

Erich Schweighofer

## Structuring Textual Feedback of E-Participation Platforms

---

In the knowledge and network society, democracy gets a strong digital element. Whereas strong reservations exist concerning e-voting, e-participation with its high degree of flexibility is much more accepted. In the E-Participation project of the KIRAS research programme, an e-participation platform demonstrator was developed. Textual feedback was so far analysed manually. In case of more than several thousand statements, text analysis tools can provide important support. Based on the example of legal data analysis, an outline of such an additional tool of e-participation platforms is presented.

---

Category: Articles

Region: Austria

Field of law: E-Democracy

Citation: Erich Schweighofer, Structuring Textual Feedback of E-Participation Platforms, in: Jusletter IT 22 September 2016

## Contents

1. Introduction
2. E-Partizipation project
3. Challenge of textual feedback of e-participation platforms
4. Legal data science: 8 views, 4 methods and 4 synthesis of the legal system
5. AI & law methods for e-participation platforms
6. Outline of an analysis of textual feedback of e-participation platforms
7. Conclusions and future work

## 1. Introduction

[Rz 1] We are living in the knowledge and network society.<sup>1</sup> We are not only moving to a «digital economy», where exchange of goods and services depends on access to the internet but also to a «digital democracy» called e-democracy.<sup>2</sup> Access to the internet, communication, intensive use of data, information and knowledge are the major element of these developments. Social life is more and more happening in networks determined by Information and Communication Technologies (ICT)<sup>3</sup>.

[Rz 2] From the perspective of human rights, access to the internet is a key element for enabling a digital life, to use the rights to free speech, expression, communication, regardless of frontiers, peaceful assembly, to knowledge and understanding and the right to participate in the conduct of public affairs, to vote and to be elected.

[Rz 3] With the internet, a fundamental change takes place in democracy: existing restrictions to the agora are disappearing. No physical meeting place is necessary; people can deliberate being anywhere. The internet provides the platform, e.g. the agora, allowing the exchange and sharing of expressions and communications, establishing a forum of deliberation and consensus building. A new era of political self-determination is possible and emerging that may become more and more a reality in the future.

[Rz 4] Until then, many technical, legal and societal obstacles have to be solved for achieving this idealist goal. Whereas technical solutions are developing and emerging, legal and societal constraints seem to grow, especially a fear of technology that cannot be easily controlled. As *Winston Churchill* said: «No one pretends that democracy is perfect or all-wise. Indeed it has been said that democracy is the worst form of Government except for all those other forms that have been tried from time to time [...]»<sup>4</sup> Thus, risks of failures of democratic procedures are common, a persistent threat that can be solved only by a risk analysis *ex ante* and *ex post*.<sup>5</sup> Democracy

---

<sup>1</sup> AHTI SAARENPÄÄ, *The Digital Lawyer. What skills are required of the lawyer in the network society?* In: Erich Schweighofer/Franz Kummer/Walter Hötendorfer (eds.), *Kooperation / Co-operation*, Tagungsband des 18. Internationalen Rechtsinformatik Symposions IRIS 2015. Wien: Österreichische Computer Gesellschaft (OCG), p. 73, 2015.

<sup>2</sup> Cf. the numerous workshops at the IRIS conferences organised by ALEXANDER PROSSER ET AL. or the CeDEM conferences at the Donau Universität Krems.

<sup>3</sup> Cf. proceedings of IRIS 2016 on networks: ERICH SCHWEIGHOFER/FRANZ KUMMER/WALTER HÖTZENDORFER/GEORG BORGES (eds.), *Netzwerke / Networks*, Tagungsband des 19. Internationalen Rechtsinformatik Symposions IRIS 2016. Wien: Österreichische Computer Gesellschaft (OCG), 2016; ERICH SCHWEIGHOFER, *Von der Wissensrepräsentation zum Wissensmanagement im e-Government*. In: Erich Schweighofer et al. (eds.), *IT in Recht und Staat, Aktuelle Fragen der Rechtsinformatik 2002*, Verlag Österreich, Wien, p. 85, 2002.

<sup>4</sup> Speech in the House of Commons, 11 November 1947.

<sup>5</sup> Cf. THOMAS PREISS, *Die Bedeutung der Risikoanalyse für den Rechtsschutz bei automatisierten Verwaltungsverfahren*, Ph.D. thesis, University of Vienna; ERICH SCHWEIGHOFER/THOMASPREISS, *Risikoanalyse im*

must face risks of manipulation of information exchange, deliberation, and vote. It is also not guaranteed that this form of crowd intelligence produces wise results but it is a good chance that representatives are elected taking care of the interest of the people. Very formalistic approaches of detailed legislation and the requirement of a near zero error rate seem to lead to a dead end of no valid elections at all.<sup>6</sup> It is evident that informational tools of democracy have to comply with constitutional law. Present strong reservations strongly restrict the use of such digital elements of democracy. This has to be accepted but communication work has to be done to overcome this obstacle to digital democracy.

