

Segmenting U.S. Court Decisions into Functional and Issue Specific Parts

Jaromir Savelka and Kevin D. Ashley

JURIX 2018

Outline

1. Background and motivation
2. Experiments
 - Data sets
 - Classification pipeline
3. Results
4. Future Work
5. Conclusions

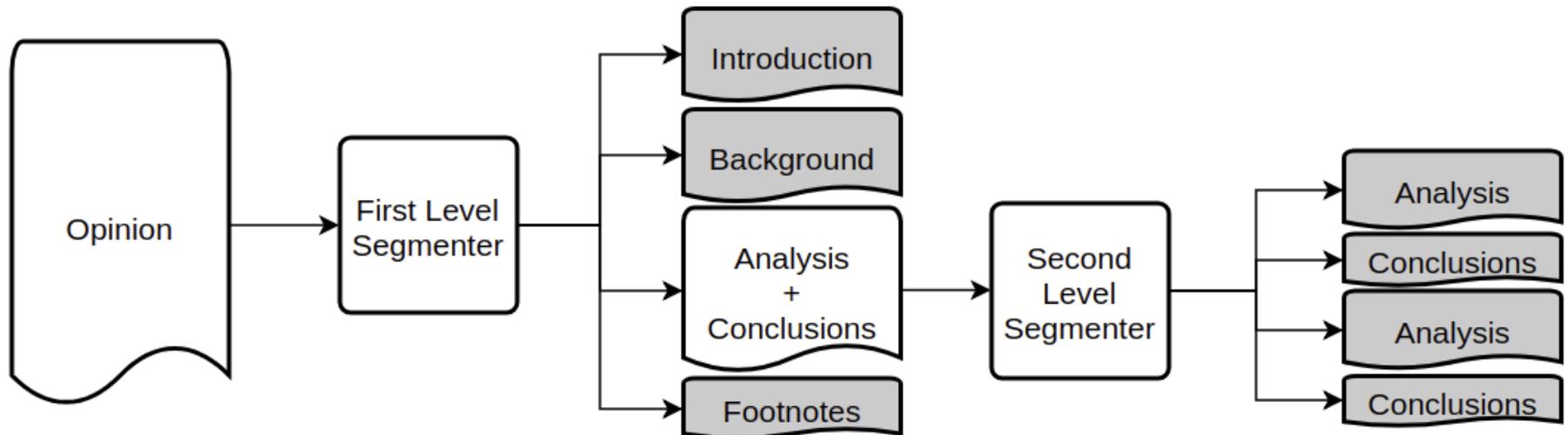
BACKGROUND AND MOTIVATION

Background

- Court opinions consist of several high-level parts each of which has a different function.
- The main **Analysis** part contains sub-parts each of which is dedicated to a different issue.
- Distinguishing the parts is crucial for a lawyer to focus attention where it matters.
- How could NLP and ML techniques be helpful in recognition of the individual parts?

Two-step Segmentation Process

- In the first step an opinion is segmented into a number of **functional** non-overlapping parts.
- In the second step the Analysis part is segmented into **issue-specific** parts.



First Step: Functional Parts

1. Introduction
2. Background
3. Analysis
4. Concurrence or Dissent
5. Footnotes
6. Appendix

Second Step: Issue-specific Parts

- The Analysis part is annotated with the **Conclusions** type.
(sentences where court states decision concerning each issue)
- The annotations are used to segment the Analysis part into specific issues.
- We adopt an assumption that reasoning wrt each issue finishes with Conclusions.
(Obviously, the situation is more complex than that.)



Second Step: Issue-specific Parts II

Examples:

Under these circumstances we cannot say that the trial court's finding that both Mills and Northrop understood the data to be confidential was "clearly erroneous."

On this aspect of the case we decide that, while the record would have supported the trial judge if he had reached a contrary conclusion, it does not show that he abused his equitable discretion in reaching the one which he did.

