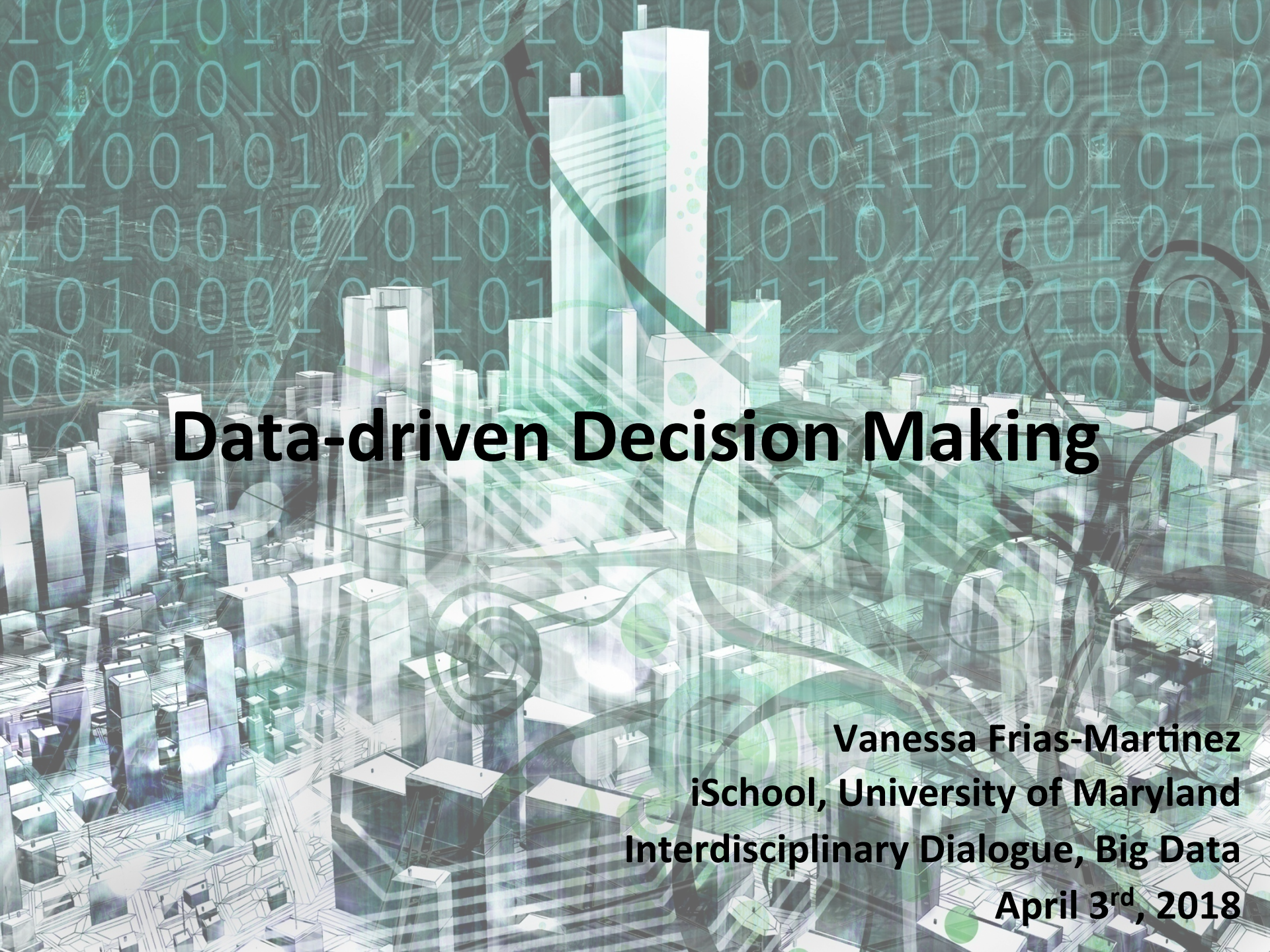
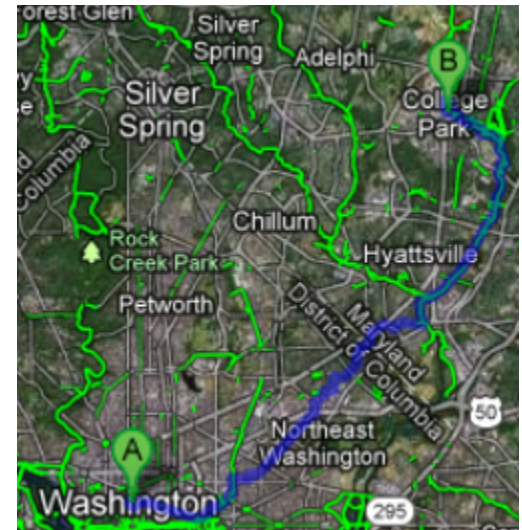
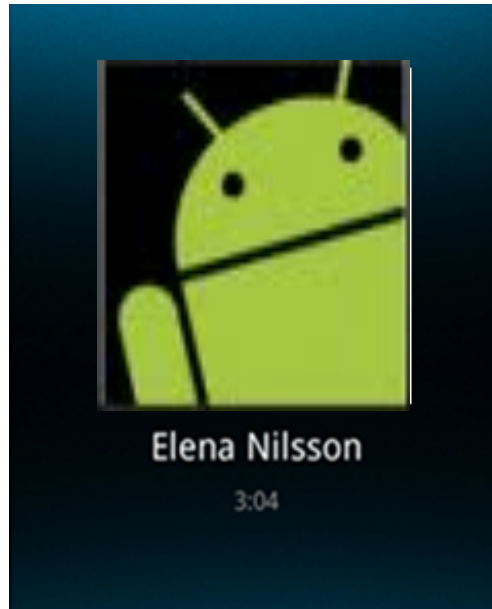


