2023

Emerging Technology’s Language Wars: AI and Criminal Justice

Carla L. Reyes
Southern Methodist University, Dedman School of Law

Author ORCID Identifier:

https://orcid.org/0000-0002-2448-8309

JOURNAL
OF
LAW & INNOVATION

VOL. 5 JANUARY 2023 NO. 1

ARTICLE

Emerging Technology's Language Wars: AI and Criminal Justice

CARLA L. REYES†

Work at the intersection of Artificial Intelligence systems (AI systems) and criminal justice suffers from a distinct linguistic disadvantage. As a highly interdisciplinary area of inquiry, researchers, lawmakers, software developers, engineers, judges, and the public all talk past each other, using the same words but as different terms of art. Evidence of these language wars largely derives from anecdote. To better assess the nature and scope of the problem, this Article uses corpus linguistics to reveal inherent value conflicts embedded in definitional differences and debates. Doing so offers a tool for reconciling specific linguistic ambiguities before they are embedded in law and ensures more effective communication of the technical pre-requisites for AI systems that, by design, seek to achieve their intended purpose while also upholding core democratic values in the criminal justice system.

† Assistant Professor of Law, Southern Methodist University Dedman School of Law; Affiliated Faculty, Indiana University Ostrom Workshop on Cybersecurity and Internet Governance; Research Associate, University College of London Blockchain Research Center. Many thanks to the Max Planck Institute Luxemburg for the opportunity to workshop an early version of the technique used in this article, to Professor David Abrams and the editors of the Journal of Law & Innovation for the invitation to contribute to this stellar volume of research, and to Professor Hannah Bloch-Wehba for extremely helpful feedback. I also offer my deepest thanks to Victoria R. Nelson, Research Assistant extraordinaire.

INTRODUCTION

In 2016, Billy Ray Johnson received a sentence of life imprisonment without parole for crimes—burglaries and sexual assaults—regarding which he claimed actual innocence.[1] Mr. Johnson's conviction rested heavily on the results of DNA testing conducted using the highly experimental algorithm TrueAllele.[2] The court denied Mr. Johnson's lawyers access to the TrueAllele source code on the grounds of trade secret protection,[3] with the result that Mr. Johnson could not adequately challenge the key piece of evidence against him[4]—evidence that by many accounts suffered from technological flaws and a high degree of inaccuracy risk due to its experimental nature.[5] In the wake

[1] Brief for the American Civil Liberties Union et al. as Amici Curiae Supporting Appellant at 11, People v. Johnson, No. F071640, 2019 WL 3025299 (Cal. Ct. App. July 11, 2019) [hereinafter Amicus Brief].

[2] *Id*. at 9–11; *see also* Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 2019 (2017) (noting that TrueAllele is regularly objected to via Frye/Daubert litigation).

[3] Amicus Brief, *supra* note 1, at 20.

[4] *Id*. at 20–26.

[5] *Id*.; *see also* Roth, *supra* note 2, at 2019–20 (discussing a case where TrueAllele returned contradicting results as another test on the same input); Vera Eidelman, *Secret Algorithms are Deciding Criminal Trials and We're Not Even Allowed to Test Their Accuracy*, ACLU: FREE FUTURE (Sept. 15, 2017) (discussing the issue with denying Johnson

of this and other similar cases in which courts denied defendants access to key data and source code essential to mounting an effective legal challenge,[6] commentators have called for greater transparency, explainability, accountability, and fairness related to algorithms and other Artificial Intelligence (AI) systems used in criminal justice administration.[7]

Indeed, a long, distinguished line of literature chronicles government's increasing use of AI systems to make decisions impacting individual rights,[8] including those rights constitutionally protected during an individual's experience in the criminal justice system.[9] Taken together, the research

---

access to the TrueAllele source code), https://www.aclu.org/blog/privacy-technology/surveillance-technologies/secret-algorithms-are-deciding-criminal-trials-and [https://perma.cc/DR8N-8LRQ].

[6] *See, e.g.*, Brief for Petitioner, Robinson v. Commonwealth, No. 25 WDM 2016 (Pa. Super. Ct. Mar. 7, 2016) (requesting review after lower court denied defendant access to the source code invented by the Commonwealth's expert witness); State v. Loomis, 881 N.W.2d 749 (Wis. 2016), *cert. denied*, 137 S. Ct. 2290 (2017) (considering the use of a risk assessment tool at sentencing); *see also* RASHIDA RICHARDSON, JASON M. SCHULTZ & VINCENT M. SOUTHERLAND, AI NOW INSTITUTE, LITIGATING ALGORITHMS 2019 US REPORT: NEW CHALLENGES TO GOVERNMENT USE OF ALGORITHMIC DECISION SYSTEMS 13–15 (2019), https://ainowinstitute.org/litigatingalgorithms-2019-us.html [https://perma.cc/9U8K-JZNR] (discussing a case denying the right to even investigate if an AI system had been used and a case denying access to data used to produce an AI system output).

[7] *See, e.g.*, Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1343–44 (2018) (arguing that trade secret law should not create privilege in criminal proceedings); Sonya K. Katyal, *Private Accountability in the Age of Artificial Intelligence*, 66 UCLA L. REV. 54, 54 (2019) (calling for methods other than public regulation to address transparency and accountability in artificial intelligence); Hannah Bloch-Wehba, *Visible Policing: Technology, Transparency, and Democratic Control*, 109 CAL. L. REV. 917, 923–924 (2021) (discussing transparency and accountability in policing technology) [hereinafter Bloch-Wehba, *Visible Policing*].

[8] *See, e.g.*, Rashida Richardson, *Defining and Demystifying Automated Decision Systems*, 81 MARYLAND L. REV. 785, 792-93 (2022) (cataloguing a variety of government uses); Hannah Bloch-Wehba, *Access to Algorithms*, 88 FORDHAM L. REV. 1265, 1273–1290 (2020) (discussing use of algorithms in Medicaid, education, policing, bail, evidence and sentencing) [hereinafter Bloch-Wehba, *Access to Algorithms*]; Carla L. Reyes & Jeff Ward, *Digging into Algorithms: Legal Ethics and Legal Access*, 21 NEV. L.J. 325, 330–331 (2020) (cataloguing a variety of government uses); Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1252 (2008) (examining the use of algorithms in administrative law); Margaret Hu, *Algorithmic Jim Crow*, 86 FORDHAM L. REV. 633, 671–79 (2017) (examining algorithms in the national security and immigration contexts).

[9] *See, e.g.*, Mirko Bagaric & Gabrielle Wolf, *Sentencing by Computer: Enhancing Sentencing Transparency and Predictability, and (Possibly) Bridging the Gap Between Sentencing Knowledge and Practice*, 25 GEO. MASON L. REV. 653, 654 (2018) (algorithmic sentencing); Wexler, *supra* note 7, at 1343 (2018) (AI systems-based evidence), Ashley Deeks, *Predicting Enemies*, 104 VA. L. REV. 1529, 1538–47 (2018) (policing) [hereinafter Deeks, *Predicting Enemies*]; Ignacio Cofone, *Algorithmic Discrimination Is an Information Problem*, 70 HASTINGS L.J. 1389, 1396–97 (2019) (parole); Meghan J. Ryan, *Escaping the*

demonstrates that although the use of AI systems may promote efficiency and potential cost savings, it is far from clear that the use of AI systems in the criminal justice process can uphold related core democratic values.[10] As a result, many scholars call for increased transparency, accountability, explainability, and fairness in the data relied upon by AI systems, the computations that form the backbone of AI systems, and the type of output generated by AI systems.[11]

In response to these calls for changes in AI systems, a variety of scholars have devoted substantial effort to untangling the legal conceptions of transparency, accountability, explanation, and fairness from the realistic present technological capacity of AI systems.[12] Some scholars point out that

---

Fingerprint Crisis: A Blueprint for Essential Research, 2020 U. ILL. L. REV. 763, 764–65 (2020) (fingerprinting); Elizabeth E. Joh, *Discretionless Policing: Technology and the Fourth Amendment*, 95 CAL. L. REV. 199 (2007) (traffic stops); Elizabeth E. Joh, *Policing Police Robots*, 64 UCLA L. REV. DISC. 516 (2016) (policing).

[10] *See* Katyal, *supra* note 7, at 105–06 ("While automation dramatically lowers the cost of decision-making, it also raises significant due process concerns, involving lack of notice and the opportunity to challenge the decision.").

[11] *See, e.g.*, Hannah Bloch-Wehba, *Transparency's AI Problem*, DATA & DEMOCRACY 18 (Tex. A&M Univ. Sch. of L., Research Paper No. 21–13, 2021) (calling for government contracting entities to take into account transparency obligations and to encourage vendors to commit to open standards); Bloch-Wehba, *Access to Algorithms*, *supra* note 8, at 1271 (arguing that disclosure through the Freedom of Information Act, its state equivalents, and the First Amendment can be used to enhance transparency and accountability in government use of AI systems); Richard M. Re & Alicia Solow-Niederman, *Developing Artificially Intelligent Justice*, 22 STAN. TECH. L. REV. 242, 285 (2019) (listing ways to entrench equitable justice values in artificial intelligence, such as by formally establishing requirements); Katyal, *supra* note 7, at 107–17 (discussing options for transparency through self-regulation); Chelsea Barabas, *Beyond Bias: Re-imagining the Terms of "Ethical AI" in Criminal Law*, 12 GEO J.L. & MOD. CRIT. RACE PERSP. 83, 84 (2020) (noting the attempts of researchers to use fairness criteria and managerial best practices to combat the biased inaccuracies in algorithmic tools used in criminal justice); Ashley Deeks, *The Judicial Demand for Explainable Artificial Intelligence*, 119 COLUM. L. REV. 1829, 1831 (2019) (discussing judicial creation of "common law of AI" as an approach to creating explainability standards for algorithms).

[12] *See, e.g.*, Solon Barocas, Sophie Hood & Malte Ziewitz, Governing Algorithms: A Provocation Piece 6–7 (Mar. 29, 2013) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2245322 [https://perma.cc/NM7V-YU4L] (posing a series of research questions related to assumptions and misunderstandings around algorithmic opacity, inscrutability, bias, and fairness); Joshua A. Kroll, Joanna Huey, Solon Barocas, Edward W. Felten, Joel R. Reidenberg, David G. Robinson & Harlan Yu, *Accountable Algorithms*, 165 U. PA. L. REV. 633, 634 (2017) (discussing technological tools for ensuring accountability in automated decision-making systems reliant on algorithms); Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085, 1089–96 (2018) (discussing inscrutability and non-intuitiveness as challenges to explaining algorithmic decision-making processes); Shira Mitchell, Eric Potash, Solon Barocas, Alexander D'Amour & Kristian Lum, *Algorithmic Fairness:*

discussions at the intersection of law and AI systems suffer from a linguistic difficulty rooted in definitional differences across disciplines.[13] In other instances, specific language is value laden, and the values attributed to the language vary, and sometimes conflict, in two or more disciplines.[14] Ultimately, it seems that the deepest felt incongruities between law and AI systems often relate to words or phrases to which two or more groups attribute *both* different definitions *and* different values.

Increasingly, the intersection of AI and criminal justice represents one area in which dueling definitions and competing values play an important role in the debate as to the appropriate use of AI systems by government.[15] Presently, the use cases for AI systems in criminal justice span from the relatively routine and mundane, to constitutionally imperative. AI systems are used to send automated hearing reminders,[16] assist in defending against

---

Choices, Assumptions and Definitions, 8 ANN. REV. STATS. & ITS APPLICATION 141, 142 (2021) (summarizing the current definitions and methods of evaluating fairness); Lilian Edwards & Michael Veale, *Slave to the Algorithm? Why a Right to an Explanation is Probably Not the Remedy you are Looking For*, 16 DUKE L. & TECH. REV. 18 (2017) (explaining why a "right to an explanation" in the GDPR may be an inadequate form of transparency for algorithmic decision-making).

