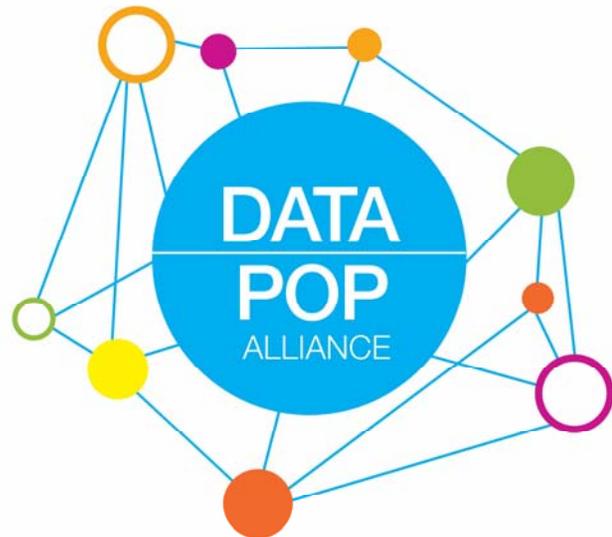


FHI 360 FIELD Conference 2014

Washington, DC.

Is Big Data the “new oil” for human development?

Reflections from the case of poverty monitoring & reduction



Emmanuel Letouzé

September 11th, 2014

Thanks!



- Paul Bundick
- Lori Reid
- Sara Seavey
- Juliene Heaney
- Lara Goldmark
-
- Alex 'Sandy' Pentland
- Participants & attendees!

Outline



1. Introduction
2. Contextualization
3. Applications
4. Implications
5. Conclusions

Introduction

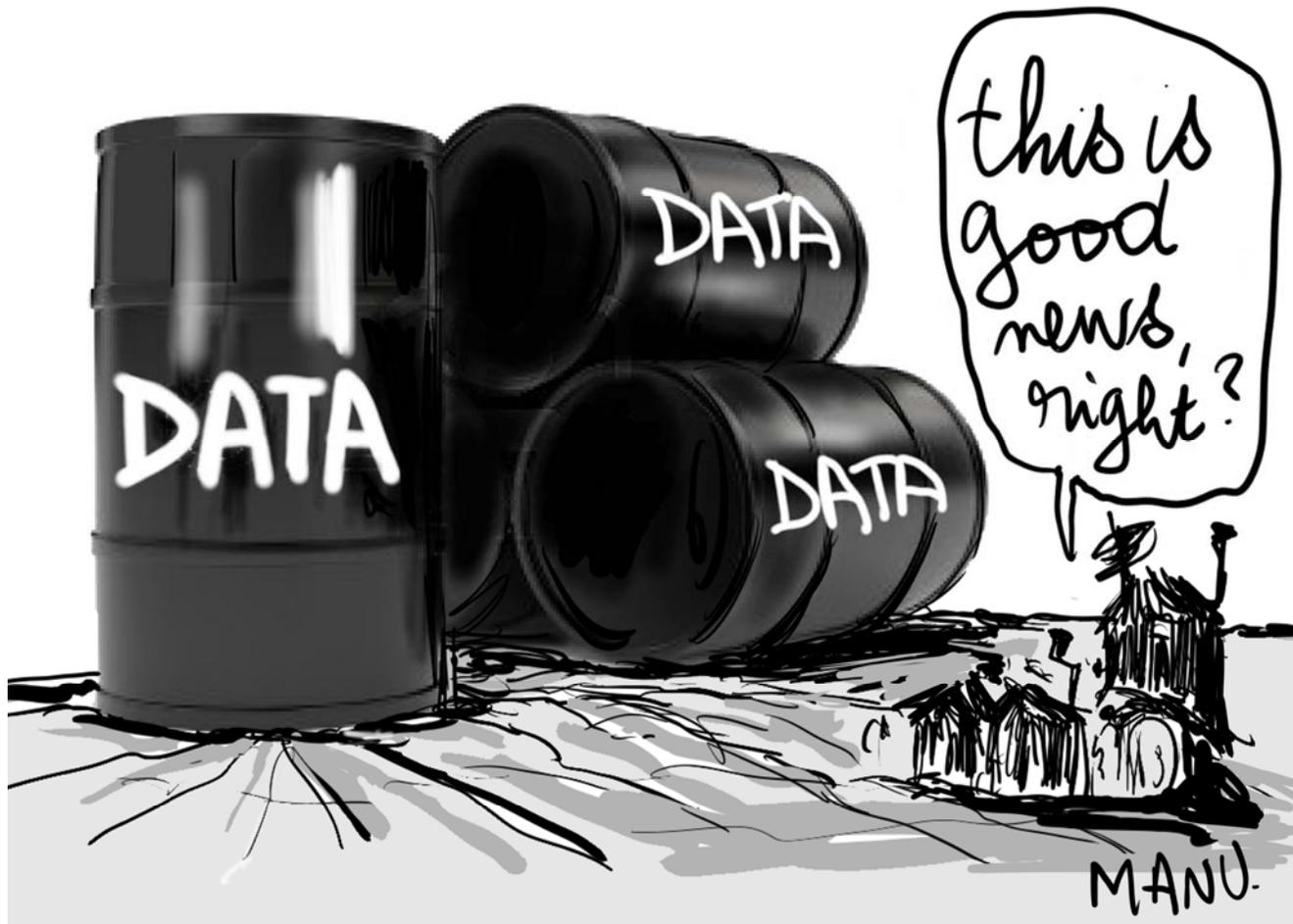


MANU.

