



Understanding Migration Flows with Mobile Data. Challenges & Opportunities

Pedro A. de Alarcon, PhD
Head of Big Data for Social Good

AI for Social Good, Doha (Qatar). Feb 2019

Our Mission: Data as a Force for the Greater Good



“When I think about social good, I think about the commitments we have all made with the UN when it comes to the **17 Sustainable Development Goals** for 2030. Forging a relationship between our big data work for social good is fundamental, especially as **80% of the 6 billion mobile phones in the world are in developing countries**, which is where we can have the greatest impact.”

—*Jose Maria Álvarez Pallete, Chairman of Telefonica*

World Economic Forum Blog.

Big data: moneymaker and force for social good?

Big Data for Social Good: Our commitment with the SDGs

Poverty & Development Metrics

- Improving SDG metrics with telco data (Spain, Central America)



Natural Disaster Response

- Magic box initiative (Colombia)
- Measuring earthquake's impact (Mexico)



Climate change in rural areas

- Internal Forced Displacements (Colombia)
- Smart Livestock Farming (Ecuador)



Air Quality in Urban Areas

- Pollution forecasting (Spain, Brazil)
- Quantifying emissions with mobility data (Germany)



Epidemics & spread forecasting

- Analyzing the spread of Zika (Colombia, Brazil)
- Analyzing the spread of Measles (Brazil)
- Mobility impact analysis of swine flu (Mexico)



