Publication of Legislative Data

Open Data, AI and Semantic Web in the Legal Domain

Enrico Francesconi

Publications Office of the EU
enrico.francesconi@publications.europa.eu

Czech Republic Conference on
Open Data in Public Administration

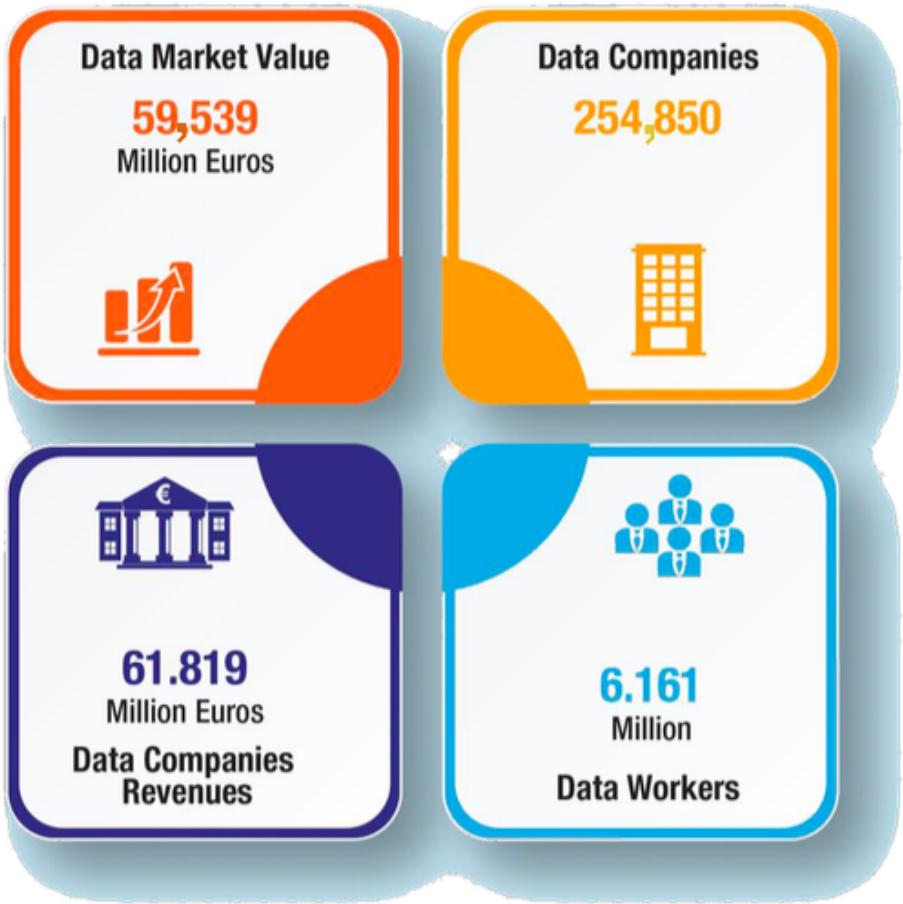
2 November 2018

Trends in IT

ARTIFICIAL INTELLIGENCE
LINKED OPEN DATA
SEMANTIC WEB
SMART DATA
ONTOLOGIES
eJustice
eLaw
NLP
OPEN DATA
WEB 1.0
WEB 2.0
WEB 3.0
WEB 4.0
WEB 5.0
Legal Semantic Web

Today's economy revolves around data

European data market in the EU28 (2016)



Cit: *"The Economic Benefits of Open Data" European Data Portal*

https://www.europeandataportal.eu/en/highlights/economic-benefits-open-data

Open Data in Economic Growth

Open Data economy is supposed to generate

- an additional growth up to 4% of the GDP by 2020
- a growth in cumulative revenues in the period 2016-2020 estimated to 110 million Euro

Benefits of Open Data

Open Data, Linked Data, and the Semantic Web

Data exploitation and mash-up on the Web need data available in a standard format, reachable, manageable and inter-linked

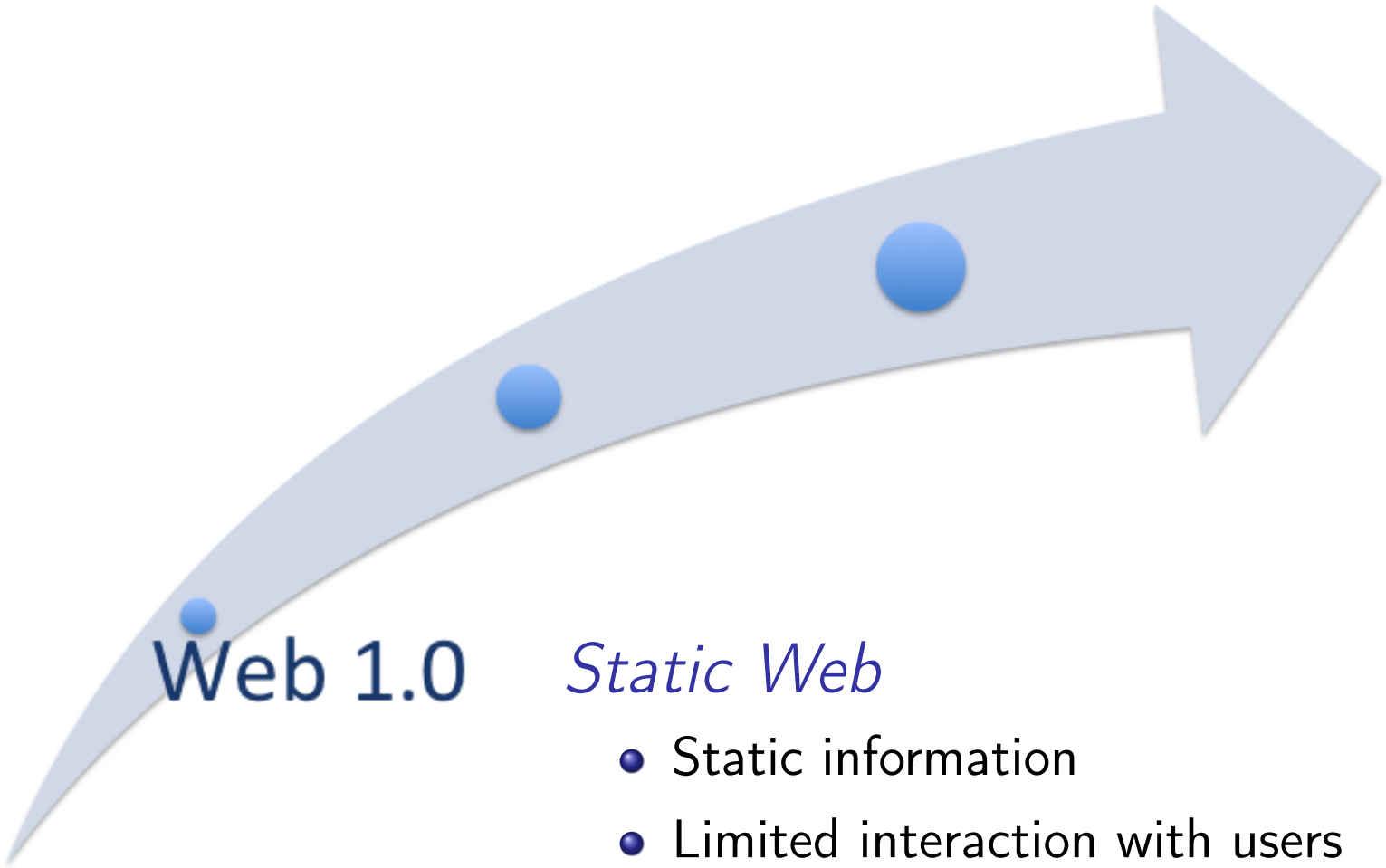
- Linked Data are structured and correlated data on the Web
- Linked Open Data (LOD) are Linked Data published with open licenses

The Semantic Web (Web 3.0) is made up of Linked Open Data

- Semantic Web is the *whole*
- Linked Open Data is the *parts*

How the Web has evolved from Web 1.0 to Web 3.0?

