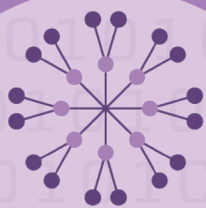


7th International Conference of  
PhD Students and Young Researchers

# LAW 2.0.

NEW METHODS, NEW LAWS

CONFERENCE PAPERS



INTERNATIONAL  
NETWORK  
OF DOCTORAL  
STUDIES IN LAW

Vilnius University Press

2019

## **INFORMATION ABOUT THE CONFERENCE:**

**Venue: Vilnius University, Vilnius, Lithuania**

Date: 25-26 APRIL 2019

### **Scientific Committee of the Conference:**

Prof. Dr. Tomas Davulis, Faculty of Law, Vilnius University

Dr. Donatas Murauskas, Faculty of Law, Vilnius University

Prof. Dr. Monika Namyslowska, Faculty of Law and Administration, Lodz University

Dr. Vigita Vėbraitė, Faculty of Law, Vilnius University

### **Conference Papers Editions composed by:**

Dr. Vigita Vėbraitė, Faculty of Law, Vilnius University

Karolina Mickutė, Faculty of Law, Vilnius University

Ieva Marija Ragaišytė, Faculty of Law, Vilnius University

**ISBN 978-609-07-0264-2**

**© Vilnius University, 2019**

Copyright © 2019 [Authors' Group]. Published by Vilnius University Press

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

# THE ROLE OF SEMANTIC WEB IN THE MANAGEMENT OF LEGAL DATA

Victor Terekhov<sup>1</sup>

## Abstract

Modern Web is the most important source of legal information: it contains legislative texts, case-law, doctrine, drafts and outdated laws from all over the world. The amount of such materials keeps growing with time, while their availability to wide audience presents a significant problem as it remains hard to find, process, analyze and systematize them. This leads to an 'information deficit' paradox: despite there being plenty of relevant data, it is at times impossible to make proper use of it. One of the possible solutions is the further development of Semantic Web – an extension of the current Internet structure relying on logical concepts and correlations between events, and making use of specific ontologies. The application of such technologies in the legal field may permit faster and more precise search, easier reasoning, storage and use of data as well as automatic resolution of disputes by relying on the previous case-law systematized in a machine-understandable format.

**Keywords:** Semantic Web, Legal Data, Linked Open Data, Legal Knowledge Interchange Format (LKIF), smart applications, e-government

## Introduction

The modern World Wide Web (WWW) is a great source of information in both general and specialized fields (including the legal domain). Speaking of the latter, one may find texts of binding legal instruments, judicial cases of various jurisdictions (resolved and in progress), drafts, soft law, outdated and historic sources ranging from the *Manusmriti* and *Hammurabi Code* to the American Declaration of Independence to the present-day blockchain and eCommerce regulations, proposed amendments and, last but not least, doctrine. All of these may be united under the common term of 'legal knowledge', since it does not differentiate between the status (binding/nonbinding), jurisdiction or type of the document, but rather underlines that it contains certain information pertaining to the law.<sup>2</sup> Most of these materials are available worldwide in common formats and may be accessed free of charge. In fact, our generation appears to be in the best position in terms of access to legal data: relevant information is at our fingertips and does not require visiting libraries or purchasing printed

---

<sup>1</sup> PhD in Civil Procedure, Master Degree in International and EU Law Vilnius University. Currently a lecturer of European Private Law and European Civil Procedure at Vilnius University. Recent research interests include comparative law, legal technology (Legal Tech) and online dispute resolution. E-mail: victor.terekhov@tf.vu.lt

<sup>2</sup> P. Casanovas, N. Castellás et al. 'An Ontology-Based Decision Support System for Judges' in: 'Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood' (Amsterdam, IOS Press 2008) 167.

copies of the documents. The Web is, without doubt, the largest repository of legal data,<sup>3</sup> and it keeps growing.

The main beneficiaries of such situation are 'law professionals': judges, arbitrators, public prosecutors, notaries, attorneys and judicial clerks, however ordinary citizens wishing to know better their rights and remedies can also make use of data stored in the Web. In fact, access to legal information is nowadays treated as a fundamental element of democracy, since, on the one hand, we do not excuse liability due to the lack of knowledge of applicable law and, on the other hand, we demand that such laws be precise and clear, so that an average individual can grasp their meaning and learn of his/her rights and obligations.<sup>4</sup>

The potential of the Web as a source of legal information is undisputed throughout the world. In many countries it even becomes common practice to treat publication of newly adopted statutes on an official website of the governmental institution as an official promulgation,<sup>5</sup> and it is hard not to see it as such since Internet tools are accessible to more citizens than paper-based newspapers or journals which were previously used for the same purpose.

