Rethinking Access Control In the Home IoT

CMSC 23210/33210 Usable Security and Privacy

Weijia He, Maximilian Golla, Roshni Padhi, Jordan Ofek, Markus Dürmuth, Earlene Fernandes, Blase Ur
From Single User to Multi User

Single User

iPhone X, macOS High Sierra - apple.com
From Single User to Multi User
Vendors Still Treat It The Old Way!

What level of access do you want to give “John”?  

- Guest  
  ✔️  
- Owner
Please Enter Your Password:
Please Enter Your Password:
Home IoT Devices

“Play music!”
“Order me a puppy!”
Just to Summarize...

<table>
<thead>
<tr>
<th>Traditional Devices</th>
<th>Home IoT</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Single User</td>
<td>Multiple Users</td>
</tr>
<tr>
<td>With Keyboard &amp; Screen</td>
<td>Often Without Keyboard &amp; Screen</td>
</tr>
<tr>
<td>Device-level Access Control</td>
<td>Capability-level Access Control</td>
</tr>
</tbody>
</table>
Research Goals

We conducted a user study to...

• Map desired access-control policies for home IoT devices
  • How policies vary by relationships and capabilities
  • Identify potential default policies
Method
Before Implementing the Survey...

• What do relationships and capabilities mean for home IoT?
6 Relationships

24 Relationships

6 Relationships

- Your Spouse
- Your Teenage Child
- Your Child in Elementary School
- A Visiting Family Member
- The Babysitter
- Your Neighbor

Icons made by Vectors Market from www.flaticon.com
22 Capabilities

1)

2)

Icon made by Those Icons, Smashicons, Vectors Market from www.flaticon.com
22 Capabilities

1) Alexa - [amazon.com](http://amazon.com), Nest Cam - [nest.com](http://nest.com), Ring Doorbell - [thewirecutter.com](http://thewirecutter.com), Husqvarna Automower - [husqvarna.com](http://husqvarna.com), Hue - [philips.com](http://philips.com)

2) Order Online

3) Mower Rule

4) Live Video

5) Lights Rule

6) Answer Door
Imagine you are the owner of a <smart device>. Using this device, some users can access the following feature: <capability>. When should <relationship> be able to use this feature?

• Always  • Sometimes  • Never
Imagine you are the owner of a Smart Voice Assistant. Using this device, some users can access the following feature:

**Make online purchases (e.g., on Amazon) on a shared household account.**

When should your spouse be able to use this feature?

- Always
- Sometimes
- Never
Are Relationships and Capabilities Enough?

5 pm – 6 pm

12 am – 1 am

LG Smart TV – lg.com; Icons made by Freepik and Pixel Buddha - CC 3.0 BY
Research Goals

We conducted a user study to...

• Map desired access-control policies for Home IoT Devices
  • How policies vary by relationships and capabilities
  • Identify potential default policies

• What contextual factors affect the user's decision?
Imagine you are the owner of a Smart Voice Assistant. Using this device, some users can access the following feature: Make online purchases (e.g., on Amazon) on a shared household account.

When should your spouse be able to use this feature?

- Always
- Sometimes
- Never
User Study

• When should they have access to this capability?
• When should they not have access to this capability?
Results
425 Participants

- 54% Male
- 46% Female

Age
- 25-34: 47%

Home IoT Device
- 44%

CS: 19%
Results

Given one particular capability, what access-control policy should be set up for whom?
<table>
<thead>
<tr>
<th></th>
<th>Spouse</th>
<th>Teenager</th>
<th>Child</th>
<th>Visiting Family</th>
<th>Babysitter</th>
<th>Neighbor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Play Music</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order Online</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spouse</td>
<td>Always</td>
<td></td>
<td>Never</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teenager</td>
<td></td>
<td>Always</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child</td>
<td></td>
<td></td>
<td>Always</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visiting Family</td>
<td></td>
<td></td>
<td></td>
<td>Always</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Babysitter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Always</td>
<td></td>
</tr>
<tr>
<td>Neighbor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Always</td>
</tr>
</tbody>
</table>

*Icon made by Alexander Madyankin, Roman Shamin, Echo Plus with built-in Hub – Black - www.amazon.com*
Capabilities Within One Device

- Answer Doorbell
- Delete Lock Log

- Spouse: Answer Door
- Teenager: Answer Door
- Child: Answer Door
- Visiting Family: Answer Door
- Babysitter: Answer Door
- Neighbor: Answer Door

- Delete Lock Log: Sometimes
- Delete Lock Log: Sometimes
- Delete Lock Log: Sometimes
- Delete Lock Log: Sometimes
- Delete Lock Log: Sometimes
- Delete Lock Log: Sometimes

