-Oriented Guidelines, which comprise Cross-Domain Concepts and related code lists, a list of Statistical Subject-Matter Domains and a Metadata Common Vocabulary.

The Content-Oriented Guidelines are designed to work within the framework of SDMX Technical Standards to produce maximum interoperability in the exchange of data and metadata. The intent is to encourage their use where possible across statistical domains in the following three areas:

- 1) Statistical concepts, as described in the "Cross-Domain Concepts" (CDC) guideline.
This guideline contains a list of statistical concepts, inter alia related to statistical processes and data quality. This list is based on the concepts used by the contributing international organizations. It is expected to grow in the future.
- 2) Classification of domains, as described in the "Statistical Subject-Matter Domains" (SMD) guideline.
This classification is based on the work of the United Nations Economic Commission for Europe (UNECE) to produce a high-level classification of statistical areas. It provides a starting point for organising the exchange of statistical data and metadata, for instance using a registry which provides information needed for locating data and metadata over the Internet.
- 3) Statistical metadata terminology, as described in the "Metadata Common Vocabulary" (MCV) guideline.

¹ This means that the standards and the guidelines, if adhered to, should make it possible to interlink statistical information systems of organizations and share or exchange data and metadata, in spite of technological or linguistic differences that might exist between them from their internal perspectives.

The terms presented in the MCV are in many cases taken from other sources, or they consist of harmonised terms used in the SDMX Technical Specifications. The MCV comprises a standard terminology related to statistical metadata across statistical domains. The MCV, like the other content guidelines, is seen as a living document which will continue to grow and will regularly be updated over time.

Statistical domains cover a very broad field of activities. Thus, there may always be some domain-specific elements in each of the three areas covered by the SDMX Content-Oriented Guidelines. It is not the intent of these guidelines to harmonize everything across all statistical domains. The guidelines provide harmonization where possible: for the instances where various domains use slightly different concepts, or classifications, or terms, the Content-Oriented Guidelines intend to provide a single, harmonized concept, classification, or term to use when exchanging data and metadata across domain boundaries.

It is important to understand what is meant by the term "cross-domain", as this appears in several places within these guidelines. By "cross-domain", the guidelines indicate that a statistical concept is used in different statistical domains in a materially similar form. This distinction is important, because of the process by which these guidelines are intended to be created and maintained. Identifying all the concepts, classifications, and terms which are potentially "cross-domain" according to this definition is a continuous task. As statistical domains change and expand, new terms and concepts and classifications may come into existence and need to be added.

2. CROSS-DOMAIN CONCEPTS

2.1. Introduction

Cross-domain concepts in the SDMX framework describe metadata concepts relevant to many statistical domains. SDMX recommends use of the concepts outlined below whenever feasible to promote re-usability and exchange of statistical information and their related metadata between national and international organizations. Whenever used, these concepts should conform to the specified names, roles, and representations defined in the SDMX Content-Oriented Guidelines.

In SDMX, the term "metadata" is very broad and a distinction is made between "structural" metadata that define the structure of statistical data and metadata, and "reference" metadata describing the actual contents (for instance, concepts and methodologies used), the data quality (e.g. accuracy and timeliness) and the production and dissemination process (e.g. contact points, release policy, dissemination formats). Reference metadata refer to specific statistical data, to entire data collections or even to the institution that provides the data.

The cross-domain concepts outlined below are used in:

- *Data structure definitions (key families)*, which define the valid content of data sets in a given domain in terms of the concepts used to define the data sets, the role and the valid content of each of the concepts when used in a data set.
- *Metadata structure definitions*, which define the valid content of metadata sets in a given domain in terms of the concepts contained in the metadata sets, the role and the valid content of each of the concepts when used in a metadata set.

The SDMX initiative expects the list of cross-domain concepts provided in Annex 1 to grow and to be regularly updated as SDMX Technical Standards and Content-Oriented Guidelines are utilized in more and more statistical domains. The Cross-Domain Concepts guideline includes not only the names of concepts and their content

description but also, where appropriate, their representation with supporting code lists and the roles they can play within the data structure definition and/or metadata structure definition.

The use of the SDMX Cross-Domain Concepts is not a prerequisite for technical conformance, but provides a framework to facilitate data and metadata sharing among those who are conformant with the Technical Standards. This promotes, in particular, the exchange of consistent metadata that can be used by different international organizations and national and regional data-producing agencies to compare concepts and practices.

