MWAIL2015

ICAIL Multilingual Workshop
on AI & Law Research

June 8, 2015, San Diego, CA

Held in Conjunction with
The 15th International Conference on Artificial Intelligence & Law
ICAIL 2015
San Diego, June 8-12, 2015

books@ocg.at
VOLUME/BAND 313

Erich Schweighofer, Fernando Galindo, Cesar Serbena (eds.)

MWAIL2015

ICAIL Multilingual Workshop
on AI & Law Research

June 8, 2015, San Diego, CA

Proceedings

Held in Conjunction with
The 15th International Conference on Artificial Intelligence & Law
ICAIL 2015
San Diego, June 8-12, 2015

1010 Wien, Wollzeile 1-3
Komitee für Öffentlichkeitsarbeit
www.ocg.at

Druck/Print: Universität Wien, Arbeitsgruppe Rechtsinformatik
1010 Wien, Schottenbastei 10-16/2/5

Volume/Band 313, MWAIL2015 Workshop, ISBN 978-3-903035-02-7

TABLE OF CONTENTS

PREFACE

As workshop organisers, we are proud to present the proceedings of the 1st ICAIL Multilingual Workshop on Artificial Intelligence & Law Research 2015 (MWAIL2015), held in conjunction with The 15th International Conference on Artificial Intelligence & Law (ICAIL 2015), San Diego, June 8-12, 2015.

The MWAIL2015 intends to reach out to non-English speaking communities worldwide, in particular Spanish and Portuguese, to present and discuss on-going research on advanced legal search engines and legal information systems, advanced applications in e-justice, e-government, e-commerce and e-democracy, legal knowledge systems and formal models of legal systems (logic applied to legal reasoning, formal models of probability etc.), discovery of electronically stored legal information (e-discovery), machine learning and data mining for legal applications, the computational study of legal reasoning and argumentation and applications of AI and automated reasoning for the legal domain.

We have invited researchers to submit their original papers on these themes. The submission language could be Spanish, Portuguese, and English, but also other languages were not excluded (e.g. Arabic, Chinese, German, French, Japanese, Russian). In any case, a submission must also be supplemented with an extended English abstract to be included in the workshop materials. The presentation can be given in the chosen language but presentation materials, in particular slides, have to be available in English in order to allow the English speaking community to get a good understanding of the research results.

MWAIL2015 is organised by three centres on legal informatics of three universities: Universität Wien (University of Vienna) in Austria, Universidad de Zaragoza (University of Zaragoza) in Spain and Universidade Federal do Paraná (UFPR) (Federal University of Paraná, Curitiba) in Brasil.

According to its leitmotiv, multilinguality plays a major part in this workshop. It is not so much presentations in other languages than English but the development of new methods in handling multilingual textual resources. Four sessions are available: knowledge management, semantic modelling, e-justice and e-discovery.

Erich Schweighofer presents a new partial theory of legal informatics, legal data science, that should help to integrate and improve the tool box of data analysis of multilingual legal sources. *Simone Trento* focusses on factual knowledge and probability. *Aline Macohin and Cesar Antonio Serbena* analyse legal precedents in complex networks. *Johannes Scharf and Erich Schweighofer* outline the concept of a rule engine for legal reasoning with time and personal scope.

Enrico Francesconi, Mark W. Küster, Patrick Gratz and Sebastian Thelen present the CELLAR repository of the Publications Office of the EU and submit ideas for its improvement. *Bernhard Waltl and Florian Matthes* offer a support for a legal subsumption process by lexical knowledge analysis. *Cristine Griffo, João Paulo Almeida and Giancarlo Guizzardi* propose UFO-L, a Legal Core Ontology based on Alexy's Theory of Fundamental Rights. *Vytautas Čyras, Friedrich Lachmayer and Erich Schweighofer* see *v*isualization as a *Tertium Comparationis* for multilingual analysis.

Mauricio Dalri Timm Do Valle and Cesar Antonio Serbena describe electronic dispute resolution within the framework of electronic process in Brazil. *Cesar Antonio Serbena* analyses the interoperability of E-Justice systems in Brasil. *Dennis José Almanza Torres and*

Flor Zúñiga Maldonado present improvements of establishing a company in Peru using E-Justice.

The workshop closes with a Roundtable on E-Discovery of Multilingual Text Corpora (jointly with the DESI VI Workshop). Panel members are: *Jason R. Baron, Jack Conrad, Enrico Francesconi, Amanda Jones, Dave Lewis, Doug Oard, Erich Schweighofer, Cesar Serbena, and Bernhard Waltl.*

Many persons worked to form the workshop and to prepare the programme and the proceedings. Special recognition has to go to the work of the Programme Committee. The work of the hosting institution, the University of San Diego, and of the Centre for Computers and Law of the University of Vienna for editing this workshop volume deserve acknowledgement.

Vienna (AT), Zaragoza (ES) and Curitiba (BR), in May 2015

Fernando Galindo, Erich Schweighofer, Cesar Serbena

Workshop Organisers

ICAIL MULTILINGUAL WORKSHOP ON AI & LAW RESEARCH (MWAIL2015)

Monday, 8 June 2015, 9-18 hours

Institute for Peace and Justice (IPJ) at the University of San Diego

held in conjunction with
The 15th International Conference on Artificial Intelligence & Law (ICAIL 2015), San Diego, June 8-12, 2015

Workshop Organisers: Fernando Galindo, Erich Schweighofer, Cesar Serbena

Opening remarks

Erich Schweighofer, Cesar Serbena, Fernando Galindo

Session 1 – Knowledge management, 9-10.30 hours

Chaired by Enrico Francesconi

1. *Erich Schweighofer,* From Information Retrieval and AI & Law to Legal Data Science *(English)*
2. *Aline Macohin, Cesar Antonio Serbena,* Legal precedents analysis through complex networks *(Portuguese with English Summary)*
3. *Johannes Scharf, Erich Schweighofer,* rOWLer - A Rule Engine for Legal Reasoning with Time and Personal Scope *(English)*

Session 2 – Semantic Modeling, 11-13 hours

Chaired by Cesar Serbena

1. *Enrico Francesconi, Mark W. Küster, Patrick Gratz, Sebastian Thelen,* Semantic Modeling of the EU Multilingual Resources *(English)*
2. *Vytautas Čyras, Friedrich Lachmayer, Erich Schweighofer,* Visualization as a *Tertium Comparationis* within Multilingual Scientific Communities *(English)*
3. *Bernhard Waltl, Florian Matthes,* Supporting the Legal Subsumption Process: Determination of Concreteness and Abstractness in German Laws using Lexical Knowledge *(English)*
4. *Cristine Griffo, João Paulo Almeida, Giancarlo Guizzardi,* UFO-L: Towards a Legal Core Ontology based on Alexy's Theory of Fundamental Rights *(English)*

Session 3 – E-Justice, 15-16.30 hours

Chaired by Erich Schweighofer

1. *Cesar Antonio Serbena, Mauricio Dalri Timm Do Valle,* Electronic dispute resolution within the framework of electronic process in Brazil *(English)*

2. *Cesar Antonio Serbena,* Quantifying the Interoperability of Systems of Electronic Judicial Processes in the Brazilian Framework *(English)*

Session 4 – Roundtable on E-Discovery of Multilingual Text Corpora (jointly with the DESI VI Workshop), 17-18 hours

Chaired by Erich Schweighofer

Panel: Enrico Francesconi, Erich Schweighofer, Cesar Serbena, Doug Oard, Bernhard Waltl

PROGRAMME COMMITTEE

Conference Chair

Prof. Dr. Dr. Erich Schweighofer, University of Vienna, Austria

Conference Co-Chair

Prof. Dr. Fernando Galindo, University of Zaragoza, Spain

Conference Co-Chair

Prof. Dr. Cesar A. Serbena, Federal University of Parana, Brazil

Program Committee

Dr. Michal Araszkiewicz, Jagiellonian University, Kraków

Prof. Dr. Danièle Bourcier, Université de Paris II & CNRS

Prof. Dr. Fernando Galindo, University of Zaragoza, Spain

Prof. Dr. Mona Al-achkar Jabbour, Lebanese University, Beirut

Prof. Dr. Ken Satoh, National Institute of Informatics, Tokyo

Prof. Dr. Giovanni Sartor, European University Institute & University of Bologna

Prof. Dr. Dr. Erich Schweighofer, University of Vienna, Austria

Prof. Dr. Cesar A. Serbena, Federal University of Paraná, Brazil

Dr. Hughes-Jehan Vibert, Paris, ITC project manager, Ministry of Justice, Paris, France and expert for the Legal and Judicial Research Center, Algeria

Dr. Yueh-Hsuan Weng, Peking University, Beijing

FROM INFORMATION RETRIEVAL AND ARTIFICIAL INTELLIGENCE TO LEGAL DATA SCIENCE

Erich Schweighofer

Professor. University of Vienna, Centre for Computers and Law
Faculty of Law, Institute of European Law, International Law and Comparative Law,
Department of International Law
Schottenbastei 10-16/2/5, 1010 Vienna, AT
Erich.Schweighofer@univie.ac.at; http://rechtsinformatik.univie.ac.at

Descriptors: ***Legal informatics, legal data analysis, artificial intelligence, information retrieval***

Abstract: *The various methods for the analysis of the legal text corpora are systemized in a theoretical model of legal data science. 8 views – text view, metadata view, network view, user view, logic view, ontological view, visualisation view, argumentation view, 4 methods – search/reading/understanding/interpretation, documentation/ search/processing, conceptual and logic analysis, facts analysis - and 4 syntheses – commentary/handbook, dynamic electronic legal commentary, citizen information, case-based synthesis - combine traditional methodology, information retrieval and artificial intelligence methods for an efficient and powerful tool box for the user-centred analysis of legal text corpora.*

1. Introduction

Lawyers got a quite early awareness of the move to an information society, now appropriately called the knowledge and network society[1]. Already in 1970, Simitis has aptly called this phenomenon "information crisis of law"[2]. A characteristic of the modern knowledge and network society is its need for deep and co-operative regulation. This “hunger for norms” has resulted in huge legal text corpora. The handling of this knowledge is now very different from previous ages. Before starting the main task of a lawyer – reading, understanding, interpreting and arguing the law – a search process has become a necessity. However, the now well-established legal search technique has never been more than an essential support tool, not a method. Without significant semantic support, the limits are more and more evident: low recall, bad precision, and insufficient ranking.[3] Thus, a more sophisticated approach with strong use of metadata is necessary: legal data analysis. Already some years ago, a new profession was created: the data scientist[4]. “Data science is, in general terms, the extraction of knowledge from data. It employs techniques and theories drawn from many fields within the broad areas of mathematics, statistics, and information technology, including signal

[1] *Saarenpää, Ahti,* The Digital Lawyer. What skills are required of the lawyer in the Network Society?. In: Erich Schweighofer, Franz Kummer, Walter Hötzendorfer (eds.), Co-operation, Proceedings of the 18th International Legal Informatics Symposium IRIS 2015. OCG Publishers, Wien, pp.73-86 (2015); ---, Regulating the Network Society. A challenge for the quality of legislation and other activities. In: Erich Schweighofer, Ahti Saarenpää, Janos Böszörmenyi, KnowRi§ght 2012 books@ocg.at, Wien, pp 97-112 (2013).

[2] *Simitis, Spiros*, Informationskrise des Rechts und Datenverarbeitung (information crisis of law and data processing), Müller, Karlsruhe (1970).

[3] Cf. For more details the PhD research of *Anton Geist*, Rechtsdatenbanken und Relevanzsortierung (legal data bases and relevance ranking, 2015, forthcoming).

[4] *Website Futurezone,* http://futurezone.at/b2b/data-scientist-der-sexiest-job-des-21-jahrhunderts/38.908.844 of December 5, 2013, last accessed: 15 May 2015 (2013).

processing, probability models, machine learning, statistical learning, computer programming, data engineering, pattern recognition and learning, visualization, uncertainty modeling, data warehousing, and high performance computing. Methods that scale to Big Data are of particular interest in data science, although the discipline is not generally considered to be restricted to such data. The development of machine learning, a branch of artificial intelligence used to uncover patterns in data from which predictive models can be developed, has enhanced the growth and importance of data science."[5] Law should take advantage of the enormous developments of big data analytics.

Legal information retrieval and AI & law are around for a long time; legal data analysis is thus not new, but a systematic and qualitative improvement is highly required.[6] Huge collections of data are now available, allowing significantly improving the methodology for its analysis. So for, not much has been done with these big data collections besides the strong use of search engines.

The methods, techniques and theories come from many disciplines, esp. the legal doctrine, the legal information science (metadata, references, layers of time, consolidations, etc.), computer-assisted language processing (document categorization, knowledge extraction, document summarization, natural language understanding, etc.), mathematics and statistics and computer science (machine learning, pattern recognition, learning techniques, visualization, data warehousing, etc.). Paraphrasing the words of *Layman Allen* in 1957,[7] legal data science provides razor sharp tools for the analysis of the legal system. "The fight for justice and law" - in the sense of *Rudolf von Jehring*[8] - goes on, but now including legal data science. The "tool box" includes by far not only the legal search technologies, but has grown to 8 views of the legal system and 4 methods of analysis, in addition to 4 variants of synthesis. Used together, these result in a legal manual or commentary that is, e.g., a comprehensive analysis of the particular area of law. The most important advantage to its traditional counterpart – the printed text book or commentary – is that the data is computer-usable; further allowing the explicit representation of a transparent traceability.

2. Legal system

With *Hans Kelsen*[9] the law is seen as a normative order that governs the behaviour of people. A norm belongs to this system when "man should behave in certain ways because of the norm." In developing the basic norm of *Hans Kelsen* with *Alfred Verdross* and *Bruno Simma*[10], law is understood as the basic consensus of the stakeholders of the legal systems on "ought to be". Legal informatics describes this body of legal rules, e.g. the legislative processes and the respective documents. These documents are the core of the legal retrieval system (legal information system). This is still primarily done in a textual way, but is increasingly being enriched with multimedia

---

[5] *Wikipedia EN,* data science, http://en.wikipedia.org/wiki/Data_science (last accessed: 15 May 2015); see also a website with a particular legal focus http://computationallegalstudies.com/2014/07/data-science-big-data-statistics-can-live-together/ (last accessed: 15 May 2015).

[6] *Bing, Jon (Ed.),* Handbook of Legal Information Retrieval, North-Holland, Amsterdam (1984); *Schweighofer, Erich,* Wissensrepräsentation und automatische Textanalyse im Völker- und Europarecht, Habilitationsschrift, Universität Wien 1996, publication in English: Legal Knowledge Representation, Automatic Text Analysis in Public International and European Law, Kluwer Law International, Law and Electronic Commerce, Volume 7, The Hague (1999); *Jackson, Peter, Moulinier, Isabelle,* Natural Language Processing for Online Applications. Text Retrieval, Extraction and Categorization, Second revised edition, John Benjamins Publishing Company (2007).

[7] *Allen, Layman,* Symbolic Logic: A Razor-Edged Tool for Drafting and Interpreting Legal Documents, in: The Yale Law Journal, Vol 66, pp. 833-879 (1957).

[8] *Jehring, Rudolf von,* Der Kampf ums Recht, Vortrag (The fight for law), Wien, 1872. Schutterwald/Baden 1997 (1872).

9 *Kelsen, Hans*, Reine Rechtslehre (Pure Theory of Law), 2nd ed., pp 196 et seq. (1960).

[10] *Verdross, A./Simma, B.,* Universelles Völkerrecht, Theorie und Praxis, (Universal International Law, Theory and Practice), Third, completely revised edition, Duncker & Humblot, Berlin (1986).

elements. This corpus is huge: at least several gigabytes of data, millions of documents over a million records in the database dictionary, etc.

3. Legal data science

The goal of legal data science is to complement the existing methodology of law with the new computer-based methods, and to bring in a theoretical framework. The new methods are presented extensively and the old methods are treated rather cursorily due to space reasons. This theory is still developing; it is at its start. The first step should be a unification of both so far more coexisting areas of information retrieval and AI & Law and a classification of the combined methodologies. The first step above all, should be a classification of its methods.

So far, the theory development has long been only in the focus by legal theory and information science. Only since the late 1950s – with the start of the research on legal information retrieval – appropriate progress can be noted (term retrieval, text retrieval, metadata, citations, search technologies, user interface, telecommunications, etc.).

Previous practice, a rather practical and cursory description of sources and methods of interpretation in textbooks, is outdated due to the new scientific analysis of law. The focus must be on the analysis of the legal multimedia corpus.

The classification distinguishes 8 views – text view, metadata view, network view, user view, logic view, ontological view, visualisation view, argumentation view, 4 methods – search/reading/understanding/interpretation, documentation & search, conceptual and logic analysis, visualisation, and 4 syntheses – commentary/handbook, dynamic electronic legal commentary, citizen information and case-based synthesis.

3.1. 8 views (representations) of the law

The "views theory" of Lu and Conrad[11] - documents view, annotation view, citation view and user view constitute a basis for this work. However, it should be extended by four other views – legal logic, legal ontologies, legal visualisation and legal argumentation. Thus, the 8 views or representations are as follows: Text (multimedia) corpus, metadata, citation network, user view, logic view, ontological view, visualisation view and argumentation view.

The basis of this methodology are the works of Sowa[12] , Fiedler[13] , Zeleznikow/Hunter[14], the authors of the book edited by Yearwood/Stranieri[15] as well as my own work on knowledge representation of the law.[16]

This formalization of the law has started in the 1970s; the research is well documented in the proceedings of the conferences ICAIL[17] and JURIX[18]. In this paper, we argue for a stronger

[11] *Lu Qiang, Conrad, Jack,* Next generation legal search - it's already there. In: Cornell Legal Information Institute. In: VoxPopuLII, https://blog.law.cornell.edu/voxpop/2013/03/28/next-generation-legal-search-its-already-here/ (2013).

[12] *Sowa, John F.* Knowledge representation: logical, philosophical, and computational foundations. Course Technology, Boston, MA (2000).

[13] *Fiedler, Herbert,* Modell und Modellbildung als Themen der juristischen Methodenlehre (Model and modeling as subjects of legal methodology). In:. Erich Schweighofer et al., e-Staat und e-Wirtschaft aus rechtlicher Sicht, Tagungsband IRIS2006 (e-government and e-business from a legal perspective, Proceedings IRIS2006), pp 275-281 (2006).

[14] *Zeleznikow, John Hunter, Dan,* Building Intelligent Legal Information Systems, Representation and Reasoning in Law, Computer Law Series 13, Kluwer, Deventer (1994)

[15] *Yearwood, John, Stranieri, Andrew (eds.),* Technologies for Supporting Reasoning Communities and Collaborative Decision Making: Cooperative Approaches, IGI Global Publishers, Hershey, PA 2011.

[16] *Schweighofer (1999),* FN. 6.

integration of methods of information retrieval, AI & law, legal language processing but also using all methods of data analysis useful for the legal discipline.

In the knowledge representation of law, it is not just about the documentation itself; each view represents further insights on the law itself. The use of all available media and methods – language, meta-knowledge, visualisation, structure, mathematics and statistics, logic, ontologies, formal structures, etc. – implies also that legal "knowledge sources" must be integrated into this knowledge structure, however, keeping the holistic and harmonious picture. Language, visualisation, structure, etc. are also essential methods of human thinking itself. The principle applies that more quantity of methods and analysis ultimately results in higher quality. The more comprehensively all media, factors and methods are used, the better is the structural analysis of the law.

3.1.1. Text corpus

The text corpus (Lu/Conrad: textually set of evidence) consists of all relevant documents of the legal system: statutes, regulations, court decisions, literature, administrative practice, as well as complementary legal practice such as briefs or draft legislation. The basis is the triad of norms, court decisions and literature, however, to be extended by the now huge body of "soft law." It has to be noted that the triad constitutes a simplification: practically there are hundreds of types of documents in any jurisdiction.

3.1.2. Meta data view

The description view contains the metadata of the law (Lu/Conrad: annotation view). Information science has developed a good methodology to describe these materials in order to classify and summarize important content or extract relevant information. Starting from the traditional library science, descriptive metadata is generated manually, semi-automatically or automatically.[19] The quality of this description is highly variable, ranging from the high-level West's Key Number System[20] or the formidable CELEX metadata system[21] to the almost negligible metadata in some commercial law information systems, but also in the applications of AustLII[22]. The meta data view itself is only useful if the user also understands the meta data. Thus, it is important to find an intelligent use of meta data in a “Google-like” search. Meta data can set the proper context of the communication, e.g. persons, roles, purpose, topic etc. that is considered as a prerequisite by the users and thus not said explicitly.

3.1.3. 3. Citation network

The citation network (Lu/Conrad: citation network view) describes the relations between the documents. It has always been a key topic of legal documentation and is and will remain indispensable. Formally, it is documented if a document cites others ((out-bound (cited) sources)

---

[17] The ICAIL proceedings are published with ACM Publishers, New York and available in the ACM Digital Library. Latest edition: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law, Rome (2013); http://dl.acm.org/citation.cfm?id=2514601 (last accessed: 15 May 2015).

[18] Website listing JURIX proceedings: http://jurix.nl/proceedings/ (last accessed: 15 May 2015).

[19] Schweighofer (1999) FN. 6.

[20] Westlaw, West Key Number System® on WestlawNext®, https://info.legalsolutions.thomsonreuters.com/pdf/wln2/L-374484.pdf (last accessed: 15 May 2015).

[21] *Schweighofer, Erich*, Wissensrepräsentation in Information Retrieval-Systemen am Beispiel des EU-Rechts, Dissertation, Universität Wien 1995, Drucklegung in erweiterter Fassung (Anhang Neuerungen Datenbanken bzw. XML), WUV, Wien (knowledge representation in information retrieval systems on the example of EU law, PhD thesis, University of Vienna 1995, published in extended version WUV publishers, Vienna (2000).

[22] AustLII website: http://www.austlii.edu.au/ (last accessed: 15 May 2015).

and if it is quoted by others (in-bound (citing) sources)). Nowadays, not only the document as a whole but also its structuring elements, e.g. articles, sections, paragraphs, lists, etc. should be cited The citations are very different; the main types are: basis of the act, cited acts in the document, citations in the operative part of the judgment, document amending other documents, document is amended by other acts, etc.[23]

3.1.4. User view

The user's perspective takes into account the opinions of legal professions, business and civil society about the document collection (Lu/Conrad: user view). Modern search engines can (semi)automatically generate user's perspectives. Based on the results of a search, the search engine generates information how often a particular document has been opened for viewing, how many times it has been printed, how many times it has been tested for its legal validity, etc. Thus, although only the main usage is analysed, information is given about the document relevancy that is otherwise difficult to obtain. Sufficient data protection has to be taken into account as it is in the practice of Westlaw. Only if sufficient quantities of data on user perspectives are generated then the data is used as then it is considered to be sufficiently anonymised .[24]

3.1.5. Logical view

The logical representation is based on the first order logic, complemented by a representation of temporal layers[25] and the personal scope[26]. The decisive factor is the rapid and efficient execution of an at least five-digit number of rules (i.e. more than 10,000 rules). These requirements have been met in practice but are still insufficiently used.[27] Logic programming restricts the complexity of the logic view. Thus, deontic logic tries to improve the theoretical basis for advanced logical representations. We share the position of *Robert Kowalski*[28] that logic programming is sufficiently powerful to be highly useful in legal environments. A co-operation between man and machine should be established leaving “hard cases” to lawyers.

Since many years, a core task of AI & law was the development of an appropriate legal logic, so far with no resounding success. Now, this line of research focusses more on the representation of legal argumentation. However, an important factor is very often overlooked. The best legal logic is useless if it cannot be converted into a powerful logic programme. Here a wide gap between the theory of legal logic and logic programming has emerged since the 1990s. This gap may not be closed in the near future. Thus, the logical view is restricted to “standard cases” leaving “hard cases” to the argumentation view.

3.1.6. Ontological view

The conceptual logical representation can draw on the experience of conceptual jurisprudence that has intensively dealt with the conceptualization and systematisation and was dominant in the 19th

[23] *Albrecht Berger* has written a classic but little-known standard work on citations: *Berger, Albrecht,* The development of references in the legislative documentation, publishing documentation, Pullach near Munich (1971).

[24] Presentation by Jack Conrad at the OCG digital2014 Conference in November 2014 in Vienna (Skype participation).

[25] See. the PhD thesis of Johannes Scharf (2015, forthcoming) and *Scharf, Johannes*, rOWLer – A hybrid rule engine for legal reasoning. In: Schweighofer, Erich et al., Co-operation, Proceedings of the 18th IRIS Conference, Salzburg, OCG Publishers, Wien 2015.

[26] Idea of Erich Schweighofer, based on relevant thoughts of *Chris Reed*, You Talkin 'to Me ?. In: Jon Bing, en hyllest, a tribute, p 154-171 (2014).

[27] A prime example is the Australian company SoftLaw; this was subsequently acquired by Oracle; the application itself is available as Oracle Business Rules.

[28] *Sergot, M., Sadri, F., Kowalski, R., Kriwaczek, F., Hammond, P., Cory, T.,* The British Nationality Act as a Logic Program, Comm ACM, Vol. 29, No. 5, pp 370-386 (1986).

century. The conceptual jurisprudence is a tool of thought for a better understanding of the legal system. The legal ontology, however, brings the automation option and a much more powerful systemisation. In case of an ontology, a much easier further analysis of the legal system can be made. In the 1990s, ontologies as a conceptualization of a domain have been recognized as the way the reorganization of the AI research[29], also in AI & law. Many legal ontologies have been developed[30] (see for an early analysis Bench-Capon/Visser 1997[31]). Now, the Universities of Amsterdam and Bologna are setting the standards with LRI Core, LKIF[32] and LegalRuleML[33]. Much research has been reused for a more convenient representation of legal knowledge.[34] Tasks of legal ontologies are conceptual information retrieval, knowledge representation, the multilingual search or exchange of information and knowledge.

The theory development of legal ontologies and also its practical implementation are sufficiently strong, however, in practice, few applications exist. Legal ontologies require a deep analysis of the legal domain that can only be done by the legal expert. So far, very few experts have moved into this new form of representation of knowledge. Thus, legal ontologies share the fate of the Semantic Web[35]. Semantic mark-up makes the text and intelligent computer-usable. Thus, a strong tool for a co-operative analysis of legal text corpora is available. However, it needs to be worked on a more efficient collaboration between man and machine in order to scale-up the application domains.

From the point of view of legal complexity, ontologies bring many benefits and flexibility. The terminology is considerably sharpened and structured and can be applied to any further computer-based applications. Comprehensive knowledge representations such as commentaries can thus be done electronically and can be updated with the much more effective means of (semi-) automated data analysis. The main advantage of legal ontologies consists in the available technical standards and the flexibility of the view. Advanced thesauri, concepts, links, rules, etc. can be displayed easily.

3.1.6.1. Legal concept ontology

Starting point of any legal ontology is the terminology of the law. Since the 19th century substantial preparatory work has been done in concept jurisprudence. Re-using this work, the respective elements of the concepts have to be transposed into a computer-readable structure, e.g. header, definition, relations (upper/lower term, antonym, related term, synonym, homonym, polysem etc.), pre-subsumption (relation between the concept of law and facts element) and comments. This

---

[29] *Gruber, T. R.,* The Role of Common Ontology in Achieving Sharable, Reusable, Knowledge Bases, in: J. Allen et al. (Eds.), Principles of Knowledge Representation and Reasoning, Proc of the Second International Conference, KR'91, Cambridge, MA, Morgan Kaufmann Publishers, San Mateo, CA, S. 601-602 (1991)..

[30] *Sator, Giovanni, Casanovas, Pompeu, Biasiotti, Maria Angela, Fernández-Barrera, Meritxell,* Approaches to Legal Ontologies: Theories, Domains, Methodologies, Dordrecht/Heidelberg/London/New York, Springer (2011).

[31] *Visser, P. R. S., /van Kralingen, R. W./Bench-Capon, T. J. M.,* A Method for the Development of Legal Knowledge Systems, in: Proceedings of ICAIL 1997, 151-160 (1997).

[32] *Hoekstra, R., Breuker, J., De Bello, M., Boer, A.,* The LKIF Core Ontology of Basic Legal Concepts. In: P. Casanovas, M.A. Biasiotti, E. Francesconi, M.T. Sagri (eds.) Proceedings of LOAIT 07, II. Workshop on Legal Ontologies and Artificial Intelligence Techniques, S. 43-64. http://www.ittig.cnr.it/loait/LOAIT07-Proceedings.pdf (2007).

[33] Website OASIS LegalRuleML, https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legalruleml (lastly accessed: 20 May 2015).

[34] *Casellas, N., Francesconi, E., Hoekstra, R., Montemagni, S. (Eds.),* Proceedings of LOAIT 2009, 3rd Workshop on Legal Ontologies and Artificial Intelligence Techniques joint with 2nd Workshop on Semantic Processing of Legal Text. Barcelona: IOT Series; *Casanovas, P., Biasiotti, M. A., Francesconi, E., Sagri, M. T. (eds.),* Proceedings of LOAIT 07, II. Workshop on Legal Ontologies and Artificial Intelligence Techniques, pp. 43-64. http://www.ittig.cnr.it/loait/LOAIT07-Proceedings.pdf (2007).

[35] *Berners-Lee, Tim et al.,* The Semantic Web. Scientific American Vol. 284, No. 5, pp 34-53 (2001).

structure is in a sense the basis of the legal concept ontology and is available for other applications such as the Dynamic Electronic Legal Comment (DynELC).[36]

3.1.6.2. Facts – world ontology

The aim of the project Cyc is to provide automated applications with a formal knowledge base of "common sense" knowledge. Currently, more than three million facts and rules that are formally represented in the Cyc knowledge base, provided as OpenCyc.[37] The enormous potential of the Cyc knowledge is still in the experimental stage. In the list of possible applications of Cyc project, law is not specifically mentioned.

3.1.6.3. Pre-subsumption: linking legal concepts with facts elements

A major advantage of an ontology constitutes the easier representation of relations between facts and legal concepts in a form of pre-subsumption. In many cases, the subsumption can therefore be left mostly to the machine; of course it requires human control as well as the use of appropriate controls for a proper deployment of simplified legal mechanisms.

3.1.7. Legal visualization view

Following *Collete R. Brunschwig*, legal visualization concerns the use of graphics, images and videos for visual representation of the law.[38] Today, it is regarded as a branch of multisensory law.[39] Visualisation draws attention to the essentials, makes these memorable, increases understanding and shows hidden connections. In the context of computer science and law, *Ralf Knackstedt* and *Marcel Heddier* link the research on visualization in business informatics with that of legal informatics. Their particular goal is the visual representation of law in operational information systems.[40]

The potential of visualizations for dialogue with the laity is undisputed. But graphical notations are also a support for the formalization of the law. The key features are represented by images or graphics, even in cases where the necessary level of abstraction for formalization is not yet reached. A good example exists in the work of *Tobias Mahler* on the risk analysis of contracts.[41]

For ontologies, graphical representations have proved to be useful in the description of the relationships of the respective descriptors. The notation is not very complex and the advantage in

[36] For a detailed description see *Schweighofer, Erich,* Indexing as an ontological-based support for legal reasoning. In: John Yearwood and Andrew Stranieri (eds.), Technologies for Supporting Reasoning Communities and Collaborative Decision Making: Cooperative Approaches, IGI Global Publishers, Hershey, PA 2011, S. 213-236 (2011).

[37] *Wikipedia EN*: http://en.wikipedia.org/wiki/Cyc (last accessed: 15 May 2015).

[38] *Brunschwig, Colette*, Multisensory Law and Legal Informatics – A Comparison of How these Legal Disciplines Relate to Visual Law. In: Anton, Geist, Colette R. Brunschwig, Friedrich Lachmayer, Günther Schefbeck (Hrsg.), Strukturierung der Juristischen Semanik – Structuring Legal Semantics, Festschrift für Erich Schweighofer, Weblaw Verlag, Bern, p. 573 et seq. (2011); *Brunschwig, Colette*, Visualisierung von Rechtsnormen, Legal Design, Diss. Zürich*; Lachmayer, Friedrich*, Visualisierung des Abstrakten, in: Erich Schweighofer et al., (Hrsg.), IT in Recht und Staat, Aktuelle Fragen der Rechtsinformatik, Wien: Band 6 der Schriftenreihe Rechtsinformatik, p. 309 (2004).

[39] *Brunschwig, Colette R.*, Multisensory Law and Legal Informatics, FN 36, p. 633 et seq. (2011).

[40] *Knackstedt, Ralf, Heddier Marcel,* Herausforderungen der Rechtsvisualisierung aus Perspektive der Wirtschaftsinformatik (Challenges of legal visualisation from the persepctive of business informatics). In: Erich Schweighofer, Franz Kummer, Walter Hötzendorfer (Hrsg.), Transformation juristischer Sprachen, Tagungsband IRIS2012, p. 355 et seq. (2012); ---, Empirische Evaluierung von Rechtsvisualisierungen am Beispiel von Handyverträgen (Empirical study of legal visualisations on the example of mobile phone contracts). In: Schweighofer, Kummer, Hötzendorfer (Hrsg.), Applikation und Applikation, Tagungsband IRIS2013, p. 413 et seq. (2013).

[41] *Mahler, Tobias,* Visualising Legal Risk. In: Schweighofer (Hrsg.), Semantisches Web und Soziale Netzwerke im Recht, Tagungsband IRIS2009, books@ocg.at, Wien, p. 315 et seq. (2009).

the representation is obvious. However, for more complex ontologies, it is necessary to consider a more complete visualization.

3.1.8. Legal argumentation view

It should be noted that the focus is not only on argumentation but also on the representation of the structure of each document in its elements and logical structure: factual information, evidence, arguments, conclusions, etc. The core is of course the legal argument, as it includes all these elements.[42]

An emphasis has been placed on the formalization of argumentation in the 1990s, after the modest success of legal logic. A fundamental work from a theoretical point of view is the book of *Robert Alexy* with the theory of legal argumentation.[43] Arguments (judgments, briefs, drafts, etc.) are represented in a hidden form. Formalisation means that these elements are systematized and placed in a logical argumentation structure. Arguments are an essential tool of legal work, because both thesis, antithesis and synthesis are relevant in each case by the predetermined dialectical roles of plaintiff, defendant and judge.

3.2. Methods

3.2.1. Searching, reading, interpreting, understanding

The basic methodology is to locate, read, interpret and understand the "legal stuff", taking into account the legal interpretation and reasoning methods in a dynamic world of concepts. This simplistic description of the legal doctrine from the perspective of legal informatics is to some extent only auxiliary in order to integrate the results of the legal doctrine into this model. Other missing elements are also considered, e.g. social context, legal authorities, sophisticated methods of interpretation etc.[44]

The most significant add-on of legal informatics is the revolution in legal search by the use of search engines. Nowadays, it is no longer a problem to search in micro seconds gigabytes of relevant documents according to predetermined similarity criteria. "Legal Googling" now belongs also to the recognized methods of finding relevant documents.

3.2.2. Legal documentation & search

Due to the abundance of the material, legal documentation has become its own independent method. However, this task is no longer done primarily by the users themselves, but legal information providers offer such services as a free or pay service. These providers follow the negotiation and adoption of new laws, rendering of new judgements or decisions, new literature etc. They take care of all relevant changes of the particular legal system, index these materials in full-text and, if possible, add metadata and provide powerful search engines. Naturally, the systems are online 7 days/24 hours.

[42] *Ashley, Kevin D.,* Modeling Legal Argument. Reasoning with Cases and Hypotheticals. Cambridge MA: MIT Press (1990).

[43] *Alexy, Robert,* Theory der juristischen Argumentation (Theory of Legal Argumentation). Frankfurt am Main (1983).

[44] Previously, the search was focussed on reading ("going into the books"), if necessary with the use of metadata. The "google-like search" needs to take advantage of meta data as the users knows little about metadata and does not master sufficiently the legal vocabulary. Newly designed legal information systems will hopefully include more semantic search.

In addition to these basic services, the documentation of the ongoing structural analysis of law by legal experts is essential. Without this analysis, the abundance of cases cannot be handled quickly and inexpensively.

3.2.3. 3. Conceptual and logical analysis

For the conceptual and logical analysis, the fundamental statement of *Sowa* applies again: the terminology is to be developed and to be brought into a convenient logical structure.

There are many ways of how this can be done. A quite useful example of logical analysis involving some concepts are the works of *Kahlig, Heindl* and *Stingl. Kahlig* has worked on logic programmes for a long time in the field of tenancy; now additions to the tax law have been also developed.[45] *Heindl* and *Stingl* are recognized legal experts in their field. Using this co-operation, very useful practical flow charts (activity diagrams) are created with logical structure. This only covers legal standard cases; disputes shall be handed with legal argumentation models. In combination, a logic programming approach to a highly complex area of law is necessary. The logic programming is powerful enough to be a decisive support tool for the implementation of laws. Well-structured flowcharts help in decision-making, esp. when thousands of legal rules must be processed.

However, two important elements should be added. Without a model of the layers of time use in a highly dynamic legal world is difficult; further, a differentiation should be made according to distinct applicability of norms on legal subjects.[46]

3.2.4. Factual elements and its links to law

The world as such and their description is considered as a "given" by the legal profession. However, this is not trivial as is clearly demonstrated by an analysis of a specific case. Company lawyers, attorneys, administrative and judicial authorities use a large part of the work for determining the legally relevant facts. In practice it is often argued about the existence or non-existence of elements of the situation. Therefore, it is very helpful to make use of existing world ontologies, because there is already a lot of basic work has been done there. The use of powerful modelling languages like UML or SysML should be also considered.[47]

In legal practice, the "Rechtssatz" (legal rule) effectively combines the elements of the situation with the conclusions of law.[48] These works should be included in the development of the elements of the situation.

In the knowledge society, the automated generation of factual elements from pictures, videos, website, intelligent forms etc. is of particular importance. Successful practice consists in tax law as well as in the e-Justice with intelligent forms. Another field of application is the interpretation of images (esp. fining of violations of speed violations in traffic).

---

[45] *Kahlig, Wolfgang,* Rechtsvisualisierung mit C.O.N.T.E.N.T. In: Erich Schweighofer, Franz Kummer, Walter Hötzendorfer (eds.), Co-operation, Proceedings of the 18th International Legal Informatics Symposium IRIS 2015. OCG Publishers, Wien, pp. 425-442 (2015).

[46] Such a model was developed by Johannes Scharf as part of his PhD research.

[47] *Weilkiens, Tim,* Systems Engineering with SysML/UML: Modeling, Analysis, Design. The MK/OMG Press (2008).

[48] *Handstanger, Meinrad,* rule sets as a cooperative type of text. In: Erich Schweighofer, Franz Kummer, Walter Hötzendorfer (eds.), Co-operation, Proceedings of the 18th IRIS Conference, OCG Publishers, Wien, pp. 391-398 (2015).

3.3. 3. Synthesis

The methods of synthesis of these results are various. Here, four main methods are singled out: commentary or manual, Dynamic Electronic Legal Commentary (DynELC), citizen information system and case-based synthesis.

3.3.1. Handbook

In the knowledge and network society, legal informatics has now a major task in legal science: the essential support of the analysis and synthesis of the law. The text corpus results in an abundance of norms, positions and opinions; these views can be properly structured in a reasoning catalogue that is helpful for the claimant, the opponent and the judge. At the moment, such handbooks are written traditionally. Authors read all materials and try to develop a well-structured analysis of this legal domain. Due to the strong dynamics of the legal system, this task is getting more and more difficult. Legal publishers also note that the depth of analysis is diminishing due to time constraints. Authors favour a more documentary approach with extensive notes.

3.3.2. Dynamic Electronic Legal Commentary (DynELC)

The legal commentary is the appropriate form of representation of the knowledge of a legal system offering in a systematic analysis all relevant elements for a comprehensive and holistic understanding of the particular area of law. This form of a commentary takes advantage of the options of legal data science and using all big data methods for analysis of the text corpus.

Doing a synopsis of these 8 views and 4 methods with traditional legal methodology, it can be described as a process of interpretation of a text corpus and its synthesis. The methods are present more or less present, but not explicitly mentioned or methodically strongly developed for the legal environment. Much experience exists in computer science and practice that can be further developed.

For a long time, we are proposing a change from the traditional system to an electronic commentary.[49] So far, the theoretical foundation has been developed that can be now applied practically.[50] The idea of dynamic electronic legal commentary (DynELC) is simple. On the one hand, the 8 views and 4 methods are presented in a structured format to create a simplified synthesis, on the other hand, the basis for further analysis by legal experts is provided. The following main methods are available:

- Classification: document categorization
- Thesaurus: semi-automatic generation of thesaurus descriptors; also considering multilingual jurisdictions (i.e. European Union, Finland, Belgium, Switzerland, Canada)
- Citations: automatic generation of hypertext links
- Temporal relations: the automatic generation of temporal relationships
- Ranking: Document vs. requirements, document in the corpus, document in the citations network, document in the timeline etc.
- Text summary: semi-automatic generation of summaries of documents
- Multilingualism: an automatic translation of documents (e.g. Google Translate)

[49] These efforts are unsuccessful to date; because the resources were not available or no suitable application area could be found.

[50] See. in detail *Schweighofer, Erich*, Indexing as an ontological-based support for legal reasoning. In: John Yearwood and Andrew Stranieri (eds.), Technologies for Supporting Reasoning Communities and Collaborative Decision Making: Cooperative Approaches, IGI Global Publishers, Hershey, PA 2011, S. 213-236 (2011).

The DynELC consists of a structured representation of the metadata and the text corpus. A major advantage is that the dynamic factor of law is taken into account. New developments can be easily included in this commentary. Further, the different layers of time can be properly represented. In addition to, the commentary as such can be adjusted by the user as the analysis parameters are sufficiently flexible.

3.3.3. Citizens Information

Citizens Information Systems use the internet to spread easily understandable public information. Focus is naturally on legal information, authority structure and citizen participation. The systems use standard technology; its clear advantage lies in the citizen-focussed description of the legal system.[51] Legal data science helps to adjust these systems to the dynamics of the law but also generate a more citizen-focused analysis.

3.3.4. Case-related synthesis

The commentary or system as well as the citizens' information are syntheses of the legal system for all. In practice, the case-based synthesis for each specific case is crucial. All existing sources and syntheses should be used to best present their own legal position in relation to key legal authorities. A further distinctive feature is the strong argumentation view. All relevant arguments and counterarguments are presented in its logical structure, however, taken into account the respective view of the claimant, the defendant or the judge.

4. Conclusions

Legal data science looks back to more than 50 years of experience in the analysis of legal text corpora (albeit under different names, in particular legal information retrieval or AI & law). Further, the long tradition of legal theory research provides some guidance. The methodology of legal data science is multidisciplinary, but, respecting the principle of *Hans Kelsen*, no syncretism should occur. Thus, the results are developed separately and only brought together by a particular synthesis with isomorphism. This helps to carry out a rapid and convenient adjustment of an assessment respecting ongoing changes in the legal system.

The analysis consists in creating 8 views and using 4 methods for the analysis. A major result would be the Dynamic Electronic Legal Commentary (DynELC) consisting in a structured representation of the metadata and the text corpus. This commentary offers also a personalised view of the legal system depending on the needs of an individual case.

The stronger use of computational methods speeds up legal analysis and allows a faster and more efficient assessment of the legal materials. The methods exist, are used every day in practice, but still need more adjustment to law. Thus, they are still not sufficiently optimized. At the moment, a transitional period can be noticed characterised by semi-automation. Computers can be very supportive but in the very end a human being has to check the proper context and quality of the work. A fruitful cooperation between men and machines has to be developed in order to deploy and improve the methodology of legal data science.

[51] *Krenmayr, Andreas, Traunmüller, Roland*, public information systems - new ideas. In: Erich Schweighofer, Franz Kummer, Walter Hötzendorfer (eds.), Co-operation, Proceedings of the 18th International Legal Informatics Symposium IRIS 2015. OCG Publishers, Wien, pp. 227-234 (2015).

O CONHECIMENTO DOS FATOS NO CONTEXTO PROCESSUAL: AS LIMITAÇÕES DA ABORDAGEM PROBABILÍSTICA NO CONTEXTO JURÍDICO

KNOWLEDGE OF FACTS IN PROCEDURAL CONTEXT: THE LIMITATIONS OF THE PROBABILISTIC APPROACH IN THE LEGAL CONTEXT

Simone Trento

Universidade Federal do Paraná
simonetrento@yahoo.com

Descriptors: ***Knowledge, process, probabilistic approach, analysis of evidence***

Extended English Abstract.

This paper discusses the limits of probabilistic approach in the analysis of evidence.

Quantitative probabilistic approaches are not capable of consistently take care about this issue, not just because of pragmatic lack of statistics, but also because of the complexity of judicial proof and because of the claim, in judicial context, for individualizing the treatment of the case.

The facts are established in judicial process just because this is necessary for deciding the case. Therefore, knowing facts in judicial process is something instrumentally oriented for deciding the case. Of course, deciding the case according to Law reflects the authority of Law. Therefore, deciding according to the true is the institutional finality of proof activities.

The logic does not help to establish what is the applicable standard of proof, but it can help reach a conclusion logically based on whether or not filled the applicable standard.

Much of the complexity of the decision on the evidence of the facts in the procedural context comes from peculiarities of this context. One of these peculiarities is that what is strictly proved in the judicial process interacts with other elements: incontrovertible facts and background knowledge, that are admitted to the procedure for reaching conclusions from the evidence.

In most legal systems today, including the Brazilian, there is explicit legal demand for motivation of judgments. Thus, results clear that the proof understood as the result of evidential activity demand not only the presentation of the reasons that led to the delivery of that decision in those terms, but not only that these reasons are listed. The duty of motivation requires favorable reasons and contrary to that decision are articulated so as to result justified that was that - and no other - the decision.

Another of the peculiarities of the procedural context of knowledge of the facts is that judicial adjudication is not just epistemologically oriented. There are also limitations coming from legal determinations (excluding some pieces of evidence, predeterminations of the evaluation of other pieces of evidences) and from pragmatic considerations (giving some time limits and demanding, even with no concluding evidence, a solution to the case).

This is why logic is a part, important, no doubt, of evidential activity, but only a part of it, that, in addition to logic, operates with regard to legal rules of procedural law and substantive law.

The rationalist conception of proof, when dealing with proof as a result of the analysis of the support that pieces of evidence give (or not) to each other, demand a conception of argumentative probability, able to explain to what extent a piece gives or withdraw support to the other and to what extent the hypothesis are supported by the whole evidence produced.

The function of logic therefore is to clarify how the inferences made from the earliest assumptions to the latest conclusions adopted are properly sustained from the logic point of view and how the pieces of evidence can be understood coherently. And if not all pieces of evidence can be understood coherently, logic can point which conflicting pieces of evidence are not consistent with the explanatory hypothesis adopted.

In the end, though, if (and why) the parts discarded do not deserve credit and if (and why) the coherent parts together deserve credit and make the event an acceptable description of the facts is a decision about which logic is not decisive, but rather the law.

The main disagreements in this endeavor are not based on logic error in the argument, but on different weights that are assigned to different factors taken into consideration.

Quando se trata da análise de evidência no contexto processual, se está a tratar da suficiência ou não da prova para o fim de se decidir o processo de forma a admitir como estando ou não provados os fatos objeto de decisão.

Mas o conhecimento dos fatos no processo, como está instrumentalmente (DINAMARCO, 2000) direcionado à decisão da causa levada a juízo, não é um conhecimento puro e simples. Ou seja, não se trata de um conhecimento que tenha as finalidades aplicáveis à pesquisa científica ou histórica. Não se trata do saber pelo saber, mas do saber para decidir uma causa jurídica. Trata-se, portanto, de um conhecimento funcionalmente orientado à tomada de decisão.

Na medida em que o processo não se satisfaz em decidir uma causa de uma forma qualquer, mas sim de acordo com a disciplina dada pelo direito material, é de enorme importância que a decisão – para se adequar à disciplina dada pelo direito material – se baseie no modo como os fatos efetivamente se passaram na realidade histórica. Erros acontecerão: dadas as limitações dos seres humanos tomadores das decisões e as limitações impostam pela própria ordem jurídica, processos serão decididos de forma equivocada, baseando-se em equívocos a respeito do modo como os fatos se passaram ou até mesmo em acertamento dos fatos de forma conscientemente sem correspondência com a realidade (é o que se dá, por exemplo, quando o ordenamento jurídico determinada a exclusão de uma prova que conduziria à verdade na fixação dos fatos). Porém a medida em que esses erros possam ser minimizados é o retrato da efetividade de uma determinada ordem jurídica. Com efeito, um ordenamento jurídico que abrisse mão de aplicar as consequências jurídicas (aplicáveis em abstrato) aos fatos efetivamente ocorridos simplesmente colapsaria (FERRER BELTRÁN, 2013, p. 31).

Muito da complexidade da decisão acerca da prova dos fatos no contexto processual advém de peculiaridades deste contexto. Uma destas peculiaridades é a de que aquilo que é estritamente provado nos autos interage com outros elementos: elementos que são tidos como comprovados independentemente de prova em sentido estrito (fatos incontroversos e fatos notórios) e elementos que provém de generalizações (regras da experiência) que são admitidas no processo para o fim de se permitir a retirada de conclusões mais avançadas a partir dos meios de prova levados ao processo e daqueles fatos que nem mesmo precisavam ser estritamente provados.

Outra dificuldade é que a palavra "prova" tem mais de um sentido: diz-se "prova" para se referir (1) à fonte de prova, (2) à atividade probatória e (3) ao resultado probatório (GASCÓN ABELLÁN, 2012, p. 30 e 33). Quando se fala a respeito especificamente do *standard* de prova, aquele umbral mínimo de prova a ser produzido para que se possa concluir suficientemente provado um fato (para o fim de se decidir o processo como se tal fato houvesse efetivamente acontecido, ou seja, tomá-lo como ocorrido), se está a tratar do resultado probatório (3).

Na maior parte dos ordenamentos jurídicos da atualidade, inclusive no brasileiro, há exigência expressa de motivação das decisões jurisdicionais (a não ser excepcionalmente, como ocorre nas decisões proferidas pelo Júri). Sendo assim, resulta patente que a prova entendida como *resultado* da atividade probatória demanda não só a apresentação das razões que conduziram à prolação daquela decisão naqueles termos, e não só que essas razões sejam elencadas. O dever de motivação exige que as razões favoráveis e contrárias àquela decisão sejam articuladas de forma a resultar *justificado* que tenha sido aquela – e não outra – a decisão tomada.

Sendo assim, há que se verificar qual teoria de "probabilidades" pode resultar adequada e este contexto de justificação (MACCORMICK, 2005, p. 271) do *resultado* da atividade probatória.

O saber probatório necessário e suficiente para se decidir uma causa jurídica favoravelmente àquele que tem o ônus da prova é aquele capaz de preencher o que se chama *standard* de prova.

