AI and Law
Semantic Web, Open Data and AI in the Legal Domain

Enrico Francesconi

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

ITTIG-CNR – Florence (Italy)
enrico.francesconi@ittig.cnr.it

Central South University, Changsha – 14 April 2019

Trends in IT


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

AI and Law: a combination that comes from afar

The **Law**

- is made of Rules
- interprets and creates Facts

AI and Law: a combination that comes from afar

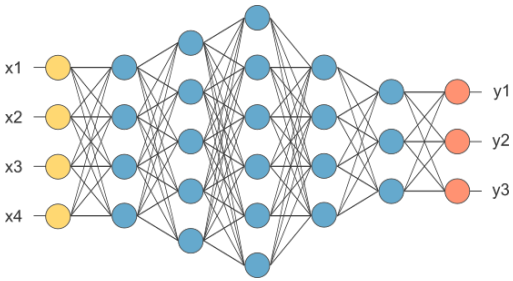
The **Law**

- is made of Rules
- interprets and creates Facts

The **Turing Machine** processes Facts (data) through Rules

- Symbolic paradigm
  - Rules expressed by symbols
  - Collection of rules and algorithms
  - ex: Expert Systems

- Sub-symbolic paradigm (connectionist models)
  - Rules as combination of elementary processing structures
  - Learning by examples
  - ex: Neural Networks



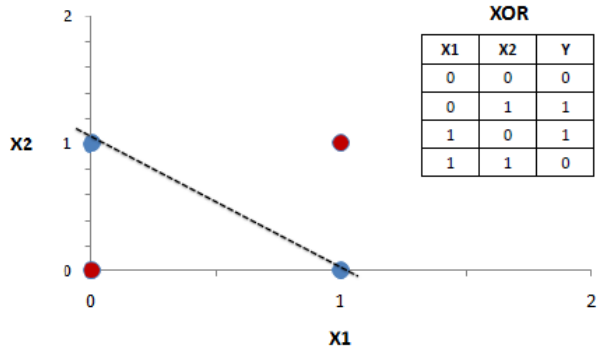
AI in the '40s-'90s: from the first results to the *AI Winter*

Limited results in the symbolic AI

- toy applications
- costs and complexity of representing and keeping information up-to-date
- not all the information can be represented in symbolic form

Limits of the computational power of the first connectionist models

- *Perceptron algorithm* (Rosemblatt)
- *XOR problem* (Marvin Minsky and Seymour Papert)



AI Winter: crisis in the Artificial Intelligence research

AI and Knowledge

Elaine Rich (Univ. Texas), Kevin Knight (Univ. South. California)

Intelligence requires Knowledge

AI Winter due to the lack of Knowledge available

Problems in managing Knowledge

- it's voluminous
- it is hard to characterize accurately
- it is constantly changing
- it requires a semantic organization

AI and the Web: the end of the *AI Winter*

On mid '90s the AI meets the Web

- Availability of large quantity of information in digital format for the development of AI systems
- Internet and the Web need advanced applications for managing data

AI and the Web: the end of the *AI Winter*

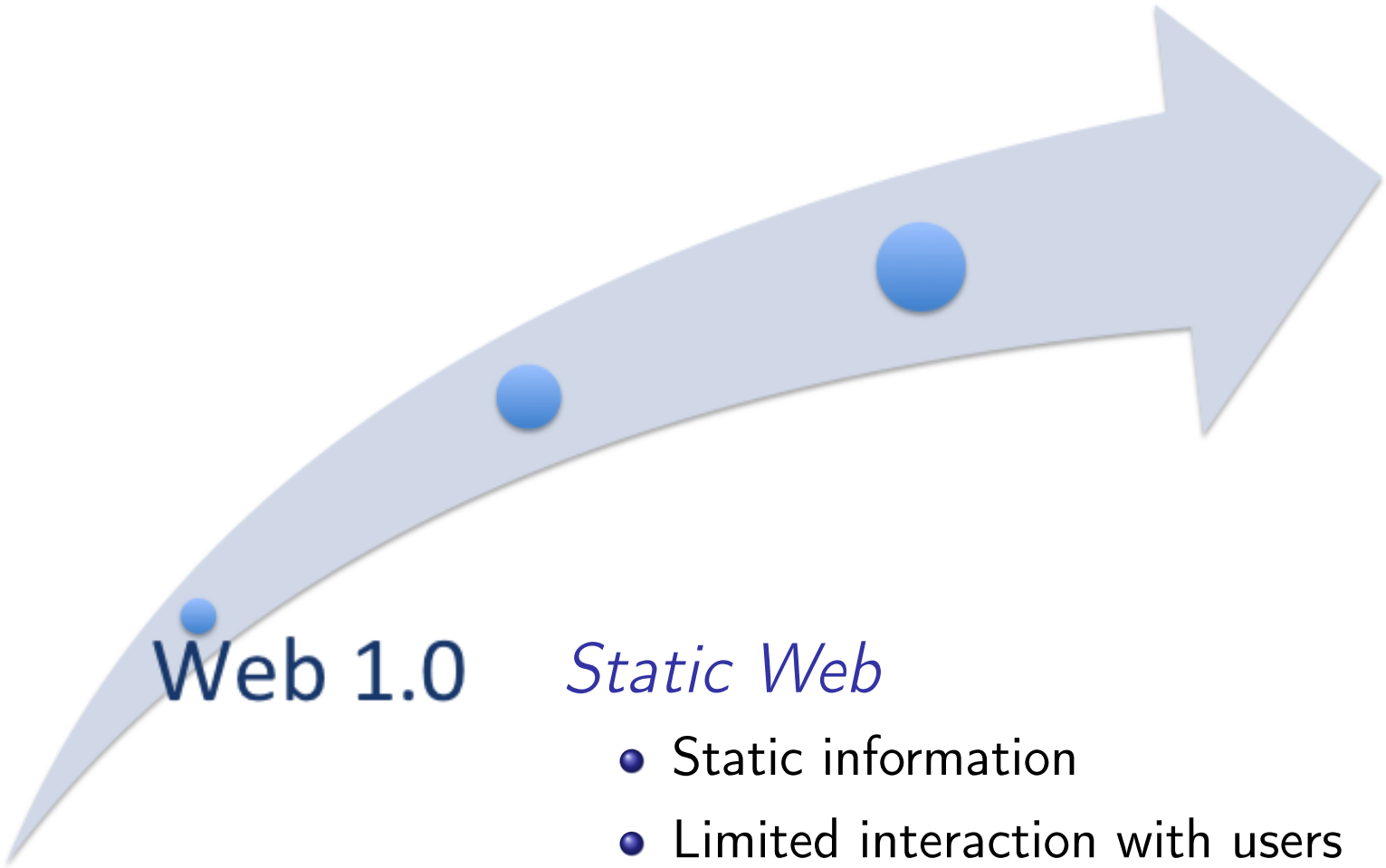
On mid '90s the AI meets the Web

- Availability of large quantity of information in digital format for the development of AI systems
- Internet and the Web need advanced applications for managing data

The evolution of the AI has followed the evolution of the Web

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

Evolution of the Web

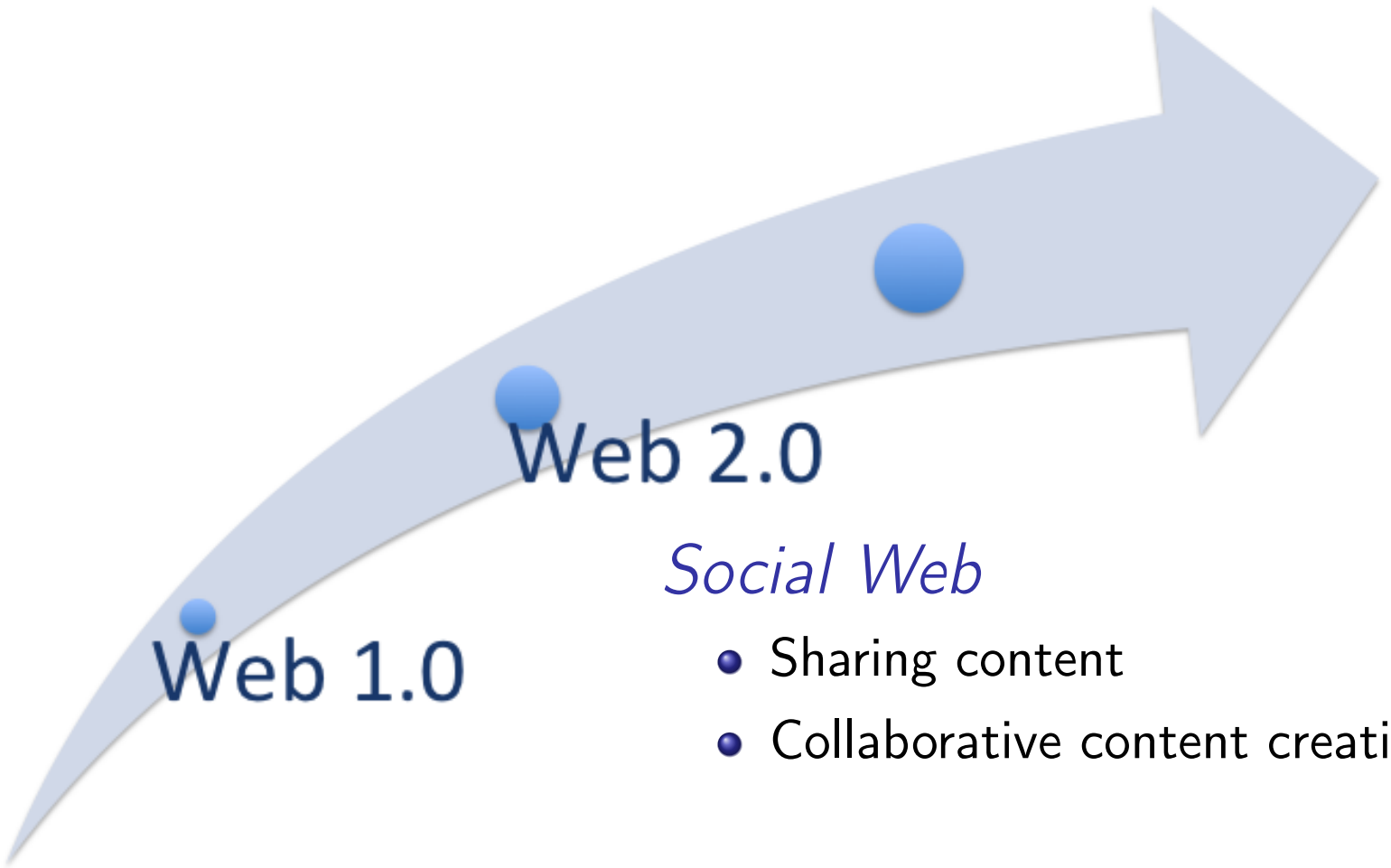


Web 1.0

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

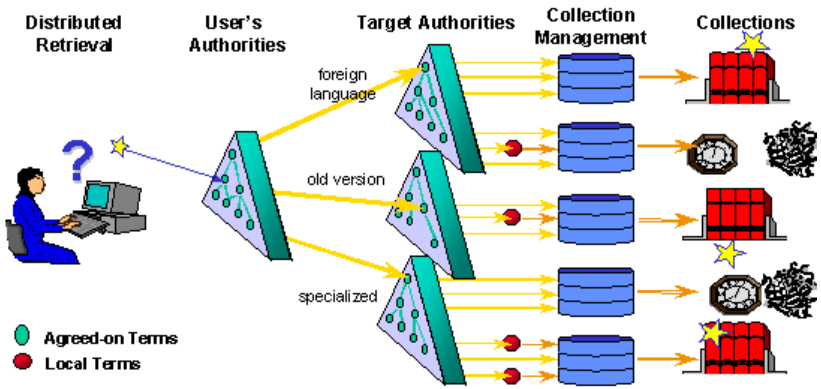
- Keywords indexing

- Query based on keywords and not on semantics

- Semantics is inferred:
  - by contexts
  - by algorithms able to infer the meaning of queries and contexts

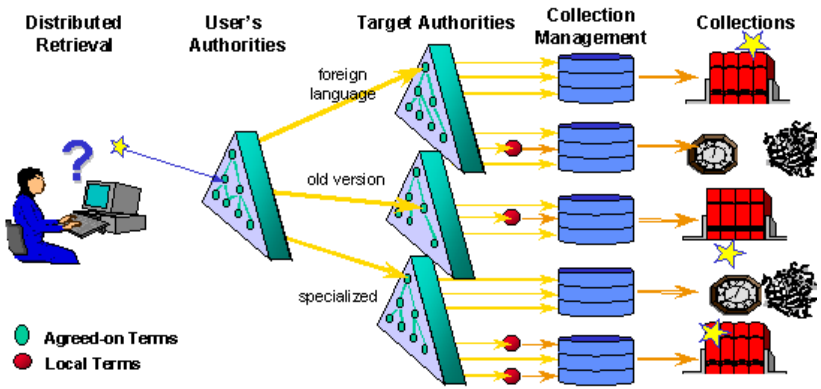
Example of Users' Information Needs

Unique point of access to resources in a distributed environment



Example of Users' Information Needs

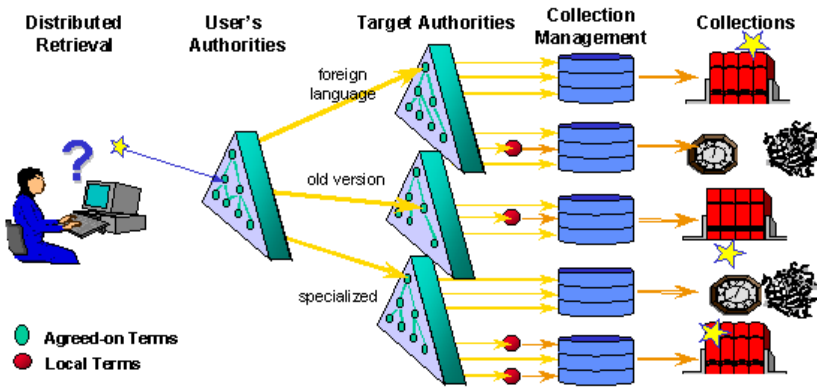
Unique point of access to resources in a distributed environment



