Natural Language Processing for Law and Social Science

13. Legal NLP

Outline

Tools for Legal NLP

Wrapping Up

Legal Texts

- Legislation
  - the statutes enacted by legislators, which are then added to a compiled code.
  - hierarchical structure, extensively cross-referenced.

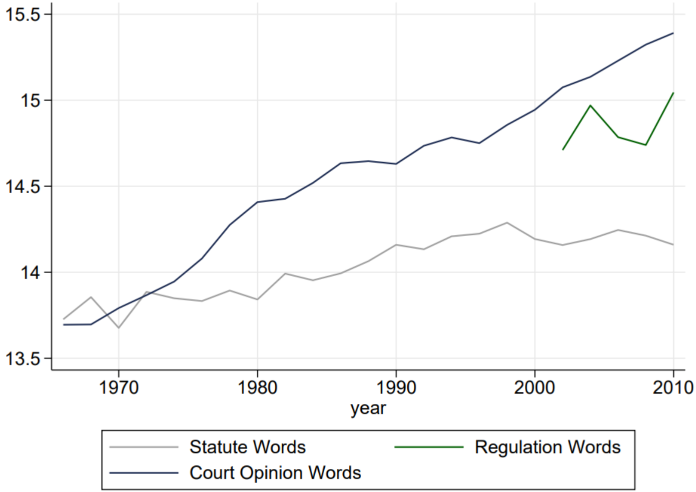
Legal Texts

- Legislation
  - the statutes enacted by legislators, which are then added to a compiled code.
  - hierarchical structure, extensively cross-referenced.
- Regulations
  - the more specific rules to implement legislation, decided by more technocratic agencies.
  - e.g., tax agency should decide whether a gift counts as income.

Legal Texts

- Legislation
  - the statutes enacted by legislators, which are then added to a compiled code.
  - hierarchical structure, extensively cross-referenced.
- Regulations
  - the more specific rules to implement legislation, decided by more technocratic agencies.
  - e.g., tax agency should decide whether a gift counts as income.
- Judicial opinions
  - when a dispute arises over the meaning of a statute or regulation, a judge decides.
  - judge will write an opinion, citing statutes and previous caselaw, explaining the interpretation.

Legal Text Output in U.S. States (Ash, Morelli, and Vannoni 2022)



note log scale – per year we see:

- ~1.3M words in statutes
- ~3.3M words in regulations
- ~4.8M words in state court opinions

Legal language is different from common language

1. legal documents tend to have more structure (e.g. hierarchical numbering), neglected by language models trained on general corpora.
2. legal language is more precise $\rightarrow$ lawyers are rewarded for reducing ambiguity.
    - however:
        - definitions are often specified elsewhere in the document
        - extensive and pivotal citations to other documents
        - when provisions are contested, ambiguity might be used to overcome conflict.

Legal language is different from common language

1. legal documents tend to have more structure (e.g. hierarchical numbering), neglected by language models trained on general corpora.
2. legal language is more precise → lawyers are rewarded for reducing ambiguity.
   - however:
     - definitions are often specified elsewhere in the document
     - extensive and pivotal citations to other documents
     - when provisions are contested, ambiguity might be used to overcome conflict.

Ambiguity in natural language can be helpful → helps explain why efforts to put law on a formal-logic basis, or to say "law is code", have failed.

Legal Ambiguity

- A plan is described in this paragraph if **substantially** all of the contributions required under the plan are made by employers **primarily** engaged in the long and short haul trucking industry.

Section	**Definition**
Title 1, § 8	*every infant member of the species homo sapiens who is born alive at any stage of development*
Title 20, § 7801	*any person within the age limits for which the State provides free public education*
Title 42, § 1901	*a legitimate child, an adopted child, and, if designated as beneficiary by the insured, a stepchild or an illegitimate child*
Title 42, § 1397jj	*an individual under 19 years of age*
Title 42, §1769(d)	*a person under the age of 18*
Title 42, §5119(c)	*a person who is a child for purposes of the criminal child abuse law of a State*

Table 1: Examples of the how the legal definition *child* is defined across the U. S. Code.

