May 28, 2020

To: 2018-2020 FOIA Advisory Committee
From: Records Management Subcommittee of the 2018-2020 FOIA Advisory Committee

Re: Records Management Subcommittee Report & Recommendations to the
          FOIA Advisory Committee

I. Introduction & Context/Background

As stated in the Archivist's Strategic Plan for 2018-2022 ("Strategic Plan"),[1] the core mission of the National Archives and Records Administration (NARA) "is to provide public access to Federal Government records," in order to "strengthen our nation's democracy." The Strategic Plan supports NARA's vision as an agency "known for cutting-edge access to extraordinary volumes of government information." One of NARA's six transformational outcomes promises that "We will embrace the primacy of electronic information in all facets of our work and position NARA to lead accordingly." In turn, NARA's Strategic Goal 1 is to "Make Access Happen," recognizing that "public access is NARA's core mission and is a higher calling that gives purposes and meaning to all our work." Strategic Goal 1 recognizes that NARA is in the business of "making access happen by delivering increasing volumes of electronic records to the American public online."

In addition, as part of Strategic Goal 3 ("Maximize NARA's Value to the Nation"), NARA states that it must provide customer agencies with policy and guidance to appropriately manage records in their custody, including email and other forms of electronic records.

The Records Management (RM) Subcommittee recognizes that the commitments made in the Strategic Plan are in harmony with recent guidance issued jointly by the Office of Management and Budget (OMB) and NARA, most notably OMB Memorandum M-19-21, "Transition to Electronic Records,"[2] dated June 28, 2019. M-19-21 directs that by December 31, 2022, all permanent records in Federal agencies will be managed electronically to the fullest extent possible, for eventual accessioning into NARA in electronic formats. It further directs that by the same date all temporary records in Federal agencies will be managed electronically to the fullest extent possible. M-19-21 also specifically reaffirms the goal of the 2012 Managing Government Records Directive[3] that Federal agencies manage all email records (both permanent and temporary) in electronic form after December 31, 2016.

We make our recommendations with an eye towards how best they can complement the important commitments set out in NARA's Strategic Plan and in M-19-21. In particular, the recordkeeping requirements set out in M-19-21 will have a profound impact on the quantity of government records in electronic form stored within agencies, a portion of which will be transferred to NARA for permanent preservation. Whether these recordkeeping policies will end up truly enhancing the public's ability to obtain agency records through FOIA remains to be determined. There are significant challenges in

---

[1] https://www.archives.gov/about/plans-reports/strategic-plan/strategic-plan-2018-2022.
[2] https://www.whitehouse.gov/wp-content/uploads/2019/06/M-19-21.pdf.
[3] https://www.archives.gov/files/records-mgmt/m-12-18.pdf.

providing access to records stored in electronic form, as the recommendations herein acknowledge and attempt to address at least in part.

II. RM Subcommittee Methodology

In connection with its ongoing work, the RM subcommittee reviewed all available FOIA Chief Officer Reports submitted to the Department of Justice's Office of Information Privacy (DOJ/OIP), as well as a sample of agency annual recordkeeping self-certification reports and Senior Agency Official for Records Management (SAORM) reports to NARA. The RM Subcommittee informally pursued lines of questioning about records management training with a liaison from NARA's Office of the Chief Records Officer. The RM Subcommittee also met with representatives of the Federal Records Officer Network (FRON), and through them conducted an informal survey on various topics of interest including recommendations on how FOIA and Federal recordkeeping practices could better be integrated. Finally, the RM Subcommittee reached its conclusions based on one or more public presentations by agency staff at FOIA Advisory Committee hearings held during this term, including in follow-up conversations with those individuals.

III. RM Subcommittee Members

Jason R. Baron, Co-Chair, Faegre Drinker Biddle & Reath LLP
James R. Jacobs, Stanford University Libraries
Ryan Law, Co-Chair, U.S. Department of the Treasury
Lee Steven, Cause of Action Institute
Thomas Susman, American Bar Association
Bradley White, U.S. Department of Homeland Security

IV. Recommendations

Proposed Recommendation #1

We recommend the Archivist request that the Department of Justice, Office of Information Policy (DOJ/OIP), issue guidance to require agencies to include records management-related materials as part of agency websites and FOIA handbooks maintained pursuant to FOIA.

Section 552(g) of FOIA, 5 U.S.C. § 552(g), requires that

> The head of each agency shall prepare and make available for public inspection in an electronic format, reference material or a guide for requesting records or information from the agency, subject to the exemptions in subsection (b), including –
>
> (1) an index of all major information systems of the agency;
> (2) a description of major information and record locator systems maintained by the agency; and
> (3) a handbook for obtaining various types and categories of public information from the agency pursuant to chapter 35 of title 44, and under this section.

Additionally, the FOIA requires that "[e]ach agency . . . shall make available for public inspection in an electronic format . . . administrative staff manuals and instructions to staff that affect a member of the public." 5 U.S.C. § 552(a)(2)(C).

The RM Subcommittee understands that agency FOIA handbooks or reference guides – which up until recently have been primarily available on agency websites – are an important resource for FOIA requesters. The establishment of the National FOIA Portal at www.FOIA.gov has in turn centralized the public's access to and has increased the availability of agency reference guides. Indeed, when submitting a FOIA request on www.FOIA.gov, an agency's reference guide is prominently featured.[4]

The RM Subcommittee reviewed samples of agency websites, as well as agency reference guides available via www.FOIA.gov, with an eye toward how agencies provide information on their internal records management practices. Our review found a wide range of differences among websites and reference guides as to the level of detail provided that would enable FOIA requesters to determine what types of agency records exist, organized by records management category. In some cases, reference guides were out of date with respect to the actual known records management practices at the agency. We also note that we failed to find any references to the fact that agencies have adopted a "Capstone" email policy requiring that all email be preserved in accordance with General Record Schedule 6.1.[5]

To maximize utility for FOIA requesters, we believe agency FOIA reference guides should contain, or provide hyperlinks to agency web pages that contain, information about:

- What records or information is publicly available without a FOIA request
- A list of major records and information systems maintained by the agency
- Descriptions of what records the agency maintains, particularly concerning frequently requested records
- Information about agency records systems and databases, not limited to Privacy Act Systems of Records
- Agency records schedules, records file plans, and records-management guidance
- Information about agency email systems, including information about the agency's Capstone email policies if applicable.

