Telcos data for development

public – private partnership for sustainable development

Zbigniew SMOREDA

Orange Labs, France

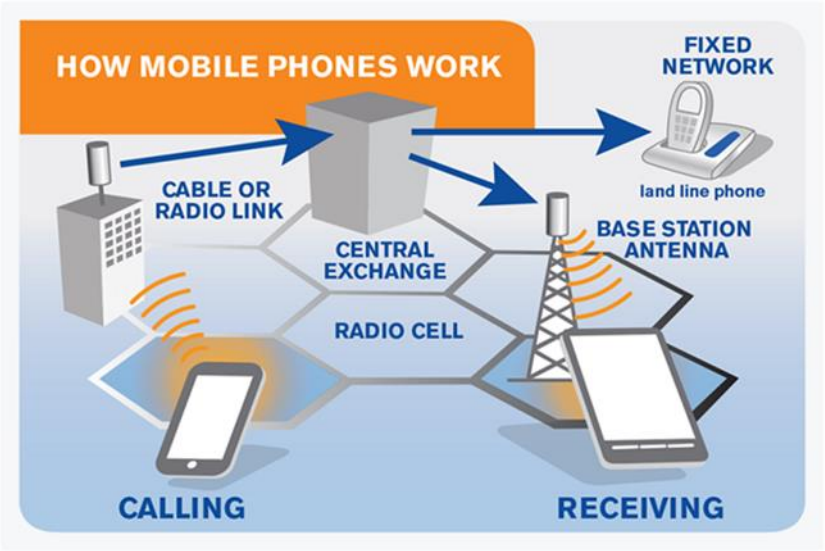
Content

- Call Detail Records (CDR) & first research trials (2006 – 2009)
- NetMob – scientific community building (2010)
- Data for Development (D4D)
  - Cote d'Ivoire (2012-2013)
  - Senegal (2014-2015)
- Open Algorithms (OPAL) project (2016)

CDRs telcos billing records

- Billing data collected automatically for all MPO's customers

timestamp	party A	party B	type	duration	cell A	cell B
16/10/30 10:01:33	0689094877	0645321101	SMS	0	123	322
16/10/30 10:01:34	0765443321	0675448765	Voice	54	233	543
16/10/30 10:01:35	0766545566	0653344567	Voice	132	435	124



CDRs telcos billing records

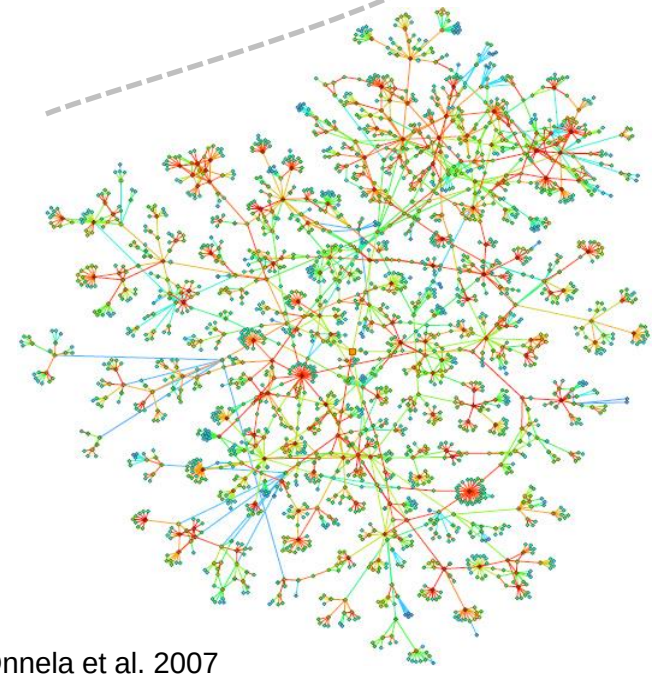
- Billing data collected automatically for all MPO's customers

timestamp	party A	party B	type	duration	cell A	cell B
16/10/30 10:01:33	ABCXRTTAAT	XYRAATRRAA	SMS	0	Lat, Lon	Lat, Lon
16/10/30 10:01:34	ZXXRTT554R	CC445EERSDA	Voice	54	Lat, Lon	Lat, Lon
16/10/30 10:01:35	776ERERREER	99TRDDAAA7	Voice	132	Lat, Lon	Lat, Lon

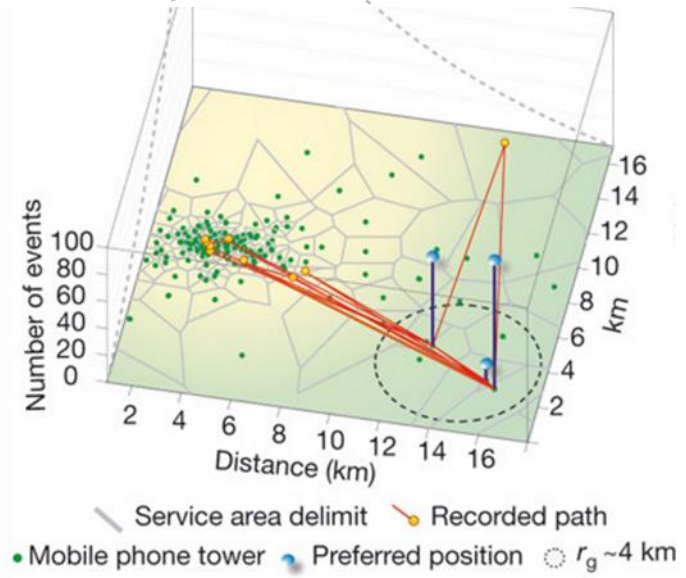
CDRs telcos billing records

- Billing data collected automatically for all MPO's customers

timestamp	party A	party B	type	duration	cell A	cell B
16/10/30 10:01:33	ABCXRTTAAT	XYRAATRRAA	SMS	0	Lat, Lon	Lat, Lon
16/10/30 10:01:34	ZXXRTT554R	CC445EERSDA	Voice	54	Lat, Lon	Lat, Lon
16/10/30 10:01:35	776EREREER	99TRDDAAA7	Voice	132	Lat, Lon	Lat, Lon



Onnela et al. 2007

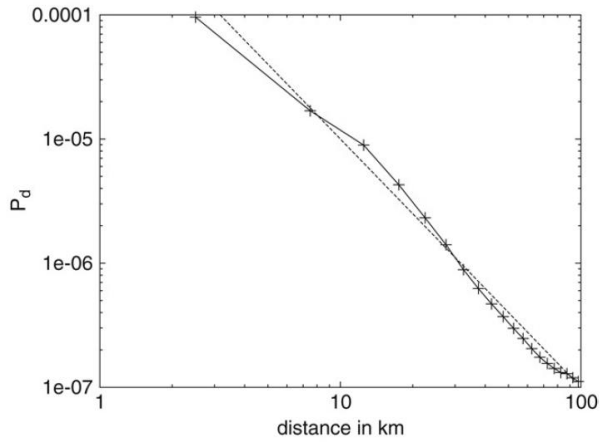


Gonzàlez et al. 2008

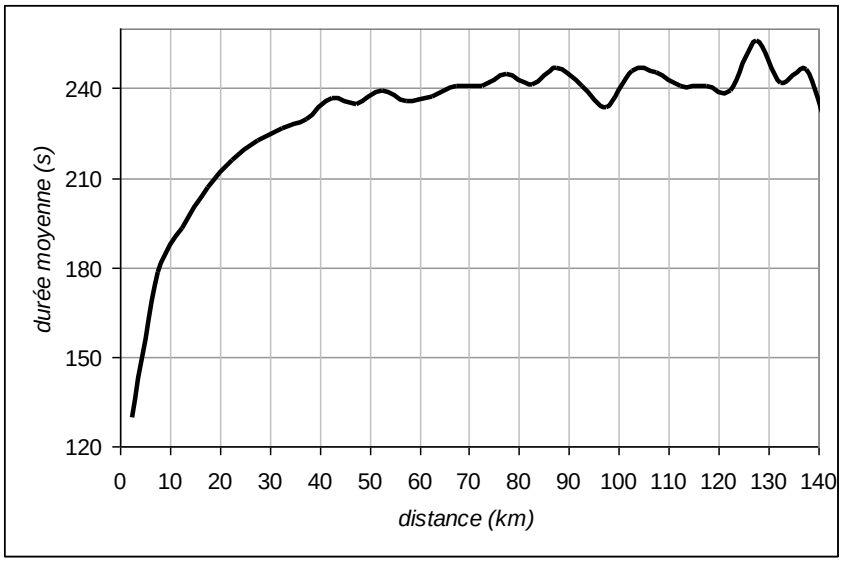
CDRs first trials

CDRs first trials

- Distance role in building social networks: towards the network geography



The probability *Pd* that two people living in Belgium at a distance *d* are connected by a mobile phone link in a log–log scale. The dashed line is the power-law d^{-2} (gravity law)

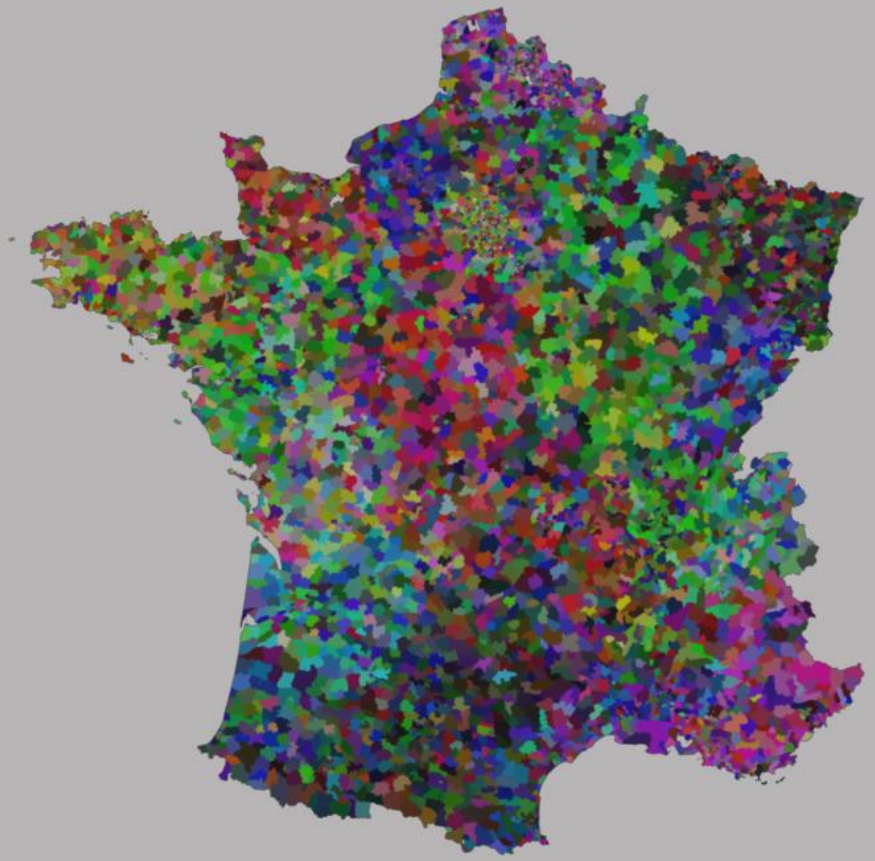


R. Lambiotte et al. 2008

CDRs first trials

Studying the structure of telephone conversations to understand how communities are spatially organized, we realized that they allow to reconstruct the French regional divisions surprisingly well.

Total coincidence or logical process?



V. Blondel et al. 2011

Quickly growing research

Cellular Census: Explorations in Urban Data Collection

Issue No. 03 - July-September (2007 vol. 6)
ISSN: 1536-1268
pp: 30-38
DOI Bookmark: http://doi.ieeecomputersociety.org/10.1109/MP

Carlo Ratti , Massachusetts Institute of Technology
Francesco Calabrese , Massachusetts Institute of Tech
Jonathan Reades , University College London
Andres Sevtsuk , Massachusetts Institute of Technology

Proceedings of the National Academy of Sciences of the United States of America

CURRENT ISSUE // ARCHIVE // NEWS & MULTIMEDIA // AUTHORS // ABOUT // COLLECTED ARTICLES // BROWSE BY

> Current Issue > vol. 106 no. 36 > Nathan Eagle, 15274–15278

Inferring friendship network structure by using mobile phone data

Nathan Eagle[a,b,1], Alex (Sandy) Pentland[b] and David Lazer[c]

Malar J. 2009 Dec 10;8:287. doi: 10.1186/1475-2875-8-287.

The use of mobile phone data for the estimation of the travel patterns and imported Plasmodium falciparum rates among Zanzibar residents.

Tatem AJ[1], Qiu Y, Smith DL, Sabot O, Ali AS, Moonen B.

