

# Aligning *legivoc* Legal Vocabularies by Crowdsourcing

Hughes-Jehan Vibert, Benoît Pin, Pierre  
Jouvelot (talk given by Laurent Daverio)

# legivoc (1/2)

- The first Internet-based platform dedicated to the diffusion, edition and alignment of legal vocabularies across countries
- Funded in part by the European Commission with with Ministries of Justice of CZ, ES, FI, FR, IT, and LU, and administered by the FR Ministry of Justice
- ARMINES and MINES ParisTech are the lead scientific advisors and implementation specialists

## legivoc (2/2)

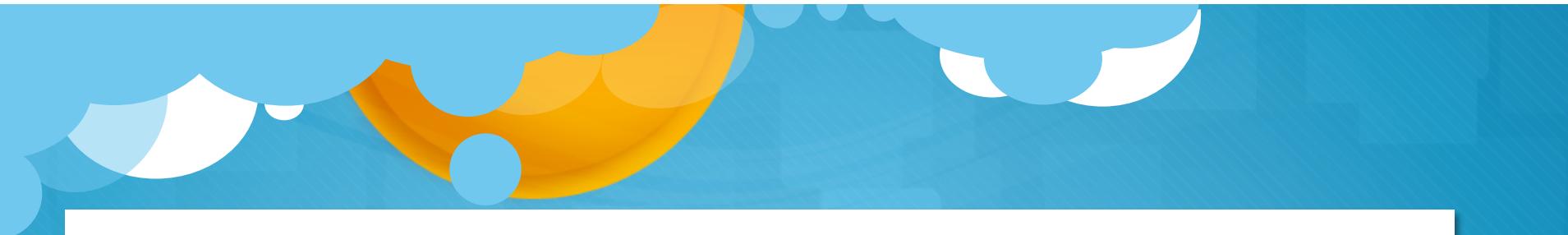
- Intended to be used directly by law and judicial experts, e.g. when dealing with cross-border legal issues or planning new legislative regulations
- Crowdsources alignment of vocabularies
  - During the alpha phase, any registered user can contribute.

# W3C Standards in legivoc

- legivoc strongly adheres to W3C standards, such as:
  - RDF (“*Resource Description Framework*”)
  - URI (“*Uniform Resource Identifier*”)
  - SKOS (“*Simple Knowledge Organisation System*”)
  - SPARQL (“*SPARQL Protocol and RDF Query Language*”)

# Accessing Vocabularies

- Concepts in vocabularies can be:
  - visualized in various forms (text, dendrogram, SKOS source),
  - edited,
  - related to more or less abstract concepts.
- In addition, SPARQL querying for IT specialists
- Registered users only



View

Edit

Dendrogram

SKOS

## Concept: eu-523

### Labels

en civil law

en ordinary law

en statutory law

### Semantic relations

#### Narrower concepts

[abuse of power](#) | [legal status](#) | [misuse of a right](#) | [liability](#) | [contract](#) | [law of obligations](#)

#### Related concepts

[civil code](#) | [private law](#)

