CS 6474/CS 4803 Social Computing: Privacy

Munmun De Choudhury

munmund@gatech.edu

Week 15 | November 26, 2018

Please take the Course Instructor Opinion (CIOS) survey!!!

http://b.gatech.edu/cios

Final presentation specs

- Next two classes
- Each team gets 15 minutes in all
  - 10-12 minutes of presentation
  - 3-5 minutes of Q&A
- Each team member needs to be present
- Structure:
  - Main idea
  - Background/Motivation
  - Research questions/Goals
  - Data/Social media platform
  - Method
  - Results
  - What you have learned

1. What can we do with data generated from social computing systems?

2. What should we **not** do with these data.

Data ex Machina:
Introduction to Big Data

Data, Privacy, and the Greater Good

Defining Privacy

- Privacy related to notion of access
- Privacy is not "being alone", but defining who has access to what
- Access
  - Physical proximity to a person
  - Knowledge about a person
- Privacy is a "zone of inaccessibility"
- Privacy violations are an affront to human dignity
  - You violate privacy when you treat a person as a means to an end.
  - Some things ought not be known – you look away when your friend is typing their password
- Too much individual privacy can harm society
- Where to draw the line?

Benefits of Privacy

- Individual growth
  - Necessary to blossom into a unique individual
- Individual responsibility
- Freedom to be yourself
  - Nobody likes to be videotaped or "watched" all the time
- Intellectual and spiritual growth
- Development of loving, trusting, caring, intimate relationships

Harms of Privacy

- Cover for illegal or immoral activities

- Hidden dysfunctional families
  - Incidents of domestic violence

- Ignored people on society's fringes
  - People with disability e.g., with mental illness

Discussion Point

Contrast with privacy on social computing systems

Rule #1

❖ It is safe to assume if you put information online it isn't 100% private.

❖ A video to get us started: http://www.youtube.com/watch?v=5P_0s1TYpJU

It's a pleasure to meet you sir. Facebook has told me so much about you.

search ID: bchn69

© Original Artist
Reproduction rights obtainable from
www.CartoonStock.com

https://www.youtube.com/watch?v=5P_0s1TYpJU

"Participant" Perceptions of Twitter Research Ethics

Casey Fiesler[1] and Nicholas Proferes[2] (iD)

Abstract
Social computing systems such as Twitter present new research sites that have provided billions of data points to researchers. However, the availability of public social media data has also presented ethical challenges. As the research community works to create ethical norms, we should be considering users' concerns as well. With this in mind, we report on an exploratory survey of Twitter users' perceptions of the use of tweets in research. Within our survey sample, few users were previously aware that their public tweets could be used by researchers, and the majority felt that researchers should not be able to use tweets without consent. However, we find that these attitudes are highly contextual, depending on factors such as how the research is conducted or disseminated, who is conducting it, and what the study is about. The findings of this study point to potential best practices for researchers conducting observation and analysis of public data.

Keywords
Twitter, Internet research ethics, social media, user studies

With your permission, you give us more permission. If you give us information about who some of your friends are, we can probably use some of that information, again, with your permission, or improve the quality of our searches. We don't need you to type at all, because we know where you are, with your permission. We know where you have been, with your permission. We an more or less guess what you are thinking about. – *Eric Schmidt, Google CEO (The Atlantic)*

Legal-ease

❖ Legally, read all platforms terms of service (TOS) for the nitty gritty, social media platforms can share some of your basic information.

❖ But why?

    ❖ Social networks that provide their services without user fees make a profit by selling advertising. This is often done through behavioral advertising, also known as targeting. Facebook Pages who boost posts and promote their brands through ads use the same targeting methods when pushing their content.

Geo-Locate Privacy?

❖ If you use Fourquare or Instagram or even have the location settings turned on for Facebook and Twitter than you are sharing your location. On Twitter you are sharing it with everyone and since it is a live update tool then you are letting everyone know exactly where you are and when and with who if you have tagged or taken a photo.

Settings vary across platforms

❖ Each social media platform has different privacy settings and they change their rules frequently. Facebook just updated their privacy settings in May of 2014, did you know? Did you just click the "Yes, I Agree" without reading?

What could Facebook possibly know?

❖ Anything you provide them. Think about it..

  ❖ Name, City of birth, City of residence, Phone, Email, Current employment, Previous employment, Relationship, Anniversary, Previous relationships, Previous names (aliases), Screen names, Address book, Family members, Birthday, Religious views, Friends, Books you've read, Movies you like, etc....

❖ Oh you thought that was all…

  ❖ What about the videos you have watched, the links you have clicked on, the comments you have left with companies, advertising you connected with and advertising that didn't intrigue you, etc.

