Towards Measures of Complexity: Applying Structural and Linguistic Metrics to German Laws

Bernhard WALTL [a,1] and Florian MATTHES [a]

[a] *Software Engineering for Business Information Systems,*
Technische Universität München, Germany

Abstract. The increasing complexity of legal systems has many origins, which are worth a deeper analysis. This paper is an attempt to unveil the complexity in legal texts driven by structural, lexical and syntactical properties. Thereby we transferred established quantitative methods from structural network analysis and linguistics into the domain of legal text analysis. Based on 3 553 German laws, respectively regulations, we calculated several structural and lexical indicators for complexity and determined highly significant correlations ($p \leq 0.01$). The papers' contribution is a set of metrics, enabling a structured and objective comparison of legal texts regarding their complexity.

Keywords. German law, complexity, legal texts, readability, measures

Introduction

Today's legal systems are complex and highly interdependent. Beside of the laws themselves, the areas they are governing are also complex. The German Civil Law Code (BGB) for example regulates the interactions between private individuals, consisting of about 2 400 different paragraphs containing more than 180 000 words, was approved in 1900. Today's legislation is dealing with highly complex areas, e.g., economy, ecology, industry, and additional regulations, with more or less effect on our daily life, are passed frequently. For example, in the legislative period from 2009 to 2013 553 different laws and between 2005 and 2009 616 laws were passed[2] [1]. The legal system can be seen as a collection of norms in legal texts, represented by laws, paragraphs, enumerations holding the regulatory information, i.e. permission, prohibition, duty and exemption.

The overall complexity of legal systems has to increase to match the complexity of society or economy [2]. Additionally, the perceived legal complexity within society increases. A survey amongst over 1 500 persons conducted 2013 in Germany states that a majority (61 %) of the German citizens think that the laws are far too complicated and not understandable to "ordinary" citizens [3]. The perceived complexity is not only

---

[1]Corresponding Author: Bernhard Waltl, Software Engineering for Business Information Systems, Boltzmannstr. 3, 85748 Garching bei München, Germany; E-mail: b.waltl@tum.de.

[2]In prior legislative periods the numbers are of the same order of magnitude, i.e. 1990-1994: 507 laws, 1994-1998: 566 laws, 1998-2002: 559 laws, 2002-2005 400 laws.

driven by the amount of laws that are continuously passed, but also due to various other reasons arising from structural and linguistic properties of laws, which are going to be discussed within this work.

This article aims to provide objective complexity metrics in legal systems. This leads to an analytical understanding of what complexity in the legal domain can be and furthermore provides an objective foundation for upcoming debates and research challenges. The following Section states research questions addressed by this paper and introduces the German law dataset used. Section 1 discusses existing approaches measuring the complexity and difficulties of texts in general and in particular of normative texts. The used methodological approach, which is a combination of structural network theoretic and quantitative linguistic approaches, is described in Section 2. Our main contribution is the provision of concrete measurements and metrics, as well as highly significant correlations ($p \leq 0.01$) between a subset of them. The paper concludes with a critical reflection of the findings in Section 5 and provides a framing and an outline of upcoming research challenges in Section 6.

1. Complexity of the Legal System

In 1992 Schuck published a work addressing the complexity of a legal system. According to him legal complexity is a composite of four qualitative dimensions, namely density, technicality, differentiation, and indeterminacy [4]. Dense rules cover a large portion of the policy space and therefore address a broad variety of cases; they are likely to conflict with other rules. Technicality is a property of rules regulating specific legal areas, like Basel II for accounting firms. These rules require sophistication or expertise of those who wish to understand and apply them. The complexity of legal rules is also influenced by their differentiation, which represents the different sources and decision structures required to create, elaborate and apply a rule. For example does product safety draw upon many sources of legitimacy, such as regulatory standards, agencies, and technical organizations. The remaining driver for complexity – according to Schuck – is the indeterminacy of legal rules, arising from vague verbalization using words like "adequate", "reasonable" and so forth. Although indeterminate rules are open-textured, their outcome is hard to predict and their application is costly. Furthermore, complexity determination can only be done in a relativistic sense: "[...] a legal rule, process, or institution is neither simple nor complex; it is only more or less simple or complex compared to some other actual or ideal one." [4, p. 5].

