Ethics of algorithms

Numbers for Policy: Practical problems in quantifications

Alexandra Theben, Open Evidence Research, Universitat Oberta de Catalunya

Algorithms decide upon an ever-increasing list of cases, such as recruiting, carriers - including of researchers, prison sentencing, paroling, custody of minors…

Brauneis, R. & Goodman, E. P. Algorithmic Transparency for the Smart City, Yale Journal of Law & Technology (2017),
Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3012499 (Accessed: 30th August 2017).

[As a **society**, we are now at a crucial juncture in determining how to deploy AI based technology in ways that promote, not hinder democratic values such as freedom, equality and transparency.]

Stanford University, One Hundred Year Study on Artificial Intelligence (AI100), August 1, 2016, https://ai100.stanford.edu.

Algorithms use by governments

AI and Big Data Analytics increasingly replace human decision-making. With ever greater frequency, governments are using computer algorithms to conduct public affairs.

Governments also have access to oceans of data. Algorithms can decide where kids go to school, how often garbage is picked up, which police precincts get the most officers, where building code inspections should be targeted, and even what metrics are used to rate a teacher (NY times).

(https://www.nytimes.com/2017/08/24/nyregion/showing-the-algorithms-behind-new-york-city-services.html)

[…] The Smart city movement worldwide impresses on local governments the importance of gathering and deploying data more effectively […]
(Brauneiss & Goodman).

BUT:
Most local governments lack the expertise to deploy data analytics on their own.

So if they want to be **smart**… they need to contract with companies, universities, and non-profits to implement privately developed algorithmic processes.

The role of private companies

[…] The result is that privately developed predictive algorithms are shaping local government actions in such areas as criminal justice, food safety, social services and transportation […]

Black box opacity

[…]Because the designing entities (private companies etc.) typically do not disclose their predictive models or algorithms, there is increased criticism of this "**black box**" opacity of these processes […] (O'Neil, 2016)

[…]In the public sector, the opacity of algorithmic decision-making is particularly problematic both because government decisions may be especially weighty, and because democratically elected governments bear special duties of accountability[…] (Brauneis & Goodman)

At the same time there are special concerns when municipal and other governments use predictive algorithms whose development and implementation neither the public nor the government *really understands* (!).

[⋯] The risk is that the opacity of the algorithm enables **corporates** capture public power, while when a government agent implements an algorithmic recommendation that he or she does not understand, the government has lost democratic accountability[⋯].

Study by Brauneis & Goodman

Robert Brauneis and Ellen P. Goodman set out to test the limits of transparency around government deployment of big data analytics, focussing on local and state government's use of predictive algorithms.

They filed 42 open record requests in 23 U.S. states seeking essential information about 6 predictive algorithm programs (identified as most common local government uses of big data prediction) developed by for-profit companies, nonprofits, and academic/private sector partnerships.

The goal was to see if they could discover what policy judgment these algorithms embody, and could evaluate their utility and fairness.

Algorithm transparency

They found that in almost every case, meaningful "algorithmic" transparency was not provided

(over-broad assertions of **trade secret** was a problem).

In this session we take a look at how they came to this conclusion …

Why do governments use algorithms?

The use of Big Data and predictive algorithms is a form of **governance** – it is a way for authorities to manage individual behaviour and allocate resources.

- Government deployment of algorithmic processes promises increased **efficiency** and fairness in the delivery of government services.

- Algorithmically informed decision-making can also help government officials **avoid biases**, explicit or implicit, that may creep into less formal, "hunch"-based decision making (decisions based on cultural or ethnic generalisations).

Implementation of algorithms at local level is part of a broader movement towards data-driven decision making and must be understood in the context of the smart city agenda.

Smart Cities en Marche

Smart city initiatives seek to harness data to rationalize and automate the operation of public services and infrastructure, such as transportation, energy, and health services…

BUT: They cannot do it without public-private partnerships, which develop the analytics and ensuing "smart" systems.

So, when algorithms are deployed in the public sphere, public authority typically yields to the private control of technology companies and other developers.

