Data Science Ethics and Policy

CIS 705

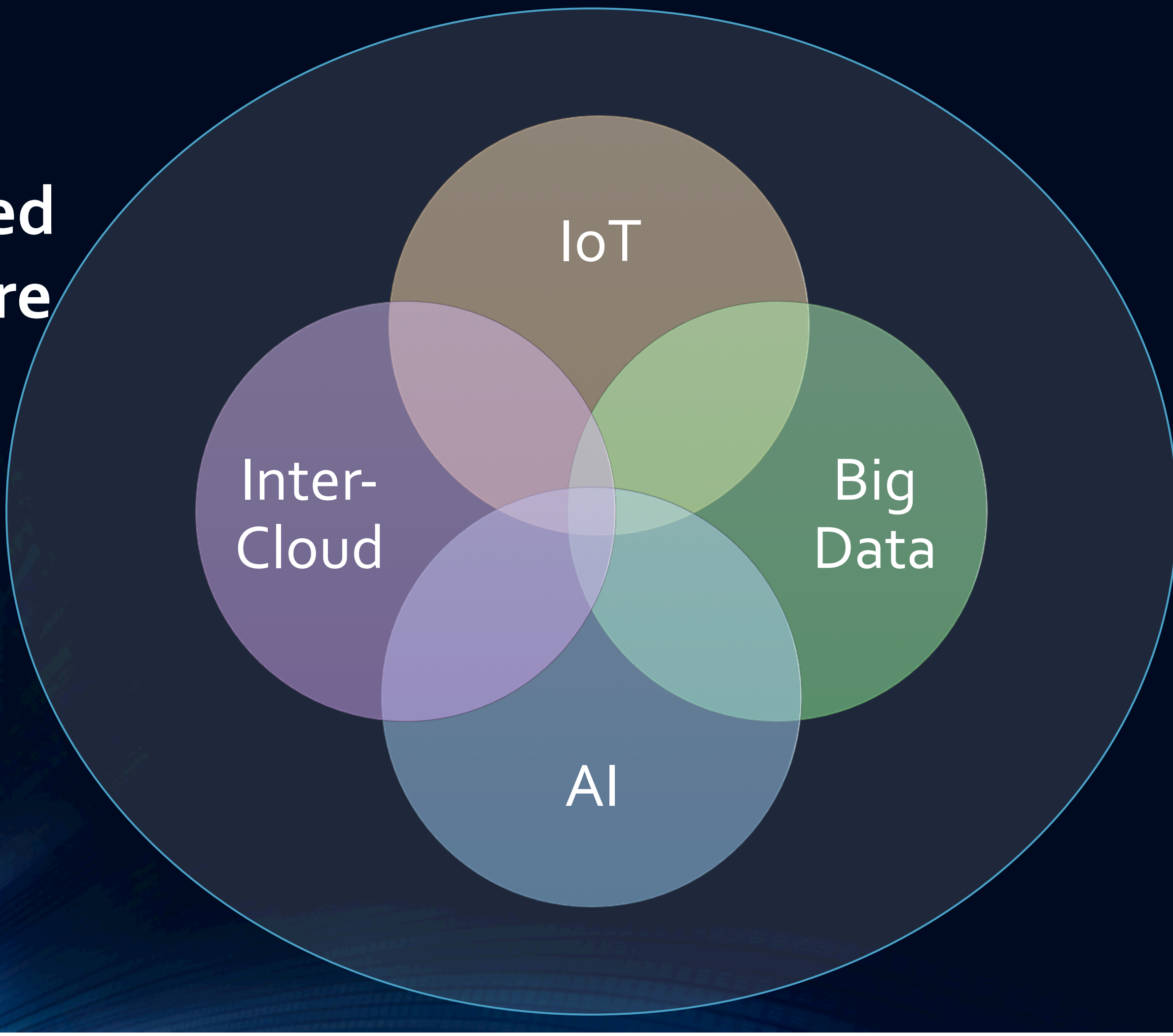
2018

From your learning outcomes

- What are the ethical, societal, and policy issues associated with the use of data science especially with big data?
- What are the privacy issues?
- What are the biases?
- How to use data science responsibly