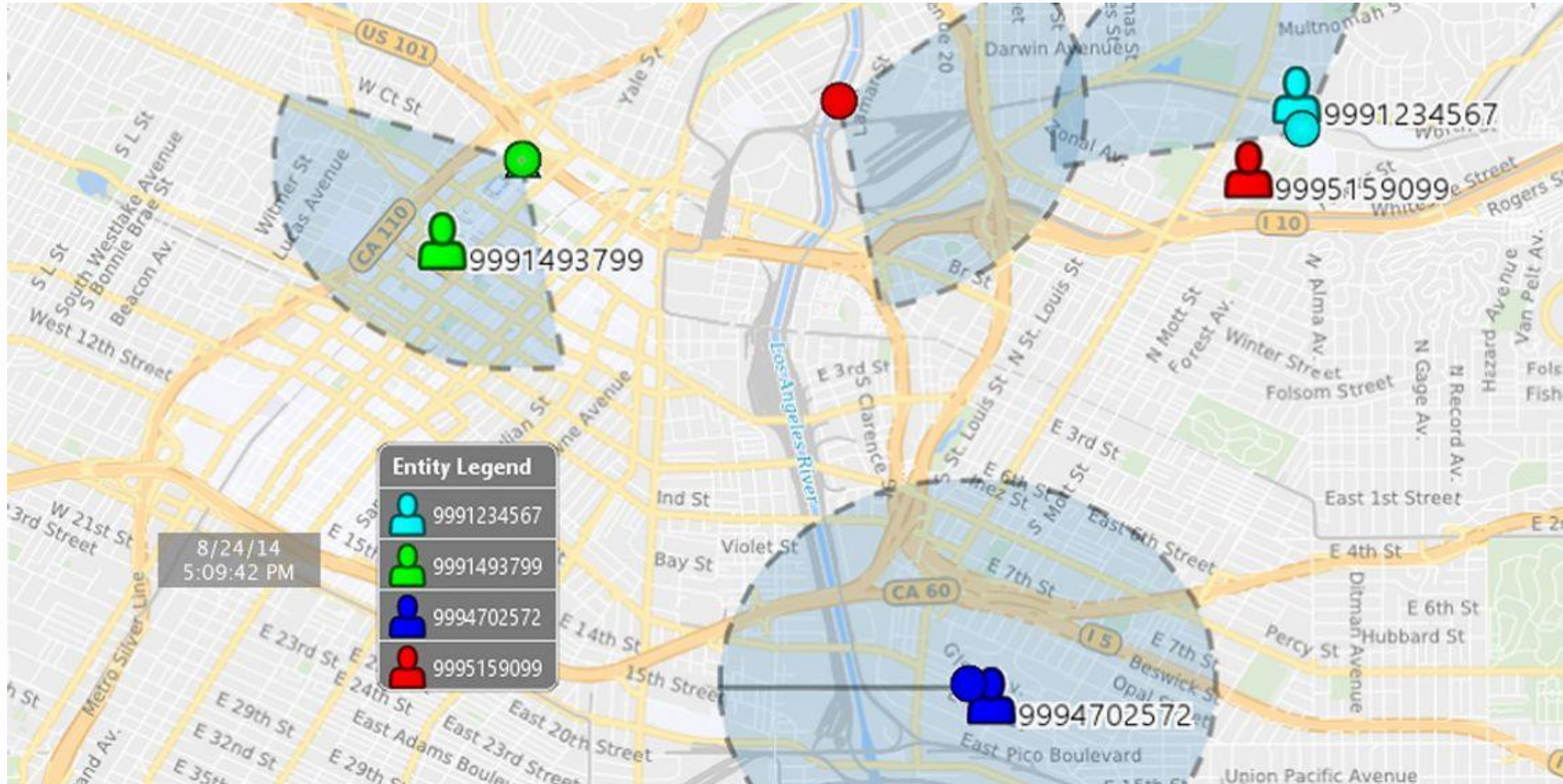
SDGs and mobile phone metadata

Knuper

The opportunities of mobile phone metadata

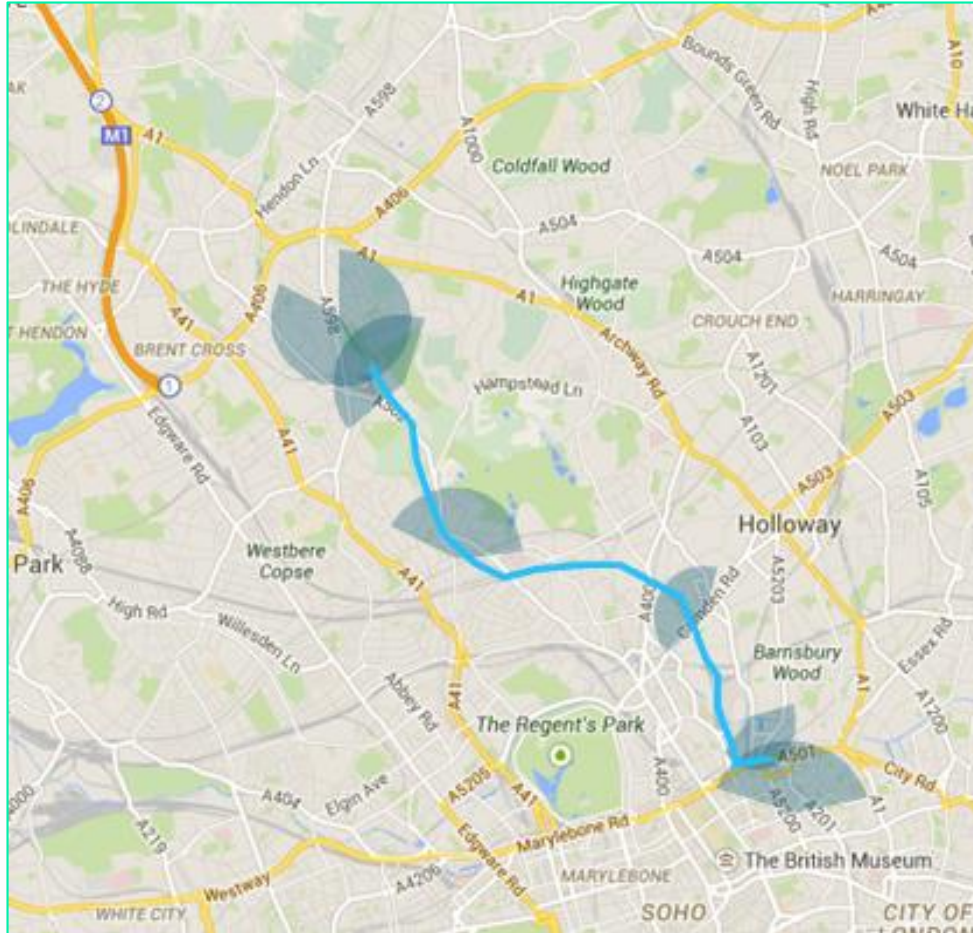
Indicator	SDGs	Proof of Concept			Disaggregation	
		Country	Administrative level	Adjusted R ²		RMSE
Proxy indicator predictable with mobile phone metadata	SDGs for which the indicator is relevant	Country in which the pilot project was carried out	The administrative level on which the groundtruthing was performed with 0 indicating the national level and higher numbers indicating smaller areas	Categorized adjusted R ² indicating how much of the variance is captured by the spatial model	Categorized Root Mean Squared Error measures the standard deviation of the unexplained residuals	Possible disaggregation into sub groups like age and gender
Multidimensional Poverty Index	1.1, 1.2	Sudan, Philippines	2 & 3	>0.75	<0.05	gender
Population Density	various	Senegal	2	pending	pending	age groups & gender
Literacy Rate	4.4, 4.6	Senegal	2, 3 & 4	>0.9	<0.05	gender
Share of Women	10.2 & various others	Senegal	3	>0.9	<0.05	n/a
Electricity Access Rate	7.1, 7.b, 9.1, 11.1	Senegal	3	>0.85	<0.15	n/a
Share of Minority Groups	10.2 & various others	Senegal	3	>0.9	<0.1	n/a
Primary Completion Rate	4.1, 4.5, 4.6, 5.1	Senegal	3	>0.9	<0.05	n/a
Migration Patterns	10.7, 8.8	Turkey	Manually defined	n/a	n/a	n/a
Commuting Patterns	10.7, 8.5, 8.8	Turkey	Manually defined	n/a	n/a	n/a

Mobile phone “metadata” typically means Call Details Records.

