*Citizen Science Q&A:* **Foundations for Evidence-Based Policymaking Act**
Published: December 23, 2020

Through the Citizen Science Association Law & Policy Working Group, the Emmett Environmental Law & Policy Clinic at Harvard Law School has volunteered to make its students available to answer questions about relevant laws and policies raised by citizen science projects. The questions below were submitted through the working group's question submission form.

*The answers below are provided for educational purposes only. Clinic students are not practicing attorneys. By answering questions, Clinic students are not providing legal advice, acting as your attorney, or serving as a substitute for the advice of an attorney. Their answers to submitted questions do not create an attorney-client relationship or a commitment to answer additional questions. Communications between individuals using the question submission form and the Clinic's students are not protected by the attorney-client privilege or work product doctrine. Readers should contact their attorney to obtain advice with respect to any particular legal matter.*

**Question: What influence is the recent "Foundations for Evidence-Based Policymaking Act" (2018) having, or likely to have, on privacy protections in citizen science (CS) projects?**

**Answer:** The Foundations for Evidence-Based Policymaking Act of 2018[1] ("Evidence Act" or "Act") promotes openness with government data while creating new privacy protections. In general, the Act aims to make government data more accessible both to the public and among federal agencies, while safeguarding confidential or personally identifiable information.

It is hard to say at this point what influences or impacts the Evidence Act's privacy provisions are having or are likely to have. This is because the Act leaves it to the Office of Management and Budget ("OMB") and other agencies to flesh out many details about the privacy protections, in the form of new rules or guidance documents. But no such documents have been published yet. Similarly, we were not able to find documentation about how any particular agency plans to implement, or has already implemented, the privacy-related sections of the Act.

Nonetheless, in this answer we summarize the Act's transparency and privacy-related provisions. Next, considering the Act more broadly, we describe opportunities the law provides for individuals and organizations to weigh in on how agencies will implement it. This answer then concludes with a list of resources for further research on these matters.

Before diving in, here are three big picture points about the Act. First, its data-related provisions apply to any data submitted to a federal agency—including citizen science data—because the Act defines data broadly. "Data" is "recorded information, regardless of form or the media on which the data is recorded."[2] The Act also uses the term "data asset," which is "a collection of data elements or data sets that may be grouped together."[3] ("Data element" and "data set" are

not defined separately.)  Therefore, although the Act does not explicitly discuss citizen science data, it applies to any such data maintained by a federal agency.  The Act does not cover data held by a private entity or a state if it has not yet been submitted to a federal agency.  But, whatever its source, any data that gets to a federal agency is subject to the Act.

Second, the Act is divided into three Titles.  The open data and privacy provisions are in Titles II and III.  Title I is therefore not directly relevant here, but it comes up in a few places below.  Title I modifies federal agencies' strategic planning processes by requiring agencies to consider in advance what data they will need to address pending policy questions.

Third, the Act's sections vary somewhat in terms of the federal agencies they cover, though many of the privacy-related provisions in Titles II and III cover agencies relevant to citizen scientists, like the U.S. Environmental Protection Agency ("EPA").  (Title I applies only to twenty-four agencies, but these also include many agencies of interest to citizen scientists, like EPA, the U.S. Department of the Interior, the National Science Foundation, etc.).[4]

The Evidence Act's Privacy Protections

*Title II – Open government requirements*

Title II is the Open, Public, Electronic, and Necessary (OPEN) Government Data Act.[5]  It applies to almost every federal agency—likely all agencies relevant for citizen scientists.[6]  It covers, for instance, the U.S. Environmental Protection Agency, the National Science Foundation, and the U.S. Department of Transportation.  Title II requires that federal agencies, by default, make all of their data assets machine-readable[7] and available in an open format.[8]  Each agency must also, "to the maximum extent practicable, develop and maintain a comprehensive data inventory that accounts for all data assets created by, collected by, under the control or direction of, or maintained by the agency."[9]

The Act directs OMB to write guidance for agencies to help them carry out these new duties, and the Act's instructions for that guidance recognize that some data is not suitable for public disclosure.  Specifically, the guidance must address, among other things, the costs and benefits of making the data publicly available in an open format and the risk of disclosure of personally identifiable information ("including the risk that an individual data asset in isolation does not pose a privacy or confidentiality risk but when combined with other available information may pose such a risk").[10]  The guidance must also take into account whether a data asset is "restricted by contract or other binding, written agreement."[11]