2.2. Cross-Domain Concepts in data/metadata exchange

As mentioned above, cross-domain concepts are used in SDMX exchange structures: the data structure definition (for data exchange) and the metadata structure definition (for reference metadata exchange).

A **Data Structure Definition** defines the information structure agreed within a specific statistical domain, thus allowing a full and complete description of a data set when the actual values are given. A limited number of specific concepts are needed for data structure definitions to function properly.

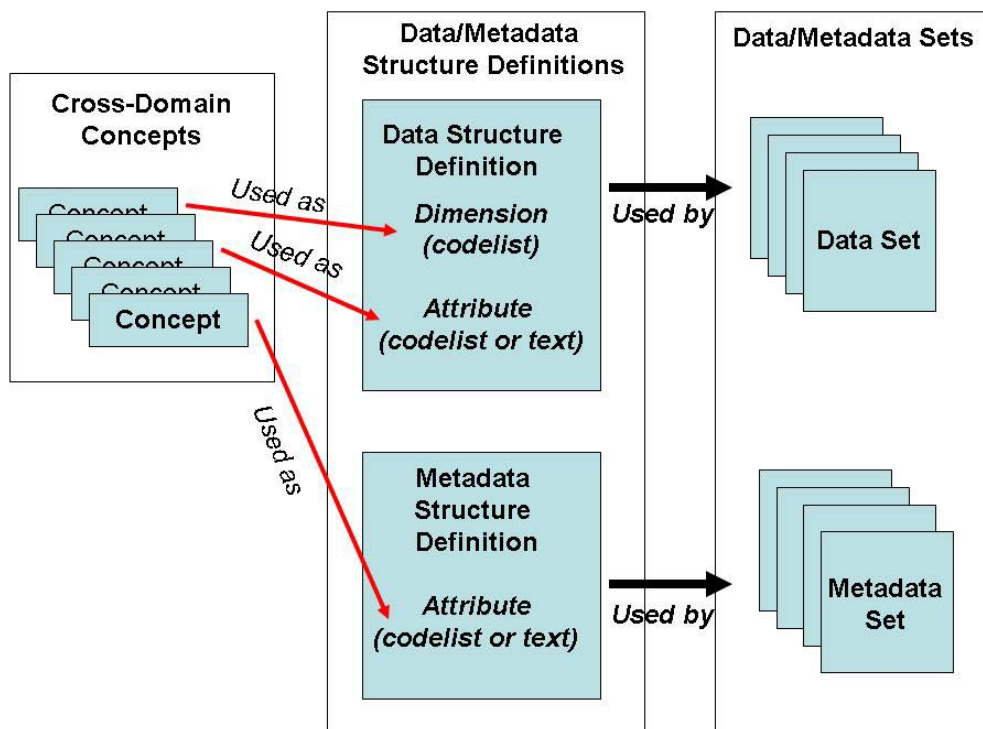
A **Metadata Structure Definition** describes how metadata sets, containing reference metadata, are organized. In particular, they define which reference metadata are being compiled, how these concepts relate to each other, how they are represented (either as free text or coded values) and with which object types (agencies, data flows, data providers, subsets of data flows, or others) they are associated.

An organization providing information about statistical data uses a set of metadata concepts (e.g. frequency of dissemination, reference area, timeliness, type of source data) in order to present the characteristics and quality of the data. These metadata concepts may be proprietary to the data provider, but interoperability will clearly be enhanced when the same concepts can be used by many exchange partners and across statistical domains. SDMX therefore recommends the use of a set of common concepts.

The cross-domain concepts presented in Annex 1 have been selected taking into account the concepts required by international organizations when collecting data from national organizations. The selection and definition of the concepts has been made by the sponsoring organizations with the assistance of some national statistical institutions from the perspectives of the metadata frameworks of these institutions.

While the cross-domain concepts include a number of broad reference metadata concepts, organizations may choose to provide information containing more detailed concepts when exchanging reference metadata. Although, in some cases, the list of cross-domain concepts presents some level of detail, it is intended that the high-level concepts, when implemented in metadata structure definitions, would support linkages to metadata available at a more detailed level. As a result, the ability to equate reference metadata sets from different domains or institutions will be enhanced, even if more detailed institution-specific metadata concepts are not always present, or are organized differently in different domains or by different institutions.

248 The following illustration provides a simplified view of how concepts are used for the
249 purposes of data and metadata exchange in the SDMX framework.