Barcelona Smart City

Barcelona has long been a leader in the smart city movement. Sometimes ranked number one – and usually in the Top 10 – it is part of an elite group of intelligent urban-planning pioneers along with such cities as Singapore, Vienna, San Francisco, and Copenhagen. Now, it is also one of a handful of smart cities trying to integrate the top-down and the bottom-up approach to urban digitalization, and boldly reach for what some are calling Smart City 3.0

According to the digital chief of Barcelona, cities can "end up with black box operating system where the city itself loses control of critical information and data that should be used to make better decisions" (Quote by Francesca Bria).

http://www.urban-hub.com/cities/smart-city-3-0-ask-barcelona-about-the-next-generation-of-smart-cities/

The role of private entities

Private entities have been at the leading edge of the entire smart city movement

It is to these companies and other private vendors that that local government officials, pressed by economic necessity, and personnel constraints, will often leave the work of data analytics.

The fear is that smart city partnerships will ultimately lead to the surrender of public services to private interests.

- *Governments are dependent on the technology provided by private companies*
- *Private companies come to own critical data*

What the public needs to know

Algorithmic governance has a politics. When private vendors control algorithmic governance, the politics of algorithms recede behind private hedges.

BUT: Judgments are encoded in the algorithmic process at all stages.

“A predictive algorithm’s recommendation actually masks an underlying series of subjective judgments on the part of the system designers about what data to use, include or exclude, how to weight the data, and what information to emphasize or deemphasize “ (Surden, 2017).

What the public needs to know

There is a strong public interest in ensuring that predictive algorithms are designed and executed justly, especially when they impact individuals.

By nature, predictive models are simplifications which do not take into account all possible relevant factors about subjects, and they therefore treat people as members of groups, not as individuals.

- Does fairness matter to the algorithm developer?
- For sensitive decisions, descision-makers (judges, social workers) are expected to exercise human judgment over algorithmic predictions

In theory, the algorithmic edict is advisory only

In practice, decision-makers place heavy reliance on the numbers, raising the stakes for their fairness.

Questions of Fairness of algorithms

The most discussed algorithm fairness question have been whether predictive algorithms are likely to introduce or perpetuate **discrimination** on the basis of race, gender, or another characteristic.

A classic example is the early Google facial recognition algorithm. It was trained on the faces familiar to the engineers who built it, which were mostly white. As a result, the program classified white-skinned human faces as human, but often classified dark-skinned human faces as animal.

Diminishing decision-making capacities of public servants

There is a further danger that algorithmic decision-making will hollow out the decision-making capacity of public servants

- The "machine says" so
- When the algorithmic output is uninterpretable – when the decision path is not explained – government officials have no way of knowing whether and how the factors they are facing accord with the factors that produced the algorithmic recommendation.
- If the algorithm is opaque, the government official cannot know how to integrate its reasoning with their own, and must either disregard it, or follow it blindly
- Over time, deference to algorithms may weaken the decision-making capacity of government officials along with their sense of engagement and agency (deskilling of human beings through automation, Carr, 2014).

Transparency

It will be possible to assess a predictive algorithm's politics, performance, fairness, and relationship to governance only with significant transparency about how the algorithm works.

Of course: there has always been risk of inefficacious or biased decisionmaking by government agents.

But predictive algorithms pose new risks of unfairness and error even if they improve overall decision making because of **scalability:** Predictive algorithms are typically used to guide decisions throughout a governmental unit and even across many local and state governments

The ability of these algorithmic processes to scale, and therefore to influence decisions uniformly and comprehensively, magnifies any error or bias that they embody, and increases the importance of rendering them transparent.

The challenge is to specify a degree and form of transparency that is meaningful for the public and practical for developers and governments.

Transparency

- Algorithms should be capable of disclosure in some combination of mathematical and logical notation and natural language.

- Access to the underlying data or at least descriptions to help us understand how strong the purported correlations actually are, what the sample size was, and other matters that affect statistical validity.

- The public purpose for which the algorithm was developed, the contract terms that govern data ownership and access, and plans for validation and follow-up.