As part of the National FOIA Portal effort, we recommend that the Archivist request to DOJ/OIP that it require Federal agencies to undertake a review and update of agency FOIA reference guides to ensure that guides are incorporating records management-related information that might be useful to the public. In support of this effort, we suggest that OIP review FOIA website guidance, update the FOIA self-assessment toolkit, and require that agencies report on their efforts in their annual Chief FOIA Officer reports to make more transparent their internal records management categories and practices. NARA's Office of Government Information Services (OGIS) should support this effort by reviewing and

---

[4] In addition to supplying a FOIA reference guide, agencies typically include links to their FOIA regulations, their agency's FOIA website, and their agency's FOIA Library. Some agencies have included information about average processing times.

[5] We note that NARA maintains an excellent FOIA reference guide that provides a link to records controls schedules for Executive branch agencies. *See* https://www.archives.gov/foia/foia-guide.

suggesting improvements to agency websites and reference guides and by identifying and highlighting best practices for integrating records management features into FOIA guidance.[6]

The RM Subcommittee notes that on August 8, 2019, the Administrative Conference of the U.S (ACUS) adopted ACUS Recommendation 2019-3, on the "Public Availability of Agency Guidance Documents."[7] The RM Subcommittee believes that our foregoing Recommendation #1 is wholly consistent with and otherwise enhances the ACUS recommendation, which states at § 7 that "[a]gencies should maintain a page on their websites dedicated to informing the public about the availability of guidance documents and facilitating public access to those documents."

When updating websites and FOIA reference guides, we suggest that OIP further instruct FOIA staff to ensure they collaborate with agency Records Management and Privacy professionals.

In conjunction with making this request to DOJ/OIP, the Archivist may also wish to consider whether under 44 U.S.C. § 2904 and other oversight provisions of Title 44, the Archivist and OMB jointly issue guidance directing agencies to publish on agency websites the records management materials discussed here.

Under the Federal Records Act, the Archivist has the general authority to "provide guidance and assistance to Federal agencies" to ensure "economical and efficient records management"; the "adequate and proper documentation of the policies and transactions of the Federal Government"; and "proper records disposition." 44 U.S.C. § 2904(a). Section 2904 goes on to state that the Archivist has the responsibility "to *direct* the continuing attention of Federal agencies and the Congress on the need for adequate policies governing records management." *Id.,* § 2904(c)(5) (emphasis added).

Additionally, as implemented through the Paperwork Reduction Act, Pub. L. 104-13, OMB, acting through its Office of Information and Regulatory Affairs (OIRA), has general oversight authority with respect to the "use of information resources to improve the efficiency and effectiveness of governmental operations . . . ." 44 U.S.C. § 3504(a). This provision further provides that the Director of OIRA "develop, coordinate and oversee the implementation of Federal information resources management policies, principles, standards, and guidelines," as well as "provide direction and oversee records management activities." *Id.*, § 3504(a)(1)(A) & B(iv).

By virtue of their statutory authorities, the Archivist and OMB from time to time have issued joint guidance to all Executive branch agencies. *See, e.g.,* M-19-21, "Transition to Electronic Records," dated June 28, 2019[8]; M-12-18, "Managing Government Records Directive," dated August 24, 2012.[9]

Accordingly, as an alternative or supplemental course of action, the Archivist may wish to consider that he and OMB jointly issue guidance directing Federal agencies to (i) undertake a review of their existing

---

[6] We understand that exceptions may be made for records management-related materials containing classified information or information withheld under a specified statute in accordance with 5 U.S.C. § 552(b)(3) of the FOIA. *See* https://www.archives.gov/records-mgmt/faqs/rcs.html#twentyhttps://www.nsa.gov/news-features/declassified-documents/records-management/.

[7] 84 Fed. Reg. 38927.

[8] *See* n.2, *supra.*

[9] https://www.archives.gov/files/records-mgmt/m-12-18.pdf.

records management policies, procedures, and directives, including but not limited to agency records schedules, records file plans, and agency records management guidance, and (ii) publish these materials online in one or more places, including on web pages devoted to records management and/or on FOIA pages, per our recommendation above. This proposed joint guidance from OMB and NARA should be continuing in nature, so as to allow for further publication as warranted whenever records-management policies are updated.

Proposed Recommendation # 2

We recommend that the Archivist of the United States direct NARA and request that DOJ/OIP offer targeted training in selected topics in Federal records management to FOIA officers and FOIA Public Liaisons in Federal agencies, and otherwise include a FOIA module in selected records management training courses open to all Federal employees.

As referenced earlier, the RM Subcommittee undertook a review of all 2019 agency FOIA Chief Officer Reports submitted to DOJ/OIP posted as of August 2019. In section I of the reports, question 4 required agencies to provide a brief description of the type of training attended or conducted and the topics covered. As a general matter, we found that DOJ/OIP provides an extensive array of FOIA courses for FOIA access professionals, FOIA Public Liaisons, and FOIA Reference Services staff – and it is clear that many agencies are taking advantage of the DOJ/OIP course offerings. Additionally, agencies conduct their own internal FOIA training and sometimes report using outside academic and commercial training offered by such organizations as the American Society of Access Professionals (ASAP) and Graduate School USA. Some of these courses may touch on records management issues in passing, but generally do not explore them in any depth.