250

251 The illustration shows that cross-domain concepts have three basic roles:

- 252 1) As *Dimensions* in a data structure definition. Dimensions, when taken together,
253 serve to **identify** each statistical observation. For example, a dimension named
254 "Reference Area" would explain which country a specific statistical observation
255 refers to (e.g. United States, New Zealand, or Italy). Dimension values are typically
256 taken from code lists (in this example, a code list of countries).
- 257 2) As *Attributes* in a data structure definition. Attributes provide information about the
258 data, thus **qualifying** the data further. For example, an attribute named "Unit of
259 Measure" might provide information about whether statistical data are measured, for
260 example, in currency units, and if so, which currency, or as a pure number. Attributes
261 are sometimes coded, but can also have free-text values. Attributes are
262 basically reference metadata that are linked directly to the data. (Note that the term
263 "attribute" as used here should not be confused with XML attributes, which are part
264 of the XML syntax used in SDMX-ML).
- 265 3) As *Attributes* in a metadata structure definition. Termed "reference metadata" in
266 SDMX, these concepts can be used to report metadata which are directly
267 associated with specific data sets or data points as well as metadata which are not
268 directly associated. This could be metadata about a data flow (which may embrace
269 a set of data sets), for example, with concepts such as timeliness, reference period,
270 classification system and data compilation. The values of these concepts may be
271 coded, but are often free text.

272 Each data set or metadata set uses a structure definition of the appropriate type, so that
273 systems which exchange data and metadata can understand what the data or metadata
274 set means.

275 Concepts may be reused in various structure definitions. Thus, a concept such as
276 "Reference Area" might be used in a data structure definition and in a metadata

structure definition. The existence of a single concept with a single representation for the purposes of exchange across several domains is crucial. Data and metadata which re-use this single concept can be easily associated, because they are referring to the same idea in the same terms.

2.3. Describing Cross-Domain Concepts

The concepts used in *data structure definitions* must have a specific set of properties according to the SDMX Information Model². These properties include:

- 1) The identifier of the concept must be unique within the list of cross-domain concepts.
- 2) There must be a description of the concept. Definitions of and comments to general SDMX cross-domain concepts for data structure definitions are included in the Metadata Common Vocabulary (MCV).
- 3) There must be an indication whether the concept is used as an attribute or as a dimension in the respective data structure definition.
- 4) If the concept is “coded”, there must be a link to a code list containing valid values that may be reported for the concept.
- 5) If the concept is used as an attribute, the attachment level must be indicated. This means an indication of the data object or structure, e.g. “time series” or “observation”, to which the concept is linked.
- 6) The Maintenance Agency maintains the concept for use within data and metadata structure definitions. For the SDMX cross-domain concepts, the maintenance agency is SDMX and they will be labelled “SDMX”. For other domain-specific concepts (not or not yet included in the SDMX cross-domain concepts) there are various different maintenance agencies.

The “Maintenance agency” concept will be important for many different purposes; therefore, it is a cross-domain concept in its own right and should be coded in a uniform way.

The underlying approach to the metadata structure definition is similar to that of the data structure definition. The concepts used in *metadata structure definitions* must have a specific set of properties according to the SDMX Information Model. These properties include:

- 1) The identifier of the concept must be unique within the list of cross-domain concepts.
- 2) There must be a description of the concept. Definitions and descriptions of general SDMX cross-domain concepts for metadata structure definitions are included in the Metadata Common Vocabulary (MCV).
- 3) In case a concept can be represented as coded, there must be a link to a code list containing valid values that may be reported.
- 4) The identity of the “key” that defines to what object or structure the concept is to be linked. The term “attached” is often used in SDMX. The “key” can refer to an agency, a data set or a dataflow, or part of a data set or dataflow defined by dimensions. The structure definition will identify the components comprising the

² For a detailed description of the SDMX Information Model, data and metadata structures, see SDMX Implementors Guide (version 2.0), chapter 3.1.

identifier or “key” of the object (this is similar in concept to the “attachment level” of attributes in the data structure definition).

- 5) The Maintenance Agency maintains the concept for use within data and metadata structure definitions. For the SDMX cross-domain concepts, the maintenance agency is SDMX and they will be labelled “SDMX”. For other domain-specific concepts (not or not yet included in the SDMX cross-domain concepts) there are various different maintenance agencies.

2.4. Cross-Domain Concepts and code lists

Annex 1 contains the set of SDMX Cross-Domain Concepts with descriptions of each term, comments and links to the Metadata Common Vocabulary.