Legal Interpretation is different from Natural Language Understanding

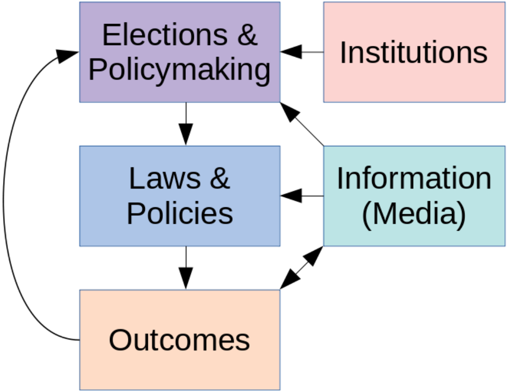
- Example from Solan (2010): whether ***or*** is interpreted as **inclusive** ("one or both") or **exclusive** ("one, not both")
  - *In U.S. v. 171-02 Liberty Ave.* (E.D.N.Y. 1989), government seized Greco's drug den under forfeiture statute for property involved in crime.
  - statute exempted crimes occurring "without knowledge **or** consent of the owner." Greco had knowledge but did not consent.

- interpretation 1:
  - lack of knowledge **or** lack of consent, **by themselves**, are sufficient for exemption $\rightarrow$ Greco wins.
- interpretation 2:
  - **both** lack of knowledge **and** lack of consent are needed for exemption $\rightarrow$ Greco loses.

Natural language understanding does not provide an answer here.

Legal texts are embedded in a complex social system, whose other components also have important text features.

- Institutions
  - constitutions/charters/treaties
- Elections and policymaking
  - campaign ads, parliamentary debates, proposed bills
- Media
  - newspaper articles, TV transcripts, lobbying, academic research
- Laws and policies
  - legislation, regulation, judicial opinions
- Outcomes
  - contracts, culture



Uses of NLP in legal practice

`https://emerj.com/ai-sector-overviews/ai-in-law-legal-practice-current-applications/`

- discovery/diligence: find relevant documents during litigation, or during company acquisitions.
- legal research: find relevant statutes/caselaw to support arguments.
- contract analysis: document templates, find unusual or missing provisions.
- question answering: match clients with a lawyer who can answer it
- judicial analytics: predict judge decisions (not really NLP focused yet)

Argument Mining

Automated extraction of inference structure in natural language (more data is needed)

Argument from example	
Premise	In this particular case, the individual a has property F and also property G.
Conclusion	Therefore, generally, if x has property F, then it also has property G.

Argument from cause to effect	
Major premise	Generally, if A occurs, then B will (might) occur.
Minor premise	In this case, A occurs (might occur).
Conclusion	Therefore, in this case, B will (might) occur.

Practical reasoning	
Major premise	I have a goal G.
Minor premise	Carrying out action A is a means to realize G.
Conclusion	Therefore, I ought (practically speaking) to carry out this action A.

Argument from consequences	
Premise	If A is (is not) brought about, good (bad) consequences will (will not) plausibly occur.
Conclusion	Therefore, A should (should not) be brought about.

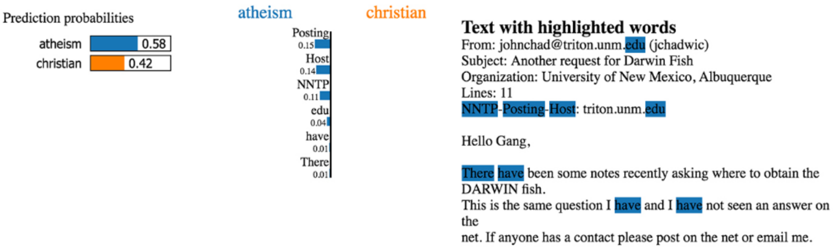
Argument from verbal classification	
Individual premise	a has a particular property F.
Classification premise	For all x, if x has property F, then x can be classified as having property G.
Conclusion	Therefore, a has property G.

