NYU Center for Data Science

The topics bring together computer science, economics, and law – and lie at the intersection of **machine learning & natural language understanding** and **behavioral economics & causal inference**. We have curated a number of legal datasets that comprise in total 12 terabytes. Cloud computing services and dropbox teams access provided. In 2016 and 2017, seven projects have resulted in peer review publications and two NIPS workshop selections. The data are extremely expensive in terms of money, time, effort, and resources. More than 10 years have been invested in these datasets. The data is strictly confidential. Anyone who wants to access needs to sign a non-disclosure agreement in terms of safeguarding the confidentiality of the data. A portion of one dataset forbids direct contact. Best practices in reproducibility are required (i.e., to submit to a conference, each group must have a codebase that is reproduceable and then must replicate another group's project).

- Three broad themes are
  - **Challenge**: obtain competitive results on classic prediction problems through methodological innovation
  - **Fairness**: ethics and machine learning
  - **Reasoning**: could AI systems learn philosophical reasoning from a diverse corpus of human text and human dialogue?

Annotated datasets include universe of US Supreme, Circuit, and District Court cases, text-aligned Supreme Court oral arguments, asylum decisions, linked data from arrest to final sentence from a federal prosecutor's office, World War I British military tribunals, and US Medicare. These and other datasets are described at the end. Discussions can lead to deviations from the proposed topics.

1 Predicting Judicial Decisions

1. Challenge: **U.S. Supreme Court Prediction** is widely considered a challenge for machine learning in law. Several attempts have been made, including by previous year's students, but the success is scarcely better than guessing the most frequent class, which is reverse. The challenge offers hundreds of years of training data and is obviously resistant to simple rule-based systems. Social science has a poor understanding when sets of features provide the greatest lift, be they political (ideology), legal (precedent), economic (considerations), psychological (biases), or sociological (group interactions). (01)

2. Challenge: **Sentencing Predictability** has motivated the US Sentencing Commission since the 1980s, and predicting punitiveness has motivated criminal justice and the social scientific study

of law since their originations. This challenge offers millions of decisions in training data. A particularly interesting dimension is see if district judge corpora of opinions in non-crime domains are reflected in their sentencing decisions. Interpretable machine learning may be feasible from the random assignment of cases to judges. Part of the challenge is to set it up in a proper sequential way, to simulate building a model that can then be used for prediction of future trials. The rest of the challenge is to find all the things (via feature generation) that could be predictive (e.g. judge ID, most recent decisions by the same judge, recent decisions in the courthouse, characteristics of the individuals the judges recently saw, etc.). What is also interesting from a social and legal perspective is whether extraneous factors like (i) characteristics of the individuals the judges recently saw, (ii) sporting events, and (iii) weather ends up predicting decisions and if so, what factors are relevant and by how much, and whether these factors affect disparities. (10)

3. Challenge: **Predict Malpractice** based on micro-level data in pharmaceutical company payments, prescribing, and patient outcomes. The challenge is policy-motivated and is made feasible through special access to the US Medicare database. Previous work lack micro-level data. A usage scenario is to consider how malpractice risk changes with medical liability law. A parallel prediction useful for end-to-end orthogonalized machine learning analysis of the causal impacts of payments is to construct a predictive model of payments using micro-level data on prescriptions and patients. Does this predictive model change before and after mandatory disclosure for doctors payments that are observed due to litigation settlement (the effects of disclosure laws on what is being disclosed are typically unknown since data on disclosed activity rarely exist in the absence of disclosure laws)? (13)

2 Reconstructing Judicial Decisions

1. Reasoning: Characterizing cases as **Liberal or Conservative** is an effective goal for the social sciences, both as an intermediate construct for end-to-end empirical analyses and for AI to understand human values. Powerful AI systems will likely need to be able to reason about moral and philosophical problems and ethical theories. Much like a recurrent neural network that aggregates information from lower chunks, the training data includes a 5% sample with 400 hand-coded features and another set of thousands of cases hand-coded for meaning (pro-plaintiff or pro-defendant, pro-business or pro-environment, pro-criminal defendant rights or pro-prosecutor, etc., in 25+ politically salient legal areas). In this challenge, one might predict the 400 labels in the larger dataset using the text or automate the coding of case category (discrimination, antitrust, environment, crime, etc - a 3-digit category is hand-labeled in the 5% sample). Critically, for end-to-end use and for AI understanding of moral reasoning, the judge identity should be excluded from the feature set. (08)