At the same time, there are some problems associated with Web as a source of legal information. Firstly, with all its petabytes of data,<sup>6</sup> Internet presents a vast 'ocean' of water where the relevant information constitutes just a small landmass ('island'). It is not easy to find exactly what you are looking for. Secondly, the Web has a lot of 'junk files': irrelevant, broken, infected, or simply incomplete and inaccurate. That sometimes casts doubt on its ability to serve as a reliable source of legal knowledge to consult with. Professionals need an instrument of high accuracy as often the lives of other individuals depend on their reasoning, and the latter may not be arbitrary. Citizens and companies also want more clarity as to their legal status, available rights, imposed obligations and procedures to follow. That is not what they get when different web portals cannot agree on such important issues as the amount of tax to be paid, full list of the documents needed to register a company or the method used by the Border Service to count how many days an individual has spent in the country in order to obtain a right to claim permanent residence there. It is not exactly legal certainty that people want: they simply want correct answers to the questions posed, and they want them to appear within the first lines of search results in Google or Bing.

This paper tries to show what can be done within the legal domain to improve the situation. The idea behind the changes is not new – it is that of Semantic Web, which was first mentioned by Sir Tim Berners-Lee back in 1994 and stands for a virtual space where most of the tasks are performed by mutually communicating machine algorithms (software agents), while humans just rely on their work.<sup>7</sup> This system is based on some new and recently popularized formal languages and builds strong ontologies for its sectors of application. In the legal field the deployment of such technologies can bring faster and qualitatively better search results, automation of some comparative analysis and reasoning

---

<sup>3</sup> E. Francesconi, 'On the Future of Legal Publishing Services in the Semantic Web' [2018] 10(6) *Future Internet* 48.

<sup>4</sup> E. Francesconi and G. Peruginelli, 'Integrated Access to Legal Literature through Automated Semantic Classification' [2008] 17 *Artificial Intelligence and Law* 31.

<sup>5</sup> In that way it is treated, inter alia, in the Russian Federation, see: Federal Law no. 289-FZ [2011] [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_120715](http://www.consultant.ru/document/cons_doc_LAW_120715).

<sup>6</sup> According to Live Counter (<https://www.live-counter.com/how-big-is-the-internet>), at the time of preparation of this paper (14 April), the size of the Internet in Petabytes was 14.964.536 (1 petabyte =  $10^{15}$  bytes).

<sup>7</sup> L. Yu, 'A Developer's Guide to the Semantic Web' (Berlin: Springer 2014) 3-5.

tasks and, to some extent, even autonomous and automatic decision-making in particular cases (fines for traffic violations, IP rights management, solution of monetary claims, etc.).

It must be observed that this paper is written from the position of the lawyer (not a computer engineer or a philologist trained to work with natural language constructions). This presupposes that the question posed by the author is 'in what way the Semantic Web can contribute to the law and the daily routine work of the lawyers with the legal sources', while answering the question of 'how exactly that may be done in terms of available programming tools' remains outside the reach of the research. Here the author cannot but note the necessity of integrated multidisciplinary approach to law in the XXI century and effective collaboration between professionals of legal, IT and (last, but not least) linguistic fields of science and practice.

## 1. Classic Web and its drawbacks

The first stage of the WWW evolution was marked by static information. It was based on simple pages that contained text, images and embedded videos and had their own Uniform Resource Locators (URLs) to be accessed. This was also known as Web 1.0 or 'documentary web'. Most users were consumers rather than creators of the content. At the beginning of the XXI century the mankind saw a move towards Web 2.0, the main feature of which was heavy reliance on social networks (Facebook, LinkedIn, Twitter) and open interaction between independent users. Users also started to actively generate content, which could be pictures of cats and memes, but on the other hand, scholar articles, commentaries to statutes, case analyzes and model acts.<sup>8</sup> At the same time the Web has become more exposed to the threat of 'fake information'.<sup>9</sup> If the content can be generated by anyone, you cannot immediately trust it, as you do not know the aim and motives behind the will to share that information and the fact of whether a person is acting in good faith. Thus, despite being the largest repository of information in all domains (including legal), Web cannot be trusted unconditionally, as many of the publications are not specifically checked for consistency and may be manipulated by the website owner in his own interest.<sup>10</sup> In some situations, it may even take place unintentionally: with purely legal issues that happens when the website contains an outdated or incomplete version of the statute, or an act matching a totally different jurisdiction. The only possible exception (when you can trust what you see) is a situation where you deal with an official page of the relevant governmental or international institution.