Na tradição anglo-americana foram sendo assentados jurisprudencialmente os *standards* de prova aplicáveis a diferentes tipos de controvérsias (HAACK, 2012, p. 211-212), de modo que como regra geral se utiliza a preponderância da prova para processos cíveis e prova além de toda dúvida razoável (*beyond resonable doubt*) para causas criminais; fala-se também em um *standard* intermediário para causas cíveis cujos erros seriam de maior potencial danoso (*clear and convincing evidence*) e de *standards* inferiores à preponderância da prova para decisões não definitivas justificadas pela necessidade de uma medida (tal como uma busca e apreensão).

Na tradição do *civil law* não é comum o uso de fórmulas para o estabelecimento dos *standards* de prova nem na legislação nem, muitas vezes, na prática forense, embora o estudo do direito comparado tenha gerado nos últimos anos uma grande repercussão das fórmulas designadoras dos *standards* de prova próprias da Inglaterra e dos Estados Unidos na doutrina dos países de tradição romano-germânica.

De qualquer sorte, quando a decisão de qualquer caso está a depender da análise da prova, há de se decidir se os elementos de juízo reunidos ao longo da instrução processual são ou não suficientes a preencher o mínimo (exigível na espécie) de "probabilidade" de que os fatos em que se funda a pretensão efetivamente tenham ocorrido. No contexto processual, não se pede ao juiz que afirme se o fato objeto da prova aconteceu, mas se o fato apurado pode – a partir da prova admitida – ser estabelecido segundo o *standard* de prova aplicável (HAACK, 2012, p. 214).

Entre o desconhecimento total e o conhecimento da verdade absoluta há vários graus, conforme os elementos de conhecimento a favor da hipótese de que a descrição corresponde à realidade (TARUFFO, 2009, p. 180). Há de se verificar, então, de que tipo de "probabilidade" se está falando neste contexto.

Teorias de probabilidades quantitativas (aplicadas ao contexto da prova judicial) enfrentaram críticas que colocam em questão sua factibilidade e correção. As limitações são diversas (v. p. ex. TARUFFO, 2009, p. 190-223): nem sempre se dispõe de estatística para aquilo que se está a investigar; muitas das teorias probabilísticas (Bayesianismo, p. ex.) partem de *prior probabilities* muitas vezes supervalorizadas – além de possivelmente arbitrárias – e que inviabilizam uma adequada valoração da prova que vem a ser reunida no processo; quando se usam probabilidades quantitativas, os *standards* de prova (por uma questão de coerência, para se concluir se aqueles X% exigidos pelo *standard* aplicável foram supridos pelo material probatório) são também expressos em formato numérico, o que sói ser arbitrário; quando se calcula a probabilidade de que dois eventos independentes tenham ocorrido, o resultado matemático da ocorrência provável de ambos simultaneamente – numericamente – diminui, o que pode ocasionar uma diminuição excessiva (uma vez que, por exemplo, 0.7 x 0.7 resultaria em uma frequência estatística de 0.49, ou seja, inferior ao *standard* normalmente exigível para causas cíveis, ao menos superior a 50%) [ALLEN, 2009, p. 322-323]; teorias quantitativas normalmente não consideram a suportividade (HAACK, 2013, P. 78) que um elemento de juízo pode conferir a outro.

Não há de se duvidar que dados estatísticos podem ser de grande importância para a definição dos fatos no âmbito processual. Tanto assim que é amplamente admitida como praticamente cabal a proba estatística da probabilidade de paternidade no processo em que se demanda o reconhecimento de paternidade. Mas isto só é assim porque neste caso o suporte do resultado de exame de DNA à hipótese de paternidade é enormemente alto e este suporte não é colocado – normalmente – em questão por prova que mereça confiabilidade à altura daquela providenciada com o exame de DNA. Além disto, a estatística a que se chega no exame de DNA não provém de frequências gerais de

classes de eventos em que simplesmente se pudesse enquadrar o caso em discussão, mas sim da fração de possibilidade específica – matemática – de que o investigado seja pai do autor. Não é o que ocorre, por exemplo, quando uma demanda indenizatória se funda exclusivamente na estatística nua segundo a qual empresa ré responde por mais de 50% da produção industrial de um determinado produto que, produzido também por empresas concorrentes, causou dano ao autor; neste caso não se admite um salto que vá diretamente da estatística à responsabilização da fabricante, fazendo-se necessária a produção de prova que vá além daquela estatística.

No contexto processual, não importa tanto que haja uma específica probabilidade numérica de que o evento X tenha ocorrido. Se isto puder ser afirmado (e, se puder, é pertinente que tal razão se inclua na fundamentação), tal afirmação terá apenas uma razão de ser instrumental para a finalidade de se poder decidir a causa tomando-se por ocorrido o fato. O que importa é haver justificativa bastante para que se possa sustentar como correta a decisão jurídica que afirmou suficiente ou insuficientemente comprovados os fatos.

O valor probatório de uma conclusão a respeito da prova é a *medida* em que ela é capaz de suportar uma evidência em um determinado contexto (ALLEN e PARDO, 2008, p. 261). Trata-se, portanto, de uma questão de *grau*. Mas esse grau ou medida não corresponde a um percentual de frequência quantitativa verificada historicamente, uma vez que o caso sob julgamento é específico em relação à generalidade dos casos que formam a regra de experiência aplicável e por isso suas especificidades precisam ser *explicadas*. Tanto assim que o valor probatório do exame de DNA não corresponde à percentagem que o perito conclui presente, sendo necessário, ainda, para se chegar ao valor probatório do exame, realizar inferências sobre o que seria capaz de explicar o resultado, considerando-se a possibilidade de erro do laboratório, a possibilidade de fraude manipulada pelas partes ou por terceiros, a possibilidade de que dois hipotéticos investigáveis fossem irmãos etc (ALLEN e PARDO, 2008, p. 263, nt. 133).

A lógica não ajuda a estabelecer qual o mínimo probatório a se fazer presente (ou seja, qual o *standard* de prova aplicável), mas ela pode ajudar a alcançar uma conclusão logicamente fundada sobre se foi ou não preenchido o *standard* aplicável (HAACK, 2012, p. 216 e ss.). Com isso tem-se que a lógica é uma parte, importante, sem dúvida, da atividade probatória, mas apenas uma parte dela, que, além da lógica (HAACK, 2007, p. 16 e 24), opera em atenção ao ordenamento jurídico processual e material.

A tradição jurídica ocidental, ao tratar da *livre* valoração da prova, deu especial acento ao "*livre*", gerando uma concepção *persuasiva* da livre valoração da prova (FERRER BELTRÁN, 2007, p. 62-65), que faz exigências fracas ou inexistentes de motivação acerca da decisão sobre os fatos (a motivação é reduzida à explicação das causas que levaram o juiz a *crer* na ocorrência do fato).

Ao lado dessa concepção persuasiva existe a concepção — hoje em vertiginosa ascensão — *racionalista* da livre valoração da prova, segundo a qual (FERRER BELTRÁN, 2007, p. 62-65): a valoração da prova dá-se com recurso ao método de corroboração e refutação de hipóteses; e há uma forte exigência de motivação. O que justifica a decisão, no modelo *racionalista*, não é a crença do destinatário da prova, mas sim que a proposição sobre o fato em questão esteja corroborada por elementos de juízo.

Esta concepção racionalista da prova, ao tratar do resultado probatório como um resultado da análise do suporte que as fontes de prova conferem uma à outra (ou retiram uma da outra), demanda uma concepção de probabilidade *argumentativa*, capaz de *explicar* em que medida uma peça de evidência dá ou retira suporte à outra e em que medida os fatos objeto de prova estão suportados pela prova produzida.

A função da lógica, assim, é esclarecer como as inferências realizadas a partir das primeiras premissas adotadas até as últimas conclusões a que se chega estão sustentadas adequadamente do ponto de vista da lógica e como as peças de evidência podem ser compreendidas coerentemente. E, caso nem todas as peças de evidência possam ser compreendidas coerentemente, a lógica pode apontar quais são as peças de evidência incompatíveis com a hipótese explicativa adotada.

Ao final, porém, se as peças descartadas não merecem crédito e se as peças coerentes entre si merecem crédito e tornam a hipótese uma descrição aceitável dos fatos é uma decisão a respeito da qual a lógica não é decisiva, mas sim o Direito.

A avaliação completa da suportividade da evidência e da medida em que estão suportadas ("warrant") as demandas requer recursos adicionais de tipo não formal (HAACK, 2007, p. 16-17). As questões relativas à vagueza e indeterminação no direito (e na valoração da prova estas são rotineiramente as questões decisivas, na medida em que se procura estabelecer se a prova produzida é "suficiente") são, no mérito, resolvidas por critérios *jurídicos*. As principais discordâncias nesta seara, destarte, não decorrem de erro lógico na argumentação, mas de diferentes pesos que são atribuídos a diferentes fatores levados em consideração (HAACK, 2008, p. 470) pelo ordenamento jurídico.

O silogismo é apenas um estágio no processo de decidir (PAVCNIK, 1997, p. 498) e de fundamentar a decisão, fornecendo a "moldura" (MACCORMICK, 2005, p. 57) dentro da qual são incluídos outros argumentos, mas essa "moldura" nada diz sobre a qualidade (a força ou a fraqueza) destes outros argumentos.

Quanto mais claro, mais fácil, menos problemático for um caso, tendencialmente mais facilmente a "moldura" silogística será capaz de resolvê-lo por si só. Quanto mais difícil e problemático for um caso, mais a "moldura" silogística haverá de se socorrer de argumentos que vão além do formato lógico-formal e que, em definitivo, serão consequencialistas (MACCORMICK, 1978, p. 163, 225, 241 e 253), no sentido de se adotar a decisão capaz de se apresentar como aquela que é a mais justificada à luz do ordenamento jurídico aplicável.

Bibliografia

ALLEN, Ronald J. Explanationism all the way down. **Episteme**, vol. 5(3), p. 320-328, 2008.

ALLEN, Ronald J.; PARDO, Michael S. Juridical Proof and the Best Explanation. **Law and Philosophy**, 27, p. 223-268, 2008.

DINAMARCO, Cândido Rangel. **A Instrumentalidade do Processo**, 8ª ed., São Paulo, Malheiros, 2000.

FERRER BELTRÁN, Jordi. La prueba es libertad, pero no tanto: una teoría de la prueba cuasibenthamiana. In: VÁZQUEZ, C. (coord.). **Estándares de prueba y prueba científica**: Ensayos de epistemologia jurídica. Madrid-Barcelona-Buenos Aires-São Paulo: Marcial Pons, p. 21-39, 2013

FERRER BELTRÁN, Jordi. **La valoración racional de la prueba**. Madri: Marcial Pons, 2007.

FERRER BELTRÁN, Jordi. **Prueba y verdad en el derecho**. 2ª ed. Madri: Marcial Pons, 2005.

GASCÓN ABELLÁN, Marina. **Cuestiones probatorias**. Bogotá: Universidad Externado de Colombia, 2012.

HAACK, Susan. El probabilismo jurídico: una disensión epistemológica. In: VÁZQUEZ, C. (coord.). **Estándares de prueba y prueba científica**: Ensayos de epistemologia jurídica. Madrid-Barcelona-Buenos Aires-São Paulo: Marcial Pons, p. 65-98, 2013.

HAACK, Susan. On logic in the law: "something, but not all". **Ratio Juris**, Vol. 20(1), p.1-31, 2007.

HAACK, Susan. The Embedded Epistemologist: Dispatches from the Legal Front. **Ratio Juris**, Vol. 25(2), p. 206-235, 2012.

HAACK, Susan. The Pluralistic Universe of Law: Towards a Neo-Classical Legal Pragmatism. **Ratio Juris**, Vol. 21(4), pp.453-480, 2008.

MACCORMICK, Neil. **Argumentação jurídica e teoria do direito**. São Paulo: Martins Fontes, 2006. Tradução de: **Legal reasoning and legal theory**, 1a. ed. 1978.

MACCORMICK, Neil. **Retórica e o Estado de Direito**. Rio de Janeiro: Elsevier, 2008. Tradução de: **Rhetoric and the Rule of Law**, 1a. ed. 2005.

PAVCNIK, Marijan. Legal decisionmaking as a responsible intelectual activity: a continental point of view. **Washington Law Review**, v. 72, p. 481-504, 1997.

TARUFFO, Michele. **La prueba de los hechos**. Tradução espanhola Jordi Ferrer Beltrán. 3ª ed. Madrid: Trotta, 2009. Tradução de: **La prova dei fatti giuridici**. Milano: Giuffrè, 1992.

ANÁLISE DE PRECEDENTES LEGAIS ATRAVÉS DE REDES COMPLEXAS

LEGAL PRECEDENTS ANALYSIS THROUGH COMPLEX NETWORKS

Aline Macohin[1], Cesar Antonio Serbena[2]

[1] Systems Analyst, Federal Data Processing Service. Lawyer.
Carlos Pioli Street 133, 80520-170 Curitiba, BR
aline.macohin@gmail.com

[2] Professor of Philosophy of Law, Law School, Federal University of Paraná-Brazil
Santos Andrade Square 50, 80020-300 Curitiba, BR
cserbena@gmail.com; http://www.ejustica.ufpr.br

Keywords: ***Legal Precedents; Graphs and citation networks; Data Mining; Legal data visualization.***

Abstract: *With the recent digitization of collections of decisions of the Brazilian courts, data visualization techniques can be applied to legal databases, in order to extract knowledge and identify relevant properties, which are not visible without the use of the appropriate techniques of Data Mining. A method already used in countries of Common Law is by modeling a network of citations of judgments by other judgments of the same court or of different courts, in which certain properties are identified and can show the importance of each precedent and how much it is. Although the Brazilian legal system is heavily influenced by Civil Law, precedents have been strengthened as a source of judicial decisions. In the new Brazilian Code of Civil Procedure, approved in March 2015, the precedent was highlighted and there was also a greater linkage of decisions of lower Courts and judges to the decisions of the Superior Courts when the substance of the cases was the same.*

The new Brazilian Code of Civil Procedure identifies the judicial precedent as an essential element of the reasons of any judicial decision (Art. 489, § 1, V, VI, art. 926, § 2, Art. 927, § 5, art. 988, IV, art. 1042, § 1, II). Due to the large number of law suits in the Brazilian courts, the obedience to precedent by judges and lower Courts has become an exigency and at the same time a necessity for the rationalization of the Brazilian judicial system. So far, the evaluation of the impact of a case to be tried by the Superior Tribunal of Justice (STJ) and its fixation as a precedent to be followed by lower Courts was intuitive. The method presented in this study provides a rational way to model the force and effects of precedents in the Brazilian judicial system, which can help the evaluation and improvement of the precedent system in Brazilian Law.

This paper presents an initial analysis of the data network formed by the citations between decisions (a decision cited by another decision) of the Superior Tribunal of Justice (STJ) between 19/09/1996 to 01/03/2015. The main function of the Superior Tribunal of Justice (STJ) in the Brazilian judicial system is to standardize the interpretation and the application of Federal Law by other judges and Brazilian courts. The research database consists of 922,333 judgments. Decisions before 2000

are only available in PDF format and decisions after 2000 are available in PDF and HTML format.

Considering each of the decisions as a node, around 4 million edges were identified (a citation from one Court case to another). In the graph network, when used as a model of networks, properties such as density, degree distribution, clustering coefficient and centrality of the nodes can be identified. As the research database is very large, the authors wish to consider subsets of decisions grouped by a specific subject or by a corpus of specific legislation in order to check how the properties of graphs can be identified and how the weight of the precedents are measured. The authors also wish, in further works, to model other networks from the same STJ database, such as the network of citations of laws, of authors and of legal theories by the decisions of the Superior Tribunal of Justice. Another model to be considered is that of the evolution and decline of precedents over time. Usually the changes of precedents are related to the enforcement of new laws in the legal systems of Civil Law.

Titulo: Análise de precedentes legais através de redes complexas

Palavras-Chave: Precedentes legais; Grafos e redes de citações; Mineração de dados; Visualização de dados judiciais.

Resumo: *Com a recente digitalização dos acervos de decisões dos Tribunais Brasileiros, podem ser aplicadas técnicas de visualização de dados aos bancos de dados judiciais, com o objetivo de extrair conhecimento e identificar propriedades relevantes, as quais não são visíveis sem o emprego de técnicas adequadas de Data Mining. Um método já utilizado nos países de Common Law consiste na modelagem de uma rede de citações das decisões judiciais por outras decisões judiciais, de uma mesma corte ou de cortes diferentes, onde certas propriedades são identificadas e elas podem conferir uma importância maior ou menor a cada um dos precedentes. Apesar do sistema jurídico brasileiro ser fortemente influenciado pelo Civil Law, os precedentes tem ganhado um peso cada vez maior como fonte das decisões judiciais. No novo Código Brasileiro de Processo Civil, aprovado em março de 2015, o precedente ganhou destaque e há igualmente uma vinculação maior das decisões dos juízes de primeira instância às decisões dos Tribunais de Recurso quando a matéria dos casos é a mesma.*

O novo Código Brasileiro de Processo Civil previu, em diversos artigos, o precedente como um elemento essencial na fundamentação de qualquer decisão judicial (arts. 489, §1º, V, VI, art. 926, §2º, art. 927, §5º, art. 988, IV, art. 1042, §1º, II). Devido ao grande número de processos tramitando nas cortes brasileiras, a obediência ao precedente pelos juízes e tribunais inferiores tornou-se uma exigência e ao mesmo tempo uma necessidade para a racionalização do sistema judicial brasileiro. Até o presente momento, a avaliação da repercussão de um caso a ser julgado pelo Superior Tribunal de Justiça e sua fixação como um precedente a ser obedecido pelas cortes inferiores foi intuitiva. O método apresentado no presente trabalho oferece uma forma racional para a modelagem da força e repercussão dos precedentes no sistema judicial brasileiro, podendo servir de auxílio à avaliação e aperfeiçoamento do sistema de precedentes no Direito brasileiro.

Neste trabalho é apresentada uma análise inicial da rede de informação formada pelas citações entre decisões do Superior Tribunal de Justiça, julgadas entre o período de 19/09/1996 à 01/03/2015. A principal função do Superior Tribunal de

Justiça no sistema judicial brasileiro é uniformizar a interpretação e aplicação da legislação federal pelos demais juízes e Tribunais brasileiros. A base de dados da pesquisa consiste em 922.333 decisões judiciais. Parte dos documentos das decisões judiciais estava disponível na linguagem HTML e as decisões anteriores ao ano 2000 estão disponibilizadas apenas em documentos no formato PDF.

Considerando cada uma das decisões como um nó, foram identificados em torno de 4 milhões de arestas (citação de um processo judicial para outro). Em uma rede de grafos, quando utilizados como modelos de redes, podem ser identificadas propriedades como densidade da rede, grau de distribuição, coeficiente de agrupamento e centralidade dos nós. Como a base de dados da pesquisa é muito grande, os autores pretendem considerar sub-conjuntos de decisões agrupadas por uma matéria específica ou por um corpo de legislação específico, a fim de verificar como as propriedades dos grafos podem ser identificadas e o que elas informam acerca do peso dos precedentes. Os autores também pretendem, nos próximos trabalhos, modelar outras redes a partir da mesma base de dados do STJ, como a rede de citações de leis e citações de autores e teorias do Direito pelas decisões do Superior Tribunal de Justiça. Outra modelagem a ser realizada será a evolução ou declínio dos precedentes ao longo do tempo. Geralmente a mudança dos precedentes ao longo do tempo está relacionada à edição de novas leis nos sistemas jurídicos de Civil Law.

Título: Análisis de los precedentes legales por medio de redes complejas

Palabras-clave: precedentes legales; Grafos y redes de citas; minería de datos; visualización de datos judiciales.

Resumen: *Con la reciente digitalización de las colecciones de las decisiones de los tribunales brasileños, las técnicas de visualización de datos pueden se aplicar a las bases de datos jurídicas, con el fin de extraer el conocimiento e identificar propiedades pertinentes, que no son visibles sin el uso de técnicas adecuadas de Minería de Datos. Un método que ya se utiliza en los países de Common Law consiste en modelar una red de citas de las decisiones judiciales por otras decisiones dentro de un mismo tribunal o diferentes tribunales, donde se identifican ciertas propiedades y ellas pueden dar más o menos importancia a los precedentes. Aunque el sistema jurídico brasileño está fuertemente influenciado por el Civil Law, el precedente ha adquirido un peso cada vez mayor como fuente de las decisiones judiciales. En el nuevo Código Brasileño de Proceso Civil, aprobado en marzo de 2015, el precedente fue destacado y también hay una mayor vinculación de las decisiones de los jueces de los tribunales inferiores a las decisiones de los tribunales de apelación cuando el asunto de los casos es lo mismo.*

El nuevo Código Brasileño de Proceso Civil predijo, en varios artículos, el precedente como un elemento esencial en la motivación de toda decisión judicial (arts. 489, § 1, V, VI, art. 926, § 2, art. 927, § 5 , art. 988, IV, art. 1042, § 1, II). Debido al gran número de casos en trámite en los tribunales brasileños, la obediencia a los precedentes por los jueces y tribunales inferiores se ha convertido en un requisito y, al mismo tiempo, en una necesidad para la racionalización del sistema judicial brasileño. Hasta la fecha, la evaluación del impacto de un caso a ser juzgado por el Superior Tribunal de Justicia (STJ) y su fijación como un precedente a seguir por los tribunales de primera instancia era intuitiva. El método presentado en este trabajo ofrece una manera racional para la modelización de la

fuerza y impacto de los precedentes en el sistema judicial brasileño, e puede ayudar la evaluación y mejora del sistema de precedentes en la legislación brasileña.
Este artículo presenta una análisis inicial de la red de datos formada por las citas de las decisiones del STJ por el período del 19/09/1996 al 01/03/2015. La función principal del STJ en el sistema judicial brasileño es estandarizar la interpretación y aplicación de la ley federal por otros jueces y tribunales brasileños. La base de datos de la investigación consiste en 922.333 decisiones judiciales. Parte de los documentos de las decisiones judiciales estaba disponible en HTML y decisiones anteriores a 2000 sólo están disponibles en formato PDF.
Teniendo en cuenta cada una de las decisiones como un nodo, se han identificado alrededor de 4 millones de bordes (cita de un caso en la corte por otro). En una red de grafos, cuando estes se utilizan como modelos de redes, se pueden identificar propiedades como la densidad de la red, el grado de distribución, coeficiente de agrupamiento y la centralidad de los nodos. A medida que la base de datos de la investigación es muy grande, los autores han deseado considerar subconjuntos de decisiones, agrupados por un tema específico, o un cuerpo de legislación específica, con el fin de verificar cómo las propiedades de los grafos se pueden identificar y lo que informan sobre el peso de los precedentes. Los autores también tienen la intención, en los próximos trabajos, de modelar otras redes de la misma base de datos do STJ, como la red de citas de leyes y citas de autores y teorías del Derecho en las decisiones do STJ. Otra modelización que los autores llevarán a cabo será la evolución o decadencia del precedente en el tiempo. Por lo general, lo cambio de precedentes con el tiempo se relaciona con la edición de nuevas leyes en los ordenamientos jurídicos de Civil Law.

1. Introdução

Seguramente o sistema judicial brasileiro avançou muito nos últimos anos com relação à informatização processual, e em termos comparativos com vários países da Europa, situa-se num patamar equivalente. A Itália, por exemplo, implementou uma ampla política de informatização processual e a colocou efetivamente em prática e de modo obrigatório apenas em 2014[1]. Assim como o Brasil, a Itália pretende com a adoção do meio eletrônico obter uma razoável melhora dos tempos de tramitação processual.

Seria a informatização processual apenas uma mudança de suporte do procedimento, do tradicional papel para o meio eletrônico, sem alterar a natureza própria da ciência jurídica processual? O presente artigo parte da premissa, compartilhada pelo processualista e teórico do Processo em Rede José Eduardo de Resende Chaves Jr., de que a resposta a esta questão é negativa. Assim como tantos setores da sociedade e da economia foram impactados pelo desenvolvimento da internet, o mais correto seria pensarmos no novo processo eletrônico como um processo em rede. O presente artigo pretende ser um exemplo de como a informatização processual abre a realidade jurídica para novas técnicas e novas disciplinas, oriundas da ciência da computação, e que permitem pensar e

[1] O leitor interessado poderá consultar o sítio do Ministério da Justiça Italiano http://pst.giustizia.it/PST/it/homepage.wp para maiores detalhes sobre o processo de informatização judicial italiano. Em 30 de junho de 2014, na Itália o registro eletrônico tornou-se obrigatório para: -requerimento de ordem de pagamento, tanto em andamento, quanto prestes a iniciar; - atos de advogados, ordens do juiz e todos os atos registrados em processos civis iniciados a partir de 01 de julho de 2014. Em 2015, o arquivo digital será obrigatório para todos os atos de processos já iniciados no dia 30 de junho de 2014, e será estendida aos processos das Cortes Superiores (segundo grau) em 30 de junho de 2015. Para um panorama da Justiça Eletrônica no Brasil, cf. Cesar Antonio Serbena (Coord.), E-Justiça: Processo Eletrônico e transparência do Poder Judiciário, Curitiba, Ed. Juruá, 2013.

reformular o próprio conceito de Direito e seus tradicionais institutos. Nosso objetivo foi desenvolver um modelo para pensar a argumentação judicial em rede, e de como ela gera uma rede de citações que é significativamente importante para analisar os precedentes judiciais.

Com a recente digitalização dos acervos de decisões dos Tribunais Brasileiros, podem ser aplicadas técnicas de visualização de dados aos bancos de dados judiciais, com o objetivo de extrair conhecimento e identificar propriedades relevantes, as quais não são visíveis sem o emprego de técnicas adequadas de *Data Mining* (mineração de dados). Um método já utilizado nos países de *Common Law* consiste na modelagem de uma rede de citações das decisões judiciais por outras decisões judiciais, de uma mesma corte ou de cortes diferentes, onde certas propriedades são identificadas e elas podem conferir uma importância maior ou menor a cada um dos precedentes. Apesar do sistema jurídico brasileiro ser fortemente influenciado pelo *Civil Law*, os precedentes tem ganhado um peso cada vez maior como fonte das decisões judiciais. No novo Código Brasileiro de Processo Civil, aprovado em março de 2015, o precedente ganhou destaque e há igualmente uma vinculação maior das decisões dos juízes de primeira instância às decisões dos Tribunais de Recurso quando a matéria dos casos é a mesma, ou nas questões envolvendo demandas repetitivas e o instituto da repercussão geral.

O novo Código Brasileiro de Processo Civil previu, em diversos artigos, o precedente como um elemento essencial na fundamentação de qualquer decisão judicial (arts. 489, §1º, V, VI, art. 926, §2º, art. 927, §5º, art. 988, IV, art. 1042, §1º, II). Devido ao grande número de processos tramitando nas cortes brasileiras, a obediência ao precedente pelos juízes e tribunais inferiores tornou-se uma exigência e ao mesmo tempo uma necessidade para a racionalização do sistema judicial brasileiro[2].

Até o presente momento, o processo de avaliação da repercussão de um caso a ser julgado pelas Cortes Superiores Brasileiras (Superior Tribunal de Justiça e Supremo Tribunal Federal) e sua fixação como um precedente a ser obedecido pelas cortes inferiores seguiu um modelo baseado na intuição e principalmente na experiência dos julgadores, sem uma métrica precisa.

Entretanto, devido ao grande volume de decisões publicadas, a análise manual é inviável, por isto estudos vêm sido desenvolvidos em busca de identificar as decisões mais influentes por período de tempo e as decisões que podem vir a se destacar no futuro. A maioria dos estudos origina-se de países em que é adotado C*ommon Law*, em que a base das decisões é fortemente ligada aos precedentes, mas mesmo em países que têm como base o C*ivil Law* o estudo dos precedentes se faz importante, pois este é fonte secundária.

Entre os diversos estudos existentes para analisar precedentes, há os que utilizam grafos[3] [4] [5] para verificar citações entre julgados e assim identificar os mais relevantes, conforme abordado a seguir.

---

[2] No plano doutrinário processual, vários processualistas defendem este ponto de vista como L.G. Marinoni, D. Mitidiero e P. Pereira. O leitor interessado em aprofundar este ponto poderá consultar as obras de Luiz Guilherme Marinoni, pela Ed. Revista dos Tribunais: A Ética dos Precedentes: justificativa do novo CPC, Precedentes Obrigatórios, O STJ enquanto Corte de Precedentes, Precedentes Jurisprudenciais, e pela Ed. Jus Podivm, A força dos Precedentes; por Daniel Mitidiero, a obra Cortes Superiores e Cortes Supremas: do controle à interpretação, da jurisprudência ao precedente, pela Ed. Revista dos Tribunais; e por Paula Pessoa Pereira, a obra Legitimidade dos Precedentes: universabilidade das decisões do STJ, pela Editora Revista dos Tribunais.

[3] Chandler, S. J. The Network Structure of Supreme Court Jurisprudence. University of Houston Law Center, 2005.

[4] Fowler, James H.; Johnson, Timothy R.; Spriggs II, James F.; Jeon, Sangick; Wahlbeck, Paul J. Network Analysis and the Law: Measuring the Legal Importance of Supreme Court Precedents. Political Analysis, v. 15, n. 3, p. 324-346, 2007.

[5] Smith, Thomas A. The Web of Law. San Diego Legal Studies Research Paper n. 06-11, 2005.

O método apresentado no presente trabalho oferece uma forma racional para a modelagem da força e repercussão dos precedentes no sistema judicial brasileiro, podendo servir de auxílio à avaliação e aperfeiçoamento do sistema de precedentes no Direito brasileiro.

1.1. Precedentes Judiciais

O precedente judicial ou jurisprudencial, segundo Reale[6], é a forma de revelação do direito que se processa através do exercício da jurisdição, em virtude de uma sucessão harmônica de decisões dos tribunais.

Ao se analisar as decisões proferidas pelos tribunais é possível identificar divergências entre sentenças relativas às mesmas questões de fato e de direito, entretanto, não pode ser considerada uma falha, mas sim a margem de poder que os magistrados possuem ao interpretar, coordenar ou preencher as lacunas das normas.

Uma das vantagens do magistrado ao se utilizar de jurisprudência é a de buscar uma equidade nas decisões e a destas inovarem em matéria jurídica, em que estabelecem normas que não estão vinculadas estritamente à lei, mas estão construídas através da conexão de dispositivos ou separação de preceitos utilizados por muito tempo. Desta forma, verifica-se que o Brasil apesar de não adotar o sistema jurídico C*ommon Law*, destaca o uso das jurisprudências em suas decisões.

1.2. Grafos

Segundo Easley e Kleinberg[7] um grafo é uma maneira de especificar relações entre uma coleção de itens. Um grafo consiste num conjunto de objetos, denominados nós, em que certos pares de nós estão conectados por meio de arestas. Estes dividem-se em dois tipos o não direcionado e o direcionado. A diferença entre o grafo direcionado e o não direcionado, é que na figura 1.a) “A” está conectado com “B” e “B” está conectado com “A”, mas na figura 1. b) “A” está conectado com “B”, mas “B” não está conectado com “A”.

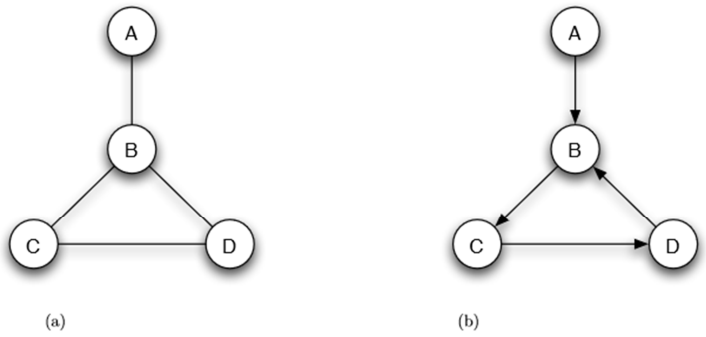


Figura 1. (a) Grafo não direcionado. (b) Grafo direcionado. Fonte: Easley e Kleinberg[8].

Os grafos são muito úteis, pois servem como modelos matemáticos de estruturas de rede[9]. Exemplo de classes de estruturas de grafos são as redes sociais, em que os nós são pessoas ou grupos de

[6] Reale, Miguel. Lições preliminares de Direito. 25ª ed. São Paulo: Saraiva. 2001.

[7] Easley David and Kleinberg Jon. Networks, Crowds, and Markets: Reasoning about a Highly Connected World. New York, NY, USA: Cambridge University Press, 2010.

[8] Ibidem.

[9] Ibidem.

pessoas e as arestas um tipo de interação social, outro exemplo, são as redes de informações, em que os nós podem ser documentos e as arestas as citações entre os documentos.

Há modelos de grafos criados empiricamente com determinadas características que identificam os tipos de grafos que podem existir, conforme detalhados abaixo. Ainda, as características de um grafo podem ser calculadas através das mais diversas métricas existentes.

Em busca de caracterizar uma rede de citações, foram escolhidas as propriedades mais utilizadas em estudos de precedentes em outros países[10] [11] [12].

1.3. Tipos de Redes

Há modelos estudados empiricamente que auxiliam no estudo das redes reais e cada tipo de rede possui propriedades estatísticas que o identificam. Entre os tipos de redes existentes estão as redes aleatórias, as de mundo pequeno e as de livre escala.

As redes/grafos aleatórios começaram a ser estudadas por Paul Erdõs e Alfréd Rényi em 1959[13]. Neste modelo cria-se uma rede com *N* nós e conecta-se cada par de nós com uma probabilidade *p*, em que se gera uma rede com aproximadamente $pN(N-1)/2$ arestas distribuídas aleatoriamente. As redes aleatórias são muito importantes, pois podem ser utilizadas como um modo de avaliar as redes reais ao se comparar com elas.

As redes de mundo pequeno representam um tipo de grafo em que a maioria das conexões é estabelecida entre os vértices mais próximos. A principal característica das redes de mundo pequeno é um caminho relativamente curto entre qualquer par de nós de uma rede. Este caminho é representado pela quantidade mínima de arestas para um nó chegar até o outro.

Com relação as redes de livre escala, estas se caracterizam por poucos nós possuírem muitas conexões e muitos nós possuem poucas conexões, além disso, um novo nó que surja na rede tem uma maior probabilidade de se conectar a um nó com mais conexões a um com menos. Barabási e Bonabeau[14] citam exemplos de redes de livre escala reais existentes, como a *World Wide Web*, em que sites com muitas citações tendem a continuar sendo amplamente citados ao longo do tempo, assim como colaborações científicas.

Já as redes complexas têm sua origem na teoria dos grafos, entretanto, a teoria dos grafos foi inicialmente focada no estudo dos grafos regulares e as redes de larga escala/complexas, desde 1950, têm sido descritas como grafos aleatórios.

Segundo Albert e Barabási[15], as redes complexas descrevem uma vasta quantidade de sistemas presentes na natureza e na sociedade. Um exemplo muito citado é o da Internet em que as redes de computadores são formadas por computadores e roteadores conectados através de cabos. E com o crescente estudo sobre as redes complexas, começou-se a questionar se estas são realmente

---

[10] Chandler, S. J. The Network Structure of Supreme Court Jurisprudence. University of Houston Law Center, 2005.

[11] Fowler, James H.; Johnson, Timothy R.; Spriggs II, James F.; Jeon, Sangick; Wahlbeck, Paul J. Network Analysis and the Law: Measuring the Legal Importance of Supreme Court Precedents. Political Analysis, v. 15, n. 3, p. 324-346, 2007.

[12] Smith, Thomas A. The Web of Law. San Diego Legal Studies Research Paper, n. 06-11, 2005.

[13] Albert, R.; Barabási, A. Statistical mechanics of complex networks. Reviews of Modern Physics, v. 74, 1. Ed., p. 47-97, 2002.

[14] Barabasi, A.-L.; Bonabeau, E. Scale-Free Networks. Scientific American, 50-59, 2003.

[15] Albert, R.; Barabási, A. Statistical mechanics of complex networks. Reviews of Modern Physics, v. 74, 1. Ed., p. 47-97, 2002.

derivadas de redes aleatórias, mas para isto foi necessário quantificar alguns princípios de organização destas redes[16], que os diferencie das características das redes aleatórias.

1.4. Métricas

Métricas são formas de se avaliar as propriedades de um grafo, para tanto, a seguir serão descritas algumas das medidas utilizadas para avaliar redes complexas.

1.4.1. Distribuição de grau

O grau de um nó se caracteriza pelo número de arestas conectadas a ele. Segundo Newman[17] o grau de distribuição divide-se em grau de entrada e saída em grafos direcionados, que significam a quantidade de citações que um nó recebeu e quantidade de citações que um nó realizou respectivamente. O grau de entrada de um nó pode ser um indicativo da importância do mesmo e implica numa probabilidade do mesmo ser citado futuramente, entretanto, há medidas que podem caracterizar esta importância de forma diferenciada como a centralidade de autovetor.

Em redes aleatórias, como as arestas são inseridas aleatoriamente, a maioria dos nós possuem a mesma quantidade de arestas e que é próximo ao grau médio da rede. A distribuição de grau em uma rede aleatória segue a distribuição de Poisson.

Já para as redes complexas, Albert e Barabási[18] verificaram que a distribuição de grau segue a distribuição de cauda longa (*power-law tail*) diferente da rede aleatória. Exemplos de redes com este tipo de distribuição são a *World Wide Web*[19] e a Internet[20].

1.4.2. Coeficiente de Agrupamento

Uma das propriedades das redes sociais é a formação de cliques, ou seja, grafos em que todos os nós estão conectados entre si e baseado nisso calcula-se o coeficiente de agrupamento[21] para estimar o quanto os nós estão agrupados.

Um exemplo citado por Albert e Barabási[22] é o de um nó *i* com k_i arestas conectadas a k_i nós, em que se os vizinhos de *i* forem parte de um clique/agrupamento terão $k_i(k_i - 1)/2$ arestas entre eles. Com base nisso, o coeficiente de agrupamento é calculado dividindo o número de arestas existentes entre os k_i nós, vizinhos de *i*, e o número total de arestas possíveis entre eles $k_i(k_i - 1)/2$.

$$C_i = \frac{2E_i}{k_i(k_i - 1)}$$

[16] Albert, R.; Barabási, A. Statistical mechanics of complex networks. Reviews of Modern Physics, v. 74, 1. Ed., p. 47-97, 2002.

[17] Newman, M. E. J. The structure and function of complex networks. SIAM Review, v. 45, n. 2, p. 167–256, 2003.

[18] Albert, R.; Barabási, A. Statistical mechanics of complex networks. Reviews of Modern Physics, v. 74, 1. Ed., p. 47-97, 2002.

[19] Albert, R.; Jeong, H.; Barabási, A.L. Internet: Diameter of the world-wide web. Nature 401, n. 6749, p. 130-131, 1999.

[20] Faloutsos, M.; Faloutsos, P.; Faloutsos, C. On power-law relationships of the Internet topology. In: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication (SIGCOMM '99). ACM, New York, NY, USA, 251-2, 1999.

[21] Watts, D. J.; Strogatz, S. H. Collective dynamics of small-world networks. Nature, v. 393, p. 440-442, 1998.

[22] Albert, R.; Barabási, A. Statistical mechanics of complex networks. Reviews of Modern Physics, v. 74, 1. Ed., p. 47-97, 2002.

O coeficiente de agrupamento divide-se em global e local, em que o primeiro fornece uma visão geral da rede e o segundo uma visão local de cada nó, conforme fórmula anteriormente citada.

Para calcular o coeficiente de agrupamento da rede inteira calcula-se a média de todos os coeficientes de agrupamentos dos nós.

Uma observação citada por Albert e Barabási[23] é a de que nos grafos aleatórios o coeficiente de agrupamento é igual à probabilidade da criação de arestas entre os nós. Os autores ainda citam que nas redes reais o coeficiente de agrupamento é tipicamente muito maior comparado ao de uma rede aleatória com os mesmos números de vértices e arestas.

1.4.3. Centralidade de autovetor

Também denominado *eigenvector centrality*, foi proposto por Bonacich e Lloyd[24], e mede a importância de um nó em função da importância de seus vizinhos. Bonanich e Lloyd[25] definem a centralidade de um elemento como uma combinação linear das centralidades dos elementos a ele conectados e é baseado no conceito de autovalores e autovetores da matriz de adjacência do grafo.

Freitas[26] cita que o método consiste em determinar o autovalor de maior valor absoluto de uma matriz e seu correspondente autovetor de maneira aproximada.

Para calcular a centralidade de autovetor de um grafo *G* deve-se executar o seguinte cálculo:

$$c_{eig} = (v_k) = x_k$$

$$x_k = \frac{1}{\rho}\sum_{j=1}^{n} a_{kj}x_j \; ; k = 1, \dots, n$$

Sendo:

n = número de vértices;

v_k = vértice que pertence a um grafo conexo *G*;

x_k = k-ésima coordenada do autovetor positivo unitário *x* associado ao índice do grafo;

Freitas[27] ainda cita que sendo a multiplicidade do raio espectral é igual a 1, todos os autovetores positivos associados ao índice serão múltiplos escalares deste. Com os autovetores positivos associados ao índice de G, $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$, terá $y = \alpha . x$, se $x_i \leq x_j$ então $y_i \leq y_j$.

Um exemplo de cálculo do autovetor é o citado por Freitas[28], de acordo com a figura 2.

[23] Albert, R.; Barabási, A. Statistical mechanics of complex networks. Reviews of Modern Physics, v. 74, 1. Ed., p. 47-97, 2002.

[24] Bonacich, P.; Lloyd, P.Eigenvector-like measures of centrality for asymmetric relations. Social Networks, v. 23, n.3, p. 191-201, 2001.

[25] Ibidem.

[26] Freitas, L. Q. de. Medidas de Centralidade em Grafos. 2010. Dissertação (Mestrado em Engenharia de Produção) - UFRJ.

[27] Ibidem.

[28] Ibidem.

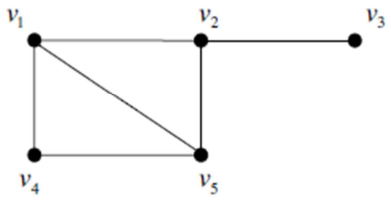


Figura 2. Exemplo de grafo com cinco nós e seis arestas. De acordo com a métrica de centralidade de autovetor os nós mais centrais são o v_1 e o v_5. Fonte: Freitas[29].

Em que o polinômio característico do grafo *G* e o respectivo espectro são:

$$\rho_G(\lambda) = \lambda^5 + 6\lambda^3 - 4\lambda^2 + 3\lambda + 2$$

$$spectG = \begin{bmatrix} 2{,}641 & 0{,}723 & -1 & -1{,}775 \\ 1 & 1 & 1 & 1 \end{bmatrix}^t$$

Logo, o autovetor positivo e de norma 1 associado a $\rho = 2{,}641$ é:

$$x = [0{,}537 \quad 0{,}474 \quad 0{,}179 \quad 0{,}406 \quad 0{,}537]$$

Por fim, para cada vértice de *G*, as centralidades de autovetor são:

$$c_{eig}(v_1) = c_{eig}(v_5) = 0{,}537$$

$$c_{eig}(v_2) = 0{,}474$$

$$c_{eig}(v_3) = 0{,}179$$

$$c_{eig}(v_4) = 0{,}406$$

Em que se verifica que os vértices mais centrais do grafo são o v_1e o v_5.

A diferença da métrica de centralidade de autovetor em relação à distribuição de graus de entrada de um nó é ilustrada na figura 3. Na figura 3.a) a circunferência do nó varia conforme a quantidade de vizinhos que possui, não considerando se estes são centrais na rede. Ainda, na figura 3.a) o nó com maior grau de entrada é o nó 10, por receber 2 arestas.

Já a figura 3.b) ilustra a métrica da centralidade de autovetor, uma vez que quanto maior a centralidade, maior é a circunferência do nó. Uma clara diferença é quanto ao nó 13 da figura 3.b), que, apesar de ter apenas uma citação, como ele é citado por um nó também de centralidade alta, acaba por ter uma centralidade maior. Já quanto ao nó 10 da figura 3.a) que apenas considera a quantidade de citações, o nó tem a circunferência menor, ou seja, não o classifica com maior centralidade.

1.4.4. Modularidade

Uma propriedade muito comum em várias redes sociais é a detecção de estruturas de comunidades, que consiste na divisão dos nós da rede em grupos com conexões internas densas e externas esparsas[30]. O estudo das estruturas de comunidades em redes sociais está relacionado com o

[29] Freitas, L. Q. de. Medidas de Centralidade em Grafos. 2010. Dissertação (Mestrado em Engenharia de Produção) - UFRJ.

[30] Girvan, M; Newman, M. E. J. Finding and evaluating community structure in networks. Phys. Rev. E, v. 69, n. 026113, 2003.

particionamento de grafos na teoria dos grafos e ciência da computação, e com o agrupamento hierárquico na sociologia[31].

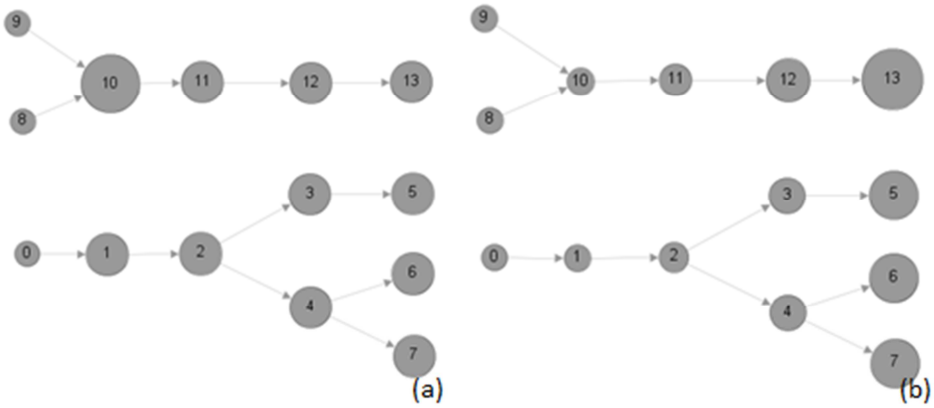


Figura 3. (a) Exemplo de grafo em que o tamanho dos nós varia conforme o grau de entrada do nó. (b) Exemplo de grafo em que o tamanho dos nós varia conforme a centralidade de autovetor. Fonte: Própria.

A definição para agrupamento segundo Jain, Murty e Flynn[32] é uma identificação não supervisionada de padrões nos dados. Esta técnica é útil para a análise de padrões, agrupamentos, tomadas de decisões, aprendizado de máquina e principalmente na exploração de relações entre nós de uma rede[33].

A modularidade é uma medida capaz de mensurar a qualidade de divisão feita em um grafo, ou seja, separar os diversos nós presentes na rede em grupos. Através do agrupamento dos nós em grupos, podem-se verificar os nós que apesar de não estarem conectados diretamente, possuem certo grau de similaridade.

A modularidade faz parte do estudo de detecção de comunidades, que por sua vez está ligado aos conceitos de teoria dos grafos e agrupamento hierárquico, pelo fato de utilizarem métodos para particionar a rede em subgrafos que representam cada comunidade existente na rede.

Quando o valor da modularidade for próximo de zero, a rede é considerada aleatória, e quando este valor se aproxima de um, as comunidades possuem estrutura forte. Contudo, na prática, segundo Girvan e Newman[34], os valores encontram-se entre 0,3 e 0,7.

Entre as técnicas para particionar o grafo em subgrafos estão as métricas de similaridade. Esta medida é fundamental para a definição dos agrupamentos segundo Jain, Murty e Flynn[35]. Devido à variedade de tipos de atributos, escalas, medidas de distância, estas devem ser escolhidas com cuidado segundo Jain, Murty e Flynn[36].

---

[31] Girvan, M; Newman, M. E. J. Finding and evaluating community structure in networks. Phys. Rev. E, v. 69, n. 026113, 2003.

[32] Jain, A.K.; Murty, M. N.; Flynn, P. J. Data Clustering: A Review. ACM Comput. Surv., v. 31, n. 3, p. 264-323, 1999.

[33] Ibidem.

[34] Girvan, M; Newman, M. E. J. Finding and evaluating community structure in networks. Phys. Rev. E, v. 69, n. 026113, 2003.

[35] Jain, A.K.; Murty, M. N.; Flynn, P. J. Data Clustering: A Review. ACM Comput. Surv., v. 31, n. 3, p. 264-323, 1999.

[36] Ibidem.

Porém, antes de abordar estas técnicas Jain, Murty e Flynn[37] citam que algumas questões podem influenciar na estrutura e operações, alterando a abordagem, como o tipo aglomerativo e divisivo.

O tipo aglomerativo pode ser exemplificado ao definir cada vértice como um agrupamento. Sucessivamente estes agrupamentos vão se unindo, até atenderem a um critério específico que os faça parar. Já o tipo divisivo define os vértices como parte de um único agrupamento que vai sendo dividido até atender a um critério.

Entre os algoritmos de agrupamento mais conhecidos estão o de Girvan e Newman[38], Newman[39], Clauset, Newman e Moore[40] e Blondel, Guillaume e Lambiotte[41].

1.5. Análise de Precedentes com grafos

O trabalho de Chandler[42] analisa a estrutura da rede de jurisprudência da Suprema Corte dos Estados Unidos. O artigo ressalta a importância do uso de modernas tecnologias para analisar estas redes, entretanto, a base para a realização dos estudos vem da física e da sociologia. Algumas das características extraídas da rede foram a densidade, grau de distribuição dos nós, grau de agrupamento e centralidade dos nós. A partir do trabalho realizado, o autor concluiu que medidas como centralidade de intermediação (*betweenness centrality*), centralidade de proximidade (*closeness centrality*) e centralidade de markov (*markov centrality*), contribuem para identificar os nós mais centrais de um tribunal. Já medidas como o coeficiente de agrupamento permite verificar o grau de interdependência dos nós. Além disso, o autor identificou comportamentos diferentes na rede dependendo do assunto dos processos/nós envolvidos.

No trabalho de Fowler et al.[43] é analisado em torno de 26 mil decisões majoritárias da Suprema Corte dos Estados Unidos e os casos que as citam entre 1791 e 2005. No artigo é descrito um método que utiliza os padrões detectados nas citações entre os processos, com o intuito de criar escores de importância que identificam os precedentes mais relevantes da Suprema Corte dos Estados Unidos a cada período de tempo analisado. O estudo se comprova válido ao demonstrar que o comportamento das futuras citações corresponde aos casos levantados como relevantes. Além disso, os autores concluem que a análise dos grafos é uma maneira viável para determinar o quanto um processo é central para o direito norte americano.

No mesmo ano, Fowler e Jeon[44] publicaram um artigo em que analisaram em torno de 30 mil decisões majoritárias e os casos que as citaram entre 1754 e 2002. Neste trabalho identificaram a evolução do "*stare decisis*" (precedentes) no século 19 e uma mudança ocorrida na corte de

---

[37] Jain, A.K.; Murty, M. N.; Flynn, P. J. Data Clustering: A Review. ACM Comput. Surv., v. 31, n. 3, p. 264-323, 1999.

[38] Girvan, M; Newman, M. E. J. Community structure in social and biological networks. Proceedings of the National Academy of Sciences, v. 99, n. 12. 2002.

