Lecture 2: Ethics/Methodology
June 27, 2019
ethics and CMC

Thought experiment (inspired by Ayers et al. 2018).

A little group of sociolinguists is thinking about doing a study of the discourse-pragmatic functions of the word ‘awesomesauce’ as used on Twitter, so they spend a couple of days playing with the Twitter API and some Python code.

It works! They scrape all of Twitter and have 12,945 tokens of ‘awesomesauce’ to play with.
@whitneyldelaney
omg awesomesauce!!!

@1990kalo1990
Terence Tao is teh awesomesauce

@inter_stitial
i told her it was awesomesauce but forgot i was talking about ketchup lol

@bowlagranola
Nah, she came up with this **AWESOMESAUCE** alternative at like the last minute and saved our butt

@00ccffblue
How do you turn weaksauce into awesomesauce?
Asking for a friend

(Note: fabricated examples with pretend usernames.)
All of these tweets have been posted publicly, and the sociolinguists’ respective IRBs (internal review boards) have no problem with this.

Just to be on the safe side, the group decides to obscure the usernames.
omg awesomesauce!!

Terence Tao is teh awesomesauce

i told her it was awesomesauce but forgot i was talking about ketchup lol

Nah, she came up with this **AWESOMESAUCE** alternative at like the last minute and saved our butt

How do you turn weaksauce into awesomesauce? Asking for a friend
The group analyses the data, does a conference presentation, writes up a manuscript, and sends it to a linguistics journal.

Neither of the reviewers is familiar with the word ‘awesomesauce.’

One of them finds it on dictionary.com and figures out what’s going on. The other fails to find it in the Oxford English Dictionary online and is confused for the entire duration of the review about whether ‘awesomesauce’ is an adjective or a noun.

Both request more actual examples in the paper. The group reviews the data and puts a few in.
The paper is accepted and published, and the group’s local media writes a little story about it.

...usually assumed to be an adjective, perhaps on the model of ‘weaksauce’. It does not have to have anything to do with the noun ‘sauce,’ as this Twitter example shows: “i told her it was awesomesauce but forgot i was talking about ketchup lol.”

A local TV show thinks is this neat, puts the quotation into Google, and sends a private message to @interstitial to comment for a story. @interstitial has never heard of the study and is alarmed.
The problem that this illustrates is that **individual accounts are identifiable** if you quote material that is indexed by a search engine.

While @inter_stitial did post their thoughts publicly, they did not consent to being **studied** or being talked about in a newspaper or on a news show.

The group of researchers **did not take adequate precautions to safeguard the privacy** of the people who used the linguistic feature they wanted to study.
ethics and CMC

This leads us to Ayers et al. (2018):

Don’t quote me: reverse identification of research participants in social media studies

John W. Ayers¹, Theodore L. Caputi², Camille Nebeker³ and Mark Dredze⁴

We investigated if participants in social media surveillance studies could be reverse identified by reviewing all articles published on PubMed in 2015 or 2016 with the words “Twitter” and either “read,” “coded,” or “content” in the title or abstract. Seventy-two percent (95% CI: 63–80) of articles quoted at least one participant’s tweet and searching for the quoted content led to the participant 84% (95% CI: 74–91) of the time. Twenty-one percent (95% CI: 13–29) of articles disclosed a participant’s Twitter username thereby making the participant immediately identifiable. Only one article reported obtaining consent to disclose identifying information and institutional review board (IRB) involvement was mentioned in only 40% (95% CI: 31–50) of articles, of which 17% (95% CI: 10–25) received IRB-approval and 23% (95% CI:16–32) were deemed exempt. Biomedical publications are routinely including identifiable information by quoting tweets or revealing usernames which, in turn, violates ICMJE ethical standards governing scientific ethics, even though said content is scientifically unnecessary. We propose that authors convey aggregate findings without revealing participants’ identities, editors refuse to publish reports that reveal a participant’s identity, and IRBs attend to these privacy issues when reviewing studies involving social media data. These strategies together will ensure participants are protected going forward.

npj Digital Medicine (2018)1:30; doi:10.1038/s41746-018-0036-2
This is a 2-page demonstration of identifying Twitter accounts from quotations in journal articles.

“One overlooked issue is the inclusion of direct quotes or usernames of social media users in academic publications. When preserved this way, the quoted material can potentially be linked back to the originating account and inferentially the account owner” (Ayers et al. 2018:1).

