



Quattrone Center for the
Fair Administration of Justice
UNIVERSITY of PENNSYLVANIA

Data Collection and Analysis:

A Resource Guide for Conviction Integrity Units and Innocence Organizations¹

¹ This project was supported by Grant No. 15PBJA-21-GK-03607-WRNG awarded by the Bureau of Justice Assistance. The Bureau of Justice Assistance is a component of the U.S. Department of Justice's Office of Justice Programs, which also includes the Bureau of Justice Statistics, the National Institute of Justice, the Office of Juvenile Justice and Delinquency Prevention, the Office for Victims of Crime, and the SMART Office. Points of view or opinions in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.

CONTENTS

INTRODUCTION	2
WHY COLLECT DATA AT ALL?	2
HOW DO WE BEGIN?	4
INITIAL DATA CONCERNS	4
MOVING TO THE NEXT LEVEL	10
APPENDIX	13



INTRODUCTION

Innocence organizations (IOs) and conviction integrity units (CIUs)² do critical work in ensuring accuracy and fairness in the criminal legal system. By reviewing past convictions to determine whether they should be adjusted or vacated, IOs and CIUs deeply investigate cases and perform new investigations to discover the truth. These reviews and investigations generate rich, voluminous data. However, few organizations are keeping track of that data to learn from past errors sufficiently to prevent future injustices. Whether a result of not enough time to put a data tracking system in place, lack of technical capacity, funding, and staff to track data effectively, or even fear of data, every program will benefit from collecting, analyzing, and reporting data.

Over two virtual meetings, we met with dozens of lawyers, investigators, paralegals, and data specialists working for innocence organizations or conviction integrity units to discuss their data keeping practices in hopes of developing best practices for data keeping and analysis. Through these meetings, we discussed the types of data groups currently keep, how they keep it, and barriers they face to keeping more robust data and using it effectively.

What has emerged, with the help of our co-facilitators, is a resource for new and established organizations to use when designing their data collection processes. While we cannot hope to answer every question in every circumstance, we aim to provide a baseline level of guidance that can benefit every office and organization.

WHY COLLECT DATA AT ALL?

1. Informed Decision-Making

Data provides a factual and evidence-based foundation upon which to build internal procedures and inform public policy. Keeping good data helps organizations make decisions about which cases should be pursued, how the organization can improve its own selection criteria, and answer questions about cases they have handled in the past.

2. Identifying Problematic Trends

Data collection can uncover disparities, biases, misconduct, systemic errors, bad actors, and bad policies and practices within an organization or jurisdiction. By analyzing data related to convictions, trials, actors, and evidence used to reach a conviction, organizations can identify instances of racial, ethnic, or gender disparities. It is also helpful to identify internal unconscious biases that impact decision-making and case selection.

3. Preventing Future Error

Once problematic patterns and trends have been identified, they can be remedied, and measures can be taken to prevent them from happening again. For example, if an exoneration identifies a bad actor, keeping accurate track of case participants can make it easy to find other cases where that individual played a role. Or, for those with access to public docketing information or a prosecutor's database, if a common forensic method from several years ago has been discredited, running a search of the database ("bitemark," "tooth," "odontologist") can be used both to review old cases, and quickly flag new cases using the same or similar methods. As we go into below, technological innovations are making it easier every day to conduct searches like this and identify large cohorts of cases quickly.

² We use the generic term "conviction integrity unit" throughout this resource to refer to small units within prosecutors' offices that investigate potential wrongful conviction or actual innocence cases. Not all units use that term: many are called "conviction review units," "justice units," or other specified names.



4. Mitigating Bias

Data collection and analysis can help mitigate biases by identifying instances where biased practices might be occurring. When data shows that certain groups are disproportionately affected by bias, measures can be taken to rectify these biases and ensure that the criminal justice system treats all individuals fairly.³

5. Evaluation of Policies

Data allows organizations to assess the effectiveness of policies and decisions. By tracking outcomes over time, staff can determine whether specific policies are achieving their intended goals and whether those policies need to be adjusted.

6. Evidence-Based Reforms

Data-driven insights guide policies and decisions that are grounded in evidence rather than anecdotal observations or subjective biases. Even deciding which cases to pursue can be driven more by unconscious bias than data, resulting in cases which should be advanced being closed too early.

7. Empirical Research Evidence

In addition to the many important uses of descriptive administrative data, which tracks information on a case-by-case basis, collecting consistent and reliable data also provides opportunities for academic researchers to causally evaluate whether policies or practices are in fact producing desired results. That kind of evidence, once peer-reviewed and published, can be used to improve the entire field, support the work of kindred offices and organizations, and seek funding and resources based upon accepted Science.⁴

8. Accountability and Transparency

Especially for prosecutors' offices, transparent and accessible data can hold institutions and individuals accountable for their actions. Collecting data and making it publicly available fosters transparency that can shed light on misconduct, corruption, or abuse of power within the criminal justice system.

9. Tracking Your Own Organizational Progress

When your unit or organization makes a change, tracking your data can give you a sense of how well that new policy is working and meeting your objective. One participant in our discussion noted that they were able to track how changes to their intake procedure—revamping an outdated questionnaire and having it translated into different languages—increased the number of applications they were receiving, giving them important insight into ensuring the most people could apply for their review.

10. Strengthening Public Trust

When the criminal legal system as a whole collects, analyzes, and acts upon data to drive reforms, it signals a commitment to fairness, transparency, and accountability. This can help rebuild public trust in the system, which is crucial for maintaining social order and cooperation.

In summary, collecting data is essential for CIUs and IOs because it provides the factual basis needed to understand, address, and rectify systemic issues within the system. Data-driven policies are more likely to be effective, fair, and capable of improving public safety.

³ Identifying racial bias, particularly unconscious bias, is, by definition, a challenging task. Data analysis can be a great place to start by identifying situations where similarly situated defendants face different treatment, while the major difference in their case is their race. However, as with all forms of analysis, data needs to be accurate to produce meaningful results. It is not uncommon for offices which rely primarily on police reports to have inaccurate race and particularly ethnicity data identifying Hispanic and Latino people. For more, see: Effect of racial misclassification in police data on estimates of racial disparities: <https://onlinelibrary.wiley.com/doi/abs/10.1111/1745-9125.12329>

⁴ Many IOs have concerns about client confidentiality or privilege issues. Research institutions (like the Quattrone Center) deal with sensitive data regularly and generally have received approval through their Institutional Review Board before they are able to proceed. Data shared with research entities can be anonymized, so no identifying data or information is shared.



HOW DO WE BEGIN?

New CIUs and IOs look at data collection with a combination of fear and obligation: the lawyers who work in these organizations understand the critical importance of collecting data on the cases they investigate, but frequently lack sufficient resources to do the job well. But here's the good news: even starting with a few data points initially can improve an organization's functioning dramatically. Funders and supporters often want to know information that can easily be tracked, and which will improve their commitment to an organization or unit. This resource will provide an overview of the few data points that should be collected for offices that lack staff capacity. It also offers information about a comprehensive suite of data and technological resources designed to help offices quickly move toward being able to collect, analyze, and report on a larger list of data points.⁵

Most IOs and CIUs have a tiny staff doing a huge job. For many, collecting data can seem like a pipe dream. But even tracking a minimal amount of data will have an impact on your work and efficiency.

