ARTIFICIAL INTELLIGENCE:

Thinking About Law, Law Practice, and Legal Education

DUQUESNE UNIVERSITY
School of Law

Presenting Sponsor: Reed Smith
Driving progress through partnership

gravity stack
Delivering Collective Intelligence in Law

LexisNexis

COZEN O'CONNOR
CONNECTING CASE TEXTS AND COMPUTATIONAL MODELS OF LEGAL REASONING

Kevin D. Ashley
Professor of Law and Intelligent Systems
University of Pittsburgh
School of Law
Intelligent Systems Program
Learning Research and Development Center
ashley@pitt.edu
OVERVIEW

1. Defining AI and Law and legal text analytics
2. Three techniques of legal text analytics with apps
   a) Machine learning from legal text collections
   b) Legal network diagrams
   c) Legal question answering
3. Three limitations of legal text analytics
4. Can computational models of case-based reasoning enable legal apps to explain and argue?
5. Meeting the need for manual text annotation
6. Conclusions
1. DEFINITIONS

• Artificial Intelligence and Law (AI & Law):
  • Computer science subarea where researchers build computational models of legal reasoning behaviors.

• Legal Text Analytics (or text mining):
  • Employs natural language processing (NLP), machine learning (ML) and other computational techniques to automatically extract meanings (or semantics) from archives of legal case decisions, contracts, or statutes.

• Argument mining:
  • Focuses on text analytic discovery of argument-related information in text corpora.
  • Premises and conclusions, argument / counter arguments, sentence roles in legal arguments, substantive strengths or weaknesses of a claim.
2. THREE TECHNIQUES OF LEGAL TEXT ANALYTICS

A. Machine learning from legal text collections:
   • Refers to computer programs that use statistical means to induce or “learn” models from data with which they can classify a document or predict an outcome for a new case.

B. Legal network diagrams:
   • Graphs of the relations between “legal objects” such as cited legal cases or statutes, concepts referred to by statutory provisions, or communications links.

C. Legal question answering (Q/A) systems:
   • Searching large text collections to locate documents, short phrases or sentences that directly answer a user’s law-related question.
A. CLASSIFICATION VIA (SUPERVISED) MACHINE LEARNING

ML classifiers are *trained* in order to later *predict* labels of documents or case outcomes.

• Training step: ML algorithm takes as input ... 
  • Chunks of text (e.g., sentences) of cases in the training set represented as *feature vectors* and a *target label* (e.g., a binary decision whether a classification applies.)
  • It statistically “learns” the correspondence between certain language features in the sentence feature vectors and the target label.

• Prediction step
  • Given the texts of new chunks (i.e., from the test set) also represented as feature vectors, it predicts the classification to assign to the sentence, if any.
  • Evaluate by comparing the learned classification to the manual one.
MACHINE LEARNING PREDICTS LEGAL OUTCOMES

• Input case features; model outputs likely result.

• Katz and Bommarito’s SCOTUS ML program.
  • Correctly forecasts 70% of case outcomes and 71% of justice level vote outcomes over a sixty-year period.
  • Using ML model (extremely randomized forest of decision trees)

• Lex Machina:
  • Statistical learning model (logistic regression)
  • predicts outcomes of Intellectual Property (IP) cases
  • based on litigation participant-and-behavior info extracted from case texts.
  • Results in 64% accuracy.

• Neither employs features that reflect substantive legal merits.

• To what extent can the features for ML be extracted from texts?
Case Information
Admin Action [S]
Case Origin [S]
Case Origin Circuit [S]
Case Source [S]
Case Source Circuit [S]
Law Type [S]
Lower Court Disposition Direction [S]
Lower Court Disposition [S]
Lower Court Disagreement [S]
Issue [S]
Issue Area [S]
Jurisdiction Manner [S]
Month Argument [FE]
Month Decision [FE]
Petitioner [S]
Petitioner Binned [FE]
Respondent [S]
Respondent Binned [FE]
Cert Reason [S]