Proceedings of the National Academy of Sciences of the United States of America

CURRENT ISSUE // ARCHIVE // NEWS & MULTIMEDIA // AUTHORS // ABOUT // COLLECTED ARTICLES // BROWSE BY TO

> Current Issue > vol. 104 no. 18 > J.-P. Onnela, 7332–7336

Structure and tie strengths in mobile communication networks

J.-P. Onnela [*, †, ‡], J. Saramäki [*], J. Hyvönen [*], G. Szabó [§, ¶], D. Lazer [‖], K. Kaski [*], J. Kertész [*, **], and A.-L. Barabási [§, ¶]

Activity-Aware Map: Identifying Human Daily Activity Pattern Using Mobile Phone Data

Authors | Authors and affiliations

Santi Phithakkitnukoon, Teerayut Horanont, Giusy Di Lorenzo, Ryosuke Shibasaki, Carlo Ratti

Community Computing: Comparisons between Rural and Urban Societies Using Mobile Phone Data

Authors: Nathan Eagle
Yves-Alexandre de Montjoye
Luis M. A. Bettencourt

2009 Article

Published in:
· Proceeding
CSE '09 Proceedings of the 2009 International Conference on Computational Science and Engineering - Volume 04
Pages 144-150

August 29 - 31, 2009
IEEE Computer Society Washington, DC, USA ©2009
table of contents ISBN: 978-0-7695-3823-5 doi>10.1109/CSE.2009.91

Bibliometrics
· Citation Count: 8
Downloads (cumulative): 0
Downloads (12 Months): 0
Downloads (6 Weeks): 0

Letter

Nature **453**, 779-782 (5 June 2008) | doi:10.1038/nature06958; Received 1 March 2008

There is an Addendum (12 March 2009) associated with this document.

Understanding individual human mobility patterns

Marta C. González[1], César A. Hidalgo[1,2] & Albert-László Barabási[1,2,3]

ELSEVIER

Physica A: Statistical Mechanics and its Applications

Volume 387, Issue 12, 1 May 2008, Pages 3017–3024

The dynamics of a mobile phone network

Cesar A. Hidalgo[a], C. Rodriguez-Sickert[b]

ELSEVIER

Physica A: Statistical Mechanics and its Applications

Volume 387, Issue 21, 1 September 2008, Pages 5317–5325

Geographical dispersal of mobile communication networks

Renaud Lambiotte[a, b], Vincent D. Blondel[a], Cristobald de Kerchove[a], Etienne Huens[a], Christophe Prieur[c], Zbigniew Smoreda[c], Paul Van Dooren[a]

NetMob – community building

NetMob

Workshop on the
Analysis of Mobile Phone Networks

A satellite workshop to NetSci 2010
Tuesday, May 11, 2010
MIT, Cambridge, MA

NetMob2011

Given the success of NetMob2010, we consider the possibility of organizing a **NetMob2011**. If you wish to be included on the NetMob mailing list, please send an email to sympa2@listes.uclouvain.be with "subscribe netmob yourname" in the subject line (where "yourname" is your first and last name). You can also subscribe/unsubscribe by going to https://listes-2.sipr.ucl.ac.be/sympa/info/netmob.

Introduction

Mobile phone datasets have become widely available in recent years and have opened the possibility to improve our understanding of large-scale social networks by investigating how people exchange information, build trust, create markets and develop social interactions. Mobile phone data is also helping us understand complex processes such as the spread of information and viruses or transportation and the use of urban infrastructures.

This workshop will consist of a number of contributed talks on the analysis of mobile phone networks. The workshop format is flexible: no registration fees, a simplified submission procedure, and the possibility to present recent results or results submitted elsewhere.

Practical information

Date: Tuesday May 11, 2010 (this is the day prior to the conference NetSci).

Location: On the sixth floor of the newly built Media Lab (building E14 on MIT campus, map available here).

Registration: Attendance is free of charge but, due to limited seating, registration is compulsory. If you wish to register please send an email to netmob@uclouvain.be. Registration will be processed on a first-come first-serve basis. Although there is no registration fee for the workshop, participants are of course encouraged to also participate (and register) in the NetSci conference.

We have have received an unexectedly **large number of registrations** to the workshop. The workshop has been moved to a larger space (the multi media hall of the Media Lab). All those who have registered by sending an email or through the NetSci website are welcome to attend.

Submissions

All contributions that deal with the analysis of mobile phone datasets are welcome.

Authors are invited to submit an abstract (one to three pages) by the deadline of March 5, 2010. Submissions should include the title, author(s), affiliation(s) and e-mail address(es) on the first page. There will be no published proceedings; the material submitted to the workshop may also be submitted elsewhere.

Electronic submission of manuscripts in PDF format is required. Please send your manuscript directly to netmob@uclouvain.be by March 5, 2010.

The evaluation of submitted abstracts will be organized by the scientific committee and decisions will be made by March 26, 2010. Once an abstract has been accepted for presentation, at least one author is required to attend the workshop and present the paper. In case too many abstracts are selected, some of these may be moved to a special session taking place the next day at the NetSci 2010 conference.

Program

The **program** is available here (PDF format).

Book of abstracts

The **book of abstracts** is available here (5.5 MB, PDF format).

NetMob – community building

NetMob

Workshop on the
Analysis of Mobile Phone Networks

A satellite workshop to NetSci 2010
Tuesday, May 11, 2010
MIT, Cambridge, MA

NetMob2011

Given the success of NetMob2010, we consider the possibility of organizing a **NetMob2011**. If you wish to be included on the NetMob mailing list, please send an email to sympa2@listes.uclouvain.be with "subscribe netmob yourname" in the subject line (where "yourname" is your first and last name). You can also subscribe/unsubscribe by going to https://listes-2.sipr.ucl.ac.be/sympa/info/netmob.

Introduction

Practical information

Submissions

Program

Book of abstracts

Scientific committee

Chair: Vincent Blondel, UCLouvain (Belgium)
Laszlo Barabasi, Northeastern University
Rob Claxton, British Telecom (UK)
Vittoria Colizza, ISI Foundation (Italy)
Massimo Colonna, Telecom Italia (Italy)
Nathan Eagle, Santa Fe Institute
Alexandre Gerber, AT&T Research
Marta Gonzales, MIT
Cesar Hidalgo, Harvard University
János Kertész, Budapest University of Technology (Hungary)
Renaud Lambiotte, Imperial College (UK)
David Lazer, Northeastern University
Jure Leskovec, Stanford University
Nuria Oliver, Telefonica Research (Spain)
Jukka-Pekka Onnela, Harvard University
Asu Ozdaglar, LIDS, MIT
Alex (Sandy) Pentland, Media Lab, MIT
Mason Porter, University of Oxford (UK)
Carlo Ratti, Senseable City Lab, MIT
Jari Saramäki, Helsinki University of Technology (Finland)
Leonardo Soto, AirSage
Zbigniew Smoreda, Orange Labs (France)
John Tsitsiklis, LIDS, MIT
Paul Van Dooren, UCLouvain (Belgium)

Organizing committee

Vincent Blondel, UCLouvain (Belgium)
Francesco Calabrese, Senseable City Lab, MIT
Gautier Krings, UCLouvain (Belgium)
Benjamin Waber, Media Lab, MIT

cial networks by investigating
nplex processes such as the

fees, a simplified submission

@uclouvain.be. Registration
o participate (and register) in

media hall of the Media Lab).

on(s) and e-mail address(es)

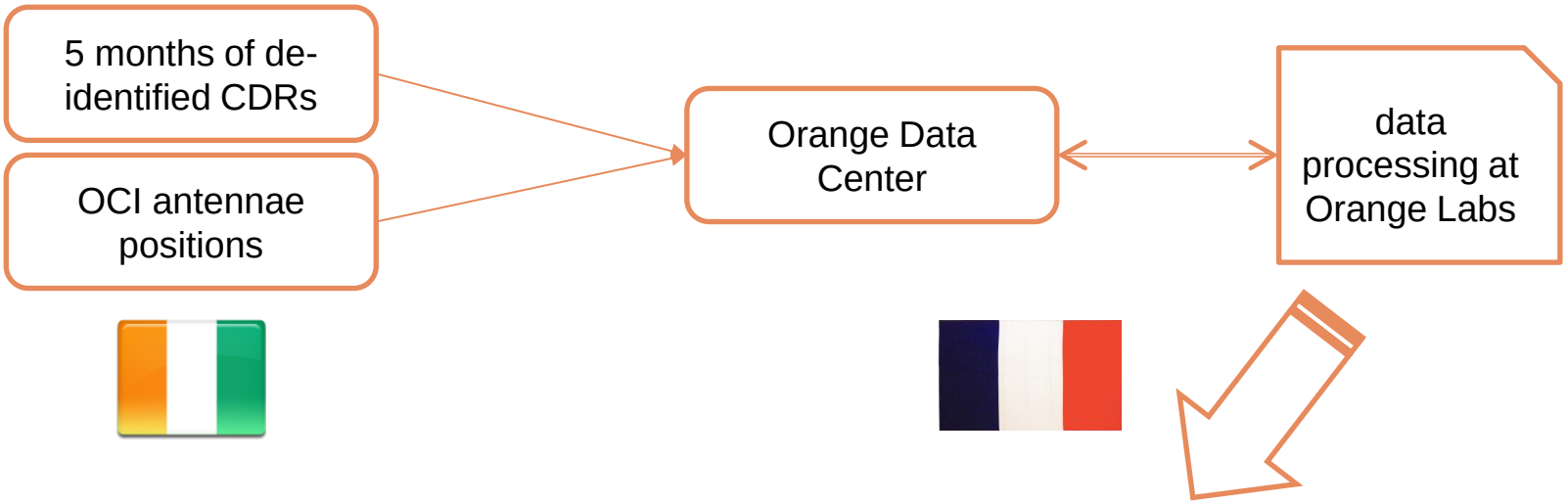
tract has been accepted for
may be moved to a special

How to accelerate?

An Open Innovation Project
with Orange Cote d'Ivoire,
Orange Marketing Vision
and
Corporate Social Responsibility

Data for Development (D4D) Cote d'Ivoire 2012

D4D: data for development Cote d'Ivoire



- Proposed datasets:
  - (1) antenna-to-antenna traffic on an hourly basis,
  - (2) individual trajectories for 50,000 randomly sampled users for two week time windows with antenna location information,
  - (3) individual trajectories for 500,000 randomly sampled users over the entire observation period with sub-prefecture location information,
  - and (4) a sample of communication graphs for 5,000 customers

D4D: data for development Cote d'Ivoire

- "Scientific challenge" - halfway between a long hackathon and a scientific conference:
  - Launched in June 2012 for 8 months only
  - Only research institutions admitted after signing the terms & conditions
  - Evaluation committee chaired by Vincent Blondel (UCL) with members from Bouake University, Global Pulse (UN), GSMA, Orange Labs, WEF and MIT
  - Selected projects presentation during the NetMob conference, May 2013 in Boston (USA)
  - Four prizes attributed (first prize, scientific, development, and data visualization prizes)

D4D: data for developme

- Côte d'Ivoire CDR samples
- Open to the researchers
- 263 projects submitted
- Over 80 reports received

D4D: data for development Cote d'Ivoire

NetMob 2013

May 1-3, 2013, MIT

> *Third conference on the Analysis of Mobile Phone Datasets*

With a special event on the Data for Development (D4D) challenge.

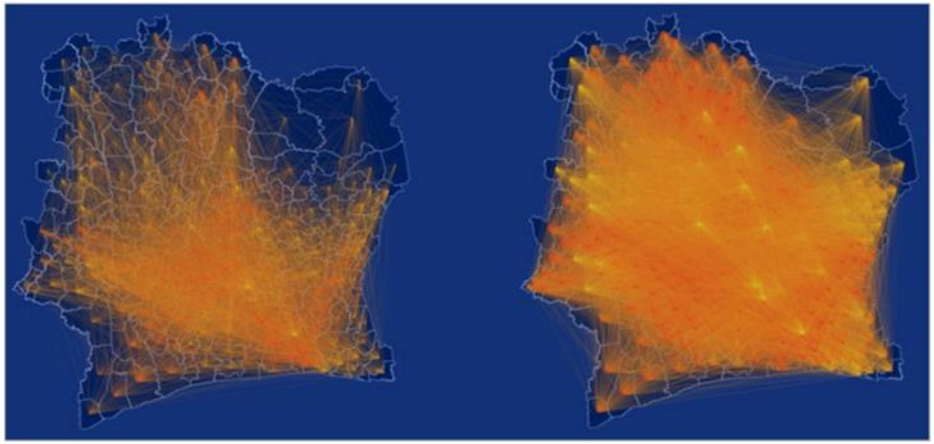
"I can't say how impressed I was with the quality and quantity of the submissions."
"There is some incredibly good work in here."

