



INTERSTATE STATISTICAL COMMITTEE OF THE COMMONWEALTH OF INDEPENDENT STATES

WELCOME TO THE STATISTICS OF THE COMMONWEALTH OF INDEPENDENT STATES

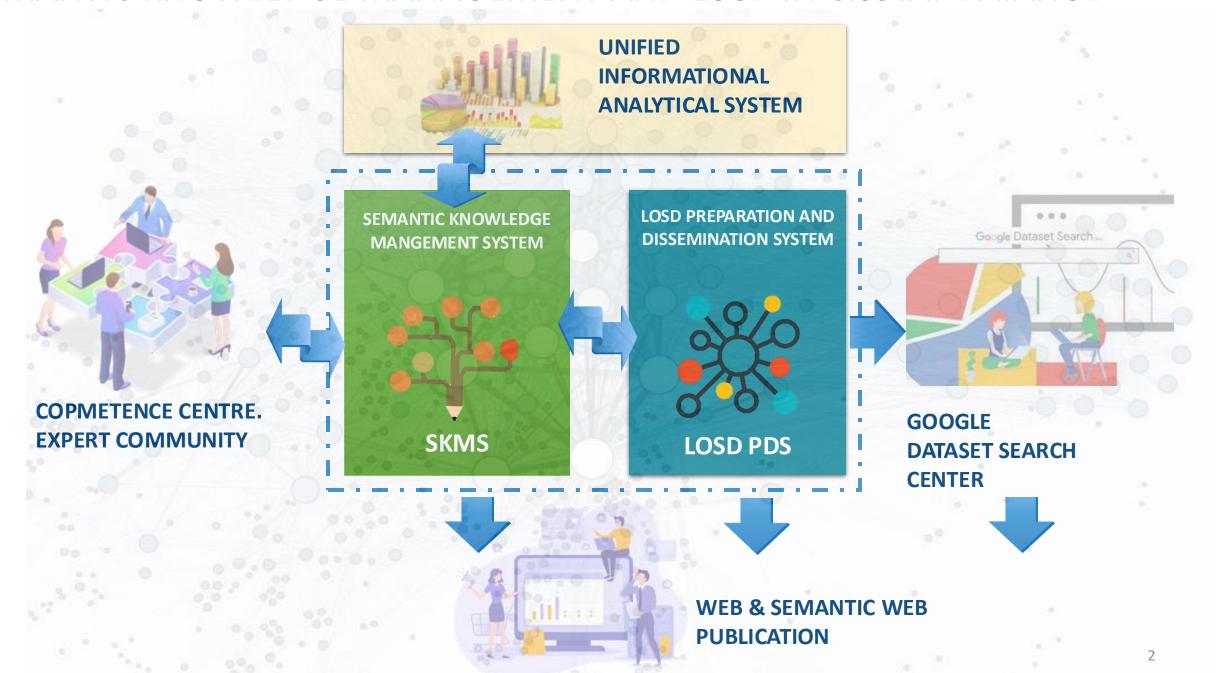
CISStat Linked Open Statistics: The Approach to Semantic Transformation of SDMX

