

The Research Frontier: AI and Legal Reasoning

Alina Petrova, John Armour, Thomas Lukasiewicz, Yordan Yordanov

AI for English Law Webinar: LawTech – Progress and Challenges

15 September 2021



**INDUSTRIAL
STRATEGY**

UK Research
and Innovation





Motivation

- Automated solutions that scale can **dramatically** lower user costs

Example: First mark-up of a standard type of contractual agreement, which might be charged at £1,000+ by a lawyer working in a top-tier law firm, can today be done by an AI system for less than £1.

⇒ For citizens, facilitate access to justice

⇒ For businesses, lower costs



Research Question

- How can we harness AI to give (better) guidance on dispute resolution?



Potential use-cases

Advice

What are my
legal rights?

What should I
do?

Strategy

Should I
settle/fund this
suit?

For how much?

Resolution

Which party
should succeed?

For what
reasons?

Review

Should a
first-instance
decision be
subject to appeal
or review?

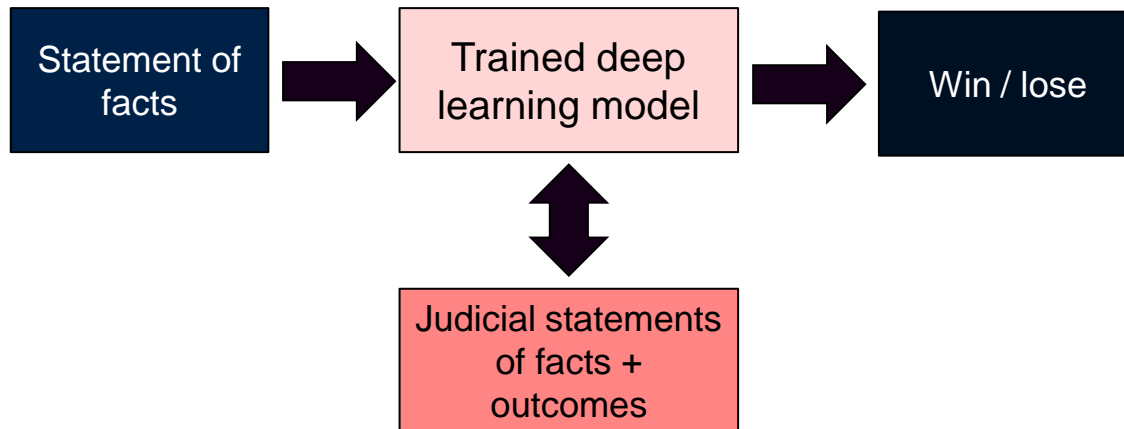


State of the art





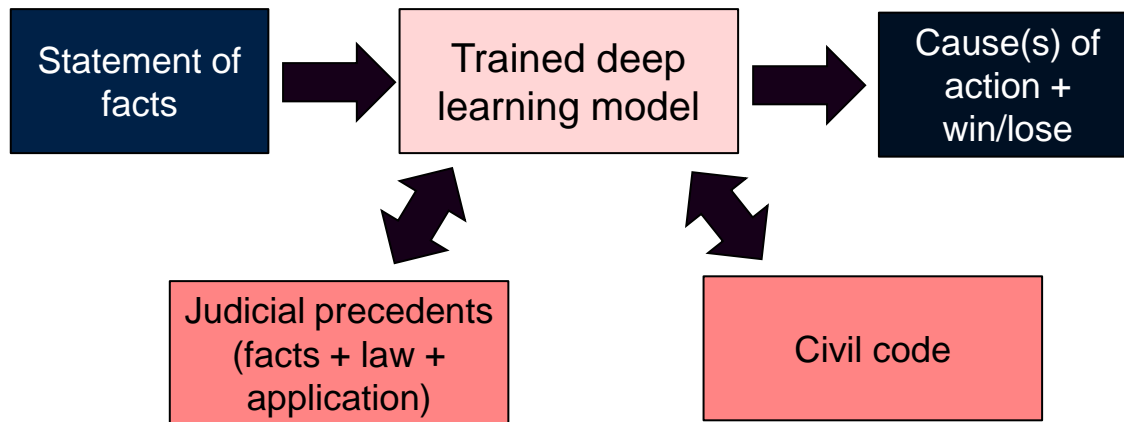
Simple outcome prediction



- Prediction = outcome: claimant win/lose (eg Aletras et al, 2016)
 - Around 70-80% accuracy can be achieved if exclude cases with multiple issues.
 - Applied to ECtHR caselaw (published in structured format)
 - No explanation for predicted outcome
- ⇒ **Research frontier:** how to apply to caselaw not published in a structured format? How to explain predictions?