Anonymous reviewers

http://perso.uclouvain.be/vincent.blondel/netmob/2013/D4D-book.pdf

D4D: data for development Cote d'Ivoire

- First prize



Geographic network obtained from mobility traces (**a**) and call logs (**b**), where nodes represent sub-prefectures. This map was generated by a custom d3 script. Map data: © OpenStreetMap contributors, available under the Open Database License.

SCIENTIFIC REPORTS

Altmetric: 6 Views: 2,318 Citations: 6 More detail »

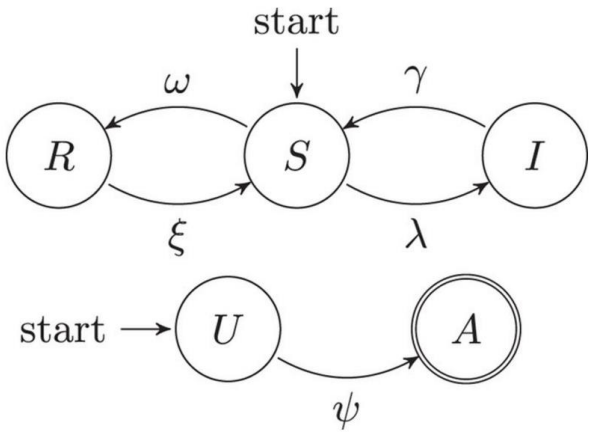
Article | OPEN

Disease Containment Strategies based on Mobility and Information Dissemination

A. Lima, M. De Domenico, V. Pejovic & M. Musolesi

Scientific Reports **5**, Article number: 10650 (2015)
doi:10.1038/srep10650

Received: 20 October 2014
Accepted: 24 April 2015
Published online: 02 June 2015



State machines describing the state transitions of a person with respect to the disease contagion (R = Resistant, S = Susceptible and I = Infected) and with respect to the information spreading (U = unaware, A = aware), respectively. A person starts in the susceptible and unaware states. We assume that aware individuals spread the information and cannot go back to the unaware state.

D4D: data for development Cote d'Ivoire

- Development prize

Joint European Conference on Machine Learning and Knowledge Discovery in Databases

ECML PKDD 2013: Machine Learning and Knowledge Discovery in Databases pp 663-666

AllAboard: A System for Exploring Urban Mobility and Optimizing Public Transport Using Cellphone Data

Authors | Authors and affiliations

Michele Berlingerio, Francesco Calabrese, Giusy Di Lorenzo, Rahul Nair, Fabio Pinelli, Marco Luca Sbodio

D4D: data for development Cote d'Ivoire

AllAboard applied to Abidjan, Ivory Coast

Input:

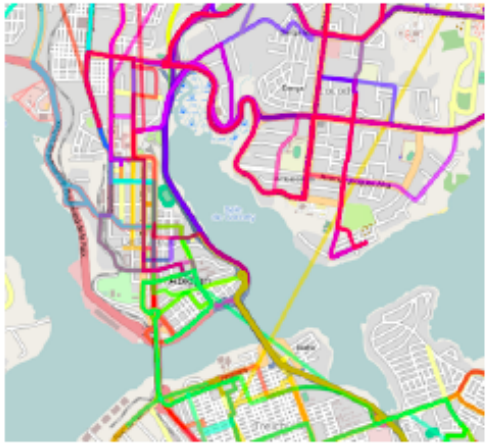
- Cellphone location data from 500,000 users in Ivory Coast(1)
- Existing transit network consisting of 17 express bus routes and 67 regular bus routes

Pre-processing

- Cellphone data was able to estimate 15,000 time-varying Origin-Destination flows
- 30 frequent location sequences were extracted

Output

- Optimization model selected 4 new routes
- 22 routes had increased ridership
- Additional met demand that resulted in citywide travel time decreased by 10%



(1) Data from Orange, under the D4D Challenge

SmarterCities Technology Centre

D4D: data for development Cote d'I

+ 50 published works!

SCIENTIFIC REPORTS

Altmetric: 7 Views: 1,782 Citations: 10 More

Article | OPEN

Inferring human mobility using communication patterns

Vasyl Palchykov, Marija Mitrović, Hang-Hyun Jo, Jari Saramäki & Raj Kumar Pan

Scientific Reports **4**, Article number: 6174

Received: 01 May
Accepted: 06 Aug

SCIENTIFIC REPORTS

Altmetric: 6 Views: 2,318 Citations: 6 More detail »

Article | OPEN

Disease Containment Strategies based on Mobility and Information Dissemination

A. Lima, M. De Domenico, V. Pejovic & M. Musolesi

Scientific Reports **5**, Article number: 10650 (2015)
doi:10.1038/srep10650

Received: 20 October 2014
Accepted: 24 April 2015
Published online: 02 June 2015

PLOS ONE | Subject Areas | Publish | About

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

Mobile Phone Call Data as a Regional Socio-Economic Proxy Indicator

Sanja Šćepanović, Igor Mishkovski, Pan Hui, Jukka K. Nurminen, Antti Ylä-Jääski

Published: April 21, 2015 • DOI: 10.1371/journal.pone.0124160

EPJ Data Science

About | Articles | Submission Guidelines

IMPACT FACTOR 1.567

REGULAR ARTICLE OPEN ACCESS

The impact of social segregation on human mobility in developing and industrialized regions

Alexander Amini, Kevin Kung, Chaogui Kang, Stanislav Sobolevsky and Carlo Ratti

EPJ Data Science 2014, 3:6 DOI: 10.1140/epjds31 © Amini et al.; licensee Springer 2014

Received: 25 January 2014 Accepted: 22 May 2014 Published: 6 June 2014

Salnikov et al. *EPJ Data Science* 2014, 3:3
http://www.epjdatascience.com/content/3/1/3

EPJ.org

REGULAR ARTICLE

The geography and carbon footprint of mobile phone use in Cote d'Ivoire

Vsevolod Salnikov, Daniel Schien, Hyejin Youn, Renaud Lambiotte and Michael T Gastner

ELSEVIER

Computer Communications

Volume 59, 15 March 2015, Pages 1–11

Large scale model for information dissemination with device to device communication using call details records

Rachit Agarwal, Vincent Gauthier, Monique Becker, Thouraya Toukabrigunes, Hossam Afifi

Andris and Bettencourt *Infrastructure Complexity* 2014, 1:1
http://www.infrastructure-complexity.com/content/1/1/1

Infrastructure Complexity
a SpringerOpen Journal

RESEARCH Open Access

Development, information and social connectivity in Côte d'Ivoire

Clio Andris and Luis MA Bettencourt

orange

D4D: data for development Cote d'Ivoire

- Many various topics addressed by the projects:
  - Health improvement (disease spread mapping or prevention)
  - Population statistics (urbanization, population, tourism and events analysis)
  - Communities understanding (diaspora cartography and needs, rural and urban customers)
  - Economic Indicators (local economic development, micro finance insight)
  - City and transport planning (transport optimization, road construction, smart city planning)
  - Emergency, Alerting & Preventing (early warning system, help distribution localization)
  - Geo-marketing (strategic points of sales)…
- Important echo among the United Nations, NGOs, development aid institutions
- New data challenges initiatives (Telecom Italia, Telefonica)
- But no project could be implemented in Ivory Cost…

Second D4D Senegal:

do it differently

sonatel

orange™

D4D Senegal

6 January 2014: Sonatel decides to launch D4D Senegal

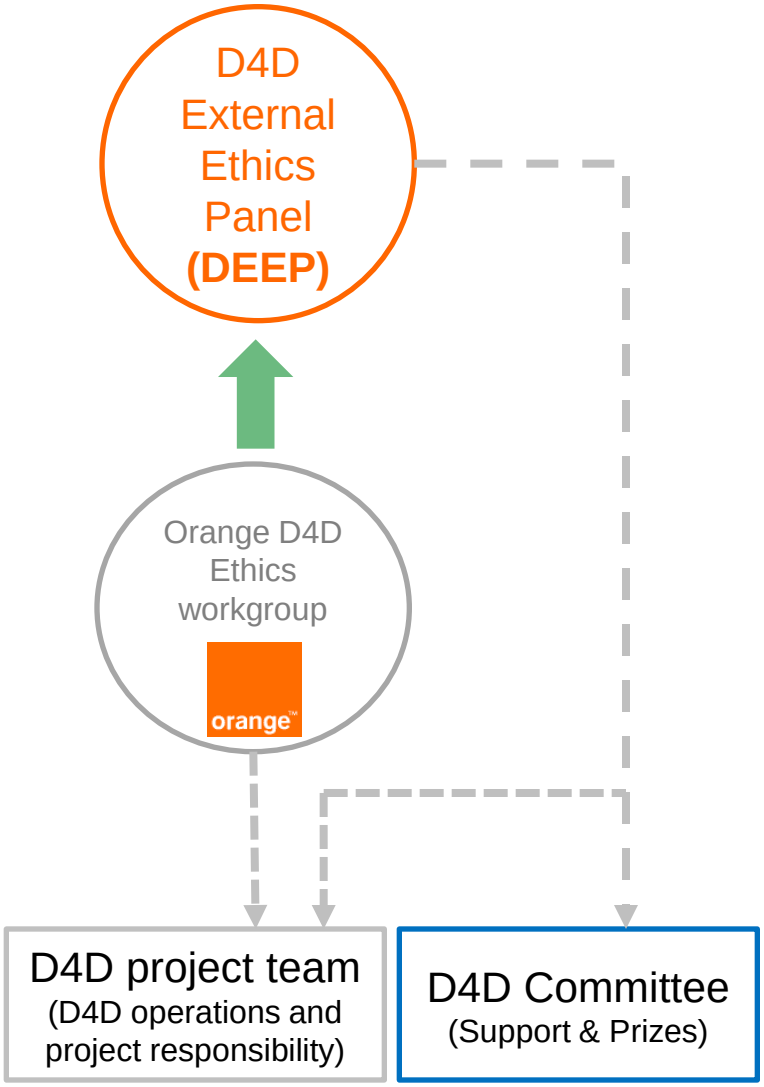
sonatel

D4D Senegal

- Improvements to bring to D4D after the Cote d'Ivoire experience

1. Focus on 5 themes
   - Health, Transport, Agriculture, Energy, National statistics
2. Involve local ecosystem
   - Questions owners: Ministries and institutions
   - Contributors: Universities, entrepreneurs…
3. Reinforce governance
   - Regulation, Ethics,…
4. Foster Data sharing
   - Find other data sources to be crossed with D4D data sets
5. Ensure an implementation of results
   - Prepare the « After D4D »

A process for Ethics review, involving both Orange staff and external experts set-up



D4D External Ethics Panel

- External advisors with balanced profile
- Provide advise/perspective to the D4D project team and the D4D Committee

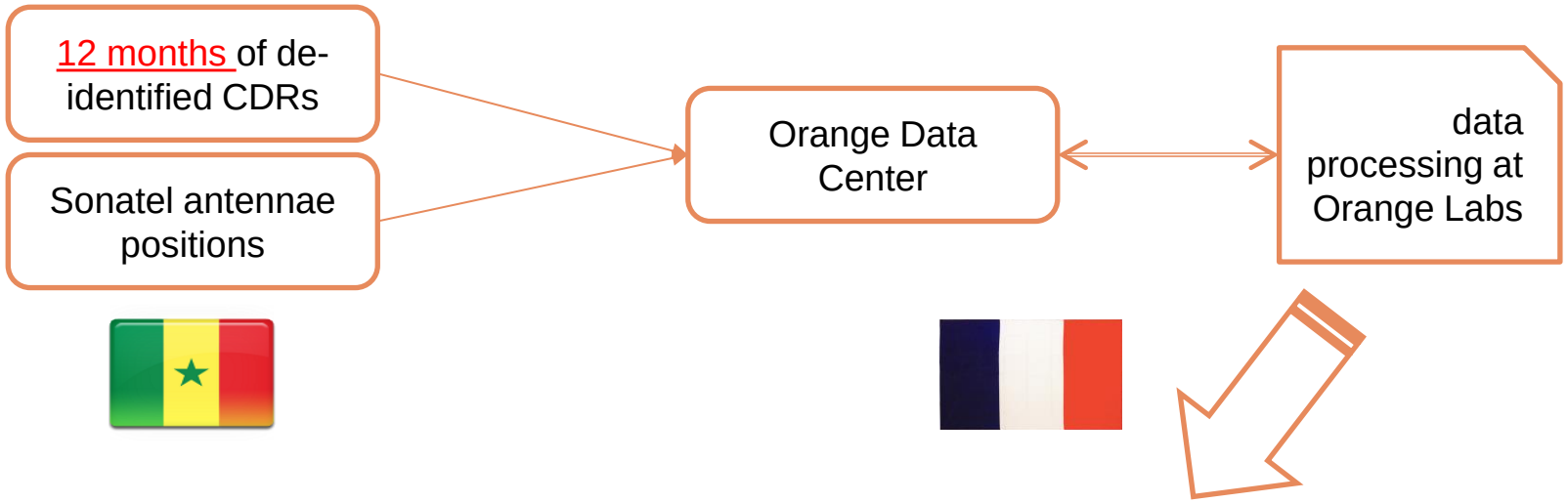
Orange D4D Ethics workgroup (internal)

