Algorithms in Policing: An Investigative Packet

Public agencies increasingly rely on algorithms to solve complex problems. While algorithms can be beneficial, they can also have significant side effects. One of the most important and controversial examples is policing algorithms. Policing algorithms are intended to help prevent or solve crime, but are well known to exacerbate biases in the criminal justice system. There are also serious worries about their efficacy and accountability.

With Sunshine Week on the horizon, now is the time for the press to delve deeply into policing algorithms in Connecticut. To help journalists in this effort, the Media Freedom & Information Access Clinic at Yale Law School ("MFIA") has prepared this investigative packet. The packet contains the following primer on policing algorithms, a template Freedom of Information request for records concerning policing algorithms, and a list of sample questions for journalists to ask government officials. MFIA has also prepared an extensive report on algorithmic accountability in Connecticut (the "MFIA Algorithmic Accountability Report") that will provide additional helpful background, available here: https://law.yale.edu/mfia/projects/government-accountability/algorithmic-accountability.

What are Algorithms? Why are They Used?

Algorithms are sets of instructions used to solve a problem. They may be simple, like instructions for ordering a list alphabetically, or complicated, like those that calculate outputs using thousands of variables and inputs. They are the heart of computing and are involved in everything from Google's driving directions to Netflix's movie recommendations. Increasingly, governments have turned to algorithms to solve public problems. For more details, see the MFIA Algorithmic Accountability Report.

The term "algorithm" encompasses a broad range of tools. Artificial intelligence ("AI") based algorithms are the cutting edge of modern algorithms and are at the heart of the most sophisticated algorithms being developed today. Artificial intelligence systems try to mimic human intelligence and can constantly improve themselves using the information they collect. They are novel in that they teach themselves how to get from the input to the output.

Artificial intelligence algorithms are not inherently good or bad. Like any tool, their value depends on their use. They can streamline mundane processes and make existing systems more efficient. In some ways, an algorithm may be significantly better than its human counterpart in accomplishing the same task. This is especially pronounced in big data analysis, where AI based algorithms can effectively and efficiently find patterns across thousands of variables.

AI based algorithms are frequently in the news, sometimes for humorous results (like the AI that misidentified a picture of a cat as guacamole). While this AI might be funny, other AI based algorithms have perpetuated discriminatory practices in housing, healthcare, and hiring, despite the programmers' intentions. To ensure that governments treat people equitably as they turn to algorithms, it is important to investigate their use and hold those using them accountable.

How Do Police Use Algorithms?

Police departments often use AI based algorithms in an effort to improve policing practices. While this goal is laudable, the algorithms they use—normally bought from private companies—have unintended drawbacks. One main drawback is bias, which occurs because the criminal justice data used to develop the algorithms are themselves biased against minority and marginalized communities. This entrenches discriminatory or disparate policing practices. Further, several cities have abandoned costly policing algorithms because they didn't work, and others have not told the public when algorithms go live. These issues illustrate the need for rigorous public oversight. We can see them by examining how agencies have used three common types of policing algorithms: predictive policing algorithms, facial recognition algorithms, and pattern recognition programs.

First, take predictive policing algorithms, which departments use to help predict and prevent future crime. Location-based predictive policing algorithms, the most common type, use existing crime data to identify areas and times with higher risk of crime. Proponents argue that they help police departments better allocate resources.

However, studies have shown that predictive policing algorithms often generate biased results from data skewed against marginalized communities and people of color. For example, one of the most common predictive policing algorithms, PredPol, uses an AI algorithm to predict areas of higher crime. A 2016 study found that PredPol led police departments to patrol already-overpoliced communities even more.[1] A 2018 study found that, if the algorithm were applied in Indianapolis, Latino and Black communities would have experienced, respectively, 200%-400% and 150%-250% greater patrol presence than white communities.[2] Many similar algorithms share these patterns.