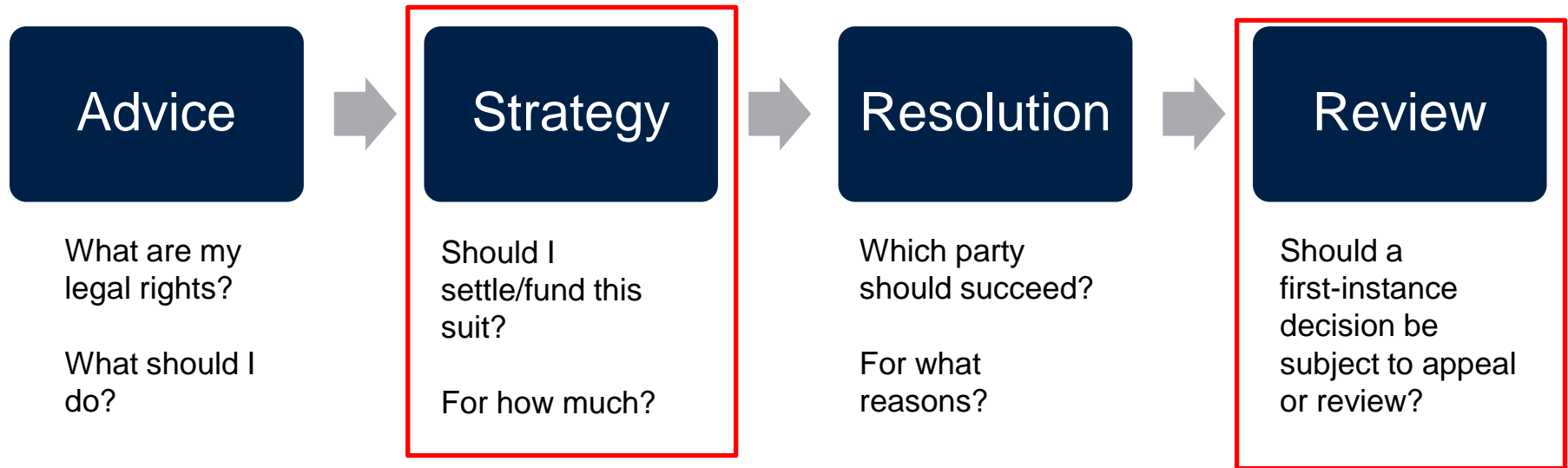


Cause of action identification



- Prediction: which causes of action are relevant
 - So far, only implemented with Chinese legal data (using civil code) (Yang et al, 2019)
 - Civil code acts as authoritative “map” of legal causes of action
- ⇒ **Research frontier:** Implement in a common law system.

Potential use-cases






Data

Approaches to digital access to caselaw



China:
Over 88m court documents freely available online at [central repository](#); state sponsorship of legal AI competitions (CAIL2018, 2019)




USA: many court pre available online; centralised repository PACER has dock information for many c but charges \$0.10/p Harvard Law School [Caselaw Access Project](#) 6.4m US cases



Data Gazette
NEWS ANALYSIS LAW PRACTICE IN-HOUSE PEOPLE
NEWS
BAILII grants access to judgments for mass AI analysis
By Michael Cross | 14 December 2020
4 Comments

England and Wales court judgments are for the first time being opened to mass analysis by artificial intelligence, the Gazette can reveal. Under an agreement announced today, the British and Irish Legal Information Institute (BAILII) has granted academics at Oxford University bulk access to its database of 400,000 judgments for research purposes.

