URI Design and Mappings in StatSpace

Ba Lam Do

Contents

A. URIs patterns for metadata description

1. Patterns

Base URI: http://statspace.linkedwidgets.org

No	Patterns	Description
1	/metadata/{datasource-dataset}	URI of a metadata
2	/dataset/{datasource-dataset}	URI of a dataset
3	/codelist/cl_{name}	URI of a code list
4	/dimension/{name}	URI of a dimension

2. Example

No	Patterns	Description
1	http://statspace.linkedwidgets.org/metadata/ONS-Population-1851-2014	URI of a metadata
2	http://statspace.linkedwidgets.org/dataset/ONS-Population-1851-2014	URI of a dataset
3	http://statspace.linkedwidgets.org/codelist/cl_area	URI of a code list
4	http://statspace.linkedwidgets.org/dimension/activity	URI of dimension of economic activity

B. List of components used to create metadata

1. Dimensions

No	URI	Label
1	http://purl.org/linked-data/sdmx/2009/dimension#refArea	Reference Area
2	http://purl.org/linked-data/sdmx/2009/dimension#refPeriod	Reference Period
3	http://purl.org/linked-data/sdmx/2009/dimension#age	Age
4	http://purl.org/linked-data/sdmx/2009/dimension#educationLev	Education Level
5	http://purl.org/linked-data/sdmx/2009/dimension#occupation	Occupation
6	http://purl.org/linked-data/sdmx/2009/dimension#currency	Currency
7	http://purl.org/linked-data/sdmx/2009/dimension#civilStatus	Civil Status
8	http://purl.org/linked-data/sdmx/2009/dimension#freq	Frequency
9	http://purl.org/linked-data/sdmx/2009/dimension#sex	Sex
10	http://statspace.linkedwidgets.org/dimension/activity	Economic Activity

11	http://statspace.linkedwidgets.org/dimension/expenditure	Expenditure

2. Measure

No	URI	Label
1	http://purl.org/linked-data/sdmx/2009/measure#obsValue	Observation

3. Attribute

No	URI	Label
1	http://purl.org/linked-data/sdmx/2009/attribute#unitMeasure	Unit of Measure

C. URI patterns of code lists

Base URIs:

- http://reference.data.gov.uk for code list of reference period dimension
- http://statspace.linkedwidgets.org/codelist for the remaining code lists

1. Code list of reference area dimension (cl_area)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_area
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_area.ttl
- Pattern

URI Pattern	Description
/cl_area/{Country}/{Area in level 2}/…/{Area in level n}	URI of a geographical area

- Example

URI	Description
http://statspace.linkedwidgets.org/codelist/cl_area/Austria/Vienna	Vienna, Austria

2. Code list of reference period dimension (cl_period)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_period
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_period.ttl
- Patterns

URI Patterns	Description
/id/gregorian-year/{year}	URI for a year
/id/gregorian-half/{year}-{half}	URI for one-half year
/id/gregorian-quarter/{year}-{quarter}	URI for a quarter
/id/gregorian-month/{year}-{month}	URI for a month
/id/gregorian-day/{year}-{month}-{day}	URI for a day
/id/gregorian-hour/{year}-{month}-{day}T{hour}	URI for a hour
/id/gregorian-hour/{year}-{month}-{day}T{hour}:{min}	URI for a minute
/id/gregorian-hour/{year}-{month}-{day}T{hour}:{min}:{sec}	URI for a second
//id/gregorian-week/{year}-{week}	URI for a week
/id/gregorian-instant/{dateTime}	URI for an instant
/id/gregorian-interval/{dateTime}/{duration}	URI for a duration

- Example

URIs	Description
http://reference.data.gov.uk/id/gregorian-year/2016	URI for year 2016
http://reference.data.gov.uk/id/gregorian-day/2016-01-01	URI for day 01/01/2016

3. Code list of age dimension (cl_age)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_age
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_age.ttl
- Patterns

URI Patterns	Description
/cl_age/Y{n}, n=0, 1, 2,…,105	URI for an individual age
/cl_age/Y{n}T{n+4}, n=0, 5, 9,…,105	URI for an age group (5 years)
/cl_age/Y{n}T{n+9}, n=25, 35,…, 95	URI for an age group (10 years)
/cl_age/Y_GE_{n}, n=65, 70,…,90	URI for an age group (equal or above a specific age)
/cl_age/Y_LE_{n}, n=15, 20	URI for an age group (under a specific age)
/cl_age/TOTAL	URI for the top concept
/cl_age/UNK	URI for an unknown age

