Structural Analysis of Contract Renewals

Frieda Josi
frieda.josi@hs-hannover.de

Christian Wartena
christian.wartena@hs-hannover.de

University of Applied Sciences and Arts Hanover
Expo Plaza 12, 30539 Hanover, Germany

Abstract

In the present paper we sketch an automated procedure to compare different versions of a contract. The contract texts used for this purpose are structurally differently composed PDF files that are converted into structured XML files by identifying and classifying text boxes. A classifier trained on manually annotated contracts achieves an accuracy of 87% on this task. We align contract versions and classify aligned text fragments into different similarity classes that enhance the manual comparison of changes in document versions. The main challenges are to deal with OCR errors and different layout of identical or similar texts.

We demonstrate the procedure using some freely available contracts from the City of Hamburg written in German. The methods, however, are language agnostic and can be applied to other contracts as well.

1 Introduction

Most contracts between insurance and reinsurance companies are updated annually. This results in many versions of a contract which are structurally and contentwise similar, but which must be completely checked again for a new contract approval. A main obstacle for efficient comparison of old and new versions of the contracts is the fact that the entire approval process is paper based. Insurance companies might send paper versions of the contracts to several reinsurance companies, each of which put stamps and signs on the contract.

Of course all contracts are scanned and stored electronically, but the paper version is in the lead. As intelligent support for the legal domain, we present an approach in which we convert contracts, based on PDF documents, into a structured XML format in order to efficiently find the changed, added or deleted clauses in the new contract version.

For all changed clauses we will predict the impact of the change, or at least determine whether the change is only a stylistic or linguistic improvement or correction or whether the interpretation of the clause is touched. Furthermore, for all changed and new clauses we will check whether the clause is part of a collection of standard clauses or was used in another contract before. In the present paper, we demonstrate a first version of the detection of changes in the contracts. Our procedure was developed and evaluated with German contract texts, but the method is language agnostic and can be applied to contracts in other languages as well.

For the development of the methods we got access to a collection of 100,000 contracts of an insurance company. Since the contracts cannot be made available publicly, we used a small set of freely available contracts for the present study.

Our approach basically consists of four steps: first we extract rectangular text areas from the PDF document. In the second step we classify all text areas into structural classes like header, footer, heading, etc. and merge some adjacent areas of the same type. On the base of this structure two documents are aligned. Finally, the aligned text areas are compared in more detail. An overview of the process flow of our structure analysis of versions of legal texts is shown in the Figure 1.

In the following we describe related work, a detailed description of the approach and an evaluation of the classifier trained for classification of the text areas.

2 Related Work

Gao et al. (Gao et al., 2011) use a method similar to ours for PDF files to analyze the structure of books. After converting the PDF, the content is extracted into a physical and logical structure, the text modules are

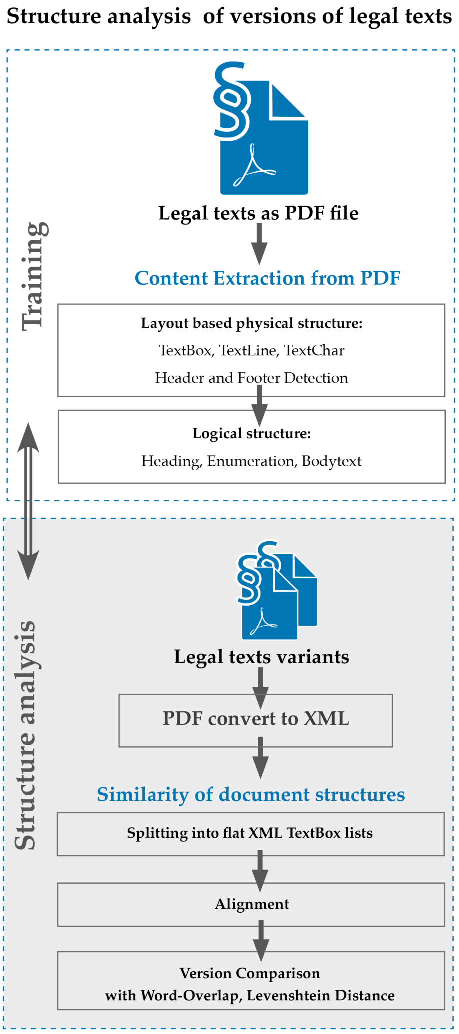

Structure analysis of versions of legal texts
Training
Legal texts as PDF file
Content Extraction from PDF
Layout based physical structure:
TextBox, TextLine, TextChar
Header and Footer Detection
Logical structure:
Heading, Enumeration, Bodytext
Structure analysis
Legal texts variants
PDF convert to XML
Similarity of document structures
Splitting into flat XML TextBox lists
Alignment
Version Comparison
with Word-Overlap, Levenshtein Distance

