LEGAL CROWDSOURCING AND RELATIONAL LAW: WHAT THE SEMANTIC WEB CAN DO FOR LEGAL EDUCATION

*POMPEU CASANOVAS**

ABSTRACT

Crowdsourcing and Relational Law are interrelated concepts that can be successfully applied to the legal domain and, more specifically, to the field of legal education. 'Crowdsourcing' means 'participation of people (crowds)' and refers theoretically to the aggregated production of a common knowledge in a global data space. 'Relational law' refers to the regulatory link between Web 2.0 and 3.0, based on trust and dialogue, which emerges from the intertwining of top-down existing legal systems and bottom-up participation (the Web of People). Legal education today has a major role to play in the broad space opened up in terms of future potential of the Semantic Web. The following paper places a lens on the educational value of crowdsourcing and the relational approach to governance and law.

I. INTRODUCTION: THE SEMANTIC WEB AND LEGAL EDUCATION

Technology is now a deeply entrenched part of modern legal academic research and legal education. A glance at annotated bibliographies shows that researchers pay considerable attention to wikis, blogs, multimedia, open-access publications, web-based tools, e-learning platforms in their teaching.[1] For the past five years technology developers, organisations, and legal educators have been highlighting the increasing potential of the Web of Data or Web 3.0,

* Director of Advanced Research at the Institute of Law and Technology, Autonomous University of Barcelona, Spain <http://idt.uab.cat>; Academic Board of Directors of the International Joint Doctorate Degree on Law, Science and Technology (EU-Erasmus Mundus 2011); 2012 Salvador de Madariaga Fellow at the Laboratory of Decision Support and Dispute Management School of Management and Information Systems, Victoria University, Melbourne.
Acknowledgments: Programa Salvador de Madariaga, (PR2011-0550). I have used also some work carried out for the EU Projects Erasmus Mundus Doctorate on Law, Science and Technology, 520250-1-2011-1-IT-ERA MUNDUS-EMJD; CAPER (Topic FP7-SECURITY-2010-1.2-1. Project 261712); and JUSTMEN (156171-LLP-1-2009-1-IT-ERASMUS). I am grateful to Marta Poblet, Meritxell Fernández-Barrera, Núria Casellas, and Albert Meroño for their kind assistance in quoting their work. I am grateful to Gina Curro for the editing of the last version of the paper. She substantially improved the original one. Last but not least, I am grateful to John Zeleznikow, Pauline Stanton, and Elizabeth Wilson-Evered for their warm reception at Victoria University.

1 See Pearl Goldman, 'Legal Education and Technology: An Annotated Bibliography' (2001) 93 *Law Library Journal* 423; 'Legal Education and Technology II: An Annotated Bibliography' (2008) 100 *Law Library Journal* 415. Automatic updating of bibliographies is a main topic for SW researchers. For some examples of re-usability in legal education related to students' generated content, see Heather Warnock, Michael C Bromby and Moira MacMillan, 'Aspects of Discrimination in Employment Law: An Annotated Bibliography' (August 1, 2011). Available at SSRN: <http://ssrn.com/abstract=2056276>; <http://michaelbromby.wordpress.com/teaching/annotated-bibliographies/>.

the next stage of the Semantic Web leaded by the W3C Consortium.[2] A number of recent studies have already addressed cognitive development in relation to educational skills and abilities fostered by the new stage of the Web.[3]

The dissemination of Law and the Semantic Web (SW) has taken place inside the boundaries of highly specialised scientific and technological communities, such as the International Association for Artificial Intelligence and Law (IAAIL), the JURIX Foundation for Legal Knowledge Systems or the Organization for the Advancement of Structured Information Standards (OASIS).[4] In peer reviewed articles of major legal journals, descriptions of the SW are usually offered in connection with specific legal topics such as intellectual property[5], copyright[6], anonymisation[7], sentencing[8], or the concerns for liberty and freedom of speech raised by the growing technical possibilities to get control over the development of the Internet. Mailland, for example, has even dramatically warned that "censorship is the semantic web's lifeblood".[9] The legal approach to the structured content of texts into the Internet has focused more on rights, liability, legal effects, the reconstruction of the commons[10], and the qualification layer of "second level agreements"[11] than on the transformative power of technology and the changing nature of regulations.

2 The Semantic Web is a web that is able to describe things in a way that computers can understand. Cfr. <http://www.w3schools.com/web/web_semantic.asp> and <http://www.semantic-web-journal.net>.The World Wide Web Consortium (W3C) is an international community that develops open standards to foster technical innovation and to ensure the long-term growth of the Web. Cfr. <http://www.w3.org/> . The Web of Data is related to the Internet of Things, the increasing possibility to link, store and retrieve identifiable objects (things) and their virtual representations in an Internet-like structure. Cf <http://www.w3.org/standards/semanticweb/>. Regarding education, see ASTD, 'Better, Smarter, Faster: How Web 3.0 Will Transform Learning In High Performing Organizations' (2011) 2 (7) *ASTD Research*; Jason Ohler 'The Semantic Web in Education. What happens when the read-write web gets smart enough to help us organize and evaluate the information it provides' (2008) 4 *Educause Quarterly* 7; Paul Anderson 'What is Web 2.0? Ideas, technologies and implications for education' JISC Technology and Standards Watch (Feb. 2007), <http://www.jisc.ac.uk/media/documents/techwatch/tsw0701b.pdf>; Paul Anderson 'Mobile technologies and their use in education - new privacy implications.' JISC Technology and Standards Watch (Jul. 2010) <www.jisc.ac.uk/media/documents/techwatch/acf11b0.pdf> .

3 Cf P Isaias et al, (eds) *Towards Learning and Instruction in Web 3.0. Advances in Cognitive and Educational Psychology* (Springer, 2012).

4 Cf < http://www.iaail.org/, http://www.jurix.nl/, https://www.oasis-open.org/> .

5 Andrew Clearwater, 'The New Ontologies: the Effect of Copyright Protection on Public Scientific Data Sharing Using semantic Web Ontologies' (2010) 10 *John Marshall Review of Intellectual Property Law* 182.

6 Julien Mailland, 'The Semantic Web and Information Flow: a Legal Framework' (2010) 11 *North Carolina Journal of Law & Technology* 269.

7 Jane Yakowitz, 'Tragedy of the Data Commons' (2011) 25 *Harvard Journal of Law & Technology* 1.

8 Giancarlo F Frosio□, 'Google Books Rejected: Taking the Orphans to the Digital Public Library of Alexandria' (2011) 8 *Santa Clara Computer and High Technology Law Journal* 81.

9 Mailland, above n 6, 294.

10 *Commons* refers to the cultural and natural resources accessible to all members of a society, held in common, not owned privately. Cf <http://en.wikipedia.org/wiki/Commons>. Referring to the SW, following the institutional analysis steps of Elinor Ostrom, cfr. Jorge L Contreras, 'Data Sharing, Latency Variables, and Science Commons' (2010) 25 Berkeley Technology Law Journal 1601; see alsoYakowitz, above n 7.

11 Second level agreements are 'preemptive licenses granted by copyright owners to platforms operators, with the purpose of ratifying the mass usage of copyrighted content by their users', see Yafit Lev-Aretz □, 'Second Level Agreements' (2012) 45 *Akron Law Review* 137.

