StatDCAT-AP

A Common Layer for the Exchange of Statistical Metadata in Open Data Portals

Makx Dekkers, Stefanos Kotoglou, Chris Nelson, Norbert Hohn, Marco Pellegrino, Vassilios Peristeras

Semstats 2016, October 18
What StatDCAT-AP is

Current status

Using StatDCAT-AP

StatDCAT-AP and SDMX

Future steps
The challenge: data silos

- The data landscape consists of many data silos:
  - Statistical data, geospatial data, legal data, research data, archival data, etc.
- Many of these silos build portals with metadata:
  - https://www.openaire.eu/ (research)
- These portals serve their goal for specific audiences, and the data set could be in a variety of formats, not necessarily RDF
- The metadata describing these datasets may be curated according to a variety of standards (DDI, SDMX, ISO 11179 etc.)
- **But**: No easy way to discover data across domains
A cross-domain standard

- Bringing together metadata from a multitude of domains in one 'general data portal' to expose domain-specific data
- Using a cross-domain description standard that is able to capture a core set of characteristics across domains: DCAT Application Profile for data portals in Europe
- Extending cross-domain standard with additional features of domain-specific data: GeoDCAT-AP, StatDCAT-AP
- Enabling creation of high-level index for the purpose of discovery across domains
- NB: Local systems and domain-specific portals continue to use specific standards: approach based on the export of metadata according to a cross-domain standard
What is StatDCAT-AP

- Application Profile of the **DCAT W3C Recommendation** for the exchange of descriptions of datasets between (open) data portals
- **DCAT-AP** was developed for specific use in Europe, among others to support the European Data Portal
- **StatDCAT-AP**: extension of DCAT-AP enabling cross-portal search for statistical data sets
- **Extend** DCAT-AP by adding:
  - Metadata elements from statistical standards (e.g. SDMX)
  - Recommendations for use of specific controlled vocabularies
What StatDCAT-AP is
Current status
Using StatDCAT-AP
StatDCAT-AP and SDMX
Future steps
Current status: public review

• Final draft of specification is available on Joinup:
  https://joinup.ec.europa.eu/node/152858

StatDCAT-AP - Draft 4

Description

Fourth editor’s draft of the StatDCAT-AP specification has been made available for public discussion.

This draft is available for public review until 23/10/2016. The following options exist:

• include your comments directly on this page; or
• create an issue using the Issue tracker; or
• contact us via the public mailing list: stat_dcat_application_profile@joinup.ec.europa.eu.

Themes

eGovernment

Distributions

StatDCAT-AP - Draft 4

Extended deadline 31 October 2016
StatDCAT-AP Working Group

• Chair: Eurostat and EU Publications Office

• Stakeholders
  The ISA and ISA² Programme of the European Commission, other Directorates General (DGs) of the European Commission, other European Union institutions, representatives of national institutes and international agencies, experts, representatives of consumers such as Digital Agenda Scoreboard, representatives of the European Data Portal.

• Meetings
  o Five meetings took place in 2016, including one face to face meeting. The next meeting will take place on 14 November 2016.
  o Presentations and minutes-discussions from the meetings are available on Joinup at https://joinup.ec.europa.eu/node/152858.
What StatDCAT-AP is
Current status
Using StatDCAT-AP
StatDCAT-AP and SDMX
Future steps
Using StatDCAT-AP in practice

- Many statistical datasets are of interest to the general data portals and their users. Using StatDCAT-AP helps general data portals to provide enhanced services for collections of statistical data.
- Statistical data providers (e.g. organisations, Member States) can increase the discoverability of their statistical datasets by including descriptions of the datasets in data portals.
- Statistical data users (e.g. national statistic officers) can explore, find, identify and select statistical datasets coming from different portals.
- StatDCAT-AP facilitates a better integration of existing statistical data portals with the open data portals, improving the discoverability of statistical datasets.
The European context

