UN Global Pulse:
Harnessing Big Data for a Revolution in Sustainable Development and Humanitarian Action

Robert Kirkpatrick
Director
@rkirkpatrick

www.unglobalpulse.org
@unglobalpulse

GLOBAL PULSE

Global Pulse

Vision: Big Data harnessed responsibly as a public good

Mission: Accelerate discovery and adoption of big data innovation for sustainable development and humanitarian action

Pulse Lab Kampala

Est. 2013

Pulse Lab Jakarta

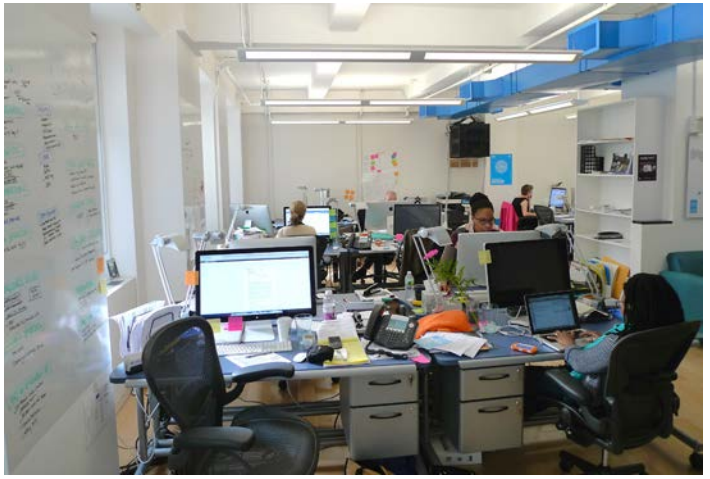
Est. 2012

Pulse Lab New York

Est. 2010

PULSE LABS:
COLLABORATIVE BIG DATA INNOVATION

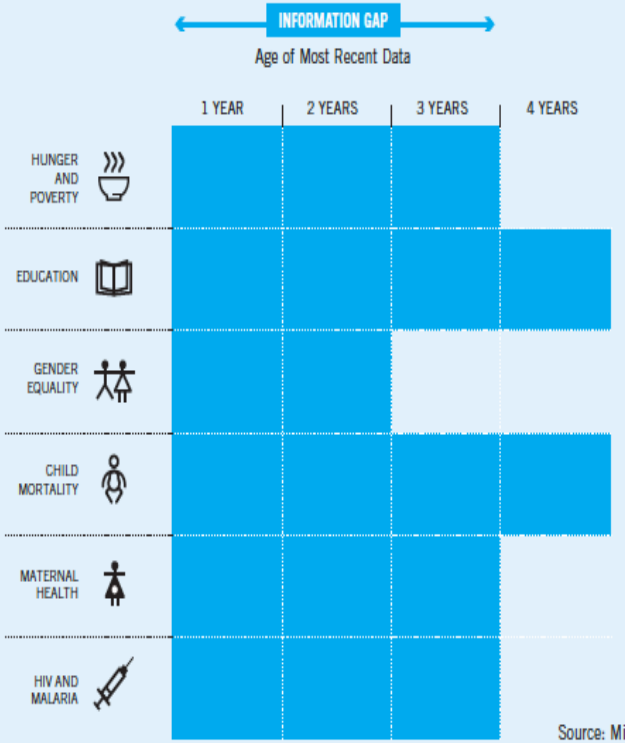
GLOBAL PULSE



Why do we need a data revolution?

The Information Gap

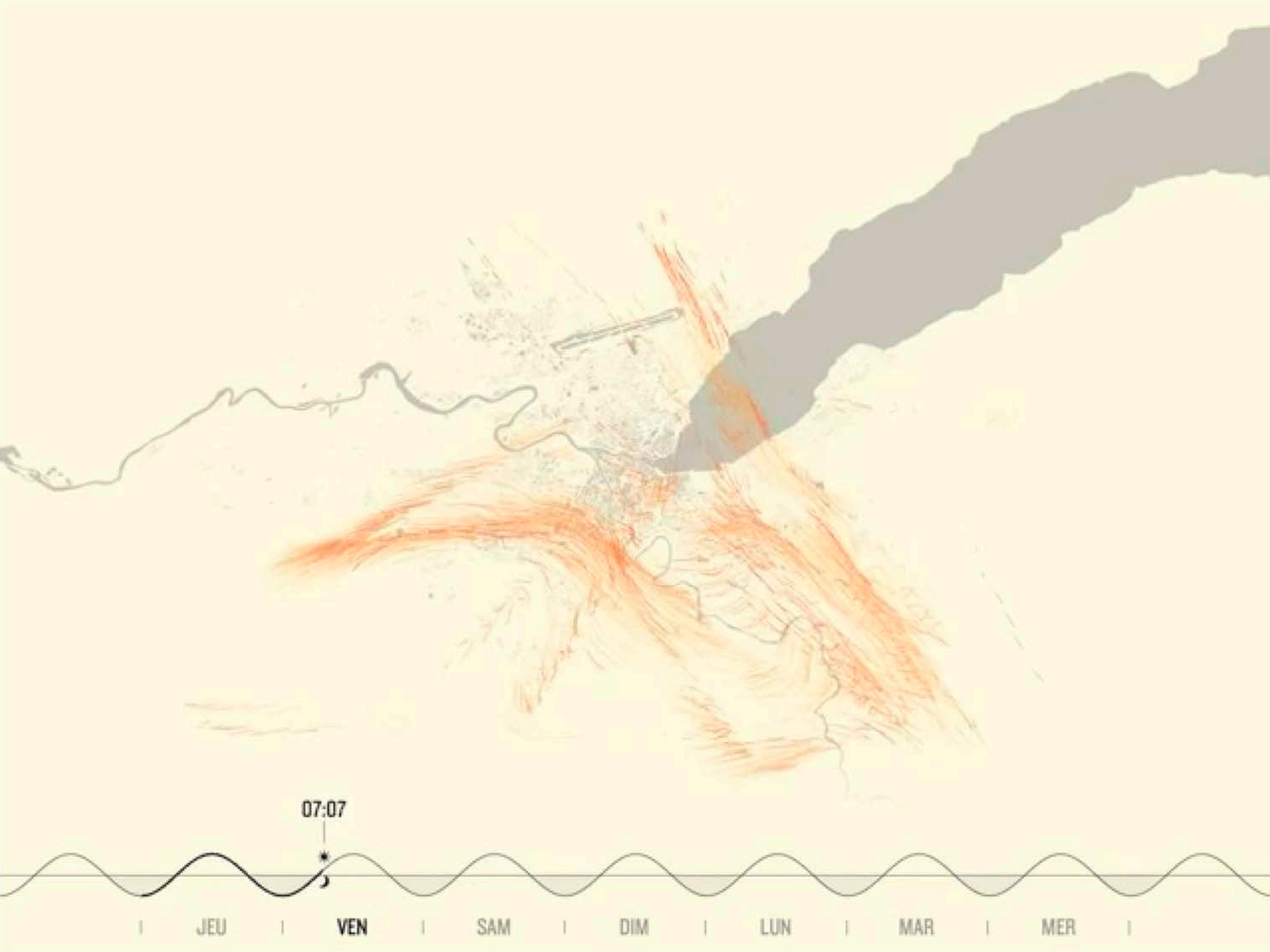
Household-level data is so hard to collect that the information being used to track development progress is frequently out of date. We know what happened in the past but not what's happening in the present. Real-time data could strengthen progress indicators by closing the information gap.



Source: Millennium Development Goals Report 2011

It is no longer enough to have processes designed for less disruptive times.

Kamal
Kamal Muara
Tanjung Priuk
Koja
Kapuk Muara
Pegadungan
Jurumudi
Kapuk
Cengkareng Timur
Kamal Muara
Penjaringan
Duri Kepa
Kebon Jeruk
Kembangan Utara
Kembangan Selatan
Kembangan Barat
Kembangan Timur


07:07
JEU
VEN
SAM
DIM
LUN
MAR
MER

4h
12h
18h
24h
senseable city lab ::: BBVA
1000€ 500€ 100€
APRIL 2011 TRANSACTIONS IN
GROCERIES GAS STATIONS FASHION BAR

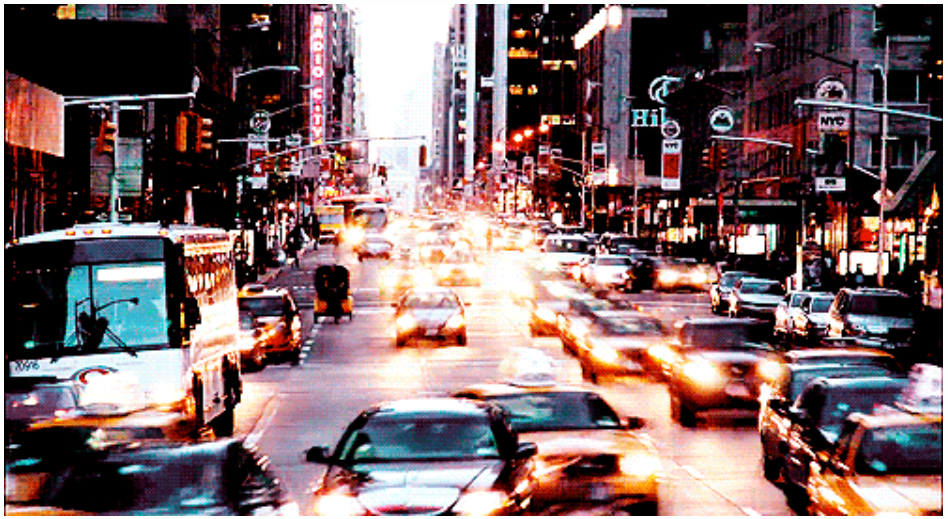
Real-time analytics allows continuous observation of human behavior.

real-time…

Real-time isn’t just faster. It’s different.