With respect to Records Management courses offered by NARA, they are open to all Federal employees. NARA's program focuses on policies and procedures unique to the Federal environment. The training program covers every aspect of Federal Records Management and represents an important step in acquiring the skills necessary to manage Federal records. NARA does not, however, offer a specific course targeted to FOIA staff, nor have we been made aware of any available government-wide statistics indicating how many Federal employees who are engaged in FOIA-related functions have undertaken records management training.[10]

Based on our review of agency reports from both DOJ/OIP and NARA, and from all information made available to the RM Subcommittee, we believe that FOIA officers would benefit from a more in-depth understanding of the current state of Federal recordkeeping within agencies, especially with respect to how agencies are managing records in electronic or digital form. Topics that we believe would be of benefit and interest to FOIA staff might include an overview of:

> (a) what "adequate documentation" of agency activities consists of under the Federal
>   Records Act and NARA regulations;
> (b) the function of agency file plans and records schedules describing individual record
>   series and their retention or disposition as temporary and permanent records of
>   agencies;

---

[10] We have been advised that between FY 2014 and FY 2020, a total of 57 Government Information Specialists (GISs) in the Record/Information Dissemination Section at the Federal Bureau of Investigation have received a NARA Certificate of Federal Records Management Training.

(c) recent initiatives involving the transition to electronic recordkeeping in government
(e.g., M-19-21)
(d) NARA's Capstone policy for email recordkeeping, including GRS 6.1; and
(e) best practices in conducting electronic records searches to find responsive agency
records.

The RM Subcommittee believes that FOIA officers handling requests that necessarily involve searches of email records, as well as other structured and unstructured electronic records, should be aware of the overall approach taken by their agency to electronic records management. Armed with this knowledge, FOIA personnel should have a greater appreciation for what constitutes an "adequate" search for agency records responsive to particular requests, and be able to better execute their overall FOIA job responsibilities.

Records management training for FOIA staff could take the form of a separate module embedded within one or more current course offerings by NARA/OGIS and/or DOJ/OIP, or could be developed as separate stand-alone training, to be conducted either in person or on-line. The RM Subcommittee hopes that OGIS staff would take the lead in continuing to work with DOJ/OIP in the latter providing an RM module as part of DOJ FOIA courses. Agencies may also wish to consider offering in-house records management training to their own agency's FOIA personnel.

We further recommend as a best practice that agencies encourage FOIA staff to attend records management training in whatever form offered, and to have FOIA staff report what records management-related course(s) they have attended in the cited section above of the annual Chief FOIA Officer reports.

Finally, we believe all Federal employees would benefit from understanding that managing Federal records includes providing access to those records, including through FOIA.

Proposed Recommendation # 3

We recommend that the Archivist of the United States request that DOJ/OIP provide further best practice guidance on the use of e-discovery tools to assist agencies in meeting their obligations to conduct an adequate search of electronic records, including but not limited to email in Capstone repositories.

The RM Subcommittee reviewed the set of agencies' 2019 Chief FOIA Officer Reports also with respect to agency responses to Section IV Question 1 (as added in response to a recommendation from the 2016-2018 FOIA Advisory Committee). Question 1 asked:

Is your agency leveraging technology to facilitate efficiency in conducting searches, including searches for email? If so, please describe the type of technology used. If not, please explain why and please describe the typical search process used instead.

Except for some of the smallest agencies, most agencies reported that they either have or are considering using some form of technology to aid in the FOIA workflow process. This has included the use of forms of automated collection software (rather than relying on manual processes by individual employees), and of various forms of FOIA redaction software. Additionally, a modest number of agencies report using some form of e-discovery search tools to conduct searches of agency records. For

the most part, the descriptions given in these reports of e-discovery search tools failed to reveal if advanced search tools utilizing software beyond relying on keyword searching were part of the software capabilities acquired.

There was no express mention by any agency in any of the Chief FOIA Officer reports of the use of "predictive coding" or "technology assisted review."[11] This technology, a form of artificial intelligence using machine learning, has been adopted over the past decade by select law firms and corporations for the purpose of conducting more accurate and efficient searches at substantially reduced cost in legal e-discovery and investigations.[12] The RM Subcommittee is aware (based on anecdotal reporting) that such technology is in use in various select components of government in connection with ongoing litigation, as well as in processing data in connection with "second requests" filed in antitrust proceedings.

Based on the latest NARA Senior Agency Official for Records Management annual reports, on the order of 200 reporting components of the Executive branch have stated that they have adopted or are in the process of adopting NARA's "Capstone" approach to managing email, which provides a way of meeting the goals of M-19-21 with respect to the management of email.[13] An agency that adopts a Capstone policy necessarily commits to managing a substantial volume of email: programmatic email from all employees will be preserved for seven years under the GRS in some form of email repository, and a subset of that email from designated senior officials (Capstone account holders) will be deemed permanent and preserved in agency repositories until such time as it is transferred (i.e., accessioned) into the National Archives at a future date as agreed to by the agency and NARA.

The RM Subcommittee believes that agencies that adopt the Capstone approach for managing their email records necessarily will see the volume of email records grow to a substantial number, potentially in the hundreds of thousands to many millions depending on existing email volumes at the agency. This future circumstance will all but necessitate serious consideration of more advanced and more efficient ways in which to search through these large volumes of records to find responsive records. Although there is limited FOIA case law on point,[14] the RM Subcommittee believes that courts will exhibit decreasing tolerance of agency arguments that because they are without e-discovery tools to