The European Data Portal

- Developed for European Commission DG CNECT
- Harvesting metadata from national data portals

http://www.europeandataportal.eu
What StatDCAT-AP is

Current status

Using StatDCAT-AP

StatDCAT-AP and SDMX

Future steps
Future steps

- Publication of final version after comments received during public review (end of 2016)
- Full support of StatDCAT-AP from EU and European Open Data Portals
- Piloting
- Building experiences
- Revising standard taking into account lessons learnt, quality aspects,...
Future work: Quality

• Quality aspects are very important for datasets in general and statistical datasets in particular
• Due to time and resource constraints, current version of StatDCAT-AP does not fully address the issue
  o Short-term: provide mechanism to link to existing quality information in StatDCAT-AP, version 1
  o Longer-term: consider integrated quality framework as basis for extensions to StatDCAT-AP, version 2
Quality, longer-term: SIMS

• Eurostat's Single Integrated Metadata Structure includes specific quality aspects:
  o e.g. Accessibility and clarity; Quality management; Relevance; Accuracy and reliability; Timeliness and punctuality; Coherence and comparability

• This set of aspects can form the basis for future extensions to StatDCAT-AP, or even to DCAT-AP
Get involved

Joinup: https://joinup.ec.europa.eu/asset/stat_dcat_application_profile/description

Visit ISA initiatives

mail@makxdekkers.com
stefanos.kotoglou@be.pwc.com
chris.nelson@metadatatechnology.com
marco.pellegrino@ec.europa.eu
norbert.hohn@publications.europa.eu
vassilios.peristeras@ec.europa.eu
The “Single Integrated Metadata Structure” (SIMS) of the European Statistical System