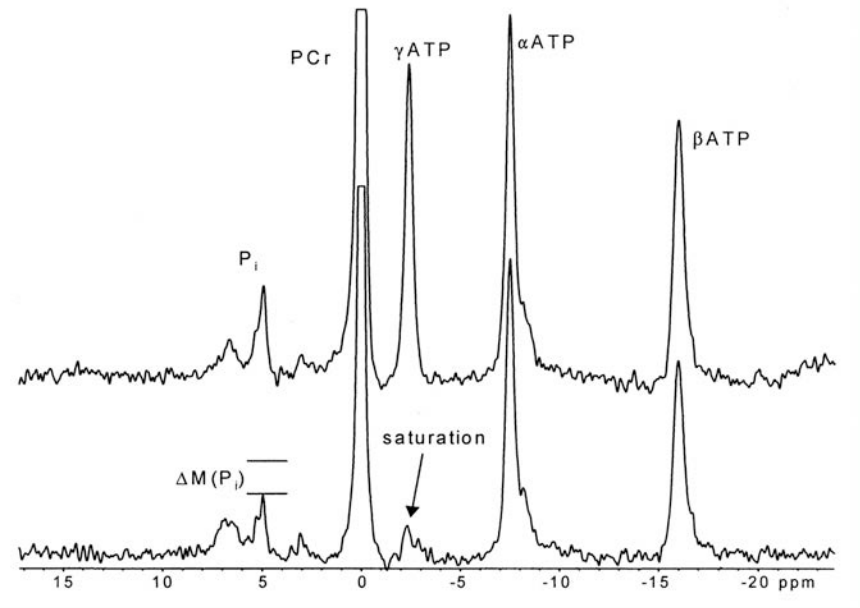
continuous…

survey data



big data

…observation…



Hypothesis:

DIGITAL SERVICES ARE SENSOR NETWORKS FOR MEASURING HUMAN WELLBEING

Data sources which reveal behavior change:

“What People Say”

Online news

Social media

Retail advertising

Radio & TV

“What People Do”

Online search

Mobile phone usage

Transaction records

Postal traffic

GLOBAL PULSE

OUR PRIVACY & DATA PROTECTION PRINCIPLES

We do not access data containing Personally Identifiable Information on any individual, without a freely given, unambiguous and informed consent

We never access the content of private communications

We never attempt to re-identify de-identified data

We ensure appropriate technical and administrative safeguards are in place to prevent unauthorized disclosure or breach of data

We access, analyze, store, transmit or otherwise use only data that has been lawfully and properly obtained from partners

We design, carry out, report and document our activities accurately, transparently and objectively

We employ even stricter standards of care while conducting research among vulnerable populations and persons at risk, children and young people

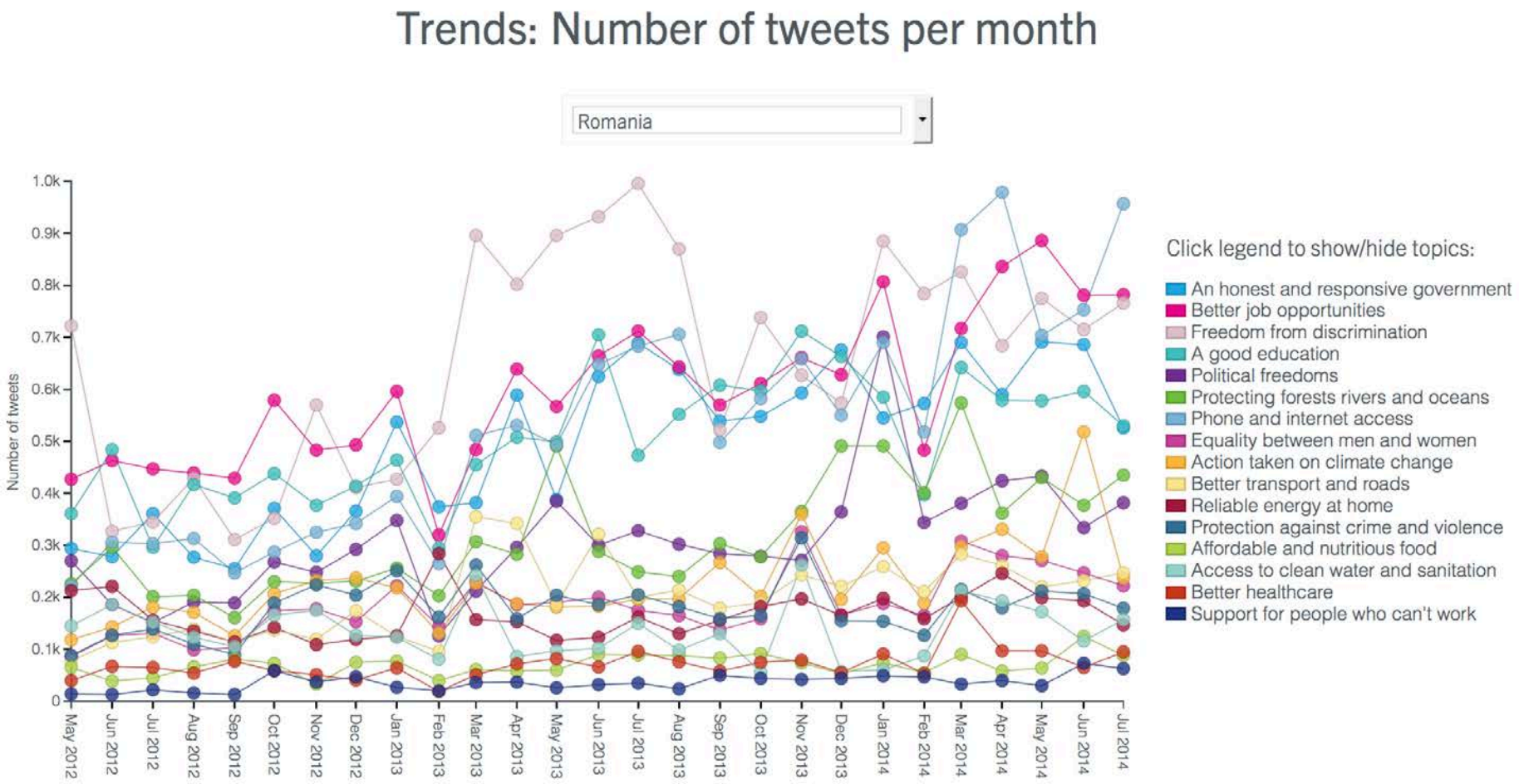
We perform due diligence when selecting data or service provider partners and ensure their activities comply with the United Nations' global mandate

We ensure that our research partners are acting in compliance with relevant law and privacy and data protection standards

GLOBAL PULSE

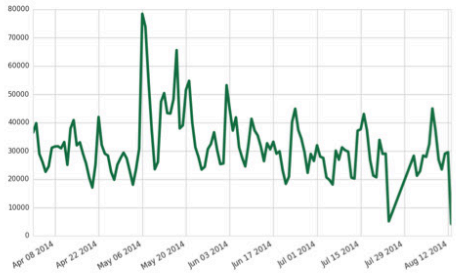
http://www.unglobalpulse.org/privacy-and-data-protection

