“Opacity in algorithmic processes, when they have real welfare effects, is a problem in all contexts. But in the case of public governance, it poses particular dangers to democratic accountability, to the efficacy and fairness of governmental processes, and to the competence and agency of government employees tasked with doing the public’s work.”

— “Algorithmic Transparency for the Smart City” by Ellen P. Goodman and Robert Brauneis in the forthcoming Yale Journal of Law and Technology
The Injustice of Algorithms

Prejudice is often coded into software, including tools used by the government.

BY ETHAN CHIEL | January 23, 2018

A CHILD ABUSE PREDICTION MODEL FAILS POOR FAMILIES

How New Jersey Is Leading the Post-Bail Revolution

VIRGINIA FUBANKS BUSINESS 01.15.18 00:00 AM

SENTINEL AND EAGLE NOV 2, 2017
Algorithmic Accountability

Inscrutability of Big Tech
• Black Box Society (Pasquale)
• Weapons of Math Destruction (O’Neil)
• The Platform is Political (Gillespie)
• AI ethics in place – Autonomous vehicles, Uber...

At the same time...

The Smart City rhetoric reprises the techno liberation creed of the 1990’s Internet. Private power? Public interest?
Local government use of predictive algorithms – what can we know?

1. About performance and fairness
2. About politics
3. About private power and control

Transparency → Accountability
Research

We filed
• 43 open records requests
• to public agencies in 23 states
• about six predictive algorithm programs:
  • PSA-Court
  • PredPol
  • Hunchlab
  • Eckerd Rapid Safety Feedback
  • Allegheny County Family Risk
  • Value Added Method – Teacher Evaluation
**Predictive Algorithms: Pretrial Disposition**

Arnold Foundation PSA - Court: Predicts likelihood that criminal defendant awaiting trial will fail to appear, or commit a crime (or violent crime) based on nine factors about him/her.

<table>
<thead>
<tr>
<th>Risk Factor</th>
<th>FTA</th>
<th>NCA</th>
<th>NVCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age at current arrest</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2. Current violent offense</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>Current violent offense &amp; 20 years old or younger</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Pending charge at the time of the offense</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>4. Prior misdemeanor conviction</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>5. Prior felony conviction</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>Prior conviction (misdemeanor or felony)</strong></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Prior violent conviction</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>7. Prior failure to appear in the past two years</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>8. Prior failure to appear older than two years</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>9. Prior sentence to incarceration</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Note: The table indicates the presence (X) or absence of risk factors for various outcomes (FTA: Failure to Appear, NCA: New Crime, NVCA: New Violent Crime).
The PSA does NOT look at any of the following factors:

- race
- gender
- income
- education
- home address
- drug use history
- family status
- marital status
- national origin
- employment
- religion
Predictive Algorithms: Child Welfare

Eckerd Rapid Safety Feedback: Helps family services agencies triage child welfare cases by scoring referrals for risk of injury or death.

The Eckerd Rapid Safety Feedback℠ model represents a pivotal shift in the approach to child safety.

MARCH 30, 2016

Using data now to keep children safe in the future

Unlike many traditional child welfare systems, which intervene only after a problem happens, some jurisdictions are using models designed to keep children safe before trouble escalates. In recognition of National Child Abuse Prevention Month in April, we’d like to look at one promising approach.
Predictive Algorithms: Policing

HunchLAB and Predpol: use historical data about where and when crimes occurred to direct where police should be deployed to deter future crimes.
The Public Interest in Knowing

Democratic accountability
  • What are the policies the program seeks to implement and what tradeoffs does it make?

Performance
  • How does the program perform as implemented? As compared to what baseline?

Justice
  • Does the program ameliorate or perpetuate bias? Systemic inequality?

Governance
  • Do government agents understand the program? Do they exercise discretion w/r/t algorithmic recommendations?
What disclosures would lead to knowing?

1. Basic purpose and structure of algorithm
2. Policy tradeoffs – what and why
3. Validation studies and process before and after roll-out
4. Implementation and training
1. What is the **problem to be solved?** What outcomes does the program seek to optimize? e.g., Prison overcrowding? Crime? Unfairness?