Legal systems can exhibit complex and dynamic behavior. Interactions can cause unpredictable effects and minor changes may have a huge impact on the outcome. Based on this assumption Bourcier and Mazzega published 2007 guidelines to measure the intrinsic complexity of legal texts, distinguishing between "structure-based" and "content-based" measurements [5]. The first approach, well known in complex systems theory, is to analyze the structure, i.e. the law as a network, text organization and quotations of a legal text. The latter investigation addresses the potential effects that laws have, including unexpected side effects. Bourcier and Mazzega emphasize that the complexity is driven by structure and content but they describe vagueness and complication as concepts, which are – according to them – converse to complexity.

As an introduction to the AICOL workshop[3] in 2009 Casanovas et al. published basic ideas of how prior concepts and understandings of complex systems apply to the law system. Legal complexity is identified as a serious problem: "[...] todays complexity of the law stands for the crisis of legal positivism and the dogma of sovereignty" [6]. Although their work presents several challenges for legal semantic web services, it misses a discussion of what possible metrics for legal complexity actually could be. Referring to the work of Bourcier and Mazzega Casanovas et al. follow their approach of "structure-based" and "content-based" measurements.

Based on a mathematical approach Bommarito and Katz measured the complexity of the law using the United States Code [7]. Their motivation was that "unnecessary complexity" increases the likelihood of "suboptimal" decisions. Therefore they argue to draft laws as simple as possible – but not simpler [2]. Although they admit that complexity of law affects the costs associated with legal compliance and litigation, they also emphasize that the analysis of structure and content in written legal texts provides useful indicators for legal complexity. They focused on analyzing structural properties (i.e. number of structural elements, element depth, citations) and lexical properties (i.e. token size, word length, word entropy) resulting in a framework for measuring legal complexity.

The analysis of related articles clearly shows that early approaches in describing legal complexity aim at providing qualitative attributes, whereas the more recent approaches try to unveil the complexity in a quantitative way, i.e. making it objectively measurable. Although quantifying complexity may lead to a controversial discussion whether something is complex or not, mathematical results are uncontroversial and allow comparison within the legal domain. Complexity indicators influence the drafting, interpretation and litigation processes in the legal domain and the resources required by those processes (e.g., time and money). Therefore, all of these processes benefit from reduction of reading complexity. Negative effects, driven by unnecessary complexity, can cause misallocation of valuable resources. Consequently, it's avoidance and early recognition contributes in economizing those resources.

1.1. Readability of Legal Texts

According to §42(5) of the Joint Rules of Procedure of the Federal Ministries (GGO) legal texts, modifications in laws or new regulations shall be written in plain and intelligible language: "Gesetzentwürfe müssen sprachlich richtig und möglichst für jedermann verständlich gefasst sein." [8] The objective is, that laws should be understandable to everyone. However, legal texts are far from understandable by everyone [9,10]. The incomprehensibility of legal texts has become a "tradition" and is not only driven by the terminology because even legal professionals are not capable to understand every normative legal text [11]. Quantifying the understandability of legal texts is a nontrivial problem, whereas the readability can be measured using metrics that take into account lexical and structural properties of the text, e.g. Flesch Reading Ease. The FRE is calculated as follows $FRE_{german} = 180 - ASL - (58.5 \cdot ASW)$; (ASL... average sentence length, ASW... average number of syllable per word [12])[4]. Interestingly, the readability metrics inherently represent far more textual properties than the formula takes

---

[3]Artificial Intelligence Approaches to the Complexity of Legal Systems

[4]The constants reflect language depending properties.

into account. This is because of the dependencies throughout textual entities, e.g. word frequency, vocabulary variety, lexical ambiguity, phonetic, etc. [13]

To determine the understandability of legal texts psychological and various other components, e.g. social, educational have to be considered, whereas readability indexes are calculated using structural-quantitative elements. Although critiques argue that the analysis of reader-text-interaction is required to make statements about the understandability, representing readability in terms of lexical and structural is still applicable.