Justice and Court Background Information
Justice [S]
Justice Gender [FE]
Is Chief [FE]
Party President [FE]
Natural Court [S]
Segal Cover Score [SC]
Year of Birth [FE]

Trends (decision directions conditioned on legal issue area, recent/prior term, cumulative terms)
Overall Historic Supreme Court [FE]
Lower Court Trends [FE]
Current Supreme Court Trends [FE]
Individual Supreme Court Justice [FE]
Differences in Trends [FE]
LEX MACHINA (Surdeanu, et al. ICAIL 2011)

• Predicts outcomes of IP claims based on corpus of all IP lawsuits in 10+ year period.
  • Focused on patent infringement cases.
  • Begun at Stanford U. by Law Prof. Mark Lemley, et al. Acquired by LexisNexis

• Represents case features as litigation participants / behavior:
  • lawsuit parties, attorneys and law firms, judges assigned to case, districts where complaints were filed.

• Most significant contributions to accuracy:
  Identities of judge and plaintiff’s law firm > defendant’s identify > district > def.’s law firm > def.’s attorney.

• Model is “agnostic to the merits of the case”!
  • Litigation participant-and-behavior features indirectly capture aspects of case merits.
  • Could it do better if features took legal merits into account?

• Fairly easy to extract automatically from case texts. 3 IP experts annotated cases as to outcomes.
• It cannot explain its predictions in terms of substantive legal merits.
B. DRAW LEGAL INFERENCES WITH NETWORK DIAGRAMS

Graphs of relations between objects.

1. *Citation network*:
   • citation relations among legal cases or statutory provisions. e.g., Ravel, Casetext, Fastcase

2. *Social network*:
   • communication relations among entities,
   • e.g., who communicated with whom about what and when:
     • connections among senders and receivers of corporate emails in e-discovery.

3. *Statutory network*:
   • relations among entities referred to by, and subject to, particular kinds of regulation across multiple statutes and jurisdictions.
LITIGATION SUPPORT WITH CITATION NETWORKS

• Citation relations among legal cases or statutory provisions.
• Ravel: U.S. case texts accessible in visual maps showing inter-case citations re a legal concept.

• CaseText’s CARA:
  • Input a brief (i.e., written memorandum of law) ...
  • CARA outputs / summarizes additional cases to cite in support of arguments in the brief.
  • Based on citation networks.

• Fastcase:
  • Interactive timeline shows search results over time, relevance to query, citation frequency both generally and within most relevant cases.

• Powerful combination of ML and network diagrams
  • Use text analytics for more info about citation links between citing/cited cases.
  • Identify topic of paragraph in citing/cited cases.
  • Finer grained links among particular paragraphs in citing/cited cases begin to represent why a case is cited.

Copyright 2019. Kevin D. Ashley
Search Results: The Visualization Map

Keys to your search results

- Your search returns a list of cases, ranked by relevance on the right hand side of the results screen. Ravel underlines your search terms in yellow.

- If you have not already filtered your search by jurisdiction, you can add a jurisdiction filter by clicking the “Jurisdictions” button near the top of the screen.

- A visual map of the 75 most relevant cases returned by your search displays on the left. There are two filters: Court and Relevance discussed below.

- Each circle in the visual map represents a case.

- The size of a circle is determined by the number of citations to that case from other cases within the search results; more citations results in a larger circle, less citations in a smaller one.

- The lines that connect two cases show a citation, arrows point from the citing case to the cited case, and the thickness of the line represents the depth of treatment.

BUCKLEY v. VALEO
424 U.S. 1 January 28th, 1976 Cited By: 2,665
United States Supreme Court

- addressed broadly the problem of political campaign financing. It wished to promote full disclosure of—Fortson, supra, at 439. 1. General Election Campaign Financing Appeals insist that Chapter 95 falls short—must also be rejected. 3. Primary Election Campaign Financing Appeals’ final challenge is to the co-administering provisions for Presidential campaign financing, the Commission may authorize convention—of Congress In this comprehensive scheme of campaign finance. By dissecting the Act bit by bit, and casting


CITIZENS UNITED v. FEDERAL ELECTION COM’N
130 S. Ct. 876 January 19th, 2010 Cited By: 485
United States Supreme Court

- flawed historical account of campaign finance laws, see Brief for Campaign Finance Scholars as Amici Curiae;—permit laws that force speakers to retain a campaign finance attorney, conduct demographic marketing research—their reach is chilled. See Part II-A, supra.

