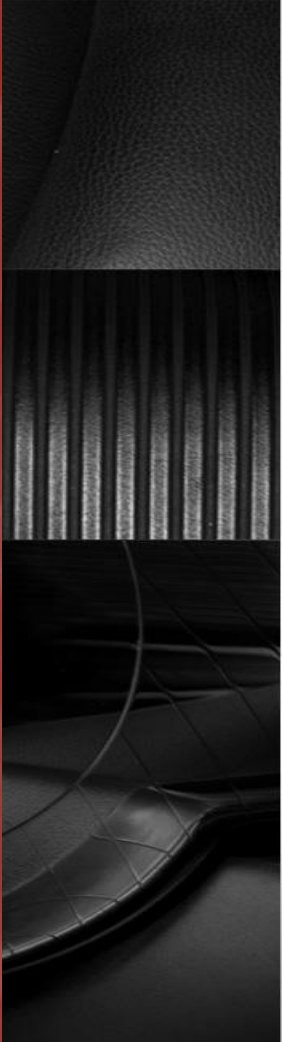


Learning to Love Big Data

The Ethics and Politics of Big Data and
Disruptive Technologies





PRESENTATION

- What is ‘Big Data?’
- Big Data as Disruptive Technology
- The Ethics and Politics of Big Data



What is Big Data?



Big Data

- *One of the biggest new ideas in computing is “big data.” There is unanimous agreement that big data is revolutionizing commerce in the 21st century. When it comes to business, big data offers **unprecedented insight, improved decision-making, and untapped sources of profit.***

40 ZETTABYTES

[43 TRILLION GIGABYTES]

of data will be created by 2020, an increase of 300 times from 2005



Volume SCALE OF DATA

It's estimated that 2.5 QUINTILLION BYTES

[2.3 TRILLION GIGABYTES]
of data are created each day



Most companies in the U.S. have at least
100 TERABYTES
[100,000 GIGABYTES]
of data stored

The New York Stock Exchange captures
1 TB OF TRADE INFORMATION
during each trading session



Velocity ANALYSIS OF STREAMING DATA

By 2016, it is projected
there will be
**18.9 BILLION
NETWORK
CONNECTIONS**
— almost 2.5 connections
per person on earth



Modern cars have close to
100 SENSORS
that monitor items such as
fuel level and tire pressure



The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume, Velocity, Variety and Veracity**

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015
4.4 MILLION IT JOBS
will be created globally to support big data,
with 1.9 million in the United States



As of 2011, the global size of
data in healthcare was
estimated to be

150 EXABYTES
[161 BILLION GIGABYTES]



**30 BILLION
PIECES OF CONTENT**
are shared on Facebook
every month



Variety DIFFERENT FORMS OF DATA



By 2014, it's anticipated
there will be
**420 MILLION
WEARABLE, WIRELESS
HEALTH MONITORS**

**4 BILLION+
HOURS OF VIDEO**
are watched on
YouTube each month



400 MILLION TWEETS
are sent per day by about 200
million monthly active users



**1 IN 3 BUSINESS
LEADERS**

don't trust the information
they use to make decisions



in one survey were unsure of
how much of their data was
inaccurate



Veracity UNCERTAINTY OF DATA

Poor data quality costs the US
economy around
\$3.1 TRILLION A YEAR





Big Data: Disruptive Technology?