[39] Newman, M. E. J. Fast Algorithm for detecting community structure in networks. Phys. Rev. E, v. 69, 6. Ed. 2004.

[40] Clauset, A.; Newman, M. E. J.; Moore, C. Finding community structure in very large networks. Phys. Rev. E, v. 70, 6. Ed., 2004.

[41] Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment, 2008.

[42] Chandler, S. J. The Network Structure of Supreme Court Jurisprudence. University of Houston Law Center, 2005.

[43] Fowler, James H.; Johnson, Timothy R.; Spriggs II, James F.; Jeon, Sangick; Wahlbeck, Paul J. Network Analysis and the Law: Measuring the Legal Importance of Supreme Court Precedents. Political Analysis, v. 15, n. 3, p. 324-346, 2007.

[44] Fowler, James H.; Jeon, Sangick. The authority of Supreme Court precedente. Social Networks, v. 30, 1 ed., p. 16-30, 2008.

Warren. É apresentado também, além dos precedentes mais relevantes, a ascensão e declínio de cada precedente ao longo do tempo.

Já no trabalho de Smith[45], é afirmado que a partir das informações disponibilizadas sobre citações de processos da justiça norte-americana forma-se uma rede com propriedades matemáticas e estatísticas assim como em redes encontradas na natureza e em relações humanas. Além disso, o trabalho ressalta a importância do estudo nesta área e identifica a rede de citações como uma rede de escala livre, em que há poucos nós com muitas citações e muitos nós com poucas conexões.

2. METODOLOGIA

Esta seção apresenta as etapas do método de pesquisa para a análise do estudo de caso proposto.

2.1. Aquisição e pré-processamento dos dados

Para analisar os dados e acompanhar as decisões proferidas é necessário escolher um tribunal que disponibilize os documentos em formato HTML (HyperText Markup Language) para permitir de forma viável a sua extração. Devido ao grande volume de documentos, deve-se escolher um grupo de processos com uma determinada característica para facilitar a compreensão da rede e realizar o estudo de caso.

Para o pré-processamento dos dados deve-se extrair as identificações dos processos (número, tipo de processo e estado) e as citações entre os processos extraídos.

Durante o pré-processamento deve ser considerado que um processo pode ter vários recursos dentro do tribunal, em que seu número identificador permanece o mesmo. Portanto, para fins de evitar equívocos durante a extração das citações, foram separados os processos por numeração, estado e tipo de processo, e quando uma citação é realizada cria-se um vínculo entre o processo citador e o processo citado. Tanto o processo citador quanto o processo citado podem abranger todos os recursos existentes nos mesmos.

Ainda, somente foram extraídas as citações dos documentos que possuíam um número seguido do estado ao lado. Foi escolhida esta forma devido a não ter sido detectado um padrão único de citações, mas esta se apresenta como majoritária. Além disso, é necessário também o estado e o tipo do processo, pois há processos com números e estados iguais, ou ainda números e tipos de processos iguais.

Um exemplo de citação é o trecho abaixo retirado da Sentença Estrangeira Contestada Nº 1 - EX (2007/0156979-5) do STJ, sendo o tipo de processo da citação o "RESP", o número o "649711" e o estado de origem do processo "BA".

> *"Recurso especial conhecido e provido." (REsp 649711/BA, Rel. Ministro ANTÔNIO DE PÁDUA RIBEIRO, Rel. p/ Acórdão Ministro CARLOS ALBERTO MENEZES DIREITO, TERCEIRA TURMA, julgado em 06/06/2006, DJ 07/08/2006, p. 219)"*

Depois de identificada a citação, busca-se o número, estado e processo na base de dados de processos extraídos. Caso não se encontre, pode-se inferir que esta citação pertence a outro tribunal, ou é uma decisão monocrática do STJ, ou ainda é um processo não disponibilizado em HTML. Este último caso é muito comum em processos julgados antes do ano 2000.

Após tratadas as citações, os dados estão preparados para formar as redes sociais com grafos para identificar propriedades, calcular métricas da rede e realizar projeções.

[45] Smith, Thomas A. The Web of Law. San Diego Legal Studies Research Paper n. 06-11, 2005.

2.2. Identificação das propriedades

Com o grafo das citações formado associa-se a cada nó a data de julgamento para possibilitar a verificação das citações ao longo do tempo e se as propriedades globais do grafo se alteraram. É possível também verificar a alteração das propriedades de cada nó ao longo do tempo, como por exemplo, se o nó passou a ter mais citações a partir de um período ou se o nó se tornou mais relevante na rede.

A análise das propriedades é essencial para acompanhar em quais períodos os nós se tornaram relevantes e ser possível inferir se este será citado futuramente. A partir destes dados, profissionais da área jurídica podem verificar se ocorreram mudanças sociais, políticas ou legislativas que justifiquem o ocorrido.

3. Dados Coletados

Como estudo de caso foram selecionados os acórdãos presentes no Superior Tribunal de Justiça, julgados por um grupo de juízes em colegiado.

Após a escolha do órgão do qual devem ser extraídas as informações e o assunto, foram extraídas de forma automática todas as decisões proferidas em colegiado do site do Superior Tribunal de Justiça (STJ) publicadas no período entre 19/09/1996 e 01/03/2015 e disponibilizadas em formato HTML. Após a extração, foram realizados os devidos tratamentos dos dados.

A análise textual do documento consistiu em identificar a numeração dos processos e tipos (recurso, agravo, habeas corpus, entre outros) e buscar citações dos mesmos dentro das outras decisões analisadas.

Ao todo foram extraídos 922.333 acórdãos e identificadas em torno de 4 milhões de citações entre as mesmas. Vale ressaltar que se uma jurisprudência foi citada mais de uma vez em um mesmo processo judicial, foi considerado apenas uma por documento.

Já o intervalo de tempo entre 1996 e 2014 foi escolhido por ser o intervalo em que foram encontrados acórdãos para *download* através do site do órgão. Como existem ainda processos sendo julgados e publicados, algumas jurisprudências deste período podem não estar inclusas, por estarem indisponíveis no momento que foram extraídos os documentos.

Como há um volume muito grande dos dados extraídos, foram selecionados processos de um determinado tema para facilitar a compreensão da rede e das mudanças que ocorreram ao longo do tempo. Para isto, foi realizada uma busca no site do STJ no dia em 20/03/2015 por jurisprudências em que estejam presentes os artigos 20 e 21 do Código Civil Brasileiro de 2002, artigos que tratam do direito à imagem[46].

A partir destes acórdãos, foram extraídas as citações de todos os acórdãos que citaram estes processos, e os acórdãos citados por estes processos, ou seja, conexões de primeiro grau. Para comparação, também foram extraídas conexões de segundo grau, ou seja, nós conectados aos nós citados ou que citam os processos escolhidos, mas que não estão conectados diretamente ao

[46] Ao todo foram encontradas 28 decisões, de acordo com o que foi disponibilizado no sistema de busca do site do tribunal. 1) AgRg no AREsp 604475/RS; 2) REsp 1457199/RS; 3) REsp 1322704/SP; 4) REsp 1307366/RJ; 5) REsp 1337961/RJ; 6) REsp 1402091/SP; 7) REsp 1209474/SP; 8) REsp 1291865/RJ; 9) REsp 1335153/RJ; 10) REsp 1334097/RJ; 11) AgRg no AREsp 16247/RJ; 12) REsp 801109/DF; 13) AgRg no AREsp 13857/RS; 14) REsp 1267232/PR; 15) REsp 997475/RJ; 16) REsp 1195995/SP; 17) REsp 1208612/RJ; 18) REsp 1020936/ES; 19) REsp 1200482/RJ; 20) REsp 1005278/SE; 21) AgRg no REsp 947368/RS; 22) AgRg no REsp 1171749/PE; 23) REsp 711644/SP; 24) REsp 812691/RJ; 25) REsp 651228/MG; 26) AgRg no Ag 684639/MS; 27) AgRg no Ag 658134/RJ e 28) REsp 142308/DF.

processo escolhido. Um exemplo de grafo com conexões de primeiro e segundo graus pode ser verificado na figura 4.

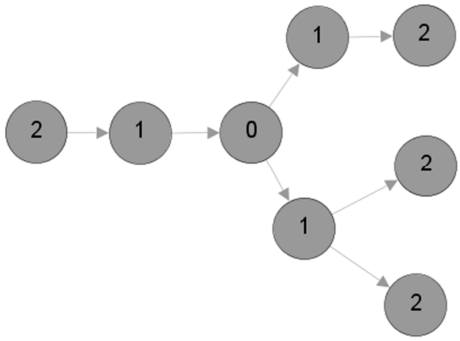


Figura 4. Exemplo de um nó (0) com conexões de primeiro e segundo graus, representados pelos nós através dos número 1 e 2 respectivamente. Fonte: Própria.

Houve também a separação por ano, para ver quantas vezes o acórdão foi citado e ou se houve alteração no valor das métricas de cada nó e nas métricas que avaliam a rede.

Diante da proposta, no total foram identificados 173 nós com conexões de primeiro grau com os processos estudados e 179 arestas entre os mesmos. Já em relação à rede com conexões de primeiro e segundo graus, foram identificados 845 nós e 961 arestas.

4. RESULTADOS

As redes de citações entre processos representam uma das maiores redes que se pode formar com dados judiciais. Os nós da rede são os processos judiciais e as arestas são representadas pelas citações aos outros processos, que representam uma forma de embasar uma opinião ao julgar um processo. As arestas são direcionadas, em vista de que um processo pode citar outro, mas o outro pode não citá-lo.

Com a limitação da rede que foi realizada, o tamanho da rede se restringiu a 845 nós e 961 arestas. Quanto à distribuição dos graus, verificou-se que a rede segue uma distribuição exponencial, pelo fato de haver poucos nós com muitas conexões e muitos nós com poucas arestas.

Diante dos dados obtidos com a rede formada, comparou-se com a tabela de redes reais citada por Albert e Barabási[47], conforme tabela 1. Pôde-se constatar que a rede de citações do STJ para o estudo de caso apresentado possui um valor baixo para o grau médio da rede, em comparação às outras redes reais, mas o coeficiente de agrupamento e caminho médio são similares e se diferenciam dos valores das redes aleatórias criadas com base no números de nós e arestas da rede original.

4.1. Propriedades da Rede

A seguir serão descritas algumas propriedades das redes obtidas a partir do estudo de caso proposto.

4.1.1. Graus de Entrada e Saída e Centralidade de Autovetor

Devido às arestas do grafo de citações serem direcionadas, verificou-se a distribuição do grau de entrada e saída (quantidade de citações que recebeu e realizou respectivamente), em que foi

[47] Albert, R.; Barabási, A. Statistical mechanics of complex networks. Reviews of Modern Physics, v. 74, 1. Ed., p. 47-97, 2002.

constatado que há muitos nós com poucas arestas de entrada e saída e poucos nós com uma grande quantidade de conexões, devido a rede seguir uma distribuição exponencial.

Tabela 1. Características de algumas redes reais. Para cada rede foi indicado o número de nós, grau médio <k>, caminho médio ℓ e o coeficiente de agrupamento C. Para uma comparação foi incluído o valor do caminho médio ℓ_{rand} e o coeficiente de agrupamento C_{rand} de um grafo aleatório com o mesmo tamanho e grau médio.
Fonte: Adaptação de Albert e Barabási[48].

Rede	Tamanho	<k>	ℓ	ℓ_{rand}	C	C_{rand}	Referência
WWW	153127	35.21	3.1	3.35	0.1078	0.00023	Adamic[49]
Internet (Domínios)	3015-6209	3.52-4.11	3.7-3.76	6.36-6.18	0.18-0.3	0.001	Yook et al.[50]
Atores de filme	225226	61	3.65	2.99	0.79	0.00027	Watts e Strogatz[51]
Coautores (SPIRES)	56627	173	4.0	2.12	0.726	0.003	Newman[52 53 54]
E. coli	282	7.35	2.9	3.04	0.32	0.026	Wagner e Fell[55]
Redes de Citações STJ – 173 nós	173	1.035	1.788	1.738	0.048	0.002	-
Redes de Citações STJ – 845 nós	845	1.137	2.897	2.076	0.026	0	-

A figura 5 representa a rede de citações considerando os processos com conexões de primeiro grau, ou seja, processos que citam ou são citados por pelo menos um dos 28 processos selecionados para o estudo de caso. Pode-se verificar também duas figuras que ilustram o mesmo grafo, entretanto, o tamanho dos nós na figura 5.a) varia conforme o grau de entrada de cada nó, já na figura 5.b) o tamanho do nó varia conforme a centralidade de autovetor de cada nó.

O processo com maior grau de entrada possui 9 citações na rede que abrange processos de primeiro e segundo graus de conexões com os processos selecionados. O valor de centralidade de autovetor para este processo foi de 0,1203, e este valor ao ser comparado com todos os demais da rede ficou em 58º lugar dos nós mais centrais.

Já o processo com maior valor de centralidade de autovetor possui apenas três citações, pelo fato de ter sido citado por nós centrais na rede. Conforme será verificado abaixo, ao a rede aumentar e ser

[48] Albert, R.; Barabási, A. Statistical mechanics of complex networks. Reviews of Modern Physics, v. 74, 1. Ed., p. 47-97, 2002.

[49] Adamic, L. A. The small world web. In: Proceedings of the Third European Conference, ECDL'99. Berlin: Springer-Verlag, 443-452, 1999.

[50] Yook, S.-H.; Jeong, H.; Barabási, A.-L. Modeling the Internet's large-scale topology. Proceedings of the National Academy of Science, v. 99, p. 13382-13386, 2002.

[51] Watts, D. J.; Strogatz, S. H. Collective dynamics of small-world networks. Nature, v. 393, p. 440-442, 1998.

[52] Newman, M. E. J. The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. USA, v. 98, n. 2, p. 404–409, 2001.

[53] Newman, M. E. J. Scientific collaboration networks: I. Network construction and fundamental results. Phys. Rev. E, v. 64, n. 016131, 2001.

[54] Newman, M. E. J. Scientific collaboration networks: II. Shortest paths, weighted networks, and centrality. Phys. Rev. E, v. 64, n. 016132, 2001.

[55] Wagner, A.; Fell, D. A. The small world inside large metabolic networks. Technical Report 00-07-041, Santa Fe Institute, 2000.

considerado também os nós de grau 2 em relação aos processos selecionados, o nó identificado como o mais central na rede da figura 5, passa a ficar na 34ª posição dos nós centrais na rede da figura 6.

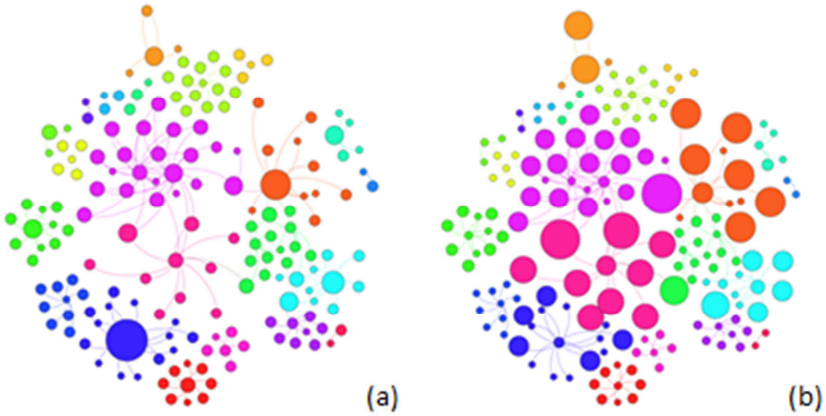


Figura 5. Comparação do grafo com citações de primeiro grau dos processo selecionados. No grafo a) pode-se verificar que o tamanho dos nós varia conforme o grau de entrada do mesmo. Já no grafo b) o tamanho do nó varia de acordo com a centralidade do autovetor. Fonte: Própria.

A figura 6 representa a rede de citações considerando os processos com conexões de primeiro e segundo graus. Pode-se verificar também duas figuras que ilustram o mesmo grafo, entretanto, o tamanho dos nós na figura 6.a) varia conforme o grau de entrada de cada nó, já na figura 6.b) o tamanho do nó varia conforme a centralidade de autovetor de cada nó.

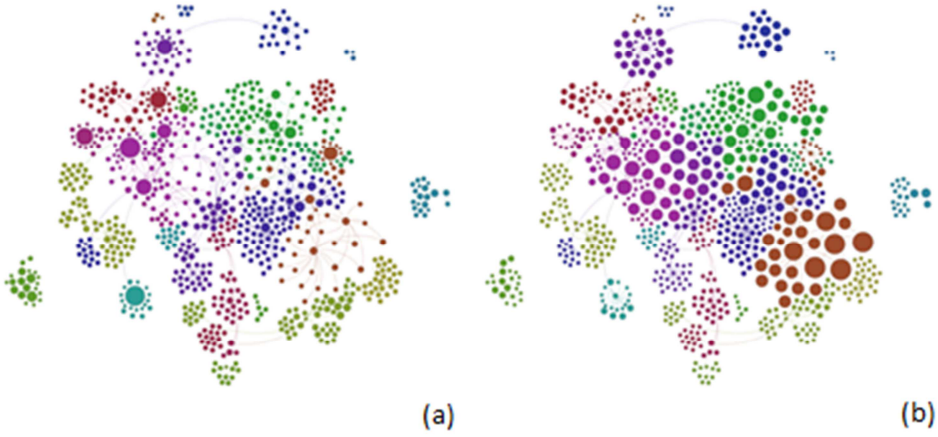


Figura 6. Comparação do grafo com citações de primeiro e segundo graus dos processo selecionados. No grafo a) pode-se verificar que o tamanho dos nós varia conforme o grau de entrada do mesmo. Já no grafo b) o tamanho do nó varia de acordo com a centralidade do autovetor. Fonte: Própria.

O processo com maior grau de entrada possui 13 citações na rede que abrange processos de primeiro e segundo graus de conexões com os processos selecionados. O valor de centralidade de autovetor para este processo foi de 0,0924, e este valor ao ser comparado com todos os demais da rede ficou em 294º lugar dos nós mais centrais. Verificou-se que apesar do nó ter bastantes citações, os nós que o citam não possuem centralidade elevada.

Já os processos com maior valor de centralidade de autovetor foram quatro e todos apenas com duas citações. Neste caso, pôde-se verificar que, apesar destes processos não abordarem sobre os artigos

do estudo de caso abordado, até por não estarem presentes ao se buscar no site do tribunal, eles possuem destaque na rede estudada.

4.1.2. Modularidade

A modularidade encontrada para a rede de citações que incluem as de primeiro e segundo graus dos processos selecionados foi de 0,908, que representa um valor alto para esta métrica. Já para a rede que considera apenas os processos ligados diretamente aos processos selecionados, foi identificada uma modularidade 0,848. Pode-se concluir que é possível dividir com sucesso os grupos de nós, o que pode facilitar na identificação de processos similares, apesar de não citados entre si.

4.2. Análise da Rede

A análise da rede consistiu em utilizar os processos selecionados a partir da busca do site do STJ juntamente com os processos com destaque em relação ao valor da centralidade de autovetor e analisá-los através das redes de citações criadas ao longo dos anos. Ao todo foram selecionados para análise 33 processos.

É possível verificar nos gráficos a mudança no valor da centralidade de autovetor de alguns processos. Ainda, em alguns casos, é possível verificar a existência de algumas linhas que começam a surgir apenas nos últimos anos, devido aos processos terem sido publicados somente nestes anos ou citados por outros nós centrais somente nesta época.

Na figura 7 é ilustrado o gráfico da rede de citações que considera apenas os processos conectados diretamente ao pelo menos um dos 28 processos selecionados. Cada linha representa um processo com seus respectivos valores de centralidade de autovetor ao longo dos anos (2001-2014).

Na linha verde verifica-se que um processo teve sua centralidade com o valor 1 desde 2001 e só teve sua centralidade reduzida em 2014. Em 2001, o processo possuía apenas uma citação de outro processo que também era citado uma única vez, esta situação se manteve a mesma até 2005, pois não houve mais citações durante este período, nem para outros processos. Já em 2006, apesar do processo representado pela linha verde não receber mais citações, outros processos passaram a receber citações, mas em igual número ao analisado.

Entre 2007 e 2013, a rede possui ainda poucas citações, entretanto, o processo analisado recebe um pouco mais de citações ao longo do tempo, mas o que mantém a sua centralidade alta é o fato dos demais nós da rede também estarem pouco citados por nós centrais.

A alteração na rede ocorre mesmo em 2014, em que houve mais citações e outros nós passaram a ganhar destaque na rede, além do processo analisado que se encontra na linha verde.

Já na figura 8 é ilustrado o gráfico da rede de citações que considera os processos conectados diretamente ao pelo menos um dos 28 processos selecionados, além dos processos com conexões de segundo grau. Cada linha representa um processo com seus respectivos valores de centralidade de autovetor ao longo dos anos (2001-2014).

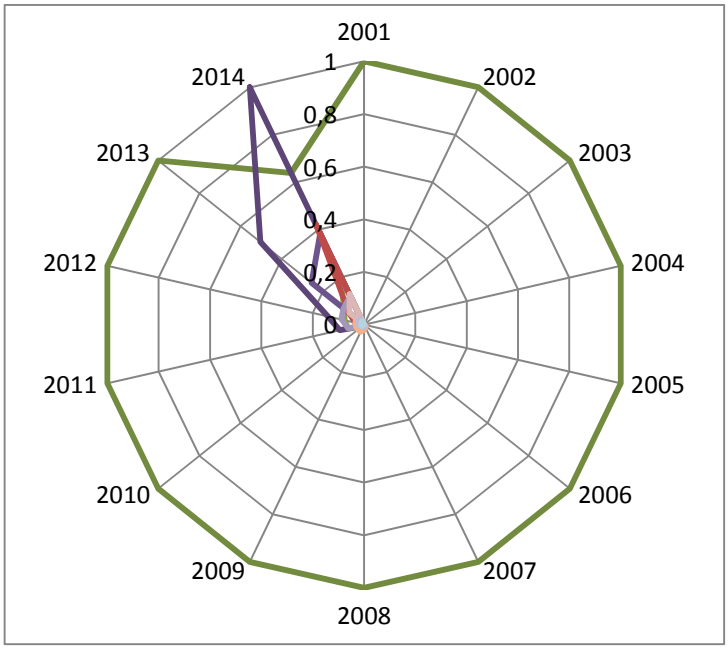


Figura 7. Rede com nós de primeiro grau em relação aos processos selecionados. Cada linha colorida representa um processo com seus respectivos valores de centralidade de autovetor ao longo dos anos (2001-2014). O valor da centralidade varia entre 0 e 1. Fonte: Própria.

Na linha verde verifica-se que um processo teve sua centralidade com o valor 1 em 2001 e teve sua centralidade mantida no valor de 0,7 entre 2002 e 2012, mas em 2013 e 2014 o valor da centralidade passou para menos de 0,4. A mudança no valor da centralidade de um processo que manteve sua centralidade alta por vários anos se dá pelo fato que houve o crescimento do número de processos e consequentemente citações nos últimos anos, favorecendo a mudança na rede. Verifica-se ainda, através de outras linhas, como a verde claro, que processos que começaram a ter sua centralidade destacada a partir de 2009, mantiveram o valor alto até 2014.

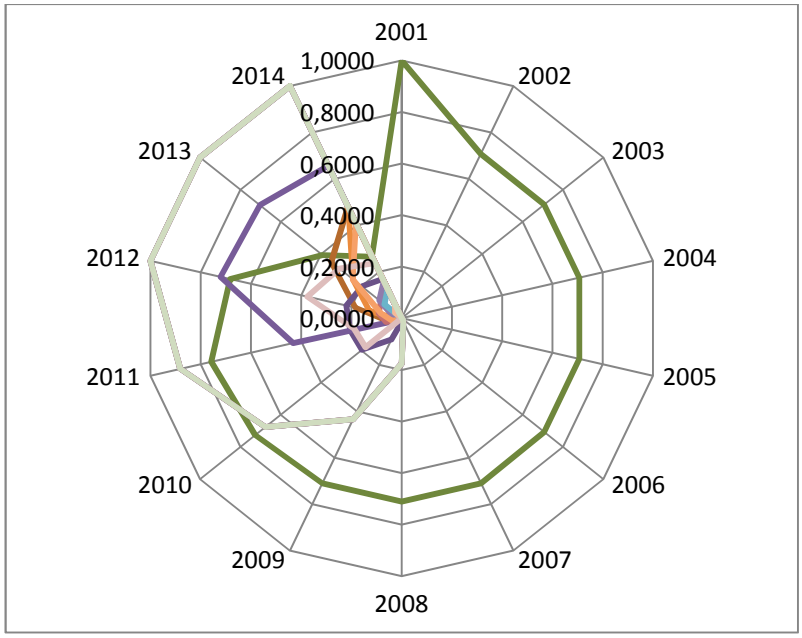


Figura 8. Rede com nós de primeiro e segundo graus em relação aos processos selecionados. Cada linha representa um processo com seus respectivos valores de centralidade de autovetor ao longo dos anos (2001-2014). O valor da centralidade varia entre 0 e 1. Fonte: Própria.

5. CONCLUSÃO E TRABALHOS FUTUROS

A partir dos dados extraídos do tribunal utilizado como estudo de caso, conclui-se que há um volume grande de informação a ser estudado e que através de técnicas computacionais é possível verificar se a rede de citações do tribunal segue algum modelo de rede real e quais os processos mais relevantes ao longo do tempo.

Com a revisão bibliográfica realizada, foi possível verificar que o estudo é viável, pois métodos similares vêm sendo aplicados em tribunais norte americanos, com o intuito de estudar como surgem os precedentes mais importantes e projetar futuras citações e comportamentos dos tribunais.

Ainda, a utilização de grafos direcionados para análise dos precedentes mostra-se a mais indicada, pois, os processos ao se conectarem por meio de citações, acabam por montar uma grande rede com características importantes para serem estudadas.

O trabalho é inovador no Brasil, pois, por adotar um sistema jurídico C*ivil Law,* a jurisprudência deveria ter menos relevância comparado aos países que adotam sistema jurídico C*ommon Law*, entretanto, é interessante verificar se as redes de ambos os sistemas jurídicos possuem semelhanças.

O trabalho também demonstra a possibilidade em se criar ferramentas para estudar a evolução do uso dos precedentes ao longo do tempo, como a ascensão ou declínio da força dos precedentes com a publicação de novos julgados.

Além disso, com o uso de métodos de detecção de comunidades, é possível agrupar processos/nós similares na rede, de acordo com as conexões existentes, o que permite identificar precedentes similares e sugestões de precedentes para cada caso.

6. Referências

Adamic, L. A. The small world web. In: Proceedings of the Third European Conference, ECDL'99. Berlin: Springer-Verlag, 443-452 (1999).

Albert, R.; Barabási, A. Statistical mechanics of complex networks. Reviews of Modern Physics, v. 74, 1. Ed., p. 47-97 (2002).

Albert, R.; Jeong, H.; Barabási, A.L. Internet: Diameter of the world-wide web. Nature 401, n. 6749, p. 130-131 (1999).

Barabasi, Albert-Laszlo; Albert, Reka. Emergence of Scaling in Random Networks. Science 286, n. 5439 (1999).

Barabasi, A.-L.; Bonabeau, E. Scale-Free Networks. Scientific American, 50-59, 2003.

Bonacich, P.; Lloyd, P. Eigenvector-like measures of centrality for asymmetric relations. Social Networks, v. 23, n.3, p. 191-201 (2001).

Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment (2008).

Chandler, S. J. The Network Structure of Supreme Court Jurisprudence. University of Houston Law Center (2005).

Clauset, A.; Newman, M. E. J.; Moore, C. Finding community structure in very large networks. Phys. Rev. E, v. 70, 6. Ed. (2004).

Easley David and Kleinberg Jon. Networks, Crowds, and Markets: Reasoning about a Highly Connected World. New York, NY, USA: Cambridge University Press (2010).

Faloutsos, M.; Faloutsos, P.; Faloutsos, C. On power-law relationships of the Internet topology. In: Proceedings of the conference on Applications, technologies, architectures, and protocols for computer communication (SIGCOMM '99). ACM, New York, NY, USA, 251-2 (1999).

Fowler, James H.; Johnson, Timothy R.; Spriggs II, James F.; Jeon, Sangick; Wahlbeck, Paul J. Network Analysis and the Law: Measuring the Legal Importance of Supreme Court Precedents. Political Analysis, v. 15, n. 3, p. 324-346 (2007).

Fowler, James H.; Jeon, Sangick. The authority of Supreme Court precedente. Social Networks, v. 30, 1 ed., p. 16-30 (2008).

Freitas, L. Q. de. Medidas de Centralidade em Grafos. Dissertação (Mestrado em Engenharia de Produção) - UFRJ (2010).

Girvan, M; Newman, M. E. J. Community structure in social and biological networks. Proceedings of the National Academy of Sciences, v. 99, n. 12 (2002).

Girvan, M; Newman, M. E. J. Finding and evaluating community structure in networks. Phys. Rev. E, v. 69, n. 026113 (2003).

Jain, A.K.; Murty, M. N.; Flynn, P. J. Data Clustering: A Review. ACM Comput. Surv., v. 31, n. 3, p. 264-323 (1999).

Kochen, M. The Small World. Ablex, Norwood, NJ (1989).

Newman, M. E. J. The structure of scientific collaboration networks. Proc. Natl. Acad. Sci. USA, v. 98, n. 2, p. 404–409 (2001).

Newman, M. E. J. Scientific collaboration networks: I. Network construction and fundamental results. Phys. Rev. E, v. 64, n. 016131 (2001).

Newman, M. E. J. The structure and function of complex networks. SIAM Review, v. 45, n. 2, p. 167–256 (2003).

Newman, M. E. J. Fast Algorithm for detecting community structure in networks. Phys. Rev. E, v. 69, 6. Ed. (2004).

Newman, M. E. J. Scientific collaboration networks: II. Shortest paths, weighted networks, and centrality. Phys. Rev. E, v. 64, n. 016132 (2001).

Reale, Miguel. Lições preliminares de Direito. 25ª ed. São Paulo: Saraiva (2001).

Smith, Thomas A. The Web of Law. San Diego Legal Studies Research Paper n. 06-11 (2005).

Wagner, A.; Fell, D. A. The small world inside large metabolic networks. Technical Report 00-07-041, Santa Fe Institute (2000).

Watts, D. J.; Strogatz, S. H. Collective dynamics of small-world networks. Nature, v. 393, p. 440-442 (1998).

Yook, S.-H.; Jeong, H.; Barabási, A.-L. Modeling the Internet's large-scale topology. Proceedings of the National Academy of Science, v. 99, p. 13382-13386 (2002).

ROWLER – A HYBRID RULE ENGINE FOR LEGAL REASONING

Johannes Scharf[1], Erich Schweighofer[2]

[1]PhD Researcher, University of Vienna, Centre for Computers and Law
Schottenbastei 10-16, 1010 Vienna, AT
johannes.scharf@gmx.at
[2]Professor, University of Vienna, Centre for Computers and Law
Schottenbastei 10-16, 1010 Vienna, AT
erich.schweighofer@univie.ac.at; http://rechtsinformatik.univie.ac.at

Keywords: ***Rules, Ontologies, OWL, Legal Reasoning, Public Administration***
Abstract: *This paper presents "rOWLer", a hybrid rule engine for legal reasoning. The engine combines the expressiveness of rules and ontologies to enable legal reasoning. It is tailored for the use in public administration (tax law, pension law, social benefits law etc.) and provides a flexible architecture, in particular concerning amendments. The proposed rule engine "rOWLer" takes into account the various time layers due to the dynamic change of the legal system as well as the different addressees of the various norms.*

1. Introduction

The development of the hybrid rule engine for legal reasoning "rOWLer" was part of the PhD thesis of the first author [29] and draws on experiences gained by modelling legal norms with Java and OWL 2. This research tries to fill the gap between the syntactical representation of norms (in XML or other formats) and the need of public administration for a powerful, easy to use and customizable legal rule engine. The architecture of rOWLer is aligned with the semantic web stack and is compatible with LegalRuleML [1], an upcoming standard for modelling legal rules. Current software solutions used in public administration could be improved by following the theoretical models available.

In this paper two extensions are presented: The temporal module for reasoning with time in a legal setting and, in particular, the personal module.

2. Motivation

The use of logic-based knowledge systems in public administration (e.g. in tax law) which formalize law by using thousands of if-then-else statements dates back to the 1970s in Austria but there is still no standard or unified methodology for implementation available. Formalization of statutes in practice happens mainly in an *ad hoc* fashion by the software expert often without considering legal theory at all.

Although the current models of law are rather useful and accepted in practice, they have several severe drawbacks. For instance they violate the isomorphism principle in a dynamic legal environment which makes maintenance a daunting task. Moreover the legal dynamics (change of law over time) caused e.g. by amendments cannot be handled appropriately. Usually, a kind of monotonic

reasoning is used which "simulates" defeasible reasoning to some extent. However, this approach is very limited in use and can only capture a few aspects of legal reasoning.

Legal theory and approaches from AI and Law can improve computable models of law used in practice today. In the long run a flexible framework for building legal expert systems is needed which builds on open standards and implements best practices to foster reuse. Such a framework would also need to be complemented by a unified methodology for formalizing legal norms.

The main contribution of this research towards a common framework is the development of a solid temporal model which is capable of handling legal change in an efficient manner, e.g. determining applicable rules according to the temporal relations of the case. This supports the development of clean and well-structured models of law and, thus, decreases maintenance costs. The technical architecture of rOWLer follows a modular approach adhering to best practices[1] from software engineering and can be perceived as an extensible framework for building legal expert systems. This complements efforts to acquire an acknowledged standard for the rule layer of the semantic web cake.

3. Architecture

The architecture of rOWLer consists of three main layers complemented by an electronic document repository, namely the process layer, the rule layer and the ontological layer. What follows is a short overview of the architectural layers of rOWLer, each providing a different view on law and legal rules.

Process Layer: The process layer formalizes the legal procedure and is responsible to handle the dialog between the applicant and the public agency. It collects the relevant facts by automatic and manual means and interacts with the rule layer to continuously provide preliminary results until the final decision. The authorizing person is asked by the system for decision if a "hard" rule should be applied.

Rule Layer: This layer contains the formal rules and the inference engine. It drives legal reasoning by retrieving necessary information like facts from the ontology and providing results to the process layer above.

Ontological Layer: The ontological layer supports the layers above by shallow reasoning on the knowledge base staying within OWL 2, preparing it for more complex reasoning using rules, especially by data completion, reasoning on material circumstances (claims, facts and proofs) and legal concepts by deriving inferences.

Electronic Document Repository: This layer complements the formal model by providing access to electronic documents in Akoma Ntoso [22]. Entities of the other layers, e.g. rules, concepts, etc., can be linked by using IRIs with legal text. This allows for supporting the decision making by the legal expert by providing statutes, commentaries and judgments as well. Moreover it fosters isomorphism of rules by linking them with their legal source.

4. Reasoning Module and Algorithm

Technically, the algorithm is encapsulated in a module which integrates the reasoner with the rest of the system and also comprises the temporal model. This thin integration layer is also responsible for deriving the parameters from the facts necessary to call the engine, e.g. the significant date. Often, it

[1] This ensures more clean and maintainable code which – at the same time – is easier to understand and read.

is required to reason over complex situations which span a longer time period[2]. Such scenarios are handled by the reasoning module which interacts with the reasoner to achieve the overall conclusion.

In the following section, the proposed algorithm for reasoning is presented. It has to be mentioned that only a rather sketchy overview is given but no complete logical formalization is provided due to space restrictions.

Basically, the algorithm is divided into two separate steps to handle temporal and legal reasoning: (1) Determine which rules are applicable to a case at a certain point in time and (2) apply the rules determined in the first step to the case using defeasible reasoning.

The distinction between temporal reasoning and legal reasoning allows for a separate treatment of both problems. In technical terms, each of the steps is encapsulated using an interface with an independent implementation. This approach reduces the complexity of the algorithm by separating the whole problem into smaller pieces, independently of each other, while at the same time fostering better integration, maintenance and testing.

5. Temporal Model and Reasoning

5.1. Theoretical Background

There are several possibilities of the legislator to reduce effort and cost of legal change management [18]. Regardless of the methodology followed by the legislator, a computable model of law has to handle changes of sources of law.

For the purposes of the current model we follow the "direct method" of [18] and assume that each change of the sources of law (e.g. by an amendment) leads to a new consolidated version of a statute containing untouched, modified and new provisions as well. The old version of the statute and its norms enter out of force before the day the new versions enter into force. This approach reduces the complexity of the temporal model.

From a theoretical perspective, this may not fully convince as only some provisions are affected by change and thus enter out of force by implicit derogation. However, if the legislator enacted an authentic consolidated version of law no such objections exist, even from a theoretical point of view.

To handle change of law two aspects need to be considered: (1) A solid naming convention for statutes and rules and (2) a versioning model which formalizes the temporal dimensions of law.

Due to limited space, only the second aspect will be discussed in the next section. It should be mentioned that the naming convention used is aligned with FRBR [27] and a simplified version of the HTTP-based syntax for IRIs of Akoma Ntoso [5, 22], compliant with CEN MetaLex [3].

5.2. Versioning model

5.2.1. Temporal dimensions

According to legal theory the temporal model distinguishes the following temporal dimensions of legal norms (e.g. [19, 20])[3]:

[2] For example due to the ruling of the Austrian Supreme Court of Justice regarding continuing obligations the time before an amendment has to be judged according to the old rules and afterwards according to the new ones.

[3] It has to be noted, that the terms are not always used homogeneously in literature and are used with different meanings. The terms "efficacy" and "applicability" refer to "Bedingungsbereich" and "Rechtsfolgenbereich" respectively in German legal theory [31].

Existence: The period in which the norm is part of the legal system, starting with the day of publication (in an official journal), ending by a subsequent normative action.

Force: When the norm is in force and thus can be applied by the judge in general. In Austria, this period usually starts after the day of publication but can be deferred by *vacatio legis*.

Efficacy: The period in which facts must have occurred in order for the rule to be applicable is called the efficacy period.

Applicability:[4] This is the period when a legal norm produces the consequences it establishes.

Usually, the periods of force coincides with efficacy and applicability of a norm. However it is possible that the effects of a norm start before its force (retroactivity) or continue after the repeal (ultra-activity). For example, the tax law of 2008 should be applied to the income earned in 2008 (efficacy), even if a case should be decided after the 31st of December (applicability)[5].

5.2.2. Versioning legal rules

The versioning model used in rOWLer is based on [21] but has been slightly modified and extended to handle not only statutes (documents) but legal rules as well and also to be capable of determining the norms which are applicable to a case at a certain point in time.

The versions of a statute are ordered linearly in so called “versioning chains” by their date of enter into force. When a new version of law is enacted it is added at the end of the chain right after the last version. The model commits itself implicitly that the periods of force of two distinct provisions never overlap. This ensures the soundness of the linear ordering and the versioning chains.

It is assumed that the time when the changes are applied to the legal text coincides with the time of enter into force of the amended provision. Moreover the publication date of the amended provision is assumed to be the same as the amending provision and is also used as the official version date of the act.

5.2.3. Retroactive modifications

Following [21], to handle retroactive modifications the timeline has to be split virtually in the past creating a new legal situation which has not existed originally in this instant in time. To avoid major change of the temporal model in case of retroactive modifications, the proposed solution is enhanced and adopted to avoid splitting of versioning chains.

Each versioning chain is identified by the publication date of the retroactive modification which is the date from that the chain is valid and hence points at the “current” legal situation. When a retroactive modification arrives the current chain gets duplicated and the new chain contains the modified provisions starting after the retroactive change is applied.

[4] This refers strictly to temporal applicability, the derogation of norms, e.g. by EU law, is tackled in the second reasoning step of the proposed algorithm.

[5] For the example we assume that the fiscal year coincides with the calendar year.

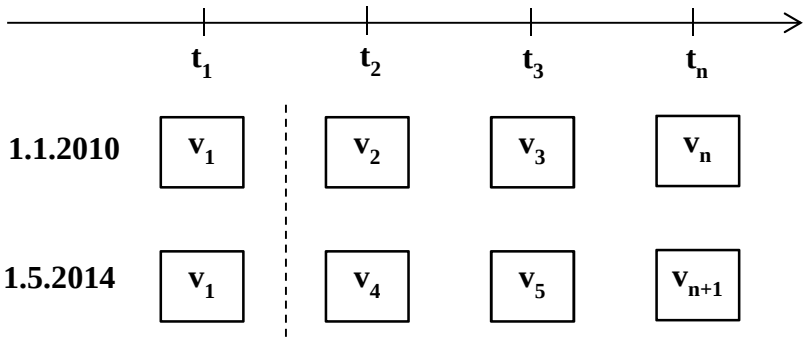


Figure 1. Example of retroactive modification (adapted from Palmirani and Brighi [21])

Figure 1 shows an example of an amendment published on January 5, 2014 which retroactively modifies v_2 at time t_1 and, thus, leads to a new versioning chain which contains the untouched v_1 followed by the amended versions. Virtually the timeline gets split after v_1 which is not affected by the modification, as indicated by the dashed line. There is no need to touch the existing chains. The retroactive change of v_2 subsequently leads to an adaption of the following versions as well, so we get the situation described above.

The versioning chains enable the reasoning engine to query the legal situation before and on (or after) the 1st of May 2014 when the retroactive amendment has been published and became part of the legal system. Further, it is possible to refer to the "current" legal situation by assigning a variable to the last chain. When the current chain needs to be put out of service due a retroactive modification the variable "current" simply refers to the new chain, without affecting the rest of the model. Unlike the approach of Palmirani and Brighi [21], there is no need to split existing chains in case of retroactive modifications.

5.3. Selecting applicable rules

Based on the versioning model and the reflections made in the previous sections a temporal reasoning engine has been designed which is able to compute the legal rules applicable to a case in a given time. In this step the algorithm deals with the "external" time of norms which guide the lifecycle of the provision and not the "internal" time which is expressed in the rule itself, e.g. when it is obligatory to use winter tires.

To figure out which rules are applicable to a case the engine needs to take the periods of efficacy and applicability into consideration. Accordingly, the temporal model needs to be queried with two dates: (1) The view point of the legal system and (2) the "significant" date of the case used to determine the applicable rules. The latter usually depends on the content and type of law (procedural or substantive law). For instance, in criminal law the date when the crime has been committed is significant and hence determines which version of law is applicable to a case.

The algorithm adopted by the temporal reasoning engine adheres to non-monotonic reasoning and roughly applies five steps to determine applicable rules: (1) Get existing norms at view date, (2) calculate temporal dimensions of norms, (3) determine applicable norms according to temporal facts of case, (4) handle suspension of norms and, finally, (5) resolve references by including referenced norms.

For the sake of brevity it is not possible to provide an in-depth discussion of this issue here. A few more details can be found in [28].

6. Personal Module

Despite the enormous size of the legal text corpora, law makers still assume that the designation of the affected addressees does not require particular focus. Context and the definitions part must be sufficiently clear. However, this is by far not sufficient. The personal scope of application must also be formalised. Clearly, human rights, civil law and criminal law are addressed as such to all. However, in business and administrative law, there are many rules and standards that are intended to concern only a selective group of addressees. *Chris Reed* describes this problem vividly in his contribution to the commemorative volume to Jon Bing [25].

The case in question is known as the Lindqvist case.[6] Mrs. Lindqvist created a website for their church in Alseda, Sweden. On the website, a note for quick recovery of a church member could be found. This was seen as data processing and Mrs. Lindqvist did not report properly to the Swedish Data Inspection Board. As a result, she was found guilty of two violations of data protection. The relevant question is clear. Why have individuals to deal with the highly complex commercial law rules? The answer is also evident but not satisfactory: The lawmaker does not explicitly exclude such cases.

Thus, we propose a personal module that explicitly describes the addressees of a rule.

6.1. Ontological Model of Persons

The aim of this model is to provide higher explicit knowledge of norm addressees, not a restriction of the law-making as such. Thus, many groups have to be identified. Further, Boolean combinations of such groups are very often required.

Persons: All natural and legal persons in this world.

Natural persons: All natural persons.

Legal persons: All legal persons.

Natural persons living in Austria: natural persons with a first residence in Austria.

Austrians: Natural persons with Austrian nationality.

EU citizens: Natural persons with a nationality of one of the 28 EU Member States.

EU resident companies: Enterprises with a legal seat or headquarters in the EU.

Minor children: Natural persons up the age of 18.

This short non-exhaustive list demonstrates that the personal regime in present legal systems is highly complicated. At least 100 different groups seem to exist. Legal dogmatics remains not much concerned with this ontological problem.

We propose an ontology identifying the various groups. Then each rule is supplemented by a time module as well as a personal module.

The advantage is obvious. Time and person are very easily and correctly identified by the addressees. Thus, the applicable rule corpus can be limited in size for the relevant groups, e.g. only citizens. As other representations than text, e.g. citizen information, visualisation, knowledge-based systems etc., are more and more important, this module allows a proper differentiation between the various groups. It must be mentioned that the attributes of the groups are objective criteria and should in no way be discriminatory.

[6] ECJ, case C 101/01, 6 November 2003, Report of Cases 2003 I-12971, ECLI:EU:C:2003:596..

Thus, law-makers have a proper tool box for appropriate identification of the addresses of the legal system as well as for the improvement of the conceptual framework of the statute.

7. Modelling norms

Following Kelsen [14], we assume in accordance with legal theory that norms have basically the following logical structure: If $A_1,\ldots,A_2$ then B; where "$A_1,\ldots,A_2$" are the conditions of the norm, "B" is the legal effect and "if…then" is a normative conditional.

Legal theory distinguishes basically between constitutive and prescriptive norms [33][7]. Constitutive norms provide definitions of the terms and concept used in a jurisdiction. On the other hand, the purpose of prescriptive norms is to regulate actions by making them obligatory, permitted or prohibited.

Besides the basic types of norms obligation, permission and prohibition, there are also more articulated effects, which follow from applying norms, e.g. those introduced by Hohfeld [11, 12][8].

Norms are formalized using rule objects[9] consisting of antecedent and consequent. Technically, rules are represented by an interface called `Rule` which provides methods for retrieving the `Antecedent` and `Consequent` of a rule as well. These are the basic buildings blocks for supporting different kinds of rules as introduced above.

To cope with different kind of rules and to allow for extension of the system, an abstraction, called `NormHandlingStrategy` is used, which itself makes use of two interfaces, namely `NormMatchingStrategy` and `NormExecutionStrategy`. The former is responsible for matching the antecedent conditions of a rule against the material facts of a case, whereby the latter is used to derive the conclusions following from a rule. This machinery allows for easily extending and tweaking the system, e.g. to support special types of norms.

7.1. Isomorphism

To foster isomorphism [2, 23], each rule is connected to its legal source by implementing a special interface, called `IsomorphicLegalEntity`.

This interface can be used to retrieve appropriate IRIs which point at the legally binding provision. Sometimes, several provisions are represented by a single rule or a single provision contains multiple rules, hence multiple IRIs are supported.

The model makes no assumptions about the format of the IRIs used, this allows for adapting the used references to the requirements of the particular use case.

7.2. Presenting Rule Priorities

In law, we have to deal with implicit (lex specialis, lex posterior) and explicit exceptions between norms. A computable model of law must be able to represent both kinds of exceptions to reflect the way statues are usually written, organized in general rules and exceptions.

In AI and Law different methods to solve conflicts between rules have been proposed, namely *specificity* [23, 30], *weight* (salience) [32] and *preference relation* [13, 15]. The model of rOWLer sup-

[7] Gordon et al. [9] identify more types of rules, but most of them can be reduced to the two types described above, because only their meaning differs from each other and not their logical structure.
[8] A rather comprehensive list of normative effects can be found in [26].
[9] We assume an object-oriented model.

ports weights and preference relations by using interfaces `WeightedRule` and `PreferenceRelation` respectively.

Conflicts between rules are resolved by ordering rules using an implementation of `RuleOrderingStrategy`. The strategy inspects all rules to order the rules supporting all of the methods above, using explicit and implicit information as well. The rules are placed in a network representing their ranking and wrapped by a dynamic proxy[10] at runtime implementing `SuperiorityRelation`.

A `SuperiorityRelation` represents an abstract concept describing the binary relationship between two rules[11], covering specificity, weight and preference relation as well. This abstraction allows for a dynamic creation of arbitrary relations between rules, e.g. of *lex superior* and *lex inferior* by inspecting the law making institutions modelled in the ontology and linked with the rules.

The model enhanced with superiority relations between rules builds the foundation for qualifying the rules as defeater, defeasible and strict in the sense of defeasible logic [17]. Further, it enables the use of a defeasible engine like SPINdle [15] for reasoning or the implementation of a custom engine built on an algorithm like [16].

7.3. Personal scope

In section 6 we argued for the urgent need to take the personal scope of norms into consideration. Here we show how that considerations translate into an object-oriented model.

Technically, the personal scope of a norm can be represented by an interface which is called `PersonallyScopedRule`. This abstraction provides a single method which returns the addressee(s) of a norm. An "addressee" in this context is an object which describes an arbitrary legal entity, e.g. a single natural person, or a group of such entities to which the particular norm should be applied in general.

With that information in place, the first step of the proposed algorithm, discussed in section 5.3. , can be extended to filter the overall set of norms which could be applied in general, by using personal scope, besides temporal aspects, of a particular norm as well.

The interested reader is referred to [29], especially for more details, on how metadata, different interpretations of norms, references between norms and contrary-to-duty obligations [24] are handled by the formal model.

8. Related Work

JBoss Drools[12] is an open-source business rule engine and as such uses production rules as data structure. Since version 6 it is based on "PHREAK" a monotonic algorithm supporting forward and backward chaining.

Although drools performs well with thousands of rules and has a nice declarative style for writing rules, it is not suited for the legal domain. First of all, it only supports monotonic reasoning and thus cannot deal with incomplete information. Defeasible reasoning can be "simulated" to some extend by using attribute "salience" on rules, determining rule order and hence allowing for representing priority relations between rules. However, this approach only captures a few aspects of defeasibility

[10] The architecture of rOWLer is consistently based on interfaces which allows for using Java's dynamic proxying facilities.

[11] In this a sense superiority relation resembles a preference relation but in contrast to the latter it is an abstraction whose instances are built dynamically at runtime by the engine.

[12] By referring to "Drools" we actually mean "Drools Expert", which is the rule engine of the Drools platform.

in law. Second, the time model of Drools does not support the temporal dimensions of law and thus would have to be extended to handle legal change over time. Compared to Drools, rOWLer adheres to defeasible reasoning and its temporal model is well suited for the legal domain.

SPINdle [15] is another open-source rule engine which supports defeasible logic and modal defeasible logic as well. Unlike Drools, which is based on a monotonic algorithm, it is capable of defeasible reasoning over theories with thousands of rules. SPINdle gives basic support for time and intervals but cannot handle the temporal dimensions of legal norms. rOWLer is built on a sophisticated versioning model supporting temporal reasoning to determine applicable provisions.