Facebook Beacon

- Fandango, eBay, and 42 other online businesses paid Facebook to do "word of mouth" advertising

- Facebook users surprised to learn information about their purchases was shared with friends

- Beacon was based on an opt-out policy

- Beacon strongly criticized by various groups

- Facebook switched to an opt-in policy regarding Beacon

Instagram's Proposed Change to Terms of Service

- Late 2012: Instagram announced changes
  - Privacy policy
  - Terms of service
- Legal experts: Instagram and Facebook would have right to use photos in ads without permission
- Instagram CEO: New policy misunderstood
- Changed advertising section of terms of service agreement back to original version

What's at Stake: Characterizing Risk Perceptions of Emerging Technologies

Michael Skirpan
University of Colorado
Boulder, CO
michael.skirpan@colorado.edu

Tom Yeh
University of Colorado
Boulder, CO
tom.yeh@colorado.edu

Casey Fiesler
University of Colorado
Boulder, CO
casey.fiesler@colorado.edu

ABSTRACT

One contributing factor to how people choose to use technology is their perceptions of associated risk. In order to explore this influence, we adapted a survey instrument from risk perception literature to assess mental models of users and technologists around risks of emerging, data-driven technologies (e.g., identity theft, personalized filter bubbles). We surveyed 175 individuals for comparative and individual assessments of risk, including characterizations using psychological factors. We report our observations around group differences (e.g., expert versus non-expert) in how people assess risk, and what factors may structure their conceptions of technological harm. Our findings suggest that technologists see these risks as posing a bigger threat to society than do non-experts. Moreover, across groups, participants did not see technological risks as voluntarily assumed. Differences in how people characterize risk have implications for the future of design, decision-making, and public communications, which we discuss through a lens we call risk-sensitive design.

ACM Classification Keywords

H.1.2 User/Machine Systems: Human Factors; H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

and behavior-driven design. These users must rely on the companies and parties to whom they have given their data (knowingly or not) to be ethical.

Yet, we already know that many impacts (e.g., privacy, ethical, legal) and constraints (e.g., protocols, technological capabilities) of online technologies are poorly understood by users [24, 8, 36, 15]. We also know that, when asked, users are often uncomfortable or find undesirable the practices of online behavioral advertising (OBA) and personalization [37, 34]. This misalignment is often framed as a consumer trade-off between privacy and personal benefit [13, 40]. Framing it this way leads to an assumption that the benefit of web services must outweigh consumer's privacy concerns since users are not opting out of services.

However, if consumers really are performing this cost-benefit analysis and making a conscious decision, then why we do we see such hype and panic around risks and harms caused by technology in the media? Daily news headlines relay injustice [19, 1, 4, 33], personal boundary violations [32], and gloom [26, 18, 14] over the impacts of technology on society. Some of these problems may indeed warrant concern from the public and social advocates; others might be overblown

What's at Stake: Characterizing Risk Perceptions of Emerging Technologies

Michael Skirpan
University of Colorado
Boulder, CO
michael.skirpan@colorado.edu

Tom Yeh
University of Colorado
Boulder, CO
tom.yeh@colorado.edu

Casey Fielser
University of Colorado
Boulder, CO
casey.fiesler@colorado.edu

	Non-Expert			Expert	
Rank	Risk	Mean Rank		Risk	Mean Rank
1	Identity Theft	5.000		Job Loss	5.769
2	Account Breach	6.101		Account Breach	6.385
3	Job Loss	7.678		Identity Theft	6.577
4	Hacktivist Leak	7.980		Technology Divide	6.923
5	Auto-Drones	8.523		Bias Job Alg	7.192
6	Harassment	9.074		Discriminatory Crime Alg	7.231
7	Undisclosed third party	9.349		Hacktivist Leak	7.231
8	DDoS	9.403		Filter Bubble	7.654
9	Nuclear Reactor Meltdown	9.644		DDoS	8.269
10	Discriminatory Crime Alg	9.758		Undisclosed third party	8.462
11	Research w/o Consent	10.141		Harassment	9.346
12	Bias Job Alg	10.154		Auto-Drones	9.808
13	Driverless Car Malfunction	10.315		Research w/o Consent	11.154
14	Technology Divide	10.765		Nude Photos	12.038
15	Plane Crash	11.060		Driverless Car Malfunction	12.269
16	Filter Bubble	11.362		Nuclear Reactor Meltdown	14.308
17	Nude Photos	11.846		Plane Crash	14.654
18	Vaccine	12.846		Vaccine	15.731

Figure 1. Average comparative risk ranking by non-experts vs experts where items with significant differences (p<.05 for two-tailed t-test) are highlighted.