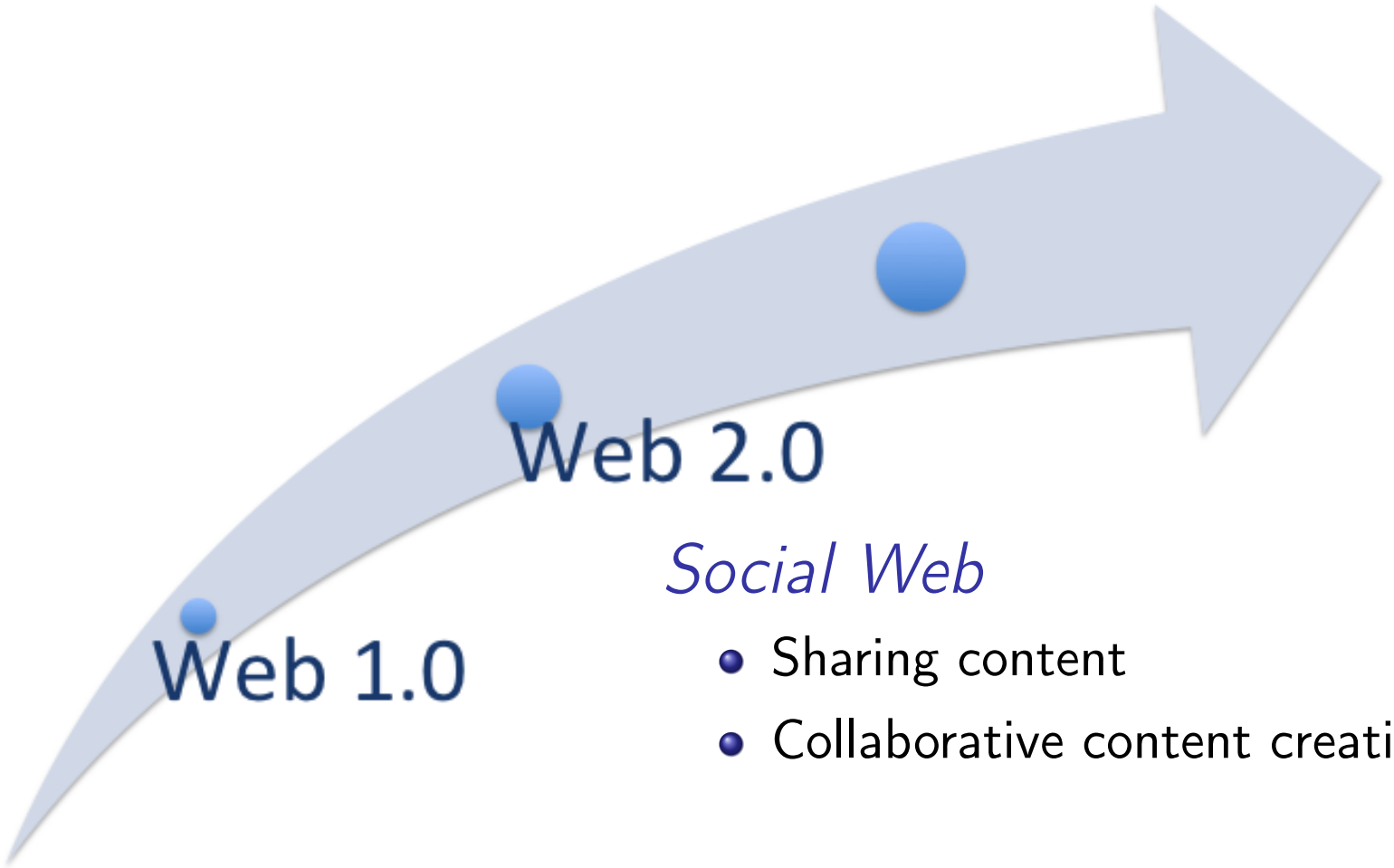
Evolution of the Web



Web 1.0 *Static Web*

- Static information
- Limited interaction with users

Evolution of the Web



Social Web

- Sharing content
- Collaborative content creation

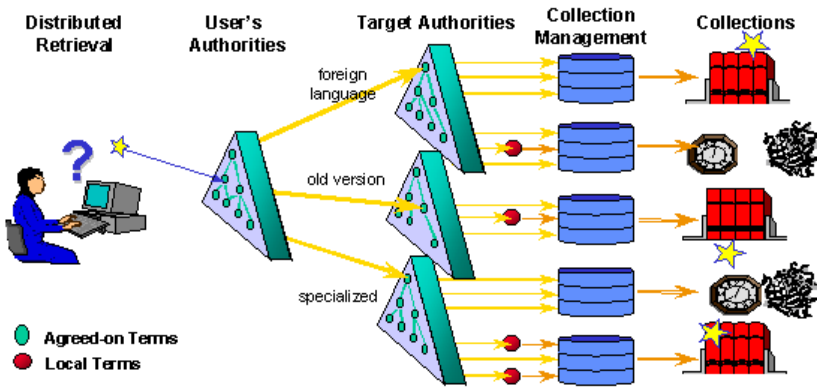
Search Engines in Web 1.0 and 2.0

- Key concept: Users' Information Needs

  a gap between what we know and what we want to know that motivates the search and this results in the formulation of a query
- Keywords indexing
- Query based on keywords and not on semantics
- Semantics is inferred:
  - by contexts
  - by algorithms able to infer the meaning of queries and contexts

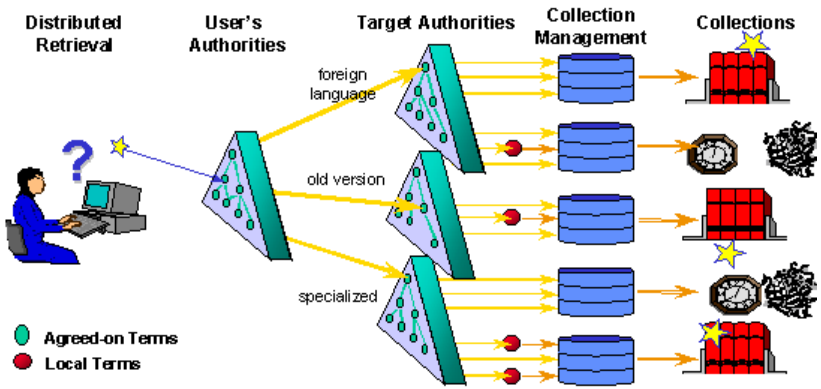
Example of Users' Information Needs

Unique point of access to resources in a distributed environment



Example of Users' Information Needs

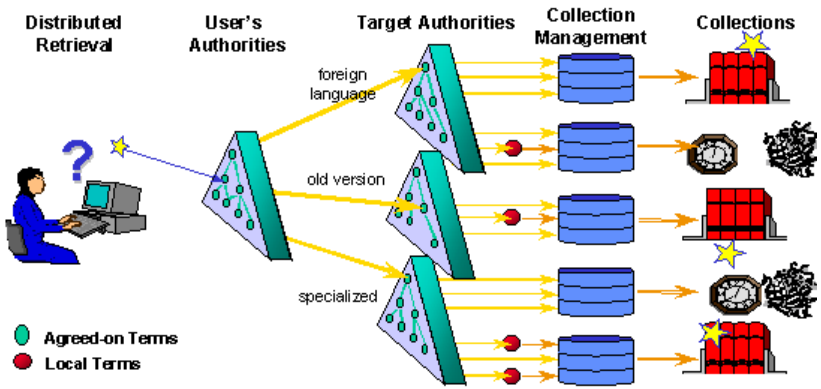
Unique point of access to resources in a distributed environment



Advanced information retrieval and reasoning services (ex. in the legal domain)

Example of Users' Information Needs

Unique point of access to resources in a distributed environment

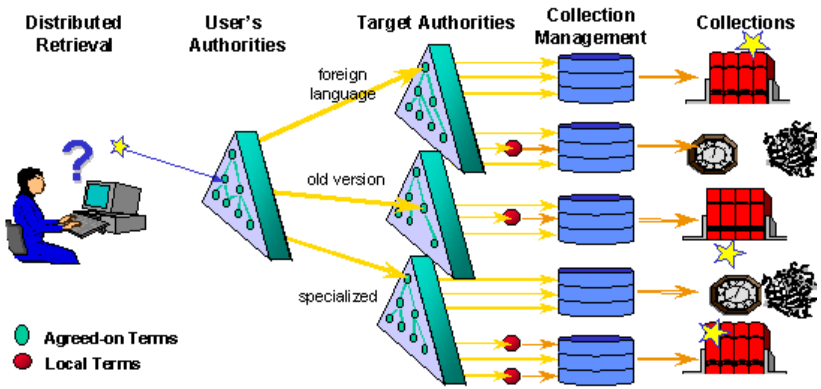


Advanced information retrieval and reasoning services (ex. in the legal domain)

- Which version of law n. 123 issued on 15 March 2007 was in force on 1 December 2010?

Example of Users' Information Needs

Unique point of access to resources in a distributed environment

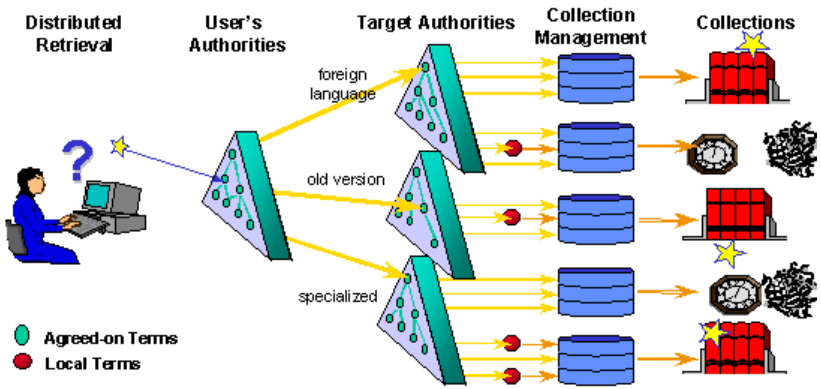


Advanced information retrieval and reasoning services (ex. in the legal domain)

- Which version of law n. 123 issued on 15 March 2007 was in force on 1 December 2010?
- In which laws Mr. XY is the first signer?

Example of Users' Information Needs

Unique point of access to resources in a distributed environment

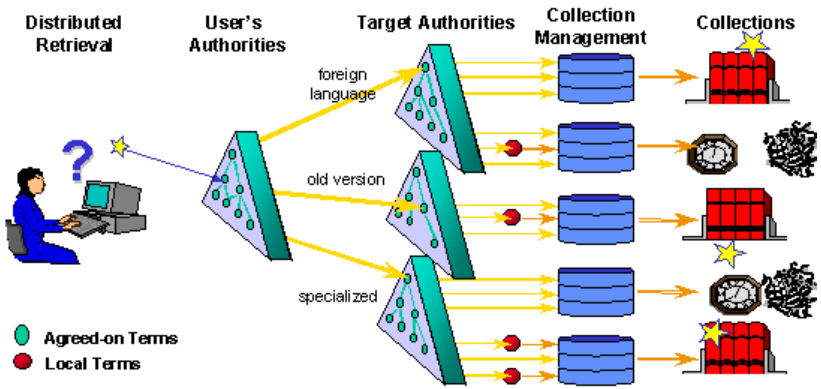


Advanced information retrieval and reasoning services (ex. in the legal domain)

- Which version of law n. 123 issued on 15 March 2007 was in force on 1 December 2010?
- In which laws Mr. XY is the first signer?
- Which laws on consumer protection do apply in a [specific region]?

