Using Large Social Data for COVID-19

Johannes Eichstaedt
(Stanford, Psychology/A.I.)

Sharath Guntuku
Garrick Sherman
(UPenn)

Andy Schwartz
Mohammadzaman Zamani
(Stony Brook)

Eichstaedt@stanford.edu
@jeichstaedt
Using Large Social Data for COVID

[0] Psychological impact of COVID

[1] Monitoring COVID prevalence
  • Symptom tracking: (e.g., Google Flu trends)
  • Finding new symptoms

[2] Adherence
  • Adherence tracking with shelter-in-place orders (location data)

[3] COVID-related discourse
  • Prevalence of public health messages
  • Tracking of Misinformation
The Psychological Impact of COVID-19

- COVID-19 outbreak and social distancing
- Uncertainty / anxiety
- Unemployment
- Loneliness
- Reduced subjective well-being & mental illness

Johannes Eichstaedt, 4/16/2020. Stanford University. eichstaedt@stanford.edu
The Psychological Impact of COVID-19

COVID-19 outbreak and social distancing → Uncertainty / anxiety → Unemployment → Reduced subjective well-being & mental illness → Loneliness
Unemployment & Life Satisfaction


eichstaedt@stanford.edu
### Unemployment & Life Satisfaction

<table>
<thead>
<tr>
<th>Rank</th>
<th>Country</th>
<th>Happiness Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.</td>
<td>United States</td>
<td>6.940</td>
</tr>
<tr>
<td>19.</td>
<td>Czech Republic</td>
<td>6.911</td>
</tr>
<tr>
<td>20.</td>
<td>Belgium</td>
<td>6.864</td>
</tr>
<tr>
<td>21.</td>
<td>United Arab Emirates</td>
<td>6.791</td>
</tr>
<tr>
<td>22.</td>
<td>Malta</td>
<td>6.773</td>
</tr>
<tr>
<td>23.</td>
<td>France</td>
<td>6.664</td>
</tr>
<tr>
<td>24.</td>
<td>Mexico</td>
<td>6.465</td>
</tr>
<tr>
<td>25.</td>
<td>Taiwan Province of China</td>
<td>6.455</td>
</tr>
<tr>
<td>26.</td>
<td>Uruguay</td>
<td>6.440</td>
</tr>
<tr>
<td>27.</td>
<td>Saudi Arabia</td>
<td>6.406</td>
</tr>
<tr>
<td>28.</td>
<td>Spain</td>
<td>6.401</td>
</tr>
<tr>
<td>29.</td>
<td>Guatemala</td>
<td>6.399</td>
</tr>
<tr>
<td>30.</td>
<td>Italy</td>
<td>6.387</td>
</tr>
<tr>
<td>31.</td>
<td>Singapore</td>
<td>6.377</td>
</tr>
<tr>
<td>32.</td>
<td>Brazil</td>
<td>6.376</td>
</tr>
<tr>
<td>33.</td>
<td>Slovenia</td>
<td>6.363</td>
</tr>
<tr>
<td>34.</td>
<td>El Salvador</td>
<td>6.348</td>
</tr>
<tr>
<td>35.</td>
<td>Kosovo</td>
<td>6.325</td>
</tr>
<tr>
<td>36.</td>
<td>Panama</td>
<td>6.305</td>
</tr>
<tr>
<td>37.</td>
<td>Slovakia</td>
<td>6.281</td>
</tr>
<tr>
<td>38.</td>
<td>Uzbekistan</td>
<td>6.258</td>
</tr>
<tr>
<td>39.</td>
<td>Chile</td>
<td>6.228</td>
</tr>
</tbody>
</table>

-0.7

That is big!

World Happiness Report, 2020

eichstaedt@stanford.edu
The Psychological Impact of COVID-19

COVID-19 outbreak and social distancing

Uncertainty / anxiety

Unemployment

Loneliness

Reduced subjective well-being & mental illness

eichstaedt@stanford.edu
The Psychological Impact of COVID-19

COVID-19 outbreak and social distancing → Uncertainty / anxiety → Unemployment → Loneliness → Reduced subjective well-being & mental illness

eichstaedt@stanford.edu
Loneliness & Life Satisfaction

Impact: \(-0.5\)

Comparable to the impact of suffering from **headaches** or being **chronically** ill.

