

DRAFT REPORT FOR THE
ADMINISTRATIVE CONFERENCE OF THE UNITED STATES

**MASS, COMPUTER-GENERATED, AND FRAUDULENT
COMMENTS**

Steve Balla
George Washington Regulatory Studies Center

Reeve Bull
Administrative Conference of the United States

Bridget Dooling
George Washington Regulatory Studies Center

Emily Hammond
George Washington Law School

Michael Livermore
University of Virginia School of Law

&

Beth Simone Noveck
The Governance Lab

This report was prepared for the consideration of the Administrative Conference of the United States. It does not necessarily reflect the views of the Conference (including its Council, committees, or members).

Recommended Citation

Steve Balla, Reeve Bull, Bridget Dooling, Emily Hammond, Michael Livermore & Beth Simone Noveck, Mass, Computer-Generated, and Fraudulent Comments (Apr. 2, 2021) (draft report to the Admin. Conf. of the U.S.).

WORKING DRAFT:
Mass, Computer-Generated, & Fraudulent Comments

*Steve Balla, Reeve Bull, Bridget Dooling, Emily Hammond, Michael Herz,
Michael Livermore & Beth Simone Noveck*

Executive Summary

This report explores three forms of commenting in federal rulemaking that have been enabled by technological advances: mass, fraudulent, and computer-generated comments. Mass comments arise when an agency receives a much larger number of comments in a rulemaking than it typically would (e.g., thousands when the agency typically receives a few dozen). The report focuses on a particular type of mass comment response, which it terms a “mass comment campaign,” in which organizations orchestrate the submission of large numbers of identical or nearly identical comments. Fraudulent comments, which we refer to as “malattributed comments” as discussed below, refer to comments falsely attributed to persons by whom they were not, in fact, submitted. Computer-generated comments are generated not by humans, but rather by software algorithms. Although software is the product of human actions, algorithms obviate the need for humans to generate the content of comments and submit comments to agencies.

This report examines the legal, practical, and technical issues associated with processing and responding to mass, fraudulent, and computer-generated comments. There are cross-cutting issues that apply to each of these three types of comments. First the nature of such comments may make it difficult for agencies to extract useful information. Second, there are a suite of risks related to harming public perceptions about the legitimacy of particular rules and the rulemaking process overall. Third technology-enabled comments present agencies with resource challenges.

The report also considers issues that are unique to each type of comment. With respect to mass comments, it addresses the challenges associated with receiving large numbers of comments and, in particular, batches of comments that are identical or nearly identical. It looks at how agencies can use technologies to help process comments received and at how agencies can most effectively communicate with public commenters to ensure that they understand the purpose of the notice-and-comment process and the particular considerations unique to processing mass comment responses. Fraudulent, or malattributed, comments raise legal issues both in criminal and Administrative Procedure Act (APA) domains. They also have the potential to mislead an agency and pose harms to individuals. Computer-generated comments may raise legal issues in light of the APA’s stipulation that “interested persons” are granted the opportunity to comment on proposed rules. Practically, it can be difficult for agencies to distinguish computer-generated comments from traditional comments (i.e., those submitted by humans without the use of software algorithms).

While technology creates challenges, it also offers opportunities to help regulatory officials gather public input and draw greater insights from that input. The report summarizes several innovative forms of public participation that leverage technology to supplement the notice and comment rulemaking process.

The report closes with a set of recommendations for agencies to address the challenges and opportunities associated with new technologies that bear on the rulemaking process. These recommendations cover steps that agencies can take with respect to technology, coordination, and docket management.

Introduction

In 2015, the Environmental Protection Agency (EPA) issued a rule defining the “Waters of the United States.” During the course of the rulemaking, the EPA received more than one million comments.¹ Over ninety percent of the comments were submitted as part of mass comment campaigns. At about the same time, a large number of mass comment campaigns was also submitted in response to the EPA’s proposed “Clean Power Plan.” These and other similar occurrences signaled the regularity of mass comment campaigns, at least in the context of highly salient, controversial rulemakings.

In February 2020, the Subcommittee on Oversight and Investigations of the U.S. House of Representatives Committee on Financial Services investigated allegations that the Securities and Exchange Commission (SEC) may have received fraudulent comments from “an Army veteran, a Marine veteran, a single mom, and a couple of retirees.”² During the committee hearing, witnesses alleged that those public comments were sent by advocacy groups and signed with the names of people who never saw the comments or who did not exist at all. Although these allegations have been disputed, the possibility of fraudulent comments in the government’s rulemaking process has prompted questions and concern.

In 2017, the Federal Communications Commission (FCC) “Restoring Internet Freedom” (i.e., net neutrality) rulemaking attracted a record number of public comments: almost 22 million by the official close of the comment period, with another three million arriving after the fact.³ Although about six percent of the comments were unique, the rest were submitted multiple times, in some cases hundreds of thousands of times.⁴ On nine different occasions, more than 75,000

¹ See Steven J. Balla et al., *Lost in the flood?: Agency Responsiveness to Mass Comment Campaigns in Administrative Rulemaking*, REGUL. & GOVERNANCE (May 20, 2020), <https://onlinelibrary.wiley.com/doi/abs/10.1111/rego.12318>; Rachel Augustine Potter, *More than Spam?: Lobbying the EPA through Public Comment Campaigns*, BROOKINGS INST. (Nov. 19, 2017), <https://www.brookings.edu/research/more-than-spam-lobbying-the-epa-through-public-comment-campaigns/>.

² *Fake It Till They Make It: How Bad Actors Use Astroturfing to Manipulate Regulators, Disenfranchise Consumers, and Subvert the Rulemaking Process: Hearing Before the H. Fin. Servs. Comm.*, 116th Congress (Feb. 6, 2020) (hereinafter *Fake It Till They Make It*).

³ In 2015, the FCC issued its “net neutrality rule,” prohibiting broadband Internet providers from blocking, degrading, or interfering with Internet traffic. Protecting and Promoting the Open Internet, 80 Fed. Reg. 19737 (June 12, 2015). The rule was upheld in *United States Telecom. Ass’n v. FCC*, 825 F.3d 674 (D.C. Cir. 2016). In 2017, the FCC proposed repealing the 2015 rule. *Restoring Internet Freedom*, Notice of Proposed Rulemaking, 32 FCC Rcd 4434 (2017). The Notice of Proposed Rulemaking was released in May and published in the Federal Register in June. 82 Fed. Reg. 25,568 (June 2, 2017). The Final Rule was promulgated in early January 2018. See *Restoring Internet Freedom*, 83 Fed. Reg. 7852 (Feb. 22, 2018). For information on the comments, see Paul Hitlin, Kenneth Olmstead, & Skye Toor, *Public Comments to the Federal Communications Commission About Net Neutrality Contain Many Inaccuracies and Duplicates*, PEW RES. CTR. (Nov. 29, 2017), https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2017/11/PI_2017.11.29_Net-Neutrality-Comments_FINAL.pdf.

⁴ *Id.*

comments were dumped into the docket at the very same second.⁵ The comments “included comments from stolen email addresses, defunct email accounts and people who unwittingly gave permission for their comments to be posted.”⁶ A consulting firm later determined that about a third of the comments were sent from temporary or disposable email domains, and about 10 million were from senders of multiple comments.⁷ FCC Commissioner Jessica Rosenworcel has stressed that 500,000 or so comments came from Russia.⁸ The New York Attorney General concluded that eight million comments were submitted with false identities.⁹ In sum, the nature of public participation—the mass occurrence of identical and near duplicate comments, the malattribution of identities, and the apparent automation of comment submission—called into question elements of the process behind the FCC’s regulation.

Partly in response to issues posed by such rulemakings, in 2019 the U.S. Senate Permanent Subcommittee on Investigations issued a staff report entitled, “Abuses of the Federal Notice-and-Comment Rulemaking Process.”¹⁰ The report identified problems associated with mass, fraudulent, and computer-generated comments, including a lack of agency processes and policies aimed at identifying, managing, and addressing such comments. This attention underscores the importance of comments for the rulemaking process, which generates thousands of regulations every year that touch many aspects of economic and social life.¹¹

Under the 1946 APA, the public has a right to participate by commenting on draft regulations in the rulemaking process, which is why it is often referred to as notice-and-comment

⁵ *Id.*

⁶ James V. Grimaldi & Paul Overburg, *Millions of People Post Comments on Federal Regulations. Many Are Fake.*, WALL ST. J. (Dec. 13, 2017), <https://www.wsj.com/articles/millions-of-people-post-comments-on-federal-regulations-many-are-fake-1513099188>. A video on the newspaper’s website summarizes: “[T]he Wall Street Journal uncovered thousands of comments from fake email addresses, abandoned or defunct email accounts, posted on behalf of unwitting participants. For example, 818,000 identical comments on the FCC site favor repealing the rules. In a random sample of people whose emails were used for those posts, 72% said they had nothing to do with them. Jack Hirsch was one of them. “I was horrified. Knowing that this is actually an issue that I cared enough to write my representatives about, and knowing that my information had been falsified to support a completely opposing view, it was really frustrating, and honestly, I felt like there was no recourse.” *Thousands of Fake Comments on Net Neutrality: A WSJ Investigation*, WALL ST. J. (Dec. 12, 2017, 12:02 PM), <https://www.wsj.com/video/thousands-of-fake-comments-on-netneutrality-a-wsj-investigation/8E52172E-821C-4D89-A2AA-2820F30B8648.html>.

⁷ *FCC Restoring Internet Freedom Docket 17-108: Comments Analysis*, EMPRATA 2 (Aug. 30, 2017), <https://www.emprata.com/emp2017/wp-content/uploads/2017/08/FCC-Restoring-Internet-Freedom-Comments-Analysis.pdf>.

