

Digital Vulnerability: Characterizing Risks and Contemplating Responses

@ChrisDiehl

July 28, 2016





Global Impact & Local Perspective



Agenda

Current Risks

Potential Responses

Existing Gaps

Future Options

Digital Vulnerability

Marginalization

Opportunity



Unwitting
Disclosure

+

Discrimination

Witting
Disclosure

+

Privilege

Disclosure Risk



Private traits and attributes are predictable from digital records of human behavior

Michal Kosinski^{a,1}, David Stillwell^a, and Thore Graepel^b

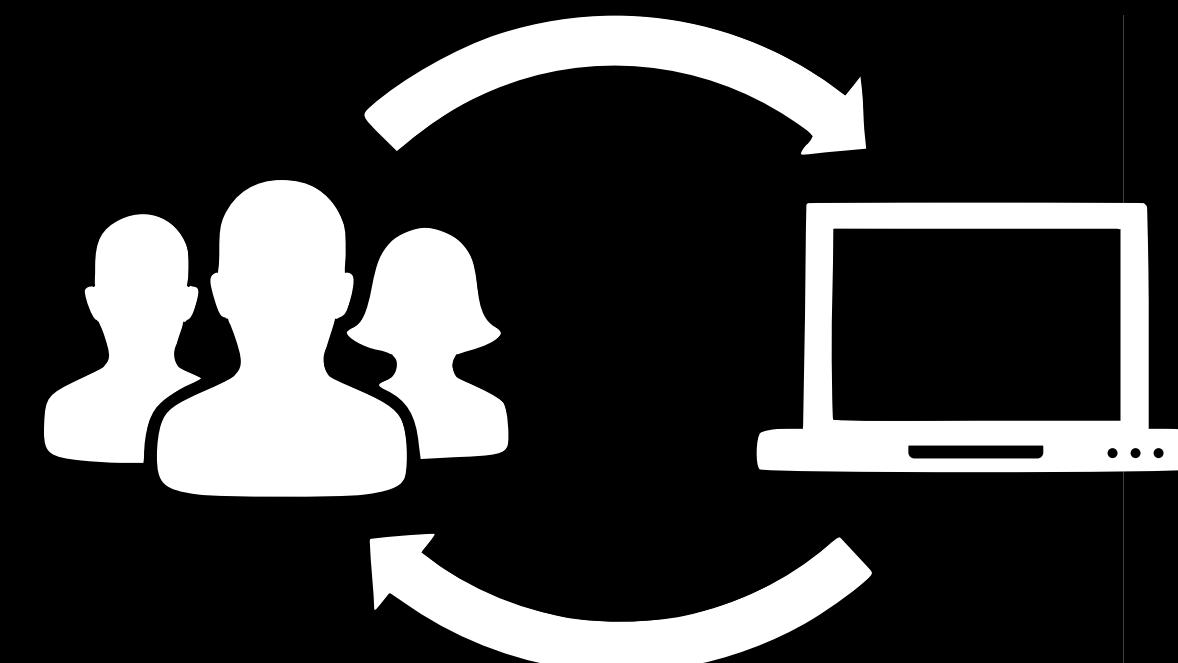


HIGH RISK

10

LOW RISK

3



Prediction Risk

Actuation Risk



Where to Start? The Design Conversation

Defining the Middle Ground

No Personalization

*“Measure and Predict
All the Things”*



*No Raw Data
Collection
and
Retention*

*Private, Fair,
Interpretable
Inference*

*Unconstrained
Raw Data
Collection and
Retention*

“Equitable Inference”

Existing Gaps
Widely Understood Options
Trusted Implementations
Design Patterns and Principles

Community Knowledge Creation and Dissemination

1. Define Algorithmic Options
2. Develop Best Practices and Open-Source Implementations
3. Teach the Approaches
4. Gather Feedback from Practitioners

Let's Do This!

[bit.ly/
equitable_inference_experiment](https://bit.ly/equitable_inference_experiment)

chris@thedataguild.com