

Is Open Data the Death of FOIA?

Beth Simone Noveck

For fifty years, the Freedom of Information Act (FOIA)¹ has been the platinum standard for open government in the United States. The statute is considered the legal bedrock of the public's right to know about the workings of our government. More than one hundred countries and all fifty states have enacted their own freedom of information laws.² At the same time, FOIA's many limitations have also become evident: a cumbersome process, delays in responses, and redactions that frustrate journalists and other information seekers.³ Politically-motivated nuisance requests bedevil government agencies.⁴ With over 700,000 FOIA requests filed every year, the federal government faces the costs of a mounting backlog.⁵

In recent years, however, an entirely different approach to government transparency in line with the era of big data has emerged: *open government data*. Open government data – generally shortened to open data – has many definitions but is generally considered to be publicly available information that can be universally and readily accessed, used, and redistributed free of charge in

1. 5 U.S.C. § 552 (2012).

2. Toby McIntosh, *Reaching 100 FOI Laws, Movement Looks to Future*, FREEDOMINFO.ORG (Sept. 18, 2014), <http://www.freedominfo.org/2014/09/reaching-100-foi-laws-movement-looks-future/> [<http://perma.cc/K4XV-5RLB>].

3. In 2015, over 59% of FOIA requests were released in part. Only 33% of requests were published in full. *FOIA*, U.S. DEP'T JUST., <http://www.foia.gov> [<http://perma.cc/HN4S-WUZR>].

4. David E. Pozen, *Freedom of Information Beyond the Freedom of Information Act* (Aug. 20, 2016) (unpublished manuscript) (on file with author).

5. U.S. DEP'T JUST., *supra* note 3.

digital form.⁶ Open data is not limited to statistics, but also includes text such as the United States Federal Register, the daily newspaper of government, which was released as open data in bulk form in 2010.⁷

To understand how significant the open data movement is for FOIA, this Essay discusses the impact of open data on the institutions and functions of government and the ways open data contrasts markedly with FOIA. Open data emphasizes the proactive publication of whole classes of information. Open data includes data about the workings of government but also data collected by the government about the economy and society posted online in a centralized repository for use by the wider public, including academic users seeking information as the basis for original research and commercial users looking to create new products and services. For example, Pixar used open data from the United States Geological Survey to create more realistic detail in scenes from its movie *The Good Dinosaur*.⁸

By contrast, FOIA promotes ex post publication of information created by the government especially about its own workings in response to specific demands by individual requestors. I argue that open data's more systematic and collaborative approach represents a radical and welcome departure from FOIA because open data concentrates on information as a means to solve problems to the end of improving government effectiveness. Open data is legitimated by the improved outcomes it yields and grounded in a theory of government effectiveness and, as a result, eschews the adversarial and ad hoc FOIA approach. Ultimately, however, each tactic offers important complementary benefits. The proactive information disclosure regime of open data is strengthened by FOIA's rights of legal enforcement. Together, they stand to become the hallmark of government transparency in the fifty years ahead.

The Impact of Open Government Data

An open data movement has taken hold in many countries, and notably in the United States. On his first day in office in 2009, President Obama called for a shift to open data when he signed the Memorandum on Transparency and Open Government. The Memorandum declared that “[i]nformation maintained by the Federal Government is a national asset,” called for the use of “new technologies to put information about [agency] operations and decisions online and [make it] readily available to the public,” and encouraged “executive

6. *Open Data: What's In a Name?*, GOVLAB (Jan. 16, 2014), <http://thegovlab.org/open-data-whats-in-a-name/> [<http://perma.cc/637P-23V2>].

7. David Ferriero, *Federal Register 2.0*, WHITE HOUSE (July 26, 2010), <http://www.whitehouse.gov/blog/2010/07/26/federal-register-20> [<http://perma.cc/X5WH-6RDW>].

