INTERNATIONAL COMPUTATIONAL LAW JOURNAL

CODEX SOURCE ARCHIVE

May it Please the Bot?

Jameson Dempsey, Gabriel Teninbaum

Computational Law Report · Release 2.0 · 2020-08-14

Reader edition prepared by The Research University under CC BY 4.0. This cover was added; the original PDF pages follow unchanged. Original authorship, publication identity and source links are preserved. This independent republication is not sponsored or endorsed by Stanford University, Stanford Law School, CodeX or MIT.

Volume 1 · Issue 2 · Systems, Rules and Public Infrastructure

Original source and license evidence are recorded on the article landing page.

MIT Computational Law Report

May it Please the Bot?

Jameson Dempsey, Gabriel Teninbaum

Published on: Aug 14, 2020

Updated on: Aug 10, 2020

License: Creative Commons Attribution 4.0 International License (CC-BY 4.0)

0. Introduction

Imagine a system where appellate attorneys could be virtually assured of how a panel of judges would decide what otherwise appears to be a close case. Imagine how the work of an overburdened legal aid attorney could be lightened if she could accurately predict the exact amount of fees that a judge would award her client in an eviction case. Imagine the benefits to a law firm that could have paralegals write the first drafts of important motions, citing only the cases preferred by the local magistrate. Finally, imagine how the work of writing judicial opinions would become easier for a judge if she was able to recognize the framework the Supreme Court would use to decide whether her decision should be overturned. These are only some of the benefits we envision for re-thinking the ways that judges write judicial opinions and encouraging the adoption of methods that better utilize modern technologies.

The American judiciary plays an integral role in defining and upholding the law. For decades,[1] the legal informatics community has sought to use information technology to search, analyze, and make predictions based on large corpora of judicial opinions.[2] Unfortunately, while these data-driven technologies have made significant progress, they face a lingering limitation: the language and structure of the opinions themselves. Judicial opinions — particularly appellate decisions —lack common standards for language, structure, and conveying critical information about the decision. Some will see this as a feature, not a bug. However, because judicial decisions come in many forms and styles, it is up to lawyers and courts to tease out essential elements of past decisions: e.g., holdings, tests, relative weights of factors.

In this essay, we argue that judges should write opinions in anticipation of later machine processing, and that in doing so they can increase the efficiency and predictability of the legal system. Below, we lay out our theory for why this is within grasp, identify the challenges we expect along the way, and describe approaches we envision.

1. What is a judicial opinion?

A judicial opinion is a document that "informs parties of the outcome of their case and articulates the legal principles on which the decision is based in order to guide the bench, the bar, academia, and the public."[3]

Opinions take a range of forms — from short to long, cursory to thorough, unpublished or published — depending on the case and the court. For example, "full-dress opinions" tend to be longer, more thorough, and published (other options are procedural or otherwise brief in nature and tend to focus on documenting the outcome dictated by the court, as opposed to explaining the legal basis for it in detail). The Judicial Writing Manual suggests that a full-dress opinion should contain the following elements:

1. An introductory statement of the nature, procedural posture, and result of the case;
2. A statement of the issues to be decided;
3. A statement of the material facts;
4. A discussion of the governing legal principles and resolution of the issues; and
5. The disposition and necessary instructions[4]

It further suggests including a "logical organization of the opinion," including "headings, subheadings, and subdivisions" to orient the reader and enable fellow judges to join or dissent from particular parts of the opinion.[5] Most important to our discussion, this type of organization "assist[s] in the indexing and classification of opinions and their retrieval by researchers."[6] In light of the prevalence and necessity of data-driven legal research software in modern litigation, we argue that judicial opinions can and should do more to assist researchers and subsequent courts.

2. Why do we have judicial opinions at all?

The role of appellate courts is to review decisions of trial judges and to determine whether the trial judge applied the law correctly in the case at hand. To memorialize the analysis, judges write an explanation. This is a judicial decision. In the appellate context, these decisions often include instructions and analytical tests — frameworks for future decision-making — that lower courts are bound to import and apply in their own cases.

