

# Social Activity Data & Predictive Analytics

## An Opportunity to Advance oSTEM

Richard Bellamy

3<sup>rd</sup> National oSTEM Conference  
Google New York, October 26<sup>th</sup> & 27<sup>th</sup>, 2013

# Social Activity Data & Predictive Analytics

We need social media platforms to have this data.



Social media and other digital platforms should be used to enhance—not substitute for—face-to-face experiences.



Privacy concerns should be focused on what companies (e.g., advertisers) are allowed to do with the information.

Social activity data enables oSTEM to add value to our social network by fundamentally changing its structure.

### **Innovation**

‘Those with many weak ties are best placed to diffuse innovations perceived as unsafe or controversial.’

### **Bridges**

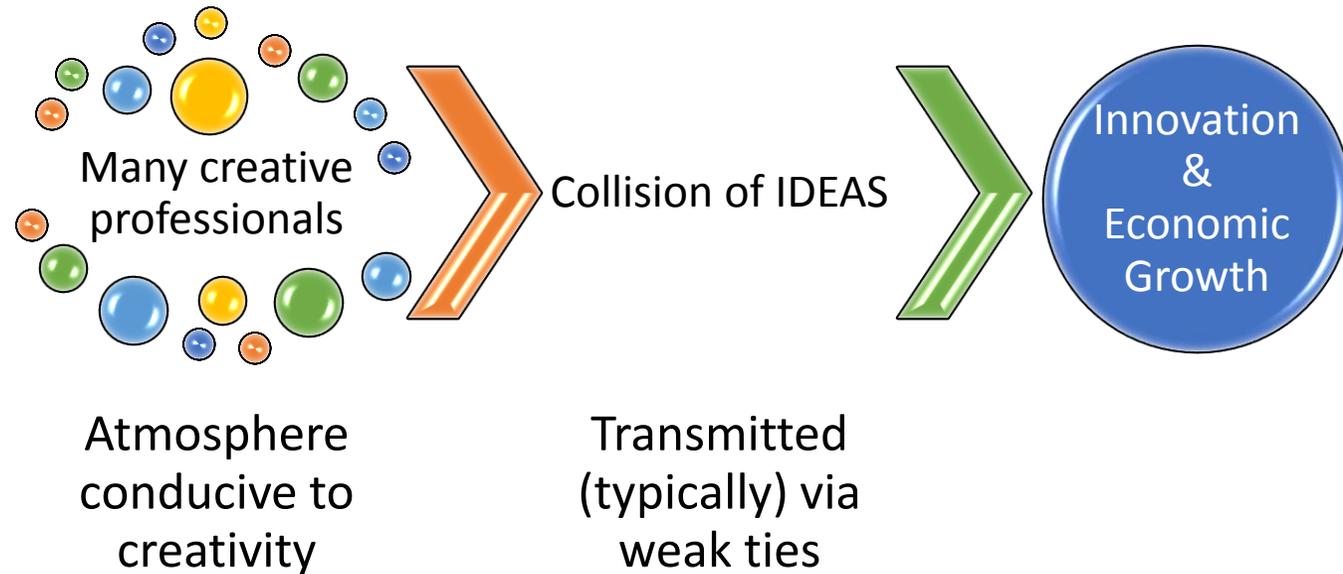
“No strong tie is a bridge.” **Do we agree?**

What about a strong, remote tie?

### **Strength of Weak Ties**

- ‘Strong ties lead to overall fragmentation.
- ‘Weak ties are indispensable to individual’s opportunities and to their integration into their communities.
- ‘A “local bridge” is the only line in a network that provides a reasonably short path between two points.’

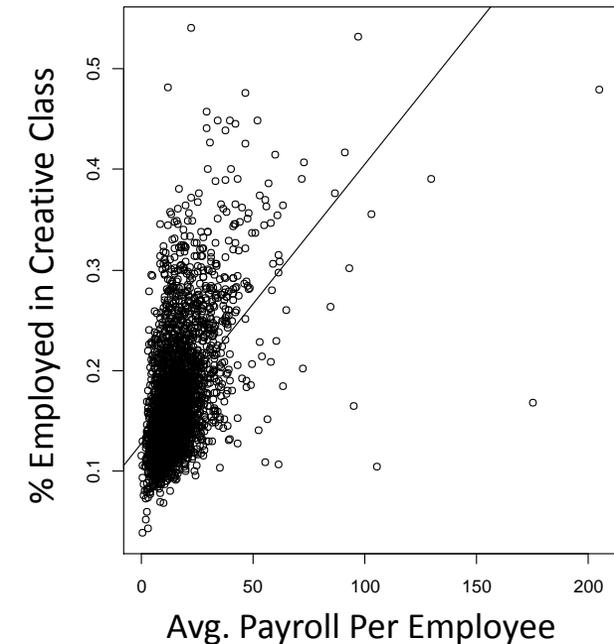
# Our network determines our level of opportunity and innovation.