⁸ Nicholas Confessore and Jeremy Singer-Vine on Request for Inspection of Records, 33 FCC Rcd. 11808 (adopted Nov. 7, 2018) (Rosenworcel, dissenting).

⁹ Letter from Eric Schneiderman, Att’y Gen., N.Y., to Thomas M. Johnson, Jr., Gen. Counsel, FCC (Dec. 13, 2017), https://ag.ny.gov/sites/default/files/ltr_to_fcc_gen_counsel_re_records_request.pdf.

¹⁰ PERMANENT SUBCOMMITTEE ON INVESTIGATIONS, U.S. SENATE COMMITTEE ON HOMELAND SECURITY AND GOVERNMENT AFFAIRS, STAFF REPORT, ABUSES OF THE FEDERAL NOTICE-AND-COMMENT RULEMAKING PROCESS (2019), <https://www.hsgac.senate.gov/imo/media/doc/2019-10-24%20PSI%20Staff%20Report%20-%20Abuses%20of%20the%20Federal%20Notice-and-Comment%20Rulemaking%20Process.pdf> (hereafter “Subcommittee Report”).

¹¹ CONG. RSCH. SERV., R43056, COUNTING REGULATIONS: AN OVERVIEW OF RULEMAKING, TYPES OF FEDERAL REGULATIONS, AND PAGES IN THE FEDERAL REGISTER (2019).

rulemaking.¹² Such participation in rulemaking enhances both the legitimacy and the quality of regulations by enabling agencies (and the executive offices and congressional committees that oversee them) to obtain information from a wide audience of stakeholders, interest groups, businesses, nongovernmental organizations (NGOs), academics, and interested individuals. Participation also provides an accountability check on the rulemaking process by ensuring public scrutiny prior to rules going into effect. Overall, the process facilitates the reason-giving requirement that is necessary for a rule to survive judicial review.

The shift over the last two decades to a digital process, in which participants submit comments via the Internet, has made commenting easier. Expanding the ability to participate in an important governance process is an overwhelmingly positive change. Rulemaking has long been criticized as an insiders' game, invisible to the general public, with modest levels of participation, almost entirely by organizations rather than individuals. But, although online participation has the potential to increase the quantity and diversity of participation, it has also inadvertently opened the floodgates to mass comment campaigns, malattributed comments, and computer-generated comments, potentially making it harder for agencies to extract the information needed to inform decision making and undermining the legitimacy of the rulemaking process.

As this discussion indicates, the technology of electronic commenting has enabled three kinds of potentially problematic comments that are the subject of this report:¹³

1. Mass comments (sometimes orchestrated as a campaign by one or more entities);
2. Fake, fraudulent, or what this report below refers to as “malattributed” comments; and
3. Computer-generated comments.

All three of these types of comments can generate serious challenges to agencies, raising a pressing set of questions concerning how best to respond while ensuring the functioning of the informal rulemaking process. The task of this report, therefore, is to evaluate whether and to what extent such submissions are problematic, and to make recommendations for how rulemaking agencies should respond using legal, policy, and technological strategies.

Our overarching conclusion is that agencies should adopt both low- and high-tech measures to limit the negative impact of these sorts of comments. Mass, malattributed, and computer-generated comments, however, do not represent a crisis for the regulatory state at this time. They have not been found to violate federal law and do not generally undermine the integrity of notice-and-comment rulemaking, and we are not aware of evidence of widespread substantive harms in particular rulemaking efforts or to the rulemaking system overall. However, appropriate responses, especially those that take advantage of new technology, could reduce the cost and negative impacts of technology-enabled comments.

Adopting such techniques could, for example, afford agency officials more time to improve the opportunity for a diverse public to participate in the rulemaking process meaningfully and to augment current practices with new forms of citizen engagement. Indeed, in addition to exploring how new technologies—the very same technologies that enable mass, fraudulent and computer-generated comments—can also help with analyzing those comments, we also explore throughout how technology can help regulatory officials make sense of public input and draw greater insights

¹² *Id.*

¹³ These terms are defined in detail below. See section II.A.

from public comments of all kinds. Finally, other jurisdictions at the state and local level and internationally are turning to the use of new technology to enable innovative forms of public participation to improve the quality of rule- and policymaking. As detailed in Part VI, these illustrate hopeful opportunities for future experimentation.

Written by seven professionals with expertise in administrative law, rulemaking practice, and new technologies, this report is informed by a review of Permanent Subcommittee's report,¹⁴ relevant law, related legal and social scientific scholarship, and a set of interviews with agency personnel with background in the rulemaking process at agencies with substantial rulemaking dockets during the summer and fall of 2020. The interviews, which were not meant to capture the views of a representative or random sample, were with staff of the EPA, the Consumer Finance Protection Bureau (CFPB), the Department of Transportation (DOT), and the FCC, as well as with officials from the General Services Administration (GSA) responsible for developing and maintaining the Federal Docket Management System (FDMS). A technical advisory group of experts drawn from government, private industry, and academia also provided feedback to the report authors, as did an additional online roundtable of agency officials with experience in the notice-and-comment process.

The report is divided seven parts. Part I provides a general introduction to notice-and-comment rulemaking and the role of technology in that process. Part II discusses recent technological development that have contributed to the growth of mass, malattributed, and computer-generated comments, and describes some of the challenges associated with these types of comments. Parts III, IV, and V focus on each of these comment types in turn. Part VI discusses technological opportunities, with a focus on current available tools that can be used to facilitate the processing of information from the notice-and-comment process or enhance supplements to the notice-and-comment process. Part VI offers draft recommendations based on the findings in this report.

I. Technology and Notice-and-Comment Rulemaking

During the latter half of the twentieth century, there was considerable growth in the use of informal rulemaking by administrative agencies.¹⁵ The procedures for informal rulemaking are set out in section 553 of the APA. A cornerstone of this process is the opportunity for members of the public to submit comments on rulemaking proposals,¹⁶ which is why it is often referred to as notice-and-comment rulemaking. For decades, organizations and individuals have availed themselves of this opportunity to help inform the process of regulatory development. Public comments on agency rulemakings take a wide variety of forms that include detailed submissions by sophisticated repeat players, short expressions of support or opposition from members of the

¹⁴ The Subcommittee Report addressed a handful of additional topics, including obscenity and copyrighted materials in public comments, that are not addressed in the current report. ACUS will provide a copy of the Subcommittee Report to the Committee on Rulemaking as they begin deliberations on this project.

¹⁵ See Christopher DeMuth, *Can the Administrative State Be Tamed?*, 8 J. LEGAL ANALYSIS 121, 126–27 (2016).

¹⁶ Cary Coglianese, *Citizen Participation in Rulemaking: Past, Present, and Future*, 55 DUKE L.J. 943, 945 (2006) (“participation in rulemaking is one of the most fundamental, important, and far-reaching of democratic rights”).

public, signed form letters in response to solicitations from NGOs, and technical reports from unaffiliated experts.¹⁷

There is considerable variation in the level of public participation from one rulemaking to another. The vast majority of rulemakings are relatively unremarked upon by the public, with—at most—participation by the stakeholders most affected by a rule.¹⁸ This level of participation is not surprising given the often highly technical and specialized nature and low visibility of many rulemakings. While federal agencies publish the opportunity to participate in the *Federal Register* (effectively, the newspaper of the federal government), they generally do not advertise rulemakings elsewhere and the public tends to have little knowledge of the right to engage unless a third party promotes the opportunity. In a small percentage of well-publicized rulemakings with particular public salience—such as those highlighted above—public participation can be orders of magnitude above the norm, with the number of comments ranging from thousands to millions.

The APA sets forth the key elements of notice-and-comment rulemaking. Subject to certain exceptions, agencies first must publish a general notice of the proposed rulemaking in the *Federal Register*.¹⁹ That notice “shall include—

- (1) a statement of the time, place, and nature of public rule making proceedings;
- (2) reference to the legal authority under which the rule is proposed; and
- (3) either the terms or substance of the proposed rule or a description of the subjects and issues involved.”²⁰

The APA further provides that “[a]fter notice required by this section, the agency shall give interested persons an opportunity to participate in the rule making through submission of written data, views, or arguments with or without opportunity for oral presentation. After consideration of the relevant matter presented, the agency shall incorporate in the rules adopted a concise general statement of their basis and purpose.”²¹ “Person” is broadly defined to include “an individual, partnership, corporation, association, or public or private organization other than an agency.”²²

¹⁷ See Michael A. Livermore, Vladimir Eidelman & Brian Grom, *Computationally Assisted Regulatory Participation*, 93 NOTRE DAME L. REV. 977 (2018).

¹⁸ David M. Shafie, *Participation in E-Rulemaking: Interest Groups and the Standard-Setting Process for Hazardous Air Pollutants*, 5 J. INFO. TECH. & POL. 399, 403–405 (2008). See also Stuart Shapiro, *When Do Agencies Change Their Proposed Rules* (Nov. 10, 2007) (presented at APPAM: Association for Public Policy Analysis & Management Conference) (unpublished manuscript), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1026066 (finding that fewer than 10% of rulemakings received more than 100 comments); John M. de Figueiredo, *E-Rulemaking: Bringing Data to Theory at the Federal Communications Commission*, 55 DUKE L.J. 969 (2006). Other agency actions that are subject to public commenting follow as similar pattern. An analysis focusing on the fifty draft guidance documents published for public comment by the Environmental Protection Agency (EPA) from 2011-2014 found that just eight received more than 5,000 comments, with five exceeding 40,000 comments. Nicholas R. Parrillo, *Should the Public Get to Participate Before Federal Agencies Issue Guidance? An Empirical Study*, 71 ADMIN. L. REV. 57, 97-98 (2019) [hereinafter Parrillo, *Should the Public Get to Participate*]; Nicholas R. Parrillo, *Federal Agency Guidance: An Institutional Perspective* (Oct. 12, 2017) (report to the Admin. Conf. of the U.S.).