Predictive policing algorithms can also be expensive even though many have proven to be largely ineffective. For example, Chicago abandoned its $2 million predictive policing algorithm because it did not lead to positive results. Palo Alto cancelled a contract for a predictive policing algorithm when it failed to reduce or solve crime. Los Angeles stopped using predictive policing algorithms when the police department's inspector general could not determine that the software reduced crime after a decade of use. For more details, see the MFIA Algorithmic Accountability Report at page 4.

Second, some police departments have recently begun using facial recognition technology to help solve crimes. This, too, is fraught with many of the same issues—indeed, the UN Committee on the Elimination of Racial Discrimination found that facial recognition and other policing algorithms risk deepening racism and xenophobia. In Detroit, a facial recognition program misidentified suspects approximately 96% of the time and led to the wrongful arrests of several Black residents.[3] Tests have also shown that facial recognition has an uneven performance across different races. For example, one found that facial recognition used in police

[1] https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2016.00960.x
[2] https://ieeexplore.ieee.org/abstract/document/8616417
[3] See MFIA Algorithmic Accountability Report at page 5.

investigations produced more false positive results when evaluating Black women.[4] Concerns over privacy and false matches led some cities, like San Francisco, to bar law enforcement from using facial recognition technology.

Third, pattern recognition programs, like the NYPD's Patternizr, sift through police data to find connections and patterns between crimes. The NYPD has tried to address concerns of racial bias by highlighting that the algorithm does not consider attributes like race, gender, and precise location. However, this is consistent with racially disparate outcomes: one study of Patternizr shows that the algorithm might be compounding implicit biases and could lead to innocent people being incarcerated.[5] And yet Patternizr has evaded scrutiny because the NYPD used it for three years (starting in 2016) before disclosing it to the public.

These are only three examples of algorithms used by police departments, but they demonstrate why algorithms should be treated with caution and subjected to public oversight. Algorithms, if used correctly, might improve law enforcement. However, the press and the public should closely scrutinize law enforcement algorithms to ensure they work well and do not discriminate against marginalized communities.

What Can Be Done?

One of the most crucial elements of effective algorithmic governance is transparency. To ensure that law enforcement and policy makers are held accountable, it is important to understand if police departments use algorithms and, if so, how. Asking the right questions and using state Freedom of Information laws are important ways to get the information needed to evaluate these programs.

We intend for the attached sample questions and template Freedom of Information request to help with this endeavor, and encourage journalists to use them in whatever ways will be most productive. The questions and template request include a large number of topics, and journalists may choose to focus on a subset. We also encourage journalists to review Part 2 of the MFIA Algorithmic Accountability Report, which describes our use of the FOI process to obtain information about Connecticut state agencies' use of algorithms outside the policing context.

For additional information, please contact MFIA students Paul Meosky (paul.meosky@ylsclinics.org) and Sruthi Venkatachalam (sruthi.venkatachalam@ylsclinics.org).

[4] https://www.nytimes.com/2020/01/12/technology/facial-recognition-police.html
[5] https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=2779&context=ulj

[Date]

[Name of Public Information Officer]
Public Information Officer
[XXXX] Police Department
[street address]
[city, CT zip]

Dear [Name],

This FOIA request is submitted on behalf of [News Organization]. In accordance with the Freedom of Information Act, Conn. Gen. Stat. §§1-200 et seq., we request copies of the following records pertaining to the XXX Police Department's use of algorithms to prevent or solve crimes.

Definitions

For the purposes of this request, the following terms shall have the following meanings:

"Algorithm" means a specific sequence of instructions, rules, or order of operations used to cause a technical tool or system to execute a set of actions, whether or not it is subject to human input in executing that set of actions.

"Policing Automated Decision-making System" means any algorithm that uses data-based analytics to help prevent or solve crime, including but not limited to one incorporating machine learning or other artificial intelligence techniques.

"Disparate impact" means any impact, including but not limited to distributional or equitable impact, that disproportionately affects individuals based upon their race, national origin, ethnicity, sex, gender identity, sexual orientation, religion, or socioeconomic status.

"Predictive Policing" means any use of algorithms to analyze datasets in order to predict and help prevent future crimes.

