



Project Summary

The use of data to inform prosecutorial decisions is a key tenet of “progressive prosecution,” a new relatively approach to criminal prosecution that seeks to reverse the role of prosecutors as drivers of mass incarceration by shifting the goals of prosecutors from conviction and punishment to fairness and accountability. The San Francisco District Attorney's Office (SFDA) is committed to this goal.

Unfortunately, the antiquated nature of criminal justice data and many criminal justice data systems makes it extremely difficult to actually use data, especially in real-time to inform daily case-specific decisions. This is especially true for Assistant District Attorneys (ADAs) in our Intake Unit, who must review extensive information about suspects and alleged criminal conduct housed in multiple data systems and in varying—and often text-based—formats to decide whether to discharge, divert, or prosecute a given case. If diversion is most appropriate, the ADA must further discern the most appropriate diversion program based on an array of person and case characteristics; similarly, if prosecution is appropriate, there are similarly difficult decisions to make, particularly for harder to identify and prosecute cases, such as human trafficking.

Despite these challenges, SFDA is committed to using data to make fairer and more equitable decisions, including better identifying cases for discharge, diversion, or prosecution. In late 2021, we will be implementing a new electronic case management system, which will put us in a prime position to achieve this goal. Nonetheless, the disparate and semi-structured nature of the data that feed into this new data system will continue to impose limitations if we do not address them. Toward this end, SFDA proposes *Justice Driven Data Science for Prosecutorial Impact*, an innovative effort to use data science and machine learning to leverage criminal justice data to inform our most critical decisions.

This two-year project will begin with a research effort to allow us to better quantify critical characteristics of cases that have been or should have been discharged, diverted, or prosecuted. Based on this, we will build a series of algorithms that can inform and improve prosecutorial decision-making by flagging cases as good candidates for specific prosecutorial action. While all cases will still be reviewed by an ADA who will make the decision to discharge, divert, or charge, by using data science to identify key case characteristics that sit in disparate places across SFDA data and flagging cases based on likely appropriate paths, this effort will serve three distinct but interrelated goals: first, reducing the prosecution of cases that pose a low risk to public safety and thus are shown by a growing body of research to be more effectively addressed without formal prosecution; second, reducing SFDA caseloads through earlier and more effective identification of cases for discharge or diversion; and three, increasing our prosecution of difficult-to-identify but high priority cases that pose a significant risk to public safety. **We believe that this pioneering effort, if successful, can transform prosecution across California and the US by creating a replicable model of true data-driven prosecution.**



Project proposal:

Problem Statement

The practice of prosecution is at a crossroads. After decades of prosecutorial practices that drove mass incarceration and exacerbated racial and ethnic disparities in the criminal justice system, the last few years have seen the emergence of a new approach to prosecution. Progressive Prosecutors are seeking to shift the focus of prosecution from conviction rates, punishment, and incarceration to “fairness, equity, compassion, and fiscal responsibility.”¹ At the San Francisco District Attorney's Office (SFDA), we take this charge seriously and are committed to avoiding unnecessary incarceration and reducing racial and ethnic disparities, while ensuring accountability and community safety.

Unfortunately, operationalizing this vision is far from easy; amid high caseloads and almost totally manual processes for case intake, review, and decision-making, it is difficult for our office to ensure that we are responding to each case—and the people impacted therein—in the most effective, equitable, and appropriate way possible. San Francisco is not alone in these challenges. Across California, district attorney's offices grapple with both heavy caseloads and highly manual intake and review processes, the latter of which is largely the consequence of outdated case management systems.²

For the SFDA much of this challenge sits within our Intake Unit, in which 7 Assistant District Attorneys (ADAs) are responsible for reviewing the 600-1200 arrests that are made in San Francisco every month and deciding whether to divert, discharge, prosecute, or respond in some other way. In 2020, SFDA's Intake Unit reviewed almost 9000 felony and misdemeanor arrests and 3600 misdemeanor citations that were presented to our office, with one misdemeanor intake) reviewing approximately 2,600 misdemeanor arrests and all non-custodial citations. Consistent with the SFDA's mission, the ADAs doing the initial review of cases strive to identify and divert all eligible cases, discharge all cases that involve racially discriminatory practices such as pretextual stops, and prosecute cases that are the source of significant community harm. There are, however, complexities that affect their ability to successfully recognize each.