¹⁹ 5 U.S.C. § 553(b).

²⁰ *Id.*

²¹ 5 U.S.C. § 553(c).

²² 5 U.S.C. § 551(2).

Consistent with the broad ideals underlying the commenting process, courts read this provision expansively.²³ Courts have also elaborated that the purpose of the notice requirement is to facilitate meaningful comments.²⁴ For example, agencies must disclose in their notice any scientific or technical details on which they base their proposed rules in order to give the public a fair opportunity to react and comment thereon.²⁵

Agencies not only must provide notice and an opportunity for public comments but also must then “consider[] . . . the relevant matter presented” in those comments.²⁶ The courts have interpreted this language to require that, in the notice of the final rule, agencies respond to “significant” comments—those that, “if true, . . . would require a change in [the] proposed rule.”²⁷ Failure to so respond is grounds for remand.²⁸ Under § 706(2)(A) of the APA, agency rulemakings will be set aside if they are “arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law.”²⁹ The failure to acknowledge and respond to substantive concerns raised in the rulemaking process is one of the grounds for a court to find that an agency’s rulemaking fails under the arbitrary or capricious standard.³⁰

Although the requirement to respond to comments is serious, it is not absolute. The “APA requirement of agency responsiveness to comments is subject to the common-sense rule that a response [is not always] necessary.”³¹ Comments that “are purely speculative and do not disclose the factual or policy basis on which they rest require no response.”³²

In recognition of the potential for information and communication technologies to facilitate broader participation in the regulatory process, Congress enacted the E-Government Act in October 2002. Among other things, the George W. Bush administration established the eRulemaking Program to spearhead the creation of an online system for conducting the notice-and-comment process at agencies throughout the federal government.³³ To the extent deemed practicable, each agency must post information required to be published in the *Federal Register*

²³ E.g., *O’Rourke v. U.S. Dep’t of Justice*, 684 F. Supp. 716, 718 (D.D.C. 1988) (holding “person” includes non-citizens and collecting cases); *Neal-Cooper Grain Co. v. Kissinger*, 385 F. Supp. 769, 776 (D.D.C. 1974) (foreign government or instrumentality thereof is a “person”).

²⁴ *Portland Cement Ass’n v. Ruckelshaus*, 486 F. 2d 375, 393 (D.C. Cir. 1976).

²⁵ *Id.*

²⁶ 5 U.S.C. § 553(c).

²⁷ *Am. Mining Cong. v. EPA*, 907 F.2d 1179, 1188 (D. C. Cir. 1990); *Home Box Office, Inc. v. FCC*, 567 F.2d 9, 35-36 (D.C. Cir. 1977) (per curiam). Without such an obligations, courts have said, the opportunity to comment would be “meaningless.” *Id.*; see *Carlson v. Postal Regulatory Comm’n*, 938 F.3d 337, 351 (D.C. Cir. 2019); *ACLU v. FCC*, 823 F.3d 1554, 1581 (D.C. Cir. 1987); *St. James Hosp. v. Heckler*, 760 F.2d 1460, 1470 (7th Cir. 1985).

²⁸ E.g., *La. Fed. Land Bank Ass’n v. Farm Credit Admin.*, 336 F.3d 1975 (D.C. Cir. 2003).

²⁹ 5 U.S.C. § 706(2)(A).

³⁰ See, e.g., *Bus. Roundtable v. SEC*, 647 F.3d 1144, 1152 (D.C. Cir. 2011) (discussing agency failure to address cost issues raised in comments).

³¹ *NRDC v. EPA*, 859 F.2d 156, 188–89 (D.C. Cir. 1988).

³² *Home Box Office, Inc. v. FCC*, 567 F.2d 9, 35, n.58 (D.C. Cir. 1977) (per curiam); see also *Pub. Citizen, Inc. v. FAA*, 988 F.2d 186, 197 (D.C. Cir. 1993).

³³ See E-Government Act § 206, 44 U.S.C. § 3501 note.

malattributed comments.¹¹⁹ Several petitions for review asserted that one of the Order’s legal defects was that it “conflicts with the notice-and-comment requirements of 5 U.S.C. § 553”¹²⁰ but no party actually raised the issue of malattributed comment to the court during briefing.¹²¹

(c) Agency Reliance on Malattributed Comments

The final question is whether it would violate the APA for an agency to read or to rely on a comment submitted under a false name. The rulemaking provision itself, § 553, in no way restricts what the agency can consider or who it can listen to. Rather, any such restriction would rest on the requirement of “reasoned decision-making” embedded in the prohibition on arbitrary and capricious agency action.¹²² Such determinations are case-specific. It would not be reasoned decisionmaking to rely on malattributed comments as a measure of public sentiment or to rely on a comment that purported to be from an authority in a relevant field when it was not. But if a comment is relevant, factually accurate, and communicates something of value, there is nothing arbitrary and capricious in an agency making use of what it has to offer, regardless of whether the sender put someone else’s name on it. Thus, there would seem to be no per se rule allowing or prohibiting agencies to rely on malattributed comments.

A. Policy Issues Raised By Malattributed Comments

In this section, we elaborate on the concerns raised by malattributed comments in particular and also discuss how agencies can discourage submission of malattributed comments and handle malattributed comments once they are discovered. At the same time, we acknowledge that a malattributed comment may nevertheless contain useful content.

1. *Misleading the Agency*

Because malattributed comments, by definition, contain a falsehood, an obvious concern is that the agency may be misled. The misleading might take either of two forms: the agency could be misled with regard to the identity of the commenter, and the agency might be misled as to public opinion, mistakenly viewing the phony comments as indicators of broader public support for a particular position than actually exists.

¹¹⁹ See, e.g., Brian Fung, *FCC net neutrality process ‘corrupted’ by fake comments and vanishing consumer complaints, officials say*, WASH. POST: THE SWITCH (Nov. 24, 2017), <https://www.washingtonpost.com/news/the-switch/wp/2017/11/24/fcc-net-neutrality-process-corrupted-by-fake-comments-and-vanishing-consumer-complaints-officials-say/> (quoting Evan Greer of Fight for the Future as stating that “this will absolutely show up in court if we get there”); see also Karl Bode, *The FCC Is Blocking a Law Enforcement Investigation Into Net Neutrality Comment Fraud*, VICE (Dec. 12, 2017), https://www.vice.com/en_us/article/wjzjv9/net-neutrality-fraud-ny-attorney-general-investigation (“Expect the agency’s failure to police comment fraud to play a starring role in these legal arguments to come.”).

¹²⁰ Pet. for Review at 2, *Ctr. for Democracy and Tech. v. FCC*, No. 18-1068 (D.C. Cir. Mar. 5, 2018); Pet. for Review at 2, *New York v. FCC*, No. 18-1055 (D.C. Cir. Feb. 22, 2018).

¹²¹ The D.C. Circuit’s opinion, which was largely but not entirely in the agency’s favor, does not mention the “fraudulent comments” issue. See *Mozilla Corp. v. FCC*, 940 F.3d 1 (D.C. Cir. 2019).

¹²² See 5 U.S.C. § 706(2)(a).

(a) Commenter Identity

With regard to the first, in general, the agency simply will not notice the name of the commenter. If the agency receives 10,000 very similar, computer-generated comments, no one is paying attention to the names under which they are submitted, whether they are false or real. One cannot be misled by something of which one is unaware. If the submission is not computer-generated—a unique comment filed under a false name—the falsehood is irrelevant for purposes of the agency’s deliberation. The agency will take the comment for what it is worth; the name adds nothing to its weight and will not affect how it is treated.¹²³

Now suppose the name is one that someone in the agency recognizes. This is the “deep fake” scenario described above. For example, it may be an important researcher or advocate in a relevant field, or the general counsel of a prominent regulated entity. In this situation, the identity of the commenter may matter for the agency’s deliberations. The agency could give particular weight to such a comment. But it would be highly unlikely for the falsehood to go unnoticed. The very background knowledge that makes the name recognizable will make it hard for someone to pull the deception off. This is especially the case in a situation where the purported commenter’s interests are well known to the agency, perhaps because of repeat interaction. However, for an agency that does not regulate often, or that regulates only in certain domains infrequently, this might be harder to ensure.

The malattributions that often grab observers’ attention involve using the name of a famous (sometimes dead) person. But these are not misleading because it is apparent that the name is false. For example, in the net neutrality rulemaking, there were multiple submissions from “Barack Obama” and from “Ajit Pai.”¹²⁴ This does not result in any actual deception; no rulemaking official would think that the former President or the FCC Chair had submitted the comment. Same for submissions from “Elvis Presley.”

Our discussions with agency officials are consistent with the foregoing. Their own sense is that consequential instances of pseudonymous submissions are extremely rare, if not nonexistent. Of course, we have not done a thorough study and by definition the victim of a successful deception is unaware of having been deceived. Nonetheless, we credit these statements because they reflect actual experience, they are consistent with what one would expect, and we are unaware of a single demonstrated instance to the contrary.

There is one possible setting, however, in which the concerns about the agency being misled may be more serious. Suppose a comment is not from a recognizable name but asserts that the submitter has particular experience that appropriately goes to the weight of the comment. For

¹²³ See, e.g., Letter from Ajit Pai, Chairman, FCC, to Rep. Michael E. Capuano (Apr. 12, 2018) (“Despite any suggestion that the public comment process was somehow ‘flawed’ or ‘tampered with’ by the alleged submission of comments under false names, any such activity did not affect the Commission’s actual decision-making”); Letter from Thomas M. Johnson, Jr., Gen. Counsel, FCC, to Eric Schneiderman Att’y Gen., New York (Dec. 7, 2017) (“[T]he Commission does not make policy decisions merely by tallying the comments on either side of a proposal to determine what position has greater support, nor does it attribute greater weight to comments based on the submitter’s identity.”).