Global Post2015 Dashboard

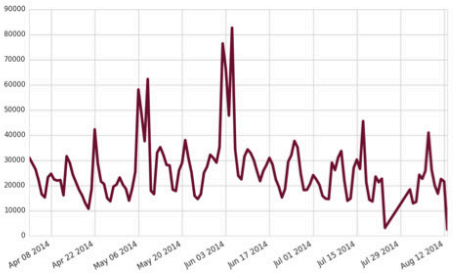


Climate monitoring dashboard

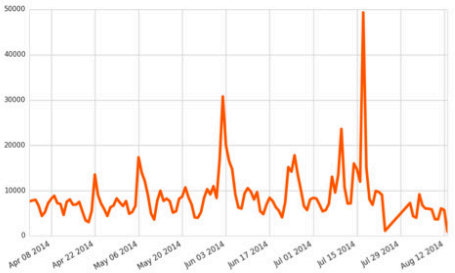
General



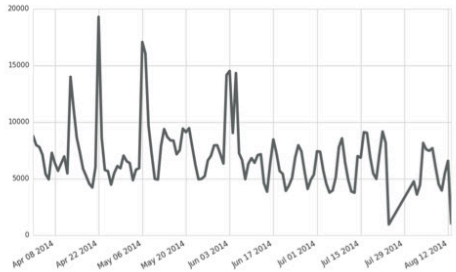
Politics/Opinion



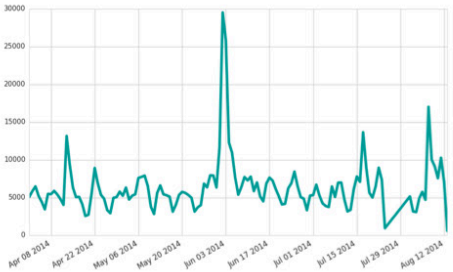
Economy



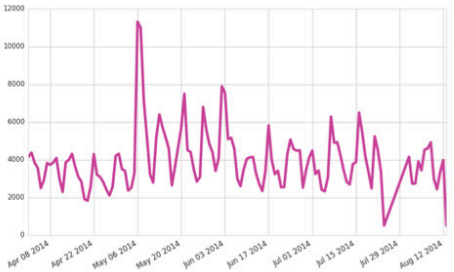
Risk/Disaster



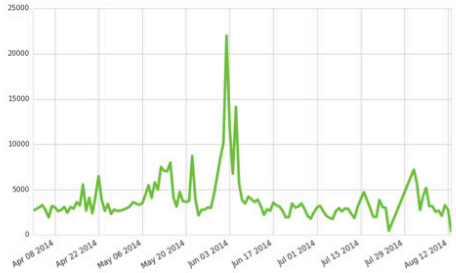
Energy



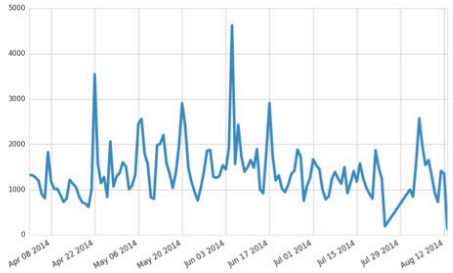
Weather



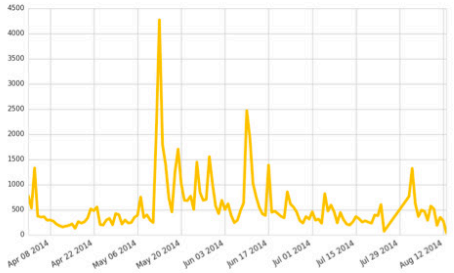
Agriculture/Forestry



Oceans/Water



Arctic



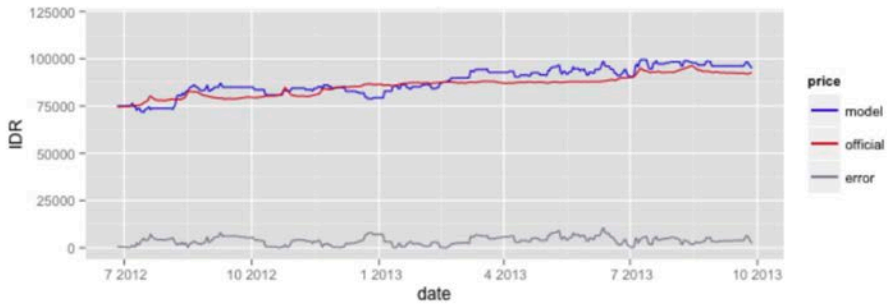
Program: Food & Agriculture

Project: *Nowcasting food prices via Indonesian Twitter*

Partners: WFP, BAPPENAS, University of Seoul

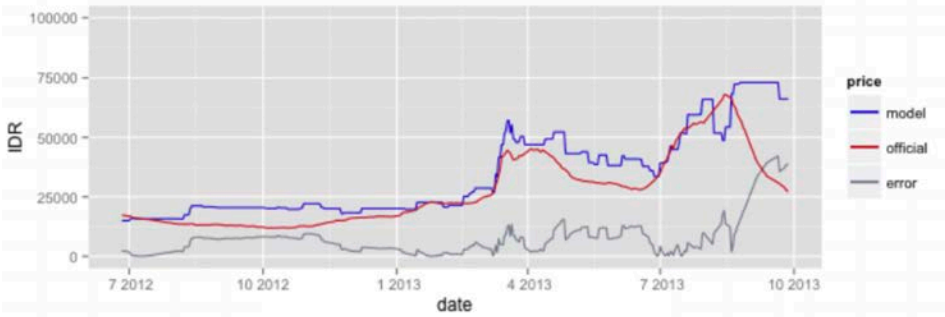
BEEF: The official beef price was relatively stable for the timeframe analysed, and the long-term price trend was accurately modeled (nowcasted prices remained close to real prices throughout the timeframe).

14,473 tweets contained price quotes on beef.



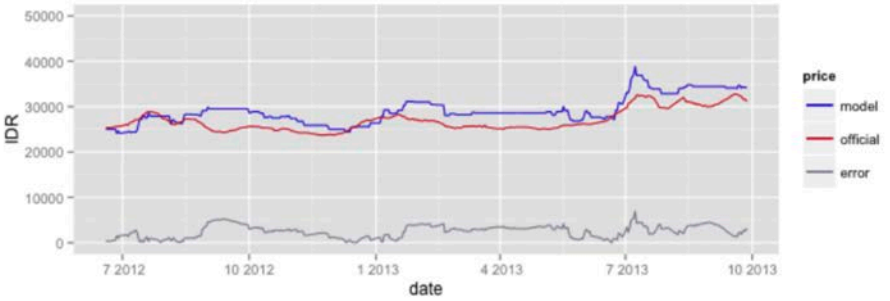
ONION: Onion ("bawang") fluctuated the most of the four commodities. The model effectively tracked price variation around April 2013, but due to the low number of tweets containing price quotes in August did not provide a successful price proxy.

1,954 tweets contained price quotes on onions.



CHICKEN: The tool successfully modeled the price of chicken ("daging ayam") with significant correlation for the duration.

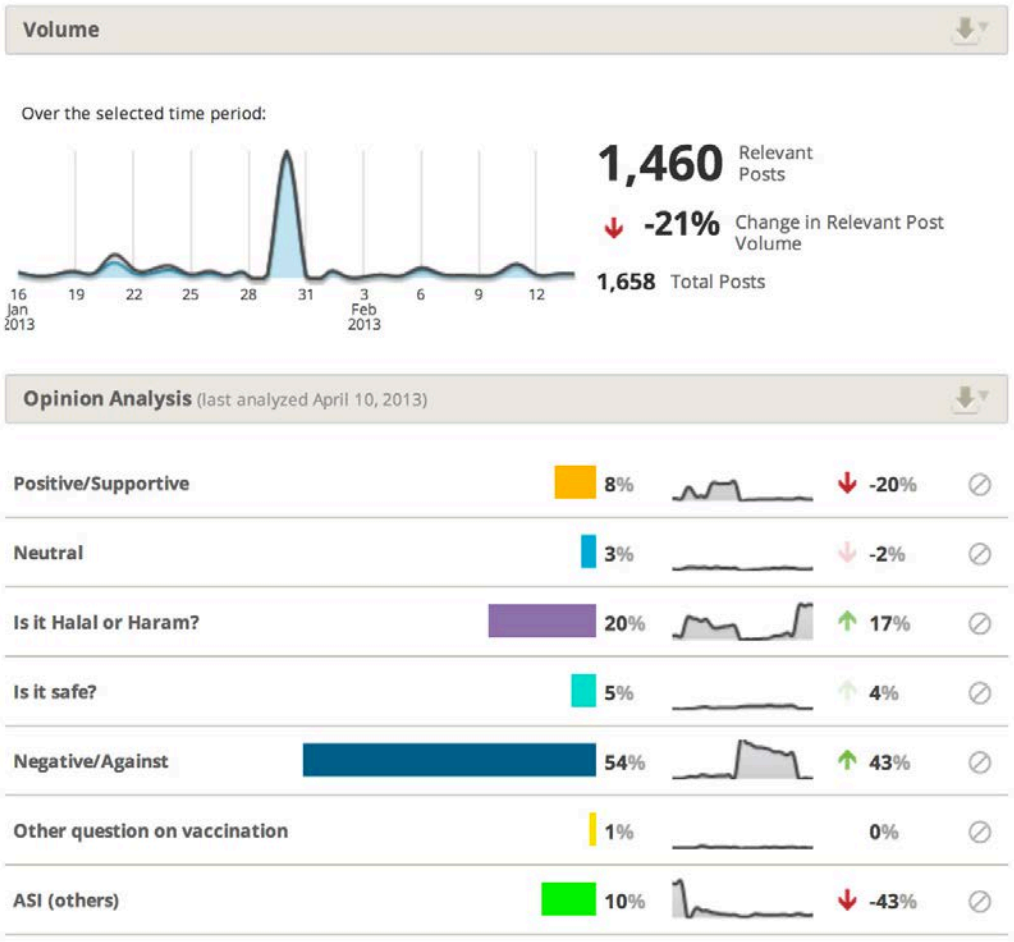
5,223 tweets contained price quotes on chicken.