In most cases though the quest for finding relevant information starts with the search portal (such as Google or Bing). These services have complex searching tools that help us find and look through the documents. However, there is one major limitation – all of the searching algorithms rely on keyword matching, so that they provide you with documentary search results that contain exactly the same words and phrases you entered in the searching box.

---

<sup>8</sup> J. Breuker, P. Casanovas et al. 'The Flood, the Channels and the Dykes: Managing Legal Information in a Globalized and Digital World' in: 'Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood' (Amsterdam, IOS Press 2008) 10.

<sup>9</sup> S. Mo Jang and J. K. Kim, 'Third Person Effects of Fake News: Fake News Regulation and Media Literacy Interventions' [2018] 80 Computers and Human Behavior 295-296.

<sup>10</sup> T. Bascik, 'Web sémantique: Quelles perspectives pour l'information juridique?' [2014] 19(1) Lex Electronica 27.

This is not bad when you are sure about what you are looking for. At the same time, more often you experience certain problems trying to get what you want. To begin with, words forming the keywords may be polysemic and homonymous, may have synonyms, acronyms and abbreviations or just be used with different meanings depending on the content. Thus, the words ‘business’, ‘trader’ or ‘entrepreneur’ are used in different EU acts on consumer protection to describe the ‘stronger’ party to a contractual relation. Some words, on the contrary, can be polysemantic and mean different things (e.g. ‘arbitration court’ in Russian law means a judicial institution that resolves disputes of economic nature between companies and/or individual entrepreneurs, but in most other states an ‘arbitration’ is seen as a form of alternative dispute resolution that takes place outside courts and is conducted by specialized individuals or institutions). Computer software can do little to help in clarification of the search, as it just provides a list of queries matching the data entered by the user. As S. Walter and M. Pinkal note, sometimes it is possible to use Boolean operators (AND, OR, NOR) in addition, but this again has to be done by the user.<sup>11</sup> In sum, we shall be very precise with our search queries, otherwise we will either get an enormous list of unwanted and irrelevant documents or an empty result.

The first of these situations leads to an ‘overload’ of irrelevant information in the Internet.<sup>12</sup> A user will get hundreds of documents with only several of them really having relevance to his query. A typical example is the search for ‘appeal proceedings’ where the person intends to find out their special features in civil procedure. However, the system may present in response documents that relate to criminal procedure, out-of-court dispute resolution, or the information on the correct area of law, but of totally different jurisdiction (e.g. Italy instead of France). The latter situation may occur due to several reasons. Some pages do not appear in the results as they do not match the keywords entered by the user or are used in a different form. Yet other places are hidden, being located in the deep Web or some professional network. A number of sources are totally impenetrable and untraceable – like images, audio/video files, zipped archives, databases information, scanned PDF documents.

Modern searching engines are also not very helpful when we need something beyond the mere text of the document. Lawyers, for example, may be interested in the following related information:

- Which edition of law [L] of [yy.mm.dd] was in force on [yy.mm.dd]?
- Which law would recognize the contract between the parties [A] and [B] as valid?
- Which law governs the legal status of company [A], a [LLC]?
- Which remedies against a [businessman] may a [consumer] use in [country A]?
- Which laws are changed or annulled by the law [X]?

As you may imagine, it is hardly possible to ask question like these to a general-purpose searching engine. Although some of them may be answered by carefully reading the text, this is not something that can be presented to us at once by the search engine.

We must also remember that users are interested in particular norms, rather than full documents. Thus, people wishing to conclude a contract for the lease of dwellings (Chapter

---

<sup>11</sup> S. Walter and M. Pinkal, ‘Definitions in Court Decisions – Automatic Extraction and Ontology Acquisition’ in: ‘Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood’ (Amsterdam, IOS Press 2008) 95.

<sup>12</sup> A. Lamandini, ‘Semantic Web’ [2011] 6(1) Ricerche di Pedagogia e Didattica 2-3.

XXXI of Lithuanian Civil Code<sup>13</sup>) will be interested in norms related to that contract and some general rules for contract law, but not the other parts of the instrument.

