

Court data sharing from scratch.



**small
scale.**



Introduction

As technology becomes more capable, courts are faced with ever more requests for data access. Large, well-resourced courts are developing sophisticated infrastructure to process and fulfill complex data requests.

This guide is not for them.

Court data sharing from scratch is designed to help small and medium-sized courts kickstart a process for responsibly handling data requests. It is meant as a complement to existing, more detailed references on building model court data policies.

Our goal is to start a conversation, rather than end it. We know every court is different, we hope that courts can use and adapt the ideas in this guide to build an approach to sharing data that works best for them. Without good data policies, we lose valuable opportunities to improve our nation's courts, and the justice system.

This guide offers a simple breakdown of the data request process, and introduces questions courts should ask at each stage. Each section includes a worksheet to help you think about how to apply the questions to real-world data requests.

At the end of the guide, you'll find worksheets and examples, along with additional NCSC resources on responsible data sharing.

Contents

Getting Started

- What counts as court data?
- What does state law require?

The Ask

- ① What data is being requested?
- ② What is the purpose of the request?
- ③ What other data will be used to fulfill the purpose?

The Review

- ④ What do we need to fulfill the request?
- ⑤ How risky is the request?
- ⑥ How could the court benefit?

The Commitment

- ⑦ Who will own the relationship?
- ⑧ What ethical commitments do we need?
- ⑨ What should be shared with the court?

Examples

Resources

Getting started

First things first: before building any data request policy or process, we need to take stock. You'll need to take inventory of the data your court might have available and any state or local laws or court rules that might regulate how you handle data requests.

What could be court data?

Court data can be a broad term. Here's our definition: court data is information about a court and its activities.

That means beyond case records, your court has other sources of data that could contribute to learning. For example:

Archives

Paper and microfilm records can be data! The right partner may even be able to help you digitize these records.

Receipts

Information about payments people make (and don't make) to the court could help illuminate the relative burden of court appearances.

Process service

You or an outside partner may be able to gather data on process service and use it to lower defaults.

Website analytics

Your website's analytics tool might have data about who visits the court website, what pages they visited, and what searches led them there.

Remote hearings

Your remote hearing software may retain data that can help reveal the extent and quality of participation in remote hearings.

Special projects

Almost every initiative your court pursues produces data. What other learning opportunities are hiding in your court?

What does state law require?

Chances are, state record laws will substantially influence how you can respond to data requests.

Laws, rules, and guidance vary widely from state to state or locality to locality about which data elements are subject to disclosure requirements, what exceptions can be made, and how quickly requests should be fulfilled. Check your local rules before building a data request process.

Even within those rules, you may have freedom to design reasonable processes for handling requests, withhold sensitive information, or even prioritize some requests over others.

If you need more help building compliant data request processes, get in touch with NCSC.

Jurisdictions where judiciary is exempt from state public records laws (as of 12/31/23):

Arizona, California, DC, Illinois, Kentucky, Maine, Massachusetts, Michigan, Minnesota, New Hampshire, New Jersey, New York, Ohio, South Dakota, Texas, Vermont, Virginia, Washington

Jurisdictions where judiciary is subject to state public records laws (as of 12/31/23):

Alabama, Alaska, Arkansas, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Idaho, Indiana, Iowa, Kansas, Louisiana, Maryland, Mississippi, Missouri, Montana, Nebraska, Nevada, New Mexico, North Carolina, North Dakota, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, Tennessee, Utah, West Virginia, Wisconsin, Wyoming

Source: [MuckRock](#)

The Ask

Inevitably, someone is going to ask your court for data. If you're starting from scratch, a request might come via email, a formal records request, or even a phone call or in-person visit.

Let's bring order to this potential chaos. Three simple questions can help you organize and manage incoming data requests.


While we're focusing on requests coming from outside the court, these questions are also a great way for internal stakeholders to formulate and answer data-driven questions.

1 What data is being requested?

We can start simple. A good data request should cover the data or information being requested, with **as much detail as possible**. Ideally, that would include not only specific records (e.g., eviction cases from 2000-2005), but specific **fields and characteristics** (e.g., addresses of eviction cases that ended in defendant default, 2000-2005).

Catalog your data

A great way to improve the quality of incoming data requests is to build a **data catalog** or data dictionary. A data catalog is a detailed description of the data you hold: including data fields, quality, and sensitivity. A public data catalog can save you from having to process data requests that you're unable to provide data for, or guess at what a request might be asking for.

 [New York State data dictionary for low-level criminal cases](#)

2 What is the purpose of the request?

In some jurisdictions, you may need to ask for a request's purpose in order to determine the requester's fee burden. But identifying the purpose of a data request can have other benefits too: it may be that you have data that is better suited to the requester's purpose, or the data requested may have quality issues that could interfere with the requester's work. Or, a requester may be pursuing a purpose that is of interest to the court, which could lead to a beneficial collaboration.