<table>
<thead>
<tr>
<th>SIMS</th>
<th>ESMS</th>
<th>ESQRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Contact</td>
<td>5.9</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Contact organisation</td>
<td>5.10</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Contact organisation unit</td>
<td>5.10.1</td>
</tr>
<tr>
<td>5.1.3</td>
<td>Contact name</td>
<td>5.10.2</td>
</tr>
<tr>
<td>5.1.4</td>
<td>Contact person function</td>
<td>5.10.3</td>
</tr>
<tr>
<td>5.1.5</td>
<td>Contact mail address</td>
<td>5.10.4</td>
</tr>
<tr>
<td>5.1.6</td>
<td>Contact email address</td>
<td>5.10.5</td>
</tr>
<tr>
<td>5.1.7</td>
<td>Contact phone number</td>
<td>5.10.6</td>
</tr>
<tr>
<td>5.1.8</td>
<td>Contact fax number</td>
<td>5.10.7</td>
</tr>
<tr>
<td>5.2</td>
<td>Metadata update</td>
<td>5.10.8</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Metadata last certified</td>
<td>5.10.9</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Metadata last posted</td>
<td>5.10.10</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Metadata ask update</td>
<td>5.10.11</td>
</tr>
<tr>
<td>5.3</td>
<td>Statistical presentation</td>
<td>5.10.12</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Data description</td>
<td>5.10.13</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Classification system</td>
<td>5.10.14</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Sector coverage</td>
<td>5.10.15</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Statistical concepts and definitions</td>
<td>5.10.16</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Statistical unit</td>
<td>5.11</td>
</tr>
<tr>
<td>5.3.6</td>
<td>Statistical population</td>
<td>5.11.1</td>
</tr>
<tr>
<td>5.3.7</td>
<td>Reference area</td>
<td>5.11.2</td>
</tr>
<tr>
<td>5.3.8</td>
<td>Time coverage</td>
<td>5.11.3</td>
</tr>
<tr>
<td>5.3.9</td>
<td>Base period</td>
<td>5.11.4</td>
</tr>
<tr>
<td>5.4</td>
<td>Unit of measure</td>
<td>5.12</td>
</tr>
<tr>
<td>5.5</td>
<td>Reference period</td>
<td>5.12.2</td>
</tr>
<tr>
<td>5.6</td>
<td>Institutional mandate</td>
<td>5.12.3</td>
</tr>
<tr>
<td>5.6.1</td>
<td>Legal acts and other agreements</td>
<td>5.12.4</td>
</tr>
<tr>
<td>5.6.2</td>
<td>Data sharing</td>
<td>5.12.5</td>
</tr>
<tr>
<td>5.7</td>
<td>Confidentiality</td>
<td>5.12.6</td>
</tr>
<tr>
<td>5.7.1</td>
<td>Confidentiality - policy</td>
<td>5.13</td>
</tr>
<tr>
<td>5.7.2</td>
<td>Confidentiality - data treatment</td>
<td>5.13.1</td>
</tr>
<tr>
<td>5.8</td>
<td>Release policy</td>
<td>5.13.2</td>
</tr>
<tr>
<td>5.8.1</td>
<td>Release calendar</td>
<td>5.13.3.1</td>
</tr>
<tr>
<td>5.8.2</td>
<td>Release calendar access</td>
<td>5.13.3.2</td>
</tr>
<tr>
<td>5.8.3</td>
<td>User access</td>
<td>5.13.3.3</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>SIMS</th>
<th>ESMS</th>
<th>ESQRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.14</td>
<td>Timeliness and punctuality</td>
<td>5.14.1</td>
</tr>
<tr>
<td>5.14.2</td>
<td>TP2. Time lag - final results for P</td>
<td></td>
</tr>
<tr>
<td>5.14.3</td>
<td>Punctuality and TP3. Punctuality - delivery and publication for U</td>
<td></td>
</tr>
<tr>
<td>5.14.4</td>
<td>TP3. Punctuality - delivery and publication for P</td>
<td></td>
</tr>
<tr>
<td>5.14.5</td>
<td>Coherence and comparability</td>
<td></td>
</tr>
<tr>
<td>5.14.6</td>
<td>Comparability - over time and CC2. Length of comparable time series for U</td>
<td></td>
</tr>
<tr>
<td>5.14.7</td>
<td>CC2. Length of comparable time series for P</td>
<td></td>
</tr>
<tr>
<td>5.14.8</td>
<td>Coherence - cross domain</td>
<td></td>
</tr>
<tr>
<td>5.14.9</td>
<td>CC1. Asymmetry for indicators - statistics - coefficient</td>
<td></td>
</tr>
<tr>
<td>5.15</td>
<td>Comparability - overall domain and national statistics</td>
<td></td>
</tr>
<tr>
<td>5.15.1</td>
<td>Coherence - internal</td>
<td></td>
</tr>
<tr>
<td>5.15.2</td>
<td>Cost and burden</td>
<td></td>
</tr>
<tr>
<td>5.47</td>
<td>Data revision</td>
<td></td>
</tr>
<tr>
<td>5.17.1</td>
<td>Data revision - policy</td>
<td></td>
</tr>
<tr>
<td>5.17.2</td>
<td>Data revision - practice and A8. Data revision - average size for U</td>
<td></td>
</tr>
<tr>
<td>5.17.3</td>
<td>A8. Data revision - average size for P</td>
<td></td>
</tr>
<tr>
<td>5.18</td>
<td>Statistical processing</td>
<td></td>
</tr>
<tr>
<td>5.19</td>
<td>Source data</td>
<td></td>
</tr>
<tr>
<td>5.19.1</td>
<td>Frequency of data collection</td>
<td></td>
</tr>
<tr>
<td>5.19.2</td>
<td>Data collection</td>
<td></td>
</tr>
<tr>
<td>5.19.3</td>
<td>Data validation</td>
<td></td>
</tr>
<tr>
<td>5.19.4</td>
<td>Data compilation</td>
<td></td>
</tr>
<tr>
<td>5.20</td>
<td>A7. Imputation - rate</td>
<td></td>
</tr>
<tr>
<td>5.20.1</td>
<td>Adjustment</td>
<td></td>
</tr>
<tr>
<td>5.20.2</td>
<td>Seasonal adjustment</td>
<td></td>
</tr>
<tr>
<td>5.20.3</td>
<td>Comment</td>
<td></td>
</tr>
</tbody>
</table>
EU Open Data Portal vs European Data Portal

Datasets of the EU institutions, agencies and other bodies (Parliament, Commission, ESTAT, JRC, Council, EEA..)