Contextualization



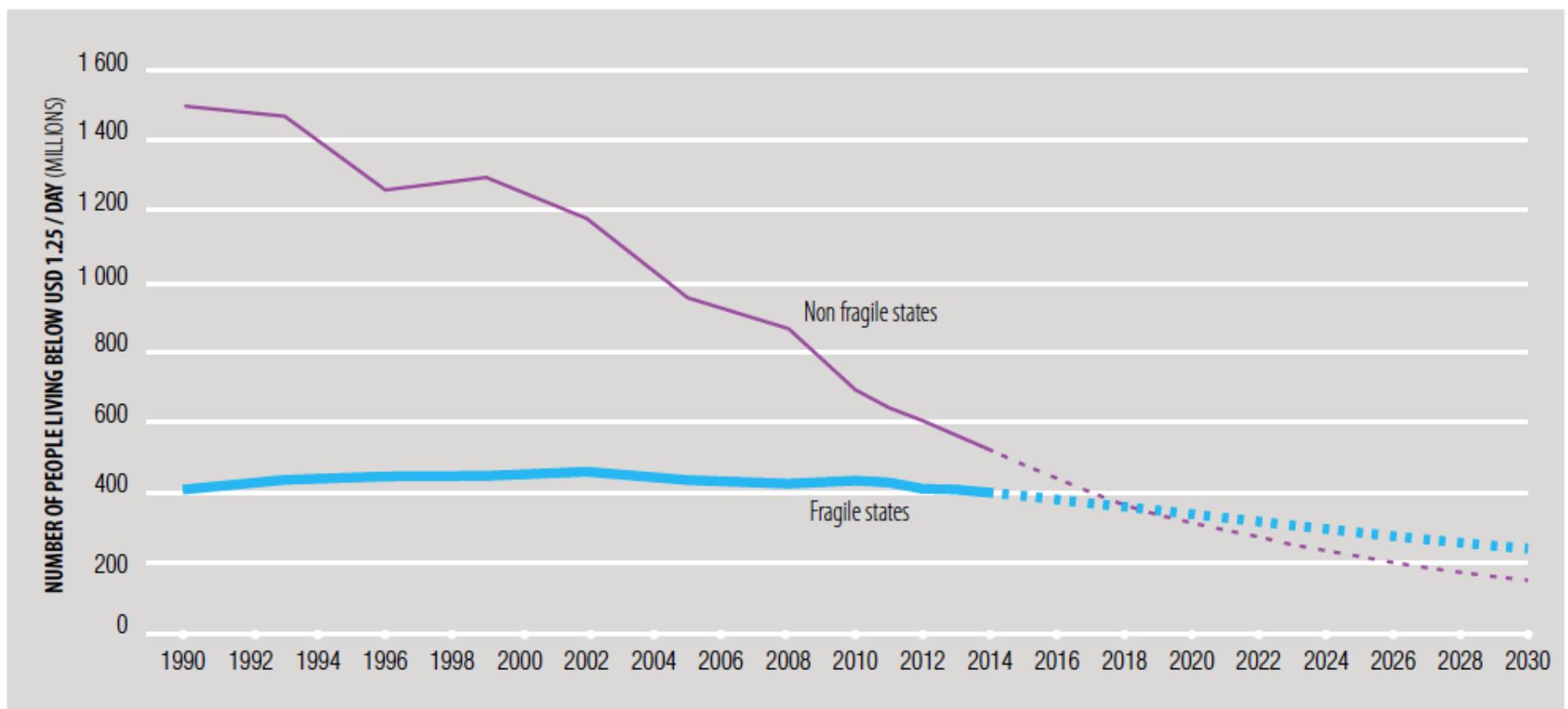
*"The bottom half of the world's population owns the same as the richest 85 people in the world"**



What do we know about poverty trends?



FIGURE 1.2 Number of people in poverty: Fragile states vs. stable developing countries, 1990-2030



Notes: The classification of countries (fragile / not fragile) is based on the list used in the 2013 Fragile States Report; classifications across years are held constant.

Source: Chandy, L., N. Ledlie and V. Penciakova (2013), *The Final Countdown: Prospects for Ending Extreme Poverty by 2030* (interactive), April 24, 2013 Brookings Institution, Washington DC, available at www.brookings.edu/research/interactives/2013/ending-extreme-poverty

