

## RULEMAKING EX MACHINA

*Melissa Mortazavi\**

## INTRODUCTION

Emerging technologies promise to expedite administrative rulemaking by analyzing public input through computerized natural language rather than clunky, old human brains. Moving far beyond software that keyword searches and deduplicates content, natural language processing (as a type of predictive coding) employs artificial intelligence that adapts and modulates depending on inputs, rendering it fluid and dynamic.<sup>1</sup> With the current concerted push to streamline agencies,<sup>2</sup> the question of how and when to use automation in rulemaking will likely be decided in the next year. Considering that recently, a single proposed rule garnered over 3.7 million public comments,<sup>3</sup> mechanisms that can make comprehending those comments “10,000%” faster have intuitive and intoxicating appeal.<sup>4</sup> Even though natural language processing software is in its infancy—its potential to impact the workings of the

---

\* Melissa Mortazavi, Associate Professor, University of Oklahoma College of Law. I’d like to extend my gratitude to Emily Bremer, Evan Way, and above all Roger Michalski.

1. See Nicholas M. Pace & Laura Zakaras, RAND Inst. for Civil Justice, *Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery* 59 (2012) (explaining the distinction between keyword-based near-duplication sorting techniques (such as Boolean searches, clustering, and email threading) and predictive coding that engages in self-learning functions to arrive at a normative assessment of the content of the document); Philip Cohen & Lauren Harrison, *Predictive Coding Is a New Tool in the E-Discovery Toolbox*, N.Y. L.J. (Mar. 19, 2012), <http://www.newyorklawjournal.com/id=1202545983136> (available in full on Lexis Advance and on file with the *Columbia Law Review*) (explaining the distinction between predictive coding and keyword searches).

2. The current administration is aggressively seeking ways to streamline and minimize administrative action. See Exec. Order No. 13,777, 82 Fed. Reg. 12,285 (Feb. 24, 2017) (“It is the policy of the United States to alleviate unnecessary regulatory burdens placed on the American people.”); Exec. Order No. 13,771, 82 Fed. Reg. 9339 (Jan. 30, 2017) (requiring that for every one new regulation issued, at least two prior regulations be identified for elimination).

3. Elise Hu, *3.7 Million Comments Later, Here’s Where Net Neutrality Stands*, NPR (Sept. 17, 2014), <http://www.npr.org/sections/Alltechconsidered/2014/09/17/349243335/3-7-million-comments-later-heres-where-net-neutrality-stands> [<http://perma.cc/V5ZV-N6RN>].

4. Fernando Hurtado, *Did You Submit a Gov’t Complaint Recently? It Likely Wasn’t Read*, Circa (Mar. 3, 2017), <http://circa.com/politics/government/it-would-take-years-for-the-govt-to-read-all-your-comments-on-govt-policies-until-now> [<http://perma.cc/YX5C-NGN2>] (quoting John Davis, Founder of Regendus, a regulatory analytics platform that “pars[es] through public comments”).

administrative state and with it, government policy and programs—is limitless.

But before embracing this high-tech panacea, it is incumbent on policymakers, scholars, and attorneys to consider how implementing such innovations could undermine or enhance existing legal systems. This Piece begins that inquiry by looking to the core of administrative policymaking. Part I will outline the requirements of the Administrative Procedure Act (APA) and specifically notice-and-comment rulemaking. Part II then proceeds to flag key ways that automation can support or hinder the legal exercise of agency action.

Such an analysis does not exist in a vacuum; legal-ethics scholars have grappled for some time with whether algorithms can approximate the work of lawyers. Often juxtaposing ethical considerations and substantive legal skills to the pragmatic needs of dealing with the explosion of e-discovery, the ethics scholarly community has engaged in a measured exploration of coding’s virtues and vices that challenges the idea that predictive coding is an “unmitigated good.”<sup>5</sup> In practice, predictive coding has taken the legal-services market by storm.<sup>6</sup> The Federal Rules of Civil Procedure, the Federal Rules of Evidence, and the American Bar Association Model Rules of Professional Conduct have hastened to respond to the changes brought by a digital age.<sup>7</sup> These ongoing conversations over emerging uses of technology in different legal fields can inform some of the debates over whether computers can do the work of administrative policymakers.