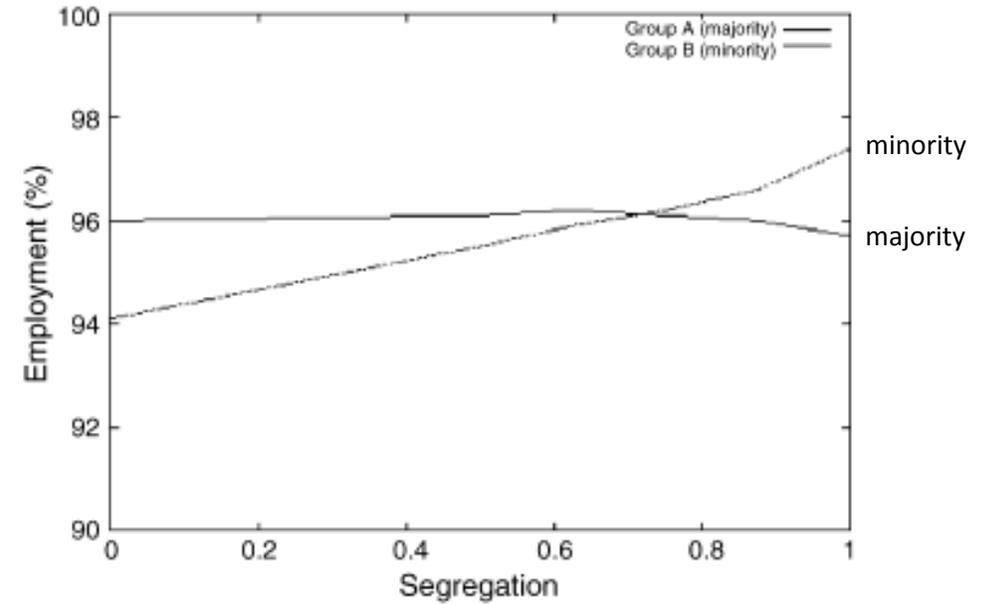
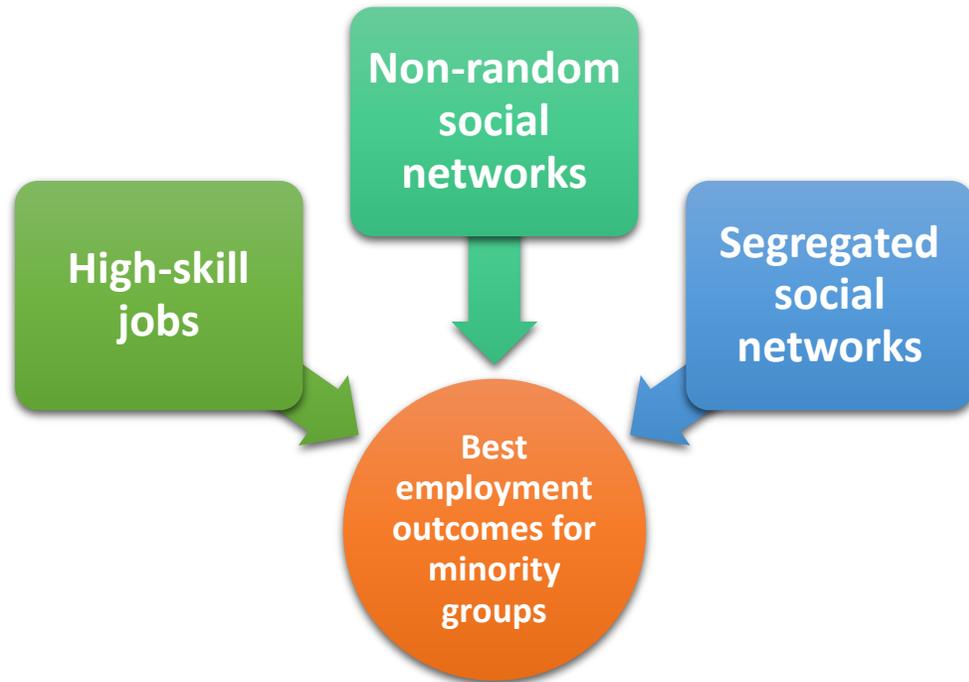
Creative class workers in same-sex couples are **2.5 times more likely** to move to a state after state-level marriage equality is enacted.