And while a single-person office can successfully collect and track data, there may be resources to help. Having technical, academic, or analytical expert guidance and support from the outset can help avoid mistakes that will become much more difficult to fix later. If you don't have someone on your staff or a resource in your office that can help with data collection, there are many options available. Public or private funders, academic institutions, and/or technical service providers may consult with you to help build the office's data and tech capacity. Having an expert alongside as you work through each of the steps below will help you make more strategic decisions and troubleshoot problems preemptively. As you become more sophisticated in your data collection practices, looking for and catching broader risks in your caseloads will almost certainly require specialized staff with more technical expertise, and it may not be realistic to pursue this without other partners or resources.

INITIAL DATA CONCERNS

STEP ONE – Assess Current Available Data

Take account of the data you're currently collecting, or that you plan to start collecting, and why. For CIUs in particular—determine what types of data your office tracks, who tracks it, and how it's tracked. When assessing data your office already collects, or when planning to collect new data, be sure to determine how you will assess whether your data are accurate and consistent. For example, your case management system may already include a check box for data entry to indicate if a case involves domestic violence. However, all staff may not be aware of this or trained to use it, so the resulting data may severely undercount the number of relevant cases.

Always remember: garbage in, garbage out. Make a plan to review the data you're collecting for accuracy at periodic intervals before you absolutely need it for data-analysis-related deliverables, particularly for crucially important data points.

In addition to making you aware of what is available, reviewing data your organization already tracks can help you avoid redundant data collection. Check with other partner organizations, courts, or other agencies that could share data with you electronically to save staff time conducting manual data entry. If your resources for data collection are limited, using your resources to track or re-enter data already collected in another system could be a large waste of time if more efficient options are available. If you receive data from other agencies, make sure to consult their data dictionaries, if any, to confirm the data definitions they use match what you're using or are planning to use.⁶

⁵ There is a link to a pre-built spreadsheet in the Appendix you can modify for your own use.

⁶ It is also best practice to consider making a “data dictionary” of your own setting out the names, descriptions, locations, sources, and other relevant properties of the data that you are collecting. For a helpful resource on making a data dictionary, see <https://help.osf.io/article/217-how-to-make-a-data-dictionary>.



Special considerations for CIUs

- » Before you start, make sure you speak with your Elected to understand their priorities for reporting to constituents or keeping data within the office. Determine if they would like for you to share any of the non-privileged data you keep with local Innocence Orgs.
 - Make sure you have a sense of what people throughout your office will embrace. It could be “nothing” because the office has never done this and is hostile to new suggestions; if so, think about outside groups you could bring in to facilitate a change in office culture around data and analysis.
- » Consult your IT team to determine whether you are permitted to access your office’s data electronically and confidentially.⁷ You may be able to auto-populate many or most of the data fields that you and your team would otherwise have to collect and input manually.⁸
- » Determine whether your office budget will allow you to hire or share staff or even purchase new technology or other resources. New technology can allow you to effectively collect and use data, but it may require outside funding and technical assistance.

For Innocence Organizations

- » Find out which types of data are already publicly available, through the courts or your local prosecutor’s office.⁹
- » Know which types of data will be required by funders and organizational supporters. Most will want to know about the numbers of cases you screen, move to investigation, investigate, and litigate. Some may have jurisdictional preferences.¹⁰
- » Set up your data to be able to provide easy answers and consider whether you want to make any of your (non-privileged) data public to help you visualize your data quickly.

STEP TWO – Determine How You Will Track Data

If a CIU, will you use a separately secured portion of the office’s CRM?

There are several basic database programs available that are free, low cost, or potentially already available to you. For example, offices with a Microsoft 365 account can use the spreadsheet program Excel or their database [Microsoft Lists](#). If your office uses Google, you have access to Sheets¹¹ and a new database program called [Google Tables](#).

If you don’t have access to those programs or want to use a commercial database program there are many options used by CIUs and IOs: [Airtable](#), [Clio](#), [FileVine](#), or [MyCase](#) are used by many.¹² Make sure the program you choose can be used by multiple people without specific training to broaden data tracking. Confirm before you begin that the solution(s) you get can provide the data analysis and reporting capabilities you want (e.g., if you eventually want to create dashboards or export data, confirm that whatever tool you are considering can do it). Before making a final decision to go with a program, software, or vendor, confirm it can satisfy any laws, rules, and security requirements for sensitive personal data in your jurisdiction.

7 CIUs, in particular, may be able to access data from outside agencies. Receiving data from the Department of Corrections, for example, can give you a head start on populating data on convicted individuals’ cases, sentences, and even demographics.

8 Many case-specific data points can be electronically auto-populated into your case management system or spreadsheet from an existing database using an API (application programming interface). Determining which of these data fields can be auto-populated and setting up the API permissions to make it happen seamlessly can be done relatively easily with the right technical support and can save your staff countless hours of manual data entry.

9 For example, a team in California is working on the creation of a federated database with members of the Innocence Network.

10 You may want to consider a public dashboard available on your website; you can save a lot of time not having to respond to queries if funders and others are able to access via a dashboard at least: 1. number of applications received by year; 2. number of cases investigated by year; 3. number of exonerations by year; 4. number of years of total wrongful incarceration overturned; 5. number of bad actors identified and/or added to Brady lists, and number of cases impacted.

11 A link to a sample spreadsheet with data to be collected is included in the Appendix.

12 A few IOs have used Salesforce – a very large platform for the sales industry. While Salesforce does offer free licenses, it is not for the non-expert; hiring an expert to put together a database for your office can cost around \$10,000.



Many organizations successfully use two (or even three) programs to keep track of and analyze data: one for case management (when a case came in, is advanced to a stage, gets closed, to whom it is assigned, etc.), one for case data (document gathering, notes for the case, memos), and even a third (for aggregate data reports and analysis).¹³ Remember, any combination of tools can be effective to gather data, as long as the approach is secure, organized, documented, and well understood by staff and partners.

» **For CIUs**

Because CIU work can be controversial within an office, it is a good practice to make sure your data is accessible only to people within your unit.¹⁴ At the same time, a CIU should be able to access data from the rest of the office to conduct its work. Being able to download case information saves the CIU time and ensures the data is consistent. In addition, if a CIU wants to evaluate cases involving particular geographic areas or individual actors, having access to the office database is critical for identifying potential cases.

» **For Innocence Organizations**

If your organization is affiliated with a university, there may be programs available for free or reduced cost. Companies like Clio offer free licenses to law schools through the Academic Access Program. Computer science departments within universities may be willing to offer course credit for students to work with you. The Innocence Network also has created resources on case management systems that are available to Network members.

Consider an academic partnership. Many universities or academic institutions will partner with government or non-profit agencies to help develop and understand data and reform the criminal legal system. Sometimes professors or centers will ask for access to data for their own research purposes as part of a partnership. These research projects can yield tremendous information and assistance—like help with collecting data you need more efficiently—in addition to reviewing larger trends.