Example of Users' Information Needs

Unique point of access to resources in a distributed environment

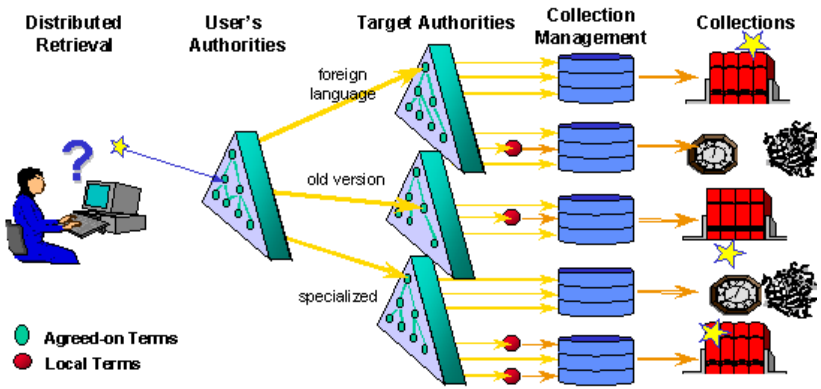


Advanced information retrieval and reasoning services (ex. in the legal domain)

- Which version of law n. 123 issued on 15 March 2007 was in force on 1 December 2010?
- In which laws Mr. XY is the first signer?
- Which laws on consumer protection do apply in a [specific region]?
- Which are the Rights of the Consumer according to the EU law?

Example of Users' Information Needs

Unique point of access to resources in a distributed environment



Advanced information retrieval and reasoning services (ex. in the legal domain)

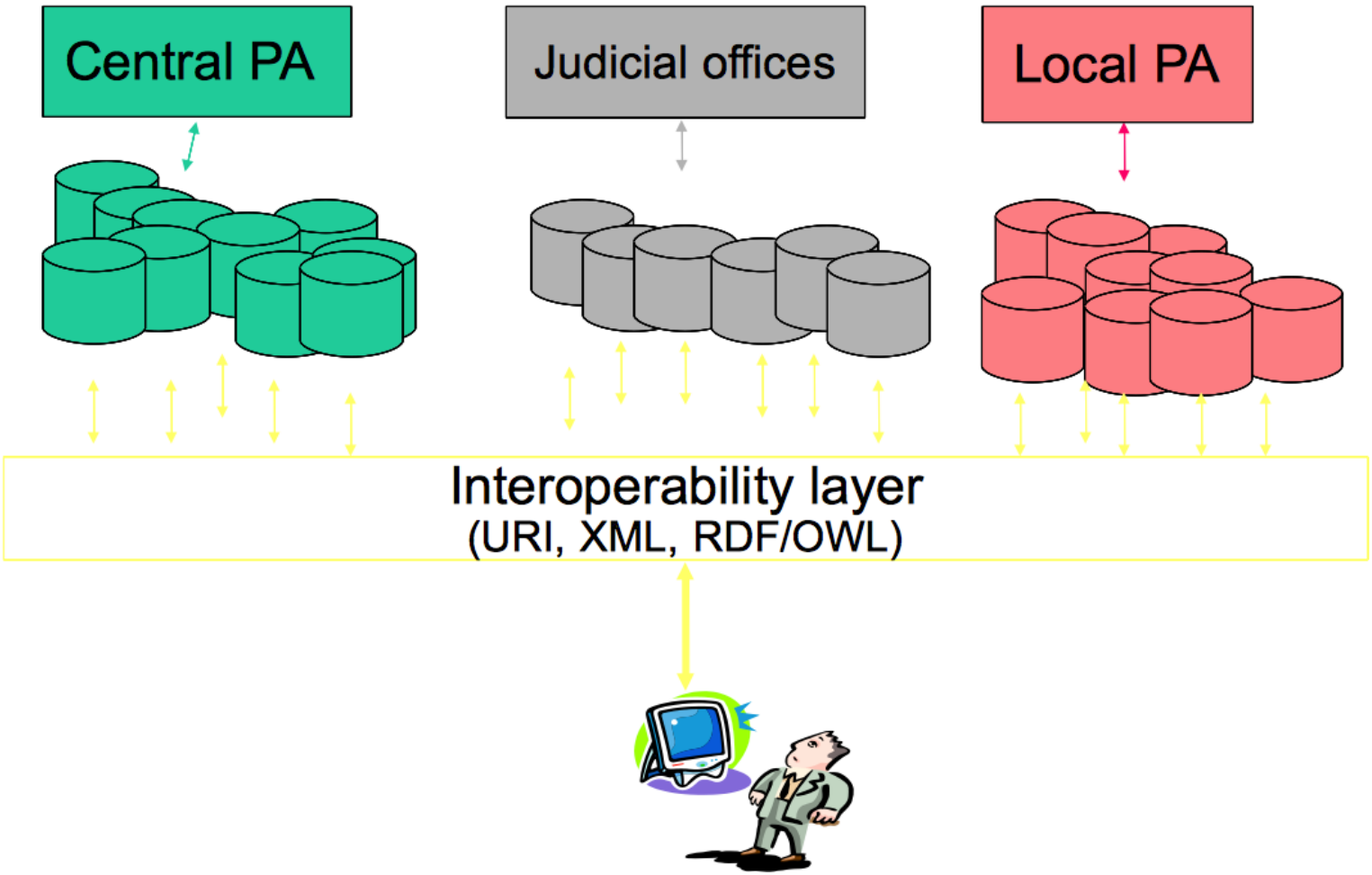
- Which version of law n. 123 issued on 15 March 2007 was in force on 1 December 2010?
- In which laws Mr. XY is the first signer?
- Which laws on consumer protection do apply in a [specific region]?
- Which are the Rights of the Consumer according to the EU law?
- Which are the Implicit Rights of the Consumer according to the EU law?

Users' Information Needs in the EU Legal Domain

Cross-Language Accessibility

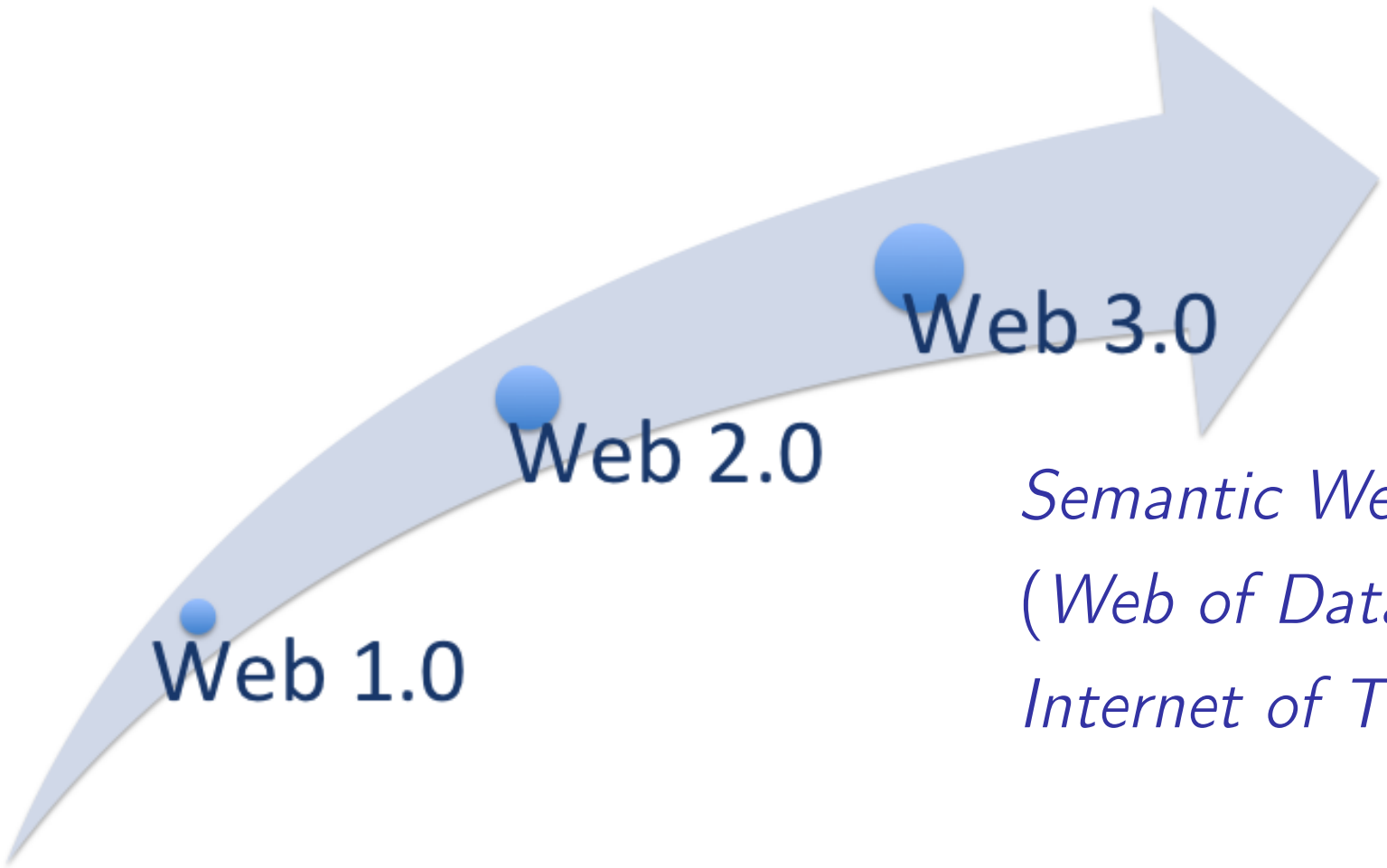
Accessing heterogeneous data sources without language barriers

