EU-IST Project IST-2003-506826 SEKT

SEKT: Semantically Enabled Knowledge Technologies

D 10.2.1 Legal Scenario

Case Study-Intelligent Integrated Decision Support for Legal Professionals

Pompeu Casanovas (UAB)
Marta Poblet (UAB)
Núria Casellas (UAB)
Joan-Josep Vallbé (UAB)
Francisco Ramos (UAB)
Richard Benjamins (iSOCO)
Mercedes Blázquez (iSOCO)
Luis Rodrigo (iSOCO)
Jesús Contreras (iSOCO)
Jesús Gorroñogoitia Cruz (iSOCO)

with legal translations from Elena Ferran
(Universitat Pompeu Fabra)

Keyword list: legal ontology, legal documents, jurisprudence, judge, textual analysis, professional knowledge, legal databases, iFAQ, software architecture.

WP10 Case Study: Intelligent integrated decision support for legal professionals.
Report PU
Contractual date of delivery: 31st December 2004
Actual date of delivery: 31st January 2005

Abstract

The aim of this Document is to offer an overview of the work done during the first year of the research by the UAB legal team and the iSOCO engineers. The cooperation between the two teams has produced the following results: (i) a reliable description of the Spanish and judicial contexts, stemming from 2004 surveys and field research at the court settings; (ii) a reliable description of the content of Spanish Legal databases; (iii) a preliminary textual analysis on legal discourse findings; (iv) definition and epistemological bases for Ontologies of Professional Legal Knowledge (OPLK); (v) knowledge acquisition and ontology building of *Iuriservice II* (second prototype of an iFAQ assistant for judges at their first appointment); (vi) use cases and software architecture design for the legal case study prototype, using NLP techniques, ontology models and ontology merging.

SEKT Consortium

This document is part of a research project partially funded by the IST Programme of the Commission of the European Communities as project number IST-2003-506826.

British Telecommunications plc.
Orion 5/12, Adastral Park
Ipswich IP5 3RE
UK
Tel: +44 1473 609583, Fax: +44 1473 609832
Contact person: John Davies
E-mail: john.nj.davies@bt.com

Empolis GmbH
Europaallee 10
67657 Kaiserslautern
Germany
Tel: +49 631 303 5540
Fax: +49 631 303 5507
Contact person: Ralph Traphöner
E-mail: ralph.traphoener@empolis.com

Jozef Stefan Institute
Jamova 39
1000 Ljubljana
Slovenia
Tel: +386 1 4773 778, Fax: +386 1 4251 038
Contact person: Marko Grobelnik
E-mail: marko.grobelnik@ijs.si

University of Karlsruhe, Institute AIFB
Englerstr. 28
D-76128 Karlsruhe
Germany
Tel: +49 721 608 6592
Fax: +49 721 608 6580
Contact person: York Sure
E-mail: sure@aifb.uni-karlsruhe.de

University of Sheffield
Department of Computer Science
Regent Court, 211 Portobello St.
Sheffield S1 4DP
UK
Tel: +44 114 222 1891
Fax: +44 114 222 1810
Contact person: Hamish Cunningham
E-mail: hamish@dcs.shef.ac.uk

University of Innsbruck
Institute of Computer Science
Techikerstraße 13
6020 Innsbruck
Austria
Tel: +43 512 507 6475
Fax: +43 512 507 9872
Contact person: Jos de Bruijn
E-mail: jos.de-bruijn@deri.ie

Intelligent Software Components S.A.
Pedro Valdivia, 10
28006 Madrid
Spain
Tel: +34 913 349 797
Fax: +49 34 913 349 799
Contact person: Richard Benjamins
E-mail: rbenjamins@isoco.com

Kea-pro GmbH
Tal
6464 Springen
Switzerland
Tel: +41 41 879 00
Fax: 41 41 879 00 13
Contact person: Tom Bösser
E-mail: tb@keapro.net

Ontoprise GmbH
Amalienbadstr. 36
76227 Karlsruhe
Germany
Tel: +49 721 50980912
Fax: +49 721 50980911
Contact person: Hans-Peter Schnurr
E-mail: schnurr@ontoprise.de

Sirma AI EOOD (Ltd.)
135 Tsarigradsko Shose
Sofia 1784
Bulgaria
Tel: +359 2 9768, Fax: +359 2 9768 311
Contact person: Atanas Kiryakov
E-mail: naso@sirma.bg

Vrije Universiteit Amsterdam (VUA)
Department of Computer Sciences
De Boelelaan 1081a
1081 HV Amsterdam
The Netherlands
Tel: +31 20 444 7731, Fax: +31 84 221 4294
Contact person: Frank van Harmelen
E-mail: frank.van.harmelen@cs.vu.nl

Universitat Autonoma de Barcelona
Edifici B, Campus de la UAB
08193 Bellaterra (Cerdanyola)
Barcelona
Spain
Tel: +34 93 581 22 35, Fax: +34 93 581 29 88
Contact person: Pompeu Casanovas Romeu
E-mail: pompeu.casanovas@uab.es

Executive Summary

This research report is divided into four main parts according to the following order: (i) *user corpora*: the Spanish judicial field, findings on judges' profile, and structure and content of Spanish documents; (ii) *legal ontology*: state of the art, epistemological grounds, the proposal of Ontologies of Professional Legal Knowledge (OPLK), and development of an Ontology of Professional Judicial Knowledge (OPJK); (iii) *use cases* collection: from question-answer to Jurisprudence Ontology and OPJK aligning; (iv) *use cases software design*: architecture diagram for *Iuriservice II* prototype.

The Report offers the preliminary results of the 2004 surveys and fieldwork on judicial behaviour, beliefs, attitudes and technological skills of newly appointed judges. The UAB researchers visited more than 100 Spanish court judicial units located in 14 (out of 17) Autonomous Communities (including the Canary Islands). The main purpose was obtaining *sur place* an accurate map of the problems faced by judges in their professional settings at their first appointment.

Then, we describe the work already done to build the first prototype of *Iuriservice*, centered on the on-duty time problems. The knowledge acquisition process for *Iuriservice* II focused mainly on the doubts, difficulties and problems faced by the newly appointed judges. This new approach made it possible the identification of more than 750 competency questions, in comparison to the 100 questions extracted from the answers to the first questionnaire.

Several changes have also been introduced regarding the ontology modeling approach: the use of KAON Oi-Modeler to input and visualize the domain ontology; different software applications (TextToOnto and ALCESTE) to extract relevant concepts from the competency questions; application of the Distributed, Loosely-controlled and evolving Engineering of oNTologies (DILIGENT) argumentation methodology during the ontology engineering process.

The use of TextToOnto has been limited as the implementation of the support for the Spanish language forces limitations to concept association extraction and the ontology pruning algorithms. However, the term extraction algorithm has proved to be useful to extract potential relevant concepts for the ontology.

Using ALCESTE [*Analyse des Lexèmes Co-occurents dans les Énnoncés Simples d'un Texte*, Analysis of the co-occurring lexemes within the simple statements of a text] we have produced an internal preliminary statistical map of the lexical worlds contained in three protocols of analysis: (i) old on-duty judicial questions belonging to the first *Iuriservice* prototype; (ii) new on-duty judicial questions produced as a result of the 2004 surveys and field research; (iii) new more general judicial questions on court proceedings and processes provided by the same field research.

TextToOnto and ALCESTE help identifying significant concepts, however, the method used in building the ontology has focused on the discussion within the UAB legal experts team over the terms which appear on the competency questions. To model the ontology, the middle-out strategy has been followed: basic terms are identified first and then specified and generalized if necessary. The evaluation of the arguments in favor or

against the introduction of a new concept in the ontology has been supported by the use of the argumentation model DILIGENT, based on the rhetorical structure theory.

The Ontology of Professional Judicial Knowledge (OPJK) developed by the legal case study team has been learnt from scratch out of nearly 200 competency questions and has, currently, nearly 50 concepts, 100 relations and more than 300 instances. The following top domain concepts have been identified: *acto_procesal*, *órgano_judicial*, *calificación_jurídica, documento_procesal, fase_procesal, jurisdicción, proceso_judicial, profesión_jurídica, rol_procesal, rol_familiar* and *sanción*.

The Legal Case Study Prototype has been designed taking into account two main considerations: (i) an accurate searching system, with advanced technology, that goes beyond traditional searching algorithms, capable of reliable search over a vast FAQ repository; (ii) a design that supports a fast, usable, modular, extensible, scalable, improving implementation.

The first point might be achieved by using some techniques like ontologies to model legal case domains and NLP techniques. The second point is achieved by leveraging on some software technology patterns, like a multistage searching cycle for successive approach to FAQ pair target or pluggable searching stage engines. The final design is flexible, modular, scalable, customizable and suitable for the prototype.

Contents

List of Figures

1 Introduction

The aim of SEKT is to develop and exploit semantically-based knowledge technologies in order to support document management, content management, and knowledge management in knowledge intensive workplaces. Specifically, SEKT aims at designing appropriate utilities to users in three main areas—digital libraries, the engineering industry, and the legal domain—providing them with quick access to the right pieces of information at the right time.

As regards the legal case study, the tasks accomplished so far provide both the quantitative and qualitative data necessary to assess both the context of users—newly recruited Spanish judges—and their specific needs with regard to the technology under development. In particular, these data give an insight on institutional, organizational, and individual constraints that could either facilitate or block the introduction of SEKT technologies in judicial units.

At this stage of the SEKT project, the main objective of this report is to offer an accurate description of the legal scenario that constitutes the empirical bases of this case study. To do so, we have divided the document into four parts. Chapter 2 presents detailed results from current fieldwork in court units, comparing them with data obtained in previous surveys. It also reviews ongoing work on modeling the acquisition and distribution of knowledge in the legal domain. Finally, it establishes a basic typology of legal documents and legal databases most relevant to the case. Chapter 3 covers the work on domain ontology building, including recent discussions on methodology and competency questions. Chapter 4 develops a set of standardized use cases, consisting of the description of several possible usage scenarios from the perspective of users. Those scenarios are described in detail from a functional perspective, and then, in Chapter 5 some are developed from a technological point of view. Also a global software architecture design is depicted.

2 User Corpora

2.1 Previous work and general framework

The precedents of the ethnographic work developed within SEKT are to be found in the research of the UAB Observatory of Judicial Culture (OJC).[1] Created in 2002, the OJC focuses on judicial behavior, reasoning and professional profiles of judges and magistrates. OJC's main purpose is to provide the Spanish judicial system with useful tools to improve the performance of its judiciary. The term "judicial culture" intends to describe the whole range of cognitive skills and technical devices displayed in courts in order to identify, organize and use the practical knowledge produced by judges in those judicial settings. The methodology put in place requires to gather good descriptions of judicial units, obtained through both ethnographic data regarding organizational issues (i.e. case management, management of the workflow knowledge) and statistical data concerning workload in courts.

[1] The Observatory results from a coordinate project between different research groups and universities, namely the Autonomous University of Barcelona (UAB), the University of Barcelona (UB), the Polytechnic University of Catalonia (UPC), Intelligent Software Components (iSOCO) and the board of directors of the Spanish Judicial School [*Escuela Judicial Española, Consejo General del Poder Judicial*].

In 2002-2003 the OJC conducted a national survey on newly recruited judges [ABAVS, 2003]. The main objective of the survey was to identify the most frequent problems that young judges usually face in their first appointment. The survey was prepared by the OCJ researchers, together with magistrates of the Judicial School and experts of the Center of Judicial Documentation of the General Council of the Judiciary (CENDOJ). The survey consisted of in depth interviews to 130 judges of the class 52[2] (out of 378 young judges who had completed their studies at the School between 1997 and 1999). The interviews were conducted by their own peers of the 49/50 class (that is to say, judges still at the Judicial School) who had to fill a detailed questionnaire after the interview as a planned school practice. The judges were taught how to perform the interviews so that they could also obtain information about what they could expect to encounter in their future workplaces. To make a comparative analysis with more experiences judges, 141 senior magistrates were also selected. All interviews to judges were made in their workplaces.

The interviews covered questions related to: (i) evaluation of the training at the Judicial School; (ii) evaluation of the CENDOJ documentation services; (iii) most relevant problems found by newly recruited judges in their professional activity; (iv) professional, institutional and social networks the judge belongs to and; (v) what Justice means to judges. The results allowed to identify three main areas in which young judges had problems: (i) the organization and management of daily relationships within the judicial units (with clerks, civil servants, etc.); (ii) the interpretation and implementation of some newly enacted procedural Spanish statutes (*Ley de Enjuiciamiento Civil*, January 2002); (iii) the "on-duty" period (called *guardia*: the week in which the entire Court in on duty tackling the preliminary investigations and procedures of the criminal cases that keep entering to the Courts).

In SEKT Deliverable 10.1.1 we presented these data as a general framework to analyze extensively the main features of the legal domain relevant to the case study, specially those that could critically influence the achievement of a successful technological application. We also compared those initial results with official data provided by the General Council of the Judiciary and the Judicial School.[3] To summarize, we first highlighted an important singularity with regard to potential users. New entrants to the Spanish judiciary are faced with a great variety of cases, procedures, hearings, decisions, and rulings. In a number of situations —especially those in which legal protocols regarding procedures were unclear or missing— asking a peer or a more experienced judge for professional advice is the most frequent way of obtaining information, even though this traditional way of proceeding has revealed uncertainty and may slow down the performance of both the judge asking for advice and the judge responding. In addition, the profile of judges as potential users showed that:

[2] Class 52 refers to the 52nd group of judges that graduated at the Spanish Judicial School. This group graduated in 2000.

[3] [CGPJ, BO 1999-2003] [EJ, 2001] [EJ, 2002] [EJ, 2003] [MV, 2004]

- Judges' use of the Internet for professional purposes is still low (or very focused to quick checks of the Official Journal of the State and the official page of the Higher Council).
- Judges' use of ICT and web services is still low, but they are willing to accept them, provided they facilitate decision-making and daily caseload. Web based services should be easy to learn and user-friendly for judges to use them.
- Judges' use of e-mail for professional purposes is still low, even though the General Council of the Judiciary provides an institutional account to all of them.
- Judges' use of legal databases on CD-ROM is widespread (more than 80 % use them regularly).

Despite being a most valuable source of information, further qualitative data focusing on both the organizational context of judges and their professional profiles were needed to define use cases with a greater detail and to design the prototype. Moreover, if the system had to be able to adapt to user requirements and provide them an efficient support in a fast and reliable way, the accuracy and the necessary validity of the FAQ repository was critical. For those reasons we planned an extended ethnographic work as a primary source of data regarding both the context of use and the contents of the questions to which the system would provide answers. Section 2.3 below describes the organization of the ethnographic work in judicial units developed under the SEKT framework during the period March to September 2004.

2.2 The Judicial Field

Doing fieldwork in Spanish courts necessarily requires the formal approval of the General Council of the Judiciary [*Consejo General del Poder Judicial*]. The Council is the institution that governs the judicial branch of the state at the national level. Both the Constitution and the 1985 Act of the Judicial Branch (LOPJ) assign to the General Council the following tasks:

(i) Selection, education, and continuing training of judges and magistrates
(ii) Promotion and appointment of judges and magistrates
(iii) Disciplinary control over the judiciary
(iv) Elaboration of the judicial budget (but administered by the Ministry of Justice)
(v) Management of the courts' system

The Spanish Constitution establishes the principle of unitary jurisdiction, which means that the court system has a national scope. The General Council of the Judiciary, therefore, is fully responsible for all decisions concerning the career of judges in any area of the peninsula. Nevertheless, the Constitution also establishes that the Autonomous Communities may have judicial competences transferred to the autonomous level. These competencies refer to management of the judicial system as an organization: human resources management, maintenance of buildings, facilities, case

management systems (CMS) and other office resources. In those Autonomous Communities which have no competences transferred to yet, the Spanish Ministry of Justice remains in charge. At present, there are eight Autonomous Communities fully responsible for the management of the judicial system within its territory: the Basque Country, Catalonia, Galicia, Andalusia, Navarra, the Community of Valencia, the Canary Islands, and Madrid. Except for the latter, we have visited all of them.

This multi-competency system adds some complexity to the organization of judicial units. Judges are a national body of civil servants governed by the General Council; judicial secretaries and prosecutors depend on the Ministry of Justice, and judicial staff is managed by the Autonomous Communities, which also are responsible for designing and developing CMS. The map below highlights the Autonomous Communities with competencies on justice matters and CMS used in each area.

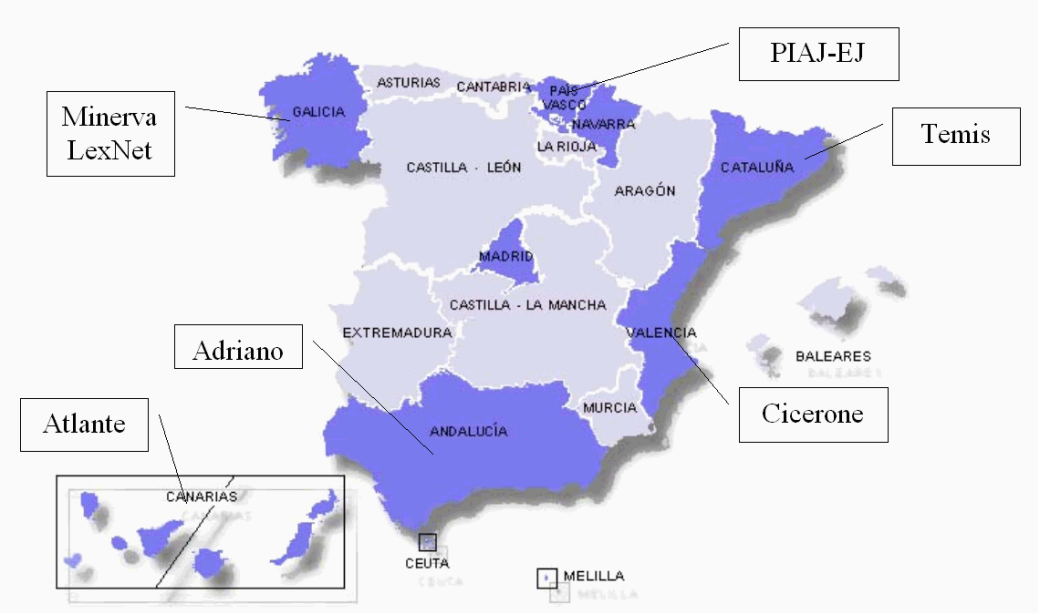


Figure 2.1: Spanish Autonomous Communities and Case Management Systems (CMS)

2.2.1 Court fieldwork

The fieldwork developed under the SEKT framework targeted the judges of the 52nd class of the Spanish Judicial School. This is a batch of 248 judges who passed the oral and written examinations in 2000, entered the School in September of that year and graduated in 2002 after having spent six months in judicial units as judges in training. They took office by early 2002. Therefore, when the UAB-GRES researchers visited and interviewed them they had already spent two years in office, were ready to move to another place and, eventually, to promote to the category of magistrate.[4] Consequently, the 52nd class fitted perfectly the two basic requirements of the ethnography: they were

[4] According to the Spanish law, 50 % of vacancies in the category of magistrate (of the civil and criminal jurisdictions) shall be covered by newly recruited judges who have already spent three years in office. Judges, therefore, are compulsory given the category of magistrate, according to their rank order. The remaining 50 % are covered by judges in competitive examinations (they do not need to spend three years) and legal professionals with at least ten years of professional experience.

newly recruited judges who, at the same time, had spent enough time in the office to provide researchers with a number of questions regarding daily problems, on-duty periods, and legal procedures at large.

Judges of the 52nd class had been sent to fill vacancies of first instance courts scattered throughout the peninsula. The Judicial School provided the official list with the destination and contact data of each of them. To create a relevant sample, we randomly selected 150 judges. The proportion of judges in each Autonomous Community was preserved in the sample, so Andalusia, Galicia and Valencia had more judges to be interviewed, mirroring the overall distribution of the class. Figure 2.2 below shows the resulting territorial distribution of the survey.

Figure 2.2: Territorial distribution of the legal case study survey

The survey covers 14 Autonomous Communities (out of 17). Before going to places we formally contacted —via telephone and official letters— the Presidents of the High Courts of each Autonomous Community —the highest judicial authority— to inform them about the characteristics of the study and check the current situation of judges (possible changes of destination, leaves of maternity, leaves of illness, etc.). We also obtained credentials from the Judicial School so that we could be accepted as members of one of the official research programs registered in the web page of the General Council.

To perform the interviewing process we organized a team of fourteen researchers. They are all trained lawyers and law professors from the UAB Law School who are used to interact with judges and master both the legal procedures and the legal language. The methodological aspects of the ethnography were discussed and set in several work sessions before the starting of the campaign (i.e. how to prepare an proceed with the interviews, how to access the judicial secretary and the staff for further information, how to record the interview or take pictures of the facilities, etc.). Besides making the interviews, every researcher had to keep a field journal where additional information

regarding the context of users or the singularities of the visited units could be preserved in a fine-grained way.

The ethnography of the SEKT legal case has been designed as a three step process involving: (i) oral interviews with judges and judicial secretaries in judicial units (whenever possible, digitally recorded with mini-disc), and informal interactions with the judicial staff working at the units; (ii) literal transcriptions of the oral interviews in a written text (that is, including the digressions or additional topics casually covered through the oral interaction with judges); (iii) literal transcriptions of questionnaires (filling the questions answered by judges and judicial secretaries to have completed questionnaires ready for statistic treatment). This is a lengthy process requiring a careful hearing of the interviews to obtain accurate transcriptions from which competency questions can be extracted. The following table reports the state of this ongoing process as of December 2004.

Team IDT – Ethnographies	Autonomous Community	Interviews	Transcriptions	Questionnaires
Casellas, Núria	València	7	7	7
Casanovas, Pompeu	Euskadi	8		
Cañabate, Josep / Paz, Olga	Castilla-La Mancha	9	9	9
Cañabate, Josep / Paz, Olga	Aragón	1		
Cañabate, Josep / Paz, Olga	Castilla y León			
Ferrera, Laura	Andalucía	11	11	11
González, Cristina	Andalucía	8	5	5
Martínez, M. Carmen	Murcia	8		
Pérez, Lorena	Asturias	8	8	8
Poblet, Marta	Canarias	11		
Ramos, Francesc	Galicia	16	16	16
Sánchez, Jorge	Catalunya	2	2	1
Sánchez, Jorge	Navarra	1	1	1
Teodoro, Emma	València	6		
Urios, Cristina	Balears			
Vallbé, Pep	Andalucía	9	9	9
Vallbé, Pep	Extremadura	6		
Vallbé, Pep	Catalunya			
Totals		**111**	**68**	**67**

Figure 2.3: Phases of the ethnographic work as of December 2004

2.2.2 Judicial units

Judicial units visited by the IDT-UAB team are the so called "Courts of First Instance and Instruction [*Juzgados de Primera Instancia e Instrucción*]" which are the lowest courts in the judicial hierarchy (some municipalities do not have those Courts but have Justices of the Peace) and constitute the entry into the Spanish judicial system. In middle size cities such as those visited, first instance and instruction courts handle most civil cases and decide on minor criminal offences.[5] They are also responsible for

[5] In contrast, large cities such as Madrid, Barcelona, Valencia or Bilbao have separate courts for civil and criminal cases, so judges are in charge of both civil and criminal cases and, correspondingly, become either judges of first instance or instructing judges.

opening preliminary proceedings in any type of criminal offence. This is the reason why newly recruited judges have to deal with a vast variety of civil and criminal cases and they cannot specialize or concentrate in a particular law area.

Typically, first instance and instruction courts are composed of ten people: the judge, the judicial secretary, and eight judicial clerks belonging to different administrative categories—two officials, four auxiliaries, and two judicial agents—in charge of all the administrative paperwork. Judicial units of first instance and instruction are usually divided in two sections—civil and criminal—in correspondence with the dual character of those courts.

The 2003 Reform of the 1985 *Ley Orgánica del Poder Judicial* (LOPG) aims at reorganizing judicial units by allowing common services and giving more competencies to judicial secretaries (in terms of management of judicial units and administrative staff) so that incoming judges may concentrate their efforts on judicial activities and decision-making. The gradual introduction of new information and communication technologies (ICTs) envisaged by the law will certainly pave the way for further reforms of judicial units (case management systems, procedures, implementation of legal decisions, relationships with citizens and external institutions, etc.). At a very practical level, therefore, new professional profiles will be required for any member of the Spanish judicial system and it is not unlikely to expect a certain clash with skills, capacities, and merits currently required of judges, secretaries, and professional staff. Even though it is too early to assess the real scope of the new legal provisions, it seems clear that the 2003 Act leaves the door open to successive reforms of the Spanish Judicial system.

2.2.3 General findings

At this stage of the process, the IDT-UAB team has made 111 interviews (out of the envisaged 150) and has gone to all Autonomous Communities except Castilla-León. Even though the results can only be considered as preliminary, there are some qualitative findings to be pointed out as framing the institutional and organizational context of users. Data sources are the field diaries of the researchers.

2.2.3.1 Urban and rural areas: contextual heterogeneity

Newly appointed judges fill the vacancies left by their senior fellows in towns and middle-size cities of the peninsula, the Balearic, and the Canary Islands. Administrative capitals of the provinces are excluded. This contextual heterogeneity of destinations does not allow depicting a single model of judicial unit. Indeed, the typology of cases that new judges will find at their arrival largely depends on whether the court is located in a rural, metropolitan, tourist, or border area. Courts in rural areas are mostly occupied in agricultural land disputes requiring the judge to put into practice the intricacies of real estate law. Judges in metropolitan areas are usually busy with economic law concerning the activities of a myriad of middle and small companies operating in that particular area. Tourist areas generate stressful work agendas in those periods of the year in which the population of the area doubles, triples, or multiples by ten. And border areas are also concerned with specific problems of smuggling, drug trafficking, illegal immigrations, etc.

Apart from dealing with very specific types of cases, living in a rural area sometimes implies personal distress to judges for a number of reasons: confusion, loneliness, lack

of social interaction, etc. In those cases, judges perceive their destination as a compulsory toll to be satisfied the sooner the better, or as a place where they cannot feel as an anonymous citizen.

> I would have though that the judge would be happy with me not visiting her (the only day I was able to go she had plenty of work). But, on the contrary, she seemed to regret it, as if she though "another one who stays in Las Palmas and Tenerife and forgets X". To my surprise, she keeps talking to me for 15 minutes, showing her desperation for the situation in X, a sort of no man's land. The prosecutor refuses going there, the facilities are in a very bad situation, there is no assistance in family cases and no one seems to hear her petitions. She feels "confined" and "prisoner for nine months" with no chance to leave the island, "even not a permit to attend continuous training, even not in August to attend the wedding of my friend, like a sister to me". Do you mean that you stay in X also during the weekends? "Of course, of course, no weekends, nothing at all. I have been denied all permissions, maybe because I'm a beginner. This is torture. Imagine yourself nine months on an island on your own!!!"[6]

> For instance, I am here, I am living here on weekdays and, even there are 30,000 inhabitants or so, this is a town and (…) I am simply a normal person, but it turns that, ah, sometimes the others do not see you as a normal person; I have had a detainee here, and I went for a drink in the evening and found him in the same café and, you know, this makes me feel anxious, because I am not an unknown person to this person.[7]

2.2.3.2 Facilities and judicial units

The transfer of competencies on judicial matters to the Autonomous Communities also implies very different conditions regarding judicial buildings and offices. Since transfers are relatively recent and judicial budgets remain moderate in most Autonomous Communities it is not difficult to find a sheer contrast between brand new facilities, well entertained and adapted to present needs, and old infrastructures which are uncomfortable for workers and lack the essentials to conceive the judicial activities as a public service to citizens.

---

[6] "Em pensava que la jutge es mostraria alleugerida quan li digués que no hi podria anar, perquè l'únic dia que a mi m'anava bé ella el tenia ple de judicis. Ben al contrari, sembla lamentar-ho, com dient "un altre que es queda a Tenerife i Las Palmas i s'oblida de X. Sorprenentment, llavors inicia un monòleg de gairebé un quart d'hora explicant-me la seva desesperació que X estigui deixat de la mà de Déu. El fiscal no hi va, els mitjans són deplorables, no hi ha cap assistència pels casos de família i no atenen les seves peticions. Ella mateixa se sent confinada i diu que l'han tingut 'reclosa durant nou mesos' sense poder sortir de l'illa. 'Ni permiso para ir a cursos, ni siquiera en agosto para ir a la boda de mi amiga, que es como una hermana'. [¿Quiere decir que sin siquiera salir los fines de semana?] Claro, claro, sin fines de semana ni nada. Me han denegado sistemáticamente los permisos, no se si porque soy novata o que. Imagínese nueve meses sola y en una isla!!!'" [Marta Poblet, Fieldwork journal].

[7] "[...] es que por ejemplo yo estoy aquí, yo vivo aquí entre semana y aunque hay treinta y tres mil habitantes o así, es un pueblo y [...] pues simplemente yo soy una persona normal, lo que pasa es que eh, los demás a veces no te ven ser normal, y yo a lo mejor he tenido aquí un detenido y por la noche me he ido a tomarme una coca cola y me lo he encontrado en la misma cafetería, y a mí eso ya me hace sentirme violenta porque no soy una desconocida para esa persona sabes [...]" [Emma Teodoro, Fieldwork journal].

> This is a little bit of chaos because judicial units are very small and when two people come in it gets crowded and the staff members find it hard to move around. The secretary has the office between the civil and the criminal sections. I learned afterwards that there are some problems between the two sections, so it's good to have such a buffer zone between them.[8]
>
> The building is two years old and the unit has the typical distribution of current "landscape units", that is, open spaces with a broad reception bar (this one is made of marble) and the staff working in front of the public. Some staff member tells me that now it's good to work here, with so many windows and lots of daylight.[9]

2.2.3.3 Technological resources (ICT)

Just as in the previous case, technological resources of judicial units largely depend on the level of investment of each Autonomous Community. Even though data drawn from the questionnaire confirm that, for the vast majority of cases and geographic areas, every member of the judicial unit has a computer set (this was not the case a few years ago), the renewal of both hardware and software systems follow different paths at different speeds, even within the same territory of an Autonomous Community. Thus, it is not difficult to see how in judicial offices old type-machines coexist with up-to-date Pentium, or how Wordperfect 5.1 lives together with new case management systems. In some other cases, the problem lies not in the lack of computer equipments or case management systems, but in the lacking of a specific training to judicial staff:

> From my conversation with the judicial staff I see that, like any other areas in Asturias, they all have new brand computers (Pentium 4 amb Windows XP) but they don't explore all the possibilities of the system. They are not very much interested in the upgrades of the system and they only use the basic utilities of Libra, those which are essential to process the cases with the computer. But they told me that Libra was very difficult to implement because elder people did not wanted to adapt to computer technology progress. They basically work with paper (…). They acknowledge their fault in their lack of interest, but they also complain that it is difficult to get used to new systems and they need computer training, especially on how to exploit all the performances of the system. They want to attend training courses in Oviedo (not in the office, with all the staff crowded in front of a computer screen). They need more material means and to attend to courses out of work time, otherwise the office is unattended. They do not have e-mail, nor are used to the Internet.[10]

[8] "És un Jutjat una mica caòtic perquè les Oficines són molt petites i quan entren dues persones ja està tot ple i els funcionaris no poden ni moure´s. El Secretari té el despatx entre l'Oficina de Penal i la de Civil, que després m´enteraré que tenen problemes entre elles, així que ja va bé que hi hagi aquesta barrera de contenció" [Lorena Pérez, Fieldwork journal]

[9] "L'edifici és nou de fa dos anys i l'oficina té la típica planta de les anomenades 'oficines paisatge' actuals, és a dir, espais diàfans, amb un mostrador ample d'atenció al públic (aquest és de marbre) i els funcionaris treballant de cara al públic. Els funcionaris em comenten que ara sí que dóna gust treballar aquí, amb tantes finestres i claror del dia." [Marta Poblet, Fieldwork journal]

[10] "De la conversa amb els funcionaris veig que tenen, a l'igual que a tot Astúries, ordinadors força nous (Pentium 4 amb windows XP) però no exploten les possibilitats del sistema. No estan molt interessats en els avenços informàtics i només utilitzen les utilitats bàsiques de Libra, aquelles imprescindibles per tramitar informàticament els procediments, però m´expliquen que aquest va ser molt difícil d'implantar

2.2.3.4 Judicial staff

One of the perennial problems of the Spanish judicial system is that first instance courts in remote areas are plagued with vacancies of the judicial staff. As said earlier, in some Autonomous Communities members of the judicial staff still depend on the Ministry of Justice and they are organized at the national level. After a compulsory period of permanency in a judicial unit, judicial staff tends to move to another area, usually closer to their homeland. This also holds true for civil servants of the Autonomous Communities with competencies over judicial staff. As a result, in remote areas judges may occasionally find at their arrival either a deserted unit (i.e. the officials have moved to another area and the new ones have not yet arrived), a unit filled with substitute and poorly trained personnel, or a unit filled with newly recruited staff who aims to spend a short period of time there before moving to another area.

2.2.3.5 The judge condition

One of the main concerns of judges at the beginning of their service is how to manage judicial staff. Although this is a task that corresponds to judicial secretaries, judges are the heads of the judicial units and they usually need to establish some organizational ground rules as well, since their main task—judging and ruling—depends completely on a proper management of cases. Most frequently, judges lack knowledge in human management resources and feel at a loss on how to deal with vices, inner conflicts, lack of personnel, etc:

> The problem here is that the organization of the unit corresponds to the judicial secretary, but when you arrive you have to do all sorts of things. There are people who do not know how to switch on the computer, to switch on the computer! And the timework, I managed to raise it to 30 hours from 25, and it should be 37 hours. But tell them that and you will see how they react. This is a responsibility of the Ministry. If the Ministry does not offer a solution I cannot do anything, I can only impose disciplinary measures. I already did it once. A judge without a good team is no one. You may be a wonderful judge, but if people do not comply what you order them… what can you do about it?[11]
>
> They send personnel with much delay. Sometimes no one is sent. It is not up to you, you can do nothing. We are flooded with work, but we are the same number of staff members and judges. If you don't endure these timetables you may end up asking for a leave. I have class fellows on

---

perquè hi havia gent gran que no volia adaptar-se als avenços informàtics. En general funcionen molt amb paper a mà. Només introduexien en el sistema els procediments, però no els escrits. (...) Reconeixen la seva part de culpa amb aquesta manca d'interés però també es queixen de que els hi resulta dificultós acostumar-se a aquestes eines i que fan falta cursets sobre informàtica i especialment sobre com treure-li tot el rendiment al sistema. Però cursets que es facin a Oviedo (i no allà al Jutjat amb tots els funcionaris apilotonats en un ordinador), amb més mitjans, i fora de l'horari laboral perquè sinó deixen l'Ofician desatesa. No tenen e-mail ni estan acostumats a treballar amb Internet." (...)[Lorena Pérez, Fieldwork journal].

[11] "Aquí el problema es que la organización corresponde al secretario por ley pero cuando llegas aquí haces del todo. Hay funcionarios que no saben encender un ordenador, encender el ordenador! Y el horario, conseguí que se subiera de 25 a 30 horas y deberían ser 37. Díselo y verás donde te mandan. Eso es responsabilidad del ministerio. Si el ministerio no pone una solución yo no puedo hacer nada, sólo puedo poner medidas disciplinarias. Ya lo hice una vez. Un juez sin un buen equipo no es nadie. Tu puedes ser un juez maravilloso, pero si no tienes la gente que haga lo que ordenas, ¿cómo lo haces?" [Nuria Casellas, Fieldwork journal]

> leave of illness. I guess it is a question of money, of asking for more people. Justice is working far too well considering how it would work if we only worked the regular hours. Flooding, the incoming cases generate a lot of paper. The solution would be to create more units. You cannot study a case, you have no time. We would like to study it, ground it well, but you can't, you have no time.[12]

2.3 Statistical Data

Sections 2.3.1 to 2.3.3 below offer a description of the main parts of the questionnaire to judges, recall the characteristics of the sample and provide some descriptive statistics relevant to the legal case study. Since the ethnographic work has not been completed yet, it is necessary to recall that both qualitative and quantitative data are to be considered as preliminary. Nevertheless, they already are already a vantage point from which user requirements may be refined.

2.3.1 Description of the questionnaire

The questionnaire of the SEKT ethnography draws from the experience of the OJC ethnographic campaign of 2002-2003. As Figure 2.4 below shows, the OJC questionnaire was organized in five sections concerning different aspects of the professional domain of judges. The initial template differs from the new one with regards to questions about the evaluation of the training at the Judicial School and the evaluation of the profession of judge have been deleted in the SEKT questionnaire. In this way, the SEKT template puts more emphasis on organizational issues and eliciting of questions regarding difficulties.

Domains	Number of questions	Examples
Professional Training	• 9 closed questions	Where did you graduate in law? Do you have postgraduate studies?
Professional Activity	• 6 closed questions • 12 open questions	What were the most complicated civil cases that you had to solve in your first year as a judge? Could you define the kind of criminal cases that have given you more work? Do you discuss a judicial case with someone else because are you are worried about its resolution? Frequency of consulting legal doctrine Do you use Internet? What would you expect from a software

[12] "También tardan mucho en mandarte personal. A veces no mandan a nadie. No depende de ti que tu no puedas hacer nada. Estamos desbordados de trabajo, seguimos con el mismo número de funcionarios y jueces. Esos horarios se aguantan o terminas con bajas. Hay gente de la promoción que está de baja. Supongo que es cuestión de dinero, pedir refuerzos. La justicia va demasiado bien para como podría ir si hiciéramos las horas normales que hace todo el mundo. Desbordamiento, tenemos un número de entradas que genera mucho papel. La solución sería la creación de más juzgados. No puedes estudiar un caso, no tienes tiempo. A nosotros nos gustaría estudiarlo, fundamentarlo bien, pero no puedes, no tienes tiempo". [Nuria Casellas, Fieldwork journal]

		supporting the task of the judge?
Professional Relationships	• 20 closed questions	Do you have any professional relation with your class fellows? Do you have any professional communication with other judges? Do you have any contact with the media? Are you a member of NGOs? Have you ever been invited to official acts as a judge of the area? Do you live in the same place where the court is located?
Quality of life	• 6 closed questions • 1 open question	Do you bring work at home? Do you work over the weekends? Could you evaluate the degree of satisfaction with your daily work? Could you explain the reasons of such evaluation?
Personal Data	• 7 closed questions	Year and place of birth, Autonomous Community, civil status, number of children, profession of parents and spouse
Data concerning the judicial unit (asked to secretaries)	• 17 closed questions • 4 open questions	Number of computers in the unit, number of Internet connections, case management system used in the unit How do you evaluate the performance of the CMS? What do you consider to be the most relevant organizational problems in your unit? How do you see the relationship with citizens? What could be improved?

Figure 2.4: The SEKT template

2.3.2 Sample and technical data

As said earlier, the sample of the SEKT survey is composed of 150 individuals (out of 248 judges of the 52nd class of the Judicial School). The method used is semi-structured oral interviews with an estimated length of one hour. Judges were granted absolute confidentiality and anonymous treatment of personal data. 90 % of the interviews have been recorded with minidisk sets to facilitate transcription protocols (except when the judge explicitly refused being recorded). At this stage of research, data from 90 questionnaires have been already coded with SPSS. Preliminary data regarding descriptive statistics of the survey are shown in sections below. In some cases, these data are compared to the total population of the class (gathered as it started the training at the Judicial School) and, additionally, to the total population of Spanish judges.

2.3.3 Statistical Results

2.3.3.1 Judge profile

As figure 2.5 shows, the vast majority of newly recruited judges of the 52nd class are at present in their early thirties (the average age at the moment of accessing the Judicial School was 28.6 years old). As regards gender —with respectj to the total N— 67 % of the judges are male and 33 % females (in general, women represent 40 % of the Spanish judiciary). Women are more frequently married than men, as figure 2.6 also shows.

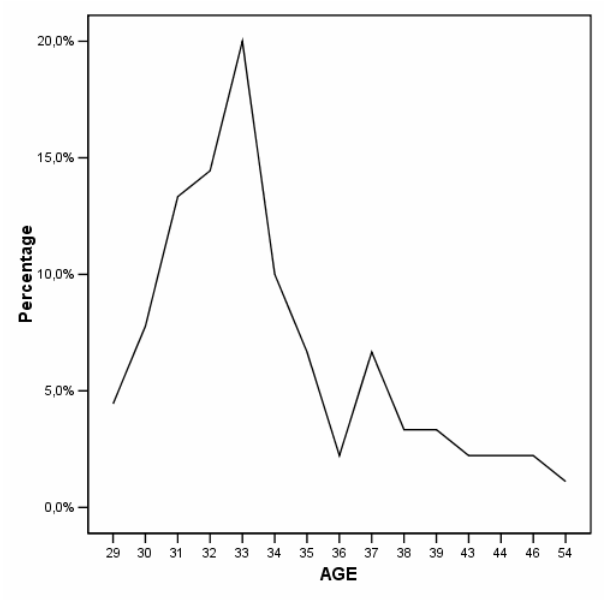


Figure 2.5: Average age of judges

		CIVIL STATUS				
		Married	Separated/ Divorced	Partner	Single	Total
SEX	Male	13	1	0	16	30
	Female	31	1	2	26	60
Total		44	2	2	42	90

Figure 2.6: Civil status and gender

The majority of newly recruited judges graduated in law by the middle nineties. This figure is coincident with data from the Judicial School, according to which successful candidates have spent up to four years after graduation preparing the competitive examination. To do so, they will have typically spent 12 to 16 hours per day in front of the textbooks and will have hired a "coach" or "*preparador*" (usually, a senior judge or prosecutor) to train them on how to prepare the oral examination before the tribunal.

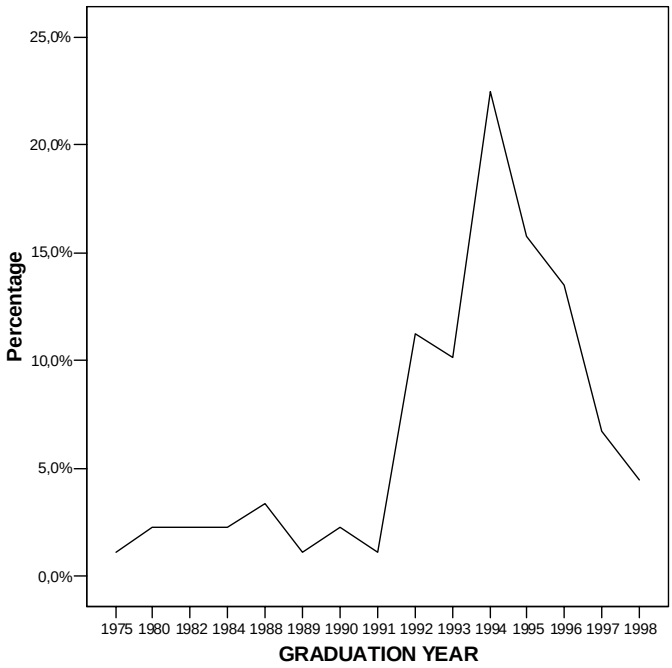


Figure 2.7: Graduation year of judges

2.3.3.2 Work environment

As regards work environment, 62.2 % of judges interviewed stated that civil cases were the ones which required their main attention at the beginning of their appointment. This is not due to any special complexity of civil issues over criminal ones, but to the fact that civil cases usually take much longer to be solved than criminal ones. In courts where judges tend to stay no longer than two years, "inherited" civil cases from previous judges are much more frequent than criminal ones.

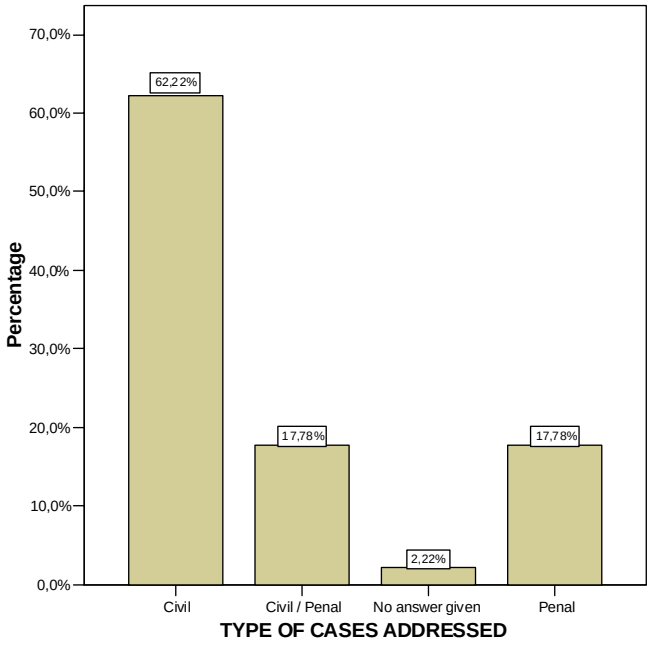


Figure 2.8: Cases requiring the main attention of the judge

As regards work caseload, almost 95 % of judges interviewed declared to bring work at home in the afternoon and 86.6 % added that they worked over the weekends as well. On average, judges work 23.89 extra hours per week (either at home or remaining at the office late in the afternoon). Little surprise, therefore, if 64.4 % of judges consider that work pressure is "high" or "very high" (Figure 2.12).

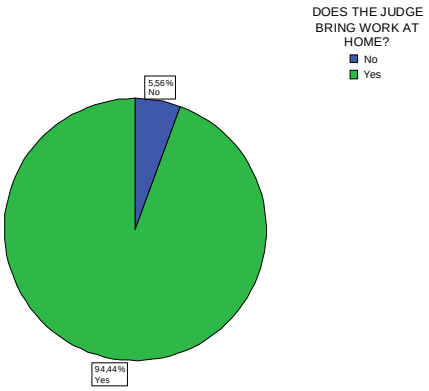


Figure 2.9: Bringing work at home

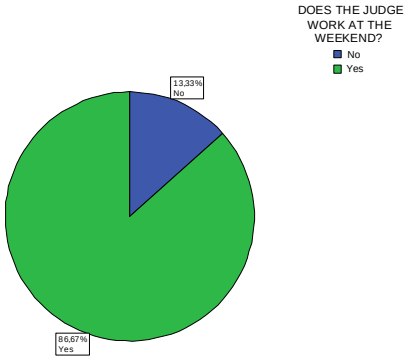


Figure 2.10: Working at weekends

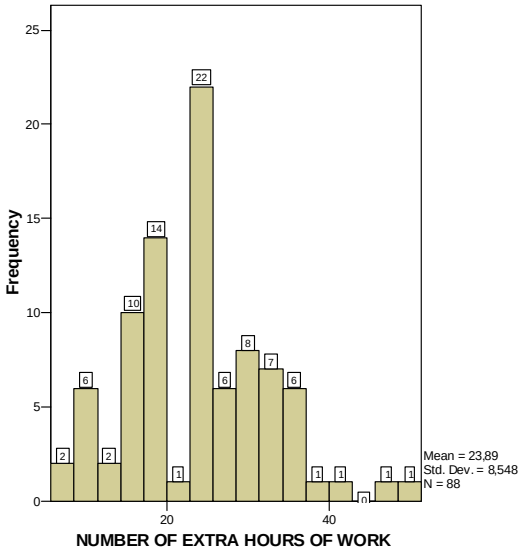


Figure 2.11: Number of extra hours at work

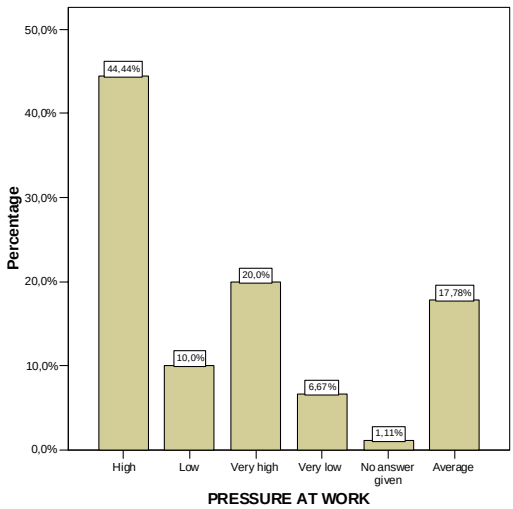


Figure 2.12: Pressure at work conditions

Another important aspect regarding work environment involves the frequency with which the new judge talks about the cases he/she is dealing with. Only 3.3 % of the

judges say that the never exchange information concerning their cases with other people. On the contrary, the vast majority of them have affirmative answers regarding information exchange. When asked with whom they talk about their cases, judges typically offer three answers: (i) senior judges or magistrates; (ii) fellows of the same graduation year, and (iii) fellows from the other judicial units of the building. This figures are consistent with one of the core hypothesis of the research: the use of senior judges' professional knowledge as a parameter when dealing with difficult or unusual cases.

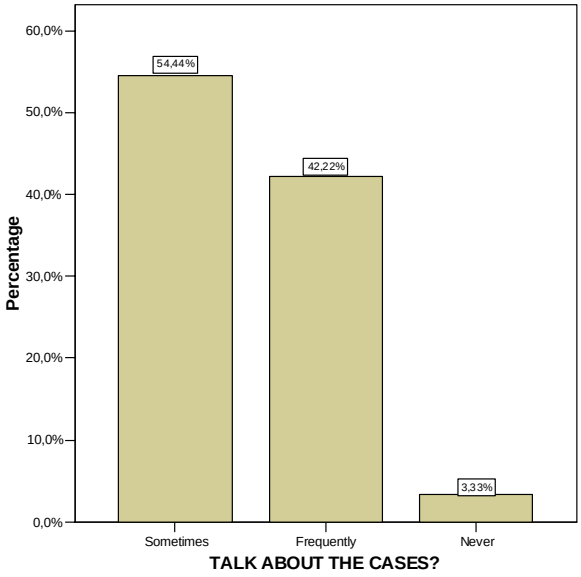


Figure 2.13: Frequency of information exchange regarding cases

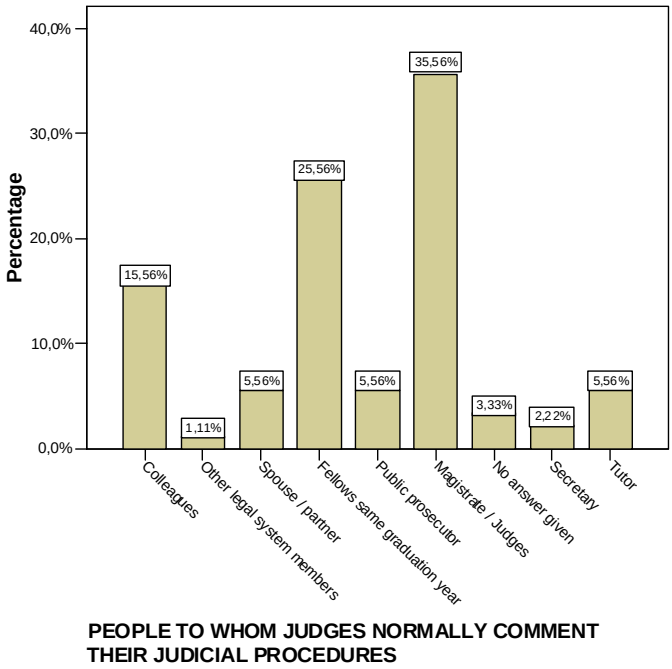


Figure 2.14: With whom judges comment their cases

As the procedure followed to comment judicial cases, judges use either the telephone or the personal interview—most usually, an informal interaction with fellow judges. Significantly enough, none of them declares using electronic mail or instant messenger systems to communicate with peers or senior fellows. Again, this is totally consistent with previous data showing a rather scarce use of ICTs in court. Therefore, it is necessary to insist on the fact already highlighted in D.10.1/ Legal Case Study Before Analysis, that is, users of the system will be judges who have medium or low technological abilities, and are not used to new technologies.

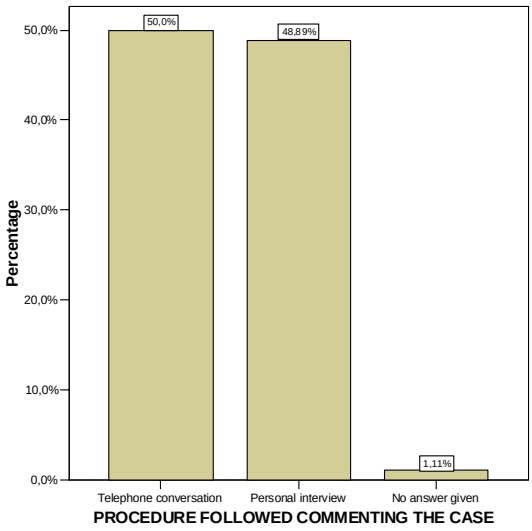


Figure 2.15: Procedure followed to comment judicial cases

2.3.3.3 Technological skills

In 2000, 47 % of members of the 52nd class declared to have no computer skills whatsoever. Those who declared to have some were mostly used with text processors. 61 % did not have a PC at home, and only 35 % of the students declared to be Internet users [ABAVAS, 2003].

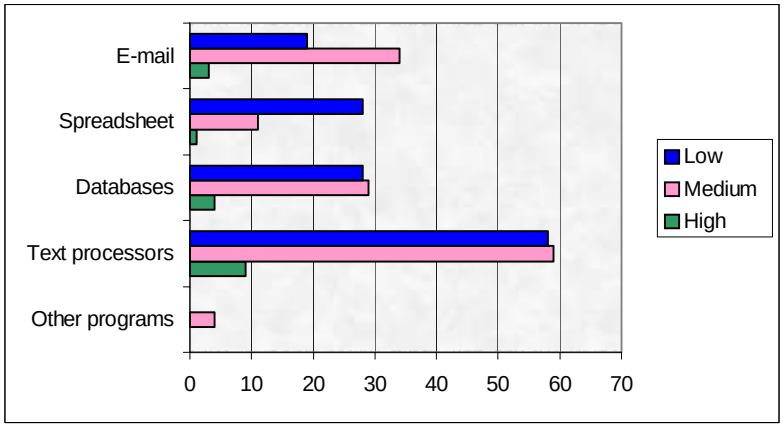


Figure 2.16: Computer skills of the 52nd class in 2000 [3]

According to the SEKT survey, nevertheless, the use of Internet among judges seems to be higher than at the beginning of their training at the Judicial School. Four years after accessing the School, the proportion is reversed: 81 % of the interviewed declare to use

Internet and only 18.8 do not use it. Clearly, having a computer set connected to Internet at the desk has fostered the daily use of the web among entrant judges.

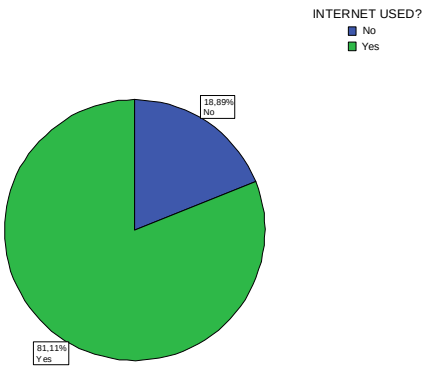


Figure 2.17: Use of Internet

As for the type of information searched through the Internet, the page of the Official Bulletin of the State is the most accessed site, followed by legal information in general. Additional uses are not frequent, and judges will typically argue that they have no time navigate through the Internet.

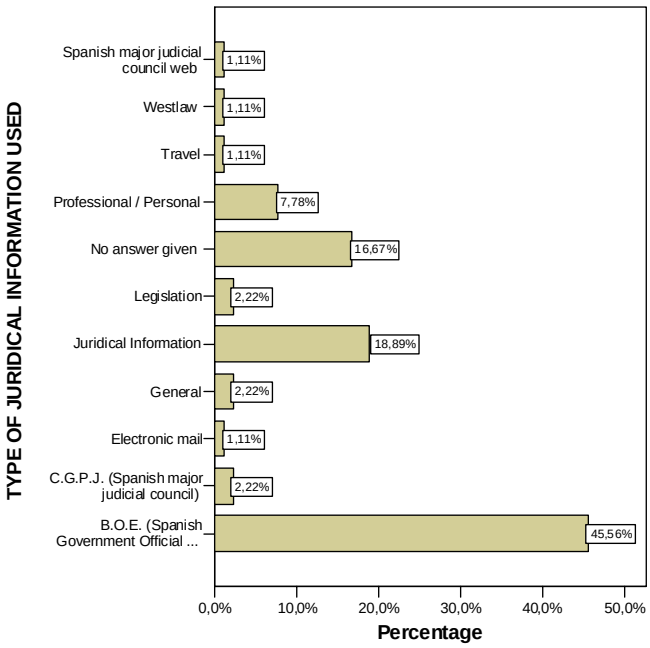


Figure 2.18: Information accessed through the Internet

Finally, judges offer an interesting answer to the question "which would you like to find if judges were given a web service system". In this case the majority of them proposed a site were doubts regarding professional cases could be put in common and discussed.