When designing an approach to collecting data or choosing a database or any other tool, offices need to consider who will be inputting the data. It's important to remember there is no "perfect" combination of tools here: using a combination of tools available based on your resources and staff availability will help guide a strategy better than trying to force a particular system to work. Some offices use a checklist or other document for screeners but ensure one or two dedicated people enter the information into the spreadsheet or data system.

Larger offices face what may feel like a prohibitively large volume of cases to review to collect data. Advanced technology and a well-run administrative database can help with tracking and sifting through these cases, reducing the amount of time spent reviewing files.

Smaller offices, on the other hand, may consider that the staff time required to make use of a database for data entry may be unrealistic given their limited capacity. In these situations, making use of simpler tools may go further and lead to higher quality data. Those tools could include tracking cases that have been reviewed in a spreadsheet or considering academic partnerships with experts in more efficient ways of extracting data from text.¹⁵

13 The capabilities of many programs vary, but so do security requirements within your offices and organizations. As much as possible, check with your IT Department staff and General Counsel to ensure your choices are complying with all local, state, and federal laws and regulations governing data storage and use.

14 Getting secured data access isn't new. Homicide units, sexual assault units, and other specialized units frequently control who in an office can view sensitive data with carefully curated permissions. CIUs can and should determine if their office's CRM has these capabilities, and if so, implement similar permission protocols, including getting alerts when unauthorized users access or attempt to access privileged CIU data.

15 Artificial Intelligence tools are being developed that can make reviewing a file dramatically shorter. Having a program that looks for specific terms or phrases can be particularly helpful. Of course, to take advantage of any AI program, all paper files must be scanned into digital files that can be searched. This takes dedication of staff or even volunteers to make sure it is done properly.



Working in partnership with other agencies can improve efficiency of data collection where the offices are willing to share scanned PDFs of case documents. Many offices and organizations have thousands of paper files in boxes that have never been scanned electronically. Critical, rich data is contained within these paper files. To the extent possible, CIUs should urge elected DAs, executive staff, and local elected officials to prioritize at minimum the electronic scanning and digitization of paper case file jackets, documents, and evidence.¹⁶ For CIUs in particular, being able to preserve and later search this data can lead to efficient identification of bad actors, misconduct trends, and more.

A key aspect to using public data is to make sure the data is correct and reliable. Like we said before—*garbage in, garbage out*. For example, say you painstakingly track race/ethnicity data from every police report in every case you handle. But the police reports incorrectly captured the race/ethnicity for 50% of those people because the police guessed at the defendant's background rather than asking them directly. It may appear you have captured excellent data when in fact you cannot rely on it with confidence.¹⁷

STEP THREE – Decide What Data to Track

Set up the parameters for the basic information you need to track for your office. We have suggestions below, which are a good starting point. But you will want to do more with your data collection. In the Appendix, we have a ready-to-go spreadsheet with more robust data collection for you to use. Generally, data will fall into two categories: case management and case-specific information.

Remember this step—in addition to deciding what you will track—fundamentally also gives you a place to decide what not to track, which is important when faced with limited resources for data collection. Particularly for process metrics related to your own organization or CIU, tracking a few key metrics accurately rather than many poorly is almost always a better approach for advocating for more resources and understanding your functioning.

Case management – organizations need to be able to track the number of applications or inquiries they get, how long it takes to process requests, where a given request is at any time, and what the outcome is. Tracking this information makes it easier to respond to requests from funders or constituents about the organization's work level and efficiency at processing cases.¹⁸

First Level Data to track

- » When an application came in
- » When it was responded to
- » When an application/questionnaire was sent and received
- » What stage the case is in: under review, under investigation, in litigation, closed¹⁹
 - Why was it closed before reaching litigation: out of jurisdiction, not meet criteria, etc.
 - Use common data points to be able to create reports

Wherever possible, use drop-down menus that force the person entering the data to make a pre-determined choice. This prevents spelling mistakes or mischaracterizations, and also makes reviewing your data far easier than having to search through text fields. You can always add a brief text box to explain something if you need.

¹⁶ In many jurisdictions, after a certain amount of time, these paper documents are destroyed. Once this data is gone, if it has never been scanned electronically, it is gone forever.

¹⁷ Many tools exist to assist with filling in incorrect or missing race and ethnicity data by making predictions based on a person's name and home address, as compared to data from the U.S census. Depending on the information available these tools can often be between 75% and 85% accurate. For more, see this recent paper with empirical findings for the extent of misclassification in police reports: <https://onlinelibrary.wiley.com/doi/10.1111/1745-9125.12329>. Notable because they were able to validate the extent of misclassification empirically, while many previous articles have speculated this was likely due to data entry processes without being able to verify it.

¹⁸ These data points may have different titles, or mean different things, in different jurisdictions.

¹⁹ Most CIUs and innocence organizations conduct their reviews in stages: an initial screen for eligibility, a deeper review to understand the case dynamics, often a third stage where all documentation for a case is gathered and reviewed, an investigation phase, and finally litigation or resolution. Each stage involves a different level of information gathering and analysis.



- When it moved from one stage to another
- How long the case was in each stage
- » Who is responsible for the case at each stage
- » What documents have been gathered and from where
- » For cases that went to litigation, track the result or outcome of the case (and why)
- » Court information (case docket number)

Consider creating **Checklists** that align with your stage of review—court information, a discovery log, file index. Checklists help routinize the information gathering to make sure information is collected the same way in each case.

Second level information to track

- » What documents have been collected?
 - When was it collected?
 - By whom?
 - From whom?
 - Have they been electronically scanned (yes or no)
- » For IOs: Is there a CIU in the jurisdiction?
 - Has the client reached out to them?
 - Has the client reached out to other innocence organizations?
- » Has IO or CIU engaged an Expert?
 - At what stage?
 - What is the substance?
 - What documents have been/need to be sent?
- » For CIUs: Have you had contact with the victim or surviving family members?
 - When?
 - What is the best contact information?
 - Has the victim ever reached out to the DAO with questions?
- » Where did the case inquiry come from?
 - Applicant, IO, defense counsel, family/friend, reporter
 - When did inquiry come in? (from applicant or other)
- » Is there a colorable innocence claim?
 - Here track those factors known to lead to wrongful convictions.²⁰
 - Include a drop-down menu with itemized issues you want to track— eyewitness case, confession, juvenile defendant, questioned science, witness recantation, DNA, etc.

²⁰ The National Registry of Exonerations tracks multiple factors associated with wrongful convictions. A good resource is found here: <https://www.law.umich.edu/special/exoneration/Pages/ExonerationsContribFactorsByCrime.aspx>.



Case-specific information – because CIUs and IOs are uniquely positioned in being able to review cases in their entirety, they can learn from past mistakes and prevent future ones. Collecting case-specific data allows organizations to look across cases to identify patterns in evidence or common factors within a jurisdiction that have contributed to wrongful convictions.

