# 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**

ISA

What StatDCAT-AP is

Current status

Using StatDCAT-AP

StatDCAT-AP and SDMX

Future steps

ISA

# 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:
  - http://ec.europa.eu/eurostat/data/database, http://stats.oecd.org (stats)
  - http://inspire-geoportal.ec.europa.eu/ (geo)
  - 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

ISA

# 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

*ISA*

# 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

ISA

# Current status: public review

- Final draft of specification is available on Joinup:

  https://joinup.ec.europa.eu/node/152858

## StatDCAT-AP - Draft 4

( ⚐ ★★★★★ )   5/5 | 1 votes |

### 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 | ⬇ PDF |
|---|---|

*Extended deadline 31 October 2016*

7

ISA

# 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
  - Five meetings took place in 2016, including one face to face meeting. The next meeting will take place on 14 November 2016.
  - Presentations and minutes-discussions from the meetings are available on Joinup at https://joinup.ec.europa.eu/node/152858.
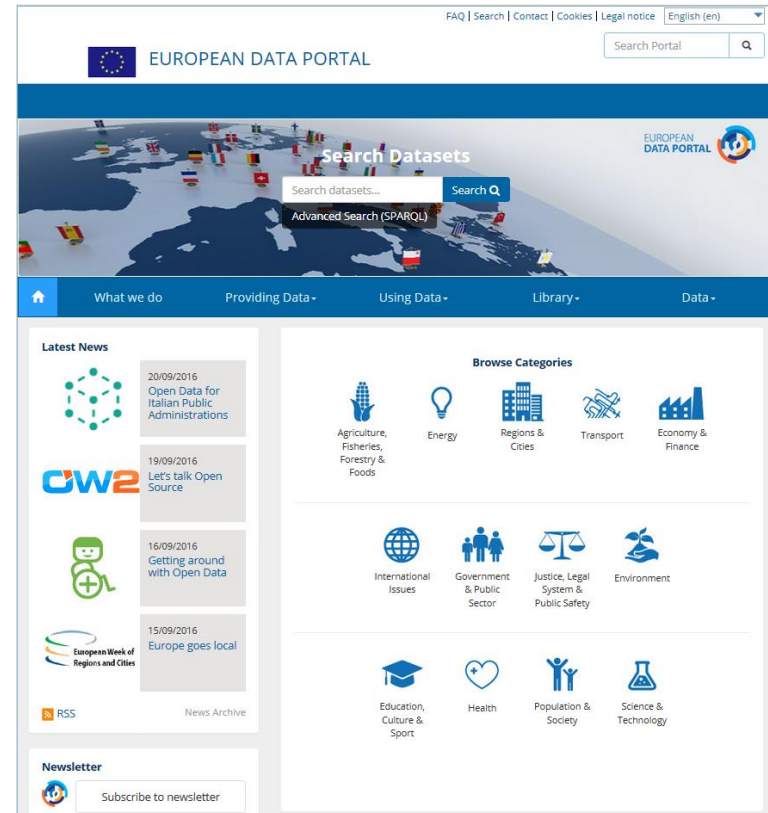
ISA

# 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.

*ISA*

# The European context



The European Data Portal

- Developed for European Commission DG CNECT

- Harvesting metadata from national data portals



http://www.europeandataportal.eu

EU Open Data Portal
data.europa.eu/euodp/

European Union Open Data Portal

EUROPA > Open Data Portal > Data > Search

**Data** | **Applications** | **Linked Data** | **Developers' corner** | **About**   🔒 *Data provider's area*

🔄 Share

migration 🔍

Show results with: ⦿ all of these words | ○ any of these words | ○ these phrase ❓  **VIDEO**

Sort by: Relevance ▼    **165** datasets found 🔊

**Suggest a dataset**
Is there data you would like to find on the portal?

*Make a suggestion>>*

**EuroVoc domains**

Social questions (94)
European Union (12)
Law (2)
Geography (2)
International relations (1)
Employment and working conditions (1)
Economics (1)
Agriculture, forestry and fisheries (1)

📊 **Health workforce migration**
Health workforce migration
ZIP  TEXT/TAB-SEPARATED-VALUES  HTML  (139 views)

📊 **Reduced migration variant – International net migration by age and sex**
Reduced migration variant – International net migration by age and sex
ZIP  TEXT/TAB-SEPARATED-VALUES  HTML  (118 views)

📊 **Migration and skills in Georgia**
The report describes the findings for the European Training Foundation migration and skills survey conducted in Georgia. Following independence in 1991, Georgia faced a series of...
PDF  (356 views)

📊 **Migration and skills in Armenia and Georgia**
The European Training Foundation developed surveys to investigate the relationship between migration, development and skills in three countries: Armenia, Georgia and Morocco. This report...
PDF  (391 views) (1 Downloads)

📊 **Crude rate of net migration plus adjustment**
The indicator is defined as the ratio of net migration (including statistical adjustment) during the year to the average population in that year. The value is expressed per 1000 persons....
ZIP  TEXT/TAB-SEPARATED-VALUES  HTML  (265 views)

**Publishers**

Eurostat (104)
Frontex (45)
Directorate-General for Communication (8)
European Asylum Support Office (3)
European Training Foundation (2)
European Parliament (1)
Directorate-General for Regional and Urban Policy (1)
Directorate-General for Health and Food

