



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

SDMX CONTENT-ORIENTED GUIDELINES:

CROSS-DOMAIN CONCEPTS

(DRAFT MARCH 2006)



- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38

39 © SDMX 2006
40 [HTTP://WWW.SDMX.ORG/](http://www.sdmx.org/)



41

42 1. INTRODUCTION 4

43 2. SCOPE OF THE CONTENT-ORIENTED GUIDELINES..... 4

44 3. CROSS-DOMAIN CONCEPTS 5

45 3.1 Use of Cross-Domain Concepts in Data and Metadata Exchange 5

46 3.2 Describing Cross-Domain Concepts 7

47 3.3 The SDMX Cross-Domain Concepts..... 8

48 4. DEPENDENCIES ON THE USE OF SDMX TECHNICAL STANDARDS..... 10

49 5. REFERENCES 11

50 APPENDIX 1: CROSS-DOMAIN CONCEPTS FOR DATA STRUCTURE DEFINITIONS 12

51 APPENDIX 2: CROSS-DOMAIN CONCEPTS FOR METADATA STRUCTURE DEFINITIONS

52 18

53 APPENDIX 3 - CODELISTS..... 24

54 APPENDIX 4 – ISO/TS 17369 (VERSION 1.0) SDMX-ML 24

55 APPENDIX 5 – VERSION 2.0 SDMX TECHNICAL SPECIFICATIONS SDMX-ML 24

56



57 1. INTRODUCTION

58 The SDMX initiative proposes content-oriented guidelines that are applicable across
59 several statistical subject-matter domains and that represent recommended practices
60 for creating interoperable data and metadata sets using the SDMX technical
61 standards. The content-oriented work focuses on the harmonization of a relatively
62 limited range of specific concepts and terminology that are common to a large number
63 of statistical domains. This harmonisation is necessary to encourage the exchange of
64 comparable statistical information (metadata) and, at this stage, it builds on the
65 experience gained in implementations to date.

66
67 In SDMX, the term “metadata” is very broad and a distinction is made between
68 “structural” metadata that define the structure of statistical data and metadata – and
69 “reference” metadata that describe the actual contents (concepts and methodologies
70 used) and the quality of statistical data. Reference metadata refer or point to data or
71 they can refer to entire collections of data or even the institution that provides the data.

72
73 The proposed cross-domain concepts outlined below are used in:

- 74
75 • *Data structure definitions (key families)* which define the valid content of a data
76 set in terms of the concepts comprising the structure of the data set, how the
77 concepts are related in terms of their role in the data set and the valid content
78 of each of the concepts when used in a data set.
- 79
80 • *Metadata structure definitions* which define the valid content of a metadata set
81 in terms of the concepts comprising the structure of the metadata set, how the
82 concepts are related in terms of their role in the metadata set and the valid
83 content of each of the concepts when used in a metadata set.

84 2. SCOPE OF THE CONTENT-ORIENTED GUIDELINES

85 This is the first draft release of the SDMX Content-Oriented Guidelines, consisting of:

- 86
87 • *SDMX Content-Oriented Guidelines: Cross-Domain Concepts;*
- 88 • *SDMX Content-Oriented Guidelines: Statistical Subject-Matter Domains;*
- 89 • *SDMX Content-Oriented Guidelines: Metadata Common Vocabulary;*

90
91 It is the intent of these guidelines to establish practices in the use of terminology, the
92 structuring of data and metadata sets, and the classification of data and metadata to
93 support the exchange of data and metadata. The content guidelines are designed to
94 work within the specified SDMX technical framework to produce maximum
95 interoperability in the exchange of data and metadata.

96
97 The intent of the SDMX content-oriented guidelines is to encourage reuse where
98 possible across statistical domains in the following three areas:

- 99
100 1. Concepts - as described in the "Cross-Domain Concepts" guideline.
- 101 2. Classifications - as described in the "Statistical Subject-Matter Domains" guideline.
- 102 3. Terminology - as described in the "Metadata Common Vocabulary" (MCV)
103 guideline.

104 3. CROSS-DOMAIN CONCEPTS

105 The list of cross-domain concepts, in the SDMX framework, contains and defines
106 metadata concepts relevant to several statistical domains. SDMX recommends use of
107 the concepts outlined below whenever feasible to promote re-usability and exchange
108 of statistical information and their related metadata between organizations.

109
110 The SDMX initiative expects the list of cross-domain concepts to grow and to be
111 regularly updated as SDMX technical standards and content-oriented guidelines are
112 utilized in more and more statistical domains. The cross-domain concepts guidelines
113 include not only the names of concepts and their content description but also, where
114 appropriate, their representation with standard codelists and the role they play within
115 the data structure definition and/or metadata structure definition.