Datasets of national public institutions (regions, ministries, etc..)

- 71 catalogues (data portals and geo portals)
- 583,727 datasets
- Reuse apps
- Quality checker
- Trainings
- Studies
emigration

Show results with: • all of these words | • any of these words | • the exact phrase

Sort by: Relevance

6 datasets found

Emigration
Total number of long-term emigrants leaving from the reporting country during the reference year

Emigration by age and sex
Emigration by age and sex

Young emigrants by sex, age and country of birth
Young emigrants by sex, age and country of birth

Emigration by five year age group, sex and citizenship
Emigration by five year age group, sex and citizenship

Emigration by five year age group, sex and country of birth
Emigration by five year age group, sex and country of birth
SPARQL

You can search for the metadata stored in the EU Open Data Portal triple store by using the SPARQL endpoint query editor below.

**Namespaces**

```
PREFIX dcat: <http://www.w3.org/ns/dcat#>
PREFIX odp: <http://data.europa.eu/euodp/ontologies/ec-odp#>
PREFIX dc: <http://purl.org/dc/terms/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
```

**SPARQL Query**

```
select distinct ?g ?o where { graph ?g {?s dc:title ?o. filter regex(?o, 'Statistics', 'i') } } LIMIT 10
```

**Format**

- HTML

**Limit results**

- 10
Linked Open Data

SPARQL
You can search for the metadata stored in the EU Open Data Portal triple store by using the SPARQL endpoint query editor below.

Namespaces *
PREFIX dcat: <http://www.w3.org/ns/dcat#>
PREFIX odp: <http://data.europa.eu/euodp/ontologies/ec-odp#>
PREFIX dc: <http://purl.org/dc/terms/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>

SPARQL Query *
```
SELECT DISTINCT ?g ?o WHERE {
  GRAPH ?g {
    FILTER REGEX(?o, 'emigration', 'i')
  }
} LIMIT 10
```
<table>
<thead>
<tr>
<th>graphURL</th>
<th>title</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset/zQallpWgMvYcWwVJYX0mWcQ">http://data.europa.eu/euodp/data/dataset/zQallpWgMvYcWwVJYX0mWcQ</a></td>
<td>&quot;Emigration by five year age group, sex and country of birth&quot;@en</td>
</tr>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset/zQallpWgMvYcWwVJYX0mWcQ">http://data.europa.eu/euodp/data/dataset/zQallpWgMvYcWwVJYX0mWcQ</a></td>
<td>&quot;Emigration par tranche d'âges de cinq ans, sexe et pays de naissance&quot;@fr</td>
</tr>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset/sqEoF5vRqOcXP8vMgn0A">http://data.europa.eu/euodp/data/dataset/sqEoF5vRqOcXP8vMgn0A</a></td>
<td>&quot;Immigration by five year age group, sex and citizenship&quot;@en</td>
</tr>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset/sqEoF5vRqOcXP8vMgn0A">http://data.europa.eu/euodp/data/dataset/sqEoF5vRqOcXP8vMgn0A</a></td>
<td>&quot;Immigration par tranche d'âges de cinq ans, sexe et nationalité&quot;@fr</td>
</tr>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset/LCyr4acr5ORc5FXixOg">http://data.europa.eu/euodp/data/dataset/LCyr4acr5ORc5FXixOg</a></td>
<td>&quot;Population par période d'immigration dans le pays, pays de naissance et région NUTS 2&quot;@fr</td>
</tr>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset">http://data.europa.eu/euodp/data/dataset</a> Pgfs44171ZVFDqgZRIKQw</td>
<td>&quot;Methods to find current job by migration status, educational attainment level and type of contract&quot;@en</td>
</tr>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset">http://data.europa.eu/euodp/data/dataset</a> Pgfs44171ZVFDqgZRIKQw</td>
<td>&quot;Methoden, mit deren Hilfe die derzeitige Stelle gefunden wurde, nach Migrationsstatus, Bildungsabschluss und Art des Arbeitsvertrags&quot;@de</td>
</tr>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset/AxGHJUlhLfcnerZCxx8LiQ">http://data.europa.eu/euodp/data/dataset/AxGHJUlhLfcnerZCxx8LiQ</a></td>
<td>&quot;Employees by migration status, educational attainment level, occupation and working time&quot;@en</td>
</tr>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset/AxGHJUlhLfcnerZCxx8LiQ">http://data.europa.eu/euodp/data/dataset/AxGHJUlhLfcnerZCxx8LiQ</a></td>
<td>&quot;Arbeitnehmer nach Migrationsstatus, Bildungsabschluss, Beruf und Arbeitszeit&quot;@de</td>
</tr>
<tr>
<td><a href="http://data.europa.eu/euodp/data/dataset/Sqek9">http://data.europa.eu/euodp/data/dataset/Sqek9</a></td>
<td>&quot;First generation of immigrants by sex, citizenship, duration and reason for migration&quot;@en</td>
</tr>
</tbody>
</table>
Emigration by five year age group, sex and country of birth

