Heather H. Hunt
Chief, Foreign Agents Registration Act Unit
Counterintelligence and Export Control Section
National Security Division
600 E Street N.W., Bicentennial Building, Room 1300
Washington, DC 20530

Daniel Schuman
Policy Director
Demand Progress

April 7, 2017

VIA EMAIL to FARA.Public@USDOJ.GOV

**RE: Proposed Collection Activities Concerning the Foreign Agent Registration Act, OMB Number 1124-0001, 1124-0002, 124-0003, 124-0004, 124-0005, 124-0006**

Dear Director Hunt:

We are writing concerning the Justice Department's proposed extension of information collection authority under the Foreign Agents Registration Act (FARA), dated February 6, 2017, concerning the following forms: Registration Statement, Supplemental Statement, Amendment to Registration Statement, Exhibit B to Registration Statement, Short Form Registration Statement, and Exhibit A to Registration Statement.[1]

---

[1] Registration Statement 1124-0001
(https://www.federalregister.gov/documents/2017/02/06/2017-02398/agency-information-collection-activities-proposed-ecollection-ecomments-requested-extension-with), Supplemental Statement 1124-0002
(https://www.federalregister.gov/documents/2017/02/06/2017-02397/agency-information-collection-activities-proposed-ecollection-ecomments-requested-extension-with), Amendment to Registration Statement 1124-0003
(https://www.federalregister.gov/documents/2017/02/06/2017-02396/agency-information-collection-activities-proposed-ecollection-ecomments-requested-extension-with), Exhibit B to Registration Statement
(https://www.federalregister.gov/documents/2017/02/06/2017-02395/agency-information-collection-activities-proposed-ecollection-ecomments-requested-extension-with), Short Form Registration Statement
(https://www.federalregister.gov/documents/2017/02/06/2017-02405/agency-information-collection-activities-proposed-ecollection-ecomments-requested-extension-with), Exhibit A to Registration Statement
(https://www.federalregister.gov/documents/2017/02/06/2017-02390/agency-information-collection-activities-proposed-ecollection-ecomments-requested-extension-with).

Our comment responds to item 7 in the requests for public comment, which states in part:

> Personnel are in the process of developing and testing new Web form versions of its current fillable FARA registration forms with the intent of providing greater standardization, improved intuitive features, and less burdensome requirements that will benefit registrants and foreign agents who are required to register under FARA. New capabilities are expected to improve online search capabilities. NSD is confident that the new features will offer an enhanced system, promoting greater transparency.

We also respond to the following prompts:

> —Evaluate whether and if so how the quality, utility, and clarity of the information to be collected can be enhanced; and

> —Minimize the burden of the collection of information on those who are to respond, including through the use of appropriate automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g., permitting electronic submission of responses.

**BACKGROUND**

The collection and disclosure of information under FARA is the subject of sustained public interest. We believe additional steps should be undertaken to modernize the implementation of FARA so that it satisfies the purpose of the statute—facilitating the evaluation by the government and the American people of the statements and activities of agents of foreign principals—in our modern digital context.

Our recommendations are based on our experiences using FARA, including efforts by three organizations to republish FARA data online in ways that support greater access and sophisticated analysis.

- The Project on Government Oversight (POGO) built the Foreign Influence Database,[2] containing informational materials filed between 2009 and 2012.
- The Sunlight Foundation built an online tool, the Foreign Influence Explorer,[3] that contained historic data between 2008 and 2013 and some more current information. (Sunlight's tool was built in collaboration with ProPublica.)

---

[2] http://www.pogo.org/tools-and-data/foreign-influence-database/
[3] http://foreign.influenceexplorer.com/

- The Center for Responsive Politics (CRP) has taken over Sunlight's project and recently unveiled Foreign Lobby Watch.[4] That site standardizes registrants, foreign principals and locations, and parses the PDF text of each filing to allow for full-text searching. In the future, it will allow CRP and others to connect the data on foreign lobbying to other related data sets, including lobbying reported under the Lobbying Disclosure Act.

POGO and the Sunlight Foundation have issued reports with recommendations on how FARA should be improved. The Sunlight Foundation report[5] focuses on how the DOJ should fix the FARA website, which is broken in many different ways, and also release the data in a structured format. POGO's report, Loopholes, Filings Failures, and Lax Enforcement,[6] does a deep dive into the lobbying data and highlights the failings of FARA and how it is enforced, with recommendations on how the underlying law should be changed. Those recommendations and additional calls for updates are summarized in the article "Tracking Lobbying by Foreign Governments,"[7] published by Demand Progress.