[Rz 5] E-democracy is the use of ICT in processes of politics and governance (e.g. policy making, decision making, governmental services, administrative powers etc.).<sup>7</sup> E-voting (or electronic voting) means voting by electronic means. Besides working IT solutions, e-voting is struggling getting legal and societal acceptance. The image of a – at least partly – law-less internet – still hinders efforts in this direction; some failures are strong arguments for the many opponents to e-democracy. Its paper variant, letter voting, gets much higher acceptance but faces similar difficulties in fulfilling constitutional requirements on voting.<sup>8</sup> E-Participation is here understood as ICT support of participatory processes in politics and governance, but much more closely related to e-government. A new quality level of two-way interaction is created between citizens and government. E-participation has the advantage of flexibility. Its main element, the electronic communication platform, is already highly accepted in society considering the strong use of social networks. For politicians, the internet, in particular the social networks, has got a similar importance as traditional communications. It is worth noticing that the strongest significance of this new form of e-democracy without particular tools could be seen in recent democratic movements against authoritarian regimes.<sup>9</sup>

[Rz 6] E-participation means a universal e-participation platform that must include all new media via open and interoperable interfaces, electronic identification and top-class security. Social networks may be strongly supportive, but in the very end cannot comply with these particular legal standards. Therefore, the following questions have to be addressed: authentication and authorisation of all users of the system, the protection of the originality (authenticity) and the confidentiality of the surveyed and processed data, e-inclusion and information security. In recent years, significant improvements of e-participation have been noticed (e.g. designing an architecture, defining software engineering tools and methods, choosing data formats and protocols). E-participation platforms must be highly robust, very usable for the e-citizen and have to respect rules on democratic participation but also privacy.

---

Recht – eine neue juristische Methodik. In: Heinrich C. Mayr/Martin Pinzger (eds.), INFORMATIK 2016, Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2016, in print.

<sup>6</sup> Cf. judgement of the Austrian Constitutional Court WI6/2016 of 1 July 2016. The decision can be found in the Austrian Legal Information System RIS, <http://www.ris.bka.gv.at>- This judgement is strongly criticized by HEINZ MAYER, Eine klare Fehlentscheidung. In: Falter 34/16, Wien 2016.

<sup>7</sup> NATALIE C. HELBIG/J. RAMON GIL-GARCIA/ENRICO FERRO, Understanding the Complexity of Electronic Government: Implications from the Digital Divide, Proceedings of the Eleventh Americas Conference on Information Systems, Omaha, NE, USA, 2005.

<sup>8</sup> Cf. a warning of the IT expert ARNIM RUPP, IT-Experte warnt vor Betrug bei der Briefwahl, in: Der Spiegel Online, 28 July 2013, <http://www.spiegel.de/spiegel/vorab/it-experte-warnt-vor-betrug-bei-der-briefwahl-a-913466.html> (all websites lastly accessed: 19 September 2016).

<sup>9</sup> The best example is the Arab Spring from 2010 to 2012, in particular in Egypt, Tunisia, Libya, Yemen, Syria and Bahrain. The Wikipedia article gives an overview: [https://en.wikipedia.org/wiki/Arab\\_Spring](https://en.wikipedia.org/wiki/Arab_Spring).

[Rz 7] Besides giving an overview of results of the E-Partizipation project, the question of analysis of textual feedback of e-participation platforms is addressed. Taking into account research results of legal data analysis that works on similar text corpora, in particular digital court files, an outline of such a system is presented.

## 2. E-Partizipation project

[Rz 8] The E-Partizipation project within the KIRAS Security Research programme of the Austrian Research Promotion Agency (FFG) has developed an E-Participation Platform Demonstrator that can strengthen direct democracy offering different participation scenarios. The consortium consists of AIT Austrian Institute of Technology GmbH, Digital Safety & Security Department, as project leader, and the following project or co-operation partners: Bundeministerium für Inneres (Federal Ministry of Interior), Donau-Universität Krems, Österreichische Staatsdruckerei GmbH, rubicon IT GmbH and Universität Wien, Arbeitsgruppe Rechtsinformatik. The project started in October 2014 and will end in October 2016. For more information, we refer to the project website.<sup>10</sup>

[Rz 9] Presently, no sufficient legal basis for e-participation in Austria exists, in particular concerning procedures, legal value and identification. Authorities may use it as a tool for participation without formal legal relevance that is not satisfactory and motivating for citizens. Examples of other European countries on e-participation applications like the e-voting solution in Estonia, the PIN-TAN procedure in Switzerland or municipal projects like the «Bürgerhaushalt» (local budget) in Cologne have been studied. Four different scenarios of participation were assessed from a legal and societal perspective. Recommendations for legal amendments and improvements were also developed.

[Rz 10] The legal analysis and the requirements of e-participation are the pillars of orientation of the developed E-Participation Platform Demonstrator. It supports different levels of participation (information, consultation, cooperation, co-decision-making) in a very user-friendly way taking into account electronic identification (eID) and social inclusion. Two components have to be mentioned in particular: Identity Anonymizer and e-participation platform.

[Rz 11] A major part of the project are eID and authentication procedures. The different forms and its different security levels like user name and password, mobile signature, citizen card («Bürgerkarte») and electronic passport with biometric tokens were analysed and classified. A strong focus was given to the existing Austria citizen card for E-Government. So far, E-Partizipation is the only project available taking into account all different eID profiles in Austria or other European countries. The anonymization procedures for citizen participation were also analyzed.