A precept of American law is that the law applies to everyone within the jurisdiction creating or interpreting it. However, it is not always clear what the law *is* in a given situation.[7] And, when deciding an appellate case, it is the role of judges to interpret ambiguities in the law so that the community can understand what behavior is allowed or disallowed. For this reason, when a judge writes an opinion, simply deciding in favor of one party or another is not enough when the law is ambiguous; rather, it is the role of the judge to lay out the reasons underlying their decision so that citizens are able to regulate their conduct according to the law.

How judges actually resolve ambiguities in law can, in practice, be messy. Even for judges working faithfully to describe their analysis of the law, it is not always easy to decide the outcome when existing statutes, regulations, or cases do not squarely address the issues at hand. At times, the language of a statute may appear to conflict with a regulation; or a case decided long ago under similar facts might appear to be at odds with modern culture. This judicial freedom can sometimes lead to unclear, inconsistent outcomes and further perpetuate a cynical view that judges are justifying their decisions without a clear approach that would be instructive to future litigants.

One approach to understanding this quandary is by thinking of legal reasoning through a pluralistic lens.[8] In short, this view suggests that any time an appellate body decides a case, it has a variety of

sources of law it can look to in order to reach a decision. Some judges might favor a given form of argument over another, but there exists a select menu of ways they can reach a decision.[9]

Although there is debate among legal scholars about exactly how to define the ways judges decide, for the purposes of this Essay, we accept Prof. Wilson Huhn's view that any appellate decision is based on at least one of the following: the law's text, intent, precedent, tradition, or policy analysis:

- *Text* looks to the language used in a legal document to glean its meaning (by, for example, looking to its plain meaning or intratextual arguments).[10]
- *Intent* looks at the intent of the drafters of a law to decide how they would want it applied.[11]
- *Precedent* requires judges to examine prior decisions on similar matters and to decide the case before them in a way that is consistent with similar historical cases.[12]
- *Tradition* is a source of law that looks to social custom and is influenced by the way that people, organizations, and governments have historically behaved.[13]
- *Policy* arguments make a prediction about the consequences that will flow from a potential decision, leaving the judge to decide the matter based on which predicted outcome is most consistent with the competing value deemed most important.[14]

Within this system, judges (or judicial systems[15]) might express individual preferences. For example, some judges might prefer to interpret ambiguity in the law based on a statute's plain language (a *text* argument in Huhn's taxonomy), while others might see law as mainly an exercise in gleaning a statute's intent. Some judges might decide in an ad hoc way, sometimes preferring intent to text, or vice versa. Still others might rely on precedent over other forms of argument and, even within the precedent form of argument, might give special weight to similar decisions by a specific judge or jurisdiction.

Currently, those who read legal opinions as a tool to predict future decisions are left relying on gut instinct. There may be anecdotal beliefs about how a given judge, court, or jurisdiction is likely to decide a case based on their past behavior, but it's seldom based on more. Courts develop reputations - "the 9th Circuit is liberal," or "Judge Smith focuses on policy," but there is often little empirical evidence to justify such assertions.

3. Why should judges write structured, machine-readable decisions?

Scholars have recognized the need for computers to enable lawyers to perform effective legal analysis since the early days of modern computing.[16] Today, effective legal research—particularly in litigation—is virtually impossible without accessing and using online case law databases. And even if analog legal research were feasible, evolving legal ethical obligations, including the duty of technology

competence,[17] make it unwise to entirely avoid using technological assistance to practice law. Some judges have limited fees for attorneys who failed to use modern computer-based legal research tools, and other commentators have asked whether the failure to use legal analytics is malpractice.[18]

But what about the judges? Are their decisions making clear the legal bases for their conclusions within the narrative? The Judicial Writing Manual counsels judges to avoid wordiness; lack of precision and clarity; poor organization; cryptic analysis; and pomposity and humor.[19] The Manual further suggests judicial writing eliminate unnecessary words, be succinct and direct, use plain English, use footnotes and citations, and be carefully edited.[20] And yet, retired Judge Richard Posner notes that precepts of "good" writing and structure often give way to personal writing style.[21] Consequently, there is still a wealth of "judicial gobbledygook."[22] Moreover, even where the style is clear, the opinions still fall short because they are viewed as works of prose, rather than data meant for computer processing. As a result, lawyers and legal technologists must fill the gaps, with mixed results.