Table 1.1: The five most frequent schemes and their definitions in Walton's scheme-set [12]

Interpreting Black Box Text Classifiers using LIME

1. Generate new texts by randomly *removing* words from the original document.
2. Form predictions $\hat{y}$ from black box model for these perturbed documents.
3. Train lasso on dataset of binary features for each word, equaling one if word appears, to predict $\hat{y}$.
   - weight by proximity to initial data point (one minus the proportion of words dropped)

```
exp = explainer.explain_instance(test_example,
classifier.predict_proba, num_features=6)
```



Pragmatics

When a diplomat says yes, he means 'perhaps';

When he says perhaps, he means 'no';

When he says no, he is not a diplomat.

Pragmatics

When a diplomat says yes, he means 'perhaps';

When he says perhaps, he means 'no';

When he says no, he is not a diplomat.

- language use depends on the context.
  - e.g. social identity, relationships, setting, conversation history, shared knowledge...
- this is mostly unexplored in NLP.

Quote Detection

Automated extraction of quotations and speaker

- Direct quotations are fully enclosed in quotation marks:
  - **X** said, "Taxes will go up next year."
- Indirect quotations paraphrase the original utterance:
  - **X** says that taxes will go up next year.
  - According to **X**, taxes will go up next year.
- Java package: `https://github.com/christianscheible/qsample`

Speech Acts

Some statements are meant to perform actions

"We hold the defendant guilty."

- **assertives** commit a speaker to the truth of the expressed proposition, e.g. reciting a creed
- **directives** cause the hearer to take a particular action, e.g. requests, commands and advice
- **commissives** commit a speaker to some future action, e.g. promises and oaths
- **expressives** express the speaker's attitudes and emotions towards the proposition, e.g. congratulations, excuses and thanks
- **declarations** change the reality in accord with the proposition of the declaration, e.g. baptisms, pronouncing someone guilty or pronouncing someone husband and wife

Important for legal NLP, but hardly any research about this

“Target-Based Speech Act Classification in Political Campaign Text”

Subramanian, Cohn, and Baldwin (2019), N = 258 docs, 6609 sentences:

Utterance	Speech act	Target party	Speaker
Tourism directly and indirectly supports around 38000 jobs in TAS.	*assertive*	NONE	LABOR
We will invest $25.4 million to increase forensics and intelligence assets for the Australian Federal Police	*commissive-action-specific*	LIBERAL	LIBERAL
Labor will prioritise the Metro West project if elected to government.	*commissive-action-vague*	LABOR	LABOR
A Shorten Labor Government will create 2000 jobs in Adelaide.	*commissive-outcome*	LABOR	LABOR
Federal Labor today calls on the State Government to commit the final $75 million to make this project happen.	*directive*	LIBERAL	LABOR
Good morning everybody.	*expressive*	NONE	LABOR
The Coalition has already delivered a $2.5 billion boost to our law enforcement and security agencies.	*past-action*	LIBERAL	LIBERAL
Malcolm Turnbull’s health cuts will rip up to $1.4 billion out of Australians’ pockets every year	*verdictive*	LIBERAL	LABOR

Table 1: Examples with speech act and target party classes. “Speaker” denotes the party making the utterance.

“Target-Based Speech Act Classification in Political Campaign Text”

Subramanian, Cohn, and Baldwin (2019), N = 258 docs, 6609 sentences:

Utterance	Speech act	Target party	Speaker
Tourism directly and indirectly supports around 38000 jobs in TAS.	*assertive*	NONE	LABOR
We will invest $25.4 million to increase forensics and intelligence assets for the Australian Federal Police	*commissive-action-specific*	LIBERAL	LIBERAL
Labor will prioritise the Metro West project if elected to government.	*commissive-action-vague*	LABOR	LABOR
A Shorten Labor Government will create 2000 jobs in Adelaide.	*commissive-outcome*	LABOR	LABOR
Federal Labor today calls on the State Government to commit the final $75 million to make this project happen.	*directive*	LIBERAL	LABOR
Good morning everybody.	*expressive*	NONE	LABOR
The Coalition has already delivered a $2.5 billion boost to our law enforcement and security agencies.	*past-action*	LIBERAL	LIBERAL
Malcolm Turnbull’s health cuts will rip up to $1.4 billion out of Australians’ pockets every year	*verdictive*	LIBERAL	LABOR