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

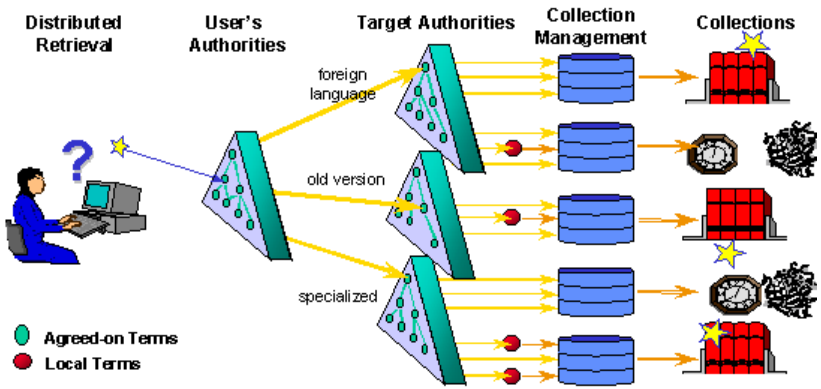


Advanced information retrieval and reasoning services (ex. in the legal domain)

- 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

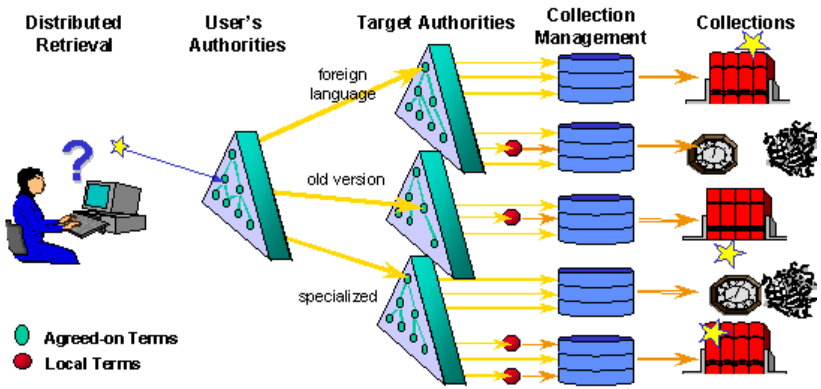


Advanced information retrieval and reasoning services (ex. in the legal domain)

- 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

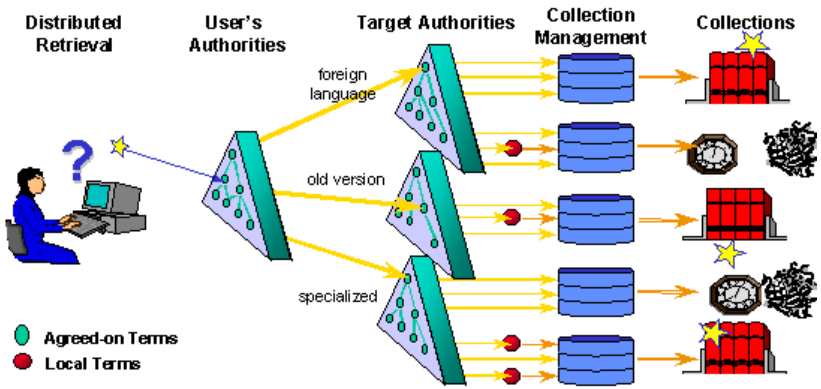


Advanced information retrieval and reasoning services (ex. in the legal domain)

- 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

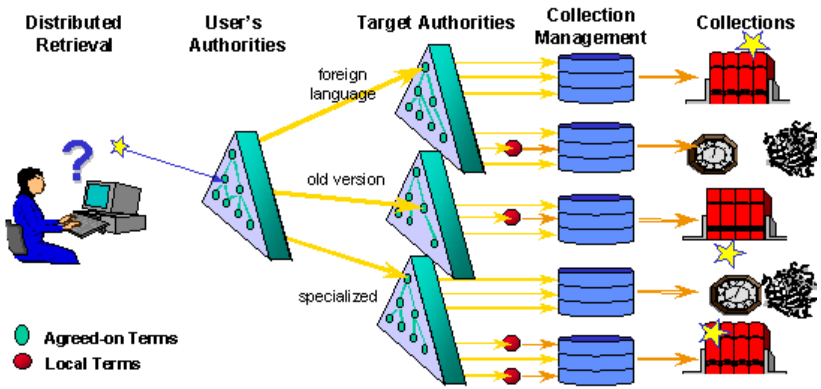


Advanced information retrieval and reasoning services (ex. in the legal domain)

- 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)

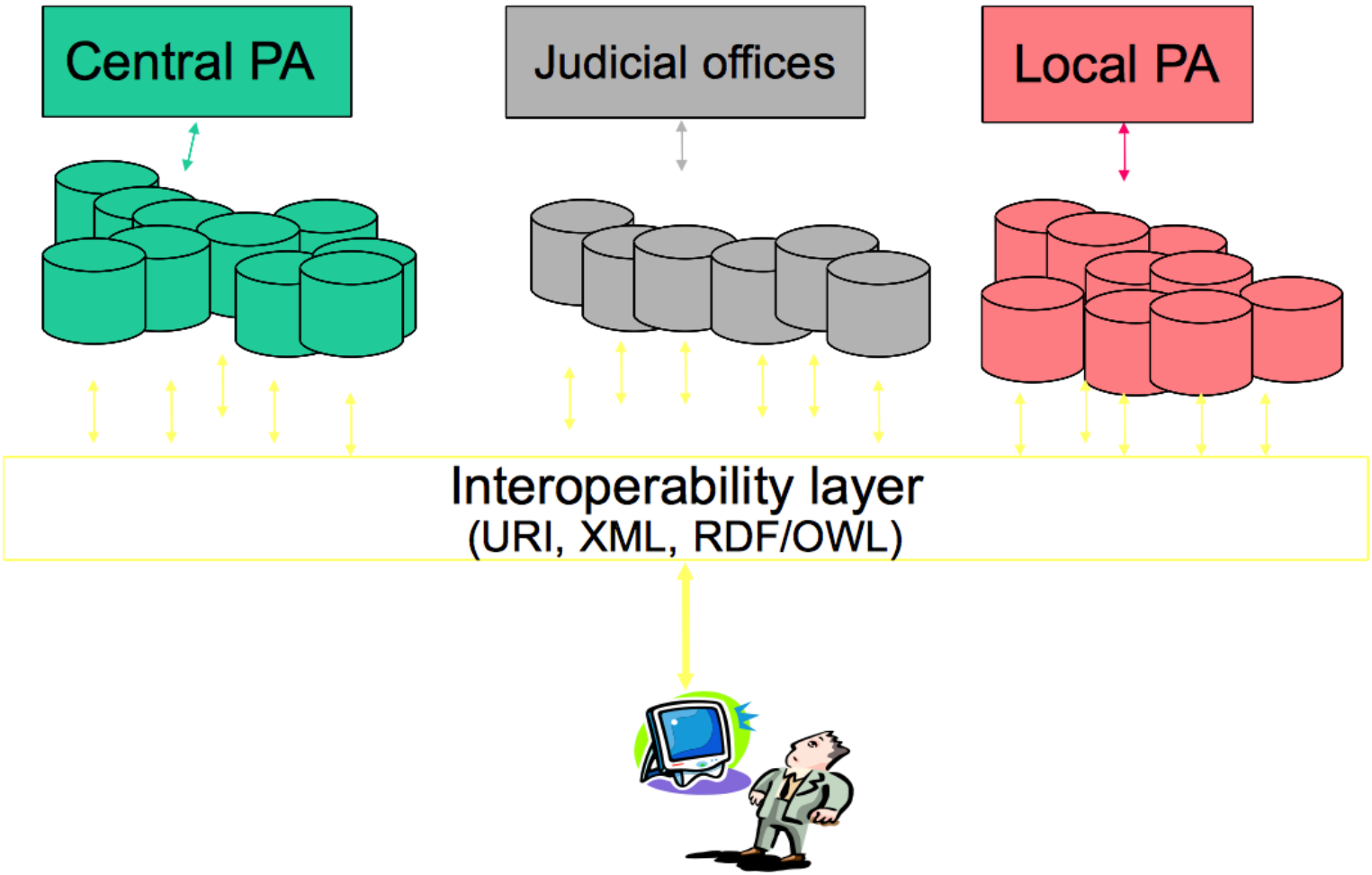
- Which version of law n. 123 issued on 15 March 2007 was in force on 1 December 2010?
- In which laws Mr. XY is the first signer?
- Which laws on consumer protection do apply in a [specific region]?
- Which are the Rights of the Consumer according to the EU law?
- Which are the Implicit Rights of the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

Users' Information Needs in the EU Legal Domain

Cross-Language Accessibility

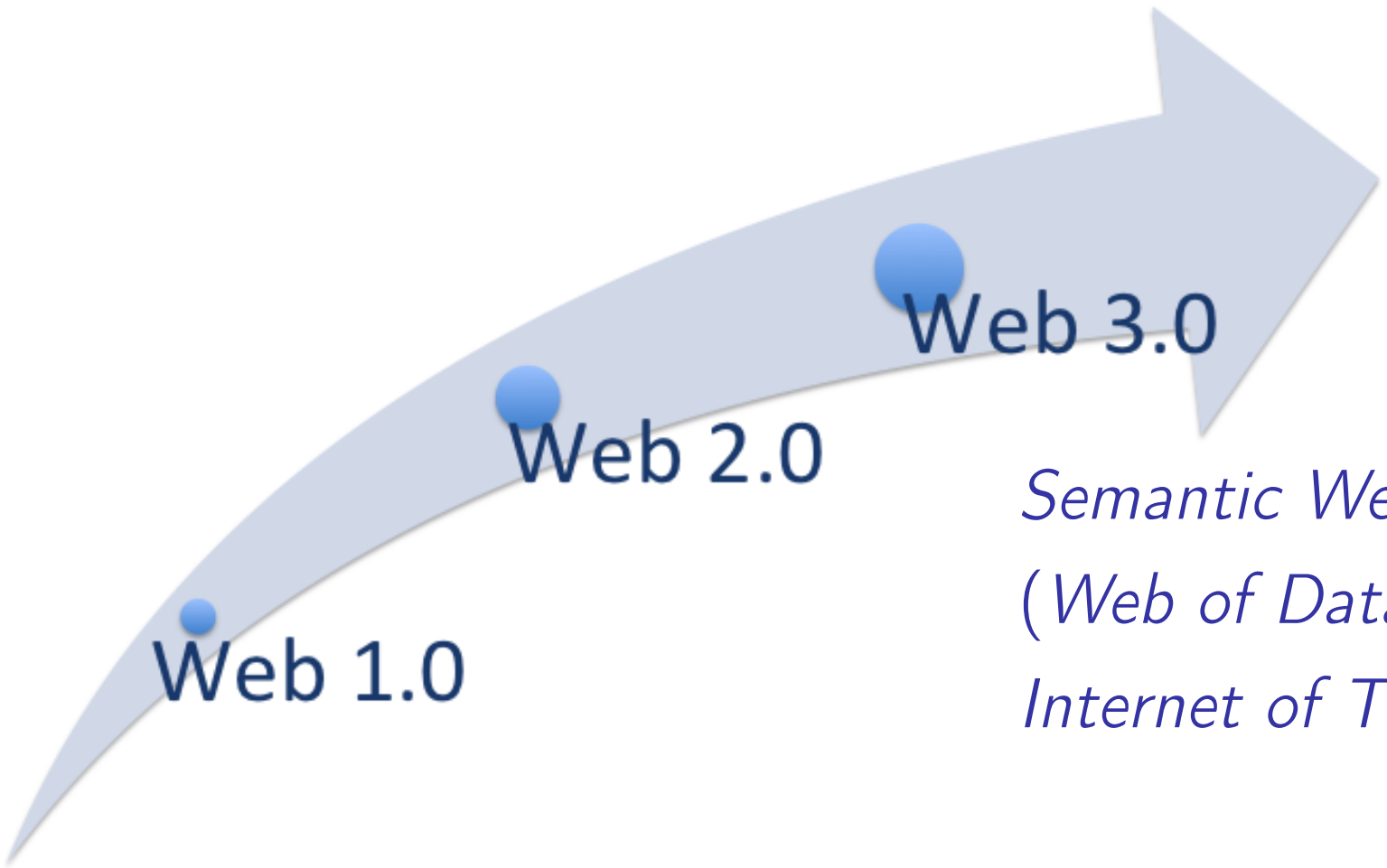
Accessing heterogeneous data sources without language barriers

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



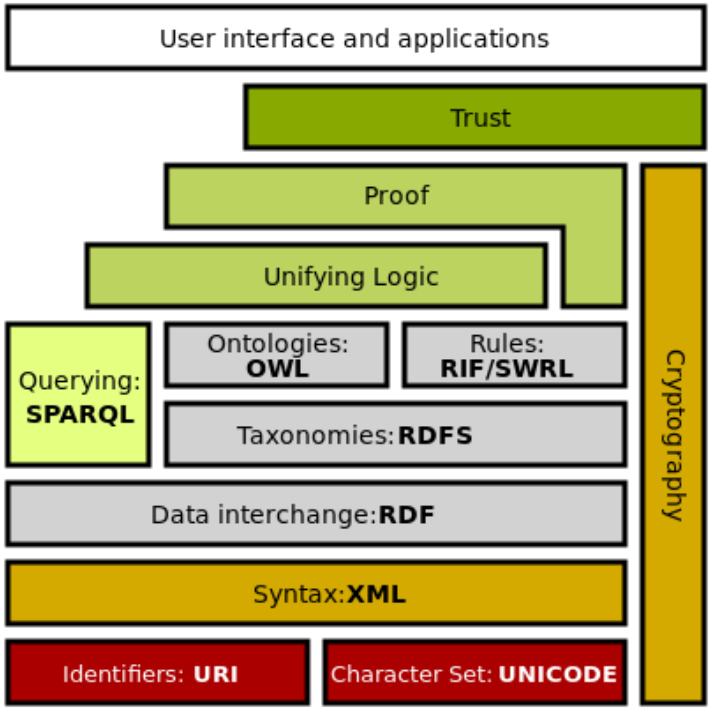
Semantic Web
(*Web of Data*,
Internet of Things)