More broadly, OMB must "develop and oversee the implementation of policies, principles, standards, and guidelines on privacy, confidentiality, security, disclosure and sharing of information collected or maintained by or for agencies."[12]  Likewise, agencies must implement and enforce those policies, principles, standards, and guidelines.[13]

Title II also touches on pledges of confidentiality, which Title III covers in more detail.  For data relating to "statistical policy and coordination," each agency must "protect respondents' privacy and ensure that disclosure policies fully honor pledges of confidentiality."[14]

Federal agencies appear to have made little progress in implementing Title II. In October 2020, the Government Accountability Office ("GAO") found that OMB had not yet issued the statutorily required guidance on comprehensive data inventories.[15] GAO reported that this guidance was in draft form at that time. (OMB has published two guidance documents under the Act, but these primarily address agencies' obligations under Title I.)[16] The Act requires federal agencies to appoint Chief Data Officers,[17] which many agencies have done.[18] Otherwise, we did not find indications that agencies have made further progress towards implementing Title II.

*Title III – Statistical agencies and confidentiality agreements*

Title III is the Confidential Information Protection and Statistical Efficiency Act of 2018.[19] It strengthens protections for data that is acquired by any agency under a pledge of confidentiality and used for "statistical purposes."[20] Specifically, "data or information acquired by an agency under a pledge of confidentiality for exclusively statistical purposes shall not be disclosed by an agency in identifiable form,[21] for any use other than an exclusively statistical purpose, except with the informed consent of the respondent."[22] Here, "agency" is defined broadly and covers just about every federal entity.[23]

Title III creates other privacy protections that apply only to "statistical agenc[ies] or unit[s]." The Act defines "statistical agency or unit" as an "agency or organizational unit of the executive branch whose activities are predominantly the collection, compilation, processing, or analysis of information for statistical purposes, as designated by" OMB.[24] There are thirteen principal federal statistical agencies.[25] Those potentially relevant for citizen science purposes include the Energy Information Administration, National Center for Health Statistics, and the National Center for Science and Engineering Statistics.

The Act provides that the "responsibilities" of each statistical agency or unit include "protect[ing] the trust of information providers by ensuring the confidentiality and exclusive statistical use of their responses."[26] Accordingly, each statistical agency or unit must "adopt policies, best practices, and appropriate procedures to implement" its statutory responsibilities.[27]

The Act also identifies three "Designated Statistical Agencies": the Census Bureau and Bureau of Economic Analysis in the Department of Commerce, and the Bureau of Labor Statistics in the Department of Labor.[28] They must "protect the confidentiality of individually identifiable information acquired for statistical purposes by adhering to [various] safeguard principles."[29]

Similar to Title II, various provisions in Title III either require or authorize OMB to issue rules or guidance related to agencies' obligations under Title III. For instance, OMB "may promulgate rules or provide other guidance to ensure consistent interpretation of [Title III] by the affected agencies."[30] And OMB "shall prescribe regulations to carry out" the section on statistical agencies' responsibilities.[31] In addition, individual agencies "may promulgate rules to implement" their obligations under Title III.[32]

We did not find any indication that OMB or any individual agency has yet issued guidance or rules that directly address their obligations under Title III.

Opportunities for Public Input

As part of their Title I planning obligations, agencies must engage in "consultation" with "stakeholders, including the public, agencies, State and local governments, and representatives of non-governmental researchers."[33] Such consultation gives the public an opportunity to influence how federal agencies conceive of citizen science data, and to raise the issue of data privacy. So far, at least the National Science Foundation has published a "request for information" under Title I of the Evidence Act; responses are due by Dec. 31, 2020.[34] Other agencies might publish similar requests.

In addition, Title II requires agencies to prepare "strategic information resources management plans." Those plans must, "to the extent practicable," include open data plans that "facilitate collaboration with non-Government entities (including businesses), researchers, and the public for the purpose of understanding how data users value and use government data."[35]

If OMB or any agencies conduct notice-and-comment rulemakings under the Evidence Act, the public will have an opportunity to provide comments on draft rules. Any such rulemakings will be published in the Federal Register (online at federalregister.gov).