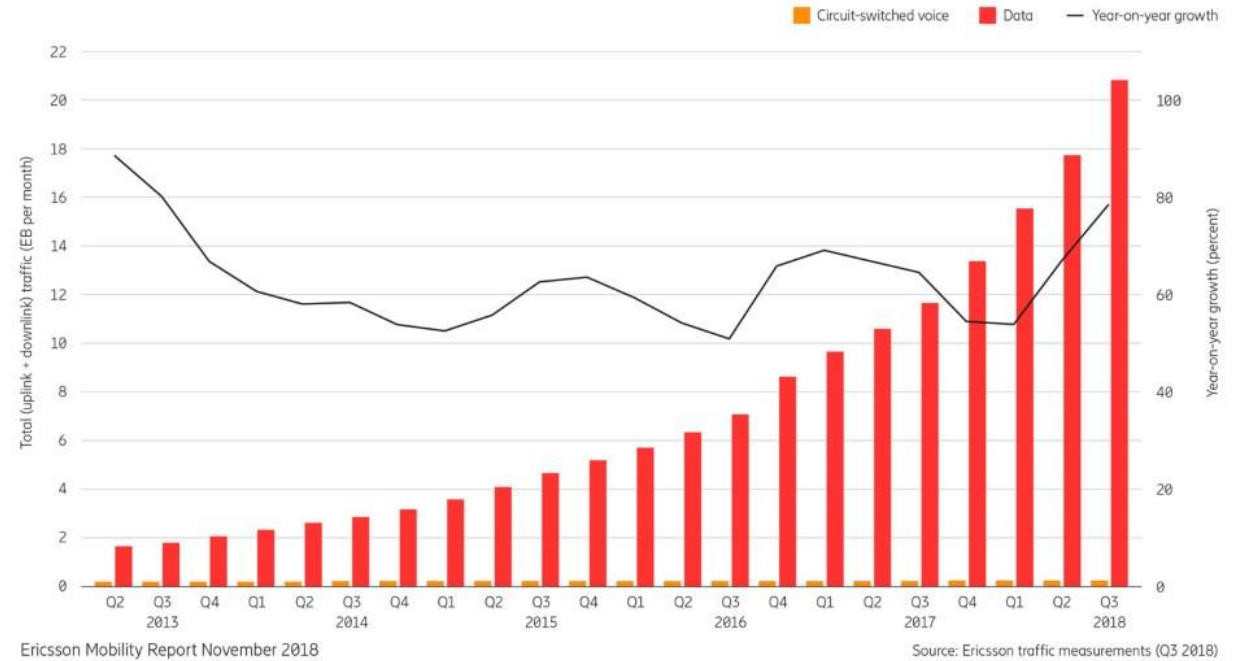


- Per antenna KPIs (activity, pop. density...)
- Social Graph
- Mobility

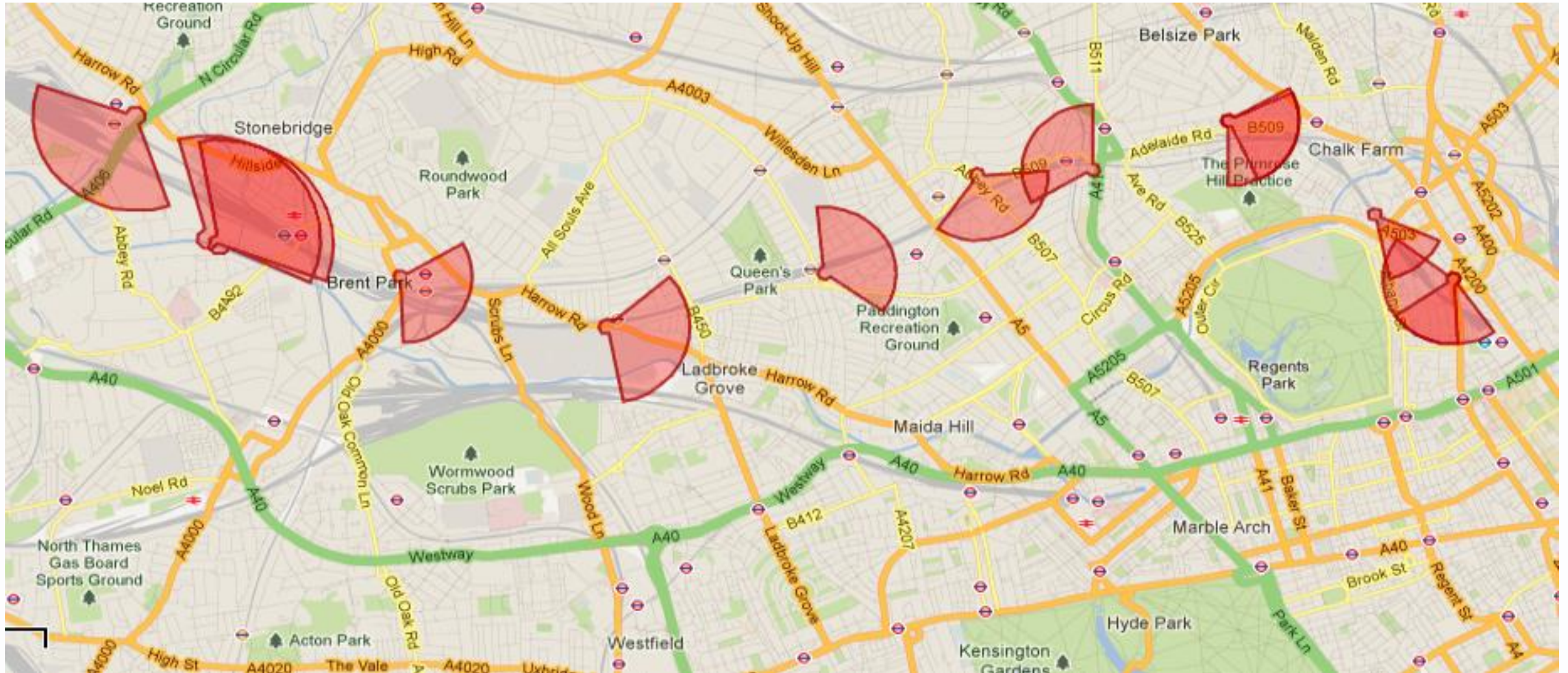
When moving, call logs are like breadcrumbs