2. Reasoning: **The Fact-Value Distinction** is widely considered the source of conflict between science and ethics, the distinction between what can be known to be true and the personal preferences of individuals, and the thin line between what is truth and what is right. The ability to make the distinction by AI is a critical first step. The court opinions have been previously annotated to distinguish between the facts and legal reasoning of a case. To illustrate the potential use of AI prediction, the product can be employed in testing for motivated reasoning. Social scientists have long speculated (and most recently articulated by Jonathan Haidt) that humans generate the reasoning for a moral reaction rather than judiciously reason towards a decision. After predicting where the fact patterns are located in the case being appealed and in the appellate decision, do differences in fact description explain the final decision? Judge Brandeis famously made up the facts of an accident case so he could make better law–his decision sparked modern tort law. Citation patterns distinguish objective (gradual, cumulative) and subjective (paradigm shifts) knowledge in the sciences and humanities; do they also reveal differences in the law (e.g., corporate law vs. civil liberties)? (03)

3. Reasoning: Universal grammar is the theory of an innate component of the language faculty, independent of sensory experience. One contested research program is universal moral grammar. **The Grammar of Law Project** exploits parallel multilingual legal databases to identify equivalent legal phrases in one prong (to identify the 'molecules' of law), linked-appealed-appellee case text to automate the detection of legal inconsistencies in another prong, and linked-citation flavor and case text to automate encoding of moral views. In a practical sense, this project has a usage scenario to predict the higher court decisions from the lower court opinion and, concomittantly, to help judges write better opinions to avoid reversal. Is there an optimal deviation in legal consistency? How much legal innovation engenders subsequent (positive) citations? Is there a trade-off between innovation and reversal likelihood? Does the legal text or the structure of its citations predict subsequent treatment in terms of importance (citations), controversy (dissent), and mistake (reversal)? (09)

4. Reasoning: For powerful AI systems to help in domains where human values are ambiguous or underdetermined, it is valuable to formalize what is ambiguous or uncertain. **Legal Ambiguity** provide a context to study the use of certainty–a more certain opinion states a clearer policy position, which makes it more attractive to cite by future judges but also more likely to be reversed–but the challenge offers very little training data. A starting point is the use of certainty words. ("High-certainty" Supreme Court opinions have been found to be cited more often.) Another is to study the polarizing effect of elections on judges as an instrument to identify when individuals perceive their views as more certain. To do so, one must isolate the self-certainty effects of electoral polarization rather than, say, the content of campaign advertisements. Other usage scenarios include studying

what policy areas have relatively more ambiguous statutory terms over time. For example, consensus issues like highway maintenance might converge on low-ambiguity clauses, while politicized/divisive issues like guns and abortion might sustain high levels of clause ambiguity.

5. Reasoning: Since the Dawkin's The Selfish Gene, it has been proposed that ideas are memes that can be analogized to genes as units of replication and propagation. At the same time, social constructionists have proposed **The Geneology of Ideology** as a leading argument against a universal moral grammar. Random assignment of judges in US Circuit and District Courts allows one to make causal inference. Computer scientists, statisticians, and genetists have offered methods to detect propagation (peer effects) when the underlying network is unobserved. Validate these methods against heuristic approaches (formulaic meme index) and causal inference. Two networks are observed–the citation network and the (randomly assigned) seating network. One way to frame the problem is predict the memetic *phrases/citations* that are likely to be passed along the network in forward citation but not appear in a distant case in the citation graph. Similarly, do the same with the seating network. Identifying memes in different legal topics can be validated against the humanistic literature of legal scholarship. Identifying thought leaders can be validated against casebooks and variation in exposure due to random assignment. Scoring judges prior to their appointment based on their lower court writings is another usage scenario. Another is to predict early death or dementia, or simply retirement, using judicial corpora. Using only the history of who the judges previously sat with on panels and how the judges' votes aligned with the panelists, predict how the judge votes on the next panel.