Annex 2 contains some examples of Cross-Domain code lists that can be used to support Cross-Domain Concepts. A selected number of code lists is presented at this stage which are to be extended in the future.

All concepts used should conform to the specified name, role, and representation. Thus, if a domain uses a materially similar concept, the one presented here should be used wherever it is applicable within the domain in question.

2.5. Representation of Cross-Domain Concepts/code lists

The list of SDMX Cross-Domain Concepts (Annex 1) and related code lists (Annex 2) are available on the SDMX website as:

- PDF
- SDMX-ML

In addition, the SDMX website provides examples of how these concepts and code lists can be used.

In particular, a mapping of the SDMX Cross-Domain Concepts to the metadata frameworks of selected international organizations will be provided.

3. STATISTICAL SUBJECT-MATTER DOMAINS

3.1. Introduction

A statistical subject-matter domain refers to a statistical activity that has common characteristics with respect to variables, concepts and methodologies for data collection and the whole statistical data compilation process. Examples of statistical domains are price statistics, national accounts, environment statistics or education statistics. The SDMX statistical subject-matter domain list is intended to cover the universe of statistical information handled by a large number of international organizations and national agencies, often referred to as official statistics³. Official statistics constitute the basic information system of a society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation.

In some cases, statistical information may be linked to several Subject-Matter Domains. This should be highlighted when the list of domains are used as a navigation aid.

³ UN Fundamental Principles of Official Statistics: <http://unstats.un.org/unsd/goodprac/bpabout.asp>

In the SDMX Content-Oriented Guidelines, the list of Statistical Subject-Matter Domains has three functions:

- 1) as a standard scheme against which similar domain lists of national and international organizations can be mapped to facilitate the exchange of data and metadata;
- 2) as an identifier framework for registering and searching statistical data on SDMX registries, the architecture of which has been developed in SDMX Technical Standards Version 2.0; and
- 3) as a navigation aide for identification and organization of corresponding “domain groups” playing an active role in the use of SDMX technical standards and content-oriented guidelines for the exchange of statistics and related metadata.

3.2. Classification of International Statistical Activities

For this part of the Content-Oriented Guidelines, the SDMX initiative took advantage of an existing categorization scheme, namely the UN Economic Commission for Europe (UNECE) Classification of International Statistical Activities and the Database of International Statistical Activities in the UNECE Region (DISA).

The UNECE framework (Version 2007⁴) has two levels of classification. The first level comprises five “Statistical Domains” that relate to the broad type or statistical activities. The second level specifies the Statistical Areas within the Domains and also provides, in some cases, more detail.

SDMX makes use of the subject-matter framework in Statistical Domains 1-3⁵, which cover:

- 1) Demographic and social statistics
- 2) Economic statistics
- 3) Environmental and multi-domain statistics

A detailed list of SDMX Statistical Subject-Matter Domains is provided in Annex 3.

3.3. Using SDMX Statistical Subject-Matter Domains

The SDMX Statistical Subject-Matter Domain classification provides a high-level scheme for organising statistical data and metadata in many types of applications.

It is anticipated that this classification will be used for one of the basic functions of SDMX at a technical level: the organization of SDMX registries (see the SDMX Technical Specifications, version 2.0). Moreover, for this purpose, it is possible that in future some registries will require a more detailed classification in their areas of focus.

The SDMX Technical Standards version 2.0 provide a mechanism for increasing the granularity (level of detail) of a classification scheme while still making clear where the SDMX Statistical Subject-Matter Domains stop and its own sub-classification begins.

⁴ <http://unece.unog.ch/disa/>

⁵ The UNECE classification was primarily created as a classification of activities. Activities that normally lead to the production of statistical data are covered by domains 1-3 (thus relevant to the SDMX Statistical Subject Matter Domains). Activities related to managerial and support activities that do not directly result in production of statistical data are contained in domains 4 and 5, so these are not relevant for the SDMX Statistical Subject-Matter Domains. Domain 4 covers “Methodology of data collection, processing, dissemination and analysis” and Domain 5 covers “Strategic and managerial issues of official statistics”.

This provides for a high-level interoperability between different SDMX registries, while allowing specific SDMX registries to have the granularity they need. The mappings can be exchanged in the form of SDMX-ML Structure Messages.