Having failed to do so, they cannot now maintain that the plaintiff's failure to move for a formal dismissal of that part of its action was so inequitable as to bar the relief for theft of trade secrets to which it is otherwise entitled.

Motivation

- Important for information retrieval, summarization, and text understanding
(Knowing in which high-level part a sentence appears will help to annotate sentences in terms of the roles they play in legal argument.)
- The annotation task may have important pedagogical potential.
(Student annotators could benefit from this kind of practice while providing training instances for machine learning and legal text analytics.)

Related and Prior Work

- **Dividing the segments by topic**

(General: Choi et al. 2001, Hearst 1997, Le 1999, Lu et al. 2011; Domain specific: Conrad&Dabney 2017, Farzindar&Lapalme 2004, Vanderbeck 2011)

- **Rhetorical roles of sentences**

(Farzindar&Lapalme 2007, Hachey&Grover 2006, Moens 2007, Saravan&Ravindran 2010)

- **Identifying characteristic features of sections**

(Barrazega&Faiz 2017, Farzindar&Lapalme 2004, Grover et al. 2003, Bansal et al. 2016, Grabmair et al. 2015, Wyner et al. 2010, Harasta et al. 2017)

- **Conditional Random Fields (the same algorithm we use)**

(Saravanan et al. 2006, Saravanan&Ravindran 2010)

Experiments

DATA SETS

Data Sets

- We downloaded 316 court decisions from Court Listener and Google Scholar.
 - 143 are from the area of cyber crime (cyber bullying, credit card frauds, possession of electronic child pornography),
 - 173 cases involve trade secrets (typically misappropriations)
- We created guidelines for annotation of the decisions with the types introduced earlier.
- Two human annotators (the authors) annotated the decisions using Gloss.
- 50 decisions were annotated by both annotators to measure inter-annotator agreement.

Annotation Example

gloss Annotator jsavelka

Data

Types 8 Annotations 8 Auto 0

431 F.2d 1012 (1970) E. I. duPONT deNEMOURS & COMPANY, Inc., Plaintiff-Appell...

This is a case of industrial espionage in which an airplane is the cloak and a camera t...

This is a case of first impression, for the Texas courts have not faced this precise fact...

We conclude that it is and that the Texas courts would so hold.

The Supreme Court of that state has declared that "the undoubted tendency of the la...

In such a situation DuPont has a valid cause of action to prohibit the Christophers fro...

We note that this view is in perfect accord with the position taken by the authors of th...

Having concluded that aerial photography, from whatever altitude, is an improper met...

Text JURIX2018_trade_secrets E I DuPont DeNemours & Company v. Christopher (None) Done Map

431 F.2d 1012 (1970)

E. I. duPONT deNEMOURS & COMPANY, Inc., Plaintiff-Appellee,
v
Rofe CHRISTOPHER et al., Defendants-Appellants.

No. 28254.

United States Court of Appeals, Fifth Circuit.

July 20, 1970. Rehearing Denied and Rehearing Denied August 25, 1970.

David J. Kreager, John G. Tucker, Orgain, Bell & Tucker, Beaumont, Tex., for defendants-appellants.

Robert Q. Keith, Mehaffy, Weber, Keith & Gonsoulin, Beaumont, Tex., William E. Kirk, Jr., Wilmington, Del., for plaintiff-appellee.

Before WISDOM, GOLDBERG and INGRAHAM, Circuit Judges.

Rehearing Denied and Rehearing En Banc Denied August 25, 1970.

GOLDBERG, Circuit Judge:

This is a case of industrial espionage in which an airplane is the cloak and a camera the dagger. The defendants-appellants, Rolfe and Gary Christopher, are photographers in Beaumont, Texas. The Christophers were hired by an unknown third party to take aerial photographs of new construction at the Beaumont plant of E. I. duPont deNemours & Company, Inc. Sixteen photographs of the DuPont facility were taken from the air on March 19, 1969, and these photographs were later developed and delivered to the third party.