By contrast, law librarians have paid close attention to the possibilities of the Web of Data. Techno-legal blogs[12]and library journals[13] have increasingly been hosting discussions and reflections on SW technologies and Linked Open Data (LOD). The link between Legal libraries, LOD and the SW has emerged as a natural one, since the work of librarians is related to the classification of what is usually known as *metadata* or *metacontent* (data providing information about one or more aspects of texts, images or multimedia items): author of data, means of creation, purpose, time and date, location, standards used.

In this paper, *crowdsourcing* and *relational law* are first defined (II). Next, the relations between the Social Web and the Web of Data are discussed (III), followed by the principles of linked open data and the concept and scope of relational law (IV). Case studies from past and present research projects are outlined (V) and lastly, the conclusions for legal education (VI) are presented.

II. Preliminary Definitions

A. *Crowdsourcing and Relational Law*

Crowdsourcing is the online resolution of micro-problems.[14] For the purposes of this paper, a more complex definition is required: the aggregated production of a common knowledge, stemming from individual contributions, in a global data space. This global space is currently estimated to contain nearly 32 billion *Resource Description Framework* (RDF) triples with half a billion links between them.[15] This was the case according to the last Statistical Report issued by the W3C in September 2011[16]. In August that year industry research reported the first load

12 See, eg, < http://legalinformatics.wordpress.com> , the Blog sustained by Robert Richards.

13 Corey A Harper and Barbara B Tillett, 'Library of Congress controlled vocabularies and their application to the Semantic Web' (2007) 43 (3/4) *Cataloging and Classification Quarterly* 47-68; Ed Summers, Antoine Isaac, Clay Redding, Dan Krech, 'LCSH, SKOS and Linked Data' (2008) *Proceedings of the International Conference on Dublin Core and Metadata Applications* (DC 2008, Berlin, 22-26 September 2008; Benjamin J. Keele and Michelle Pearse, 'How Librarians Can Help Improve Law Journal Publishing' 104 *Law Library Journal* 383. 'SKOS' stands for 'Simple Knowledge Organization System', a common data model for sharing and linking knowledge organization systems through the Web: 'Using SKOS, concepts can be identified using URIs, labelled with lexical strings in one or more natural languages, assigned notations (lexical codes), documented with various types of note, linked to other concepts and organized into informal hierarchies and association networks, aggregated into concept schemes, grouped into labelled and/or ordered collections, and mapped to concepts in other schemes.' Cf < http://www.w3.org/TR/skos-reference/>.

14 See Part IV A below in this paper.

15 RDF is 'a language for representing information about resources in the World Wide Web. It is particularly intended for representing metadata about Web resources, such as the title, author, and modification date of a Web page, copyright and licensing information about a Web document, or the availability schedule for some shared resource' <http://www.w3.org/TR/rdf-primer/> with several examples of how to formalise concepts like 'author John Smith' or 'town of New York'. Any expression in RDF is a collection of 'triples', each consisting of a subject, a predicate and an object. Cf on triples <http://www.w3.org/TR/rdf-concepts/>, especially s 3.1, 6.1 and s 6.2. RDF refers to the metadata bases for the Semantic Web. An easy way to be introduced to Semantic Web technologies is through the hierarchy of languages of the Semantic Web Stack, where each layer exploits and uses capabilities of the layer below. See <http://en.wikipedia.org/wiki/Semantic_Web_Stack> .

16 See Chris Bizer, Anja Jentzsch and Richard Cyganiak, *State of the LOD Cloud* (Version 0.3, 09/19/2011), <http://www4.wiwiss.fu-berlin.de/lodcloud/state/>.

and query of 1 trillion RDF triples.[17] Actually, the number of triples was estimated to be more than 52 billion in October 2012.[18]

The term, *relational,* is concerned with a common feature that emerges from the existing social and economic bonds among companies, providers, customers, consumers, citizens (digital neighbors) or teachers and students. It refers to the capacity to set up a common space of mutual relations —a shared regulatory framework— in which some reciprocity is expected with regard to goods, services, attitudes and actions. Thus, 'relational law' is more connected to trust and dialogue than to enactment of formal procedures or on the enforcement of sanctions.

Both crowdsourcing and relational law are concepts that have application to the legal domain, especially legal education. A number of platforms such as W3C Semantic Education, Linking Open Data (LOD) and the recent Linked Legal Data (LII Cornell) lend themselves well to investigation, especially in terms of improving communication and enhancing the relationship between Web 2.0 (the Social Web) and 3.0 (the Web of Data). The aim of this paper is to demonstrate that legal education has a major role to play in this new broad space of structured and manageable data.

B. *The Social Web and the Web of Data*

Through the Internet, information management has infiltrated our lives like never before. We rely on it for producing and reproducing new knowledge in different ways. This knowledge is general, but also local and *personal.* It is developed through a dynamic and complex network of individual, collective and sometimes coordinated interactions within multiple changing environments. From a tiny fragment or a slight nebulous idea about something wanted, a process of refinement and discovery is regularly produced through queries and the interface with the web.[19]

Today, browsing the web can be overwhelming. Typically, knowledge must be carved from an excessive amount of information. This is a social, proactive and dynamic process users cannot skip. Today's web users have to enter into a dialogue with themselves and with the *social knowledge* produced and unevenly distributed on the web; it is up to them to filter, select, aggregate and eventually mash it up.[20] When users browse and query the web, they expect it to understand them, as if the system is able to speak their natural language and refer intentionally to the same cognitive objects they are referring to. 'People don't want to search'.[21]

Actually, this is what the Semantic Web (SW) is all about —the smart interface between systems and users. The Web is not the Internet; the Web organises and processes the information being transformed into knowledge through the interaction with the end users who both consume and produce it (*prosumers*)[22]. Someone finds a song, or a videotape (or an interesting document) they like after a query process. They might upload it again and a transformative process occurs by which not only some comments and ideas are exchanged. However people might use, re-use and work out the item itself. This process can be described as follows: (i) A song linked with

17 Total load was 1,009,690,381,946 triples in just over 338 hours for an average rate of 829,556 triples per second, using AllegroGraph, <http://www.w3.org/wiki/LargeTripleStores>.

18 To be precise, 52,381,770, 554, Virtuoso (CoRelational) DBMS Benchmarks -- LOD Cloud Cache Instances (8-node cluster with 48GB Ram Per Node) <https://docs.google.com/spreadsheet/ccc?key=0AihbIyhlsQSxdHIxc3hhdk82UFdYd1ppaGw3WDNrVGc#gid=0>. This is referred only to DBpedia (the RDF companion to Wikipedia, not updated yet). The 2000 USA census alone contains 1 billion RDF triples, cf Joshua Tauberer, <http://www.rdfabout.com/demo/census/>. The semantic layer of the Internet remains largely unknown.

19 See Ricardo Baeza-Yates and Prabhakar Raghavan, 'Next Generation Web Search' in S Ceri and M Brambilla (eds), *Search Computing* (Springer Verlag, LNCS 5950, 2010) 11-23.

20 Cf <http://en.wikipedia.org/wiki/Mashup>.

21 See Ricardo Baeza-Yates, 'People don't want to search' (2009) *TNW. The Next Web*, 16 April 2009, <http://thenextweb.com/2009/04/16/ricardo-baezayates/> .

