Definition Extraction and Annotation in Sources of Law

Master's Thesis
Information Science
University of Groningen

Author: A. Znaor
Supervisor: G.J.M. van Noord
Reader: J. Bos
October 2013
Abstract

Methods to identify, annotate and extract definitions from legal texts are investigated in this thesis. Literature suggests that there is no common model of definitory characteristics in general, and legal definitions in particular. Definitions in law tend to have some distinct properties such as a surface form, coverage and scope of applicability. Therefore, a novel model is proposed to capture these particular characteristics of legal definitions. A possible user interface to present definitions to end-users is also suggested.

Definition extraction experiments are performed on a dataset of Dutch laws. This is firstly done using a rule-based approach where rules were formulated from another dataset. Supervised machine-learning using support vector machines was conducted afterwards in order to improve the rule-based results. The best setup achieved an F-score of 0.72 with precision 0.79 and recall 0.68. Varying the dataset size showed no significant improved in the overall scores.
4.2.2 General and partial definitions ................................................................. 17
4.2.3 Deeming provision .................................................................................. 17
5 Definition model ............................................................................................ 19
5.1 Surface form ............................................................................................... 19
  5.1.1 To-be definitions .................................................................................... 19
  5.1.2 Verb definitions ..................................................................................... 19
  5.1.3 List definitions ....................................................................................... 20
  5.1.4 Name definitions ................................................................................... 20
5.2 Coverage ...................................................................................................... 20
  5.2.1 Complete ................................................................................................. 20
  5.2.2 Broadening .............................................................................................. 21
  5.2.3 Narrowing ............................................................................................... 21
  5.2.4 Exhaustive ............................................................................................... 21
5.3 Scope of applicability .................................................................................. 22
  5.3.1 Default .................................................................................................... 22
  5.3.2 Present law ............................................................................................... 22
  5.3.3 Local ......................................................................................................... 22
5.4 Excluded characteristics ............................................................................. 23
6 Dataset ........................................................................................................... 24
  6.1 Data format ................................................................................................. 24
    6.1.1 Structure ................................................................................................ 24
    6.1.2 Framework ............................................................................................. 24
6.2 Natural language processing ....................................................................... 25
  6.2.1 Tokenisation ............................................................................................ 25
  6.2.2 Sentence boundary disambiguation ....................................................... 25
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2.3</td>
<td>Part-of-speech tagging</td>
<td>25</td>
</tr>
<tr>
<td>6.3</td>
<td>Annotation</td>
<td>26</td>
</tr>
<tr>
<td>7</td>
<td>Experiments</td>
<td>28</td>
</tr>
<tr>
<td>7.1</td>
<td>Candidate extraction</td>
<td>28</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Engine</td>
<td>28</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Patterns</td>
<td>29</td>
</tr>
<tr>
<td>7.1.3</td>
<td>Results</td>
<td>29</td>
</tr>
<tr>
<td>7.2</td>
<td>Machine-learning</td>
<td>29</td>
</tr>
<tr>
<td>7.2.1</td>
<td>Engine</td>
<td>30</td>
</tr>
<tr>
<td>7.2.2</td>
<td>Setup</td>
<td>30</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Error analysis</td>
<td>30</td>
</tr>
<tr>
<td>7.2.4</td>
<td>Results</td>
<td>32</td>
</tr>
<tr>
<td>7.3</td>
<td>Varying dataset size</td>
<td>32</td>
</tr>
<tr>
<td>7.3.1</td>
<td>Setup</td>
<td>32</td>
</tr>
<tr>
<td>7.3.2</td>
<td>Results</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td>Presentation</td>
<td>35</td>
</tr>
<tr>
<td>8.1</td>
<td>Enhanced definitions</td>
<td>35</td>
</tr>
<tr>
<td>8.1.1</td>
<td>Extraction</td>
<td>35</td>
</tr>
<tr>
<td>8.1.2</td>
<td>Automatic annotation</td>
<td>35</td>
</tr>
<tr>
<td>8.2</td>
<td>User interface</td>
<td>36</td>
</tr>
<tr>
<td>9</td>
<td>Conclusion</td>
<td>37</td>
</tr>
<tr>
<td>9.1</td>
<td>Research questions</td>
<td>37</td>
</tr>
<tr>
<td>9.1.1</td>
<td>Definition model</td>
<td>37</td>
</tr>
<tr>
<td>9.1.2</td>
<td>Definition extraction</td>
<td>37</td>
</tr>
<tr>
<td>9.1.3</td>
<td>Presentation</td>
<td>37</td>
</tr>
</tbody>
</table>
1 Introduction

The motive for legal definition extraction and the need and use of definitions in law is discussed in the problem description. Afterwards, the research scope is outlined and research questions formulated. Finally, an overview of the subsequent chapters is given.

1.1 Problem description

When reading legal texts such as laws and contracts, numerous words are encountered that have a specific meaning in the law. By this, it is meant that these words might deviate from common linguistic meaning. Specific definitions are often introduced to counterbalance ambiguities that arise when interpreting such terms. Yet, difficulties might occur when a legal practitioner encounters legislation that is unfamiliar to him or her. It will not always be clear whether a term is defined in the legislation itself or whether case law and legal literature have to be consulted in order to interpret the meaning of the encountered words. Even if a definition is defined in a legal text, it might not be present in the underlying texts, but defined in another source of law.

The main motive of this thesis is to help lawyers in solving these problems. To facilitate this, an information system that automatically extracts definitions is build. In addition to this an experimental graphical user interface which presents the extracted definitions to the end user is implemented.

1.2 Research scope

During the writing of this thesis, one of the difficulties encountered was the exact ‘definition’ of definitions in legal texts. At first, an intuitive distinction was made between sentences (and broader units such as articles) that did and did not contain definitions. For most sentences it was clear what had been defined, while others contained constructions where it could not be decided whether or not they contained definitions.¹ This uncertainty led to a deeper examination of the theoretical background of definitions and this extended the initial research scope. Three research questions emerged from this:

1. What constitutes a legal definition and how can legal definitions be modelled?
2. What is the performance of automatic definition extraction methods for legal texts?

¹ Fahmi and Bouma (2006) choose to label some sentences as ‘undecided’ when unsure whether they contained a definition.
3. How could defined terms be used to annotate legal texts using the extracted definitions?

The first question is answered using theoretical insights, whereas the second question is empirically tested. The third question is answered by demonstrating a possible user interface for end users.

A novel model to represent definitions in legal texts is described by using insights from philosophical and linguistic definition theories. Methods from legal theory are also examined and incorporated into this model. The novel model integrates these different views in a coherent theory about legal definitions. This means that each definition has a surface form, coverage and scope of applicability. By using this model, automatic definition extraction experiments are conducted. The retrieval score of these experiments is reported to answer the possible usefulness of such a system.

1.3 Thesis overview

Chapter 2 Related definition extraction work is investigated.

Chapter 3 The basic structure of definitions is examined, alongside classification from a philosophical and linguistic stance.

Chapter 4 The peculiarities of legal language and legal definitions are shown.

Chapter 5 A novel model is proposed to represent legal definitions. This consists of their surface form, coverage and scope of applicability.

Chapter 6 The data format of a dataset of Dutch law is described. These texts are afterwards processed and annotated to be suitable for the definition extraction experiments.

Chapter 7 The engine, setup and results of the definition extraction experiments are presented and discussed.

Chapter 8 Enhancing markup and presentation of the extracted definitions is discussed. An experimental user interface is shown.

