

# Organisatorisches

# Über mich



Sonja Wang, Msc. in Computer Science mit  
Schwerpunkt Mensch-Maschine-Interaktion und  
Verteilte Systeme

# Ablauf & Voraussetzungen

- Modul findet freitags von 10-14 Uhr über Zoom statt (siehe Isis)
  - Abstimmung: bis 13:30 und dafür zwischendurch nur 15 Minuten Pause?
- Gute Englischkenntnisse sind Voraussetzung für die Teilnahme, da die Texte auf Englisch sind
- Kenntnisse aus den Modulen Diskrete Strukturen und ISDA werden vorausgesetzt
- Das Modul besteht aus mehreren Portfolioelementen, die über das Semester verteilt sind

# Portfolioelemente

## Abschluss des Moduls

### ★ Benotung

benötigt

### ☛ Prüfungsform

Portfolioprüfung

### ─ Art der Portfolioprüfung

100 Punkte insgesamt

### 🌐 Sprache

Deutsch

### ☰ Prüfungselemente

| Name                 | Punkte | Kategorie   | Dauer/Umfang |
|----------------------|--------|-------------|--------------|
| Essay 1              | 15     | schriftlich | 2 Seiten     |
| Essay 2              | 40     | schriftlich | 2 Seiten     |
| schriftlicher Test 1 | 5      | flexibel    | 30 Minuten   |
| schriftlicher Test 2 | 20     | schriftlich | 60 Minuten   |
| Vortrag              | 20     | mündlich    | 15 Minuten   |

### ★ Notenschlüssel

Dieses Prüfung verwendet einen eigenen Notenschlüssel (siehe Prüfungsformbeschreibung).

### ☰ Prüfungsbeschreibung (Abschluss des Moduls)

Note: 1.0 1.3 1.7 2.0 2.3 2.7 3.0 3.3 3.7 4.0

Punkte: 95.0 90.0 85.0 80.0 75.0 70.0 65.0 60.0 55.0 50.0

# Ablaufplan (ggf. noch geringe Änderungen)

<https://isis.tu-berlin.de/mod/page/view.php?id=1709020>

- **Abgabe Essay 1:** 24.11.23 (Analyse einer wissenschaftlichen Publikation)
- **Schriftlicher Test 1:** 07.12.23 (Fünf Fragen über den ersten Teil des Moduls)
- **Schriftlicher Test 2:** 16.02.23 (Fragen über kennengelernte Methoden)
- **Abgabe Essay 2:** 15.03.23 (Eigenes Projekt mit Datenanalyse durch eine der kennengelernten Methoden und dem zugehörigen Text)

# Social Data: Einführung

Sonja Wang, Technische Universität Berlin, 27.10.2023

---

# Danksagung

Die Folien der heutigen Vorlesung wurden inspiriert von einem Tutorial von Dr. Alexandra Olteanu.

Hier sind die Original-Folien zu finden:

<http://www.aolteanu.com/SocialDataLimitsTutorial/>

# Ziele des Moduls

- Herausforderungen aufzeigen, die bei der Sammlung und Analyse von Social Data auftreten
- Erkennen und Verstehen von Limitationen bei der Sammlung und Analyse von Social Data und deren Konsequenzen
- Kritischer Blick auf wissenschaftliche Arbeiten und Forschungsdesigns
- Kennenlernen und Anwenden von Methoden der Analyse von Social Data

# Inhalt

- Definitionen
- Anwendungsgebiete
  - Diskriminierung in Online-Werbung
- Datensammlung
- Diskussion über APIs
- Plattformspezifische Eigenschaften
- Eigenschaften von Nutzer:innen

# Anwendungsgebiete

- Erkennen und Verfolgen von Katastrophen/Epidemien
- Verstehen von psychischen Erkrankungen
- Beschreibung von linguistischen Phänomenen
- Verstehen von Suchmaschinennutzung zur Verbesserung von Resultaten
- ...

# Definitionen

## Social Data ([Olteanu et al., 2016](#))

- Alle Arten von digitalen Spuren (“*digital traces*”) produziert von oder über User:innen, mit Fokus auf Inhalten, die ausdrücklich mit der Absicht geschrieben wurden mit anderen zu kommunizieren oder interagieren
- Kommt von Social Software

## Digital Trace Data

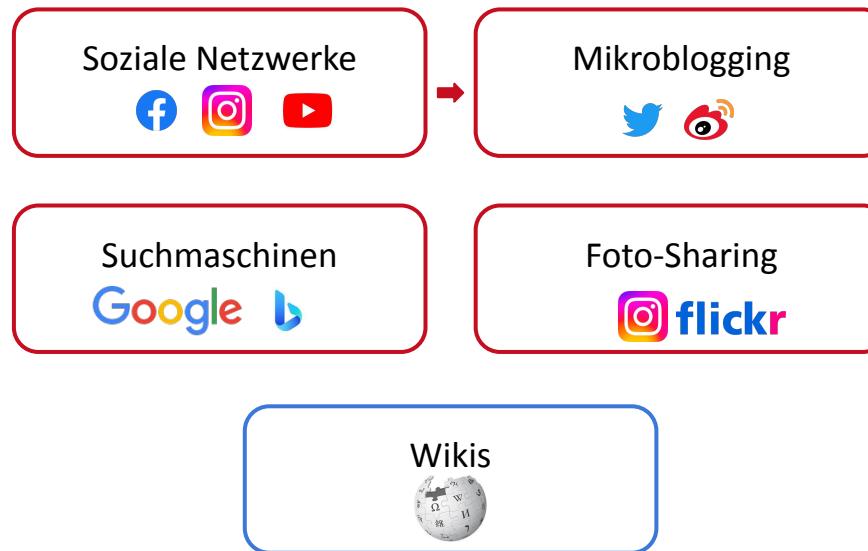
- Nebenprodukt von Aktivitäten und nicht durch bestimmtes Forschungsinstrument erzeugt (im Gegensatz zu Daten aus Umfragen, Interviews, …)
  - Müssen für Forschungszwecke angepasst werden
- Keine direkte Angabe zu sozialer Beziehung, muss von Forscher:innen interpretiert werden

# Definitionen

## Social Software

- Software, die menschliche Kommunikation, Kollaboration und Interaktion (in großen Gemeinschaften) fördert
- Umfasst eine Vielzahl von
  - Zwecken (z. B. Kommunikation, Pflege von Freundschaften, Selbstdarstellung)
  - Plattformen (z. B. soziale Medien und Netzwerke, Empfehlungs- und Frage- und Antwortseiten)
  - Interpretationsmöglichkeiten (z. B. Klicks, Likes, Shares, Profilbesuche, Nachrichten)

# Beispiele für Social Software



- Einblicke über Phänomene auf individueller Ebene, aber auch großflächige Phänomene

# Beispiele für Social Software

Kategorien nach Dr. Olteanu:

- Soziale Netzwerke (z. B. Facebook, Twitter, Weibo),
- Suchmaschinen (z. B. Google, Bing),
- Foto-Sharing (z. B. Instagram, Flickr),
- Websites mit Bewertungen (z. B. Yelp, Amazon, Douban),
- Websites für spezielle Themen oder Geschäftskontakte (z. B. LinkedIn für Jobs, BeerAdvocate für Bier),
- Standortnetzwerke (z. B. Foursquare),
- Video-Sharing (z. B. TikTok, YouTube, Vimeo),
- Crowdsourcing (z. B. Amazon Mechanical Turk)
- ...

# Suchmaschinen

Google

sonja mei wang

All Images News Videos Maps More Tools

About 1.220.000 results (0,98 seconds)

[https://www.researchgate.net › profile › Sonja-Wang](https://www.researchgate.net/profile/Sonja-Wang) :: Sonja Mei Wang, Technische Universität Berlin | TUB

Sonja Mei Wang, Technische Universität Berlin | TUB · Department of Telecommunication Systems. Master of Science. Contact.

You've visited this page 2 times. Last visit: 8/6/22

<https://www.ias.tu-berlin.de › menue › Translate this page> :: IAS: Team - Fachgebiet Internet und Gesellschaft - TU Berlin

7 Jul 2022 — Wissenschaftliche Mitarbeiter:innen: Sonja Mei Wang, Studentische Mitarbeiter:innen: Kadircan Aksoy, Sekretariat: Evelyn Adams

You've visited this page 2 times. Last visit: 5/2/22

Images for sonja mei wang :: View all →

<https://independent.academia.edu › SonjaMeiWang> :: Sonja-Mei Wang - Academia.edu

Sonja-Mei Wang studies Creative and Cultural Industries, Classical Hollywood, and National Cinemas.