12

What StatDCAT-AP is

Current status

Using StatDCAT-AP

StatDCAT-AP and SDMX

Future steps

ISA

# 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,…

ISA

# 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
  - Short-term: provide mechanism to link to existing quality information in StatDCAT-AP, version 1
  - Longer-term: consider integrated quality framework as basis for extensions to StatDCAT-AP, version 2

ISA

# Quality, longer-term: SIMS

- Eurostat's Single Integrated Metadata Structure includes specific quality aspects:
  - 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

16

ISA

# Get involved

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

# Visit ISA initiatives

**ADMS**
ASSET
DESCRIPTION
METADATA
SCHEMA

**StatDCAT-AP**
FOR
STATISTICAL
DATASETS

**GeoDCAT-AP**
FOR
GEOSPATIAL
DATASETS

**DCAT-AP**
FOR
DATA PORTALS
IN EUROPE

**CORE**
PUBLIC
ORGANISATION
VOCABULARY

**CORE**
PERSON
VOCABULARY

**REGISTERED**
ORGANISATION
VOCABULARY

**CORE**
CRITERION &
EVIDENCE
VOCABULARY

**CORE**
LOCATION
VOCABULARY

**CORE**
PUBLIC
SERVICE
VOCABULARY

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 1**

| | SIMS | |
|---|---|---|
| | ESMS | ESQRS |
| S.1 | Contact | |
| S.1.1 | Contact organisation | |
| S.1.2 | Contact organisation unit | |
| S.1.3 | Contact name | |
| S.1.4 | Contact person function | |
| S.1.5 | Contact mail address | |
| S.1.6 | Contact email address | |
| S.1.7 | Contact phone number | |
| S.1.8 | Contact fax number | |
| S.2 | Metadata update | |
| S.2.1 | Metadata last certified | |
| S.2.2 | Metadata last posted | |
| S.2.3 | Metadata last update | |
| S.3 | Statistical presentation | |
| S.3.1 | Data description | |
| S.3.2 | Classification system | |
| S.3.3 | Sector coverage | |
| S.3.4 | Statistical concepts and definitions | |
| S.3.5 | Statistical unit | |
| S.3.6 | Statistical population | |
| S.3.7 | Reference area | |
| S.3.8 | Time coverage | |
| S.3.9 | Base period | |
| S.4 | Unit of measure | |
| S.5 | Reference period | |
| S.6 | Institutional mandate | |
| S.6.1 | Legal acts and other agreements | |
| S.6.2 | Data sharing | |
| S.7 | Confidentiality | |
| S.7.1 | Confidentiality - policy | |
| S.7.2 | Confidentiality - data treatment | |
| S.8 | Release policy | |
| S.8.1 | Release calendar | |
| S.8.2 | Release calendar access | |
| S.8.3 | User access | |

**Table 2**

| | SIMS | |
|---|---|---|
| | ESMS | ESQRS |
| S.9 | Frequency of dissemination | |
| S.10 | Accessibility and clarity | |
| S.10.1 | News release | |
| S.10.2 | Publications | |
| S.10.3 | On-line database | |
| S.10.3.1 | AC1. Data tables - consultations | |
| S.10.4 | Micro-data access | |
| S.10.5 | Other | |
| S.10.5.1 | AC 2. Metadata - consultations | |
| S.10.6 | Documentation on methodology | |
| S.10.6.1 | AC 3. Metadata completeness - rate | |
| S.10.7 | Quality documentation | |
| S.11 | Quality management | |
| S.11.1 | Quality assurance | |
| S.11.2 | Quality assessment | |
| S.12 | Relevance | |
| S.12.1 | User needs | |
| S.12.2 | User satisfaction | |
| S.12.3 | Completeness and R1. Data completeness - rate for U | |
| S.12.3.1 | R1. Data completeness - rate for P | |
| S.13 | Accuracy and reliability | |
| S.13.1 | Overall accuracy | |
| S.13.2 | Sampling error and A1. Sampling errors - indicators for U | |
| S.13.2.1 | A1. Sampling errors - indicators for P | |
| S.13.3 | Non-sampling error and A4. Unit non-response - rate for U and A5. Item non-response - rate for U | |
| S.13.3.1 | Coverage error | |
| S.13.3.1.1 | A2. Over-coverage - rate | |
| S.13.3.1.2 | A3. Common units - proportion | |
| S.13.3.2 | Measurement error | |
| S.13.3.3 | Non response error | |
| S.13.3.3.1 | A4. Unit non-response - rate for P | |
| S.13.3.3.2 | A5. Item non-response - rate for P | |
| S.13.3.4 | Processing error | |
| S.13.3.5 | Model assumption error | |