Chapter 9 In conclusion, the research questions are answered alongside suggestions for further research.
2 Related work

Related work on definition extraction has been studied by both Natural Language Processing (NLP) practitioners and researchers in the Law & Artificial Intelligence (Law & AI) community. Whereas NLP research has focused on automatic definition extraction in general texts, the Law & AI community has primarily focused on modelling legal reasoning and inference. An overview of some endeavours in both fields is given below.

2.1 Definition extraction

Automatic definition extraction has been addressed using pattern-based and machine learning approaches (E. N. Westerhout, 2010, p. 24). Joho and Sanderson (2000) use regular expressions to extract descriptive phrases, while Fahmi and Bouma (2006) implement several machine learning methods to extract definitions from parsed text. A similar approach is conducted by Westerhout (2010) where definition patterns found by regular expressions are filtered through machine learning methods to minimise false positives. A rule-based approach to extract definitions from German court decisions has been used by Walter and Pinkal (2006).

2.2 Legal modelling

Maat and Winkels (2010) have developed a methodology to model fragments for sentences in Dutch law. Their approach uses a frame-like representation for various models. Wyner and Peters (2011) use similar methods to extract rules from regulations. Their model includes an agent and theme, deontic modals and verbs, main verbs and exception clauses. Translation of legal texts into First-order Logic (FOL) formulas has been attempted by Wyner et al. (2012).

2.3 Findings

The work examined offers a firm foundation to base research on. This is especially true for the definition extraction literature in which the extraction is extensively described. Legal modelling research offers a useful insight for definition modelling. However, these models do not account

\[ \text{\footnote{For an overview of research in this area see Burns (2007, Chapter 1) and a review of her dissertation in Mountain (2010). The term Law & AI has been described using various terms "including information technology for lawyers, artificial intelligence and law, legal informatics and the computerisation of law" (Burns, 2007, p. 16). This field is known in Dutch as 'rechtsinformatica'.}} \]
for some specific properties of legal definitions. These shortcomings are overcome by proposing a novel definition model (Chapter 5).
3 Definitions

The nature of definitions, starting with Plato and Aristotle, has been discussed for nearly two and a half millennia (Robinson, 1968, p. 1). Since then, contemporary writings about this subject have been spread across philosophy, logic, cognitive science and linguistics. This chapter starts with examining the basic structure of definitions. Then, philosophical and linguistic categorisations are analysed. Investigation of these categories was helpful for devising the definition model described in Chapter 5.

3.1 Basic structure

Before elaborating on the different kinds of definitions, the basic structure of definitions is explained. Textbooks state that every definition consists of at least a definiendum and a definiens. The definiendum is that what is to be defined, whereas the definiens is that what fixes meaning to the definiendum. A typical definition is given in Example (1). The word “LAKE” is the definiendum, and “LARGE, LANDLOCKED, NATURALLY OCCURRING STRETCH OF WATER” the definiens.

(1) A LAKE IS A LARGE, LANDLOCKED, NATURALLY OCCURRING STRETCH OF WATER.  

One of the main characteristics of definitions is the interchangeability between definiendum and definiens. This means that usage of definitions can reduce the total amount of text by avoiding repetition. This is especially useful for texts where the defined term is frequently mentioned after it has been defined.

3.2 Philosophical classification

Most authors begin their elaboration on definitions by examining classical texts, especially Aristotle. He stated that a definition is a phrase indicating the essence of something (Parry & Hacker, 1991, p. 80). Since then, other writers have made an attempt to classify them, but there is no author that has provided a complete overview of definition types (E. N. Westerhout, 2010, p. 33).

3 Westerhout (2010) also considers the connector an essential element. Her hypothesis is examined in §3.3.

3.2.1 Real and nominal definitions

One of the typical distinctions is between so-called nominal definitions and real definitions. The former define a ‘word’, whereas the later define a ‘thing’. A ‘word’ is a word, phrase or conventional sign, a ‘thing’ anything else (Parry & Hacker, 1991, p. 83). Gupta (2012) states that “to discover the real definition of a term X one needs to investigate the thing or things denoted by X; to discover the nominal definition, one needs to investigate the meaning and use of X”. A real definition is given in Example (2), whereas a nominal definition is given in Example (3).

(2)  WATER IS H₂O.

(3)  “P.N.” IS AN ABBREVIATION FOR “PROMISSORY NOTE”.

Robinson (1968) has further divided nominal definitions in purpose-based and method-based, and Westerhout (2010) has heavily drawn upon this classification. Parry and Hacker (1991) have made a similar differentiation between functional types of definitions and ways of defining. This classification is also followed in this chapter, but only definition types that are considered relevant for legal texts are shown.

3.2.2 Purpose-based definitions

Lexical definition

A lexical definition (also known as dictionary definition or reportive definition) is “a nominal definition intended to report a conventional meaning of the definiendum, that is, to report a meaning established for some group of users of the language” (Parry & Hacker, 1991, pp. 90–91). These kind of definitions prevail in dictionaries.

Stipulative definition

A stipulative definition (also known as working definition or operational definition) is “a definition that tells what one intends a ‘word’ to mean” (Parry & Hacker, 1991, pp. 91–92). These kind of definitions are context-bound and are of great importance in law.

---

5 Both examples taken from Parry and Hacker (1991, pp. 82–83).
3.2.3 Method-based definitions

Westerhout (2010) has distilled eight different method-based definition classes. Only the difference between intensional and extensional definitions will be discussed here. A complete overview is given in Appendix A.

3.2.3.1 Intensional definitions

Parry and Hacker (1991, p. 102) consider an intensional definition (also known as coative definition) a definition whose definiens is a set of properties that is intended to be the intension of a concept or word.

Aristotelian definition

The most typical example of an intensional definition is the Aristotelian definition (also known as analytical definition or definition per genus proximum et differentiam specificam). The definiendum is hereby defined according to its category (genus) and difference between it and other terms in that category (differentiae). In the earlier Example (1), “STRETCH OF WATER” is the genus and “LARGE”, “LANDLOCKED” and “NATURALLY OCCURRING” the differentiate.

Synonymous definition

The notion of a synonymous definition is quite intuitive. Another word is given that can be used instead of the definiendum (Example (4)).

(4) GRANDMA MEANS THE SAME THING AS GRANDMOTHER.

If the definiendum and definiens originate from different languages this type of definition is sometimes called a translational definition. Duk (1999) speaks of an abbreviating definition if the definiendum abbreviates the term(s) in the definiens.

3.2.3.2 Extensional definition

Extensional definitions (also known as denotative definitions or definitions by example) are definitions made primarily or entirely by indicating individual objects in the extension of the definiendum (Parry & Hacker, 1991, p. 113).

---

6 Duk (1999, p. 14) calls this generalising definition.
**Ostensive definition**

When some or all of the objects denoted by the definiendum are produced, presented or shown we speak of an *ostensive definition* (also known as *exemplifying definition* or *denotative definition*) (Parry & Hacker, 1991, p. 114). This technique of defining is shown in Example (5). The meaning of “FRUIT” is conveyed by pointing out examples of things that are considered instances of the definiendum.

(5) **A FRUIT IS A THING SUCH AS AN APPLE, BANANA OR ORANGE.**

**Enumerative definition**

An *enumerative definition* (also known as *specifying definition*) lists all possible things that fall under the definiendum.7 These sorts of definitions are often encountered in law. This type of definition is used in Example (6) to define public holidays. Sometimes, enumerative definitions are explicitly written down as lists.