Resources for Further Research

OMB, *Evidence and Evaluation*, https://www.whitehouse.gov/omb/information-for-agencies/evidence-and-evaluation
- This page includes the following OMB guidance documents under the Evidence Act:
  - *M-19-23 Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance* (July 10, 2019), https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf
  - *M-20-12 Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices* (March 10, 2020), https://www.whitehouse.gov/wp-content/uploads/2020/03/M-20-12.pdf
- Any future guidance under the Evidence Act will likely appear on this website.

Government Accountability Office, *Open Data: Agencies Need Guidance to Establish Comprehensive Data Inventories; Information on Their Progress is Limited* (October 2020), *available at* https://www.gao.gov/assets/720/710068.pdf.

President's Management Agenda, *Federal Data Strategy 2020 Action Plan*, https://strategy.data.gov/assets/docs/2020-federal-data-strategy-action-plan.pdf.

Resources.data.gov ("an online repository of policies, tools, case studies, and other resources to support data governance, management, exchange, and use throughout the federal government").

Nick Hart & Nancy Potok, *Modernizing U.S. Data Infrastructure: Design Considerations for Implementing a National Secure Data Service to Improve Statistics and Evidence Building* (Data Foundation, July 2020), *available at* https://www.datafoundation.org/modernizing-us-data-infrastructure-2020.[36]

Data Coalition, *Open, Public, Electronic and Necessary (OPEN) Government Data Act*,
http://www.datacoalition.org/wp-content/uploads/2018/12/Summary-
OPENGovtDataAct.Section-by-Section-2018.pdf.

- This fact sheet explains each section of Title II of the Evidence Act.

---

[1] Pub. L. No. 115-435, 132 Stat. 5529 (Jan. 14, 2019) [hereinafter "Evidence Act"]. The Act's full text is available at Congress.gov: https://www.congress.gov/bill/115th-congress/house-bill/4174/text. This document generally cites the United States Code (e.g., "44 U.S.C. § 3502"). The Act's text identifies which code sections it is amending, though the bill text sometimes leaves out the periods or section symbol (e.g., "44 USC 3576").

[2] 44 U.S.C. § 3502(16).

[3] *Id.* § 3502(17).

[4] Title I, 5 U.S.C. § 311(1), borrows the definition of "agency" from another law, the Chief Financial Officers Act, which only applies to the twenty-four agencies listed in 31 U.S.C. § 901(b): the departments of Agriculture, Commerce, Defense, Education, Energy, Health and Human Services, Homeland Security, Housing and Urban Development, Interior, Justice, Labor, State, Transportation, Treasury, and Veterans Affairs, as well as the Environmental Protection Agency, the National Aeronautics and Space Administration, the Agency for International Development, the General Services Administration, the National Science Foundation, the Nuclear Regulatory Commission, the Office of Personnel Management, the Small Business Administration, and the Social Security Administration.

[5] Evidence Act § 201.

[6] In Title II, "agency" means "any executive department, military department, Government corporation, Government controlled corporation, or other establishment in the executive branch of the Government (including the Executive Office of the President), or any independent regulatory agency, but does not include-- (A) the Government Accountability Office; (B) Federal Election Commission; (C) the governments of the District of Columbia and of the territories and possessions of the United States, and their various subdivisions; or (D) Government-owned contractor-operated facilities, including laboratories engaged in national defense research and production activities." 44 U.S.C. § 3502(1).

[7] "Machine-readable" data is "data in a format that can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost." *Id.* § 3502(18).

[8] *Id.* § 3506(b)(6).

[9] *Id.* § 3511(a)(1).

[10] *Id.* §§ 3504(b)(6)(A), 3511(a)(2)(E)(i).

[11] *Id.* §§ 3504(b)(6)(E)(iii), 3511(a)(2)(E)(v)(III). In addition, the OMB guidance on comprehensive data inventories must take into account that some data assets may not be "subject to open format or open license requirements due to existing limitations or restrictions on government distribution of the asset." *Id.* § 3511(a)(2)(A)(iii)(I)(dd). "Open format" is not defined separately, but "open license" means "a legal guarantee that a data asset is made available (A) at no cost to the public; and (B) with no restrictions on copying, publishing, distributing, transmitting, citing, or adapting such asset." *Id.* § 3502(21).