**Table 3**

| | SIMS | |
|---|---|---|
| | ESMS | ESQRS |
| S.14 | Timeliness and punctuality | |
| S.14.1 | Timeliness and TP2. Time lag - final results for U | |
| S.14.1.1 | TP1. Time lag - first results for P | |
| S.14.1.2 | TP2. Time lag - final results for P | |
| S.14.2 | Punctuality and TP3. Punctuality - delivery and publication for U | |
| S.14.2.1 | TP3. Punctuality - delivery and publication for P | |
| S.15 | Coherence and comparability | |
| S.15.1 | Comparability - geographical | |
| S.15.1.1 | CC1. Asymmetry for mirror flows statistics - coefficient | |
| S.15.2 | Comparability - over time and CC2. Length of comparable time series for U | |
| S.15.2.1 | CC2. Length of comparable time series for P | |
| S.15.3 | Coherence- cross domain | |
| S.15.3.1 | Coherence - sub annual and annual statistics | |
| S.15.3.2 | Coherence- National Accounts | |
| S.15.4 | Coherence - internal | |
| S.16 | Cost and burden | |
| S.17 | Data revision | |
| S.17.1 | Data revision - policy | |
| S.17.2 | Data revision - practice and A6. Data revision - average size for U | |
| S.17.2.1 | A6. Data revision - average size for P | |
| S.18 | Statistical processing | |
| S.18.1 | Source data | |
| S.18.2 | Frequency of data collection | |
| S.18.3 | Data collection | |
| S.18.4 | Data validation | |
| S.18.5 | Data compilation | |
| S.18.5.1 | A7. Imputation - rate | |
| S.18.6 | Adjustment | |
| S.18.6.1 | Seasonal adjustment | |
| S.19 | Comment | |

# European Union Open Data Portal

**Data** | **Applications** | **Linked Data** | **Developers' corner** | **About** 🔒 *Data provider's area*

🔁 Share

emigration 🔍

**Show results with**: ⦿ all of these words | ○ any of these words | ○ the exact phrase ⑦

Total datasets available: 9069 📶

## Suggest a dataset

Is there data you would like to find on the portal?

*Make a suggestion>>*

## Most viewed datasets                          �III *view all »*

- III **DGT–Translation Memory**
  (14665 views)

- III **[LATEST VERSION] Elevation map of Europe**
  (8917 views)

- III **CORDIS – EU research projects under FP7 (2007–2013)**
  (7774 views)

- III **Tenders Electronic Daily (TED) – public procurement notices from the EU and beyond**
  (7686 views)

## Browse datasets by subject or groups
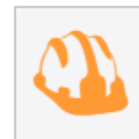
**Employment and working conditions**

**Social questions**

**Economics**

**Finance**

**Trade**

**Industry**

**Education and communications**

**Science**

# 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

**European Union Open Data Portal**

**Data** | **Applications** | **Linked Data** | **Developers' corner** | **About**

emigration

**Show results with**: ● all of these words | ○ any of these words | ○ the exact phrase ?

Sort by: Relevance ▼                          **6** datasets found
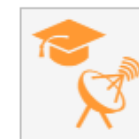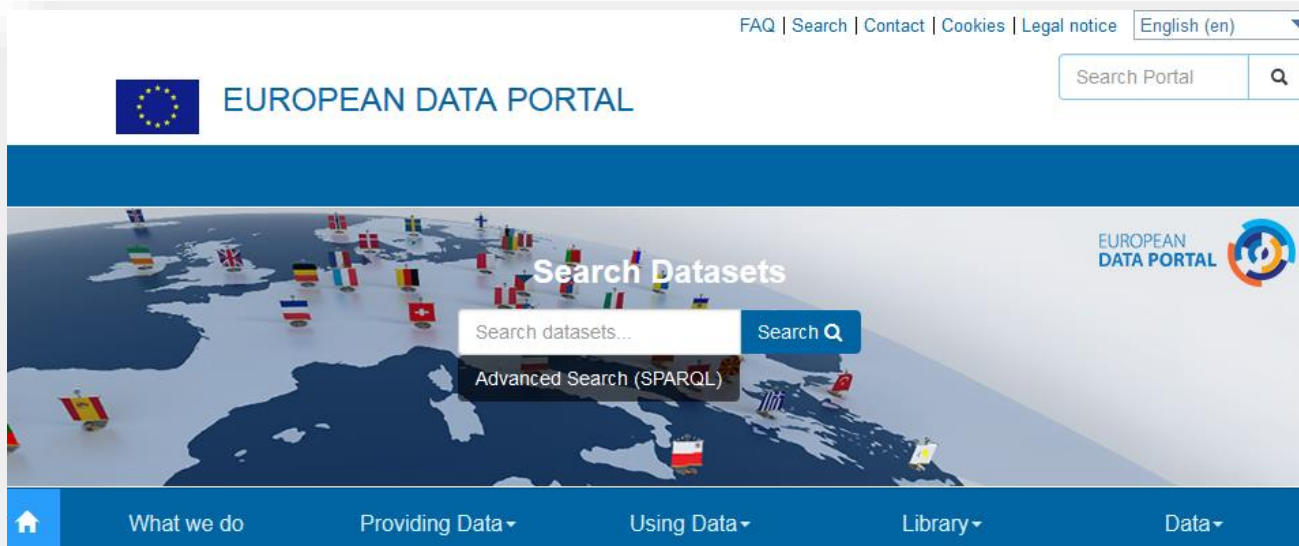
**Emigration**
Total number of long-term emigrantsleaving from the reporting country during the reference year
ZIP | TEXT/TAB-SEPARATED-VALUES   (153 views)