The Embedded Infosphere

(Taylor, 2016)



EI Features, Functions, and New Governance Problems

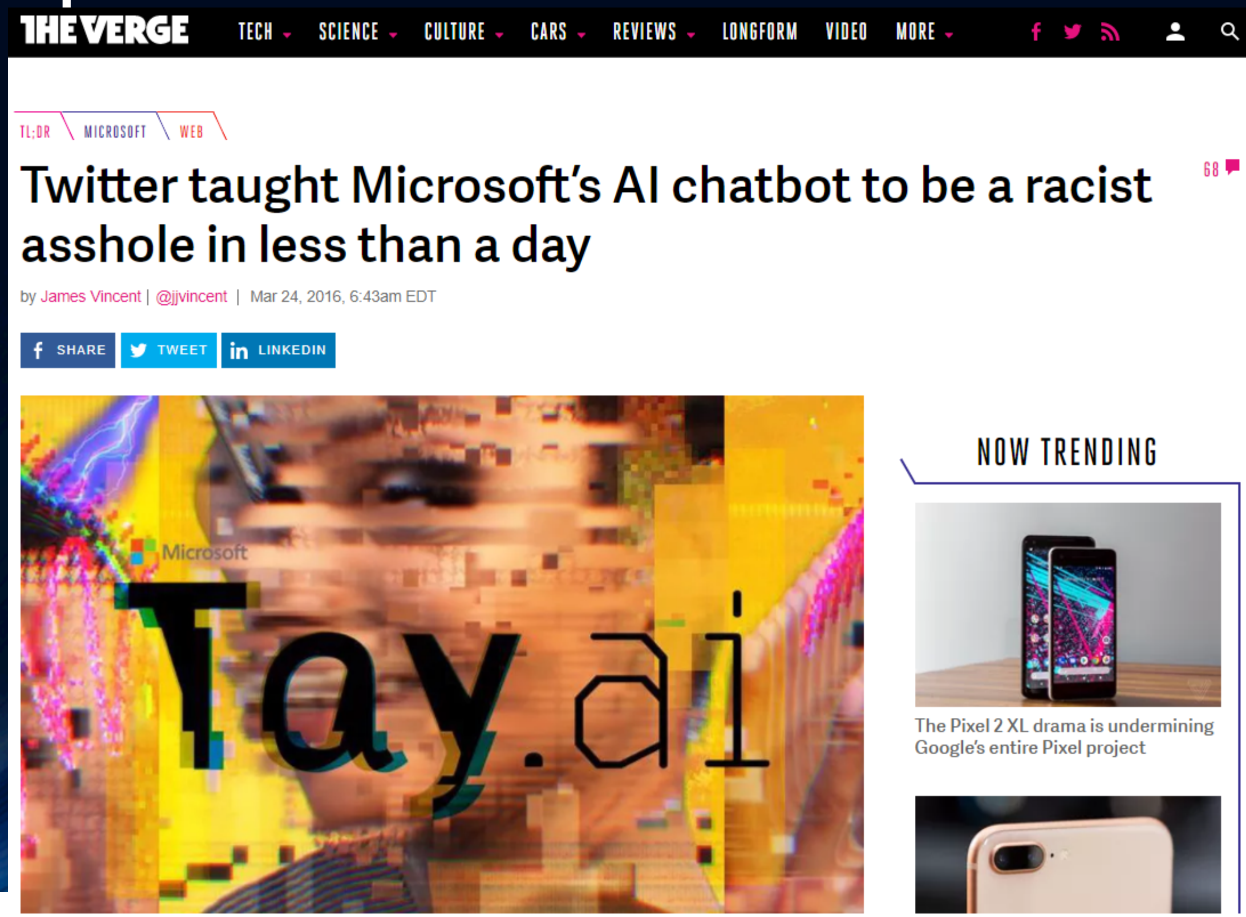
- Each of the components of the EI, individually and working together, have features that raise old and new problems, which are not currently addressed by existing global venues.
- Governance of this area is much more complex and is different than governance of the Internet, which is primarily technical and related to the Domain Name System and other technology.

