

THOUGHTS ON CHAPTER & THEME “GROUP PRIVACY IN THE AGE OF BIG DATA” —THE CASE OF MOBILE PHONE DATA

Emmanuel Letouzé | Director & Co-Founder
Amsterdam Privacy Week | Oct. 25, 2015

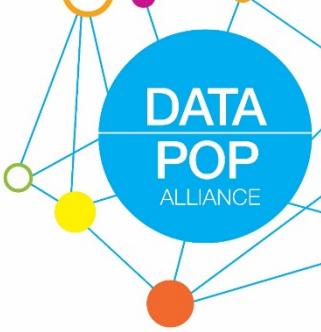


HARVARD
HUMANITARIAN
INITIATIVE



DATA-POP
ALLIANCE

GROUP PRIVACY IN THE AGE OF BIG DATA



DATA-POP
ALLIANCE

GROUP PRIVACY IN THE AGE OF BIG DATA

October 24, 2015

DRAFT V2 FOR DISCUSSION

October 2015

By

Lanah Kammourieh*
Thomas Baar
Jos Berens
Emmanuel Letouzé
Julia Manske
John Palmer
Patrick Vinck

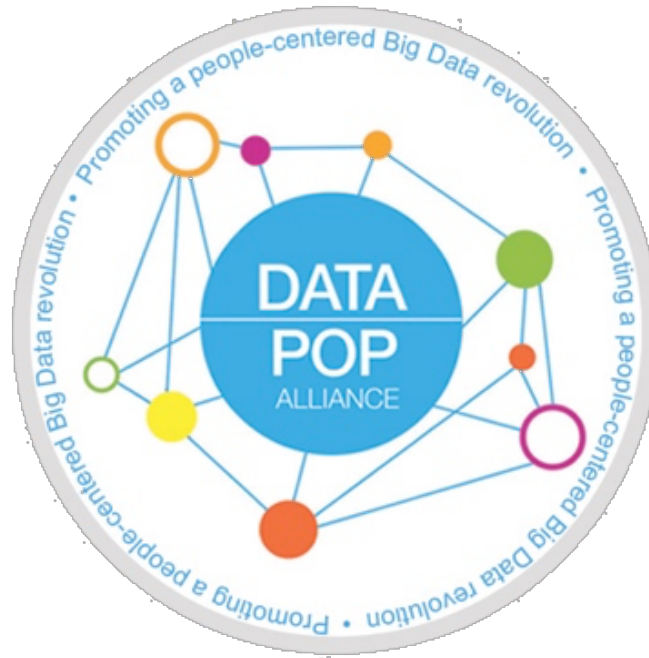
With contributions from

Augustin Chaintreau
Yves-Alexandre de Montjoye
Augustin Chaintreau
Natalie Shoup

What we (aim to) do



HARVARD
HUMANITARIAN
INITIATIVE



CYBERSECURITY Saving Big Data from Itself

A three-step plan for
using data right in an age
of government overreach

By Alex "Sandy" Pentland

For the first few decades of its existence, the National Security Agency was a quiet department with one primary job: keeping an eye on the Soviet Union. Its enemy was well defined and monolithic. Its principal tools were phone taps, spy planes and hidden microphones.

After the attacks of September 11, all of that changed. The NSA's chief enemy became a diffuse network of individual terrorists. Anyone in the world could be a legitimate target for spying. The nature of spying itself changed

as new digital communication channels proliferated. The exponential growth of Internet-connected mobile devices was just beginning. The NSA's old tools apparently no longer seemed sufficient.

In response, the agency adopted a new strategy: collect everything. As former NSA director Keith Alexander once put it, when you are looking for a needle in a haystack, you need the whole haystack. The NSA began collecting bulk phone call rec-

Research:

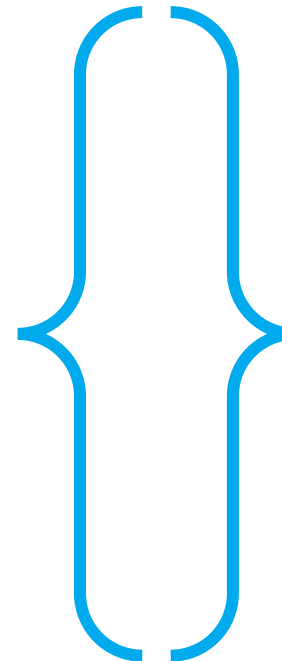
Advancing Knowledge and Innovation

Training:

Building Capacities and Connections

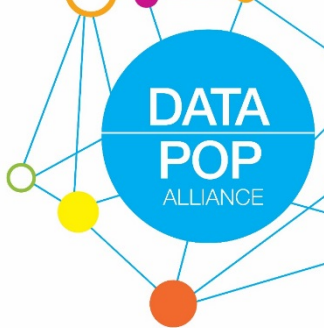
Engagement:

Crafting Ethical Systems and Strategies

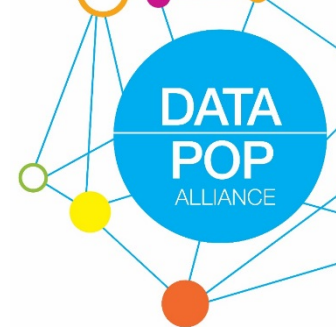


1. Politics and Governance
2. Official and Population Statistics
3. Peacebuilding and Violence
4. Climate change and Resilience
5. Data Ethics and Rights