3

What other data will be used to fulfill the purpose?

As you know, the data you hold may not always tell the full story. Cases are settled out of court, some records aren't retained, mistakes are made. Asking requesters for more detail about how they hope to use court data may help calm fears that data is being taken out of context or misinterpreted.

In addition, this question can help you evaluate risks of disclosing sensitive data, and whether confidential information may be reidentifiable using linked datasets.

What is data linking?

Linking refers to the practice of combining multiple related datasets in order to gain additional insights. Linking datasets can also make it easier to reidentify individuals in a dataset, or uncover other sensitive information that was previously obscured.

For example, imagine sharing an eviction case with the party names redacted. If the address is included in the shared data, someone can use other data to reidentify both parties.

Historical property records are publicly searchable in most jurisdictions. A search of the case address could help reidentify the landlord.

Many (but not all) states also allow for the purchase of **voter files**, which could include names and addresses. If the tenant is (or was) registered to vote, searching the voter file with the case address could be used to reidentify a potential tenant.

Worksheet: Data Request

Data

Describe the data you are requesting. Be as specific as possible.

If you are requesting records that are described in our Data Catalog, your request must refer to specific records and fields as described in the Catalog.

If you are requesting records that are not described in our Data Catalog, please provide as much detail as possible for your request.

1

Purpose

Describe the purpose of your request. What will you use the data for?

Your answer here will help us determine whether the data requested will be responsive to your needs. We may use this answer to suggest alternative or additional data that may be helpful.

2

Linking

Describe or list any other data you are using to achieve the purpose.

Your answer here will help us identify risks to data subjects from linking court data to outside datasets.

3

The Review

Even if local laws afford you very little discretion to refuse data requests, you can still benefit from a review process. A review process can help you better understand the demand for data and the resources it requires from your court. It can help you identify risks that may inform policy changes to protect people from harm. And it creates an opportunity to identify collaborations that may add value to the court.

Your review process need not be hermetically sealed from the outside world. Engaging with the requester and the community around the court can help you build a data ecosystem that benefits everyone.

④ What do we need to fulfill the request?

Often, the most important factor for prioritizing requests is how long the request will take to fulfill. Every request is different, but you can prepare by identifying **internal stakeholders** who would be responsible for fulfilling the request: whether clerks or IT staff.

Beyond a simple calculation of how many hours of staff time a request might take, try to learn what factors could help reduce the complexity or cost of a request. Remember, data usability issues don't just affect outside parties; they affect your ability to understand how well the court is operating.

Can a requester help?

In some cases, your requester might be able to commit resources to overcome data quality or accessibility challenges. In exchange, they may be able to provide the court with usable, if unofficial datasets that can be reused or reshared.

 The Law Library Microform Consortium (LLMC) helps digitize thousands of historical state records stored on microfilm.

5 How risky is the request?

People who are included in data requests may face ongoing risks from use or publication of data. Once data is disclosed, you may lose the ability to control how it is used or reshared.

Assessing the risk level of data can help inform the steps you take to mitigate that risk. In many cases, this might require you to redact or deidentify information or fields within a record. Where deidentification isn't possible, you may need to aggregate data in order to minimize risk of harm.

Risk levels

Many states have started to build frameworks to help agencies identify the sensitivity of data ([WA](#), [NC](#)). Adopting your state's framework—or our example one—can help you think through how to mitigate disclosure risks when sharing data.

High-risk data is data that, if disclosed, would put a person at risk of substantial harm. For example, someone's personal information—a name paired with an address, a driver's license, a social security number—could be considered high-risk data.

Medium-risk data is data that is considered sensitive, for official use, or that might otherwise be protected from disclosure by law. Definitions of this middle category vary widely from jurisdiction to jurisdiction—it's up to you to decide what fits here.

In general, data that isn't high-risk or medium-risk is **public data**. While your local rules may not require you to disclose this information, improving the availability of public data can save you time processing low-risk data requests, and ultimately improves the quality of information about the court.

6

How could the court benefit?

One tool for managing the flow of data requests is prioritization. If there are data requests that align with areas that are important to the court, you could prioritize reviewing and fulfilling those requests. This doesn't mean you have to relax your review standards—just that you can deploy your limited resources in a way that maximizes benefit to your court.

In other cases, a data requester's project may be of interest to the broader community. Taking time in the review process to identify opportunities for collaboration can help maximize the benefit of sharing data.

Community engagement

Engaging with community stakeholders can help facilitate safe and effective use of data.