Critiques warn that analyses exclusively considering readability metrics have to be treated with caution [11, pp.337]. Nevertheless, to compare different texts and to classify them based on a uniform method is an accepted method in other disciplines, like evaluating software requirements specifications [14,15] or in the domain of knowledge transfer [13]. Over 40 different measures exist to determine the readability of texts, whereas the Flesch Reading Ease is one of the most common ones [16,17]. Within Section 4.1 we will provide calculated information about the readability of laws, not using it as an absolute measure but as a relative measure for comparison and correlation with readability indicators.

2. Research Method: Combining Structural Analysis and Linguistic Approaches

Based on the investigations of prior approaches we address the structural and linguistic complexity indicators in German legal texts with the following research questions:

RQ1 How is textual difficulty contributing to the complexity of laws?
RQ2 How does the length of laws, respectively paragraphs, influence the difficulty?
RQ3 How does the variety of the vocabulary contribute to the difficulty?
RQ4 How can the indeterminacy of laws be measured?

Measuring the complexity in terms of quantitative research is a challenging task, especially within the legal system [2]. It can be seen from two points of view: the quantitative linguistic perspective and the structural network perspective. Therefore, we argue that the methodological transfer of complexity determining methods from related disciplines is applicable to derive complexity indicators for the legal domain. Based on this, the research method is quantifying the German law, using common methods from complexity theory combined with network and linguistic theory (see Table 1).

Structure In various disciplines complexity driven by structure is well-studied [18]. Relying on the fact, that network structure is jointly responsible for the complex behavior of a system it is a reasonable to assume that the same holds also for the legal system. With respect to the results from interdisciplinary research on complexity theory the statement of Bourcier and Mazzega holds that "Law can be considered as a complex dynamical system evolving from network structures" [5]. Within this work we will extend the "structure-based" approach in determining complexity in the legal domain by quantifying it.

Linguistic In the linguistic domain, there is no consensus about how complexity of language can actually be measured. Thereby different approaches exist, mainly focusing on counting occurrences on hand-picked properties or complexity indicators [19]. Approaches from formal linguistics emphasize and recommend the us-

age of quantitative measurements in order to produce falsifiable results [20]. Resulting in quantitative linguistics a variety of metrics exist, representing the readability of texts, which are in particular relevant to the domain of knowledge transfer [13]. Other approaches like morphological or phonological analysis could not easily be applied, since we operate on a large dataset [19].

Neglecting the fact that legal complexity is also influenced by the consequences, that legal statements may have (see Section 1), we focus on objectively measurable indicators. Predicting the possible effects, that laws may have is challenging [21,2] and definitely controversial. Consequently, our approach focuses on uncovering the aspects of complexity analyzing the network and linguistic structure of legal texts, exemplary with the German law.

Name	Abbreviation	Indicated Complexity	Literature
Paragraph Count	# §	Linguistic & Structural	[20,2]
Sentence Count	# S	Linguistic & Structural	[22,19]
Word Count	# W	Linguistic & Structural	[22,2]
Structural Depth	D	Structural	[18,2]
Number of Outgoing References (internal)	INT	Structural	[18,2]
Number of Outgoing References (external)	EXT	Structural	[18,2]
Vocabulary Variety	V	Linguistic	[22,20,19]
Indeterminacy	I	Linguistic	[22,19]
Readability (Flesch-Reading-Ease)	FRE	Linguistic	[17,12,13]

Table 1. Quantitative indicators for complexity in legal texts derived from literature.

As stated above, our approach aims to provide a method to quantitatively evaluate legal texts on a structural and lexical level. This is as a foundation for the assessment of the complexity of legal texts, and allows further applications in various domains, e.g., drafting, educating, communicating etc.

Our research is based on the analysis of German laws, which we retrieved from the platform *www.gesetze-im-internet.de* hosted and maintained by the Federal Ministry of Justice, represented by *Kompetenzzentrum Rechtsinformationssystem* (CC-RIS). To run deeper analysis we imported all available laws into a local information system. At the importing date (13th June, 2014) we imported 6 015 laws and regulations, which represent according to the platform "almost the complete and current federal law" [23]. Since we are performing several algorithms we only considered those texts, with at least 200 words[5], leading to a dataset with 3 553 distinct legal texts.