6. Reasoning: **Utilitarian vs. Deontological** modes of reasoning is a classic divide in moral philosophy. Economics and law, too, is divided between the utilitarian view that optimal policy should be based on calculations of costs and benefits and a non-consequentialist view that policy should be determined deontologically: from duties judges derive what is the correct law–what is right and just. To begin understanding human values, AI systems will likely need to be able to detect and annotate when an argument is utilitarian or deontological. One approach might be to classify assignments of obligations and authority: e.g., "the right is vested in party 1", "party 1 has the right", "the duty is assigned to party 1", "party 1 must", etc., or to classify conditional language: "if A, then B" constructions. Another approach might be to utilize data on the timing of judge's attendance in a controversial economics training program. A third could be to observe whether economics-trained judges use different features of defendants consistent with cost-benefit analysis when obtaining sentences. Consider building a classifier for political speech seeking to expand/restrict policy and use this on the congressional record floor speech and court opinions. Consider building a classifier for analogical or logical modes of reasoning and use this on several hundred years of Confuscian examinations.

7. Fairness: Recent advances in natural language understanding suggest AI's ability for **Measuring Implicit Bias From Semantics** derived automatically from corpora. A challenge is to identify what sample size is sufficiently large to classify individual implicit bias. The goal of this project is to conduct an in-depth investigation on the properties of measuring judicial implicit bias. How robust are the measurements? Can they be validated against actual judicial decisions (e.g., in discrimination or market competition cases)? Can judicial pre-appointment text be used to predict their future writings and decisions? What is the causal impact of judicial implicit bias on society (e.g., word embeddings quantify gender and ethnic stereotypes, but what is the direction of causality?)? What is the causal impact of randomly assigning a woman (black) judge to a panel on the verdict's gender (race) implicit bias? When the opinion is authored by the woman (black) or by a man (white)? Are implicit attitudes of district judge corpora in non-crime domains reflected in their sentencing decisions? (04)

8. Fairness: Do judge's vocal intonations reflect implicit bias? The goal of this project is also to assess the robustness of **Measurements of Vocal Implicit Bias**. Using 15 years of Supreme Court oral arguments, it appears that vocal intonation of gendered words (e.g., actor vs. actress) classify vocal intonations of neutral words into stereotypically male (e.g., logical, ability, think) and female (e.g., looking, cook, goodwill). Documenting this in a rarefied setting like the Supreme Court is surprising, and suggests the relevance of people's perceptions of gender being revealed in how people speak. Furthermore, the vocal intonations of judges' speaking these words are predictive of their decisions. These results complement other work indicating that perceptions of gender improve predictions of Supreme Court outcomes and continue to play a role in a manner more complex and nuanced than conventionally perceived. Do vocal intonations of judges' speaking these words explain the gender gap in lawyer win rates, or why Democrat judges vote against masculine-sounding males? Other validation possibilities are audio data from police stops of motorists. (02)

9. Reasoning: **Judge Embeddings** as inputs for end-to-end machine learning estimates of the causal impacts of law is a widely sought-after goal, whether to simply score judges on an ideological multi-dimensional spectrum or to use in high-dimensional instrumental variables impact analysis of court decisions. Related goal is to analyze the legal text as a high-dimensional treatment for causal analysis and to see what parts of the opinion impact outcomes. The project builds "belief vectors" for judges based on the opinions they have written. To evaluate the model, the belief vectors can be used as input for a variety of different supervised and unsupervised machine learning tasks to validate whether the information recovered by our model provides a meaningful signal about a judge's judicial beliefs. Can we score judges on statutory interpretation, like textualism or originalism, critical legal studies, etc.? (05)

10. Reasoning: Whether AI systems can learn philosophical reasoning from a diverse corpus of human

text and human dialogue includes querying whether the manner of speech matters. A starting point is to **Predict Ideology from Audio** holding fixed the words spoken and a speaker's demographic characteristics. Several studies have documented that short-term convergence in speech (and audio) is predictive of votes and that judicial ideology displays longitudinal movements over time - is this reflected in their audio beyond the text? Speech variation beyond word choice, that is, fluctuations in the way one speaks holding the words fixed of lawyers appears to double the prediction accuracy of their political ideology relative to the text alone. Examine whether the model fits political ideology of televised political speech, sermons, Buddhist talks, etc. Another starting point is to expand to the universe of data analyses of whether judges' questions, interruptions, or how lawyers' respond to questions are predictive of outcomes. (11)