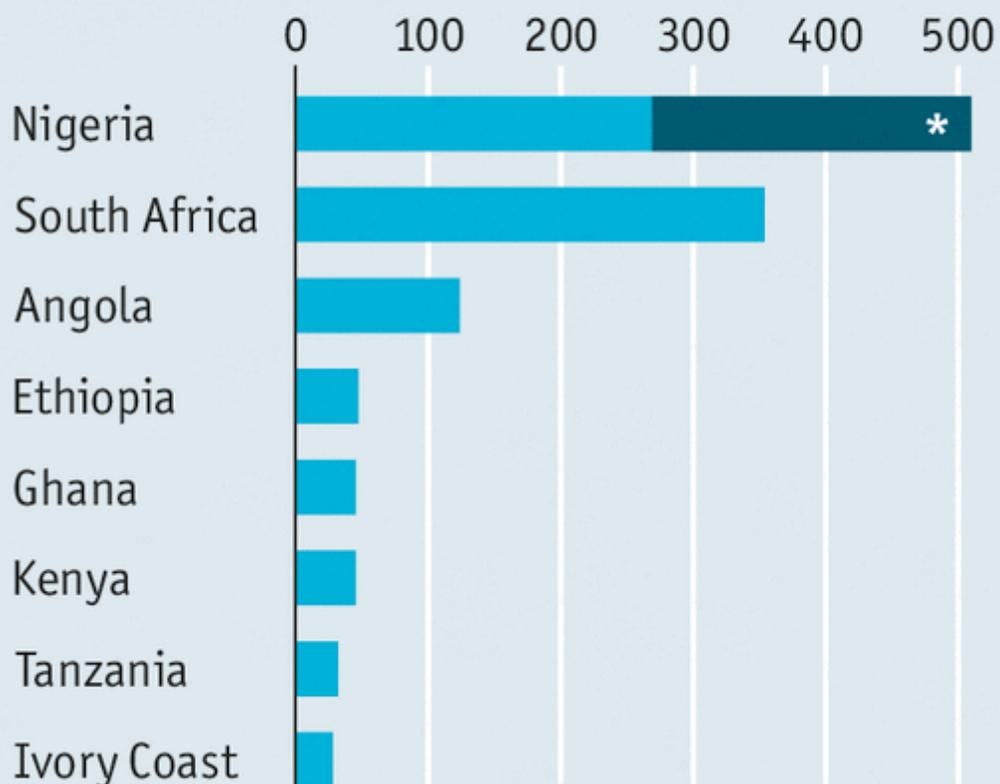
Source: Letouzé, 2014, based on primary and secondary sources



Vaulting ahead

Biggest sub-Saharan economies

GDP, 2013 estimates, \$bn

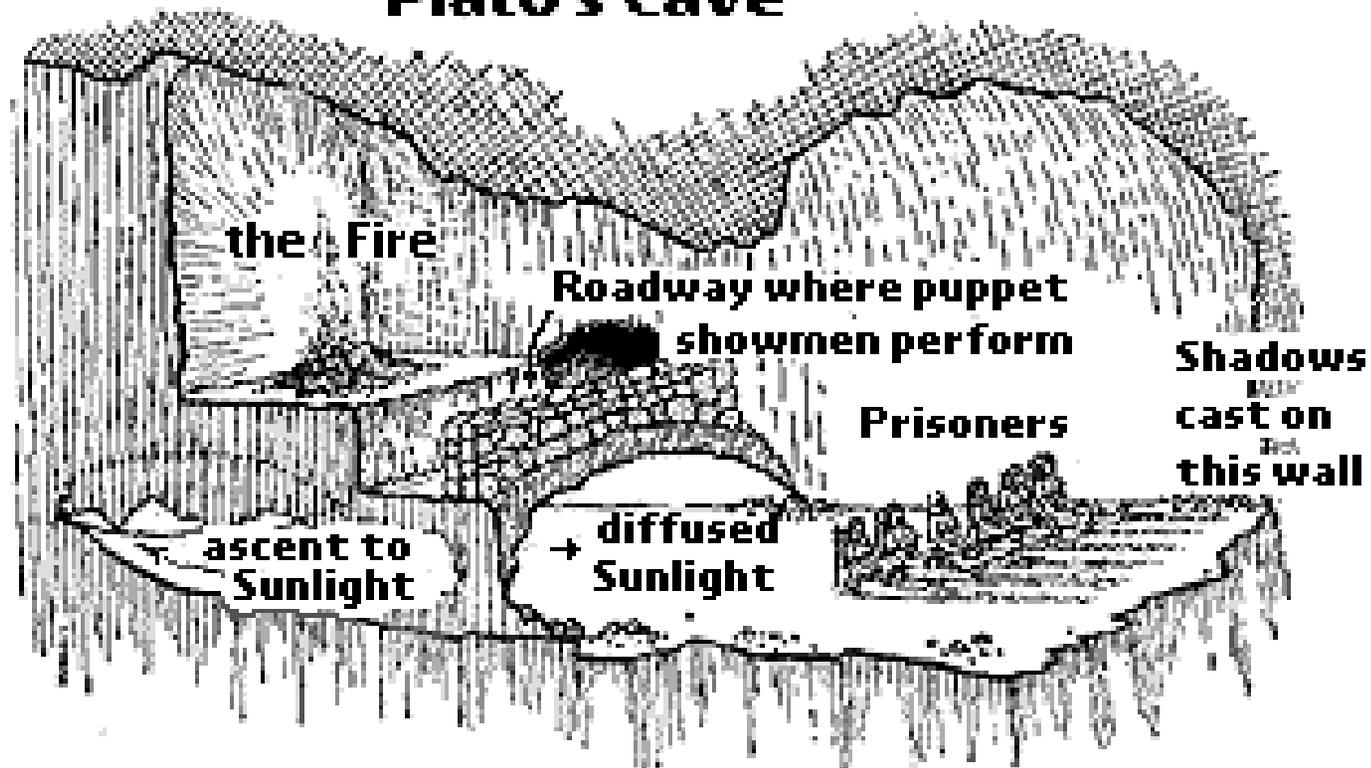


Sources: IMF; national statistics

*After revision



Plato's Cave



Are official stats more than shadows in the cave?

Africa's statistical tragedy



SUBMITTED BY **SHANTA DEVARAJAN** ON THU, 10/06/2011



Marcelo Giugale

World Bank's Director of Economic Policy and Poverty Reduction Programs for Africa

GET UPDAT



Fix Africa's Statistics

December 18, 2012

Cell-Phone Data Might Help Predict Ebola's Spread

Mobility data from an African mobile-phone carrier could help researchers recommend where to focus health-care efforts.