[13] *See, e.g.*, Harry Surden, *Artificial Intelligence and Law: An Overview*, 35 GA. ST. U. L. REV. 1305, 1308 (2019) ("When many people hear the term 'AI' they imagine current AI systems as thinking machines. A common misperception along this line is that existing AI systems are producing their results by engaging in some sort of synthetic computer cognition that matches or surpasses human-level thinking. The reality is that today's AI systems are decidedly not intelligent thinking machines in any meaningful sense."); Andrew D. Selbst, Danah Boyd, Sorelle Friedler, Suresh Venkatasubramanian & Janet Vertesi, *Fairness and Abstraction in Sociotechnical Systems*, *in* FAT* '19: PROCEEDINGS OF THE CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 61–62 (2019) (examining how technical computer science understandings of fairness can misalign with the fairness needs of the social context in which the AI system is deployed); Andrew D. Selbst, *A Mild Defense of Our New Machine Overlords*, 70 VANDERBILT L. REV. EN BANC 87, 90–91 (2017) (critiquing a law review article because its premise rests on a misunderstanding of the capacities and limits of AI system technology).

[14] *See, e.g.*, Andrew D. Selbst & Julia Powles, *Meaningful Information and the Right to Explanation*, 7 INT'L DATA PRIVACY LAW 233, 236 (2017) (explaining that law and society attribute at least two different values to the right to explanation); Reyes & Ward, *supra* note 8, at 356–57 (describing a system designed to identify gaps between context ideals and current circumstances).

[15] *See, e.g.*, Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 GA. L. REV. 109, 123–26 (2017) (discussing issues in defining "discrimination," "fairness," and "bias"); K. Brennan-Marquez, *Plausible Cause: Explanatory Standards in the Age of Powerful Machines,* 70 VANDERBILT L. REV. 1249, 1258–60 (2017) (contrasting the terms "probable" and "plausible"); Bloch-Wehba, *Visible Policing*, *supra* note 7, at 923–24 (examining critiques of "transparency" as an effort to weaken agencies and programs).

[16] MARCY R. PODKOPACZ, USING REMINDERS TO REDUCE FAILURE TO APPEAR IN COURT, FOURTH JUD. DIST. OF MINN. 7–9 (2019), https://www.mncourts.gov/mncourtsgov/media/fourth_district/documents/Research/Henne

parking tickets,[17] fight financial crime,[18] identify crime hot spots within cities,[19] set policing priorities,[20] set bail,[21] help determine sentencing,[22] provide or analyze key evidence,[23] among other use cases.[24] A growing body of academic literature calls attention to the need to balance technological design, the contracts under which the technology is procured by government bodies, and the legal rules, norms, and values that apply to the technology's use in a given context.[25] One suggested approach to mitigating the potential harms from value clashes between AI system design and the public functions of government has been to create procurement standards that encourage heightened attention to transparency, accountability, fairness, and

---

pin-County-Court-eReminders-Project-September-2019.pdf; Sarah Jane Staats, *New Automated Text Messages Remind Santa Clara County Public Defender Clients of Court Dates to Reduce Pretrial Incarceration*, STANFORD IMPACT LABS (Mar. 23, 2021), https://impact.stanford.edu/news/new-automated-text-messages-remind-santa-clara-county-public-defender-clients-court-dates [https://perma.cc/KR2S-MJGU].

[17] Erwin van den Burg, *Chatbot Delivers Justice, One Parking Ticket at Any Given Time*, SCIENCE BRIEFS (Nov. 20, 2021), https://sciencebriefss.com/innovation/chatbot-delivers-justice-one-parking-ticket-at-any-given-time/ [https://perma.cc/8TMS-WZ78].

[18] European Banking Federation, *Using Machine Learning and Artificial Intelligence to Strengthen the Fight Against Financial Crime*, LINKEDIN (Feb. 23, 2022), https://www.linkedin.com/pulse/using-machine-learning-artificial-intelligence-/ [https://perma.cc/6WT7-8EVT].

[19] Deeks, *Predicting* Enemies, *supra* note 9, at 1530 ("In the criminal justice context, federal, state and local law enforcement officials increasingly rely on computer algorithms to help them predict how dangerous certain people and certain physical locations are, so as to make more objective judgments about whom to keep in custody and how to use policing resources most efficiently.").

[20] *Id.*

[21] Bloch-Wehba, *Access to Algorithms*, *supra* note 8, at 1284–86.

[22] *Id*. at 1288–90.

[23] *Id*. at 1286–88; *see also* Ryan, *supra* note 9, at 807–09 (discussing the use of algorithms in verifying fingerprints); Sabine Gless, *AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials*, 51 GEO. J. INT'L L. 195, 198 (2020) (introducing the evidentiary issues associated with AI-generated data).

[24] *See, e.g.*, RICHARDSON ET AL., *supra* note 6, at 13–14 (discussing use in gang databases).

[25] *See, e.g.*, Bloch-Wehba, *Transparency's AI Problem*, *supra* note 11, at 18 (stating that contracting entities should consider whether vendors who claim openness will assert trade secret protections or circumvent governmental transparency obligations, and counseling that government contracts should hold these vendors to standards of openness developed publicly by multiple stakeholders); Bloch-Wehba, *Access to Algorithms*, *supra* note 8, at 1271 (discussing the conflict between transparency and trade secrecy in litigation involving algorithms); Re & Solow-Niederman, *supra* note 11, at 254 (arguing that standardized algorithm may undermine discretionary reasoning); Katyal, *supra* note 7, at 107–17 (describing how existing law meant to address algorithmic bias has failed to protect civil rights); Barabas, *supra* note 11, at 84; Deeks, *supra* note 11, at 1831 (explaining that courts developing case law in the area of algorithms are developing the "common law of AI").

explainability by design.[26] Even as this approach gains traction among state legislatures,[27] it remains unclear whether such efforts simply side-step the hard definitional and value conflicts that have come to define the space, or whether they further misunderstandings of how technological design and important legal values fit together. Do the legislative proposals impose technologically achievable requirements for AI systems? Do they adequately balance the competing values at play?

Scholars also seek to clarify technical misunderstandings about AI systems and expose how those misunderstandings might misdirect policy concerns, push law toward poorly understood AI uses, or lead to inadequately designed precautionary measures.[28] Despite these warnings that law and AI might be suffering from a misalignment due to clashes in linguistic meaning, growing anecdotal evidence suggests that the linguistic difficulties so thoroughly warned against in the literature remain ill-understood or ignored by policy and lawmakers.[29] Ultimately, however this sentiment is just an uneasy feeling based on anecdotal evidence. That is, until now.

This Article uses corpus linguistics to demonstrate that conversations at the intersection of AI systems and criminal justice administration suffer from

[26] *See* Bloch-Wehba, *Access to Algorithms*, *supra* note 8, at 1308 ("Acting as consumers, governments can therefore require more demanding contract terms that bring their procurement processes into alignment with due process and transparency requirements.").

[27] *See* S.B. 5116, 67th Leg., Reg. Sess. (Wash. 2021) ("An Act Relating to establishing guidelines for government procurement and use of automated decision systems in order to protect consumers, improve transparency, and create more market predictability . . . ."); H.B.1323, Reg. Sess. (Md. 2021) ("For the purpose of requiring a State unit to purchase a product or service that is or contains an algorithmic decision system that adheres to responsible artificial intelligence standards . . . ."); H.B. 263 (Vt. 2021) ("This bill proposes to require: . . . the Secretary of Digital Services to adopt standards and practices on the development, use, and procurement of automated decision systems by the State . . . .").

[28] *See, e.g.*, Rebecca Crootof & BJ Ard, *Structuring Techlaw*, 34 HARV. J.L. & TECH. 347, 365–66 (2021) (discussing how ambiguity in legal fields, caused by technology, can lead to problematic results); Juliet M. Moringiello & Christopher K. Odinet, *The Property Law of Tokens*, 74 FL. L. REV. 607 (2022) (explaining how misunderstandings around NFTs and tokenization are leading to misdirected policy concerns); Andrew Verstein, *The Misregulation of Person-to-Person Lending*, 45 U.C. DAVIS L. REV. 445, 447–48 (2011) (arguing that misunderstanding P2P lending led the SEC to inappropriately assert its jurisdiction); Devin R. Desai & Joshua A. Kroll, *Trust But Verify: A Guide to Algorithms and the Law*, 31 HARV. J.L. & TECH. 1, 5 (2018) ("Put simply, current calls for algorithmic transparency misunderstand the nature of computer systems. . . . We believe this problem is aggravated because although algorithms are decidedly not mystical things or dark magic, algorithms are not well understood outside the technical community.").

[29] *See, e.g.*, State v. Loomis, 881 N.W.2d 749, 772 (Wis. 2016), *cert. denied*, 137 S. Ct. 2290 (2017) (holding that a court could consider COMPAS risk assessment scores in sentencing); Lola v. Skadden, Arps, Slate, Meagher & Flom LLP, 620 F. App'x 37, 43–45 (2d Cir. 2015) (opining that engaging in a task that could be entirely performed by a machine cannot be said to be engaging in the practice of law).

clashes in language that reveal interdisciplinary disagreements over both definitions and values. In doing so, this Article demonstrates an alternative use for corpus linguistics in legal analysis than that which has been focused on in the relevant literature to date. Namely, this Article does not attempt to identify an "ordinary meaning" of any particular term, nor does it seek to advocate for one meaning of a term over another. Rather, the Article argues that the more basic task of simply identifying the various meanings of interdisciplinary terminology used in discussions about AI systems and criminal justice can help move law reform and policy changes forward. To do so, the Article begins by considering the uneasy relationship between law and other disciplines, such as computer science, when it comes to language. Using AI systems and criminal justice as a case study, the Article briefly reviews some of the debate around accountability, transparency, explainability, and fairness. The Article then presents the results of a corpus linguistic study of these terms—revealing how such contested terminology and its fuzzy meanings impact the broader discourse around AI systems and criminal justice, and, perhaps more problematically, how it increasingly clouds well-meaning legal reform attempts. The Article concludes by considering the implications of law and technology's language wars for the use of AI systems in criminal justice.

I. LAW AND TECHNOLOGY'S LANGUAGE WARS: A COLLISION OF EQUALLY VALID TERMS OF ART

Lawyers are word smiths, and law prominently uses terms of art to create legal ideas, legal doctrine, and advocate for clients.[30] But law is not the only discipline that adopts and uses terms of art. Indeed, recognizing the limits of law to account for, understand, and accommodate terms of art used in technology-related disciplines, law reformers count the principle of technology neutrality among the foundational principles of law-making.[31] Even so, when crafting law applicable to emerging technologies, legal language commonly collides with technical language.[32] Failure to recognize

---

[30] *See* Daniel Martin Katz, Michael James Bommarito, Julie Seaman, Adam Candeub & Eugene Agichtein, *Legal N-Grams? A Simple Approach to Track the 'Evolution' of Legal Language*, 2011 PROC. OF JURIX 1 (citing E. MERTZ, THE LANGUAGE OF LAW SCHOOL (2007)) ("Modern legal systems feature bodies of highly technical vocabulary that lawyers, judges and legal academics use to construct legal ideas and legal doctrines.").

[31] *See* Brad A. Greenberg, *Rethinking Technology Neutrality*, 100 MINN. L. REV. 1495, 1495 (2016) ("Scholars and legislators have overwhelmingly adopted the latter mode—'technology neutrality'—based on the assumption it promotes statutory longevity and equal treatment of old and new technologies.").

[32] *See, e.g.*, Desai & Kroll, *supra* note 28, at 4 (discussing the clash of legal expectations

this collision can undermine the effectiveness of law as it relates to emerging technology. Indeed, recognition of the collision between legal language and technical language, and blind insistence that legal terms of art trump the meanings those words hold in other disciplines may also undermine effective law-making.

Law does not exist without language.[33] Indeed "enacting legislation is generally recognized as an act of communication" by which a "legal 'message' is 'transported' in a one-sided 'flow model' of information, that is, from 'law-giver' to 'law-taker', from sender to receiver."[34] This one-sided flow model assumes, of course, that both the sender and receiver share a common understanding about the content of the message.[35] For example, the ordinary meaning cannon of statutory interpretation rests on this foundational premise of shared linguistic meaning.[36] This form of statutory interpretation requires that the statute be understood by looking "at the statutory structure and hear[ing] the words as they would sound in the mind of a skilled, objectively reasonable user of words."[37] What happens when reasonable minds of "skilled, objectively reasonable user[s] of words" differ? In particular, what happens when different understandings are shared widely by a community of like-minded users of words? Or when the differences are a result of specific training in a particular field or profession?