(6) **GENERALLY RECOGNISED HOLIDAYS FOR THE PURPOSE OF THIS LAW ARE:**


**Recursive definition**

Duk (1999, p. 15) mentions *recursive definitions* (also known as *inductive definitions*) as definitions that somehow resemble enumerative definitions with the exception that the species are infinite or undefined. Example (7) gives a (fictional) definition of descendants.

(7) **SOMEONE'S DESCENDANTS ARE:**

   A. EACH OF HIS CHILDREN;
   
   B. THE DESCENDANTS OF EACH OF HIS CHILDREN.

---

7 Loth (1991) uses the term *enumerative definition* (in Dutch: 'opsommingsdefinitie'), while Duk (1999, pp. 14–15) uses *specifying definition* (in Dutch: 'specificerende definitie'). According to the description, it can be assumed that they refer to the same concept.
3.2.3.3 Other definitions

Two other types of definitions are the precising definition and the persuasive definition. They are mentioned separately because they do not fall under the distinction of intensional or extensional definitions.

Precising definition

Hurley (2002, pp. 95–96) mentions precising definitions as those definitions whose purpose is to reduce the vagueness of words. He notes that “[t]hey resemble stipulative definitions, but differ from a stipulative definition in that the latter involves a purely arbitrary assignment of meaning, whereas the assignment of meaning in a precising definition is not at all arbitrary”.

Persuasive definition

The persuasive definition is used to persuade a particular meaning to a term. This kind of definition is not encountered in written law, but may be used by attorneys to convince judges to explain a term in a particular way.

3.3 Linguistic classification

Linguistic classification of definitions deal with the word types used to categorise utterances as definitory. Westerhout (2010) and Walter and Pinkal (2006) have made such classifications.

3.3.1 Westerhout

Let us recall that the constituents of a definitions are at least the definiendum and definiens. Westerhout (2010, pp. 16–17) also considers the connector an essential part of definitions. This is mostly a verbal phrase or punctuation character that relates the definiendum and definiens. In Westerhout and Monachesi (2007) she has distinguished five types of ‘definitory contexts’ which are grouped according to the connector used.8 These are definitory contexts:

1. in which a form of the verb ‘to be’ is used as connector verb;
2. in which other verbs are used as connector;
3. having specific punctuation features;
4. in which the layout plays an important role;

8 There is also a category ‘other’ that is used as a rest category.
5. in which relative and demonstrative pronouns are used to point back to an earlier used defined term.

This classification is used as an inspiration for the ‘surface forms’ of legal definitions in §5.1.

3.3.2 Walter

Another attempt to linguistically classify definitions was carried out by Walter (2008) and Walter and Pinkal (2006). They classified German case law texts into four classes:

1. **classificatory**: e.g. copular sein – be, classificatory verbs such as **fallen unter – fall under** or (less neutrally) **gelten als – be considered as**;

2. **meta-linguistic**: used to speak directly either about word meaning or conditions of applicability (e.g. **bedeuten – mean** or **vorliegen – 'be existent'**);

3. **interpretational**: referring to aspects important to the process of legal interpretation (e.g. **fordern – require, darstellen – constitute**);

4. **feature-specific**: naming a specific type of feature used in the respective definition (e.g. **dienen zu – serve as, schützen – protect**).

Here the first class roughly resembles the ‘to be’ and ‘verb’ classification of Westerhout (2010). Definitions in the second class could also fall under the ‘verb’ class. The third and fourth class point out how terms should be used rather than defining them by themselves. They should therefore not be considered as real definitions in my opinion.
4 Definitions in law

Many laws and contracts start with defining particular terms. These definitions determine the use of a particular phrase in the context of the legal source being referred to. They can be thought of as overruling, and often narrowing or broadening, the common sense linguistic meaning of words.

4.1 Legal language

Before we dive deeper into the peculiarities of legal definitions, a general note about legal language is given.9 According to Termorshuizen-Arts (2003, p. 30) legal language can be described as general language in which legal concepts are embedded. These legal concepts can be thought of as concepts with special, legal meaning. Although these terms often originate from ordinary language, their meaning is in general different from language used outside the law. Thus, legal concepts often concretise or differentiate the meaning of general concepts. Lawyers sometimes want to fixate them to facilitate legal certainty. This is where legal definitions come into play.

4.2 The nature of legal definitions

Philosophical and linguistic characteristics of definitions, as seen in previous classifications, are not adequate to fully explain the nature of legal definitions. The law is an autonomous system with its own rules and corresponding peculiarities.

4.2.1 Formulation

The need for definitions in law can arise because of various reasons. Hospers (Hospers, 1997, p. 13) states that “[o]ordinarily we don’t attempt to give the definition of a word unless some dispute arises involving its use”. Lawmakers will especially want to explicitly fix meaning to particular words in legal texts to restrict ambiguity.10

For example, in the Netherlands principles for legislative drafting have been written down in the ‘Guidelines for Legislation’.11 These instructions enforce uniformity during law-making and civil servants are obliged to follow them when drafting legislation. Recommendation 121 of the Guidelines states that terms that are too specific or deviate too much from

---

9 Tiersma (2010, p. 32) describes law itself as a “textual enterprise”.
10 Restrict, not eliminate, as both natural language and legal concepts are inherently ambiguous.
regular meaning should be defined. These definitions should not diverge significantly from normal language. Due to their nature, there exist many ways to define them. Walter (2008) states that

“[o]ne important reason for this high degree of variability in formulation lies in the specific role of definitions in court decisions. Scientific and technical terminology is often built up using more or less context-free general definitions assigning new terms to places within a given taxonomy. In contrast, defining statements in verdicts are parts of coherent texts and do not only serve as specifications of terms, but also as arguments for or against their application in a specific case.”

We can see that their meaning should be based on regular language, but on the other hand also serve an autonomous purpose. The intent is twofold: explain meaning and provide a framework for lawyers to work with. This is achieved by their authoritativeness of being issued by a legislator or court.

4.2.2 General and partial definitions

After the need for introducing new terms arises, a general or partial definition may be given (Eijlander & Voermans, 1999, pp. 232–238). If a definition captures the whole meaning of a term, we speak of a general definition. If the meaning of an existing terms is broadened or narrowed, we speak of a partial definition.13

4.2.3 Deeming provision

Concepts that somehow resemble definitions are deeming provisions. These kind of legislative building blocks are introduced when the lawmaker wants to link different parts of legislation or create a legal fiction. Eijlander and Voermans (1999, pp. 242–243) distinguish between legal identity and legal identification. The former is used when two ‘factual entities’ (in Dutch: feitencomplexen) are declared the same; the latter when they ought to be treated the same. A deeming

---

12 Aanwijzing 121

1. Termen die een te weinig bepaalde of een van het spraakgebruik afwijkende betekenis hebben, worden gedefinieerd.
2. In een begripsbepaling wordt aan een term geen sterk van het normale spraakgebruik afwijkende betekenis gegeven.

13 De Maat (2012, p. 52) uses the term type extensions for these kind of definitions.
provision containing a legal identity is shown in Example (8), whereas legal identification is shown in Example (9).

(8) IN THIS LAW AND THE PROVISIONS BASED ON IT IS UNDERSTOOD:

(...)