Konstantin Laykam
Chairman of
CIS Statistical Committee

Yury Akatkin Interoperability Basis Slovenia

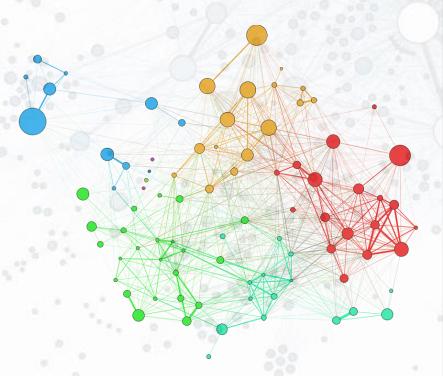


SEMANTIC KNOWLEDGE MANAGEMENT AND LOSD IN CISSTAT DATAHUB



CISSTAT LINKED OPEN STATISTICS

SEMANTICALLY RICH INTERPRETATION ENVIRONMENT



IMPROVE THE QUALITY OF STATISTICAL DATA AND METADATA

HARMONIZE STATISTICAL TERMINOLOGY AND ALIGHN METHODOLOGY

COMPLY WITH FAIR PRINCIPLES

PROVIDE SEMANTIC INTEROPERABILITY

FACILITATE (META)DATA
RELEVANT INTERPRETATION



SEMANTIC MODELS



SMART METADATA

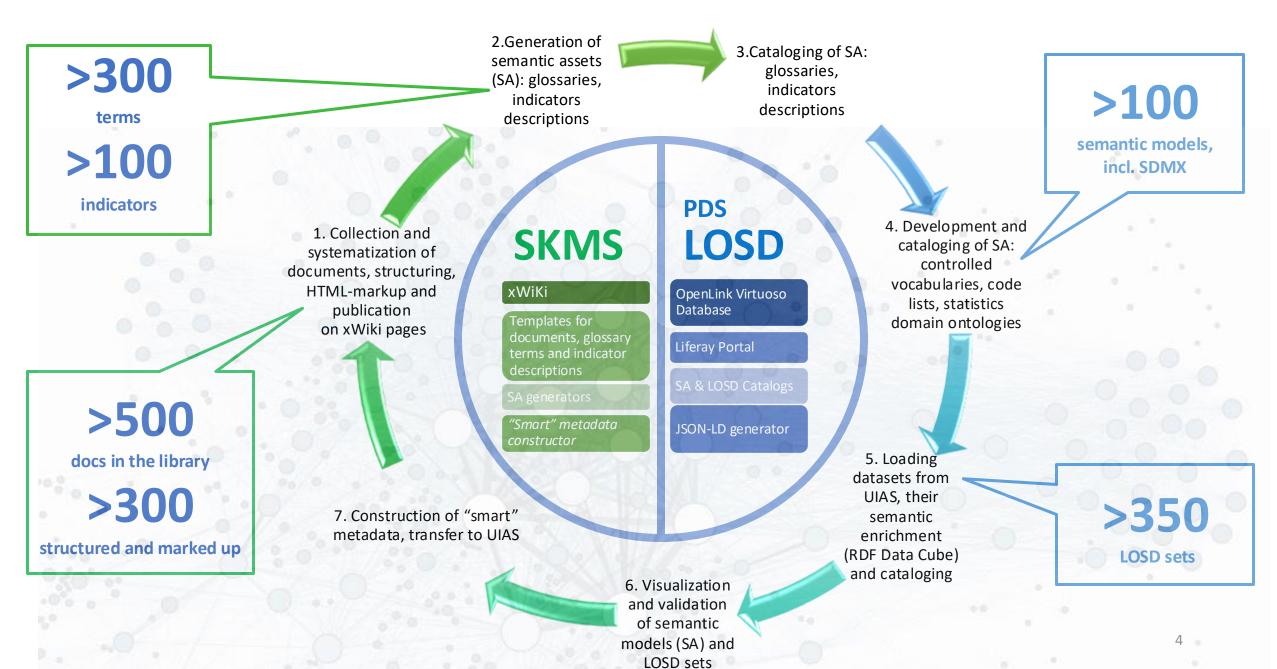


SEMANTICALLY RICH LOSD

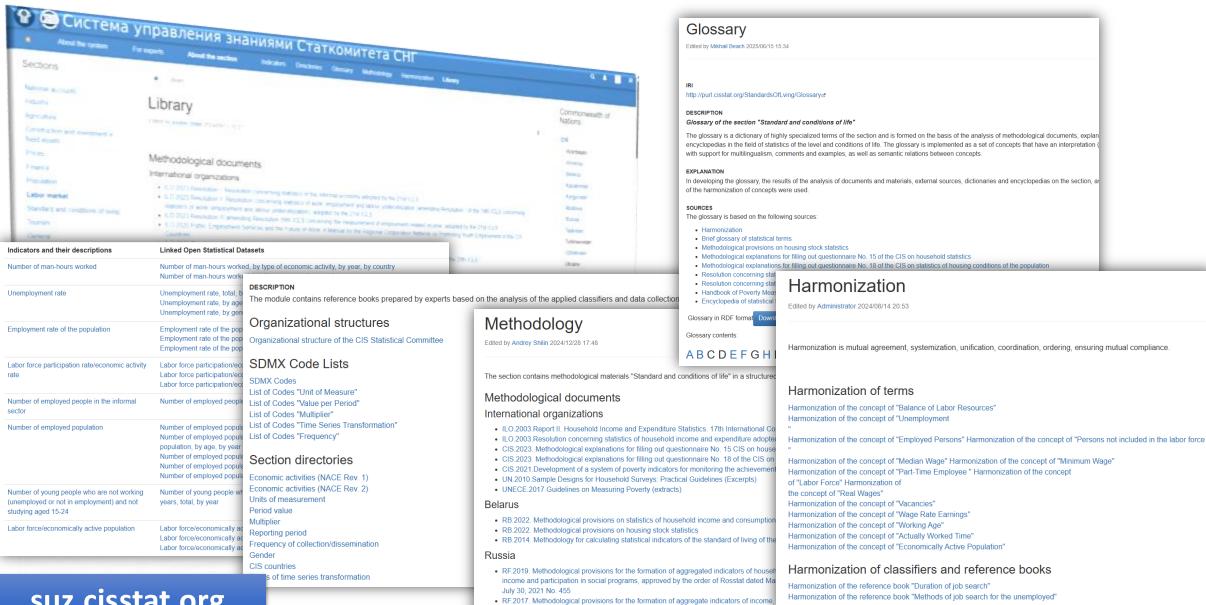


VISUALISATION FOR EXPERT VALIDATION

CISSTAT. TOWARDS 15 STATISTICAL DOMAINS



CISSTAT SEMANTIC KNOWLEDGE MANAGEMENT SYSTEM



. RF.2011. Housing stock and public utilities (methodological provisions)

household budgets

suz.cisstat.org

THE ROLE OF SDMX IN LOSD



Information Model

RDF DATA CUBE VOC

Semantic interoperability, Flexibility, Crossdomain use, Semantic Web dissemination