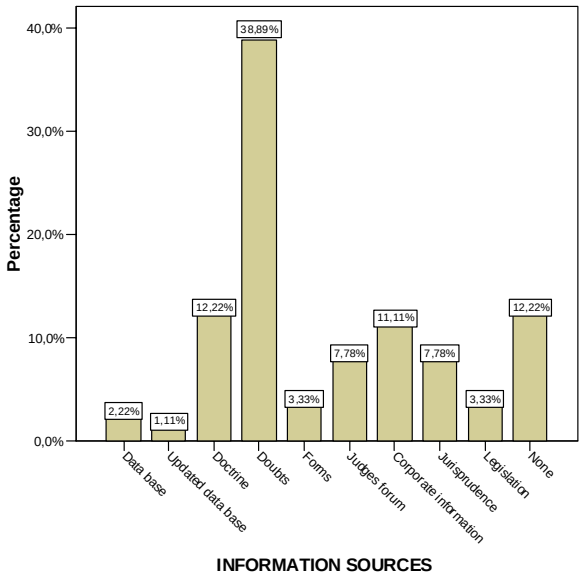


Figure 2.19: Type of desired information in a web service

2.4 Expert Knowledge

2.4.1 Transcription protocols

There are three kinds of documents that constitute the written (transcript) results of the ethnographical campaign throughout the Spanish courts during 2004:

2.4.1.1 Literal transcripts of the interviews

Most of all interviews were recorded (using Minidisk), so that every researcher could transcript them literally afterwards. This type of protocol has given us an important number of qualitatively rich ethnographic documents regarding the way the judges speak about their daily work. This protocol offers us many opportunities to perform our textual analysis. On the one hand, as has been showed before, the length of the questionnaire —due to the open-ended questions— allows us to regard every single transcript as a textual corpus to be analyzed by itself. Moreover, we could also build and analyze a large unified corpus with all these literal transcripts.

What kind of matter has drawn most of your attention in your first year as a judge, either civil or criminal?
Well, Civil, you have to study a lot because of the judgment, but there are also criminal cases where I had to spend a long time. I had a case at a Jury, although it was later suspended; it was already waiting when I started, then I had to study a lot. Afterwards I had cases on drugs and fraud with credit cards, that is, offences that have to do with organized delinquency, where I have had prisoners, and to which I have dedicated a long time.

To sum up: the cases that have required more attention, not so much because of the burden of work or study, but rather.... are criminal cases
Criminal proceedings because in comparison with others the criminal court is almost the same as a Court in Torremolinos. We have 5.000 preliminary actions and if we want this not to get blocked, last year we closed with more than 5.000. And if you want it not to get too blocked you have to press a lot and dedicate your time to the truly important issues. Then, to achieve this I have three persons only dedicated to preliminaries, and petty offences.

Figure 2.20: Example of a literal transcript.

2.4.1.2 Completed questionnaires

Besides the literal transcripts, every researcher has filled in the questionnaire with all the judges' answers, avoiding (slightly correcting) natural language incoherences (grammatical, syntactical, etc.). The answers have been inserted according to the order set up by the questionnaire. This factor (the order of questions and answers) will give us the opportunity to analyze, for instance, all the answers of a single question among all the questionnaires (this is done with ALCESTE). Moreover, we also have the results of the closed-ended questions which can be analyzed with other statistical programs (e.g. SPSS).

What documents do you usually look up when solving a matter?: *(several possible answers)*

1. Case-Law in paper
2. **Case-Law in the data bank**
3. Jurisprudence
4. Statistics
5. Sociological reports
6. Other
7. DON'T KNOW/NO ANSWER

How often do you read jurisprudence?

1. **Rarely**
2. Sometimes
3. Regularly
4. Frequently
5. Very frequently

Do you use Internet?

1. Yes
2. **No**
3. DON'T KNOW/NO ANSWER

Figure 2.21: Example of a completed questionnaire

2.4.1.3 Extracted questions

At a particular moment of the interview (see part II.2 of the questionnaire), the researcher asked the judge to formulate concrete questions about the main problems faced in the daily work concerning civil, criminal jurisdiction and the on-duty period. Actually, only some of the judges formulated short and concrete questions; others preferred to explain their doubts or problems in great detail. UAB researchers extracted all those questions formulated by judges. Once organized, this constitutes another important corpus to be analyzed.

If a person appears to make a report about himself, in what capacity should the judge consider he is reporting?

A problem of procedural nature: an individual commits an offence within a court district, and he is arrested in the next district (or else in a different one), before which judge should the person arrested appear?

If the court receives a testimony or police report on repeated telephone calls that do not amount to a criminal offence (the calls are then called "ill intended"), what kind of proceedings should be started and thus, how should you handle the inclusion of such calls within the circuit of ill-intended calls?

When the judge is on duty he receives a call from the Hospital Clínico informing that there is a sexual abuse. The victim has not reported the facts yet. Actions to be taken. Where should you place them?

How should I organize a round of recognition of suspects if there are no people available?

Which are the actual functions / competences the judge as compared with those of the secretary?

Figure 2.22: Example of extracted questions.

2.4.2 Expert knowledge

2.4.2.1 Acquisition

The acquisition of the judicial knowledge and understanding of the problems faced by newly recruited judges in their daily work is essential to *Iuriservice* (prototypes I and II) as a web based application that retrieves answers to questions in the judicial domain [CBCLU, 2003] [BCCABLU, 2003]. It provides judges with access to frequently asked questions (FAQ) through a natural language interface. The system responds with a list of similar question-answer pairs that offer solutions to the problem faced by the judge. This application can also be used as a traditional FAQ system, by selecting the appropriate question from a list.

Thus, both prototypes, *Iuriservice I* and *II*, provide answers to problems dealt by judges in their daily professional activity. The judicial knowledge, the answers to these problems, are mainly based on experience and peer-to-peer transmission. As it is referred to in Chapter 2, this knowledge has been acquired through an ethnographic

process designed by the UAB team (experts from different disciplines e.g. Law, Sociology and Anthropology) to efficiently obtain useful and representative information from questionnaire-based interviews. For this, *Iuriservice I* and *II* keep significant points in common; the main ones being the use of real information obtained on the field and the use of ontologies for information retrieval.

2.4.2.1.1 First prototype (Iuriservice I)

a) Knowledge acquisition process

The knowledge stored in the FAQ repository in *Iuriservice I* was extracted from the answers given by judges to the *Spanish Young Judges Survey 2002*. It consisted of two sets of interviews. In the first one the interviews were questionnaire-based. Once the main difficulties faced by judges had been identified, the second set of interviews tackled these problems directly [BCBRCP, 2004].

The main objective of the research was to implement a technological network to support the newly appointed judges in their daily decisions. The survey was designed to find out the problems that they could have at the beginning of their professional career. *Iuriservice I* was, therefore, concerned with the questions regarding the doubts faced in their first appointment.

Questions:

- *What kind of judicial cases (civil or criminal cases) were more frequent in your first destination?*
- *What were the two most important doubts that you had in your first three months as a judge?*
- *What were the most complicated civil cases that you had to solve in your first year as a judge? Why?*
- *What were the most complicated criminal cases that you had to solve in your first year as a judge? Why?*
- *Could you define the civil case that has given you more work?*
- *Could you define the criminal case that has given you more work?*
- *Which are the most frequent problems that you have found in the hearings?*

Figure 2.23: The first and second questionnaire (professional activity)

The open-ended questions listed in Figure 2.23 were analyzed with textual statistics methods to extract the domains where the difficulties lied. In general, the problematic domains were mentioned in the answers by the use of substantives. E.g. to the question "tell us the two main doubts that arose during your first three months as a judge", the over-represented substantives in the young judges' subcorpus (called *JJ*) were found

significant.

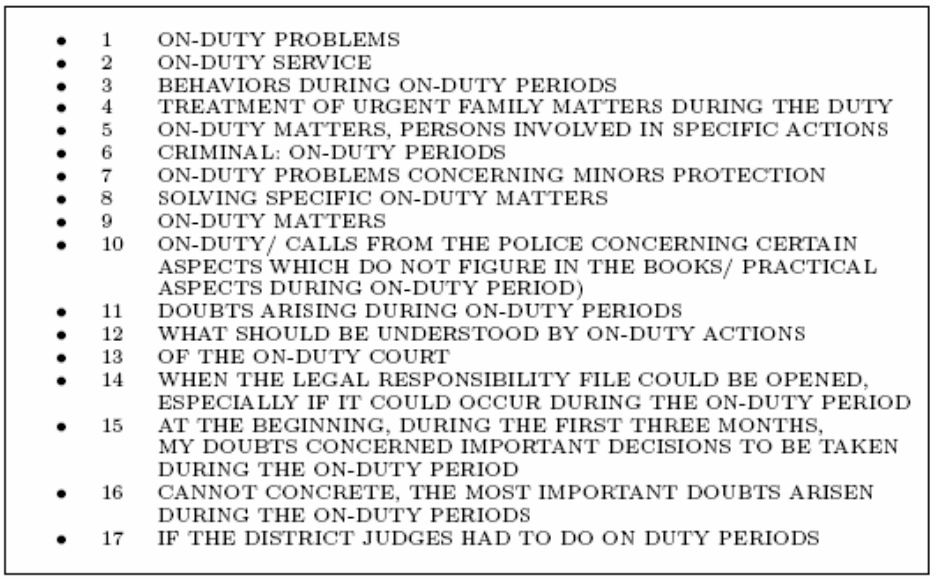
For each word *i*, the OCJ team leaded by M.Ayuso and M. Bécue [ABAVS, 2003] [BCCABL, 2003] [AAB, 2005] had to assess a statistical conclusion about the significance of the count *kij* , the frequency of the word *i* in part *j* of the corpus. For example, 18 occurrences of the substantive *guardia (on-duty)* were used by the judges. The count *kij* had to be compared to the other counts obtained with all the possible samples comprised of *k:j* occurrences extracted from the whole corpus without replacement. The probability distribution for a sampling without replacement under the hypothesis of independence is the hypergeometric distribution, with the parameters:

- *k..*, the size of the whole corpus (the corpus contained 2665 occurrences)
- *ki.*, the frequency of the word *i* in the whole corpus (20 occurrences of *guardia* in the whole corpus)
- *k.j*, the size or number of occurrences of part *j* (the subcorpus *JJ* had a global size of 1331 occurrences)

Prob (*k.., ki., k.j, n*), as calculated using the hypergeometric distribution, is the probability of obtaining exactly *n* occurrences of the word *i* when extracting a sample without replacement of size *k.j* among a population of total count *k..*, knowing that there are *ki.* repetitions of the word *i* in the whole corpus. Then, the substantive *guardia* covers 1.35% (= 18/1331) of the area of the subcorpus but only 0.75% (= 20/2665) of the area of the whole corpus. This word is over-represented in the subcorpus. To assess if this over-representation is significant, Bécue, Alvarez and Ayuso had to calculate the probability of having a number of occurrences of this word *i* greater than or equal to *kij* among the *k.j* occurrences randomly chosen. This probability is given by:

$$\sum_{n=k_{ij}}^{k_{.j}} \mathrm{Prob}(k_{..}, k_{i.}, k_{.j}, n) = \frac{\binom{k_{i.}}{n}\binom{k_{..}-k_{i.}}{k_{.j}-n}}{\binom{k_{..}}{k_{.j}}}$$

In case of *guardia*, the corresponding *p*-value was equal to 0.0002. The conclusion was that the high frequency of *guardia* was very significant in the subcorpus *JJ* and the topic *guardia* raised many problems and doubts (a list of answers mentioning *guardia* in the subcorpus JJ are listed in Figure 2.24). For that, *guardia* could be identified as an important topic to consider in the FAQ system. A more detailed statistical analysis can be found in [ABAVS, 2003] and [BCCABL, 2003].



- 1 ON-DUTY PROBLEMS
- 2 ON-DUTY SERVICE
- 3 BEHAVIORS DURING ON-DUTY PERIODS
- 4 TREATMENT OF URGENT FAMILY MATTERS DURING THE DUTY
- 5 ON-DUTY MATTERS, PERSONS INVOLVED IN SPECIFIC ACTIONS
- 6 CRIMINAL: ON-DUTY PERIODS
- 7 ON-DUTY PROBLEMS CONCERNING MINORS PROTECTION
- 8 SOLVING SPECIFIC ON-DUTY MATTERS
- 9 ON-DUTY MATTERS
- 10 ON-DUTY/ CALLS FROM THE POLICE CONCERNING CERTAIN ASPECTS WHICH DO NOT FIGURE IN THE BOOKS/ PRACTICAL ASPECTS DURING ON-DUTY PERIOD)
- 11 DOUBTS ARISING DURING ON-DUTY PERIODS
- 12 WHAT SHOULD BE UNDERSTOOD BY ON-DUTY ACTIONS
- 13 OF THE ON-DUTY COURT
- 14 WHEN THE LEGAL RESPONSIBILITY FILE COULD BE OPENED, ESPECIALLY IF IT COULD OCCUR DURING THE ON-DUTY PERIOD
- 15 AT THE BEGINNING, DURING THE FIRST THREE MONTHS, MY DOUBTS CONCERNED IMPORTANT DECISIONS TO BE TAKEN DURING THE ON-DUTY PERIOD
- 16 CANNOT CONCRETE, THE MOST IMPORTANT DOUBTS ARISEN DURING THE ON-DUTY PERIODS
- 17 IF THE DISTRICT JUDGES HAD TO DO ON DUTY PERIODS

Figure 2.24: List of the answers given by newly recruited judges containing the substantive *on duty* (first questionnaire)

Finally, after the analysis had been performed on all the questions, legal experts were able to derive a large set of questions regarding the difficulties faced in the judicial activity, e.g. the judicial process, all those problems arising in the on duty period, etc.

b) Domain ontology

The design of the ontology started from the information and the FAQs provided by the Spanish young judges survey, using the "competency approach" [GF, 1995] to identify relevant aspects and the coverage of the ontology. Nearly one hundred competency questions were extracted from the ethnographic work.

As we have stated above, the doubts regarding the judicial process and questions arising in the on duty period were the most relevant. The judge on duty has to make quick decisions about the facts of a case, relevant measures to establish or the applicable procedure for a given case. Therefore, the most usual set of questions take the form of "what I should do in such a situation?" More experienced peers are often consulted to reply to these questions; the professional legal knowledge contains a repository of know-how solutions, next steps, ready made procedural and practical knowledge, for a huge amount of similar cases which are not covered by statutory provisions.

Our first Ontology for Professional Legal Knowledge (OPLK) was based on the common ground of knowledge that any young inexperienced judge shared with the more experienced ones. That is to say, we inferred some matching concepts from the bulk of materials that we had before us (hard cases, rare cases, legal interpretations, legal analogies, professional attitudes, and common standards). The most general concept we found in the judicial criminal field was *proceso* (process, trial, procedures), the Spanish procedural notion that stands for all kinds of proceedings under the Spanish law. This notion constitutes the kernel of a wide network of related concepts that shape the backbone of the judicial culture. A possible representation is offered below:

1. Ordinary Trial: a) beginning + b) agents
2. Preliminary Investigation: a) Building of the Records: (i) findings (ordering) + (ii) personal area (ordering, rights) + (iii) liability + (iv) secondary liability, b) End of the Records (file OR start proceedings)
3. Criminal Hearing (summary trial OR instruction)
4. Misdemeanor Trial
5. Preliminary Investigation of the Jury Trial + Jury Trial.

Figure 2.25 represents this professional knowledge graphically. It is a mixture of taxonomic (first level), part-of (second level) and decision knowledge (third level). The most important decision to be taken after the preliminary investigation is if the proceedings are taken to trial or filed stating that there is no criminal case to be ruled.

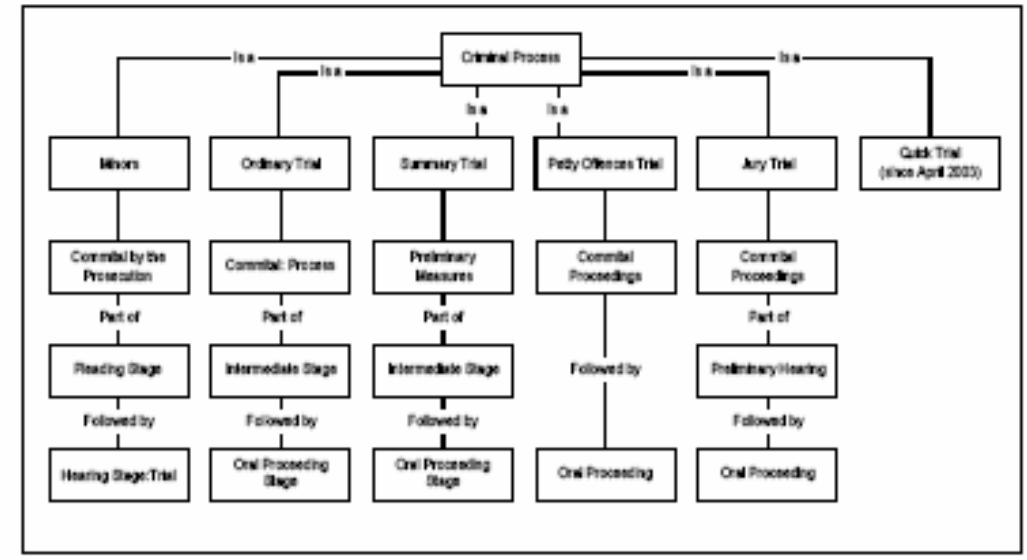


Figure 2.25: Representation of Processes Types in Spain (English).

To identify all the "competency questions" that the ontology had to take into account, this dynamic flow had to be captured. Judges use it as a kind of cognitive tool for a quick understanding of the facts that are submitted to them. They can select the appropriate legal procedure through this framework. Therefore, going along of these guidelines, they may think of what to do first.

We described this complex conceptual structure (*proceso*) as triggering general cognitive *schemas* and *scripts* or *prototypes*. A *schema* is an organized framework of objects and relations that has yet to be filled in [S, 1990]. A *prototype* is created through the filling in of the slots of a schema with an individual's standard default values [DR, 1995].

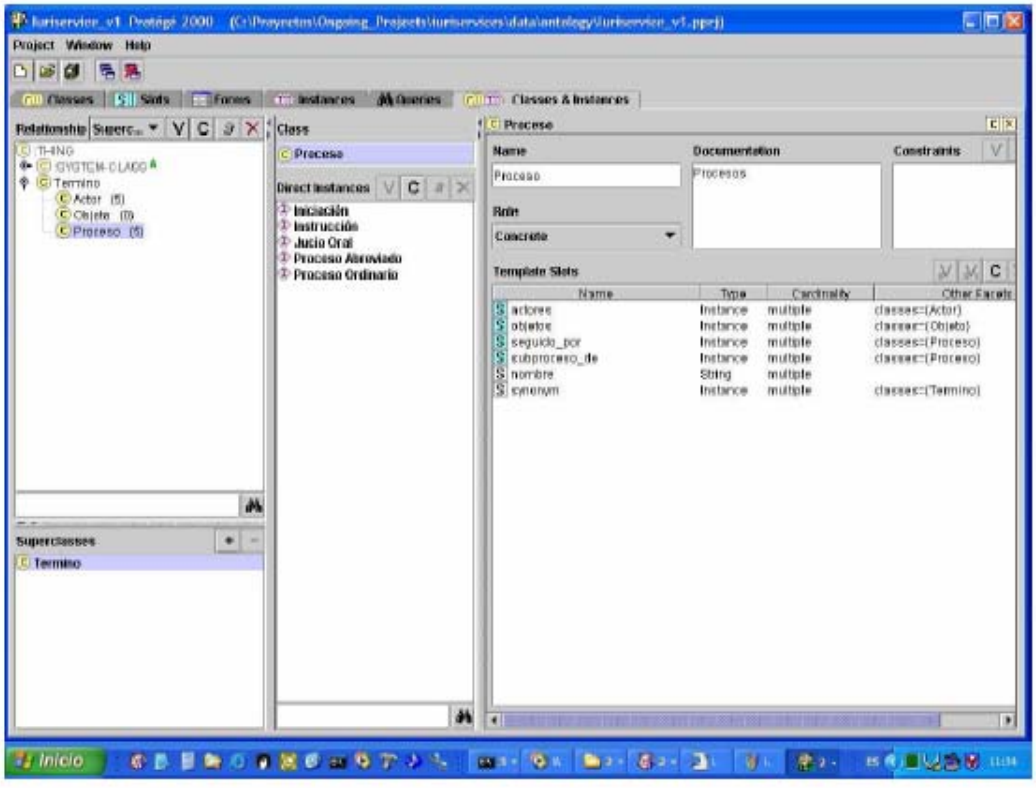
We assumed that this preliminary ontology for professional legal knowledge, even if still lightweight and only formulated in a semiformal language, captured the templates that judges had to fill in almost automatically with the bulk of cases and situations that they encountered while being on duty. Therefore, this structure allowed the system to reply through the same set of basically related concepts that users (young judges) had

in mind in their consultations. The first version of the ontology for professional legal knowledge (implemented in the first *Iuriservice* prototype using Protégé 2000[13] is shown in Fig. 2.26, and includes the following terms:

- *Process* and its instances: different trial processes or their parts
- *Object* and its instances: physical or abstract inanimate objects representing documents, information, physical items used by a process or an actor, as an input or result of a process
- *Actor* and its instances: Persons or organizations able to execute changes within the model. This concept was similar to *agent* as in [Mi, 1995].

These concepts were linked through the following attributes, which represented relations:

- *Generalization*: the *is-a* relation that allows representing that one concept is more general than another, e.g. an actor *is a* person. Can be applied to any concept.
- *Equivalence*: allows relating two concepts that are synonyms in this domain. Can be applied to any concept.
- *Actor*: process instances are associated with actors that participate in that process. The link is made through this (*actor*) attribute.
- *Follows*: attribute for processes to determine the logical or temporal order for processes or their parts.
- *Part of*: applied to processes or objects to represent when one concept is a part of another, e.g. instruction is *part of* the trial process



Figu

Figure 2.26: Screenshot of the first version of the legal ontology in Protégé

[13] http://protege.stanford.edu

The first *Iuriservice* prototype is now being revised not only towards the integration of the technology being developed in SEKT but also towards the improvement of the ontology used for the FAQ retrieval system. This last issue will be introduced below and further discussed in Chapter 3.

2.4.2.2 Second prototype (Iuriservice II)

The *Iuriservice* prototype *II* is a web-based FAQ application, which also provides answers to problems dealt by judges in their daily professional activity. There are, however, important differences regarding the knowledge acquisition process and the methodology used in the construction of the domain ontology, hence, there are significant differences regarding both, the quality of the professional knowledge stored within the system and the ontology itself.

a) Knowledge acquisition process

For the construction of this second prototype, the knowledge acquisition process was more focused on the doubts, difficulties and problems faced by the newly appointed judges. The questionnaire was modified in order to give more importance to those questions which were useful to obtain information regarding all those problems. This focus could be established due to the fact that the previous questionnaire had pointed out the areas where the difficulties were located, e.g. "on-duty", (civil and criminal) *process*, etc.

This new approach made possible, for example, to identify more than 750 competency questions, in comparison to the 100 questions that were extracted from the answers to the first questionnaire. Thus, the quality of the FAQ repository improves. Not only the system has more question-answer pairs to offer, but also the typology of problematic domains develops into such a fine detail that allows better replies on one specific problem (e.g. 163 of the new competency questions are related to the *on duty* period alone).

b) Domain ontology

Several changes have been introduced regarding ontology modeling issues. First, several KAON applications, provided by the SEKT partner AIFB, like the ontology editor Oi-Modeler, have been used in order to model and visualize the domain ontology[14].

Secondly, we have used two different software applications to analyze the competency questions (or extracted questions' protocol as explained above) and extract the relevant concepts: TextToOnto and ALCESTE.

- TextToOnto is a tool embedded in the Oi-Modeler platform which supports the semi-automatic creation of ontologies by applying text mining algorithms[15]. Although TextToOnto is not currently provided with textual analyzer components in the Spanish language, it is able to identify important concepts and instances that the judicial domain ontology has to take into

[14] http://kaon.semanticweb.org
[15] http://km.aifb.unikarlsruhe.de/kaon2/Members/rvo/Module.2002-08-22.4934

account. This version of TextToOnto will not be further developed, but as a result of the SEKT Project, Spanish GATE components will be implemented into Text2Onto in the near future, a new tool, designed by the same team, with improved features.[16] Then, the competency questions will be analyzed accordingly and more information will be retrieved to refine the existing ontology, so the legal case study can benefit directly from the development of SEKT technology. The conceptual analysis obtained with TextToOnto is described in below (2.4.2.3.1).

- ALCESTE is a software used to perform automatic analysis of textual data, developed by the Centre National de la Recherche Scientifique (CNRS), in order to quantify a text and extract the most significant structures[17]. It's findings are also specified in 2.4.2.3.2.

Finally, we have also followed the Distributed, Loosely-controlled and evolving Engineering of oNTologies (DILIGENT) argumentation model during the ontology engineering process, provided by AIFB (within WP7 of the SEKT Project) [PST, 2004]. This argumentation model is based on the rhetorical structure theory. The visualization of the arguments took place on a wiki-based environment which also allows them to be traced.

The use of the DILIGENT methodology in the legal case study implied the analysis of the content of each competency question, the identification of the relevant concepts and possible relations and the controlled discussion of each of the arguments used in favor or against the introduction of a concept *X* as part of the domain ontology. This discussion lead to the current version of the domain ontology with 50 concepts, more than 300 instances and 100 relations. The particular uses of the DILIGENT argumentation and the domain ontology itself, the Ontology of Professional Judicial Knowledge, are detailed below (Chapter 3).

2.4.2.3 Distribution

2.4.2.3.1 Conceptual distribution

As it has been stated above, TextToOnto, despite its limitations, has been used in order to extract significant concepts from the competency questions for the domain ontology. The algorithms used in the extraction of concepts by this application are: the term extraction algorithm, the concept association extraction algorithm and the ontology pruning algorithm. The use of Spanish language introduces limitations to the concept association extraction and the ontology pruning algorithms. Besides, the term extraction algorithm has proved to be useful, even though the Spanish language components are not yet integrated.

The term extraction algorithm extracts sets of terms that might be considered potential concepts to be included in the ontology. We used two corpuses in order to extract this terms. The first corpus consisted of all the questions related to the on duty period. The term extraction resulted as follows:

[16] http://ontoware.org/projects/text2onto
[17] http://www.image.cict.fr/english/index_alceste.htm

Word	Frequency	TFIDF	Entropy	C-value
juez	61	3.738	1	-18.675
o	56	3.738	1	-17.675
con	54	3.738	1	-17.675
guardia	40	3.738	1	-16.675
su	38	3.738	1	-15.675
debe	38	3.738	1	-15.675
si	36	3.738	1	-15.675
cómo	35	3.738	1	-15.675
prisión	30	3.738	1	-15.675
auto	28	3.738	1	-14.675
policía	26	3.738	1	-14.675
me	23	3.738	1	-14.675
protección	21	3.738	1	-14.675
pero	19	3.738	1	-13.675
declaración	19	3.738	1	-13.675
caso	18	3.738	1	-13.675
está	18	3.738	1	-13.675
como	16	3.738	1	-13.675
puede	15	3.738	1	-13.675
contra	14	3.738	1	-13.675
cuando	14	3.738	1	-13.675
art	14	3.738	1	-13.675
mujer	13	3.738	1	-13.675
hago	13	3.738	1	-13.675
orden	13	3.738	1	-13.675
dice	12	3.738	1	-13.675
son	12	3.738	1	-13.675
instrucción	12	3.738	1	-13.675
sin	12	3.738	1	-13.675
actuar	11	3.738	1	-13.675
quiere	11	3.738	1	-13.675
hay	11	3.738	1	-13.675
supuesto	11	3.738	1	-13.675
qué hago	11	3.738	1	7.625
hecho	10	3.738	1	-13.675
a	10	3.738	1	-13.675
otro	10	3.738	1	-13.675

Figure 2.27: Screenshot of term extraction based on the on-duty corpus using TextToOnto

TextToOnto is also useful because at the engineer's request the relevant terms (in Fig. 2.27 are selected in blue) can be transferred to the Oi-Modeler and visualized. In Fig. 2.28 we can see the most relevant terms for the on-duty competency questions corpus. From this graph, concepts such as *hecho* = *supuesto* [fact], *acto* (from *actuar*) [act], *proceso* (from *instrucción*) [process] and *guardia* [on-duty] can be easily identified and confirm the previous work done in the domain ontology used in *Iuriservice I.*

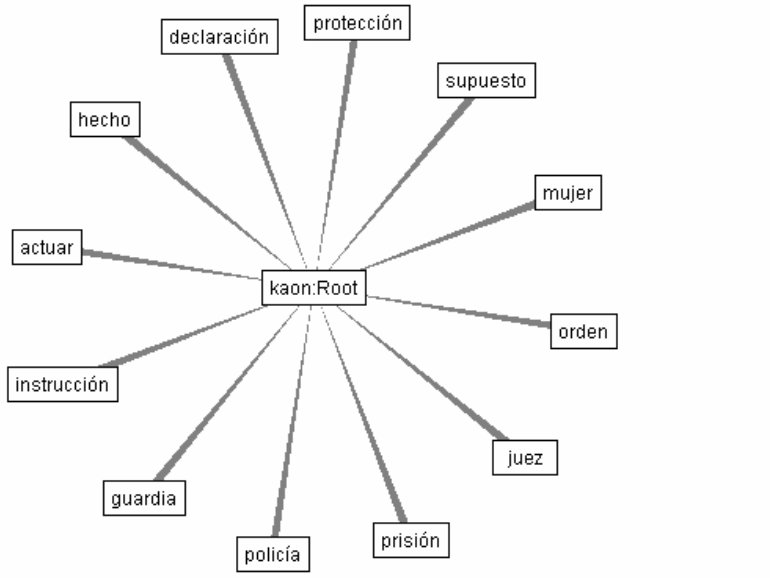


Figure 2.28: Term extraction from on-duty corpus into transferred into Oi-Modeler's graph.

Once the analysis was performed on a corpus based on all the reminder of the 600 competency questions (not on the on-duty questions), the term extraction performed by TextToOnto and visualized by the graph in Fig. 2.29 offered the following results:

- First, it confirms that the concepts inferred from the on duty corpus (*hecho* = *supuesto* [fact], *acto* -from *actuar*- [act], *proceso* -from i*nstrucción* and *proce*(-*imiento*)- [process] and *guardia* [on duty]) remain significant and should be included in the domain ontology.
- Secondly, some concepts appear to have clear logical relations between them, and could be organized into ontological subdomains. For example, *policía* [policeman,], *juez* [judge], *mujer* [woman], *testigo* [witness], *procurador* [court representative], *forense* [forensic doctor], *señor* [man, mister], could be considered Roles that agents might play when participating in a judicial process. In the same line, *orden* [order], *resolución* [resolution], *auto* [type of judgment] and *demanda* [civil/public lawsuit] could be considered procedural documents.

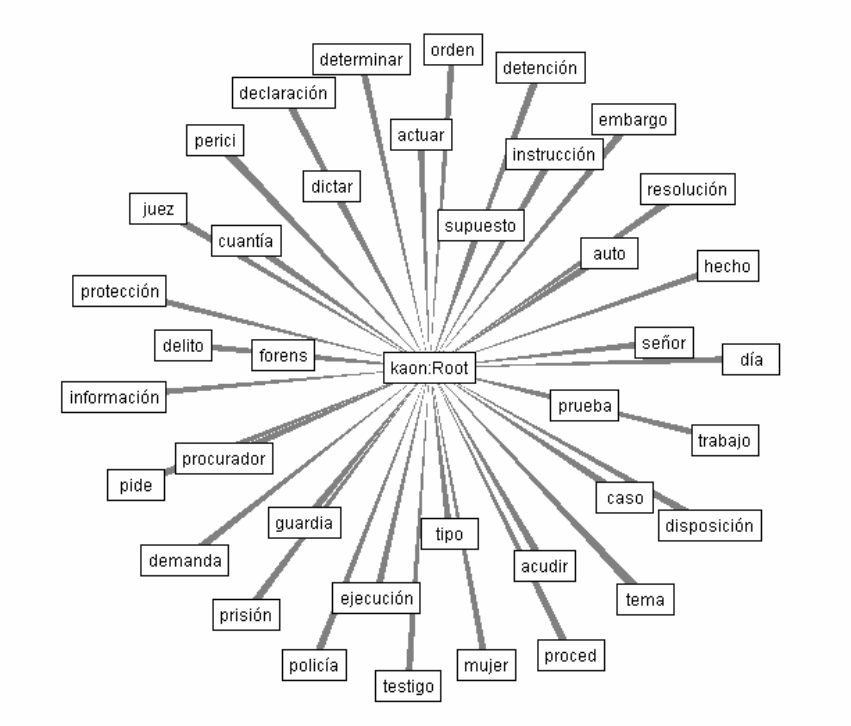


Figure 2.29: Term extraction graph from the corpus containing all questions (not on duty)

These ontological subdomains inferred intuitively from the graphs represented above, are further explored and analyzed by ALCESTE (below) leading to interesting findings.

2.4.2.3.2 Textual statistics distribution (ALCESTE)

As it was said in the introduction to this chapter, we observed that in order to acquire some rich information about judges' expert knowledge it would be useful to use a textual statistics program. This program would provide us with the significant structure of judges' discourse so that we could propose basic hypothesis and heuristics for the construction of the ontology. This section is dedicated to the description of the method we have used, the description of the corpora we have created and analyzed and the presentation of some preliminary results.

a) The method

ALCESTE stands for *Analyse des Lexèmes Co-occurents dans les Énnoncés Simples d'un TExt* [Analysis of the co-occurring lexemes within the simple statements of a text] and its algorithm was created by Max Reinert at the CNRS —partly based on Bénzecri's important contributions in textual statistics [Be, 1982]— with the support of ANVAR. The company IMAGE has formatted this software program to be compatible with Macintosh and PC formats.

As it is commonly known, textual data analysis (or textual statistics analysis) is a method by which we acquire the essential information from a text. This text can be presented in many formats (interview trancripts, answers to open questions of different

kinds of surveys, collections of texts, etc.). The goal of ALCESTE is to quantify the text in order to extract its most significant structures. These identification and extraction processess are based on a general principle concerning discourse: the discourse is not analyzed depending on its representations but according to the activity that takes place in it. Reinert terms this activity *repetition* [R, 2003]. *Repetition* is supposed to be the main factor for the stabilization of a discursive activity.

ALCESTE represents statistically repetition in many ways (contingent tables, chi-square analysis, etc.). The program does a particular analysis of the "topography of a discourse" by creating, confronting and representing different *lexical worlds*. In other words, it is assumed that the most significant structures of a text are deeply linked to the distribution of words within this same text, and this distribution is rarely a matter of chance. Thus we can model the trace of the meaning of a text as the trace of a discourse activity (production and repetition of signs) (B, 2002).

The method used by ALCESTE is known as *Hierarchical Decreasing Classification* (HDC) The corpus under analysis is successively split up in text segments; then the program lists the distribution of the *most significant words* within each segment; finally it extracts the most representative words from the text.

It has to be said that a *most significant word* is not a lexeme and that a segment does not need to be a complete (or full-meaning) statement or sentence. As a program of distributional statistical analysis, its working mechanisms are independent from meaning. This means that there is no need to have precise units to represent a concrete object [R, 2003]. What really matters is the presence or absence of words. The co-occurrence of *most significant words* translates the "associative background" (or topic background) that operates within the text: the proximity (temporal in the statement or spatial in the text) of the *most significant words* [B, 2002].

The program, then, classifies the segments of the corpus (called *elemental context units* [ECU]) according to the distribution of the vocabulary which appears within these context units. The program finds the vocabulary in the different context units and relates them. In other words, it connects those contexts having common words, finds the strongest vocabulary oppositions and extracts some categories of representative statements.

For the relevance of the statistical distribution, the volume of the corpus has to be large enough, and ALCESTE also requires that the text under analysis has a certain degree of thematic coherence.

How it works

As already said, ALCESTE breaks down the text into different segments or context units, since it recognizes the forms in the context units. (Each series of character sets inside of the limiting characters is an occurrence; identical occurrences are forms.)

We can distinguish between two kinds of contexts:

1. Initial Context Units (ICU). They are the different parts of the corpus and the biggest context unities. They may be texts (e.g. magazine articles) or different parts of the same text (e.g. book chapters, different answers to open questions, etc.). Variables attributed to these text segments enable to cross them quickly. These variables (expressed either by single words or sentences) can be tagged with asterisks [*] so that the program knows that tagged words or sentences will be considered out of the corpus. This first fragmentation is not compulsory: one can treat a corpus without tagging any variable before the program analysis.

2. Elemental Context Units (ECU). The second fragmentation of the corpus is made by ALCESTE defining the "sentences" according to which it will perform the analysis. This fragmentation is eventually based upon punctuation —if any— and the number of words.

Another important type of unit for the program concerns the lexical unit. The program identifies the occurrences of every form through a dictionary. As an example, the figure below shows a vocabulary classification of one of the seven classes created by ALCESTE for one of our protocols.[18]

Chi-2	ECU in the class	Formes réduites	Formes complètes
47.88	22	tratos [treatments]	tratos(22)
44.11	22	malo+ [ill]	malos(22)
37.11	8	marido [husband]	marido(10)
27.45	6	tratador [treater]	tratador(6)
22.52	14	mujer+ [woman]	mujer(14)
19.60	14	proteccion[protection]	proteccion(16)
18.53	10	senor+ [man]	senor(12)
			senora(10)
18.05	4	trabaj+ [work]	trabajaba(2)
			trabajar(2)
18.05	4	insul+ [insult]	insulta(2)
			insultos(2)
18.05	4	ido [gone]	ido(4)
18.05	4	empresa [company]	empresa(8)
18.05	4	cautelar+[preventive]	cautelar(2)
			cautelares(2)
17.83	6	viviend+ [house]	vivienda(4)
			viviendo(6)
17.83	6	tema+ [topic]	tema(2)
			temas(4)
17.83	6	psicologico+ [psychological]	
			psicologico(2)
			psicologicos(4)

Figure 2.30: Representative forms of Class number 1 for questions posed by judges concerning problems they have to face during their *on duty* period (Corpus 2)

The sign [+] indicates reduction to the root of a word. For instance, in Fig. 2.30 we see that the root 'trabaj+' (work) (in the column *Formes reduites*) comprises two kinds of complete words in the column *Formes complètes*. This means that every word that appears in this class pertaining to the same family counts as an occurrence of this form, while the column in the right shows the different forms that the word takes in the text segment. This process is known as *lemmatization*. Following this example, the forms

[18] The number in parenthesis after a form means its frequency in a class.

‘trabaja’ [works], ‘trabajó’ [worked] and ‘trabajando’ [working] would be reduced to ‘trabaj+’ [work+]. However, the column on the right tells us that the only two forms which appear in this part of the corpus are ‘trabajaba’ [worked] (twice) and ‘trabajar’ [to work] (twice).

From this partition of the context units and forms, the corpus is modeled through a table of data that comprises the chunks in lines and the most significant words in columns. Then different classifications are performed. ALCESTE crosses the context units and the presence/absence of forms and it further creates the classes according to the context units containing the same words. In an iterative way, it changes the number of words for every context unit, compares the obtained classes and keeps the classes that are associated with the greatest number of context units.

Finally we obtain —through the Decreasing Hierarchical Classification— a certain number of coherent classes of words representative of the analyzed text. These classes present the same type of “symbolic repetition”. The program thereby shows the main “lexical worlds” of the corpus [R, 2000, 2003], that is to say, the sets of words that are most particularly associated to a class.

The steps of analysis

ALCESTE performs the analysis of the corpus in 4 steps which are subdivided in some other operations (the “analysis plan”).

Step 1. Firstly, the program recognizes the ICU (the context units given by the analyst) and the tagged words or sentences (variables), if any. After that, it carries out three successive treatments:

a) it cuts down the corpus into forms;
b) it performs a syntactical categorization (*mots-outils*);
c) it makes a lemmatization (*mots-pleins*).

Fig. 2.31 and Fig. 2.32 show some results of this first step.

```
General information

Name of the corpus                          Guardia-nuevas.txt
Name of the analysis plan                   Nuevas.pl

Vocabulary analysis

Number of initial context units (UCI) 1
Number of occurrences                       8970
Number of different forms                   1035
Average Frequency per form                  9
Number of nonce/hapax words                 0

After the lemmatization

Number of reduced forms                     295
Number of supplementary words               95
Number of elemental context unities (ECU)   210
```

Figure 2.31: Basic information provided by ALCESTE for the Class 1 of our Corpus 2

Grammatical category	Indicator
verbos modales [modal verbs]	Supplémentaire
preposiciones [prepositions]	Supplémentaire
conjunciones [conjunctions]	Supplémentaire
interjecciones [interjections]	Supplémentaire
articulos [articles]	Supplémentaire
pronombres personales [personal pronouns]	Supplémentaire
ponombres posesivos [possessive pronouns]	Supplémentaire
pronombres demostrativos [demonstrative pron.]	Supplémentaire
pronombres relativos [relative pron.]	Supplémentaire
conceptos que expresan cantidad [concepts of quantity]	Supplémentaire
numerales y numeros [numerals and numbers]	Supplémentaire
adverbios de lugar [location adverbs]	Supplémentaire
adverbios de tiempo [time adverbs]	Supplémentaire
adverbios de modo [mode adverbs]	Supplémentaire
adverbios de duda [doubt adverbs]	Supplémentaire
negaciones [negations]	Supplémentaire
auxiliares [auxiliaries]	Supplémentaire
Locuciones [idioms]	Supplémentaire
Numeros en cifras [numbers]	Eliminée
Palabras mayusculas [words in capital letters]	Supplémentaire
Palabras no encontradas en DICIN [words not found in the dict.]	Eliminée
Verbos [verbs]	Analysée
Nombres [nouns]	Supplémentaire
Palabras no reconocidas y frecuentes [not found but frequent words]	Supplémentaire
Palabras no reconocidas [not recognized words]	Analysée

Figure 2.32: Syntactical categorization for the Class 1 of our Corpus 2

Step 2. The second step comprises the fragmentation of the corpus in Elemental Context Units (ECU) and the conformation of the classes according to their distribution. ALCESTE thus constitutes the classes on the basis of the lexical content of each ECU: it assembles the ECU which contain the same lexical forms.

Step 3. In this step, the results of the former steps are showed in some files: the obtained classes, the most frequent forms of every class, etc.

Step 4. During this last step, some complementary operations are performed:

a) two kinds of three-crossing (one part of the text is crossed with a variable or a particular word);
b) a factorial analysis of correspondences (crossing of the vocabulary and the classes) that can be seen with the help of useful graphical representations;
c) a hierarchical increasing classification (HIC) that shows the more or less close links that words have among each other.

These analysis help to the interpretation of the statistical results and to the description of classes.

b) Description of the corpora

We have considered three different text corpora:

a) Scholar "on duty" questions (**Corpus 1** = 99 questions). This corpus is the result of the interaction between a very experienced judge (Magistrate Antonio Doñate) and their students of the Spanish Judicial School. In this exercise, the judge and the students tried to ask and answer questions they believed reflected typical problems during the *on duty* period of their colleagues (newly recruited judges in their first appointment). That's why we call these corpus "Scholar *on duty* questions". Obvioulsly 99 three-line questions do not constitute a big corpus. Therefore, we doubled each question in the corpus without producing any change on its content (in terms of presence/absence of words, etc.).

b) Practical "on duty" questions (**Corpus 2** = 163 questions). After the ethnographic campaign, UAB researchers extracted questions from the questionnaires and the literal transcripts. These questions were posed by the judges when asked for the main problems during their actual *on duty* period (which comprises, between others, most questions about gender violence). 163 questions still do not constitute a big corpus for ALCESTE, so we performed the same doubling operation.

c) All practical questions (**Corpus 3** = 756 questions). This is the whole group of questions extracted from the questionnaires and the literal transcripts of the interviews. This is a bigger corpus, therefore no change has been made.

c) Preliminary results

Corpus 1: Scholar "on duty" questions

A double HDC —Hierarchical Decreasing Classification— has been performed. The volume of the analyzed elemental context units (ECU) changes (13 words in the first classification and 15 in the second). On 196 defined ECU, 136 have been associated to a class, which represents a 69.39% of the ECU. As shown in Fig. 2.33, 2.34 and 2.35, the program identified 7 classes and performed a factorial analysis of correspondences (crossing the vocabulary and the classes) that is represented graphically.

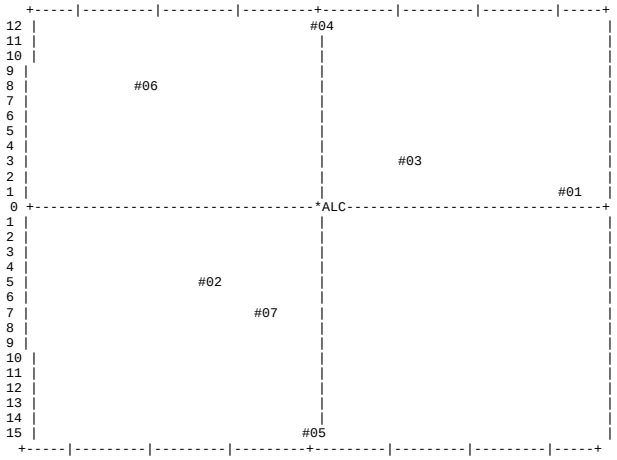


Figure 2.33: Class projection of Corpus 1

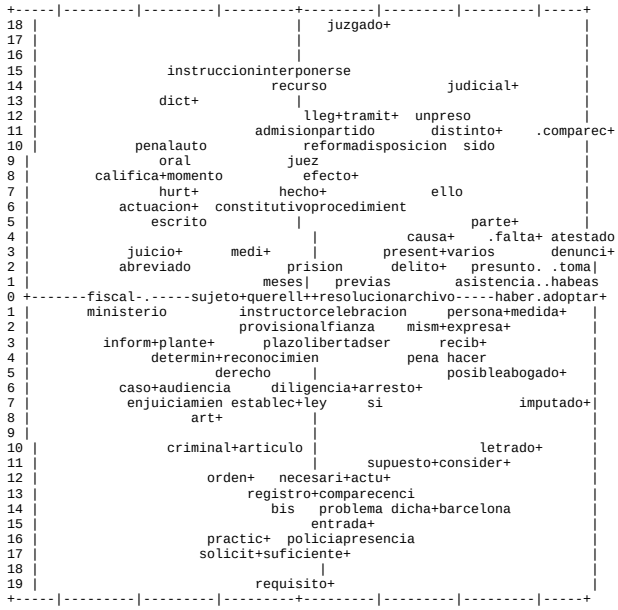


Figure 2.34: Word projection of Corpus 1

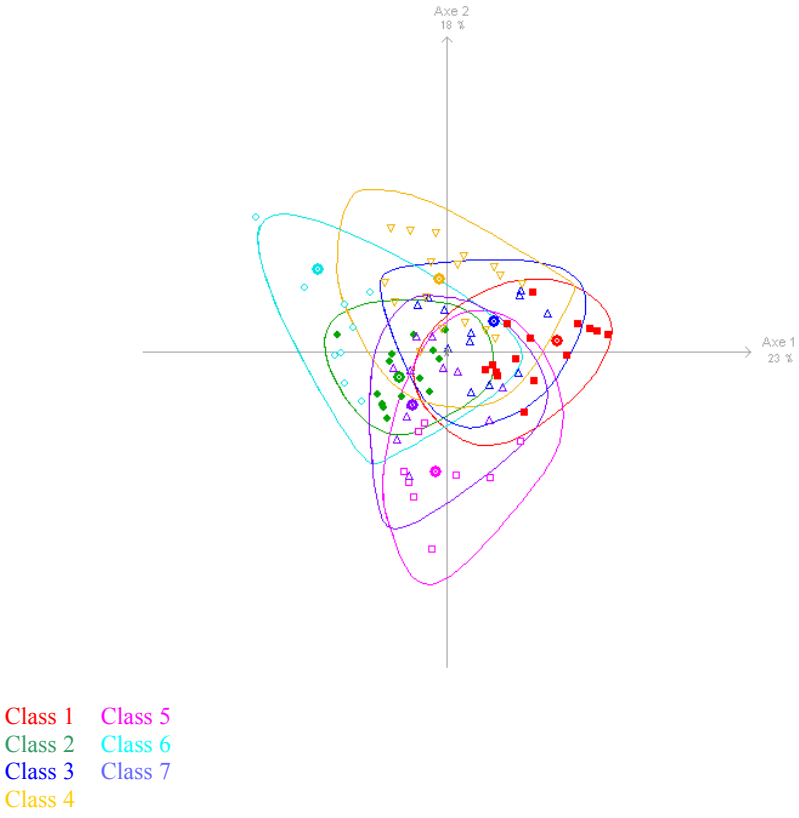


Figure 2.35: Factorial plan of Corpus 1

Observing the results for the Corpus 1, our interpretation of the meaning of X and Y axis is that axis X represents the linear dimension of legal proceeding: it starts with a law suit (Class 1 in 2.33 and 2.35) and ends up with the qualification of the facts (Class 6). This represented process involves an important physical and practical element, which is represented by words related to the physical locations of the process and

judicial acts.

With respect to the axis Y, we interpret it again as a gradual signification of the legal process but in a more abstract way. Classes 5 and 4 are strongly opposed in the graphic: they represent the dimension of the process from the (hypothetical) facts to the legal qualification of these facts. On the one hand, Class 5 implies the representation of the proceedings related to external elements to the Court (evidence elements and facts) while class 4, on the other hand, corresponds to decisions made in the initial steps of the proceedings (internal decisions). It has an abstract element: the steps of the process (facts, demand, legal qualification, judgment).

According to this interpretation, each class has been labeled with reference to the contained vocabulary and to its place within the graphical representation of the factorial analysis.

Class 1: INITIAL STEPS OF THE PROCEEDINGS (ON DUTY)[19]
declaracion+(14), adoptar(8), asistencia(6), cuenta(6), medida+(8), personal+(6), toma(8), abogado+(8), buque(4), denunci+(8), detenido+(10), habida(4), lesiones(4), letrado+(10), oficio(4), retras+(4), asist+(2), atestado(4), caracter(2), casa+(2), caus+(2), corpus(2), excepcion+(2), exist+(4), forma(2), guardia(8), habeas(2), haber(2), hacer(6), halla(2), hora+(4), malos(2), nacional+(4), persona+(6), precisa+(2), presunta+(2), presunto(2), renunci+(2), tipo(2), trafico(2), tratos(2), varias(2), parte+(4), posibilidad(4), posible(4), amotin+(2), bordo(2), contigua(2), correct+(2), disciplinaria(2), expresa+(2), fuerza+(2), imputado+(6), mar(2), punto+(2), resolver(2), territorial(2), zona(2);

Figure 2.36: Specific vocabulary of class 1

According to this, we have labeled class 1 with the title “Initial steps of the proceedings (on duty)”. This class represents 20.59% of the classed speech segments. Fig. 2.36 shows its specific vocabulary. It represents the physical and practical elements of the initial steps of a proceeding during the on duty period. The most frequent (and representative) words of this class are deposition [*declaración*], arrested [*detenido*], lawyer [*letrado*], to adopt [*adoptar*], suit [*denunci+*] and on duty [*guardia*].

Class 2: ELEMENTS OF CRIMINAL PROCEDURE AND CRIMINAL LAW
criminal+(16), enjuiciamiento(14), ley(16), art+(10), instructor(8), determin+(4), establec+(6), hecha(4), investigacion(4), junto+(4), lugar+(4), medi+(8), poder(4), procurador(4), public+(4), sobreseimiento(4), sujeto+(6), abreviado(4), admitirse(2), aport+(2), autoriz+(2), codigo(2), derecho(2), entenderse(2), especial(2), estim+(2), inform+(2), limites(2), objeto(4), practiqu+(2), reconocimiento(2), responsabilidad+(2), telefonica+(2), validez(2), ya(4), diligencia+(6), necesari+(6), procedimiento(8), caso+(6), querell+(6);

Figure 2.37: Specific vocabulary of class 2

[19] It has to be noted that in Corpus 1 and 2 the number of word occurrences —number between parenthesis— are doubled because we doubled the two corpora.

Class 2 has been labeled “Elements of criminal procedure and criminal law”. It represents 22.06% of the classified ECU and it has to be considered the most representative class, although there is not an important percentual difference between Class 2 and 1. Class 2 is quite opposed to Class 1 in the X axis but not in the Y axis. This lexical world (2.37) represents the final steps of a legal process (X axis) with some relation to procedural elements. Its most representative words are criminal [*criminal*+], adjudication [*enjuiciamiento*+], law [*ley*+], article [*art*+], instructor [*instructor*], establish [*establec*+], mean [*medi*+], proceeding [*procedimiento*], place [*lugar*+], etc.

Class 3: DOUBTS AND DECISIONS ABOUT THE FACTS OF THE CASE
dias(4), dinero(4), ello(6), hecho+(10), mism+(8), admisible(2), carecen(2), conocimiento(2), demas(2), falta+(4), manifiestamente(2), policial(2), posterior+(4), previas(4), relativa+(2), sede(2), ser(4), si(8), unic+(2), varios(4), archivo(4), constitutivo+(4), recib+(4), base(2), conoce(2), contaminacion(2), delito+(6), inici+(2), jueces(2), manera(2), negarse(2), ocurre(2), principio(2), procesal+(2), secretario(2), trata(2), tribunal(2), comparec+(2), efecto+(2);

Figure 2.38: Specific vocabulary of class 3

Class 3 represents 13.24% of the classed ECU. We have called it “Doubts and decisions about the facts of the case” for it is located after the first step of the process but before the final actions (X axis), and it corresponds to an internal dimension of the process (Y axis) with respect to the judge himself (doubts and decisions). As it may be observed in Fig. 2.38, its most representative words express a hesitant-like moment of a process (constituting [*constitutivo*+], to be [*ser*], fact [*hecho*+], crime [*delito*+], previous [*previas*+], etc.

Class 4: STARTING OF THE PROCEEDINGS (LOCATION JUDGE)
recurso(12), reforma(12), auto(14), admision(4), admitiendola(4), correspondiente(4), dict+(10), habiendose(4), interponerse(6), interpon+(4), aforado(2), ano+(2), aparec+(2), cliente(2), condicion(2), const+(2), dicho(2), disposicion(2), instruccion(8), juez(12), juzgado+(10), lleg+(2), notific+(2), partido(2), prescripcion(2), presentacion(2), preso(4), puesto+(2), tiempo(2), tramit+(4), un(14), judicial+(4), present+(8), resolucion(4), causa+(4), admit+(2), distinto+(2), motiv+(2), periodo(2), produc+(2), sido(2), testigo+(2);

Figure 2.39: Specific vocabulary of class 4

Class 4 is an extreme case with respect to the Y axis, but stands in the middle of the linear way that represents the X axis. It represents 14.71% of the classed ECU. As shown in Fig. 2.39, it covers concrete location (court [*juzgado*], judicial district [*partido*], the intermediate steps of a process (appeal [recurso+], reform [*reforma*+], interlocutory decision [*auto*+], admission [*admisión*], decide [*dict*+], lodge [*interpon*+], proceedings [*instrucción*], judge [*juez*], etc.).