- Example

URIs	Description
http://statspace.linkedwidgets.org/codelist/cl_age/Y80	URI for age 80
http://statspace.linkedwidgets.org/codelist/cl_age/Y80T84	URI for an age group from 80 to 84

4. Code list of education level dimension (cl_educationLev)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_educationLev
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_educationLev.ttl
- URIs of values

URIs	Description
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L0	Pre-primary education
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L1	Primary education
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L2	Lower secondary
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L3	Upper secondary
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L4	Post-secondary non-tertiary education
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L5	Short-cycle tertiary education
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L6	Bachelor or equivalent
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L7	Master or equivalent
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L8	Doctoral or equivalent
http://statspace.linkedwidgets.org/codelist/cl_educationLev/L9	Not elsewhere classified

5. Code list of occupation dimension (cl_occupation)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_occupation
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_occupation.ttl
- Pattern

URI Pattern	Description
/cl_occupation/{code}	URI of an occupation

- Example

URIs	Description
http://linkedwidgets.org/resource/codelist/cl_occupation/OC1	URI for Managers
http://statspace.linkedwidgets.org/codelist/cl_occupation/OC11	URI for Chief executives, senior officials and legislators

6. Code list of currency dimension (cl_currency)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_currency
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_currency.ttl
- Pattern

URI Pattern	Description
/cl_currency/{ISO 4217 code}	URI for a currency

- Example

URI Pattern	Description
http://statspace.linkedwidgets.org/codelist/cl_currency/AED	URI for United Arab Emirates dirham
http://statspace.linkedwidgets.org/codelist/cl_currency/EUR	URI for Euro

7. Code list of civil status dimension (cl_civilStatus)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_civilStatus
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_civilStatus.ttl
- URIs of values

URIs	Description
http://statspace.linkedwidgets.org/codelist/cl_civilStatus/D	Divorced person
http://statspace.linkedwidgets.org/codelist/cl_civilStatus/E	Person whose registered partnership was legally dissolved
http://statspace.linkedwidgets.org/codelist/cl_civilStatus/L	Leggaly separated person
http://statspace.linkedwidgets.org/codelist/cl_civilStatus/M	Married person

http://statspace.linkedwidgets.org/codelist/cl_civilStatus/P	Person in Registerd partnership
http://statspace.linkedwidgets.org/codelist/cl_civilStatus/Q	Person whose registered partnership ended with the death of the partner
http://statspace.linkedwidgets.org/codelist/cl_civilStatus/S	Single person
http://statspace.linkedwidgets.org/codelist/cl_civilStatus/W	Widowed person

8. Code list of freq dimension (cl_freq)

- URI of code list: http://purl.org/linked-data/sdmx/2009/code#freq
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_freq.ttl
- URIs of values

URIs	Description
http://purl.org/linked-data/sdmx/2009/code#freq-H	Hourly
http://purl.org/linked-data/sdmx/2009/code#freq-D	Daily
http://purl.org/linked-data/sdmx/2009/code#freq-N	Minutely
http://purl.org/linked-data/sdmx/2009/code#freq-S	Half yearly, semester
http://purl.org/linked-data/sdmx/2009/code#freq-A	Annual
http://purl.org/linked-data/sdmx/2009/code#freq-Q	Quarterly
http://purl.org/linked-data/sdmx/2009/code#freq-M	Monthly
http://purl.org/linked-data/sdmx/2009/code#freq-B	Daily-business week
http://purl.org/linked-data/sdmx/2009/code#freq-W	Weekly

9. Code list of sex dimension (cl_sex)

- URI of code list: http://purl.org/linked-data/sdmx/2009/code#sex
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_sex.ttl
- URIs of values

URI Pattern	Description
http://purl.org/linked-data/sdmx/2009/code#sex-M	URI for Male gender
http://purl.org/linked-data/sdmx/2009/code#sex-F	URI for Female gender
http://purl.org/linked-data/sdmx/2009/code#sex-T	URI for Total
http://purl.org/linked-data/sdmx/2009/code#sex-U	URI for Unknown gender

http://purl.org/linked-data/sdmx/2009/code#sex-N	URI for not applicable gender

10. Code list of activity dimension (cl_activity)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_activity
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_activity.ttl
- Pattern

