BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Anxiety, Crisis, and a Computational Future for Journalism

Momin M. Malik, PhD <momin_malik@cyber.harvard.edu>
Data Science Postdoctoral Fellow
Berkman Klein Center for Internet & Society at Harvard University

University of Maryland, Philip Merrill College of Journalism / College of Information Studies
November 27, 2018

Slides: https://mominmalik.com/merrill.pdf (full reference list and image credits)


#NOTHELPING
WHAT WILL BECOME OF US

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

3 futures for computation + journalism

Journalism about computation: Informed reporting about AI should treat it skeptically, breaking down its language and claims

Computation *about* journalism: At-scale studies to support qualitative and small-scale inquiry into the news ecosystem

Journalism *with* computation: Not just stories in data, but the interrogating the data as the story

(Computation *for* journalism: limited, I argue)

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

My story

- DEPARTMENT OF THE HISTORY OF SCIENCE HARVARD UNIVERSITY

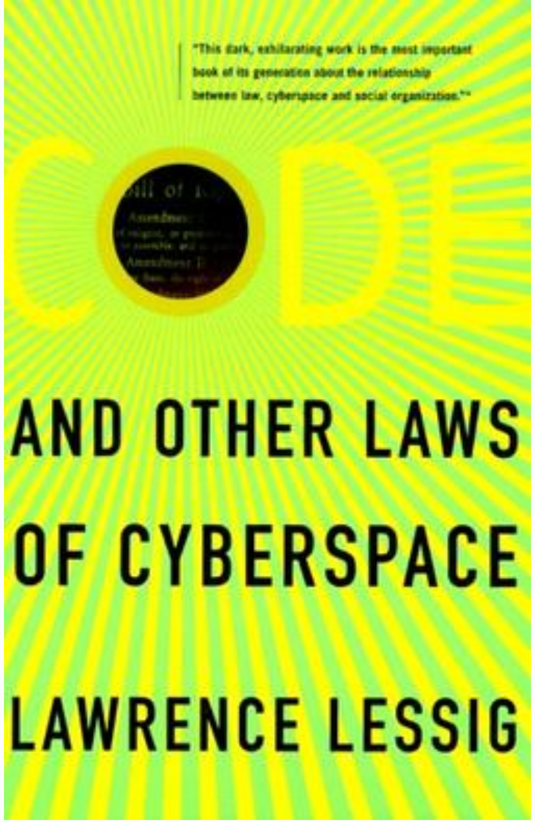
"Everything you think you know about science is wrong."

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

My story

- DEPARTMENT OF THE HISTORY OF SCIENCE HARVARD UNIVERSITY
- Berkman The Berkman Center for Internet & Society at Harvard University



My story

- DEPARTMENT OF THE HISTORY OF SCIENCE HARVARD UNIVERSITY
- Berkman
  The Berkman Center for Internet & Society
  at Harvard University
  - *Reacting* to the creations of technologists

My story

- DEPARTMENT OF THE HISTORY OF SCIENCE, HARVARD UNIVERSITY
- Berkman — The Berkman Center for Internet & Society at Harvard University
  - *Reacting* to the creations of technologists
- OXFORD INTERNET INSTITUTE — UNIVERSITY OF OXFORD

Sociology
Copyright © 2007
BSA Publications Ltd®
Volume 41(5): 885–899
DOI: 10.1177/0038038507080443
SAGE Publications
Los Angeles, London, New Delhi and Singapore

The Coming Crisis of Empirical Sociology

Mike Savage
University of Manchester

Roger Burrows
University of York

ABSTRACT
This article argues that in an age of *knowing capitalism*, sociologists have not adequately thought about the challenges posed to their expertise by the proliferation of 'social' transactional data which are now routinely collected, processed and analysed by a wide variety of private and public institutions. Drawing on British examples, we argue that whereas over the past 40 years sociologists championed innovative methodological resources, notably the sample survey and the in-depth interviews, which reasonably allowed them to claim distinctive expertise to access

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

My story

- DEPARTMENT OF THE HISTORY OF SCIENCE, HARVARD UNIVERSITY
- Berkman — The Berkman Center for Internet & Society at Harvard University
  - *Reacting* to the creations of technologists
- OXFORD INTERNET INSTITUTE — UNIVERSITY OF OXFORD

BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

My story

- DEPARTMENT OF THE HISTORY OF SCIENCE HARVARD UNIVERSITY
- Berkman — The Berkman Center for Internet & Society at Harvard University
  - *Reacting* to the creations of technologists
- OXFORD INTERNET INSTITUTE — UNIVERSITY OF OXFORD
- Berkman — The Berkman Center for Internet & Society at Harvard University

My story

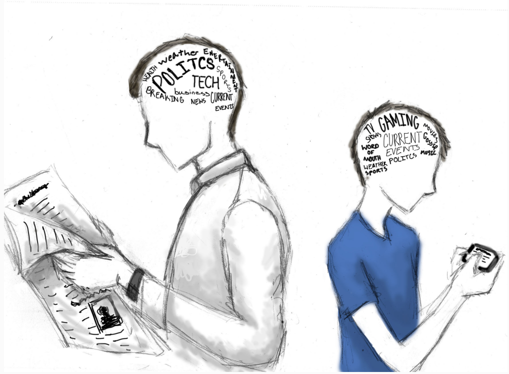
- DEPARTMENT OF THE HISTORY OF SCIENCE, HARVARD UNIVERSITY
- Berkman — The Berkman Center for Internet & Society at Harvard University
  - *Reacting* to the creations of technologists
- OXFORD INTERNET INSTITUTE — UNIVERSITY OF OXFORD
- Berkman — The Berkman Center for Internet & Society at Harvard University

Youth and Media — Berkman, The Berkman Center for Internet & Society at Harvard University

The Challenges of Defining 'News Literacy'

July 8, 2013

Momin Malik, Sandra Cortesi, and Urs Gasser *



* Momin Malik, Research Assistant at the Berkman Center for Internet & Society.
Sandra Cortesi, Director of Youth and Media at the Berkman Center for Internet & Society.
Urs Gasser, Executive Director of the Berkman Center for Internet & Society.

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

My story

- DEPARTMENT OF THE HISTORY OF SCIENCE, HARVARD UNIVERSITY
- Berkman: The Berkman Center for Internet & Society at Harvard University
  - *Reacting* to the creations of technologists
- OXFORD INTERNET INSTITUTE; UNIVERSITY OF OXFORD
- Berkman: The Berkman Center for Internet & Society at Harvard University
- Carnegie Mellon University School of Computer Science

- Law. Library science. Sociology. Journalism.
- Anxiety about being *replaced* by computer science.
- Crises of jobs and funding.

The desert and the swamp (Arthur Leff)

Cameron Wears, https://flic.kr/p/yQ7xZd

Ekaterina Vladinakova, https://flic.kr/p/UKbXfk

Journalism about computation

- Reporting on "prediction" in machine learning and AI

Predict...

MIT Technology Review

Topics+ The Download Magazine Events

The Download

What's up in emerging technology

April 13, 2018

Facebook is using AI to predict users' future behavior and selling that data to advertisers

In confidential documents seen by the *Intercept*, Facebook touts its ability to "improve" marketing outcomes with what it calls "loyalty prediction."

Newspeak: The AI software that powers this capability, called "FBLearner Flow," was first announced in... **Read more**

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Predict speech

MIT Technology Review

Topics+ The Download Magazine Events

Rewriting Life

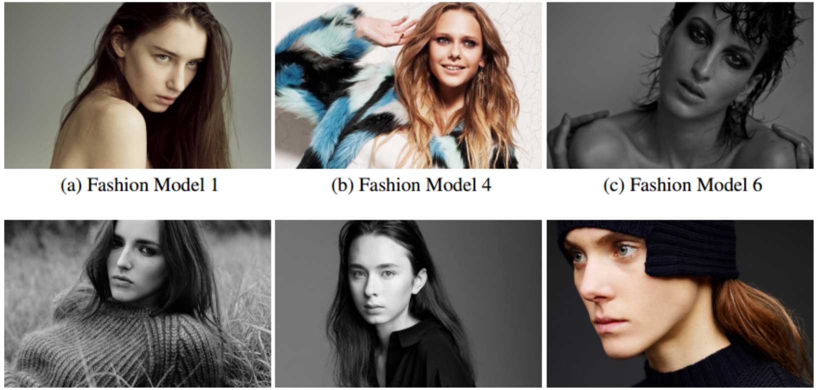
Scientists Can Read a Bird's Brain and Predict Its Next Song

Next up, predicting human speech with a brain-computer interface.

by Antonio Regalado October 11, 2017

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Predict fashion models success



(a) Fashion Model 1 (b) Fashion Model 4 (c) Fashion Model 6
(d) Fashion Model 7 (e) Fashion Model 8 (f) Fashion Model 9

MIT Technology Review

Topics+ The Download Magazine Events

A View from **Emerging Technology from the arXiv**

Machine Learning Algorithm Predicts Which New Faces Will Make It as Fashion Models

A machine-learning algorithm picks out the fashion models most likely to succeed.