22 Cf <http://en.wikipedia.org/wiki/Prosumer>.

users as a *document*; (ii) people being linked by the song adding and sharing knowledge about it through an *interactive network*; (iii) and the song itself as a semantic object being *structured and linked as data* to other semantic objects. The Web, the Social Web and the Web of Data, all use written text and documents to communicate or *convey* meaning.

Extracting, *using* and *reusing* this shared meaning are different aspects of the acquisition, storage, retrieval and transformation of information into knowledge. These processes are intricate. Songs, paintings, or writings on the Web consist of *structured abstract objects* that cannot be confused with the structure of the physical objects. The same could be said about legal knowledge. On the one hand, knowing means *representing* and *processing* semantic entities in order to be able to reproduce and manipulate them; on the other hand, knowing means *acting* upon them to enrich the whole process with additional information.[23]

Web 2.0[24] and Web 3.0[25] are abbreviated forms broadly used to refer to new extensions of the Web. For example, blogs (and Blawgs), Wikis, Podcasts/Video Blogs (A/V Blogs), tags, mashups (AJAX), and Web Services (API) are among these new technologies that users switch between.

From a technical point of view the Semantic Web (SW) consists of a number of computer languages.[26] These semantic languages are capable of modelling data, annotating and relaying information —RDF [*Resource Description Framework*] and OWL [*Ontology Web Language*]. RDF facilitates the description of knowledge using triples, encoding factual and linguistic knowledge; OWL facilitates simple deductive reasoning through sets and properties that model formal concepts, relationships and instances. These are graph-languages on XML [*eXtended Mark-up Language*], "serialized" and *representing* data in files (using Turtle or Phyton, e.g).[27] The result presents information management and processing as *knowledge —hypertext* links, connection of objects, and information retrieval from the Web using not keywords (terms), but concepts.

Semantics are used to harness the human-machine interface. OWL formalizes what is known as ontologies. 'Ontology' is a philosophical term that refers to a systematic account of Existence.[28] Ontologies, plural, is nowadays a common term in engineering and computational science — because 'for knowledge-based systems, what 'exists' is exactly that which can be represented'.[29] What is meant by ontologies, then, is the explicit and formal specification of

23 See Ricardo Baeza-Yates and Yoelle Maarek. 'Web retrieval: the role of users' in ACM (2011) *Proceedings of the 34th international ACM S,GIR conference on Research and development in Information Retrieval* 1303-1304.

24 This term from the late nineties became widely used after the 2003 Conference on Web 2.0, hosted by O´Reilly Media and Medialive. See the contrasting description between Web 1.0 and Web 2.0 by Tim O'Reilly in 'What is Web 2.0. Design Patterns and Business Models for the Next Generation of Software', 30 September 2005, at <http://oreilly.com/web2/archive/what-is-web-20.html>.

25 This term was coined in 2006 within the discussions on Web 2.0 applications between Jeffrey Zeldman and Nova Spivak. It was popularized by the journalist John Markoff in 'Entrepeneurs See a Web Guided by Common Sense', *The New York Times, 12 November 2006*, <http://www.nytimes.com/2006/11/12/business/12web.html?_r=1&pagewanted=1&ei=5088&en=254d697964cedc62&ex=1320987600>.

26 See above n 15.

27 'Serialization is the process of converting the state of an object into a form that can be persisted or transported. The complement of serialization is deserialization, which converts a stream into an object. Together, these processes allow data to be easily stored and transferred.' <http://msdn.microsoft.com/en-us/library/7ay27kt9%28v=vs.80%29.aspx>.

28 "Ontology as a branch of philosophy is the science of what is, of the kinds and structures of objects, properties, events, processes, and relations in every area of reality." See Barry Smith, 'Ontology', in Luciano Floridi (ed) *Blackwell Guide to the Philosophy of Computing and Information* (Blackwell, 2003) 155–166.

29 See Tom Gruber, 'A Translation Approach to Portable Ontologies' (1993) 5 (2) *Knowledge Acquisition* 199-220.

the conceptual structure of a given knowledge.[30] It is through ontologies that the Web is able to "understand" the meaning of queries formulated by users in their natural language (such as English, Spanish or French).

Legal ontology engineering, the formal structuring of concepts, is a well-trodden path, as ontologies have been built up for twenty years or more to be applied to security, e-libraries, e-commerce, e-court, e-government and e-administration.[31] Five years ago it was still possible to summarise and organise legal ontologies according to their technical features and degree of development.[32] However this task is not possible anymore because their use has become so common in the legal domain. In addition, it is important to keep in mind that ontologies are intrinsically plural and intentional, and there is more than one way to build them as they depend upon different and equally acceptable theoretical approaches (cognitive, linguistic, normative, socio-legal), according to the different perspectives and final ends for which they are designed.[33]

III. Linked Open Data

In a presentation at the W3C[34] in July 2006 Tim Berners-Lee advocated four basic principles for Linked Data to operate effectively:

1. Use URIs [*Uniform Resource Identifiers*] 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 [*SPARQL Protocol and RDF Query Language*])
4. Include links to other URIs, so that they can discover more things.

Three years later at a TED talk, he presented the proposal in an even simpler way: *All kinds of conceptual things now have names that start with HTTP.*[35] This means relationships among subjects, data and metadata —the Web of Data— can be viewed as one big relational database of knowledge. As observed by Heath and Bizer (2011), the Web of Data is an additional layer that is tightly interwoven with the classic Web document and has many of the same properties:

1. The Web of Data is generic and can contain any type of data.
2. Anyone can publish data on the Web of Data.
3. The Web of Data is able to represent disagreement and contradictory information about an entity.
4. Entities are connected by RDF links, creating a global data graph that spans data sources and enables the discovery of new data sources. This means that applications do not have to be implemented against a fixed set of data sources, but they can discover new data sources at run-time by following RDF links.

---

30 Ontologies can be defined as an explicit specification of a shared conceptualization. They consist of concepts (classes), relationships (properties), instances and axioms. See Gruber, above n 29.

31 Legal ontologies can be divided into core-ontologies (legal theory, structure of rules etc.) and domain-ontologies (for example: criminal law, civil law). For a systematic overview see Núria Casellas, *Legal Ontology Engineering: Methodologies, Modelling Trends, and the Ontology of Professional Judicial Knowledge* (Springer, LGT Series, 2011).

32 See André Valente, 'Types and Roles of legal Ontologies', in Victor R Benjamins et al (eds), *Law and the Semantic Web* (Springer, LNAI 3369, 2005) 65-76; Joost Breuker et al (eds), *Law, Ontologies and The Semantic Web. Channelling the Legal Information Flood* (IOS Press, 2009). See the most complete existing table of legal ontologies so far in Casellas, above n 31, 109 (table pp. 147-150).

33 See Giovanni Sartor et al (eds), *Approaches to Legal Ontologies. Theories, Domains, Methodologies* (Springer, LGT Series, 2011).

34 See Tim Berners-Lee, 'Linked Data', 27 July 2006 (last change 18 June 2009) <http://www.w3.org/DesignIssues/LinkedData.html>.

35 See Tim Berners-Lee, 'Tim Berners-Lee on the next Web', *TED Talk* (Videoconference) February 2009 <http://www.ted.com/talks/tim_berners_lee_on_the_next_web.html>.

