COMPARISON OF LAW TEXTS: AN ANALYSIS OF GERMAN AND AUSTRIAN LAW TEXTS REGARDING LINGUISTIC AND STRUCTURAL METRICS

Bernhard Waltl[1], Florian Matthes[2]

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

Keywords: ***Legislation Analysis, Complexity Indicators, Language, Linguistics, Text Comparison, Measurement, German Laws, Austrian Laws, Big Data***

Abstract*:* *Understanding legal texts is a non-trivial task. This has various causes arising from different origins, such as structural or linguistic properties of legal texts. In this article, we extend prior results of structural and linguistic metrics, to deepen the understanding of legal texts' complexity. Since complexity can be observed on various levels of legal systems and law texts, we restricted ourselves to the complexity of language, including structural and lexical properties. In particular, we focused on the usage of general clauses, indicated by indeterminate legal terms. Thereby, we differentiated and classified indeterminate legal terms regarding their properties. We performed an analysis on a dataset containing 3 553 German laws, respectively regulations. The papers' result is the extension of an existing complexity indicating set of metrics, including a quantitative, comprehensive and data-intensive analysis in German law texts. Finally, we used the metrics to compare the Austrian and German version of the act governing the liability for a defective product.*

1. Introduction

The complexity of legal systems was addressed by Schuck in 1992 (Schuck 1992). He provided four criteria, which compose legal systems' complexity: density, technicality, differentiation and indeterminacy. Since, Schuck did not intend to quantify the given dimensions; he rather resided on a qualitative level. In 2007 Bourcier and Mazzega distinguished two forms of complexity in legal systems: "structure-based" and "content-based" (Bourcier, Mazzega 2007). Whereas the first one emphasizes the network-like structure of the legal systems, mainly arising through references and quotations, the latter one addresses the (un-)intended effects that laws and changes in laws may have. Legal complexity is known to be a problem. Casanovas et al. stated out in 2007 that "[...] todays complexity of the law stands for the crisis of legal positivism and the dogma of sovereignty" (Casanovas et al. 2007).

There are approaches to improve the understandability of legal texts on various levels. Nevertheless, a comprehensive method to analyze and quantify the quality of legal texts is still missing. This paper is an attempt to quantify the complexity of legal texts (see Section 2). In particular, we analyze legal texts regarding textual vagueness. Thereby, it focusses on linguistic properties, indicating understandability and readability, in German and Austrian laws (see Section 2.3.). Section 3 explains

the used research method in detail and introduces the used set of complexity indicating metrics (see Section 3.3.) and the used dataset, existing of 3 553 German laws. The paper continues with an analysis of German laws (see Section 4). Based on a selection of legal texts, differences in understandability, vagueness and complexity regarding the metrics is shown (see Section 5). Because Austria and Germany share the same language and their legislations have several commonalities, we did a detailed analysis and comparison. Finally, we critically reflect our work (see Section 6), summarize the outcomes and sketch future research directions (see Section 7).

2. Complexity, Vagueness and Comparison of Legal Texts

2.1. Complexity Research for Legal Texts

In order to address the phenomena of textual complexity, one can choose different approaches. It is possible to determine the textual difficulty of a text using text-reader interaction models (Schendera 2004). Schendera argues for a psychological differentiation of readers, covering the spectrum from active-elaborated readers to passive-determinate readers. Whereas the first model describes readers, which are eager to understand the content and critically reflect it, the latter ones are passive and willing to believe everything that is written in the text, without elaborated reflection. Beside of the approaches addressing psychological effects and phenomena, Bane compared different languages regarding their complexity using quantification, i.e. measures (Bane 2008). Based on Shannon's entropy introduced 1948 and the Kolmogorov's complexity from 1965 Bane developed the morphological complexity of languages and texts, which consists of metrics integrating distribution of stems, affixes and signatures of words. The discussion about readability metrics is held widely, and there are some drawbacks, like lack of psychological considerations. However, metrics certainly provide important advantages like predictions and comparison; furthermore, they are commonly used in other domains like military, journalism, health care, etc. A comprehensive overview focusing on advantages and disadvantages was given by DuBay in 2004 (DuBay 2004).