A big problem is also lack of real connection between separate documents in the Web. Hyperlinks are intended to be used by the humans, not the machines. That is why a broken link (a well-known 404 error) presents a fatal thing that a computer cannot fix on its own.

Consequently, the search phase is only a first step in the intellectual work of a law professional – the next stage is processing and analyzing of these results where the person in charge picks necessary documents in order to apply them in his case. The last (third) stage will be actual reasoning with the help of information duly obtained.<sup>14</sup> It must be observed that computers only help users during the first stage, while the second and the third are totally left to human operators. They have to invest sufficient time and effort in completing these tasks which until now was almost taken for granted.

The reason the computers do so little to assist humans lies in the nature of the Web. Its main tool is the Hypertext Markup Language (HTML) which describes in its elements (known as *tags*), how the page shall be structured, e.g. which part of it must be treated as a heading, paragraph, image, embedded video and so on. However, this information only instructs user browser how to display content. Meanwhile, it does not help computers to understand the *contents* of a page. Whether the text displayed constitutes a draft version of the Declaration of Independence of the USA, a binding EU Regulation on Data Protection (GDPR) or an apple pie recipe – the machine will just treat it as a sequence of zeroes and ones. Computer programs and algorithms cannot share and perform tasks on such data. The only thing they can help us with is the actual search, but even here they are far from perfect.

Of course, there is a number of commercial databases and web services specifically designed for legal researchers. They have more 'intellectual' searching tools and the possibility to perform advanced searches, indicating authority, date of adoption of the legal instrument, its type and serial number, etc. It can give a better result when compared with mainstream solutions and even provide for a possibility to answer some of the questions specified above. However, there are also many important drawbacks. Firstly, such databases are for professional use only and even lawyers need preliminary practice before mastering them. This cannot be compared to general-purpose searching tools that are familiar to all Internet-users. Secondly, the databases are paid, which limits access to them of general society, including unemployed, migrant workers, poor and other socially vulnerable people. Thirdly, they have significant jurisdictional limitations. Thus, Westlaw and Lexis only provide data on major western jurisdictions. Such systems as 'Garant' and 'Consultant Plus' contain great collection of Russian and Post-Soviet laws, but are practically unknown elsewhere in the world and are never used in reasoning there. Moreover, one of the limitations of proprietary databases is their reliance on patented technology which makes it almost impossible to add new blocks without the agreement of the right holder. Consequently, they remain limited to what has been included to the database by the company's employees.

## 2. Semantic Web and its solutions for the legal sector

---

<sup>13</sup> Lietuvos Respublikos Civilinis Kodeksas (Civil code of the Republic of Lithuania) [2000] Valstybės žinios 74-2262.

<sup>14</sup> G. Shankhdhar, V.K. Singh and M. Darbari, 'Legal Semantic Web – a Recommendation System' [2014] 7(3) International Journal of Applied Information Systems 31-32.

One of the possible solutions to the problems named above is the progressive development of what is known as Semantic Web. It is not considered a different kind of Internet, but rather an extension of the latter. While the original Web was known as 'Web of documents', this one presents 'Web of data'.<sup>15</sup> In essence, it is a project of World Wide Web Consortium (W3C) to promote common data formats and exchange protocols. The general idea is that the Internet can be used to share not only the information, but also the *meaning*. In that way web pages and other materials distributed over the network will become understandable by the machines, and the latter will gain ability to perform various functions upon them (read, process, analyze, store, compare, use in equations and deliver them to end-users).<sup>16</sup> This project seems to partially solve the problem of information overload in the Internet, as it assigns a more active role to computers, allowing them not only to conduct a more thoughtful and thorough search, but also to assist lawyers and other seekers of legal knowledge at a later stage (analysis and reasoning). Full implementation of the semantic technologies will lead us to Web 3.0, a next stage in its progressive evolution.

The central concept of the project is '*semantics*', which in linguistics stands for an area of study devoted to the meaning of words and phrases. Currently, the Web provides meaning only to the human-reader, not to the machine. The latter needs an instruction in a formalized, instead of a natural language. Such languages, indeed, have been created and implemented in different sectors.