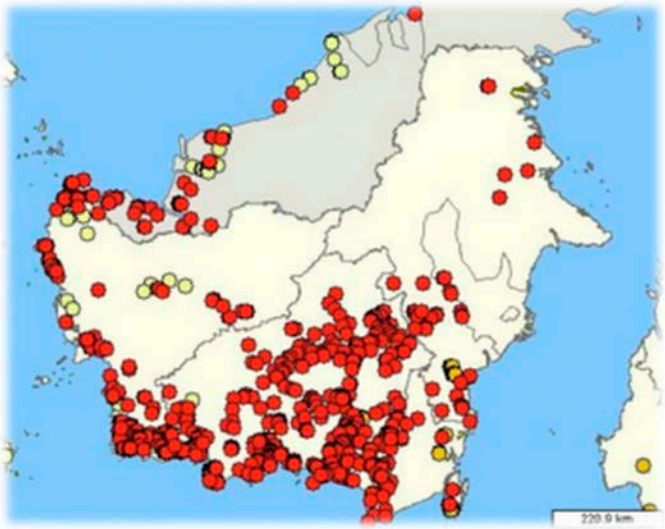


COMMODITY	PEARSON CORRELATION COEFFICIENT
Beef	0.87
Chicken	0.81
Onion	0.85

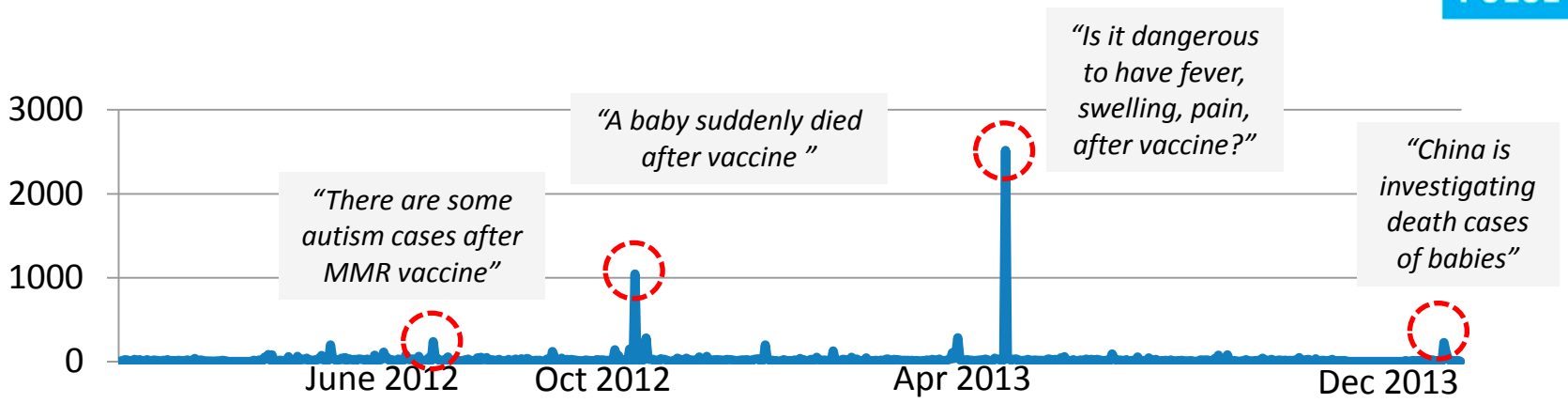
http://www.unglobalpulse.org/nowcasting-food-prices

Program: Public Health
Project: *Perceptions of Vaccination in Indonesia*
Partners: Ministry of Health, WHO, UNICEF
Product: Real-time map and trends in Twitter user sentiment toward vaccination and disease risk





SITUATIONAL AWARENESS – IMMUNIZATION DISCUSSION IN INDONESIA



Rank	2012-06-20	2012-10-08	2013-04-28	2013-12-23
1	Autism (213)	Death(1030)	Fever (1498)	Death (224)
2	Death (5)	Fever (14)	Swelling (1494)	Fever (3)
3	Sick (4)	Sick (4)	Pain (1491)	Crying (1)
4	Fever (2)	Crying (3)	Autism (1011)	Autism (1)
5	Crying (1)	Fever (3)	Fever (4)	-

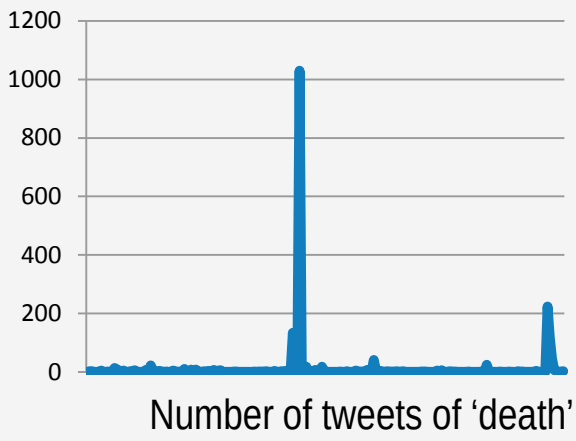
Monitoring a few keywords in digital media allows to understand up-to-date situation in real-time; an addition to existing MoH media monitoring for immediate action

http://www.unglobalpulse.org/immunisation-parent-perceptions

EARLY WARNING AND RAPID RESPONSE

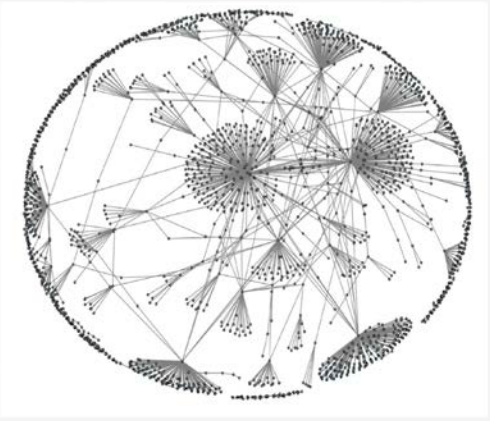
Early warning → **Rapid response with actionable plan**

Detect people concerned about death after vaccine from Twitter



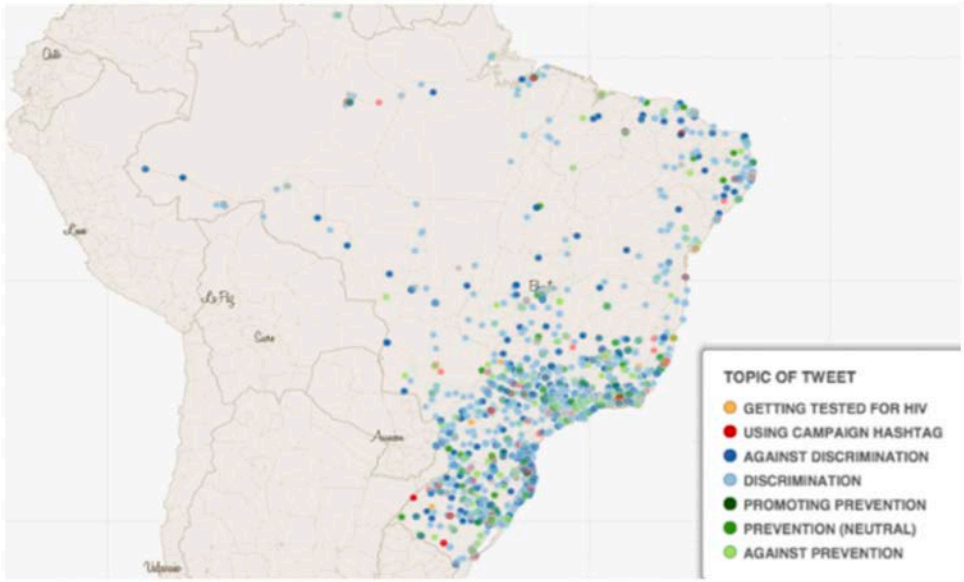
Disseminate correct information through Twitter via influential users