2.2. Vagueness in Legal Texts

Unveiling a words' meaning is in general not a trivial task, especially not to algorithms. Depending on the particular word, humans can be very good at determining the meaning of a word in its context, nevertheless it requires a long learning phase. Although humans can determine the meaning behind the word "bank" easily, there are words, which meaning cannot be determined easily, even if the context is known. A well-known concept in the legal domain are the so-called indeterminate legal terms[1]. Commonly, an adjective is used in combination with a noun, e.g. adequate waiting time (StGB, § 142). §142 of the German criminal code regulates the required waiting of a person involved in an accident. Trivially, the legislation cannot provide a concrete number, specifying the minutes and hours to wait. This is because of the complex nature of the regulated area. § 142 of the criminal code was analyzed and different criteria contributing to the waiting time were identified, using prior judgments. Gerathewohl did a comprehensive analysis in 1987, and although he provided eleven criteria, e.g. daytime, place, damage, …, the determination of the waiting in particular cases remains complex (Gerathewohl 1987).

The phenomena of indeterminate, respectively vague words is well-studied in legal science and it is furthermore a basic and necessary concept in legislation. Hart describes it as the "open texture of law", whereas he argues for its necessity: "[…] the law must predominantly […] refer to classes of person, and to classes of acts, things and circumstances" (Hart, H. L. A 2012). Additionally, Hart states out, that law depends on the capability of language to express general rules, standards, and

[1] The German translation: „unbestimmte Rechtsbegriffe".

principles but cannot exclusively work by giving directives to each individual separately. Prior work shows that vagueness in legal language is indicated by vague words and terms (Bhatia et al. 2005; Mellinkoff 2004; Endicott 2000). In this paper, we are particularly interested in the vague words used within German acts and a possible classification of those (see Section 3.3.2.).

2.3. Comparison of Acts

The comparison of legislations is an accepted method throughout legal sciences (Rusch 2006; Zweigert, Kötz 1996). Different sub-disciplines within the legal sciences, like criminal law and private law, use the comparison of legislations and laws as an additional information source. Based on this extended information basis. it is possible to gain additional insights into various legal domains, such as the style of judgments, the codification process, dependencies etc. Thereby the usage of the comparison of legislations as a method depends on the research interest. In this paper, we compare different but related acts, such as the act governing the liability for a defective product (orig. Produkthaftungsgesetz), regarding linguistic properties. In particular we are interested in the investigation of the laws regarding their vagueness and readability. Other approaches to compare legislations, respectively legal texts, focus on the so-called functional comparison, which aims to investigate differences regarding problems and their solutions in different legislations of countries and cultures (Rusch 2006). The functional aspect starts from an existing problem in society or economy and analyzes the differences and commonalities in the solutions that different countries have produced. Consequently, legal experts are performing these analyses and they have a strong focus on the different functionality of the solutions (Rusch 2006; Rösler 1999).

3. Research Approach: Quantitative Analysis and Comparison

We use quantitative measurements as an inter-subjective comparison decreasing the influence of highly personal and subjective opinions. This paper analyzes the linguistic features in legal texts as well as their comparison. Thereby, a set of linguistic metrics using existing law texts is calculated. The basic method is derived from quantitative linguistics (Köhler 2005).

3.1. Research Objectives

Based on the research method used within this paper, we address several research questions, which we answer by using the quantification of metrics approach, based on a large amount of legal texts. The research questions are as follows:

1. How to objectively measure linguistic properties of laws, such as indeterminacy, vocabulary variety and readability?
2. What are relevant linguistic properties extending the existing set of metrics to represent textual complexity of acts?
3. Can complexity indicators be used to compare different but related acts of distinct legislations, e.g., Germany and Austria, on a linguistic and structural level?

3.2. An Existing Set of Linguistic Metrics

The set of metrics, that we use as a base line for our analysis, was introduced in 2014 (Waltl, Matthes 2014). The nine different metrics were derived from the scientific domain of linguistic and structural network analysis are shown in Table 1.

Name	Abbreviation	Indicated Complexity
Paragraph Count	# §	Linguistic & Structural
Sentence Count	# S	Linguistic & Structural
Word Count	# W	Linguistic & Structural
Structural Depth	D	Structural
Number of outgoing internal references	INT	Structural
Number of outgoing external references	EXT	Structural
Vocabulary Variety	V	Linguistic
Indeterminacy	I	Linguistic
Flesch-Reading-Ease	FRE	Linguistic

Table 1: Metrics indicating linguistic and structural complexity (Waltl, Matthes 2014).

The table shows nine metrics categorized into the complexity category they indicate. The classification differentiates metrics regarding linguistic and structural complexity. Paragraph, sentence and word count contribute to linguistic as well as to structural complexity. In short, this is due to the idea, that a high number of paragraphs, sentences, and words indicate an increased complexity. There can be various origins for the increased complexity and it is unlikely that a single influence factor can be identified. Results from complexity theory have shown that in complex systems it is not reasonable to assume monocausality. It is more appropriate to assume that a larger set of parameters contribute to the emerging behaviour of a system. Consequently, we have different factors leading to the complex system's behaviour.