Technische Universität Berlin

See photos See outside

Berlin Institute of Technology

Website Directions Save Call

Public university in Berlin

The Technical University of Berlin is a public research university located in Berlin, Germany. It was the first German university to adopt the name "Technische Universität". [Wikipedia](#)

Address: Straße des 17. Juni 135, 10623 Berlin

You visited 7 years ago

Phone: 030 3140

President: Geraldine Rauch

Total enrollment: 33,933 (2015)

Founded: April 1, 1879, Berlin

Campus: Urban

School type: Public university

Gpa

Ranking

# Diskriminierung in Online-Werbung

- Harvard-Professorin Latanya Sweeney untersuchte in einer Publikation von 2013 den Unterschied in Online-Werbung bei “white-sounding names” und “black-sounding names”
- Annahme: Personalisierte Werbung für Vorstrafen unterscheidet sich nicht zwischen *Races*



**Google ads, black names and white names, racial discrimination, and click advertising**

Latanya Sweeney

# Diskriminierung in Online-Werbung

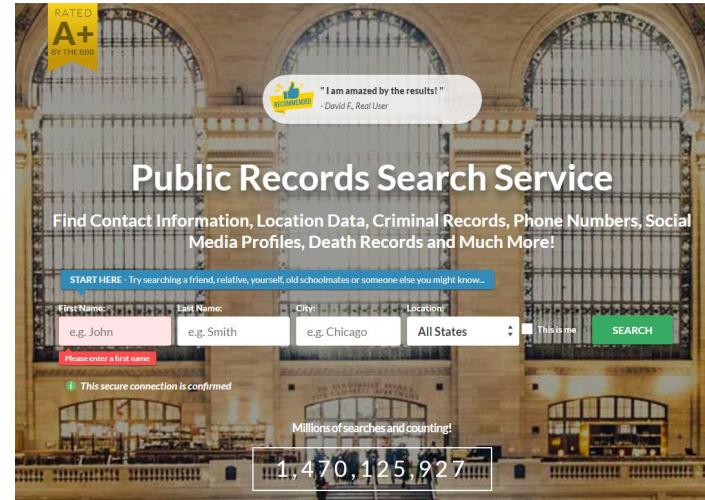
- Es wurde nach *Latanya Farrell*, *Latanya Sweeney* und *Latanya Lockett* gesucht
- Die Werbung erschien bei Google.com und auf der News Website Reuters.com
- Die Werbung scheint zu suggerieren, dass *Latanya Farrell* verhaftet wurde
  - Bei Klick auf den Link und Bezahlung erscheint jedoch, dass *Latanya Farrell* keine Vorstrafen hat

Ads by Google

[Latanya Sweeney, Arrested?](#)  
 1) Enter Name and State. 2) Access Full Background Checks Instantly.  
[www.instantcheckmate.com/](http://www.instantcheckmate.com/)

[Latanya Sweeney](#)  
 Public Records Found For: Latanya Sweeney. View Now.  
[www.publicrecords.com/](http://www.publicrecords.com/)

[La Tanya](#)  
 Search for La Tanya Look Up Fast Results now!  
[www.ask.com/La+Tanya](http://www.ask.com/La+Tanya)



The screenshot shows a search results page for "Latanya Sweeney" on a website with a background image of a grand hall. At the top, there's a yellow badge saying "RATED A+ BY THE BBB". Below it, a testimonial says "I am amazed by the results!" - David K., Real User. The main heading is "Public Records Search Service" with the subtext "Find Contact Information, Location Data, Criminal Records, Phone Numbers, Social Media Profiles, Death Records and Much More!". A search bar at the bottom has fields for First Name (e.g. John), Last Name (e.g. Smith), City (e.g. Chicago), Location (All States), and a checkbox for "This is me". A green "SEARCH" button is to the right. A message at the bottom left says "Please enter a first name" and "This secure connection is confirmed". A large counter at the bottom right shows "1,470,125,927". The URL [instantcheckmate.com](http://instantcheckmate.com) is visible at the bottom right.

# Diskriminierung in Online-Werbung

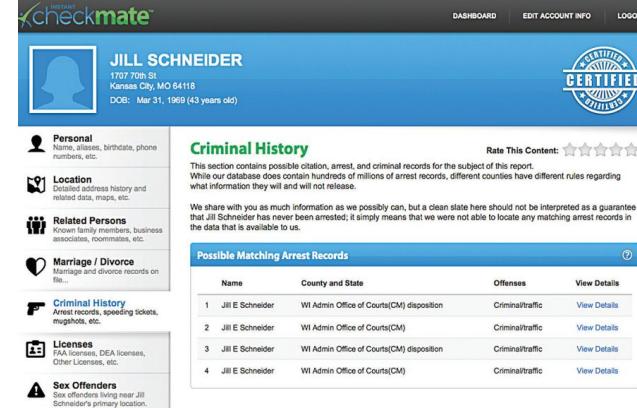
- Dasselbe Vorgehen mit *Jill Foley, Jill Schneider, Jill James*: alle drei hatten Vorstrafen
  - Im Gegensatz zu Latanya kam das Wort arrested aber nicht vor
- Die Autorin ging nun davon aus, dass Namen eine Proxy-Variable für *Race* ist

Ads related to Jill Schneider ⓘ

**Jill Schneider Art**  
[www.posters2prints.com/](http://www.posters2prints.com/)  
 Custom Frame Prints and Canvas. Shop Now, SAVE Big + Free Shipping!

**We Found Jill Schneider**  
[www.intelius.com/](http://www.intelius.com/)  
 Current Phone, Address, Age & More. Instant & Accurate Jill Schneider  
 10,256 people +1'd this page  
 Reverse Lookup - Reverse Cell Phone Directory - Date Check - Property Records

**Located: Jill Schneider**  
[www.instantcheckmate.com/](http://www.instantcheckmate.com/)  
 Information found on Jill Schneider Jill Schneider found in database.

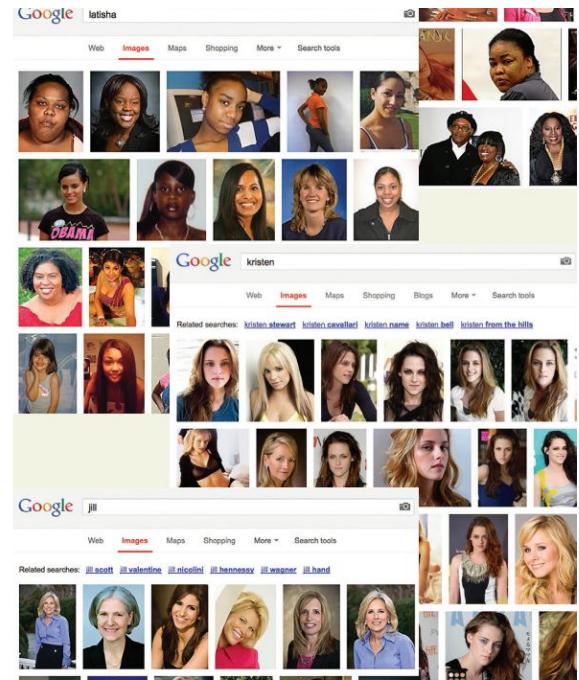


The screenshot shows a search result for 'Jill Schneider'. At the top, there's a profile picture placeholder and the name 'JILL SCHNEIDER' with an address: '1707 10th St Kansas City, MO 64118' and a birthdate: 'DOB: Mar 31, 1969 (43 years old)'. To the right is a 'CERTIFIED' badge. Below the main info are sections for 'Personal' (name, aliases, birthdate, phone numbers), 'Location' (address history, maps), 'Related Persons' (family members, business associates), 'Marriage / Divorce' (marriage records), 'Criminal History' (arrest records, speeding tickets, mugshots), 'Licenses' (FAA, DEA, Other), and 'Sex Offenders' (offenders near primary location). A sidebar on the left lists 'Ads related to Jill Schneider'. At the bottom, a table titled 'Possible Matching Arrest Records' lists four entries for 'Jill E Schneider' from 'WI Admin Office of Courts(CM)' with offenses like 'Criminal/traffic' and links to 'View Details'.

| Name               | County and State                          | Offenses         | View Details                 |
|--------------------|---|------------------|------------------------------|
| 1 Jill E Schneider | WI Admin Office of Courts(CM) disposition | Criminal/traffic | <a href="#">View Details</a> |
| 2 Jill E Schneider | WI Admin Office of Courts(CM)             | Criminal/traffic | <a href="#">View Details</a> |
| 3 Jill E Schneider | WI Admin Office of Courts(CM) disposition | Criminal/traffic | <a href="#">View Details</a> |
| 4 Jill E Schneider | WI Admin Office of Courts(CM)             | Criminal/traffic | <a href="#">View Details</a> |

# Diskriminierung in Online-Werbung

- Google Suche nach Bildern für Latanya, Latisha, Kristen und Jill zeigt, dass Latanya und Latisha mit schwarzen Gesichtern assoziiert wird, während Bilder von Kristen und Jill überwiegend weiß sind
- Unterschied von Google Suchen zu Zeitungswerbung ist, dass Werbung personalisiert ist (Interessen, geografischer Standort, ...)
  - Zwei User sehen daher eventuell unterschiedliche Werbung
- Unternehmen gibt Google u.a. Suchkriterien, Kopien von möglicher Werbung für diese



[Sweeney, 2013](#)