The background is a complex digital collage. It features a city skyline with various skyscrapers in shades of blue and white. Overlaid on this are numerous green binary digits (0s and 1s) of varying sizes. There are also green circular patterns, some resembling orbits or data paths, and a large, faint green 'X' shape. The overall aesthetic is high-tech and data-oriented.

Data-driven Decision Making

Vanessa Frias-Martinez
iSchool, University of Maryland
Interdisciplinary Dialogue, Big Data
April 3rd, 2018



Urban Computing Lab

We design, develop and evaluate
computational tools
that extract insights from large-scale data
to enhance decision-making processes

Cost
Frequency
New Insights

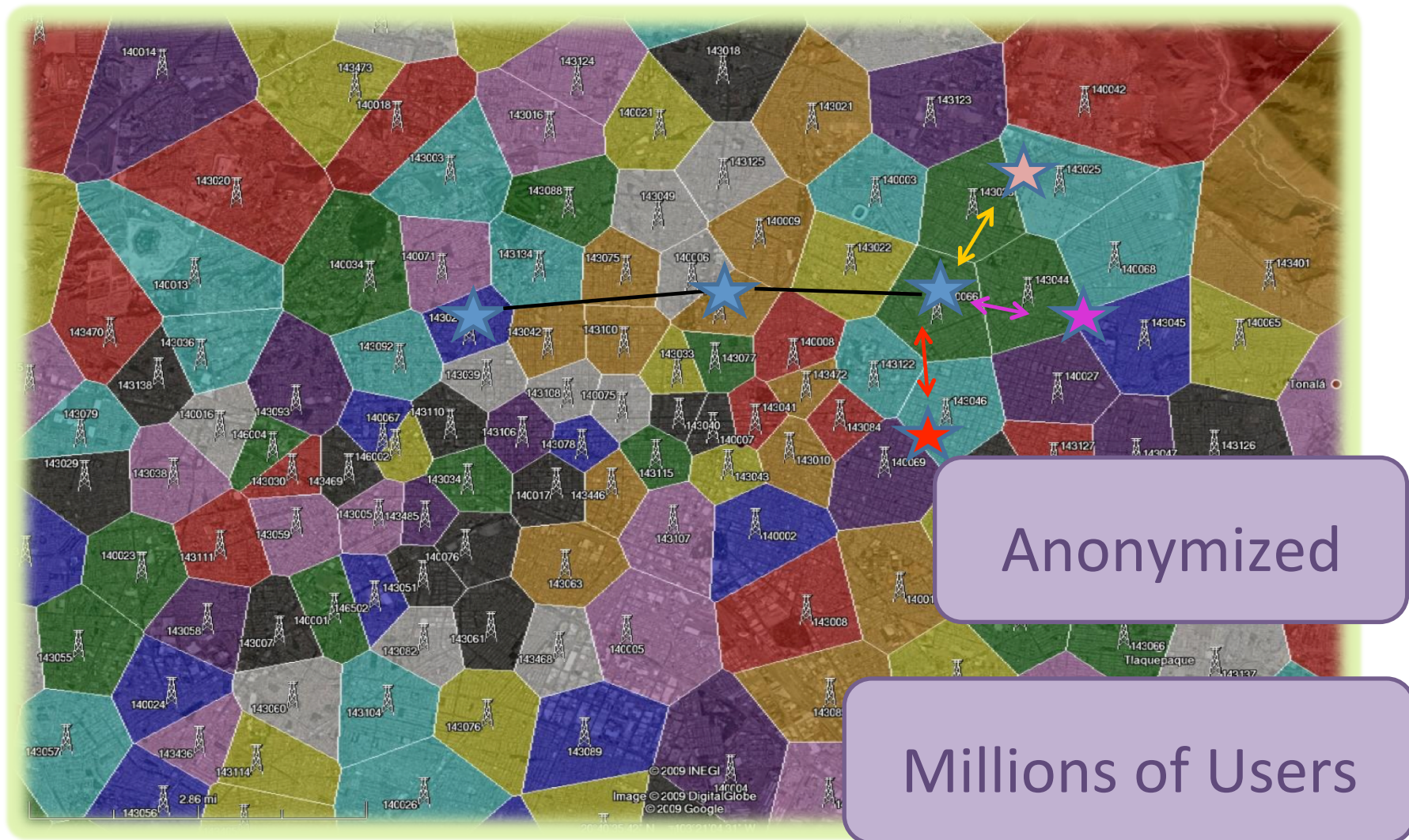
Societal Impact

- Urban Planning
- Economic development
- Emergency response
- Transportation
- Epidemiology

We collaborate with UN, WB and telecommunication companies interested in
“Data/AI for social good”

DATA

Call Detail Records



Modeling Human Behavior



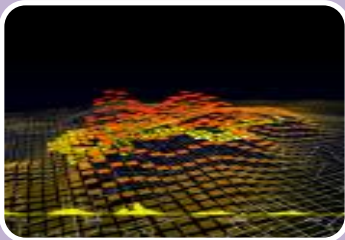
Consumption

- Number calls, duration, frequency, SMS/MMS/voice
- Expenses
- Handset Type and Features



Social

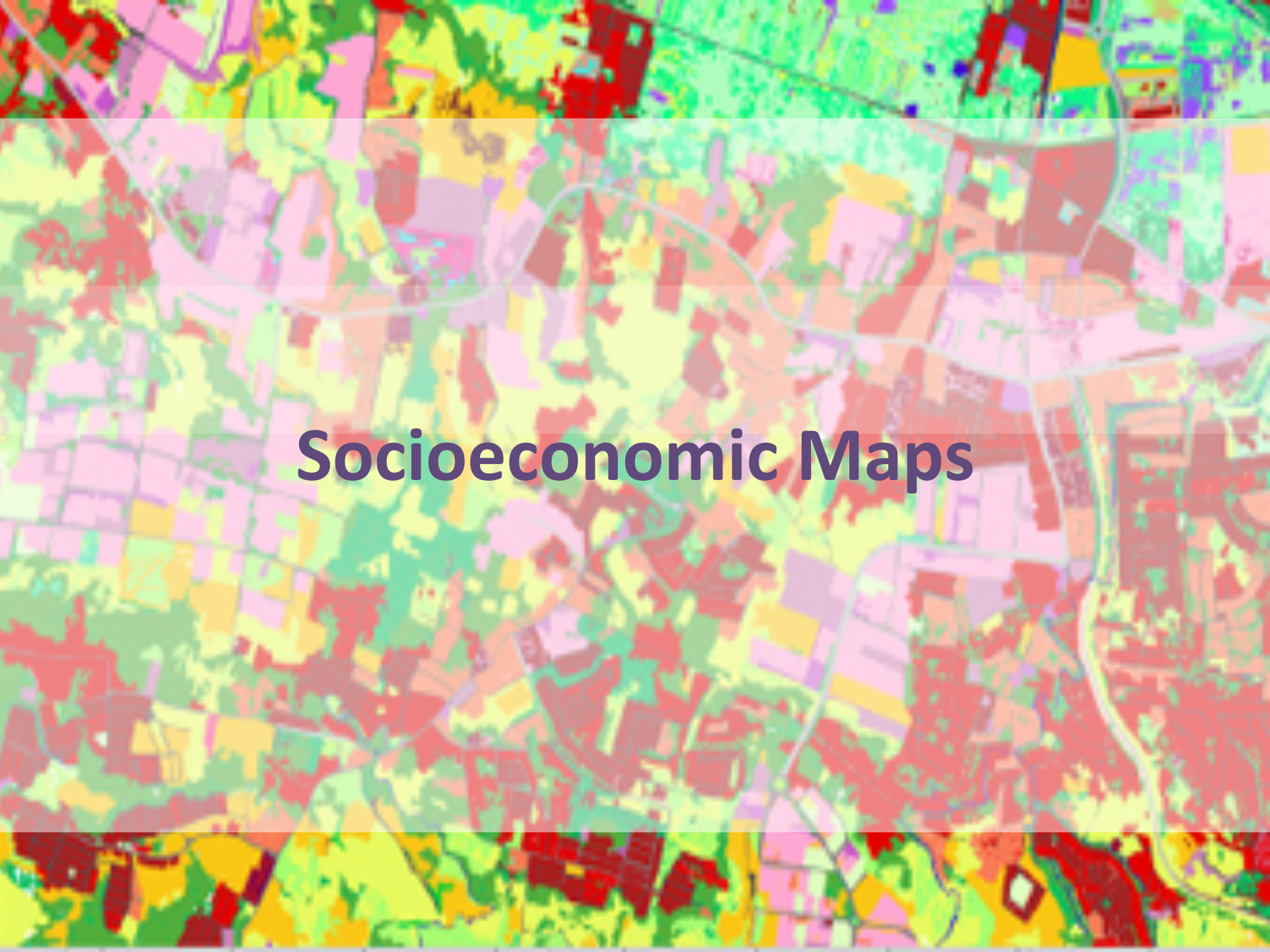
- Degree of the social network
- Strength of the contacts (Reciprocity & Frequency)
- Geography of the social contacts