DuPont employees apparently noticed the airplane on March 19 and immediately began an investigation to determine why the craft was circling over the plant. By that afternoon the investigation had disclosed that the craft was involved in a photographic expedition and that the Christophers were the photographers. DuPont contacted the Christophers that same afternoon and asked them to reveal the name of the person or corporation requesting the photographs. The Christophers refused to disclose this information, giving as their reason the client's desire to remain anonymous.

Having reached a dead end in the investigation, DuPont subsequently filed suit against the Christophers, alleging that the Christophers had wrongfully obtained photographs revealing DuPont's trade secrets which they then sold to the undisclosed third party. DuPont contended that it had developed a highly secret but unpatented process for producing methanol, a process which gave DuPont a competitive advantage over other producers. This process, DuPont alleged, was a trade secret developed after much expensive and time-consuming research, and a secret which the company had taken special precautions to safeguard. The area photographed by the Christophers was the plant designed to produce methanol by this secret process, and because the plant was still under construction parts of the process were exposed to view from directly above the construction area. Photographs of that area, DuPont alleged, would enable a skilled person to deduce the secret process for making methanol. DuPont thus contended that the Christophers had wrongfully appropriated DuPont trade secrets by taking the photographs and delivering them to the undisclosed third party. In its suit DuPont asked for damages to cover the loss it had already sustained as a result of the wrongful disclosure of the trade secret and sought temporary and permanent injunctions prohibiting any further circulation of the photographs already taken and prohibiting any additional photographing of the methanol plant.

The Christophers answered with motions to dismiss for lack of jurisdiction and failure to state a claim upon which relief could be granted. Depositions were taken during which the Christophers

Data Sets Statistics

| | Cyber Crime | | | Trade Secrets | | | Total | | |
|--------------|-------------|----------|-------|---------------|----------|-------|----------|----------|-------|
| | Filtered | Original | Agree | Filtered | Original | Agree | Filtered | Original | Agree |
| Documents | 139 | 143 | - | 158 | 173 | - | 297 | 316 | - |
| Introduction | 139 | 143 | .965 | 158 | 173 | .934 | 297 | 316 | .947 |
| Background | 133 | 139 | .758 | 152 | 172 | .787 | 285 | 311 | .774 |
| Analysis | 139 | 143 | .932 | 158 | 173 | .936 | 297 | 316 | .935 |
| Conclusions | 398 | 429 | .689 | 833 | 909 | .763 | 1231 | 1338 | .732 |
| Concurr/Diss | 0 | 18 | 1.0 | 0 | 44 | 1.0 | 0 | 62 | 1.0 |
| Footnotes | 95 | 97 | .983 | 103 | 113 | .982 | 198 | 210 | .982 |
| Appendix | 0 | 7 | - | 0 | 6 | - | 0 | 13 | - |

Experiments

CLASSIFICATION PIPELINE

Segmentation into Functional Parts

1. Each document is split into individual paragraphs and sentences.
2. These are transformed into vectors of paragraph level features.
(e.g., lower-case tokens and POS-tagged lemmas, the position of a paragraph within a document, its length as well as the average length of sentences)
3. The first and the last 5 tokens in a paragraph are described through more detailed features
(e.g., token's signature, length, and type - digit, case, white space)