Figure 1: Procedure of structure analysis of versions of legal texts.

parsed and displayed. However, since these are books, the Gao et al. could assume that all pages have the same layout. This enabled the definition of global typographies. The authors divided the logical structure into a page level and a document level. The page level contains the hierarchical order of the text elements, the header, figures, tables and footnotes. The document level included the writers' chapter structure and metadata. For the extraction of the logical structure at page level, the texts and individual letters were extracted from these text blocks to obtain additional characteristics such as boldface for a heading. For example, the extraction of the logical structure at document level contained the title of the book. For header and footer recognition we use a layout-based approach similar to that of Dèjean and Meunier (2006).

This approach is based on the use of geometric coordinates. In addition, they use the occurrence of digits as an indicator for a text element in the header or footer and the length of the text. With the coordinates of the text blocks in the PDF files a structural sorting per page is possible. The recognition and merging of contiguous text blocks from extracted PDF files is e.g. used by Ramakrishnan et al. (2012). There is some work dealing with extracting named entities (such as companies, persons, places, etc.) from legal texts or finding references to laws (Dozier et al., 2010; Schweighofer, 2010; Nanda et al., 2017). In (Nanda et al., 2017) the vocabulary IATE (Inter-Active Terminology for Europe) is used to create an annotated corpus of named entities and to use it for the NER for European and British legal documents. Chalkidis et al. (2017) use a combination of state-of-the-art methods (such as word embeddings, and part-of-speech tag embeddings) to extract typical contract elements from contract texts. The conversion of content from the layout format of a PDF file to the structured format of an XML file with a small amount of human interaction is done as described by Paick and Zhang (2004). The similarity of the contract versions is compared with the text blocks of the XML output. The word overlap is used as a measure for the agreement between two text blocks of the contract changes. This approach is described by Klampfl et al. (2014).

3 Legal text structure analysis

This section describes our approach to analyze the PDF structure and finding the differences between contract versions.

A simple line by line comparison of documents makes no sense, since the addition of a single word already can change the position of line or page breaks. Furthermore, contracts are usually highly structured texts with lists of definitions, figures, headers and footers on each page. Figure 2 gives an example page of one of the contracts we used. A simple extraction of all text will disturb the natural text flow and insert header and footer text at arbitrary points in the contract text. Thus, we prefer to extract blocks of texts, align the blocks of two documents and compare the document block by block.

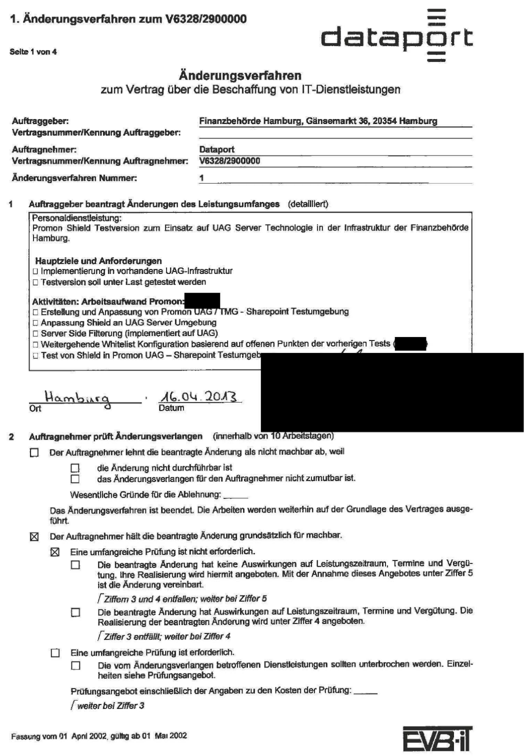
3.1 Document collection

For training a classifier we use 4 non-public insurance documents and 3 publicly available contracts. These contracts are part of the open data strategy of the City Administration Hamburg[1]. These 7 PDF documents consist in total of 198 pages.