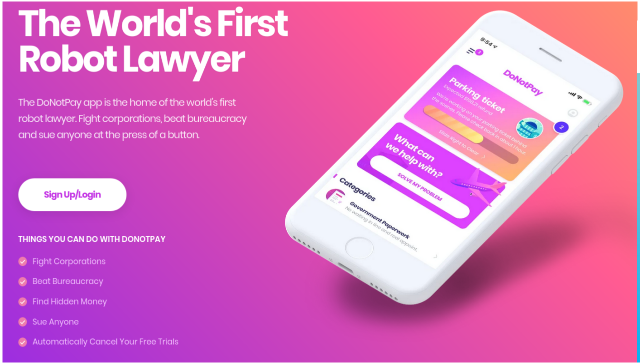
Table 1: Examples with speech act and target party classes. “Speaker” denotes the party making the utterance.

Speech act	%	Kappa (κ)
assertive	40.8	0.85
commissive-action-specific	12.4	0.84
commissive-action-vague	6.6	0.73
commissive-outcome	4.9	0.72
directive	1.7	0.92
expressive	1.9	0.88
past-action	6.3	0.76
verdictive	25.4	0.82

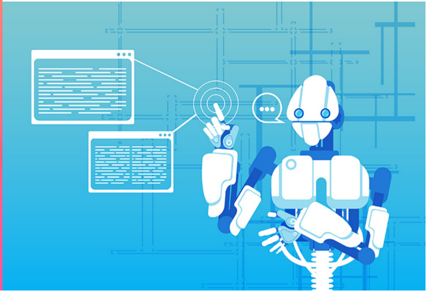
Table 3: Speech act agreement statistics

Speech act	MLP_{ELMo}	Our approach
assertive	0.77	0.80
commissive-action-specific	0.65	0.69
commissive-action-vague	0.45	0.48
commissive-outcome	0.28	0.39
directive	0.58	0.59
expressive	0.55	0.58
past-action	0.45	0.48
verdictive	0.48	0.61

Table 6: Speech act class-wise F1 score.



Your Court-Appointed Chatbot – Is Artificial Intelligence Threatening the Legal Profession?



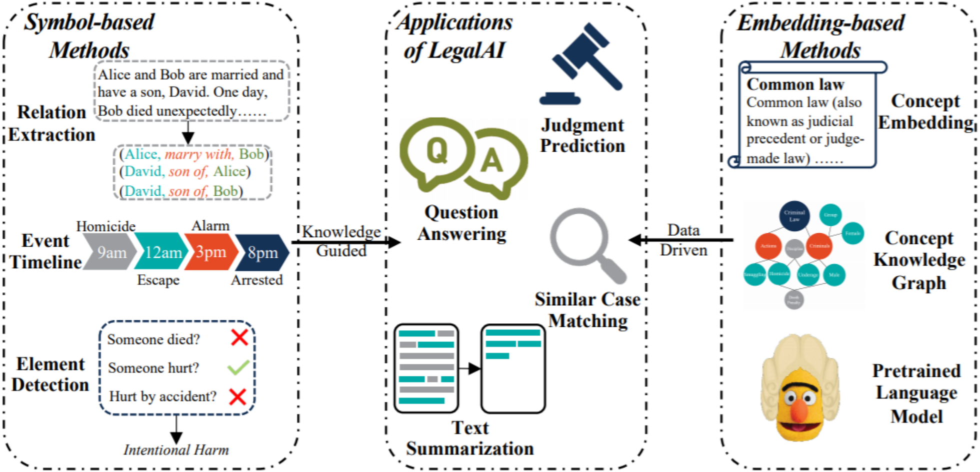


Figure 1: An overview of tasks in LegalAI.

https://arxiv.org/pdf/2004.12158.pdf

Style Transfer (Wegmann et al 2022)