Visualise this dataset with   Graph ⌕ ,  Map ⌕ ,  Table ⌕

**Emigration by age and sex**
Emigration by age and sex
ZIP | TEXT/TAB-SEPARATED-VALUES | HTML   (225 views)

**Young emigrants by sex, age and country of birth**
Young emigrants by sex, age and country of birth
ZIP | TEXT/TAB-SEPARATED-VALUES   (229 views)
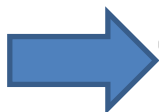
**Emigration by five year age group, sex and citizenship**
Emigration by five year age group, sex and citizenship
ZIP | TEXT/TAB-SEPARATED-VALUES | HTML   (235 views)

**Emigration by five year age group, sex and country of birth**
Emigration by five year age group, sex and country of birth
ZIP | TEXT/TAB-SEPARATED-VALUES | HTML   (112 views)

22

# European Union Open Data Portal

EUROPA > Open Data Portal > Data > Publisher > Eurostat > Health workforce migration

**Data** | **Applications** | **Linked Data** | **Developers' corner** | **About**

## 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 ▼

**Run Query**     **Get query URL**

23

# Linked Open Data
## http://data.europa.eu/euodp/en/linked-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 {?s dc:title ?o. filter regex(?o,
'emigration', 'i') } }  LIMIT 10
```
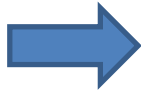
**Format** *

HTML ▼

**Format** *

HTML ▼

HTML
Spreadsheet
XML
JSON
Javascript
NTriples
RDF/XML
CSV

**Limit results** *

10 ▼

**Run Quer**      **et query URL**

24

| graphURL | title |
|---|---|
| http://data.europa.eu/euodp/data/dataset/zQallpwGmYClwVJVX0nWcQ | "Emigration by five year age group, sex and country of birth"@en |
| http://data.europa.eu/euodp/data/dataset/zQallpwGmYClwVJVX0nWcQ | "Emigration par tranche d'âges de cinq ans, sexe et pays de naissance"@fr |
| http://data.europa.eu/euodp/data/dataset/B6PcO66PJmNRJmKybasPiQ | "Emigration"@fr |
| http://data.europa.eu/euodp/data/dataset/B6PcO66PJmNRJmKybasPiQ | "Emigration"@en |
| http://data.europa.eu/euodp/data/dataset/sqeEoF5vRq0cXPBvMgr0A | "Immigration by five year age group, sex and citizenship"@en |
