

Freedom of Information Act (FOIA) Officer Council (CFOC) Technology Committee Search Working Group – Key Challenges and Findings

Last Updated September 29, 2021

The FOIA Searches Working Group (Working Group) under the Chief FOIA Officers Council (CFOC) Technology Committee presents the following challenges and findings regarding search and technology considerations in FOIA case processing as noted in the group's charter, which is publicly available at <https://www.archives.gov/ogis/about-ogis/chief-foia-officers-council/technology-committee/wg/searches-charter>.

Given the complexity of some of these issues, the different technology used at agencies, the different governance and management of technology at agencies (*i.e.*, some FOIA offices have their own IT support, some rely on central agency IT support, etc.), this paper does not seek to resolve the issues presented but rather seeks to raise awareness of them for ongoing discussion and collaboration that is focused on resolving these and other challenges. It also offers a few practical tips to assist agencies in conducting searches.

In light of the challenges of remote work during the COVID-19 pandemic, this Working Group met remotely through the fiscal year to develop the findings in this paper. The Working Group will focus on how to leverage technology to address these key challenges that have been identified and continue to serve as a central contact point for FOIA professionals with questions or challenges involving FOIA search and technology.

Research and Analysis

The FOIA Searches Working Group has monthly meetings. For this paper, the Working Group reviewed publicly available Chief FOIA Officer reports, conducted a survey of approximately 30 federal agencies, met with representatives from the FOIA Advisory Committee (including the private sector) to discuss questions and concerns about FOIA search capabilities at federal agencies, and reviewed current agency practices at over 10 federal agencies based on survey responses and the experiences of the Working Group's members from five different federal agencies.

To encourage candid responses from agencies, the Working Group agreed to discuss findings and challenges without attribution to specific agencies. This approach resulted in very frank discussions and findings about agencies' abilities to conduct searches.

The questions asked in the survey are below.

- Does your agency have a centralized or decentralized approach to searching electronic archives (or a combination of both)?
- Which electronic search tools does your agency use?
- What are their limitations?
- What kind of IT features do your search tools use such as Boolean, AI, etc.?
- Share up to three examples of any FOIA Search issues that you or your agency has experienced by using technology to conduct searches.

- Share up to three examples of any FOIA Search solutions that you or your agency has developed by using technology to conduct searches.
- Do you use any electronic search questionnaires? If so, how is the data used? Is its use automated?
- Do you use tailored search instruction forms if records custodians are searching?

Provided below are key findings and challenges that were identified from this effort.

Key Findings

While there are several key findings from this effort, there are three major ones: (1) there is a significant gap in the public's understanding of agencies' abilities to search their electronic archives and databases; (2) the strategies and tools used by federal agencies to conduct searches vary greatly; and (3) searching email is still a challenge.

For the first point, it became apparent from discussions and meetings that many FOIA requesters think that agencies can perform a "Google-like" search of their IT systems and databases for FOIA requests in a centralized way. That is not the case at most federal agencies that were part of this review. There are sometimes dozens of IT systems that need to be searched to find records. Debunking this myth has been an important discussion topic for this group. There are some simple options available to agencies to address these misconceptions. For example, agencies can provide guidance about their federal records and databases on their public FOIA websites. This guidance can provide links to information about records systems and tips on how to request information in them. Agencies can also invest in new technology to search across electronic record and data sets. Engagement with requesters on these topics via calls or email can also assist.

Additionally, from the vendor community, there are often aggressive pushes for technical solutions to eDiscovery or other search challenges without taking the time to understand the additional IT and technical needs of an agency including the agency's IT network, infrastructure, governance, or management. The Working Group found that there are plenty of technical tools and solutions available to address the primary search needs of agencies; however, these tools may not be approved for use on government systems, may be expensive, or may not be compatible with existing technology for FOIA case processing. As a result, search "solutions" can resolve the problem of search but create new technology FOIA case processing problems.

For the second point, the Working Group reviewed at least three different approaches to searches at agencies including an outline of search considerations when considering how to conduct electronic (or other) searches. Forms exist within some agencies to assist in conducting searches but they were often simply outlining where to search such as email accounts, shared drives, paper files, etc. Some agencies shared that they have an approach or methodology that asks about the content sought by the requesters if a request is overly broad to work with the requester to clarify or negotiate the scope of requests; having such a methodology can be helpful in narrowing very broad requests or when a request is highly itemized with key words or terms to be used in conducting searches.

Search parameters vary from agency to agency based on the data, records, and available technology within the organization. For agencies with diverse and complex records, the Working Group found that having a plan on how to approach searches within the agency can help to navigate different IT systems, which are often managed by different offices with their own IT support. One challenge noted was how to electronically move certain data and record types from one system to another. Agencies may also be dependent on subject matter experts to conduct searches, identify responsive records, and evaluate search results. In other instances, some IT systems can only be searched using rigid methodologies based on the system itself. Data integrity and standards vary in the record systems including visualization techniques. These standards also vary regarding the FOIA case processing platform being used by agencies (and some agencies use more than one case processing platform). These findings led to discussions by the larger CFOC Technology Committee about creating a data standardization working group in 2022.