C. FAMILY:
1°. THE MARRIED TOGETHER;
2°. THE MARRIED WITH THEIR DEPENDENT CHILDREN;
3°. THE SINGLE PARENT WITH HIS DEPENDENT CHILDREN;

(9) FOR THE PURPOSES OF STATUTORY REGULATIONS GOVERNING OBJECTIONS AND APPEALS, THE FOLLOWING SHALL BE EQUATED WITH AN ORDER:

A. A WRITTEN REFUSAL TO MAKE AN ORDER, AND
B. FAILURE TO MAKE AN ORDER IN DUE TIME.

Constructions such as the above provide a convenient way to equate terms and concepts whose meaning differs. However, the distinction between legal identity and other types of definitions cannot always be sharply drawn. Looking at philosophical classifications of §3.2 it could be argued that legal identity qualifies as an enumerative definition. Legal identification should not be considered a definition in my opinion.
5 Definition model

A novel model is proposed to annotate definitions in the dataset. This is necessary because existing classifications do not capture all characteristics of legal definitions. The model was constructed by looking at legislation that is different from the gold-standard corpus used in the extraction experiments (Chapter 7). Each definition within this framework has three attributes: a surface form, coverage and scope of applicability.

5.1 Surface form

Four types of definitions are distinguished according to their surface form. These are to-be definitions, verb definitions, list definitions and name definitions. Surface forms resemble linguistic classifications as seen in §3.3. They describe the word order of utterances that classify as definitions.

5.1.1 To-be definitions

The to-be definitions roughly correspond to the first type of Westerhout and Monachesi (2007) and the ‘classificatory’ class of Walter (2008). They consist of the verb ‘to be’ in the typical form [DEFINIENDUM] IS/ARE [DEFINIENS] as seen in Example (10).

(10) ‘THINGS’ ARE TANGIBLE OBJECTS THAT CAN BE CONTROLLED BY HUMANS.

5.1.2 Verb definitions

As verb definitions are considered definitions that contain other verbs than the verb ‘to be’ as a connector. By taking in account the different theories and classifications of definition, four verbs have been distilled that indicate these. This is only achieved by using the past participle of the verbs CONSIDERED, REGARDED, COMPREHENDED and UNDERSTOOD (Example (11)).

(11) AS A FOSTER CHILD IS CONSIDERED THE CHILD THAT IS MAINTAINED AND BROUGHT UP AS ITS OWN CHILD.

Other verbs such as EQUATED and DESIGNATED are not considered to indicate definitions. Formulations such as these qualify as deeming provisions. However, as has been noted before, the line between deeming provisions and definitions can sometimes be hard to draw.
5.1.3 List definitions

List definitions consist of a header and subsequent individual definitions. The header mostly begins with the typical formulation “IN THIS LAW (AND THE PROVISIONS BASED ON IT) IS UNDERSTOOD:” which explicitly states the scope of applicability for the defined terms. Individual definitions consist of the form [DEFINIENDUM]: [DEFINIENS] (Example (12)). Such formulations provide a convenient way to group all definitions at the beginning of a legal act or at the beginning of chapters and sections.

(12) IN THIS LAW AND THE PROVISIONS BASED ON IT IS UNDERSTOOD:

(...)

B. PRODUCER OF A DATABASE: THE PERSON WHO BEARS THE RISK OF THE INVESTMENT FOR THE DATABASE;

(...)

5.1.4 Name definitions

When a phrase of the form “UNDER THE NAME [DEFINIENDUM]” occurs it will be considered a name definition (Example (13)). Such constructions somehow resemble verb definitions, but the four distinctive verbs (CONSIDERED, REGARDED, COMPREHENDED and UNDERSTOOD) are not used.

(13) UNDER THE NAME TAX ON TAP WATER IS A TAX LEVIED ON TAP WATER.

5.2 Coverage

All definitions are either complete, broadening, narrowing or exhaustive according to their coverage. This is a novel distinction of this model that has not been described by other authors.

5.2.1 Complete

Most of the examined definitions are so-called complete definitions. By this are meant definitions in which the definiendum is a novelty and does not derogate or broaden the meaning of terms (Example (14)).

(14) UNDER TRADE NAME IS UNDERSTOOD BY THIS LAW THE NAME UNDER WHICH AN ENTERPRISE IS DRIVEN.
We have seen in §4.2.2 that Eijlander and Voermans (1999) speak of general definitions in this case. The broadening, narrowing and exhaustive definitory types could be considered partial definitions according to their classification. Complete definitions are also a residual category for definitions that cannot be placed into another category.

5.2.2 Broadening

Broadening definitions are definitions in which the definiens broadens the meaning of the definiendum (Example (15)). This is often expressed with the word also.14 An earlier complete definition might as well be given, but this is not mandatory.

(15) UNDER EARLIER CONVICTION IS ALSO UNDERSTOOD A CONVICTION BY A CRIMINAL COURT IN ANOTHER MEMBER STATE OF THE EUROPEAN UNION FOR SIMILAR FACTS.

5.2.3 Narrowing

Definition statements can, as well as including, also exclude concepts, thus narrowing them. This is the case with formulations such as “UNDER [DEFINIENDUM] IS NOT UNDERSTOOD” (Example (16)). Another common expression is “notwithstanding”.15 As in the case of broadening definitions, an earlier complete definition of the term might already be given.

(16) AS SEPARATE DISCLOSURE IS NOT UNDERSTOOD THE RETRANSMISSION OF A PROGRAMME BY THE SAME ORGANISATION THAT MADE THE ORIGINAL BROADCAST.

5.2.4 Exhaustive

When a definiens is said to apply “IN ANY CASE”16 to a particular definiendum, we speak of an exhaustive definitions. Definitions containing such formulations assert the exclusive applicability of the definiens (Example (17)). In other words, when the situation described in the definiens occurs, it shall always be classified as the definiendum.

---

14 In Dutch: ‘mede’.
15 In Dutch: ‘in afwijking van’.
16 In Dutch: ‘in ieder geval’.
In the case of the example, this means that “THE DIRECT GUIDANCE REGARDING THE MEDICAL, NURSING AND ECONOMIC AFFAIRS OF THE HOSPITAL FACILITY” shall always qualify as a “DALLY BOARD”. However, another “DIRECT GUIDANCE” inside a hospital might also qualify as such.

5.3 Scope of applicability

Definitions within sources of law contain a scope of applicability which tells us where the defined term is used. This is often bound to the present law or the definition’s use is only local. If the scope is not explicitly declared we speak of a default scope and the definition might apply to legislation outside the one in which it is mentioned.

5.3.1 Default

A definition has a default scope when the applicability is not explicitly stated. Usage is inferred from context, meaning and other factors known to legal practitioners. In such a case, a legal scholar will mostly determine the pertinence by putting the current legislation in perspective of the whole legal system, and by accessing case law and literature. A typical example is given in the earlier Example (11) where the scope of “FOSTER CHILD” is implicit, thus initially unrestricted. This might make it applicable in other legislation.

5.3.2 Present law

A common scope is the restriction of a definitional phrase to the present law. This is accomplished by restricting the scope to this law such as in Example (14). The construction “THIS LAW AND THE PROVISIONS BASED UPON IT” is also seen, which indicates that terms mentioned in lower legislation do not need separate definitions.

5.3.3 Local

When a definition only applies to a smaller subset of the present law, we speak of a local scope. Such a scope is typically restricted to a single article or heading inside a law (Example (18)). This scope of applicability is convenient when the same term has to be defined differently across the same legal act.
5.4 Excluded characteristics

Although the model captures the most important characteristics of legal definitions, it does not account for the vast amount of definitionary properties as discussed in Chapter 3. This means that common semantic relationships between the definiendum and definiens are not included. This is the case with hyponymy (is-a relations) and meronymy (part-of relations). For instance, in Example (6) every listed holiday is-a “GENERALLY RECOGNISED HOLIDAY” (hyponymy).
6 Dataset

A corpus of Dutch laws is analysed and its data format described. This dataset is afterwards processed to be suitable for manual annotation. This is necessary for the later definition extraction experiments.