Class 5: GATHERING OF EVIDENCE
dicha+(6), entrada+(6), registro+(6), formal+(4), policia(6), practicarla+(4), provincia(4), ampliatoria+(2), barcelona(4), considerarse(2), contenido(2), cualquier+(2), domicilio(4), encuentra(4), frecuencia(2), intercept+(2), numero(2), pericial(2), presencia(4), problema(2), prueba(2), sustancia(2), actu+(4), requisito+(4), solicit+(6), alegan(2), colegio(2), competente+(2), concurr+(2), consider+(2), implic+(2), interven+(2), madr+(2), practic+(2), respecto(2), solicitud(2), suficiente+(2), vist+(2), entender(2), llev+(2);

Figure 2.40: Specific vocabulary of class 5

Class 5 represents 13.24% of the classified textual segments and as has been said before, it is located in the opposite side of Class 4. While Class 4 evokes the internal elements of the process (the court, the judge…), Class 5 brings us outside the court and shows the external actors of the process and their most typical moves. As shown in Fig. 2.40, its most frequent words are entry [*entrada*+], search or register [*registro*+], police [*policía*], house [*domicilio*], finds [*encuentra*], act [*actu*+], requisite [*requisito*+], ask for [*solicit*+], proof or evidence [*prueba*], etc.

Class 6: QUALIFICATION OF THE FACTS
apertura(4), califica+(6), hurt+(4), oral(6), robo(4), actuacion+(6), acusacion(4), remit+(4), escrito(4), fiscal(6), juicio+(6), ministerio(6), momento(2), organ+(2), provincial(4), anterior+(2), audiencia(4), defensa(2), formular(2), intimidacion(2), jurisdiccional(2), nulidad(2), particular(2), penal(4), peticion(2), plante+(2), violencia(2);

Figure 2.41: Specific vocabulary of class 6

Class 6 is the least representative of all (7.35% of the ECU). It brings us back to the internal space of the proceedings (with words like trial [*juicio*+], ministry [*ministerio*], court [*audiencia*], prosecution [*acusación*], defense [*defensa*], see Fig. 2.41). It represents the final steps of a legal process: the legal qualification of the facts. We interpret in this sense the appearance of words such as: qualify [*califica*+], theft [*hurt*+], oral [*oral*], theft [*robo*], intimidation [*intimidación*], and violence [*violencia*].

Class 7: INTERLOCUTORY DECISIONS
comparecencia(8), prision(12), provisional(8), decretar(4), celebracion(6), acusadores(2), arresto+(2), articulo(6), bis(4), captura(4), imprescindible(2), mayor(2), pidan(2), plazo(4), fianza(4), libertad(4), supuesto+(4), dado(2), detencion(2), meses(2), original+(2), pena(2), prevista+(2), reo(2), sentencia(2), situacion(2), termino+(2), acordad+(2), realiz+(2), orden+(2);

Figure 2.42: Specific vocabulary of class 7

Finally, Class 7 is a small class as well, for it represents 8.82% of the ECU. It is located near Class 2 in the factorial plan. It represents a very concrete and central aspect —also graphically— of the legal process: judicial interlocutory decisions. As shown in Fig. 2.42, some of its most representative words are appearance [*comparecencia*], prison

[*prisión*], provisional [*provisional*], order [*decretar*], celebration [*celebración*], capture [*captura*], bail [*fianza*], release [*libertad*], accused [*reo*], judgment [*sentencia*]…

d) Corpus 2: Practical "on duty" questions

The number of classed ECU in the second HDC is 180 out of 210, which represents 85.71%. The program has identified 7 classes which are spatially represented in Fig. 2.43 (see the word distribution in Fig. 2.44 below).

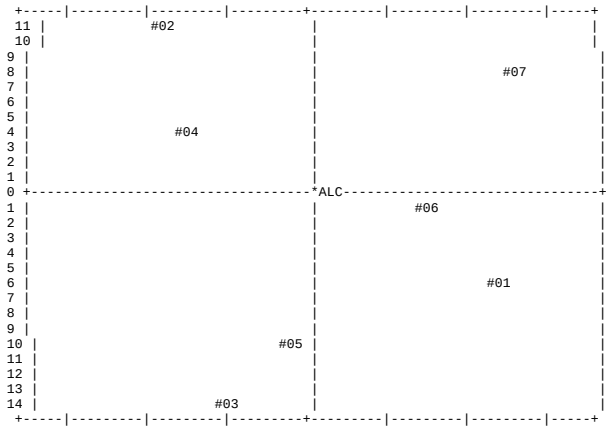


Figure 2.43: Class projection of Corpus 2

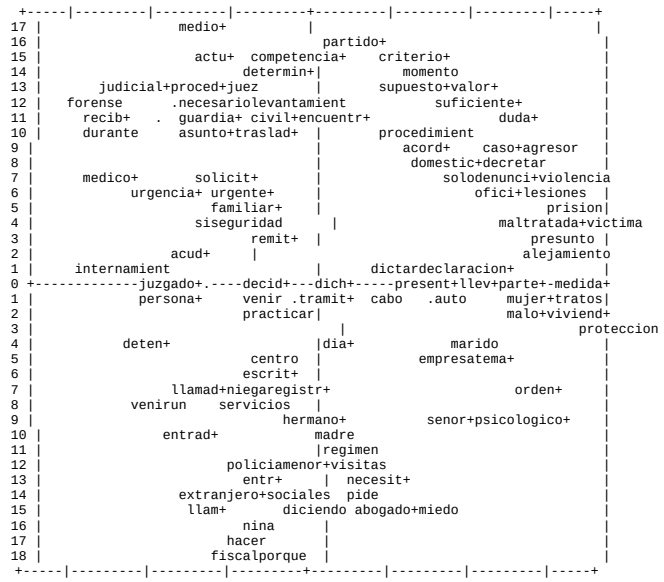
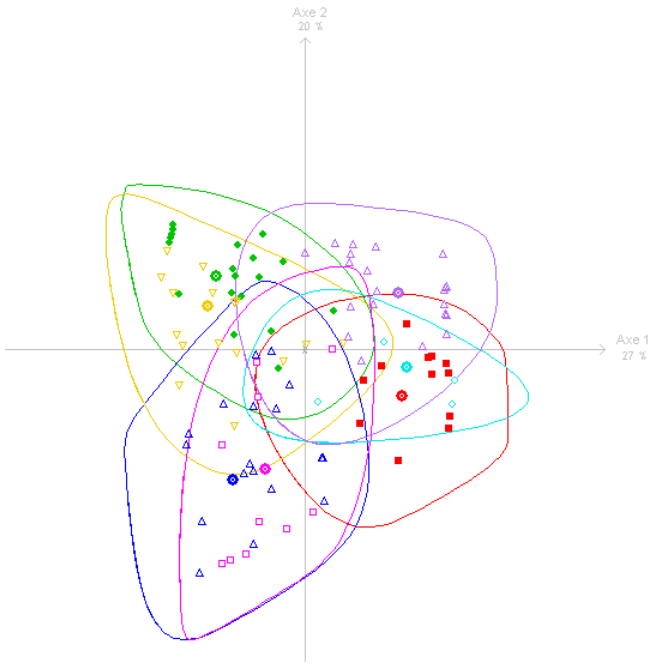


Figure 2.44: Word projection of Corpus 2



Class 1 Class 5
Class 2 Class 6
Class 3 Class 7
Class 4

Figure 2.45: Factorial plan of Corpus 2

From Fig. 2.43 to Fig. 2.45, the X axis represents the abstract phases of the judicial process (from the description of the facts to the legal measures and qualification). Furthermore, X axis seems to express a graded and nuanced progress between two important fields: legal system (left side) and family (right side).

Y axis, on the other hand, represents those phases of the legal process which have more in common with physical elements and actors of the judicial process. However, as it will be shown below, all the classes of this corpus (because of its origins) express practical problems.

Class 1: GENDER VIOLENCE (DENOUNCE AND INMEDIATE MEASURES)
tratos(20), malo+(20), senor+(14), tratador(6), cautelar+(4), empresa(4), ido(4), insul+(4), mujer+(14), proteccion(14), psicologico+(6), tema+(6), trabaj+(4), casa(4), orden+(20), pedido(4), bebe(2), comparecencia(4), dejar(2), explotacion(2), pega(2), pida(2), presunto(6), tomarse(2), viviend+(4), dar(2), dispone(2), habido(2), peticion(2), previa+(2), psiquiatr+(2);

Figure 2.46: Specific vocabulary of Class 1

Class 1 represents 16.67% of the classed ECU and expresses the first step of a gender violence case (suit and immediate measures). Its most representative words (Fig. 2.46) also show some of the principal actors and elements of this kind of cases: molestation [*malos tratos*], man [*señor*+], molester [*maltratador*], insult [*insult*+], woman [*mujer*+],

protection [*protección*], psychological [*psicológico+*], order [*orden*], house [*casa*]…

Class 2: CRIMINAL PROCEDURE ACTIONS (ON DUTY)
cadaver(22), juez(24), levantamiento+(18), guardia+(24), autopsia+(6), encuentr+(6), estando(6), fuerza+(6), funcionario+(6), proced+(12), si(22), telefonica+(6), traslad+(6), urgencia+(8), acta(4), criminal+(4), exig+(4), finaliz+(4), forense(8), intervencion+(4), judicial+(10), necesaria+(4), necesario(6), necesidad(4), plantea(4), pregunt+(4), reconocimiento(4), requisitori+(4), respond+(4), secretario(4), actuacion+(4), acud+(6), civil+(6), competencia+(4), conduc+(2), determin+(4), esquizofren+(2), firma(2), grave+(4), hacerse(4), ley(2), libertad(2), lugar(4), mental+(4), muerte(2), penal(4), recib+(4), situacion(2), solicitud(4), solicit+(6), violent+(4), un(28), actu+(8), medio+(4), acord+(6), in(2), indicios(2), materia(2), seguir(2), turno(2);

Figure 2.47: Specific vocabulary of Class 2

Class 2 is strongly opposed to Class 1 (see Fig. 2.47 above) in relation to the field in which the situations expressed are located in. It represents 21.11% of the classified segments. While Class 1 expresses the outsider elements of a gender violence case, Class 2 represents those actions of procedural law which have to be typically performed during the *on duty* period. In this sense, some of its most representative vocabulary is quite illustrative: corpse [*cadáver*], judge [*juez*], removal (of the corpse) [*levantamiento+*],[20] on duty [*guardia*], autopsy [*autopsia+*], civil servant [*funcionario+*], proceed [*proced+*], removal [*traslad+*], urgency [*urgencia+*], criminal [*criminal*], necessity [*necesidad*]…

Class 3: PROBLEM MANAGEMENT OF INCOMING CASES DURING "ON DUTY" PERIOD (INMIGRATION, MINORS…)
porque(14), entr+(6), llam+(10), sociales(6), policia(14), cargo(4), gas(4), piso(4), sola+(4), diciendo(6), entrad+(6), llamad+(6), otorg+(4), pide(10), registr+(6), servicios(6), telefono+(4), vive(4), abogado+(4), miedo(4), necesit+(4), niega(4), fiscal(4), persona+(8), acceder(2), asist+(2), atender(2), ayud+(2), cerrado+(2), mand+(2), nino+(2), profesional+(2), salir(2), sujeto(2), tenido(2), venido(2), alguien(2), decid+(2), hacer(12);

Figure 2.48: Specific vocabulary of Class 3

Class 3 represents 13.33% of the ECU and has to do with problem management of incoming cases during the *on duty* period. It is strongly opposed to Class 2 in the sense that Class 3 expresses an activity which constitutes the very first step of a legal process during the *on duty* period (the income of cases) and shows us the typical problems the judge has to face in relation to those situations (minors, etc.). As shown in 2.48, some of its most representative words are call [*llam+*], social [*social+*], services [*servicios*], police [*policía*], entry [*entrad+*], flat [*piso*], demands [*pide+*], register [*registr+*], lawyer [*abogado*], fear [*miedo*], need [*necesit+*], person [*persona+*], help [*ayud+*], etc.

[20] In Spanish Criminal law, the judge or the forensic doctor control and certificate the process of retiring a dead body from the place where it has been found before it is brought to the morgue. This action is known in Spanish as 'to remove the corpse'.

Class 4: URGENT DECISIONS
urgente+(10), competente+(6), juzgado+(12), pidiendo(4), razones(4), remit+(6), seguridad(4), tarde(4), tramit+(6), venir(6), dich+(6), internamiento+(10), medico+(10), familiar+(4), permanente(2), rapidos(2), abiert+(2), activ+(2), alcoholemia(2), anterior(2), asunto+(2), autorizacion(2), cuenta(2), deten+(4), materiales(2), piden(2), poblacion(2), prueba(2), publica(2), salg+(2), venir(4), vez(2), durante(4), dictar(4), consider+(2), juicio+(2), ser(2);

Figure 2.49: Specific vocabulary of Class 4

Class 4 has been labeled "Urgent decisions" and represents 10.00% of the ECU. Fig. 2.49 shows its most representative vocabulary: urgent [*urgente*+], competent [*competente*+], court [*juzgado*+], security [*seguridad*], internment [*internamiento*+], doctor [*médico*], blood-level of alcohol [*alcoholemia*], etc.. According to its location at the factorial plan, it refers to the actors, to the appropiate judicial space and also to the typical urgent decisions the judge has to make during the *on duty* period.

Class 5: PROBLEM RAISING IN FAMILY LAW DURING ON DUTY PERIOD
madre(8), menor+(10), aportado(4), cierto(4), donacion(4), exploracion(4), tomar(4), demand+(4), escrit+(4), extranjero+(4), regimen(4), visitas(4), autoridades(2), compraventa(2), entrega+(2), herencia(2), hermano+(2), mismo(2), ostenta(2), permit+(2), progenitor(2), puest+(2), saber(2), voluntad(2), cambi+(2), centro(4), conocimiento(2), discute(2), edad(2), ingreso(2), nina(2), practicar(4), present+(4), trastorno(2), informacion(2), posible+(2), resulta(2);

Figure 2.50: Specific vocabulary of Class 5

Class 5 is as much representative as class 4 (10% of the classed ECU), but it is opposed to class 4 with respect to the temporal stage of the judicial processes regarding Family Law cases. This class offers some information about the private actors of a Family case (Fig. 2.50): mother [*madre*], minor [*menor*], brother [*hermano*], girl [*niña*], foreigner [*extranjero*], exploration [*exploración*], visiting [*visitas*], alienation [*transtorno*], legacy [*herencia*], admission [*ingreso*].

Class 6: GENDER VIOLENCE (DESCRIPTION OF FACTS)
cabo(6), dictad+(4), llev+(6), retir+(4), concreta+(2), dia+(4), marido(4), adoptar(2), detiene(2), diligencia+(2), estar(2), libr+(2), permiso(2), proteger(2), sabado(2), supone(2), vuelv+(2), ya(2), algun(2), limites(2), pareja(2), problema(2);

Figure 2.51: Specific vocabulary of Class 6

Class 6 represents 6.67% of the ECU and, like Class 1, it represents gender violence from the point of view of the description of the facts. Its most representative words: decided [dictad+], retire [retir+], day [día+], husband [*marido*+], adopt [*adoptar*], arrests [*detiene*], protect [*proteger*], Saturday [*sábado*], couple [*pareja*], problem [*problema*].

Class 7: GENDER VIOLENCE (CRIMINAL AND PROCEDURE TREATMENT)
domestic+(16), violencia(16), agresor(10), denunci+(14), abus+(6), alejamiento(18), caso+(24), lesiones(8), prision(10), supuesta+(6), adopcion(4), existe(4), facilit+(4), gravedad(4), imputado(4), ofici+(6), sentido(4), separacion(4), suficiente+(6), supuesto+(8), tambien(4), version+(4), victima(10), abusos(2), acordarse(2), amenazas(4), apreci+(4), asistencia(4), auto(4), conoc+(4), criterio+(6), declaracion+(8), decretar(2), delito+(4), derecho+(2), distinto+(2), hecho+(4), horas(2), inmediat+(2), maltratada+(6), medida+(10), misma+(4), momento(4), parte+(8), peligro(2), personal+(2), preventiva(2), procedimiento(4), sexuales(2), sido(2), sospecha(4), valor+(8), viven(2), duda+(4), partido+(4), solo(4);

Figure 2.52: Specific vocabulary of Class 7

Finally, class 7 represents a third point of view about gender violence: its treatment from Criminal Procedural Law. This class represents 22.22% of the classed ECU and helps us to interpret the main problems the judges have with this kind of cases, represented by classes 1, 6 and 7 (placed at the right side of the factorial plan). In this sense, on the one hand, judges seem to be more concerned about practical problems that take place during the first stages of this type of criminal case (case entry and first measures) and, on the other hand, during the proceedings. Class 6 (the least representative of all classes) shows that the description of the facts is probably not a big problem for judges, while the real practical problems come when decisions have to be made. Returning to Class 7, its most representative words show some doubts and problems about the legal treatment that has to be implemented (domestic [*doméstic*+], violence [*violencia*], agressor [*agresor*], denounce [*denuncia*], abuse [*abus*+], injuries [*lesiones*], case [*caso*], supposed [*supuesta*+], deposition [*declaración*], interlocutory decision [*auto*], etc.

Corpus 3: All practical questions

With respect to the third corpus, 431 ECU out of 500 have been put in classes, which represents 86.20%. Four stable classes have been identified and represented, as shown in Fig. 2.53 and Fig. 2.54. Fig. 2.53bis shows the word distribution of Corpus 3.

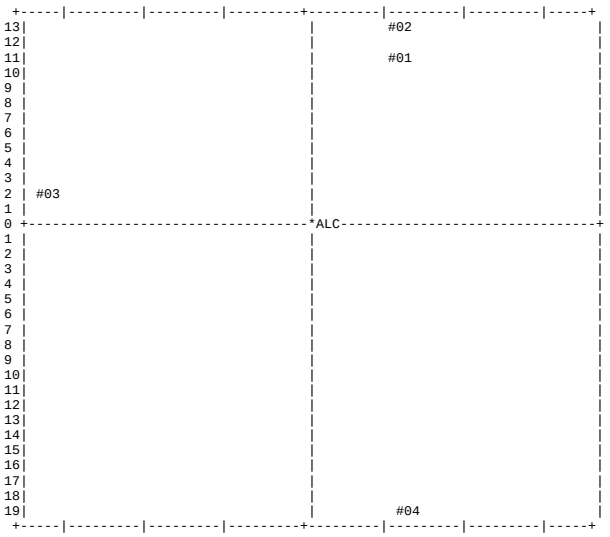


Figure 2.53: Class projection of Corpus 3

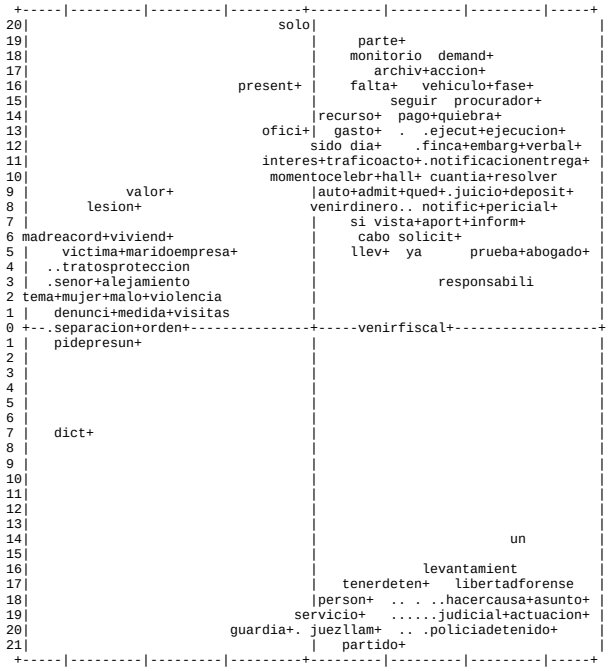


Figure 2.53bis: Word projection of Corpus 3

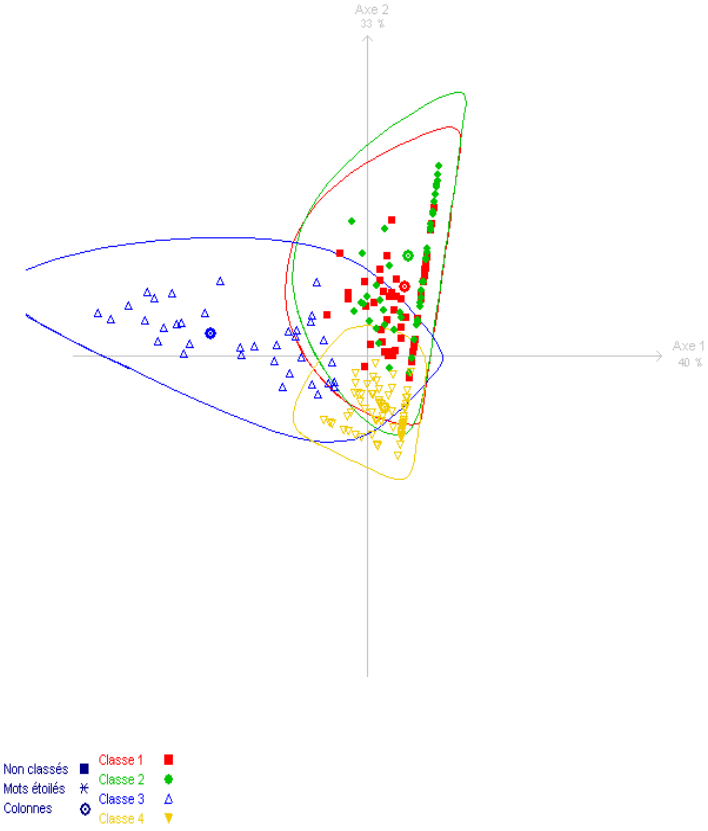


Figure 2.54: Factorial plan of Corpus 3

As we saw in Corpus 2, the X axis can be interpreted as well as representing the path from the private world (family, left side) to the public world (legal system, right side in a word projection graphic). Again, the Y axis would somehow represent the boundaries between these two fields. In particular, distribution of words through the Y axis refers to the interlocutory decisions the judge has to make *on duty*, to the proceedings (lower half of the graphic), and to the legal qualification, trial and judgment (upper half of the graphic).

Class 1: PROCEEDINGS AND TRIAL
juicio+(37), verbal+(16), cuantia+(9), instruccion(13), admit+(7), circunstancia+(8), fase+(8), inform+(9), ordinario+(5), pericial+(6), procedimiento+(18), vista+(13), acto+(7), acumularse(4), aport+(6), cada(4), celebr+(5), dia+(11), escrit+(6), interrogatorio(3), momento(7), monitorio(5), oral+(5), parec+(4), parte+(25), procurador+(8), prueba+(11), recurso+(5), resolver(6), suspenderse(4), venir(9), celebracion(3), decision+(3), escritura+(2), fiscal+(11), ofici+(6), prev+(2), producid+(3), responsabilidad+(5), seguir(6), trafico(5), venir(4), venta+(2), abogado+(8), acumul+(3), antigua+(3), citar(3), distinta+(3), falta+(7), meses(3), oposicion(3), present+(8), reivindicatoria+(3), si(24), suspend+(3), tomar(3), alcoholemia+(1), diferencia+(2), injusto(2), pendiente(2), propon+(2), recurr+(4), senal+(3), acumulacion(2), admision(2), cambiar+(2), derecho+(3), entend+(2), fax(2), haciend+(2), ley(5), mismo+(5), ocurr+(2), piden(2), preceptiv+(2), presencia(2);

Figure 2.55: Specific vocabulary of Class 1

Class 1 represents 18.10% of the ECU and has been labeled "Proceedings and trial", for it mostly refers to the final steps of the judicial process in the context of the judicial system (not private field), as it may be observed above. Fig. 2.55 shows its most representative vocabulary, containing words such as trial [*juicio*+], verbal [*verbal*+], quantity [*cuantía*+],[21] proceeding [*instrucción*, *procedimiento*+], circumstance [*circunstancia*+], phase [*fase*+], ordinary [*ordinario*+], expert [*pericial*+], act [*acto*+], party [*parte*+], attorney [*procurador*], etc.

Class 2: ENFORCEMENT (JUDGEMENT)
ejecucion+(15), ejecut+(15), embarg+(10), finca+(9), depositario(5), notificacion+(7), pago+(5), demand+(14), interes+(7), quiebra+(6), sentencia(10), accion+(4), acordarse(4), archiv+(6), bienes(4), cabo(12), coche(3), condenad+(3), deposit+(6), edicto+(2), entrega+(4), fallo(3), gasto+(5), hall+(5), imposibilidad(3), llev+(12), modific+(3), multa(3), nombr+(6), pretension+(2), sido(6), sociedad(4), solicit+(9), suspension(4), valida(2), vehiculo+(3), acreedor+(2), auto+(7), cantidad(3), conocimiento+(2), costas(2), dano+(2), dinero(4), fecha(3), notific+(4), obra+(2), personal+(5), privativ+(2), qued+(4), solo(5), subvencion+(2), veces(2), ya(6), administracion+(3), afect+(3), forma(3), grabad+(1), interpretacion+(1), subasta(1), deslind+(2), despach+(2), diligencia+(9), haber(3), habido(4), hacerl+(3), impon+(2), indemnizacion+(2), pasiv+(2), pena(2), propietari+(2), reclama+(2), represent+(2), sobreseimiento(2), volver+(2), acced+(3), anterior+(3), dich+(5);

Figure 2.56: Specific vocabulary of Class 2

[21] Referred to money.

Class 2 represents 14.39% of the classed ECU and occupies nearly the same place as Class 1 at the word distribution graphic. However, Class 2 expresses a proper lexical world by itself since it refers to a very specific (final) part of the legal proceeding that causes doubts and problems to inexperienced judges: judgment enforcement. Its specific vocabulary (Fig. 2.56) is specially illustrative: enforcement [*ejecución*+], enforce [*ejecut*+], garnishment [*embarg*+], property [*finca*], payment [*pago*], demand [*demand*+], interest [*interés*], bankruptcy [*quiebra*+], judgment [*sentencia*+], administration [*administración*+], etc.

Class 3: FAMILY LAW (GENDER VIOLENCE, DIVORCE, SEPARATION…)
alejamiento(21), malo+(22), medida+(14), mujer+(15), orden+(24), proteccion(17), senor+(14), tratos(22), victima+(10), domestic+(9), padre+(7), violencia(9), denunci+(13), madre(7), marido(6), pension(5), agresor(4), alimentos(4), hijo+(4), lesion+(7), maltratada+(4), nino+(5), separacion+(5), viviend+(5), acogida(2), acord+(8), casa(3), cautelar+(3), dict+(9), divorcio+(2), empresa+(4), otorg+(3), pareja(4), psicologico+(4), regimen(4), visitas(4), abuso+(2), amenaz+(3), hermano+(2), nina(2), pide(9), presun+(6), proceso+(3), sus(3), tema+(6), valor+(5), adopt+(2), alguien(2), apreci+(2), conoc+(2), familia(3), prision(7), proteger(2), sexual+(2), suficiente+(3), supuesto+(6), venido(3), dejar(2), ingreso+(2), tambien(2), asistencia(1), ayud+(2), base+(2), comparecencia+(3), consegu+(2), declaracion+(3), desahuci+(1), duda+(3), familiar+(2), imputado+(3), insult+(1), pedido(2), posterior+(1), reforma+(1), terminos(1), via+(2);

Figure 2.57: Specific vocabulary of Class 3

Class 3 represents 13.46% of the ECU and refers quite clearly to the middle stages of Family Law proceedings. In general, it somehow contains external elements to this type of process (Fig. 2.57): withdrawal [*alejamiento*+], ill [*malo*+], treatment [*tratos*], measure [*medida*+], woman [*mujer*+], order [*orden*+], protection [*protección*], man [*señor*], victim [*víctima*+], domestic [*doméstic*+], father [*padre*], violence [*violencia*], denounce [*denunci*+], mother [*madre*], husband [*marido*], pension [*pensión*], aggressor [*agresor*], food [*alimentos*], son [*hijo*+], injury [*lesión*+], ill-treated (she) [*maltratada*+], child (he) [*niño*+], separation [*separación*], divorce [*divorcio*], prison [*prisión*], protect [*proteger*], sexual [*sexual*], etc.

Class 4: "ON DUTY" ACTIONS
juzgado+(64), juez(48), cadaver+(19), detenido+(18), funcionario+(20), guardia+(30), internamiento+(18), judicial+(38), llam+(26), medico+(23), policia(29), actuacion+(10), acud+(20), asunto+(21), causa+(8), comision(8), competente+(10), deten+(11), disposicion(9), funcion+(10), hacer(91), hospital(8), informacion+(8), levantamiento+(11), mand+(17), oficina(11), organo+(9), pais(8), partido+(19), person+(34), problema+(14), rogatoria(8), servicio+(14), autopsia+(7), destino+(6), diciendo(9), esper+(6), forense(11), identific+(6), libertad(10), papel(6), preso+(6), registro+(6), requer+(6), secretari+(10), seguridad+(7), tard+(6), tener(9), traslad+(10), un(145), adjunto(5), autoridad+(5), busc+(5), conten+(5), durante(12), entrada+(5), extranjeros(5), gente(5), internacional(5), investigacion(5), jueces(5), jurisprudencia(5), municipal(5), otros(8), public+(12), rapido+(18), reconocimiento(8), situacion+(5), social+(10), tecnicos(5), urgencia+(5), actu+(13);

Figure 2.58: Specific vocabulary of Class 4

And finally we have Class 4, which represents 54.06% of the ECU and expresses specific actions the judge has to perform during the *on duty* period. This class is strongly opposed to classes 1 and 2, for it refers to the very first steps of any incoming case during *on duty* period and it is quite opposed to Class 3 because it brings us back from the private (Family law) to the public field (legal system). The words that represent this class express (Fig.2.58) typical *on duty* activities (internment [*internamiento*+], judicial [*judicial*+], call [*llam*+], action [*actuación*+], go [*acud*+], perpetration [*comisión*+], do [*hacer*], information [*información*+], send [*mand*+], autopsy [*autopsia*], identification [*identific*+], register [*registro*+], urgency [*urgencia*+], etc.), the main actors (judge [*juez*], arrested [*detenido*+], civil servant [*funcionario*+], doctor [*médico*+], forensic doctor [*forense*], corpse [*cadáver*], police [*policía*], person [*person*+], law clerk [*secretario*], experts [*técnicos*], etc.) and the places where these actions take place (court [*juzgado*+], on duty [*guardia*+], hospital [*hospital*], office [*oficina*], country [*país*], judicial district [*partido*], appointment [*destino*+], etc.).

2.4.3 Ongoing work: the team of the Spanish Judicial School

All these results are preliminary. However, they may be used to understand the particular shape of the problems, doubts and questions that judges have expressed from their first court units. They are especially useful to redesign the legal areas in which these problems are located, because the internal structure of the classes plotted by ALCESTE draws *from the inside* a professional mapping of the problems which does not match exactly with the same problems viewed from the outside. This means that practical problems are not necessary doctrinaire or dogmatic. We think that only a judge or a magistrate that has faced these same problems in court or during his on-duty time himself can offer a reasonable answer.

Therefore, we set up a judicial team of magistates to work out this kind of problems with the UAB researchers. The former results are being analyzed from the professional point of view by a team of three Magistrates and one Prosecutor belonging to the Judicial School.[22] This team has been especially entitled by the School Commission of the General Council of the Judiciary since July 2004 to answer the questions extracted from the questionnaires.

For the ontology building task, paying a special attention to the vocabulary is also worthwhile. This vocabulary is full of technical terms that refer to the most general language of law. But it contains other terms as well which belong properly to the judicial domain, and even more specifically, to the procedural part of the trial called *instrucción* [proceedings]. This is not surprising at all. The first step in a Spanish judge career takes place in the lower courts, where judges must face all the typical problems of instructing the incoming cases since the beginning.

Problems seem to be centered in three different domains: (i) judges' on-duty time ; (ii) court proceedings; (iii) family conflicts and gender violence. From the judicial point of view, there is always a procedural approach with a social dimension. The legal task of qualifying crimes, the assessment of a "legal type" to some particular facts or findings - what is called in Spanish "tipo penal"- seems to be doubled by a real worry for the

[22] Magistrates Manuel Bellido, Pascual Ortuño, and Antonio Doñate; Prosecutor Manuel Estrampes.

effect of the judge's intervention in solving the situation ("how can I help this man or this woman?", "what I am going to do with this man"?…). What is interesting here is the position of judges reflecting on their own duty and the possible impact of the mediate and intermediate decisions they must make to instruct certain findings as a legal case.

2.5 Legal Documents

In this section, we are going to introduce the different kinds of judicial documents that a court produces and, in particular, focus on the types of judicial rulings that a judge or a court makes. We will also address the main differences and similarities between the rulings of courts in different jurisdictions and hierarchical positions. Finally, we will present some statistical data regarding the volume of judicial rulings in Spain and their distribution through commercial legal databases.

The goal of this section is to explore the possible ways to integrate the existing case law and legal databases on a large scale into the FAQ system and illustrate the answers offered by the system with relevant references to the case law.

2.5.1 Typology of Spanish judicial documents

A court and its judicial office produce a great variety of documents, which may be called "judicial documents" for simplicity. We will exclude from this category documents produced by the parties to a case, third parties or other public officials, and use it exclusively to refer to documents produced by a "court". Based on the function that a document plays, one may distinguish at least three types of judicial documents: judicial communications, judicial decisions and judicial records.

2.5.1.1 Judicial Communications

In spite of the great possibilities to communicate that new technologies offer, a judicial office still uses mainly written documents to communicate with parties, public officials and other courts. To communicate with private parties, a judicial office may issue a notification (*notificación*), a summon (*citación* or *emplazamiento*), or a court order (*requerimiento*). To communicate with public officials, a court may use an official letter (*oficio* or *exposiciones*) or a writ of mandamus (*mandamiento*). To communicate with other courts, a judge may formulate a judicial request (*exhorto*), to address another national court, or a rogatory letter (*comisión rogatoria*), to address a foreign court.

2.5.1.2 Judicial Decisions

An essential part of the role of a court is to take decisions regarding cases in its docket. Those decisions are sometimes rendered orally, and then documented, but most of the time they are directly written. There are basically three types of judicial decisions, based on the subject matter of the issue with which they deal, on whether they put an end to a given issue, and on whether they are motivated: judicial orders (*providencias*), interlocutory decisions (*autos*) and judgments (*sentencias*). Since those are the judicial documents that are most relevant for our project, we will come back to them below.

2.5.1.3 Judicial Records

Finally, another type of judicial document that a court produces is that which is intended to document and record any type of judicial act or hearing. As a matter of fact, the main task of a Judicial Secretary is to validate any judicial act or hearing and to protect the court's records. As of today, most judicial records are still in writing. However, in civil courts, for example, since 2001 ample use is made of video taping to record hearings and court procedures. Due to the great success of this system, in the near future, one should expect the use of video taping to extend to other jurisdictions.

2.5.2 Typology of judicial rulings (tipos de sentencia)

Because of their role in providing guidance to other parties and public officials on how to accommodate their behavior to what the law requires, judicial decisions regarding the substance of cases are the most important judicial documents for society. As we have already pointed out, in Spanish procedures, judges express themselves basically through three types of judicial documents: judicial orders, interlocutory decisions and judgments. Here we will briefly review their essential characteristics, which are regulated in Chapter IV, Title III, Book III of the Spanish Organic Act 6/1985 on the Judicial Power.

*2.5.2.1 Judicial order (*Providencia*)*

A judicial order is the typical form of a decision that bears on the procedural management of a case: this is, it is the usual form of a decision regarding the procedures to be followed, acts to be performed and all motions as to how the case is to be handled.

A judicial order is the simplest type of decision in that it must merely contain the actual decision of the judge, the date in which it is rendered, and the signatures of the Judge and the Secretary. Characteristically, it does not need to be motivated, that is, it does not need to state the grounds for the decision, although it may, if the judge deems it appropriate.

While judicial orders are quantitatively the most important type of decision, because of their merely procedural nature, they are not usually of interest to third parties beyond the parties to the case, and, therefore, they are neither collected nor published. They can be basically disregarded for the purposes of our project.

*2.5.2.2 Interlocutory Decision (*Auto*)*

An interlocutory decision is the typical form of a ruling that decides (1) on appeals against judicial orders, (2) on prejudicial issues that may have been raised —for example, whether the victim and the injurer where married to apply specific provisions for domestic violence or not—, (3) procedural requirements —for example, whether the court has jurisdiction to hear the case, whether the case is *res judicata*, whether there is another case pending on the same issue (*lis pendes*), etc.—, (4) the invalidation of previous acts performed in violation of basic procedural requirements or fundamental rights, and (5) any other issue, when the law requires that the decision take this form.

Interlocutory decisions are more complex than judicial orders: they must be motivated and, therefore, they must contain a history of the case (*Antecedentes de Hecho*), any findings of fact (*Hechos Probados*) if the decision requires the legal grounds for the decision (*Fundamentos de Derecho*), and finally the decision itself (*Parte Dispositva* or *Fallo*), in separated and numbered paragraphs. They must also be signed by the Judge and the Secretary.

Since interlocutory decisions sometimes touch upon important issues, for example, on the availability of provisional measures, or the admissibility of an appeal, some of them are collected and reported in databases.

*2.5.2.3 Judgment (*Sentencia*)*

A judgment is qualitatively the most important type of judicial decision. It is the usual form of a decision that puts an end to a case. In this sense, it can be grounded either on the merits of the case or on procedural issues, although usually they are decisions on the merits.

Judgments are much like interlocutory decisions in that they must also be motivated and, therefore, must contain a heading identifying the case, the history of the case, the findings of fact —we will see that, depending on the jurisdiction, facts are presented in different ways—, the legal grounds of the decision and, finally, the decision itself. They must also be signed by the Judge and the Secretary.
When a court sits in panels of judges, sometimes a judgment may be accompanied by a dissenting opinion by one of the judges. Dissenting opinions are considered to be part of the judgment: they are attached to it and both must be published together.

In each court or tribunal, a record of all judgments, interlocutory decisions, and dissenting opinions is kept by the Secretary. This is called the Book of Judgments (*Libro de Sentencias*). Once a judgment or interlocutory decision is duly signed, it is numbered correlatively and incorporated into the Book of Judgments, and can then be accessed by any interested party.

2.5.3 Statistical framework of legal rulings in Spain (CGPJ):

Having described the basic kinds of judicial decisions and their differences and commonalities, it might be interesting to survey the volume of decisions rendered by Spanish Courts each year and the way in which they are made available to the general public.

Fig. 2.59 shows the number of judgments and interlocutory decisions rendered by each type of court during 2002, as well as the total number of cases terminated. As one can see, Spanish Courts overall may render around 1.4 million judgments a year and around 600.000 interlocutory decisions. Approximately 43% of all judgments are rendered by criminal courts, 32% by civil courts, 16% by labor courts, and 9% by administrative courts. Interlocutory decisions of civil courts represent 79% of all interlocutory decisions, those of administrative courts 13%, those of criminal courts 7%, and those of labor courts 1%.

JURISDICTIONS	JUDGMENTS	INTERLOCUTORY DECISIONS	CASES TERMINATED
Civil (1)			
First I. Courts (without family jurisdiction)	86.243	164.341	258.996
First I. Courts (with family jurisdiction)	9.256	10.317	21.185
Family Courts	41.687	21.431	66.197
First Instance and Investigation Courts	211.736	267.648	496.712
Court of Appeals (Civil Sections)	49.623	15.079	64.991
Court of Appeals (Merged Sections)	38.466	10.492	49.240
Superior Courts of Justice (Civil Sections)	131	123	257
Supreme Court (First Section)	1.306	4.795	6.073
TOTAL	**438.448**	**494.226**	**963.651**
Criminal (2)			
Investigation Courts	128.309	–	1.991.138
First Instance and Investigation Courts	253.804	–	3.073.669
Juvenile Courts	15.216	–	31.850
Penitentiary Courts	–	–	190.221
Criminal Courts	125.794	–	136.586
Central Investigation Courts	–	–	3.655
Central Criminal Courts	113	–	114
Central Juvenile Courts	17	–	15
Courts of Appeal (Crim. Sections)	35.963	21.514	57.990
Courts of Appeal (Merged Sections)	32.654	16.980	51.161
Superior Courts of Justice (Crim. Section)	143	574	727
National Court (Criminal Section)	117	2.749	1.623
Supreme Court (Second Section)	1.605	3.380	4.985
TOTAL	**593.805**	**45.197**	**5.543.734**
Administrative			
Central Administrative Court	2.050	1.104	3.169
Administrative Courts	45.728	27.403	72.941
Superior Courts of Justice (Adm. Section)	69.861	43.132	99.338
National Court (Administrative Section)	8.281	3.136	11.622
Supreme Court (Third Section)	4.685	5.759	11.537
TOTAL	**130.605**	**80.534**	**198.607**
Labor			
Labor Courts	162.201	–	259.674
Superior Courts of Justice (Labor Section)	60.584	3.104	64.432
National Court (Labor Section)	91	80	241
Supreme Court (Fourth Section)	1.067	4.007	5.074
TOTAL	**223.943**	**7.191**	**329.421**
Military			
Supreme Court (Military Section)	179	128	307
TOTAL	**179**	**128**	**307**
Supreme Court (Special Sections)	21	69	85
TOTAL	**21**	**69**	**85**
TOTAL JURISDICTIONS	**1.387.001**	**627.345**	**7.035.805**

(1) Not including acts of voluntary jurisdiction
(2) Not including indeterminate proceedings
(Source: Memoria CGPJ 2003)

Figure 2.59: Decisions per Type of Court per Year (2002)

And 78% of all judgments and also all interlocutory decisions are handed down by lower courts.[23] These percentages, however, do not necessarily reveal the importance of the judgments and decisions for society as a whole. For instance, while criminal courts render more judgments than other jurisdictions, this is due to the fact that any conviction requires a judgment, even if it is a petty offence and even if a plea bargain between the prosecutor and the accused has been reached, and therefore, most of those judgments are not necessarily interesting to report. Similarly, while lower courts of all jurisdictions render most decisions, their decisions are not as significant as those of higher courts since most of their time is devoted to fact-finding rather than rule-elaboration.[24]

The collection and distribution of judicial rulings is done by the CENDOJ, and they are made available to the public through commercial legal databases. The CENDOJ (*Centro de Documentación Judicial* or Judicial Documentation Centre) is an official centre dependent on the CGPJ, which works as a national repository of judicial rulings and a centre for their distribution.[25] We have seen before that each court or tribunal keeps a Book of Judgments in which judgments and interlocutory decisions are compiled. Some types of courts, namely the Supreme Court, the National Court, Superior Courts of Justice and Courts of Appeal, have an obligation to send a copy of all their judgments to the CENDOJ on a monthly basis.[26] Other lower courts may be required to send some of their decisions by request of the CGPJ. Fig. 2.60 gives an idea about the number of judgments received by the CENDOJ in the years 2001 and 2002 and the format in which they were collected.

[23] We shall consider that the following courts are "lower courts": Civil: First Instance Courts (no family matters), First Instance Courts (with family matters), Family Courts, First Instance and Investigation Courts; Criminal: Investigation Courts, First Instance and Investigation Courts, Juvenile Courts, Criminal Courts, Central Criminal Courts, Central Juvenile Courts; Administrative: Central Administrative Court, Administrative Courts; Labor: Labor Courts.

[24] On the other hand, note that since the FAQ system is intended primarily for judges sitting in lower courts, it may be that the rulings passed by other lower court judges may be more interesting to other lower court judges than to society as a whole.

[25] It also has a library that can be used by judges and scholars and publishes many books and journals for judges and the general public.

[26] The latest regulation on the issue is Agreement the CGPJ of April 9, 2003, which adopts Regulation 4/2003 on the collection of judicial decisions by the CGPJ for their compilation and treatment by the CENDOJ (BOE num. 104, May 1, 2003).

2001	COURT	PAPER	DISK	LOTUS NOTES	TOTAL
	Supreme Court	0	0	30080	30080
	National Court	270	212	7642	8124
	Superior Courts of Justice	73650	44602	2140	120392
	Courts of Appeal	117285	35604	8408	161297
	TOTAL JUDGMENTS	191205	80418	48270	319893
2002	**COURT**	**PAPER**	**DISK**	**LOTUS NOTES**	**TOTAL**
	Supreme Court	0	0	30008	30008
	National Court	266	106	7885	8257
	Superior Courts of Justice	61563	43612	2223	107398
	Courts of Appeal	105614	34268	3133	143015
	TOTAL JUDGMENTS	167443	77986	43249	288678

(Source: Memoria CGPJ 2003)

Figure 2.60: Format of Judgments Received at the CENDOJ (2001-2002)

As one can see, in 2002, around 15% of the judgments were in lotus notes format (database), 27% were sent in a computer disk, and 58% were still in paper format. The large amount of documents that are received in paper creates a problem of storage, because of the large amount of space required, and of handling, since pages must be unbound for photocopying and then rebound, and a general handicap for the swift treatment of the judgments, because the CENDOJ is supposed to convert all those into electronic format and compile a single database.[27] All names and personal identifications are also removed from the documents before they are delivered to the public to preserve the right to privacy of the persons involved. This is another major problem since this process consumes a considerable amount of time and resources.

Finally, judges and third parties may request judgments to the CENDOJ on specific issues or wholesale. In fact, the CENDOJ has agreements with several legal publishers and many local bars to periodically distribute the judgments that it receives.

[27] See more on this by Concha Alvaro Bermejo, Iñigo Sanz de Ormazabal and Marina Cueto Aparicio, Centro de Documentación Judicial (CENDOJ) del Consejo General del Poder Judicial, http://fesabid98.florida-uni.es/Comunicaciones/c_alvaro/c_alvaro.htm (last visited December 10, 2004).

BUSINESSES	QUANTITY
A & F Abogados	133
Auloce, S.A.	285
Centro Europeo Estudios y Formación Empresarial	2.285
Civitas Ediciones, S.L.	282
Difusión Jurídica y Temas de Actualidad Económica	27.577
Editorial Agal-57, S.L.	498
Editorial Aranzadi, S.A.	307.663
Editorial Bosch, S.A.	84.144
Editorial Lex Nova, S.A.	20.090
Ediciones Netlex, S.L.	3.683
El Derecho Editores, S.L.	88.194
Europea de Derecho, S.A.	23.307
Facultatis Iuris	2.216
Financial-Tax Ediciones, S.L.	1.186
Francisco Javier Messia de la Cerda	2
Gage Datadiar, S.L.	14.627
Gomylex, S.L.	39
José Luis Encinas Pardo	1.090
Juego Legal, S.L.	221
Jurcom y Legiscom, S.L.	2.208
Jurisoft	16.900
La Ley Actualidad, S.A.	242.707
M.N.M. Programación, S.L.	12.458
Normacef, S.L.	22.883
Portal Derecho, S.A.	35.374
Resoluciones Judiciales, S.L.	6.255
Revista General del Derecho, S.L.	38.556
Servicio Propiedad Inmobiliaria, S.L.	13.168
Vlex Networks,, S.L.	264
TOTAL	**968.295**

(Source: Memoria CGPJ 2002)

Figure 2.61: Decisions served by the CENDOJ to Businesses (2001)

Fig. 2.61 shows the number of judgments served by the CENDOJ to businesses in 2001. The data illustrates that the major business clients of the CENDOJ —and consumers of judicial rulings— are Aranzadi (32% of all documents requested by businesses), La Ley (25%), followed by El Derecho Editores (9%) and Editorial Bosch (9%). Fig. 2.62 contains the same type of data for non-business entities.

OTHER ENTITIES	QUANTITY
Spanish Association of Family Lawyers	3.991
Baleares Bar Association	13.585
Badajoz Bar Association	6.491
Barcelona Bar Association	104.862
Castellón Bar Association	4.981
Cuenca Bar Association	683
Granada Bar Association	25.945
Guipúzcoa Bar Association	11.501
La Rioja Bar Association	2.368
Sevilla Bar Association	27.278
Valencia Bar Association	46.780
Valladolid Bar Association	2.436
Zaragoza Bar Association	14.885
Social Graduates Bar Association	56.585
Labor Commissions (CC.OO.)	60.177
Counsel of Andalucía Bars	16.084
Drugs Observatory National Plan	883
General Workers Union (UGT)	60.177
TOTAL	**459.692**

(Source: Memoria CGPJ 2002)

Figure 2.62: Decisions served by the CENDOJ to Other Entities (2001)

Among other entities that request decisions from the CENDOJ, attention must be called to the Barcelona Bar (23% of all documents requested by non-businesses), and two labor unions, Labor Commissions, and General Workers Union (both 13%). These numbers and percentages reveal that there are two major players in the market for legal databases (Aranzadi and La Ley). The SEKT project will be working together with La Ley, who will provide access to their databases.

To conclude this section, a short reference must be made to the work of the Constitutional Court because of its qualitative importance. While the Court is called upon to rule on many cases, it produces a limited number of decisions.

	JUDGMENTS	INTERLOCUTORY DECISIONS	CASES TERMINATED
Constitutional Court	230	429	6735

(Source: Memoria Tribunal Constitucional 2003)

Figure 2.63: Decisions of the Constitutional Court (2003)

The Constitutional Court is not formally part of the Judiciary Power and, for this reason, has its own separate administration, compilation and distribution channels. Its judgments are actually published in the Spanish Official Journal (BOE).

2.5.4 The common structure of judicial final rulings

In this section we are going to address the main differences and similarities between the judgments that several types of courts hand down. The purpose of this is to allow the automatic recognition of the different parts of a judgment and their electronical systematic processing.

2.5.4.1 Basic Similarities

As one could already infer from the description of the different types of rulings, interlocutory decisions and judgments of all courts in all jurisdictions have a basic common structure with the following elements:

(A) The **Title** of the document (Interlocutory Decision or Judgment) and the number of the decision in the court's records.

(B) A **Heading** which refers
   a. the date and location of the decision
   b. the court hearing the case at present & other courts having heard the case in other instances
   c. the docket number of the case or appeal (various numbers if the case has been through several instances)
   d. the parties to the case (the plaintiff, the defendant, the Public Prosecutor, and other intervening parties)
   e. the representatives (*Procurador*) and counsels (*Abogado* or *Letrado*) of each of the parties (if applicable)
   f. the subject matter of the case (charges, object, etc).
   g. the author of the opinion.

(C) The **History of the Case** (*Antecedentes de Hecho*), this is, the basic steps through which the case has been, in numbered and separated paragraphs:
   a. the initiation of the proceedings (complaint, police report, etc.), the basic contents of the complaint or report, and the date.
   b. the response of the defendant or the accused to the complaint or the report, and the date.
   c. any hearings that may have taken place, issues that may have been raised, and any decisions that may have been taken, and their date.
   d. the previous decisions on procedural issues or on the merits of the case, this is, the exact contents of those decisions, which usually entails a literal reproduction of the decision (*Fallo* o *Parte Dispositva*).
   e. if the case is on appeal, the grounds of appeal and the response to those grounds by the opposing party.

(D) The **Findings of Fact** (*Hechos Probados*): a clear statement of the facts that the judge or panel finds to have been proved —the standard varies depending on the jurisdiction—, in numbered and separated paragraphs.

(E) The **Grounds of the Decision** (*Fundamentos de Derecho*), which usually means that for each type of issue raised in the case there will be:
   a. A paragraph that determines the applicable legal provisions and relevant case law. Judges like to reproduce here lengthy considerations which many times are *obiter dicta* (considerations that are not strictly relevant to the case at hand).
   b. A paragraph that applies the legal rules and precedent to the facts of the case and derives a conclusion.
   c. One should also note that procedural issues will be treated first and substantive issues later.

d. Finally, the opinion will address the allocation of litigation costs (if applicable).[28]

(F) The **Decision** (*Fallo* or *Parte Dispositiva*), which has to determine, in separated and numbered paragraphs:
   a. Whether the complaint is allowed or dismissed, or whether the appeal is allowed or rejected, partially or in its entirety.
   b. If the complaint is allowed, then any declarations that may have been requested and/or the type of behavior that is ordered, as well as the parties that the decision addresses.
   c. An allocation of litigation costs (if applicable).

(G) **Appeals**: whether the decision can be appealed or not, what types of appeals are allowed and the term to formulate an appeal.

(H) **Signatures** by the Judge or Judges to the decision and the Secretary.

Some of these basic elements can be identified on the following figures, which partially reproduce a judgment of the Barcelona Court of Appeals in a domestic violence case.
Fig. 2.64 is the first part of the Judgment and contains the Heading, and the History of the Case. There are also some database references that have been added by the publisher but are not on the original document.

[28] This is usually the last paragraph before the actual Decision.

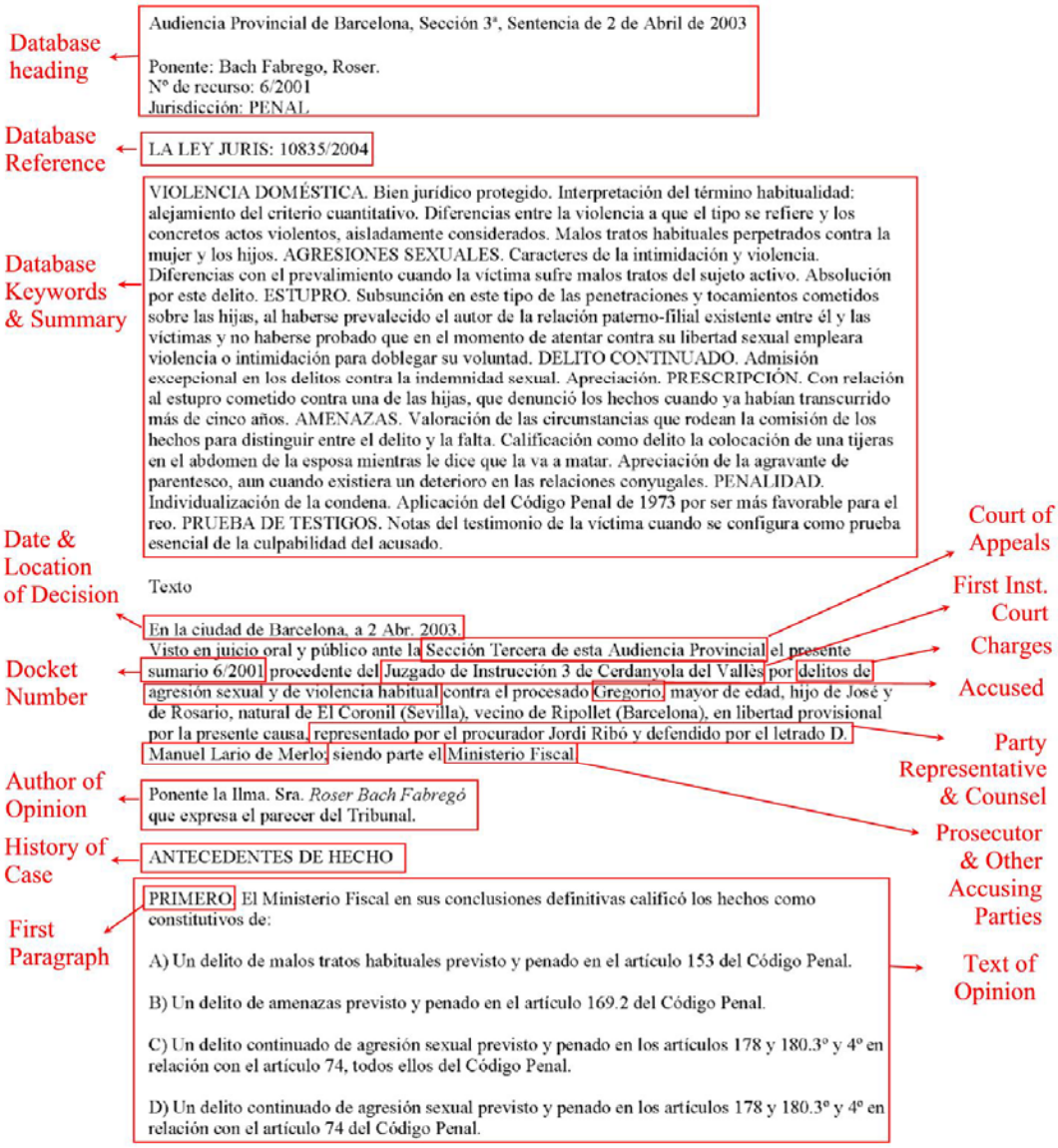


Figure 2.64: Example of Decision (Part 1)

Fig. 2.65 shows the Findings of Fact on which this same decision is based and then the first part of the Grounds of the Decision. And finally, contains the last part of the judgment, in which one can see the Decision and the contents of that decision, a reference to litigation costs, and finally a statement regarding whether the decision can be appealed and the signature of the Judges and the Secretary.

