



Annotation schema for legal doctrine: a case study on DoGi database

Tommaso Agnoloni, Elisabetta Marinai,
Ginevra Peruginelli, Maria Teresa Sagri,
Daniela Tiscornia

The context

Legal information has specific features due to its nature, different purposes and uses made of it, while preserving a strong requirement for integration of its different sources: legislation, case law and doctrine. Not only have these three sources different functions and purposes, but they require peculiar solutions for ensuring their access. In spite of this diversity there is a strong need, both at national and international level, of new forms of integration between these different sources to ensure effective sharing and interconnection of legal data. In traditional legal information systems data-sets are stored, organized and classified according to the classical distinction among types of legal sources. Apart from the European system (Eur-Lex) and few commercial initiatives the provision of integrated access to data originating from a wide range of distributed, multilingual and heterogeneous data sources is still scarce. Therefore,

users are obliged to access several providers of legal information and perform their searches step by step, by identifying specialized data repositories, assessing their own search strategy, browsing several databases, and establishing conceptual and legal connections by themselves. Of course there is quite a number of providers of legal information users can turn to in their research activity. As a matter of fact a user should first examine the legal sources hosting the desired information and then locate and check the availability of the resource. In particular, legal doctrine plays a key role in integrating various types of information. It comprises a vast amount of scientific literature produced by a variety of legal sources and meets essential requirements such as interpretation and understanding of rules and judgements as clarified and commented by legal scholars. As such, legal literature does reflect the legal debate within an ever growing and constantly changing society. By merging conceptual synthesis and critical analysis, legal literature can be seen as the final step in the creative process of the legal discourse, after the regulative step, the interpretative reconstruction and the comparison between the factual and the abstract description of a case.

The linked data approach in testing the DoGi-Legal Doctrine database

In the area of law the application of linked open data technology provides several advantages over the current management of legal data, while offering an opportunity for the development of new information systems to legal users.¹ In fact, the linked open data

¹«A new approach for the notion of legal information is necessary, [...] by adapting the term to the new technological possibilities whilst ensuring the authority of the information as a legal source, law could become more of a logical networked system of rules and concepts, in line with that it is assumed to be» (Kirchberger).

environment provides the design principles and the technical infrastructure to publish structured data on the web so as to allow their interconnection and, finally, integration into a larger and distributed data space. The possibility to aggregate different data sets based on the systematic relationships between classes of sources, to meet and combine them regardless of their format and origin, is particularly interesting for the reconstruction of the complex structure of legal discourse. In the legal domain, where knowledge emerges from the connection among legislative, jurisprudential and bibliographic data,² the linked data methodology allows the creation of new repositories of knowledge. Furthermore, the development of a collaborative connection between different types of legal information contributes to ensure open access to digital legal knowledge and to develop a *seamless web* for the legal domain. Legal data sets, produced and published independently, are typically structured resources that can be enriched by extra-legal information content such as bibliographic citations related to authors' biographies. Therefore, linked data technologies can provide a standardized mechanism for publication of structured legal doctrine, with the opportunity to gather and explore external legal resources providing an interpretational context. The choice to use legal doctrine for testing a linked data approach in a legal environment is not accidental. For this study the DoGi-Legal Doctrine database³ has been retained as a valuable resource due to its importance in documenting the evolution of the Italian legal science. The database is, at present, the most comprehensive source of online information on Italian doctrine. The Institute of Legal Information Theory and Techniques of the National

²«The needs of information retrieval of the jurist who faces a legal query consists in the selection of all the relevant legislation to this case, the largest possible number of opinions from experts and the largest possible number of significant case law on domain» (Rescigno).

³<http://nir.ittig.cnr.it/dogi>.