6.1 Data format

The structure of the original dataset is examined and the findings of it described. Subsequently, the text manipulation framework is presented.

6.1.1 Structure

At the time of writing, nearly all legislation in the Netherlands was available as structured XML-files. The structure of these electronic documents is marked up according to the Basis Wetten Bestand (BWB) criteria. This is an official content model that contains principles and rules for consolidating and maintaining legislation files. Relations among various legal documents and distinct versions are defined in the model (Redactie Wetgeving, 2011). This predefined markup relieves the burden of low-level parsing of the structure of documents such as recognising headings and lists.

Each file consists of a root node <wetgeving> with a unique attribute bwb-id that identifies the legislation. Afterwards, the title and provisions follow. These provisions consists of a preamble, the main section and a closure with signatures of the ministers. Text within this main section is always contained in an <al> tag. Internal and external references which point to other named articles in the same text or another document are explicitly marked. Articles are the primarily content unit which can reside within one or more headings. Other elements such as lists are also explicitly declared.

6.1.2 Framework

The unedited laws are loaded as structured XML-files and preprocessed within the General Architecture for Text Engineering (GATE) (Cunningham, Bonceva, & Maynard, 2011). This is a text processing software suite which incorporates ready-made components to process texts in various ways. A distinct feature of this program is the use of stand-off markup where multiple

---

17 Exceptions are legal documents issued by lower administrative bodies such as municipalities.

18 http://koop.overheid.nl/producten/basis-wetten-bestand
layers of annotation can be added to the same text (Wilcock, 2009, pp. 11–18). This is achieved through serialising the XML-tags. Such an approach is in contrast with traditional in-line markup where text and annotations are mixed together in the same document.

6.2 Natural language processing

The definition annotation task is implemented as an NLP-pipeline where the output of one process is used as input for the next one. The documents are processed to be suitable for the definition extraction task. This consists of tokenisation, sentence boundary disambiguation and part-of-speech tagging.

6.2.1 Tokenisation

Tokenisation is performed using the ‘GATE Unicode Tokeniser’ which is a generic, not language specific tokeniser that splits words using a syntax that resembles regular expressions. Each sequence of characters that is identified as a token is assigned features such as length, orthography and kind (word, punctuation, number). The tag name used for tokens is `<Token>`, whereas space and line-breaks are marked by `<SpaceToken>`.

6.2.2 Sentence boundary disambiguation

The text is segmented at the sentence level and every sentence found is marked with the tag `<LegalSentence>`. Sentence boundaries are detected by a grammar that considers every full stop (.) as a split. This is justified by the fact that, in general, no abbreviations and other anomalies are encountered in legal texts. Colons (:) and semicolons (;) are also used to indicate sentence boundaries in lists. Such an approach does not always produce grammatically correct splits but is necessary to tackle long sentences. Otherwise, the part-of-speech tagger would have a hard time analysing sentences containing hundreds of words.

6.2.3 Part-of-speech tagging

All legal sources are fully parsed using the Alpino parser for Dutch. This is a “wide-coverage computational analyser of Dutch which aims at full accurate parsing of unrestricted text, with coverage and accuracy comparable to state-of-the-art parsers for English” (G. van Noord & Malouf, 2004). During parsing, Alpino outputs dependency-nodes which are assigned part-of-

---

speech (POS) tags. Following Westerhout (2010) only the POS output is used to extract definitions.20

Because Alpino only accepts single sentences per line as input, every sentence in a document had to be transformed to the desired format (G. J. M. van Noord, 2013). Each line consists of an identifier, followed by a pipe (|). This was achieved by writing a Groovy script that could be embedded into the GATE pipeline. To improve the speed of the parser, only a single (first) parse is used.

6.3 Annotation

Seventeen laws containing definitions were chosen from a larger set of Dutch legislation. From this, a gold-standard corpus was manually annotated by marking each sentence containing a definition. A sentence was only considered as such if it is a ‘legal sentence’ found during the sentence disambiguation task. The definitions were annotated only if they fitted into the definition model of Chapter 5.

A total number of 145 definitions were found in the dataset. This means that each law contains on average 8.53 definitions. As can be seen in Table 1, these definitory sentences constitute between 2% and 12% of all sentences in a particular document.

<table>
<thead>
<tr>
<th>Name</th>
<th>Number of definitions</th>
<th>Number of sentences</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algemene Kinderbijslagwet</td>
<td>14</td>
<td>259</td>
<td>0.05</td>
</tr>
<tr>
<td>Algemene termijnwet</td>
<td>1</td>
<td>16</td>
<td>0.06</td>
</tr>
<tr>
<td>Boxwet</td>
<td>6</td>
<td>53</td>
<td>0.11</td>
</tr>
<tr>
<td>Dienstenwet</td>
<td>26</td>
<td>278</td>
<td>0.09</td>
</tr>
<tr>
<td>Handelsnaamwet</td>
<td>3</td>
<td>50</td>
<td>0.06</td>
</tr>
<tr>
<td>Kaderwet zelfstandige bestuursorganen</td>
<td>2</td>
<td>124</td>
<td>0.02</td>
</tr>
<tr>
<td>Paspoortwet</td>
<td>17</td>
<td>325</td>
<td>0.05</td>
</tr>
<tr>
<td>Prijzenwet</td>
<td>4</td>
<td>33</td>
<td>0.12</td>
</tr>
<tr>
<td>Reglement van orde voor de ministeraad</td>
<td>3</td>
<td>91</td>
<td>0.03</td>
</tr>
<tr>
<td>Wet giraal effectenverkeer</td>
<td>11</td>
<td>209</td>
<td>0.05</td>
</tr>
<tr>
<td>Wet Infrastructuurfonds</td>
<td>11</td>
<td>102</td>
<td>0.11</td>
</tr>
<tr>
<td>Wet melding collectief ontslag</td>
<td>6</td>
<td>65</td>
<td>0.09</td>
</tr>
<tr>
<td>Wet op de kansspelen</td>
<td>21</td>
<td>446</td>
<td>0.05</td>
</tr>
<tr>
<td>Wet op de parlementaire enquête 2008</td>
<td>5</td>
<td>161</td>
<td>0.03</td>
</tr>
</tbody>
</table>

20 We could also have used a POS tagger which does not provide a full parse. However, Alpino was chosen due to the high correctness of the POS tags.
<table>
<thead>
<tr>
<th>Law Description</th>
<th>ID</th>
<th>Characters</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet op de overkoepse</td>
<td>1</td>
<td>10</td>
<td>0.10</td>
</tr>
<tr>
<td>Wet tot beheer van cultuurbezit</td>
<td>11</td>
<td>114</td>
<td>0.10</td>
</tr>
<tr>
<td>Winkeltijdenwet</td>
<td>3</td>
<td>54</td>
<td>0.06</td>
</tr>
</tbody>
</table>

*Table 1: Annotated laws in the dataset*
7 Experiments

Following Westerhout (2010) and Fahmi and Bouma (2006), definition extraction is approached as a two-track problem. Firstly, candidate extraction patterns are identified that could refer to definitions. Secondly, machine learning techniques are used to filter out false positives to maximise precision, while retaining high recall. Finally, variation in dataset size is examined to show the effect of adding more documents to train on.