# Diskriminierung in Online-Werbung

- Weitere Diskussion über Google AdSense, Suchkriterien für Werbung (siehe Paper)
- Die Autorin stellte eine Liste von “black and white-identifying names” zusammen, die sie aus Geburten von 1974 bis 1979 übernahm und einer Liste aus dem Buch Freakonomics + die Namen Latanya und Latisha
- Mehr als >2000 Namen (Vorname+Nachname) wurden gesucht und 78% hatten mindestens eine Werbung für Daten in „Public Records“ dieser Person

| White Female | Black Female | White Male | Black Male |
|--------------|--------------|------------|------------|
| Allison      | Aisha        | Brad       | Darnell    |
| Anne         | Ebony        | Brendan    | Hakim      |
| Carrie       | Keisha       | Geoffrey   | Jermaine   |
| Emily        | Kenya        | Greg       | Kareem     |
| Jill         | Latonya      | Brett      | Jamal      |
| Laurie       | Lakisha      | Jay        | Leroy      |
| Kristen      | Latoya       | Matthew    | Rasheed    |
| Meredith     | Tamika       | Neil       | Tremayne   |
| Molly        | Imani        | Jake       | DeShawn    |
| Amy          | Ebony*       | Connor     | DeAndre    |
| Claire       | Shanice      | Tanner     | Marquis    |
| Emily*       | Aaliyah      | Wyatt      | Darnell*   |
| Katie        | Precious     | Cody       | Terrell    |
| Madeline     | Nia          | Dustin     | Malik      |
| Katelyn      | Deja         | Luke       | Trevon     |
| Emma         | Diamond      | Jack       | Tyrone     |
|              | Latanya      |            |            |
|              | Latisha      |            |            |

# Diskriminierung in Online-Werbung

**Was hat solche Werbung für Auswirkungen auf Betroffene?**

**Welchen Einfluss kann diese Publikation auf dieses Problem haben?**

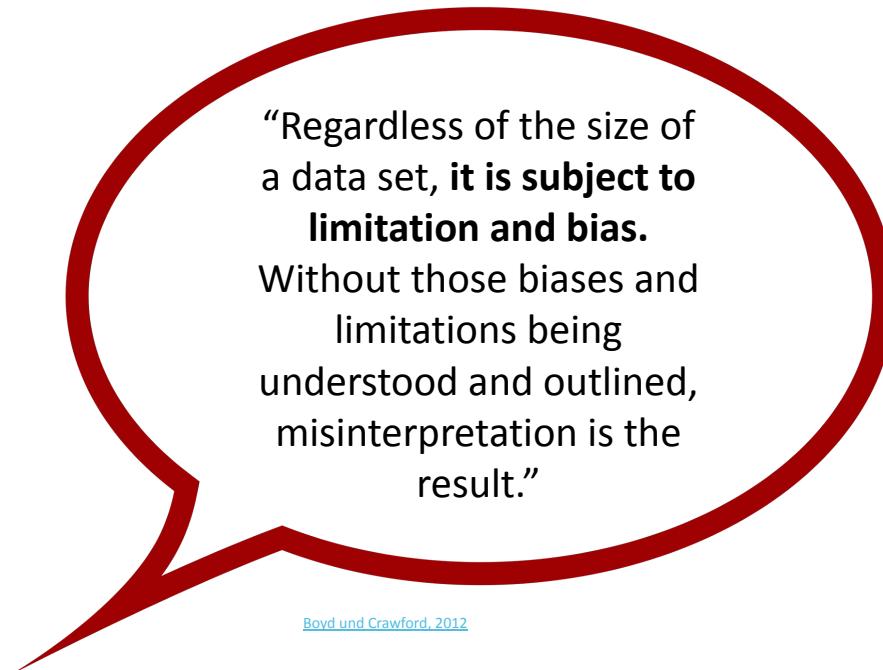
# Datensammlung

**Scientists Warn About Bias In The Facebook And Twitter Data Used In Millions Of Studies** [Parnell, 2014](#)



Probleme mit: Repräsentation von Populationen, Repräsentation von menschlichem Verhalten, Methoden

# Bias (Verzerrung, Voreingenommenheit, ...)



“Regardless of the size of a data set, **it is subject to limitation and bias.** Without those biases and limitations being understood and outlined, misinterpretation is the result.”

[Boyd und Crawford, 2012](#)

# Datensammlung

Plattformen haben nicht immer eine Schnittstelle, um Daten abzufragen oder verbieten die Abfrage explizit (z.B. LinkedIn und TripAdvisor).

**LinkedIn:** “[You agree that you will] not develop, support or use software, devices, scripts, robots or any other means or processes (including crawlers, browser plugins and add-ons or any other technology) to scrape the Services or otherwise copy profiles and other data from the Services”

[User Agreement | LinkedIn](#)

# Datensammlung

Limitationen auf Plattformen (alte Twitter-API für verschiedenen Personengruppen, z.B. Forscher:innen, *Hydrating* von Tweets):

| Key benefits                  |  |
|-------------------------------|--|
|                               | Access Twitter's real-time and historical public data with additional features and functionality that support collecting more precise, complete, and unbiased datasets. <a href="#">More details on included endpoints</a> |
| <b>Tweet cap</b>              | 10 million Tweets / month  |
| <b>Query rules</b>            | 1024 characters, 1000 streaming rules  |
| <b>Streaming rates</b>        | 50 requests / 15 minutes, per app  |
| <b>Technical support</b>      | Developer documentation, tutorials, support content, and community forums  |
| <b>Cost</b>                   | Free   |
| <b>Ready to move forward?</b> | <a href="#">Apply →</a>  |

# Datensammlung

API-Zugänge können sich verändern oder ganz wegfallen

The change in question was part of a project to shut down free access to the Twitter API, *Platformer* can now confirm. On February 1st, the company announced it will no longer support free access to its API, which effectively ended the existence of third-party clients and dramatically limited the ability of outside researchers to study the network. The company has been building a new paid API for developers to work with.

[The Verge, 2023](#)

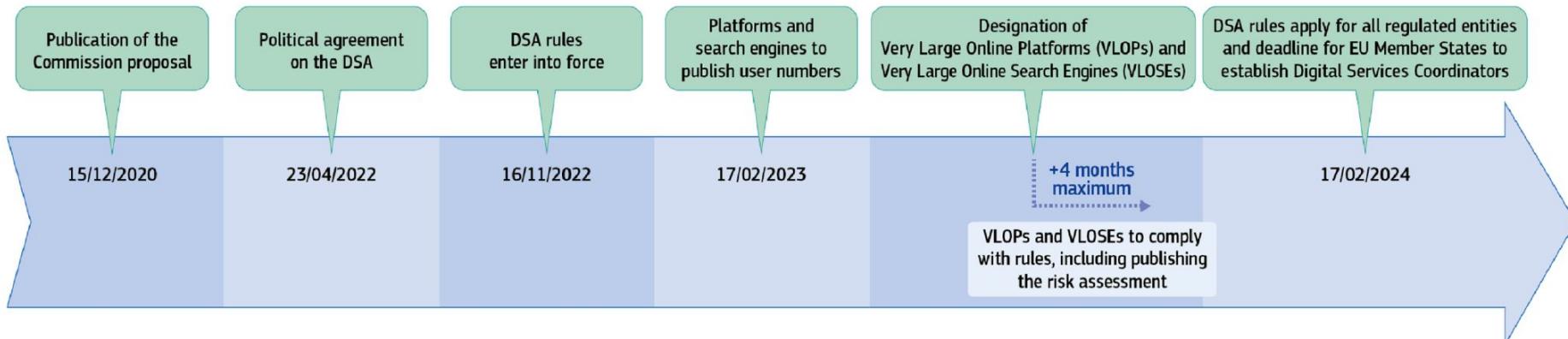
# Digital Services Act als “Rettung”?

In the next few weeks, we will start rolling out a new suite of tools for researchers: Meta Content Library and API. The Library includes data from public posts, pages, groups, and events on Facebook. For Instagram, it will include public posts and data from creator and business accounts. Data from the Library can be searched, explored, and filtered on a graphical user interface or through a programmatic API. Researchers from qualified academic and research institutions pursuing scientific or public interest research topics will be able to apply for access to these tools through partners with deep expertise in secure data sharing for research, starting with the University of Michigan’s Inter-university Consortium for Political and Social Research. These tools will provide the most comprehensive access to publicly-available content across Facebook and Instagram of any research tool we have built to date and also help us meet new data-sharing and transparency compliance obligations.

[Meta, 2023](#)

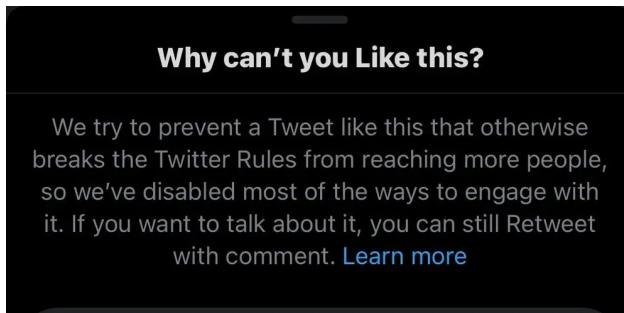
# Digital Services Act als “Rettung”?