Research Council (ITTIG) produces, manages and distributes the database since 1970, the year of its foundation. The corpus includes approximately 250 print and electronic legal journals published by Italian publishers. Here articles; case notes, reviews / comments of legislation and case law, reports on conferences and critical reviews are analytically indexed. This large information resource including approximately 400,000 records consists of bibliographic references of each contribution; abstracts listing the topics discussed by the author, one or more classification codes that identify the legal topics, references of the main sources such as legislation, regulations and case law cited by the author of the contribution with a link to the full text of the legal source⁴ The value of the database compared to other available information resources provided by commercial publishers, but also produced on the basis of voluntary institutional initiatives,⁵ consists in the fact that DoGi resource provides access to literature not only through bibliographic references, but also through references to legislation and case law cited in the text. The richness and quality of the available information offer new opportunities to create relationships between entities and resources. Giving access to legislation and case law cited in the article, as well as providing the user with a very analytical legal classification, lawyers, as well as ordinary citizens have the opportunity to obtain an overview of legal issues and have access to complete documentation to resolve legal cases of day life. This allows users to know, understand and use data for specific purposes, both professional and academic, or just for simple interest. In this context, linked data technologies can contribute to the development, application and sharing of DoGi records on the web. Following the increasing adoption of advanced web technologies by governments, academia and professionals, the

⁴For further information see: <http://nir.ittig.cnr.it/dogiswish/brochureSwish.htm>.

⁵Associazione ESSPER: <http://www.biblio.liuc.it/essper/default.asp>.

intent is to model the DoGi resources reusing existing descriptive standard schemas such as BiBo, FOAF, DC-Terms and SKOS eventually extended to adapt to the DoGi conceptual structure. The two main objectives of this study are the following:

1. creation of a data model for the DoGi resource as a first work in progress for the legal community in order to lay the ground work for new forms of integration of bibliographic data, legislation and case law available on the web;
2. conversion of the DoGi classification scheme into SKOS and its alignment with other thesauri (such as the BNCF "Nuovo Soggettario", Eurovoc). This involves: *a*) enrichment of the collections through the use of controlled vocabularies; *b*) retrieval and integrated search of heterogeneous collections through the use of vocabularies; *c*) analysis of specific topics by navigating through controlled vocabularies; *d*) cross-collection and cross language retrieval, *e*) link between domains, disciplines, different communities.

Definition of the DoGi Data model

As already mentioned, due to its richness in connections to heterogeneous possibly distributed sources, the DoGi database is especially suited to a conversion to linked data formats. One of the peculiarities of the DoGi database is in fact that it can be seen halfway between a bibliographic archive (accessed by bibliographic metadata e.g. author, title, journal..) and a legal sources archive (accessed by legal sources through legal metadata e.g. enacting authority, year, number ...) with the doctrine article acting as a connection hub among legal subjects and the legal documents commented, discussed or supporting the authors argumentation throughout the paper.

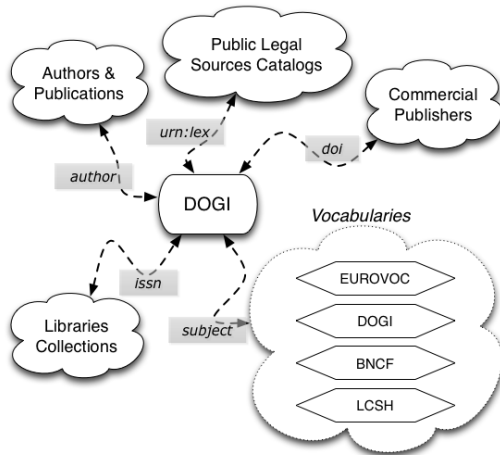


Figure 1: Potential connection of the DoGi dataset in the linked data cloud.

In figure 1 sketches the potential enrichment of the existing DoGi dataset, both in content and access possibilities, made possible by the integration with data published by third parties. In a symmetric way DoGi data would contribute to enrich related data collections. The work carried on consisted in the first place in identifying the standard vocabularies and domain specializations able to model the entities in the DoGi database. Based on the identified models and on the definition of a unified integrated data model, we have been able to define the data mappings necessary to open up to RDF format conforming to the model an excerpt of the DoGi dataset (the bibliographic records, authors and referred legal sources) and of the DoGi classification scheme used for its content categorization in legal subjects. Data were originally made available as XML dumps from the database. Before that, first class resources in the dataset

should be assigned globally unique identifiers to be referenced in the linked data space by defining namespace prefixes to prepend to database entities IDs; In the conversion from the database we used the following namespaces as prefixes to corresponding resources identifiers:

<http://www.ittig.cnr.it/dogi/periodical/>
<http://www.ittig.cnr.it/dogi/issue/>
<http://www.ittig.cnr.it/dogi/record/>
<http://www.ittig.cnr.it/dogi/legalsource/>
<http://www.ittig.cnr.it/dogi/person/>
<http://www.ittig.cnr.it/dogi/agent/>
<http://www.ittig.cnr.it/dogi/descriptor/>
<http://www.ittig.cnr.it/dogi/term/>

Modeling DoGi Bibliographic records

As regards bibliographic informations, standard vocabularies can be effectively reused for DoGi data to a large extent. Dublin Core Terms metadata vocabulary, BibO the bibliographic ontology, Foaf friend of a friend ontology have been used as sketched in figure 2 on the next page to represent bibliographic records, bibliographic resources and authors respectively. A mapping of the DoGi records fields (left column) to Dublin Core metadata and bibliographic ontology properties (right column) as used for the conversion of data, is reported in the following tables. The mapping adopted the three distinct hierarchical levels: journal, journal issue and individual article level.

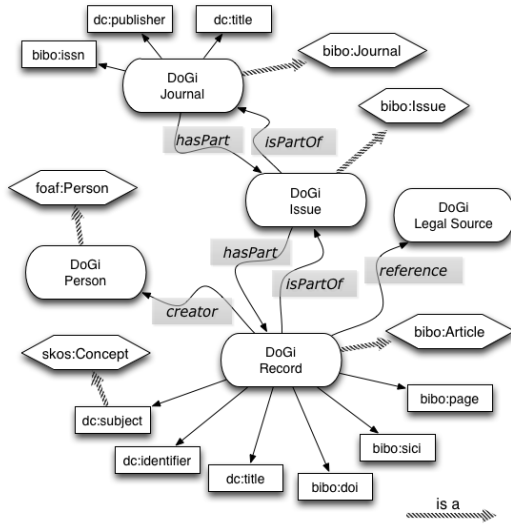


Figure 2: Modeling DoGi metadata reusing DC Terms, BibO, Foaf.

record ID	dc:identifier
	bibo:doi
	bibo:sici
Title	dc:title
Author	dc:creator
Editor	dc:contributor
Language of the original contribution	dc:language
Classification	dc:subject
Outline	dcterms:tableOfContents
Summary	dcterms:abstract
Source of the article (e.g. <i>Decision Court of ...</i> , <i>International conference on...</i> , etc.)	dc:source
Appeared in	dcterms:isPartOf
Pages	bibo:pageStart, bibo:pageEnd
Referenced Legal sources (legislation, case law)	dcterms:references

Table 1: DoGi Record / Articolo.

issue ID	dc:identifier bibo:doi
Year of publication	dcterms:issued
Issue number	bibo:issue
Volume number	bibo:volume
Editor	dc:contributor
In Journal	dcterms:isPartOf
Contains article	dcterms:hasPart

Table 2: DoGi Issue

Journal ID	dc:identifier
ISSN code	bibo:issn
Journal title	dc:title
Description	bibo:shortDescription
Publisher	dc:publisher
Director	bibo:director
Contains issue	dcterms:hasPart

Table 3: DoGi Journal

The possible kind of doctrinal contributions catalogued in DoGi have been modelled as specializations of generic bibliographic resources in the dogi namespace (v. figure 3 on page 238).

Listing 1: Different types of doctrinal contribution examined in the DoGi archive.

```

<!-- JOURNAL -->
<rdf:Description rdf:about='http://www.ittig.cnr.it/dogi/
  periodical/1486''>
  <rdf:type rdf:resource='bibo:Journal' />
  <bibo:issn>urn:issn:1591-5611</bibo:issn>
  <dc:title>Diritto penale e processo</dc:title>
  <bibo:shortDescription>Mensile di giurisprudenza,
    legislazione e dottrina</bibo:shortDescription>
  <dc:publisher rdf:resource='http://www.ittig.cnr.it/dogi/
    organization/ipsoa''>
</rdf:Description>

<!-- ISSUE -->
<rdf:Description rdf:about='http://www.ittig.cnr.it/dogi/
  issue/1486-13-1''>
  <rdf:type rdf:resource='bibo:Issue' />
  <dcterms:identifier>1486-13-1</dcterms:identifier>
  <dcterms:isPartOf rdf:resource='http://www.ittig.cnr.it/
    dogi/periodical/1486''/>

