

# 3 Empirically-Grounded Development of Legal Ontologies: a Socio-Legal Perspective

Pompeu Casanovas, Núria Casellas and Joan-Josep Vallbé

Institute of Law and Technology, Faculty of Law, Universitat Autònoma de Barcelona 08193 (Barcelona), Spain (pompeu.casanovas, nuria.casellas, pep.vallbe)@uab.es

Pre-publication version. Please quote: G. Sartor, P. Casanovas, M.A. Biasotti, M. Fernández-Barrera (Eds.) *Approaches to Legal Ontologies. Theories, Domains, Methodologies*. Law, Governance and Technology Series, Vol. 1. Springer Verlag, Dordrecht, Heidelberg, London, New York, 2011, pp. 49-68

ISBN 978-94-007-0119-9, e-ISBN 978-94-007-0120-5, DOI10.1007/978-94-007-0120-5

**Abstract** This paper shows the multiple relationships between empirical data and semantic content in the legal field. One of the well-known problems of ontology construction is the “knowledge acquisition bottleneck problem” pointed out by Edward Feigenbaum and others, many years ago. In the next generation of Semantic Web developments this problem has not been completely solved. It is our contention that an accurate description of the legal environment, and well-grounded previous sociological studies may help to face it in a more satisfactory way. This means adopting a user-centered approach for legal ontologies, in what we will call an “iterative and integrated pragmatic circle” involving legal theorists, socio-legal researchers, professional people (lawyers, magistrates, prosecutors...) and computer scientists. We put the example of how the ontology of IURISERVICE was built up.

## 3.1 Introduction: wrestling with the angel

Dealing with data is far from easy. The definition of objects, measures and research processes have to be added to the problem of knowledge representation. There is at present a great deal of meth-

odologies for ontology building. CommonKADS, DOGMA, KACTUS, METHONTOLOGY, TOVE or DILIGENT, are well-known approaches to ontology engineering. All of them use to distinguish analytically types of activities, and phases or successive stages in the building process. METHONTOLOGY, e.g., includes the identification of the ontology development process, a lifecycle based on evolving prototypes, and techniques to carry out each activity in the management, development-oriented, and support activities (Gómez-Pérez et al., 2003: 125). DILIGENT separates as well these three types of distinct activities: ontology management, ontology development activities, an ontology support activities (Sure et al. 2006: 173).

Those are useful guidelines. However, as we have already stated (Casanovas et al. 2007), if ontology building is to be considered not only as an engineering process, but as a process of knowledge integrated in a wider field of research, sometimes these successive stages tend to be blurred into the dynamics of field research and the knowledge acquisition process.

We think that this is a difficult problem, because building an ontology does not mean that the process of knowledge acquisition can be taken for granted or finished once and for all. Therefore, lifecycles follow other patterns which are not predicted in advance. At least, this is our experience in doing social science research on legal behavior and legal knowledge. Therefore, grounding empirically the legal knowledge to be modeled through ontologies is an evolving and ongoing process.

This is not new. Experts systems faced a similar problem regarding the representation of common knowledge. Ed Feigenbaum called it the “knowledge acquisition bottleneck problem” (1977), and later on, “wrestling with the angel” (1992). Reflecting on their own experience, Forsyth and Buchanan (1989) suggested going beyond the traditional engineering methods and following the eliciting techniques of ethnomethodology and qualitative sociology.

More recently, ethnography has been seriously considered as a starting point. The assumption is that shared and implicit knowledge can be reassumed in *patterns of cooperative interaction* elicited by researchers to model and design computing e-government tools (Martin et al., 2002). These patterns consist of examples of similar

social phenomena that serve as resources for defining, generalizing and reusing design concepts (e.g. “telling a story about the workplace”, and they complement findings of field research). They may be defined as “regularities in the organisation of work, activity, and interaction among participants, and with, through, and around artefacts” (Martin and Sommerville, 2004: 59).

In the Pragma-dialectics field, a similar ethnographic trend has been taken by Jackson (1989) and Jacobs (2002) to produce detailed accounts on interactions in mediation processes. They reconstruct the *collaborative design of the disagreement space*. “We do not take the central problem of pragmatics to be how communicators assign functional meaning to specific messages or disambiguate speaker intention, but how is that people mutually negotiate social activity with language and thus participate in everyday life” (Aakhus and Aldrich, 2002).

Common language, tacit and implicit knowledge, and the content of shared experiences and routines in everyday life seem to be the problems to be tackled before modeling. In this chapter, we will summarize the lessons learned in the construction of the ontology of IURISERVICE, an i-FAQ for judges in their first appointment. We could have taken other examples as well, stemming from other projects (NEURONA, E-SENTENCIAS or ONTOMEDIA ontologies). But, for the sake of synthesis, we will concentrate only on a single one, just to widen up the field in the last section.

In Section 2 we will define the socio-legal approach, including the conceptual problems of the sources and what are the boundaries of the legal domain. In Section 3, we will introduce the construction process of IURISERVICE. We will conclude redefining the social-legal approach in section 4.