	Features	Functions	New Governance Issues (examples)
IoT	Billions of networked, smart devices	Perform applications, collect data, interconnect to the cloud, AI and Big Data servers	Privacy; Data security; Unconsented data collection; Data sale and resale; Liability; Data reuse/resale
Big Data	Geographically distributed server farms all over the world	Private companies offer services to public to store, process and apply huge databases	Unjust discrimination; Discriminatory algorithms; Discriminatory data sets; Misuse of data; Transparency of algorithms; Disclosure of protected information (e.g., genetic, health, medical); Appeal/correction of records
Intercloud	Mesh of cloud computing resources geographically distributed across server farms all over the world	Private and public clouds exchange bodies of data directly over private networks	Data stewardship; Data security; Data location ("localism"); Access to data; Legal access to records (extraterritorial jurisdiction); Contracts and jurisdiction; Transborder data flows
AI	Increasing global growth of devices and networks with capacity to autonomously perform activities to achieve goals, or to modify their own behavior in ways that are not externally transparent	This "intelligence" can be built into a device, or accessed over a network. Personal agents, e.g., Alexa	Impacts on industries/employment/government revenues; Legal status of agents for contracts; Legal status of autonomous agents; Duty of safety – precautionary principle; Locus of liability for injury; Jurisdiction; Ethics Code for AI research; Transparency; Non-weaponization; AI rights (potential)

Ethical Questions about Data Analytics

- Who governs data analytics? Who sets the rules? What oversight is there?
- Will the overall benefits of data analytics outweigh social and environmental detriments?
- Will data analytics give one group unfair (and perhaps catastrophic) advantages over another?
- What rules is AI learning as a basis for its decision-making?
- Who is responsible, legally and morally, for harms caused by data analytics?

Example 1: Data and embedded cultural values



THE VERGE TECH SCIENCE CULTURE CARS REVIEWS LONGFORM VIDEO MORE

TL;DR MICROSOFT WEB

68

Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day

by James Vincent | @jjvincent | Mar 24, 2016, 6:43am EDT

SHARE TWEET LINKEDIN

NOW TRENDING

The Pixel 2 XL drama is undermining Google's entire Pixel project

Example 2: The Intelligent Web/Big Nudging

- “You’re reading this from a network, using software, on a device, all of which rely fundamentally on algorithms and AI. The vast portion of the software iceberg that lies beneath the surface, doing its clever but invisible thing, the real building blocks of contemporary computing – are algorithms and AI. Whenever you search, get online recommendations, engage with social media, buy, do online banking, online dating, see online ads; algorithms are doing their devilishly clever work” (Clark, 2017).

Big Nudging

- AI can shape user behavior through “choice architectures” in the design of information systems, influencing their behavior indirectly by “nudging” them towards certain actions (Thaler et al., 2014).
- For example, the combination of nudging and big data analytics (“big nudging”), might be used to motivate citizens to use less energy and protect the natural environment, or it could also be used to influence voters in a way that undermines the democratic process (Helbing et al., 2017).

Example 3: Royal Free and Google DeepMind

- In 2015, Royal Free, one of the largest hospitals in the UK NHS, began working with Deep Mind Health (DMH), a technology startup acquired by Google to develop health applications using advanced analytics and data mining approaches.
- In fall 2015, Royal Free transferred PHI on 1.6 million patients to DMH to test its first application, Streams. Streams provides intelligent alerts, clinical noting, and task management on patients with Acute Kidney Injury (AKI) to physicians and nurses via iPads and iPhones.

Royal Free and Google DeepMind

- This and similar applications developed with DMH have the potential to improve healthcare delivery and services to patients while also measuring and monitoring organizational performance to reduce undesirable variability in health delivery quality (Independent Review Panel, 2017)
- However, there is also a concern about private companies having access to health data. This information is often intensely personal. (Independent Review Panel, 2017, p. 3)

Royal Free and Google DeepMind

- "One of the basic steps in the design of an AI system is the setting of a goal for an agent," argues Rakhal Gaitonde, senior scientist at the Indian Institute of Technology, Madras, and a member of The BMJ's patient advisory group. "Playing Atari, the goal is to maximise your points with least effort," he says. "What would be an equivalent goal in patient care? A sense of wellbeing? The normalising of a set of biochemical parameters? In a situation where we are all only too aware of the way in which corporate interest has conspired to influence the definition of disease, how would one set the goal within the sustainable business model Deep Mind have discussed?" (Armstrong, 2016)

