computational social media

lecture 1: introduction

daniel gatica-perez

21.02.2020

this course

2006

2010
http://content.time.com/time/specials/packages/article/0,28804,2036683_2037183_2037185,00.html
today

course goals, syllabus, evaluation, logistics

setting the stage: who studies social media?

how did we get here?  
the triple revolution and networked individualism

goals of the course

present a human-centered view of social media

provide a coherent historical & conceptual background for multidisciplinary research on social media: media studies, HCI, machine learning, multimedia, network science

present fundamentals to understand motivations & analyze phenomena in Facebook, Twitter, YouTube, Instagram (and Flickr, Foursquare)

critique seminal research

develop hands-on experience with homeworks & student projects
**definition of social media**

“a group of internet-based applications that build on the **ideological and technological** foundations of Web 2.0, and that allow the creation and exchange of **user-generated content**”


**types of social media (in theory)**

**social network sites** *(SNSs)*:
“promote interpersonal contact; forge personal, professional, or geographic connections”
Facebook, Twitter, LinkedIn, Google+, Foursquare

**user-generated content** *(UGC)*:
“support creativity and foreground cultural activity; promote exchange of amateur or professional content”
YouTube, Instagram, Flickr, Pinterest, Vine, Wikipedia

**trading and marketing sites** *(TMSs)*:
“support exchange and sales of products”
Amazon, eBay, GroupOn, Craigslist, Etsy

**play and gaming sites** *(PGSs)*:
“social games”
FarmVille, Angry Birds

**types of social media (in practice)**

- social network site
- user-generated content
- marketing site
- gaming site

Van Dijck: “the entire ecosystem of interconnected platforms and applications has been in flux and will remain volatile”

**who studies social media?**

- computer science & engineering
- psychology
- sociology
- communication
- economics
- law
- management
- marketing
computational social media

Use social media data for CS problems
- natural language processing
- text mining
- computer vision and multimedia
- network science, graph theory
- machine learning
- human-computer interaction

Automate tasks for social media analysis
- sentiment analysis
- user modeling
- social network analysis
- social search & recommendation
- misinformation detection

Systems: volume, variety, velocity, complexity
- storage
- management
- sharing
- streaming

computational social media

Use social media as traces to study behavior in everyday life
- Friendship
- Communities
- Influence
- Mobility

Study phenomena within social media sites and communities

Computational social science

psychology
sociology
communication
economics
law
political science
digital humanities
Social media as a lens of society

"new opportunities to study human behavior that previously had to rely in behaviors difficult to assess (like making friends and chatting)"

"measurable behavioral traces with levels of ecological validity that are hard to match in most common research settings"

"the popularity of [social media] makes it a topic worthy of study in its own right"

"in addition to reflecting existing social processes, they are spawning new ones by changing the way people relate to one another"

"the rise of [social media] brings both new benefits and dangers to society"

computational social media

Effects of social media in other media production and consumption
- News
- Citizen journalism
- Entertainment

Marketing
- Attention
- Virality
- Branding

Corporate communication
- Community management
- Reputation management

example: topics in AAAI ICWSM
Int Conf. on Web and Social Media: http://www.icwsm.org/2018/index.php

Studies of digital humanities (culture, history, arts) using social media
Psychological, personality-based and ethnographic studies of social media
Analysis of the relationship between social media and mainstream media
Qualitative and quantitative studies of social media
Centrality/influence of social media publications and authors
Ranking/relevance of social media content and users
Credibility of online content
Social network analysis; communities identification; expertise discovery
Trust; reputation; recommendation systems
Human computer interaction; social media tools; navigation and visualization
Subjectivity in textual data; sentiment analysis; polarity/opinion extraction
Text categorization; topic recognition; demographic/gender/age identification
Trend identification and tracking; time series forecasting
Measuring predictability of real world phenomena spanning politics, finance, health
New social media applications; interfaces; interaction techniques
Engagement, motivations, incentives, and gamification
Social innovation and effecting change through social media
Social media usage on mobile devices; location, human mobility, and behavior
Organizational and group behavior mediated by social media
Interpersonal communication mediated by social media
course syllabus