Class Exercise I

As a social media designer, what additional elements would you incorporate on Facebook so that people are more aware of their privacy settings? (People often complain about Facebook changing privacy related setting too often)

But it is not just the third party "bad actors"; what happens when the risk of privacy lies in the hands of the service provider themselves?

Support The Guardian

Contribute → Subscribe →

Search jobs | Sign in | Search | US edition

The Guardian

News | Opinion | Sport | Culture | Lifestyle | More

US World Environment Soccer US midterms 2018 Business Tech Science

Head quarters
Science

Facebook fiasco: was Cornell's study of 'emotional contagion' an ethics breach?

A covert experiment to influence the emotions of more than 600,000 people. A major scientific journal behaving like a rabbit in the headlights. A university in a PR tailspin
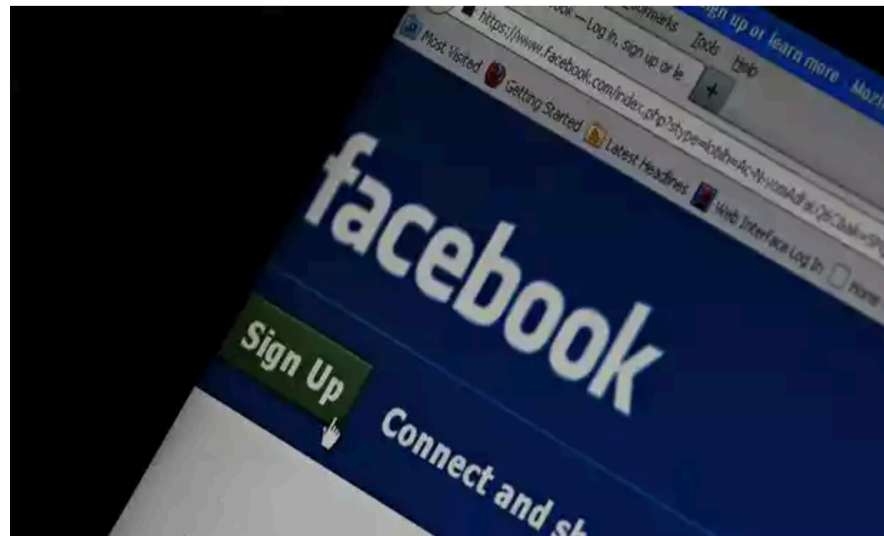
Chris Chambers

🐦 **@chrisdc77**

Tue 1 Jul 2014 02.00 EDT

1,012   17

Class Exercise II

In the aftermath of the controversial Facebook contagion study, how do you think people's privacy perceptions may have changed? Or did they at all?

Secondary information use of social media data and privacy

Google's Personalized Search

- Secondary use: Information collected for one purpose use for another purpose

- Google keeps track of your search queries and Web pages you have visited

  - It uses this information to infer your interests and determine which pages to return

  - Example: "bass" could refer to fishing or music

- Also used by retailers for direct marketing

Collaborative Filtering

- Form of data mining

- Analyze information about preferences of large number of people to predict what one person may prefer

  - Explicit method: people rank preferences

  - Implicit method: keep track of purchases

- Used by online retailers and movie sites

Social Network Analysis

- Data mining now incorporating information collected from social networks

- Examples
  - Cell phone companies in India identify "influencers" – provide discounts
  - Police predict locations of big parties
  - Banks evaluate the riskiness of loans

Facebook created an AI tool that can prevent suicide, but won't talk about how it works

f **Share on Facebook**   🐦 **Share on Twitter**   +

IMAGE: VICKY LETA / MASHABLE

Class Debate

Social media monitoring and health insurance

But it is not just the third party "bad actors"; what happens when the government or other similar authorities start to make use of people's online data?

China's Social Ranking System Is Getting Closer to Becoming a Terrifying Reality

Catie Keck
Thursday 5:00pm • Filed to: SOCIAL CREDIT ⌄

The lifelong social ranking system is set to be adopted in Beijing in 2021, Bloomberg reported Tuesday, with residents to be judged on data based on their social standing by the end of 2020. The program would essentially mark any individuals found to have violated laws or social codes and restrict their access to services like travel or certain programs.

China's Social Ranking System Is Getting Closer to Becoming a Terrifying Reality

Catie Keck
Thursday 5:00pm • Filed to: SOCIAL CREDIT ⌄

The capital city will pool data from several departments to reward and punish some 22 million citizens based on their actions and reputations by the end of 2020, according to a plan posted on the Beijing municipal government's website on Monday. Those with better so-called social credit will get "green channel" benefits while those who violate laws will find life more difficult.

The Beijing project will improve blacklist systems so that those deemed untrustworthy will be "unable to move even a single step," according to the government's plan.

Class Exercise III

Government use of social media data for social surveillance.