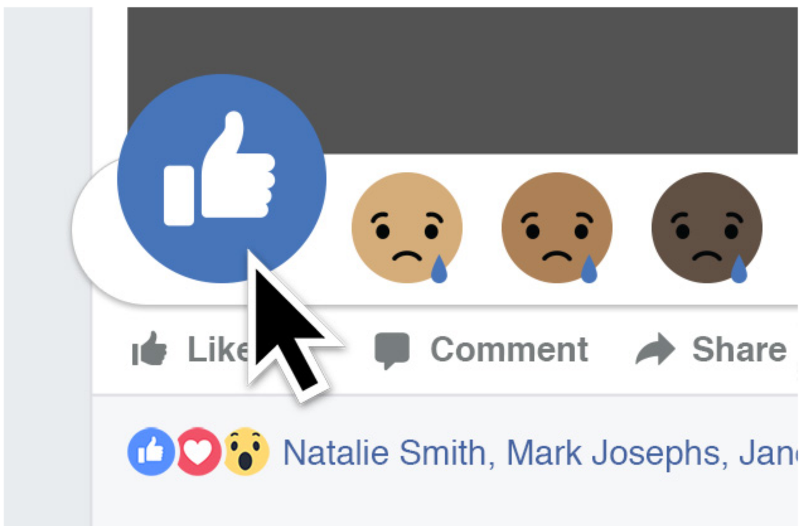
Example 4: Facebook targeted advertising

Facebook Lets Advertisers Exclude Users by Race

Facebook's system allows advertisers to exclude black, Hispanic, and other "ethnic affinities" from seeing ads.

by Julia Angwin and Terry Parris Jr.
ProPublica, Oct. 28, 2016, 8 a.m.