**Friending.** A human-centered review of Facebook research. Users, communities, and networks. The real-name web

**Tweeting.** From random chatter to worldwide pulse. Followers, hashtags, topics, events. Analyzing phenomena on Twitter

**Shooting.** Photo sharing. Flickr, Instagram, Snapchat. Media, user, community analysis enabled by photo sharing

**Watching.** YouTube as a media phenomenon. Video blogging. Video analysis techniques

**Moving.** Location-based social networks. Large-scale phenomena revealed by mobility data

**Crowdsourcing.** Crowdsourcing and social participation. Uses in social media research. Crowdsourcing models

**Burning out.** Effects on society, privacy, fairness

course activities

**lectures**
- instructor: Daniel Gatica-Perez
- teaching assistants: Trung Phan
- session AM: lectures
- session PM: papers, assignments, projects

**papers**
- student presentations of papers that complement lectures
- discussion sessions: analyze & critique papers
- 3 roles: presenter, discussant, scribe
- number of papers depends on number of students

**homeworks**
- qualitative: reflection about social media phenomena
- quantitative: analyze social media data

**project**
- to be developed throughout the semester
- work in small groups
- presentation and report at the end of the semester
course evaluation

homeworks (45%)

paper presentation & participation in discussion (15%)

project (40%)

communication

office hours
+ fridays 14:30-15:30 (by appointment) INN 138
+ email: daniel.gatica-perez@epfl.ch

course website
www.idiap.ch/~gatica/teaching-csm/computational-social-media.html

lecture slides
+ on website, the evening before lecture

paper and homework assignments
+ announced in class and online

project
+ you will decide specific topic early on
+ we will monitor progress: short presentation around mid-semester
+ presentation & report at end of semester
a brief history of social media

1. what got us here?
networked individualism

Sociologist, NetLab, U. Toronto
Founder, Int. Network for Social Network Analysis (1977)

Director, Pew Research Center's Internet & American Life Project
Non-profit, "fact tank" that studies the social impact of the internet

MIT Press, 2012

Networked Individualism
The move to looser, far-flung networks

social life arranged around the individual rather than social units (e.g. family)
people operate more as networked individuals, less as members of tight groups
each person operates their own network
everyday life: online, mobile, social opportunities & challenges

slide by Lee Rainie: http://www.pewinternet.org/2013/11/07/networked-worlds-networked-enterprises/
Personal networks are:

+ more fragmented
+ differently composed
+ more specialized
+ more ephemeral
+ performing many functions: helpers, connectors, evaluators, audience
+ … and just as important as in the past

adapted from slide by Lee Rainie: http://www.pewinternet.org/2013/11/07/networked-worlds-networked-enterprises/

the triple revolution

social network revolution

internet revolution

mobile revolution

credit: Occupy Wall Street, http://www.flickr.com/photos/shankbone/6183443813 (cc)
I. the social network revolution

- widespread connectivity
- weaker group boundaries
- increased personal autonomy

credit: http://www.flickr.com/photos/hanspoldoja/5001818922 (cc)
1. widespread connectivity

1.1. “automobiles and airplanes have made travel wider-ranging and broadly affordable, helping spread social networks worldwide”

![Graph showing percentage of U.S. households by number of vehicles.](image1)


1. widespread connectivity

![Graph showing per capita airline boardings in the United States.](image2)

1. widespread connectivity

1.2. “growth of computing and communications has made communicating more powerful and more personal”

Figure 2.3
U.S. billed international calls.
Source: U.S. Federal Communications Commission.


1. widespread connectivity

Figure 2.4
Number of landlines and mobile phones in the United States (per 100 inhabitants).
Source: U.S. Federal Communications Commission.