### **3.2 The Socio-Legal Approach: pluralism and legal culture**

Legal scholars use to identify many sources of law. Discourses, rules, acts, events or documents are considered to be *legal*. Law and

non-law are split up according to several criteria of identification or *validity*. Legal material, then, is classified according to the validity of the source, and according to three main positions –monism (normative positivism), dualism (facts/norms), and pluralism (plurality of legal sources).

Typically, socio-legal scholars embrace a pluralist perspective and, following the first legal realism, they do not refer to a validity criterion or a validity rule to describe norms or rules as social artifacts.<sup>1</sup> They use to refer to law as a “political arena”, “legal field”, “political domain”, “power construction” or “professional field”.<sup>2</sup> This is a heritage from legal realism as well: what lawyers do (or what legal academics do) —as Trubek, Dézalay, Buchanan or Garth put it— reflects “what judges do” or “what officers do”, the old motos by Holmes and Llewellyn.

From this behavioral perspective, there are actually many *legal pluralisms* based on multiple regulatory forms referring to a plurality of aspects, according to different authors. Very broadly: (i) negotiations on rules, norms and rights; (ii) mechanisms stemming from different legal orders (especially in post-colonial states); (iii) selection of different jurisprudential places and procedural subjects (*forum shopping*); (iv) selection of different legal systems or norms (*répertoire normatif du juge*); (v) non-legal sources vs. legal sources (grounded on the state); (vi) increasing complexity of transnational orders and structures; (vii) semi-autonomous social fields (social ordered organisms or institutions) in industrial societies; (viii) differential regulations depending on class, race or gender; (ix) dialogue among cultures to grasp the different symbolic feature of regulations; (x) self-organizational or self-regulatory social or normative systems or sub-systems; (xi) implicit cultural models producing some social or institutional; (xii) folk models of law, institutions and rights.<sup>3</sup>

---

<sup>1</sup> See Friedman et al. (1995), and Abel (1995) for consistent readings on the field.

<sup>2</sup> The legal field is defined, e.g., as “the ensemble of institutions and practices through which law is produced, interpreted, and incorporated into social decision-making. Thus, the field includes legal professionals, judges, and the legal academy.” (Trubek et al. 1994: 411)

<sup>3</sup> See for a summary of different kinds of pluralism, Casanovas (2002).

All these objects and aspects lead to different social approaches and methodologies too, opposed to the normative, deontic or “formalistic” methods of legal philosophy. During the 20<sup>th</sup> century, this relationship has been sometimes difficult.<sup>4</sup> Law is not conceived as “a mirror of society that functions to maintain legal order” (Tamanaha, 2001: ix).

However, monism and pluralism face similar problems from the technological point of view. Legal knowledge representation, legal knowledge acquisition, and modelization have to be faced independently of which conception of law has been chosen.

Measures are another common problem. Ontology construction requires some kind of control of the different versions and steps of the process, and decisions have to be made at each step. There are no “neutral” ontologies (Bench-Capon, 2003). This means that, even for pluralists, some conception of law and regulation is *always* implemented. And, in the other way round, perhaps legal theory scholars have a natural trend to take ontology exclusively in a philosophical sense,<sup>5</sup> without taking into account that the whole lifecycle comprehends walkthroughs and tests of prototypes with end users.<sup>6</sup>

Besides, from an empirical point of view, there is a preliminary problem before any modelization takes place. What does it mean “legal data”? How the construction of legal data is solved?

### 3.2.1 *Legal culture*

Socio-legal scholars, coming from sociology or anthropology do not need, such a previous qualification. Meaning is ot

- the discourse of the legislator (laws and regulations);

---

<sup>4</sup> See Abel (1995:1): “When asked what I study, I usually respond gnomically: everything about the law, except the rules.”

<sup>5</sup> “What does exist and what can exist? What is the essence of things, and what the conditions of their existence?” (McCormick, 1991: vii).

<sup>6</sup> See the conception of “top ontology” developed by Jaap Hage (2002).

- the discourse of the judges (judgements and other judicial decisions);
- the discourse of the doctrine (studies on several legal subdomains, systematising legislator and judges' discourses);
- the discourse of legal theory (legal works having a general content, not addressing a particular legal system).

### **3.3 An Ontology-Enhanced Decision Support System for Judges: JURISERVICE**

In this section we will briefly describe the methodological approach for the design of the JURISERVICE system, and the process of construction of the Ontology of Professional Judicial Knowledge (OPKK), a legal professional ontology grounded on empirical data and expert knowledge. The JURISERVICE application was designed to provide Spanish judges in their first appointment with on-line access to an FAQ (Frequently Asked Questions) system, which contains a repository of practical questions (problems) with their corresponding answers. The aim of the system was to discover the best semantic match between the user's question [input question] (formulated in natural language) and a stored question, so as to offer an answer that satisfied the user. Time and accuracy were critical issues and, to that end, the main research had been based on the possibility of modelling the legal professional knowledge contained in the repository of questions through the use of ontologies. In order to find the question-answer pair that best matches the input question, the system was enhanced with an ontology (the Ontology of Professional Judicial Knowledge) and semantic distance calculation.