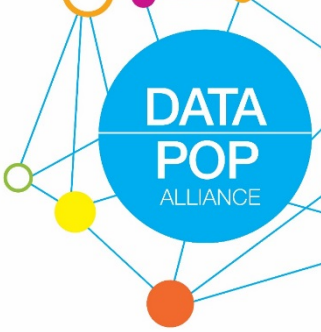
Data-Pop Alliance



Co-authors & contributors

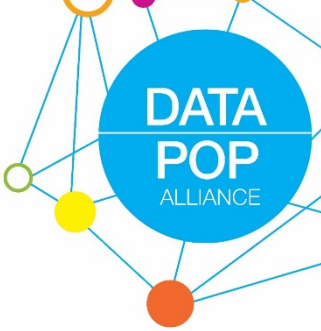


Starting Point & Main Questions



1. Big Data allows governments and businesses to track the habits and movements of individuals, combine and recombine people into categories, and analyze and attempt to predict their behavior.
2. Individual data is no longer only useful for gaining information about and targeting the individual, but also – and perhaps above all – for gaining information about and targeting groups.
3. The questions raised in this chapter flow naturally from these observations:
 1. *Is there such a thing as 'group privacy'*
 2. *Is 'group privacy' a workable concept?*
 3. *If so, should it be a legally enforceable right, and how can it be protected?*
 4. *How about a regulatory approach and broader social requirements?*

Outline

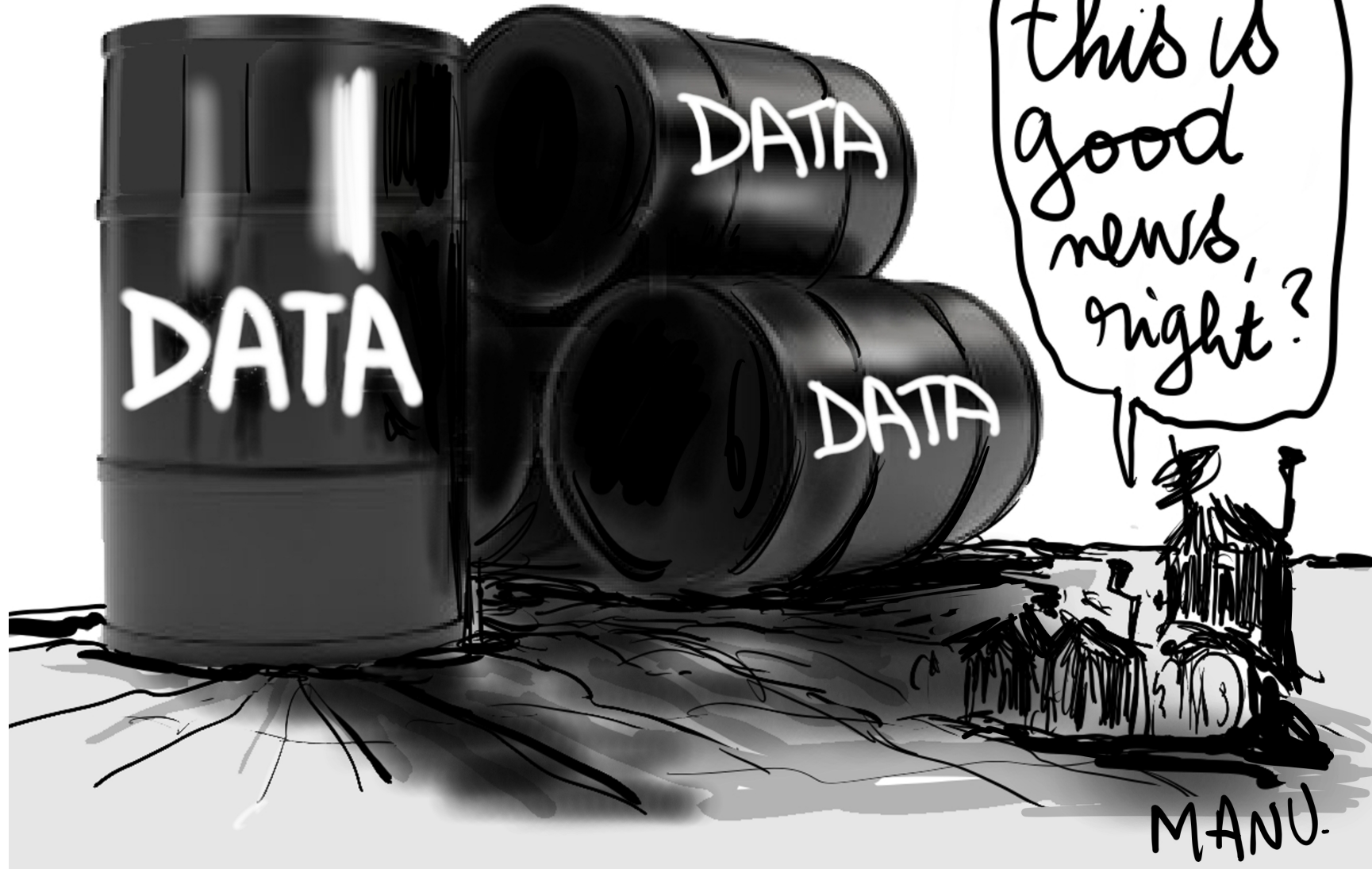
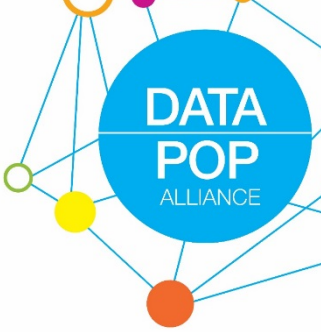