In such cases, when shared understandings of language do not exist, the legal rules communicated via legislation are likely to be rendered ineffective.[38] The problem, of course, is that meaningful "rule of law requires

---

of transparency and the reality of algorithms); Angela Walch, *The Path of the Blockchain Lexicon (and the Law)*, 36 REV. BANKING & FIN. L. 713, 713–14 (2017) (examining how technical words like "immutability" creep into statutes without actual understanding of what the term means in the technical community).

[33] *See* Deborah Cao, *Legal Speech Acts as Intersubjective Communicative Action*, *in* INTERPRETATION, LAW AND THE CONSTRUCTION OF MEANING 65, 65 (Anne Wagner, Wouter Werner & Deborah Cao eds., 2007) ("Law relies on language and particularly, it relies on the performative nature of language use."); Judith N. Levi, *The Study of Language in the Judicial Process*, *in* 5 LAW, SOCIETY, AND POLICY: LANGUAGE IN THE JUDICIAL PROCESS 4, 4 (Judith N. Levi & Anne Graffam Walker, eds. 1990) ("[L]anguage is the vehicle by means of which law is transmitted.").

[34] Hanneke Van Schooten, *Law as Fact, Law as Fiction: A Tripartite Model of Legal Communication*, *in* INTERPRETATION, LAW AND THE CONSTRUCTION OF MEANING 3, 3–4 (Anne Wagner, Wouter Werner & Deborah Cao eds., 2007).

[35] *Id*. at 7–8.

[36] *See* BJ Ard, *Interpreting by the Book: Legislative Drafting Manuals and Statutory Interpretation*, 120 YALE L.J. 185, 194 (2010) ("Textualism is a theory of statutory interpretation premised on the idea that the courts should act as faithful agents of Congress by interpreting each law to give effect to its ordinary meaning.").

[37] Frank H. Easterbrook, *The Role of Original Intent in Statutory Construction*, 11 HARV. J.L. & PUB. POL'Y 59, 65 (1988).

[38] *See* Van Schooten, *supra* note 34, at 7 (explaining the way that the imaginary reality

that governing rules provide advance notice to enable people to plan their affairs with knowledge of the legal consequences of their actions."[39] As such, when laws use interdisciplinary language that may be intended one way, and received another, the mismatch poses a genuine problem for upholding key rule of law principles.[40] At a minimum, then, those advancing new or changed law in an area of interdisciplinarity must tread particularly carefully.

Indeed, some scholars have observed that law's track record with interdisciplinarity leaves much to be desired in a quest for shared meaning.[41] Some argue that, by and large, when lawyers and legal academics reach out to other disciplines, it is to co-opt lessons, or, in the case of emerging technology, words, and then use them for the lawyer or academic's own purpose.[42] Generally, that purpose is to offer the most persuasive interpretation of the law in order to either further the interest of a client, or for the legal academic, to further the aims of a coherent, rational and just legal order.[43] The concern shared by many scholars working at the intersection of AI systems and criminal justice is that when it comes to law and technology, co-opting technology terms for a lawyer or legal academic's own purpose can have the perverse effect of making the law less coherent and understandable, rather than more.[44] Adopting technology terminology in a legal context to address legal questions[45] or as a solution to a problem at the intersection of law and technology can make communicating shared meaning through law more difficult rather than less.[46]

---

of a rule's message can fail to materialize).

[39] Stefan Th. Gries & Brian G. Slocum, *Ordinary Meaning and Corpus Linguistics*, 2017 BYU L. REV. 1417, 1427 (2017) (citing Richard H. Fallon, Jr., *"The Rule of Law" as a Concept in Constitutional Discourse*, 97 COLUM. L. REV. 1, 7–8 (1997)).

[40] *See generally* Reyes & Ward, *supra* note 8 (investigating interdisciplinary clash of values at the intersection of the regulation of the practice of law and algorithms).

[41] *See* Jack M. Balkin & Sanford Levinson, *Law and the Humanities: An Uneasy Relationship*, 18 YALE J.L. & HUMANS, 155, 173 (2006) ("Although it is impossible to understand contemporary legal theory without recognizing its strongly interdisciplinary character, law has a most curious relationship to interdisciplinarity."); *see also* Elizabeth Mertz, *The Role of Social Science in Law: Introduction*, *in* THE ROLE OF SOCIAL SCIENCE IN LAW xxix, xxix (Elizabeth Mertz ed., 2008) (discussing the troubled relationship between law and social science).

[42] Balkin & Levinson, *supra* note 41, at 173 ("Law seems endlessly to poach upon other disciplines and absorb many of their insights while still remaining law.").

[43] *Id*. at 178 ("Interdisciplinarity has made gains in law to the extent that it has allowed lawyers and legal scholars to do what they had already been doing—making persuasive arguments for the justification, change or interpretation of legal norms.").

[44] *See, e.g.*, Selbst, *supra* note 13, at 98–99 (describing the difficulty of explaining machine learning systems to satisfy the Fourth Amendment).

[45] This is referred to by Balkin and Levinson as "the internalist" approach to interdisciplinarity. *See* Balkin & Levinson, *supra* note 41, at 161.

[46] *See* Desai & Kroll, *supra* note 28, at 4 ("A more recent fear is that the rise of large

To cross the chasm of interdisciplinary discussion in law and emerging technology, industry, lawyers, and legal academics often turn to metaphors.[47] Metaphors, of course, have limits, and ultimately offer an incomplete mechanism for developing shared meaning in the law applicable to emerging technology.[48] Eventually, as metaphors break down, those on either side of a debate will start to dismiss or deride the language used by the other as "rhetoric," accusing each other of hiding meaning behind terms of art.[49]

---

data sets combined with machine learning . . . might allow those who use such techniques to wield power in ways society prohibits or should disfavor, but which society would not be able to detect. . . . The standard solution to this general problem is a call for transparency, which in this context has been called 'algorithmic transparency.' We argue that although the problems are real, the proposed solution will not work for important computer science reasons.").

[47] *See, e.g.*, T.J. McIntyre & Colin Scott, *Internet Filtering: Rhetoric, Legitimacy, Accountability and Responsibility*, *in* REGULATING TECHNOLOGIES 109, 112 (Brownsword & Yeung eds., 2008) (explaining the metaphorical development of the term "filtering"); Norberto Nuno Gomes de Andrade, *Technology and Metaphors: From Cyberspace to Ambient Intelligence*, 4 OBSERVARIO J. 121, 122 (2010) (noting that the development let to a wide array of new metaphors like "information highways," "metaverse," and "cyberspace"); Sam Harnett, Words Matter: How Tech Media Helped Write Gig Companies into Existence 4 (Aug. 7, 2020) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3668606 [https://perma.cc/8WMZ-5NKU] (discussing the importance of what shorthand is used in reporting).

[48] *See, e.g.*, Rebecca Crootof, *Autonomous Weapon Systems and the Limits of Analogy*, 9 HARV. NAT'L SEC. J. 54, 55–56 (2018) (showing the limitations of using weapon and combatant analogies for autonomous weapons); Lex Gill, *Law, Metaphor, and the Encrypted Machine*, 55 OSGOODE HALL L. J. 440, 455–56 (2018) (noting that the metaphors used in law are emotionally and ideologically loaded, and that overtime it becomes less clear that the terms are metaphors); Ryan Calo, *Robots as Legal Metaphors*, 30 HARV. J.L. & TECH. 209, 210 (2017) (arguing that judges use the term "robot" to justifying removing agency from people); Amy E. Sloan & Colin P. Starger, *New Wine in Old Wineskins: Metaphor and Legal Research*, 92 NOTRE DAME L. REV. ONLINE 1, 2 (2016) (showing the dangers of metaphor through the example of the "War on Drugs"); Neil M. Richards & William Smart, *How Should the Law Think About Robots?*, *in* ROBOT LAW 3, 16 (Ryan Calo, Michael Froomkin & Ian Kerr eds., 2016) ("In designing and implementing new technologies, we must be mindful of the metaphors we use to understand the technologies."); Lyria Bennett Moses, *Recurring Dilemmas: The Laws Race to Keep Up With Technological Change*, 2007 ILL. J.L. TECH. & POL'Y 239, 242 (2007) (commenting that there is no literature to explain why the use of metaphors are appropriate to reify technology and law); I. Glenn Cohen & Jonathan H. Blavin, *Gore, Gibson, and Goldsmith: The Evolution of Internet Metaphors in Law and Commentary*, 16 HARV. J.L. & TECH. 265, 268 (2002) ("By failing to adopt appropriate metaphors in regulating new technologies, courts risk creating bad law."); Joshua Fairfield, *The Magic Circle*, 14 VAND. J. ENT. & TECH. L. 823, 825 (2012) (arguing that it is a fallacy to distinguish between the "real" world and the "virtual" world).

[49] *See, e.g.*, Neil M. Richards & Jonathan King, *Three Paradoxes of Big Data*, 66 STAN. L. REV. ONLINE 41, 45 (2013) (concluding that those in the debate around the promises and perils of big data use "rhetoric of big data, in which utopian claims are being made that overstate its potential and understate the values on the other side of the equation").

Perhaps, however, even with their limited nature, metaphors are sufficient to convey the requisite shared meaning for effective legislation in a particular area of intersection between law and technology—for example, in law at the intersection of AI systems and criminal justice. We simply cannot be precise about the limits of metaphors and other common mechanisms for bridging meaning in interdisciplinary discussions. Without information beyond anecdote, it is impossible to tell whether metaphors suffice or whether more precision is necessary. A more evidence-driven understanding of the effect that interdisciplinary language usage has on law and policy-making at the intersection of AI systems in criminal justice proceedings is needed in order to improve outcomes for those individuals in criminal proceedings impacted in some way by AI systems.

II. AI AND CRIMINAL JUSTICE: A CURRENT LINGUISTIC BATTLEFIELD

Government entities at every level (city, state, and federal) increasingly employ AI systems in the administration of the criminal justice system.[50] Although the ostensible goal of such uses of AI systems is to increase efficiency and reduce bias,[51] a variety of scholars raise concerns about AI systems' ability to achieve such ideals.[52] For example, many researchers worry that the use of AI systems can reinforce and entrench bias in the criminal justice system rather than eliminate it, pointing to bias in the

[50] Alyssa M. Carlson, *The Need for Transparency in the Age of Predictive Sentencing Algorithms*, 103 IOWA L. REV. 303, 313 (2017) ("While risk assessment tools first emerged as a method of weighing parole decisions, the Justice Department's National Institute of Corrections now promotes the use of predictive algorithms for all phases of criminal cases, including sentencing. . . . [As of 2010] almost every state uses an assessment tool at one or more points in the criminal justice system to assist in the better management of offenders in institutions and in the community.").

[51] *See* Deeks, *Predicting Enemies*, *supra* note 9, at 1539–41 (discussing the use of algorithmic risk assessments to attempt to predict recidivism in an empirical manner).

[52] *See, e.g.*, *id*. (explaining that some scholars criticize algorithmic risk assessment tools as having biased inputs and a lack of transparency on how the algorithm operates); s*ee also* Anna Maria Barry-Jester, Ben Casselman & Dana Goldstein, *The New Science of Sentencing: Should Prison Sentences Be Based On Crimes That Haven't Been Committed Yet?*, THE MARSHALL PROJECT (Aug. 4, 2015), https://www.themarshallproject.org/2015/08/04/the-new-science-of-sentencing [https://perma.cc/EAV8-8TCE] (noting the heightened values at stake and querying whether predictive algorithms can uphold such values); Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing [https://perma.cc/7EPX-YKJC] (discussing concerns over inaccuracy of risk scores).

underlying data that fuel AI systems.[53] Because many governments acquire AI systems from the private sector,[54] other researchers document the myriad problems resulting from the inability of the defendant or the public to inspect and stress-test the computational components of proprietary algorithms.[55] In particular, many raise due process concerns when a defendant cannot review the computational components and raise questions about their accuracy.[56] Moreover, restricting public access to the computational components of AI systems reduces the ability to audit the system for accuracy and other bugs, like embedded bias.[57] Lastly, even when defendants and the public obtain access to AI systems for inspection, many observe that the black box nature of such systems—the idea that an AI system's complexity and ability to produce emergent results makes AI systems opaque—prevent understanding even when inspection is available.[58] Such commentators ask, if a defendant or the public cannot understand what it can inspect, what good does access to AI system code achieve on its own?[59]

Some measure of consensus seems to have developed around the idea that

---

[53] *See, e.g.*, Deeks, *Predicting Enemies*, *supra* note 9, at 1540 ("Others, however, have expressed a variety of concerns about the use of risk assessment algorithms, including worries about the use of flawed or biased data inputs, lack of transparency about how the algorithm is assembled and trained, and the difficulty in holding people accountable for flawed algorithm-driven decisions."); Carlson, *supra* note 50, at 312 (discussing the ProPublica study of COMPAS that determined the risk assessment algorithm was biased).