3 Predictability of Judicial Decisions

1. Fairness: Why do individuals want laws to be predictable but eschew replacing judicial decision-makers by algorithms? We distinguish between **Early Predictability** (using features available to a judge prior to the opening of a case) and predictability (using features available to a statistician over the course of a case but not observing the decision itself). Early predictability suggests ignoring information about the case. The project has a behavioral component of analyzing information acquisition of judges (for example, using textual data from Wikileaks). Refugees seeking asylum are assigned to a randomly chosen judge. The refugee is either granted asylum or not. For each refugee, we have a record of the judge they were assigned to and the timestamp of the trial, as well as information about the refugee (e.g. nationality, defensive/affirmative, lawyer). The obvious prediction problem here is whether or not the refugee is granted asylum. Part of the challenge of this problem is to set it up in a proper sequential way, to simulate building a model that can then be used for prediction of future trials. The rest of the challenge is to find all the things (via feature generation) that could be predictive (e.g. judge ID, presumed skintone of refugee based on nationality, time-ofday, most recent decisions by the same judge, recent decisions in the courthouse, recent successes of the lawyer, characteristics of the individuals the judges recently saw, etc.), and whether these factors affect disparities. (12)

2. Fairness: A perennial question is how much law clerks affect judicial opinions. Several prominent studies suggest traces of the writing styles of clerks are reflected in the final opinions. This project examines **Legal Schools of Thought** using the law school training of clerks. Several episodes of unraveling in the judicial clerkship market (where judges had to make hires with less information about clerks) allow studying the efficiency and equity consequences of the current clerkship market system. A parallel question is whether judges' financial conflicts of interest is an important, predictive feature of how judges' vote, write, and cite legal precedent.

3. Fairness: **What Do Prosecutors Maximize**? Do they seek to minimize recidivism, maximize conviction rates, maximize sentence lengths, or minimize time until trial completion? How do prosecutors compare to a predicted prosecutor? Part of the challenge of this problem is to set it up in a proper sequential way, to simulate building a model that can then be used for prediction of future cases. The rest of the challenge is to find all the things (via feature generation) that could be predictive (e.g. judge ID, most recent decisions by the same judge, recent decisions in the courthouse, characteristics of the individuals the judges recently saw, characteristics of the screener, what the screener recently saw, characteristics of the prosecutor, etc.). What is interesting from a social and legal perspective is whether extraneous factors like (i) the screener identity, (ii) prosecutor identity, and (iii) judge ends up predicting decisions and predicts inequality of treatment of individuals by race and gender, and if so, what factors are relevant and by how much, and whether these factors affect disparities. Developed further, the project could have several important policy implications: it might identify defendant characteristics that are particularly 'noisy' to prosecutors; it could suggest ways of alleviating criminal caseloads without increasing crime rates; and it might provide important insights into how a prosecutor's background relates to the quality and nature of their charging decisions. (14)

4. Fairness: Construct topographical map of WW1 loyalty using universe of courts martials and see if it explains deterrent or anti-deterrent effect of death penalty across different military crimes. Investigate the long-run impacts of "missing men" (geocoded casualties) on demographic and socio-economic outcomes in the UK. Analyze judicial response to executions: **Endogenous Justice** to "we made an example". Analyze physical descriptions of soldiers (youth, complexion, looks, occupation), especially outcome heterogeneity, e.g., do young soldiers respond more to executions? Model geographic spillover / news of executions using geo-location of military units.

4 Impacts of Judicial Text

1. Fairness: A platform for analyzing the **Causal Impacts of Court Decisions** has been developed that employs a machine learning first stage using features that are exogenous (randomly assigned) to simulate a retrospective clinical trial. The second step analyzes the correlation between the predicted court decision and society-wide outcomes. A variety of policy questions are available for analysis in tax, crime, immigration, environment, labor, civil rights, and societal attitudes. (06, 07)

2. Fairness: A platform for analyzing the **Stock Market Impacts of Text** (e.g., in Wikileaks, SEC corporate filings, etc.). Predict changes in state GDP, employment, profits, and wages from text of session laws, and use topic models to interpret the prediction. Compare with predicting changes in the same outcomes from the text of legal precedent.

3. Fairness: Extract features of teacher collective bargaining agreements associated with better or worse student test scores. Predict test scores and other outcomes out of sample. Analyze differences in contracts across areas with different socioeconomic status and school density.

5 Datasets

1. U.S. Circuit Courts

   (a) Digitized universe since 1891 of all 380,000 cases, 1 million judge votes, across 94 hand-labeled legal topics, engineered into 2 billion N-grams of length eight, and 5 million citation edges across cases.

   (b) This is merged with the 268 judges who served during this time period, 250 biographical features, a 5% random sample with 400 hand-labeled features (330-paged codebook), and 6000 cases handcoded for meaning in 25 polarized legal areas.