## Timeline for Digital Services Act

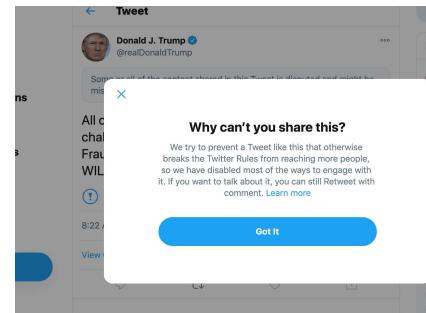


# Datensammlung

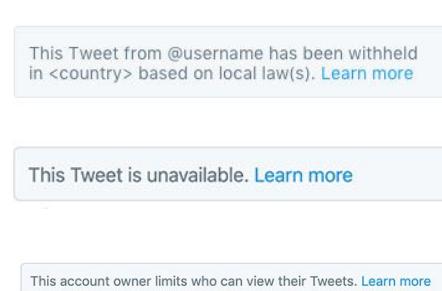
- Datensätze enthalten normalerweise nur öffentliche Inhalte (keine die gelöscht wurden oder deren Sichtbarkeit auf privat gesetzt wurde)
- Einfluss von Tweets oder Posts eventuell nicht nachvollziehbar, da bestimmte Funktionen gesperrt sind



[Perrett, 2020](#)



[Graham, n.d.](#)



[Notices on Twitter](#)

# Datensammlung

- Bestimmte Daten wie Views, Klicks und Keystrokes in der Regel nicht für Forscher:innen zugänglich, aber andere wie Retweets oder Likes schon
  - Ausnahmen für Forschung, die direkt bei Facebook etc. stattfindet
  - Studie zu *self-censorship*: Nachrichten



Figure 1. Screenshot of the “composer” (top) and comment box (bottom) HTML elements on Facebook.

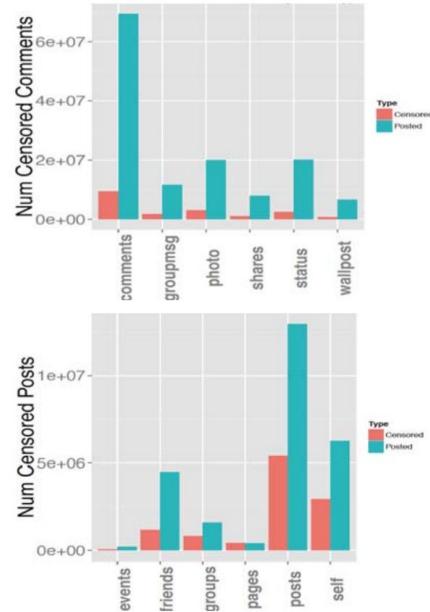


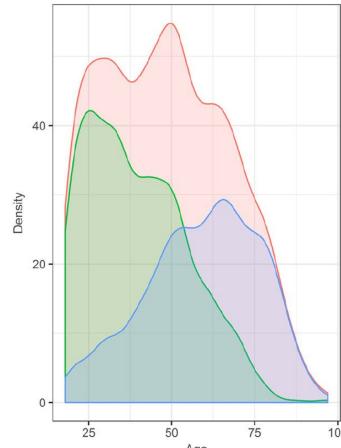
Figure 3. Number of censored (red) and shared (blue) comments (top) and posts (bottom) grouped by location. Location represents where the content was shared (e.g., “groupmsg” means comments on group posts). The “posts” and “comments” locations are the values aggregated across all locations.

# Datensammlung

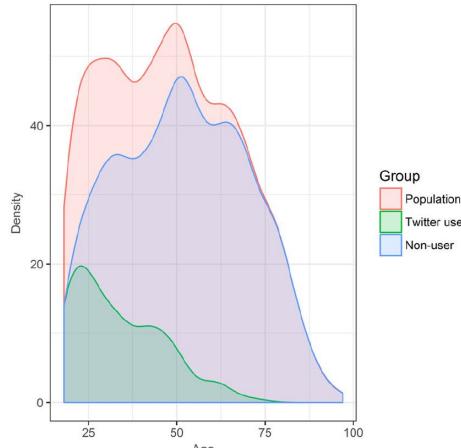
- Art der Suchanfragen beeinflusst Ergebnisse, denn Anfragen erfolgen über
  - Schlüsselworte
  - Usernamen
  - Zeitliche Intervalle
  - Ortsinformationen
    - 2014: nur 2.9% der Twitter-User:innen verwendete geolocation, 1% auf Facebook
    - eventuell falsch angegeben auf Profil

# Datensammlung: Selection Bias

- Repräsentativität der Stichprobe beachten
- Instagram besonders attraktiv für 18- bis 29-jährige, African-Americans, Latinos, Frauen und Stadtbewohner:innen
- Pinterest-User:innen vor allem weiblich im Alter von 25-34, mit durchschnittlichem Jahreseinkommen von 100.000\$

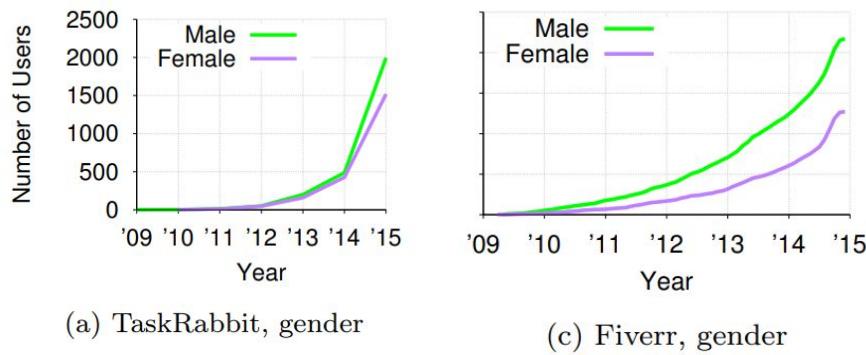


Pew Research Center, 2021



# Datensammlung: Temporal Bias

Die Demographie von User:innen kann sich über die Zeit verändern.



[Hannák et al., 2017](#)

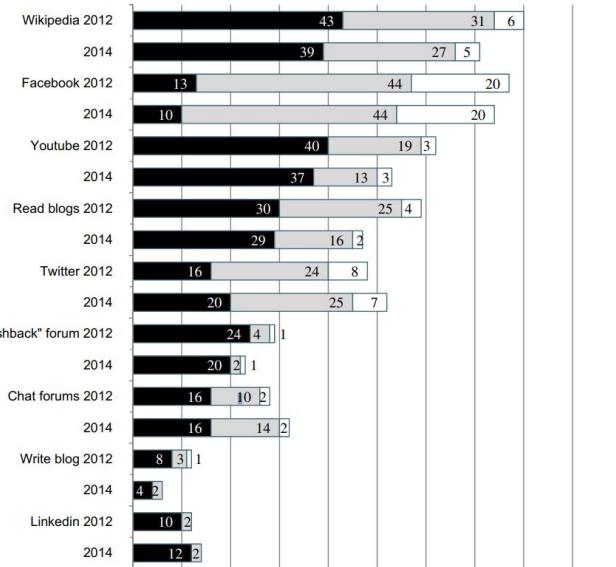


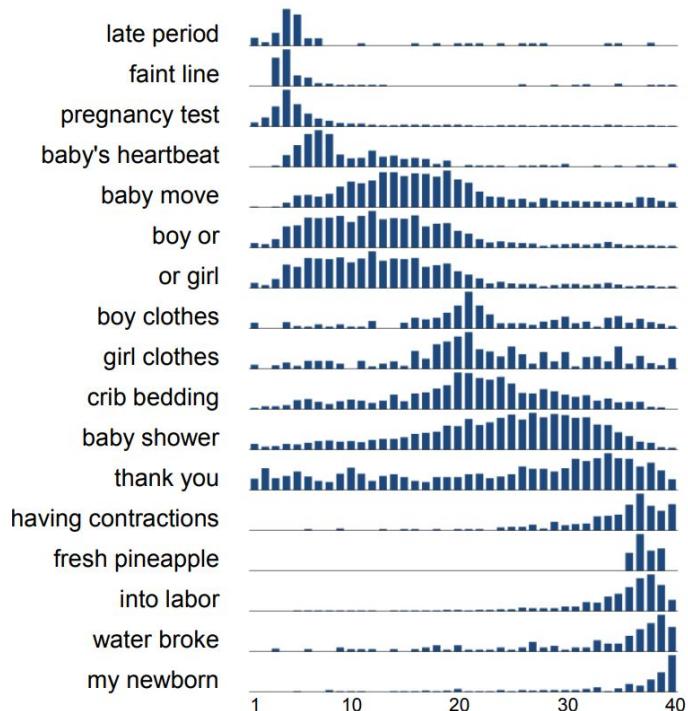
FIGURE 1

The use of different social media (%) in 2012 ( $N = 1305$ ) and 2014 ( $N = 957$ ). Black series: weekly use; grey series: daily use; white series: 24/7 use. The respondents were asked: "How often do you use...?" The response alternatives were: "Never", "Now and then each month", "Now and then each week", "Now and then every day", and "All of the time" (24/7). The "Flashback" forum is a Swedish forum targeting a broad range of topics, including "national politics" (often with a strong populist and anti-immigration sentiment), "ongoing crime news" (often speculative and sensationalist) and "entertainment gossip".