[54] *See* Carlson, *supra* note 50, at 315 ("As a result of the growing trend to implement actuarial risk assessment in sentencing, risk assessment has become a competitive industry with . . . for-profit businesses developing instruments.").

[55] *See, e.g.*, *id*. at 315–316 (". . . a massive disadvantage of proprietary risk assessment tools is that for-profit companies do not publicly disclose the formulas used to arrive at a risk score, so neither the defendants nor the public are privy to the calculations.").

[56] *See, e.g.*, *id*. at 320 (discussing Loomis as one of such examples).

[57] *See id*. at 323 (discussing a study at the Center for Criminal Justice Research at the University of Cincinnati which found only 30% of agencies using risk assessment tools had validated them on the local population); Edward J. Latessa, Richard Lemke, Matthew Makarios, Paula Smith & Christopher T. Lowenkamp, *The Creation and Validation of the Ohio Risk Assessment System (ORAS)*, 74 FED. PROB. 16, 17 n.3 (2010) (noting space constraints limited presentation methodology involved in the development of ORAS).

[58] *See* STEPHEN C. REA, A SURVEY OF FAIR AND RESPONSIBLE MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE: IMPLICATIONS OF CONSUMER FINANCIAL SERVICES 3 (2020), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3527034 [https://perma.cc/B3HD-CGT9] ("[T]he inherent complexities of ML algorithms that defy explanation even for the most expert practitioners can make it difficulty, if not impossible, to identify the root causes of unfair decisions. That same opacity also presents an obstacle for individuals who believe that they have been evaluated unfairly, want to challenge a decision, or try to determine who should—or even could—be held accountable for mistakes.").

[59] *See id*. at 31 ("[M]odel opacity inhibits the possibility of remedying an adverse decision, not to mention how difficult it is in practice for many on the receiving end to pursue actionable recourse.").

ameliorating problems related to the use of AI systems in criminal justice should involve one or more key concepts: transparency, explainability, accountability, and fairness. Designing an algorithm with attention to fairness, the argument goes, should, for example, encourage AI system creators to intentionally protect against the use of biased data.[60] In this context, legal commentators often use the term fairness to mean unbiased, process abiding, and independent.[61] Those concerned with access to information about AI systems in order to assist in a proper defense emphasize that due process requires transparency—including notice and the opportunity to challenge.[62] Others use the term transparency to refer to concepts like fish bowl transparency and reasoned transparency, proposing that public access to deeper information about the inner-operations of AI systems would enable greater accountability.[63] In this context, the term accountability seems to be connected to relative level of transparency, harnessing concepts of competence, diligence, and predictability.[64] Finally, many argue that to overcome the inherent black-box issues in complex algorithms, AI systems must be designed as explainable AI—developed in ways that lend themselves toward explanation that non-developers can understand.[65]

Recognizing that the opportune moment to embed features that uphold principles of transparency, explainability, accountability, and fairness occurs at the time of AI system design and creation, some scholars suggest regulating the procurement process in ways that incentivize attention to these principles.[66] The idea is that "[a]cting as consumers, governments can . . .

---

[60] *See* Solon Barocas & Andrew Selbst, *Big Data's Disparate Impact*, 104 CAL. L. REV. 671, 672 (2016) (arguing that here is no obvious method to correct bias in data used, but the result can be altered after it is completed).

[61] *See, e.g.*, Reyes & Ward, *supra* note 8, at 369; *see also* Deirdre K. Mulligan, Joshua A. Kroll, Nitin Kohli & Richmond Y. Wong, *This Thing Called Fairness: Disciplinary Confusion Realizing a Value in Technology*, 3 PROC. ACM HUM.-COMPUT. INTERACT. 119:1, 119:8–9 (2019) (noting non-discrimination and due process as concepts of fairness in law); Harry Surden, *Ethics of AI in Law: Basic Questions*, *in* OXFORD HANDBOOK OF ETHICS OF AI 719–20 (2020) (noting the use of fairness in due process).

[62] *See* Mike Ananny & Kate Crawford, *Seeing Without Knowing: Limitations to the Transparency Ideal and its Application to Algorithmic Accountability*, NEW MEDIA & SOC'Y 1, 2 (2016) ("Transparency concerns are commonly driven by a certain chain of logic: observation produces insights which create the knowledge required to govern and hold systems accountable.").

[63] *See* Cary Coglianese & David Lehr, *Transparency & Algorithmic Governance*, 71 ADMIN L. REV. 1, 19 (2019) (arguing fishbowl transparency "prioritizes the disclosure of information about what government is doing" while reasoned transparency "aims to promote an understanding of why government does what it does.").

[64] *See* Reyes & Ward, *supra* note 8, at 369 (listing core values of fairness in providing due process and integrity in decision making).

[65] *Id*.

[66] *See* Bloch-Webha, *Visible Policing*, *supra* note 7, at 973; Bloch-Webha, *Access to*

require more demanding contract terms that bring their procurement processes into alignment with due process and transparency requirements."[67] Indeed, this approach seems to be gaining traction as states take up the call to amend their procurement requirements. In Washington a legislative proposal would have required "that automated decisions systems have several transparency and accountability-enhancing features, including that they be open to audit and inspection by state agencies and third parties and that they be capable of giving intelligible explanations for the decisions they reach."[68] Vermont introduced a bill in 2021 that seemed to echo the concerns of Washington legislators, but extended the inquiry into an audit of "all automated decisions systems that are developed, used or procured by the State," with an eye to understanding the sources of data, the extent of prior audits and accuracy testing, and whether and to what extent the AI systems and state uses of those systems uphold core values like due process and fairness.[69] Maryland also considered legislation in 2021 that would build "responsible artificial intelligence standards" and require that vendors prove adherence to such standards in the procurement process, require further regulation of public use of AI systems, and expand the definition of "discriminatory act" to include actions taken through AI systems.[70]

As such laws have only just been introduced, the question remains: will designers of AI systems understand the message that the legislatures intend to convey with these new "responsible artificial intelligence standards"? Even if AI system designers do understand the message, can they technically achieve that which is required of them? Let's consider, for example, interest in ensuring that AI systems used in the criminal justice system are "fair" by design. The legal conception of fairness generally ties to antidiscrimination statutes and due process requirements, and stands as "a core principle in the goal of society-wide equilibrium of rights, opportunities, and resources."[71] Meanwhile, AI system creators have *at least* twenty-one different technical meanings of fairness to choose from when designing a system that is "fair by design."[72] If the law instructs the AI system designer to preference one such

Algorithms, *supra* note 8, at 1271.

[67] Bloch-Webha, *Access to Algorithms*, *supra* note 8, at 1308.

[68] *Id*. (citing H.R. 1655, 66th Leg., Reg. Sess. (Wash. 2019)).

[69] Vt. H.B. 263, *supra* note 27.

[70] Md. H.B. 1323, *supra* note 27.

[71] REA, *supra* note 58, at 21.

[72] *See* Arvind Narayanan, *Tutorial: 21 Fairness Definitions and Their Politics*, YOUTUBE (Mar. 1, 2018), https://youtu.be/jIXIuYdnyyk [https://perma.cc/3F8D-R4FY]; *see also* REA, *supra* note 58, at 22 ("Complicating matters further is the fact that ML communities themselves do not have a consensus definition of fairness. . . . [This] underscores how optimizing for fairness depends largely on the goals and desired outcomes for specific use cases."); SOLON BAROCAS, MORITZ HARDT & ARVIND NARAYANAN,

meaning over another, that requirement may result in a technical trade-off that legislatures neither contemplated nor intended.[73] For example, if a legislature requires statistical parity as an anti-bias measure, it may force a trade-off in accuracy.[74] The same definitional difficulty exists for each of the core concepts legal scholars hope will help shore up gaps between the use of AI systems in criminal justice and important legal and constitutional norms like due process.[75]

If these definitional difficulties exist, one of two possibilities must be true. One possibility is that lawmakers knew of the definitional differences and did not care, electing to try and force AI systems to conform to legal terms of art even when those terms inherently conflicted with technical realities and understanding. Alternatively, lawmakers failed to understand the mismatch in the terminology, and had these laws been enacted, certain difficulties in implementation would have emerged. Although the debates of these issues suggest the latter scenario is at play, it does so only anecdotally. Moving beyond anecdote may enable deeper understanding of the linguistic conflict, its implications, and anticipated consequences.

III. CORPUS LINGUISTICS AS A TOOL FOR CLARIFYING LAW AND POLICY AT THE INTERSECTION OF AI SYSTEMS AND CRIMINAL JUSTICE

Anecdotal evidence in the literature investigating the use of AI systems and criminal justice suggests that participants in the law-making process talk past one another, producing sub-optimal legal outcomes.[76] But anecdote and metaphor only move the needle so far.[77] In an attempt to arrive at a more evidence-driven understanding of the role of linguistic conflict in law-making related to AI systems and criminal justice, this section introduces the prospect of using corpus linguistics.[78] "Corpus analysis is a form of text analysis which

---

FAIRNESS AND MACHINE LEARNING 45 (2021) ("Many fairness criteria have been proposed over the years, each aiming to formalize different desiderata.").

[73] *See* REA, *supra* note 58, at 22 ("Satisfying different fairness conditions almost always entails some degree of trade-off with respect to accuracy . . . . There are also trade-offs involved in trying to balance the conditions for individual versus group fairness.").

[74] *See* Cofone, *supra* note 9, at 1434 (discussing factors affecting algorithmic unfairness and bias); BAROCAS ET AL., *supra* note 72, at 31 (noting the focus on accuracy and accuracy-fairness trade-off in machine learning discussions).

[75] *See, e.g.*, BAROCAS ET AL., *supra* note 72, at 56 (discussing recall-precision trade-off).

[76] *See supra* Part I.B.

[77] *See supra* Part I.A.

[78] *See* Heather Froehlich, *Corpus Analysis with Antconc*, PROGRAMMING HISTORIAN (Nov. 5, 2020), https://programminghistorian.org/en/lessons/corpus-analysis-with-antconc [https://perma.cc/MZ7W-G3X8] ("Corpus analysis is especially useful for testing intuitions

allows you to make comparisons between textual objects at large scale."[79] Corpus linguistics focuses on accuracy in describing language,[80] and to that end, embraces "complexity and variation as inherent in language."[81] Together, these elements of corpus linguistics naturally lend the discipline to a focus on "describing the use of language as a communicative tool."[82] In particular, one common methodological approach in corpus linguistics involves "the study of what is termed 'genre variation,' i.e. how language usage varies according to the context in which it occurs."[83] It is precisely this type of data-driven investigation which promises to shed light on whether those making law adequately convey the message they intend to those organizing their affairs under the law.

Applying corpus linguistics to law is not new. In recent years, a movement emerged encouraging use of corpus linguistics as a method for uncovering the "plain meaning" of ambiguous words in statutes.[84] Like the push to use AI systems to reduce the opportunity that bias might influence a decision-maker's use of discretion, the hope for corpus linguistics centers on offering judges a more empirical, more transparent, and more neutral and

---

about texts and/or triangulating results from other digital methods.").

[79] *Id*.

[80] *See* CHARLES F. MEYER, ENGLISH CORPUS LINGUISTICS 4 (2002) (discussing how corpus linguists prioritize descriptive adequacy over explanatory adequacy in their studies).

[81] *Id*. at 3.

[82] *Id*. at 5.

[83] *Id*. at 18.