Applied standard, Commonly used, Efficient sharing across NSS





Rich semantics & context, LOSD interpretation environment,

Expert community

CHALLENGES OF SDMX IMPLEMENTATION FOR LOSD

RDF DATA CUBE VOCABULARY

Based on an earlier, now **outdated version** of the **SDMX Information Model** (specification published in 2009)

No namespace maintenance since 2013 and no alignment with best practices of **URI persistence** promoted by interoperability experts (e.g., in EU ISA²)

Limited set of semantic models, not supporting the full range of classifications required for LOSD development (only 9 code lists and 75 codes)

INTEROPERABILITY BASIS

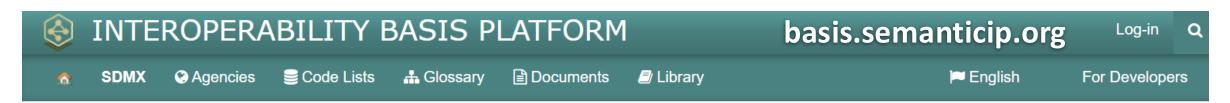
Open, non-profit initiative aimed at overcoming technological and organizational barriers that hinder the effective exchange and dissemination of Linked Data



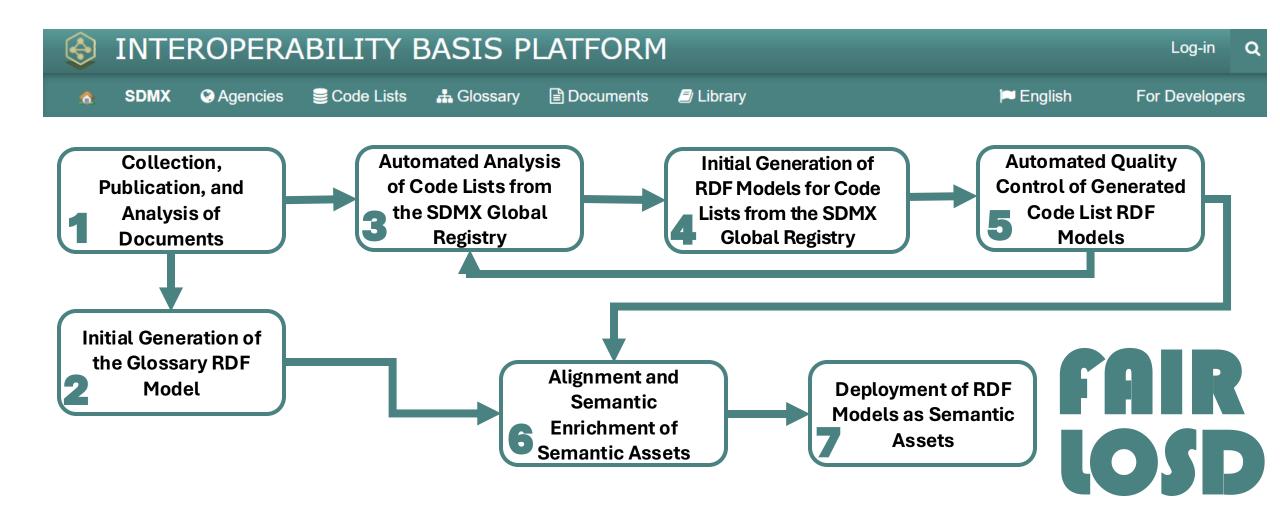
To integrate existing data exchange standards, classifications, and reference systems into the Semantic Web environment to achieve **sustainable semantic** interoperability across a wide range of user scenarios