It is worthwhile to emphasize here that technical work was embedded since 2001 into a larger research on judges and the judiciary in Spain. This research started with the Observatory of Judicial Culture, and was followed up in the subsequent years by several national and international projects.<sup>7</sup> In all of them, the pragmatic focus was

---

<sup>7</sup> See the acknowledgments section.

the structuring of practices and procedures of the judicial culture: how judicial professional knowledge is produced, shared and organized through formal and informal means (Poblet and Casanovas, 2005). In this approach, the implication of judges and magistrates themselves turned out to be crucial.<sup>8</sup> But, as we will see, judges are not researchers. Thus, relying on a user-centered approach has undoubtedly positive effects on technology construction, but it has to be mentioned that a higher level of uncertainty on the final result has to be assumed as well.

### ***3.3.1 Empirical-Based Design and Knowledge Acquisition***

The need for the JURISERVICE system and its initial design was established as a result of a thorough ethnographic survey carried out with the collaboration of the Spanish General Council of the Judiciary [*Consejo General del Poder Judicial*, CGPJ]. The Spanish Young Judges [*Jueces Jóvenes en España*] survey<sup>9</sup> was conducted during 2002, it involved five Spanish Universities,<sup>10</sup> and its main objective was to gather information towards the implementation of a support network for judges.

In-depth interviews were made by judges still in training to a set of 129 judges with less than 4 years of experience, out of the total set of 352 young judges who had completed their studies at the

---

<sup>8</sup> “The practical implication of cognitive apprenticeship is to refocus instructional research on the design process itself: We should design computer systems in partnership with students, teachers, and practitioners in the context of use, so we can produce programs that people can afford and want to use, that promote creativity, and that relate in an honest, pragmatic way to everyday life” (Clancey, 1992: 139).

<sup>9</sup> Detailed information regarding this survey can be found at Ayuso et al. (2003); Álvarez et al. (2005). Also Casanovas et al. (2004) includes some references to the data.

<sup>10</sup> Universitat Autònoma de Barcelona (UAB), Universitat de Barcelona (UB), Universitat Politècnica de Catalunya (UPC), Universidad de León, and Universidad de Burgos.

Spanish School of the Judiciary between 1997-1999 and occupied at the time their first appointment. The questionnaire was designed by senior judges of the Spanish School of the Judiciary, experts of the Documentation Office and team members of the research group. To perform a comparative analysis a set of 139 senior judges was selected.

The questionnaire contained both open-ended and closed questions tackling several areas. Table 1 below includes some of the questions included in the questionnaire, adapted from Ayuso et al. (2003) and Álvarez et al. (2005).

**Table 1** Questionnaire

Domains	Number of questions	Examples
Training evaluation	18 closed questions and 3 open-ended questions	3 What is your opinion about the education received at the Law School? What is your opinion regarding the current system of access to the profession? What changes do you suggest in the training at the Spanish School of the Judiciary? Have you used the on-line continuous training system of the CGPJ?
Professional activity	13 closed questions and 16 open-ended questions	4 What was the most complex professional problem that you had during the first 3 months of appointment? Do you comment your cases if complex with other peers? Do you use Internet?
CENDOJ services	5 closed questions	Do you use the personal attention service of CENDOJ? Do you use legal databases?
Relationships	26 closed questions and 4 open-ended questions	4 Do you think that people are right when they say that "Justice is very slow"? Do you keep professional relationships with judges from your own class?
Comments on the profession	7 closed questions and 3 open-ended questions	3 What do you think it is a "good judge"? Why did you become a judge? Do you take your work home?

Some of the results provided some insight towards which problems could a system for judges in their first appointment offer support. The analysis of the open-ended question "Explain the two most important doubts that you had during the first three months as a judge", reported that their questions referred to mostly to the on-call



period [*guardia*] (Benjamins et al., 2005). During that period, usually a week, the judge must be available 24 hours a day for any case that reaches the judicial office.

Thus, this on-call period doubts seem to refer mostly to practical situations regarding, for example, who is to keep the belongings of a detainee or a corpse? Or what is to be done when the prosecutor or the coroner does not attend an appointment? These on-call problems did not appear in the responses of the interviewed senior judges, suggesting that this type of knowledge was probably acquired with the day-to-day practice of the profession. Accordingly, the theoretical training that applicants endure to access judicial appointments does not contain this more 'practical' knowledge; neither does the training at the Spanish School of the Judiciary. This is also consistent with the findings of the survey regarding the changes proposed to the training received in the Spanish Judicial School at Barcelona: judges in their first appointment suggested educational changes towards offering a more practical teaching approach rather than the focus on theoretical study (Ayuso et al., 2003).

This, together with the fact that most judges in their first appointment declared to comment with peers (especially more experienced peers) their cases frequently (11.8%) or sometimes (72.7%), was thought to provide a ground for such a support system. Problems regarding on-call situations at late hours were difficult to consult or comment with others and access to a Frequently-Asked Questions (FAQ) repository containing this type of material could be of use, especially during the first months of appointment.