```

```
<dcterms:hasPart rdf:about=''http://www.ittig.cnr.it/dogi/
  record/2012-G0162'' />
<dcterms:issued>2012</dcterms:issued>
<bibo:volume>13</bibo:volume>
<bibo:issue>1</bibo:issue>
</rdf:Description>

<!-- DoGi RECORD -->
<rdf:Description rdf:about=''http://www.ittig.cnr.it/dogi/
  record/2012-G0162''>
  <rdf:type rdf:resource=''bibo:Article'' />
  <rdf:type rdf:resource=''dogi:IndependentContribution'' />
  <dc:identifier rdf:resource=''http://www.ittig.cnr.it/dogi/
    record/2012-G0162'' />
  <bibo:doi>10.3280/DUDI2012-001001</bibo:doi>
  <bibo:sici>1591-5611(201201)1:1%69:MEDP%2.0.TX;2-4</bibo:
    sici>
  <dc:creator rdf:resource=''http://www.ittig.cnr.it/dogi/
    persons/BartoliRoberto'' />
  <dc:title>Mobbing e diritto penale</dc:title>
  <dc:type>Contributo dottrinale indipendente</dc:type>
  <dc:type rdf:resource=''http://purl.org/dc/dcmitype/Text'' />
  <dcterms:language>it<dcterms:language>
  <dcterms:isPartOf rdf:resource=''http://www.ittig.cnr.it/
    dogi/issue/1486-13-1'' />
  <bibo:pageStart>85</bibo:pageStart>
  <bibo:pageEnd>94</bibo:pageEnd>
  <dcterms:description>Bibliografia: a pie' di pagina o nel
    corpo del testo</dcterms:description>
  <dcterms:abstract>Il fenomeno del mobbing risulta molto più
    complesso di quanto si possa credere, verificandosi
    all'interno di dinamiche relazionali senza dubbio
    tradizionali [...] fatti offensivi.</dcterms:abstract>
  <dcterms:tableOfContents>Mobbing e reati relazionali, tra
```

```

    passato e modernità. I modelli classici di tutela penale
    a contrasto del mobbing.</dcterms:tableOfContents>
<!--. classification -->
<dc:subject rdf:resource=''http://www.ittig.cnr.it/dogi#
    s0920'' />
<!-- TUTELA DELLA DIGNITA DEL LAVORATORE -->
<dc:subject rdf:resource=''http://www.ittig.cnr.it/dogi#
    d4764'' />
<!-- Tutela della personalita morale del lavoratore (mobbing
    ) -->
<!--.Legal sources -->
<dcterms:references rdf:resource=''http://www.ittig.cnr.it/
    dogi/legalsource/22367902'' />
<dcterms:references rdf:resource=''http://www.ittig.cnr.it/
    dogi/legalsource/35366803'' />
    <...../>
</rdf:Description>

<!-- DoGi AUTHOR / ORGANIZATION -->
<rdf:Description rdf:about=''http://www.ittig.cnr.it/dogi/
    persons/BartoliRoberto''>
    <rdf:type rdf:resource=''foaf:Person'' />
    <foaf:givenname>Roberto</foaf:givenname>
    <foaf:family\_name>Bartoli</foaf:family\_name>
    <foaf:mbox>xxx@yy.com</foaf:mbox>
    <foaf:phone/>
    <foaf:topicInterest/>
    <foaf:depiction/>
    <foaf:homepage/>
    <foaf:weblog/>
    <foaf:publications/>
    <dogi:hasCreated rdf:resource=''http://www.ittig.cnr.it/dogi
        /record/2012-G0162'' />
</rdf:Description>

```

```
<rdf:Description rdf:about='http://www.ittig.cnr.it/dogi/
  organization/ipsoa'>
  <rdf:type rdf:resource='foaf:Organization' />
  <foaf:name>IPSOA<foaf:givenname>
  <address:localityName>Rome</address:localityName>
  <foaf:mbox>zzz@kk.com<foaf:mbox>
  <foaf:homepage/>
  <foaf:phone/>
  <foaf:depiction/>
  <dogi:hasCreated rdf:resource='http://www.ittig.cnr.it/dogi
    /periodical/1486'>
</rdf:Description>
```

- ▼ ⓘ dogi:BibliographicResource
 - dogi:BookReview
 - dogi:CommentOnCaseLaw
 - dogi:CommentOnLaw
 - dogi:IndependentContribution
 - dogi:SeminarReport
 - dogi:WorkshopReport

Figure 3: Dogi type of contributions classification.

