Data Science to Improve Population Health

Elaine Nsoesie, PhD
Assistant Professor of Global Health
Email: onelaine@bu.edu / Twitter: @ensoesie
Network Individual-based Model

Household
Network Individual-based Model

Household 1

Household 2

Household 3
Network Individual-based Model

- Home
- School
- Bus
- Work
- Lunch
- Work
- Carpool
- Home
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
-Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Daycare
- Car
- Shopping
- Car
- Dayca
Disease model

- Susceptible: Contact occurs with infectious person
- Exposed: Disease progresses in infected person
- Recovered: Person recovers after disease has completed its course
- Infectious
Day: 1

Number of components: 0
Sizes: []
Forecasting disease trends

Number or proportion of reports

Time in weeks
Digital Epidemiology: new data streams

"Got my flu shot this morning and now my throat is sore."

"Stomach flu & normal flu in the same month. I'm officially a germaphobe."

"Such an upset stomach today. I hope it's just a bug and not the Truvada."

"My weight: 170.1 lb. 10.1 lb to go. #raceweight @Withings scale auto-tweets my weight once a week http://withings.com"

Text
Images
Videos
Sounds
Location
Biological data
etc.

Slide by Marcel Salathé
Studies have explored a variety of health questions

How can we monitor the spread of influenza using Internet search data?

Can Instagram photos provide insights into changes in mental health?

What do postings on social media tell us about the stigma associated with women’s health issues?

How does misinformation about vaccines spread through social networks?

How can discussions on radio stations lead to early identification of a Cholera outbreak?

Is there seasonality in online searches for chicken pox and can we use that to predict outbreaks?

Can postings on discussion forums reveal new insights into adverse drug events?

Are there geographical and racial differences in hospital care experiences?

Do filtered images on Instagram distort the self-image of teenagers?

Can news reports mined from online sources help us detect Ebola outbreaks faster?
Fig 1. The CDC’s %ILI (Influenza like illnesses), the performance of the 5 available predictors, the baseline predictions, and the performance of the best ensemble method for last week’s predictions are displayed as a function of time (top). The errors associated with each weak predictor and the ensemble approach are shown (bottom).
Research and Implementation

How can we use data from the Internet and digital devices to improve individual and population health?

Process data and develop tools for health departments, global health organizations, and foundations for decision making.
Satellite Images
Monitoring Social Disruption in South America with Satellite Parking Data

Butler et al. IEEE. 2014
Monitoring Social Disruption in South America with Satellite Parking Data

(a) Mexico
(b) Chile
(c) Argentina
- Obesity is a complex health issue
- Multiple factors have been linked to obesity
- In the US, obesity affects about one-third of the adult population
Many environmental factors have been linked to obesity

But there are differences in measures and measurements making it difficult to compare findings across cities
Census Tract 47157009600

Images show green cover in neighborhood and walkways in field

Census Tract 47157000800

Industrial area, presence of vehicles and sparse greenery
Transfer Learning with VGG-CNN-F

Transfer of Knowledge

Predictive Model

Feature Maps for Domain B

Model trained for Domain A
Built environment information is extracted from satellite images. POI indicates point of interest.
Social Media
Social media provides a window into the built environment, interactions, access, likes/dislikes, lifestyle, behaviors, choices
<table>
<thead>
<tr>
<th>EXPLICIT DATA ABOUT USER</th>
<th>EXPLICIT DATA ABOUT TWEET</th>
<th>EXTRACTABLE DATA ABOUT USER</th>
<th>EXTRACTABLE DATA ABOUT TWEET</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twitter handle</td>
<td>Time/date of tweet</td>
<td>Disease state</td>
<td>Content</td>
</tr>
<tr>
<td>Language</td>
<td>Hashtag</td>
<td>Gender</td>
<td>Sentiment</td>
</tr>
<tr>
<td>Time zone</td>
<td>URL</td>
<td>Marital status</td>
<td>Image Analysis</td>
</tr>
<tr>
<td>Location</td>
<td>Geotag</td>
<td>Political party</td>
<td>Language</td>
</tr>
<tr>
<td>Date account created</td>
<td>App used to send tweet</td>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>User profile</td>
<td>Number of retweets</td>
<td>Occupation</td>
<td></td>
</tr>
<tr>
<td>User profile photo</td>
<td>Number of favorites</td>
<td>Interests</td>
<td></td>
</tr>
<tr>
<td>Number of followers</td>
<td>Total number of tweets</td>
<td>Religion</td>
<td></td>
</tr>
<tr>
<td>Number following</td>
<td>Linked images</td>
<td>Income</td>
<td></td>
</tr>
</tbody>
</table>
It’s complicated.
Social media data raises important ethical and regulatory questions.
The Impact Health of Misinformation on Public Health

- Fake audio warning against the dangers of the yellow fever vaccine contributed to vaccine hesitancy in Brazil
- Misinformation about the spread of 2014-2016 Ebola epidemic in West Africa
- Increased incidence of measles outbreaks in the US and Europe linked to misinformation on vaccines

Using Twitter Data to Monitor Reports of Foodborne Illness in US Cities

1. Reporting food poisoning on Twitter
2. HealthMap showing reports from St. Louis
3. Targeted restaurant inspections and surveillance
How do people talk about women’s health and infertility on social media? And why does it matter?
Using digital data for health applications in Africa requires a rethinking of who, what and how.
Health in African Cities

Dar es Salaam
Travel Advisor
About 55% of the world’s population lives in urban areas. This number is expected to increase to 68% by 2050.

United Nations
The field of urban health started around the turn of the 21\textsuperscript{st} century is focused on the study of how and why cities influence health.
The foods we have access to

The foods we eat

Images from unsplash.com
All neighborhoods are not created equal
Drivers of health in cities

The physical environment
The social environment
Access to health and social services
What types of information are Africans seeking online and how are trends/patterns evolving with changing disease patterns?
Nutritional Composition

- 34% bread, cereal, rice, pasta, potatoes
- 26% fruits, vegetables
- 22% desserts, candy, cakes, ice cream
- 20% dairy, milk, cheese
- 18% chicken, eggs, poultry
- 13% lamb, beef, pork
- 6% fish, seafood, shellfish
- 4% snacks, chips, pretzels
- 6% soda
- 4% juice
- 12% alcohol

Nutritional Composition

- 21% homemade/fresh
- 23% heavily processed
- 24% fried, heavy oil content, butter
- 23% heavy salt content

Food & Beverage on Instagram

Slide by Raina Merchant, MD
Call for collaborators
Contributors
Nina Cesare, Boston University
Pallavi Dwivedi, University of Maryland
Quynh Nguyen, University of Maryland
Adyasha Maharana, SCIOME
Bryan Lewis, University of Virginia
Boston University Spark!
John Brownstein & Jared Hawkins, Harvard Medical School
Olubusola Oladeji, Boston University
Kadija Ferryman, Data & Society
Raina Merchant, University of Pennsylvania

Thank you!

Contact
Email: onelaine@bu.edu
Twitter: @ensoesie