Semantic Web (Web 3.0, Web of Data, Internet of Things)

The **Semantic Web**

The process of embedding in the World-Wide Web information that is

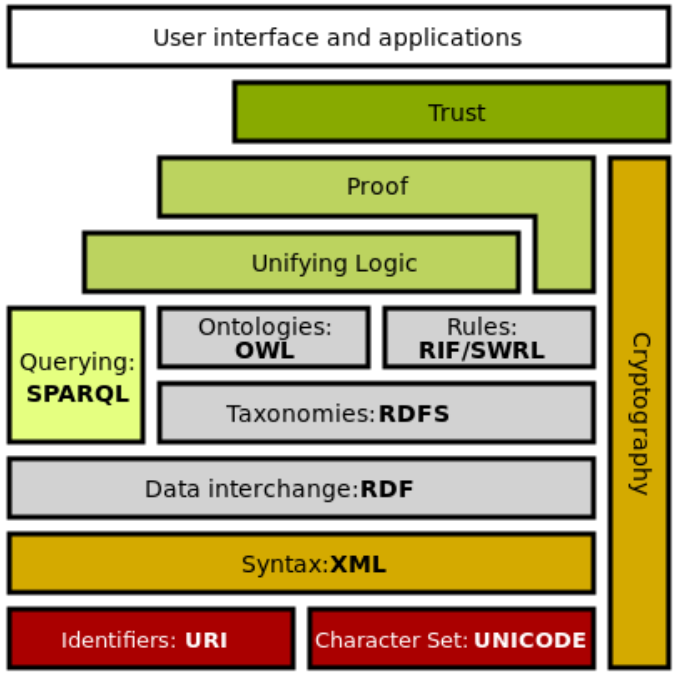
- understandable by humans
- processable and understandable by machines



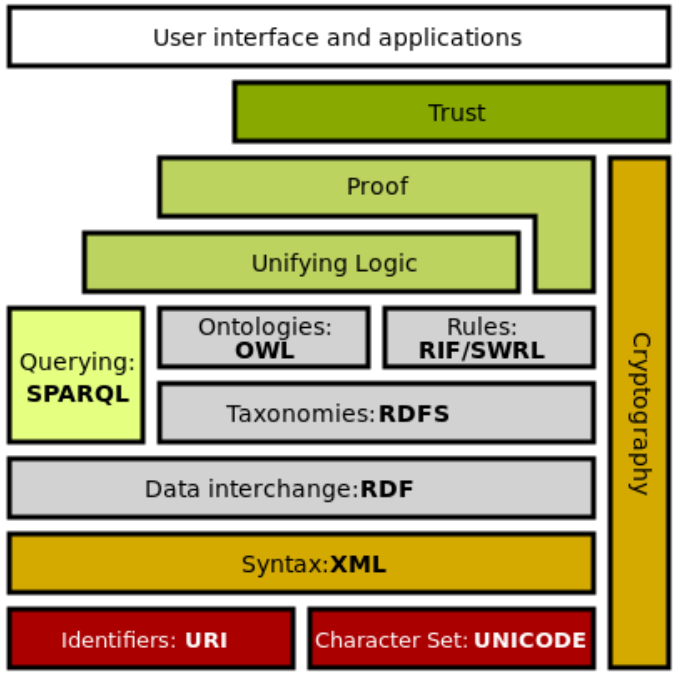
Objectives

- Technological, Semantic and Multilingual Interoperability between information systems
- Advanced access services

The Semantic Web Layers



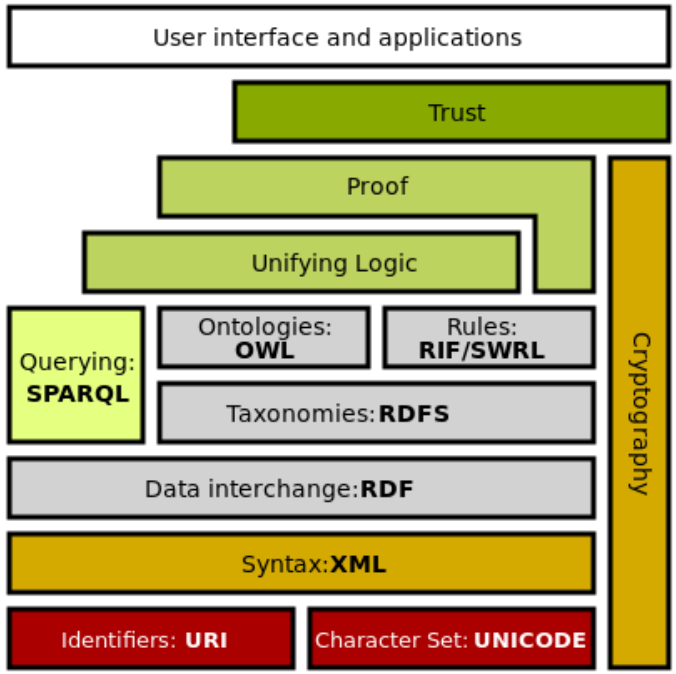
The Semantic Web Layers



- URI: uniform resource identifier in the Web

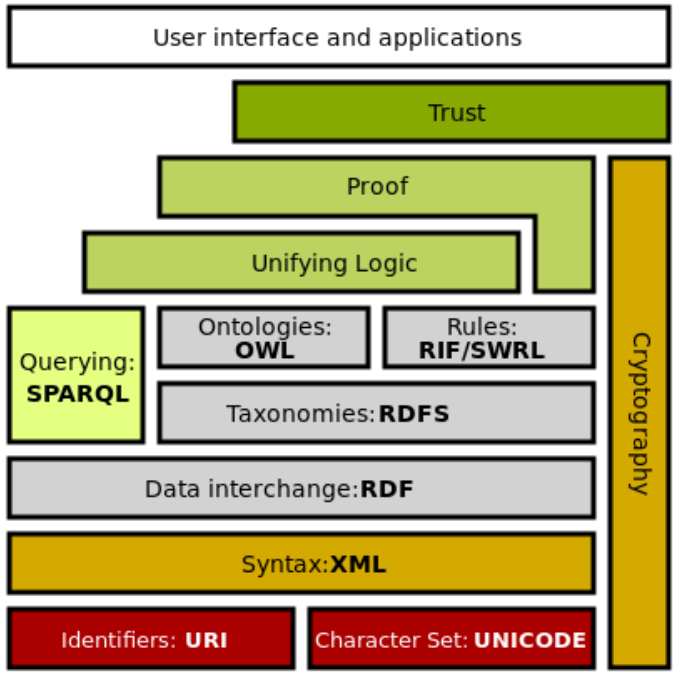
The Semantic Web Layers

- XML: markup syntax for representing structured information
- URI: uniform resource identifier in the Web



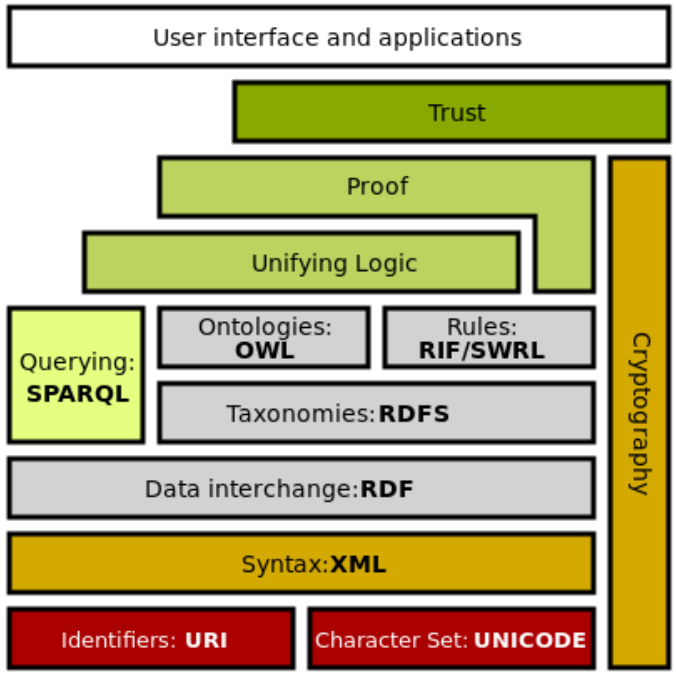
The Semantic Web Layers

- RDF: is a framework for creating statements about resources in a form of triples
- XML: markup syntax for representing structured information
- URI: uniform resource identifier in the Web



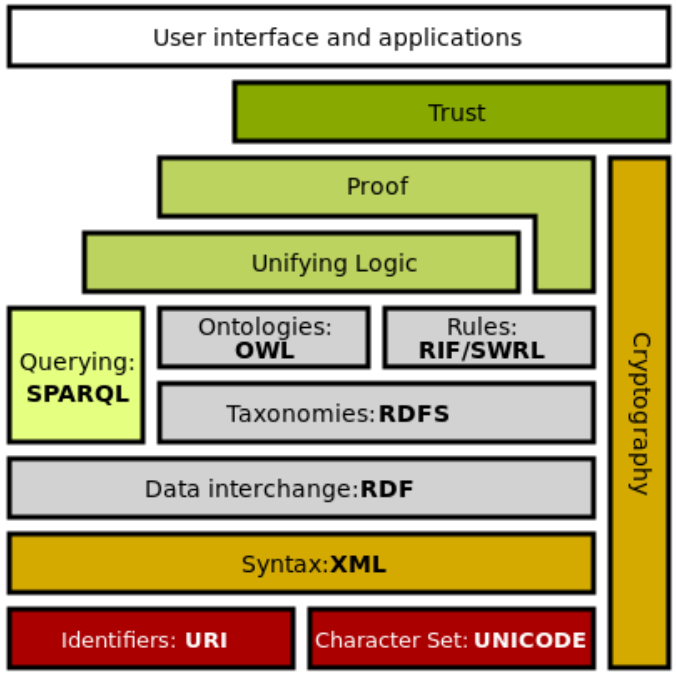
The Semantic Web Layers

- RDFS/OWL: provide basic vocabulary (hierarchies of classes and properties (RDFS), cardinality, properties such as transitivity (OWL)) for RDF
- RDF: is a framework for creating statements about resources in a form of triples
- XML: markup syntax for representing structured information
- URI: uniform resource identifier in the Web



The Semantic Web Layers

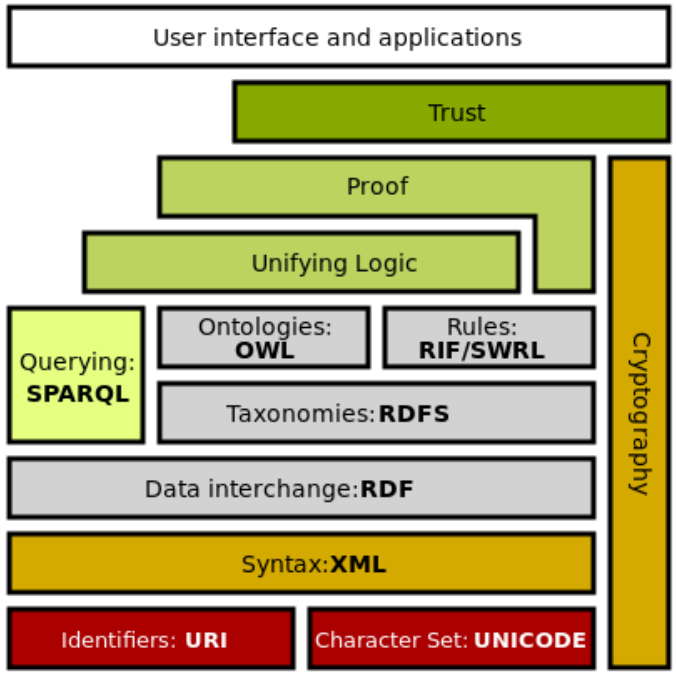
- Logic (RIF or SWRL): bring support to describe rules
- RDFS/OWL: provide basic vocabulary (hierarchies of classes and properties (RDFS), cardinality, properties such as transitivity (OWL)) for RDF
- RDF: is a framework for creating statements about resources in a form of triples
- XML: markup syntax for representing structured information
- URI: uniform resource identifier in the Web



The Semantic Web Layers

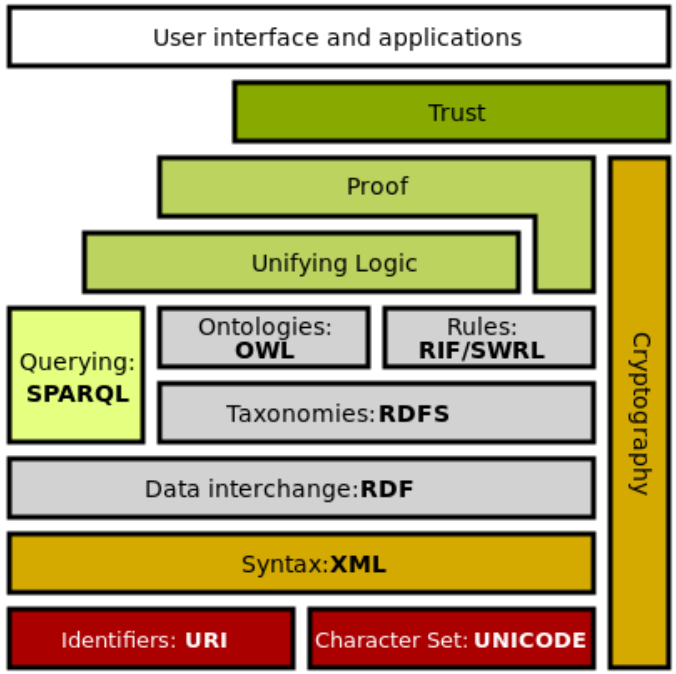
- Proof: proof that an answer found in the Semantic Web is correct:
  - how has it been derived - logic
  - on which data - sources
  - by whom - chain of data providers needs to be considered

