

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

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		

74 **1. INTRODUCTION**

75 **1.1. Background**

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

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

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

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

93 **1.2. Cross-domain and domain-specific concepts**

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

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

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

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

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

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

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

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

137 **1.3. Scope of the Content-Oriented Guidelines**

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

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

145 1) Statistical concepts, as described in the "Cross-Domain Concepts" (CDC) guideline.

146 This guideline contains a list of statistical concepts, inter alia related to statistical
147 processes and data quality. This list is based on the concepts used by the
148 contributing international organizations. It is expected to grow in the future.

149 2) Classification of domains, as described in the "Statistical Subject-Matter Domains"
150 (SMD) guideline.

151 This classification is based on the work of the United Nations Economic
152 Commission for Europe (UNECE) to produce a high-level classification of statistical
153 areas. It provides a starting point for organising the exchange of statistical data and
154 metadata, for instance using a registry which provides information needed for
155 locating data and metadata over the Internet.

156 3) Statistical metadata terminology, as described in the "Metadata Common
157 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.

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

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

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

179 **2. CROSS-DOMAIN CONCEPTS**

180 **2.1. Introduction**

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

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

194 The cross-domain concepts outlined below are used in:

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

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

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

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

214 **2.2. Cross-Domain Concepts in data/metadata exchange**

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

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

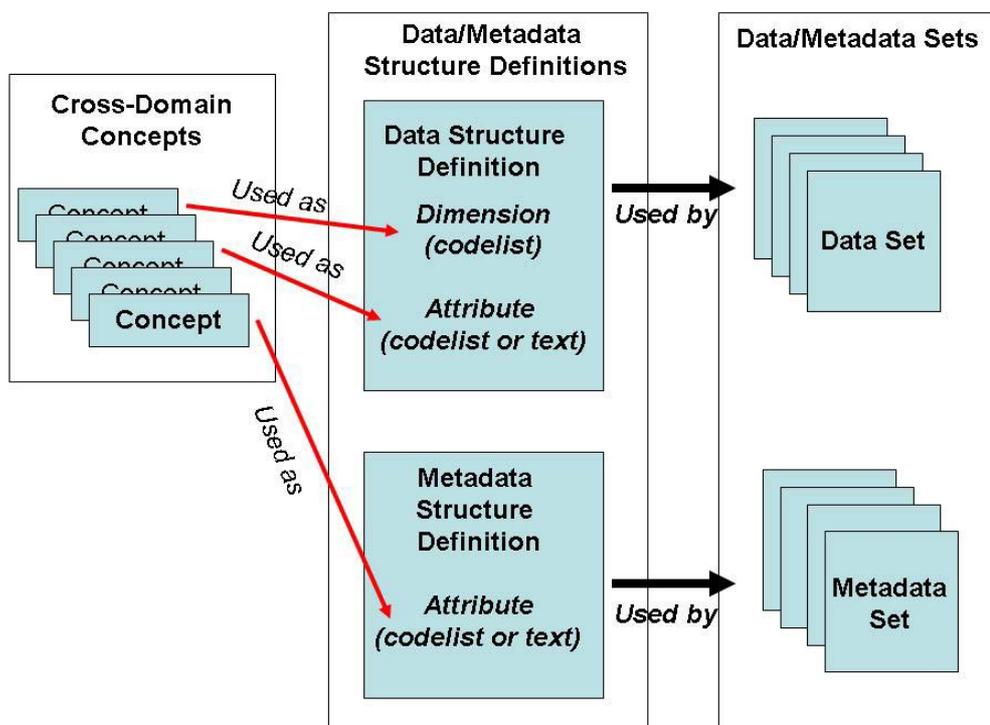
222 A **Metadata Structure Definition** describes how metadata sets, containing reference
223 metadata, are organized. In particular, they define which reference metadata are being
224 compiled, how these concepts relate to each other, how they are represented (either as
225 free text or coded values)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