From these pages we extracted 4046 text boxes using PDFMiner[2] and classified them by hand. Figure 3

[1] Transparenzportal Hamburg: http://transparenz.hamburg.de/

[2] PDFMiner: https://pypi.org/project/pdfminer/



1. Änderungsverfahren zum V6328/2900000

dataport

Seite 1 von 4

Änderungsverfahren
zum Vertrag über die Beschaffung von IT-Dienstleistungen

Auftraggeber: Finanzbehörde Hamburg, Gänsemarkt 36, 20354 Hamburg
Vertragsnummer/Kennung Auftraggeber:
Auftragnehmer: Dataport
Vertragsnummer/Kennung Auftragnehmer: V6328/2900000
Änderungsverfahren Nummer: 1

1 **Auftraggeber beantragt Änderungen des Leistungsumfanges** (detailliert)

Personaldienstleistung:
Promon Shield Testversion zum Einsatz auf UAG Server Technologie in der Infrastruktur der Finanzbehörde Hamburg.

Hauptziele und Anforderungen
□ Implementierung in vorhandene UAG-Infrastruktur
□ Testversion soll unter Last getestet werden

Aktivitäten: Arbeitsaufwand Promon: [redacted]
□ Erstellung und Anpassung von Promon UAG / TMG - Sharepoint Testumgebung
□ Anpassung Shield an UAG Server Umgebung
□ Server Side Filterung (implementiert auf UAG)
□ Weitergehende Whitelist Konfiguration basierend auf offenen Punkten der vorherigen Tests ([redacted])
□ Test von Shield in Promon UAG – Sharepoint Testumgeb [redacted]

Hamburg , 16.04.2013
Ort Datum

2 **Auftragnehmer prüft Änderungsverlangen** (innerhalb von 10 Arbeitstagen)

□ Der Auftragnehmer lehnt die beantragte Änderung als nicht machbar ab, weil

□ die Änderung nicht durchführbar ist
□ das Änderungsverlangen für den Auftragnehmer nicht zumutbar ist.

Wesentliche Gründe für die Ablehnung: ______

Das Änderungsverfahren ist beendet. Die Arbeiten werden weiterhin auf der Grundlage des Vertrages ausgeführt.

☒ Der Auftragnehmer hält die beantragte Änderung grundsätzlich für machbar.

☒ Eine umfangreiche Prüfung ist nicht erforderlich.

□ Die beantragte Änderung hat keine Auswirkungen auf Leistungszeitraum, Termine und Vergütung. Ihre Realisierung wird hiermit angeboten. Mit der Annahme dieses Angebotes unter Ziffer 5 ist die Änderung vereinbart.

⎡Ziffern 3 und 4 entfallen; weiter bei Ziffer 5

□ Die beantragte Änderung hat Auswirkungen auf Leistungszeitraum, Termine und Vergütung. Die Realisierung der beantragten Änderung wird unter Ziffer 4 angeboten.

⎡Ziffer 3 entfällt; weiter bei Ziffer 4

□ Eine umfangreiche Prüfung ist erforderlich.

□ Die vom Änderungsverlangen betroffenen Dienstleistungen sollten unterbrochen werden. Einzelheiten siehe Prüfungsangebot.

Prüfungsangebot einschließlich der Angaben zu den Kosten der Prüfung: _____

⎡weiter bei Ziffer 3

Fassung vom 01 April 2002, gültig ab 01 Mai 2002

EVB-IT

Figure 2: Example page from the contract texts for the prediction model.

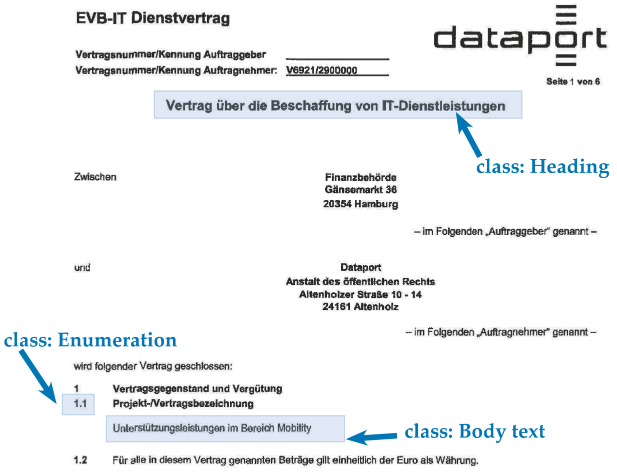
shows an example.