"Predictive Policing Product" means any tool that implements Predictive Policing. They include but are not limited to both location-based tools, which predict where and when crime is likely to occur, and person-based tools, which predict who is likely to be involved in future criminal activity. PredPol and COMPAS are examples of Predictive Policing Products.

"Facial Recognition Program" means any tool meant to identify individuals using images of faces. Clearview AI is an example of a Facial Recognition Program.

"Pattern Recognition Program" means any tool that uses data to find patterns or similarities to help solve crimes that have been committed. New York City's Patternizr is an example of a Pattern Recognition Program.

Documents Requested

1. All documents relating to the procurement or development of any Predictive Policing Product, Facial Recognition Program, Pattern Recognition Program, or other Policing Automated Decision-making System, including but not limited to any Requests for Proposals or communications with Clearview AI or PredPol.

2. All agreements between the Department and any entity or entities for the acquisition of a Predictive Policing Product, Facial Recognition Program, Pattern Recognition Program, or other Policing Automated Decision-making System.

3. Documents sufficient to disclose each category of data collected or used in connection with any Predictive Policing Product, Facial Recognition Program, Pattern Recognition Program, or other Policing Automated Decision-making System.

4. The source code of any Predictive Policing Product, Facial Recognition Program, Pattern Recognition Program, or other Policing Automated Decision-making System.

5. All documents evaluating the predictive accuracy of policing outcomes resulting from the use of any Predictive Policing Product, Facial Recognition Program, Pattern Recognition Program, or other Policing Automated Decision-making System.

6. All documents assessing the relative improvement and/or decline in policing success under the Department's use of any Predictive Policing Product, Facial Recognition Program, Pattern Recognition Program, or other Policing Automated Decision-making System.

7. All documents describing any disparate impact potentially or actually caused by the Department's use of Predictive Policing Product, Facial Recognition Program, Pattern Recognition Program, or other Policing Automated Decision-making System.

8. All training materials, including but not limited to manuals or handbooks, pertaining to the use or implementation of any Predictive Policing Product, Facial Recognition Program, Pattern Recognition Program, or other Policing Automated Decision-making System.

9. All correspondence between the Department and its vendors referring to the purchase, implementation, or use of any Predictive Policing Product, Facial Recognition Program, Pattern Recognition Program, or other Policing Automated Decision-making System.

Because FOIA governs information recorded "by any . . . method," Conn. Gen. Stat. § 1-200(5), your search for records should include electronic as well as tangible sources, that is, all records or communications preserved in electronic or written form, including but not limited to correspondence, interoffice memoranda, intraoffice memoranda, documents, data, videotapes, audio tapes, mails, faxes, files, guidance, guidelines, evaluations, databases, instructions, analyses, memoranda, agreements, notes, order, policies, procedures, protocols, reports, rules, technical manuals, technical specifications, training manuals, or studies.

We have addressed this request to you in the belief that you are the custodian of such records. If you are not, please forward this request to the proper custodian of such documents and inform us of who the proper custodian is.

We request a waiver of any fees for searching or copying these records because disclosure of the requested information is in the public interest and will contribute significantly to the public's understanding of Connecticut's use of predictive analytics or automated decision making in policing, and this information is not being sought for commercial purposes. If you do not grant a waiver, please inform us if the fees will exceed $100, before incurring them.

The Connecticut Freedom of Information Act requires a response within four business days. If access to the records we are requesting will take longer, please let us know when we can expect to receive copies or be permitted to inspect the requested records.

If you deny any or all of this request, please cite each specific exemption you feel justifies the refusal to release the information and notify us of the appeal procedures available to us under the law. If an otherwise public record has a portion that is exempt from disclosure, you should redact the exempt portion and release a copy of the rest of the record together with a notation identifying the specific exemption that you believe applies to the portion withheld. If you have questions about this request, please feel free to contact us.

If any requested record no longer exists, we request a copy of the destruction order. Further, under Records Retention Schedule #16-7-1R (which states that "records subject to pending or active Freedom of Information Act (FOIA) requests . . . may not be destroyed"), if any record is subject to destruction but has not yet been destroyed, it may not be destroyed upon receipt of this FOIA request.

