Appeared online at CEUR-WS.org/Vol-1105

JURIX2013-DoCoPe
Doctoral Consortium and Poster Sessions

Proceedings of the First JURIX Doctoral Consortium and Poster Sessions
in conjunction with the 26th International Conference on Legal Knowledge and Information Systems

Bologna, Italy, December 11th-13th, 2013.

Edited by

Monica Palmirani *
Giovanni Sartor **

* University of Bologna, CIRSFID, Faculty of Law, Bologna, Italy
** European University Institute, Law Department, Fiesole, Italy

Table of Contents

Session 2: Poster Session

- Legal Conflict Detection in Interacting Legal Systems
  Tingting Li, Tina Balke, Marina De Vos, Julian Padget, Ken Satoh, Dept. of Computer Science, University of Bath, UK, Centre for Research in Social Simulation, University of Surrey, UK,Principles of Informatics Res. Division, National Institute of Informatics, Japan

- From Oral Hearing to Opinion in The U.S. Supreme Court
  Latifa Al-Abdulkarim and Katie Atkinson, Trevor Bench-Capon, Department of Computer Science, University of Liverpool, UK

- Addressing Argumentation Puzzles with Model-based Diagnosis
  Giovanni Sileno, Alexander Boer, Tom Van Engers, Leibniz Center for Law, University of Amsterdam

- Legal Rules Modelling Meets the Web
  Monica Palmirani, Luca Cervone, Octavian Bujor, CIRSFID, University of Bologna

- #Unfair #Law: Folksonomies & Law between Openness and Knowledge
  Federico Costantini, Dipartimento di Scienze Giuridiche, Università degli Studi di Udine

- Network Analysis to Support Navigation and Use of Ever-Changing Legislation
  Guido Boella, Luigi di Caro, Livio Robaldo, Andrea Violat, Department of Computer Science, University of Turin and Nomotika s.r.l.

- Automated Methods for Extracting and Expanding Lists in Regulatory Text
  Alan Buabuchachart, Nina Charness, Charness, Katherine Metcalf, Leora Morgenstern, Leidos, Arlington

- Social Media Policy and Non-Profit Organization: Exploring the Legal Ramifications of Being Social
  Naomi Augara, Nussen Ainswortha, Marta Poblet, John Zeleznikow, Victoria University, Melbourne Australia and RMIT University, Melbourne, Australia

---

Preface

This volume collects the fourteen contributions selected for the First International Jurix Doctoral Consortium and Poster Sessions, organized in conjunction with the 26th International Conference on Legal Knowledge and Information Systems, held in Bologna on December 11–13, 2013.

This is the first year that Jurix has hosted Doctoral Consortium and Poster Sessions, and the aim is to provide a proper space for young Ph.D. researchers in AI&Law, while encouraging a constructive and fruitful dialogue between the senior scholars and the emerging generation of researchers. The doctoral consortium enables students to interact with academics and experts in the field who can evaluate their research projects from both a theoretical and an applicative point of view. Young researchers have an opportunity to present and discuss their ideas in a dynamic and friendly setting, while the AI&Law community can support the new generation of researchers in carrying forward the interdisciplinary method.

We have organized the proceedings in two sessions in accordance with the conference program: a Doctoral Consortium session and a Posters session.

Included here are the six papers from the doctoral consortium: three are selected from the Computer Science background and the other three from the Law background. In this way we achieve a good balance between the legal side and the technical side, integrating the two so as to firm up the interdisciplinary foundation within the AI&Law community. The topics addressed include legal argumentation, online dispute resolution, eDiscovery and NLP techniques, and legal-knowledge modelling.

The Poster Session includes eight extended abstracts outlining corresponding posters presented at the Jurix2013 conference: three posters present application tools and demos; the other five present theoretical investigations supported by a strong evaluation phase. The topics include legal argumentation, legal-conflict detection, legal-rule modelling tool via the Web, social-network analysis applied to the legal

system, NLP parsing for detecting lists in regulation, social-media policy, and nonprofit organizations.

We would like to warmly thank all students, supervisors, and referees and all the members of the program committee and the organising team, for they have made the First International Jurix Doctoral Consortium and Poster Sessions a huge success and an excellent opportunity to enrich the AI&Law community with new emerging ideas.

Monica Palmirani
Giovanni Sartor

Dialogue Interactions in Oral Hearings

Latifa Al-Abdulkarim

Department of Computer Science
University of Liverpool
latifak@liverpool.ac.uk

Abstract. The human reasoning process used in conducting arguments to resolve conflicts and reach a decision is an interdisciplinary study. Modelling argumentation has a great impact on the development of theories and applications in AI especially in critical domains that involve richness of reasoning such as Law. Therefore, argumentation has been recognised as a core topic in AI and Law. Developing computational argumentation systems that simulate legal reasoning requires resolving distinctive challenges that concern contextual rules, procedural issues and the interpretation of different elements.
This PhD research aims to provide a supportive computational model for analysing stages in the Supreme Court starting from the oral hearings. In particular, a key aim is analysing the social values of the legal arguments from different perspectives, and finding the relation between the Court opinion and the components constructed in the Court oral hearings. This paper gives an overview of the current PhD research proposing the problem, the overall aims and the approach required to fulfil the aims. Furthermore, it provides a summary of the background in argumentation in the domain of AI and Law and presents the deliverables obtained so far.

1 Research Problem and Aims

This work aims to develop a framework for deliberation through which Supreme Court oral hearings can be analysed in order to identify the components from which arguments are constructed for delivering the opinion. Following on from the analysis, a dialogue system will be defined to capture the moves made during the hearings. This dialogue system will be validated through application to selected cases in various courts. It will then be investigated how the framework can be generalised so that it can be applied to other domains in which deliberative reasoning occurs.

In comparison to other contexts, dialogues in the legal domain combine arguments from different sources, i.e. argument about the case evidence and facts, argument from legal rules, argument from precedent cases, argument from hypothetical tests and others which are required to resolve the ambiguity of the conflict issues. However, the structure of exchanging arguments in legal dialogues is not clear, the argument types are interleaved and there is no particular order for the parties to pose arguments which makes the analysis of the oral hearings more complicated.

Furthermore, coming to a decision in a legal case dialogues is a separate process that requires legal analysis in order to derive the case facts, apply the facts to the current law, which intended to reflect the values of society, and announce the decision that is limited

to two outcomes (deciding for plaintiff and decide for defendant), in a form of Court opinion that explores the arguments supporting the decision. Modelling these aspects provide challenges in the computational development in the domain of AI and Law.

Thus, in particular, this PhD research is initially attempting to fulfil the following aims:

- Define a representation based on conflicts in social values that enable to show the components of the arguments for the U.S. Supreme Court oral hearing dialogues.
- Provide a full analysis of the oral hearing dialogues by studying a particular legal case study using the defined representation.
- Develop a dialogue system by defining speech acts and a dialogue protocol .
- Analyse the subsequent Court opinion arguments by finding the relation between the argument components that emerge from the oral hearings through selecting and justifying the options.
- Demonstrate the model using other Supreme court cases and other jurisdiction.
- Generalise the model to handle deliberation dialogues in non-legal contexts.

2 Background Research

Extensive work has enriched the domain of argumentation in AI and Law over the last 25 years [8]. The nature of legal dialogues emerge different types of arguments which result in different types of argumentation schemes such as argument from analogy, argument from expert opinions, argument from rules and others. These argumentation schemes are used extensively in modelling legal reasoning.

Thus, modelling reasoning with legal cases has been a central topic of AI and Law from the beginning, and there is now a good degree of consensus, especially with regard to the main elements involved. This consensus can be expressed as a tree of inference with a legal decision as the root and with evidence as the leaves. Between the two there are a number of distinct layers.

Immediately below the decision there is a level of issues [10], or values [6], which provide the reasons why the decision is made. The idea here is that laws are made (and applied) so as to promote social values: whether a value is promoted or not is an issue. Where more than one value is involved and they point to different decisions, the conflict needs to be resolved. Sometimes it is appropriate to give priority to one value over another (as in [6]), sometimes a balance needs to be struck (as in [10]). Note that the relation between issues may be seen as a matter of ordering, or requiring a balance between the values: there is as yet no consensus on this point [7].

At the next level down there are a number of *factors* [3]. Factors are stereotypical fact patterns which, if present in a case, favour one side or the other by promoting a value, and so are used to resolve the issues and permit comparison between the cases. Sometimes (as in [3]) it may be convenient to group several factors together under more abstract factors, so that we may have two or three layers of factors, moving from the base level factors through more abstract factors, before reaching the issues.

Below the factors there are the fact patterns used to determine their presence. These facts supply reasons for and against the presence of the factor which need to be considered and weighed to make a judgement. At the lowest level there is the evidence.

Facts are determined by particular items of evidence, and where evidence conflicts a judgement will need to be made: often this judgement is made by a jury of lay people rather than lawyers. In the lower courts there will be real items of evidence, but by the time a case reaches the Supreme Court, the facts are usually considered established and beyond challenge. one example of work concerned with this stage includes Gordon's Pleadings Game [11], which identifies which facts are agreed by the parties and which will require resolution in the trial itself.

Thus a complete argument for a case will comprise a view on what can be considered as evidence for relevant facts: what facts are required to establish the presence of various factors, and how they relate; how the factors can be used to determine the issues; and, where issues and values conflict, how these conflicts should be resolved. In the first stage of this research we will show how these elements relate to the individual and collective goals of the oral hearing dialogues.

3 Research Methodology

Towards the main aim of establishing a deliberative reasoning framework, this research is intended to follow a qualitative method as shown in the following description.

Oral Hearings Analysis and Representation The research begins by investigating the dialogue interactions in the legal domain, particularly in the U.S. Supreme Court oral hearings. After that, we analyse the oral hearing transcript using a case study from the domain of AI and Law and provide a preliminary representation of the arguments components exchanged at this stage. Based on the defined components, we also define the speech acts required to capture the moves in the oral hearings and construct the arguments representation for every party in each oral dialogue.

Opinion Analysis and Representation Following the oral hearing, we define a representation for the court opinion and analyse the decision by navigating through the arguments constructed from the oral hearings.

Dialogue System Toward automation for this analysis, we develop a dialogue system to support the process of constructing the argument components representation using the defined dialogue moves. Throughout this stage, we need to define the legal case ontologies for the components to provide the grammar that set out the rules for how the components can be combined and constructed. Further more, we will define the algorithm that is required to navigate through the trees to provide the analysis for the decision made and propose *a protocol* for the oral hearings dialogue.

Evaluation In order to evaluate the system, we apply the implemented system to a number of supreme court cases and cases from other jurisdictions to identify required modifications of the ontologies and representation.

Generalisation Finally, the work elements will be combined together, *generalise* the developed system to include deliberative dialogues in non-legal contexts, and conduct an empirical evaluation and theoretical analysis on the final version.

4 Research Results

This section shows an overview of the findings we obtained so far. First we will investigate the dialogues in the oral hearings stage and identify a normative set of speech acts which could be used by counsel and Justices to achieve their dialogue goals. After that, we explain how the illocutionary force of these acts can be represented in terms of a developing *Arguments Component Tree* (ACT). For illustration, we use the transcript of the oral hearing in *California v Carney* , and give the resulting ACTs. Finally, we relate the ACTs to the majority and minority opinions in *Carney*.

4.1 Dialogues in Oral Hearings

There are three nested dialogues in the main oral argumentation dialogue of the Supreme Court. Each of the three dialogues will involve a counsel and nine justices. Prior to analysing the arguments in the oral hearings dialogues, we will describe the initial situation, the individual goals and the collective goal for Oral hearings accordingly.

In the *initial state* of the petitioner presentation, briefs from the petitioner, respondent and any "friends of the Court" are available. These will set out (and justify) a set of tests forming candidate arguments: the arguments of each counsel will, if accepted give rise to a decision for their clients. These briefs will also state the accepted facts of the case, and draw attention to relevant precedent cases. The *collective goal* is to obtain a clear statement of the argument for the petitioner. Individually *the counsel* will wish to state his argument and answer any critical questions satisfactorily: modifying his tests if necessary. *The justices* will wish to clarify any points that had not been made clear in the original brief, and to pose challenges arising from other briefs.

The *collective goal* of the second dialogue, the respondent presentation, is to obtain a clear statement of the argument for the respondent. The respondent dialogue differs in its *initial state* because the petitioner has already presented. Thus as well as presenting his own argument, *counsel for the respondent* may wish to rebut the argument proposed by the petitioner, and so will have the goal of questioning the petitioner's argument as well as presenting his own argument. *The justices* remain interested in clarification and eliciting answers to questions arising from the other briefs.

While the *collective goal* of the rebuttal dialogue is again a clear statement of the arguments, the *initial state* now also contains the respondent's argument and the individual goal of *the counsel* is to pose questions against this argument. *Justices* usually say very little during this stage, but they may seek clarification of the exact questions being posed.

The goal of the three dialogues together is to provide a clear statement of the arguments for the petitioner and the respondent to provide a basis for the justices to decide the case.

4.2 Speech Acts Used in Oral Hearings

The goal of the dialogues is to establish the various components, and the connections between them, expressed as clearly and unambiguously as possible, which can be used by the justices in constructing the tests that will provide arguments to resolve the case.

The following speech acts will thus need to enable such components to be proposed, and a set of critical questions challenging the components, or seeking additional components to be posed (see [1] for fuller discussion).

- **Values Assertion**: The following values are relevant to decide the legal question. *Law Enforcement and Privacy are the values relevant to determining whether a case falls under the automobile exception.*
- **Issues Assertion**: The values require consideration of these issues. *The issues are whether there was sufficient exigency (so that Law Enforcement is promoted) and insufficient expectations of privacy (so that Privacy is not demoted) to permit a search without a warrant.*
- **Issues Linkage Assertion**: The issues should be considered collectively as follows. *The issues are related as Sufficient Exigency* $\vee$ *Insufficient Privacy.*

We then have a number of moves to introduce factors relating to the issues.

- **Factors for Issue Assertion**: The following factors are relevant to resolving the issue. *Vehicle Configuration and Location are relevant to resolving Sufficient Exigency.*
- **Factor Linkage Assertion**: The factors relevant to the issue should be considered collectively as follows. *Sufficient Exigency is resolved by considering Vehicle Configuration* $\wedge$ *Location.*

We need a number of assertions to identify the facts relevant to the various factors:

- **Facts for Factor Assertion**: The following facts are relevant to determining whether a factor is present. *Wheels and Means of Propulsion are relevant to determining Vehicle Configuration.*
- **Fact linkage Assertion**: The facts relevant to the issue should be considered collectively as follows. *The presence of Vehicle Configuration is determined by considering ((Wheels* $\wedge$ *Engine)* $\vee$ *Self propulsion)* $\vee$ *(Vessel* $\wedge$ *(Motor* $\vee$ *Oars)).*

The structure as a whole is meant to provide a *test*. The test can be challenged at all three levels to question the relevance of the components, the completeness of the asserted components and how the components relate. There is no evidence level, as the facts have been determined by the lower court, but whether such facts are observable by a person applying the test in practice [1] does need to be considered.

In the course of the hearing the various components of the proposed tests emerge. The dialogue is often not well structured: the challenges are not posed in any particular order, and may be interleaved with the presentation of the proposal, so that the proposal is modified as it is presented. None the less, the aim of each counsel is to present and defend the components required for a test which will decide the case for their client, and the Justices aim to get a clear statement of the various components which they can use to build the arguments in their opinions.

4.3 Argument Components Tree (ACT)

We can now organise the argument components identified in the speech acts as an Argument Component Tree (ACT). For each dialogue in the oral hearing we form one ACT

for the counsel and one for the Justices (we do not distinguish individual Justices). Each ACT is constructed starting from the issues. Issues may be *conjunctive* so that all issues must be considered. Or they may be *disjunctive* so that the issues are independent, and one positive will suffice. These are shown in the ACT using "$\wedge$" and "$\vee$" respectively. Sometimes, however, the relationship is not truth functional: like factors, all must be considered, but none is necessary or sufficient (see [5] for a fuller discussion of these relationships). The non-truth functional relation is shown in the ACT using "+".

Throughout the dialogue, the participants' ACTs are updated by the assertion of new factors to resolve issues, or facts that indicate the presence of factors or the linkage between them in order to construct a test. These links may also be truth functional conjunction or disjunction, or reasons that must be considered, essentially the standard factor/abstract relation of [3]. These are shown in the ACT as arrows from children to parent and the factors and facts which attracted the most attention in the dialogue are indicated with an "R".

All the facts mentioned in the oral hearing are *underlined*. Furthermore, the ACT distinguishes several types. Facts which are true of the current case are indicated using an *asterisk (*)*; facts which are not true in the current case but could be used in future tests are indicated by a *question mark (?)*, while an *exclamation mark (!)* is used for facts which could not be used in practice, perhaps because they are not directly observable. By the end of the dialogue, each ACT shows a *complete* representation of a perspective on the components exchanged in the course of the dialogue. The next section provides the ACTs constructed in the case study, *California v Carney*.

4.4 California v Carney: A Case Study

This case is concerned with whether the exception for automobiles to the protection against unreasonable search provided by the Fourth Amendment applies to mobile homes, in particular motor homes in which the living area is an integral part of the vehicle. The Fourth Amendment protects the "right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures." A search is considered reasonable if a warrant has been obtained.

California v Carney has often been used in AI and Law to explore Supreme Court oral argument (e.g. [12], [4]), and to consider the interaction of two competing values (e.g. [9]). In *Carney*, the competing values are enforceability of the law, which makes exigency important, and citizens' rights, which include the right to privacy [7]. In the following sections we provide an example of the construction of the ACTs in the oral hearing dialogues of *Carney*.

Dialogue One - Petitioner Oral Hearing In this dialogue the petitioner states that the exigency is sufficient in *Carney* regardless of any expectations of privacy. This position was based on the *inherent mobility* of the motor home, together with its location offering ready access to the highway, as the factors satisfying the automobile exception, as illustrated in the petitioner ACT of Figure 1.

For the Justices privacy is also an issue that needs consideration. As the vehicle was not actually moving on the road, they suggest tests give a bright line (BL) for

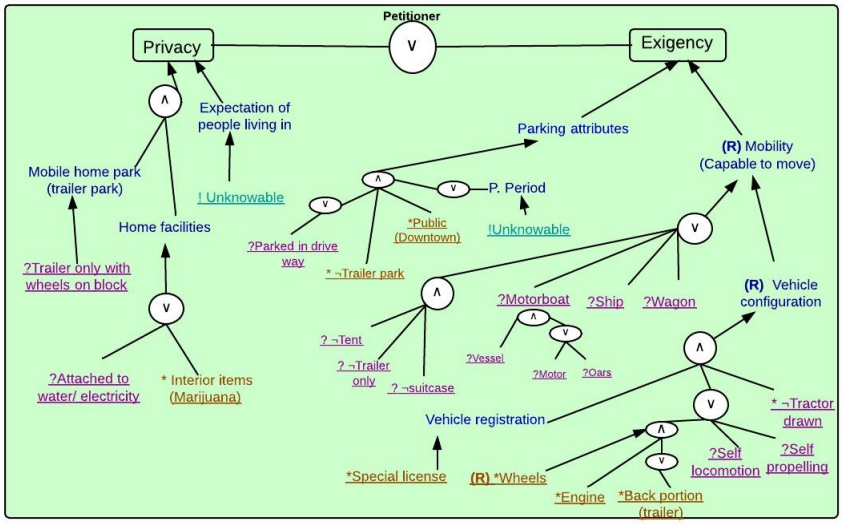


Fig. 1. Petitioner Dialogue - Petitioner ACT

applying the automobile exception to parked vehicles. They propose certain facts to test the privacy degree including the type of the parking lot and attachment to home utilities (water /electricity) relating to the expectations of people living in a mobile home. Furthermore, the Justices challenge the sufficiency of the mobility factor for exigency by considering the vehicle license type (which is different for motor homes and regular cars) and discussing other mobile things such as tents, suitcases, trailers with wheels, houseboats and regular automobiles.

In response to these challenges, the petitioner maintained that exigency is the sole issue and it overrides any expectation of privacy. The petitioner indicates that mobility involves any vehicle, wagon, ship or motorboat but not a mobile item such as a tent, suitcase or trailer, stressing the importance of self-propulsion for the automobile exception. This position thus stresses the significance of the vehicle's configuration and its ability to move quickly on a public highway, which is not true of any of the other mobile objects.

The petitioner accepts the need to consider parking location claiming that if a vehicle is in a residential location (such as a mobile home park) and/or attached to home utilities such as (water/electricity) it *might* not be considered inherently mobile, whereupon issues of privacy would become relevant, but claims that a vehicle in a regular parking lot can always be considered inherently mobile. Figure 1 also presents these components and the relation between them in the petitioner ACT.

Dialogue Two - Respondent Oral Hearing The respondent in contrast insists that *both* exigency and privacy issues need to be considered. The respondent accepts that the exigency is indicated by mobility, but says that this mobility is limited to vehicles that are actually moving on the highway: thus the exigency was insufficient in *Carney* because the mobile home was inoperable (because there was no driver and the curtains were drawn). Moreover, it was parked not far from a courthouse so obtaining a warrant was possible.

Furthermore, the respondent claims that the mobile home attracts sufficient expectations of privacy. He states that such expectations can be indicated through the configuration of the mobile home which involves a living compartment that contains furniture such as bed, refrigerator and other attributes indicating a residence. Moreover a separate class of vehicle known as a *house car* is recognised and defined in the California vehicle regulation code. In addition, the respondent states that privacy interests of a mobile home arise from its use for the storage and transportation of personal effects, and so it should be respected as much as a suitcase, which had previously been held to attract Fourth Amendment protection (see *US v Chadwick*). Figure 2 shows the respondent's ACT.

The Justices defend the petitioner exigency factors, i.e. parking location and vehicle configuration to insist that the vehicle was able to move quickly and thus falls within the automobile exception, giving an example of a crashed car (*Cady v Dombrowski*). The Justices do consider the privacy of home attributes and personal effects, but argue that it is not possible to determine these factors from outside the vehicle, so that no bright line test is given.

Dialogue Three - Petitioner Rebuttal Towards the end of the oral hearing, the petitioner attempts to maintain his position and rebut the elements introduced by the respondent by showing the inapplicability of the tests to prove sufficient privacy.

According to the respondent test above, the fact that the living quarters are an integral part a vehicle should attract sufficient privacy expectations. The petitioner claims, however, that it is not possible to determine the required residential facts, and anyway in *Carney* there was no evidence of food or personal items inside the motor home (except marijuana!) as shown in Figure 1. Moreover, the petitioner states that the definition of "house car" is not used to indicate a dwelling like a house, but to permit the regulation of this type of automobile, as shown by the same definition applied to burglary, aligning house cars with vehicles rather than houses. The new components are used to update the petitioner ACT.

4.5 Relating the Oral Hearing to the Opinions

After the oral hearing, there are four ACTs. These will set out the available facts, factors and issues, and possible linkages between them. The task now is to merge these alternatives to produce an answer for the current case, and a test applicable to future cases. This is the role of the Justices' conference stage, and, given the (competing) ACTs, could be done by proceeding top down, choosing the desired elements, and evaluating the resulting structure using the facts of the case. Thus while all four trees identify privacy and

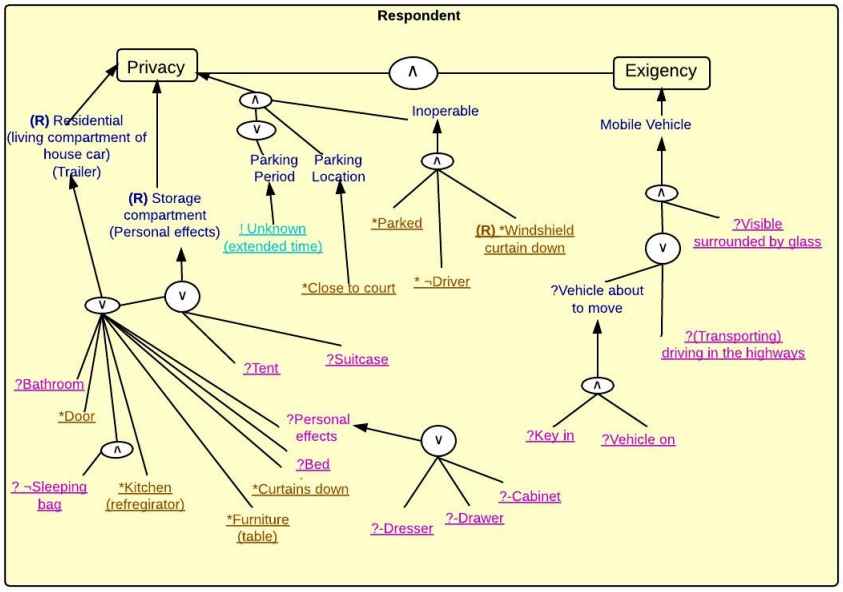


Fig. 2. Respondent Dialogue - Respondent ACT

exigency as issues, all three ways of linking them are available, and must be chosen between. Having identified *exigency* as an issue, a selection from the proposed factors must be made, and so on. Different Justices may make different choices, which may lead Justices to write individual opinions, either dissenting from the majority, or expressing a different view of the appropriate tests. For example, the Justices must decide upon the role that the factor 'parking location' plays. In the petitioner ACT it affects mobility, since the location determines how readily the vehicle can become mobile on the highway, whereas in the respondent ACT it affects privacy by indicating the current use of the vehicle.

From the analysis of *Carney*'s opinion, we find that the opinions offer different navigations through the components that have been presented in the oral hearing ACTs: *all* the components used in the opinions can be found in the ACTs. Some elements form the basis of the court opinion tests. Some of the remaining facts, although not true of *Carney*, are mentioned as potentially pertinent, and so may still provide tests in future cases.

5 Conclusion

In this paper we have provided an overview about the problem concerning the development of a deliberative reasoning framework through the investigation of the dialogue interactions in the Supreme Court oral hearings. In particular, we have focused on the

analyses of the social values of the legal arguments from different perspectives in order to find the relation between the Court opinion and the components constructed in the Court oral hearings. Throughout the paper, we have presented the main aims of the research in defining a proper representation of the Oral hearing, developing a dialogue system for constructing this representation, analysing the opinion by navigating through the tree representation and validating the application of the dialogue using selected cases from various courts. Furthermore, we have presented the reasoning model and show how we moved from the oral hearing transcripts to ACTs, through the use of a set of defined speech acts. Now that we have established a framework for conducting this analysis task, the next step will be to move towards automation.

Acknowledgement Some parts of this work was previously reported in [1] and [2] in collaboration with my supervisors Katie Atkinson and Trevor Bench-Capon, with thanks.

References

1. L. Al-Abdulkarim, K. Atkinson, and T. J. M. Bench-Capon. Dialogues in US supreme court oral hearings. In *Proceedings of CMNA 2013*, 2013. In press, see: http://sinatra.cirsfid.unibo.it/jurix2013.
2. L. Al-Abdulkarim, K. Atkinson, and T. J. M. Bench-Capon. From oral hearing to opinion in the us supreme court. In *Proceedings of JURIX 2013*, 2013. In press, see: http://www.cmna.info/CMNA13.
3. V. Aleven. *Teaching case-based argumentation through a model and examples*. PhD thesis, University of Pittsburgh, 1997.
4. K. D. Ashley. Teaching a process model of legal argument with hypotheticals. *Artif. Intell. Law*, 17(4):321–370, 2009.
5. K. Atkinson and T. J. M. Bench-Capon. Legal case-based reasoning as practical reasoning. *Artif. Intell. Law*, 13(1):93–131, 2005.
6. T. Bench-Capon and G. Sartor. A model of legal reasoning with cases incorporating theories and values. *Artificial Intelligence*, 150(1-2):97–143, 2003.
7. T. J. M. Bench-Capon. Relating values in a series of supreme court decisions. In *Proceedings of JURIX 2011*, pages 13–22. IOS Press, 2011.
8. T. J. M. Bench-Capon, M. Araszkiewicz, K. D. Ashley, K. Atkinson, F. Bex, F. Borges, D. Bourcier, P. Bourgine, J. G. Conrad, E. Francesconi, T. F. Gordon, G. Governatori, J. L. Leidner, D. D. Lewis, R. P. Loui, L. T. McCarty, H. Prakken, F. Schilder, E. Schweighofer, P. Thompson, A. Tyrrell, B. Verheij, D. N. Walton, and A. Z. Wyner. A history of ai and law in 50 papers: 25 years of the international conference on ai and law. *Artif. Intell. Law*, 20(3):215–319, 2012.
9. T. J. M. Bench-Capon and H. Prakken. Using argument schemes for hypothetical reasoning in law. *Artif. Intell. Law*, 18(2):153–174, 2010.
10. S. Brüninghaus and K. D. Ashley. Predicting the outcome of case-based legal arguments. In *ICAIL'03: Proceedings of the 9th International Conference on Artificial Intelligence and Law*, pages 233–242. ACM Press: New York, NY, 2003.
11. T. F. Gordon. The pleadings game: Formalizing procedural justice. In *Proceedings of the Fourth ICAIL*, pages 10–19, 1993.
12. E. L. Rissland. Dimension-based analysis of hypotheticals from supreme court oral argument. In *Proceedings of the Second ICAIL*, pages 111–120, 1989.

An Intelligent Hybrid Approach for Improving Recall in Electronic Discovery

Eniafe Festus Ayetiran

CIRSFID, University of Bologna, Bologna-Italy

eniafe.ayetiran2@unibo.it

Abstract. In this work, we propose a hybrid method for improving recall in electronic discovery proceedings. This approach takes ideas from Natural Language Processing (Word sense disambiguation) and Information Retrieval in enhancing retrieval of responsive documents using the semantics of query terms instead of direct text matching. Preliminary results from disambiguation of user queries show that this approach is promising to improve recall at the same time maintaining high degree of precision in the retrieval of relevant documents to help lawyers and their clients during litigations.

Keywords: eDiscovery, artificial intelligence, information retrieval, natural language processing

1 Introduction and General Background of the Study

There have been studies as early as the 1950s comparing automated methods for classification of documents [3]. eDiscovery is an emerging problem domain that calls for solutions provided from two separate disciplines: Law and Information Systems [3]. The term eDiscovery refers to electronically stored information (ESI) sought by an opposing party during litigation [2], is an important area that poses difficulties for lawyers, litigants and the entire court all alike. Discovering and producing required document(s) among huge volume of data created and stored electronically in various formats in repositories is a big challenge which needs to be addressed. It can be viewed as a form of legal research, which is the process of identifying and retrieving information necessary to support legal decision-making. For many years, lawyers and their clients have relied upon manual and physical methods for retrieving and providing requested documentation during litigations.