¹²⁴ See, e.g., “Barack Obama,” ID 1051157755251, Restoring Internet Freedom, WC Docket 17-108 (May 11, 2017), <https://www.fcc.gov/ecfs/filing/1051157755251> (submission from “Barack Obama” of “1600 Pennsylvania Avenue, Washington, DC,” objecting to the “unprecedented regulatory power the Obama Administration imposed on the internet” and “Obama’s . . . power grab”).

example, the commenter claims to have done research in the area, or to possess “situated knowledge,”¹²⁵ or to have had direct personal experience, or to be a person who will be directly regulated or benefitted by the proposed rule. All of those people possess information that members of the general public do not and that the agency may find valuable. They may also have a stake that should counsel caution in taking their assertions at face value. For both reasons, the agency would want to know who the source of the comment is. An anonymous comment that claimed to be from a person in such categories would be somewhat suspect; a signed comment may carry more weight. If the name is a malattribution, and the *actual* submitter does not have the qualifications claimed, there is a real risk of inappropriate reliance on the comment. Moreover, suppose a rulewriter found that comment helpful but wanted to double-check its provenance. An internet search might reveal the falsehood, but it might reveal nothing, or might appear to confirm the biographical claims made in the comment.

This risk seems real but slim. We are not aware of real-world examples of such submissions. That does not mean they have not occurred. In the real world, the malattributed comments that have gotten attention were duplicative rather than bespoke; they do not make individualized claims about the submitter. In addition, the real problem here is not the malattribution so much as the biographical misrepresentation. The malattribution may make it harder to uncover the relevant falsehood but is not itself misleading. Thus, the problem here is actually the distinct one of accuracy in the assertions within comments. It is entirely possible for commenters submitting under their own name to misrepresent their experiences, expertise, or even views. The SEC proxy rule proceeding is an example.¹²⁶

(b) The Weight of Public Support or Opposition to the Proposed Rule

The second concern is that the agency will be misled as to public sentiment. Malattributed comments are often, though not necessarily, a form of computer-generated comments. Such was the case in the Net Neutrality rulemaking, for example. Millions of individuals did not sit down and prepare comments that they submitted under someone else’s name. A handful may have done so, but presumably almost all the malattributed submissions involved a computer taking a prepared text, or writing a text, and then randomly attaching actual names and email addresses to the comment. As with computer-generated comments, part of what observers object to here is that what looks like a set of mass comments submitted by millions of concerned individuals is in fact just the effort of a single submitter. To the extent this is the concern, the *malattribution* is largely irrelevant. Perhaps, however, an agency might think the 100,000 identical comments with different names are more likely to be from different individuals than are 100,000 identical anonymous comments, in which case malattributed computer-generated comments are more misleading. This is especially problematic to the extent that public comments are understood by agencies as providing insight into public sentiment.

¹²⁵ See Cynthia R. Farina, Dmitry Epstein, Josiah Heidt, & Mary J. Newhart, *Knowledge in the People: Rethinking “Value” in Public Rulemaking Participation*, 47 WAKE FOREST L. REV. 1185, 1187–88, 1197 (2012) (describing “situated knowledge” as “information about impacts, ambiguities and gaps, enforceability, contributory causes, unintended consequences, etc. that is known by participants because of their lived experience in the complex reality into which the proposed regulation would be introduced”).

¹²⁶ See *supra* text accompanying notes **Error! Bookmark not defined.**–**Error! Bookmark not defined.**

2. Harms to Individuals

Unlike mass and computer-generated comments, malattributed comments can have impacts outside the agency and the rulemaking process, imposing harms on the people whose names and email addresses are used without permission. Many or most will never be aware that they have supposedly submitted a comment in a federal rulemaking, and many or most may not care. Even if using someone's name and address on a comment does not constitute identity theft under federal law,¹²⁷ it still may be harmful to the person whose name is used in this manner.

Two sorts of harms can be imagined. The first is psychological. It would be understandable that a person who learned that their name was used to submit a comment would be annoyed or angry, especially if they disagreed with the content of the comment. The harm is somewhat abstract; unlike standard identity theft, the victim's bank account is intact. But for some people, the distress or anger will be quite real.¹²⁸ The second possible harm is reputational. For a malattributed comment in a regulatory docket to cause reputational harm, it would have to be noticed by someone who changes their opinion of the purported commenter for the worse. The obscurity of the rulemaking process may make this unlikely, and we are not aware of any instance in which it has occurred. Still, all it takes is one viral tweet by someone with a large following about a comment considered benighted or outrageous to do serious harm to the ostensible author of the comment. Quantifying these harms may well be impossible. Individual views on how seriously to take them will vary. While some people will not care or perceive themselves to be harmed at all, others may see themselves as victims of identity theft.

3. Discouraging Malattributed Comments

The e-rulemaking program has taken several recent steps to discourage the submission of malattributed comments. For example, the user notice on regulations.gov now includes the following under the heading "Terms of Participation":

Public comments help agencies develop regulations; we encourage comments from all viewpoints. Comments submitted to Regulations.gov *should be the submitter's own comments or be submitted with the commenter's permission*. The development of federal regulations is within the jurisdiction of the U.S. Government's executive branch agencies. *It is a violation of federal law to knowingly and willfully make a materially false, fictitious, or fraudulent statement or representation including false statements about your identity or your authority to submit a comment on someone*

¹²⁷ See Herz, *supra* note 43, at 51–55.

¹²⁸ For examples, see Letter from Brittany Ainsworth et al., to Ajit Pai, Chairman, FCC (May 25, 2017) (letter from 27 individuals whose names and email addresses were used to submit comments without their involvement or permission complaining that "someone stole our names and addresses, publicly exposed our private information without our permission, and used our identities to file a political statement we did not sign onto" and calling on the agency to remove these "fraudulent comments" from the docket and notify "all proper authorities"); Bode, *supra* note 119 (complaining that "the agency told me there was nothing it could do after someone hijacked my identity to claim I falsely supported killing net neutrality protections"), <https://www.vice.com/en/article/wjzjv9/net-neutrality-fraud-ny-attorney-general-investigation>; Press Release, A.G. Schneiderman Releases New Details On Investigation Into Fake Net Neutrality Comments (Dec. 13, 2017), <https://ag.ny.gov/press-release/2017/ag-schneiderman-releases-new-details-investigation-fake-net-neutrality-comments> (quoting, among others, unidentified individual as saying "I'm sick to my stomach knowing that somebody stole my identity and used it to push a viewpoint that I do not hold").

else's behalf, in relation to the development of such federal regulations, including through comments submitted on Regulations.gov. See 18 U.S.C. § 1001.

Subject to 18 U.S.C. § 1028(c), *it is also a violation of federal law to knowingly use, without lawful authority, a means of identification of another person in connection* with the violation of any federal law or the commission of a felony under state or local law. See 18 U.S.C. § 1028(a)(7).

By clicking the submit button, you are verifying that *you are not making any materially false, fictitious, or fraudulent statement or representation regarding your identity* or your authority to submit on someone else's behalf with regard to the comment you are submitting on Regulations.gov, and that you are not using, without lawful authority, a means of identification of another person, real or fictitious, in connection with any comment you are submitting on Regulations.gov.¹²⁹

This notice implies that a malattributed comment could be a federal crime. This may be part of a deterrence strategy on the part of the government to discourage anyone from sending malattributed comments. Whether users are likely to see this, in a user notice that contains several other paragraphs of policies, is uncertain. Moreover, although it may discourage some individual submitters from using a false name, it is unlikely to have any impact on large-scale operations.

Agencies have to make decisions about how to treat malattributed comments, once suspected or discovered. Because of the novelty of this issue, many agencies do not have protocols for how to resolve whether a comment is malattributed and, if so, policies on how to handle that comment in the docket. Question that such policies would address include whether an agency should strive to resolve a question about a comment's provenance, or merely flag the potential issue. Also, and in line with the user notice described above, to the extent that criminal action is under consideration for particular malattributed comments, agencies may need to make staff available to assist with any investigations or prosecutions.

B. Technological Responses to Malattributed Comments

The technology readily exists to authenticate users and is in widespread use in many contexts. Common techniques include secure login, two-factor authentication, biometric authentication using facial recognition or fingerprint, answering security questions or verifying names against a database (as is the case in voter registration), or clicking an additional "I agree" button to acknowledge and agree to terms of service. However, agencies currently do not have the technology in place to authenticate those filing comments in the way a government department authenticates someone applying for a driver's license or a commercial website authenticates someone buying a product to prevent credit card fraud.

While tools are not in place to authenticate someone's identity, agencies do use tools to ensure that a commenter is a human instead of a bot. These tools are primarily a response to

¹²⁹ *User Notice*, REGULATIONS.GOV, <https://www.regulations.gov/user-notice> (emphases added).

computer-generated comments but, by imposing a “speed bump” on the commenting process, they may also help to reduce malattributed comments.