5. Data publishers are not constrained in their choice of vocabularies with which to represent data.
6. Data is self-describing. If an application consuming Linked Data encounters data described with an unfamiliar vocabulary, the application can dereference[36] the URIs that identify vocabulary terms in order to find their definition.
7. The use of HTTP as a standardized data access mechanism and RDF as a standardized data model simplify data access compared to Web APIs, which rely on heterogeneous data models and access interfaces.[37]

Since 2007, a number of ongoing projects have involved recollecting and developing these ideas. The W3C project *Linking Open Data* updates the link between data using RDF.[38] Perhaps the most popular visualisation of linked data is the one produced by the DBpedia Project on all the information contained in the databases feeding the Wikipedia[39] [See Fig. 1]. DBpedia extracts information and links new sites as information resources describing millions of things.[40] Visualization allows checking the shared elements present at different databases (thick arrows indicate a greater degree of linkage; bidirectional arrows point at the coexistence of elements).

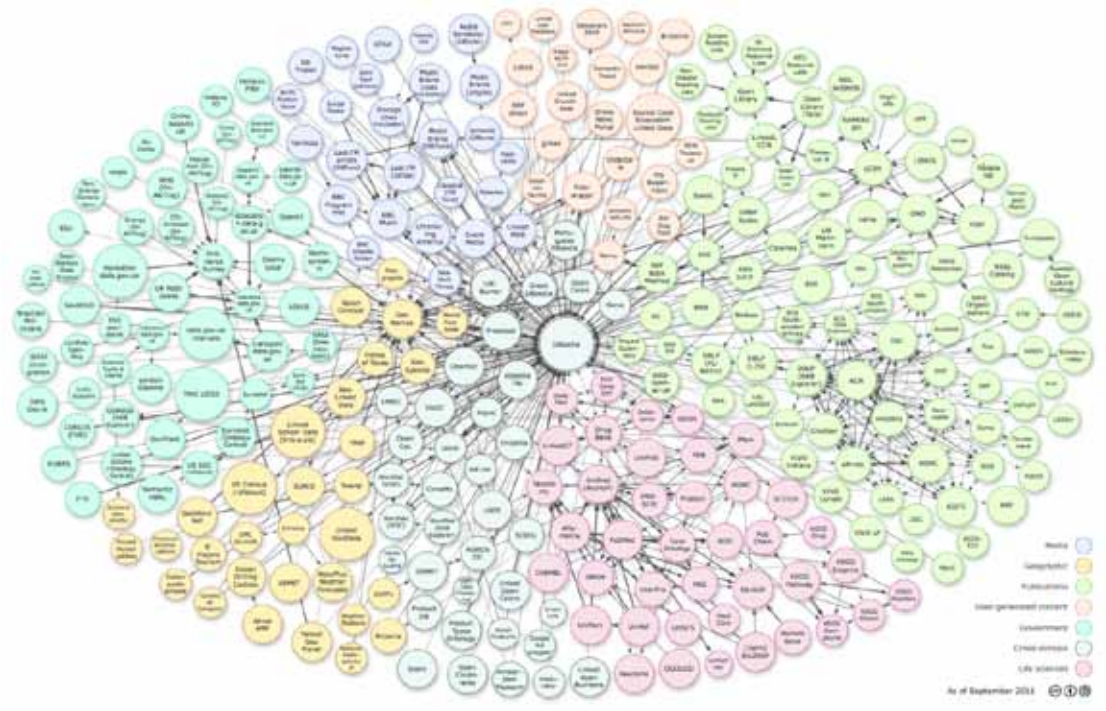

Fig. 1. Cloud Diagram of DBpedia *Linking Open Data***. Source: Richard Cyganiak and Anja Jentzsch, http://lod-cloud.net/** *instance linkages within the linking open data datasets* **[quoted with permission].**

36 'The act of retrieving a representation of a resource identified by a URI is known as *dereferencing* that URI. Applications, such as browsers, render the retrieved representation so that it can be perceived by a user. Most Web users do not distinguish between a resource and the rendered representation they receive by accessing it.' Rhys Lewis (ed), *Dereferencing HTTP URIs*, 31 May 2007, W3C, <http://www.w3.org/2001/tag/doc/httpRange-14/2007-05-31/HttpRange-14>.

37 Tom Heath and Christian Bizer, *Linked Data: Evolving the Web into a Global Data Space* (2011), <http://linkeddatabook.com/editions/1.0/#note3>.

38 See <http://www.w3.org/wiki/SweoIG/TaskForces/CommunityProjects/LinkingOpenData> .

39 See <http://dbpedia.org/About> .

40 Cf <http://en.wikipedia.org/wiki/DBpedia> See for a description of the information extraction method, C Bizer et al, 'DBpedia - A Crystallization Point for the Web of Data' (2009) 7 (3) *Journal of Web Semantics*: *Science, Services and Agents on the World Wide Web* 154–165.

> The term Linked Data refers to *a set of best practices* (emphasis added) for publishing and connecting structured data on the Web. These best practices have been adopted by an increasing number of data providers over the last three years, leading to the creation of a global data space containing billions of assertions —the Web of Data.[41]

Although the Web of Data is a 'network of items in the world, described as data on the Web',[42] it actually links beyond data to end users, citizens, *prosumers:* people:

> There are cries from the heart (e.g. The Open Social Web Bill of Rights for my friendship, that relationship to another person, to transcend documents and sites).There is a "Social network Portability" community. It's not the Social Network *Sites* that are interesting —it is the Social Network itself. The Social Graph. *The way I am connected, not the way my Web pages are connected* [emphasis added]. We can use the word *Graph*, now, to distinguish from *Web.* I called this graph the Semantic Web, but maybe it should have been Giant Global Graph![43]

The *Web of People* or the *Giant Global Graph* is connected to law in many ways. Firstly it introduces shared ordering into the personal mashups used in 'social machines'; secondly, creating SW standards on linking data means following general principles —such as transparency and accountability— to protect people; thirdly, 'bringing humanity fully into the information loop requires data structures and computational techniques that enable us to treat social expectations and legal rules as first-class objects in the new Web architecture (to create a declarative rule-based infrastructure that is appropriate for the Web).'[44] These are the grounds for Government Linked Open Data - the practice of publishing public sector information on the Web using Linked Data.[45] There is a parallel move towards establishing a regulatory layer for the Internet, based on concerns regarding trust, security and privacy expressed by many lawyers and engineers. This is termed Privacy by Design.[46] It should be pointed out that issues related to embodying legal rules into a Web standard language are beyond the scope of this paper.

IV. RELATIONAL LAW AND LEGAL EDUCATION

Both Privacy by Design and Linked Open Data promote the *empowerment* of people, in a highly connected world —*humanity in the loop*, to use Hendler and Berners-Lee´s expression. Therefore, the role of law in this changing environment is evolving with more complex regulation patterns, involving protocols and good practices (*soft law*), governance (the relationships between citizens and public or private organizations), and ethics now playing a much bigger part.

A. *The Concept And Scope Of Relational Law*

For the purposes of this paper the above regulatory phenomena are referred to as *relational law* —the allocation of behavioral expectations (social assignment of rights and obligations),

---

41 C Bizer, T Heath and T Berners-Lee, 'Linked Data - The Story So Far", (2009) 5 (3) *International Journal on Semantic Web and Information Systems* 1-22.