| http://data.europa.eu/euodp/data/dataset/sqeEoF5vRq0cXPBvMgr0A | "Immigration par tranche d'âges de cinq ans, sexe et nationalité"@fr |
| http://data.europa.eu/euodp/data/dataset/R6vj4arsBORRcjFXixiOg | "Population par période d'immigration dans le pays, pays de naissance et région NUTS 2"@fr |
| http://data.europa.eu/euodp/data/dataset/Pgfs44171ZIVFdqJZRIGKw | "Methods to find current job by migration status, educational attainment level and type of contract"@en |
| http://data.europa.eu/euodp/data/dataset/Pgfs44171ZIVFdqJZRIGKw | "Methoden, mit deren Hilfe die derzeitige Stelle gefunden wurde, nach Migrationsstatus, Bildungsabschluss und Art des Arbeitsvertrags"@de |
| http://data.europa.eu/euodp/data/dataset/AxGHJUIhLFcnerZCxzBlQ | "Employees by migration status, educational attainment level, occupation and working time"@en |
| http://data.europa.eu/euodp/data/dataset/AxGHJUIhLFcnerZCxzBlQ | "Arbeitnehmer nach Migrationsstatus, Bildungsabschluss, Beruf und Arbeitszeit"@de |
| http://data.europa.eu/euodp/data/dataset/Sqek9 | "First generation of immigrants by sex, citizenship, duration and reason for migration"@en |

# 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 | Download dataset in TSV format  `ZIP` |
| ⬇ DOWNLOAD | Download dataset in TSV format (unzipped)  `TEXT/TAB-SEPARATED-VALUES` |
| ⬇ DOWNLOAD | Download dataset in SDMX-ML format  `ZIP` |

## Documentation

| | |
|---|---|
| ↗ VISIT PAGE | ESMS metadata (Euro-SDMX Metadata structure) HTML  `HTML` |
| ⬇ DOWNLOAD | ESMS metadata (Euro-SDMX Metadata structure) SDMX |
| ⬇ DOWNLOAD | More information on Eurostat Website |

*URI*

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

*Status*

26

# eurostat
## Your key to European statistics

Type a keyword, a code, a title...

**European Commission** > Eurostat > Products Datasets

| News | Data | Publications | About Eurostat | Help |
|------|------|--------------|----------------|------|

DATASET DETAILS