The addition of reCAPTCHA to regulations.gov is intended to help to “improve[] the integrity of the commenting process.”¹³⁰ CAPTCHA is an example of a “Turing Test”—a thought experiment developed by Alan Turing to evaluate artificial intelligence—and stands for “Completely Automated Public Turing Test to tell Computers and Humans Apart.”¹³¹ With CAPTCHA, users are presented an image of a set of visually distorted letters and numbers and asked to enter the same characters into a textbox. When CAPTCHA was invented nearly two decades ago, it was believed that machines would not be able to complete this task since only humans would be able to interpret what the distorted characters were. With advances in computing power this is no longer true and techniques to defeat CAPTCHA have been created. CAPTCHA, however, has also been reinvented to protect against these attacks. In 2018, Google announced “reCAPTCHA v.3” which eliminates the need for any human interaction with CAPTCHA at all. By using risk analysis algorithms that assign a “risk score” to every person browsing a website using the tool, the software alerts administrators if fraudulent activity is detected.¹³²

Also, regulations.gov now includes a comment application programming interface (API) to allow authorized entities to post mass comment campaigns to Regulations.gov if they have been verified by GSA using a commercial identity validation service. In the press release announcing these changes, GSA indicated that this was “to assure such entities ‘are who they say they are.’”¹³³ The service does not aim to verify the identities of individual commenters, however.

The public prominence of malattributed comments prompts a fresh look at whether agencies should verify commenter identity, either on the front-end or after either an internal or external review flags a comment as potentially malattributed. While authentication is a common practice and technically straightforward in many circumstances, the practice would be in tension with agency policies to permit anonymous comments.

V. Computer-Generated Comments

A. Legal Issues Raised by Computer-Generated Comments

The APA requires agencies to provide an opportunity to comment to “interested persons.”¹³⁴ The term “interested” is undefined and is generally understood not to limit the scope

¹³⁰ Press Release, Gen. Servs. Admin., GSA Launches Updated Regulations.gov to Improve the Integrity (Feb. 17, 2021), <https://www.gsa.gov/about-us/newsroom/news-releases/gsa-launches-updated-regulationsgov-to-improve-the-integrity-of-public-commenting-02172021#:~:text=WASHINGTON%20%E2%80%94%20The%20U.S.%20General%20Services.gov%20launching%20February%2018%2C%202021.&text=%E2%80%9CThe%20new%20Regulations.gov%20re, and%20mobile%2Dfriendly%20interface.%E2%80%9D>.

¹³¹ The “imitation game” experiment proposed by Turing was invented as a way of approaching the hard question of “Can machines think?”. See A.M. Turing, *Computing Machinery and Intelligence*, 59 MIND 433 (1950).

¹³² reCAPTCHA: Easy on Humans, Hard on Bots, <https://www.google.com/recaptcha/intro/v3.html?ref=techmoon> (last visited Apr. 1, 2021).

¹³³ See Gen. Servs. Admin., *supra* note 130.

¹³⁴ 5 U.S.C. § 553(b).

of potential commenters.¹³⁵ The term “persons” is defined as follows: “person includes an individual, partnership, corporation, association, or public or private organization other than an agency.”¹³⁶ When Congress passed the APA, it would not have contemplated that a computer might send a comment. But the definition is instructively broad; it is not limited to natural persons, and courts have read the word capaciously.¹³⁷ Moreover, because a person must set a computer-generated comment in motion, the section 551 definition is arguably met in any event.

As described above, agencies are required to respond to significant issues raised in comments. As of this writing, no courts appear to have interpreted this requirement in light of computer-generated comments. During the interviews, agency staff expressed skepticism that a computer-generated comment would bring content or issues to the rulemaking docket that were not otherwise raised by other comments. But these staff also expressed their commitment to reviewing all comments, regardless of origin, to ensure compliance with their obligations to consider and respond to comments.

It is theoretically possible—if highly unlikely at this time—that a person would challenge an agency action on the basis of its failure to adequately account for the substance of a computer-generated comment that was not otherwise presented in other comments. Should such circumstances arise, courts may determine whether the suit should move forward based on factors such as whether the petitioner can demonstrate the reliability or authenticity of the computer-generated comment.¹³⁸ It is also possible that authentication technology might exclude either computer-generated comments or ordinary comments that raise unique significant issues. If the agency’s obligation to consider and respond to significant comments does not change in such circumstances, technological means of identifying computer-generated comments would have to account for this overarching obligation.

B. Policy and Technical Issues Raised by Computer-Generated Comments

The policy issues raised by computer-generated comments overlap significantly with those already identified for mass and malattributed comments. For example, the presence of computer-generated comments may undermine public confidence in the rulemaking process or draw down agency resources. Many of the issues presented by computer-generated comments, however are technical. First, one issue is the ability of agencies to identify computer-generated comments. In 2019, an experiment demonstrated the ease with which bots mimic human speech, therefore making it difficult to distinguish computer-generated comments from comments directly submitted

¹³⁵ See Herz, *supra* note 73, at 357–58.

¹³⁶ 5 U.S.C. § 551(1).

¹³⁷ See *e.g.*, O’Rourke v. U.S. Dep’t of Justice, 684 F. Supp. 716, 718 (D.D.C. 1988) (holding that “person” includes non-citizens and collecting cases); Neal-Cooper Grain Co. v. Kissinger, 385 F. Supp. 769, 776 (D.D.C. 1974) (foreign government or instrumentality thereof is “person”).

¹³⁸ In the evidentiary context, courts have managed to assess the admissibility of electronically stored information (which may include computer-generated information) on the basis of the Federal Rules of Evidence; for example, proponents must demonstrate the information’s relevance, reliability, authenticity, and so on. See, *e.g.*, Lorraine v. Markel Am. Ins. Co., 241 F.R.D. 534, 538 (D. Md. 2007). This is not to suggest that the Federal Rules apply to administrative records; rather, this example is offered to demonstrate that courts may find useful analogies that may be applied consistently with their equitable powers and authority under the APA.

by persons.¹³⁹ The focus of this experiment was a comment period on a waiver from federal requirements requested by the Idaho Medicaid program. A text generation model was utilized to submit one-thousand comments on the proposed waiver. The inputs for this model were thousands of comments submitted in response to Medicaid waivers previously requested by a number of other states. These inputs were used to train the model to employ search-and-replace techniques as a means of generating comments, which were submitted automatically to the Centers for Medicare and Medicaid Services at random intervals.

Following the submission of the computer-generated comments, subjects were recruited to judge whether particular comments in the docket were submitted by a bot or human. On average, the respondents—all of whom had previously demonstrated competency in identifying conspicuous bot texts—correctly classified less than half of the comments. Performance was particularly poor in the context of computer-generated comments, in that less than one-third were correctly recognized. These results indicate that the computer-generated comments were as a general matter plausibly human, therefore making consistent sorting of such submissions a non-trivial exercise for the agency. At the conclusion of the experiment, the researcher revealed the computer-generated comments and requested that CMS withdraw the bot submissions from consideration.

These results are consistent with assessments of computer-generated comments that have occurred outside of the context of experimentation. A variety of analyses have emphasized that search-and-replace algorithms, and the resulting comment-to-comment variation in content, enhance the difficulty of identifying computer-generated comments. As a result, “analysts have struggled to pinpoint” the precise frequency with which computer-generated comments occur.¹⁴⁰

Are there approaches for identifying computer-generated comments in a systematic manner? The FCC’s net neutrality policy is a good place to turn in this regard, as researchers have expended considerable energy identifying computer-generated comments that were submitted in this particular rulemaking. Note that these approaches entail identifying computer-generated comments in hindsight, as opposed to screening for such comments during the intake process.

One analysis focused on the text of net neutrality comments, searching for expressions regularly contained in submissions.¹⁴¹ The analysis discovered combinations of phrases consistent with the automated deployment of search-and-replace algorithms. Take, for example, the following comment excerpt: “Americans, as opposed to Washington bureaucrats, deserve to enjoy the services they desire.” This sentence repeatedly appeared in comments in numerous other permutations, with “Americans” replaced by terms such as “people like me” and “individual citizens.” Similarly, “the FCC,” “so-called experts,” and other analogous phrases substituted for “Washington bureaucrats.” One result of this automation was the submission of large numbers of comments that, while not identical, conveyed essentially equivalent sentiments. Another

¹³⁹ Max Weiss, *Deepfake Bot Submissions to Federal Public Comment Websites Cannot Be Distinguished from Human Submissions*, TECH. SCI. (Dec. 17, 2019), <https://techscience.org/a/2019121801/>.

¹⁴⁰ Katherine Krems, *supra* note 116, at 71.

¹⁴¹ Jeff Kao, *More Than a Million Pro-Repeal Net Neutrality Comments Were Likely Faked*, HACKERNOON (Nov. 22, 2017), <https://hackernoon.com/more-than-a-million-pro-repeal-net-neutrality-comments-were-likely-faked-e9f0e3ed36a6>.

characteristic of this process was the brevity of the resulting computer-generated comments.¹⁴² Increases in comment length multiply opportunities “for the appearance of ‘tells’ (e.g., repeated words, incorrect grammar, nonsensical sentiment) that the comment was not created by a human.”¹⁴³

Other analysis has examined over-time patterns of the submission of net neutrality comments with identical and near-duplicate content, an approach that is useful for identifying mass comment campaigns (regardless of human or computer submission). One such pattern is the receipt of large numbers of comments at precisely the same moment.¹⁴⁴ Researchers discovered, for example, that on “nine different occasions, more than 75,000 comments were submitted at the very same second—often including identical or highly similar comments.”¹⁴⁵ Another pattern is embodied by the submission of the following comment excerpt: “The unprecedented regulatory power the Obama Administration imposed on the internet is smothering innovation, damaging the American economy and obstructing job creation.” This text occurred in approximately a half-million comments.¹⁴⁶ These comments were submitted at near-constant rates for given periods, which were punctuated by interludes during which no such comments were received.¹⁴⁷ This cycle suggests that bots were turning on and off at specified intervals.