### Mappings

#### Close matches

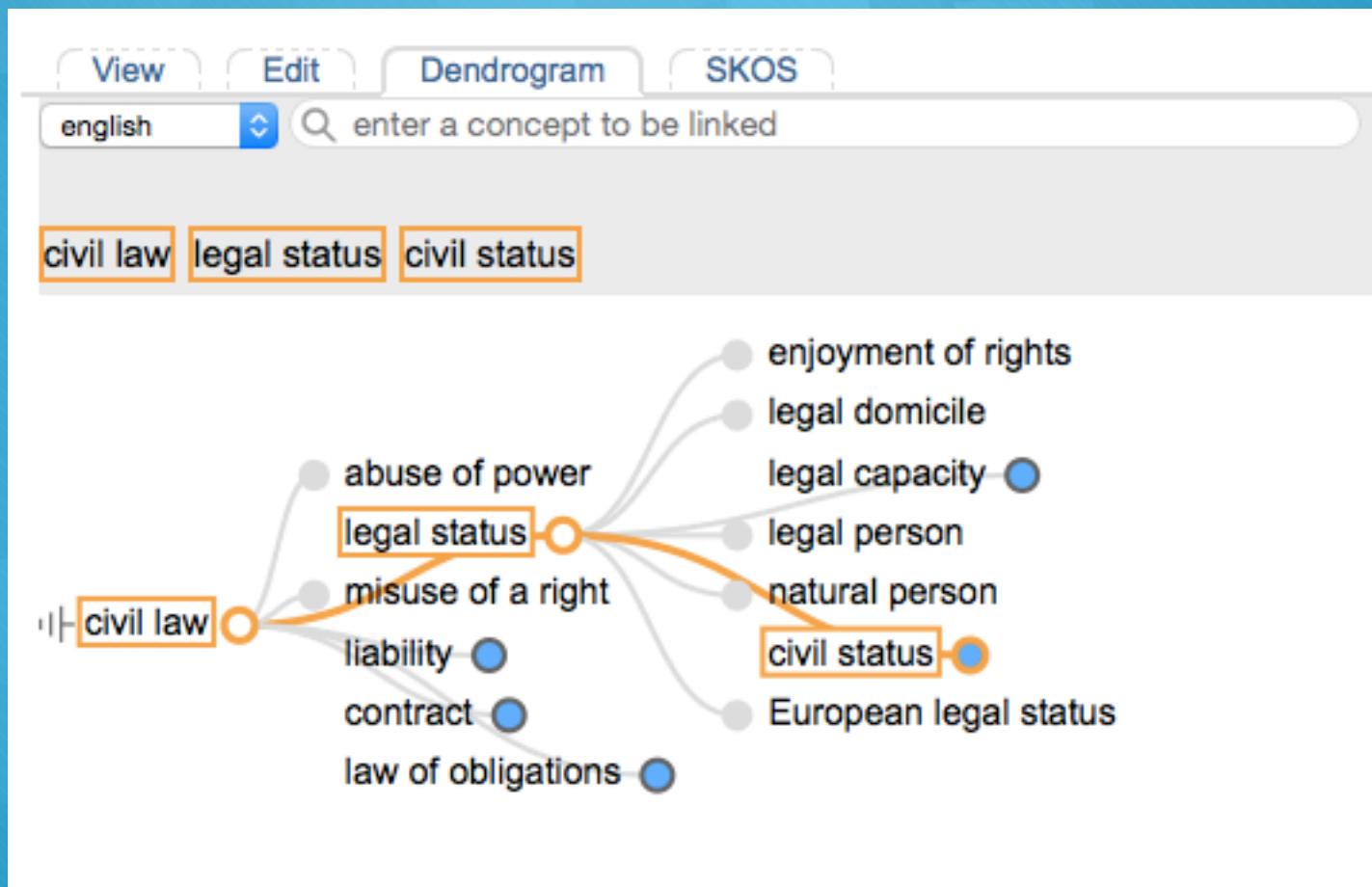
 [Civil law](#) |  [civil law](#) |  [civil law](#) |  [civil law](#) |  [Civil law](#)

[View](#)[Edit](#)[Dendrogram](#)[SKOS](#)

## Edit concept: eu-523

### Labels

bg	гражданское право	<input type="button" value="pref"/>
es	Derecho civil	<input type="button" value="pref"/>
es	Derecho común	<input type="button" value="alt"/>
cs	občanské právo	<input type="button" value="pref"/>
cs	civní právo	<input type="button" value="alt"/>
da	civilret	<input type="button" value="pref"/>
da	borgerlig ret	<input type="button" value="alt"/>
de	Bürgerliches Recht	<input type="button" value="pref"/>
de	allgemeines Recht	<input type="button" value="alt"/>
de	Zivilrecht	<input type="button" value="alt"/>
et	tsiviilõigus	<input type="button" value="pref"/>
et	üldõigus	<input type="button" value="alt"/>
el	αστικό δίκαιο	<input type="button" value="pref"/>
el	Koivó Δίκαιο	<input type="button" value="alt"/>
en	civil law	<input type="button" value="pref"/>
en	ordinary law	<input type="button" value="alt"/>
en	statutory law	<input type="button" value="alt"/>
fr	droit civil	<input type="button" value="pref"/>