116
117 The use of the SDMX Cross-domain Concepts where appropriate is not a requisite part
118 of technical conformance, but provides a framework to promote interoperability among
119 those who are conformant with the technical standards. Such interoperability includes
120 the exchange of consistent metadata that can be used to compare concepts and
121 practices in the statistical output of different international organisations and national
122 data-producing agencies.

123 3.1 Use of Cross-Domain Concepts in Data and 124 Metadata Exchange

125
126 As mentioned above, cross-domain concepts are used in SDMX exchange structures;
127 the data structure definition (for data exchange) and the metadata structure definition
128 (for metadata exchange).

129
130 A **Data Structure Definition** (key family) includes information about how concepts are
131 used as dimensions (to identify data) and how other concepts are used as attributes
132 (to qualify data), thus contributing to a full and complete description of a data set when
133 the actual values are given.

134
135 Some types of concepts are needed by data structure definitions to function properly.
136 In those data structure definitions which support time-series data, there will always be
137 a concept reflecting time, and another reflecting frequency. All data structure
138 definitions which describe data sets of any kind will use concepts for observation
139 values and statuses. This guideline provides a specific convention for the use of these
140 types of concepts. All other SDMX cross-domain concepts are concepts that are
141 almost always used the same way in existing data structure definitions. The SDMX
142 guidelines recommend that the proposed cross-domain concepts be used instead of
143 any equivalent concept, although they are not technically required for a data structure
144 definition to function properly. Whenever used, these concepts should have the
145 standard names, roles, and representations defined in the SDMX content-oriented
146 guidelines.

147
148 A **Metadata Structure Definition** refers to the set of structural metadata which
149 describes how a metadata set is organized. These metadata identify what reference
150 metadata concepts are being reported, how these concepts relate to each other, how
151 they may be represented (as free text, as coded values, etc.) and what type of thing
152 they describe – that is, with which object types (agencies, data flows, data providers,
153 or others) they may be associated.

154

155 *Reference metadata* are themselves content metadata that give information which
 156 supports statistical interpretation, making it more meaningful than what is provided in
 157 the data structure definition alone. Among other things, reference metadata describe
 158 concepts, collection and processing methodology and provide information regarding
 159 the accuracy of the statistics they refer to.

160

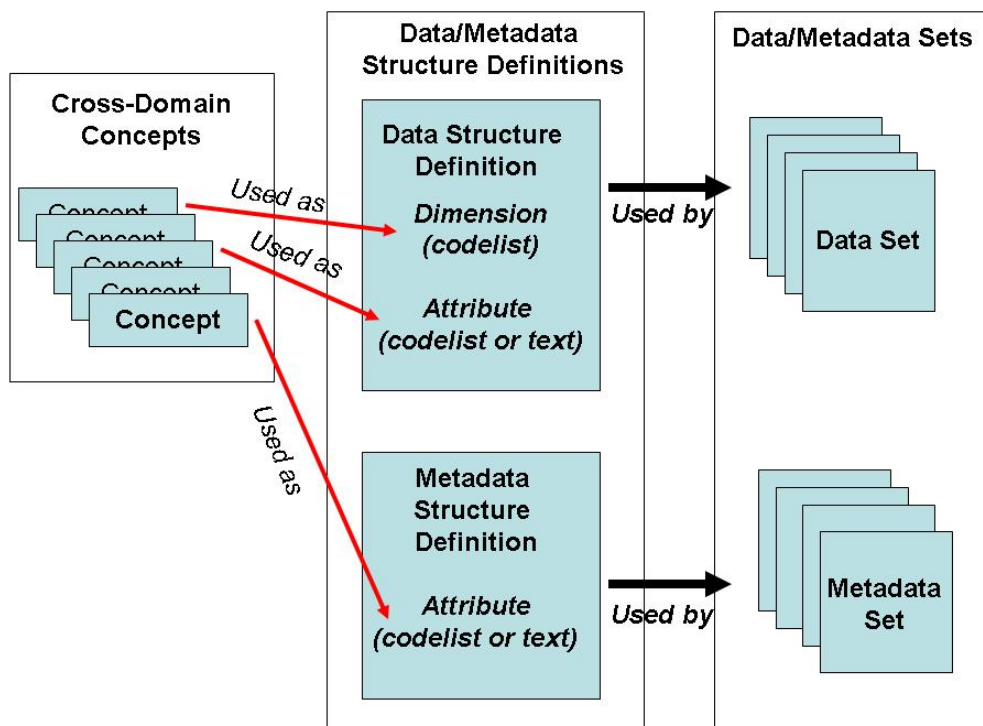
161 An organisation providing information about statistical data must use a set of metadata
 162 concepts (such as source data, periodicity) in order to present the characteristics and
 163 quality of the data. These metadata concepts may be proprietary to the data provider,
 164 but interoperability will clearly be enhanced to the extent the same concepts can be
 165 used by many exchange partners and across many statistical domains. Therefore,
 166 SDMX recommends the use of common concepts, and the concepts provided in these
 167 guidelines are the first step in this direction. The SDMX “cross-domain concepts” are
 168 not technically required in a metadata structure definition but should be used whenever
 169 exchanging information about domains in which they are applicable.