Another indication of the submission of computer-generated comments was repetition in email addresses, in particular domains and locations exhibiting behavior inconsistent with human messaging activity. The FCC, for example, determined that millions of comments were the product of websites that produce one-off emails and are unable to receive messages. The agency also discovered that hundreds of thousands of emails originated “from the same address in Russia.”¹⁴⁸

The regular submission of computer-generated comments was also suggested by the nature of the information submitted along with the comments themselves. When humans fill out information, the resulting inputs are typically inconsistent. For example, name, address, and email fields are often left blank, and individuals utilize varying formats. In the context of the “unprecedented regulatory power” comments referenced earlier, however, fewer than ten submissions failed to contain complete information. Furthermore, these names, addresses, and emails exhibited unusual similarity in presentation. Finally, exceedingly few comments requested that the FCC provide email confirmation of receipt. These attributes suggest that algorithms, rather than humans, were the immediate sources of the submitted information.¹⁴⁹

¹⁴² See Krems, *supra* note 116, at 75.

¹⁴³ Weiss, *supra* note 139.

¹⁴⁴ DAVID FREEMAN ENGSTROM ET AL., GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 60 (2020), <https://www-cdn.law.stanford.edu/wp-content/uploads/2020/02/ACUS-AI-Report.pdf>.

¹⁴⁵ Hitlin, Olmstead, & Toor, *supra* note 4, at 3.

¹⁴⁶ See Chris Sinchok, *An Analysis of the Anti-Title II Bots*, MEDIUM, <https://medium.com/@csinchok/an-analysis-of-the-anti-title-ii-bots-463f184829bc>.

¹⁴⁷ See *id.*

¹⁴⁸ Grimaldi & Overberg, *supra* note 6.

¹⁴⁹ See Sinchok, *supra* note 146.

1. *Current Agency Practices*

In the interviews, agency staff expressed their awareness of computer-generated comments having been submitted in a few rulemaking proceedings. Despite this awareness, the staff we interviewed did not report systematic approaches to identify computer-generated comments. One agency discovered computer-generated comments through a *Wall Street Journal* report on the rulemaking, as well as the rulemaking team's identification of a number of unusual comments. These comments consisted of strings of nonsensical words, which made the agency suspicious that the submissions were not generated by humans.

Despite the availability of tools discussed earlier under mass and malattributed comments, implementing approaches to systematically identify computer-generated comments was not a high priority for the agencies we interviewed. Agency staff characterized the discovery of computer-generated comments as requiring substantial effort, a resource-intensive undertaking that is not worth the dedication of agency bandwidth. In general, the interviews revealed that agencies are not focused on the issue of computer-generated submissions in and of themselves. Rather, they indicated greater concern about mass and malattributed comments whose detrimental attributes may be deepened by computer generation. Despite this concern, the agency staff we interviewed reported as their primary concern the need to identify and respond to significant issues that comments raise, regardless of a given comment's source.

One reason for the lack of attention to computer-generated comments in and of themselves may be that agencies are already using de-duplication tools to address mass comments. Computer-generated comments, in other words, are not seen as creating problems in rulemaking proceedings other than increasing the volume of comments received by agencies—thereby turning the matter into one of mass comment management. As an agency official put it during an interview, computer-generated comments essentially present agencies with a de-duplication task. With the utilization of de-duplication software, the unique content in computer-generated comments can be readily identified.

As this perception indicates, agency staff generally expressed that the generation of comments by computers is not, in and of itself, an important attribute of submissions. The point was repeatedly made during the interviews that it is the substance of comments that matters, as opposed to the identity of submitters or the volume of comments. Agencies were not overly concerned that computer-generated comments convey insights that have not separately been communicated through human comments. That said, agencies emphasized that they would not exclude a computer-generated comment on the basis of the source of the submission, but would consider whether it raised significant issues requiring agency consideration.

2. *Threat Versus Practice*

In sum, there is currently a disjuncture between conceptions of computer-generated comments from the vantage points of technologists and the agency staff we interviewed. Technologists warn of a present—and especially future—in which computer-generated comments effectively mimic human content, thereby making prevention and detection an impossibility. The agency staff we interviewed, by contrast, were not overly concerned with such scenarios at this time. They saw de-duplication tools as being adequate for the task and did not seem anxious to experiment with additional technologies to streamline the comment review process. Notwithstanding current perceptions, in the years ahead, it will be important to monitor whether

the technologies that enable mass, malattributed, and computer-generated comments threaten to undermine the perceived legitimacy of the notice-and-comment process and the ability of agency officials to make sense of and consider comments thoughtfully.

VI. Innovations to Enhance Participation and Commenting

The foregoing discussion has identified some of the risks associated with mass, malattributed, and computer-generated comments. These risks are real, and agencies must undertake appropriate measures to ensure that they protect the integrity and value of the notice-and-comment process. But technology can present opportunities as well as challenges. As we have already seen, agencies extensively use de-duplication software to help them process mass comment campaigns. And this is only a preview of what agencies can accomplish with newly emerging technologies. This section explores technologies that are not yet in widespread use but that might enhance and supplement the notice-and-comment process.¹⁵⁰

A. Summarization Technologies and Enhancing the Value of Public Commenting

Although our focus has been on mass, malattributed, and computer-generated comments, an additional salient and urgent opportunity for regulators is using new technologies to enhance the process of reviewing public comments.¹⁵¹

While our interviews found that agencies are not currently using such tools, affordable technology is on the horizon to help agencies more easily make sense of public comments, helping those reading the rules, not simply to save time, but to better analyze, spot patterns in, and understand public comments.

The GSA’s innovative technology unit 18F comments: “It takes enormous amounts of staff time, resources, and taxpayer dollars to manually analyze written public comments submitted to agencies through various committees and many other channels.” Their experts go on to recommend the need to “explore if Natural Language Processing and Artificial Intelligence can semi-automate, streamline, and expedite the public comment process, and whether any additional policy or guidance might be required to create a standard approach.”¹⁵²

NLP techniques can help comment reviewers both summarize and sort comments, helping them to extract the most important substantive information from the comments.¹⁵³ For example, these techniques can be used to identify those parts of comments that bear on questions that are of

¹⁵⁰ See generally Beth Simone Noveck, *The Innovative State*, DAEDALUS (forthcoming 2021) (special issue on the administrative state in the United States in the twenty-first century).

¹⁵¹ See *Fake It Till They Make It*, *supra* note 2 (testimony of Beth Simone Noveck), <https://financialservices.house.gov/uploadedfiles/hhrg-116-ba09-wstate-noveckb-20200206.pdf>.

¹⁵² Aditi Rao, Ben Peterson, & Andrew Suprenant, *Synthesizing Public Comments: Phase 2 Report* (on file with authors).

¹⁵³ See Livermore, Eidelman, & Grom, *supra* note 17, at 980 (discussing “needle-in-the-haystack” and “forest-for-the-trees” challenges of mass rulemakings).

particular interest to rule-writers, or that contain relevant legal, technical, or operational information.¹⁵⁴

While still a challenging task, researchers and entrepreneurs have developed tools for summarization, including shortening and extracting the most relevant portions of documents. To sort information, one technique that can be used is topic modeling. In brief, a topic model is a computational text analysis technique that extracts patterns in the semantic content in a corpus of documents, generating a list of topics (which are distributions over the vocabulary in a corpus) and characterizing every document as a distribution over those topics.¹⁵⁵ Topic modeling makes it possible to automatically and quickly sort textual information into semantic categories.

Both Google and Microsoft announced in 2019 that they had built systems capable of summarizing an enormous range of texts, including news, fictional stories, instructions, emails, patents, and legislative bills.¹⁵⁶ The MIT Center for Constructive Communication conducted

¹⁵⁴ The field of NLP encompasses a wide range of technologies that use computational tools to convert natural language artifacts into a format that can be processed and analyzed using computational and statistical tools. NLP techniques include deduping software as well as:

- Flesch-Kincaid Readability: a measure of the difficulty or clarity of written English. The readability score of a text is based on the average number of words per sentence and the average number of syllables per word. Other readability metrics include the Gunning Fog Index and the Spache Index.
- Linguistic Inquiry and Word Count (LIWC): a software application that counts “words in psychology-relevant categories,” such as whether words are associated with honesty or deception or track individual thinking styles.
- Plagiarism Detection: a technique for detecting similarity in written texts, with the goals of identifying plagiarism or copyright infringement.
- Automated Document Summarization: an application that processes larger texts, or multiple texts, as inputs with the goal of generating summary texts that convey a condensed version of the original input texts.
- Sentiment Analysis: a measure of words based on positive or negative valence, as a way to estimate the opinions or attitudes expressed in a written text.
- Topic modeling: a family of computational tools used to discover the latent thematic structure within a collection of documents.
- Word Embeddings: a technique for mapping words or phrases into a vector space that compactly represents semantic content. One technique for generating word embeddings involved “skip-gram” where a model is trained to use a word to predict surrounding words in a document.

¹⁵⁵ For key works in the relevant NLP literature, see, e.g., Kincaid et al., *Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel*, Research Branch Report 8-75 (1975); Adam Feldman, *Opinion Clarity in State and Federal Trial Courts*, in *LAW AS DATA: TEXT, COMPUTATION, AND THE FUTURE OF LEGAL ANALYSIS* 407, 415–18 (Michael A. Livermore & Daniel N. Rockmore eds., 2019); Yla R. Tausczik & James W. Pennebaker, *The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods*, 29J. LANGUAGE & SOC. PSYCH. 24 (2010); BING LIU, *SENTIMENT ANALYSIS: MINING OPINIONS, SENTIMENTS, AND EMOTIONS* (2015). David M. Blei, Andrew Y. Ng, & Michael I. Jordan, *LATENT DIRICHLET ALLOCATION*, 3 J. MACH. LEARNING RES. 993 (2003); Mikolov et al., *Distributed Representations of Words and Phrases and Their Compositionality*, in *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS* 26, 3136 (2013).