That is also big!

\(N=700,000\)

Kahneman & Deaton (2010) *PNAS*
The Psychological Impact of COVID-19

COVID-19 outbreak and social distancing → Uncertainty / anxiety

Uncertainty / anxiety → Unemployment

Unemployment → Reduced subjective well-being & mental illness

Unemployment → Loneliness

Loneliness → Reduced subjective well-being & mental illness

Digital social transition

eichstaedt@stanford.edu
We need to measure the impact of COVID on well-being and mental health.
How?
National Measurement with surveys

Expected Outcomes of Social Distancing Practices

- Already experiencing
- A few more weeks
- A few more months
- As long as is necessary

How long can you follow social distancing practices and business/school closures before ...

- ... your physical health suffers?
  - 6%
  - 12%
  - 15%
  - 68%
- ... experiencing significant financial hardship?
  - 9%
  - 15%
  - 22%
  - 54%
- ... your emotional or mental health suffers?
  - 15%
  - 18%
  - 19%
  - 48%

GALLUP PANEL, APRIL 6-12, 2020

eichstaedt@stanford.edu
Daily Emotional Experiences of U.S. Adults, Tended (January 2018–April 2020)

Did you experience each of the following a lot of the day yesterday?

- % Enjoyment
- % Worry
- % Stress

Data from January 2018 to August 2019 are based on the Gallup National Health and Well-Being Index; data from March 6-20 and March 21-April 5, 2020, are from the Gallup Panel.

GALLUP
1.53 billion geotagged tweets 2009 to 2015 + text analysis A.I.


Community measurement with Social Media

Life Satisfaction
Predicting **Life Satisfaction**

Language Prediction Model

- Life satisfaction
- Income

Eichstaedt et al., in preparation
Accuracy: Predicting County Life Satisfaction

Prediction Accuracy (Pearson r)

- Log. Income + Twitter
- Twitter alone
- Log. Income

Eichstaedt et al., in preparation
Using Large Social Data for COVID

[0] Psychological impact of COVID

[1] Monitoring COVID prevalence

[2] Adherence

[3] COVID-related discourse
[1] COVID prevalence tracking

Main difficulty: no specificity of COVID symptoms

<table>
<thead>
<tr>
<th>COVID-19 symptoms</th>
<th>Cold symptoms</th>
<th>Flu symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fever</td>
<td>Sore throat</td>
<td>Fever or chills</td>
</tr>
<tr>
<td>Cough</td>
<td>Runny nose</td>
<td>Cough</td>
</tr>
<tr>
<td>Shortness of breath</td>
<td>Cough</td>
<td>Sore throat</td>
</tr>
<tr>
<td></td>
<td>Sneezing</td>
<td>Runny or stuffy nose</td>
</tr>
<tr>
<td></td>
<td>Headaches</td>
<td>Body aches</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Headaches</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tiredness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vomiting and diarrhea</td>
</tr>
</tbody>
</table>

Learn more about symptoms from the CDC.

Dependable COVID baselines through Google etc. are difficult.
[1] COVID prevalence tracking

Source: Google | By The New York Times

eichstaedt@stanford.edu
Using Large Social Data for COVID

[0] Psychological impact of COVID

[1] Monitoring COVID prevalence

[2] Adherence

[3] COVID-related discourse
We are seeing the biggest change in social norms / behaviors since World War 2.
When average distance traveled first fell below 2 miles

“Where America Didn’t Stay Home Even as the Virus Spread,” New York Times, 04/02/2020
+50% more movement than usual

San Francisco metro

-50% less movement than usual

-100%

+50% more movement than usual

25 largest metro areas

~50% less movement than usual

~100%

Americans' Readiness to Return to Normal Activities

Once government restrictions on social contact are lifted and businesses and schools start to reopen, how quickly do you think you would return to your normal activities, including interacting with people in public?

