DUAL TEXTUALITY OF LAW

Vytautas Čyras / Friedrich Lachmayer

Associate Professor, Vilnius University, Faculty of Mathematics and Informatics
Naugarduko 24, 03225 Vilnius, LT
Vytautas.Cyras@mif.vu.lt; http://www.mif.vu.lt/ cyras/

Professor, University of Innsbruck
Tigergasse 12/12, 1080 Vienna, AT
Friedrich.Lachmayer@gmx.at; http://www.legalvisualization.com

Abstract: *This paper treats legal informatics as a scientific bridge between law and legal documentation. Ten layers are distinguished in the realm of law, e.g., phrase, sentence, sentence formation, and legal act. Four kinds of representation are distinguished in legal documentation (a document representation, a syntactic one, a semantic one, and a pragmatic one). A granularity question is discussed: what is the smallest entity – a document, a sentence or a word? Different software systems are built within each layer. Such applications link different layers of law with different representations.*

1. Textuality of Law

This paper addresses the textual granularity of representation of law. Thus, both law and legal documentation are concerned. The word granularity questions what the smallest entity is, for example, the whole document of a legal source, a sentence, or a word. Furthermore, two standpoints are concerned – law and legal informatics – and building a bridge between these two realms is discussed.

This paper is related to written language, and not speech. Textuality of law deals with different layers; see layers from 1 to 10 below (Figure 1). In legal informatics, layers 4 to 9 are important, whereas layers 1 to 3 are not. Different applications can be built within each of the layers.

Layer 1 – Dot. Dimension is zero. Dot properties such as colour may be important in a technical domain, but not in legal informatics.

Layer 2 – Letter. No semantics are assigned to letters. However, in a technical domain different meanings can be assigned, for example, to letters in different fonts and sizes, such as bold or 10 pt/12 pt.

Layer 3 – Syllable. Syllables are important in acoustics. They form a layer of analysis in speech processing.

Layer 4 – Word. The word is a syntactical entity of the language. The word is the unit of thinking. A word is the smallest element that can be uttered in isolation with practical meaning. The word is the substrate whereas the term is a semantical entity; the term is related with its content. For example, numbers: there are Arabic numbers (1, 2, 3, etc.), Roman numbers (I, II, III, etc.), or numbers on a dice (•, ••, •••, etc.). However, in mathematics they denote the same abstractions. Ontologies are built on words, and the word is the minimal unit to translate.

Layer 5 – Phrase. In linguistic analysis, a phrase is a group of words that functions as a constituent in the syntax of a sentence. Phrases are more important in professional languages and legitimation than in common language. A phrase, its synonyms, and translations can be distinguished.

Layer 6 – Sentence. The sentence is the unit of communication. The sentence can serve as a unit of legal documentation; see, e.g., [REIMER 2008] about the Tasmanian legislation, https://www.legislation.tas.gov.au/. In mark-up languages such as XML, sentences are separated.

Layer 7 – Sentence formation. A sentence formation is a kind of formal sentence configurations, such as abstracts, head notes, and even norms. Sentence formations are not completely standardised in the legal domain. In arts, Haiku is an example. (Haiku is a very short Japanese poem with seventeen syllables and three verses.)

Layer 8 – Form (*Formular*). This is a more standardised formation of fields; see, e.g., [GANTNER 2010]. Forms can be represented in mark-up languages; an example is a paragraph.

Layer 9 – Legal act. Here, speech acts are also comprised. Institutional meaning is assigned on this layer, and as such pragmatics is a characteristic. The meaning of the word «ought» is raised from grammatical meaning (content) to legal meaning. A legal act is a container that comprises its institutional meaning and content meaning.

Layer 10 – Langue. This is a meta-level. Langue is viewed as a linguistic system; one speaks in parole that is ruled by the rules of langue.