8. Lexi Pandell, *How the Good Dinosaur Got Its Stunning Scenery*, WIRED (Nov. 23, 2015), <http://www.wired.com/2015/11/pixar-sharon-calahan/> [<http://perma.cc/7NQP-APT6>].

departments and agencies [to] solicit public feedback to identify information of greatest use to the public.”⁹ Later the same year, the Office of Management and Budget (OMB) published a directive explaining to federal agencies that transparency was not limited to information about the workings of government, but also included “high value” information, which they defined as information that “can be used to increase agency accountability and responsiveness; improve public knowledge of the agency and its operations; further the core mission of the agency; create economic opportunity; or respond to need and demand as identified through public consultation.”¹⁰ In other words, the directive suggested that the public benefits of certain kinds of data, such as locations of reported crimes or weather information, or information that could foster new businesses, justify the effort to publish the data openly.

High value data, therefore, goes beyond the data that the government produces about its own operations (e.g., budgets, salaries of public officials, or names of those who attend meetings at the White House) and includes data and facts that public institutions collect in their role as regulators (e.g., workplace safety and injury records, airplane flight on-time logs or doctors’ prescribing records) or information gathered in their capacities as scientific research organizations (e.g., weather data or information about the human genome). This open data policy—part of a broader set of open government mandates—set in motion a process, enabled by new technology, for agencies to inventory and disclose the information they collect and to move toward the proactive publication of classes of information in their entirety, such as spending records, visitor logs, air and water quality indicators, and measures of hospital infection rates.

Since 2011, seventy countries, including the United States, have signed onto the Open Government Partnership Declaration. The Declaration, which copies the U.S. framework, calls for governments to commit to “pro-actively provide high-value information, including raw data, in a timely manner, in formats that the public can easily locate, understand and use, and in formats that facilitate reuse.”¹¹ Fifteen countries have adopted the International Open Data Charter, which goes further by calling for making government data open in digital formats by default and for investing in the creation of a culture of openness.¹²

9. Memorandum from President Barack Obama to the Heads of Exec. Dep’ts and Agencies on Transparency and Open Government, 74 Fed. Reg. 4685 (Jan. 21, 2009).

10. Memorandum from Peter R. Orszag to the Heads of Exec. Dep’ts and Agencies on the Open Government Directive, WHITE HOUSE (Dec. 8, 2009), <http://www.whitehouse.gov/open/documents/open-government-directive> [<http://perma.cc/WK8V-PW98>].

11. *Open Government Declaration*, OPEN GOV’T PARTNERSHIP (Sept. 2011), <http://www.opengovpartnership.org/about/open-government-declaration> [<http://perma.cc/V3UG-5NNH>].

12. OPEN DATA CHARTER, <http://www.opendatacharter.net> [<http://perma.cc/Q552-H2Z4>].

Since the organization that collects and maintains information is not always in the exclusive position to use the information well, the ability of third parties to access and use open data is what makes the concept truly transformative. Open data facilitates reuse by public institutions as well as entrepreneurs, activists, researchers, and students who can analyze, visualize, and combine open government data—both data the government collects about its own workings and data it collects in its role as regulator or researcher—to achieve impressive results, including the creation of useful tools in the public interest.

For example, the Fire Department of San Ramon, California created the PulsePoint app to enable the public to save lives in medical emergencies using a real-time feed of open data from emergency 911 calls. Effective CPR administered immediately after a cardiac arrest can potentially double or triple a victim's chance of survival, but less than half of victims receive that immediate help because they have to rely upon the arrival of small number of official first responders, such as paramedics or police, to administer aid. PulsePoint takes this function typically performed exclusively by government officials and decentralizes it by enabling trained citizens to serve as first responders.¹³

In communities where emergency calls are published as open data, PulsePoint is able to notify registered and trained CPR users—off-duty doctors, nurses, police, and trained amateurs, for example—near the victim to come to the aid of stricken neighbors. PulsePoint sends them a text message in capital letters: “CPR Needed.” The app has been used in eleven thousand cases to activate citizen responders in response to cardiac emergencies.¹⁴

One consequence of opening government data, when there is a desire to do so, is the ability to improve transparency and accountability in government. In the United States, at the federal level, open data facilitated the creation of USASpending.gov, a set of online tools for exploring the federal budget.¹⁵ At the local level, open data drives various “open checkbook” websites;¹⁶ in Aus-

13. BETH SIMONE NOVECK, SMART CITIZENS, SMARTER STATE 39–40 (2015).

14. *Id.*; see also Beth Simone Noveck, *Swipe Right To Fix the World: Can Tinder-Like Tech Match Solutions to Problems?*, THE GUARDIAN (Dec. 29, 2015), <http://www.theguardian.com/commentisfree/2015/dec/29/swipe-right-to-fix-the-world-can-tinder-like-tech-match-solutions-to-problems> [<http://perma.cc/JA7Z-F9UG>] (describing how the PulsePoint application operates).