A multilingual extension of such a representation is straightforward and natively supported by the RDF data model.

Modeling DoGi legal sources

As regards legal sources referred in DoGi records, these are extremely heterogeneous covering both legislation and case law and national, international, European sources. We reproduced this classification introducing the classes reported in figure 4 on the next page

in the DoGi annotation schema.



Figure 4: Dogi legal sources classification.

Legislation

Concerning legislative sources, despite several related initiatives in the legal documentation field, no shared reusable vocabulary for the attribution of legislative metadata covering different legislative sources typology (national, European, International) and for different countries' national legislations, is currently available. However, for the purpose of legislative documents identification the urn:lex schema is increasingly being adopted. urn:lex is a proposed Internet standard for legal document identifiers.⁶ The urn:lex namespace aims to facilitate the process of creating URIs for legal sources independent of a document's online availability, location, and access mode. "Sources of law" include any legal document within the domain of legislation (including bills), case law and administrative acts or regulations. This identifier will be used as a way to represent references (and more generally, any type of relation) among the various sources of law. Being a transparent identifier, urn:lex

⁶<http://tools.ietf.org/html/draft-spinosa-urn-lex-06>.

can be seen as a serialization of the basic metadata able to describe a legislative act (the country where the legal source has effect, the enacting authority, the type of document, its number, an effective date..). We adopted an extension of this set of basic metadata for the description of referred legislative sources in DoGi documents along with the urn:lex specification for their unique identification (Tab. 4).

Listing 2: descrizione delle risorse legislative riferite nei record DoG

```
<!--cited legislation: art.55 d.lgs. 9 aprile 2008, n.81 -->
<rdf:Description rdf:about=
  ''http://www.ittig.cnr.it/dogi/legalsource/22367902''>
  <rdf:type rdf:resource=
    ''http://www.ittig.cnr.it/dogi/NationalLegislation''/>
  <dcterms:identifier rdf:resource=
    ''urn:nir:stato:decreto.legislativo:2008-04-09:81#art55''>
  <dcterms:isPartOf rdf:resource=
    ''urn:nir:stato:decreto.legislativo:2008-04-09:81''>
  <dc:type>text</dc:type>
  <dc:title>Sanzioni per il datore di lavoro
    e il dirigente
  </dc:title>
  <dcterms:coverage>it</dcterms:coverage>
  <dcterms:language>IT</dcterms:language>
  <dogi:enactingAuthority>Stato</dogi:enactingAuthority>
  <dogi:legislativeDocumentType>Decreto
    Legislativo
  </dogi:legislativeDocumentType>
  <dcterms:issued>2008-04-09</dcterms:issued>
  <dogi:publicationDate>2008-04-30</dogi:publicationDate>
  <dcterms:publisher>Gazzetta Ufficiale
  </dcterms:publisher>
  <dogi:textualCitation>art. 55 d.lgs. 9 aprile 2008,
    n. 81</dogi:textualCitation>
  <dcterms:subject rdf:resource=
```

```

''http://www.ittig.cnr.it/dogi#s0920''/>
<dogi:text_url>
http://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:
decreto.legislativo:2008-04-09;81~art55
</dogi:text_url>
<dogi:isReferencedBy rdf:resource=
''http://www.ittig.cnr.it/dogi/record/2012-G0162''/>
</rdf:Description>

```

Legislative document ID	dc:identifier (urn:lex Schema)
Document title	dc:title
Territorial jurisdiction	dc:coverage
Document type	dogi:legislativeDocumentType
Enacting Authority	dogi:enactingAuthority
Date of enactment	dcterms:issued
Publication Date (Efficacy)	dogi:publicationDate
Publisher	dcterms:publisher
Textual Citation	dogi:textualCitation
Classification	dcterms:subject
Document URL	dogi:text_url
Reference to partition container (e.g. Article part of Section)	dcterms:isPartOf
References to other resources	dcterms:references
Resources referring the legislative source (e.g. a DoGi record)	dogi:isReferencedBy