# Datensammlung: Temporal Bias

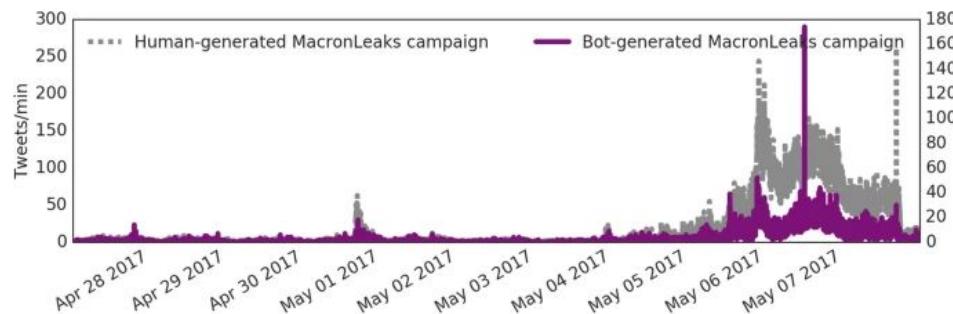
Manche digital traces treten nur zu bestimmten Zeiten/Lebensabschnitten auf.

- Schwangerschaft, Jobsuche, Erkrankungen wie Krebs

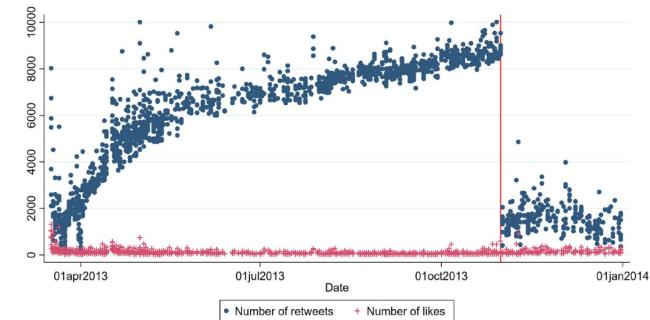


# Datensammlung: Bots

- Schätzungen, dass 9-15% Accounts auf Twitter Bots sind (2017)
- Tritt vor allem im politischen Zusammenhang immer wieder auf (“Manipulation und Täuschung” von User:innen)
- Vermutung, dass Bots Kaskaden von *disinformation* hervorrufen können



Ferrara, 2020



Likes und Retweets zu Präsident Maduros Tweets

Morales, 2019

# Datensammlung: kulturelle Unterschiede

Tweets verschiedener Sprachen weisen unterschiedliche Eigenschaften auf.

- 18% der deutschen Tweets verwenden Hashtags, aber nur 5% der japanischen
- 39% indonesischer Tweets waren retweets, aber nur 7% der japanischen
- Deutsche Tweets enthalten deutlich öfter URLs als Tweets anderer Sprachen

| Language   | URLs | Hashtags | Mentions | Replies | Retweets |
|------------|------|----------|----------|---------|----------|
| All        | 21%  | 11%      | 49%      | 31%     | 13%      |
| English    | 25%  | 14%      | 47%      | 29%     | 13%      |
| Japanese   | 13%  | 5%       | 43%      | 33%     | 7%       |
| Portuguese | 13%  | 12%      | 50%      | 32%     | 12%      |
| Indonesian | 13%  | 5%       | 72%      | 20%     | 39%      |
| Spanish    | 15%  | 11%      | 58%      | 39%     | 14%      |
| Dutch      | 17%  | 13%      | 50%      | 35%     | 11%      |
| Korean     | 17%  | 11%      | 73%      | 59%     | 11%      |
| French     | 37%  | 12%      | 48%      | 36%     | 9%       |
| German     | 39%  | 18%      | 36%      | 25%     | 8%       |
| Malay      | 17%  | 5%       | 62%      | 23%     | 29%      |

# Datensammlung: APIs

- Funktionsweisen der APIs nicht immer bekannt  
(Twitter Streaming API vs. kostenpflichtige Firehose API)
- Unterschiedliche Datensätze nach Anwendung verschiedener APIs möglich

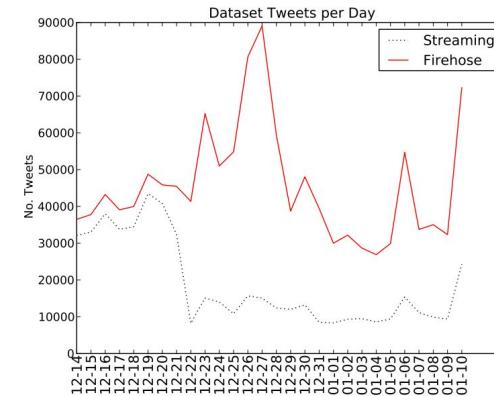


Figure 2: Raw tweet counts for each day from both the Streaming API and the Firehose.

# Diskussion (10 Minuten)

Stellen Sie sich vor, dass Sie herausfinden möchten, wie viele Prozent der User:innen auf GitHub weiblich sind anhand ihrer Namen. Dazu verwenden Sie eine Name-zu-Gender-API. Auf der nachfolgenden Slide sehen Sie einige APIs im Vergleich. Bitte diskutieren Sie in den Kleingruppen (4 Personen) folgende Fragen:

- 1) Welche API würden Sie verwenden und warum?**
- 2) Welche Biases können durch die Verwendung der API auftreten?**
- 3) Wofür könnte die Verwendung von Gender APIs sonst noch genutzt werden?**

# Diskussion

|  | <a href="#">Gender API</a> | <a href="#">Genderize</a>         | <a href="#">NameAPI</a> | <a href="#">Namsor</a> |
|--|----------------------------|-----------------------------------|-------------------------|------------------------|
| Datenbankgröße                             | 6.084.389                  | 114.541.298                       | ?                       | 7.106.962.120          |
| Beachtung von Nachnamen                    | Ja                         | Nein                              | Ja                      | Ja                     |
| Beachtung von Land                         | Ja                         | Ja                                | Ja                      | Ja                     |
| Unterstützung nicht-lateinischer Alphabete | ?                          | ?                                 | Ja                      | Ja                     |
| Kostenlose Abfragen/Monat                  | 500/Monat                  | 1000/Tag,<br>unlimitiert im Monat | 10.000/Monat            | 5.000/Monat            |
| Abonnementkosten/Monat                     | 67,15€/100.000 req.        | 9\$/100.000 req.                  | 50€/100.000 req.        | 130\$/100.000 req.     |

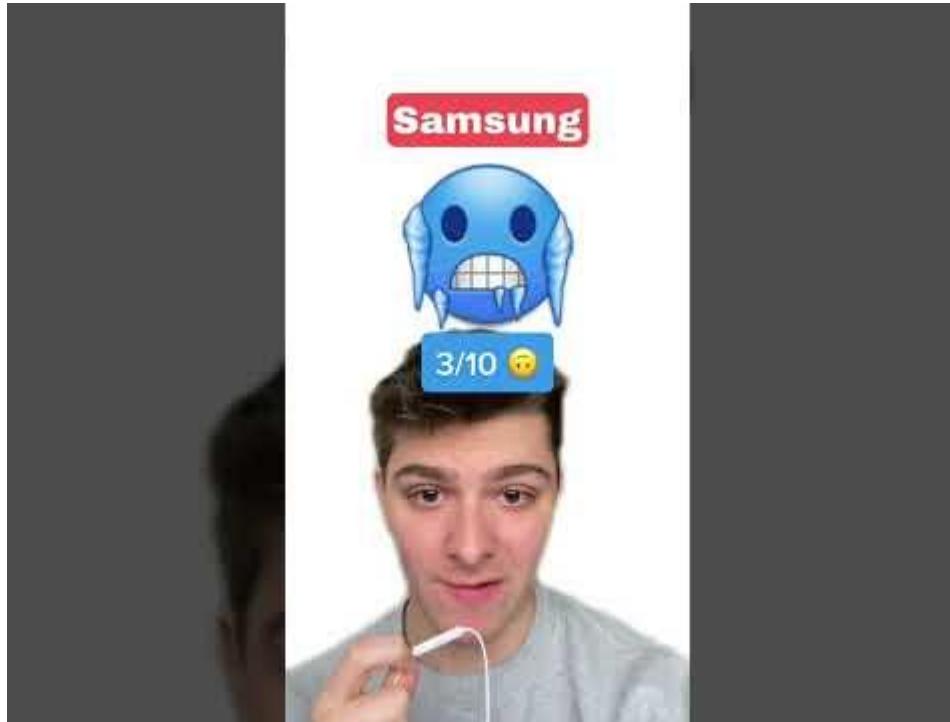
# Zusammenfassung

- Die Analyse von Social Data kann positive (und negative) Auswirkungen auf bestimmte Personengruppen haben
- Während der Analyse treten Biases auf (also auch schon bei der Datensammlung)
- Nicht alle Typen von erfassbaren Daten können auch von Forscher:innen ausgewertet werden
- Die Demographie von User:innen unterscheidet sich auf verschiedenen Plattformen