At present, the process is commonly carried out mostly through the use traditional technologies such as keyword searching to speed up the process due to advent and subsequent ubiquitous use information systems.

Recently, there have been a lot of research efforts in Machine Learning to improve the present situation. Machine learning is a branch of artificial intelligence which concerns the construction and study of systems that can learn from data and using the knowledge learned on some other new data.

For instance, the 3 emerging Artificial Intelligence techniques for eDiscovery proposed by [5] all of which fell in the line of Machine Learning. These techniques include: (1) Machine learning to extend and apply theories of relevance (2) Generalizing relevance theories with a hypothesis ontology (3) Social network analysis to apply relevance theories

Here we present a proposal which attempts produce a novel approach to eDiscovery by combining techniques from Natural Language Processing and traditional Information Retrieval in overcoming the problems in the existing methods. Natural Language Processing (NLP), a field classified under Artificial Intelligence and Linguistics. NLP enables computers to derive meaning from human or natural language. The idea is to learn from the user queries to improve recall and high degree of precision yet economical.

2 Electronic Discovery

Electronic Data Discovery or eDiscovery is any process (or series of processes) in which electronic data is sought, located, secured, and searched with the intent of using it as evidence in a civil or criminal legal case [10]. eDiscovery, born on April 12, 2006 as a result of the approved amendments to the Federal Rules of Civil Procedure by The United States Supreme Court governing the discovery of electronically stored information (ESI). These amendments took effect on December 1, 2006. It has been the major decisive factor in many cases. According to a report by Socha and Gelbmann [11], the consensus among legal consumers is that 60% of today's cases warrant some form of eDiscovery activity. This percentage will continue to grow over the next several years. Regarding EDD content, according to Corporate Counsel, at least 50% of eDiscovery documents will be in the form of e-mail, with another large chunk coming in the form of office documents (e.g Word, spreadsheets, etc.), together with small databases (e.g MS Access) or larger databases (e.g Oracle), as well as less conventional forms of digitized data (e.g., software code) or other forms (e.g voice mail or video clips) [1]. Today eDiscovery has spread to different parts of the world including Australia, United Kingdom (eDisclosure) and parts of Asia.

2.1 Electronic Discovery and Information Retrieval

eDiscovery is a form of information retrieval. In any Information Retrieval system there is always a trade-off between precision and recall. eDiscovery is a recall-centred task because under production and over production of responsive may have effects on the litigation process as there have been several cases where these situations have been penalized. Although the legal community is familiar with key word search, which historically has been the foundation of case law and statutes searching, standard key word search alone is inadequate for obtaining complete, high recall solutions. There is a wide spectrum of eDiscovery software and service providers today, many that rely on conventional IR techniques, while others harness alternative technologies such as machine learning or concept search along with more standard techniques

2.2 Critical Problems in Existing/ State-of-the-art Approaches and Motivation for Research.

Keyword search, which uses direct text matching between query terms and terms in the document collection, does not provide an intelligent search approach that can cater for the requirements of eDiscovery as the search results includes too many false hits in terms of irrelevant documents. This is because the two foundational issues which arise when searching in an unstructured information domain has not been addressed. The first is the synonym problem – words having the same meaning. The second problem is known as "polysemy," - many words having more than one meaning [9]. Synonyms and polysemies are two factors that reduce the power and accuracy of information retrieval systems. Hence the present generic tools cannot be effectively used to discover relevant documents electronically. Hence, there is need for more intelligent approach.

Machine Learning, an intelligent approach provides a good search that can cater for the requirements of the present day eDiscovery by training a system on a set of data and applying it to new set of data to predict an outcome. One major concern about using Machine Learning is how to get a wide coverage of data enough to cover reasonable level for a problem like eDiscovery may be an almost impossible task knowing the fact testing a Machine Learning system on an entirely different data domain for which it has not been trained may lead to poor results. The big issue is what can be done about this since discovery documents cut across all areas of human endeavour and not limited to a particular domain.

Therefore, we see this as more of a human language problem and propose an intelligent system which learns from the user queries may be a better approach. Computing the actual meaning of each query terms used in context can greatly help improve the overall retrieval process.

2.3 Research Questions

The research questions to be addressed are as follow:

1. How can we conduct an intelligent search and improve recall with only the user query?
2. How do we produce a scalable system to handle large volume of documents usually involved in eDiscovery?
3. How we handle the heterogeneous nature of document formats within the document collection indexing and retrieval?

3 Research Methodology

We present in Figure 3.1 below the proposed architecture of the eDiscovery system.

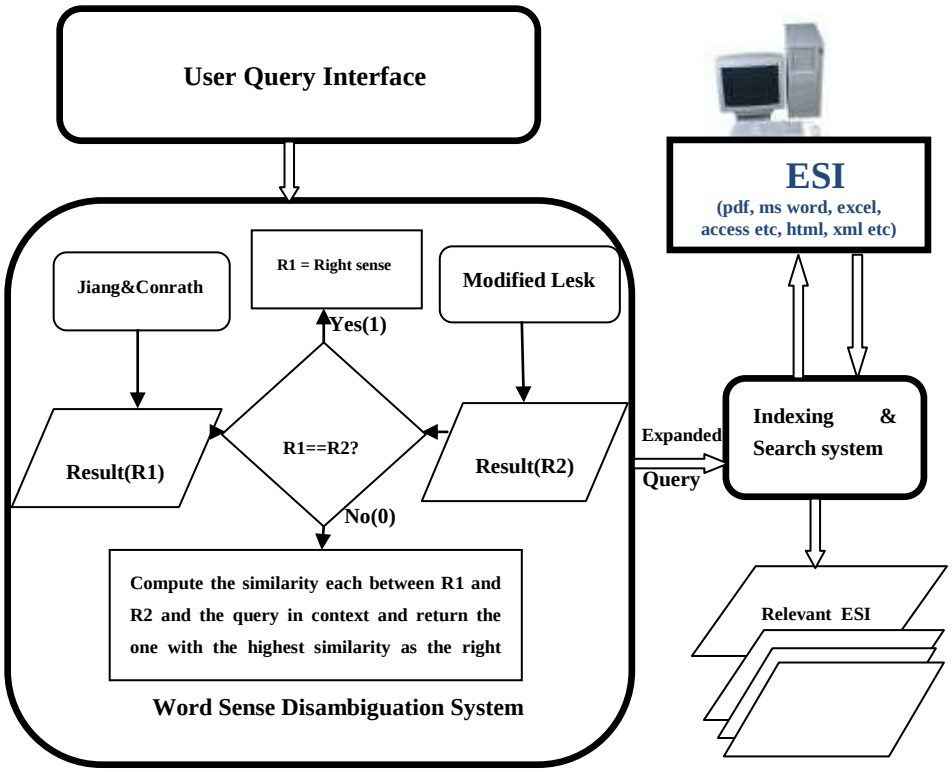


Fig. 1. General Architecture of the Proposed eDiscovery System

Below is an outline of the methodology as depicted in Figure 3.1 above:

a. Sense disambiguation of user query.
b. Expansion of query with semantically related terms to the query terms
c. Development of format-independent indexing and search system using vector space classification
d. Classification and retrieval of responsive documents by the indexing and search system using the expanded query

The whole idea is to compute the meaning of each query terms using word sense disambiguation techniques. The disambiguation will lead to the production of other semantically related words to each of the query terms. The query terms and their semantically related words will serve as input to the indexing and search system which will then classify the documents and subsequently retrieve the responsive documents. The indexing and search system is an accumulation of various technologies that can handle documents of several formats as each format have their own characteristics and tools to index them.

3.1 Word Sense Disambiguation

Ambiguity is a fundamental characteristic of every language of which the English language is not an exception. A considerable number of English words have more than one meaning. The meaning of word intended by a particular user can be inferred considering the context of usage.

For example: (a) *I have a permit to stay in the lodge* (b) *A permit was brought from for dinner preparation*. Based on the context of the usage of the word, *permit,* in the two sentences above, we can infer that the first instance (sentence (a)) is referring to a legal document or an authority to do something and the second instance (sentence (b)) is referring to a large game fish found in the waters of the west Indies. However, human identification of the right word sense is relatively simple compared to machines which need to process large unstructured textual information, carrying out complex computations in order to determine the sense of a word used in a particular context.

The computational identification of meaning of words in context is called Word Sense Disambiguation (WSD) also known as Lexical Disambiguation. Considering the instances in the examples above, the sentences can be sense-tagged as follows: (a) I have the permit/authority/license to stay in the lodge (b) A permit/fish was caught in the Indian Ocean. Basically, the output of any word sense disambiguation system with the right synonymous word (if any). Word Sense Disambiguation relies on knowledge. This means, it uses a knowledge source or knowledge sources to associate the most appropriate senses with words in context. Ideally, Word Sense Disambiguation is a means to an end but not usually the end itself, enhancing others tasks in different fields and application development such as parsing, semantic interpretation,

machine translation, information retrieval and extraction, text mining, and lexical knowledge acquisition. “Polysemy” means to have multiple meanings. It is an intrinsic property of words (in isolation from text), whereas “ambiguity” is a property of text. Whenever there is uncertainty as to the meaning that a speaker or writer intends, there is ambiguity. So, polysemy indicates only potential ambiguity, and context works to remove ambiguity.

In our approach, we have employed a method of inter-technical cross validation of two widely used techniques in the field leveraging on their strengths. These algorithms are the Modified Lesk algorithm – a modified version of the original Lesk algorithm and the Jian & Conrath algorithm. Both algorithms are forms of knowledge-based approach based to WSD.

3.1.1 The original Lesk Algorithm.

A basic knowledge-based approach relies on the calculation of the word overlap between the sense definitions of two or more target words. This approach is named *gloss overlap* or the *Lesk* algorithm after its author [6]. It is one of the first algorithms developed for the semantic disambiguation of all words in unrestricted text. The only resource required by the algorithm is a set of dictionary entries, one for each possible word sense, and knowledge about the immediate context where the sense disambiguation is performed. The idea behind the Lesk algorithm represents the starting seed for today’s corpus-based algorithms. Almost every supervised WSD system relies one way or another on some form of contextual overlap, with the overlap being typically measured between the context of an ambiguous word and contexts specific to various meanings of that word, as learned from previously annotated data.

The main idea behind the original definition of the algorithm is to disambiguate words by finding the overlap among their sense definitions. Namely, given two words, W1 and W2, each with NW1 and NW2 senses defined in a dictionary, for each possible sense pair W1i and W2 j, i = 1…...NW1, j = 1…..NW2, we first determine the overlap of the corresponding definitions by counting the number of words they have in common. Next, the sense pair with the highest overlap is selected, and therefore a sense is assigned to each word in the initial word pair. The Algorithm is summarized in Listing 2.1 below:

1. *for each sense i of* W_1
2. *for each sense j of* W_2
3. *compute Overlap(i,j), the number of words in common between the definitions of sense i and sense j*
4. *find i and j for which Overlap(i,j) is maximized*
5. *assign sense i to W1 and sense j to W2*

Listing 3.1: The Original Lesk Algorithm

3.1.2 Jiang & Conrath Algorithm

Jiang & Conrath propose a combined model that is derived from the edge-based notion by adding the information content as a decision factor. The model is based on the lexical taxonomy of the lexicon and statistics in the information content. In particular, attention is given to the determination of the link strength of an edge that links a parent node to a child node. Jiang and Conrath [4] (Equation 3.1) uses the difference in the information content of the two concepts to indicate their similarity. He used the information content defined by Resnik[8] and augmented it with notion of path length between concepts. This approach includes the information content of the concepts themselves along with the information content of their lowest subsumer.

$$\text{Similarity} = 2 \times \text{IC}(\text{LCS}(C_1, C_2)) - \text{IC}(C_1) + \text{IC}(C_2) \quad (3.1)$$

Where IC is the information content, LCS is the lowest common subsume, C_1 and C_2 are the concepts under consideration

3.1.3 Inter-technical Cross Validation Algorithm

Our technique has been derived from the two algorithms discussed above using WordNet [7] as the knowledge resource. We have modified the original Lesk algorithm adopting WordNet lexical and semantic taxonomy and direct implementation of the Jiang & Conrath algorithm using all the words in context as the window size. In the Modified Lesk implementation, we have not considered the glosses of only the target word and that of their surrounding neighbours, but also that of their semantically related ones in the WordNet taxonomy and these include the hypernyms, hyponyms, meronyms, antonyms etc. We then cross validate the results produced by both Modified Lesk and the Jiang and Conrath algorithms with query terms in context. The main idea is that the glosses of the right sense and that of their semantically related ones in the WordNet hierarchy should be similar as much as possible with the query. The process starts by tokenizing the query with each term in the query as a token and tagging the terms into their part of speech based on the usage in the query. That is for a set of terms, Ti ϵ Qi , where Qi is the query, tag T ϵ Ti into their part of speech based on the usage in the query. For monosemous terms, return the sense accordingly. For polysemous tokens, obtain the synsets from the WordNet with the sense definitions, the lemma names, semantic relations i.e hypernyms, hyponyms, meronyms, etc and examples. We consider the sense definitions of each synset with their associated lemma names, their glosses, glosses of their hypernyms, hyponyms, meronyms etc. We compute the initial score based on the overlap of terms in the gloss of the target word, its hypernyms, hyponyms etc, and that of each of the surrounding words. The overall score for each senses of a term is obtained by summing the all the initial scores with other words in the window size (in this case, all the terms in the sentence). We chose the sense with the highest score as the appropriate sense for the Modified Lesk algorithm.

In the same manner, we compute initial semantic similarity scores for the target word in the query with each of the terms in the query using Jiang & Conrath method. Compute final semantic similarity scores for the target word from the addition of all initial semantic similarity scores. Again, we chose the sense with the highest final semantic similarity score as the appropriate sense for the Jiang & Conrath method.

Finally, we then compare the senses returned by Modified Lesk and Jiang & Conrath algorithms for agreement. We chose the sense for which they agreed as the right sense, otherwise where they disagree, we compute score based on the overlap of their glosses, that of their hypernyms, hyponyms etc with the original query in consideration. The sense with highest score between the two senses is selected as the right sense.

3.2 Vector Space Model

Our aim is to classify documents in the collection or repository into relevant (responsive) and irrelevant (non-responsive) and retrieve the relevant once based on a determined threshold in the weighting and scoring of terms in the expanded query terms and terms in the document collection.

The representation of a set of documents as vectors in a common vector space is known as the vector space model and is fundamental to a host of information retrieval operations ranging from scoring documents on a query, document classification and document clustering. In a typical setting we have a collection of documents each represented by a vector, a free text query represented by a vector, and a positive integer K. We seek the K documents of the collection with the highest vector space scores on the given query.

3.3 Innovation of Research Methodology

Why *Disambiguation, Expanded Query, Indexing and Retrieval Instead of Directly Using Latent Semantic Indexing?*

Latent Semantic Indexing [9] is a method for automatic indexing and retrieval taking into account the issues of synonyms and polysemies. The approach is to take advantage of implicit higher-order structure in the association of terms with documents ("semantic structure") in order to improve the detection of relevant documents on the basis of terms found in queries. The particular technique used is singular-value decomposition (SVD), in which a large term by document matrix is decomposed into a set of ca. 100 orthogonal factors from which the original matrix can be approximated by linear combination. However, the computational cost of the SVD is significant; LSI works best in applications where there is little overlap between queries and documents. Also, it is most suitable where small number of documents are involved.

Hence, it is not suitable for eDiscovery where we have to deal with large volume of data.

Furthermore, the original LSI works with clustering but not with statistical/probabilistic techniques (classification) used for scoring and ranking in information retrieval. eDiscovery is purely a classification rather clustering hence direct implementation of LSI for this type of problem may not be a suitable idea.

Finally, implementing the solution through a method of disambiguation, query expansion, indexing and scoring documents for retrieval brings about the solution to scalability problem while also taking into account the problems of polysemy and synonyms.

4 Preliminary Results and Discussion

We implemented the inter-technical cross validation algorithm and evaluated with the Semeval 2007 coarse-grained English All-words dataset. The result produced 76.516% accuracy (F1 score). The results from this will be used to expand the query which will serve as input to the indexing and retrieval system.

5 Conclusion

With this high performance result of semantic determination of query terms, we believe is a good performance result that will positively enhance the entire retrieval system. In the preceding phase of this research, we hope to effectively adopt the results as an expended query to the indexing and retrieval system using the techniques we discussed previously.

References

1. Counsel C: The American Bar Association (ABA), section of litigation, committee on Corporate Counsel. http://www.abanet.org/litigation/committees/corporate/((2006)
2. D. Oard, B. Hedin, S. Tomlinson, J. R. Baron, "Overview of the TREC 2008 Legal Track." TREC Conference 2008, Proceedings
3. H.S. Hyman, "Designing a Text Mining Artifact for eDiscovery Solutions." Working Paper (2010)
4. Jiang, Jian & David Conrath.: Semantic similarity based on corpus statistics and lexical taxonomy. *Proceedings of the International Conference on Research in Computational Linguistics*, Taipei, Taiwan.(1997)
5. Kelvin. D. Ashley & Will Bridewell: Emerging AI & Law Approaches to Automating Analysis and Retrieval of Electronically Stored Information in Discovery Proceedings. In: ICAIL 2009 Global eDiscovery/eDisclosure Workshop (2009)

6. Lesk, Michael.: Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. In: *Proceedings of the ACM-SIGDOC Conference*, Toronto, Canada, 24–26 (1986)
7. Miller, George.: Wordnet: A lexical database. *Communications of the ACM*, 38(11): 39–41.(1995)
8. Resnik, Philip. 1995. Using information content to evaluate semantic similarity. *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, Montreal, Canada, 448
9. S. Deerwester, S. T. Dumais, G.W. Furnas, T. K. Landauer, R. Harshman, "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science*. (Sep. 1990)
10. SearchFinancialSecurity.com: Definitions: Electronic Discovery. http://searchfinancialsecurity.techtarget.com (2009)
11. Socha G, Gelbmann T: The 2006 Socha-Gelbmann electronic discovery survey report. Socha Consulting LLC and Gelbmann & Associates, MN ((2006)

Online Dispute Resolution for Cloud Computing Services

Dusko Martic

Law Science and Technology Joint Degree EM program,
IDT-Universitat Autonoma Barcelona, Bellaterra, Barcelona
dusko.martic@gmail.com
dusko.martic@unibo.it

Abstract.

The traditional concepts of practicing the law do not follow the pace of the development of the new technologies. Development of cloud computing services last decade raised new issues of applicable law, jurisdiction, access to justice, the legal nature of the disputes, consumer protection etc. At the same time, we are witnessing parallel attempts from several public bodies and international organisation to introduce (on national and global level) the legal framework for the application of the online dispute resolutions. EU parliament has recently voted in favour of the proposal on the ADR and ODR regulations for consumer disputes. These proposals in principle are focusing on e-commerce aspects of dealing with low-cost consumer/seller disputes. This research answers under which circumstances ODR mechanisms are the most suitable means to resolve conflict coming out of provision of cloud computing services in the EU and globally. Building on existing knowledge of ODR, it goes beyond and provide applicable proposals for redress in growing industry of cloud services.

Keywords: Online dispute resolution, ODR, ADR, Cloud computing, cloud services

1 Introduction.

Poles on usage of cloud computing services display constant increase in adoption of this technologies and steady growth of industries providing this kind of services [3][25]. Most of the leading cloud service providers are US-based and even though the nature of cloud computing is to provide service globally, contracts framing these services tend to be in the favor of providers. Cloud contracts usually contain provi-

adfa, p. 1, 2011.

sions stating exclusive jurisdiction of certain US state or specific US courts (where is the corporate seat of the company) and law of the same state as applicable law[2]. It indicates misbalance in negotiation power between cloud providers and users on global level and immaturity of cloud market.

European Union recognized the importance of the cloud technology in its strategic document – "Unleashing the Potential of Cloud Computing in Europe"[4]. The strategy points to the key issues and necessary steps that should be taken in order to remove the barriers and increase economic benefits of cloud computing. As one of the key actions, European commission recognized the importance of safe and fair terms and conditions in cloud contracts. Further elaborating this key action, commission plans to develop model contract for cloud services where it will propose "fair" mechanisms for dispute resolution in case of conflict between provider and user of cloud services [5]. Online dispute resolutions (ODR)[1] are one of the possible mechanisms for redress in case of cloud provider-user disputes. However there are some regulative inconsistencies as we will see in following pages.

In the first part-introduction of the paper we state principal research question. The second part illustrates the problem of redress for cloud computing service disputes from the aspect of contracting. We introduce ODR as potential solution in third chapter, discuss some initial findings in forth part and methodology of research is discussed in fifth. In conclusion we initiate the discussion about research. Although we do not intend to answer all research questions within this paper, as this is much wider research that could not fit in the limits of the paper, we will indicate some preliminary findings to further the discussion on the topic. We will primarily focus on some legal aspects of cloud computing services here and some regulatory responses of the EU. At this point, we will not discuss technical aspects of cloud or ODR as this will come in later phases of research. Even though the research is from a global point of view in this paper we will discuss some preliminary findings related to the EU law, since EU has made some regulatory advancement in ODR field.

1.1 Research question.

The prinicple research question that guides this research is stated as follows:

- Under which circumstances ODR mechanisms are the most suitable means to resolve conflicts coming out of provision of cloud computing services?

The research also aims to answers following set of questions that are subject of particular focus/chapter:

1. What are legal protection issues in cloud environment? What is the current way users seek redress in this types of conflicts? What are the alternatives?

---

[1] ODR as a term has been accepted by practitioners, although many names have been used to describe the same concept: Electronic Dispute Resolution, Online Alternative Dispute Resolution, Internet Dispute Resolution…

2. What kind of legal framework is most appropriate for developing online dispute resolution for cloud services in the EU and globally?
3. What conditions led to the successful ODRs in the past? What are the factors for adoption of these models? Which kind of ODR model has proven to be efficient in comparable services?
4. What ODR supporting technologies are most appropriate for resolving cloud conflicts?
5. Under which conditions ODR could lead to successful online resolution of selected typical issues/use cases for cloud services?

2 Redress for cloud services in contracts.

Cloud computing legal issues illustrate clearly the mismatch between technological advances and the laws regulating society. Certain legal institutions, with long tradition, that were developed over the course of years seemed to be challenged by the technological advancements of last 20 years. Simple fact that in these days it is possible to provide highly specialized on-demand service on global market with low-cost, scalable and easily accessible computing power (for which there is no need high infrastructural investment), changes the markets significantly.

Defining cloud computing is not the easiest task[12]. Cloud computing in simplified terms could be understood as the storing, processing and use of data on remotely located computers accessed over the internet.[4]. More commonly as a starting point authors take broad NIST definition: cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction[17]. In the same NIST document cloud definition is described with five essential characteristics: on-demand self-service (automatic provisioning of computing), broad network access (capabilities are available over a networked infrastructure), resource pooling (resources are pooled together to serve multiple consumers using a multi-tenant model), rapid elasticity (rapid and elastic provisioning of capabilities to quickly scale up or down as required) and measured service (automatic control and optimization of resources utilizing a pay-per-use model)[12]. NIST also recognizes four deployment types (private, public, community and hybrid cloud) and three service models[17]:

- Infrastructure-as-a-Service (access to remote physical or virtual machines model of service) or IaaS
- Platform-as-a-Service (typically including operating system programming language execution environment, database, and web server) or PaaS
- Software-as-a-Service (access to application software and databases) or SaaS

One of more prominent characteristic of cloud services is a shift in payment model to pay-per-use, which compared to similar IT infrastructure investments and software licensing brings significant savings to enterprises and consumers. It also cuts the costs of upgrading needs of hardware and software. Based on previous lowering of prices of some of the biggest cloud providers[24] coupled with influence of Moore's law[2] and Kryder's law[3], we also point to the likelihood of increase in offer of low cost services and high utilization of free model(or freemium[4]) for certain cloud services. The connection to this observation will be explained in fourth chapter.

To obtain cloud computing services users generally accept predefined contract of adhesion, where the terms should be accepted on "as is" basis [2] [20][22]. We have examined contracts offered by 40 big cloud providers (more than 60 cloud services) which indicated to certain regularities in their composition. They usually comprise: Terms of Service (and-or conditions), Acceptable Use Policies, Service Level Agreements and Privacy policies.

Cloud providers - total 40	*Applicable law(in contract)*	*Jurisdiction (contract)*
21	US-California	California courts
8	US-Others	US courts
7	UK and Wales	UK and Wales
5	EU (without UK)	Member state
1	Swiss, Canada	Swiss, Canadian
2	Brazil	Brazil
8	Mandatory Arbitration	AAA rules
4	Possible Arbitration	AAA (+ 1 other)

Table 1. Illustration of survey of cloud providers' contracts[5]

In-depth surveys[2] has been conducted before with similar findings. The negotiations over contracts are more plausible for big corporations and public bodies, while service providers are less prone to offer negotiation for SMEs and consumers [10].

[2] Observation that processing power is doubling every 18 months accredited to Gordon Moore

[3] Observation by Mark Kryder that storage capacity is doubling every 18 months or less, described in http://www.scientificamerican.com/article.cfm?id=kryders-law

[4] Model where provider offers basic services at no cost and charger for upgrades or has alternative way for creating profit out of the usage of the free service[20]

[5] Surveyed cloud providers: Google Cloud(Drive, Docs, Gmail…),Apple Icloud, Evernote, Dropbox, Box, Amazon, Skydrive (Microsoft),Microsoft Azure, SoundCloud, Spotify, Mendeley, CloudON, Zoho, SAP, MicrosoftOffice365, Salesforce, GoogleAppEngine, Coursera, Fuzbox, GoGrid, Rackspace, Joyent, Enomaly, Appistry, Engineyard, ThinkGrid, Opsource, HP cloud, Lunacloud, Nephoscale, Adrive, Mozy/Decho, Softlayer, Symantec, PayPal, Intycascade, Flipboard, Netflix, EDX, Prezi,Trustmarque, Servicemesh

Even though the nature of their service could be global, the terms in contracts are set favoring local jurisdiction and choice of law of provider. In practice, this means that for example SME[6] from Indonesia, using SaaS paying 100 dollars per year, could have a dispute in front of California court and potentially pay approximately ten or twenty times more for fees and expenses, and then dependant on case backlog wait a while for the court deliberation on the issue.

Significant problem in the market is the legal uncertainty when it comes to certain cloud issues (about applicable law and possible enforcement) and lack of appropriate redress in disputes for consumers and SMEs. By appropriate redress we assume redress for smaller claims - fast, low cost dispute resolution and for high level claims - expert deliberation within appropriate time frame. Strategic documents of European Union confirmed this problem [5] and concluded that it leads to the lack of trust in cloud services.

3 The promise of ODR.

Online dispute resolution is a method of resolving disputes using technology as a facilitator or as a "fourth party"[14] in the dispute. While it resembles to be natural extension of ADR[7], since it includes online negotiation, mediation arbitration, ODR has also developed innovative methods using technology such as double blind bidding, visual blind bidding and assisted negotiation. It has proven to be difficult to precisely define the characteristics and types of ODRs, but there is a consensus that we can divide them on adjudicative (i.e. online arbitration, UDRP) and consensual (i.e. mediation, assisted negotiation).

Proponents of ODR claim advantages such as: accessibility, speed of process, asynchronous communication, lower costs, flexibility, etc. However, regardless of corresponding disadvantages (confidentiality issues, higher privacy risks, lack of human "feel "…), after the initial rise of providers of ODR, following the dot-com bubble, the number of active providers has diminished and only a handful selected ODR providers can claim successful practice.[13]

Recently ODR development has entered into the new face with new public support on the horizon. EU has recognized the potential of ODR and chose to connect existing network of ADRs in member states through ODR platform on the EU level[8]. At the same time, UNCITRAL Working party III on ODR is trying to design, global redress

[6] Small and medium enterprises

[7] Alternative Dispute Resolution – all dispute resolution outside of judicial process

system for consumer complaints. Both of initiatives envisioned system for solving high-volume low-value buyer/seller disputes. Even though UN proposal is far from consensus on one model(or two), and the EU model is subject to certain criticism[11], we could claim public bestowing of trust in vision of ODR.

Having in mind advantages and characteristics of ODR relevant authors in the field distinguish ODR for its potential suitability for e-commerce fully-online disputes[19, 23], and consumer protection [7][6]. However, it has not been thoroughly researched from a legal point of view, or successful in practice on a global scale, except from notable cases of EBay, PayPal, Square trade and few other providers of ODR. Even in those cases, e-commerce giants EBay and PayPal have been the providers of ODR and not direct parties of the disputes.

We would argue, having in mind practices of ODR so far, usefulness for e-service disputes and that it had proven itself, especially with parties with equal or similar negotiation power. Nevertheless, serious research needs to be taken of cases where there is a huge discrepancy in negotiation power on the global scale, such as over providing cloud services. Also sometimes failure of one cloud service (of different company) can have cascading effects on other services. The end-user has no relationship with IaaS and his redress is based on his contract with SaaS. This research has this relationship in mind, in order to extrapolate most useful use case scenarios and applications on ODR for cloud services. Even though there are cases where cloud services, engage in arbitration, online or off-line, question remains, is it most appropriate choice of dispute resolution for the other party. To answer the principal research question we need to examine all the positions of parties in dispute and to propose a solution that balances protection of rights and interests of all parties.

The important factor of the solutions could be the costs of process and accessibility. ODR costs also depend of technological developments supporting dispute resolution[15], whether by using agreement technologies[16] that improve the process or having enabling devices widely available[18]. On the table below we illustrate the similarities in prominent characteristics of service domain and instrument for dispute resolution; ODR is perceived much more flexible environment compared to court/ADR procedure.

Cloud computing services	**Online dispute resolution (services)**
On demand	On demand/Asynchronous/Synchronous
Elastic/ Scalable	Flexible/Certain ODR software scalable
Automatic	For certain disputes possible automatic
Pooled resources	Consensual/Flexible adjudication
Measured service/pricing model	Measured dispute/ pricing model

Table 2. Comparison of characteristics of services

The table does not compare services but illustrates similarities in approach of cloud computing services and ODR and change they proposed to previous models. Just the

mere speed of provisioning of cloud services could indicate needs of industry for a faster resolution, and in our opinion it does not sound appealing or appropriate to exchange claims written on paper and send them physically, to resolve disputes for online services that could have a quick life cycle, high volume and sometimes could be highly technical.