15. USA SPENDING, <http://www.usaspending.gov/Pages/Default.aspx> [<http://perma.cc/L96L-3HAN>].

16. *Checkbook NYC*, N.Y.C. COMPTROLLER'S OFF., http://www.checkbooknyc.com/spending_landing/yeartype/B/year/118 [<http://perma.cc/T38J-JUKU>].

tria, for example, over eight hundred municipalities have made their spending data more transparent and easy to visualize online.¹⁷

But open data is not limited to information about spending. Opening local government data about public works in Zanesville, Ohio revealed a fifty-year pattern of discriminatory water service provision and led to a successful civil rights lawsuit against Zanesville in 2008.¹⁸ While access to clean water from the City of Zanesville water line spread throughout the rest of Muskingum County, residents of the predominantly African-American area of Zanesville, Ohio were able only to use contaminated rainwater or to drive to the nearest water tower in order to truck water back to their homes.¹⁹

Although shining a light on government is important, the potential impact of open data is far broader than FOIA's paradigmatic goals of improving government transparency.²⁰ Open data can improve the accountability of private organizations and institutions. For example, several states are moving to release data collected on doctors about their opioid pain medication prescribing patterns.²¹ By showing doctors their own practices in comparison to that of others to create an incentive for behavior change, states like Arizona are already showing a ten percent reduction in opiate prescriptions and a four percent reduction in overdose deaths in those counties that used open data in this fashion over those counties that did not.²²

Open data, however, does more than promote accountability. It also enables the creation of tools to improve consumer choice and citizen decision-making. The data collected by the government from universities, for example, has been transformed into a calculator—The College Scorecard—to help parents and

17. OPEN BUDGET, <http://www.offenerhaushalt.at/> [<http://perma.cc/CZ4D-MVYU>]; see also Bernhard Krabina, *Austria's Award-winning Open Spending Portal Turns Two*, FISCAL TRANSPARENCY (Oct. 1, 2015), <http://www.fiscaltransparency.net/use/austrias-award-winning-open-spending-portal-turns-two/> [<http://perma.cc/5577-43CE>].

18. *Kennedy v. City of Zanesville*, No. 2:03-CV-1047, 2008 WL 2036713 (S.D. Ohio May 9, 2008); see also Christina Rogowski, Stefaan Verhulst & Andrew Young, *Open Data's Impact: Ohio, USA: Kennedy v. City of Zanesville: Open Data as Evidence*, GOVLAB (Jan. 2016), <http://odimpact.org/static/files/case-studies-us-kennedy.pdf> [<http://perma.cc/6PBE-KAPA>].

19. Rogowski, Verhulst & Young, *supra* note 18, at 3.

20. For more about the impacts of open data, see *Open Data's Impact: Open Data Is Changing the World in Four Ways*, GOVLAB, <http://www.ODimpact.org> [<http://perma.cc/K3WB-YEPP>].

21. Scott Calvert, *Doctors' Individual Opioid Prescription 'Report Cards' Show Impact*, WALL ST. J. (Sept. 2, 2016, 6:50 PM), <http://www.wsj.com/articles/doctors-individual-opiate-prescription-report-cards-show-impact-1472856624> [<http://perma.cc/NJ8S-A4L2>].