September 1, 2015

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Predict *news*

MIT Technology Review

Topics+ The Download Magazine Events

Intelligent Machines

Software Predicts Tomorrow's News by Analyzing Today's and Yesterday's

Prototype software can give early warnings of disease or violence outbreaks by spotting clues in news reports.

by Tom Simonite February 1, 2013

A method of using online information to accurately predict the future could transform many industries.

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Predict… the future?

Predicting the Future With Social Media

Sitaram Asur
Social Computing Lab
HP Labs
Palo Alto, California
Email: sitaram.asur@hp.com

Bernardo A. Huberman
Social Computing Lab
HP Labs
Palo Alto, California
Email: bernardo.huberman@hp.com

Mar 2010

***Abstract*—In recent years, social media has become ubiquitous and important for social networking and content sharing. And yet, the content that is generated from these websites remains largely untapped. In this paper, we demonstrate how social media content can be used to predict real-world outcomes. In particular, we use the chatter from Twitter.com to forecast box-office revenues for movies. We show that a simple model built from**

This paper reports on such a study. Specifically we consider the task of predicting box-office revenues for movies using the chatter from Twitter, one of the fastest growing social networks in the Internet. Twitter [1], a micro-blogging network, has experienced a burst of popularity in recent months leading to a huge user-base, consisting of several tens of millions of

Predicting the Future — Big Data, Machine Learning, and Clinical Medicine

Ziad Obermeyer, M.D., and Ezekiel J. Emanuel, M.D., Ph.D.

By now, it's almost old news: big data will transform medicine. It's essential to remember, however, that data by themselves are useless. To be useful, data must be analyzed, interpreted, and acted on. Thus, it is algorithms — not data sets — that will prove transformative. We believe, therefore, that attention has to shift to new statistical tools from the field of machine learning that will be critical for anyone practicing medicine in the 21st century.

First, it's important to understand what machine learning is not. Most computer-based algorithms in medicine are "expert systems" — rule sets encoding knowledge on a given topic, which are applied to draw conclusions

1216 N ENGL J MED 375;13 NEJM.ORG SEPTEMBER 29, 2016

Merriam-Webster SINCE 1828

predict *verb*

pre·dict | \pri-ˈdikt \

predicted; **predicting**; **predicts**

Definition of *predict*

transitive verb

: to declare or indicate in advance

especially : foretell on the basis of observation, experience, or scientific reason

intransitive verb

: to make a prediction

Other Words from *predict*

Synonyms

Choose the Right Synonym

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Prediction is not prediction

"I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper"

A Balanced Survey on Election Prediction using Twitter Data

Daniel Gayo-Avello
dani@uniovi.es
@PFCdgayo
Department of Computer Science - University of Oviedo (Spain)

May 1, 2012

Abstract

Predicting X from Twitter is a popular fad within the Twitter research subculture. It seems both appealing and relatively easy. Among such kind of studies, electoral prediction is maybe the most attractive, and at this moment there is a growing body of literature on such a topic.

This is not only an interesting research problem but, above all, it is extremely difficult. However, most of the authors seem to be more interested in claiming positive results than in providing sound and reproducible methods.

It is also especially worrisome that many recent papers seem to only acknowledge those studies supporting the idea of Twitter predicting elections, instead of conducting a balanced literature review showing both sides of the matter.

After reading many of such papers I have decided to write such a survey myself. Hence, in this paper, every study relevant to the matter of electoral prediction using social media is commented.

From this review it can be concluded that the predictive power of Twitter regarding elections has been greatly exaggerated, and that hard research problems still lie ahead.

Introduction

In the last two years a number of papers have suggested that Twitter data has an impressive predictive power. Apparently, everything from the stock market

"*It's not prediction at all!* I have not found a single paper predicting a future result. All of them claim that a prediction could have been made; i.e. they are *post-hoc* analysis and, needless to say, negative results are rare to find."

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

What "prediction" really is

- Prediction = "Fitted value"
- If a model generalizes, then fitted values can predict
- But lots can go wrong
- "Prediction" is done based entirely on correlations.
- Spurious (non-causal) correlations can fit really well
- But are also fragile.

POLICYFORUM

BIG DATA

The Parable of Google Flu: Traps in Big Data Analysis

Large errors in flu prediction were largely avoidable, which offers lessons for the use of big data.

David Lazer,[1,2]* Ryan Kennedy,[1,3,4] Gary King,[3] Alessandro Vespignani[5,6,3]

In February 2013, Google Flu Trends (GFT) made headlines but not for a reason that Google executives or the creators of the flu tracking system would have hoped. *Nature* reported that GFT was predicting more than double the proportion of doctor visits for influenza-like illness (ILI) than the Centers for Disease Control and Prevention (CDC), which bases its estimates on surveillance reports from laboratories across the United States (*1*, *2*). This happened despite the fact that GFT was built to predict CDC reports. Given that GFT is often held up as an exemplary use of big data (*3*, *4*), what lessons can we draw from this error?

The problems we identify are not limited to GFT. Research on whether search or social media can predict *x* has become commonplace (*5*–*7*) and is often put in sharp contrast with traditional methods and hypotheses. Although these studies have shown the value of these data, we are far from a place where they can supplant more traditional methods or theories (*8*). We explore two issues that contributed to GFT's mistakes—big data hubris and algorithm dynamics—and offer lessons for moving forward in the big data age.

Big Data Hubris

"Big data hubris" is the often implicit assumption that big data are a substitute for, rather than a supplement to, traditional

ability and dependencies among data (*12*). The core challenge is that most big data that have received popular attention are not the output of instruments designed to produce valid and reliable data amenable for scientific analysis.

The initial version of GFT was a particularly problematic marriage of big and small data. Essentially, the methodology was to find the best matches among 50 million search terms to fit 1152 data points (*13*). The odds of finding search terms that match the propensity of the flu but are structurally unrelated, and so do not predict the future, were quite high. GFT developers,

run ever since, with a few changes announced in October 2013 (*10*, *15*).

Although not widely reported until 2013, the new GFT has been persistently overestimating flu prevalence for a much longer time. GFT also missed by a very large margin in the 2011–2012 flu season and has missed high for 100 out of 108 weeks starting with August 2011 (see the graph). These errors are not randomly distributed. For example, last week's errors predict this week's errors (temporal autocorrelation), and the direction and magnitude of error varies with the time of year (seasonality). These patterns mean that GFT overlooks considerable information that could be extracted by traditional statistical methods.

Even after GFT was updated in 2009, the comparative value of the algorithm as a stand-alone flu monitor is questionable. A study in 2010 demonstrated that GFT accuracy was not much better than a fairly simple projection forward using already available (typically on a 2-week lag) CDC data (*4*). The comparison has become even worse since that time, with lagged models significantly outperforming GFT (see the graph). Even 3-week-old CDC data do a better job of projecting current flu prevalence than GFT [see supplementary materials (SM)].

Considering the large number of approaches that provide inference on influ-

"Wishful mnemonics" of AI

My story

Reporting on "prediction" in machine learning/AI

At-scale study of news media on Twitter

Modeling algorithmic user manipulation

Priorities, current work, and goals

References

ARTIFICIAL INTELLIGENCE MEETS NATURAL STUPIDITY

Drew McDermott
MIT AI Lab Cambridge, Mass 02139

As a field, artificial intelligence has always been on the border of respectability, and therefore on the border of crackpottery. Many critics <Dreyfus, 1972>, <Lighthill, 1973> have urged that we are over the border. We have been very defensive toward this charge, drawing ourselves up with dignity when it is made and folding the cloak of Science about us. On the other hand, in private, we have been justifiably proud of our ideas, because pursuing them is the onl

Unfortunately, the necessity for s the culture of the hacker in computer to cripple our self-discipline. In a young field, self-discipline is not necessarily a virtue, but we are not getting any younger. In the past few years, our tolerance of sloppy thinking has led us to repeat many mistakes over and over. If we are to retain any credibility, this should stop.

This paper is an effort to ridicule some of these mistakes. Almost everyone I know should find himself the target at some point or other; if you don't, you are encouraged to write up your own favorite fault. The three described here I suffer from myself. I hope self-ridicule will be a complete catharsis, but I doubt it. Bad

Wishful Mnemonics

ble.

m crowd for

Wishful Mnemonics

A major source of simple-mindedness in AI programs is the use of mnemonics like "UNDERSTAND" or "GOAL" to refer to programs and data structures. This practice has been inherited from more

Compare the mnemonics in Planner <Hewitt,1972> with those in Conniver <Sussman and McDermott, 1972>:

Planner	Conniver
GOAL	FETCH & TRY-NEXT
CONSEQUENT	IF-NEEDED
ANTECEDENT	IF-ADDED
THEOREM	METHOD
ASSERT	ADD

It is so much harder to write programs using the terms on the right! When you say (GOAL . . .), you can just feel the enormous power at your fingertips. It is, of course, an illusion.

When you say (GOAL . . .), you can just feel the enormous power at your fingertips. It is, of course, an illusion.