In terms of diversion, identifying appropriate cases is surprisingly difficult, as well as time consuming. SFDA has 10 pre-filing diversion programs in addition to the San Francisco Superior Court's 7 post-filing collaborative courts. Each program has distinct eligibility criteria, some of which are tied to suspect/defendant characteristics, such as age, neighborhood of residence, prior criminal history; and some of which are tied to offense characteristics, such as violent or nonviolent, drug-related, etc. All of the information related to these characteristics is stored in disparate locations across police reports, the California Law Enforcement Telecommunications System (CLETS), and the SFDA data system. Moreover, because the San Francisco Police Department (SFPD), the primary arresting agency in the City and County of San Francisco, does not have any law enforcement-led diversion programs, it is likely that the SFDA receives a higher than average proportion of cases that are good candidates for diversion.

Identifying cases for discharge can be difficult as well, particularly for cases that SFDA discharges because evidence was collected during a pretextual stop. These cases usually involve vehicle or pedestrian stops of young Black or Latino men who are not involved in any criminal conduct at the time

¹ <https://fairandjustprosecution.org/about-fjp/our-work-and-vision/>

² See *The California Criminal Justice Data Gap*, 2019. Stanford Criminal Justice Center. (https://www-cdn.law.stanford.edu/wp-content/uploads/2019/04/SCJC-DatagapReport_v07.pdf)



San Francisco District Attorney's Office

of the stop but who are subsequently determined through a search to be in possession of weapons or controlled substances. Because the relevant information that can alert intake ADAs to a pretextual stop is often buried in a police report as well as across SFDA data elements, these cases are not always caught at Intake.

Missing or misidentifying these cases can have critical consequences for people's lives and for society more generally. Prosecuting cases that would have been more appropriately responded to via diversion and/or social service provision can saddle people with criminal convictions and all of the downstream collateral consequences thereof. Prosecuting cases based on pretextual stops exacerbates the already dire racial/ethnic disparities in the criminal justice system, while sanctioning and even encouraging racist law enforcement practices.

Prosecuting these cases can also have serious consequences for public safety. Consistent with a sizeable body of juvenile justice research showing that contact with the juvenile justice system actually increases the likelihood of future delinquent or criminal conduct, recent research indicates that the prosecution of lower level incidents substantially increases the risk for future arrest and prosecution.³

The unnecessary—and counter-productive—prosecution of cases also exacerbates caseloads that already far exceed recommended standards. ADAs in our general felony units handle approximately 253 cases a year, 69% greater than the 1973 standard, while ADAs in our misdemeanor unit handle approximately 290 cases a year (roughly equal with the revised standard). In a 2011 Northwestern University Law Review article, authors rightly point out that “the ramifications of excessive prosecutorial caseloads extend throughout the criminal justice system and, perhaps surprisingly, are most harmful to criminal defendants. Excessive caseloads lead to long backlogs in court settings, including trials, and bottom-line plea bargain offers.”

One of the biggest consequences of SFDA's high caseloads is the limited capacity ADAs have for identifying and prosecuting more complex crimes, despite the fact that many of those crimes can be the source of significant community harm. For example, cases that involve human trafficking/commercial sexual exploitation, are often presented to our office due to arrests for other allegations, such as theft or drug related charges. It is incumbent upon the Intake ADA reviewing the allegations to identify characteristics of the suspect and incident that indicate a potential link to human trafficking and/or commercial sexual exploitation so that the case can be passed onto the Human Trafficking Unit for further review, investigation, and prosecution. Similarly, domestic violence incidents in which there is a high risk of subsequent—and potentially lethal—violence require additional review at intake and attention by our Domestic Violence Unit to ensure reviewing ADAs notice the warning signs and respond appropriately.