We note the Justice Department's 2014 Open Government plan 3.0, which includes the following commitment (emphasis added):

> In the process of implementing the Open Government Plan, the FARA Registration Unit, in conjunction with the National Security Division Information Technology Section, has begun to assess the feasibility of generating additional features to the current online portal, which will enable the public to search, sort, and print information from the database more easily. Over the next two years, the Department will continue to review the FARA website and electronic filing system, while soliciting reasonable and concrete suggestions and feedback from the public, and will work to make feasible and appropriate modifications to the database. **Throughout this process, the Department will specifically investigate collecting and publishing registration information as structured data in a machine-readable format.**

We also note the September 2016 Justice Department Inspector General report "Audit of the National Security Division's Enforcement and Administration of the Foreign

---

[4] https://www.opensecrets.org/fara
[5] https://docs.google.com/document/d/19S-UJNx2KyJLOBRQA8P_jUbiPV1zlS1GZR7MgBRaD38/edit
[6] http://www.pogo.org/our-work/reports/2014/loopholes-filing-failures-lax-enforcement-how-the-foreign-agents-registration-act-falls-short.html
[7] https://medium.com/demand-progress/tracking-lobbying-by-foreign-governments-a8d870a12c95

Agents Registration Act,"[8] which included fourteen recommendations to improve NSD's enforcement and administration of FARA. In particular, see the Inspector General's comment with regard to e-filing:

> We continue to believe that e-file presents opportunities to better manage and ultimately improve registrant timeliness, and recommend that e-file develop with timeliness as a consideration.

## RECOMMENDATIONS

We welcome the news that the FARA Unit is developing new online capabilities (webforms) to enhance the quality, utility, and clarity of the information it gathers and releases. We believe that collecting and publishing FARA information in a machine-readable format is essential. As that topic has been under study by the FARA Unit for a number of years, we look forward to its implementation.

We make the following recommendations that we believe will enhance the quality, utility, and clarity of the information collected by the FARA Unit, and also minimize the burden of collection by moving to online mechanisms that reduce the likelihood of error and increase the reliability of the data.

### *Move Online*

We recommend the FARA Registration Unit move to a system that *requires* all filers to use its webforms, submit data in structured formats, and that outputs that data in nonproprietary, machine-readable formats. Where this data is published in bulk, it should utilize file formats designed for bulk, structured data, such as the CSV format. Data can be further improved by using templates and forms wherever possible.

### *Data Formats and Information Security*

The current e-filing system requires filers to submit data in either image or PDF formats. Yet, much of the supplemental, registration, and amendment information is originally produced in electronic formats, such as CSV files. Image and PDF formatted files destroy critical aspects of the included data, and cannot be marked for sensitive information or be used for automated calculations.  It is virtually impossible to transform a PDF into a structured spreadsheet that supports analysis and reuse of the information.

---

[8] https://oig.justice.gov/reports/2016/a1624.pdf

Users should be required to submit information in structured data formats unless they indicate it would be a hardship, i.e., they do not have access to a computer with an internet connection. FARA should provide a sample template that would allow registrants to standardize their responses in formats that could easily be processed by computers. This would allow for easy processing of data by the FARA Unit, faster review, and more uniform reporting.

Examples of problems that would be ameliorated by this change:
- Prevention of accidental release of sensitive personal information, such as bank account numbers and judicial death notices;
- Prevent submission of inaccurate reports to Congress caused by miscalculations or incomputable data;
- Datasets (spreadsheets) broken up over multiple pages, that must be then taped together.

While it is a simpler system, by way of comparison, new registrations under the Lobbying Disclosure Act *must* be filed to Congress electronically and are immediately converted to structured data.[9]

We recognize that, with respect to supplemental information, the FARA Unit may have concerns about information security. A computer virus can hide inside most types of file formats (including the popular PDF and DOCX). When accessing those files, it is important to have an up-to-date virus scanner. However, there are file formats, known as "plain-text," that are generally recognized as safe. For spreadsheets, this includes files in CSV (comma-separated value) format. For documents, this includes documents in TXT format. In addition, generating spreadsheets as CSV files, and documents as TXT files, is widely supported across many platforms, including Microsoft Office and Google Docs, so it is easy for submitters to generate the files in the required format.

In addition, problems of incomplete and inaccurate data can be addressed through use of better formats in which the data is filed, including the use of forms that ensure that users file valid information.

Examples:
- Totals of a filer's income and spending could be generated automatically.

---

[9] http://lobbyingdisclosure.house.gov/faq.html

- Registrants may indicate on a form that they have included an attachment, but then failed to submit the file. In those circumstances, registrants could be reminded to attach the file prior to submission.

*Unique Identifiers*.

Wherever possible, the FARA Unit should use unique identifiers to represent data, and wherever possible draw the identifiers from unique IDs used by other government entities or create crosswalks.[10] This allows for the interoperability of data, but also allows for the FARA Unit to validate the data upon entry and mechanically compare it to prior entries. Adding additional IDs may sound like it is harder on registrants, but if the system takes advantage of an API driven architecture and takes advantage of open data, it will increase consistency and accuracy, while making the process easier on the filer.