Table 4: dogi:Legislation

Notice that, in a linked open data environment, the national publication offices should be themselves responsible for the open publication of national legislation catalogs in open formats along with the associated RDF metadata as open data sets in the linked data cloud. In such a foreseeable scenario the urn:lex identifier would be by itself sufficient to establish a machine readable connection to the official legislative metadata catalogs, automatically augmenting the DoGI dataset with the official descriptive metadata of the legal sources referred in DoGI doctrinal articles. In this respect the most advanced initiative of open legislation publishing is so far the UK

legislation portal⁷ part of the British open government data strategy⁸ where all the best practices of XML publication of legislative documents along with associated RDF metadata have been put in practice.

In legislation.gov.uk a more sophisticated metadata model – incorporating FRBR,⁹ the CEN MetaLex vocabulary,¹⁰ Dublin Core Terms,¹¹ and the Crown Legislation Markup Language – enabling advanced version control and output of descriptive metadata have been adopted providing also all the metadata able to implement a point in time legislative system.

A rapid change of the scenario in this direction can be foreseen in the near future both at the national and European level under the impulse of the open government data movement and of revised discipline in favour of the right of access to public sector information. A significant example of this trend is the experimentation currently carried on in the EU LATC Project of transition to linked open data of the whole Eur-Lex service of European Union Law along with its connection to national implementing legislation.¹²

Case law

Similar considerations hold for the identification of a set of basic metadata able to identify and describe case law documents. In this case we based on the recommendation establishing Council of Europe conclusions inviting the introduction of the European Case Law Identifier (ECLI) along with minimum set of uniform metadata

⁷www.legislation.gov.uk.

⁸<http://www.data.gov.uk>.

⁹<http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records>.

¹⁰<http://www.metalex.eu>.

¹¹<http://www.dublincore.org/documents/dcmi-terms>.

¹²<http://eur-lex.publicdata.eu>.

for case law. The following is an RDF serialization of the metadata related to a Jurisprudential source from the DoGi archive according to the annotation defined in Tab. 5.¹³

```

<!--cited decision: Cass. sez. VI pen. 13 gennaio 2011, n. 685
-->
<rdf:Description rdf:about= ''http://www.ittig.cnr.it/dogi/
legalsource/35366803''>
  <rdf:type rdf:resource= ''http://www.ittig.cnr.it/dogi/
NationalCaseLaw''/>
  <dcterms:identifier rdf:resource= ''urn:lex:it:corte.
cassazione:sentenza:2011-01-13;685''/>
  <dcterms:identifier rdf:resource=''ECLI:IT:CCA
:2011-01-13:685''/>
  <dcterms:creator>Corte di Cassazione sezione VI penale</
dcterms:creator>
  <dogi:court>Corte di Cassazione</dogi:court>
  <dogi:courtSection>VI</dogi:courtSection>
  <dogi:judicialDivision>penale</dogi:judicialDivision>
  <dc:type>judicial decision</dc:type>
  <dc:title/>
  <dcterms:coverage>it</dcterms:coverage>
  <dcterms:language>IT</dcterms:language>
  <dcterms:date>2011-01-13</dcterms:date>
  <dcterms:issued/>
  <dcterms:publisher/>
  <dcterms:contributor>Serpico</dcterms:contributor>
  <dcterms:contributor>Milo</dcterms:contributor>
  <dogi:textualCitation>
    Cass. sez. VI pen. 13 gennaio 2011, n. 685
  </dogi:textualCitation>
  <dc:subject rdf:resource=''http://www.ittig.cnr.it/dogi#

```

¹³<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2011:127:0001:0007:IT:PDF>.

```

    d4764'' />
<dcterms:abstract>
    Il mobbing è solo vagamente assimilabile alla previsione di
        cui all'art. 572 c.p....
</dcterms:abstract>
<dogi:text_url/>
<dogi:isReferencedBy rdf:resource= ''http://www.ittig.cnr.it
    /dogi/record/2012-G0162'' />
</rdf:Description>