- Always
- Sometimes
- Never
Spouse Can Do Almost Everything

Access Control Preference for Different Relationships/Capabilities

Spouse

Software Update
Play Music
Order Online
Temperature Log
Mower On/Off
Mower Rule
Lock Log
Lock State
Lock Rule
Answer Door
Delete Lock Log
Lights State
Lights On/Off
Lights Rule
Light Scheme
New Device
New User
Live Video
Facial Recognition
Delete Video
Camera On/Off
Camera Angle
Neighbor Can Do Nothing

Access Control Preference for Different Relationships/Capabilities

- Spouse
  - Software Update
  - Play Music
  - Order Online
  - Temperature Log
  - Mower On/Off
  - Mower Rule
  - Lock Log
  - Lock State
  - Lock Rule
  - Answer Door
  - Delete Lock Log
  - Lights State
  - Lights On/Off
  - Lights Rule
  - Light Scheme
  - New Device
  - New User
  - Live Video
  - Facial Recognition
  - Delete Video
  - Camera On/Off
  - Camera Angle

- Neighbor
  - Software Update
  - Play Music
  - Order Online
  - Temperature Log
  - Mower On/Off
  - Mower Rule
  - Lock Log
  - Lock State
  - Lock Rule
  - Answer Door
  - Delete Lock Log
  - Lights State
  - Lights On/Off
  - Lights Rule
  - Light Scheme
  - New Device
  - New User
  - Live Video
  - Facial Recognition
  - Delete Video
  - Camera On/Off
  - Camera Angle
Other Relationships Are More Complex

<table>
<thead>
<tr>
<th>Access Control Preference for Different Relationships/Capabilities</th>
<th>Teenager</th>
<th>Child</th>
<th>Visiting Family</th>
<th>Babysitter</th>
<th>Neighbor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software Update</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Play Music</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order Online</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temperature Log</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mower On/Off</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mower Rule</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lock Log</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lock State</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lock Rule</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Answer Door</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delete Lock Log</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lights State</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lights On/Off</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lights Rule</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light Scheme</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Device</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New User</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Live Video</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facial Recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delete Video</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Camera On/Off</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Camera Angle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Always
- Sometimes
- Never
Teenager vs. Child

“At 16 they would be able to pick their own things to buy but the final purchase should be ultimately my choice and need my authorization.”
Teenager vs. Child

“At 16 they would be able to pick their own things to buy but the final purchase should be ultimately my choice and need my authorization.”

“They are in no way responsible enough at this age.”
# Relationships Matter... But Are Not Enough

## Access Control Preference for Different Relationships/Capabilities

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Spouse</th>
<th>Teenager</th>
<th>Child</th>
<th>Visiting Family</th>
<th>Babysitter</th>
<th>Neighbor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software Update</td>
<td>❌     </td>
<td>❌     </td>
<td>❌    </td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
</tr>
<tr>
<td>Play Music</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Order Online</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Temperature Log</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
</tr>
<tr>
<td>Mower On/Off</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mower Rule</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lock Log</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lock State</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lock Rule</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Answer Door</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Delete Lock Log</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lights State</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lights On/Off</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lights Rule</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Light Scheme</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>New Device</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>New User</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Live Video</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Facial Recognition</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Delete Video</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Camera On/Off</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Camera Angle</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Temperature Log</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
<td>❌</td>
</tr>
<tr>
<td>Order Online</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mower Rule</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lock State</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lock Rule</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Answer Door</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Delete Lock Log</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lights State</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lights On/Off</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Lights Rule</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Light Scheme</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>New Device</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>New User</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Live Video</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Facial Recognition</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Delete Video</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Camera On/Off</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Camera Angle</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Legend: **Always**, **Sometimes**, **Never**
Relationships Matter...But Are Not Enough

What does *sometimes* mean?
Contextual Factors
“I would not want anyone trying to use the mower at night. The neighbors would most likely get mad.”
Factor: People Around

“They would be allowed to use it whenever I am home with them.”
Factor: Location of User

Why not play some music?

What is going on???

“Why do you need to use it if you aren’t close?”
Factor: Location of Device

“If it is used in the bedroom then it would matter who has access.”
Factor: Explicit Permission

"Can I...?"

“When they are authorized by the owner.”
Factor: Consequences
“They shouldn't use the lights if they are using them too frequently.”
Factor: Understanding

“I would need to teach her how to first.”
Factor: Help

“If they want to come over to mow the lawn, then why not?”

Thank you!!!
Recap: Missing From Current Systems

Relationships

Capabilities

Contextual Factors
Design Implications
Current: Guest vs. Owner

What level of access do you want to give “John”? 