[12] *Id.* § 3504(g)(1).

[13] *Id.* § 3506(g)(1).

[14] *Id.* § 3506(e)(3). A "respondent" is "a person who, or organization that, is requested or required to supply information to an agency, is the subject of information requested or required to be supplied to an agency, or provides that information to an agency." *Id.* § 3561(9).

[15] Government Accountability Office, *Open Data: Agencies Need Guidance to Establish Comprehensive Data Inventories; Information on Their Progress is Limited*, at 14 (October 2020), *available at* https://www.gao.gov/assets/720/710068.pdf [hereinafter "GAO Report"].

[16] OMB, *M-19-23 Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance* (July 10, 2019), *available at* https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf; OMB, *M-20-12 Phase 4 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Program Evaluation Standards and Practices* (March 10, 2020), *available at* https://www.whitehouse.gov/wp-content/uploads/2020/03/M-20-12.pdf. In these memos, OMB talks about "learning agendas," which are the same as the "agency evidence-building plans" required by the Evidence Act, 5 U.S.C. § 312(a). President's Management Agenda, *Federal Data Strategy 2020 Action Plan*, at 22 n.7, *available at* https://strategy.data.gov/assets/docs/2020-federal-data-strategy-action-plan.pdf.

[17] 44 U.S.C. § 3520(a).

[18] GAO Report, at 11.

[19] Evidence Act § 301.

[20] "Statistical purpose" means "the description, estimation, or analysis of the characteristics of groups, without identifying the individuals or organizations that comprise such groups; and . . . includes the development, implementation, or maintenance of methods, technical or administrative procedures, or information resources that support [those] purposes." 44 U.S.C. § 3561(12).

[21] "Identifiable form" means "any representation of information that permits the identity of the respondent to whom the information applies to be reasonably inferred by either direct or indirect means." *Id.* § 3561(7).

[22] *Id.* § 3572(c)(1).

[23] In Title III, "agency" means "any entity that falls within the definition of the term 'executive agency' … or 'agency', as defined in" Title II. *Id.* § 3561(1). "Executive agency," in turn, means "a department, agency, or instrumentality in the executive branch of the United States Government." 31 U.S.C. § 102.

[24] 44 U.S.C. § 3561(11).

[25] OMB, Statistical Programs and Standards, https://www.whitehouse.gov/omb/information-regulatory-affairs/statistical-programs-standards; *see also* White House, *Strengthening Federal Statistics*, https://www.whitehouse.gov/wp-content/uploads/2018/02/ap_15_statistics-fy2019.pdf, at 219 (Table 15-1 lists the thirteen principal federal statistical agencies).

[26] 44 U.S.C. § 3563(a)(1)(D).

[27] *Id.* § 3563(a)(2).

[28] *Id.* § 3576(e).

[29] *Id.* § 3576(b)(3). The "safeguard principles" are "(A) emphasizing to [the Designated Statistical Agencies'] officers, employees, and agents the importance of protecting the confidentiality of information in cases where the identity of individual respondents can reasonably be inferred by either direct or indirect means; (B) training their officers, employees, and agents in their legal obligations to protect the confidentiality of individually identifiable information and in the procedures that must be followed to provide access to such information; (C) implementing appropriate measures to assure the physical and electronic security of confidential data; (D) establishing a system of records that identifies individuals

accessing confidential data and the project for which the data were required; and (E) being prepared to document their compliance with safeguard principles to other agencies authorized by law to monitor such compliance." *Id.*

[30] 44 U.S.C. § 3562(a).

[31] *Id.* § 3563(c).

[32] *Id.* § 3562(b).

[33] 5 U.S.C. § 312(c).

[34] National Science Foundation, *Request for Information*, 85 Fed. Reg. 63,300 (Oct. 7, 2020), *available at* https://www.federalregister.gov/documents/2020/10/07/2020-22141/request-for-information. The website says that comments received after Dec. 31, 2020 "will be considered to the extent practicable."

[35] 44 U.S.C. § 3506(b)(2)(B)(i)(II).

[36] Before you can download this report, the Data Foundation requires you to sign up for its email list.