

# Profiling and Particularity

Solon Barocas  
Microsoft Research and Cornell University

March 25, 2019

# Basis upon which to judge other factors

- Individual?
  - 'Sins of the father'
  - 'The company you keep'
- Volitional?
- Consensual?
- Logical relationship?

# Sentencing, by the Numbers

By SONJA B. STARR AUG. 10, 2014

984356000794286582133521  
875909650987304000000209  
315827315927187590965098  
730400000020935372398435  
600079412865000002000000  
034123435912718759096509  
873040011000020935372398  
410794286582133521875909  
650987304000000233117359  
271187590965098730400000  
020935372398413560007942  
810000



Oliver Munday

certainly unconstitutional. It contravenes the principle that punishment should depend on what a defendant did, not on who he is or how much money he has.

ANN ARBOR, Mich. — IN a recent letter to the United States Sentencing Commission, Attorney General Eric H. Holder Jr. sharply criticized the growing trend of [evidence-based sentencing](#), in which courts use data-driven predictions of defendants' future crime risk to shape sentences. Mr. Holder is swimming against a powerful current. At least 20 states have implemented this practice, including some that require risk scores to be considered in every sentencing decision. Many more are considering it, as is Congress, in pending sentencing-reform bills.

Risk-assessment advocates say it's a no-brainer: Who could oppose "smarter" sentencing? But Mr. Holder is right to pick this fight. As currently used, the practice is deeply unfair, and almost

# *Bearden v. Georgia*

“lumping him together with other poor persons and thereby classifying him as dangerous ... would be little more than punishing a person for his poverty.”

## US 9100400 B2

In particular embodiments, a method includes accessing a graph structure comprising a plurality of nodes and edges where each node represents a user, receiving a request to transmit content related to a first user to a second user, and prohibiting transmission of the content to the second user if the first user and the second user are connected in the graph structure through a series of edges and nodes that comprises an unauthorized node.

**External Links:** [USPTO](#), [USPTO Assignment](#), [Espacenet](#)

## IMAGES (9)

[illegible]

### DESCRIPTION

### CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation under 35 U.S.C. §120 of U.S. patent application Ser. No. 10/897,766, filed Jul. 22, 2004.

## BACKGROUND OF THE INVENTION

## 1. Field of the Invention

The present invention generally relates to authorizing activities and authenticating individuals, and more specifically, to methods for authorizing information flow to an individual and authenticating an individual for access to

## CLAIMS (16)

What is claimed is:

1. A method comprising:

receiving, by one or more computing devices, a request to send to a first user content relating to a second user;

accessing, by the one or more computing devices, a graph structure comprising a plurality of nodes and a plurality of edges connecting the nodes, wherein:

each edge connects two nodes and represents a single degree of

DJIA ▲ 22928.04 0.25%

S&P 500 ▲ 2555.54 0.09%

Nasdaq ▲ 6614.80 0.14%

U.S. 10 Yr ▼ -5/32 Yield 2.297%

Crude Oil ▲ 51.91 0.89%

Euro ▼ 1.1803 -0.15%

# THE WALL STREET JOURNAL.

[Subscribe Now](#) | [Sign In](#)

**Join Now: 50% Off**

[Home](#) [World](#) [U.S.](#) [Politics](#) [Economy](#) [Business](#) [Tech](#) **[Markets](#)** [Opinion](#) [Life & Arts](#) [Real Estate](#) [WSJ. Magazine](#)

Search 



Financial Stocks  
Push U.S. Indexes  
Higher



Silicon Valley Vs.  
Wall Street: Can the  
New Long-Term  
Stock ...



J.P. Morgan's  
Dimon May Hate  
Bitcoin, but He Loves  
Blockchain



As Edward Jones  
Tops \$1 Trillion in  
Assets, It Seeks Street  
Cred



Social  
Number  
Hated—  
Irreplacable



## MARKETS



## Facebook Isn't So Good at Judging Your Credit After All

Lenders drop plans to use social media to gauge creditworthiness as regulators balk; plus, one startup says, 'It's creepy'

By [Telis Demos](#) and [Deepa Seetharaman](#)

Feb. 24, 2016 5:30 a.m. ET

In the growing business of online lending, Facebook was supposed to be the new FICO. No longer.

Regulatory hurdles have stymied efforts by online lenders and credit-data providers to use information from social media to judge American borrowers' creditworthiness.