- Senior managers, most of them not involved in D4D
- Propose the Sonatel-Orange preliminary position
- Initiate actions in case of obvious need

D4D Committee (external)

- 13 external members
- Advisor to D4D team and independent Prize allocation

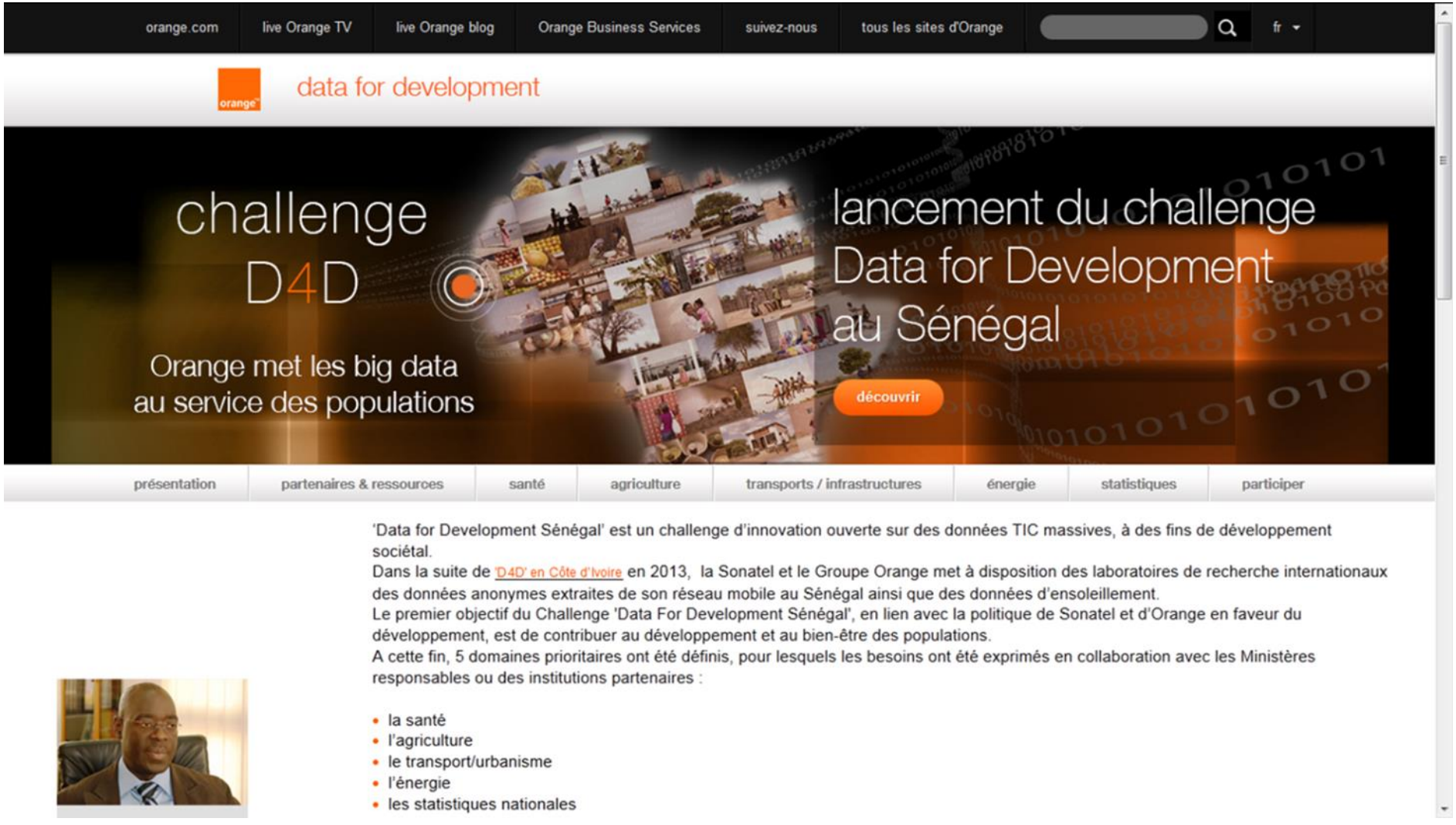
D4D Senegal



- Proposed datasets:
  - (1) antenna-to-antenna traffic on an hourly basis,
  - (2) individual trajectories for 300,000 randomly sampled users for two week time windows with antenna location information,
  - (3) individual trajectories for 150,000 randomly sampled users over the entire observation period with arrondissement location information,
  - ~~and (4) a sample of communication graphs for 5,000 customers~~
  - Bandicoot individual indicators with (2) & (3) http://bandicoot.mit.edu/

D4D Senegal

April 2014: challenge launched



D4D Senegal

April to August: more data and resources from donors

The Global Fund
To Fight AIDS, Tuberculosis and Malaria

- provided socio-demographic data on HIV, Malaria and Tuberculosis and the related expenditures
- Data are available on the website

TERALAB

- provides computing power to 5 teams
- Email sent to the teams that have submitted a project
- The opportunity is visible on the website

INSTITUTE FOR DISEASE MODELING
INTELLECTUAL VENTURES Laboratory

- provided a 1 km gridded estimate of rainfall, relative humidity and mean temperature measures
- Intellectual Ventures provided us several links for the resources
- Email sent to all the teams

DigitalGlobe

- will provide satellite images in selected regions
- That will be integrated to maps by OpenStreetMap
- OpenStreetMap will enrich the maps with data asked for by the teams
- Email sent to the teams that have submitted a project

ANSD
Agence Nationale de la Statistique et de la Démographie

- provided its databases, publications and reports
- Link on the website to these resources

CSE

- provided its databases and documents bases
- Link on the website to these resources

MINISTÈRE DE LA SANTÉ ET DE L'ACTION SOCIALE

- provided its databases and reports
- Link on the website to these resources

Open Knowledge Foundation

- OSM is detailing chosen areas with field works
- Link on the website to these resources

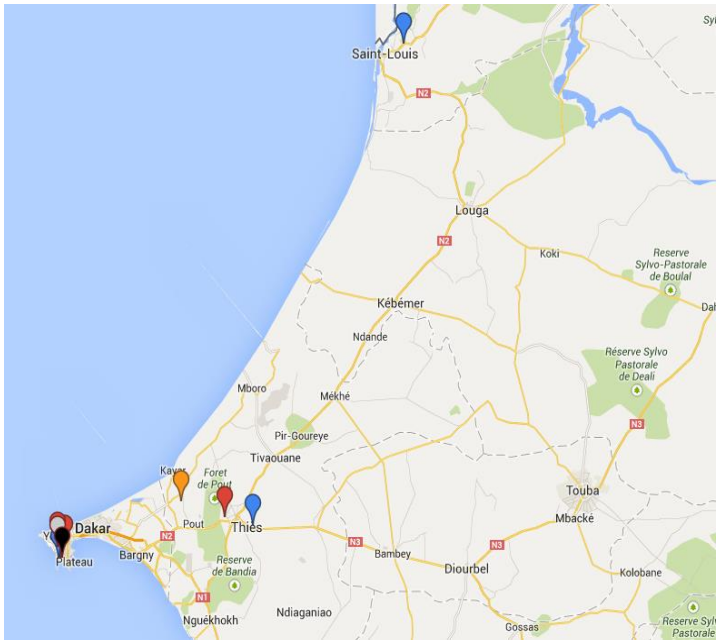
ESRI France
L'avantage géographique

- provide vouchers to use their services

D4D Senegal - June 2014: Launch in Dakar

D4D Senegal

Projects proposals came from all over the world,
11 from Senegalese universities



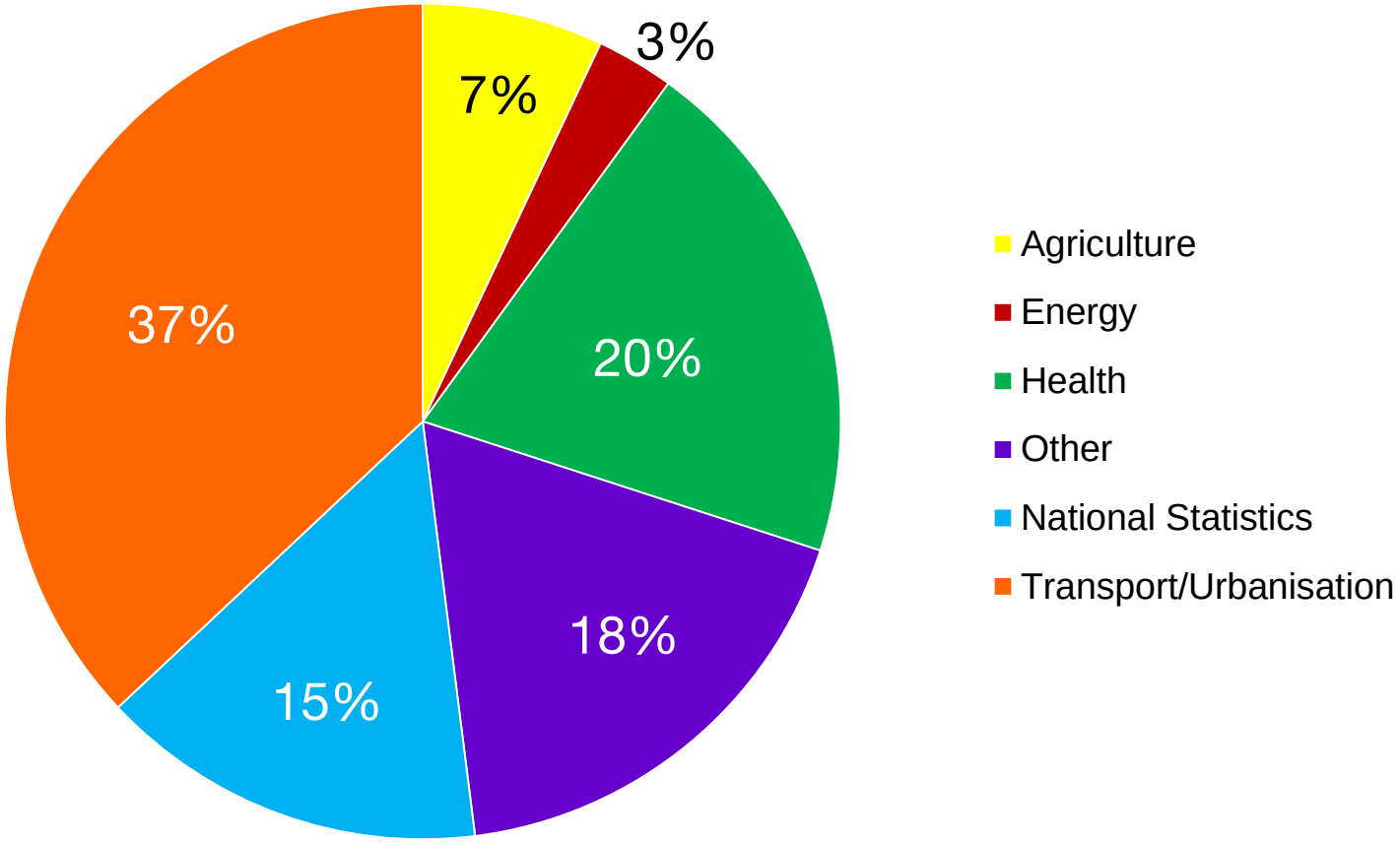
Statistics Health Energy Agriculture Transport/Urban Data Visualization Others

D4D Senegal

About 60 high quality submissions on time for the Challenge

D4D Senegal

All themes are represented, with a majority of Transports, Health and National Statistics



sonatel

orange

D4D Senegal results announcement

NetMob 2015 Program

7-10 April 2015
MIT MediaLab
http://netmob.org

School // Conference // D4D Challenge

Editors: Esteban MORO, Yves-Alexandre de MONTJOYE, Vincent BLONDEL, Alex 'Sandy' PENTLAND, Nicolas DE CORDES

Organized by

Universidad Carlos III de Madrid

MIT

UCL Université catholique de Louvain

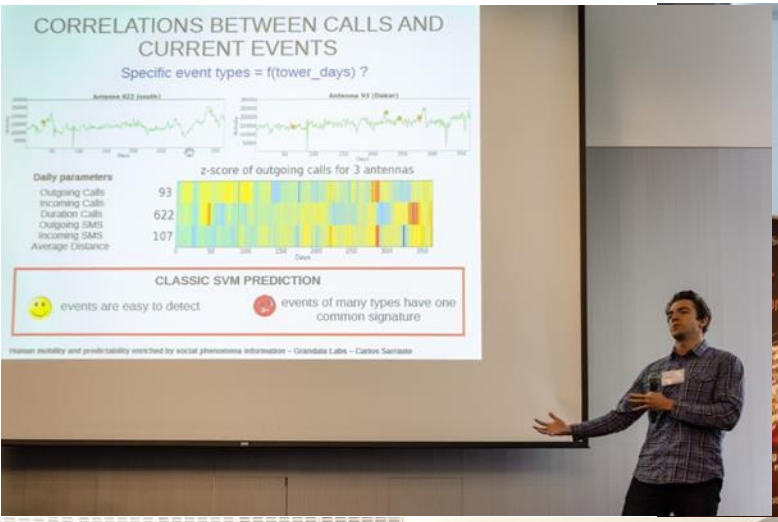
Sponsored by

orange

RealImpact ANALYTICS

orange

The event mixed presentations, workshops, datathon social events and the announcement of the winners



the winners

First Prize and Energy Prize: Using mobile phone data for electrification planning

E.A. Martínez-Ceseña [1], P. Mancarella [1], M. Ndiaye [2], and M. Schläpfer [3]

Knowledge of local energy needs is crucial for the electricity infrastructure planning of a country. We have shown that mobile phone data are an accurate proxy of the energy needs and can be used to develop bottom-up demand models. The new methodology supports and prioritizes the electrification plans in areas with scarce information on local activities and energy consumption.