42 Ibid.

1. 43 Tim Berners-Lee, **'**The Giant Global Graph' *Semantic Web Technologies*, 21 November 2007, http://dig.csail.mit.edu/breadcrumbs/node/215>

44 Jim Hendler and Tim Berners-Lee, 'From the Semantic Web to social machines: A research challenge for AI on the World Wide Web' (2010) 174 *Artificial Intelligence* 158-161.

45 <http://www.w3.org/2011/gld/wiki/Main_Page>; cf also Li Ding al, 'TWC LOGD: A portal for linked open government data ecosystems' (2011) 9 (3) *Web Semantics: Science, Services and Agents on the World Wide Web* 325-333.

46 Ann Cavoukian, '7 Laws of Identity: The Case For Privacy-Embedded Laws of Identity in the Digital Age' (2006) *Technology*, Ontario Information and Privacy Commissioner, 1-24 October; 'Privacy by Design. The 7 Foundational Principles. Implementation and Mapping of Fair information Practices' (2010) Information and Privacy Commissioner, Ontario, Canada.

in terms of a shared technological framework.[47] Interactions among people, programs and human-machine interfaces create an aggregated value that facilitates the emerging bottom-up connection between Web 2.0 and Web 3.0. Trust, security, confidence, and not only the identity of users, matter and must be assumed as features of the regulatory models at stake. This constitutes the ecological niche in which different types of technology and behaviors (human or artificial) converge for a common result.

The Web of Data (WD) and Privacy by Design (PbD) appeared at the same time.[48] In fact 2006 was the year in which the second version of *Code* was delivered by Lawrence Lessig as well.[49] Lessig´s contributions to a better regulation of the Internet are invaluable —Creative Commons[50], and more recently, his fight against corruption are broadly known and continue to influence the field. As a result of the metasystem layer being shaped by big companies and governments, Lesssig has been working on overcoming obstacles which undermine the web from reaching its full potential.

When Lessig wrote *Code Version 2.0* he adopted a different approach to that used for the first version.[51] Through a wiki tool he encouraged contributions from interested students and others to reshape his book.[52] This represents an example of *educational crowdsourcing*, and demonstrates how people can contribute to a common outcome through a shared process of learning.

Originally, the term *crowdsourcing* was introduced by Jeff Howe in 2006 referring to "the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call".[53] Different types of crowdsourcing have recently been discussed in the literature.[54] Wikipedia, an example of crowdsourced encyclopedia, defines the term as a non-profit collective aggregation of information from micro-tasks widely distributed across the Web, and freely performed by people.[55] Therefore, crowdsourcing implies much more than a new way to collect information or to respond to labor offers or contests, following the Amazon Mechanical Turk or Microworker marketplace models.[56] Within the Semantic Web community, wikis were used to produce shared

47 For more detailed explanations, see Pompeu Casanovas 'Agreement and Relational Justice: A Perspective from Philosophy and Sociology of Law' in Sascha Ossowski (ed), *Agreement Technologies* (Springer, LGT Series, 2012) 19-42. Also see *'The future of Law: Relational Justice and Next Generation of Web Services',* in Meritxell Fernández-Barrera et al (ed), *Law and Technology. Looking into the Future. Selected Essays* (European Press Academic Publishing, 2010) 137-158.

48 See Ann Cavoukian and Jeff Jonas, 'Privacy by Design in the Era of Big Data', June 8 2012, <http://privacybydesign.ca/content/uploads/2012/06/pbd-big_data.pdf >.

49 Lawrence Lessig, *Code and other laws of the Cyberspace. Version 2.0*. (Basic Books, 2006) <https://www.socialtext.net/codev2/table_of_content>.

50 For an international version of CC compatible with the technical points of view sustained in this paper, see Danièle Bourcier et al (eds), *Intelligent Multimedia. Managing Creative Works in a Digital World* (European Press Academic Publishing, 2010).

51 Cf Lawrence Lessig, *Code and Other Laws of the Cyberspace* (Basic Books, 1999).

52 See the *Code Version 2.0* Website <http://code-is-law.org>. In addition, the Spanish version of *Code 2.0* was also cooperatively translated by students from *Tecnología de la Comunicación Audiovisual* at the University of Málaga (Spain), Cfr. <http://es.wikipedia.org/wiki/El_Código_2.0>.

53 Jeff Howe, 'The rise of crowdsourcing', *Wired*, June 14 2006, <http://www.wired.com/wired/archive/14.06/crowds.html>. Cf also James Surowieki, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations* (Brown, 2004).

54 See David Geiger et al, 'Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes' (2011) *AMCIS-Proceedings of the Seventeenth Americas Conference on Information Systems* (Detroit, Michigan August 4th-7th) 8.

55 Cf < http://en.wikipedia.org/wiki/Crowdsourcing>.

56 Cf <https://www.mturk.com>, <http://microworkers.com/>.

knowledge and keep track of the discussions required to build up ontologies.[57] Crowdsourcing has a more public and open dimension that cannot be ignored, because of the *personalization* of services and applications, and its link to Web 2.0 and 3.0.

Crowdsourcing therefore, creates the conditions to aggregate individual information into collective, common knowledge,[58] and it contributes to broaden and enhance democratic ways of living and acting in the global world. The challenge for education lies in the technological empowerment that the Semantic Web, mobile technologies, grid computation and cloud computing afford.[59] How does legal education relate to this? Does "the cloud" make any difference with respect to previous educational processes? In particular, does cloud computing have any effect on the learning outcomes of students in Law Schools today?

The differences between virtual courses and conventional ones are obvious, but actually large universities remain competitive by offering online courses similar to those of the Open Universities. The recent examples of Coursera and OpenLearning demonstrate that massive numbers of students can engage in online courses, making it possible for a new business model to challenge the traditional one.[60] These are becoming to be known as Massive Online Open Courses (MOOC). They are mainly focused on delivering content but, as Professor Buckland states, it is not clear yet how they will relate with other academic values as community building, learning from peers, tutorials, practical work, and motivation to study and progress.[61]

One of the many criticisms of this type of learning is that it may compromise academic integrity, especially in terms of assessment with students engaging in copy and paste practices or plagiarism.[62] Students are always a step ahead of the game; they know how to browse the web, to exploit the most common tools in order to get the information they need; and they have extended horizontal links within the Social Web.[63] Why would students in legal learning be an exception to these practices? How students and teachers overcome these practices is highly dependent on the organisational and technological context in which they operate. What are the sources, mashups and intelligent tools which retrieve information and transform it into *personal* knowledge? And how can these applications create the right type of environment in which to share discoveries?

---

57 Pompeu Casanovas et al, 'OPJK and DILIGENT: Ontology Modelling in a Distributed Environment' (2007) 15 (2) *Artificial Intelligence and Law* 171-186.

58 Collective knowledge is related, but not identical, to collective rationality —Condorcet's paradox, Arrow's impossibility theorem, discursive dilemma. I use the term as the result of *collective intelligence*, 'a shared or group intelligence that emerges from the collaboration and competition of many individuals', The MIT Center for Collective Intelligence, *Handbook of Collective Intelligence*, <http://scripts.mit.edu/~cci/HCI/index.php?title=Main_Page>.

59 Crowdsourcing can be expanded into *crowdservicing*. See Joseph Davies, 'From Crowdsourcing to Crowdservicing' (2011) 15 (3) *IEEE Internet Computing* 92-94.