170

171 The proposed SDMX concepts are broad cross-domain reference metadata concepts.
 172 Organisations might provide information contained within more granular (detailed)
 173 concepts when exchanging reference metadata, in addition to the concepts provided
 174 here. The broad set of cross-domain concepts provides the list of high-level concepts
 175 that, when implemented in metadata structure definitions, would support linkages to
 176 information available at a more granular level. As a result, the ability to equate
 177 reference metadata sets from different domains or institutions is enhanced, even if
 178 some detailed metadata concepts are not always present, or are organized differently
 179 in different domains or by different institutions.

180

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



183

184 In the illustration, it can be seen that there is a group of concepts – the cross-domain
185 concepts – which provides a set of named concepts and their definitions. These
186 concepts can be used in three basic ways:

187

188 1. As *Dimensions* in the description of a data structure. Dimensions, when taken
189 together, serve to identify each statistical observation. For example, a
190 dimension named “Reference Area – Country” would explain which country a
191 specific statistical observation refers to, whether this is for the US, New
192 Zealand, or Italy, etc. Dimension values are typically taken from codelists (in
193 this example, a codelist of countries).

194

195 2. As *Attributes* in the description of a data structure. Attributes provide some
196 information about the data which uses that structure. For example, an attribute
197 named “Unit of Measure” might provide information about whether statistical
198 data are measured in currency units or as a pure number. Attributes are
199 sometimes coded, but can also have free-text values. Note that the term
200 “attribute” as used here should not be confused with XML attributes, which are
201 part of the XML syntax used in SDMX-ML.

202

203 3. As *Attributes* in the description of a metadata structure. Termed “reference
204 metadata” in SDMX, these concepts can be used to report metadata which are
205 not directly associated with specific data sets or data points. This could be
206 metadata about a data flow, for example, with concepts such as Timeliness.
207 These values may be coded, but are very often textual.

208

209 Each data set or metadata set uses a structural definition of the appropriate type, so
210 that systems which exchange the data and metadata can understand what the data or
211 metadata set means.

212

213 It is important to note that concepts may be reused in various structural definitions.
214 Thus, a concept such as “Reference Area – Country”, might be used in a data
215 structure definition and in a metadata structure definition. The importance of having a
216 single concept with a single representation for the purposes of exchange across
217 several domains is easy to understand: data and metadata which re-use this single
218 concept can be easily associated, because they are referring to the same idea in the
219 same terms.

220 **3.2 Describing Cross-Domain Concepts**

221 The concepts used in *data structure definitions* must have a specific set of properties
222 according to the SDMX Information Model version 1.0 and version 2.0. These
223 properties include:

224

225 1. Identifier – an identifier of the concept, unique among the concepts maintained by
226 that agency.

227 2. Description / definition – a definition of the concept. SDMX cross-domain concepts
228 for data structure definitions are included in the Metadata Common Vocabulary
229 (MCV).

230 3. Agency - the agency that maintains the concept for use within data structure
231 definitions and elsewhere (a single concept may be used both within a data
232 structure definition and in the reporting of reference metadata using a metadata
233 structure definition). For the SDMX cross-domain concepts, the maintenance

234 agency will always be “SDMX”, while for other concepts, the agency maintaining
235 the concepts will vary.

236

237 The underlying approach to the *metadata structure definition* is similar to that of the
238 data structure definition. The commonality of these structures facilitates software
239 development because similar software components can be used to maintain or
240 process data structure definitions and metadata structure definitions, and to maintain
241 the links between data and metadata.

242

243 A metadata structure definition comprises a number of concepts (examples: source
244 data, periodicity). These concepts can be taken from one or more concept schemes.
245 The following information is defined in the metadata structure definition for each
246 concept:

247

- 248 1. the identity of the concept;
- 249 2. the link to a code list containing valid values that may be reported for the
250 concept; it may be uncoded (free text field) in which case no code list is
251 identified;
- 252 3. whether the reporting of the concept is “mandatory” or “conditional”;
- 253 4. the identity of the key that defines both to what object or structure the metadata
254 reported is to be linked – the term “attached” is often used in SDMX - and the
255 components comprising the identifier or “key” of the object (this is similar in
256 concept to the “attachment level” of attributes in the data structure definition).