---

[11] *See, e.g.,* The Sedona Conference, Best Practices Commentary on The Use of Search and Information Retrieval Methods in E-Discovery, 15 Sedona Conf. J. 217 (2014), www.thesedonaconference/publications/publications.
[12] *See* Nicolas M. Pace & Laura Zakaras, "Where the Money Goes: Understanding Litigant Expenditures for Producing Electronic Discovery" (RAND Corp. 2012), https://www.rand.org/content/dam/rand/pubs/monographs/2012/RAND_MG1208.pdf.
[13] https://www.archives.gov/records-mgmt/resources/saorm-reports.
[14] *See, e.g., Open Society Justice Initiative v Central Intelligence Agency, et al.* 399 F. Supp. 3d 161, 168 (S.D.N.Y. 2019) (where the Department of Defense (DOD) claimed an unreasonable burden in processing 5000 pages per month, the Court observed that a declarant from DOD's Office of Information Counsel "notes the remarkable fact that the . . . DOD Department that oversees FOIA requests . . . does not possess eDiscovery software," finding that the means for searching was "antiquated" and going on to hold that the Court "must focus on a reasonable agency's technological capability" in assessing search adequacy); *Leopold v. National Security* Agency, 196 F. Supp. 3d 67, 76 (D.D.C. 2016) (in a case where DOJ had existing e-discovery software, the Court ordered the DOJ Office of Legal Counsel "to use the Clearwell eDiscovery tool . . . to search the email files of departed OLC attorneys, as well as any attachments to those emails, for any draft legal memoranda or opinions relating to the propriety of surveilling federal or state judges."); *Nat'l Day Laborer Org. Network v. U.S. Immigration & Customs Enf't Agency,* 877 F. Supp. 2d 87, 95 (S.D.N.Y. 2012) ("If [the parties] wish to and are able to, then they may agree on predictive coding and other more innovative ways to search."); *cf. Bigwood v. U.S. Department of Defense*, 132 F. Supp. 3d 124, 140, 142 (in response to plaintiff's arguments including that DOD had employed an "ineffective search

perform efficient searches of email and other forms of electronic records, they should continue to be allowed to conduct searches at the pace of manual efforts from an earlier day. Regardless, however, of future case precedent, efficient government fairly demands that agencies consider a range of available tools to conduct faster and more efficient searches against the growing volumes of electronic records they manage, including but not limited to in Capstone repositories.

The RM Subcommittee recommends that the Archivist make a request to DOJ/OIP to issue guidance by way of a FOIA bulletin or otherwise that commends agency adoption of e-discovery tools and encourages agencies to become aware of advanced search methods that may enhance the ability to find responsive agency records on a more expedited basis.

The RM Subcommittee further recognizes that newer forms of electronic messaging available online are increasingly being used in the workplace as a means of conducting official government business. As part of any guidance DOJ/OIP may issue regarding search technologies, DOJ/OIP should also consider reminding agencies of their obligations to conduct adequate searches of all records created using electronic messaging that relate to the conduct of government business. *See* 44 U.S.C. § 2911 (electronic messages sent on commercial servers constituting Federal records should be forwarded or copied to a ".gov" or ".mil" account).

In Appendix A to this Report, the RM Subcommittee proposes a brief set of "best practice" suggestions with respect to using e-discovery tools to conduct more efficient searches for responsive records. We recommend that as part of any guidance to be issued, DOJ/OIP and NARA/OGIS consider adopting or further publicizing this checklist for use in the greater FOIA community of agencies. Agencies may also deem Appendix A useful in developing RFIs or RFPs, including under current and future GSA schedules.

The RM subcommittee understands that not all Federal agencies may desire or otherwise be in need of e-discovery software to assist in conducting FOIA searches. In particular, smaller agencies and agencies that receive few FOIA requests annually may not yet be seeing a volume of electronic records that would justify the use of more automated methods. On the other hand, the RM subcommittee believes that larger Cabinet departments and agencies should give serious consideration to using more advanced technology to increase their efficiency and lower their overall cost burden in responding to large-volume FOIA requests.

Proposed Recommendation # 4

We recommend that as part of the Federal Electronic Records Modernization Initiative (FERMI), the Archivist of the United States direct NARA to incorporate and further develop the idea of public access to Federal records, including through FOIA.

In undertaking FERMI, NARA is attempting to provide the Executive branch with standardized and interoperable records management solutions and services to Federal agencies. NARA recognizes that

---

strategy" by using compound phrases in keyword searching, the Court found that a "FOIA petitioner cannot dictate the search terms for his or her FOIA request; that "[w]here the search terms are reasonably calculated to lead to responsive documents, a court should neither 'micromanage' nor second guess the agency's search," and that an "agency's search description [is] sufficient despite the absence of information concerning any Boolean operators or connectors used to facilitate an electronic search," citing *Citizens for Responsibility and Ethics in Washington v. National Indian Gaming Commission,* 467 F. Supp. 2d 40, 50 (D.D.C. 2006)).

agencies have common needs for managing their electronic records. The RM Subcommittee believes that a critical aspect of managing electronic records is providing for economical and efficient public access to those records. The RM Subcommittee further believes that incorporating the idea of public access more expressly into the FERMI's baseline requirements may be one way in which the issue of access to a rapidly rising volume of agency records in electronic or digitized form can be addressed.

Within FERMI, the Universal Electronic Records Management (ERM) Requirements comprise six sections based on the lifecycle of electronic records management: capture, maintenance and use, disposal, transfer, metadata, and reporting. "Access," while not expressly identified as a category, is implicitly subsumed within the second category, "Maintenance and Use."[15] Under this category, one of its subcomponent requirements as detailed in an accompanying spreadsheet provides for the following specification:

> 2.01 *Records of current and former employees must be managed in a manner that supports searching in response to information requests, including FOIA and agency business needs*.[16]

The RM Subcommittee believes NARA should consider further highlighting and developing what constitute "Access" requirements that are common to Federal agencies. One way of doing so is in the development of a FERMI "Use case" involving FOIA access, in order to make clear what additional functional requirements are necessary to efficiently process FOIA requests for responsive records found in large electronic or digital agency repositories. These requirements should include, but not be limited to, the use of efficient forms of search technologies for locating responsive records.