The insurance contracts are written in English, the contracts from Hamburg in German. Since our approach is completely language agnostic, the documents can be mixed for training without any problem.

For the evaluation of the alignment and comparison of contract versions we used 5 documents from the City Administration Hamburg for which at least two versions are available. In the process, care was taken to ensure that there were different degrees of change. The selected contract versions were:

- **HH1a/HH1b:** version with additions
- **HH2a/HH2b:** very different (by many handwritten notes)
- **HH3a/HH3b:** very similar contracts with different contractual partners
- **HH4a/HH4b:** same, but different scanned at an angle
- **HH5a/HH5b:** year variants

The exact names and URLs of all test documents used are given in the Appendix.



EVB-IT Dienstvertrag

dataport

Vertragsnummer/Kennung Auftraggeber ____________
Vertragsnummer/Kennung Auftragnehmer: V6921/2900000

Seite 1 von 6

Vertrag über die Beschaffung von IT-Dienstleistungen

class: Heading

Zwischen

Finanzbehörde
Gänsemarkt 36
20354 Hamburg

– im Folgenden „Auftraggeber" genannt –

und

Dataport
Anstalt des öffentlichen Rechts
Altenholzer Straße 10 - 14
24161 Altenholz

– im Folgenden „Auftragnehmer" genannt –

class: Enumeration

wird folgender Vertrag geschlossen:

1 **Vertragsgegenstand und Vergütung**
1.1 **Projekt-/Vertragsbezeichnung**

Unterstützungsleistungen im Bereich Mobility

class: Body text

1.2 Für alle in diesem Vertrag genannten Beträge gilt einheitlich der Euro als Währung.

Figure 3: Example of Manual Classification

3.2 Detection of structural elements

The 4046 text boxes from the contract texts were classified with the following classes: header, heading, enumeration, body text and footer. The twenty features extracted or calculated for each text box are:

- the coordinates of the lower left and upper right corners of the text box (x1, y1, x2, y2)
- the free margin on each side (m1, m2, m3, m4)
- the fact whether there is a neighboring text box on each side (nb1, nb2, nb3, nb4)
- the font styles bold and upper (bold, upper)
- enumeration elements in text box (enum)
- the size of the text box (area)
- height and width of the text box (height, width)
- number of letters in text box (length)
- fraction of special characters in text box (spec)

We obtain the coordinates of the text boxes, the font (bold and upper) and the text of each box from the parse of PDFMiner. The other features were calculated based on this information. The feature "enumeration" indicates whether the text of the box matches the following regular expression (in Perl Syntax):

$$"\backslash(?([0-9]+|[A-Za-z])(\backslash.([0-9]+|[A-Za-z]))*\backslash)?\$"$$

The distances to the adjacent elements were calculated from their horizontal and vertical overlapping of the coordinates and their distances to the right and left element. The distance to the margins and the size of the text field were also computed. The features for bold and for uppercase indicate whether all characters in a text box are typeset in the respective way. Since headings are often written in this way, we expect this to be a useful feature. From the text we calculate also the fraction of special (non alpha-numeric) characters. Finally, we have calculated the size, width and height of the individual text fields.

A SVM (Support Vector Machine) classifier with RBF Kernel was trained with this data set. The parameters used for this are $\gamma = 0.1 \cdot 10^{-5}$ and penalty parameter $C = 10$. In addition we have calculated a logistic regression model. The performance results of SVM and logistic regression were almost identical. The forecast values of the logistic regression are shown in the "Evaluation and Results" section.

3.3 Alignment

For layout-based structure analysis, we have sorted the text elements on each page from top to bottom and from left to right if they elements are placed next to each other. Adjacent elements that have the same class and have a margin between the areas that is smaller than the height of a text line are merged. Thus, we correct a number of anomalies introduced by the detection of text areas. E.g., in many cases the last line of a paragraph is detected as a separate area, if it has only one or two words.

For the alignment of the text boxes we consider insertions, deletions and substitutions. For insertions and deletions we assign a penalty of 1. The penalty for a substitutions of text t_1 with t_2 is defined as

$$D(t_1, t_2) = 1 - \frac{v(t_1) \cap v(t_2)}{v(t_1) \cup v(t_2)}$$

where $v(t)$ denotes the set of words, excluding stop words, of t. Using dynamic programming we find the alignment with the minimum sum of penalties.