View Edit Dendrogram SKOS

## 523

```
@prefix dct: <http://purl.org/dc/terms/> .  
@prefix legivoc: <http://legivoc.org/namespaces/2014/legivoc#> .  
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .  
@prefix skosxl: <http://www.w3.org/2008/05/skos-xl#> .  
@prefix xml: <http://www.w3.org/XML/1998/namespace> .  
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .  
  
<http://eurovoc.europa.eu/523> a skos:Concept ;  
    skos:closeMatch <http://legivoc.org/be/932>,  
                    <http://legivoc.org/fi/1619>,  
                    <http://legivoc.org/fi/2270>,  
                    <http://legivoc.org/nl/3792>,  
                    <http://legivoc.org/si/13> ;  
    skos:narrower <http://eurovoc.europa.eu/164>,  
                  <http://eurovoc.europa.eu/186>,  
                  <http://eurovoc.europa.eu/3497>,  
                  <http://eurovoc.europa.eu/4274>,  
                  <http://eurovoc.europa.eu/5598>,  
                  <http://eurovoc.europa.eu/96> ;  
    skos:related <http://eurovoc.europa.eu/5496>,  
                 <http://eurovoc.europa.eu/576> ;  
    skosxl:altLabel <http://eurovoc.europa.eu/523/label/11>,  
                     <http://eurovoc.europa.eu/523/label/13>,  
                     <http://eurovoc.europa.eu/523/label/15>,  
                     <http://eurovoc.europa.eu/523/label/16>,  
                     <http://eurovoc.europa.eu/523/label/19>.
```

## SPARQL form

vocabulary: gr.rdf

```
1 PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
2 PREFIX skos:<http://www.w3.org/2004/02/skos/core#>
3 PREFIX skosxl:<http://www.w3.org/2008/05/skos-xl#>
4
5 SELECT *
6 WHERE {
7   ?conceptUri skosxl:prefLabel ?labelUri .
8   ?labelUri skosxl:literalForm ?label .
9 }
10 LIMIT 10
11 |
```

Get response as XML

conceptUri	labelUri	label
<a href="http://legivoc.org/gr/163">http://legivoc.org/gr/163</a>	<a href="http://legivoc.org/gr/163/label/0">http://legivoc.org/gr/163/label/0</a>	ΧΑΡΤΟΓΡΑΦΗΣΗ ΔΑΣΩΝ
<a href="http://legivoc.org/gr/6610">http://legivoc.org/gr/6610</a>	<a href="http://legivoc.org/gr/6610/label/0">http://legivoc.org/gr/6610/label/0</a>	ΕΞΑΦΑΝΙΣΗ-ΑΝΑΚΛΗΣΗ-ΜΕΤΑΡΡΥΘΜΙΣΗ-ΑΠΟΚΛΕΙΣΜΟΣ ΑΠΟΕΛΛΑΣΗΣ
<a href="http://legivoc.org/gr/2465">http://legivoc.org/gr/2465</a>	<a href="http://legivoc.org/gr/2465/label/0">http://legivoc.org/gr/2465/label/0</a>	ΦΟΡΟΛΟΓΙΑ ΧΡΕΟΓΡΑΦΩΝ
<a href="http://legivoc.org/gr/5039">http://legivoc.org/gr/5039</a>	<a href="http://legivoc.org/gr/5039/label/0">http://legivoc.org/gr/5039/label/0</a>	ΤΑΜΕΙΟ ΕΠΙΚΟΥΡΙΚΗΣ ΑΣΦΑΛΙΣΕΩΣ ΚΑΙ ΠΡΟΝΟΙΑΣ ΤΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ ΥΠΟΧΡΕΩΤΙΚΗΣ ΑΣΦΑΛΙΣΗΣ
<a href="http://legivoc.org">http://legivoc.org</a>	<a href="http://legivoc.org/gr/7199">http://legivoc.org/gr/7199</a>	ΓΕΝΙΚΗ ΣΥΝΕΛΕΥΣΗ ΑΣΤΙΚΩΝ ΣΥΝΕΤΑΙΡΙΣΜΩΝ