<table>
<thead>
<tr>
<th>Response</th>
<th>March 27-29 %</th>
<th>April 3-5 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediately</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td>Wait to see what happens with the coronavirus before resuming</td>
<td>69</td>
<td>71</td>
</tr>
<tr>
<td>Continue to limit your contact with other people and daily activities indefinitely</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

GALLUP PANEL, 2020


Travel fell dramatically in Seattle, from 3.8 miles... to an average of 61 feet.

But in Daytona Beach, Fla. travel only fell from 4.4 miles... to an average of 1.9 miles.

cellphone data from 15 million people

eichstaedt@stanford.edu
### U.S. Adults' Adoption of Social Distancing Practices

There are some things people may do because of their concern about the coronavirus. For each one of the following, please indicate if this is something you have done, are considering doing or have not considered.

<table>
<thead>
<tr>
<th></th>
<th>Avoided crowds</th>
<th>Avoided mass transit/air travel</th>
<th>Avoided small gatherings</th>
<th>Avoided public places</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Mar 27-29</td>
<td>--</td>
<td>89</td>
<td>83</td>
<td>78</td>
</tr>
<tr>
<td>Mar 23-26</td>
<td>--</td>
<td>90</td>
<td>78</td>
<td>77</td>
</tr>
<tr>
<td>Mar 20-22</td>
<td>92</td>
<td>87</td>
<td>68</td>
<td>72</td>
</tr>
<tr>
<td>Mar 16-19</td>
<td>79</td>
<td>75</td>
<td>46</td>
<td>54</td>
</tr>
<tr>
<td>Mar 13-15</td>
<td>59</td>
<td>55</td>
<td>23</td>
<td>30</td>
</tr>
</tbody>
</table>

Avoided going to events with large crowds, such as concerts, festivals or sporting events; Avoided traveling by airplane, bus, subway or train; Avoided small gatherings of people, such as with family or friends; Avoided going to public places, such as stores or restaurants

GALLUP PANEL, 2020

eichstaedt@stanford.edu
Using Large Social Data for COVID

[0] Psychological impact of COVID

[1] Monitoring COVID prevalence

[2] Adherence

[3] COVID-related discourse
2 million Tweets w/ COVID hashtags
Feb 26th to Mar 26th

N = 1,015 counties

eichstaedt@stanford.edu
Urban counties

Methods: N = 1,015 counties, 2 million Tweets (02-27 to 03-26, no retweets or replies). Correlation with % Pop in Urban Area, all p’s < .05, B.H.-controlled.

Don’t touch face

Wash Hands

Panic buying

Cancelled Events

Wash your hands

Work from home
Educated counties

“Government”

“Will be canceled”

“Testing”

Methods: N = 1,015 counties, 2 million Tweets (02-27 to 03-26, no retweets or replies).
Correlation with % Bachelor, all p’s < .05, B.H.-controlled.

eichstaedt@stanford.edu
Older counties

Methods: N = 1,015 counties, 2 million Tweets (02-27 to 03-26, no retweets or replies).
Correlation with % age < 18 or >65, all p’s < .05, B.H.-controlled.

Younger counties

Wash hands

Economic impact
Voted for Trump

“like the flu”

Methods: N = 1,015 counties, 2 million Tweets (02-27 to 03-26, no retweets or replies). Correlation with Trump 2016 Vote Share, all p’s < .05, B.H.-controlled.
Most negative sentiment :( 

“Scared”  

Economy  

Methods: 200k Tweets (02-27 to 03-26, no retweets or replies). Correlation with NRC-Sentiment, all p’s < .05, B.H.-controlled.


eichstaedt@stanford.edu
Takeaways:

We need to measure the well-being impact of COVID! (and think about scalable mental health care.)

Signs of COVID adjustment in urban & educated communities, not yet in rural areas.

We can use geotagged Twitter to monitor progress in adherence and behaviors!

COVID tracking purely through digital traces is difficult.