- Logic (RIF or SWRL): bring support to describe rules
- RDFS/OWL: provide basic vocabulary (hierarchies of classes and properties (RDFS), cardinality, properties such as transitivity (OWL)) for RDF
- RDF: is a framework for creating statements about resources in a form of triples
- XML: markup syntax for representing structured information
- URI: uniform resource identifier in the Web



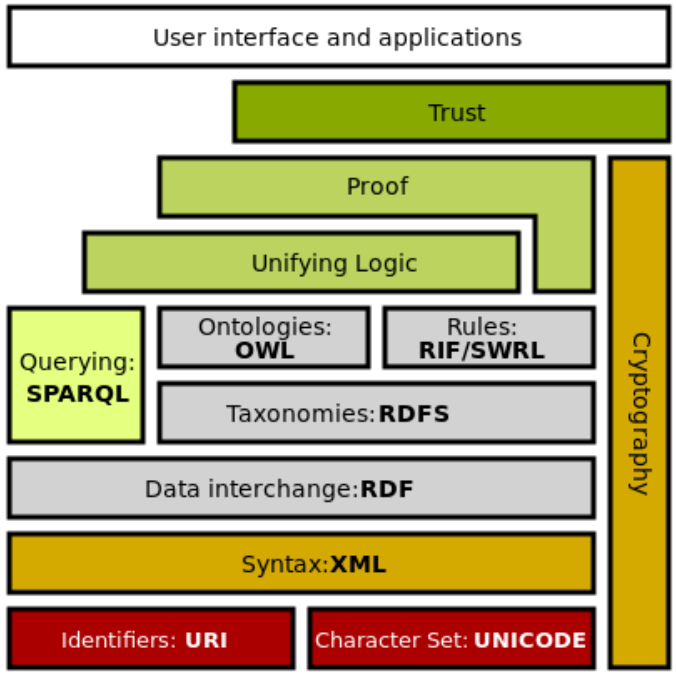
The Semantic Web Layers

- Trust: how to enable autonomous communicating parties to achieve bi-literal agreements? No single authority, no single operational model of trust management in the Web (digital signature, block chain, etc.)
- Proof: proof that an answer found in the Semantic Web is correct:
  - how has it been derived - logic
  - on which data - sources
  - by whom - chain of data providers needs to be considered
- Logic (RIF or SWRL): bring support to describe rules
- RDFS/OWL: provide basic vocabulary (hierarchies of classes and properties (RDFS), cardinality, properties such as transitivity (OWL)) for RDF
- RDF: is a framework for creating statements about resources in a form of triples
- XML: markup syntax for representing structured information
- URI: uniform resource identifier in the Web



The Semantic Web Layers

- Trust: how to enable autonomous communicating parties to achieve bi-literal agreements? No single authority, no single operational model of trust management in the Web (digital signature, block chain, etc.)
- Proof: proof that an answer found in the Semantic Web is correct:
  - how has it been derived - logic
  - on which data - sources
  - by whom - chain of data providers needs to be considered
- Logic (RIF or SWRL): bring support to describe rules
- RDFS/OWL: provide basic vocabulary (hierarchies of classes and properties (RDFS), cardinality, properties such as transitivity (OWL)) for RDF
- RDF: is a framework for creating statements about resources in a form of triples
- XML: markup syntax for representing structured information
- URI: uniform resource identifier in the Web



Voluntary adhesion up to the most convenient interoperability level

Semantic Web Initiatives in the Legal Domain

- Eur-Lex (Formex)
- National initiatives
  - NormeInRete - Normattiva (Italy)
  - JURICONNECT (The Netherlands)
  - LexDania (Denmark)
  - CHLexML (Switzerland)
  - eLaw (Austria)
- Extra-European initiatives
  - Senado Federal do Brasil
  - AkomaNtoso (Pan African Parliaments)
  - En-Act project (Tasmanian Government)
- Pan-European initiative
  - CEN-Metalex

MetaLex/CEN

Legal Identifiers

- URN-LEX (URL-LEX) naming convention

- AkomaNtoso naming convention

- ECLI and ELI

Legal XML Schemas

- Formex

- CEN Metalex

- AkomaNtoso

Excerpt of the Structure of Directive 2002/65/EC, represented in CEN Metalex compliant XML

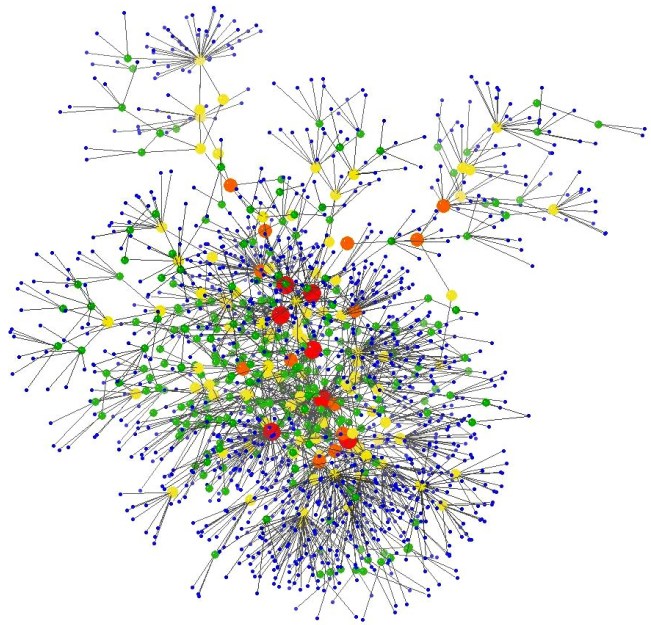
```
<article id="art5">
  <paragraph id="art5-par1">
    1. The supplier shall communicate to the consumer all the
    contractual terms and conditions and the information referred
    to in Article 3(1) and Article 4 [...]
  </paragraph>
  <paragraph id="art5-par2">
    2. The supplier shall fulfil his obligation under paragraph 1
    immediately after the conclusion of the contract, if the
    contract has been concluded at the consumer's request using
    a means of distance communication which does not enable
    providing the contractual terms [...]
  </paragraph>
  <paragraph id="art5-par3">
    3. At any time during the contractual relationship the
    consumer is entitled, at his request, to receive the
    contractual terms and conditions on paper. [...]
  </paragraph>
</article>

<article id="art6">
  <paragraph id="art6-par1">
    1. The Member States shall ensure that the consumer shall have
    a period of 14 calendar days to withdraw from the contract
    without penalty and without giving any reason [...]
  </paragraph>[...]
</article>
```

Modeling Legal Concepts

- Legal concepts modeling is essential for implementing the Semantic Web in the legal and multilanguage domain

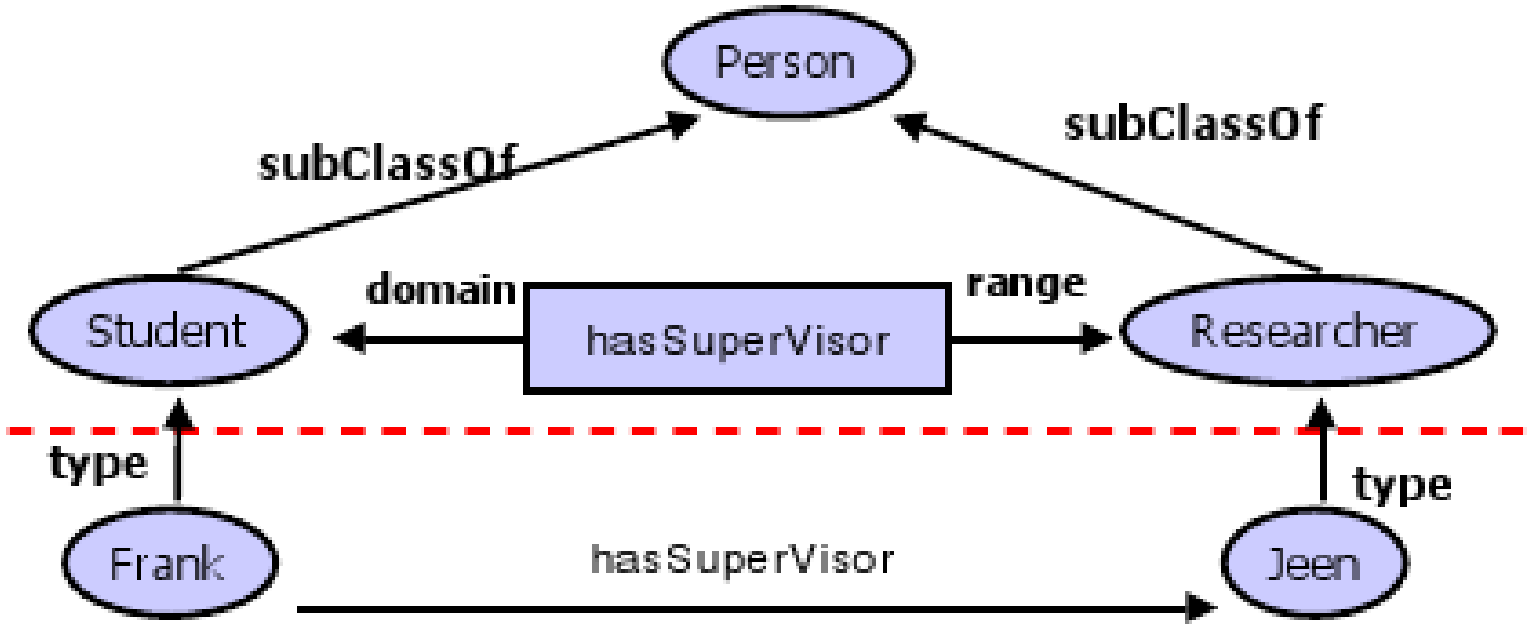
- Knowledge organization systems
  - Thesauri
    (Eurovoc, ETT, Eclas, Gemet, etc.)
  - Semantic lexicons
    (WordNet, Syllabus, etc.)
  - Legal Ontologies
    (LRI-Core, LKIF, CLO, Dalos, etc.)

- Modeling strategy in a multilingual and multicultural domain
  - Collaborating platform connecting
    - Legal comparatists
    - Translators
    - Ontology developers



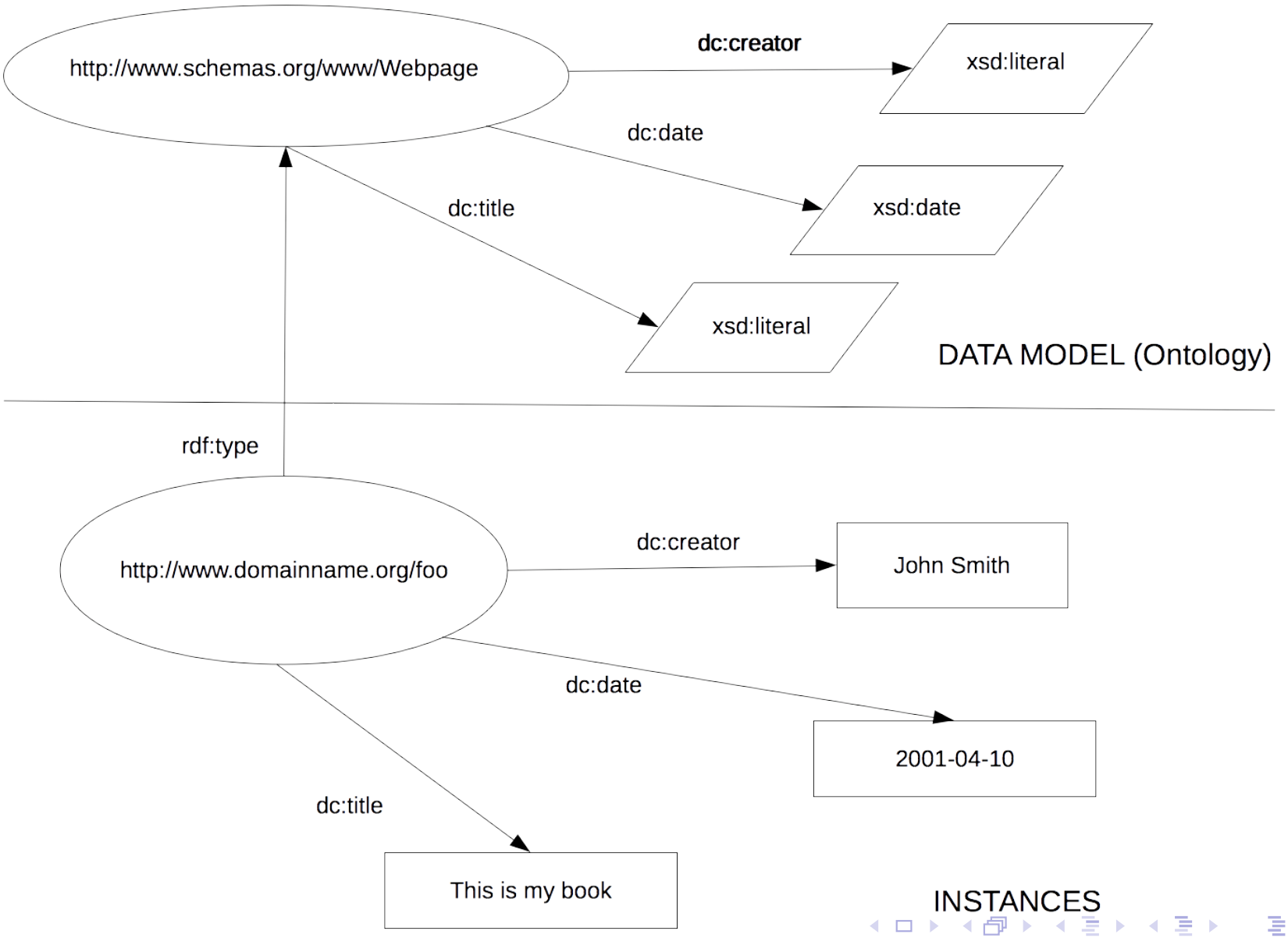
Knowledge Models and Instances

RDFS/OWL (Knowledge Models / Ontologies)