22. *Id.*

students make more informed financial decisions about their college education.²³

Sometimes the benefits of open data ripple out beyond the immediate motivations for its publication. For instance, while publishing government contracts can boost public integrity, it can also catalyze greater business competition and entrepreneurship.²⁴ A study by McKinsey estimates that open data will end up creating \$3 billion in economic value.²⁵ The Open Data 500 tracks thousands of companies worldwide that use open government data as a core business asset. For example, the New York-based start-up Aidin shows patients released from the hospital the best available post-acute care providers, using information published by the Department of Health and Human Services.²⁶

We are undoubtedly living in the middle of an open data revolution, which is only in its infancy. One thousand people attended the 2015 International Open Data Conference, an event dedicated to showcasing the impact of opening governing data, and one thousand five hundred in 2016.²⁷ In the United States, an Open Data Executive Order (2013) augments the earlier Open Government Memorandum and Directive by effectively declaring that open data will become the new normal for the United States government.²⁸ Recent domain-specific legislation and regulation include promises of more real-time open data, as in the DATA Act, which calls for publishing federal government spending data as open data in standardized formats.²⁹ Legislation passed by Congress for an evidence-based policy commission in 2016 shows a growing

23. *College Scorecard*, U.S. DEP'T EDUC., <http://collegescorecard.ed.gov/> [<http://perma.cc/SF5H-4JSA>].

24. See OPEN CONTRACTING PARTNERSHIP (2016), <http://www.open-contracting.org/> [<http://perma.cc/78ZR-TMZ3>].

25. James Manyika et al., *Open Data: Unlocking Innovation and Performance With Liquid Information*, MCKINSEY GLOBAL INST. (Oct. 2013), <http://www.mckinsey.com/business-functions/business-technology/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information> [<http://perma.cc/DKP3-3ATK>].

26. The Open Data 500, *The OD500 Global Network*, THE GOVERNANCE LAB, <http://www.opendata500.com> [<http://perma.cc/KD4M-H457>]. The OD500 Global Network is an international network of organizations that seek to study the use and impact of open data.

27. *Today Begins the 4th International Open Data Conference*, IO DC16, <http://opendatacon.org/today-begins-the-4th-international-open-data-conference/> [<http://perma.cc/H8SG-QL7G>] (“The fourth edition of IO DC16 convenes more than 1,500 guests from a hundred countries . . .”).

28. Exec. Order No. 13642, 78 Fed. Reg. 28111 (May 9, 2013).

29. Digital Accountability and Transparency (DATA) Act of 2014, Pub. L. No. 113-101, 128 Stat. 1146.

interest in basing policies on empirical research and data, and is fueling demand for more access to administrative information of all kinds, including the data that agencies collect about companies, workplaces, the environment, and the world beyond government.³⁰ In criminal justice, there is a national policy initiative to open up access to administrative data from criminal justice and related agencies—not only about those agencies’ practices but also about crime more broadly—to enable empirically informed innovation and reform.

Open Data and FOIA: Competitors or Partners?

The explosion of newly available data coupled with the availability of technologies to make use of these large quantities of data—so-called Big Data—gives rise to the question of whether open data, as a policy framework, should supplant the FOIA system. In practice, open data: a) promotes broad scale transparency; b) simplifies the disclosure process; c) requires publication in reusable electronic formats; and, above all, d) focuses on disclosure of information collected by government as regulator and researcher, not exclusively on data created by government about its own workings. As a result, open data promotes a collaborative rather than an adversarial process and catalyzes co-creation of solutions by government working with the public. At the same time, FOIA indispensably *complements* open data by a) providing a legal right of action to compel disclosure; b) suggesting the kinds of data to prioritize releasing; and c) disclosing who is using what data how.

Open data encourages proactive disclosure and publication in open formats on a centralized data portal such as Data.gov. FOIA, by contrast, imposes limited affirmative obligations on agencies,³¹ emphasizes ex post disclosure, and, in most cases, releases information on paper only to the individual requestor³² rather than to the general public. “The process of drafting and submitting FOIA requests and then waiting for the agency’s response,” writes David C. Vladeck, “is a breeding ground for delay and cynicism over the Act’s efficacy.”³³ Under FOIA, there is no centralized “reading room”: each agency must create its own website and post only those documents that are the result of three or

30. Evidence-Based Policymaking Commission Act of 2016, Pub. L. No. 114-140, 130 Stat. 317.

31. Michael Herz, *Law Lags Behind: FOIA and Affirmative Disclosure of Information*, 7 CARDOZO PUB. L. POL’Y & ETHICS J. 577, 578-79 (2009) (“FOIA’s fundamental limitation is its failure to impose affirmative responsibilities on agencies. In particular, (a) it does not require agencies to generate information, and (b) it imposes only minimal (and frequently disregarded) obligations to disseminate information without being asked.”).