- Wegmann et al (2022) use contrastive author prediction while controlling for content, to isolate the style dimensions in text:

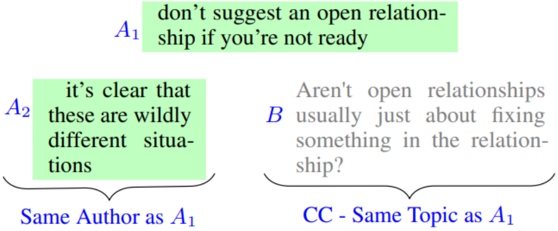


Figure 1: **Contrastive Authorship Verification (CAV) Setup and Content Control (CC) Variable.** The CAV task is to match A_1 with the utterance A_2 that was written by the same author. Contrary to the traditional authorship verification task (AV), this is complemented by a third "constrastive" utterance that was written by a different author (B). In addition to the CAV variation to AV, we experiment with content control (CC) by selecting B and A_1 to have the same approximate content with the help of a topic proxy. As topic proxies we use conversation and domain information.

Dangers of Legal NLP systems

- We discussed previously how GPT might flood the internet with machine generated text, e.g. fake news
  - is there a similar risk with legal language models?

Limits of Judicial Support Systems

Limits of Judicial Support Systems

- (Lack of) transparency in judicial support systems:
  - Closed-source algorithms result in "black box justice" and could be abused by insiders.
  - But open-source algorithms are prone to gaming: savvy attorneys could "trick" the algorithm.

Limits of Judicial Support Systems

- (Lack of) transparency in judicial support systems:
  - Closed-source algorithms result in “black box justice” and could be abused by insiders.
  - But open-source algorithms are prone to gaming: savvy attorneys could “trick” the algorithm.
- NLP systems will only attend to evidence that appear in a lot of cases, will ignore special/mitigating circumstances.

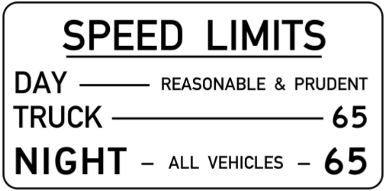
Limits of Judicial Support Systems

- (Lack of) transparency in judicial support systems:
  - Closed-source algorithms result in “black box justice” and could be abused by insiders.
  - But open-source algorithms are prone to gaming: savvy attorneys could “trick” the algorithm.
- NLP systems will only attend to evidence that appear in a lot of cases, will ignore special/mitigating circumstances.
- NLP systems do not generalize to new types of cases.
  - e.g., judicial prediction systems would not account for new laws/legislation.

Limits of Judicial Support Systems

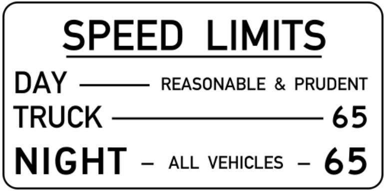
- (Lack of) transparency in judicial support systems:
  - Closed-source algorithms result in "black box justice" and could be abused by insiders.
  - But open-source algorithms are prone to gaming: savvy attorneys could "trick" the algorithm.
- NLP systems will only attend to evidence that appear in a lot of cases, will ignore special/mitigating circumstances.
- NLP systems do not generalize to new types of cases.
  - e.g., judicial prediction systems would not account for new laws/legislation.
- Teaching a legal NLP system to understand rare evidence, and to understand new laws, would require something much closer to **legal artificial intelligence**.

Legal Vagueness and Value Judgments



- Even if the AI could read new laws, there is the problem of legal vagueness:
  - How will the AI decide in this circumstance?

Legal Vagueness and Value Judgments



- Even if the AI could read new laws, there is the problem of legal vagueness:
  - How will the AI decide in this circumstance?

- Making choices in the presence of vagueness or indeterminacy requires value judgements.
  - What counts as a "good" outcome? Is it even measurable?

Philosophical Issues

- What does it mean to surrender the implementation of legal interpretation and judicial decision making to machines?
- What are the long-term implications for the system and its adaptiveness to change?
  - what are the political and cultural impacts?
  - how does it affect trust in the system and motivation to appeal?