   (c) This is also merged with administrative data (date of key milestones, e.g., oral arguments, when was the last brief filed, etc.), publicly available U.S. Supreme Court datasets, U.S. District Court datasets, geocoded judge seats, biographies of judicial clerks, and oral arguments' audio files.

   (d) The identities of randomly assigned judges sitting on 3-judge panels (who is authoring the opinions, writing dissents, or writing concurrences) render a random seating network among the judges.

   (e) 25 polarized legal areas have in addition been collected and hand-coded: sexual harassment, eminent domain, free speech, abortion, church-state separation, affirmative action, gay rights, disability rights, campaign finance, capital punishment, criminal appeals, desegregation, sex discrimination, punitive damages, federalism, National Labor Review Board, environmental protection, National Environmental Policy Act, Federal Communications Commission, Title VII, First Amendment, Eleventh Amendment, standing, contracts, and corporate veil piercing.

2. U.S. District Courts

   (a) Digitized universe of millions of criminal sentencing decisions across 94 U.S. District Courts from 1991 (with randomly assigned judges), hand-labeled biographical data of judges, and 83-paged codebook from the U.S. Sentencing Commission.

   (b) Linkages to judge identity were obtained (not publicly available) and hand-labeled biographical data of judges incorporated. Data linkages have been made to daily weather and local sporting events.

   (c) Text of opinions are available since 1923.

3. U.S. Supreme Court

   (a) Digitized speech patterns in oral arguments since 1955– longitudinal data on speech intonation (linguistic turns) are rare.

   (b) Linked to hand-labeled oral advocates' biographies, lawyers' faces, clipped identical introductory sentences*, ratings of their traits, and publicly available U.S. Supreme Court databases containing dozens of additional features and preceding U.S. Circuit Court data.

   (c) *Lawyers always use the exact same sentence when they introduce themselves to the Supreme Court: "Mr. Chief Justice, (and) may it please the Court." We have clipped this data for 1955-2013 comprising over 8000 audio recordings, spoken by many different lawyers over time. Mechanical turk workers have rated 1999-2013 sample (2000+ recordings) based on whether they sound "confident", "trustworthy", "attractive", "masculine", etc. We also have data on the Mturk workers, the Supreme Court cases, and the Supreme Court oral advocates (including their faces).

   (d) Actual analysis of speech patterns is statistically challenging, since speech is modified dynamically. A common measure for variation in speech patterns considers resonances of vowel sounds. In order to properly measure these, the starting locations of all distinct vowel sounds have been manually flagged. An algorithm then measured vowel resonances and assigned to each vowel sound a multidimensional continuous quantity. Therefore, the size of the oral data set is much larger than the size of the underlying text.

   (e) Text is traditionally treated with discrete models. Speech measurements (for example resonances) by contrast are continuous.

4. U.S. State Supreme Court

   (a) Digitized universe for 1947-1994 (roughly 400,000 cases), identities of judges sitting on the panels, hand-labeled biographies, citation network, and original text.

   (b) Some of these judges run for election.

5. EOIR Immigration Courts

   (a) Digitized universe of administrative data on 1 million refugee asylum and 15 million hearing sessions and their time of day across 50 courthouses and 20 years (with randomly assigned judges), hand-labeled biographical data of judges, and dozens of features on the case and the defendant.

   (b) We know when the asylum case was assigned, whether the hearing was an individual hearing or whether multiple individuals were scheduled in the same session, how many cases were scheduled for sessions during a day for that judge, whether this was an in person hearing or

by audio or video, whether it was a written or oral order, whether there are other related applications for relief filed by the individual and the judge's ruling on each, ethnicity of the applicant, the reason for the case and the judge.

(c) Data linkages have been made to daily weather and local sporting events.

6. Vertical linkages from arrest to final sentence

   (a) Digitized universe of individuals in a federal prosecutor's office over a decade with many stages of random assignment (screener, federal prosecutor, and judge). New Orleans is the largest city and metropolitan area in the state of Louisiana. The Orleans Parish District Attorney's Office and its prosecuting attorneys are responsible for enforcing state criminal laws and local ordinances to protect and serve the citizens of New Orleans and surrounding areas.

   (b) The current data set is from 1988 to 1999 and provides detailed information on approximately 430,000 charges and 280,000 cases (involving 145,000 defendants) filed or adjudicated during this timeframe. The data collected also contains detailed information regarding each individual offender, such as social security number and the corresponding prosecutor and judge.