[Rz 12] E-participation requires the use of highest security standards for the protection of privacy as sensible data about political opinions and voting behaviour are collected. Thus, the principle of Privacy by Design (PbD) is getting more and more important. This principle implies addressing privacy and data protection during the entire technology lifecycle.<sup>11</sup> In a software development

---

<sup>10</sup> Website E-Partizipation <https://www.epartizipation.info/>.

<sup>11</sup> European Union Agency for Network and Information Security enisa, Privacy and Data Protection by Design – from policy to engineering. 2014, [https://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/privacy-and-data-protection-by-design/at\\_download/fullReport](https://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/privacy-and-data-protection-by-design/at_download/fullReport); OLIVER TERBU/WALTER HÖTZENDORFER/MARIA LEITNER/ARNDT BONITZ/STEFAN VOGL/SEBASTIAN ZEHETBAUER, Privacy and Security by Design im agilen Soft-

process, privacy and data protection have to be part of the whole process. The new General Data Protection Regulation (GDPR) of the EU enhances this principles, in particular data minimisation.<sup>12</sup>

[Rz 13] A further important aspect was e-inclusion. Conceptualisation of the platform and the inherent design principles («Design for All») aiming at enhancing the inclusion of people with disabilities such as visual, auditory, physical, cognitive, learning and neurological disabilities as well as non-native speakers and elderly people were examined and identified as a central aim for official e-participation projects.<sup>13</sup>

[Rz 14] Two scenarios were evaluated in practice with focus groups: works council elections and citizen's participation on the use of a public space in Vienna. Users were highly satisfied with its friendly interface and the level of participatory elements.

[Rz 15] The platform was also presented and discussed in interviews with members of the Austrian Parliament and high-level representatives of the administration. As e-participation is mostly an activity sponsored by the respective authority, or a party, or a member of parliament or another council, the possible use of familiar logos and colours was appreciated. Self-chosen personal identification in e-participation seems to be good solution, allowing respect of privacy but also self-chosen visibility. The secrecy of vote is strongly respected. The organiser of a participation platform must get sufficient information about the electorate of the participation process for accessing its value as a democratic participation process. This tricky question must be solved by balancing between openness of participation processes and privacy.

[Rz 16] The following applications seem to be promising in the future: city planning, participatory budgeting, structured discussion about politics, voting (legally non-binding), outcome focused discussions (e.g. about Codes of Conduct), works council elections in companies and issue focused voting in parties, NGOs, but also citizen movements etc. In most of these scenarios, the organiser of the participation process is not legally obliged to accept the result of this e-participation process. After some experience in the next years, sufficient practice may exist that allows developing principles and rules that guide the e-participation organiser but also the respective redress procedures.

[Rz 17] As result of the participation process, e-participation platforms produce statistical data but also textual (and visual) statements. An opinion, a vote, but also with reasons: a statement that overcomes the rigid limitations of present voting procedures, a more or less binary choice. This new complexity is an unsolved challenge of assessment of the will of the people. Thousands of solutions may be submitted, with variable expressions of support or amendments. Whereas many tools are available for statistical analysis, no proper tools exist for text analysis. Within the project tests, the text corpus was sufficiently small to allow a manual assessment of the textual feedback. However, if several thousands, ten-thousands or even hundred thousands are partici-

---

wareentwicklungsprozess. In: Erich Schweighofer/Franz Kummer/Walter Hötzendorfer/Georg Borges (eds.), *Netzwerke / Networks*, Tagungsband des 19. Internationalen Rechtsinformatik Symposions IRIS 2016. Wien: Österreichische Computer Gesellschaft (OCG), p. 457, 2016.

<sup>12</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4 May 2016, p. 1, Articles 5 and 25 GDPR.

<sup>13</sup> VINZENZ HEUSSLER/JUDITH SCHOSSBÖCK/JANOS BÖSZÖRMENYI, *Aspekte der Inklusion aus Sicht der E-Partizipation*. In: Erich Schweighofer/Franz Kummer/Walter Hötzendorfer/Georg Borges (eds.), *Netzwerke / Networks*, Tagungsband des 19. Internationalen Rechtsinformatik Symposions IRIS 2016. Wien: Österreichische Computer Gesellschaft (OCG), p. 213, 2016.

pating, often with more than one statement, manual analysis without ICT supports seems to be a nearly impossible task. Here, ICT can provide a tool strongly supporting an objective assessment process. The research results on analysis of similar text corpora in law are taken as a start. This paper summarizes results of the workshop and related discussions with the Hebrew University Jerusalem and Bar Ilan University, Ramat Gan.

### 3. Challenge of textual feedback of e-participation platforms

[Rz 18] Textual feedback refers to any kind of statements, e.g. written text or a visual sign submitted in e-participation platforms. It consists of highly unstructured texts, not very long, but often strongly interconnected. In length, structure, language etc., the feedback is quite similar to comments on newspaper articles. The quality of statements is very diverse, it can be specific, detailed, well founded, in good or bad language, in standard language or also in dialects or highly specialised language. Very often, the text is strongly emotional to the level of being un-objective or even offending. The quantity can be enormous. A realistic scenario may be an e-participation process with more than 100'000 participants with up to 10 comments meaning about 500'000 statements. Given the time, this text corpus can be read and analysed. It is happening now with participatory processes without a two-tier interaction (e.g. European Commission). However, in e-participation, the platform should be interactive and dynamic, allowing all to follow the discussion. Citizens may read some statements, even up to 1'000 statements in exceptional cases, but then it starts to become tricky, arbitrary and unbalanced. Therefore, IT support for this task is strongly needed in future e-participation platforms.