Until now, BAILII, the principal online repository of case law, has refused to allow its database to be downloaded in bulk or 'scraped' by software. The fear is that this would enable the development of AI systems for predicting the outcome of cases on a judge-by-judge basis. It is understood that the agreement between BAILII and Oxford's AI for Law project excludes such use. Rather, the agreement will 'unlock new research insights into English case law' and help the development of ways to improve access to legal information, a statement said. The negotiations with BAILII took over a year and encompassed review by stakeholders at the Ministry of Justice and the judiciary, it added.



on using data in (Loi t 33)), o 5 years nt

Most open



Least open

Stakeholders at the Ministry of Justice and the judiciary, it added.



Data Challenges

Hope:

Reality:

Case description: On July 7, 2017, when the defendant Cui XX was drinking in a bar, he came into conflict with Zhang XX..... After arriving at the police station, he refused to cooperate with the policeman and bited on the arm of the policeman.....

Result of judgment: Cui XX was sentenced to 12 months imprisonment for creating disturbances and 12 months imprisonment for obstructing public affairs.....

- Charge#1 creating disturbances term 12 months
- Charge#2 obstructing public affairs term 12 months

Held: The court found in favor of the plaintiff, assigning numerous errors to the findings of fact made by the Trial Court and its refusal to make certain findings as to the matters in issue. The court found that the findings made by the court were fully justified by the evidence, and contain all the material facts necessary to an adjudication of the matters in issue. This court has affirmatively established the rule that where the findings of fact of the Trial Court are based upon substantial evidence to sustain them, they will not be disturbed by this court. The statement of facts above given is substantially as found by the Trial Court, and there seems to be no dispute as to the correctness of its conclusions of law as to the invalidity of the tax sales of the property to Henry and to his assignors, or as to the invalidity of the deed from the Arrijo's to Henry, which was placed in Moore's hands, but never delivered to Henry with Arrijo's consent. These conclusions of the Trial Court and that part of the decree setting them aside will be set aside as error and consequently will not be considered by this court. The only remaining questions are: 1st, whether there was a failure, or partial failure, of the consideration of the \$2,829.68-note from Arrijo and wife to Henry; and, 2nd, whether the computation made by the court of the amount due Henry from the Arrijos was correct. 1. The Trial Court in its decree as to the note for \$2,829.68 says: (P. 248 Tr.) 'said note is valid and given upon a valuable consideration to the extent of \$1,122.44, being the sum of \$444.00 paid to said Johnston by said Henry on April 3rd, 1897, and \$691.11, the amount paid by said Henry at said tax sale on July 6th, 1894, with interest thereon at 6 per cent per annum, to the date of said note, and is invalid and without consideration as to the remainder thereof,' and this note is treated in the decree as though the principal sum at that date was for said amount of \$1,122.44, instead of \$2,829.68, as appears upon the face of the note. But was the remainder of the principal expressed invalid and without consideration? The evidence shows that at the time this note was given the two former notes were long past due and a suit to foreclose the first two deeds of trust was then pending in the District Court of Bernalillo County, New Mexico, these cases were subject to dismissal for failure to file copies of his notes and trust deeds, but the way for their foreclosure was still open and Henry had at least shown his intention to foreclose. The tax certificates which he held against, these premises, with the three per cent interest added in, both he and Arrijo believed to be valid claims against the Arrijos, and liens upon the property as well, as the certificate held by Johnston. An attempt was made by these parties to adjust these tax sale claims. Arrijo testified that he wanted me to give a note for twelve per cent; it was too much to be paying three per cent a month and I signed it to stop three per cent a month. (P. 128 Tr.) There is no question but a valid tax sale certificate was at that time did drop three per cent, per month and there is no claim that Henry had any knowledge of the invalidity of the tax sale. The question of its validity was equally open to both parties and there is no claim of any fraud or misrepresentation on the part of Henry, or that there was an unfair advantage taken of Arrijo. Nor is it claimed that there was any mistake in the calculation of the amount agreed upon as due Henry. Had Arrijo been able to have paid the money at that time he would have unquestionably paid the amount named in the note and taken up the certificates, which, according to Henry's undisputed testimony, was the original intention. It was a complete settlement of these tax sale certificates between the parties. According to Henry's testimony the certificates were turned over to Arrijo's agent, and we think we are justified in so finding. It is difficult to see then wherein there was a partial failure of consideration for this note. These certificates were outstanding against his property and had not been declared void by any court, nor could they have been without a suit in equity brought by him for that purpose. He obtained six months time in which to pay them off at a reduced rate of interest, as shown by the note itself, which extension of time was in itself a good consideration. (P. 171, and cases cited. Equity will not interfere and declare a failure of consideration in whole or in part except in cases where the money could have been recovered back if paid. It is settled in law, and the rule has been followed in equity, that money paid under a mistake