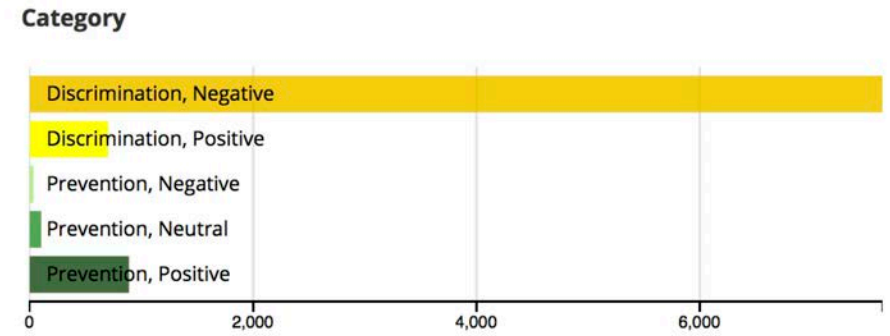
@dr_piprim
@dirgarambe
@blogdoktor
⋮

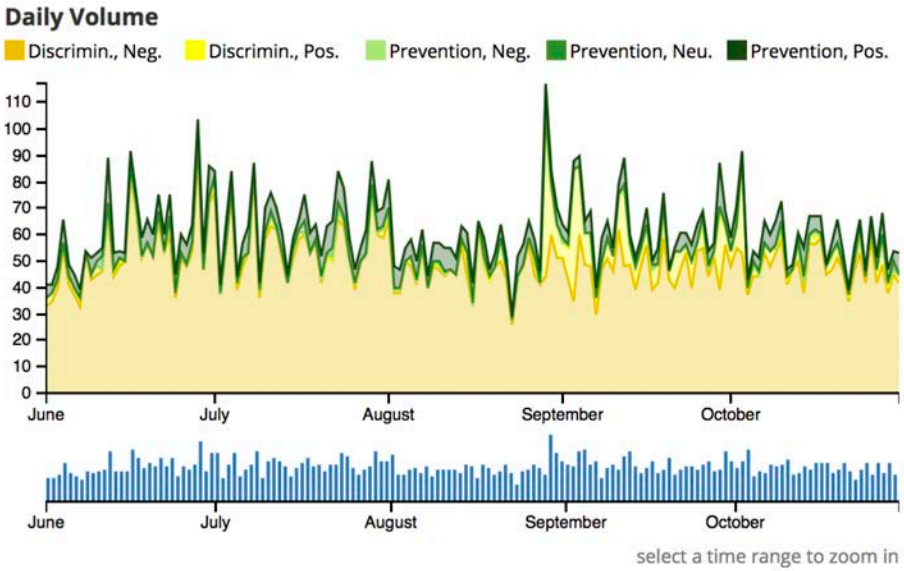


http://www.unglobalpulse.org/immunisation-parent-perceptions

Program: Public Health
Project: *HIV Risk Awareness at the World Cup*
Partners: UNAIDS





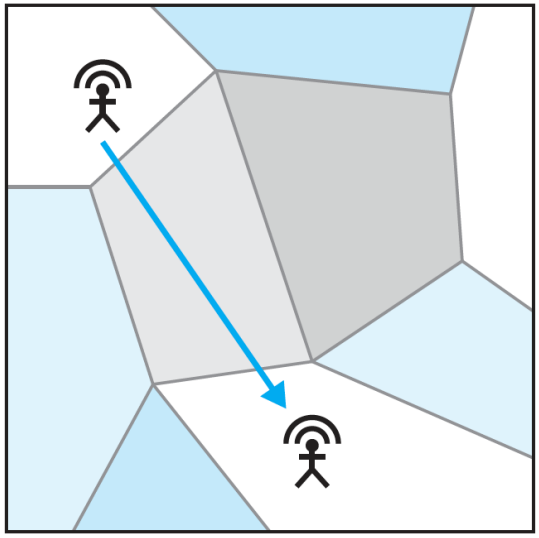


Call Detail Records

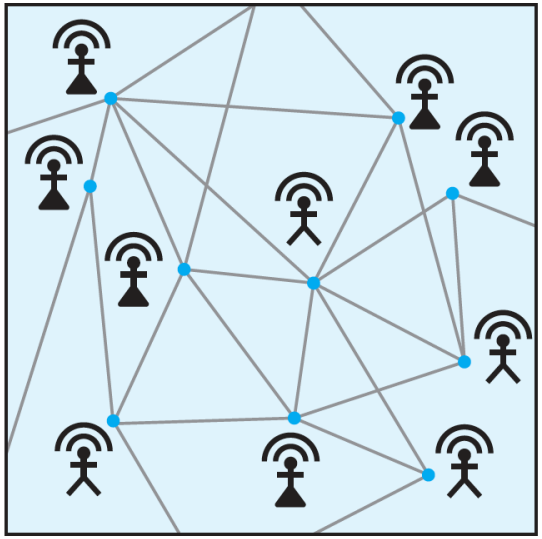
CDR Format:

CALLER ID	CALLER CELL TOWER LOCATION	RECIPIENT PHONE NUMBER	RECIPIENT CELL TOWER LOCATION	CALL TIME	CALL DURATION
X76VG588RLPQ	2°24' 22.14", 35°49' 56.54"	A81UTC93KK52	3°26' 30.47", 31°12' 18.01"	2013-11-07T15:15:00	01:12:02

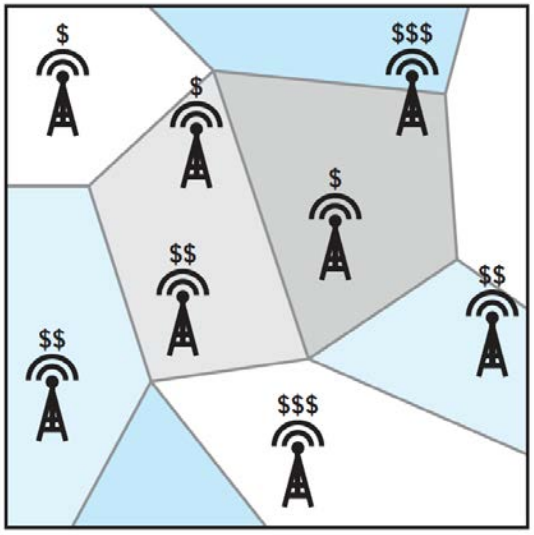
Mobility



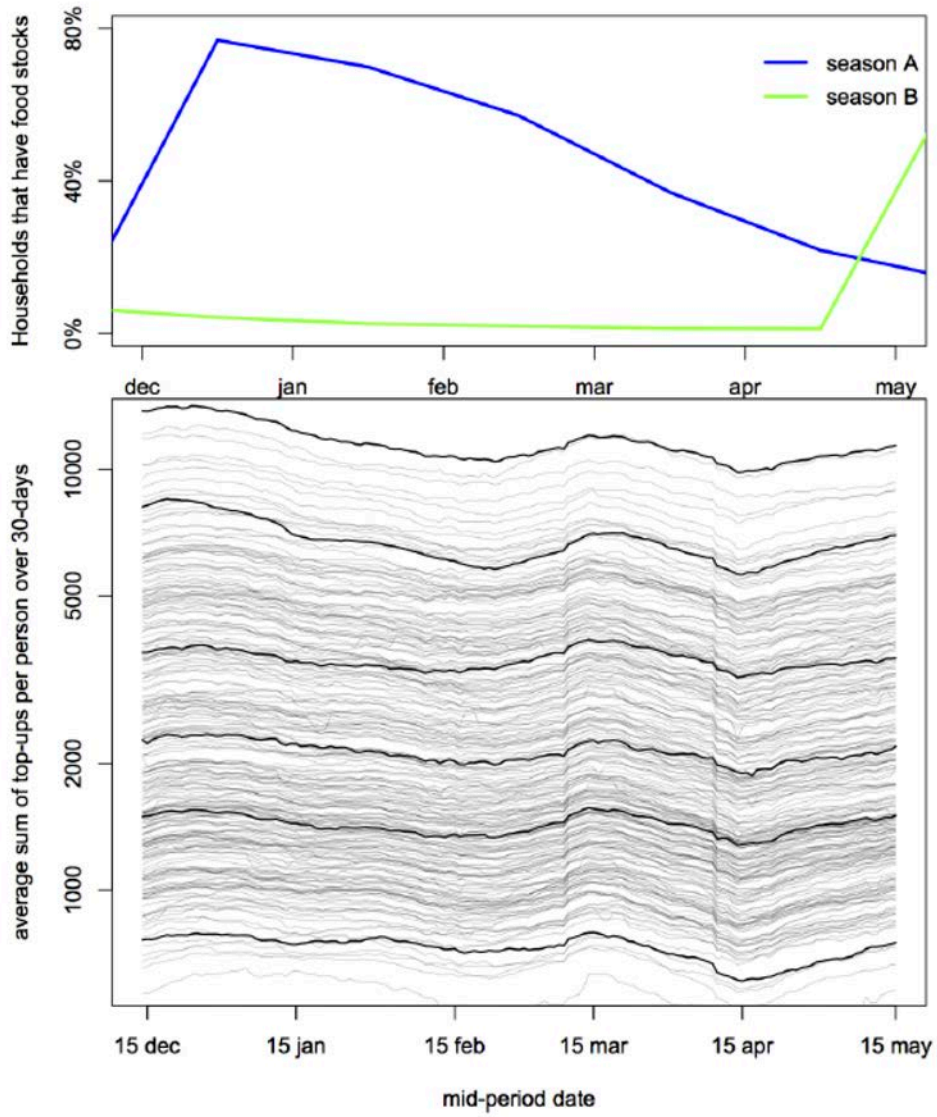
Social Interaction



Economic Activity



Real-time Estimation of Household Food Expenditures Using Anonymized Mobile Phone Airtime Expenditures