Similar to the challenges identifying cases that are likely candidates for diversion or discharge, identifying cases requiring further review and possibly specialized prosecution is complicated by the antiquated format of information to be reviewed and the consequently highly manualized process ADAs must use to discern relevant case, defendant, and victim characteristics to make a charging decision. Police reports, although received electronically, are essentially text documents with minimal structure with which to organize and extrapolate the information of interest to ADAs. Similarly, RAP sheets and

³ Agan, AY; Doleac, JL; Harvey, A. 2021. “Misdemeanor Prosecution.” *NBER Working Paper Series*. National Bureau of Economic Research: Cambridge, MA.



San Francisco District Attorney's Office

other criminal history information as electronic text documents that list suspects' prior arrests, prosecutions, and convictions, but do not in any aggregate this information or create variables that could inform SFDA's decision to prosecute. Even SFDA's own data system is currently not set up to aggregate and highlight relevant suspect and/or case characteristics that might inform the appropriate action to take on a given case. A modular case management system that was initially implemented in 2003, the SFDA's DAMION case management system has limited functionality and no ability to "flag" cases based on designated variables across various data elements.

In late 2021, however, our office is set to deploy a new case management system with significantly improved functionality and greater capacity to pull in additional data from other criminal justice agencies in the City and County of San Francisco. This puts us at the perfect juncture to better leverage data science and information technology to inform our discharge, diversion, and prosecution decisions and, in so doing, reduce time spent on cases that pose a low risk to community safety while increasing time available to prosecute cases that are the source of significant harm. Toward that end, the SFDA proposes Data Science for Prosecutorial Impact: a two-year effort to leverage research and technology to proactively identify cases as likely 1) candidates for discharge based on evidence available; 2) eligible for diversion, or 3) requiring special attention and possibly additional investigation and prosecution. Identifying cases for the first two categories will support our efforts to safely reduce caseloads and thus support the third, which will increase our ability to enhance community safety

The Policy Idea

Justice Driven Data Science for Prosecutorial Impact will use qualitative and quantitative research, machine learning, and data science to build a series of algorithms that can inform and improve prosecutorial decision-making by flagging cases as good candidates for specific prosecutorial action. While all cases will still be reviewed by an ADA who will make the decision to discharge, divert, or charge, by using data science to identify key case characteristics that sit in disparate places across SFDA data and flagging cases based on likely appropriate paths, this effort will serve three distinct but interrelated goals: first, reducing the prosecution of cases that pose a low risk to public safety and thus are shown by a growing body of research to be more effectively addressed without formal prosecution; second, reducing SFDA caseloads through earlier and more effective identification of cases for discharge or diversion; and three, increasing our prosecution of difficult-to-identify but high priority cases that pose a significant risk to public safety.

Theory of Change and Expected Outcomes

At its core, our theory of change is that we can use research and technology to better determine the best action for different cases that are presented to our office. In so doing, we can reduce harmful prosecutions and increase appropriate prosecutions, thus increasing the "fairness, equity, compassion, and fiscal responsibility" with which we operate. The direct outcomes we expect to see are an increase in cases identified for and referred to diversion, an increase in discharges of cases involving pretextual stops, and increases in prosecution of complex and harmful cases including human trafficking and domestic violence. More distally, this effort can reduce the number of people entangled in the criminal legal system, reduce racial disparities, and increase community safety and wellbeing. Moreover, should this effort succeed in San Francisco, it will be replicable in prosecutors' offices across California and the United States.



San Francisco District Attorney's Office

Given the SFDA's role at the forefront of California's progressive prosecution movement, we are particularly well-situated to disseminate this intervention, should it be successful. In 2020, the San Francisco District Attorney's Office became one of the founding members of the Prosecutor's Alliance of California, a membership organization composed of California prosecutors committed to criminal justice reform. Through this organization—whose members lead district attorney's offices that are responsible for more than one-third of California's felony prosecutions—reform-oriented prosecutors collaborate to share programs and strategies that promote their shared interest in reform.