3. Structural Metrics of the German Law

3.1. Structural Properties & Depth

The legal system in Germany is characterized by text-intensive legal documents. For example the Civil Law Code (BGB)[6] or the Criminal Code (StGB) cover a broad spectrum

---

[5]Text corpus and Readability metrics should not be applied on texts with less than 200 words [16].

[6]The provision of further explanation to the law abbreviations is omitted within this work, since it is not necessary to understand the conclusion, but would consume valuable space.

of different topics. As in legal codes of other countries (see United States Code [2]), the textual representation follows a tree-like structure: law, chapter, section, paragraph, sentence and enumeration[7]. However, since paragraphs are consecutively numbered within a law, those numbers are sufficient to uniquely locate a reference. Consequently, the structural depth within the German law is 4 (law $\rightarrow$ paragraph $\rightarrow$ sentence $\rightarrow$ enumeration). The structural depth does therefore not influence complexity, nor understandability.

3.2. Word & Paragraph Count

Using the mentioned dataset, we counted the number of words and paragraphs in the different legal texts and sorted them accordingly. The paragraphs were counted individually, although they may have the same number (e.g., §319 and §319a are counted separately). In Table 2 the ten topmost legal texts, according to the sorting criteria (see arrow symbol), are shown. The last column of each table holds the average size of paragraphs (in words). This provides an insight into the size of legal texts in the German laws. So one can clearly see that legal texts are can quite long, although their paragraphs stay in a manageable length, so a comprehensive understanding of laws is challenging.

As already stated out, a text corpus is the "container" for the different laws and a paragraph is a segment which tends to hold highly related information. Consequently, long paragraphs are likely to contain more information and thereby tend to consist of more complex definitions, differentiations, regulations and content (see also Section 4.1).

Law	#W ↓	#§	# W/§
BGB	185 751	2 381	78
SGB 5	147 621	456	323
KAGB	113 166	356	317
KredWG	91 145	208	438
HGB	90 877	643	141
ZPO	90 421	1 003	90
SGB 6	84 683	413	205
AMG	77 002	216	356
StPO	74 887	644	116
StGB	62 986	518	122

Law	#W	#§ ↓	# W/§
BGB	185 751	2 381	78
ZPO	90 421	1 003	90
StPO	74 887	644	116
HGB	90 877	643	141
StGB	62 986	518	122
FamFG	47 272	499	94
SGB 5	147 621	456	323
AO 1977	61 122	423	144
SGB 6	84 683	413	205
InsO	37 876	380	100

Table 2. An overview of the largest German laws, according to word count and paragraph count.

3.3. Outgoing References: Internal and External Dependencies

The references between different laws and paragraphs within laws, are causing the network-like dependencies throughout the legal system (see Table 2). To understand or interpret a law, respectively a paragraph, it can be necessary to be aware of the laws, to which it's referring. Consequently, the number of outgoing references influences the complexity of a law, whether they are pointing to another law or to another paragraph within the same law. Within our analysis, we distinguish between internal and external outgoing references. Whereas the internal outgoing references are those where one para-

---

[7]Illustrative examples: BGB, StGB or KAGB

graph, e.g. §46b StGB, refers to a paragraph of the same law, e.g. to §49 StGB. External outgoing references are dependencies which point to another law, e.g. §46b StGB refers also to §100a of the StPO.

Law	#W	#§	INT↓		Law	#W	#§	EXT↓
BGB	185 751	2 381	4 489		KredWG	91 145	208	666
SGB 5	147 621	456	4 391		KAGB	113 166	356	529
KAGB	113 166	356	2 988		EnergieStV	36 167	137	401
HGB	90 877	643	2 238		SGB 5	147 621	456	400
AMG…	77 002	216	2 165		KStG 1977	26 430	40	388
KredWG	91 145	208	2 147		FamFG	47 272	499	328
StPO	74 887	644	2 080		EnWG 2005	56 501	206	292
ZPO	90 421	1 003	2 007		SGB 6	84 683	413	274
VAG	59 246	257	1 800		SVG	33 470	149	271
StGB	62 986	518	1 791		AMG 1976	77 002	216	266

Table 3. An overview of the ten most referring laws, sorted by number of internal and external references.

We omit the analysis of a laws' effects and do not consider the incoming references. The incoming references determine the consequences of a law, and since we do not analyze them it is not required to understand the "depending" laws.