(1)University of Manchester, UK - (2) Ecole supérieure polytechnique de Dakar UCAD, Senegal - (3) Santa Fe Institute, USA

Agriculture Prize: Genesis of millet prices in Senegal: the role of production, markets and their failures

D.C. Jacques [1], R. d'Andrimont [1], J. Radoux (1), F. Waldner [1], and E. Marinho [2]

Information asymmetries are responsible for price differentials in only the few areas where the mobile phone coverage has not yet reached its full potential, which damages both poor producers and food insecure consumers. To address this issue, we have integrated it in a spatially explicit model that simulates the functioning of agricultural markets.

(1) Earth and Life Institute, Université Catholique de Louvain, Belgium - (2) Independent researcher, Rio de Janeiro, Brazil

Health Prize: Uncovering the impact of human mobility on schistosomiasis...

L. Mari [1], R. Casagrandi [1], M. Ciddio [1], S.H. Sokolow [2], G. De Leo [2], and M. Gatto [1]

Schistosomiasis is water based parasitic worm infection with debilitating symptoms affecting millions of people. We show that a relatively simple model can reliably reproduce regional patterns of schistosomiasis prevalence across the country. We use the model to study the role of human mobility on disease dynamics and to analyze intervention strategies aimed at reducing disease burden.

(1)Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Italy - (2) Hopkins Marine Station, Stanford University, USA

National Statistics Prize: Virtual Networks and Poverty Analysis in Senegal

N. Pokhriyal, W. Dong, and V. Govindaraju
Computer Science and Engineering, State University of New York at Buffalo, USA

Poverty is a complex phenomenon, but can be approximated by observing mobile phone usages and mobility at regional level and extrapolated at more granular level. Poverty maps showcasing multiple perspectives can provide policymakers with better insights for effective responses for poverty eradication.

Transport Prize: National and Regional Road Network Optimization for Senegal Using Mobile Phone Data

Y. Wang [1], G. Homem de Almeida Correia [1], and Erik de Romph [1,2]

Anonymous mobile phone traces can be filtered with an algorithm to generate a proxy for a trip origin-destination matrix. This is used to develop a gravity model that predicts the future mobility in the country dependent on travel time and number of calls and messages between the departments. This information is then used to improve decision making for road network planning.

(1) Department of Transport and Planning, Delft University of Technology, The Netherlands - (2) DAT.mobility, The Netherlands

Data Crossing Prize: Using mobile phone data for Spatial Planning simulation and Optimization Technologies (SPOT)

S. Gueye [1], B.M. Ndiaye [3], D. Josselin [3], M. Poss [5], R.M. Faye [2], P. Michelon [1], C. Genre-Grandpierre [3], and F. Ciari [4]

We propose a methodology of location and relocation of amenities (home, shop, work, leisure places) for urban planning decision. Our methodology exploits mobile phone data and other variables and point of interest on maps to propose optimal amenity locations to reduce the overall travel time or travel distance.

(1) LIA, Université d'Avignon, France - (2) LTI, ESP - Université de Cheikh Anta Diop, Senegal - (3) LMDAN, FASEG-Université de Cheikh Anta Diop, Senegal - (4) Institute for Transport Planning and Systems (IVT), Zurich, Switzerland - (5) UMR ESPACE, CNRS, Avignon, France

Data Visualization Prize: Data for Development Reloaded: Visual Matrix Techniques for the Exploration and Analysis of Massive Mobile Phone Data

S. van den Elzen, M. van Dortmont, J. Blaas, D. Holten, W. van Hage, J-K. Buenen, J.J. van Wijk, R. Spousta *, S. Sala *, S. Chan *, A. Kuzmickas * University of Technology SynerScope BV Sensemaking Fellowship

Eindhoven University of Technology & SynerScope BV, The Netherlands

* Sensemaking Fellowship (MIT, Harvard University)

In our Visual analytics techniques for the exploration and analysis of massive mobile phone data, users are enabled to identify both temporal and structural patterns such as normal behavior, outliers, anomalies, periodicity, trends and counter-trends.

Practical Application Prize: Mobile Data as Public-Health Decision Enabler: A Case Study of Cardiac and Neurological Emergencies

E. Mutafungwa [1], F. Thiessard [2], M. Pathé Diallo [2], R. Gore [3], V. Jouhet [2], C. Karray [4], N. Kheder [4], R. Saddem [4], J. Hämäläinen [1], G. Diallo [1]

The objective of the study is to show the areas in which the absence of a nearest hospital can result in death or serious squeals. The identification of areas at high risk in case of stroke of myocardial infarction, requiring rapid intervention, could help Public Health decision makers to priorize investments.

(1) Department of Communications and Networks, Aalto University School of Electrical Engineering, Finland - (2) ERIAS INSERM U897, ISPED, Université de Bordeaux, France - (3) Virginia Modeling Analysis and Simulation, Old Dominion University, USA - (4) Faculté des Sciences de Tunis, University of Tunis, Tunisia

Scientific Prize and Ethics Mention: Construction of socio-demographic indicators with digital breadcrumbs

F. Bruckschen [1], T. Schmid [2], T. Zbiranski [1]

We show that socio-demographic indicators such as population, age, literacy, poverty, religion, ethnicity, electricity supply and others can be estimated in unprecedented detail and virtually ad-hoc using antenna-to antenna traffic data only. We offer a uniform approach that can be easily extended to other variables. Results are tested for spatio-temporal robustness and visualized as heat maps.

(1) Humboldt Universität Berlin, Germany - (2) Freie Universität Berlin, Germany

D4D Senegal

Results are published on the Web:

D4D scientific contributions
http://netmob.org/assets/img/NetMob%202015_D4D%20Challenge%20Senegal_Sessions_Scientific_Papers.pdf (298 pages, 83MB)

D4D posters
http://netmob.org/assets/img/NetMob%202015_D4D%20Challenge%20Senegal_Sessions_Posters.pdf (56 pages, 53 MB)

The datasets for D4D are described here: http://arxiv.org/abs/1407.4885

D4D Senegal was made possible thanks to…

Core team Sonatel

Sponsor:	Jerôme
PMO:	Coumba
BI:	Daouda
Legals:	Jamil, Oumar, Emilie
Com:	Racky & team

Core Team Orange

Sponsor	Pascal, Brigitte
PMO:	Stephanie
Labs:	Zbigniew & Cezary
Trainees:	Vladimir then Laetitia
CSR	Ludovic
Legals:	Dominique
Com:	Philippe
Web:	Serge

Anonymisation, Synthetic CDR, Weather

Anne-Sophie, Cezary, Romain, Emmanuel, Valérie, Dominique, Michel, CNRS team

D4D Committee

Vincent, Sandy, Jake, Bill, Robert, Johannes Elizabeth, Ngalla, Toguebaye, Omar, Zouhair, Marie-Claude, Nicolas

Ministries & Institutions

Commission des Données Personnelles
Recherche et Education supérieure
Plan Sénégal Emergent
Health and Social Action
Agriculture
Energie
Transport
National Statistics - ANSD
Centre de Suivi Ecologique
Telecom regulator

Prize donors

The Bill & Melinda Gates Foundation
Orange, Sonatel

Resources donors

Tera Labs

Data donors

WHO, Digital Globe, The Global Fund
IDM, Centre de suivi écologique,

D4D External Ethics Panel (DEEP)

Lucy Bernholz
Philippa Foster Back
William Hoffman
Johannes Jutting
Robert Kirkpatrick
Emmanuel Lulin
Ulrich Mans
Mark Nelson
Yaye Fatou Camara Niang
Nuria Oliver
Juliana Rotich
Olivier Sagna
Jean-Philippe Vanot
Pat Walshe

Internal Ethics Workgroup

Pierre Petillault
Fabrice André,
Bénédicte David
Annick Lelièvre
Valérie Peugeot,
Kitty Boullé
J-Y Leonnec
Ludovic Levy
Olivier Ondet
Oumar Sidibe
Catherine Fauvel

Hackathon, Datathon

OBS: Jean-Pierre, Stephane
Valérie, Axelle, Philippe
Boston: Yves-Alexandre, Zbigniew
Paris: Ludovic, Cezary, « Le Simplon »
Dakar: Ndongo & Technocentre

NetMob @ MIT

Yves-Alexandre

NetMob

Esteban

Orange Communication

Vanessa, Nicole, Cecile,
Sylvie, Christine, Liza, Xavier

Thanks to B&M Gates Foundation grants three projects were launched to explore further the use of CDRs for:

Health: Disease modeling
National Statistics: Proxy for indicators
Agriculture: Food security

orange™

D4D Senegal: Grants for implementation

Health	National Statistics	Agriculture
Uncovering the impact of human mobility on schistosomiasis... R Casagrandi et al., Politecnico di Milano, Italy	**Construction of socio-demographic indicators with digital breadcrumbs** F Bruckschen & T Zbiranski, Freïe Universität Berlin, Germany	**Genesis of millet prices in Senegal: the role of production, markets and their failures** D Jacques, UC Louvain, Belgium
Quantifying effect of movement due to holidays on malaria prevalence S Milusheva, Brown University, USA	**A multidimensional analysis of poverty and its determinants in Senegal** N Pokhriyal, State University of New York at Buffalo, USA	**Mobility profiles and calendars for food security and livelihoods analysis** PJ Zufiria et al., Universidad Politécnica de Madrid, Spain

(See: Annex)

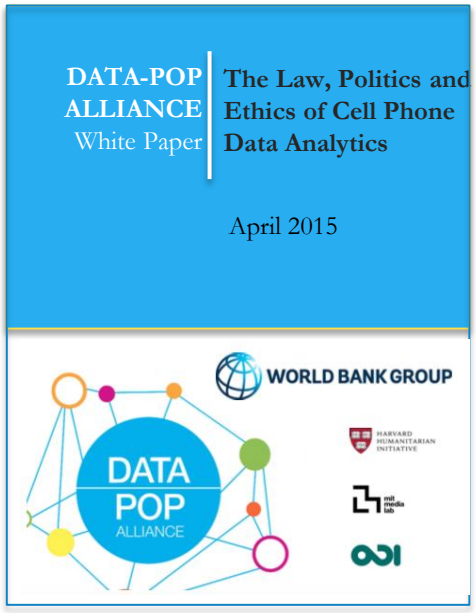
granted access to the complete database

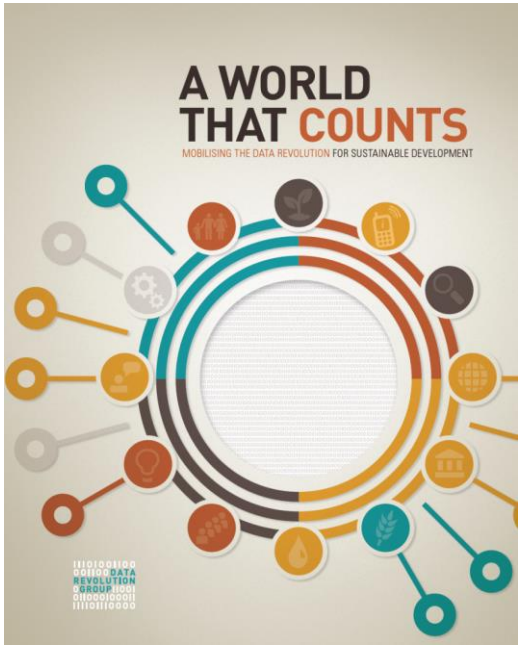
D4D Senegal: 6 projects implementation

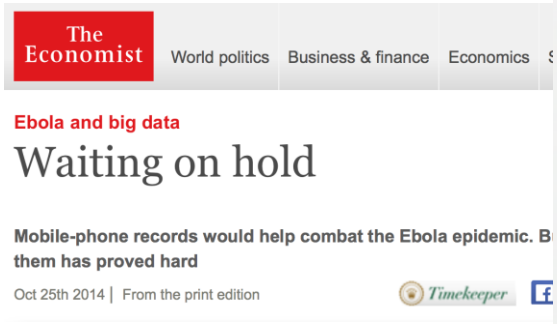
- Special event in Dakar (Radisson Sonatel event, 6th June 2015) and teams field visit to Senegal
- Collaborations with local actors:
  - ANSD (National Statistics)
  - Ministry of Health
  - World Food Program Dakar
  - DAPSA Senegal (Agriculture Statistics)
  - PNLP (National Malaria Control Program)
  - …
- Help of international development bodies (OECD, WFP, WHO, UN…)