# 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

**News**
News releases
Release calendar

**Data**
Database
Statistics by theme
Statistics A to Z

**Publications**
Statistics Explained
Recent publications

**About us**
Overview
How to find us
Contact

**Opportunities**
Calls for tenders
Grants
Share

European Commission | eurostat

27

# DCAT main entities



Catalogue

Dataset
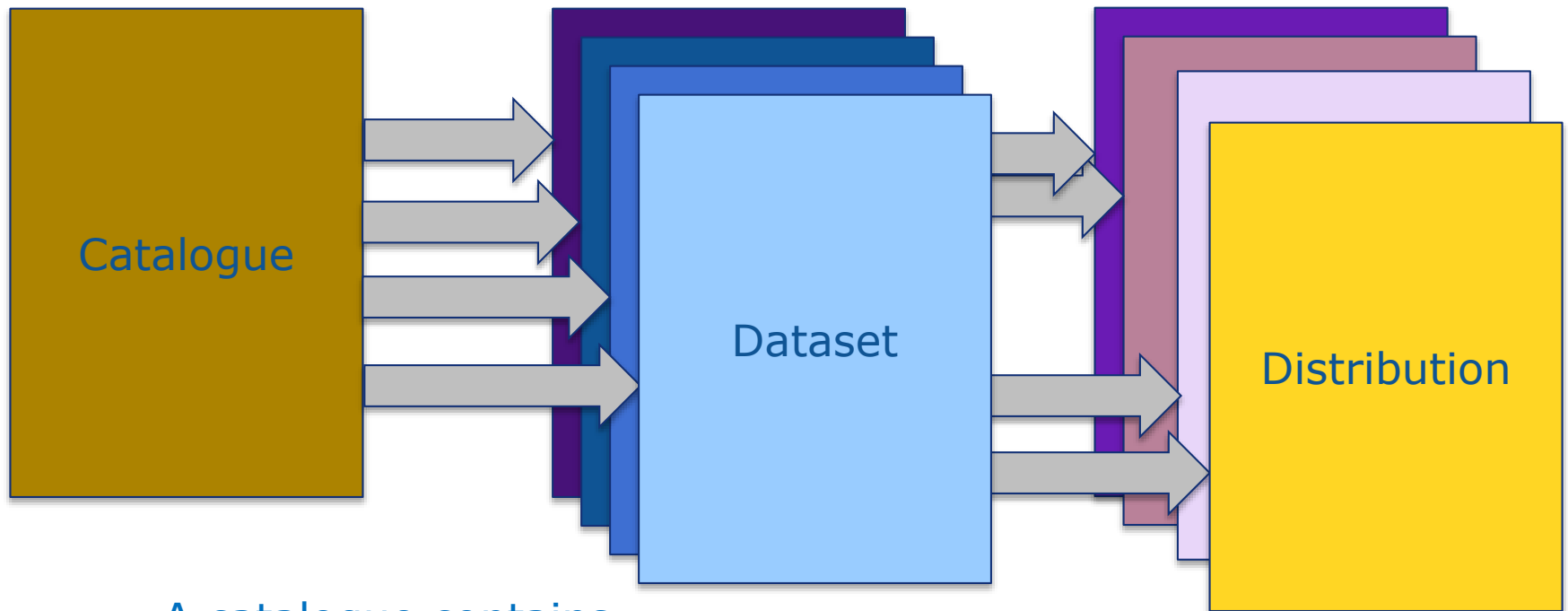
Distribution

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 |

# eurostat
## Your key to European statistics

Legal notice | RSS | Cookies | Links | Contact        English

Type a keyword, a code, a title...

**European Commission** > Eurostat > Products Datasets

| News | Data | Publications | About Eurostat | Help |
|---|---|---|---|---|

**DATASET DETAILS**

# 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                          Themes: Population and social conditions
Last update: 22/04/15
Oldest data: 2001
Most recent data: 2012
Number of values: 447

**News**
News releases
Release calendar

**Data**
Database
Statistics by theme
Statistics A to Z

**Publications**
Statistics Explained
Recent publications

**About us**
Overview
How to find us
Contact

**Opportunities**
Calls for tenders
Grants

European Commission | eurostat

# 18-year-olds in education

Dct:title

**Publisher**

Dct:publisher

*Eurostat »*

Licence:
**Legal Notice**

**Description**

Dct:description

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.

**Keywords**
education
student

Dcat:keyword

Dcat:keyword

**EuroVoc domains**

Dcat:theme

Education and communications, Employment and working conditions

**EuroVoc concepts**
education
vocational training

**Resources**

Dcat:distribution

**Catalogue record**
open-data.europa.eu
2014-12-22
Updated on open-data.europa.eu
2015-05-18

Views: 486

| | | |
|---|---|---|
| ⬇ DOWNLOAD | Download dataset in SDMX-ML format | ZIP |
| ⬇ DOWNLOAD | Download dataset in TSV format | ZIP |

**Documentation**

dqv:hasQuality Annotation