# Gestaltung von Plattformen

---

# Unterschiede in Emojis

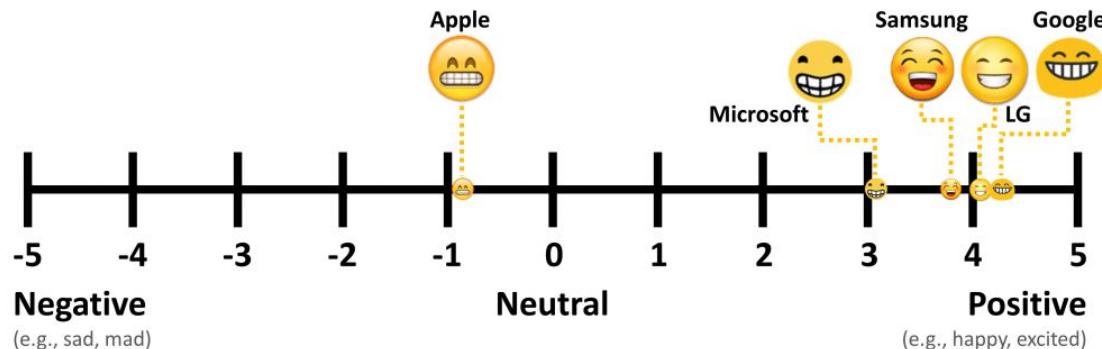


# Unterschiede in Emojis

Plattformspezifische Eigenschaften haben einen Einfluss auf Social Data, zum Beispiel unterschiedliche Darstellung von Emojis.

## Same Emoji + Different Smartphone Platform = Different Emotion

*For example, if you send the Apple emoji to a Google Nexus, they'll see the Google emoji, and vice versa!*



[Miller, 2016](#)

# Unterschiede in Emojis

Plattformspezifische Eigenschaften haben einen Einfluss auf Social Data, zum Beispiel unterschiedliche Darstellung von Emojis.

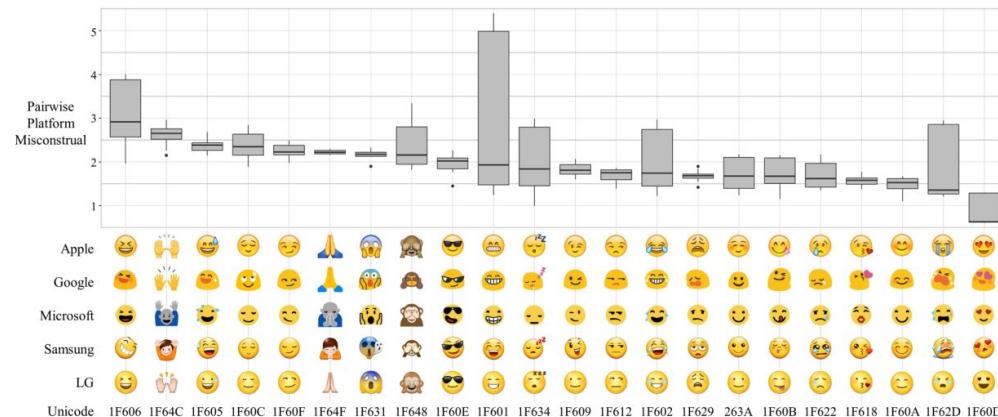


Figure 1. Across-platform sentiment misconstrual scores grouped by Unicode. Each boxplot shows the range of sentiment misconstrual scores across the five platforms. They are ordered by decreasing median platform-pair sentiment misconstrual, from left to right.

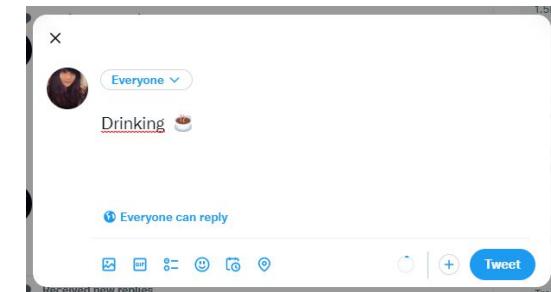
[Miller et al., 2016](#)

# Analyse von Emojis

- Emojis können je nach Kontext unterschiedlich interpretiert werden
- Es gibt Methoden wie Sentiment Analysis, die sich damit beschäftigen
- Nicht-antropomorphe Emojis können auch Wörter ersetzen

**Table 1:** Tweet examples with emojis. The sentiment ground truth is given in the second column. The examples show that inconsistent sentiments exist between emojis and texts.

| Emoji | Sentiment | Tweets  |
|-------|-----------|---|
| 😊     | Positive  | Good morning Justin!!! I hope u have an amazing Friday :) Don't forget to smile 😊.<br>That's awesome :) I'm super keen to hear/see it all 😊.  |
|       | Negative  | I really hate times square personally it's too busy (I'm claustrophobic 😩).<br>Not very good when your sat waiting for your food and there is a queue forming to complain to a manager 😩. |
| 😢     | Positive  | This weather is perfect! 😌 It's just the change I needed.<br>The dresses I ordered arrived this morning and they're so pretty 😌.  |
|       | Negative  | Worst headache ever and feel so sick, mum where are you 😢.<br>This nurse always comes mad early in the morning I'm mad tired 😢.   |



# Neue Funktionen verändern Daten

Die Funktionen von Plattformen können sich über die Jahre verändern.

- Facebook fügt 2008 die “People you may know”-Funktion hinzu
- Dadurch änderten sich Freundschaftsbeziehungen

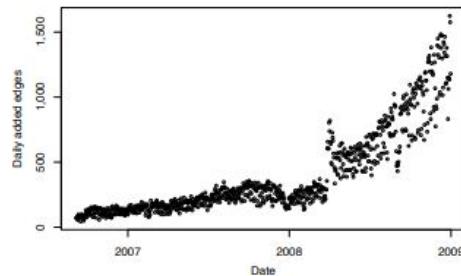


FIGURE 2.1: Observed edges added (friendship ties made) in Facebook New Orleans data.

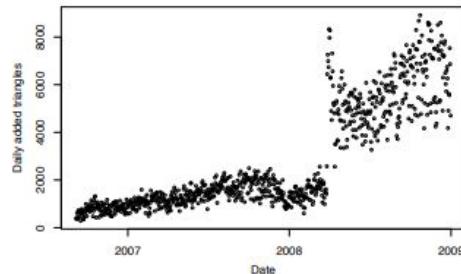


FIGURE 2.2: Triangles created with the added edges in Facebook New Orleans data.

[Malik, 2018](#)

# Social Influence Bias

Recommender Systems zeigen durchschnittliche Ratings in der Regel an, bevor neue Ratings angenommen werden, was diese neuen Ratings beeinflusst.

- Tendenz sich an eine wahrgenommene Norm in der Community anzupassen

## ERGEBNISSE

Erfahre mehr über diese Ergebnisse.



COMPO SANA Grünpflanzerde und Palmenerde mit 12 Wochen Dünger für alle Zimmerpflanzen, Balkonpflanzen, Palmen und Farne,...  
 ★★★★☆ ~ 7.703

6,99€ (0,70 €/l) UVP: 7,49€  
 ✓prime GRATIS 1-Tages-Lieferung  
 Lieferung bis Morgen, 26. Oktober



Flora&gärd Aktiv Garten-Pflanzerde 50 Liter - Erde mit Langzeitdünger zum Pflanzen von Gehölzen, Sträuchern, Hecken und Stauden  
 ★★★★★ ~ 349

11,49€ (0,23 €/l)  
 KOSTENLOSE Lieferung bis Dienstag, 1. November für Prime-Mitglieder  
 ⓘ Klein-Unternehmen



Pflanzerde 70 Liter Blumenerde Plantop  
 ★★★★★ ~ 431

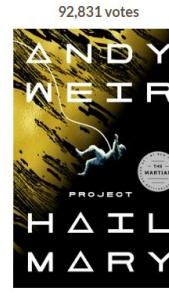
15,10€ (0,22 €/l)  
 Lieferung Donnerstag, 27. Oktober – Freitag, 28. Oktober  
 KOSTENLOSE Lieferung  
 Andere Angebote: 14,75 € (6 neue Artikel)



COMPO BIO Anzuchterde und Kräutererde, Erde für Aussaaten, Kräuter, Stecklinge und Jungpflanzen, Blumenerde, Torffrei,...  
 ★★★★★ ~ 6.222

7,29€ (1,46 €/l)  
 ✓prime KOSTENLOSE Lieferung bis Freitag, 28. Oktober  
 Andere Angebote: 6,95 € (6 neue Artikel)

All Nominees • 281,584 votes total



92,831 votes

Want to Read ▾  
 Rate this book  
 ★★★★★



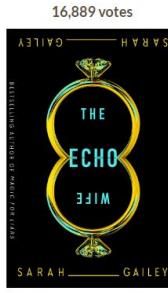
56,284 votes

Want to Read ▾  
 Rate this book  
 ★★★★★



16,980 votes

Want to Read ▾  
 Rate this book  
 ★★★★★



16,889 votes

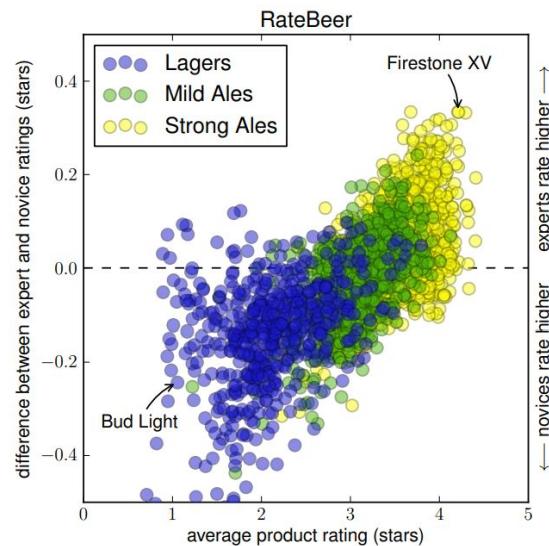
Want to Read ▾  
 Rate this book  
 ★★★★★

# Eigenschaften von Nutzer:innen

---

# Expertenwissen

- Um Produkte zu empfehlen sollten nicht nur “Geschmäcker” betrachtet werden, sondern auch Erfahrung und Expertise
- Novizen und Experten bewerten Produkte anders
  - Annahme: “acquired taste”, ändert sich über die Zeit



**Figure 1:** Beloved products are most beloved by experts; hated products are most hated by experts.

# Soziale Normen

Inhalte, die User:innen teilen, unterscheiden sich auf verschiedenen Plattformen.