Interoperability Basis Platform (IoBP) supports semantic alignment, enrichment, and publication of existing standards using a knowledge management system, modeling tools, namespace control, and persistent URI infrastructure



SEMANTIC TRANSFORMATION PIPELINE. SDMX DOMAIN



Interoperability Basis (C) https://basis.semanticip.org/xwiki/bin/view/Main/

BENEFITS OF USING IOBP IN CISSTAT

SDMX context navigation; interactive discovery and exploration; linked elements and backlink analysis

▼ / SDMX ▼ / SDMX Documents ▼ / SDMX 3.1 Standards. Section 1. Summary of Major Changes and New Functionality ▼

SDMX 3.1 Standards. Section 1. Summary of Major Changes and New Functionality

Last modified by Artur on 2025/09/13 19:12

Document card

Download document

Contents

- 1. Overview
- 2. Summary of Breaking Changes in 3.1
 - 2.1 Removal of Advanced Release Calendar
 - 2.2 Removal of Version on Categorisation
- 3. Information Model
 - 3.1 Horizontally Complex Data Structure Definitions
 - 3.2 Constraint Cohesion

1 Overview

SDMX 3.1 is a minor revision to the SDMX 3.0 Standard which introduces a limited set of changes, which cover the following:

Information Model

- . Support for <u>Dataflows</u> to reference a subset of <u>Dimensions</u> from a <u>Data Structure Definition</u>
- . Simplification to Data Constraints
- · Addition of Availability Constraints

Documentation

Registering Reference Metadata removed from d

Breaking Changes

- Remove <u>version</u> property on Categorisation
- Removal of Advanced Release Calendar

Data structure definition

Set of structural metadata associated to a data set, which includes information about how concepts are associated with the measures, dimensions, and attributes of a data cube, along with information about the representation of data and related descriptive metadata.

reference a Metadata Provision, ar

ALTERNATIVE NAME 🥒

DEFINITION 🥒

Set of <u>structural metadata</u> associated to a <u>data set</u>, which includes information about how <u>concepts</u> are associated with the <u>measures</u>, <u>dimensions</u>, and <u>attributes</u> of a data cube, along with information about the <u>representation</u> of data and related descriptive metadata.

COMMENT /

A DSD defines the structure of an organised collection of data (Data Set) by means of concepts with specific roles, and their representation. In order to exchange or disseminate statistical information, an institution needs to specify which statistical concepts are necessary for identifying the series (and for use as dimensions) and which statistical concepts are to be used as attributes and measures. These definitions form the data structure definition. In a data collection scenario the specification of the data structure definition is often a collaborative venture between the collecting institution and its partners. There are three types of construct in the DSD: Dimension, Attribute, and Measure. Each of these combines a Concept with its representation (this can be either a reference to a Code list or a non-coded data type such as "integer", "string", "date/time"). The roles of the three types of construct (Dimension, Attribute, and Measure) are as follows: A Dimension is an identifying component, sometimes referred to as a "classificatory variable". When a value is given to each of the Dimensions in a data set (this is often called a "key" or a "series") the resulting key, when combined with a time value, uniquely identifies an observation. For instance, country, indicator, measurement unit, frequency, and time dimensions together identify the cells in a cross-country time series with multiple indicators (for example, gross domestic product, gross domestic debt) measured in different units (for example, various currencies, percent changes) and at different frequencies (for example, annual, quarterly). The cells in such a multi-dimensional table contain the observation values. The DSD construct that specifies the Concept and expected representation of an observation is called a Measure. The semantics of the measure are derived from the Dimensions or a sub set of them and, if not specified in a Dimension, an Attribute indicating the measurement unit e.g. indicator and measure unit (gross domestic product perce

BROADER /

Data set

NOTATION /

urn:sdmx:org.sdmx.infomodel.conceptscheme.Concept=SDMX:CROSS_DOMAIN_CONCEPTS(2.0).DSD €