7.1 Candidate extraction

Various articles in approximately thirty laws, not contained in the dataset, were examined for definition patterns. The extraction engine is presented and extraction patterns discussed. At last, the results are presented.

7.1.1 Engine

Definition extraction patterns are implemented using the Java Annotation Patterns Engine (JAPE) inside GATE. This is a finite state transducer over annotations based on regular expressions (Cunningham et al., 2011). Every JAPE grammar consists of a set of phases, each of which consists of a set of pattern/action rules. These phases run sequentially and make up a cascade of finite state transducers over annotations. Each pattern/action rule consists of a left-hand-side (LHS) and right-hand-side (RHS). The LHS consist of an annotation pattern description, whereas the RHS consists of action statements. These actions can consist of labels that are attached to pattern elements or regular Java code. An example is given in Figure 2.

```
Phase: tokens
Input: Token
Options: control = appelt

Rule: colon
{Token.kind != "number"}
(   {Token.string == ":",
       Token.sentencePosition != "end"}
): candidateToken
{Token.kind != "number"}
--->
:candidateToken.candidateToken = {type = "colon"}
```

*Figure 1: JAPE example*
7.1.2 Patterns

Around thirty Dutch laws were examined to construct definition patterns for. None of them were present in the dataset of Table 1. Candidate extraction patterns are selected by using distinct words contained in the surface forms of §5.1. This consists of two phases in which positive examples are extracted and some negative afterwards eliminated. These extraction patterns are the following:

1. **To-be definitions:** every sentence containing the root of the word TO BE and the Alpino lexical category ‘smain’. The preceding word should not be punctuation or the article THE.\(^{21}\) The following word should also not contain punctuation.

2. **Verb definitions:** every sentence that contains the words CONSIDERED, REGARDED, COMPREHENDED and UNDERSTOOD not followed by the word UNDER.

3. **List definitions:** every sentence containing a colon (:) which is at the end of the sentence. The surrounding words should not be numbers.

4. **Name definitions:** every sentence containing the sequence of words ‘UNDER THE NAME’.\(^{22}\)

7.1.3 Results

A total of 374 candidate extraction patterns (sentences) have been identified using the described setup. Of the 145 sentences containing definitions, 9 were not identified as such. This gives us an initial recall score of 0.94. The precision score is 0.28 using this setup. Some definitions did not precisely follow the extraction patterns which explains why the recall score is not higher. However, the recall is satisfying enough for the subsequent machine-learning phase.\(^{23}\)

7.2 Machine-learning

Machine-learning techniques were used to filter out false positive results. The engine and different setups are firstly discussed. After that, the results are shown.

---

\(^{21}\) In Dutch: ‘het’.

\(^{22}\) In Dutch: ‘onder de naam’.

\(^{23}\) Modifying the extraction patterns on the left-out set of documents showed no significant improvement in recall, while lowering the precision significantly.
7.2.1 Engine

A language model is trained using the seventeen laws from the annotated gold-standard corpus (Table 1). This is implemented using a machine learning algorithm based on support vector machines (SVM). This SVM engine is a component inside the GATE-framework. All experiments are performed using a binary SVM with a linear kernel.

The model is validated using k-fold cross-validation, where k is 5. This means that subsequently five times 80% of the data is randomly selected to train on. Testing is done on the remaining 20% of the data in the dataset. Results are reported as the average of these five runs.

7.2.2 Setup

Machine-learning features (labels) are selected for each sentence in the corpus. Labels are created per instance and can be unigrams, bigrams and lexical categories (POS). Unigrams and bigrams are constructed from the root form of tokens as outputted by the Alpino parser. Lexical categories of words corresponding to the pos attribute of Alpino. An additional label denotes the class to be learned (definition or non-definition).

Four setups are examined to verify the contribution of different features for machine learning (Table 2). They are divided between unigrams and bigrams, both with and without the use of POS-data.

<table>
<thead>
<tr>
<th>Setup</th>
<th>Unigrams</th>
<th>Bigrams</th>
<th>POS</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2: Machine-learning setups

7.2.3 Error analysis

To get a better understanding of the results, an error analysis is performed by investigating the most significant machine-learning features. Here for, the setup which uses most features is examined, containing unigrams, bigrams and POS.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><em>unigram</em>:'_punct&lt;&gt;</td>
</tr>
<tr>
<td>2</td>
<td><em>unigram</em>:'_punct&lt;&gt;</td>
</tr>
<tr>
<td>3</td>
<td><em>bigram</em>:'_punct&lt;&gt;het_det&lt;&gt;</td>
</tr>
</tbody>
</table>

Table 2: Machine-learning setups
When looking at the four most significant positive features (Table 2), it can be noticed that colons (:) and semicolons (;) are among the most informative ones. This corresponds to list definitions. The presence of the pronoun \textsc{who} (\textit{die}) and the comparative \textsc{as} (\textit{als}) is encountered in definitions such as Example (11). However, the presence of features containing the nouns \textsc{minister} and \textsc{client} (\textit{afnemer}) is specific to the dataset. This makes them probably not generalizable to other texts.

\begin{table}[h]
\centering
\begin{tabular}{ll}
\hline
\textbf{Rank} & \textbf{Features} \\
\hline
1 & \texttt{unigram\_ben\_verb}\> \\
2 & \texttt{unigram\_\_punct}\> \\
3 & \texttt{unigram\_te\_comp}\> \\
4 & \texttt{unigram\_het\_det}\> \\
5 & \texttt{bigram\_in\_prep}\>\texttt{het\_det}\> \\
6 & \texttt{unigram\_de\_det}\> \\
7 & \texttt{unigram\_toepassing\_adj}\> \\
8 & \texttt{bigram\_van\_adj}\>\texttt{toepassing\_adj}\> \\
9 & \texttt{unigram\_jaar\_noun}\> \\
10 & \texttt{unigram\_dat\_comp}\> \\
\hline
\end{tabular}
\caption{Most significant positive features}
\end{table}

Amongst features with negative weights are unfortunately the root form of the verb \textsc{to be} (\textit{ben}) which corresponds with to-be definitions (Table 3). They are probably wrongly classified due to there being very few instances of them. The presence of a full stop (.) is explained by the absence of it in list definitions. Determiners (\textit{de, het}) are more present in non-definitions.

\begin{table}[h]
\centering
\begin{tabular}{ll}
\hline
\textbf{Rank} & \textbf{Features} \\
\hline
1 & \texttt{bigram\_\_punct}\>\texttt{de\_det}\> \\
5 & \texttt{bigram\_minister\_noun}\>\texttt{\_punct}\> \\
6 & \texttt{unigram\_die\_pron}\> \\
7 & \texttt{unigram\_afnemer\_noun}\> \\
8 & \texttt{unigram\_als\_comp}\> \\
9 & \texttt{bigram\_\_punct}\>\texttt{die\_pron}\> \\
10 & \texttt{unigram\_minister\_noun}\> \\
\hline
\end{tabular}
\caption{Most significant negative features}
\end{table}
than sentences containing definitions. The adjective APPLICABLE (toepassing and van toepassing) is a typical constituent of legal texts and often used to link legislation. The noun YEAR (jaar) is probably specific for this dataset and most likely not generalizable.

7.2.4 Results

Results of the machine learning experiment are summarised in Table 4. The best score is achieved using unigrams in combinations with POS features. This score has a precision of 0.79 and recall of 0.68. The harmonic mean of these measurements is the F-score of 0.72.