Findings of Fact

HECHOS PROBADOS

El acusado Gregorio, mayor de edad y sin antecedentes penales, contrajo matrimonio con Mónica en el año 1972, de cuya unión nacieron sus cuatro hijos. Desde dicha fecha ha agredido de forma reiterada a su esposa, y asimismo de forma continuada le ha proferido expresiones ofensivas dirigidas a amedrentarla, llegando en algunas ocasiones a esgrimir tijeras, cuchillos y destornilladores para conseguir imponer su voluntad, incluso en presencia de los hijos.

Concretamente, el día 28 Feb. 1999 sobre las 10 h, cuando Mónica se encontraba en el cuarto de baño, su esposo tras decirle «eres una perra, lo que tienes que hacer es ir a trabajar», le dio un fuerte empujón tirándola contra la bañera, sin causarle lesiones. Cuando Mónica consiguió levantarse del suelo se dirigió a la cocina, donde entró también el acusado llevando en la mano unas tijeras que le colocó en el abdomen mientras le decía que la iba a matar.

El acusado también ha agredido de forma reiterada a sus hijos y también de forma continuada les ha proferido expresiones dirigidas a amedrentarles, esgrimiendo cuchillos y tijeras con ánimo de conseguir su propósito.

Esta situación se prolongó en el tiempo hasta el año 2000 en que los cónyuges se separaron.

El acusado desde el año 1985 hasta el año 1991, en numerosas ocasiones entraba en la habitación de su hija Carla, nacida el 22 Nov. 1978 y de su hija Paula, nacida el 9 Jul. 1976, haciendo a ambas objeto de tocamientos e instándolas a que le masturbasen, llegando con relación a Carla, a penetrarla analmente en cuatro ocasiones distintas, a finales del año 1990. La conducta del acusado sobre su hija Paula finalizó en el año 1989. El procesado las amenazaba con matarles a ellas o a su madre si contaban los referidos hechos.

No obstante Carla explicó los hechos a la persona con la que mantenía una relación sentimental, y posteriormente denunció los mismos.

La denuncia se formuló en fecha 1 Mar. 1999.

Grounds of Decision

FUNDAMENTOS DE DERECHO

First Paragraph

PRIMERO. La convicción del Tribunal se ha conformado de acuerdo con la prueba practicada en el acto del juicio oral con todas las garantías y bajo los principios de inmediación contradicción.

En primer término en lo que se refiere al delito de violencia habitual en el ámbito familiar, debe ponerse de manifiesto que, además de haber sido admitido por la defensa en sus conclusiones definitivas, el propio procesado en el acto del juicio ha reconocido la conducta que integra la referida conducta delictiva. Ha admitido que es una persona agresiva y que tiene «pronto aunque después se le pasa», que a veces no se puede frenar y después se siente arrepentido y pide disculpas. Afirmó asimismo haber dado «tortazos» a sus hijos y a su esposa y haber proferido expresiones amenazantes y haberles lanzado objetos.

Figure 2.65: Example of Decision (Part 2)

Decision

Vistos los preceptos legales citados y demás de general y de pertinente aplicación,

FALLO

Convicted Person

Declaration of Guilt

Que debemos condenar y condenamos al procesado Gregorio como autor criminalmente de un delito de malos tratos habituales sin la concurrencia de circunstancias modificativas de la responsabilidad criminal a la pena de dos años de prisión con la accesoria de inhabilitación especial para el derecho de sufragio pasivo durante el tiempo de la condena; como autor criminalmente responsable de un delito de amenazas con la concurrencia de la circunstancia modificativa de la responsabilidad criminal agravante de parentesco a la pena de año y seis meses de prisión con la accesoria de inhabilitación especial para el derecho de sufragio pasivo durante el tiempo de la condena; y como autor criminalmente responsable de un delito continuado de estupro de prevalimiento sin la concurrencia de circunstancias modificativas de la responsabilidad criminal a la pena de seis años de prisión, con la accesoria de inhabilitación especial para el derecho de sufragio pasivo durante el tiempo de la condena; y al pago de las tres cuartas partes de las costas procesales, declarándose la cuarta parte restante de oficio.

Crimes & Sanctions

Litigation Costs

Damages: quantities & beneficiaries

Asimismo el acusado deberá indemnizar a Mónica en la cantidad de 6.000 euros, a Luis Antonio en la cantidad de 6.000 euros, a Ángel en la cantidad de 6.000 euros, a Paula en la cantidad de 6.000 euros y a Carla en la cantidad de 24.000 euros.

Appeals

Notifíquese que contra la presente resolución cabe interponer recurso de casación por infracción de ley y por quebrantamiento de forma en el plazo de cinco días.

Signature of Judges & Secretary

Así por nuestra sentencia de la que se unirá certificación a las actuaciones, lo acuerdan y firman los Sres. Magistrados anotados al margen; doy fe.

Figure 2.66: Example of Decision (Part 3)

2.5.4.2 Main Differences

But judgments are also different in several respects when one looks across courts and across jurisdictions. To be sure, these differences are minimal, but they are worth noting.

(A) Heading. Differences in the Heading of a decision allow a quick determination of several basic aspects of the case. Indeed, this information is provided in a rather standard way in all decisions:
   a. The name of the court, which also usually allows to determine the jurisdiction and the stage of the proceedings.
   b. The name of the proceedings (*proceso ordinario*, *proceso verbal*, *recurso de apelación*, *recurso de queja*, *recurso de casación*, *recurso por infracción procesal*, etc.), which also allow to discriminate between jurisdictions and hierarchical positions.
   c. The names of the parties, which vary across jurisdictions (civil or labor terms such as *demandante*/*demandado* or *parte demandante*/*parte demandada*, vs. criminal terms such as *parte acusadora* or *ministerio fiscal* or *denunciante* or *acusador particular*/*acusado* or *procesado* or *imputado*). One can also determine the relations between the parties (plaintiff vs. defendant or appellant vs. appellee).
   d. The subject matter of the case. For example, debt collection (*reclamación de cantidad*), consumer protection (*acción colectiva de protección de los consumidores*), car accident (*accidente de tráfico*), job accident (*accidente laboral*), administrative sanctions (*sanciones administrativas*), permits (*licencia*), manslaughter (*homicidio imprudente*), murder (*asesinato*) and a long etcetera.[29]

(B) Findings of Fact. Not all decisions contain a part devoted to enunciate Findings of Fact. This varies across decisions, jurisdictions and the judicial hierarchy. First, an interlocutory decision may not refer any Findings of Fact if the type of decision does not require it. Then, as a general matter, civil and administrative courts do not refer separate Findings of Fact but they usually devote the first paragraphs of the Grounds of the Judgment to the determination of the facts of the case, instead.[30] Lastly, first instance criminal and labor courts do enunciate Findings of Fact, but courts on appeals in these jurisdictions may not, if the facts of the case are not the object of the appeal.

(C) Legal References. Of course, one of the most visible differences between the judgments of different courts are the substantive legal provisions that are cited. It turns out that by identifying the basic norms that are cited in an opinion, it is possible to discriminate between several possibilities. For instance, Code of Civil Procedure (*Ley de Enjuiciamiento Civil* or *LEC*) vs. Code of Criminal Procedure (*Ley de Enjuiciamiento Criminal* or *LECr*) vs. Code of Labor Procedure (*Ley de Procedimiento Laboral* or *LPL*) vs. Code of

[29] These are like keywords that can be used to classify opinions.
[30] On the other hand, some recent decisions in civil cases and administrative cases are beginning to incorporate a separate part devoted to the Findings of Fact.

Administrative Procedure (*Ley de la Jurisdicción Contencioso-Administrativa* or *LJCA*).[31]

(D) *Cassation* Proceedings. Supreme Court decisions are peculiar in another important way in civil, criminal and administrative cases. *Cassation* proceedings are based on the existence of concrete "grounds of appeal" (*Motivos de Casación*) that are formally articulated by the appellant. The Supreme Court therefore has to examine whether any "ground of appeal" is admissible and its opinions are structured around those: each paragraph usually corresponds to a given "ground of appeal", that is enunciated first and then examined by the Court and either accepted or rejected. Decisions of other courts do not have this structure.[32]

(E) New Judgment. When the Supreme Court quashes a decision by a lower court it may directly produce a new judgment for the case. This means that, after the decision on appeal, it will issue a New Judgment, which will have the structure of a judgment: a heading, a history of the case, the grounds of the judgment and a new decision.[33] Other court decisions never have this feature because they do not quash decisions, but rather review their substance.

To conclude, it seems that the structure of judicial rulings is fairly similar across jurisdictions and courts and that while there are some minor formal variations, it is the substance of the decisions that really allows us to distinguish between the decisions of different courts. Hence the importance of key expressions and words that the opinions use in building a system to process decisions automatically.

2.5.5 Spanish Legal Databases

The Spanish market for legal databases was spurred by the incorporation of new technologies during the 1990s. Before that, only Constitutional Court and Supreme Court decisions were widely available in paper format, while other court decisions were randomly published by journals and law reviews from time to time. As a result, lawyers seldom used the case law of the highest courts, because "paper searches" were very time consuming, and barely cited other cases, because they remained unreported.

By the end of the 1990s this changed dramatically as several important publishers began to collect large amounts of decisions systematically across jurisdictions. As of today, there are several commercial databases available, which reasonably cover all courts and jurisdictions.

The table below contains a list of homologated commercial databases by publisher. By "homologated databases" one should understand databases that are the result of

[31] Note that by identifying the number of the rule that is cited and knowing the substance of the rule one could automatically determine the kind of issue raised in the case. The problem is that often times opinions cite many provisions, some of them actually not applicable to the case at hand.

[32] Neither do the Judgments of the Supreme Court in Labor cases because Cassation proceedings in labor cases are not based on the existence of "grounds of appeal".

[33] Of course, this New Judgment is usually shorter because it refers in many respects to the previous judgment on the appeal.

agreements between the CGPJ-CENDOJ and private publishers and which, as a result, collect large amounts of rulings on a given topic or in general. They are not "official" databases.[34]

PUBLISHER	PRODUCT
GENERAL DATABASES	
ARANZADI	General
EL DERECHO EDITORES	General
LA LEY	General
SPECIALIZED DATABASES	
ARANZADI	Criminal Law Specialized Database
ARANZADI	Civil Law Specialized Database
ARANZADI	Labor Law Specialized Database
ARANZADI	Administrative Law Specialized Database
ARANZADI	Tax Law Specialized Database
BOE	IberLex
CISS PRAXIS	Labor Atlas
CISS PRAXIS	Tax Atlas
DIFUSIÓN JURÍDICA Y TEMAS DE ACTUALIDAD	Innova Jurídica Databases (several)
EDITORIAL BOSCH	Actionis Database of Civil Actions
EDITORIAL BOSCH	Private Law Database
EDITORIAL BOSCH	Bosch Database of Legislation & Cases
EDITORIAL BOSCH	Bosch Database of Cases
EL DERECHO EDITORES	Horizontal Property
EL DERECHO EDITORES	Urban leases
EL DERECHO EDITORES	Family Law
EUROPEA DE DERECHO	Civil Law Counselor
EUROPEA DE DERECHO	Commercial Law Counselor
EUROPEA DE DERECHO	Administrative Law Counselor
EUROPEA DE DERECHO	Labor Law Counselor
EUROPEA DE DERECHO	Criminal Law Counselor
FRANCIS LEFEBVRE	Mementis
FRANCIS LEFEBVRE	Tax Nautis
FRANCIS LEFEBVRE	Labor Nautis
JURISOFT	Database of Cases of Court of Appeals and Superior Courts of Justice of Castilla y León, Cantabria and La Rioja
LA LEY	Civil Law News
LA LEY	Administrative Law News
LA LEY	Criminal Law News
LA LEY	Labor Law News
LA LEY	Taxes
LEX NOVA	Labor Law Database (Legislation & Cases)
LEX NOVA	Tax Law Database (Legislation & Cases)
LEX NOVA	Civil Law Database (Legislation & Cases)
NORMACEF	Labor Law Normacef
NORMACEF	Tax Law Normacef
QUANTOR	Labor Law Quantor
QUANTOR	Tax Law Quantor
SEPIN	Urban Leases & Horizontal Property
SEPIN	Code of Civil Procedure 1/2000
SEPIN	Family Law

[34] Homologated databases, however, do advertise the fact that they are homologated.

VLEX	vLex Database of New Technologies
VLEX	vLex Database of Labor Law
VLEX	vLex Database of Civil-Commercial Law
VLEX	vLex Database of Administrative Law

(Source: Memoria CGPJ 2003)

Figure 2.67: Homologated Commercial Databases (2003)

As one can see, there are some general databases (La Ley, Aranzadi, and El Derecho Editores) and several specialized databases by practice area, territory, and contents. By practice area, 20% of the specialized databases focus on labor law, 18% on private-civil-commercial law, 16% on tax law, 7% on criminal law, 7 % also on horizontal property and urban leases, 4 % on family law, 2% on procedural law and another 2 % on new technologies. There is another 11% that cuts across practices areas because they focus on territory and contents.[35]

General databases essentially offer three types of contents: legislation, cases and bibliography. The figure below compares the format and contents of the commercial databases La Ley, Aranzadi (which is now called Westlaw) and El Derecho. The table shows that all databases are comparable regarding their contents. All of them offer access to all national, regional and European legislation in force. The main differences

here lie on the availability of older legislation, the ability of the system to keep track of the changes in specific provisions, and the integration between different provisions through hyperlinks. Moreover, the three databases gather all Supreme Court and Constitutional Court decisions and then a selection of decisions of other courts such as Superior Courts of Justice and Courts of Appeal.[36] The differences among them arise out of the comprehensiveness of the database in terms of the number of decisions of lower courts, their classification of judgments by subject matter and keywords, and the integration of decisions and legislation through hyperlinks. Finally, both Aranzadi and La Ley offer access to their own scholarly articles, but not access to the articles of other publishers. This is unlike legal databases in other countries – (American) Westlaw and Lexis-Nexis for example– and, in this sense, online full-text access to legal scholarship is still limited.

[35] By contents is meant legislation only or cases only.
[36] A subjective estimation suggest that around 10% of the rulings of those courts are now reported, but the numbers are rapidly increasing.

	LA LEY	WESTLAW (ARANZADI)	EL DERECHO
FORMAT	Internet and DVD	Internet	DVD
LEGISLATION			
National	All legislation in force	All legislation since 1930, full text since 1978 and Consolidated Codes	All legislation in force
Autonomous Communities	All legislation in force	All legislation since their creation	All legislation in force
European Community	All legislation in force	Selected legislation since 1952 and other decisions published in the Official Journal of the European Communities	All legislation in force
Collective Negotiation Agreements	All national, regional and provincial agreements since 1998 full text. References and summaries to agreements before 1990 and below the provincial level.	Published in the Spanish Official Journal, the Official Journal of the Autonomous Communities and Provincial Official Journals	
CASES	More than 386.000	Unknown	More than 500.000
Constitutional Court	All decisions since 1981	All decisions since 1981	All decisions since 1981
Supreme Court	All decisions full text since 1985	All decisions full text since 1979 (selected cases on civil matters since 1930)	*All* decisions (1)
Superior Courts of Justice	Selected decisions	Selected decisions	Selected decisions
National Court	Selected decisions	Selected decisions	Selected decisions
Court of Appeals	Selected decisions	Selected decisions	Selected decisions
Courts of First Instance		Selected decisions	
European Court of Human Rights	Selected decisions	All decisions since 1960s	*All* decisions (1)
Court of Justice of the European Communities	Selected decisions	All decisions since 1980s	*All* decisions (1)
General Directorate of Registries and Notaries	All decisions since 1985	Selected decisions	*All* decisions (1)
Antitrust Court	Selected decisions	Selected decisions	*All* decisions (1)
Central Tax Court	Selected decisions before 1998. All decisions since 1999	Selected decisions	*All* decisions (1)
General Tax Directorate	Selected decisions before 1998. All decisions since 1999	Selected decisions before 1998. All binding decisions since 1998.	
Accounting Court	Selected decisions		
General Prosecutor		Selected decisions	
BIBLIOGRAPHY			
	Full text articles of La Ley since 1998, references of articles of other Publishers	Full text articles of Aranzadi, references of articles of other Publishers	

(1) Based on advertisement.

Figure 2.68: Comparison of Main Legal Databases

Specialized databases do not seem to be much different from general databases. In fact, as a general practice, they are based on a restriction of the type of legislation, the cases, and other documents that they make available to their subscribers. Several other interesting and practical aspects may be added to these basic services, such as:

- Commentaries of specific legal provisions and cases.
- Daily news in the field
- Specialized bibliographical references
- Standard forms, contracts and documents used in the field
- Protocols on dealing with other entities such as the Administration
- Solved examples or practical cases
- Encyclopedia or dictionary of specific terms
- Schedules relevant to the field[37]
- Information regarding aids and other sources of financing relevant to the field[38]
- Links to specialized web pages

Figure 2.69: La Ley Cases Main Search Page

Searches for cases both in the general and specialized databases can be performed in a variety of standard ways.[39] One may look for the reference of the case in the database, the reference of the case in the judicial records, the date of the decision, the type of decision, the court, the jurisdiction, the judge writing the opinion for the court, a

[37] For example, tax schedules.
[38] For instance, State aids in labor matters.
[39] Searches in specialized databases are performed in exactly the same way as in general databases, except for a more context specific set of options.

summary of the case, keywords, and legislation, codes and rules cited in the opinion. The text of the opinion can also be searched. Fig. 2.69 shows the main search page of La Ley.

As of today, publishers are gathering more decisions than they are able to process and, as a matter of fact, not all this information is always available. In particular, opinion reported often lack a summary of the case and the keywords associated with the ruling. Since these two elements are critical to efficient searches, it is usually necessary to fall back on less efficient free-text searches for an exhaustive exploration of the case law.

The results window in the databases will usually show a list of documents, if the search criteria render more than one possible target document, or a single document. A "treated" opinion –an opinion that has been analyzed by the database's experts– will usually show, apart from the full text of the opinion, a summary of the opinion regarding the main issues of the case, an outline of the opinion to navigate its text, and relevant links to rules and cases cited by the opinion. But many opinions are still "untreated" and in this sense the results may still be quite unsatisfactory as one will have to skim through the text of the opinions. Fig. 2.70 shows the final results page of La Ley.

Figure 2.70: La Ley Final Results Page

Among the interesting features of the results window, and one which owes much to the legal databases of common law countries that pay high regard to precedent, but that has been adapted to the Spanish continental tradition, is the "flagging" of

opinions, or simply the introduction of references to other judicial decisions, and some times even legal scholarship, that either cite the opinion favorably or criticize it. While this may be a very useful tool to locate other relevant information, the way in which it is implemented in Spanish databases is not as functional as in foreign databases.[40]
Once an interesting case is located, the user is able to either "copy & paste" the text of the opinion, download the opinion to a disk, print it, or send it by email. Lists of cases can also be downloaded, printed or sent by email.

With regard the costs of a subscription to those legal databases, it is apparent that the price of the subscription is a function of the contents of the database, the number of users of the access point and the type of client (law firm, academic institution, association, etc). Fig. 2.71 lists the costs of subscribing to several databases that could be found on publicly listed catalogues.[41]

[40] We call it "flagging" because usually colored flags are used as symbols for different types of citations. Westlaw, for example, uses a red flag to show that the opinion has been reversed on appeal, and a yellow flag to show that the appeal has been dismissed. In comparison to how these flags are used in common law databases, in Spanish databases they have a much more limited nature and therefore are not intended to reveal the existence of any trends in the case law. Another shortcoming of the system is that not all decisions are flagged, usually because cases have not been treated yet.
[41] Information on all homologated databases could not be found. Some publishers do not list their prices because of the differences in price due to the variables mentioned. Prices can be obtained by getting in touch with the publisher. For example, a subscription to the Westlaw.es database by the Regional Administration of Navarra costed 89.063 €. See Decision of February 14, 2002 of the General Director of Organization and Information Services of the Foral Community of Navarra, in Official Journal of Navarra num. 69 (07/06/2002), available at http://www.cfnavarra.es/bon/026/02607012.htm.

DATABASE	FORMAT	COST (1)
Westlaw.es	Online	3.300 € (2)
Dat@ley General Database	Online	1.307 €
La Ley Civil Law News	Online	866 €
La Ley Labor Law News	Online	881 €
La Ley Administrative Law News	Online	794 €
La Ley Criminal Law News	Online	785 €
La Ley Taxes	Online	580 €
Labor Law Atlas	Online	807 €
Tax Law Atlas	CD/DVD	860 €
Innova Jurídica Civil Law	CD/DVD	118 €
Innova Jurídica Civil & Commercial Law	CD/DVD	118 €
Innova Jurídica Labor Law	CD/DVD	118 €
Innova Jurídica Courts of Cataluña	CD/DVD	70 €
Bosch Online Database	Online	373 €
Bosch Legislation and Cases	CD/DVD	210 €
Bosch Cases Database	CD/DVD	536 €
Bosch Actionis Database of Civil Actions	CD/DVD	536 €
Bosch Private Law Database	CD/DVD	580 €
Mementis	CD/DVD	139 €
Nautis Labor Law	CD/DVD	650 €
Nautis Tax Law	CD/DVD	920 €
Normacef Tax Law	CD/DVD	348 €
Normacef Labor Law	CD/DVD	348 €
Normacef Zoning Law	CD/DVD	116 €
Lex Nova Database (Labor, Tax & Civil Law)	Online	626 € (3)
vLex Premium	Online	1.330 €
VLex Civil-Commercial Database	Online	499 €
vLex Tax Law Database	Online	449 €
vLex Labor Law Database	Online	499 €
vLex New Technologies Database	Online	399 €
vLex Criminal Law Database	Online	299 €
vLex Administrative Law Database	Online	449 €

(1) Based on a 1 year subscription
(2) Not publicly listed.[42]
(3) Based on a 1 year subscription and 500 judgments max.

Figure 2.71: Legal Databases Subscription Costs

In general, the prices of general databases are higher (above 1000 €) than those of specialized databases (below 1000 €).

At this stage of the research, we are about closing a research agreement of cooperation with La LEY[43] by which we will have access to its databases. Fig. 2.72 show the amount of stored judgements at December 2004. In this way, we will have the opportunity to explore during the next two years a significant volume of edited judicial rulings to check the meta-search engine of *Iuriservice II*.

[42] Based on: http://cein.es:81/ceinews.nsf/0/b64ae21cffc0700dc1256b03004afb7b?OpenDocument.
[43] The agreement is being signed by the UAB, La Ley and iSOCO within the framework of the SEKT Project.

L A L E Y N E X U S Diciembre 2 0 0 4	
JURISPRUDENCIA	
TRIBUNALES	
Tribunales Europeos	5.929
Tribunal Constitucional	14.752
Tribunal Supremo	200.384
Audiencia Nacional	26.668
Tribunales Superiores de Justicia CC.AA.	83.693
Audiencias Provinciales	120.357
D.G.R.N.	15.685
Resto Tribunales	14.608
TOTAL	**482.076**
LEGISLACIÓN	
ÁMBITO	**dic-04**
Internacional	3.713
Europea	5.265
Estatal	95.959
Autonómica	57.607
Resto	229
TOTAL	**162.773**
BIBLIOGRAFÍA	
	dic-04
TOTAL	**72.607**

Figure 2.72: Content of La LEY Databases in December 2004 (data provided by La Ley).

3 Legal Ontology

3.1 Legal Ontologies

The legal domain has been of interest to Artificial Intelligence since long. Pamela N. Gray (1997) has pointed out that the theory and the tools of Artificial Legal Intelligence have developed in corresponding leaps, with the following progression of themes: (i) legal language, (ii) deontic logic, (iii) rule processing, (iv) case processing, (v) stratification of reasoning, (vi) procedural reasoning, (vii) co-ordination of multiple tasks.

Legal ontologies have been a part of such a process. Many legal ontologies have been built so far. One current way of describing the actual state of the art is identifying the main current legal ontologies [RBCP, 2004] [GB, 2002] [VB, 1998]. (We offered already a more detailed description in [RBCP 2004]):

(1) LLD [Language for Legal Discourse: [Mc, 1989], based on atomic formula, rules and modalities;
(2) NOR [Norma: [St, 1991, 1996] based on agents behavioural invariants and realizations;
(3) LFU [Functional Ontology for Law: [V, 1995] based on normative knowledge, world knowledge, responsibility knowledge, reactive knowledge and creative knowledge;

(4) FBO [Frame-Based Ontology of Law, [K, 1995], [Vi, 1995], based on norms, acts and descriptions of concepts;
(5) LRI-Core Legal Ontology [BEPW, 2002], based on objects, processes, physical entities, mental entities, agents and communicative acts ;
(6) IKF-IF-LEX Ontology for Norm Comparaison [GPS, 2001], based on agents, institutive norms, instrumental provisions, regulative norms, open-textured legal notions and norm dynamics.

All these ontologies have been discussed within the intellectual community of AI and Law, and presented at the regular annual meetings of JURIX and ICAIL. Those are some of the specific projects in which legal ontologies have been worked out [W, 2003]:

- e-POWER (2001-2003). An International EC 5th framework project (IST-2000-28125) aimed at improving quality, maintenance and access to legislation by formal specification. Core (normative, legal) and domain ontologies (tax and pension regulations) are distinguished [EV, 2002].
- e-COURT (2001-2003). An International EC 5th framework project (IST-2000-28199) aimed at speeding up the search and retrieval of data in criminal trials by using multi-media data bases through inter- and intra-net. Ontologies of courtroom procedures and parts of the Dutch penal code are provided [BEPW, 2003].
- CLIME (1998-2001). An International ESPRIT project (P25.414) aimed at building intelligent legal information servers with a restricted natural language interface and WWW-based used interface. A large ontology for the domain of ship classification was built in this project.
- KDE (1999-2001). An international ESPRIT project (P28.678) aimed at developing a methodology and tools for the management and access to corporate knowledge and information sources. In this project the CLIME ontology was reused.

Independently from the the Semantic Web and its applications, legal ontologies are being applied in a variety of settings, which present typical problems that Breuker has summarized as follows [BCBG, 2005]:

- Creation of regulatory metadata and content standardization (e.g. Legal-XML/LeXML/MetaLEX, ADR/ODR-XML, etc.). Ontologies can be used for metadata creation or standardization, and existing metadata and standards can be used as sources for ontology building. Ultimately, some envision a future ontology-driven legal data management.
- Information extraction from legal documents: patterns of textual chunks or of syntactic constructs can be matched against semantic patterns derived from legal ontologies.
- Regulatory compliance: case matching against existing jurisprudence, compatibility of norms from laws pertaining to different time, orderings or systems, for comparison or harmonization purposes.

- Modelling legal reasoning (from the typically reified ontological viewpoint): the epistemological aspects of Law have been pointed out as a necessary target for legal ontologies. Therefore, case-based reasoning and reasoning with uncertainty may be hybridised with legal ontologies.
- Management of workflows based on legally-defined tasks.
- Decision-support for legal advice, e-Government, e-Commerce.

3.2 Epistemology and Legal Ontologies

Legal ontologies differ from other types of domain ontologies in two special features. The first one is the bulk of common sense notions that are carried out within the legal domain. Legal statutes, legal judgments and jurisprudence are written both in natural and in a more technical language. But practically all the common sense notions and connections among them that people use in their everyday life are embodied in the legal domain.

The second special feature has to do with the fact that the strategy of ontology building must take into account the particular model of law that has been chosen. This occurs in a middle level that is possible to skip in other ontologies based in a more contextual or physical environment.

When the task to be done involves a whole process, then it is possible to figure out the ontology following some particular cognitive models that can be used as templates to build the basic ontological concepts and their relationships. In this case only two levels are really needed: the domain conceptual level and the upper ontological level.
The following figures [3.1-3.3] show in which way a plant oil battery and a smart home environment have been plotted into cognitive maps leading to two ontological levels.

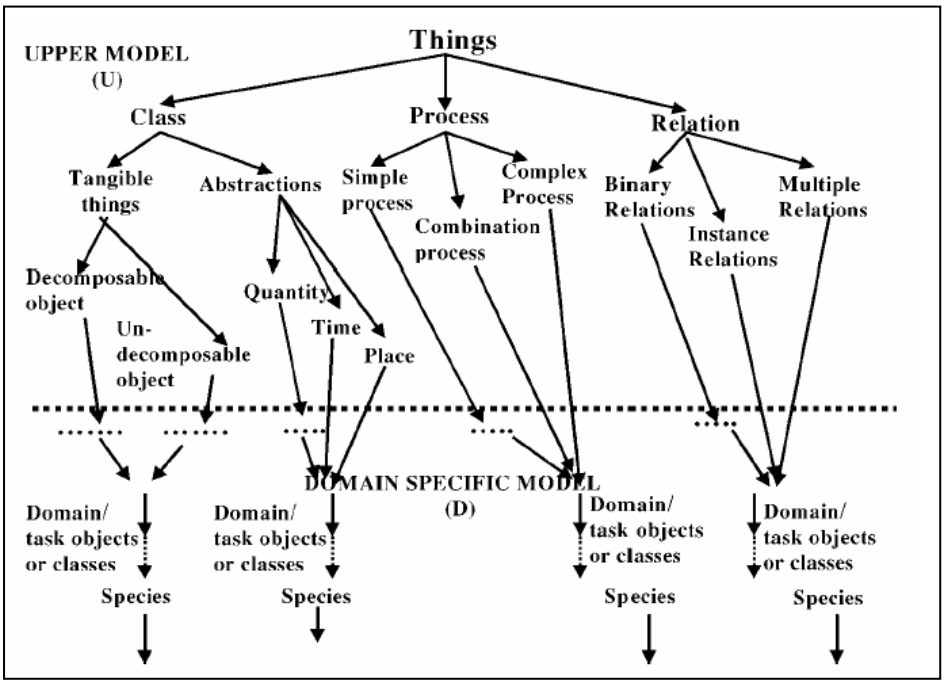


Figure 3.1: Two level ontology structure. Christine W. Chan [C, 2003: 276].

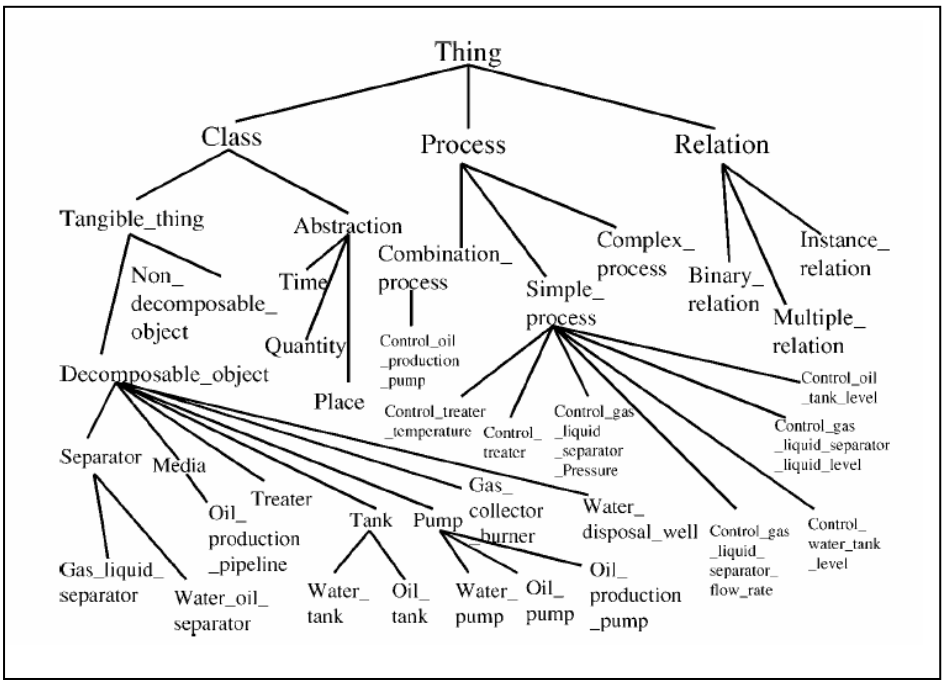


Figure 3.2: Classification hierarchy of classes in oil battery domain. Christine W. Chan [C, 2003:277]

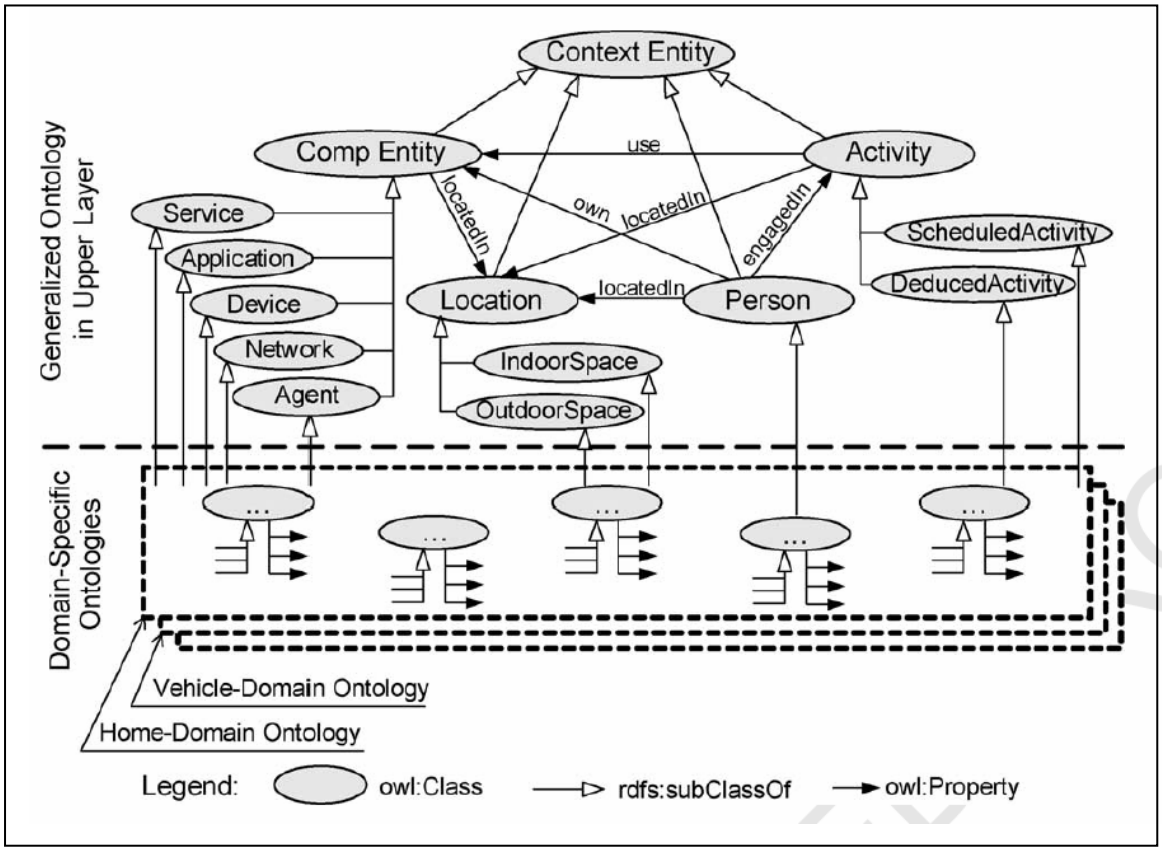


Figure 3.3: Class hierarchy of the upper ontology of a smart home environment. T. Gu, H. Peng Kung [GP, 2004].

In the legal field, the modeling process usually requires an intermediate level in which several concepts are implicitly or explicitly related to a set of decisions about the nature of law, the kind of language used to represent legal knowledge, and the specific legal structure covered by the ontology. There is an *interpretative* level that is commonly linked to general theories of law.

The former assertion requires further specification. This intermediate level is a well known layer between the upper top and the domain specific ontologies, especially in "practical ontologies":

> "An *interpretation* is the mapping (semantics) from one application instance (conceptual schema) syntactically described in some language into the ontology base, which is assumed to contain conceptualizations of all relevant elementary facts. (…).
> The interpretation layer constitutes an intermediate level of abstraction through which ontology-based applications map their syntactical specification into an implementation of an ontology 'semantics'." [JM, 2001]

We may also implicitly find this distinction between an ontology layer and an application layer in cognitive modeling, in which categories, concepts and instances are distinguished:

> "Cognitive informatics is the study of the cognitive structure, behavior, and interactions of both natural and artificial computational systems, and emphasizes both perceptual and information processing aspects of cognition. (…).
> Constructing the mental model of human expertise within the context of a particular problem-solving task is referred to as cognitive or conceptual modeling. (…).
> An ontology can also be regarded as a description of the most useful, or at least most well trodden organization of knowledge in a given domain". [C, 2003: 269-70]

But the most striking feature of the legal ontologies constructed so far is that the intermediate layer is explicitly occupied by a kind of high conceptual constructs provided by general theories of law instead of empirical or cognitive findings. Therefore, the link between epistemology and ontology is filled up with some intuitive or philosophical assumptions about the nature and function of law. Of course this is not necessarily a critical issue –there is a free space for epistemological assumptions in ontology building- but it seems to us that this fact may broaden the gap between users' needs and expected solutions. This lack of sociological or anthropological knowledge in ontology engineering has already been noticed [P 2001].

General theories of law seek to map the most general legal concepts into a single coherent body aiming to represent legal knowledge. In this way, several formulations have been focused. E.g. the difference between primary and secondary rules —

Herbert A. Hart—, the legal hierarchical levels —typical of the Hans Kelsen's "pure theory of law"— or, more recently, the assumption of the existence of an institutional dimension in which facts and rules differ from the behavioral dimension – MacCormick institutional theory of law, based on J.R. Searle's philosophy of language.

Law conceived (i) as a set of related static and dynamic norms, (ii) as a set of interrelated rights and duties, (iii) as a set of institutional rules and facts, or (iv) as a set of states of affairs, events and rules, has been a source of inspiration for legal ontology building.

In this sense, L.T. McCarty [Mc, 2002] has formalized Hohfelds' legal fundamental conceptions (1919) to model property rights (Fig. 3.4), J. Haage and B. Verheig [HV, 1999] have modeled a formal theory of law stemming from the causal or ruled link between events and state of affairs (Fig. 3.5), A. Gangemi et al. [GST, 2002] have conceptually represented the law implementation or application process as a relationship between legal normative descriptions and cases (Fig. 3.6), and G. Boella and L. van der Torre [BT, 2004] have constructed a normative multiagent system based on regulative and constitutive norms (Fig. 3.7).

Rights	Correlatives	Opposites
claim	duty	no-claim
liberty	no-claim	duty
power	liability	no-power
immunity	no-power	liability

Figure 1. Hohfeld's fundamental legal conceptions.

The basic strategy for a computational analysis of the ownership relation is: (1) to formalize Hohfeld's system in *LLD*; and (2) to write down the incidents of ownership in the formalized Hohfeldian language. To do this, we need to examine the modalities over actions that are available in *LLD*. The most prominent modalities are deontic: *permitted*, *forbidden*, *obligatory*. For expository purposes, we will write these in two ways. When a short form is needed, we will use (roughly) the notation in McCarty (1983): $\mathbf{P}\langle\phi \mid \alpha\rangle$, $\mathbf{F}\langle\phi \mid \alpha\rangle$, $\mathbf{O}\langle\phi \mid \alpha\rangle$. For example, $\mathbf{O}\langle\phi \mid \alpha\rangle$ means: "Under the condition ϕ, the action α is obligatory". When we want to write out a complete representation of the conditions and actions, however, we will use *LLD* syntax:

```
(Permit 'PE-1
  {condition . . . }
  {action . . . })

(Forbid 'FO-1
  {condition . . . }
  {action . . . })

(Oblige 'OB-1
  {condition . . . }
  {action . . . })
```

Figure 3.4: L.T. McCarty's formulation of the ownership relation based on hohfeldian concepts [Mc, 2002].

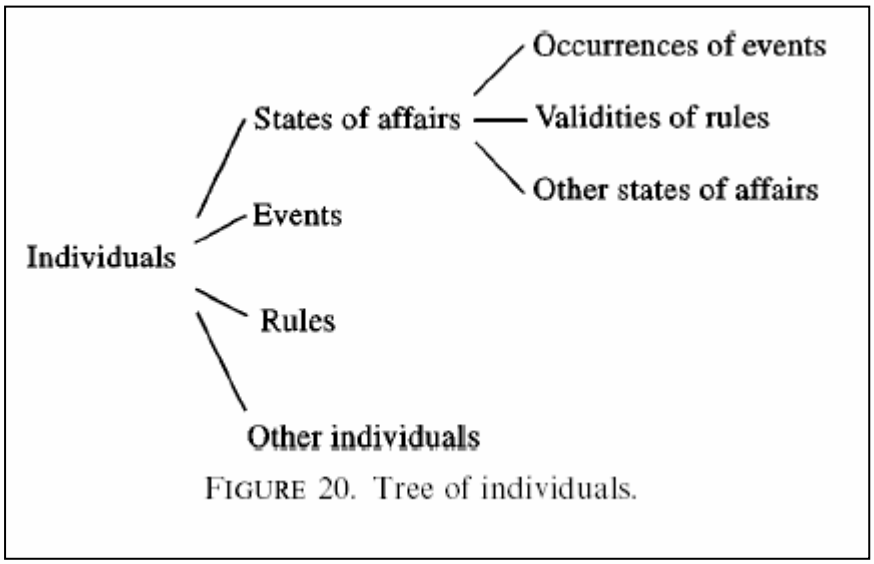


FIGURE 20. Tree of individuals.

Figure 3.5: Individuals: primitives of the top model legal theory. J.Hage, B.Verheig [HV, 1999].

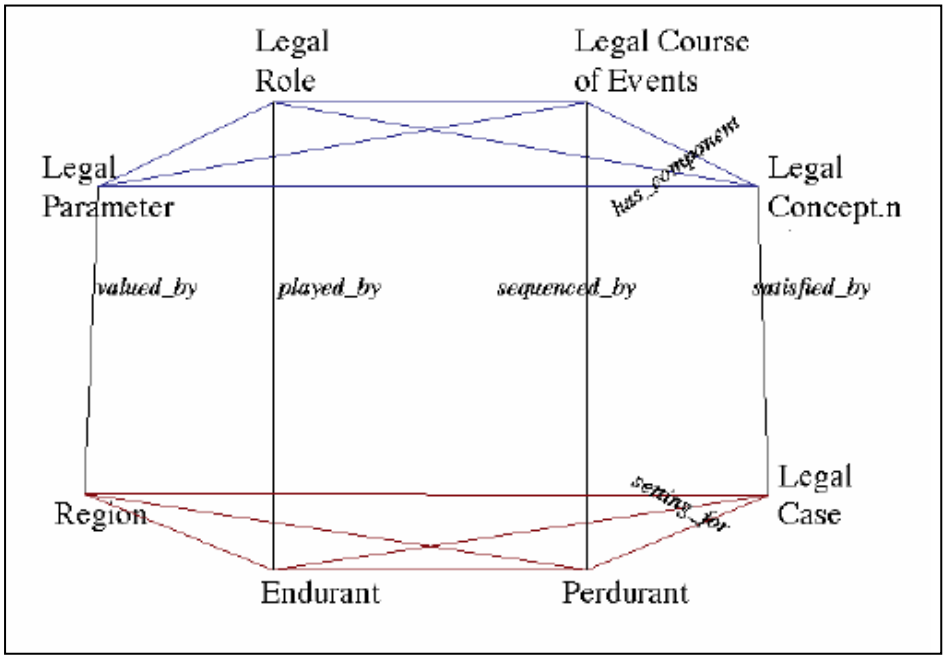


Figure 3.6: A conceptual template for legal descriptions (conceptualisations) and situations (cases). The current version of the *legal core ontology* assumes the distinction between the legal and non-legal worlds, and maps it into the DOLCE+ distinction between descriptions (in this domain *legal conceptualisations*) [laws, norms, regulations, crime types, etc.] , and situations (*legal facts* or *cases*). Aldo Gangemi, Maria T. Sagri, Daniela Tiscornia [GST, 2002].

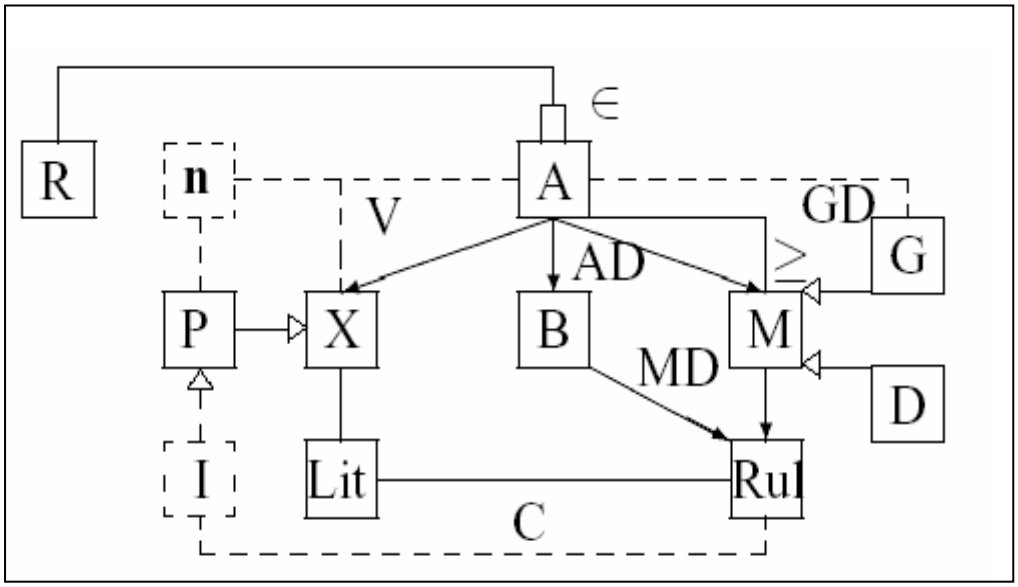


Figure 3.7: Conceptual model of the normative multiagent system. The multiagent system is represented in straight lines, while additions for the normative system are represented in dotted lines. The authors distinguish between regulative norms that describe obligations, prohibitions and permissions, and constitutive norms that regulate the creation of institutional facts as well as the modification of the normative system itself. (A) means Agents. Beliefs (B), desires (D) and goals (G) are represented by different sets representing the epistemic and motivational states of the agent. The variables (*X*) are either *decision variables* of an agent, which represent the agent's actions and whose truth value is directly determined by it, or *parameters* (*P*), which describe both the state of the world and *institutional facts*, and whose truth value can only be determined indirectly. Guido Boella, Leenert v. der Torre [BT, 2004].

The interpretative middle level in which all fundamental concepts are defined is usually known as the *Legal-Core Ontology*. J. Breuker and R. Winkels [BW, 2003] have recently distinguished between legal ontologies originally based on normative knowledge (legal theory) and legal ontologies –or "with an ontological flavour"- in which modalities play the role of knowledge categories. This would be the case for McCarty's LDD or for deontic logic formulations applied to the legal domain. However, in both cases, the fundamental concepts are epistemologically set within a Legal-Core Ontology, that is to say, the ontological representation of basic legal knowledge.

Figure 3.8 shows Valente's and Breuker's representation of the three level ontology for FOL. Between the upper and the core ontology there is an overlapping area in which an "anchoring" process is produced into the foundational (common) level.

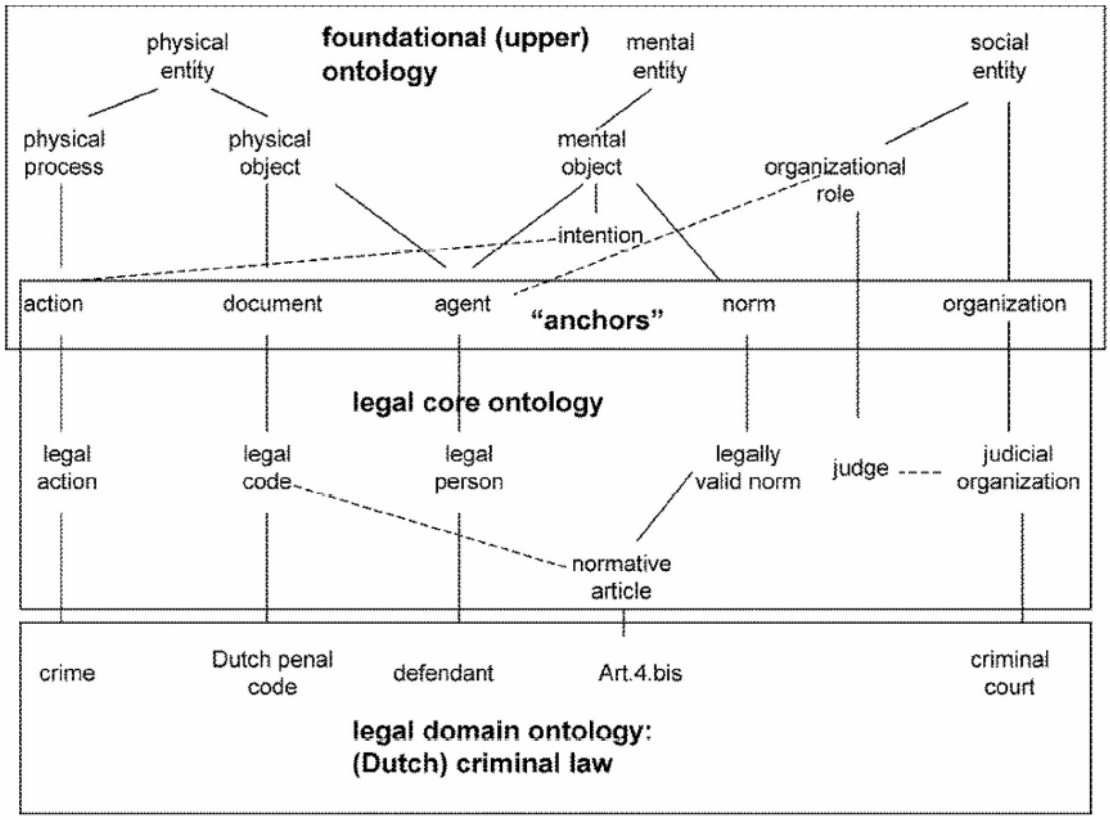


Figure 3.8: The "cascading" three level structure of the Functional Ontology for Law (FOL). The middle level has been developed as LRI-Legal Core Ontology. In LRI-Core, social roles and social functions are distinguished. Social functions are external roles of organizations. Social roles make up the functional internal structure of an organization [BW, 2003]. André Valente, Jost Breuker, Bob Brouwer [VBB, 1999].

We are not going to discuss this FOL schematic representation now. We only would like to point out that the link between the instances of concepts in the domain-specific layer and the LRI-Core is epistemologically guided by a representation of legal knowledge which is not based into an inductive acquisition learning process. We think that there is room for a complementary approach based on data about the professional use of legal knowledge. This is why we have turned to the professional use of procedural and legal concepts to build our middle-out ontology. Professional knowledge has to absorb dogmatic or more doctrinaire formulations and transform them into a more flexible knowledge to adapt itself to changing contexts and incoming cases.

3.3 Ontologies of Professional Legal Knowledge (OPLK)

3.3.1 Professional Knowledge (PK)

Professional knowledge (PK) as encoding a specific kind of knowledge related to particular tasks, symbolisms and activities has been described many times in the literature. Actually, the sociology of professions is a well established subdomain of the sociological studies.

Since Berger and Luckman (1966), sociologists have seen the development of a professional identity as an aspect of the development of a "secondary socialization". Primary socialization refers to a person becoming a part of society through the educational process (child learning). Secondary socialization comes later with the internalization of "institutional worlds" involving the acquisition of specialized knowledge (*professional knowledge*).

Recent studies have focused on: (i) the organizational negotiation of professional identities, (ii) their regulatory character, (iii) the globalized hi-tech side of present-day professions [DW 2002].

Legal professions have changed dramatically under the impact of globalization. In a global world, the creation of a new *lex mercatoria*, the expansion of law firm sizes, the consolidation of law firms as professionally managed business organizations and the leading position of international American and UK firms change the contexts of practicing law and the ways of litigation and conflict resolution. Studies show the recent rise of "transnational legal networks" and their link to strategies of knowledge management. The expected percent growth of US law firms up to 2007 is 21%, with an expected value of 200 billion $ [CP, 2005].

Professional knowledge has been defined as possessed by professionals which enables them to perform their work with a quality standard [E, 1992]. Professional knowledge then includes: (i) propositional knowledge (knowing that); (ii) procedural knowledge (knowing how); (iii) personal knowledge (intuitive, pre-propositional); (iv) principles related to morals or some kind of deontological code.

A *knowledge base* is generally distinguished which is being developed through practice and experiential learning. But do notice that globalization produces pervasive and unexpected paradoxes.

It would be inaccurate, for instance, to talk about a unified or common professional field. The orientation to the market of the ancient liberal professions and their trend to commodate their outcomes (medicine, law…) produce big differences inside the same collective. Moreover, in the law field, there is a growing gap between the institutionalized professions closer to the state (judges, prosecutors) and those operating within the legal markets.

We could say that a counsel shares with the judge, the prosecutor or other court staff only a portion of the legal knowledge (very likely the legal language and the most general acquaintance of statutes and previous judgments). But there is another kind of legal knowledge, the one having to do with personal behavior, practical rules, corporate beliefs, effect reckoning and perspective on similar cases, that remain implicit and tacit within the relation among judges, prosecutors, attorneys and lawyers. It may take long, even for an experienced lawyer, to flesh out what kind of professional guidelines are being followed by a judge in his daily management of cases.

3.3.2 Professional Legal Knowledge (PLK)

The legal ontologies described above have been built up with several purposes: information retrieval, statute retrieval, normative linking, knowledge management or legal reasoning. Although the legal domain remains very sensitive to the features of regional or national statutes and regulations, some of the Legal-Core Ontologies (LCO) are intended to share a common kernel of legal notions. However, LCO remain in the domain of a general knowledge shared by legal theorists, national or international jurists and comparative lawyers.

Our data indicate that there is a kind of specific legal knowledge, which belong properly to the expert domain and that is not being captured by the current LCO.

Consider the following problems, extracted from different kinds of transcriptions of the research protocols, contained in the Figures below:

"I have the following problem, let's see if you come up with something: one woman files a suit (she went to hospital to get care for the bruises) but then she forgives her husband, tells us that they both were drunk that night but are very happy (to show us how happy they are she even insists on remaining in the room while he declares). She keeps saying no way, she is not going to denounce her husband, she has forgiven him.
Since it's a public offence I go ahead and then the prosecutor (fiscala [fem.]) gets angry with me because she appoints him to court (*lo persona*) and wants me to appoint her wife to instruct her on her rights (*instruirle de sus derechos*). The issue has no objective criminal entity (*entidad penal objetiva*); to criminalize those little things seems to me really nonsense, it may even be worse regardless of the prosecutor moving forward." [May 2004, personal communication]

Figure 3.9: Literal transcription of a practical procedural problem on gender violence. Pompeu Casanovas. [personal e-mail communication , May 2004, reproduced with the permission of the sender]

"In villages there is no attorney. You have to decide whether to phone him or not and if you phone him, then you have to tell him on the phone the whole story, and on the phone he cannot see the file either, and all that has been done; of course you can tell him what has been done, but he does not see things as though he was there with you, either, as if he had a look at the file with you, and of course this has caused us problems quite often, in cases where I thought the case deserved prison but the attorney did not actually want to come to the court appearance. Well, he thought that with the information I gave him, well he as not going to petition imprisonment. Then, you cannot put him into prison, you take an ordinary declaration and do not convene the appearance, or at most you call the appearance and you write down that the Attorney does not come if you want to remain safe"
Question: What decision is to be taken when the attorney does not attend and yet the judge considers that he must order imprisonment?

Figure 3.10: Literal transcription and reformulation of the competency question. Cristina Urios [Mallorca, summer 2003, previous ethnographic fieldwork]

"Today, a person came to me when I was on duty with a piece of paper that said ", that is, a piece of paper from the doctor that said "this person has to be taken to the psychiatric hospital urgently"; in theory it is them who should take him to hospital and yet they do not do it, they come to us with the piece of paper. It is an unclear issue. There is a protocol, but it is not clear, either. If I give my authorization, I should see him. The problem is that the mother wants me to take him to hospital and the doctor...".

Question: When an ill person enters a psychiatric hospital, which is the job of the judge on duty, to ratify the urgent admission carried out by the doctor after verifying the state of health of the person admitted into hospital or rather to order admission into hospital without seeing the person?

Figure 3.11: Literal transcription and reformulation of the competency question. Cristina Urios [Mallorca, summer 2003, previous ethnographic fieldwork]

1.I have ordered a injunction of protection [orden de alejamiento] in favour of a woman, and after some days she comes back asking me to cancel or remove it. What should I do? Do I always have to cancel or remove it? What may I advise her?