**Publisher**

*Eurostat*

**Description**

Emigration by five year age group, sex and country of birth

**eurovoc domains**

Social questions

**Resources**

- [Download dataset in TSV format](#)
- [Download dataset in TSV format (unzipped)](#)
- [Download dataset in SDMX-ML format](#)

**Documentation**

- [ESMS metadata (Euro-SDMX Metadata structure) HTML](#)
- [ESMS metadata (Euro-SDMX Metadata structure) SDMX](#)
- [More information on Eurostat Website](#)

**URI**

http://ec.europa.eu/eurostat/web/products-datasets/-/migr_emi4ctb
Emigration by five year age group, sex and country of birth

View table
Download table
Show table location in data tree
Metadata

Additional information
Code: migr_emi4ctb
Last update: 27/09/16
Oldest data: 2002
Most recent data: 2014
Number of values: 2306069

Themes: Population and social conditions
A catalogue contains one or more datasets

A dataset has one or more distributions
## 5 star-schema of Linked Open Data

| ★ | Make your stuff available on the Web (whatever format) under an open license |
| ★★ | Make it available as structured data (e.g., Excel instead of image scan of a table) |
| ★★★ | Use non-proprietary formats (e.g., CSV instead of Excel) |
| ★★★★ | Use URIs to denote things, so that people can point at your stuff |
| ★★★★★ | Link your data to other data to provide context |
18-year-olds in education

Participation rates, all levels (%)

This indicator gives the percentage of all 18-year-olds who are still in any kind of school (all ISCED levels). It gives an indication of the number of young people who have not abandoned their efforts to improve their skills through initial education and it includes both those who had a regular education career without any delays as well as those who are continuing even if they had to repeat some steps in the past.

View table

Download table

Show table location in data tree

Metadata

Additional information

Code: tps00060
Last update: 22/04/15
Oldest data: 2001
Most recent data: 2012
Number of values: 447