<table>
<thead>
<tr>
<th>Features</th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>unigram</td>
<td>0.78</td>
<td>0.62</td>
<td>0.68</td>
</tr>
<tr>
<td>unigram + POS</td>
<td>0.80</td>
<td>0.69</td>
<td>0.72</td>
</tr>
<tr>
<td>unigram + bigram</td>
<td>0.79</td>
<td>0.63</td>
<td>0.69</td>
</tr>
<tr>
<td>unigram + bigram + POS</td>
<td>0.77</td>
<td>0.65</td>
<td>0.69</td>
</tr>
</tbody>
</table>

Table 5: Machine-learning results

These overall scores are not very satisfactory and possibly occur due to the high variability of definition types. Other authors have only included definitions with a complete coverage. The definitions in this dataset also contain broadening, narrowing and exhaustive types (see §5.2).

7.3 Varying dataset size

Due to the unsatisfactory results of the machine-learning experiments, the effect of the number of documents in the dataset was investigated.

7.3.1 Setup

A new setup was devised whereby the machine-learning experiments of §7.2 were repeated with varying datasets. Each dataset was a randomly selected subset of the original dataset of 17 laws. The smallest dataset consisted of four laws, whereas the largest contained sixteen laws. Size increment was done by two, making a total of seven datasets (containing 4, 6, 8, 10, 12, 14 and 16 documents respectively). Cross-validation was still 5-fold, which means that each dataset

---

24 This was also reported by Fahmi and Bouma (2006).

25 Note that the actual scores are lower due to the initial recall of 0.94 found using pattern-based extraction.

26 The results in the table are directly taken from GATE. They seem incorrect if the F-score is calculated using the formula $F = \frac{2 \times (precision \times recall)}{precision + recall}$. These deviations are probably due to rounding errors of floating-point numbers in computers.
was five times randomly divided into an 80% training set and tested on the remaining 20% of the documents.

7.3.2 Results

The results of varying datasets cannot be unambiguously interpreted (Table 6; Figure 2). The smallest dataset containing four documents clearly performs the worst. However, the second smallest dataset containing six documents has the best overall recall score of 0.67. The best overall precision score of 0.87 is achieved with a dataset containing ten documents. The F-score does not fluctuate a lot between datasets with ten and sixteen documents.

<table>
<thead>
<tr>
<th>Dataset size</th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.60</td>
<td>0.17</td>
<td>0.27</td>
</tr>
<tr>
<td>6</td>
<td>0.82</td>
<td>0.67</td>
<td>0.70</td>
</tr>
<tr>
<td>8</td>
<td>0.73</td>
<td>0.43</td>
<td>0.48</td>
</tr>
<tr>
<td>10</td>
<td>0.87</td>
<td>0.62</td>
<td>0.70</td>
</tr>
<tr>
<td>12</td>
<td>0.85</td>
<td>0.51</td>
<td>0.60</td>
</tr>
<tr>
<td>14</td>
<td>0.82</td>
<td>0.63</td>
<td>0.69</td>
</tr>
<tr>
<td>16</td>
<td>0.78</td>
<td>0.62</td>
<td>0.68</td>
</tr>
</tbody>
</table>

*Table 6: Results of varying datasets*27

These results do not suggest that adding more documents to the datasets improves the overall result. However, it cannot be definitely concluded that a larger corpus will not perform better. Subsequent experiments with corpora of substantial size will have to show this.

---

27 See footnote 26.
Figure 2: Results of varying datasets
8 Presentation

Presenting the extracted definition to users could be a step towards building an information system for lawyers. This means that every occurrence of defined terms has to be marked as such. Unlike the previous experiments, definitions in this presentation are not automatically annotated. This could however be implemented by using the output of the extraction model in a future system. For now, a prototype of an enhanced markup model is described alongside the demonstration of a possible user interface.

8.1 Enhanced definitions

A markup model is shown that uses identified definitions. Subsequently, a proposal for automatic annotation is made.

8.1.1 Extraction

Enhanced documents should contain, besides the original markup, definition tags. These tags are of the form <Definition>…</Definition> which contain text marked as a definition. Sentences containing definitions should be further split into their constituents: the definiendum, definiens and possibly the connector. The <Definiendum>, <Definiens> and <Connector> tags should be separately annotated inside the <Definition> tag. An example might look like this:

```xml
<Definition>
  <Definiendum> Things </Definiendum>
  <Connector> are </Connector>
  <Definiens> tangible objects that can be controlled by humans </Definiens>
</Definition>
```

This extracted metadata should be stored alongside the original documents or in a database.

8.1.2 Automatic annotation

Each document should be automatically searched and annotated for occurrences of defined terms from the extracted data. The enhanced text should include a reference to each encountered defined term. This terms should be marked with the <Defined> tag that links to the place where the term was defined. In this way, it reassembles a hyperlink similar to cross-references in legislation.
8.2 User interface

The final document with the definitions that have been found should be presented to the user as a webpage in a web browser or another graphical user interface. To demonstrate this approach, some documents of the gold-standard corpus were annotated following the model to display defined terms. The definitions are highlighted throughout the text. When the user hovers over a defined term, a window with the text of the definition is shown. This window displays the location where the definiendum is defined (with a hyperlink) alongside the type and scope of the definition (Figure 2).

![Figure 3: User interface](image)
9 Conclusion

The results of the three research questions are answered and discussed. Some hints for further research are also given.

9.1 Research questions

A conclusion is drawn about the three research question presented in the introduction. These are the questions about the definition model, definition extraction and presentation.

9.1.1 Definition model

The newly developed definition model proved to capture the nature of legal definitions very well. This was especially useful during annotation of the gold-standard corpus. To build the model, literature from different fields has been researched. The drawback of this is the prevalence of theoretical, rather than experimental research in this thesis. A “bycatch” of this is the definition overview in Appendix A.

9.1.2 Definition extraction

The retrieval scores of the definition extraction experiments are too low for a practically useful system. The cause of this is possibly not only due to the absence of a larger dataset. Further research is required to demonstrate the peculiarities of the definition model used in this thesis. The resources to build this were not feasible for a master thesis. Still, they provide a framework to conduct further experiments.

9.1.3 Presentation

A possible enhanced markup has been described to facilitate automatic annotation of defined terms in a legal document. An information system containing a user interface could be constructed around this notion.

9.2 Further research

Further research should concentrate on building a larger dataset to repeat the performed experiments. To do this, a larger gold-standard corpus would have to be annotated. Another aspect is the deconstruction and extraction of the definitions that are found. This would be necessary to build a complete information system to assist lawyers. Afterwards, empirical research involving
human subjects (legal scholars and/or layman) could be carried out to assure that such an information system would have practical value. A possible setup would be to let subjects solve legal problems using the information system and physical books and afterwards compare the results.
10 Bibliography

Burns, C. V. (2007). *Online legal services – a revolution that failed?* Faculty of Law, UNSW.