[84] *See, e.g.*, Stephen C. Mouritsen, Note, *The Dictionary is Not a Fortress: Definitional Fallacies and a Corpus-Based Approach to Plain Meaning*, 2010 BYU L. REV. 1915, 1919 (2010) (advocating for the use of a corpus-based approach to interpret legal language when contextual cues and legislative definitions do not help); Gries & Slocum, *supra* note 39 (arguing that corpus analysis and similar empirical based study should be used to help judicial interpretation of legal language); Thomas R. Lee & Stephen C. Mouritsen, *Judging Ordinary Meaning*, 127 YALE L.J. 788, 788 (2018) (proposing the use of corpus linguistics to resolve the indeterminacy of ordinary meaning); Thomas R. Lee & James C. Phillips, *Data-Driven Originalism*, 167 U. PA. L. REV. 261, 262 (2019) (using corpus linguistics to uncover the original communicative content of the Constitution); Stephen C. Mouritsen, *Contract Interpretation with Corpus Linguistics*, 94 WASH. L. REV. 1337, 1341 (2019) (offering corpus linguistic as a middle ground between formalism and contextualism for the purpose of interpreting contractual language); Jennifer L. Mascott, *Who are Officers of the United States?*, 70 STAN. L. REV. 443, 453 (2018) (using corpus linguistics to determine whether the term "officer" is consistent with the term's original public meaning); Lawrence M. Solan, *Can Corpus Linguistics Help Make Originalism Scientific?*, 126 YALE L.J. FORUM 57, 57–58 (2016) (proposing corpus linguistic as a research tool to analyze the original public meaning during the Founding Era); Lawrence M. Solan & Tammy Gales, *Corpus Linguistics as a Tool in Legal Interpretation*, 2017 BYU L. REV. 1311, 1312–13 (2017) (arguing that corpus linguistic is a useful tool in constructing the ordinary meaning when such meaning is legally relevant).

consistent method for interpreting and applying statutes.[85] The approach encounters a variety of critiques.[86] For example, some argue that the emphasis on the frequency with which words appear in a corpus is an improper emphasis for assessing plain meaning.[87] Others argue that those using corpus linguistics often make inferential errors that reduce its usefulness in uncovering plain or ordinary meaning.[88] Putting this debate in the context of law's history with the humanities, we might query whether those advancing the role of corpus linguistics in judicial interpretation of statutory language are adopting an internalist approach to law and interdisciplinarity.[89] Namely, in applying corpus linguistics to the interpretation of the plain meaning of ambiguous words in a statute, proponents of the tool seek to borrow from another discipline in order to solve a traditional legal question.[90] Viewed in this light, many of the critiques leveled against the use of corpus linguistics in law are aimed squarely at the appropriateness of an internalist approach to law and interdisciplinarity.

Gratefully, this Article need not take a position in the debate over the use of corpus linguistics to identify the plain meaning of words in a statue. This Article tests a research question vastly different than uncovering the plain meaning of a word used in a statute.[91] The anecdotal evidence of

---

[85] *See* Thomas R. Lee & Stephen C. Mouritsen, *The Corpus and the Critics*, 88 U. CHI. L. REV. 275, 282, 285 (2021) (illustrating importance of having a common linguistic legal directive and the deficiencies of using pure intuitions in interpretation).

[86] *See, e.g.*, Carissa Byrne Hessick, *Corpus Linguistics and the Criminal Law*, 2017 BYU L. REV. 1503, 1511 (2017) (arguing against the use of corpus linguistics as a new interpretive theory in criminal adjudication because the tool impedes the public notice requirement inherent in certain law (like criminal statutes)); Evan C. Zoldan, *Corpus Linguistics and the Dream of Objectivity*, 50 SETON HALL L. REV. 401, 431–34 (2019) (challenging the validity of speech community selection when analyzing legal text); Kevin P. Tobia, *Testing Ordinary Meaning*, 134 HARV. L. REV. 727, 753–77 (2020) (using survey results to argue that corpus linguistics incorrectly focuses on prototypical meaning); Matthew Jennejohn, Samuel Nelson & D. Carolina Nunez, *Hidden Bias in Empirical Textualism*, 109 GEO. L.J. 767, 771 (2021) (arguing that the Corpus of Historical American English (COHA) is sexist); Francis J. Mootz III, *Corpus Linguistics and Vico's Lament: Against Vivisectional Jurisprudence*, 20 NEV. L.J. 845, 845–47 (2020) (arguing that use of corpus linguistics to devise plain meaning reinforces an incorrect theory of the relationship between law-making and law enforcement); Anya Bernstein, *Democratizing Interpretation*, 60 WM. & MARY L. REV. 435, 444 (2018) (challenging the claim that corpus linguistics eliminates all manner of discretion in determining ordinary meaning).

[87] *See, e.g.*, Hessick, *supra* note 86, at 1514 ("Corpus linguistics tells us that the ordinary meaning of a statutory term ought to be resolved by looking to the frequency with which a term is used a certain way. This is a problematic theory for the interpretation of criminal laws because it creates problems of notice and accountability.").

[88] *See* Tobia, *supra* note 86, at 794–797 (listing fallacies of legal corpus linguistics).

[89] *See* Balkin & Levinson, *supra* note 41, at 161-164.

[90] *Id*. at 163–64.

[91] My hope, then, is that this modest Article will: (1) move the discussion around AI

miscommunication at the intersection of AI systems and criminal justice suggests that miscommunication occurs long before words are arranged in a statute and adopted as new law. Determining the extent and nature of miscommunication around key terms resembles the type of inquiry commonly undertaken by linguists—namely, genre variation: how different speech communities use the same words. Rather than use a technique from a different discipline to solve traditional legal inquiries, this Article uses an interdisciplinary technique to answer an interdisciplinary research question that might provide insight into how to improve law-making at the intersection of AI systems and criminal justice.[92]

IV. PATTERNS OF LANGUAGE USAGE REVEAL MISCOMMUNICATION AND VALUE CONFLICT AT THE INTERSECTION OF AI SYSTEMS AND CRIMINAL JUSTICE

Every corpus linguistic research investigation evolves out of a core research goal and linguistic hypothesis.[93] In this section, the Article presents the results of a genre variation study of the words transparency, accountability, explainability, and fairness in the context of five genres pertinent to the law-making process. The analysis presented here is a product of my efforts to test the linguistic hypothesis[94] that different communities in discussions about AI systems and criminal justice use the terms transparency, accountability, explainability, and fairness differently—corresponding both to different definitions and different values. Long before a judge will ever consider the plain or ordinary meaning of one of these terms, the term must first be used in a statute. If, at the time that lawmakers write, discuss, and vote to adopt a statute containing those terms, they rely upon incoherent discussions with stakeholders using the same words but different meanings, the resulting law will likely underperform in its role as communicator of clear rules as part of an effective rule of law system.

---

systems and accountability, transparency, and explainability in the context of the criminal justice system, and (2) offer an alternative use case for corpus linguistics in law.

[92] Such an approach might be categorized as an externalist approach to the intersection of law and other disciplines. *See* Balkin & Levinson, *supra* note 41, at 161–64 (discussing the differences between internalist and externalist approaches to law and legal education).

[93] *See* Meyer, *supra* note 80.

[94] Corpus linguistics is often criticized for simply counting how frequently a given linguistic construction occurs in any given corpus. *Cf.* Meyer, *supra* note 80, at 102 ("To move beyond simply counting features in a corpus, it is imperative before undertaking a corpus analysis to have a particular research question in mind, and to regard the analysis of a corpus as both 'qualitative' and 'quantitative' research—research that uses statistical counts or linguistic examples to test a clearly defined linguistic hypothesis.").

To test my hypothesis, I conducted collocation analyses[95] and concordance line analyses[96] of the terms transparency, accountability, explainability, and fairness[97] using corpora representing each of five different stakeholder groups that are involved in the development of law at the intersection of AI systems and criminal justice: legal academia, computer science and engineering academia, lawmakers, judges, and the general public.[98] The collocation analysis offers insight into "which words tend to occur next to or close to [the] search term and sort[s] those results by frequency."[99] The concordance line analysis, for its part, provides further insight into the collocation results by providing evidence of the context in which the words appear.[100]

In terms of the data studied, considering the approaches of various stakeholders required the study of various corpora. To uncover how legal academics and lawmakers use these terms, I sourced and created my own corpora. The legal academic corpus contains the text of every law review article using the terms transparency, accountability, explainability and fairness with or without a connection to AI systems since 2018.[101] To consider the use of these terms by state lawmakers, I sourced and created a corpus consisting of every proposed or adopted legislation relating to AI

[95] Collocation analysis gives the linguist "a sense for which words tend to occur next to or close to your search term and sort those results by frequency." *Quickstart Guide to AntConc*, MCGRAW CTR. FOR TEACHING & LEARNING, https://mcgrawect.princeton.edu/guides/Quickstart-Guide-AntConc.pdf [https://perma.cc/L94B-6P43].

[96] "A concordance lists the occurrences of certain words in the corpus ordered by how frequently those words are used as well as the context in which those terms appear." *Id.*

[97] I note the fantastic suggestion of *Journal of Law & Innovation* Symposium participants to expand the list of words in the study beyond these four examined here. Although outside the scope of this Article, the suggestion represents an area for future fruitful research.

[98] Every corpus linguistics investigation must begin by answering certain threshold questions: (1) "What is the relevant speech community I want to investigate?", and (2) "What is the relevant time period I want to investigate?" Lee & Mouritsen, *supra* note 85, at 293–94.

[99] MCGRAW CTR. FOR TEACHING & LEARNING, *supra* note 95.

[100] *Id.*

[101] As to the mechanics of this, I searched SSRN and HeinOnline for law review articles using the four terms of interest in connection with AI systems. Due to the volume of articles, I limited the results to those that hit on the search terms and were published in the last five years. I downloaded the pdfs, and then uploaded them to AntFile Converter, which converted each document into a plain text format compatible with the corpus linguistics software AntConc. *See* Laurence Anthony, *AntFile Converter Homepage*, https://www.laurenceanthony.net/software/antfileconverter/ [https://perma.cc/9AF4-AL4D] (conversion software). I used the same approach to create a second corpus of articles that hit on the search terms and were published within the last five years but were not articles focused substantively on AI systems.

systems at the state level since 2018.[102] For federal legislative discussions, I performed the analysis using the Corpus of the Current US Code (COCUSC), supplemented by consideration of federal legislation relating to AI systems that was introduced but not adopted in 2018 or later.[103] To consider the voice of more technical disciplines developing AI systems, I sourced and created a corpus containing the text of academic articles or whitepapers using the terms transparency, accountability, explainability, and fairness in connection with AI systems since 2018.[104] To look at the way judges use these terms when deciding cases involving AI systems, I first conducted my analysis using the Corpus of US Caselaw (CUSC),[105] and then used a self-created corpus to reflect judicial decisions using the four terms in the last 10 years.[106] Considering these two corpora together offers insight into how judicial use of the four terms changed over time. And to uncover whether any of the uses suggested by the other four corpora resembled general public understanding of those terms, I turned to an analysis of the News on the Web (NOW)

---

[102] As to the mechanics of this, I pulled every adopted or proposed bill from Westlaw and LegiScan that hit on several variations of the term AI systems (including artificial intelligence, AI, AI system, and automated decision system) since 2019 at both the state and federal levels. I used AntFile Converter in the same manner as described *supra* note 101 to convert the bills from pdfs to plain text that AntConc could use.

[103] For current federal law, I used the Corpus of Current US Code (COCUSC), BYU LAW, https://lawcorpus.byu.edu/. [https://perma.cc/7FPF-B3Q4] Recognizing the debate about the extent to which the metric is useful at all, I note here that the terms in the study are used with the following frequency in the COCUSC: transparency: 4,553; accountability: 20,124; explanation: 11,764 (including variations of explain and explainable); fairness: 1,808. The mechanics of creating a supplemental corpus for federal legislation relating to AI systems that had been introduced but never adopted are similar to that described for state laws, *supra* note 101.

[104] The mechanics of this are identical to those described *supra* note 69, except that I conducted my search in discipline-appropriate periodical and research databases, including SSRN, https://www.ssrn.com/index.cfm/en/ [https://perma.cc/8TBP-2VZS] FRONTIERS, https://www.frontiersin.org/ [https://perma.cc/PUT7-M6LT] and the historical records of papers presented at the Fairness, Accountability, and Transparency in Machine Learning conference (FaT), https://www.fatml.org/ [https://perma.cc/6ECW-YPAS] and the ACM Conference on Fairness, Accountability, and Transparency (FaCCT), https://facctconference.org/ [https://perma.cc/VH44-SQUZ].