2. There is a couple and an injunction of protection against the husband, but the police knows that they are living together and they told me that. Any time she gets angry with him or they have some trouble she uses the injunction, the police detains him and I have to organize a hearing…just to find them together again next morning. What can I do? Can I modify or cancel the injunction?

3. I have been asked to dictate an injunction of protection, but this implies to leave a man on the street without a living. What can I do? How can I help this man?

4. One woman asks me for an injunction of protection because of psychological abuse, but it turns out that she's never gone either to the psychologist or the psychiatrist. Should I dictate an injunction of protection?

Figure 3.12: Reformulated questions extracted from literal transcriptions of the questionnaires answered during the fieldwork [Francisco Ramos, Courts of Galicia, May 2004]

Technically speaking, these problems are not complex. However, they are difficult to solve. The judges' original questioning cannot be answered by simply pointing out to a particular statute or legal doctrine. This is not only an issue of normative information retrieval.

Judges are experts: they take for granted the acquaintance with legal texts, textbooks or former legal decisions. What is at stake here is a different kind of legal knowledge, a *professional legal knowledge* (PLK) [BCBRCP, 2004]. What judges really seek are some clues, some hints or well-grounded practical guidelines that refer to the problem they have before them when they ask the question or start the query.

In this regard, the design of legal ontologies requires not only to represent the legal, normative language of written documents (decisions, judgments, rulings, partitions…) but also those chunks of professional knowledge in which the daily practice at courts consists of.

From this point of view, professional knowledge of a legal topic (such as e.g. gender violence) involves a particular knowledge of: (i) statutes, codes, and legal rules; (ii) professional training; (iii) legal procedures; (iv) public policies; (v) everyday

routinely cases; (vi) practical situations; (vii) people's most common reactions to previous decisions on similar subjects.

We may point out several properties of *Professional Legal Knowledge* (PLK). PLK is: (i) shared among members of the judicial professional group (e.g. judges, attorneys, prosecutors…); (ii) learned and conveyed formally or most often informally in specific settings (e.g. the Judicial School, professional associations –the Bar, the Judiciary…-); (iii) expressible through a mixture of natural and technical language (legalese, legal slang); (iv) non-equally distributed among the professional group; (v) non-homogeneous (elaborated on individual bases); (vi) universally comprehensible by the members of the profession (there is a sort of implicit identification principle).

One of the main features of PLK is that it is context-sensitive, anchored in courses of action or practical ways of behaving. In this sense, it implies: (i) the ability to discriminate among related but different situations (e.g. when is an injunction of protection really needed or required to prevent a woman of being injured or murdered by her husband?); (ii) the practical attitude or disposition to rule, judge or make a decision; (iii) the ability to relate new and past experiences of cases; (iv) the ability to share and discuss these experiences with the peer group.

We have already described the original process of knowledge acquisition for the two prototypes *Iuriservice-I* and *Iuriservice-II* (see sections 2.4.2.1.1. and 2.4.2.1.2.). The problem we have to tackle now is how *to represent* such a knowledge in a feasible ontology that preserves its main features.

3.3.3 Ontologies of Professional Legal Knowledge (OPLK)

In order to build OPLK, we believe that we have to take into account the kind of *situated knowledge* that judges put into practice when they store, retrieve and use PLK to make their most common decisions. We use "situated knowledge" in a similar way in which W.J. Clancey and T.Menzies talk about "situated cognition": the concrete use of knowledge which is partially shared and unequally distributed through a certain "community of practice" who is able to use and reuse this same knowledge while transforming it.

> "*Situated cognition* is an approach for understanding cognition that seeks to relate social, neural, and psychological views. From the social perspective, situated cognition provides insights about *the content of knowledge*, namely how people conceive of what they are doing in terms of their contribution to a community of practice (Wenger, 1997) and how this affects their attention and priorities over time. From the neural perspective, situated cognition provides insights about *the physical structure of knowledge*, namely how perception, conception and motor action are related through a self organizing coordination process with a memory. From a psychological perspective, situated cognition provides insights about *how behaviour is improvised* by resequencing and recomposing previous behaviours." [CSSH, 1998: 836]

> "*Situated cognition* explores a host of fundamental assumptions about artificial intelligence and the process of building expert systems. These considerations are much more complicated than the traditional symbolic view of knowledge and include the following: (1) The social context of knowledge, e.g. decision making is grounded in organizational identity and norms. (2) The structural aspects of human memory that allow for self-organization and reconstruction of ideas, e.g. behaviour may be conceptually coordinated without describing either the world or the behaviour. (3) The manner in which previously articulated heuristics, designs and policies are reinterpreted in practice.
> Situated problem solving often involves the reconception of meanings and goals, especially in multidisciplinary pursuits. For example, the relation of medicine, economics, lifestyles and ethical policy in the practice of medicine. Consequently, we cannot expect to write one symbolic model once and reuse it in all situations to come. Symbolic models, such as those found in expert systems, are *tools* that people can use, not the mechanism by which human perception, conception and action are actually coordinated in the brain."
> [MC, 1998: 767-8]

The main idea is that PLK is always *situated* in a particular context in which the judge or the lawyer needs to complete the information they possess about a particular case or problem to trigger or put into practice the basic knowledge that they already have. In this sense, they do not need to be provided with a complete legal reasoning, but only with some reliable information that they may use as a comparative parameter. They seek, so to speak, "another opinion", an external interlocutor to follow the full reasoning process that they build up any time they have to make a new decision. This is the reason why they use to comment the case with their peer colleagues: not really to be helped in the decision making but to double check the decision they are going to make.

Legal reflection is eventually almost an automatic process in which lawyers or judges are involved without being fully aware of the devices they are using. But our data analysis makes clear that this is a *collective and interactive process*, even if it is usually performed on individual bases.

Other related concepts close to "situated knowledge" are the ideas of "situated communities", "situated meaning", "organizational memory" and "corporate ontologies".

On the one hand, for all practical purposes there is no such thing as absolute *meaning*: everything must ultimately be the result of agreements among human agents such as designers, domain experts and users [JM, 2001: 3]. On the other hand, in ontology knowledge modeling a concept is neither a class or a set: the concepts which represent the term's meaning are structured into binary trees based on couples of opposite differences [Ro, 2000: 188].

OPLK models the situated knowledge of professionals at work. In our particular case we have before us a particular subset of PLK belonging specifically to the judicial field. Therefore, we will name our conceptual specifications of the knowledge contained in our empirical data Ontology of Judicial Professional Knowledge (OJPK).

To build the ontology of the second prototype (*Iuriservice-II*), we have used so far only the terms and related concepts used by the interviewed young judges in their answers transcripted in our analysis protocols. However, Figures 3.13 and 3.14 show that the answers of the more experienced Magistrates of the Judicial School to the professional questions formulated by the young ones should be also taken into account in a next stage to refine the central concepts of the intermediate layer.

The doctor has sent an urgent internment to me ¿What am I supposed to do with the doctor?

Art. 763 LEC establishes as a general rule that the internment of a presumably disabled person requires judicial authorization. Nevertheless, it also considers the exception of the urgency of the internment. In this case the judicial authorization comes after the measure being adopted, which needs to be communicated to the judge as soon as possible and, in any case, no later than 24 hours. In those cases, the doctor is in the best position to evaluate the presence of physical and psychical circumstances requiring internment. Thus, in this particular case the doctor could have ordered the internment and inform the judge as soon as possible so that the latter could ratify the measure. However, in many occasions doctors do not order internment, either because they fear to assume responsibility or due to difficulties at the effective execution of the measure. In those cases they prefer to ask for a judicial internment, communicating its urgent character.

Answering to the question raised, if the judge considers that, given the circumstances of the case, doctor's behaviour has been incorrect, the wisest way to do is to have a conversation with him in order to establish the way to proceed in further cases. By no means it can be forget that the prevalent interest is the health and integrity of the presumably disabled person and the one of the people who could be affected by his acts. Consequently, and regardless of the doctor behaviour, the judge has to decide the sooner the better on the petition of internment.

Figure 3.13: Judicial answer from the Judicial School team to young judges' questioning.

I have a phone call from the doctor telling me that he has someone who is in a very bad situation, who probably needs internment, and he asks me for an order of internment by phone. Can I do it?

All judicial decisions, and particularly those with an especial relevance for the rights and freedoms of individuals—like the internment of a presumably disabled person—need to be written, following the provisions of procedural norms. Exceptionally and very restrictively, the judge may adopt a verbal decision in extremely urgent and serious cases, and put it in a document later on.

In the present case, the doctor may adopt the decision on his own if the internment is urgent, following the provision of art. 763 of the LEC. He will have to communicate the internment to the judge as soon as possible and, in all cases, no later than 24 hours, so that the judge can ratify the measure as the aforementioned article establishes. If the doctor does not order the interment—a circumstance which would cast a doubt on the presumed urgency—the right way to proceed for the judge is to produce a written authorization, as stated in art. 763 LEC, in the least possible time, according to the presumed urgency of the case.

Figure 3.14: Judicial answer from the Judicial School team to young judges' questioning.

3.4 Methodology

As it has been presented in Chapter 2, several modeling changes have been introduced towards the construction of the ontology used by the second *Iuriservice* prototype. Nevertheless, the "competency approach" [GF, 1995], adopted for the construction of the ontology used in *Iuriservice-I,* continues to be followed as it represents the best method to capture professional knowledge. Thus, we have extracted significant terms for the ontology and their relations from discussions based on the 756 competency questions provided by the new knowledge acquisition approach. The ontology is still under construction, but more than 200 questions have already been discussed.

3.4.1 Competency questions discussion method

TextToOnto and ALCESTE provided a good basis for regarding some terms as significant and their conclusions have proved to be really useful to both feed and control the modeling process. However, the method used in building the ontology has focused on the discussion within the UAB legal experts team over the terms which appear on the competency questions. This method has several phases.

First, it basically consists in selecting (underlying) all the nouns (*usually* concepts) and adjectives (*usually* properties) contained in the competency questions. Below, there is an example of selected terms (in bold) in some competency questions.

> ¿Cuál es el tratamiento de las **denuncias** manifiestamente inverosímiles o relativas a **hechos** que evidentemente carecen de tipicidad?
>
> ¿Y si se trata de una **querella** que reúne todos los demás **presupuestos procesales** pero los **hechos** objeto de la misma carecen de relevancia penal o manifiestamente falsos?
>
> ¿Qué ocurre si **comparece** en el **juzgado** una **persona** que quiere **denunciar hechos** difícilmente creíbles, sin relación entre sí, dudándose por el **juez** de la capacidad mental del **denunciante**?
>
> ¿Ante quién debe **interponerse** el **recurso de reforma** contra la **prisión**, delante del **juez de guardia** o del **juez** que dictó el correspondiente **auto de prisión**?

Once the terms had been identified, the team discussed the need to represent them within the ontology and their organization within taxonomies. The relevant relations between those terms also have to be identified (mainly *is_a* and *instance_of*).

Accordingly, we followed the *middle-out strategy* [GCF, 2002]. With this strategy, the core of basic terms are identified first and then they are specified and generalized if necessary.

As an example, and in relation to the competency questions analyzed above, modelers considered that the concepts *auto* [interlocutory decision], *recurso* [appeal], *demanda* [private/civil lawsuit] and *querella* [public/criminal lawsuit] needed to be represented in the ontology. Moreover, a concept *documento* [document] had to be created as all terms: *auto*, *recurso*, *demanda* and *querella* describe documents. The result was the construction of a more general concept from those specific terms found in the competency questions.

However, the team also agreed that *demanda, auto, recurso* and *querella* were not only instances of *documento* but also constituted a specific class of documents used only within the judicial process. For that reason, *documento_processal* [procedural document] had to be created as a subconcept of *documento*. At the same time, there are different types of appeals and court orders stated in the questions, that have to be considered instances of *recurso* and *auto*. In this case, the terms where specified, not generalized. This is a clear example of the use of the middle-out strategy in the legal case study ontology.

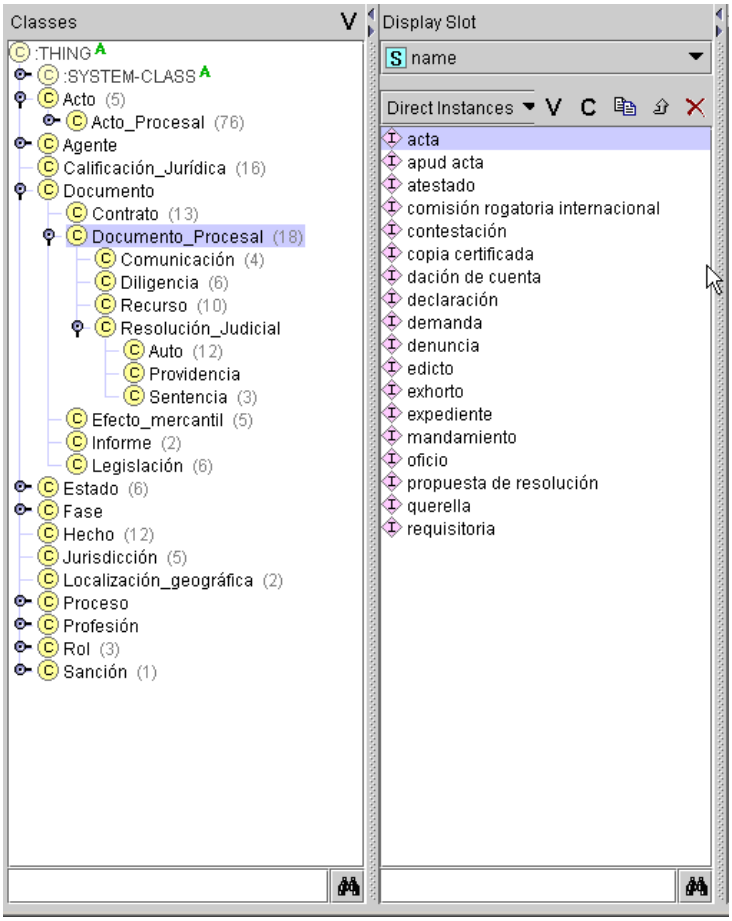


Figure 3.15: Screenshot Documento Procesal in Protégé

Furthermore, some other relations (different from *is_a* and *instance_of*) were also identified: someone creates those documents (*juez*, *denunciante*, *persona*), thus *document has_author*.

However, difficulties in reaching consensual decisions and the lack of traceable lines of argumentation was slowing down the construction of the ontology.

3.4.2 DILIGENT argumentation framework

The ontology modeling process was quite slow due to the difficulties faced by the team in reaching consensual decisions. For that reason, the introduction in the engineering process of the argumentation framework DIstributed, Loosely-controlled and evolving Engineering of oNTologies (DILIGENT), based on the rhetorical structure theory [PST, 2004], offered a reliable basis for a controlled discussion of the arguments in favor and against modeling decision. The introduction of DILIGENT proved the need to count on evaluation measures for the decision-making process within ontology design and the effect of the use of DILIGENT pro and against arguments (such as: elaboration, evaluation/justification, alternatives, examples and counter examples) was the speeding up of the modeling process, as decisions were more easily reached. Below there is an example of the argumentation stack used to establish *documento_procesal* [procedural document] as a concept of the OPLK, derived from the competency questions analyzed above.

Documento
Pro:Justification The questions contain different types of documents that should be included in the ontology as they are specific for judicial settings.

Pro:Example *Denuncia, demanda, recurso* y *auto* are documents that should be represented.

Pro:Elaboration These documents should not be only considered instances of *Documento* but also they should be included in a specific class of documents, those part of a judicial process.

Documento_Procesal
Pro:Justification The terms in the questions refer to a certain kind of documents, those documents produced in relation to a judicial process. So, that subclass of *Documento* should be created.

From AIFB partners (WP7) it was also suggested that the modelers could use a wiki tool to visualize, organize and trace the arguments used in the discussion. This tool was set up on the SEKT Project web (internal pages) and it proved to be very useful. The legal case ontology discussion wiki made all decisions transparent, traceable and available to all members of the team at the same time. However, the tool did not provide several features such as: visualization of the graphical representation of the ontology being built or a system of e-mail notifications when arguments had been added. To solve that lack of graphical visualization, we extended the wiki with snapshots from the relevant parts of the ontology build with KAON Oi-Modeler. Nevertheless, the lack of a notification system and of organizative procedures proved to be a problem with regards to regulate the input of comments.

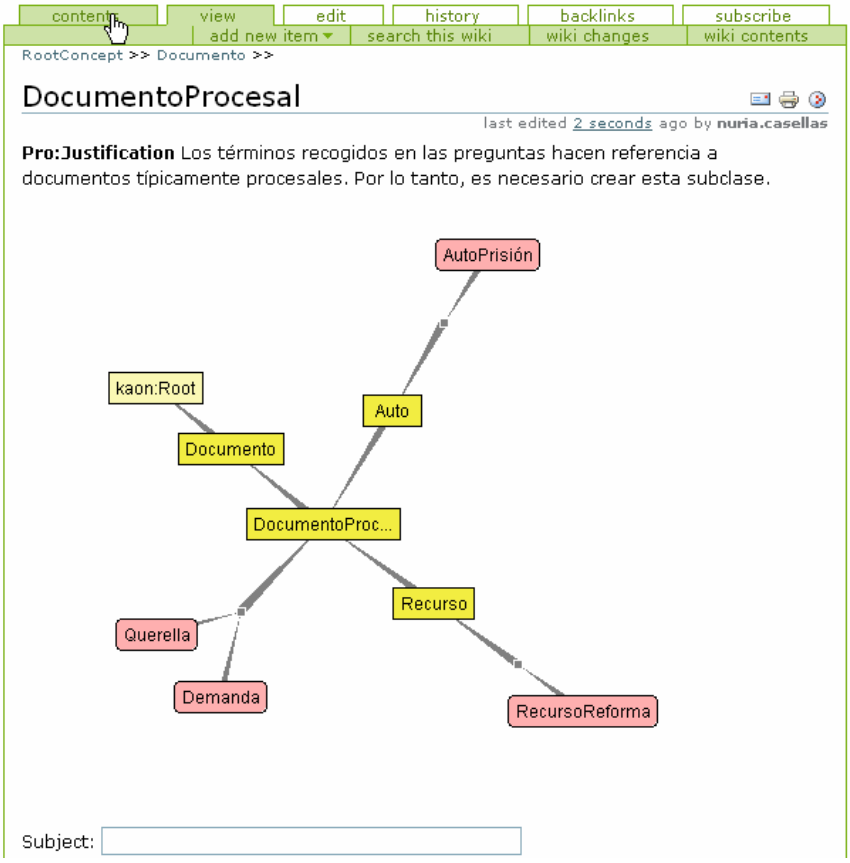


Figure 3.16: Screenshot of the legal case study wiki in the SEKT webpage

Finally, the partners from WP7 also suggested that in order to build a consistent ontology we could use the OntoClean approach, a "methodology for testing the ontological adequacy of taxonomic links" [GW, 2002]. Once the ontology has been build to a certain extent, this approach "cleans it up", meaning that the OntoClean systematization establishes ontology modeling constraints to analyze the efficiency and rearrange the taxonomy of nouns of our ontology. The OntoClean approach was also used in the restructuring of WordNet's Top Level ontology [OGGM, 2002].

These constraints take into account, for example, *essence*, *identity*, *rigidity* and *unity* as properties of a concept; then, when given *p* and *q*, where *p* subsumes *q*:

1. If *q* is anti-rigid, then *p* must be anti-rigid
2. If *q* carries an identity criterion, then *p* must carry the same criterion
3. If *q* carries a unity criterion, then *p* must carry the same criterion
4. If *q* has anti-unity, then *p* must also have anti-unity [GW, 2004].

One important thing that has to be remembered when developing our ontology of professional legal knowledge is the need to avoid the confusion between *identity* and *unity* in relation to modeling the judicial process. As Guarino and Welty point out, time intervals are not time durations, time intervals *have* time durations; "the duration is a component of an interval, but it is not the interval itself" [GW, 2004].

Due to time constraints, the OntoClean analysis hasn't been performed in full yet and will continue to be implemented during all the development of the ontology of professional judicial knowledge.

3.5 SEKT Ontology of Professional Judicial Knowledge (OPJK)

3.5.1 Introduction

The Ontology of Professional Judicial Knowledge developed by the legal case study team is learnt from scratch out of the competency questions posed by the judges during their interviews. Modeling this professional judicial knowledge demands the description of this knowledge as it is perceived by the judge and the abandonment of dogmatic legal categorizations. TextToOnto and ALCESTE have proved to be useful tools to visualize and analyze this knowledge (or lack of it) and to show that it is organized within distinct subdomains (domestic violence, on-duty period, procedural doubts, imprisonments,…). However, modelers have not used any automated or semi-automated ontology learning software, and have only relied on the methodologies described above.

The Ontology of Professional Judicial Knowledge has been extracted from the selection of relevant terms from nearly 200 competency questions and has, currently, nearly 50 concepts, 100 relations and more than 300 instances. This is result of a choice to minimize the concepts at the class level when possible in favor of creating instances and relations.

Nonetheless, the integration of the legal case study ontology into PROTON (Proto Ontology),[44] as part of the integration of SEKT technology, has created constrains towards the engineering process. This integration implies that the Ontology for Professional Judicial Knowledge should include the System Module and the Top Module from PROTON.

The System Module includes *Entity, EntitySource, LexicalResource, Alias, SystemPrimitive, TransitiveOver* and the Top Module includes *Abstract, Agent, ContactInformation, Document, Event, Group, Happening, InformationResource, JopPosition, Language, Location, Number, Object, Organization, Person, Role, Situation, Statement, Topic, TimeInterval*, and their correspondent relations. PROTON is a domain independent ontology and the specificity of the OPJK might require rearrangements. For that reason, although it is important to keep this two Modules in mind, it is essential for the OPJK to model judicial knowledge as perceived by judges and that point of view has to be maintained when possible.

At the moment, the ontology is being built without considering in depth its connection to PROTON super-classes, although the System Module can be fully integrated and some classes such as *Entity, Agent, Document, Event, Organization, Person, Role and TimeInterval* are either being used already or could be easily integrated.

[44] http://proton.semanticweb.org/

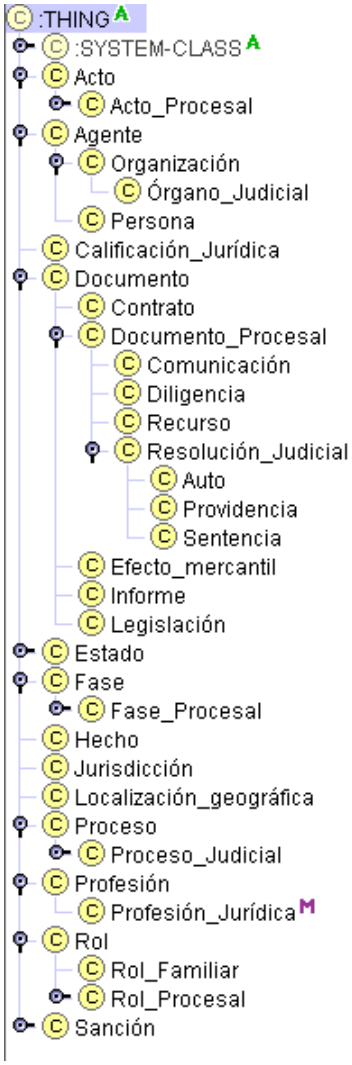


Figure 3.17 Screenshot of the current status of the OPJK in Protégé.

3.5.2 OPJK classes

Different top classes of the domain ontology have been identified: *acto_procesal*, *órgano_judicial*, *calificación_jurídica*, *documento_procesal*, *fase_procesal*, *jurisdicción*, *proceso_judicial*, *profesión_jurídica*, *rol_procesal*, *rol_familiar* and *sanción*.

- *Acto_procesal* [procedural act] is a subclass of *acto* [act] and represents a specific action taking place in the course of a judicial procedure. A subclass of *acto_procesal* is *acto_de_comunicación* [communication act], a class that includes all those acts of communication made by the court.

ACTO
--- ACTO_PROCESAL
--- levantamiento
--- Vista
--- ...
--- ACTO_DE_COMUNICACIÓN
--- citar
--- emplazar
--- notificar
--- requerir
--- ...

– *Órgano_judicial* [court] is a subclass of *agente* [agent], who can perform actions with or without consciousness. It is a subclass of *organización* [organization] and can perform actions with or without consciousness. *Persona* [person] is also a subclass of agent.

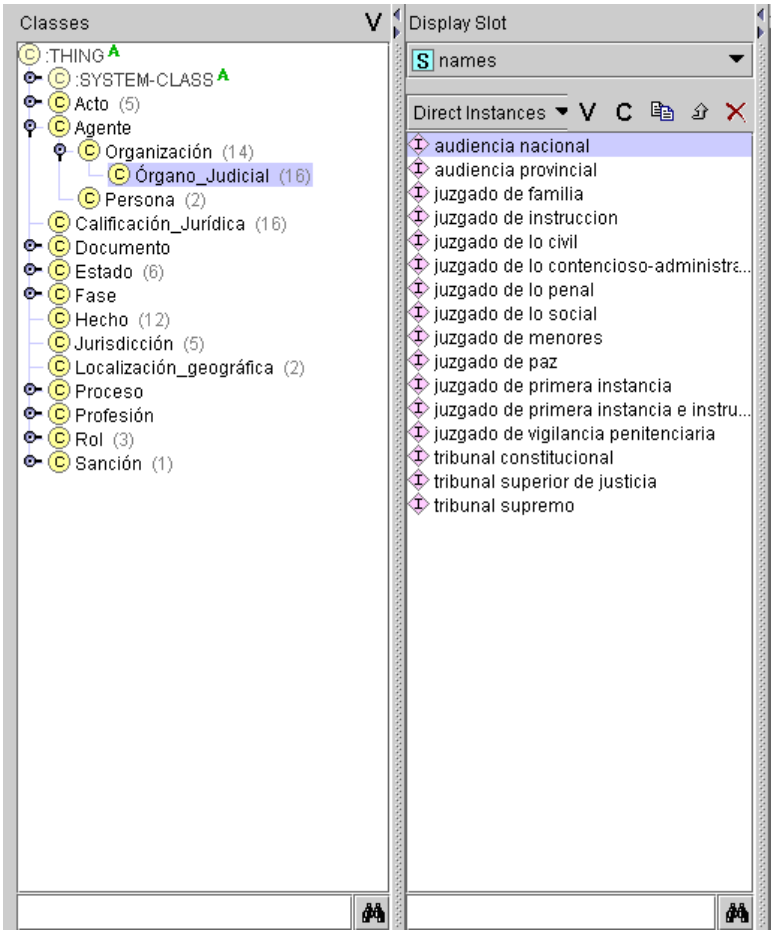


Figure 3.18: Screenshot of instances of órgano_judicial

– *Calificación_jurídica* [legal status] is a necessary class which consists of all those types of crimes, felonies, misdemeanors or legal status regulated by norms or established by final rulings.

CALIFICACIÓN_JURÍDICA
- homicidio
- hurto
- malos tratos
- violencia doméstica
- ...

– *Fase_procesal* [procedural phase] is an important concept for the OPJK ontology as it represents the time phases in relation to the judicial process. This concept is subclass of *Fase* [phase] and has several subclasses itself:

FASE_PROCESAL
- FASE_DE_DECLARACIÓN
  - fase de alegaciones
  - FASE_DE_PRUEBA
    - admisión de la prueba
    - proposición de la prueba
    - práctica de la prueba
  - fase de conclusiones
- fase de impugnación
- fase de ejecución (civil)
- fase de instrucción
- fase de juicio oral
- fase de ejecución (penal)

– Accordingly, *proceso_judicial* [judicial process] is a key concept for the OPJK ontology, as most of the questions are somehow related to procedural problems during on-duty periods or during normal opening hours.

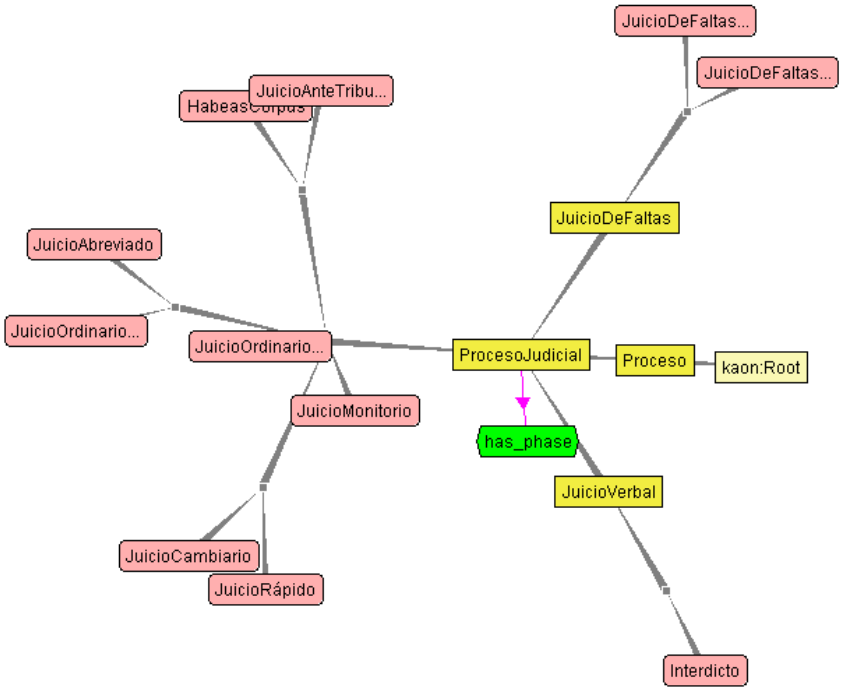


Figure 3.19: Screenshot of proceso_judicial modeled in KAON.

– *Rol_procesal* [procedural role] is a class of role. A role is the part that an agent plays in a specific situation. This class was newly introduced in Proton Top Module after its discussion in a working session within SEKT. Proton inherited KIMO classes such as *Situation*, *Profession* and *Agent* but during the competency questions discussion, the legal case modelers found out the importance of this concept altogether with *profesión_judicial* [legal profession].

  The need for the *role* concept within the legal domain had also been contemplated in other relevant legal ontologies. In Breuker's et al. Legal Core Ontology, the LRI-Core is equipped with *role*, a subclass of *mental_entity*, described as a functional view on a *physical_object*, *agent_behaviour* or *mental_process*. For these authors, *roles* are played by *persons* who are *agents* [BW, 2003].

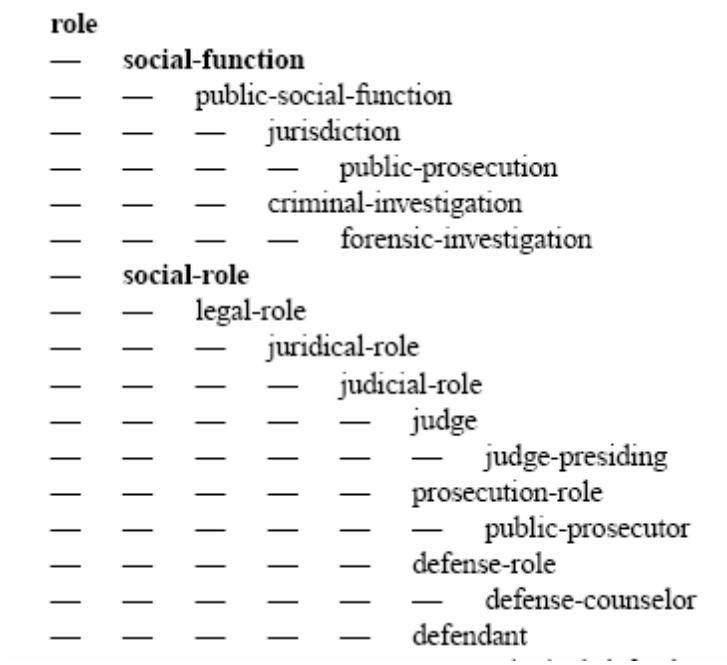


Figure 3.20: Example from the modelization of roles and functions in LRI-Core

Another approach to model *role* is the one presented by Gangemi et al. in the construction of the Jur-(Ital)Wordnet (Jur-IWN) project, an extension to the legal domain of the Italian version of EuroWordnet. Jur-IWN has been based on the DOLCE foundational ontology. In the preliminary linking of legal concepts to DOLCE+, JurWordNet, contains that *natural_person* (considered a *physical_object*) is separated from functional roles. Under this point of view, judge, defendant and prosecutor would be functional roles, whether or not they are physical objects [GST, 2003].

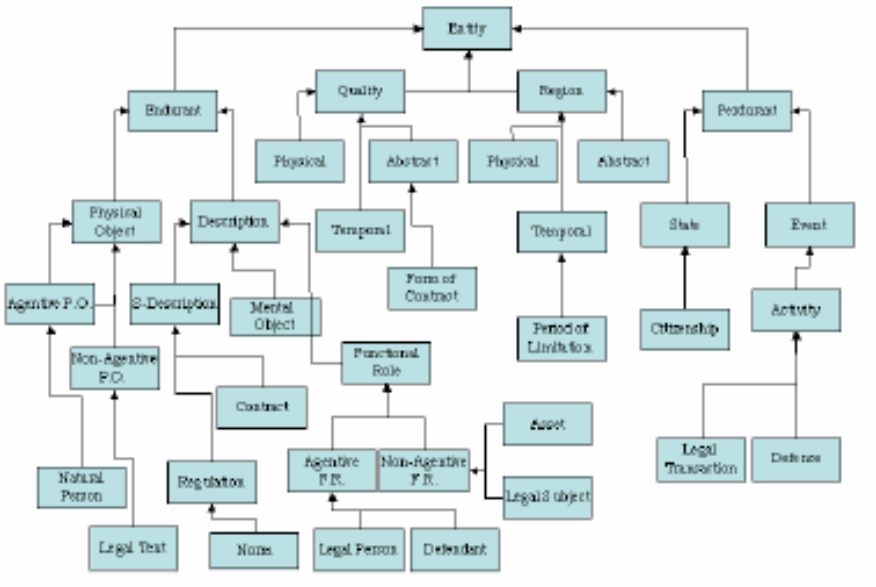


Figure 3.21: A preliminary linking of legal concepts to DOLCE+ in JurIWN.

We believe *role* to be a central concept to the PJK ontology, although because of its complexity, it is still under revision. One *agente* [agents] might play several *roles* during a process or might have several opened processes where the agent plays different *roles*. Another subclass of *rol* is *rol_familiar* [family role]. This *role* is important to model all that questions related to domestic violence and family proceedings. The *role* played by the *agent* in the family has significance in the establishment of sanctions or the legal status of a fact.

ROL		
	ROL_PROCESAL	
		absuelto
		acusador
		demandado
		víctima
	ROL_FAMILIAR	
		compañero
		cónyuge
		esposa
		ex novio
		hijo
		madre
		marido
		novio
		padre
		pareja
		pareja de hecho

- *Documento_jurídico* [procedural document] is a subclass of *documento* [document]. The argumentation has been already discussed above.

- Finally, *Jurisdicción* [jurisdiction] and *Sanción* [sanction] are relevant concepts regarding the geographical distribution of courts and the different types of sanctions (derived from civil or criminal liability), respectively.

3.5.3 OPJK relations

Some properties/attributes of concepts and relations between concepts have also been identified and they are summarized in the following chart:

Acto_procesal – *has_document* {instances of *documento_procesal* class}
Agente – *has_role* {instances of *rol* class} – *is_involved_in* {instances of *hecho* [event] class} – *has_state* {instances of *estado* [status, situation] class} – *has_location* {instances of *localización* [loction] class} – *can_be* {string of attributes: *competente, incompetente, no competetente, distinto, otro, igual*).
Órgano_judicial – *a profesión_jurídica_works_in* {instances of *profesion_jurídica* class} – *has_jurisdiction* {instances of *jurisdicción* class}
Persona – *has_age* – *has_profession* {instances of *profesión* class}
Calificación_jurídica – *has_degree_of_comission* (string of attributes: *consumado, continuado, tentativa*) – *has_will* (string of attributes: *culpa, dolo,*

imprudencia, negligencia)
Documento – *has_date* – *has_location* {instances *localización* class} – *has_author* {instances of *agente* class} – *has_addressee*
Documento_procesal – *is_issued_in* {*Proceso_judicial* class}
Resolución_judicial – *is_issued_by* {instances of *órgano_judicial* class}
Fase_procesal – *begins_with* – *ends_with* – *followed_by* {instances of *fase_procesal* class} – *has_time_interval*
Proceso_judicial – *has_phase* {*Fase_procesal* class}
Rol_procesal – *played_by* {instances of *agente* class & instances of *profesión_jurídica* class} – *has_time_interval*

Above we have described the main classes, concepts, instances, attributes and relations contained in the current version of the Ontology of Professional Judicial Knowledge. This ontology is still under development; on the one hand, there are still more than 500 competency questions to be analyzed. On the other, the ontology, once

integrated into the *Iuriservice II* prototype, will be tested for its efficiency in relation to the iFAQ retrieval system. That will surely lead to an in-depth refinement process.

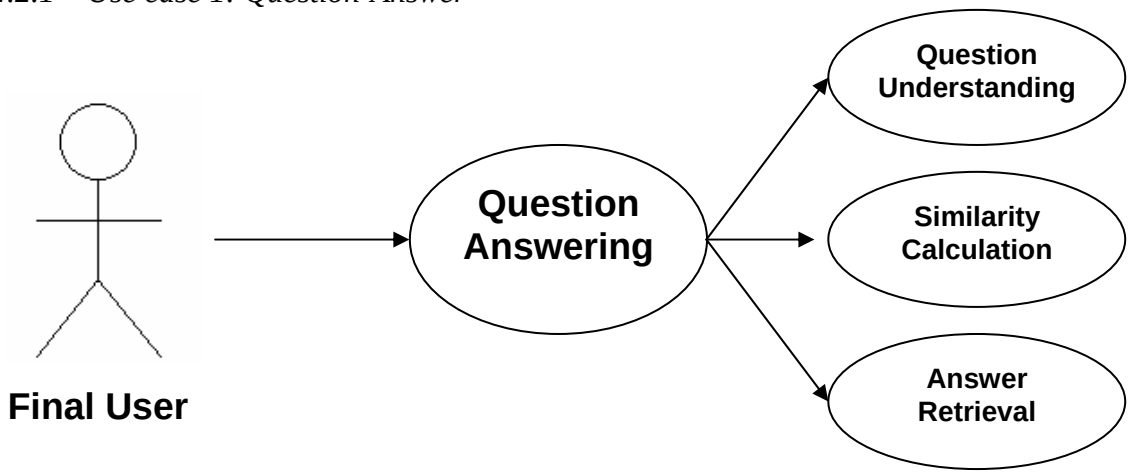
4 Use Cases

4.1 Introduction

The aim of this document is to collect some use cases in order to capture the desired behavior of the system. This will be of help to identify how each partner can contribute with its technology to the case study, and to start a discussion that ends up in a first design of the architecture of the system.

4.2 Use cases collection

4.2.1 Use case 1: Question-Answer



4.2.1.1 Description

An inexperienced judge faces a new problem and does not know how to proceed. Instead of calling a more experienced colleague, decides to give an opportunity to the new system installed in his computer. She/he opens the system and, in the box that appears in her screen, types the question in Spanish.

Once the question is typed, the system first parses it, trying to identify important concepts present in the question. To do so, it relies on the ontology, and on simple natural language techniques, such as morphology tagging or chunk parsing. Once this information has been extracted from the user sentence, it is compared with all the corresponding information of each of the questions contained in the FAQ, which have been previously pre-processed, so that they do not need to be parsed each time a question is posed. The most similar question is selected, and the question-answer pair is provided to the user.

In case the information present in the questions of the FAQ repository is not enough to do the matching, also the corresponding answers could also be considered.

If there are several question-answer pairs that offer a good matching, all of them could be provided as the system response.

4.2.1.2 Required methods and technology

The system needs the following processes to be available:

- Spanish NLP techniques.
- Ontology API.
- FAQ repository, with pre-processed questions and answers.
- Similarity calculation.
- Ontology of Professional Legal Knowledge.

4.2.1.3 Expected benefits

The inexperienced judge will get a quick answer based on the experience of other judges, in almost real time.

All the judges will get the same kind of answers to the same problems.

4.2.1.4 Potential pitfalls

There are several potential sources of problems in this process, none of them with a high risk associated.

First of all, if the user question processing does not retrieve enough or correct information, it can be due to a lack of vocabulary in the detection process or a lack of domain knowledge in the ontology.

Another potential source of problems can be the similarity calculation algorithm, which may not associate the question to the most related questions in the repository properly.

Finally, there can be a lack of knowledge in the FAQ repository and the user question may not be covered by the system.

The most probable one out of these possibilities is the last one, but the inclusion of new questions to fill these knowledge gaps is considered as another of the possible use cases for this case study.

4.2.1.5 Assignment to SEKT technology

This is a simple use case in terms of its relation with the rest of the SEKT technologies.

The process will need to access the domain ontology several times, so the ontology API will be needed.

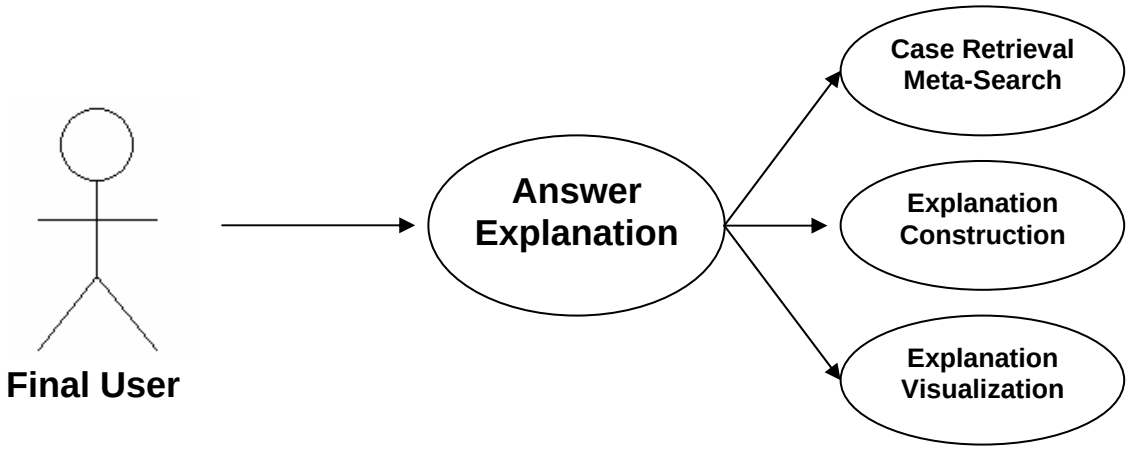
At this point the domain ontology is considered to be ready to use, so no input from workpackages 1 and 2 is expected.

From WP3, the ontology API will be used.

Only one ontology is involved in the process, so no input from WP4 is planned.

Possibly some of the technology developed in WP5 will be of use to access the knowledge, or maybe to generate some sentences as an explanation. This will be decided after fixing the user interface.

4.2.2 Use case 2: Answer explanation



4.2.2.1 Description

This use case starts where the use case number one ends, so the user has posed a question to the system, and it has provided an answer. Once the judge gets an answer, she/he feels the need for some more information regarding the decision recommended, and chooses to obtain an explanation for it.

To select the corresponding cases associated to the answer provided, the question and the answer are linked to the corresponding concepts in the Ontology of Professional Legal Knowledge. The set of concepts that represent the question and the answer is then transformed into the corresponding set of concepts that appear in the Jurisprudence ontology. With this set of concepts, which also may appear in any of the databases ontologies, and taking into account which of them belong to the question and which of them belong to the answer, the cases that are representative are retrieved and presented to the judge.

The number of cases that is returned as a result is not limited.

4.2.2.2 Required methods and technology

The system needs the following processes to be available:

- Ontology API.
- FAQ repository, with pre-processed questions and answers.
- Cases databases.
- Ontology of Professional Legal Knowledge.
- Ontologies representing the cases databases.
- Jurisprudence ontology.
- Ontology of Professional Legal Knowledge and Jurisprudence ontology alignment.

4.2.2.3 Expected benefits

Apart from the quick answer to a problem, the judge can obtain some jurisprudence related to the concrete problem. The location and access to this relevant cases will be much faster than the usual search a judge performs.

4.2.2.4 Potential pitfalls

This process relies on some work done in previous steps, such as the design and building of ontologies, mapping, etc., however these processes are detailed in other use cases, and their risks evaluated there.

The process in itself does not present high risks. The main problem that could be identified by now comes from the fact that a cases database may contain millions of them, and the selection may be a time costly operation.

4.2.2.5 Assignment to SEKT technology

This use case relies on previous work to be done perfectly.

First of all, it needs the Ontology of Legal Professional Knowledge built and ready. In order to access it (and all the rest of ontologies involved), the Ontology API is also needed.

The process also needs another ontology to be built from each of the databases of cases (one ontology per database) representing the knowledge contained in the cases, and the Jurisprudence ontology, including all the knowledge in all the database ontologies.

In order to connect the concepts in the two ontologies involved in the use case, a mapping is needed.

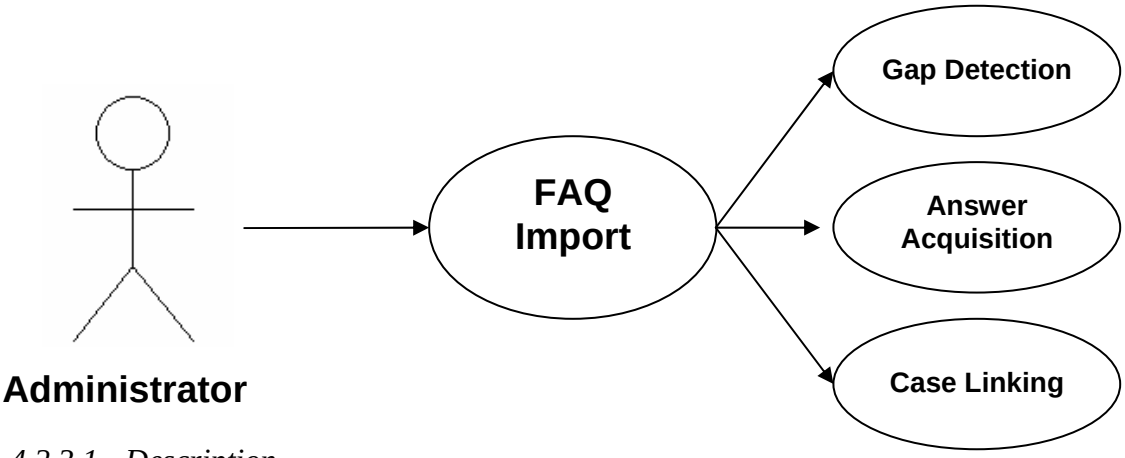
The same happens with the concepts relevant in each Question-Answer pair and in each case, they are pre-calculated for efficiency reasons. Every time a new Question-Answer or a new set of cases is included in the system it should be processed to extract these information.

From WP3, the ontology API will be used.

Possibly some of the technology developed in WP5 will be of use to access the knowledge, or maybe to generate some sentences as an explanation. This will be decided when the user interface design is clarified.

From the rest of the technical WPs some work is needed as a prerequisite, but it will be detailed in different use cases.

4.2.3 Use case 3: FAQ updating



4.2.3.1 Description

After a number of interactions with real users, a gap in the knowledge covered by the frequently asked questions is detected.

The domain experts design a question (or a set of them) to fill that gap of knowledge. Each question is answered by experienced judges, and the question-answer pair is delivered to the system administrator.

The question-answer pair has to be pre-processed to extract the relevant concepts present and incorporate the knowledge to the repository.

4.2.3.2 Required methods and technology

The system needs the following processes to be available:

- Log containing the questions with low matching, which may constitute pointers to possible knowledge gaps.

4.2.3.3 Expected benefits

This use case allows semi-automatically detecting gaps in the knowledge contained in the frequently asked questions, and updating the professional knowledge present in the system as needed by the users.

4.2.3.4 Potential pitfalls

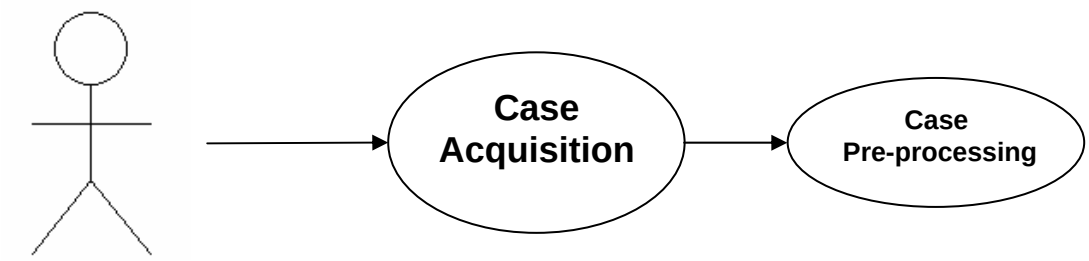
The main problem that may arise from this use case would be an incorrect pre-processing of the new knowledge to be included in the system. This would cause

problems trying to answer future user questions regarding the point addressed by the new question-answer. However, as this pre-processing is semi-automatic, the administrator should personally supervise its correctness.

4.2.3.5 Assignment to SEKT technology

No special assignments regarding the technology developed by the technical partners is involved in this use case.

4.2.4 Use case 4: Cases updating



4.2.4.1 Description

Periodically, the set of available cases in the databases will be updated, and therefore this knowledge needs to be included in the system.

When new cases are included in the databases, this update should be detected or notified, either manually or automatically, and the cases pre-processed in order to extract the relevant concepts that appear in them. These concepts, if necessary, must be included in the ontologies related to the databases. This update of the ontologies may also imply a modification in the global Jurisprudence ontology. If this modification happens, it may be the case that the alignment with the Ontology of Professional Legal Knowledge may also need to be revised.

Finally, the judicial concepts appearing in the new cases should be linked to the corresponding instances in the ontology.

4.2.4.2 Required methods and technology

The system needs the following processes to be available:

- Access to cases databases.
- Notification/Detection of updates.
- Pre-processing of cases.
- Ontology learning.
- Ontology feeding.
- Ontology merging.
- Annotation.

4.2.4.3 Expected benefits

This use case allows to update the jurisprudence, so that the judges can have available the latest cases to justify the answers provided by the system.

As new cases are produced everyday in all the judicial units, there is a need for a constant updating of the knowledge in the system, or it would become obsolete in a short to medium term.

4.2.4.4 Potential pitfalls

This is a complex use case, and therefore the potential pitfalls are more numerous than in other use cases.

First, it would be a problem if the addition of new cases in the system would not be correctly detected or notified, as these cases would not update the ontology, or be linked to any concepts appearing on it, and so they would never be used for justifying any of the answers provided by the system.

An incorrect pre-processing of the new knowledge to be included in the system would cause unpredictable behaviour trying to justify future answers with the new cases.

If the concepts in the cases are not correctly updated in the corresponding ontologies, or the Jurisprudence ontology is not updated adequately, the new knowledge will not be used to justify the answers of the system.

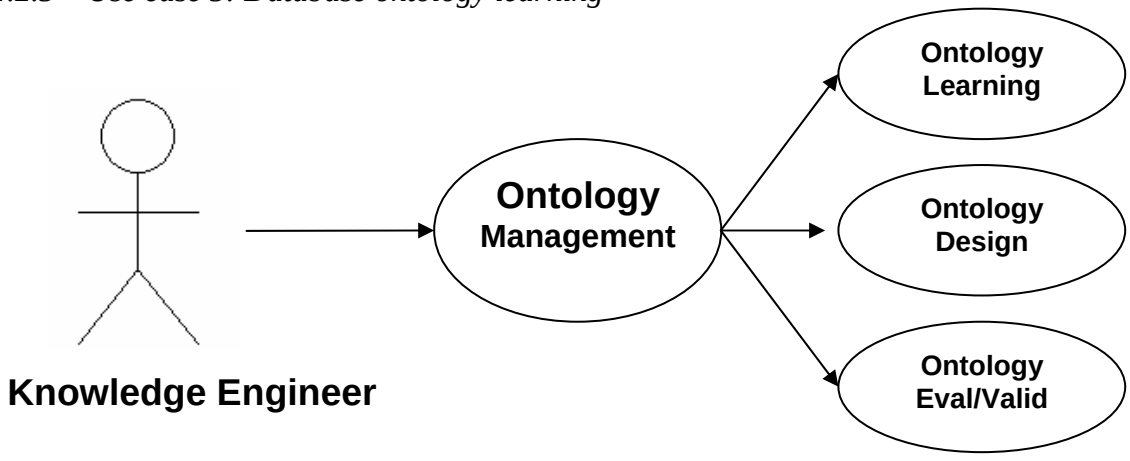
Finally, the last step of the process may also carry out some problems, if the alignment of the two main ontologies of the system is not correctly done, the appropriate cases will not be used to justify the answers that the systems provides.

4.2.4.5 Assignment to SEKT technology

This use case implies many of the technologies included in the project.

The detection of new knowledge and necessity of changes is very close to some of the objectives of WP3. The ontology modification to cover these changes could be covered by some of the tasks included in WPs 1 and 3. The supervision of the ontology merging and possible re-calculation is included in task 1 of WP4. The same applies to ontology alignment, in task 2 in the same WP.

4.2.5 Use case 5: Database ontology learning



4.2.5.1 Description

One of the first stages of the system will be the creation of an ontology containing the knowledge in each of the possible databases of cases available.

The cases in the databases will be processed to infer an structure of the ontology, that should be evaluated (automatically, semi-automatically or manually) and corrected until a satisfactory version is reached. Once this ontology is designed, the knowledge present in the cases is extracted and formulated as an ontology, resulting in an ontology reflecting the knowledge present in the database.

4.2.5.2 Required methods and technology

The system needs the following processes to be available:

- Ontology learning.
- Ontology evaluation.
- Ontology feeding.
- Annotation.

4.2.5.3 Expected benefits

The system will have available a set of ontologies (one per database) reflecting the knowledge included in the databases, and ready to work with.

4.2.5.4 Potential pitfalls

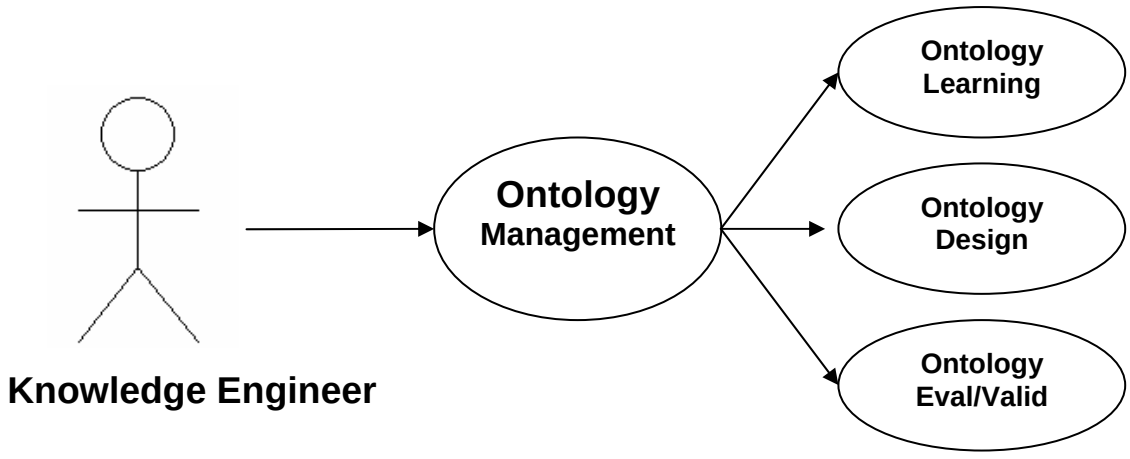
Spanish judicial language is difficult to process using NLP.

Apart from this, the risks are the usual risks associated to the automatic learning and feeding of ontologies, and should be minimized by the knowledge engineer supervision of the processes.

4.2.5.5 Assignment to SEKT technology

This use case implies tasks directly related to WP1, with a small help from the rest of the workpackages.

4.2.6 Use case 6: Jurisprudence ontology learning



4.2.6.1 Description

Once the ontologies for each database have been created, their knowledge is combined to build a general ontology covering all the jurisprudence knowledge included in the system.