Mobility

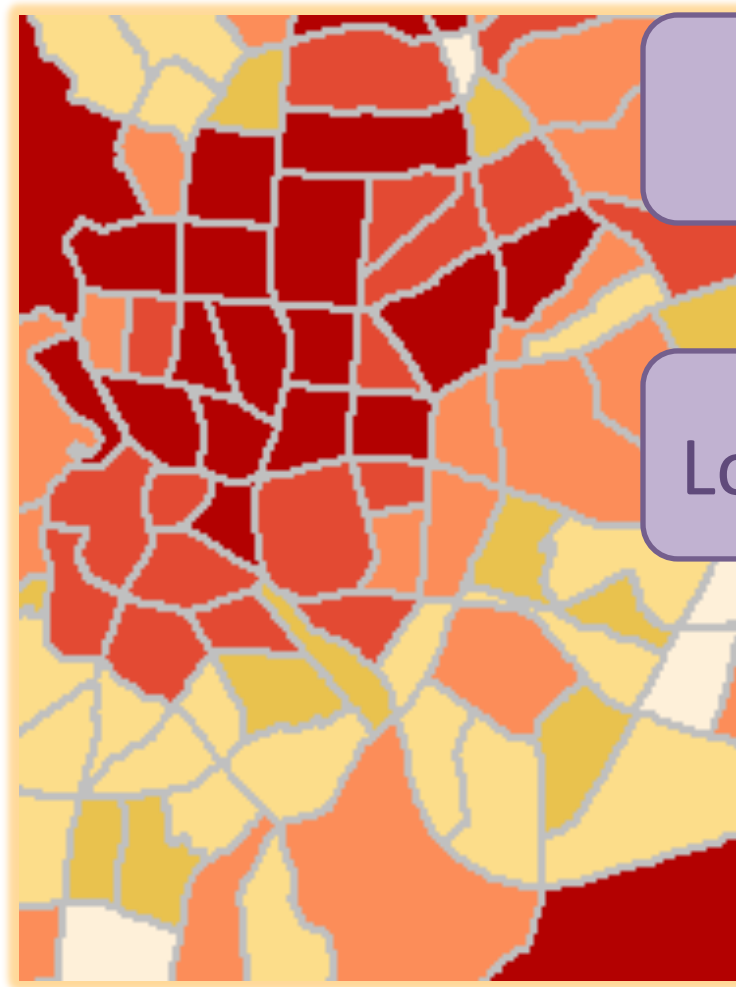
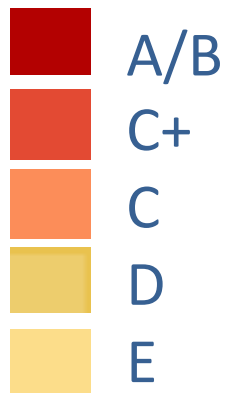
- Mobility Patterns (Entropy)
- Diameter of mobility
- Radius of gyration (Home/Work)