However, even lessons learned from existing civil litigation challenges provide only limited insight into predictive coding’s application to tasks unlike discovery’s finite and predetermined scope. Using automation to expedite rulemaking is fundamentally different from the tasks of client representation. In discovery, parties are better situated to structure

---

5. See, e.g., Dana A. Remus, *The Uncertain Promise of Predictive Coding*, 99 *Iowa L. Rev.* 1691, 1706, 1708–10 (2014) (discussing the potential erosion of core ethical values such as cooperation, the unauthorized practice of law, and court processes); Charles Yablon & Nick Landsman-Roos, *Predictive Coding: Emerging Questions and Concerns*, 64 *S.C. L. Rev.* 633, 637 (2013) (arguing that these technologies cannot supplant lawyers as they are unable to “assemble theories of a case . . . or even decide whether a document is helpful or hurts a particular side’s case”). But cf. Aaron Goodman, *Predictive Coding: A Better Way to Deal with Electronically Stored Information, Litigation*, Fall 2016, at 23 (advocating strongly for the use of predictive coding to ease discovery burdens).

6. See, e.g., Bennett B. Borden & Jason R. Baron, *Finding the Signal in the Noise: Information Governance, Analytics, and the Future of Legal Practice*, *Rich. J.L. & Tech.*, Mar. 14, 2014, at 1, 3–14 (2014) (discussing the history of how predictive coding has been used within the legal field).

7. See, e.g., Fed. R. Civ. P. 37(e) (outlining specific electronic discovery requirements); Fed. R. Evid. 502 (allowing more lenient clawback procedures for inadvertent disclosure in response to the use of predictive coding); Model Rules of Prof’l Conduct r. 1.6(c) cmt. 19 (Am. Bar Ass’n 2015) (discussing the duty of confidentiality, electronic security, and data privacy measures).

data processing because they are searching for specific facts and materials. Notice-and-comment rulemaking is predicated on the unknown—the most valuable and useful comments are those that were previously unanticipated. In applying data-processing technology to administrative rulemaking, the march of computer automation soldiers onward to uncharted and dangerous terrain: not the execution of laws and regulations, but their very creation.<sup>8</sup> This leap is not one to make haphazardly.

### I. THE LEGAL STRUCTURE OF RULEMAKING

The APA binds federal agency action. Passed in 1946, the APA outlines uniform procedural expectations for agency action, allowing the public to better understand and predict how agencies will behave and modify responses accordingly.<sup>9</sup> It has remained extraordinarily static since its passage, given its centrality in administrative law.<sup>10</sup> Born of concerns that the administrative state was increasingly unwieldy and anti-democratic, the procedural norms created through the APA were a political compromise, balancing efficiency and individual rights through public interaction, information gathering, and increased predictability across and within agencies.<sup>11</sup> The APA outlines procedural requirements for formal adjudications,<sup>12</sup> formal rulemaking,<sup>13</sup> and informal rulemaking (known colloquially as “notice-and-comment” rulemaking),<sup>14</sup> and it provides the terms of judicial review of agency action.<sup>15</sup> Formal rulemaking requires extensive additional procedures<sup>16</sup> and is exceedingly rare,<sup>17</sup> as it

---

8. The creation of law is grounded in imagining new systems and new solutions, something that is not necessarily aided by computerized processing patterns. See Laura Pappano, Learning to Think Like a Computer, N.Y. Times (Apr. 4, 2017), <http://www.nytimes.com/2017/04/04/education/edlife/teaching-students-computer-code.html> (on file with the *Columbia Law Review*) (“There is no reliable research showing that computing makes one more creative or more able to problem-solve.”).