In addition to its key role in supporting categorization of data and metadata flows, the development of a list of statistical domains within the SDMX initiative allows the identification of “domain groups”. These groups comprise organizations, working parties, expert groups, task forces, intersecretariat working groups, UN city groups, etc, that are responsible for the development of statistical guidelines and recommendations and identification of best practice for statistics falling within the scope of a particular statistical domain. Working with the UNECE framework should facilitate identifying current or potential participants in various subject-matter domain groups. In particular, one of the objectives of the UNECE framework is the promotion of close co-ordination of statistical activities among international organizations active in the UNECE region. It achieves this close coordination by providing an extensive list of the domain groups, identifying their areas of interest as well as their activities during the year. Activities of the leading international organizations are presented in the UNECE database, whether they relate to the UNECE region or have a broader (worldwide) scope. This guarantees a good level of coverage and a basis for composing the domain groups.

3.4. Representation of Statistical Subject-Matter Domains

The list of SDMX Statistical Subject-Matter domains is available on the SDMX website in the following representations:

- PDF
- SDMX-ML

4. METADATA COMMON VOCABULARY

4.1. Introduction

The Metadata Common Vocabulary contains concepts and related definitions used in structural and reference metadata of international organizations and national data producing agencies.

The MCV covers a selected range of metadata concepts:

- 1) General metadata concepts, mostly derived from ISO, UNECE and UN documents, useful for providing a general context to metadata (for example: classification, metadata registry, statistical metadata, statistical production);
- 2) Metadata terms describing statistical methodologies (for example: frequency, data collection method, data revision, source, adjustment);
- 3) Metadata for assessing quality (for example: accuracy, timeliness);
- 4) Terms referring specifically to data and metadata exchange (terminology from the SDMX information model and from existing data structure definitions, etc., for example: bilateral exchange, gateway exchange).

More specifically, the MCV provides:

- ISO/IEC 11179-compliant definitions for a wide range of statistical metadata terms, which may be used directly, or against which other terminology systems may be mapped. This set of terms is inclusive of the terminology used in the SDMX Technical Standards;

- Definitions for terms on which the SDMX cross-domain concepts work is built;
- Other terminology used within the SDMX initiative.

The MCV is not intended to cover the whole range of statistical terminology, as this area is already covered by other general and domain-specific glossaries. The focus of the MCV is largely those terms that are normally used for building and understanding metadata systems and SDMX data exchange arrangements. A change in the Content-Oriented Guidelines involving SDMX cross-domain concepts implies updating the MCV to reflect these SDMX concepts. In addition, since the cross-domain concepts will be revised and expanded, the MCV will have to follow as new terms need to be included, existing definitions need to be refined and more detailed information need to be added.

A value added of the MCV is also in the opportunity of having one single entry point for accessing a variety of terms, sometimes not available or hard to find on the Internet. In some cases, the MCV deliberately presents one definition linked to several context explanations, always quoting the respective source, sometimes providing additional explanations, other times highlighting peculiarities in how a certain definition is applied within a certain domain or geographical context. Users can live with different metadata models, as long as each concept is well identified and transparent to users. In other words, transparency is a prerequisite for a correct interpretation (and for convergence) of the different statistical frameworks.

4.2. Structure of the MCV

The MCV is built on work already undertaken by several organizations. Where possible, definitions have been drawn from existing international standards or from recommended statistical practices. Where standard definitions were not available or needed adjustment, suitable national definitions have been considered or new definitions formulated.

For each term the following detailed information is provided:

- 1) term
- 2) definition
- 3) context
- 4) source
- 5) related terms

As mentioned above, "context" information is provided extensively throughout the glossary, sometimes to offer additional explanations, other times highlighting peculiarities in how a certain definition is applied within a particular domain or geographical context.

In particular, the MCV also provides information authored as "SDMX", e.g. terms used within the SDMX Technical Specification and Cross-Domain Concepts. This subset of terms is directly maintained by SDMX, while the wider set of metadata terms of more general use with more detailed explanations is linked to external sources and glossaries.

Annex 4 of this document contains the complete listing of the current version of the MCV.

4.3. Representation of Metadata Common Vocabulary

The MCV is envisaged to be made available on the SDMX website in the following representations:

- PDF file
- SDMX-ML

A Word file of the MCV is also available on request from the SDMX Secretariat (secretariat@SDMX.org).

In addition, as described above, web glossaries such as CODED (Eurostat concepts and definitions database) and the OECD Glossary of Statistical Terms contain the MCV terms in a consistent manner.

