

Breaking the Lock: Accessing Public Records to Map Systems, Algorithms and Data

CENTER FOR
CONSTITUTIONAL
RIGHTS

Northeastern University
School of Law
Center for Law, Information
and Creativity

AINOW



Public records requests have been a useful tool in identifying where and how government agencies are using **algorithmic tools**, especially in regards to surveillance technology and databases. Requesting these records can provide useful information to help challenge government use of these types of tools and technologies. However, it is not always apparent what to ask for, what information is subject to public records laws, and what information is useful for civil liberties analysis or for movement advocacy campaigns.

“Algorithmic tools” is a term that encompasses a broad range of technologies and systems. An algorithm is a series of steps designed for a system to perform tasks, or support or make decisions. Therefore, the term “algorithmic tool” can include spreadsheets, databases, machine learning technologies, and systems performing predictive or matching functions, such as face recognition or “predictive policing.” This resource has been created to support activists, lawyers or anyone interested in filing an open records request to determine what to look and ask for in order to assess and potentially challenge government use of algorithmic systems.

You should be aware that these types of public records requests can become resource intense. Some of the information you will request may be held by a private company that a government agency has contracted with, so the government agency may suggest that they do not have responsive records to your request because they are not technically retained by the agency. Or, the agency is unable to obtain such information because vendors claim that such information includes proprietary trade secrets. In these circumstances, you will need to be prepared to challenge the government's response through an [administrative appeal](#) or often [a lawsuit](#). You can also consider advocating for legislative reform of your state's open records law to address this problem.

Steps to Take Before Drafting Your Request



1. **Review the federal Freedom of Information Act (FOIA) or Your State's Open Records Law**– If you are filing a federal FOIA request, the Center for Constitutional Rights has many useful tips available as part of our FOIA Basics Guide for Activists available at <http://foiabasics.org>. While many state public records laws are very similar to the federal law, several vary drastically so it is important follow this guidance in context of your state law. Two helpful resources are the [Open Records Guide](#) provided by the Reporters' Committee for Freedom of the Press and also [Muckrock](#).
2. **Compile Sources or References to the System** – Try to tailor your request in accordance with known or speculative information about the system. Potential sources of references to systems may include recent news articles, public statements or press releases by public officials, agency or legislative budgets, agency or legislative public hearing notes or minutes, past releases of FOIA documents or relevant databases (e.g. [MuckRock Project Public Records Request Archive](#)). For example, after the murder of George Floyd in 2020, New York state issued Executive Order 103, which ordered all law enforcement agencies to review their individual policies and procedures and draft a reform plan for the future. Many of these plans were made

public and may provide important information for your records request. Compiling these reference sources in advance may make your request more specific and help if you have to appeal or challenge the response to your public records request. Try to review the most recent information as possible.

3. **Draft Public Records Request** – You can draft your request using the [annotated sample FOIA](#) or or [annotated sample](#) public records request as a template but make sure to use your own sources or reference documents to narrow and craft your request. In addition to being specific, you should try to include examples of what you are looking for so that it is as clear as possible to the public records officer making determinations about what is responsive to your request. You can provide cites or links to these examples in your request, such as providing a link to a news article, or you can attach them to your request as exhibits if they aren't too long.

Key Parts of Public Records Requests for Algorithmic Systems and Databases:

Most of the provisions in this model request are in reference to a software based algorithmic-system. Be mindful that some systems may function as a hybrid, where government officials perform certain tasks or functions of the algorithm. Therefore, you must modify the language to capture these distinctions so your request is not rejected (in part or whole) because of semantics or misinterpretations.

Below is an annotated list of key parts of a request for **algorithmic systems**, and a similar list focusing on **database systems**. You don't need to include every numbered section below, but some or all of the sections may be helpful for the information you are seeking. Remember to tailor your request to the specific system(s) you are asking for information about.

In addition to the general templates below, you should also review **two annotated sample requests** we have included in this resource:

- [Annotated Tech FOIA Request](#)
- [Annotated Tech State FOIA Request](#)

These examples use semi-fictional information but are based off real-world public records and FOIA requests [AI Now](#). Prof. Rashida Richardson, the Center for Constitutional Rights have filed in the past.