Campaign finance regulations now impose “unique and complex See B. Smith, Unfree Speech: The Folly of Campaign Finance Reform 23 (2001). Yet not until 1947 did Congress—associations under a regulatory scheme for campaign financing”). Seizing on this aside in Bellotti’s footnote—

FOOTNOTES: — (interest’s); Strauss, Corruption, Equality, and Campaign Finance Reform, 94 Colum. L.Rev. 1369, 1369, and n. of its continued application. It is gutting campaign finance laws across the country, as the Court does scholarship challenging the historical account of campaign finance law given in United States v. Automobile Workers—cited by the majority is correct that certain campaign finance reforms were less deliberate or less benign—no response to any legislature that takes campaign finance regulation seriously. It merely illustrates
Query: “Allergy” + Complaint seeking damages for a death caused by exposure to peanut oil
How this document has been cited

- "...the Supreme Court noted in *Citizens United* that the suppression of political speech harms not only the speaker, but also the public to whom the speech would be directed: "The right of citizens to inquire, to hear, to speak, and to use information to reach consensus is a precondition to enlightened self-government and a necessary means to protect it."" 11
- In *Sindicato Puertorriqueño v. Fortuño*, 2012 and 124 similar citations
- "...sufficient governmental interest justifies limits on the political speech of nonprofit or for-profit corporations." 11
- In *WESTERN TRADITION v. Attorney General*, 2011 and 195 similar citations
- The Court found that a later case had "little relevance" to the issue at hand because it "involved contribution limits, which, unlike limits on independent expenditures, have been an accepted means to prevent quid pro quo corruption" 11
- In *EX PARTE ELLIS*, 2010 and 69 similar citations
- *Citizens United* invalidated a ban on independent expenditures by corporations and unions for certain types of political speech. 11
- In *US v. Haloran*, 2016 and 64 similar citations
- "...it held that "the Government may not suppress political speech on the basis of the speaker's corporate identity."" 11
- In *IN RE INVESTIGATIVE SUBPOENAS*, 2010 and 187 similar citations
- This Court has reviewed First Amendment challenges to disclosure requirements in the electoral context under an "exactng scrutiny" standard, requiring "a 'substantial relation' between the disclosure requirement and a 'sufficiently important' governmental interest." 11
- In *John Doe No. 1 v. Reed*, 2010 and 57 similar citations
- "...an 'antidistortion' interest—the government's preventing the "corrosive and distorting effects [in elections] of immense aggregations of wealth that are accumulated with the help of the corporate form and that have little or no correlation to the public's support for the corporation's political ideas."" 11
- In *Yamada v. Karamoto*, 2010 and 74 similar citations
- "...laws that burden political speech are subject to strict scrutiny, which requires the Government to prove that the restriction furthers a compelling interest and is narrowly tailored to achieve that interest."" 11
- In *State v. Terrant*, 2012 and 218 similar citations
- "...certainty was not in conflict with Justice Stevens's conclusion on this particular question about foreign influence." 11
- In *Blum v. FEDERAL ELECTION COMMISSION*, 2011 and 33 similar citations
- "...Premised on mistrust of governmental power, the First Amendment stands against attempts to disfavor certain subjects or viewpoints." 11
- In *Jenkins v. KURTNYTH*, 2015 and 125 similar citations

Cited by

- 1A Auto, Inc. v. Director of the Office of Campaign and Political Finance
- Independence Institute v. Williams
- INDEPENDENCE INSTITUTE v. Williams
- Protect My Check, Inc. v. Dilger
- McCutcheon v. Federal Election Commission

Related documents

- Buckley v. Valeo
- McConnell v. Federal Election Commission
- Austin v. Michigan Chamber of Commerce
- Fed. Election Comm’n v. Wisc. Right to Life
- First Natl. Bank of Boston v. Bellotti
- all related documents
Search large text collections to locate documents, short phrases or sentences that directly answer a user’s question.