5. GOVERNANCE, MAINTENANCE AND OUTREACH

5.1. SDMX Cross-domain developments

As part of the efforts to strengthen well-established governance processes and to foster the sustainability of SDMX, a Memorandum of Understanding⁶ (MOU) was signed in March 2007 by all members of the SDMX Sponsors Committee. The MOU provides a section about SDMX products, including Content-Oriented Guidelines.

In line with international best practices, the MOU notes that SDMX will consult widely and openly on the development and maintenance of its various products. The Sponsors Committee will establish adequate processes for the development and maintenance of SDMX products, including consultation and the placing of draft documents related to SDMX products on the SDMX website for public comment.

As a practical matter, the SDMX Secretariat will provide forms on the SDMX website so that it can receive at any time comments or suggestions concerning SDMX Content-Oriented guidelines from the international statistical community (e.g. national statistical agencies, central banks, international organizations, groups) or other interested parties.

On a regular basis, the SDMX Secretariat, with the approval of the SDMX Sponsors Committee, will release suggested amendments to the guidelines for public comment via the SDMX website. As a general principle, following a comment period, the Sponsors Committee will review the changes prepared by the SDMX Secretariat and then approve the latest version of the SDMX Content-Oriented guidelines for release on the SDMX website.

5.2. Domain-specific developments

More generally, SDMX will facilitate the development of domain-specific content-oriented guidelines making use of the SDMX cross-domain guidelines. The Secretariat will provide forms for information to be supplied and then posted on the SDMX website, in order to foster awareness about these SDMX-conformant domain activities as well as to encourage collaboration among statistical experts in national and international statistical agencies.

Broad-based collaboration among institutions and statistical experts has to be ensured, especially to:

⁶ <http://sdmx.org/wp-content/uploads/2007/05/sdmx-memorandum-of-understanding-mou-2007.pdf>

- foster good practices for the development of domain-specific terminology for concepts and code lists;
- facilitate awareness of important issues and possible mapping principles that can be applied to existing classification schemes and systems of countries and international institutions.

Information about domain-specific developments are organised along the lines of the SDMX Statistical Subject-Matter Domains guidelines. Statistical domains conforming to SDMX good practices would be expected to proceed along the following lines:

- 1) Identifying and defining concepts used in the domain and distinguishing between:
 - Concepts that are cross-domain and which should be found in, or proposed to be added to, the list of SDMX cross-domain concepts; and
 - Concepts that are domain-specific and can be articulated by the respective domain group.
- 2) Providing domain-specific data and metadata structure definitions.
- 3) Indicating the various code lists used for the various concepts of the domain.
- 4) Preparing and maintaining a list showing the statistical agencies that agree to use SDMX standards for the statistics of the domain, in particular, distinguishing those that agree to use the data and/or metadata structure definitions provided by the domain.
- 5) Indicating where all the above information is maintained and made available to users by the domain.

5.3. Outreach

In formulating an outreach strategy to involve the international statistical community in the development of SDMX-conformant products, including content-oriented guidelines, the SDMX initiative seeks to consult widely, especially with those concerned with official statistics. SDMX also regularly reports on its activities to the UN Statistical Commission (UNSC)⁷, as all the sponsoring organizations do within their respective working structures.

In addition, SDMX reviews its plans and achievements with the Committee for the Coordination of Statistical Activities (CCSA), which involves more than 25 international organizations and which has adopted SDMX. CCSA also reports on SDMX developments to the UNSC.⁸

6. REFERENCES

- ISO/TS 17369:2005 Statistical Data and Metadata Exchange (SDMX)*, [version 1.0](#), ISO, April, 2005.
- SDMX Technical Standards*, [version 2.0](#), SDMX, November, 2005.
- SDMX Draft [User Guide](#)* (2007)
- Towards Implementation of SDMX: [Conference and Capacity-Building Materials](#)* (2007)

⁷ See Reports to <http://unstats.un.org/unsd/statcom/doc08/2008-13-SDMX.pdf>, <http://unstats.un.org/unsd/statcom/doc07/2007-26e-SDMX.pdf>.

⁸ See Reports to <http://unstats.un.org/unsd/statcom/doc08/2008-26-CCSA-E.pdf>, <http://unstats.un.org/unsd/statcom/doc07/2007-24e-CCSA.pdf>.

556 **7. ANNEXES**

557 Annex 1: Cross-Domain Concepts

558 Annex 2: Cross-Domain Code Lists

559 Annex 3: Statistical Subject-Matter Domains

560 Annex 4: Metadata Common Vocabulary

561 Annex 5: SDMX-ML for Content-Oriented Guidelines (zip file)