- And finally a « twin project » on the National Statistics Senegal - Cote d'Ivoire is about to start

After D4D: the world of development needs Big Data







Privacy warning: impossible CDRs anonymization

- The notion of **anonymity** is one of the cornerstone of data protection
- But an individual spatiotemporal trace becomes very quickly unique, even in the largest datasets


SCIENTIFIC REPORTS

Altmetric: 1,481 Views: 271,913 Citations: 112 More detail »

Article | OPEN

Unique in the Crowd: The privacy bounds of human mobility

Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen & Vincent D. Blondel

Scientific Reports 3, Article number: 1376 (2013)
doi:10.1038/srep01376

Received: 01 October 2012
Accepted: 04 February 2013
Published online: 25 March 2013

- CDRs from an European country (1.5 M people, 15 months)
- Points: antenna tower / time step: one hour

$$\mathcal{E}_4 = .95$$


REPORT

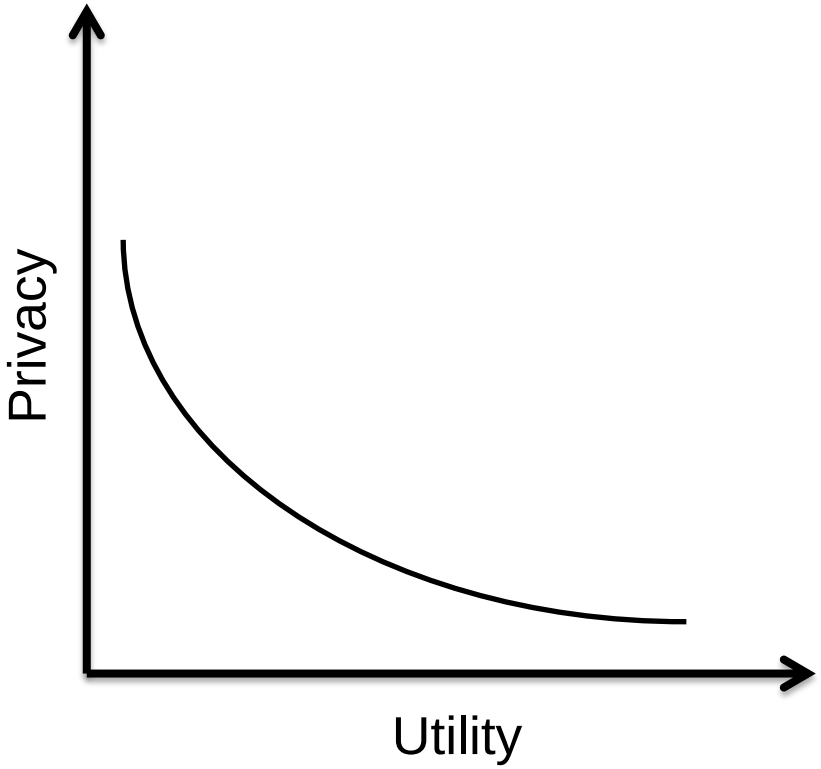
Unique in the shopping mall: On the reidentifiability of credit card metadata

Yves-Alexandre de Montjoye[1,*], Laura Radaelli[2], Vivek Kumar Singh[1,3], Alex "Sandy" Pentland[1]

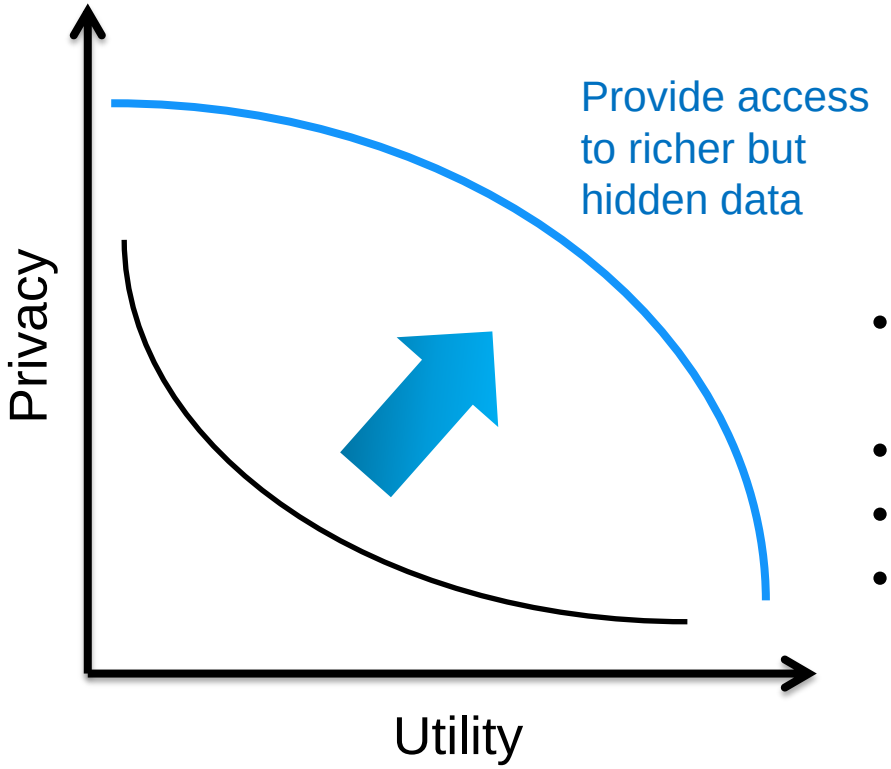
[1]Media Lab, Massachusetts Institute of Technology (MIT), 20 Amherst Street, Cambridge, MA 02139, USA.
[2]Department of Computer Science, Aarhus University, Aabogade 34, Aarhus, 8200, Denmark.
[3]School of Communication and Information, Rutgers University, 4 Huntington Street, New Brunswick, NJ 08901, USA.
*Corresponding author. E-mail: yvesalexandre(at)demontjoye.com

Science 30 Jan 2015:
Vol. 347, Issue 6221, pp. 536-539
DOI: 10.1126/science.1256297

The “privacy-utility” trade-off



A paradigm-shift in data protection



OPAL
bring the code to the data

- secure platform with pseudonymized data
- bring the code to the data using APIs
- create a bank of open algorithms
- set-up a governance board (local and global)

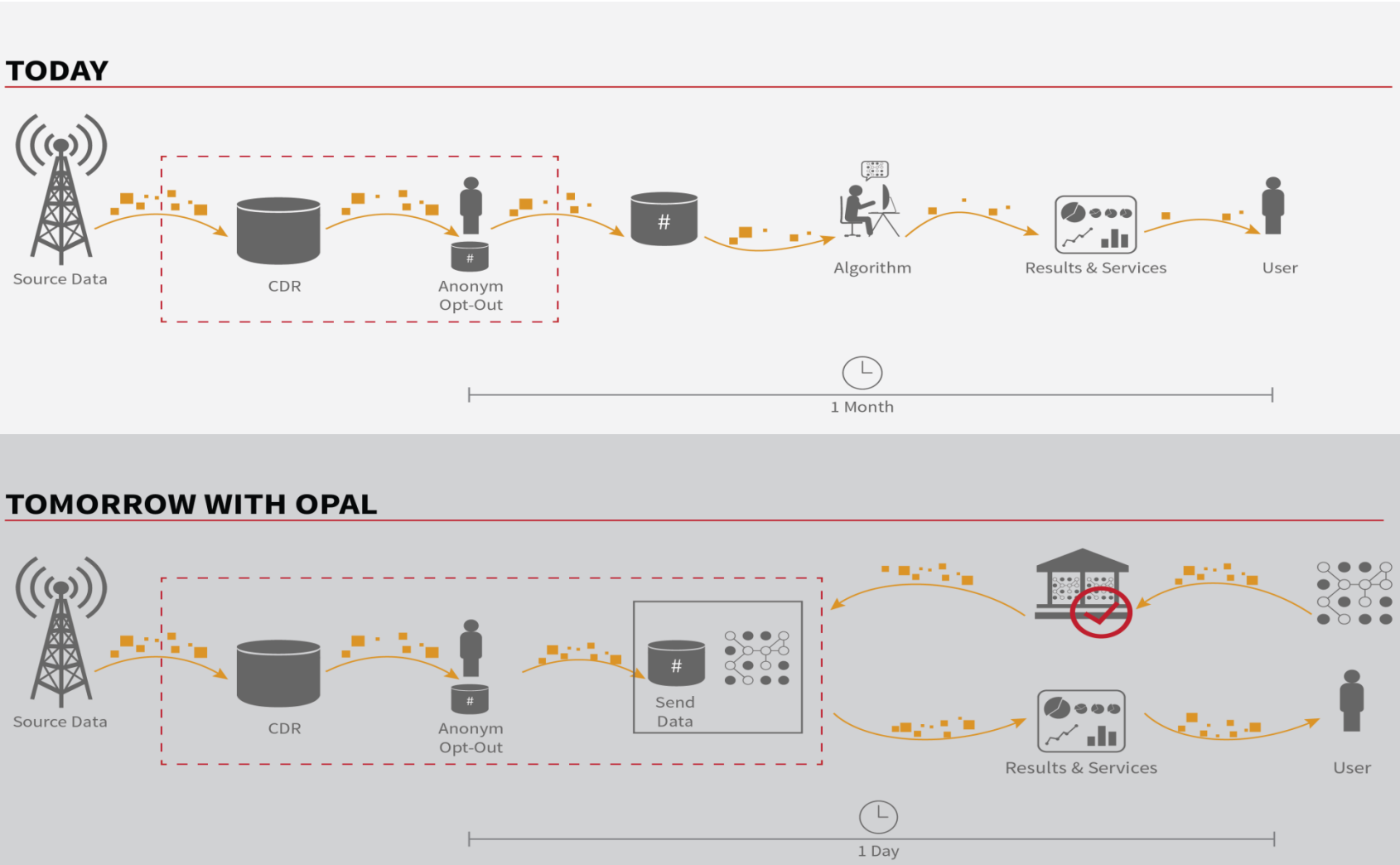
OPAL : a multiple privacy preserving platform

- OPAL project principles:
  - Use of Robust and State of the art technologies: Hadoop, Spark…
  - Leverage Open Software and a developer community
  - Develop User interfaces that make it easy to develop and use algorithms
  - Develop Security and Auditing with Best of Breed and BlockChain