Take citators as one example. Historically, case law publishers have relied on human editors to review cases and make determinations about whether a case has been overruled, distinguished, or otherwise negatively treated. Unfortunately, analyses of these citators have revealed errors within and conflicts between them.[23] In the last several years, legal technology companies have developed algorithms to determine whether a particular case has been negatively treated.[24] While these citators may make the process of flagging documents more efficient, they still rely on the language in the primary material, i.e., the way judges describe what they are doing. If judges use inconsistent or unclear language when describing what they are doing (or what previous judges have done), the citators—along with the lawyers relying on them—suffer.

Judicial tests provide another example. Judges throughout American history have adopted analytical frameworks to justify their opinions and help guide future courts in their decision making. In some cases, these frameworks are required by statute. In others, judges have developed their own multipart tests, which function as rules for future courts to adopt when conducting their analyses. While some have criticized these multipart tests as complex, confusing, misleading, or improper,[25] others have viewed them more favorably.[26] Whether the rules are statutorily prescribed or judicially invented, they are a core feature of judicial opinions that inform not only judicial decision-making itself, but also how attorneys research and structure their arguments. Unfortunately, there is no standard for how to structure judicial tests, making computer-assisted comparison across cases difficult. As a result, attorneys must either conduct a detailed analysis and comparison of cases or rely on secondary sources such as proprietary treatises.

By changing the way these decisions are written, there is an opportunity to dramatically improve the level of clarity and predictability within the judicial process. In short, it would increase the ability of

those governed by the decisions to understand and make predictions about new cases, as well as speed up research and lower legal costs.

4. How might a judge make her opinions more machine readable?

Marshall McLuhan famously stated, "the medium is the message."[27] It stands to reason, then, that when representing the information of the law certain mediums will be more amenable to certain outcomes. Extending this further and thinking about opportunities to address existing problems of transparency, accessibility, and predictability, machine readable judicial opinions would appear to be a step in this direction. This section seeks to answer the question, if judges should draft opinions in anticipation of computer processing, what would those opinions look like?

a. Consistent, simplified language

Perhaps the most straightforward way for judges to increase the machine readability of their opinions is to strip opinions of legalese.[28] For example, judges could score opinions based on readability before publishing them[29] Alternatively, imagine an "autocorrect" for legalese that suggests plain language replacements for common turns of legalese (a Garnerizer).[30] Or judges could explicitly identify which legal precedents or facts were persuasive or unpersuasive using a common lexicon.[31] In fact, legal research tools have had success identifying fundamental legal principles by searching for the phrase "it is well settled."[32] Using more consistent and simplified language could improve machine readability by making it easier for general-purpose natural language processing models trained on non-legalese text (such as BERT) to understand legal opinions.[33]

b. Common Structures

Beyond language, requiring judges to adopt consistent structures for opinions would enhance their amenability to machine processing. For example, courts could adopt local rules and templates for specific types of opinions that constrain the ordering, numbering, and naming of sections, with enough flexibility to accommodate complex cases. For simple and routine cases, these common structures could be simple online forms or question-and-answer prompts, with judges deciding which boxes to tick.

c. Opinion Markup

In addition to writing clearer opinions and distilling key information in the opinion in a structured appendix, judges could tag the key elements of their opinions, treating facts, issues, cited cases, or other elements of the case as pieces of data. Judges could leverage the significant work that has already been done to develop ontologies and markup standards for legal documents.[34] This approach would promote machine readability while retaining flexibility for judges to write expressively and provide context. In practice, during or after drafting the opinion, the judge or her clerk could insert

tags for relevant information within the opinion (e.g., parties, holding, citations, etc.). In theory, the courts could employ a common WYSIWYG platform or integrated development environment for law that would enable judges to quickly markup their opinions using shortcuts.[35] This could even be done in a hybrid form, where judges wrote a decision as an unstructured text, and clerks (or other court personnel) then attached a separate document that created a structured summary, identifying the key elements to the decision.