Having pointed to some appealing characteristics, we have to point out also that the use of ODR for cloud services has not occurred yet. That fact alone test the assumption within ODR community that disputes that have occurred online should be resolved online[14, 19, 23]. Nevertheless, it is not the technical aspects alone that drive the adoption of ODR, but we have to thoroughly examine legal and other factors that are relevant when we are to decide the course of action after dispute occurred. At this moment cloud providers and users rarely consider ODR as viable option for dispute redress.

We will illustrate this point with one of our observations that came after analysis of recent ODR/ADR EU regulation on consumer disputes that will create ODR platform as an instrument of consumer protection by the end of 2015. EU ADR Directive on consumer protection defines: "service contract means any contract other than a sales contract under which the trader supplies or undertakes to supply a service to the consumer and the consumer *pays or undertakes to pay* the price thereof."[8][21]. This effectively leads to situation that consumers can not send complaints about free online services (which are becoming common) to the EU ODR platform. But, if user pays 1 euro or less he/she will be eligible for online dispute resolution!

This inconsistency in explainable only as oversight of legislator who did not think through the concept of services (or possible evolution) and is opposed to the principles of consumer protection that EU promotes. The need of consumer protection from certain cloud services has been raised before [1] and even EU has undertaken the task of proposing and recommending model law for cloud services with terms that could be considered fair from consumers' point of view[5].

From theoretical point this raises interesting question: are users of free service deprived of legal rights simply because they are not paying for service. The French court of Cassation ruled differently in a recent *Mr. Sebastian R v Facebook* case, claiming that since users are important source of funding (freemium model[9]) and their use of service has economic value, they should be under (certain) consumer legal protection[9].

[8] Italics by author
[9] Italics and comment by author, freemium model described in page 4.

4 Preliminary conclusions and ideas.

Preliminary idea of research is to match deficiency within possibilities in redress caused by misbalance in negotiation power between cloud providers and cloud users, and which manifested in terms of service, with the proposal for online dispute resolution circumventing certain problems with unfair terms and conditions. For certain disputes, as illustrated before, ODR could be filling some gaps that legal system is not paying attention enough. Further discussion about distinguishing right based approach from interest based approach in ODR for these purposes is needed. The research so far indicates that:

- Under certain circumstances ODR could be very effective tool and response for the lack of consumer protection vis-à-vis certain cloud services, as well as effective resolution of B2B disputes while maintaining somewhat the relationship.
- The speed, low costs, access and privacy that ODR offers, resonates both with providers and users of cloud services
- If EU decides to fully extend the use of ODR on cloud services it has to structure the ODR system to provide some incentives for cloud providers (especially big ones) to adhere to the schema
- EU and other ODR provider should rethink de minimis rule for services, to address the growth in providing free online services.
- ODR is part of the solution for the questions of jurisdiction and applicable law. It could be also very helpful tool for assisting judicial processes.
- Sector specific ODR for the disputes of cloud service should be looked into as a form of more competent, specialized forum for cloud disputes.
- For certain disputes over SLAs blind bidding assisted negotiation correspond in the sense of savings in time, costs, human involvement, consensual agreement etc.

We would suggest as preliminary observation that it would be logical to include within the scope of ODR/ADR regulation for consumer protection definition of services including free (online) services, at least where there is considerable economic exploitation of users. In fact, we could say that there are lots of concerns about protection of users in these services, not properly addressed, as opposed to selling goods online or providing more traditional services that have been regulated in some other manner. However we should specify in that case: how could we value these cases from ODR aspect, but also to leave possibility to exclude frivolous claims?

From the point of view of cloud services ODR is offering new, unused, cheaper ways for solving dispute with huge number of users (if needed) while achieving increase in trust, loyalty, feedback on service. Even though the market may not be mature enough for some service and there are not enough competitors for user to have alternative, it does not mean it will always remain this way.

From ODR perspective, the applicability to cloud disputes have not been researched and certainly not considered much for free products or services. Even courts do not consider (or quickly decide upon) the lowest claim based on de minimis rule. ODR tend to be focused on value such as price to select most appropriate tools and mechanisms. But in digital world free services have evolved to a business model and have significant place in users' eyes. Users have duties as well, and ODR tool could be also applied to negotiate or enforce current or existing obligations. We believe that there is potential for innovation especially within the possibility of integration of ODR in cloud service, especially when they are completely software based.

All these aspects as well as opposite end of the spectrum in the form of ODR for high value disputes (online arbitration) will be researched thoroughly.

5 Methodology of research.

This research tends to focus primarily on legal analysis combined with data gathering from selected cloud service providers and ODR providers (that correspond to proposed cloud dispute use cases) and finally we analyze the state-of-the-art of ODR supporting technologies and technologies used in selected ODR providers.

In the first phase of research we introduce legal analysis of the provisioning of cloud services. We examine the private and public laws that shapes cloud computing services. Cloud services are based on contracts usually comprised of four parts: terms of services (ToS), acceptable use policies (AUP), privacy/security policy and service level agreements (SLAs). In order to illustrate the points more clearly we will construct four cloud dispute use cases corresponding to the four usual parts of the contracts. Each dispute illustrates typical problem that could be the cause of dispute. These use cases will serve as binding element for connecting the parts of research into coherent body of work with concrete solutions.

Second phase is dedicated to analysis of international legal framework for online dispute resolutions and cloud service offerings. In the third phase we conduct in-depth semi-structured interviews and analyze data gathered from selected ODR providers (based on previously formulated criteria for providers that offer corresponding or comparable solutions to cloud disputes use cases), in order to thoroughly examine best practices of selected ODR providers. In the fourth phase, we examine state-of-the-art in ODR supporting technologies and cross-examine practices of observed ODR providers. We propose directions for future research based on observed needs of all parties. In final phase based on conclusions from previous chapters we will deduct possible scenarios under which ODR is the most appropriate means to resolve cloud computing disputes.

6 Conclusion.

Cloud computing is a relatively new technology, with high adoption rate and trends that enable even further innovation in ICT. This research sheds a new light on ways we could solve some of the legal issues in cloud computing environments and seeks to find optimal ways to use ODR in cloud services. By constructing use cases of the disputes, which will be connected to every aspect of research, it will give rise to the possibility of practical solutions to certain cloud disputes. Research answers the question under which circumstances ODR could be the most appropriate solution for cloud computing disputes and in that way it could be a starting point of a new research and development of ODR technologies for e-services. This paper illustrates certain oversight by EU legislators who cannot consider all circumstances and situation while designing the dispute resolution system. However, they should rely on independent research to go beyond existing concepts of application.

7 Acknowledgments

I would like to acknowledge the guidance and to express gratitude to my supervisor's prof. Marta Poble of IDT-UAB and prof. Ronald Leenes of TILT, Tilburg University. I would also like to express gratitude for initial, but non-the-less important guidance of professors Monica Palmirani of CIRSFID and Guido Boela of Torino, as well as useful advices of prof. Pompeu Casanovas and prof. Antoni Roig Battala of IDT-UAB. Finally I would like to thank Law, Science and Technology PhD (LAST-JD) program which allowed me to pursue my research.

8 References

1. BEUC-Bureau Européen des Unions des Consommateurs: EU Cloud Computing Strategy-Beuc Position Paper, http://www.beuc.org/BEUCNoFrame/Docs/2/NJBAILHAMLKJKPAJEOPIJMAFPDW69DBYG69DW3571KM/BEUC/docs/DLS/2013-00143-01-E.pdf.
2. Bradshaw, S. et al.: Contracts for Clouds : Comparison and Analysis of the Terms and Conditions of Cloud Computing Services Contracts for Clouds : Comparison and Analysis of the Terms and Conditions of Cloud Computing Services 1. 63, (2010).
3. Cloud Industry Forum: UK Cloud adoption and trends for 2013. 8, (2013).

4. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Unleashing the Potential of Cloud Computing in Europe. COM(2012) 529 final,.
5. Accompanying document to - Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Unleashing the Potential of Cloud Computing in Europe. {COM(2012) 529 final}. SWD(2012) 271 final, (2012).
6. Cortes, P.: Developing Online Dispute Resolution for Consumers in the EU: A Proposal for the Regulation of Accredited Providers. Int. J. Law Inf. Technol. 19, 1, 1–28 (2010).
7. Cortés, P.: Online Dispute Resolution for Consumers in the European Union (Routledge Research in Information Technology and E-Commerce Law). Routledge (2010).
8. Regulation (EU) No 524/2013 of the European Parliament and of the Council of 21May 2013. 1–12 (2013).
9. Cunningham, A.: Caveat Consumer?–Consumer Protection and Cloud Computing–Part 1. Consum. Prot. Cloud Comput. 130, (2013).
10. Hon, W. et al.: Negotiating Cloud Contracts: Looking at Clouds From Both Sides Now. Stan. Tech. L. …. 79–129 (2012).
11. Hörnle, J.: Encouraging Online Dispute Resolution in the EU and Beyond-Keeping Costs Low or Standards High? Queen Mary Sch. Law Leg. Stud. Res. …. 122, (2012).
12. Ivanov, I. et al. eds: Cloud Computing and Services Science. Springer New York, New York, NY (2012).
13. Katsh, E.: ODR : A Look at History. Online Dispute Resolution: Theory and Practice: A Treatise on Technology and Dispute Resolution. pp. 21–33 (2012).
14. Katsh, E.E. et al.: Online Dispute Resolution: Resolving Conflicts in Cyberspace. (2001).
15. Lodder, A., Zeleznikow, J.: Artificial Intelligence and Online Dispute Resolution. 73–94 (2012).
16. Ossowski, S. ed: Agreement Technologies. Springer Netherlands, Dordrecht (2013).
17. Peter Mell, Grance, T.: The NIST Definition of Cloud Computing, Recommendations of the National Institute of Standards and Technology, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.
18. Poblet, M. ed: Mobile Technologies for Conflict Management. Springer Netherlands, Dordrecht (2011).
19. Rule, C.: Online Dispute Resolution For Business: B2B, ECommerce, Consumer, Employment, Insurance, and other Commercial Conflicts. Jossey-Bass (2002).
20. Scruggs, R. et al.: Contracting for Cloud Services. (2011).
21. The European Parliament and the Council of European Union .: Directive 2013/11/EU of the European Parliament and of the Council of 21 May 2013 on Alternative Dispute Resolution for Consumer Disputes and amending Regulation (EC) No 2006/2004 and Directive 2009/22/EC, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:165:0063:0079:EN:PDF.
22. Vulder, K. De, Group, S.: Legal considerations Cloud computing contracts. June, (2010).
23. Wahab, M.S.A. et al.: Online Dispute Resolution: Theory and Practice: A Treatise on Technology and Dispute Resolution. Eleven International Publishing (2012).
24. Cloud Economics: Visualizing AWS Prices over Time | eScience Institute, http://escience.washington.edu/blog/cloud-economics-visualizing-aws-prices-over-time.
25. Gartner Says Worldwide Public Cloud Services Market to Total $131 Billion, http://www.gartner.com/newsroom/id/2352816.

Modeling Legal Documents as Typed Linked Data for Relational Querying

Nada Mimouni

LIPN, Paris 13 University – Sorbonne Paris Cité & CNRS (UMR 7030)
F-93430, Villetaneuse, France
nada.mimouni@lipn.univ-paris13.fr

Abstract. Access to legal knowledge is particularly challenging to information retrieval systems. Not only is legal knowledge usually expressed in linguistically complex forms, but it is also structurally sophisticated (e.g pieces of legislation applicable to a case, version in force of a legal document, other related sources). Modeling the collection of documents in such complex domains requires taking into account the semantic content of the documents as well as their relational structure since documents are usually related to each other by various types of links. In this paper we describe two approaches for modeling and querying a collection of interlinked legal documents. The first approach is based on Formal Concept Analysis and Relational Concept Analysis to model and query the collection of documents. The second approach uses semantic web techniques (RDF, OWL and SPARQL). Different types of relational queries are discussed.

Keywords: Information Retrieval, Linked documents, Relational queries, Formal Concept Analysis, Relational Concept Analysis, Ontology, Semantic web.

1 Introduction

A collection of documents is usually represented as a set of documents. This is a very simplified view since in reality documents are get in a set of intertextual relations that condition their interpretation : a document should not be interpreted solely but with reference to the texts it cites, from which it derives or which derive from it. In the legal domain, documents are linked to each other with amendment, transposition, complementation, jurisprudence relations, etc. These links are not only made for documentary purposes. They also determine the legal validity of documents. For example, in French law, codification is the strong process of structuration of information and the links between documents must be explicitly expressed [1]. Consolidation requires that a legal act makes explicit reference to its successive amendements. Legal information access tools should take into account this complexity of legal material.

XML based standards have been defined to normalize the structure of legal texts, in order to facilitate the access and management of these data. The trend

is to use those standards in the process of legal drafting so as to solve the interoperability issue, which usefulness is obvious. In parallel, open government data initiatives are increasing (e.g. UK Government Linked Data) and many legal information access portals offer querying and search features on this data. However, the data made available are often underused.

Accessing such complex data, characterized by the extra abundance of cross references between legal texts (regulations, laws), requires a querying model integrating both semantic features and intertextual links. Our requirement analysis showed that the need for relational querying is critical from a legal point of view ("find by which texts a given order have been applied?", "what are the local texts that talk about noise that are valid at a given date?", "what are the texts that modify a given text?").

In this work, we propose two approaches which allow representing and querying in a unified manner the semantic content of documents, their structure and their intertextual relations. The proposed approaches are based on Formal and Relational Concept Analysis (FCA, RCA), and on semantic web techniques applied to documentary objects.

The paper first reviews the existing solutions (Sec. 2) and explains the requirements for relational querying (Sec. 3). Sections 4.1 and 4.2 describe the proposed approaches and show how the collections and queries are modeled. Those approaches are finally discussed in Section 5.

2 Related Work

In most specialized domains, documents, such as regulations or laws in the legal domain, must not be interpreted in isolation but in relation with other documents, with which they form "a collection of documents". Legal documents are linked to each other through various types of relations (*e.g.* amendment, transposition, implementation, etc.) and these links often determine their legal validity. We define a collection as a set of documents with semantic descriptors, associated metadata and various types of semantic links between them. Law corresponds thus to a large and highly interconnected network of documents. IR systems should make full use of the afforded richness when processing such complex data, thus exploiting the links, the documents structure as well as their semantic content.

Many efforts have been made to take intertextual links into account in an IR process. Semantic and relational search is handled by both general search engines and specialized legal IR systems in different ways: classic IR on semantic content then navigation with hyperlinks, boolean IR on semantic content then filtering according to links or semantic and intertextual queries.

2.1 Intertextuality in Existing IR Systems

Suppose we have a relational query of the form "what are documents (d') having a given type of link (l) with a document (d) talking about a given subject (s)?". Let's consider how the above types of systems deal with such a query.

- Generalist IR systems such as Google use the most trivial way to deal with intertextuality. The query is treated into two steps: a simple query on the semantic content (s) returns the document (d) and the user can then navigating the hyperlinks according to the type of link (l) to find the set of answers (d'). This category of systems do not allow for relational queries.
- In the second category we classify all systems that allow relational queries via attributes in the query such as XML native databases (queried with XPath, XQuery) and RDF data (queried with SPARQL). The query is treated in a first step as a boolean query on the semantic content (s) to find the set of d, then a filtering step is performed according to the XML elements specified in the query (for XML native repositories) or the set of constraints (in the case of SPARQL queries).
- The third category of systems consists of relational systems such as relational databases and relational concept analysis. Both types of systems allow encoding the references between documents in the model level and also formulating relational queries. The originality of this approach is that the documents collection is structured prior to being queried. For instance, in the case of RCA, a set of conceptual structures (called a relational lattice family) is build upon the semantic content of the documents and the links they hold between them. Then the query is executed against these relational structures to find relevant answers. The advantage of this approach is to allow for navigating within the created lattices to specialize or generalize the query if no exact answer is found.

2.2 Legal IR Systems

Legislative portals or legal access systems (*e.g.* Legifrance[1]) exist in most countries to enable a large and public access to the law. Based on XML standards, they offer rich functionalities such as hyperlink navigation, point-in-time access to historical and repealed documents versions[2], interactive generation of user-oriented up-to-date information[3].

However, so far, legal links between documents have been exploited in a limited way by IR systems. For instance, in Legifrance, explicit links are mainly dealt with manually. Some of them are included in the content of the data base (hyperlinks) and others are implemented as document attributes when the data base content is managed[4]. The UK Legislation site allows to search for changes made in the legislation since 2002[5]. The user can query the database either by specifying the modified legislation or the legal source that introduces the change. Whereas the system treats the general link "modifies/modified-by" as a relation

[1] www.legifrance.fr
[2] *e.g.* UK legislation (http://www.legislation.gov.uk/search/point-in-time)
[3] *e.g.* New South Wales legislation website (http://www.legislation.nsw.gov.au).
[4] Force (V), With force term (VT), Delayed effect (VD), Repealed (Ab), Canceled (A), Disjoint (D), Modified (M), Implied repeal (P), Substituted (S), Transferred (T).
[5] http://www.legislation.gov.uk/changes

between documents, more specific types of modifications are represented as document attributes. The Italian website Normattiva enables point-in-time access to legislation too[6], allowing thus to retrieve versions of a document in force at different dates.

The analyzed systems do not exploit explicit links between legal documents to their full potential. In order to illustrate this point we can think of a *continuum* from less to more operational representations of links in legal IR systems:

- Links are represented as strings in the text of the document: usually they appear in the final part of the document and are added manually (by an editorial team).
- Hyperlinks between documents: links are references that point to objects in the collection (other legal documents or fragments of those documents).
- Links are queryable as attributes: legal relations between documents are represented as attributes of the linked documents.
- Relational query: links are modeled as relations between documents in the collection. This allow for relational querying.

If we compare this continuum to the categories presented in the section 2.1, we notice that systems of the first and the second items belong to the first category, the third item systems belong to the second category and the last one corresponds to the third category. Our goal is to exploit the further end of this continuum, namely, the representation of *various types* of legal links as relations between documents in the collection. It is our assumption that such representation mirrors more precisely the way legal professionals conceive the network of legal provisions and will thus enable a more natural interaction between the user and the system.

3 Requirements for Relational Querying

Legal expert common queries show that it is important to distinguish and exploit different types of inter-document links. The query may deal with the case of application of a law text (for example: "find all application cases of a given order"), a validity date (for example : "which local texts deal with noise and are valid in a given date?") or modification links (for example: "which are the texts that modify another text?"). Table 1 give more examples of relational queries. To overcome such limitations, legal IR system should deal with the rich typology of relations linking the documents of a legal collection in order to enable relational querying.

4 Proposed Approaches

To meet these requirements, we propose two different approches for relational modeling and querying. The proposed approaches allow answering simple and

[6] http://www.normattiva.it/ricerca/avanzata/vigente

What conventions implement the recommendations that talk about termination ?
Which recommandation about benzene are implemented by conventions on occupational Cancer ?
Does a law text has been applied? and in which cases (give examples of case law) ?
What recommandations are implemented by conventions on air pollution ?
Given an order, what are the legal texts that it develops ?

Table 1. A sample of relational queries expressed by legal experts

relational queries on a collection of linked legal documents. This work is part of the LEGILOCAL project [7]. The collections of legal documents we are dealing with are characterized by:

- Different types of documents (laws, codes, editorial documents, etc.).
- A specific internal structure for each document type (sections, paragraphs, etc.).
- Various types of links between the different types of documents.
- Semantic descriptors annotating the documents w.r.t a semantic resource.

The first approach [2], based on FCA and RCA, creates classes of documents using their semantic contents and the links between them. Despite its consistency from a formal point of view, a major limit of this solution is the size of the created conceptual structures when applied to a big collection of documents. To tackle this problem, we proposed a second solution [3, 4], based on semantic web techniques (RDF, OWL, SPARQL), which is scalable and nevertheless adresses the problem of relational querying.

4.1 Conceptual Classification based on FCA and RCA

Figure 1 gives an overview of our approach, composed of four main steps:

Semantic content modeling: the semantic content of the documents is annotated and binary contexts are extracted based on those annotations allowing formal concept lattices to be build.

Intertextual structure modeling: the links between documents are identified and relational contexts are extracted based on those links allowing enriched relational lattices to be build.

Relational querying: the user creates a query, possibly as a combination of semantic descriptors and cross-references constraints.

Search and results: the search algorithm analyses the query and looks for relevant answers on the lattices. The user can get traditional list or graphs of result documents. Alternatively, he can directly visualize results in the lattice structure which can be further explored to get approximated results.

[7] LEGILOCAL is an FUI project 2010-13. See http://www.mondeca.com/fr/R-D/Projets/LegiLocal-Projet-FUI-9-Cap-digital-2010-2013.

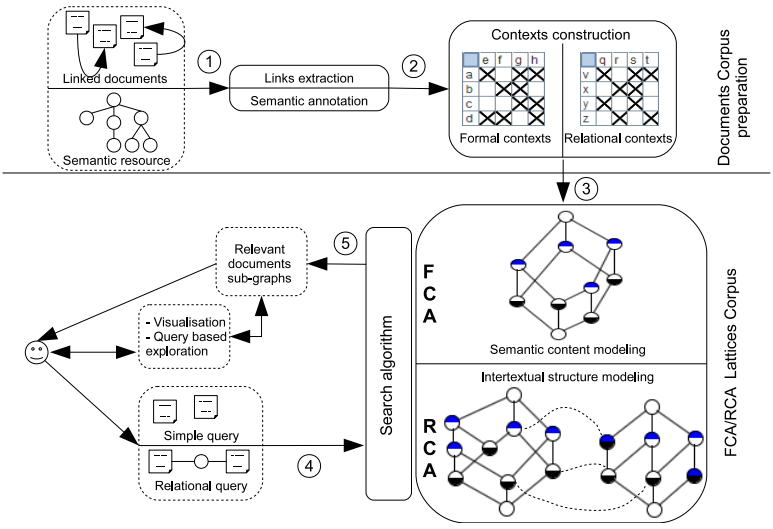


Fig. 1. Overview of the relational IR approach

The semantic content of documents is first modeled as a formal context which describes a binary relation between a set of objects and a set of attributes *(object x attributes)*. The objects correspond to documents. The attributes correspond to the semantic descriptors characterizing the content of these documents. In an information retrieval (IR) perspective, the lattice built by the FCA on binary contexts gathers all possible combinations of documents attributes. These combinations are represented by the intentions of concepts having as extensions all the documents sharing these properties. To answer a query, the search algorithm identifies the class of documents sharing the maximum number of attributes with the query.

We use RCA, the relational extension of FCA, to take into account the cross-references dimension in the modeling of the collection. The approach builds a family of relational contexts, from binary contexts *(documents* $\times$ *semantic descriptors)* and a relation represented separately in a new context defining a type of relation between documents *(documents* $\times$ *documents)*. This family of contexts forms the starting point for the creation of corresponding conceptual structures called Relational Lattice Family [5]. RCA is able to take into consideration different types of links, which are represented by different relational contexts.

Simple Queries We call "simple queries" the queries that are expressed as a set of semantic descriptors. For example *"Which orders talk about abnormally annoying noise (bag) and sound disturbance (ns)?"*. The key words "abnormally annoying noise" and "sound disturbance" are considered as semantic descriptors

annotating the documents which type is "orders". The initial lattice built with FCA represents the set of all the simple queries based on semantic descriptors combination which are satisfiable, i.e. return orders (all descriptors combinations associated to a non null extension). If the query corresponds to the intension of a concept having an extension, documents of this extension are returned as an answer to the query. If the query corresponds to an intension without a proper extension, we can propose a specialization or generalization of the query: this is the advantage of the proposed approach of relational information retrieval.

Relational Queries Our model allows also to answer relational queries. Relational queries contain not only a set of semantic descriptors but also relational indicators between documents. The relational indicators express one or different types of cross references between one or more types of documents. For example *"Which orders talk about abnormally annoying noise (bag) and make reference to decrees talking about soundproofing (ip)?"*. The key words "noise" and "acoustic pollution" are considered as the semantic descriptors annotating respectively the documents which are of type "orders" and "decrees". Different types of relational queries can be handled: legal text to legal text relational query, legal text to semantic category relational query and semantic category to semantic category relational query. Answers to these types of queries are graphs of linked documents.

4.2 An Ontology of Legal Documents Collection

In the second approach we propose an ontology based document model to support the sharing of documents of French local administrations. This ontology has been designed on the basis of Legilocal requirements analysis. It allows to represent all information on legal documents: 1) the structure of a document (sections, paragraphs, etc.), 2) the time frame in which it is registered, 3) the semantic description of its content using concepts or entities in the considered domain, 4) its type (law, decree, etc.) and 5) its relationships with other documents (modification, repeal, transposition, etc.). Our document ontology is structured into three main modules which allow to model the above properties : document module (properties 1 and 2), the semantic module (property 3) and the collection module (properties 4 and 5). Details of each module are given in the following.

Types and Structures of Documents In the Metalex ontology, resources are typed according to the FRBR convention as work, expression, manifestation and item. In our model, we focus on the two upper levels, namely work and expression, in order to represent the different versions of articles and documents. Moreover, those documents have different types (French legislation, court decisions, local acts as well as editorial documents). These various documents have different structures and are characterized by different metadata. Indeed, to prepare a municipal act on a particular subject, local administrators have to investigate national legislation and case law on the same subject. In order to

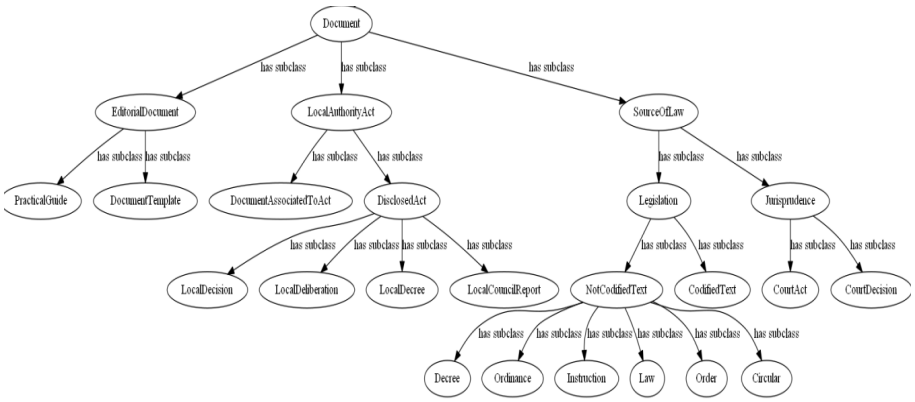


Fig. 2. Hierarchy of documents types

help them, our approach aims to provide semantic search in national legislation and case law, as well as in local acts of other municipalities on the same subject and even some editorial documents. These semantic search facilities require that the documents be annotated with both topics and interdependencies.

As Metalex ontology was firstly designed to model legislation, we extend it with a document typology (see figure 2) that enables us to describe specific properties for each type of documents. For example, we want to be able to specify the structures for certain local acts in order to check their conformance, and some related properties such as the local organization and the person in charge of the document which are specific for each local act. We propose a fine-grained description for legislation text in which the basic unit is the article (which has an independent life-cycle, and could be cited and returned as answer to a user query). On the contrary, for local acts, we do not go through fine-grained description and keep coarse decomposition.

Documents Relationships We want to answer queries such as: "*What are the judgments that implement articles 4 and 5 of the law on minor work?*" or "*Which amendments are made to the article7 of law1955?*"). To reach this goal we propose to model a collection of documents as a semantic network based on a fine description of the types of citations. Our reference model differs from the Metalex one in two respects. On the one hand, we refine the generic reference notion. A broad distinction opposes the citations that refer to a textual object and the semantic annotations that refer to non-textual objects, but we also introduce various semantic citation subtypes (see Fig. 3). On the other hand, we simplify the event-based model of Metalex by encoding references as direct relations, that are directly exploitable for search and visualization purposes.

In our model, each type of reference property is associated to specific domain and range, which allows to specify not only to which types and parts of texts (for

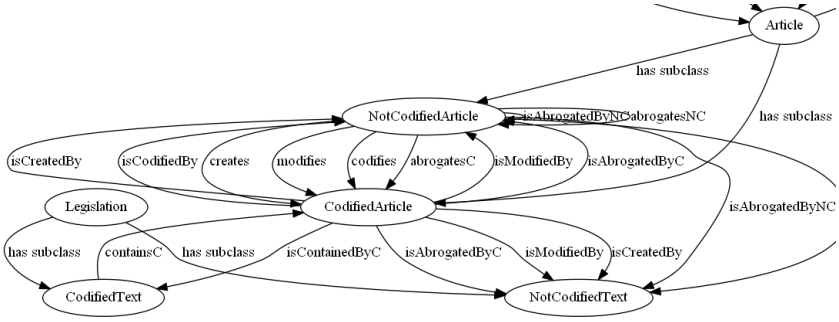


Fig. 3. Various types of links between different types of documents

citations) or semantic classes (for annotations) it refers but also in which types of texts and parts of texts it may appear. Actually, we introduce an opposition between document fragments and units to distinguish the document parts that are citable (units or `CitableBibliographicObject` in Metalex ontology) from those that are not (mere fragments). For instance, we consider whole documents and articles as units but not the preamble of a law. The same opposition holds for the search results: only graphs of document units can be returned to answer a relational query. On the contrary, semantic annotations can be attached to any fragment of text.

Semantic Annotations In this work, the term "semantic annotation" denotes the references that are not citations. We define semantic annotations as references referring to ontological entities that do not represent documents or parts of documents.

5 Discussion

In this work we have introduced a solution to the problem of the complexity of legal sources. Using semantic content descriptors, documents typology and cross references between documents, we have introduced two approaches to model and search within a collection of interlinked legal documents. This allows to answer relational queries and return graphs of linked documents. The first approach is based on FCA and RCA to model the collection as conceptual structures. We have experimented relational queries to explore and query this relational model and return relevant documents or graphs of documents. A more operational solution based on semantic technologies (RDF modeling and SPARQL querying) is introduced as a second approach. We propose an ontology-based model to tackle the complexity of legal sources and to model collections of interlinked legal documents. Beyond traditional legal search, those models already support fine-grained semantic and relational IR functionalities.

Adopting an integrated document model to encode the structure of the documents, their semantic annotations and the semantic structure of the collection enables to process complex queries combining structural, intertextual and content search criteria. For instance, if a local administrator wants to find examples of local acts dealing with "rural roads" and based on a particular decree d, he can express a query combining constraints on semantic annotation (`refers to` the class `chemin rural`) and document references (`cites` the decree d). Our future research will include conceiving user friendly interfaces, allowing to easily create a relational query based on the collection characteristics (semantic descriptors, documents types, references), and also to display results returned as graphs of documents.