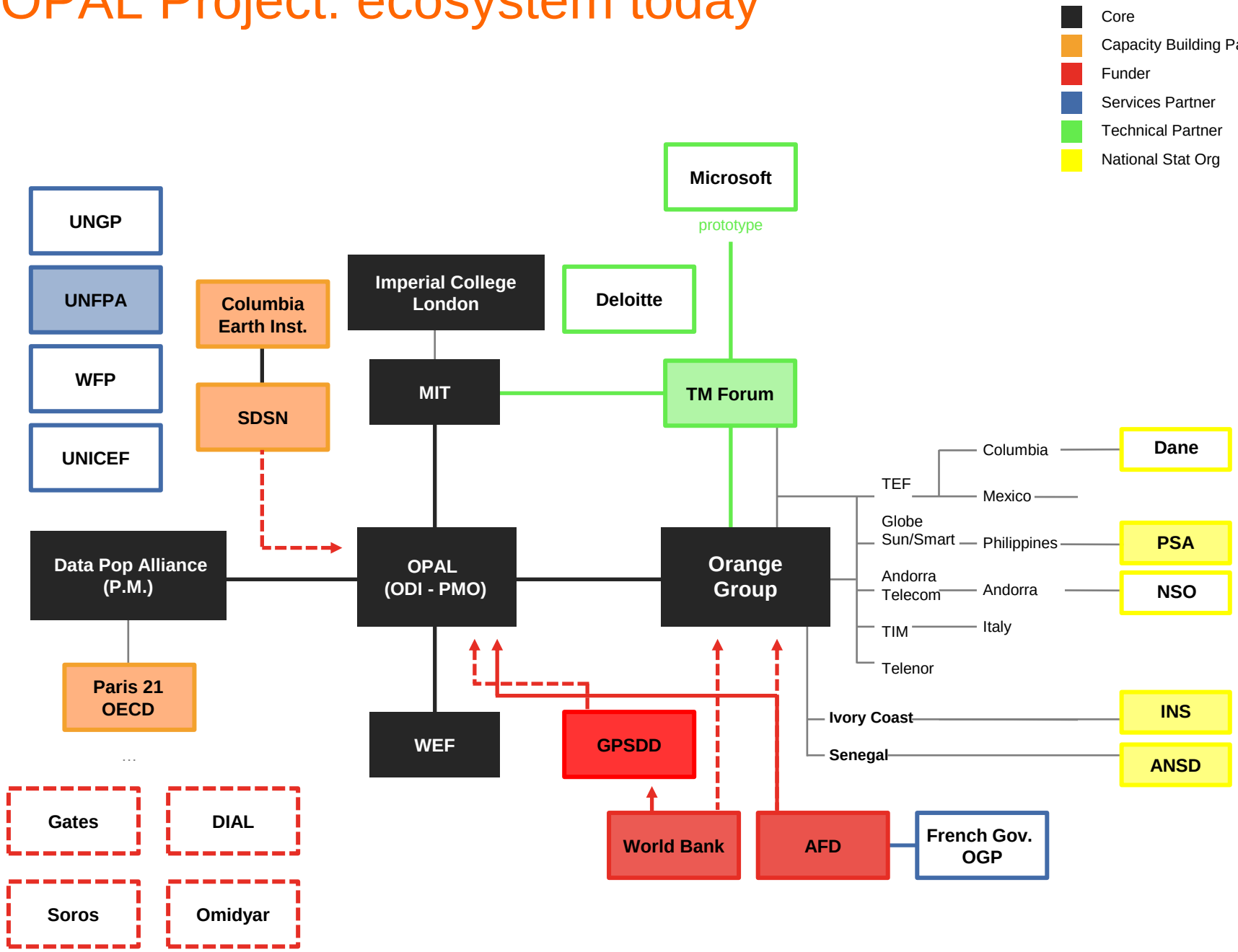
- OPAL Data Governance Board (local, multi-party: Civil Society, Academia, Government, Data Providers…) which check and certify algorithms to secure:
  - Privacy of individuals
  - Privacy of social groups, tribes, nations
  - “Privacy” of data-provider (business risks)

- A central support from OPAL team with a Data Ethics Network coordination

OPAL Project: query private data in a safe way



OPAL Project: ecosystem today



OPAL project

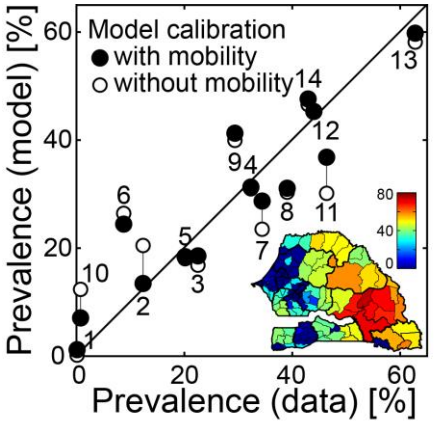
- **The Official Launch of the OPAL Project** took place on 8 December 2016 during the Open Government Partnership Global Summit in Paris

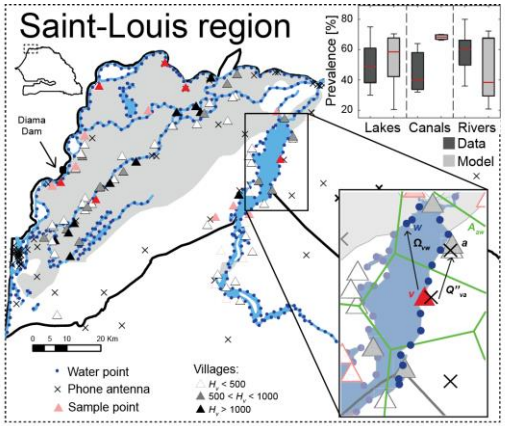
- OPAL contact: Natalie Shoup (Program Manager) nshoup@opalproject.org

Annex:
B&M Gates Foundation founded D4D Senegal projects in detail

D4D Challenge Senegal, Grant from the B&M Gates Foundation

Uncovering the impact of human mobility on schistosomiasis





Researchers: Renato Casagrandi, Manuela Ciddio, Marino Gatto, Lorenzo Mari – Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Italy

Contact: renato.casagrandi@polimi.it

Links to publications:

- Ciddio, Mari, Sokolow, De Leo, Casagrandi, Gatto (2016). The spatial spread of schistosomiasis: a multidimensional network model applied to Saint-Louis region, Senegal. Advances in Water Resources, http://dx.doi.org/10.1016/j.advwatres.2016.10.012
- Mari, Casagrandi, Ciddio, Dia, Sokolow, De Leo, Gatto. Big-data-driven modeling unveils country-wide drivers of endemic schistosomiasis, submitted

Conferences where presented:

- National Congress of the Italian Ecological Society, Milan (IT), Sept 2016
- International Workshop DeMagma at Univ Pierre et Marie Curies, Paris (FR), 2-4 Feb 2016
- 5th Int Conf on Infectious Diseases, Epidemics 5, Clearwater Beach (FL-USA), 1-4 Dec 2015
- 9th Eu Congr on Tropical Medicine ECTMIH 2015, Basel (CH), 6-10 Sep 2015
- MeetMeTonight 2015, Milano (IT), 25-26 Sep 2015, [short and long video]
- Radisson Sonatel event, Dakar (SN), June 2015
- Impact of Env Changes on Infectious Diseases, IECID 2015, Sitges (ES), 23-25 Mar 2015

Local collaborations during the research:

Ministry of Health: Dr. Elhadji Daouda Dia

Objective of Project & key results:

▪ Schistosomiasis is water based parasitic worm infection affecting millions of people. We build large-scale, data-driven models aimed at understanding, predicting and possibly controlling country-wide schistosomiasis transmission in Senegal.

▪We show that a relatively simple model can reliably reproduce regional patterns of schistosomiasis prevalence across the country and assist in health care planning. The inclusion of human mobility is an important factor for the realism of model predictions. Also, when tailored for specific regions of Senegal, the model can help identify the focal transmission hotspots of the disease.

Possible use for development:

▪ We use the model to study the role of human mobility on disease dynamics and to analyze intervention strategies aimed at reducing disease burden.

▪ The model can also be used to map infection risk at country and regional scales, so as to inform communication campaigns and disease control.

Further Research needs:

▪ CDRs for other years and/or countries to validate model predictions
▪ High-res schistosomiasis data for fine-tuning of the model
▪ Ecological data to describe the dynamics of intermediate snail hosts

Local next steps in progress:

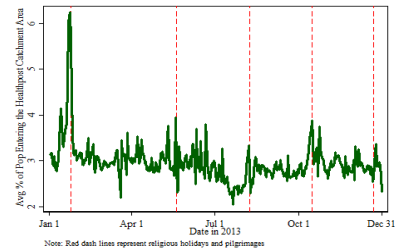
Project **MASTR-SLS** (**MA**pping **S**chistosomiasis **T**ransmission **R**isk in **S**aint-**L**ouis, **S**enegal), funded by Politecnico di Milano through the PoliSocial initiative (www.polisocial.polimi.it) . The project (Jan 2017 – Dec 2018) will consist of three main tasks: i) collection of epidemiological (school surveys) and ecological data (snail sampling) in the region; ii) elaboration of quantitative tools for transmission risk mapping; iii) design of communication strategies to help prevent disease spread in a highly endemic area.

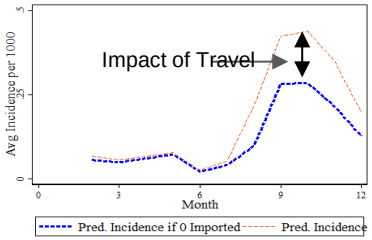
Local Customer(s) for delivery of results:

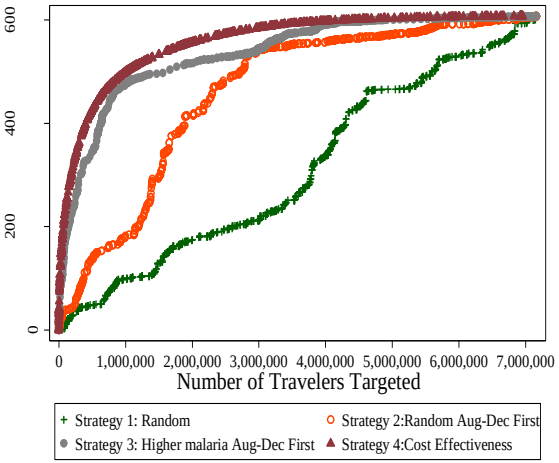
Ministry of Health: Docteur Marie Khemesse Ngom NDIAYE, Directeur de la Lutte Contre la Maladie

D4D Challenge Senegal, Grant from the B&M Gates Foundation

Quantifying the effect of population movement on malaria incidence using cell phone data







Researchers: Sveta Milusheva, Brown University, USA

Contact: svetoslava_milusheva@brown.edu

Links to publications: (Job Market Paper, will submit for publication in the Spring)

https://drive.google.com/open?id=0B8oQfBs38UHGZ01fMnh2bFBmZWc

Conferences where presented:

North East Universities Development Consortium Conference (Boston, Nov 5 2016), Population Health Sciences Research Workshop (Boston, Sept 23 2016), PopPov Conference on Population, Reproductive Health and Development (Sept 8 2016), World Bank ABCDE Conference (June 21 2016), Population Association of America Conference (April 1 2016); Radisson Sonatel event, Dakar, 6 June 2015

Local collaborations during the research:

PNLP: Omar Sarr

MACEPA: Philippe Guinot

Objective of Project & key results:

- Develop model of malaria for the Ministry of Health using Mobile Meta Data to derive patterns of mobility from and to locations where the disease is present
- Demonstrate the usefulness of CDR to improve malaria diffusion modeling, improving the potential for targeting efforts especially in regions with low prevalence
- Used modeling to show that each infected traveler contributes to 1.20 new cases of malaria in the location they enter
- Demonstrate the efficiency gain from targeting travelers most likely to be bringing malaria infections
- Presented results to PNLP policy strategy meeting Jan 2016, and being asked to develop a service by MACEPA

Possible use for development,

- Improve Policy making for national program (PNLP)
- Develop a service/App for weekly information to health facilities in the field
- Additional mobile phone data to validate the model and test out of sample predictions
- Weekly data to test development of a weekly information app

Local next steps in progress:

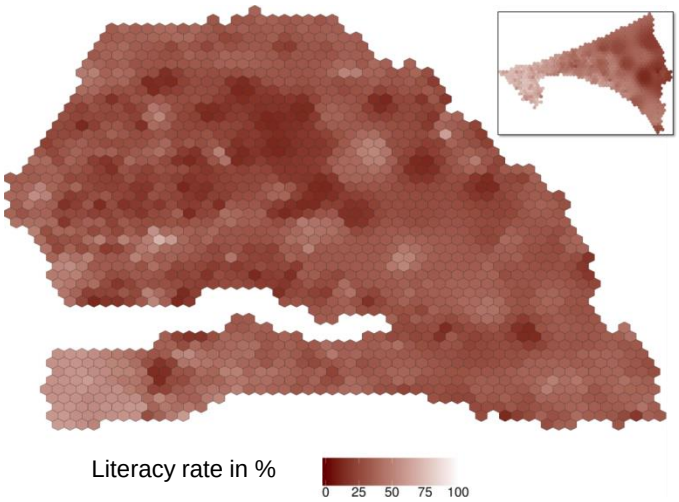
- 2017 service pilot for health workers in Richard Toll for Malaria elimination before scaling. Proposal submitted to the GPSDD innovation fund on Sept 1st 2016.
- Create SMS program for travelers when entering a new location concerning staying healthy and potentially provide an incentive in the SMS to visit a facility and get tested if coming from a high malaria area

Local Customer(s) for delivery of results:

Ministry of Health: Dr N'diaye, Directeur de la lutte contre la maladie

D4D Challenge Senegal, Grant from the B&M Gates Foundation

Construction of socio-demographic indicators with digital breadcrumbs



Researchers: Fabian Bruckschen, Till Zbiranski, Freïe Universität Berlin, Germany
Contacts: fabian.bruckschen@gmx.de, till.zbiranski@cms.hu-berlin.de

Objective of Project & key results:

- To capture so far hidden local socio-demographic heterogeneity, such as pockets of illiteracy or poverty. We offer a uniform approach that can be easily extended to other variables. It is based on aggregated antenna traffic data only - data that is less prone to privacy concerns than e.g. mobility patterns, thus facilitating implementation.
- Models are fitted to geocoded survey data and used for prediction on the tower level. Results are tested for spatiotemporal robustness and visualized as heat maps.