Data for Outcome Measurement

The primary outcomes of interest for this project are the increases in diversion, discharge, and prosecution of appropriate cases for each of those actions. The primary data sources for each will be derived from the SFDA case management system, in which we track all cases presented to our office by law enforcement agencies as well as how we respond to/act on each. Because we want to track increases in appropriate identification (as opposed to the overall number of cases identified for each action), we will also want to analyze historic data to determine how many cases were accurately identified and how many potentially eligible cases were missed for each category.

In addition, qualitative data collection such as interviews with attorneys from different units will provide additional information regarding the efficacy with which case flags are identifying cases. Toward this end, we will work with our evaluation partner to identify appropriate respondents in Intake who make action recommendations, as well as attorneys who oversee our diversion programs and those who prosecute cases flagged for prosecution.

Capacity to Carry it Out

The SFDA's well-established commitment to data-driven decision making and our forthcoming implementation of a new and more flexible case management system make this the perfect project at the perfect time. Currently the SFDA analyst team, IT unit, policy director, and ADAs responsible for overseeing different units and divisions work closely together to review data on a regular basis in order to track and assess our decision-making. While these reviews are always intended to inform policy and practice, the data system limitations and other challenges discussed above limit our ability to use data to inform our decision in real time.

In addition, our data and analytics team currently lacks the staff capacity to implement an ambitious effort such as this, including conducting both primary and secondary qualitative and quantitative research and developing programs to implement algorithms based on those analyses.

To implement this project, SFDA will need to collect both primary and secondary research to identify key characteristics of cases that are likely eligible for diversion or discharge, or that require additional review and consideration for prosecution. This will include 1) review of best practice research on diversion and prosecution, with special attention to the prosecution of complex or hard to identify cases involving vulnerable victims, such as human trafficking and high lethality domestic violence; 2) interviews with SFDA's Intake Unit and Diversion Unit attorneys and paralegals, as well as with attorneys who focus on prosecuting domestic violence and human trafficking cases, and with staff from the Victims' Services Division who provide services to human trafficking and domestic violence survivors; 3) developing coding schemas to use machine learning and/or data science to pull relevant case characteristics from text-based documents; and 4) quantitative analysis of historical cases in each of these categories.



San Francisco District Attorney's Office

Quantitative analyses will examine those cases identified as fitting within their respective categories at initial intake as well as those not initially identified but subsequently discharged, diverted, or moved into a special prosecution to see if there are different characteristics that define easier and harder to identify cases. Having conducted these research steps, the project lead will work with the SFDA's IT Department and case management system vendor to develop a combination of machine learning programs and of business process rules from which we can flag cases based on these characteristics. The project lead will then work with ADAs in SFDA's Intake Unit and specialized vertical prosecution units to test and refine the deployment of these case flags.



Project Timeline and Milestones:

The SFDA *Data Science for Prosecutorial Impact* (DSPI) proposes the following goals, objectives, and milestones, outlined in the Chart below, to successfully implement this project.

Goal	Objective	Milestones
Phase 1: Months 1-4		
1. Lay groundwork for successful project implementation	<ul style="list-style-type: none"> Objective 1.1. Complete City and County of San Francisco Accept and Expend Process allowing for release of funds Objective 1.2. Hire Project Director 	<ul style="list-style-type: none"> Project director job description approved by SFDA HR Project director hired
Phase 2: Months 5-10		
2. Reduce SFDA caseloads by increasing identification of cases for discharge or diversion at initial intake.	<ul style="list-style-type: none"> Objective 2.1. Establish common characteristics of cases for discharge. Objective 2.2. Establish common characteristics of cases eligible for SFDA pretrial diversion programs and/or San Francisco collaborative courts. 	<ul style="list-style-type: none"> Review of best practices in prosecution and discharge completed. Interviews conducted with intake unit ADAs and case carrying ADAs re: commonly discharged cases. Primary discharge categories identified, such as insufficient evidence; interest of justice; pretextual stop. Established list of case characteristics for primary categories of discharged cases. Review of best practices in pretrial and collaborative court diversion. Interviews conducted with diversion unit ADAs and program partners. List of case and defendant characteristics for each diversion program.
Phase 3: Months 11-15		
3. Improve identification of cases involving harder-to-identify characteristics that pose a high risk to public safety and	<ul style="list-style-type: none"> Objective 3.1. Establish common characteristics of cases involving human trafficking, including commercial sexual 	<ul style="list-style-type: none"> Review of best practices in human trafficking identification and prosecution. Interviews conducted with specialized unit ADAs and Victim Service Division staff. Interviews conducted with experts in human trafficking.