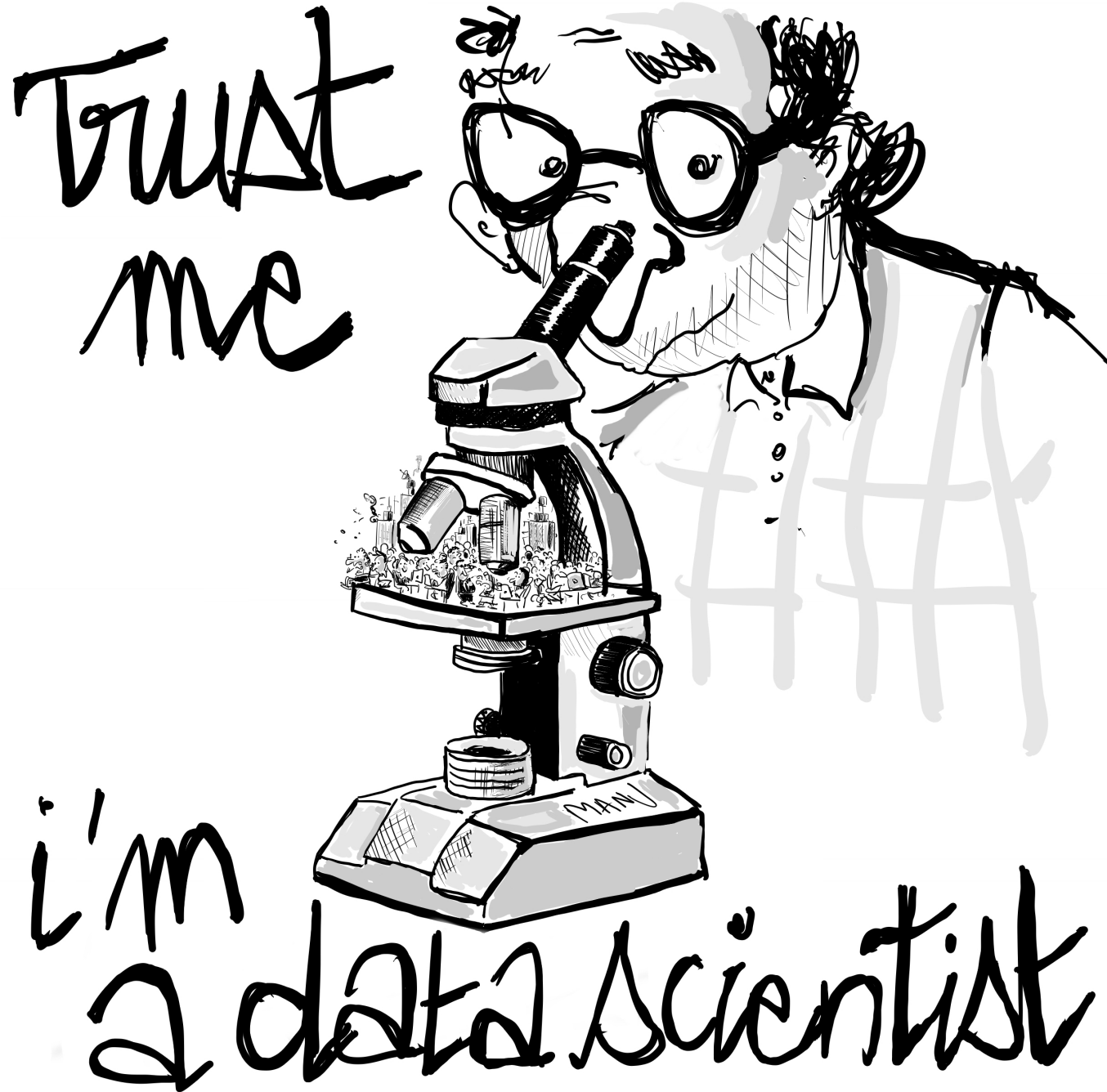
I.	What is a group?	6
A.	Traditional notions of groups: self-proclaimed and self-aware	6
B.	Big Data: new grounds for identifying groups	7
II.	What is group privacy?	9
A.	Challenging traditional notions and protections of privacy	10
	<i>The shifting ontology of privacy</i>	10
	<i>Data, information, and knowledge</i>	11
	<i>Privacy protection in the age of Big Data</i>	12
B.	The content and protection of group privacy	13
	<i>Group privacy is not reducible to individual privacy</i>	13
	<i>Group privacy as dignity: concepts of self-determination and data sovereignty as applicable to groups</i>	15
III.	Affirming and protecting group privacy	16
A.	Through traditional levers of power: developing a legislative and regulatory framework for group privacy	16
B.	Through a more harmonized international regime for data management by users	18
C.	By improving data security and accountability for breaches	19
D.	Through improving awareness and 'data literacy'	20