584 Comments | Print



David Sleight/ProPublica

Follow ProPublica

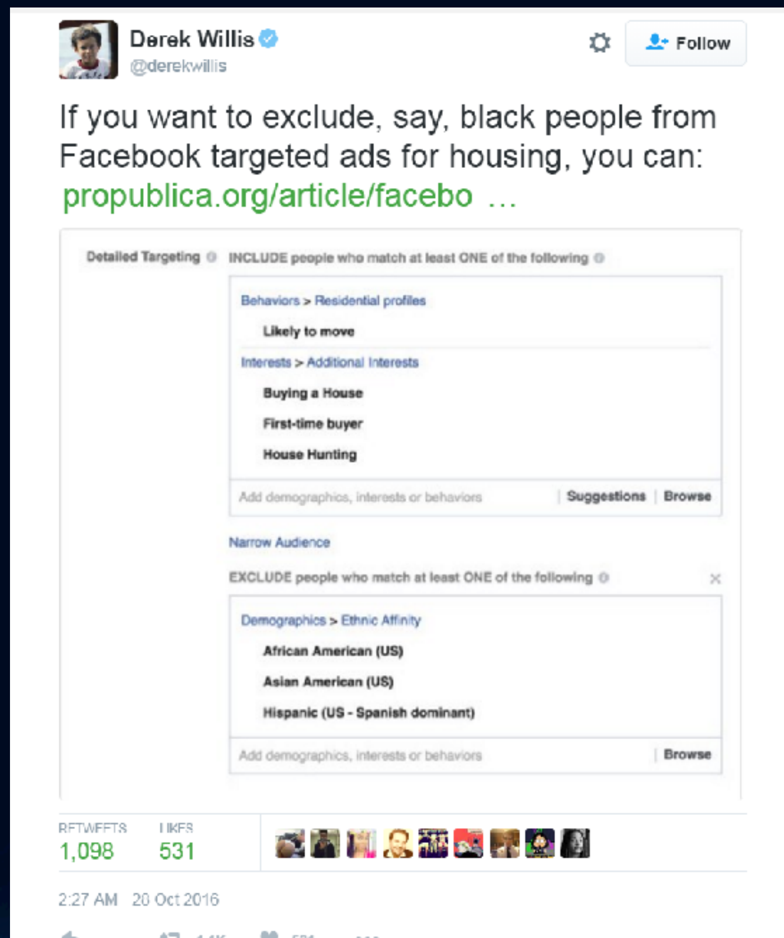
Twitter Facebook
Podcast RSS

Updates by email

Email address

Zip-code optional

SUBSCRIBE


Derek Willis
@derekwillis
Follow
If you want to exclude, say, black people from Facebook targeted ads for housing, you can:
propublica.org/article/facebo …
Detailed Targeting
INCLUDE people who match at least ONE of the following
Behaviors > Residential profiles
Likely to move
Interests > Additional Interests
Buying a House
First-time buyer
House Hunting
Add demographics, interests or behaviors
Suggestions
Browse
Narrow Audience
EXCLUDE people who match at least ONE of the following
Demographics > Ethnic Affinity
African American (US)
Asian American (US)
Hispanic (US - Spanish dominant)
Add demographics, interests or behaviors
Browse
RETWEETS
1,098
LIKES
531
2:27 AM 20 Oct 2016
1.1K
531

Example 5: China's Social Credit System

- China has noted AI (and IoT) as a key next-generation technology in its 13th Five-Year Plan. China has been extremely successful investing in AI, with several large companies, such as Baidu and Tencent, establishing productive research labs (Condliffe, 2017).
- There has been a focus on economic innovation, as well as **security** enabled through AI-intensive tools like **facial recognition** (Li, 2017).

Mashable VIDEO ENTERTAINMENT CULTURE TECH SCIENCE BUSINESS SOCIAL GOOD MORE

World

FOLLOW MASHABLE >

China is using facial recognition to shame jaywalkers

208 SHARES

Share on Facebook Share on Twitter

IMAGE: SHUTTERSTOCK / JOEYPHOTO

The Visibility of Social Class From Facial Cues

RT Bjornsdottir et al. J Pers Soc Psychol 113 (4), 530-546. 2017 May 29. more

Get Full Text: Journal site

Abstract

Social class meaningfully impacts individuals' life outcomes and daily interactions, and the mere perception of one's socioeconomic standing can have significant ramifications. To better understand how people infer others' social class, we therefore tested the legibility of class (operationalized as monetary income) from facial images, finding across 4 participant samples and 2 stimulus sets that perceivers categorized the faces of rich and poor targets significantly better than chance. Further investigation showed that perceivers categorize social class using minimal facial cues and employ a variety of stereotype-related impressions to make their judgments. Of these, attractiveness accurately cued higher social class in self-selected dating profile photos. However, only the stereotype that well-being positively relates to wealth served as a valid cue in neutral faces. Indeed, neutrally posed rich targets displayed more positive affect relative to poor targets and perceivers used this affective information to categorize their social class. Impressions of social class from these facial cues also influenced participants' evaluations of the targets' employability, demonstrating that face-based perceptions of social class may have important downstream consequences. (PsycINFO Database Record

Facial recognition

Face-reading AI will be able to detect your politics and IQ, professor says

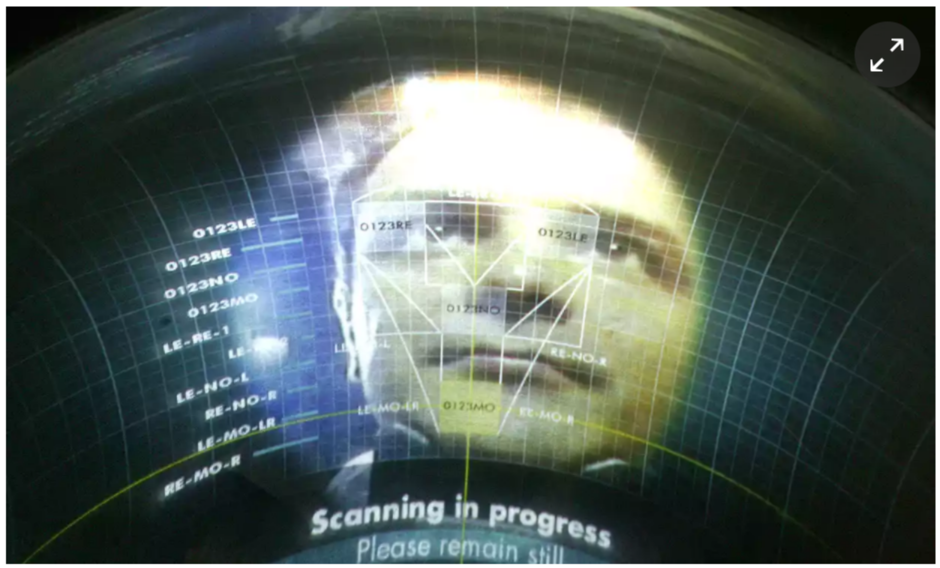
Professor whose study suggested technology can detect whether a person is gay or straight says programs will soon reveal traits such as criminal predisposition