The table continues with three metrics, contributing to structural complexity. Structural depth describes the tree like structure of the German law. In German laws, it is sufficient to name the act, section, sentence and number, and one can uniquely reference to any part in the law corpus. Outgoing references are distinguished between internal and external. This contributes to the fact, that a reference can refer to a paragraph or section of the same law, then it would be an internal reference, and on the other hand, it can refer to another act, then it is an external reference. If one would determine a graph, representing the network-like structure of the law, then the references would represent the edges between the vertices, representing a law, respectively section.

The remaining three metrics, namely vocabulary variety, indeterminacy, and readability in terms of the Flesch-Reading-Ease, contribute linguistic complexity. The vocabulary variety indicates the number of distinct words that are contained within a law text. The indeterminacy metric counts the number of indeterminate words like adequate, appropriate, etc. This metric also counts indeterminate words, even if their meaning is well known in the context and to legal experts. An example would be term "Absicht" (engl. intention). Legal experts, which are aware of various comments, judgements and legal literature, know the meaning of the word quite well, and can decide what makes an action intentionally. Nevertheless, to lay persons, who are not as experienced as the experts, the term "Absicht" is undetermined, as must not coincide with the understanding that the legislator has about the meaning. Therefore, this can lead to misunderstandings and the purpose and consequence of the law can hardly be understood by lay persons due to the undetermined terms.

3.3. Extending the Existing Set of Metrics

In this paper, we further investigate the capability of metrics in order to determine the linguistic and structural complexity of law texts. Thereby, we extended the existing set regarding the linguistic metrics from Table 1, with vocabulary variety, indeterminacy, and readability.

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" (BMJ 2014). Since we are performing several algorithms we only considered those texts, with at least 200 words, leading to a dataset with 3 553 distinct legal texts.

3.3.1. Vocabulary Variety

In order to investigate a language and to understand the relationship between different parts of speech, it is common practice to count and measure their distribution within a text corpus (Köhler 2005; Ruoff 1981). The linguistics differentiates between several different and distinct categories of words like, nouns, verbs, adjectives, adverbs, prepositions, etc. Nouns are known to be very fundamental in every language, since those are the references to objects and entities in the real world. Therefore, they are the main carrier of information and can together with adverbs or adjectives, make specified statements about objects. Adverbs and adjectives provide then additional information, which can either enhance or restrict the informational content of nouns. A third main category of written or spoken parts-of-speech are verbs. Verbs define actions, occurrence, or state of beings and can therefore describe processes and time-dependent changes of real world objects, i.e. nouns.

According to Ruoff et al. the frequency of nouns, adjectives, respectively adverbs, and verbs – if summed up – make about 44.88 % of the parts of speech used in the German language (Ruoff 1981, pp. 19–26). Thereby nouns make 10.81%, verbs 21.19% and adjectives occur to 12.88% (adverbs included). The distribution was determined on 500 000 German words, which were classified manually. Based on this we counted the three parts of speech that are significantly contributing to the textual information (see Section 4.1.).

3.3.2. Indeterminacy

Gotti summarized his analysis about vague terms in legal texts and proposed a possible classification (Gotti 2005), see Table 2. The table classifies vague terms, which are also called weasel words, into several categories. Using this categorization, we derived three main categories, namely evaluation, quantification and time. These categories are not disjoint in a mathematical sense; consequently, a word could in principle be assigned to multiple categories. E.g., one can assign the word "almost" to the category "Quantification" but also to "Evaluation". This is because the "Evaluation" category in this case is the more general. Nevertheless, the word has to be assigned to the more specific category and if that is not possible, to the more general category. Throughout German acts, we identified 170 different indeterminate terms, which we assigned manually to one category of Table 2. The majority (80%) of words were assigned to "Evaluation", which makes 136 words out of 170.

Category	**Description**	**Example**	**Identified words**
Evaluation	Expression that introduce a number of other interpretations. The evaluation of the term is left to the arbitrator.	allgemein verständlich, anders, angemessen, angestrebt, aufhebbar, ausdrücklich, …	136 out of 170 (80%)

Quantification	Decodification of a concrete number based on words like, some, sufficient amount.	anteilig, ausreichend, genügend, geringfügig, …	15 out of 170 (9%)
Time	Timespans and properties based on flexible words concerning temporal attributes.	dauerhaft, dauernd, demnächst, gleich, häufig, …	19 out of 170 (11%)

Table 2: Differentiation of vague terms (weasel words) in law texts (Gotti 2005).