Research on governance suggests that the people should have a role in governing initiatives that are likely to affect them. Here, sharing court data can impact more than just the court. It can have real impact on the community surrounding the court—people may benefit from improved policies or court processes, or be harmed via discrimination or loss of privacy.

How can you put this into practice? One way is to assemble a **data advisory council**—a mix of internal and external stakeholders that can help the court build a data sharing strategy, or comment on specific requests. Or, you could put together a **data jury**: an ad hoc group to provide a more nuanced perspective on a given data request or initiative.

Worksheet: Request Review

Effort

What effort is needed to fulfill the request?

To estimate this, try breaking down the request into discrete tasks: what will court staff have to do in order to fulfill the request, and how long will it take them?

An effort calculation can also make the case for investing in data infrastructure, which can bring the effort required to fulfill certain requests down to zero.

1

Task	Who performs the task?	Effort (hrs)

Risk

Does the request include high-risk data? Would disclosing the data threaten a person's safety or well-being?

Identify potential risks related to this request, along with possible mitigations.

A mitigation could include rescoping the data request, or securing commitments from the requester to adequately protect data.

2

Data	Risk level and explanation	Mitigation

Value

How could the court benefit from this request?

Are there additional or related questions we could ask the requester to pursue, as part of a collaboration?

3

Community

What stakeholders around the court might have an interest in this request?

How might this data request affect them?

4

Stakeholder	How could this request affect the stakeholder?

The Commitment

Data requests are sometimes memorialized with a short agreement or memorandum of understanding (MoU). In addition to including boilerplate disclaimers about data quality and accuracy, you can use MoUs to set expectations about ethical data use or encourage requesters to share findings back with the court.

7 Who will own the relationship?

Be sure to designate a court point of contact for each request. This is especially important if you ask requesters to make commitments about how they'll use data, or what they'll share back to the court.

8 What ethical commitments do we need?

In some cases, you may want to prevent requesters from attempting to reidentify people in the dataset you provide, or from linking your dataset with other data that could enable reidentification. Alternately, if data is being collected to prepare an in-person study, you may want to ensure that the study is supervised by an Institutional Review Board (IRB).

9 What should be shared with the court?

Data sharing need not be a one-way process, with outside parties extracting data from the court. Many data requests are in service of research questions that may be of interest or value to the court.

Implementing “share back” requirements or incentives can help ensure that the court maximizes learning from its own data, and can build stronger relationships with researchers and data users. Over time, smart resharing can help expedite future requests by providing you with cleaned and refined datasets to share with others.

What could be shared back?

Consider asking for cleaned datasets, secondary datasets or data derivatives, prepublication drafts of research findings, or final research findings. In some cases, you may need to embargo shared documents until findings have been published.

Worksheet: Commitments

Resharing

What limitations are there on resharing or selling provided data?

May data be reshared or sold?

- Yes No With restrictions

1

Details on sharing restrictions:

Contact

What limitations are there on contacting people in the provided dataset?

May persons in the dataset be contacted for research purposes?

- Yes No

If yes, include information on supervising IRB:

2

May persons in the dataset be contacted for other purposes, such as marketing?

- Yes No

Reidentification

May the requester attempt to reidentify persons in the dataset?

3

- Yes, reidentification permitted.
 No, reidentification is not permitted.

Sharing

What should the requester share back with the court? Check all that apply.

4

- Cleaned data sets
 Secondary or derived datasets
 Preprint or draft report
 Final results or report
 Other

Examples

Try using the worksheets to think about how your court might respond to the following example data sharing projects.


Healthcare for justice-involved individuals


Pretrial services wants to share information about justice-involved individuals with the state Medicaid office, in order to identify whether those individuals might qualify for coverage under Medicaid.

 [Pretrial Services Division expands MassHealth Enrollment](#)

Eviction dashboard

A local nonprofit wants to collect daily data on eviction cases in order to power a dashboard that tracks aggregate evictions over time.

 [Harris County \(TX\) Eviction Dashboard](#)

 [Indiana Eviction Dashboard](#)

Text message evaluation

A university researcher wants to collect data from the court's text message reminder system and associated cases in order to study the reminder system's effectiveness

 [Court Date Notification Best Practices](#)

Resources

Ready for the next step? Check out these resources.

Developing CCJ/COSCA Guidelines for Public Access to Court Records: A National Project to Assist State Courts (NCSC)

Civil Court Modernization Toolkit (Pew Charitable Trusts)

Handbook on Using Administrative Data (MIT)

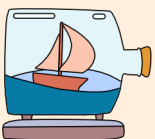
Best Practices for Court Privacy Policy Formulation (NCSC/SJI)

Court Data: Open, With Care (COSCA)

National Open Court Data Standards (NCSC)

Data Governance Policy Guide (NCSC)

Justice Counts (Council Of State Governments Justice Center)



**small
scale.**