By David Talbot on August 22, 2014

Big data: The next frontier for innovation, competition, and productivity

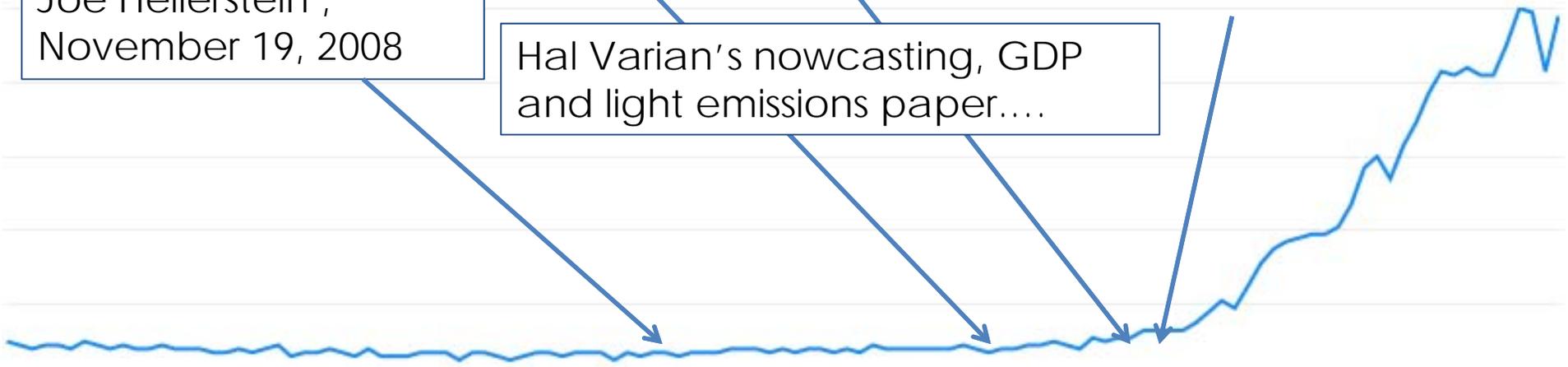


Big Data, Big Impact:
New Possibilities for International Development

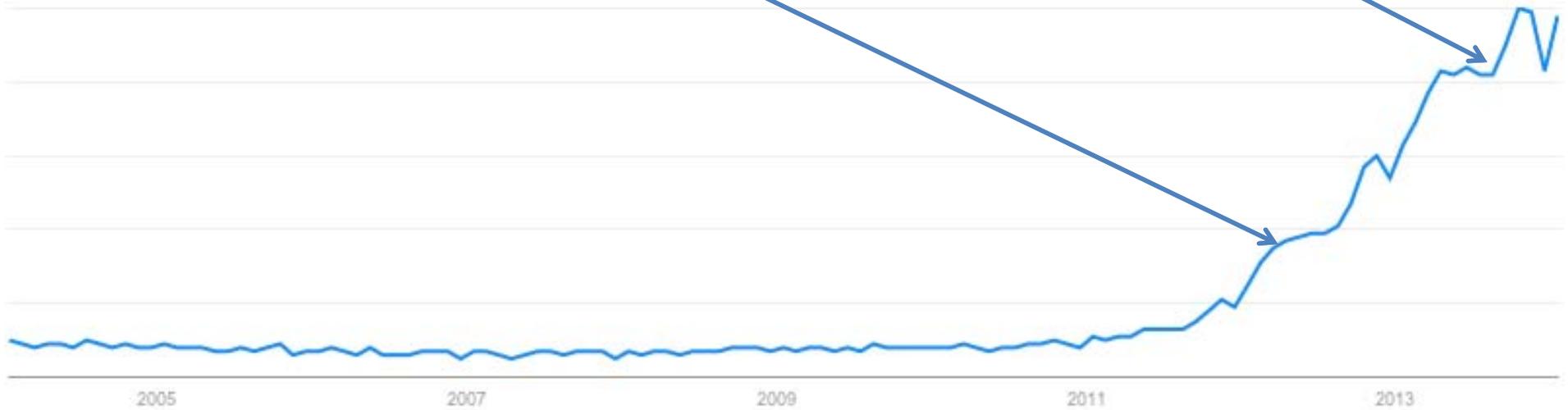
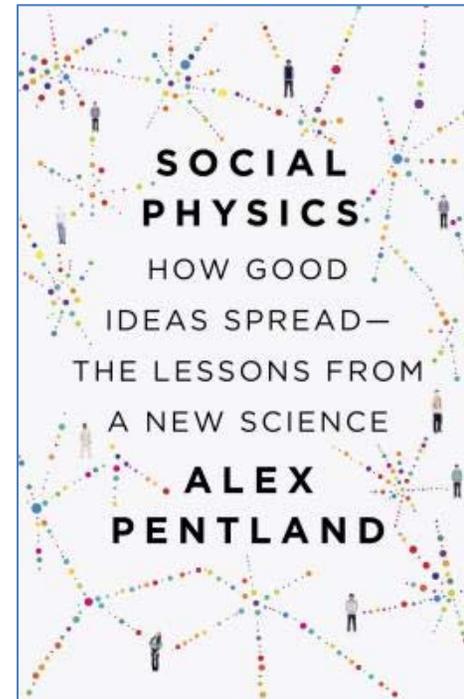
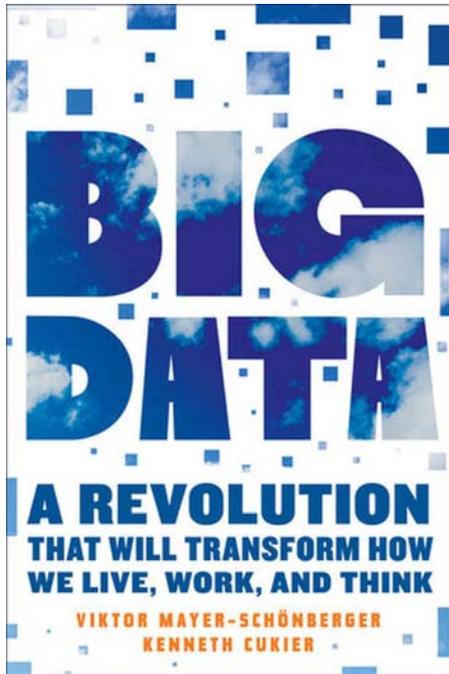
*"We are at the beginning of what I call The **Industrial Revolution of Data.**"*
Joe Hellerstein ,
November 19, 2008