URI Pattern	Description
/cl_activity/{code}	URI of an economic activity

- Example

URI Pattern	Description
http://statspace.linkedwidgets.org/codelist/cl_activity/A	URI for activity of Agriculture, forestry, and fishing
http://statspace.linkedwidgets.org/codelist/cl_activity/A01	URI for activity of Crop and animal production, hunting and related service activities

11. Code list of expenditure dimension

- Contains 4 code lists
  - Classification of individual consumption by purpose (COICOP),
  - Classification of the functions of government (COFOG),
  - Classification of the purposes of non-profit institutions serving households (COPNI) and,
  - Classification of outlays of producers by purpose (COPP).

11.1. Code list of COICOP (cl_coicop)

- URI of code list COICOP: http://statspace.linkedwidgets.org/codelist/cl_coicop
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_coicop.ttl
- Pattern

URI Pattern	Description
/cl_coicop/{code}	URI for expenditure of an individual

- Example

URI	Description
http://statspace.linkedwidgets.org/codelist/cl_coicop/CP01	URI for expenditure of Food and non-alcoholic beverages

11.2. Code list of COFOG (cl_cofog)

- URI of code list COFOG: http://statspace.linkedwidgets.org/codelist/cl_cofog
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_cofog.ttl
- Pattern

URI Pattern	Description
/cl_cofog/{code}	URI for expenditure of government

- Example

URI	Description
http://statspace.linkedwidgets.org/codelist/cl_cofog/GF01	URI for expenditure of government for General public services

11.3. Code list of COPNI (cl_copni)

- URI of code list COPNI: http://statspace.linkedwidgets.org/codelist/cl_copni
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_copni.ttl
- Pattern

URI Pattern	Description
/cl_copni/{code}	URI for expenditure of non-profit organization

- Example

URI	Description
http://statspace.linkedwidgets.org/codelist/cl_copni/PN1	URI for expenditure of non-profit organization for Housing

11.4. Code list of COPP (cl_copp)

- URI of code list COPP: http://statspace.linkedwidgets.org/codelist/cl_copp
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_copp.ttl
- Pattern

URI Pattern	Description
/cl_copp/{code}	URI for expenditure of producer

- Example

URI	Description
http://statspace.linkedwidgets.org/codelist/cl_copp/PP1	URI for expenditure of producer for Outlays on infrastructure

12. Code list of unit of measure (cl_unitMeasure)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_unitMeasure
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_unitMeasure.ttl
- Pattern

URI Pattern	Description
/cl_unitMeasure/{unit}	URI of a unit

- Example

URI Pattern	Description
http://linkedwidgets.org/resource/codelist/cl_unitMeasure/P1	URI for unit "People"
http://statspace.linkedwidgets.org/codelist/cl_unitMeasure/TU	URI for unit "Twenty-Foot Equivalent"

13. Code list of subject (cl_subject)

- URI of code list: http://statspace.linkedwidgets.org/codelist/cl_subject
- Download RDF data at: http://statspace.linkedwidgets.org/code/cl_subject.ttl
- Pattern

URI Pattern	Description
/cl_subject/{code}	URI of a subject

- Example

URI Pattern	Description
http://statspace.linkedwidgets.org/codelist/subject/AG.SRF.TOTL.K2	URI for subject "Surface area (sq. km)"
http://statspace.linkedwidgets.org/codelist/subject//SP.POP.TOTL	URI for subject "Population, total"

D. Methods for identifying co-reference

1. Identifying co-reference URIs for dimensions

Input:
- o a URI and its label
- o e.g., ex:ref-Area, Reference area

Output:
- o co-reference URI of this URI
- o e.g., http://purl.org/linked-data/sdmx/2009/dimension#refArea

No	Dimension	Methods
1	Reference Area	- Define a set of possible names for representing this dimension. - Check the existence of one name in the input URI or label - {"ref-area", "refarea", "country", "refdistrict", "refstate", "place", "geocode", "region", "reference area"}
2	Reference Period	- {"ref-period", "ref-date","ref-year","refperiod", "timeperiod", "date", "year", "time-period", "time period"}
3	Age	- {"/age","_age", "#age","refage"}
4	Education Level	- {"educationlev", "edulev", "education level"};
5	Occupation	- {"occupation"}
6	Currency	- {"currency"}
7	Civil Status	- {"civil", "status"}
8	Frequency	- {"freq"}
9	Sex	- {"sex", "gender"}