Facebook Inc. itself made it tougher for outsiders to tap its treasure trove of...

[TO READ THE FULL STORY](#)

[SUBSCRIBE](#)

[SIGN IN](#)

### Recommended Videos

1. **Trump Calls on Democratic Leaders to Negotiate Health Care**



2. **Police Camera Captures Dramatic Wildfire Evacuation**

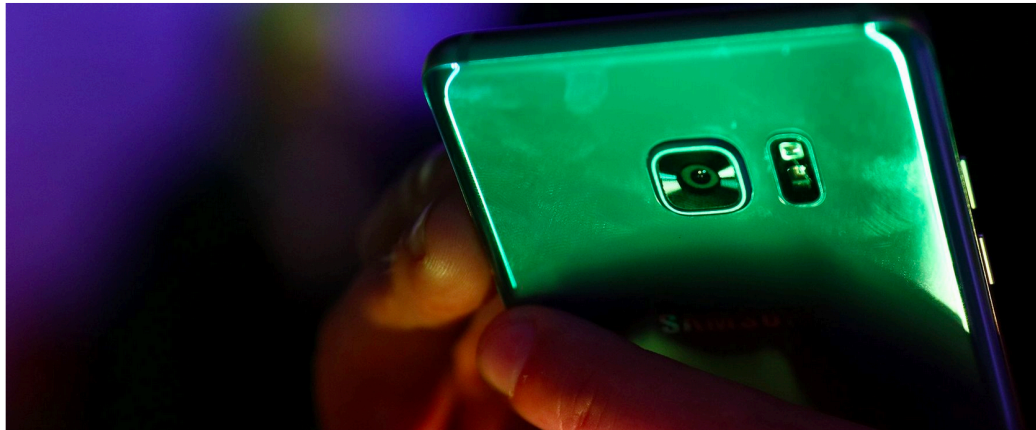


# No Credit History? No Problem. Lenders Are Looking at Your Phone Data

By **Olga Kharif**

November 25, 2016, 5:00 AM EST

- FICO, Equifax stike partnerships to expand access to loans
- 'The way you use the phone is a proxy for the way you live'



## Most Read

- 1 A 'Crazy' Stock Market Is Punishing Sellers
- 2 Yellen Calls Inflation the 'Biggest Surprise' in the Economy
- 3 London House Prices Fall Most Since Financial Crisis
- 4 Storm Ophelia Kills 3 in Ireland's Worst Weather in Decades
- 5 Facebook Is Looking for Employees With National Security Clearances







"All data is credit data"

oscar





JOURNAL REPORTS: LEADERSHIP

## Big Data Uncovers Some Weird Correlations

There's a Link Between Sales and Phases of the Moon, Among Other Things

By [Deborah Gage](#)

Updated March 23, 2014 4:36 p.m. ET

Are sales deals affected by the cycles of the moon? Is it possible to determine credit risk by the way a person types?

### JOURNAL REPORT

- Insights from [The Experts](#)
- Read more at [WSJ.com/LeadershipReport](http://WSJ.com/LeadershipReport)

### MORE IN UNLEASHING INNOVATION:




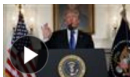

#### BIG DATA

- [People Battle to Regain Online Privacy](#)
- [Companies Embrace Real-Time Marketing](#)
- [The Data Companies Wish They Had](#)

Fast new data-crunching software combined with a flood of public and private data is allowing companies to test these and other seemingly far-fetched theories, asking questions that few people would have thought to ask before.

By combining human and artificial intelligence, they seek to uncover clever insights and make predictions that could give businesses an advantage in an increasingly competitive marketplace.

### Recommended Videos

1. [Jeffrey Katzenberg Calls Bob Weinstein 'Abusive'](#) 
2. [Top Women at Microsoft, JP Morgan Discuss Gender Equality](#) 
3. [Sheryl Sandberg on Harvey Weinstein](#) 
4. [What's Next For the Iran Nuclear Deal?](#) 
5. [The Unnamed Podvideocast With Jason Gay and Dana White](#) 

### Most Popular Articles

# Basis upon which to judge other factors

- Individual?
  - 'Sins of the father'
  - 'The company you keep'
- Volitional?
- Consensual?
- Logical relationship?