4. Linguistic Metrics of the German Law

4.1. Readability

Although their usage is controversial (see Section 1.1), we used readability metrics to compare laws with each other. Figure 1 shows two correlations whereas the left one visualizes the readability depending on the number of words a law has. Hereby we observed a not significant correlation ($r = 0.016$, $p > 0.01$). Interestingly, the FRE correlates with the average word length of a paragraph. Hereby we observed a negative correlation ($r = -0.322$, $p \leq 0.01$). This example shows that even though the absolute length of a law only has a minor impact on the readability, the average length of paragraphs seems to have a major impact. This result is not obvious, since the FRE metric depends on language specific constants, the average sentence length and the average number of syllables per word, which are independent of the number of words in a law text.

4.2. Indeterminacy

Since laws intentionally try to cover many different cases and situations their formulation requires abstract concepts and word [24]. E.g. §142 StGB states out that someone involved to an accident is allowed to leave after an "adequate" (Original: "angemessen") waiting time without committing a hit-and-run offense. Indeterminate words and concepts, like "adequate", represent situations not limited to a single circumstance. This technique ensures on one hand the applicability throughout a variety of cases but on the other hand increases the difficulty of interpretation. Consequently, those words indicate the complexity of legal texts. Their understandability and applicability is influenced by
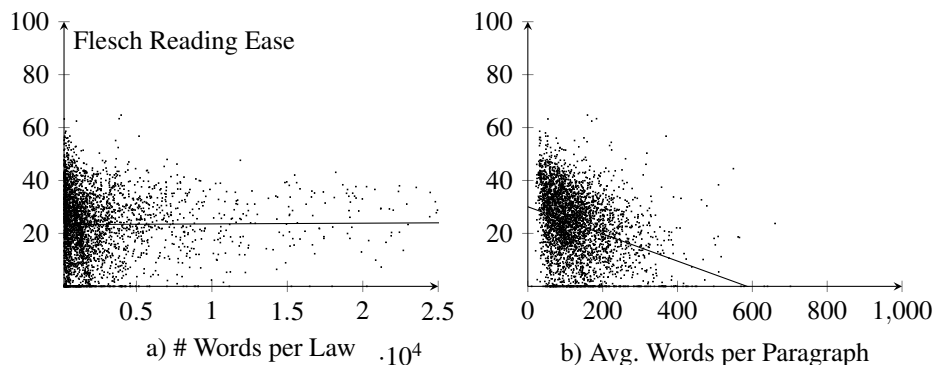
a) # Words per Law $\cdot 10^4$

b) Avg. Words per Paragraph

Figure 1. Increasing number of words in the paragraphs cause less readability.

indeterminate terms. Therefore, we identified 62 indeterminate legal terms, which are due to limited space not listed here, and counted their appearance throughout German law texts. The result is shown in Table 4 and Figure 2.

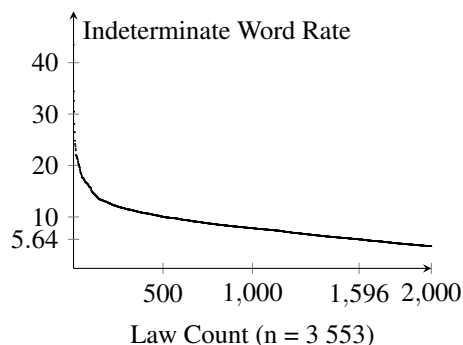| Law	#W	I↓	I Rate
BGB	185 751	1 505	8.10
KAGB	113 166	1 285	11.36
SGB 5	147 621	1 173	7.95
KredWG	91 145	938	10.29
StPO	74 887	735	9.81
HGB	90 877	700	7.70
ZPO	90 421	640	7.08
StGB	62 986	638	10.13
VAG	59 246	635	10.72
AMG	77 002	566	7.35

Table 4 & Figure 2. The German laws containing most indeterminate words and their relative frequency (per 1 000 words). The Figure visualizes the distribution of indeterminate legal terms (mean = 5.64).

Although we only considered a subset of the possible indeterminate legal terms it can be observed, that some laws tend to have a high occurrence rate of those words. Figure 2 unveils the relative frequency (per 1 000 words) of those legal terms. Whereas the average usage is 5.64 indeterminate words per 1 000 words, more than 1 596 legal texts have a higher usage of those complexity-driving terms.