60 See <https://www.coursera.org/>; cf John Markoff, 'Online Education Venture Lures Cash Infusion and Deals With 5 Top Universities' *New York Times* 18 April 2012. All over the world, tens of thousands of students can follow the same online course.

61 See Charis Palmer, 'OpenLearning launches into competitive MOOCs marquet" *The Conversation*, 15 October 2012, <http://theconversation.edu.au/openlearning-launches-into-competitive-moocs-market-10155>.

62 Cf eg, Beverley Oliver, 'Credentials in the cloud: how will MOOCs deal with plagiarism?' *The Conversation*, 10 September 2012, <http://theconversation.edu.au/credentials-in-the-cloud-how-will-moocs-deal-with-plagiarism-8581>.

63 *Patatabrava*, the most popular social network among Spanish graduate students has grown up to more than 300.000 members, see <http://www.patatabrava.com/> (*Una manera diferente de vivir la Universidad*)

B. *Steps Towards SW Educational Services*

The Web is oriented to Web-services while the Semantic Web is oriented to SW services.[64] Computer scientists usually represent SW applications as retrospective linear stages, and in doing so they set the stage for new challenges, applications and prototypes.[65] Dicheva classifies the generations of Web-based educational systems into three broad categories[66]: (i) the first generation systems which provide a "centralized (typically client–server) architecture" and a proprietary format for representing learning resources, (ii) the second generation which add "domain conceptualization and concept-based presentation of the maintained resources"; (iii) the third generation systems which enable "ontology-aware software, reusability, exchangeability, and interoperability of the maintained learning resources and components".

Rajiv and Manohar Lal recently summarised certain characteristics of the Web 3.0[67]: (i) *intelligence* (documents in different languages to be intelligently translated into other languages, including natural languages), (ii) *personalization* (individual preferences such as information processing, search, formation of personalized portal on the web) (iii) *interoperability* (collaboration and reusability interrelated with what is called the *pervasive web*, in which applications are easy to customize and run on different kind of devices) and (iv) *virtualization* (high speed internet bandwidths and High end 3D Graphics). Accordingly, Rajiv and Manohar Lal single out certain toolkits: (i) 3D-Wikis / Virtual 3D Encyclopedia e.g Copernicus-3D Wikipedia[68]; (ii) learning with 3D Virtual worlds and Avatars (Second Life, IMVU, Active Worlds, Red Light Center...); (iii) intelligent search engines; (iv) online 3-D Virtual Labs / Educational labs / Simulations or 3D Web.

In the above applications, we do not see any reference to the social and educational conditions in which they were produced. What is missing is the treatment of the day-to-day activities involving teaching and learning, either on- or offline, in the toolkits. Legal innovative toolkits can help bridge the gap and there are three stages in their development. The first is the Free Access to Law Movement,[69] with platforms focused on making legal data available, organised in Legal Information Institutes (LIIs) that constitute the World Legal Information Institute.[70] Next

64 Semantic Web services are 'the server end of a client–server system for machine-to-machine interaction via the World Wide Web, using markup which makes data machine-readable in a detailed way' Cf <http://en.wikipedia.org/wiki/Semantic_Web_Services>. Cf also Dieter Fensel et al, *Semantic Web Services* (Springer, 2011).

65 Cf D Dicheva, 'Ontologies and Semantic Web for E-learning' in Heimo H Adelsberger et al, *Handbook on Information Technologies for Education and Training* (Springer, 2008) 47–65; H Rego et al, 'Metadata and Knowledge Management Driven Web-based Learning Information System Towards Web/e-Learning 3.0' (2010) 5 (2) *International Journal of Emerging Technologies in Learning* (iJET) 36–44; Burasakorn Yoosooka and Vilas Wuwongse, 'Linked Open Data for Learning Object Discovery: Adaptive e-Learning Systems', *INCOS '11 Proceedings of the 2011 Third International Conference on Intelligent Networking and Collaborative Systems*, IEEE Computer Society (2011) 60-67; Rajiv and Manohar Lal, 'Web 3.0 in Education & Research' (2011) 3 (2) *BVICAM's International Journal of Information Technology (BIJIT);* <http://www.bvicam.ac.in/bijit/Downloads/pdf/issue6/01.pdf>; N Rubens, D Kaplan, and T Okamoto, 'E-Learning 3.0: anyone, anywhere, anytime, and AI' *International Workshop on Social and Personal Computing for Web- Supported Learning Communities*' SPeL 2011, December 2011.

66 Dicheva, above n 65 49-50.

67 Lal, above n 48.

68 See <http://copernicus.deri.ie>.

69 Graham Greenleaf, 'The global development of free access to legal information', in A Paliwala (ed), *A history of legal informatics* (Lefis, 2010) 53-82; Tom R Bruce, 'Foundings on the Cathedral Steps' in Ginevra Peruginelli and Mario Ragona (eds) *Law via the Internet. Free Access, Quality of Information, Effectiveness of Rights* (EPAP, 2009) 441-422; Daniel Poulin, 'Fifteen Years of Free Access to Law' ibid. 15-32. See also Michael W Carroll, 'The Movement for Open Access Law' (2006) 10 *Lewis & Clark Law Review* 741.

70 <http://www.worldlii.org/>

is the effort to apply Artificial Intelligence to legal programming, helping students to master legal case-based reasoning and argumentation.[71] The final stage is the Semantic Web, linking Web 2.0 and 3.0. These three stages have evolved concurrently providing flexible, shareable and secure ways of communicating and publishing.[72]

We should distinguish between (i) producing, implementing and consuming resources for final users (Semantic Web techniques applied to Web-learning), and (ii) the converging process of developing technologies as a teaching resource. The former is captured by technological life-cycles in which technologies are tested, evaluated and eventually implemented, whereas the latter requires the combination of research and educational models, fostering the imagination of students and teachers alike towards active participation in hybrid, "mixed up" experiences.[73]

For the past ten years, Law Schools have increasingly incorporated a virtual side into regular courses, and conceived a space in which students and teachers can communicate either privately or publicly with specific dashboards (wikis, communication rooms, downloadable materials). This might occur with different types of crowdsourcing, in different scales, from local experiences to global responses, as exemplified by Coursera. More importantly, it can be argued that web science and technological tasks introduced into the legal curricula prepares law students in learning how to overcome the challenges in SW languages, ontologies, and Internet scenarios. Students enter the legal learning and research process faster and more effectively when introduced into legal and technological content simultaneously.