257

258 A Metadata Structure Definition defines the components of the “key” to which metadata
259 can be attached. The full key of the object is defined (Target Object Identifier), and
260 optionally any number of partial keys can be defined (Partial Target Object Identifier).
261 Each key and partial key also identifies the object type that it is describing (e.g. a
262 dataflow, a data structure definition, a component in a data structure definition).

263 **3.3 The SDMX Cross-Domain Concepts**

264 The following two tables show how SDMX Cross-Domain Concepts are typically used.
265 The first table provides a list of concepts usually found in data structure definitions,
266 while the second table shows concepts usually found in metadata structure
267 definitions.¹

268

269 All concepts are given a standard name, role, and representation for *whenever they*
270 *are needed*. Thus, if a domain uses a materially similar concept, it should be the one
271 presented here, wherever it is applicable within the domain in question.

272

273 The description of each term, using definitions and context explanations derived from
274 the Metadata Common Vocabulary, is provided in Appendix 1 and 2.

275

¹ In these lists, the distinction between concepts usually found in data structure definitions and metadata structure definitions is largely based on institutional practice of agencies already using GESMES/TS (SDMX-ED) for data exchange. This distinction is provided as an illustration of how cross-domain concepts could be used in SDMX exchange. Concepts usually found in metadata structure definitions may well be reported using a data structure definitions and vice-versa.

276
277
278
279
280
281

Table 1. Cross-domain concepts principally used in data structure definitions

Concept	Representation
Adjustment	Coded
Adjustment detail	Textual
Collection	Coded
Compilation	Coded
Currency	Coded
Decimals	Coded
Description	Textual
Embargo time	Textual
Frequency*	Coded
Frequency detail	Coded
Local name	Textual
Name	Textual
Observation comment	Textual
Observation confidentiality	Coded
Observation status*	Coded
Observation value*	Textual
Pre-break value	Textual
Reference area.country **	Coded
Reporting area.country **	Coded
Sensitivity	Coded
Time*	Textual
Time format*	Coded
Title	Textual
Unit – unit of measurement	Coded
Unit detail	Textual
Unit multiplier	Coded
Vis-à-vis area.country **	Coded

282
283
284
285
286
287
288
289
290
291
292
293
294
295

* These concepts present a conventional approach to meeting the basic needs of data structure definitions. For data structure definitions which support time series data, the concepts of Time, Time format, and Frequency are used. For all data structure definitions, the concepts of Observation value and Observation status are used.

** These concepts are hierarchical – that is, the label after the “.” is a sub-class of the overall concept. This is important because a standard representation can only be assigned at the level of the sub-concept.

296
297
298
299
300
301

Table 2. Cross-domain concepts principally used in metadata structure definitions

Concept	Representation
Accessibility of documentation	Textual
Accounting conventions/basis	Textual
Accuracy	Textual
Classification systems	Textual
Comparability/Coherence	Textual
Confidentiality	Textual
Contact	Textual
Data presentation	Textual
Date of update	Textual
Dissemination formats	Textual
Frequency and Periodicity	Textual
Institutional framework	Textual
Professionalism and ethical standards	Textual
Quality management (including resource management)	Textual
Release calendar	Textual
Relevance	Textual
Revision policy and practice	Textual
Scope / coverage	Textual
Simultaneous release	Textual
Source data	Textual
Statistical concept	Textual
Statistical processing	Textual
Supplementary data	Textual
Timeliness and punctuality	Textual
Transparency	Textual
Validation	Textual

302
303

304 **4. DEPENDENCIES ON THE USE OF SDMX TECHNICAL** 305 **STANDARDS**

306 There are dependencies between the SDMX technical standards and the content-
307 oriented guidelines. The technical standards provide a framework within which the
308 content-oriented guidelines are meaningful, although they also may be useful outside
309 the SDMX technical framework. However, the difficult task of agreement and
310 harmonization of concepts will be facilitated by having a known technical framework
311 within which this process takes place. The technical standards provide "ground rules"



312 for the functioning of systems which helps to define the characteristics of meaningful
313 concepts (metadata) harmonization.

314

315 The content-oriented guidelines outlined in this document may be used in addition to
316 the SDMX technical standards. There is a dependency on the use of the technical
317 standards but it is possible to be conformant with the SDMX technical standards
318 (ISO/TS 17369:2005 SDMX and version 2.0 SDMX technical specifications) and yet
319 not implement the recommendations contained in the content-oriented guidelines. The
320 use of these guidelines is in addition to conformance with the technical standards.