[105] Corpus of US Caselaw (CUSC), BYU LAW, https://lawcorpus.byu.edu/ [https://perma.cc/7FPF-B3Q4].

[106] As to the mechanics of this, I pulled every judicial decision since 2010 at both the federal and state level that hit on a search of each of the four terms on Westlaw. I downloaded the decisions as pdf files and uploaded them to AntFile Converter, which converted each document into a plain text format compatible with the corpus linguistics software AntConc. *See* Laurence Anthony, *AntFile Converter Homepage*, https://www.laurenceanthony.net/software/antfileconverter/ [https://perma.cc/9AF4-AL4D] (conversion software).

Corpus[107] and the Corpus of Contemporary American English (COCA).[108] The below discussion presents the results of the analysis of each stakeholder group and its related corpus.

A. *Legal Academics*

Legal academics use the terms transparency, accountability, explainability, and fairness in a variety of contexts, including, but certainly not limited to, the creation, use, and application of AI systems in public administration settings such as the criminal justice system. In order to get a more complete sense of how legal academics use these terms, I conducted two separate collocation analysis. First, I looked at the use of those terms in legal academic articles that had some connection to AI systems specifically.[109] Second, I performed a collocation analysis of the four terms across any law review article in which they appeared—whether the article mentioned AI systems or not.[110] Comparing the two sets of results gives an indication as to whether legal academics maintain a primary set of definitions and values of these terms and intend them to apply across settings, or whether there is some, even minuscule, recognition that such terms may embody different definitions and values in the context of AI systems.

When considered in the context of articles that touch on, even tangentially, AI systems, the results of the collocation analysis, presented in Table 1 below, reveal that the legal literature views transparency as closely associated with accountability, explainability, and fairness. Meanwhile, legal academics seem to strongly associate accountability with transparency and fairness, but not explanation. The results are similar for fairness—highly associated with transparency and accountability, but not explanation. Indeed, the other three reform terms do not appear in the top collocate results for explanation at all. A concordance analysis of the corpus confirms this assessment of the collocation results. The concordance analysis for transparency reveals that many of the transparency-accountability collocates are in sentences explicitly recognizing a relationship between the two

---

[107] NOW Corpus (News on the Web), https://www.english-corpora.org/now/ [https://perma.cc/SWU7-7SNC**]**.

[108] Corpus of Contemporary American English, https://www.english-corpora.org/coca/ [https://perma.cc/7ZQB-3555].

[109] The corpus that resulted from the procedure explained *supra*, note 101, remains on file with the author. Note that within that corpus of articles, the term "explanation" appeared 1,880 times, the term "fairness" appeared 1,241 times, the term "accountability" appeared 1,106 times, and the term "transparency" appeared 2,768 times.

[110] The corpus that resulted from the procedure explained *supra* note 101, remains on file with the author. Note that within that corpus of articles, the term "explanation" appeared 883 times, the term "fairness" appeared 911 times, the term "accountability" appeared 381 times, and the term "transparency" appeared 268 times.

principles or arguing that one exists. According to the results, transparency makes accountability possible.

In contrast, the words most often associated with explanation focus on the circumstances in which the law expects an explanation to be given, and what the explanation should contain.[111] The concordance line analysis offers insight as to why the results return this way. The concordance line analysis shows that the literature ties the ability of the government to be transparent, accountable and fair, to the sufficiency of the explanation that it can provide for consequential decisions that it makes. Many of the references to AI systems and explanation in this context relate to concern that if an AI system explanation does not demonstrate certain characteristics, the government will be unable to fulfill expectations of transparency, accountability, and fairness. Ultimately, this indicates that, at least in the context of decisions by AI systems, explanations of decisions and the reasoning behind decisions sits at the foundation of core legal values embedded in the terms transparency, accountability, and fairness. Those values are also revealed by the corpus analysis itself: transparency should be reasoned and is for the benefit of the public; accountability should be systemic, applies equally to public institutions, is for the benefit of the public, and may require collaboration; and fairness requires accuracy, removal of bias, and applies in both group and individual settings.

Table 1: Collocates of AI Reform Terms by Legal Academics in AI-Related Legal Research

Transparency		Accountability		Explanation		Fairness	
253	accountability	253	transparency	487	right	157	algorithmic
224	algorithmic	165	algorithmic	139	decision	120	transparency
163	report(s)	87	public	86	ex post	107	notion(s)
120	fairness	71	fairness	77	why	71	accountability
106	fishbowl	65	AI	74	provide	63	accuracy
99	reasoned	57	governance	68	GDPR	57	bias
93	government	52	systems	65	what	56	AI
84	governance	40	conference	64	making	52	group
83	public	31	GDPR	57	should	47	individual
78	between	31	systemic	50	based	45	different
67	lack	29	mechanisms	49	decisions	42	measures
64	explainability	26	collaborative	48	specific	40	conference

[111] A concordance analysis of the corpus confirms this assessment of the collocation results. When it comes to explaining "how" an algorithm reached its decision, the literature suggests explanations for everything from how the model functions and generates predictions, to how the algorithm is influenced by an input or changes to inputs, and strongly emphasizes the need for the explanation to be "human-understandable."

When considered in the context of legal academic articles unrelated to AI systems, the results shift somewhat. As shown in Table 2 below, the highest collocates for transparency contain only accountability and fairness, and not explanation. And while transparency remains a frequent collocate for accountability, none of the other AI reform terms appear in the most frequent collocates for explanation or fairness. Further, the collocation results indicate that the broader legal literature calls for greater transparency in certain contexts (government, public, use of force, procurement); sees accountability as an aspect of justice, with an emphasis on certain contexts (delictual, criminal, children, political, human rights, public, democratic process); recognizes that different situations call for different types of explanations: legal, best, reasoned, theory, inference, proof, and contrastive; and views fairness as intricately connected to procedural, criminal, racial, and ethnic justice, while also simultaneously acting as a pervasive concept in the business law and contracts context (entire fairness doctrine, reasonableness of provisions, and insider trading). Interestingly, the general legal literature does not place such an emphasis on explanation in relation to accountability, transparency, and fairness, as the legal literature on AI systems does.[112]

Table 2: Collocates of AI Reform Terms by Legal Academics in General Research

Transparency		Accountability		Explanation		Fairness	
29	accountability	60	delictual	112	law	97	doctrine
25	government	44	capacity	104	best	94	ex ante
23	public	43	criminal	101	reasoned	80	justice
22	lack	38	children	44	theory	77	procedural
22	greater	29	transparency	43	nature	64	criminal
12	increase	26	political	40	requirement	51	legitimacy
10	force	25	rights	38	APA	45	insider trading
9	fairness	23	justice	37	inference	43	racial
8	increased	23	public	36	proof	42	ethnic
8	procurement	22	human	35	juridical	32	reasonableness
7	documentation	13	cities	32	tort	31	jurisdiction
6	UNCITRAL	11	democratic	30	contrastive	26	commission

[112] Again, the concordance line analysis confirms this interpretation of the collocation results. The concordance line analysis shows that in more general legal literature, many of the issues related to explanation relates to the role explanation should play in a specific decision-making context (criminal, tort, etc.), the relative weight to be put on specific types of explanations when offered as evidence, and other related issues. This may reflect the different starting point of the inquiries in the AI-related literature (how does the use of AI in public administration affect rule of law) and the more general legal literature. Again, a corpus linguistic study can only tell part of the story.

Analyzing the use of the terms transparency, accountability, explanation, and fairness in the legal academic literature suggests that law's emerging consensus around the need for AI systems to comply with four key concepts—transparency, accountability, explanation, and fairness—actually represents demands that AI systems fulfil a rather significantly varied number of values embedded in law's use of those terms. Yes, the literature on AI systems tailors use of the four terms slightly in comparison to their use in more general substantive areas of law. However, with one exception, the difference seems to be in form rather than in substance, with the values represented by each term remaining constant across corpora. The one exception, of course, is explanation. Somehow, explanation represents a core building block to achieving transparency, accountability, and fairness in the AI systems context, but not in the more general legal context. Perhaps, this difference represents the relative difficulty law encounters as it attempts to understand the technical functions of AI systems.[113] Nevertheless, when the shoe is on the other foot, the law seems to expect that those building AI systems will easily parse the appropriate values from law's use the terms, including how those values vary with the context in which the AI systems are put to use. This expectation, of course, assumes shared values between legal systems and socio-technical systems. It seems rather paradoxical for law to insist on this assumption for builders of technical systems when the collocate results suggest that law does not always understand those technical systems. Because this expectation persists, however, it is useful to better understand whether AI system designers do share the law's understanding of and the legal values embedded in the four key AI system legal reform terms. To do so requires consideration of how researchers in the fields that build AI systems use the terms transparency, accountability, explanation and fairness in their own work.

B. *Non-Legal Researchers*

To investigate the context in which researchers in the technical sciences use the terms transparency, accountability, explanation, and fairness when discussing and building AI systems, I built a corpus of articles and whitepapers written by computer scientists and engineers.[114] The collocation

---

[113] At least one study found that "AI literacy significantly influence[s] the perceived informational fairness" of an AI system. Jakob Schoeffer, Niklas Kuehl & Yvette Machowski, *"There is Not enough Information": On the Effects of Explanations on the Perceptions of Informational Fairness and Trustworthiness in Automated Decision-Making*, 2022 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY 1616, 1616, https://doi.org/10.1145/3531146.3533218 [https://perma.cc/QE26-F7WS].

[114] The group of technical science literature with which legal academics most frequently interact in this area publish in the proceedings of the Association of Computing Machinery

analysis performed on this corpus offers a window into the words statistically most likely to appear in the same context as transparency, accountability, explanation, and fairness from within the discipline building the systems that legal academics seek to influence. That snapshot, captured in Table 3, below, indicates that, like their law-focused counterparts, those building AI systems view transparency, accountability, and fairness as intricately connected. The focus for explanation, however, appears to vary significantly from that discussed in the legal literature. Namely, the collocation and concordance line analysis suggest that those developing AI systems simply want to figure out how to explain AI system outputs in useful ways, and have not yet arrived at a phase of development where questions of values expected to flow from explanations can be interrogated. Further, explanation does not appear among the highest collocates for any of the other three terms, signifying that those developing AI systems believe transparency, accountability, and fairness can be achieved without explanation. Indeed, the concordance line analysis revealed that at least some technical researchers worry that the expanding legal obsession with explainable AI as a way to improve transparency and accountability may create a type of "transparency fallacy" in which explainable AI systems are automatically the best AI systems, irrespective of other important technical performance benchmarks. This stands in stark contrast to the legal literature on AI systems, which viewed explanation as a foundational element to achieving transparency, accountability, and fairness.

The results also suggest that those building AI systems worry about slightly different issues than those prevalent in the legal literature. For example, in the context of providing adequate transparency, technical researchers worry about the effect of the technical process that makes transparency possible: privacy, tracking, and trust. Similarly, technical researchers place an emphasis on the data and mechanisms that enable accountability,[115] show high interest in the constraints on and metrics for

(ACM). To attempt to better understand how well legal and non-legal academics in this space are communicating with each other, I therefore drew the material for this corpus from the last five years of proceedings of the ACM Conference on Fairness Accountability and Transparency (ACM FAccT), https://facctconference.org/index.html [https://perma.cc/QNE8-CJWH]. I downloaded the pdfs and converted them to txt files using AntConverter. The resulting corpus remains on file with the author. Within the corpus, the frequency of the four AI reform terms that are focus of this study are as follows: transparency: 1,133; accountability: 1,258; explanation: 1,023; and fairness: 7346.

[115] Part of this emphasis may reflect the technical need for different data depending on the type of accountability sought. *See, e.g.*, Joseph Donia, *Normative Logics of Algorithmic Accountability*, 2022 ACM CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY 598, 598, https://doi.org/10.1145/3531146.3533123 [https://perma.cc/R7LK-H57W] ("[A]cademic, policy, and public discourse has increasingly emphasized accountability as a desirable, if not elusive, feature of system design, and component of effective governance. Accountability, however, is a versatile concept that has been operationalized in a number of

fairness,[116] and evidence interest in the models, methods, features, and maps of explanations stand out as collocates absent from the legal literature corpus.