From the experience of this previous Spanish Young Judges survey, a further questionnaire and ethnographic campaign were designed, and performed during 2004. 124 newly appointed judges around Spain conformed the sample (from a total of 248 judges of the 52th Class), and the semi-structured interviews were entirely carried out by the research team of the Institute of Law and Technology at the Universitat Autònoma de Barcelona (UAB). The new questionnaire was also organized in 5 sections, concerning professional training, professional activity, professional relationships, quality of life, and personal data. This time, the questionnaire contained some of the 2002 questions, together with questions directed towards gathering information on the requirements that a would-be system

ought to have. Information regarding complex cases in civil or criminal law was included again, together with the inquiry regarding their comments with peers about the cases, and the use of Internet. However, new questions such as, “Could you explain specific doubts or problems that came up during the on-call period?”, “What kind of professional information do you usually look for in the Web? ”, and “What would you expect from a web service/software that would provide professional assistance to judges? ” were added.

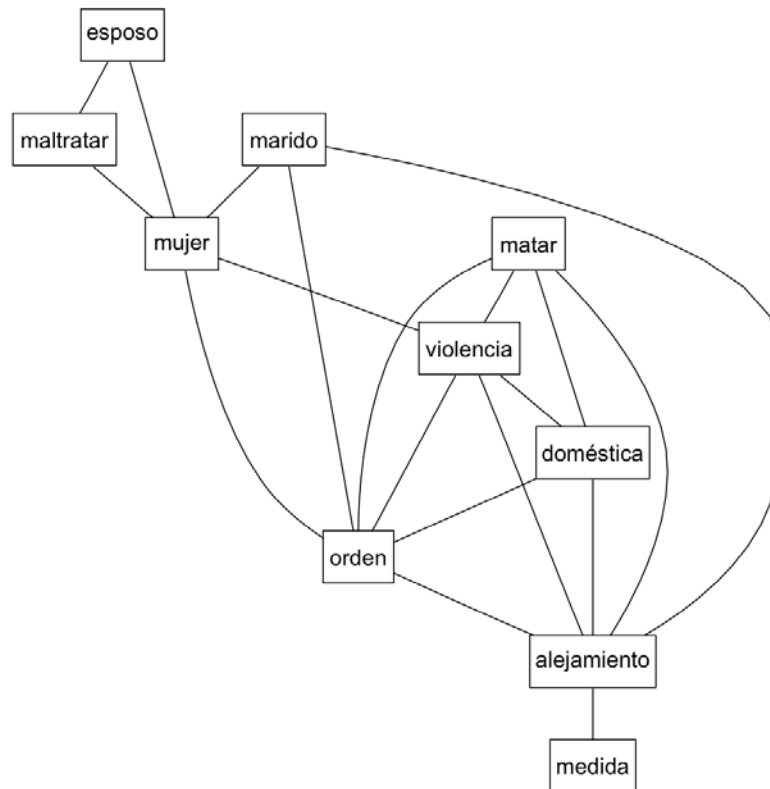
According to the findings presented in Vallbé (2009) (initially analysed in Casanovas et al. (2004)), the use of the Internet was widespread among the interviewed judges, and the information search was mostly job-related, being the official websites (official legislative publications, the judicial power website, etc.) the sites most visited. Computer skills were generally at user level, and mainly regarding the use of text processors, databases and e-mail. The majority of interviewees also desired a system that could solve “doubts”, although “corporate information”, “judges’ forum” and “doctrine” were also considered issues that a system could also offer.<sup>11</sup> Finally, in this survey, the interviewed judges were also asked to provide a list of problems (in the form of questions) that they had faced during their first appointment.

Moreover, different text-based analyses on junior judges responses were carried out in order to verify the practical nature of these on-call problems and the domain for legal professional knowledge acquisition. In effect, through the combination of simple term-frequency lists, text-mining techniques, and text multivariate statistics—e.g., correspondence analysis—on judges’ responses, it was concluded that the relevant terms used by judges pointed to actual references in the judicial world, in the sense that these terms did not point to abstract instances such as *justice*, *good* or *evil*, but to instances that have a representation in the world, whether they be particular actors (e.g., forensic doctor) or actions (e.g., arresting a person).<sup>12</sup>

---

<sup>11</sup> The most up-to-date analysis of the data is contained in Vallbé (2009), although more information regarding the data and the results may be found in Casanovas et al. (2004); Casanovas et al. (2005).

<sup>12</sup> Correlations among terms are based on similarity measures between objects within a dissimilarity matrix (Feinerer, 2008). The search for correlations is carried out in the vector



**Fig. 1** Representation of the terms that describe a typical problem during on-call services (with a  $r=.4$  correlation threshold).

Figure 1 maps the correlation of a cluster of terms referring to a typical problem arisen during on-call services, namely conflicts regarding domestic violence. In it we can firstly observe the central role played by the terms *violencia* [violence], *orden* [order, injunction], and *mujer* which may be interpreted as representing three different dimensions of those situations. These analyses led us to the

---

space computing the cosine between vectors interpreted as the *normalized correlation coefficient* (Manning and Schütze, 1999)—with values between 0 and 1.

conclusion that problems junior judges do face when on call may be regarded as being of a behavioral, practical nature.