The rules in SPINDle are heavily based on literals. Basically, the conclusion of a rule is a literal or its negation. To formalize norms we need a representation of a rule which allows for representing richer conclusions, e.g. a calculation or the inclusion of other norms in case of references. Therefore, rOWLer supports a richer object model supporting different kind of rules which are executed by using an appropriate strategy. However, it would be convenient if SPINdle could be used as defeasible rule engine embedded inside rOWLer.

Another system worth mentioning is Carneades [7, 8] which provides support for constructing, evaluating and visualizing arguments. Arguments can be constructed from ontologies, rules and cases. Carneades is based on a formal model of argumentation, designed for supporting real-world argumentation in practice. At the moment Carneades lacks support for the temporal dimensions of law and deontic operators [4]. The purpose of rOWLer is limited to apply simple rules to cases but builds on a sophisticated temporal model that copes with legal change over time.

9. Conclusions and Future Work

Compared to present approaches in public administration, rOWLer is aligned with legal theory and fosters defeasible reasoning while maintaining isomorphism with the sources of law. To cope with legal change over time a solid temporal model has been developed with formalizes the temporal dimensions of law and, further, is able to decide which norms should be applied to a case at a certain point in time. By using a viewpoint the model is also capable of handling retroactive modification by providing the historic and current version of a statute after the amendment. Present implementations used in practice lack a sophisticated temporal model for handling legal change which increases code complexity and leads to severe maintenance problems.

At the moment, rOWLer is designed as a single-agent system and the reasoning engine is optimized to deal with statutes with a rather mathematical content like tax law or “easy” cases[13] in the terminology of Hart. The model of rOWLer is flexible enough to be extended in the future to handle “hard” cased as well, e.g. by providing the legal expert with different alternatives for decision making and integrating more sophisticated argumentation systems like Carneades [6].

In the future, the conceptual model needs to be refined, especially with regard to the representation of norms and defeasible reasoning. Feasibility of the theoretical approach should be evaluated by developing a prototype in Java which has become the “mainstream” programming language nowadays.

[13] “Easy” cases can be largely decided “mechanically” by deducing the required result from the rule and the facts. “Hard” cases are ones for Hart in which the facts fall within the “penumbra” of the meaning of the words in the applicable rule. These cases require the judge to exercise discretion [10].

10.References

[1] Athan, T. et al. 2013. OASIS LegalRuleML. *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Law - ICAIL '13* (New York, USA, 2013), 3–12.

[2] Bench-Capon, T.J.M. und Coenen, F.P. 1992. Isomorphism and legal knowledge based systems. *Artificial Intelligence and Law*. 1, 1 (1992), 65–86.

[3] Boer, A. und van Engers, T.M. 2011. A MetaLex and Metadata Primer: Concepts, Use, and Implementation. *Legislative XML for the Semantic Web*. G. Sartor et al., Hrsg. Springer Netherlands. 131–149.

[4] Ceci, M. 2013. *Interpreting Judgements using Knowledge Representation Methods and Computational Models of Argument*. University of Bologna.

[5] Francesconi, E. 2011. Naming Legislative Resources. *Legislative XML for the Semantic Web*. G. Sartor et al., Hrsg. Springer Netherlands. 49–74.

[6] Gordon, T.F. 2010. An Overview of the Carneades Argumentation Support System. *Dialectics, Dialogue and Argumentation. An Examination of Douglas Walton's Theories of Reasoning*. C. Reed und C.W. Tindale, Hrsg. College Publications. 145–156.

[7] Gordon, T.F. 2011. Analyzing open source license compatibility issues with Carneades. *Proceedings of the 13th International Conference on Artificial Intelligence and Law - ICAIL '11* (New York, 2011), 51–55.

[8] Gordon, T.F. 2011. Combining Rules and Ontologies with Carneades. *Proceedings of the 5th International RuleML2011@ BRF Challenge* (Fort Lauderdale, Florida, USA, 2011).

[9] Gordon, T.F. et al. 2009. Rules and Norms: Requirements for Rule Interchange Languages in the Legal Domain. *Rule Interchange and Applications*. G. Governatori et al., Hrsg. Springer. 282–296.

[10] Hart, H.L.A. 1994. *The Concept of Law*. Clarendon Press.

[11] Hohfeld, W.N. 1917. Fundamental Legal Conceptions as Applied in Judicial Reasoning. *The Yale Law Journal*. 26, 8 (Juni 1917), 710.

[12] Hohfeld, W.N. 1913. Some Fundamental Legal Conceptions as Applied in Judicial Reasoning. *The Yale Law Journal*. 23, 1 (Nov. 1913), 16.

[13] Johnston, B. und Governatori, G. 2003. Induction of defeasible logic theories in the legal domain. *Proceedings of the 9th international conference on Artificial intelligence and law - ICAIL '03* (New York, New York, USA, 2003), 204.

[14] Kelsen, H. 1979. *Allgemeine Theorie der Normen*. Manz.

[15] Lam, H.-P. und Governatori, G. 2009. The Making of SPINdle. *Rule Interchange and Applications*. G. Governatori et al., Hrsg. Springer. 315–322.

[16] Maher, M.J. 2004. Propositional defeasible logic has linear complexity. *Theory and Practice of Logic Programming*. 1, 06 (Juni 2004), 691–711.

[17] Nute, D. 2003. Defeasible Logic. *Web Knowledge Management and Decision Support*. O. Bartenstein et al., Hrsg. Springer Berlin Heidelberg. 151–169.

[18] Palmirani, M. 2011. Legislative Change Management with Akoma-Ntoso. *Legislative XML for the Semantic Web*. G. Sartor et al., Hrsg. Springer Netherlands. 101–130.

[19] Palmirani, M. et al. 2011. Modelling temporal legal rules. *Proceedings of the 13th International Conference on Artificial Intelligence and Law - ICAIL '11* (New York, 2011), 131–135.

[20] Palmirani, M. et al. 2010. Temporal Dimensions in Rules Modelling. *Legal Knowledge and Information Systems, JURIX 2010: The Twenty-Third Annual Conference* (Amsterdam, 2010), 159 – 162.

[21] Palmirani, M. und Brighi, R. 2006. Time Model for Managing the Dynamic of Normative System. *Electronic Government*. M.A. Wimmer et al., Hrsg. Springer Berlin Heidelberg. 207–218.

[22] Palmirani, M. und Vitali, F. 2011. Akoma-Ntoso for Legal Documents. *Legislative XML for the Semantic Web*. G. Sartor et al., Hrsg. Springer Netherlands. 75–100.

[23] Prakken, H. und Schrickx, J. 1991. Isomorphic models for rules and exceptions in legislation. *Legal knowledge based systems JURIX 91: The Foundation for Legal Knowledge Systems* (Lelystad, 1991), 17–27.

[24] Prakken, H. und Sergot, M. 1996. Contrary-to-duty obligations. *Studia Logica*. 57, 1 (Juli 1996), 91–115.

[25] Reed, C. 2014. You Talkin' to Me? *Jon Bing – en hyllest/a tribute*. D.W. Schartum et al., Hrsg. Gyldendal. 154–171.

[26] Sartor, G. 2006. Fundamental legal concepts: A formal and teleological characterisation. *Artificial Intelligence and Law*. 14, 1-2 (Sep. 2006), 101–142.

[27] Saur, K.G. 2009. *Functional Requirements for Bibliographic Records: Final report.* IFLA Study Group on the Functional Requirements for Bibliographic Records.

[28] Scharf, J. 2014. rOWLer - A Hybrid Rule Engine for Legal Reasoning. *Proceedings of the Semantic Web for the Law and Second Jurix Doctoral Consortium Workshops* (Kraków, 2014).

[29] Scharf, J. 2015. *Wissensrepräsentation und automatisierte Entscheidungsfindung am Beispiel des Kriegsopferversorgungsgesetzes*. PhD thesis, Universität Wien.

[30] Van de Ven, S. et al. 2008. Automated Legal Assessment in OWL 2. *Legal Knowledge and Information Systems - JURIX 2008: The Twenty-First Annual Conference on Legal Knowledge and Information Systems* (Amsterdam, 2008), 170 – 175.

[31] Walter, R. et al. 2007. *Grundriss des österreichischen Bundesverfassungsrechts*. Manzsche Verlags- und Universitätsbuchhandlung.

[32] Walton, D. und Gordon, T.F. 2012. The Carneades model of argument invention. *Pragmatics & Cognition*. 20, 1 (2012), 1–31.

[33] Von Wright, G.H. 1963. *Norm and Action*. Routledge.

SEMANTIC MODELING OF THE EU MULTILINGUAL RESOURCES

Enrico Francesconi[1], Mark W. Küster[1]
Patrick Gratz[2], Sebastian Thelen[2]

[1] Publications Office of the EU, 2 rue Mercier, L-2985 Luxembourg, LU
{enrico.francesconi, mark.kuster}@publications.europa.eu
[2]infeurope S.A., 62 rue Charles Martel, L-2134 Luxembourg, LU
{patrick.gratz, sebastian.thelen}@infeurope.lu

Keywords: ***Multilingual documents, Semantic indexing, Knowledge modeling***
Abstract: *The Publications Office of the European Union has the responsibility to make available and disseminate the official publications and bibliographic resources produced by the institutions of the European Union. The central component of its information system is the CELLAR repository, providing semantic indexing, advanced search and data retrieval for multilingual resources. This paper gives an overview of the semantic modeling approach for CELLAR, based on semantic web technologies. Moreover, a proposal for a possible evolution aiming to improve the modularity and facilitating the general management of the model is shown.*

1. Introduction

A multilingual access to the documents produced by the European institutions is a fundamental democratic right for citizens. The Publications Office of the European Union is the inter-institutional structure of the European Commission in charge of making available and disseminating the official publications of the EU and other bibliographic resources.

As for official publications[1], this right is guaranteed by their availability in the 24 official languages spoken in the 28 Member States of the European Union, while other publications (like tendering documents, general publications and information on EU-funded research projects) are mainly available in the 3 working languages (English, French and German).

With the authentic and legally binding publication of the electronic edition of the Official Journal (e-OJ) from 1 July 2013, the on-line accessibility of legal resources has become an essential requirement, guaranteed by the EUR-Lex service[2].

In the last couple of years most efforts of the Publications Office (OP) were focusing on a project aiming to transform the archival and dissemination architecture, based on different systems, into a federative architecture based on a common archival service, providing also a common interface for disseminating materials to the users (Fig. 1).

[1]Treaties, International agreements, Legislation, Complementary legislation, Preparatory acts, Case-law, National implementing measures, References to national case-law concerning EU law, Parliamentary questions, Consolidated legislation, Other documents published in the Official Journal C series, EFTA documents

[2]http://eur-lex.europa.eu

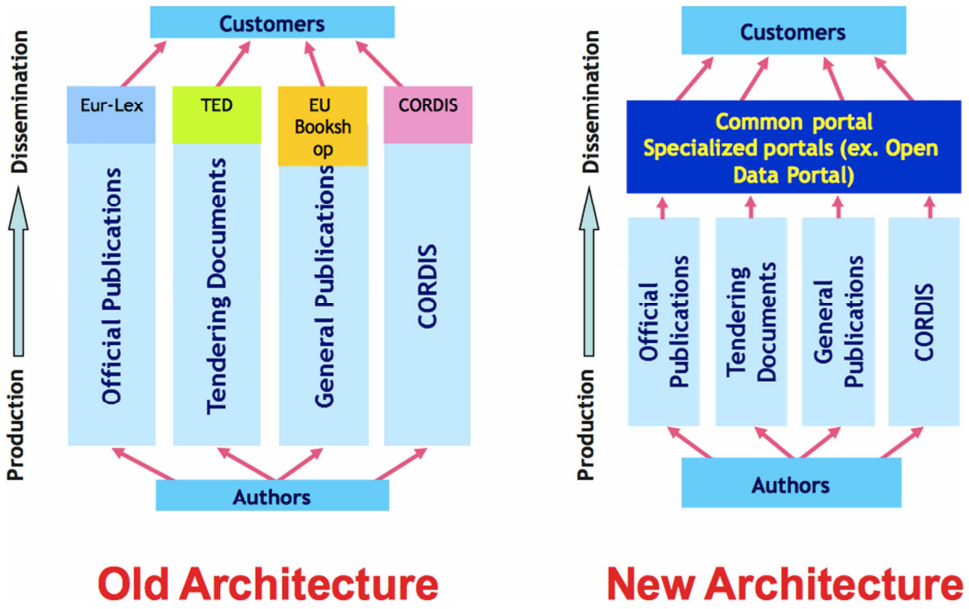


Figure 1: The Publications Office archival and dissemination transformation programme

The central component of this architecture is CELLAR, a content and metadata repository containing documents coming from the production and postproduction services (including content validation and metadata production). They are available for long term preservation, open data, indexing, as well as advanced search and retrieval services. CELLAR resources are semantically described by an ontology, which represents the Common Metadata Model (CDM) of the OP resources.

This paper is focused on the description of this ontology as well as on possible developments. It is organized as follows: in Section 2 the architecture of the CELLAR platform is described; in Section 3 the CELLAR multilingual semantic approach, represented by CDM, is presented; in Section 4 a possible evolution of CDM is illustrated and in Section 5 the advantages of the proposed evolution are discussed. Finally in Section 6 some conclusions are reported.

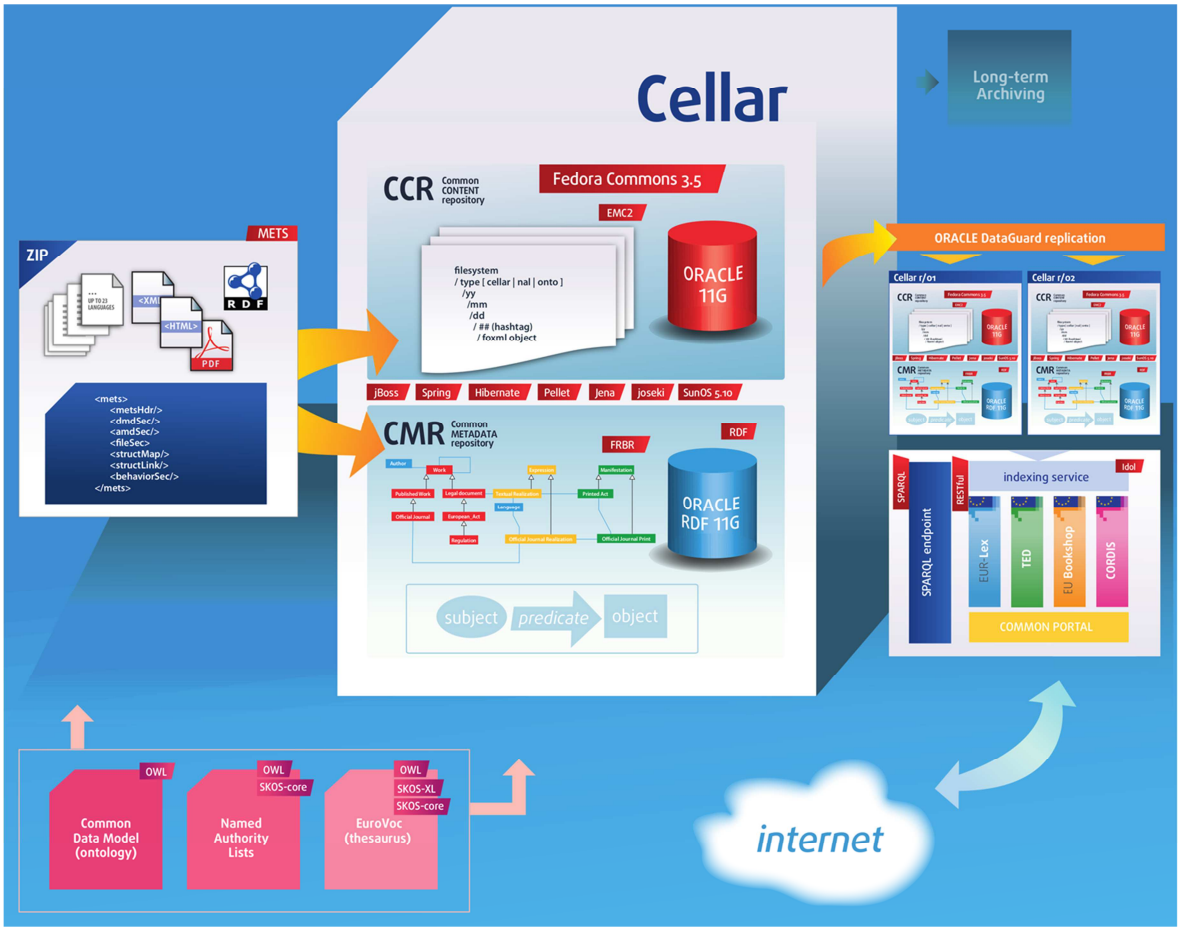
2. The CELLAR architecture

CELLAR represents the central hub of the whole information system of the OP. It is based on a Fedora digital objects repository[3], organized in two logical units including Oracle database technologies: content is stored in the CELLAR Common Content Repository (CCR) currently[4] including about 152 million documents in 24 languages; metadata in as many languages are stored in the CELLAR Common Metadata Repository (CMR) described by semantic web technologies, resulting in about 1100 million triples, stored in an RDF triple store. Currently CELLAR receives about 5 million requests per day, providing information results for the EUR-Lex service and for the SPARQL endpoint service recently exposed in order to complement linked open data services to potential consumers.

Other services and types of resources, like TED for tendering documents, EU Bookshop for general publications, CORDIS for information on EU-funded research projects will be served by CELLAR in the near future (Fig. 2).

[3]http://www.fedora-commons.org

[4]March 2015



Concerning disaster recovery and emergency management a proper data replication service for the production database has been put in place as shown in Fig. 2.

Figure 2: The CELLAR architecture and services5

Based on CDM, the Common Metadata Repository (CMR) represents the essential asset to guarantee multilingual semantic access services to the CELLAR contents.

The following section depicts how the CDM allows to describe, from a semantic point of view, all the OP resources.

3. Common Metadata Model (CDM)

The current CDM is an ontology based on the FRBR[6] model (3), described by RDF(S)/OWL technologies, able to represent the relationships between the resource types managed by the OP and their views according to the FRBR model in terms of Work, Expression, Manifestation and Item.

In the current CDM organization, the FRBR hierarchy represents a sort of pivot knowledge organization system, according to which resource types (general publications, legal resources, legislation, case law, etc.) and FRBR views (ex: general publication expression, case law expression, official journal manifestation, etc.) are organized through sub-class relationships (Fig. 3).

[5]courtesy F. Sanmartin

[6]Functional Requirements for Bibliographic Records

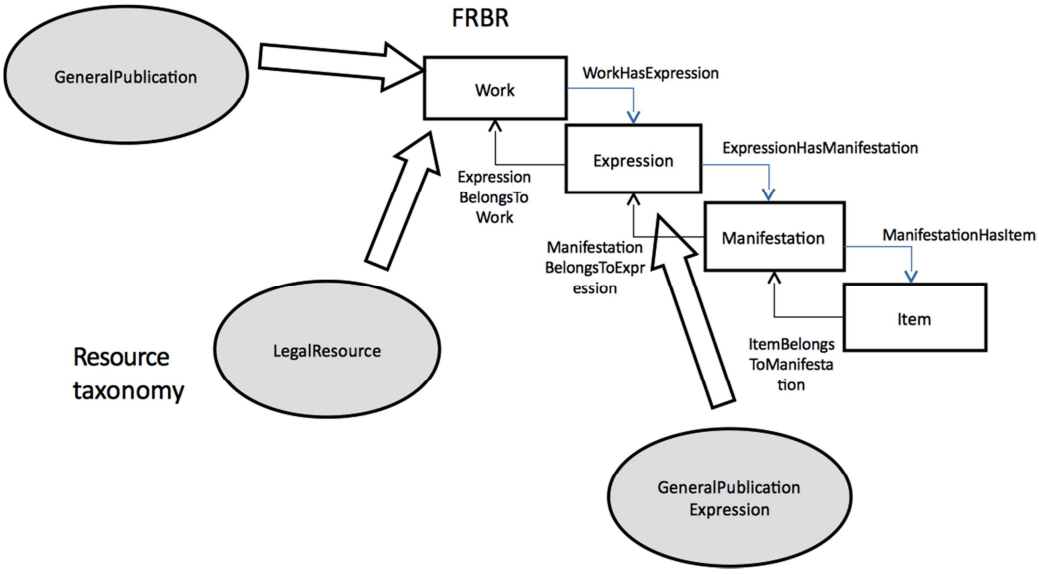


Figure 3: The current CDM organization system

Resources are identified by URIs classified according to the FRBR hierarchy, thus organizing the objects managed by CELLAR at different FRBR abstraction levels.

Such URIs have for example `http://publications.europa.eu/resource/oj/` as namespace of the official journal resources, followed by an ID created as concatenation of metadata values at each FRBR level (see Tab. 1 for some examples).

Resource FRBR type	**Resource ID**
Work (Regular OJ publication n. 26 of 2015)	JOL_2015_026_R_0001
Expression (English variation)	JOL_2015_026_R_0001.ENG
Manifestation (PDF/A-1a format)	JOL_2015_026_R_0001.FRA.pdfa1
Item	JOL_2015_026_R_0001.FRA.pdfa1a.l_02620150131fr00010002.pdf

Table 1: URIs at different FRBR levels

Based on commonly known best practices for linked data, CELLAR enables clients to retrieve various resource representations via content negotiation. Each resource that represents an entity of the FRBR hierarchy can be considered a non-information resource, redirecting clients to one of its specific representations based on parameters (i.e., Accept header, Accept-Language, etc.) passed in the request.

Of all existing representations, the so called tree notice of a FRBR hierarchy is the representation that best describes the bibliographic record from a "web of data" point of view, since it provides the entire set of metadata at each level of the FRBR hierarchy in a single RDF serialization. In CELLAR, tree notice URIs follow the pattern `cellar:[hash-value]/rdf/tree/full`[7]

[7]where `cellar:` represents the CELLAR namespace
`http://publications.europa.eu/resource/cellar/`

This CDM version is currently in production providing detailed views, in particular regarding language versions and formats, of the OP resources, for both documents and metadata search and retrieval services, as well as for the OP common portal.

In the context of a recent activity a review of the current CDM in order to reduce complexity of the query framework was undertaken. During this review the following shortcomings of the current model have been revealed: 1) the mixture in the same taxonomy of resource types and FRBR classes, 2) the need to follow complex paths to reach different FRBR views of the same resource type (see General Publication type: GeneralPublication → Work → Expression → GeneralPublication-Expression).

These issues result in certain limitations of the framework. For instance that, given a resource type, the access to the different levels of the FRBR model is not direct. Moreover, it is necessary to know the type of a resource at query level in order to retrieve metadata at each level of the FRBR model, while it would be more simple that, given a resource, there is a common query to access metadata at different FRBR levels, irrespective of the resource type.

In the next sections an overview of the current discussion about a possible CDM evolvement is presented.

4. Proposal for CDM evolvement

A proposal for possible evolvement of the current CDM approach aims firstly to keep a distinction between the taxonomy of the resources and the FRBR model.

A *Resource* in the ISBD[8] sense is defined as "*an entity, tangible or intangible, that comprises intellectual and/or artistic content and is conceived, produced and/or issued as a unit, forming the basis of a single bibliographic description"*. Therefore, resources are actually not equivalent to, or subclass of, any individual FRBR classes (1).

As pointed out in (2) each FRBR classes *reflects* one aspect of a resource, seen as a bibliographic entity at different levels of abstraction.

A *Resource* (in the ISBD sense) has the same intention as the combined attributes of the FRBR model (2), therefore it can be considered as the disjoint union of the Work, Expression, Manifestation and Item levels in FRBR model, as expressed by:

$$Resource = Work + Expression + Manifestation + Item$$

The relationship between the two domains (resource taxonomy and FRBR model) is therefore of *part-of*/*aspect*. In this context, every FRBR level is an *aspect* of a current resource and can be considered as collector of the metadata able to describe a resource at that level.

Therefore, a resource and its FRBR model can be viewed as aspects of the same reality in two perspectives (1):

1. The "web of data" perspective
2. The "bibliographic data" perspective

A resource identified by a specific URI represents an entity of the "web of data". The resources published by the OP are basically bibliographic entities. Therefore, they can be described according to the FRBR model.

[8]International Standard Bibliographic Description

Works, Expressions, Manifestations and Items of the FRBR model are also type of entities of the web of data, but they can also be viewed as a specific aspects of a bibliographic resource, therefore viewed in the "bibliographic data" perspective.

This distinction provides the main motivation for improving CDM with the goal to simplify the query framework, thus improving the accessibility of the resources. To achieve this goal, the following actions have been undertaken:

1. Introduction of a logical separation between the taxonomy of the OP resources and the FRBR model, therefore avoiding any subClass relations between them;

2. Introduction of `cdm:has[FrbrClass]Aspect`[9] relations between a classes of the OP resource taxonomy and their aspects as FRBR classes (e.g.: `cdm:hasWorkAspect`, `cdm:hasExpressionAspect`, etc.);

3. Introduction of a `rdfs:subPropertyOf` relation between `cdm:has-[ResourceTypeFrbrClass]Aspect` at different levels of the taxonomies.

In Fig. 4 a sketch of the OP resource taxonomy (limited, for simplicity, to the root and one subclass) and its relationships with the FRBR model at each taxonomy level is represented. In particular the generic class of `OPBibliographicResource` is linked with `cdm:hasWorkAspect`, `cdm:hasExpressionAspect`, `cdm:hasManifestationAspect`, `cdm:hasItem-Aspect` to the corresponding classes of the FRBR model.

Sub classes in the resource taxonomy, like `SourceOfLaw`, are linked to the corresponding classes of the FRBR model with similar specific properties (as `cdm:hasSourceOfLawWorkAspect`, `cdm:hasSourceOfLawExpressionAspect`, etc.). Such "aspect" properties are organized in pure taxonomic relationships (`subPropertyOf`) for each level of the FRBR model (`cdm:hasSourceOfLawWorkAspect` is a sub property of `cdm:hasWorkAspect`, and so on).

[9] where `cdm:` is the CDM namespace and `[FrbrClass]` is one of Work, Expression, Manifestation or Item classes

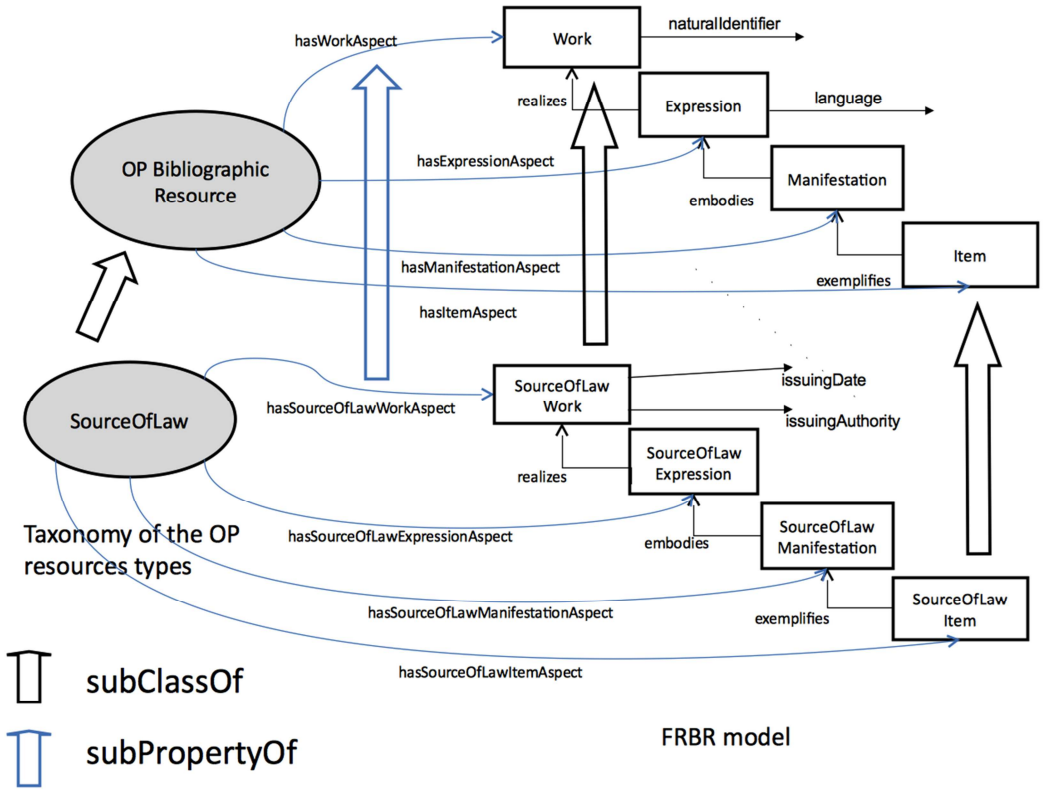


Figure 4: A proposal for CDM organization system development

The FRBR classes are collectors of resource metadata at their specific taxonomy level: for example (see Fig. 4) at `Work` aspect level, a resource will have for example `cdm:naturalIdentifier` as generic metadata, described by object or datatype properties, shared by all the OP bibliographic resources. Similarly at `SourceOfLawWork` aspect level, specific metadata shared by all the sources of law, are given, as for example `cdm:issuingDate` and `cdm:issuingAuthority` of a legal measure. The same holds for the other FRBR classes at each level of the OP resource taxonomy.

In this CDM model, CELLAR tree notice URIs are the identifiers of the resource taxonomy entities linked with `cdm:has[ResourceTypeFrbrClass]Aspect` relation to the FRBR classes, since they provide the entire set of FRBR metadata in a single RDF serialization (an excerpt of it for a `SourceOfLaw` is the following[10]):

```
<rdf:Description rdf:about="cellar:58da3a99-a91d-11e4-8e01-01aa75ed71a1/rdf/tree/full">
 <rdf:type rdf:resource="cdm:SourceOfLaw"/>
 <cdm:hasSourceOfLawWorkAspect rdf:resource="ojns:JOL_2015_026_R_0001"/>
 <cdm:hasSourceOfLawExpressionAspect rdf:resource="ojns:JOL_2015_026_R_0001.ENG"/>
 <cdm:hasSourceOfLawExpressionAspect rdf:resource="ojns:JOL_2015_026_R_0001.FRA"/>
 <cdm:hasSourceOfLawManifestationAspect rdf:resource="ojns:JOL_2015_026_R.ENG.pdfa1a"/>
 <cdm:hasSourceOfLawManifestationAspect rdf:resource="ojns:JOL_2015_026_R.FRA.pdfa1a"/>
 <cdm:hasSourceOfLawItemAspect
      rdf:resource="ojns:JOL_2015_026_R_0001.ENG.pdfa1a.l_02620150131en00010002.pdf"/>
 <cdm:hasSourceOfLawItemAspect
      rdf:resource="ojns:JOL_2015_026_R_0001.FRA.pdfa1a.l_02620150131fr00010002.pdf"/>
```

[10] `ojns:` is the namespace `http://publications.europa.eu/resource/oj/`

```
</rdf:Description>
```

Moreover, the `SourceOfLaw` in the previous example has metadata (properties) related to its corresponding FRBR aspects, as well as the metadata of the FRBR aspects of its super classes. An excerpt of its metadata at its Work level is the following:

```
<rdf:Description rdf:about="ojns:JOL_2015_026_R_0001">
 <rdf:type rdf:resource="cdm:SourceOfLawWork"/>
 <cdm:naturalIdentifier>L 26/1</cdm:naturalIdentifier>
 <cdm:issuingDate rdf:datatype="&xsd;dateTime">2015-01-26T00:00:00</cdm:issuingDate>
 <cdm:issuingAuthority>Council of the European Union</cdm:issuingAuthority>
</rdf:Description>
```

The described approach has been implemented as proof of concepts in RDF(S)/OWL, resulting in the OWL-DL profile, thus available for deriving inferences by using DL reasoners like Pellet[11] or HermiT[12].

5. Benefits of the approach

The proposed CDM modeling approach has several advantages with respect to the existing one.

First of all it allows a direct constant access to the FRBR levels through the properties `cdm:has[ResourceTypeFrbrClass]Aspect`, while in the existing CDM the FRBR levels have to be navigated until reaching the expected one. In the existing CDM in fact there is no resource in the metadata that identifies the actual bibliographic entity (`SourceOfLaw`), therefore the resource can be either a Work, Expression, Manifestation or Item. Consequently a complex property path is necessary to navigate to the suitable FRBR entity and, in order to access to the Expression of a `SourceOfLaw`, for example, the following query is needed:

```
SELECT ?uri WHERE
{
  ?resource  cdm:item_belongs_to_manifestation?/
  cdm:manifestation_manifests_expression?/
  cdm:expression_belongs_to_work?/
  ^cdm:expression_belongs_to_work ?uri
}
```

On the contrary, in the new model the same result can be obtained by the following, more simple query:

```
SELECT ?uri WHERE
{
   ?resource rdf:type cdm:SourceOfLaw .
   ?resource cdm:hasSourceOfLawExpressionAspect ?uri
}
```

A similar query can be created to access all the FRBR aspects of an OP resource.

Another important advantage of this architecture is that the queries for retrieving metadata of a resource are independent of its resource type. In fact, the inheritance mechanism on properties allows

[11] `http://clarkparsia.com/pellet/`

[12] `http://hermit-reasoner.com`

us to express queries at the top level of the hierarchy, independently from the resource type, while in the existing model the resource type is to be known to retrieve its metadata, having as many queries as resource types. For example, in the new model given the `SourceOfLaw` URI considered in Section 4, the following query will retrieve the URI of its Work aspect:

```
SELECT DISTINCT ?uri WHERE
{
  cellar:58da3a99-a91d-11e4-8e01-01aa75ed71a1/rdf/tree/full
  cdm:hasWorkAspect ?uri
}
```

Note that this query does not contain any reference to the type of the OP resource, therefore it is valid for every type of resources and the query framework of the system becomes more simple. Similarly, a query able to retrieve the Expression aspect in the English language is the following:

```
SELECT DISTINCT ?uri WHERE
{
  cellar:58da3a99-a91d-11e4-8e01-01aa75ed71a1/rdf/tree/full
  cdm:hasExpressionAspect ?uri .
  ?uri cdm:language "en"^^xsd:string
}
```

Also in this case no reference to the type of OP resource is contained in the query.

An additional advantage of this modeling approach is the possibility to obtain a simplified management of the resource metadata, since they are organized in terms of properties of the FRBR classes, distributed at different levels of the resource taxonomy. This allows us, for example, to query the CDM model asking for all the Work metadata (i.e. `owl:DatatypeProperties`) of a generic `SourceOfLaw`, as follows:

```
SELECT DISTINCT ?property WHERE
{
   ?property rdf:type owl:DatatypeProperty .
   ?property rdfs:domain ?class .
   cdm:SourceOfLawWork rdfs:subClassOf* ?class
}
```

or to query the CDM model just selecting the specific metadata at `SourceOfLawWork` level:

```
SELECT DISTINCT ?property WHERE
{
   ?property rdf:type owl:DatatypeProperty .
   ?property rdfs:domain cdm:SourceOfLawWork
}
```

6. Conclusions

CELLAR represents the central information system of the OP, providing storage as well as advanced semantic indexing and access facilities to all the dissemination portals.

The CDM semantic approach for the CELLAR resources is able to greatly improve accessibility of the OP multilingual documents.

The proposed revision of the current CDM architecture, in particular, has the benefit of providing modularity and flexibility to the CDM approach, thus facilitating the management and extension of such knowledge organization system, as well as to simplify the query framework.

7. References

(1) *Bianchini, Carlo / Willer, Mirna* ISBD resource and its description in the context of the semantic web - In Cataloging & Classification Quarterly, *52:869–887, 2014.*

(2) *Dunsire, Gordon,* Resource and work, expression, manifestation, item. - *Amended October 6, 2013, following comments by Patrick Le Boeuf and discussion at IFLA 2013, July 28 2013.*

(3) *Study group on IFLA.* Functional requirements for bibliographic records - Technical report, International Federation of Library Associations and Institutions, 1998. http://www.ifla.org/VII/s13/frbr/frbr.pdf.

SUPPORTING THE LEGAL SUBSUMPTION PROCESS: DETERMINATION OF CONCRETENESS AND ABSTRACTNESS IN GERMAN LAWS USING LEXICAL KNOWLEDGE

Bernhard Waltl[1], Florian Matthes[2]

[1]Research Associate, Technische Universität München, Department of Informatics,
Software Engineering for Business Information Systems
Boltzmannstraße 3 85748 Garching bei München, DE
b.waltl@tum.de; https://wwwmatthes.in.tum.de/
[2]Professor, Technische Universität München, Department of Informatics,
Software Engineering for Business Information Systems
Boltzmannstraße 3, 85748 Garching bei München, DE
matthes@in.tum.de; https://wwwmatthes.in.tum.de/

Keywords: ***German Laws, Subsumption, Lexical Knowledge, GermaNet***

Abstract: *Determining, whether an act is applicable or not, is a non-trivial task. This is strongly associated with the interpretation of acts and the subsumed objects. Although subsumption is a complex process, it is also well-studied process and a central part of the legal theory and practice. Words are used as a base line during subsumption and allow for taxonomic structuring amongst itself, using hyper- and hyponym relationships. E.g., the word "energy" is a hypernym to "electricity". This paper determines the application scope of acts by accessing real-world knowledge stored in a German lexical knowledge database, called GermaNet. Based on the set of the ten largest German law texts we determine the average level of abstractness over a huge set of norms. Our research shows that words used in German acts are either a very high or a very low abstractness. Furthermore, we compared highly related laws from distinct countries, namely Austria and Germany, namely the act governing the liability for a defective product. We are able to automatically determine differences in the application scope of acts, respectively norms.*

1. Introduction

Subsumption is a fundamental process in the legal domain. It is necessary to determine whether case facts are within the scope of a particular legal act or not. The logic reasoning process behind the subsumption process in the legal domain is the well-studied syllogism (Larenz, Canaris 1995; Raabe et al. 2012). Syllogism is one kind of logical argument, which reasoning nature is deductive. The base line for the reasoning are two asserted true propositions. One proposition is the major premise, whereas the second is the minor premise. A famous example is about the mortality of people. Knowing that all people are mortal (major premise), and knowing that men are people (minor premise), the syllogism now allows us to make the logical conclusion, that all men are mortal. The so-called "middle (M)", namely "people", is the key, connecting the major and minor premise. Consequently, if a reasoner decides about the mortality of something, he could automatically decide for everything that is subsumed within people is mortal. In order to refine this structure, advanced taxonomies can be defined. E.g., if one would add the premise that a bachelor is a men, the automati-

cally bachelor is mortal. This transitivity in the subsumption process calls for complex structures, representing real-world knowledge and allowing advanced reasoning process.

Obviously, different nouns refer to different real-world objects. It is the very nature of the word itself and of course the usage of the word in a particular context, which determines how many objects of the real world are affected. Natural language offers us mechanisms to address many objects at once, using a common word. E.g. the words “organism“ or “person“ refer to many different real-world objects. The implicit abstractness of those words is the actual key to subsume various objects. Again, real-world knowledge is necessary to determine the implicit relationship if “is-a“ between “person“ and “organism“. However, making this implicit relationship explicit is one of the major challenges to make a next step towards computer-assisted subsumption.

This paper analyzes the subsumption of words and their corresponding objects used in acts (see Section 2). Thereby, machine-readable real-world knowledge as described in Section 3 will be accessed. Section 4 continues by referring to existing work and aligning to related and prior approaches. Section 5 introduces the research objectives, and the used data. A definition about abstractness and concreteness of normative texts will be provided in Section 6. The paper's contribution, namely the automated measurement of abstractness and concreteness in German law texts, is in Section 7. Strength and limitations are critically reflected in Section 0 A usage scenario, namely law comparison, is given in Section 7.3. Finally, the work concludes and shows further research directions in Section 9.

2. Subsumption in Legal Theory and Legal Practice

Several challenges, mainly addressing the limited expressiveness of natural language and the correspondence to real world situations, exist during the subsumption process of legal norms. However, the problem is well studied in legal theory and the subsumption process allows different techniques to provide solutions to the fundamental problems. The subsumption process, which is essential during the application of a norm has, according to Larenz and Canaris (see Larenz, Canaris 1995), four different dimensions, which are shown in the Figure below.

Subsumption			
Grammatical	***Systematical***	***Historical***	***Teleological***
A word’s meaning decides about the application scope of a particular act. Thereby, the structure of the language, i.e. the relationships between words and phrases is considered.	The construction and subsumption principle does not allow contradicting norms. Circumstances are excluded from the application scope if systemic validity (consistency) is threatened.	The subsumption process is guided by the reconstruction in the light of modern (current) situations. Thereby, it is necessary to reconstruct the legislators usage of language.	The application of a particular act, is determined based on originally intended purpose, pursued by the legislator. The underlying motivation and intention is important (“ratio legis”).
E.g. Gun is a Weapon; Fist is not a Weapon	**E.g.** claim of “outstanding debts” is no fundamental right (see § 823 Abs. 1 BGB)	**E.g.** the general freedom of action (see Art. 2 Abs. 1 GG)	**E.g.** “owner” includes “tenants” and “usufructuaries” (see German Federal Mining Act)

Figure 1: Differentiation of legal subsumption according to Larenz and Canaris

The essential differentiation of the subsumption process in legal theory and practice shows the great potential but also the great challenges for automated or semi-automated algorithmic reasoning approaches. The flexibility, and adaptability of language is its main strength but also its main weakness (Larenz, Canaris 1995). This research aims to support the grammatical subsumption process and evaluates the usage of lexical knowledge as introduced in the next section.

3. Representation of Real-World Knowledge

In order to enable computer-assisted subsumption over objects, an adequate representation, i.e. machine-readable form, of real-world knowledge is required. Thereby different approaches exist, whereas structuring real-world data in taxonomic structures is one of the most promising organizing principles (Hutchison et al. 2005). Enhancing those taxonomies with functionality regarding semantic constraints, lead to ontologies.

From early stages on, scientists and philosophers tried to set up a complete and comprehensive taxonomy, in which every observation, that can be addressed using words, has its unique place. Those approaches were introduced in the domains of nature and life sciences, i.e. biology. The taxonomy thereby is the organizing principle whereas each entity is classified regarding to its properties, so that it can be either distinguished or combined with other entities that are already in the taxonomy. The linguistics answer to the biological classification is of course a taxonomy over words of a language. Two prominent representatives are WordNet, „a large lexical database of English“ (see George 1995), and GermaNet, the German pendant to WordNet (see Hamp, Feldweg 1997). Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept.

GermaNet is based on the same structure as WordNet, in which words are the key elements. Those are linked to each other regarding various relationships. Many words are synonyms to other words, such as „shut“ and „close“ or „car“ and „automobile“, therefore they are stored in a common set, namely the synset. Hence, the synonym relationship is expressed using a common set as storage. The most common relationship between those synsets is the super-subordinate relation, which defines hyper- and hyponym or ISA relation between words. Thereby general synsets such as „energy“ are linked to more specific synsets „electricity“ or „heat“. Beside of this relation type other relationships between words are stored in GermaNet, which are not relevant to us in this particular research. Integrating a huge amount of words into this structure leads to a comprehensive tree-like structure, whereas general and specific words can uniquely be identified. Due to the fact, that we solely investigated German and Austrian acts, we limited ourselves to the usage GermaNet (see Section 5.1.).

4. Related and Prior Work

The usage of ontologies is very common at the intersection of law and informatics. The guiding rationale is the creation of domain specific knowledge with proper semantic constraints, which allows the modeling of an excerpt of the real-world with regards to specified problems (Bench-Capon et al. 2012; Sartor et al. 2013). The principle behind ontologies addresses different aspects of modeling. The mentioned lexical knowledge databases WordNet and GermaNet are ontologies integrating lexical information and semantic relationships, such as hypernymity and homonymity.

The creation of ontologies to support legal information retrieval across different language barriers was the objective of LOIS (Lexical Ontologies for Information Sharing) by Tiscornia et al. (Tiscornia 2006). The main idea was to describe the legal domain of six different European languages and link the concepts between them. Words as placeholders for legal concepts are linked with each other and the resulting semantic lexicon supports multi-lingual information retrieval. WordNets architecture of ordering lexical information was the template for the architecture used in LOIS.

Textual representation as the interface between normative regulation through legislation and the effect on real-word problems inevitably calls for linguistics, since it is the science dealing with language in particular. From a linguistic point of view, several properties of legal texts could analytically be investigated (see McNamara et al. 2014; Köhler 2005). The main aspects that the analytical

and quantitative linguistic is dealing with, concerns structure, coherence, hierarchy, etc. of text. Using linguistic methods to analyze words in legal texts and acts inevitably leads to an overlap to legal sciences. In legal sciences, the word is the basic information entity to communicate and interpret norms. Consequently, the wording is more crucial in legal sciences than it is in any other discipline. Which did other legal experts and philosophers already express „Law is a profession of words“ (Mellinkoff 2004).

The usage of lexical knowledge as a grounded measure for the abstractness, respectively generality, was also used in the domain of social environments. Thereby, Benz et al. used the measurement of a words position within the taxonomic tree as a comparison to other competing abstract metrics (Benz et al. 2011). During the analysis of generality of error- and noise-prone tags of social information systems, such as folksonomies, the semantic information contained in „well-defined“ semantic repositories, namely WordNet and GermaNet, served as base line and gold standard for the comparison.

4.1. Grammatical Subsumption

The normative character of laws consequently leads to abstract norms and regulations. A general formulation of norms is required not to only describe allowed and prohibited actions on a level of single and isolated actions and tasks, but to provide statements about a set of actions. To determine the application scope of norms several mechanisms exist (Larenz, Canaris 1995; Mellinkoff 2004). Within in this paper we are particularly interested in the processes based on the wording of an act.

The usage of words and nouns, that refer to many objects in the real world. The rationale behind is that words can be more or less concrete, whereas concrete words can be subsumed beneath different, more general nouns. E.g., „electricity“ can be subsumed under „energy“. This arises from straight-forward linguistic definitions of abstractness as exemplary given by Brown (Brown 1958). Wording is what Larenz and Canaris call the grammatical subsumption principle and is the starting point of every subsumption principle. Larenz and Canaris argue, that it is obvious because the legislator uses the words in a common sense, so that the citizens and addressees can read it and determine whether they are effected or not (Larenz, Canaris 1995, p. 141).

Raabe et al. also use the subsumption of legal terms starting from the wording argument. They argue, that it is essential during the reconstruction of a legislative term, to start with the wording of a term and then – if necessary – progress to the more elaborated subsumption processes (Raabe et al. 2012). Raabe et al. also used during their research ontologies to provide domain specific structure and relationships, such as DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) (Gangemi et al. 2002). According to Raabe et al. the ontological categorization of DOLCE can be extremely helpful in order to extend a words meaning, which could not be captured by its original sense.

The wording mechanism has a strong focus on the usage and analysis of text that is used to express the meaning and intention of norms. Within this work, we will focus on the determination of the abstractness and generality of words within law texts, which is due to the objectivity of text suitable for algorithmic and computer-supported analysis.

5. Research Objectives

The paper summarizes a quantitative and descriptive empirical research. Thereby, we used publicly available German law texts and applied information extracted from a lexical knowledge database, i.e. GermaNet, to them. The resulting values are metric indicators for the textual abstractness and concreteness. The indicators allow for a comprehensive and effective analysis and comparison of acts and their application scope. The quantitative research based on textual information with con-

sideration of lexical knowledge is addressing a few research questions. That aim to determine whether the measurement of abstractness and concreteness of concept used in law is in principle possible or not. The concrete questions are as follows:

1. What does abstractness and concreteness mean in the domain of legal language, i.e. text?
2. How to measure abstractness and concreteness formally and objectively and what is a possible quantification?
3. What is the distribution of abstract and concrete words used in German acts?
4. What are the limitations of the usage of lexical knowledge for quantifying textual properties?

5.1. Data

To perform the proposed analysis our research requires two different dataset, a German law text corpus and a lexical database containing real-world knowledge. The German law corpus was retrieved on the 10th of October 2014 from the platform *www.gesetze-im-internet.de* hosted and maintained by the Federal Ministry of Justice, represented by *Kompetenzzentrum Rechtsinformationssystem* (CC-RIS), which represents „almost the complete and current federal law“ (BMJ 2014). The second dataset is the lexical database GermaNet. GermaNet is the German pendant to the English WordNet. The number of different words distinguished by their meaning is called lexical unit, of which 121 810 are contained within GermaNet.

6. Concreteness and Abstractness in German Laws

The question what makes a law abstractness or concreteness, cannot easily be answered. As we have already stated out in the introduction, this paper addresses legal norms on a textual level, namely the level of grammatical, linguistic representation, i.e. words. There are three major reasons for these decisions and we will shortly summarize them:

1. Text is an objective artefact. As an artefact, it does not contain subjective biases. This certainly changes during the interpretation by a reader (for a discussion about text-reader interaction models see Schendera 2004), nevertheless the text itself remains unchanged.
2. The subsumption process is heavily determined by the words that are used to express a norms application scope. Although there are concept during the interpretation process like teleological reduction or teleological expansion, which influence - from case to case - the scope of a norm, the act and the concepts described with words, i.e. nouns, remain unchanged.
3. Due to its accessibility, the text is suitable for automatic processing. Beside of the fact, that natural language processing is challenging (see Section 0), algorithms allow to process a huge amount of data, which can create useful insights and reliable results.

In the following, we explain in more detail what abstractness and concreteness of nouns can be and we will give constructive definitions how to measure them. Furthermore, it will become clear how the used nouns and their usage within ordinary language and how this can be accessed by algorithmic and automated approaches effect the scope of norms.

6.1. Abstractness

Using a lexical knowledge database, the calculation of the abstractness of a word can be measured in various ways. For our approach, we decided to use a straightforward approach, namely counting

the number of child nodes. Figure 2 (right side) visualizes the rationale behind our definition of abstractness.