Problem Statement

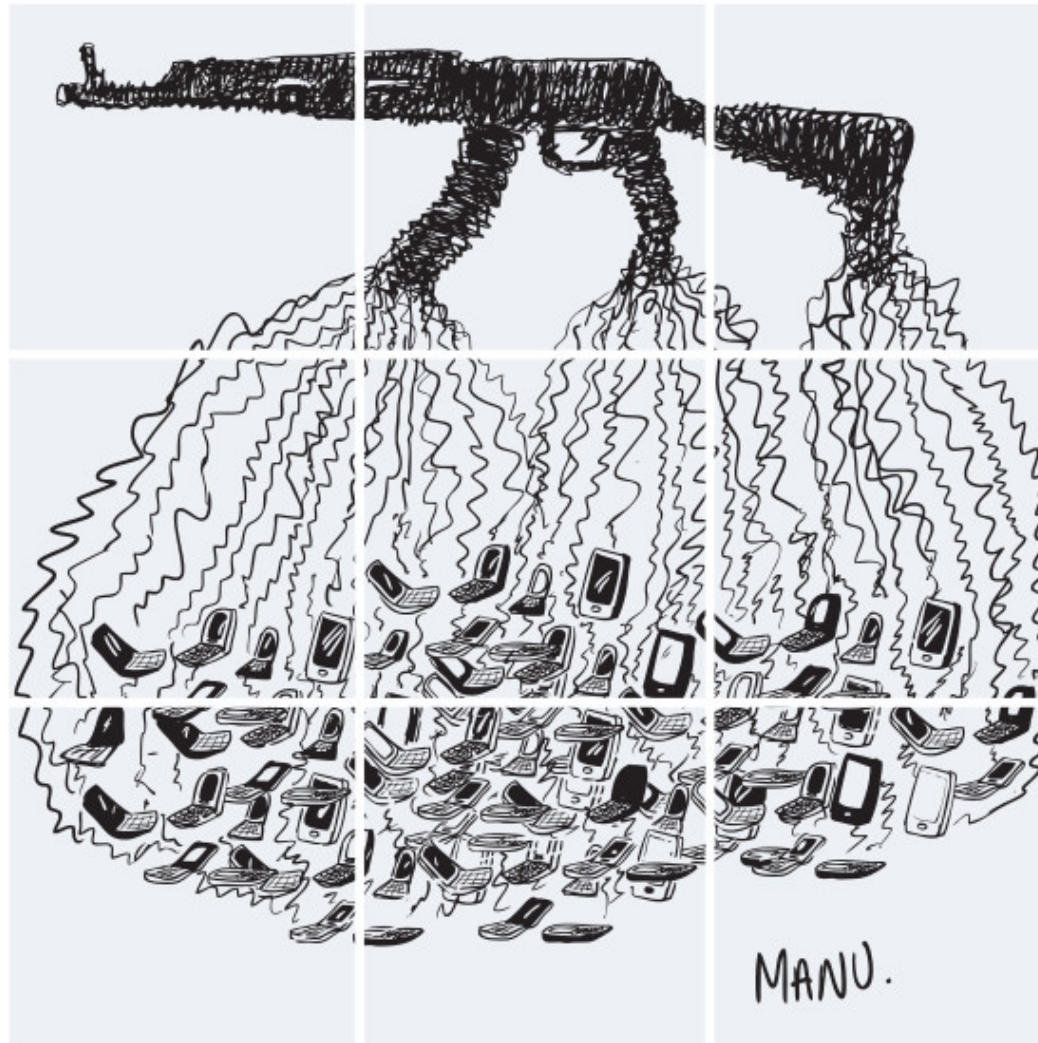
Data is the new oil





New Technology and the Prevention of Violence and Conflict

EDITED BY FRANCESCO MANCINI



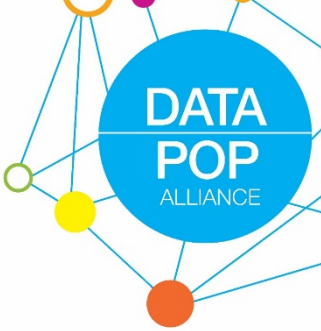
*Empowered lives.
Resilient nations.*

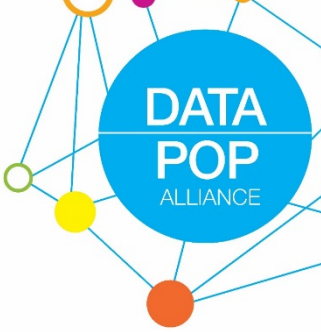


USAID
FROM THE AMERICAN PEOPLE

IPI

INTERNATIONAL
PEACE
INSTITUTE





Taxonomies of applications of Big Data:

1. Descriptive
2. Predictive:
 - i) forecasting
 - ii) inference
3. Prescriptive—causal inference
4. Discursive

an illustrated introduction to Predicting socioeconomic levels through cell-phone data

Question:



so, how is it possible to predict an area's socioeconomic - or poverty- level from the cell - phone data it emits?

step ①

first find or collect actual survey data..