- » Applicant information – name (including aliases), age, gender, race, language spoken, immigrant status (for issues pertaining to potential release), and location
 - If English is not the applicant's first language, was there an interpreter at trial, interrogation?
 - Age at incident, age at arrest, interrogation, trial, sentencing
 - Where is that information gathered from?
 - Any potential communication or disability issues
- » Charges being contested
 - Case tracking number(s) for trial, appeal
 - Arrest location, date; Incident date
- » When the sentence will be complete (for jurisdictions where post-conviction relief is limited to those serving a sentence)
- » Case actors—investigating police, prosecutors, judges, defense counsel, informant
- » Any co-defendants and the status of those cases²¹
- » Victim information—name, gender, race, age
 - Known to applicant?
 - Relationship to applicant (if any)?
- » Red flag concerns: did the case involve DNA, bad science, misconduct potential, eyewitness ID, confession, or other canonical factor?
 - Where did that come from—applicant or case review
- » County/geographic information
- » *Brady* disclosures from the record
- » How did case resolve (trial/plea)

Information on processing

A key factor for many organizations is documenting the need for more resources. One way to do that is to be able to show funders and supporters the length of time it takes to move a case from one stage to another: how long it takes to decide whether to move an initial application forward, gather documents, review documents, conduct initial investigation, etc. By showing a funder how long those phases take, organizations can make a solid data-driven case for additional resources.

It is fairly easy to incorporate timing into data keeping. For example, Excel, Google, and other available spreadsheets have a simple time calculation formula that can be used to automatically compare data in one column to another, including the days between dates.

Whenever possible, type data into fields in structured forms consistent in language, terminology, appearance, layout, sizing, and formatting. This way, if you ever need to scan documents, it will be much easier for OCR scanners to extract the data. Create a data dictionary and try to consistently use the same data definitions in everything you collect and analyze.

²¹ For innocence organizations this is particularly important to avoid any issues of conflict among prospective clients.



STEP FOUR – Train Staff and Volunteers on Data Keeping

Train staff—including student volunteers or interns—on proper data keeping, and the importance of tracking data for case information as well as case management. Lots of offices have a negative attitude toward tracking data: it's seen as too time-consuming or just not important. As your staff understands the importance of clear data tracking, you can move the culture of the entire office. Having and even analyzing data within a small unit is great but not as helpful as it could be if it doesn't reach decision-making level for your office. For CIUs especially, promoting transparency and good data tracking is a critical part of the community's faith in their work.

It is important to ensure that anyone working with data has a clear understanding of best practices and quality standards for that work. Set expectations about what good data entry looks like (e.g., standard double-check at point of first entry; two-person review; general reinforcement of attention-to-detail) and lay out your expectations of everyone clearly. Consider possible training exercises or create opportunities for experienced staff to share tips.

Additionally, consider building in time to your workflow every week, month, or quarter to assess the data you're tracking for quality and accuracy. Whenever manual entry is used for data collection, some errors are almost guaranteed. It is much easier to resolve problems with missing or inaccurately input data when there are relatively few entries and the time between data entry and review is quick, rather than waiting several months or years to address quality concerns.

MOVING TO THE NEXT LEVEL

When to start tracking information

A key question for participants was when to start tracking information—upon an initial inquiry, when a questionnaire or other application has been received, etc. The consensus was that information should be tracked as it comes into the organization and built up from there.

At a minimum, most organizations want to know the applicant's background (race, age, gender, potential disability), the victim's background, the investigating officers, prosecutors, defense counsel. Most of this information can be obtained from public court dockets or documents.

Tracking where data comes from for reliability

Where data comes from matters; some sources are more reliable than others. An applicant may have inaccurate information about their own case, or what a witness did or did not testify to at trial. Organizations should be aware of the source of information and that it may not be fully reliable.

Relatedly, knowing why you're collecting data helps with understanding accuracy: you may not want to know self-reported red flags, for example, or at least record it as reliable until it's been verified.

Integrating intake and data collection

Consider developing an intake tool that allows for digital upload and integration into your data program. This can help reduce input errors as well as staff time.

Digitizing old files and making sure they can be searched will yield information on decisions you've made and potentially identify cases that should be reopened. This will also allow you to quickly search for and isolate cohorts of cases that involve specific bad actors, rather than requiring manual paper file reviews. Outside partnerships may be particularly effective at assisting with this.



Increasing information tracked to get better answers

The types of data you track will determine the types of questions about your cases you can ask. If you want to understand whether there are particular bad actors or more problematic geographic areas, you must be tracking that information. Consider the questions you want to be able to answer from the cases you're reviewing:

- » How long does it take us to respond to a request?
 - You'll have to track time increments from receipt of application to response, to decision.
- » What types of methods are police using in identifying suspects?
 - Tracking the types of identification procedures used with detail including blinded administration, warnings to witnesses, post-identification feedback, or other issues
- » How many problematic cases involve alleged confessions?
 - Knowing which cases included an alleged confession from police, how long that interrogation lasted, how young the suspect was at the time, what types of questions were asked can all yield important information for developing new trainings with detectives.
- » What are the racial disparities in our cases?
 - Tracking the race of the convicted individual, victim(s), police, prosecutors, defense counsel, judges, and jury makeup and comparing that to the type of error (official misconduct, bad forensics, etc.)

Keep track of developing technology

Technology is constantly changing—both inside offices and organizations and on the market. Very recent advances in the ability to extract data from documents without manual review, as well as advances in large language models and other forms of AI's ability to sort and classify documents means there may have never been a better time to more efficiently collect data to improve the criminal justice system.²² You should always be asking: what tools can we use to quickly identify entire cohorts of cases for review?

Keep up on internal data and tech resources within your office or organization

Identify external academic partners, technical assistance providers, and funders that may be able to provide personnel, technology, and assistance in data collection, analysis, implementation of technical solutions, data dashboards, and other tools to assist you.

Some questions to ask:

- » What is the data transparency culture of my office/organization/institution (e.g., is there an appetite to share data freely internally and/or externally)?
- » What is the data capacity of my overarching office/organization/institution? Does my overarching office/organization/institution have the capacity to assist me (and can I trust them to keep my work confidential?) as I build out the data and tech capacity needed to identify, triage, and investigate cases?
- » Do I need a full-time team to work alongside me, or can I bring in a data and tech provider one time to implement needed data and technology interventions?

²² Exciting projects are happening all over the country in the criminal justice field. Here are a few to check out:

The Possibility Lab at Berkeley is looking at ways to use data to help with prosecutor-led resentencing: <https://possibilitylab.berkeley.edu/project/can-we-make-it-easier-for-prosecutors-to-re-sentence-people-in-prison/>

The Human-Centered Artificial Intelligence program at Stanford University used machine learning to search through 35,000 parole hearing transcripts to look for information on race and attorney representation data to inform the process: <https://law.stanford.edu/event/reading-35000-parole-hearing-transcripts-a-new-direction-for-machine-learning-in-criminal-law/>



- » Are there municipal, county, or state data and tech resources that can support me as I build out the office's data and tech capacity?
- » Are there public or private funders, academic institutions, and/or technical service providers that I may consult to help me build the office's data and tech capacity?
- » Are there any law firms or private companies in my area that can help scan large quantities of paper documents?
- » Are there any networking entities (such as Quattrone) that can help me identify needed data and tech resources, and form partnerships to help me build the office's data and tech capacity?
- » What are the most up-to-date tech innovations and tools that I can use to save time and enhance my work?