9. Pub. L. No. 79-404, 60 Stat. 237 (1946) (codified as amended at 5 U.S.C. §§ 551–559, 701–706 (2012)).

10. William H. Allen, The Durability of the Administrative Procedure Act, 72 Va. L. Rev. 235, 235–48 (1986) (“The conventional wisdom is that, since its enactment . . . the Administrative Procedure Act has been unusually impervious to change.”).

11. See George B. Shepherd, Fierce Compromise: The Administrative Procedure Act Emerges from New Deal Politics, 90 Nw. U. L. Rev. 1557, 1681–82 (1996) (explaining that the APA was a compromise between liberals and conservatives).

12. 5 U.S.C. §§ 554, 556, 557.

13. Id. §§ 553, 556, 557.

14. Id. § 553.

15. Id. §§ 702–706.

16. Id. §§ 556, 557.

17. Aaron L. Nielson, In Defense of Formal Rulemaking, 75 Ohio St. L.J. 237, 253 (2014) (discussing how since 1973 formal rulemaking has “become almost extinct”).

is triggered only by the presence of specific statutory language in an enabling act.<sup>18</sup>

In the absence of this statutory language, notice-and-comment rulemaking procedures are the default. There are limited statutory exemptions for foreign and military affairs, matters regarding agency personnel or management, and public-property issues.<sup>19</sup> The APA general rulemaking procedures also do not apply to “interpretive rules, general statements of policy, or rules of agency organization, procedure, or practice.”<sup>20</sup> Notice-and-comment rulemaking is the most common type of quasi-legislative agency action that the APA governs and the primary vehicle for agencies to create legally binding regulations.<sup>21</sup>

The statutory requirements for the informal rulemaking process begin with a mandate that federal agencies publish a Notice of Proposed Rulemaking (NPR) publicly in the Federal Register and provide the NPR directly to interested parties. The original purpose of this requirement was to “fairly apprise interested parties of the issues involved, so that they may present responsive data or argument[s].”<sup>22</sup> Then, the agency must “give interested persons an opportunity to participate in the rulemaking through submission of written data, views or arguments.”<sup>23</sup> After “consideration of the relevant matter presented,” the agency may promulgate a final rule that includes an explanation of how it addresses important comments submitted.<sup>24</sup> The validity of such agency action is evaluated on judicial review based on the agency’s record under the “arbitrary and capricious” standard.<sup>25</sup>

The benefit of taking and considering public comments is multifaceted. Defenders of notice-and-comment rulemaking laud its deliberative democratic qualities: its ability to engage the public in the administrative process, collect and vet ideas, and protect agencies from capture.<sup>26</sup> The comment process requires that agencies read, consider, and respond to input. In doing so, it engages the public in deliberations that build legiti-

---

18. 5 U.S.C. § 556(c) (requiring rulemaking to be “on the record after opportunity for an agency hearing”); *United States v. Fla. E. Coast Ry.*, 410 U.S. 224, 237–38 (1973) (stating that specific triggering language is required to mandate formal rulemaking).

19. See 5 U.S.C. § 553(a)(1)–(2).

20. *Id.* § 553(b)(3)(A).

21. Sidney A. Shapiro & Richard W. Murphy, *Eight Things Americans Can’t Figure Out About Controlling Administrative Power*, 61 *Admin L. Rev.* 5, 13 (2009) (stating that notice-and-comment rulemaking is “the default mode in the federal government for making ‘binding’ legislative rules”).

22. S. Rep. No. 79-752, at 14 (1945).

23. 5 U.S.C. § 553(c).

24. *Id.* § 553(b)–(c).

25. *Id.* § 706(2)(a).