Name of variable (question: "how many times have you eaten [item] in the last 7 days?" (answers between 0 and 7))	Cor. sum.
carrot, orange sweet potato (vitamin rich orange vegetables)	0.82
rice, wheat and other cereals	0.76
mandazi/chapatti/bread	0.76
sugar and sweets	0.70
flesh meat	0.69
eggs	0.59
orange coloured fruits	0.57
oil, fat, butter, ghee (including palm oil)	0.54
milk and milk products	0.50
organ meat	0.48
sorghum	0.43
ground nuts and seeds	0.37
other vegetables	0.35
fish	0.30
other fresh fruits	0.28
cooking banana	0.21
dark green leafy vegetables	0.18
beans, peas and other pulses	0.09
condiments	0.07
maize/ maize meal	0.04
other white roots and tubers	0.02
pumpkin, squash and other orange vegetables	0.01
cassava	-0.04
white sweet potato	-0.41

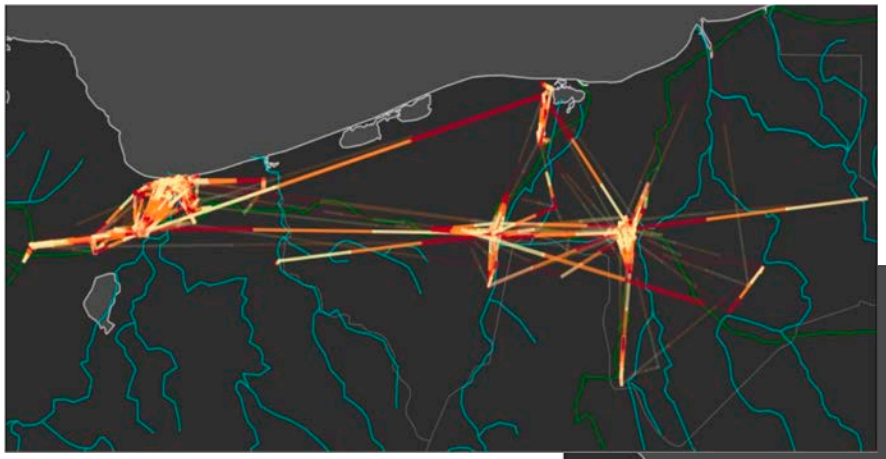
TABLE I
CORRELATION BETWEEN THE CONSUMPTION OF DIFFERENT TYPES OF FOOD AND THE SUM OF AIRTIME EXPENSES.

Program: Humanitarian Action

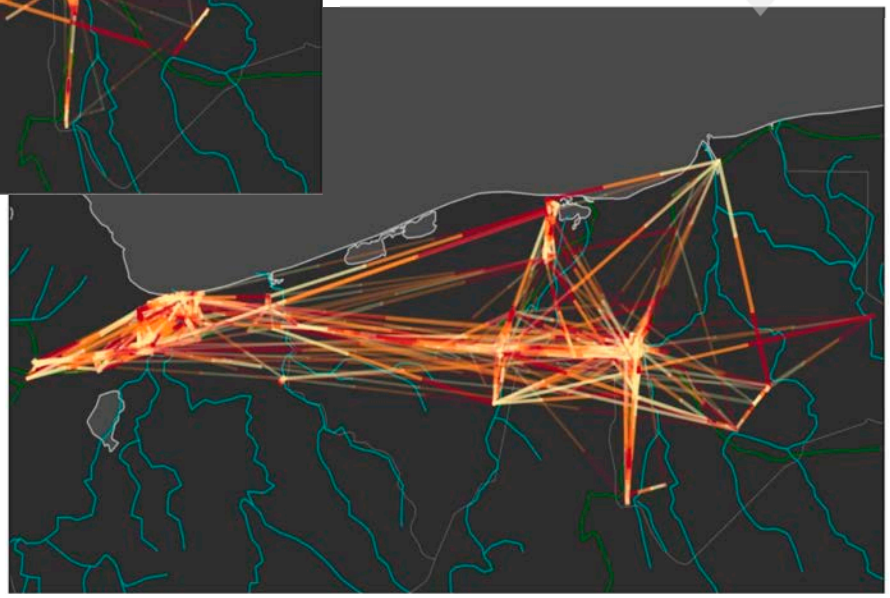
Project: *Visualizing displacement due to floods via mobile phone data*

Partners: WFP, Polytechic University of Madrid, Telefonica Research

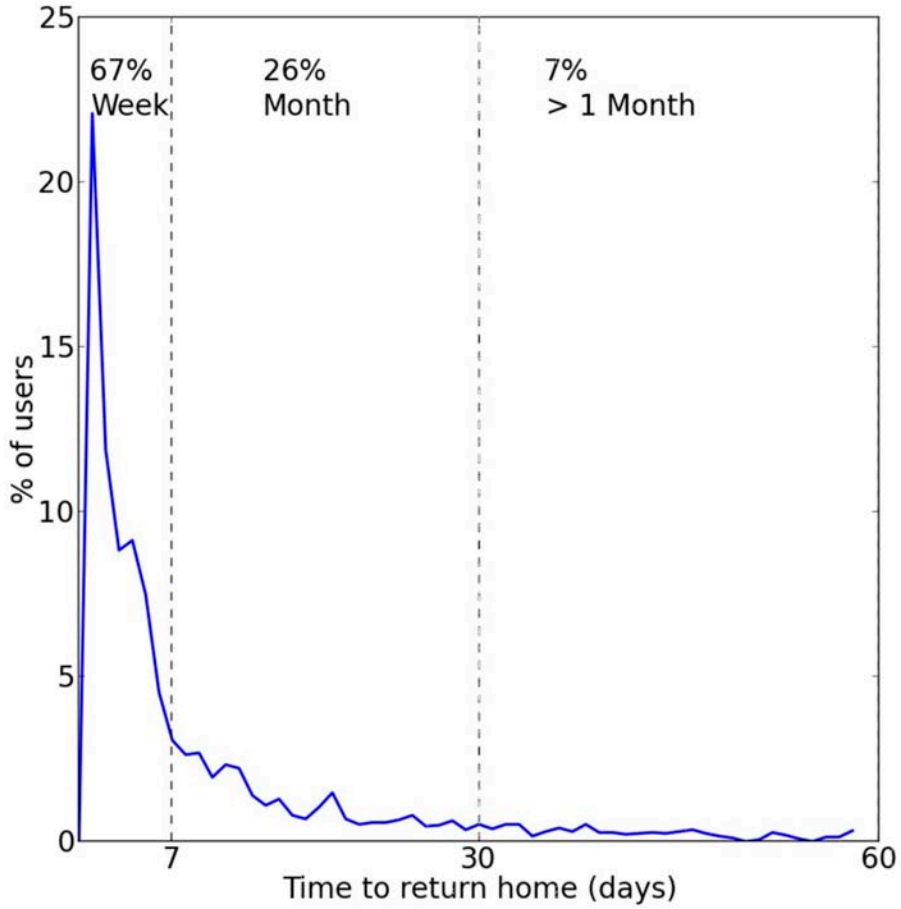
Control night (random)



Network of night displacements during floods



http://www.unglobalpulse.org/tabasco-floods-CDRs



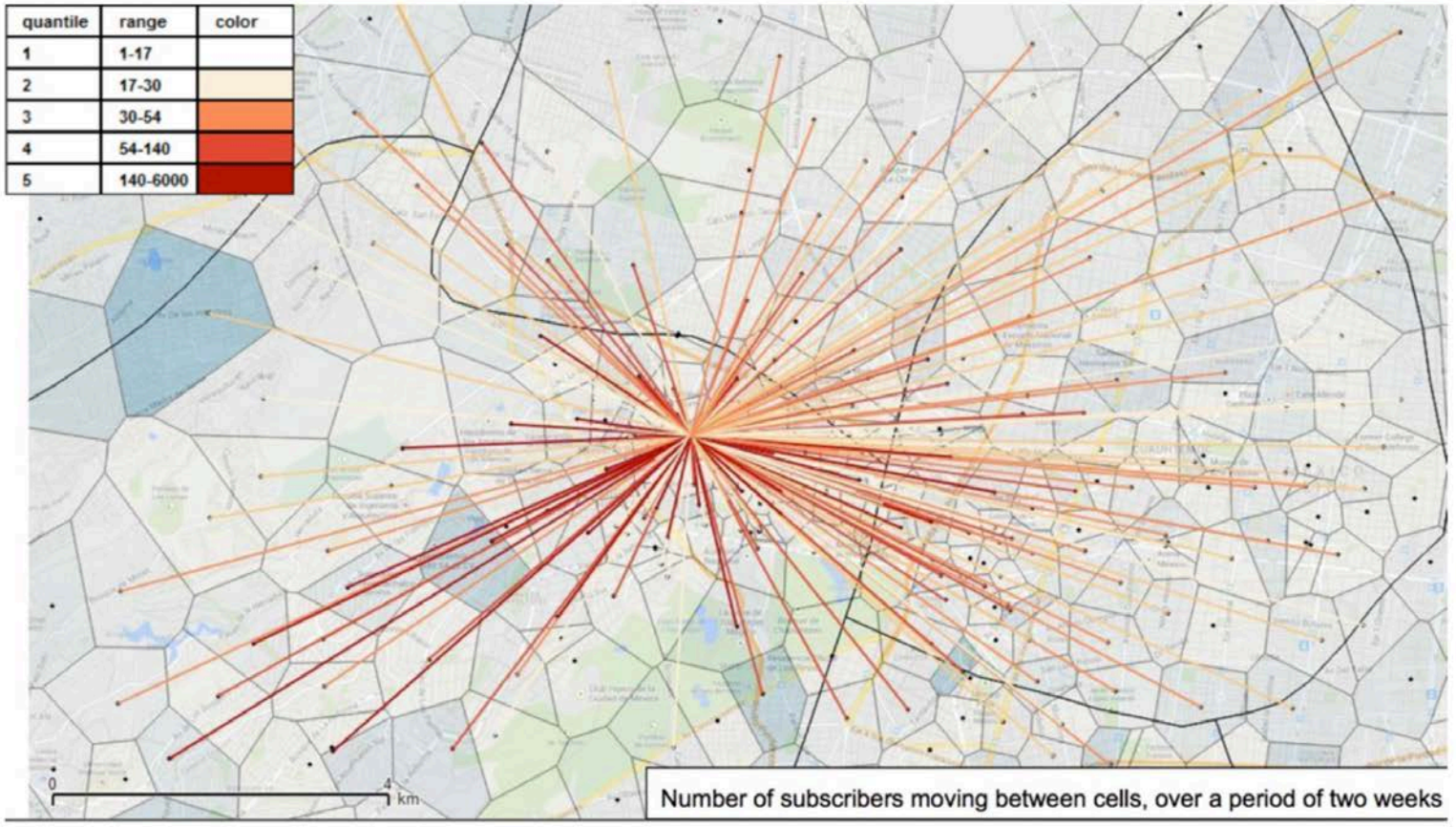
http://www.unglobalpulse.org/tabasco-floods-CDRs