## Creative Class Theory

“Creative people are attracted to places most conducive to creative activity[, which] . . . increases local economic dynamism.”

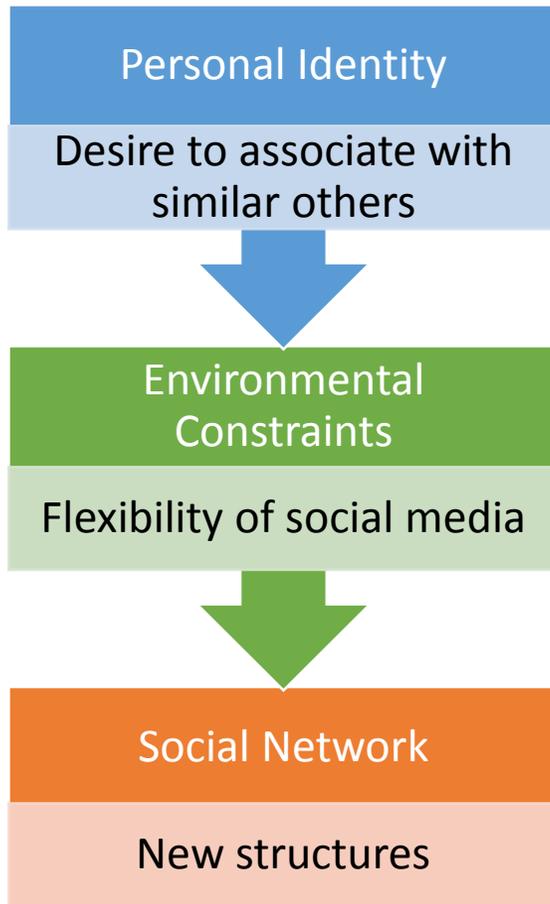


# We benefit from placing oSTEM at the center of our network.



The employment rate of a minority group with non-random social networks increases as segregation increases.

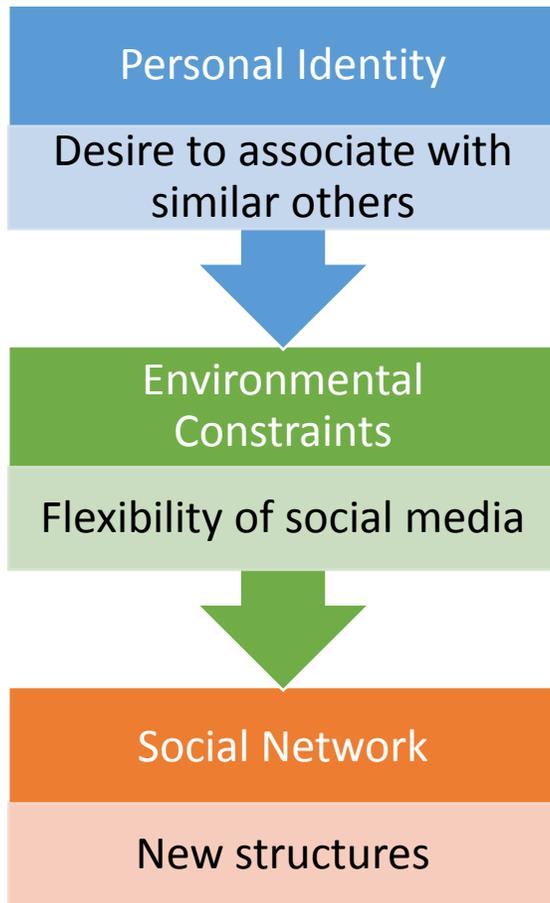
# Social media platforms reduce geographic constraints on our network.



Social media reduces geographic constraints.