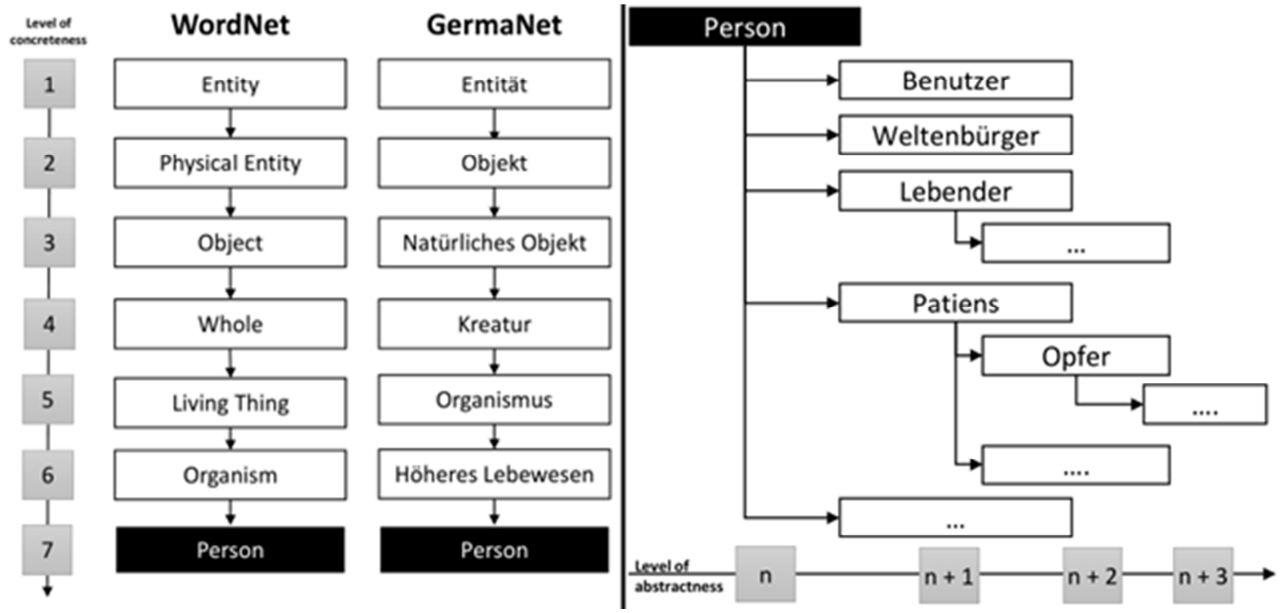


Figure 2: Concreteness (left side) and abstractness (right side) of the noun "Person" in GermaNet

Figure 2 starts with the root node „Person“ and determines all the nouns, that have a ISA (hyponym) relation to the noun. In this particular case, this are several words like “Benutzer“, “Weltenbürger“, “Lebender“, etc. However, the abstractness of words -- as we use it in our research -- does not only consider the child nodes but also the child nodes of the child nodes and so forth. Consequently, the whole subtree underneath one node is determined. This of course could cover several levels until the algorithm terminates in the leafs of the tree.

The abstractness of a term is not determined on the number of levels the corresponding node has to its leafs (depth of inheritance), but is the number of child nodes that are below the node. The idea is that the abstractness is determined in a mathematical sense by counting number of possible meanings, respectively words. Therefore, the more child nodes a node has, the more meanings are subsumed underneath the node and therefore, the more abstract the node actual is.

The example given in Figure 2 shows the noun “Person“ with some of its child nodes. The lexical database GermaNet knows 49 direct hyponyms for the noun „Person“. This means, that on the n+1 level of abstractness 49 different specialized words for „Person“ exist, such as “Benutzer“, etc. The tree structure of the lexical database can lead to an exponential growth of the number of words while moving down-wards the hyponym relations. Consequently, the following definition emerges:

Definition Abstractness: *The total number of nouns* α_n *which can be reached by a hyponym relation starting from the noun n will be defined as n's abstractness. Reoccurring hyponyms are only counted once.*

Using the definition above, we are now able to determine the abstractness of “Person“ explicitly, objectively and quantifiable. Querying GermaNet returns a total number of $\alpha_{Person} = 12\,628$ distinct nouns, which can be reached by the noun “Person“ solely by using hyponym relation.

6.2. Concreteness

In contrast to the abstractness of a noun, the concreteness determines the specificity of a noun, which is semantically connected to the abstractness. It is possible to derive an objective measurement for the noun's concreteness by using a lexical knowledge database, it. Thereby, again the hy-

pernym relation provided by the lexical databases is used. Figure 2 shows an excerpt of WordNet and GermaNet hypernym relations starting from the noun “Person“. The Figure furthermore provides a comparison between WordNet and GermaNet regarding the level of concreteness. Following the noun “Person“ upwards in the hypernym tree, WordNet returns us the nouns “Organism“, “Living Thing“, “Whole“, “Object“, “Physical Entity“ and the least concrete concept “Entity“.

The inheritance tree on the left side was derived from WordNet and GermaNet for the right inheritance tree. Both lexical databases return the same depth of inheritance for the English and German noun „Person“. The depth of inheritance is seven in both cases. This means that seven steps along the hypernym relation are required to reach the root node “entity” (“Entität”).

Using GermaNet or WordNet as a base line to determine the concreteness of a noun has the advantage, that every noun, that is stored in the database has a hypernym relationship, as long as it is not the root itself. In GermaNet, the root is called “GNRoot”. From this root node downward, the nouns are placed hierarchically using the hypernym relationship. “GNRoot” has several child nodes, such as “Zustand“ (state), „Attribut“ (property), etc. “GNRoot” is the most generic concept since it does not have any hypernyms, obviously it is a fictional concept inserted for technical reasons. Based on our prior investigations of the concreteness of a noun, we are now defining the concreteness as an objective measure that can be used in all lexical knowledge databases:

Definition Concreteness: *The total number of hypernyms β_n that exist between the noun and the root node of the lexical database will be defined as concreteness.*

Problematically, the words stored in WordNet and GermaNet, are organized in synsets, in order to enable polysemy. Language allows several meanings for the same word depending on its usage and its context (polysemy). For example, the word „bank“ can have several meanings. Those meanings are represented in so-called synsets, which contain an entry for all the different meanings a word can have. Each of the meanings can have a different concreteness. It might be the case, that the financial institution „bank“ has more hypernyms until the root node is reached than other meanings of the same word. As a possible workaround, we determine the average of all different concreteness measures starting from a particular noun:

Definition Average Concreteness: *The average concreteness of a noun is the average length $\bar{\beta}_n$ of all possible hypernym paths starting from a given noun n.*

To illustrate the difference, we will have another look at the example in Figure 2. Using GermaNet, $\beta_{Person} = 7$ but using the same dataset $\bar{\beta}_n = 7.5$. Based on our definition, this means that the average length of all paths from the noun „Person“ to the root node „Entity“, considering polysemy, consists of 7.5 vertices. Taking into account the polysemy is necessary in order to do not make systemic errors. Furthermore, if an analysis is done on a sufficiently large dataset the error becomes very small.

6.3. Generality as the Synthesis of Concreteness and Abstractness

The determination for the abstractness and the concreteness of a noun in legal texts, using lexical knowledge, are two diverging approaches. Both measurements consider different aspects, namely generality and specificity as a linguistic phenomena. However, in order to fully understand a terms generality, respectively specificity, both measurements have to be considered simultaneously. An integrative indicator, combining both values is the synthesis of two opposing and intrinsic properties of a linguistic term, which can be defined as follows:

Definition Generality: *The generality γ_n of a noun is the noun's abstractness α_n divided by it's average concreteness $\bar{\beta}_n$.*

$$\gamma_n = \frac{\alpha_n}{\bar{\beta}_n}$$

Consequently, the generality for „Person“, as used in the prior Sections, is as follows:

$$\gamma_{Person} = \frac{\alpha_{Person}}{\bar{\beta}_{Person}} = \frac{12\,628}{7.5} = 1\,683.73$$

The generality for the noun “Person” is high, since several thousand hyponyms ($\alpha_{Person} = 12\,628$) are stored in GermaNet, whereas the average concreteness is $\bar{\beta}_{Person} = 7.5$. Analyzing a second example, the German word for “law”, i.e. “Gesetz”, the generality is quite different.

$$\gamma_{Gesetz} = \frac{\alpha_{Gesetz}}{\bar{\beta}_{Gesetz}} = \frac{131}{6.67} = 19.65$$

The two examples show that the combination of abstractness and average concreteness as defined above, combined in a division gives an overview of the nouns generality. Based on the measurement on the dataset of GermaNet, we continue to automatically analyze the application scope of legal norms and legal texts in general (see Section 7.1.). Thereby, the measurement serves as a heuristic to compare different but related legal texts, such as the German and Austrian definition of products given in the act governing the liability for a defective product (see Section 7.3.).

7. Algorithmic Determination of Abstractness and Concreteness

7.1. German Laws

From the introduced dataset of German laws, we selected the ten acts, containing the most words. Based on this selection, we automatically analyzed 1 018 448 words. From this 1 018 448 words, 221 985 are nouns (21.78%). Considering only distinct nouns, we aggregated those without stemming and finally retrieved the number of distinctive nouns for each law (see Table below).

Law	**Distinct Nouns**	**Recognized norm**	**Recognized stem.**	**Recognized brute-f.**
AMG	2079	0,48	0,66	0,75
BGB	3399	0,46	0,63	0,71
HGB	2429	0,47	0,65	0,74
KAGB	2097	0,46	0,64	0,72
KredWG	2628	0,42	0,60	0,69
SGB 5	4004	0,39	0,56	0,65
SGB 6	2173	0,45	0,61	0,69
StGB	1898	0,56	0,75	0,83
StPO	2101	0,51	0,69	0,76
ZPO	2464	0,47	0,64	0,72
MEAN	2527	0,47	0,64	0,73
SD	636,69	0,04	0,04	0,05

Table 1: GermaNet noun recognition rate

Table 1 shows the ten German laws with the most words ordered regarding their total word count. The first column gives the name of the corresponding act. Due to lack of additional space, the Table only shows abbreviations. The second column refers to the number of distinct nouns contained in the law. This number represents the overall number of nouns, determined by the Stanford POS tagger (Toutanova et al.). No aggregation regarding noun stem or any other pre-processing is performed. The next three columns show the recognition rates of the different approaches used to determine whether a noun can be found in GermaNet or not. Thereby, we have implemented three different algorithms. The first approach (3rd column) just considers the noun as it appears in the text and compares it to the nouns provided by GermaNet. No stemming or other pre-processing is done. The second approach (4th column) stems the noun - if it is not found as is - using a porter-stemming algorithm and afterwards the retrieved word stem is again compared to GermaNet. The third „brute-force" approach (5th column) firstly searches for the word as provided from the text. If it cannot be found, it stems the word and searches again in GermaNet. If the search is again without success, the algorithm takes the word-stem and iterates over all nouns in GermaNet (93 631). In case there is noun in GermaNet, that starts with the word-stem, the algorithm terminates and returns the determined noun. This algorithm is vulnerable to errors, because it would reduce the word „Mitteilungen" to its stem „Mitteil" which afterwards matches to „Mitteilungsblatt", which is of course wrong. Consequently, it rather be used as decision support and proof-of-concept, than as exact result.

The recognition rates shown in column 3-5 of Table 1 give an overview of the nouns used throughout German law texts and their correspondence in GermaNet. The recognition rate 0.46 of the German Civil Code (BGB) represents the fact, that 46% of the nouns as used in the law text are contained in GermaNet. Using a stemming algorithm, a recognition rate of 63% is achieved. Furthermore, the vulnerability to errors is also kept low since it mainly removes lexical post-fixes like, -s, -er, -es, -en, etc. The usage of the brute-force algorithm also leads to a higher recognition rate (71%), but the failure rate also increases (see Section 7.2.). The before mentioned problem remains: making the „starts-with" criteria sufficient for the assignment of nouns is problematically. The sources of possible errors and the error rate are in detailed discussed in the next Section 7.2.

	Nouns	α	$\overline{\beta}$	γ
Normal	1547	260.50	7.37	63.95
Stemming	2141	264.10	7.32	66.47
Brute-Force	2417	245.55	7.35	61.44

Table 2: Abstractness, average concreteness and generality of the German Civil Code (BGB)

Using the noun as is allows to find 1547 nouns (46%) in GermaNet. Based on every single noun, the average concreteness, abstractness and generality as defined above is calculated. The same procedure was done using the stemming and brute-force method to increase the recognition rate of nouns in GermaNet. Table 2 shows the overview of the average measurements performed. The average concreteness, the distance to the GermaNet root node, does not differ significantly, but is at about 7.3. Comparing the different abstractness rates, the number of hyponyms of a noun, the difference is again not very large. The normal method leads to an average noun abstractness of 260.50, the stemming method give 264.10 and the brute-force method results in 245.55 average noun abstractness. Calculating the standard deviation on the concreteness and abstractness offers a greater insight into the distribution of the determined measurements. The standard deviation of the concreteness ranges from 2.14 (normal) to 2.19 (stemming) which is relatively low. Considering the standard deviation from the abstractness measurements the situation is different. The standard deviation ranges from 1530.24 (brute-force) to 1594.56 (stemming). The detailed investigation of the

large standard deviation of the abstractness, 6-7 times as large as the average, was done with histograms (see Figure 3), whereby the focus was set on the distribution of the abstractness and concreteness as defined in Section 6. The distribution of the nouns' generality γ_n does not much differ from a qualitative perspective from the abstractness. Therefore, we omitted the visualization in this paper.

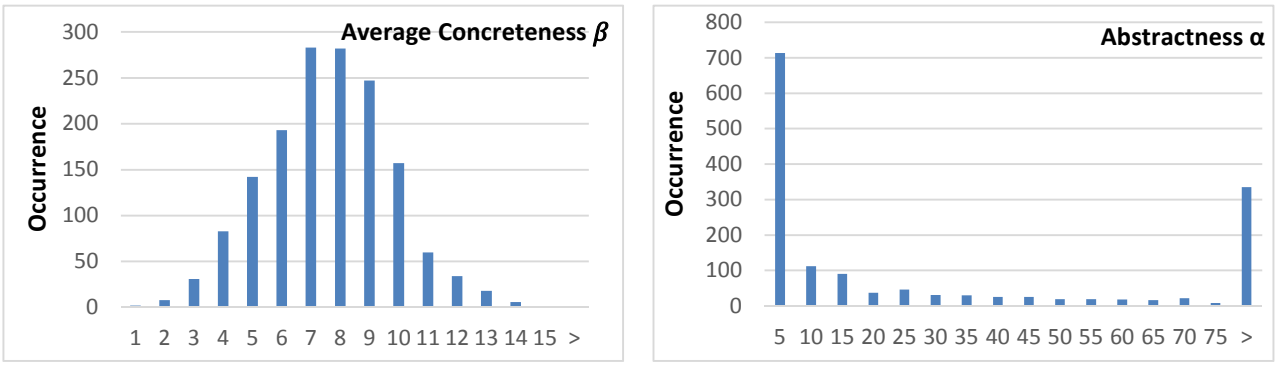


Figure 3: Histogram of nouns' concreteness and abstractness

The distribution of the concreteness of the different nouns in the German Civil Code is shown as histogram in Figure 3. Thereby, the concreteness of those nouns detected using the normal approach, are visualized. Consequently, 1549 different nouns with $\bar{\beta}_n = 7.37$ and standard deviation of 2.14 (see Table 1) were considered. The Figure shows, that the distribution is similar to a Gaussian distribution, although the imaginary bell shaped curve would not fit perfectly on the dataset. The decrease of nouns frequency towards higher concreteness is faster, than to lower concreteness. Mathematically speaking, the distribution has a non-zero and positive skewness (right-tailed). Splitting the measurements into three equal distance clusters, shows, that in the cluster $1 < \bar{\beta} \leq 5$, 17% of all nouns, in the cluster $5 < \bar{\beta} \leq 10$, 75% of all nouns and in the upper third $10 < \bar{\beta}$ 8% of all nouns are contained. Interestingly, the first cluster containing less concrete nouns has two times more nouns than the third cluster.

The situation is different if we analyze the distribution of the nouns' abstractness. We already observed that the average abstractness is 260.50 with a standard deviation of 1548.80. The histogram of the nouns' abstractness measurements is also shown in Figure 3. At a first glance, one can already see the distribution is completely different. The analysis of the abstractness measurements offers, that a noun used in the German Civil Code are either very abstract ($\alpha_n > 75$) or are the quite opposite ($\alpha_n \leq 5$). This also explains the high standard deviation that was measured. Obviously, the calculated mean is not an adequate representative for the determined nouns' abstractness since it does not represent the information about distribution.

7.2. Evaluation

Since NLP is error prone and we also faced some drawbacks using algorithms to automatically determine words and their POS. We identified three different error sources for possible errors during the overall processing the German law texts:

1. NLP techniques: The POS tagger does not always determine right and comprehend results. Some words are determined to be nouns, although they are something different, like adjectives or verbs. The misspelling and orthographic errors of the text can be neglected, since law texts are mostly free from those errors due to their high textual quality.
2. Pre-processing techniques: Due to the lack of processing the determined noun, it cannot be found in GermaNet. In some cases, it is not possible to just stem the word, because the

stemming does not always deliver correct and useful results. The introduced brute-force method to boost the recognition rate increases the error rate so that it is unusable.

3. Incompleteness of GermaNet: Some nouns used in the law are not contained in the GermaNet since there first usage ever is the law text. The German legislator acts as the creator of new and artificial words like „Leibrentenversprechen“, „Zahlungsauthentifizierungsinstrument“, or „Verfahrensbeteiligter“. Hence, the lexical database lacks of a law specific vocabulary.

To evaluate the usage of NLP technologies, we performed a manual evaluation of the retrieved nouns using the Stanford log-linear POS Tagger. Table 3 summarizes the result of the manual evaluation part of the German Civil Code (BGB). The Stanford POS tagger determined 3399 distinct nouns, from which we checked 1000 randomly selected nouns. The result is, that out of the selection of 1000 different nouns 982 words are nouns, whereas only 18 are no nouns and tagged wrongly. This leads to a precision rate of 98.20%, which is quite high.

We also analysed the recall of the Stanford POS tagger manually. Therefore, we randomly selected 200 nouns from the law text and looked them up in the list of nouns that were determined. The resulting recall was quite surprising: 78.50%. Many words could not be determined as nouns by the POS tagger. From the arbitrary selection of 200 nouns, the algorithm also recognized only 157. One possible explanation of this phenomenon is, that the vocabulary contains nouns that are not used in common language like complex composite words, e.g. „Leibrentenversprechen“, etc. A further explanation would be that the German Civil Code, created 1896 and promulgated 1900, uses a vocabulary, which is nowadays outdated in some cases. Therefore, current training models for POS tagger might not be able to recognize all those words.

	Amount	**Percentage**
Precision	982 out of 1000	98.20%
Recall	157 out of 200	78.5%

Table 3: Precision and recall for noun recognition in the German Civil Code

We also analyzed the nouns in the German Civil Code, that could not be recognized by GermaNet in a first step and on which either the stemming or the brute-force method leads to a successful identification. The following Table 4 summarizes the precision rates, which were checked manually.

	Amount		**False Positives**	
Normal	1 548	45.54%	0	0.00%
Stemming	592	17.42%	19	3.21%
Brute-Force	276	8.12%	153	55.43%
Not found	983	28.92%	-	-

Table 4: Recognition error rates for the Civil Code

The table above shows the error rates we manually detected after the matching of nouns and GermaNet. Of the 3399 nouns determined using the normal approach of processing, 1548 (45.54%) could be found in GermaNet. Additional pre-processing like stemming and the mentioned brute-force searching increased the number of matched nouns by 17.42%, respectively 8.12%. Nevertheless, 983 (28.92%) of the determined nouns could not be found in GermaNet. On the other hand, pre-processing also increases the vulnerability to errors. Out of the 592 stemmed nouns, 19 were

wrongly determined (3.21%). The error rate using the brute-force method was very high. Out of the 276 nouns, 153 (55.43%) were wrongly classified. Obviously, the usage of pre-processing can really boost the recognition rate, but has to be used with care, because it dramatically increases the error rates.

7.3. Act Governing the Liability for a Defective Product: Germany vs. Austria

Due to the political situation, the European Union has an impact on the national legislation, resulting in Council Directives that have to be adopted and promulgated by its Member States. A common example is the Council Directive 85/374/EEC which governs the liability for defective products. In the years after its entrance into force, the German and the Austrian legislation also promulgated their versions of the corresponding act. Below, both articles are given:

Austria ProdHaftG §4: *Produkt ist jede bewegliche körperliche Sache, auch wenn sie ein Teil einer anderen beweglichen Sache oder mit einer unbeweglichen Sache verbunden worden ist, einschließlich Energie.*

Germany ProdHaftG §2: *Produkt im Sinne dieses Gesetzes ist jede bewegliche Sache, auch wenn sie einen Teil einer anderen beweglichen Sache oder einer unbeweglichen Sache bildet, sowie Elektrizität.*

Based on this selection of two different but related norms, we did an analysis regarding the concreteness, abstractness and generality of the nouns.

Germany	**Austria**	α	$\bar{\beta}$	γ
Produkt		4224	5.00	844.80
Sache		23573	3.50	6735.14
Teil		11506	4.20	2739.52
	Energie	118	6.50	18.15
Elektrizität		58	8.00	7.25

Table 5: Comparison of the acts' nouns

Table 5 lists the nouns of the Austrian and Germany act governing the liability for a defective product. Both acts share three different nouns, which appear in both acts, namely “Produkt“ (product), “Sache“ (thing), and „Teil“ (part). Additionally, the table provides information about the average concreteness $\bar{\beta}$, abstractness α and generality γ. Interestingly, the two acts differ regarding their application scope. Whereas the Austrian act also includes energy (Energie) as a product, the German act only electricity (Elektrizität). This difference in the application scope can automatically be determined using lexical knowledge. Table 5 holds both nouns and their respective information from GermaNet. As one can clearly see, „Energie“ has an abstractness $\alpha_{Energie} = 118$, which means that GermaNet knows 118 different nouns that are specific forms of energy. The electricity as used in the German act, has an abstractness $\alpha_{Elektrizität} = 58$. Furthermore, using lexical knowledge it is also possible to determine the hypernym relationship between both nouns.

The ISA relationship between energy and electricity is mapped in GermaNet. Consequently, the usage of this lexical knowledge allows the subsumption in a restricted way, namely the subsumption according to the words sense (see Section 2). The ISA relationship between “Elektrizität“ and “Energie“ cannot be determined by solely looking at the corresponding α, β, $\bar{\beta}$, or γ values. This information is a relation between two separate words, i.e. nouns. This information is stored in GermaNet. Nouns, that are subsumed as energy but not as electricity are concepts like “Primärenergie“

(engl. primary energy), “Wärmeenergie“ (engl. thermal energy), or “Arbeit“ (engl. work). The ISA relationship is transitive; consequently, every noun that is a hyponym to electricity is also a (inherited) hyponym of energy. This inheritance and transitivity is a basic principle and enables the subsumption in the sense of word meaning.

8. Challenges and Limitations

As we have shown above, lexical knowledge can be used to support the comparison of laws regarding the application scope of legal norms and acts. The lexical knowledge databases thereby serve as information provider making the implicit semantic relationships between words of a language explicit and accessible to algorithms. This Section summarizes the challenges arising during the processing of texts, accessing a lexical knowledge and objectively measuring concreteness, abstractness and generality.

Lexical knowledge: GermaNet. Although the usage of lexical knowledge allows an extensive analysis of semantic relationships between words, some drawbacks remain. Firstly, GermaNet lacks of comprehensiveness especially with regard to the vocabulary used in the legal domain.. Secondly, the determination of concepts, represented in language as bigrams, such as „natural phenomena“ or „living thing“, harden the problem of determination the semantics. At last, the problem of polysemy detection, as already observed by Gangemi et al. (Gangemi et al. 2002) exists. If a noun has several meanings, such as the noun „bank“ it is unclear, which word sense is the right one.

Natural language processing (NLP). The algorithmic processing of natural language is known to be challenging but promising. Especially in the research domain of legal texts and legal informatics, the usage of NLP technologies is common. Legal texts are usually well written and without orthographical or grammatical mistakes. This positively contributes to the precision rates. Nevertheless, the complex sentence structures and word compositions introduced by the German legislation are major drawbacks.

Domain specificity of legal practice and theory. The applicability of automatically processed legal texts and acts depend on the intended usage. Thereby, the field of application determines the requirements and use cases that decide about the usefulness of information. This variation also effects the words and their meanings. This complex, and to a certain extend social phenomena, challenges the automated determination and usage of a words meaning even more.

We briefly sketched the limitations of the usage of automatically derived information from legal texts as well as their combination with existing lexical real-world knowledge. The paper now proceeds with a conclusion, summarizing the papers' contribution.

9. Conclusion and Outlook

Our paper is a contribution to the investigation of the application scope of legal texts, based on the subsumption principle, which is a complex and well-studied field in legal theory. We restricted the subsumption process on its word-sense, i.e. grammatical, driven process. We analyzed German law texts regarding their nouns and proposed a theory regarding concreteness, abstractness and generality of nouns, using a lexical knowledge database, namely GermaNet, the German pendant to WordNet.

We exhaustively analyzed the nouns of the German Civil Code. Based on our analysis we can draw conclusion regarding the content and the used method. Whereas the concreteness of nouns used in the law text follows a bell-shaped distribution, the abstractness behaves opposite. Nouns are either very abstract $\alpha > 75$ or it not very abstract $\alpha < 5$, but it's unlikely to be in between. The processing of natural language has some drawbacks, contributing to the recognition, precision and recall, which

we measured and discussed in detail. Our approach also allows for the comparison of application scopes of different acts and norms. We exemplary showed this on the German and the Austrian version of the act governing the liability for a defective product, with the algorithmically reproducible result, that the Austrian act defines a more abstract ($\alpha_{Energie} > \alpha_{Elektrizität}$) application scope for products than the German version.

These measurements give insights into the structure and usage of words, especially nouns, in the domain of law texts. During the subsumption, this could serve as a base line heuristic for decision support, but also for automated comparison the application scope of acts and norms.

10.**Acknowledgement**

This research was sponsored in part by the German Federal Ministry of Education and Research (BMBF) (project “Software Campus (TU München)”, grant no. 01IS12057).

11.Publication bibliography

Bench-Capon, Trevor; Araszkiewicz, Michał; Ashley, Kevin; Atkinson, Katie; Bex, Floris; Borges, Filipe et al. (2012): A history of AI and Law in 50 papers: 25 years of the international conference on AI and Law. In *Artificial Intelligence and Law* (20), pp. 215-319.

Benz, Dominik; Körner, Christian; Hotho, Andreas; Stumme, Gerd; Strohmaier, Markus (2011): One tag to bind them all: Measuring term abstractness in social metadata. In : The Semantic Web, pp. 360–374.

BMJ (2014): Juris. Gesetze im Internet. Available online at http://www.gesetze-im-internet.de/, updated on 7/22/2014, checked on 7/22/2014.

Brown, R. (1958): Words and things: The Free Press.

Gangemi, Aldo; Guarino, Nicola; Masolo, Claudio; Oltramari, Alessandro; Schneider, Luc (2002): Sweetening Ontologies with DOLCE. In *Knowledge Engineering and Knowledge Management*, pp. 166-181. DOI: 10.1007/3-540-45810-7_18.

George, A. Miller (1995): WordNet: a lexical database for English. In *Commun. ACM* 38 (11). DOI: 10.1145/219717.219748.

Hamp, Birgit; Feldweg, Helmut (1997): GermaNet - a Lexical-Semantic Net for German. In *Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*.

Hutchison, David; Kanade, Takeo; Kittler, Josef; Kleinberg, Jon M.; Mattern, Friedemann; Mitchell, John C. et al. (Eds.) (2005): Law and the Semantic Web. Berlin, Heidelberg: Springer Berlin Heidelberg.

Köhler, Reinhard (2005): Quantitative Linguistik. Berlin [u.a.]: De Gruyter (Handbücher zur Sprach- und Kommunikationswissenschaft, 27).

Larenz, Karl; Canaris, Claus-Wilhelm (1995): Methodenlehre der Rechtswissenschafft. Berlin [u.a.]: Springer.

McNamara, Danielle S.; Graesser, Arthur C.; McCarthy, Philip M.; Cai, Zhiqiang (2014): Automated evaluation of text and discourse with Coh-Metrix: Cambridge University Press.

Mellinkoff, D. (2004): The language of the law: Resource Publications.

Raabe, Oliver; Wacker, Richard; Oberle, Daniel; Baumann, Christian; Funk, Christian (2012): Recht ex machina. Berlin, Heidelberg: Springer Berlin Heidelberg.

Sartor, Giovanni; Casanovas, Pompeu; Biasiotti, Mariangela; Fernandez-B., Meritxellx (2013): Approaches to Legal Ontologies: Theories, Domains, Methodologies: Springer Publishing Company, Incorporated.

Schendera, Christian F. G. (2004): Die Verständlichkeit von Rechtstexten. In Kent D. Lerch (Ed.): Die Sprache des Rechts: Recht verstehen. Berlin, New York: De Gruyter, pp. 321–373.

Tiscornia, Daniela (Ed.) (2006): The LOIS project: Lexical ontologies for legal information sharing.

Toutanova, Kristina; Klein, Dan; Manning, Christopher D.; Singer, Yoram: Feature-rich part-of-speech tagging with a cyclic dependency network. In : Proceedings of the HLT-NAACL Conference 2003, pp. 173–180.

TOWARDS A LEGAL CORE ONTOLOGY BASED ON ALEXY'S THEORY OF FUNDAMENTAL RIGHTS

Cristine Griffo[1], João Paulo A. Almeida[2], Giancarlo Guizzardi[2]

[1] Postgraduate student, Federal University of Espírito Santo, Ontology and Conceptual Modeling Research Group (NEMO)
Av. Fernando Ferrari, 514, CT-VII, 29075-910, Vitória, Brazil
E-mail: cristine.griffo@aluno.ufes.br
[2] Professor, Federal University of Espírito Santo, Ontology and Conceptual Modeling Research Group (NEMO)
Av. Fernando Ferrari, 514, CT-VII, 29075-910, Vitória, Brazil
E-mail: jpalmeida@inf.ufes.br, gguizzardi@inf.ufes.br

Keywords: ***Legal core ontology, legal theory, foundational ontology***
Abstract: *Ontologies have been used in recent decades as a conceptual modeling tool in different areas of knowledge. In Law, legal core ontologies (LCO) are proposed as a means of computational representation of essential concepts in order to construct legal domain ontologies and applications for the legal world. A relevant source of legal concepts is the legal theory. However, there are divergences between legal theories about what is law. This divergence should be taken account by ontologists because of their consequences to the usefulness of the concepts. In the last decades, legal theories have proposed solutions for modern social claims. These legal theories have the potential of producing a LCO that is more suitable for the current society. An example of these theories is Alexy's Theory of Fundamental Rights. In this paper, we explore an initial ontological model for rights based on Alexy's Theory of Fundamental Rights in order to build a consistent LCO grounded in Unified Foundational Ontology (UFO). We aim to build up this LCO such that it can become a basis for building domain ontologies, languages, knowledge bases, and applications of the legal world.*

1. Introduction

Research in Computer and Law has its roots in the 1960s. In 1957, Mehl *apud* Bing [1] published a paper about automated legal decisions and initiated this new research trend. Since then, the transdisciplinary area of Computer and Law has remained in the spotlight, with different research niches investigating the various aspects of the field. One of the niches that has received special attention in recent decades is the one of *Legal Ontologies*.

The importance of understanding the universe of norms has to do with the broad spectrum of roles that norms play in society. As well observed Bobbio [1], individuals, from birth to death, live in a world of norms, which direct their actions. In recent decades, ontologies have been used as a proposal of conceptual models in the Computer and Law area to represent this domain. In fact, some authors have argued, "ontologies could be the 'missing link' between Legal Theory and Artificial Intelligence" [2].

In this paper, our goal is to present the outline of a *legal core ontology* (LCO), which represents essential concepts of the Law based on Alexy's Theory of Fundamental Rights. We call this legal core ontology UFO-L (Ontology of Legal Concepts) as it is proposed as a layer built on top of the *Unified*

Foundational Ontology (UFO). We take into account two aspects: 1) Legal theories and 2) foundational ontologies. Also, we defend the convergence between Legal Theory and Artificial Intelligence in the same line of Valente et al [2], i.e., the understanding that "to create or use ontologies without regard to Legal Theory is a certain path to reinvent the wheel".

Legal theories. The representation of a complex domain such as the Law, with polysemic concepts, and several theories for defining *what is Law*, motivates the ontologist to investigate not only computational issues but also existing legal theories. In a simple definition, a legal theory is a body of systematically arranged fundamental principles in order to describe Law under a perspective.

For the investigation of legal theories, we considered that an ontologist should know the *Ontological Problem of Law*, whereas the question *"what is Law?"* has a significant influence on the development of particular LCO. For instance, concepts in a LCO based on Kelsen's Theory [3] differ from those in a LCO based on Cossio's Egological Theory of Law [4] or on Alexy's Theory of Fundamental Rights [5]. In turn, domain ontologies and applications based on LCOs with different underlying doctrinal perspectives will also reflect these differences. In this context, the problem with using, for instance, theories based on Legal Positivism (*e.g.* Kelsen's Pure Theory of Law) is that they do not include modern concepts of the Law introduced by the explicit countenance of a social reality. We advocate here that modern theories can produce a *legal core ontology* more suitable for our current society. Examples of theses legal theories are Reale's Three-Dimensional Theory of Law [6], and Alexy's Theory of Fundamental Rights. In this research work, we choose Alexy's Theory of Fundamental Rights because two aspects. The first aspect is the classification of *norms* as *rules* and *principles*. The idea to understand *norms* in this way was firstly proposed by Dworkin [7], but Alexy gave new contours, defining principles in satisfaction degrees as opposed to *rules*, which are satisfied or not. With this understanding, Alexy's pointed out the impossibility of representing (and solving) every legal problem only with Classic Logic. The second aspect is the relational perspective and the use of Hohfeld's works (correlatives and opposites legal relations) as a base to extend the understanding of legal relations.

Foundational ontologies. In Computer Science, ontologies are used to represent categories and their ties that are countenanced to exist in a conceptualization of given subject domain. Traditionally, these concrete artifacts explicitly representing and underlying conceptualization have been successfully employed over decades to support reuse and sharing of knowledge. An important kind of ontology is a *foundational ontology*. In a *foundational ontology*, this system of categories and their ties is a domain-independent one (representing the most general aspects of reality) and should be built with the explicit support of theories from Formal Ontology in philosophy. Moreover, in the particular case of the so-called descriptive foundational ontologies, theories from areas such as cognitive science and linguistics should also be seriously taken into account [8]. The *Unified Foundational Ontology* (UFO) is an example of a descriptive foundational ontology that has been constructed for more than a decade employing results from formal ontology, cognitive psychology, linguistics, philosophical logics, but also significant accumulated empirical and theoretical results from the area of conceptual modeling in computer science.

UFO-L uses domain-independent concepts of domain provided by the *Unified Foundational Ontology* (UFO). Extending these concepts, a conceptualization for legal domain is built, which can be used in other particular domain ontologies, legal knowledge bases and so on. The use of UFO ensures an ontological consistency due to the observance of principles and postulates that it has (dealing, for example, with relations such as identity, parthood, dependence, etc.). Our choice for UFO is motivated by: (i) our experience with its successful application in a large number of domains ranging from natural

science domains such as Petroleum and Gas and Electrophysiology of the heart to social domains such as organizations, services and software; (ii) the fact that UFO comprises a rich theory of relations and complex relational properties that is absent in other foundational ontologies [9], [10], [11].

The structure of this paper is as follows. In section 2, a background describing both Computer and Law contexts. In section 3, we present the initial steps of the legal core ontology modeling. In section 4, we trace final considerations.

2. Background

2.1. The Legal Theory Perspective

In a timeline of legal theories, we notice that the Classic Legal Positivism was a successful dividing line between Natural Law Theory and Legal Positivism [12], especially with the Kelsen's Pure Theory of Law. Since then, several legal theorists have discussed the *ontological problem of law* under different perspectives, from *social thesis* of Hart [13], *legal Interpretivism* of Dworkin [7] to Alexy's *Theory of Fundamental Rights,* and *Theory of Argumentation* [5], [14].

Postpositivist theories have dealt in a general way with two problems not solved satisfactorily by Classic Legal Positivism: *determining the law in hard cases* and *judge's discretionary power*. Postpositivist theorists have also criticized the *exacerbated legalism* inherited from Legal Positivism. Researchers in AI & Law have pointed out that *Law* is more than a "set of rules or cases" [15], [2], highlighting the due importance of *legal theories* for building *legal ontologies*. Thus, we emphasize this point and add another one: the importance of choosing a legal theory coherent with the current legal reality. Using only positivist theories to build legal ontologies we run the risk of propagating the problem of *legalism exacerbated* to technological artefacts.

We have investigated the use of legal theories in studies about legal ontologies. Our systematic mapping of *legal core ontologies* indicated that 45% of the selected studies used a Legal Positivist approach; 8% used a non-positivist legal theory and 47% did not use as primary source any legal theory. For instance, Valente [16] uses Hohfeld's theory, Kelsen's theory and Hart's theory to build the FOLaw ontology. Breuker et al. [17] follow the same legal theories to develop the LRI-Core. Shaheed et al. [18] uses Hohfeld's theory and McCarthy's Discourse Theory to build the NM-L core ontology. Lu and Ikeda [19] use Kelsen's Theory to build the International Copyright Ontology. Schweighofer and Liebwald [20] use Hohfeld's theory to propose a legal information retrieval application. Palmirani et al. [21] uses Kelsen's theory to propose an ontology of time, and Scharf [22] uses Kelsen's theory to propose a rule engine for legal reasoning (rOWLer).

In addition to issues about the *ontological problem of Law,* described in section 1, and to the question of "*which legal theory is more suitable for the current legal reality?"*, presented in this section, there is another aspect to consider in building legal core ontologies: *what kind of representation fields in Law will be modeled.* We identified two fields: 1) the Science of Law and 2) the Law as a particular Legal Order. *The Science of Law* studies the existing concepts in the Law through a scientific method and descriptive language. It is concerned with the general notions of the Law and not with particular norms [11]. In contrast, *Law as a particular Legal System (*or *Legal Order*) is a specific system of legal norms (rules and principles) established by a competent authority. Because of its coercive nature, it has a prescriptive language [23]. For instance, the Brazilian Legal System, the Legal System of England and Wales. The Brazilian Legal System is an example of *civil law legal system.* In this system, the foundation is the written *law*. On the other hand, The Legal System of England and Wales is an

example of *common law legal system* and its foundation is the *common law*, which means that law is built by judges case by case and a judge is generally bound on a prior case. However, it is important to emphasize that this distinction is not watertight, as an example, the binding precedents and jurisprudence in the Brazilian Legal System have become very important in contemporary judgments [24].

2.2. Alexy's Theory of Fundamental Rights

Alexy's theory of Constitutional Rights or Alexy's theory of Fundamental Rights [5] (called here Alexy's theory) addresses some problems of Legal Positivism by proposing the *Weighing and Balancing structure* based on The *Lüth case* [25]. We present a brief overview of two aspects of the Alexy's theory: 1) Norms as *rules* and *principles,* and 2) *legal positions* (some other important discussions in this theory, *e.g. balancing system and weighing formula*, are not mentioned due to space constraints).

The first aspect of Alexy's theory is about *(legal) norm.* A *norm* is defined as "the meaning of a normative enunciation". *Norms* are classified as *deontological norms* and *axiological norms*. *Deontological norms* are, in turn, classified as *rules* and *principles*. *Principles* are optimization requirements, which have different degrees of satisfaction (degree of fulfillment) depending on both factual and legal aspects. On the other hand, *rules* are norms, which are or fulfilled or not [5].

The second aspect is concerned with *legal positions*. *Legal positions* are defined by Alexy [5] as situations in which a subject, in a legal relation, has a *right (lato sensu)* against other subject. In that sense, every *legal position* is a *relation* between two subjects and an object. Alexy's system of *basic legal positions* divide *rights* in *rights to something, liberties,* and *competences (legal power).* In turn, *rights to something* are divided in *rights to negative acts (non-obstruction of acts, non-affecting of characteristics and situations, non-removal of legal positions),* and *rights to positive acts (factual act, normative act).*

According to Alexy's theory, the legal positions of the sort r*ights to something* can be represented using the logical connections between legal relations from Hohfeld's Theory [26], [27]. According to Hohfeld [27], the legal relations are grouped in a "convenient scheme of opposites and correlatives" as follows. Jural opposites (*right, no-right, duty, no-duty or permission),* and jural correlatives (*right, duty, no-right, permission),* as shown in figure 1.

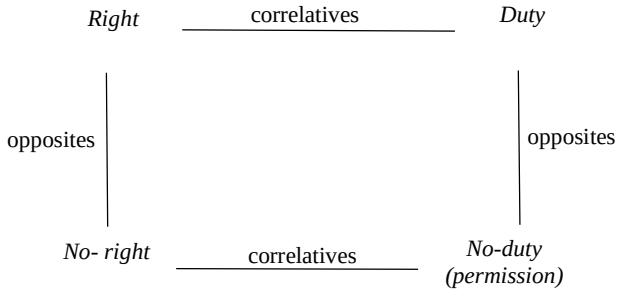


Figure 1. Hohfeld's scheme *apud* Alexy [5]

The use of Alexy's theory is justified because of its structural and relational aspects, providing a "basis and a framework for everything else that follows" [5]. For our work, we have divided Alexy's theory in two parts as follows. The first part is concerned with the system of *basic legal positions* (fundamental

rights as subjective rights). The second part refers *Weighing and Balancing* applied to the interpretation of the legal norms and the structure of norms (deontological and axiological norms). This paper deals with the first part of Alexy's theory.

2.3. Legal Ontologies Perspective

Since the 60's, many studies on AI & Law, conceptual modeling and Law, information retrieval in Law, among others, have contributed to solve the initial problem of representing and retrieving legal knowledge [14]. Regarding legal knowledge bases, Bing [1] accurately predicted in 1992 the strengthening of research related to the legal knowledge bases and legal philosophy or jurisprudence. Since then, a number of important research efforts concerning legal bases knowledge have appeared Casellas [21], Agnoloni and Tiscornia [22], Poblet et al. [23], Breuker et al. [24].

Nowadays, it is clear that is not enough to represent the syntax of legal entities; it is necessary to represent their semantics as well as their mutual relationships. In a globalized world, it is not enough that there is a legal knowledge base; it is necessary that this base interoperates with others existing knowledge bases. Ontologies applied to Law aim at addressing these representation and interoperable problems.

The concept of ontology has its origins in Philosophy (as a field of study and as a system of categories and their ties). However, in the past 2-3 decades, it has been adapted to Computer and Information Science to mean frequently a formal representation of a particular system of categories and their ties [8], [28]. From this convergence, Guarino [28], Gruber [29], and Staab [30] proposed definitions, methodologies and classifications of ontologies.

According to Gangemi *apud* Oberle [31], ontologies are classified either by their specificity or by their purpose. Related to specificity, ontologies are: 1) foundational ontology; 2) core ontology; and 3) domain ontology. Related to purpose, ontologies are: 1) reference ontology; and 2) application ontology.

A *foundational ontology* defines a set of domain-independent ontological categories. In turn, a *core ontology* defines a set of fundamental concepts of a field of knowledge (e.g. services, collaboration, law, organizations, software) that are still general concepts that occur across multiple domains. Core ontologies are often built by reusing and/or extending a foundational ontology [32]. Finally, a *domain ontology* defines a set of concepts from a specific domain (e.g. Brazilian law). *Foundational ontologies*, such as UFO [8] and DOLCE [33] are useful in building LCOs because they can help to bring both ontological consistency and completeness to the process [13]. For instance, the OPJK ontology [34] used concepts as *agent, role, document, process,* and *act* from DOLCE Lite + CLO, SUMO, and PROTON.

Core Ontologies that represent legal domain-independent concepts in Law are denominated *Legal Core Ontologies* (LCO). In this paper, a *LCO* is defined as a cohesive and coherent set of concepts, properties and relations that exist in the legal universe. A LCO can be used as basic structure in legal domain ontologies, frameworks, and application ontologies.

In the literature, the expression "*legal core ontology"* began to be used in middle 90' by Valente et al [35], and Breuker et al [36]. Among the most cited *legal core ontologies* in the literature we have:

- **Frame-Based Ontology (FBO)** published in 1993 by van Kralingen et al [37], based on legal positivism (Hart, Kelsen, van Wright, and Ross theories) and written in ONTOLingua. It is a mix of foundational categories and legal core concepts. The core of this ontology is the concept

of *norm* and concepts related to it, such as *norm subject, legal modality,* and *description of the act.*

- **Functional Ontology of Law (FOLAW)** published in 1994 by Valente [16], written in ONTOLingua, it is based on Kelsen, Hart and Bentham theories, and has a functional perspective and knowledge-oriented (*normative knowledge, responsibility knowledge, reactive knowledge, creative knowledge, and meta-level knowledge).* As this ontology is based on Kelsen's theory, basically, *norms are rules, which are either observed or violated.*
- **Hage and Verheij's Ontology**. Published in 1999, and written in First-Order Logic, it is an ontology based on Dworkin and Alexy's theories of norms classification (*norms are rules and principles).* For them, a legal ontology is an interconnected dynamic system of state of affairs. The principal categories of this ontology are *individuals (state of affairs, events, and rules)* [38], and similar with FBO's ontology, it mixes foundational concepts with legal core concepts.
- **Core Legal Ontology (CLO)** published in 2003 by Gangemi et al [39] and written in OWL-DL, it is the first LCO built grounded in an explicitly defined foundational ontology (DOLCE). There is, however, no explicitly defined primary legal theory source on which this ontology is based.
- **LRI-CORE** built by Leibniz Center for Law Research Group [40], published in 2004, and written in OWL+DL, it is grounded in different foundational ontologies (DOLCE, SUO, John Sowa's ontology). It has later evolved to **LKIF-CORE,** which has been built by the same group (2007).
- **PROTON+OPJK** is a combination of ontologies built inside the SEKT European project, PROTON is a foundational ontology based on commonsense concepts. Casellas's ontology (OPJK) [34] is an ontology which contains relevant legal domain specific knowledge. Although, at first sight OPJK can be considered a legal domain ontology, it also contains several generic concepts that can be reuse in different legal domain ontologies (*e.g. judicial organization, judicial role),* giving to it a nature of core ontology.

Other works related with legal domain representation cited in the literature, are: LEGOL, the seminal work, by Stamper [41], Hafner's semantic work [42], McCarty's language [43], Mommer's ontology [44], among others.

2.4. Unified Foundational Ontology (UFO)

Following a well-documented trend in the ontology engineering literature, we here strongly subscribe to the practice of using foundational ontologies as a central methodological tool for building core and domain ontologies. In particular, we employ the foundational ontology UFO as a basis for our work.

The *Unified Foundational Ontology* (UFO) was initially proposed by Guizzardi and Wagner [13], permits the building of an ontology reusing some generic concepts (*e.g. kind, sub kind, relator, role, role mix*). The ontologist does not need to rebuild these concepts. For instance, Lopes et al [46] grounded the Civil Law domain ontology in UFO, using the ontology modeling language OntoUML (containing ontological notions such as *kind, sub kind, phase, mix, relator, role).*

The *foundational ontology* UFO has three layers. UFO-A (*ontology of endurants*) is the UFO core, and includes terms as *universal, relator, role, intrinsic moment*. UFO-B (*ontology of perdurants*) is a layer built on the UFO-A, and relates terms as *event, state, atomic event, complex event*. UFO-C is built on UFO-B and UFO-A and represents *the social reality,* which relates categories such as *social agent,*

social object, *social role*, *and*, *normative description*. Figure 2 shows the fragment of UFO and some UFO-C categories are described as follows.

A *normative description* defines one or more rules/norms recognized by at least one *agent* and that can define nominal *universals* such as *social moment universals* (*e.g.*, social commitment types), *social objects* (the crown of the king of Spain) and *social roles* (president, prime minister, PhD candidate or pedestrian) [47]. For instance, consider the rules of hopscotch game Even in an informal social context, there is a set of rules (in general sense) being observed by its participants. Breaking these norms will result on penalties (exclusion of the player, the game ends) or a social imbalance (the conflict). Brazilian Constitution, ICAIL 2015 Regulations are examples of *normative description.*

Agents are *substantials* capable of bearing special kinds of moments named *intentional moments.* Examples of agents include Barack Obama and the Brazilian Federal Republic. *Agents* can bring about *actions* (*intentional events*). According to Almeida and Guizzardi [48], they are *substantials* capable of bearing special kinds of moments named *intentional moments*. *Agents* may play *social roles*, such as *husband* and *wife* in the context of a marriage (a *social relator*), as well as, *student* and *professor* in the context of an enrollment (a *social relator*). *Social agents* are those defined by a normative description, e.g., the Brazilian Federal Republic, as opposed to a *Human (or physical) Agent*. A *social object* is a category of UFO-C that defines *non-agentive substantials* produced in a social context. For instance, the crown of the king of Spain object defined in the context a certain geo-political entity. For a detailed description of UFO categories, we refer Almeida and Guizzardi's paper [4], which describes UFO categories, especially UFO-C categories, pointing out the nature of each one of them.

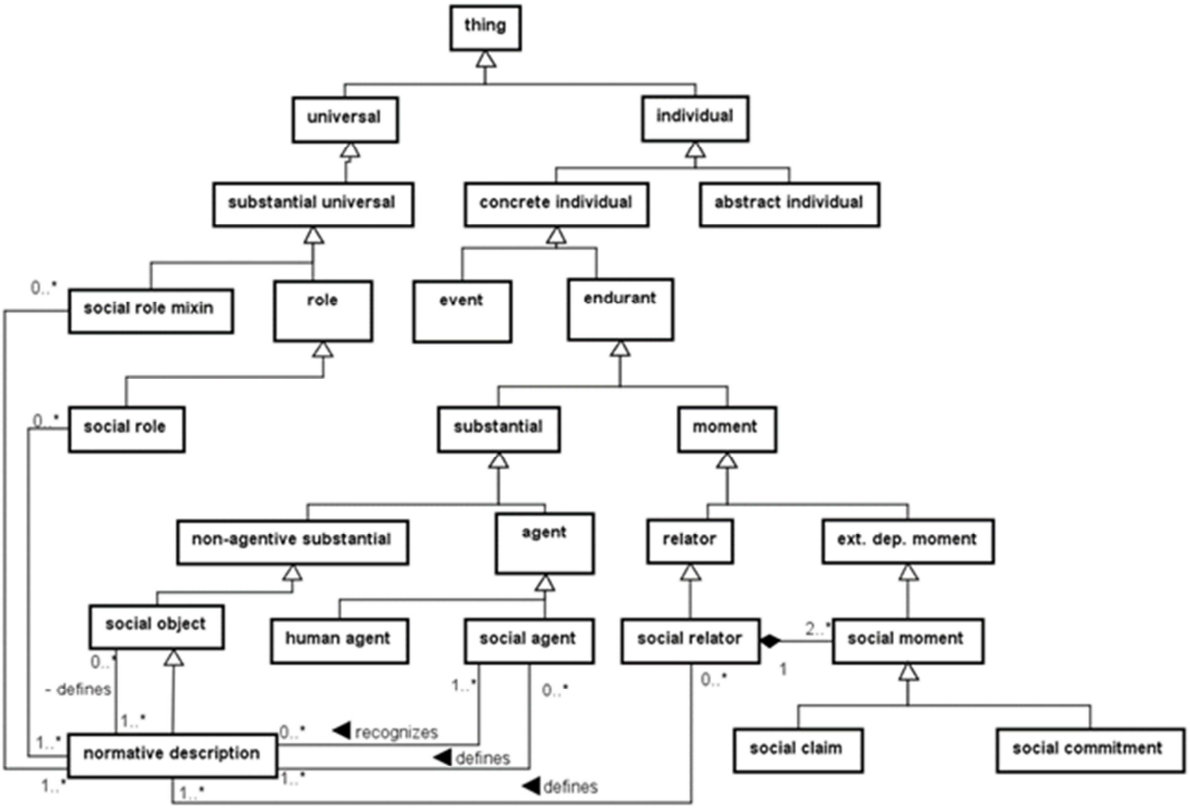


Figure 2. Fragment UFO (Almeida and Guizzardi [4])

3. A Model of Alexy's notion of rights to something

In this initial work, we focus on Alexy's notion of *rights to something*. Other rights (*protected liberty, non-protected liberty, citizen competence,* and *state competence)* will be represented in future work. These concepts are described as follows.

