**Semantic web**

- **Linked Data:** Describes a method of publishing structured data so that it can be interlinked and become more useful. It builds upon standard Web technologies such as HTTP, RDF and URIs, but rather than using them to serve web pages for human readers, it extends them to share information in a way that can be read automatically by computers. This enables data from different sources to be connected and queried.
Open Data requires that certain data are of free access to all, with no technical or legal limitations, with licences and terms of use subject to laws on the reuse of public sector information (PSI).

Free of conditions as long as the data are kept intact and not manipulated, the source is cited and the date the data were last updated is indicated.

Tim Berners-Lee (2009) defined the principles for publishing RDF (Resource Description Framework) data on the Web:

- Use URIs (Uniform Resource Identifier) as unique names for things
- Use HTTP URIs so that people can look up those names
- When someone looks up a URI, provide useful information, using Web standards such as RDF, and SPARQL
- Include links to other URIs, so that they can discover more things

Usually LOD is grouped in datasets, equipped with ontology.
“With linked data, when you have some of it, you can find other, related, data.”

Tim Berners-Lee, “Linked Data Design Issues”
http://www.w3.org/DesignIssues/LinkedData.html

The Open Definition sets out principles that define “openness” in relation to data and content.

“A piece of data or content is open if anyone is free to use, reuse, and redistribute it - subject only, at most, to the requirement to attribute and/or share-alike.”

http://opendefinition.org/

• Revision June 2013

  – introduces a genuine right to reuse by making reusable all content that can be accessed under national access to documents laws;
  – lowers the upper ceiling for charges on reuse applicable in standard cases to marginal costs;
  – expands the scope of application of the Directive to certain cultural institutions such as libraries (including university libraries), museums and archives, but making them subject to a number of different rules;
  – reinforces the obligation to be transparent on conditions and on charges applied to reuse;
  – invites MSs to make more documents available in machine-readable and open formats.
Legal Linked Open Data

• Legal domain is suitable for LOD:
  – Holistic character, meaning depends on many sources: regulations, case law, doctrine.
  – Network of documents explicitly referred to in other texts
  – Semi-structured character of documents
  – Terminologies, taxonomies and ontologies.

• It can benefit from these technologies:
  – Millions of documents
  – Multi-language
  – Technical language
The problem
EUCases: the problem

Legal professionals (judges and court staff, barristers and solicitors, in-house legal counsel, policy makers and legislators, legal information managers, paralegals, academics, legislators, compliance managers, etc.) operating in Europe more and more experience the need for accessing case law in the different member states of the EU to carry out their work.
When deciding a case which involves applying EU law, a national judge is acting as a guarantor of the European legal order itself. His ruling becomes a decision not only on national laws but also on the application of EU legal provisions, and as such, (s)he needs access to the relevant case law of other MSs.

In its Resolution of 9 July the European Parliament determines the national judges as “the keystone of the European Union judicial system” who play “a central and indispensable role in the establishment of a single European legal order”.
Linking legal open data in Europe
Duration: **2 years (1.10.2013 – 30.09.2015)**

Budget: **€ 1 860 840**

**Call:** FP7 SME-DCA Data analytics initiatives for the SMEs

**Partners:**

- empirica GmbH (Bonn, Germany) - *Coordinator*
- Apis-Hristovich EOOD (Sofia, Bulgaria)
- Institute of ICT (Sofia, Bulgaria)
- Universita degli Studi di Torino (Italy)
- Averbis GmbH (Freiburg, Germany)
- Nomotika SRL (Torino, Italy)
EUCases’ aim

• The goal of the project is to create a unique Pan-European legal platform – **EUCases Linking Platform**, and services linking law and case law of Member States to European Union legal system, thus providing access and search facilities at the European level.

• It transforms multilingual *legal open data* into **linked open data** after **semantic and structural analysis** to improve their usability and retrieval.

• State of the art Natural Language Technologies will be used.
Legal documents

- It will download and reuse the millions of legal documents and open access doctrinal work: European and national legislative and case law (high courts and appeal courts) portals of 6 Member States (BG, DE, FR, IT, UK, and AT): EUR-Lex (EU), Legifrance (FR), legislation.gov.uk (UK), Supreme Court (UK), Bundesgerichtshof (DE), Gesetze im internet (DE), Rechtsinformationssystem (AT), Supreme Court of Cassation (BG), Supreme Administrative Court (BG), Normattiva (IT), Consiglio di Stato (IT), DOAJ (Directory of Open Access Journals) and others.
A technology transfer project

• Outcome for the participating SMEs working in the legal information provision sector
  – a portfolio of innovative data analysis and human language technology components for elaborating legal documents,
  – resulting from the collaboration with the research partners and the advisory board members
  – in order to create innovative high tech value-added services and products in their market
  – addressing the needs of nearly a million users:
    • legal professionals (judges, lawyers, etc.) and
    • legal information providers/publishers redistributing the linked open data
  – market analysis and exploitation plan
The technologies
Technologies used

• Structuring legal documents:
  – Legal XML: Akoma Ntoso
  – Naming of documents: ELI/ECLI
  – NLP parsers for translation in XML and linking

• Semantic enrichment:
  – Multilingual ontologies (Eurovoc and Legal Taxonomy Syllabus)
  – Text to ontology linking
  – Automated classification
  – Automated summarization
  – Automated translation

• Open data: SPARQL
Overview

EUCases Linking Platform

Language technologies
- Structural interlinking
  - Identifiers: ELI, ECLI, national
  - Tools for linking, LT2XML converters
  - Metadata extraction
  - Tools for case annotation
- Semantic interlinking
  - Eurovoc thesaurus
  - Syllabus ontology, ontology indexer
  - Ontology-to-text processing
  - Multilingual access
  - Classification

Legal doctrine

Multi-lingual database
- EN, DE, FR, IT, BG

Legal open data

Data analysis and linking

Value added tools & services

SMEs' value added products
- ConsumerCases
  - Web Application
  - Consumer protection law
  - Access to a multilingual collection of tens of thousands cross-linked resources

EULinksChecker
- Add-in tools
  - for browsers, text editors and viewers

Legal linked open data
- Courts, law firms, data providers, publishers, open data re-users

User testing & validation/Exploitation plan