d. Computable Judicial Tests

Perhaps judges could take things one step further, making certain elements of the opinion—such as judicial tests—"computable." Appellate judges in the US create new tests all the time, and judges are often called to apply tests within statutes in their judicial analyses. As stated above, these tests often function as rules, with varying levels of prescriptiveness. How might a computable judicial opinion work? As one example, when establishing a new test, an appellate court could issue a question-and-answer workflow or a form using an open source tool like docassemble. A judge applying the test would complete the workflow to reach a result, including relevant facts and narrative explanation as necessary. As more lower-court cases rely on the test, attorneys and legal researchers could perform legal and data analyses across cases. A requirement to more explicitly define legal tests could improve the quality of those tests, encouraging judges to avoid vague standards and balancing tests in favor of explicit rules. It could also serve as a pre-check for future appellate courts faced with questions whether lower courts fully engaged with required analyses.

5. What are the limitations of machine readable judicial opinions?

While we have argued here that there would be considerable benefits to making decisions computable, there are also notable downsides that deserve acknowledgment. We should begin with culture: after all the jurist, Oliver Wendell Holmes, famously said "the life of the law has not been logic; it has been experience." Beyond the fact that this is not a promising worldview if we are to influence courts to adopt this approach,[36] there are plenty of reasons to agree with Holmes.

The narrative form of decision-making has been a tradition of American courts since the nation's founding (and, before that, an important part of English judicial opinion-writing). In many instances, the language of judges is nothing less than prose, with famous lines from decisions serving as inspiration for the society they govern. This aspect of our legal culture, a broader cultural context, may be lost by the adoption of some of the more restrictive formats of writing that have been proposed in this article.

Narrative, of course, plays another important role in the context of judicial opinions: it makes things easier to follow. If we were to reduce judicial decisions to a fill-in-the-box approach, or anything like

it, important context about the case at hand would be lost, making it more difficult to read decisions with focus and to follow the actions they analyze.

Moreover, there are also a number of practical challenges with computable judicial opinions. While many of the suggestions above could be accomplished using existing tools and resources, to function well, computable judicial opinions would require at least (1) a common ontology, processes, and tools to ensure consistency across courts and cases, (2) new technical infrastructure to host and analyze computable judicial opinions, and (3) initial and ongoing training for judges and clerks to ensure they are using the tools correctly.

Additionally, there is a strong possibility that normative challenges arise from the use of computable judicial opinions. Poorly written or biased judicial tests could wreak havoc on lower court decisions without the opportunity for lower-court judges to deviate from unfair results, a criticism that many have leveled against criminal sentencing guidelines.

Finally, there is also the possibility that our ideas to use various technologies to make judicial opinions easier to analyze and predict will seem quaint and antiquated in just a short time: an efficiency expert in the 1950s arguing that secretaries should start using electric typewriters to save time would certainly look quaint to those who came of age in the era of powerful computers and smartphones. It is entirely possible that the approach proposed here could be leapfrogged in a similar way by the tools of the future in a relatively small window of time. Rather than point to one specific technology, we instead intend to raise a question about what efforts will be made to develop methods for parsing judicial opinions that face the challenges we now face. What such tools might look like remains to be seen.

6. A Call To Action

The legal field has been famously hostile toward change, with adoption of digital transformation being a particular sticking point; this in spite of clear evidence from consulting firms that doing so is an integral strategic priority and improved efficiency.[37] In some ways, this is why we believe such an opportunity exists for data scientists, web developers, and others familiar with the tools that can help realize this idea.

In an era in which U.S. jurisdictions are creating "regulatory sandboxes"[38] that leverage the use of new tools and methods to encourage better functioning legal systems, technologists should advance the opportunity to encourage new approaches to judicial decisionmaking. We do not propose a single framework (though have identified several possibilities, above) simply because this is an idea that would benefit from widespread testing using different approaches. No doubt, some approaches will be more efficient and effective than others and comparing outcomes among various approaches will help us to determine which will fit best.