# Alignment Management

- Based on SKOS' mapping properties
  - *exactMatch, broadMatch, narrowMatch, relatedMatch*
  - *Too many concepts for a text-based approach*
    - *A possible approach would be machine learning*
    - *But legivoc chose a different approach :*
      - *Crowdsourcing, human computation, gamification*



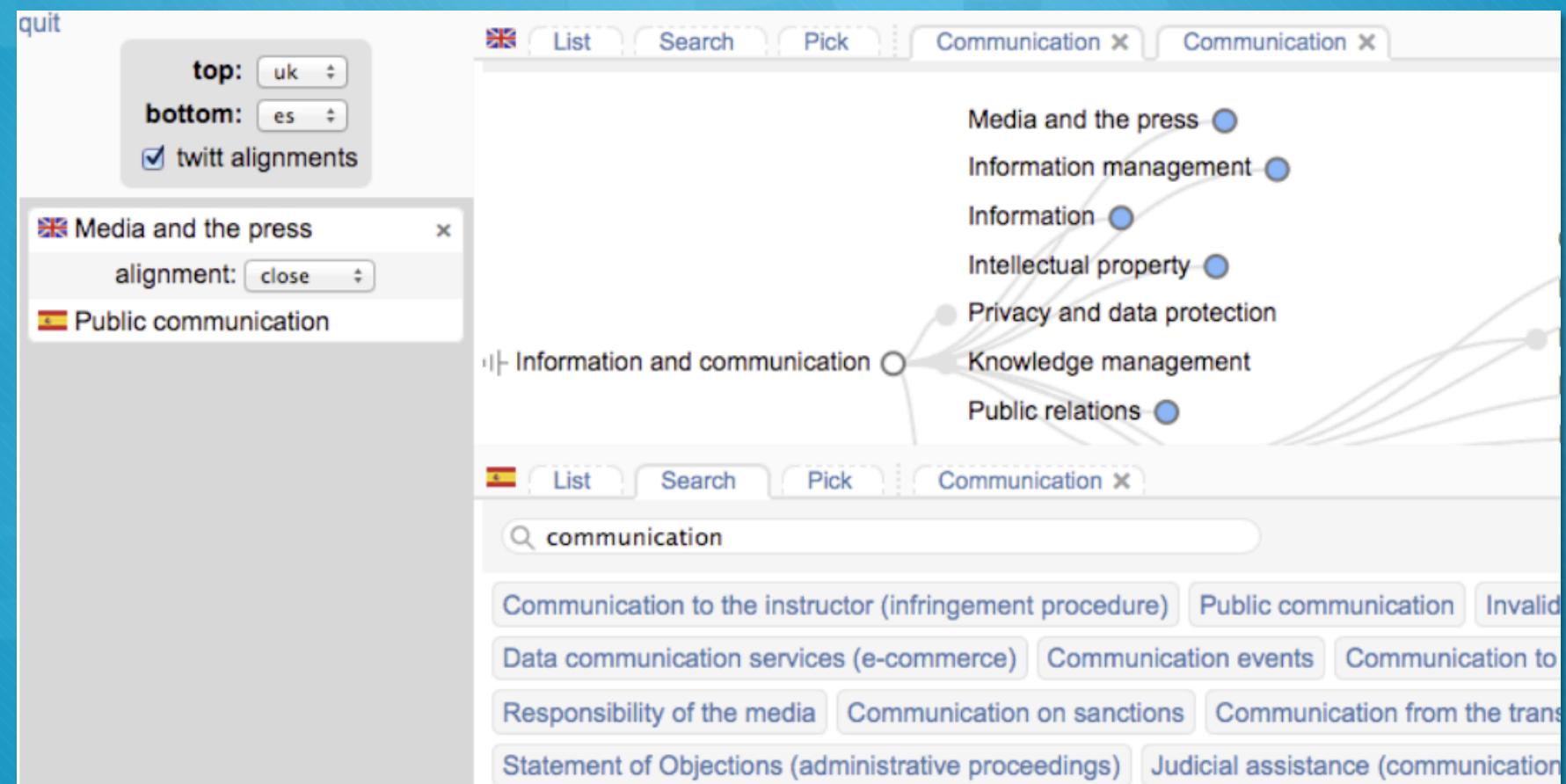
# Aligning Vocabularies in legivoc (1/2)