RESEARCH PROJECTS



Socioeconomic Maps

Important Data Comes at a Price

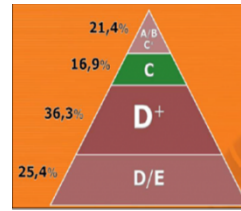
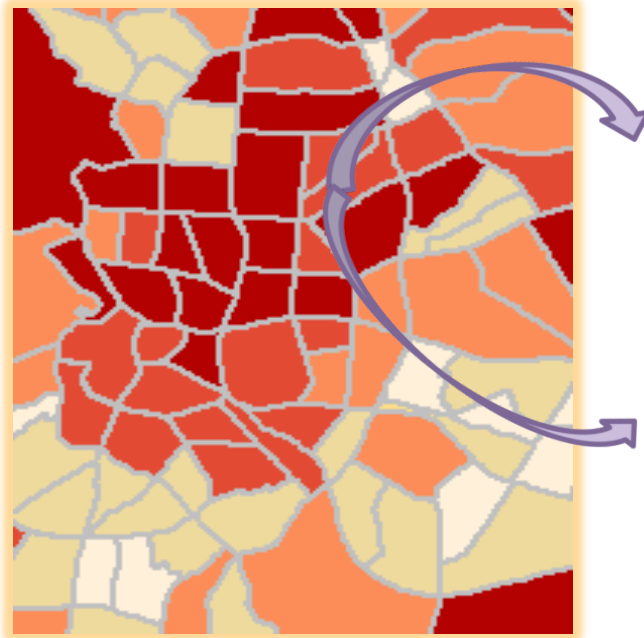


Expensive

Low Frequency

Can human behaviors extracted from Call
Detail Records be used to forecast
regional socioeconomic information?

Methodology



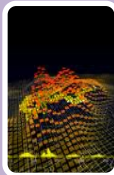
SEL (NSI)



Consumption



Social



Mobility



C
L
A
S
S
I
F
I
E
R

Proxies

- Higher SELs associated to longer commuting distances 
- Higher SELs associated to more reciprocal cell phone relationships 
- Lower SELs associated to larger volumes of voice vs SMS 

Results

- Accuracies of up to 82% per individual SEL
- Aggregated classification with privacy preserving data: 78% accuracy
- WB currently working in Guatemala to scale to multiple companies

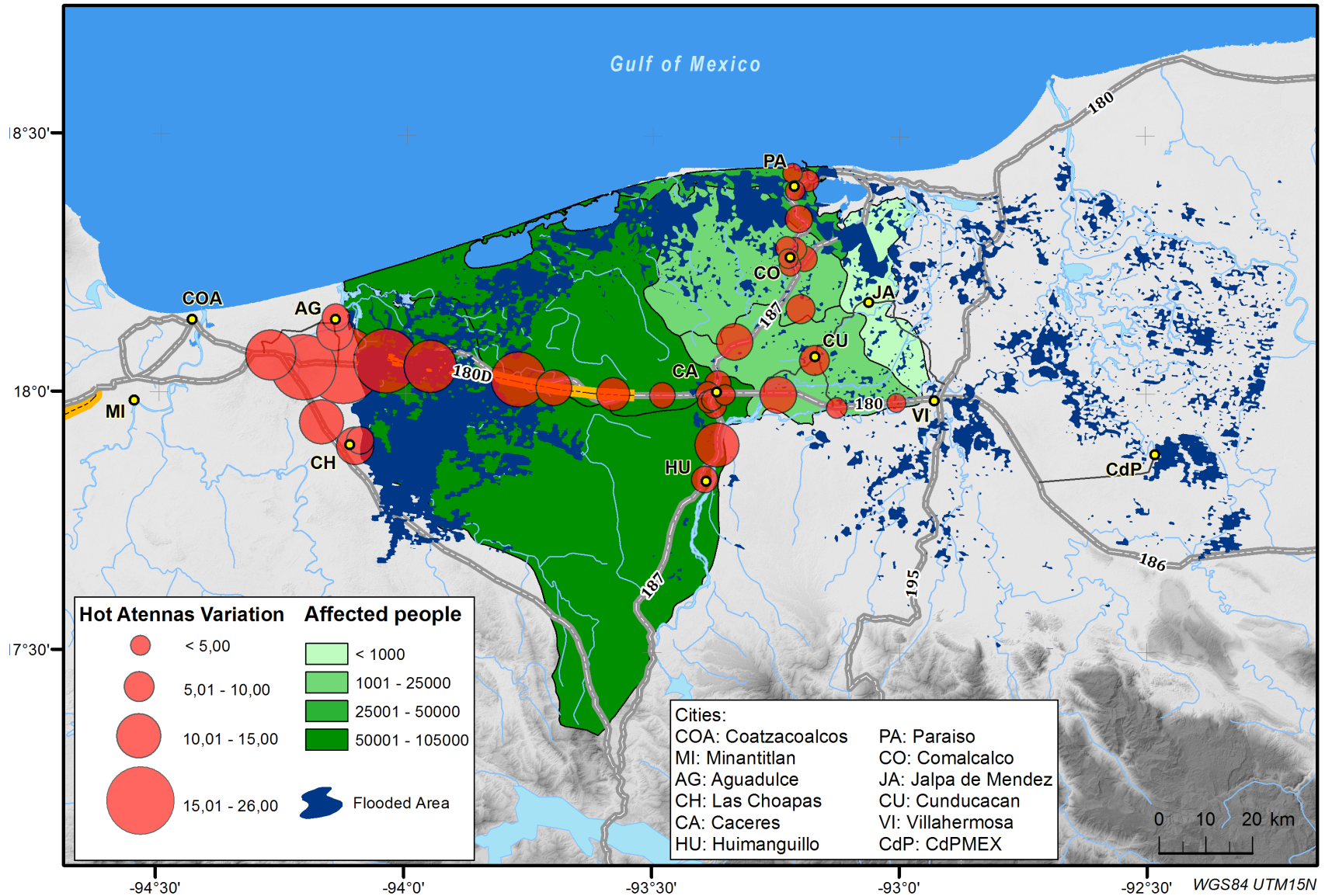
An aerial photograph showing a severe flood that has inundated a large area of land. A multi-lane road runs diagonally from the top center towards the bottom right. Several vehicles are visible on the road, including a red bus and a white van. The surrounding fields and forests are completely submerged in murky, brown floodwater. Only the tops of trees and some small patches of vegetation are visible above the water level. The overall scene depicts a major natural disaster, likely a river overflow or heavy rainfall event.