of law with respect to the liability to make payment, but with full knowledge, or with means of obtaining knowledge of all the circumstances cannot be recovered back.' 2 Pom. Eq. Jurisprudence (3rd Ed.) Sec. 851; Painter v. Park Co., 81 Ia. 242; Alton v. First Nat. Bank, 137 Mass. 341; Frain v. Nicoll, 39 Minn. 401; Gillson v. Alford, 40 Tex. 207; Board v. Barr, 23 W. Va. 456. In Perkins v. Trumb, 38 Minn. 243, plaintiff held a tax deed upon lands occupied by defendants and which he claimed as owner, plaintiff compromised by giving a note secured by a mortgage on the land. The Supreme Court of Minnesota afterwards declared tax deeds such, as plaintiff held void and as vesting no title and this was set up as a defense to the note, but the Supreme Court said: 'Whether it is any defense that it was afterwards judicially determined that tax deeds of this form are void, where parties whose rights are questionable and doubtful, and who have equal means of ascertaining what their rights are, come together and settle these rights among themselves, a court must enforce the agreement' to which they may fairly come at the time, although a judicial decision should afterwards be made showing that these rights were different from what they then really had, nothing in the fore-going.' True, the compromise by which the note was given was made after suit was brought for possession under the void tax deed, but I can see no difference in principle between it and the case at bar. In each case the parties attempted to and did settle their supposed rights between themselves, waiving any legal rights either party might have claimed, and it can make no difference what might have been established by a judicial determination of their claims. The agreement between Henry and Arrijo was in the nature of a compromise, and an agreement between two or more persons, and it is not a law suit, amicably settled their differences on such terms as they can agree upon.' 6 Am. and Eng. Enc. of Law (2nd Ed.) 418. Arrijo testifies (P. 158 Tr.): 'He said he was entitled by law to collect but was willing to take the low per cent a year; and that for that reason he executed the note in question to stop the three per cent per month Henry then believed Henry to have above quoted language presumably to avoid litigation. It is almost universally held that such a consideration is good, in the absence of fraud. Northern Liberty Market Co. v. Kelly, 113 U. S. 159; Booth v. Dexter Steam Fire Engine Co., 118 Ala. 369; Richardson v. Constock, 21 Ark. 69; Kane v. Barnes, 181 Ia. 382; French v. French, 81 Ia. 565; Keyes v. Mum, 81 Ia. 566; Cobb v. Arnold, 8 Neb. 489; Pickett v. St. L. Chamber of Commerce, 88 Mo. 65; Housatonic Nat. Bk. v. Foster, 85 Mo. 376; White v. Nutt, et al., 72 W. Va. 365. This case does not fall within the rule laid down in Briscoe v. Knealy, 8 Mo. App. 67, for in that case the portion of the note held to be void for failure of consideration was an amount added to the settlement and no part of the settlement itself, while in the case at bar the whole sum included in the note was included in the settlement and agreed to by the parties before the note was executed. Waters, 38 Ga. 344. Cited by the appellee, in point, for in that case the contract declared to be void was separate and distinct from the main contract and based upon an entirely separate consideration. In our judgment the whole sum is but one consideration, the amount agreed upon as owing from Arrijo to Henry and the entire note must be taken together and stand or fall together as to failure of consideration. It is contended that the three per cent allowed in the settlement on the tax sale certificates is usury, and that the note is therefore tainted with usury, and void for that reason as to the amount of the three per cent, penalty. But the agreement upon which this note is founded lacks the essential element of usury, that is, the intent to exact more than legal interest. 29 Am. and Eng. Enc. of Law (2nd Ed.) 461; Kane v. Wagner, 9 Peters 378; Hamilton v. Palmer, 5 Wall. 684; Call v. Palmer, 116 U. S. 98. At the time the note was made both parties believed Henry was entitled to the three per cent, per month interest, or rather penalty on the same paid at the tax sale, and there was no intent on the part of either to take or pay any sum not allowed by law, and it is said that the question whether a contract is usurious is to be decided with reference to the time when it was entered into. 29 Am. and Eng. Enc. of Law (3rd Ed.) 448; Pollard v. Baylors, 6 Huff. (Va.) 431. Appellee is inconsistent in the claim, for he concedes the justice of the claim for \$444.00 (the amount paid for the Johnston certificate), though that amount contains the three per cent, from the date of sale, and is to that extent usurious, if it is usury. We conclude, therefore, that the note for \$2,829.68 was a valid note for the consideration expressed on its face, and that Henry should have decreed the entire sum, with interest at twelve per cent, per annum, as provided by its terms. 2. The rule by which the court arrived at the amount due Henry is that known as the Massachusetts Rule, and is the correct one for computing the interest and applying the payments made by Arrijo, as we understand the Trial Court's method. This court 50 held in the recent case Jones, Down & Co. v. Chandler, 85 Pac. 392, and we are content with the doctrine there laid down. 3. There 'is a cross appeal on the part of Arrijo from the decree of the court allowing ten per cent of the amount found due to attorney's fees, and the allowance of interest after its computation, and the allowance of interest after May 25th, 1894, the date of the alleged tender from Arrijo to Henry. The first contention on the part of cross-appellant is that attorney's fees should have been allowed as a lien against the property, for the reason that the deeds of trust do not provide for attorney's fees, citing Cichel v. Carrillo, 42 Cal. 493, and Stever v. Joneyacke, 9 Kan. 367, as authority therefor. The former case does not discuss