Interoperability in a Multilanguage and Distributed Environment



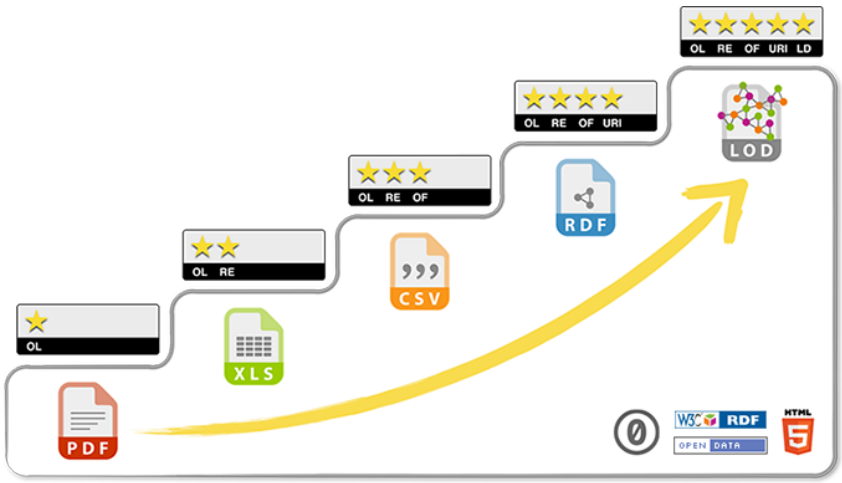
Re-organization of information in a distributed environment by an infrastructure based on standards

Evolution of the Web



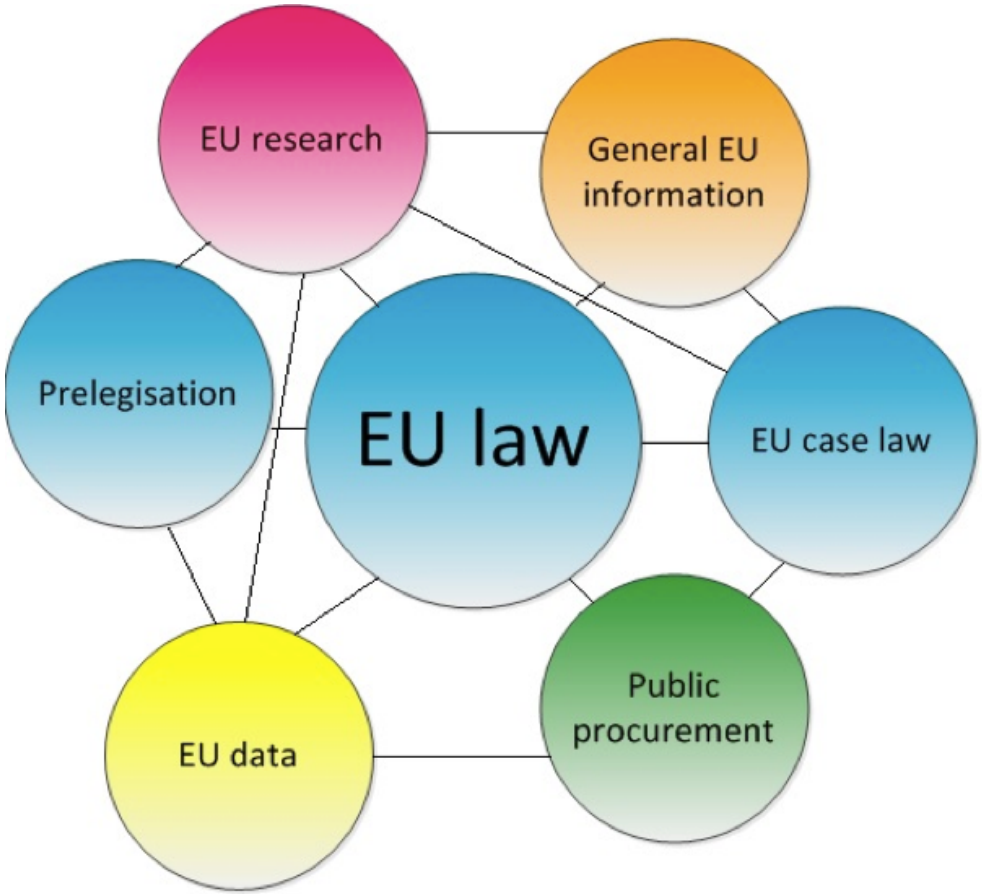
Semantic Web
(*Web of Data*,
Internet of Things)

5-star Rating Scheme for Linked Open Data



- ⋆ ⋆ ⋆ ⋆ ⋆ Link your data to other people's data to provide context
- ⋆ ⋆ ⋆⋆ Use open standard from W3C (URI, RDF and SPARQL) to describe things, so that people can point at them
- ⋆ ⋆ ⋆ Use non-proprietary format (e.g. CSV instead of excel)
- ⋆⋆ Available as machine-readable structured data (e.g. excel instead of image scan of a table)
- ⋆ Available on the Web (whatever format) but with an Open License (to be Open Data)

Publications Office Web 3.0 services – Open Data cloud



Cellar

Central hub of the OP publications and data, composed by:

1. Common Content Repository (CCR)
2. Common Metadata Repository (CMR)

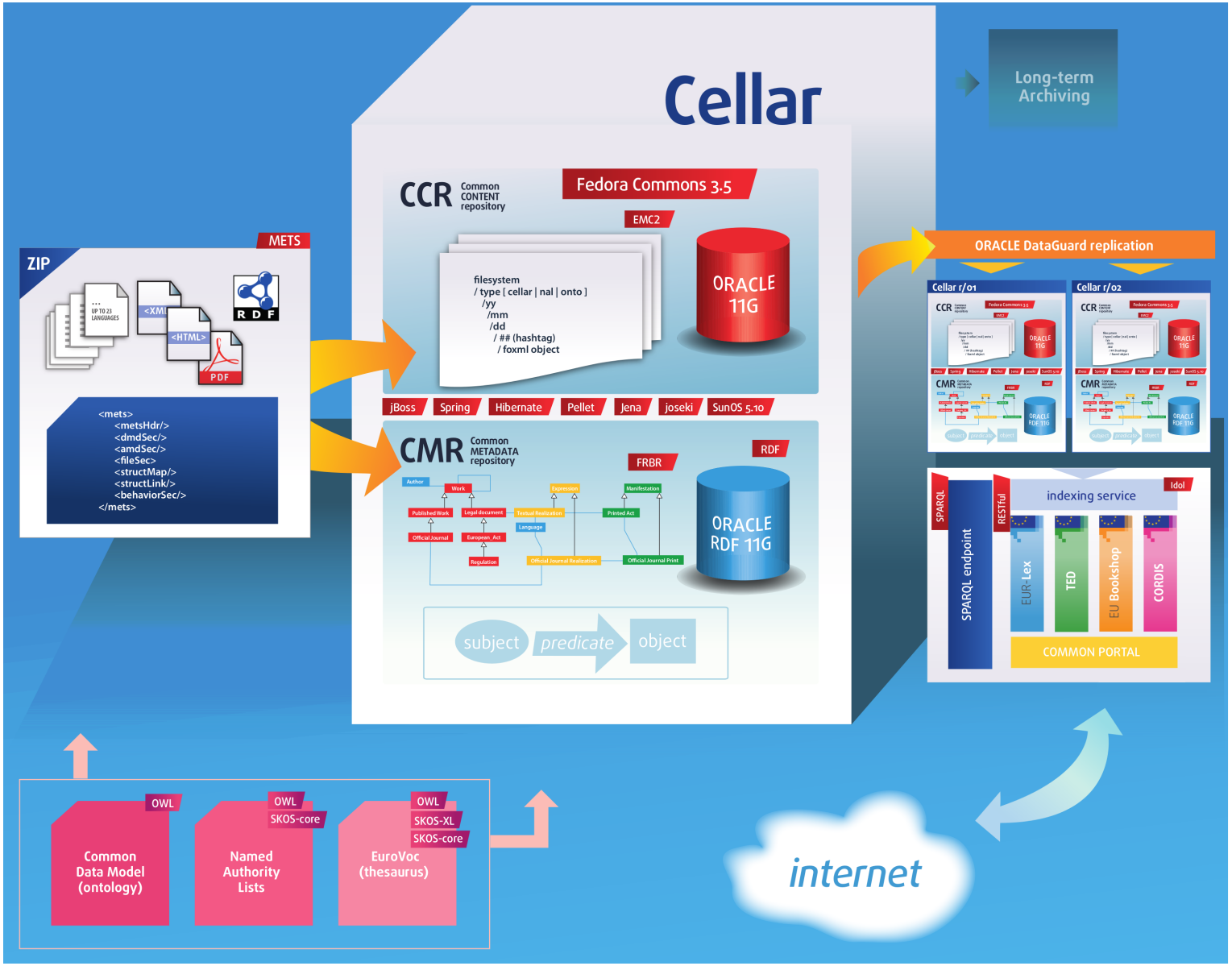
Docs are available for

- long term preservation
- open data
- indexing and advanced search & retrieval services

Resources are semantically described by an ontology:
Common Metadata Model (CDM)

Cellar SPARQL end-point publicly available

Cellar architecture



Cellar figures

- ∼250 million documents in 24 languages
- Metadata in 24 languages
- ∼1500 million semantic statements (triples)
- 10 million requests per day

Manual for reusers
https://publications.europa.eu/en/publication-detail/-/publication/50ecce27-857e-11e8-ac6a-01aa75ed71a1/language-en