¹⁵⁶ Patrick Fernandes, Miltiadis Allamanis, & Marc Brockschmidt, *Structured Neural Summarization*, published at Int’l Conf. on Learning Representations (Feb. 2019), <https://arxiv.org/pdf/1811.01824.pdf>; see also Peter Liu & Xin Pan, *Text summarization with TensorFlow*, GOOGLE AI BLOG (Aug. 24, 2016), <https://ai.googleblog.com/2016/08/text-summarization-with-tensorflow.html>.

research on large-scale Twitter data sets.¹⁵⁷ Its Electome project, for example, extracts semantic content from the entire corpus of Twitter data—billions of tweets—in order to summarize the core political messages of the day and help drive election coverage.¹⁵⁸

Such summarization and sorting processes sometimes combine automation with human intelligence to make quick work of large data stores and overcome the biases that arise from using automation alone. Journalists took advantage of such tools, for example, when they needed to rapidly sift through the 13.4 million documents that comprised the so-called “Paradise Papers.”¹⁵⁹ Public institutions have also used natural language data analytical techniques to make sense of social media data. To help UNICEF and other actors craft more effective pro-immunization messaging programs, researchers set out to monitor social media networks, including blogging platforms, forums, Facebook, Twitter, Tumblr, and YouTube. They sought to analyze prevalent conversation themes according to volume, types of engagement, and demographics; to identify influencers across languages and platforms; and to develop specific recommendations for improving messaging strategies across languages, platforms, and conversation themes. The research methodology involved scraping text from conversations on social media platforms in English, Russian, Polish, and Romanian, in order to be able to summarize them and identify key patterns.¹⁶⁰

A recent State Department project offers a simple illustration of how agencies could make sense of rulemaking comments using a combination of artificial intelligence and human oversight. In 2016, the State Department sought to improve its passport application and renewal process in anticipation of an increase in the number of passport application and renewal forms. It ran an online public engagement process to ask people what improvements they wanted, receiving almost 1,000 suggestions. In order to make rapid sense of those submissions, it used a third-party software company, which applied a text-mining algorithm that scanned the highlighted text for responses containing similar keywords in order to create summaries. The public was invited to proof and make suggestions for how to improve those highlights, adding accountability but in a way that is efficient. The combination of human and machine intelligence made it faster and easier to summarize content than using an algorithm alone.

To date, application of NLP to public comments in administrative rulemaking has been largely limited to de-duplication. While still under development, more advanced NLP techniques could eventually assist agency personnel in identifying relevant substantive content within comments and summarizing the information presented across a broad spectrum of comments. One of the challenges for deploying summarizing technology in the context of rulemaking is that there is often domain specific language that requires retraining the relevant models. However, for

¹⁵⁷ The Laboratory for Social Machines, which carried out some of this research, was incorporated into the MIT Center for Constructive Communication in 2021. See LABORATORY FOR SOCIAL MACHINES, MIT MEDIA LAB, <https://www.media.mit.edu/groups/social-machines/overview/>.

¹⁵⁸ See THE ELECTOME, <http://www.electome.org/>.

¹⁵⁹ Fabiola Torres López, *How They Did It: Methods and Tools Used to Investigate the Paradise Papers*, GLOB. INVESTIGATIVE JOURNALISM NETWORK (Dec. 4, 2017), <https://gijn.org/2017/12/04/paradise-papers/>.

¹⁶⁰ STEFAAN G. VERHULST & ANDREW YOUNG, THE POTENTIAL OF SOCIAL MEDIA INTELLIGENCE TO IMPROVE PEOPLE’S LIVES (Sept. 24, 2017) (report for The Governance Lab), <http://www.thegovlab.org/static/files/publications/social-media-data>.

important rulemakings likely to receive a large number of comments, this investment may well be worth it. Furthermore, the addition of human oversight can provide a check on the performance of machine learning applications, making it possible to evaluate and confirm the reliability of new tools for summarization. While NLP tools can be used to augment, rather than replace, human review, the agency staff we interviewed expressed concerns related to how the use of new technologies might interact with their legal obligation to review and respond to comments. This legal uncertainty creates the risk that agencies may innovate slowly. Depending on their risk tolerance, it may prevent them from adopting these technologies at all.

B. CrowdLaw: Innovations in Equitable Participation

In addition to improving the commenting process ex-post using new technologies, agencies could also explore using complementary platforms and processes—ones already well-honed and tested by other governments—to create new opportunities for public engagement, especially to solicit information and expertise from more diverse and varied audiences as a complement to notice-and-comment. Building on ACUS’ earlier work, we conclude our discussion of public participation in rulemaking by looking at several contemporary examples of how governments are enhancing citizen participation using new technology.¹⁶¹

Over the last decade, federal agencies have expanded citizen engagement through the use of prize-backed challenges or what is sometimes called open innovation via the Challenge.gov website. Since 2011 with the reauthorization of the America Competes Act,¹⁶² a hundred federal agencies have run online challenges to tap the intelligence and expertise of the public.¹⁶³ NASA has regularly used prize-backed challenges to spur crowdsourcing of innovative solutions from the public. The Asteroid Grand Challenge, for example, was focused on finding all asteroid threats to human populations.¹⁶⁴ Prize-backed challenges require agencies to articulate and define exactly what information they need from the public and provide very transparent and specific criteria for evaluating public submissions. With ten years of experience with prized backed challenges, there may be useful insights for federal agencies to draw about how to improve public participation in agency decision making.

¹⁶¹ As Michael Herz wrote in a 2013 ACUS report:

[T]he online world in general has come to be increasingly characterized by participatory and dialogic activities, with a move from static, text-based websites to dynamic, multi-media platforms with large amounts of user-generated content. At the heart of this move to “Web 2.0” have been social media, blogs, Twitter, Facebook, YouTube, IdeaScale, wikis, Flickr, Tumblr, and the like. Outside the rulemaking setting, federal, state, and local governments have enthusiastically jumped on the social media bandwagon.

Michael Herz, USING SOCIAL MEDIA IN RULEMAKING: POSSIBILITIES AND BARRIERS (Nov. 2013) (report to Admin. Conf. of the U.S.), <https://www.acus.gov/sites/default/files/documents/Herz%20Social%20Media%20Final%20Report.pdf>.

¹⁶² America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Reauthorization Act of 2010, Pub. L. No. 111-358, 124 Stat. 3982 (Jan. 4, 2011).

¹⁶³ *About*, CHALLENGE.GOV, <https://www.challenge.gov/about> (last visited Apr. 1, 2021).

¹⁶⁴ *Asteroid Grand Challenge*, NASA.GOV, <https://www.nasa.gov/content/asteroid-grand-challenge> (last visited Apr. 1, 2021).

Challenge.gov is one example of institutionalized public engagement or what is sometimes referred to as “CrowdLaw,” namely the use of technology to engage the public in law-, rule-, or policy-making. It is the idea that public institutions work better when they increase citizen engagement by using new technologies to obtain diverse sources of information, insight and expertise at each stage of the law and policymaking cycle to improve the quality as well as the legitimacy of the resulting laws, regulations, and policies, especially by engaging with underrepresented communities.¹⁶⁵ CrowdLaw does not describe one form of participation. Rather, it describes a variety of different methods, tools and platforms that institutions use.

Expert sourcing, where officials crowdsource expert advice, is one example of how government bodies are implementing more citizen engagement. The Federation of American Scientists’ Congressional Science Policy Initiative invites hundreds of scientists to help draft questions for Members of Congress to ask of committee witnesses. Such crowdsourcing, facilitated by new technology, helps beleaguered staffers write more informed questions.¹⁶⁶ The Governance Lab at NYU uses videoconferencing to help coordinate online dialogues among experts to advise government officials on a variety of topics. In Fall 2020, for example, it ran six deliberative sessions at the behest of seven governments in Latin America to help them develop implementable strategies for responding to specific public health challenges, including the improvement of mental health services and combating misinformation.¹⁶⁷

Some jurisdictions have used online collaborative drafting processes and platforms to write policies and rules *with* the public, especially with expert members of the public. Instead of an advisory committee or hearing with a handful of experts or writing rules entirely behind closed doors, online collaborative annotation makes it possible to hear from a broader and deeper range of experts and to focus their participation on specific comments on a document. In 2018, the German government used an annotation platform to “expert source” feedback on its draft artificial intelligence policy. By putting the draft on Hypothes.is, a free and open-source annotation tool, the German Chancellor’s Office, working in collaboration with Harvard University’s Berkman Center for Internet and Society, was able to solicit the input of global legal, technology and policy experts. Using an annotation platform also made it possible for people to see one another’s feedback, instead of a series of disconnected comments. One could envision an agency using collaborative annotation to invite experts to annotate and comment on the text of a draft rule.

Many governments are experimenting with the use of random samples of members of the public as a mechanism to obtain more legitimate forms of participation. New technology is making it easier to assemble these representative samples of citizens, known as mini-publics, to weigh in on a governing process. Small groups are known as citizen juries while larger random samples are called citizen assemblies. For example, in the Brussels-Capitol region, a random sample of citizen

¹⁶⁵ Victòria Alsina & José Luis Martí, *The Birth of the CrowdLaw Movement: Tech-Based Citizen Participation, Legitimacy and the Quality of Lawmaking*, 40 ANALYSE & KRITIK 337 (2018); see also Beth Simone Noveck, *CrowdLaw*, in THE PALGRAVE ENCYCLOPEDIA OF INTEREST GROUPS, LOBBYING AND PUBLIC AFFAIRS (Phil Harris et al. eds.) (forthcoming 2021).

¹⁶⁶ CONGRESSIONAL SCIENCE POLICY INITIATIVE, <https://fas.org/congressional-science-policy-initiative/> (last visited Apr. 1, 2021).