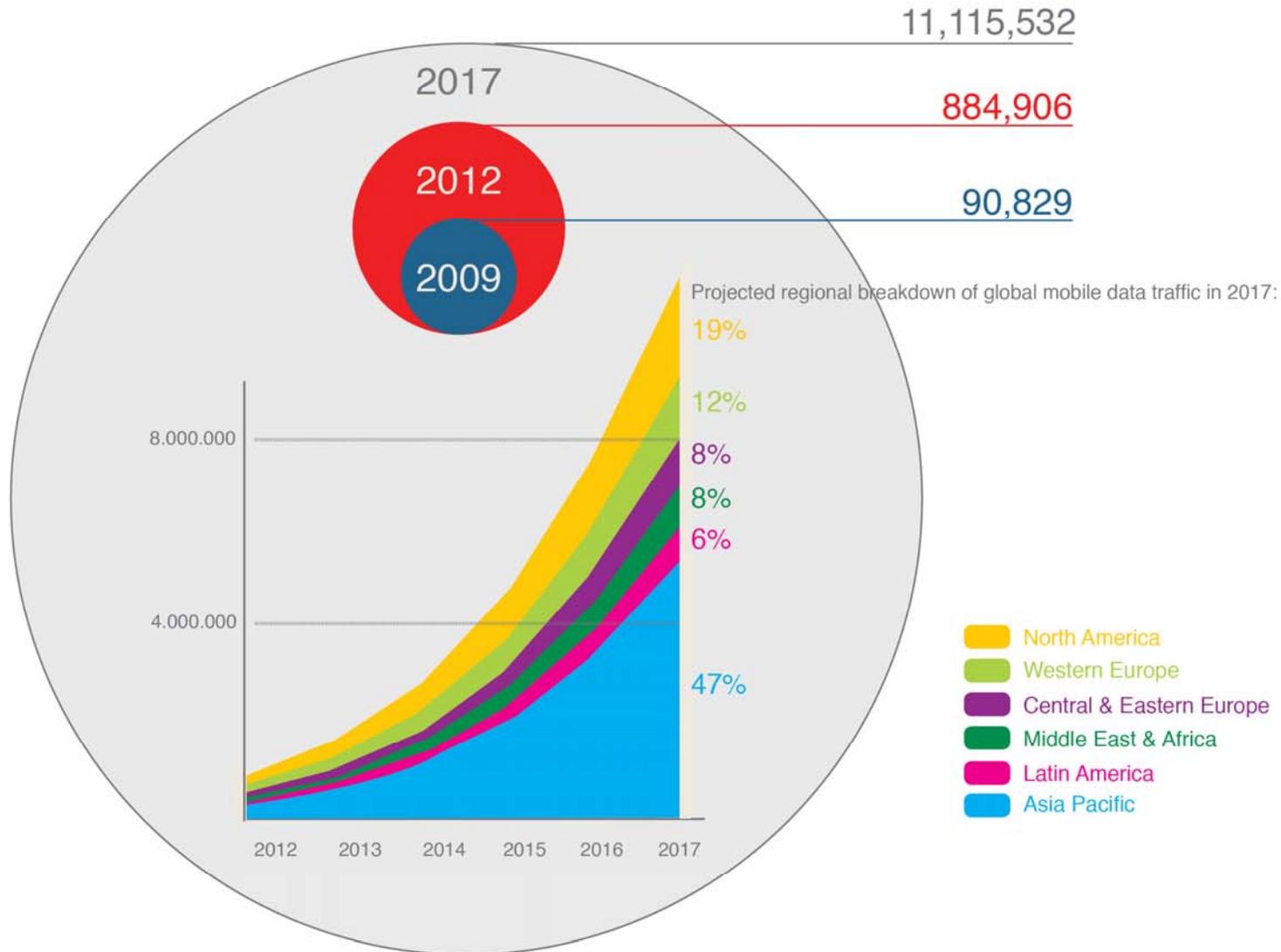
Hal Varian's nowcasting, GDP and light emissions paper....