26. See *Sierra Club v. Costle*, 657 F.2d 298, 400 (D.C. Cir. 1981) (“[T]he very legitimacy of general policymaking performed by unelected administrators depends in no small part upon the openness, accessibility, and amenability of these officials to the needs and ideas of the public . . .”).

macy for democratic government institutions and the resulting regulations.<sup>27</sup>

Second, comments build agency expertise by expanding access to data and cognizance of impacted groups.<sup>28</sup> This not only increases accountability by foreclosing the argument that agencies were unaware of consequences, but it also aids in the public-education functions of agencies. Finally, public comment collection and examination creates a record for effective judicial review of agency decisionmaking. So far, the picture painted is a rosy one; notice-and-comment rulemaking sounds, in theory, like a good legislative solution to quality, capture, uniformity, and responsiveness concerns overshadowing agency action.

But to its critics, notice-and-comment rulemaking has not delivered on these promises. In practice, the notice-and-comment process has become increasingly burdensome, expensive, and time consuming, as judges have interpreted the APA's statutory requirements to place substantial procedural burdens on agency action.<sup>29</sup> Current case law requires that a final rule be set aside unless it is a "logical outgrowth" of the proposed rule.<sup>30</sup> To show this, agencies must articulate their "unspoken thoughts" in the NOPR and offer some "persuasive evidence that possible objections to its final rules have been given sufficient consideration."<sup>31</sup> To withstand judicial scrutiny, a valid final rule often must also include a comprehensive and meticulous treatment of facts and arguments considered and the reasoning motivating agency action.<sup>32</sup>

This has led some to view notice-and-comment rulemaking as a sluggish exercise in popular window dressing. In 1960, administrative

---

27. See Nicholas Bagley, Remedial Restraint in Administrative Law, 117 Colum. L. Rev. 253, 265 (2017) ("Broadly speaking, notice-and-comment rulemaking serves both *informational* and *participatory* functions: It assures that agencies incorporate all relevant information into their decisionmaking and guarantees that members of the public have a voice in the decisions that affect their lives.").

28. See Cass R. Sunstein, Democratizing Regulation, Digitally, Democracy (Fall 2014), <http://democracyjournal.org/magazine/34/democratizing-regulation-digitally/> [<http://perma.cc/4RHH-GP78>] (identifying how public participation can improve the quality and legitimacy of rules by making rulemakers accountable).

29. These developments began with *Abbott Labs. v. Gardner*, 387 U.S. 136, 136–37 (1967) (allowing pre-enforcement review of rulemaking on the basis of the NOPR, comments from the public, and the final rule).

30. *Small Refiner Lead Phase-Down Task Force v. EPA*, 705 F.2d 506, 546–47 (D.C. Cir. 1983). The "logical outgrowth" test is used interchangeably with the "sufficiently foreshadowed" test. See *Horsehead Res. Dev. Co. v. Browner*, 16 F.3d 1246, 1267–68 (D.C. Cir. 1994) (using the two tests interchangeably).

31. *Shell Oil v. EPA*, 950 F.2d 741, 751–52 (D.C. Cir. 1991).

32. Richard J. Pierce, Jr., *Administrative Law* 68 (2d ed. 2012) (explaining that circuit courts require agencies to include detailed discussions of the reasoning behind their courses of action).

agencies issued notices of 498 proposed rulemakings annually;<sup>33</sup> by 2008, that number ballooned to over 2,475 per year.<sup>34</sup> Regulations.gov estimates that federal agencies now issue nearly 8,000 regulations per year.<sup>35</sup> The number of comments a proposed rule receives can be enormous. In 2012, the Environmental Protection Agency (EPA) was inundated with a deluge of over 2.5 million comments on its proposed rule regarding greenhouse gas performance standards.<sup>36</sup> Today, the legal requirements of the APA, coupled with increased submissions, have left chronically under-resourced agencies in a no-win situation.