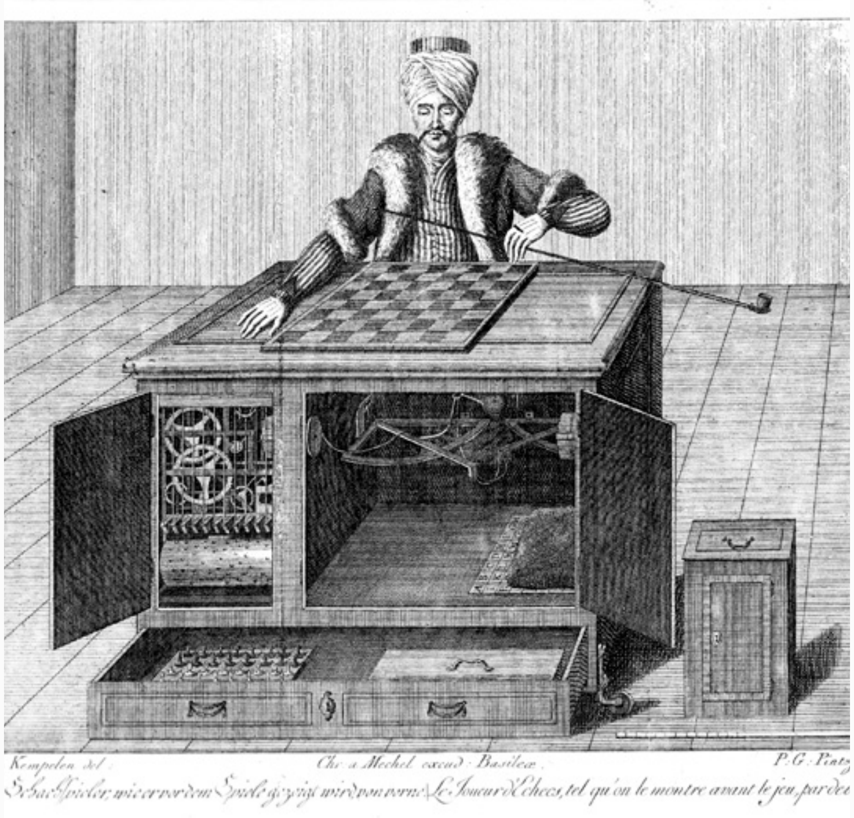
1965> What if atomic symbols had been called "concepts", or CONS had been called ASSOCIATE? As it is, the programmer has no debts to pay to the system. He can build whatever he likes. There are some minor faults; "property lists" are a little risky; but by now the term is sanitized.

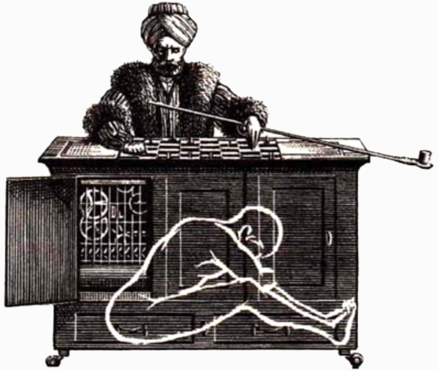
Resolution theorists have been pretty good about wishful mnemonics. They thrive on hitherto meaningless words like RESOLVE and PARAMODULATE, which can only have their humble, technical meaning. There are actually quite few pretensions in the resolution literature. <Robinson, 1965> Unfortunately, at the top of their intellectual edifice stand the word "deduction". This is *very* wishful, but not entirely their fault. The logicians who first misused the term (e.g., in the "deduction" theorem) didn't have our problems; pure resolution theorists don't either. Unfortunately, too many AI researchers took them at their word and assumed that deduction, like payroll processing, had been tamed.

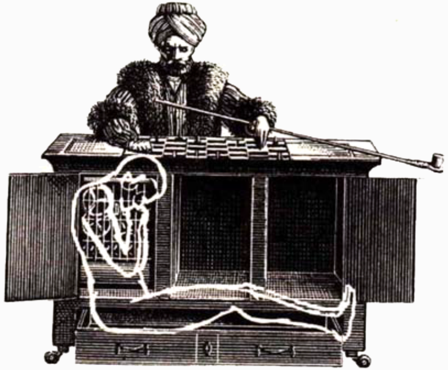
Of course, as in many such cases, the only consequence in the long run was that "deduction" changed in meaning, to become something narrow, technical, and not a little sordid.


#NOTHELPING
WHAT WILL BECOME OF US

The "Chess Turk" as precedent





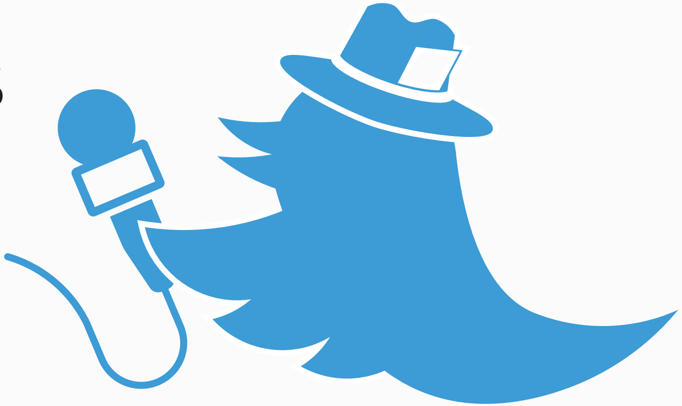


BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

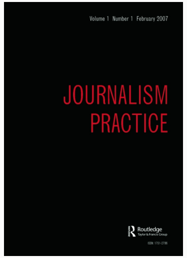
Computation about journalism

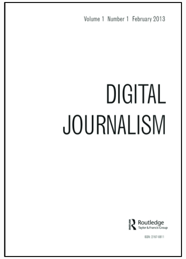
- Studying the news media's use of Twitter at scale



With Twitter, study news on Twitter

- First, use "decahose" (10%) access to estimate volume/focus of news media-related content
- 





Second: What do academic journalism researchers care about?
- Target: "Push" model of Twitter usage (Armstrong and Gao 2010; Holcomb, Gross, and Mitchell 2011; Vis 2013)

Study of news on Twitter at scale

- *Digital Journalism*, 2016
- Bob Franklin, "The Future of Journalism: Risks, Threats, and Opportunities" (2016):

  "the first issue of *Digital Journalism* reported a study involving an automated content analysis of 2,490,429 articles... three years later, [Malik & Pfeffer] presented findings based on a sample of 1.8 billion tweets..."

A MACROSCOPIC ANALYSIS OF NEWS CONTENT IN TWITTER

Momin M. Malik and **Jürgen Pfeffer**

Previous literature has considered the relevance of Twitter to journalism, for example as a tool for reporters to collect information and for organizations to disseminate news to the public. We consider the reciprocal perspective, carrying out a survey of news media-related content within Twitter. Using a random sample of 1.8 billion tweets over four months in 2014, we look at the distribution of activity across news media and the relative dominance of certain news organizations in terms of relative share of content, the Twitter behavior of news media, the hashtags used in news content versus Twitter as a whole, and the proportion of Twitter activity that is news media-related. We find a small but consistent proportion of Twitter is news media-related (0.8 percent by volume); that news media-related tweets focus on a different set of hashtags than Twitter as a whole, with some hashtags such as those of countries of conflict (Arab Spring countries, Ukraine) reaching over 15 percent of tweets being news media-related; and we find that news organizations' accounts, across all major organizations, largely use Twitter as a professionalized, one-way communication medium to promote their own reporting. Using Latent Dirichlet Allocation topic modeling, we also examine how the proportion of news content varies across topics within 100,000 #Egypt tweets, finding that the relative proportion of news media-related tweets varies vastly across different subtopics. Over-time analysis reveals that news media were among the earliest adopters of certain #Egypt subtopics, providing a necessary (although not sufficient) condition for influence.