APPENDIX



DATA TO COLLECT - OUTLINE

CASE MANAGEMENT INFORMATION

Original Court Case Number: (the main identifier)

CIU File Number:

MDOC Number:

Last Name

First Name

DOB

Highest Education level [drop down]

Was applicant a juvenile? [multi drop down: at incident, at arrest, at interrogation, at sentencing]

Gender [drop down]

Ethnicity [drop down – Black, White, Hispanic, Asian, Latinx]

Native Language [drop down]

Represented by counsel on application? [y/n]

If yes, name of counsel [text]

Innocence organization? [y/n]

Waiver obtained? [y/n/partial]

Mental Health Diagnosis? [y/n with explanation for yes]

Medical/Cognitive/Emotional Disability? [drop down]

Prior Criminal Record? [y/n with explanation for yes]

Case Status

In queue to be screened

Open

Closed

Date Initial Letter Received (date)

How Received? Select One (radio buttons or drop down)

Case meets CIU Basic Criteria? (y/n/idk)

If no, reason Select One (radio buttons or drop down)

Unfounded,

Not claiming innocence

Not post-conviction

police misconduct

Date Closed (date)

Closing letter sent (y/n/IDK/NA)

Closing letter date

Time to evaluate [calculate time b/w letter received and decision made]

Connected to Grant? (y/n)

Currently incarcerated? [y/n]

Sentence to serve [text]



Where incarcerated [drop down with institutions]

Unit Conflicts (y/n/idk – default to idk)

If yes, describe – (open text)

Application form sent – (y/n default to n)

If yes, date sent (date)

Application received by CIU – (y/n)

If yes, date application received (date)

If yes, length of time to receive [calculate date b/w sent and received]

Assigned for CIU screening – (y/n)

If yes, date assigned for CIU screening (date)

If yes, identify CIU screener (open text)

Date CIU screening completed (date)

Time between application sent and assignment [calculate time]

Time between screening assigned and completed [calculate time]

Credible Claim of Innocence? (y/n/maybe)

Date determined (date)

Case Flag: *INCLUDE ANY FLAGS YOU WANT TO TRACK*

Case Flag: police misconduct

Case Flag: systemic issue

Case Flag: arson

Case Flag: bitemark

Case Flag: DNA

Case Flag: Other forensic flag

Case Flag: informant testimony

Case Flag: false confession

Case Flag: eyewitness identification

Case Flag: prosecutorial misconduct

Case Flag: recanting witness

Case Flag: discredited expert

Case Flag: Brady claim

Case Flag: ballistics

Case Flag: serology

Case Flag: Hair microscopy

Case Flag: Fiber evidence

Case Flag: SBS/AHT

Case Flag: Ineffective Defense Counsel

Unsubstantiated reason: (open text box)

Assigned? (y/n) (if yes)

CIU Attorney (name)

Date assigned to CIU Attorney

Time b/w screening and assignment [calculate time b/w dates]



Review completed date

Time b/w review assignment and completion [calculate time b/w dates]

Outcome of CIU Review: select one

Relief Denied - checkbox

Reason (Select one):

unsubstantiated

new evidence supports guilt

Improvidently reviewed (should have been screened out)

other

Other: text box

Relief Recommended

DA Relief Memo Submission Date

DA Memo Decision Date

DA Decision: Select one

Exoneration

No Exoneration, New trial, case dismissed

No Exoneration, New trial, case not dismissed

No Exoneration, No Relief

Case outcome: Select one

Exoneration

No Exoneration, New trial, case dismissed

No Exoneration, New trial, case not dismissed

No Exoneration, No Relief

Offer a plea

Back for investigation



CASE INFORMATION

Offense Date

Offense Location [police district drop down]

Officer in Charge

Other Officers 1

Other Officers 2

Other Officers 3

Other Officers 4

Other Officers 5

Arrest charges

Number of Victims [text for all names and ages]

Victim Name

Victim Age at time of offense

Victim race (self-reported or otherwise? Y/N/DK)

Victim gender (self-reported or otherwise? Y/N/DK)

Victim relationship to applicant

None/Stranger

Intimate Partner

Family

Friend

Other (open text)

Number of co-defendants

Names of co-defendants

Any convicted? y/n

If yes, convicted of lesser charge? (y/n)

Any tried by different jury? (y/n)

Any co-defendant acquitted? (y/n)

Arrest Date



Calculate time from incident to arrest date

Arrest Location
Original Trial Prosecutor Name
Defense Trial Counsel Name
Appointed counsel (y/n)
Retained counsel (y/n)
Any bar complaints? [y/n]
Judge Name
Trial? (y/n)
Length of trial
Jury
Bench
People that Testify (embedded table)
Pleaded Guilty (y/n)
Date
Pleaded no contest (y/n)
Date
Conviction date
Conviction offense (highest)
Other conviction flags: sexual assault component
Other conviction flags: arson
Other conviction flags: shootings
Highest Sentence
Years (range)
Life with parole
Life, mandatory
LWOP
Sentencing date
Age of person convicted at time of offense
Forensic evidence collected (y/n/idk)
Forensic evidence analyzed (y/n/idk)
Items that linked person convicted to the crime: (checkboxes)
Eyewitness
Ballistics
Serology
Arson or Fire Scene investigation
DNA
Hair evidence
Fiber evidence
Fingerprint evidence



Photo or video

Other media

Co-defendant or informant testimony

Other: (open text)

Experts that prepared a forensic report [room for 10, check for Defense or Prosecution]

Eyewitness identification [y/n]

If yes: number of witnesses

stranger to applicant? [y/n]

Procedure [drop down: showup, photo array, live lineup, none]

Flags on procedure: [drop down: not blind, no warnings, suggestive nature, other]

Interrogation (y/n)

Confession (y/n)

If yes: age at time of interrogation

Length of interrogation

Interrogation recorded? [y/n]

Use of deception evident? [y/n]

Informant information: [alias, convicted name]

Alibi (y/n)

Presented at trial? [y/n]

Investigated by police/prosecution? [y/n]

Applicant testified at trial? [y/n]

Crimestoppers tip? [y/n]

Conviction Appealed? (y/n)

Post-conviction appeal? [y/n]

Federal habeas? [y/n]

All prosecutors on appeal:

Documents located:

police file

trial transcripts

laboratory reports

appellate file

Appellate/federal court files

Documents shared with defense: [drop down]

Police file

Prosecutor file

Forensics/Lab file

Other investigative materials

Dates documents shared

How shared [drop down – electronic, hand delivery, mail]



Physical evidence exists (y/n/IDK)

If exists, forensic evidence retesting requested? (y/n)

Ballistics

Serology

Arson or Fire Scene investigation

DNA

Hair evidence

Fiber evidence

Fingerprint evidence

Photo or video

Other media

Outcome of new forensic evidence testing (open text)

Inculpatory

Exculpatory

Indeterminate

Notes on new forensic testing (open text)

WITNESSES

Trial witnesses: [list all witnesses from investigation/trial]

Current locations: [present locations]

Outreach? [y/n]

When contacted?

How/Who contacted?