Line shows returns for "Big Data" on Google Trends between 2007 and 2014; 100=maximum value



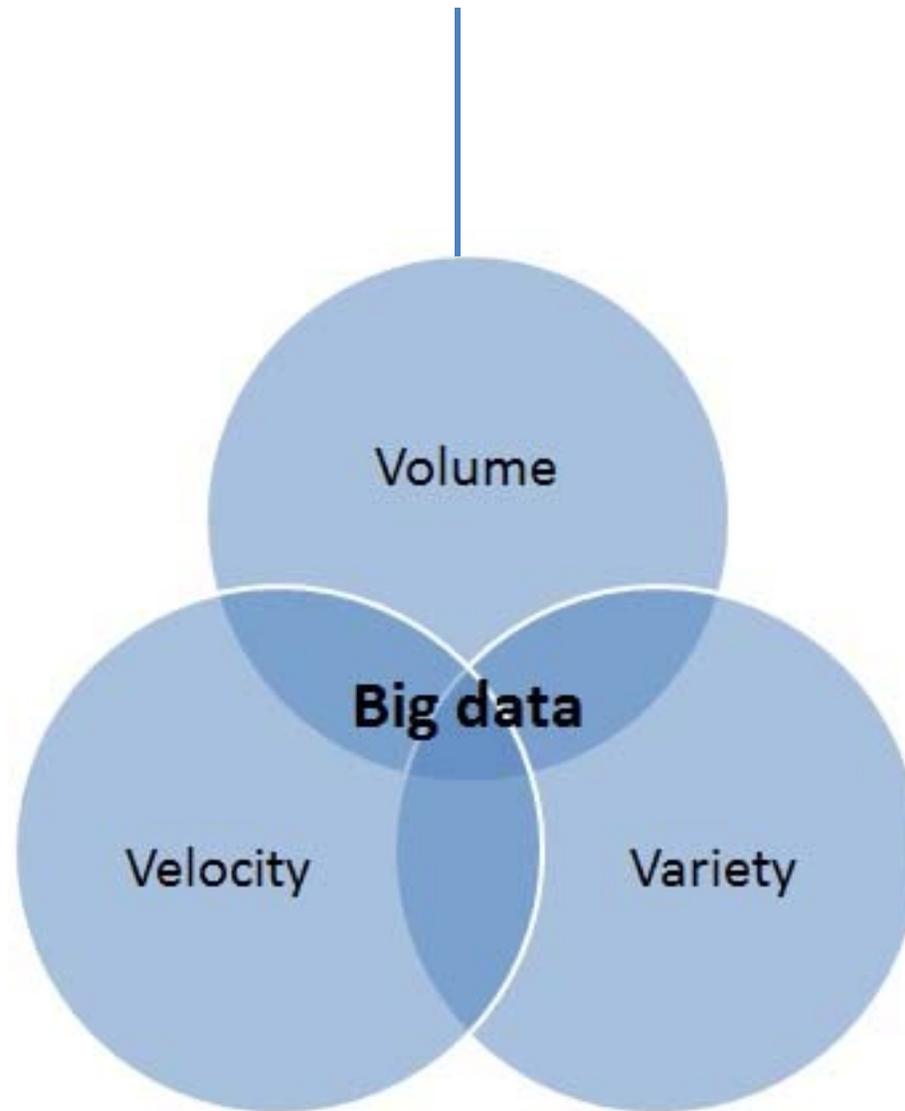
Global Mobile Data - Traffic growth & forecast (terabytes per month)





Zettabyte (ZB)	1,000EB, or 2^{70} , bytes	It is estimated that in 2013, humanity generated 4-5ZB of data, which exceeds the quantity of data in 46 trillion print issues of <i>The Economist</i> . If that many magazines were laid out sheet by sheet on the ground, they would cover the total land surface area of the Earth.
----------------	------------------------------	--

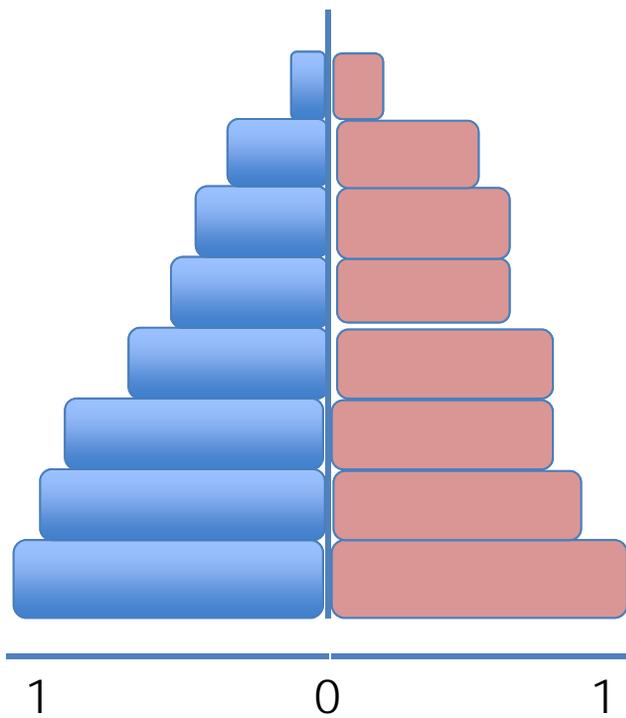
Circa 2010-12: the 3 Vs of Big Data



Stock of world data?