RELATED

Attribute, Dimension, Measure

Used in the following terms: Attachment level, Attribute relationship, Code list, Comment, Component

More (14)

Backlinks: 1 Introduction, 1 Introduction, 1 Introduction, 1 Introduction, 1 Purpose and Structure

More (93)

1. Purpose and Structure, 1. Purpose and Structure, 10 Community, Management, 10 Constraints, 10 Constraints, 10 Constraints, 10 Validation and Transformation Language (VTL), 11 Annex 1 — Content Oriented Guidelines (COG), 11 Transforming between versions of SDMX, 11 Transforming between versions of SDMX, 12 Annex 2 — SDMX Business Process Model, 12 Constraints, 12 Validation and Transformation Language, (VTL), 12 Validation and Transformation Language, 13 Structure Mapping, 13 Validation and Transformation Language, 14 Annex 4 — Data Reader and Data Writer Functions, 15 Validation and Transformation Language, 15 Validation and Transformation Language, 16 Annex 6: Worked Use Case, 2 Actors and Use Cases, 2 Actors and Use Cases, 2 General Notes on This Document, 2 General Notes for Implementers, 4 Specific Item Schemes, 4 Specific Item Schemes, 5 Specific Item Schemes, 4 Specific Item Schemes, 5 Specific Item Schemes, 6 Specific Item Schemes, 8 Specific Item Schemes, 9 Specific

Content of the Document

BENEFITS OF USING IOBP IN CISSTAT

Semantic SDMX Glossary with human-readable views; exploring relationships, context and usage across SDMX versions

Accounting conventions

Last modified by Artur on 2025/08/04 16:06

STATUS Upload

https://purl.semanticip.org/linked-data/sdmx/concept/ACC CONV

INTERNATIONAL NAME

Accounting conventions

DEFINITION

Practical procedures, standards and other aspects used when compiling dat

COMMENT

This metadata element refers to descriptions of the types of prices used to v of measurements used for recording the phenomena being observed. time of recording of the flows and stocks or the time of recording of other phe are used. Accounting conventions may refer to whether the data are recorde calendar year) employed. The description could also include how consistent SNA (System of National Accounts) - or good practices.

BROADER

Unit of measure, Reference period

urn:sdmx:org.sdmx.infomodel.conceptscheme.Concept=SDMX:CROSS_DO

Recommended representation: String

Used in the following terms:

Backlinks: -



Fags: Accounting conventions

Glossary terms in the table

Last modified by Artur on 2025/07/14 10:16

lame	URI	name	Definition	Comment	
Accounting conventions	https://purl.semanticip.org/linked- data/sdmx/concept/ACC_CON Ye		Practical procedures, standards and other aspects used when compiling data from diverse sources un- der a common methodological framework.	This metadata element refers to descriptions of the types of prices used to value flows and stocks, or other <u>units of measurements</u> 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 <u>reference period</u> employed, and the grossing/netting procedures that are used. Accounting conventions may refer to whether the data are recorded on a cash/accrual or mixed accounting basis, the time of their recording and the <u>reference period</u> (fiscal or calendar year) employed. The description could also include how consistent the practices used are with internationally accepted standards - such as the Balance of Payments Manual or SNA (System of National Accounts) - or good practices.	-
Accuracy.	https://purl.semanticip.org/linked-data/sdmx/concept/ACCURAC Yv		Closeness of computations or estimates to the unknown exact or true values that the statistics were intended to measure.	The accuracy of statistical information is the degree to which the information correctly describes the ph statistical estimates are variance (random errosures of accuracy (nur data) or qualitative ass the major sources of e pling, non-response, reof the data, which is disubsequent estimated	
				Ctatistical data and materials avalages	Ĺ

Alternative

I or SNA (System of National Accounts) - or good practices. Frequency of terms mentioned in the table Last modified by Artur on 2025/07/14 10:16

Broader

Unit of measure,

Reference period

Interchangeable

(skos:exactMatch)