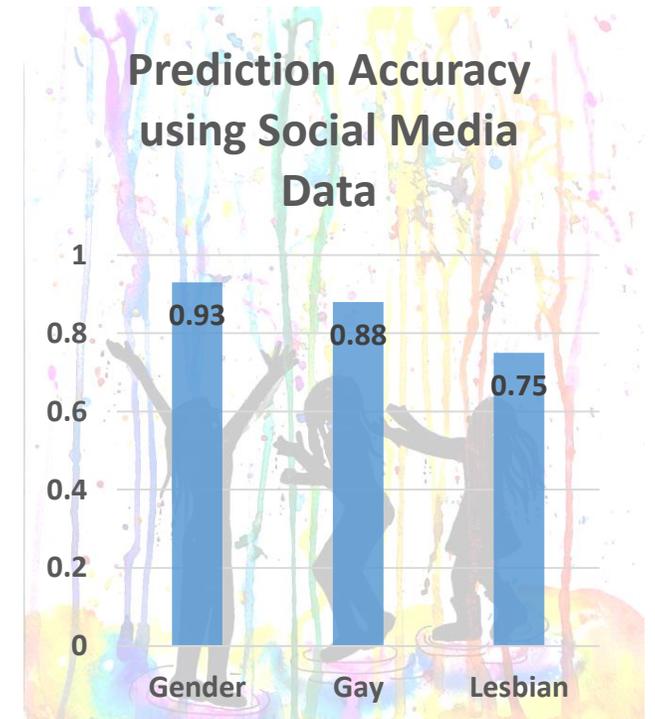
- Users of social networking services are 30% **less likely to know** their neighbors.
- Internet users are 26% **less likely to rely on** their neighbors for help with small services.
- Yet they **remain as willing to help** their neighbors with the same activities.

# Predictive analytics using social activity data places similar others at the center of our network.



Social media content streams are designed to maximize each user's engagement.

- Instead of **democratization**, individuals use social media to associate with **similar others**.
- User engagement is dependent on content that **stimulates** the user, regardless of the content creator's **authority**.



# Social Activity Data & Predictive Analytics

We need social media platforms to have this data.



Social media and other digital platforms should be used to enhance—not substitute for—face-to-face experiences.



Privacy concerns should be focused on what companies (e.g., advertisers) are allowed to do with the information.

The benefits that oSTEM provides are maximized when oSTEM is integrated with our local experience.