Chen et al. Charge-Based Prison Term Prediction with Deep Gating Network. 2019



Challenges (cont'd)

Data challenges:

- ⑩ No access to claim forms/pleadings or settled cases.
- ⑩ Lack of metadata and annotations.
- ⑩ Information extraction: outcomes, facts, legal reasons.

Modelling challenges:

- ⑩ Manual annotation leads to small sample size.
- ⑩ Most data is out-of-sample: different judges, courts, countries.
- ⑩ Lack of first instance cases; appellate data means less informative.











Experiments and Results





Step 1: Annotation

-  *facts* (what initially happened)
-  *procedural history* (inc. what lower court ruled)
-  *relevant precedents* (which prior precedents applied)
-  *application of law* (how law is applied to the facts)
-  *outcome*
-  *other discussion*





Manual annotation

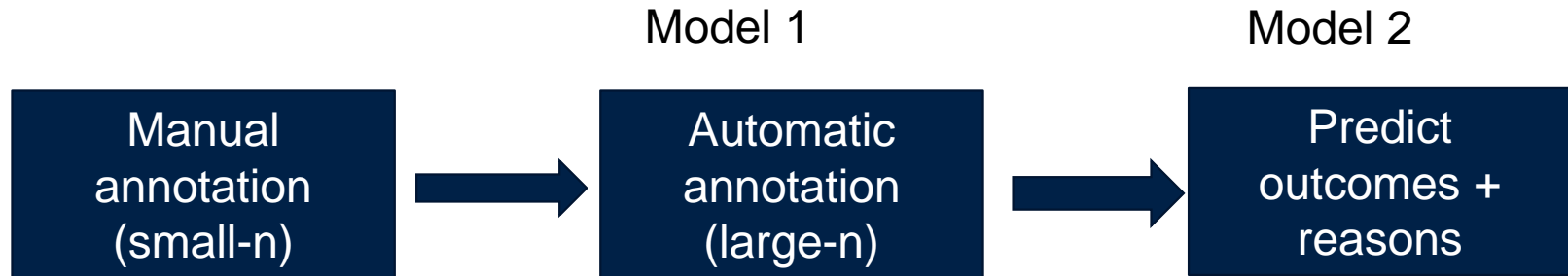
- Team of 3 RAs (law research students)
- Multiple iterations to converge on equivalent coding
- Coding protocols continuously developed
- Weekly meetings / problem workshops