- **Ross**: based on IBM Watson [rossintelligence.com](http://rossintelligence.com)
  - Developed by U. Toronto law students.
- **Input questions** in plain English, e.g.:
  - In New York, what is secondary liability with respect to copyright infringement and how is it established?
- **Outputs answers**, citations, suggested readings, updates.
  - Cites cases with answers in order of relevance, e.g., Arista Records LLC v. Usenet.com, Inc. (SDNY 2009)
  - Quotes textual answers:
    - “… A party is liable for contributory infringement if, ‘with knowledge of the infringing activity,’ it ‘induces, causes, or materially contributes to the infringing conduct of another.’ Gershwin Pub’l’g....”
Arista Records LLC v. Usenet.com, Inc.

... Contributory copyright infringement is a form of secondary liability with roots in tort-law concepts of enterprise liability and imputed intent." Perfect 10, Inc. v. Visa Int’l Serv. Ass’n, 494 F.3d 788, 794-95 (9th Cir. 2007), cert. denied, ___ U.S. ___, 128 S.Ct. 2871, 171 L.Ed.2d 811 (2008). A party is liable for contributory infringement if, “with knowledge of the infringing activity,” it “induces, causes, or materially contributes to the infringing conduct of another.” Gershwin Publ’g ... read case →

Metro-Goldwyn-Mayer Studios Inc. v. Grokster, Ltd.
Sup. Ct. | June 20, 2005 | 546 U.S. 57, 126 S.Ct. 2764, 162 L.Ed.2d 781

... Despite the currency of these principles of secondary liability, this Court has dealt with secondary copyright infringement in only one recent case, and because MGM has tailored its principal claim to our opinion there, a look at our earlier holding is in order. In Sony Corp. v. Universal City Studios, supra, (125 S.Ct. 2777) this Court addressed a claim that secondary liability for infringement can arise from the very distribution of a commercial product. There, the product, novel at the time ... read case →

Agence France Presse v. Morel

... secondary liability, a copyright holder need not join all infringers as defendants in order for the Court to consider the actions of the non-party infringers in determining where, within the permissible scale, a statutory damages award should fall. See, e.g., Arista Records LLC v. Usenet.com, Inc., No. 07 Civ. 8822 (TIB) (THK), 2010 WI 3629688, at *5 (S.D.N.Y. Feb. 2, 2010) (during an inquest on damages, after granting summary judgment in plaintiffs’ favor on direct and secondary liability, the court ... read case →
1. ML model learns from training set of Q/A pairs how to assess likelihood that short text answers new query.

2. Learns to recognize if version of question is one it can answer.
   - Experts provide natural language [legal] practice questions for which a paragraph is correct answer.
   - System learns weights associated with features of training instances to distinguish pos. / neg. instances of question.
   - Weights inform system’s level of certainty that it “understands” the user’s question.

3. User feedback updates Ross’s confidence in responsiveness of its answer to user’s version of a question.
   - “press thumbs up if the response is accurate” or
   - “press thumbs down for another response.”
3. THREE LIMITATIONS OF LEGAL TEXT ANALYTICS

a) Apps cannot *read* legal texts like lawyers can.
   • Using statistical methods, app can (only) extract some semantic information (i.e., meaning) from legal text.
   • App uses extracted meanings to improve retrieval and ranking.
   • AI & Law cannot yet extract legal rules in logical form from statute texts.

b) ML yields answers but not explanations:
   • ML cannot explain its answers to legal questions or provide arguments.
   • Cannot reason robustly about how different circumstances would affect its answers.

c) Need for manual annotation for supervised learning.
   • There are limits on what can be annotated.
4. CAN CASE-BASED REASONING (CBR) ENABLE LEGAL APPS TO EXPLAIN AND ARGUE?