Life-Centric Benefits

**Strong, Remote Ties**

More cohesive **national** LGBTQA community

More **influence** over societal views

Greater access to **opportunity**

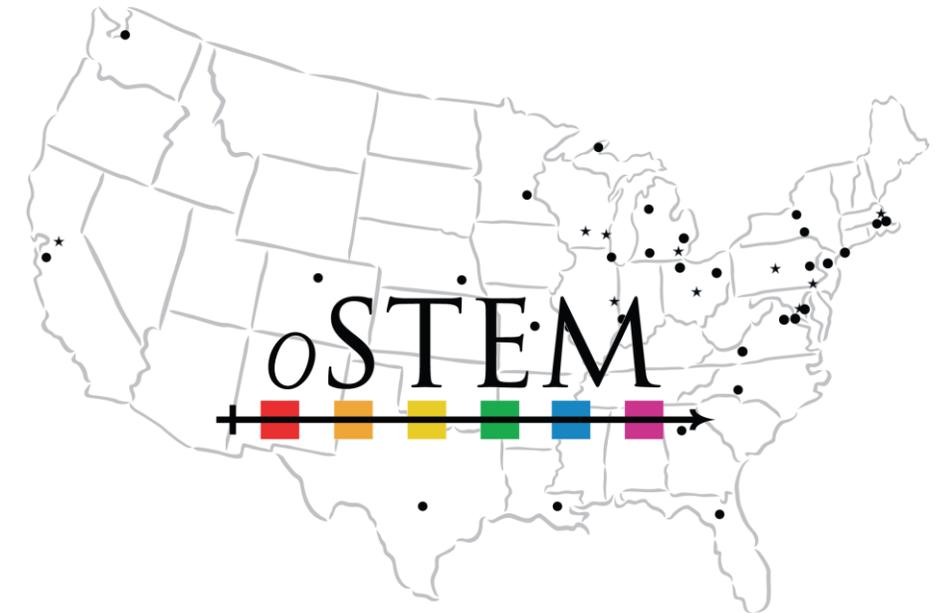
Locality-Centric Benefits

**Improved Identity Compatibility**

Increased **cognitive capacity**

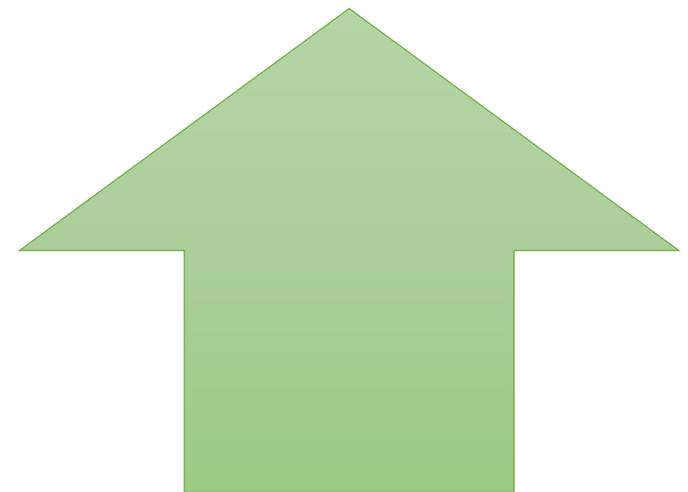
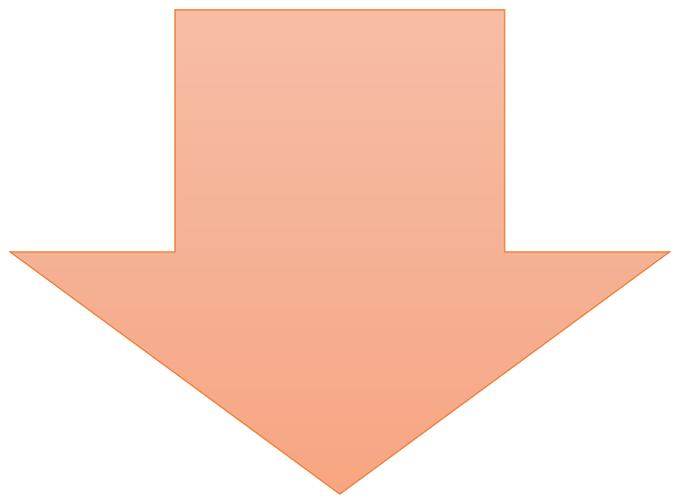
Stronger motivation to pursue **multiple goals**

Increased **interpersonal problem solving**



Relying on social media platforms as a substitute for face-to-face experiences can have negative consequences.

Increased time spent on a social media platform during a two-week period was correlated with a decrease in satisfaction with life.



Social media platforms facilitate building new ties.

High intensity use of a social media platform enabled students with lower self-esteem or satisfaction with life to build more new ties.

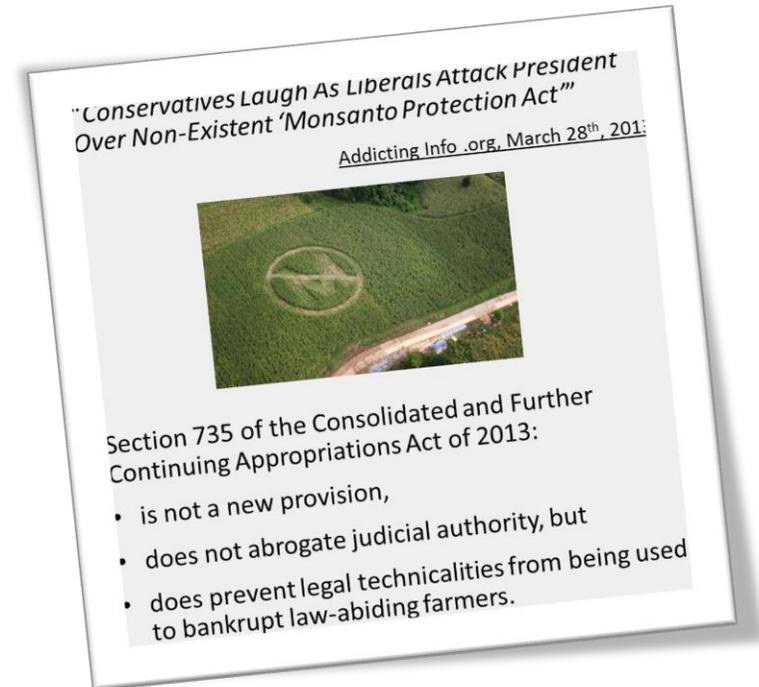
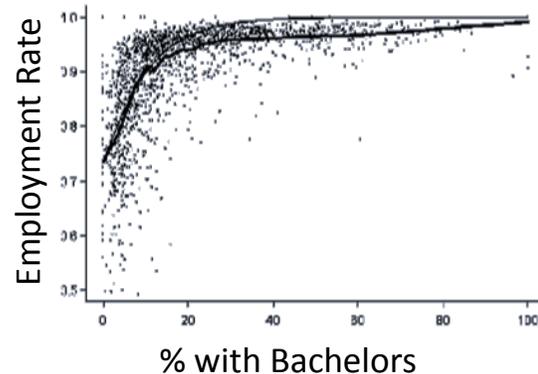
Intensity of usage of a social media platform in year one was correlated with new ties in year two.