Adoption of relevant standards

METS (Metadata encoding and transmission standard)

- ingestion protocol for the repository and the long-term preservation archive

Dublin Core

- core metadata definition

FRBR (Functional Requirements for Bibliographic Records)

- data model/ontology

RDF(S)/OWL (Linked Open Data (LOD))

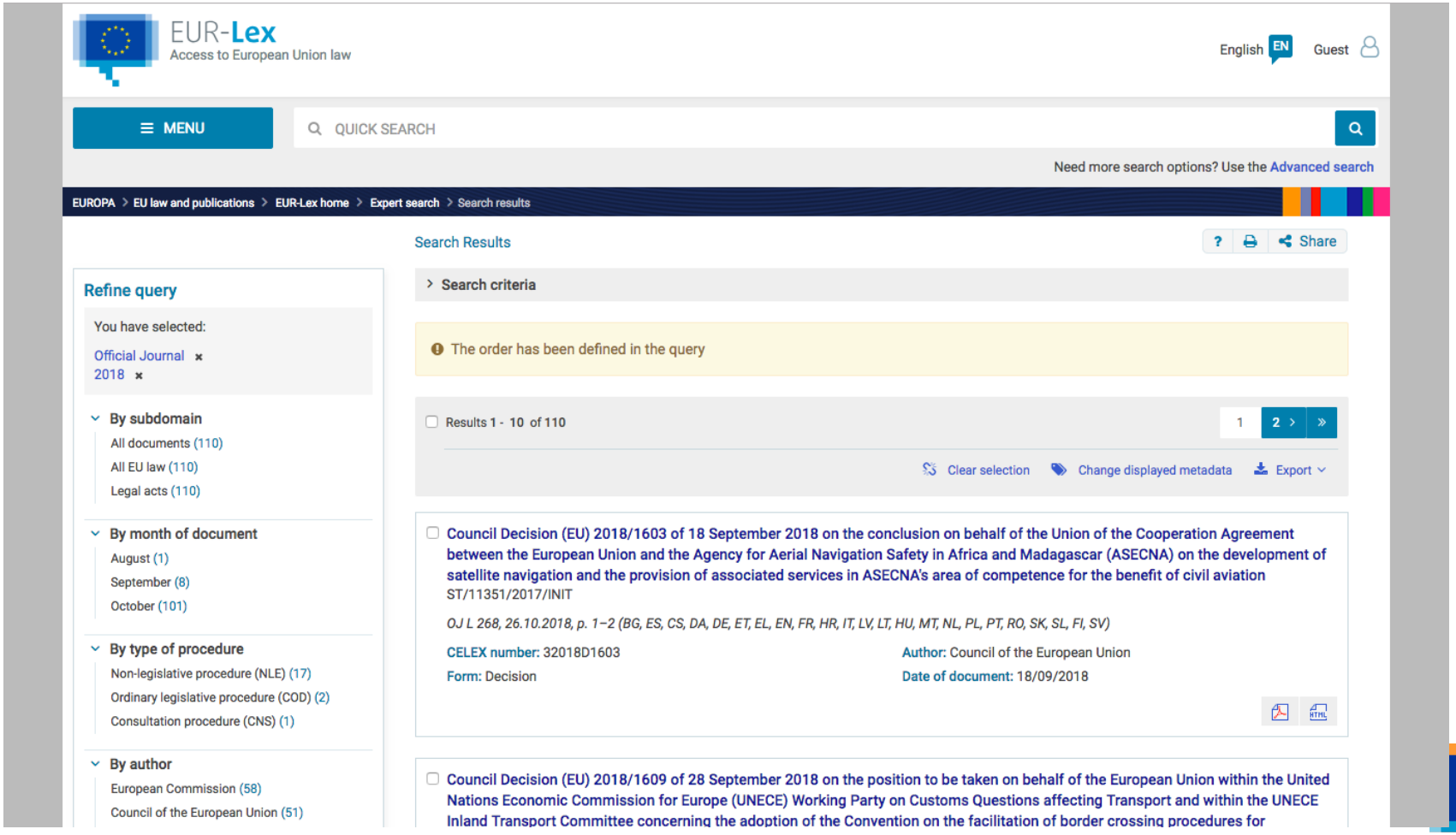
- access/reuse

DCAT (Data Catalog Vocabulary), DCAT-AP (Application profile)