Ultimately, we think it worthwhile for jurisdictions to boldly attempt this, and to do so with an attitude of experimentation and openness to the future. The prize for those who succeed in this endeavor will be a more open, accessible understanding of the application of the laws that guide society.

Gabe Teninbaum is the Assistant Dean for Innovation, Strategic Initiatives, & Distance Education, as well as a Professor of Legal Writing, at Suffolk University Law School in Boston, Massachusetts. He can be reached at gteninbaum@suffolk.edu.

Jameson is a technology attorney and community builder based in the Bay Area. He currently serves as a fellow at CodeX – The Stanford Center for Legal Informatics and on the board of Legal Hackers.

The authors would like to thank Dyane O'Leary, Pablo Arredondo, Adam Ziegler, Daniel Hoadley, Dazza Greenwood, Brian Ulicny, and Bryan Wilson for helpful feedback.

Footnotes

1. In the early 19th century, Simon Greenleaf lost a case relying on precedent that, unbeknownst to him, had been recently overruled. Dabney, Laura C., "Citators: Past, Present, and Future" (2008). Articles by Maurer Faculty. 2113. available at: https://www.repository.law.indiana.edu/facpub/2113.

   Greenleaf saw a need, and decided to try and fill it by creating an alphabetical list of overruled cases. Id. In 1873, the "Shepard's" product was introduced to fill a similar function, the offspring of which still are used in law practice today. Id. ↩
2. Judicial decisions, also called "opinions," are written records created by judges to explain the basis for the outcomes of cases. ↩
3. Federal Judicial Center, "Judicial Writing Manual: A Pocket Guide for Judges" ("Judicial Writing Manual") pp. 3 (2013). ↩
4. *Id.* at 13. ↩
5. *Id.* ↩
6. *Id.* ↩
7. For example: a City statute states that "any sandwich vendor operating in City Park must be licensed with the City to sell sandwiches" or the vendor faces a $100 fine. No doubt, someone selling grilled cheese or peanut butter and jelly on white bread within City Park should seek a license. But what of the hot dog vendor? The taco truck? Are *those* sandwiches? Here, if the hot dog vendor were

to be issued a citation, she could appeal the citation to a trial judge by arguing hot dogs are not sandwiches. The trial judge would be expected to look to the language of the statute, and to any other similar cases interpreting the statute, to decide whether to dismiss the citation or uphold it. If the trial judge decided against the hot dog vendor, she could appeal the decision. The appellate judge's role would be to decide if the trial judge applied the law correctly or not. Traditionally, the appellate judge would write an opinion explaining the basis for that decision, together with instructions to trial judges (and by extension, the community) as to how to handle similar matters in the future). And, for every future burrito-maker, ice cream sandwich stand, and empanada stall owner, it is the role of attorneys to advise them on the likelihood of their being fined based on the language of the law and the court's subsequent interpretations of it. ↩

8. Wilson R. Huhn, Teaching Legal Analysis Using a Pluralistic Model of Law, 36 Gonz. L. Rev. 433, 436 (2001). ↩

9. Looking forward, the ability to identify the form of legal argument a judge applies is one way to get a toe-hold in creating computational judicial decisions. *See* Section 4 *infra*. ↩

10. Huhn, Note 8 infra at 441. ↩

11. Id. at 443. ↩

12. Id. at 444. ↩

13. Id. at 445 ↩

14. Id. at 448-49. ↩

15. For example, continental European legal systems typically give very little weight to precedent, but rely heavily on the text and intent of statutory law. ↩

16. William B. Eldridge and Sally F. Dennis, The Computer As A Tool For Legal Research, 28 Law & Contemp. Prob. 78, 78 (1963). Available at: https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=2949&context=lcp. ↩

17. Model Rule of Professional Conduct 1.1 Comment 8. ↩

18. Robert Ambrogi, "This Tech Can Turn The Tables in Litigation" Above the Law, December 3, 2018. Available at: https://abovethelaw.com/2018/12/this-tech-can-turn-the-tables-in-litigation/ ↩