The first example is eXtensible Markup Language (XML), which complements classic HTML and also relies on *tags*.<sup>17</sup> Unlike the latter, this language does not influence the way humans see the page, but provides commands for the computers. Its tags refer to objects, persons and relations between them. It may include such statements as <author>, <title>, <year>, <price> and many others. Moreover, it allows nesting, i.e. one element can be placed within the other, thus becoming its 'child' and acquiring relation with its 'parent'. With the above-mentioned tags that may happen when all of them are united together under a common tag of <book>. A great feature of XML language is that it allows users to define tags of their own, so that they can take advantage of all the peculiarities of their informational needs. We clearly see that XML may be used to describe legal concepts – parliamentary statutes (which all have their titles, reference numbers, sections and subsections and, of course, valid and outdated redactions). XML may also describe people (parties to a contract), places (jurisdiction, country of origin of a product) and more abstract things. At the same time, this language just puts meta-annotations to existing blocks of text – it does not do anything with them on its own.<sup>18</sup> In order to perform certain tasks on the data, there must be an additional application instructed to use it in a prescribed way.

Another important technology of Semantic Web is Resource Description Framework (RDF). It is a methodology, or a data model, for conceptual description or modelling of information in the Web through the use of different syntax notations and data serialization formats. In essence, it is used to describe objects (resources) and relations between them. This one indeed helps to catch semantic meaning of the data and represent it in an XML-based syntax. It also relies on specialized vocabularies defined by the users. This is done with the help of special RDF Schemas. RDF allows making statements about various

---

<sup>15</sup> Semantic Web [2018]: <https://www.w3.org/standards/semanticweb/>

<sup>16</sup> A. Patel and S. Jain, 'Present and Future of Semantic Web Technologies: a Research Statement' [2019] International Journal of Computers and Applications (<https://doi.org/10.1080/1206212X.2019.1570666>).

<sup>17</sup> C. Fong, 'What is the Semantic web?' [2011] 30 TALL Quarterly 13.

<sup>18</sup> Introduction to XML [2019] [https://www.w3schools.com/xml/xml\\_what\\_is.asp](https://www.w3schools.com/xml/xml_what_is.asp)



concepts in expressions of the form *subject-predicate-object* (known as *triples*). Such statements may be found in most legal texts, including Art. 25 of Lithuanian Constitution: “[e]veryone shall have the right to have his own convictions”.<sup>19</sup> “Everyone” here is a subject, “shall have the right to have” is a predicate and “his own convictions” is an object. This statement in its formalized form may be understood by the software agent, which draws the relationship between the subject and the object. It must be noted that RDF is not intended for the human eye, just for the machine.

The last part of Semantic Web is Web Ontology Language (OWL) which is used to standardize the definition of real-world concepts. The main notion here is that of ontology, a term coming from philosophy and meaning identification in the most general terms of the kinds of objects that virtually exist and ways of their description. An ontology is an explicit and formalized specification of conceptualizations.<sup>20</sup> OWL can help to describe properties and classes, relations between classes (e.g., disjointness, cardinality, equality, symmetry). The relationships also need to include hierarchy of classes. In practice, ontology shall consist of a finite list of terms and the relationships between them. These terms have to be the most important concepts of a given area. The core notions in Law include: norm, case, person, agent, role, status, responsibility, property, etc. What is important here is that the general notions established by the ontology are shared for the whole domain and will still do their job even if different terminology is used.

The following example from Civil Procedural Law<sup>21</sup> demonstrates application of the named technologies to the legal sphere:

```
<p>Natural and legal entities may be parties to a civil procedure: plaintiff
<!DOCTYPE rdf:RDF [<!ENTITY law "http://domain.tld/otherpath/law#" >]>
<owl:Class rdf:ID="Plaintiff">
<rdfs:subClassOf rdf:resource="#Person"/>
</owl:Class> <owl:inversOf>
<owl:ObjetcProperty rdf:ID="Defendant"/>
</owl:inversOf>
or defendant
<!DOCTYPE rdf:RDF [<!ENTITY law «http://domain.tld/otherpath/law#» >]>
<owl:Class rdf:ID="Defendant">
<rdfs:subClassOf rdf:resource="#Person" />
</owl:Class>
<owl:inversOf>
<owl:ObjetcProperty rdf:ID="Plaintiff"/>
```

<sup>19</sup> Lietuvos Respublikos Konstitucija (Constitution of the Republic of Lithuania) [1992] Valstybės Žinios 33-1014.

<sup>20</sup> T. Gruber, ‘A Translation Approach to Portable Ontology Specifications’ [1993] 5 Knowledge Acquisition 199.

<sup>21</sup> Lietuvos Civilinio Proceso Kodeksas (Code of Civil Procedure of Lithuania) [2002] Valstybės žinios, 36-1340 (Art. 41(1)).