32. Pozen, *supra* note 4, at 6.

33. David C. Vladeck, *Information Access—Surveying the Current Legal Landscape of Federal Right-to-Know Laws*, 86 TEX. L. REV. 1787, 1789 (2008).

more FOIA requests.³⁴ To be sure, there is a certain incentive for some requestors, such as the journalist looking for a scoop, to be the only one to see the information that results from a FOIA request. But that benefit is far outweighed by the gains to be had from broad information disclosure to a larger audience who can make use of the information.

In the era of big data technologies, when information storage is cheap and plentiful, if the goal is to promote greater transparency, it is hard to imagine why we should continue to invest in the legal framework and its attendant practices for demanding data after the fact when, instead, we can build the platforms and policies to ensure proactive and prospective publication of government information in reusable formats online. Notwithstanding frequent calls by Attorneys General³⁵ for more rapid disclosure under FOIA, the largely paper-based process, whereby the seeker has to file a written request specifying the desired information, is inherently fraught with delays and backlogs.

By contrast, with open data, technology helps to transform information transparency from a legal principle into a practical reality. In essence, whereas FOIA is a legal regime, open data is a set of technology standards and practices. The need to populate Data.gov drives agencies to automate the process of inventorying all of the data sets stored on their servers without needing to know who created or currently maintains the data, thereby accelerating the process of discovering the data sets to which Data.gov can then point. When Data.gov launched in May 2009, it made forty-seven datasets searchable; that number skyrocketed to 186,000 in 2016.³⁶

Even more important than the volume of raw information is the positioning of open data as a collaborative rather than an adversarial process. FOIA is contentious, giving rise to litigation when the government refuses a request and perversely reinforcing a culture of closed-door governing. Lawyers and specially trained FOIA officials acculturated in this cat-and-mouse process dominate the FOIA process. Chief Innovation, Chief Data, or Chief Infor-

34. FOIA Improvement Act of 2016, Pub. L. No. 114-185, 130 Stat. 538 (June 30, 2016) (to be codified at 5 U.S.C. § 552) (requiring each agency to post “copies of all records, regardless of form or format—i) that have been released to any person under paragraph (3); and ii) (I) that because of the nature of their subject matter, the agency determines have become or are likely to become the subject of subsequent requests for substantially the same records; or (II) that have been requested 3 or more times”).

35. See, e.g., Memorandum from Eric Holder, Att’y Gen., to Heads of Exec. Dep’ts and Agencies (Mar. 19, 2009), <http://www.usdoj.gov/ag/foia-memo-march2009.pdf> [<http://perma.cc/43R9-SV3Z>].

36. DATA.GOV, <http://www.data.gov> [<http://perma.cc/72V6-A2S3>]. As of this writing, Data.gov contains 186,467 published datasets; see also *Twenty-First Century Government*, WHITE HOUSE, <http://www.whitehouse.gov/21stcenturygov/tools/data-gov> [<http://perma.cc/S58S-EURN>] (“In just one year, Data.gov has grown from 47 datasets.”).

mation Officers, in contrast, are tasked with identifying and posting open data, which is structured for computability and reusability and which by its nature is designed for collaboration. The Commerce Department's Data Service website not only makes the federal department's information more searchable; it provides training, tutorials, and guidance precisely in order to help the public use its data to generate useful insights, products, and tools.³⁷

As a practical matter, the open data disclosure process affords various new efficiencies and improvements. For example, why take time to black out sensitive passages and redact personally identifiable information, when government can invest in technologies for automatically de-identifying Social Security numbers when storing information so it can be automatically released? In an open data process, managed by technologists, there is greater attention to the application of new technology, which can enhance the publication of appropriately public information while protecting the privacy of personally identifiable information. In place of lawyers who have to eyeball documents manually, open data creates the impetus to hire computer scientists to write software that strips Social Security numbers and other sensitive personal data automatically, as well as social and data scientists who can extract value from large scale administrative data sets and translate those findings into improvements for government and society.