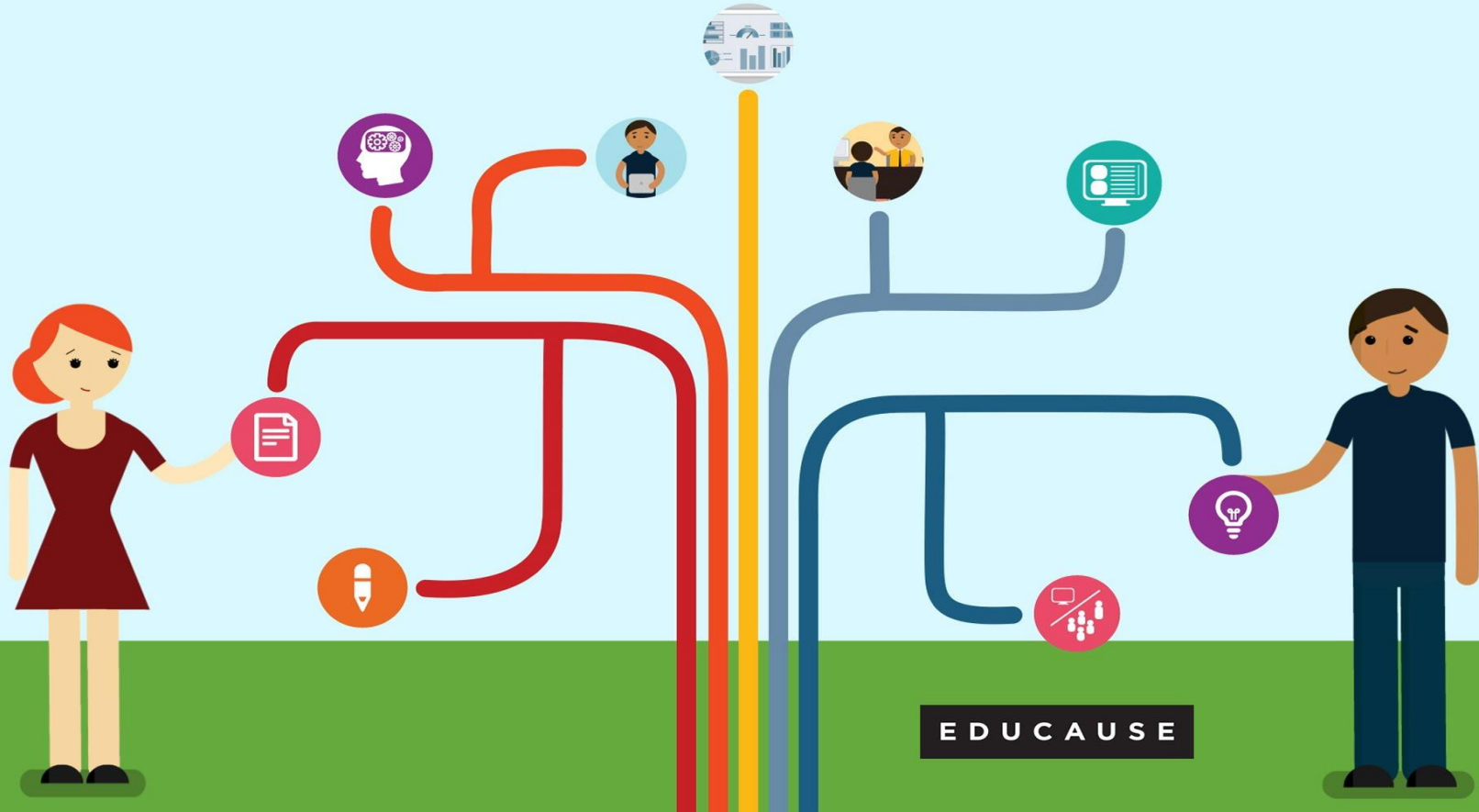
Big Data: Disruptive Technology

- Big data is **changing the competitive landscape**.
- Those who are in a position to take advantage of it often get to market faster with products and services that are **better aligned with customer needs and desires**.
- In some cases could upend established industries altogether.

WHY ROBO ADVISOR IS BETTER FOR YOUR MONEY?



WHAT IS PERSONALIZED LEARNING?





Big Data & the Future of Agriculture



DATA
is the new **OIL!**



ETHICS AND POLITICS OF BIG DATA



The Ethics and Politics of Big Data

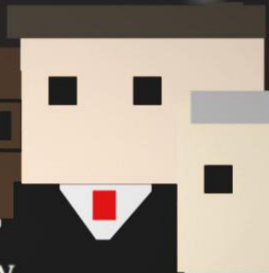
- What is Big Data doing for us?
- What **might** it do for us?
- What should we **not** entrust it to do?
- Who and what is represented by these algorithms?
- How does this technology shape what and how we see?