They searched for all studies new to PubMed in 2015 or 2016 that mentioned relying on Twitter.
Results: 112 papers from 2015-2016.

21% of these included usernames, making the participants immediately findable.

72% included at least one tweet.

60% included at least one tweet which, when plugged into a search engine, led directly to the user who was quoted.

= 81% of these studies made someone who’d been unwittingly quoted in a medical paper very easy to find online.
“Publication in the biomedical literature is permanent and removes control from the poster” (Ayers et al. 2018:1).
Only **40%** of the studies referenced an IRB application.

**One** of the 112 papers had IRB **approval** to gather and to disclose this much information.

In other words: 4/5 of the papers pointed to identifiable users on Twitter, “nearly all without consent and most occurring outside IRB review” (Ayers et al. 2018:1).
ethics and CMC

What to do instead?

Researchers, editors, and IRBs can all independently mitigate the risk of harm.

“We propose that authors convey aggregate findings without revealing participants’ identities, editors refuse to publish reports that reveal a participant’s identity, and IRBs attend to these privacy issues when reviewing studies involving social media data. These strategies together will ensure participants are protected going forward” (Ayers et al. 2018:1).
One of the big ideas we’re getting at is that someone consenting to posting something publicly on the Internet is not necessarily consenting to being in an academic study.

Individuals, groups, or broad subsets of humans could all be harmed through careless handling of identifying information (Hoffman and Jones 2016).
ethics and CMC

For a sense of the disparity here in users’ perceptions, let’s turn to Fiesler and Proferes (2018):

“Participant” Perceptions of Twitter Research Ethics

Casey Fiesler¹ and Nicholas Proferes²

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.
“As the research community works to create ethical norms, we should be considering users’ concerns as well” (Fiesler and Proferes 2018:1).

Most Twitter users are not aware of APIs, web-scraping, and automatic archiving of tweets.

A few don’t realize in the first place that their tweets are public (Proferes 2017).

Some are careful, but then other people are careless with their information – what if that is captured? (Mao et al. 2011).
Fiesler and Proferes (2018) surveyed Twitter users (via Amazon Mechanical Turk), presenting them with a range of hypothetical situations.

More than 60% of the respondents stated that they were unaware that tweets get used in research.

About 16% of respondents realized that this was possible, but thought it was not permitted.
ethics and CMC

There is abundant unease.

Table 2. Comfort Around Tweets Being Used in Research.

<table>
<thead>
<tr>
<th>Question</th>
<th>Very uncomfortable</th>
<th>Somewhat uncomfortable</th>
<th>Neither uncomfortable nor comfortable</th>
<th>Somewhat comfortable</th>
<th>Very comfortable</th>
</tr>
</thead>
<tbody>
<tr>
<td>How do you feel about the idea of tweets being used in research? (n = 268)</td>
<td>3.0%</td>
<td>17.5%</td>
<td>29.1%</td>
<td>35.1%</td>
<td>15.3%</td>
</tr>
<tr>
<td>How would you feel if a tweet of yours was used in one of these research studies? (n = 267)</td>
<td>4.5%</td>
<td>22.5%</td>
<td>23.6%</td>
<td>33.3%</td>
<td>16.1%</td>
</tr>
<tr>
<td>How would you feel if your entire Twitter history was used in one of these research studies? (n = 268)</td>
<td>21.3%</td>
<td>27.2%</td>
<td>18.3%</td>
<td>21.6%</td>
<td>11.6%</td>
</tr>
</tbody>
</table>
Table 4. “How Would You Feel If a Tweet of Yours Was Used in a Research Study and...” (n=268).