But even if all this information was provided, it may be difficult to understand the results of the algorithmic process

- Difficult to understand whether the algorithm correlates with our sense of fairness

- Difficult for government officials to assess the algorithmic output in light of their own sense of a situation

Algorithmic accountability in the public sphere requires that government actually be held accountable for the algorithms it deploys.

RESULTS OF OPEN RECORDS REQUESTS FOR ALGORITHMS by Brauneis and Goodman

They filed open records requests (in absence of "push" transparency as codified under the Federal Freedom of Information Act, FOIA*) covering six different programs featuring predictive algorithms.

The six programmes are:

- Public Safety Assessment;
- Eckerd Rapid Safety Feedback;
- Allegheny Family Screening Tool;
- PredPol;
- HunchLab;
- New York City Value-Added Measures.

*Governments should reveal the relevant structures, logic and policies of the algorithms voluntarily from the outset (Push method). Open records requests are "pull" requests.

Public Safety Assessment

Public Safety Assessment (PSA) is a pre-trial risk assessment tool developed by the Laura and John Arnold Foundation, designed to assist judges in deciding whether to detain or release a defendant before trial.

PSA includes three different risk assessment algorithms, which are intended to assess the risks that a released defendant will, respectively, fail to appear for trial; commit a crime while on release; and commit a violent crime while on release.

Eckerd Rapid Safety Feedback

Eckerd Rapid Safety Feedback (RSF) is a risk assessment process designed to identify child welfare cases with a high probability of serious child injury or death.

RSF was developed by Eckerd Kids (Eckerd), a nonprofit family and child services organization, and Mindshare Technology, a for-profit software company.

Eckerd identified the greatest risk factors contributing to child injury or death, namely, "a child under the age of three, a paramour in the home, substance abuse, and domestic violence history, and a parent who had previously been placed in foster care."

Allegheny Family Screening Tool

Like Eckerd RSF, The Allegheny Family Screening Tool (AFST) was developed to facilitate the sighting of child welfare cases. AFST was developed by a consortium led by the Centre for Data Analytics at the Auckland University of Technology (the "Auckland Consortium"), in cooperation with the Allegheny County Department of Human Services.

While Eckerd RSF is apparently used on an ongoing basis to monitor cases within the child welfare system, AFST is applied at the time an initial call is made to report child maltreatment.

It assists in determining whether the report warrants a formal investigation.

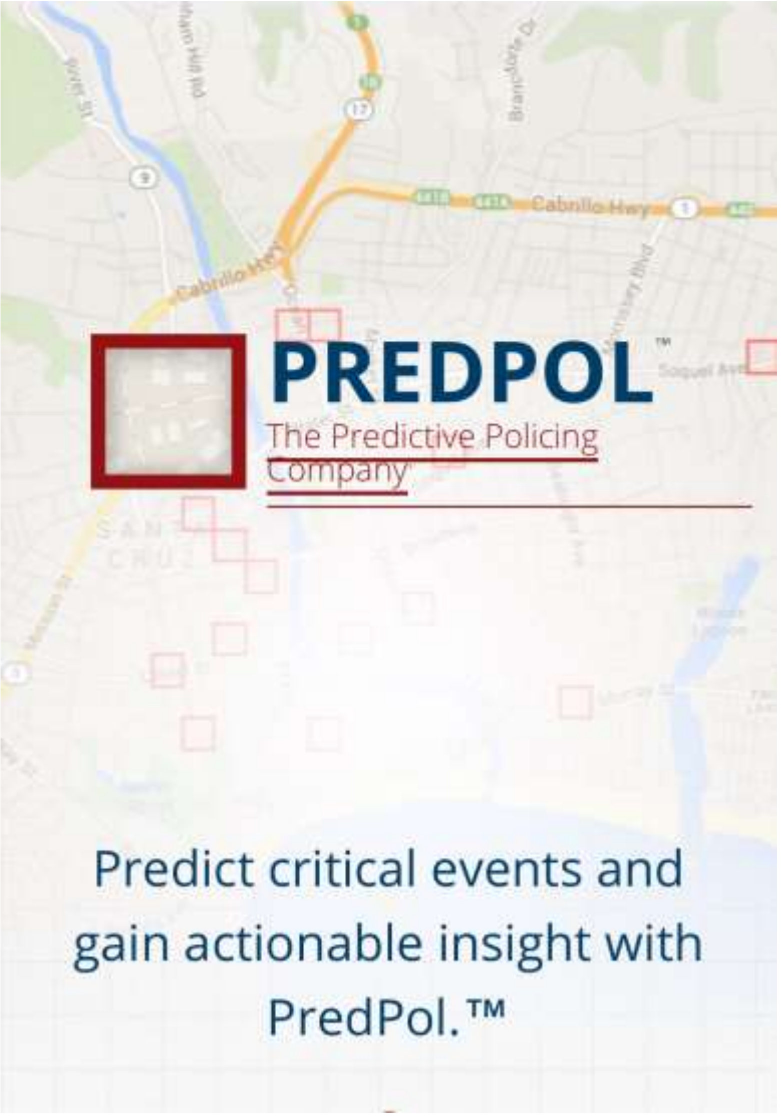
Public PredPol– Predictive Policing Assessment