- Working total: c. 500 cases





Contributions





Automating annotation: outcomes

Extracting Outcomes from Appellate Decisions in US State Courts

Alina PETROVA¹, John ARMOUR and Thomas LUKASIEWICZ
University of Oxford, UK

Abstract. Predicting the outcome of a legal process has recently gained considerable research attention. Numerous attempts have been made to predict the exact outcome, judgment, charge, and fines of a case given the textual description of its facts and metadata. However, most of the effort has been focused on Chinese and European law, for which there exist annotated datasets. In this paper, we introduce CASELAW4 — a new dataset of 350k legal case reports from the U.S. Caselaw Access Project, of which 250k are automatically annotated with the binary outcome labels of AFFIRM or REVERSE. To our knowledge, it is the first attempt to perform outcome extraction (a) on such a large volume of English-language judicial opinions, (b) on the Caselaw Access Project data, and (c) on US State Courts of Appeal cases, and it paves the way to large-scale outcome prediction and advanced legal analytics using U.S. Case Law. We set up baseline results for the outcome extraction task on the new dataset, achieving an F-measure of 82.32%.

Keywords. legal analytics, outcome extraction, legal reasoning, outcome prediction

Petrova et al (2020)

- 500 randomly-selected cases manually annotated for outcome (affirm, reverse, mixed)
- Deep learning model trained using hand-coded sample
- 250,000 cases automatically annotated with 82.3% accuracy



Automating annotation: facts

Legal or Non-Legal: Using External Data for Legal Fact Extraction

Abstract

Legal outcome prediction is the task of forecasting the outcome of a court case given the information about the case. The bottleneck of outcome prediction is automatically identifying facts in legal texts, which has not yet been solved for the English language case law. While generating high-quality annotated legal data is prohibitively expensive, we explore data augmentation with additional legal and general domain data sources in the context of legal fact extraction. We demonstrate that both types of external data can improve fact extraction, and we achieve a 7.8% and 5% improvement in model F1 score by adding out-of-domain legal cases and Wikipedia articles, respectively, to the training data.

the outcome is formally stored in the metadata. On the other hand, resources such as the U.S. Caselaw Access Project² or the UK's BAILII³ provide raw, unstructured texts with very limited metadata.

Extracting facts and outcomes from a court case proceeding is the bottleneck of legal outcome prediction and other analytics tasks (Medvedeva et al., 2020; Chalkidis et al., 2021; Liu et al., 2019b; Thomas and Sangeetha, 2021). While outcome extraction has been successfully attempted before (Petrova et al., 2020), legal fact extraction remains an unsolved problem for the English-speaking legal systems. The problem is not trivial: facts should be separated from procedural history and legal discussion so that the training data for a prediction model is distilled from legal arguments, outcomes or tes-

Petrova et al (2021)

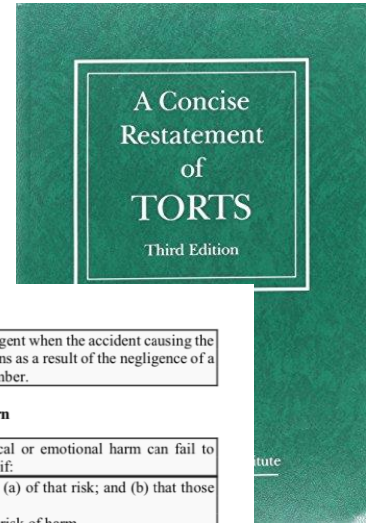
- 70 negligence cases from CAP, annotated for legal facts
- Data augmentation with ECHR cases, news reports, Wikipedia articles
- Fact extraction accuracy improves by 7.8% with additional legal data and by 5% with non-legal data



Predicting Outcome + Reasons

Step 2: Giving context via:

- ⑩ Legal issues
what legal issues are relevant to the facts? (Restatement)
- ⑩ Prior caselaw
which prior cases are relevant to the facts?
- ⑩ Reasons: application of law to facts
what reasons are given for application of law to facts in similar cases?