Most of the provisions in these model requests are in reference to a software based algorithmic-system. Be mindful that some systems may function as a hybrid, where government officials perform certain tasks or functions of the algorithm. Therefore, you must modify the language to capture these distinctions so your request is not rejected (in part or whole) because of semantics or misinterpretations.

Potential Requests for Algorithmic Systems

1. All records including information relating to the algorithm that [DESCRIBE WHAT IT DOES] within [NAME OF AGENCY, DIVISION OR FACILITY], including but not limited to its source code, models, developer documentation, and operator manuals.

This provision seeks the actual algorithm and other relevant technical information. It is important to specific to the best of your knowledge what the system does. Questions to consider when describing the system:

-Does it predict behavior or actions?

-Is it used to determine what resources a person will receive or how they may be treated?

-Is it classifying an individual or group of people?

2. All records relating to the training data used to develop, or train, the algorithm.
3. All records, including but not limited to documentation or internal communications, about the traits, characteristic, or factors used to develop the data fields in the System.
4. All records showing the full list of the data fields in the System.
These provisions seek information about the variables a system may use to produce an output (e.g. a prediction or determination).
5. All records of de-identified input data in the System.
6. All de-identified records of algorithm outputs, including but not limited to [ADD A SPECIFIC EXAMPLE OF SOMETHING YOU ARE SEEKING].
Use your source references to try to identify specific outputs of a system to specify here because if this provision is too broad the public records officer may reject the request because they do not know what may be response.
7. All records showing how [NAME OF AGENCY] staff use algorithm outputs to determine [INSERT ANY KNOWN PREDICTIONS OR DETERMINATIONS MADE BY THE SYSTEM].
This provision seeks information to help you determine what level of deference agency staff may give to the system's outputs. Typically, the lack of any information, directives, and/or guidances suggests a fairly subjective process.
8. All records of, including communications regarding, audits, internal reviews, or validation studies of the System.
This provisions seeks information on whether the agency or vendor has audited or tested the system for bias and errors.
9. Any internal policies, practices, procedures, memoranda and training materials for using the System, and for storing, accessing, and sharing data inputs and analysis created by the System.
This provision seeks information on how the system is used, as well as how information may be shared. Ideally, this information will help identify security features of the system and agencies procedures to prevent or address misuse, abuse, etc.
10. Any internal policies, practices, procedures, memoranda and training materials for sharing data inputs and outputs created by the System with entities outside of [NAME OF AGENCY], including the [SPECIFY AN OUTSIDE ENTITY OF CONCERN].
This provision seeks information about policies or procedures for external sharing. You should specify external entities of concern like law enforcement or federal agencies.
11. Any records showing which entities outside of [NAME OF AGENCY] have accessed, used or requested to use, the System.
This provision seeks information about whether information has ACTUALLY been shared externally, regardless of whether there are policies or procedures and whether they have been followed.
12. Any records reflecting any agreements for or permission to develop, use, test, or evaluate an algorithmic system used to [SPECIFY FUNCTIONS AND/OR OUTPUTS] and services with any third-party vendor or consultants, including the [SPECIFY THIRD-PARTY THAT MAY HAVE COLLABORATED OR REFERENCED IN SOURCE DOCUMENTS].
This provision is seeking information about contracts with consultants or other third-parties.
13. Any Records referencing the public process preceding the procurement or acquisition of the System, including public meeting agendas or minutes, public notice, analyses, or communications between [NAME OF AGENCY] and elected officials or other public servants.
This provision seeks to identify if there was any public communications about the system. This information can be useful

for any advocacy regarding public procurement loopholes or problems.