Sam Levin in San Francisco

@SamTLevin

email

Tuesday 12 September 2017 03.00 EDT



Smile for the Camera: Privacy and Policy Implications of Emotion AI

Elaine Sedenberg, John Chuang
School of Information
University of California, Berkeley
{elaine, chuang}@ischool.berkeley.edu

Abstract

We are biologically programmed to publicly display emotions as social cues and involuntary physiological reflexes: grimaces of disgust alert others to poisonous food, pursed lips and furrowed brows warn of mounting aggression, and spontaneous smiles relay our joy and friendship. Though designed to be public under evolutionary pressure, these signals were only seen within a few feet of our compatriots — purposefully fleeting, fuzzy in definition, and rooted within the immediate and proximate social context.

The introduction of artificial intelligence (AI) on visual images for emotional analysis obliterates the natural subjectivity and contextual dependence of our facial displays. This technology may be easily deployed in numerous contexts by diverse actors for purposes ranging from nefarious to socially assistive — like proposed autism therapies. Emotion AI places itself as an algorithmic lens on our digital artifacts and real-time interactions, creating the illusion of a new, objective class of data: our emotional and mental states. Building upon a rich network of existing public photographs — as well as fresh feeds from surveillance footage or smart phone cameras — these emotion algorithms require no additional infrastructure or improvements on image quality.

Privacy and security implications stemming from the collection of emotional surveillance are unprecedented — especially when taken alongside other signals including physiological biosignals (e.g., heartrate or body temperature). Emotion AI also presents new methods to manipulate individuals by targeting political propaganda or fish for passwords based on micro-reactions. The lack of transparency or notice on these practices makes public inquiry unlikely, if not impossible.

Social Credit System

- Chinese government recently developed a “social credit” (shehui xinyong 社会信用) system
- “AI, like most other technologies, can serve the interests both of promoting individual autonomy and government accountability and of entrenching authoritarian social control” (Chen et al., 2018)

Basic Information

Full Name
Car(s) Info
ID number
ID photo
Headshot
Email Address
Education
Provident Fund Info
Login Info (Device & Network)

Products / Services

Shopping
Car Rental
Transfer/Payments
Loan Transactions
Credit Card Transactions

Government (Payment Records)