3.3.3. Readability

As mentioned above, readability metrics are discussed controversially. However, the literature supports the usage of metrics as an objective comparison of related texts (DuBay 2004; Best 2006). Consequently, we extended the usage of a metric, namely Flesch-Reading-Ease (FRE) (Amstad 1978), by two more metrics applicable to the German language, namely the "Wiener Sachtextformel (WSTF)" (Bamberger, Vanecek 1984) and the "Lesbarkeitsindex (LIX)" (Anderson 1981). The explanation of the rationale behind each of every mentioned readability metric would exceed this papers' length, we thereby refer to the original literature.

FRE_{Ger}	=	180 – avg. number of words per sentence – (58,5 × avg. number of syllabus per word)
WSTF	=	0,1935 × (percentage of words > 2 syllabus) + 0,1672 × avg. number of words per sentence + 0,1297 × percentage of words > 6 characters – 0,0327 × percentage of words with 1 syllabus - 0,875
LIX	=	percentage of words > 6 characters + avg. number of words per sentence

Name	**Differentiation**	**Abbreviation**	**Literature**
Vocabulary Variety	Noun	# VV-N	(Ruoff 1981) (DuBay 2004)
	Adjectives or Adverbs	# VV-AA	(Ruoff 1981)
	Verb	# VV-V	(Ruoff 1981)
Indeterminate Words	Evaluation	# I-EVAL	(Bhatia et al. 2005) (Gotti 2005)
	Quantification	# I-QUANT	
	Time	# I-TIME	
Readability	Flesch-Reading-Ease	R-FRE	(Amstad 1978)
	Wiener Sachtextformel	R-WSTF	(DuBay 2004)
	Lesbarkeitsindex	R-LIX	(Anderson 1981)

Table 3: Resulting extension of linguistic metric for legal texts.

4. Applying Metrics to German Laws

Based on the identified metrics, derived from linguistic and legal literature, we applied those to the dataset (see Section 3.3.). The following section summarizes the main outcome and contribution.

4.1. Vocabulary Variety

As stated above, the distribution of different parts-of-speech throughout a language is common practice in linguistics. We analyzed German laws regarding the distribution of nouns, adjectives, respectively adverbs, and verbs. Table 4 shows the results of the computer-supported classification.

Category	POS	Count	Percentage
# VV-N	Noun	2 537 561	23,54 %
# VV-AA	Adjective or Adverb	772 363	7,16 %
# VV-V	Verb	549 988	5,10 %
Rest	Proposition, Particle, Pronoun, Determiner, ...	6 932 151	64,20 %
Total		10 789 063	100 %

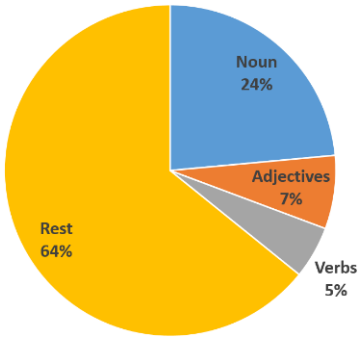

Table 4: Distribution of parts-of-speech throughout the German law

The table states out the distribution with significant differences between ordinary usage of language and the usage of words in law texts. The amount of nouns in law texts is 2.2 times higher than in the ordinary usage of language (see Ruoff 1981). Nouns are known to be representatives of objects in the world, this can be explained by the normative character of acts. According to our interpretation, the increased usage of nouns is indicating the increased interference to real-world objects.

4.2. Indeterminate Words

The usage of indeterminate words throughout acts is a necessary technique to formulate legal norms in an abstract and general way, so that they can be applied to various occasions and cases. Nevertheless, they make the interpretation of legal norms difficult, since they require more effort to be understood and to be interpreted correctly.

	Sum	Percentage
# I-EVAL	172 278	1,60 %
# I-QUANT	8 590	0,08 %
# I-TIME	25 580	0,24 %
Total	206 448	1,91 %

	Average
# I-EVAL Rate	13,38
# I-QUANT Rate	0,66
# I-TIME Rate	1,78
Total	15,82

Table 5: Indeterminate word counts and rates (per 1000 words)

Table 5 shows the determined amount of indeterminate words throughout the German law texts. Not surprisingly, the indeterminate words of the evaluation category is represented highest. This is due the fact that most indeterminate words belong to this category. The right table shows, the average occurrence per 1000 words. In average, 13,38 out of 1000 words are known to be indeterminate. Since we know from our analysis, that the average sentence is about 30 words long, we can say, that on average in every 2,46 sentences one indeterminate word is contained.