Number of mentions of

sip-sdmx-concept:ACC_CONV a skos:Concept ;
rdfs:comment "This metadata element refers to descriptions of the types of prices used to value
flows and stocks, or other <a norefe<="" rel='\"noopener' target='\"_blank\"' td="">
rrer\" href=\"https://basis.semanticip.org/xwiki/wiki/sdmx/view/Glossary/Unit%20of%20measure/\">units
of measurements used for recording the phenomena being observed; the time of recording of t
he flows and stocks or the time of recording of other phenomena that are measured, including the <span< td=""></span<>
class=\"wikiexternallink\"> <a href='\"https://basis.seman</td' noreferrer\"="" rel='\"noopener' target='\"_blank\"'>
ticip.org/xwiki/wiki/sdmx/view/Glossary/Reference%20period/\">reference period employed; an
d the grossing/netting procedures that are used. Accounting conventions may refer to whether
the data are recorded on a cash/accrual or mixed accounting basis, the time of their recording and the
<pre><a href='\"https://basi</pre' noreferrer\"="" rel='\"noopener' target='\"_blank\"'></pre>
s.semanticip.org/xwiki/wiki/sdmx/view/Glossary/Reference%20period/\">reference period (fisc
al or calendar year) employed. The description could also include how consistent the practices used ar
e with internationally accepted standards - such as the Balance of Payments Manual or SNA (System of N
ational Accounts) - or good practices.";
swvs:term_status "upload" ;

ational Accounts) - or good practices." ; swvs:term status "upload" ; skos:broader sip-sdmx-concept:REF_PERIOD, sip-sdmx-concept:UNIT MEASURE :

refix rdfs: <http://www.w3.org/2000/01/rdf-s

@prefix swvs: <http://www.w3.org/2003/06/sw-vocab-status/ns#> .

@prefix sip-sdmx-concept: <https://purl.semant

@prefix skos: <http://www.w3.org/2004/02/skos/

skos:definition "Practical procedures, standards and other aspects used when compiling data fro m diverse sources under a common methodological framework.";

skos:inScheme sip-sdmx-concept:CROSS_DOMAIN_CONCEPTS;

skos:notation "urn:sdmx:org.sdmx.infomodel.conceptscheme.Concept=SDMX:CROSS_DOMAIN_CONCEPTS(2.0).A CC_CONV";

skos:note "Recommended representation: String"; skos:prefLabel "Accounting conventions"; skos:related sip-sdmx-concept:REF_PERIOD,

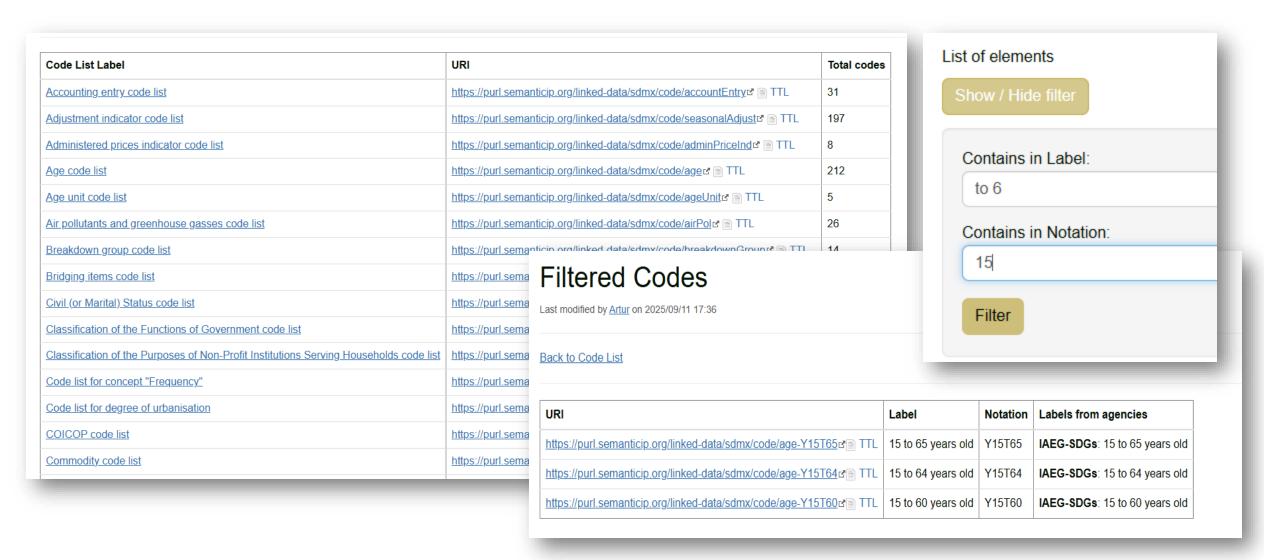
sip-sdmx-concept:UNIT_MEASURE skos:topConceptOf sip-sdmx-concept:CROSS_DOMAIN_CONCEPTS .