   (c) Linkages have been made to voting records, bankruptcy, and foreclosure. The dataset is rare: vertical linkages from the time of arrest, including those sent home without a trial, otherwise do not exist. There is a 594-paged codebook.

7. World War I British courts martial

   (a) Digitized World War I British archival datasets, including universe of deserters (including names and often their birthplace) reported in military diaries, police gazettes, and handwritten military trials, commuted and executed capital sentences (which historians believe was random), geocoded casualties, maps, officer lists, and order of battle.

8. US Medicare

   (a) Linked administrative Medicare data to industry-physician relationships cleaned from litigation settlements and through the Affordable Care Act, comprising 30 million payments including the date of payment and affiliated drug code.

   (b) Linked to biographical characteristics of physician and some patient demographics.

9. Other datasets include judges' financial disclosures, many social and economic datasets, and other cleaned legal datasets:

   (a) Corporate Filings: The text of disclosures, contracts, and charters submitted to the SEC by publicly listed companies for 1996-2016.

(b) Federal and State Legislation: We have the full history of federal and state laws enacted up until 2012.

(c) UN Parallel Texts and Hong Kong Laws.

(d) Party Manifestos Corpus: Large corpus of political party manifestos across the world, with rich metadata.

(e) U.S. Congressional Record: All speeches by congressman and senator for 1880-2015, with metadata.

(f) Teacher's union contracts: Large corpus of 6000 teachers contracts for all districts in the state of michigan since 2010.

10. Over 1000 legal databases tagged and linked including all federal (supreme, appellate, district, bankruptcy, tax, patent, trade, customs, claims, unpublished) and state (supreme, appellate, district, tax, chancery, family, labor, unpublished) court cases to the earliest available date (some as early as 1778).

    (a) Types of databases include code, statutes, bills, regulations, bulletins and notices, commission decisions, Attorney General opinions, rulings, statements, opinion letters, bill tracking, workers' compensation decisions, municipal codes, physician discipline decisions, market conduct examinations, issuances, directives, public health reports, FTC, IRS, EEOC, Department of Labor, Department of Defense, EPA, SEC, Federal Reserve, contract appeals decisions, legislative service, manuals, etc.

6 Prior Years' Projects

1. Algorithms as Prosecutors: Lowering Rearrest Rates Without Disparate Impacts and Identifying Defendant Characteristics 'Noisy' to Human Decision-Makers (D. Amaranto, E. Ash, L. Ren, C. Roper)

    (a) NIPS workshop paper in 2017

    (b) Conference on Law and STEM

2. What Matters: Agreement Between U.S. Courts of Appeals Judges (*Journal of Machine Learning Research*; TSE Working Paper No. 16-747; X. Cui, L. Shang, J. Zheng)

    (a) NIPS workshop paper in 2016

    (b) Conference on Empirical Legal Studies (CELS)

3. The Genealogy of Ideology: Identifying Persuasive Memes and Predicting Agreement in U.S. Circuit Courts (*Proceedings of the ACM Conference on AI and the Law*, forthcoming; A. Parthasarathy, S. Verma)

(a) cited in 2017 AEA presidential address by Nobelist Bob Shiller

(b) ACM Conference on AI and the Law

(c) Conference on Empirical Legal Studies (CELS)

4. Early Predictability of Asylum Court Decisions (*Proceedings of the ACM Conference on AI and the Law*, forthcoming; M. Dunn, L. Sagun, H. Sirin)

   (a) ACM Conference on AI and the Law

5. Can Machine Learning Help Predict the Outcome of Asylum Adjudications? (*Proceedings of the ACM Conference on AI and the Law*, forthcoming; J. Eagel)

   (a) ACM Conference on AI and the Law

6. Is Justice Really Blind? And Is It Also Deaf? (in Computational Analysis of Law, Santa Fe Institute Press, forthcoming, ed. M. Livermore and D. Rockmore)

   (a) Conference on Computational Analysis of Law

7. Judge Embeddings: Vector Representations of Legal Belief (in Computational Analysis of Law, Santa Fe Institute Press, forthcoming, ed. M. Livermore and D. Rockmore; E. Ash)

   (a) Conference on Computational Analysis of Law

8. Is Justice Really Blind? And Is It Also Deaf? (in Computational Analysis of Law, Santa Fe Institute Press, forthcoming, ed. M. Livermore and D. Rockmore)

   (a) Conference on Computational Analysis of Law