</owl:inversOf>

</p>

This code makes explicit statements about 'plaintiff' and 'defendant', marking them up for the software agents to be noticeable, and showing to which class the two belong and what is their relation towards each other. There is no doubt that in practice it will be necessary to perform more complex work by trying to establish meaningful connections between more than two persons and objects in real life (companies, state institutions, rights and legal titles, etc.).

Together the three named components will enable more effective structuring, publication and referencing to legal documents and consequently will save time and money of those working with the legal data. Semantic technologies will enable cooperation between the computer and human beings at all stages of the legal research. There may be two possible scenarios for such work: interaction between machine and human (M2H), where the computer program will help to clarify the searching issues and present the most relevant result. Another option is machine-to-machine (M2M) operation, where separate programs (software agents) will exchange data to solve a certain problem.

In practice we see some specific projects, developing semantic technologies for the legal sector. One of the most prominent in Europe is the Legal Knowledge Interchange Format (or, LKIF), which was developed by ESTRELLA project and intends to establish uniform standards for representing and interchanging data on law, judicial cases and governmental policy.<sup>22</sup>

Another feature of the Semantic Web is its intended decentralization: there are clear intentions to move away from quasi-monopolistic position of Google and Facebook to the status quo where there is no need to depend on any major organization for the Web to function properly.<sup>23</sup> Private parties will not only create content (as it happens in Web 2.0), but also fill it with meaning. Such concept as Linked Open Data helps to implement this desire as it allows to publish structured data and provide for cross-references through hyperlinks. Each resource has its unique Uniform Resource Identifier (URI) and can be accessed from outside and contain its own links to other resources and objects.

### **3. Prospects of the semantic technologies**

As was noted from the very beginning, the first thing where Semantic Web can do its job is the searching mechanisms. With content enhanced by semantic annotations it will be easier for machines to understand search queries and present an optimal result. In fact, specialists hope that instead of getting a set of hyperlinks to follow in response for your question typed into a search box, computers will be ready to propose a final and definite answer.<sup>24</sup> In many situations significant computing power will have to be used in order to perform a statistical query. You may ask the computer to find out how many states still retain the death penalty, or where in the world you will get the highest fine for driving drunk. These examples may present a special interest to academics, but other law professionals also may benefit from them. Thus, a judge can consult a specialized database in order to get an

---

<sup>22</sup> Th. Gordon, 'An Overview of the Legal Knowledge Interchange Format' in: 'Business Information Systems Workshop' (Berlin, Springer 2010) 240-241.

<sup>23</sup> J. Mailland, 'The Semantic Web and Information Flow: a Legal Framework' [2010] 11(2) North Carolina Journal of Law and Technology 269.

<sup>24</sup> E. Francesconi, *Ibid.*, 49.

insight of which punishment to choose for the criminal, by comparing his case with previous convictions for similar offences. Attorneys will be able to better help their clients if all the details of the problem will be taken into consideration and a full list of applicable statutes, regulations and case-law will be delivered by a software agent.

The other important thing is to keep the existing databases up-to-date, so that you will get the law in force instead of some old and irrelevant edition. However, the outdated versions shall not be simply put aside as in many instances they present an interest for the lawyer (the solution of the case may depend on the act that was in force on a certain date). For that reason, it must be easy enough to get a correct redaction just by giving the computer an exact date. Again, this may be done only by supplementing the relevant document with the necessary time tags.

Semantic Web can also contribute to the establishment of more advanced data-management systems. Since RDF and OWL permit us to define subjects and assign various roles to them, it becomes possible to determine who has access to particular documents and their parts. This is quite promising for working with classified information and even trade or state secrets, so that only a duly authorized person may view the corresponding file. Another available feature is permission to modify and update content. If that is given to a limited number of actors, it increases the credibility of the system and allows other users to rely on the data contained therein. As we may access a certain legal rule from any application using semantic technologies, it is extremely relevant for us, where does this rule come from. Where the author and source are encoded together with the rule, we have no reasons to worry and question its authority.