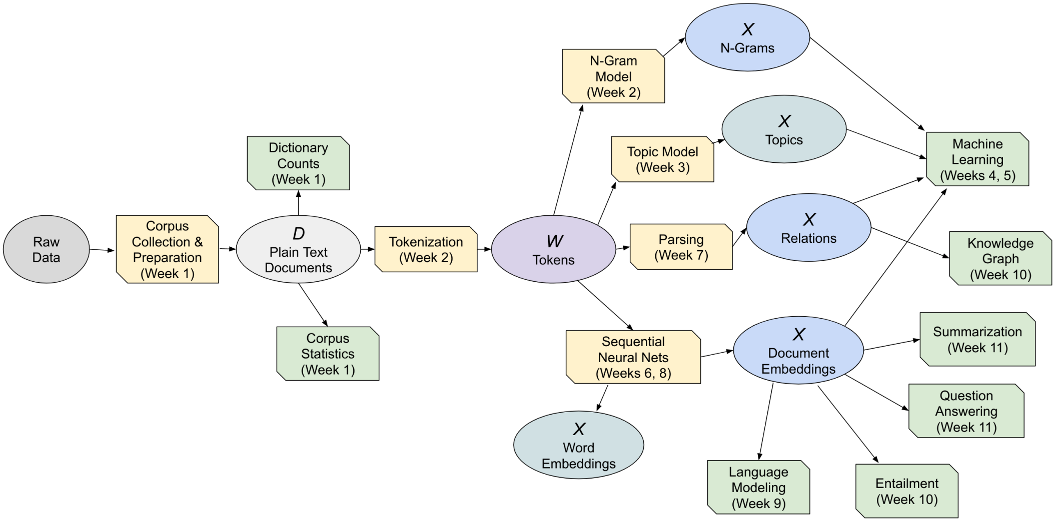
Outline

Tools for Legal NLP

Wrapping Up

- We focused on **natural language processing** in **law** and **social science**.

- We focused on **natural language processing** in **law** and **social science**.
- Learning objectives:
  1. **Implement and evaluate natural language processing pipelines.**
  2. **Apply NLP tools to support legal practice.**
  3. **Understand how (not) to use NLP tools for measurement in social science.**



Raw
Data
Corpus
Collection &
Preparation
(Week 1)
D
Plain Text
Documents
Dictionary
Counts
(Week 1)
Corpus
Statistics
(Week 1)
Tokenization
(Week 2)
W
Tokens
N-Gram
Model
(Week 2)
X
N-Grams
Topic Model
(Week 3)
X
Topics
Parsing
(Week 7)
X
Relations
Sequential
Neural Nets
(Weeks 6, 8)
X
Document
Embeddings
X
Word
Embeddings
Machine
Learning
(Weeks 4, 5)
Knowledge
Graph
(Week 10)
Summarization
(Week 11)
Question
Answering
(Week 11)
Language
Modeling
(Week 9)
Entailment
(Week 10)

Final Assignment

- content based on the slides and required readings
- If you have been keeping up, it should not take more than 2-3 hours.
- otherwise, could take longer.
- but in any case you will have a week to do it:
  - posted at noon June 6th, closes at noon June 13th.
  - can ask clarifying questions to me by email before noon June 14th, I will provide answers to the whole class.

Next Term: "Building a Robot Judge" Course

- In the fall term, I teach a complementary course focusing on machine learning and causal inference:
  - "Building a Robot Judge: Data Science for Decision-Making" (851-0760-00L)
- Not a lot of overlap:
  - non-text data (tabular datasets, computer vision)
  - a lot more on causal inference
  - focus on how predictions / causal estimates can support decision-making
- Similar setup in terms of course credits:
  - 3 credits for the lectures/assignments, 2 additional credits for a project.

Stay in touch

- e.g. add me on LinkedIn
- let me know if anything in this course helps you later on!
- can provide references for your work in the course.

Stay in touch

- e.g. add me on LinkedIn
- let me know if anything in this course helps you later on!
- can provide references for your work in the course.

Thanks!