$$\frac{\max_k(P_k) - \min_k(P_k)}{\sum_{k=1}^K P_k}$$

$$P = \frac{1}{K} \sum_{i=1}^N p_i.$$

$$\frac{\min_k(P_k)}{P_k}$$

$$C_{IPQ} = \frac{4\pi A}{P^2}.$$



$$P = \frac{1}{K} \sum_{i=1}^N p_i.$$

$$R(V)$$

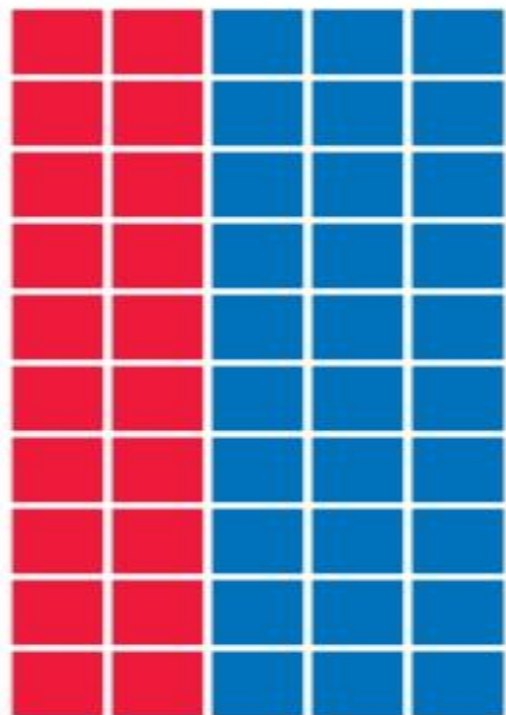
$$R(V) = \frac{dE(S|V)}{dV}$$

**The algorithm that could help
end partisan gerrymandering**

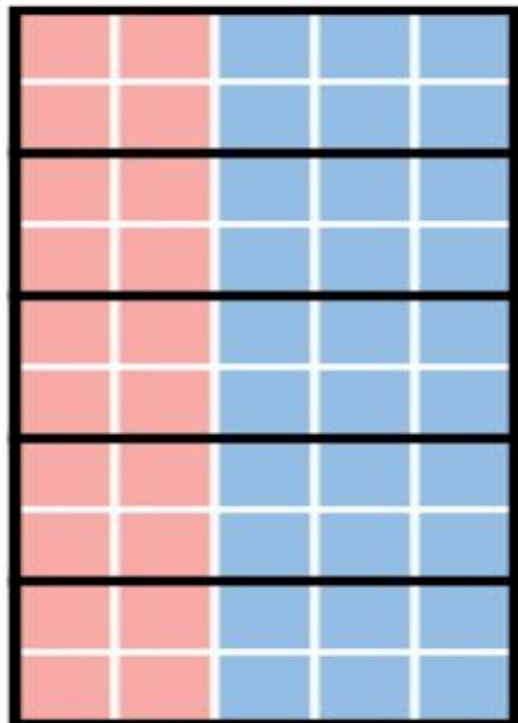
$$P = \frac{1}{K} \sum_{i=1}^N p_i.$$

$$p = \frac{\max_k(P_k) - \min_k(P_k)}{\sum_{k=1}^K P_k}$$

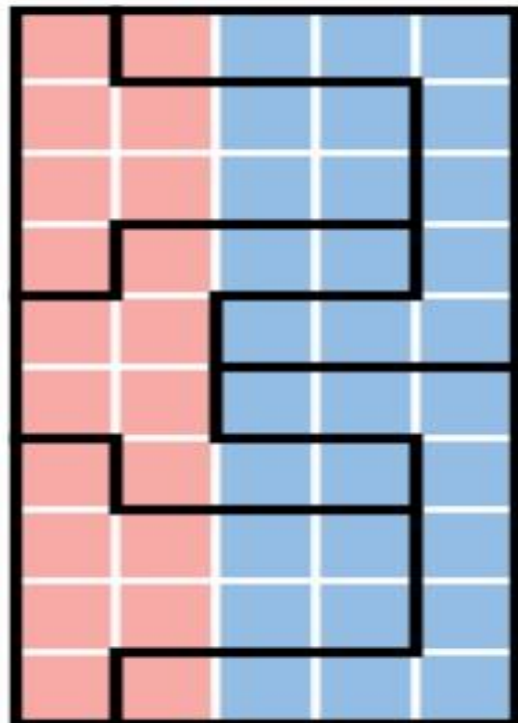
HOW TO STEAL AN ELECTION



50 PRECINCTS
60% BLUE
40% RED



5 DISTRICTS
5 BLUE
0 RED
BLUE WINS



5 DISTRICTS
3 RED
2 BLUE
RED WINS



THE SURVEILLANCE STATE



Home

Products ▾

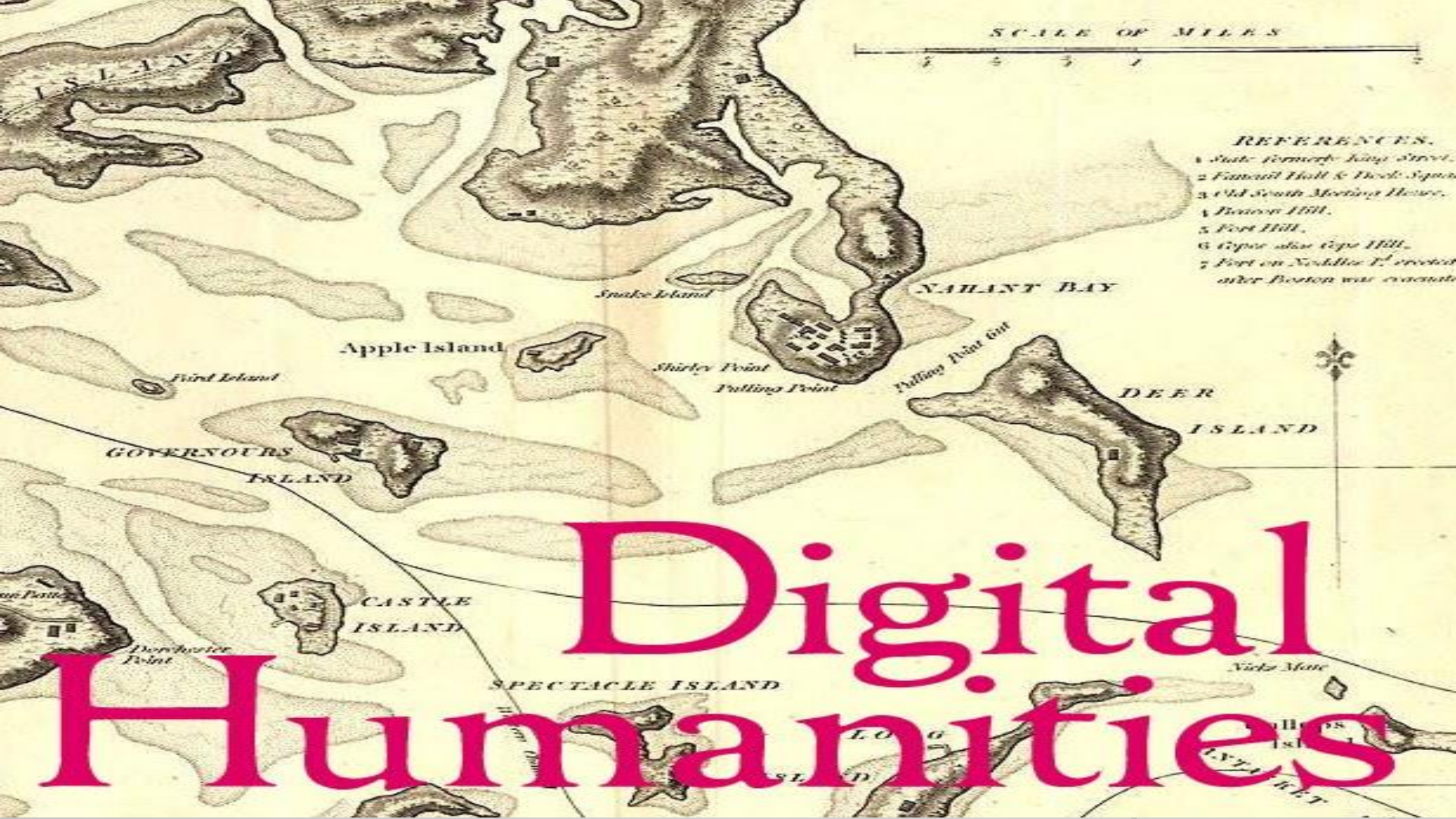
About Gild ▾

**Liberating you from the
challenge of finding developers
using the power of big data.**

[Learn More](#)

LEHMAN BROTHERS

NEW YORK



SCALE OF MILES

REFERENCES.

1. State formerly King Street.
2. Faneuil Hall & Dock Square.