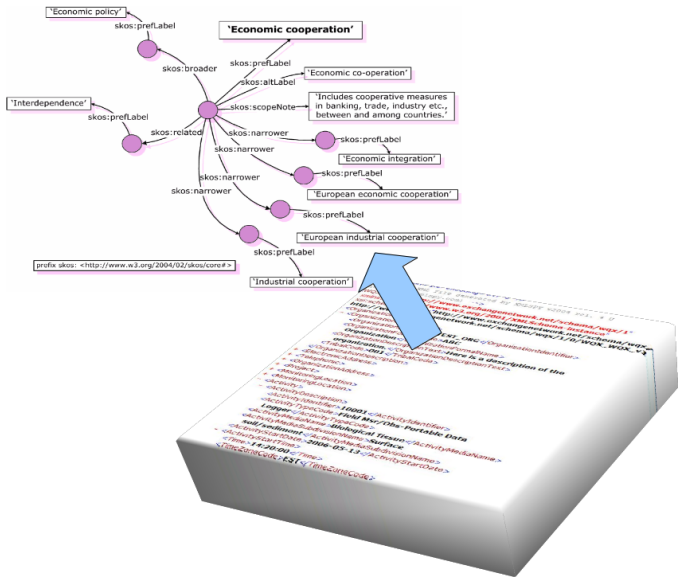
RDF (Instances / Individuals)

Knowledge Models and Instances



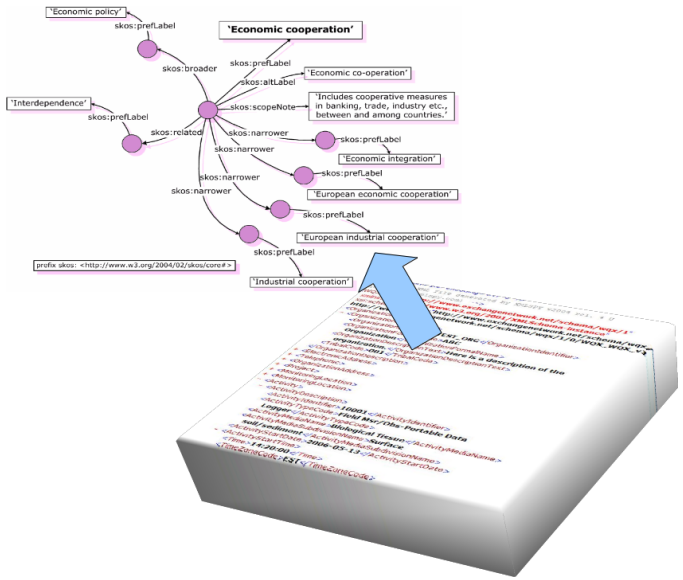
Semantic Web to overcome semantic and language barriers

Semantic mark-up of multilingual documents

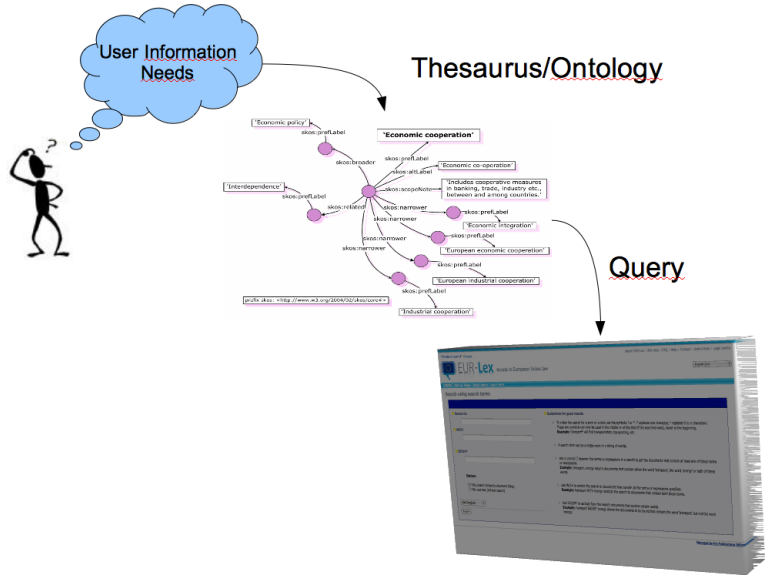


Semantic Web to overcome semantic and language barriers

Semantic mark-up of multilingual documents



Semantic tools (thesauri, ontologies) to express the semantics of users' information needs



Semantic Web: The crisis of the Top-Down approach

The top-down approach needs:

- Standardization activities in working groups
- Wide coordination and economic efforts of the involved actors to adopt the proposed standards and models

Benefits		**Drawbacks**
Relevance of research achievements	$\Longleftrightarrow$	Results on data exposition not comparable to such achievements

Semantic Web: the Bottom-up approach (Linked Data)

Based on exactly the same technological stack and principles of the Semantic Web

Guidelines for implementing the Semantic Web to enable data sharing and reuse on a massive scale:

1. Use URIs as names for things.
2. Use HTTP URIs, so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL).
4. Include links to other URIs, so that the user can discover more things.

Linked Data objectives

1. Exposing data according to the Linked Data guidelines at the available level of interoperability
2. Including links where possible
3. Leaving the effort of semantic enrichment and interconnection to a virtuous trend stimulated by data consuming

Linked Open Data (LOD)

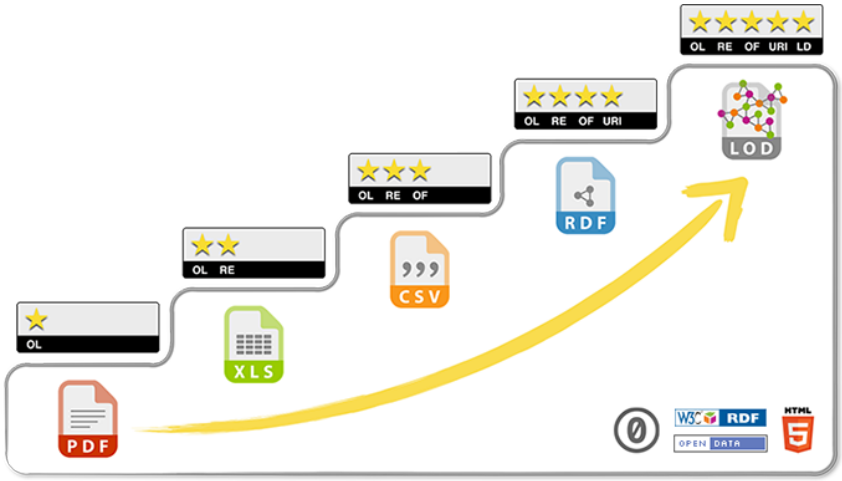
Linked Data released under an open licence, which does not impede their reuse for free.
Creative Commons CC-BY-SA is an example open license

Open Data, Linked Data, and the Semantic Web

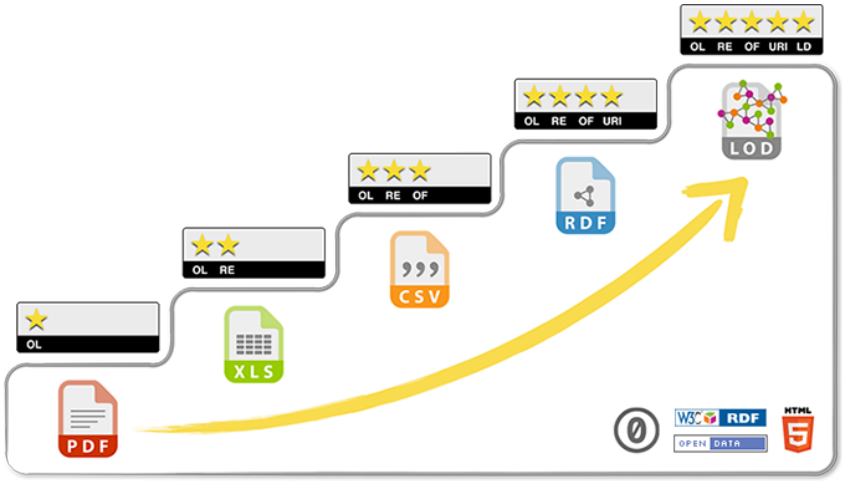
The Semantic Web (Web 3.0) is made up of Linked Open Data

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

5-star Rating Scheme for Linked Open Data

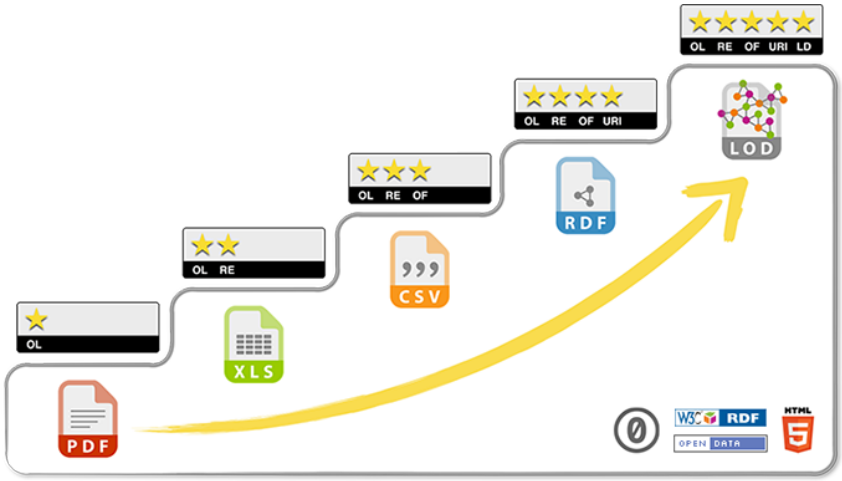


5-star Rating Scheme for Linked Open Data



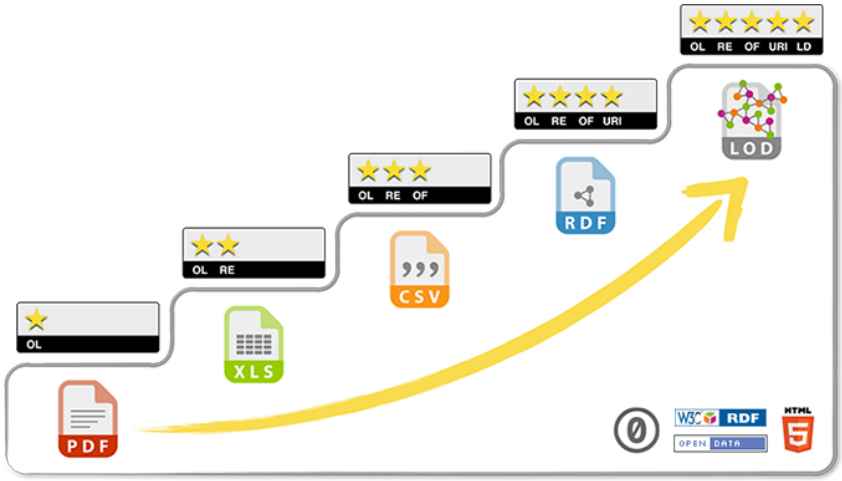
- Available on the Web (whatever format) but with an Open License (to be Open Data)

5-star Rating Scheme for Linked Open Data



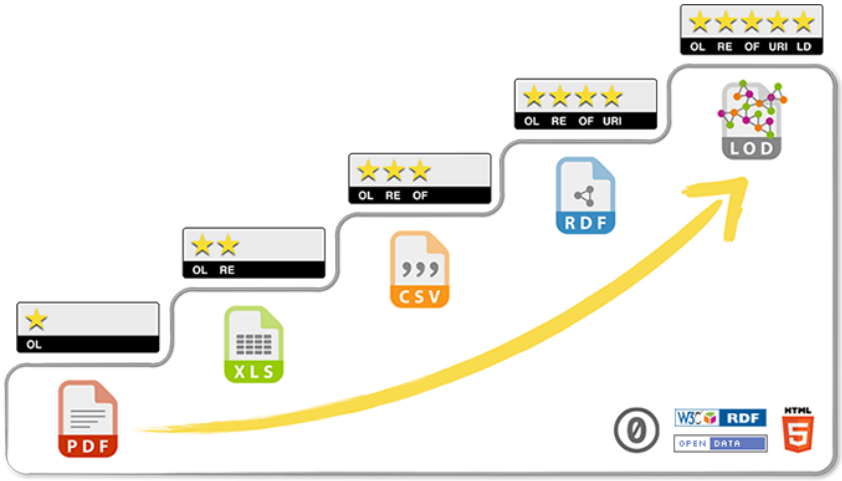
- ⋆⋆ Available as machine-readable structured data (e.g. excel instead of image scan of a table)
- ⋆ Available on the Web (whatever format) but with an Open License (to be Open Data)