PredPol is software that predicts where and when crimes of various types are likely to occur, and thus assists police forces in plotting their patrols to deter those crimes.

It was originally developed by mathematicians and behavioral scientists from UCLA and Santa Clara University in collaboration with crime analysts and officers from the Los Angeles and Santa Cruz Police Departments, but is now managed by a for-profit company, PredPol Inc.

The creators of PredPol determined that the three most important types of information or "data points" for predicting crime are crime type, crime location, and crime date and time.

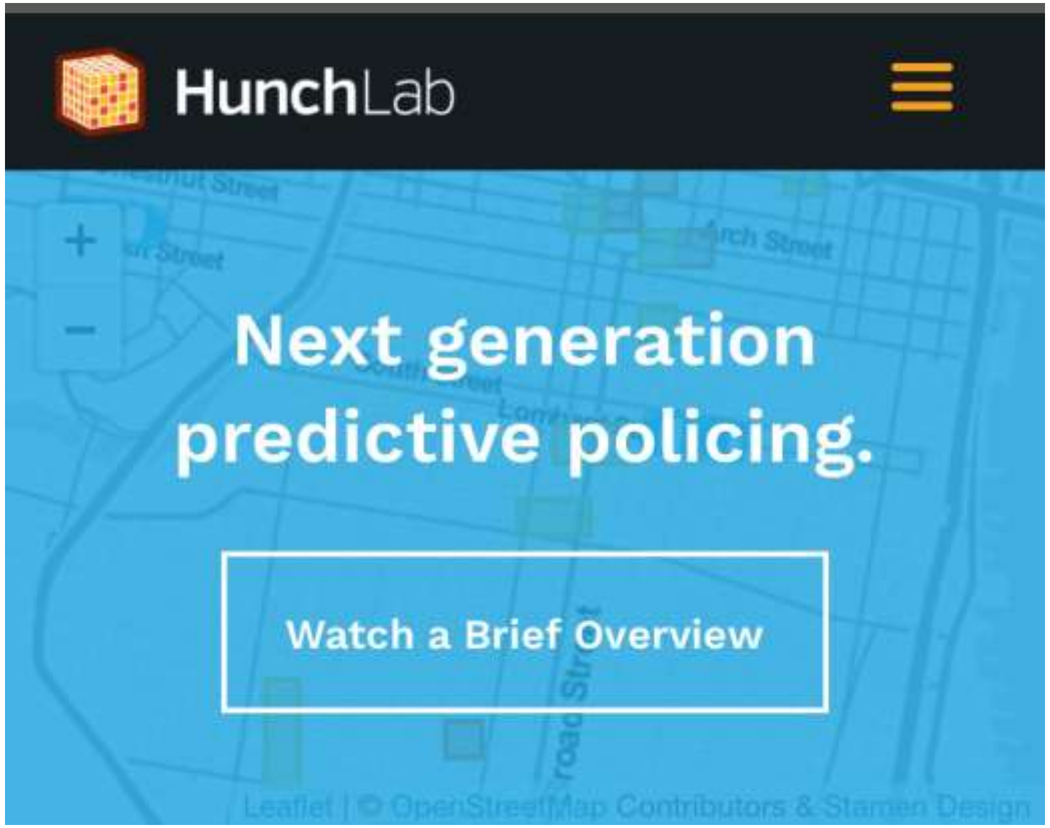
PredPol feeds data about past patterns of criminal activity into an algorithm that predicts where and when new crimes will be committed.