The RM Subcommittee believes that NARA should be open to considering additional ways in which access issues can be highlighted in connection with records management initiatives. For example, NARA may wish to consider asking agencies to assist in responding to informal supplemental questions to the annual Federal Email Management Reports,[17] Senior Agency Officials for Records Management Annual Reports,[18] and Records Management Self-Assessment Reports[19] aimed at eliciting how electronic records are being accessed within agency repositories. For example, just with respect to Capstone email policies, NARA could ask agencies (i) to report on the estimated volume of email records being managed in a Capstone email repository, and (ii) to provide updates on how Capstone repositories are being searched and which types of software are being used for these searches. Through this process, NARA may be able to improve its own policy guidance on the available means for capture and management of email records via Capstone programs.

Proposed Recommendation # 5

We recommend that the Archivist of the United States make a formal request to the Chair of the Council of the Inspectors General on Integrity and Efficiency (CIGIE)[20] that CIGIE consider designating

---

[15] Accord 44 U.S.C. § 3102(1) (each Federal agency head shall establish "effective controls over the creation and over the *maintenance and use* of records in the conduct of current business") (emphasis added).

[16] https://www.archives.gov/records-mgmt/policy/universalermrequirements (see accompanying spreadsheet).

[17] https://www.archives.gov/records-mgmt/resources/email-mgmt-reports.

[18] https://www.archives.gov/records-mgmt/resources/saorm-reports.

[19] https://www.archives.gov/records-mgmt/resources/self-assessment.html.

[20] https://ignet.gov/.

as a cross-cutting project or priority area the issue of how agencies are doing in providing FOIA access to agency records in electronic or digital form.

Based on a set of presentations by representatives of the greater Inspector General (IG) community at the FOIA Advisory Committee meeting of November 29, 2018,[21] and a further review of selected reports from agency IGs on the subject records management, the RM Subcommittee believes that CIGIE is well-positioned to serve in an oversight capacity to monitor how well agencies are doing in providing access to agency records, especially in electronic repositories. CIGIE is an independent entity established within the Executive branch that addresses integrity, economy, and effectiveness issues "that transcend individual Government agencies."[22]

The FOIA Advisory Committee heard several representatives of IG offices provide updates on recent audits involving the controls agencies had put into place on the subject of electronic records management. In response to a question posed by the FOIA Advisory Committee, one of the IG representatives indicated that the areas of records management and FOIA were "potential area[s] for CIGIE to convene a group to do a cross-cutting project or compile FOIA and records management audit reports."[23] Archivist David Ferriero later indicated in the meeting that "for several years, NARA's Inspector General has proposed that CIGIE take up records management as a cross-cutting project."

The RM Subcommittee recommends that the Archivist of the United States, either on his own or in conjunction with NARA's IG, should formally request that CIGIE elevate the issue of how agencies are doing in providing access to records managed in electronic form. Initially, this could take the form of an audit or review of how agencies are planning to meet the goals set out in M-19-12, including the 2022 deadline for ensuring that all permanently appraised records are accessioned into NARA in electronic or digitized form.

The RM Subcommittee further believes that IG offices should consider establishing a point of contact within each office who will function as an in-house subject matter expert on records management and access issues and who will keep track of any audits conducted with respect to records management and access issues.

Proposed Recommendation # 6

We recommend that the Archivist of the United States direct NARA and request that DOJ/OIP each establish a liaison with the newly created Chief Data Officer (CDO) Council, for the purpose of ensuring that CDO officials understand the importance of Federal recordkeeping and FOIA requirements and how such laws apply to the maintenance of data within agencies.

Under the OPEN Government Data Act, enacted on January 14, 2019,[24] each agency was to have a designated CDO, and that these positions together would comprise a newly-formed CDO council.[25] Pursuant to further implementing guidance issued by OMB, CDO designations were to be made by July

---

[21] https://youtu.be/32o6mqQ2Nw0.

[22] *Id.*

[23] *See* Minutes of Meeting, at 4, https://www.archives.gov/files/ogis/assets/foiaac-mtg-min-11-29-2018.pdf.

[24] *See* Title II of the Foundations for Evidence-Based Policymaking Act, Pub. L. 115-435, 132 Stat. 5529, 5534.

[25] 44 U.S.C. §§ 3520 & 3520A.

2019, and CDOs are expected to lead the creation of a new Data Governance Body at each agency. *See* OMB Memorandum M-19-23, dated July 10, 2019.[26]

Senior agency officials who are to serve on each agency's Data Governance Body are expected to "set and enforce priorities for managing data as a strategic asset to support the agency in meeting its mission. . . ."[27] Appendix C to the OMB Memo states that each agency's Data Governance Body is to include the agency's General Counsel, Chief Freedom of Information Act Officer, and Senior Agency Official for Records Management, among many others.

The RM Subcommittee notes that neither the Evidence Act itself, nor OMB Memo M-19-23 in providing implementing guidance, expressly references the Federal Records Act (FRA) and FOIA as existing statutory authorities that play important roles in the management of Federal data. In light of this fact, the RM Subcommittee believes an opportunity exists for NARA, through OGIS and other components, and for DOJ/OIP, to play an important, continuing role in educating the greater "open data" community regarding how Federal data assets are currently managed under the FRA and FOIA.

More generally, the RM Subcommittee recommends that both DOJ/OIP and NARA/OGIS should work to align FOIA policy and FERMI with the Administration's overall Federal Data Strategy,[28] including as set out in the President's Management Agenda[29] and Reform Plan.[30] In particular, we believe it will be increasingly important that Federal staff understand that federally created data in numeric form should be included within FOIA and Federal records management workflows. Aligning these policies will in our view facilitate the proactive release of Federal data, expand on the available catalog of data on Data.gov, standardize the release of agency data in open standards and machine-readable formats, and enable better public use of agency-created data from across the Federal government.

Proposed Recommendation # 7

We recommend that the Archivist of the United States work with other governmental components and industry in promoting research into using artificial intelligence (AI), including machine learning technologies, to (i) improve the ability to search through government electronic record repositories for responsive records, and (ii) identify sensitive material for potential segregation in government records, including but not limited to material otherwise within the scope of existing FOIA exemptions and exclusions.