Suggest a dataset

Is there a dataset from the EU that you could not find in this portal?

*Please request the dataset >>*

| | | |
|---|---|---|
| ⬈ VISIT PAGE | ESMS metadata (Euro-SDMX Metadata structure) HTML | |
| ⬇ DOWNLOAD | ESMS metadata (Euro-SDMX Metadata structure) SDMX | |
| ⬇ DOWNLOAD | More information on Eurostat Website | |

*URI*
http://ec.europa.eu/eurostat/web/products-datasets/-/tps00060

*Status*
Completed

*Identifier*
tps00060

Dct:identifier

*Type of Dataset*
Statistical

*Modified Date*
2015-04-22

Dct:modified

*Temporal Coverage From*
2001

Dct:temporal

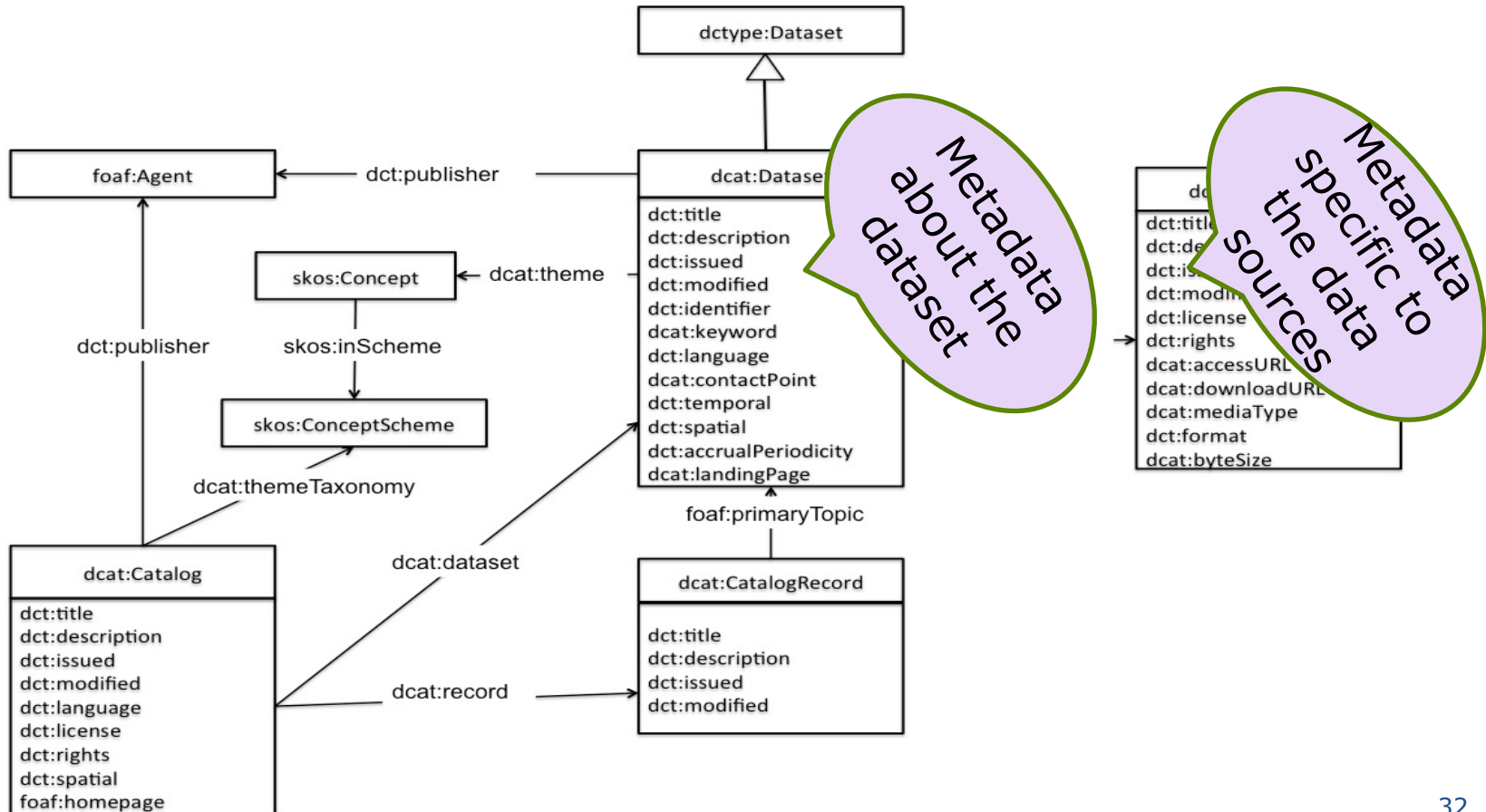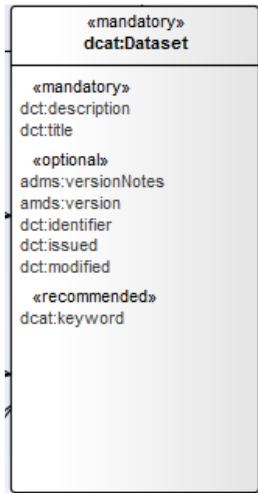*Temporal Coverage To*
2012

Dct:temporal

*Contact*
**Eurostat, the statistical office of the European Union**

Dcat:contactPoint

Share

31

# DCAT-AP (core) model

# Dataset



| Mandatory | Recommended | Optional |
|-----------|-------------|----------|
| dct:description<br>dct:title | dcat:contactPoint<br>dcat:distribution<br>dcat:keyword<br>dct:publisher<br>dcat:theme | adms:identifier<br>adms:sample<br>adms:versionNotes<br>dcat:landingPage<br>dct:accessRights<br>dct:accrualPeriodicity<br>dct:conformsTo<br>dct:hasVersion<br>dct:isVersionOf<br>dct:identifier<br>dct:issued<br>dct:language<br>dct:modified<br>dct:provenance<br>dct:relation<br>dct:source<br>dct:spatial<br>dct:temporal<br>dct:type<br>foaf:page<br>owl:versionInfo |

dqv:hasQualityAnnotation

stat:attribute
stat:dimension
stat:numSeries
stat:unitMeasure

StatDCAT-AP to add optional properties:

# 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 | Download dataset in TSV format  `ZIP` |
| ⬇ DOWNLOAD | Download dataset in TSV format (unzipped)  `TEXT/TAB-SEPARATED-VALUES` |
| ⬇ DOWNLOAD | Download dataset in SDMX–ML format  `ZIP` |

### Documentation

| | |
|---|---|
| ⬈ VISIT PAGE | ESMS metadata (Euro–SDMX Metadata structure) HTML  `HTML` |
| ⬇ DOWNLOAD | ESMS metadata (Euro–SDMX Metadata structure) SDMX |
| ⬇ DOWNLOAD | More information on Eurostat Website |