## II. THE BIG DATA SOLUTION: PITFALLS AND POTENTIAL

With notice-and-comment straining under the weight of its procedural burdens, it is no wonder that commentators and policymakers are looking for creative solutions. Enter big data, the newest proposed solution to the comment kerfuffle, allowing the public to have their cake and comment on it too. Computerized processing techniques target notice-and-comment rulemaking not only because it is the most common type of quasi-legislative agency action contemplated under the APA but also because it is costly and time-consuming.<sup>37</sup>

However, notice-and-comment rulemaking is where natural language processing is most likely to hit legal hurdles. Purveyors of the software argue that it will decrease the time that it takes for agencies to review comments and make rulemaking faster and cheaper.<sup>38</sup> Leaving aside the obvious complications of time spent on training, developing inputs, and quality control measures, a more substantive question lingers:

---

33. See Reuel E. Schiller, *Rulemaking's Promise: Administrative Law and Legal Culture in the 1960s and 1970s*, 53 *Admin. L. Rev.* 1139, 1147 (2001) (documenting the rapid and exponential rise of informal rulemaking).

34. See Maeve P. Carey, Cong. Research Serv., R43056, *Counting Regulations: An Overview of Rulemaking, Types of Federal Regulations, and Pages in the Federal Register* 18 (2016), <http://fas.org/sgp/crs/misc/R43056.pdf> [<http://perma.cc/C35G-5PQU>].

35. Site Data: Your Voice in Action, Regulations.gov, <http://www.regulations.gov/siteData> [<http://perma.cc/V4Z6-TM3H>] (last visited July 28, 2017).

36. Greenhouse Gas New Source Performance Standard for Electric Generating Units, Regulations.gov, <http://www.regulations.gov/docket?D=EPA-HQ-OAR-2011-0660> [<http://perma.cc/U6C4-BW3A>] (last visited July 28, 2017). In the same year, the EPA received over 300,000 comments in relation to vehicle emissions. See EPA/NHTSA Joint Rulemaking to Establish Light-Duty Vehicle GHG Emissions Standards and CAFE Standards for Model Year 2017 and Later, Regulations.gov, <http://www.regulations.gov/docket?D=EPA-HQ-OAR-2010-0799> [<http://perma.cc/GJE8-6NGD>] (last visited July 28, 2017).

37. Cf. Maeve P. Carey, Cong. Research Serv., R44348, *Methods of Estimating the Total Cost of Federal Regulations 2* (2016), <http://fas.org/sgp/crs/misc/R44348.pdf> [<http://perma.cc/8W7P-ADD9>] (discussing two different approaches to calculating the high cost of regulations issued by federal agencies).

38. Hurtado, *supra* note 4 (highlighting software proponents' claims that software can read and analyze comments "10,000%' faster" and at a much lower cost).

Is it possible to do the required work of notice-and-comment rulemaking with a computer making the first (and perhaps last) determination of what is important for bureaucrats (people) to read? To withstand judicial scrutiny, an agency's decision must stand based on the agency record—what did they actually review? The automated culling process leaves important holes that eviscerate administrative benefits to the process and render administrative action subject to remand for procedural deficiency.

Take “[analyzing] sentiment,” the example given by the software design firm currently wooing the administrative state rulemaking.<sup>39</sup> This computerized sorting process allows “[p]olicy makers [to] instantaneously detect positive and negative expressions to determine issue-specific public sentiment.”<sup>40</sup> Here, forcing rulemaking into a binary sorting process presupposes that content submitted by complex humans contains internal and verbal sentiment coherence that the program can recognize. However, many nuanced and useful comments are a mix of both positive and negative. A writer may claim that she speaks in opposition, yet the information submitted may speak in favor of the rule. Other submissions may lack a clear positional stance. Moreover, as a policy matter, the popularity of a rule does not determine its validity or the utility of the comment.

For the most part, useful and important comments contain information—opinions, data, examples, or experiences—that allows the agency to base its decision on expertise and place the rule in factual context. Rulemaking is about discovering stakeholders and inadequacies, not anticipating them. The act of reading comments grouped as “for” or “against” (or no position) may also impact the agency factfinder in the presentation of the information she is assimilating. For machines to do this work, the devil is in the details: How would an agency choose to sort the comments? How would processing be sufficiently tailored to help agencies expedite the comment process without sacrificing content?