§ 17 Res Ipsa Loquitur

The factfinder may infer that the defendant has been negligent when the accident causing the plaintiff's harm is a type of accident that ordinarily happens as a result of the negligence of a class of actors of which the defendant is the relevant member.

§ 18 Negligent Failure to Warn

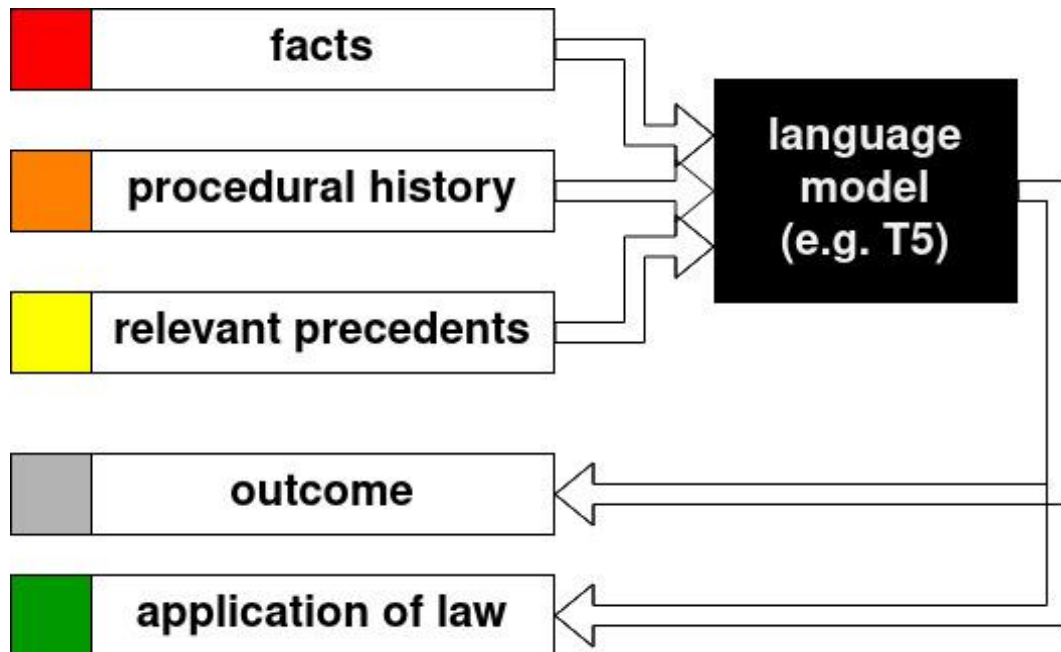
(a) A defendant whose conduct creates a risk of physical or emotional harm can fail to exercise reasonable care by failing to warn of the danger if:

- (1) the defendant knows or has reason to know: (a) of that risk; and (b) that those encountering the risk will be unaware of it; and
- (2) a warning might be effective in reducing the risk of harm.

(b) Even if the defendant adequately warns of the risk that the defendant's conduct creates, the defendant can fail to exercise reasonable care by failing to adopt further precautions to protect against the risk if it is foreseeable that despite the warning some risk of harm remains.



Predicting Outcome + Reasons



The language model is pre-trained on a large corpus of multi-domain text.

Relevant precedent retrieval could be an intermediary step.



Applications?





Potential use-cases

Advice

What are my
legal rights?

What should I
do?

Strategy

Should I
settle/fund this
suit?

For how much?

Resolution

Which party
should succeed?

For what
reasons?

Review

Should a
first-instance
decision be
subject to appeal
or review?