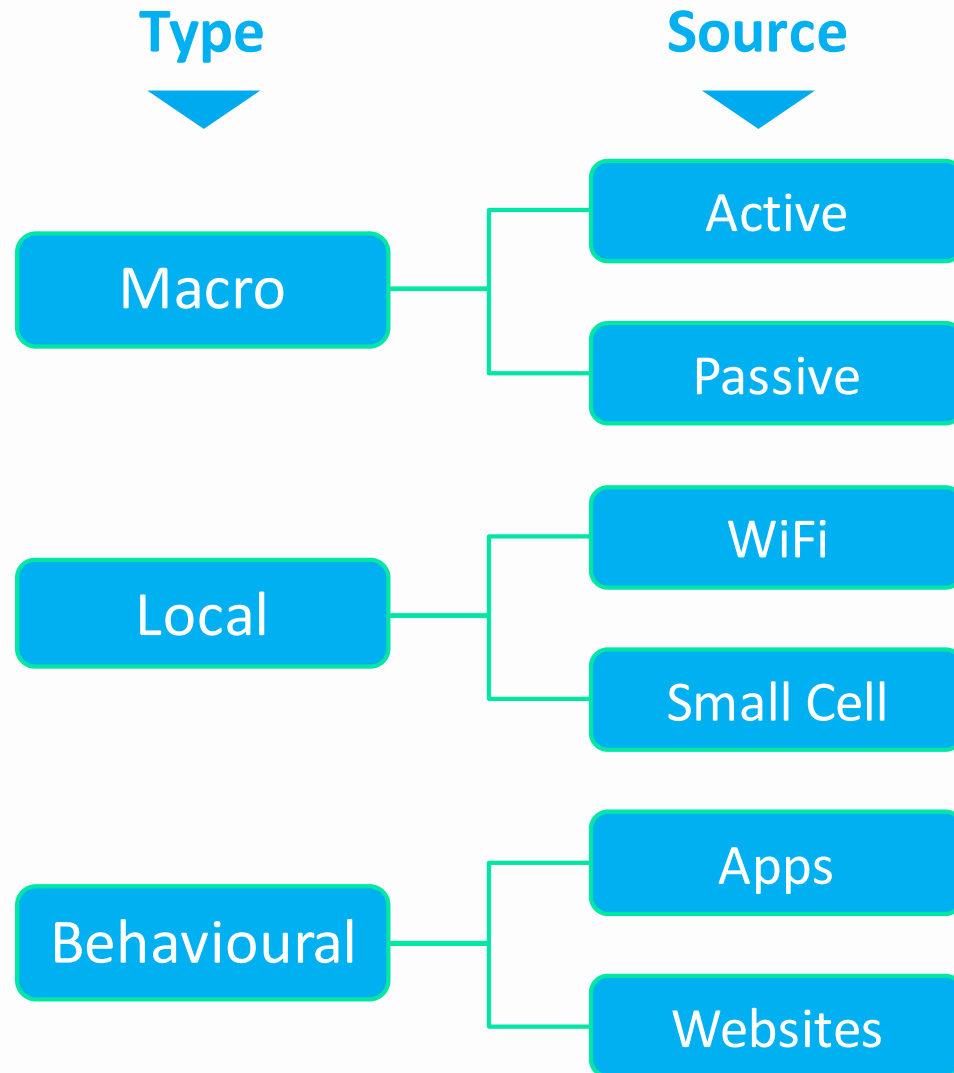
But...