The literature on technology-assisted legal work makes much of the distinction between traditional keyword searches and predictive coding.<sup>41</sup> While machine learning advances the frontier of what computers can do (think Google Translate), it cannot overcome its fundamental limitation: Machines can only classify (to a point) new material based on a stock of

---

39. Regendus, <http://www.regendus.com/> [<http://perma.cc/M3TY-DFFM>] (last visited July 28, 2017) (highlighting “analyzing sentiment” as a way to determine issue-specific public sentiment or develop more effective comment strategies).

40. *Id.*; see also Hurtado, *supra* note 4 (showing a screenshot of a public comment analyzed using natural language processing in which software underlined “pro” language in green and “con” language in red.).

41. Predictive coding is a huge academic field and an even bigger market for software companies. Predictably (no pun intended), the predictive coding approaches of different companies vary significantly. This Piece discusses predictive coding only in general terms.

pre-coded examples that help to “train” the algorithm.<sup>42</sup> This limitation has important consequences in the context of automated rulemaking.

First, the algorithm must be trained with pre-coded examples.<sup>43</sup> A different starting stock of examples leads to a different algorithm. Mistakes or omissions in examples lead to mistakes later. If inputs change between training and application, the algorithm will mispredict and need to be repeatedly recalibrated.<sup>44</sup> Agencies must be aware of these limitations and retain full control over “training” an algorithm. Second, current machine learning is geared toward sorting material into categories. However, predictive coding struggles to create new categories and may simply miss that hundreds of publicly submitted comments raise a concern not anticipated by the agency or by the material used to train the algorithm. Third, because their power stems from simulating or resembling complex, multilayered neural networks, sophisticated machine learning algorithms are unable to explain why or how they make categorization decisions. Such networks are calibrated through the previously mentioned “training” that adjusts connections between network nodes. Understanding the algorithm’s “decisionmaking” process would require an analysis of each node, its connections, and its relative weight. This is an increasingly daunting and impossible task.

Finally, algorithms do not avoid normative judgments but embed them deep in the decisionmaking process. Humans must still decide whether a concern that a comment raises is worthy of attention and consideration. Algorithms can do this work (as they do every day—think of a Facebook feed), but they still make decisions. Just because an algorithm decided something does not mean it was a “disinterested” decision. The normative decisions embedded within the algorithm, combined with a tendency to view numbers as neutral, renders predictive coding dangerous because important decisions that agencies have made in the past may be outsourced to private subcontractors that are not subject to accountability standards.<sup>45</sup>

Despite these inherent limitations, some uses of automated technology in rulemaking might support agency action without violating

---

42. See Ralph C. Losey, *Predictive Coding and the Proportionality Doctrine: A Marriage Made in Big Data*, 26 *Regent U. L. Rev.* 7, 21–23 (2014) (detailing ways in which humans can “train” machines).

43. Machines “learn” from an interactive process with a human “subject matter expert” who demonstrates how to categorize documents by doing a sample of the work. *Id.* at 21. The machine then extrapolates from these inputs to create its own analysis based on different methodological sampling approaches. *Id.* at 21–22.

44. See Yablon & Landsman-Roos, *supra* note 5, at 639–41 (discussing the iterative process of training predictive coding systems); see also Pace & Zakaras, *supra* note 1, at 60 (discussing the problems and difficulties associated with training predictive coding systems).

45. One way to alleviate this concern is to require that all predictive coding algorithms used by agencies be open-sourced, trained, and supervised by agency employees rather than private subcontractors.

the statutory requirements of the APA. For example, removing duplicate submissions, when truly identical, appears to save time with little substantive loss.<sup>46</sup> This is a mechanical process, the equivalent of a keyword search, which is a fundamentally different process from using an automated analysis to sort comments based on fluid and adapting criteria. Using processing to create broad data analytics that study participant data, and not the rough content of the submissions, might provide insight into agency capture. However, even such an analysis is complicated by trade associations, joint submissions, and the choice (which must be made by the human training the software) of whether a 100-page submission should count in the same way as a one-line opinion.<sup>47</sup> Here, humans again must make hard decisions.