Possible use for development:

- In the short-term, uncovering local socio-demographic heterogeneity at little costs can facilitate timely & targeted relief.
- In the medium-term, variables can be modelled reliably, collected less frequently, thereby reducing the scope and thus the costs of surveys.

Links to publications:

Conferences where presented:

Radisson Sonatel event, Dakar, 6 June 2015

Local next steps in progress:

- 2017 key socio-demographic indicators generated from CDRs for the 17 Sustainable Development Goals

Local collaborations during the research:

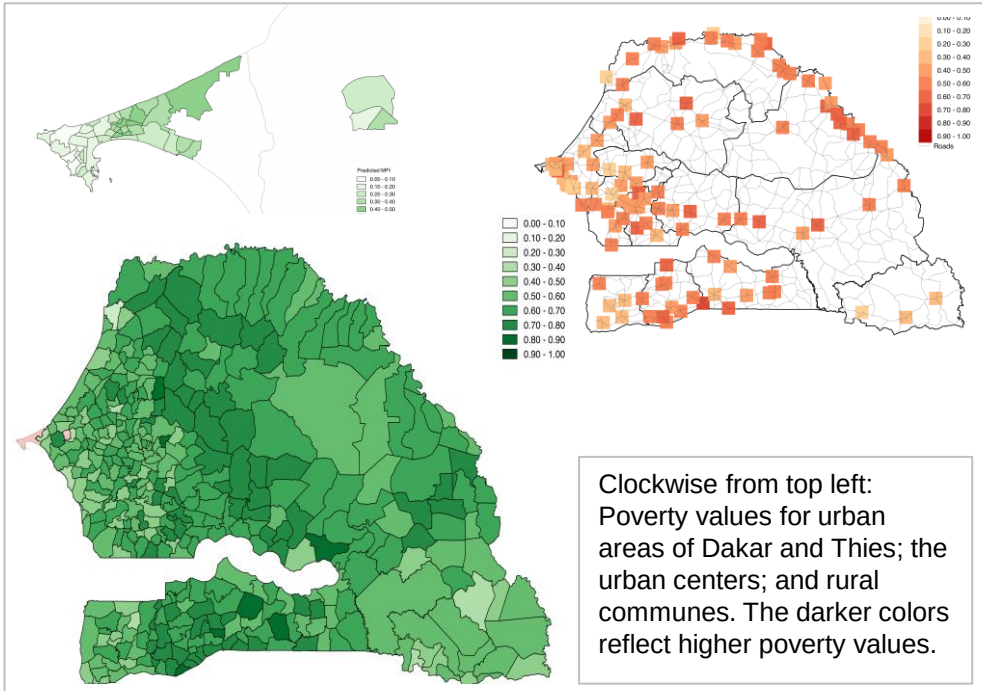
PARIS21/OCDE: Johannes Jutting

Local Customer(s) for delivery of results:

ANSD: Mr. Mamadou NIANG, Informatique Statistique
INS : Mr. Be, Directeur de l'INS (Cote d'Ivoire)

D4D Challenge Senegal, Grant from the B&M Gates Foundation

A multidimensional analysis of poverty and its determinants in Senegal



Clockwise from top left: Poverty values for urban areas of Dakar and Thies; the urban centers; and rural communes. The darker colors reflect higher poverty values.

Objective of Project & key results:

- To develop highly accurate and spatially-detailed, 'commune-level', poverty maps using mobile phone data.
- We provide such maps ensuring full geographical coverage of Senegal. Our model estimates of poverty have been validated using census data provided by ANSD.
- We study the robustness of our model predictions, and its generalization capability to data-scarce locations.
- We integrate mobile phone data with census, to build a model that predicts individual determinants of poverty in the dimensions of health, education and standard of living.

Possible use for development,

- Spatially detailed policy-planning level poverty maps will assist in targeted policies for inclusive growth of all the micro-regions.
- These maps are disaggregated into individual determinants of poverty like deprivations in health, education, assets, further assisting in appropriate policy intervention.
- These maps can be generated at any temporal resolution, as they are based in mobile phone data.

Researchers:

Neeti Pokhriyal, State University of New York at Buffalo, USA

Contacts:

neetipok@buffalo.edu

Further Research needs:

- Combining mobile data with other diverse sources of data, like satellite imagery to better predict poverty.

Links to publications and conferences where presented:

1) A Computational Approach to Poverty Mapping; International Conference on Computational Sustainability at Cornell University in July 2016.
2) Virtual Network and poverty analysis in Senegal; Radisson Sonatel event, Dakar, 6 June 2015

Local next steps in progress:

- A twin project considering the synergies of the model in Ivory Coast
- Newer mobile phone data for study temporal stability of the model

Local collaborations during the research:

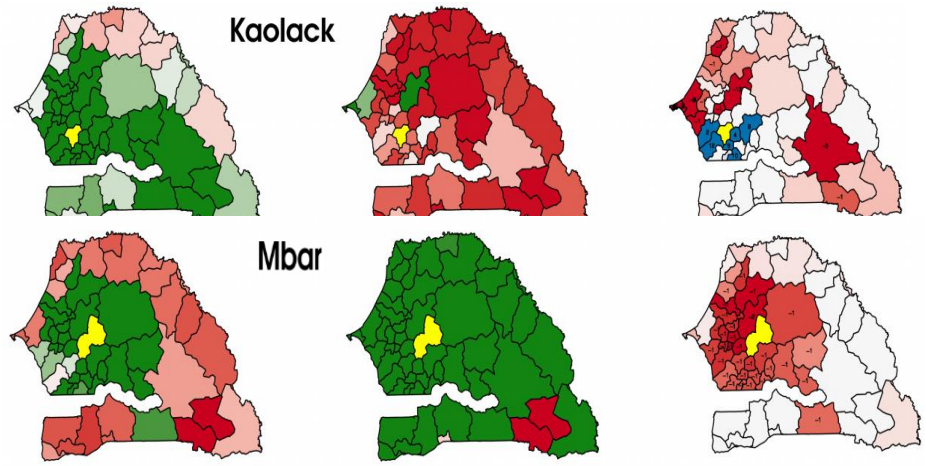
PARIS21/OCDE: Thilo Klein, OCED

Local Customer(s) for delivery of results:

ANSD – Sene Papa Ibrahima, ANSD

D4D Challenge Senegal, Grant from the B&M Gates Foundation

Genesis of millet prices in Senegal: a trade flow approach



Objective of Project & key results:

Monitoring and forecasting of staple prices are capital to formulate adequate policies in favor of food security in Africa. Vegetation index derived from Earth Observation satellite data combined with a unique dataset of mobile phone calls between 9.000.000 users support the development of a new conceptual approach to estimate monthly millet prices for 3 contrasted years in 57 markets in Senegal. We introduce a mechanistic model that simulate trade flows between markets taking into account transportation cost and information asymmetry. Prices are accurately estimated up to 10 months after the harvest. Although the distance between the markets is the main factor explaining price differentials, the social network also appeared as a significant proxy of trade intensity.

Researchers:

Damien Jacques, Université catholique de Louvain (UCL), Belgium

Contact:

damienjacquesUCL@gmail.com

Possible use for development,

These results can be used in the assessment of the social welfare impacts of the further development of both road and mobile phone networks in the country. The model could be further developed as a valuable tool for the forecasting of staple prices in the country.

Links to publications:

Main publication will be submitted to PNAS week of 17th of October 2016
http://bit.ly/1QoXNpu, http://bit.ly/1TO9wDT

Conferences where presented:

Radisson Sonatel event, Dakar, 6 June 2015
AGU Fall Meeting, San Francisco, 14-18 December 2015

Local next steps in progress:

Potential support for operational monitoring still to be discussed

Further Research needs:

New findings are expected from the use of several years of mobile phone data and the expansion of the model to other Sahelian countries.

Local collaborations during the research:

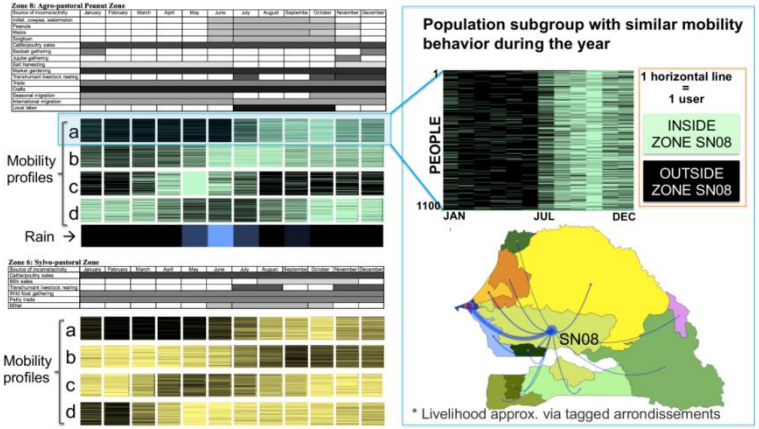
WFP Dakar: Wilfred NKWAMBI
CSE: Bamba Diop
DAPSA: Mame Nogaye Fall
CSA: Mr. Aly MAR, Directeur Général

Local Customer(s) for delivery of results:

WFP: Mr. Simon RENK, Regional Market Advisor

D4D Challenge Senegal, Grant from the B&M Gates Foundation

Mobility profile and calendars for food security and livelihood analysis



Researchers: Pedro J. Zufiria, David Pastor-Escuredo, Luis A. Úbeda-Medina, Miguel A. Hernández-Medina, Iker Barriales-Valbuena, Alfredo Morales-Guzmán, Universidad Politécnica de Madrid, Spain; Wilfred Nkwambi, WFP; John Quinn, Paula Hidalgo-Sanchís, Miguel Luengo-Oroz, UN-Global Pulse

Contact: pedro.zufiria@upm.es

Links to publications:

https://www.dropbox.com/sh/x05dz5e3turyhk9/AABKJ8R2DguPi9Ox_3rRFZrxa?dl=0

Conferences where presented:

Radisson Sonatel event, Dakar, 6 June 2015
Orange Spain, Invited lecture, Madrid, 24 June, 2015
Big Data, Good Data. Workshop. Madrid , 12-13 November 2015
Big Data and Climatic Change, Workshop FRA, 29 Feb. 2016
Data Beers, 2 Invited talks, Madrid, 25 January and 12 March, 2016

Local collaborations during the research:

WFP Dakar: Wilfred NKWAMBI
DAPSA: Mama Nogaye Fail
CSA: Mr. Aly MAR, Directeur Général
CSE: Mr. Mouhamadou Bamba DIOP
SECNSA: Mr. Cheikh Sadibou Pene

Objective of Project & key results:

- To create a mobility analysis system leading to a mobility baseline at different levels (markets, cities, administrations, production areas).
- To use the system for estimating and/or forecasting crops yields based on farmers prediction capabilities and consequent movements.
- To develop an food security early warning system based on the estimated indicators.
- Provide accurate estimates and forecasts of crop area and yield to policy makers and integrate mobility information with NDVI and rainfall data. Integration with knowledge provided by local actors..

Possible use for development,

- Use the mobility analysis system for estimating coping strategies and the impact of early humanitarian actions.
- Use the mobility analysis system for food security early warning. This application can be used by the WFP.
- Use the system to characterize pastoralism (CSA and CIRAD).
- Use the system to estimate early indicators in order to improve market prices estimates (in collaboration with D Jacques).

Further Research needs:

- Define a baseline of regular behavior based in CDRs corresponding to years 2014 and 2015 (new available data).
- Assessment of the system with ground truth information and expertise of local actors.
- Improve space resolution of estimates.

Local next steps in progress:

- SECNSA, CSE and CSA contacts are expected to collaborate with African Risk Capacity.
- Cheetah Accelerator (Mariéme Jamme) can provide local grassroots software developers.

Local Customer(s) for delivery of results:

WFP: Mr. Simon RENK, Regional Market Advisor
CSE and CSA are interested in the applicability of the system