For example:
- Donations to campaign committees should use the FEC identifier on supplemental question 15(c) and short form question 15.[11]
- Specific contacts with members of Congress (often reported in the supplemental, usually as an attached document)[12] should use the congressional bioguide IDs[13] and common ID for the committees[14].
- Bills should use a common format.[15]

---

[10] This information is available from official sources inside Congress. Outside entities, like the Congressional Data Coalition, may also help you locate this data.

[11] Campaign committee information is published by the FEC here.
http://www.fec.gov/data/CommitteeSummary.do?format=html

[12] For an example of a supplemental, see
https://www.fara.gov/docs/3492-Supplemental-Statement-20160129-25.pdf

[13] The bioguide IDs are officially published at http://bioguide.congress.gov/biosearch/biosearch.asp. GovTrack publishes the Bioguides as a spreadsheet file at
https://www.govtrack.us/data/congress-legislators/legislators-current.csv for current members and
https://www.govtrack.us/data/congress-legislators/legislators-historic.csv for former members. That information is also available via API from Govtrack at https://www.govtrack.us/developers/api and ProPublica at https://www.propublica.org/datastore/api/propublica-congress-api. The Bulk Data Task Force, led by the Deputy Clerk of the House of Representatives, Robert Reeves, may be an invaluable point of contact for the ongoing work inside the House, Senate, Office of Legislative Counsel, Government Publishing Office, Library of Congress, and elsewhere.

[14] The FEC publishes FED committee IDs at
https://api.open.fec.gov/developers/#!/committee/get_committees. General information is at
http://fec.gov/data/DataCatalog.do?cf=downloadable, which also is published via API at
https://api.open.fec.gov/developers/.  There is also a typeahead integration where partial names can generate suggestions with standardized committee names and committee ids, available at
https://api.open.fec.gov/developers/#!/search/get_names_committees.

[15] See, for example, the Government Publishing Office's *User Guide* that explains and implements common standards to describe legislation.
https://www.gpo.gov/fdsys/bulkdata/BILLS/resources/BILLS-XML_User-Guide-v2.pdf

- OMB is in the process of publishing a machine-readable map to federal agencies, which will include unique IDs for each department, agency, and program.
- Staff who are also registered under the Lobbying Disclosure Act should be required to include their ID number.
- You should add an unique client Ids that could identify the same client across registrations. When financial and contact information is tied to a unique client ID, that would make the information easier to analyze. This is helpful for clarity in the records in the event that the client changes names or the name is reported differently across registrants.

**New fields**
Filers should be required to clearly state the original recipients of the documents and the original date of distribution.

**Autocompletion**
Files often must repeatedly file the same information for each cycle. For data that is unlikely to change, a webform could be automatically populated with that data so it need not be retyped. The user could then accept or correct that information. This may save significant time and effort for filers. It would also make review by the FARA Unit faster, as the information likely had been reviewed before and checked for errors.

**ADDITIONAL RECOMMENDATIONS**
While not squarely within the scope of updating these particular forms, we think it is important to raise the following issues as relevant to new information collections and dissemination to the public.

***Repair Bulk Data Publication***
On the FARA website, the bulk data output function in the "document search" currently exports a CSV file that apparently is broken. While it intends to include a column that hyperlinks to a document available from the FARA public office, it fails to actually link to the target document.

To view the problem, go to [http://www.fara.gov/search.html](http://www.fara.gov/search.html),  search for documents, and then click "Click here for DOWNLOADABLE DATA spreadsheet."

This can be fixed by including the actual link instead.

***Changelog***
When data on the website is changed, publish a changelog—a file that identifies only

information that has been changed and when it was changed. Currently, for users who want to know when a FARA filing has been changed/updated, they must download all the files and compare to what they currently have. A changelog will decrease the burden on the database by allowing power users to know when a file has been changed so they only have to download that particular file.

### *Termination Form*
Add a termination form, for when a registrant no longer represents a foreign entity. This will allow users to know when a registrant no longer is active.

### *Inadvertent Disclosure of Sensitive Information*
The FARA database occasionally contains sensitive information that likely should not have been filed,such as bank account and routing numbers, judicial death notices, etc. Requiring the use of electronic forms and moving away from PDFs will make it possible to prevent the release of certain data and ensure that only inappropriate information is removed if accidentally published.

However, until that comes to pass, vulnerable parties should know when (legally mandated) filings expose them to liability. In addition, the public should be aware of how frequently these inadvertent disclosures occur.

### CONCLUSION
Thank you for the opportunity to submit our recommendations concerning FARA. We would welcome the opportunity to discuss this further. Please contact Daniel Schuman, policy director at Demand Progress, at [daniel@demandprogress.org](mailto:daniel@demandprogress.org) or 202-792-4550.

Sincerely yours,

Campaign Legal Center
Center for Responsive Politics
Citizens for Responsibility and Ethics in
Washington
Demand Progress
Every Voice
Government Accountability Project
Issue One
Project On Government Oversight
Public Citizen

Sunlight Foundation
The Revolving Door Project
Norm Eisen
Richard Painter
Zephyr Teachout
James Thurber