# It is possible for a network to be *too* similar.

There exists a **critical level of human capital**:

- below which no group member will be employed, and
- near which a **small change** in human capital can have a **large effect** on employment outcomes.

Social networks explain high unemployment rates among certain minority groups.



**Groupthink Experiment:** Subject, placed in a group with four people, each giving the same *clearly wrong* answer. **1 in 3** people will give that same wrong answer.

# Social Activity Data & Predictive Analytics

We need social media platforms to have this data.



Social media and other digital platforms should be used to enhance—not substitute for—face-to-face experiences.

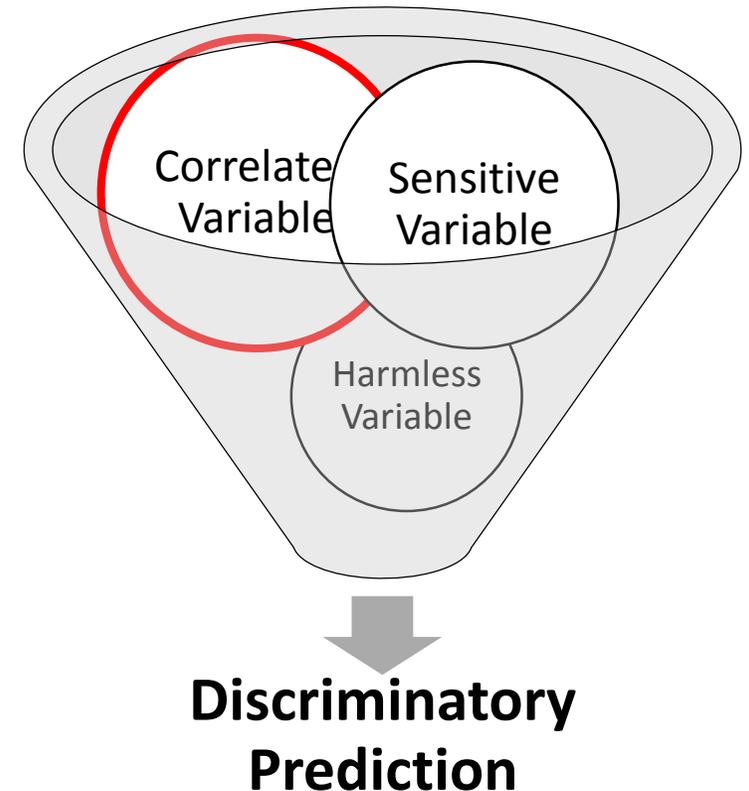


Privacy concerns should be focused on what companies (e.g., advertisers) are allowed to do with the information.

Withholding information from corporations will cause discriminatory analytics.

**Redlining:**

Removing a sensitive variable increases discrimination when a correlated variable remains.