For the 10 test documents we find on average 24 text blocks per page after merging adjacent blocks.

3.4 Version Comparison

Once two texts are aligned, we can start comparing the documents. At the moment we do not analyze insertions and deletions. With a simple heuristic we try to classify pairs of aligned text fragments. We distinguish between:

- **Identical:** Texts are identical up to white spaces
- **OCR Errors:** Texts are identical, but there are differences due to OCR errors
- **Small Differences:** At most 5 words inserted, deleted or substituted
- **Different:** More than 5 words are changed

To decide whether there are real differences or OCR differences we align the texts two times. First we tokenize the text and compute the edit distance based on words (i.e. the minimum number of words that have to be inserted, deleted or changed to obtain the new version from the old one). Then we compute the character based edit distance. If the character based edit distance is at most 2.5 times larger than the word based edit distance, all changes in the words are just small changes, replacing 2 or 3 characters. In this case we assume that all changes are due to OCR errors. However, we did not (yet) determine an optimal value for this threshold.

Table 1: Confusion Matrix from logistic regression

Pred. Real	Header	Heading	Enum.	Text	Footer
Header	701	4	3	8	0
Heading	12	359	8	217	1
Enum.	5	13	429	39	0
Text	7	161	40	1891	5
Footer	0	0	3	7	133

Table 2: Per class results from logistic regression

Class	Precision	Recall	f1-score
Header	0.97	0.98	0.97
Heading	0.67	0.60	0.63
Enum	0.89	0.88	0.89
Text	0.87	0.90	0.89
Footer	0.96	0.93	0.94
Overall	0.87	0.87	0.87

4 Evaluation and Results

4.1 Classifier

Using 10-fold cross validation the accuracy of the classifier (logistic regression) is 87%. The accuracy of the majority classifier, that assigns each element to the class body text, is 52%.

As we can see from the confusion matrix (Table 1) and per class results (Table 2) the best results are achieved for the most important classes: the header and footer. These classes contain text that is not part of the contract text and has to be separated clearly. Most problems arise from confusion between headings and body text.

The contribution of each feature for the logistic regression model is given in Figure 4. The boolean value for an enumeration, the features indicating whether there is a text element above and below (nb1+nb2) and the fraction of special characters in a text element (spec) are used most strongly. Interestingly, the position on the page and the margins around a text box are hardly used.

4.2 Comparison

We use the logical structure of the contracts (heading, enumeration, body text) converted into an XML format for the comparison of contract renewals. The results of the comparison for the test data can be seen in Table

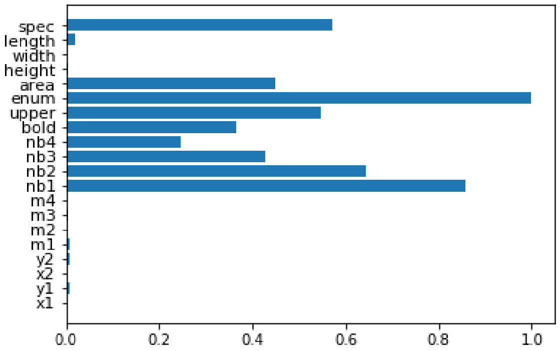


Figure 4: Relative Feature Importance

Table 3: Evaluation version comparison

	HH1a/ HH1b	**HH2a/ HH2b**	**HH3a/ HH3b**	**HH4a/ HH4b**	**HH5a/ HH5b**
Inserted	5	50	79	8	20
Different	4	65	228	43	104
Identical	15	75	186	58	24
OCR Diff.	3	39	43	49	8
Deleted	1	16	25	14	41
Total text boxes	28	245	561	172	197
Fraction identical	0.26	0.16	0.31	0.17	0.012

3. The extracted text boxes are compared as described in section 3.4. As we can see here, for the text pair HH3a/HH3b, e.g., our method found 186 identical text boxes with a text length (measured in characters) of 30% of the contract. These two contracts consist of a very similar structure but with different contractual partners. This means that the underwriters no longer have to check these passages in the text of the contract for consistency, thus making their work more efficient.