HunchLab

Like PredPol, HunchLab is software that predicts where and when crime will occur, with a cartographic output indicating areas at higher risk for certain types of crimes over certain time periods.

HunchLab is developed and maintained by Azavea, Inc., a for-profit corporation.
HunchLab uses a wide range of inputs to predict risks of crime, and allows individual police departments to prioritize for selected crimes.



New York City Value-Added Measures

New York City and the State of New York are among the jurisdictions that have adopted a Value Added Model (“VAM”) method for evaluating teachers.

Value Added Model algorithms compare test scores of students at the beginning and end of a given year in order to measure the progress of those students.

Those results are then adjusted to try to account for factors other than teacher effectiveness, such as socioeconomic status, that might be responsible for the students’ progress or lack thereof.

The adjusted results for the students that are taught by a particular teacher are then used to produce an evaluation of that teacher’s effectiveness

Show picture of NY times “Showing Algorithms Behind New York City Services”

In New York, where algorithms are used by the administration for a large array of decisions, the mayor has decided to pursue legislation for “algorithmic audits”.

The New York Times

Showing the Algorithms Behind New York City Services

About New York
By JIM DWYER AUG. 24, 2017

3

Let us say that James Vacca is not necessarily the first person you’d think would begin a deeply necessary revolution to peel away some of the secrecy around technology that shapes government decisions. In the 1980s, Mr. Vacca admitted, he told an aide that it would be a waste of money to replace office typewriters with

About New York
Twice a week, a chronicle of New York and New Yorkers.

Online Chats Tie Politicians to the Troll of Staten Island SEP 7

Dwyer J. Showing the Algorithms Behind New York City Services - The New York Times. New York Times Aug. 24, (2014).

Results of the request

The author's efforts to learn about predictive algorithms through open records requests were in many respects frustrating….

- Either no response or
- Claims to be either generally exempt from open records acts (as, for example, courts) or beneficiaries of specific exemptions, such as those for trade secrecy.

The results suggest that transparency is a choice that jurisdictions and their vendors make – a choice having less to do with immutable trade secrets or confidentiality concerns than with a culture of disclosure.

Very little information was collected about the development of algorithms (probably because governments were never even in its possession)

.

One exception:

Allegheny County, which contracted for the development of a predictive algorithm from scratch by a consortium of university researchers, was the biggest exception, because it commissioned and possessed reports that detailed the development of its algorithm and disclosed the algorithm itself.

Impediments to transparency

Three principal impediments to making government uses of big data prediction transparent:

1. The absence of appropriate record generation practices around algorithms processes
2. Insufficient government insistence on appropriate disclosure practices
3. The assertion of trade secrecy or other confidential privileges by government contractors .

Impediments to transparency

Brauneis & Goodman find that publicly deployed algorithms will be sufficiently transparent only if:

1. Governments generate **appropriate records** about their objectives for algorithmic processes and subsequent implementation and validation;

2. Government contractors reveal to the public agency **sufficient information** about how they developed the algorithm; and

3. Public agencies and courts **treat trade secrecy claims as the limited exception** to the public disclosure that the law requires (which is currently not the case).

What can be done?

Governments should consciously generate – or demand that their vendors generate – records that will further public understanding of algorithmic processes.

This seems to be what is contemplated by the European Union General Data Protection Regulation (came into force recently in 2018), which stipulates that the function of an algorithm must be made understandable to the public.