[Rz 19] So far, no research exists concerning these text corpora in e-participation platforms. However, related research gives importance guidance how to move on in this direction, to name the most important: natural language processing, e-discovery<sup>14</sup> and legal data analysis. In this paper, the focus will be on legal data analysis. As a first step, adding semi-automatically metadata to these statements seems to be most promising. Such metadata allows the structured and graphical representation of the participation process in topical maps. A more detailed analysis of the content of the texts faces the high challenges of a more powerful natural language processing.

[Rz 20] One of the scenarios of legal data analysis seems to be in particular appropriate: the legal file. It can be highly voluminous, containing statements of many persons with very different quality and purpose. Here, legal data science aims to identify personal and topical relations between the various arguments concerning facts or legal rules.

---

<sup>14</sup> Cf. the ICAIL workshops on discovery of electronically stored information (DESI), on using machine learning and other advanced techniques to address legal problems in e-discovery and information governance, the latest in 2015, <http://www.umiacs.umd.edu/~oard/desi6/>.

#### 4. Legal data science: 8 views, 4 methods and 4 synthesis of the legal system

[Rz 21] The goal of legal data science is to complement the existing methodology of lawyering with computer-based methods. Paraphrasing the words of LAYMAN ALLEN in 1957<sup>15</sup>, legal data science provides razor sharp tools for the analysis of the textual element of the legal system. «The fight for justice and law» – in the sense of RUDOLF VON JHERING<sup>16</sup> – goes on, but now including legal data science. For a long time, work on legal knowledge has been only in the focus of legal theory and information science. Only since the late 1950's – with the start of research on legal information retrieval – appropriate progress can be noted (term retrieval, text retrieval, metadata, citations, search technologies, user interface, telecommunications, etc.).

[Rz 22] For a long time, only the legal text corpora was available and the subject of legal text analysis. With the growing digitalisation, electronic files are more and more emerging. The tools remain but the focus is on creating a map of relations between the various people, facts, rules and arguments. Here, the similarity with the analysis of textual feedback of e-participation platforms is obvious.

[Rz 23] The theory shows that law can be seen very differently, distinguishing 8 views – text view, metadata view, network view, user view, logic view, ontological view, visualisation view and argumentation view. It should be noted that in the knowledge representation of law, it is not solely about the documentation; each view represents further insights on the law itself.<sup>17</sup>

[Rz 24] Three different steps can be distinguished: text corpora and legal information retrieval, (semi)automatic metadata generation and detailed structured analysis of concepts, facts, rules and arguments.

[Rz 25] The first step, legal information systems with powerful search engines, has been around for a long time. Besides nice user interfaces, getting sufficient recall and precision remains a challenge. Ranking seems to be the best solution. The «views theory» of LU and CONRAD<sup>18</sup> - documents view, annotation view, citation view and user view describes this extension and constitute an important step on legal data analysis. However, it should be extended by four other views – legal logic, legal ontologies, legal visualisation and legal argumentation. Further, the works of SOWA<sup>19</sup>, FIEDLER<sup>20</sup>, ZELEZNIKOW/HUNTER<sup>21</sup>, the results of the book edited by

---

<sup>15</sup> LAYMAN ALLEN, Symbolic Logic: A Razor-Edged Tool for Drafting and Interpreting Legal Documents. In: The Yale Law Journal 66, p. 833, 1957.

<sup>16</sup> RUDOLF VON JHERING, Der Kampf ums Recht, Vortrag (The fight for law), Wien, 1872. Schutterwald/Baden 1997, 1872.

<sup>17</sup> ERICH SCHWEIGHOFER, Rechtsdatalytik – Versuch einer Teiltheorie der Rechtsinformatik. In: Erich Schweighofer/Franz Kummer/Walter Hötendorfer (eds.), Kooperation / Co-operation, Tagungsband des 18. Internationalen Rechtsinformatik Symposions IRIS 2015. Wien: Österreichische Computer Gesellschaft (OCG), p. 61, 2015.

<sup>18</sup> QIANG LU/JACK G. CONRAD, 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.

<sup>19</sup> JOHN F. SOWA, Knowledge representation: logical, philosophical, and computational foundations. Course Technology, Boston, MA, 2000.

<sup>20</sup> HERBERT FIEDLER, Modell und Modellbildung als Themen der juristischen Methodenlehre (Model and modeling as subjects of legal methodology). In: Proceedings of the International Legal Informatics Conference IRIS 2006, OCG Publishers, Vienna, p. 275, 2006.

<sup>21</sup> JOHN ZELEZNIKOW/DAN HUNTER, Building Intelligent Legal Information Systems, Representation and Reasoning in Law, Computer Law Series 13, Kluwer, Deventer, 1994.

YEARWOOD/STRANIERI<sup>22</sup> as well as our own work on legal knowledge representation<sup>23</sup> have been taken into account.