19. Judicial Writing Manual *infra* 3 at Section V. ↩

20. *Id.* at 25. ↩

21. Richard A. Posner, "Judges' Writing Styles (And Do They Matter?)," 62 U. of Chicago L. Rev. 1420 (1995). ↩

22. *See* Whalen, Ryan, Judicial Gobbledygook: The Readability of Supreme Court Writing, Yale L. J. Forum (2015). ↩

23. Aaron S. Kirschenfeld, "Yellow Flag Fever: Describing Negative Legal Precedent in Citators" (2016). Library Staff Productions. 108. Available at: https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=1020&context=faculty_publications; Hellyer, Paul, "Evaluating Shepard's, KeyCite, and Bcite for Case Validation Accuracy" (2018). Library Staff Publications. 131. Available at: https://scholarship.law.wm.edu/libpubs/131. ↩

24. Jake Heller, "You're probably citing bad law -- here's how to avoid that" Casetext, May 23, 2019. Available at: https://casetext.com/blog/how-to-avoid-citing-bad-law/; https://www.fastcase.com/blog/badlawbot/. ↩

25. Stephen M. Johnson, The Changing Discourse of the Supreme Court, 12 U.N.H. L. REV. 29 (2014), available at http://scholars.unh.edu/unh_lr/vol12/iss1/4. ↩

26. Frederick Schauer (1995) "Opinions as Rules," *University of Chicago Law Review*: Vol. 62 : Iss. 4 , Article 11.

Available at: https://chicagounbound.uchicago.edu/uclrev/vol62/iss4/11. ↩

27. See e.g., Marshall McLuhan, "Understanding Media: The Extensions of Man" MIT Press (1994). Available at: https://mitpress.mit.edu/books/understanding-media. ↩

28. *See e.g.,* Resources at PlainLanguage.gov. Available at: https://www.plainlanguage.gov/resources/content-types/legal-profession/ ↩

29. *Id.* ↩

30. For ways to eliminate legalese from legal writing, see Bryan Garner, "Ax these terms from your legal writing" ABA Journal, April 1, 2014. Available at: http://www.abajournal.com/magazine/article/ax_these_terms_from_your_legal_writing. ↩

31. There are already conventions in this respect that legal technologists have identified, e.g. the fastcase Bad Law Bot. Available at: https://www.fastcase.com/blog/badlawbot/ ↩

32. *See* Valerie McConnell, "What is the Black Letter Law section in search results" Casetext. Available at: https://help.casetext.com/content-on-casetext/what-is-the-black-letter-law-section-

in-search-results ↩

33. *See* Emad Elwany, Dave Moore, Guarav Oberoi, "BERT Goes to Law School: Quantifying the Competitive Advantage of Access to Large Legal Corpora in Contract Understanding" Available at: arXiv:1911.00473; *See also* Laura Safdie, Khalid A. Al-Kofahi, Daniel Hoadley, Anne Tucker, "Is Law's Moat Evaporating?" CodeX Future Law Conference (2020). Available at:

https://conferences.law.stanford.edu/futurelaw2020/sessions/bert-and-the-future-of-legal-analytics/ podcast beginning at 31:20 (Dan Hoadley: "[T]to get good results on legal text, the model does really need to be trained from scratch on legal text, so the BERT based model, which was trained on general web language doesn't really get the job done too well. . . .") ↩

34. *See* Ceci, Marcello & Palmirani, Monica. (2011). Ontology Framework for Judgment Modelling. 116-130. 10.1007/978-3-642-35731-2_8; *See also* LegalXML, http://www.legalxml.org/. ↩

35. Michael Jeffery, "What Would an Integrated Development Environment for Law look like?" MIT Computational Law Report (2020). Retrieved from https://law.mit.edu/pub/whatwouldanintegrateddevelopmentenvironmentforlawlooklike. ↩

36. The Supreme Court thinks highly enough of its slow and steady approach that it has - literally - included images of the tortoise and the hare in the frieze over the Court's entrance. ↩

37. *See, e.g.*, Mark A. Cohen, "Law is Lagging Digital Transformation - Why it Matters," Forbes, Dec. 20, 2018. ↩

38. *See, e.g.*, Utah's Regulatory Sandbox. ↩