Program: Privacy

Project: Data privacy risk and the limits of decision making

Partners: MIT

Mapping movements: Record which areas people move between by recording locations of successive calls



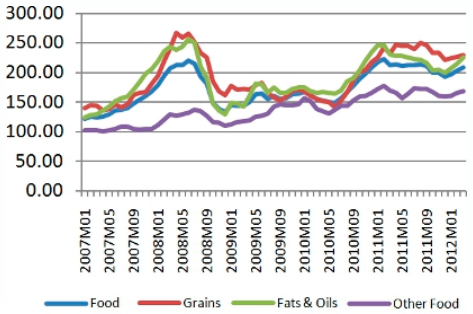
Speech-to-Text Radio Mining

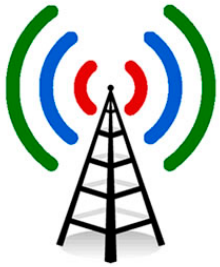
Toolkit

Speech-to-text engine + near real-time analysis of live public FM radio conversation in Uganda

Features

location, volume, mood, associated topics, anomaly detection



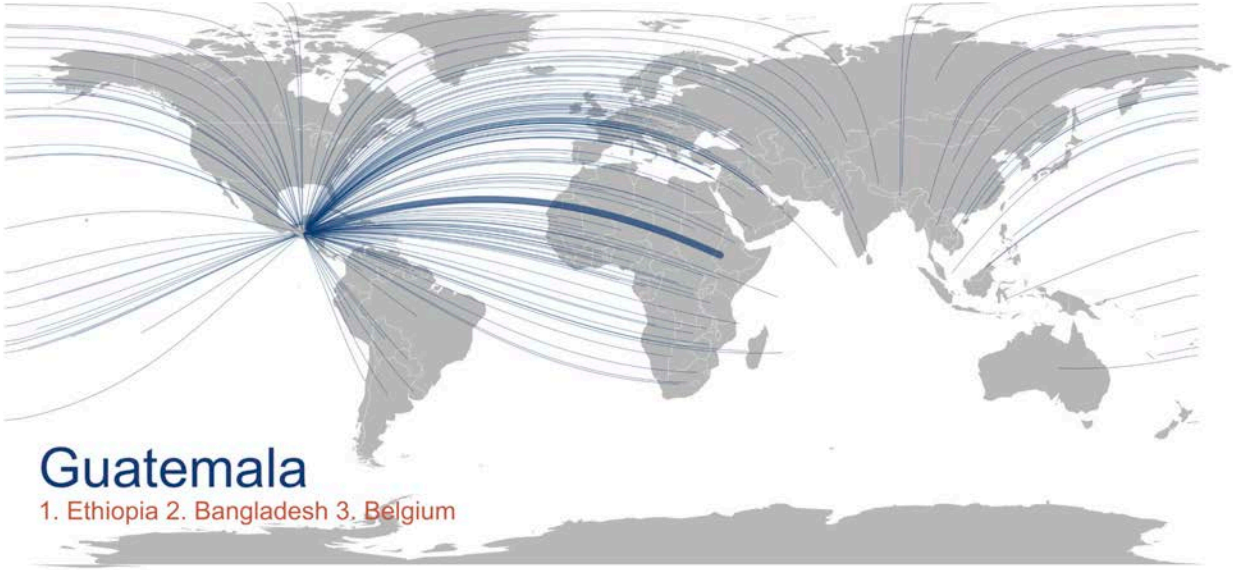


Record Audio → Identify Language (Luganda, Acholi) → Convert Speech to Text Transcript → Filter for Topics (health, food prices, jobs, crops, etc.) → Analyze volume, mood, associated topics → Visualize on dashboard

PARTNERS: Leiden University, Makerere University, Sheffield University

Postal Flows

UPU UNIVERSAL POSTAL UNION



GLOBAL PULSE

Big Data for Development: 3 Opportunities

1. **Enhanced Early Warning:** Earlier detection of anomalies and events allows rapid response to crises.

1. **Real-Time Awareness:** Real-time trend analysis of population activities and dynamics can inform the design and targeting of programmes and policies.

1. **Real-time Evaluation:** Real-time feedback from citizens, and measurement of behavior change, allows for adaptive course corrections in programmes and policies.

GLOBAL PULSE

Advantages of using big data

New insights

New sources provide data historically unavailable, yielding new insights

Cost of data collection

Digital systems can be significantly less resource intensive than traditional statistics

Risk of data collection

Allows remote analysis, allowing data to be tracked in risky or unstable locations