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Very Uncomfortable</th>
<th>Somewhat Uncomfortable</th>
<th>Neither Uncomfortable nor Comfortable</th>
<th>Somewhat Comfortable</th>
<th>Very Comfortable</th>
</tr>
</thead>
<tbody>
<tr>
<td>... you were not informed at all?</td>
<td>35.1%</td>
<td>31.7%</td>
<td>16.4%</td>
<td>13.4%</td>
<td>3.4%</td>
</tr>
<tr>
<td>... you were informed about the use after the fact?</td>
<td>21.3%</td>
<td>29.1%</td>
<td>20.5%</td>
<td>22.0%</td>
<td>7.1%</td>
</tr>
<tr>
<td>... it was analyzed along with millions of other tweets?</td>
<td>2.6%</td>
<td>18.7%</td>
<td>25.5%</td>
<td>30.0%</td>
<td>23.2%</td>
</tr>
<tr>
<td>... it was analyzed along with only a few dozen tweets?</td>
<td>16.5%</td>
<td>30.3%</td>
<td>24.0%</td>
<td>20.2%</td>
<td>9.0%</td>
</tr>
<tr>
<td>... it was from your “protected” account?</td>
<td>54.9%</td>
<td>20.5%</td>
<td>13.8%</td>
<td>6.0%</td>
<td>4.9%</td>
</tr>
<tr>
<td>... it was a public tweet you had later deleted?</td>
<td>31.3%</td>
<td>32.5%</td>
<td>20.5%</td>
<td>10.4%</td>
<td>5.2%</td>
</tr>
<tr>
<td>... no human researchers read it, but it was analyzed by a computer program?</td>
<td>2.6%</td>
<td>14.3%</td>
<td>30.5%</td>
<td>32.3%</td>
<td>20.3%</td>
</tr>
<tr>
<td>... the human researchers read your tweet to analyze it?</td>
<td>9.7%</td>
<td>27.6%</td>
<td>25.0%</td>
<td>25.4%</td>
<td>12.3%</td>
</tr>
<tr>
<td>... the researchers also analyzed your public profile information, such as location and username?</td>
<td>32.2%</td>
<td>23.2%</td>
<td>21.0%</td>
<td>13.9%</td>
<td>9.7%</td>
</tr>
<tr>
<td>... the researchers did not have any of your additional profile information?</td>
<td>4.9%</td>
<td>15.4%</td>
<td>25.1%</td>
<td>34.1%</td>
<td>20.6%</td>
</tr>
<tr>
<td>... your tweet was quoted in a published research paper, attributed to your Twitter handle?</td>
<td>34.3%</td>
<td>21.6%</td>
<td>21.6%</td>
<td>13.1%</td>
<td>9.3%</td>
</tr>
<tr>
<td>... your tweet was quoted in a published research paper, attributed anonymously?</td>
<td>9.0%</td>
<td>16.8%</td>
<td>26.5%</td>
<td>28.4%</td>
<td>19.4%</td>
</tr>
</tbody>
</table>
79.5% of the users surveyed wanted to be notified if their own tweets were used in a study.

65% of the users felt, more broadly, that researchers should have to ask for permission before using tweets in research.

That said, the answers depended heavily on what kind of study, what size of study, the type of data analysis (human or automated), etc.
Table 3. Percentage of Respondents Checking Each Contextual Factor, Ordered Highest Percentage to Lowest Percentage. 
\( n = 268 \).

<table>
<thead>
<tr>
<th>Condition</th>
<th>Total checked</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Would any of the following conditions change how you feel about a tweet of yours being used in a research study?”</td>
<td></td>
</tr>
<tr>
<td>Whether or not you are asked permission</td>
<td>67.4%</td>
</tr>
<tr>
<td>Whether or not you are informed before the research took place</td>
<td>59.6%</td>
</tr>
<tr>
<td>What the study is about</td>
<td>56.6%</td>
</tr>
<tr>
<td>The size of the dataset (i.e., is your tweet one among millions or are only a small number of tweets being analyzed)</td>
<td>45.7%</td>
</tr>
<tr>
<td>Whether the researchers are also analyzing other information about you from Twitter, such as your profile information or geo-location information</td>
<td>48.3%</td>
</tr>
<tr>
<td>Who is doing the research</td>
<td>44.9%</td>
</tr>
<tr>
<td>The type of analysis (i.e., whether a human is reading your tweet or it is only being analyzed by a computer program)</td>
<td>39.7%</td>
</tr>
<tr>
<td>Whether your tweet is quoted verbatim in a research paper</td>
<td>38.6%</td>
</tr>
</tbody>
</table>
Suggestions include "being careful about anonymization, never using real names, and making certain that nothing could link published data back to a Twitter account...publication of user identity should only occur when the benefits of doing so clearly outweigh the potential harms, or with user permission" (Fiesler and Proferes 2018:10).