KEYWORDS computational; news media; social media; topic modeling; Twitter

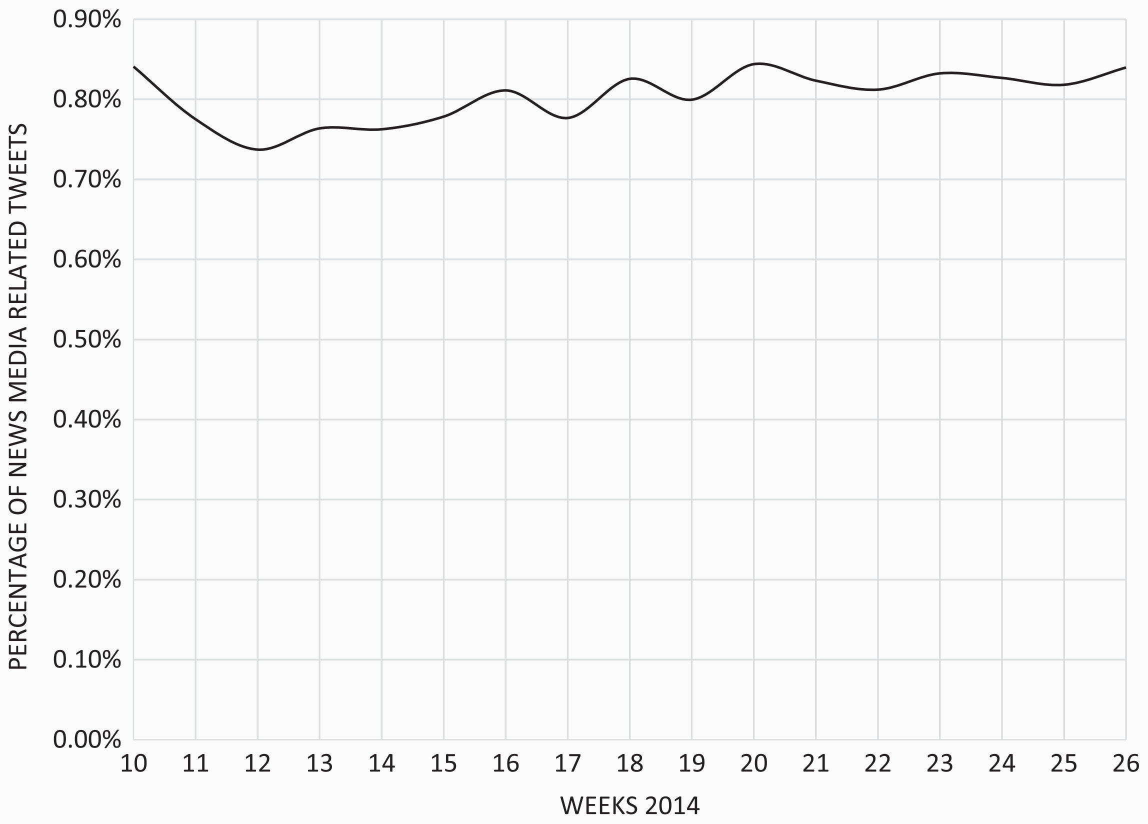
Introduction

The rise of the internet and social media has been involved with a crisis in journalism (Chouliaraki and Blaagaard 2013; Franklin 2012; Hirst 2010; McChesney 2012; Picard

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

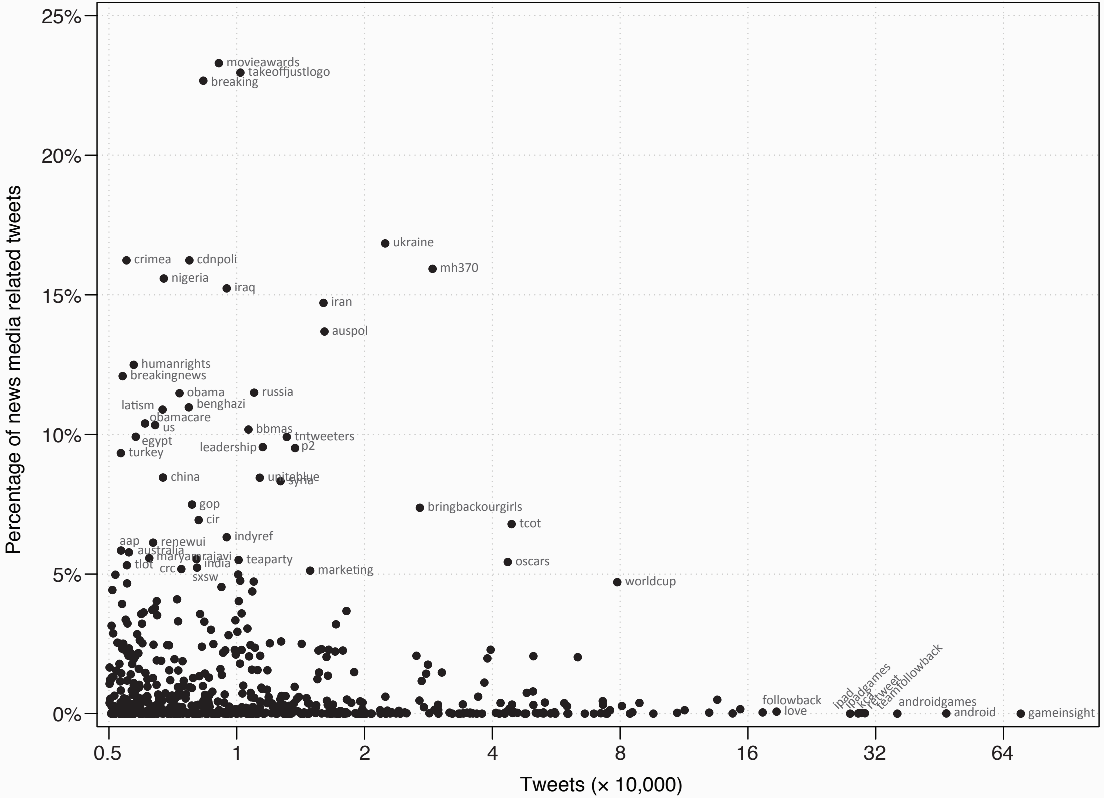
- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