Therefore, the conceptualization process of the Ontology of Professional Judicial Knowledge was based on this conclusion and on this previous and careful knowledge acquisition stage, which comprehended the acquisition of the list of questions and the treatment of this corpus in order to obtain the relevant terminology related to practical problems faced by judges in their first appointment, through term extraction and ontology learning from the corpus of questions—problems—faced by judges. The treatment of the corpus and the extraction of relevant terms was preceded by the establishment of the ontological requirements—including competency questions, other knowledge sources such as materials used by judges for their knowledge acquisition process: course syllabus, legislation, and doctrine, the study of existing legal ontologies towards reuse, etc. Detailed information regarding this process may be found at Casellas (2008). The methodological steps recommended by most ontology development methodologies (Noy and McGuinness, 2001; Sure, 2003; Gómez-Pérez et al., 2003) and knowledge acquisition techniques (Schreiber et al., 1999; Sure, 2003; Milton, 2007) were followed and accounted for: 1) preparatory phase (specification of ontology requirements), 2) development phase (knowledge acquisition—experts, documents, reuse—, conceptualization—classes, relations, properties, instances—, validation and formalization), and 3) evaluation phase.

### ***3.3.2 The Ontology of Professional Judicial Knowledge***

The lists of questions gathered from these interviews provided the input list of questions for the system and, together with the answers that senior judges of the Spanish School of the Judiciary gave to these questions, they conform the repository of the system. As the aim of the system was to discover the best semantic match between the user's question or input question (formulated in natural language) and a stored question, so as to offer an answer that satisfies the user, the ontology to be developed was to provide better search capabilities than the mere keyword search and to be designed towards se-

semantic indexing and search. This ontology, therefore, ought to represent the relevant concepts related to the problems that take place during the on-call period, the knowledge contained in the list of questions.

This main corpus for judicial professional knowledge modelling, which was acquired through semi-structured interviews, is constituted by nearly 800 practical questions formulated by the newly recruited judges. The interviews were recorded by the team of researchers, with the consent of the judges involved. Later, the recorded interviews were transcribed by the team. This corpus, the set of questions, contains the professional judicial knowledge gathered during daily practice at courts and constituted the repository of the application. The questions contain mainly problems or doubts arisen during the on-call period, although they also include other complex cases that junior judges had to face during their first year of practice. As an example,

- A doctor phoned to inform of someone who is not quite well and that would require internment (confinement). He asks for a court order on the phone. Can I grant it?
- Police is asking for a search warrant to enter a property to unblock a drain-pipe, as the owner does not let them in. Should I grant that warrant?
- What is to be done if, while on-call, a corpse removal needs to be performed and there is not forensic doctor available?

Several tools were used in order to extract information regarding subdomains of knowledge and relevant terms of those domains, included in the questions. Correlations shown in Figure 1, together with results from ALCESTE and OntoGen supplied subdomain information. Also TextToOnto, Text2Onto, AntConc, and Yoshikoder were used as tools to support term and ontology extraction.

For example, in order to gain some more insight towards the general contents of these questions, ONTOGEN was used on the corpus of questions to suggest concepts and relations, while the instances of those concepts were the questions themselves. This semi-automatic classification of the questions into different concepts (or topics) produced the following main topics: *Oficina\_Judicial* [court office], *Violencia\_Domestica* [gender violence],

Extranjeria [immigration], Proceso [process], and Familia [family]. A total of 17 classes (root, main topics and sub-topics) were semi-automatically learnt by the OntoGen tool (see, for example, Table 2). This topic ontology (Question Topic Ontology) was used to support question classification within the IURISERVICE system. See, for more details, Blázquez et al. (2005) and Casellas et al. (2007).

**Table 2** QTO topic and subtopic classification

Topic	Subtopics
Process	Competence Conflicts, Enforcement Proceeding, Quick Trial, Comision Rogatoria
Judicial Office	Public Prosecution, Hearing Video Recording
Gender Violence	Restraining Order
Immigration	Expulsion and Extradition
Family	Internment and Incapacitation, Autopsy and Corpse Removal, Minors

Regarding relevant terminology, Yoshikoder and AntConc offered the initial list of 477 terms (later extended to include more than 900 terms), which supplied the initial terminological knowledge for term grouping, conceptual modelling and ontology formalization. This initial knowledge acquisition and grouping was informally validated by legal experts.<sup>13</sup>

Once the conceptual extractions were performed and as much information as possible was acquired on the corpus of questions, we proceeded at grouping and organizing the concepts in a taxonomy, taking into account the content of the corpus of questions (practical

<sup>13</sup> With Yoshikoder, the analysis of the document containing the full set of questions obtained an initial list 1,998 terms for the lemmatized text. To gather an initial more manageable set of terms, a threshold of 5 occurrences was established, 452 terms were obtained. AntConc obtained a similar list with 455 terms. The 455 list of terms from the AntConc analysis on the lemmatized corpus was manually revised to offer a first working set of terms, including a revision on multiple terms (N+Adj, N+prep+N, and N+prep+N+Adj forms).

problems), the context of the questions (the judicial setting) together with background theoretical knowledge acquired during the training period of judges (from academic textbooks, legislation and examination and training course syllabuses), the established competency questions, and the insights provided by the analysis of several upper and core ontologies Casellas (2008).