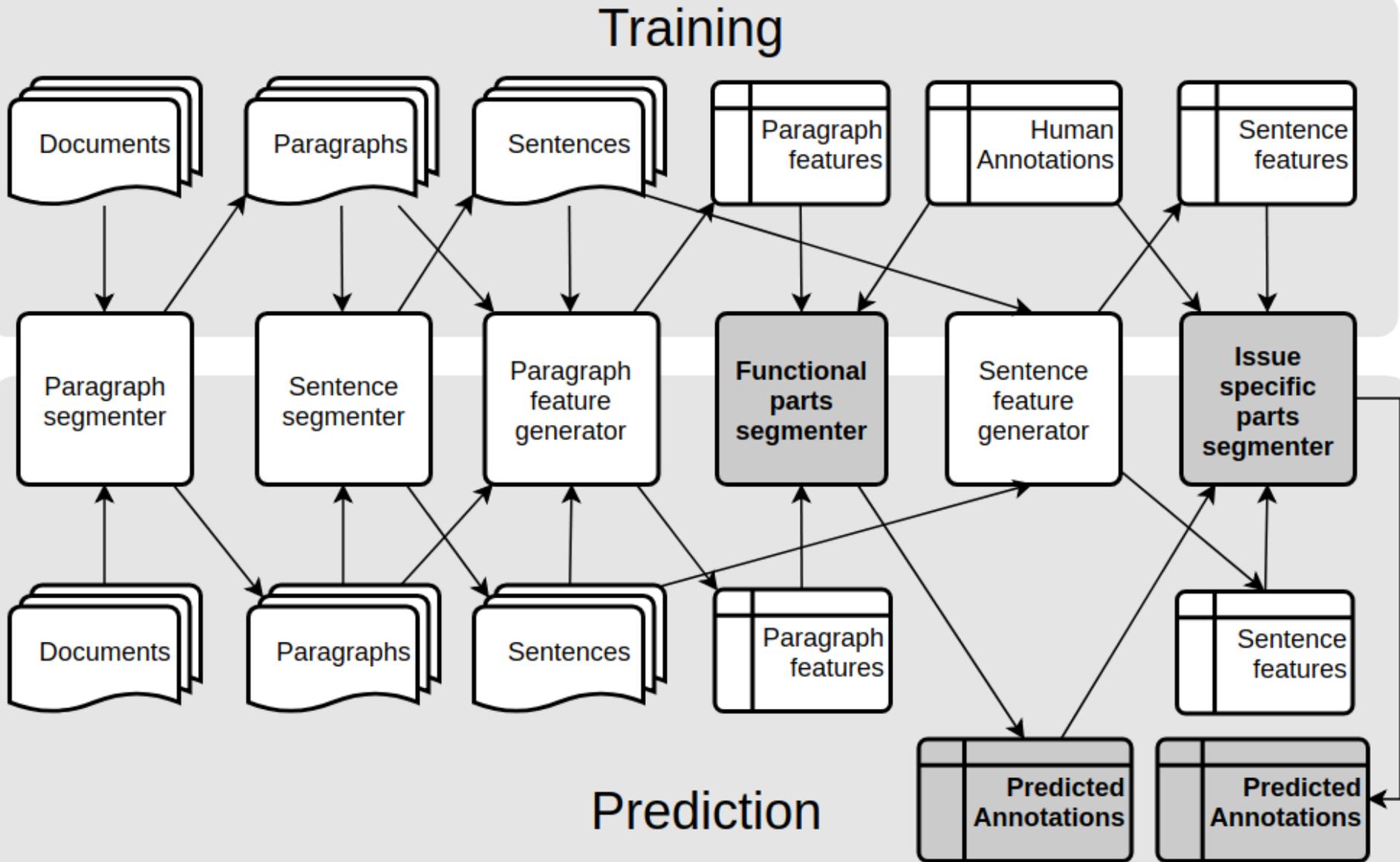
Segmentation into Functional Parts II

4. The feature vectors and human created annotations are then used in the training.
 - i. The model for recognizing the **Introduction** type is trained on full texts.
 - ii. The model separating the **Background** type from the rest.
 - iii. The model for finding the boundary between the **Analysis** and **Footnotes**.

Segmentation into Issues

1. Each document is split into individual paragraphs and sentences.
2. These are transformed into vectors of paragraph level features.
(e.g., lower-case tokens and POS-tagged lemmas, the position of a sentence within a paragraph, its length, token's signature, length, and type - digit, case, white space)
3. The feature vectors and human created annotations are then used in the training.
(The Conclusions recognizer consist of a single CRF model.)