For the third point, email continues to be a challenge due to volume and the ability to conduct searches for and of emails. The increasing number of emails each year, not to mention other forms of electronic communication, presents an ongoing and cumulative volume issue. Some agencies have more advanced means of searching emails than others. For example, while some agencies can conduct centralized searches of emails, others are reliant on email account holders to conduct searches. This can cause delays in response times and inconsistent approaches to the overall search if the agency is not providing specific guidance on how each individual should conduct their search.

Some agencies also expressed concern about the inability to deduplicate email search results. For example, at one agency, after an email search is completed, a FOIA professional needs to review manually all of the emails found, even if there are hundreds or thousands of duplicates in the case. This lack of ability to efficiently narrow search results by eliminating duplicates translates into longer processing times. Additionally, email search results often produce a clutter of group messages, marketing, newsletters, and other content that have terms associated with a request but may not ultimately be responsive.

Applying consistent redactions on large volumes of emails was also cited as an ongoing issue and concern. Agencies have very different tools and abilities to review emails and apply redactions using FOIA exemptions. These latter two points are outside the scope of this Working Group.

Provided below are the Working Group's other key findings:

- At the 10 agencies that provided survey responses, most searches of individual archives and databases are manual.
- Leveraging consistent data standards, APIs, and interoperable IT solutions could assist agencies in both records management and conducting searches of agency records, including for FOIA. Funding for such efforts or commitment to these approaches, which can take years to develop and implement, are often limited or not available.

- Absent interoperable systems or consistent data standards in agency IT systems, it will be challenging to leverage AI for FOIA at many agencies in a holistic way.
- Some agencies are using Boolean logic to conduct searches for records and in their FOIA case processing system; not all agencies are using Boolean.¹
- At least one agency that was surveyed is using trained librarians to develop search strategies, conduct searches of central archives, and provide guidance to agency employees conducting searches of other IT tools.
- Requesters sometimes provide specific Boolean strings or certain terms with “and/or” search requirements. For agencies that use Boolean, this approach can work. For other agencies, the “and/or” requirement may not be useful for certain IT systems and could result in no records being identified even if there may be responsive material in that database or archive.
- If a requester is seeking a specific electronic record or report, the request can be processed faster if the underlying records or emails are not part of the request. In some cases, these requests for any email or records on a topic could delay processing times.

Challenges

Throughout the year, the Working Group identified the challenges below and wanted to raise awareness of these findings for the federal FOIA community. The Working Group will continue to work with agency-specific issues and seek to identify shared tools and approaches to address these challenges.

- There are issues at some agencies involving the scale and volume of records found using tools and the ability of those tools to process large datasets. Some IT tools used by agencies to conduct searches of records have data storage limits. After a search produces a certain size search result, the ability of the tool to function well is limited.
- Support for FOIA IT and technology needs vary by agencies.
- Available financial and human resources may also be a challenge in developing technology search solutions.
- Searches of text messages and social media tools continue to present challenges at agencies.

¹ A Boolean search is a search that uses the logical (i.e., Boolean) operators (AND, OR, NOT, -) in addition to the keywords.

AND: The AND operator tells the search engine to return only documents with all the keywords you entered. This operator narrows the search and returns fewer search results

OR: The OR operator tells the search engine to return documents if they contain one or more keywords.

NOT: The NOT operator tells the search engine to exclude documents from a search if they contain the keywords.

- Operator: The "-" operator is the same as the NOT operator and tells the search engine to exclude documents from a search if they contain the keywords.

Note: Boolean operators are not case sensitive. For example, entering and or AND will return the same results.

<https://www.govinfo.gov/help/search-operators>

Tips for Requesters

The Working Group recognized that agencies are working to meet requester needs. As with most things, communication is critical to achieving that goal. Below are some tips that were discussed over the past year that may be helpful to requesters based on discussions with federal agencies.

- If a requester wants a specific record the requester should limit the request to that specific record or records.
- Be specific in what is being sought. Less can be more in writing a request.
- Research agency record types. This could help requesters understand the kinds of records held by agencies and what may be able to be produced in response to a FOIA request.
- Try to avoid long lists of terms or overly broad topics; some agencies shared examples of getting lists of hundreds of terms in one request. A more focused list can help to reduce processing times.
- Search online FOIA Libraries, NARA websites, or online search platforms like Google before placing a FOIA request; there is a lot of information already released and available, which could either negate the need for a new FOIA request or help develop and scope a new one. A link to each agency's FOIA Library can be found on FOIA.gov.