References

1. Bourcier, D., Mazzega, P.: Codification, law article and graphs. In: Lodder, A., (eds.), L.M. (eds.) Legal Knowledge and Information Systems, JURIX. pp. 29–38. IOS Press (2007)
2. Mimouni, N., Fernàndez, M., Nazarenko, A., Bourcier, D., Salotti, S.: A relational approach for information retrieval on xml legal sources. In: Francesconi, E., Verheij, B. (eds.) ICAIL. pp. 212–216. ACM (2013)
3. Mimouni, N., Nazarenko, A., Salotti, S.: Une ontologie documentaire pour la recherche d'information relationnelle. In: Actes du 5ème Atelier Recherche d'Information SEmantique (IC-RISE 2013). Lille, France (Juillet 2013)
4. Mimouni, N., Salotti, S., Paul, E.: Modeling collections of french local administration documents. In: Proceedings of Jurix 2013 (accepted). Bologna, Italie (December 2013)
5. Rouane, M.H., Huchard, M., Napoli, A., Valtchev, P.: A proposal for combining formal concept analysis and description logics for mining relational data. In: ICFCA, pp. 51–65. Springer (2007)

Legal Knowledge Framework for Identifying Water, Energy, Food and Climate Nexus

Md Mizanur RAHMAN[1]
Erasmus Mundus Doctoral Fellow, LAST-JD program, CIRSFID, University of Bologna

Abstract. The paper inclines to briefly articulate major problems and critical questions exist in water, energy, food and climate (WEFC) nexus from artificial intelligent and law perspective. Then a legal knowledge framework, based on computational ontology, Akoma Ntoso and LegalRuleML standard, is proposed to identify WEFC nexus. It also presents a brief use case on the existing legislations of quality standard of drinking water from EU and UK that covers WEFC nexus, where the proposed framework will be used. At the end, it states briefly fundamental methodologies for the proposed framework and their strengths, related tools for environmental decision support systems and their limitations, and other related works like "Fill the Gap" project in order to rationalize the degree of innovativeness and necessity of this proposed legal knowledge framework for identifying WEFC nexus.

Keywords. Legal knowledge framework, WEFC nexus, artificial intelligent and law, Akoma Ntoso, LegalRuleML.

Introduction

Traditionally water, energy and food regulations are managed in separate legislative branches due to their sectorial approach [1]. Therefore it is not easy to detect the implications that a legal textual provision of one domain could have over the others. Additionally the collective approach of mutli-sectorial legal rules of WEFC nexus is often neglected in the public policy analysis, particularly in the case of favoring technical requirements (e.g. soil characteristics, energy plant requirements) of one domain to others [2][3]. Similarly when a policy of one domain is adopted and implemented, it is also difficult to maintain aligned policies of same or other domains within the legislative system in order to not create paradoxical situations in other legislative areas (e.g. taxation policy) that could be against the WEFC nexus's approach [3][4]. Besides, when once it is possible to detect WEFC nexus and make evident the relationships, the next further difficulty is to resolve the conflicting rules that exist within WEFC domain in order to decide the best policy to adopt. For these reasons, a legal knowledge framework might be useful to understand and manage WEFC nexus in a better way as well as to simulate multi-sectorial scenarios of WEFC domain, these all scenarios are equally legally valid in scope and nature but depending to a specific expert interpretation or operative implementation.

Considering this context, the paper presents a comprehensive legal knowledge framework for detecting WEFC nexus based on the possibility to use original legal texts and to formalize the legal knowledge of WEFC domain for permitting legal reasoning among different rules (normative, social, technical, ethical, cultural) with the help of the legal knowledge engineer.

In definitional point of view, legal knowledge framework encompasses the scope for utilizing legal knowledge formalization by implementing three indispensable chronological technological requirements: (a) to systematically documentize the content of legislations in such a way that makes machine to understand the process, e.g. implementing Akoma Ntoso standard [5], (b) to use computational ontology [6], that is legal, social, ethical and scientific-information based ontology, in order to make machine to understand meaning of the prescribed content or

[1] Corresponding Author: Erasmus Mundus LAST-JD program, CIRSFID, University of Bologna, Via Galliera, 3 IT – 40121 Bologna, Italy; E-mail: mdmizanur.rahman2@unibo.it.

document, and (c) to use operational-rules-based-logic, e.g. implementing LegalRuleML standard [7], expecting machine to apply legal logic for enhancing evidence-based-hybrid-reasoning.

The fundamental aspiration beyond the legal knowledge framework is that, from the rule of law perspective, nothing is above the rule of law [8]. Everything that happens within the WEFC domain must be comply with the rules prescribed in some forms of legal instruments. In contrast, if there is any legal structural constraints within WEFC nexus, it must be first detected and resolved before further proceedings in order to make system effective and efficient with sustainable efficacy within WEFC domain by avoiding consequential loops of disorder. Therefore under WEFC nexus discourse it is very crucial to detect how existing rules of each domain, e.g. water or carbon tax, affect and reinforce other domains.

Moreover, the paper uses a brief use case on "Quality Standard of Drinking Water of EU and UK" and how proposed legal knowledge framework will be useful to detect WEFC nexus within that problematic context.

1. Brief State of Art

During last five years, responsive citizens, researchers, environmentalists, lawyers and policy makers have demonstrated a great level of importance to understand dynamics and complex relationships between WEFC nexus [1] [9]. On December 10, 2012, the National Intelligence Council (NIC) of USA mentioned in its report "Global Trends 2030: Alternative Worlds" that growing food, water and energy nexus is one out of four megatrends in coming transformative world, which will be responsible for major power shifts, human insecurity and geopolitical risks [10]. European report on Development 2012 recommended that a radical transformation is needed to cope with nexus's requirements [11]. As a result, EU has already taken a number of initiatives in order to support activities related with nexus [12].

Rules of these three sectors, water, energy and food, interact and reinforce each other. Water, for example, is used for fuel extraction, refining and production. It also generates electricity and cools power plants. Scarcity of water affects food processing, generating electricity, crop and livestock yields. The overuse of water also affects negatively quality of crop, soil and other elements of social and environmental interaction. Energy is required for water transfer and treatment. Food prices increase as fuel, fertilizer and transportation costs rise [13]. In order to understand better, however, the concept of nexus, following statistics might be helpful:

- Global population has been increasing by some 80 million in a year. By 2030, it is expected that the total global population would be 9 billion. That will need 30% more water, 40% more energy and 50% more food in order to survive [14].
- To manage the demand of global drinking water, energy and food, the human community will face 40% water gap by 2030 [15][16].
- 1.1 billion people have lack of clean drinking water, 1.3 billion are living without electricity, and more than one million are hungry [17].
- 20% of world's electricity comes from hydroelectric power, e.g. 99% in Norway and 50% in developing countries. 70% of global freshwater withdrawals accounts for agriculture [18]. 32 million to 54 million barrels of oil was used to generate energy to produce amount of bottled water consumed in US in 2007 [19].
- The number of middle class consumer will increase from 1.8 billion to 3 billion by 2030 [20]. That will reinforce water, energy and food market.
- The water is required for producing food is 70% times greater than the water is used for domestic uses like drinking, bathing and washing [21].

However, it is very little that policy makers or citizens know about this nexus's complexities, specially how rules and/or legislations of each domain affect and reinforce other domains [22][23][24]. It is also noteworthy to mention that in the current state of art of WEFC nexus there is no such constructive case yet has been developed in line with examining utility and efficacy of the legislation of one sector of WEFC over the others and using artificial intelligent and law.

2. Major Critical Problems and Questions in WEFC Nexus

2.1. Major Critical Problems of WEFC Nexus

Traditional way of formulating policy documents, legislative, administrative and institutional rules within WEFC domains, as a resource of policy, are still segregated and sectorial wise [1] [3]. The substantive and institutional rules of respective sector of WEFC that help to install the political administrative programing and arrangement for implementing public policies are isolated too [3], as in Figure 1. However, in the context of WEFC nexus, from artificial intelligent and law perspective, the following major problems have been detected in order to framing a legal knowledge framework:

2.1.1. Lack of Detection Mechanism for Revealing Legal Textual Implications of One Domain over the Others

The implication of legal textual provision of one domain of WEFC nexus, generally, affects and reinforces other domains, which is not easy to detect [25].

2.1.2. Difficulties to Maintain the Rules of Policies and/or Legislation of One Domain with Aligned Polices and/or Legislation of Other Domains

There are always a number of rules of one policy documents and legislations have legal relationship with other aligned policy documents and legislations. But generally these rules are subject of different institutions to implement and to prepare necessary financial allocations. That plays an important role to make difficulties in the process of detecting WEFC nexus. So, in order to simplify the process, it is essential to maintain linked-rules of different policy documents and legislation [26].

2.1.3. Lack of Mechanism for Integrating between "Related Institutional Rules" and "Rules Coming from Policy-based Legislations" of WEFC Domain

The rules of game of WEFC domains are not only determined by policy documents and legislations, rather institutional rules play an important role in determining ideas, interests, process, content and what need to be done at the ground time to time of WEFC domains [27]. Hence making functional links between institutional rules and rules coming from policy documents and legislations may help to coordinate legal knowledge of WEFC domains more efficiently.

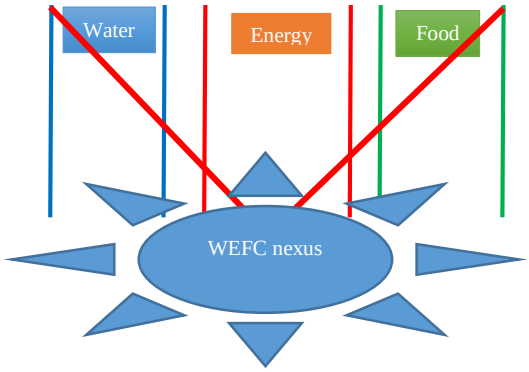


Figure 1. WEFC nexus in political administrative arrangement

2.1.4. Lack of Mechanism for Cross Compliance Check among Rules (Legal, Institutional, Social, Cultural, Ethical and Technical) of WEFC Domain

Cross compliance check between and within policy documents and legislations of WEFC domains is not enough. Besides, it is also very crucial to have integrated mechanism for cross compliance checking among rules coming from different sources such as legal, institutional, social, cultural, ethical and technical perspectives of WEFC domains [28]. These two types of cross compliance checking jointly are very requisite not only for detecting the WEFC nexus as well as for legal reasoning in favor of WEFC domain.

2.1.5. Unresolved Conflicting Rules within WEFC Nexus

Detecting of conflicting rules within WEFC nexus adds another degree of efficiency in order to adopt most appropriate set of rules for WEFC domain [29].

2.1.6. Lack of Collective Approach based on Multi-sectorial-Linked-Legal-Rules of WEFC Nexus

Legal rules of one particular domain of WEFC might create goal-conflicts to the legal rules of the others. Hence it is necessary to have collective approach based on multi-sectorial-linked-legal-rules for detecting WEFC nexus in a synchronized way [30].

2.1.7. Absence of Implication of Technical Rules of One Domain over the Others and Detecting Contradicting Technical Rules

Generally, but not always, technical rules are guided by legal rules within a specific domain of WEFC. Hence traditional way of applying technical rules is limited to the respective domain. But, in order to detect WEFC nexus cautiously, technical rules of one domain must need to apply to the others in its appropriate scope and context. However, in the case of not having appropriate technical rules within the policy documents and legislations of one specific domain of WEFC, it is necessary to include technical rules from scientific investigations [31]. Moreover, detecting of contradicting technical rules is too essential for detecting WEFC nexus in most appropriate way.

2.1.8. Absence of Standardized and Systematized Documentation of Contents and Rules (Legal, Institutional, Social, Ethical and Technical) of WEFC Domain

Policy documents, legislations, authoritative reports and other legal documents of WEFC domains are not systematically documentized in according to any international standard such as Akoma Ntoso. Hence it is very difficult to process mechanically the contents and rules of WEFC domains, which can be considered as a fundamental obstacle for detecting WEFC nexus automatically.

2.1.9. Lack of Formalization of Legal Knowledge of WEFC Domain Using Computational Ontology and Standardized Legal Reasoning Approach

Once standardized and systematized documentation of contents and rules of WEFC domains are processed, it is required to formalize legal knowledge of WEFC domains using computational ontology and standardized legal reasoning approach in order to detect the WEFC nexus spontaneously in real time application and with legal reasoning for legitimizing detection of WEFC nexus.

2.1.10. Lack of Legal Knowledge Network for detecting WEFC Nexus

Existing networks of WEFC nexus's initiatives are neither based on Akoma Ntoso and LegalRuleML standard nor use computational ontology. These networks are merely preserving information in pdf or html format and also not independent from technology, language, machines and platform. Most importantly these networks are not designed for formalizing legal knowledge of WEFC domains. Therefore, usages of these networks are very limited.

2.1.11. Lack of Rule-based Simulation of Multi-Sectorial Scenarios of WEFC Nexus

Existing simulation techniques for Environmental Decision Support Systems (EDSS) are mainly based on mathematical models [32], but the rules of the game for WEFC domains are based on mostly legal and institutional rules including other relevant rules such as social, cultural, ethical and technical. Therefore, in order to simulate WEFC nexus pragmatically with legal reasoning, it is required to simulate based on all available exiting and legally valid rules.

2.1.12. Lack of Change Management within WEFC Nexus

Existing WEFC domains, on the one hand, are mainly closed and non-adaptive in nature towards the changes transported by new rules coming from new legislation, institutional, social, cultural, ethical and technical requirements. On the other, there is no the best solution for detecting the WEFC nexus, but the most appropriate one based on available open-linked-data. Because rules of game for WEFC domains get changes over time to time.

Therefore, flexibility and adaptability must be ensured towards the new rules in order to update the detection of WEFC nexus.

2.2. Major Critical Questions of WEFC Nexus

Two major questions is expected to resolve are: (a) How is it possible to use artificial intelligent and law using Akoma Ntoso and LegalRuleML standards, and computational ontology for performing a legal knowledge framework for detecting WEFC nexus? And (b) what functionalities or systems or sub-systems should be designed in order to resolve following major problems?

3. Legal Knowledge Framework for Identifying WEFC nexus: Main Pillars and Features

The proposed legal knowledge framework for identifying WEFC nexus is based on three main pillars:

- *Akoma Ntoso standard*, that is a machine readable and technology neutral XML standard for digital representation of substantive and institutional regulations and policy based legislation and documents. It is for systematizing documentation of related legal documents of WEFC domains.
- on a *Computational Ontology*, that is to represent the main concepts and relationships of the WEFC domains, and
- *LegalRuleML standard*, that is for modeling rules for formalizing legal knowledge related with WEFC domains using logic-based theory of legal and evidence-based hybrid reasoning. It is also intended to use for legitimizing the identifications of WEFC nexus by proving legal reasoning.

This framework is also intended to provide the following features:

- *A Knowledge network*, that is for connecting legal texts relevant in WEFC domains aiming to create a knowledge network that could help the legislator and policy makers to maintain updated legal knowledge of WEFC domains over time in a coordinated way;
- *Identification of WEFC nexus*, that is by modelling rules for detecting WEFC nexus not immediately explicit;
- *Evidence based Hybrid Reasoning* [33] that is for using non-monotonic logic reasoning (defeasible logic) in order to manage the conflicts among the above mentioned rules and to provide different scenarios where the decision maker and the policy maker could use for evaluating the impacts on the WEFC.

4. Use Cases on Legislations of Quality Standard of Drinking Water in EU and UK and identification of its WEC nexus

4.1 Brief Background Information

UN Water's statistics inform that the fresh and drinkable water is only 3% of total world's water. Out of which, over 2.5% is frozen and not available to human being and rest .5%, equivalent to 200,000 square km, is for the survival of humanity [34]. Generally, however, the legal rules related with quality standard of drinking water dominate the massive market of bottle water as well as water treatment and reuse that typically ingests 1 to 2% of GDP [35]. Energy is required for water transport and treatment and carbon is released when water is supplied to where the demand is.

There are particular enforced legislations in EU and UK in order to guide the water industry, market and community people to be aware about it. However, these legal rules of these legislations do not concern about energy consumption and related carbon emissions in order to water transfer, treatment and reuse.

4.2 Major Functionalities

There are two major functionalities, as it is showed in Table 1:

- Twofold legal compliance checks – (a) between EU's directive and UK's legislation related with quality standard of drinking water, and (b) between the legal quality standard of drinking water and the water citizen uses to drink from market,

- Simulation based on non-binding technical rules of (a) the energy and drinking water transfer and treatment nexus, (b) water-energy nexus, when water required for producing electricity in order to water transfer, treatment and reuse, (c) simulation of supplied water-carbon nexus.

Table 1: Quality Standard of Drinking Water and its Energy-Carbon Nexus

Stages	***Functionality***	***Legal and technical Rules***		***Rule's type***
First	Legal Compliance checking	EU Quality Standard of Drinking water		Legal rules based on legislation
		UK quality standard of Drinking water		
Second	Simulation of required energy for water transfer and treatment	***Source of drinking water and required energy [36]***		Technical rules based on social and scientific investigation
		When water comes from	***Then required energy***	
		lake/Water	.37 kWh/m3	
		Ground water	0.48 kwh/m3	
		wastewater treatment	0.62 to 0.87 kwh/m3	
		Waste water reuse	1.0 to 2.5 kwh/m3	
		Sea water	2.58 to 8.5 kwh/m3	
Third	Simulation of water required to produce electricity for water transfer and treatment	***Types of energy plant and required water [37]***		
		When energy comes from	***Then required water***	
		Solar plant with dry cooling	80 gallons Mwah/m3	
		Nuclear plants (with closed-loop cooling)	700-1100 gallons Mwah/m3	
		Nuclear plants (with open-loop cooling)	25,000-60,000 gallons Mwah/m3	
		Coal-fired plants (closed-loop)	500-600 gallons Mwah/m3	
		Coal-fired plants (open-loop)	20,000-50,000 gallons Mwah/m3	
		Biomass (crops grown for the purpose of fuel)	40,000 to 100,000 gallons Mwah/m3	
		Natural gas fracking	2-10 million gallons per well	
Fourth	Simulation of carbon emission responsible for maintaining and delivery of drinking water	Carbon emission of every litter of water supplied 0.29 g/co2 [38]		

4.3 Targeted Legal Documents

Target legal documents and rules for the use case are: (a) Article 5 and Annex 1 (Part A and B) of European Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption [39], (b) Schedule 2 of the Water Supply (Water Quality) Regulations 2000 of UK [40].

5 **Methodology**

The methodology for proposed legal knowledge framework for identifying WEFC nexus can be compartmentalized into following three major segments, where each of these segments has its own objectives and desired tools and languages to be used, as it is illustrated in Table 2:

5.1 Documentation Stage followed by Akoma Ntoso Standard

In order to documentize systematically the targeted legal documents of quality standard of drinking water of WU and UK followed by Akoma Ntoso standard, following tools is preferred to be used:

- Functional requirements for bibliographic records (FRBR) system based URI (Uniform Resource Identifier) is intended to be used to identify a uniform name of a web resource of each targeted legal documents, e.g. article, section, which will enable interactions between content of each resources over the proposed knowledge network using specific protocols of World Wide Web (WWW) [41].

- XML (EXtensible Markup Language) [42] that is to transport and store data and its related metadata of respective targeted legislations in a software-and-hardware independent and machines understandable way, and XML schema [42] that is to describe the structure of the targeted legislation.
- RDF (Resource Description Framework) that is to describe resources of target legislative documents on the web written in XML, and RDF schema (RDFS) [43] that is to extend RDF vocabularies in order to allow describing taxonomies of classes and properties of targeted legislative documents.

Table 2: Methodology for proposed Legal Knowledge Framework for identifying WEFC nexus

Major Stages	***Objectives of the methodology***	***Desired Tools and Languages to be used***	***Expected Outcomes***
Documentation stage based on Akoma Ntoso Standard	To documentize systematically the specific content of legislations of Quality Standard of Drinking Water of EU and UK following Akoma Ntoso standard.	URI, XML, XML Schema, RDF, RDF Schema, Akoma Ntoso, LIME editor	Cross compliance check and simulation of water-energy-carbon nexus
Computational Ontology stage	To represent main concepts and relationships within specific legal rules of Quality Standard of Drinking Water of EU and UK, and other related technical rules.	OWL (Web Ontological Language)	
Hybrid Reasoning stage	To model defeasible logics of legal rules, coming from specific content of legislation of Quality Standard of Drinking Water of EU and UK, and non-binding technical rules, coming from scientific communities, following LegalRuleML standard.	SPINDLE engine for hybrid reasoning and simulation and RAWE (an editor for rule markup of legal texts)	

- LIME, that is an open source based the Language Independent Markup Editor developed by CIRSFID at University of Bologna, will be used to structurize the targeted legislations maintaining Akoma Ntosao standard [44].
- Akoma Ntoso 3.0 Schema will be used as the standard for documenting targeted legislations in an XML based document format.

5.2 Computational Ontological Stage

OWL (Web Ontology Language) Full [45] will be used for legal and technical knowledge representation of related terms and concepts of targeted legislations. That will help to use the predefined relevant vocabularies stored in RDF.

Even through the state of art of computational environmental ontology is very new and on-growing, there is no such computational ontology for WEFC nexus has been yet developed. In recent literature, following types of ontologies have been evolved for expressing environmental terms and concepts, but it is noteworthy to mention that all of these ontologies are based on specific purpose or sectorial wise which are far behind the WEFC nexus's terms and concepts:

- *XeO (XEML Environmental Ontology)* expresses terms and concepts related with plant in order to help plant scientists [46].
- *Ontologies for Energy Efficiency* is dedicated exclusively to the terms and concepts of energy supply chain [47].
- In *EcoLexicon*, the terms and concepts are structured by terminological knowledge base (TKB) which is hosted in a relational database. The basic environmental conceptual underpinning are taken from the environmental event (EE) which represents the location of conceptual sub-hierarchies [48].
- *EnvO (the Environmental Ontology)* contains a comprehensive controlled and structured vocabulary of terms and concepts related with biomes, environmental features, and environmental materials [49].
- *Biome* articulates terms and concepts connected with particular patterns of ecological succession and climax vegetation [50].

These above mentioned examples give a strong observational result is that it is fundamental requirement to develop computation ontology for WEFC nexus. In the case of formalizing terms and concepts related with the above mentioned use case of WEFC domains, the differential ontological model [51] is intended to use.

5.3 Evidence based Hybrid Reasoning Stage

Following LegalRuleML, RAEW editor [52], a web editor for rule markup in LegalRuleML, and SPINDLE engine [53] will be used for evidence based hybrid reasoning [54].

5.4. *Schema for the Legal Knowledge Framework for Ex-ante and Ex-post of Policy Life Cycle*

The following, as in Figure 2, schema will be used in order to help the process each stages (analysis of the requirement, draft of the policy, implementation of the policy, monitoring of the policy and then the refinement of the policy) of entire policy life cycle of WEFC domains from standardized and systematized documentation to simulation of the multi-sectorial scenarios. The simulation and evidence based reasoning are jointly expected to play a crucial role by using norms and rules coming from various sources at every stages of WEFC domain in order to adopt the most appropriate rules and norms for policy.

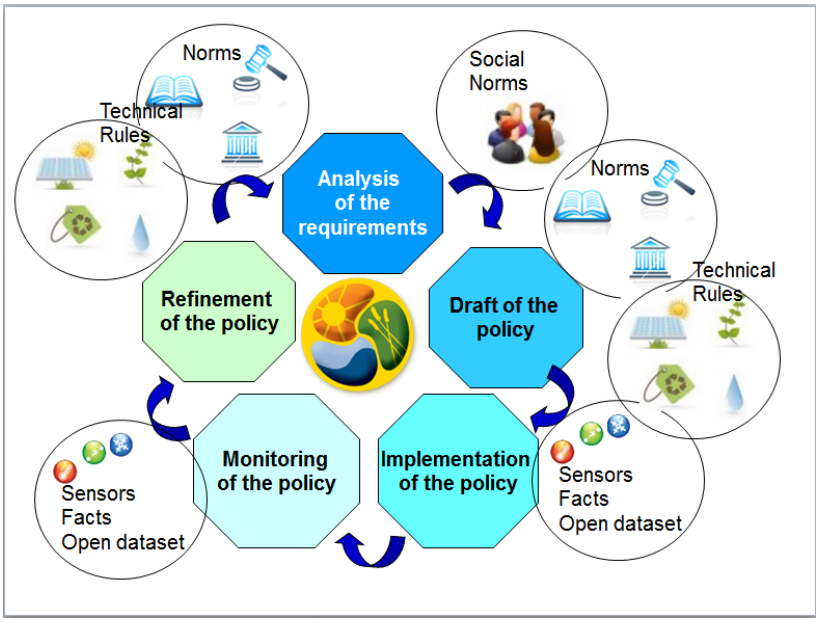


Figure 2. Schema for proposed legal knowledge framework

6 Related Works

Even though in the state of art of Environmental Decision Support Systems (EDSS), there are many useful tools, but they are very limited in scope and their functionalities, in order to simulate scenarios of different policy decisions, these tools can be clearly distinguished from this proposed framework in following ways:

- Existing EDSS tools are based on mathematical models that does not comply with legal rules, and with other relevant rules, of WEFC domain. In some extend, EDSS also integrates geographic information systems (GIS), mathematical process models, monte carlo simulation, linear programing optimization, and expert systems etc [55].
- Human rules coming from legal, institution, society, culture, ethics and news scientific discoveries usually only considered in ad hoc ways. Therefore, historically, EDSS has very limited success despite considerable effort has been made in the development of EDSS during last 25 years [56].
- They are not independent from jurisdiction, machine, language and platform. Hence these tools are not useable as anywhere policy makers want to use [57].
- They are not designed for evolutionary and evidence-based hybrid logic reasoning and creating a knowledge network for WEFC domain [58].
- They are also not designed for standardized and systematized documentation of legal documents.

However, many important learning can be shared, in the development of this proposed legal knowledge framework, from “Fill the Gap” project organized, led and funded by CIRSFID-University of Bologna [59]. Because this project

has designed an information system based on XML standards to store, in an integrated way, legal resources and rules in order to serve important roles for supporting legal knowledge engineers and end-users.

7 Critical Issues Encountered

The most critical issues encountered are : First, in the case of computational ontological representation of different terms and concepts coming from legal, institutional, social, cultural, ethical and technical perspectives of WEFC domains, the critical issue encountered is to maintain hierarchy among representations of related terms and concepts, e.g. legally binding and non-binding terms and concepts. Second, in the case of formalizing rules, using LegalRuleML standard, how the legal status of each rules coming from different legal and non-legal sources of WEFC domains will be maintained in the process of applying these different rules for evidence-based hybrid reasoning.

8 Conclusion

The paper presented a very primary idea of a legal knowledge framework for identifying WEFC nexus based on Akoma Ntoso, computational ontology and LegalRuleML standard. This proposed framework is intended to establish a knowledge network using systematized original legal documents integrated with other relevant institutional, technical, social, ethical rules of WEFC domains in order to simulate multi-sectorial scenarios of WEFC nexus with evidence-based hybrid reasoning. A use case from legislations on quality standard for drinking water of EU and UK is taken to show the possible implications of this proposed legal framework.

Acknowledgement. An especial thanks to Prof. Monica Palmirani, the Director of Erasmus Mundus LAST-JD program and Associate Professor at Department of Law in University of Bologna and also supervisor of this doctoral project, who has been providing a great level of aspirations to think in emergent and transformative way to solve the most critical problems of WEFC nexus using artificial intelligent and law.

References

[1] Hoff H. 2011. Understanding the nexus: Background paper for the Bonn2011 Nexus Conference: the water, energy and food security nexus. Stockholm, Stockholm Environment Institute. 5 p.
[2] Bazillian M, Rogner H, Howells M, Hermann S, Arent D, Geilen D, Steduto P, Mueller A, Komor P, Tol RSJ. Yumkella KY. 2011. Considering the energy, water and food nexus: towards an integrated modelling approach. Energy Policy 39(12):7896-7906.
[3] Knoepfel P, Larrue C, Varone F, Hill M. 2007. Public policy analysis. Bristol: The Policy Press, University of Bristol. 151 p.
[4] Keulertz M. 2011. The water, food, trade, energy nexus – the epitome of the next phase of green capitalism. In Kowarsch M, editor. Water management options in a globalized world: proceedings of an international scientific workshop. Munich: Munich school of philosophy, Institute for Social and Development Studies. p 134-142.
[5] Palmirani M, Vitali F. 2013 July 23. Release CSD05 - Akoma Ntoso 3.0 Schema. < http://www.akomantoso.org/release-notes/akoma-ntoso-3.0-schema/schema-for-AKOMA-NTOSO-3.0>. Accessed 2013 Oct 25.
[6] Kishore R, Sharman R, Ramesh R. 2004. Computational ontologies and information systems: foundation. In Communication of the Association for Information Systems 4: 158-183.
[7] Palmirani M, Governatori G, Rotolo A, Tabet S, Boley H, Paschke A. 2011. LegalRuleML: XML-based rules and norms. The 5th internaltion RuleMLsymposium on rules, 2011 Nov 3-5, Florida, USA. RuleML America. 298-312 p.
[8] Roberto R, Rodrik D. 2005. Rule of law, democracy, openness and income – estimating the interrelationships. In the Economics of Transition 13(3): 533-564.
[9] Waughray D, editor. 2011. Water security: the water food energy climate nexus – the World Economic Forum Water Initiative. Washington: Island press. 272 p.
[10] [NIC, USA] National Intelligence Council, US. 2012. Global trends 2030: alternative worlds. Washington: Office of the Director of National Intelligence. 5 p.
[11] [EU] European Union. 2012 May 16. European report on development 2011/2012: towards better management of water, energy and land. < http://ec.europa.eu/europeaid/what/development-policies/research-development/erd-2011-2012_en.htm>. Accessed 2013 Oct 21.
[12] The water energy and food security resource platform. 2013 Jan 1. EU support for water food and energy nexus. <http://www.water-energy-food.org/en/news/view__1191/eu-support-for-water-food-and-energy-nexus.html>. Accessed 2013 Oct 22.
[13] Hellegers P, Zilberman D, Steduto P, McCornick P. 2008. Interactions between water, energy, food and environment: evolving perspectives and policy issues. Water Policy 10(1): 1-10.
[14] Campbell A. 2008. Food, energy, water: conflicting in securities (and the rare win-wins offered by soil stewardship. Journal of soil and water conservation 63(5): 149-151.
[15] Shahbaz K, Hanjra MA. 2009. Footprints of water and energy inputs in food production: global perspectives. Food Policy 34(2): 130-140.
[16] Berger M. 2009 Nov 23. Development: water gap to widen dramatically by 2030. <http://www.globalissues.org/news/2009/11/23/3618>. Accessed 2013 Oct 24.
[17] Orts E, Spigonardo J. 2013. Special report: the nexus of food, energy and water. Pennsylvania: Institute for Global Environmental Leadership, the Wharton School, University of Pennsylvania. 5 p.
[18] Siegfried T, Sobolowski S, Raj P, Fishman R, Vasquez V, Narula K, Lall U, Modi V. 2010. Modeling irrigated area to increase water, energy and food security in semiarid India. Weather, Climate and Society 2(4): 255-270.