Pipeline Schema



RESULTS

Results: Functional Parts

| | Cyber Crime | | | Trade Secrets | | | Total | | |
|--------------|-------------|------|----------------|---------------|------|----------------|-------|------|----------------|
| | P | R | F ₁ | P | R | F ₁ | P | R | F ₁ |
| Introduction | .926 | .945 | .935 | .907 | .947 | .926 | .914 | .947 | .930 |
| Background | .634 | .752 | .689 | .640 | .775 | .701 | .638 | .767 | .697 |
| Analysis | .922 | .879 | .900 | .948 | .871 | .908 | .939 | .873 | .905 |
| Footnotes | .874 | .963 | .916 | .845 | .981 | .908 | .855 | .973 | .910 |

- The performance of the models differs considerably across the types.
- It correlates across the two different domains (and with inter-annotator agreement).
- Due to data sparsity we did not attempt to predict the Appendix and the Concurrence and Dissent types.

Results: Issue-specific Parts

| | Cyber Crime | | | Trade Secrets | | | Total | | |
|-------------|-------------|------|----------------|---------------|------|----------------|-------|------|----------------|
| | P | R | F ₁ | P | R | F ₁ | P | R | F ₁ |
| Conclusions | .521 | .461 | .489 | .576 | .493 | .532 | .552 | .482 | .515 |

- Although, the performance is promising the experiments confirm that this task is very challenging.
- From the inter-annotator agreement it appears that this task is the most challenging one as well.

Error Analysis: Issue Segmentation

- The sentences that were missed by the system were often quite short
(“There is no error.”)
- Some of the shorter length sentences were recognized correctly
(“The judgment of the district court is affirmed.”)
- For false positives it appears that attribution is a major challenge.
(Some sentences included verbs like “conclude” or “hold” which are likely very suggestive for a sentence to be classified as the Conclusions.)

Error Analysis: Issue Segmentation

- Words such as “therefore” or various forms of “find” often indicate the Conclusion type.
(But not always: “Therefore, we now address appellants' claims,” “hereby finds the following facts and state separately its conclusions of law,” or “because there was no jury finding against her.”)
- The error analysis also confirmed the task is very challenging even for humans.
(For example, when are findings of fact conclusions as to an issue? In addition, a court may break an issue down into multiple sub-issues like whether a rule applies or whether evidence supports a conclusion)

FUTURE WORK

Future Work

- Higher level features
- One of the problems that we detected is data sparsity in case of certain types.
 - We hypothesize that law students can annotate legal texts as a useful pedagogical exercise.
 - 10 students in Ashley's IP course employed Gloss to annotate 4 trade secret law cases
 - we plan students to annotate legal decisions in terms of key aspects of the reasoning in a case beyond high-level parts and conclusions

CONCLUSIONS

Conclusions

- We examined the possibility of automatically segmenting court opinions into
 - high-level functional (step 1) and
 - issue specific (step 2) parts.
- The functional parts could be predicted in a quality not too far from human performance.
- For issue-specific parts there is a gap between a machine and a human annotator.

Thank you!

Questions, comments and suggestions are welcome now or any time at jas438@pitt.edu.

References

A. Bansal, Z. Bu, B. Mishra, S. Wang, K. Ashley, and M. Grabmair. “Document Ranking with Citation Information and Oversampling Sentence Classification in the LUIMA Framework.” *JURIX*, v. 294, p. 33. IOS Press, 2016.

Berrazega, I., and Faiz, R. “Automatic structuring of Arabic normative texts.” *Proceedings of the Computational Methods in Systems and Software* 238-249. Springer, 2017.

Choi, F. Y., Wiemer-Hastings, P., and Moore, J. “Latent semantic analysis for text segmentation.” *Proceedings of the 2001 Conference on Empirical Methods in Natural Language Processing*, 2001.

Conrad, J. and Dabney, D. “A cognitive approach to judicial opinion structure: applying domain expertise to component analysis.” *Proc. 8th Int'l Conf. on Artificial intelligence and Law* 1-11. ACM, 2001.