10	Economic Activity	- {"activity" ,"economy"}
11	Expenditure	- cofog = {"funcofgov", "function of government", "functions of government"}; - coicop = {"indvcons", "individual consumption"}; - copp = {"outlayofproducer", "outlay of producer", "outlays of producer"}; - copni= {"purposeofnpi", "purpose of non-profit institution", "purposes of non-profit institution" };

2. Identifying co-reference URIs for values of a dimension

Input:
- o a URI and its label
- o e.g., ex:AT, Austria

Output:
- o co-reference URI of this URI
- o e.g., http://statspace.linkedwidgets.org/codelist/cl_area/Austria

No	Value of dimension	Methods
1	Reference Area	- Detect hierarchical relation among areas in the input dataset - Algorithm:presented in Semantics paper 2015
2	Reference Period	- Use Patterns e.g., Interval: [1-9][0-9]{3}-[1-9][0-9]{3} Year: [1-9][0-9]{3} Month:[1-9][0-9]{3}-[0-1][0-9] Quarter: [1-9][0-9]{3}-Q[1-4] Date: [1-9][0-9]{3}-[0-1][0-9]-[0-3][0-9]
3	Age	- Literal values e.g., {Value}^^http://www.w3.org/2001/XMLSchema#long => identify value, then builing corresponding URI - URIs e.g., ex:Y{Value1}-Y{Value2}, ex:{Value}%2B => identify age group, then building correpsonding URI
4	Education Level	- Compare label and code of the URI with values in the corresponding code list. - For example: if the URI ends with code "L0" or its label contains "Pre-primary education" => map to http://statspace.linkedwidgets.org/codelist/cl_educationLev/L0

5	Occupation	- Compare label and code of the URI with values in the corresponding code list. - For example: if the URI ends with code "OC11" or its label contains "Chief executives, senior officials and legislators => map to http://statspace.linkedwidgets.org/codelist/cl_occupation/OC11
6	Currency	- Compare label and code of the URI with values in the corresponding code list. - For example: if the URI ends with code "EUR" or its label contains "EURO" => map to http://statspace.linkedwidgets.org/codelist/cl_currency/EUR
7	Civil Status	- Compare label and code of the URI with values in the corresponding code list. - For example: if the URI ends with code "D" or its label contains "Divorced person" => map to http://statspace.linkedwidgets.org/codelist/cl_civilStatus/D
8	Frequency	- Compare label and code of the URI with values in the corresponding code list. - For example: if the URI ends with code "H" or its label contains "Hourly" => map to - http://purl.org/linked-data/sdmx/2009/code#freq-H
9	Sex	- Compare label and code of each URI with values in the corresponding code list. - For example: if the URI ends with code "M" or its label contains "Male" => map to http://purl.org/linked-data/sdmx/2009/code#sex-M
10	Economic Activity	- Compare label and code of each URI with values in the corresponding code list. - For example: if the URI ends with code "A" or its label contains "Agriculture, forestry, and fishing" => map to http://statspace.linkedwidgets.org/codelist/cl_activity/A
11	Expenditure	- Compare label and code of each URI with values in the corresponding code list. - For example: if the URI ends with code "GF01" or its label contains "General public services" => map to http://statspace.linkedwidgets.org/codelist/cl_cofog/GF01

3. Identifying co-reference URIs for unit of measure

Input:
- o a URI and its label
- o e.g., http://dd.eionet.europa.eu/vocabulary/eurostat/unit/1000PERS, 1000 persons

Output:
- o co-reference URI of this URI
- o e.g., http://statspace.linkedwidgets.org/codelist/cl_unitMeasure/P1.3

Method:
- o Step 1. Identify unit (i.e., P1)

  Use step of possible names for a unit to identify the co-reference URI for the input URI. For example, to detect unit "People", our set is as follows:

  {"people", "person", "worker", "population", "migration", "migrant", "labor", "births", "adults"}

- o Step 2. Indentify scale/power of unit (i.e., 3)

  Detect value in the input label, then identifying scale of this unit. For example, if a label contains "1000" or "1,000" or "1.000" or "thousand", returns value 3.