Term	Number of articles mentioning the term
Statistical data and metadata exchange	157
<u>Data set</u>	120
Data structure definition	117
<u>Attribute</u>	113
Dimension	108
Code list	106
Code	98
Component	95
Artefact	89
Dataflow	87
Reference metadata	76
Structural metadata	74
<u>Measure</u>	72

CC CONV

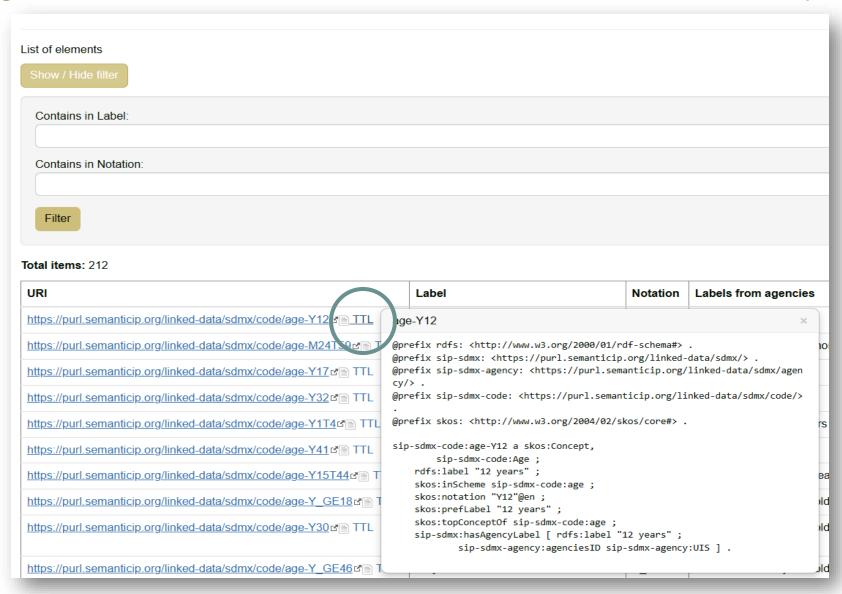
BENEFITS OF USING JOBP IN CISSTAT

Interactive visualization, navigation and filtering of SDMX codes in code lists



BENEFITS OF USING IOBP IN CISSTAT

Dereferencing and one-click access to a machine-readable model of the code (TTL)



APPLICATION OF SEMANTIC SDMX IN CISSTAT LOSD

Number of employed population

INTERNATIONAL NAME

Number of employed

SUBSECTIONS

Labor resources

DESCRIPTION OF THE INDICATOR

The indicator characterizes the number of employed people (aged 15 years and older , established for measuring the lab

According to the document Methodological Explanations for Completing Questionnaire No. 14, the employed populatio all sectors of the economy. Thus, the number of employed should include persons working in state enterprises and organi (peasant) households, as well as those engaged in individual labor activity, in personal subsidiary farming and for individual

LINKS TO REGULATORY DOCUMENTS

- . Resolution I of the 19th ICLS on statistics of work, employment and labour underutilization
- · Methodological explanations for filling out questionnaire No. 14

DATA SOURCES

Form Table 14.1b. Distribution of the employed population by type of economic activity on average per year (persons)

Form Table 14.10. Economic activity of the population/labor force (people)

DATA COLLECTION METHODOLOGY

The labour force survey is conducted according to the methodology of the International Labour Organization (ILO) of in ac

The data are collected using forms. Table 14.10, and Table 14.1b, which are completed based on data from the Labor Fig.

Recommendations for filling out are presented in the document. Methodological explanations for filling out questionnaire

VALUES FOR THE PERIOD

Average for the period of

UNITS OF MEASUREMENT

Humant

PERIODICITY (FREQUENCY) OF DISTRIBUTION

Annually™

FREQUENCY OF COLLECTION

Annually™

THE SYSTEM OF CLASSIFIERS

- 1. Reporting periods
- 2. Types of economic activities (NACE Rev. 2)15
- 3. Floor®

4. CIS countries re

6. Age group ™

- Level of education
- 7. Employment status of

CIS countries

Description

URI

http://purl.cisstat.org/Common/vocabs/country/1.0 2

SOURC

The CIS Country Directory (hereinafter the Directory) is prepared on the basis of the Classifier of Countries of the World (hereinafter, the CIS CCM). The Directory is linked to the elements of the SDG reference area code list (SDMX). using the Interoperability Basis Platform. Links have been established with the elements of DBPedia, EU Vocabularies, GeoNames, OASIS GeoLang TC.