- Guest
- Owner
Future: Designing for Relationships

Adding a new user:

is:

- spouse
- teenage child
- young child
- visiting family member
- babysitter
- neighbor
Future: Designing for Relationships

Adding a new user:

is

- spouse
- teenage child
- young child
- visiting family member
- babysitter
- neighbor
Future: Designing for Relationships

Smart Home

Adding a new user:

is a young child

Next
Future: Relationships and Capabilities

Smart Home

Default Settings for a Young Child

Voice Assistant

Lights

Thermostat
Future: Relationships and Capabilities

Default Settings for a Young Child

Voice Assistant

With permission, allowed to play music

Never allowed to order online
Current: Full Access or Temporary Access

Set Access Time

- **Start Date**: Thu, 19 July 2018
- **Start Time**: 06:00 PM
- **End Date**: Thu, 19 July 2018
- **End Time**: 06:00 PM

ADD SCHEDULE

Set individual date and time to allow users to access the door temporarily.

- **Jan 19 to 21 July**

ACCESS TIME

Thu, 19 Jul 2018 02:00 PM to Sat, 21 Jul 2018 12:00 PM
Future: Contextual Factors

Your young child can have access when
- I’m around
- I’m not around
Future: Device Context

Is your camera an...

Indoor Camera  Outdoor Camera
Future: Device Location

Is your camera placed in...

Living Room

Bedroom
Rethinking Access Control and Authentication for the Home Internet of Things

Weijia He, Maximilian Golla, Roshni Padhi, Jordan Ofek, Markus Dürmuth, Earlence Fernandes, Blase Ur
Fairness and Machine Learning

Galen Harrison
Julia Hanson
Usable Security and Privacy CMSC 23210/33210
1. Why does this matter?

2. What is machine learning?

3. Why should we be worried about whether or not it's fair?

4. What are some techniques for making machine learning fair?
Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016
Machine Bias
There's software used across the country to predict future criminals. And it's biased against blacks.
by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

Car Insurance Companies Charge Higher Rates in Some Minority Neighborhoods
First-of-its-kind data analysis finds price differences that can't be explained by risk alone

By Julia Angwin, Jeff Larson, Lauren Kirchner, and Surya Mattu of ProPublica
Last updated: April 21, 2017
Machine Bias
There’s software used across the country to predict future criminals. And it’s biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

Car Insurance Companies Charge Higher Rates in Some Minority Neighborhoods
First-of-its-kind data analysis finds price differences that can’t

Amazon’s Face Recognition Falsely Matched 28 Members of Congress With Mugshots

By Jacob Snow, Technology & Civil Liberties Attorney, ACLU of Northern California
JULY 26, 2018 | 8:00 AM

TAGS: Face Recognition Technology, Surveillance Technologies, Privacy & Technology
Machine Bias

There’s software used across the country to predict future criminals. And it’s biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

Car Insurance Companies Charge Higher First-of-its-Kind Rates in

Amazon’s Members of Congress With Mugshots

The false matches were disproportionately of people of color, including six members of the Congressional Black Caucus, among them civil rights legend Rep. John Lewis (D-Ga.). These results demonstrate why Congress should join the ACLU in calling for a moratorium on law enforcement use of face surveillance.

By Jacob Snow, Technology & Civil Liberties Attorney, ACLU of Northern California

JULY 26, 2018 | 8:00 AM

TAGS: Face Recognition Technology, Surveillance Technologies, Privacy & Technology
Amazon scraps secret AI recruiting tool that showed bias against women

By Jacob Snow, Technology & Civil Liberties Attorney, ACLU of Northern California
JULY 26, 2018 | 8:00 AM

TAGS: Face Recognition Technology, Surveillance Technologies, Privacy & Technology
What is Machine Learning?

• Problem

• There is some unknown function \( f : A \rightarrow B \)

• Examples

\( A = \{ \text{pictures} \}, \ B = \{ \text{is face} \} \)

\( A = \{ \text{chess board} \}, \ B = \{ \text{optimal move} \} \)
What is Machine Learning?

- Can’t find $f$ directly, but have examples of $(\vec{a}, f(\vec{a}))$
- Can approximate $f$
- $\vec{a}$ could be pixels of picture
- in income prediction (age, education, …)
Linear Regression
Regression

“Most machine learning is actually regression” - Someone

Key idea: find the right $\vec{w} = (w_1, w_2, \ldots, w_k)$

Such that $\sum_{i=1}^{n} (w \cdot x_i - y_i)^2$
Other Machine Learning Techniques

- Logistic regression
- Support Vector Machines (SVM)
- Deep learning (aka neural networks)
Key Questions

• These will be more relevant later!

• Does the type of model applied to the problem matter? If so, when?

• When does the machine learning problem matter?