[19] Andrea, T. 2009 Mar 18. The energy footprint of bottled water. <http://www.livescience.com/3406-energy-footprint-bottled-water.html>. Accessed 2013 Oct 22.
[20] Pezzini M. 2012. An emerging middle class. < http://www.oecdobserver.org/news/fullstory.php/aid/3681/An_emerging_middle_class.html >. Accessed 2013 Oct 23.
[21] Jain SK. 2012. Sustainable water management in India considering likely climate and other changes. Current Science 102(2): 177-188.
[22] Partzsch L, Hughes S. 2010. Food versus fuel: governance for water rivalry. In Gottwald FT, Ingensiep HW, Meinhardt M, editors. Food ethics. New York: Springer. 153-166 p.
[23] Glassman D, Wucker M, Isaacman T, Champilou C. 2011. The water energy nexus: adding water to the energy agenda. New York: World Policy Institute. 6 p.
[24] Hoff H. 2012 May 31. Translating Nexus research into nexus policy making. < http://www.water-energy-food.org/en/news/view__630/translating_nexus_research_into_nexus_policy-making.html >. Accessed 2013 Oct 25.
[25] Hussey K, Pittock J. 2012. The energy water nexus: managing the links between energy and water for a sustainable future. In Ecology and Society 17(1): 31.
[26] UNWater. 2013. Water security and the global water agenda: a UN-Water analytical brief. Ontario: United Nations University Press. p 26.
[27] Bizikova L, Roy D, Swanson D, Venema HD, McCandless M. 2013. The water energy food security nexus: towards a practical planning and decision support framework for landscape investment and risk management. Manitoba: International Institute for Sustainable Development. p 14.
[28] Lindstrom A, Granit J. Weinberg J. 2012. Large scale water storage in the water, energy and food nexus: perspectives on benefits, risks and best practices. Stockholm: SIWI paper 21. p 19.
[29] Andrews-Speed P, Bleischwitz R, Boersma T, Johnson C, Kemp G, VanDeveer SD. 2012. The global resource nexus: the struggles for land, energy, food, water and minerals. Washington: Transatlantic Academy. p 55.
[30] Greenwood R, Wills R, Hoenig M, Pegram G, Baleta H, Morrison J, Schulte P, Farrington R. 2012. Guide to water related collective action. New York: The CEO Water Mandate. p 31.
[31] UN-ESCAP. 2013. The status of the water, food, energy nexus in Asia and the pacific. Bangkok: UN publication. p 38.
[32] McIntosh BS, Ascough JC, Twery M, Chew J, Elmahdi A, Haase D, Harou JJ, Hepting D, CUddy S, Jakeman AJ, Chen S, Kassahun A, nd others. 2011. Environmental decision support systems development – challenges and best practices. In Environmental Modelling and Software 26(12): 1389-1402 p.
[33] Ashley KD. 2009. Ontological requirements for analogical, teleological and hypothetical legal reasoning. In ICAIL 2009: 1-10.
[34] Fry A. 2006. Facts and trends water. Geneva: World Business Council for Sustainable Development. 2 p.
[35] [WWAP] World Water Assessment Programme. 2009. The United Nations world water development report 3: water in a changing world. Paris: UNESCO Publishing. p 80-95.
[36] [WBCSD] World Business Council for Sustainable Development. 2009. Water, energy and climate change: a contribution from the business community. Geneva: WBCSD. p 16.
[37] Civil Society Institute. 2012 Sep 19. U.S. Energy policy is all wet when it comes to hidden costs. <http://www.civilsocietyinstitute.org/media/091912release.cfm> Accessed 2013, Oct 19.
[38] Hackett MJ, Gray NF. 2009. Carbon dioxide emission savings potential of household water use reduction in the UK. In Journal of Sustainable Development 2(1): 36-43.
[39] EUR-Lex. 1998 Nov 03. Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31998L0083:en:NOT> Accessed 2013 October 06.
[40] Legislation.gov.uk.2000. TheWater Supply (Water Quality) Regulations 2000. < http://www.legislation.gov.uk/uksi/2000/3184/made> Accessed 2013 October 07.
[41] Ford K. 2013 Jan 28. When URIs become authority: benefits and challenges of library linked data. < http://www.loc.gov/bibframe/pdf/ALAmw2013-sac_Ford.pdf> Accessed 2013 Oct 07.
[42] Stanford.edu. XML introduction: data model, DTDs, XML schema. < http://www.stanford.edu/class/cs145/pdf/cs145xml.pdf> Accessed 2013 Oct 19.
[43] W3C. 2000 Mar 27. Resource description framework schema specification 1.0. < http://www.w3.org/TR/2000/CR-rdf-schema-20000327/> Accessed 2013 October 20.
[44] LIME-CIRSFID, 2013. Lime – the language independent markup editor. < http://lime.cirsfid.unibo.it/ > Accessed 2013 Sept 23.
[45] W3C. 2004 Feb 10. OWL web ontology language overview. < http://www.w3.org/TR/owl-features/> Accessed 2013 Oct 11.
[46] XEML environmental ontology. 2009 Sep 9. Welcome to the wiki for XEML environmental ontology. < http://xeo.codeplex.com/> Accessed 2013 Oct 24.
[47] e-SAVE. 2012. Ontologies for energy efficiency. Luxembourg: e-SAVE. p 11-13.
[48] Arauz PL, Redondo PJM. 2010. EcoLexicon: contextualizing an environmental ontology. In Proceedings of the Terminology and Knowledge Engineering (TKE) Conference 2010. Dublin: Dublin City University.
[49] [Envo] Environmental ontology. 2013. About EnvO. < http://environmentontology.org/home/about-envo > Accessed 2013 Oct 28.
[50] Environment ontology. 2012 Nov 19. Biome. < http://bioportal.bioontology.org/ontologies/ENVO?p=classes> Accessed 2013 Oct 27.
[51] Cisney VW. 2013 Jun 28. Differential Ontology. In Internet Encyclopedia of Philosophy. < http://www.iep.utm.edu/diff-ont/> Accessed 2013 Oct 25.
[52] Palmirani M, Cervone L, Bujor O, Chiappetta M. 2013. RAWE: a web editor for rule markup in LegalRuleML. In CEUR workshop proceedings. < http://ceur-ws.org/Vol-1004/paper4.pdf > Accessed 2013 Oct 25.
[53] Lam HP, Governatori G. 2009. The marking of SPINdle. In RuleML 09 proceedings of the 2009 international symposium on rule interchange and applications. Berlin: Springer. p 315-322.
[54] Xu F, Zheng X, Zhang J. Fu Z, Zhang X. 2010. A hybrid reasoning mechanism integrated evidence theory and set pair analysis in Swine-Vet. In Journal of Expert System with Application 37(10): 7086- 7093.

[55] Sullivan T. 2002. Evaluating environmental decision support system. New York: Brookhaven Science Associates for the US Department of Energy. p 02-43.
[56] Frysinger SP, Gmbh C. 2010. Environmental decision support systems: a human factors perspective. < http://sudplan.eu/polopoly_fs/1.26071!03%20-%20EDSS_SFrysinger%20%5BKompatibilitetsl%C3%A4ge%5D.pdf> accessed 2013 Oct 26

[57] Poch M, Comas J, Rodríguez-Roda I, Sánchez-Marre M, Cortes U. 2003. Designing and building real environmental decisión support system. In Enviromental Modelling and Software. < http://www.lsi.upc.edu/~ia/articulos/MPochetal-EMS.pdf> Accessed 2013 Oct 27.
[58] Sánchez-Marre M, Gibert K, Sojda RS, Steyer JP, Struss P, Rodríguez-Roda I, Comas J, Brilhante V, Roehl EA. 2008. Intelligent environmental decision support systems. In Envrionmental Modelling, Software and Decision Support Research paper 190: 119-144.
[59] Palmirani M, Ognibene T, Cervone L. . Legal rules, text and ontologies over time. 2013. In CEUR workshop proceedings. < http://ceur-ws.org/Vol-874/paper7.pdf> accessed 2013 Oct 28.

Increasing media richness in Online Dispute Resolution and the need for personal data protection

Cristiana SANTOS[a]
[a] *Phd candidate of the Joint International Doctoral Degree in Law, Science and Technology and the Universitat Autònoma de Barcelona, Institute of Law and Technology, Spain*

Abstract. This contribution introduces a new approach to online dispute resolution (ODR) and provides a portrayal of the performativity that Ambient Intelligence systems seem to convey to ODR: substantial enrichness and levels of support to the decision-making process with the provision of meaningful context information. We will portray the main issues and concerns addressed to Ambient Intelligence and we conceptualize them in the prism of online mediation. We will detail an analytical approach towards deconstructing the AmI scenario envisioned in online mediation context. We will frame privacy and data protection in the prospect of the emerging challenges raised by the development of information and communication technologies and through the filter of the ODR Regulation.

Keywords. ODR, Ambient Intelligence, Media Richness, Context-awareness, Data Protection, Privacy, Regulation

Introduction

Leading research has shown that mediation[1], as a consensual method of dispute resolution, appears to be particularly suitable to manage and solve consumer disputes [1] and has become a legal functionality [2] incorporated in the daily legal routine. According to the recent conclusions and current applications in the domain of Online Dispute Resolution[2] [3], emotions emerging in online interactions can be identified as "social functions", "contextual cues" or "indexes" in virtual environments (such as

[1] Mediation means a structured process, however named or referred to, whereby two or more parties, on a voluntary basis, try to reach an agreement on the settlement of their dispute with the assistance of a mediator. This process may be initiated by the parties or suggested or ordered by a court or prescribed by the law of a Member State, as stated in Article 3 (a) of the Directive 2008/52/EC, of the European Parliament and of the Council of 21 May 2008 on certain aspects of mediation in civil and commercial matters (OJ L136/3).

[2] Regulation n. º 524/2013 of the European Parliament and of the Council on online dispute resolution for consumer disputes (Regulation on consumer ODR), hereinafter termed simply as ODR. We consider ODR as a communicative process involving the parties engaged in an interactive decision-making task, as a mean for consumer redress. Therefore, "emotions" are an essential component in any online dispute process. Emotions have interpersonal effects on mediators that monitor the parties' emotions and use them to estimate their limits, to adjust their demands and anticipate possible obstacles to conflict resolution, therefore, shaping individual's attitudes towards the communicative and informational flow.

facial gestures, voice inflection, intonation, etc). These conclusions propose that online communication culture has parameterized its own "paralinguistic cues to express emotions (i.e. through special characters, emotions, use of capital letters, etc)." Recent findings on ODR embrace that ODR is not "emotionally limited" and may moderate major concerns about ODR as an impersonal environment, where emotions cannot be used as contextual or interactive cues. Empirical studies conclude that ODR "allows disputants to be more thoughtful in their submissions, to evaluate their emotions and express them rationally and engage at their own pace" [3]. Moreover, research has shown that pre-communication reframing and caucusing with the participants can sustain a balanced communication within a given dispute. By contrast, the most frequently concerns about ODR skeptics[3] consists that online processes cannot match the richness of the face-to-face interactions and commentators often define that parties communicating screen-to-screen are likely to experience low levels of interpersonal trust, and raise concerns about confidentiality, security, identity and higher rates of deterioration than those engaged in face-to-face interaction [5]. Cognitively, we posit that online dispute resolution "situates and intensifies the strength and the content of the communication flow"[6].

The performativity of AmI systems[4] seems to convey substantial enrichment and higher levels of support (as a serviceable tool) to online dispute resolution [7]. Thus, ODR services and technology must be constructed in such a way that their interveners will trust them as an efficient and effective way of managing their disputes [8].

Nevertheless, the fact that this environment surrounds the users and constantly acquires information about them and their context of interaction, by means of regular devices with computational power (e.g., touch screens, video cameras, accelerometers, PDAs), brings along legal requirements concerning the consent of the users and the finalities of the use of the collected data that we propose to analyze. To acquire maximum advantage from ambient intelligence, it becomes compulsory to forecast and respond to possible drawbacks and threats emerging from the new technologies[5], in

[3] Online interactions, when compared to face-to-face communication, are seen as impersonal, lacking human interaction and unable to express non-verbal cues (such as the variable tone, pitch and volume).

[4] The term Ambient Intelligence (AmI) was coined by Emile Aarts and taken up by the Advisory Group to the European Community's Information Society Technology Program (ISTAG) as the convergence of ubiquitous computing, ubiquitous communication and interfaces adapting to the user. The concept of AmI depicts a vision of the future information society, where the emphasis is on greater user-friendliness, more efficient services support, user empowerment, and support for human interactions. As an illustrative instance, during 2008, the number of things connected to the internet exceeded the number of people on earth and "these things are not just smartphones and tablets. Increasingly, the objects in our lives can now talk to us and this isn't just about health, it's also about manufacturing, the auto industry, business, government, science and everyday life. In the not too distant future, everybody, everything and every object will become a communication platform. These things are tracking our lives, giving us data about things we've never measured before. In 2014, there will be 400 million of these devices (...)", Rachel Kalmar, Data Scientist at Misfit Wearables, http://www.slideshare.net/kalmar1. As a new trend and wave of the nascent marketable technologies, the future of networked computing is called "Body Computing" and regards the wireless and mobile devices that are implanted in human bodies or wearable, both aesthetically and practically, that will one day control the future of health, lifestyle management and communication, among other things, in http://project10x.com/.

[5] Recent prototypes try to apply emotions in computer-mediated-communication, such as linguistic models to tag chat conversation with emotion tags; or even through information visualization interfaces, that enables a user to input a real-time continuous flow of their predominant emotion, by using a color spectrum which provides an insight into when, how and with what degree of certainty opinions were developed and changed over time.

order to devise and furnish appropriate safeguards regarding privacy and data protection.

The foreseen concerns unfold towards the "*homo-conectus*"[9] as the technology develops. In fact, the realm of AmI is reconfiguring and blurring the definition of the private-public space continuum, allowing the erosion of privacy. Entering in an AmI scenario appears to entail the loss of control over personal information: "the constitutive ideas of AmI, such as pervasiveness, invisibility of information systems, constant and automatic recording of events etc. render highly implausible that the user will retain control over what and how information is processed"[10]. The development of value-sensitive perceptual interfaces in pervasive and context-aware information systems, requires "design guidelines that are both specific enough to provide meaningful direction and that are sufficiently flexible to be used across systems"[10] (as the ODR system). We will seek if these new trend of specific wearable technology devices convey the meaningful and actionable data about the parties behavior within the data ecosystem.

In this line, ODR studies [11] assert that IT is not fully employed within the current ODR systems [6] [12, 6]. Conversely, the incorporation of new technologies with high penetration in different world areas may facilitate the development of ODR services: mobile penetration has grown dramatically over the past decade, and it seems that services attached to mobile devices will increase accordingly [13]. In fact, mobile artifacts are portable, durable, basic and relatively low-cost, whereas they employ easy-to-use technology and have far-reaching functionalities [14]. These characteristics might suggest that mobile devices (which incorporate sensors) may be particularly appropriate for empowering consumers in the ODR process within an AmI scenario. Moreover, empirical evidence concluded that synchronous online communication (such as chats or video-conference that are proposed in this paper) had a much higher rate of win-win solutions compared to delayed communication (asynchronous tools) [3].

In this paper, we introduce a new approach to online dispute resolution. We will describe the main issues and concerns addressed to AmI and we conceptualize them in the prism of online mediation. We will detail an analytical approach towards deconstructing AmI scenario envisioned in online mediation context. Consequently, we will frame privacy and data protection *i)* in the prospect of the emerging challenges raised by the development of information and communication technologies (on the threshold of an "ambient intelligence era"); and *ii)* through the filter of the ODR Regulation. We will assess the relevance, applicability and adequacy of the European privacy and data protection legal frameworks (encompassing the European Union Article 29 Working Party contributions and clarifications) towards these unprecedented challenges.

[6] In order to have a more forthcoming and practical insight, we quote the author excerpt "(...)The Wikipedia dispute resolution system is perhaps one of the few hallmarks of ODR 2.0: processes are highly flexible, interactive, and collaborative. But, how may other ODR initiatives benefit from both the trends and opportunities of Web 2.0? Colin Rule predicted in 2006 that ODR would be one of the biggest beneficiaries of these new technologies, because they are squarely aimed at ODR's core functionality areas: communication, collaboration, and interactivity. However, he also warned that too many ODR providers rely on outdated platforms and technology because they are reluctant to make the investments in time and resources needed to bring their platforms up to Web 2.0 standards. Colin Rule also asserts that costs have an impact on not only access but also to perceptions of distributive justice. If ODR is less expensive than other alternatives, it enhances access. Outside big marketplaces, however, there are few business models for sustainable ODR systems" [6, 12].

1. Are online mediation AmI systems compatible with privacy?

In this section we will try to expose some of the critics addressed to AmI and allocate them in the ODR framework. We will also try to respond to some primary questions on privacy and ubiquitous computing: which differences will an ubiquitous computing environment shift in our concrete lives? Is technology not only limiting, but also altering privacy? What myths and grounded concerns can be unraveled? Hereby we will try to evaluate possible answers.

Current EU legal framework differentiates the categories of data and applies a stricter protection regime towards the sensitive data. The category of sensitive data, as depicted in article 8.º of the EU Data Protection Directive (Directive 95/46/EC - hereinafter termed simply "the Directive"), makes it illegal to process personal data[7] revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and data concerning health or sex life. This is relevant particularly due to the potential threat of AmI applications to sensitive data. Images collected from the parties predictably provide information about their private life, their racial or ethnic origin; profiling parties on the basis of the very nature of the conflict (such as a consumer conflict) may transmit specifications about those persons' philosophical beliefs and status, while videos disclose visual cues (like dress, physical condition or body language and personal characteristics, as age, sex). The multitude of linking data from different sensors implies that data collected by the ubiquitous computing, are, in principle, personal data[8] or as it is also defined, "personally–identifiable information"[15].

The online mediation approach tries, in principle, to comply with data protection requirements. Processing of personal data is enclosed in Article 12 (1) of the Regulation on consumer ODR that says that access to information, including personal data, related to a dispute and stored in the ODR database shall be granted only to the ADR and ODR entities to which the dispute was transmitted. Also, all the sensitiveness of the personal data processing is submerged to the confidentiality principle settled in Article 7 of the ADR Directive (Directive 2013/11/EU): the third neutral can not reveal data conveyed during the mediation and that the parties did not authorize to disclose. [9]

It is argued that this new tools that "potentially reconfigure human experience, may also interfere with the process through which individuals come to build their own personality (process of "subjectivation") [10] and individual autonomy[10] [16 at 3] (as the "freedom from unreasonable constraints on the construction of one's own identity"). Reconducting AmI within the online mediation environment may encompass users´ intentionality, decisional power and a sense of control. The completion of a voluntary agreement in mediation, gives ODR participants a greater control over the results and

[7] The Data Protection Directive applies to the processing of "personal data", defined as any information relating to an identified or identifiable natural person ("data subject"). Article 2 of the Directive defines an identifiable person as one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his psychic, psychological, mental, economic, cultural or social identity.

[8] Article 29 Data Protection Working Party. Opinion 4/2007 on the Concept of Personal Data.

[9] Within this purpose, the Commission shall take the appropriate technical and organizational measures to ensure the security of information processed under the ODR Regulation, including appropriate data access control, a security plan and a security incident management, Article 12 (2).

[10] "(...) control over personal information is control over an aspect of the identity one projects to the world and the right to privacy is the freedom from unreasonable constraints on the construction of one's own identity [16].

control over their personal information, increasing the options for resolving conflicts without the limitations imposed by law and ensuring greater possibility of its fulfillment. ODR mediation system stems on the free-will of the parties and consists in a self-composed model. The processing of conflict resolution is based on the principle that only the parties will conduct all the process and will operate in the realm of maximization of their interests (in an interest-based approach) and are the ones who control the terms of the process and its results [17] within an inter-party trust dynamic.

The information systems involved in AmI visions intends to observe the unique complexity of each individual human being, which makes possible to emulate and produce knowledge about their users (by profiling). But in regarding to the acquired knowledge, it might typify individuals in a variety of heterogeneous categories, according to their conflict resolution style [10]. Regarding the assessment of personal conflict resolution styles (by analyzing the behavior of the disputant parties whilst they are interacting), we tend to classify their response according to the individual's assertiveness and cooperativeness. The apprehension here is to identify *if* the behavior [18] of a targeted party (its traces taken as object of scientific inquiry, monitorized, characterized, matched with other information and therefore classified), may possibly *affect* and constrain the people's classified behavior and actions (up to the standards accepted by the majority), and the effects on the people, in turn, change the classifications", and if it may possibly reconfigure human experience, revealing the "looping effect"[11] [10], or the "chilling" effect" [19] and the "making up of people" [12] [20] result. AmI systems vocation is described as "human centered", reactive to the individual's choices and needs and oriented towards empowering their users (therefore evades from the "Kafka metaphor"[13]) [21]. The engines involved in an AmI scenario provide real assets but aren't envisioned nor conceived for "making up" the intervenors of a mediation process, nor to create or mold behavioral patterns of either party in a given conflict, but for observation of contextual meaning in order to help to facilitate communication for the conflict manager, advise upon potential solutions and enhance the mediator's performance to obtain a better framed and realistic decision in real-time.

Another alleged concern relies in the possibility that the deployment of AmI technologies might *construct* or *produce* meaningful knowledge, even from trivial and fugitive image, sound or movement captured voluntarily or involuntarily released by the users, and within this conceptualization, would epitomize the "frame of the user's environment in ways that would impact and interfere on their self-perception (...), and their capacity for self-determination[14] [22]. In this line of research we aim to empower ODR settings with estimated information about the levels of stress of the parties *rather than* extensive profiling, retrieving generic or trivial data, or even waiving the users' control upon their data. We acknowledge that the ability of a mediator to form rapport

[11] "(...) AmI visions rely on systems capable of 'learning' from occurring events and incrementally self-adjusting to respond optimally to human 'needs' whereas these "needs, are decreasingly defined by the concerned 'users' themselves, but increasingly defined according to the system's interpretations of whatever happens in the contexts, and of whatever users do or even, increasingly, of what their facial expressions and body motions are"[10, at 13].

[12] (...) "They are moving targets because our investigations interact with them, and change them. And since they are changed, they are not quite the same kind of people as before. The target has moved. I call this the 'looping effect'. Sometimes, our sciences create kinds of people that in a certain sense did not exist before. I call this 'making up people'"[20].

[13] The author explains the "Kafka metaphor" through the idea of the helplessness and the vulnerability that individual's face regarding the powerful bureaucracies that handle their personal data.

[14] In article 35. º of the Portuguese Constitution is depicted the right to informational self-determination.

with parties has been found to be the most important skill a mediator can possess [23] in order to accomplish an integrative, win-win outcome. It has been observed that the social rapport and the physical and emotional cues, on the bases of correlated data, will enhance the naturalness of the emerging negotiation dialogues and thus increasing the richness of the communication medium. Whenever the mediator notices a significant change in the interaction, it induces to the rethinking of the strategies defined and to the re-orientation of the focus of the conflict resolution process in order to keep the parties interested in its resolution and to find more suitable ways of achieving an outcome. Specifically, whenever the mediator feels that it is necessary, he may choose to adapt these strategies. In order to decide when and how to perform this adaptation, the mediator interprets the information provided by an intelligent environment about the context of interaction, including the levels of escalation, the attitudes, the personal conflict styles, the emotional state (e.g. passive or emotionless behavior) or the levels of stress. This process goes on until a party leaves the process or a successful agreement is reached [41].

With this high level of "informational, emotional, relational privacy" [25] data, mediators can provide better support, enabling parties foreseeing their decisions and subjects can be more cooperative, which can result in more reliable data. In this way we argue that privacy can enhance data reliability [26] in AmI by alignining technology with the parties' interests. This endeavor approaches to traditional processes in which people communicate face-to-face and make use of the perceived feedback of the context. As soon as the relevant information is made available during the negotiating spectrum, the content of the agreement will be more consensual ("interested-based approach"), reducing the information gap that may exist and its "negative expected value"[27].

Moreover, a variety of communication methods are currently used during the mediation process and differ according to the (in)formality of the sessions, the constancy of the state of mind of the parties and the balance of power. The cadence of communication is thus adjusted to achieve the best results in the dialectic composition. It is envisaged the coexistence of a performative and fluid balance: the mediator and the parties can make use of the joint sessions, such as mediation rooms [28], online *caucus* [29] and follow-up interface (even to clarify and deepen latent inaccuracies that were detected and plausibly induced by the analysis of these parameters provided by intelligent artifacts). This personalization/customization inherent to online mediation paradigm -*mediate-centered* approach -, allows for better weighting, pondering and accuracy as to the authenticity and trustworthiness of the compiled data (and thus avoiding risks of overly weighting some reactions over others).

The applied apparatus (tactile screen and cameras) simplify users' experience so that the parties can feel they are in control of their data and that this linking and merging of data is managed in an accountable way. If users are not willing to be involved in the active protection and management of their digital assets, the trusted third party (mediator) could do this on their behalf and could provide them with easy-to-use tools to monitor and keep the situation under control.

We are already living in a world of "ubiquitous data availability" and we need pervasive privacy preserving solutions. As ambient intelligence challenges existing legal protection of privacy, conceiving and designing privacy friendly legal safeguards systems has become a priority [30], as enclosed by the "Security Safeguards Principle" that states that "personal data should be protected by reasonable security safeguards against such risks as loss or unauthorized access, destruction, use, modification or

disclosure of data"(Paragraph 11, OECD Guidelines). Regarding implementations of new and mindful security proposals, we are thus overcoming the "titanic phenomenon"[31][15] and anticipating the consequences of what can go wrong before embracing a new technology. The existing legal framework contains some important safeguards for privacy and data protection: "by default, privacy law protects the *opacity* of the individual, while data protection, also by default, calls for *transparency* of the processor of personal data"[16][19]. "However, we envisage a new paradigm where the default position will be the use of transparency tools. If the goal of regulation is to control or channel the exercise of power rather than to restrict it, then transparency tools seem more appropriate than opacity tools. In such situations, the collection and processing of data would thus be allowed, but made controllable and controlled" [19]. Therefore, transparency tools could offer a solution to some of the legal problems raised by AmI. Also, a global technical standard of data protection is needed, to support data protection laws [31].

2. Relevance, applicability and adequacy of the European privacy and data protection legal frameworks to the challenges in AmI

The instantiation of privacy and data protection is based in the European Data Protection Directive and in the important provisions listed in the OECD Privacy Guidelines that compound the classical "Fair Information Practice" principles of data protection law. In this regard and for our purposes, we will proceed with its essential legal features and we will compare with some inherent textures of computing systems, although mitigated with ODR principles. We will only give particular attention to those that exert influence and intersect to the configuration of ODR in an AmI environment.

2.1. Collection Limitation Principle and Consent (Article 6 (c) of the Data Protection Directive)

The envisioned non-invasive (and transparent to the user) approach appears to comply with the above principle. AmI, viewed contingently, purports the massive collection, aggregation and algorithmic analysis of data on everyone and everything ("dataveillance" [17] [19] or "panoptic society").[18] But the ensuing analysis of this data, in the ODR perception, has to be enforced by the data minimization principle (that allows collecting as little data as necessary for a given purpose) and shaped by the principle of

[15] The impressive term coined by Solove consists in the premise that due to the rapid pace of innovation, new mindful and technological artifacts should be first pondered and evaluated in practical terms before its deployment [31, at 199].

[16] As stated in Article 12 (4) of the ODR Regulation, each ODR advisor shall be regarded as a controller with respect to its data processing activities under this Regulation, in accordance with point (d) of Article 2 of Directive 95/46/EC, and shall ensure that those activities comply with national legislation adopted pursuant to Directive 95/46/EC.

[17] "(...) The lifeblood of AmI is, the massive collection, aggregation and algorithmic analysis of data on everyone and everything. Dataveillance brings about the second big challenge for privacy protection: the blurring of boundaries between what is private and what is public. In an AmI environment, different spaces and activities overlap. How (or even if) we can distinguish between what is private and what is not, and how privacy can be protected when its boundaries are increasingly blurred?" [19, at 2].

[18] Concept developed by Jeremy Bentham. See: http://en.wikipedia.org/wiki/Panopticon

proportionality (here implied, generally recognized as having three prongs: *(i)* suitability; *(ii)* necessity; and *(iii)* non-excessiveness).

The last part of the principle refers to the awareness and informed consent of the person whose data are being collected (Article 2 (h) and Article 7 of the Directive). For the legitimate processing of their personal data, and as a general ground for lawfulness, [19] it is legally required that the involved parties give their "unambiguous and informed consent". ODR is rendered in accordance with the principle of appropriateness, which reveals the manifestation of desirable technological neutrality of the law: the same rules apply to legal relationships online and offline [33]. The general conditions for the validation of the consent are foreseen in the Directive and apply both in the offline/online world.[20] Consequently, before entering in an AmI ODR process, parties will engage in signing a "consent term"; this form is a document that brings together all the principles inherent in the process of mediation,[21] establishing itself as a formality required for the initiation of the process. It should therefore be read and signed by all parties before mediation begins.[22] In this consent form, parties are knowledgeable of the deployment of AmI technologies framework and about the potential usage of the monitoring system and how it will be used by the mediator. This implies that all this necessary information must be given at the moment the consent is requested (information addressing the substantive aspects of the processing that the consent is intended to legitimize, such as the elements of information and transparency listed in Article 10 of the Directive). Notwithstanding, concerning the validity of individual consent, the disputants will need to provide unambiguous, specific (intelligible consent that specifies the exact purpose of the processing), expressed [22] [23] and informed consent; therefore, it is patent the requirement of "granularity" of the consent[24] with regard to the different elements that constitute the data processing. To signify this consent, the data subject will deliberately fill in offline or online forms (on a contract based form) before the processing starts; it could include a handwritten signature affixed at the bottom of a paper form or by using electronic or digital signatures (through the use of "advanced electronic signature"). Regarding sensitive data, as it is our case, it is widely admitted that the signed consent is required *ad validitatem* [34]. In principle, it should be sufficient for the data controllers to obtain consent only once, according to the specific purpose of the data processing and according to the reasonable expectations of the parties. Renewed consent is not needed from the subscriber if it is guaranteed that the data in question will not be used for other purposes other than those that were defined.[25]

---

[19] Article 29 Data Protection Working Party. Opinion 15/2011 on the definition of consent.