Legal relation is a *bond* between subjects achieved by the existence of a *legal fact*. In other words, it is the *social relation* typified in a *legal norm* according to Larenz [49], and Reale [6]. Since this research work is guide by the relationship perspective, this legal concept is the main concept represented in our model, using the notion of *legal relator*. According to Almeida and Guizzardi [48], a ***social relator*** (figure 2) is a *relator* "composed of two or more pairs of associated *social moments (social commitments; social claims)*". In turn, a ***legal relator*** is a specialization of *social relator*, dependent on a number of other *individuals* or *universals* that play *legal roles* (which are universals that agents instantiate contingently when bound by the legal relator). Figure 3 shows the taxonomy of *legal relators* according to Alexy's theory. The taxonomy shows specializations of the UFO notion of relator.

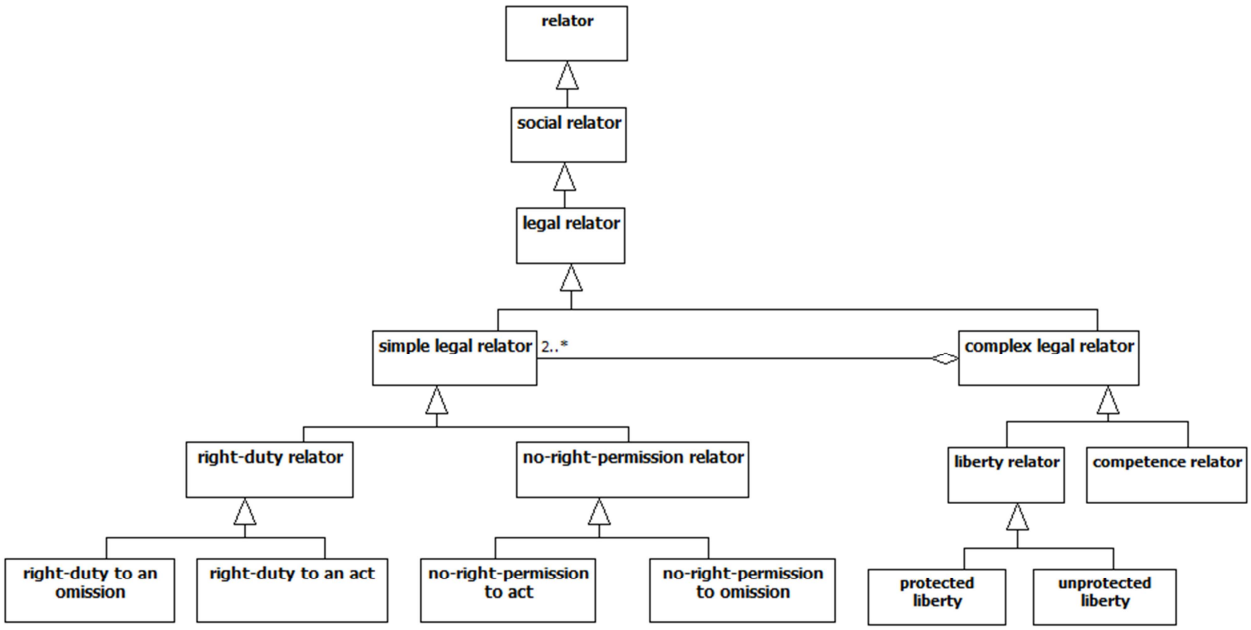


Figure 3. Taxonomy of relators

A *legal relator* is specialized in *simple legal relator* and *complex legal relator.*

Simple legal relator. A *simple legal relator* represents Alexy's concepts of *rights to something.* It uses pairs of legal fundamental concepts (right–duty, no-right–permission). A s*imple relator* may be classified as *right–duty relator* or *no-right–permission relator*.

Right–duty relator. A *right–duty legal relator* uses the *legal relation* right–duty (correlative) to bind *right holder* and *duty holder.* A *right holder* is someone who has a right to something against a *duty holder* (*e.g.* a citizen as right holder has a right to vote against the state as duty holder). A *duty holder* is someone who has the duty to materialize the right of a *right holder.* Table 1 discusses further examples

of this type of legal relator according to their specializations (right–duty to an omission or right–duty to an act).

Rights to an omission – right-duty relator		
Legal Position	Description	Example
Rights to the non-obstruction of acts	The duty-holder must not prevent or hinder certain acts of the right-holder.	*The right to express an opinion*. Right-holder: person (citizen, non-citizen); Duty-holder: State Act: non-obstruct person to express an opinion. If "Person **a** has the **right**, against **s**, to express an opinion" then "State **s** has the **duty** not to obstruct **a** in expressing an opinion".
Rights to the non-disruption of characteristics and situations	The duty-holder must not adversely affect certain characteristics and situations of right-holder.	*The inviolability of the confidentiality of correspondences.* Right-holder: person (citizen, non-citizen) Duty-holder: State Act: non-disrupt characteristics and situations If "Person **a** has the **right**, against s, to the inviolability of the confidentiality of correspondences" then "State **s** has the **duty** not to disrupt the characteristic (or situation) of **a** to inviolability of the confidentiality of correspondences".
Rights to non-removal of legal positions	The duty-holder must not remove certain legal positions of the right-holder. The existence of a legal position means that a corresponding norm is valid. Removing a certain legal position of right-holder is similar to derogating particular norm.	*The right to express an opinion*. Right-holder: person (citizen, non-citizen) Duty-holder: State Act: non-remove legal positions If "Person **a** has the right, against **s**, to express an opinion" then "Person **a** has the right, against **s**, that s should not remove from **a** legal position to express an opinion".
Rights to an act – right-duty relator		
Rights to factual act	The duty-holder must act when a fact exists.	*The right to education.* Right-holder: child; Duty-holder: State Act: to educate a child If "Child **a** has, against **s**, the right to education" then "State **s** has the **duty** to undertake the positive factual act in order for **a** to be educated".
Rights to normative act	The duty-holder must create certain legal norms.	*The right to have a legal norm regulating the right to strike.* Right-holder: public employee Duty-holder: State Act: to create a legal norm regulating the right to strike If "Public employee **a** has, against s, the **right** to have a legal norm prescribing the right to strike" then "State **s** has the **duty** undertake the positive normative act of creating the legal norm which regulates the right to strike of the public employee **a**".

Table 1: Rights to something – right-duty relator

No-right–permission relator. A *no-right–permission legal relator* uses the *legal relation* no-right–permission (correlative) to bind *permission holder* and *permitter*. A *permission holder* is someone who has a permission (no–duty), against the *permitter*, to do (or not to do) something. Table 2 discusses examples of this type of legal relator according to their specializations (no-right–permission to an act or no-right–permission to an omission).

Legal Position	Description	Example
Permission to act	Permission to act is the no-duty not to act. The permission-holder may act.	*Permission to smoke in open place*. Permission-holder: smoker Permitter: State Act: smoke in an open place. "Smoker **a** has **permission** to smoke in an open place, against to State **s"** iff "State **s** has **no-right** to obligate smoker **a** not to smoke in open place".
Permission to omission	Permission to omission is the no-duty to act. The permission-holder may omit an act.	*Permission to do not join an association.* Permission-holder: person Permitter: State Act: not to join an association 'Person **a** has, against State **s, permission** not to join an association" iff State **s** has **no-right** to obligate person **a** to join an association.

Table 2: Permission to something – no-right-permission relator

4. Final Considerations

In this paper, we outlined the fragment of a legal core ontology grounded in a *foundational ontology* as well as based on Alexy's Theory of Constitutional Rights. The focus is on the representation of *rights to something* on a relational perspective, specifying concepts and its legal relations. For this, we used categories from UFO, especially, *relators* to represent the triadic relations between *holders* and *rights*, *duties*, *no-rights* and *permissions.* Although the theory is directed to *constitutional rights*, it was possible to use the structure of legal relations to model generic legal relations.

We presented the theoretical issues about legal theories and ontologies, discussing the importance of *foundational ontology* for building both core and domain ontologies, as well as the importance to use a legal theory as basis for legal ontologies. We emphasized that the choice of a legal theory should take into account the reality of our contemporary society. Nowadays, legal theories propose different solutions to solve problems not addressed by Logical-Normative Positivism.

As future work, we will extend the formalization of *rights (liberties* and *competences*). In addition, we will extend the study to the second part of Alexy's theory (*Weighing and Balancing).* Finally, we intend to validate the LCO using existing domain ontologies.

Acknowledgements
This research is funded by the Brazilian Research Funding Agencies CNPq (grants number 311313/2014-0 and 485368/2013-7) and CAPES/CNPq (402991/2012-5). Cristine Griffo is funded with a CAPES grant.

5. References

1. *Bobbio*, N.: Teoria da norma jurídica. EDIPRO, Bauru, SP (2001).
2. *Valente*, A., *Breuker*, J.: Ontologies: the Missing Link Between Legal Theory and AI & Law. Leg. Knowl. based Syst. JURIX 94 Found. Leg. Knowl. Syst. 138–149 (1994).
3. *Kelsen*, H.: Pure Theory of Law. The Lawbook Exchange, LTD, New Jersey (2005).
4. *Cossio*, C.: Panorama de la Teoria Egologica del Derecho. Rev. Trimest. Cult. Mod. pp–67–94 (1948).
5. *Alexy*, A.: Teoria dos Direitos Fundamentais. Ed. Malheiros, São Paulo (2011).
6. *Reale*, M.: Lições Preliminares do Direito. Ed. Saraiva, São Paulo (2009).
7. *Dworkin*, R.M.: Taking rights seriously. Harvard University Press, Cambridge, MA (1978).
8. *Guizzardi*, G.: Ontological Foundations for Structural Conceptual Model. Universal Press, Veenendaal, The Netherlands (2005).
9. *Lopes*, M., *Guizzardi*, G., *Baião*, F.A., *Falbo*, R.: Reverse Engineering A Domain Ontology To Uncover Fundamental Ontological Distinctions An Industrial Case Study in the Domain of Oil and Gas Production and Exploration. (2006).
10. *Guizzardi*, G.: Some Applications of a Unified Foundational Ontology in Business Modeling. Appliations of a Unified Foundational Ontology. pp. 345–367 (2005).
11. *Guizzardi*, R. A., *Falbo*, G., *Guizzardi*, R.S.S.: A importância de Ontologias de Fundamentação para a Engenharia de Ontologias de Domínio: o caso do domínio de Processos de Software. Ieee Lat. Am. Trans. 6, 244–251 (2008).
12. *Bix*, B.: On the Dividing Line Between Natural Law Theory and Legal Positivism. Notre Dame Law Rev. 75, 1613–1624 (2000).
13. *Hart*., H.: O Conceito de Direito. Fundaçao Calouste Gulbenkian, Lisboa, Portugal (1994).
14. *Alexy*, R.: Teoria da argumentação jurídica: a teoria do discurso racional como teoria da justificação jurídica. Landy Livraria e Editora, São Paulo (2001).
15. *Casanovas*, P.: A Note on Validity in Law and Regulatory Systems. Quad. Filos. i ciència. 42, 29–40 (2012).
16. *Valente*, A.: Legal Knowledge Engineering; A Modelling Approach. IOS Press, Amsterdam (1995).
17. *Breuker*, J., *Valente*, A., *Winkels*, R.: Legal Ontologies in Knowledge Engineering and Information Management. Artif. Intell. Law. 12, 241–277 (2006).
18. *Shaheed*, *Jaspreet*, *Alexander Yip*, and J.C.: A Top-Level Language-Biased Legal Ontology. ICAIL Workshop on Legal Ontologies and Artificial Intelligence Techniques (LOAIT) (2005).
19. *Ikeda*, M.: A uniform conceptual model for knowledge management of international copyright law. J. Inf. Sci. 34, 93–109 (2007).
20. *Schweighofer*, E., *Liebwald*, D.: Advanced lexical ontologies and hybrid knowledge based systems: First steps to a dynamic legal electronic commentary. Artif. Intell. Law. 15, 103–115 (2007).
21. *Palmirani*, Monica, Tommaso Ognibene, and L.C.: Legal Rules, Text and Ontologies Over Time. RuleML (2) (2012).
22. *Scharf*, J.: rOWLer - A hybrid rule engine for legal reasoning. Second Doctoral Consortium, 27th International Conference on Legal Knowledge and Information Systems (2014).
23. *Carvalho*, P.B.: Curso de Direito Tributário. Ed. Saraiva, São Paulo (2000).
24. *Cortes*, O.M.P.: Súmula Vinculante e Segurança Jurídica. Editora Revista dos Tribunais, São Paulo (2008).
25. *Alexy*, R.: Constitutional Rights, Balancing, and Rationality. Ratio Juris. 16, 131–140 (2003).
26. *Hohfeld*, W.N.: Some Fundamental Legal Conceptions. Yale Law J. 23, 16–59 (1913).
27. *Hohfeld*, W.N.: Fundamental Legal Conceptions as Applied in Judicial Reasoning. Fac. Scholarsh. Ser. Paper 4378, (1917).
28. *Guarino*, N.: Formal Ontology in Information Systems. Formal Ontology in Information Systems (FOIS). pp. 3–15. IOS Press, Trento, Italy (1998).
29. *Gruber*, T.: Toward principles for the design of ontologies used for knowledge sharing. Int. J. Hum. Comput. Stud. 43, 907–928 (1995).
30. *Staab*, S., *Studer*, R., *Schnurr*, H.P., *Sure*, Y.: Knowledge processes and ontologies. IEEE Intell. Syst. Their Appl. 16, 26–34 (2001).
31. *Oberle*, D.: Semantic Management of Middleware. Vol. 1. Springer Science & Business Media (2006).

32. *Nardi*, J.C., *Falbo*, R.D.A., *Almeida*, J.P. a, *Guizzardi*, G., *Pires*, L.F., *Van Sinderen*, M.J., *Guarino*, N.: Towards a commitment-based reference ontology for services. Proc. - IEEE Int. Enterp. Distrib. Object Comput. Work. EDOC. 175–184 (2013).

33. *Masolo*, C., *Borgo*, S., *Gangemi*, A., *Guarino*, N., *Oltramari*, A.: IST Project 2001-33052 WonderWeb Deliverable D18. Ontology Infrastructure for the Semantic Web. (2003).

34. *Caralt*, N.C.: Modelling Legal Knowledge through Ontologies. OPJK: the Ontology of Professional Judicial Knowledge, (2008).

35. *Valente*, A., *Breuker*, J.: Towards Principled Core Ontologies. Proceedings of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems (1996).

36. *Breuker*, J., *Muntjewerff*, A., *Bredewej*, B.: Ontological modelling for design educational systems. Proceedings of the AI-ED 99 Workshop on Ontologies for Educational Systems (1999).

37. *Kralingen*, R. *Van*: A Conceptual Frame-based Ontology for the Law. Proceedings of the First International Workshop on Legal Ontologies. pp. 6–17 (1997).

38. *Hage*, J., *Verheij*, B.: The law as a dynamic interconnected system of states of affairs : a legal top ontology -. Int. J. Human-Computer Stud. 51, 1043–1077 (1999).

39. *Gangemi*, A.: Design patterns for legal ontology construction. LOAIT. pp. 65–85 (2007).

40. *Breuker*, J., *Hoekstra*, R.: Epistemology and ontology in core ontologies: FOLaw and LRI-Core, two core ontologies for law. Proceedings of the EKAW*04 Workshop on Core Ontologies in Ontology Engineering (2004).

41. *Stamper*, R.K.: The LEGOL 1 prototype system and language. Comput. J. 20, 102–108 (1977).

42. *Hafner*, C.D.: Representation of knowledge in a legal information retrieval system. Proceedings of the 3rd annual ACM conference on Research and development in information retrieval. pp. 139–153 (1980).

43. *McCarty*, L.T.: A language for legal Discourse I. basic features. Proceedings of the 2nd international conference on Artificial intelligence and law (1989).

44. *Mommers*, L.: Knowing the law. Legal Information Systems as a Source of Knowledge. (1999).

45. *Rosemann*, M., Vessey, I., Weber, R.: Alignment in Enterprise Systems Implementations: The Role of Ontological Distance. ICIS 2004 Proc. (2004).

46. *Machado*, A.L., *Oliveira*, J.M.P. de: A Legal Ontology of Relationships for Civil Law System. Proceedings of the 1st Joint Workshop ONTO.COM / ODISE on Ontologies in Conceptual Modeling and Information Systems Engineering (2014).

47. *Guizzardi*, G., *Falbo*, R., *Guizzardi*, R.S.S.: Grounding Software Domain Ontologies in the Unified Foundational Ontology (UFO): The case of the ODE Software Process Ontology. CIbSE. pp. 127–140 (2008).

48. *Almeida*, J.P. a, *Guizzardi*, G.: An ontological analysis of the notion of community in the RM-ODP enterprise language. Comput. Stand. Interfaces. 35, 257–268 (2013).

49. *Larenz*, K.: Metodologia da Ciência do Direito. Fundaçao Calouste Gulbenkian, Lisboa (1997).

50. *Goffman*, E.: A Representação Do Eu Na Vida Cotidiana. Vozes, Petrópolis (2002).

VISUALIZATION AS A TERTIUM COMPARATIONIS WITHIN MULTILINGUAL SCIENTIFIC COMMUNITIES

Vytautas Čyras[1], Friedrich Lachmayer[2], Erich Schweighofer[3]

[1] Associate Professor, Vilnius University, Faculty of Mathematics and Informatics
Faculty of Mathematics and Informatics, Naugarduko 24, 03225 Vilnius, Lithuania
Vytautas.Cyras@mif.vu.lt; http://www.mif.vu.lt/~cyras/
[2] Professor, University of Innsbruck, Faculty of Law
Innrain 47, 6020 Innsbruck, Austria
Friedrich.Lachmayer@uibk.ac.at; http://www.legalvisualization.com
[3] Professor, University of Vienna, Faculty of Law, Centre for Computers and law
Schottenbastei 10-16/2/5, 1010 Vienna, Austria
Erich.Schweighofer@univie.ac.at; http://rechtsinformatik.univie.ac.at

Keywords: ***Legal informatics, legal visualization, knowledge visualization, legal meaning, data science, knowledge representation, semantic networks***

Abstract: *Legal data science as part of modern legal informatics serves as the integrative model of computer-supported representation and analysis techniques. A major part of the analysis consists of structural visualization, which deals with logical diagrams and represents the semantics of law.*

Visualizing legal meaning differs from representing it as text, because its greater and easier expressiveness make it able to capture structural relations between documents, legal concepts or events. Visualizations of timelines, events and concepts are commonly used, but only hint at the great potential of visualization. Results from legal theory research, in particular tertium comparationis, *are not well known but are highly relevant. Relations between two entities can be manifold and are often insufficiently expressed in legal language. Visualization as* tertium comparationis *represents these relations but also constitutes an intermediate step towards a formal and computer-useable representation.*

Our own model on multilingual legal systems takes into account that only English is now the reference language for translations. Thus, the legal requirement of the equal treatment of all languages is disregarded in practice. Therefore, we propose the visualization of meaning as tertium comparationis *to act as a common element for all 24 different linguistic versions. We see two discourse patterns. The use of English shows the top-down pattern. Another concept,* tertium communicationis, *denotes the third part of the communication between two agents who speak languages A and B, respectively. The use of other languages – e.g., German – shows the bottom-up pattern. Besides visualization, other intermediate formats such as XML schema are targeted. We aim to use* tertium communicationis *as a conceptual definition that improves communication. We see two directions for the development of ideas on visualization: 1) from the natural language to a professional legal language and then to a formal technical language, and 2)* vice versa. *In the transition text-visualization-model-metamodel, we see two ways of producing* tertium comparationis*: via visualization and via model/metamodel. Next, we classify legal relations according to Is-Ought combinations and approach an ontology of legal relations. The applications show the already high but not well recognized potential of visualization*

in representing the deep structure of legal systems, and its relations with sociology, economics and public policy.

1. Introduction

The achievements of research in AI and Law are not much recognized in law schools. Even in the knowledge and network society, legal methodology has not much changed at all. The trias of legal dogmatics, handcraft and art covers the insufficient methodological state of the legal discipline. insufficient account is given to legal theory and modern technologies. An example can be seen in the use of legal information systems. Practical training is now standard, but scientific reflection is still insufficient.

Legal informatics has developed new methods for representation, analysis and synthesis of legal materials. Schweighofer (2015) has structured these analytical tools as legal data science. His model of 8 views, 4 methods and 4 synthesis describes the eight different representations of a legal system, four computer-supported methods of analysis, which lead to synthesis, a consolidated and structured analysis of a legal domain, either a commentary, an electronic legal handbook or commentary [Schweighofer 2011], a citizen representation or a case-based synthesis. A more detailed description can be found in Section 2.

This paper describes in detail the method of visualization, and in particular visualization as *tertium comparationis* in legal informatics and in multilingual scientific communities. *Tertium comparationis* (Latin – the third [part] of the comparison) is the quality that two things that are being compared have in common (http://en.wikipedia.org/wiki/tertium_comparationis, Fig. 1).

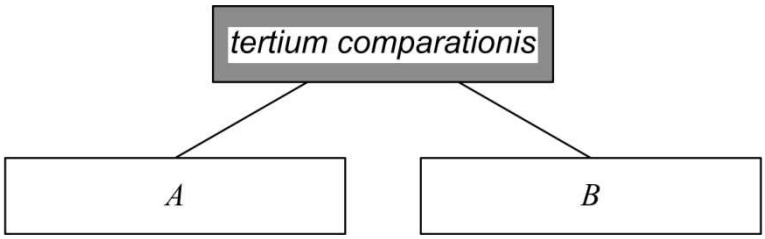


Fig. 1: An indirect relation between A and B through tertium comparationis, the common property

On the one hand there are formal notations that go beyond the textual ones; on the other hand, there are visual representations that also occur in competition with the text. In turn, two different types of visualizations can be distinguished: first, visualizations formed according to strict formal rules; and second, more intuitive pictures that can describe situations better.

The remainder of this contribution is structured as follows: related work (Section 2), visualization as *tertium comparationis* (Section 3), top-down and bottom-up communication (Section 4), legal relations (Section 5) and conclusions (Section 6).

2. Related Work

In the **legal data science** ("Rechtsdatalystik") model of Schweighofer (2015), the four views of [Lu and Conrad 2012, 2014] are extended by adding four other views, four methods and four synthesis. The basis is the textual representation, the text (multimedia) corpus, which consists of primary sources (e.g., statutes, regulations, court cases, and administrative decisions), and secondary sources (e.g., descriptive and analytical legal publications). Secondly, the annotation view consists of legal documentation (bibliographical data, topical classifications, thesaurus descriptions, and expert

annotations (e.g., Westlaw's headnotes[2]), which relies on a legal taxonomy. Thirdly, using long-standing experience in cross-references, the multiplicity of both out-bound (cited) sources and in-bound (citing) sources can also be exploited as the citation network view. Advanced citation does not stop with the document but goes to the granularity of these citations at a document segmentation level (e.g., articles, sections, lists etc.). Such citations can be weighted by citation frequency (citing or cited). Fourthly, a modern search engine can aggregate user behavior. Respecting data protection and thus disregarding individual behavior, the accumulated evidence represents the numbers of views, prints, citation checks etc. for a document. Lu and Conrad's list is extended by adding the logical view, the ontological view, visualization and the argumentation view. Logical representation describes the legal system as a set of first order logic statements, structured in time layers (one per day) and quantifiers identifying the (possible) persons concerned. It is strongly linked to logic programming, [cf. Sergot et al. 1986]. The main advantage of logical representations lies in the potential for the automation or semi-automation of case handling. Using intelligent forms or digital pictures for the descriptions of facts, the logic program can automatically apply rules for a given date and particular persons. Ontological representations are computer-useable conceptualizations of the domain. In law, ontologies describe both the legal conceptualization and the factual conceptualization, i.e., both a legal ontology and a world ontology (e.g. a "common sense" ontology like Cyc) are relevant. Ontologies enhance legal analysis with computer-useable concepts and their relations [Schweighofer 2011]. Legal visualization concerns the use of graphics, images and videos for visual representation of the law [Brunschwig 2011]. The potential of visualizations for citizens information is obvious. Graphical notations are also a strong support for a formalised view of the law. The key features are represented by images or graphics, even in cases where the necessary level of abstraction for formalization is not yet reached. In recent years, the field of AI and Law has strongly concentrated on the formalization of arguments. This case-based reasoning approach started in the 1980s, and culminated in Ashley's book (1990). Gordon added a more theoretical approach with his pleadings game and his formalization of Alexy's theory of legal argumentation [Gordon 1995]. Taking into account the dialectical nature of the legal process – thesis (plaintiff), antithesis (respondent), synthesis (judge) – a representation of possible arguments is very important and useful. Recent research can be found in the ICAIL 2013 and JURIX 2014 proceedings. Argumentation representation is also part of the textual view, but can be much less explicit, and is found in judgments and legal briefs.

Legal methods start with reading, finding, understanding and interpreting the law. For this manual process, books are sufficient, but a legal information system provides a much better and more efficient knowledge platform. This documentation method consists in collecting all relevant sources, adding metadata and making the documents available on the Internet, with or without requiring payment of a fee. The main methodological add-on of legal informatics in the interpretation of the law is the search engine. Modern search techniques are indispensable for finding the only appropriate document in a collection of millions of documents. Search is based on an understanding of legal vocabulary, combined with metadata. The popularity of search engines means that legal searches are too often made easy rather than powerful. New approaches try to include elements of semantic searching, following Google search techniques. Legal search is an important IT support in the interpretation process, because it finds and analyzes relevant documents. A developing area that is becoming more important can be seen in the ranking of legal documents. The structural analysis consists in the re-writing of rules as logical statements or conceptual structures. In both cases, paper and electronic representations can be used. Without appropriate and fine-tuned conceptual structures and rule frames such as decision trees, the application of rules remains cumbersome and time-consuming. For any well-defined process, this analysis is

[2] West's Key Number System: http://info.legalsolutions.thomsonreuters.com/pdf/wln2/L-374484.pdf (accessed 30 April 2015).

indispensable for automation or semi-automation. Further, (semi-)automated linguistic methods can be very helpful. In this paper, we deal only with one part of legal visualization, the graphical representations, in particular with the visualization of the abstraction of the law [Lachmayer 2002], as an analytical tool. An overview of legal visualization can be found in [Röhl and Ulbrich 2007]. The use of legal visualization in citizen's information systems is obvious (e.g., Europe Direct or the Austrian HELP.gv.at[3]). The complexity of legally relevant events, actions, and documents is structured and put into a proper timeline that is sufficiently clear for laypeople in such situations. The private sector, however, is much more advanced, using legal visualizations for highly complex regulations like rent law or tax law [Kahlig & Stingl 2011].

The synthesis changes dramatically in the "knowledge and network" society because of the much more powerful views and methods. It is no longer only text that has to be interpreted and analyzed. All eight views and the four analytical methods have to be taken into account. The methods of synthesis of these results are various. Here, four main methods are singled out: manual commentary, Dynamic Electronic Legal Commentary (DynELC), citizen information system and case-based synthesis. The legal commentary is the appropriate form of representation of the knowledge of a legal system offering in a systematic analysis all relevant elements for a comprehensive and holistic understanding of the particular area of law. The concept of an electronic generation of these data in a more formalistic way already exists with the Dynamic Electronic Legal Commentary (DynELK) [Schweighofer 2011]. Metadata for the text corpus are generated (semi-)automatically, and added in a computer-useable way. This process comprises document categorization, semi-automatic generation of thesaurus descriptors, automatic generation of hypertext links, and automatic generation of temporal relations. Ranking means comparing the document with the search request, and considering the document in the text corpus, the document in the citations network and the document in the timeline. Text summary comprises the semi-automatic generation of summaries of documents. Multilingualism comprises automatic document translation (e.g., Google Translate). The legal subsumption is supported by an inference machine. Citizens Information Systems use the internet to spread easily understandable public information. In practice, the case-based synthesis for each specific case is crucial. All existing sources and syntheses should be used to best present their own legal position in relation to each key authorities.

Legal visualization is a view of the legal system but also a powerful method. For analytical purposes, it is used for legal risk management, describing relations between business and law in business information systems and legal theory. Visualizing legal risk is addressed in the works of [Mahler 2010, 2013]. For the visualization of factual risk he adapts the CORAS graphical language [Vraalsen et al. 2007]. According to [Susskind 2013], legal problem solving will be much less significant in the future. The emphasis will shift towards legal risk management: risks should be understood and identified, and also controlled, before there is an escalation. Various relations between events, legal provisions, and risks can be described much better using visualizations. Enterprise software is the core of the knowledge systems of the business world. [Heddier and Knackstedt 2012, 2013] have used visualization to describe relations between the business world and the legal system. Because they are potentially much easier to understand, visualizations are a good choice for the representation of law in business information systems. An important potential consists in the consideration of the conversion of visualizations into computer code and *vice versa*. We use *tertium comparationis* to describe the quality that two things to be compared have in common. A major field of research are multilingual relations.

[3] Websites http://europa.eu/europedirect/index_en.htm and https://www.help.gv.at/Portal.Node/hlpd/public/en (accessed 30 April 2015).

3. Visualization as *Tertium Comparationis*

Tertium comparationis describes the quality that two things that are being compared have in common. Two different types can be distinguished: first, visualizations formed according to strict formal rules; and second, the more intuitive pictures that can describe situations better. Normally, visualization is a prerequisite for a more formal representation.

There are also quite different approaches to visualization – through semiotics (Fig. 2), for instance. The classical philosophy of law, however, as approximately represented by Arthur Kaufmann [Lachmayer 2005], has provided a methodological introduction to visualization with the thought pattern of *tertium comparationis*. In the European Union with its many official languages, in particular, visualization, which appears as a *tertium*, can form a mental bridge between the different languages.

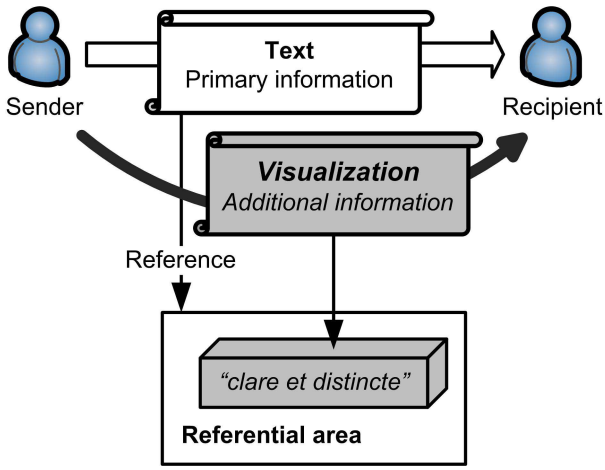


Fig. 2: A text is communicated from a sender to a recipient. A visualization refers to clear and distinct knowledge that contributes to understanding.

The lack of pictures in jurisprudence becomes a learning obstacle [Röhl & Ulbrich 2007, pp. 15–17]. A starting position is "Law is text", and therefore law is always textual for jurists. Hence, there are reasons for jurists' reluctance to visualize. Pictures can have drawbacks, such as redundancy, a low level of abstraction, trivialization, and emotions [ibid., pp. 18–25, 100–102]. However, the use of logical pictures (*logische Bilder*) can bring advantages. Metaphors and symbols can be employed to represent norms, and thus pictorial two-dimensional representations emerge [ibid., pp. 42–62]. Communicating the meaning of law to the human user is of primary importance in legal education. The visual structure is a diagram, which represents the meaning. Diagrams serve well as visualizations of legal norms [*Rechtsnormbilder*, ibid., pp. 109–111]. Besides pictorial visualizations, logical diagrammatical visualizations such as argumentation graphs, storytelling, and legal workflow, including info-graphics, are widely used to represent legal content.

4. Top-down and Bottom-up Communication

We see two communication patterns in multilingual discourses: *top-down* communication and *bottom-up* communication. Different languages can be used in scientific discourse. Therefore, two situations arise regarding the discourse language. On the one hand, English, a global language, can be used. (Other standards such as Latin could also be used, and thus the role of *lingua franca* emerges.) This is the *top-down* pattern. On the other hand, other (working) languages, such as German or French, can also be used. This is the *bottom-up* pattern. Native languages allow a scientist to unfold his ideas more naturally, and the discourse becomes more creative and productive. Hence, the bottom-up approach is also more meaningful than the top-down one.

Translation. The use of different languages brings translation problems. Therefore, dictionaries and translation machines, such as IATE (formerly Eurodicautom) emerge.[4] Currently, semi-automatic extraction from legal texts is being addressed in various projects [cf. Francesconi 2012 and Yoshida et al. 2013; one of the first extensive works in the legal area was Schweighofer 1999].

Visualization supplements translation. It is quite possible to go a long way around from one language into another language by going via a third language. This bridge language is then the *tertium translationis*. Examples of this bridge language being visualization can be found in books for visualized learning, where illustrations complement word translation; for instance, from the English 'table' to the German 'der Tisch' (Fig. 3). In this way, visualization supplements translation and brings an additional syntactic dimension to natural languages. Vividness is increased in the course of translation, so speakers obtain additional contemplation capabilities, and their discourse becomes more efficient. The more often use of visual dictionaries, in particular for languages like Japanese or Mandarin, is evident. Thus, visualization is shown to be important beyond legal informatics.

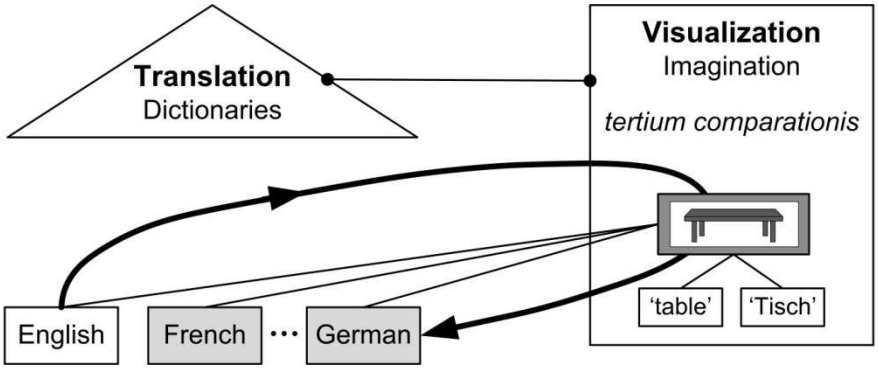


Fig. 3: Translation with visualization

Lettering. A special situation occurs with worded visualizations. So far as the pictures are involved, no translation is required, since the pictures can be more or less "read" in all languages. If a visualization is to be offered in another language, the wording must be replaced. Here, the *tertium comparationis* consists either in a text system or in the visual elements themselves, because they have a common reference to the different language versions.

Wording brings semantics to visualization and may have various forms, such as figure captions, explanations, footnotes, labels, inscriptions, etc. A picture without a description is simply a graphic structure and can be viewed as mere visual chaos without semantics; it is therefore not acceptable in a discourse. The description could be in English and in other languages. Thus, the top-down and the bottom-up approaches can also be used in wording.

4.1. *Tertium Communicationis* in Communication

Tertium communicationis is not a word play: we are introducing a new term to denote the third part of communication (Fig. 4). Suppose a translation from language *A* to language *B* is being performed. Besides visualization, other intermediate formats can be employed in translation. Nowadays, an intermediate format can be an XML Schema.[5] We use *tertium communicationi*s as a

[4] Eurodicautom, created in 1975, was the pioneering terminology database of the European Commission. In 2007 Eurodicautom was replaced by Inter-Active Terminology for Europe (IATE), the inter-institutional terminology database of the European Union (http://iate.europa.eu).

[5] XML, Extensible Markup Language, is a markup language for encoding documents in a format that is both human-readable and machine-readable.

conceptual definition of something that improves communication between human beings or machines. This communication need not be visual. Text is not just verbal and in the end a textual document has a layout, its graphic structure. The question "Which formats contribute to better communication?" depends on various factors, such as the document type and the communication task, and is worth a separate study. Intermediate formats have their syntax and semantics.

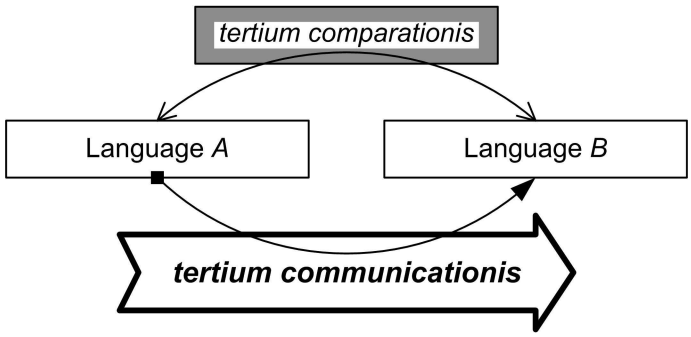


Fig. 4: *Tertium communicationis* as an intermediate format

Converting a *tertium comparationis* into a *tertium communicationis* can make an indirect relation more dynamic and personal. This conversion leads further, to *tertium identificationis* and *tertium socialisationis*.

4.2. Two Directions: from Natural Language to Professional Juristic Language and *Vice Versa*

In the projects that produce legal visualizations, we single out two directions for the development of ideas: first, from the natural language to a professional language (legal language) and then to a formal technical language (Fig. 5 a), and, second, vice versa, from a professional legal language to the natural language (Fig. 5 b). Laypeople speak the natural language and jurists speak their professional language(s).

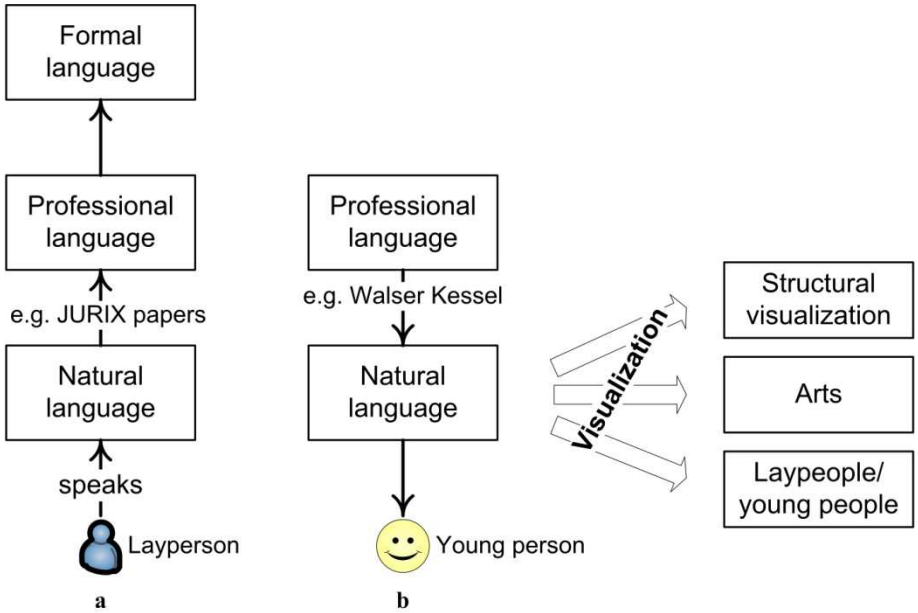


Fig. 5: Two directions: a) from a natural language to a professional language, b) vice versa

The first direction can be observed in various projects [e.g. from the annual JURIX conferences, Francesconi 2012 and Yoshida et al. 2013]. The second direction is demonstrated in Walser Kessel's (2011) informative book about law for young people.

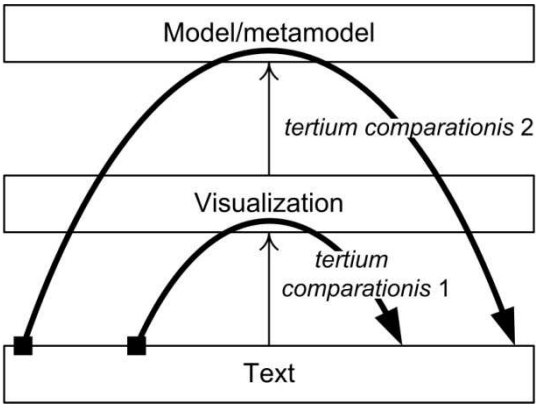


Fig. 6: Two ways of producing *tertium comparationis*: via a visualization and via a model or a metamodel

We point to three kinds of legal visualization:

1. *Structural visualization*; see Lachmayer's PowerPoint presentations, http://jusletter-it.weblaw.ch/visualisierung/
2. *Arts*. Examples are novels and films about legal matters and also pictures and statues of Themis, etc.
3. *Explaining law to laypeople or young people*.

A topic to explore is the transformation of syntax when a diagram is produced from a text. For example, the text layout and font have to be changed to communicate legal content for young people.

4.3. From Text to Visualization and to Model

There are two ways to move from a text in one language to a text in another language. One way is via visualization, as we have discussed above. This path is shown in Fig. 4 and also Fig. 6 as the *tertium comparationis* 1 arch. However, there is another way – via the level of model/metamodel [Fill 2014a; Fill 2014b]. This way is shown in Fig. 6 as the *tertium comparationis* 2 arch, and uses a model of the text, an ontology or a higher-level model, a metamodel.

4.4. Text-Visualization Correspondence

We see a correspondence between the textual world and the world of visualization. This correspondence is shown in Fig. 7, where the traditional model-driven development infrastructure, which is addressed by Atkinson & Kühne (2003), is taken into account. An example of a modeling language in software development is UML [Booch et al. 2005] or SysML [Weilkiens 2014].

We now explain the correspondence. Let us start from the world of textuality (Fig. 7). Metadata descriptors are extracted from texts. Next, thesauri appear beyond texts and metadata. Then, beyond thesauri we place legal ontologies [cf. Guarino et al. 2009].

The visualization world is shown on the right in Fig. 7. Pictures, photos, and other visually sensed raw materials correspond to texts. Above them we place *structural visualization*, which denotes the graphical representation of the legal meanings of the texts. Above that we place *meta-visualization*, which addresses the methods of visualization and their components [cf. Fill & Karagiannis 2013].

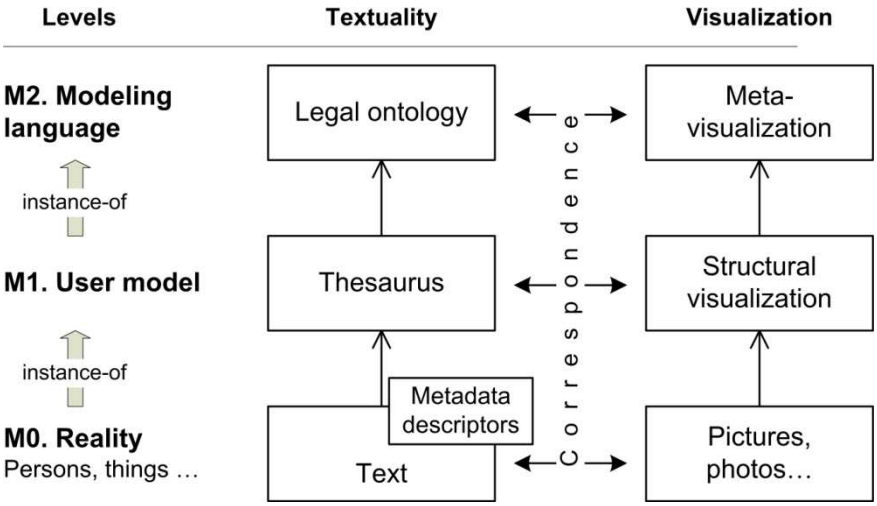


Fig. 7: Correspondence between textuality and visualization at different levels of abstraction

4.5. Visual Products as *Tertium Comparationis*

We can see different examples of visualizations that can serve as *tertium comparationis* products in law. A starting point is verbal metaphors. A pyramid represents the hierarchical structures of the branches of law or legal sources. Then comes a bridge (e.g., connecting the banks of law and technology), step working, etc. Here we can revert to the point of view that legal terms are also metaphors and have a specific meaning [Lakoff & Johnson 2003].

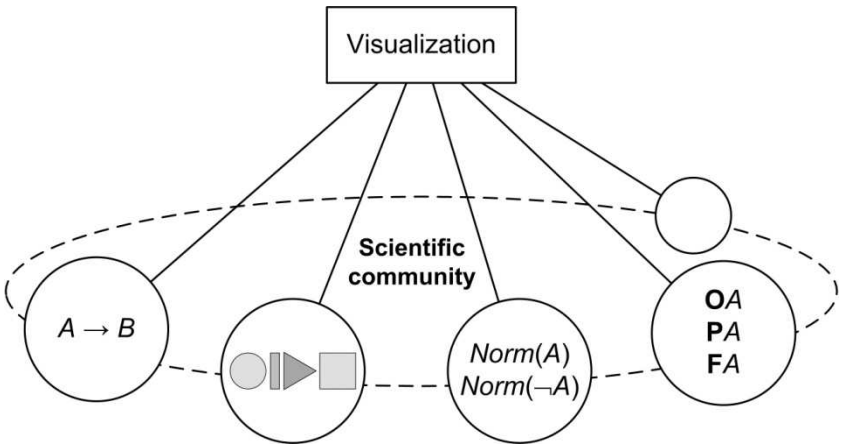


Fig. 8: A metaphor of different models for legal visualization

As ideal visual models we would mention the globe, the solar system, the atom model that is composed of a nucleus made of protons and neutrons surrounded by a cloud of electrons, and molecule models. Here we stress that we are talking of pictorial models and not formal graphic models. There are different types of models, depending on the legal task, the domain of law, and the scientific community (Fig. 8).

5. Legal Relations

Arthur Kaufmann replaces ontologies of substances by ontologies of relations [Lachmayer 2005]. Legal relations are relations between different kinds of elements, for example, a) civil obligations between persons, *vinculum juris,* i.e. "bonds of law," b) relations between movable/fixed assets, and c) relations between the facts of a matter and a circumstance (in German *Tatbestand (Tatsache) und Sachverhalt*). It is not straightforward to model a legal relation as a mathematical relation. A

relation R over the sets $X_1,\ldots, X_n$ is defined as a subset of its Cartesian product, written $R \subset X_1 \times \ldots \times X_n$. A relation can be represented as a table. The legal meaning of legal relations differs from the concept of a (relational) structure in philosophy and from the concepts of an extensional relational structure, an intensional relational structure, and an ontology in computer science [cf. Guarino et al. 2009].

Next, we take into account explicit and implicit relations and also direct and indirect relations.

5.1. Indirect Relations and *Tertium Comparationis*

Indirect relations. *Tertium comparationis* is the case of a relation that does not lie directly between one element and another, but goes through a third. However, a course through *tertium comparationis* modifies the relation. With *tertium comparationis* one deals not with a direct relation between two elements, but, rather, with an indirect relation between them that is mediated over a third element. This indirect relation is a reflected relation and can also be characterized as a broken relation. A broken relation, a direct one, is replaced by two relations. For instance, a translation from Portugal into Lithuanian would be performed not directly, but through English. Another example is making two information systems interoperable. Interoperability needs making a bridge between the systems.

In order to explain the nature of *tertium comparationis* we provide the following example. Suppose four apples are being brought into relation with four pears. This is about the number, in this case about the number four, which occurs as *tertium comparationis*. It does not compare apples with pears but compares four elements with four other elements. A comparison can be performed through other common qualities such as “fruitiness”.

Overcoming barriers with *tertium comparationis*. A reflected or broken *tertium comparationis* is able to make a connection through walls or other barriers. The situation is similar to a mirror, which allows one to survey areas that cannot be viewed directly. In this way, one can see not only the present, but also the past and the future. *Tertium comparationis* is a suitable technique to make connections in the unconscious, as they cannot be made directly.

Projecting a relation. Legal relations are generally not simple matters. In most cases a relation is not like a bridge between two banks because it is not even observable in the outside world. Often, relations are projected and a relation becomes the result of projecting. Hence, projection is the content of a thought act, a speech act, or a legal act.

Comparison. A comparison also concerns relations. Various elements can be compared and hence brought into a relation. If a relationship is projected, the elements that are connected in the relation are also projected. Hence, a) Is can be compared with Sense (*Sinn*), b) Sense with Is, and c) Sense with Sense.

Interpretation and comparison. A classical usage is a relation between the facts (*Tatbestand*) of a matter and the circumstances (*Sachverhalt*), which are expressed in a norm. It is meaningful to examine this relation because it usually appears in judgments, i.e. legal acts. We hold that interpretation precedes comparison. The fact and the circumstance are compared not directly but through their sense that is projected onto the fact and the circumstances, respectively (Fig. 9).

In legal language, it is not the case that a fact (which appears in the Is world) is compared directly with the circumstances of a norm, but the interpreted fact is compared with the sense of the norm’s circumstances (which appears in the Ought world). The interpretation (*Deutung*) is a prerequisite. The comparison compares the sense-structure of the fact with the sense-structure of the circumstances. Legal terms serve as *tertium comparationis*.

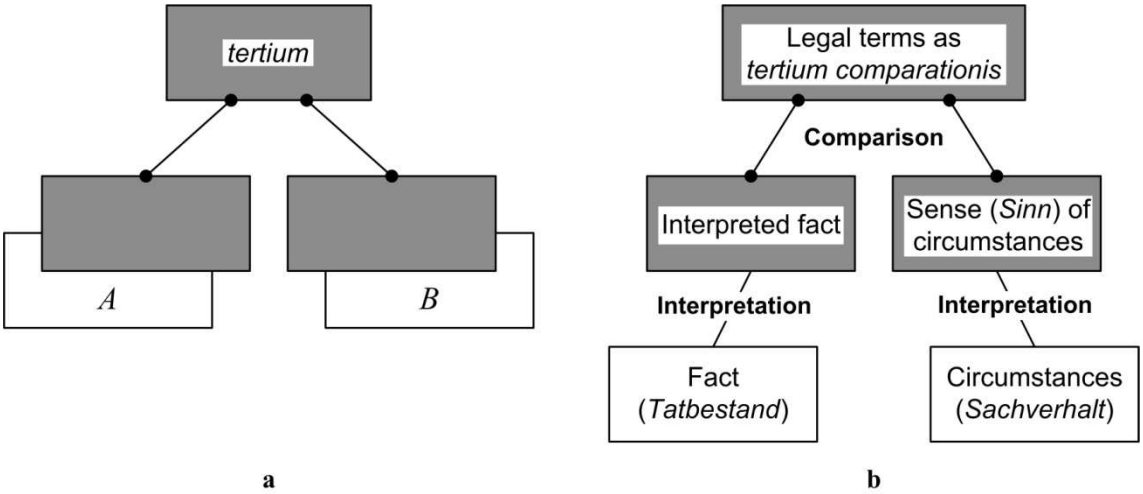


Fig. 9: An indirect relation between *A* and *B* through a common quality *tertium comparationis*: a) pattern and b) explanation

Pretextual universals. A textual culture dominates in law and, therefore, there is little that is pretextual or non-verbal. However, there are also normative approaches that are centrally non-verbal. Examples are the simulated measurement units of the body, such as the radius or the cubit, the foot or the step. Hence, there are archetypes that are non-linguistic and have a social normative effect.