4.3. Readability

The usage and impact of Flesch-Reading-Ease as a representative readability metrics in the legal context is discussed in Waltl, Matthes (2014). The extension of expressing readability using three different metrics allows us to analyze whether there are quantitative differences and whether the readability metrics lead to different or comparable results. We therefore determined the readability of every German law, which has more than 200 words and calculated the correlation between the resulting metric values.

	FRE	**WSTF**	**LIX**
FRE	-	-0,776	-0,727
WSTF	-0,776	-	0,979
LIX	-0,727	0,979	-

Name	**FRE**	**WSTF**	**LIX**
Civil Code (BGB)	39,31	14,38	58,88
Penalty Law (StGB)	32,21	14,96	65,40
Capital Investm. Code (KAGB)	8,52	18,03	75,83
Banking Act (KWG)	8,85	18,41	79,50
Liability of Prod. (ProdHaftG)	39,38	14,16	60,71

Table 6: Correlation between the used readability indexes and five exemplary laws and their readability

Table 6 shows the highly significant correlations ($p < 0,01$) between the different metrics. The high correlation values show, that there are only minor differences between the metrics. The negative correlation between FRE and WSTF, respectively LIX, is because the FRE is low for difficult texts, whereas WSTF and LIX have high values for difficult texts. Consequently, only the absolute value is representative.

5. Law Comparison regarding Metrics: Liability of Defective Products

Using the proposed metrics, we analyzed two different but related laws from different countries. We considered the Austrian and the German act governing the liability for a defective product. The council regulation 85/374/EEC is the foundation for both laws. Table 7 shows the comparison of the two laws, regarding the derived metrics.

Metric	**AUT**	**GER**
# W	1 328	1 445
# §	21	19
# S	45	47
EXT	1	6
INT	9	11
R-FRE	41,29	39,38
R-WSTF	14,31	14,16
R-LIX	60,01	60,71
#I	19	28
#I-TIME	1	4

Metric	**AUT**	**GER**
#I-QUANT	17	22
#I-EVAL	1	2
#VV-DISTINCT	371	378
#VV-N	291	321
#VV-N-DISTINCT	139	145
#VV-AA	76	87
#VV-AA-DIST.	52	52
#VV-V	91	106
#VV-V-DISTINCT	54	61

Table 7: Comparison of linguistic and structural complexity metrics for the Austrian and German law

The comparison shows the differences between the two acts. Interestingly, the readability of the Austrian act is higher regarding the Flesch-Reading-Ease and the LIX. However, the metric "Wiener Sachtextformel" is slightly better for the German version. Furthermore, the Austrian act contains less indeterminate words. 28 indeterminate words are contained in the German act, whereas the Austrian only contains 19. The size of the vocabulary used is almost the same, 371 different words in the Austrian and 378 words in the German version. 291 nouns are used in Austrian, whereas 321 nouns are used in Germany. Considering only the distinct nouns, in the Austrian version 139 different nouns and 145 nouns in Germany remain. The same differentiation between total usage and distinct amount of parts of speech is done with adjectives and verbs.

Overall, the German version tends to use slightly more words. Consequently, the amount of distinct words is higher within the formulation of the law. The question whether this makes the law more complex or not cannot be answered by solely analyzing the numbers, but it can be an indicator whether the law can be formulated more accessible, using less words and lower vocabulary variety.

6. Critical Reflection

Although the used method of quantifying linguistic properties can be used for inter-subjective comparison, a few drawbacks remain. The first drawback is the processing of natural language. This task is known to be error-prone. Algorithm determination cannot achieve a perfect accuracy. This circumstance contributes to the detection of nouns, adjectives and verbs. The second drawback concerns the determination of indeterminate words. Even if the meaning of those words could be determined using comments, cases or judge decisions, at first, they are indeterminate and therefore increase the difficulty of interpretation. The third drawback of the paper addresses certainly the usage of metrics in order to determine an act's complexity. Indicators based on linguistic and structural properties can only assist during the analysis phase and point out weaknesses and problematic properties of the law. Since complexity of legal texts has multiple facets, we know that our approach is just one out of several possible and necessary approaches to fully address legal text complexity.