The 2016-2018 FOIA Advisory Committee made two specific recommendations related to search technologies. The first recommendation resulted in the creation of a Technology Committee of the Chief FOIA Officers Council to study the utilization and deployment of FOIA technology across agencies and to identify best practices and recommendations that could be implemented. A Technology

---

[26] https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf.

[27] *Id.*

[28] https://strategy.data.gov/overview (explaining that the mission of the Federal Data Strategy is "to leverage the full value of Federal data for mission, service and the public good . . . .")

[29] https://www.whitehouse.gov/wp-content/uploads/2018/03/Presidents-Management-Agenda.pdf.

[30] https://www.whitehouse.gov/wp-content/uploads/2018/06/Government-Reform-and-Reorg-Plan.pdf.

Committee was formed, and its final report was published on February 14, 2020.[31] As referenced above in Recommendation # 3, a second recommendation resulted in OIP's collecting detailed information, as part of each agency's CFO report, regarding specific methods and technologies agencies are using to search their electronic records, including email.

As discussed above, based on its review of the 2019 Chief FOIA Officer Reports the RM Subcommittee believes that Federal agency FOIA staff do not appear to be well-versed in how AI and machine learning technologies may improve the efficiency of FOIA searching in ever-growing digital repositories. Nevertheless, the RM Subcommittee understands that such advanced tools and technologies are readily available by solution providers, to assist FOIA staff in undertaking complex searches of large repositories of electronically stored information. This software has not, however, been generally deployed in the context of FOIA searching, nor has it been developed with an eye towards the types of record content with a range of sensitivities (e.g., personally identifiable information) found within components of the Federal government. See Technology Committee Report at 5 (quoting agency report saying "Often the technology that FOIA professionals leverage to search and review records are not specifically designed with FOIA in mind. . . ").

The RM Subcommittee notes that on February 11, 2019, an "Executive Order on Maintaining American Leadership in Artificial Intelligence" (EO 13859) was issued, calling for AI efforts within government to be coordinated through the National Science and Technology Council (NSTC) Select Committee on Artificial Intelligence. Among other things, the EO states that in accordance with the President's Management Agenda, "agencies shall identify opportunities to use new technologies and best practices to increase access to and usability of open data and models, and explore appropriate controls on access to sensitive or restricted data and models, consistent with applicable laws and policies. . . . "

The recommended initiative here would be focused on promoting the use of advanced search capabilities to help solve issues that Federal agencies are only now beginning to confront, including searching large repositories of preserved emails pursuant to NARA's Capstone policy, and identifying or filtering sensitive content – including but not limited to FOIA-exempt materials – so as to more timely and efficiently respond to access requests of all kinds. The initiative would be completely in keeping with NARA's announced strategy to "[e]xplore cutting-edge technologies such as advanced search to automate processing of large volumes of electronic records."[32]

We wish to emphasize that the intended use of advanced AI technologies such as machine learning would be to supplement – not replace – the continued involvement of FOIA subject matter experts in segregating FOIA-exempt material in specific documents.

The RM Subcommittee recommends that the Archivist, either acting through or in conjunction with NARA's CDO, seek input from a variety of governmental components, including research components of NARA, the General Services Administration, the Networking and Information Technology Research and Development Program (NITRD), and the above-referenced NSTC. The initiative could be coordinated with OMB's CIO and CDO councils, OSTP's CTO, NIST, and designated Chief Privacy Officers from selected

---

[31] *See* Report of the Technology Committee of the Chief FOIA Officers (CFO) Council – Best Practices and Recommendations ("Technology Committee Report"), https://www.archives.gov/files/ogis/assets/cfoc-tech-comm-final.report-02-14-2020.pdf.

[32] Strategic Plan, Goal 1.1.

agencies. Consideration should also be given to establishing private-public partnerships to work with the commercial sector on any such initiative.

Proposed Recommendation # 8

We recommend that the Archivist of the United States request that NARA/OGIS work together with DOJ/OIP to encourage agencies to work towards the goal of collecting, describing, and giving access to FOIA-released records in one or more central repositories in standardized ways, in addition to providing access on agency websites.

As amended by the FOIA Improvement Act of 2016, FOIA states that agencies "shall make available for public inspection . . . copies of all records, regardless of format . . . that have been released to any person," where "because of the nature of their subject matter, the agency determines that [the records] have become or are likely to become the subject of subsequent requests for substantially the same records," or that otherwise "have been requested three or more times."[33]

DOJ/OIP currently hosts guidance for agencies regarding FOIA websites, under the title "Agency FOIA Websites 2.0."[34] As the Guidance explains:

Agency websites have evolved as technical capabilities have improved over time and internet use has become a primary method for information dissemination. While agencies have updated their websites over the years to account for new guidelines or policies, two of the most visible reflections of this evolution are in the websites' style and design and the volume of material posted on them. The style of agency websites has gone from simple pages, containing a small amount of information and links to the far more comprehensive sites seen   today that are designed with a uniform style and contain a wealth of information about agency operations, including administration of the FOIA. One of the benefits of this evolution is that agencies can create a unique online identity that is easily recognized and best serves their community of users.

The DOJ/OIP Guidance does a careful job in providing detailed guidance on how agency FOIA websites should be constructed, including with respect to what key information and resources should be contained on the website. In addition, following the issuance of DOJ's Open Government Plan 2.0, in 2013 DOJ provided guidance on "Using Metadata as the Foundation for a Government-wide FOIA library."[35] As the Guidance itself acknowledges, however, agencies continue to have wide leeway in how their online web pages are structured, including what formats are used, how or even if documents are described (e.g., using keywords and other forms of metadata tagging), and what if anything they post to these sites. Further, these sites may be difficult for the public to find, navigate, and search. Content description is especially important for public access. For example, Data.gov uses a robust metadata schema[36] with a set of required fields (Title, Description, Tags, Last Update, Publisher, Contact Name, etc.) for every data set. Describing FOIA content across federal agencies in a similar, standardized way would facilitate search and thefaceting of search results, as well as enhancing application programming interfaces (APIs) and bulk access to FOIA content.