[Rz 26] The next step – metadata – is already here but in many different forms. Classification and summarization of important content or extraction of relevant information is a standard of legal information science (e.g. high-level West's Key Number System<sup>24</sup> or the formidable CELEX metadata system<sup>25</sup>). Contrary to these documentation systems, metadata in legal data science is generated mostly automatically. Given some input with proper heuristics, use of learning techniques and then matching of the text corpus with these pre-defined structures, a fine set of metadata can be achieved. Considering the strong dynamics of legal systems – like textual feedback of e-participation platforms – a main advantage is evident: after sufficient training and regular updating, the analysis tool can be used for a longer time, making the analysis of new documents or new perspectives much easier than the usual retrieval and reading. Three main types of metadata can be distinguished: general metadata, citations and user feedback. General metadata concerns document dates (creation, validity, etc.), document type and annotations with thesauri or classifications, also considering multilingual jurisdictions. Citations describe the relations between 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) and if it is quoted by others (in-bound [cited] sources). It is important to specify citations, e.g. referring also to the structuring elements of a document, e.g. articles, sections, paragraphs, lists, etc. The main types of citations 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.<sup>26</sup>. The similar automatic generation of temporal relationships describes the different states of a document, e.g. creation, validity, applicability etc. Data-friendly generation of user behaviour gives strong insights of their assessment. The user's perspective takes into account the opinions of legal professions, business and civil society about the document collection. Sufficient data protection remains a prerequisite of such metadata, in particular anonymization of user data. Metadata is not restricted to providing additional information only. It also allows the semi-automatic generation of summaries of documents, supporting a much easier and more efficient analysis of documents. The user needs are taken into account in sophisticated ranking algorithms: document vs. requirements, document in the corpus, document in the citations network, document in the timeline etc.

[Rz 27] The last step is the most difficult one and requires strong tools of linguistic analysis. The text has to be segmented in facts, logical rules and arguments. Presently, tools are developed and tested in small environments with promising results. Here, the pre-defined models of legal

---

<sup>22</sup> JOHN YEARWOOD/ANDREW STRANIERI (eds.), *Technologies for Supporting Reasoning Communities and Collaborative Decision Making: Cooperative Approaches*, IGI Global Publishers, Hershey, PA 2011.

<sup>23</sup> ERICH SCHWEIGHOFER, *Indexing as an ontological-based support for legal reasoning*. In: John Yearwood/Andrew Stranieri (eds.), *Technologies for Supporting Reasoning Communities and Collaborative Decision Making: Cooperative Approaches*, IGI Global Publishers, Hershey, PA 2011, p. 213, 2011.

<sup>24</sup> Westlaw, *West Key Number System* on WestlawNext, L-374484.pdf  
<https://info.legalsolutions.thomsonreuters.com/pdf/wln2/L-374484.pdf>.

<sup>25</sup> ERICH SCHWEIGHOFER, *Wissensrepräsentation in Information Retrieval-Systemen am Beispiel des EU-Rechts*, Dissertation, Universität Wien 1995, (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.

<sup>26</sup> ALBRECHT BERGER, *The development of references in the legislative documentation*, Verlag Dokumentation, Pullach near Munich 1971.



knowledge are much more sophisticated requiring intensive use of natural language processing. The existing prototypes are explorative with different complexity. The logical representation is based on the first order logic, complemented by a representation of temporal layers<sup>27</sup> and the personal scope<sup>28</sup>. 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).<sup>29</sup> The legal ontology brings the automation option and a much more powerful systemisation. Since the 1990's, many legal ontologies have been developed<sup>30</sup>; now a strong standard exists with LRI Core, LKIF<sup>31</sup> and LegalRuleML<sup>32</sup>. The respective elements of the concepts have to be transposed into a computer-readable structure.<sup>33</sup> Facts, e.g. a world ontology, can be taken from projects like Cyc.<sup>34</sup> A major advantage of an ontology is the easier representation of relations between facts and legal concepts in a form of pre-subsumption. Legal visualization concerns the use of graphics, images and videos for visual representation of the law.<sup>35</sup> Visualisation is especially powerful for representing the essentials, making them memorable, increasing understanding and showing hidden connections. Graphical notations are also a support for the formalization of the law, also in cases where the necessary level of abstraction for formalization is not yet reached. The focus of the argumentation view is twofold: representation of legal arguments but also the structure of each document in its elements and logical structure: factual information, evidence, arguments, conclusions, etc.<sup>36</sup>. Following the legal theory work on theories on legal argumentation (e.g.<sup>37</sup>), formalisation of legal arguments stays in the focus since the 1990's. The elements of argumentation are systematized in a logical argumentation structure. Arguments are an essential tool of legal work, because thesis, antithesis and synthesis are relevant in each case due to the dialectical roles of plaintiff, defendant and judge.

---

<sup>27</sup> JOHANNES SCHARF, Wissensrepräsentation und automatisierte Entscheidungsfindung am Beispiel des KOVG (Knowledge representation and automated decision making), PhD thesis, University of Vienna, 2015.

<sup>28</sup> Idea of Erich Schweighofer, based on relevant thoughts of Ch. Reed, You Talkin' to Me?. In: Jon Bing, en hyllest, a tribute, p. 154–171, 2014.

<sup>29</sup> A prime example is the Australian company SoftLaw; this was subsequently acquired by Oracle; the application itself is available as Oracle Business Rules.