Thank you for your attention to our request. Given the current disruptions due to COVID-19, we would appreciate if your responses were sent to [email address].

Sincerely,

[Name]
[Organization]
[Mailing Address]
[City, ST zip]
Tel: [telephone number]
Email: [email address]

Investigative Guide on Algorithms and Policing

The following sample questions are provided to start conversations between journalists and public officials regarding the use of policing algorithms in their community. The first set of questions is directed towards police departments, and the second set of questions is directed towards local representatives, such as state senators, mayors, or council members.

Questions for Police Departments

1. Does your department currently use any algorithms to assist in either (a) identifying criminals, (b) solving crimes, or (c) allocating resources to prevent future crimes?

IF "no"…

2. Are you considering the future use of such algorithmic policing tools? For what purposes?
3. How will you involve the public in the procurement of policing AI tools? What information will you make available to the public about your procurement and use of such algorithmic tools?
4. Public oversight for policing algorithms is often frustrated when private companies provide those algorithms and refuse to share the underlying source code. Would you require private companies to waive trade secret protections before using their AI products?

IF "yes"...

2. When you interact with a member of the public as a result of a policing AI tool—for example, when a policing algorithm identifies a high-risk suspect—do you inform them of that fact?
3. Protections for commercial information like trade secrets stop the public from knowing exactly how their data is being used to make policing decisions. How can the public rest assured that commercial tools like PredPol use citizens' data responsibly to produce just outcomes?
4. The bias in certain algorithms developed for use in the Criminal Justice System is well-established. How do you control for potential bias in the algorithms currently used?
5. Citizens are increasingly concerned about the disparate impact of predictive policing algorithms like PredPol on minority communities. What have you done to address these concerns?
6. Feeding bad data into an algorithm produces bad results. What data is used by the algorithms currently used by your Department? 911 calls? Officer reports? Arrests? Convictions? What evidence do you have that using this data produces meaningful, unbiased information?
7. Has your department ever shared criminal or non-criminal data with any private firm, such as Palantir? Would your department ever do so to obtain access to algorithmic policing tools? What protections are in place to prevent private companies from misusing data provided by your department?

8. Los Angeles, Palo Alto, and Chicago have all tried predictive policing algorithms and found them ineffective and expensive. How does your department independently verify the accuracy, reliability and value of such tools? How much does your department pay to use each AI tool? And how do you justify the cost?
9. What training do you provide your officers to ensure they use AI policing tools effectively and fairly?
10. All policing methods make mistakes sometimes, but those mistakes can erode public trust in law enforcement. What measures can your department take to maintain the department's legitimacy when policing algorithms fail? Particularly among minority and impoverished communities?
11. We all know that there are many ways for a department to acquire new tools and programs. Does every AI tool used by your department go through the public procurement process? If not, why did you exclude the public?
12. Where can the public learn more about the tools your department uses now or may consider using in the future?
13. *See Table 1 below before asking this question*. Last year, BuzzFeed released a database listing taxpayer-funded entities, including local police departments, that had made at least one facial recognition scan on Clearview AI as of February 2020. Your department was on the list. What were you using Clearview AI for and are you still working with Clearview AI or any other facial recognition system?

Questions for Government Leaders

1. Hamden's Legislative Council voted unanimously to ban all use of facial recognition by the town and its officials. Are you considering any similar bans against facial recognition or other policing AI?
2. Does your town currently pay for any policing AI tools?
3. Los Angeles, Palo Alto, and Chicago have all tried predictive policing algorithms and found them ineffective and expensive. In 2020, Santa Cruz became the first city to ban predictive policing outright. Would you endorse a similar measure? If not, how would you justify the cost of using such algorithms? If so, how might the town spend the money more effectively?
4. Reporters and researchers seeking more information about how artificial intelligence informs government decision-making have faced resistance. What measures are you considering to ensure democratic oversight for the algorithms our agencies use to make life-changing decisions?
5. The Data Accountability and Transparency Act, as proposed in Ohio, would require agencies to assess each algorithm for accuracy and bias before acquiring the algorithm and every year thereafter. The agencies would also be required to recommend measures to minimize inaccuracies and bias. Would you support similar legislation here? Or is there an alternative model you would support?
6. The value of machine learning algorithms is often the ability to draw patterns that humans would not see themselves. Given the "black box" nature of these algorithms, how

can the public trust that policing AI uses personal data responsibly to guarantee just outcomes?