XDRs mean more breadcrumbs (up to 5x)

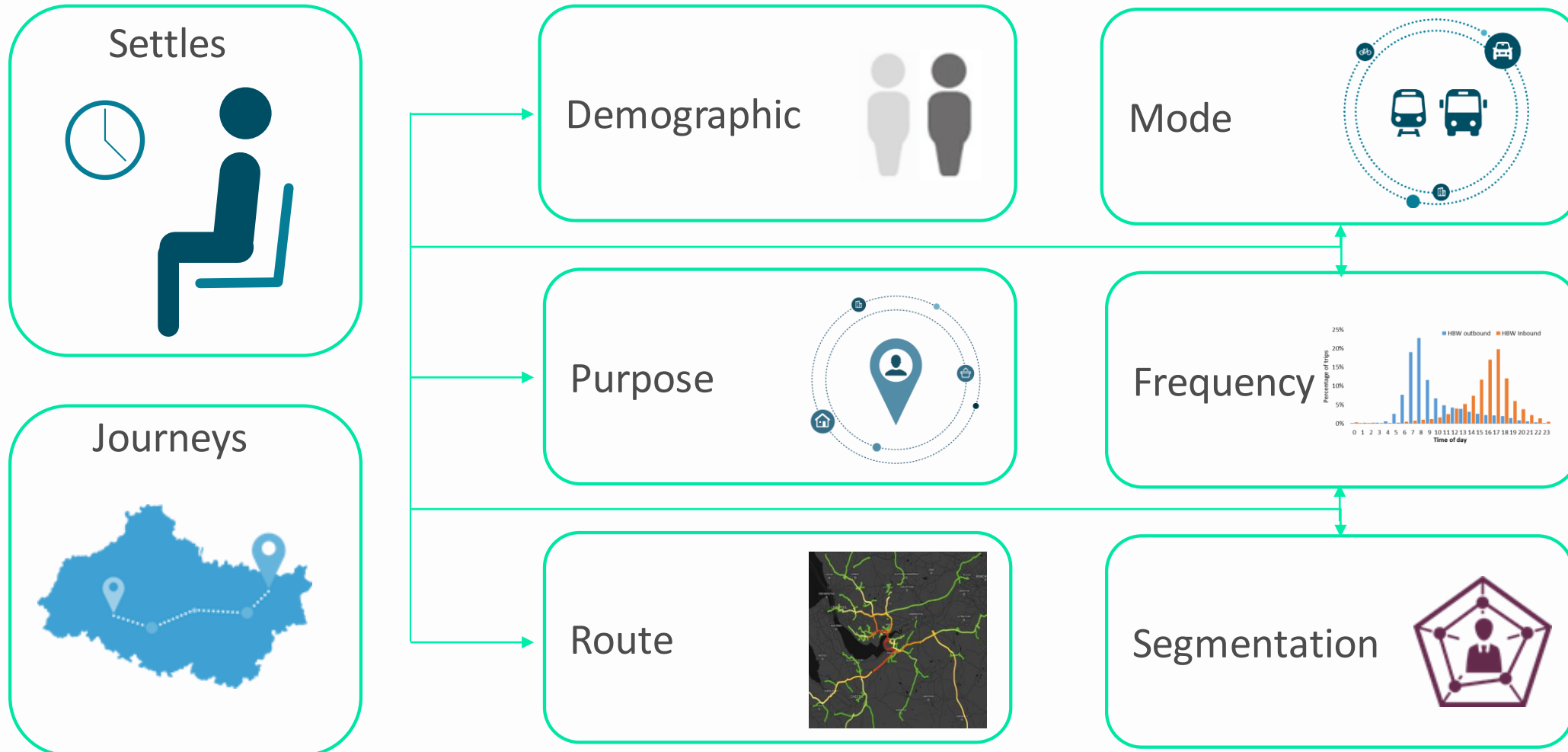


Mobility Insights from Telco Data

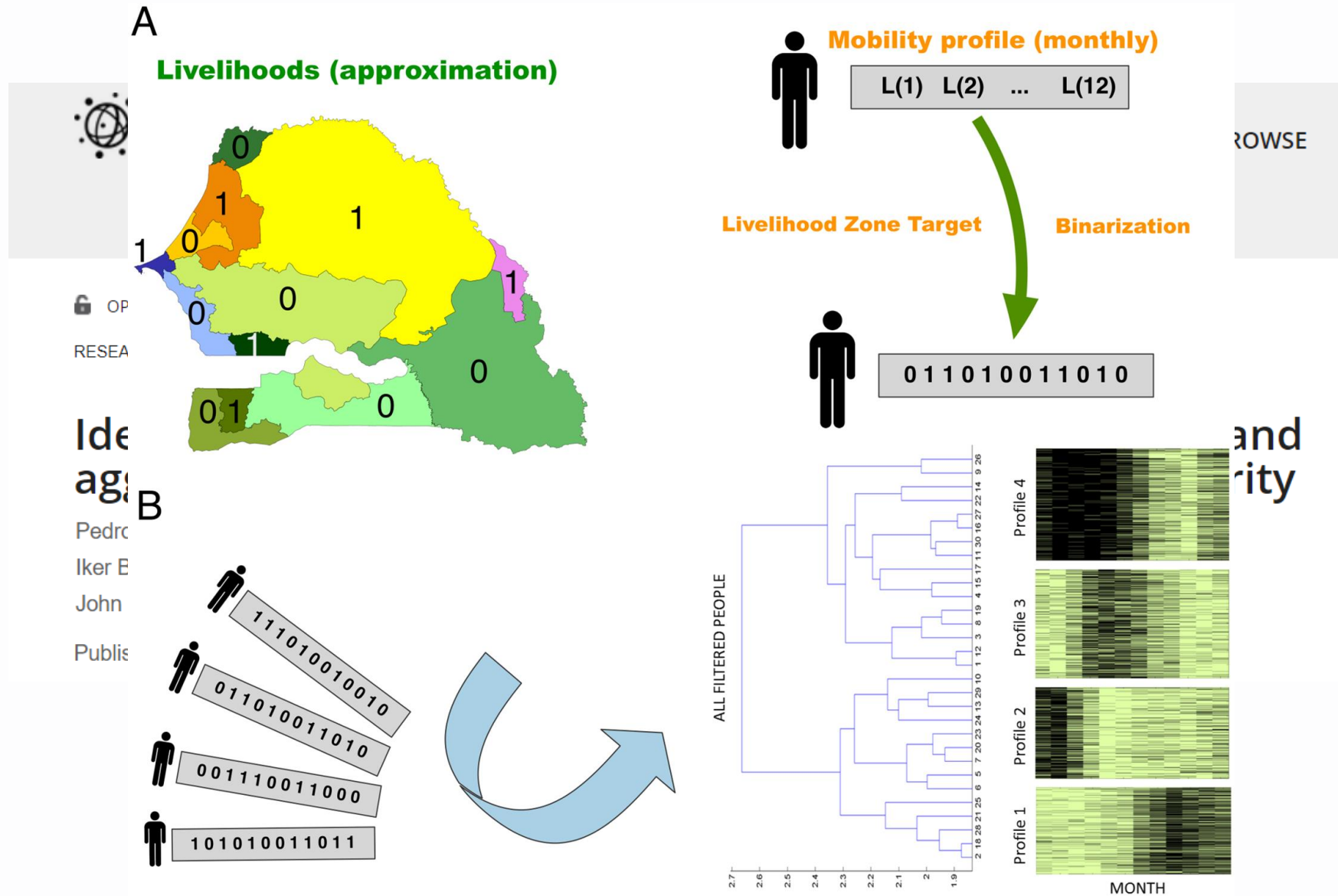


**Billions
of network events
24 / 7 / 365**

Mobile metadata + CRM + Spatial Semantics



The Individual Trajectories Matrix




The Individual Trajectories Matrix

 Adjust temporal resolution to days, weeks, months...


Full
customer
base
(>11M
users)

PhoneID	Week1	Week2	...	WeekN
sdqadw	Bogotá	Bogotá	..	Bogotá
fw4efef	Cali	Cali	...	Bogotá
...				