2. What **input data** (e.g., arrests, geographic areas, etc.) were considered relevant to the predicted outcome, including time period and geography covered.

3. **Refinements.** Was the data culled or the model adjusted based on observed or hypothesized problems?
Policy Tradeoffs Reflected in Tuning

Predictive models are usually refined by minimizing some cost function or error factor. What policy choices were made in formulating that function?

For example, a model will have to trade off false positives and false negatives.

Philadelphia’s APPD decided on a cost ratio where false negatives were 2.6 times more costly than false positives.

(Adult Probation and Parole Department) from https://www.nij.gov/journals/271/pages/predicting-recidivism.aspx
Validation Process

1. It is standard practice in machine learning to withhold some of the training data when building a model, and then use it to test the model.

   Was that “validation” step taken, and if so, what were the results?

2. What steps were taken or are planned after implementation to audit performance?
Implementation and Training

Interpretation of results: Do those who are tasked with making decisions based on predictive algorithm results know enough to interpret them properly?

PSA-Court

- High risk
- Low risk

Philadelphia APPD

- High
- Medium
- Low

Risk

Open Records Responses

• 25 either did not provide or reported they did not have responsive documents
• 5 provided confidentiality agreements with the vendor
• 6 provided some documents, typically training slides and materials
• 6 did not respond
• 1 responded in a very complete way with everything but code – has led to an ongoing collaboration on best practices
Impediments

1. Open Records Acts and Private Contractors
2. Trade Secrets / NDAs
3. Competence of Records Custodians and Other Government Employees
4. Inadequate Documentation
5. [Non-Interpretability, Dynamism of Machine Learning Algorithms]
Impediment 1: Private Contractors

- Algos developed by private vendors
- Vendors give very little documentation to governments
- Open records laws typically do not cover outside contractors unless they are acting as records managers for government
Impediment 2: Trade Secrets/NDAs

• Mesa (AZ) Municipal Court (PSA-Court): “Please be advised that the information requested is solely owned and controlled by the Arnold Foundation, and requests for information related to the PSA assessment tool must be referred to the Arnold Foundation directly.”

• 12 California jurisdictions refused to supply Shotspotter data – detection of shots fired in the city – even though it’s not secret, and not IP

• Overbroad TS claims being made by vendors, and accepted by jurisdictions
Impediment 3: Govt. Employees

Records custodians are not the ones who use the algorithm

Those who use the algorithm don’t understand it
Impediment 4: Inadequate Records

Jurisdictions have to supply only those records they have (with some exceptions for querying databases). Governments are not insisting on obtaining, and are not creating, the records that would satisfy the public’s right to know.
FIXES

Government procurement: don’t do deals without requiring ongoing documentation, circumscribing TS carve-outs, data and records ownership
Data Reasoning with Open Data
Lessons from the COMPAS-ProPublica debate

Anne L. Washington, PhD
NYU - Steinhardt School

Sunday February 11, 2018
Regulating Computing and Code - Governance and Algorithms Panel
Silicon Flatirons 2018 Technology Policy Conference
University of Colorado Law School
DATA SCIENCE REASONING

Can you argue with an algorithm?
Reasoning

• Arguments
  ▪ Convince, Interpret, or Explain
  ▪ Arguments logically connect evidence and reasoning to support a claim
    • Quantitative Statistical Reasoning
    • Inductive Reasoning
    • Data Science Reasoning
SUPREME COURT OF WISCONSIN

2015AP157-CR

State of Wisconsin,
   Plaintiff-Respondent,
v.
Eric L. Loomis,
   Defendant-Appellant.

ON CERTIFICATION FROM THE COURT OF APPEALS

July 13, 2016

BRIEFS:
April 5, 2016
¶ 49 The Skaff* court explained that if the PSI Report was incorrect or incomplete, no person was in a better position than the defendant to refute, supplement or explain the PSI. (State v. Loomis, 2016)


.. but what if a Presentence Investigation Report ("PSI") is produced by an algorithm?
Algorithms in Criminal Justice

Predictive scores

- A statistical model of behaviors, habits, or characteristics summarized in a number
- **Risk/Needs Assessment Scores**
  - Determines potential criminal behavior or preventative interventions
THE DEBATE