Similarly, by requiring publication of digital records, open data enables publication of information that was "born digital," in contrast to FOIA, which has traditionally focused on distribution of paper copies of records. Why be content with receiving paper when information—especially information that is created digitally in the first place—could and should be disseminated digitally to enable more meaningful and productive reuse? In June 2016, President Obama signed into law amendments to the Freedom of Information Act striking the phrase "for public inspection and copying" and inserting "for public inspection in an electronic format."³⁸

In light of FOIA's contentiousness and its narrow focus on information about government's workings, open data would seem to be a natural evolution from, and improvement over, FOIA.³⁹ However, there are strong arguments in favor of blending the two approaches and drawing on the best of FOIA to improve how open data works, treating one regime as a supplement rather than a substitute.

37. *Commerce Data Service*, DEP'T OF COMMERCE, <http://www.commerce.gov/dataservice/> [<http://perma.cc/C2NN-VXMV>].

38. FOIA Improvement Act of 2016, Pub. L. No. 114-185, 130 Stat. 538 (June 30, 2016) (to be codified at 5 U.S.C. 552).

39. Herz, *supra* note 31 at 596-98.

First, where gaps exist in the open data regime, FOIA provides the legal right of action to fight for the data that government refuses to disclose when it should. The data to be found on Data.gov is not always the contentious nuclear secrets, budget models, or national security information most in demand by journalists, activists, and researchers. In 2013, transparency activist Carl Malamud had to use FOIA to request nine nonprofit tax returns from the IRS. Although nonprofit returns are required to be disclosed by law and the filers submitted those returns electronically, the IRS wanted to send Malamud image files of the returns. He sued for the digital originals and won.⁴⁰ The significance of the decision is that the IRS will now make all electronically filed nonprofit tax returns—about 60% of those filed since 2011—digitally downloadable as open data.⁴¹ The IRS did not release the nonprofit tax returns as downloadable data until Malamud filed suit and won.

Despite seven years of open data policy in the United States, FOIA requests have risen from under 600,000 to well over 700,000 during the same period.⁴² It is hard to know whether that increase is due to greater secrecy on the part of the Obama administration, decreased efficiency on the part of its FOIA officers or, in fact, a growing culture of openness simply spurring more requests as a result of the coexistence of the two regimes. But, even discounting for spurious requests (the JFK assassination and Area 51 conspiracy theories remain popular subjects for requests), open data has not yet prompted the online release of everything that needs to be published either to improve government transparency or effectiveness. As a last resort, the legal right of action to sue for information under FOIA is essential.

Second, FOIA doctrine articulates clear priorities. FOIA emphasizes the release of information about the workings of government, such as the activities of government officials, and focuses on one-off demands by requestors who have to guess what information the agency has and then contend with denials and delays in the processing those requests. Open data, meanwhile, requires inventorying *all* data and creates an opportunity for reasoned debate between the

40. *Public.Resource.org v. U.S. Internal Revenue Serv.*, 78 F. Supp. 3d 1262 (N.D. Cal. 2015), *appeal dismissed* No. 15-15607 (9th Cir. June 24, 2015) (holding that the IRS must produce digital and not paper-based files of electronically-filed nonprofit tax returns).

41. *IRS Makes Electronically Filed Form 990 Data Available in New Format*, INTERNAL REVENUE SERV. (June 16, 2016), <http://www.irs.gov/uac/newsroom/irs-makes-electronically-filed-form-990-data-available-in-new-format> [<http://perma.cc/4AFM-YU8Q>]; see also Beth Simone Noveck & Daniel L. Goroff, *Information for Impact: Liberating Nonprofit Sector Data*, THE ASPEN INSTITUTE (Sept. 2013), http://assets.aspeninstitute.org/content/uploads/files/content/docs/pubs/Information_for_Impact_Report_FINAL_REPORT_9-26-13.pdf [<http://perma.cc/28TP-4A7F>].

42. *FOIA at a Glance: Requests Received*, U.S. DEP'T OF JUSTICE, <http://www.foia.gov/> [<http://perma.cc/U86U-RHEV>].

public and the agency about what to publish, with what frequency, and in what formats. Open data focuses on publishing whole classes of information in bulk. Given that data often has to be cleaned and formatted to make it suitable for downloading as open data, creating open data requires an investment of both time and money to release. Hence, prioritization is important. Drawing upon the lessons of FOIA—notably, the emphasis on disclosure of data by government about government—open data policy, too, should evolve to articulate normative guidance to agencies about what should be published online and when.