4.3. Vocabulary Variety

The vocabulary variety drives the inhomogeneity of texts and is an indicator for the number of distinct topics covered by a law. Therefore, we analyzed the number of distinct words within a law text. Since we are primarily interested in the number of words contributing to the content, we removed the stop words of the text. Table 5 shows the German laws with the largest vocabulary variety, the number of distinct words per 1 000 words, and their corresponding readability index. Trivially, the total number of words correlates with the number of distinct words ($r = 0.902$, $p \leq 0.01$). Interestingly, the readability

is not influenced by the number of distinct words used within a legal text (r = 0.010, p > 0.01). This relationship is shown in Figure 3. Hence, the conclusion that vocabulary variety contributes to the complexity of legal texts, can not be drawn. Furthermore, the significance is above 0.01, which also indicates the lack of dependence between both measured parameters.

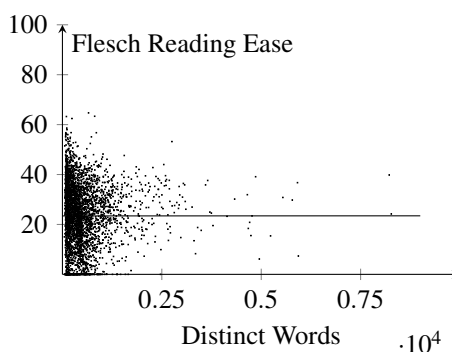| Law	V ↓	V/#W	FRE
SGB 5	8 271	56.03	24.21
BGB	8 219	44.25	39.82
KredWG	5 934	65.11	7.35
ZPO	5 912	65.38	36.74
HGB	5 781	63.61	29.78
AO 1977	5 544	90.70	30.81
EnWG 2005	5 235	92.65	15.42
KAGB	4 952	43.76	6.15
StPO	4 859	64.88	39.12
SGB 3	4 769	84.61	23.42

Table 5 & Figure 3. An overview of the German laws with largest vocabulary and their readability. The Figure visualizes that there is no measurable correlation between vocabulary size and readability.

5. Critical Reflection

The proposed approach of determining complexity indicators via structural and linguistic metrics allows quantified statements about German law texts. However, a few drawbacks remain. First, the law texts in general follow a specialized terminology and it is not sure whether linguistic metrics with educational and journalistic background can be applied properly. An interesting follow-up research question would be whether this framework reflects the perceived complexity of legal professionals and citizens and may be answered by conducting an empirical study. Second, the data source does not provide perfect and noise-free texts but some small text chunks not applicable for analysis. Therefore we restricted our analysis from originally 6 015 imported laws to 3 553 laws. Third, natural language processing approaches hardly achieve perfect accuracy while determining linguistic properties, such as counting words or other text corpus attributes.

6. Conclusion & Outlook

This paper is an attempt to unveil complexity indicators within legal texts. Thereby we focused on quantitative measurements of objective metrics. Based prior results legal systems' complexity and considering that those are mainly restricted to qualitative criteria, we objectively analyzed structural and linguistic properties. Thereby we derived several measurements and identified their contribution in textual difficulty. We showed that the average paragraph length of legal texts correlates with the readability, represented by Flesch-Reading-Ease. Making statements about the absolute complexity may be controversial, but it's a base for comparisons of legal corpora, or even juridical cultures.

Thereby not only different laws could be compared, but also the complexity changing over time or due to adoptions within legal texts.

Therefore, we consider our results as a next step towards an extensive understanding of complexity in the legal domain. Those results can furthermore serve as a base for legal drafting in order to avoid unnecessary complexity and ambiguity and to support the understanding and interpretability of legal texts, making them more accessible to people not familiar with the legal domain.

References

[1] Deutscher Bundestag, "Deutscher Bundestag - Neue Ausgabe des Datenhandbuchs zur Geschichte des Deutschen Bundestages," 15.04.2014. [Online]. Available: https://www.bundestag.de/datenhandbuch

[2] D. M. Katz and M. J. Bommarito, "Measuring the Complexity of the Law: The United States Code," *SSRN Electronic Journal*, 2013.