Some other ways in which Semantic Web can be useful are provided in legal literature. Thus, P. Casanovas et al. speak about an ontology-based decision support system for young judges (*luriservice*). Its necessity is justified by a number of problems that inexperienced judges of their first appointment may face while performing their functions. They have numerous questions, most of which are of para- or meta-legal nature (Which procedures to follow? How to document an interim decision?). Most of them may be answered by their more experienced colleagues. However, the latter do not usually have enough time to do that, or may already be retired. Consequently, *luriservice* tries to combine all the wisdom in a sort of database with restricted access. The judges will be able to ask direct questions in natural language and get the response in it from the machine. Behind the curtain is a complex process of finding and matching the relevant question with the most probable and appropriate answer.<sup>25</sup>

A semantic approach to copyright management is presented by R. García and R. Gil. In their article the authors speak of a system that would help people establish the copyright conditions for their content. OWL-based technologies would allow checking if a certain action is granted by a specific license, as well as incorporating penalties for copyright violation directly into the system.<sup>26</sup>

Within a sector of private law software agents can look through a vast collection of previously drafted contracts and propose a set of terms that will mutually benefit the two parties entering into a particular relationship. It will do so basing its solution on law and

---

<sup>25</sup> P. Casanovas, N. Casellas and J-J. Vallbé, 'An Ontology-Based Decision Support System for Judges' in: 'Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood' (Amsterdam, IOS Press 2008) 165.

<sup>26</sup> R. García and R. Gil, 'Copyright Licenses Reasoning using an OWL-DL Ontology' in: 'Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood' (Amsterdam, IOS Press 2008) 161-162.

regulations, previous practice and usages. In certain situations of B2B relationships the contracts can even be concluded automatically if certain conditions are satisfied.

Finally, semantic technologies may be important for online dispute resolution. They can both assist the judge in reaching a decision on the matter and do the job instead of him in an automatic fashion. While a human in charge of resolving a dispute may be able to consult tens or hundreds of cases, the machines can do much better. The only problem is that modern judgments and arbitration awards do not usually have the most suitable form. They all have similar structure, but the reasoning part, where the competent authority cites legal norms and applies them to the exclusion of the others is more creative and cannot be as easily reduced to simple equations. That is why semantic technologies can help solve all the nuances of a dispute if it is a repeating and standard problem (late payments, non-enforcement of monetary obligations, etc.), but will be of less importance in cases with wide judicial discretion where decision is rather based on intuition than an analysis of previous practices.

#### **4. Challenges for the Semantic Web**

Common problems include vast amount of data that is not already in relevant formats. There are also many imprecise concepts in law, such as 'unpleasant emotional experiences' (which are elements of non-pecuniary damage in Civil Law)<sup>27</sup> or 'impeccable reputation' as a precondition for a person to become a judge.<sup>28</sup> In fact, law is not 'black and white' and has many nuances. Quite difficult will be to conceptualize the fundamental principles (such as 'fair trial', 'equality', 'rule of law') which are rather abstract and subject to controversy in doctrine as well as in practice when it comes to their precise meaning.

Another big problem is that there is no universal standard for development. Instead, several projects are run side-by-side in different institutions and they are poorly coordinated. What happens is that their end results suit only marginal policy goals and do not change the whole picture. At the same time, it may be presumed that such situation is common for transition and experimentation period and in the end, we will combine all the best solutions into the uniform standards of future Web 3.0 (or, even 4.0).

A frightening issue was revealed by J. Mailland, according to whom the situation where the Semantic Web is built from the bottom-up (by private individuals putting tags on the information) gives more opportunities to representatives of the Western nations. They are in a better position to impose familiar linguistic, cognitive and ideological frames as universal norms, while the rest of the world may be left behind.<sup>29</sup> The same author also fears the increased possibilities of censorship when all the information in the Web is 'labelled' with semantic tags.<sup>30</sup>

#### **Conclusions**

---

<sup>27</sup> Lietuvos Respublikos Civilinis Kodeksas (Civil Code of the Republic of Lithuania) [2000] Valstybės žinios 74-2262 (Art. 6.250).

<sup>28</sup> Lietuvos Respublikos Teismų įstatymas (Law on Courts of the Republic of Lithuania) [1994] Valstybės žinios 46-851 (Art. 22).

<sup>29</sup> J. Mailland, *Ibid.*, 285.

<sup>30</sup> *Ibid.*, 286.