- Answering legal questions and predicting outcomes should include reasons.
  - Q/A systems do not understand legal reasons, legal reasoning, or argument.
  - Statistical ML learns feature weights, not reasoning.
- Approach: Connecting computational models of case-based legal reasoning to case texts
  - Identify useful types of argument-related information in legal cases.
  - Annotate training sets of case texts with arg-related info.
  - Train ML to identify arg-related information in case texts.
  - Apply extracted arg-related info for intelligent legal IR: argument retrieval (AR).
  - Enable computational models to predict outcomes and construct explanations and arguments.
• Certain argument structures (Mochales and Moens 2011)
  • Argumentative propositions (80% accuracy on ECHR corpus w/max. entropy classifier.)
  • If proposition is premise or conclusion (F1: 68% for premises and 74% for conclusions with SVM).
  • Nested argument structures (60% acc. detecting tree structures; F1: 70% for premises and conclusions with rules.)

• Instances of certain argument schemes (Feng and Hirst, 2011)
  • E.g., argument by example (90.6% acc. using decision trees)

• Sentences that play certain roles in legal arguments (LUIMA: Bansal, et al. 2016)
  • Legal rule sentence (F1 of 66% using log. regression)
  • Evidence-based finding sentence (F1 of 48% using log. regression)

• Legal factors in trade secret law (Falakmasir and Ashley, 2017) (F1 of 65% using word embedding and SVM)
  • Factors: stereotypical patterns of fact that strengthen or weaken a legal claim.
  • A standard representation for case-based argument modeling in AI & Law
From the *Mason* case, a **trade secret dispute** concerning the recipe for a cocktail, Lynchburg Lemonade:

**F6: Security-Measures (pro-plaintiff):** He testified that he told only a few of his employees—the bartenders—the recipe. He stated that each one was specifically instructed not to tell anyone the recipe. To prevent customers from learning the recipe, the beverage was mixed in the “back” of the restaurant and lounge.

**F15: Unique-Product (pro-plaintiff):** It appears that one could not order a Lynchburg Lemonade in any establishment other than that of the plaintiff.

**F16: Info-Reverse-Engineerable (pro-defendant):** At least one witness testified that he could duplicate the recipe after tasting a Lynchburg Lemonade.

**F21: Knew-Info-Confidential (pro-plaintiff):** On cross-examination Randle agreed that he had been under the impression that Mason’s recipe for Lynchburg Lemonade was a secret formula.
HOW COULD LEGAL APPS EXPLAIN AND ARGUE?

• Some AI & Law computational models implement case-based legal argument schemes
  • e.g., CATO, IBP, VJAP, AGATHA:
    • Generate legal predictions and explain them with case-based arguments attorneys can understand.
    • Employ legal factors and relate them to legal values and effects.
  • If legal text analysis programs can learn how to automatically identify factors in case texts, then
    • these computational models could accept case texts as inputs and
    • output predictions and arguments.
• Even without arguments, factors can help to explain likely results.
• Active areas of research in AI&Law
In DYNAMICS, plaintiff’s product information is not sufficiently valuable because the plaintiff has taken such little efforts to maintain the secrecy of the information.

Specifically, regarding the maintenance of secrecy by the plaintiff, the public disclosure amounts to such a clear waiver of property interest, a scenario where usability of public information is critical and such a clear waiver of confidentiality interest regarding the lack of maintenance of secrecy by the plaintiff that the lack of value of the information must be deemed sufficiently established despite the lack of strong evidence for the defendant and the fact that the product information was unique.

A similar inter-issue tradeoff was made in NATIONAL-REJECTORS, which was decided for defendant. There, regarding the maintenance of secrecy by the plaintiff, the disclosure to outsiders amounted to such a clear waiver of property interest and such a clear waiver of confidentiality interest, the public disclosure amounted to such a clear waiver of property interest, a scenario where usability of public information is critical and such a clear waiver of confidentiality interest and the absence of security measures amounted to such a clear waiver of property interest and such a clear waiver of confidentiality interest that the reverse-engineerability qualified as the lack of value of the information despite the fact that the product information had been unique.
USING FACTORS TO PREDICT AND ANALYZE LANDLORD-TENANT DECISIONS TO INCREASE ACCESS TO JUSTICE

• JusticeBot Project, CyberJustice Lab, U. Montreal

• Large corpus of case decisions by Régie du logement du Québec.