```

Case law ID	dcterms:identifier (ECLI)
	dcterms:identifier (urn:lex)
The full name of the court	dcterms:creator
Name of the court	dogi:court
Section (e.g. number)	dogi:courtSection
Judicial division (e.g. criminal, civil)	dogi:judicialDivision
The country in which the court or tribunal is seated	dc:coverage
The date of the decision	dcterms: date
Language of the instance document	dcterms:language
Document type (e.g. Judicial Decision)	dcterms:type
Title (Preferably the name of the parties or an alias according to national practice)	dcterms:title
Date of deposit	dcterms:issued
The (commercial or public) organization responsible for the publication	dcterms:publisher
Names of judges, Advocate general or other staff involved	dcterms:contributor
Classification (field of law)	dcterms:subject
Abstract or summary of the case	dcterms:abstract
Textual Citation	dogi:textualCitation
Document URL	dogi:text_url
References to other resources	dcterms:references
Resources referring the judicial decision (e.g. a DoGi record)	dogi:isReferencedBy

Table 5: dogi:CaseLaw

Similar remarks made for legislation hold for publication of European and national case law in public portals: the European e-Justice portal and its national sections should be in the near future responsible for the publication, conforming to the aforementioned standards, of judicial decisions in Europe.

Opening DoGi classification scheme

Even more important in order to share with external data providers the conceptual organization of the legal doctrine materials stored in DoGi is the opening of its classification scheme. DoGi classifies its resources according to a consolidated classification scheme covering all areas of national (Italian) law and international law. It is divided into 24 micro thesauri covering general areas of law, each structured in a three level hierarchy of systematic descriptors. Additionally a flat list of analytical terms related to systematic descriptors is provided for further specification of the articles classification. SKOS (the Simple Knowledge Organization Systems) is a W3C Recommendation¹⁴ that provides a means for representing knowledge organization systems (including controlled vocabularies, thesauri, taxonomies and folksonomies) in a distributed and linkable way. From the XML dump of the conceptual scheme stored on the DoGi database, containing for each descriptor its alphanumeric identifier and its lexicalization, we provided a conversion to SKOS preserving the original structure and content. The descriptors have been assigned a globally unique identifier by prefixing their original id with the DoGi prefixes

<http://www.ittig.cnr.it/dogi/descriptor/> for systematic descriptors

<http://www.ittig.cnr.it/dogi/term/> for analytical terms

obtaining for example the following

Listing 3: SKOS/RDF representation for a DoGi description.

```
<skos:Concept rdf:about='http://www.ittig.cnr.it/dogi/descriptor#s1018'>
```

¹⁴<http://www.w3.org/TR/skos-reference>

```
<rdf:type rdf:resource=''http://www.w3.org/2002/07/owl#
  NamedIndividual''/>
<skos:prefLabel xml:lang=''it''>Responsabilita penale}</skos:
  prefLabel>
<skos:broader rdf:resource=''http://www.ittig.cnr.it/dogi/
  descriptor#c0014''/>
<skos:related rdf:resource=''http://www.ittig.cnr.it/dogi/
  term#d4977''/>
<!-- mapping to BNCf Nuovo Soggettario -->
<skos:exactMatch rdf:resource=''http://purl.org/bncf/tid
  /12679''/>
</skos:Concept>
```

The SKOS representation allows DoGi descriptors to be referred by external sources on the web as well as to provide translation in different languages and semantic mapping to external vocabularies. As a first step the correspondences provided by an ongoing joint effort with BNCf (National Library of Florence) of mapping legal descriptors of the "Nuovo Soggettario" with DoGi descriptors, have been integrated in the SKOS version of the concept scheme using `skos:semanticRelation` (`exactMatch`, `broaderMatch`, `narrowerMatch`) properties. This lays the foundations for example to use the DoGi vocabulary published on the Linked data space as a bridge for multilingual access, programmatic access from third parties archives to DoGi data by legal subject and fetching resources from external archives linked with mapped descriptors.

Part of the experimentation consisted in loading the SKOS version of the DoGi vocabulary in the open source platform for collaborative multilingual thesaurus editing LegalVocbench. LegalVocbench is based on the web platform Vocbench,¹⁵ originally developed by FAO for the management of the agricultural thesaurus AGROVOC and

¹⁵<http://aims.fao.org/tools/vocbench-2>.

recently specialized by ITTIG for the management of legal vocabularies. This will eventually allow further enrichment of the DoGI vocabulary in its content (e.g. multilingual labelling, adding legal definition), structure (e.g. adding semantic relations among descriptors to further enhance of "related material" retrieval) and alignment to other legal or general purpose value vocabularies available on the web.