Researchers “should also consider taking steps toward informing users about the research and providing them with opt-out options, if it would not compromise the research” (Fiesler and Proferes 2018:11).
This is just a start. Questions remain: e.g. “do people feel different about their data being collected from Reddit, Tumblr, Instagram, or Facebook as opposed to Twitter? What factors affect these potential differences? How much do perceptions of the ‘publicness’ of data impact comfort with the idea of being a research subject? In addition, how does the context of the use change perception?”

(Fiesler and Proferes 2018:11).
Now let’s turn to Bolander and Locher (2014), who outline some of the methodological/ethical issues likely to arise in the study of sociolinguistics and CMC:

“[U]nderstandings of public and private have changed” (Bolander and Locher 2014:17).

The distinction is “gradable and not absolute” (Bolander and Locher 2014:17).
New double dissociation between **access** and **content**:

<table>
<thead>
<tr>
<th>Private access (e.g. restricted by password)</th>
<th>Private content (sensitive information)</th>
<th>Public content (non-sensitive information)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environment</td>
<td>Emailing with a financial advisor</td>
<td>Closed group for <em>Lord of the Rings</em> fanfiction</td>
</tr>
<tr>
<td>Public access (anyone can see it)</td>
<td>Public blog of an abuse survivor</td>
<td>Most of Twitter</td>
</tr>
</tbody>
</table>
For all these reasons, guidelines are better than inexorable rules when it comes to ethics and CMC. There needs to be flexibility in order to accommodate:

- A broad range of platforms, networks, and situations – guiding principles may apply differentially.
- Unusual cases.
- Gray area.
- Pace of change online as technology continues to shift.
IRBs are already behind when it comes to what technology can do. Guidelines are already behind. What we can do with technology is constantly leaping forward. The discussion of what we should do with technology is not always so advanced.

Here’s an example.
Earlier this month, news outlets reported on an experiment done by a bunch of machine-learning folks on a very large number of YouTube videos.

The paper (Oh et al. 2019) is not yet peer-reviewed but as of May 23rd has been published to STEM manuscript repository arXiv.org, and is readily available:

“"In this paper, we study the task of reconstructing a facial image of a person from a short audio recording of that person speaking. We design and train a deep neural network to perform this task using millions of natural Internet/YouTube videos of people speaking. During training, our model learns voice-face correlations that allow it to produce images that capture various physical attributes of the speakers such as age, gender and ethnicity. This is done in a self-supervised manner, by utilizing the natural co-occurrence of faces and speech in Internet videos, without the need to model attributes explicitly"” (Oh et al. 2019:1).
ethics and CMC
ethics and CMC

They do have a section about ethics, where they argue that their experiment is about capturing generalities, not predicting individual faces.

However...
ethics and CMC

From the Fast Company report (https://www.fastcompany.com/90357561/this-ai-guesses-human-faces-based-only-on-their-voices):

“As Slate reported, not everyone that was part of the dataset was thrilled to become an unwitting part of the project. Nick Sullivan, a technology researcher at Cloudflare, tweeted out about discovering his face and voice were in the paper, and his attempt to learn how he became part of it. Many public and non-public face recognition databases rely on faces scraped from the web. For now, that kind of data harvesting may be protected by law: YouTube content is considered publicly available data, and any claims to copyright could likely be countered with a fair use argument.”
ethics and CMC

RUN THIS BY THE LEGAL DEPARTMENT, BUT RUN SUPER FAST SO THE ETHICS DEPARTMENT DOESN'T SEE IT.
Where can we start when it comes to best ethical practices for an Internet that, more than ever, is full of humans – and lots of types thereof?
Six main ideas advanced by Markham et al. (2012):

1. “[T]he greater the vulnerability of the community/author/participant, the greater the obligation of the researcher to protect the community/author/participant.”

Anyone who is young, part of a marginalized population, and/or in a dangerous situation needs to be treated with extra care.

Note that this may extend to people mentioned in the online text you are studying who are not members but who are being gossiped about.
ethics and CMC

2. Harm must be minimized, and harm is defined in context.

3. Even if an IRB says it doesn’t care, thinking of CMC research as research on human subjects is appropriate. It is easy to forget (Bolander and Locher 2014:24), but human subjects have created the data that we are working with. Particularly important in small-group settings or where people are potentially identifiable.

4. There must be balance between the rights of human subjects and the benefits of the research. In some cases, the research just isn’t worth it.
ethics and CMC

5. Deal with **every stage** of the research ethically (from selecting a community and data source through the methodology all the way to publication).