<sup>30</sup> GIOVANNI SARTOR/POMPEU CASANOVAS/MARIANGELA BIASIOTTI/MERITXELL FERNÁNDEZ-BARRERA (eds.), Approaches to Legal Ontologies: Theories, Domains, Methodologies, Dordrecht/Heidelberg/London/New York, Springer, 2011.

<sup>31</sup> RINKE HOEKSTRA/JOOST BREUKER/MARCELLO DI BELLO/ALEXANDER BOER, The LKIF Core Ontology of Basic Legal Concepts. In: Pompeu Casanovas/Mariangela Biasiotti/Enrico Francesconi/Maria Teresa Sagri (eds.), Proceedings of LOAIT 07, II. Workshop on Legal Ontologies and Artificial Intelligence Techniques, p. 43–64, 2007.

<sup>32</sup> Website OASIS LegalRuleML, [https://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=legalruleml](https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legalruleml).

<sup>33</sup> NÚRIA CASELLAS/ENRICO FRANCESCO/RINKE HOEKSTRA/SIMONETTA MONTEMAGNI (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, 2009; POMPEU CASANOVAS/MARIANGELA BIASIOTTI/ENRICO FRANCESCO/MARIA TERESA SAGRI (eds.), Proceedings of LOAIT 07, II. Workshop on Legal Ontologies and Artificial Intelligence Techniques, 2007.

<sup>34</sup> Wikipedia EN, Cyc, <https://en.wikipedia.org/wiki/Cyc>.

<sup>35</sup> COLETTE BRUNDSCHWIG, 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 (eds.), *Strukturierung der Juristischen Semantik – Structuring Legal Semantics*, Festschrift für Erich Schweighofer, Editions Weblaw, Bern, p. 573, 2011.

<sup>36</sup> KEVIN D. ASHLEY, Modeling Legal Argument. Reasoning with Cases and Hypotheticals. Cambridge MA: MIT Press, 1990.

<sup>37</sup> ROBERT ALEXY, Theory der juristischen Argumentation (Theory of Legal Argumentation). Frankfurt am Main, 1983.

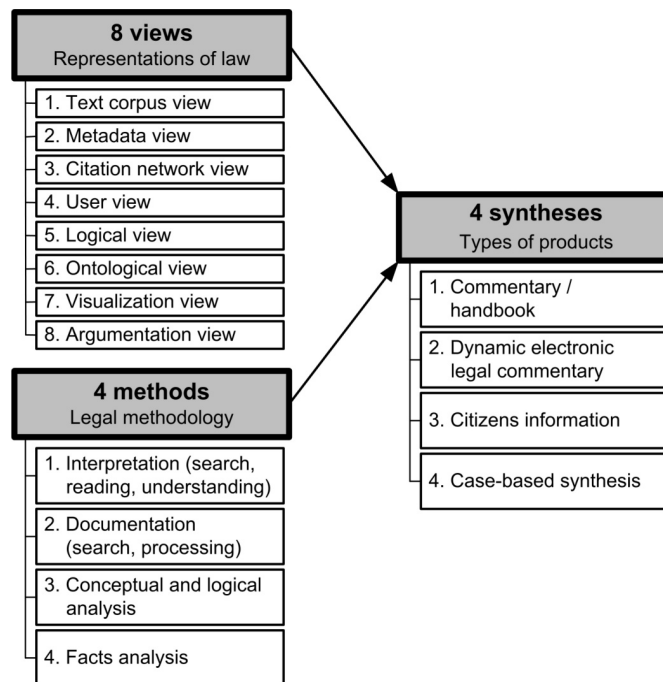


Figure 1: Legal data science<sup>38</sup>

[Rz 28] For providing these views, 4 legal methods are available: legal dogmatics, in particular interpretation, documentation, conceptual and logical analysis and fact analysis. The standard 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. Other elements are also considered, e.g. social context, legal authorities, sophisticated methods of interpretation etc. However, the very high costs can be financed only in exceptional circumstances. Thus, many views are not available at all. Stronger use of AI & law can be the solution to this knowledge acquisition problem. Legal documentation is the first major achievement solving the challenge of up-to-date storage and retrieval of documents. Legal information 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. 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 strong logical structure.

[Rz 29] The 8 views are more an intermediate result that has to be transformed into a legal product, e.g. syntheses: commentary/handbook, dynamic electronic legal commentary, citizen's information and case-based synthesis. AI & law methodology allows a representation according to the different needs of the users: Handbook, Dynamic Electronic Legal Commentary (DynELC), citizen information or case-related synthesis. Starting from the text corpus in XML format, legal data science methods drawn from many fields within the broad areas of mathematics, statistics, and information technology: pattern recognition and learning, machine learning, probability models,

<sup>38</sup> Source: VYTAUTAS YRAS/FRIEDRICH LACHMAYER/ERICH SCHWEIGHOFER, Network of Legal Metalevels. In: Erich Schweighofer/Franz Kummer/Walter Hötzenedorfer/Georg Borges (eds.), Netzwerke / Networks, Tagungsband des 19. Internationalen Rechtsinformatik Symposions IRIS 2016. Wien: Österreichische Computer Gesellschaft (OCG), p. 83, 2016.

statistical learning, visualization, data warehousing, etc. AI & law has developed and refined these methods for legal output. It has to be noted that the required accuracy for practical use has very often not been achieved yet. Recent research on network analysis in law can be considered as an important step forward.<sup>39</sup>

## 5. AI & law methods for e-participation platforms

[Rz 30] Due to space reasons, only a short summary of the state of the methodology can be given. The focus is on pattern recognition, thesaurus generation, machine learning, natural language processing, ranking, legal ontologies, legal logic, legal argumentation and visualization.