7. Protections for commercial information like trade secrets present major obstacles to public oversight and government transparency. How can your government guarantee the accuracy and value of these commercial tools? Would you limit trade secret protections for such algorithms? And is there any plan for developing a publicly-sourced alternative?
8. Even when models are accurate, humans can misuse the results. What measures can we take to ensure policing algorithms aren't arbitrarily applied to target certain groups and individuals? And how do we hold the humans accountable when they try to blame the machine?
9. Would you support legislation requiring police departments to post information about their algorithms on their websites?
10. Some police departments have made backroom deals with companies like Palantir trading free access to criminal and non-criminal data in exchange for AI tools. By classifying the deal as "philanthropic" or "research," departments can evade public procurement processes. What steps are you taking to ensure the public is involved in all decisions relating to the acquisition and implementation of policing AI?

Table 1. Buzzfeed Record of all Clearview AI Requests through February 2020

Name	Number of Face Searches in Clearview AI	Response
Avon Police Department	11-50	"[Officers] learned of it through a training class." – James Rio, director of police services
Branford Police Department	11-50	No response
Bridgeport Police Department	101-500	No response
Bristol Police Department	11-50	No response
Connecticut Intelligence Center	11-50	No response
Connecticut Office of Adult Probation Services	51-100	No response
Dansbury Police Department	101-500	"We have not used Clearview AI" – Mark Williams, Detective
East Haven Police Department	51-100	No response
Enfield Police Department	11-50	No response
Farmington Police Department	1-5	"We do not have facial recognition technology and I don't know of any Farmington uses" – Paul J. Melanson, Police Chief
Glastonbury Police Department	11-50	No response
Greenwich Police Department	101-500	No response
Guilford Police Department	11-50	No response
Hamden Police Department	11-50	No response
Hartford Police Department	101-500	No response
Madison Police Department	11-50	No response
Manchester Police Department	11-50	No response
Mashantucket Pequot Tribal Police Department	51-100	No response
Middletown Police Department	11-50	No response

Naugatuck Police Department	6-10	"I am sorry to inform you that our department merely scheduled a meeting to get a product demo with Clearview AI and decided against purchasing the software at this time due to budgetary reasons." —Lt. Antonio Bastos
New Canaan Police Department	11-50	No response
New Haven Police Department	11-50	No response
Newington Police Department	11-50	No response
North Branford Police Department	11-50	No response
North Haven Police Department	1-5	No response
Southern Connecticut State University	101-500	No response
Stamford Police Department	101-500	"In July 2019 a victim was beaten unconscious and robbed by a male escort. Investigators used Clearview facial rec on the image from the internet escort profile of the suspect. Clearview returned a possible match from an image of a Washington Police Metro wanted poster for an unarmed kidnapping. The wanted flyer had a photo of their suspect whose identity was unknown at the time of its distribution. Stamford investigators contacted Washington Metro and learned that they had since identified their suspect." —Thomas Wuennemann, assistant police chief
State of Connecticut	501-1000	No response
Stratford Police Department	11-50	No response
Torrington Police Department	101-500	"We had purchased the tool in the beginning of the year, however with the pandemic and everyone wearing face masks we have not had an opportunity to use it and therefore will most likely not be renewing the contract in the upcoming year." —Lt. Brett Johnson
Waterbury Police Department	11-50	No response
West Hartford Police Department	51-100	No response
West Haven Police Department	6-10	No response

Westport Police Department	1-5	No response
Wethersfield Division of Police	11-50	"I understand some of my officers have tried the software with good results but I have not authorized the purchase of any software of this nature." —James Cetran, police chief
Wilton Police Department	51-100	No response