321

322 The dependencies on the SDMX Technical standards typically involve the use of
323 SDMX-ML to express the structural metadata which they standardize. To assist in the
324 use of these guidelines, SDMX-ML files are provided as appendices to the appropriate
325 documents.

326

327 **5. REFERENCES**

328 *SDMX Content-Oriented Guidelines: Framework*, SDMX, March 2006.

329

330 *SDMX Content-Oriented Guidelines: Statistical Subject-Matter Domains*, SDMX, March
331 2006.

332

333 *SDMX Content-Oriented Guidelines: Metadata Common Vocabulary*, SDMX, March
334 2006.

335

336 *ISO/TS 17369:2005 Statistical Data and Metadata Exchange (SDMX), version 1.0*,
337 ISO, April, 2005.

338

339 *SDMX Technical Standards, version 2.0*, SDMX, November, 2005.

340 APPENDIX 1: CROSS-DOMAIN CONCEPTS FOR DATA STRUCTURE DEFINITIONS

341
342 The following table provides details of the cross-domain metadata concepts for data structure definitions.

343
344 In the table below, concepts with * indicate concepts that are needed for a data structure definition to function properly. The concepts of Time,
345 Time format, and Frequency are only needed for those data structure definitions which support time-series data. Concepts with ** indicate
346 concepts for which descriptions or contextual information have not yet been included in the MCV.

Name	ID	Description	Role	Typical attachment level, if attribute	Representation (code list)
Adjustment	ADJUSTMENT	Types of adjustment used for time series MCV: Adjustment methods: The set of procedures employed to improve coverage / classification / timing / valuation of the data or to conform to an accounting/recording basis or address data quality differences in compiling specific data sets.	Attribute	Group	CL_ADJUSTMENT
Adjustment Detail	ADJUSTMENT_DETAIL	A textual explanation of the adjustment methods MCV: Adjustment methods: The set of procedures and/or employed to improve coverage/classification/timing/valuation of the data or to conform to an accounting/recording basis or address data quality differences in compiling specific data sets.	Attribute	Group	Free text
Collection **	COLLECTION	Explanation on how observations are made (e.g. beginning, middle or end of period), indication of whether data are averages, highest or lowest in period, etc.	Attribute	Group	CL_COLLECTION
Compilation	COMPILATION	MCV: The processes for condensing information by manipulating or classifying statistical data into various categories or groups with the object of producing statistics.	Attribute	Group	Free text



Name	ID	Description	Role	Typical attachment level, if attribute	Representation (code list)
Currency **	CURRENCY	Currency of the data.	Attribute or dimension	Group	CL_CURRENCY 3 char. ISO 4217 codes
Decimals **	DECIMALS	Indicates the number of decimals given for the values of the observations of the series.	Attribute	Group	CL_DECIMALS
Description **	DESC	Description of the time series.	Attribute	Group	Free text
Embargo time **	EMBARGO_TIME	A specification of the exact time at which the data could be made available to the public. Note: OPTIONAL ATTRIBUTE WITH A DATE ATTACHED AT ANY LEVEL. Currently under review by SDMX.	Attribute	Group	TIME STAMP
Frequency *	FREQ	Observation frequency of the series MCV: The rate at which something happens or is repeated. Context: If a time series has a constant time interval between its observations, this interval determines the frequency of the time series (e.g. monthly, quarterly, yearly). Note: Only required if the data structure definition can be used for time-series data.	Dimension		CL_FREQ
Frequency detail	FREQ_DETAIL	A further specification of the frequency to include detailed types of frequencies not commonly used. MCV: Frequency: The rate at which something happens or is repeated. Context: If a time series has a constant time interval between its observations, this interval determines the frequency of the time series (e.g. monthly, quarterly, yearly).	Attribute / dimension	Series	CL_FREQ_DETAIL
Local name **	LOCAL_NAME	The series_ID as found in the originating database	Attribute	Series	Free text
Name **	NAME	Name of the time series. NOTE: under discussion.	Attribute	Group	Free text



Name	ID	Description	Role	Typical attachment level, if attribute	Representation (code list)
Observation Comment **	OBS_COM	A textual comment on the observation.	Attribute	Observation	Free text
Observation Confidentiality	OBS_CONF	Information about whether or not the observation can be made known outside the receiving institution MCV: Confidentiality: A property of data, usually resulting from legislative measures, which prevents it from unauthorised disclosure.	Attribute	Observation	CL_OBS_CONF code list under review
Observation Status */ **	OBS_STATUS	Information on quality of value or an unusual or missing value.	Attribute	Observation	CL_OBS_STATUS
Observation Value *	OBS_VALUE	Value of observation. Observation value is the field which holds the exchanged (or provided) data. MCV: Observation: The value, at a particular period, of a particular variable (sometimes called the economic phenomenon).	Attribute	Observation	
Pre-Break Value	PRE_BREAK_VALUE	At a time series break (period) two observations may be recorded: the pre-break value as measured by the “old” standard and the post-break value as measured by the “new” standard. MCV: Time series break: Breaks occurred when there is a change in the standards for defining and observing a variable over time. Such changes may be the result of a single change or the combination of multiple changes at any one point in time of observation of the variable.	Attribute	Observation	Free text