Finally, the taxonomical structure was formalized in OWL using the Protégé ontology editor to allow future reuse or enrichment.<sup>14</sup> Two versions of OPJK, regarding their computational complexity, have been produced in order to facilitate computation capabilities and to obtain significant technical evaluation results in the future.

OPJK version 1.0 includes 74 classes, 73 `rdfs:subClassOf` relations and 912 instances, together with a total of 31 `owl:ObjectProperty` axioms (14 `owl:subPropertyOf` and 15 `owl:inverseOf`), 1 transitive and 1 functional `owl:ObjectProperty`. OPJK version 2.0 includes, as well, 1 `owl:equivalentClass` and 75 `owl:disjointWith` axioms, around 100 multiple class instantiation constructs, and, finally, 53 `owl:sameAs` axioms.<sup>15</sup>

---

<sup>14</sup> Versions 3.3.1, 3.4 (beta) and 4.0 (beta) were used.

<sup>15</sup> OPJK versions 1.0 and 2.0 have a DL expressivity of  $\mathcal{ALSHIF}^+$  and  $\mathcal{SHOIF}$ , respectively, as detected by Pellet in Protégé 4.0.

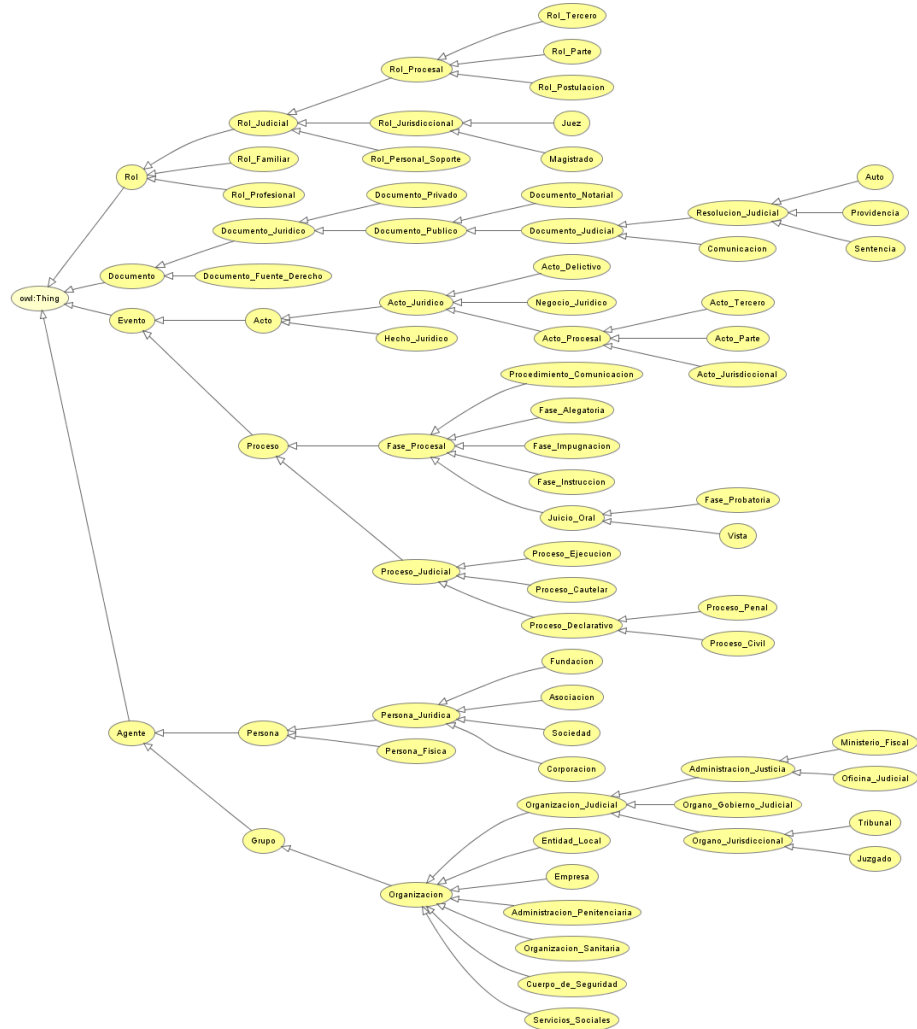


Fig. 2 OPJK v1.1 class structure in Protégé

### 3.3.3 User-Centered Approach: Expert Involvement

Evaluation activities, both during conceptualization and after formalization are central to the development of the Ontology of Profes-



sional Judicial Knowledge. This empirical-based ontology represents legal professional domain knowledge, therefore the participation and consultation of legal experts (academics and professionals), to validate the knowledge extracted, the modelling decisions taken, and the final OPJK ontology was key. There are different moments and levels of involvement of experts in the development process of the Ontology of Professional Judicial Knowledge. First, as explained above, the research team gathered the knowledge that constituted the corpus during the interviews with judges in their first appointment. Second, a first validation regarding the grouping of terms was performed with experts and, finally, the formalized ontology was evaluated by experts and refined accordingly.