Interviewed? [y/n]

Substance [text]

Police witnesses:

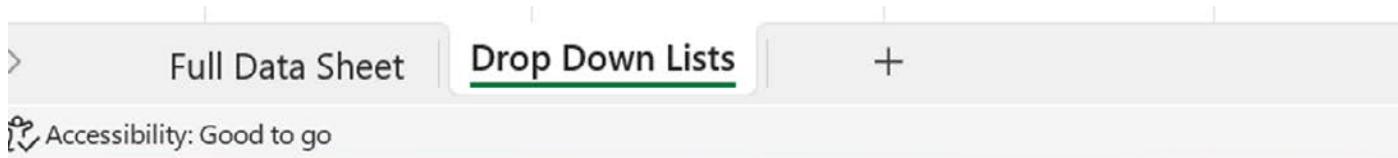
Outreach? [y/n]

Interviewed? [y/n]



GUIDE TO CREATE AND ADD DROP-DOWN ITEMS TO EXCEL

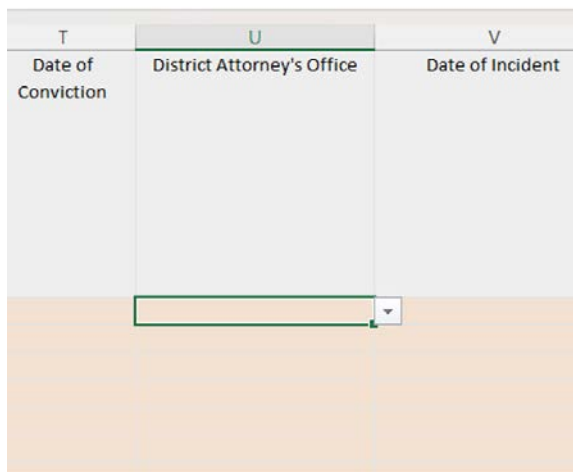
- 1) To make sure everything is uniform, create a separate sheet that has items on it to fill in drop down menus.



- 2) Create lists of the data you want to use in drop down menus.

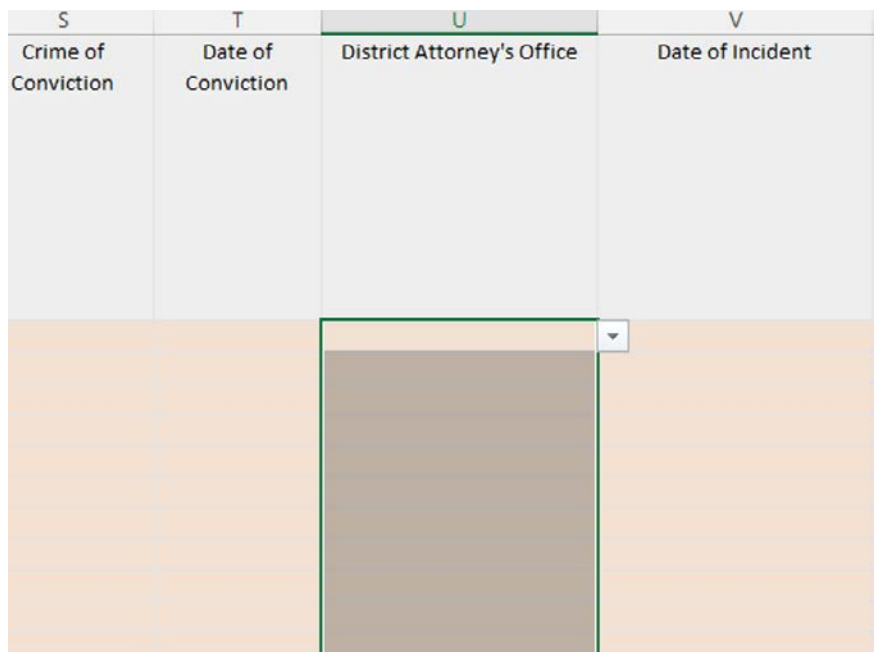
District Attorney's Office	Race	Education Level	Yes/No/Unknown/NA
County 1	Black	K-8th Grade	Yes
County 2	White	Some High School	No
County 3	LatinX	High School Graduate	Unknown
County 4	Asian	GED	N/A
County 5	Cape Verdean	Some College	
County 6	American Indian	College Graduate	
County 7	Pacific Islander	Beyond	
County 8	Other	Unknown	
County 9	Unknown		
County 10	N/A		
County 11			

- 3) Back on your Data Sheet, highlight first cell below column header where you want a drop down.

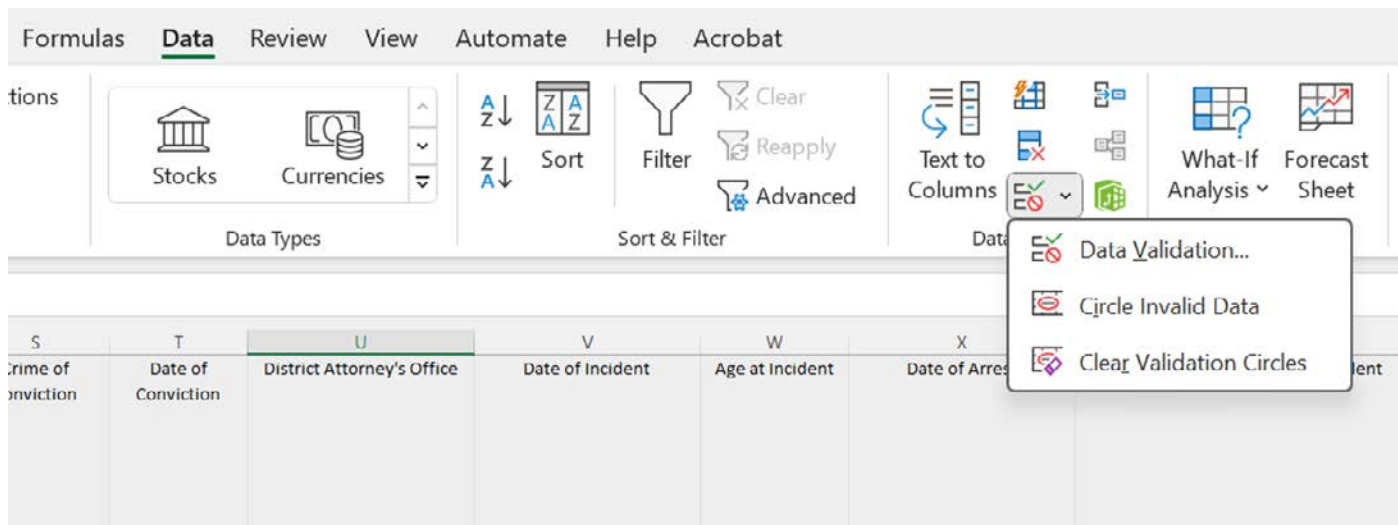




4) To select all the cells in the row, hold down SHIFT + CONTROL + Down Arrow.