5-star Rating Scheme for Linked Open Data



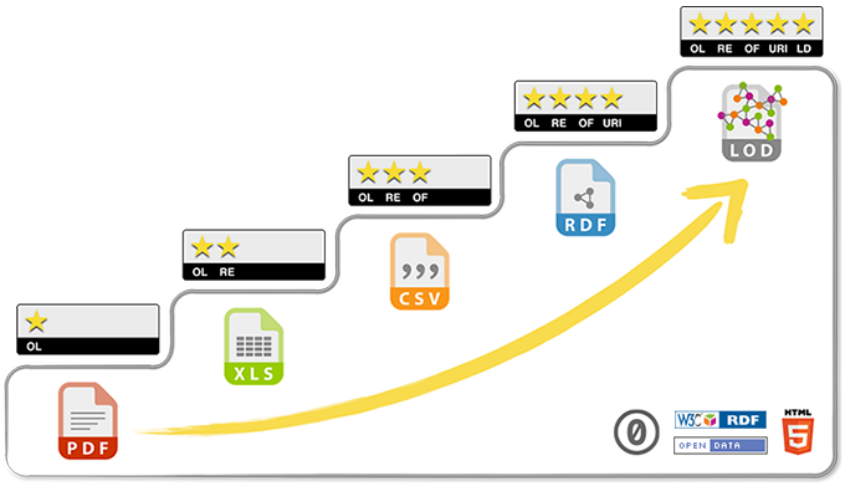
- ⋆ ⋆ ⋆ Use non-proprietary format (e.g. CSV instead of excel)
- ⋆⋆ Available as machine-readable structured data (e.g. excel instead of image scan of a table)
- ⋆ Available on the Web (whatever format) but with an Open License (to be Open Data)

5-star Rating Scheme for Linked Open Data



- ⋆ ⋆ ⋆⋆ 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)

5-star Rating Scheme for Linked Open Data



★ Available on the Web (whatever format) but with an Open License (to be Open Data)

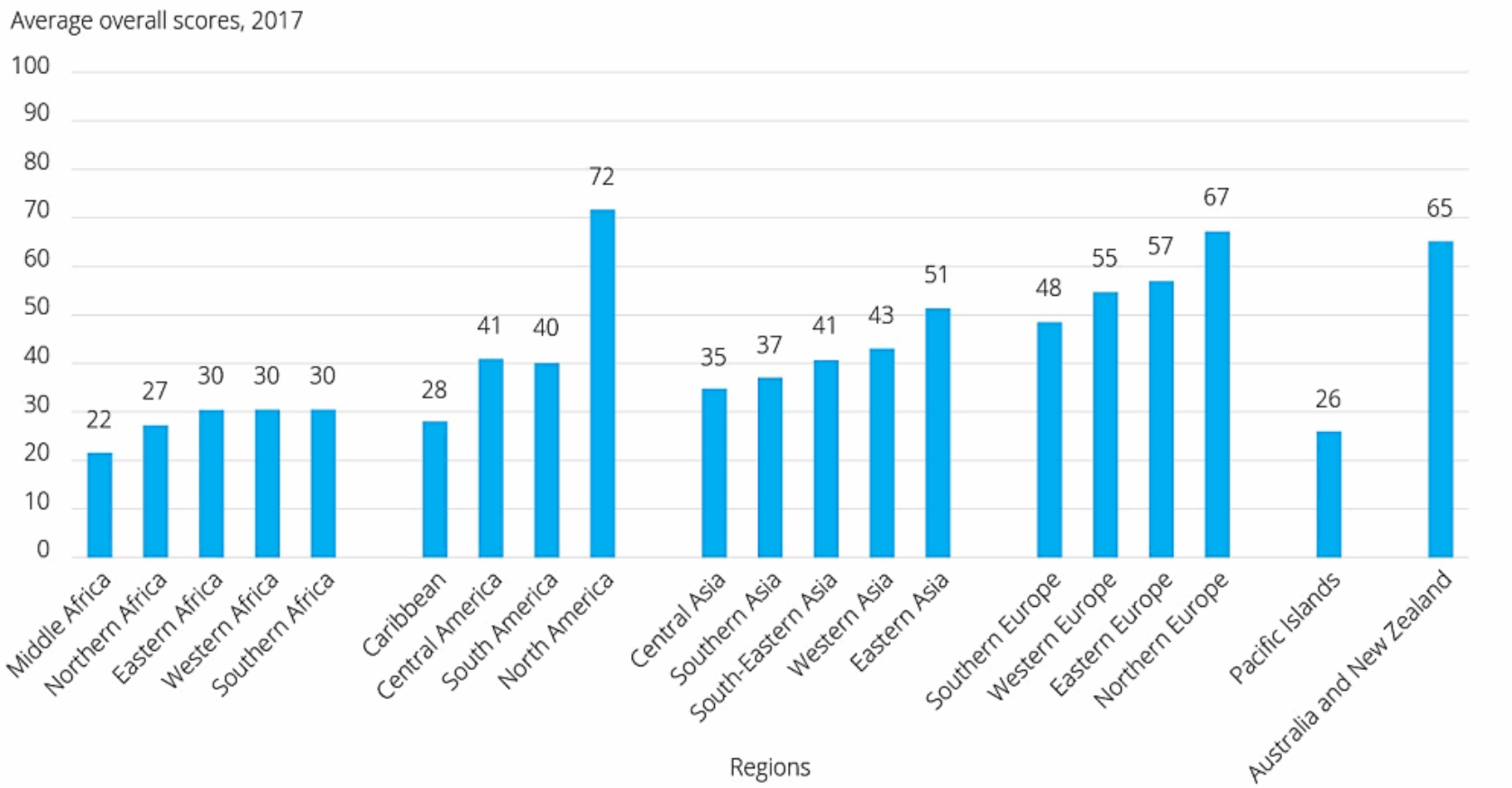
★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table)

★★★ Use non-proprietary format (e.g. CSV instead of excel)

★★★★ Use open standard from W3C (URI, RDF and SPARQL) to describe things, so that people can point at them

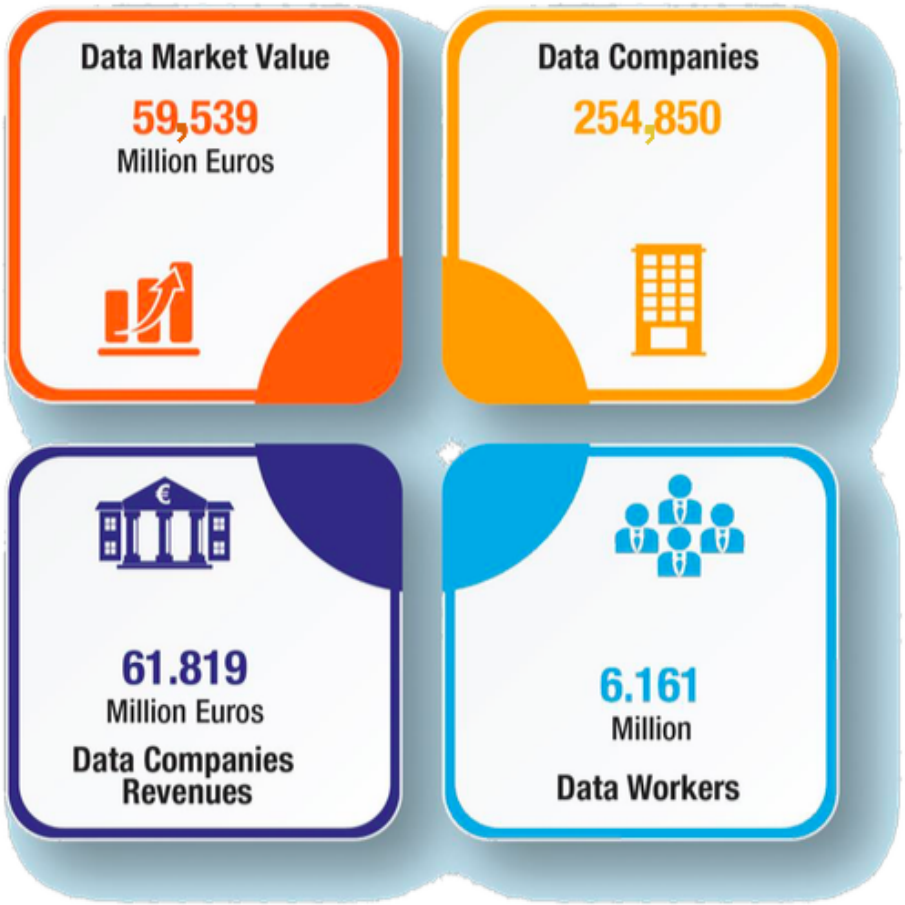
★★★★★ Link your data to other people's data to provide context

Open Data Inventory 2017



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

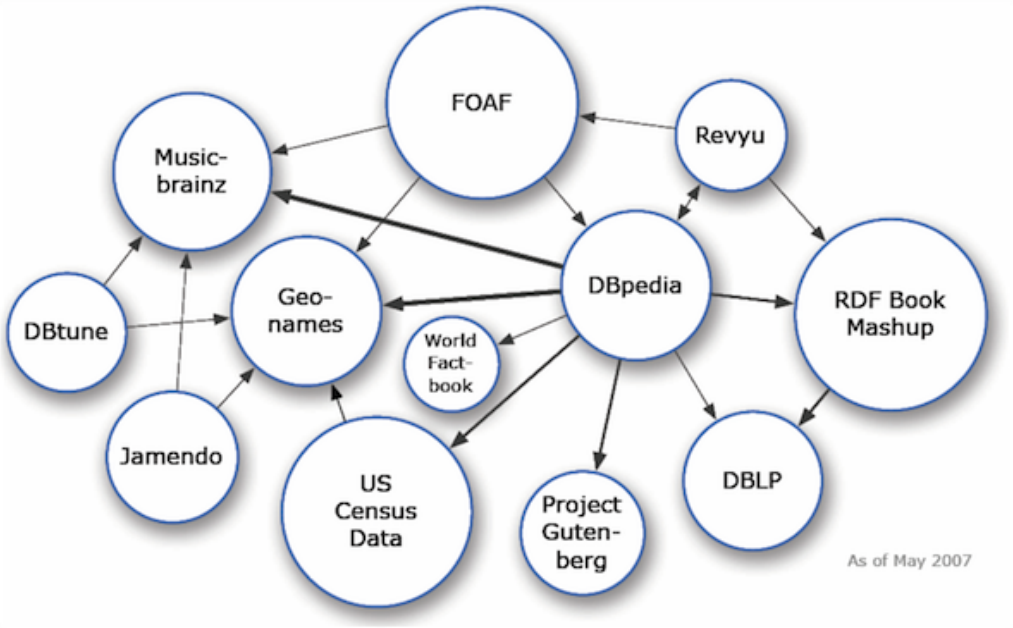
Open Data in Economic Growth

Open Data economy is supposed to generate

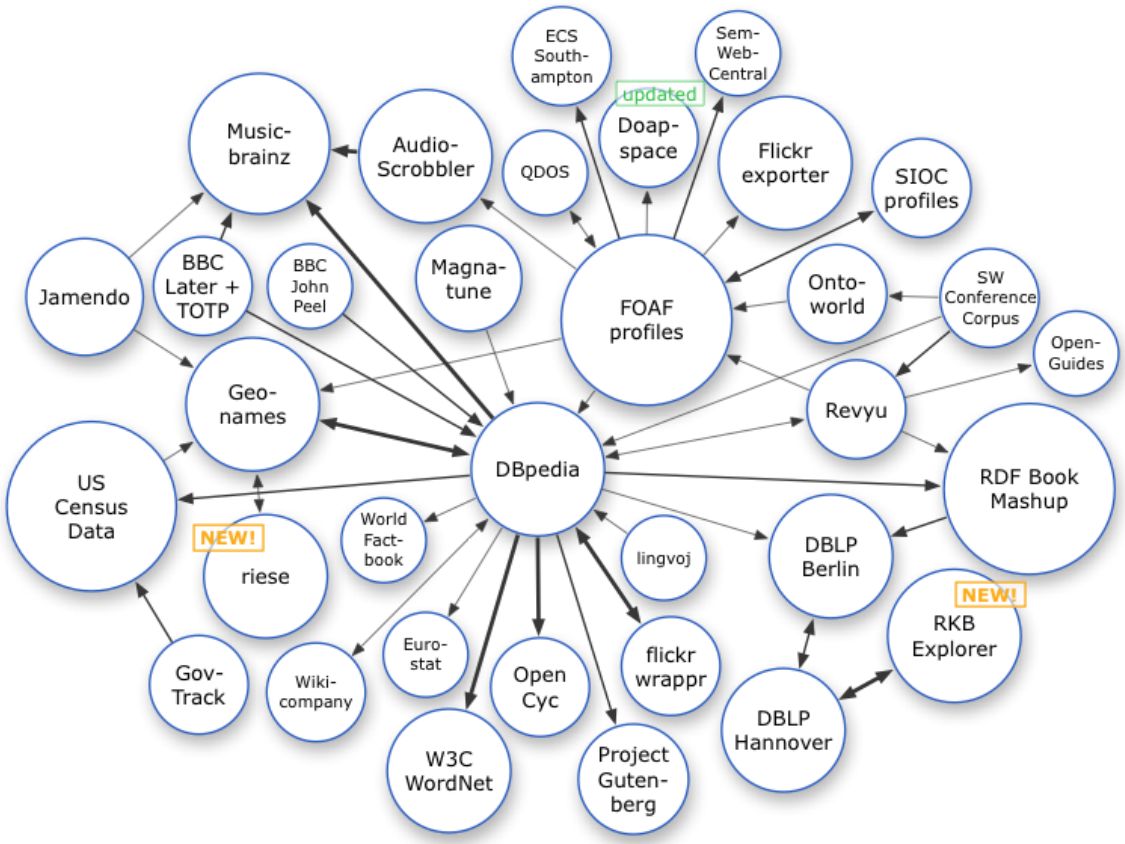
- an additional growth up to 4% of the GDP by 2020
- a growth in cumulative revenues in the period 2016-2020 estimated to 110 million Euro