~0.8% of Twitter is news media related



BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

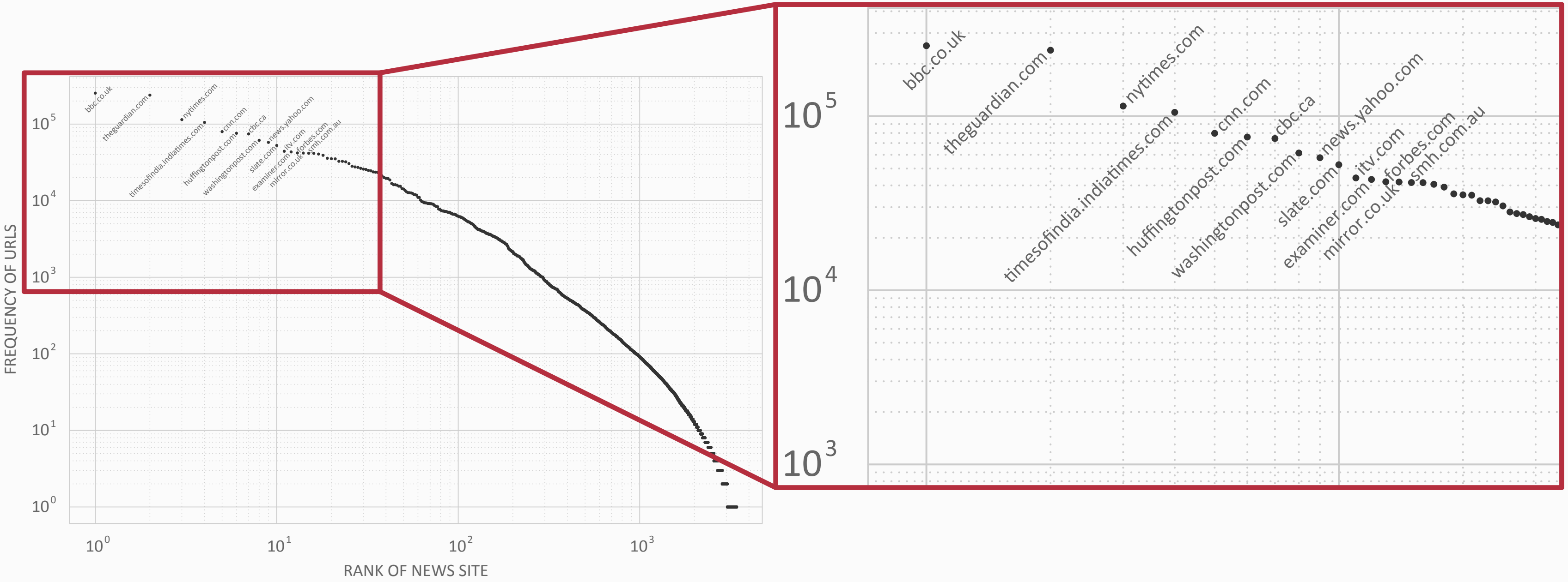
News covers totally separate topics



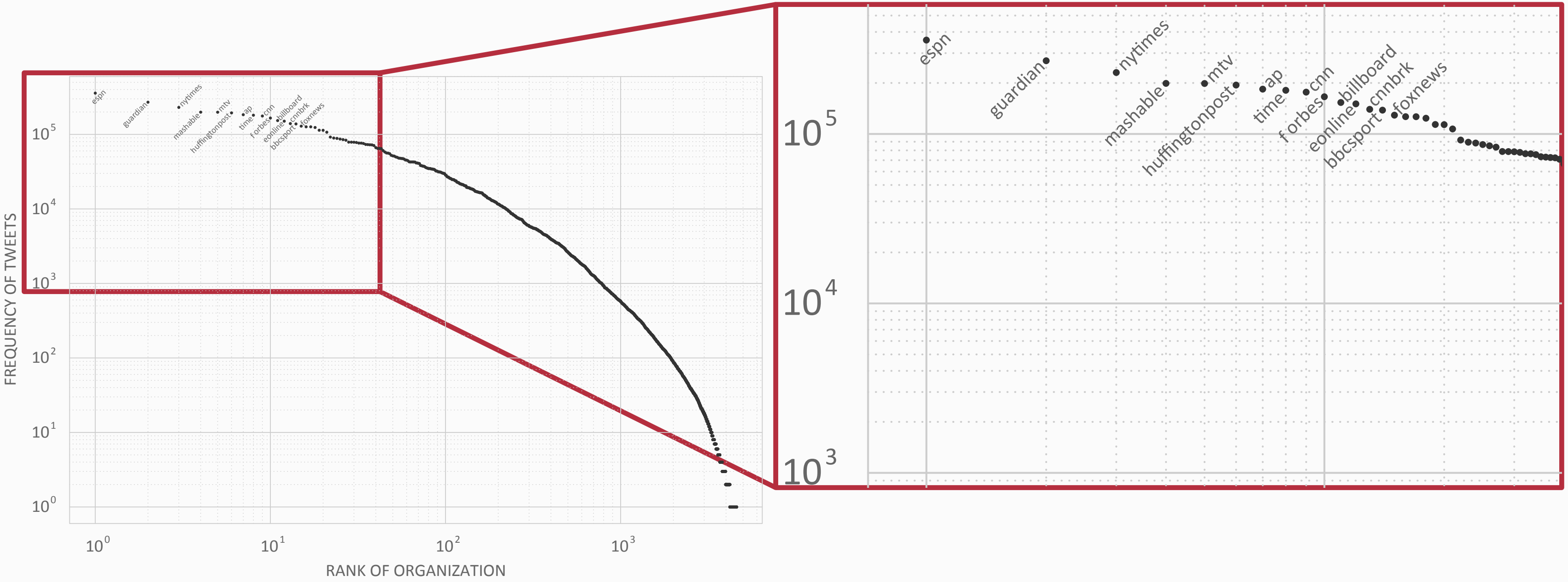
BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

URLs: Major newsrooms dominate



Activity: Top 51 account for about 50%



Evidence at scale for the "push" model

- Sample 1000 tweets from the top 51
- 89.7% mention (@) the organization itself, or link to the organization's website
- Confirms, via sampling from the entire population, that the largest-volume news org use of Twitter is to drive traffic to websites

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Journalism with computation

- Modeling to show "algorithmic user manipulation"

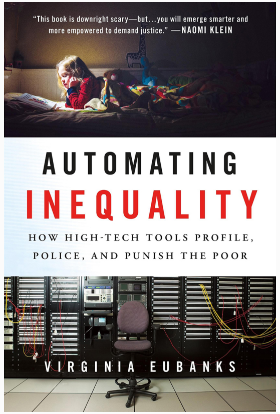
- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

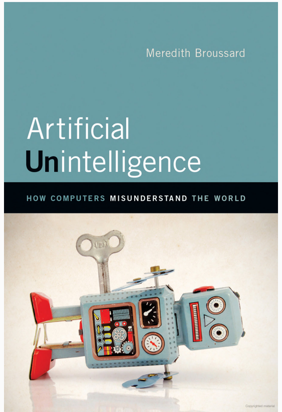
"Computational Social Science"

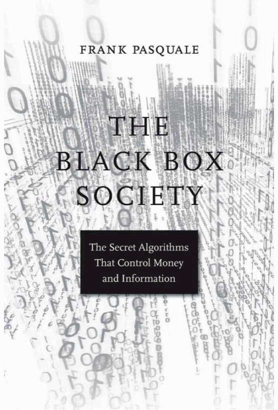
Eric Fisher. European detail map of Flickr and Twitter locations, 2011. https://flic.kr/p/a1vp4W

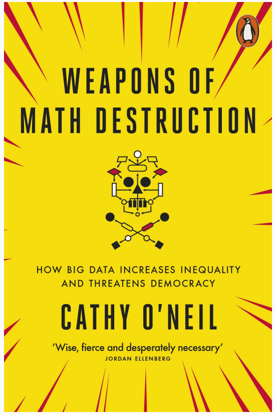
- Much computational social science looks for stories in the data
- But data are never neutral. They are produced in specific contexts, for specific purposes.
- The data contain a story, but their genesis is also a story

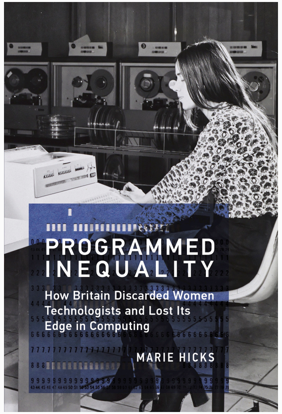
Great work telling the story of data

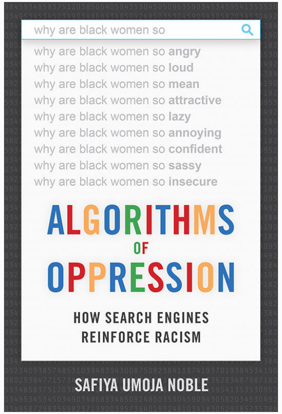












- But none of this work explores opportunities for modeling

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- **Modeling algorithmic user manipulation**
- Priorities, current work, and goals
- References

Studying algorithmic manipulation

- *The platform* (that produces the data) is the story
- Using data and modeling "reflexively": apply to themselves

Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)

Identifying Platform Effects in Social Media Data

Momin M. Malik[1] and **Jürgen Pfeffer**[1,2]
[1]Institute for Software Research
School of Computer Science
Carnegie Mellon University
[2]Bavarian School of Public Policy
Technical University of Munich

Abstract

Even when external researchers have access to social media data, they are not privy to decisions that went into platform design—including the measurement and testing that goes into deploying new platform features, such as recommender systems, seeking to shape user behavior towards desirable ends. Finding ways to identify platform effects is thus important both for generalizing findings, as well as understanding the nature of platform usage. One approach is to find temporal data covering the introduction of a new feature; observing differences in behavior before and after allow us to estimate the effect of the change. We investigate platform effects using two such datasets, the Netflix Prize dataset and the Facebook New Orleans data, in which we observe seeming discontinuities in user behavior but that we know or suspect are the result of a change in platform design. For the Netflix Prize, we estimate user ratings changing by an average of about 3% after the change, and in Facebook New Orleans, we find that the introduction of the 'People You May Know' feature locally nearly doubled the average number of edges added daily, and increased by 63% the average proportion of triangles created by each new edge. Our work empirically verifies several previously expressed theoretical concerns, and gives insight into the magnitude and variety of platform effects.

Introduction

In social media data, the design and technical features of a given platform constrain, distort, and shape user behavior on that platform, which we call the *platform effects*. For those inside companies, knowing the effect a particular feature has on user behavior is as simple as conducting an A/B test (i.e., a randomized experiment), and indeed such testing is central to creating platforms that shape user behavior in desirable ways. But external researchers have no access to the propriety knowledge of these tests non-embedded researchers having access to the data (Savage and Burrows 2007; Lazer et al. 2009; Huberman 2012; boyd and Crawford 2012), but also that even when researchers have access, without full knowledge of the platform engineering and the decisions and internal research that went into design decisions, the data can be systematically misleading.

One way to study and quantify platform effects as an external researcher is to look for available data that include a significant platform change. Making the assumption that, in absence of the exogenous shock (the change) the previous 'trend' would have remained the same, we can apply the observational inference method of *regression discontinuity design* (Imbens and Lemieux 2008; Lee and Lemieux 2010; Li 2013). While not as certain as experimental design, observational inference methods are the best available way for outside researchers to understand the effects of platform design.