Appendix A  Types of definitions

A.1 Westerhout (2010)

<table>
<thead>
<tr>
<th>Classification of definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real</td>
</tr>
<tr>
<td>Purpose-based</td>
</tr>
<tr>
<td>Nominal</td>
</tr>
<tr>
<td>Method-based</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

A.2 Loth (1991)

<table>
<thead>
<tr>
<th>Category</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>origin</td>
<td>lexical definition</td>
</tr>
<tr>
<td></td>
<td>stipulative definition</td>
</tr>
<tr>
<td>description of definiens</td>
<td>intensional definition</td>
</tr>
<tr>
<td></td>
<td>extensional (enumerative) definition</td>
</tr>
<tr>
<td>definition per genus proximum et differentiam specificam</td>
<td></td>
</tr>
</tbody>
</table>
### A.3 Parry and Hacker (1991)

<table>
<thead>
<tr>
<th>Type</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>functional type</td>
<td>lexical (reportive) definition</td>
</tr>
<tr>
<td></td>
<td>stipulative definition</td>
</tr>
<tr>
<td></td>
<td>precising definition</td>
</tr>
<tr>
<td>ways of defining</td>
<td>intensional definition</td>
</tr>
<tr>
<td></td>
<td>definition by genus and difference</td>
</tr>
<tr>
<td></td>
<td>definition by species</td>
</tr>
<tr>
<td></td>
<td>operational definition</td>
</tr>
<tr>
<td></td>
<td>synonymous definition</td>
</tr>
<tr>
<td></td>
<td>extensional definition (definition by example)</td>
</tr>
</tbody>
</table>

### A.4 Hurley (2002)

<table>
<thead>
<tr>
<th>Classification of definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>purposes</td>
</tr>
<tr>
<td>stipulative definition</td>
</tr>
<tr>
<td>lexical definition</td>
</tr>
<tr>
<td>precising definition</td>
</tr>
<tr>
<td>defintional techniques</td>
</tr>
<tr>
<td>extensional (denotative) definition</td>
</tr>
<tr>
<td>intensional (connotative definition)</td>
</tr>
<tr>
<td>synonymous definition</td>
</tr>
<tr>
<td>operational definition</td>
</tr>
<tr>
<td>definition by genus and difference</td>
</tr>
</tbody>
</table>

### A.5 Duk (1999)

<table>
<thead>
<tr>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>generalising definition</td>
</tr>
<tr>
<td>specifying definition</td>
</tr>
<tr>
<td>abbreviating definition</td>
</tr>
<tr>
<td>recursive definition</td>
</tr>
<tr>
<td>syntactical definition</td>
</tr>
<tr>
<td>restrictive definition</td>
</tr>
<tr>
<td>extending definition</td>
</tr>
</tbody>
</table>
### A.6 Eijlander and Voermans (1999)

<table>
<thead>
<tr>
<th>Type</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>general definitions (with subtypes from Duk (1999))</td>
<td></td>
</tr>
<tr>
<td>partial definitions</td>
<td></td>
</tr>
<tr>
<td>definitions inside definitory provisions</td>
<td></td>
</tr>
<tr>
<td>definitions outside definitory provisions (incidental definitions)</td>
<td></td>
</tr>
</tbody>
</table>

### A.7 Definitional types not investigated

<table>
<thead>
<tr>
<th>Type</th>
<th>Author(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>conventional definition</td>
<td>Parry and Hacker (1991)</td>
</tr>
<tr>
<td>persuasive definition</td>
<td>Hurley (2002); Parry and Hacker (1991)</td>
</tr>
<tr>
<td>facetious (humorous) definition</td>
<td>Parry and Hacker (1991)</td>
</tr>
<tr>
<td>conceptual (explicative) definition</td>
<td>Parry and Hacker (1991)</td>
</tr>
<tr>
<td>definition by essence</td>
<td>Parry and Hacker (1991)</td>
</tr>
<tr>
<td>theoretical (Copi’s) definition</td>
<td>Hurley (2002); Parry and Hacker (1991)</td>
</tr>
<tr>
<td>definition by subclass</td>
<td>Hurley (2002); Parry and Hacker (1991)</td>
</tr>
<tr>
<td>equational (quantitative) definition</td>
<td>Parry and Hacker (1991)</td>
</tr>
<tr>
<td>contextual definition</td>
<td>Parry and Hacker (1991)</td>
</tr>
<tr>
<td>citational definition</td>
<td>Parry and Hacker (1991)</td>
</tr>
<tr>
<td>definition by paradigm example</td>
<td>Parry and Hacker (1991)</td>
</tr>
<tr>
<td>verbal definition</td>
<td>Parry and Hacker (1991)</td>
</tr>
<tr>
<td>etymological definition</td>
<td>Hurley (2002)</td>
</tr>
</tbody>
</table>
Appendix B  Original Dutch Sources

Example (6)  Algemeen erkende feestdagen in de zin van deze wet zijn: de Nieuwjaarsdag, de Christelijke tweede Paas- en Pinksterdag, de beide Kerstdagen, de Hemelvaarts- dag, de dag waarop de verjaardag van de Koning wordt gevierd en de vijfde mei (artikel 3 lid 1 van de Algemene termijnenwet).

Example (7)  Iemands nakomelingen zijn:
   a.  elk van zijn kinderen;
   b.  de nakomelingen van elk van zijn kinderen (Duk, 1999, p. 15).

Example (8)  In deze wet en de daarop berustende bepalingen wordt verstaan onder:
   (…)
   c.  gezin:
      1°.  de gehuwden tezamen;
      2°.  de gehuwden met de tot hun last komende kinderen;
      3°.  de alleenstaande ouder met de tot zijn last komende kinderen (artikel 4 onder c van de Algemene bijstandswet);

Example (9)  Voor de toepassing van wettelijke voorschriften over bezwaar en beroep worden met een besluit gelijkgesteld:
   a.  de schriftelijke weigering een besluit te nemen, en
   b.  het niet tijdig nemen van een besluit (artikel 6:2 van de Algemene wet bestuursrecht).

Example (10)  Zaken zijn de voor menselijke beheersing vatbare stoffelijke objecten (artikel 3:2 van Boek 3 van het Burgerlijk Wetboek).

Example (11)  Als pleegkind wordt beschouwd het kind dat als eigen kind wordt onderhouden en opgevoed (artikel 4 lid 3 van de Algemene Kinderbijslagwet).

Example (12)  Voor de toepassing van het bij of krachtens deze wet bepaalde wordt verstaan onder:
   (…)
   b.  producent van een databank: degene die het risico draagt van de voor de databank te maken investering; (artikel 1 lid 1 aanhef en onder b van de Databankenwet)
**Example (13)** Onder de naam belasting op leidingwater wordt een belasting geheven op leidingwater (artikel 13 van de Wet belastingen op milieugrondslag).

**Example (14)** Onder handelsnaam verstaat deze wet de naam waaronder een onderneming wordt gedreven (artikel 1 van de Handelsnaamwet).

**Example (15)** Onder vroegere veroordeling wordt mede verstaan een vroegere veroordeling door een strafrechter in een andere lidstaat van de Europese Unie wegens soortgelijke feiten (artikel 7 lid 3 van de Handelsnaamwet).

**Example (16)** Als afzonderlijke openbaarmaking wordt niet beschouwd de heruitzending van een programma door hetzelfde organisme dat dat programma oorspronkelijk uitzendt (artikel 2 lid 9 van de Wet op de naburige rechten).

**Example (17)** Onder dagelijks bestuur wordt in ieder geval verstaan de directe leiding met betrekking tot de medische, de verpleegkundige en de economische aangelegenheden van de ziekenhuisvoorziening (artikel 15 lid 2 van de Wet zorginstellingen BES).

**Example (18)** Voor de toepassing van dit artikel worden onder octrooien mede begrepen kwekersrechten (artikel 12b lid 3 van de Wet op de vennootschapsbelasting 1969).