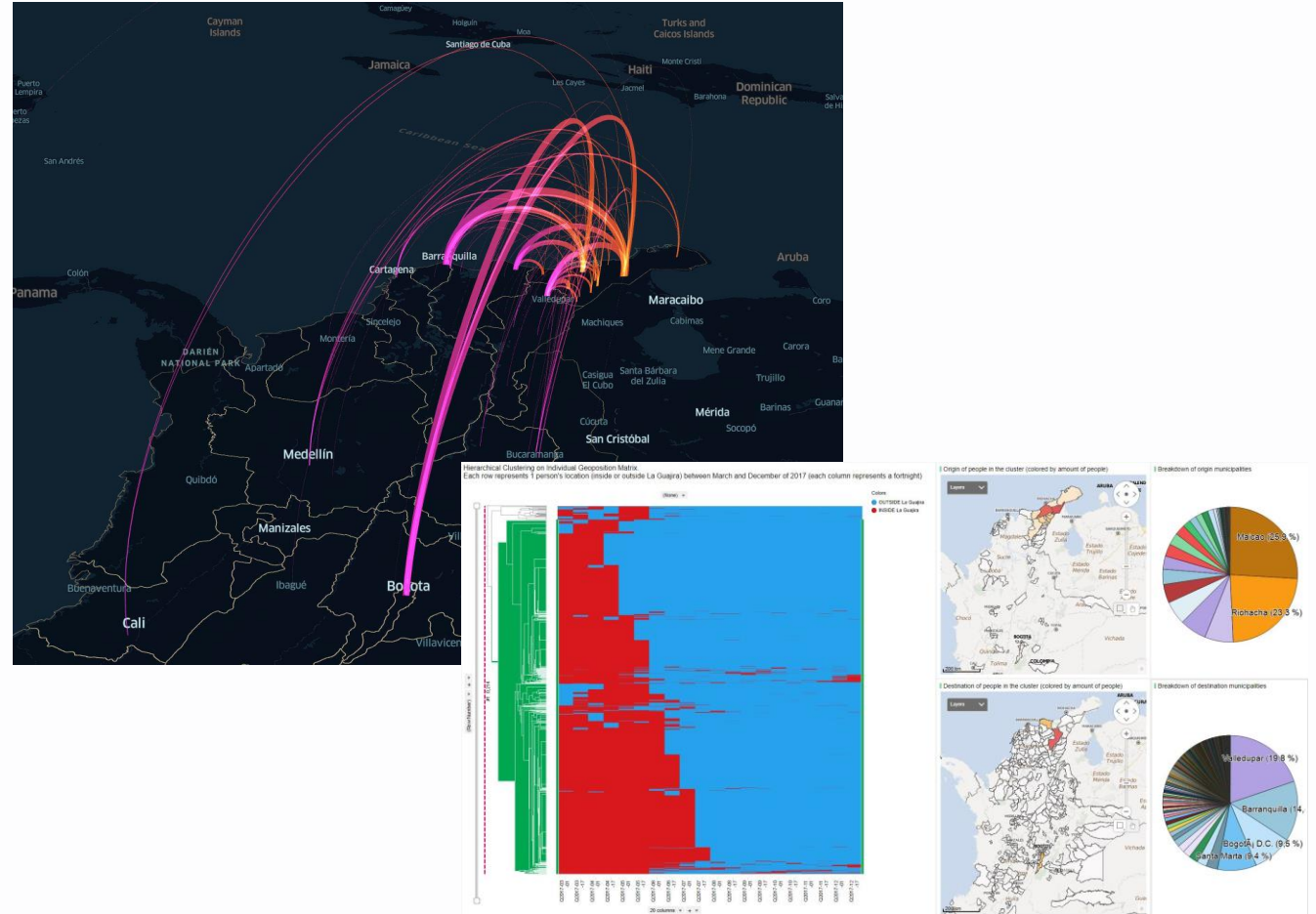

Most common location:
Adjust spatial resolution to
district, municipality,
Department...

Climate change is exacerbating **extreme climate phenomena**.
Rural populations are specially impacted and, in many cases, forced to migrate to urban areas.



Analysing forced displacements due to climate variability in Colombia

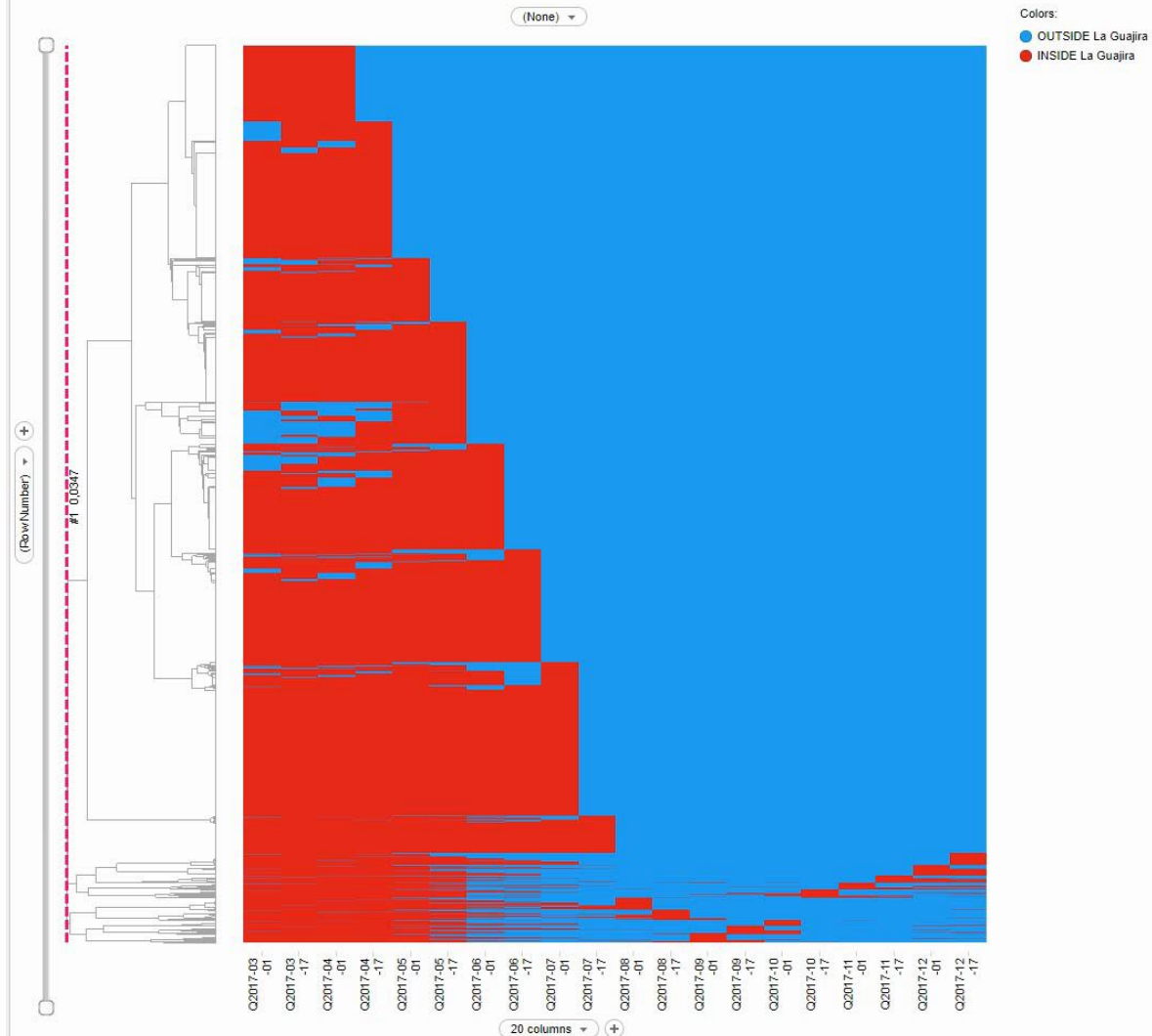
- Telefónica leveraged open datasets from SIAC (Environmental Information System of Colombia) to **identify regions with high vulnerability to drought conditions** and limited ability to recover, focusing on departments **La Guajira, Tolima and Huila**
- Mobility insights from anonymized mobile network data helped to identify **clusters of users with specific long-term movement patterns**, indicative of internal displacement.
- We combined these movement patterns with census data from Colombia's National Administrative Department of Statistics to **estimate the number of people displaced from the focal region** during the period of the drought.