San Francisco District Attorney's Office

Goal	Objective	Milestones
<p>community well-being.</p>	<p>exploitation and labor trafficking.</p> <ul style="list-style-type: none"> • Objective 3.2. Establish common characteristics of domestic violence cases with a high risk for escalation, especially lethality. 	<ul style="list-style-type: none"> • List of case and defendant characteristics indicating high likelihood of human trafficking nexus. • Review of best practices in high-lethality domestic violence identification and prosecution. • Interviews conducted with specialized unit ADAs and Victim Service Division staff. • Interviews conducted with experts in domestic violence, especially high lethality domestic violence. • List of case and defendant characteristics indicating high likelihood of severe domestic violence cases, especially high lethality risk.
<p>Phase 4: Months 16-24</p>		
<p>4. Use data science, machine learning, and IT business rules to flag cases for likely discharge, diversion, or additional attention at Intake.</p>	<ul style="list-style-type: none"> • Objective 4.1. Match relevant case characteristics to available data elements in SFDA case management system. • Objective 4.2. Develop scripts to flag characteristics from text-based materials • Objective 4.3. Develop algorithms to flag cases for special review at intake based on relevant characteristics. • Objective 4.4. Test, algorithms. • Objective 4.5. Refine and redeploy algorithms. 	<ul style="list-style-type: none"> • Initial map of data elements/values available in SFDA case management system. • Scripts to pull data from text-based materials. • Preliminary diversion algorithm. • Preliminary discharge algorithm. • Preliminary high safety risk case algorithm. • Tests of each algorithm. • Revised and deployed algorithms.



Key Participating Staff

Mikaela Rabinowitz, PhD. | SFDA Director of Data, Research, and Analytics

This project will be led by Dr. Mikaela Rabinowitz, SFDA's Director of Data, Research, and Analytics. As project lead, Dr. Rabinowitz will oversee all aspects of project planning and implementation, working with a to-be-hired research associate to collect and analyze the qualitative and quantitative data necessary to identify characteristics for different case processing paths and then with a data engineer and with SFDA IT staff to coordinate the development and implementation of appropriate algorithms.

Dr. Rabinowitz brings 15 years of experience in using data to inform criminal justice decisions and worked in a number of criminal justice research and advocacy positions prior to joining SFDA. In her role at SFDA, Dr. Rabinowitz oversees all aspects of data collection, processing, and analysis, including working closely with SFDA's IT Department to plan for the implementation of the office's new case management system.

Beth Munger | SFDA Principal IT Business Analyst

Beth Munger, SFDA's Principle IT Business Analyst, will work closely with Dr. Rabinowitz and a data engineer to support the integration of business rules and algorithms to flag cases. As the project manager for SFDA's new case management system, she will also act as the liaison between the case management system vendor and SFDA staff during the implementation of this project.

Ms. Munger has been an IT Business Analyst with SFDA for more than six years, during which time she has served as the single point of contact for all technical implementations and business streamlining opportunities. In this role, she has partnered with attorneys, analysts, and support staff to understand the existing business environment, identify opportunities for streamlining, and facilitate the transition to new business processes. Prior to her work at SFDA, Ms. Munger was a product manager and senior business analyst in San Francisco's Human Services Agency for more than a decade. She holds a Business Analyst Certification from George Washington University and a Project Management Professional (PMP) certification from UC Berkeley.