Regarding validation, difficult and complex modelling decisions were discussed with the team of domain experts and ontology engineers, prior to formalization. Yoshikoder and AntConc offered the initial list of 477 terms, which supplied the initial terminological knowledge for term grouping, conceptual modelling and ontology formalization. This initial knowledge acquisition and grouping was informally validated by legal experts. Taking into account most suggestions and revising the classifications offered by the legal experts, some changes were introduced to the groups towards conceptualization and further relevant extracted terms from the initial list were classified.

Finally, the evaluation of the Ontology of Professional Judicial Knowledge by experts included the evaluation of the OPJK classes, subclass relationships, properties and instances and a more general and experimental evaluation based on a usability questionnaire. The results of these evaluations, a 72.92% and 69.44% of agreement with the ontology, respectively, suggested that, although there was general agreement, there was also some room for improvement regarding class conceptualization which could offer more granularity and foster understanding and shareability amongst experts. With the evaluation results and the expert's suggestions the Ontology of Professional Judicial Knowledge was refined, several classes were added or modified (e.g., relabelled), some instances were redistributed, properties and disjoint axioms were modified or added.

### 3.4 Final Remarks

#### Acknowledgments

The research presented in this paper has been developed within the framework of the following projects: TSI-020501-2008, 2008-2010; CSO-2008-05536-SOCI.

#### References

- Aakhus, M., Aldrich, A. (2002). Crafting Communication Activity: Understanding Felicity in ‘I wish I...’ Compliments”. *Research on Language and Social Interaction* 35, 4, pp. 395-425.
- Abel, R.L. (Ed.) (1995). *The Law & Society Reader*. New York: New York University Press.
- Abel, R.L. (1995). “What We Talk About When We Talk About Law”, in R. Abel, *The Law & Society Reader*, pp. 1-10.
- Álvarez, R., M. Ayuso, and M. Bécue (2005). Statistical study of judicial practices. In V. R. Benjamins, P. Casanovas, J. Breuker, and A. Gangemi (Eds.), *Law and the Semantic Web. Legal Ontologies, Methodologies, Legal Information retrieval, and Applications*, Volume 3369 of *Lecture Notes in Artificial Intelligence*, pp. 25–35. Springer-Verlag Berlin Heidelberg.
- Ayuso, M., M. Bécue, R. Álvarez, O. Valencia, R. Álvarez, M. L. Hernández, and M. Santolino (2003, Septiembre). Jueces jóvenes en españa (2002). análisis estadístico de las encuestas a los jueces en su primer destino (promociones 48/49 y 50). análisis comparativo con jueces de mayor experiencia. SEC-2001-2581-C02-01/02 informe interno Report n.2, Consejo General del Poder Judicial [General Council of the Judiciary].
- Benjamins, V. R., P. Casanovas, J. Contreras, J. M. López-Cobo, and L. Lemus (2005). Iuriservice: An intelligent frequently asked questions system to assist newly appointed judges. In V. R. Benjamins, P. Casanovas, J. Breuker, and A. Gangemi (Eds.), *Law and the Semantic Web. Legal Ontologies, Methodologies, Legal Information retrieval, and Applications*, Number 3369 in *Lecture Notes in Computer Science*, pp. 205–222. Springer-Verlag Berlin Heidelberg.
- Blázquez, M., R. Peña-Ortiz, J. Contreras, R. Benjamins, P. Casanovas, J.-J. Vallbé, and N. Casellas (2005, December). D10.3.1 legal case study: Prototype. Sekt ist-2003-506826 deliverable, SEKT, EU-IST Project IST-2003-506826, Intelligent Software Components S.A. (iSOCO) and Universitat Autònoma de Barcelona (UAB).
- Casanovas, P. “Dimensiones del pluralismo jurídico”, IX Congrés d’Entropologia FAAEE, Barcelona, 2002, available at