# Basis upon which to judge other factors

- Logical relationship?
  - Department of the Treasury “will conclude a variable [...] does not warrant further scrutiny if the variable is statistically related to loan performance, and has an understandable relationship to an individual applicant's creditworthiness.”

# Basis upon which to judge other factors

- Individual?
  - 'Sins of the father'
  - 'The company you keep'

# Basis upon which to judge other factors

- Individual?
  - 'Sins of the father'
  - 'The company you keep'
  - 'Reference group'

The right to  
be treated as an individual



# The right to be treated as an individual

...and not simply a member of a group

# The right to be treated as an individual

...and not simply a member of a ~~legally protected~~ group

# Profiling and particularity

- Deindividualization (Vedder)
- Statistical discrimination (Lippert-Rasmussen)
- Intuitive notion of fairness: everyone should be assessed on her or his individual characteristics and merits—in her or his particularity
- Complementary intuition: Individuals should not be assumed to possess the quality ascribed to the group to which he or she belongs—fit the profile

# Quick sidebar

- Data mining
- Knowledge discovery in databases
- Big data
- Machine learning
- Artificial intelligence



FREDERICK SCHAUER

PROFILES

PROBABILITIES

AND

STEREOTYPES

# The problem with profiling

- 'Non-distributive group profiles' (Vedder)
- 'Statistically sound but nonuniversal generalizations' (Schauer)

What does it mean to treat  
someone as an individual?

# Sentencing, by the Numbers

By SONJA B. STARR AUG. 10, 2014

984356000794286582133521  
875909650987304000000209  
315827315927187590965098  
730400000020935372398435  
600079412865000002000000  
034123435912718759096509  
873040011000020935372398  
410794286582133521875909  
650987304000000233117359  
271187590965098730400000  
020935372398413560007942  
810000



Oliver Munday

certainly unconstitutional. It contravenes the principle that punishment should depend on what a defendant did, not on who he is or how much money he has.

ANN ARBOR, Mich. — IN a recent letter to the United States Sentencing Commission, Attorney General Eric H. Holder Jr. sharply criticized the growing trend of [evidence-based sentencing](#), in which courts use data-driven predictions of defendants' future crime risk to shape sentences. Mr. Holder is swimming against a powerful current. At least 20 states have implemented this practice, including some that require risk scores to be considered in every sentencing decision. Many more are considering it, as is Congress, in pending sentencing-reform bills.

Risk-assessment advocates say it's a no-brainer: Who could oppose "smarter" sentencing? But Mr. Holder is right to pick this fight. As currently used, the practice is deeply unfair, and almost



**FICO®**

oscar



# Schauer's argument

- “[A]cknowledging the way in which seemingly direct observation involves a process of inference and generalization enables us to appreciate that even the processes that initially appear to us to be ‘direct,’ ‘actual,’ or individualized turn out to rely far more on generalizations from past experience than is often appreciated. Once we see that all evidence is in the final analysis probabilistic, the distinction between the probabilistic and the ‘real,’ ‘direct,’ or ‘actual’ emerges as even more of an anomaly [...] the cumulative set of inferences that produces a purportedly ‘direct’ conclusion or observation is nothing more than a collection of inferences drawn from generalizations known to be reliable”

How can failing to treat  
someone as an individual be  
objectionable if it's impossible?

# A matter of degree?

Coarse-grain

Fine-grain

# A matter of degree?

Coarse-grain

Fine-grain

Race

Gender

Age

...

Dog breed

## Lippert-Rasmussen's argument

- “X treats Y as an individual if, and only if, X's treatment of Y is informed by all relevant information, statistical or non-statistical, reasonably available to X.”

## Lippert-Rasmussen's formulation

- “X treats Y as an individual if, and only if, X’s treatment of Y is informed by all relevant information, statistical or non-statistical, ***reasonably available*** to X.”



## Lippert-Rasmussen's justification

- “But obtaining information is costly, so it is morally justified, all things considered, to treat people on the basis of statistical generalizations even though one knows that, in effect, this will mean that one will treat some people in ways, for better or worse, that they do not deserve to be treated”

## My restatement of this formulation

- “[T]he perceived legitimacy seems to depend on a number of factors: (1) whether the errors seem avoidable because (2) gaining access to additional or more granular data would be trivial or (3) would not involve costs that (4) outweigh the benefits.”

What kinds of outcomes might a cost/benefit analysis permit?

Is a more fine-grain form of  
statistical discrimination always  
preferable?