[20] *Idem*.

[21] The alternative dispute resolution principles are stated in Directive 2008/52/EC of the European Parliament and of the Council, of 21 May 2008, on certain aspects of mediation in civil and commercial matters and on the ADR Directive (Directive 2013/11/EU); the following principles are: Transparency principle, Independence principle, Impartiality principle, Fairness principle, Effectiveness principle, Legality principle and Liberty principle.

[22] Although contract law can protect privacy within relationships formed and articulated between parties, it does not redress privacy invasions by third parties outside of the contractual bonds.

[23] The Portuguese Data Protection Authority requires a written consent.

[24] Article 29 Data Protection Working Party. Opinion 15/2011 on the definition of consent.

[25] "(...) the need for granularity in the obtaining of consent should be assessed on a case-by-case basis, depending on the purpose(s) or the recipients of data". See Article 29 Data Protection Working Party. Opinion 15/2011 on the definition of consent.

Either disputant can veto the use of the monitoring system (the so called "zero-option" or "opt in or opt out system") [21][26] and withdraw formerly given consent, at any time. It is worthy to emphasize that the knowledge gleaned during the process is destroyed once the dispute is concluded, in order to have an accountable[27] system that protects the fundamental rights of the users. The use, non-use or shift to more privacy respecting technologies remains under the discretion of each individual user[28] [21]. This right can be preserved concerning the concrete architecture and design of ODR system.[29]

In the light of the above, regarding the cognition that the ubiquitous, proactive computing systems are so embedded in daily lives that they will literally "disappear" from users consciousness [35], so that individuals will not even necessarily be conscious of their presence and will sign gladly and willingly contract clauses, consenting to the collection, cannot be envisaged nor conceived within the online mediation configuration. Conversely, under this rights-based approach, data subjects will not become de-sensitized. The technology employed in the ODR vision is accepted as being inherently control-friendly that can comprise the "reasonable expectations" of the parties, concerning their privacy [35]. On the view just presented, the potential operational data is conveyed and gathered in one specific context: only when each party accesses the ODR system, through synchronic or assynchronic communication tools, intelligent platforms and visible devices that are activated for that purpose only (each mediation session and during each mediation process); and therefore are disconnected when that mediation process is concluded.[30] These detectable devices are relatively easy-to-apply and constitute personal devices[31] [37].

We can, at some extent, concede that the obtaining of consent, in the online mediation context, is not defined in a mechanical or perfunctory manner, or as a "routinization" of consent; and this precludes the "Fallacy of Necessity" that consists in considering legitimate, justified and proportional the necessity to have an agent's consent before an action that impacts on the agent's plans and preferences [35].

---

[26] However, there are too many collectors of information for a right to opt-out to be effective. Without a centralized mechanism for individuals to opt-out, individuals would have spend much of their time guarding their privacy.

[27] Withdrawal is not retroactive, but it should, as principle, prevent any further processing of the individual´s data by the controller, Article 29 Data Protection Working Party. Opinion 15/2011 on the definition of consent.

[28] "(...)Paul Schwartz notes how consent screens on a website asking users to relinquish control over information often do so on a "take-it-or-leave-it basis" resulting in the "fiction" that people have "expressed informed consent to [the website's] data processing practices." Individuals are often presented with an all-or-nothing choice: either agree to all forms of information collection and use or to none whatsoever. Such a limited set of choices does not permit individuals to express their preferences accurately. Individuals frequently desire to consent to certain uses of their personal information, but they do not want to relinquish their information for all possible future uses (...)"[21, at 85].

[29] The ODR system, as a matter of good practice, should endeavor to review, after a certain time, the individual's choices, e.g., by offering the possibility to either confirm or withdraw. See Article 29 Data Protection Working Party. Opinion 15/2011 on the definition of consent.

[30] A basic awareness is still achievable, e.g. through clearly visible warning tags indicating that ubiquitous computing is in use.

[31] "(...) personal recording devices still lack the full surveillance capability (...).They still miss the full ability of spontaneous networking and access to data stored anywhere, and they do not possess all analytical capacities, like dataveillance, to explore the past, or profiling to generate statements and predictions about the present and the future", [37, at 145].

2.2. Data Quality Principle (paragraph 8, OECD Guidelines)

This principle computes two dimensions: *i)* the relevance of the data for the intended purpose (which locates in close relation to the further principles that we will confer); and *ii)* the exactness, completeness and topicality of the data. In order to get more accurate data, there must be regular controls and corrections as well. In this regard, each ODR advisor shall be regarded as a controller with respect to its data processing activities, in accordance with point (d) of Article 2 of the Directive, (article 12 (4) of the ODR Regulation). In this line, Article 12 (3) regarding the Regulation on ODR, establishes that personal data related to a dispute shall be kept in the ODR database only for the *time necessary* to achieve the purposes for which they were collected and to ensure that data subjects are able to access their personal data in order to exercise their rights, and shall be automatically deleted, at the latest, *6 (six) months after the date of conclusion of the dispute* which has been transmitted to the ODR platform. Thus, it is avoided the perpetuating of the appropriation of personal data.

2.3. Purpose Specification Principle (paragraph 9, OECD Guidelines)

This principle conveys the idea that at least at the time of data acquisition, the purposes are known and identifiable. In the Data Protection Directive (Article 6 (b)), it is further specified that personal data must be collected for specified, explicit and legitimate purposes. The Purpose Specification Principle is also correlated to the "Use Limitation Principle" that specifies that "personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with Paragraph 9, except: (a) with the consent of the data subject; or (b) by the authority of law," paragraph 10 OECD, Guidelines. The legitimacy of the finality of data processing and its compatibility with these purposes can be assessed in relation to the specific and pre-definable purpose: to enrich and increase the efficiency of the communication process within a conflict resolution system with the provision of meaningful context information, and only applicable with determined and foreseeable devices. The definition of this specific purpose is the criteria for the evaluation of the lawfulness of data collection. Further, the contents and the context in which this knowledge is applied is clear at the time of collecting the data.

In our line of conceptualization, the completion of the specification purpose principle also comprises the transparency principle. The online mediator will abide to the requirements of the Directive regarding transparency of the processing of personal data. Article 10.º and 11.º demands the controller (the mediator in our case) to provide to the data subject, from whom data relating to himself are collected, with: (a) the identity of the controller; (b) the purposes of the processing for which the data are intended; (c) any further information such as the recipients or categories of recipients of the data.

The "Openness Principle"[32] is also visualized in this scenario, and is related to creating awareness about the presence of intelligent technologies. To engender online mediation in the field of ubiquitous computing systems, the release of data comprises

[32] "There should be a general policy of openness about developments, practices and policies with to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller." Paragraph 12, OECD Guidelines.

the activities the data subjects are conscious of, and therefore are, in principle, under the individual's control. The impact of online mediation AmI upon privacy is rendered evident from an analysis of particular technologies and visible terminals that are known (and some are personal) to the users (mobile devices, tactile screens and video cameras) within context-aware virtual negotiation environment, that are profile-based (in order to provide useful, non-trivial information). Moreover, it remains precise and foreseeable the moment that the monitoring system begins to operate. This substantiates our claim for the transparent extraction of features that are feasible through these devices. Hence, the knowledge derived from online mediation AmI systems matches the intentions, expectations or interests of the concerned citizen-consumers, which in term may help to mitigate information asymmetries or imbalances between the data controllers, the processors and the data subjects. It is thus evident that contextual compiled data will emerge in a transparent way to the users rather than in an unobtrusive, automatic and invisible mode.

3. Conclusions: the need for a regulatory pluralism

In this paper we suggest that the most usual criticisms to data protection envisioned in the AmI environment can be counterweighted with the specific context of ODR (within its process-centered principles and premises that are depicted in the ODR Regulation). Increasing media richness in ODR by the emotion-approach epitomizes and predicts that users will want to approach pleasant, stimulating, and controllable virtual environments and thus, emotions influence and contextual information provided by ODR processes may possibly render end-users a comfortable asset for the disclosure of their data. Even though ODR research confirms the existence of different type of services that are offered, different mechanisms employed, different IT tools used as well as the lack of interoperability services or the lack of web 2.0, web 3.0 and mobile web tools, it can be affirmed that "ODR providers understand that parties prefer to use consensual, win to win methods that entitle them to retain the ultimate decision of the controversy. Moreover, consensual methods seem to be less expensive than litigation or arbitration. Therefore, it seems that consensual-based services will increase and this seems a trend for the near future. We presume that the substantial growth of mobile penetration worldwide and some of this device's features, such as its portability, durability and relatively low-cost, suggest that mobile devices (with its incorporated sensors) might be suitable media devices.

In this article we assume that AmI leads the existing legal framework to reassess a new cognition and ponderation towards privacy and data protection law, considering the online mediation environment. The challenges of the advanced information society and the unprecedented character of a world of ubiquitous computing and ambient intelligence in an ODR scenario will mitigate and menorize the automatic collection, analysis and mining of information about the parties and contexts that may pave the way for personally–identifiable information protection, and the implied values of autonomy and self-determination.

As there is no monolithic perspective on privacy, there are multiple stake-holders and multidisciplinary endeavor in this instantiation of AmI [26]. This design conveys the proposal of relational justice, which is defined as a "bottom-up justice, produced through cooperative behavior, agreement, negotiation or dialogue" [38]. In this line, it is required adequate and matured responses from every infrastructure for fair

information practices. As such, a better regulation approach to data protection [31] would take advantage of market-based and self-regulatory institutions (including Public Law, Private Law, Soft Law, Self-Regulation, Social Norms and Technical Standards). This new approach to privacy and data protection is desirable, based on control and responsibility rather than on restriction, prohibition or "privacy myopia"[33] [39], due to the vulnerabilities affecting privacy and data protection in AmI: “the crucial issue is not the abuse but rather the fact that we have no effective means of knowing whether and when profiles are used or abused” [40]. In assessing the scope of privacy and data protection that are pertinent in the context of wearable computing, we consider transversal concern[34] [10]. A new regulatory metabolism or "regulatory pluralism" [31] including law, technology developers, ICT stakeholders and societal deliberation will need to be activated within this conceptualization of relational justice. Law and ODR may have to evolve to accommodate the new challenges raised by AmI in a communicative and evolutionist perspective. This regulatory pluralism, originating from a mainframe computer paradigm, with "built-in flexibility", can be adapted and reformed to cope with the new challenges, in order to encompass and open up the design for more privacy friendly systems and compliant tailored infrastructures.

In the present stage of research (combining AmI and ODR), it’s intricate to provide something more than simplistic and naive answers, but only modest views for the revision of EU policies and regulations, as the “Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data” (General Data Protection Regulation). The particular display of configuring ODR in AmI needs more empirical research in data protection to be fully understood, as it is necessary to be cautious about its results, since further empirical studies, tests and models are required to contrast or confirm their validity in a more general level. Nevertheless, this new advent is a promising line of research for the future of ODR.

References

1. Poblet M. and others, "Tecnologías para la mediación en línea, estado del arte, usos y propuestas", in Casanovas, P., J. Magre and M. E. Lauroba (eds), Libro Blanco de la mediación en Catalunya, Departament de Justícia, Generalitat de Catalunya edn, Huygens, 2011.
2. Poblet, M., Casanovas, P., López-Cobo, J.M., Cabrerizo, A., Prieto, J.A., Mediation, ODR, and the Web 2.0: A Case for Relational Justice. Lecture Notes in Computer Science, Vol. 6237, 2010.
3. Poblet, M., Casanovas, P., Emotions in ODR, International Review on Law, Computers, and Technology, Vol. 21(2), 2007.
4. Hammond, Anne-Marie G., How do you write “yes”? A study on the effectiveness of online dispute resolution, Conflict Resolution Quarterly, Volume 20, Issue 3, Spring 2003.
5. Ebner, Noam. ODR and Interpersonal Trust. In Wahab, Mohamed S. Abdel; Katsh, Ethan; Rainey D. (Eds.) ODR: Theory and Practice. The Hague: Eleven International Publishing, 2012, in http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2167856

[33] As Froomkin evocatively lists, "privacy myopia" consists in the fact that controllers often fail to comply with their duties under data protection laws (e.g. those arising pursuant from Articles 10 and 11 of the Directive). The author contends that data subjects suffer from inability to properly value the worth of their data in market terms, or to properly gauge the long-term significance of their consent, in terms of the impact on their privacy and autonomy [39, at 161].

[34] "(...) from the new complexities facing legal regulation of unpredictable technological developments, policy and technology have become increasingly interdependent. Legal principles, to be efficient, may need to be "embedded" in the technology itself (the development, encouraged by the European Commission, of privacy-enhancing technologies (PET’s), attests of the new distribution of regulating power between law and technology), with the implication that lawyers and engineers must engage in dialogue"[10, at 19].

6. Casanovas, P.; Poblet, M.; López-Cobo, J.M.. Relational Justice: Mediation and ODR through the World Wide Web. In F. Steiner (ed.) Archivfürrechts-und sozialphilosophie, ARSP, 2011.
7. Friedewald, M.; Vildjiounaite, E.; Punie, Y.; Wright, D., The Brave New World of Ambient Intelligence: An Analysis of Scenarios regarding Security, Security and Privacy Issues. In: Clark, J. A.; Paige, R. F. et al. (Hrsg.), Security in Pervasive Computing. Proceedings of the Third International Conference, SPC 2006, York, UK, 2006. Berlin, Heidelberg, New York: Springer (Lecture Notes in Computer Science, 3934).
8. Wahab, Mohamed S. Abdel, Does Technology Emasculate Trust? Confidentiality and Security Concerns in Online Arbitration, ICC Bulletin Special Supplement on Using Technology to Resolve Business Disputes, No. 667, 2004.
9. Andrade, Francisco, Comunicações Electrónicas e Direitos Humanos: o perigo do "homo conectus", in Direitos Humanos e sua efetivação na era da Transnacionalidade, Outubro 2012, Juruá Editora.
10. Rouvroy, Antoinette, Privacy, Data Protection and the Unprecedented Challenges of Ambient Intelligence, Studies in Ethics, Law and Technology, Vol. 2, Iss. 1, Article 3, 2008, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1013984.
11. Suquet, J; Poblet, M.; Noriega, P.; Gabarró, S., Online Dispute Resolution in 2010: a Cyberspace Odyssey? Proceedings of the 6th International Workshop on ODR 2010, CEUR Workshop Proceedings, Vol. 684, 2010.
12. Poblet M. and others, "Tecnologías para la mediación en línea, estado del arte, usos y propuestas", in Casanovas, P., J. Magre and M. E. Lauroba (eds), Libro Blanco de la mediación en Catalunya (Departament de Justícia, Generalitat de Catalunya edn, Huygens, 2011.
13. Suquet, J,"Mobile Technology and Consumer Empowerment: An Application for Online Consumer Mediation in Catalonia (Geoconsum)", European Journal for Law and Technology, Vol. 3, No. 2, 2012.
14. Poblet M., "Introduction to Mobile Technologies, Conflict Management, and ODR: Exploring Common Grounds" in Marta Poblet (ed), Mobile Technologies for Conflict Management. Online Dispute Resolution, Governance, Participation (Springer 2011).
15. Schwartz, Paul M.; Solove, Daniel J., The PII Problem: Privacy and a New Concept of Personally Identifiable Information, New York University Law Review, Vol. 86. ; UC Berkeley Public Law Research Paper No. 1909366; GWU Legal Studies Research Paper No. 584, 2011. available at SSRN: http://ssrn.com/abstract=1909366
16. Agre, Philip E.; Marc Rotenberg (eds.), Technology and Privacy. The New Landscape, MIT Press, 1998.
17. Rule, Colin, Online Dispute Resolution for Businesses. B2B, E-Commerce, Consumer, Employment, Insurance, and Other Commercial Conflicts, San Francisco, Jossey-Bass, 2002.
18. Lahlou, S, Langheinrich, M., Rocker, C., Privacy and Trust Issues with Invisible Computers, Communications of the ACM, Vol. 40, No. 3, 2005.
19. De Hert, Paul; Gutwirth, Serge; Moscibroda, Anna; Wright David; Gonzalez-Fuster, Gloria. Legal Safeguards for Privacy and Data Protection in Ambient Intelligence, http://works.bepress.com/serge_gutwirth/4/.
20. Hacking, Ian, "Making Up People", London Review of Books, 26 (16), 17 August 2007.
21. Solove, Daniel J., The Digital Person: Technology and Privacy in the Information Age, New York University Press, 2004.
22. Castro, Catarina Sarmento ,"Direito da Informática, privacidade e dados pessoais", Almedina, 2005.
23. Goldberg, Stephen B.. The Secrets of Successful Mediators, Negotiation Journal (2005) 21, 3; S.B. Goldberg & M.L. Shaw, The Secrets of Successful (and Unsuccessful) Mediators Continued: Studies 2 and 3, Negotiation Journal, 23, 2007.
24. Novais P., Carneiro D., Neves J., Incorporating Stress Estimation into User-Centred Agent-Based Platforms, in Advances on Practical Applications of Agents and Multi-Agent Systems (PAAMS 2012), Yves Demazeau, Jörg P. Müller, Juan M. Corchado Rodríguez, and Javier Bajo Pérez (Eds.), Springer - Series Advances in Intelligent and Soft Computing, vol. 155, 2012.
25. Sartor, Giovanni. Privacy, Reputation, and Trust: Some Implications for Data Protection European University Institute (EUI), Department of Law, EUI-LAW Working Papers 01/2006, in http://papers.ssrn.com/sol3/papers.cfm?abstract_id=891123
26. Jean, Camp; Kay, Connelly. "Beyond Consent", Chapter in Digital Privacy: Theory, Technologies and Practices, Alessandro Acquisti, Sabrina De Capitani di Vimercati, Stefanos Gritzalis, Costas Lambrinoudakis(eds). Auerbach Publications (Taylor and Francis Group), 2007, also available in http://www.cs.indiana.edu/~connelly/Papers/B1_BeyondConsentChapter.pdf
27. Fernando Araújo, Teoria Económica do Contrato, Almedina, 2007.
28. Benford, S.; Bullock, A; Cook, N.; Harvey, P.; Ingram, R.; Lee, O.. "From Rooms to Cyberspace: Models of Interaction in Large Virtual Computer Spaces", Interacting with Computers, Vol. 5, Issue 2, Elsevier Science, 1993.

29. Kohler, Gabrielle Kaufmann. "La resolución de los litigios en línea – perspectivas y retos del contencioso internacional contemporáneo", Revista Latino-Americana de Mediación y Arbitraje, vol. III – Nº 4, 2003.
30. Wright D.; Gutwirth S.; Friedewald M.; Punie, Y. Vildjiounaite, E., (eds.) Safeguards in a World of Ambient Intelligence, Springer Press, Dordrecht, 2008.
31. Daniel J. Solove, Nothing to Hide: The False Tradeoff between Privacy and Security, Yale University Press, 2011.
32. Winn, Jane K., Technical Standards as Data Protection Regulation, S. Gutwirth et al. (eds.) Reinventing Data Protection? Springer Science+Business Media B.V. 2009, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1118542.
33. Vicente, Dário Moura. Meios extrajudiciais de composição de litígios emergentes do comércio electrónico, in Direito da sociedade da informação, Volume V, Coimbra Editora, Coimbra, 2004.
34. Le Métayer, Daniel; Monteleone, Shara. Computer Assisted Consent for Personal Data Processing. In Proc. of 3d Conference on Legal Security and Privacy Issues in IT (LSPI'2008), p.8, in http://pop-art.inrialpes.fr/~lemetayer/lspi2008.pdf
35. Mark Weiser, Computer Science Problems in Ubiquitous Computing, Commun, ACM 36, ACM Press, 1993.
36. Brownsword, Roger, Consent in Data Protection Law: Privacy, Fair Processing, and Confidentiality. Reinventing Data Protection? Ed. Serge Gutwirth; Yves Poullet; Paul de Hert; Cecile de Terwangne; Sjaak Nouwt. Springer, 2009.
37. Cas, Johann, Ubiquitous Computing, Privacy and Data Protection: Options and Limitations to Reconcile the Unprecedented Contradictions. Computers, Privacy and Data Protection, 2011.
38. Casanovas, Pompeu; Poblet, Marta, "Concepts and Fields of Relational Justice". Computable Models of the Law, Languages, Dialogues, Games, Ontologies, 2008.
39. Bygrave, Lee Andrew; Schartum, Dag Wiese. Consent, Proportionality and Collective Power, In Serge Gutwirth; Yves Poullet; Paul De Hert; Cécile de Terwangne & Sjaak Nouwt (ed.), Reinventing Data Protection? Springer Science+Business Media B.V, 2009.
40. Hildebrandt, M., "Profiling and the identity of the European citizen." *in* Profiling the European Citizen: Cross-Disciplinary Perspectives, edited by M. Hildebrandt and S. Gutwirth. Dordrecht, Springer, 2008.
41. Carneiro D., Novais P., Andrade F., Zeleznikow J. and Neves J., Context-aware Environments for Online Dispute Resolution, in GDN 2012 - The 12th international annual meeting of the Group Decision and Negotiation conference, Recife, Brasil, 2012.

Legal Conflict Detection in Interacting Legal Systems

Tingting LI [a], Tina BALKE [b,a], Marina DE VOS [a], Julian PADGET [a] and Ken SATOH [c]

[a] *Dept. of Computer Science, University of Bath, UK*
[b] *Centre for Research in Social Simulation, University of Surrey, UK*
[c] *Principles of Informatics Res. Division, National Institute of Informatics, Japan*

Keywords. legal conflict detection, interacting legal systems, answer set programming

1. Introduction

Social rules such as laws, conventions and contracts prescribe and regulate human behaviour. It is also possible for us to break these rules at our discretion and face the consequences. In the normative multiagent systems community, normative frameworks are a way to identify a set of norms that describe the ideal behaviour of agents by specifying what is permitted, obliged and empowered within a given normative context. In [1] it was demonstrated how these frameworks could be applied to the legal domain. In [2], using the same technology, we showed that conflicts between legal specifications can be detected and resolved. Two types of conflicts are distinguished: (i) *weak* conflicts capture situations where an event/action is permitted by one legal system but prohibited by another, and (ii) *strong* conflicts between an obligation to perform an action in one and a prohibition on the action in another. An important assumption in [2] is that the specifications are independent of one another in that there is no connection between their respective transition rules. This is a strong assumption which is rare in reality, where such interactions will surely occur, intended or otherwise. This is why we here extend the work to the detection of conflicts between *interacting* legal specifications.

2. Methodology

We begin by modelling individual legal specification using an event-driven approach [1] in which a legal specification is defined over a trace of exogenous events. Starting from an initial state, each event brings about a state change, through the initiation and termination of fluents (i.e. permission, power, obligation and domain). From such a trace, we can compute a sequence of states that constitute the model of the legal specification. This process is automated through the encoding of the formal model in a computational framework built on Answer Set Programming, which is then combined with a similar translation of the different legal specifications.

Consequently, we can address the matter of the combination of a set of legal specifications – *interacting legal specification*, denoted C_I – to examine how individual specifications may interact with one another, such that either an event or a state change in one institution can trigger an event or state change, respectively, in another. We introduce two special rules in order to describe this *interaction*: (i) cross-specification generation rules provide a bridge for event generation between specifications; (ii) cross-specification consequence rules update the states of different specifications. By means of these rules, the occurrence of an external event may trigger all the constituent individual specifications to compute their next state – or if the event is not recognised by a specification, its next state is the same as its current state. Therefore, a sequence of *combined models* can be obtained from a given trace, where the combined model comprises the models of each individual specification.

We view each event trace as characterising a particular case that an interacting legal specification C_I may encounter. Conflicts are then detectable by comparing fluents from the individual models at a given time point. The whole detection procedure is implemented as an *AnsProlog* program and each generated answer set that contains an atom `conflict(X,Y,I,F)` represents the occurrence of a conflict between specification `X` and specification `Y` at time `I` with respect to fluent `F`. Furthermore, by testing all possible cases a C_I may encounter, we can identify whether the C_I is in general conflict-free.

We demonstrate this mechanism with a case study from a topical issue related to digital civil rights in Europe. The Irish Data Protection Authority (ODPC) has recently ruled that the Irish subsidiaries of Facebook are not breaking EU laws by sharing data with the NSA. The subjects involved are: Facebook Ireland, EU privacy law and US surveillance law. The data sharing activities of Facebook triggered a legal conflict between EU privacy law and US surveillance law. On the one hand, EU law states that exporting data to another country is legal only if adequate protection is provided. On the other hand, US law requires US companies to cooperate when data collection for surveillance purposes. As a subsidiary of Facebook, Facebook Ireland is placed in a dilemma, as it should abide by both US and EU law. The discussion of the ruling is out of the scope of this paper, but this case itself fits the characteristics of interacting legal specifications, in that a data sharing event without user's consent by Facebook generates the event of data exportation for EU privacy law and the event of data collection for US surveillance law respectively, leading to a state change for both EU and US legal positions. The resulting states of the EU and US in turn influence the state of Facebook with regard to the permission and obligation of sharing data. The initial state specifies that the NSA is a trusted party but there is no acceptable protection can be guaranteed by the NSA. Therefore, EU privacy law and US surveillance law disagree on the permission and obligation of Facebook's data-sharing action, resulting in a legal conflict. Such a conflict can be detected by our system automatically when given an event describing the scenario.

References

[1] M. De Vos, J. Padget, and K. Satoh. Legal modelling and reasoning using institutions. In T. Onoda, D. Bekki, and E. McCready, editors, *New Frontiers in Artificial Intelligence (JSAI-isAI 2010 Workshop)*, pages 129–140. Springer, 2011.

[2] T. Li, T. Balke, M. De Vos, J. Padget, and K. Satoh. A model-based approach to the automatic revision of secondary legislation. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law*, ICAIL '13, pages 202–206, New York, NY, USA, 2013. ACM.

Legal Conflict Detection in Interacting Legal Systems

Tingting Li [1], Tina Balke [2,1], Marina De Vos [1], Julian Padget [1] and Ken Satoh [3]

{t.li, mdv, jap}@cs.bath.ac.uk, t.balke@surrey,ac,uk, ksatoh@nii.ac.jp

[1]Department of Computer Science, University of Bath, UK

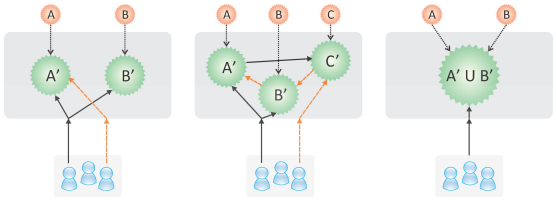
[2]Centre for Research in Social Similautaion, University of Surrey, UK

[3]Principles of Informatics Res. Division, National Institute of Informatics, Japan

Abstract

Acting under several jurisdictions at the same time is becoming the norm rather than the exception, certainly for companies but also (sometimes without knowing) for individuals. In these circumstances disparities among the different laws are inevitable. Here, we present a mathematical and a computational model of ***interacting legal specifications***, along with a mechanism to find conflicts between them. We illustrate the approach by a case study using European Privacy law.

Cooperating Legal Specifications



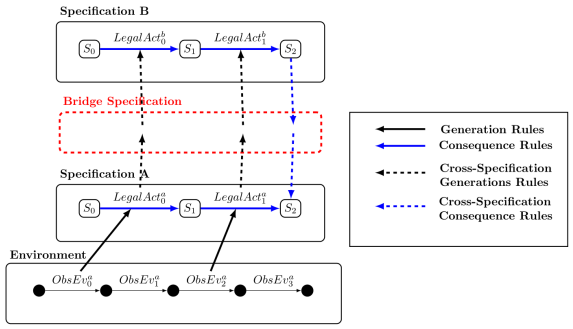
(a) (b) (c)

Three Forms of Combining Legal Specifications

(a) Comparative[2]: a set of peer legal specifications combined together to form a common governance scope, but states evolve independently.
(b) Interacting: an interlinked structure of a set of legal specifications, in which one might change the states of another. *(Focus of this work)*
(c) Merged: all the laws of each legal specification are merged to form a completely new specification.

Interacting Legal Specification

- **Bridge Specification:** separates connecting rules from the main specifications to make individuals oblivious to their interaction partners. Maintaining the flexibility and reusability of the structure.
- **Cross-specification Generation Relation:** an event in one specification triggers one or more events in one or more legal specifications:
  * $\mathbb{G}^x : \mathcal{X}^m \times \mathcal{E}^m \rightarrow \langle 2^{\mathcal{E}_1}, \ldots, 2^{\mathcal{E}_n}, 2^{\mathcal{E}^m} \rangle$
  * generation power: *gpow(source, event, destination)*
- **Cross-specification Consequence Relation:** a state change of one specification may result in a state change of another specification:
  * $\mathbb{C}^x : \mathcal{X}^m \times \mathcal{E}^m \rightarrow \langle 2^{\mathcal{F}_1}, \ldots, 2^{\mathcal{F}_n}, 2^{\mathcal{F}^m} \rangle \times \langle 2^{\mathcal{F}_1}, \ldots, 2^{\mathcal{F}_n}, 2^{\mathcal{F}^m} \rangle$.
  * initiation power: *ipow(source, fluent, destination)*
  * termination power: *tpow(source, fluent, destination)*



Legal Conflict Detection

1. Formalize and model individual legal specifications
2. Combine them to form an interacting specification
3a. *User-lead detection*
- give a particularly interesting event trace
- trigger states change of all participating specifications.
- obtain combined state model - a sequence of states.