Falakmasir, M.H. and Ashley, K.D. “Utilizing Vector Space Models for Identifying Legal Factors from Text.” *JURIX* 183-192, 2017.

References II

- A. Farzindar and G. Lapalme. "Letsum, an automatic legal text summarizing system." *Legal knowledge and information systems, JURIX* 11-18, 2004.
- Grabmair, M., K. Ashley, R. Chen, P. Sureshkumar, C. Wang, E. Nyberg, and V. Walker. "Introducing LUIMA: an experiment in legal conceptual retrieval of vaccine injury decisions using a UIMA type system and tools." In *Proc. 15th Int'l Conf. on Artificial Intelligence and Law*, pp. 69-78. ACM, 2015.
- Grover, C., Hachey, B., Hughson, I., and Korycinski, C. "Automatic summarisation of legal documents." *Proc. 9th Int'l Conf. on Artificial intelligence and law* 243-251. ACM. 2003.
- B. Hachey and C. Grover. "Extractive summarisation of legal texts." *Artificial Intelligence and Law* 14,4 305-345, 2006.
- Harasta, J., F. Kasl, J. Misek, and J. Savelka. "Segmentation of Czech Court Decisions into Subtopic Passages." *CEILI Workshop on Legal Data Analysis, JURIX*, 2017.

References III

Hearst, M. "TextTiling: Segmenting text into multi-paragraph subtopic passages." *Computational Linguistics* 23.1, 33-64. 1997.

Lafferty, J., A. McCallum, and F. Pereira. "Conditional random fields: Probabilistic models for segmenting and labeling sequence data." 2001.

Le, T. T. N. "A study on Hierarchical Table of Indexes for Multi-documents." *Ph.D. Diss., Japan Advanced Inst. of Science and Technology* 1-47. 1999.

Lu, Q., Conrad, J., Al-Kofahi, K., and Keenan, W. "Legal document clustering with built-in topic segmentation." *Proc. 20th ACM Int'l Conf. on Information and Knowledge Management* 383-392. ACM. 2011.

M.-F. Moens. "Summarizing court decisions." *Information Processing and Management* 43, 6, 1748-1764. 2007.

References IV

Okazaki, N. “Crfsuite: a fast implementation of conditional random fields.” 2007.

Saravanan, M., and B. Ravindran. “Identification of rhetorical roles for segmentation and summarization of a legal judgment.” *Artificial Intelligence and Law* 18.1 45-76. 2010.

Saravanan, M., B. Ravindran, and S. Raman. “Improving legal document summarization using graphical models.” *JURIX* 2006.

Savelka, J., V. Walker, M. Grabmair, and K. Ashley. “Sentence Boundary Detection in Adjudicatory Decisions in the United States.” *Traitement Automatique des Langues* 58.2 21-45. 2017.

Savelka, J., and K. Ashley. “Detecting Agent Mentions in US Court Decisions.” *JURIX*. 2017.

References V

Savelka, J., and K. Ashley. "Extracting case law sentences for argumentation about the meaning of statutory terms." *Proc. of the Third Workshop on Argument Mining (ArgMining2016)*. 2016.

Shulayeva, O., Siddharthan, A., and Wyner, A. "Recognizing cited facts and principles in legal judgements." *Artificial Intelligence and Law* 25(1),107-126. 2017.

Schweighofer, E. "The Role of AI and Law in Legal Data Science." *JURIX* 191-192. 2018.

Vanderbeck, S., Bockhorst, J., and Oldfather, C. "A Machine Learning Approach to Identifying Sections in Legal Briefs." *MAICS* 16-22. 2011.

Wyner, A., Mochales-Palau, R., Moens, M. F., & Milward, D. Approaches to text mining arguments from legal cases. In *Semantic processing of legal texts* (pp. 60--79). Springer, Berlin, Heidelberg. 2010.