As we can see in the Table 3 there are many text boxes that have received the comparison degree "Different". Again, these are often OCR errors, but they are too numerous to be classified as "OCR errors" (see the first example in Table 4 class "Different"). The second example in the class "Different" shows that errors in segmentation and hierarchical sorting also lead to the classification "Different". Another problem is that the text boxes recognized by PDFMiner are not always the same in the two versions and merging does not entirely compensate for this, e.g. because one of the elements was classified incorrectly.

5 Discussion and Future Work

In this paper we have shown that modifications in contract renewals can be identified and analyzed using supervised learning and text alignment.

We want to continue this approach in further work and improve the classification of the classes heading, body text and enumeration. In addition, we want to implement the recognition of named entities, as described e.g. in (Nanda et al., 2017). Furthermore, the text structure can be subdivided in more detail and further structural elements such as text boxes containing handwritten notes can be included. We will improve our approach by carrying out further tests with a larger training corpus, making further parameter settings and adding additional features such as font size. During the course of the project, the existing XML structure also will be transformed into a standardized legal XML structure, as proposed by "OASIS LegalXML Electronic Court Filing TC".[3] On this basis we plan the clause analysis in the contract texts. The recognized clauses will be checked against a collection of model clauses and the occurrence of the same or almost same clause in other contract will be checked. We plan to visualize the status of each clause, like unchanged, found in another contract, etc.

Table 4: Examples version comparison for HH2a vs. HH2b. Differences are marked in the text.

Identical	für die Leistungen nach 3.2 die Kostenschätzung. für die Leistungen nach 3.2 die Kostenschätzung.
OCR Errors	6.1.3 (1) Grundlagenermittlung 6.1._3 (1) GrundlageAermittlung ·
Small Diff.	Fertigstellung der leistungen dieses Vertrages bis Ende Juli 2012 Fertigstellung der Leistungen dieses Vertrages bis Ende Oktober 2013
Different	2.4 ··i Die Baumaßnahme untCFliegt dem ZustimFF1uF1gS'1cffelhreF1 Flach § 84 HBauO. Die für die veranhvortliene Leitung zuständige Person wird der bzw. dem AN sehriftlieh be nannt. 2.4 ¿ Die Baumaßnahme uAterliegt dem Zustimmungsvcrfahren nach § 64 HBauO. Die für eli9e verantwoftliehe Leitung zuständige Person wird der bzw. dem AN schriftlich be ft8flflt: § 8 - Ergänzende Vereinbarungen und anderen fachlich Beteiligten

With the visualization of the changes in the contract renewals, a tool can then be implemented that provides valuable support for underwriters and other legal entities in their daily work and simplifies and improves their daily work in the long term.

Acknowledgements

The authors would like to thank Fabian Schmieder for many helpful discussions and pointing us to the publicly available contracts of the City of Hamburg.

[3] OASIS LegalXML: https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=legalxml-courtfiling

References

Chalkidis, I., I. Androutsopoulos, and A. Michos (2017). Extracting contract elements.

Dèjean, H. and J.-L. Meunier (2006). A system for converting PDF documents into structured XML format. In *Document Analysis Systems VII*, Lecture Notes in Computer Science, pp. 129–140. Springer, Berlin, Heidelberg.

Dozier, C., R. Kondadadi, M. Light, A. Vachher, S. Veeramachaneni, and R. Wudali (2010). Named entity recognition and resolution in legal text. In *Semantic Processing of Legal Texts*, Lecture Notes in Computer Science, pp. 27–43. Springer, Berlin, Heidelberg.

Gao, L., Z. Tang, X. Lin, Y. Liu, R. Qiu, and Y. Wang (2011). Structure extraction from PDF-based book documents. In *Proceedings of the 11th Annual International ACM/IEEE Joint Conference on Digital Libraries*, JCDL '11, pp. 11–20. ACM.

Klampfl, S., M. Granitzer, K. Jack, and R. Kern (2014). Unsupervised document structure analysis of digital scientific articles.

Nanda, R., G. Siragusa, L. Di Caro, M. Theobald, G. Boella, L. Robaldo, and F. Costamagna (2017). Concept recognition in european and national law. In A. Z. Wyner and G. Casini (Eds.), *Legal Knowledge and Information Systems - JURIX 2017: The Thirtieth Annual Conference, Luxembourg, 13-15 December 2017*, Frontiers in Artificial Intelligence and Applications, pp. 193. IOS Press.

Paick, Y. Y. K. and Y. P. Y. Zhang (2004). PDF2xml: Converting PDF to XML.