Layers 1 to 3 have no content meaning, whereas layers 4 to 8 are attributed to content meaning and therefore are most important for legal informatics. Different software systems can be built with each layer and various informational processes can be computerised. For example, ontologies traditionally concern words (layer 4), but they can also be built on phrases (layer 5).
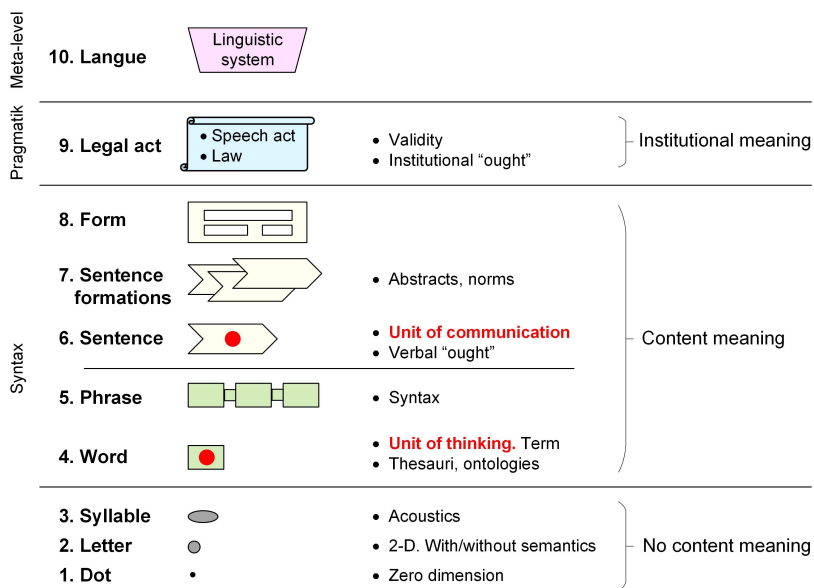

Figure 1: Layers of textuality in law

2. Legal Informatics – a Bridge from Law to its Representation

Legal documentation is not equal to legal informatics. Legal informatics is a science concerned with bridging law and legal documentation. Hence, a transformation «law → legal documentation» is studied. Legal documentation speaks about different representations of law. Here, a duality of word and sentence has to be taken into account: the word is the unit of thinking, whereas the sentence is the unit of communication.

We distinguish the following representations of law:

– A – *Document representation*. The granularity question is a theme: what is the smallest unit – a sentence, a paragraph, or a document?

– B – *Syntactic representation*. Is it represented with a text again, a notation of formal logic, a mark-up language, or a graphical notation?

– C – *Semantic representation*. It deals, among other issues, with the relation between the factual and the normative.

– D – *Pragmatic representation*. It concerns a relation between an act of legal information and a legal act. Treating an authentic electronic document as a legal act is a theme, i.e. 9=D.

Building a bridge from law to legal informatics is about connecting layers 1 to 10 with representations A to D (Figure 2).
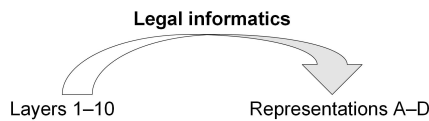
Legal informatics

Layers 1–10   Representations A–D

Figure 2: Legal informatics is a science concerned with building a bridge from law to representation of law

Different links between layers 1-10 and representations A–D can be distinguished; see Figure 3. For example, while speaking about document representation A, different granularity options can be studied: sentence (the link from layer 6 to representation A, 6-A for short), form (link 8-A), or document (link 9-A). While speaking about semantic representations, «if…then…» rules, subsumption and situative representations are tacked. While speaking about pragmatic representation, the theme of authentic electronic documents is tackled (link 9-D, which can be marked with the identity 9=D).
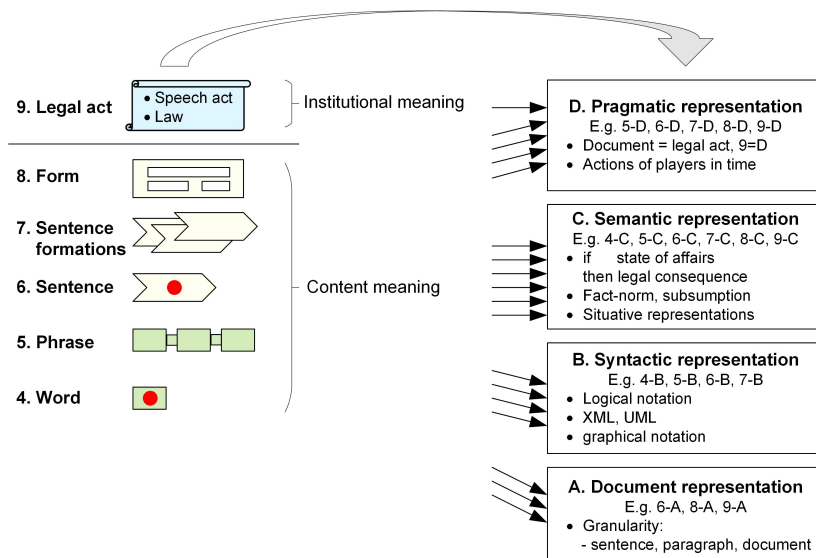
9. Legal act
- Speech act
- Law

Institutional meaning

D. Pragmatic representation
E.g. 5-D, 6-D, 7-D, 8-D, 9-D
- Document = legal act, 9=D
- Actions of players in time

8. Form

7. Sentence formations

6. Sentence

Content meaning

C. Semantic representation
E.g. 4-C, 5-C, 6-C, 7-C, 8-C, 9-C
- if   state of affairs
  then legal consequence
- Fact-norm, subsumption
- Situative representations