However, those priorities should take account of the unique properties of open data. Because open data lends itself to analysis, visualization, and the creation of algorithms to drive decision-making, efforts to publicize the calendars of cabinet secretaries or the salaries of government officials online—very much FOIA priorities—are misplaced for open data because, absent gross malfeasance, such disclosures will not drive immediate changes in how government operates. Naked disclosures alone are not enough; rather, change depends upon the actions that are taken following the publication of data.

Open data can be such a powerful tool because the proactive disclosure is often—although not always—accompanied by a plan for how to use the data. For example, Brazil has long published government spending data—an initiative that has led to a decrease in certain public expenditures because of the ability to analyze the entire corpus of credit card records. At the same time, the publication of spending data has not prevented the country's endemic corruption.⁴³ Clearly, the Brazilian transparency website is better than nothing, but the example highlights that it is as important to act upon the data as to publish it. Whether disclosed as a result of a FOIA petition or an open data policy push, steps must be taken to make use of the data to enable more evidence-based practices.

In some areas of government operations, open data can immediately improve how services are delivered both because the data is available and because incentives are aligned between those inside and outside of government to act upon the data to change how government works. For example, opening up the entire class of data about food-borne illnesses in Chicago prompted the creation of an algorithm to re-organize how the city allocates scarce resources to

43. Auralice Graft, Stefaan Verhulst & Andrew Young, *Brazil's Open Budget Transparency Portal: Making Public How Public Money is Spent*, GOVLAB (Jan. 2016), <http://odimpact.org/case-brazils-open-budget-transparency-portal.html> [<http://perma.cc/S4A6-LBRS>]; see also *Brazil's Gold Medal for Corruption*, N.Y. TIMES (June 6, 2016), <http://www.nytimes.com/2016/06/06/opinion/brazils-gold-medal-for-corruption.html> [<http://perma.cc/97U3-3Q2N>].

doing restaurant inspections.⁴⁴ For open data, like for FOIA, priorities should be articulated, and should emphasize improvement of policies and services.

Third, FOIA requests are counted and logged through a process of agencies reporting to the Department of Justice, which then reports FOIA statistics on the FOIA.gov website. By contrast, open data turns on a data fire hose. One of the unintended consequences of making data freely available for reuse without restriction or registration is that it is very difficult to trace the downstream uses of such data. Although there is mounting evidence of the positive impact of open data, it is hard to know with certainty how valuable any given dataset is. We can be fairly certain that agencies post a lot of data on Data.gov that no one wants, knows exists, or uses. Thus, ideally, the hi-tech world of open data should be married to the accounting practices of FOIA. For example, the *Learning Registry* is a collaborative effort between the Departments of Education and Defense and the White House to create “data about data” and make it possible to learn who is using open learning content on the web.⁴⁵

Conclusion: The Way Forward

Of course, FOIA and open data are similar. They both emphasize disclosure to the public of information created or collected by the government. Because open data has its roots in the data processing technologies of the big data era, however, most open data projects focus on the analysis of facts collected by the government in its role as regulator and researcher. As a result, open data substitutes a utilitarian rationale for transparency in place of a justification based on moral obligation.

In other words, open data is rooted in a theory about government effectiveness whereas FOIA is grounded in a theory of governmental legitimacy. The narrative underlying FOIA—though hardly borne out in practice—suggests that shining the light of transparency on the workings of government—knowing the calendar of a cabinet secretary or what the postman earns—will de facto lead to a decrease in corruption, a resulting improvement in government accountability, and therefore greater legitimacy on the part of public institutions. In the FOIA narrative, this is an almost automatic process; hence, the law enshrines the right to ask for information for any reason or no reason at all but focuses on individual requests for specific facts and data. As a result, writes David Pozen, FOIA creates an “entitlement program with no eligibility criteria . . . [that] establishes nondisclosure as the default norm.”⁴⁶

44. Julian Spector, *Chicago Is Predicting Food Safety Violations. Why Aren't Other Cities?*, ATLANTIC CITY LAB (Jan. 7, 2016), <http://www.citylab.com/cityfixer/2016/01/chicago-is-predicting-food-safety-violations-why-arent-other-cities/422511> [<http://perma.cc/94VJ-JQ4L>].