• Identified 44 factors in disputes where tenant seeks remedy for apartment problems.
  • E.g., existence of bedbugs, high noise levels or problems with insulation
  • Used these factors to tag 149 cases.

• How many factors found in a case correlated with award of rent reduction and amount.

Westermann, et al. ICAIL 2019
PROGRESS ON LEARNING TO IDENTIFY FACTORS AUTOMATICALLY

• Assembled 1600 trade secret cases from CourtListener
• Represented contextual semantic information using Doc2Vec
• Focused on 179 cases in VJAP corpus
• Trained ML model for each factor (Support Vector Machine (SVM) using OneVsRest)
• Predicted factors that apply in each case document.
• Evaluation:
  - 70% percent of documents in VJAP corpus as training set.
  - 30% of documents as hold-out test set.
  - Results: F1 = .69/.65 (arithmetic mean of precision and recall)

(Falakmasir & Ashley, Jurix 2017)
5. NEED FOR MANUAL TEXT ANNOTATION

- Manually annotating cases essential to mine arguments.
  - *Supervised* machine learning programs need sets of training instances manually annotated by humans.
  - *Unsupervised* ML from legal text collections (e.g., clustering) possible but so far not detailed enough.

- Text annotation:
  - Marking-up texts of case decisions (contracts, or statutes) to identify instances of semantic types of information.

- Type system:
  - Ontological hierarchy of concepts / relations so annotation software pipeline can automatically assign semantics to regions of text.
  - E.g., types of argument schemes, sentence roles in legal argument, stereotypical strengths/weaknesses (i.e., factors).
WHO WILL PERFORM LEGAL ANNOTATION?

• Annotation for ML is expensive.

• Recent Ross platform: lawyers upload briefs; platform:
  1. Parses and analyzes brief,
  2. Determines if cited legal authorities are sound, and
  3. Provides feedback on quality of legal analysis.

• In exchange, platform asks users to annotate data.
  • Annotations involve highlighting decision, key facts, etc.

• Access to high quality data sets is key for ML.
  • Expert crowd-sourced data projects could address that need.
  • Win/win situation for Ross and clients.
• **Hypothesis:** Law students could learn valuable lessons from annotating legal texts.
  
  • Annotating task draws students’ attention to key aspects of the reasoning in a legal case:
    • roles that certain sentences play in legal argument,
    • structural features of argumentation, and
    • substantive strengths and weaknesses of a legal argument involving a particular area of law.

• **Law students are inveterate “highlighters”**.
  • As they highlight, students would produce useful data with which
    • machine learning programs can learn to annotate texts automatically.

• **Maintaining interrater agreement is key**:
  • Average Kappa agreement between pairs of labelers
[1][22 P.2d 884]; Ruffin v. Coca Cola Bottling Co., 311 Mass. 514 [42 N.E.2d 259]; Slack v. Premier-Pabst Corporation, 40 Del. 97 [5 A.2d 516]; Wheeler v. Laurel Bottling Works, 111 Miss. 442 [71 So. 743, L.R.A. 1916E 1074]; Seven-Up Bottling Co. v. Grestes, Va., [27 S.E.2d 92]; Dail v. Taylor, 151 N.C. 284 [95 S.E. 135, 28 L.R.A.N.S. 948]. It would serve no useful purpose to discuss the reasoning of the foregoing cases in detail, since the problem is whether under the facts shown in the instant case the conditions warranting application of the doctrine have been satisfied.

[1] Res ipsa loquitur does not apply unless (1) defendant had exclusive control of the thing causing the injury and (2) the accident is of such a nature that it ordinarily 458,459 could not occur in the absence of negligence by the defendant. (Honea v. City Dairy, Inc., 22 Cal.2d 614, 616-617 [140 P.2d 369], and authorities there cited; cf. Hind v. Wheadoson, 19 Cal.2d 458, 461 [121 P.2d 724].) Prosser on Torts (1941), 293-301.)