3b. *Full-diagnose detection*
- compute all possible event traces
- for each trace, run **3a.**
- obtain combined state model in response to each trace

4. find conflicts by comparing fluent values between specifications at each state, by means of the detection program below:

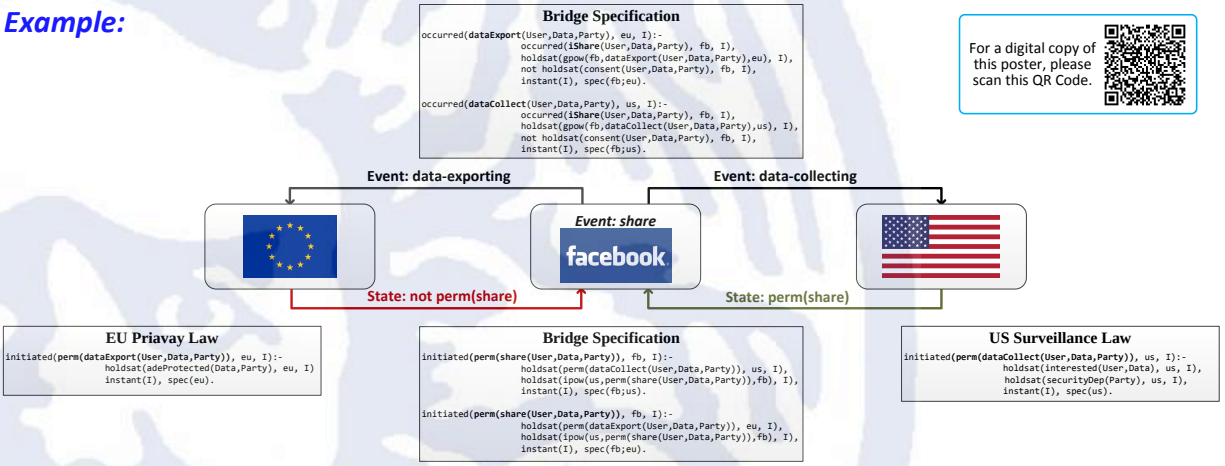
```
weakConflict(SpecX, SpecY, I, F):- holdsat(F,SpecX,I), not holdsat(F,SpecY, I),
                                   ifluent(F,SpecX), ifluent(F,SpecY),
                                   instant(I), spec(SpecX; SpecY).

strongConflict(SpecX, SpecY, I, E):- holdsat(obl(E,D,V), SpecX, I),
                                     not holdsat(perm(E), SpecY, I),
                                    ifluent(obl(E,D,V),SpecX),ifluent(perm(E),SpecY),
                                    instant(I), spec(SpecX; SpecY).
```

Reference:

[1] M. De Vos, J. Padget, and K. Satoh. Legal modelling and reasoning using institutions. In T. Onoda, D. Bekki, and E. McCready, editors, New Frontiers in Artificial Intelligence, pages 129–140. Springer, 2011.
[2] T. Li, T. Balke, M. De Vos, J. Padget, and K. Satoh. A model-based approach to the automatic revision of secondary legislation. In Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law, ICAIL'13, pages 202–206. ACM, 2013.

Example:



For a digital copy of this poster, please scan this QR Code.

Tingting Li
t.li@cs.bath.ac.uk
University of Bath

From Oral Hearing to Opinion in The U.S. Supreme Court

Latifa Al-Abdulkarim and Katie Atkinson and Trevor Bench-Capon [1]
Department of Computer Science, University of Liverpool, UK

1. Introduction

This paper provide a structured analysis of US Supreme Court Oral Hearings to enable identification of the relevant issues, factors and facts that can be used to construct a test to resolve a case. Our analysis involves the production of what we term 'argument component trees' (ACTs) in which the issues, facts and factors, and the relationship between these, are made explicit. We show how such ACTs can be constructed by identifying the speech acts that are used by the counsel and Justices within their dialogue. We illustrate the application of our analysis by applying it to the oral hearing for the case of *Carney v. California*, and we relate the majority and minority opinions delivered in that case to our ACTs. The aim of the work is to provide a formal framework that addresses a particular aspect of case-based reasoning: enabling the identification and representation of the components that are used to form a test to resolve a case and guide future behaviour.

2. The Supreme Court Process

The Supreme Court receives a number of *certiorari* requests from parties who are not satisfied with lower court decisions. Normally, when the *certiorari* is accepted, the petitioner, respondent and third parties write briefs to prepare the Justices for the oral hearings. When the Justices have considered all the briefs, the oral hearings take place. The total time for the oral hearings is just one hour, thirty minutes for each party. Normally the petitioner will begin, reserving some of his thirty minutes for rebuttal. The respondent will follow for thirty minutes, and the petitioner will finish taking the remaining time for rebuttal. Following the oral hearing, the Justices meet in conference to discuss and vote on the case. Following this the opinion arguments are prepared.

As part of the Supreme Court procedure, there are three nested dialogues in the main oral argumentation dialogue. The overall goal of the main dialogue is to establish the various components, and the connections between them, expressed as clearly and unambiguously as possible, which can be used by the justices to construct the arguments they will use in their opinions. The table in figure 1 describes the initial situation and the individual goals of each dialogue in the oral hearing which will help to drive our analysis of the dialogues [1].

[1]Department of Computer Science, University of Liverpool, Ashton street, L69 3BX Liverpool, UK. E-mail: [latifak,katie,tbc]@liverpool.ac.uk.

3. Models of Reasoning

Modelling reasoning with legal cases can be expressed as a tree of inference with a legal decision as the root and evidence as the leaves with a number of distinct layers in between. Immediately below the decision we have a level of issues, or values, which provide the reasons why the decision is made. At the next level down there are a number of *factors*. Factors are stereotypical fact patterns which, if present in a case, favour one side or the other by promoting a social value, and so are used to resolve the issues. Below the factors we have the fact patterns used to determine their presence. At the lowest level there is the evidence, which has been already considered by the time a case reaches the Supreme Court (see [1] for more discussion). Thus a complete argument for a case will comprise a view on what can be considered as evidence for relevant facts: which facts are required to establish the presence of various factors, and how they relate; how the factors can be used to determine the issues; and, where issues and values conflict, how these conflicts should be resolved.

4. Speech Act and Argument Components Tree

To enable the tree components to be proposed, we need to define speech acts for the oral hearing dialogues together with a set of critical questions challenging the components, or seeking additional components to be posed (see [1] for fuller discussion). In this poster we briefly identify the moves, and organise the argument components identified in the speech acts as an Argument Component Tree (ACT) as shown in the poster (Figure 1). For each dialogue in the oral hearing we form one ACT for the counsel and one for the Justices. Each ACT is constructed starting with the issues and it gets updated throughout the dialogue by the assertion of new factors and facts. By the end of the dialogue, each ACT shows a *complete* representation of a perspective on the components exchanged in the course of the dialogue.

5. Illustration with *California v. Carney*: From Oral hearings to Opinion

This case is concerned with whether the exception for automobiles to the protection against unreasonable search provided by the Fourth Amendment applies to mobile homes which the living area is an integral part of the of the vehicle [2]. Using the oral hearings transcript of *Careny* we applied manual analysis to propose the speech acts and construct the ACTs. One example of a petitioner ACT is shown in the poster (Figure 1).

After the oral hearing, we get four ACTs. The task now is to merge these alternatives to produce an answer for the current case, and a test applicable to future cases. This is the role of the Justices' conference stage, and, given the (competing) ACTs, could be done by top down traversal of the trees, choosing the desired elements, and evaluating the resulting structure using the facts of the case. Thus while all four trees identify privacy and exigency as issues, all three ways of linking them are available, and must be chosen between. Having identified *exigency* as an issue, a selection from the proposed factors must be made, and so on. Different Justices may make different choices, which may lead Justices to write individual opinions, either dissenting from the majority, or expressing a

different view of the appropriate tests. From *Carney's* decision, we find that the opinions offer different navigations through the components presented in the oral hearing ACTs: *all* the components used in the opinions can be found in the ACTs. Some elements form the basis of the court opinion tests. Some of the remaining facts, although not true of *Carney*, are mentioned as potentially pertinent, and so may still provide tests in future cases. Our current work concerns automation of the ACT construction and traversal.

References

[1] L. Al-Abdulkarim, K. Atkinson, and T. J. M. Bench-Capon. Dialogues in US supreme court oral hearings. In *Proceedings of CMNA 2013*, 2013. In press, see: http://www.cmna.info/CMNA13.

[2] L. Al-Abdulkarim, K. Atkinson, and T. J. M. Bench-Capon. From oral hearings to opinion in the US supreme court. In *Proceedings of Jurix 2013*, 2013. In press, see: http://sinatra.cirsfid.unibo.it/jurix2013/.

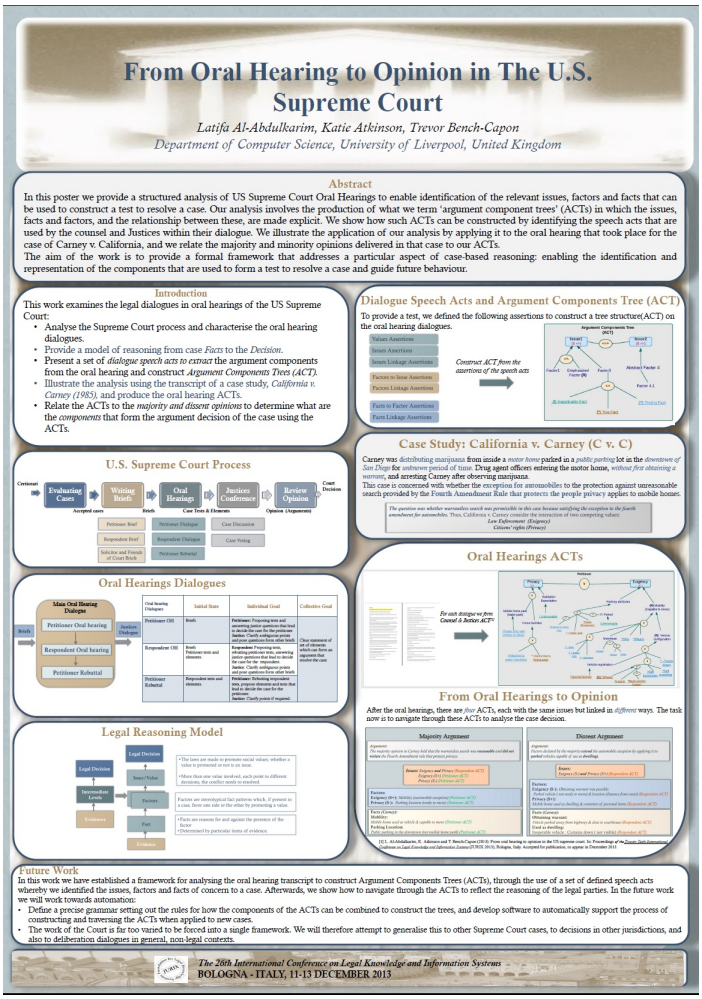


Figure 1. Poster

Addressing argumentation puzzles with model-based diagnosis (extended abstract)

Giovanni SILENO [a,1], Alexander BOER [a] and Tom VAN ENGERS [a]
[a] *Leibniz Center for Law, University of Amsterdam*

Keywords. Explanation, Model-based diagnosis, Argumentation, Agent-roles

1. Introduction

In law, and, consequently, in AI & Law, argumentation, scenario-modeling, and the combination of both, are the traditional ways of theorizing about judicial reasoning and legal truth, while probabilistic reasoning has traditionally been treated with suspicion[2]. Nevertheless, because of the growing relevance of forensic scientific evidence, a proper integration of probabilistic reasoning into the argumentation process is increasingly a debated problem.

Pollock presents in [1] a lucid philosophical critique on how probabilistic methods approach the problem of *justification*, in the form of some interesting legal puzzles. He gives the following case: *Jones says that the gunman had a moustache. Paul says that Jones was looking the other way and did not see what happened. Jacob says that Jones was watching carefully and had a clear view of the gunman.* This is an example of "collective defeat" (Paul vs Jacob), which results in a "zombie argument" (Jones'). From this story, Pollock targets some intuitive properties. (1) Given the conflict of witnesses, we should not believe to Jones' claim carelessly. (2) If we consider Paul more trustworthy than Jacob, Paul's claim should be justified, but to a lesser degree. (3) Conversely, if Jacob had confirmed Paul's claim, its "degree of justification" should have increased.[3] Pollock gives then a preliminary, elaborated proposal for degrees of justification, based on "probable probabilities". Working with a different – in one sense opposite – perspective, we have found an alternative solution to his quest.

2. Methodology

Argumentation is generally perceived as operating at a meta-level, concerned with support and attack relationships between *claims*, rather than between *messages* and *explana-*

[1] Corresponding author: g.sileno@uva.nl.

[2] In Nulty & Ors v Milton Keynes Borough Council [2013] the court puts the point concisely and – to many indignant scientists – provocatively: "you cannot properly say that there is a 25 per cent chance that something *has happened*: Hotson v East Berkshire Health Authority [1987]. Either it has or it has not".

[3] We slightly changed the third one, in order to make use of the same story.

tions. In fact, common argumentation theories treat messages directly as claims, i.e. constructions based only on the *story* level of narrative acts [2]. We propose to consider the relation between an individual message and an explanation, and the space of hypothetical explanations.

Messages are *speech acts*, and as such, they are generated and interpreted depending on the knowledge and intentions of the participants. Thus, the quest for a solution to a case requires not only an investigation into the structures and processes that made the occurrence of the case possible, but also into the process of elicitation and evaluation of explanations of the case.

In our approach, we emphasize agents' *positions*. We encourage the modeller to consider scenarios from the perspective of the participants, through the elicitation of *agent-roles*, which refer to prototypical patterns of behaviour in the target social domain.[4] Some of them represent normal behaviours, while others are associated to faulty, non compliant ones, in the sense of being at fault with the (normatively characterized) behaviour of the social system.[5]

Fundamental concepts An *observation O* consists of three elements: 1) a set of *scenario agents*, including an observer, 2) a set of messages between the observer and other agents, and 3) a temporal ordering relationship on messages (e.g. indexed on reception time). An observation becomes a *diagnostic problem* if it is surprising/alarming to the observer. Given a certain social context, an *explanation E* (or interpretation) is a multi-agent system, and consists of three elements: 1) a set of scenario agents, embodying *agent-roles*, 2) a set of messages between the agents, and 3) a (partial) temporal ordering relationship on messages. Given an observation, the observer/interpreter should be able to generate a set of explanations. An explanation may include 1) additional agents beyond the observed ones, 2) the merging of multiple agents into one agent, or 3) the splitting of an observed agent into multiple agents. To determine the relative value of an explanation E, given O, we calculate the confirmation value of O for explanation E with the measure proposed in [5], permitting ordinal judgments about explanations[6]:

$$c(O,E) = \frac{P(O|E) - P(O|\neg E)}{P(O|E) + P(O|\neg E)} \tag{1}$$

Operationalization Our methodology can be applied in three steps. First, we create executable models of the prototypical agent-roles. Second, we generate all explanatory hypotheses, allocating known agent-roles to the scenario agents. Third, we evaluate all explanations, given the messages reported to the observer/interpreter.

[4] An *agent-role* is a social intentional entity provided with certain beliefs, rationality and goal-oriented plans of actions, dual to specific social dispositions. It may be epistemically associated to multiple identities. An agent-role may produce unsuccesful outcomes too, because of faulty inputs, incomplete knowledge or wrong processing.

[5] It is worth to observe that compliance and non-compliance are qualifications relative to the position of the diagnostic agent in the social system. In a world of liars, people telling the truth would fail in respect to the social practice of systematically lying.

[6] $p(\neg E)$ is the probability that E is not the case. If $c(O,E)$ approaches 1 (-1), the observation O confirms (disconfirms) the explanation E. If c is equal to 0, the observation O is irrelevant. Put in words, with this measure, an observation confirms an explanation if it is predicted by the explanation *and* discriminates the explanation from its alternatives.

	(1) Jacob attacks Paul			(2) Jacob attacks Paul			(3) Jacob supports Paul		
	E_9	E_{24}	E_{26}	E_9	E_{24}	E_{26}	E_{13}	E_{20}	E_{30}
Jones tells the truth	true	false	false	true	false	false	true	false	false
Paul tells the truth	false	true	false	false	true	false	false	true	false
Jacob tells the truth	true	false	true	true	false	true	false	true	false
Jones saw the gunman	true	false	true	true	false	true	true	false	true
gunman had a moustache	true	false	false	true	false	false	true	false	false
$P(\texttt{k}(\texttt{Paul}))$		0.5			0.8			0.5	
$P(\texttt{k}(\texttt{Jacob}))$		0.5			0.5			0.5	
$c(O_2, E)$	0,72	0,72	0,72	0,76	0,81	0,76	0,72	0,72	0,72
$c(O_3, E)$	0.88	0.88	0.88	0.88	0.95	0.88	0.88	0.88	0.88

Table 1. Confirmation factors in Pollock's puzzles

3. Results

We apply this method to Pollock's puzzle. We consider $2^5 = 32$ possible scenarios, three scenario agents (Jones, Paul, Jacob) and two agent-roles: truth-tellers (`k`) or liars (`-k`). The outcome is summarized on Table 1, reporting only explanations confirmed by the complete observation. The following results show how we have obtained the properties targeted in the introduction. (1) Assuming indifference toward hypotheses, our approach confirms *to the same degree* hypotheses in which the gunman has a moustache, and not. (2) Using for instance $P(\texttt{k}(\texttt{Paul})) = 0.8 > P(\texttt{k}(\texttt{Jacob})) = 0.5$, the hypothesis in which Paul is telling the truth is the one confirmed to the greater degree. (3) Seeing that Jacob confirms what said by Paul, we observe that the confirmation factor of the hypothesis they both support increases, *just as much* as the hypotheses in which they are both lying. The third point is an important consequence of indifference towards prior probabilities. For instance, it allows us to consider – with the same strength – the possibility of organized crime schemes.

Obviously, our easy solution does not solve Pollock's argumentation puzzles within the rules of his game, but it clearly demonstrates the added value of our model-based diagnosis framework, proposed first in [3,4], in the field of law.

References

[1] J. L. Pollock. Reasoning and probability. *Law, Probability and Risk*, 6(1-4):43–58, 2007.

[2] G. Sileno, A. Boer, and T. van Engers. Analysis of legal narratives: a conceptual framework. In *JURIX 2012: 25th Int. Conf. Legal Knowledge and Information Systems*, 143–146, 2012.

[3] A. Boer and T. van Engers. Diagnosis of multi-agent systems and its application to public administration. *Lecture Notes in Business Information Processing*, 97:258–269. Springer, 2011.

[4] A. Boer and T. van Engers. An agent-based legal knowledge acquisition methodology for agile public administration. *ICAIL 2011: 13th Int. Conf. AI & Law*, 171–180, 2011.

[5] K. Tentori, V. Crupi, N. Bonini, and D. Osherson. Comparison of confirmation measures. *Cognition*, 103(1):107–119, 2007.

Legal Rules Modelling Meets the Web

Monica Palmirani[1], Luca Cervone[1] and Octavian Bujor[1]

[1]*CIRSFID, University of Bologna.*
{monica.palmirani, luca.cervone, octavian.bujor}@unibo.it

Introduction

This abstract presents RAWE, a Web editor for helping the legal knowledge engineer marking up legal document's using the Akoma Ntoso [2][13] XML standard, and also to model legal rules using a logic formalism and convert them into LegalRuleML [1][11]. The main goal of the RAWE Web editor is to provide a tool capable of managing in an integrated way the legal source text and the legal rules. It offers the advantages of the Akoma Ntoso and of LegalRuleML, applying the isomorphism principle [3][6][10] to connect, as far as possible, legally binding textual provisions with the logic formalism expressed using rules. AI&Law tools [8] usually are too focused on the task of applying a logic formalism to achieve isomorphism (e.g. often they use plain text, paraphrase techniques or simplified English text—ACE [1]), but the legal experts (judges, lawyers, and administrators) are interested in verifying the results of the legal reasoning engine and in finding evidence in the legally binding text that more and more, nowadays, is available on the web in digital format[2].

Secondly, a legal text changes over time, and so the rules need to be updated accordingly. If the isomorphism principle is not applied properly, it is quite difficult to determine whether those rules need to be updated [12]. The RAWE editor helps the legal knowledge engineers to maintain text and rules aligned and to minimize manual mark-up activity during the lifecycle of the legal documentation.

Thirdly, the aim of the RAWE is also to meet the Semantic Web techniques, for so it converts all the legal knowledge embedded in Akoma Ntoso and in LegalRuleML in RDF serialization to favour Linked Open Data interoperability with other legal open resources available in the cloud (e.g. geoNames [3], organizations [4], crowd-sourcing annotation, journalism, etc.).

1. RAWE Functionality

RAWE[5] is the web editor that permits the abovementioned mechanism and it provides the following functionalities: i) Authentication of the end-user and customization of the environment according with the personal profile (e.g., legal system, legal tradition, legal guidelines); ii) Multilanguage interface and environment; iii) Customized interface and buttons on the basis of the user profile and of JSON configuration files; iv) Mark-up of a legal text with Akoma Ntoso standard using parsers to automatically detect the normative references, dates, metadata, and structure of legal documents; v) Record of the XML files in the eXist repository [9]; vi) Tree of the marked-up elements; vii) On-the-fly view in Akoma Ntoso and in LegalRuleML; viii) Conversion and export in PDF, XML, ePub, or RDF format; ix) Web editor environment with WYSIWIG interface; x) Contextual functionalities based on the XML tree and XML-schemas; xi) Mouse-over for detecting the metadata of a portion of legal text and reuse for modelling legal

[1] ACE—Attempto Controlled English: http://attempto.ifi.uzh.ch/site/

[2] Euro-Lex from July 1st provides on the web the legal binding electronic publication of the Official Journal of the European Union: http://new.eur-lex.europa.eu/legal-content/EN/TXT/?qid=1374823435988&uri=CELEX:32013R0216

[3] http://www.geonames.org/

[4] http://www.w3.org/TR/vocab-org/

[5] http://sinatra.cirsfid.unibo.it/rawe/

rules; xii) Toolbar for marking up the document's structure; xiii) Toolbar for marking up legal rules.

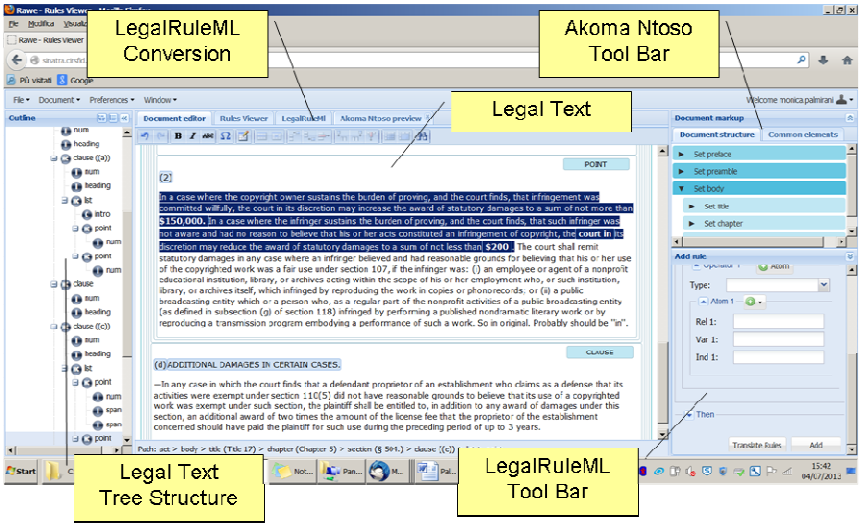


Figure 1. *RAWE* Web editor for marking up legal texts and normative rules.

There are some special functionalities that we have faced in the RAWE implementation using HCI techniques for coping with the isomorphism mechanism.

Contextual Composition of the Rules. In LegalRuleML we have five groups of statements: PrescriptiveStatement, ConstitutiveStatement, PenaltyStatement ReprationStatement and FactStatements. Each group allows some particular modeling following the legal logic theory (e.g. Prescriptive rule is a sequence of deontic operators). For this reason RAWE is based on the LegalRuleML prescriptive grammar constraints and it lead the end user to compose the rules correctly.

Reparation is a binary relationship between a penalty and a prescriptive rule or violation. So we found a smart interface mechanism to select the two parts of the relationship and to connect them to each other using drag and drop function.

2. Future Work

A first pilot case was marked up with RAWE (section 504, title 17, US code, 7 versions over time). This pilot case, even if modest, has made evident some critical points:

Metadata in Context. If we need to refine or readjust the context of the rule and the related metadata, we need a new toolbar and panel.

Extra isomorphism rules. Sometimes we need to include extra rules not directly linked to the legal text.

Ontology. Some elements of the rule modeling need to be enriched with the definitions of an external vocabulary or ontology (e.g. LKIF [5] [7]).

Meta-Rules. Meta-rules (rules about other rules), need to find a mechanism for linking rules as antecedents and consequents.

Multiple interpretation. In this version of the editor is not possible to have multiple interpretations of the same legal textual document fragment [4]. It is a crucial feature for guarantee the multiple annotation.

Granularity. For now the granularity of the isomorphism is on the rule. In the future we will be able to also manage the same functionality on the body, head, and atom.

References

[1] Athan T., Boley H., Governatori G., Palmirani M., Paschke A., Wyner A.: OASIS LegalRuleML. In Bart Verheij, ed, Proceedings of 14th International Conference on Artificial Intelligence and Law (ICAIL 2013). ACM, 2013.

[2] Barabucci G., Cervone L., Palmirani M., Peroni S., Vitali F.: Multi-layer Markup and Ontological Structures in Akoma Ntoso. In: LNCS 6237/2010, pp. 133-149, Springer, 2010.

[3] Bench-Capon T. and Coenen F.: Isomorphism and legal knowledge based systems. Artificial Intelligence and Law, 1(1):65–86, 1992.

[4] Boella G., Governatori G., Rotolo A., Torre L.V.D.: A Formal Study on Legal Compliance and Interpretation. ;In AICOL Workshops(2009), Springer, 162-183, 2011.

[5] Breuker J., Boer A., Hoekstra R., Van Den Berg C.: Developing Content for LKIF: Ontologies and Framework for Legal Reasoning, in Legal Knowledge and Information Systems, JURIX 2006, pp.41-50, ISO Press, Amsterdam, 2006.

[6] Gordon T. F., Governatori G., Rotolo A.: Rules and Norms: Requirements for Rule Interchange Languages in the Legal Domain. RuleML 2009: pp. 282-296, Springer, 2009.

[7] Gordon T. F.: Constructing Legal Arguments with Rules in the Legal Knowledge Interchange Format (LKIF). In: Computable Models of the Law, Languages, Dialogues, Games, Ontologies (2008), pp. 162-184, Springer, 2008.

[8] Gordon, T. F.: The Carneades web service. In Computational Models of Argument – Proceedings of COMMA 2012, B. Verheij, S. Szeider, and S. Woltran, Eds., IOS Press, pp. 517-518.

[9] Palmirani M., Cervone L.: Legal Change Management with a Native XML Repository. A cura di G. Governatori. Legal Knowledge and Information Systems. JURIX 2009. The Twenty-Second Annual Conference. Rotterdam. 16th-18th December 2009,pp. 146-156, Amsterdam: ISO press, 2009.

[10] Palmirani M., Contissa G., Rubino R: Fill the Gap in the Legal Knowledge Modelling. In Proceedings of RuleML 2009, pp. 305-314, Springer, 2009.

[11] Palmirani M., Governatori G., Rotolo A., Tabet S., Boley H., Paschke A.: Legal-RuleML: XML-Based Rules and Norms. RuleML America, Springer, 2011, pp.298-312.

[12] Palmirani M.: Legislative Change Management with Akoma-Ntoso, in Legislative XML for the Semantic Web, Springer, Law, Governance and Technology Series Volume 4, 2011, pp 101-130.

[13] Vitali F., Palmirani M.: Akoma Ntoso Release Notes. [http://www.akomantoso.org]. Accessed 5 July 2013.

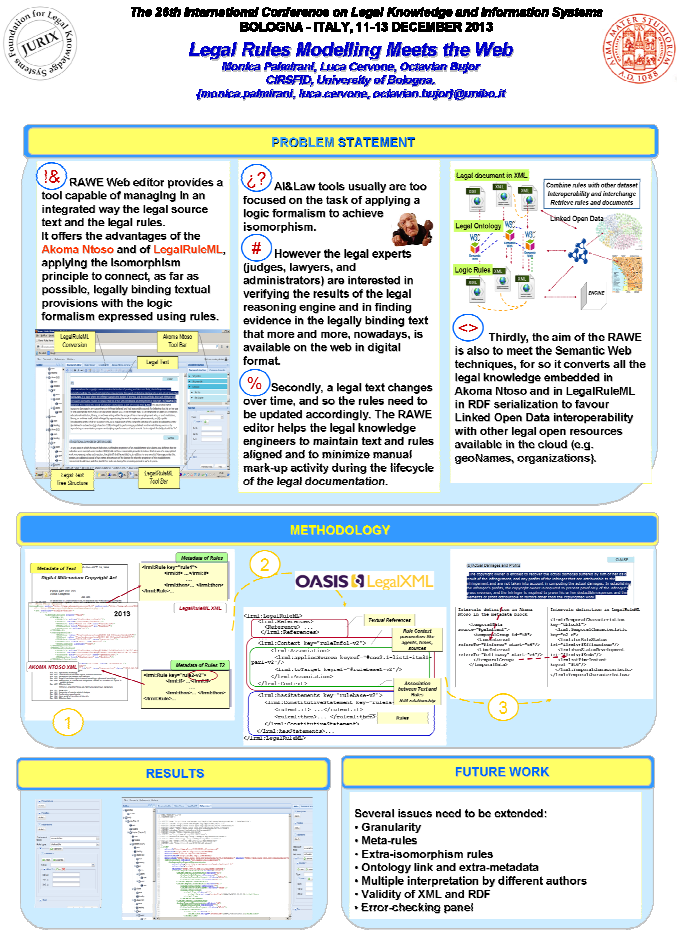


Figure 2. Poster.

#Unfair #Law: Folksonomies & Law between Openness and Knowledge

Federico COSTANTINI[a]
[a] *Dipartimento di Scienze Giuridiche, Università degli Studi di Udine*

Abstract. In this poster I suggest that folksonomies could be fruitfully used in legal information management as a collective process of "codification" carried out by the users of legal documents available on line. In this sense, through the "lattice" topology of collective tagging systems could arise a synthesis between "openness" and "knowledge", legal information retrieval and legal artificial reasoning.

Keywords. Folksonomies, Legal ontology, Legal artificial reasoning, Semantic Web

1. Background

From data to metadata: collective tagging systems. Today each Internet user might be aware that the tag is identified by the "#" (hashtag) and is associated with a hyperlink. By tagging we can: (1) describe the contents of an object, (2) label the item freely, without having to follow a preset taxonomy, (3) use any lexical expression, even belonging to natural language, (4) allocate many tags to an object or assign the same tag to different objects, and (5) share or recommend our choices and preferences.