45. LEARNING REGISTRY, <http://learningregistry.org/> [<http://perma.cc/9BVA-U3ZL>].

46. Pozen, *supra* note 4, at 3.

Controversially and by contrast, the open data narrative is largely unrelated to accountability. Instead, open data is a tool for use by diverse stakeholders to make government more effective at solving problems but also to solve social problems, create jobs, and generate entrepreneurship.⁴⁷ Open data emphasizes the instrumental value of information as an asset for evidence-based decision-making, service delivery, and economic growth. The Department of Education posts data about the costs of college tuition so that the agency can create a financial aid calculator. Municipalities release a feed of 911 call data so that PulsePoint can operate a citizen first responder system.⁴⁸ Hence certain classes of information—what Cass Sunstein calls output transparency—should be disclosed a priori whether or not people request it to improve government effectiveness.⁴⁹

Some view this utilitarian calculus of the usefulness of open data for solving problems as a dilution of FOIA's sui generis goals of transparency for its own sake. But the open data movement has progressed far enough worldwide to demonstrate that it is not going away. Arguably, the greatest advances in transparency in recent years have come either as a result of extra-legal disclosures such as WikiLeaks or the Snowden revelations, or as a result of the publication of open data.

As open data's scope continues to expand and its culture becomes a comfortable part of government agencies' behavior, policymakers have to re-evaluate the role of FOIA. FOIA officials currently convene as a government-wide community, but they need to talk and work more closely with Chief Information Officers, Chief Data Officers, and Chief Innovation Officers to ask and answer how the FOIA process should take advantage of the technological affordances of open data and, above all, to embrace more of the collaborative nature of open data practices, which emphasize problem solving. At the same

47. The various rationales and impacts of open data are discussed at length at GOVLAB, *supra* note 20.

48. This utilitarian thinking is further justified by the argument that taxpayer dollars paid for the collection of the data, which cannot be restricted by copyright law in the United States, where the federal government may not assert intellectual property rights. See Subject Matter of Copyright: United States Government Works, 17 U.S.C. § 105 (1976) ("Copyright protection under this title is not available for any work of the United States Government.").

49. Cass R. Sunstein, *Output Transparency vs. Input Transparency* (Aug. 18, 2016) <http://dx.doi.org/10.2139/ssrn.2826009> [<http://perma.cc/27J8-J7BH>], at 3. In fact, the FOIA Improvement Act of 2016 takes steps in this direction toward creating a legislative basis for open data by calling for the creation of "procedures for identifying records of general interest or use to the public that are appropriate for public disclosure, and for posting such records in a publicly accessible electronic format." FOIA Improvement Act of 2016, Pub. L. No. 114-185, § 4, 130 Stat. 538 (June 30, 2016) (to be codified at 5 U.S.C. § 552).

time, those responsible for the implementation of open data stand to learn from fifty years of FOIA practice about the urgency of responsive information disclosures for achieving good government and greater value to society. Evolving our legal and policy framework for public information collection and publication will engage the legal profession, technologists, and policymakers—as well as the general public that has benefitted from the open data movement and from FOIA.

Beth Simone Noveck is the Florence Rogatz Visiting Clinical Professor of Law, Yale Law School, Jerry M. Hultin Global Network Professor of Engineering, New York University and director of The Governance Lab. Her most recent book is Smart Citizens, Smarter State: The Technologies of Expertise and the Future of Governing (Harvard University Press, 2015). She tweets @bethnoveck. The author is grateful to David Pozen and Michael Schudson of Columbia University for organizing an excellent conference on FOIA's fiftieth anniversary in June 2016 at which many of the ideas in this Essay were presented and to David Schulz of the Media Freedom and Information Access Clinic at Yale Law School for help refining them.

Preferred Citation: Beth Simone Noveck, *Is Open Data the Death of FOIA?*, 126 YALE L.J. F. 273 (2016), <http://www.yalelawjournal.org/forum/is-open-data-the-death-of-foia>.