[Rz 31] Document categorisation, document segmentation, citation analysis<sup>40</sup>, temporal relations<sup>41</sup> and thesaurus generation rely on pattern matching, machine learning and natural language processing<sup>33</sup>. The representation of documents in XML format has to be refined by use of a standard part-of-speech tagger. A knowledge base with patterns of document types, document segments, citation structures, temporal relations and concepts is matched with the text corpus resulting in a raw analysis of a document. The knowledge base has to be refined and improved using machine learning techniques. Further, the output has to be presented in a formal model (e.g. types of citations with relevance ranking). This semi-automatic approach delivers also relevant feedback from authors and users for improving the knowledge base.<sup>42</sup> Model-based visualizations are based on grammars, i.e. strict specifications for composing visualizations based on sets of pre-defined elements and relation (e.g. the PICTMOD of *Fill*<sup>43</sup>). Ranking is used as a mechanism for establishing a relationship of relevance for a particular information or analysis. In law, it is used in legal information retrieval, however, with still insufficient success compared to the Google algorithm.<sup>44</sup> Ranking gives some guidance concerning the relations of a document vs. requirements, document in the corpus, document in the citations network, document in the timeline etc.

---

<sup>39</sup> JÖRG LANDTHALER/BERNHARD WALTL/FLORIAN MATTHES, Unveiling references in legal texts: implicit versus explicit network structures. In: Erich Schweighofer/Franz Kummer/Walter Hötendorfer/Georg Borges (eds.), *Netzwerke / Networks*, Tagungsband des 19. Internationalen Rechtsinformatik Symposions IRIS 2016. Wien: Österreichische Computer Gesellschaft (OCG), p. 71, 2016.

<sup>40</sup> ERICH SCHWEIGHOFER/DIETER SCHEITHAUER, The Automatic Generation of Hypertext Links in Legal Documents. In *Proceedings of Database and Expert Systems Applications, 7th International Conference, DEXA'96, Zurich 1996*, Lecture Notes in Computer Science 1134, p. 889, Springer, Berlin.

<sup>41</sup> JOHANNES SCHARF, Wissensrepräsentation und automatisierte Entscheidungsfindung am Beispiel des KOVG (Knowledge representation and automated decision making). Ph.D thesis, University of Vienna, OCG Publishers, 2015.

<sup>42</sup> ERICH SCHWEIGHOFER, Indexing as an ontological-based support for legal reasoning. In: John Yearwood/Andrew Stranieri (eds.), *Technologies for Supporting Reasoning Communities and Collaborative Decision Making: Cooperative Approaches*, IGI Global Publishers, Hershey, PA 2011, p. 213–236, 2011.

<sup>43</sup> HANS-GEORG FILL, Bridging Pictorial and Model-based Creation of Legal Visualizations: The PICTMOD Method. In: Erich Schweighofer/Franz Kummer/Walter Hötendorfer (eds.), *Kooperation / Co-operation*, Tagungsband des 18. Internationalen Rechtsinformatik Symposions IRIS 2015. Wien: Österreichische Computer Gesellschaft (OCG), p. 443–447, 2015.

<sup>44</sup> ANTON GEIST, Relevanzsortierung bei Rechtsdatenbanken (relevance ranking in legal databases), Ph.D thesis (to be submitted), University of Vienna, 2016.

[Rz 32] The logical analysis and the argumentation analysis require a rule and an argumentation model. Some support can be achieved if text is transferred into logical rules (e.g. the well-known example of Softlaw/Oracle<sup>45</sup> or recently by ISLAM and GOVERNATORI<sup>46</sup>). Flow-charts (activity diagrams) are also very helpful in practice. Recent work on argumentation analysis can be found in the ICAIL2015 proceedings.<sup>47</sup> 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<sup>48</sup>; further, a differentiation should be made according to distinct applicability of norms on legal subjects.

[Rz 33] All these tools are highly helpful for analysis of textual feedback of e-participation platforms. However, computers can be only very supportive. 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.

## 6. Outline of an analysis of textual feedback of e-participation platforms

[Rz 34] A sketchy outline of a first analysis of textual feedback of e-participation platforms is described here. The text corpora are different, containing shorter documents than in law. The content and language can be very diverse, e.g. legal jargon, dialect, etc. Symbolic expressions are more common, e.g. a like or dislike button. In the first step, statements are analysed and metadata is added, using data analysis tools. This allows the visualisation of supporting tools for text analysis, e.g. the structure of a statement, its citations, its temporal relations, relevant user feedback etc. A logical and ontological analysis is not envisaged yet. All these metadata can be used to produce fine visualisations of the e-participation process.

[Rz 35] *Semantic web*: Textual feedback exists mostly in HTML files that have to be transferred to XML allowing adding of metadata with RDF or OWL. The metadata is best described by the 4 views of the legal system.