Natural Disasters

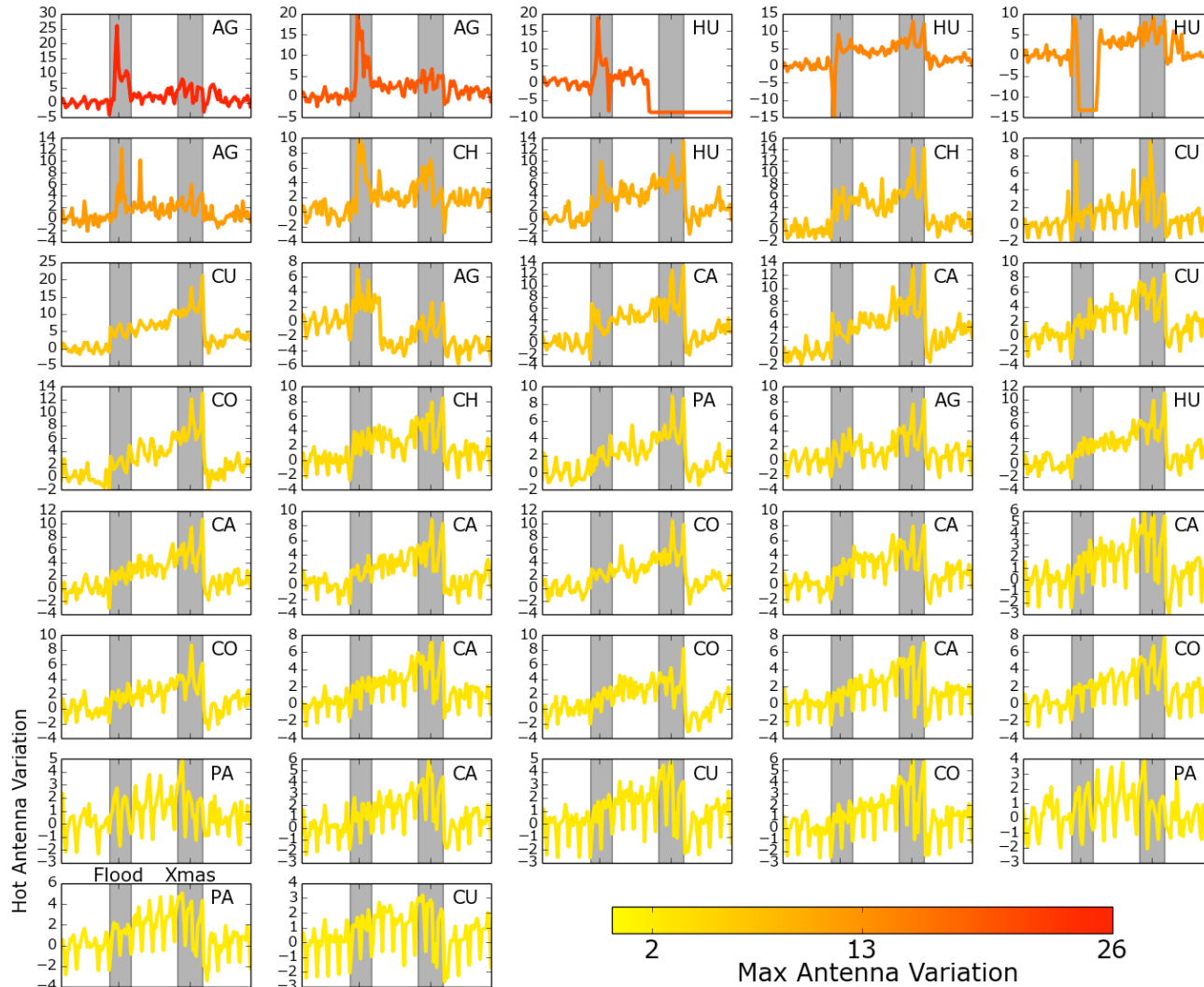
Decision Making for Emergency Preparedness and Response