V. Case Studies In Legal Education

According to Vladan Devedžić, web-based education offers the following features: (i) the separation of teachers and learners; (ii) the influence of an educational organization; (iii) the use of Web-technologies to use or distribute some educational content; (iv) the provision of two-way communication via the Internet.[74]

> Educational material is still highly unstructured, heterogeneous, and distributed as everything else on the Web, and current learning and lecturing tools offer limited support for accessing and processing such material. The main burden of organizing and linking the learning contents on the Web, as well as extracting and interpreting them, is on the human user.[75]

71 Vid. in the USA the work carried out by Kevin Ashley and Vincent Aleven for more than twenty years (with CATO). Cf V Aleven 'Using Background Knowledge in Case-Based Legal Reasoning: A Computational Model and an Intelligent Learning Environment' (2003) 150 *Artificial Intelligence* 183–237; Niels Pinkwart et al, 'Toward Legal Argument Instruction with Graph Grammars and Collaborative Filtering Techniques' in Mitsuru Ikeda, Kevin Ashley and Tak-Wai Chan (eds) *Intelligent Tutoring Systems: Proceedings of the 8th International Conference* ITS 2006, Jhongli, Taiwan, 26-30 June, 227–236. See also the interesting discussion between Ashley and Cash Sunstein in Eric Engle, 'Smoke and Mirrors or Science? Teaching Law with Computers—A Reply to Cass Sunstein on Artificial Intelligence and Legal Science' (2002-03) 9 *Richmond Journal of Law and Technology* 9. In Europe, cf A Muntjewerff and J Groothuismink, 'PROSA. A Computer Program as Instructional Environment for Supporting the Learning of Legal Case Solving' in J C Hage et al (eds), *Legal Knowledge and Information Systems*, Jurix 98, GNI, Nijmegen 85-100; A J Muntjewerff et al, 'Case Analysis and Storage Environment CASE' (2003), in Danièle Bourcier (ed), *Legal Knowledge and Information Systems JURIX* (IOS Press, 2003) 1-10.

72 See papers and materials presented at the recent LIIs Conference on Free Access to Law, Cornell University, NY, 7-9 October 2012, <http://blog.law.cornell.edu/lvi2012/>.

73 Cf Feona Sayles and Ina Te Wiata, 'Mixing it up: Experiences with the Combined Use of Technology and Other Methods to Enhance Learning', (2011) 4 (1/2) *Journal of the Australasian Law Teachers Association* 65.

74 Vladan Devedžić, *Semantic Web and Education* (Springer, 2006) 1.

75 Ibid. 27

While there have been some significant advances in technological support for e-learning and Web-based education,[76] legal data requires careful representation into Web languages and needs to be coupled with metadata. Ongoing projects face the specific problem of building up legal XML standards and best practices for the use of XML in legislative, regulatory, and judiciary documents. This is a matter for future research[77], with increased educational projects on Law and the Semantic Web.[78] Following the distinction made by the W3C between 'case studies' and 'use cases',[79] SW Core-semantics Educational Projects are distinguished from SW Educational Applications. There are 32 accepted case studies and 12 use cases (prototypes). Only one full-fledged case study in the education and learning technology area exists to date.

A. *SW Core-Semantics Educational Projects (CED)*

A CED may be defined as a structural project covering all the four features of Web-based education, and embedding semantic tools (RDF, ontologies e.g.) into architecture to offer semantically-based web services. In response to the unstructured material problem pointed out by Devedžić, in 2007 the Talis group and the University of Plymouth started a project to be implemented into the internal architecture of services of UK and Eire Universities. They focused on collections of text books, journal articles, Web pages and/or audio visual content defined by instructors, intended to be companions for students to degree courses, modules or assignments known as Resource Lists (RL). A RL ontology was created to unify the descriptions of existing resources[80], linked open data principles improved the interoperability of the data, and students and instructors were encouraged to annotate and enrich the data to enable context-aware recommendation functionality.[81]

Clarke and Greig define this CED as follows:

> Once obtained, the metadata is stored in the instructor's library as RDF using the Bibliographic Ontology increasing the interoperability of the harvested data with other systems and workflows. [...]
>
> By storing metadata about the resource being described, rather than the page describing it, more resilient strategies can be employed to ensure content links do not break if the library decides in

76 A broad comprehensive survey can be found in Information Resources Management Association, (IRMA) *Web-based Education: Concepts, Methodologies, Tools, and Applications* (IGI Global, 2010) 3 vols.

77 See the ongoing development of CEN-MetaLex and AkomaNtoso, the two main trends in creating such standards, <http://doc.metalex.eu/>, <http://www.akomantoso.org/> and <https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legaldocml>).

78 There are several trends to teach and disseminate *legislative XML*, a markup language defining a set of rules for encoding legal documents in a format that is both human-readable and machine-readable. Cf the Handbook by G Sartor et al, *Legal XML for the Semantic Web* (2011, Springer). Since 2007 onwards, Prof. Monica Palmirani, from CIRSFID (Bologna), organises the Legislative XML Summer Schools (LEX) with the support of OASIS organization (USA), the European University Institute (Florence), the Leibniz Center for Law (Amsterdam), the UAB Institute of Law and Technology (Barcelona), among other Institutes for Legal Informatics. See <http://summerschoollex.cirsfid.unibo.it>, and <http://www.legalxml.org/governance/>.

79 See <http://www.w3.org/2001/sw/sweo/public/UseCases/>: 'Case studies include descriptions of systems that have been deployed within an organization, and are now being used within a production environment. Use cases include examples where an organization has built a prototype system, but it is not currently being used by business functions.'

80 Cf <http://vocab.org/resourcelist/schema-20080519.html> .

81 Cf Chris Clarke and Fiona Greig, 'Case Study: A Linked Open Data Resource List Management Tool for Undergraduate Students', June 2009, <http://www.w3.org/2001/sw/sweo/public/UseCases/Talis/>.

> future to change supplier. […]. The philosophy of this feature is to get the user as close to the resource as possible.[82]

This Resource List Management (RLM) tool is the origin of Talis Aspire, a cloud-based reading-list application used by most UK (and one Australian) universities.[83] The philosophy of "getting the user as close to the resource as possible" is complemented by a clear *crowdsourced* orientation, because the tasks of re-using, sharing and transforming content are faced from two standpoints facilitating both personal as well collective learning.[84]

B. *SW Educational Applications (EA)*

Finally, three ongoing trends relevant to educational purposes and at the use case stage will be discussed. This is not the first time that IDT SW researchers and lawyers have collaborated on educational projects. From 2004 to 2006, and under the framework of the EU SEKT Project[85] an i-FAQ system was designed for the Spanish Judiciary School to support newly appointed judges to solve practical issues not covered by the law.[86] In the same judicial area, we developed multimedia tools to annotate and manage legal videos.[87] These are first generation SW projects in which semantics and ontologies played a major role attempting to improve end users' performances. However these initiatives did not contribute to the updating of the system or to the interplay of the community.

On the contrary, crowdsourcing is being applied in the ongoing EU Project *Justmen. Menu for Justice – Toward a European Curriculum Studiorum on Judicial Studies*.[88] One of the most interesting outcomes of the Project is the crowdsourced map of Law Schools and legal educational units in Europe, with aggregated information (reports) from several different sources [Fig. 2].[89]

---

1. 82 Ibid. Cf also Chris Clarke, 'A Resource List Management Tool for Undergraduate Students Based on Linked Open Data Principles' (2009), Luis Aroyo et al, (eds) (ESWC 2009) 6-7.

83 Cf <http://campus.talisaspire.com/>.

84 Nadeem Shabir and Chris Clarke, 'Using Linked Data as a basis for a Learning Resource Recommendation System', *1st International Workshop on Semantic Web Applications for Learning and Teaching Support in Higher Education* (SemHE'09), ECTEL'09, Nice, France.

85 See <http://www.sekt-project.com/>.

1. 86 See P Casanovas, 'Use Case: Helping New Judges Answer Complex Legal Questions', May 2007,

<http://www.w3.org/2001/sw/sweo/public/UseCases/Judges/>; P Casanovas et al, 'Supporting newly-appointed judges: a legal knowledge management case study' (2005) 9 (5) *Journal of Knowledge Management* (Special issue on SEKT Technologies) 7-27. See Casellas, above n 31.