Ideally, relevant stakeholders could produce a set of best practices for documenting the creation and implementation of predictive algorithms.

Eight categories of information should be considered according to Brauneis & Goodman:

1. The algorithmic model's general predictive goal;
2. Relevant, available, and collectable data;
3. Considered exclusion of data;
4. Specific predictive criteria;
5. Analytic techniques used;
6. Principal policy choices made;
7. Results of validation studies and audits;
8. Explanation of the predictive algorithm and the algorithm output

[…] Virginia Eubanks's *Automating Inequality* identifies […] profound problems in governmental use of algorithmic sorting systems […] stories of individuals who lose benefits, opportunities, and even custody of their children, thanks to algorithmic assessments that are inaccurate or biased in profound ways […]Were we to approach the problem from a purely technical perspective, we might promote more and better data gathering about the struggling individuals she describes, to ensure that they are not misclassified. But Eubanks argues that complex benefits determinations are not something well-meaning tech experts can "fix." Instead, the system itself is deeply problematic, constantly shifting the goal line (in all too many states) to throw up barriers to access to care […]

(Frank Pascale, Odd Numbers, Why algorithms can't meaningfully hold other algorithms accountable, http://reallifemag.com/odd-numbers/)

Make the public aware of the use of algorithms: Media Literacy and Online Empowerment issues raised by Algorithm-Driven Media Services

Media Literacy and Online Empowerment issues raised by Algorithm-Driven Media Services

Open Evidence has been commissioned by the European Commission (DG CONNECT) to conduct the study on “Media Literacy and Online Empowerment issues raised by Algorithm-Driven Media Services”.

This study will be a first step for EU media policy to analyse the underlying issues posed by algorithm-driven media services and explore the problem area in a structured way for the benefit of policy makers and stakeholders.

The overall objectives of the study are to:

1. Analyse the issues posed by algorithm-driven media services and explore problem areas in a holistic and cross-disciplinary way;
2. Identify gaps in understanding and research;
3. Identify best practices from a media literacy perspective;
4. Include a substantial stakeholder involvement;
5. Include a fundamental rights perspective (freedom of expression/media freedom and pluralism).

The study is not limited to algorithm transparency, but explores benefits and downsides arising from use of algorithms in connection with media services.

The study complements ongoing work by DG JUST on consumer rights together with the European Parliament pilot project, Algorithmic Awareness Building Initiative, and DG CNECT's parallel study 'The Mechanisms that shape Social Media and their Impact on Society,' and any other initiative in the same field at EU or Member State level.

#EUalgorithms

Home About News Contact

For more information check www.open-evidence.com

An **algorithm** is a set of “encoded procedures for transforming input data into a desired output, based on specified calculations” (Gillespie, 2014).

[⋯] Like a recipe, it provides instructions for transforming ingredients into a simple or complex product [⋯]

(Brauneis & Goodman)

[⋯]**Predictive algorithms** are created through analysis of large datasets, typically with the aid of machine-learning processes, to reveal correlations between various features (of a person, circumstance or activity) and desired, objectionable outcomes.
Those patterns can be used to create a model that will estimate the likelihood of future behaviour or events (the output) when given relevant facts (the input)[⋯].

[⋯] An **algorithmic process** will typically involve:

i)the construction of a model to achieve some goal, based on analysis of collected historical data,
ii) the coding of an algorithm that implements this model,
iii) collection of data about subjects to provide inputs for the algorithm,
iv) application of the prescribed algorithmic operations on the input data, and
v)outputs in the form of predictions or recommendations based on the chain of data analysis[⋯]

Example provided by Brauneis & Goodman:

"For example, a government may want to know how likely a prisoner is to commit a crime if paroled, or how likely an admitted student is to enroll in a state university if offered a scholarship of a certain amount. By correlating a set of characteristics of past parolees with their subsequent criminal histories, or of past admitted students with their enrollment decisions, data scientists can build a predictive model. The government can then apply that model to current parolees or admitted students, and predict their behavior."