Name	ID	Description	Role	Typical attachment level, if attribute	Representation (code list)
Reference Area **	REF_AREA.xx x The concept will be hierarchical and have many representations REF_AREA.CTY	The country, or geographical/political group of countries that the measured phenomenon relates to.	Dimension		CL_ AREA.xxx CL_ AREA.CTY The code list provided is one of the sub-concepts of area
Reporting Area **	REP_AREA.xx x The concept will be hierarchical and have many representations REP_AREA.CTY	The country reporting the data.	Attribute / Dimension	Group	CL_ AREA.xxx CL_ AREA.CTY The code list provided is one of the sub-concepts of area



Name	ID	Description	Role	Typical attachment level, if attribute	Representation (code list)
Sensitivity and restrictions **	SENSITIVITY	Information on how widely the data can be made available. NOTE: Proposed to replace the “availability” attribute currently used in GESMES/TS	Attribute	Series	CL_SENSITIVITY code list under review
Time *	TIME ²	Identification of time for the measured phenomenon. MCV: Time period: The time interval of single repetition of a varying quantity of a motion or phenomenon which repeats itself regularly. Context: The period is the reciprocal of the frequency. More loosely, the expression is used to denote the time interval or average interval between identifiable points of recurrence, e.g. between peaks or troughs of the series (month, quarter, year,). In GESMES/TS, a period is a time reference. (GESMES/TS User Guide", Release 3) Note: Only required if the data structure definition can be used for time-series data.	Dimension		
Time Format * / **	TIME_FORMAT	Format that the time is represented in. Note: Only required if the data structure definition can be used for time-series data and is intended for use with or transformation into SDMX-EDI formats.	Attribute	Series	CL_TIME_FORMAT List not complete
Title **	TITLE	Title of the time series	Attribute	Group	Free text
Unit (of measure)	UNIT	MCV: The actual unit in which the associated values are measured. (Unit of measure)	Attribute	Group	CL_UNIT code list under review

² TIME APPLIES TO TIME SERIES EXCHANGE ONLY. Time is a required dimension for all data structure definitions which are not strictly for cross-sectional use. It is considered good practice to include time even in those data structure definitions where the intended use of the data does not include time-series presentations, as others may wish to use such presentations of the data.



Name	ID	Description	Role	Typical attachment level, if attribute	Representation (code list)
Unit Detail **	UNIT_DETAIL	Additional information on unit.	Attribute	Group	Free text
Unit Multiplier **	UNIT_MULT	Exponent in base 10 specified so that multiplying the observation numeric values by $10^{\text{UNIT_MULT}}$ gives a value expressed in the UNIT NOTE: Use of intervals in version 2.0 instead of code lists has been suggested.	Attribute	Group	CL_UNIT_MULT
Vis-à-vis Area **	VIS_AREA.xxx The concept will be hierarchical and have many representations VIS_AREA.CTY Y	The country to which the measured data is in relation to.	Attribute / Dimension	Group	CL_AREA.xxx CL_AREA.CTY The code list provided is one of the sub-concepts of area

349

APPENDIX 2: CROSS-DOMAIN CONCEPTS FOR METADATA STRUCTURE DEFINITIONS

350

351

Name	ID	Description and Content	Role	Typical attachment level	Representation
Accessibility of Documentation	Accessibility_of_documentation	Refers to the availability of documentation of various aspects of the data (sources and methods documents) and the content of such documentation.	Attribute	Data flow	Free text
Accounting conventions/basis	Accounting_conventions_basis	Refers to descriptions of the types of prices used to value flows and stocks or other units of measurements used for recording the phenomena being observed; the time of recording of the flows and stocks or the time of recording of other phenomena that are measured, including the reference period employed; and the grossing/netting procedures that are used. The description could also include how consistent the practices used are with internationally accepted standards, guidelines, or good practices.	Attribute	Data flow	Free text
Accuracy	Accuracy	Refers to the provision of either measures of accuracy or precision (numerical results of the methods/processes for assessing the accuracy or precision of data) or qualitative assessment indicators. It may also be described in terms of the major sources of error that potentially cause inaccuracy. It includes providing the results of the assessment of source data for coverage, sampling error, response error and non-sampling error.	Attribute	Data flow	Free text
Classification systems	Classification_systems	Classification systems refer to a description of the classification systems being used and how they conform with internationally accepted standards guidelines, or good practices. It also refers to the description of deviations of classification systems compared to accepted statistical standards, guidelines, or good practices, when relevant.	Attribute	Data flow	Free text