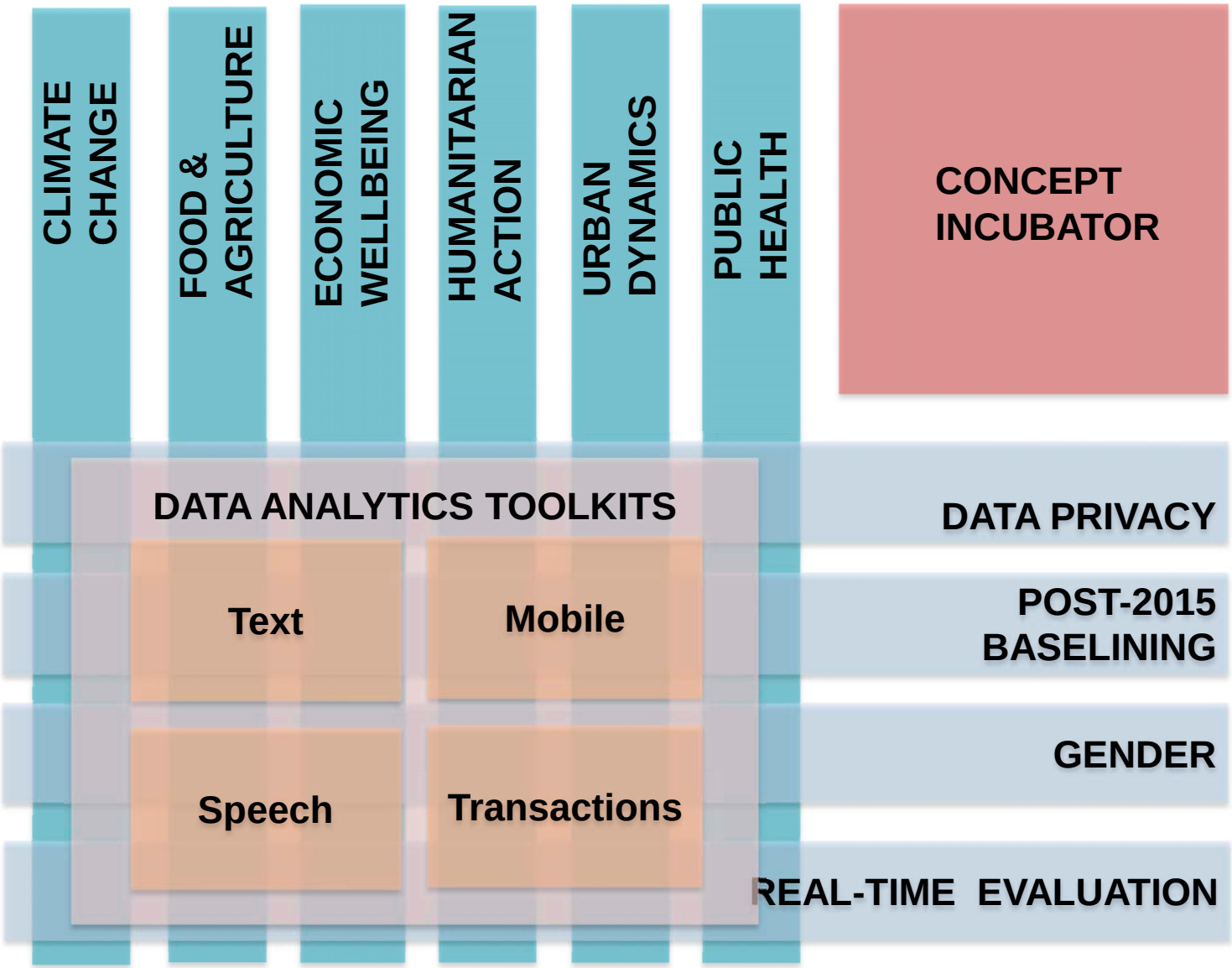
Speed of response

Response can significantly improve on lag in traditional statistics

Adaptive execution

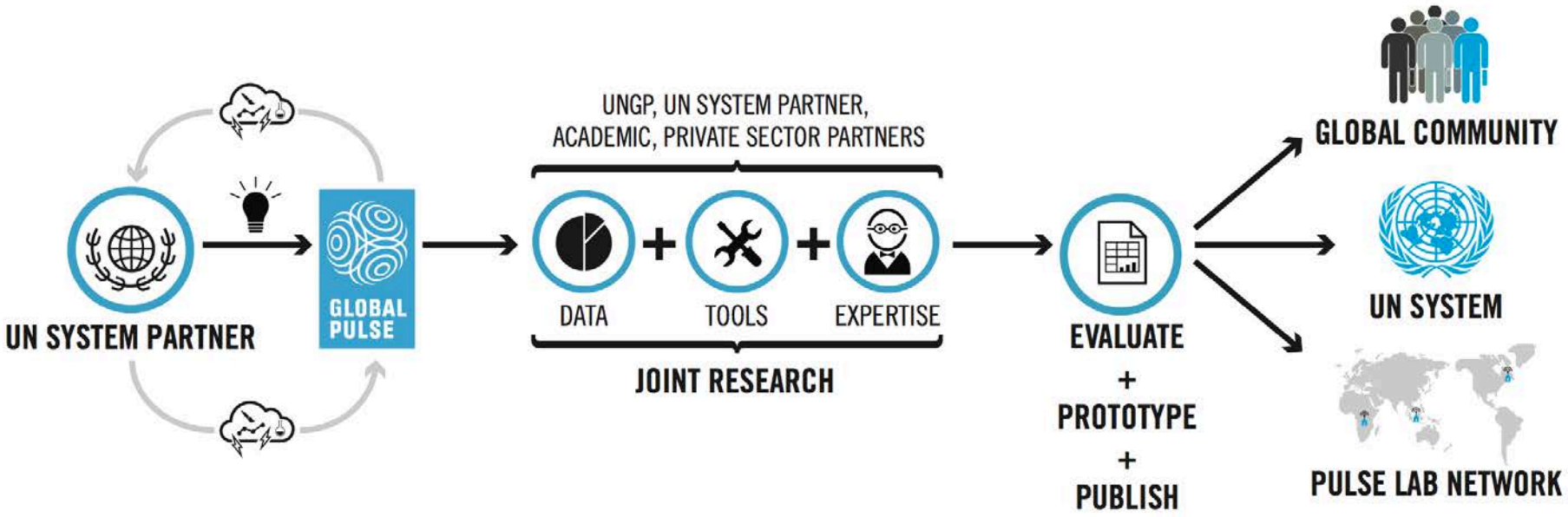
Continuous real-time feedback allows strategy to evolve with changing realities on the ground

2014-2015 Track 1 Innovation Programs



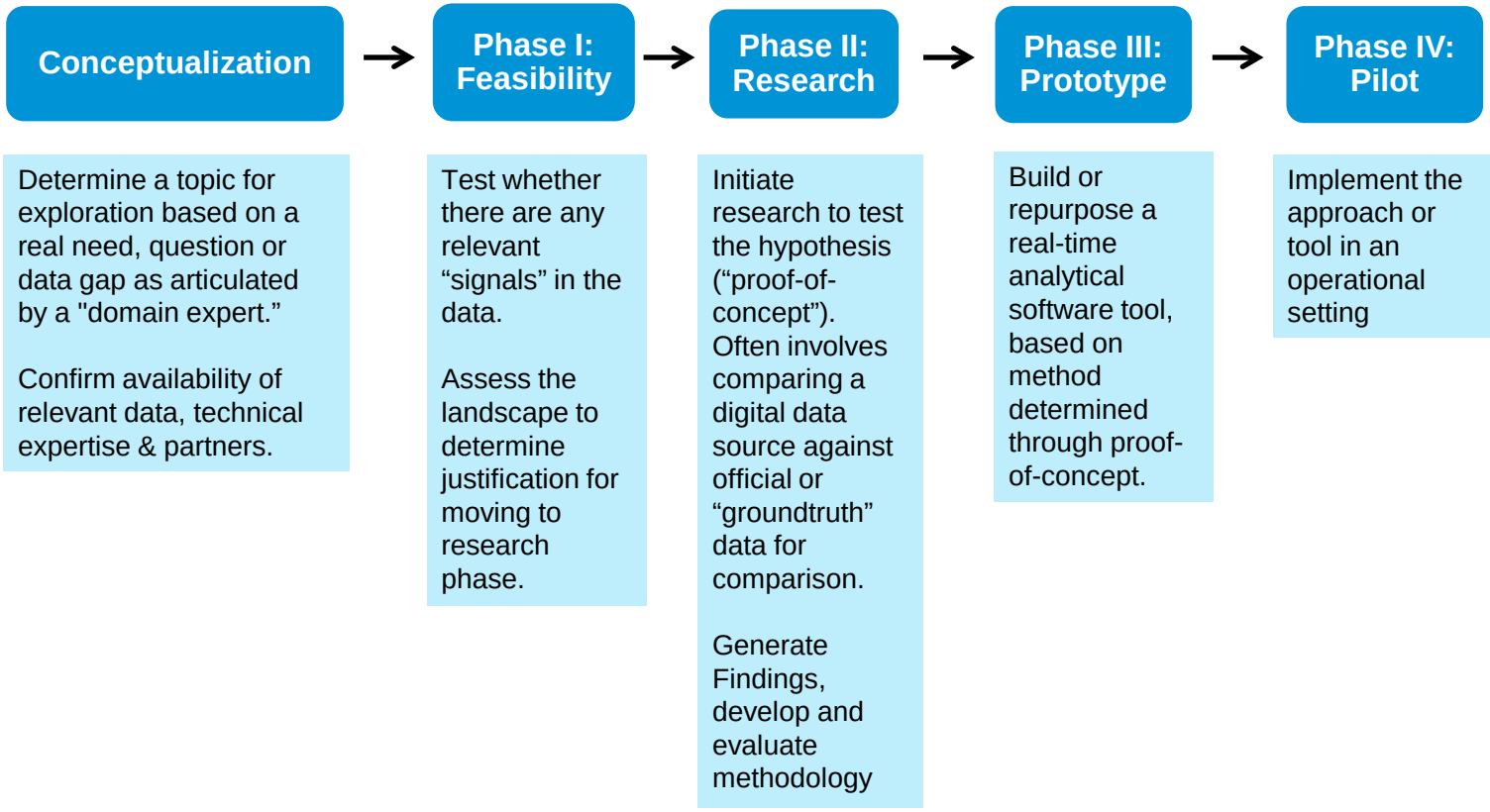
GLOBAL PULSE

INNOVATING TOGETHER:
Typical cycle of a joint project



GLOBAL PULSE

INNOVATING TOGETHER:
Typical cycle of a joint big data project



GLOBAL PULSE

You Can’t Do This Alone

“Data philanthropy” – a new form of shared value corporate social responsibility in which private sector companies participate in making big data and real-time analytics available for social good.

Big Data Access	• Twitter • Orange France Teleco • Telenor • Telefonica • Real Impact • Universal Postal Union
Data Mining & Analysis Technologies	• Amazon Web Services (supercomputing) • DataSift (data filtering) • SAS (analytics & data visualization) • Crimson Hexagon (data analysis)
Data Science Expertise	• Université catholique de Louvain • Institut des Systèmes Complexes de Paris Ile-de-France • Universidad Politécnica de Madrid • Stockholm University • Karolinska Institutet • University of Sheffield • Microsoft Research • Google

Catalyzing the formation of a new information ecosystem

GLOBAL PULSE

Washington, DC

Tuesday 10:00 AM
Scattered Clouds

29 °F | °C

Precipitation: 0%
Humidity: 30%
Wind: 19 mph

Temperature | **Precipitation** | **Wind**

Tue	Wed	Thu	Fri	Sat	Sun	Mon	Tue
36° 25°	37° 34°	46° 28°	37° 28°	50° 39°	66° 54°	66° 54°	61° 43°

Thank You!

kirkpatrick@un.org

www.unglobalpulse.org