Ramakrishnan, C., A. Patnia, E. Hovy, and G. A. Burns (2012). Layout-aware text extraction from full-text PDF of scientific articles.

Schweighofer, E. (2010). Semantic indexing of legal documents. In *Semantic Processing of Legal Texts*, Lecture Notes in Computer Science, pp. 157–169. Springer, Berlin, Heidelberg.

Appendix: Used Contracts

Training Documents

Reference	File name	URL
HHTrain1	Akte_611.10-13(1).pdf	http://suche.transparenz.hamburg.de/dataset/oeffentlich-rechtlicher-vertrag-gehrecht-bebauungsplan-harburg-59-theodor-york-strasse?forceWeb=true
HHTrain2	Akte_FB2a.809.13-25_4(1).pdf	http://suche.transparenz.hamburg.de/dataset/aenderungsverfahren-fuer-vertrag-6328-zuvex-weitere-schritte-zur-anbindung-externer-nutzer?forceWeb=true
HHTrain3	Akte_FB2a.800.01-2_3(1).pdf	http://suche.transparenz.hamburg.de/dataset/v6921-unterstuetzungsleistung-mobility-vertrag?forceWeb=true
Train data1-4	*4 non public reinsurance contracts*	

Test Documents

Reference	File name	URL
HH1a	Aenderungsbescheid.pdf	http://suche.transparenz.hamburg.de/dataset/3-planen-zur-temporaeren-anbringung-an-einem-baugeruest-zur-bewerbung-von-mietwohnungen?forceWeb=true
HH1b	Befristete_Genehmigung_nach_HBauO.pdf	http://suche.transparenz.hamburg.de/dataset/3-planen-zur-temporaeren-anbringung-an-einem-baugeruest-zur-bewerbung-von-mietwohnungen1?forceWeb=true
HH2a	Akte_000.00-04.pdf	http://suche.transparenz.hamburg.de/dataset/vertrag-spielplatz-voigtstrasse-ii?forceWeb=true
HH2b	Akte_000.00-04(1).pdf	http://suche.transparenz.hamburg.de/dataset/vertrag-spielplatz-voigtstrasse?forceWeb=true
HH3a	Akte_FB63.51-06(1).pdf	http://suche.transparenz.hamburg.de/dataset/bezirk-eimsbuettel-vereinbarung-ueber-die-erstmalige-endgueltige-herstellung-von-erschl-02-2014?forceWeb=true
HH3b	Akte_FB63.51-06(3).pdf	http://suche.transparenz.hamburg.de/dataset/bezirk-hamburg-nord-vereinbarung-ueber-die-erstmalige-endgueltige-herstellung-von-ersch-02-2014?forceWeb=true
HH4a	Akte_G103-36.01_06-10-.pdf	http://suche.transparenz.hamburg.de/dataset/aenderungsvertrag-zum-vertrag-zwischen-der-freien-und-hansestadt-hamburg-fhh-und-dem-ha-12-20161?forceWeb=true
HH4b	Akte_G103-36.01_06-10-(1).pdf	http://suche.transparenz.hamburg.de/dataset/aenderungsvertrag-zum-vertrag-zwischen-der-freien-und-hansestadt-hamburg-fhh-und-dem-hamburger-?forceWeb=true
HH5a	entwurf-eines-gesetzes-zu-dem-abkommen-zur-dritten-änderung-des-abkommens-über-das-deutsche-institut-für-bautechnik.pdf	http://www.buergerschaft-hh.de/ParlDok/dokument/53849/entwurf-eines-gesetzes-zu-dem-abkommen-zur-dritten-%c3%a4nderung-des-abkommens-%c3%bcber-das-deutsche-institut-f%c3%bcr-bautechnik.pdf
HH5b	entwurf-eines-gesetzes-zu-dem-abkommen-zur-zweiten-änderung-des-abkommens-über-das-deutsche-institut-für-bautechnik-und-zum-erlass-des-bauprodukte-mar.pdf	http://www.buergerschaft-hh.de/ParlDok/dokument/37131/entwurf-eines-gesetzes-zu-dem-abkommen-zur-zweiten-%c3%a4nderung-des-abkommens-%c3%bcber-das-deutsche-institut-f%c3%bcr-bautechnik-und-zum-erlass-des-bauprodukte-mar.pdf