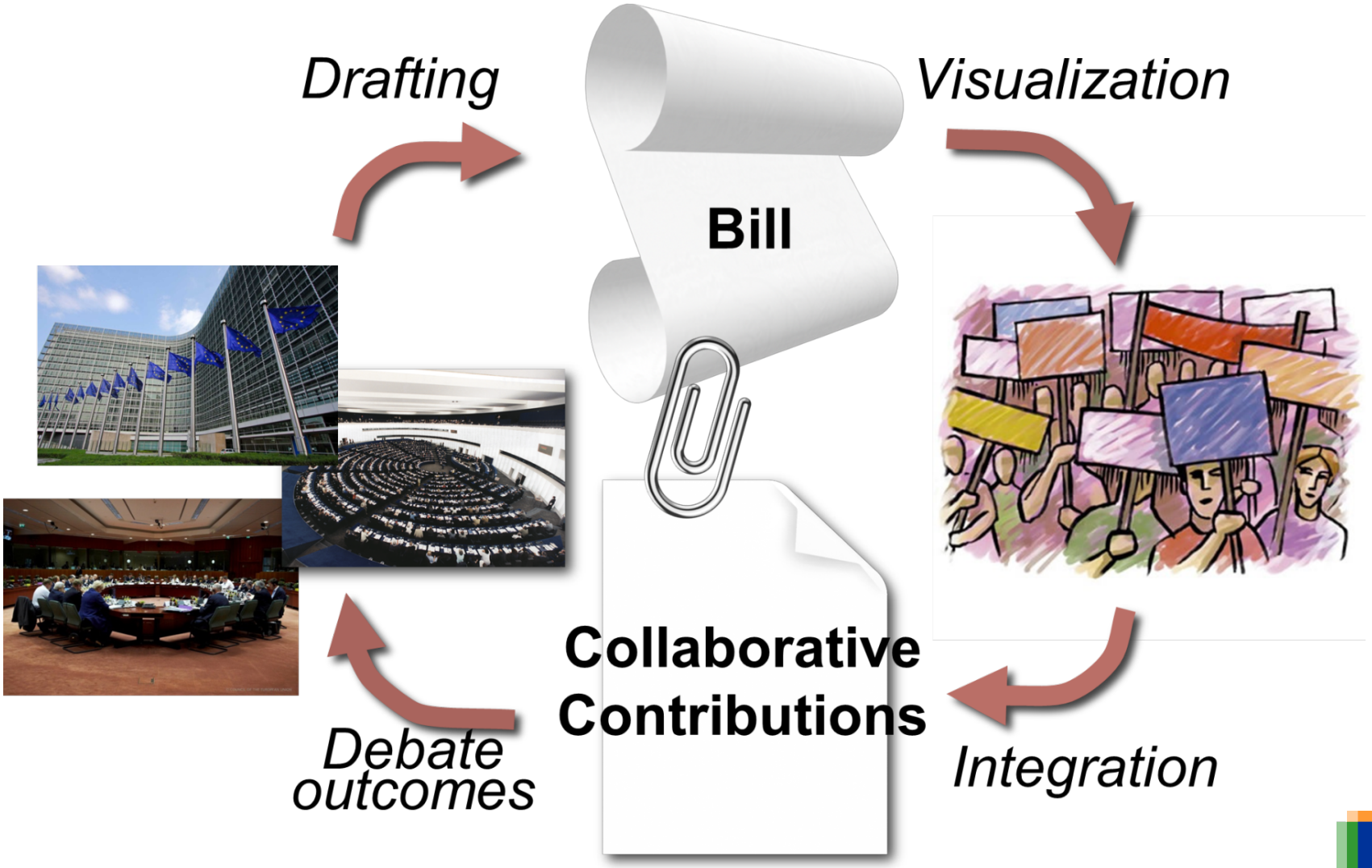
- reuse

Eur-Lex

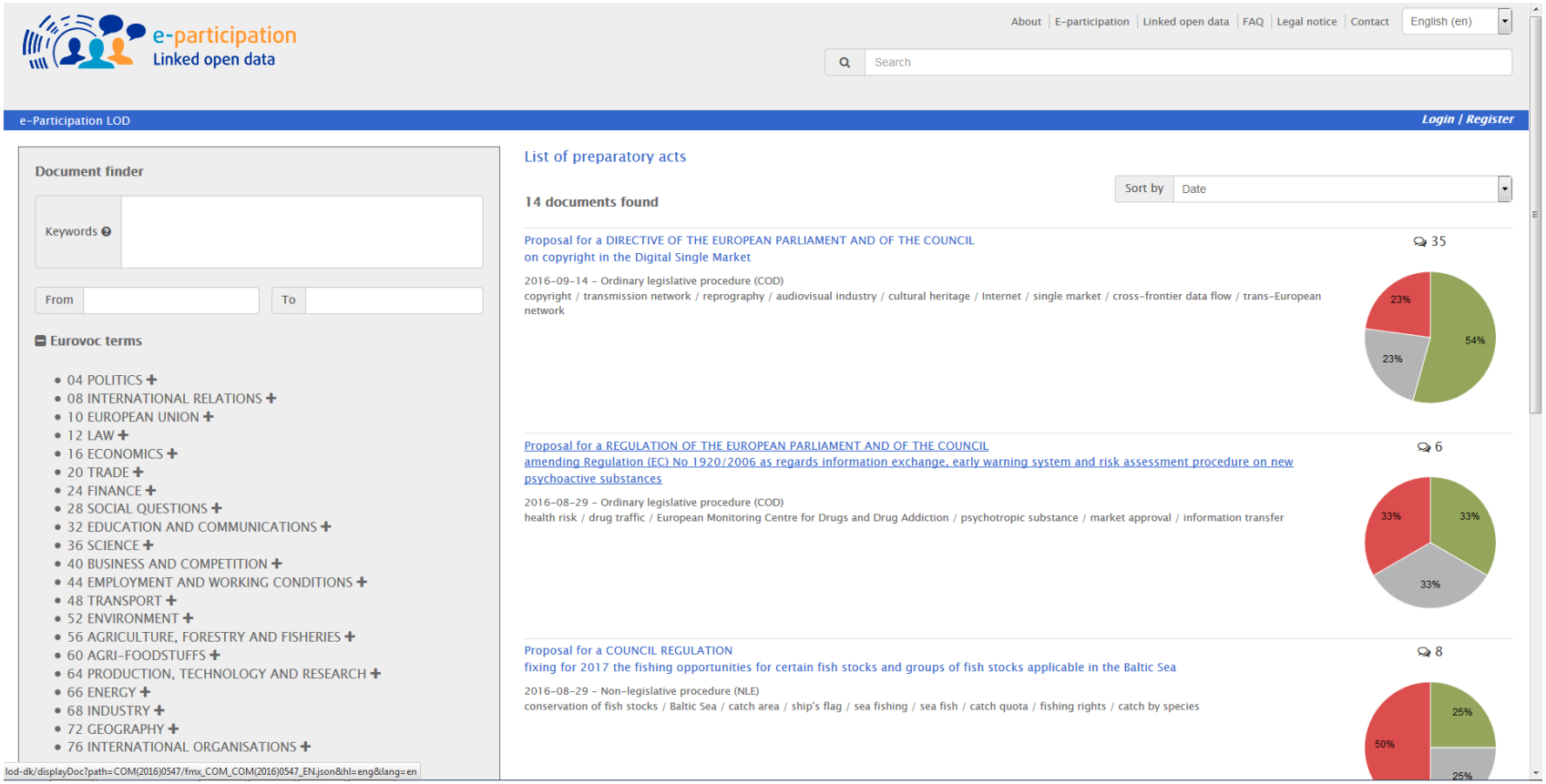
Cellar provides content for different services as Eur-Lex



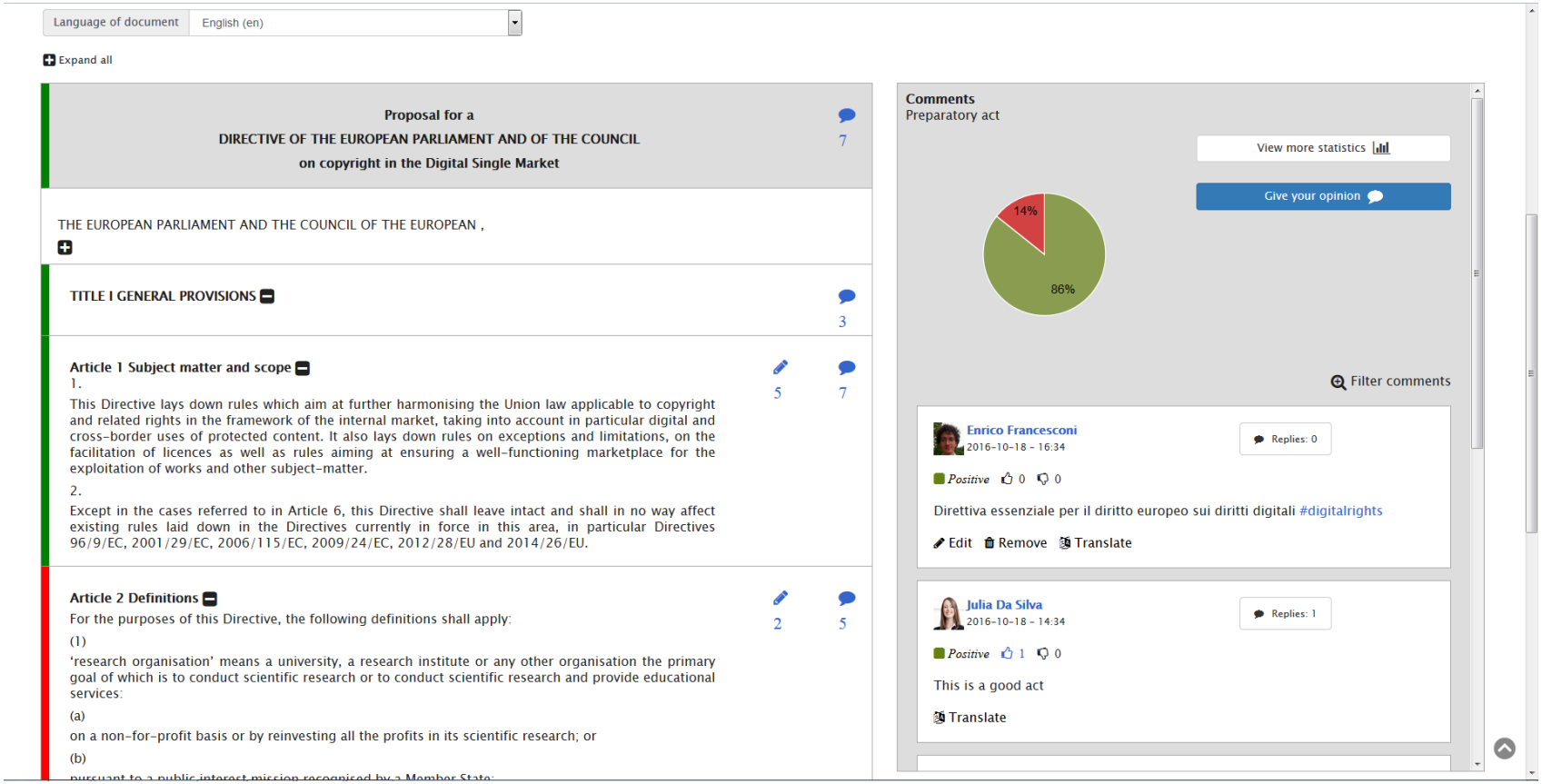
OP pilot project: LOD and e-Participation



Home page


e-participation
Linked open data
About | E-participation | Linked open data | FAQ | Legal notice | Contact
English (en)
Search
e-Participation LOD
Login / Register
Document finder
Keywords
From
To
Eurovoc terms
04 POLITICS
08 INTERNATIONAL RELATIONS
10 EUROPEAN UNION
12 LAW
16 ECONOMICS
20 TRADE
24 FINANCE
28 SOCIAL QUESTIONS
32 EDUCATION AND COMMUNICATIONS
36 SCIENCE
40 BUSINESS AND COMPETITION
44 EMPLOYMENT AND WORKING CONDITIONS
48 TRANSPORT
52 ENVIRONMENT
56 AGRICULTURE, FORESTRY AND FISHERIES
60 AGRI-FOODSTUFFS
64 PRODUCTION, TECHNOLOGY AND RESEARCH
66 ENERGY
68 INDUSTRY
72 GEOGRAPHY
76 INTERNATIONAL ORGANISATIONS
lod-dk/displayDoc?path=COM(2016)0547/fmx_COM_COM(2016)0547_EN.json&hl=eng&lang=en
List of preparatory acts
14 documents found
Sort by
Date
Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
on copyright in the Digital Single Market
35
2016-09-14 – Ordinary legislative procedure (COD)
copyright / transmission network / reprography / audiovisual industry / cultural heritage / Internet / single market / cross-frontier data flow / trans-European network
23%
23%
54%
Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
amending Regulation (EC) No 1920/2006 as regards information exchange, early warning system and risk assessment procedure on new psychoactive substances
6
2016-08-29 – Ordinary legislative procedure (COD)
health risk / drug traffic / European Monitoring Centre for Drugs and Drug Addiction / psychotropic substance / market approval / information transfer
33%
33%
33%
Proposal for a COUNCIL REGULATION
fixing for 2017 the fishing opportunities for certain fish stocks and groups of fish stocks applicable in the Baltic Sea
8
2016-08-29 – Non-legislative procedure (NLE)
conservation of fish stocks / Baltic Sea / catch area / ship's flag / sea fishing / sea fish / catch quota / fishing rights / catch by species
25%
50%
25%

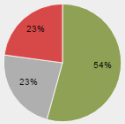
The discussion forum (comments, sentiments, amendments)



Statistics on users' sentiments per document and section

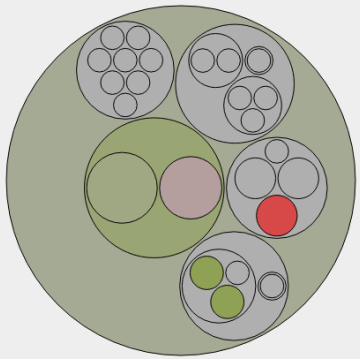
Proposal for a
DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
on copyright in the Digital Single Market

Interinstitutional procedure code : 2016/0280/COD
CELEX : 52016PC0593
Creation date of document : 2016-09-14
Legal procedure type : **Ordinary legislative procedure (COD)**
Document directory identifiant : 132060
Keywords : *copyright / transmission network / reprography / audiovisual industry / cultural heritage / Internet / single market / cross-frontier data flow / trans-European network*

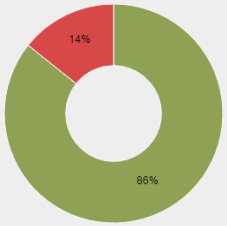


35

Statistics about the Act | Statistics about the Act and its sections | Comments | Amendments | Back to the document