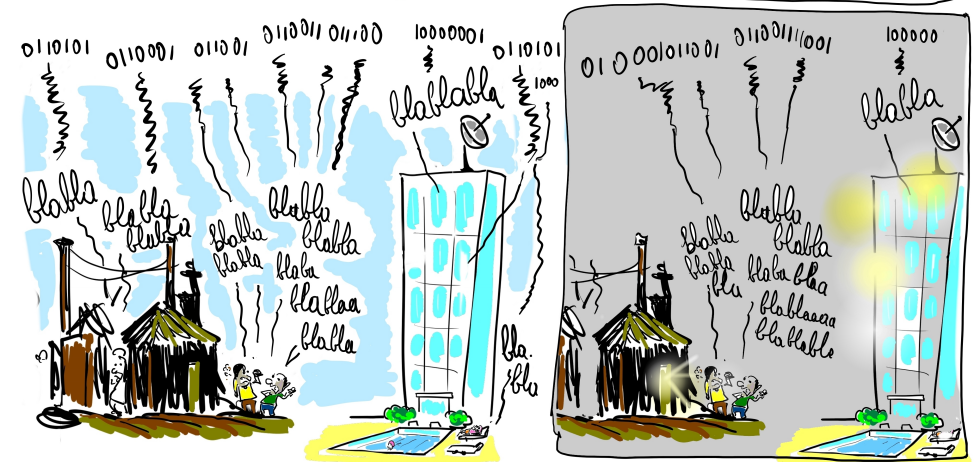


hello, we are conducting an official survey: are you poor or rich?

step ②



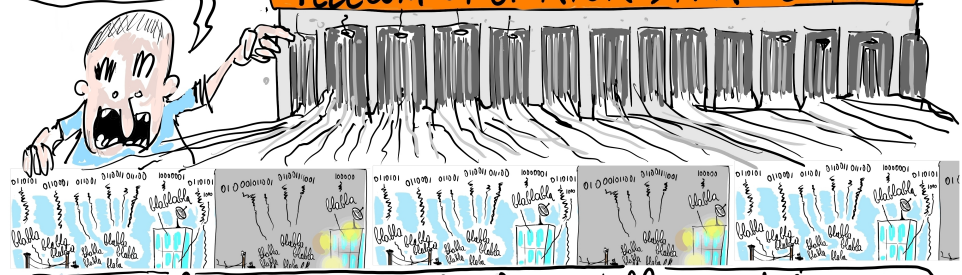
then notice how cell phone users leave digital traces, day & night..



"these digital traces, recorded by every telecom operator, are << Call Detail Records >> or CDRs, metadata that look like that"

CALLER ID	LOCATION	RECIPIENT ID	RECIPIENT LOCATION	CALL TIME	CALL DURATION
X368729748Y	2°24'22" 35°49'56"	A8C492TC73646	3°38'49" 31°12'12"	2014.04.01 ET 17 12	01.12.27

TELECOM OPERATOR DATA CENTER



"and these CDRs will show differences in calling patterns between different areas ..."

POOR AREA

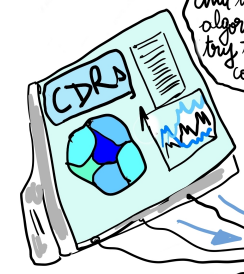
RICH AREA

their "digital signatures" will differ

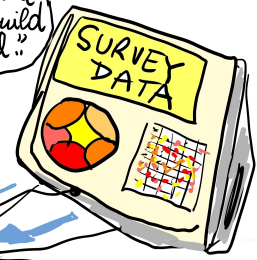
maybe people in richer areas will stop calling sooner at night?

and using computers and algorithms, researchers try to find the strongest correlations

between CDRs and survey data to build a predictive model..

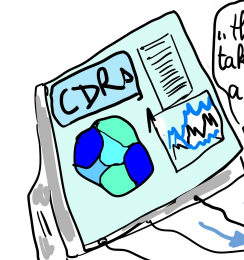


PREDICTIVE MODEL?

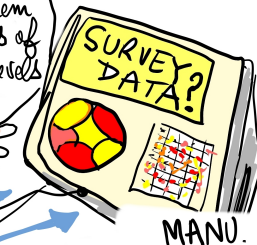


"that can then take CDRs from a later time or different area"

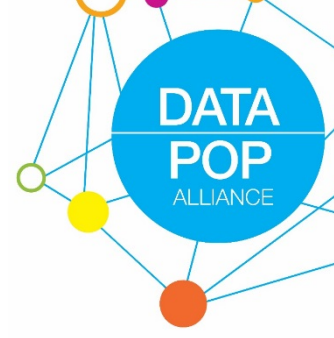
"and turn them into estimates of socioeconomic levels without a survey!"

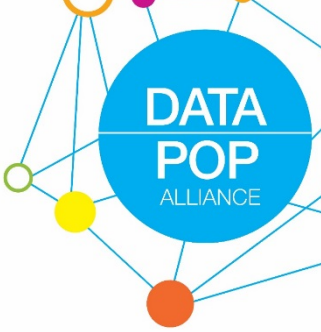


PREDICTIVE MODEL



MANU.