[3] ROLAND Rechtsschutz-Versicherungs-AG, "ROLAND Rechtsreport 2014," Köln, 2014.

[4] P. H. Schuck, "Legal complexity: some causes, consequences, and cures," *Duke Law Journal*, 1992.

[5] D. Bourcier and P. Mazzega, "Toward measures of complexity in legal systems," in *ICAIL '07 Proceedings of the 11th international conference on Artificial intelligence and law*, 2007, pp. 211–215.

[6] P. Casanovas, U. Pagallo, G. Sartor, and G. Ajani, "Introduction: Complex Systems and Six Challenges for the Development of Law and the Semantic Web," in *AI Approaches to the Complexity of Legal Systems*, P. Casanovas, U. Pagallo, G. Sartor, and G. Ajani, Eds. Springer Berlin Heidelberg, 2010.

[7] Bommarito II, Michael J. and D. M. Katz, "A mathematical approach to the study of the United States Code," *Physica A: Statistical Mechanics and its Applications*, vol. 389, no. 19, pp. 4195–4200, 2010.

[8] Bundesministerium des Innern, "Gemeinsame Geschäftsordnung der Bundesministerien," Berlin.

[9] K. D. Lerch, Ed., *Die Sprache des Rechts: Recht verstehen*. Berlin and New York: De Gruyter, 2004.

[10] ——, *Die Sprache des Rechts: Recht vermitteln*. Berlin and New York: W. de Gruyter, 2005.

[11] C. F. G. Schendera, "Die Verständlichkeit von Rechtstexten," in *Die Sprache des Rechts: Recht verstehen*, K. D. Lerch, Ed. Berlin and New York: De Gruyter, 2004, pp. 321–373.

[12] T. Amstad, *Wie verständlich sind unsere Zeitungen?* Zürich: Studenten-Schreib-Service, 1978.

[13] S. Wichter and A. Busch, Eds., *Wissenstransfer - Erfolgskontrolle und Rückmeldungen aus der Praxis: Band 5*. Frankfurt am Main: Peter Lang GmbH, 2006.

[14] William M. Wilson, Linda H. Rosenberg, and Lawrence E. Hyatt, "Automated analysis of requirement specifications," in *Proceedings of the 19th international conference on Software engineering*. Boston, Massachusetts, USA: ACM, 1997, pp. 161–171.

[15] H. A. Kanter, T. J. Muscarello, and C. Ralston, "Measuring the Readability of Software Requirement Specifications: An Empirical Study," *Information Systems Control Journal*, vol. 2008, no. 1, 2008.

[16] A. C. Graesser, D. S. McNamara, M. M. Louwerse, and Z. Cai, "Coh-Metrix: Analysis of text on cohesion and language," *Behavior Research Methods, Instruments, & Computers*, vol. 36, no. 2, pp. 193–202, 2004. [Online]. Available: http://dx.doi.org/10.3758/BF03195564

[17] R. Flesch, "A new readability yardstick," *Journal of Applied Psychology*, 1948.

[18] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175–308, 2006.

[19] M. Bane, "Quantifying and measuring morphological complexity," *Proceedings of the 26th West Coast Conference on Formal Linguistics*, pp. 69–76, 2008.

[20] R. K. Shosted, "Correlating complexity: A typological approach," *Linguistic Typology*, 2006.

[21] D. M. Katz, "Quantitative Legal Prediction – or – How I Learned to Stop Worrying and Start Preparing for the Data Driven Future of the Legal Services Industry," *Emory Law Journal*, vol. 2013, no. 62, 2011.

[22] Holmes, D. I., "The Analysis of Literary Style–A Review," *Journal of the Royal Statistical Society. Series A (General)*, vol. 148, no. 4, pp. 328–341, 1985. [Online]. Available: http://www.jstor.org/stable/2981893

[23] Bundesrepublik Deutschland, "juris BMJ - Startseite: Gesetze im Internet." [Online]. Available: http://www.gesetze-im-internet.de/

[24] P. Gerathewohl, "Erschließung unbestimmter Rechtsbegriffe mit Hilfe des Computers: Ein Versuch am Beispiel der angemessenen Wartezeit bei § 142 StGB: Dissertation," Tübingen, 1987.