---

[33] *See* 5 U.S.C. § 552(a)(2)(D)(i)-(iii).

[34] https://www.justice.gov/oip/oip-guidance/OIP%20Guidance%3A%20%20Agency%20FOIA%20Websites%202.0.

[35] https://www.justice.gov/oip/blog/using-metadata-foundation-government-wide-foia-library.

[36] Data.gov follows the DCAT-US Schema v1.1 (Project Open Data Metadata Schema).

The unfortunate consequences of non-standard posting and non-standard metadata description – or lack thereof – of FOIA records, coupled with the myriad agency FOIA libraries (formerly known as electronic reading rooms) that exist, may include (i) agencies ending up duplicating their efforts by receiving unnecessary queries for FOIA records already released, and (ii) the public not otherwise being provided adequate access to FOIA-requested records.

A small number of federal agencies currently post records released via FOIA requests in a central digital repository, FOIAonline.gov.[37] Additionally, some agencies support web pages on FOIA.gov that act as links or "pointers" to agency FOIA websites. These pointers are not, however, standardized in a way that optimizes public access to the underlying materials or works toward a government-wide FOIA library.

The RM Subcommittee recommends encouraging more agencies to (i) use an existing centralized FOIA records portal like FOIAonline.gov, (ii) develop metadata standards building on the above-referenced DOJ/OIP metadata guidance, and (iii) upgrade and standardize their links in FOIA.gov to documents posted on agency FOIA web pages. These actions will serve dual purposes: First, doing so will save agency time in FOIA processing and FOIA Library administration. Second, it will facilitate and expand public access to proactively released records from across the government, especially if accompanied by more standardized metadata.[38]

Proposed Recommendation # 9

We recommend that the Archivist of the United States request NARA/OGIS to work together with DOJ/OIP to encourage agencies to release FOIA documents to the public on their FOIA Websites, and in FOIA portals in open, legible, machine-readable and machine-actionable formats, to the extent feasible.

The current means of FOIA access for researchers and the public at large is still primarily through receipt of paper-based documents. Although enactment of the Electronic Freedom of Information Act Amendments of 1996 required Federal agencies to honor FOIA requests for records in electronic form,[39] in the vast majority of cases record requesters asking for documents in electronic form often obtain scanned documents that have not undergone optical character recognition (OCR) processing. Scanned textual documents of all types, along with handwritten documents, images, and photographs, are all very difficult to search, analyze (*i.e.*, use for analytics), or otherwise manipulate in any fashion. They also may have legibility issues due to poor reproduction.[40] Subsequent amendments to the FOIA requiring the uploading of documents to electronic reading rooms and agency websites have not in the main changed agency practices with respect to providing the public with access to scanned documents or images.

By contrast, documents in "machine readable" formats are structured in a manner that can be processed by a computer. As defined in the OPEN Government Data Act, data considered to be "machine readable" is "in a format that can be easily processed by a computer without human

---

[37] https://foiaonline.gov/foiaonline/action/public/about.

[38] *See, e.g.,* Data.gov as one possible model of a central standardized metadata repository pointing to documents on agency websites.

[39] 5 U.S.C. § 552(a)(3)B), as amended by Pub. L. 104-231 (1996).

[40] We note also that sometimes original source documents have legibility issues themselves.

intervention while ensuring no semantic meaning is lost."[41] A further definition of "machine readable" contained in a former version of OMB A-11 explains in greater detail that a "machine readable format" is a:

> Format in a standard computer language (not English text) that can be read automatically by a web browser or computer system (e.g., xml). Traditional word processing documents and portable document format (PDF) files are easily read by humans but typically are difficult for machines to interpret. Other formats such as extensible markup language (XML), (JSON), or spreadsheets with header columns that can be exported as comma separated values (CSV) are machine readable formats. It is possible to make traditional word processing documents and other formats machine readable but the documents must include enhanced structural elements.[42]

For our purposes here, the use of the further term "machine-actionable" highlights the key aspect that structured formats allow for greater manipulation of the content of data and records through use of software.[43]

The Federal government has in the last several years moved inexorably toward more transparent government in the form of embracing "open" data. In 2009, the White House issued an Open Government Directive,[44] which instructed agencies to "take prompt steps to expand access to information by making it available online in open formats," defined as "platform independent, machine readable, and made available to the public without restrictions that would impede the re-use of that information."[45] Along similar lines, on May 9, 2013, President Obama signed Executive Order 13642 "Making Open and Machine Readable the New Default for Government Information." These policies served to focus federal agencies on the task of publishing a greater amount of their data and records in open, machine-readable and machine-actionable formats.

Under the OPEN Government Data Act, each agency is to ensure that its "public data assets" are machine-readable, and available as an open Government data asset.[46] Without our taking a position on whether the definition of "public data asset"[47] in the law applies to agency FOIA responses, the RM Subcommittee nevertheless note that the new law underscores that the government at large is increasingly being encouraged and expected – through both legislation and Executive branch policies – to generally embrace open, machine-readable formats when releasing information to the public.[48]

---

[41] *See* 44 U.S.C. § 3502(18).

[42] https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/a11_current_year/s200.pdf.

[43] *See, e.g.*, https://ddialliance.org/taxonomy/term/198 (defining "machine actionable" as "information that is structured in a consistent way so that machines, or computers, can be programmed against the structure").

[44] https://obamawhitehouse.archives.gov/open/documents/open-government-directive.

[45] Examples of "open" file formats include, but are not limited to, XML, TXT, HTML, PDF, CSV, and MP3.

[46] *See* 44 U.S.C. § 3506(b)(6) & 3506(d)(5), as added by Pub. L. 115-435, Title II, § 202(c)(1) & (3), 132 Stat. 5536, 5537. These provisions became effective one year after the original enactment date of the Act. 44 U.S.C. § 3506 Note.