1. widespread connectivity

1.3. “general outbreak of peace and spread of trade have driven commercial and social connectedness”

![Graph showing U.S. imports and exports as percentage of GDP. Source: The World Bank Group, the Organisation for Economic Co-operation and Development, and the U.S. Department of Commerce Bureau of Economic Analysis.]


2. weaker group boundaries

2.1. “family composition, roles, and responsibilities transformed households from groups to networks”

![Graph showing distribution of households in the United States as percentage. Source: U.S. Bureau of Labor Statistics.]

2. weaker group boundaries

2.2. “structured voluntary organizations are replaced by more open and informal networks of civic and religious practice”

Table 2.1

<table>
<thead>
<tr>
<th>Religions</th>
<th>Childhood</th>
<th>Current</th>
<th>Net</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baptist</td>
<td>20.9</td>
<td>17.2</td>
<td>3.7</td>
</tr>
<tr>
<td>Methodist</td>
<td>8.3</td>
<td>6.3</td>
<td>2.1</td>
</tr>
<tr>
<td>Nondenominational</td>
<td>1.5</td>
<td>4.5</td>
<td>3.0</td>
</tr>
<tr>
<td>Lutheran</td>
<td>5.5</td>
<td>4.6</td>
<td>0.9</td>
</tr>
<tr>
<td>Presbyterian</td>
<td>3.6</td>
<td>2.7</td>
<td>0.9</td>
</tr>
<tr>
<td>Pentecostal</td>
<td>3.0</td>
<td>4.4</td>
<td>1.4</td>
</tr>
<tr>
<td>Anglican/Episcopal</td>
<td>1.8</td>
<td>3.5</td>
<td>1.7</td>
</tr>
<tr>
<td>Catholic</td>
<td>31.4</td>
<td>23.9</td>
<td>7.5</td>
</tr>
<tr>
<td>Mormon</td>
<td>1.8</td>
<td>1.7</td>
<td>0.1</td>
</tr>
<tr>
<td>Jehovah’s Witness</td>
<td>0.4</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>Jewish</td>
<td>1.5</td>
<td>1.7</td>
<td>0.2</td>
</tr>
<tr>
<td>Muslim</td>
<td>0.3</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>Hindu</td>
<td>0.4</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>Other faiths</td>
<td>1.3</td>
<td>1.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Unrelated atheist, agnostic, &quot;nothing in particular&quot;</td>
<td>7.1</td>
<td>18.1</td>
<td>8.8</td>
</tr>
</tbody>
</table>


2.3. “common culture spread through few mass media firms has shifted to fragmented culture spread through more channels & hardware”

Figure 2.10
Mean number of televisions per U.S. household.
Source: U.S. Census Bureau.

3. increased personal autonomy

3.1. “work became flexible in developed world, from pushing atoms in manufacturing to pushing bits in white-collar work”


3. increased personal autonomy

3.2. “american society has become less bounded by ethnicity, gender, religion, sexual orientation”

II. the internet revolution

**trends enabling the internet revolution**

1. “government and scientists were first users; light regulatory role once popular

2. tech improved dramatically: computing power grew, prices dropped, bandwidth increased, storage improved

3. internet remained an interconnected network

4. applications led people to embrace it”

credit: http://www.flickr.com/photos/teflon/3190769121 (cc)
computers support networked individualism

**affordances:** “possible actions a person can perform on an object” (Don Norman)

design choices in hardware/software fostered networked individualism

1. personal
2. connected
3. humanized
4. private
5. decentralized
6. open to choice
7. asynchronous

credit: http://www.flickr.com/photos/pocait/2634190989 (cc)

---

**how americans use the internet**

Digital Divide? Where is the other 11%?

http://www.pewinternet.org/fact-sheet/internet-broadband/
how americans use the internet

% of U.S. adults who use the internet, by education level

- Less than high school graduate
- High school graduate
- Some college
- College graduate

credit: joan isaac biel

net culture: creating & sharing material online

“internet users…

- blog, upload images/videos, tweet
- belong to online support groups
- critique, rank, and rate books, movies, people
- advocate for political and social cases
- mash up existing media into parodies
- chronicle their travels
- give tips about their hobbies and passions”

credit: joan isaac biel
III. the mobile revolution

mobile phone trends

Cellphone Ownership, 2004-2013
Percentage of American adults who own a cell phone

Source: Pew Research Center’s Internet & American Life Project, April 17-May 19, 2013 Tracking Survey. Interviews were conducted in English and Spanish and on landline and cell phones. Margin of error is +/-2.3 percentage points based on all adults (n=2,232).