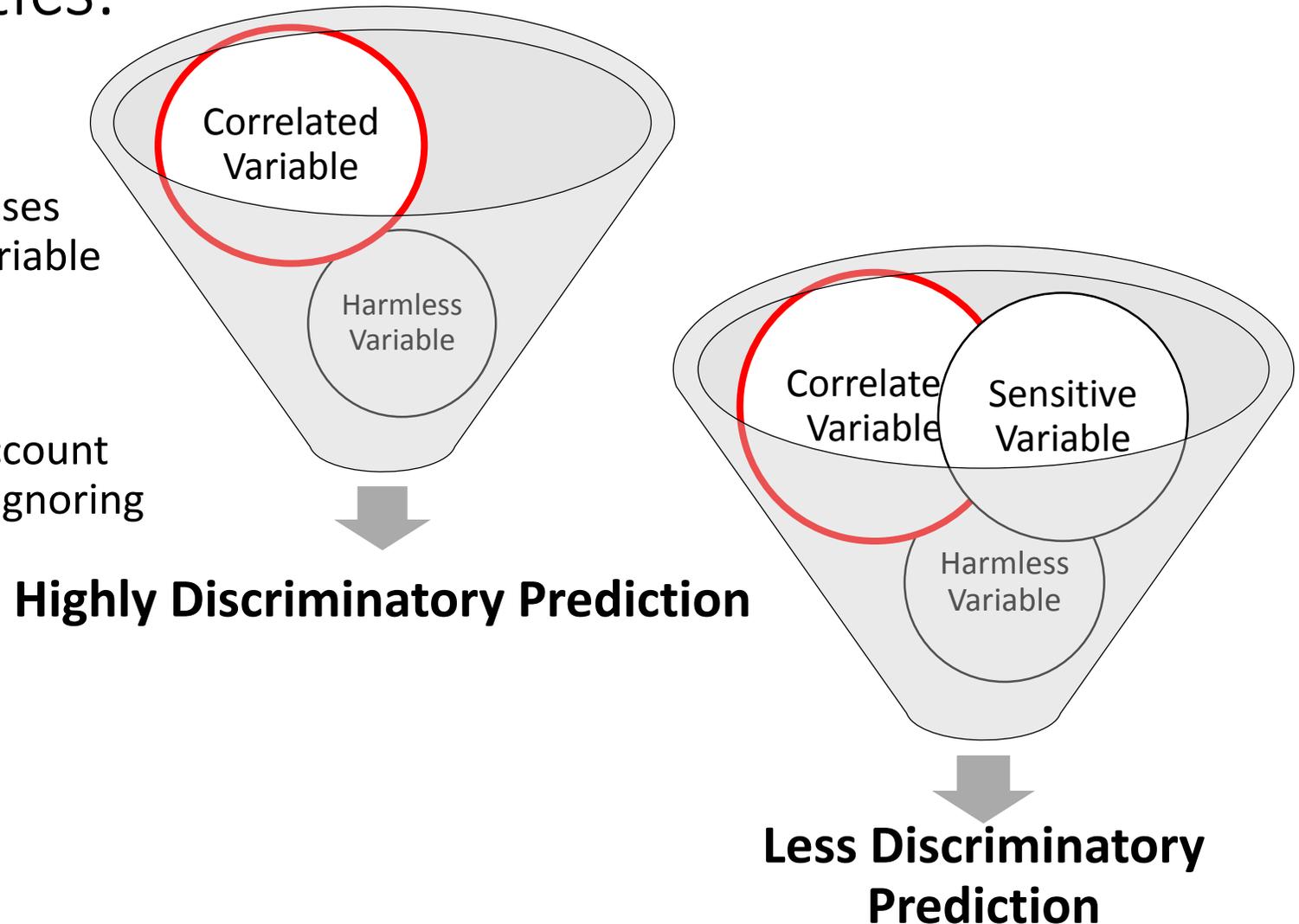


# Withholding information from corporations will cause discriminatory analytics.

## Redlining:

Removing a sensitive variable increases discrimination when a correlated variable remains.

**Discrimination-aware algorithms** account for the sensitive variable instead of ignoring it.



# Withholding information from corporations will cause discriminatory analytics.

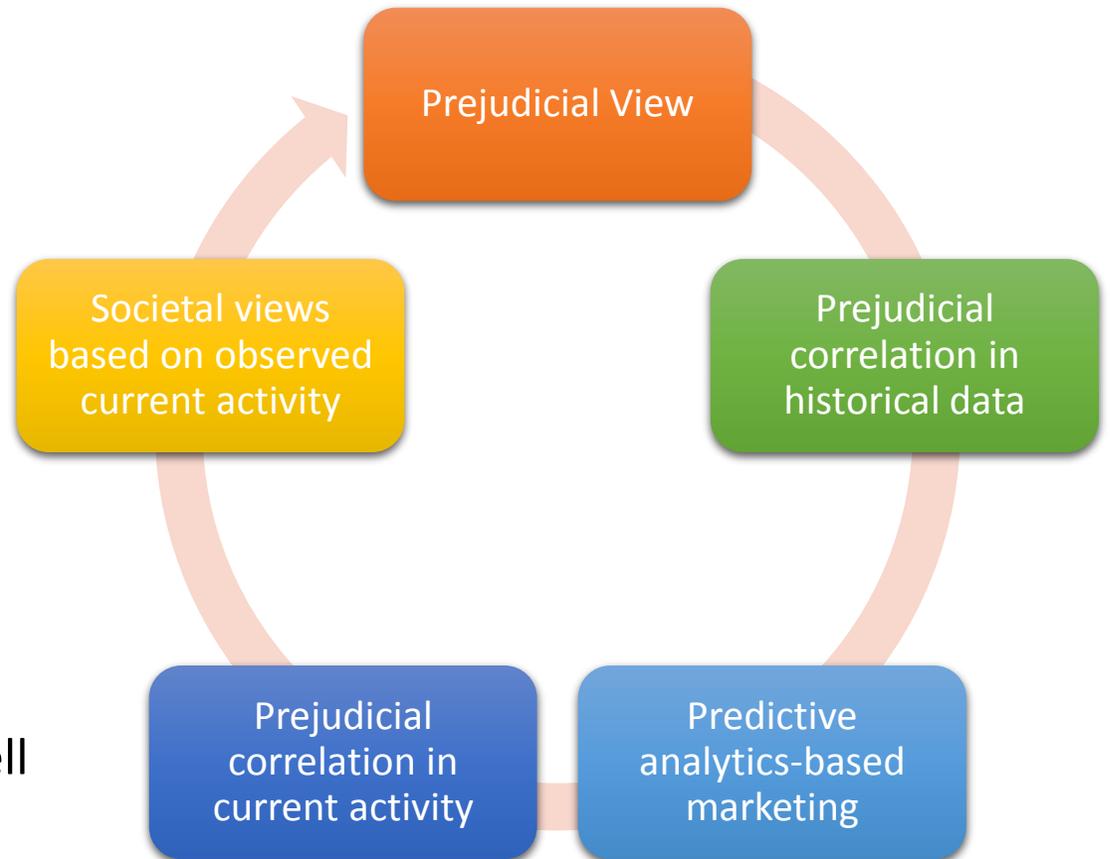
## Redlining:

Removing a sensitive variable increases discrimination when a correlated variable remains.

**Discrimination-aware algorithms** account for the sensitive variable instead of ignoring it.

Credit risk decisions are being made using **capitalization** of names, and **pre- vs. post-paid** cell phones.

*Do we know if these variables are correlated with sensitive variables if sensitive variables are not in the dataset?*



**By understanding the effect that our social data has,  
and sharing our data with social media platforms  
conscious of that effect, social media platforms can be  
the most impactful resource available to us for  
strengthening our community.**

# Sources

## Structure Matters

Slide 4: Gates, *Marriage Equality and the Creative Class*, The Williams Institute (May 2009)

Slide 3: Granovetter, *The Strength of Weak Ties*, 78 American J. of Sociology 1360 (May 1973)

Slide 5: Tassier & Menczer, *Social network structure, segregation, and equality in a labor market with referral hiring*, 66 J. Econ. Behavior & Organization 514 (2008)

Slide 4: Wojan, Lambert, & McGranahan, *Emoting with their feet: Bohemian attraction to creative milieu*, 7 J. of Econ. Geography 711 (2007)

## Social Media Platforms

Slide 7: boyd, "Streams of Content, Limited Attention" Web2.0 Expo, November 17, 2009

Slide 10: Ellison, Steinfield & Lampe, *The benefits of Facebook "friends:" Social capital and college students' use of online social network sites*, 12 J. Computer-Mediated Communication 1143 (2007)

Slide 6: Hampton, et al, *Social Isolation and New Technology*, Pew Internet & American Life Project (Nov. 4, 2009)

Slide 7: Kosinski, Stillwell, & Graepel, *Private traits and attributes are predictable from digital records of human behavior*, 110 Proceedings of the National Academy of Sciences of The United States of America, 5802 (April 9, 2013)

Slide 10: Kross, et. al., Facebook Use Predicts Declines in Subjective Well-Being in Young Adults. PLoS ONE 8(8): e69841. doi:10.1371/journal.pone.0069841 (Aug. 14, 2013)

Slide 10: Steinfield, Ellison, & Lampe, *Social capital, self-esteem, and use of online social network sites: A longitudinal analysis*, 29 J. of Applied Psychology 434 (2008)

## Benefits & Risks

Slide 12: Calders & Verwer, *Three naïve Bayes approaches for discrimination-free classification*, 21 J. of Data Mining and Knowledge Discovery 277 (Sept. 2010)

Slide 14: Carney, *Flush with \$20M from Peter Thiel, ZestFinance is measuring credit risk through non-traditional big data*, PandoDaily (July 31, 2013)

Slide 11: Downes, *Conservatives Laugh As Liberals Attack President Over Non-Existent 'Monsanto Protection Act'*, Addicting Info (Mar. 28, 2013)

Slide 11: Krauth, *A dynamic model of job networking and social influences on employment*, 28 J. of Econ. Dynamics & Control 1185 (2003)

Slide 11: Morris & Miller, *The Effects of Consensus-Breaking and Consensus Preempting Partners on Reduction of Conformity*, 11 J. of Experimental Social Psychology 215 (1975)

Slide 9: Rothbard & Ramarajan, *Checking Your Identities at the Door? Positive Relationships Between Nonwork and Work Identities in Exploring Positive Identities and Organizations* (Roberts & Dutton eds., 2009)