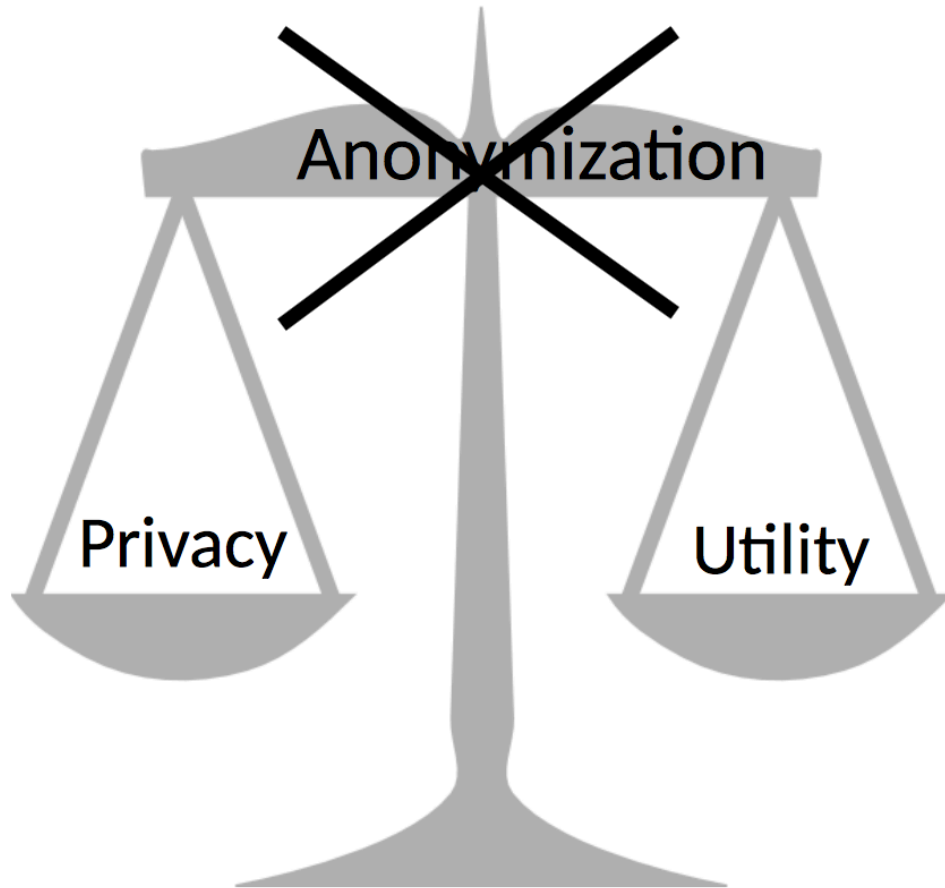
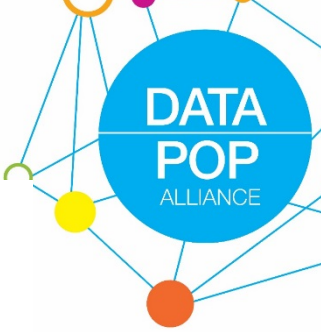


Computational Privacy

How human behavior bounds privacy
and what we can do about it

Yves-Alexandre de Montjoye
Twitter: yvesalexandre
MIT Media Lab – Human Dynamics

Big Data all but kills anonymity

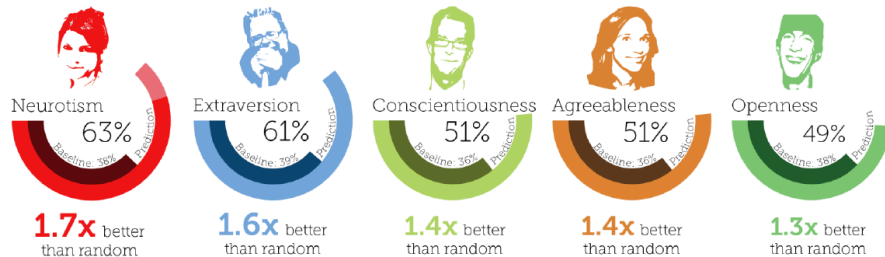


1. It is **possible** to re-identify mobile phone metadata, credit card metadata (even if there is no name, phone number, etc)
2. It is **not simply** a question of coarsening the data (we'd just need a few more points or more information)
3. It is **not "just"** metadata or what is directly visible in the data (e.g. one might use it to predict your personality)

Big Data also provides new approaches and grounds upon which groups can be identified, formed, monitored and targeted

Group privacy isn't just the sum of individual privacies

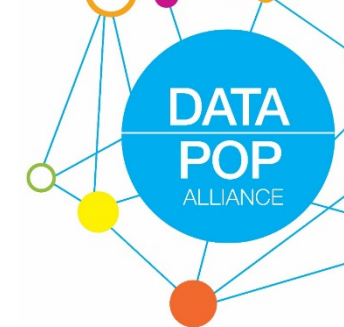
Predicting **personality** using **metadata**



de Montjoye, Y. A., Quoidbach, J., Robic, F., & Pentland, A. S. (2013). Predicting personality using novel mobile phone-based metrics. In Social Computing, Behavioral-Cultural Modeling and Prediction (pp. 48-55). Springer Berlin Heidelberg.