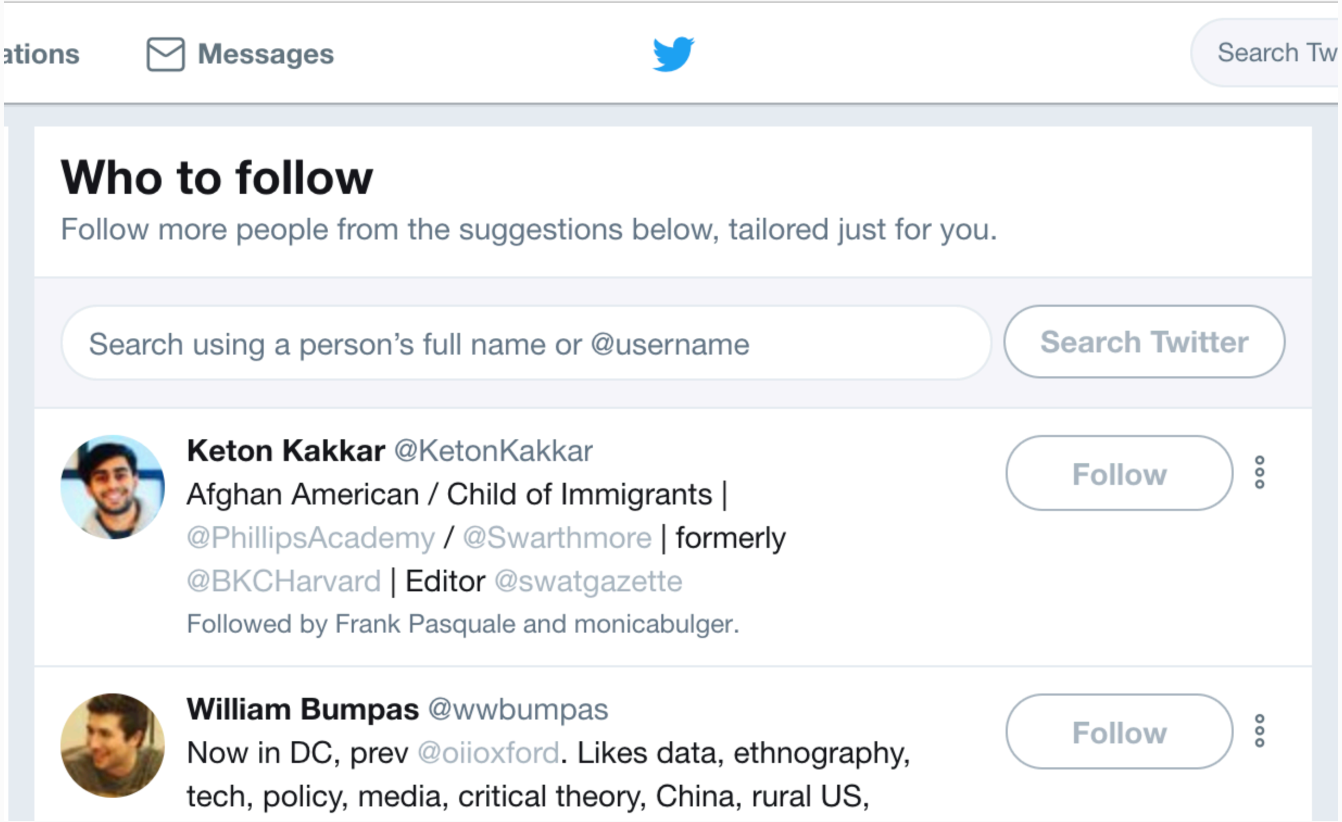
We select two data sets: the Facebook New Orleans data collected by Viswanath et al. (2009), and the Netflix Prize data, described by Koren (2009b). This is no longer publicly available since the close of the Netflix prize, although the terms of use do not mention any expiration on use for those who have already downloaded it.

In the Netflix Prize data set, Koren (2009b), a member of the team that ultimately won the prize (Koren 2009a), points out a curious spike in the average ratings in early 2004. As such a change has modeling implications (previous data should be comparable in order to properly use for training purposes), he explores the possible reasons for this, ultimately identifying an undocumented platform effect as the most likely driver. Then, the Facebook New Orleans data contains an identified, and ideal, example of a platform effect: a clear exogenous shock and a dramatic difference af-

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

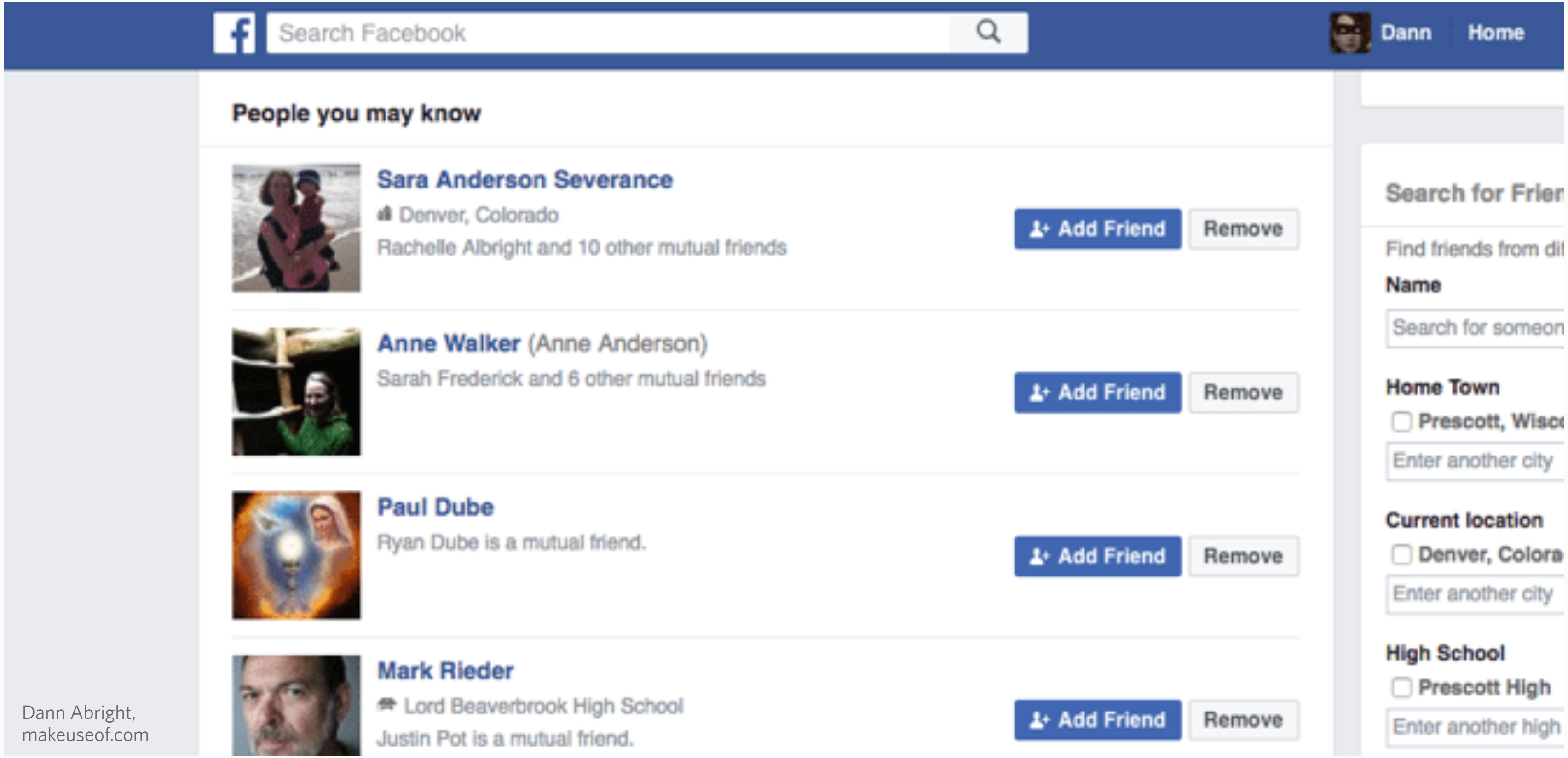
- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Have you used 'Who to follow'?



What about 'People you may know'?

My story
Reporting on "prediction" in machine learning/AI
At-scale study of news media on Twitter
Modeling algorithmic user manipulation
Priorities, current work, and goals
References