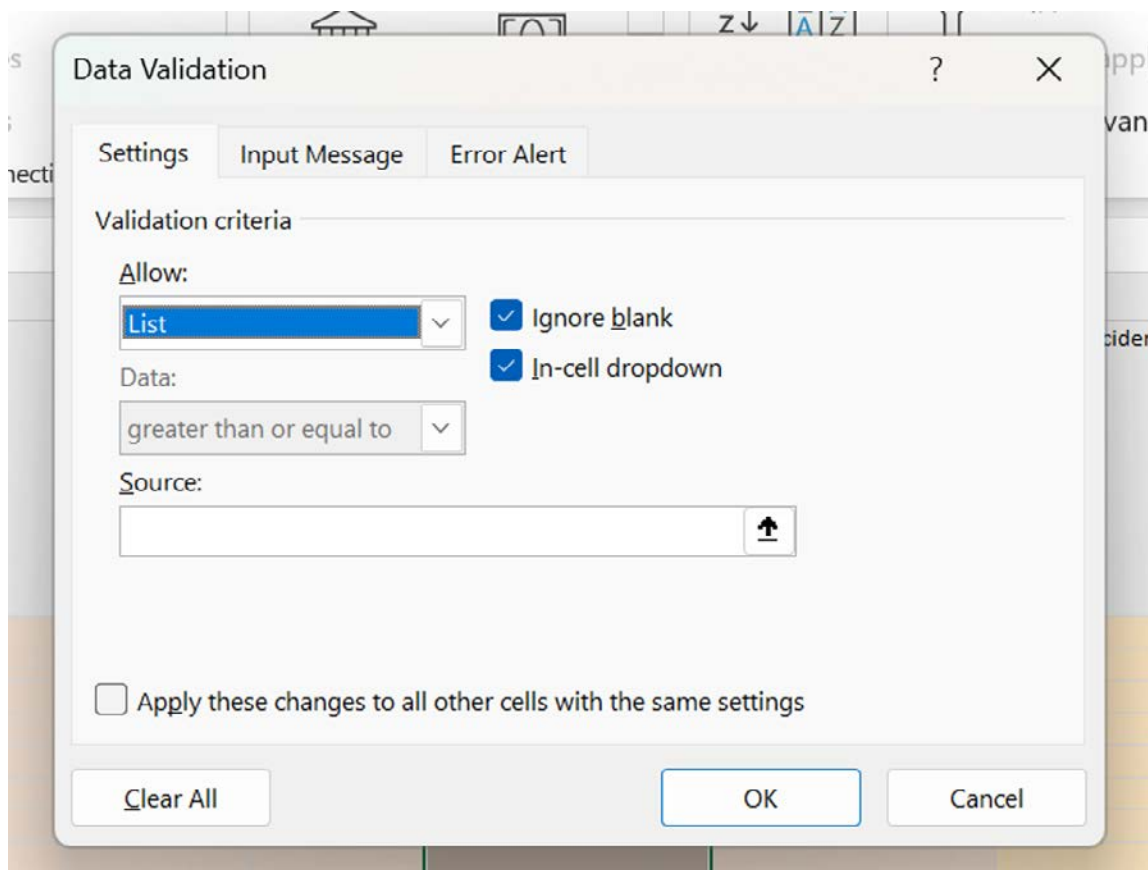


5) Under DATA tab, Click on Data Tools then select Data Validation.





6) Under Allow: select List.



7) In the Source field, you can add your own selections to be displayed separated by commas (Yes, No, Unknown) as choices OR go back to your Drop Down Sheet and choose what you want to include as possible responses. Click on the first item in the list (not the header) and, holding down the cursor, drag down to the last item.

District Attorney's Office
County 1
County 2
County 3
County 4
County 5
County 6
County 7
County 8
County 9
County 10
County 11



8) Hit OK to accept the data fields.

District Attorney's Office	Race
County 1	Black
County 2	White
County 3	LatinX
County 4	Asian
County 5	Cape Verdean
County 6	American Indian
County 7	Pacific Islander
County 8	Other
County 9	Unknown
County 10	N/A
County 11	

Data Validation

Settings | Input Message | Error Alert

Validation criteria

Allow: List Ignore blank In-cell dropdown

Data: greater than or equal to

Source: ='Drop Down Lists'!\$D\$2:\$D\$14

Apply these changes to all other cells with the same settings

Clear All OK Cancel

9) If you ever want to change the options, rather than going into each column individually on the Drop Down Sheet and make the changes there. It will automatically update any column using that data on the Data Sheet. Just make sure you don't add any rows or they won't appear; you'll have to redo the process.