Name	ID	Description and Content	Role	Typical attachment level	Representation
Comparability/Coherence	Comparability_coherence	<p>Comparability refers to the extent to which differences between statistics from different geographical areas, non-geographical domains, or over time, can be attributed to differences between the true values of the statistics. The coherence of statistics refers to their adequacy to be reliably combined in different ways and for various uses.</p> <p>Context: The coherence of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytic framework and over time. The use of standard concepts, classifications and target populations promotes coherence, as does the use of common methodology across surveys. Coherence does not necessarily imply full numerical consistency.</p>	Attribute	Data flow	Free text
Confidentiality	Confidentiality	Refers to the legislative measures or other formal provision which prevent unauthorised disclosure of data that identify a moral or physical person either directly or indirectly. Also refers to the procedures in place to prevent disclosure of confidential data, including rules applying to staff, aggregation rules when disseminating data, provision of unit records, etc.	Attribute	Agency, Data flow	Free text
Contact	Contact	Describes contact points for the data or metadata, including how to reach the contact points.	Attribute	Data flow	Free text
Data presentation	Data_presentation	Refers to descriptions of the ways the data are presented. It would include descriptions of the table contents, with their data breakdowns, etc. It should also include summary information on units of measurement, time span covered, adjustments to data (e.g., seasonal adjustments for time series) and availability of textual analysis of current-period development with the dissemination of the data	Attribute	Data flow	Free text
Date of update	Date_of_update	Refers to the date on which the metadata element was inserted or modified in the database. It can be further detailed in: a) last update of content; b) last certified without update; c) last posted on web site.	Attribute	Data flow	Free text



Name	ID	Description and Content	Role	Typical attachment level	Representation
Dissemination formats	Dissemination_formats	Refers to the various means of dissemination used for making the data available to the public. It would include a description of the various formats available, including where and how to get the information (paper, electronic formats, longer time series).	Attribute	Agency, Data flow	Free text
Frequency and Periodicity	Frequency_periodicity	<p>Frequency refers to the time interval between the observation of a time series. Periodicity refers to the frequency of compilation of the data (e.g., a time series could be available at annual frequency but the underlying data are compiled monthly, thus have a monthly periodicity).</p> <p>Context: The periodicity of a particular data category is determined by several factors, including the ease of observation or compilation and the needs of analysis. Periodicity is usually expressed in terms of divisions of the calendar (e.g. monthly, quarterly). Periodicity of original data refers to the frequency of compilation of data by the source agency, i.e. the national agency or international organisation that provided the information.</p>	Attribute	Data flow	Free text
Institutional framework	Institutional_framework	Refers to a law or other formal provision that assign primary responsibility as well as the authority to an agency for the collection, processing, and dissemination of the statistics; it also includes arrangements or procedures to facilitate data sharing and coordination between data producing agencies.	Attribute	Agency, Data flow	Free text



Name	ID	Description and Content	Role	Typical attachment level	Representation
Professionalism and ethical standards	Professionalism_ethical_standards	Describes the elements providing assurances that statistics are produced on an impartial basis; elements providing assurances that the choices of sources and statistical techniques as well as decisions about dissemination are informed solely by statistical considerations; elements providing assurances that the recruitment and promotion of staff based are based on relevant aptitude; elements providing assurances that the statistical entity is entitled to comment on erroneous interpretation and misuse of statistics, guidelines for staff behaviour and procedures used to make these guidelines known to staff; other practices that provide assurances of the independence, integrity, and accountability of the statistical agency.	Attribute	Agency	Free text
Quality management (including resource management)	Quality_management	Quality management refers to processes in place to focus on quality, to monitor the quality of the statistical programs, to deal with quality considerations in planning the statistical programs. It also includes how well the resources meet the requirements of the program(s), and measures to ensure efficient use of resources (staffing, facilities, computing resources, and financing of statistical programs).	Attribute	Agency	Free text
Release calendar	Release_calendar	Describes the policy regarding the release of statistics according to a preannounced schedule and its availability. It also contains the release calendar information.	Attribute	Data flow	Free text
Relevance	Relevance	Refers to the processes for monitoring the relevance and practical utility of existing statistics in meeting users' needs and how these processes inform the development of statistical programs.	Attribute	Agency , Data flow	Free text
Revision policy and practice	Revision_policy_practice	Describes the data revision policy; the policy and practice for identifying the data revision status of available data; the availability of revision studies and analyses.	Attribute	Data flow	Free text