87 P Casanovas et al, 'The e-sentencias Prototype: a Procedural Ontology for Legal Multimedia Applications in the Spanish Civil Courts' (2009), in Breuker et al, above n 32, 199-219.

88 See <https://www.academic-projects.eu/menuforjustice/default.aspx>.

89 See <https://legaleducationineurope.crowdmap.com/>.

Fig. 2. Screenshot from the EU education crowdsourced map. Source: https://legaleducationineurope.crowdmap.com/ (mainly managed by Sílvia Gabarró)

Another step in the same direction is the Stanford Project on constitutional tagging.[90] Its main goal is to construct the *Constitution Explorer,* a structured database of Constitutions to enable people to compare and contrast Constitutions from other countries. Rather than going manually through the text, semantic searches are enabled by a legal taxonomy. This is an object of collaborative design through mindmaps. Volunteers from different countries review and complete the taxonomy, tagging and annotating national Constitutions with relevant cases. Students from the UAB Law School (alongside students from Stanford and the University of Edinburgh) performed this task on November 12th 2011.[91] This is an example of creative academic endeavour in which more than one-hundred students communicated in different languages to reflect on the conceptual difficulties of comparative constitutional law. We discovered that discussions in cross-cultural and political tagging are analogous to expert discussions in ontology building. In both cases, cross-fertilization and contrasting opinions helped to create an added-value in learning from the existing materials.

In the final case study the work of the legal scholar Núria Casellas, is cited. Casellas and graduate students from the Department of Computer Sciences are responsible for the Linked Legal Data Project (Project) at Cornell Legal Information Institute. Although the Project is still in its early stages it represents a rigorous attempt to bring LOD to law and law libraries:

> With this project, we will enhance access to the Code of Federal Regulations (a text with 96.5 million words in total; ~823MB XML file size) with an RDF dataset created with a number of semantic-search and retrieval applications and information extraction techniques based on the development and the reuse of RDF product taxonomies, the application of semantic matching algorithms between these materials and the CFR

90 Program on Liberation Technology at Stanford's Center on Democracy, Development, and the Rule of Law, Stanford University, <http://liberationtechnology.stanford.edu/> (Program Managers: Vivek Srinivasan and Sarina A Beges).

91 Antoni Roig, Marta Poblet, and Meritxell Fernández-Barrera coordinated the IDT-UAB contribution to the Stanford Project. See Marta Poblet, 'Tagging Constitutions Online' (2011), <http://serendipolis.com/2011/11/29/tagging-constitutions-online-constitution-day-in-barcelona/>

content (Syntactic and Semantic Mapping), the detection of product-related terms and relations (Vocabulary Extraction), obligations and product definitions (Definition and Obligations Extraction).[92]

One of the goals of this Project is to create an RDF dataset of the Code of Federal Regulations (CFR) —structure, vocabulary, definitions, obligations— and link content to other collections of data (DrugBank, DBpedia e.g.). One of the expected outcomes is the development of a standardized SKOS vocabulary for the CFR.[93]

These three projects: Justmen, Constitution Explorer, Legal Linked Data— are substantially different. Only the last one implies the construction of a full SW application. The tasks performed by students are of different levels of complexity. The common element is the cooperative work of students, researchers and teachers alike from the legal field. These are the products of collective, *crowdsourced* knowledge, at the crossroads of shared understandings of technological and legal knowledge. These innovations spring from the ongoing dynamics of individual experiences, legal expertise, and technological skills.

VI. CONCLUSIONS

To a large extent regulatory systems rely on legal systems. By the same token, we need legal knowledge in order to structure and model legal data. We now have many ways at our disposal to aggregate, organise, re-use and improve knowledge which the synergy between SW technologies and educational experiences can provide. Research is being enhanced by crowdsourcing contributions. Since 2006, we have been witnessing the blossoming of a second Semantic Web generation of projects.[94] While some of these projects are still in their infancy and caution is due, it is my contention that legal initiatives, innovative trends and regulatory requirements of Web 2.0 and Web 3.0 are creating the future for Web-based learning. Linked Open Data (LOD), Linked Government Data (LGD) and Linked Legal Data (LLD) have now entered the field and are here to stay.

New concepts such as *crowdsourcing*, *crowdservicing*, *networked governance*, *data protection governance*, *hybrid open access publishing*, *semantic web services*, *legal mashups*, *dereferenced law*, *legal XML*, have recently emerged. New principles and standards are being formulated to stabilise languages and protocols, and to harmonise them with legal systems known as *global law* (including actors other than national states and international official organizations —NGOS, companies, global institutions, digital neighbourhoods etc.).

The concept of *relational law* in this paper refers broadly to the regulatory link between Web 2.0 and 3.0, based on trust and dialogue, and which emerges from the intertwining of top-down existing legal systems and bottom-up participation (the Web of People). Relational law constitutes a new challenge in the development of democracy, providing further opportunities for the educational skills and programs upon which democracy is based.

In this paper, a distinction has been made between SW Core-semantics Educational Projects and SW Educational Applications. When web science and research tasks are integrated into the legal educational curricula students are provided with the opportunity to actively participate, as

92 Núria Casellas, Sarah Bouwman, Dallas Dias, Jie Lin, Sharvari Marathe, Krithi Rai, Ankit Singh, Debraj Sinha, Sanjna Venkataraman, 'Linked Legal Data. Improving Access to Regulatory Information', Legal Information Institute, Cornell Law School, March 2012.

93 See Núria Casellas, 'Linked Legal Data: A SKOS Vocabulary for the Code of Federal Regulations' (forthcoming), *Semantic Web Journal* (2013 Special Issue on Semantic Web for the legal domain: from text to knowledge) <http://www.semantic-web-journal.net/>. To date it has been explored in particular Title 21, Food and Drugs of the CFR. 'SKOS' stands for 'Simple Knowledge Organization System', see above n 12.

94 Cfr. Mathieu D'Aquin et al, 'Toward a New Generation of Semantic Web Applications', (2008) 23 (3) *IEEE Intelligent Systems* 20-28.

shown in SW Educational experiences[95]. As Jason Ohler has said: "15 years ago, the Web was science fiction to most. Today it is taken for granted. Eventually, we will take Semantic Web for granted as well".[96]

If young lawyers are expected to operate in this technological environment, then web science needs to have a greater presence in course curriculum and design in legal education. In Europe, we have tried to give a preliminary answer to this need at the doctorate level, combining law, science and technology courses and framing this learning into research programs. [97] It seems reasonable to expect new developments of the synergy between Semantic Web and legal education experiences in the next future.

95 See above n 48, 78, and s 4 A, s 4 B, and s 5 B.

96 See Ohler, above n 2, 9.

97 See the recent Erasmus Mundus European Joint Doctorate on Law, Science and Technology [University of Bologna, University of Torino (Italy), IDT-Autonomous University of Barcelona (Spain), Mykolas Romeris University (Lithuania), University of Luxembourg, University of Tilburg (The Nederlands); with Associated partners (NICTA, IIIA-CSIC); and Industrial Partners (IBM, S21, Vicomtech, ASCAMM, NOMOTIKA) <http://www.last-jd.eu/>.