6. Do not make decisions lightly – discuss, consult, think, rethink, and choose carefully. Note that there may be additional **legal** questions (e.g. copyright).
Rachael Tatman’s Twitter ethics guidelines
(adapted from Tatman 2018):

More acceptable:
- large datasets
- analyzed automatically
- users informed about research
- anonymizing any tweets quoted

Less acceptable:
- small datasets
- analysis done by hand
- tweets that are protected or have been deleted
- quoting tweets verbatim and/or with citations
“[M]ore discussion on how to quote from one’s data in publications is needed in light of research ethics. Without quotes, linguists cannot exemplify their results and provide support for their arguments, yet quotes can easily be traced via google searches, rendering the practice of anonymization a pro forma act.”

(Bolander and Locher 2014:24)

Tatman (2018): “enough words should be changed that a reverse search isn't possible.”
What changes when we move over to a more closed platform such as Facebook?

Let’s have a look, via D’Arcy and Young (2012).

“Facebook is distinct from other online spaces (e.g. Twitter) in that its content is not aimed at the Net but rather is geared toward a constellation of known actors within the networks of individual users.”

(D’Arcy and Young 2012:533)

In other words, they outline a framework meant to apply to platforms that are shaped more by internal connections and social structure.
Facebook is much less loosely organized than Twitter.

Originally set up in 2004 for college students in the U.S. Each college where ‘thefacebook’ was available had its own inventory of groups, and you could see friend networks and class schedules.

Entire profiles, rather than mere microblogging. The ‘News Feed’ did not even emerge until 2006, and was very limited for a while.
Still lots of features Twitter lacks: groups (public or private), fan/business pages (official or not), events, **threaded replies**, different types of emotional response buttons, maps + recommendations, nuanced privacy controls.

**Messenger** is now somewhat disconnected from Facebook: used to be heavily integrated (worth noting – this has changed since D’Arcy and Young 2012).
ethics and CMC

Facebook is more of a gathering place ("agora" – D'Arcy and Young 2012:532) for groups of people.

Less anonymity: most people use their real names (D'Arcy and Young 2012:533).

Groups may exist offline and set up a Facebook page as an extension of their network (e.g. a group of 35 friends who met at a summer camp in 2003).

Or they may get to know each other and develop networks through Facebook (e.g. friends who meet in a group devoted to cosplay for Game of Thrones).

We will be discussing how networks and groups operate online more next class!
But most of the take-aways are the same as what Bolander and Locher (2014) concluded for Twitter.

D’Arcy and Young: “visibility is a continuum, and any analytical apparatus must be shaped by the structure of a given site” (2012:533).

“[I]ndividuals targeted for research in online social media are human subjects” (2012:535).

“[U]sers do not respond well to the idea of their posts being used for research without either their knowledge or consent” (2012:535).

In the case of Facebook, content might be intended only for viewing by friends, or friends-of-friends!
Specific advice for Facebook-based projects:

**Seek permission; don’t eavesdrop.** “[A]ppropriating an eavesdropper role for Facebook research is no different from surreptitiously recording conversations in non-virtual spaces” (D’Arcy and Young 2012:537).

**Allow participants to opt in** via a landing page (D’Arcy and Young 2012:539-540).

**Use your own name** and account and make it easy for participants to ask questions (D’Arcy and Young 2012:538).

**Specify straightforwardly** what will count as granting consent (D’Arcy and Young 2012:539-540).
Be thoroughly cautious with data; “anonymity and confidentiality are extremely difficult to maintain” (D’Arcy and Young 2012:537).

At the end of the study, close everything down carefully, including severing ‘friendship’ connections and informing everyone that the data collection has concluded (D’Arcy and Young 2012:541-542).
The ethics of CMC data collection is **elaborate** (D’ArCY and Young 2012:542) and cannot be taken lightly.

What are some considerations that might be necessary for newer or more **multimodal** platforms?

What if someone wants to do an ethnographic study of a set of **YouTubers** as they get to know each other online and start a group channel?

As with most Twitter users, YouTubers typically know they are talking to the public, but this does not necessarily mean they have consented to being in a sociolinguistics study!
Next class:
Monday, July 1

Preview the Lecture 3 readings and read **one or more** of them.
Post two **discussion questions** on Orbund.
**Homework 1** is due tomorrow night (submitted online).