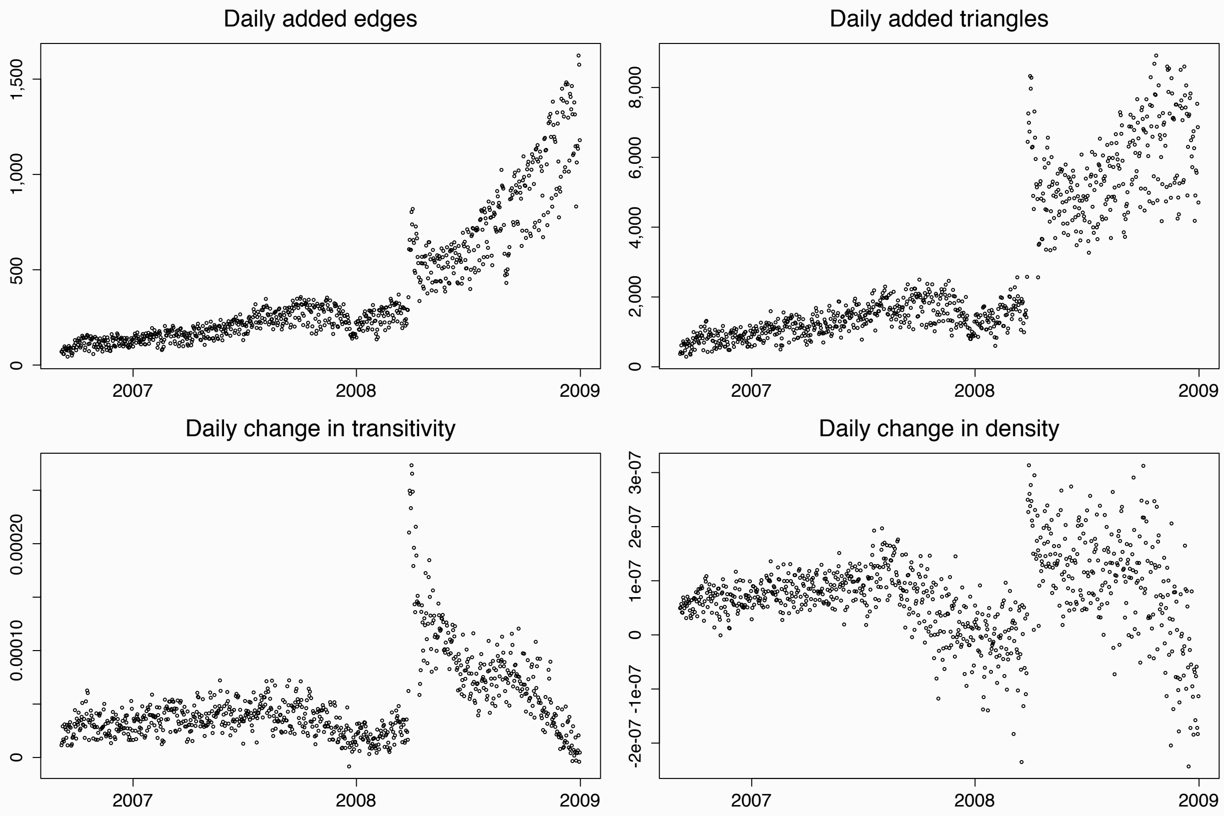


Dann Abright, makeuseof.com

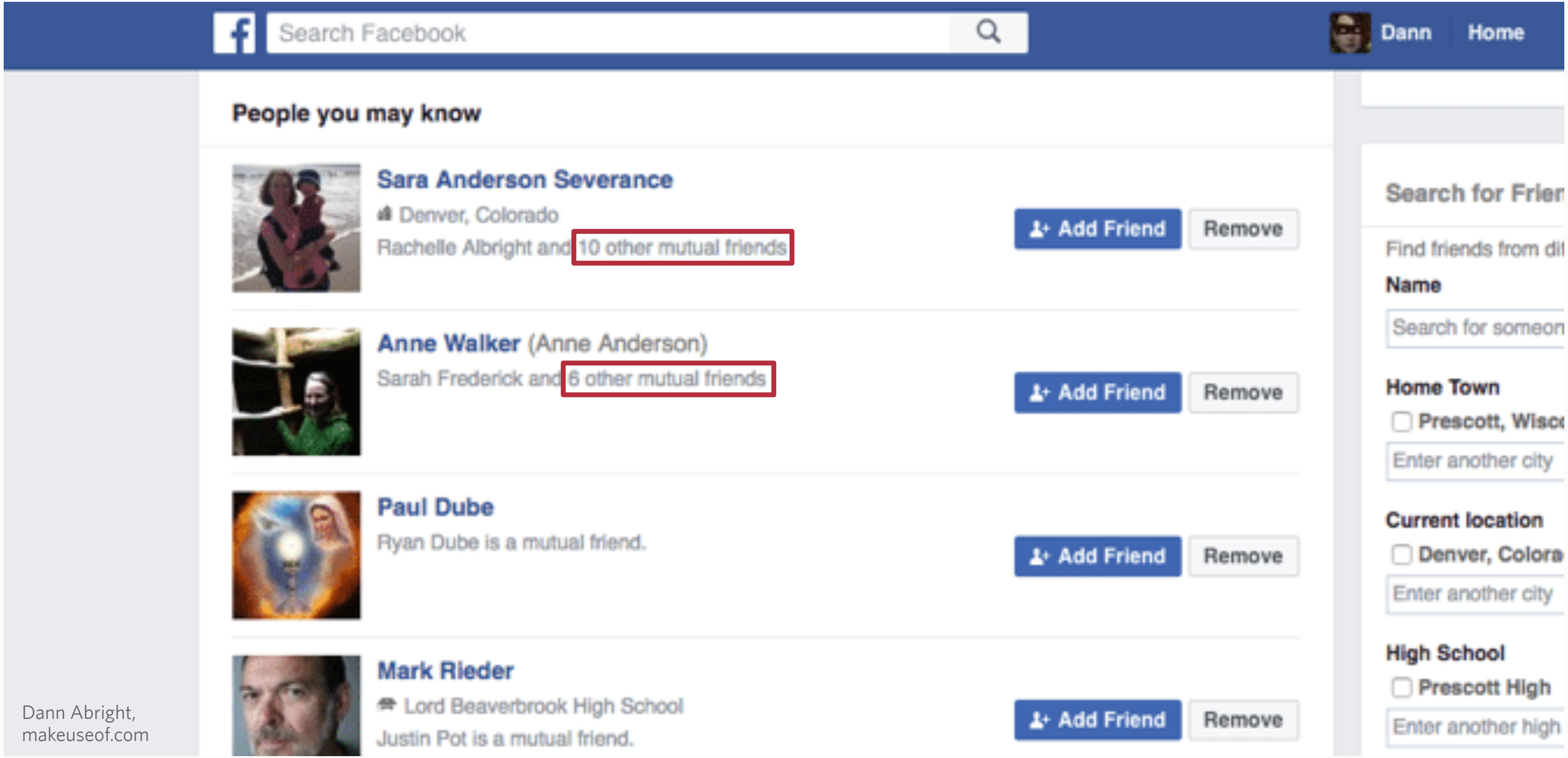
- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Data artifacts can reveal inner workings

Facebook's 'People you may know'



Based on *number of mutual friends*

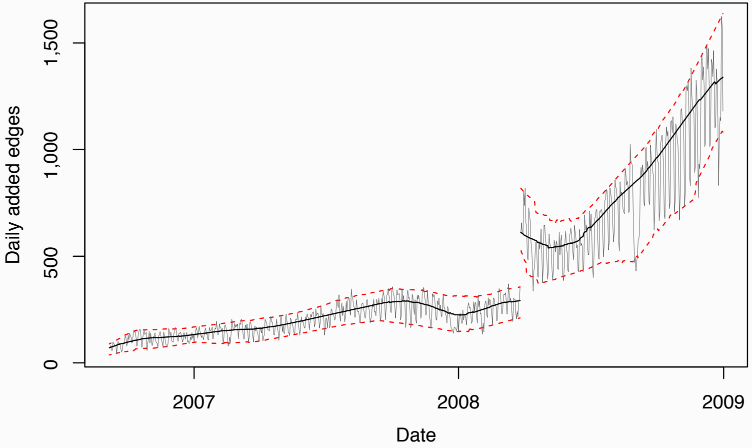


Dann Abright, makeuseof.com

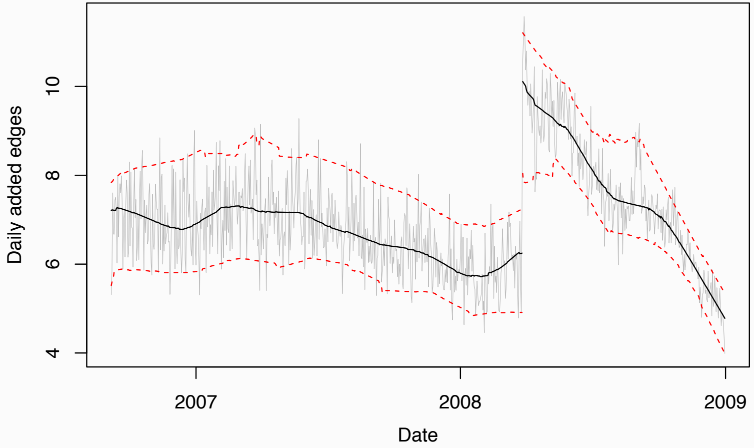
- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Recommendations change the network!

- Facebook links: +300 new edges per day (~200%)



- Triangles: +3.8 triangles per edge (~64%)



BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Manipulations have consequences

Identifying Networks of Criminals

"Facebook has helped me by identifying suspects that were friends or associates of other suspects in a crime and all brought in and interviewed and later convicted of theft and drug offenses."

"My biggest use for social media has been to locate and identify criminals. I have started to utilize it to piece together local drug networks."

LexisNexis® Risk Solutions (2014). Survey of law enforcement personnel and their use of social media.

Goal:	Predict, Monitor, and Prevent Risk In/Around Protests
Anticipated Activity:	Protests, Riots, Looting
Overt Threats:	Unions, Activist Groups, Etc.
Locations:	Schools, Public Spaces, Malls, High-Rent Districts
Actions Taken:	During Event(s), Post-Event

Nicole Ozer. (2016, September 23). Police use of social media surveillance software is escalating, and activists are in the digital crosshairs. ACLU of Northern California.

Authorization and authentication based on an individual's social network

US 9432351 B2

ABSTRACT

In particular embodiments, a method includes receiving a request for a first user to access a loan from a lender, the request identifying a user identifier (ID) of the first user; determining whether the first user is authorized to access the loan based at least in part on a gray list comprising user IDs of the users who are not authorized to access loans, wherein the gray list is based on a black list; and permitting the loan to be accessed by the first user if the first user is authorized to access the loan based on the gray list.