Lessons & reflections from the Ebola Outbreak

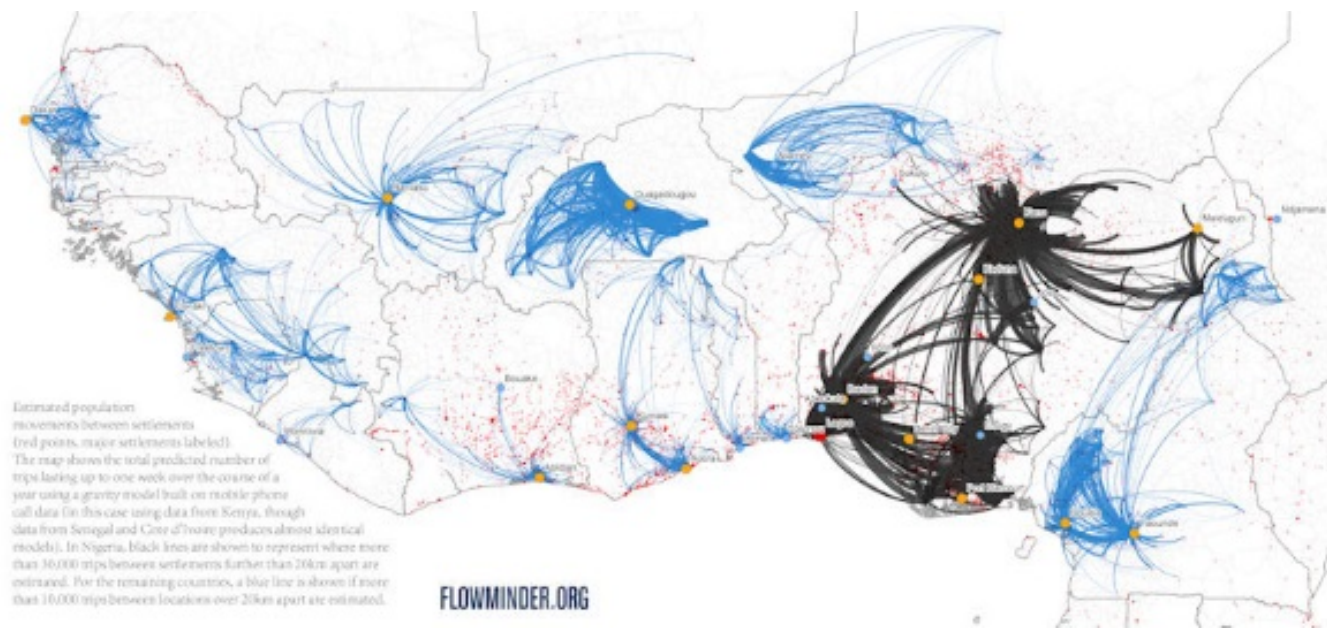
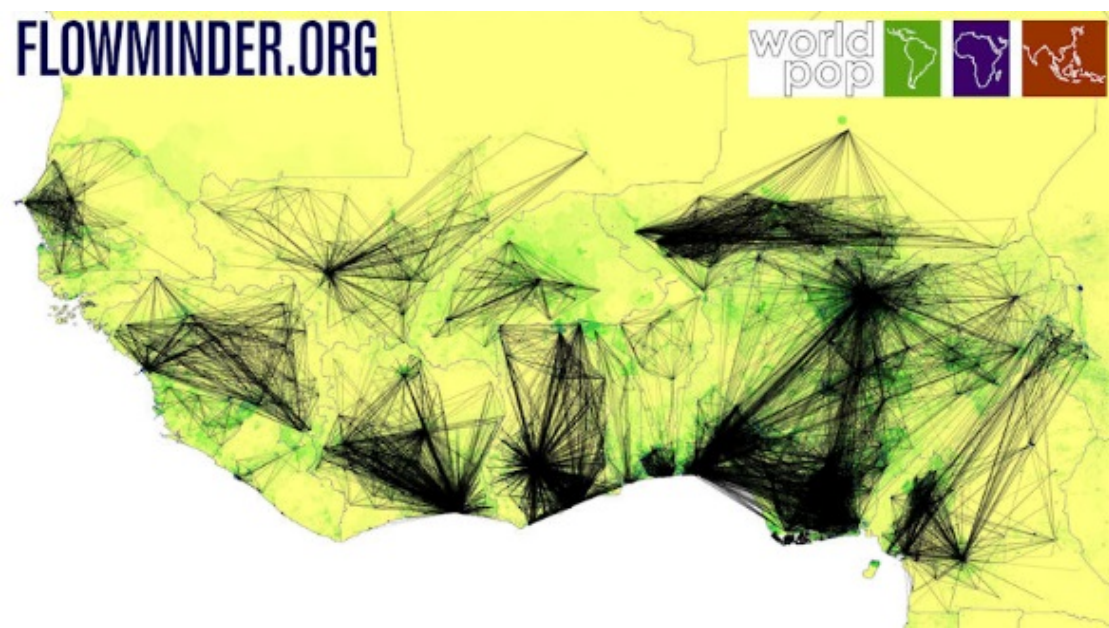


[Home](#) [Aims & Scope](#) [Review Board](#) [Authors ↓](#) [Resources](#) [About](#)