[2] Many authorities state that the happening of the accident does not speak for itself where it took place some time after defendant had relinquished control of the instrumentality causing the injury. Under the more logical view, however, the doctrine may be applied upon the theory that defendant had control at the time of the alleged negligent act, although not at the time of the accident, provided plaintiff first proves that the condition of the instrumentality had not been changed after it left the defendant’s possession. (See cases collected in Honea v. City Dairy, Inc., 22 Cal.2d 614, 617-618 [140 P.2d 369].) As said in Dunn v. Hoffman Beverage Co., 126 N.J.L. 556 [20 A.2d 352, 354], ‘defendant is not charged with the duty of showing affirmatively that something happened to the bottle after it left its control or management;... to get to the jury the plaintiff must show that there was due care during that period.’ Plaintiff must also prove that she handled the bottle carefully. The reason for this prerequisite is set forth in Prosser on Torts, supra, at page 300, where the author states: ‘Allied to the condition of exclusive control in the defendant is that of absence of any action on the part of the plaintiff contributing to the accident. Its purpose, of course, is to eliminate the possibility that it was the plaintiff who was responsible. If the boiler of a locomotive explodes while the plaintiff engineer is operating it, the inference of his own negligence is at least as great as that of the defendant, and res ipsa loquitur will not apply until he has accounted for his own conduct.’ (See, also, Olson v. Whitthorne & Swan, 203 Cal. 206, 206-209 [263 P. 518, 58 A.L.R. 129].) [4] It is not necessary, of course, that plaintiff eliminate every remote possibility of injury to the bottle after defendant lost control, and the requirement is satisfied if there is evidence permitting a reasonable inference that it was not accessible to extraneous harmful forces and that it was carefully handled by plaintiff or any third person who may have moved or touched it. (Cf. Prosser, supra, p. 300.) If such evidence is presented, the question becomes one for the trier of fact (see, e.g., 459,459 MacPherson v. Canada Dry Ginger Ale, Inc., 129 N.J.L. 365 [29 A.2d 868, 869], and, accordingly, the issue should be submitted to the jury under proper instructions.

In the present case no instructions were requested or given on this phase of the case, although general instructions upon res ipsa loquitur were given. Defendant, however, has made no claim of error with reference thereto on this appeal. [5] Upon an examination of the record, the evidence appears sufficient to support a reasonable inference that the bottle here involved was not damaged by any extraneous force after delivery to the restaurant by defendant. It follows, therefore, that the bottle was in some manner defective at the time defendant relinquished control, because sound and properly prepared bottles of carbonated liquids do not ordinarily explode when carefully handled.

[6] The next question, then, is whether plaintiff may rely upon the doctrine of res ipsa loquitur to supply an inference that
LIMITS ON SEMANTIC ANNOTATION

• Information has to be expressed fairly directly in the texts for:
  • humans to able to annotate it reliably,
  • pipeline techniques to annotate automatically.

• **Annotation techniques ineffective if:**
  • info must be inferred indirectly or from multiple passages scattered across the text.
  • patterns in legal texts are too:
    • Fine-grained,
    • general,
    • rare, or
    • complex
    • for text analytic techniques to identify them well enough for a CMLA to apply

• features are engineered
  • e.g., based on behavioral trends in decisions of individual justices, the Court, and lower courts.
6. CONCLUSIONS

1. Legal text analytics is a burgeoning new arena in AI and Law:
   • Employ ML, legal network diagrams, Q/A.
   • Cannot read but can extract some useful semantic information from legal texts.
   • Cannot explain answers or make legal arguments.

2. Goal: to connect computational models of case-based legal reasoning to enable explanation and argument:
   • Extracting argument-related information from legal case texts such as factors
   • Applying extracted legal semantic info for:
     • Intelligent legal IR, i.e., argument retrieval
     • Aid in predicting outcomes, explaining predictions, and making arguments.