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

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

239 While the cross-domain concepts include a number of broad reference metadata
240 concepts, organizations may choose to provide information containing more detailed
241 concepts when exchanging reference metadata. Although, in some cases, the list of
242 cross-domain concepts presents some level of detail, it is intended that the high-level
243 concepts, when implemented in metadata structure definitions, would support linkages
244 to metadata available at a more detailed level. As a result, the ability to equate
245 reference metadata sets from different domains or institutions will be enhanced, even if
246 more detailed institution-specific metadata concepts are not always present, or are
247 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.
261 Attributes 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

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

281 **2.3. Describing Cross-Domain Concepts**

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

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

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

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

- 308 1) The identifier of the concept must be unique within the list of cross-domain
309 concepts.
- 310 2) There must be a description of the concept. Definitions and descriptions of general
311 SDMX cross-domain concepts for metadata structure definitions are included in the
312 Metadata Common Vocabulary (MCV).
- 313 3) In case a concept can be represented as coded, there must be a link to a code list
314 containing valid values that may be reported.
- 315 4) The identity of the “key” that defines to what object or structure the concept is to be
316 linked. The term “attached” is often used in SDMX. The “key” can refer to an
317 agency, a data set or a dataflow, or part of a data set or dataflow defined by
318 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.

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

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

326 **2.4. Cross-Domain Concepts and code lists**

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

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

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

335 **2.5. Representation of Cross-Domain Concepts/code lists**

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

- 338 • PDF
- 339 • SDMX-ML

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

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

344 **3. STATISTICAL SUBJECT-MATTER DOMAINS**

345 **3.1. Introduction**

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

355 In some cases, statistical information may be linked to several Subject-Matter Domains.
356 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>

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

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

368 **3.2. Classification of International Statistical Activities**

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

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

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

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

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

383 **3.3. Using SDMX Statistical Subject-Matter Domains**

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

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

390 The SDMX Technical Standards version 2.0 provide a mechanism for increasing the
391 granularity (level of detail) of a classification scheme while still making clear where the
392 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”.

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

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

411 **3.4. Representation of Statistical Subject-Matter Domains**

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

- 414 • PDF
- 415 • SDMX-ML

416 **4. METADATA COMMON VOCABULARY**

417 **4.1. Introduction**

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

421 The MCV covers a selected range of metadata concepts:

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

431 More specifically, the MCV provides:

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

- 436
- Definitions for terms on which the SDMX cross-domain concepts work is built;
- 437
- Other terminology used within the SDMX initiative.

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

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

455 **4.2. Structure of the MCV**

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

461 For each term the following detailed information is provided:

- 462 1) term
- 463 2) definition
- 464 3) context
- 465 4) source
- 466 5) related terms

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

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

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

478 **4.3. Representation of Metadata Common Vocabulary**

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

- 481 • PDF file
- 482 • SDMX-ML

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

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

488 **5. GOVERNANCE, MAINTENANCE AND OUTREACH**

489 **5.1. SDMX Cross-domain developments**

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

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

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

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

509 **5.2. Domain-specific developments**

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

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

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

518 • foster good practices for the development of domain-specific terminology for
519 concepts and code lists;

520 • facilitate awareness of important issues and possible mapping principles that
521 can be applied to existing classification schemes and systems of countries and
522 international institutions.

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

526 1) Identifying and defining concepts used in the domain and distinguishing between:

527 • Concepts that are cross-domain and which should be found in, or proposed to
528 be added to, the list of SDMX cross-domain concepts; and

529 • Concepts that are domain-specific and can be articulated by the respective
530 domain group.

531 2) Providing domain-specific data and metadata structure definitions.

532 3) Indicating the various code lists used for the various concepts of the domain.

533 4) Preparing and maintaining a list showing the statistical agencies that agree to use
534 SDMX standards for the statistics of the domain, in particular, distinguishing those
535 that agree to use the data and/or metadata structure definitions provided by the
536 domain.

537 5) Indicating where all the above information is maintained and made available to
538 users by the domain.

539 **5.3. Outreach**

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

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

550 **6. REFERENCES**

551 *ISO/TS 17369:2005 Statistical Data and Metadata Exchange (SDMX), [version 1.0](#)*, ISO,
552 April, 2005.

553 *SDMX Technical Standards, [version 2.0](#)*, SDMX, November, 2005.

554 *SDMX Draft [User Guide](#) (2007)*

555 *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)