As was shown in this paper, Semantic Web is a promising solution to the problems faced by law professionals and ordinary citizens while dealing with information overload in the vast expanses of the Internet. It is not as easy as it may seem, since requires additional and complicated work on establishing relations between concepts and ontology-based vocabularies. A big dilemma is that lawyers generally are not perfect at Web-technologies, while IT specialists know little about fundamental legal concepts and the relations between them. Another significant issue is the lack of uniformity with the development of Semantic Legal Standards for the Web. Akoma Ntoso, NormInRete, LexML, MetaLex, LexDania and the others function independently (although consult each other from time to time), pursue different goals and used their own modifications of the standard XML/RDF/OWL technologies. In the long run there is a possibility of their integration, but now it seems that each of the projects is only capable of solving small problems identified by their leaders. At the same time, the very idea of introducing legal meaning to the world of the Internet and making the machines smart enough to help us with daily routine and really complex cases is brilliant and requires universal support.

## **Bibliography**

### **Articles:**

1. Bascik T., 'Web sémantique: Quelles perspectives pour l'information juridique?' [2014] 19(1) Lex Electronica.
2. Breuker J., Casanovas P. et al., 'The Flood, the Channels and the Dykes: Managing Legal Information in a Globalized and Digital World' in: 'Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood' (Amsterdam, IOS Press 2008).
3. Casanovas P., Castellás N. et al., 'An Ontology-Based Decision Support System for Judges' in: 'Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood' (Amsterdam, IOS Press 2008).
4. Casanovas P., Casellas N. and Vallbé J-J., 'An Ontology-Based Decision Support System for Judges' in: 'Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood' (Amsterdam, IOS Press 2008).
5. Fong C., 'What is the Semantic web?' [2011] 30 TALL Quarterly.
6. Francesconi E., 'On the Future of Legal Publishing Services in the Semantic Web' [2018] 10(6) Future Internet.
7. Francesconi E. and Peruginelli G., 'Integrated Access to Legal Literature through Automated Semantic Classification' [2008] 17 Artificial Intelligence and Law.
8. García R. and Gil R., 'Copyright Licenses Reasoning using an OWL-DL Ontology' in: 'Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood' (Amsterdam, IOS Press 2008).
9. Gordon Th., 'An Overview of the Legal Knowledge Interchange Format' in: 'Business Information Systems Workshop' (Berlin, Springer 2010).
10. Gruber T., 'A Translation Approach to Portable Ontology Specifications' [1993] 5 Knowledge Acquisition.

11. Lamandini A., 'Semantic Web' [2011] 6(1) Ricerche di Pedagogia e Didattica.
12. Mailland J., 'The Semantic Web and Information Flow: a Legal Framework' [2010] 11(2) North Carolina Journal of Law and Technology.
13. Mo Jang S. and Kim J.K., 'Third Person Effects of Fake News: Fake News Regulation and Media Literacy Interventions' [2018] 80 Computers and Human Behavior.
14. Patel A. and Jain S., 'Present and Future of Semantic Web Technologies: a Research Statement' [2019] International Journal of Computers and Applications (<https://doi.org/10.1080/1206212X.2019.1570666>).
15. Shankhdhar G., Singh V.K. and Darbari M., 'Legal Semantic Web – a Recommendation System' [2014] 7(3) International Journal of Applied Information Systems.
16. Walter S. and Pinkal M., 'Definitions in Court Decisions – Automatic Extraction and Ontology Acquisition' in: 'Law, Ontologies and the Semantic Web: Channeling the Legal Information Flood' (Amsterdam, IOS Press 2008).
17. Yu L., 'A Developer's Guide to the Semantic Web' (Berlin: Springer 2014).

**Legal acts:**

18. Federal Law of the Russian Federation no. 289-FZ [2011] [http://www.consultant.ru/document/cons\\_doc\\_LAW\\_120715](http://www.consultant.ru/document/cons_doc_LAW_120715).
19. Lietuvos Respublikos Konstitucija (Constitution of the Republic of Lithuania) [1992] Valstybės Žinios 33-1014.
20. Lietuvos Respublikos Civilinis Kodeksas (Civil code of the Republic of Lithuania) [2000] Valstybės žinios 74-2262.
21. Lietuvos Civilinio Proceso Kodeksas (Code of Civil Procedure of Lithuania) [2002] Valstybės žinios, 36-1340.
22. Lietuvos Respublikos Teismų įstatymas (Law on Courts of the Republic of Lithuania) [1994] Valstybės žinios 46-851.

**Internet sources:**

23. How Big is the Internet? Live Counter (<https://www.live-counter.com/how-big-is-the-internet>).
24. Introduction to XML [2019] [https://www.w3schools.com/xml/xml\\_what\\_is.asp](https://www.w3schools.com/xml/xml_what_is.asp)
25. Semantic Web [2018]: <https://www.w3.org/standards/semanticweb/>