7. Conclusion and Outlook

The paper addresses linguistic phenomena on a large dataset, namely 3 553 German law texts. Thereby it uses quantitative linguistics to measure metrics indicating readability, indeterminacy and vocabulary variety. The metrics extend and existing set of metrics, representing legal complexity, and address properties of texts and their words with respect to vocabulary variety, indeterminacy, and readability. Furthermore, the paper uses the overall set of metrics to compare different but related acts, namely the liability for a defective product.

The insights of the paper address textual quality on a low level with a strong focus on words. As DuBay summarizes the interpretation problem and an approach for a possible solution: "[…] improve the text on the level of words and sentences, the first causes of reading difficulty" (DuBay 2004). This paper could be a baseline for the investigation of evolution of laws and their structural and linguistic complexity. Furthermore, a next research question should be whether the proposed metrics represent the perceived complexity of legal texts.

8. Publication bibliography

StGB (4/23/2014): Strafgesetzbuch. Available online at http://www.gesetze-im-internet.de/stgb/, checked on 11/11/2014.

Amstad, T. (1978): Wie verständlich sind unsere Zeitungen? Zürich: Studenten-Schreib-Service.

Anderson, Jonathan (1981): Analysing the Readability of English and Non-English Texts in the Classroom with Lix. In *ERIC*.

Bamberger, R.; Vanecek, E. (1984): Lesen-Verstehen-Lernen-Schreiben: die Schwierigkeitsstufen von Texten in deutscher Sprache: Jugend und Volk. Available online at http://books.google.de/books?id=TElTAAAACAAJ.

Bane, Max (2008): Quantifying and measuring morphological complexity. In *Proceedings of the 26th West Coast Conference on Formal Linguistics*, pp. 69–76.

Best, Karl-Heinz (2006): Sind Wort- und Satzlänge brauchbare Kriterien der Lesbarkeit von Texten? In Sigurd Wichter, Albert Busch (Eds.): Wissenstransfer - Erfolgskontrolle und Rückmeldungen aus der Praxis. Band 5. Frankfurt am Main: Peter Lang GmbH, pp. 21–31.

Bhatia, Vijay; Engberg, Jan; Gotti, Maurizio; Heller, Dorothee (Eds.) (2005): Vagueness in Normative Texts. Berlin: Peter Lang GmbH (23).

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

Bourcier, Danièle; Mazzega, Pierre (2007): Toward measures of complexity in legal systems. In : ICAIL '07 Proceedings of the 11th international conference on Artificial intelligence and law.

Casanovas, Pompeu; Biasiotti, Maria Angela; Francesconi, Enrico; Sagri, Maria Teresa (Eds.) (2007): Proceedings of LOAIT ’07. II Workshop on Legal Ontologies and Artificial Intelligence Techniques.

DuBay, William H. (2004): The Principles of Readability. Education Resources Information Center.

Endicott, Timothy A. O. (2000): Vagueness in Law: Oxford University Press.

Gerathewohl, Peter (1987): Erschließung unbestimmter Rechtsbegriffe mit Hilfe des Computers. Ein Versuch am Beispiel der angemessenen Wartezeit bei § 142 StGB. Dissertation. Eberhard-Karls-Universität, Tübingen.

Gotti, Maurizio (2005): Vagueness in the Model Law. In Vijay Bhatia, Jan Engberg, Maurizio Gotti, Dorothee Heller (Eds.): Vagueness in Normative Texts. Berlin: Peter Lang GmbH (23).

Hart, H. L. A (2012): The concept of law. Third edition (Clarendon law series).

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

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

Rösler, Hannes (1999): Rechtsvergleichung als Erkenntnisinstrument in Wissenschaft, Praxis und Ausbildung. In *JuS - Juristische Schulung* (11), pp. 1084–1086.

Ruoff, A. (1981): Häufigkeitswörterbuch gesprochener Sprache: gesondert nach Wortarten, alphabetisch, rückläufig alphabetisch und nach Häufigkeit geordnet: Niemeyer (Idiomatica Series).

Rusch, Arnold F. (2006): Methoden und Ziele der Rechtsvergleichung. In *Jusletter* (13). Available online at http://jusletter.weblaw.ch/juslissues/2006/362.html.

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

Schuck, Peter H. (1992): Legal complexity: some causes, consequences, and cures. In *Duke Law Journal*.

Waltl, Bernhard; Matthes, Florian (2014): Towards Measures of Complexity: Applying Structural and Linguistic Metrics to German Laws. In *Jurix 2014: Legal Knowledge and Information Systems.*