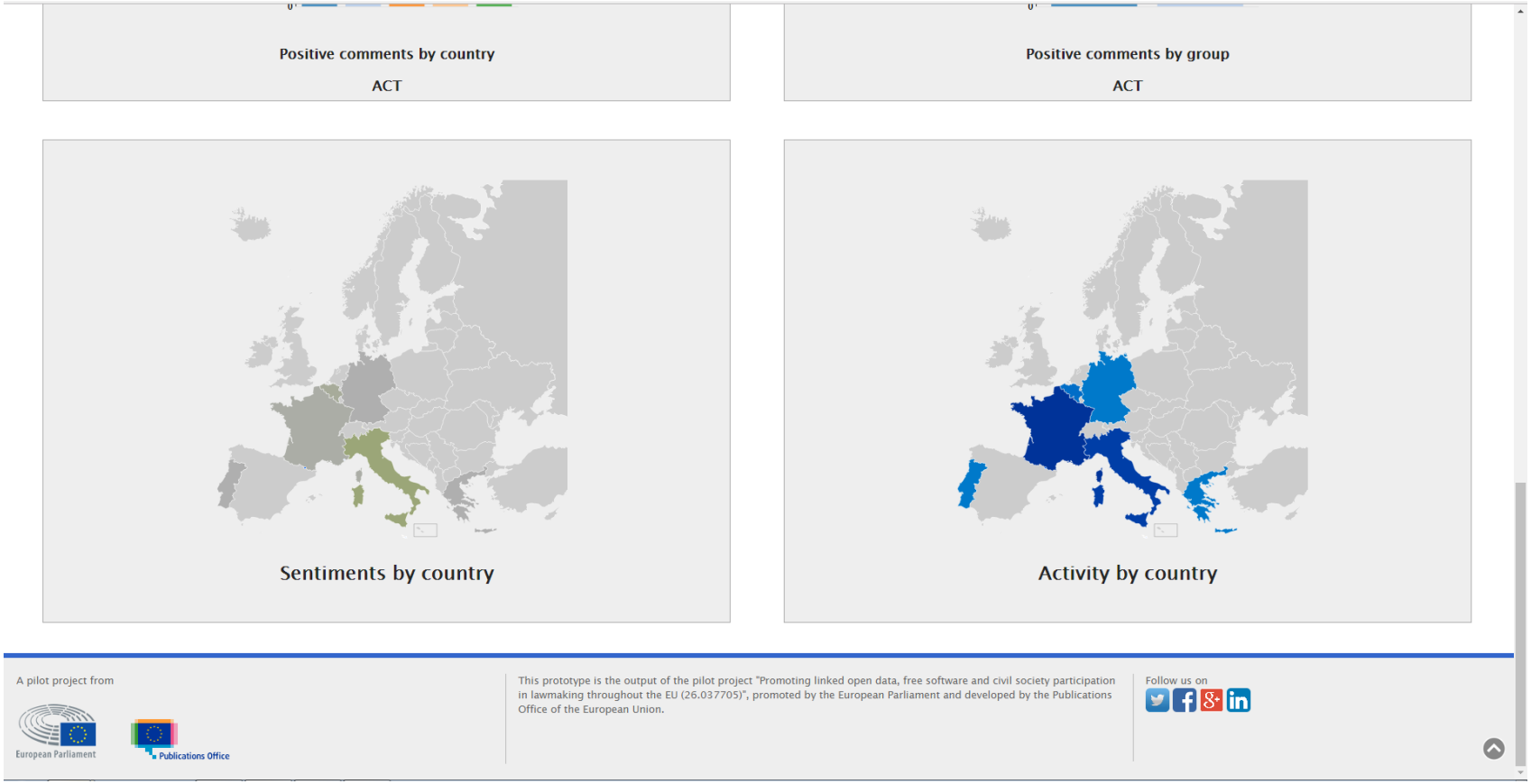


Sentiments per section



Comments division
ACT

Maps of the sentiments and activities per country

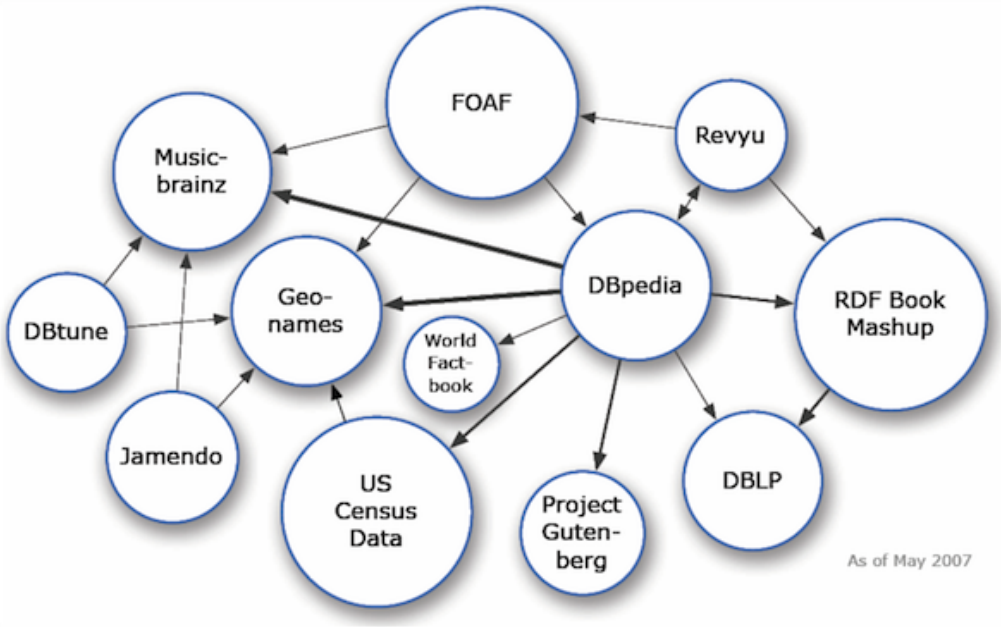


Role of Artificial Intelligence in the Semantic Web

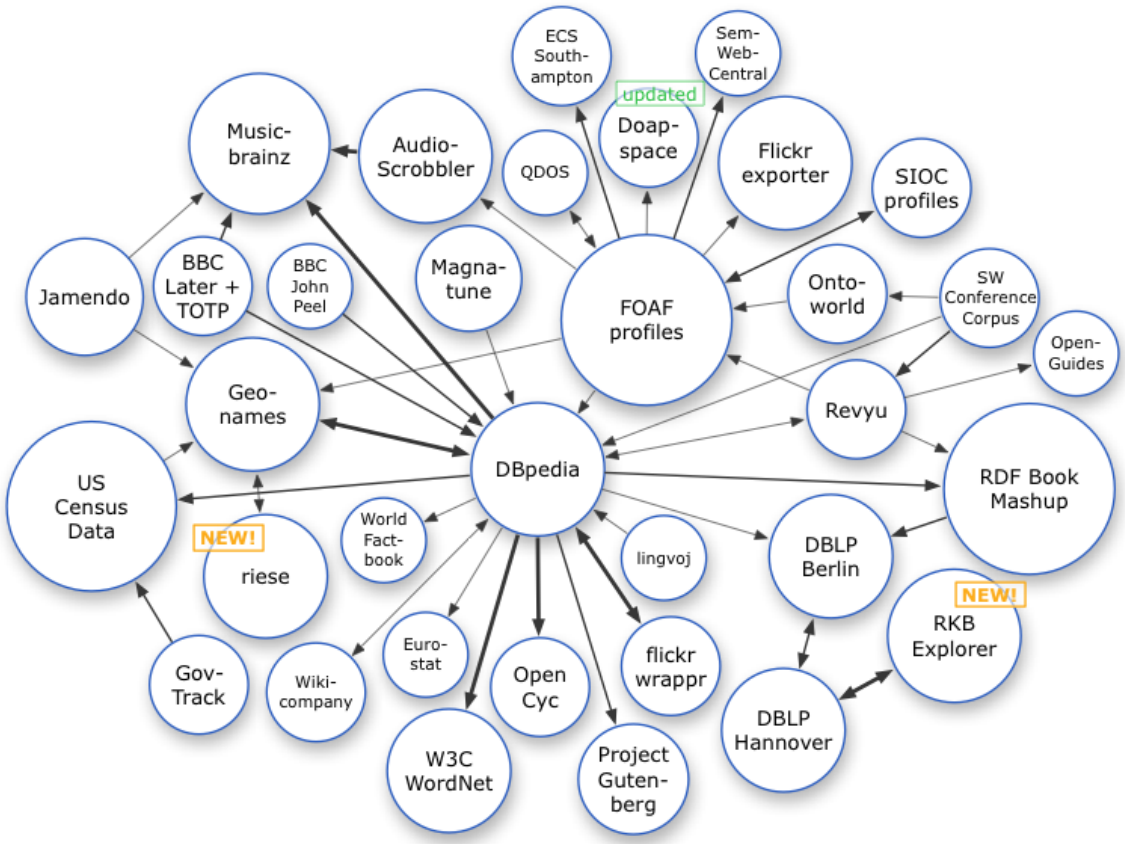
Semantic Web is an infrastructure for _Artificial Intelligence_ (AI)

- AI (NLP, machine learning) helps to translate the language into machine-understandable data (*Smart Data*)
- AI exploits Smart Data to implements advanced reasoning
  - Premises $\Longrightarrow$ Conclusions
  - Accessing Implicit Knowledge