Introducing folksonomies. Sets of categories resulting from the use of tags in the description of resources are commonly defined as "folksonomies". In folksonomes, the spontaneous activity of users generates information. Let us assume that collective tagging systems consist of three elements: (1) the users of the system (people who actually do the tagging), (2) the tags themselves, and (3) the resources being tagged.

Empirical findings on legal information management in the Internet. A few features should be addressed: (1) the relationship between legal texts and legal concepts, (2) multilingual contexts, as, for example, in the European Union, (3) transposition in different characters, such as those of Chinese, and (4) technical difficulties that affect availability of documents.

2. Theoretical framework

The theme can be addressed taking into consideration four theoretical aspects. For each level we can focus on three key concepts. Among them we can establish cross-cutting relationships.

Metaphorical level: "bottom-up", "top-down", "lattice". The figure of the "network" is often used in contemporary thought to represent the ideal synthesis between two functional patterns, the "bottom-up" and the "top-down". The pattern of

the human mind, the topology of social relations, and the logical structure of computer networks are all represented with the "lattice structure".

Epistemological level: inferential logics, deductive systems, complexity theory. The "bottom-up" model provides a reliable empirical analysis but fails to provide a satisfactory synthesis. The "top-down" model, on the contrary, allows achieving a rigorous classification of the data but excludes those which leak from *a priori* categories. Through the "complexity theory" have been developed patterns suitable to organize the data into information constructing flexible representations, that is, systems that can adapt their structure to changes in the environment.

Philosophy of law: codification of sources of law (French Civil Code), codification of legal reasoning (German Civil Code), codification as process (contemporary complex legal systems). The most recent applications of the "complexity theory" to the law are trying to overcome the limitations of the modern conception of "system" combining the theory of the sources of the law with the theory of legal reasoning. This is done by means of a "lattice" logic structure that has two main functions: (1) to open the system to the changing influences of its context, and (2) to articulate the information in a permanent organization.

Legal informatics: inferential theories (openness), legal ontologies (knowledge), folksonomies. There are two key aspects: the sharing of resources by Internet users and the representation of data in a logical-mathematical structure. In overall terms, I may refer to the former element as the "openness" and to the second as "knowledge". As of "openness", it is worth highlighting the efforts to increase as much as possible the interaction of the legal system with the social environment. Concerning "knowledge", it should be considered that the widest amount of data remains meaningless if not organized.

Considering legal information management, the two issues above outlined affect both its main research fields: legal information retrieval and legal artificial reasoning. As regards the first aspect, nowadays information technologies enable us to access not only to the legal documents, but also to the data held by public institutions (Legal Open Data) (in Italy, see http://www.dati.gov.it.). With regard to the second aspect, the amount and diversity of data that we face is such as to overwhelm not only our ability to understand but also processing capabilities of the computer. The application of folksonomy to the law allows the interaction of "openness" and "knowledge" through users activity of tagging.

3. Main issues on folksonomies and law

Several remarks have to be made. (1) Law aggregates vast communities of users, since there is an obvious interest in that matter. (2) Users belong from different cultures, backgrounds, skills and jobs. The essential difference between "experts" and "novices" nowadays seems to fade, especially if we consider the ongoing process of specialization sustained by the legal professions. (3) It could be acknowledged that law has a taxonomy that is understood or that can be learned in its broad lines by all users without special endeavour. (4) Certainly there is a huge amount of disparate legal documents (for example, legislative texts, judicial decisions, regulations, comments, scientific research, manuals, notes, but also video footages, audio tracks, and even images or pictures). (5) It seems that the description of the legal documents by users can be made more efficient and effective with some simple measures, such as

integration with legal ontologies or the suggestion of labels by users who release the documents on line (the drafters of the legislative texts, for example), namely by applying "narrow" folksonomies. (6) Through folksonomies it would be possible elaborate qualitative elements that would hardly be considered otherwise: individual feelings and beliefs, collective principles, ethical values, legal arguments. These elements, suitably treated, could be useful for the assessment of what is identified as "implicit knowledge" of the legal system and thus for building legal ontologies and graduating defeasibility in the representation of the rules. (7) Final addressees of legal system may directly influence the creation of the legal ontology that is its logic representation, in this way pushing the legal system to an effective "openness".

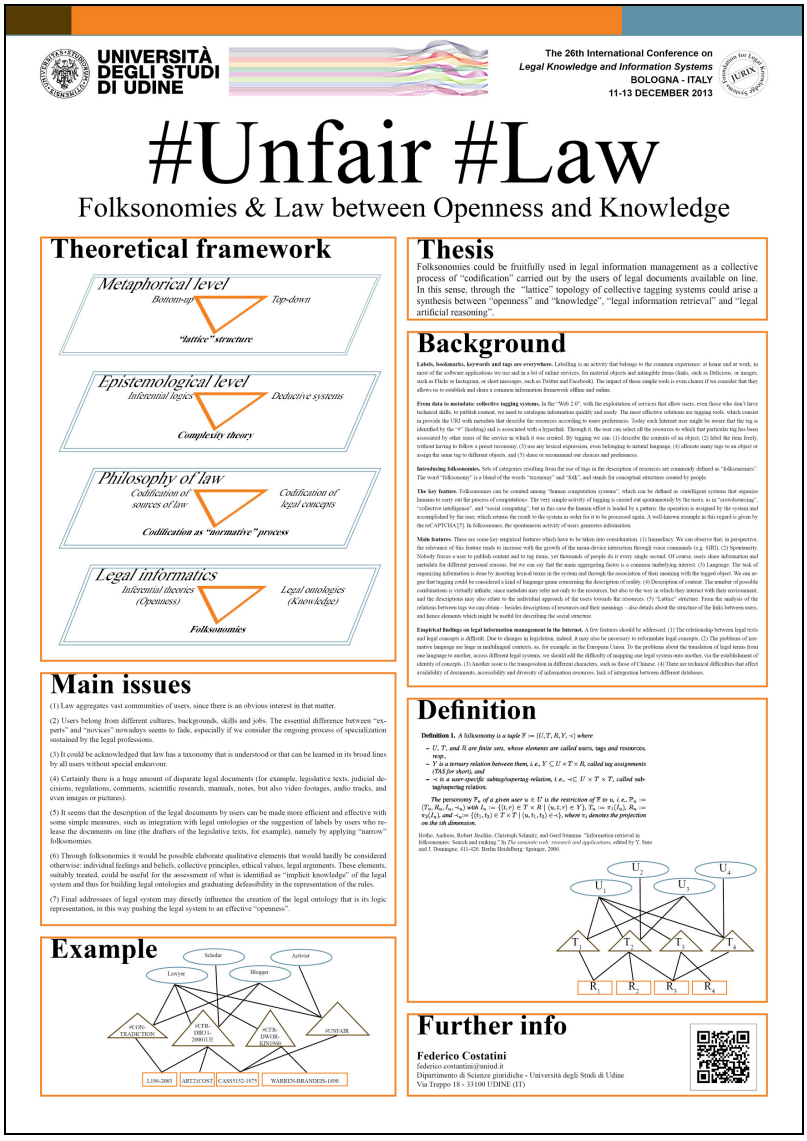


Figure 1. Poster image.

Network Analysis to Support Navigation and Use of Ever-Changing Legislation

GUIDO BOELLA [a] LUIGI DI CARO [a] LIVIO ROBALDO [a] and ANDREA VIOLATO [b]

[a] *Department of Computer Science, University of Turin*
[b] *Nomotika s.r.l.*

1. Introduction

In this abstract we introduce the working context related to the understanding of an heterogeneous network of references contained in the Italian regulatory framework. We then present an extended analysis of a large network of laws, providing several types of analytical evaluation that can be used within a legal management system for understanding the data through summarization, visualization, and browsing.

The importance of Information Technology tools lies on the ability of limiting the well-known complexity of the regulatory framework. Starting from the language used to express concepts and rules that often create serious problems of interpretation, *scientia iuris* is full of innumerable aspects that make it difficult to manage by domain experts. The use of computer systems can reduce this complexity for those who have to do with a legal reality that results to be more and more interdisciplinary, international and multifunctional. All this is even more true by taking into account one of the most complex aspects of any modern legal text, which is the extensive use of *references*. It is therefore the intention of this paper to first investigate the need for the development of tools capable of assisting the employee to manage the particular regulatory complexity characterized by cross-references, and then to simulate possible technological solutions.

The presence of thousands of stratifications in the Italian legal sources over the years has enabled the strengthening of the use of cross-references between legal texts in order to complete the content. This scenario is even more complex when considering a second and unique aspect of this discussion: the temporal dynamism that characterizes the use of normative references in any legal context. In fact, it is not a rarity that, over the years, the legislature put hand to the subjects adapting the contents to the stimuli given by socio-economic factors and the surrounding cultural context. This natural evolution is often mediated through a surgical use of references which permit to individuate the precise points under modification. In the Italian scenario, this assumption is far from being uncommon. Many, in fact, are the examples of legal decrees subsequently converted (even with modifications) into laws. Finally, it is worth noting how the use of normative references can also take place outside the purely legislative context. In fact, it is quite common to find a citation of a law inside a legal text that is not used with explicit references, but also by case law of the courts, which, in exercising a peculiar hermeneutical

activity of this rule, it implicitly complements the content, clarifying the more conceptual and obscure aspects and thus bringing important suggestions to the reader.

In [2] we introduced the Eunomos software, which is being developed in the context of the ICT4LAW project[1]. Eunomos is an advanced legal document management system based on legislative XML representation of laws which are retrieved automatically from institutional legislative portals, and incorporates a tool for building legal ontologies called Legal Taxonomy Syllabus [1].

2. Techniques and Applications

In legislation, all the data can be also viewed as complex networks where nodes are laws and links represent kinds of relationship like "*modification*", "*implementation*", "*substitution*", and so on. These information are often complex to treat, organize, and use since they are many and continuosly changing over time. Manual intervention, indeed, is often required in classical tasks because of the difficulty to get the "big picture", that is, to have an at-a-glance overview over the data under evaluation.

Social Network Analysis is quite a new field that inherits methods from Phisics (i.e., Complex Networks) and Mathematics (i.e., Theory of Graphs) to face problems related to the huge amount of data coming from social networks like Twitter[2], Facebook[3], Flickr[4], and so forth. These algorithms are useful to capture statistics about the type of the networks along their properties. More in detail, it is possible to analyze a network in terms of its evolution over time, important nodes, implicit relationships, etc. All these evaluations can be helpful also in a network of laws. [3] presents a large overview over the mathematical properties of graphs. In this section we present a set of analyses that may help the jurist to deal with these networks.

The first statistical analysis of a graph is given by the distribution of the number of neighbours of a node, also called *degree*. In real networks, the degree distribution has a tail that often follows a power law, that means that it contains many nodes with low degree and some node with large degree. This characteristic of real networks is called *community structure*. Such communities are also commonly named clusters, and they represent nodes that play similar roles within the graph. A common and often useful analysis of a network looks at such latent community structures, and it is based on some clustering approach. In our test we made use of the concept of *edge betweenness* of an edge, i.e., the extent to which it lies along shortest paths between all pairs of nodes in the network. This algorithm works by iteratively following two steps: computing the edge betweenness for all edges in the current graph and then removing the edge with the highest betweenness value. This analysis helps finding communities, since it iteratively removes central nodes to separate the graph in distinct subnetworks.

Relationships between nodes of a network may have a precise direction, that needs to be taken into account to understand the system as a whole. PageRank [4] is an algorithm

[1] ICT4LAW: ICT Converging on Law: Next Generation Services for Citizens, Enterprises, Public Administration and Policymakers funded by Regione Piemonte 2008-2013, call Converging Technologies 2007, website: http://www.ict4law.org

[2] http://www.twitter.com/

[3] http://www.facebook.com/

[4] http://www.flickr.com/

that assigns a weight to each node of a network in the World Wide Web domain, with the purpose of quantifying its relative importance within it. In spite of its original use on hyperlinks, it can be useful for several other domains dealing with directed graphs. Generally speaking, in the legal domain, if one law modifies a law B, one usually does not find on B a link back to A. In some case, there can be few relationships that may be reciprocal (i.e., citations). The PageRank algorithm measures the importance of a node by considering that of the nodes that link to it. Thus, this indicates a finer node evaluation with respect to the edge betweenness. An accurate analysis of the most important nodes within a network of laws can be helpful to understand and use the data.

The *diameter* of a graph can give useful insights. In detail, it is the largest number of nodes which must be traversed in order to travel from one node to another. In a network of laws it can be used to estimate the maximum complexity of modification/citation paths. Since a network of law is often disconnected, it is constituted by many subnetworks. In this case, the diameter considers the maximum distance found in all subnetworks.

3. Data and Results

In this section we present an analysis of the data coming from Eunomos, i.e., a network of around 10K laws interconnected by 7K links. The dataset contains different types of laws, as for instance "*stato:legge*", that indicates the principal type of law of Italy. The resulting distribution of the degree levels demonstrated that most of the laws have few connections with other laws, whereas some other laws have a large connectivity within the graph. This is in line with the majority of the networks automatic systems have to deal with.

The diameter of the network (actually, the maximum diameter of all the disconnected subnetworks in it) is 8. This means that, in the worst case, if a jurist has to navigate and understand a path between two laws in the network, it could go through other eight laws. Still, links can have different meanings, so the process may include multiple cateogories of operations. Having a tool that can support this navigation may represent a dramatically help in such process.

Finally, the distribution of the PageRank scores over the nodes/laws of the network gives an interesting result since it seems to perfectly split the important nodes (also called hubs) from the others. A system that makes use of these scores can filter out unimportant information rather than letting emerge crucial points within the regulatory framework.

References

[1] Gianmaria Ajani, Guido Boella, Leonardo Lesmo, Alessandro Mazzei, Daniele P Radicioni, and Piercarlo Rossi. Legal taxonomy syllabus: Handling multilevel legal ontologies. *Proceedings of Langtech 2008*, 2008.

[2] G. Boella, L. Humphreys, M. Martin, P. Rossi, and L. van der Torre. Eunomos, a legal document management system based on legislative XML and ontologies (position paper). *Applying Human Language Technology to the Law*, page 41, 2011.

[3] Béla Bollobás. *Modern graph theory*, volume 184. Springer, 1998.

[4] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: bringing order to the web. 1999.

Automated Methods for Extracting and Expanding Lists in Regulatory Text

Alan BUABUCHACHART, Nina CHARNESS,
Katherine METCALF, and Leora MORGENSTERN

Leidos, 4001 N. Fairfax Drive, Arlington, VA 22203

Abstract. This is a highly condensed version of a paper [1] that presents automated methods to accurately transform regulations with bulleted lists into sets of complete sentences that include their proper context. We discuss the technical challenges addressed, including extracting intended structure from HTML documents, and correctly distributing preambles over nested text. Our work has been used to preprocess the corpus used for our experiments in classifying paragraphs in regulatory documents by several categories, including illocutionary point, regulation type, and reference structure. That work is presented in a companion paper published in the JURIX 2013 proceedings.

Keywords. text analysis, regulation, preprocessing, classification

1. Introduction

Many regulatory documents contain or are largely composed of bulleted lists. Such lists break up complex text and increase comprehension for human readers, who are good at distributing preambles over bulleted text as they read. However, automated processing of such documents is difficult, especially if bulleted units are not complete sentences, and if there are multiple levels of nesting. We develop automated methods to transform text that contains bullets to text in which the bulleted text is fully expanded, with preambles distributed over the bulleted text.

This work is a preliminary step in our study of the feasibility of automating the translation of regulatory text into formal, executable rules. Our approach to this general problem involves both machine learning and deep parsing techniques; we have found that distribution is a necessary first step for both tasks. As discussed in [2], both the consistency of annotation/training data and the performance of clustering algorithms is superior when using expanded text rather than standard bulleted text, or bulleted text to which sentence splitting techniques [3] have been applied. Moreover, the loss of context inherent in sentence splitting suggests that expanded text will lead to more accurate parsing.

2. Motivating Example: The importance of context in reading bulleted lists

Domain and corpus: We are working with a corpus of 250 United States financial regulation units. Consider, e.g., the initial fragment of FINRA Rule 3240:

```
3240. Borrowing From or Lending to Customers
(a) Permissible Lending Arrangements; Conditions
 No person associated with a member in any registered capacity may borrow money from or lend money
 to any customer of such person unless:
    (1) the member has written procedures allowing the borrowing and lending of money between
such registered persons and customers of the member;
    (2) the borrowing or lending arrangement meets one of the following conditions:
      (A) the customer is a member of such person's immediate family;
      (B) the customer (i) is a financial institution regularly engaged in the business of
       providing credit ...
       and (ii) is acting in the course of such business;
      (C) the customer and the registered person are both registered persons of the same member; .....
```

Bulleted structure aids human comprehension by breaking up text. We understand that bullet (a) lists ways in which lending is allowed; that subbullet (2) specifies alternative necessary conditions constraining the relationship between customer and lender. As we read the text we must keep *context* in mind.

[3] and [4] advocate processing bulleted text by using punctuation cues to do sentence splitting. This yields sentences such as *the member has written procedures allowing the borrowing and lending of money between such registered persons and customers of the member*. Such sentences are missing context and are therefore difficult to understand.

3. Extracting from HTML, Tree Building, Distributing Preambles

In developing our technical approach, we address two hard problems: (1) extracting bulleted structure from available text; (2) building a tree structure that supports expansion and distribution of parent preambles over child bullets for arbitrarily deep levels of nesting. We can then traverse the tree to obtain the distributed text.

We recovered bulleted structure using HTML files from 6 different online law sources. Utilities like jsoup facilitate detection of paragraphs and indentation. HTML tags facilitate getting rid of junk text. Unfortunately, no source HTML files use standard bulleting tags (e.g., <ol>,) to indicate bullets in the text. Recovering the bulleted list structure is challenging because each website has its own conventions for representing lists, necessitating customized analysis. One source often has several nested labels appearing in a single line, which makes it difficult to distinguish bullet labels from references to other regulation parts and introduces potential error. For all sources, it is difficult to determine if a label like "(i)" acts as a letter or a Roman numeral, which could introduce error when multiple levels of nesting are present.

The extraction step outputs a set of labels, each of which is assigned a label type (e.g., uppercase letter, Arabic numeral) and is attached to a chunk of text in the document. The tree is then built by traversing the document:

For each paragraph

If the label type is different than the previous label type

If the label type is not on the stack
Create a new node and add it as a child of the previous node
Save previous node as the parent of this node
Put this label type on the stack
Else
Remove everything above this label type from the stack
Find the parent of the current label type
Create a new node and add it as a child of that parent
Else
Create a new node and add it as a child of the same parent of previous node

It is then easy to distribute preambles over bullet content: every path in the tree corresponds to one fully expanded bullet. One need only read out the text associated with the nodes in the path to obtain the fully expanded and distributed bullet. The text associated with all ancestors of the bullet is concatenated with the text of the bullet itself. A sample of the results for the distributed version of our example is shown below.

3240. Borrowing From or Lending to Customers (a) Permissible Lending Arrangements; Conditions No person associated with a member in any registered capacity may borrow money from or lend money to any customer of such person unless: (2) the borrowing or lending arrangement meets one of the following conditions: (B) the customer (i) is a financial institution regularly engaged in the business of providing credit

4. Results and Utility

We have achieved near perfect results in distribution of bulleted text. We have used this method to preprocess our corpus of 250 regulation units, and have found that annotation and clustering algorithms are markedly superior when working on text in which bullets have been expanded [2]). When using sentence splitting methods, we could identify definitions with an average F1 score of barely .8. (Recall was relatively low since many definitions were identified as regulations.) Using the expanded, distributed text, the F1 score rose to .95. Certain classification experiments were impossible before bullet expansion. For example, we could not annotate regulation types after sentence splitting, since the lines of text often had too little context; these difficulties disappeared once bulleted lists were expanded.

5. Acknowledgements

The research described in this paper is supported by the Intelligence Advanced Research Projects Activity (IARPA) via the Air Force Research Laboratory (AFRL) contract number FA8750-13-C-0085. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, AFRL, or the U.S. Government.
Thanks also to Adam Wyner, Ron Keesing, and Ted Senator for helpful ideas and suggestions.

References

[1] A. Buabuchachart, N. Charness, K. Metcalf, and L. Morgenstern, Automated Methods for Extracting and Expanding Lists in Regulatory Text, Working paper, at http://cs.nyu.edu/leora/papers .
[2] A. Buabuchachart, K. Metcalf, N. Charness, and L. Morgenstern, Automated Classification of Regulatory Text by Discourse Structure, Reference Structure, and Regulation Type, *JURIX 2013*.
[3] F. Dell'Orletta, S. Marchi, S. Montemagni, B. Plank, and G. Venturi, The SPLeT-2012 Shared Task on Dependency Parsing of Legal Texts, *SPLeT 2012, Workshop on Semantic Processing of Legal Text*, at *LREC 2012*, Istanbul.
[4] A. Wyner and W. Peters, On Rule Extraction from Regulations, *JURIX 2011*.

Social Media Policy and Non-Profit Organizations: Exploring the Legal Ramifications of 'Being Social'

Naomi Augar[a], Nussen Ainsworth[a], Marta Poblet[b] and John Zeleznikow[a]
[a] *Victoria University, Melbourne Australia* [b] *RMIT University, Melbourne, Australia*

1. Introduction

In recent years individuals, businesses and non-profit organizations have embraced the use of social media, online tools that support the quick and easy creation and sharing of content by users including blogs, social network sites [1]. Increasingly, organizations are noting the imperative to be where their target market is, and are creating a social media presence.

This paper contrasts three non-profit organizations' social media policies and guidelines. It identifies critical legal and ethical points for organizations to consider when devising a social media policy. While various guidelines on writing social media policies exist, the cases examined here underlined the variations in organizational approach, which may in some cases expose the employer to liability and potential litigation. In addition, non-profits have context specific considerations such as neutrality and crisis care that should be noted explicitly in their guidelines and policy for the benefit of all stakeholders. As more and more social media and employment cases are heard by tribunals and courts, it is important for all organizations to stay up to date with social media technology and the law, to ensure that their interests are protected.

Klang and Nolin provide a useful model for classifying social media policies and guidelines based on their foundation position [2]. Policies or guidelines that specify and offer different guidelines or clauses based on a specific social media platform basis (e.g. clauses tailored to Facebook and Twitter) are considered *heterogeneous*. Those that are generic, providing one policy or guideline to suit all platforms are *homogenous*. Additionally, policies or guidelines can be written so as to indicate social media is a *problem* to be managed or a *possibility* to be explored [2]. These differences are presented in a quadrant model classifying social media policy as having a *Bureaucratic foundation (Homogenous-Problem), a Branding foundation (Homogenous Possibility), a Disciplining foundation (Heterogeneous-Problem) or a Participatory foundation (Heterogeneous-Possibility)* [2]. This model will be used to classify the cases introduced in the paper.

2. Methodology

The research used an exploratory archival analysis based multi-case study [3], it allowed for cross case comparison, strengthening the insights gained from the research in comparison to a single case approach [4]. To complete the comparative document analysis the authors identified three suitable cases for analysis as part of the research via a purposive web search. Cases were restricted to non-profit organizations from the health and well-being domain, where a copy of their social media policy was available online. The sample was restricted further so as to feature one Australian based organization, one European organization, and one International organization. This criterion aimed at supporting the identification of similarities and differences that might be attributed to the relevant laws considered in devising the respective policies.

The search for cases uncovered two different approaches used by non-profits; some adopted policies and others used guidelines. The sample therefore includes both guidelines and policy to contrast these approaches to managing employee and volunteer use of social media by non-profit organizations. Inclusion of such polar cases supports comparison aimed at identifying emergent patterns [4]. This sampling approach is considered appropriate as the aim of the research is to develop insights into non-profit use of social media policy and guidelines rather than to test a theory; consequently a representative, generalizable sample is not required [4].

The selected case policies and guidelines were critically evaluated to identify the distinctions between policy and guidelines; the inclusions and gaps in each case; organization domain specific considerations and the relationship with policy/guidelines other organizational policies and the law. Alongside policy/guideline analysis, the authors reviewed the web and social media presence for each organization to identify the organizational context and to characterize their use of social media.

3. Policy and Guideline Case Analysis and Discussion

Three cases were selected for examination as part of the research: Lifeline, the Royal National Lifeboat Institution (RNLI) and the International Federation of Red Cross and Red Crescent Societies (IFRC). Lifeline is an Australian based charity that operates a 24 hour telephone and web chat crisis support service [5]. They use YouTube, Facebook and Twitter to promote their services, raise awareness about mental health issues, to support fund raising and to engage with the community.

The RNLI is a UK based charity. It supports a network of local UK organizations that provide services such as lifesaving and rescue and education and information about maritime safety [6]. Volunteers make up a considerable proportion of their work force. They use YouTube, Facebook and Twitter pages to engage with the community, raise awareness about their activities and services and to provide information to the public. Local organizations in the network may also maintain a social media presence.

The IFRC is a humanitarian network that links national Red Cross and Crescent organizations around the world as they provide health and wellbeing services to those in need [7]. They use YouTube, Facebook, Twitter, Google+ and LinkedIn to raise awareness and provide information and education about humanitarian issues, crisis, and the Societies' role in supporting those in need. Posts from or about local Red Cross and Crescent organizations are promoted as are fundraising and humanitarian initiatives. However, the IFRC primarily use LinkedIn to publicize job opportunities and requests for tenders for research programs.

A review of the social media policies and guidelines for the selected cases was conducted to classify the content into key areas of coverage. These areas included an introduction that explained social media and relevant terms, explicit guidance on appropriate and inappropriate personal and professional use of social media by employees and notes on branding including use of logos. However, areas specific to health and wellbeing organizations also featured, such as guidance on posts responding to requests for support and counseling, or posting about political or world events. The results of the policy and guideline classification are presented in the full paper.

The RNLI and IFRC tried to develop their guidelines using a *possibility* approach [2], recognizing the value that could come from promotion of the organization through personal and professional use of social media. The RNLI guidelines empowered volunteers to act by encouraging them to reflect on their use of social media and self-moderate. The IFRC guidelines also did this to a lesser degree. Thus, these policies fell into the *Branding foundation* quadrant of Klang and Nolin's model [2]. However, both guidelines veered into a *problem* approach [2] when discussing social media risks and how to manage them.

The Lifeline policy adopted a clearer *problem* approach, evidenced through a focus on management of use, the brand and discipline for misuse that was only tempered through brief encouragement of participation as long as it adhered to *guiding principles*. This approach indicated that a primarily *Bureaucratic foundation* [2] was adopted in writing the policy. All policies noted that the organization had a dedicated division or team of staff who focused on managing the official social media presence. All provided direct contacts for employees and volunteers to discuss social media with if they had specific questions or concerns. Much of the narrative around professional use focused on the need to be accurate and factual in official postings and on brand in all communications including the use of approved logos.

The IFRC guidelines stressed the importance of neutrality when discussing core business such as politically related events. This was echoed to some degree in the Lifeline policy and RNLI guidelines that discouraged specification of political affiliation.

Lifeline provided scripted examples of posts that could be tailored for use in specific instances on social media. These sample posts were for use in directing subscribers seeking counseling to contact the official help line or emergency services. Similarly, the RNLI provided some examples of what and how to post about services, even providing tips on how to boost subscriber engagement.

Interestingly, Lifeline and the RNLI had an explicit non deletion position advising staff that the integrity of social media was undermined if they deleted negative posts. This point shows considerable insight from Lifeline and the RNLI into the nature of social media and the potential backlash from the public if negative feedback is seen to be suppressed. Both guidelines indicated that posts should only be deleted in cases where the usage guidelines had been breached (e.g. discrimination). In such cases, Lifeline noted a staff member should contact the person concerned to explain the deletion whereas the RNLI suggested a sample statement advising that the post was inappropriate and had been removed.

Both Lifeline and the RNLI included points of caution relating to endorsement or advertisement of products or organizations. The RNLI guidelines cautioned that posting about partner organizations may be perceived by subscribers as advertising, which they felt may be off putting for some. However, Lifeline was specifically concerned with staff not appearing to endorse a business on Lifeline's behalf. While it has been addressed, both organizations would benefit from further detail to clarify the ramifications for employees.

Lifeline and the IFRC sanctioned employees posting the organization name as their employer in their personal social media profiles and all condoned staff posting about their work. However, they also suggested the inclusion of a disclaimer indicating that all views expressed were personal and did not reflect the organization. Interestingly, the Lifeline policy explicitly encouraged employees to re-tweet or re-post official Lifeline posts.

All three highlighted the need to protect professional confidentiality when using social media. Lifeline and the IFRC both promoted staff to be mindful of privacy protection, while only the IFRC made reference to IT security. All three also mentioned a process for reporting inappropriate posts from other staff.

The Lifeline policy was the only case where the process to be followed on detection of inappropriate staff use of social media was specified. The IFRC explicitly linked to their staff codes of conduct, employment handbook and IT use policy. Neither of the guidelines nor the policy explicitly identified the law that underpinned and related to the document. However, all included reference to common themes in employment law when specifying inappropriate use, including: a) Bullying, stalking and harassment; b)Hate speech, discrimination or other such inappropriate content; and c) Negative postings about the employer, the organization or fellow employees/volunteers

The RNLI was unique in providing a detailed section focusing on transparency of communication with children via social media. They referenced UK Government guidelines discouraging sharing or requesting information with or from children. Given that staff from all three organizations may interact with children online, this is a prudent and necessary inclusion. Aside from this reference to UK guidelines, no distinction between the cases could be made based on country of origin or international scope.

References

[1] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," *Business Horizons*, vol. 53, pp. 59-68, 2010.

[2] M. Klang and J. Nolin, "Disciplining social media: An analysis of social media policies in 26 Swedish municipalities," *2011*, 2011.

[3] R. K. Yin, *Case Study Research: Design and Methods* vol. 5. Thousand Oaks: Sage, 2009.

[4] K. M. Eisenhardt and M. E. Graebner, "Theory building from cases: Opportunities and challenges," *Academy of Management Journal*, vol. 50, pp. 25-32, 2007.

[5] Lifeline Australia. (2013, 20 August 2013). *Want to know more about Lifeline?* Available: http://www.lifeline.org.au/About-Lifeline/Want-to-know-more-about-Lifeline-/Lifeline-Information

[6] RNLI. (2013, 28 August 2013). *RNLI About Us*. Available: http://rnli.org/aboutus/Pages/About-us.aspx

[7] International Federation of Red Cross and Red Crescent Societies. (2013, 20 August 2013). *Who we are*. Available: http://www.ifrc.org/en/who-we-are/