Table 3: Collocates of AI Reform Terms by Technical Science Researchers

Transparency		Accountability		Explanation		Fairness	
611	accountability	611	transparency	175	model(s)	608	transparency
608	fairness	595	fairness	90	method(s)	595	accountability
471	conference	505	conference	85	feature(s)	575	conference
46	algorithmic	147	algorithmic	83	algorithm(s)	521	learning
41	privacy	45	machine	77	counterfactual	467	algorithmic
30	tracking	29	relationships	72	map(s)	437	group
30	report	28	framework	57	local	421	constraint(s)
27	computing	27	data	50	global	356	notion(s)
22	trust	26	impact	45	decision	329	machinery
21	machinery	25	public	43	right	295	metrics
720	towards	25	towards	42	different	258	accuracy
18	development	24	mechanisms	38	why	210	individual

This comparison of the four terms most used by legal academics in an effort to reform the use of AI systems in public administration, including in the criminal justice context, with the way those building the systems use the same terms begins to hint at evidence of potential miscommunication. Legal academics focus on the values the terms transparency, accuracy, explanation, and fairness represent in the legal system. In particular, the legal concepts underpinning these terms seem to place a rather significant value on the role of explanation. Meanwhile, those building AI systems view the same four

ways across different use-contexts, policy settings and research disciplines. . . . [T]his article introduces five normative logics underpinning discussions of algorithmic accountability that appear in the academic research literature: (1) accountability as verification, (2) accountability as representation, (3) accountability as social license, (4) accountability as fiduciary duty, and (5) accountability as legal compliance.").

[116] It is worth noting that the concordance line analysis reveals that much of the literature worries about a privacy-fairness trade-off when attempting to pursue both values in the code of a given AI system. In at least one instance, authors explicitly recognized that "legal concerns may hinder one or the other of privacy and fairness when they are both pursued." Michael D. Ekstrand, Rezvan Joshaghani, Hoda Mehrpouyan, *Privacy for All: Ensuring Fair and Equitable Privacy Protections*, 81 PROC. MACH. LEARNING RSCH. 35, 35 (2018). Further, at least some suggest that although the technical literature has proposed a variety of algorithmic fairness metrics in recent years, and "[w]hile fulfilling these metrics is typically intuitively appealing, the literature has also shown that they are often mutually incompatible. This gives rise to questions such as which metrics should be evaluated and whether fulfilling any of these metrics is a necessary condition for fairness." Eleonora Viganò, Corinna Hertweck, Christoph Heitz & Michele Loi, *People are Not Coins: Morally Distinct Types of Predictions Necessitate Different Fairness Constraints*, 2022 ACM CONF. ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 2293, 2293.

terms through a technical lens, considering transparency, accountability, and fairness as metrics that can be achieved even in the absence of explanation—a concept the technical literature seems unsure how to fulfill, and even worries will dilute focus on true understanding of the AI system. Lastly, the fact that the results from the technical literature return issues that appear only sparingly in the AI system legal academic literature results might signal areas law is largely overlooking, such as data integrity, and privacy-transparency trade-offs.

C. *Lawmakers*

To consider the context in which federal lawmakers use the terms transparency, accountability, explanation and fairness, I used the collocation function of the Corpus of Current US Code (COCUSC).[117] Doing so provides a snapshot of the words that are statistically most likely to appear in the same context as transparency, accountability, explanation, and fairness in the entirety of the US Code as it existed as of July 2019.[118] Excluding conjunctions, the most common collocates for each of these four terms are quite varied, as evidenced below.

Table 4: Collocates of AI Reform Terms in the US Code

Transparency		Accountability		Explanation		Fairness	
151	accountability	1109	office	139	reasons	15	equity
72	act	1054	government	136	written	13	regulatory
44	wall	293	accounting	120	any	12	business
44	street	284	general	114	how	12	fee
28	public	231	substituted	106	detailed	11	ensure
27	federal	161	health	99	include	11	small
25	increase	160	portability	83	basis	10	procedural
24	promote	155	performance	57	provide	10	such
23	government	153	insurance	53	together	8	cabin
23	improve	151	transparency	50	report	8	enforcement
21	ensure	93	congressional	40	determination	8	notice
21	market	89	measures	31	provided	8	promote
19	information	85	under	31	justification	8	user
18	price	83	referred	30	action	7	cited
16	transparency	77	public	30	challenges	7	considered
15	financial	68	report	30	notice	7	convenience

[117] Anyone can use COCUSC. *See* Corpus Linguistics, BYU LAW, https://lawcorpus.byu.edu/ [https://perma.cc/7FPF-B3Q4]. To replicate the results presented take the following steps: (1) Select 'Collocates'; (2) Enter "TRANSPARENCY_n" (or whichever word you want to replicate results for) in the field for a word or phrase; (3) Enter * in the collocates section, and (4) Initiate the search.

[118] *Id.*

Transparency and accountability are highly connected to one another. Both are also frequently collocated with issues relating to government, government offices, federal legal institutions (congress), money, and finance. Further, transparency's frequent colocation with words like improve, increase, and ensure, evidence a strong desire to achieve deeper levels of information sharing around issues of governance and finance.[119] The term explanation, for its part, clearly relates to law's demand that decision-makers provide reasons (often written), that offer detailed reports of the how, the basis, the justification, the action, that may have been taken. The analysis also reveals a clear tie-in between explanation and fairness: both are frequently collocated with the word "notice," evidencing the level at which the law tends to attach due process requirements to concepts of fairness and the imperative of providing an explanation.

State legislation echoes the connections between transparency and accountability, but both transparency and accountability also appear to be key aspects of state conceptions of fairness. Interestingly, state legislation seems to view explicability and oversight through audits as key components of true transparency and accountability, while primarily interested in fairness as a process concern associated with rooting out bias and ensuring accuracy. State legislatures seem less clear on what to expect from the concept of explanation, with only four words collocated with sufficient frequency to be worth reporting, and even then, only at very low levels. The emphasis from this very small amount of evidence appears to be on requiring "vendors" to "create" explanations of the outcomes AI systems reach.

Table 5: Collocates of AI Reform Terms in State Legislation

Transparency		Accountability		Explanation		Fairness	
11	prosecutorial	19	office	3	vendor	4	transparency
7	accountability	18	government	3	create	3	accountability
4	fairness	7	accounting	1	thorough	3	digital
4	securities	5	general	1	seemingly	3	process
3	undermine	4	substituted			2	concerns
2	explicability	4	health			2	crimination
2	auditability	3	portability			2	accuracy
2	associations	3	approved			2	bias
2	adequate	3	fairness				
		2	explicability				
		2	auditability				
		2	concerns				

[119] Notably, transparency is collocated another nineteen times with improving and another fifteen times with increasing. The law is preoccupied with radical transparency of systemically important institutions like the federal government and Wall Street.

Considering the use of the terms transparency, accountability, explanation, fairness in federal and state legislation together, and in comparison with the way that legal and non-legal academics use those terms, several lessons emerge. First, the use of these four terms in legislation meant to regulate technical requirements of AI systems procured for government largely rely on legal or legal-related meanings to the exclusion of technical understandings of the same terms. The federal and state statutory collocates with the transparency, accountability, explanation, and fairness demonstrates striking similarities to the collocates of those same terms in the legal academic corpora.[120] As a result, the federal and state statutory corpora reflect the same disconnect with the use of these terms by those building AI systems as the legal academic corpora.[121] The distance between the results for explanation in the state statutory corpus and the technical science research corpus is particularly concerning. State legislatures are moving to require vendors to create or otherwise provide explanations for AI system outputs without any reference to the issues that those building AI systems actively struggle to resolve. What kind of explanation would satisfy the legal need for an explanation? How can the technical explanation translate into something sufficiently useful for justifying decision-making, providing notice, and fulfilling values of due process?

In other words, these laws are written with lawyer-audiences in mind, not with software developer audiences in mind. As such, emerging reforms intended to guide responsible design of socio-technical systems such as AI systems used in the administration of criminal justice and other government services may provide very little technical design guidance at all. In fact, in light of the stark difference in meanings of the terms between law and more technical disciplines, the use of terms like transparency, accountability, explanation and fairness, without further elaboration of their intended technical import may create rules that make responsible design more difficult rather than less.

D. *Judges*

To consider the context in which judges use the terms transparency, accountability, explanation and fairness, I first used the collocation function of the Corpus of US Caselaw (COUSC).[122] The COUSC aims to "expand[] public access to U.S. law."[123] Eventually, it hopes to make all published U.S.

---

[120] *See* Table 1 and Table 2, *supra*.

[121] *See* Table 3, *supra*.

[122] Corpus of U.S. Caselaw (COUSC), BYU LAW, https://lawcorpus.byu.edu/ [https://perma.cc/7FPF-B3Q4]. Note the following frequency of the collocation of these terms: transparency: 0; accountability: 32; explanation: 432; and fairness: 376.

[123] *Id*.

court opinions available online for free.[124] For now, however, it holds all the published U.S. court cases from 1760-1799. Since language changes over time, and since the years at issue in COUSC are quite a long time ago, I constructed an additional corpus of cases in the last 10 years that hit on one of the four AI reform terms.[125]

The first interesting aspect of these results lies in the relative scarcity of collocates for transparency and accountability in judicial decisions. The query returns no hits for transparency, and only 32 for accountability in the older cases found in the COUSC corpus, and while the terms appeared more frequently in the corpus of more recent cases, both terms appeared only about one-fourth as frequently as the terms explanation and fairness. Further, accountability and transparency do not appear together or in association with each other in COUSC, and act as collocates 35 times in the corpus of more recent cases. Given the high frequency of appearance of these terms in federal and state legislation, including the frequency in which those terms are collocated in those texts, obtaining this result from judicial decisions begs for further investigation. Indeed, even the collocates that do appear most frequently with the four terms suggest a very different emphasis from that which appears in the context of legal academic and state statutory discussion of these terms in connection with AI systems. Perhaps, at the very least, these results suggest that the legislative and policy preoccupation with these principles have not yet made their way into litigated disputes. Perhaps that result obtains from a practice of settlement or private arbitration. Ultimately this study only reveals part of the story. This analysis does, however, certainly signal that a story exists worth exploring.

Table 6: Collocates of AI Reform Terms in US Case Law – CUSC

Accountability		**Explanation**		**Fairness**	
2	office	4	given	11	justice
2	health	3	before	6	trial
2	portability	3	comment	4	enquire
1	approved	3	construction	4	fraud
1	fairness	3	first	4	transaction
1	explicability	3	give	3	defendant
		3	principles	2	acquit
		3	testaments	2	agreement

[124] *Id.*

[125] The corpus remains on file with the author. Note the following frequency of the collocations of the terms among the more recent cases: transparency: 552; accountability: 260; explanation: 3,007; and fairness: 1,797.

Table 7: Collocates of AI Reform Terms in US Case Law – Last 10 Years

Transparency		Accountability		Explanation		Fairness	
55	lack	42	act	461	provide(d)(s)	536	entire
36	price	35	transparency	341	how	345	standard
35	accountability	26	health	280	why	175	hearing
35	rule	25	portability	279	any	141	court
26	requirements	24	insurance	275	without	124	see
22	public	22	project	174	offer(ed)(s)	122	plaintiffs
21	full	19	government	112	plaintiff	117	procedural
19	online	18	department	103	further	104	settlement
17	process	16	measure	103	evidence	102	class
17	about	16	office	96	what	89	merger
17	system	16	murder	92	decision	89	review
16	greater	16	HIPAA	76	detailed	86	action

E. *General Public*

Performing a collocate analysis on the News on the Web (NOW) corpus offers a sense of how the general public contextualizes the words transparency, accountability, explanation, and fairness. "The NOW corpus . . . contains 15.0 billion words of data from web-based newspapers and magazines from 2010 to the present time."[126] The general public, like federal legislation, closely associate transparency and accountability. Interestingly, the general public, like state legislation, also associates fairness with transparency. The public also generally seems to feel the need to "ensure," "commit to," "bring," and "increase" transparency, suggesting that the public feels these characteristics remain lacking at the present. The target of these efforts includes "government," "governance," and "process." Likewise, the public targets "national" "government," "bureaus," "courts," the "police," "government office," and "governance" as institutions in need of accountability. Importantly, the general public seems to think these entities could benefit from further transparency even before they adopt AI systems.