Subject-internal tertium comparationis. We spoke above about the abstract structural background that lies behind universal interpretation schemas such as language, types, and terms, and that thus lies behind supposed objectivity. However, another course can be followed to facilitate *tertium comparationis*, specifically through the subject. Universals can also be derived from the subject. There are *universalia in rem* that are internally in the subject (Fig. 10); they differ from *universalia ante rem* that are in the objective area before the subject and the thing. These universals can, but need not, be formulated verbally. Such indirect relationships can be produced in the subject for a preliminary understanding. Since we hold that language is a distinct human competence, the pre-verbal ability may be associated with the development stages before humans. A comparison is also possible, to a certain extent, and thus a thought. The big advantage of language is less in the standardization in the projected sense, but rather in the inter-subjectivity.

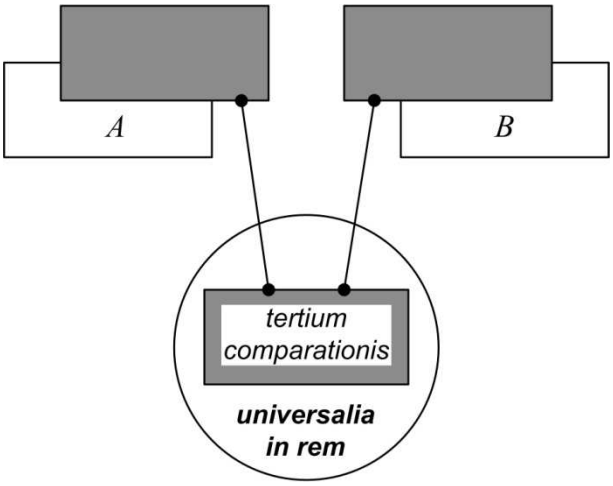


Fig. 10: An indirect relation between A and B following the course universalia in rem

Two poles of *tertium comparationis*. There are thus two poles of *tertium comparationis* – namely, *universalia ante rem* that is assigned to the objective and *universalia post rem* that is attributed

subjectively. Although you can find such comparison measures in different areas, they are still functionally lifted from the things whose conceptual link they make possible.

Relations and personality. Relations are assigned to the sense level. There are many different types of relations, especially in the area of law. If a case is brought into a relation with a norm, the projecting onto the relation of correspondence is performed. However, it is different with complementary roles. Here there is something like a *vinculum juris* between people. The personal relation of the complementary roles of two or more persons is probably what Arthur Kaufmann had in mind when he developed his theory of the person.

Substance of *tertium comparationis*. The question "What is the substance of *tertium comparationis*?" is not trivial. A *tertium comparationis* such as the meter or the kilogram (of the International Bureau of Weights and Measures) can be assigned a concrete substance. However, the substance of *tertium comparationis* can be weakened; think, for example, of merely projected units of measurement. Here the substance is not as clear as in the case of concrete *universalia in rem* examples like the meter or a cardboard/computer model of a house that is built by an architect.

5.2. Ontology of Relations

Stressing the ontology of relations is a radical step that is interesting from a linguistic viewpoint. However, the practical consequence of this step has not been sufficiently considered. Is it in fact the case that only relations, and not the substances that are associated with them, are real? Through the elimination of the substances one falls into a bottomless abyss, and the relations alone are not able to slow down this fall. An attempt to visualize the ontology of relations is shown in Fig. 11. This ontology can be treated as a classification of relations, which are grouped according to 'Is'/'Ought'-'Is'/'Ought' combinations. The proposed concept of the ontology of relations is at a very abstract level, and does not conform entirely to the treatment of ontologies in computer science [cf. Guarino et al. 2009].

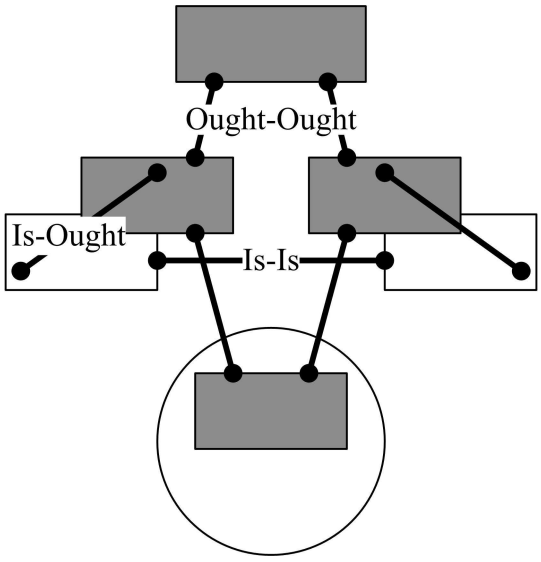


Fig. 11: Towards an ontology of relations

Arthur Kaufmann made a radical change to Aristotle's category. Relation is a category for Aristotle. Aristotle replaced one category by a different category. Like Arthur Kaufmann, Hans Kelsen stressed this relational character at least of subjective law, in which he defined the person as an embodiment of rights and obligations. For Kelsen this was possibly an attempt to reconsider the traditional concept of the person in its figure (*Gestalthaftigkeit*) and to suspend it dialectically, especially in order to understand it from his ideology-critical approach.

6. Conclusions

Modern legal informatics theory developed legal data science for computer-supported representation and analysis techniques. A major part of the analysis consists of structural visualization that deals with logical diagrams and represents the semantics of law.

In this contribution, we focus on visualizations that can serve as *tertium comparationis*. In a multilingual scientific discourse we see two communication patterns: top-down and bottom-up. Next, we introduce the concept of *tertium communicationis*, which facilitates communication between human beings or machines. In the production of legal visualizations, we single out two directions for the development of ideas: 1) from the natural language to a professional language (legal language) and then to a formal technical language, and 2) *vice versa*. We see two ways of producing *tertium comparationis*: 1) via visualization and 2) via a model/metamodel. Therefore, we show the correspondence between textuality and visualization at different levels of abstraction. Next, we provide a classification of legal relations based on 'Is'-'Ought' combinations. We conclude that the substance of *tertium comparationis* may not be trivial, as in the case of units of measurement.

7. Acknowledgement

V. Čyras has been supported by the project "Theoretical and engineering aspects of e-service technology development and application in high-performance computing platforms" (No. VP1-3.1-ŠMM-08-K-01-010), which is funded by the European Social Fund.

8. References

Ashley, K. D. 1990. *Modeling Legal Argument – Reasoning with Cases and Hypotheticals*. Artificial Intelligence and Legal Reasoning, MIT Press, Cambridge, MA.

Atkinson, C. and Kühne, T. 2003. Model-Driven Development: A Metamodeling Foundation. *IEEE Software* 20(5) (September 2003), 36–41. DOI= http://dx.doi.org/10.1109/MS.2003.1231149.

Booch, G., Rumbaugh, J., and Jacobson, I. 2005. *The Unified Modeling Language User Guide*, 2nd edition. Addison-Wesley, Reading.

Brunschwig, C. R. 2014. On visual law: visual legal communication practices and their scholarly exploration. In *Symbol and Magic of Law. Liber Amicorum Friedrich Lachmaye*r (Zeichen und Zauber des Rechts), Schweighofer, E., Handstanger, M., Hoffmann, H., Kummer F., Primosch, E., Schefbeck, G., and Withalm, G., Eds. Editions Weblaw, Bern, pp. 899–933. http://ssrn.com/abstract=2405378.

Brunschwig, C. R. 2011. Multisensory Law and Legal Informatics – A Comparison of How these Legal Disciplines Relate to Visual Law. In *Strukturierung der Juristischen Semanik – Structuring Legal Semantics, Festschrift für Erich Schweighofer*, Geist, A., Brunschwig, C.F. Lachmayer, F. Schefbeck G.(Hrsg.), Weblaw Verlag, Bern,. 573-667 (in German).

Fill, H.-G. 2014a. Transitions between syntax and semantics through visualization. In *Symbol and Magic of Law*, Schweighofer, E. et al., Eds. (Zeichen und Zauber des Rechts). Editions Weblaw, Bern, 935–944 (in German).

Fill, H.-G. 2014b. Abstraction and transparency in meta modeling. In *Transparency*, E. Schweighofer et al., Eds. OCG Vienna and *Jusletter IT*, February 2014, 47–54. http://jusletter-it.weblaw.ch/issues/2014/IRIS.html.

Fill, H.-G. and Karagiannis, D. 2013. On the conceptualization of modeling methods using the ADOxx meta modeling platform. *Enterprise Modelling and Information Systems Architectures* 8, 1 (March 2013), 4–25.

Francesconi, E. 2012. Supporting transnational judicial procedures between European member states: the e-Codex project. In *Legal Knowledge and Information Systems, JURIX 2012: The Twenty-Fifth Annual Conference*, Schäfer, B., Ed. IOS Press, Amsterdam, 41–50.

Gordon, T. F. 1995. *The Pleadings Game – An Artificial Intelligence Model of Procedural Justice*. Kluwer Academic Publishers, Dordrecht.

Guarino, N., Oberle, D., and Staab, S. 2009. What is an Ontology? In *Handbook on Ontologies*, S. Staab and R. Studer, Eds. Springer, Berlin Heidelberg, 1–17.

Heddier, M., and Knackstedt, R. 2012. Challenges of legal visualization from business informatics perspective (Herausforderungen der Rechtsvisualisierung aus Perspektive der Wirtschaftsinformatik). In *Transformation of Legal Languages*, Proc. of the 15th International Legal informatics Symposium IRIS 2012, Schweighofer, E., Kummer, F., Hötzendorfer, W., Eds. OCG, Vienna, 355–363 (in German).

Heddier, M., and Knackstedt, R. 2013. Empirical evaluation of legal visualization on example of handy contracts (Empirische Evaluierung von Rechtsvisualisierungen am Beispiel von Handyverträgen). In *Abstraction and Application*, Proceedings of the 16th International Legal informatics Symposium IRIS 2013, Schweighofer, E., Kummer, F., Hötzendorfer, W., Eds. OCG, Vienna, 413–420 (in German).

Kahlig, W. and Stingl, W. 2011. *Tax Law for Real Estate* (Immobilien-Steuerrecht). Manz, Vienna (in German).

Lachmayer, F. 2002.Visualization of abstract (Visualisierung des Abstrakten). In *IT Law and State* (IT in Recht und Staat, Aktuelle Fragen der Rechtsinformatik), Schweighofer, E., Menzel, T., Kreuzbauer, G., Eds. Schriftenreihe Rechtsinformatik, vol. 6. Vienna 309–317 (in German).

Lachmayer, F. 2005. Tertium comparationis in law. (Das tertium comparationis im Recht. Variationen zu einem Thema von Arthur Kaufmann). In *Responsible Law: the Philosophy of Law of Arthur Kaufmann* (Verantwortetes Recht – die Rechtsphilosophie Arthur Kaufmanns), U. Neumann, W. Hassemer, U. Schroth, Eds. Archiv für Rechts- und Sozialphilosophie, ARSP, vol. 100. Franz Steiner Verlag, Wiesbaden, 67–77 (in German).

Lakoff, G. and Johnson, M. 2003. *Metaphors We Live By*. University of Chicago Press, London.

Lu, Q. and Conrad, J. G. 2012. Bringing order to legal documents: an issue-based recommendation system via cluster association. In *Proc. of the Fourth International Conference on Knowledge Engineering and Ontology Development (KEOD 2012)*. SciTePress DL, 76–88.

Lu, Q. and Conrad, J. 2014. Next Generation Legal Search – It's Already Here. VoxPopuLII blog, Cornell Legal Information Institute, 28 March 2013, http://blog.law.cornell.edu/voxpop/2013/03/28/next-generation-legal-search-its-already-here/ (accessed 30 April 2015).

Mahler, T. 2010. *Legal Risk Management*. Ph.D. thesis, University of Oslo, 2010.

Mahler, T. 2013. A graphical user-interface for legal texts? In *Internationalisation of Law in the Digital Information Society: Nordic Yearbook of Law and Informatics 2010–2012*, Svantesson, D. J. B. and Greenstein, S., Eds. Ex Tuto Publishing, Copenhagen, 311–327.

Moody, D. 2009. The "physics" of notations: towards a scientific basis for constructing visual notations in software engineering. *IEEE Transactions of Software Engineering* 35(5), 756–778.

Röhl, K. F. and Ulbrich, S. 2007. *Law Graphically. Visualization in the Education of Jurists* (Recht anschaulich. Visualisierung in der Juristenausbildung). Halem, Köln (in German).

Schweighofer, E. 1999. *Legal Knowledge Representation. Automatic Text Analysis in Public International and European Law*. Kluwer Law International, Law and Electronic Commerce 7, The Hague.

Schweighofer, E. 2011. Indexing as an ontological-based support for legal reasoning. In *Technologies for Supporting Reasoning Communities and Collaborative Decision Making: Cooperative Approaches*, Yearwood, J. and Stranieri, A., Eds. IGI Global Publishers, Hershey, PA, 213–236.

Schweighofer, E. 2015. Legal data science – outline of a theory of legal informatics (Rechtsdatalystik – Versuch einer Teiltheorie der Rechtsinformatik). In *Proceedings of the 18th International Legal informatics Symposium IRIS 2015*, Schweighofer, E., Kummer, F., and Hötzendorfer, W., Eds., OCG, Vienna, 61–72 and *Jusletter IT*, February 2015, http://jusletter-it.weblaw.ch/issues/2015/IRIS.html.

Sergot, M. J., Sadri, F., Kowalski, R. A., Kriwaczek F., Hammond, P., and Cory, H. T. 1986. The British Nationality Act as a logic program. *Communications of the ACM* 29(5), 370–386.

Susskind, R. 2013. *Tomorrow's Lawyers: An Introduction to Your Future*. Oxford University Press, Oxford.

Vraalsen, F., Mahler, T. Lund, M. S., Hogganvik, I., den Braber, F., and Stølen, K. 2007. Assessing enterprise risk level: the CORAS approach. In *Advances in Enterprise Information Technology Security*, Khadraoui, D. and Herrmann, F., Eds. Information Science Reference, Hershey, New York, 311–333.

Walser Kessel, C. 2011. *Do You Know Law?* (Kennst du das Recht?) Editions Weblaw, Bern (in German).

Weilkiens, T. 2014. Systems Engineering mit SysML/UML. Dpunkt.verlag, Heidelberg.

Yoshida, Y., Honda, K., Sei, Y., Nakagawa, H., Tahara, Y., and Ohsuga, A. 2013. Towards semi-automatic identification of functional requirements in legal texts for public administration. In *Legal Knowledge and Information Systems, JURIX 2013: The Twenty-Sixth Annual Conference*, Ashley, K. D., Ed. IOS Press, Amsterdam, 175–184.

ELECTRONIC DISPUTE RESOLUTION WITHIN THE FRAMEWORK OF ELECTRONIC PROCESS - EP - IN BRAZIL

Cesar Antonio Serbena[1], Mauricio Dalri Timm do Valle[2]

[1]Professor of Philosophy of Law, Faculty of Law, Federal University of Paraná-Brazil
Praça Santos Andrade, 50, 80020-300 Curitiba, BR
cserbena@gmail.com; http://www.ejustica.ufpr.br
[2]Professor of Tax Law of the University Center Curitiba - Brazil-UNICURITIBA,
PhD in Law, Federal University of Paraná-Brazil
Praça Santos Andrade, 50, 80020-300 Curitiba, BR
mauricio_do_valle@hotmail.com; http://www.ejustica.ufpr.br

Keywords: ***Small Claims, Online Dispute Resolution, Alternative Dispute Resolution***

Abstract: *The present paper seeks to demonstrate that it is fully possible to replace the preliminary hearing, whose main purpose is to attempt conciliation in the Electronic Process - Pje - in which the parties could seek reconciliation by a platform and thereby avoid unnecessary realization of the hearing provided for in article 331 of the Civil Procedure Code, saving, therefore, time and money. The final part of the chapeau of art. 331 makes clear that the main scope of the preliminary hearing is to attempt conciliation, which provides that the parties will "... summoned to appear, and may be represented by an attorney or representative with authority to settle." It will prove, furthermore, that, despite a strong policy of encouragement by the National Council of Justice - CNJ - so that the parties undertake agreements, the so-called "Reconciliation Week", the number of reconciliations effectively achieved is relatively unsatisfactory. In the cases of "Reconciliation Week", performing a large number of hearings does not secure proportionality with the agreements effectively obtained. One imagines that the same occurs with the conciliation hearings provided for in article 331 of the CPC. Using these data, the second part of this research is to conduct a survey with members of the Judiciary, looking for the designated number of hearings, the number of hearings held, the reason for missing the agreement prior to or previously expressed disinterest in the agreement, the number of agreements reached in the hearing, the time of each hearing, and officials involved in the designation of the summons of the parties and the hearing activities. This is done for one simple reason: seeking to obtain data supporting the deployment of the electronic method of conflict resolution, in order to substitute the conciliation hearing. Law n. 11.419, of December 19, 2006, focuses on the computerization of the judicial discipline process. For this reason, it is known as the "Law of Electronic Process". This Law regulates the completion of legal proceedings by electronic means. Electronic means being any form of storage or digital documents and files (art. 1, § 2, I). Procedural Law establishes the possibility of sending petitions and performing acts procedures by electronic means through the use of an electronic signature - which can be based on a digital certificate or registration by the user in the Judiciary - where there are, in both cases electronic signatures, the prior registration in the Judiciary, which should be done in person (article 2). This permits, therefore, that all stakeholders in the process - parties, magistrates, clerks etc. - receive access to the system and thereby can practice procedural acts. There is,*

however, an issue that deserves careful attention. Article 8 of the Act provides that the various organs of the judiciary "... may develop electronic systems for processing lawsuits through wholly or partly digital file, using, preferably, a worldwide network of computers and access through internal and external networks". Finally, we seek to draw attention to the need to develop a survey with the staff of the Judiciary in order to collect data on the cost of these hearings in which the agreement is not obtained.

Título: Resolução Eletrônica de Conflitos no Âmbito do Processo Eletrônico – PJe – no Brasil

Palavras-chave: Pequenos litígios. Resolução de conflitos em linha. Resolução alternativa de litígios.

Resumo: *O presente trabalho visa a demonstrar que é totalmente possível substituir a audiência preliminar, cujo objetivo principal é buscar a conciliação, por uma plataforma no Processo Eletrônico - Pje - em que as partes poderiam buscar a conciliação e, assim, evitar a realização desnecessária da audiência prevista no artigo 331 do Código de Processo Civil, economizando, portanto, tempo e dinheiro. A parte final do caput do art. 331 deixa claro que o escopo principal da audiência preliminar é a tentativa de conciliação, que prevê que as partes "...serão intimadas a comparecer, podendo fazer-se representar por procurador ou preposto, com poderes para transigir". O artigo pretende provar, além disso, que, apesar de uma forte política de incentivo pelo Conselho Nacional de Justiça - CNJ - de modo que as partes possam realizar acordos, a chamada "Semana da Reconciliação", o número de conciliações efetivamente alcançadas é relativamente insatisfatório. Nos casos de "Semana da Reconciliação", a realização de um grande número de audiências não assegura, proporcionalmente, a obtenção efetiva de acordos. Imagina-se que o mesmo ocorra com as audiências de conciliação previstas no artigo 331 do CPC. Com estes dados, a segunda parte da pesquisa é realizar um levantamento com membros do Poder Judiciário, em que se solicitarão dados relativos ao número de audiências designadas, o número de audiências realizadas, o motivo da sua não realização (acordo entre as partes antes ou em que anteriormente expressaram desinteresse no acordo), o número de acordos alcançados na audiência, o tempo de cada audiência, os funcionários envolvidos na sua designação, intimação das partes e as atividades realizadas na audiência. E isso por uma razão simples: procura-se obter dados que apoiem a implantação do método eletrônico de resolução de conflitos, a fim de substituir a audiência de conciliação. A Lei n. 11.419, de 19 de dezembro de 2006, disciplina a informatização do processo judicial. Essa a razão pela qual é conhecida como a "Lei do Processo Eletrônico". Esta Lei regula a tramitação de processos judiciais por meios electrónicos. Por meio eletrônico entende-se qualquer forma de armazenamento de documentos e arquivos digitais (art. 1, § 2, I). A Lei estabelece a possibilidade de envio de petições e a realização de atos ou procedimentos por meio eletrônico por meio da utilização de uma assinatura electrónica - baseada em um certificado digital ou em um registro por parte do utilizador no Judiciário – onde há, em ambos os casos, o registo prévio das assinaturas no Poder Judiciário, o que deve ser feito pessoalmente (artigo 2). Isso permite, portanto, que todas as partes interessadas no processo - partes, magistrados, funcionários etc. - recebam o acesso ao sistema e, assim, possam praticar atos processuais. Há, no entanto, uma questão que merece atenção especial. O artigo 8 da lei prevê que os vários órgãos do Poder Judiciário "...poderão*

desenvolver sistemas eletrônicos de processamento de ações judiciais por meio de autos total ou parcialmente digitais, utilizando, preferencialmente, a rede mundial de computadores e acesso por meio de redes internas e externas". Finalmente, procuramos chamar a atenção para a necessidade de desenvolver uma pesquisa com os funcionários do Poder Judiciário, a fim de recolher dados sobre o custo dessas audições em que o acordo não seja obtido.

Título: Resolución Eletrónica de conflictos en el ámbito de Proceso Electrónico - PJe - en Brasil

Palabras-clave: Pequeños litigios. La solución de controversias en línea. Resolución alternativa de conflictos.

Résumen: *Este trabajo tiene como objetivo demostrar que es totalmente posible reemplazar la audiencia preliminar, cuyo principal objetivo es buscar la reconciliación, por una plataforma en el Proceso Electrónico - Pje - en la cual las partes podrían buscar la reconciliación y evitar, así, la realización de la innecesaria audiencia prevista en el artículo 331 del Código de Procedimiento Civil, ahorrando así tiempo y dinero. La parte final del caput del artículo 331 deja claro que el principal propósito de la audiencia preliminar es la conciliación, que dispone que las partes "... serán convocados a comparecer y pueden ser representados por un abogado, con el poder de transigir" Este trabajo muestra, además, que, a pesar de un fuerte incentivo por parte del Consejo Nacional de Justicia - CNJ - de modo que las partes puedan llegar a acuerdos, la llamada "Semana de la Reconciliación", el número de conciliaciones logradas con eficacia es relativamente insatisfactorio. En los casos de la "Semana de la Reconciliación", la realización de un gran número de audiencias no asegura proporcionalmente el alcance de acuerdos efectivos. Se cree que lo mismo ocurre con las audiencias de conciliación previstas en el artículo 331 del CPC. Con estos datos, la segunda parte de la investigación es realizar un estudio con los miembros del Poder Judicial, en lo cual se solicitarán los datos sobre el número de audiencias, el número de audiencias realizadas, la razón para su no realización (acuerdo entre las partes antes o donde anteriormente expresado anteriormente desinterés en el acuerdo), el número de acuerdos alcanzados en la audiencia, el tiempo de cada audiencia, los funcionarios involucrados en la cita, convoca a las partes y las actividades llevadas a cabo durante la audiencia. Y esto por una razón simple: tratar de obtener datos para apoyar la aplicación del método electrónico de resolución de conflictos con el fin de sustituir a la audiencia de conciliación. La ley n. 11.419, de 19 de diciembre de 2006, regula la informatización del proceso judicial. Es por eso que se conoce como la "Ley del Proceso electrónico ". Esta Ley regula la conclusión de procedimientos judiciales por medios electrónicos. Electrónicamente se entiende cualquier forma de almacenamiento de documentos y archivos digitales (art. 1, § 2, I). La Ley establece la posibilidad de enviar las peticiones y la realización de actos o procedimientos por vía electrónica a través del uso de una firma electrónica - sobre la base de un certificado digital o un registro por parte del usuario en el poder judicial - donde hay, en ambos casos, el registro previo de las firmas en el Poder Judicial, que debe hacerse en persona (artículo 2). Esto permite, por lo que todos los interesados en el proceso - partes, jueces, empleados, etc. - recibir el acceso al sistema y por lo tanto para la práctica de actos procesales. Hay, sin embargo, un tema que merece especial atención. El artículo 8 de la Ley establece que los diversos órganos del poder judicial "... pueden*

desarrollar sistemas electrónicos de procesamiento de demandas a través de autos total o parcialmente digitales, utilizando preferentemente la World Wide Web a través de redes de acceso y interno y externo". Por último, se busca llamar la atención sobre la necesidad de desarrollar una investigación con los empleados del poder judicial con el fin de recopilar datos sobre el costo de esas audiencias en las que no se obtenga el acuerdo.

1. Introduction

There is no doubt that in recent years the volume of cases pending before the judiciary in Brazil, considering the regular courts, the Federal Court and the Labor Court, have increased and have not seen processing. The increased volume results in an increase of process time and process costs. It seems to us that the use of electronic tools such as the development of the EO platform, enabling parties to try to celebrate through reconciliation, without the need for a hearing, would be a breakthrough in the quest for speed and procedural economy. These are goals, that even of the National Council of Justice, which determined the "Weeks of Reconciliation", are aimed to relieve the Judiciary policy. There are also concrete data on how much time and money is spent on conducting preliminary hearings, survey consisting of the second part of this research.

2. The Brazilian Code of Civil Procedure and the Conciliation

The Code of Civil Procedure - CCP - contains prescriptive statements that lead to the conclusion that the reconciliation should be sought by the judge. The art. 331 states that in not occurring the chances of dismissal - provided in arts. 267 and 269, II-V, the CPC (Article 329.) - And early trial of the suit, in which the judge will know directly from the application, uttering the sentence - where the issue of merit is only right or, also being in fact, there is no need to produce evidence in court, and also when the default (article 330, I and II.) occur - and, finally, whether the cause be about rights which allow the transaction, judge must designate the called preliminary hearing. The final part of the chapeau of art. 331 makes clear that the main scope of the preliminary hearing is to attempt conciliation, which provides that the parties will "... summoned to appear, and may be represented by an attorney or representative with authority to settle." There are few who recognize the importance of reconciliation. Some, like Luiz Guilherme Marinoni and Sergio Cruz Arenhart emphasize that reconciliation is an opportunity to eliminate the conflict more quickly and more cost effectively, and enable the "... restoration of harmonious relations between the parties", since it enables the elimination of conflict both in the sociological level as the legal.[1] If obtained in conciliation hearing, it will be reduced to term and approved by sentence (article. 331, § 1, CPC). However, if conciliation is not reached, the judge should fix the controversial points - which have a close connection with the distribution, in this case, the burden of proof - decide the outstanding procedural issues and, ultimately, determine the evidence be produced, if necessary, designating hearing and trial. The paragraph 3 of article. 331 provides that "the right not to admit the dispute transaction, or if the circumstances of the case be unlikely evidencing their achievement ...", the judge may sanitize the outset the process and order the production of evidence, pursuant to paragraph 2 the same article. Encouraging reconciliation, it seems evident.

[1] Processo de conhecimento. 11. ed. rev. e atual. RT. São Paulo. p. 243-244. (2013)

3. The National Council of Justice - CNJ - and their Functions

Article 2 of the Brazilian Constitutional Amendment. 45, of December 30, 2004 included, into the Federal Constitution, article 103-B, which establishes the National Council of Justice - CNJ - consisting of fifteen (15) members, each for a term of two (2) years, being allowed one (1) reappointment. The primary responsibility of CNJ is the control of the administrative and financial operations of the judiciary as well as the compliance of the judges, and their functional duties. Those that are interested in our study, however, are those set out in sections VI and VII of § 4 of art. 103-B of the Constitution, namely, to "prepare semiannual statistical reports on processes and sentences handed down, by the state, and the different organs of the judiciary" and to "prepare an annual report, including the measures it deems necessary, in the case of the judiciary in the country and the activities of the Council, which must integrate the message of the President of the Supreme Court which must be sent to Congress, at the opening of the legislative session.

Given these two items, and also considering the "urgent need of obtaining statistical data for compliance with these constitutional powers," the CNJ at the time chaired by Minister Nelson Jobim, created through Resolution n. 4 of 16 August 2005, the Statistical System of the Judiciary - SIESPJ. Article 1 of this Resolution establishes that the SIESPJ focuses and analyzes the data reported by all judicial organs of the country, whose referral is required. True that such data, which is obtained after, must be analyzed. As a result, the CNJ at the time chaired by then Minister Ellen Gracie, issued Resolution n. 49, of December 18, 2007, whose object is to discipline the organization of the Statistical and Strategic Management Center in the Judiciary according to article 92, II, III, IV, V, VI and VII of the Constitution. According to Article 1 of this Resolution, an administrative unit responsible for the compilation of statistics and strategic management plans of the respective court should be created. This core - that will be permanent and will assist the court in streamlining the institutional modernization process - should consist of public servants with degrees in law, economics, management, information science, and atleast one of them must necessarily be trained in statistics.

The data obtained from these cores should be sent to the CNJ. Inside the CNJ, the Statistical and Strategic Management Commission oversees the SIESPJ, this commission is responsible, with the advice of the Department of Judicial Research, "to aggregate statistical data sent by core statistical and strategic management of the courts." Currently, in view of the need for critical analysis of the statistical data of each organ of the judiciary and also the importance of such data as input for the decision on public policy to be adopted in the Judiciary, the SIESPJ was regulated by Resolution n. 76 of May 12, 2009, the CNJ at the time chaired by the Minister Gilmar Mendes, whose scope was on the principles, establishes indicators, deadlines and furthermore determine penalties. According to article 1 of this Resolution, the SIESPJ will be coordinated by the CNJ and will consist of the Superior Court of Justice, the Federal Regional Courts, the Labor Courts, the Electoral Courts, the Military Courts and the Courts of the States and the Federal District and Territories. The SIESPJ, in accordance with the provisions of article 2, will be "... governed by the principles of publicity, efficiency, transparency, mandatory information and statistical data received under the presumption of veracity of the statistical data reported by the Courts and the constant updating of indicators as improving the management of the courts."

All statistics, which should necessarily be transmitted to the CNJ and the Presidency of the Court, will be by means of electronic transmission by the online system, be available on the website https://estatistica.cnj.jus.br. Such data are received by the Department of Judicial Research, which operates under the supervision of the Statistics Committee and Strategic Management (art. 7), which presents them to the National Council of Justice in a reported form, which will contain "statistical data on processes and sentences handed down by state courts or in different organs of the judiciary, each semester, pursuant to art. 103-B, § 4, VI, "the annual consolidation, including the

statistical data collected in the previous year" and "the consolidated historical series, covering no more than the previous ten years, where available." (article 8).

4. The Law n. 11.419, of December 19, 2006 and the Resolution of the CNJ n. 185 of December 18, 2013

Law n. 11.419, of December 19, 2006, focuses on the computerization of the judicial discipline process. For this reason, it is known as the "Law of Electronic Process". This Law regulates the completion of legal proceedings by electronic means. Electronic means being any form of storage or digital documents and files (art. 1, § 2, I). Traffic Law establishes the possibility of sending petitions and performing acts procedures by electronic means through the use of an electronic signature - which can be based on a digital certificate or registration by the user in the Judiciary - where there are, in both cases electronic signatures, the prior registration in the Judiciary, which should be done in person (article 2). This permits, therefore, that all stakeholders in the process - parties, magistrates, clerks etc. - receive access to the system and thereby can practice procedural acts. There is, however, an issue that deserves careful attention. Article 8 of the Act provides that the various organs of the judiciary "... may develop electronic systems for processing lawsuits through wholly or partly digital file, using, preferably, a worldwide network of computers and access through internal and external networks". And, in the "General and Final Provisions", the Law establishes (article 14) that such systems "... should preferably open source programs, continuously accessible through the World Wide Web, prioritizing its standardization." This is due to concern about "interoperability" between systems, mainly to avoid situations like the one provided for in article 12, § 2 of the Law, according to which, if the electronic process is to be remitted to another court or authority which does not have a compatible system, the records must be printed on paper and presented as a physical process. For example, the process that was electronic returns to be physical, due to the incompatibility of systems. Something that, in our view, is true nonsense, to the extent that this is far from the main scope of the electronic process, namely, procedural efficiency and speed of its passage.

Currently, the CNJ, through Resolution n. 185 of December 18, 2013, determined the unification of the information processing system, which must occur by 2018 (Article 34, § 3) called Judicial System Electronic Case - Pje. The resolution previously points out, a total of 13 "recitals" or justifications, among them, the most important for the object of our study are: i) "the benefits arising from the replacement of the conduct of proceedings in the physical medium by electronic means, as an instrument of speed and quality of adjudication "; ii) "the need to rationalize the use of budgetary resources by the Judiciary"; and iii) "the need to regulate the implementation of the Electronic Judicial Process system - EO on the Judiciary, to give it uniformity." The text of the Resolution is to establish clarity (Article 44). From its validity, it is expressed that it is "prohibited ... the creation, development, procurement and deployment of diverse electronic systems or prosecution of the EO modules ...", except in cases of "corrective maintenance and upgrades necessary to run the systems" and also in the case prescribed in the enigmatic article 45, according to which the Plenary of the CNJ can relativize the seal, providing the following conditions: i) any request by the Court; and ii) when "... understood and justified by circumstances or specific locals ...".[2] It is clear, therefore, that the problem of interoperability, at least normatively, is solved.

[2] http://www.cnj.jus.br/atos-administrativos/atos-da-presidencia/resolucoespresidencia/27241-resolucao-n-185-de-18-de-dezembro-de-2013.

5. The results obtained with the "Reconciliation Week"

Following the posture adopted by the Code of Civil Procedure - to encourage reconciliation - CNJ issued Resolution n. 125, of November 29, 2010, which provides for a National Policy of adequate treatment of conflicts of interest within the Judiciary. The resolution previously points out, a total of 10 "recitals" or justifications, among which those that are of interest for this study are: i) that falls to the CNJ to control the administrative and financial operations of the Judiciary; ii) that it is up to the judiciary to establish public policy for proper treatment of legal problems and conflicts of interest that occur in large and increasing scale in society, in order to organize at the national level, not only the services provided in the court processes, as well as those which may be so through other mechanisms of conflict resolution, in particular consensual, such as mediation and conciliation; iii) that it is necessary to consolidate permanent public policy on the encouragement and improvement of consensual dispute resolution mechanisms; iv) that conciliation and mediation are effective instruments of social peace, solution and prevention of disputes, and that the appropriate discipline in programs already implemented in the country has reduced excessive judicialization of conflicts of interest, the amount of resources and execution of sentences; v) that it is relevant and necessary to the organization and standardization of conciliation, mediation and other consensual methods of dispute resolution, services to prevent them from disparities on guidance and practices as well as to ensure the proper implementation of public policy, respecting the specificities of each segment of Justice; and finally, vi) that the organization of conciliation, mediation and other consensual methods of dispute resolution services should be the principle and basis for the creation of Courts of alternative dispute resolution, real specialized courts in the matter.

The text resolution is clear in stating that the judiciary should provide mechanisms of dispute resolution, especially consensual, such as mediation and conciliation, and provide care and guidance to citizens, as part of this policy (article 1, § 1), whose implementation will observe the specific statistical monitoring (article 2, III). Note that the resolution itself states that the Courts should create and maintain databases on the activities of its Center for Conciliation, on which we will discuss later (article 13) and also that the CNJ compiles "information about the public services consensual resolution of controversies in the country and on the performance of each of them, through the DPJ, constantly keeping updated databases."

The "Centers of Judicial Conflict and Citizenship" mentioned – which focus on the conduct of conciliation and mediation sessions, and whose discipline is through article 8 of the Resolution - should be installed by the "Permanent Cores on Consensual Dispute Resolution Methods", whose creation, by the Courts, is mandatory (article 7, IV). These "centers" must necessarily rely on at least three sectors: i) pre-trial settlement of disputes; ii) procedural conflict resolution; and finally iii) citizenship (article 10).

Remember that article 4 of Resolution prescribes that the CNJ has the responsibility of "organizing programs aiming to promote action to encourage self-composition disputes and social peace through conciliation and mediation," such as "Reconciliation Week", whose results should be available on the Portal of the Conciliation on the website of the CNJ (article 15, VI). And precisely the analysis of statistical data[3] - 2006[4], 2007[5], 2008[6], 2009[7], 2010[8], 2011[9] and 2012[10] - we spent this time taking

[3] http://www.cnj.jus.br/programas-de-a-a-z/acesso-a-justica/conciliacao/semana-nacional-de-conciliacao.
[4] http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2006-semana_conciliacao_2006.pdf
[5] http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2007-semana_conciliacao_2007.pdf
[6] http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2008-semana_conciliacao_2008.pdf
[7] http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2009-relatrio_semana_pela_conciliao_07a11_dez09.pdf
[8] http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2010-dadosestatisticos.pdf
[9] http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2011/Semana_Conciliacao_20-01-2012.pdf

into account the activities from December 8th, 2006, from the 3rd to December 8th, 2007, from the 1st to the 5th of December 2008, from the 7th to December 11th, 2009, from November 29th to December the 3rd, 2010, from November 28th to December 2nd, 2011 and the 7th through November 14th, 2012.

Table 1 Comparison between the percentage of hearings and the percentage of agreements reached, considering the overall activity of Reconciliation Week, ex. the activity of all participants Courts.

Year	nº of days	Participating Courts	Desginated Hearings	% Hearings Realized	Agreements Reached	% Reached
2006	1	55	112.112	74,913%	46.493	55,29%
2007	6	53	303.638	74,946%	96.492	42,40%
2008	5	56	398.012	76,779%	130.848	42,80%
2009	6	56	333.324	78,127%	122.943	47,20%
2010	5	53	439.180	82,414%	171.637	47,40%
2011	5	54	434.479	80,467%	168.841	48,30%
2012	8	49	419.031	83,979%	175.173	49,78%

Analyzing the table, one can seen that between 2006 and 2012 there was a considerable increase in the number of designated audiences, however, the percentage of hearings actually performed did not exceed 84%. Moreover: of the hearings actually performed, only about half of them were able to reach an agreement.

Table 2 Comparison between the percentage of hearings and the percentage of agreements reached at the Ordinary Courts

Year	Desginated Hearings	Hearings Realized	% Hearings Realized	Agreements Reached	% Reached
2006	82.523	58.981	71,472%	31.223	52,96%
2007	199.347	137.426	68,938%	59.736	43,47%
2008	253.634	178.830	70,507%	74.215	41,50%
2009	221.120	165.159	74,692%	79.458	48%
2010	316.113	252.405	79,846%	122.683	48,60%
2011	303.625	241.172	79,431%	119.840	49,69%
2012	336.123	295.175	87,818%	155.717	52,75%

Analyzing the table of Reconciliation Week activities of the Common Justice, one can noticed that the percentage of hearings actually performed did not exceed 88%. Moreover: of the audiences actually performed, just about half of them, as well as when considering all the Courts, obtained an agreement.

Table 3 Comparison between the percentage of hearings and the percentage of agreements reached in Federal Court.

Year	Designated Hearings	Hearings Realized	% Hearings Realized	Agreements Reached	% Reached
2006	16.917	13.893	82,124%	9.198	66,21%
2007	20.217	17.428	86,205%	10.725	61,54%
2008	28.652	25.661	89,561%	16.446	64,10%
2009	15.325	13.464	87,856%	7.739	57%
2010	31.956	25.980	81,299%	14.991	57,70%
2011	28.937	23.619	81,622%	16.385	69,37%

[10] http://www.cnj.jus.br/images/programas/conciliacao/2012/relat%C3%B3rio_final_Conciliacao2012.pdf

2012	11.446	7.624	66,608%	5.886	77,20%

Analyzing the table of Reconciliation Week activities of the Federal Court, it can be noticed that the percentage of hearings conducted effectively came very close to 90%. Moreover: of the hearings actually performed approximately 77% of them reached agreements, a considerable percentage of effectiveness.

Table 4 Comparison between the percentage of hearings and the percentage of agreements reached within the Labour Court.

Year	Designated Hearings	Hearings Realized	% Hearings Realized	Agreements Reached	% Reached
2006	12.292	11.113	90,408%	6.072	53,98%
2007	58.727	54.654	93,065%	21.883	39,97%
2008	115.726	101.100	87,362%	40.187	39,70%
2009	221.120	81.793	36,990%	35.746	44%
2010	91.111	83.560	91,712%	122.683	48,60%
2011	101.917	84.822	83,227%	332.616	38,45%
2012	71.462	49.099	68,706%	13.570	27,64%

Analyzing the table of Reconciliation Week activities of the Labour Court, it can be noticed that the percentage of hearings conducted effectively reached a significant level in 2007, reaching 93%, suffering a significant drop in 2009, nearing 37%, partially recoving in 2012, when it reached close to 69%. With regards to the arrangements effectively achieved, the percentage is of concern, inasmuch as in 2012, for example, audiences actually carried out only about 28% of those reaching agreements.

As seen, even in cases where the parties are summoned to attend a specific purpose of trying to resolve their conflicts of interest, the so-called "Reconciliation Week", show the results to be satisfactory. Note that in those years, a total of 2,439,776 audiences, only 1,941,014 were made, for example only 79.56%.

6. Conclusion: Possibility of Electronic Conflict Resolution through Creation Platform in EP

As mentioned above, the Code of Civil Procedure (article 331, § 3) allows the judge not to perform [dispense] the preliminary hearing, which is becoming common in the scope of conciliation in cases where "the right at issue does not admit transaction "or even" if the circumstances of the case are unlikely evidencing their achievement."

As seen in the cases of "Reconciliation Week", performing a large number of hearings does not secure proportionality with the agreements effectively obtained. One imagines that the same occurs with the conciliation hearings provided for in article 331 of the CPC. However, according to these data, the second part of this research is to conduct a survey with members of the Judiciary, in which prompts the designated number of hearings, the number of hearings held, the reason for missing the (agreement prior to or previously expressed disinterest in the agreement), the number of agreements reached in the hearing, the time of each hearing, officials involved in the designation, summons of the parties and the hearing activities. And this for one simple reason: seeking to obtain data supporting the deployment of the electronic method of conflict resolution, in order to substitute the conciliation hearing.

Remember that, within short time , the problem of interoperability does not exist – and has been resolved, as seen normatively - and even the judiciary, by express provision of § 3 of article 10 of Law n. 11 419/2006, shall maintain the structure, such as scanning equipment and computers with

access to the worldwide web, available to interested parties for the distribution of pleadings and also article 41 of Resolution no. 185/2013 determines that since the implementation of the EO, the Courts shall maintain "structures of care and support to users," it is fully possible to create a platform within the EO, for the parties to seek reconciliation, rather than being necessary to conduct the conciliation hearing.

Nothing prevents the application of the Online Dispute Resolution (online dispute resolution) - ODR[11] - by downloading it from a physical environment to an "electronic location". In an excellent article in the Portuguese and Spanish doctrines, that explains the use of ODR, offers the parties a tool for dealing with conflicts in an easy way, providing opportunities, perhaps, saving time and money.[12] This, at first, with one person holding the position of conciliator one must not forget, however, that there is a strong tendency in that conciliators are replaced by software. This is the second generation of ODR, which has, according to the authors, three essential characteristics: i) the scope of present solutions to the dispute; ii) reduce human intervention, increasing the software; and finally iii) include the use of software agents. These ODR processes are known by the parties as the Best Alternative to a Negotiated Agreement - BATNA and Worst Alternative to a Negotiated Agreement - WATNA. Within these levels - for a better or worse alternative - is the Zone of Possible Agreement - ZOPA. This knowledge enables the parties to assess the proposed agreement more rationally.

New technologies and methods available for conciliation and resolving Online conflict has been underutilized in Brazil. The Brazilian judicial system is on the verge of exceeding 100 million cases, according to the latest reports already disclosed in “Justice in Numbers”. Thus, the use of new technologies is not only for the Electronic Process systems design, which is already well developed in Brazil, but also urgent and necessary in systems or electronic platforms for reconciliation.

The Brazilian judicial system already has a reasonable degree of maturity in the legal regulation of new ICTs. The "Civil Mark of the Internet", as it is known in Brazil, is the most recent and important legislation on ICT and on the web environment of the Brazilian legal system. Some have called it the "Bill of Rights" of the Brazilian Internet (Federal Law No. 12,965, from April 23rd, 2014). We must now move to a second step in the opposite direction, which is not just regular ICTs, but the need to employ ICTs, software, and new technologies, to improve and enhance the practice of professional and legal practitioners.

A wide field of new technologies is emerging and it is necessary to open a legal practice to do this job, by using: mobile technologies, cloud computing, social networks, the development of social Web 2.0 and the Semantic Web 3.0, etc. Certain fields of legal information are being developed rapidly and can find wide use in mediation practices, such as the recovery of legal information systems and extraction systems and visualization of information expressed or implied, contained in legal texts as positive norms and agreements.

---

[11] About ODR, here are the words of Francisco Andrade Carneiro Pacheco, Davi de Carneiro and Paulo Novaes: "This new model will significantly broaden the scope of intervention of alternative systems of dispute resolution and, through the introduction of more sophisticated mechanisms such as expert systems ("expert systems"), enhance the generation of answers and possible solutions to the needs and aspirations of the parties. Online conflict resolution is made possible through the use of the most common technological means, such as electronic messages or conversations ("Instant Messaging"), e-mail, video conferencing, electronic forums, mailing lists etc. These are some of the technologies that will allow the parties easier and faster communication in synchronous or asynchronous mode (3), even if they are not, or are unable to be in each other's presence. "- F. Andrade, Ram D., Novais P., Artificial Intelligence in Online Dispute Resolution, Scientia Ivridica - Volume LIX, No. 321, (2010), p. 2.

[12] *Barbieri D., Carneiro D., Andrade F., Novais P.*, Resolução de Conflitos em Linha - Uma aplicação ao direito do consumo. Scientia Ivridica, Tomo LIX – nº 323, p.292.

Brazilian courts need to be concerned not only in developing procedural systems of e-Justice, but also in developing and implementing electronic systems of mediation, online dispute resolution and crowdsourcing[13]. In regards to crowdsourcing, our courts should implement cooperation and use collective production strategies of knowledge and information, by adding this information to remove common properties, which are invisible in massive sets of information. Specific techniques of Big Data and Data Mining in this sense are essential, but so far are absent from debates about the procedural congestion of the Brazilian legal system.

Regarding the use of ICTs to reconcile environment, the two main strategies are the ADR (Alternative Dispute Resolution) and ODR (Online Dispute Resolution).

The development of ADR had a strong role in the field of Consumer Law, and was mainly due to the need for an alternative system of justice, with simple, fast and inexpensive procedures, accessible to millions of consumers. To the extent that the implementation of ADR has become mandatory in Consumer Law, mainly within the European Community, where the national justice systems had to deploy the ODR at the same time.

The term Online Dispute Resolution - ODR, refers to the application of ICTs to the resolution of disputes and litigation between conflicting parties. There are several systems of ODR, which have a variety of technological complexities. According Immaculada Barral Viñals[14], the term ODR was coined originally in 2001 by Ethan Katsh and Janet Rifkin.

The positives of ODR operate on two levels: first, the ODR allows for the modeling of several negotiation processes, such as: arbitration, conciliation, etc., which also allows to identify your flowchart and point out to what extent the neutral third party, or the negotiator, must participate in the process. At a second level, the ODR also allows for fully automated decisions with the use of artificial intelligence techniques, although there is still strong resistance regarding their usage[15].

It seems that there are no other more effective solutions than the use of ODR in the case of small claims. A massive number of cases, the low monetary value involved in the cases, and the close similarity between them are strong reasons for its adoption[16]. The e-Justice in Brazil has been well

[13] About Mediation Theory and the actual state of research about the use of ICT in Judicial Mediation, cf. CASANOVAS, 2014.

[14] Online Dispute Resolution and Small Claims, 2014.

[15] "In this sense, the virtuality of ODRs is the intense interaction between technical tools and how to resolve disputes especially when they offer means that in the offline environment simply do not exist (POBLET et al., 2009).
This versatility is largely a result of the technologies of Web 2.0 and Web 3.0 which offer real opportunities for development to the three major areas of the ODR: communication, collaboration and interactivity. Translated into legal terms, and as indicated by Benyekhlef and Gélinas (2005), the technology currently provides mechanisms using a technological infrastructure to automate some functions, modeling the dispute resolution process and provide an interface that allows for compliance, document and archive stages in the resolution process. It is in this definition which is the creative potential of ODRs as they are able to automate processes that should connect the neutral or third party and should also model the whole process with tools to help in the resolution of the conflict.
First, at the first level, we should emphasize the versatility of electronic media to offer so-called multidoor resolution processes, which are offered in different stages of ADR processes, which is one of their highest virtues as they are able to offer different procedures in which the parties have successive resolution mechanisms at their disposal: negotiation, mediation and arbitration, for example, assigning a neutral third party.
At the second level, there are fully automated electronic procedures that the field of consumer complaints are well known, such as auto-negotiation mechanisms and charge back. Automatic trading systems are carried out without human intervention and are a major innovation of the ODR. Its scope is to monetary claims where there is no dispute over the amount of compensation that the consumer should receive. (...) As you can imagine, this system is extremely flexible and very inexpensive, although the absence of human intervention can damage the goodness of certain decisions. It assumed the epitome of connection between new technologies and the conflict resolution network." Viñals, p. 400-401.

[16] "The commitment to these interactive platforms in consumer disputes rests on the idea that the ODR seem particularly effective in an environment of claims as described: similarity of conflicts and complaints of small

developed in relation to the computerization of the common Civil Procedure over the last ten years. Now it is necessary that the same technological investment be made in platforms for resolving small claims, which will fulfil an important role in combating procedural congestion. It seems that there has to be a more efficient and effective form of utilizing ODRs in cases with small claims.