Are risk assessment scores biased?
Summer 2016

- **US Congress**
  H.R 759 Corrections and Recidivism Reduction Act

- **Wisconsin v Loomis**
  881 N.W.2d 749 (Wis. 2016)

- **Machine Bias**
  ProPublica Journalists

- **COMPAS risk scores**
  Correctional Offender Management Profiling for Alternative Sanctions
The Public Debate: ProPublica vs COMPAS

- Machine Bias
  www.propublica.org
  - By Angwin, Larson, Mattu, Kirchner

- COMPAS Risk Scales:
  volarisgroup.com
  - By Northpointe (Volaris)
    Correctional Offender Management Profiling for Alternative Sanctions
The Scholarly Debate: Is COMPAS fair?

Open data

ProPublica

Data repository

github.com/propublica/compas-analysis

- From May 2016 – Dec 2017
- nearly 230 publications
- cited Angwin (2016), Dieterich (2016), Larson (2016) or ProPublica's github data repository

Fairness requires interpretation

Kleinberg (2016)

• No mathematical ideal choice
  ▪ Not possible to satisfy the three constraints simultaneously
  ▪ Algorithmic estimates are generally not pure yes-no decisions


Berk (2017)

• Impossible to maximize accuracy and fairness at the same time


Seven types of Fairness

• 1. Overall accuracy equality
• 2. Statistical parity
• 3. Conditional procedure accuracy
• 4. Conditional use accuracy equality
• 5. Treatment equality
• 6. Total fairness
AN INTERPRETIVE ADVANTAGE

Why the defense had no ability to refute, supplement, or explain without comparative data
Who will commit crime?
Risk Assessment:
Who is likely to commit crime?
Risk Scores:
Who was a threat to public safety?

3 5 1 4 2 8
9 3 4 7 1 2

Highlighted:
9 3 4 7 1
Can we predict new scores?

3 5 1 4 2
9 3 4 7 1 3
Needs Assessment:
Who needs help to succeed?
Why the court has additional information
Judging the judge’s scales
... with open data

1 Verification with test data
2 Proof of verification with open data
LESSONS

What can we learn from the debate?
Innovating bureaucracy

USB Typewriters created by Jack Zylkin  [https://www.usbtypewriter.com/collections/typewriters/products](https://www.usbtypewriter.com/collections/typewriters/products)
Data transparency

- Analytics does not provide “an answer”
- Data science requires interpretation

trust in allah, but tie your camel’s leg at night
Доверяй, но проверяй
Data Science Reasoning
Anne L. Washington, PhD
washingtona@acm.org

Assistant Professor of Data Policy
Steinhardt School, New York, NY
New York University
http://annewashington.com

FUNDING

2016-2017 Fellowship
New York, NY

Currently funded under National Science Foundation
What if the score conflicts with other indicators?
¶30 "This court reviews sentencing decisions under the erroneous exercise of discretion standard." An erroneous exercise of discretion occurs when a circuit court imposes a sentence "without the underpinnings of an explained judicial reasoning process."

McCleary v. State, 49 Wis. 2d 263, 278, 182 N.W.2d 512 (1971); see also State v. Gallion, 2004 WI 42, ¶3, 270 Wis. 2d 535, 678 N.W.2d 197.


Data Science Reasoning

Anne L. Washington, PhD
anne.washington@nyu.edu
washingtona@acm.org

Assistant Professor of Data Policy
Steinhardt School, New York, NY
New York University

http://annewashington.com
1. Title >>
2. Data Science Reasoning >>
3. The Debate >>
4. Court Advantage >>
5. Lessons >>
6. Appendix >>
FAREWELL TO THE FIRST GOLDEN ERA

THE MAMAS & THE PAPAS

MONDAY MONDAY - CALIFORNIA DREAMIN' - TWELVE THIRTY (Young Girls Are Coming To The Canyon)
DEDICATED TO THE ONE I LOVE - I CALL YOUR NAME - WORDS OF LOVE - I SAW HER AGAIN LAST NIGHT - CREEQUE ALLEY
LOOK THROUGH MY WINDOW - DANCING IN THE STREET - GO WHERE YOU WANNA GO - GOT A FEELIN'

D 50625