Potential Requests for Database Systems

1. All records including information relating to the database that [DESCRIBE WHAT IT DOES] within [NAME OF AGENCY, DIVISION OR FACILITY], including but not limited to the screenshots of web-portals or interfaces, applications, any additional tools of analysis (including the source code and models for such analysis), developer documentation, and operator manuals.
This provision seeks the actual algorithm and other relevant technical information. It is important to specific to the best of your knowledge what the system does. Questions to consider when describing the system:
 - Does it predict behavior or actions?
 - Is it used to determine what resources a person will receive or how they may be treated?
 - Is it classifying an individual or group of people?
2. All records maintained in the database.
It may or may not be strategic to request ALL records maintained in the database. If it is a large database, an agency may respond and ask you to narrow your request, or there may be costs or other challenges. However, if you are looking to do a detailed audit of a database it can be a useful request.
You can also modify this request to certain "all" of certain types of records maintained by the database.
3. All records, including but not limited to documentation or internal communications, about the factors and data sources used to develop and determine the data fields in the Database.
4. All records showing the full list of the data fields in the Database.
These provisions seek information about the fields or variables a database may use to produce an output (for example, a prediction or determination).
5. All records of policies regarding database use, management, maintenance, and compliance with relevant laws. This can include internal guidelines or memorandum regarding department audit or notification procedures or records of database maintenance or system updates.
6. All records showing how [NAME OF AGENCY] staff access and use the database. This can include user manuals, training documents, internal directives, and procedures[INSERT ANY KNOWN PREDICTIONS OR DETERMINATIONS MADE BY THE SYSTEM].
This provision can include requests for purging, notification, and auditing requirements, if any.
This provision seeks information to help you determine what level of deference agency staff may give to the system's outputs. Typically, the lack of any information, directives, and/or guidances suggests a fairly subjective process.
7. All records of, including communications regarding, audits, system test, internal reviews, or intragovernmental investigations or inquiries of the database.
This provision seeks information on whether the agency or vendor has audited or tested the database for bias and errors
8. Any internal policies, practices, procedures, memoranda and training materials for using the Database, and for storing, accessing, and sharing data analysis created by the Database.
9. Any internal policies, practices, procedures, memoranda and training materials for sharing data inputs and outputs created by the Database with entities outside of [NAME OF AGENCY], including the [SPECIFY AN OUTSIDE ENTITY OF CONCERN].
This provision seeks information about policies or procedures for external sharing. You should specify external entities of concern like law enforcement or federal agencies.

10. Any records showing which entities outside of [NAME OF AGENCY] have accessed, used or requested to use, the Database.
This provision seeks information about whether information has ACTUALLY been shared externally, regardless of whether there are policies or procedures and whether they have been followed.
11. Any records reflecting any agreements for or permission to develop, use, test, or evaluate a database used to [SPECIFY FUNCTIONS AND/OR OUTPUTS] and services with any third-party vendor or consultants, including the [SPECIFY THIRD-PARTY THAT MAY HAVE COLLABORATED OR REFERENCED IN SOURCE DOCUMENTS].
This provision is seeking information about contracts with consultants or other third-parties.
12. Any Records referencing the public process preceding the procurement, acquisition, or development of the Database, including public meeting agendas or minutes, public notice, analyses (including fiscal or costs analysis), or communications between [NAME OF AGENCY] and elected officials or other public servants.
This provision seeks to identify if there was any public communications about the system. This information can be useful for any advocacy regarding public procurement loopholes or problems.

Suggested Terms and Definitions

It is often helpful to define certain terms you might be using in your public records requests, especially in larger FOIA requests. We have provided some suggested definitions of different terms relating to databases and algorithmic systems below.



- **Automated Decision System/ Automated Decision-making** – In this request the term “Automated Decision System/ Automated Decision-making” refers to any system, software, or process that uses computation to aid or replace government, decision, judgments, and/or policy implementation that impact opportunities, access, liberties, rights, and/or safety. This can include the use of spreadsheet software like, Microsoft Excel.
- **Database** – In this request the term “databases” refers to information systems, either digital and manually maintained, that are used to compile, analyze and share information regarding individuals and/or groups to aid or inform criminal investigation.
- **Access** – In this request the term “access” refers to the ability to do one or more of the following: view, query, add, delete, alter or retrieve records in the database or system.
- **Audit** – In this request the term “audit” means collecting data about the behavior of an algorithm in the relevant use context, and using data to assess the algorithm’s behavior, patterns, and anomalies. Audits are performed to assess the risks and errors an algorithm may present in the relevant use context, including whether the algorithm’s behavior is negatively or adversely affecting the rights or interest of people affected by the algorithm.
- **Source Code** – In this request the term “source code” means the instructions written by a human to create a computer program, software, or application.
- **Validation Study or Analysis** – In this request the term “validation study or analysis” refers to studies regarding the accuracy, validity, reliability or other factors related to a database or systems performance.

Last updated in March 2022.

This resource was authored by

Prof. Rashida Richardson of Northeastern University

Amba Kak, Director of Global Policy at AI Now

Ian Head, Open Records Coordinator at the Center for Constitutional Rights