Name	ID	Description and Content	Role	Typical attachment level	Representation
Scope / coverage	Scope_cov erage	Describes the scope/coverage of the statistics and how consistent it is with internationally accepted standards, guidelines, or good practices. The scope/coverage include a description of target population, and geographic, sector, institutional, item, population, product, and other coverage.	Attribute	Data flow	Free text
Simultaneous release	Simultaneous_r elease	Describes the policy for release of the data to the public, how the public is informed that the data are being released, and whether the policy provides for the dissemination of statistical data to all interested parties at the same time. It also describes the policy for briefing the press in advance of the release of the data.	Attribute	Agency, Data flow	Free text
Source data	Souce_data	Refers to a description of the data collection programs and their adequacy for the production of statistics, including meeting the requirements for methodological frameworks, scope, classifications systems, and basis for recording.	Attribute	Data flow	Free text
Statistical concept	Statistical_conc ept	Refers to the internationally accepted statistical standards, guidelines, or good practices on which are based the concepts and definitions that are used for compiling the statistics. It also refers to the description of deviations of the concepts and definitions from accepted statistical standards, guidelines, or good practices, when relevant.	Attribute	Data flow	Free text
Statistical processing	Statistical_proc essing	Refers to a description of the data compilation practices and other statistical procedures to deal with intermediate data and statistical outputs (e.g., data adjustments and transformation, and statistical analysis).	Attribute	Data flow	Free text
Supplementary data	Supplementary _data	Refers to a description of data not routinely disseminated that are made available to users upon request. May include customized tabulations that can be provided (perhaps for a fee) to meet specific requests. Also include information on procedures for obtaining these supplementary data.	Attribute	Data flow	Free text



Name	ID	Description and Content	Role	Typical attachment level	Representation
Timeliness and punctuality	Timeliness_punctuality	<p>Timeliness: refers to the speed of dissemination of the data - i.e., the lapse of time between the end of a reference period (or a reference date) and dissemination of the data. It reflects many factors, including some that are related to institutional arrangements, such as the preparation of accompanying commentary and printing.</p> <p>Punctuality: Punctuality refers to the possible time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar, laid down by Regulations or previously agreed among partners.</p>	Attribute	Data flow	Free text
Transparency	Transparency	<p>Describes the policy on the availability of the terms and conditions under which statistics are collected, processed, and disseminated. It also describes the policy of providing advanced notice of major changes in methodology, source data, and statistical techniques; the policy on internal governmental access to statistics prior to their release; the policy on statistical products' identification</p>	Attribute	Agency, Data flow	Free text
Validation	Validation	<p>Describes methods/processes for routinely assessing source data—including censuses, sample surveys, and administrative records; describes how the results of the assessments are monitored and made available to guide statistical processes. Also describes how intermediate results are validated against other information where applicable; describes how statistical discrepancies in intermediate data are assessed and investigated; describes how statistical discrepancies and other potential indicators or problems in statistical outputs are investigated. All the controls made in terms of quality of the data to be or already published are included in the validation process. Also describes the results of studies and analysis of revisions and how they are used to inform the statistical processes.</p>	Attribute	Data flow	Free text

352

353



354 **APPENDIX 3 - CODELISTS**

355
356
357
358
359

NOTE: A document containing codelists will be provided as a separate document of the *SDMX Content-Oriented Guidelines: Cross-Domain Concepts* for review purposes.

360 **APPENDIX 4 – ISO/TS 17369 (VERSION 1.0) SDMX-ML**

361
362
363
364
365
366
367
368
369
370
371
372
373
374
375

[Note: This appendix will contain the canonical SDMX-ML formats for expressing appropriate pieces of this guideline. The contents are all structural metadata, which will be expressed as two structure messages, containing:

- *Concepts for Cross-Domain Concepts for Data Structure Definitions*
- *Codelists*

These will not be attached until after review of the concepts and codelists has been completed.]

376 **APPENDIX 5 – VERSION 2.0 SDMX TECHNICAL SPECIFICATIONS SDMX-ML**

377
378
379
380
381
382
383
384
385
386
387
388
389

[Note: This appendix will contain the canonical SDMX-ML formats for expressing appropriate pieces of this guideline. The contents are all structural metadata, which will be expressed as three structure messages, containing:

- *Concepts for Cross-Domain Concepts for Data Structure Definitions*
- *Concepts for Cross-Domain Concepts for Metadata Structure Definitions*
- *Codelists*

These will not be attached until after review of the concepts and codelists has been completed.]