?

- LinkedIn als “Facebook in a suit” bezeichnet
- Hauptsächlich benutzt, um Fähigkeiten zu hervorheben und Stärken zu bewerben

# Lurker

- Strategien für Bereinigung von Daten beinhalten:
  - Entfernen von Ausreißern (nicht aktive oder sehr aktive User:innen),  
Entfernen von Bots, Filterung nach Sprache
- Nicht/wenig aktive User:innen werden auch als “Lurker” bezeichnet
- Interessant für zielgerichtete Werbung und Erfassung der Stimmung in der Gesamtbevölkerung z.B. bei Wahlen (geringe Vorhersagekraft bei Umfragen auf Social Media)

# Lurker

- Annahme: User:innen praktizieren oft Selbstzensur; Interesse, trotz inaktivem Verhalten
- In Posts nur Meinungen über Themen von denjenigen, die auch Posts verfassen  
→ Analyse von “i-silent” und “i-active” User:innen, i steht für “issue specific”

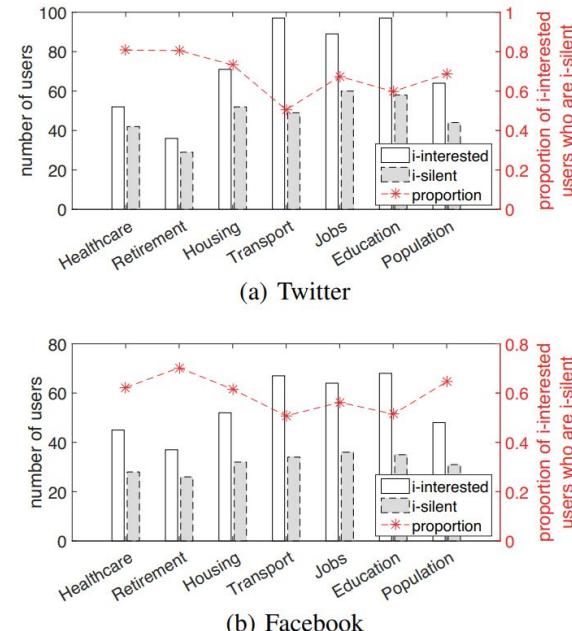


Figure 2: The proportion of  $i$ -interested users who are silent on  $i$ .

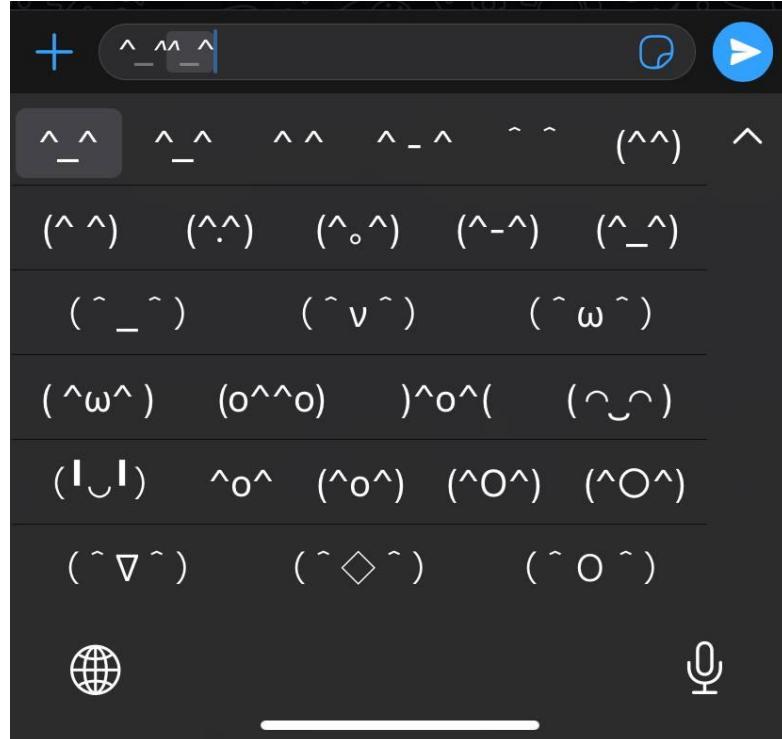
# Emoticons: Kulturelle Unterschiede

Vertikale Emoticons werden in asiatischen Ländern mehr verwendet als horizontale.  
 (~58% der Daten aus USA, 1,45% aus Japan, 0,3% aus Südkorea)

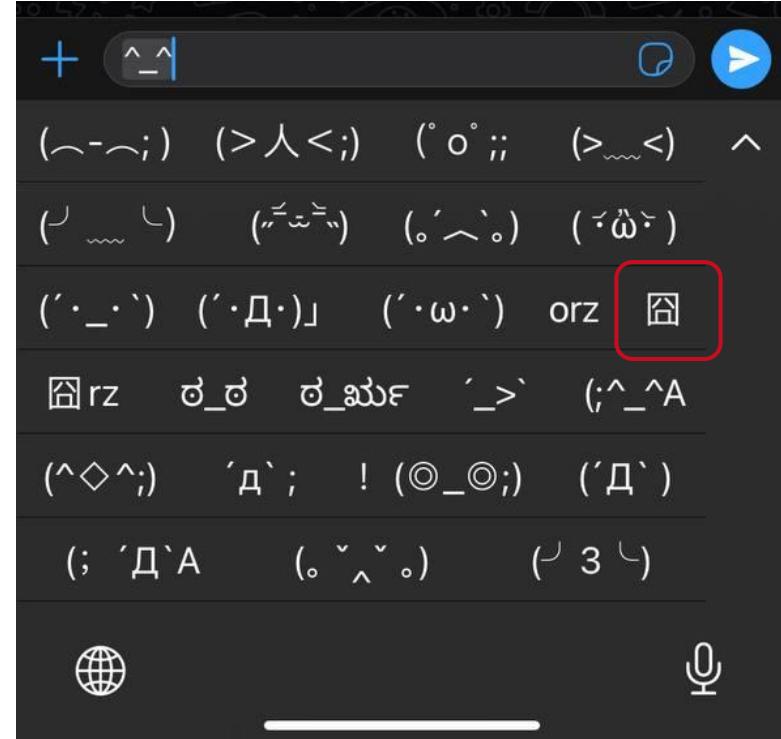
| Style   | Normative form | Affect   | Meaning             | Variant examples          |
|---|----------------|----------|---------------------|---------------------------|
| Horizontal<br>(expression based on the mouth shape) | :)             | positive | happy               | wink ;)                   |
|   | :(             | negative | sad                 | mouth :)) :(((            |
|   | :o             | neutral  | surprise            | nose :-) :-(- :-[         |
|   | :P             | positive | tongue sticking out | tear :'( :*(              |
|   | :D             | positive | laugh               | forehead or hair ¿:( =:-) |
| Vertical<br>(expression based on the eye shape)     | ^^             | positive | happy               | chin (^)                  |
|   | T_T            | negative | sad                 | mouth ^__^ T__T           |
|   | @@             | neutral  | surprise            | nose ^.^ ^.^ T.T          |
|   | -_-            | negative | absent-minded       | sweat ^; -_-;;            |
|   | o.o            | positive | curious, amazing    | eyebrow _-_-^             |

Table 2: Two different styles of emoticons: horizontal (popular in western countries) and vertical (popular in eastern countries).