This is done by merging all the ontologies built automatically resembling the knowledge included in each of the cases databases.

This process is represented in the architecture figure with the name "Ontology merging" (see chapter 3).

The knowledge engineer should check the result of the merging.

4.2.6.2 Required methods and technology

The system needs the following processes to be available:

- Ontology merging.
- Ontology evaluation.
- Ontology API.
- Ontology editor.

4.2.6.3 Expected benefits

The system will put together in one ontology the knowledge contained in a set of different databases, which will constitute a central point of the system reasoning.

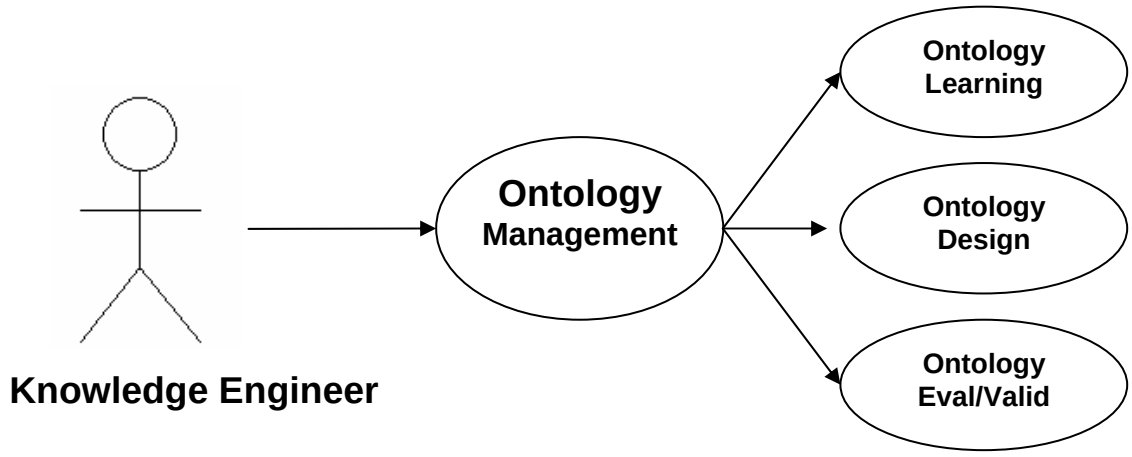
4.2.6.4 Potential pitfalls

The pitfalls are the risks associated to a normal process of ontology merging, and should be minimized by the intervention of the knowledge engineer supervising the results of this process.

4.2.6.5 Assignment to SEKT technology

Ontology merging is the objective of task 1 of workpackage 4.

4.2.7 Use case 7: Jurisprudence Ontology and Ontology of Professional Legal Knowledge (OPLK) aligning.



4.2.7.1 Description

When the two main ontologies of the system are completed, they need to be aligned, in order to connect the two different kind of knowledge present in the system, the everyday professional knowledge (represented by the OPLK), and the (more theoretical) jurisprudence knowledge.

This use case is represented in the architecture figure with the name “Ontology alignment”.

As well as in the previous use case, the knowledge engineer should revise the results of the process.

4.2.7.2 Required methods and technology

The system needs the following processes to be available:

- Ontology aligning.
- Ontology API.
- Ontology editor.

4.2.7.3 Expected benefits

The system will connect the knowledge expressed by the repository of frequently asked questions, a representation of the expertise gained by a experienced judge, and the Jurisprudence ontology, an ontology that represents the knowledge present in the real cases verdicts, much more based in law theory. This will allow the association of the frequently asked questions and the related cases that are available in the system.

4.2.7.4 Potential pitfalls

One source of problems would come from the fact that some of the concepts in the Ontology of Legal Professional Knowledge were not linked to any concepts in the Jurisprudence ontology. However, the knowledge engineer or a domain expert should supervise the alignment, and this kind of problems would be manually corrected.

4.2.7.5 Assignment to SEKT technology

Ontology aligning is the objective of task 2 of workpackage 4.

4.3 First draft of the architecture

In order to better understand the use cases, an initial high level draft of the architecture of the system has been designed.

The system manages two independent kinds of knowledge, represented by two subsystems. On the left, the expert knowledge subsystem contains the expertise of experienced judges. This knowledge is represented by a repository of FAQ (questions frequently asked by novice judges) and the ”Ontology of Professional Legal Knowledge” (OPLK), which holds and structures all this knowledge. On the right hand side, the existing body of law is found, represented by databases with existing rulings. The knowledge of each database is represented in an ontology, and all these ontologies are merged in order to get a single ontology that comprises all the knowledge in the databases. This ontology is aligned with the OPLK in order to be

able to map the concepts that are used in the two different subsystems, what allows connecting the questions and answers of the repository in the "expert knowledge" subsystem with the cases in the databases that can be relevant in order to justify the decisions recommended in the answers.

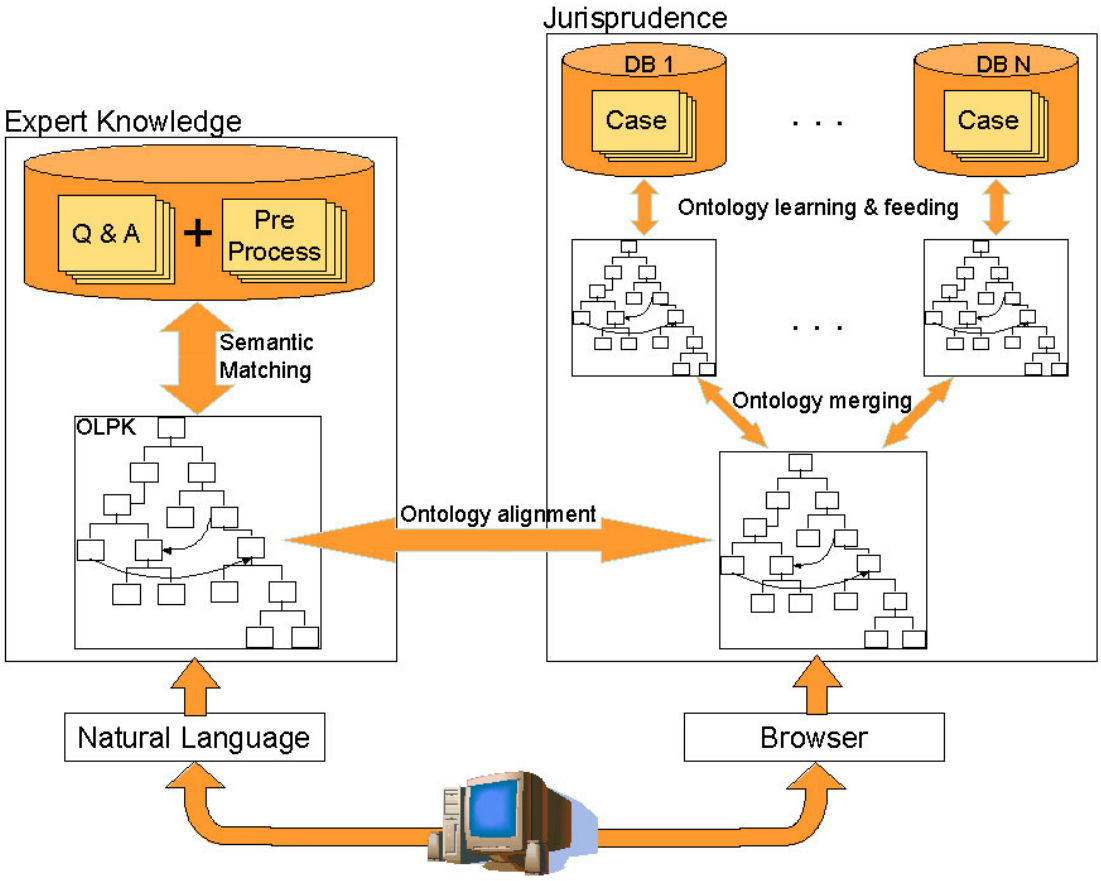


Figure 4.1: Preliminary *Iuriservice* architecture

It could be desirable that the user could navigate through the "jurisprudence" ontology and find cases related to the legal concepts she/he is interested in, without asking any question.

5 Use Cases software design.

5.1 Introduction

This section describes the Legal Case Study Prototype software architecture. The Legal Case Study Prototype system is an integrated decision support system for legal professionals. System users will use it when they are confronted with situations in which they are not sure what to do. The most frequent use case of this system consist of a user asking a question and the system trying to find out among a vast legal FAQ repository a few candidates having the best question-answer match to the question. Then all this candidate FAQ items are shown to the user ordered by a score given by the system.

The main challenge of this project is to go further than conventional searching system based on metadata or keyword. A step beyond will be to use AI techniques like

Ontologies or Natural Language Processing (NLP) to assist this system to select the best FAQ question-answer item candidate among the complete FAQ repository. To accomplish that, a deep understanding of the FAQ knowledge domain must be acquired of an ontology.

We have design a open software architecture that may be plugged with several AI modules like those cited in the previous paragraph, ie. NLP and Ontology processing. Nevertheless, other AI technologies to assist in the searching process may be implemented and plugged seamlessly like a module in this architecture..

In the rest of the chapter we use the word FAQ in the following contexts:

- A FAQ, or FAQ item is a object built by a legal question and its answer. Therefore, we will talk about FAQ question and FAQ answer, too.
- A FAQ set or list is a group of FAQ items that can be recover from and save into a FAQ repository.
- A FAQ repository store the whole set of FAQs used by the system. A FAQ database is a synonym used in the text.

5.2 Technological considerations

The main goal of the Legal Case Study Prototype is to build an integrated decision support system for legal professionals. System users will use it when they are confronted with situations in which they are not sure what to do and they need some advice from expert knowledge.

Our improved FAQ searching system is based in several key ideas considerations we summarize in the following items:

- Accurate search. This system will be able to find out the best possible matching FAQ question with user question.
- Efficiency. Searching must not take a long time to find a result successfully and must scale well with FAQ repository size.
- Ontology based searching technologies.
- Spanish Natural Language Processing (NLP) keyword detection.

Due to these considerations, we have valued a compromise between accuracy and efficiency, both often incompatible.
This compromise is achieved by means of some techniques we have applied in our design:

- Background calculation of ontology and NLP based search helper data.
- Caching of background calculated data.
- Multistage searching approach with progressive delimiting of FAQ set target.

5.3 Use cases

The most important and frequent use cases designed for this FAQ searching system are described in detail in the next sections.
This next use case UML diagram (5.1) depicts the most important cases:

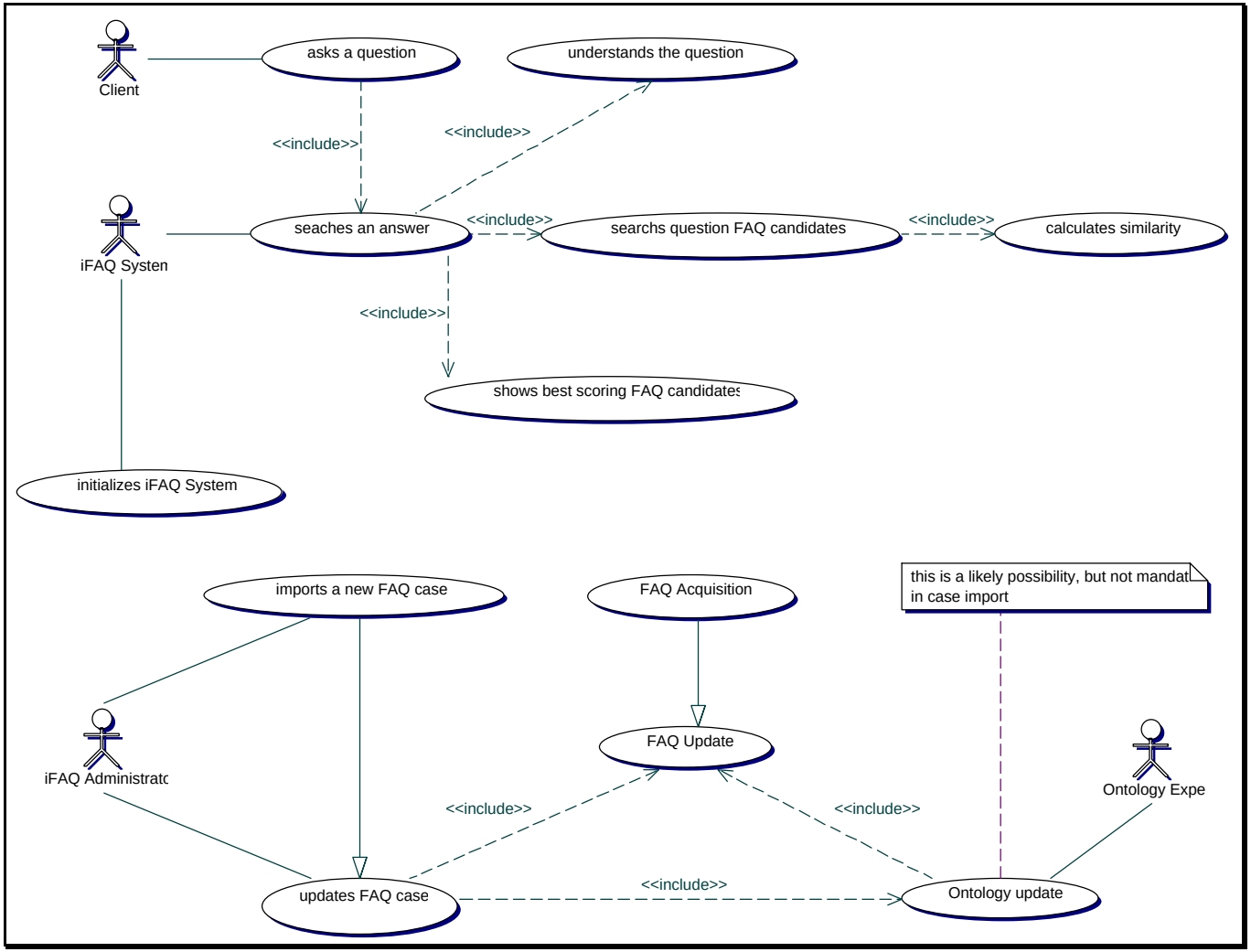


Figure 5.1: Legal Case Study Prototype main use cases.

5.3.1 User question search use case

Use case description:
This is the most frequent and important use case found in our system. A system user types a question that he expects to find at a FAQ repository so question meaning is as close as possible to FAQ question. The system will use all its technologies available to try to find out the best scoring FAQ item.

Several searching and score algorithms have been designed to achieve this use case goal. They can be applied consecutively in a multistage cycle that approaches to best matching FAQ item.

There are some functional considerations behind this software design:

- Exhaustive search can be very time consuming.
- Multistage searching system lets the system stop when FAQ target set has been reduced considerably. Therefore not all stages are exhausted and the computational cost is reduced and adapted to search features.
- Besides, multistage searching allows configuring the system to show the best score results to the user before completing all available searching stages, so users can decide to accept the result or continue with the remaining search stages.

5.3.2 Overall search system

In Figure 5.2 we depict an overall description of multistage search system. This software architecture uses a Factory pattern to build a FAQSearchEngine implementation suitable for the ontology search purpose. In our case ontology based FAQSearchEngine is to be created, but other FAQSearchEngine could be used, if necessary. FAQSearchEngine will determine from configuration or from demand what search engine to use among the available ones.

Next, three stages will be processed in order to reduce our FAQ searching target set, with the compromise that the best score FAQ item related with the user question must be included in this target. This last point is very important because if previous searching stages mistakenly determine a FAQ target set, next searching stages will be unprofitable.

At each searching stage, the input will be a FAQ subset determined at a previous stage, and the output a reduced FAQ subset as an outcome of a narrower search.
We postulate two assertions in our design:

- The best user question matching FAQ item must be included in the outcoming FAQ subset.
- Each stage is more restrictive than the previous one, that is, the FAQ result set is narrowing in each stage.

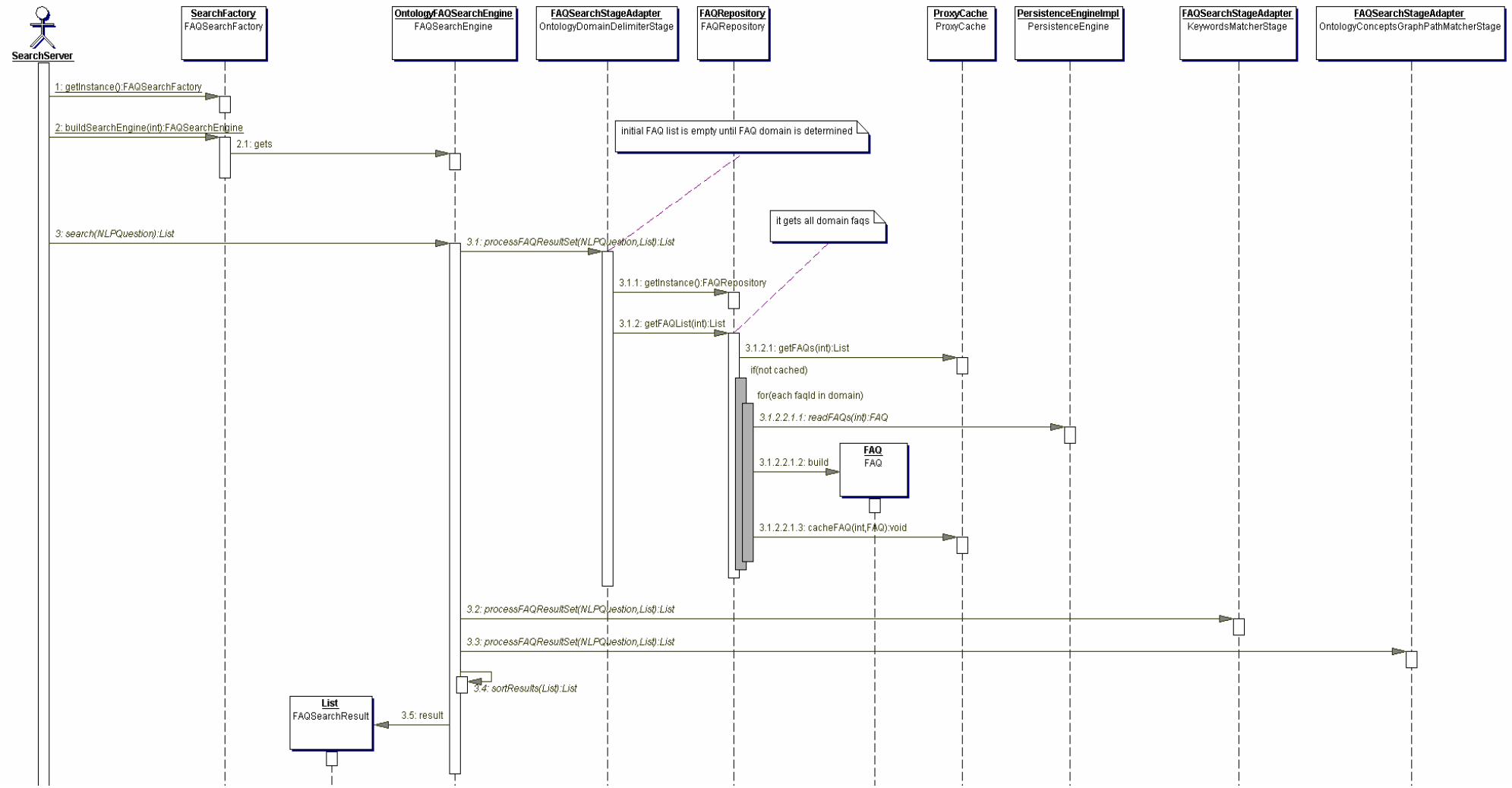


Figure 5.2: Overall search system UML sequence diagram.

This multistage protocol allows us to choose at which stage to stop searching, if we achieved a good reduced subset of FAQ candidate items. Besides, it enables us to insert new searching stages at any point of the process.

We have designed three searching stages that leverages on ontology and NLP technologies:

- Ontology domain detection stage. The main purpose of this stage is to determine the FAQ domain target set, based on user question analysis, for use at a later stage. Complete FAQ database is made up over some domain FAQ databases as building blocks. Reducing the FAQ database domain target is the first consideration to take into account, so the first search stage focuses on it. This goal is achieved by a statistical recount of occurrences of all relevant concepts of the user question among the different domain ontologies.
- Keyword matching stage. At this stage, the user question is word-like tokenized and system tries to match each token with each FAQ question token. Both exact and synonymy matching is tried. As a result of this stage, a narrower list of candidate FAQ items will be supplied. This might seem similar to other standard searching systems using keywords or metadata. Here, the difference lies on the use of morphological parsing of the user question that discards non-relevant words and the use of synonymy.
- Ontology concept graph path matching stage. This is the most time consuming stage, and therefore it is left to the final stage, when the FAQ target database has already been very reduced by previous searching stages. At this stage, grammatical patterns will be detected from user question. Then, these patterns are searched in the Ontology to build a graph path or trace. Finally, the system tries to match this user question ontology graph path with a reduced FAQ target subset graph paths previously calculated in the background using some semantic distance algorithms.

In the picture there is also another technology we will use frequently in our design: a cache system for rapid access to frequently used data, such as, for example, the FAQ database. Performance is the main motivation for using this cache-based design. This searching system uses intensely certain data that is not necessary to recalculate more than once. But as cache memory is limited, a compromise between memory consumption and efficiency has to be managed. There are several open source caching frameworks,, like JCS[45], OSCache[46], JOC[47], etc., or commercial products, like SpiritCache[48], Coherence[49], etc. These frameworks works fine with Java objects and supply different algorithms for storage and recovery. They also supply different algorithms to solve memory limitations like LFU (least frequently used), LRU (least recently used), MRU (most recently used), FIFO, etc.

As a result of a completed staged searching phase, a FAQ list ordered by score will be provided.

[45] http://jakarta.apache.org/turbine/jcs
[46] http://www.opensymphony.com/oscache
[47] http://jocache.sourceforge.net
[48] http://www.spiritsoft.com/products/cache/introducing.shtml
[49] http://www.tangosol.com/coherence.jsp

5.3.3 Ontology domain detection stage

The next sequence UML graph (Figure 5.3) depicts in detail the workflow followed at this stage. This is our first searching stage. Its goal is to reduce the FAQ database target of our searching system to improve performance in the other phases. The FAQ database can be very vast. Indeed, this FAQ database is composed of several smaller domain databases. Therefore, as the user question will likely belong to a knowledge domain, it is only necessary to search in that specific domain database, not in the rest.
In this sequence diagram, the searching system is built over some important components that will often appear:

- NLPEngine: responsible of morphologic and grammatical parsing of questions.
- MorphologicalEngineAdapter: a helper class used by NLPEngine, specialized on morphological analysis. This adapter leverages on specialized external morphological analyzers supplied by other companies.
- ThesaurusEngineAdapter: a helper class used by NLPEngine, specialized on synonyms searching. This adapter leverages on specialized external thesaurus engines supplied by other software manufactures.

This design, based on adapters, allows us to replace those specialized helper classes with other alternatives provided that these new helper classes fulfill the adapter interface. So adapters allow us to use plug-ins for special tasks.

First of all, a NLP engine dissects the user question and detects all its relevant words, analyzing them syntactically. For this work, NLPEngine leverages in other helper classes, that offer special functions. A morphological engine adapter processes each user question word, detecting the POS, genre, number, etc. of each meaning of the word, and all relevant words are filtered. This task is very computational consuming so the most frequently processed words are cached in memory. Later accesses to that data are collected from the cache memory instead of morphologically processing the user question again.

After that, we have collected a list of meanings of words relevant to the user question. We also need all the synonyms of these relevant words, because it is possible not all of the relevant meanings appear in the ontology, but some of their synonyms may do. To achieve this, a thesaurus engine is used, leveraged on some external thesaurussoftware like IBM JADT[50] or similar.

[50] https://secure.alphaworks.ibm.com/tech/jadt

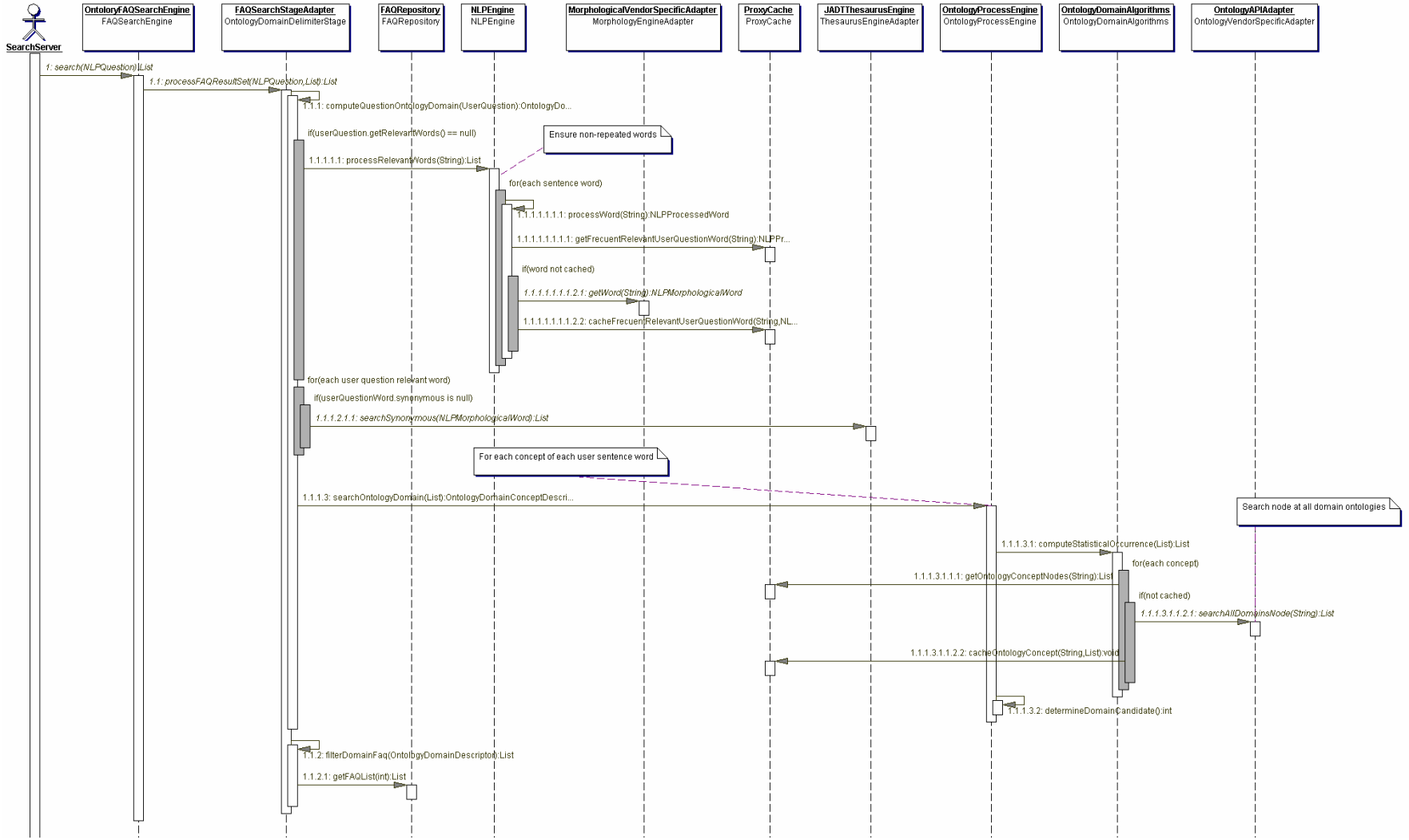


Figure 5.3: Ontology domain detection stage UML sequence diagram.

Now, we have got all needed processed data from the user question to find out the best candidate of ontology domain fitted to our user question. Then, the next task is to statistically detect all those concepts in each ontology domain. An ontology process engine leveraged on some ontology algorithms and on an ontology API adapter, calculates all these statistical occurrences. To avoid reiterate searches of the most frequent ontology nodes, these can be recovered from cache if the performance of the ontology searching system is not as expected. The outcome is an ontology domain candidate with the best statistical score.

Once the ontology domain has been determined, the stage final outcome is built, consisting in the complete FAQ list of the candidate ontology domain.
Another design aspect considered is that all data calculated at one stage must be kept over the complete multistage cycle, because it is likely to be reused again at the following stages.

For example, the relevant words meanings of a user question are kept in cache for later use.

5.3.4 Keyword matching stage

The second searching stage (depicted at Figure 5.4) receives a FAQ list processed at a previous stage. The responsibility of this stage is to filter this input to reject non-matching FAQ subset, considering matching at keyword level. This phase punctuates each FAQ item with a score. All FAQ items with scores beneath a threshold will be eliminated from the candidate list.

If the FAQ candidate list is reduced to a short number of elements, the system can decide not to pass throughout the next stage, showing the result to the user. This increases the overall performance. System configuration can be used to customize this behavior. Normally, system will not stop until all stages are finished.

The key concept at this phase, which differentiates it from other standard searching system, is the use of morphological meanings and synonyms in this keyword-matching algorithm. This avoids the use of exact or partial word matching. The use of only relevant words also reduces the emergence of many false candidates.

After processing or caching the relevant words of the user question and its synonyms using an NLPEngine, the searching stage pursues iterating over each input filtered FAQ item.

For each FAQ item, the following strategy is used to determine a matching score between user question and FAQ question: first, the relevant words of the FAQ question are recovered from cache (otherwise, they will be recovered from persistence RDBMS). Then, two inner loops iterate along all the relevant words of the user's question and FAQ question. For each couple of both meanings, an exact matching is tried and if it fails, a synonymy matching is processed.

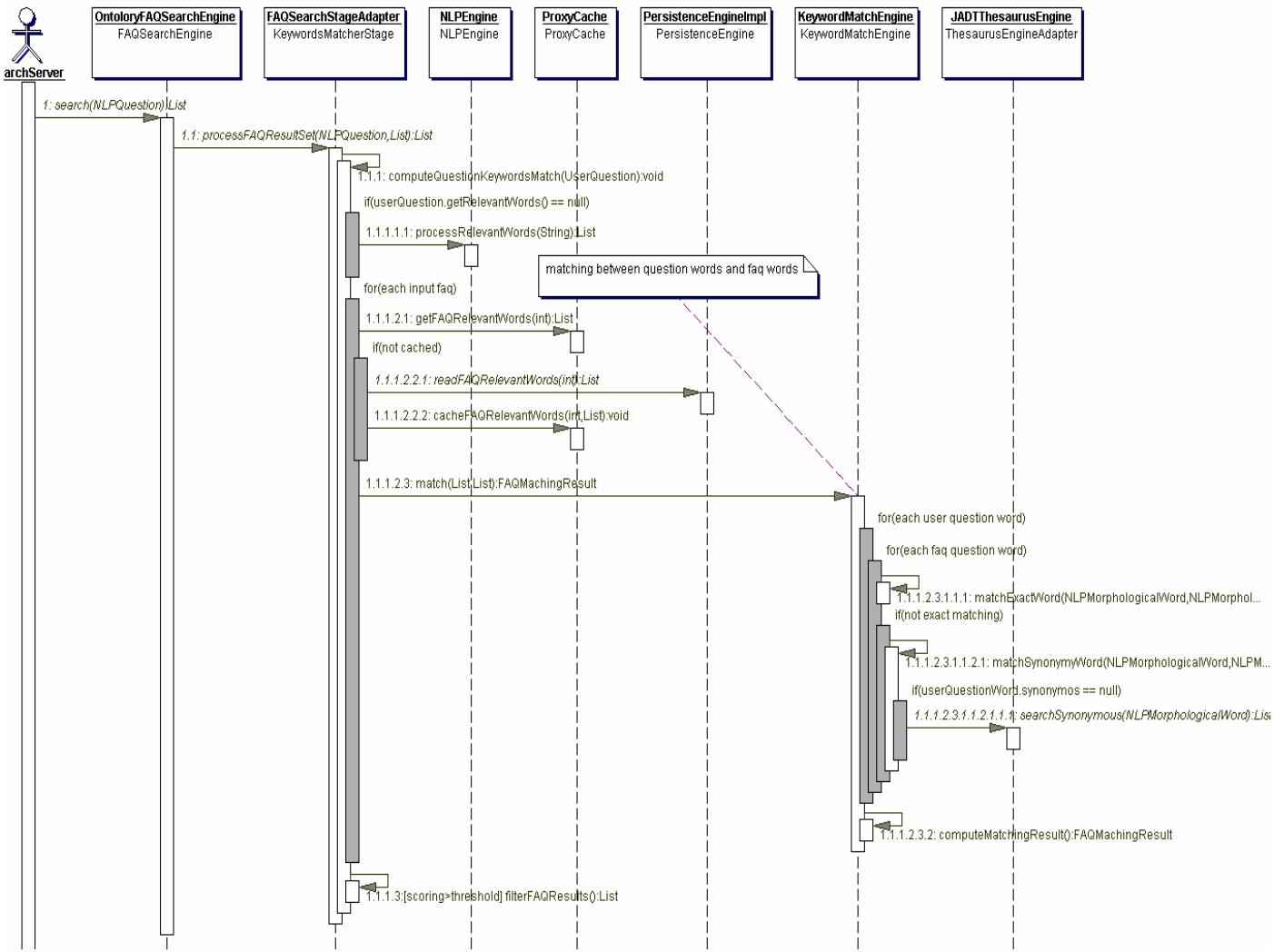


Figure 5.4: Keyword matching stage UML sequence diagram.

Finally, all collected matched hits are processed to assign a final score to that FAQ item.
As a result, we have got a list of punctuated FAQ items. Each FAQ item with a score below a threshold will be discarded.

5.3.5 Ontology concept graph path matching stage

This last stage is the most time consuming so it is important to start it with a very reduced input FAQ candidate list. The work made at this phase can be summarized roughly as follows:

- Parsing of the user question to detect grammatical patterns.
- Identifying these grammatical patterns in the ontology, implying:
  - Searching for the concepts of the grammatical patterns.
  - Finding the paths that connect those concepts.
- Computing the minimum distance between grammatical pattern paths of user question and FAQ question.

In followings paragraphs we will describe these items more deeply.
The next sequence UML diagram (5.5) depicts this stage in detail.

Some other important helper classes and third party components used at this stage:

- A syntactic engine adapter that leverages on a third party syntactic parsers.
- Ontology process engine that leverages on a third party ontology API and on libraries of ontology algorithms. Different types of algorithms will be needed, for ontology searching, path distance calculations, etc.

A grammatical analysis is carried out over the user question, by using a syntactic engine adapter (leveraged on external syntactic software like Schug parser, from the Esperonto Services IST-2001-34373[51]). The objective is to dissect the sentence into smaller meaningful grammatical units. For each unit, its role and position in the sentence is determined, besides other syntactical aspects. If the sentence is composed then it is split in simpler sentences. The outcome is a list of grammatical patterns (for each simple sentence), each of them consisting in a list of grammatical concepts. That means that a complex sentence is split in simpler ones that are processed individually. With this information we construct the patterns. A common pattern would be like: subject, verb, direct object, etc.

[51] http://www.esperonto.net/semanticportal/jsp/frames.jsp

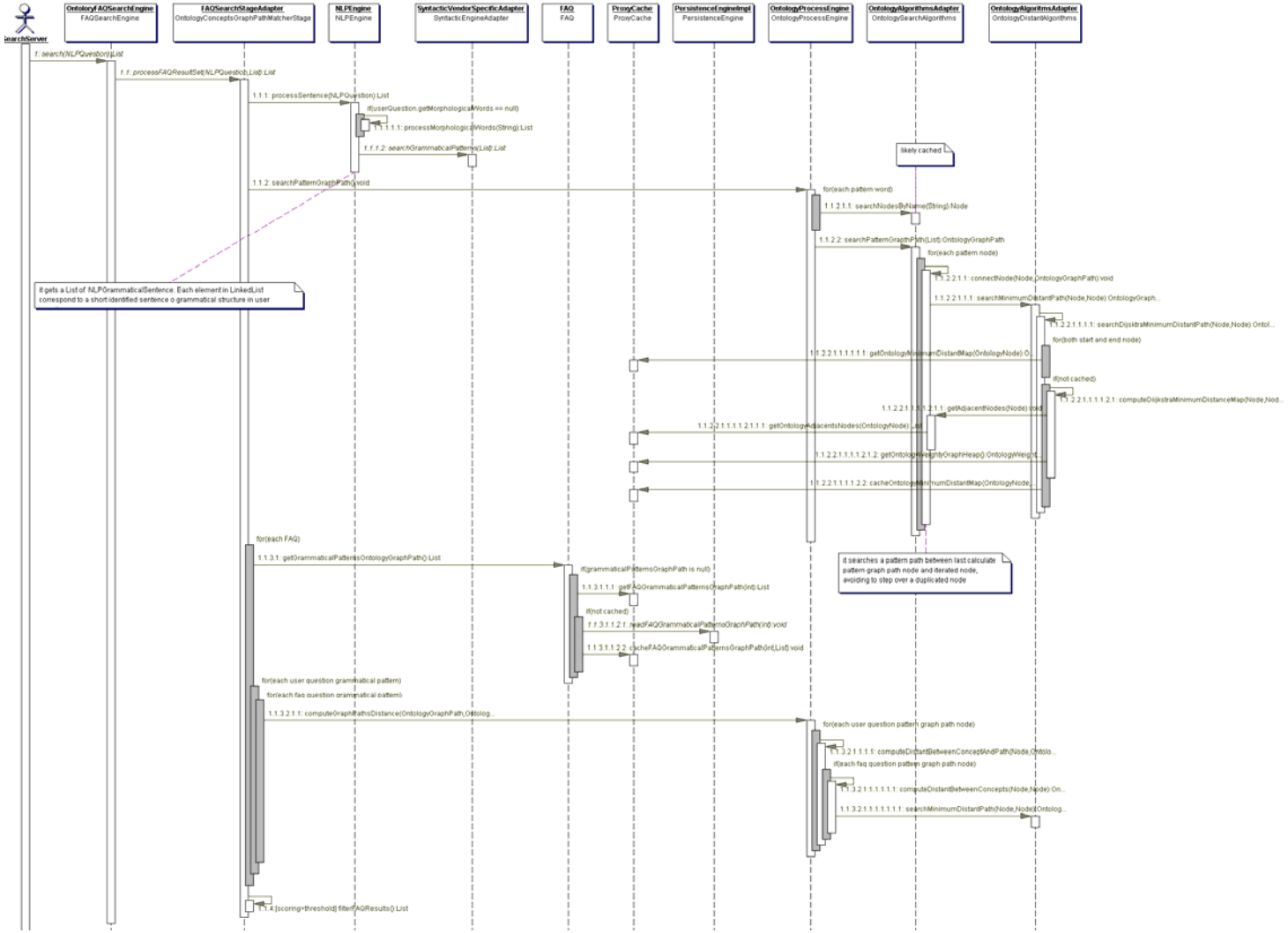


Figure 5.5: Ontology concepts graph path-matching stage UML sequence diagram.

Next step consists of searching those previously calculated grammatical patterns in the domain ontology. This is a complex process consisting of different steps, briefly enumerated in the following list, and deeper developed later in this section:

- Searching of the ontology nodes that match any grammatical concept of the user's question.
- Searching of the minimum distance path that connects each pair of consecutive concepts.
- Building a minimum path by connecting some or all the pattern concepts in the correct order.

The first task consists on locating the ontology nodes corresponding to each concept of the grammatical pattern in the correct order. The system can locate these nodes from the cache (most frequent used nodes), or by searching them with the help of an ontology API framework. Some nodes will be located, others won't. For located nodes the system will build a path that connect them. Non-existing nodes will be marked for score considerations.

Next tasks build a connecting path as follows: from each node the system will find a minimum path to the next node, and so on. By doing so, the path that connects all nodes will be completed. Sometimes, it will not be possible to connect a pair of nodes, so the resulting path will not be fully connected.

Now, we face the challenge of finding a minimum distance path between two nodes. A minimum distance path between two ontology concepts is that with the smallest accumulated weight, considering that between two connected nodes there is an edge with a slot weight assigned. We have opted for a Dijkstra's algorithm with a Fibonacci heap representation of the ontology graph. Fibonacci heap or binary heaps are a suitable way to represent a weighted graph where a minimum node must be quickly extracted. Also these representations reduce the algorithm's running time from $\Theta(n^2)$ to $\Theta(m + n \log n)$, where n stands for number of nodes, and m for the number of edges . Other algorithms can be used, but this one is very appropriate for our purpose and very efficient computationally. This algorithm uses also some suitable data other representation structures, like adjacency matrix and adjacency lists, which improve its efficiency as well. Both the Fibonacci heap and adjacency structures are previously calculated during the system initialization phase and are recovered from cache memory (in other words, we make these data persistent). Dijsktra's algorithm allows us to calculate, with the same computational effort, the minimum distance path from one node to the rest of them (minimum distance path map from that node). Minimum distance path between two ontology nodes is very time consuming. Therefore, for performance reasons it would be desirable to keep as many as possible in the cache. This will be made when a FAQ item was registered (adding) in the system. So we arrange to have all minimum distance path maps from any concept node of all FAQ items cached in memory or persistence on a RDBMS. As a result of this step we get a list of ontology paths for each grammatical pattern detected on the user question. Also from cache we have got similar ontology paths for all FAQ grammatical patterns. In short, we have acquired all data needed for the next step.

In this last step we use an algorithm for estimating the minimum distance between two ontology paths: one path from the user question pattern, the other from the FAQ candidate question pattern.

The key idea for this design is as follows: the user question and its best FAQ candidate item are supposed to have very similar grammatical patterns. This similarity is translated to the ontology domain in such way that both grammatical patterns must be similar as ontology graph paths. A measured distance between both graphs is an indicatory of this similarity.

Several algorithms can be considered and used to estimate this distance. Here we depicted a very simple algorithm: for each user question pattern path node the system calculates the minimum distance to all FAQ question path nodes and annotate the minimum. Proceeding with the rest of user question pattern path nodes, we calculate the sum of all of them.

This partial sum will be completed with all combinations of grammatical patterns from user and FAQ questions.

It's pending to design a score policy that determines the best matching FAQ item based on ontology distance between grammatical pattern paths.

We wish to point out some important aspects of using of Dijkstra's or similar algorithms to compute minimum distance paths. First some ontological considerations:

- The relations between concepts in the ontology are weighted, so a minimum distance path is calculable. That means that all arcs from one node to another have a weight slot.
- Ontologies used are bidirectional in the sense of one can walk from one node to another along the same path or vice versa.
- Algorithms like Dijkstra calculate with similar computational effort both the minimum path from one node to another and all minimum paths from the same node to the rest of them (the minimum path map for initial node).

These considerations let us save for all FAQ question nodes their minimum path map in a database repository and keep them cached on memory for fast access. Then, when the system needs to get the minimum path from one question user node to a FAQ question node, it proceeds in inverse order, because that info is available from cache.

We are studying some algorithms designed to estimate the minimum distance between two graph paths in ontologies [FGRT, 1992], and [CDF, 2004]. Each graph path represents a grammatical pattern from both the user and FAQ question. Here in these documents we have taken a glimpse on one of those algorithms, having taken ideas from the literature. Others algorithms can be incorporated to this prototype to improve its accuracy, efficiency or both.

5.4 Search system initialization use case

Search system leverages on some preliminary calculations made at startup or in background during normal working or with a schedule of low system activity, for example when an administrator adds a new FAQ item in the system or when the ontology has changed and it needs to be updated (these use cases will be treated later). When system starts up it must executed some initialization tasks. The following sequence's UML diagram (Fig. 5.6) summarizes this task.

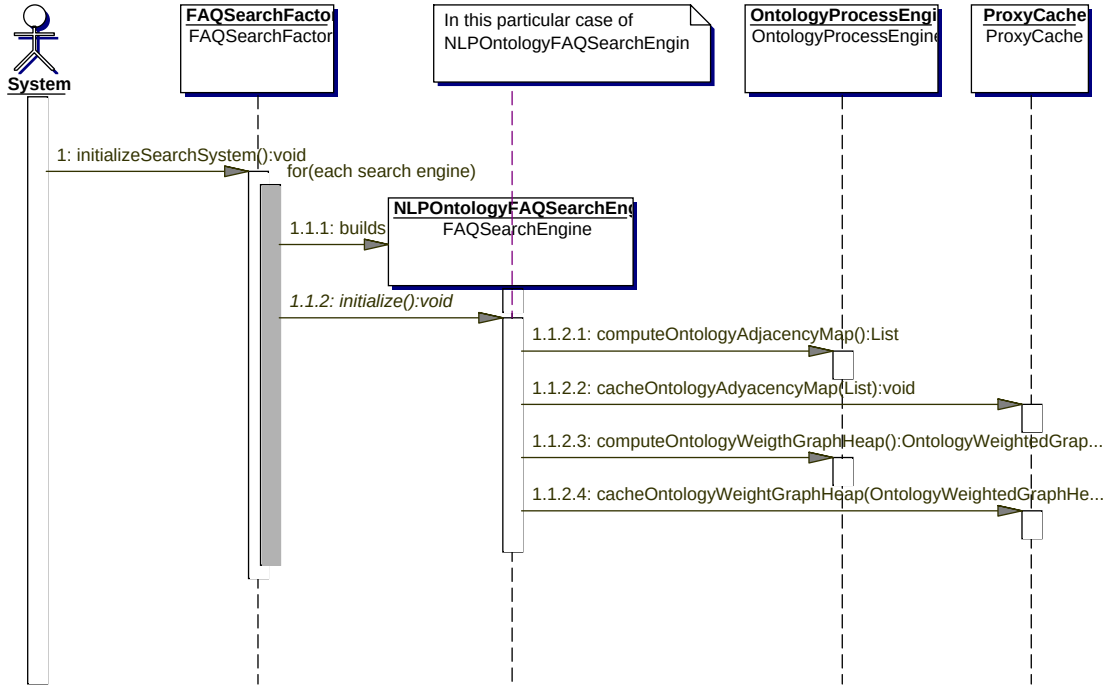


Figure 5.6: Search system initialization use case UML sequence diagram.

At startup phase system uses the FAQSearchFactory to build all available FAQSearchEngine. After building these engines the system initializes them.
In the particular case of an NLPOntologyFAQSearchEngine (our particular engine we use in this prototype) the initialization consists on the following steps:

- Compute the adjacency data structures for all ontology concepts by using an OntologyProcessEngine. This adjacency matrix is necessary for searching a minimum path between two nodes by using a Dijsktra's algorithm. It can be calculated or recovered from persistence database.
- Compute the Fibonacci or binary heap representation of the weighted ontology graph by using also an OntologyProcessEngine. This heap is used by Dijsktra's algorithm to improve its computational efficiency.
- Cache both objects in memory for later use. These two objects are save wholly in memory for efficiency reasons while other cached objects can be retired from memory if they are less demanded in a reduced memory cache system.

Other needed data will be recovered from a persistence repository and stored in the cache memory when they were requested.

5.5 New FAQ item acquisition use case

This use case and the next one are considered for maintenance purposes. Normally a legal domain expert revises both the FAQ repository and domain ontology. As a consequence the FAQ searching system evolves with time, incorporating new FAQ items or updating some of them. Also the domain ontologies are modified or extended, that's the reason why these use cases need to be incorporated in the system. Then, the administrator uses this use case to maintain the system to date.

In this section we consider the case when a new FAQ item needs to be added to the repository, so some preliminary calculations have to be undertaken and saved for later use.

FAQRepository is the class responsible of managing the adding, updating and deleting of the FAQ database. In order to add a new FAQ item into the repository, the FAQ question and answer must be supplied. But before storing this new FAQ item into the repository, the system runs some tasks to calculate additional information. For performance considerations, these tasks have to be run only once and the produced information is being saved into a persistence repository. The following list summarizes this calculated additional info:

- A NLP engine is used to produce morphological and grammatical analysis of the FAQ question: its relevant words and its grammatical patterns. Both calculations are time consuming.
- An Ontology process engine is used to search the matching grammatical pattern of the FAQ question in the ontology graph. This calculation implies to search a minimum distance path among all question nodes. As an intermediate result the minimum distance map for all FAQ question nodes will be determined. This map is a very important result, because with it we can easily calculate the minimum distance between the graph path of user and FAQ questions, avoiding complex and time consuming calculations.

Finally, after saving all this information into a persistence repository, it is also saved into a cache memory for later use.

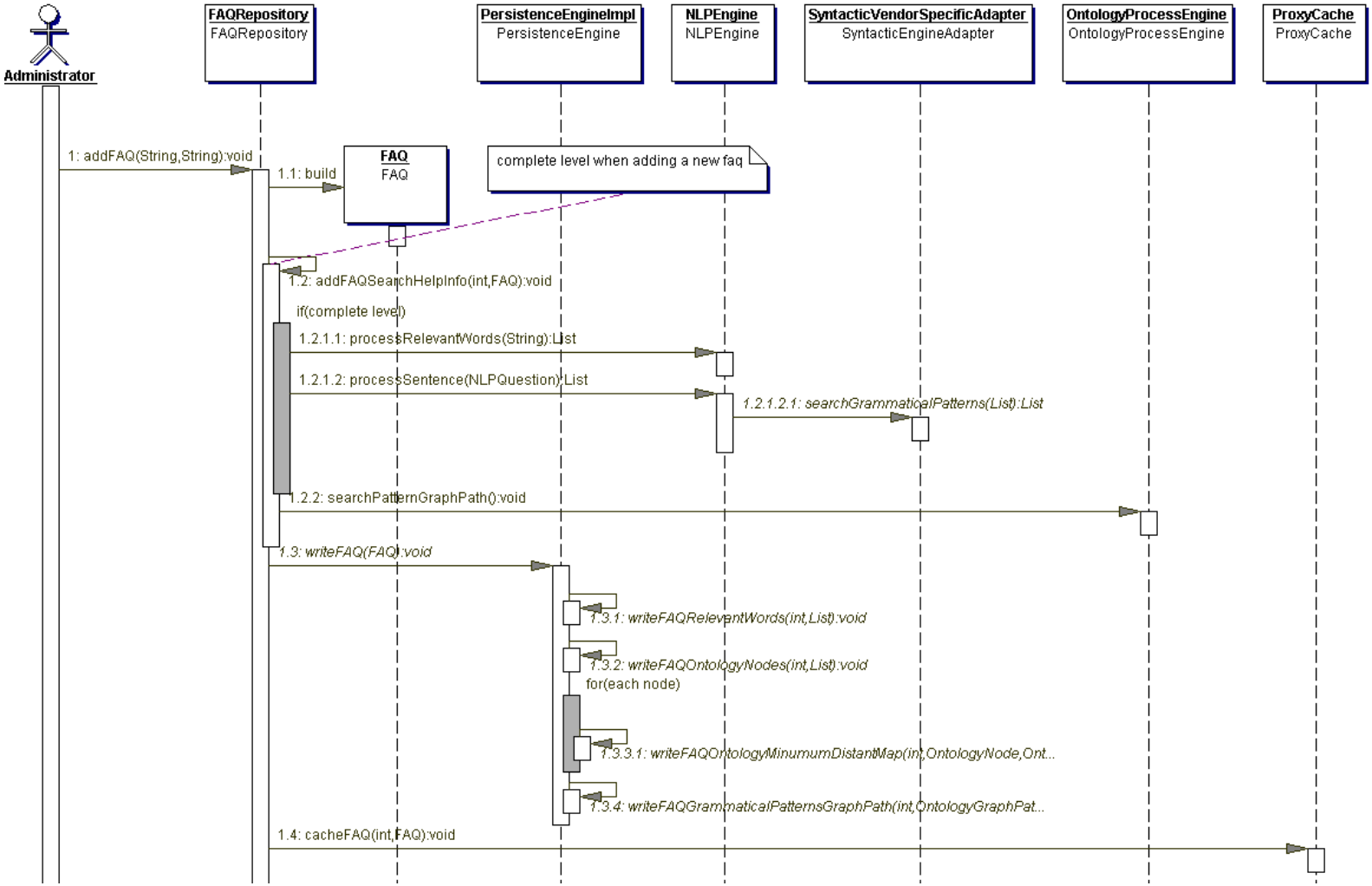


Figure 5.7: New FAQ item acquisition use case UML sequence diagram.

5.6 Existing FAQ item updating use case

This use case has to be run in two possible situations:

- A domain expert modifies some data of a FAQ question, a FAQ answer or both so a system administrator must update that FAQ item.
- A domain expert modifies one or more domain ontologies, invalidating all FAQ pre-calculations, so updating the whole FAQ set is mandatory. This is a new use case that includes this one. It will be developed further in the next paragraph.

Next sequence UML diagram (Fig. 5.8) depicts this updating process:

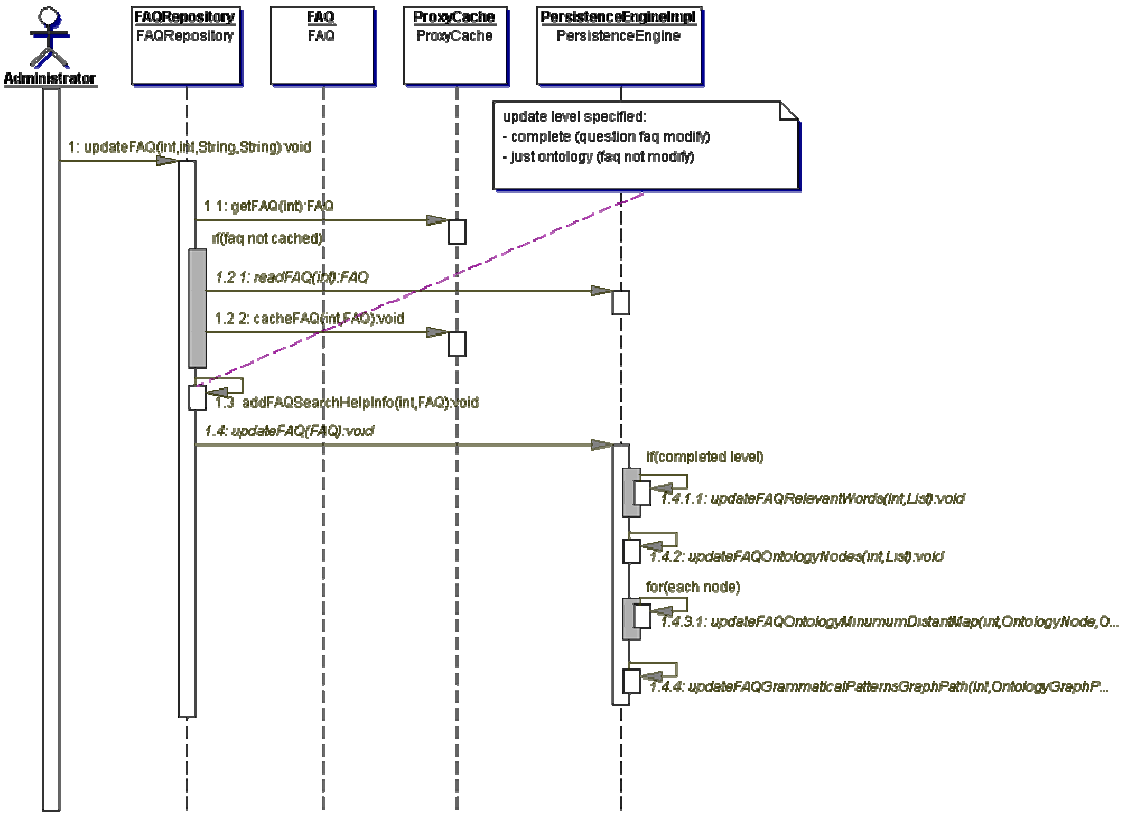


Figure 5.8: Existing FAQ item updating use case UML sequence diagram.

Some tasks in this flow graph can be tackled or not depending on whether the FAQ question has changed or only the ontology:

- If FAQ question has changed, no matter if the ontology does, when the system calls a FAQ repository to search FAQ helper info (see previous use case for details) a NLP processing must be redone. At this phase the system process the FAQ question relevant words and grammatical patterns again.
- Furthermore in any case, an ontology engine is used to search the corresponding graph path of all grammatical patterns located from FAQ question.

After all this preprocessing has finished, the system makes all these updated calculations persistent. The updated FAQ item is available from cache repository automatically.

5.7 Ontology updating use case

This use case includes the previous use case so it will not be explicitly developed in this paragraph in form of an UML diagram. In short, when a domain expert has modified the ontology, all FAQ items need to be preprocessed again to update their internal data structures. Therefore, the previous use case, "existing FAQ item updating use case" will be used iterated over all FAQ items. In this particular case the FAQ item updating is not at completed level. That means:

- NLP processed data is not affected by ontology modifications so relevant words and grammatical patterns do not change. Therefore they will not be recalculated.
- Ontology graph path patterns are modified obviously by ontology changes so they will be recalculated again.