7. References

1. *Andrade F., Carneiro D., Novais P.*, A Inteligência Artificial na Resolução de Conflitos em Linha, Scientia Ivridica – Tomo LIX, n.º 321, ISSN: 0870-8185, pp 137-164 (2010).
2. *Barbieri D., Carneiro D., Andrade F., Novais P.*, Resolução de Conflitos em Linha - Uma aplicação ao direito do consumo, Scientia Ivridica, Tomo LIX – nº 323, pp 581-607 (2010).
3. http://www.cnj.jus.br/atos-administrativos/atos-da-presidencia/resolucoespresidencia/27241-resolucao-n-185-de-18-de-dezembro-de-2013.
4. http://www.cnj.jus.br/atos-administrativos/atos-da-presidencia/323-resolucoes/12243-resolucao-no-125-de-29-de-novembro-de-2010.
5. http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2006-semana_conciliacao_2006.pdf
6. http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2007-semana_conciliacao_2007.pdf
7. http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2008-semana_conciliacao_2008.pdf
8. http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2009-relatrio_semana_pela_conciliao_07a11_dez09.pdf
9. http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2010-dadosestatisticos.pdf
10. http://www.cnj.jus.br/images/programas/movimento-pela-conciliacao/2011/Semana_Conciliacao_20-01-2012.pdf
11. http://www.cnj.jus.br/images/programas/conciliacao/2012/relat%C3%B3rio_final_Conciliacao2012.pdf
12. http://www.cnj.jus.br/programas-de-a-a-z/acesso-a-justica/conciliacao/semana-nacional-de-conciliacao.
13. *Marinoni, L. G., Arenhart, S. C.* Processo de conhecimento. 11th. Ed. RT, São Paulo (2013).
14. *Barral Viñals, I.* Online Dispute Resolution and Small Claims. Democracia Digital e Governo Eletrônico, Florianópolis, n° 10, p. 394-415 (2014).
15. *Casanovas, Pompeu.* Ethics in Mediation and Ethics of Mediation: the redefinition of public space. 14. Democracia Digital e Governo Eletrônico, Florianópolis, n° 10, p. 416-432 (2014).

amount; and offer the advantages of greater speed and lower costs. Thus, in the analysis of the relationship between the amount and the suitability of the conflict resolution, the following stands:
• in claims with small amounts, the option for the ADR / ODR is not an alternative, but is the only one that arises in many cases, in terms of choice.
• that despite the prejudice to each consumer in particular have little value, the sum of the performance against all consumers can have great proportions, so the lack of means of claim leads to injustice." Viñals, p. 412-413.

QUANTIFYING THE INTEROPERABILITY OF SYSTEMS OF ELECTRONIC JUDICIAL PROCESSES IN THE BRAZILIAN FRAMEWORK

Cesar Antonio Serbena

Professor of Philosophy of Law, Law School, Federal University of Paraná-Brazil
Santos Andrade Square 50, 80020-300 Curitiba, BR
cserbena@gmail.com; http://www.ejustica.ufpr.br

Keywords: ***Interoperability, Electronic Judicial Processes, Measurement.***

Abstract*: Interoperability can be described in four levels: legal interoperability, organizational interoperability, semantic interoperability and technical interoperability. The legal level establishes the legislation and legal rules for data exchanging; the organizational level coordinates processes in which organizations achieve agreed and mutually beneficial goals; the semantic level states more precisely the meaning of exchanged information for both or different organizations; and the technical level plans technical issues involved in linking computer systems and services. In the Brazilian framework, the four levels are not equally developed. Legal, semantic and technical interoperability have well established rules, but the main problem is located in organizations. Not all Brazilian Courts agreed to cooperate and adopt the Brazilian National Model of Interoperability. Also the gap to achieve a total interoperability is the user's perspective or citizen's perspective. Public organizations can operate in these four levels, but these are not sufficient to offer good judicial services to citizens. Most commonly users have indirect benefits if public organizations can interoperate between them, but the ideal one-stop-shop delivery of public services is far from been implemented today. The main objective of this paper is to present a survey to quantify the interoperability of systems of the Electronic Judicial Process from the user's point of view. This paper is a part of a research project that will measure all the Brazilian systems of Electronic Judicial Process. This first stage of the research is to develop the metric to measure the systems. In 2015 the group will apply the survey and will collect the answers. The result of the research project will be a complete and national ranking of the Brazilian systems from the user's perspective.*

Título: Quantificando a Interoperabilidade dos Sistemas de Processos Judiciais Eletrônicos no âmbito Brasileiro

Palavras-chave: interoperabilidade, processos judiciais eletrônicos, Mensuração.

Resumo: *A interoperabilidade pode ser descrita em quatro níveis: interoperabilidade jurídica, interoperabilidade organizacional, interoperabilidade semântica e de interoperabilidade técnica. O nível legal estabelece a legislação e normas legais para a troca de dados; o nível organizacional coordena os processos nos quais as organizações alcançam metas acordadas e mutuamente benéficas; o nível semântico torna preciso o significado das informações trocadas por duas ou mais diferentes*

organizações; o nível técnico planeja questões técnicas envolvidas na conexão dos sistemas e serviços de informática. No âmbito brasileiro, os quatro níveis não são igualmente desenvolvidos. As interoperabilidades Legal, semântica e técnica possuem regras bem estabelecidas, mas o principal problema está localizado nas organizações. Nem todos os Tribunais brasileiros concordaram em cooperar e adotar o Modelo Nacional de Interoperabilidade. Também a lacuna para conseguir uma interoperabilidade total é a perspectiva do usuário ou perspectiva do cidadão. As organizações públicas podem operar nesses quatro níveis, mas estes não são suficientes para oferecer bons serviços judiciais para os cidadãos. Muito comumente os usuários têm benefícios indiretos se as organizações públicas podem interoperar entre elas, mas a entrega ideal dos serviços públicos está longe de ser sido plenamente implementada. O principal objetivo deste trabalho é apresentar um estudo para quantificar a interoperabilidade dos sistemas do Processo Judicial Eletrônico do ponto de vista do usuário. Este artigo é parte de um projeto de pesquisa que vai mensurar todos os sistemas brasileiros de Processo Judicial Eletrônico. A primeira etapa da pesquisa é desenvolver a métrica para os sistemas. Em 2015 o grupo irá aplicar a pesquisa e coletar as respostas. A conclusão do projeto de pesquisa será uma avaliação completa e nacional dos sistemas brasileiros a partir da perspectiva do usuário.

Título: La cuantificación de la interoperabilidad de los Sistemas de Procesos Judiciales Electrónicos en la realidad brasileña

Palabras clave: interoperabilidad, procesos judiciais electrónicos, medición.

Resumen: *La interoperabilidad se puede describir en cuatro niveles: interoperabilidad jurídica, interoperabilidad organizativa, semántica y técnica. El nivel legal establece las leyes y normas legales para el intercambio de datos; el nivel de organización coordina los procesos en los que las organizaciones logran metas mutuamente beneficiosas e convenidas; el nivel semántico hace preciso el significado de la información intercambiada por dos o más organizaciones diferentes; el nivel técnico son las cuestiones técnicas involucradas en la conexión de los sistemas y servicios informáticos. En el contexto brasileño, los cuatro niveles no están igualmente desarrollados. Las interoperabilidades legal, semántica y técnica tienen reglas bien establecidas, pero el principal problema se encuentra en las organizaciones. Ni todas las cortes brasileñas acordaron en cooperar y adoptar el Modelo Nacional de Interoperabilidad. También la brecha para lograr la plena interoperabilidad es la perspectiva del usuario o la perspectiva del ciudadano. Las organizaciones públicas pueden operar en estos cuatro niveles, pero éstas no son suficientes para proporcionar buenos servicios legales para los ciudadanos. Con demasiada frecuencia, los usuarios tienen beneficios indirectos si las organizaciones públicas pueden interoperar entre ellos, pero la administración óptima de los servicios públicos es lejos de ser aplicado en su totalidad. El objetivo de este trabajo es presentar un estudio para cuantificar la interoperabilidad de los sistemas de Proceso Judicial Electrónico del punto de vista del usuario. Este artículo es parte de un proyecto de investigación que medirá todos los sistemas brasileños de Proceso Judicial electrónicos. La primera etapa de la investigación es desarrollar métricas para los sistemas. En 2015 el grupo aplicará la investigación y vá recoger las respuestas. La conclusión del proyecto de investigación será una evaluación completa y nacional de sistemas de Brasil desde la perspectiva del usuario.*

1. Introduction

According to European Interoperability Framework – EIF v.2, Interoperability can be described in four levels: legal interoperability, organizational interoperability, semantic interoperability and technical interoperability [1]. The legal level establishes the legislation and legal rules for data exchanging; the organizational level coordinates processes in which organizations achieve agreed and mutually beneficial goals; the semantic level states more precisely the meaning of exchanged information for both or different organizations; and the technical level plans technical issues involved in linking computer systems and services[1]. A gap to achieve a total interoperability is the user's perspective or citizen's perspective. Public organizations can operate in these four levels, but these are not sufficient to offer good judicial services to citizens. Most commonly users have indirect benefits if public organizations can interoperate between them, but the ideal one-stop-shop delivery of public services is far from been implemented today. The main objective of this paper is to present a survey to quantify the interoperability of systems of the Electronic Judicial Process in Brazil from the user's point of view. This paper is a part of a research project that will measure all the Brazilian systems of Electronic Judicial Process. This first stage of the research is to develop the metric to measure the systems. In 2015 the group will apply the survey and will collect the answers. The conclusion of the research project will be a complete and national ranking of the Brazilian systems from the user's perspective[2].

2. 1. Timeline of legislations on E-Justice and Interoperability in Brazil

Articles 8 and 18 of the Brazilian Legal Framework on *Computerization of the Judicial Process* (Federal Statute No. 11.419/2006) states that "*The organs of the Judiciary* (i.e., Courts) *may develop electronic systems for processing lawsuits through total or partially digital files, preferably using the worldwide network of computers and access through internal and external networks"* and that they *"(...) shall regulate this Act, when applicable, within their respective competences"*.

Considering that the Brazilian Judiciary Branch is composed of 5 higher courts, 27 state and district appellate courts, 5 federal appellate courts and 24 labor appellate courts, it is clear that, according to the aforementioned articles it would be legally possible to have at least more than 50 different systems of *judicial electronic process*.

Nowadays several systems of the Brazilian Courts coexist. Among others, it is possible to mention, for example, the PJe, eProc, Projudi and E-SAJ systems, some of which have been in operation for more than half a decade, after several years of investment, development and improvement by Courts.

The computerization of the Brazilian Courts had its legal beginning with Federal Statutes No. 11.280 and mainly with No. 11.419, both from 2006. Since 2006, the Brazilian Courts began to develop their own systems of EJP – Electronic Judicial Process. The CNJ – National Council of Justice - formulated its own system called PJe ("Processo Judicial Eletrônico" in Portuguese). It was officially launched on June 21st, 2011, and it is proposed to be adopted by all Brazilian Courts. In many cases and even before the PJe, several other systems of EJP were developed by several Brazilian courts, which continue to be in use today.

1 More details about the European Framework on Interoperability can be found at [2].

2 The present article is a new development of a previous paper entitled *Electronic Judicial Process and Interoperability: Current State of Affairs in Brazil and Comparative Law* published [3], and it is a part of a Research Project financed by the Brazilian National Council of Justice – CNJ Acadêmico and Capes, a Brazilian Public Agency for Scientific Development.

These systems have a considerable size and power to process millions of judicial processes. For instance, the following table show some data about them:

E-SAJ - In use in the State Courts of Sao Paulo and 7 other States. (Data of Sao Paulo State Courts from Dec. 2014)
1,973,337 digital processes (active and stored)

PROJUDI - In use in the State Courts of Parana and 19 other States (Data of Parana State Courts from Dec. 2014)
Total registered cases: 3,856,418
Total distributed processes: 3,827,265
Total active processes: 1,521,535
Total distributed appells: 192,092
Total active lawyers: 61,958
Total parties with access to the system: 67,131
Projudi is implemented in all Parana State Districts.

E-PROC - In use in the Federal Courts of the 4th Region (Federal Justice Circuits for the States of Parana, Santa Catarina and Rio Grande do Sul) and State Courts of Tocantins - (Data from Dec. 2014)
3.031.102 digital processes

In the Brazilian framework of judicial computerization, the development of interoperable e-Justice systems was a constant concern. Along with the legal rules of judicial computerization, the resolutions below of the CNJ created a National policy implementation of ICT in Brazilian courts and implemented a National Model of Interoperability for the Brazilian Judiciary:

CNJ Resolution No. 12, 2006: created the Interoperability Group (G-INP) of Brazilian Judicial Branch.

CNJ Resolution No. 90, 2009 and Resolution No. 136, 2011: Establishes and regulates the ICT infrastructure and ICT management throughout the Brazilian Judiciary.

CNJ Resolution No. 91, 2009: Establishes the MoReq-Jus (Model Requirements for Computerized Systems of Document and Process Management of the Judiciary) and the obligation of its use in the development and maintenance of computerized systems for the judicial and administrative activities within the Judiciary Branch.

CNJ Resolution No. 99, 2009: Establishes the Strategic Planning of ICT in the framework of the Brazilian Judiciary.

CNJ Resolution No. 100, 2009: States the official communication by electronic means within the Judiciary Branch.

In the framework presented, it is clear that Brazil has established rules for legal, semantic and technical interoperability. The next challenge is to integrate the various courts so that the organizational interoperability would be achieved.

From the point of view of lawyers/users which operate the systems, there are several problems concerning interoperability: there are different systems of EJP working in several levels (State level, Federal Level and National Level) and in several Courts (Labor, Civil and Criminal, Federal, Military and Higher Courts). In summary, the main sources of problems related to interoperability in Brazil are:

1 - Brazil has dozens of systems of Electronic Judicial Process systems;

2 – Systems still are not interoperable between themselves and between Lower Courts, Appellate Courts and Superior Courts;

3 - Interoperability gaps between systems are mostly solved through solutions planned for paper based procedures;

4 - Lawyers must operate the same legal procedures in different systems

3. The Brazilian model of Interoperability: Moreq-Jus

Brazil already has a National Model of Interoperability for the Executive Branch, the E-PING (Interoperability Patterns of Electronic Government). The model of Interoperability of the Judicial Branch – MNI (National Model of Interoperability) – is under development by the CNJ. It should be noted, however, that since 2009 Brazilian Courts have followed the system of Documental Management called MoReq-Jus (Model Requirements for Computerized Systems of Document and Process Management of the Judiciary). The last version is 2.2.2 of July 2014.

The MoReq-Jus resulted from the necessity to establish minimum requirements for computerized systems of the Brazilian Judiciary, to ensure the reliability, authenticity and accessibility of documents and processes managed by these systems. The Brazilian model of interoperability was based on the Model Requirements for the Management of Electronic Records (MoReq), prepared by the Data Exchange program between Administrations (IDA) of the European Commission and on its latest version, MoReq2 and on the PREMIS (Preservation Metadata: Implementation Strategies) Working Group, version of March 2008[3]. These models have the common goals to provide requirements for the acquisition, development and evaluation of the management systems of processes and documents:

Digital - the metadata and the documents themselves are entered into the system;

Non-digital - the system records only the metadata of the documents; and

3 See details on the websites http://ec.europa.eu/idabc/en/document/413.html and http://www.oclc.org/research/activities/pmwg.html?urlm=159816

Hybrids – it enables the management of non-digital and digital documents.

The MoReq-Jus sets minimum conditions to be fulfilled in the production, of processing, custoding, storaging, preserving, archiving or receiving documents, by the management systems of processes and digital, non-digital or hybrids documents, in order to ensure its reliability, authenticity and access.

The MoReq-Jus evaluates the Computerized Systems of Management of Processes and Documents (GestãoDoc), regardless of the technology platform in which they have been developed and deployed.

The MoReq-Jus model is composed of 17 areas, as listed below. The model has specific modules, with 17 checklists for the evaluation of each area of the management systems of processes and digital, non-digital and hybrid documents:

a. Organization of Institutional Documents

b. Capture

c. Storage

d. Preservation

e. Security

f. Processing and Workflow

g. Evaluation and Destination

h. Search, Location and Document Presentation

i. Administrative Functions

j. Usability

k. Interoperability

l. Availability

m. Performance and Scalability

n. Metadata

    i. content

    ii. preservation

    iii. auditing

    iv. security

From this model, the CNJ certifies management software of digital processes of the Brazilian Courts. From our point of view, along with the certification, it is necessary to evaluate the systems over its operation and mainly from the user's perspective. Constant research about user's satisfaction also should be included in the methodology of evaluation.

The research project described in this article intended to accommodate the user's perspective with respect to the requirement of level and interoperability capability of the system to be assessed. One big problem for lawyers and law firms in Brazil is the need to install several specific applications on their desktops for each computerized system and for each court in which they operate.

3. Quantifying the interoperability of systems of EJP

The research project aims to do a mapping of the main electronic processes systems in use in Brazil from the perspective of interoperability among those most relevant and widespread. In 2015 the group will apply the following survey with the following questions, to be answered by the ICT teams of the Brazilian Courts. Each question is assigned with a weight between 1 and 3 (weight 3 is more important and influential for the final result of the evaluation).

Questions:

1. Is the system interoperable with the EJS of superior and lower judicial instances? (3)

2. Is the system interoperable with software aimed at aiding the visually or hearing impaired? (3)

3. Is the system interoperable with several operational systems like Windows, OS X and Linux? (3)

4. Is the system interoperable with any web browser? (3)

5. Is the system interoperable with mobile phones and tablets (Android and iOS)? (3)

6. Is the system interoperable with the EJS of other Courts? (2)

7. Is the system externally interoperable with registers and other auxiliary organs to the Judiciary? (3)

8. Is the system interoperable with systems of Judicial fees payment? (3)

9. Which is the access of the user to the system: via login/password or digital certificate? (1)

10. Does the system allow for a query of documents? (3)

11. Does the system allow for retrieval of documents? (3)

12. Does the system allow for the importing of documents? (3)

13. Does the system allow for the exporting of documents? (3)

14. Does the system support one of the following document formats: XML versions 1.0 or 1.1 (.xml) or XSL (.xsl), or Open Document (.odt) or PDF open version PDF/A or pure text (.txt) or HTML version 4.01 (.html or .htm)? (3)

15. Spreadsheet files: Does the system support Open Document (.ods) spreadsheet files ? (3)

16.Presentation files: Does the system support Open Document (.odp) or HTML (.html ou .htm) presentation files ? (3)

17. Files of database for workstations: Does the system uses XML files versions 1.0 or 1.1 (.xml) or MySQL Database (.myd, .myi) generated in MySQL formats, version 4.0 or higher or pure text (.txt) or (.csv) or Base archive (.odb)? (3)

18. Exchange of graphical information and static images: Does the system use PNG (.png) or TIFF (.tif) or SVG (.svg) or JPEG File Interchange Format (.jpeg, .jpg or .jfif) or Open Document (.odg)? (3)

19. Vector graphics: Does the system use one of the formats SVG (.svg) or Open Document (.odg)? (2)

20. Specification of animation patterns: Does the system use the SVG (.svg)? (2)

21. Audio files and Video files: Does the system use .mpg or MPEG-4 (.mp4) or MIDI (.mid) or Ogg Vorbis I (.ogg) or Audio-Video Interleaved (.avi) with Xvid encoding? (3)

22. Compression commonly used files: Does the system use one of the ZIP formats (.zip) or GNU ZIP (.gz) or TAR Package (.tar) or compressed TAR package (.tgz or .tar.gz) or BZIP2 (.bz2) or TAR pack compressed with BZIP2 (.tar.bz2)? (3)

23. What is the periodicity of the updates planned for the system? (3)

24. Does the system interface follow established and consolidated standards as good graphic design practices, are they scientifically validated? (3)

25. Does the display interface of institutional documents provide drag and drop resources, where it is appropriate in the operating environment of the system? (1)

26. Does the system allow for its use and that it is not mandatory to use specific selector devices (e.g., the mouse)? (3)

27. Does the system allow to perform the most frequently performed transactions or tasks with a small number of iterations (e.g. mouse clicks)? (1)

28. Is the system integrated with the standard system of editing documents? (3)

Mandatory standards for Organization and Exchange of Information:

29. Language for data exchange: XML - yes or no? (3)

30. Data transformation: XSL or XSL Transformation (XSLT) - yes or no? (3)

31. Definition of data to exchange: XML, XML Schema Part 0: Primer; XML Schema Part 1: Structures; XML Schema Part 2: Datatypes, UML (Unified Modeling Language) - yes or no? (3)

Integration areas for Electronic Government:

32. Legislation, Legal Decisions and Legislative Proposals - mandatory use of LexML v. 1.0 http://projeto.lexml.gov.br: yes or no? (3)

33. Georeferenced Information - Interoperability between geographic information systems: Does the system use one of the following standards: WMS or WFS or WCS CSW or WFS-T or KT/WKB? (1)

Through these questions the Research will measure the interoperability of the EJS of Brazil and will also rank the most and least interoperable EJS.

4. References

[1] European Interoperability Framework for Pan-European eGovernment Services. Luxembourg: Office for Official Publications of the European Communities, 2004. ISBN 92-894-8389-X. Available at http://ec.europa.eu/idabc/en/document/3473/5887.html

[2] CONTINI, Francesco; LANZARA, Giovan Francesco. Beyond Interoperability: Desiging Systems for European Civil Proceedings Online. In: CONTINI, Francesco; LANZARA, Giovan Francesco (Org.). *Building Interoperability for European Civil Proceedings Online*. Bologna: CLUEB, 2013.

[3] SERBENA, C.A. and KRASSUSKI, L.H. Electronic Judicial Process and Interoperability: Currently State of Affairs in Brazil and Comparative Law. In SCHWEIGHOFER, E., KUMMER, F., HÖTZENDORFER W. Co-operation IRIS 2015. Proceedings of the 18th International Legal Informatics Symposium. Vien: Austrian Computer Society, 2015. p. 291-300.

Abbreviations

EJS – Electronic Judicial Systems - www.cnj.jus.br/programas-de-a-a-z/sistemas/processo-judicial-eletronico-pje

CNJ – National Council of Justice – www.cnj.jus.br

LA CONSTITUCIÓN DE EMPRESAS EN LÍNEA: NUEVAS TECNOLOGÍAS, ESTADO Y EMPRESA

ON-LINE ESTABLISHMENT OF COMPANIES: NEW TECHNOLOGIES IN GOVERNMENT AND BUSINESS

Dennis José Almanza Torres[1], Flor Zúñiga Maldonado[2]

[1] PhD student, Federal University of Paraná (Brazil). Praça Santos Andrade, 50, 3° and – Centro, Curitiba – PR. – Brasil. almanzadennis@gmail.com.
[2] Master of law, Federal University of Paraná (Brazil). Praça Santos Andrade, 50, 3° and – Centro, Curitiba – PR. – Brasil. Florzuniga10@gmail.com

Keywords: ***Company, business establishment, e-governance***

Abstract: *Market opening and facilitating free trade are guiding principles of the current Peruvian Political Constitution of 1993. These principles work as a basic fundament for the government to facilitate processes having the objective of enhancing trade. These principles served well to increase commercial activities in recent years. A high quantity of business partnerships has been founded. However, one of the main reasons of the lack of formal business establishment is tax evasion.*

The refusal to pay taxes has a negative impact on the welfare of the population. Business owners frequently avoid the registration of their business due to the tedious amount of steps and time that the process of business formalization takes as well as the costs involved in this process.

In order to deal with this situation, the Peruvian government decided to move to new technologies of information and communication (ICT). These mechanisms intend to facilitate the process of business formalization. The Peruvian government, through the joint efforts of the Presidency of the Council of Ministers, the Association of Notaries of Lima, the Andean Development Corporation (CAF) and the support of the National Supervisory Authority of Tax Administration (SUNAT), the National Supervisory Authority of Public Records (SUNARP) and the National Registry of Identification and Civil Status (RENIEC), finalized the Online System of Business Formation. This electronic system facilitates the process of establishment of Peruvian business companies; streamlining notary and registration procedures thought the use of the internet and digital signature.

With the use of this new system, the process of establishment is performed in a maximum of three working days (72 hours) - due to the majority of the process is made online, compared to the previous process when it could take approximately around 120 days.

The new system of business registration has some important advantages. This new system does not request the physical presence of the applicant. The process monitoring is done online. Information is recorded only at the beginning. The process also reduces the chances of typing errors, as well as counterfeiting and falsification of documents. Additionally, the costs associated with lawyer's staff, processing and transport are reduced. Notary costs are diminished by the standardization of formats. The registry operations decreased by the use of pre-approved formats.

1. Introducción

La vigente Constitución Política peruana de 1993 tiene como uno de sus principios rectores la apertura al mercado y la facilitación del libre comercio. Este dispositivo sirve como base para que el Estado pueda facilitar procesos que tengan como objetivo el incremento del comercio. Si bien este dispositivo sirvió para que las actividades comerciales se incrementen durante los últimos años, no se ha observado un notorio aumento en la cantidad de sociedades empresariales formalmente constituidas. Una de las principales consecuencias de la no formalización de las empresas es la evasión tributaria. La negativa a pagar impuestos por parte de los empresarios afecta las funciones del estado como ente que busca el bienestar social de la población.

Son diversos los motivos por las que las empresas optaban por no formalizarse, dentro de estos destaca el oneroso y demorado trámite administrativo que este proceso implicaba. Los costos que se deben pagar por formalización de una actividad comercial, por lo general, son los principales obstáculos. Frente a ese panorama, el gobierno peruano decidió actuar valiéndose de las nuevas tecnologías de la información y comunicación (TICS). Con estas herramientas se busca que la formalización no represente mayores gastos de tiempo y dinero.

El gobierno peruano, a través de sus organismos como la Presidencia del Consejo de Ministros, el Colegio de Notarios de Lima, la Corporación Andina de Fomento (CAF) y con el apoyo de la Superintendencia Nacional de Administración Tributaria (SUNAT), la Superintendencia Nacional de Registros Públicos (SUNARP) y el Registro Nacional de Identificación y Estado Civil (RENIEC), concretó el Sistema de Constitución de Empresas en Línea. Éste sistema surgió como una forma de facilitar al ciudadano la formalización de su empresa; agilizando trámites notariales y registrales de manera electrónica con el uso de internet y la firma digital.

Con este proceso, el trámite de formalización se realiza en un tiempo máximo de tres días hábiles (72 horas) – antiguamente este proceso demoraba en torno de 120 días aprox. - ya que casi todos los procesos se realizan en línea. El solicitante tiene una presencia física mínima, se realiza el seguimiento del proceso vía online, solo se consigna información al inicio.

Con este proceso se reducen errores en digitación, se minimiza la falsificación y adulteración de documentos, los costos asociados a los abogados, tramitadores y al transporte son reducidos. Asimismo, los costos notariales son disminuidos por la estandarización de formatos, las operaciones registrales disminuyen por el uso de formatos pre aprobados.

En el presente trabajo pretendemos explorar este nuevo proceso, determinar aspectos relacionados a su implantación y los beneficios que la constitución de empresas en línea trae para los empresarios y para el Estado.

2. Las nuevas tecnologías y la Administración Pública

Diversos factores influyen para que en la actualidad sea común referirse a nuestra sociedad como una “sociedad tecnológica”, esta denominación surge por la predominancia e influencia que ejerce la tecnología de la información en nuestros quehaceres diarios. Ciñendo la definición de lo que sería la tecnología de la información, se puede decir que esta está referida al conjunto de recursos tecnológicos empleados para el uso de la información o aun, de recursos no humanos dedicados al almacenamiento, procesamiento y comunicación de la información.[1]

[1] WACHOWICZ, Marcos e CASAGRANDE, Thais de Santos. A inclusão digital dos advogados: Gestao da tecnologia da informação e comunicação nos escritórios de advocacia. In: _____. (coord.) *Direito da sociedade da informação & propriedade intellectual,* p. 95.

Los orígenes de esta expresión pueden ser hallados en 1958, en un artículo de Leavith y Whisler titulado "Administrando en los años 80", allí, estas innovaciones fueron definidas como los medios utilizados por las empresas productivas para promover y potencializar el proceso de creación y desarrollo de la capacitación tecnológica.

Si bien en un inicio estas tecnologías estaban referidas a la información bajo un enfoque clásico (entendido como transmisión de datos sobre algo o alguien, por lo general noticias que llegaban al receptor con cierto intervalo de tiempo), actualmente esta idea se ha visto alterada, pues modernamente la palabra información está relacionada al contenido y la integración de diferentes fuentes de conocimiento.[2]

Las Nuevas Tecnologías de la Información pretenden o fueron creadas para acelerar el desarrollo de la sociedad con base en el conocimiento. Todo ello fue entendido por el Estado quien decidió valerse de estos conocimientos a través de reformas administrativas y la creación de un nuevo concepto, el Gobierno Electrónico.

En el Perú, como en otros países de la región, la modernización del Estado que se dio durante los últimos años, estuvo orientado hacia el mercado, se buscó una gestión práctica con miras al cumplimiento de metas y a la obtención de mayor eficiencia. La reforma del Estado obligó a revisar roles, funciones y mecanismos de funcionamiento dentro del aparato estatal para luego adecuarlos a las exigencias de la Nueva Gestión Pública.[3]

El *New Public Management*, cuyo origen se remonta a las reformas administrativas realizadas en Gran Bretaña en la década de los 80, se refiere a las mudanzas organizativas, directivas y operacionales que se dieron en el sector púbico de varios países durante esa década y años posteriores. Esta *Nueva Gestión* tiene como una de sus principales características "el énfasis en la observancia de las "E" (economía, eficiencia y eficacia) en las operaciones administrativas, que conduce a reformas profundas del sistema presupuestal y del control interno de gestión, con el apoyo de la incorporación de las nuevas tecnologías de información".[4]

Fue en el seno de la Nueva Gestión Pública que surgió el Gobierno Electrónico. El Gobierno Electrónico, según lo define la Organización de las Naciones Unidas (ONU), es el uso de las Tecnologías de la Información y la Comunicación (TIC), por parte del Estado, para brindar servicios e información a los ciudadanos, aumentar la eficacia y eficiencia de la gestión pública, e incrementar sustantivamente la transparencia del sector público y la participación ciudadana.

Este sistema atraviesa un proceso evolutivo según la adopción de sus herramientas por parte de un gobierno, la ONU ha distinguido cinco niveles evolutivos de acuerdo al grado de desarrollo en el que se encuentran, estos son: (i) *Presencia emergente*, cuando las herramientas del gobierno electrónico se utilizan solamente para brindar información a través de Internet, (ii) *Presencia ampliada*, cuando la presencia se expande por medios sofisticados, incluyendo búsquedas en *web sites*, y comunicación por medio de correos electrónicos, (iii) *Presencia interactiva*, las organizaciones gubernamentales están presentes masivamente en las páginas web, otorgando algunos servicios como el llenado de formularios electrónicos, (iv) *Presencia transaccional*, cuando el Estado ofrece transacciones completas y seguras propias de la Administración, como el otorgamiento de pasaportes o de certificados. (v) *Integración total*, cuando existe una relación

[2] FREITAS, Cinthia O. de A. Redes sociais: Sociedade tecnológica e inclusão digital. In: Wachowicz, Marcos. (coord.) *Direito da sociedade da informação & propriedade intellectual*, p. 50.
[3] MEZZAROBA, Orides; BERNIERI, Juliana y BIER, Clerilei. Os desafios da governança no novo século, as reformas estatais e a accountability. In: ROVER, Aires et. al. *Direito, governança e tecnologia: princípios, políticas e normas do Brasil e da Espanha*, p. 23,24.
[4] AGUILAR VILLANUEVA, Luis F. *Gobernanza y gestión pública*, p. 144-150.

integrada entre el usuario y el Estado, pues este no percibe la diferencia entre los servicios *on line* y los físicos.[5]

Si bien el gobierno electrónico, ofrece innumerables ventajas para el usuario, debe tenerse en cuenta que para el correcto funcionamiento de este sistema es necesario satisfacer ciertos requerimientos entre ellos de interoperabilidad, infraestructura, legales y de seguridad sobre servicios transaccionales.

La superación de estas etapas hizo posible que se concrete el sistema de ventanilla única de constitución de empresas en línea. La convergencia de especialistas multidisciplinares – legales, tecnológicos y administrativos -, hizo posible la integración tecnológica para la creación de este sistema. Para conseguir este objetivo, fue necesaria una reforma en el sistema jurídico comercial del estado peruano, reforma que aun está en curso debido a las constantes mudanzas en este campo, especialmente por el incremento en la utilización de las nuevas tecnologías de la información y de comunicación.

3. El derecho comercial peruano y su adecuación a un nuevo entorno

El Derecho Comercial se caracteriza por no ser un derecho estático, su transformación hasta alcanzar su fase actual es el resultado de una serie de factores económicos, políticos y sociales; los cuales influyeron decisivamente para el surgimiento de un marco legislativo especial que sea capaz de albergar situaciones nuevas como la aparición de nuevas técnicas de comercialización, la internacionalización del comercio y el uso de las nuevas tecnologías en las diferentes etapas de la actividad empresarial.[6]

Los códigos fueron quienes, en un inicio, regulaban las operaciones comerciales, sin embargo, las frecuentes transformaciones económicas que se intensificaron a partir de los años 60 redujeron su función normalizadora, sustrayendo materias enteras que fueron sometidas a textos específicos, lo que se refleja en las leyes especiales. Es por ello que a lo largo del tiempo, surgieron numerosos estatutos con características propias, diferentes a la legislación codificada, no guardando el carácter universal que estos mantenían, vehiculando normas de derecho material (y procesal) que se adapten fácilmente a la nueva realidad.[7]

En el estado peruano, las normas que regentan actualmente las actividades comerciales tienen como base lo establecido en el vetusto Código de Comercio promulgado el 15 de Febrero de 1902, el mismo que empezó a regir el 1° de Julio del mismo año. El Código de Comercio de 1902 que agrupa las materias en cuatro libros,[8] es el instrumento central de la regulación de las relaciones comerciales. El Código civil peruano de 1984 también tiene un papel destacado como fuente reguladora de los contratos comerciales, tal y como lo establece su artículo 2112.[9]

---

[5] NASER, Alejandra y CONCHA, Gastón. *El gobierno electrónico en la gestión pública*, p. 15-16. Disponible en: http://repositorio.cepal.org/bitstream/handle/11362/7330/S1100145_es.pdf?sequence=1. Acceso en: 18/05/2015.

[6] GONCALVES NETO, Alfredo de Assis. *Direito de empresa: comentários aos artigos 966 a 1.195 do Código Civil*, 2008, p. 33.

[7] TEPEDINO, Gustavo. O Código Civil, os chamados microssistemas e a Constituição: premissas para um reforma legislativa. In: _____ (Coord.) *Problemas de direito civil-constitucional*, p. 4,5. Al respecto Lorenzetti agrega: En algunas situaciones, normas generales se fraccionaron gradualmente hasta desvincularse totalmente del código, en otros casos el surgimiento de nuevas operaciones comerciales y sus complicaciones, fueron la causa para que surjan regulaciones propias. (LORENZETTI. Ricardo Luis. *Razonamiento judicial. Fundamentos de Derecho Privado*, p. 29.)

[8] (i) De los comerciantes y del comercio, (ii) De los contratos especiales de comercio, (iii) Del comercio Marítimo, (iv) De la suspensión de pagos, de las quiebras y de las prescripciones.

[9] Art. 2112. Regimen Unificado. Los contratos de compraventa, permuta, mutuo, deposito y fianza de naturaleza mercantil, se rigen por las disposiciones del presente Código, Quedan derogados los artículos 297° a 314°, 320 a 341° y 430° a 433° do Código de Comercio,

Las constantes variaciones en el sistema económico y social, fueron la causa para que operadores del derecho se enfrenten a situaciones nuevas no previstas en ley, en razón a ello periódicamente fueron creándose nuevas leyes especiales las cuales en la actualidad sirven para regular nuevas figuras comerciales o para actualizar o ampliar las normas del Código de Comercio peruano.[10] Como resultado del proceso de reforma de las leyes comerciales se observó el surgimiento de microsistemas que de alguna forma restaron protagonismo al Código Comercial. Así, surgieron leyes como la Ley General de Sociedades (Ley N° 26887) y la Ley de Títulos Valores (Ley N° 27287), estos microsistemas, fundamentales para el derecho comercial, actualmente constituyen las columnas donde descansa el derecho comercial peruano.[11] Además de estas normas, se tiene la Ley de Promoción y Formalización de la Micro y Pequeña Empresa (Ley N° 28015, de 03 de Julio de 2003), la Ley de la Empresa Individual de Responsabilidad Limitada (Decreto Ley N° 21621, del 14 de setiembre de 1976 modificada por última vez pela Ley 27075 de 26 de marzo de 1999), la Ley General del Sistema Concursal N° 27089 que reemplazó la Sección Primera del libro cuarto del Código de Comercio; y más unas decenas de leyes que regulan situaciones especificas del derecho comercial peruano.[12]

La constitución de empresas en el ordenamiento jurídico peruano tiene como base lo establecido en los artículos quinto, sexto y séptimo de la Ley general de sociedades,[13] y los artículos decimo tercero y decimo cuarto de la Ley de la Empresa Individual de Responsabilidad Limitada,[14] donde, en líneas generales destacan dos momentos: (i) La elaboración de la escritura pública de constitución y (ii) La inscripción en el Registro correspondiente.

En tal sentido se concreta la constitución de la sociedad con la respectiva Inscripción en el Registro correspondiente, a partir de allí – con el inicio de la personalización de la sociedad empresaria – surgen determinados derechos, deberes y obligaciones para este nuevo sujeto de derecho.[15] Entre ellas, la obligación de tributar.

El estado peruano, con la finalidad de ampliar la base tributaria, es decir obtener mayor recaudación buscando que las personas naturales y jurídicas tributen o cumplan con sus obligaciones, sea por

[10] MONTOYA MANFREDI, Ulises. *Derecho Comercial*, p. 61,62.
[11] MONTOYA MANFREDI, U. *Derecho Comercial*, p. 57-59.
[12] Al respecto se recomienda ver: MONTOYA MANFREDI, U. *Derecho Comercial*, p. 62-66.
[13] Art. 5. La sociedad se constituye por escritura pública, en la que está contenido el pacto social que incluye el estatuto. Para cualquier modificación de éstos se requiere la misma formalidad.
En la escritura pública de constitución se nombra a los primeros administradores, de acuerdo con las características de cada forma societaria.
Los actos referidos en el párrafo anterior se inscriben obligatoriamente en el Registro del domicilio de la sociedad.
Cuando el pacto social no se hubiese elevado a escritura pública, cualquier socio puede demandar su otorgamiento por el proceso sumarísimo.
Art. 6. La sociedad adquiere personalidad jurídica desde su inscripción en el Registro y la mantiene hasta que se inscribe su extinción.
Art. 7. La validez de los actos celebrados en nombre de la sociedad antes de su inscripción en el Registro está condicionada a la inscripción y a que sean ratificados por la sociedad dentro de los tres meses siguientes. Si se omite o retarda el cumplimiento de estos requisitos, quienes hayan celebrado actos en nombre de la sociedad responden personal, ilimitada y solidariamente frente a aquellos con quienes hayan contratado y frente a terceros.
[14] Art. 13. La empresa se constituirá por escritura pública otorgada en forma personal por quien la constituye y deberá ser inscrita en el Registro Mercantil.
La inscripción es la formalidad que otorga personalidad jurídica a la Empresa, considerándose el momento de la inscripción como el inicio de las operaciones.
Art. 14. La validez de los actos y contratos celebrados en nombre de la Empresa antes de su inscripción en el Registro Mercantil, quedará subordinada a este requisito. Si no se constituye la Empresa, quien hubiera contratado a nombre de la Empresa será personal e ilimitadamente responsable ante terceros.
[15] COELHO, Fabio Ulhoa. *Curso de direito comercial. Direito de empresa*, p. 34,35.

medio de la creación de la conciencia tributaria en el ciudadano o facilitando la formalización de las empresas informales, emitió diversas normas[16] que buscan conseguir este objetivo.[17]

A todo ello, debe considerarse que una de las razones por las que varios emprendedores opten por no formalizarse – no crear una sociedad regular – es por el costo que representa este procedimiento. Costo referido tanto al valor económico que el empresario debe invertir como el costo representado en el tiempo que utiliza para concretar la inscripción.

Con el objetivo de disminuir las etapas inherentes a la creación de una empresa de manera tradicional, el Estado peruano, valiéndose de las nuevas tecnologías de la información, ha desarrollado sistemas que permiten agilizar estos trámites.

4. Las Nuevas tecnologías y las actividades empresariales

Con el advenimiento de la era electrónica, diversos conceptos y figuras deben ser revisados y replanteados, pues esta evolución de la tecnología informática trae consigo una transformación en costumbres y procedimientos que la sociedad tenían (o tienen) arraigadas. En el derecho, esto se refleja en temas como la desmaterialización de los títulos ejecutivos, los remates judiciales por internet o los interrogatorios *on line* o por video conferencia.[18]

De la misma forma, la automatización de los procesos de la toma de decisiones dentro de la empresa resulta absolutamente necesaria en un entorno dinámico como el actual, esto ha propiciado el empleo progresivo de la tecnología informática como medio para la toma de decisiones rápidas y de gran importancia para la empresa.

Como el conocimiento formalizable matemáticamente representado por la automatización de tareas operativas no ha sido capaz de solucionar problemas cuya salida no resultaba de la aplicación de algún proceso algorítmico, surgió un nuevo núcleo de conocimiento denominado Inteligencia Artificial (IA)[19] donde, además de otras, se desarrollaron programas informáticos que emulan el comportamiento de expertos humanos en la resolución de problemas (Sistemas Expertos, SS.EE.),[20]

Las acepciones que nos permiten identificar cuando se está frente a un sistema que emule determinadas funciones de la inteligencia humana parte de tres criterios: (i) un criterio débil, que consiste en cualquier operación que no aparente ser demasiado rutinaria, se citan como ejemplos la búsqueda de una información entre muchas otras partiendo de varios gran cantidad de datos no relacionados entre sí, o la representación de un dialogo con el operador a través de un sistema complejo de opciones.

---

[16] Muestra de ello es la ley N° 28932, emitida por el Congreso de la República, donde se delegan facultades legislativas al Poder ejecutivo con la finalidad de que pueda ampliar la base tributaria, lograr mayor eficiencia, equidad y simplicidad en el Sistema Tributario Nacional; y dotar al país de un Sistema Tributario predecible que favorezca el clima de inversión.

[17] ZAVALETA ALVAREZ, Michael y otros. Servicios digitales en el impuesto a la renta peruano. Interpretación a la luz de la reforma tributaria. In: *Revista peruana de derecho tributario*, N° 2, p. 3,4.

[18] Una sucinta revisión sobre el tema puede ser observado en: GRECO, Leonardo. O proceso eletrônico. In: SILVA Jr. Roberto. *Internet e Direito. Reflexões doutrinarias*, p. 13-19.

[19] Guiborg define a la inteligencia artificial como "la capacidad de un sistema informático (hardware más software) para reproducir, imitar o ejercer alguna de las funciones que suelen atribuirse a la inteligencia humana". GUIBORG, Ricardo, *Informática jurídica*, p. 795. Disponible en: http://biblio.juridicas.unam.mx/libros/8/3875/25.pdf , Acceso en: 21/05/2015.

[20] Como explica Fortuna Lindo: "Los sistemas expertos, con el transcurso del tiempo, han supuesto la ayuda o la automatización de forma real y contrastada, de los más variados problemas de decisión empresarial: selección de inversiones, reclutamiento de personal, diseño de cadenas de distribución, transmisión de información relevante para la dirección, concesión de créditos, planificación estratégica, etc.; destacando el entorno financiero como uno de los campos con mayor numero de aplicaciones". FORTUNA LINDO, José María. Una introducción a los sistemas expertos en la empresa. Disponible en: https://dialnet.unirioja.es/descarga/articulo/786048.pdf , p. 50. Acceso en 22/05/2015.

Una segunda acepción utiliza la expresión para designar cualquier programa capaz de almacenar numerosos datos de una determinada especialidad del conocimiento y presentarlo, frente a una consulta, aproximadamente como lo haría un experto humano. En este grupo se pueden incluir sistemas expertos en medicina, en geología o en ingeniería.

Una tercera acepción, y la que goza de mayor aceptación en la actualidad, es la que reserva el nombre de inteligencia artificial para los sistemas capaces de aprender de la realidad, por medio del proceso del ensayo y del error. "Tales sistemas no dependen enteramente de la fijación de pautas por el hombre: una vez establecido un objetivo e introducidos algunos parámetros acerca del tipo de procedimientos admisibles para alcanzarlo, el programa intenta distintos caminos o combinaciones de caminos, retiene los que dan mejor resultado y desecha los que se demuestran inútiles. De un sistema así es posible afirmar, en alguna medida, que investiga por su cuenta y pone en práctica los resultados obtenidos".[21]

Sin embargo, si la inteligencia artificial es analizada desde una perspectiva amplia, se observa que todo esto hace parte de un proceso continuo que se fue dando a través del tiempo. Dentro de las funciones de determinado ordenador, la comparación o emisión de juicios siempre estuvo presente manifestándose por medio de un lenguaje sencillo y entendible para el usuario. Solamente cuando los juicios no son sencillo o no se conocen las reglas que lo rigen, la función de las maquinas aparentan ser misteriosas. Claro esta, está posición es bastante discutida.

De todas maneras, procesos que envuelven algún tipo de elección, a pesar de no basarse en coleccionar y procesar grandes volúmenes de datos mediante procedimientos complejos que impliquen cálculos científicos y técnicos avanzados, poseen determinadas particularidades que hacen posible enmarcarlos como parte de un sistema experto.

Tal sería el caso del servicio de constitución de empresas en línea disponible en algunas ciudades del estado peruano.[22]

La constitución de empresas en línea se realiza, casi en su totalidad prescindiendo de la presencia física. Este proceso se inicia con la reserva del nombre de una empresa. Esta reserva se realiza personalmente (aproximándose a las oficinas de la Superintendencia Nacional de Registros Públicos - SUNARP), o de manera virtual desde cualquier espacio físico, el pago por la realización de esta reserva se puede concretar por medio de tarjetas de crédito.

Posteriormente es necesario registrarse en el *web site* de dicha institución (www.sunarp.gob.pe), en el módulo "Registro de usuario SUNARP", la inscripción y el llenado de datos en dicho *site* genera una cuenta de usuario y una seña, con la cual se puede continuar con el procedimiento.

En el módulo "solicitud de trámite de Constitución de empresa", que es el siguiente paso, se aceptan (o no) los términos y condiciones que establece este organismo para la creación de las empresas *on line*. Luego de aceptado, se elige alguna Notaría (*Notary´s office*) entre las que el sistema proporciona, en seguida se selecciona el tipo de persona jurídica que se desea constituir, para luego completar los datos del acto constitutivo que el sistema ha generado para el usuario, entre estos, son necesarios: datos del solicitante, datos de la empresa a constituir, datos de la ubicación geográfica, datos del capital social y datos de los participantes.

Concluido el llenado de estos documentos, el acto de constitución ha sido generado completamente, faltando corroborar la información y si está conforme, proceder con la opción finalizar, a partir de allí, los documentos y toda la información son encaminados virtualmente a la notaría que se

[21] GUIBORG, Ricardo, *Informática jurídica*, p. 796. Disponible en: http://biblio.juridicas.unam.mx/libros/8/3875/25.pdf , Acceso en: 21/05/2015.

[22] Estas son Lima metropolitana, Callao, San Martin, Lambayeque, La Libertad, Arequipa, Cusco, Madre de Dios, Puno, Tumbes y Piura.

seleccionó en un inicio. Finalmente se emitirá una constancia de constitución de empresa que será enviada a su correo electrónico.

Concluida esta etapa, el usuario dispone de un plazo de treinta días para aproximarse a la Notaría seleccionada para firmar el respectivo documento, con este acto se finaliza una primera etapa la cual, tradicionalmente concluía con la suscripción de la escritura pública por parte del (los) interesado(s). Restando la inscripción en la oficina de Registros Públicos (SUNARP), para ello, por correo electrónico se informará al usuario, la fecha y hora de presentación del parte notarial a esta oficina, así como el resultado de la calificación registral, dando fin, con ello, a la etapa de constitución de sociedades.

Como se explicó en un inicio, lo que se busca con la utilización de las nuevas tecnologías es facilitar las actividades de los ciudadanos a efecto de disminuir o reducir costos agilizando la concretización de trámites que de manera tradicional implicaban mayor gasto físico y económico, este sistema permite cumplir con estos objetivos valiéndose para ello de las NTI.

5. Referencias Bibliográficas

AGUILAR VILLANUEVA, Luis F. *Gobernanza y gestión pública*. México: Fondo de cultura económica, 2006.

BARRAGAN, Julia. *Informática y Decisión Jurídica*. México: Fontamara, 2008.

COELHO, Fabio Ulhoa. *Curso de direito comercial. Direito de empresa*, Sao Paulo: Saraiva, 2013.

FORTUNA LINDO, José María. Una introducción a los sistemas expertos en la empresa. Disponible en: https://dialnet.unirioja.es/descarga/articulo/786048.pdf , p. 4-66. Acceso en 22/05/2015.

FREITAS, Cinthia O. de A. Redes sociais: Sociedade tecnológica e inclusão digital. In: Wachowicz, Marcos. (coord.) *Direito da sociedade da informação & propriedade intelectual.* Curitiba: Juruá, 2012, p. 43-65.

GRECO, Leonardo. O proceso eletrônico. In: SILVA Jr. Roberto. *Internet e Direito. Reflexões doutrinarias,* Rio de Janeiro: Lumen Juris, 2001, p. 11-31.

GUIBORG, Ricardo, *Informática jurídica*, p. 791-893. Disponible en: http://biblio.juridicas.unam.mx/libros/8/3875/25.pdf , Acceso en: 21/05/2015.

GONCALVES NETO, Alfredo de Assis. *Direito de Empresa. Comentários aos artigos 966 a 1.195 do Código Civil*. São Paulo: Revista dos Tribunais, 2007.

IRTI, Natalino. *La edad de la descodificación*. Barcelona: Bosh, 1992.

LORENZETTI, Ricardo Luis. *Razonamiento Judicial. Fundamentos de Derecho Privado*. Lima: Grijley, 2006.

MEZZAROBA, Orides; BERNIERI, Juliana y BIER, Clerilei. Os desafios da governança no novo século, as reformas estatais e a accountability. In: ROVER, Aires et. al. *Direito, governança e tecnologia: princípios, políticas e normas do Brasil e da Espanha*. Florianopolis: Conceito editorial, 2014, p. 17-38.

MONTOYA MANFREDI, Ulises. *Derecho Comercial, Tomo I, Parte General, Derecho de Sociedades, Derecho Concursal, Derecho del consumidor, Derecho de la Competencia.* Lima: Grijley, 2004.

NASER, Alejandra y CONCHA, Gastón. *El gobierno electrónico en la gestión pública.* Santiago de Chile: Instituto Latinoamericano y del Caribe de Planificación Económica y Social (ILPES), 2011. Disponible en: http://repositorio.cepal.org/bitstream/handle/11362/7330/S1100145_es.pdf?sequence=1. Acceso en: 18/05/2015.

TEPEDINO. Gustavo. O Código Civil, os chamados microssistemas e a Constituição: premissas para uma reforma legislativa. In: _____ (Coord.) *Problemas de Direito Civil-Constitucional*. São Paulo: Renovar, 2000, p. 1-16.

WACHOWICZ, Marcos e CASAGRANDE, Thais de Santos. A inclusão digital dos advogados: Gestao da tecnologia da informação e comunicação nos escritórios de advocacia. In: _____. (coord.) *Direito da sociedade da informação & propriedade intelectual.* Curitiba: Juruá, 2012, p. 93-117.

ZAVALETA ALVAREZ, Michael y otros. Servicios digitales en el impuesto a la renta peruano. Interpretación a la luz de la reforma tributaria. In: *Revista peruana de derecho tributario*, N° 2, año 1, p. 2-15. (2007)