Themes: Population and social conditions
DCAT-AP (core) model
## Dataset

<table>
<thead>
<tr>
<th>Mandatory</th>
<th>Recommended</th>
<th>Optional</th>
</tr>
</thead>
<tbody>
<tr>
<td>dct:description</td>
<td>dcat:contactPoint</td>
<td>adms:identifier</td>
</tr>
<tr>
<td>dct:title</td>
<td>dcat:distribution</td>
<td>adms:sample</td>
</tr>
<tr>
<td></td>
<td>dcat:keyword</td>
<td>adms:versionNotes</td>
</tr>
<tr>
<td></td>
<td>dcat:publisher</td>
<td>dcat:landingPage</td>
</tr>
<tr>
<td></td>
<td>dcat:theme</td>
<td>dcat:accessRights</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:accrualPeriodicity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:hasVersion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:isVersionOf</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:identifier</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:issued</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:language</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:modified</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:provenance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:relation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:source</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:spatial</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:temporal</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:type</td>
</tr>
<tr>
<td></td>
<td></td>
<td>foaf:page</td>
</tr>
<tr>
<td></td>
<td></td>
<td>owl:versionInfo</td>
</tr>
</tbody>
</table>

### StatDCAT-AP to add optional properties:

- dqv:hasQualityAnnotation
- stat:attribute
- stat:dimension
- stat:numSeries
- stat:unitMeasure
Emigration by five year age group, sex and country of birth

**Publisher**

*Eurostat*

**Description**

Emigration by five year age group, sex and country of birth

**Dimensions**

Country of birth, sex, 5 year age groups

**Eurovoc domains**

Social questions

**Resources**

- [Download dataset in TSV format](#)
- [Download dataset in TSV format (unzipped)](#)
- [Download dataset in SDMX-ML format](#)

**Documentation**

- [ESMS metadata (Euro-SDMX Metadata structure)](#)
- [ESMS metadata (Euro-SDMX Metadata structure) SDMX](#)
- [More information on Eurostat Website](#)

**URI**

http://ec.europa.eu/eurostat/web/products-datasets/-/migr_emi4ctb

**Status**

Completed
## Distribution

<table>
<thead>
<tr>
<th>Mandatory</th>
<th>Recommended</th>
<th>Optional</th>
</tr>
</thead>
<tbody>
<tr>
<td>dcat:accessURL</td>
<td>dct:description</td>
<td>adms:status</td>
</tr>
<tr>
<td></td>
<td>dct:format</td>
<td>dcat:byteSize</td>
</tr>
<tr>
<td></td>
<td>dct:license</td>
<td>dcat:downloadURL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dcat:mediaType</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dct:conformsTo</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dct:issued</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dct:language</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dct:modified</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dct:rights</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dct:title</td>
</tr>
<tr>
<td></td>
<td></td>
<td>foaf:page</td>
</tr>
<tr>
<td></td>
<td></td>
<td>spdx:checksum</td>
</tr>
</tbody>
</table>

StatDCAT-AP to add optional property:

- **dct:type**
What StatDCAT-AP is

Current status

Using StatDCAT-AP

StatDCAT-AP and SDMX

Future steps
Use case: StatDCAT-AP ‘users’

- Search / discovery of data existence
- Process-oriented
- Ad-hoc

Definition of StatDCAT-AP

Evaluation of STAT DCAT-AP
SDMX Information Model: Schematic View

- **Data Flow**
  - **Data Providers**
  - **Data Structure Definition**
    - **Concepts**
    - **Codes**
  - **Codelists**
  - **Registered Data Source**
    - **Data Sources and Indexed Content**
      - **Concepts and Coding Schemes**
        - **Concept Schemes**
        - **Category Schemes**
          - **Concepts**
          - **Coding Schemes**
            - **Codelists**
  - **Content Constraint**
  - **Provision Agreement**
  - **Publishers**
  - **Data Providers Scheme**
Publishing StatDCAT-AP from SDMX: Requirements

Ability to

• Combine metadata from a variety of sources
  o SDMX registry
  o Excel or CSV files
  o Metadata Repository

• Validate the metadata
  o Mandatory/Conditional
  o Representation (URL, text, code)
  o Multiple or single occurrence
  o Hierarchy

• Output StatDCAT-AP RDF

• Submit Catalogue metadata to the portal
StatDCAT-AP
A New Dawn in Statistical Data Discovery