Benefits of Open Data

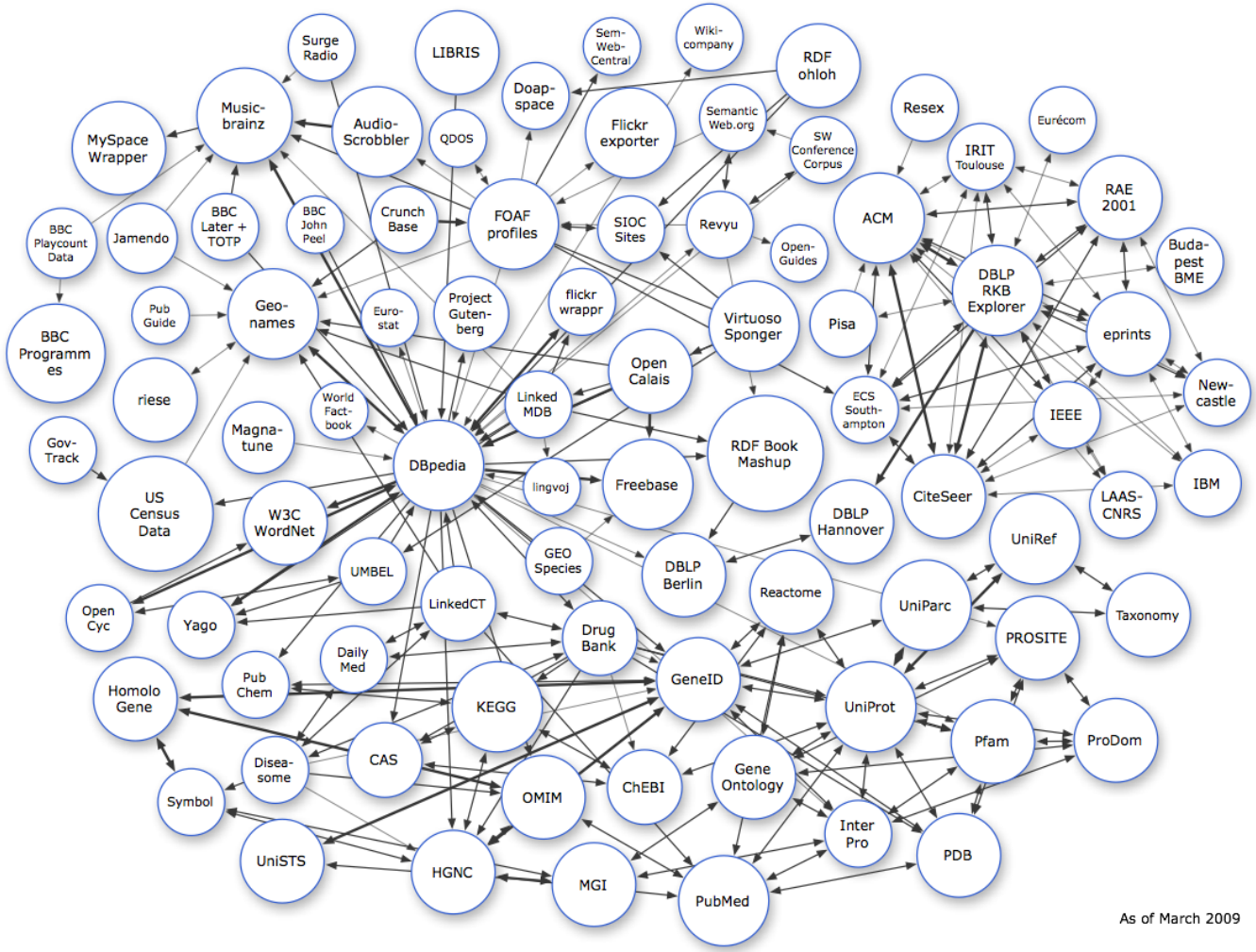
Linked Open Data growth



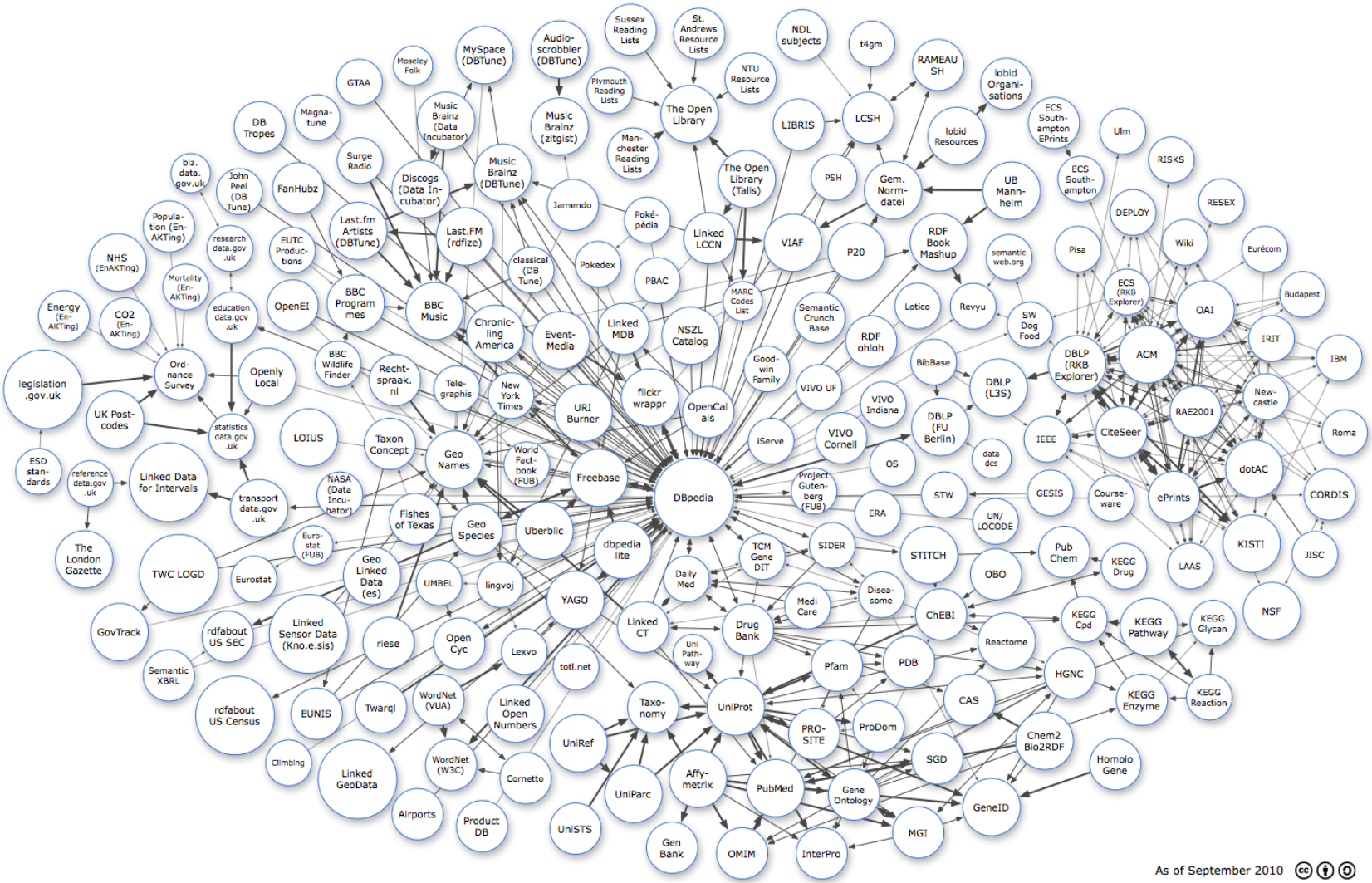
Linked Open Data growth



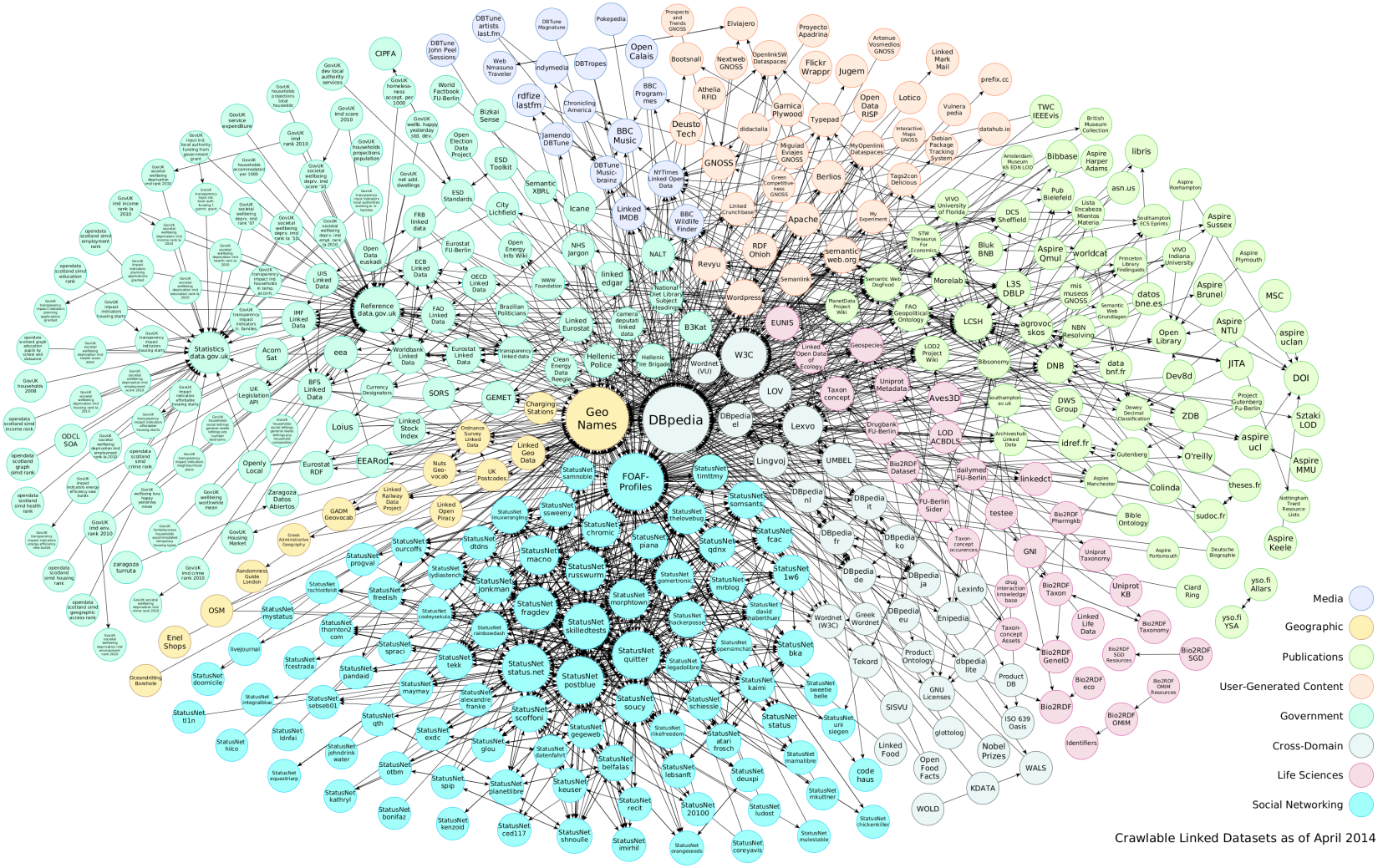
Linked Open Data growth



Linked Open Data growth



Linked Open Data growth



Crawlable Linked Datasets as of April 2014

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)

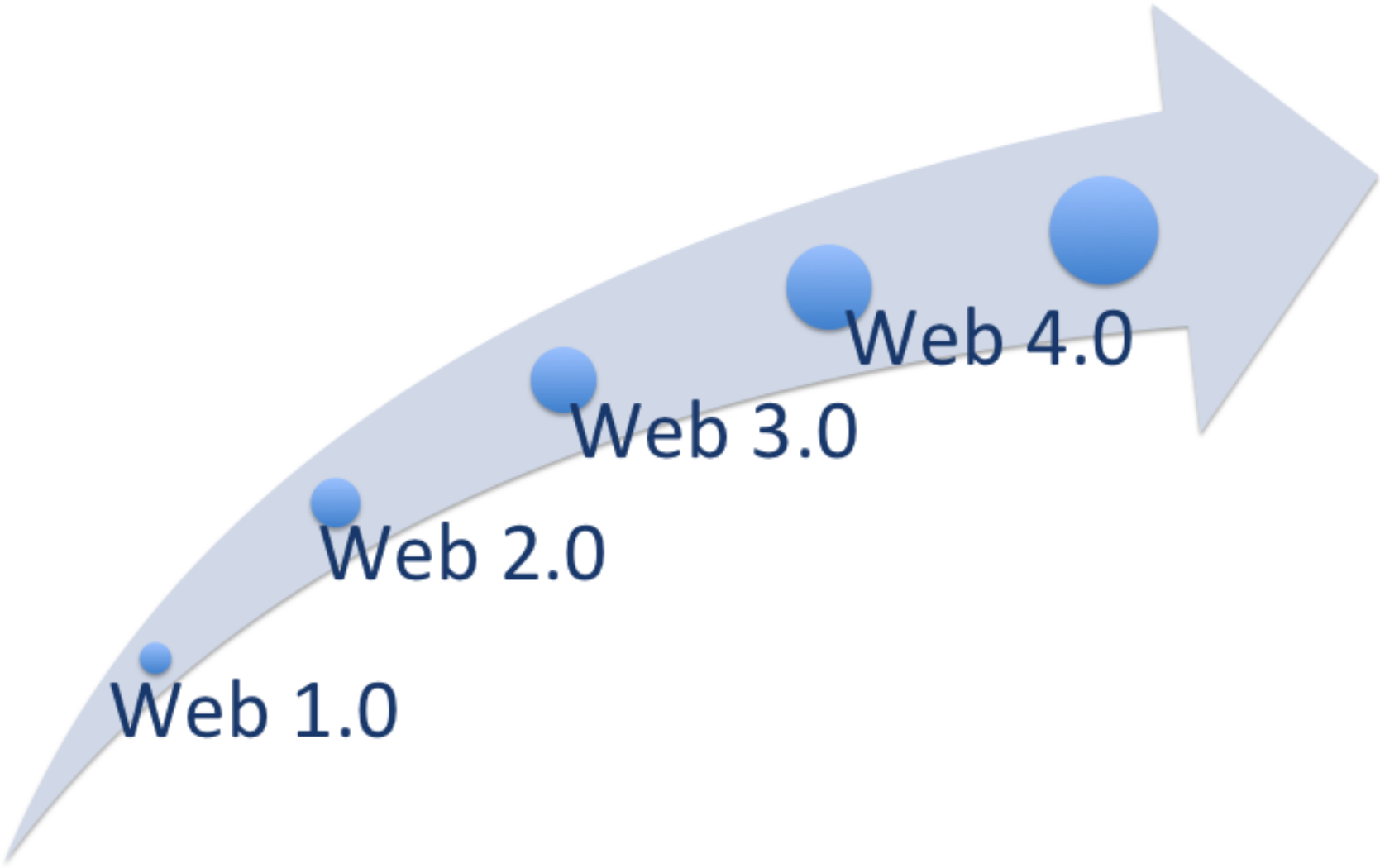
In conjunction with eGov initiatives, they are about: Legal Data Accessibility, Quality, Reusability for Citizens