Finally all recalculated data will be updated both in the repository database and cached memory.

5.8 Class diagrams

As another result of our technological design, we have built some class diagrams that depict the main classes and their relationships. For sake of simplicity, we have split those diagrams into one big class diagram and some smaller type diagrams:

- Overall class diagram: it represents the complete system, to sight a general idea of the system.
- Types class diagrams: they collect the main types involved in this architecture. We have collected them in separated diagrams to simplify the overall class diagram.

5.8.1 Overall class diagram

The main classes of this system are collected in the following UML class diagram (Fig. 5.9):

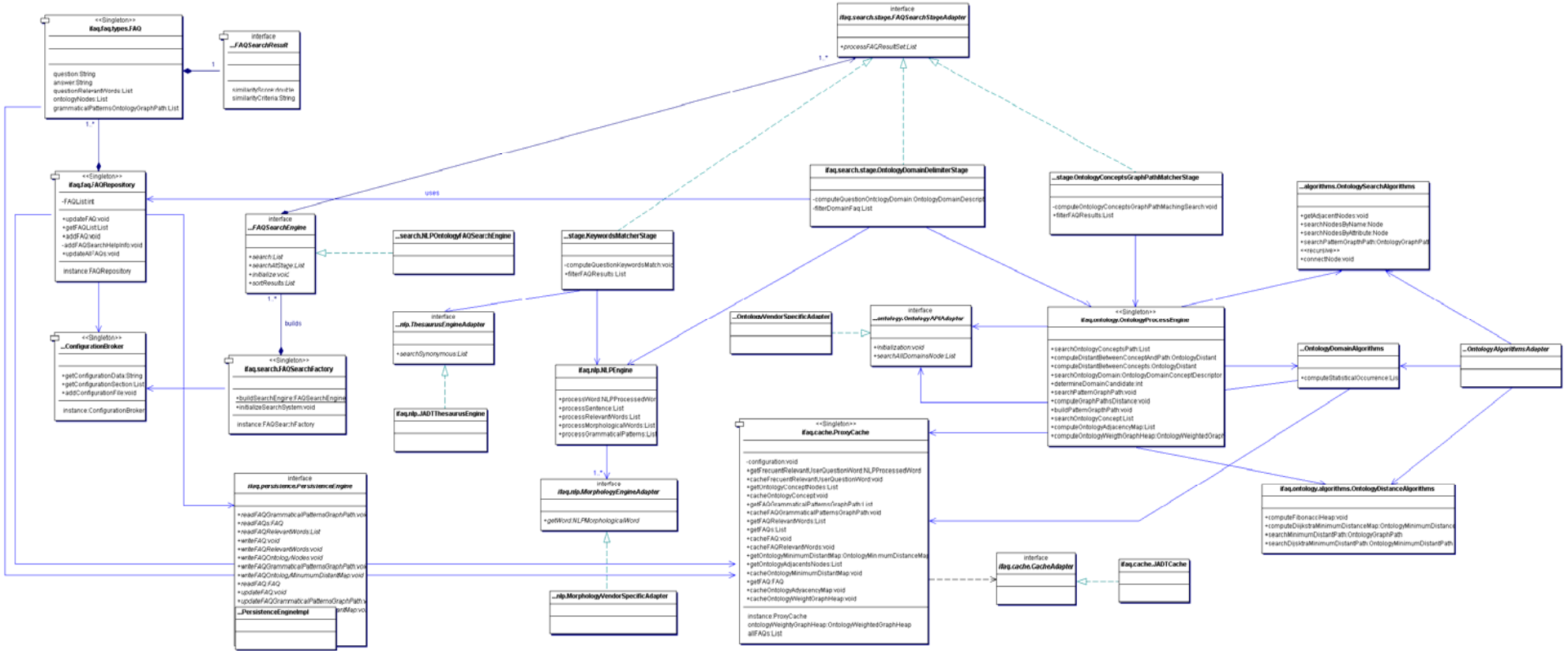


Figure 5.9: Overall class diagram.

It contains these main groups of classes:

- Searching subsystem. This constitutes the main interface of the system. We have used the factory pattern to allow many different searching engines to be built with the same interface. In this group fall FAQSearchFactory, FAQSearchEngine and its implementations like NLPOntologyFAQSearchEngine. This subsystem holds a group of classes (KeywordsMatcherStage, OntologyConceptsGraphPathMatcherStage, OntologyDomainDelimiterStage) that support multistage searching by a common interface: FAQSearchStageAdapter. FAQSearchFactory builds, under demand, objects that implement the interface FAQSearchEngine. FAQSearchEngine iterates over all objects that implement FAQSearchStageAdapter, to complete the searching cycle.
- NLP subsystem. The sybsystem helps in all the Natural Language Processing tasks. It is based on one main engine, NLPEngine and some helper class adapters for specialized work: MorphologyEngineAdapter, SyntacticEngineAdapter, ThesaurusEngineAdapter. These adapters allow the use of different technologies from different manufacturer vendors: JADTThesaurusEngine, SchugSyntacticEngine, MorphologyVendorSpecificAdapter, and allow to plug them into NLPEngine.
- Ontology subsystem. This subsystem manages all ontology related work. It is based on a main engine, OntologyProcessEngine, a main ontology API adapter, OntologyAPIAdapter and its implementations like OntologyVendorSpecificAPI. To this subsystem belong also all ontological algorithms classes, used by the ontology engine: OntologyAlgorithmsAdapter, OntologyDistanceAlgorithms, OntologyDomainAlgorithms, OntologySearchAlgorithms.
- FAQ subsystem. This subsystem supplies the FAQ repository management, with a main class: FAQRepository.
- Cache subsystem. This subsystem supplies the memory cache management with a main class: ProxyCache, CacheAdapter and its implementations like JADTCache.
- Persistence subsystem. This subsystem supplies the persistence management, mainly over RDBMS support, with a main class: PersistenceEngine.

5.8.2 Types class diagrams

The business classes described above work with data represented by some system defined types that we depict in the following UML class diagrams.

5.8.2.1 NLP types class diagram

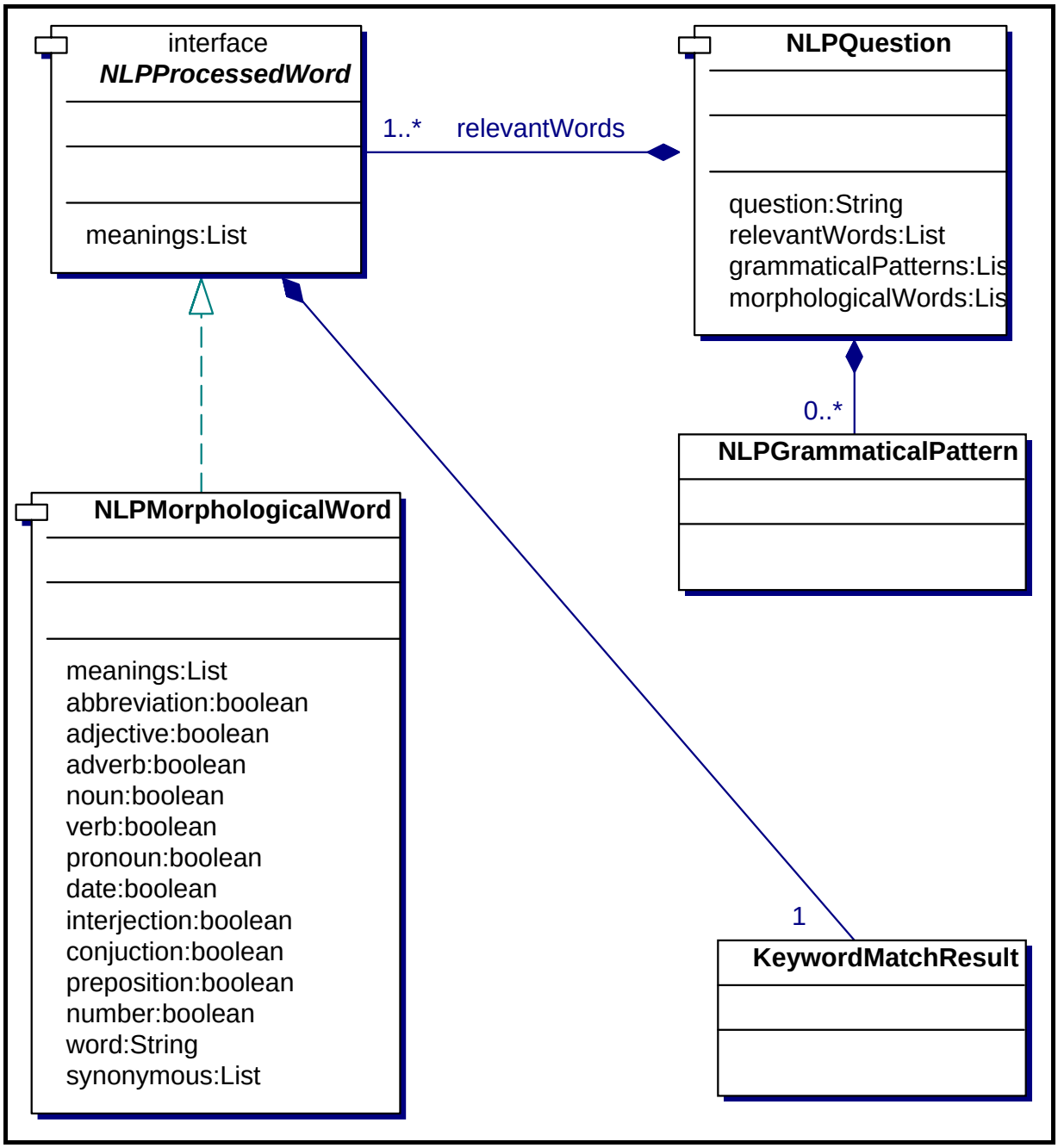


Figure 5.10: NLP types class diagram.

These classes are used mainly by the search subsystem and the NLP subsystem. The class NLPQuestion aggregates other data structures produced by NLP processing:

- NLPProcessedWord: it contains the processed morphological meanings of a word (from user's question or FAQ's question). Specialized implementations as NLPMorphologicalWord collect additional morphological data.
- NLPGrammaticalSentence: it contains the grammatical pattern detected by SyntacticEngineAdapter.

- KeywordMatchResult it contains the score data given by keyword matching stage. These data depend on the algorithm used to punctuate each FAQ candidate item.

5.8.2.2 Ontology types class diagram

These ontology type classes are tightly bound to the ontology API used. Therefore, to avoid dependencies from external API we have to design certain helper data structures to accommodate to those external data. As a consequence, a new ontology type has been considered, besides those considered for our system:

OntologyNode, OntologyClass, OntologyInstance and OntologySlot constitute themselves as completed components of the ontology framework.

- OntologyGraphPath: a LinkedList of OntologyNode that builds a path in an ontology graph.
- OntologyMinimumDistanceMap: a graph representation of a minimum distance map for a given node. It is the outcome of Dijkstra's algorithm computation for that node.
- OntologyDomainConceptDescriptor: a statistical score for an ontology concept in a selected domain. It is the result of computing the occurrence of a concept in the domain ontology.
- OntologyDomainDescriptor: an aggregate of OntologyDomainConceptDescriptor as a final score for a domain ontology candidate, used as the result of the detection stage of the ontology domain.
- OntologyWeightedGraphHeap: a helper data structure used by Dijkstra's algorithm. It is an alternative view of the ontology as a binary heap or Fibonacci heap.

At this design stage, some of these structures are only concepts to be further developed.

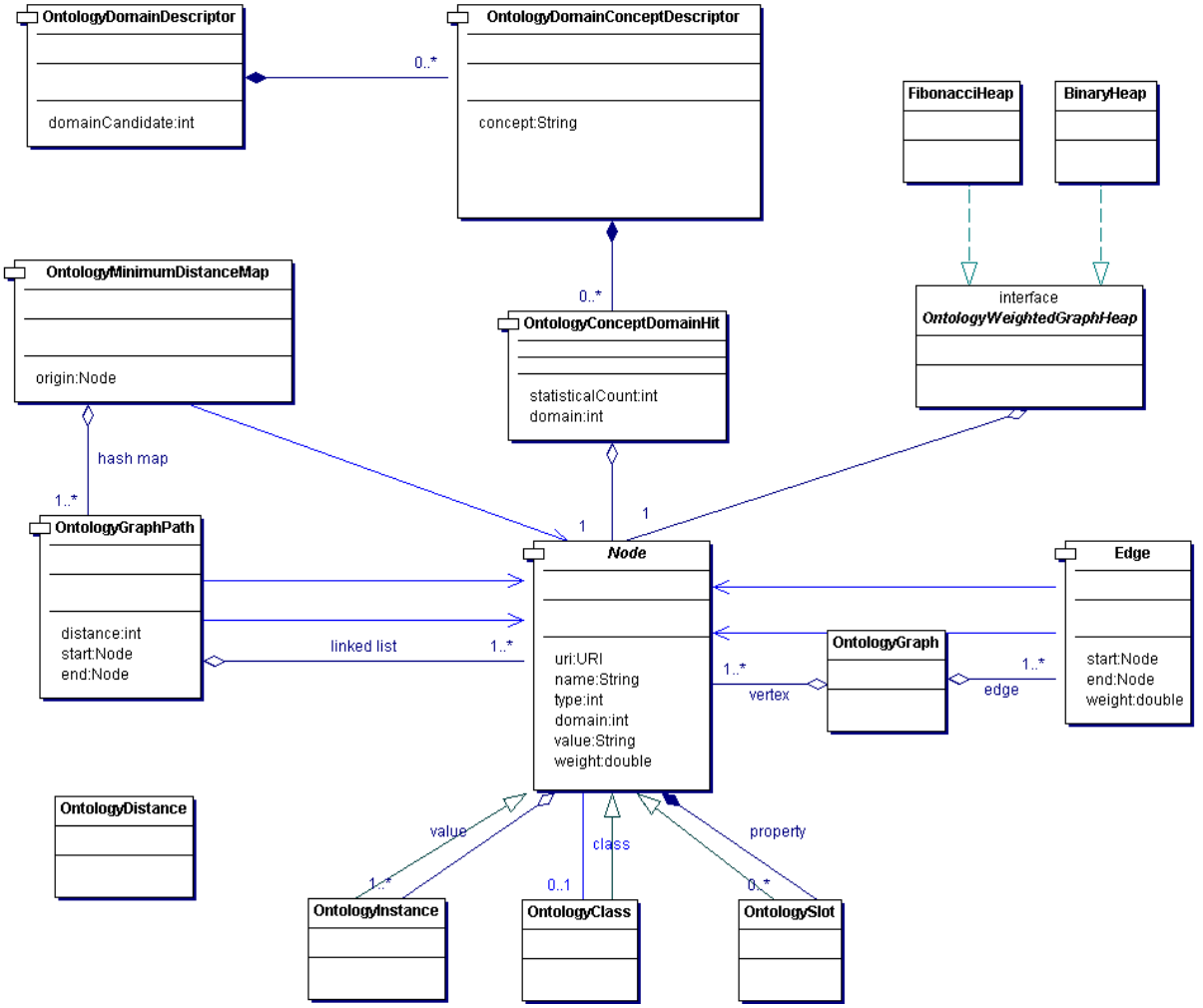


Figure 5.11: Ontology types class diagram.

5.8.3 FAQ types class diagram

In this group of types belong all structures related with question-answer FAQs items and all pre-calculated structures needed for the searching process:

- FAQ: it contains the question-answer pair, besides other pre-calculated data like morphological, grammatical and ontological structures of the question.
- FAQSearchResult: it contains search-score structures. These structures depend on the stage that produces them, therefore their interfaces will be instantiate by some helper derived classes:
    - Similarity score: punctuation given to that FAQ item.
    - Similarity criteria: explicative reasoning to assign it that score.

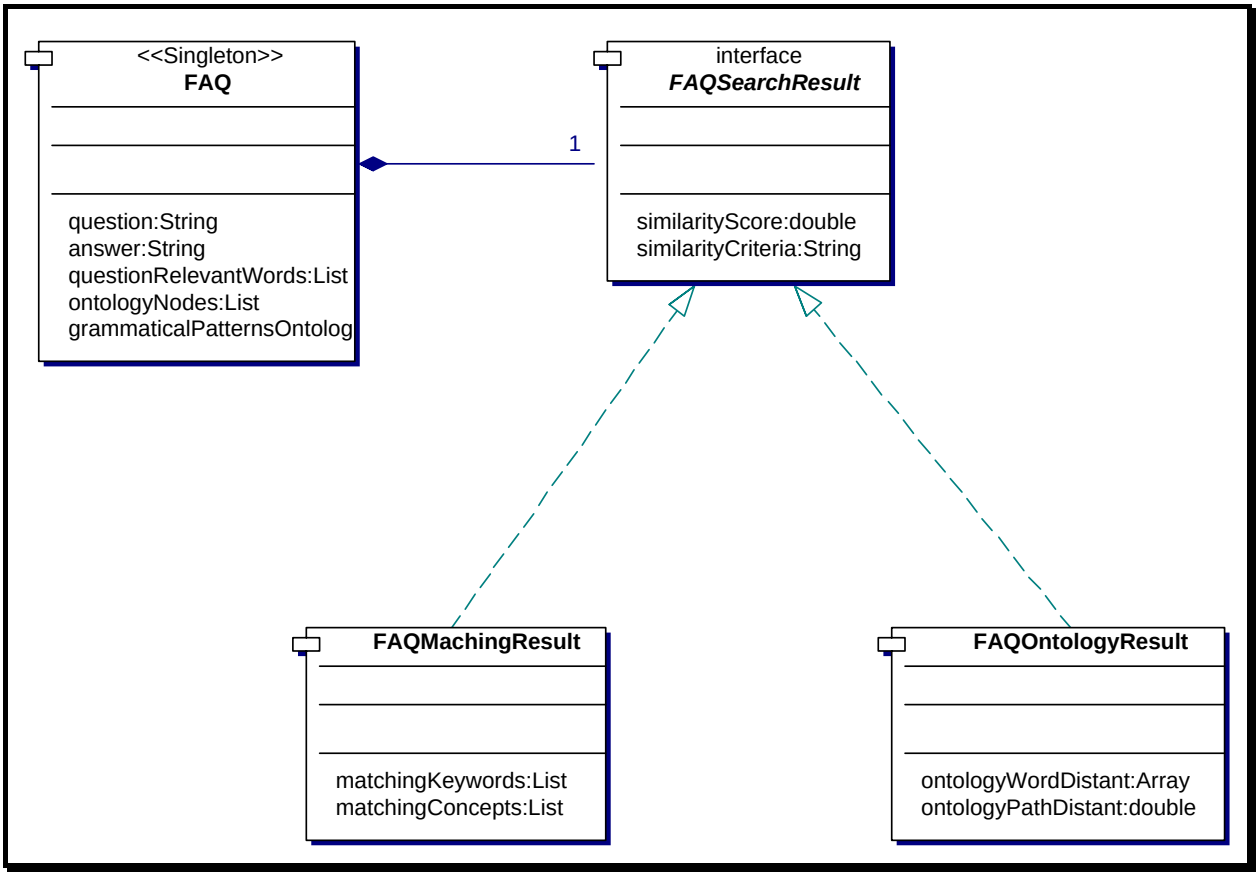


Figure 5.12: FAQ types class diagram.

5.9 Component Diagram

Finally, we have collected all this classes and we have grouped them in packages that constitute independent components. A software component is a group of clases that work together in similar tasks and it may be reused in another developments. Those components build up the whole search system. Otherwise, they can be considered as independent blocks that can be plugged in this system or in another as necessary. Some components, grouped together, constitute a subsystem when they play an important role in the system. The process to divide between components and subsystems is subjective and constitutes a simplified classification of the software.

The main subsystems are:

- NLPSubsystem.
- OntologySubsystem.
- Search subsystem.

The other components are not so important to be group in a subsystem, but they act as a helper frameworks for business logic.

Our search system can offer its services to its clients as long as they send a request. Among them, we enumerate the following clients:

- Webservice: the client web service connects to our system to request some search services.
- Instant messaging Client: it receives the user question and uses our system to get the best answer, in a similar way.
- Web client: it manages as the other examples.

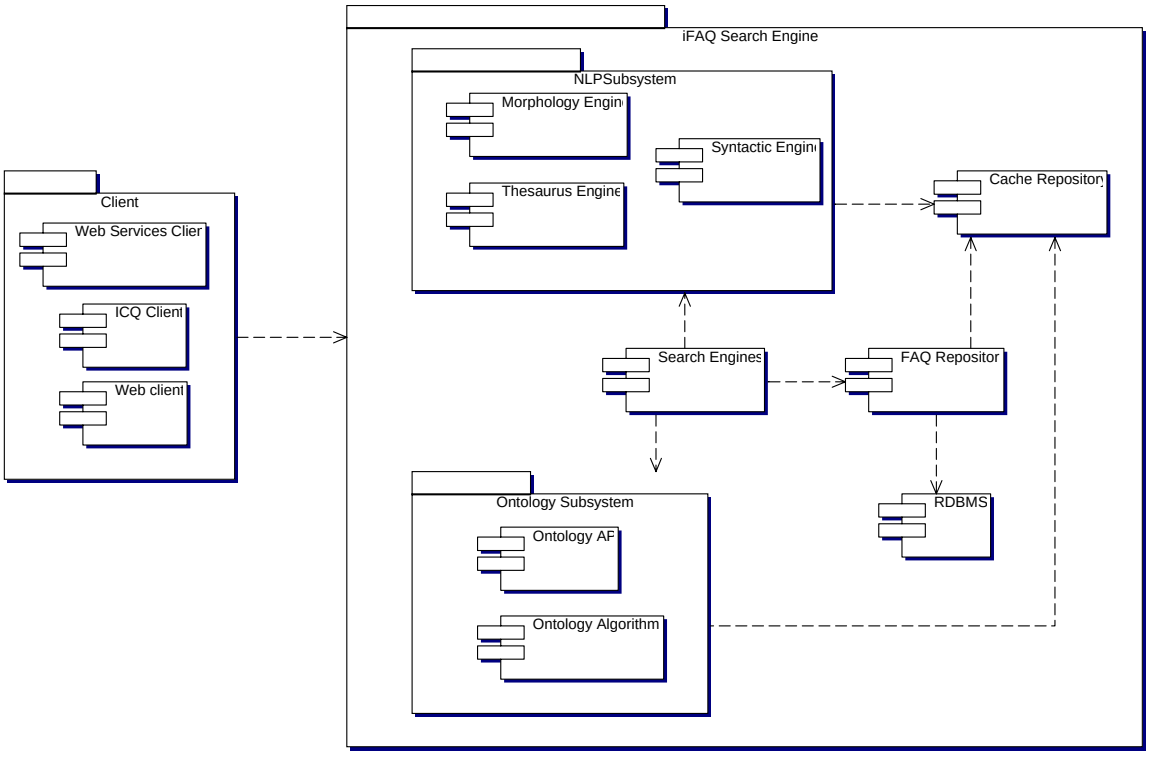


Figure 5.13: Component Diagram.

5.10 Architecture Diagram

Considerations of scalability, efficiency and memory limitation have led us to consider a distributed architecture with different nodes to process the user searching request:

- One or more nodes for NLP processing.
- One or more nodes for Ontology processing.

- One node to gather user requests as a main dispatcher.

All nodes exhibit their services with web services public interface. All connections between nodes use that architecture. This allows us to improve the system efficiency in terms of global memory and processing to fit itself to increasing demand of services. It is important to note that NLP and especially ontology processing tend to consume a lot of memory.

Next diagram depicts this architecture.

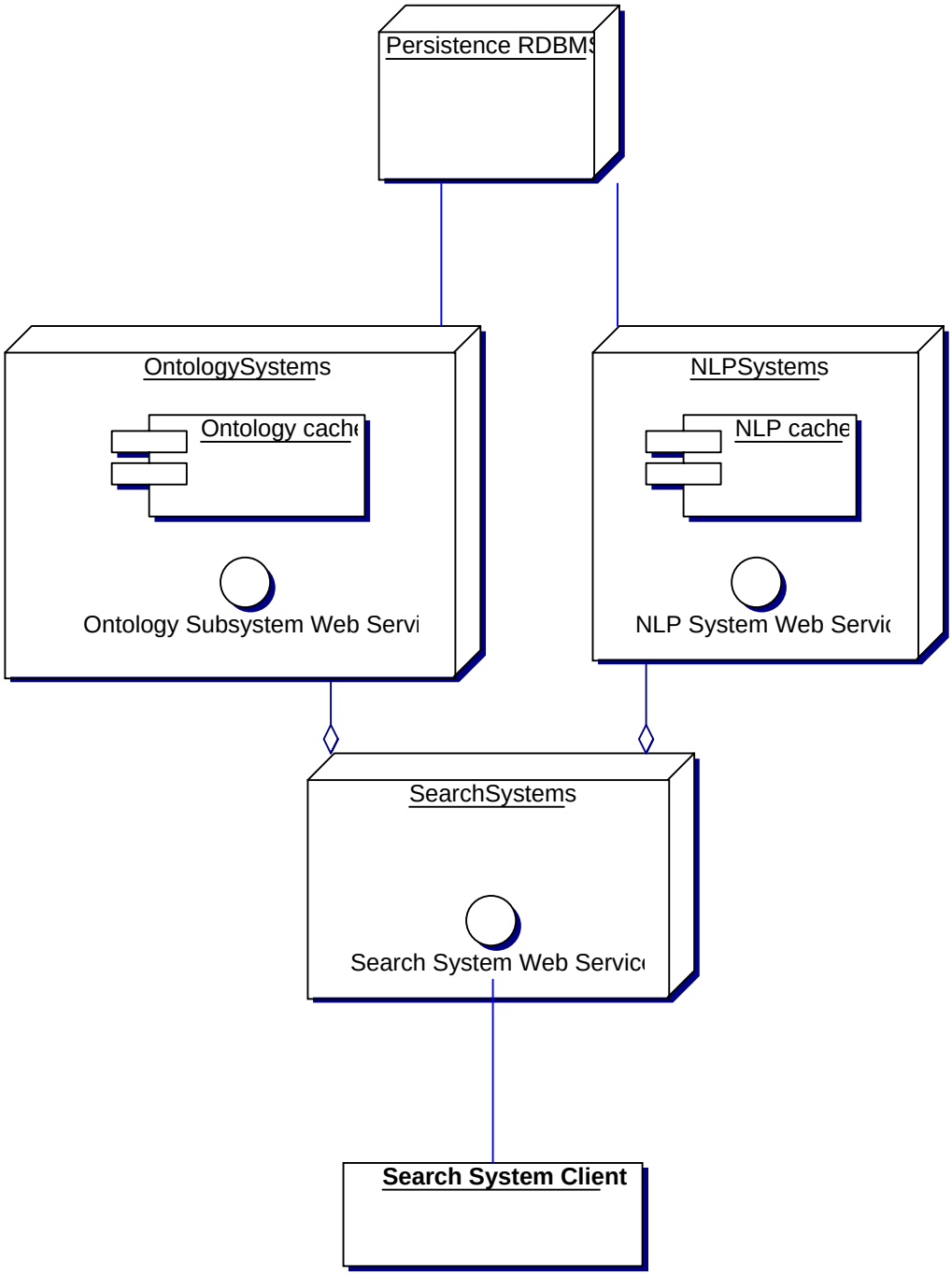


Figure 5.14: Architecture Diagram.

5.11 Ontology considerations

The Legal Case Study Prototype works with a special case of ontologies, that can be regarded as weighted graphs. That means that any arc from one node to another has a weight associated. Figure 5.15 depicts this idea:

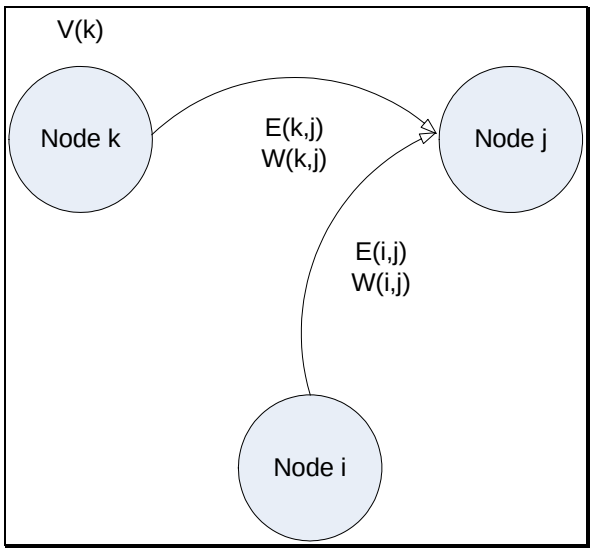


Figure 5.15: Weighted graph.

V{} stands for the graph vertex set, and E{} for the graph edge set. From the node i to the node j (both belong to V) there are only one edge: E(i,j). Every edge has a weight W(i,j).

In ontology language, the relationship from one concept (class or instance) to another is implemented with the use of a slot of a particular type, a relationship implemented by another concept with a slot that has a weight annotation. Next picture represent this idea:

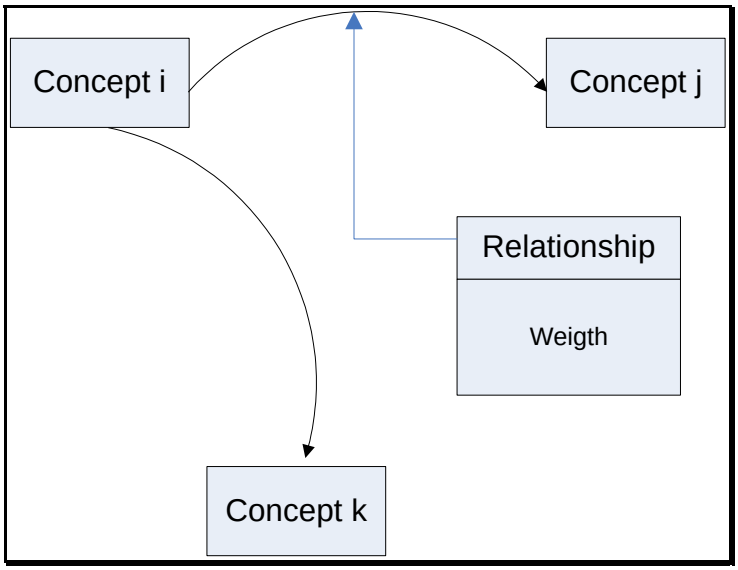


Figure 5.16: Weighted ontology.

5.12 Summary

Legal Case Study Prototype has been designed taking into account two main considerations:

- Accurate searching system, with advanced technology, that goes beyond traditional searching algorithms, capable of reliable search over a vast FAQ repository.
- A design that allows a fast, usable, modular, extensible, scalable, improving implementation.

The first point might be achieved by using some techniques like:

- Ontologies to model Legal Case domains. The system uses some specialized sub -domain ontologies that group themselves together to form a complete legal domain knowledge. These ontologies assist the system to achieve some intelligent knowledge of user questions. That extra knowledge will be used by the system to find the correct FAQ question-answer item.
- Natural Language Processing techniques, to analyze the user question and get morphological and syntactical information that help the system to understand the question in order to get more comprehension from ontology.

Second point is achieved by leveraging on some well-established software technology patterns:

- A multistage searching cycle for successive approach to FAQ target item.
- Pluggable searching stage engines.
- Use of previous computed cached data.
- Use of adapters for AI technologies like NLP o Ontology Processing.

We think this design is flexible, modular, scalable, customizable and suitable for this prototype. This design allows us to extend it with the use of pluggable engines to help to improve its searching capabilities. The use of engines pluggable also let us to improve its performance and accuracy by using new ones with improved technology.

6 Conclusions

In this Document we have depicted the legal scenario to develop SEKT technologies.

First, we have summarized our sociological findings on the Spanish legal system, the judicial context, real users (judges in their first appointment) and their technological skills. In the same Chapter, using ALCESTE, we have offered a preliminary analysis of the nature and structure of the questions we have recollected to feed the second prototype of *Iuriservice*.

In Chapter 2, we have summarized the state of the art and the ongoing research on legal ontologies. We have defined our epistemological starting point and we have established the general guidelines to build what we call Ontologies of Professional

Legal Knowledge (OPLK). We have described also the existing databases of Spanish legal documents (judgments, statutes, doctrine) and we have depicted in detail the internal composition and structure of the documents containing judicial rulings.

In Chapter 3, we have described more specifically the middle-out strategy that we are following to build our Ontology of Professional Judicial Knowledge (OPJK) and our use of some methodological and ontology editing tools (PROTÉGÉ, KAON, DILIGENT, ONTOCLEAN…).

In Chapters 4 and 5, we have developed a software architecture design for the Legal Case Study Prototype that has some main features:

- It is designed to be accurate and technological advanced by using NLP and ontological techniques.
- It is designed to be efficient, extensible, customizable, scalable, etc.
- It makes use of incremental search as a process of narrowing the solicited FAQ set.
- It uses a variety of pluggable searching algorithms.

This flexibility allows us to customize the prototype's behavior to achieve both the accuracy and efficiency desired.

Finally, in Chapter 8, we have translated from Spanish three kind of useful materials: (i) one recent Supreme Court Sentence (Tribunal Supremo) (8.1); (ii) one recent Sentence by the Provincial Appeal Court of Barcelona (Audiencia Provincial); (iii) the official Questionnaire we have used in our 2004 fieldwork research.

7 References

[A, 2004] Alba, M. de (2004): "El método ALCESTE y su aplicación al estudio de las representaciones sociales del espacio urbano: el caso de la ciudad de México". *Papers on Social Representations*. Volume 13, pp. 1.1-1.20. http://www.psr.jku.at/PSR2004/13_01Alb.pdf

[AAB, 2005] Alvarez, R.; Ayuso, M.; Bécue, M. (2005): "Statistical Study of Judicial Practices", in V.R.Benjamins, P.Casanovas, J.Breuker, A.Gangemi (eds.) *Law and the Semantic Web*, Springer-Verlag, London, Berlin (in press).

[ABAVAS, 2003] Ayuso, M.; Becue, M.; Alvarez, R.; Valencia, O; Alvarez, M.; and Santolino, M. (2003): "Análisis estadístico de las encuestas a los jueces en su primer destino", *Internal Report for the General Council of the Judiciary*. , Escuela Judicial de Barcelona, 2002. SEC-2001-2581-C02-01/02.

[B, 2002] Bastin, G. (2002): "Note sur la méthode Alceste : mondes sociaux et mondes lexicaux",
November 28 2002 http://www.melissa.ens-cachan.fr/imprimer.php3?id_article=200

[BF, 2004] Bedin, V.; Fourniet, M. (2004): "Les représentations des cadres de la net-économie sur les fonctions de la formation continue dans des contextes professionnels vécus comme «à risque»". *7ème Biennale Iternationale de l'Éducation et de la Formation*, april 2004.

[BCBG, 2005] Benjamins, V.R. ; Casanovas, P.; Breuker, J.; Gangemi, A. (2005): "Law and the Semantic Web, an Introduction", in Benjamins et al. (Ed.) *Law and the Semantic Web*, Springer Verlag, London, Berlin.

[BCBRCP, 2004] Benjamins, V.R.; Contreras, J.; Blázquez, M.; Rodrigo, L.; Casanovas, P.; Poblet, M. (2004): "The SEKT Legal Use Case Components: Ontology and Architecture", in Thomas F. Gordon, *Legal Knowledge and Information Systems. JURIX 2004: The Seventeenth Annual Conference*. IOS Press, Amsterdam, pp. 69-77.

[BCCABLU, 2003] Benjamins, V.R; Contreras, J.; Casanovas, P.; Ayuso, A.; Becue, M.; Lemus, L.; Urios, C. (2003) : "Ontologies of Professional Legal Knowledge as the Basis for Intelligent IT Support for Judges", *Workshop on Legal Ontologies and Web-based Legal Information Management*, held at *ICAIL 2003*, Edinburgh, June 2003.

[Be, 1982] Bénzecri, J.P. (1982): *L'analyse des donnés*. 4th ed. 2 Vols. Dunod, Paris.

[BT, 2004] Boella, G.; van der Torre, L. (2004): "Regulative and Constitutive Norms in Normative Multiagent Systems". http://di.unito.it/PS/Boella-vanderTorre-kr04.pdf

[BEPW, 2002] Breuker, J. ; Elhag, A.; Petkov, E.; Winkels, R. (2002): "Ontologies for Legal Information Serving and Knowledge Management". Legal Knowledge and Information Systems. Jurix 2002: The Fifteenth Annual Conference. Amsterdam. IOS Press, 2002, pp. 73-82.

[BEPW, 2003] Breuker, J.; Elhag, A.; Petkov, E.; Winkels, R. (2003): "IT Support for the Judiciary: Use of Ontologies in the e-Court Project".
http://www.lml.bas.bg(iccs2002/acs/breuker.pdf.

[BW, 2003] Breuker.J; Winkels R. (2003): "Use and reuse of legal ontologies in knowledge engineering and information management", ICAIL03 *ICAIL 2003 Workshop on Legal Ontologies & Web Based Legal Information Management*, Edinburgh, http://lri.jur.uva.nl/~winkels/legontICAIL2003.html

[CBCLU, 2003] Casanovas, P.; Benjamins, V.R.; Contreras, J.; Lemus L.; Urios, C. (2003): "*Iuriservice*: Un FAQ Inteligente Para Los Jueces En Su Primer Destino", in *Internet y Pluralismo Jurídico: Formas Emergentes de Regulación*, Pompeu Casanovas (Ed.), Editorial Comares, Granada.

[C, 2003] Chan, C.W. (2003): "Cognitive Modeling ad Representation of Knowledge in Ontological Engineering". *Brain and Mind* 4: 269-282.

[CSSH, 1998] Clancey, W.J.; Sachs, P., Sierhus, M.; Hoof,R.v. (1998): "Brahms: simulating practice for work systems design". *International Journal of Human-Computer Studies* 49: 831-865.

[CGPJ, 1999, 2003] Consejo General del Poder Judicial, Barómetros de opinión 1999 and 2003 [available at http://www.poderjudicial.es].

[CP, 2005] Contreras, J.; Poblet, M. (2005): "NetCase: An Intelligent System to Assist Legal Services Providers in Transnational Legal Networks". In: Benjamins et al. (eds.) *Law and the Semantic Web*. Springer, London, Berlin (in press).

[CDF, 2004] Corby. O., Dieng-Kuntz R., Faron-Zucker C. (2004): "Querying the Semantic Web with CORESE search engine". *Proceedings of the 16th European Conference on Artificial Intelligence*, ECAI'2004, including Prestigious Applicants of Intelligent Systems, PAIS 2004, Valencia, Spain, August 22-27. http://www-sop.inria.fr/acacia/pub/2004/corby-pais2004.pdf.

[D, 1995] D'Andrade, R. (1995): *The development of cognitive anthropology*. Cambridge University Press.

[DW, 2002] Dent, M.; Whitehead, S. (eds.). (2002): *Managing Professional Identities: Knowledge, Performativity and the 'New' Professional*. Routledge, London.

[EV, 2002] Engers, T.M van; Vanderbergue, R.A.W. 2002. The POWER light-version: Improving Legal Quality under Time Pressure". http://www.lri.jur.uva.nl/epower/publications/KnGov2002-5.pdf.

[E, 1992] Eraut, M. (1992): "Developing the knowledge base: a process perspective on professional education". In: Barnett, R. (ed.) *Learning to effect*. Open University Press, Buckingham, pp. 98-18.

[EJ, 2001] Escuela Judicial (2001): Datos Estadísticos de los Jueces en Prácticas (Curso 2000-2002).

[EJ, 2002] Escuela Judicial (2002) :Datos Estadísticos de los Jueces en Prácticas (Curso 2001-2003).

[EJ, 2003] Escuela Judicial. (2003): Datos Estadísticos de los Jueces en Prácticas (Curso 2002-2004).

[FGRT, 1992] Foo N., Garner B.J., Rao A., Tsui E. (1992): "Semantic distance in conceptual graphs". *Current Directions in Conceptual Structure Research*. (eds) P Eklund, T Nagle, J Nagle and L Gerhotz, Ellis Horwood, pp. 149-154. http://www.cse.unsw.edu.au/~ksg/Abstracts/Conf/sem_dist.html

[GB, 2002] Gangemi, A. ; Breuker, J. 2002. *Harmonising Legal Ontologies*. Ontoweb. IST Project 2000-29243. http://ontoweb.aifb.uni-karlsruhe.de/About/Deliverables.

[GPS, 2001] Gangemi, A. ; Pisanelli, D.M., Steve, G. (2001): "A formal Ontology Framework to represent Norm Dynamics." *Proceedings of the Second International Workshop on Legal Ontologies*, Amsterdam, 2001.

[GST, 2003] Gangemi, A.; M. Sagri, M.; Tiscornia, D. (2003): "Metadata for Content Description in Legal Information", in ICAIL 2003 Workshop on Legal Ontologies & Web based legal information management, June 2003, Edinburgh, Scotland, UK. http://www.lri.jur.uva.nl/~winkels/LegOnt2003/Gangemi.pdf .

[GCF, 2002] Gómez-Pérez, A.; Corcho O.; Fernández-López, M. (2002): *Ontological Engineering: With Examples from the Areas of Knowledge Management, E-Commerce and Semantic Web (Advanced Information and Knowledge Processing)*, Springer-Verlag, London.

[G, 1997] Gray, P. N. (1997): *Artificial Legal Intelligence*. Aldershot: Dartmouth.

[GF, 1995] Grüninger, M.; Fox, M. (1995): "Methodology for the design and evaluation of ontologies", in *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing* held in conjunction with IJCAI-95, Montreal, Canada, 1995.

[GP, 2004] Gu, T.; Pu, H.K. (2004): "Providing middleware support for context-aware services". *Journal of Network and Computer Applications* (1-17, in press).

[GW, 2002] Guarino, N; Welty C. (2002): "Evaluating Ontological Decisions with OntoClean", in Communications of the ACM, 45 (2002) p. 61-65.

[GW, 2004] Guarino, N. and Welty, C. (2004): 'An Overview of OntoClean', in S. Staab, R. Studer (eds.), *Handbook on Ontologies*, Springer Verlag, London, pp. 151-159.

[HV, 1999] Hage, J.; Verheig, B. (1999): "The law as dynamic interconnected system of states of affairs: a legal top ontology". *International Journal of Human-Computer Studies* 51: 1043-1077.

[JM, 2001] Jarr, M.; Meersman, R. (2001): "Practical Ontologies and their Interpretations in Applications –the DOGMA experiment". http://www.starlab.vwb.ac.be/publications/STAR-2001-04.pdf .

[K, 1995] Kralingen, R.W. van (1995). *Frame-based Conceptual Models of Statute Law*, Computer/Law Series, No. 16, Kluwer Law International, The Hague, The Netherlands.

[L, 1995] Lahlou, S. (1995): "Vers une théorie de l'interprétation en analyse des données textuelles". *JADT 1995. 3rd International Conference on Statistical Analysis of Textual Data*. S. Bolasco, L. Lebart, A. Salem (eds.). CISU, Roma, 1995, Vol. I, pp. 221-228.

[La, 2002] Lancia, F. (2002) : *The Logic of a Text-Scope*. October 28. http://www.tlab.it/en/textscope.htm

[M, 2004] Martínez-Vares, V (2004): "Nuevas tecnologías para la Justicia," in *Diario Expansión*, March 16th, 2004: 47.

[Mc, 1989] McCarty, L.T. (1989): "A language for legal discourse, I. Basic features". *Proceedings of the Second International Conference on Artificial Intelligence and Law*, pp 180-189, Vancouver, Canada.

[Mc, 2002] McCarty, L.T. (2002): "Ownership: A case study in the representation of legal concepts". *Artificial Intelligence and Law* 10: 135-161.

[MC, 1998] Menzies, T.; Clancey W.J. (1998): "Editorial: the challenge of situated cognition for symbolic knowledge-based systems". *International Journal of Human-Computer Studies* 49: 767-769.

[Mi, 1995] Miller, G. (1995): "Wordnet: A lexical database for English", in *Communications of the ACM* 38 (11): 39-41.

[Mo, 2001] Moscarola, Jean (2001): "Contributions des méthodes de l'analyse qualitative à la recherche en psychologie interculturelle. Sphinx et MCA". *Congrès ARIC.* Genève.

[OGGM, 2002] Oltramari, A.; Gangemi, A.; Guarino, N.; Masolo; C. (2002), "Restructuring WordNet's Top-Level: The *OntoClean* approach" http://www.loa-cnr.it/Papers/Oltramari2.pdf.

[P, 2001] Paliwala, A. (2001): "An Intellectual Celebration: A Review of the Jurix Legal Knowledge Based System Scholarship". *Artificial Intelligence and Law* 8: 317-335.

[P-G, 2000] Peyrat-Guillart, D. (2000) : "Une application de la statistique textuelle à la gestion des ressources humanies : appréhender le concept d'implication au travail de façon alternative". *JADT 2000 : 5es Journées Internationales d'Analyse Statistique des Données Textuelles*.

[PST, 2004] Pinto, S.; Staab, S.; Tempich, C. (2004): "DILIGENT: Towards a fine-grained methodology for Distributed Loosely-controlled and evolvInG Engineering of oNTologies", in *Proc of ECAI-2004*, Valencia, August, 2004.

[Po, 2004] Pommier, J.L. (2004): "Des variables tensives inscrites dans le texte: une interprétation dynamique de l'A.F.C. dans l'analyse d'Alceste". *JADT 2004: 7es Journées internationales d'Analyse statistique des Données Textuelles.*

[R, 2002] Reinert, M. (2002): "La tresse du sense et la méthode Alceste : Application aux «Rêveries du promeneur solitaire»". *JADT 2000 : 5es Journées Internationales d'Analyse Statistique des Données Textuelles*.

[R, 2003] Reinert, M. (2003): "Le rôle de la répétition dans la représentation du sens et son approche statistique par la méthode ALCESTE", *Semiotica* 147 – ¼, 389-420.

[Ro, 2000] Roche, C. (2000): "Corporate ontologies and concurrent engineering". *Journal of Material Processing Technology* 107: 187-193.

[RBCP, 2004] Rodrigo, L.; Blázquez, M.; Casanovas, P.; Poblet, M. (2004): D10.1.1. *Before Analysis. Case Study- Intelligent Integrated Decision Support for Legal Professionals. State of the art*. SEKT. EU Project IST-2003-506826.

[S, 1990] Schank, R. (1990): *Tell me a story. Narrative and Intelligence*, Northwestern Univ. Press.

[St, 1991] Stamper, R.K. (1991): "The role of semantics ni legal expert reasoning and legal systems". *Ratio Juris* 4(2): 219-244.

[St, 1996] Stamper, R.K. (1996): "Signs, Information, Norms and Systems". In B. Holmqvist and P. Andersen (eds.), *Signs of Work*, Berlin: De Gruyter.

[V, 1995] Valente, A. (1995): *A Modelling Approach to Legal Knowledge Engineering*. Amsterdam, Tokyo: IOS Press.

[VBB, 1999] Valente, A.; Breuker, J. ; Brouwer, B. (1999): "Legal modeling and automated reasoning with ON-LINE". *International Journal of Human-Computer Studies* 51: 1079-1125.

[Vi, 1995] Visser, P.R.S. (1995): *Knowledge Specification for Multiple Legal Tasks; A Case Study of the Interaction Problem in the Legal Domain*, Computer / Law Series, No. 17, Kluwer Law International, The Hague, The Netherlands.

[VB, 1998] Visser, P.R.S.; Bench-Capon, T. J.M. (1998): "A Comparaison of Four Ontologies for the Design of Legal Knowledge Systems". *Artificial Intelligence and Law* 6: 27-57.

[W, 2003] Winkels, R. (2003): Legal Ontologies. The development and re(use) of legal ontologies. http://www.lri.jur.uva.nl/~winkels/LegalOntologies.html .

[ZLY, 2002] Zhu, H.; Zhong, J.; Li J. and Yong, Y. (2002): "An approach for Semantic Search by Matching RDF graphs". http://www.dit.hcmut.edu.vn/~tru/SPECIAL-STUDIES/rdf-semantic-matching.pdf.

8 Appendix

8.1 Suppreme Court Sentence

Supreme Court, Second Criminal Division, Sentence of 3rd of May 2004.

Speaker: García Ancos, Gregorio.
Nº of sentence: 592/2004
Nº of appeal 1085/2003
Jurisdiction: CRIMINAL LAW

THE LAW JURIS: 13072/2004

HOMICIDE: The accused hit his wife, from whom he had separated as a matter of fact, with a blunt and elongated object at the back part of the head, causing her death. DOMESTIC VIOLENCE. Habitual physical abuse. Requirements of the *actus reus*. For the purposes of assessing whether the conduct is habitual, previous offences are taken into consideration, notwithstanding whether they have prescribed or not.

THE JURY. Offences under its competence. CRIMINAL PROCEEDINGS. Right of double instance.

In the city of Madrid, on the 3rd of May 2004

SENTENCE

In the appeal of cassation for Breach of Form, Breach of Law and Breach of Constitutional Rule we are dealing with, as brought by the accused Carlo María, against the sentence issued by the Provincial Appeal Court of Almería that found him guilty of the offence of physical abuse within the family and homicide; the Second Division of the Supreme Court, composed by the Illustrious Judges mentioned on the margin, have gathered to issue the Vote and Sentence, under the Chairmanship of the first of the persons mentioned, Mr. Gregorio García Ancos being the speaker. The Public Prosecutor is also a party to the proceedings and Isabel, Irene and Julia are also the private prosecution, as represented by the Court Attorney Mrs. Mª del Mar Hornero Hernández. The appellant party is represented by the Court Attorney Mrs. Consuelo Rodríguez Chacón.

1. ANTECEDENTS

1.- The Examining Court number 1 of Huercal Overa, commenced the examination under file number 1/2002 and when it finished, it raised it to the Provincial Court of Almería, on the 25th of September 2003, which delivered the sentence that contains the following proven facts:

"Proved that the accused Carlos María, of age, with no criminal records married Olga 25 years back. After a short time, the marriage started having problems due to the aggressive attitude of the accused, who even battered continually his wife, insulting her repeatedly, and threatened her loudly constantly. It becomes clear, for example, that on a certain not established date, but five or six years ago he hit her in the head

with a lantern so that she lost conscience for two hours without that worrying the accused, who left the place without giving her any assistance; about 3 years before he attacked her with a pair of scissors, bursting and cutting the dressing gown she was wearing, although the injuries caused for that reason are not known since Olga did not inform of such facts. On another occasion he tried to assault her with an axe and a knife, not attaining his purpose because their daughters present there prevented him from doing so. On another occasion he was going to set on fire the house where his wife was at that moment, but he did not succeed because his brother Felix stopped him. Two years before, already separate, when he was trying to take the car keys from his wife, which he indeed managed to do, although, as she resisted, he started throwing bricks at her, but he did not reach her but rather the car. He stopped attacking her on that occasion thanks to the intervention of the brother of the accused. The reiterated incidents continued with no interruption making Olga leave for Tarragona on three occasions, although she would return to the family domicile until the moment she reported to the police (to withdraw the charges later) against Carlos María on the 16th of May 1997 as a result of the threats suffered and the damages caused by her husband in the dwelling. Since it was impossible to live with the accused, she decided to leave in December of such year the domicile located in the neighbourhood called Palaces, in the municipality of Zurgena. She took the three daughters of the marriage with her, Isabel, Julia and Irene, who had seen the violent incidents of his father on their mother and who suffered also beating, threats and insults from their father. After this factual separation consented by the parties, the accused for quite some time and with no justification whatsoever, suspected that his wife had a relationship with another man and persisted in the same conduct and imposed on her a constant surveillance wherever she was, either in Zurgena or in the place she was working, with a threatening attitude towards her. Everybody became aware of his attitude in a place of little population, such as the Neighbourhood of Palaces. As things stood, on the 25th of December 2000, Olga, when coming from Zurgena, arrived in Palaces at 17,30 hours, remaining some minutes at Nieves home. Afterwards she went to a terrace planted with oranges that belonged to her, although a neighbour she held a short conversation with warned her that her husband could be nearby. She answered that she did not care and so she went on. The accused then saw her. He kept watering another terrace owned by both of them; then he approached her and, when he reached her, with no evidence as to whether they argued or not, at a certain stage, actually carried out the threats to death he had repeated so insistently, after a wrestle with the victim, as the external injuries of the corpse showed. The injuries consisted of three excoriations with the skin coming off in the second finger of the right hand, four excoriations with the skin coming off in the third finger of the right hand, haematoma in the internal side of the left wrist, in the internal side of the left shoulder two ecchymosed of 5 per 0,5 and 4,5 per 1,5 cm, three ecchymosis of 1cm that would converge, at the back of the neck several linear ecchymosis that take up all its width, injury incise contuse of 2cm of triangular morphology, in the left occipital area and pre orbitary haematoma and left continuous othorragy. He attacked her with a blunt and elongated object of unknown measures since he hid it, and he hit her in the back part of the head, causing her cranium encephalic traumatism affecting vital centres that caused death around 18 p.m. of that day. Later on the accused went back to his mother's whom he was living with, leaving his brother Félix first in the parking, next to the door, an unusual place, the rubber boots and a hoe remaining in the same place until the said brother appeared explaining what had happened, without

him showing any feeling whatsoever, not leaving the bathroom (where he was) at first but only later, when the Civil Guard came.

2.- The Instance Court gave the following sentence:

"SENTENCE.- We must condemn and we condemn Carlos María, as author criminally liable for the offences defined, one offence of continuous `physical abuse within the family and one of homicide, with the concurrence of any modifying circumstance that may modify his criminal liability to the following punishment:

A) For the first offence to TWO YEARS AND SIX MONTHS' IMPRISONMENT, as well as the cumulative punishment of specific disqualification to vote during the term of the conviction.

B) For the second offence, the punishment of 13 YEARS IMPRISONMENT, with the cumulative punishment of absolute disqualification for the term of the conviction.- We bar the convicted from returning to the place of the facts or to approach or relate to the victims, his daughters, by any means whatsoever, including the telephone for a term of FIVE YEARS, from the moment he leaves the penitentiary centre (which fact, the relevant authority shall notify them). As well as to the payment of the court costs, including those accrued by the private prosecution and to indemnify the persons damaged, Isabel, Julia and Irene in the amount of 60.101,12 euros increased in the legal interests.- The time he has been in prison by reason of these proceedings (unless applied to cancel other liability) is to be deducted from the time of conviction and it has to be proved when enforcing the sentence. – The separate piece leading to establish the relevant civil liability is to be sent to the examining judge so that he may proceed in accordance with the law."

3.- The sentence was notified to the parties and the appeal of cassation was prepared for Breach of Form, Breach of Law and Breach of Constitutional Rule, by the representation of the accused, Mr. Carlos María. This Second Division of the Supreme Court received the relevant certification in order to proceed with the proceedings, to form the relevant file and to formalise the relevant Appeal.

4.- The appeal brought by the representation of the accused Carlos María is grounded on the following motives of cassation: FIRST MOTIVE: BREACH OF FORM of paragraph 1 of section 851 of the Law on Criminal Proceedings. The sentence appealed does not express in a clear and definitive way the real moment when the accused made the actions he is accused for. In the "proven facts", there are expressions such as "on a certain date but five or six years back, "about three years before" "on another occasion" telling later on the different incidents of the factual story. These are time spaces fully uncertain, not fixed in the course of time. Fixing the time is very important to be able to determine, for example, which is the legal rule to apply or whether there is prescription, which amount to lack of clarity in the proven facts.-

SECOND MOTIVE.- BREACH OF LAW of number 2 of section 99 of the Law on Criminal Proceedings. Since it was taken as a proven fact by the sentence that the

person I am representing was the author of Olga's death, notwithstanding the Report on the autopsy, as well as the experts' reports carried out by the National Institute of Toxicology, we do not find any data that might relate to the person I represent with this fact, which means a clear mistake in the pondering of the evidence analysed, evidence that has not been contradicted by other evidence.-

THEIR MOTIVE.- BREACH OF CONTITUTIONAL RULE pursuant to number 852 of the Law on Criminal Proceedings and 5.4. of the Organic Law 5/1985.- The constitutional principles provided for in Section 24 of the Constitution have been breached and more specifically the right to grant an effective judicial protection with a due defence and the presumption of innocence. On the one hand, the evidence taken into consideration lacks any reasonable ground to deduct the conviction of the person I represent and, on the other hand the evidence taken into consideration by the First Instance Court only indicates that the facts took place as told in the list of proven facts and that there are not enough data to conclude that the person I represent is the author of the facts that are attributed to him and therefore, in this case, there are several possible explanations, all of them logical and possible as to how the facts happened and thus the most favourable interpretation should be adopted.