Tax
Social Insurance
Provident Fund

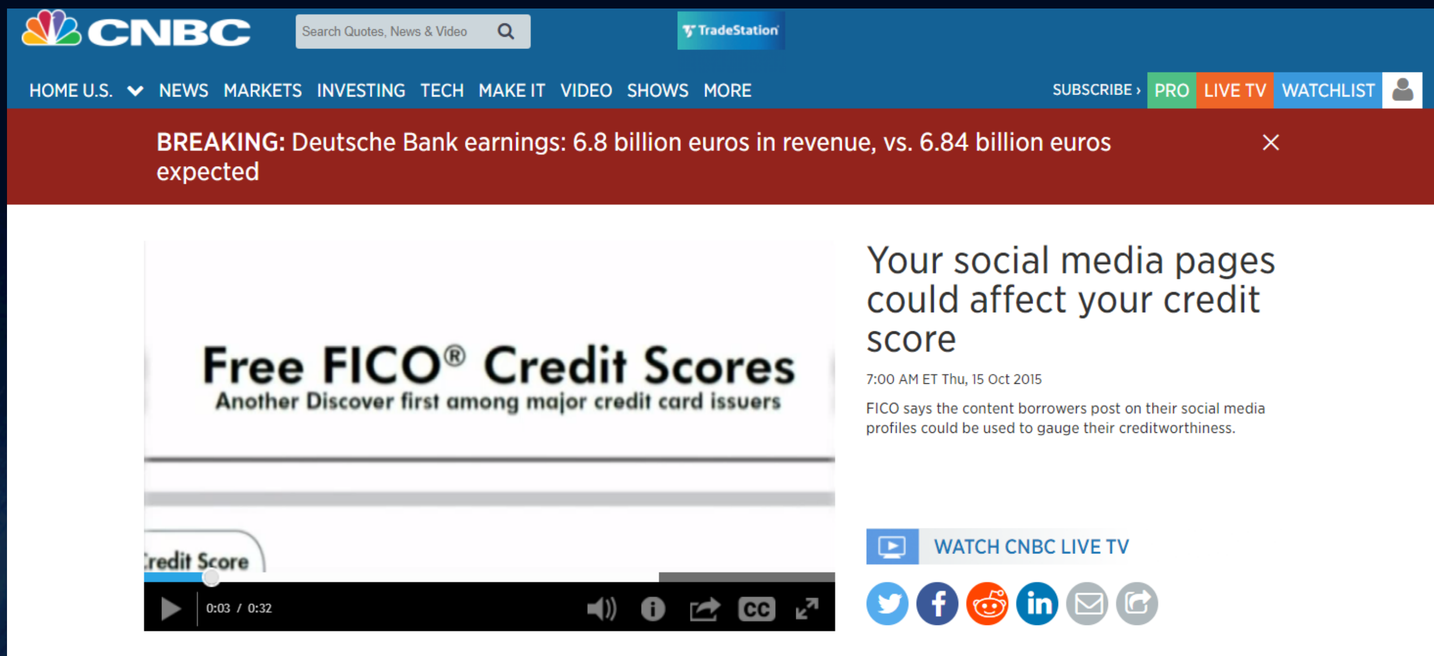
The Judiciary & Government

A list of citizens who refuse to repay debt

A list of citizens who have committed offences

Data such as income, savings, securities, commercial insurance, assets and tax info can only be collected with users' explicit written consent. No attempt is made to collect data on users' religion, genes, fingerprints, blood type, diseases, medical record, social media posts, messaging or voice call record.

Source: Zhima Credit User Agreement


CNBC
Search Quotes, News & Video
TradeStation
HOME U.S.
NEWS
MARKETS
INVESTING
TECH
MAKE IT
VIDEO
SHOWS
MORE
SUBSCRIBE ›
PRO
LIVE TV
WATCHLIST
BREAKING: Deutsche Bank earnings: 6.8 billion euros in revenue, vs. 6.84 billion euros expected
Free FICO® Credit Scores
Another Discover first among major credit card issuers
0:03 / 0:32
Your social media pages could affect your credit score
7:00 AM ET Thu, 15 Oct 2015
FICO says the content borrowers post on their social media profiles could be used to gauge their creditworthiness.
WATCH CNBC LIVE TV

Example 5: COMPAS

Key Challenges

- Scale and scope - Deep learning algorithms rely on massive data sets to train and improve AI models. The growing stockpiles of digitized data available through research and commercial arrangements provide essential fuel for deep learning innovation in several sectors.
- Deep learning algorithms also create a whole new category of predictive data about individual and group behaviors.

Key Challenges

- Opacity – Deep learning is a "black box" , with its inner workings obscured by the opacity and complexity of algorithms.
  - In some cases, opacity may be an "intentional corporate or institutional self-protection and concealment" (Burrell, 2016, pp. 1-2), particularly to protect corporate intellectual property and competitive advantage.
  - In others, those examining an AI system may lack the specialized coding skills to understand its processes.
  - A third type of opacity, characteristic of deep learning, relates to the scale and complexity of machine learning, and thus to humans' difficulty understanding an algorithm in action as it reads real-time data and adapts (Burrell, 2016).

How to use data science responsibly?

Discussion