*URI*

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

*Status*

# Distribution
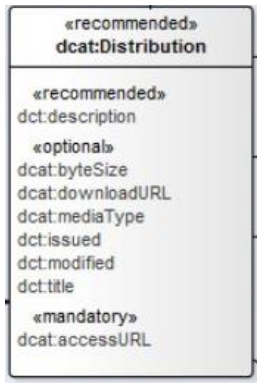


| Mandatory | Recommended | Optional |
|-----------|-------------|----------|
| dcat:accessURL | dct:description<br>dct:format<br>dct:license | adms:status<br>dcat:byteSize<br>dcat:downloadURL<br>dcat:mediaType<br>dct:conformsTo<br>dct:issued<br>dct:language<br>dct:modified<br>dct:rights<br>dct:title<br>foaf:page<br>spdx:checksum |

StatDCAT-AP to add optional property:

dct:type

# Use case: StatDCAT-AP 'users'

**Consumer**

DCAT-AP — Search / discovery of data existence

**producer**

SDMX — Process-oriented — SDMX/StatDCAT

**producer**

CSV — Ad-hoc

Definition of StatDCAT-AP

Evaluation of STAT DCAT-AP

ISA
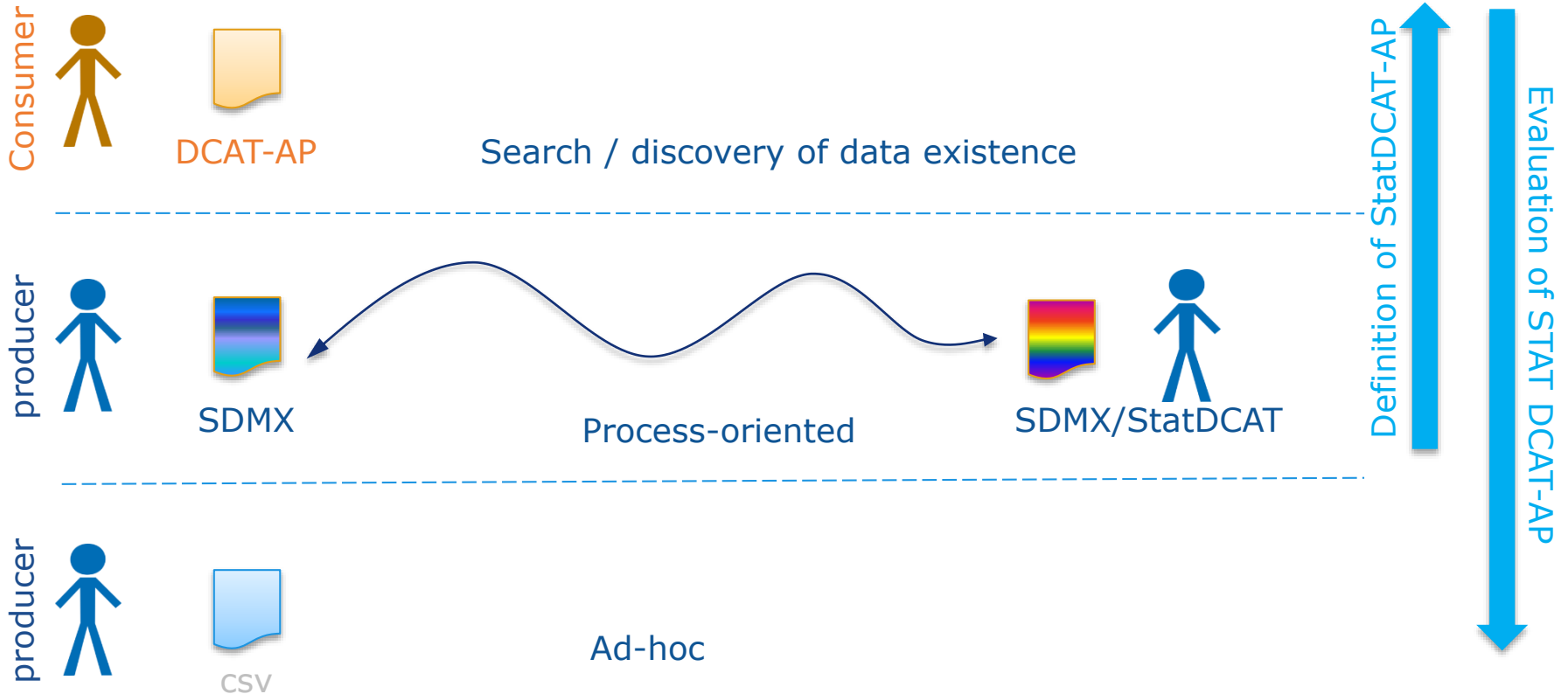
# SDMX Information Model: Schematic View

Concepts

Codes

Concepts and Coding Schemes used to Publish Data Sets

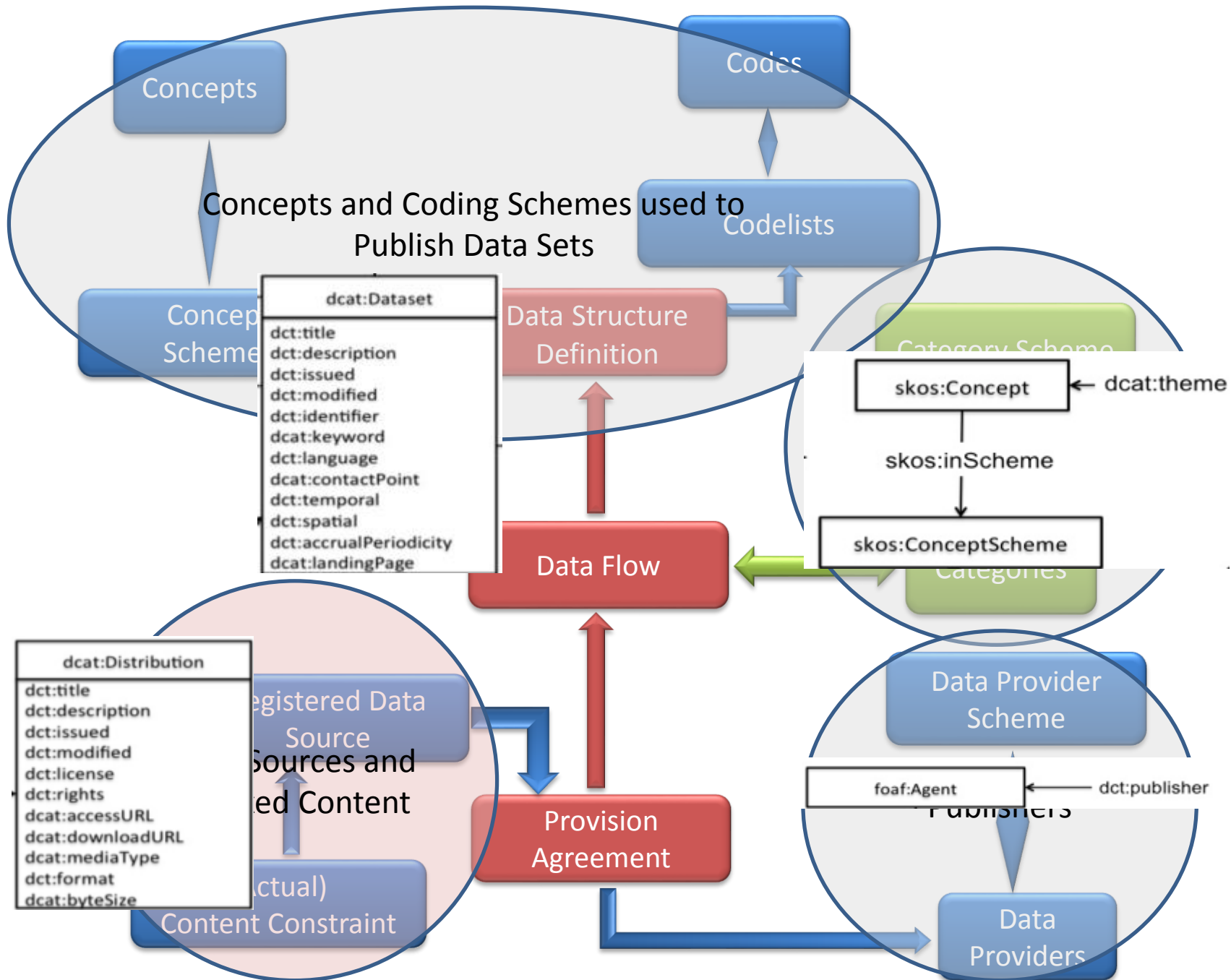Codelists

Concept Scheme

**dcat:Dataset**
- dct:title
- dct:description
- dct:issued
- dct:modified
- dct:identifier
- dcat:keyword
- dct:language
- dcat:contactPoint
- dct:temporal
- dct:spatial
- dct:accrualPeriodicity
- dcat:landingPage

Data Structure Definition

Category Scheme

skos:Concept ← dcat:theme

skos:inScheme

skos:ConceptScheme

Categories

Data Flow

**dcat:Distribution**
- dct:title
- dct:description
- dct:issued
- dct:modified
- dct:license
- dct:rights
- dcat:accessURL
- dcat:downloadURL
- dcat:mediaType
- dct:format
- dcat:byteSize

Registered Data Source

Sources and ed Content

Data Provider Scheme

foaf:Agent ← dct:publisher

Publishers

Provision Agreement

(Actual) Content Constraint

Data Providers

# Publishing StatDCAT-AP from SDMX: Requirements

*Ability to*

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

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

- Output StatDCAT-AP RDF

- Submit Catalogue metadata to the portal

ISA

StatDCAT-AP
A New Dawn in Statistical Data Discovery