¹⁶⁷ See *Smarter Crowdsourcing: Coronavirus*, THE GOVERNANCE LAB, <https://coronavirus.smartercrowdsourcing.org> (last visited Apr. 1, 2021).

representatives serves on each parliamentary committee. Citizens ask questions and provide advice.¹⁶⁸ These processes could also be designed to elicit expertise and know-how relevant to agency decision makers.

Similarly, some have suggested ideas such as administrative agencies empaneling a thousand randomly selected citizens to provide oversight over agency decisionmaking.¹⁶⁹ A variation on this idea would use citizen juries to solicit information on agency agenda setting and priorities,¹⁷⁰ providing the citizen jurors with background materials generated by deliberative polling before their discussions.¹⁷¹

Finally, instead of selecting a random sample, other institutions have relied on self-selected participation using a variety of tools. In month-long online exercises known as “Evidence Checks,” UK parliamentary committees invite experts, stakeholders, and members of the public to comment on the validity of evidence on which a policy is based. The process begins when government departments supply information to their respective committees about an issue. Each committee publishes the information on a parliament.uk web page, and it is scrutinized by a wider pool of invitees. The committee also presents specific questions and problems that it would like participants to address. In contrast to a representative sample, this process allows a group of people with relevant experience and expertise to identify gaps in research that require further review.¹⁷²

Another example of self-selected participation was initiated by the New Jersey Department of Education in March 2021 when that agency invited students, parents and educators across the state to help inform the Department’s policymaking by responding to questions via All Our Ideas, a free platform developed at Princeton University. All Our Ideas has been used in over 18,000 citizen engagement projects.¹⁷³ The owner of the consultation uses the platform to write a series of statements that are then randomly presented to the participant. People select the response they prefer (or “I can’t decide” as a third answer) or they may submit their own response. As people are repeatedly selecting between two randomly generated options, it is a faster and easier mechanism for responding to a series of questions. This so-called “wiki survey” method of showing people two pieces of information and having them choose between them and/or submit a new item offers efficiency benefits over open-ended commenting and can be designed to draw on participant expertise.

¹⁶⁸ The Governance Lab, *Belgian Sortition Models: Institutionalizing Deliberative Democracy*, CROWDLAW FOR CONGRESS, <https://congress.crowd.law/case-belgian-sortition-models.html> (last visited Apr. 1, 2021).

¹⁶⁹ See David R. Arkush, *Direct Republicanism in the Administrative Process*, 81 GEO. WASH. L. REV. 1458 (2013).

¹⁷⁰ BETH SIMONE NOVECK, SMART CITIZENS, SMARTER STATE: THE TECHNOLOGIES OF EXPERTISE AND THE FUTURE OF GOVERNING 220 (2015).

¹⁷¹ See Reeve T. Bull, *Making the Administrative State “Safe for Democracy”*: A Theoretical and Practical Analysis of Citizen Participation in Agency Decisionmaking, 65 ADMIN. L. REV. 611 (2014).

¹⁷² Nesta, *UK Parliament Evidence Checks*, <https://www.nesta.org.uk/feature/six-pioneers-digital-democracy/uk-parliament-evidence-checks> (last visited Mar. 28, 2021).

¹⁷³ MATTHEW SALGANIK, BIT BY BIT: SOCIAL RESEARCH IN THE DIGITAL AGE 111–15 (Princeton Univ. Press, 2017).

Recommendations

To help facilitate committee deliberations, we offer the following draft recommendations. The general categories are technology, coordination and training, docket management, and transparency. We could conceive of many variations of the ideas below, but offer these to help the committee formulate its views.

Based on our research, we believe that mass, malattributed, and computer-generated comments do not, at least currently, fundamentally undermine the notice-and-comment process. However, such comments raise issues of sufficient significance that steps can and should be taken to mitigate the difficulties emanating from them. Technology also presents opportunities for enhancing public engagement in rulemaking.

The following recommendations lay out a variety of immediate and long-run actions for reducing the challenges of mass, malattributed, and computer-generated comments and taking advantage of technology-enabled participation.

Technology

Mass, malattributed, and computer-generated comments raise challenges for agencies at two stages of the rulemaking process.

The first stage is comment submission. At this stage, one difficulty stems from the submission of large numbers of comments. The GSA has recently taken important actions to help agencies manage mass comments. For example, the current version of Regulations.gov includes an API that facilitates the submission of comments in bulk. Another difficulty concerns the authentication of the identity of the commenter. In this regard, Regulations.gov has implemented two features. One feature is identity validation in the API, which enables authorized users to submit comments in bulk. The other feature is reCAPTCHA, which is designed to screen out commenters who are not humans.

We **recommend** that agencies, both those that use Regulations.gov and those that do not, consider utilizing bulk submission, identity validation, reCAPTCHA, or similar tools in their comment submission processes.

We **recommend** that agencies and relevant coordinating bodies stay abreast of developments in the submission of mass, malattributed, and computer-generated comments, so that approaches to combating difficulties arising from such developments can be implemented as needed.

The second stage at which mass, malattributed, and computer-generated comments raise challenges is the processing of comments. One challenge in this regard is the submission of large numbers of duplicate or near-identical comments. Our research indicates that it is commonplace for agencies to use de-duplication software to identify the unique content. An additional challenge concerns the identification of comments submitted under false identities and by computers. Our research suggests that agencies to this point have devoted little attention to identifying comments in their dockets that are malattributed or computer-generated.

We **recommend** that agencies continue to (or, if they have not already, begin to) utilize de-duplication software to identify the unique content in submitted comments.

We **recommend** that agencies publish policies regarding the posting of duplicate and near-identical comments. These policies should balance values such as user-friendliness, transparency, and informational completeness. Options that could be considered include: posting

a single representative example with the count of the duplicates received and an option to view all comments; breaking out non-identical content; and providing enhanced search options based on the unique information content of comments.

We **recommend** that agencies and relevant coordinating bodies should encourage and stay abreast of technology for identifying malattributed and computer-generated comments in the docket.

We **recommend** that agencies and relevant coordinating bodies stay abreast of technologies that can facilitate public participation outside the notice-and-comment process. Agencies often find that supplemental public participation processes can be useful, and a wide range of technologies can be used to structure meaningful dialogue between agencies and relevant publics.

Coordination and Training

We **recommend** that agencies and relevant coordinating bodies share best practices and relevant innovations for addressing challenges and opportunities connected with mass, malattributed, and computer-generated comments, and technologies related to supplemental public participation processes.

We **recommend** that agencies work closely with relevant coordinating bodies to improve existing technologies and develop new technologies to address issues associated with mass, malattributed, and computer-generated comments.

- The eRulemaking Program should provide a common de-duplication platform for agencies to use, though agencies should be free to modify it or use another platform as appropriate.
- The eRulemaking Program and other relevant coordinating bodies should work with agencies and private sector experts and vendors to develop technologies that respond to common issues associated with mass, malattributed, and computer-generated comments.

We **recommend** agencies offer opportunities for ongoing training and staff development to respond to the rapidly evolving nature of technologies related to mass, malattributed, and computer-generated comments, and supplemental public participation processes.

Docket Management

We **recommend** agencies post comments promptly to the rulemaking docket. If a comment is to be included in the docket, it is important that it be posted promptly upon receipt. Evaluation and control of mass, malattributed, and computer-generated comments depends in part on public responses to them. Most obviously, someone whose identity has been used in a malattributed comment may notice and flag the malattribution. Such external correction cannot occur unless the problematic submission is made available.

We **recommend** that, if an agency decides to exclude or remove some or all duplicate, malattributed, or computer-generated comments from the docket, it articulate such a policy in advance [or provide a reasoned explanation after excluding]. The presumption is that an agency must place all comments in the docket. To exclude a document requires an affirmative justification that is articulated in law or in an agency policy. Existing examples include certain agency prohibitions on threatening language or profanity, on anonymous comments, and on copyrighted material.

We **recommend** that an agency policy against submission of malattributed comments provide that if the agency is aware that it has received such a comment, it either retain the comment in the docket but remove the malattribution (i.e., render it an anonymous submission) or remove the comment from the docket altogether. Even if agencies adopt technological barriers to the submission of malattributed comments, those methods are not likely to be perfect. Our analysis finds that agencies do not have an obligation to affirmatively search the docket for malattributed comments. But agencies are free to set reasonable policies concerning the public comment process and reject comments that violate their policies. Agencies can also rely on comments that violated their commenting policies (e.g. late comments) in some circumstances. If an agency determines that a malattributed comment will remain in the docket, anonymization should be used to protect the person whose identity has been used.

We **recommend** that agencies not discard computer-generated comments that it receives, although those comments may be segregated and treated separately.

We **recommend** that any duplicative, malattributed, or computer-generated comment on which an agency actually relies be placed and retained in the rulemaking docket. As noted, agencies may choose to anonymize malattributed comments, and to segregate or flag computer-generated comments, that are retained in the docket.

We **recommend** that, to the extent practicable, agencies should provide opportunities (including potentially after the comment deadline) for individuals whose names have been attached to comments they did not submit to identify and request removal of such comments from the docket.

We **recommend** that agencies consider taking affirmative steps to identify comments that are malattributed [or computer-generated] comments.

We **recommend** that, if an agency flags a comment as malattributed [or computer-generated] or removes such a comment from the docket and the submitter provided electronic contact information, the agency should notify the submitter of the agency's action.

Transparency

We **recommend** agencies and relevant coordinating bodies consider providing materials that explain to prospective commenters what information is useful to an agency in a public comment. This could include various formats to reach different audiences (e.g., videos, FAQs).

We **recommend** that in NPRMs, NOIs, and ANPRs agencies ask specific questions and identify particular information that would be useful in developing the proposal.

We **recommend** that, when publishing a final rule, agencies state whether they removed from the docket any malattributed and/or computer-generated comments.