- [http://www.ub.edu/reciprocitat/GER\\_WEB\\_CAS/Actividades/Actividades%20Simposio%202002/Ponencia-Casanovas.pdf](http://www.ub.edu/reciprocitat/GER_WEB_CAS/Actividades/Actividades%20Simposio%202002/Ponencia-Casanovas.pdf) (accessed 10/5/2010)
- Casanovas, P., M. Poblet, N. Casellas, J. Contreras, R. Benjamins, and M. Blázquez (2005). Supporting newly-appointed judges: A legal knowledge management case study. *Journal of Knowledge Management* 9(5), 7–27.
- Casanovas, P., M. Poblet, N. Casellas, J. Vallbé, F. Ramos, R. Benjamins, M. Blázquez, L. Rodrigo, J. Contreras, and J. Gorroñoigoitia-Cruz (2004, December (January 2005)). D10.2.1 Legal Case Study: Legal scenario. Sekt ist-2003-506826 deliverable, SEKT, EU-IST Project IST-2003-506826, Intelligent Software Components S.A. and Universitat Autònoma de Barcelona.
- Casellas, N. (2008, December). *Modelling Legal Knowledge through Ontologies. OPJK: the Ontology of Professional Judicial Knowledge*. Ph. D. thesis, Departament de Ciència Política i Dret Públic, Facultat de Dret, Bellaterra (Barcelona).
- Casellas, N., P. Casanovas, J.-J. Vallbé, M. Poblet, M. Blázquez, J. Contreras, J. M. López-Cobo, and V. R. Benjamins (2007). Semantic enhancement for legal information retrieval: Iuriservice performance. In *Proceedings of the Eleventh International Conference on Artificial Intelligence and Law. ICAIL 2007, June 4-8, Stanford Law School, California*, pp. 49–57. Association for Computing Machinery.
- Clancey, W.J. (1992). "Representations of knowing: In defense of cognitive apprenticeship." *Journal of Artificial Intelligence in Education* 3 (2): 139-168.
- Feigenbaum, E.A. (1977). The art of Artificial Intelligence: I. Themes and Case Studies of Knowledge Engineering. STAN-CS-77-621 Heuristic Programming Project Memo 77-25.
- Feigenbaum, E. A. (1992). "A Personal View of Experts Systems: Looking Back and Looking Ahead", Knowledge System Laboratory, Report n. 92-41 KSL, Stanford.
- Feinerer, I. (2008). *tm: Text Mining Package* (R package version 0.3-3 ed.).
- Forsyth, D.E. ; Buchanan, B. (1989). "Knowledge Acquisition for Expert Systems: Some Pitfalls and Suggestions", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, n.3, pp. 435-442.
- Friedman, L.M.; MacAulay, S.; Stookey, J.A. (Eds.) (1995). *Law and Society Reader: Readings on the Social Studies of Law*, New York: Norton and Co.
- Gómez-Pérez, A., M. Fernández-López, and O. Corcho (2003). *Ontological Engineering. With examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Advanced Information and Knowledge Processing. London: Springer-Verlag.
- Jacobs, S. (2002). Maintaining neutrality in dispute mediation: managing disagreement while managing not to disagree, *Journal of Pragmatics* 34, pp. 1403-1426.
- Jackson, S. (1998), "Disputation by Design", *Argumentation* 12, pp.183-198.

- Manning, C. D. and H. Schütze (1999). *Foundations of Statistical Natural Language Processing*. Cambridge (MA) / London, UK: The MIT Press.
- Martin, D.; Rouncefield, M.; Sommerville, I. (2002) "Applying patterns of Cooperative Interaction to Work (Re)Design: E-Government and Planning", Proceedings of CHI 2002. Minneapolis, Minnesota, Publications of the ACM.
- Martin, D.; Sommerville, I. (2004) "Patterns of Cooperative Interaction: Linking Ethnomethodology and Design", ACM Transactions on Computer-Human Interaction, Vol. 11, No. 1, March 2004, pp. 59-89.
- Milton, N. (2007). *Knowledge Acquisition in Practice. A Step-by-step Guide*. Decision Engineering. London: Springer-Verlag.
- Noy, N. F. and D. L. McGuinness (2001). Ontology development 101: A guide to creating your first ontology. Technical Report SMI-2001-0880, Stanford University School of Medicine.
- Poblet, M.; Casanovas, P. (2005). "Recruitment, Professional Evaluation and Career of Judges and Prosecutors in Spain", in G. di Federico (Ed.), *Recruitment, Professional Evaluation and Career of Judges and Prosecutors in Europe: Austria, France, Germany, Italy, The Netherlands and Spain*. IRSIG-CNR, Lo Scarabeo, Bologna, pp. 185-214.
- Schreiber, G., H. Akkermans, A. Anjewierden, R. de Hoog, N. Shadbolt, W. V. de Velde, and B. Wielinga (1999). *Knowledge Engineering and Management. The CommonKADS Methodology*. Cambridge, Massachusetts. London, England: A Bradford Book. The MIT Press.
- Sure, Y. (2003). *Methodology, Tools and Case Studies for Ontology Based Knowledge Management*. Ph. D. thesis, Fakultät für Wirtschaftswissenschaften der Universität Fridericiana zu Karlsruhe.
- Sure, Y.; Tempich, C.; Vrandečić, D. (2006). "Ontology Engineering Methodologies", in J. Davies et al. *Semantic Web Technologies. Trends and Research in Ontology-based Systems*, John Wiley, Chichester, 2006, pp. 171-190.
- Tamanaha, B.Z. (2001). *A General Jurisprudence of Law and Society*. Oxford: Oxford University Press.
- Trubek, D.M.: "Back to the future: the short, happy life of the Law and Society Movement", *Florida State University Law Review* vol. 18 n. 1/1990, p. 1-55.
- Trubek, D. M.; Dézalay, Y.; Buchanan, Ruth; Davis, John, R.: "Global Restructuring and the Law: Studies of the Internationalization of Legal Fields and the Creation of Transnational Arenas", *Case Western Reserve Law Review*. vol. 44, n. 2/1994, p. 407-498.
- Vallbé, J.-J. (2009, July). *Models of Decision-Making: Facing Uncertainty in Spanish Judicial Settings*. Ph. D. thesis, Departament de Dret Constitucional i Ciència Política. Universitat de Barcelona, Barcelona.