[Rz 36] *Identification*: The participation author can be anonymous or pseudonymised; self-declared discussant (e.g. local political), participation class, nick name etc. Further, additional information can be stored if available, in particular location data, sex, profession, education etc.

---

<sup>45</sup> SURENDRA DAYAL/MICHAEL HARMER/PETER JOHNSON/DAVID MEAD, Beyond Knowledge Representation: Commercial Uses for Legal Knowledge Bases. In: Proceedings of the Fourth International Conference on Artificial Intelligence and Law ICAIL1993, p. 167–174, ACM, New York.

<sup>46</sup> MOHAMMAD BADIUL ISLAM/GUIDO GOVERNATORI, RuleOMS: A Rule-Based Online Management System. In: Proceedings of the Fifteenth International Conference on Artificial Intelligence and Law ICAIL 2015, ACM, New York, p. 187, 2015.

<sup>47</sup> JOONSUK PARK/CHERYL BLAKE/CLAIRE CARDIE, Toward Machine-assisted Participation in eRulemaking: An Argumentation Model of Evaluability. In: Proceedings of the Fifteenth International Conference on Artificial Intelligence and Law, ICAIL2015, ACM, New York, p. 206, 2015.

<sup>48</sup> JOHANNES SCHARF, rOWLer – A hybrid rule engine for legal reasoning. In: Erich Schweighofer/Franz Kummer/Walter Hötendorfer (eds.), Kooperation / Co-operation, Tagungsband des 18. Internationalen Rechtsinformatik Symposions IRIS 2015. Wien: Österreichische Computer Gesellschaft (OCG), p. 155, 2015.

[Rz 37] *Classification of statement*: Proposals, comments, remarks on comments etc.

[Rz 38] *Document segmentation*: A very important part is the segmentation of a statement in its structural elements, e.g. author or author class, title (if available), facts, arguments, positions, emotions etc.

[Rz 39] *Citations*: The automatic generation of hypertext links describes the network between documents but also its relevance. It shows the relations to other documents (comments, remarks on comments etc.), discussion topics, etc.

[Rz 40] *Temporal relations*: The automatic generation of temporal relationships analyses the temporal dimensions: existence, efficacy etc.

[Rz 41] *Document description and summary*: A lexical ontology, e.g. a thesaurus, remains critical for a condensed short content description. Multilingual versions like Eurovoc can be very helpful for first handling of documents in a foreign language. Without question, an e-participation organiser must develop his own lexical ontology, taking into account language use in the constituency.

[Rz 42] *Citizen's assessment*: E-participation platforms can (semi)automatically generate summaries of perspectives of citizens on a statement. The feedback is mostly statistical data on likes, dislikes, votes etc.

[Rz 43] *Logical and argument analysis*: Given the document segmentation, textual rules and arguments can be also formally represented in a pre-defined model, e.g. in very useful practical flow charts (activity diagrams) or argumentation structures. Here, more research is required to achieve a higher quality of transfer of text into formal logical models or argumentation structures. A major part would be the extension of the lexical ontology to a domain specific ontology on the relevant participation process.

[Rz 44] *Visualisations*: Recently, traditional pictorial representations are supplemented with model-based visualizations. The semantic text corpora can be used to automatically produce user-defined visualisations of situations, events etc. The use of ranking algorithm in visualisation describes the importance of the relevant information.

[Rz 45] The output would be mostly the result of matching the text with pre-defined analytical structures, extraction of information elements and some (light) interpretation. It must be mentioned that transparency of these tools is highly important. The e-participation organiser must have consensus in his constituency about the use of these tools in order to achieve an analysis that can be accepted by all stakeholders.

## 7. Conclusions and future work

[Rz 46] E-participation platforms face the same challenges as legal text corpora. Size, diversity, different language use and dynamic development strongly constrain intellectual analysis, in particular in real time. 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). Present analysis consists in creating 8 views using AI & law methodology for producing automatically or semi-automatically analytical maps. The same methodology can be used to speed up textual feedback analysis of e-participation platforms. A first outline is given in this paper, focussing on generating semi-automatically metadata of statements: identification, type, temporal and pragmatic relations, summary and assessment. The metadata is visualised in

topical maps according to the needs of the e-participation organiser. In the future, this outline will be further developed and, hopefully, added to the E-Participation Platform Demonstrator.

---

ERICH SCHWEIGHOFER, Professor, University of Vienna, Centre for Computers and Law (DEICL/SIL), Department of European, International and Comparative Law, Schottenbastei 10-16/2/5, 1010 Vienna, AT. [Erich.Schweighofer@univie.ac.at](mailto:Erich.Schweighofer@univie.ac.at); <http://rechtsinformatik.univie.ac.at/> / <https://epartizipation.info>; <https://deicl.univie.ac.at>.

Acknowledgements: The E-Participation project is funded by the Austrian security research programme KIRAS of the Federal Ministry for Transport, Innovation and Technology (bmvit). It has developed an E-Participation Platform Demonstrator that can strengthen direct democracy offering different participation scenarios. The support of Vinzenz Heußler and Janos Böszörményi is gratefully appreciated.<sup>49</sup>

---

<sup>49</sup> HEUSSLER/SCHOSSBÖCK/BÖSZÖRMENYI (note 13).