Food and Agriculture
Organization of the
United Nations

Visualization tool to navigate & discover insights

Hierarchical Clustering on Individual Geoposition Matrix.
Each row represents 1 person's location (inside or outside La Guajira) between March and December of 2017 (each column represents a fortnight)

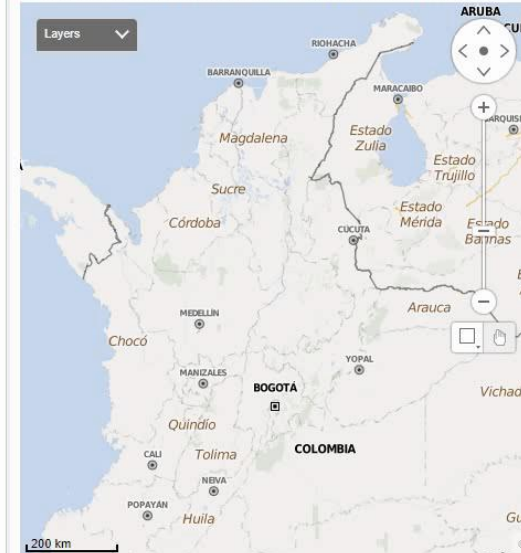


Origin of people in the cluster (colored by amount of people)



Breakdown of origin municipalities

Destination of people in the cluster (colored by amount of people)



Breakdown of destination municipalities

Challenges

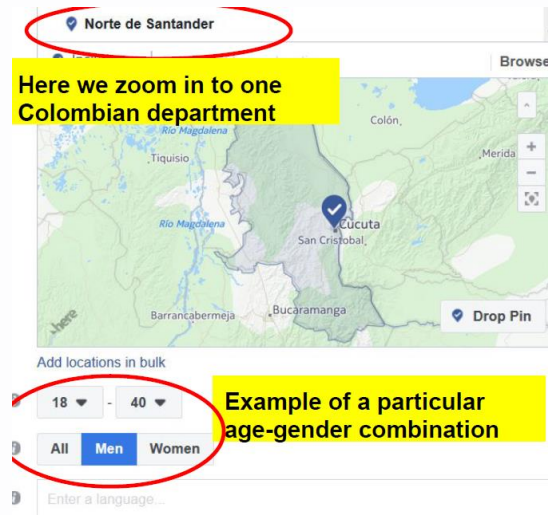
- Limitations to create group profiling:
 - The new personal data privacy regulation (GDPR).
 - Most users are pre-paid subscribers (scarce information from CRM)
 - Addressing different market share rates within the same country
 - Filtering out people < 18yo
- Very few official and reliable **ground truth sources** to compare with.

Opportunities

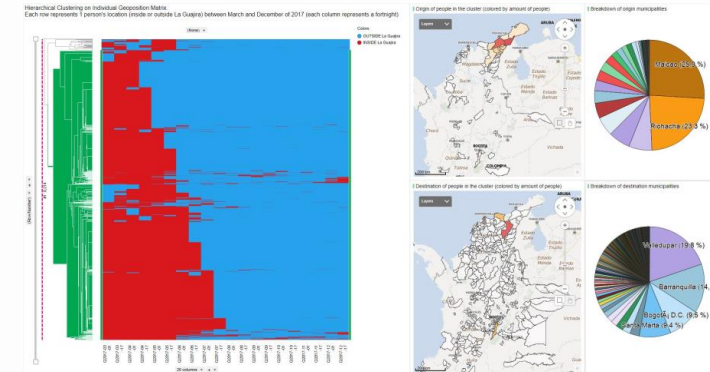
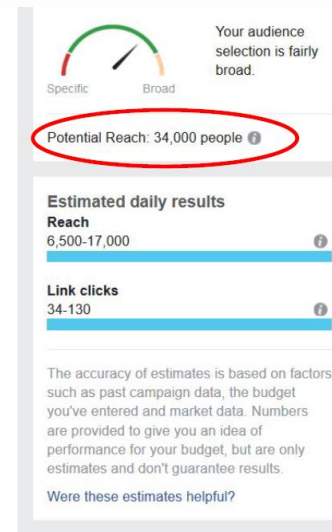
Building robust IDP indicators from multidimensional data integration (from public and private sources)



Satellite Imagery & Computer Vision
Digital Globe



Social Media
Facebook and others



Telco Data
Telefonica

Thank you!

www.luca-d3.com

Follow us on Twitter: @pdealarcon @LUCA_D3