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 implement advanced reasoning
  - Premises $\Longrightarrow$ Conclusions
  - Accessing Implicit Knowledge

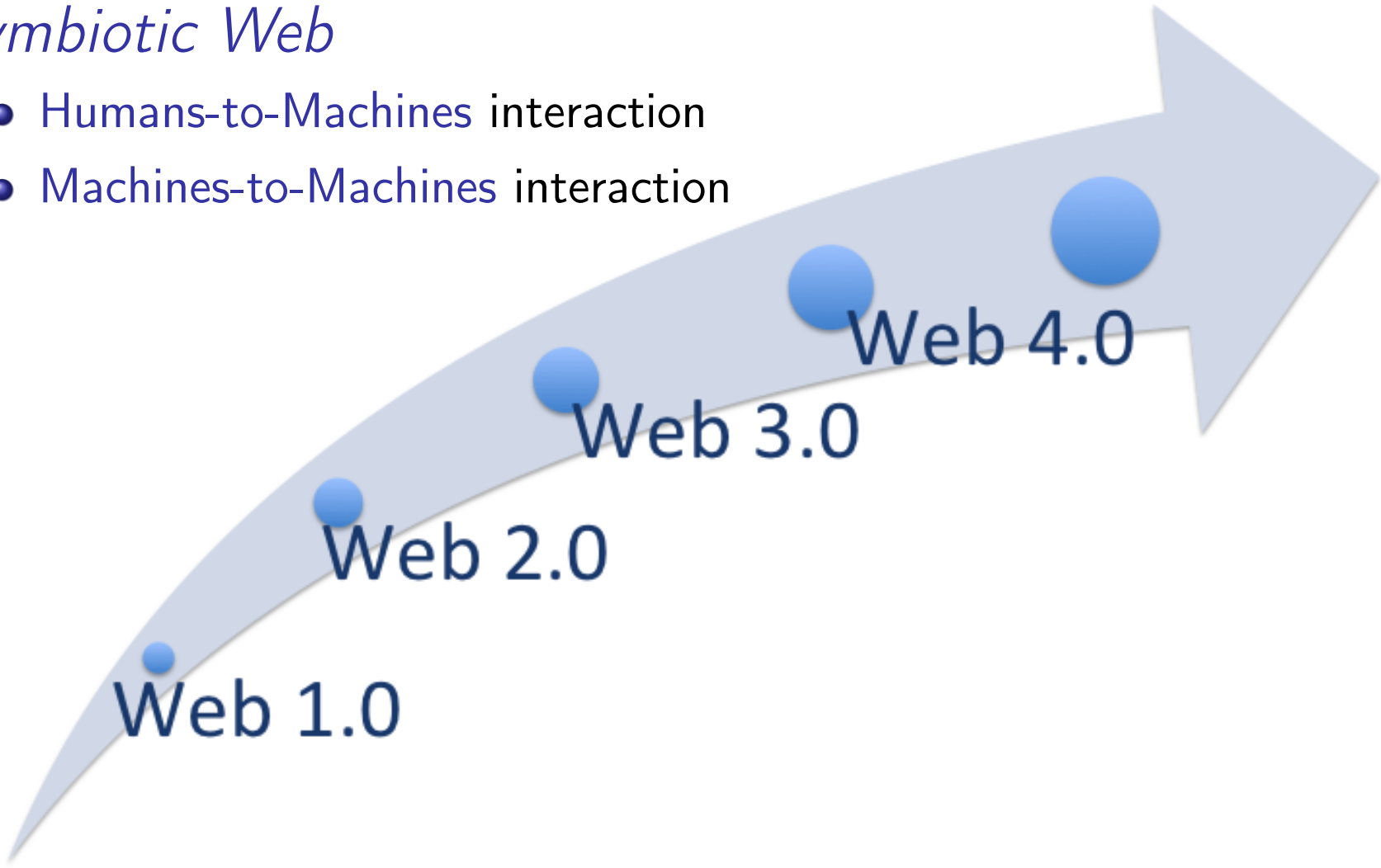
What's next? The Web 4.0



What's next? The Web 4.0

Symbiotic Web

- Humans-to-Machines interaction
- Machines-to-Machines interaction



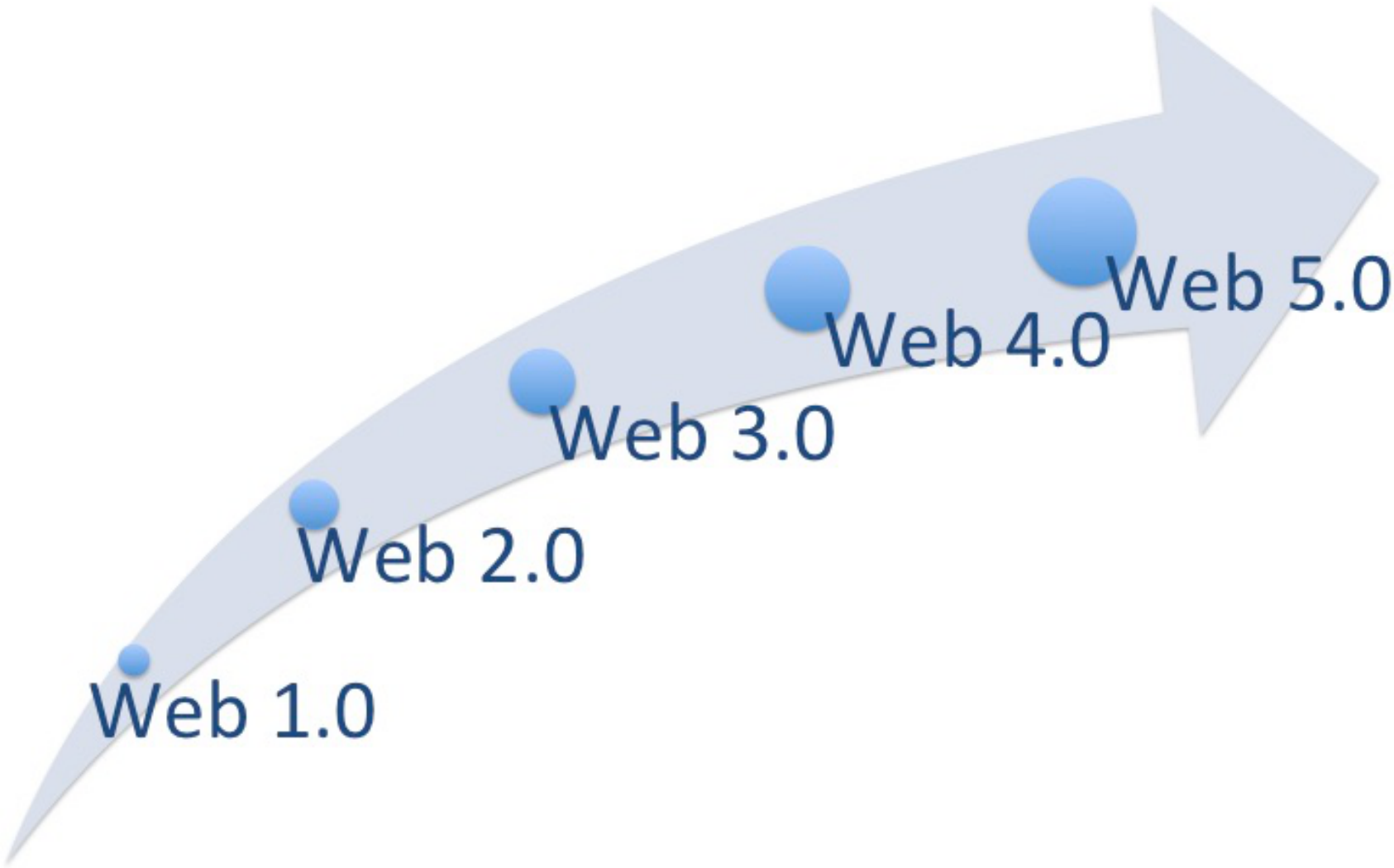
Web 4.0 for eLaw and eJustice

- In Web 3.0

  Law understandable and processable by machines

- In Web 4.0
  - Intelligent Agents for Legal data mining and e-Discovery
  - Digital Judges with knowledge of personal profiles, specific cases and laws, taking decisions on legal disputes

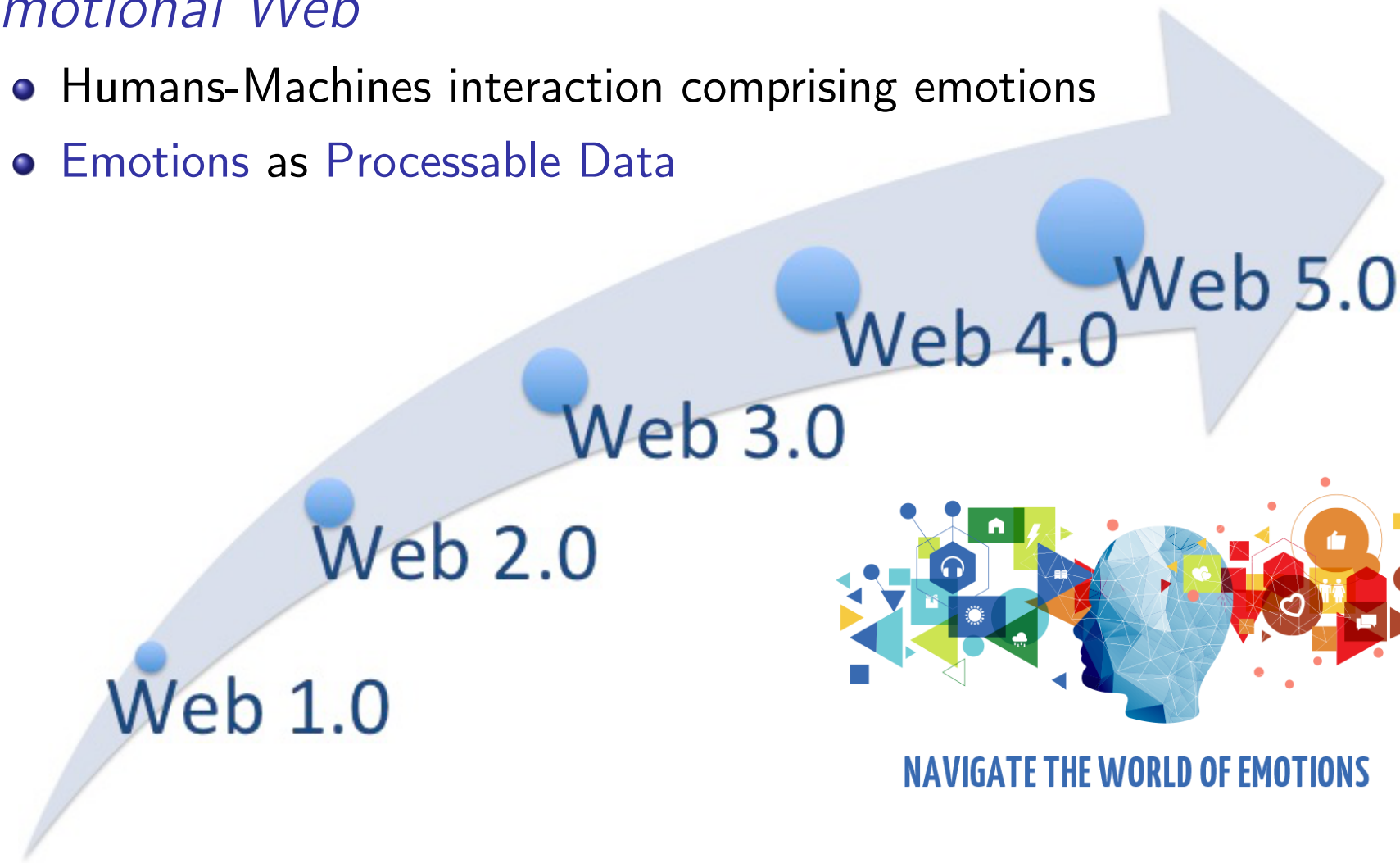
Nowadays the Web is *emotionally neutral*: next? Web 5.0



Nowadays the Web is *emotionally neutral*: next? Web 5.0

Emotional Web

- Humans-Machines interaction comprising emotions
- Emotions as Processable Data



NAVIGATE THE WORLD OF EMOTIONS

In the Web 5.0 how will you persuade a Digital Judge?

Thanks for your attention!

enrico.francesconi@ittig.cnr.it
enrico.francesconi@publications.europa.eu