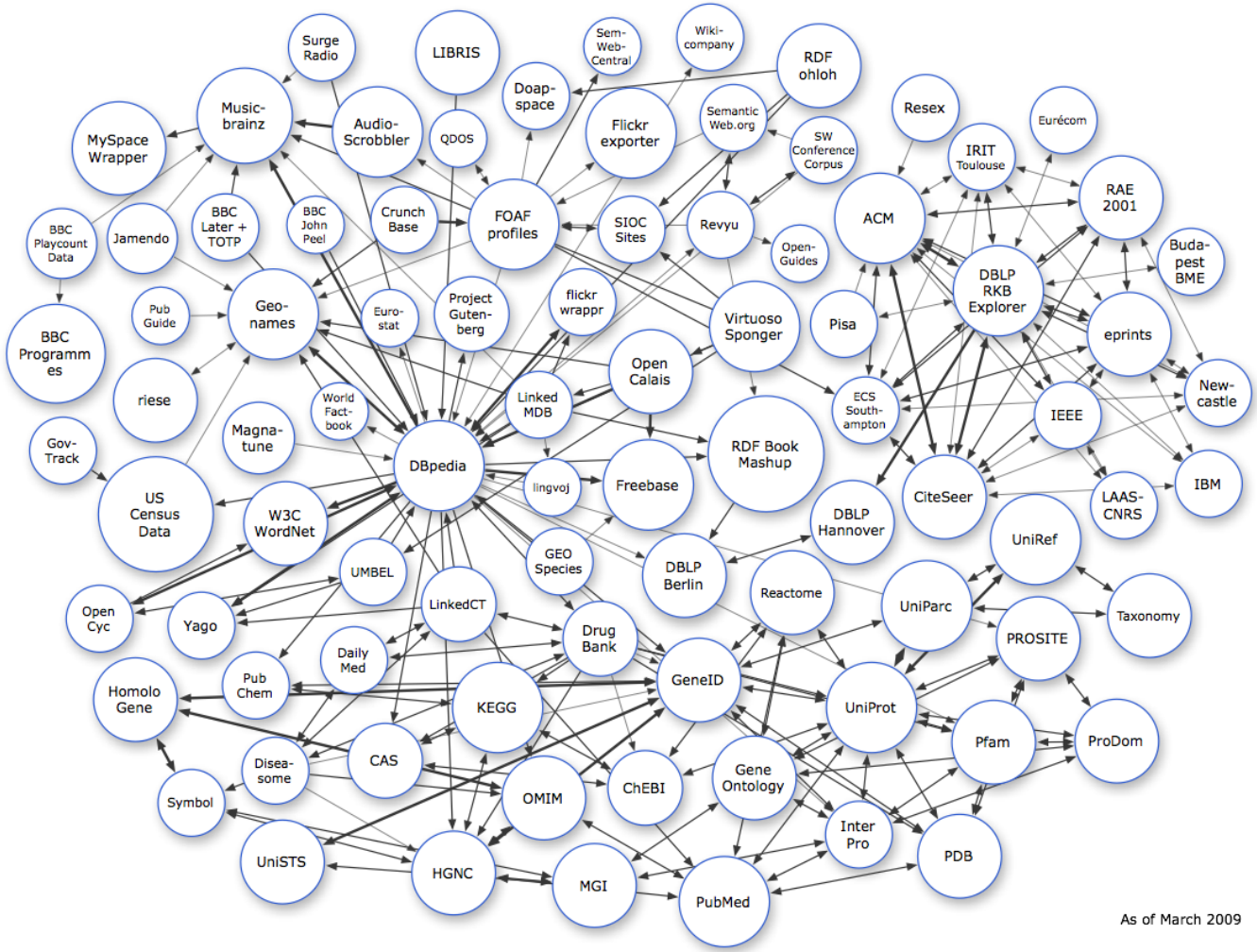
Linked Open Data growth



Linked Open Data growth

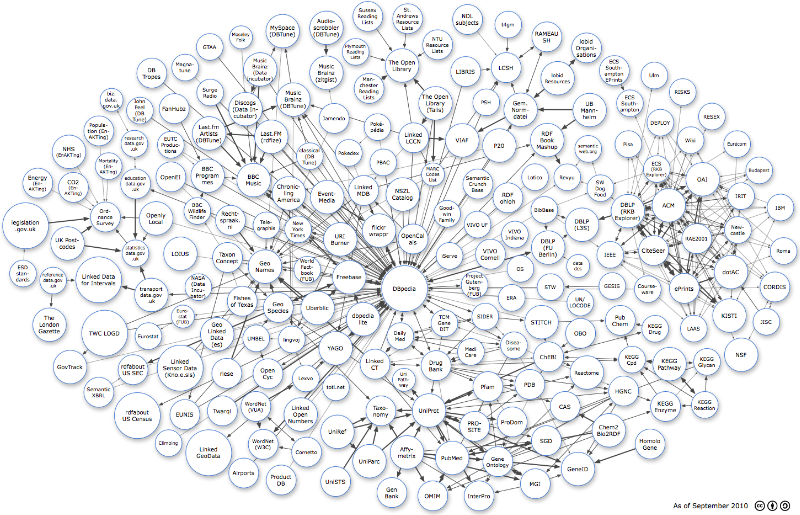


Linked Open Data growth

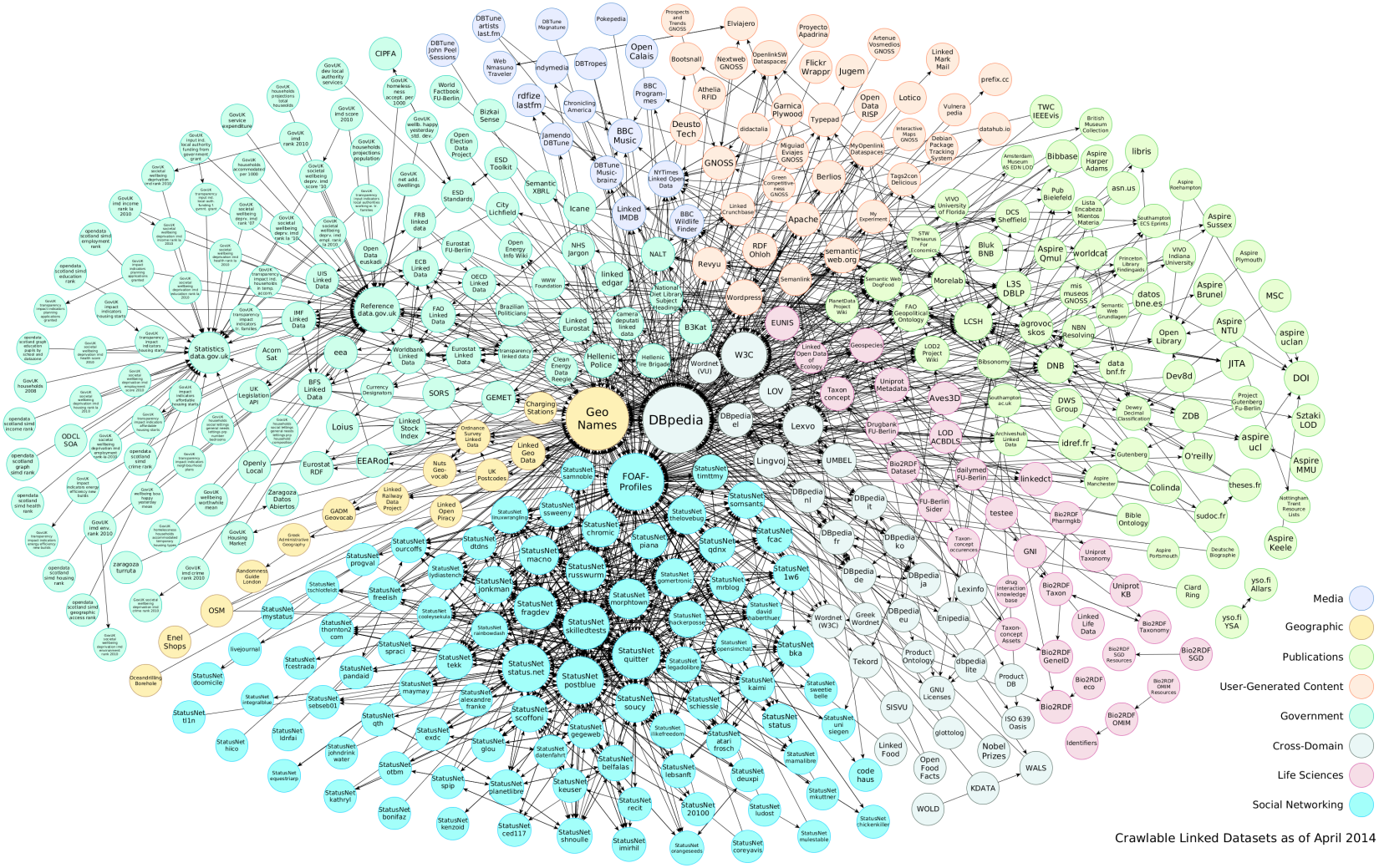


As of March 2009

Linked Open Data growth



Linked Open Data growth



What Open Data and Semantic Web are about?

They have do with Business Models for data consuming and sharing (Web services and Apps)

What Open Data and Semantic Web are about?

They have do with Business Models for data consuming and sharing (Web services and Apps)

In conjunction with eGov initiatives:

- Memorandum on Transparency and Open Government (Barack Obama's administration, 2009)
- Open Data institutional initiatives (European Union)

What Open Data and Semantic Web are about?

They have do with Business Models for data consuming and sharing (Web services and Apps)

In conjunction with eGov initiatives:

- Memorandum on Transparency and Open Government (Barack Obama's administration, 2009)
- Open Data institutional initiatives (European Union)

They are about:
Legal Data Accessibility, Quality, Reusability for Citizens

Therefore a promotion of:
Freedom, Justice and Democracy

Thanks for your attention!

enrico.francesconi@publications.europa.eu