Publication number	US9432351 B2
Publication type	Grant
Application number	US 14/299,391
Publication date	Aug 30, 2016
Filing date	Jun 9, 2014
Priority date	Jul 22, 2004
Also published as	CN101036366A, 26 More »
Inventors	Christopher Lunt
Original Assignee	Facebook, Inc.
Export Citation	BiBTeX, EndNote, RefMan

Patent Citations (81), Non-Patent Citations (13), Classifications (23)

External Links: USPTO, USPTO Assignment, Espacenet

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

My priorities, current work, and goals

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

My priorities

- The best way to overcome anxiety is to gain mastery over the threat.
- Individual tools are fleeting. I'm interested in principles and training in foundations.
- Apply a critical lens to computational; conversely, apply computation to a critical lens

My current work

- BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY
- Serving social scientists, librarians, lawyers, journalists, etc
- "Ethics and Governance of AI" project: governance people don't know AI and vice versa
- Also, materials aimed at computer scientists to be more faithful, reflective communicators
- MEDIA CLOUD Self-collected database of online news to study disinformation and misinformation
- Work on sensors/IoT and social networks

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- **Priorities, current work, and goals**
- References

My goals

Marsh Arab village, 1974. Nik Wheeler / Corbis. Via *Al Jazeera.America.*

Carolyn Drake / Panos. Via *National Geographic.*

- Save the swamp, but prepare for the desert
- Empower "swamp-dwellers," convert "desert-dwellers"
- Dream: teach a hybrid course
- Produce *critical* training material on data, modeling
- Further empirical study about the limits of data and modeling
- *Produce transformative understandings*


#NOTHELPING
WHAT WILL BECOME OF US


#MUCHBETTER


#THANK YOU

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

Additional credits/citations

Credits:

Robot holding skull: Cover image of "What Will Become of Us?", *New York Times Magazine* (The Tech & Design issue), 14 November 2018. Concept by delcan & company. Photo illustration by Jamie Chung. Prop styling by Pink Sparrow. C.G. work by Justin Metz. https://www.nytimes.com/2018/11/14/magazine/behind-the-cover-what-will-become-of-us.html.

Newspaper graphic: "Newspaper" by art shop from the Noun Project. https://thenounproject.com/term/newspaper/847413.

Computer graphic: "Computer" by Komkrit Noenpoempisut from the Noun Project. https://thenounproject.com/term/computer/2056682.

Arthur Leff on the desert and the swamp: quoted in Robert C. Ellickson, "A Critique of Economic and Sociological Theories of Social Control." *Journal of Legal Studies* 16 (January 1987): 67–99.

Mechanical crystal ball: by Zhivko Terziivanov, cover image of Eric Siegel, *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die.* John Wiley & Sons, 2013.

Chess Turk: from Tom Standage, *The Turk: The Life and Times of the Famous Eighteenth Century Chess Playing Machine*. Walker & Company, 2002.

Microphone: "reporter" by Aldric Rodríguez from the Noun Project. https://thenounproject.com/term/reporter/813747

Hat: "reporter" by Bonegolem from the Noun Project. https://thenounproject.com/term/reporter/373481.

Black cat/déjà vu: The Matrix/Warner Bros, 1999.

Terminator skull: Nemesis Now Ltd, Terminator Skull Box T-800 (18CM). https://www.menkind.co.uk/terminator-t800-skull-box.

Hand: Hayati Kayhan, Holding human skull in hand, Conceptual image (Shakespeare's Hamlet scene concept). 19 October 2014.

Citations:

Armstrong, Corey L., and Fangfang Gao, "Now Tweet This: How News Organizations Use Twitter." *Electronic News* 4, no. 4 (2010): 218–235. https://dx.doi.org/10.1177/1931243110389457.

Holcomb, Jesse, Kim Gross, and Amy Mitchell. "How Mainstream Media Outlets Use Twitter: Content Analysis Shows an Evolving Relationship." The Project for Excellence in Journalism, Pew Research Center, and the George Washington University School of Media and Public Affairs, 14 November 2011. http://www.journalism.org/2011/11/14/how-mainstream-media-outlets-use-twitter/.

Vis, Farida. "Twitter as a Reporting Tool for Breaking News: Journalists Tweeting the 2011 UK Riots." *Digital Journalism* 1, no. 1 (2013): 27–47. https://dx.doi.org/10.1080/21670811.2012.741316.

Bob Franklin, "The Future of Journalism." *Digital Journalism* 4, no. 7 (2016): 806–808, https://dx.doi.org/10.1080/21670811.2016.1197638; *Journalism Studies* 17, no. 7 (2016): 798–800, https://dx.doi.org/10.1080/1461670X.2012.712301; *Journalism Practice* 10, no. 7 (2016): 805-807, https://dx.doi.org/10.1080/17512786.2016.1197640.

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- My story
- Reporting on "prediction" in machine learning/AI
- At-scale study of news media on Twitter
- Modeling algorithmic user manipulation
- Priorities, current work, and goals
- References

References for pictured books/papers

Lawrence Lessig, *Code and Other Laws of Cyberspace*. Basic Books, 1999.

Urs Gasser, Sandra Cortesi, Momin Malik, and Ashley Lee, "Youth and Digital Media: From Credibility to Information Quality," Berkman Center for Internet & Society Research Publication No. 2012-1 (2012). http://ssrn.com/abstract=2005272.

Momin Malik, Sandra Cortesi, and Urs Gasser, "The Challenges of Defining 'News Literacy'." Berkman Center for Internet & Society Research Publication No. 2013-20 (October 2013). http://ssrn.com/abstract=2342313.

Roger Burrows and Mike Savage, "The Coming Crisis of Empirical Sociology." *Sociology* 41, no. 5 (2007): 885–899. https://dx.doi.org/10.1177/0038038507080443.

Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution that will Transform How We Live, Work, and Think*. Eamon Dolan/Houghton Mifflin Harcourt, 2013.

Martin Hirst, *News 2.0: Can Journalism Survive the Internet?* Allen & Unwin, 2011.

Jackie Snow, "The Download: Facebook is using AI to predict users' future behavior and selling that data to advertisers." *MIT Technology Review*, 13 April 2018. https://www.technologyreview.com/the-download/610888/facebook-is-using-ai-to-predict-users-future-behavior-and-selling-that-data-to/.

Antonio Regalado, "Rewriting Life: Scientists Can Read a Bird's Brain and Predict Its Next Song." *MIT Technology Review*, 11 October 2017. https://www.technologyreview.com/s/609032/scientists-can-read-a-birds-brain-and-predict-its-next-song/.

"A View from Emerging Technology from the arXiv: Machine Learning Algorithm Predicts Which New Faces Will Make It as Fashion Models." *MIT Technology Review*, 1 September 2015. https://www.technologyreview.com/s/541036/machine-learning-algorithm-predicts-which-new-faces-will-make-it-as-fashion-models/.

Tom Simonite, "Intelligent Machines: Software Predicts Tomorrow's News by Analyzing Today's and Yesterday's." *MIT Technology Review*, 1 February 2013. https://www.technologyreview.com/s/510191/software-predicts-tomorrows-news-by-analyzing-todays-and-yesterdays/.

Sitaram Asur and Bernando A. Huberman, "Predicting the Future With Social Media." In *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (WI-IAT '10), 492–499. 2010. https://dx.doi.org10.1109/WI-IAT.2010.63.

Ziad Obermeyer and Ezekiel J. Emanuel, "Predicting the Future — Big Data, Machine Learning, and Clinical Medicine." *New England Journal of Medicine* 375, no. 13 (2016): 1216–1219. https://dx.doi.org/10.1056/NEJMp1606181.

Merriam-Webster. "predict". Merriam-Webster.com, 2018. https://www.merriam-webster.com/dictionary/predict.

Daniel Gayo-Avello, "'I Wanted to Predict Elections with Twitter and all I got was this Lousy Paper': A Balanced Survey on Election Prediction using Twitter Data." arXiv, 28 April 2012. https://arxiv.org/abs/1204.6441

Daniel Gayo-Avello, "No, You Cannot Predict Elections with Twitter." *IEEE Internet Computing* 16 (2012): 91–94. https://dx.doi.org/10.1109/MIC.2012.137.

David Lazer, Ryan Kennedy, Gary King, and Alessandro Vespignani, "The Parable of Google Flu Trends: Traps in Big Data Analysis." Science 343 (14 March 2014): 1203–1205. https://dx.doi.org/10.1126/science.1248506.

Drew McDermott, "Artificial Intelligence meets Natural Stupidity." *SIGART Bulletin* 57 (April 1976): 4–9.

Momin M. Malik and Jürgen Pfeffer, "A Macroscopic Analysis of News Content on Twitter." *Digital Journalism* 4, no. 8 (2016): 955–979. http://dx.doi.org/10.1080/21670811.2015.1133249.

Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press, 2018.

Meredith Broussard, *Artificial Unintelligence: How Computers Misunderstand the World*. MIT Press, 2018.

Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, 2015.

Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown, 2016.

Marie Hicks, *Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing*. The MIT Press, 2017.

Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press, 2018.

Momin M. Malik and Jürgen Pfeffer, "Identifying Platform Effects in Social Media Data." In *Proceedings of the Tenth International AAAI Conference on Web and Social Media* (ICWSM 2016), 241–249. 2016.

Christopher Lunt. "Authorization and authentication based on an individual's social network." US Grant US9100400B2, 9 June 2014. https://patents.google.com/patent/US9432351B2/en.