- CDR behaviors to provide emergency responders with information beyond the one they can collect manually
- A couple of preliminary projects:
 1. Number of displaced individuals during disaster
 2. Mobility disruption and resilience to disasters

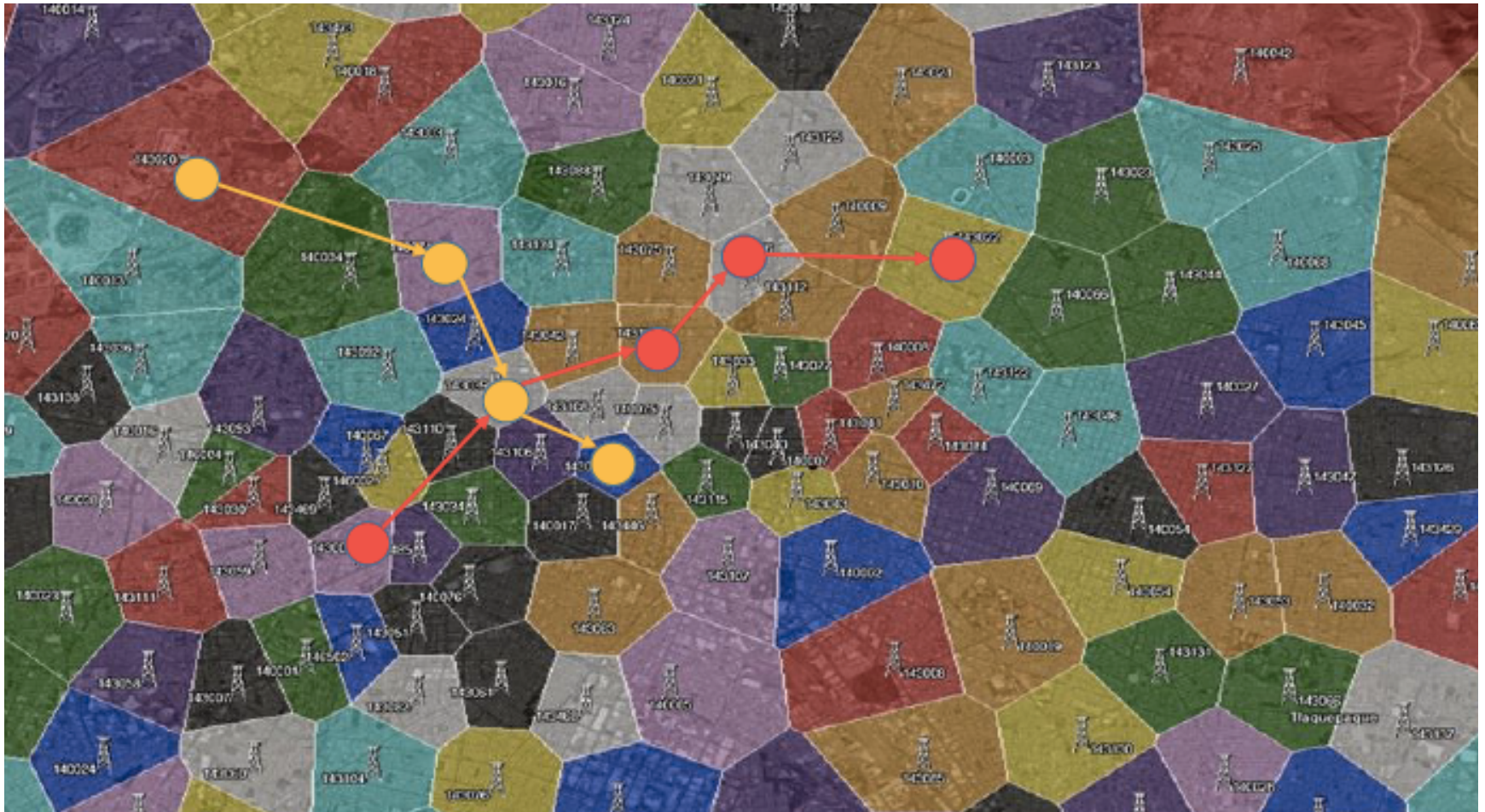
Displaced Individuals- Tabasco, Mexico



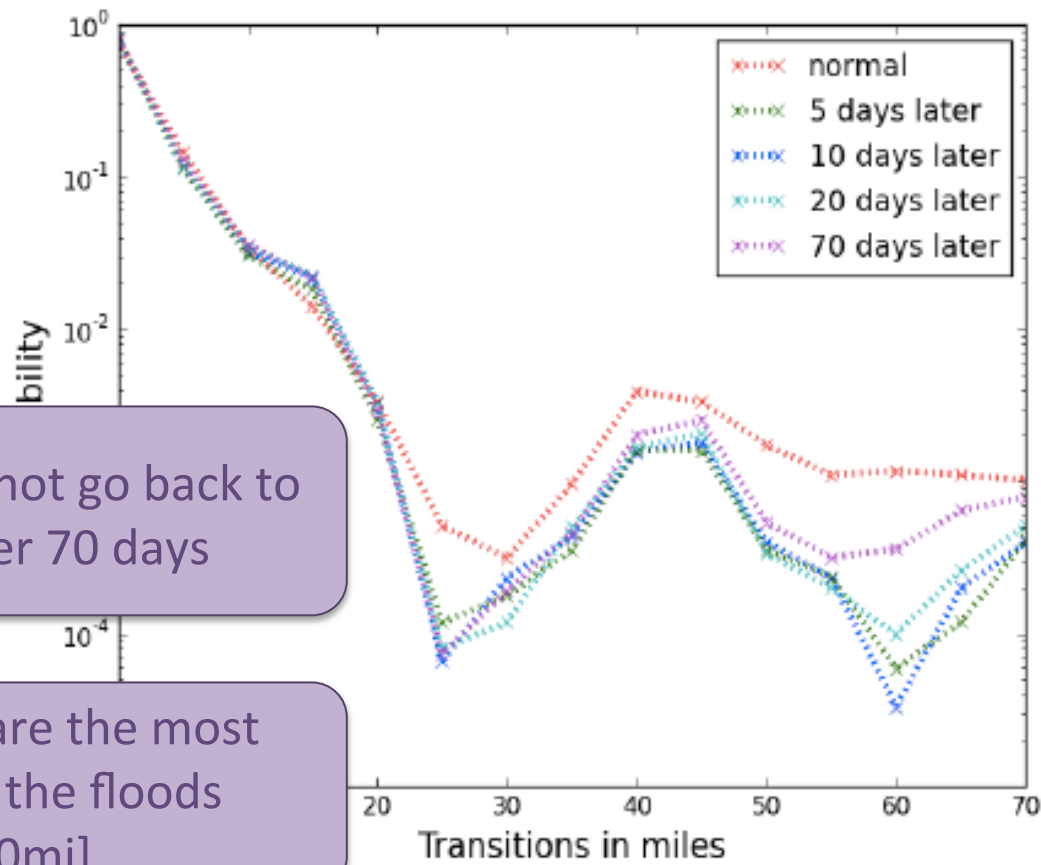
Displaced Individuals – Signatures



Mobility Disruption and Resilience



Disruption and Resilience – Musanze (Rwanda)



Current and Future Work

- Biases in data
- Working with decision makers:
 - Data, from collection to sampling
 - Tool usage, from modeling to analysis
- Working with communities:
 - Data collection and ownership
 - Transparency of models
- Interdisciplinary research takes time and effort



THANK YOU!!!

vfrias@umd.edu