• What, if anything, makes the use of data for ML different from other ways of making decisions?
Returning to Compas

Machine Bias

There’s software used across the country to predict future criminals. And it’s biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica
May 23, 2016

Propublica reporting on Northpointe risk assessment tool
Risk Assessments

- Risk assessments = Predictive algorithms

- **In bail hearings:** “will this person commit a crime or fail to appear in court?”

- **At sentencing:** “will this person commit crime in the future?”

- Theoretical goal: reduce the number of individuals behind bars before trial without increasing risk to the public
Returning to Compass

137 questions, 10 topics

<table>
<thead>
<tr>
<th>Current criminal charges</th>
<th>Criminal attitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criminal history</td>
<td>Neighborhood safety</td>
</tr>
<tr>
<td>Substance abuse</td>
<td>Criminal history of friends and family</td>
</tr>
<tr>
<td>Stability of employment</td>
<td>Quality of social life</td>
</tr>
<tr>
<td>Personality</td>
<td>Education and behavior in school</td>
</tr>
</tbody>
</table>

No questions about sensitive features!
Consider the following

\[ PPV = \frac{TP}{TP + FP} \]

\[ FPR = \frac{p}{1 - p} \frac{1 - PPV}{PPV} (1 - FNR) \]

• If \( p \) differs between two groups, then equal PPV implies differing FPR rates

\[
FPR = \frac{p}{1 - p} \frac{1 - PPV}{PPV} (1 - FNR)
\]
Other ways bias can arise

• Pre-existing bias

  • Individual - individual people within system design, implementation, use are biased

  • Societal - society as a whole has biases (e.g. a loan system that uses zip codes, reinforcing redlining)

Friedman and Nissenbaum, 1996
Other ways bias can arise

- Technical Bias
  - Computer tools
  - Decontextualized algorithms
  - Random number generation
  - Formalization of human constraints

Friedman and Nissenbaum, 1996
Other ways bias can arise

- Emergent Bias
- New Societal Knowledge
- Mismatch between users and system
  - Different expertise
  - Different values

Friedman and Nissenbaum, 1996
Another perspective

(1) *Fair*: lacking biases which create unfair and discriminatory outcomes;
(2) *Accountable*: answerable to the people subject to them;
(3) *Transparent*: open about how, and why, particular decisions were made.

By assuring these conditions are met, we can rest easy, threatened no more by the possibility of an algorithm producing harmful outcomes.
Another perspective

(1) *Fair*: lacking biases; (2) *Accountable*: answerable to the people subject to them; (3) *Transparent*: open about how, and why, particular decisions were made. By assuring these conditions are met, we can rest easy, threatened no more by the possibility of an unethical, biased, or harmful outcome; therefore, researchers are encouraged to ensure these conditions are met, for example, by directly “auditing” algorithms to identify whether they are biased [5, 6, 28], or not, before being rendered. In an attempt to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered. In an effort to move past this, LNI has developed a new, automated approach. An algorithm, provided with vast amounts of social media and communications data, identifies subjects to be rendered.

Logan-Nolan Industries

*Helping Humanity Make Ends Meat*

Figure 1: A publicity image for the project, produced by Logan-Nolan Industries
Key Questions

• What is the specific problem that we’re trying to solve?

• How much responsibility does the data scientist/machine learning developer have for the broader effects of their work?

• Should we be concerned with fairness per se? Or is justice/control/equity a better framing?
Possible solutions

- Technical
- Design
- Regulatory
Individual Fairness

Idea: treat similar people in a similar manner

\[ M : V \rightarrow \Delta(A), \ d_1, d_2 \text{ metrics in } V \text{ and } \Delta(A) \text{ respectively} \]

\[ d_2(M(x), M(y)) \leq d_1(x, y) \]

What intuitions does this encode? What might be some problems?

Fairness Through Awareness, Dwork et al. 2011
Disparate Impact

- Equal Employment Opportunity Commission interprets to say that if a facially neutral test selects a group at 80% of the rate for other groups, then it is discriminatory according to Title VII of the Civil Rights Act § 2000e-2(a) (2)

- Generalize to

\[
\frac{Pr(C = 1 | X = 0)}{Pr(C = 1 | X = 1)} \leq \tau
\]

Certifying and Removing Disparate Impact, Feldman et al. 2014
Process Fairness

• Idea: Some features may be fair to use, others may not be

• Base feature use fairness through a survey

• Examples

  • Current charges

  • Criminal History: self

  • Criminal History: social circle

  • Education and school behavior

Human Perceptions of Fairness in Algorithmic Decision Making: A Case Study of Criminal Risk Prediction, Grgić-Hlača et al. 2018
Questions? Comments?
Additional Resources

- ACM Conference on Fairness Accountability and Transparency (ACM FAT*) https://fatconference.org/

- FAT/ML http://www.fatml.org/