- Any registered user can suggest alignments
  - New alignments are posted on Twitter
    - Build a sense of community
    - Build motivation
  - Alignments are checked *a posteriori* by experts
  - Incorrect/conflicting information can be reverted
    - A similar approach to *Wikipedia*



# Aligning Vocabularies in legivoc (2/2)

- Concepts to be aligned can be chosen in various ways
  - List: alphabetical list
  - Search
  - Pick: random suggestion
- The page is split in two halves, and matching can be done by a simple drag-and-drop between dendrogram nodes



# legivoc as a Web Service

- Two types of querying:
  - *Structured*: SPARQL queries
    - Output format : XML
  - *Non-structured* : textual searches on concepts
    - Output format: JSON
- Authenticated access only
- Currently tested within the *legicoop* network



# SKOS (*Simple Knowledge Organisation System*)

- An application of RDF for representing *thesauri, taxonomies, classification schemes*, etc.
- SKOS handles *concepts*
  - identified by *URIs*
  - characterised by natural-language *labels*
  - Linked by *relationships (broader/narrower, related)*



# **SPARQL (*SPARQL Protocol and RDF Query Language*)**

- A query language for RDF “triples” (subject – predicate – object)
- Similar to query languages for key-value datastores (NoSQL)
- Offers analytics tools such as joins and aggregates
- A SPARQL server, Fuseki, part of the Apache Jena project, allows to interact with RDF data over HTTP

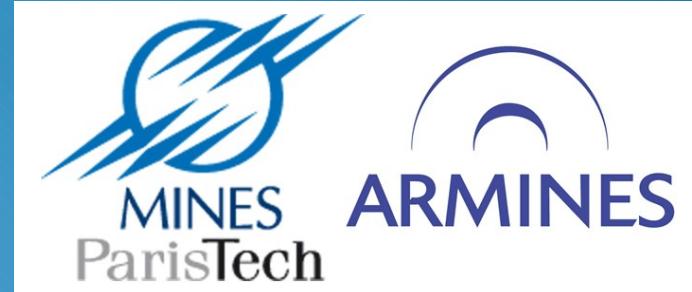
# Using Existing Legal Vocabularies

- National legal vocabularies from 13 EU members+ Switzerland, and Eurovoc
  - Note: Eurovoc is not a “legal” thesaurus
- Need to handle a variety of input formats:
  - ASCII, Excel, Word, XML, SKOS, ...
- Output:
  - Encoding into legivoc-specific SKOS format
  - Could be the base for a future standard

# Conclusion

- Current goals have been reached
  - Based on W3C standards (SKOS, SPARQL)
  - Intuitive and fun to use — communication techniques: Twitter
  - Widely accessible
- Future work
  - More legal vocabularies to come soon
    - Croatia, several African and Arab countries
  - Reasoning over alignment information
  - More gamification?

# Contact



- Hughes-Jehan Vibert  
Ministère de la Justice, France
  - [hughes-jehan.vibert@justice.gouv.fr](mailto:hughes-jehan.vibert@justice.gouv.fr)
- Benoît Pin, Pierre Jouvelot, Laurent Daverio  
MINES ParisTech, PSL Research University, France
  - [{benoit.pin, pierre.jouvelot, laurent.daverio}@mines-paristech.fr](mailto:{benoit.pin, pierre.jouvelot, laurent.daverio}@mines-paristech.fr)