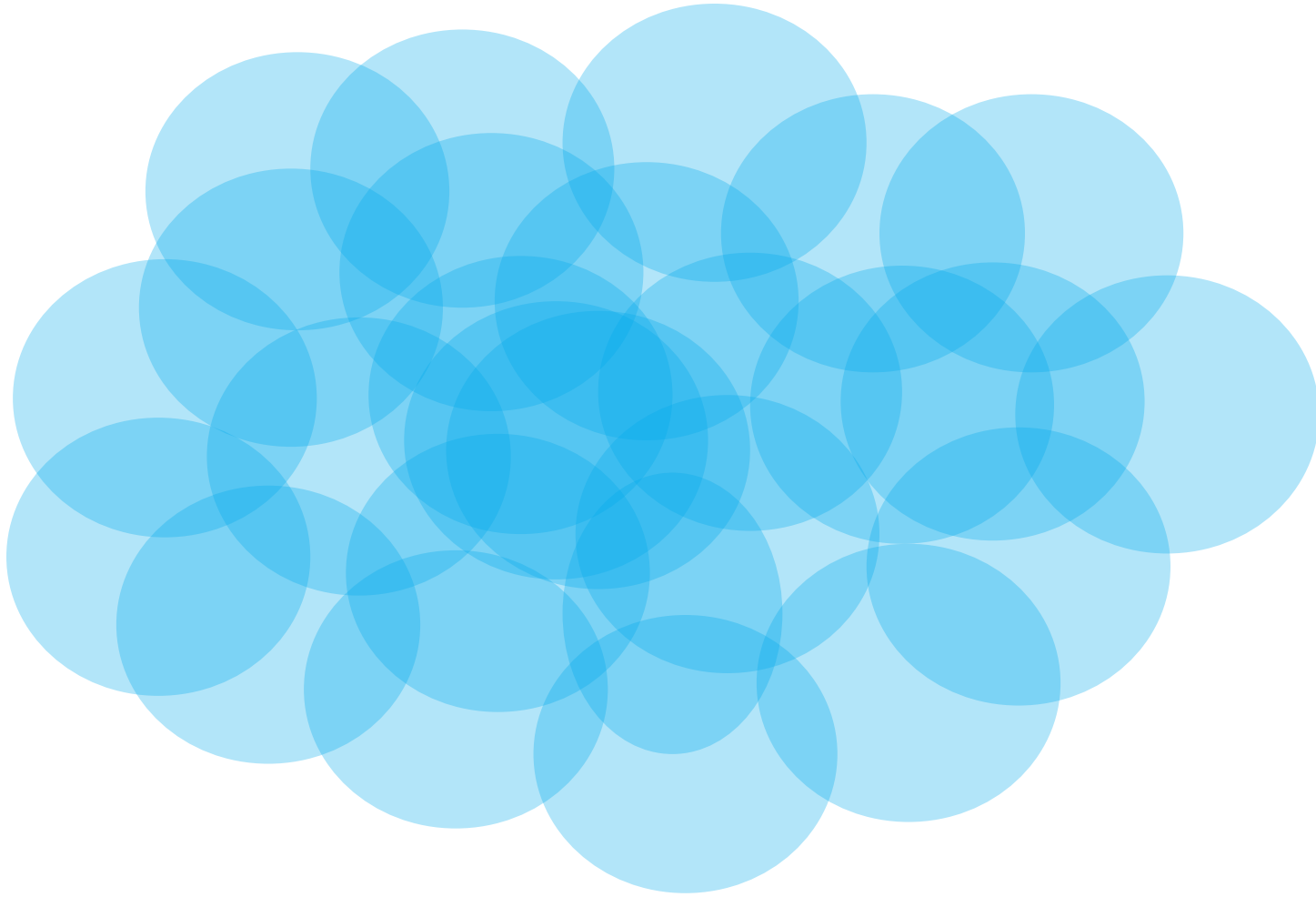
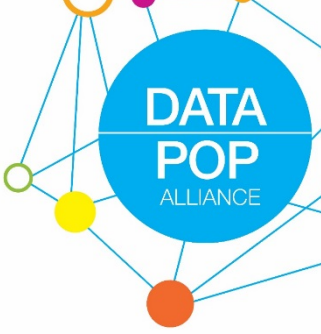
Commentary: Containing the Ebola Outbreak – the Potential and Challenge of Mobile Network Data

SEPTEMBER 29, 2014 • COMMENTARY

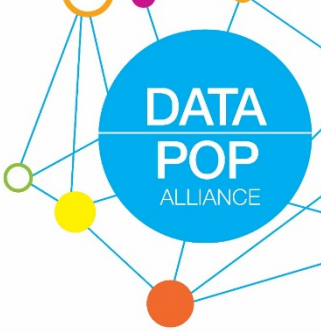
Privacy as dignity



Individuals and groups intersect; so do their privacies

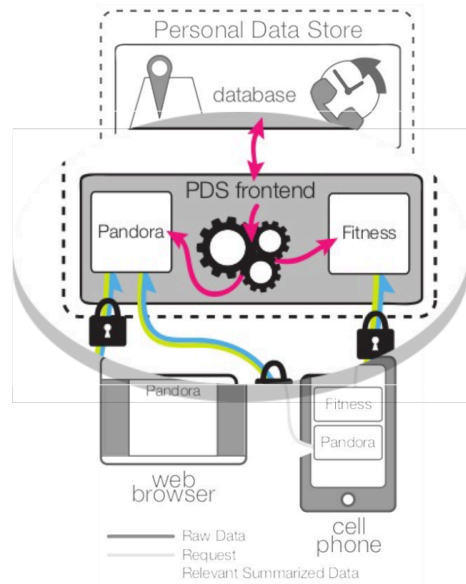


Race Between Technology, Policy & Society



Technological, Legal and Regulatory options

Dynamic systems



OpenPDS/SafeAnswers at individual level and
Differential Privacy at group level

Pro

- Keep the raw data in a safe environment
- Monitor and audit data use
- Turn a privacy problem into a security one

Cons

- New ways of working and protecting privacy
- Need for infrastructure

Use cases

- Individual-level
- Use of large-scale dataset



Data Control:
My data, our data

Data Stewardship:
Burden of responsibility
is on collector



Thank you

eletouze@datapopalliance.org

www.datapopalliance.org