# Emoticons: Kulturelle Unterschiede



Screenshot von japanischer iPhone-Tastatur



Screenshot von chinesischer iPhone-Tastatur

# Anonymität und Meinung

- Twitter: Einige User:innen identifizierbar, einige anonym
  - Tweets über sensible Themen wie Pornographie, Sexualität, Religion, Rassismus ... untersucht
- Korrelation zwischen Sensibilität des Inhaltes und Anonymität des/der User:in

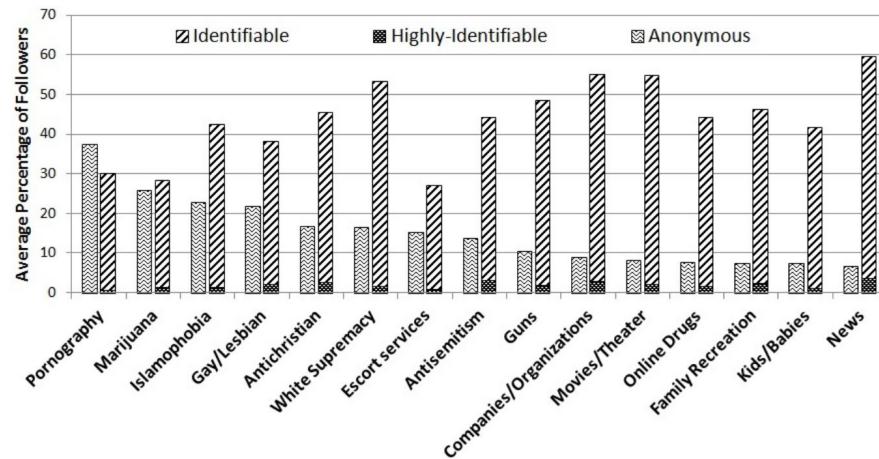


Figure 1: Sensitive and Non-Sensitive Twitter Account Categories: Follower Distribution

# Trolling

**Wenn ihr den Begriff Trolling hört, was versteht ihr darunter?**

# Trolling - Gruppendiskussion (10min)

[Paper für Gruppe 1](#), [Paper für Gruppe 2](#)

Lesen Sie sich den Abstract und die Einleitung durch und beantworten Sie folgende Frage:

- 1) Was verstehen die Forscher:innen unter Trolling?

Schauen Sie in den Methodenteil und beantworten Sie folgende Frage:

- 2) Welche Daten verwenden die Forscher:innen?

Kann aus dem Abstract oder der Konklusion entnommen werden:

- 3) Zu welchem Ergebnis kommen die Forscher:innen?

# Definitionen von Trolling (Fortsetzung)

Aus Paper 1: Hardaker 2010: trolls are characterized by four distinct traits: “aggression, deception, disruption, and success.”

Trolling was **operationalized** as comments that “purposely or deliberately commented to get a rise out of the other people on the subreddits, whether it was successful or not.” Successful trolling were troll comments that received replies, while unsuccessful trolling comments did not.

# Definitionen von Trolling (Fortsetzung)

**Aus Paper 2:** Russische - Pro-Trump, Iranische: Anti-Trump, aber: nicht klar, wie sie diese Informationen einschätzen/keine richtige Definition von Trolling gegeben (lediglich Bezug auf Datensätze von Twitter und Reddit).

# Daten Paper 2: Trolling?

**Twitter.** On October 17, 2018, Twitter released a large dataset of Russian and Iranian troll accounts [18]. Although the exact methodology used to determine that these accounts were state-sponsored trolls is unknown, based on the most recent Department of Justice indictment [9], the dataset appears to have been constructed in a manner that we can assume essentially no false positives, while we cannot make any postulation about false negatives. Table 1 summarizes the troll dataset.

| Platform | Origin of trolls | # trolls | # trolls<br>with tweets/posts | # of tweets/posts |
|----------|------------------|----------|-------------------------------|-------------------|
| Twitter  | Russia           | 3,836    | 3,667                         | 9,041,308         |
|          | Iran             | 770      | 660                           | 1,122,936         |
| Reddit   | Russia           | 944      | 335                           | 21,321            |

**Table 1: Overview of Russian and Iranian trolls on Twitter and Reddit.** We report the overall number of identified trolls, the trolls that had at least one tweet/post, and the overall number of tweets/posts.

## Update on Twitter's review of the 2016 US election

By [Twitter Public Policy](#)

Friday, 19 January 2018



Updated on January 31, 2018

We have expanded the number of people notified about interactions with Twitter accounts potentially connected to a propaganda effort by a Russian government-linked organization known as the Internet Research Agency. Our notice efforts are focused on certain types of interactions, and they will not encompass every person that ever saw this content. Our goal in providing these notifications is to advance public awareness of and engagement with the important issues raised in our blog post, and provide greater transparency to our account holders and the public.

# Andere Definitionen von Trolling

Broniatowski et al., 2018 (Über Impfungen auf Twitter) Much health misinformation may be promulgated by “bots”—accounts that automate content promotion—and “trolls”—**individuals who misrepresent their identities with the intention of promoting discord.** One commonly used online disinformation strategy, amplification, seeks to create impressions of false equivalence or consensus through the use of bots and trolls.

## Andere Definitionen

[Zannettou et al., 2020](#) (Russische Trolls auf Twitter): State-sponsored organizations are increasingly linked to efforts aimed to exploit social media for information warfare and manipulating public opinion. Typically, their activities rely on a number of social network accounts they control, aka **trolls, that post and interact with other users disguised as “regular” users. These accounts often use images and memes, along with textual content, in order to increase the engagement and the credibility of their posts.**

# Andere Definitionen von Trolling

Saeed et al., 2022 (State-Sponsored Trolls auf Reddit): Disinformation campaigns are often orchestrated by state actors, with the goal of polarizing public discourse or pushing talking points to favor particular agendas [37, 38]. To do so, **malevolent actors instrument so-called troll accounts to engage in discussions among each other and with real users, pushing certain narratives and sharing false information** [65].

- Trolls als “manually-controlled” Accounts
- Paper enthält eine ausführliche Beschreibung über Schritt-für-Schritt-Vorgehen in der Detektion von Trolls

# Konstrukte

- “Trolling” ist zunächst einmal ein Konstrukt
- Ein **Konstrukt** ist ein nicht empirisch erkennbarer Sachverhalt innerhalb einer wissenschaftlichen Theorie
  - sind gedanklicher/theoretischer Natur
- Der Sachverhalt existiert, aber wird aus anderen, messbaren Sachverhalten (**Indikatoren**) erschlossen
- Dieser Prozess der Erschließens heißt **Operationalisierung**
- Beispiel: das Konstrukt Intelligenz kann nicht direkt beobachtet werden, aber über Indikatoren wie Leistung in Intelligenztestaufgaben gemessen werden

# Konzepte

- Der Begriff Konstrukt ist eng verwandt mit dem Begriff Konzept
- Das Konzept betont stärker, dass es sich um einen wissenschaftlichen oder theoretischen Begriff handelt
- Beim Konstrukt liegt die Betonung auf der Nicht-Beobachtbarkeit
- Es gibt verschiedene Arten von Definitionen
  - Nominaldefinition
  - Realdefinition
  - Operationale Definition

# Operationale Definition

- Verknüpfung eines Begriffs mit einem empirischen Sachverhalt über Korrespondenzregeln
- In welchen (messbaren) Variablen äußert sich das Konstrukt?
- Wie kann man messen, dass etwas vorliegt/inwieweit es vorliegt und was muss man dafür tun (die Operationen des Messens)
- Messen kann qualitativ oder quantitativ sein

# Codebücher

- Codebuch ist ein Dokument, welches die Konstrukte, Variablen und Messanleitungen in einer Weise als Anleitung formuliert, dass verschiedene Coder dies mit möglichst gleichen Resultaten messen können
- Codebücher sind für jede empirische Forschung wichtig, besonders aber in der Inhaltsanalyse
  - z.B. für die Annotation von Tweets als “trolling” oder “nicht trolling”

# Codebücher

- Ein Codebuch zu erarbeiten, ist ein wesentlicher Schritt vor der konkreten Inhaltsanalyse
- Dient dazu, Merkmale aus Texteinheiten (oder sonstigen Medien) zu identifizieren und Kategorien zuzuordnen
- Enthält konkrete Anweisungen für den Codiervorgang, zum Beispiel wichtige Informationen zum Verständnis von Kategorien und deren Ausprägungen
- Beispiele:
  - Zu [Fake News](#)
  - Zu [Corona Sentiment auf Twitter](#)

# Codebücher

- Generell gibt es keine festen Regeln oder Vorgaben für die Gestaltung eines Codebuchs, es lassen sich aber Muster erkennen, nach dem ein Codebuch üblicherweise aufgebaut ist
  - Learning by doing
- In den letzten Jahren haben immer mehr Nicht-Forscher:innen, z.B. in der Industrie, in großem Stil Daten gesammelt, annotiert und ausgewertet
- Mit Hilfe dieser manuell erstellten Daten („Trainingsdaten“) KI-Modelle gelernt und angewendet –oft mit fragwürdigen Konsequenzen

# Codebuch zu Trolling

- Codieranleitung:
  - Trolling: “comments that ‘purposely or deliberately commented to get a rise out of the other people on the subreddits, whether it was successful or not”
  - Successful trolling: troll comments that received replies
  - Unsucessfultrolling:troll comments that received no replies

—> Ist das ein gutes Codebuch? Warum/warum nicht? Inwiefern ja/Inwiefern nein?

Was würdet ihr (anders) machen?

**Vielen Dank für die Aufmerksamkeit!**  
**Gibt es Fragen?**