FOURTH MOTIVE.- BREACH OF LAW 849.1 of the Law on Criminal Proceedings by undue implementation of section 153 of the Criminal Code because the requirements set by the authors and by case-law have not been met in this case to gather that there is an offence of habitual Physical abuse within the family.-

FIFTH MOTIVE.- BREACH OF A CONSTITUTIONAL RULE under number 852 of the Law on Criminal Proceedings section 5.4 of the Organic Law on the Judicial Power because we consider that the constitutional rules contained in section 24 of the Spanish Constitution, more specifically the right to a process with all its guarantees and the right of the ordinary judge s predetermined by Law, since this proceedings were followed by the ordinary procedural steps, the court dealing with the matter being the Provincial Appeal Court of Almeria when the proceedings should have been followed in accordance with the Law on the Jury.-

SIXTH MOTIVE.- BREACH OF CONSTITUTIONAL RULE. Under number 852 of the Law on Criminal Proceedings and 5.4. of the Organic Law 5/1985. This party understands that in this case the right to an effective judicial protection has been infringed and the right to a process with all the necessary guarantees as well, since pursuant to section 14.5 of the International Agreement on Civil and Political Rights (PIDCP) this party should have been entitled to appeal in an effective way against the convicting sentence and the penalty imposed.

5.- The parties and the Criminal Prosecutor have learnt about the appeal, the Proceedings being concluded to give a date to issue the Sentence, as the turn indicates.

6.- Once the date for the Sentence was given, a vote foreseen for the 29th of April 2004 was held.

II. LEGAL GROUNDS

FIRST.- We sustain the initial motive of cassation, breach of form, under section 851.1º of the Law on Criminal Proceedings because the appellant thinks that it is clear that there was no clarity in the account of the proven facts.
This procedural shortcoming basically derives from lack of detail when putting forward expressions such as "on a certain date", "but five or six years back", "about three years before", "on another occasion", etc.

It is obvious that such expressions do not amount to the lack of clarity observed and they only mean and that when the evidence was analysed it was not possible to indicate with the necessary certainty the dates of the series of incidents that amounted to the mentioned offence of physical abuse. It would be totally absurd to pretend to require the court to express things in detail when such details are not known or they are only known approximately.

Furthermore, from the drafting of the motive, it is to be observed that the argumentation used in it refers to a question of substance, a dialectic which is not allowed when you are within the cassation path of the breach of form.

The motive is dismissed.

SECOND.- The correlative is done under section 849.2 of the Law on Criminal Proceedings by reason of mistake of fact when assessing evidence based on documents included in the file proceedings that show the mistake made by the judge and are not contradicted by other pieces of evidence.

Generally speaking we have to indicate that as case-law has repeatedly said (cases of 28th of November, 2003 and 31st of March 2004) one of the essential requirements of this cassational path is that the document or documents where the mistake on the facts (error facti) is to be based on, besides having a documentary nature, have to provide evidence by themselves of the fact that there was a mistake on some data or element that amounts to a direct proof, that is to say, it is not necessary to add any other piece of evidence, or have to resort to "circumstantial evidence or complex argumentations" or, likewise, among the facts explained in the sentence there appear elements in contradiction with those that the document proves by reason of its only nature and contents.

Another element necessary for the documentary evidence to be effective regarding any mistake claimed is that the documents should not be contradicted by other means of evidence practiced in the proceedings. The document has not value either because it has already been taken into consideration by the Court when describing the facts: otherwise, we would be before a new and different assessment of the evidence by the Court of First Instance, an assessment that corresponds in this case to this Court pursuant to section 741 of the Procedural Law, an important section regarding the principle of immediacy.
In the case we are dealing with, the appellant basis his claim mainly on the report on the autopsy and on reports carried out by the National Institute of Toxicology that sow, according to its thesis, that there is no data that relates the appellant with offence of homicide he was convicted for. Therefore, in order to deal with this claim we have to say the following: a) The autopsy is not a documentary piece of evidence for the purposes of the cassation appeal, since (although it has great importance as means of

evidences in general) it is a mere act that is documented as far as it is a result of the actual proceedings. As the appellant acknowledges, the only means of evidence susceptible to open the path of the mistake of section 849.2 are the documents that are to appear as submitted to the proceedings, the following not being considered as such: those stemming from the proceedings". b) With regard to the experts' reports, although they may have the document nature required it is necessary to assess them in contrast with other means of evidence, an assessment that has to be carried out in the most suitable way in order to preserve the presumption of innocence we will be dealing with in the next motive.

The motive is dismissed

THIRD.- For this reason, section 5.4. of the Organic Law on the Judicial Power serves to put forward that section 24.2 of the Constitution has been infringed as far as it proclaims the presumption of innocence.

Over and over again this Division and that of the Constitutional Court has said that in order to accept this presumption it is necessary to notice that the proceedings have not served to provide the necessary evidence, either for lack of means of evidence, or because they were obtained unlawfully, or else because the interpretation of such evidence was made by the authorised person in an irrational or illogical way in such a way that they decay or break in for of evidence that provides certainty beyond reasonable doubt PRUBAS DE CARGO or direct evidence or simply circumstantial evidence with adequate credibility to determine conviction.

In the case judged there is a lot of circumstantial evidence that invalidates the presumption of innocence invoked and that proves the authorship of the appellant in the facts. Some of those pieces of evidence are prior to the death of the victim, other are coetaneous and there are later ones. We can summarise them as follows:

As antecedents of the case we have repeated threats to death from the accused to his wife, the physical abuse he frequently exercised on her, as well as the fact that he kept on following and controlling her after she left with her daughters the domicile of the marriage. It has become clear (through the testimonies of the three daughters Isabel, Julia and Olga, who witnessed directly the series of acts of violence suffered by their mother over time, as well as through the declarations of Félix and Encarna, the brother and the sister in law of the accused, respectively, of Nieves, the sister of the deceased, of other family members and other various witnesses) that are the neighbours within a small neighbourhood.

-As proof related to the same day the incident took place, we have the information that the accused was that day (the 25th of December 2000) carrying out the task of watering the terrace owned by both spouses, located near the place the dead body was found at. This circumstance was proved by several witnesses that provided their testimony, with the relevant guarantees of oralness, contradiction and immediacy, who are the following: a) The marriage composed of Félix and Lorenza said in a clear way with no appreciable contradiction or inconsistency that on that day, about 17.30 (the death, according to the autopsy report, occurred around 18 p.m) they noticed that the accused was watering his terrace so that when they were coming from the property nearby, they met Olga (the victim) a quarter of an hour later, warning her that her

husband was in the thereabouts. They tried to warn her of the danger thast might affect her. However, Olga did not pay any notice to that and she went to the terrace she owned. b) Soledad said that on that day, around 17 p.m, Carlos María (the accused) was at her home where he would say that he was doing the watering but that "he had not finished yet", he would stay on chatting away for five minutes and he would leave just afterwards c) The accused was not at home between 17.30 p.m. and 19 p.m., according to his sister-in-law, Encarna. D) From other series of pieces of evidence gathered in the sentence it is possible to infer that the accused was in the place of the facts when they occurred.

--As subsequent facts or data, we have the statements of the brother of the appellant, Felix and his sister-in-law, who indicated that he had used his brother's boots because his were broken; he left them in an unusual place in his brother's garage. Likewise, Félix told that the accused was not surprised when he visited him to tell him about the facts and that he closed himself in the bathroom, not coming out despite the fact that he required him to. He only came out of the bathroom when the Civil Guard got there and asked him to.

In view of the above, we consider that the evidence proposed in order to release him from liability in the appeal is not enough. Such pieces of evidence have consisted of the following: a) the statements of the mother of the accused who, although she was ill and in bed, said that she knew his son had been at home while the incident happened, since she had heard him coming in and out of the rooms. These statements were not taken into account by the First Instance Court quite rightly, not only since the witness was ill, but also because of her lack of objectiveness. B) as to the autopsy, it is obvious that it is not enough to infer the author's liability but it is not enough either to infer that he is innocent, but rather the injuries caused, the characteristics of the weapon used in the aggression and other similar data of the utmost interest to show the facts as considered objectively, but that does not help to come to a conclusion on the authorship. C) as to the experts' report we have referred to in the previous motive, it is true that the experts have not detected that in the incident took part the accused, since there were no traces of him in the gadgets examined, but it is also true that this fact does not prove that, by way of exclusion of some specific data, that he is not the author of the death, when there is clear evidence to the contrary, as we have put forward before. D) finally the fact that the killing weapon was not found does not mean anything in favour of the murderer, since hiding the weapon or making it disappear is very frequent in these cases.

To conclude: the court "a quo" was right in its assessment of the evidence it examined since it did it according to logic and the rules of experience within the competence granted for the purpose by section 741 of the Law on Criminal Procedure in accordance with the principle of immediacy.

The motive is dismissed.

FOURTH. This fourth motive is put forward pursuant to section 849.1 of the Law on Criminal Procedure by undue application of section 153 of the Criminal Code that categorises the offence of physical abuse.

If we do not consider now the monographic survey made by the appellant in the appeal very smartly but inappropriately (as checked in the rest of the pleadings), the truth is that the objections are essentially reduced to two issues that are not included in the facts: the requirement of the marriage relationship between the passive and the active subject and, most of all, the requirement that one of the elements of the legal definition of the offence, that is, "habituallity" is not there.

On this issue, initially considered and accepted as a whole, we can say as a previous thing that, when putting forward the argumentation the proven facts of the sentence are not respected, but rather they are contradictory, which entailed their dismissal "a límine" from the motive, in view of the fact that the cassational path was used in his defence pursuant to section 884.3 of the Procedural Law.

However, we would like to answer appropriately this issue. If we stick to the facts of the sentence, as it is due, in the conduct of the accused over time we can see the existence of the requirements and elements of the legal definition of physical abuse since:

1. It is a proven fact, as so has been said in the section of proven facts, that the accused subjected his spouse for a long time, both before their separation as after it, to physical abuse consisting of constant battery, threats to death, beating her on one occasion with a lantern which caused her to lose conscience for two hours. On another occasion he attacked her with a pair of scissors, tearing and cutting the dressing gown she was wearing, on other occasions he tried to attack her with an axe and a knife and on another occasion he tried to set fire on the house where his wife was at the time, not managing to do so as a third person intervened, etc.

2. The passive subject of the action or victim was the spouse of the accused at the time the actions took place. The spouse to spouse relationship did not change after the separation of both (the attacker and the victim) as put forward in the appeal. It is obvious that one thing is the separation of fact and a another different thing the breaking of the marriage. Apart from that, it is clear that most of the attacks carried out took place before the separation, that is to say, when the marriage subsisted.

3. The requirement of habituallity, as a necessary element integrating the legal definition of the offence is the one the appellant questions the most in his pleadings. As a main point he claims that it is not possible to talk about habituallity since according to the description of the facts, many aggressive actions should be considered to be prescribed in view of their date. To oppose that we have to say that one thing is that certain facts that amount to a criminal offence may be considered as prescribed and another very different thing that those facts and actions may not have the consideration of evidence showing habituallity in the abuse. That is to say, since the accused is not exempted from criminal liability, we may deduce that the physical reality of his aggressive attacks may be blurred and have no effect of proof and therefore his insistence and habituallity in his actions.

   By reason of the statements made so briefly, we have to dismiss the motive.

FIFTH.- According to section 5.4 of the Organic Law on the Judicial Power, a claim has been made for breach of section 24 of the Constitution with regard to the right every citizen has to be judged by the ordinary Judge predetermined by law. More specifically, it is claimed that the competent body to deal with the facts was the Jury and not the Provincial Appeal Court.

The appellant says that the offence of homicide he has been convicted for corresponds to the mentioned Court and that the offence of physical abuse he has also been convicted for should have been dealt with in different proceedings since there is no connexion between the two offences.

We do not agree with that.:

1- The list of offences attributed to the Jury is a "numerus clausus" and cannot be extended to other offences, with the exception of the connected offences section 5 of the Organic Law on the Jury refers to, also in a list. On these lines, section 1.2.of the Organic Law specifically lists the jurisdiction of the Jury, and although with the list you can find homicide, sections 138 and 140 of the Criminal Code do not refer to the offence of physical abuse defined in section 153 of the Criminal Code, since, it is automatically excluded from such scope of competence.

2- The fact that the two offences have been judged in the same proceeding is fully justified in view of the fact that they are obviously connected since according to the description of the facts made in the sentence, the acts of violence and abuse of the accused on his spouse are clearly not preparatory actions but rather the antecedents that lead to death. The connexion between both actions is therefore obvious, which led to considering them in the same proceedings.

3- As we have said, the concept of "connection" is found in this case, such connection may not be included in the set of cases where the connection amount to extending the competence of the Jury pursuant to section 5 of the regulatory Law, since that Section lists the following cases in a closed list: a) the case where two o more persons in a meeting make the different offences simultaneously, b) the case where two or more persons make more than an offences in the different places or times, if they had come to a previous agreement; and c) the case where any of the offences was made in order to make others, facilitate their execution or obtain impunity. It is obvious that in the case we are dealing with, the first two cases are not applicable and as for the third one, although there is a correlation between both offences, such correlation does not amount to the physical abuse meant to allow homicide at a later time, nor was it intended to facilitate its execution or impunity.

The motive is dismissed.

SIXTH.- The last motive comes also under section 5.4. of the Organic Law on the Judicial Power because we consider that the constitutional principles of section 24 of the Spanish Constitution have been breached, more specifically the right to judicial protection and the right to a due process.

In essence, it is claimed that there is a breach of section 14.5 of the International Convention on Civil and Political Rights of 19th of December 1966, ratified by Spain on the 13th of April 1977 which recognised the right to appeal in criminal proceedings.

The Constitutional Court has dealt with the same very problem hereby raised in several decisions, mainly in the decision of 3rd of April 2002, as well as in the interlocutory decisions of this Supreme Court on the 14th of December 2001, 23rd of April 2002 and 16th of February 2004. In all those decisions there is a doctrine against the claims of the appellant that we may sum up as follows:

a) It is true that section 14.5 of such Convention, formally ratified by Spain recognises the right to a second instance where it says that "any person found guilty of a criminal offence shall have the right to have a higher court consider again the sentence and penalty imposed on him in accordance with the provisions of the Law." Equally, it is true that there is a report by the Committee of Human Rights of the Unities Nations of 11th of August 2000 according to which Spain would have breached this provision considering that the cassation appeal does not allow for the revision of the sentences and penalty imposed as required by the Convention.

b) However, although the provision of the Convention has become part of our internal law, it is not enough in itself to create non-existent appeals, most of all if we take into consideration that the cassation appeal in criminal matters may meet the requirements of the convention "provided that the possibilities of revision of the cassation appeals are interpreted widely and that the right recognised in the Convention is interpreted not as the right to have a second instance with full repetition of the proceedings, but rather as the right to have a Higher Court control the adequacy of the reasoning carried out in the firs instance, revising the correct implementation of the rules that have led to the declaration of guilt and the imposition of the penalty in each specific case". On these lines, we may indicate that the cassation appeal in its original form did not meet the requirement since it was sunk in a rigid formalism that prevented any revision of the proof, except for the one derived (with an exceptional nature) from the contents of some document that may provide proof with no contradiction of the mistake of the court of instance. Today, however, the Constitution and section 5.4. of the Organic Law on the Judicial Power have opened big expectations grounded on the breach of the rights of the individuals, particularly the infringement of the right to effective judicial protection and the presumption of innocence. We have evidence of that in the judgement we are considering which solves the problem of the presumption of innocence in its point THREE.

c) With regard to the report of the Human Right Committee of the United Nations, the mentioned case-law tells us that the "remarks" that the Committee issues in its Opinion do not amount to a judicial decision, since such Organ

> does not have judicial powers as can be deduced from sections 41 and 42 of the Convention, and its Opinions, thus, the authentic interpretation of such Convention, since neither in it nor in the facultative protocol of 16th of December 1966 it is granted such competence. Therefore, "if through its opinions the Committee decided to redefine the contents of the Convention, interpreting section 14.5 as the right to a second instance in a strict sense, questioning this way the internal system of appeal of one of the Member states and compelling it to enact new legislation in accordance with such interpretation, we would have to remember that , pursuant to the Resolution of the European Court of Human Rights of 30th of May 2000, the Member States have the power to decide the modalities of the right to re-examine and may restrict its scope".

In view of the above considerations, the last motive is dismissed.

III. SENTENCE

We must DECLARE AND WE HEREBY DECLARE THAT THE CASSATION APPEAL IS DISMISSED as brought by the representative of the accused Carlos María, against the judgement of the Provincial Appeal Court of Almeria of 25th of September 2003, in the proceedings followed against him for the offence of homicide and physical abuse within the family.

We hereby condemn the appellant to pay the judicial costs accrued in these proceedings. Give notice of this decision to the mentioned Court for the relevant legal purposes with return of the file if it happened to send it at the time.

Therefore, by this judgement, which is to be published in the Legislative Directory we hereby do, order and sign José Antonio Martín Pallín, Joaquín Giménez García, Andrés Martínez Arrieta, Francisco Monterde Ferrer, Gregorio García Ancos.

PUBLICATION

Read and published the above judgement by the Speaking Magistrate the Illustrious Mr. Gregorio García Ancos, in open court on the day and date, the Second Division of the Supreme Court, which I certify as Secretary.

The motive is dismissed.

8.2 Provincial Appeal Court of Barcelona

Provincial Appeal Court of Barcelona, 3rd Division, Sentence of 2nd of April 2003

Speaker: Bach Fabrego, Roser
Appeal Number: 6/2001
Jurisdiction: criminal

THE LAW JURIS: 10835/2004-12-17

DOMESTIC VIOLENCE: Protected legal interest. Interpretation of the term "habituallity": departing from the quantitative criterion. Differences between the violence the legal definition refers to and the specific violent actions, considered one by one Habitual physical abuse against his wife and children.

SEXUAL ABUSE. Traits of the intimidation and violence.
Draw differences vis-à-vis abuse of authority when the victim suffers physical abuse from the active subject. Declared innocent for such an offence. RAPE TO A MINOR. Placing in this legal definition the penetrations and touching performed on the daughters, by the author by abusing his parenthood and by not having been proved that at the time of attacking their sexual freedom he used violence or intimidation to annul their will. CONTINUOUS OFFENCE. Exceptional admission in offences against sexual integrity. Assessment. PRESCRIPTION. With regard to the rape to a minor against one of the daughters, who informed of the facts when more than 5 years had elapsed. THREATS. Assessment of the circumstances that surround the actions in order to make a distinction between an offence and a small offence. The fact that he put a pair scissors in the abdomen of his wife is categorized as a criminal offence because he threats that he is going to kill her. The aggravating circumstance of the persons been relatives is bound to exist, although there was a deterioration of the marriage. PUNISHMENT. Individualization of the punishment. Section 1973 of the Criminal Code applies, since it is the most favourable rule for the culprit. WITNESSES. Notes of the testimony of the victim when such piece of evidence is taken to be essential to establish the liability of the convict.

Text

In the city of Barcelona, on the 2nd of April 2003.
Seen in the public trial before the Third Division of the Provincial Appeal Court, these proceedings 6/2001 stemming from the Examining Court 3 of Cerdanyola del Vallés for the offence of sexual abuse against the accused Gregorio, of age, the son of José and Rosario, born in El Coronil (Sevilla), a neighbour of Ripollet (Barcelona), on pre-trial release in these proceedings, represented by the attorney Jordi Ribó and defended by the lawyer Mr. Manuel Lario de Merlo, the Criminal Prosecutor being a party to the proceedings.

Speaker: Mrs. Roser Bach Fabregó
She expresses the opinion of the Court

FACTS OF THE CASE

FIRST.- In his final conclusions the criminal prosecutor categorized the facts under the following offences:

Habitual physical abuse envisaged and punished by section 153 of the Criminal Code.
The offence of threats envisaged and punished in section 169.2 of the Criminal Code.
A continuous offence of sexual abuse envisaged and punished by sections 178 and 180.3 and 180.4 in relation with section 74 of the Criminal Code.
A continuous offence of sexual abuse envisaged and punished by sections 178 and 180.3 and 180.4 in relation with section 74 of the Criminal Code.
Four offences of sexual abuse envisaged and punished by sections 179 and 180.3 and 180.4, in relation to section 74 of the Criminal Code.

Alternatively to D) and E) a continuous offence of sexual abuse envisaged and punished by sections 179 and 180.3 and 180.4 in relation to section 74 of the Criminal Code.

The offence was found guilty of the mentioned offences as author, the offence of threat, the aggravating circumstance of relationship and he applied for the imposition of the following penalties, with the restriction set in section 76 of the Criminal Code.

For the offence of paragraph A) the penalty of two years imprisonment with the cumulative penalty of special disqualification of the right to vote for the term of the conviction.

For the offence of paragraph B) the penalty of two years imprisonment with the cumulative penalty of the right to vote for the term of the conviction.

For the offence of paragraph C) the penalty of ten years imprisonment with the cumulative penalty of the right to vote for the term of the conviction.

For the offence of paragraph D) the penalty of ten years imprisonment and absolute disqualification for the term of the conviction.

For the offence of paragraph E) the penalty of 15 years imprisonment and absolute disqualification for the term of the conviction.

Alternatively, the penalty of fifteen years imprisonment for the time of the conviction.

Likewise, the accused shall indemnify Mónica in the amount of 6.000 euros, Luis Antonio in the amount of 6000 euros, Carla in the amount of 6000 euros for the physical abuse and in the amount of 18.000 euros for the sexual assault and Paula in the amount of 6.000 euros for the physical abuse and in the amount of 18.000 euros for the sexual assault, and shall pay the judicial costs.

SECOND.- The defence of the accused considered that the facts amount to an offence of habitual physical abuse under section 153 of the Criminal Code, being liable as author. The defence applied the imposition of a penalty of one year imprisonment and special disqualification for the right to vote for the term of the conviction and considered that he shall indemnify in amounts going from 600 euros to 1.200 euros to

the informer DENUNCIANTE Mónica and the same amount to each one of his children Luis Antonio, Angel, Paula and Carla.

Likewise, he was found innocent of the other offences he has been accused of by the Criminal Prosecutor.

PROVEN FACTS
The accused Gregorio, of age, with no criminal records, married Mónica in 1972. From this marriage were born four children. From this date he has attacked his wife repeatedly and, likewise, he has repeatedly addressed insulting expressions to her in order to frighten her, taking out pair of scissors, knives and screwdrivers to manage to impose his will on her even in the presence of his children.

More specifically, on the 28th of February 1999, at 10. a.m., when Mónica was in the bathroom, his husband, after telling her "you are a bitch, what you have to do is go to work", he strongly pushed her, making her fall against the bath, without causing her any injury. When Mónica managed to get up from the floor, she went to the kitchen where the accused went in carrying a pair of scissors in his hands which he placed in her abdomen while he kept telling her that he was going to kill her.

The accused has also assaulted his children repeatedly and likewise he has thrown out expressions addressed to threaten them, taking out knives and scissors so as to attain his goal.

This situation lasted through out time until the year 2000 when the spouses separated.

From 1985 to 1991, the accused on several occasions went into his daughter's room, born on the 22nd of November 1978 and also into Paula's room, born on the 9th of July 1976, subjecting them to his touching and making them masturbate. In the case of Carla, he even came to the point of penetrating her through the anus on four different occasions, at the end of 1990. The conduct of the accused on his daughter Paula finished in 1989. The accused threatened to kill them and their mother, if they happened to tell the mentioned actions.

However, Carla told about it to the person she had a sentimental relationship with and later she also informed the police about it.

She informed the police on the 1st of March 1999.

LEGAL GROUNDS

FIRST.- The judgement of the Court was built on the basis of the evidence carried out in the oral trial with all the guarantees and in compliance with the principles of immediacy and contradiction.

In the first place, as regards the offence of habitual violence within the family, besides having been accepted by the defence in its final conclusions, the accused himself in the act of the judgement has acknowledged that the conduct amounts to the mentioned criminal offence. He has accepted that he is an aggressive person and that he will have an outburst and then he calm down", and sometimes he cannot control himself, he

feels sorry and he apologizes. He also said that he had given "big slaps" to his children and wife and that he had uttered threatening words to them and thrown gadgets to them.

Likewise, the wife's and children's declarations coincide in that they all say that they lived under constant pressure and aggressiveness provoked by the accused. Thus, Monica said that the accused had beaten her and threaten her on many occasion and said that he would usually beat his sons with a belt. More in detail, he said that in relation with the incident of the 28th of February 1999 his mother was in the bathroom, about to leave, when the accused prevented her from leaving, he pushed her against the bath and said to her that she was a bitch and when Mónica went into the kitchen, the accused followed her and he placed a pair of scissors in her abdomen saying that he was going to kill her. Such an incident was also accepted by the accused in his declaration in the trial, although he denied having threatened his wife with a pair of scissors admitting having held them in his hands.

Likewise, the sons of the accused have described the situation of constant violence he had placed his family in. Angel talked about the continuous arguments he has been witnessing for as long as he can remember and that he used to beat him and his brothers and sisters with a belt. When they were at the table, their father made them stay silent and he would show a knife, also saying that it was normal and that the son only realised of the situation when he left home and he met other families.

Carla and Paula said that their father used to beat and threatened them as a usual thing, that is, all of them: the daughters, the sons and the mother.

As for Luis Antonio, he said that he had witnessed how his father threatened his mother with knives and similar gadgets, at least on three or four occasions and that the accused had beaten him by using the belt leaving him with traces in his buttocks and that he remembers the situation has remained the same ever since he can remember. He also said the violent atmosphere has caused him psychological personality problems and that he has even required treatment.

As for offences against sexual freedom attributed to the accused, it is necessary to underline that this type of offence (since it takes place in secrecy by its own nature) counts on visual witnesses of the facts and therefore, the declarations made by the victim become vital and overriding, the Constitutional Court having repeatedly affirmed that the statements of the victim amount to a testimonial proof as long as it is carried out with due guarantees and that it has the virtue, as sole piece of evidence to contradict the constitutional presumption of innocence. However, when the testimony of the victim amount to an essential proof of the fact that the accused is guilty, it has to have the following requirements: 1) absence of subjective incredibility derived from the previous relationship of the accused with the victim that underline the motives that may raise doubts about the sincerity of the testimony, causing a state of uncertainty incompatible with the creation of a condemnatory conviction based on solid grounds; 2) Credibility of the testimony, that is to say, verifying that concur objective peripheral corroborations that guarantee what amounts to "the statement of the party, since the victim may appear as private prosecutor or civilly damaged in the proceedings (sections 109 and 110 of the Law on Criminal Procedure); and 3) persistence of the offence, which has to last over time, it has to be repeatedly

expressed and explained with no ambiguity or contradiction: since it is the only piece of evidence confronted to the declaration of the accused who proclaims his innocence, virtually, the only possibility of preserving the principle of due defence is to allow him to effectively question such an statement, underlining any contradictions that indicate that they are not true.

Pursuant to the case-law provided, the Court grants full credibility to the declarations made by the damaged female persons, first of all, because the Court had a direct and personal perception of the testimonies in the trial, which deserved a special credibility and, secondly, because the Court considers that all the requirements exposed above are met in this case.

First of all, there is a total absence of subjective incredibility in his statements. Precisely, the accused says that his daughters attributed him sexual assaults because when the marriage separated they supported the mother. That has no ground at all and there is no indication in the statements made by Carla and Paula that they wanted to take revenge, most of all if we take into consideration how difficult it was for her to make their statements before the Court in the trial. In the case of Carla, she even had to stop her statement for a moment and be assisted by the forensic doctor. Obviously, there is a bad relationship between the accused and Carla and Paula (perfectly logical if we stick to the facts, but it is not this bad relationship that has determined the charges they have made, the Court not having intended but to give an account of the facts as they indeed happened.

Secondly, the statements made by Carla and Paula have been coherent at all times. Indeed, on all the occasions where they explained the facts that amount to the charges of the case, they have been coherent, except for some discrepancies in Paula's declaration in the trial with regard to the charges made during the committal proceedings that she justified at the time when she could not remember some issues that came up during the interrogatory.

Lastly, the declarations made by the victims are trustworthy since they are corroborated by certain objective peripheral data. First of all, the accused himself, although he has denied at any time that he has not maintained any sexual relationship with his daughters, he accepted that perhaps he had been "too warm with his daughters", that sometimes he would sleep with them and he would caress them but never with a "sexual intention". Secondly, the accused wife, Mónica, said in the trial that it was habitual for her husband to get up from bed during the night, but she did not know what he was up to; likewise, Angel said in the trial that he remembers that at night as a matter of course, when he was in the corridor, he would see his father leaving the daughters' room.
If we consider the above, as the Court has pointed out before, this Court considers the declarations of the victims to be completely trustworthy and that they have the virtue of counteracting the presumption of innocence.

In the statement made in the trial, Carla (who places the actions of the accused within the period comprised between the ages of 7 and 13) said that at first her father would start touching her and when she was somewhat older he started penetrating her through the anus, pointing out that she could remember 4 penetrations of such a kind

in the late 90s. Also, she said that her father threatened her permanently, since he said that if she explained something, he would kill her and that when she was little she was not aware of the fact that her father's acts were wrong and that it was only after the anal penetrations that she thought "that can't be", since it was very painful.

As for Paula, she said in the trial that the abuse of his father started when she was very little (she was around 5). She also said that like her sister he started touching her and later she started to make her touch his penis and he would then place himself over her ejaculating, although she could nor remember whether there have been penetrations. She said that sometimes she had placed a chare to prevent her father form opening the bedroom's door, but he would get in anyway.

SECOND. The proven facts amount to the offence of habitual physical abuse typified in section 153 of the Criminal Code.

Quite precisely, the account of the "proven facts" shows that such kind of offence took place, particularly if we bear in mind that during the proceedings the accused accepted the categorization of the facts pursuant to the mentioned section of the Criminal Code. In fact, although only a violent incident has been declared as proven fact, the truth is that the accused had sunk the family in an atmosphere of continuous aggressiveness and violence, traits that characterise the offence described. Therefore, case-law has dealt with the requirement of "habituallity" departing from the quantitative criterion, considering that it concurs as long as the abuse remains and considering that "violence" in the legal definition described by the code means something different from the acts of violence considered as an isolated thing and that the legal interest protected is much wider and relevant than a mere attack on integrity, the values of the person being substantially affected (cases of the High Court of 24th of June 2000 and of 22nd of January 2002).

Certainly, the "physical abuse of a psychological nature" was included in the legal definition of section 153 of the Criminal Code after informing the police of the offence. However, the proven facts allow the categorization made since, as has been said, there was a conduct amounting to physical abuse against members of the family.

THIRD.- The proven facts amount to the offence of threats as described and punished by section 169.2 of the Criminal Code. As explained before, on the 28th of February 1999 the accused placed a pair of scissors in his wife's abdomen telling her that he was going to kill her. Fitting the facts of the case into the rule is clear-cut: the facts amount to threats and not to a small offence of the kind. In order to draw the difference between a small offence and an offence it is necessary to take into consideration the circumstances of the facts, such as, on which occasion are the threatening expressions made, the prior and simultaneous facts, in order to conclude that the aggressive and violent conduct of the accused on the members of the family determines that the evil promised appears as serious and possible.

Case-law defines the offence of physical abuse within the family as a ALIUD and an extra different from the actual acts of aggression, stating that the specific acts of violence serve only to prove the attitude of the aggressor and for this reason they are punished separately.

FOURTH.- The Criminal Prosecutor, in his final conclusions, charged the accused with the continuous offence of sexual assault (sections 178 and 180.3 and 4 of the Criminal Code) in relation to his wife Paula. A continuous offence of sexual assault of sections 178, 180.3 and 4 of the Criminal Code, all of the legal offences that are characterised for attacking sexual freedom by means of violence or intimation.

However, we consider that it is a proven fact that there was sexual abuse with abuse of parental authority, and not an offence of sexual assault: from the evidence practiced we cannot gather that at the time of attacking the sexual freedom of the minor, the accused did use violence or intimation to compel them, but rather at a first stage of the conduct followed by the accused, the daughters' will was not there in view of their age, and later on their will was vitiated and conditioned by the situation of superiority deriving from the parental relationship.

For that matter, in the trial, Carla said that when her father started to have sexual intercourses with her, she did not think it was a bad thing and she did not rebel against it, since she thought it was not possible for her father to do any wrong.

As for Paula, in the trial, she said that when her father did the touching to her or else made her touch her, she felt embarrassed, but she did not oppose to that due to her age and also because after such actions the accused would take her by the neck and threatened her so that she would say nothing.

In view of that, it is obvious that there was no violence or intimation in the sexual abuse of his daughters to compel them, but rather (as put forward before) the accused acted from a situation of superiority that conditioned essentially the discerning capacity of his daughters. And this Court acknowledges that there was constant tension and aggressiveness in the home by reason of the accused behaviour. Furthermore, although the two victims have said that after performing the actions described, the accused threatened to kill them so that they would not tell anybody about it, such conduct amounts to an examination ex post facto that, as such, may not amount to the typical intimation of the offence of sexual assault. On these lines, case-law has set that to meet the legal category of rape it is necessary for the intimation or violence to be rational and immediate, in such a serious and immediate way that it is used to overcome the victim's will. Therefore, it is not possible to consider whether it has gone further than the abuse of authority when the victim is influenced by a generic and diffuse threat, deriving from a prior abnormal **living together** with repeated threats of physical abuse, generally speaking, since the legal provision requires each sexual intercourse to be caused by the immediate and specific intimation leading to overcoming the opposing will shown by the victim in the coetaneous sexual relationship occurred. (Judgement of the High Court of 6th of October, 1998).

In view of the above considerations, this Court considers that the proven facts of the case amount to the offence of rape to a minor with abuse of authority pursuant to section 434 of the Criminal Code of 1973 in force at the moment the facts took place. More so because the facts against sexual freedom with sexual intercourse have been established as a result of the statements of the victim when she was twelve years old. The acts carried out by accused on her daughter Paula amount to a continuous offence of rape to a minor with the abuse of authority typified in section 436 of the Criminal Code.

It is true that case-law has been objecting to the continuity of the offence in offences against sexual freedom and, yet, it states that every time that there is an action against sexual freedom, (even where the victim is the same), there is a different offence and therefore there is a different legal action. There is a case-law line more modulated than this one that allows an exception to the general rule: it is reasonable to conclude that there is a continuous offence when there is homogeneity of the proven facts and there is an absolute impossibility to specify the occasion where the offences were performed, since it is more like the reality of the facts to group all of them in accordance with their material structure and their objective seriousness. More even so because the continuous offence is not envisaged exclusively as a mere legal fiction destined to solve the problems of implementation of the penalties when several legal categories of offence apply to the facts of the case, but rather it is envisaged as a real legal institution that allows us to build unitary proceedings over a plurality of actions that present certain objective and subjective unity (Judgement by the High Court of 7th of November 2000). Therefore, judgements dating from 21st of January and 23rd of March 1999 allow for the implementation of the application of the continuous offence when there is a homogeneity of the actions that derive from an only plan by the author as presided over by a sole intent that projects itself equally over the actions on a single victim in similar circumstances. Such case-law establishes that there has to be a sexual relationship over time that derives from a single intent or a single end or from profiting from similar occasions by the active subject that damages the same only victim.

In the case we are now analysing, the succession of actions against sexual freedom carried out by the accused against his daughters Carla and Paula indicate that there is a unitary intent, as well as the fact of taking advantage of similar occasions clearly indicates that there is a legal unit integrated by a series of typical actions that, in the case of Carla, the ones of a lesser seriousness, that is, the abuses with no sexual intercourse, must be integrated and absorbed in the continuous offences provided for in section 434 of the Criminal Code before mentioned, so that it is not possible to rely on the categorization made by the Criminal Prosecutor so as to punish as independent units the abuses without penetration and the abuses with penetration.

Notwithstanding the above, we have to point out that the proven facts in relation with Paula finished in 1991. As the charges were made in 1999, the offence has prescribed, since the term of 5 years envisaged in section 113 of the Criminal Code has elapsed in excess. That is so because initially calculating the term of prescription from the acquisition of legal age of the victim in the case of offences such as the ones dealt with in these proceedings was introduced in the Criminal Code of 1995 by the reform carried out by the Organic Law 14/1999 of 9th of June.

The judgement of the facts it to be carried out, pursuant to the legislation inf force at the time the facts took place, pursuant to the transitional provision nº 1 of the Criminal Code of 1995 in its original drafting. However, if we consider paragraph 2 of the mentioned section, it is necessary to examine whether any later legislation is more favourable. The Criminal Code of 1995 in force in its original drafting is the most favourable to the convict and in section 182 includes sexual abuse consisting of sexual intercourse with abuse of authority and attributes it a punishment of prison from one to six years in its higher half (from three yeas and six months to six years)

pursuant to paragraph one of the same provision, in view of the kinship relationship between the accused and the victim, which is also to be imposed in its superior half pursuant to section 74, which amount to imposing the punishment of imprisonment for a term of four years, nine months and six years.

In the Criminal Code of 1973, section 434 the offence is punished with probation isión menor, in the highest degree, since the accused is the ancestor ascendant of the victim, section 69 **bis** indicating that the punishment shall be imposed to the most serious infringement in any of its degrees so that it may be increase up to the medium degree of the highest penalty.

The defence of the accused did not choose one or another provision, but rather it restricted itself to indicate that the most favourable legislation shall be applied. Thus, the Court considers that it has to assess for the purposes of such a comparison the specific penalties it has to set and it considers that it is most favourable for the convict to apply the Criminal Code of 1973, since both legal texts provide for the maximum penalty of six years-imprisonment, since there are no motives to make use of the power provided for in section 69 bis as to imposing the highest penalty in degree and applying the text in force at the time the actions took place allows for the reduction of the penalty according the redemption of the penalty for work

FIFTH.- In the offence of threats the aggravating circumstance of the criminal liability, that of kinship, as set out in section 23 of the Criminal Code, a circumstance of a mixed nature that operates as aggravating in the offences that have a personal nature, such as in the case of threat and that according to repeated case-law it is to be applied even though the personal relationship between the spouses is being undermined (Decision of the High Court of 10th of February 2000).

In the other offences no modifying circumstance concurs.

SIXTH.- For the offence of physical abuse pursuant to section 153 of the Criminal Code, the penalty of two years of prison entails, as put forward by the Criminal Prosecutor, if we bear in mind that the conduct proved was serious and lasted over time and that there are several persons damaged by the offence, that is, the four children and the wife of the accused.

For the offence of threat, since the aggravating circumstance of kinship concurs, the accused deserves the punishment of one year and six months of prison.

For the continuous offence of rape to a minor with abuse of authority the accused deserves the maximum penalty by law, that is, six years of prison, if we bear in mind that the category of “continuous offence” covers conducts that the accused started when his daughter Carla was only five years old and that have continued until she was twelve years old.

SEVENTH.- Pursuant to section 116 of the Criminal Code, any person criminally liable is also liable under civil law, a liability that is expressly set in section 193 for the offences against sexual freedom.

If we take into account the nature of the crimes the accused has been convicted for and the moral damages that these type of offences causes, the amounts claimed by the Criminal Prosecutor are considered to be adequate, and the accused shall indemnify Mónica for the amount of 6.000 euros, Luis Antonio in the amount of 6.000 euros, Anges in the amount of 6.000 euros, Paula in the amount of 6.000 euros and all of them for the damages caused for the habitual physical abuse suffered by Carla in the amount of 6000 euros as compensation for the offence of habitual physical abuse and in the amount of 18.000 euros for the continuous offence of rape to a minor.

EIGHTH.- It is necessary to impose on the accused the payment of ¾ of the judicial costs, since he is declared innocent of one of the offences of abuse of authority, declaring the rest ex-officio.

As we have verified the legal provisions mentioned and the rest generally applicable

SENTENCE

That we find the accused Gregorio guilty as author of the criminal offence of habitual physical abuse without any modifying circumstances of criminal liability concurring) to the punishment of two years in prison with the cumulative punishment of specific for the right of vote for the term of the conviction; as author criminally liable for the offences of threats with the concurrence of the aggravating circumstance of kinship to the punishment of one year an six months of prison, with the cumulative punishment of special disqualification for the right of vote for the term of the convictions; and as author criminally liable for a continuous offences of rape to a manor with abuse of authority without concurring any modifying circumstances of criminal liability , to the punishment of six years imprisonment, with the cumulative punishment of special disqualification of the right of vote for the term of the conviction; and to the payment of ¾ of the judicial costs, the forth part being declared ex-officio.

Likewise, the accused shall indemnify Mónica in the amount of * euros, Luis Antonio in the amount of *euros, Angel in the amount of *, Paula in the amount of * and Carla in the amount of 24.000 euros.

Service is to be made that against this decision it is possible to bring an appeal of cassation for breach of law and breach of form within the term of five days.

We hereby sentence by agreement of the Magistrates noted in the margin of the page, a certificate of this sentence to be joined to the file; we bear witness.

8.3 Questionnaire

Information required from Judges trained in the Judicial School of Barcelona in their first post. Standard questionnaire as a guide for the interviews in field research.

TYPE OF COURT ORGAN			
Provincial Court		**Court**	
Division Nº:		First Instance	
Division		Committal proceedings	
Criminal Division I		First Instance and Committal proceedings	
Civil Division		Criminal	
Mixed Division		Contentious-Administrative	
		Social	
		Minors	
		Penitentiary surveillance	
		Civil Registry	

Note: The location of the Court, the personal and professional traits, as well as the personal opinions of the Judge are secret and confidential. These data are required merely to control the information, but shall not be rendered to the CGPJ.[52] The individual shall not be identified. Therefore, the answer given by the Judge is secret and anonimous, in order to guarantee the full reliability of the results.

This is a standard questionnaire deviced to guide the semistructured interviews in the place of work. The answers have to be provided in writing, independently of the use of other means of recording the information. The questionnaire contains a part about the court office (for the secretary) and another one, at the end of it, of a personal and confidential nature.

[52] CGPJ

Information required from the Judge or Magistrate

I Data about your training

a. The university where you did your Law degree:

b. Year when he finished the Law degree:

c. Before passing the contest did you practice any legal profession?:
(several possible answers)

(i) Lawyer
(ii) Court secretary
(iii) Prosecutor
(iv) Deputy Prosecutor
(v) Deputy Judge
(vi) Civil servant for the Administration of Justice
(vii) Other:_____
(viii) DON'T KNOW/NO ANSWER

d. Has he done postgraduate studies?
1. Yes
2. No
3. DON'T KNOW/NO ANSWER

e. In which field of law?:
(several possible answers)

(i) Criminal
(ii) Civil
(iii) Commercial
(iv) Labor
(v) Tax
(vi) Administrative
(vii) International
(viii) Other_____

(ix) DON'T KNOW/NO ANSWER

f. When did you pass the Contest to become a Judge?

g. Which year did you finish the Judicial School?

h. Do you belong to any Association for Judges and Magistrates?

1. Yes
2. No (*go* II. Personal Assessment)
3. DON'T KNOW/NO ANSWER

i. Which one?

(i) Professional Association of the Magistracy
(ii) Judges for Democracy
(iii) Francisco de Vitoria
(iv) Other:
(v) DON'T KNOW/NO ANSWER

II Professional Activity

1. Insertion

a. How did you feel when you first arrived at your Judicial Office?

b. Which are the most important problems that the court (or procedural unit) showed? Did you observe any problem of non-coordination between the procedural units of the different Courts?

c. What professional problems have you had in your Court Office? (Have you had any problem that affected you personally?)

2. Judicial matters

a. What kind of matters have required your attention the most in your first post?

1. Civil
2. Criminal
3. DON'T KNOW/NO ANSWER

b. What kind of civil judicial matter you have found the most difficult to solve during your first year?

c. Can you put forward the difficulties (dogmatic, technical or professional) you have had when solving these issues in the fashion of a **specific question**?

d. What kind of criminal judicial matter you have found the most difficult to solve during your first year?

e. Can you put forward the difficulties also in the fashion of a specific question?

f. Can you make a specific question about any queries or problems that have come up when you have had to take decisions when **on duty**?

g. What are the most frequent issues in your Court? (both in the first post and in the present one).

h. Do you need to discuss any court issue because your are worried about either the facts of the case or the legal solution?

1. Frequently
2. Only sometimes
3. Never (go to paragraph *s.*)
4. DON'T KNOW/NO ANSWER

i. In the affirmative, who do you usually discuss it with?

j. Which method do you usually follow to verify such an issue?:
(several possible answers)

(i) Personal interview
(ii) Telephone conversation
(iii) Seek advice through a letter
(iv) E-mail
(v) Other_____

k. What documents do you usually look up when solving a matter?:
(several possible answers)

Case-Law in paper
Case-Law in the data bank
Jurisprudence
Statistics
Sociological reports

Other_____
(vi) DON'T KNOW/NO ANSWER

l. How often do you read jurisprudence?

1. Rarely
2. Sometimes
3. Regularly
4. Frequently
5. Very frequently

m. Do you use Internet?

1. Yes
2. No
3. DON'T KNOW/NO ANSWER

n. What kind of information do you usually look for in Internet? Do you look for legal information that is useful to establish the facts of the case or its legal grounds?

o. What kind of information would you look for if you had a computer software that would d you to do it? What would you expect from a computer software that would provide professional assistance to the CGPJ?

III. Interprofessional relationships

a- Bearing in mind the following scale of assessment: 1, very negative; 2, negative; 3, medium; 4, good ; 5, very good; 6, DON'T KNOW/NO ANSWER, assess the professional relationship with the persons specified below:

Members of the Criminal Prosecution	1	2	3	4	5	6
Lawyers	1	2	3	4	5	6
Court Secretaries	1	2	3	4	5	6

The rest of the court personnel	1	2	3	4	5	6

b- How would you assess the professional relationship between judges, prosecutors and lawyers?:

1. Very negative
2. Negative
3. Medium
4. Good
5. Very good
6. DON'T KNOW/NO ANSWER

c- Do you keep professional contact with your contemporaries in the profession?
1. Yes
2. No (*go to paragraph e*)
3. DON'T KNOW/NO ANSWER

d- In the affirmative, how often?

6. Rarely
7. Sometimes
8. Regularly
9. Frequently
10. Very frequently
11. DON'T KNOW/NO ANSWER

e- Do you keep a professional contact with judges or lawyers from other?

1. Yes
2. No (*go to paragraph g*)
3. DON'T KNOW/NO ANSWER

f- In the affirmative, how often?

12. Rarely
13. Sometimes
14. Regularly
15. Frequently
16. Very frequently
17. DON'T KNOW/NO ANSWER

g- Can you tell us the profession of your best three friends?

	1	2	3
Magistrate or Judge from the year you passed the contest			
Magistrate or Judge from other years			
Criminal prosecutor			
Secretary			
Lawyer			
Other legal professions			
Liberal profession, doctor, engineer or technician, teacher			
Military man			
Executive, big entrepreneur, industrialist			

Salesman, small business man, artisan			
Civil Servant for the State, the community or the local administration			
Farmer, stock breeder, grape breeder, fisher...			
Qualified or non-qualified manual worker			
Service employee/ clerical work/ salesman			
Other (Which?)			

Interinstitutional relations

a. Bearing in mind the scale of assessment: 0, I do not have any relation/ I do not collaborate; 1, very bad quality; 2, deficient quality; 3, medium quality; 4, good quality; 5, Very good quality; 6, he does Don't know/he does not answer)

a.1- Can you tell us whether you have a professional relationship with the following institutions and assess the quality of this relationship?

Teaching corps of the Judicial School	0	1	2	3	4	5	6
Law Society of your district	0	1	2	3	4	5	6
Prison civil servants	0	1	2	3	4	5	6
National Police Force	0	1	2	3	4	5	6
Civil Guard	0	1	2	3	4	5	6
Autonomous Police	0	1	2	3	4	5	6
Local Police	0	1	2	3	4	5	6
Forensic Doctors	0	1	2	3	4	5	6
Psychiatrists (Non forensic)	0	1	2	3	4	5	6
Educators in Penitentiary Centers	0	1	2	3	4	5	6
Social Workers	0	1	2	3	4	5	6
Foreign language interpreters	0	1	2	3	4	5	6
Interpreters into autonomic languages	0	1	2	3	4	5	6
Other experts	0	1	2	3	4	5	6
Municipal bodies	0	1	2	3	4	5	6
Autonomic organs	0	1	2	3	4	5	6

a.2- Can you tell us whether you collaborate with teh following network of social resources and assess the quality of the collaboration?:

Rehabilitation centers	0	1	2	3	4	5	6
Battered women shelter	0	1	2	3	4	5	6
Asylum seeker center	0	1	2	3	4	5	6
ONGs	0	1	2	3	4	5	6

b. Have you been invited to any official act in the location of your District?

1. Yes
2. No (go to *7, relacions with different social sectors*)
3. DON'T KNOW/NO ANSWER

c. In the affirmative, where did the invitation come from and how often were you invited? (1, Rarely; 2, Sometimes; 3, Regularly; 4, Frequently; 5, Very frequently; 6, DON'T KNOW/NO ANSWER)

Autonomic powers	1	2	3	4	5	6
Local powers	1	2	3	4	5	6
Civil associations	1	2	3	4	5	6
Professional societies	1	2	3	4	5	6
Other (specify)	1	2	3	4	5	6

IV. Relations with different social sectors

a- As judge or magistrate, do you keep any kind of relationship with the media?

1. Yes
2. No
3. DON'T KNOW/NO ANSWER

b- What do you thing about the treatment the media are doing of the work carried out by the judges?

1. Very negative
2. Negative
3. Medium
4. Positive
5. Very positive
6. DON'T KNOW/NO ANSWER

c- As a citizen, do you keep any relationship with ONGs?

1. Yes
2. No *(go to paragraph f)*
3. DON'T KNOW/NO ANSWER

d- If the answer is affirmative, what type of relationship? (1: Sympathizer, 2: Collaborator, 3: Active collaborator, 4: Of direct legal support, 5: Of participation in the organization and direction, 6: Don't know/no answer)

Associations in aid of alcoholics and drug addicts	1	2	3	4	5	6
Associations in aid of battered women	1	2	3	4	5	6
Associations of immigrants	1	2	3	4	5	6
Associations in aid of minors	1	2	3	4	5	6
Neighborhood associations	1	2	3	4	5	6
Ecologists associations	1	2	3	4	5	6
Associations to cooperate with development	1	2	3	4	5	6
Associations for human rights	1	2	3	4	5	6
Health Associations	1	2	3	4	5	6
Other	1	2	3	4	5	6

e- Do you have your residence in the same location where the court you work at is placed?
1. Yes
2. No
3. DON'T KNOW/NO ANSWER

f- Where do you usually spend your weekends?

1. In the region where you live by reason of your job
2. In the region where you studied or else in the region where you were born
3. In another place _____

h- If you could choose, where would you practice as a judge?:

1. In the region where you life by reason of your job
2. In the region where you studied or else in the region where you were born
3. In another place_____

i- Do you belong to any club or association in the location where your Court is placed?

1. Yes
2. No *(go to 8, quality of life)*
3. DON'T KNOW/NO ANSWER

j- In the affirmative, what kind?:
(several possible answers)

(i) cultural recreational or musical
(ii) sportive
(iii) social
(iv) gastronomic
(v) religious
(vi) other_____

Quality of life

a. Do you usually take your work home?

1. Yes
2. No
3. DON'T KNOW/NO ANSWER

b. Do you usually work at weekends?

1. Yes
2. No
3. DON'T KNOW/NO ANSWER

c. What is the average of hours of work per week you spend apart from your timetable and time on duty):_____Hours/week

d. Do you think your work is well paid as compared with other jobs of similar qualification within the Administration?

1. Yes
2. No
3. DON'T KNOW/NO ANSWER

e. Assess the pressure you get from your work conditions (hearings, drafting the judgments):

1. Very low
2. Low
3. Medium
4. High
5. Very high
6. DON'T KNOW/NO ANSWER

f. Could you assess the degree of satisfaction you experiment with your daily work (as you live it)?

1. Not satisfied at all
2. Little satisfied
3. Quite satisfied
4. Satisfied
5. Very satisfied
6. DON'T KNOW/NO ANSWER

g. Why? Could you indicate the reasons for the above assessment?

V. Personal Data

a. Year you were born: _____

b. Sex: Male _____ Female____

c. Civil Status (including circumstances of fact): _____

d. Number of children: _____

e. Autonomous Community of your birth: _____

f. Size of the town of your birth:

1. less than 5000 inhabitants
2. from 5001 to 50000 inhabitants
3. from 50001 to 100000 inhabitants
4. from 100001 inhabitants to 500000 inhabitants
5. more than 500000 inhabitants

Can you indicate your mother's and father's profession (the one they dedicated to most of their lives' time)? And the profession of your couple, if you have one.

	father	mother	couple
Magistrate or Judge			
Prosecutor			
Secretary			
Lawyer			
Other legal professions			
Liberal profession, doctor, engineer or technician, teacher			
Military man			
Executive, big entrepreneur, industrialist			
Salesman, small business man, artisan			
Civil Servant for the State, the Community or the local administration			
Farmer, stock breeder, grape breeder, fisher			
Qualified or non-qualified manual worker			
Service employee/ clerk/ salesman			
Housewife			
Other			

Information required from the Judicial Secretary

The goal of this second part of the work is to gather information about the different aspects related to the infrastructure, organization and the Court Office or Procedural Unit where the interview takes place. The data gathered are to be completed with the information that the Center of Judicial Documentation has in the CGPJ.

Organizational modality of the Court Office (or Procedural Unit)

1. Physical space

a. Number of dependencies of the Office or Unit:
b. Square meters of the Office:
c. Number of tables:
d. Number of computers in the Office:
e. Number of computers that have a connection with Internet:
f. Number of printers:
g. Number of typing machines:
h. Is the follow-up of the proceedings computerized?
    1. Yes
    2. No
    3. DON'T KNOW/NO ANSWER

i. How often is such a system used?
    1. Always
    2. Sometimes
    3. Only Sometimes
    4. It is not used
    5. DON'T KNOW/NO ANSWER

j. What kind of computerized system does your Court use?
    1. LIBRA
    2. TEMIS
    3. ADRIANO
    4. PIAJ [EJ]
    5. MINERVA
    6. Other:

k. Are you satisfied with the performance of the system?

l. What specific problems have you had with the system or what services do you consider that could be improved?

m. Is there a multimedia system for the recording of the hearings?
   1. Yes
   2. No
   3. DON'T KNOW/NO ANSWER

n. Is there a deposit of conclusive pieces of evidence and files?
   1. Yes
   2. No
   3. DON'T KNOW/NO ANSWER

o. Is there any space shared with other bodies (oficinas paisaje)?
   1. Yes
   2. No
   3. DON'T KNOW/NO ANSWER

p. Type of Hearing Room: Own____ Shared____

q. How would you assess the maintenance condition of the facilities of your Court Office?:
   1. Very bad
   2. Bad
   3. Medium
   4. Good
   5. Very good
   6. DON'T KNOW/NO ANSWER

2. Division of the working time

a. The persons mentioned below have a working burden that obliges them to work apart from their timetable at the court or on duty? (circle the correct answer)

Court Secretary	YES	NO	DON'T KNOW/NO ANSWER
Officers	YES	NO	DON'T KNOW/NO ANSWER
Auxiliary officers	YES	NO	DON'T KNOW/NO

			ANSWER
Usher	YES	NO	DON'T KNOW/NO ANSWER

b. Approximately, what is the average of working hours per week done apart from the work within the timetable or while on duty?:
   1. Court Secretary: _____(hours/week)
   2. Officers: _____(hours/week)
   3. Auxiliaries: _____(hours/week)
   4. Ushers: _____(hours/week)

3. Organization and level of communication

a. In your opinion, what are the most important problems regarding organization and functioning of the Court Office (special units of direct support, common procedural services, administrative units)?

b. How do you see the relationship of your work with the citizen? What kind of complaints do citizens make when they do make one? What do you think could be improved?

For the interviewer

General comments on the development of the interview

_____
___