PURPOSE

The reference book is used in statistical data sets that contain data across the CIS countries.

STRUCTURE

The directory structurally consists of a list of elements, each of which includes blocks: identification, description and URI.

Identification includes a three-digit numeric code, two-letter and three-letter ISO codes, and the URI of the element. Description includes the names of the countries that are part of the CIS. URI includes related external URIs.

The Handbook also shows the composition of the CIS

The list of countries presented in the Directory includes countries (territories, regions) from Table 1 of the CIS CSM.

DESIGNATION

С

Directory in RDF for

COMMONWEALTH OF INDEPENDENT STATES

URI: http://purl.cisstat.org/Common/vocabs/country/1.0#CIS &

Digital code: 172

Equivalents:

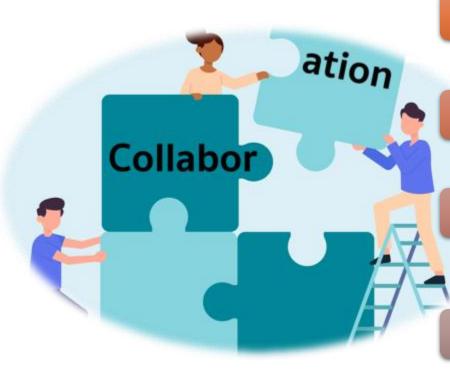
https://dbpedia.org/page/Commonwealth_of_Independent_States@,

https://purl.semanticip.org/linked-data/sdmx/code/area-172d,

https://purl.semanticip.org/linked-data/sdmx/code/area-R14rd

DISCUSSION. SDMX SEMANTIC GAPS







Missing links between related code lists

Code lists from different agencies (e.g., CL_AREA from IMF, ESTAT, UNSD, SDMX, UIS, IAEG-SDGs) are not semantically connected, despite overlaps

Lack of descriptions, purposes, and sources for code lists

 Code lists lack description, purpose, and source. Users cannot interpret code semantics without additional references

Limited glossary semantics

 XML-version of glossary only uses RELATED_TERMS, which does not reflect full relationships between terms

Inconsistent source references

• In the XML version of the SDMX Glossary, most concepts are linked to sdmx.org as the only source. In contrast, its HTML version points to concrete sources, including SDMX, SIMS, and DQAF (IMF) documents. It prevents correct RDF mapping without expert involvement

Semantic "loops" in concept definitions

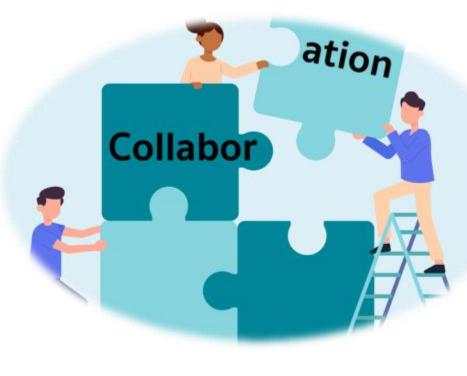
 Some concepts are defined through other concepts, requiring expert correction to resolve circular references

Unused glossary terms

• 83 of 229 glossary terms do not appear in any published SDMX documentation (v2.1, 3.0, 3.1)

IOBP FOR SDMX SEMANTIC TRANSFORMATION







Our proposal: enhancement and semantic transformation of SDMX in collaboration with SDMX developers for making SDMX Al-ready

Bridging semantic gaps — jointly transforming SDMX components into the Semantic Web environment, consolidating the efforts of international initiatives (EU SDMX Registry, ShowVoc, IoBP and others)

Alignment and enrichment — Expert collaboration to refine and enrich glossaries, code lists, and semantic links across agencies

Advancing glossary and code lists — enhancing definitions, applicability, and linking for practical use in LOSD

Iterative review of semantics — jointly assessing and improving RDF models to deliver reusable semantic assets

Automating generation and publishing of unified code lists, enriched glossary terms, and RDF/TTL assets with increasing semantic quality

COMMUNICATION

ANY QUESTIONS? We are open for cooperation cisstat@cisstat.org info@semanticip.org





INTERSTATE STATISTICAL COMMITTEE OF THE COMMONWEALTH OF INDEPENDENT STATES