5. Phrase

4. Word

B. Syntactic representation
E.g. 4-B, 5-B, 6-B, 7-B
- Logical notation
- XML, UML
- graphical notation

A. Document representation
E.g. 6-A, 8-A, 9-A
- Granularity:
  - sentence, paragraph, document

Figure 3: Links between law and representations of law

3. Schweighofer's 8 Views/4 Methods/4 Syntheses Approach

In this paper, we identified 4 representations of law, A to D, whereas SCHWEIGHOFER [2015] proposes 8. In his 8 views/4 methods/4 syntheses approach, Schweighofer proposes a methodological framework in legal in-

formatics. He has structured newly developed methods for the representation, analysis, and synthesis of legal materials as legal data science. His model describes the 8 different representations of a legal system and 4 computer-supported methods of analysis, which leads to a synthesis, a consolidated and structured analysis of a legal domain, either (1) a commentary, an electronic legal handbook, (2) a dynamic electronic legal commentary DynELC, (3) a representation for citizens, or (4) a case-based synthesis; see Figure 4. The 8 views (or representations of law) are: (1) text corpus, (2) metadata view, (3) citation network view, (4) user view, (5) logical view, (6) ontological view, (7) visualisation view, and (8) argumentation view. The 4 methods are: (1) interpretation (search, reading, and understanding), (2) documentation (search and processing), (3) structural analysis (conceptual and logical), and (4) fact analysis.
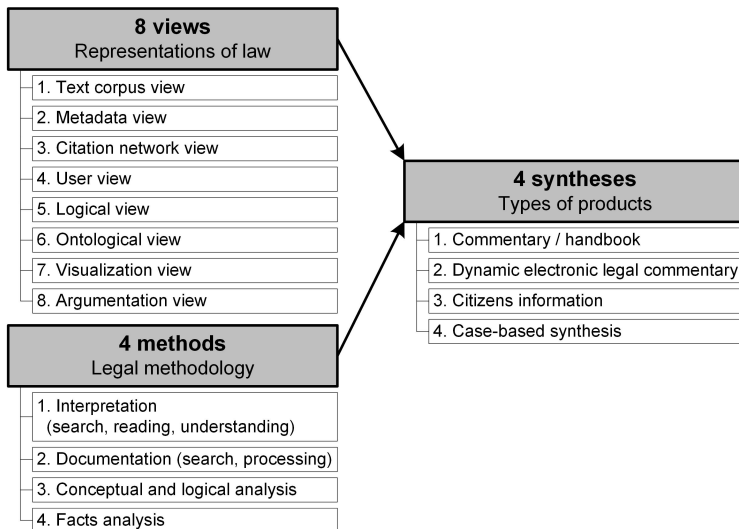
8 views
Representations of law
- 1. Text corpus view
- 2. Metadata view
- 3. Citation network view
- 4. User view
- 5. Logical view
- 6. Ontological view
- 7. Visualization view
- 8. Argumentation view

4 syntheses
Types of products
- 1. Commentary / handbook
- 2. Dynamic electronic legal commentary
- 3. Citizens information
- 4. Case-based synthesis

4 methods
Legal methodology
- 1. Interpretation (search, reading, understanding)
- 2. Documentation (search, processing)
- 3. Conceptual and logical analysis
- 4. Facts analysis

Figure 4: Schweighofer's 8 views/4 methods/4 syntheses model to legal data science

In our classification of the 4 representations A–D, we view from another standpoint than Schweighofer and see no conflict herein. Numerous stakeholders can be identified in the legal domain and legal informatics, therefore perspectives from different standpoints make sense. There are stakeholder views which are not elaborated on in this paper.

4. References

GANTNER, FELIX, Theorie der juristischen Formulare, Schriften zur Rechtstheorie, Heft 252, Duncker & Humblot, Berlin 2010.

REIMER, SEBASTIAN, Rechtssetzung zwischen Österreich und Tasmanien, in: Komplexitätsgrenzen der Rechtsinformatik. Tagungsband des 11. Internationalen Rechtsinformatik Symposions IRIS 2008, Boorberg, Stuttgart 2008, pp. 277–286.

SCHWEIGHOFER, ERICH, From Information Retrieval and Artificial Intelligence to Legal Data Science, in: Schweighofer, E./Galindo, F./Cerbena, C. (eds.), Proceedings MWAIL2015, ICAIL Multilingual Workshop on AI & Law Research, held within the 15th International Conference on Artificial Intelligence & Law (ICAIL 2015), series books@ocg.at vol. 313, OCG, Vienna 2015, pp. 13–23. http://fedora.phaidra.univie.ac.at/fedora/get/o:399570/bdef:Content/get.