[47] The law defined "data asset" to mean "a collection of data elements or data sets that may be grouped together." *Id*., Title II, § 202(a). In turn, the term "public data assets" means "a data asset, or part thereof, maintained by the Federal government that has been, or may be released to the public, including any data asset, or part thereof, subject to disclosure under section 552 of title 5 [the FOIA]."

[48] We defer to guidance expected to be issued by OMB on the scope of newly amended 44 U.S.C. § 3506(b) & (d).

There are already examples where both NARA and DOJ are embracing machine readable formatting of government information. NARA recently posted to Data.gov searchable and manipulable versions of the General Records Schedules (GRS) in comma-separated value (.csv) format, to test the usefulness of records schedule data in this format.[49]

The *FOIA Annual Report XML schema*[50] is another leading example of the use of a machine-readable format to publish information otherwise contained in FOIA annual reports. Using a similar XML schema would be one way in which agencies could accomplish making FOIA documents searchable, sortable, downloadable in bulk, and machine-actionable. By publishing FOIA documents in this way, the government chooses to facilitate machine-learning, artificial intelligence and advanced indexing, and allows the public to find and use FOIA documents in new and more efficient and effective ways – leveraging technology to improve FOIA administration across the government. Combining machine-readable and actionable formats with standardized descriptive metadata will move the government one step closer to what some see as the aspirational goal of a government-wide FOIA library.

The existing universe of FOIA responses simply does not leverage or take advantage of "open" formats of software Internet technologies. To bring FOIA in line with open government best practices, we recommend that NARA/OGIS work with DOJ/OIP to encourage agencies to release their FOIA documents in both human-readable and structured, machine-readable and machine-actionable formats, to the extent feasible using available technology.[51]

---

[49] https://records-express.blogs.archives.gov/2020/02/10/grs-uploaded-to-data-gov/
[50] https://www.foia.gov/developer/.
[51] https://obamawhitehouse.archives.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-

2018-2020 FOIA Advisory Committee
Records Management Subcommittee Report

Appendix A

E-Discovery Best Practices Checklist for Use in Responding to FOIA Requests

The following consists of a simplified checklist of best practices that agencies may find useful when considering the use of e-discovery tools to perform FOIA searches. References to selected further reading providing more comprehensive e-discovery protocols and workflows are also included.[52]

1. *Preservation*

An agency's responsibility for taking reasonable steps to preserve records subject to a pending FOIA request includes ensuring that electronic records are preserved. Custodians of responsive electronic records should be identified and placed on notice of the need to take preservation actions. As appropriate, this may include suspending auto-delete functions for email and elsewhere, as well as taking reasonable steps to preserve select backup media.[53]

2. *Collection and Search*

Agency records in electronic form may be found on network servers, local servers, databases, portable media (e.g., laptops, mobile devices and phones), in "the cloud," and on agency-controlled social media. Email may reside in a Capstone repository.[54] All forms of documents (structured and unstructured text, audio, video, graphics, foreign language, etc.) may be within the scope of a request. Documents collected in native form should include their associated metadata. The use of optical character recognition (OCR) software promotes search capability for scanned documents.

Keyword searching should be employed with Boolean operators ("and," "or," and "not").

Errors in keywords being misspelled should be accounted for through the use of wildcards (e.g., "*" for alternative individual letters, and "!" for alternative extensions of truncated keywords).

Predictive coding (also known as technology assisted review) should be considered when a review is to be conducted of very large universes of documents (exceeding tens of thousands). These advanced search methods use human reviewers to first code agency records as responsive, for the purpose of training software to identify further responsive records out of a much larger universe of documents.

Any process employed should be designed in a repeatable or standardized way that can be documented for further explanation to a requester or a tribunal.

---

[52] Agency FOIA professionals should also consult with their respective records management officials for additional agency-specific guidance.

[53] Examples of "backup media" include backup tapes, optical drives, flash drives, and cloud storage.

[54] Under NARA's "Capstone approach" for managing email, email is managed at the account level, where selected senior officials have their email accounts designated as permanent records, with other staff's email accounts considered temporary records. *See* https://www.archives.gov/files/records-mgmt/grs/grs06-1.pdf.

Any process employed should include quality control features, such as random sampling of the portion of the universe of documents searched that are not initially considered "responsive."

3. *Culling/De-duplication*

Culling (also referred to as filtering) documents by custodian or date range will assist in filtering requests. As appropriate, determine whether certain file types can be culled as non-responsive (e.g., .dll, .db, .jpeg, etc.)

The use of de-duplication software with any of the above search methods will reduce the burden and cost of further review.

4. *Redacting*

Use of automated redaction software is encouraged to expedite the processing of responsive documents containing FOIA exempt.

Consideration can be given to the use of automated software that filters regular "expressions" containing personally identifiable information (PII).[55]

5. *Production*

Determine whether production will be in native or imaged formats. If the latter is deemed appropriate, consider using some form of Optical Character Recognition (OCR) to render documents searchable.
References

E-discovery General Protocols
https://www.mdd.uscourts.gov/sites/mdd/files/ESIProtocol.pdf
https://www.nap.edu/read/22986/chapter/13
https://www.wawd.uscourts.gov/sites/wawd/files/61412ModeleDiscoveryProtocol.pdf

Keyword searching
https://www.edrm.net/resources/project-guides/edrm-search-guide/search-methodologies/
https://www.edrm.net/collections/edrm-search-guide-glossary/

Predictive Coding / Technology Assisted Review Guidelines
https://www.edrm.net/wp-content/uploads/2019/02/TAR-Guidelines-Final.pdf
The Sedona Conference Best Practices Commentary on the Use of Search and Information Retrieval Methods in E-Discovery (2014), www.thesedonaconference.org/publications

E-discovery Production Guide
https://www.edrm.net/resources/frameworks-and-standards/edrm-model/production/

---

[55] Examples of potentially exempt PII contained in regular expressions may include social security numbers, numbers on credit cards, passports, and driver's licenses, telephone numbers, email addresses, etc.