Conclusions

The interoperability of DoGi records on the web is a first step to reach different communities (institutional data providers such as governments, libraries, commercial publishers), but of course this evolution involves an effort of standardization to be shared among various communities. This work has been carried out starting from legal literature with the intent to make legal documentation stakeholders aware of this need, identifying metadata models and schemas, standards and protocols for interoperable information systems and supporting the creation of new forms of *alliances*. This evolution can contribute to a strong development of legal information systems on the web, encourage the emergence of digital content and thus supporting advanced shareable searching of legal data. Furthermore, the adoption of the Library Linked Data approach in the field of legal doctrine offers a unique opportunity to provide stability and persistence of these information resources regardless of format and software tools used for their production, releasing institutions from dependence on a limited number of technical providers.

References

- Kirchberger, Christine. "The 'I' in Legal Information Retrieval". *Vem reglerar informationssamhället?* Stockholm: Jure AB, 2010. <http://ssrn.com/abstract=1957558>. (Cit. on p. 228).
- Rescigno, Ugo. "Il giurista come scienziato". *Rivista di Diritto Pubblico* 3. (2003): 833–864. (Cit. on p. 229).

TOMMASO AGNOLONI, Istituto di Teoria e Tecniche dell'Informazione Giuridica del CNR (ITTIG), tommaso.agnoloni@ittig.cnr.it

ELISABETTA MARINAI, Istituto di Teoria e Tecniche dell'Informazione Giuridica del CNR (ITTIG), elisabetta.marinai@ittig.cnr.it

GINEVRA PERUGINELLI, Istituto di Teoria e Tecniche dell'Informazione Giuridica del CNR (ITTIG), ginevra.peruginelli@ittig.cnr.it

MARIA TERESA SAGRI, Istituto di Teoria e Tecniche dell'Informazione Giuridica del CNR (ITTIG), m.t.sagri@ittig.cnr.it

DANIELA TISCORNIA, Istituto di Teoria e Tecniche dell'Informazione Giuridica del CNR (ITTIG), daniela.tiscornia@ittig.cnr.it

Agnoloni, T., E. Marinai, G. Peruginelli, et al. "Annotation schema for legal doctrine: a case study on DoGi database". *JLIS.it*. Vol. 4, n. 1 (Gennaio/January 2013): Art. #5480, p. 1–249. DOI: [10.4403/jlis.it-5480](https://doi.org/10.4403/jlis.it-5480). Web.

ABSTRACT: Interoperability today is the key term for the enhancement of databases published on the web: the data, when isolated, have little value, on the contrary, their value increases significantly when different datasets, produced and published independently by different providers, can be reused and freely mashed by third parties. The use of data for new purposes not foreseen by organizations and individuals who publish "raw data" is one of the advantages of linked open data model. To achieve these benefits content and relationships between the entities described in the dataset should be explicitly represented in standard web formats (XML, RDF,

URI). The DoGi-Legal Literature database, one of the most valuable sources for online access to legal doctrine, created and managed by the Institute of Legal Information Theory and Techniques of the CNR is following this direction. This paper will define the schema of the data representing the database in RDF format. This will make the DoGi database interoperable between different data and service providers (libraries, publishers, information services for accessing national and European legal information), allowing the creation of new advanced services to be made available on the web of data. In particular, the contribution will focus on the goal to promote semantic interoperability between the DoGi classification schema and other semantic indexing tools in legal domain.

KEYWORDS: Library linked data; DoGi-Legal Literature database

ACKNOWLEDGMENT: Maria Teresa Sagri, Ginevra Peruginelli, Daniela Tiscornia and Elisabetta Marinai are authors of paragraphs at pages 1–5; Tommaso Agnoloni is author of paragraphs on pages 5–21; Ginevra Peruginelli is author of conclusions on page 21.

Submitted: 2012-04-25

Accepted: 2012-08-31

Published: 2013-01-15