3. Old South Meeting House.
4. Beacon Hill.
5. Fort Hill.
6. Cape also Cape Hill.
7. Fort on Noddies I., erected after Boston was evacuated.

Apple Island

Snake Island

NANTUCKET BAY

Shirley Point

Pulling Point

Nantucket Point

DEER

ISLAND

GOVERNORS

ISLAND

CASTLE
ISLAND

SPECTACLE ISLAND

Nicks Mate

Digital
Humanities



Kjell-Morten

@kjellmorten



Following

Remember the word "dataism" - the (exaggerated) belief in data's ability to give a true description of our reality. @jilltxt at #tedxbergen

← Reply ↻ Retweeted ★ Favorite ... More

RETWEETS

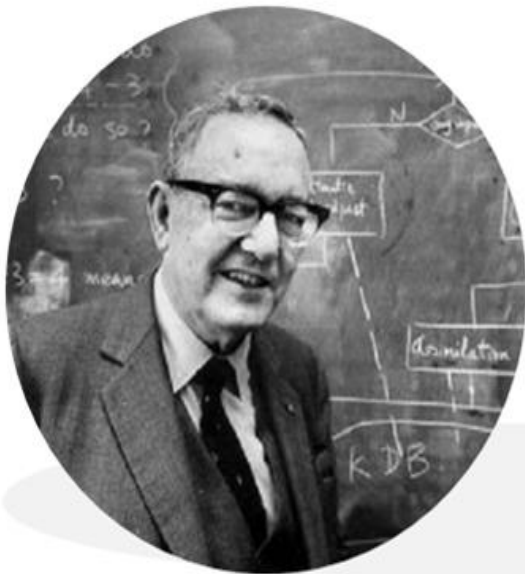
13

FAVORITES

5



2:32 PM - 4 Oct 2014



“

What information consumes is rather obvious. It consumes the attention of its recipients.

Hence, a wealth of information creates a poverty of attention.

Herbert A. Simon

”

JIMI HENDRIX (1942-1970)

You've got to know much more than just the technicalities of notes; you've got to know what goes between the notes."