This begs the question: Is the role of data processing in administrative action doomed to undermine agency effectiveness in the name of expediency? Not necessarily. With the correct inputs, it could be highly advantageous in areas in which the purpose of agency action is not open communication, policy formation, and knowledge acquisition, but information retrieval and organization. The Freedom of Information Act (FOIA) grants the public presumptive access to all agency records unless requested material falls into limited statutory exemptions.<sup>48</sup> Under FOIA, any person, whether an individual or a corporation, may request and obtain existing, identifiable, and unpublished agency records on any topic.<sup>49</sup> In 2014 alone, the federal government received 714,231 FOIA requests, and this number continues to rise.<sup>50</sup> Recent data places the estimated cost of FOIA-related activities for all federal departments and agencies at \$429.6 million.<sup>51</sup> FOIA responsiveness is an area in which carefully crafted predictive coding could save agency resources. With correct human oversight, natural language processing could also capture

---

46. See Admin. Conference of the U.S., Administrative Conference Recommendation 2011-1: Legal Considerations in E-Rulemaking 4 (2011), <http://www.acus.gov/sites/default/files/documents/Recommendation%202011-1%20%28Legal%20Considerations%20in%20e-Rulemaking%29.pdf> [<http://perma.cc/9SKW-7X7V>] (“While 5 U.S.C. § 553 requires agencies to consider all comments received, it does not require agencies to ensure that a person reads each one of multiple identical or nearly identical comments.”).

47. See Carey, *supra* note 34, at I (noting that for congressional analysis, the Congressional Research Service uses two metrics: “The number of federal rules issued annually and the total number of pages in the *Federal Register*”).

48. There are nine specifically exempted categories of information. See 5 U.S.C. § 552(b) (2012).

49. See *id.* § 552(a)(3)(A).

50. Wendy Ginsberg & Michael Greene, Cong. Research Serv., 97-71, Access to Government Information in the United States: A Primer 5 (2016), <http://fas.org/sgp/crs/secretary/97-71.pdf> [<http://perma.cc/CF3P-4E65>] (noting that the total number of FOIA requests in 2014 was 9,837 more than in 2013). The 2014 total number of FOIA requests excludes over 150,000 backlogged requests from 2014. *Id.*

51. Wendy Ginsberg, Cong. Research Serv., R41933, The Freedom of Information Act (FOIA): Background, Legislation, and Policy Issues 11 (2014), <http://fas.org/sgp/crs/secretary/R41933.pdf> [<http://perma.cc/65V9-BL5N>].

useful data analytics on who is making requests and with what frequency, thereby clarifying instances of agency capture and other lopsided participation.

#### CONCLUSION

Automation can do many things and its proponents make many claims. Companies' broad assertions of efficiency—which promise a lot but can only deliver so much—should not woo lawmakers, judges, and the public. To use data analytics responsibly, agencies need Congress to provide guidance on acceptable uses as soon as possible.<sup>52</sup> Lawmakers ought to devise stringent quality controls and build open-source requirements into agency protocols to maintain agency independence and increase transparency of embedded flaws and value assumptions. Agencies could legitimately use automation to supplement policy work, increase accountability through data diagnostics, and search for specific facts in their archives.

However, agencies should not use automation to supplant required review of public comments. Notice-and-comment rulemaking is grounded in democratic deliberation and the creative process of lawmaking. Trimming too much through natural language processing also renders rulemaking vulnerable to upheaval on judicial review. While efficiency gains from automation are unclear, this much is certain: Remanding and reinitiating the rulemaking process is less efficient than just doing it right the first time.

---

52. The natural home for this would be an amendment to the E-Government Act of 2002. Pub. L. No. 107-347, 116 Stat. 2899 (codified as amended in scattered titles of the U.S.C.).