PEW RESEARCH CENTER

http://pewinternet.org/Infographics/2013/Cell-Phone-Ownership.aspx
1-in-10 adults are smartphone-only internet users: they do not use home broadband service

% of U.S. adults who do not use broadband at home but own smartphones, by age

- 2013:
  - 18-29: 17%
  - 30-49: 13%
  - 50-64: 11%
  - 65+: 7%

Source: Surveys conducted 2013-2016. Data for each year based on a pooled analysis of all surveys containing broadband and smartphone questions fielded during that year.

http://www.pewinternet.org/fact-sheet/internet-broadband/

The income divide

% of U.S. adults who do not use broadband at home but own smartphones, by income

- 2013:
  - Less than $30,000: 21%
  - $30,000-$49,999: 12%
  - $50,000-$74,999: 10%
  - $75,000+: 5%

Source: Surveys conducted 2013-2016. Data for each year based on a pooled analysis of all surveys containing broadband and smartphone questions fielded during that year.

http://www.pewinternet.org/fact-sheet/internet-broadband/
2.

a social media timeline
participation in social media

% of U.S. adults who use at least one social media site

69% as of January 2018

http://www.pewinternet.org/fact-sheet/social-media/

participation in social media (2)

% of U.S. adults who use at least one social media site, by age

88%
78%
64%
37%

http://www.pewinternet.org/fact-sheet/social-media/
participation in social media (3)

http://www.pewinternet.org/fact-sheet/social-media/

participation in social media (4)

Among the users of each social media site, the % who use that site with the following frequencies

http://www.pewinternet.org/fact-sheet/social-media/
the social media paradox

Mark Zuckerberg:

“We’re not thinking about ourselves as a community — we’re not trying to build a community — we’re not trying to make new connections. [...] What we’re trying to do is just make it really efficient for people to communicate, get information and share information. We always try to emphasize the utility component.” Jul. 2007

“Maybe electricity was cool when it first came out, but pretty quickly people stopped talking about it because it’s not the new thing, the real question you want to track at that point is: are fewer people turning on their lights because it’s less cool?” [...] Our society needs a new digital social fabric...We can help build it.” Sep. 2013

J. Constine, Facebook Doesn’t Want To Be Cool, It Wants To Be Electricity, Facebook Doesn’t Want To Be Cool, It Wants To Be Electricity, Cruncher, Sep 18 2013, http://techcrunch.com/2013/09/18/facebook-doesnt-want-to-be-cool/
Jack Dorsey:

"I think Twitter’s a success for us when people stop talking about it, when [...] people **just use it as a utility, use it like electricity** [...] It fades into the background, something that’s just a part of communication…. That’s where we want to be." Jun. 2009

---

**social media as a utility, really?**

"Social media as **neutral** platforms upon which users freely interact, much like the Web itself – an infrastructure that transports streams, regardless of who its users are and indifferent to the contents they exchange"

"…yet social media do not just transport streams of live tweets; neither the platforms are simple carriers of information. Streams of data **are engineered to promote certain uses and users over others**, challenged by the pressure to **make content streams profitable**."

J. van Dijck, The culture of connectivity, Oxford University Press, 2013
the social media paradox

enable **connectedness** while engineering **connectivity**

propagate **neutrality** while securing **profitability**

daniel.gatica-perez@epfl.ch
references and further reading


J. van Dijck, The Culture of Connectivity, A Critical History of Social Media, Chapter 1-2, Oxford University Press, 2013