3. Law students can annotate case texts for machine learning and legal pedagogy.
The field of artificial intelligence (AI) and the law is on the cusp of a revolution that began with text analytic programs like IBM's Watson and Debater and the open-source information management architectures on which they are based. Today, new legal applications are beginning to appear and this book - designed to explain computational processes to non-programmers - describes how they will change the practice of law, specifically by connecting computational models of legal reasoning directly with legal text, generating arguments for and against particular outcomes, predicting outcomes and explaining these predictions with reasons that legal professionals will be able to evaluate for themselves. These legal applications will support conceptual legal information retrieval and allow cognitive computing, enabling a collaboration between humans and computers in which each does what it can do best. Anyone interested in how AI is changing the practice of law should read this illuminating work.
REFERENCES


Casetext CARA: https://casetext.com/


Fastcase: https://www.fastcase.com/


Grabmair, M. Predicting trade secret case outcomes using argument schemes and learned quantitative value effect tradeoffs, Proc. 16th International Conference on Artificial Intelligence and Law (ICAIL 2017), 89-98, ACM, 2017.


Kira: https://kirasystems.com/


Ravel: http://ravellaw.com/

Ravn: https://imanage.com/product/artificial-intelligence/

Ross: http://www.rossintelligence.com

U. Pitt School of Public Health LENA Project Legal Network Analyzer (LENA) (http://www.phdl.pitt.edu/LENA/) and Emergency Law Database (ELDB) (http://www.phasys.pitt.edu/default.aspx)

PRIORITIES FOR LAW FIRM’S AI STRATEGY

a) Establish a firm-wide AI committee
   • Include attorneys \textit{and} computer technical staff
   • Survey firm’s use of AI (e.g., predictive coding in e-Discovery)

b) Develop culture of legal process engineering
   • Institutionalize habits of managing firms’ data sources
     • (e.g., briefs, memoranda, employment data, time on task information) and of
     • identifying how text analytic techniques could add value.

c) Support technology education and research
   • Connect with CS university researchers in ML and AI&Law.
   • Support law faculty in teaching students and new associates about:
     • advances / limitations of AI technology and
     • how to evaluate legal text analytic tools.
d) Establish sourcing strategy:

   • Employing external vendor:
     • Does firm have access to code? Control of data? Need to depend on vendor’s expertise and rely on vendor’s representations?
   • Developing in house:
     • Easier to arrange required maintenance and continual adaptation of tools. Firm may license software. Do AI text analytic experts fit into firm’s employment structure?

d) Establish criteria for choosing / purchasing AI tools.

   • Understand assumptions on which the AI/ML analysis is based and data uses.
   • Inspect models for features that impact predictions.
   • Try out programs on known data and results.
   • Understand how system is evaluated and what evaluations signify.
   • Error analysis may indicate systematic miscalculations / causes.
• “LawGeex, an Israel-based contract analytics company, recently published a study that dramatically demonstrated how its product was more accurate than lawyers at performing contract-assessment work on a sample set of non-disclosure agreements (“NDAs”), with an average accuracy of 94% versus an average accuracy rate of 85% for the lawyers. Even more impressive was how much faster the LawGeex system was than the lawyers; while lawyers took an average of ninety-two minutes to review five NDAs, the AI system needed only twenty-six seconds.”* p. 282


• Task: identify “issues” in five nondisclosure agreements (NDAs).

• Issues \equiv thirty types of provisions in NDA contracts.

• e.g., ‘Definition of Protected Information’ is the ‘issue’ of the provision in NDA that defines:

  • “The information that is protected including the form of the information and the way it is disclosed.”

Nota bene:

• Issue identification does not mean reading a textually described problem or scenario and identifying and analyzing the NDA-related legal issues it presents.
• Not too surprising that program outperforms humans at identifying instances of 30 types of provisions.

• But ask program if an NDA covers a particular item or type of information or raises issues in a particular scenario...
  • Human attorney will probably do better.

• Program will know
  • *that* a provision is NDA’s Definition of Protected Information
  • *but not* what the text of the provision *means*
  • *or how* it relates to the various legal issues it may raise.