In the context of accountability, the words "lack," "ensure," and "demand," like the words frequently used in context of transparency, signal that the public feels a need for heightened accountability. In terms of explanation, the most frequently associated word is "why." And the "why" should be "given," "offered," and "provided." The words that may signal expected characteristics of explanations include "possible," "simple," "detailed," and "plausible." And, once again, words like "more," "demand," and "need" indicate that the public places a heightened emphasis on receiving explanations for "why" government entities make decisions and take action.

[126] NOW Corpus, *supra* note 108.

Lastly, the pubic tends to associate fairness with "justice," "equity," "honesty," "independence," "neutrality," "objectivity," "equality," "respect," and "integrity."

Table 8: AI Systems Reform Terms in the General Public – NOW

Transparency		Accountability		Explanation		Fairness	
44201	accountability	44207	transparency	21449	why	12797	justice
24517	lack	21211	national	19372	give	10923	transparency
18015	ensure	19071	bureau	15016	offer	8238	equity
17596	great	16384	court	12818	provide	7180	honesty
17358	more	16015	government	8758	possible	6627	independence
14398	international	14706	NAB	6319	some	6516	neutrality
12332	report	11968	ensure	6081	about	4950	ensure
10740	bring	10055	public	5581	simple	4020	objectivity
10682	fairness	9413	lack	5397	other	3796	equality
10421	government	9123	police	5158	more	3195	balanced
10176	commit	8815	office	5002	demand	3170	reports
9548	provide	7019	demand	4753	detailed	3132	woven
8729	process	6754	governance	4726	only	2673	tax
7959	governance	6172	great	4214	need	2503	respect
7723	independence	5847	justice	3814	find	2430	sense
7551	increase	5641	take	3794	plausible	2232	integrity

To confirm these results reflect popular usage in the broadest possible sense, a collocate analysis of the four terms was also performed on the Corpus of Contemporary American English (COCA).[127] COCA is a genre-balanced corpus of American English, meaning that its text is pulled from eight genres of language usage: "spoken, fiction, popular magazines, newspapers, academic texts, and TV and movies subtitles, blogs and other web pages."[128] As detailed in Table 9, below, as with the collocate analysis of the NOW corpus, the COCA results suggest that the general public closely associates accountability and transparency (like federal legislation), and that the general focus of those two terms remains on government and public administration. The general desire from the public that transparency and accountability increase in these institutions also remains evident in the COCA results. Again, it seems the general public believes government institutions could benefit from greater transparency and accountability even without considering any of the concerns that legal academics raise related to increased opacity from use of AI systems.

In the context of explanation, the most frequently associated word in COCA remains the same as in NOW—"why." The "why" should be "offer[ed]," "provide[ed]," "simple," "plausible," "and yet also "scientific," "detailed," and "logical." The emphasis on "require[ing]" explanation

[127] Corpus of Contemporary American English, *supra* note 109.
[128] *Id*. The corpus contains more than one billion words of text. *Id*.

confirms the earlier assessment that the general public places high importance on receiving they "why" for decisions that affect them. Finally, the COCA results for fairness confirm the snapshot of public use of that term provided by the NOW results. The public connects fairness to "justice," "equality," "perception," "accuracy," "equity," and "process," and views fairness as "basic" and "fundamental."

Table 9: AI Reform Terms in the General Public – COCA

Transparency		Accountability		Explanation		Fairness	
567	accountability	851	government	2153	why	381	justice
485	lack	386	transparency	1655	possible	260	sense
340	great	324	lack	1581	offer	234	issue
289	government	276	more	1295	provide	185	doctrine
195	international	275	about	915	simple	272	equality
133	process	240	public	736	only	162	tax
133	provide	216	greater	598	alternative	146	act
133	increase	193	some	549	plausible	125	question
120	openness	187	responsibility	486	scientific	124	perception
119	full	169	school	452	detailed	112	accuracy
113	public	152	demand	429	logical	109	equity
107	promote	133	standards	415	likely	105	process
101	level	133	state	393	require	104	basic
96	information	131	democratic	381	far	104	fundamental
88	corruption	115	health	373	rational	96	ensure
84	financial	114	personal	352	reasonable	95	economic

The results from NOW and COCA suggest that public perception of the four terms is shaped heavily by the legal values embedded in those terms. Perhaps most tellingly, nowhere do the results from NOW or COCA suggest that the public worries at all about the impact of AI systems on transparency, accountability, explanation, or fairness. The terms AI, artificial intelligence, algorithm, data, privacy, and tracking do not appear in the collocation results here the way that they do in the legal academic and state legislative corpora. Determining why these words do not appear in the NOW and COCA collocation results lies beyond the capacity of the corpus linguistic approach. As with the absence of terms from the corpora of U.S. cases, this study can only provide part of the story. It could be that the public, broadly speaking, is only tangentially aware of the risks posed by AI systems. Or, perhaps public awareness of the issues exists, but apathy reigns. In either case, many of the issues that preoccupy both legal academics, state legislatures, and those seeking to build responsible AI systems do not appear to weigh as heavily on the minds of the public.

F. *Lessons Suggested from Comparing the Results*

The results from the corpora reflect five stakeholders in the discussion around the use of AI systems in areas of consequential legal decision making, such as criminal justice. Comparing these results leads to two interrelated conclusions. First, parties engaging in discussions at the intersection of AI systems and criminal justice use the same words but have different terms of art in mind. Legal academics and lawmakers tend to take the general legal doctrinal meaning of the terms transparency, accountability, explanation, and fairness, and port them over to discussions about AI systems and the use of such systems in the administration of the criminal justice systems and other areas of public administration.[129] Meanwhile, researchers from within the disciplines with the technical expertise to build AI systems remain preoccupied with entirely different questions inherent in their understanding of the terms transparency, accountability, fairness, and explanation as they apply in the context of software development.[130]

Second, these terminology conflicts represent more than mere definitional differences. Rather, the collision between equally valid terms of art represents core value conflicts that need to be reconciled before effective law-making can truly be undertaken at the intersection of AI systems and criminal justice. Language is "always laden with value judgments and carrying attitudes."[131] Part of the miscommunication occurring at the intersection of AI systems, criminal justice, and law centers on the different values embedded in the way each discipline uses the terms explainability, transparency, accountability, and fairness. To move the law forward in the incredibly important arena of AI systems and criminal justice, lawmakers and lawyers must do more than merely co-opt technical terms and ascribe to them legal values. Rather, law must systematically seek to reconcile the value conflict inherent in interdisciplinary language clashes.[132]

Last, part of the difficulty in understanding the differences in linguistic meaning may stem from a lack of incentive to do so. The general public's use of the terms transparency, accountability, explanation, and fairness reflects many of the same understandings and embedded values as the uses evidenced

---

[129] *See* Table 1 and Table 2, *supra*.

[130] *See* Table 3, *supra*.

[131] Clarke Rountree, *Kairos and American Legal Praxis*, 20 NEV. L.J. 855, 863 (2020) (citing RICHARD M. WEAVER, LANGUAGE IS SERMONIC: RICHARD M. WEAVER ON THE NATURE OF RHETORIC 18 (Richard L. Johannesen, Rennard Strickland & Ralph T. Eubanks eds., 1970)).

[132] For a proposed approach to designing legal systems that undertake such reconciliation, see the proposal for Algorithmic Systems Query (ASQ) in Reyes & Ward, *supra* note 8.

by the legal-academic and lawmaker corpora.[133] Further, the two terms that frequently arise in the judicial context—explanation and fairness—evidence strong meaning and value alignment between judges, legal academics, and lawmakers, but a significant disconnect from the meaning and values attached to those terms by those building AI systems. When everyone else seems to find alignment on the definitions and embedded values for (at least two) of the terms, why shouldn't the law just make the technical sciences come into alignment? Why go to the trouble of reconciling definitions and value conflict between the creator of AI systems and everyone else—systematically or otherwise—when the technical sciences seem to be the only stakeholders with a different view? The simple answer lies in the limits of technology. The development of AI systems must follow certain rules, if the AI system will work at all,[134] let alone work for a particular nuanced purpose.[135] The preoccupation with AI system constraints, models, methods, data, accuracy, etc., should not be misconstrued as failure to be concerned with values such as non-discrimination, equity, notice, etc. Rather, they reflect the technical reality that frames the possibilities for AI system design and use. In that regard, those developing new law to address real concerns regarding bias, discrimination, and privacy in connection with the application of AI systems in the criminal justice system and in other areas of public administration must seek a deeper understanding of AI systems and approach the inquiry with a framework that helps bridge the communication gap.

CONCLUSION

This Article presented linguistic evidence confirming the anecdotal intuition that interdisciplinary miscommunication impacts the current efforts

---

[133] *See* Table 1, Table 2, Table 4 and Table 5, *supra*.

[134] "Deployed AI systems often do not work. They can be constructed haphazardly, deployed indiscriminately, and promoted deceptively. However, despite this reality, scholars, the press, and policymakers pay too little attention to functionality. This leads to technical and policy solutions focused on 'ethical' or value-aligned deployments, often skipping over the prior question of whether a given system functions, or provides any benefits at all." Inioluwa Deborah Raji, Aaron Horowitz, I. Elizabeth Kumar & Andrew D. Selbst, *The Fallacy of AI Functionality*, 2022 ACM CONF. ON FAIRNESS, ACCOUNTABILITY, & TRANSPARENCY 959, 959, https://doi.org/10.1145/3531146.3533158 [https://perma.cc/DGK9-N78B].

[135] *See, e.g.*, Sebastian Bordt, Michele Finck, Eric Raidl & Ulrike von Luxburg, *Post-Hoc Explanations Fail to Achieve Their Purpose in Adversarial Contexts*, 2022 ACM CONF. ON FAIRNESS, ACCOUNTABILITY, & TRANSPARENCY 891, 891, https://doi.org/10.1145/3531146.3533153 [https://perma.cc/KB7D-8WHD] ("[P]ost-hoc explanation algorithms are unsuitable to achieve the transparency objectives inherent to the legal norms.").

of states to address known problems with the government's use of AI systems in the administration of criminal justice. Importantly, the evidence goes further, demonstrating that the miscommunication does not stem solely from definitional differences, but also from differences in the underlying values imbued in the core terms accountability, transparency, explainability, and fairness. Armed with such knowledge, law and policy-makers might consider devising new processes in the law-making and rule-making process beyond changes to procurement law. In particular, legislative bodies and rule-making authorities might engage in a formal value-alignment analysis, considering the needs of the legal system alongside the needs of the technical system and finding reconciliation mechanisms that can be reflected in the law-making process.[136]

Other communities participating in policy discussions at the intersection of AI systems and criminal justice might also make use of this linguistic study. For example, lawyers practicing in this area, whether as a prosecutor, defense attorney, or counsel to the companies developing the AI system might now understand their duty of technological competence[137] to include a deep dive into technical definitions of otherwise commonly used terms, in order to understand related value conflicts and operational impacts when engaged in client consultation. Attorneys often balk at the idea that their duties might include a deep-dive into the inner-workings of socio-technical systems, however, when those systems are employed to make consequential decisions touching on issues of life and liberty, such discomfort must be set aside. Unless the miscommunication and value conflict evidenced by this study can be resolved, further efforts to mitigate harms from the use of AI systems in the administration of criminal justice may compound known problems or surface new problems. Law-making in an area of high interdisciplinarity must move forward carefully and intentionally, keeping language wars front and center in the quest to communicate understandable and useful legal rules to a diverse set of recipients.

---

[136] See, for example, the analysis tool ASQ as described in Reyes & Ward, *supra* note 8.

[137] *See* MODEL RULES OF PRO. CONDUCT r. 1.1 cmt. 8 ("To maintain the requisite knowledge and skill, lawyers should keep abreast of changes in the law and its practice, including the benefits and risks associated with relevant technology."); Agnieszka McPeak, *Disruptive Technology and the Ethical Lawyer*, 50 U. TOLEDO L. REV. 457, 468–470 (2019) (discussing lawyers' duty to understand technology and its ethical implications when working with clients).