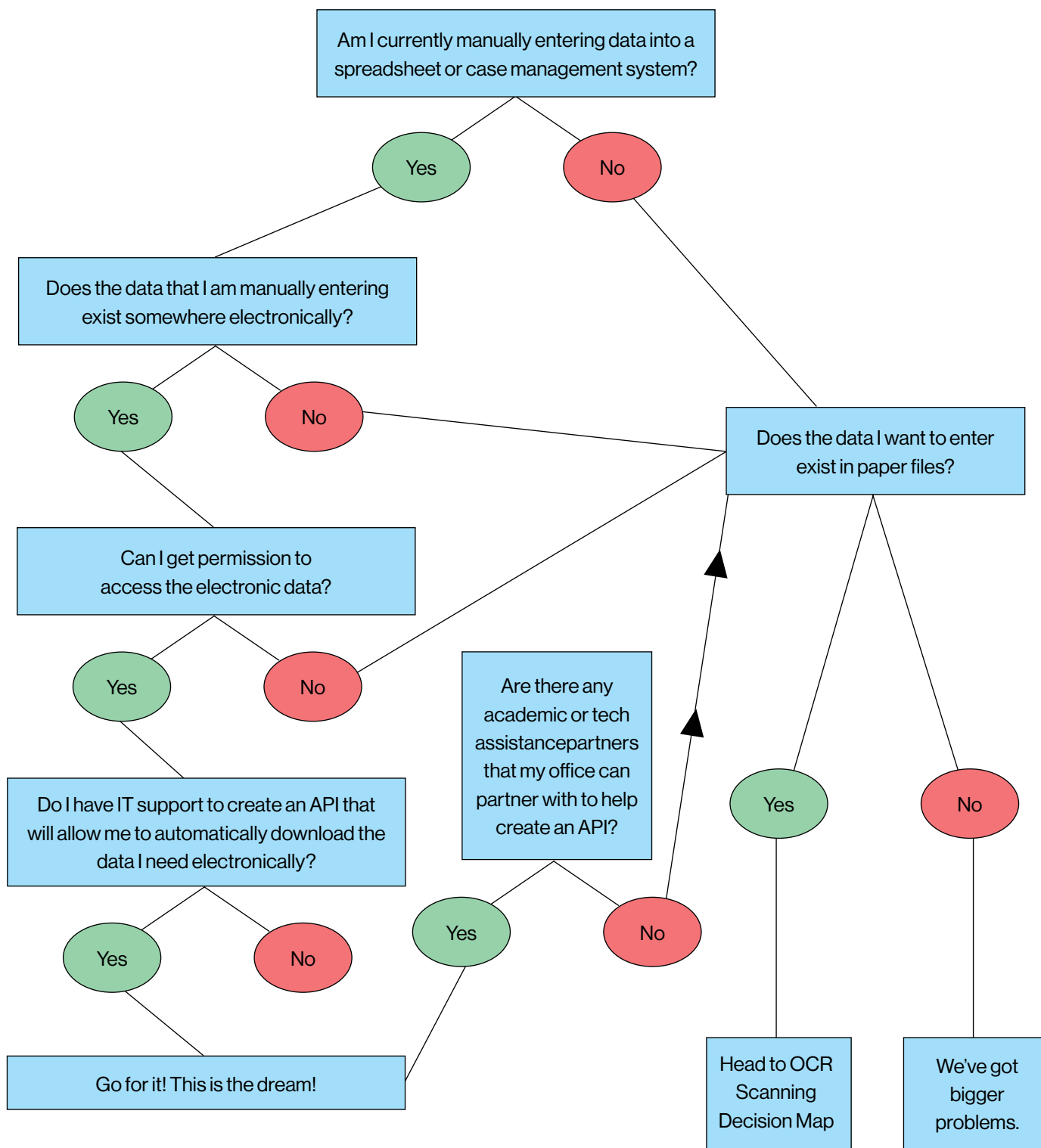


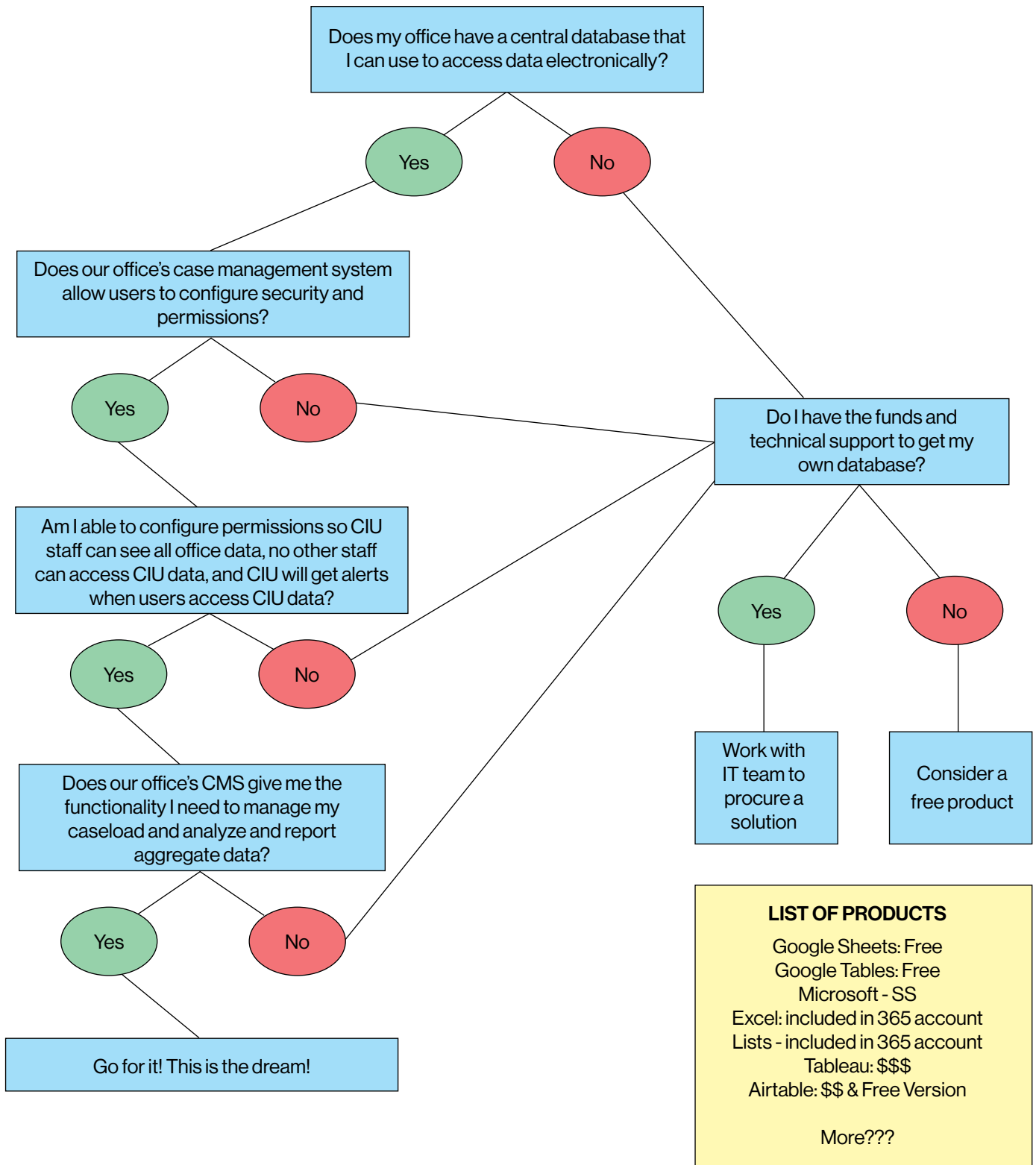
LINK TO EXCEL SHEET INCLUDING DROP DOWN ITEMS

<https://www.dropbox.com/scl/fi/vkwpta9za3kt730ujpw61/SAMPLE-DATA-SHEET.xlsx?rlkey=0krr8wq5xdqfz1ojvv2gdipeu&dl=0>



DECISION TREES







DISCUSSION PARTICIPANTS

CONSULTANTS

Andrew Taylor, Vera Institute

Bobby Constantino

PARTICIPANTS

Nancy Adduci, Cook County (IL) Conviction Integrity Unit

Ally Bennett, Middlesex County (MA) Conviction Integrity Unit

Jennifer Bergeron, Ohio Innocence Project

Adele Bernhard, New York Law School Innocence Clinic

Harry Breault, Institute for Innovation in Prosecution

Barbara Buhai, Lake County (IL) Conviction Integrity Unit

Tiffany Bush, Georgia Innocence Project

Isabel Corn, Georgia Innocence Project

Sara Desimone, Middlesex County (MA) Conviction Integrity Unit

Demarris Evans, Alameda County (CA) District Attorney's Office

Robyn Frankel, Michigan Attorney General's Office Conviction Integrity Unit

Patrick Gilbert, Wicomico County (MD) Conviction Integrity Unit

Jaime Henderson, Innocence Project

Libby Huegetz, Committee for Public Counsel Services (MA) Innocence Program

Ayyub Ibrahim, Innocence Project New Orleans

Lisa Lazzari-Strasiser, Pennsylvania Attorney General Conviction Integrity Unit

Jennifer Kamorowski, MidAtlantic Innocence Project

Lorna McClusky, Shelby County (TN) Justice Review Unit

Vanessa Meterko, Innocence Project

Becky Michaels, Northwestern (MA) County Conviction Integrity Unit

James Moody, MidAtlantic Innocence Project

Kimberlee Moran, New Jersey Innocence Project

Valerie Newman, Wayne County (MI) Prosecutor's Office

Gail Pamukov, Macomb County (MI) Conviction Integrity Unit

Brad Roach, Pima County (AZ) Conviction Integrity Unit

Barry Scheck, Innocence Project

Nyssa Taylor, New Jersey Innocence Project

Patrick Tran, Illinois Attorney General's Office

Leticia Uzueta, Minnesota Attorney General Conviction Review Unit



Quattrone Center for the
Fair Administration of Justice
UNIVERSITY *of* PENNSYLVANIA

3501 Sansom Street
Philadelphia, PA 19104-6204