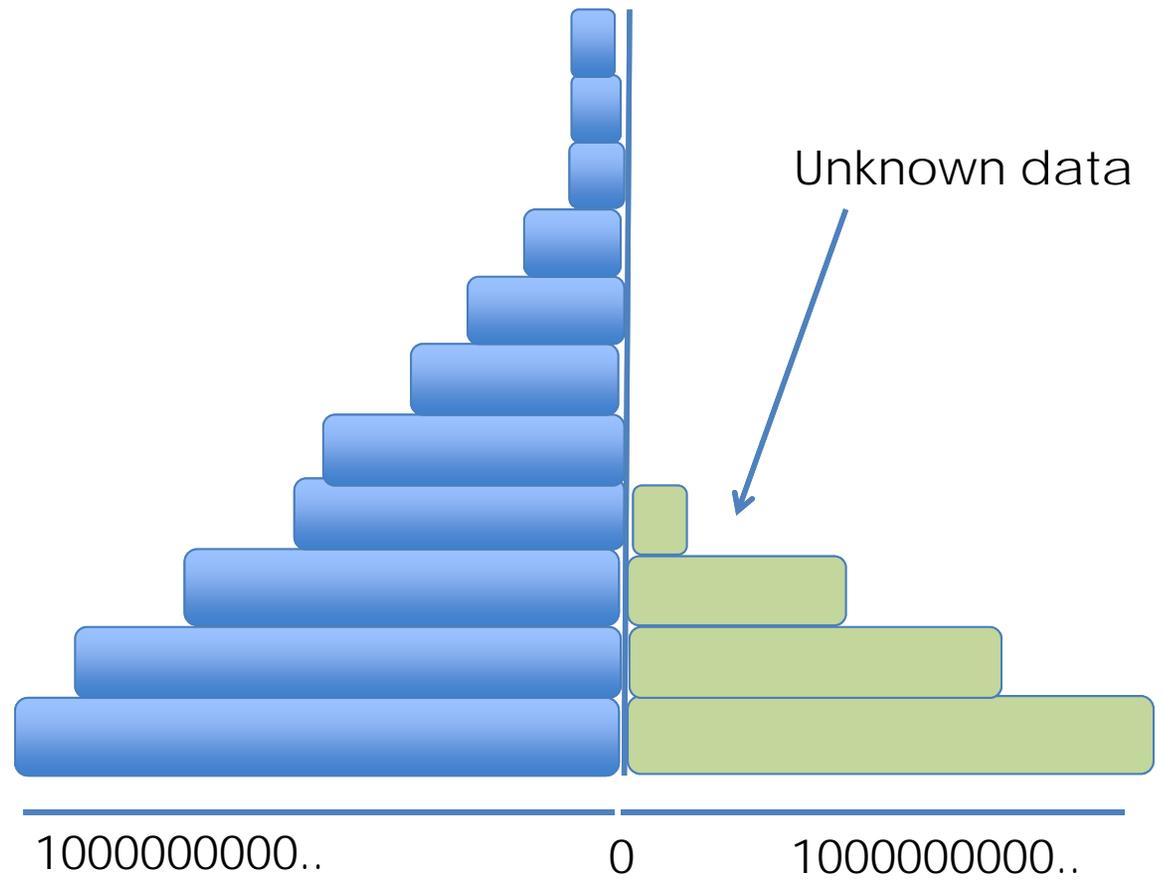


50 years



Circa
2000

90 days



Circa
2050

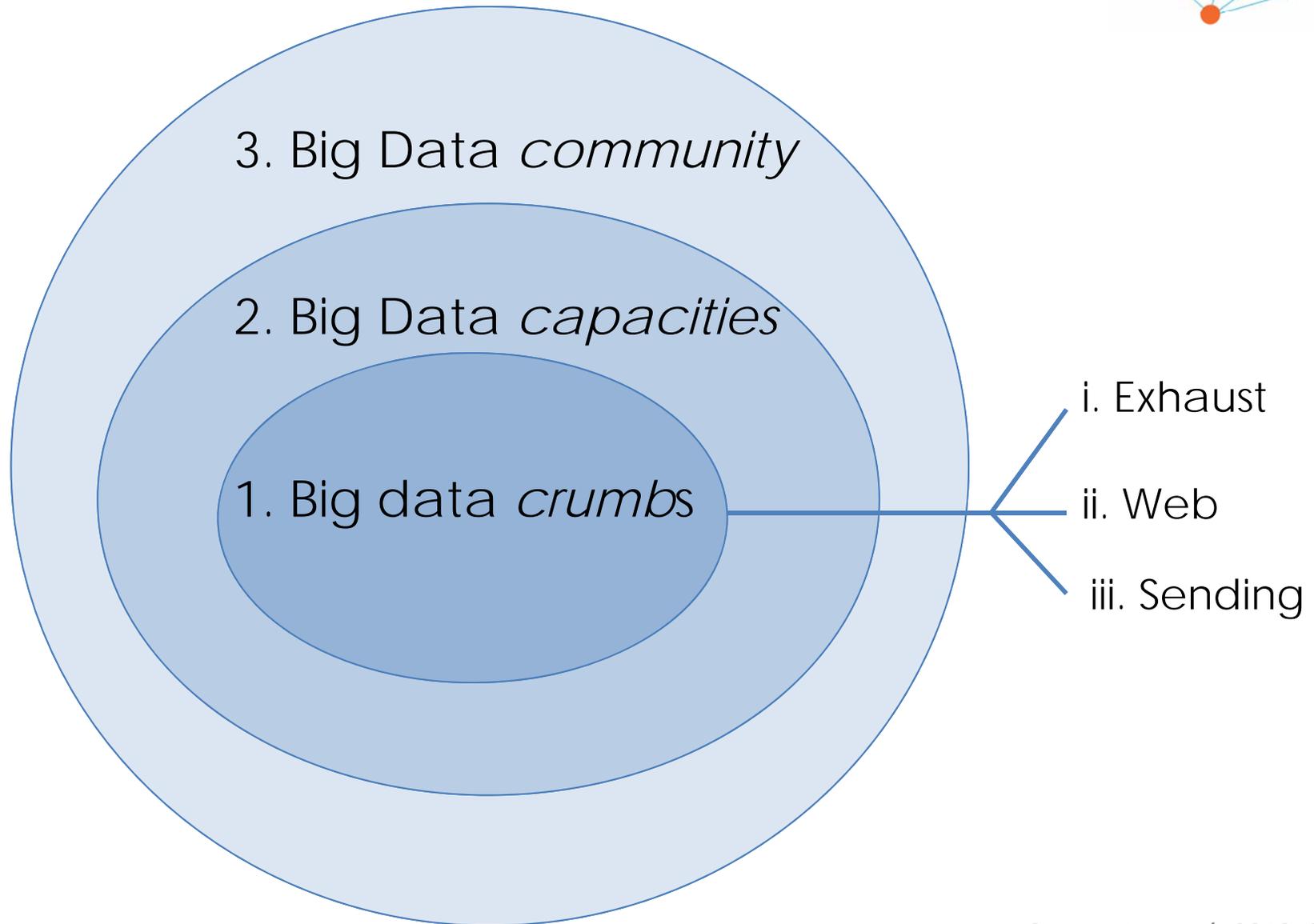


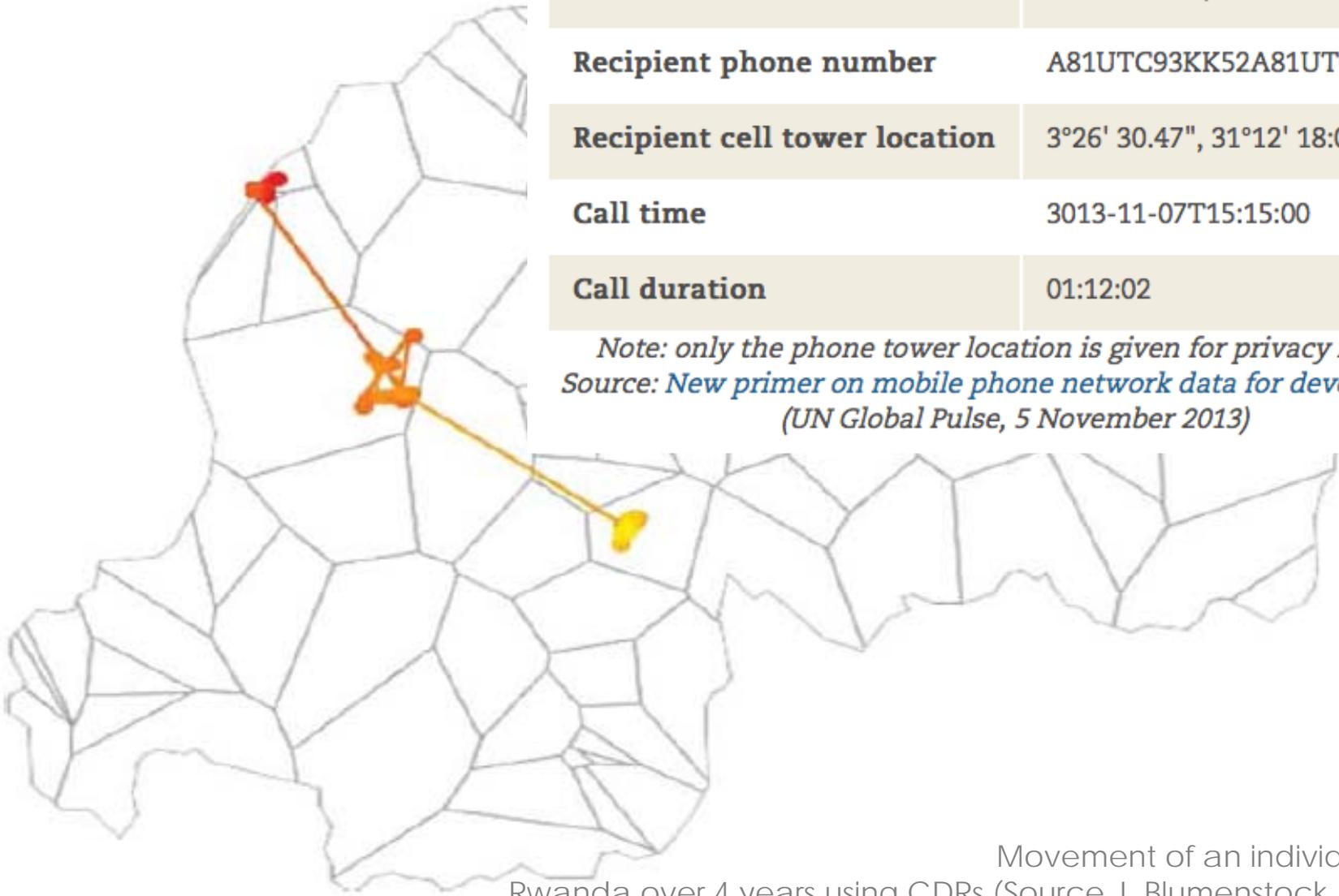
Big Data is Not About the Data!

Gary King¹

Institute for Quantitative Social Science
Harvard University

Now: the 3 Cs of Big Data





Variable	Data
Caller ID	X76VG588RLPQ
Caller ID tower location	2°24' 22.14" , 35°49' 56.54
Recipient phone number	A81UTC93KK52A81UTC93KK52
Recipient cell tower location	3°26' 30.47" , 31°12' 18:01"
Call time	2013-11-07T15:15:00
Call duration	01:12:02

*Note: only the phone tower location is given for privacy reason.
 Source: New primer on mobile phone network data for development.
 (UN Global Pulse, 5 November 2013)*

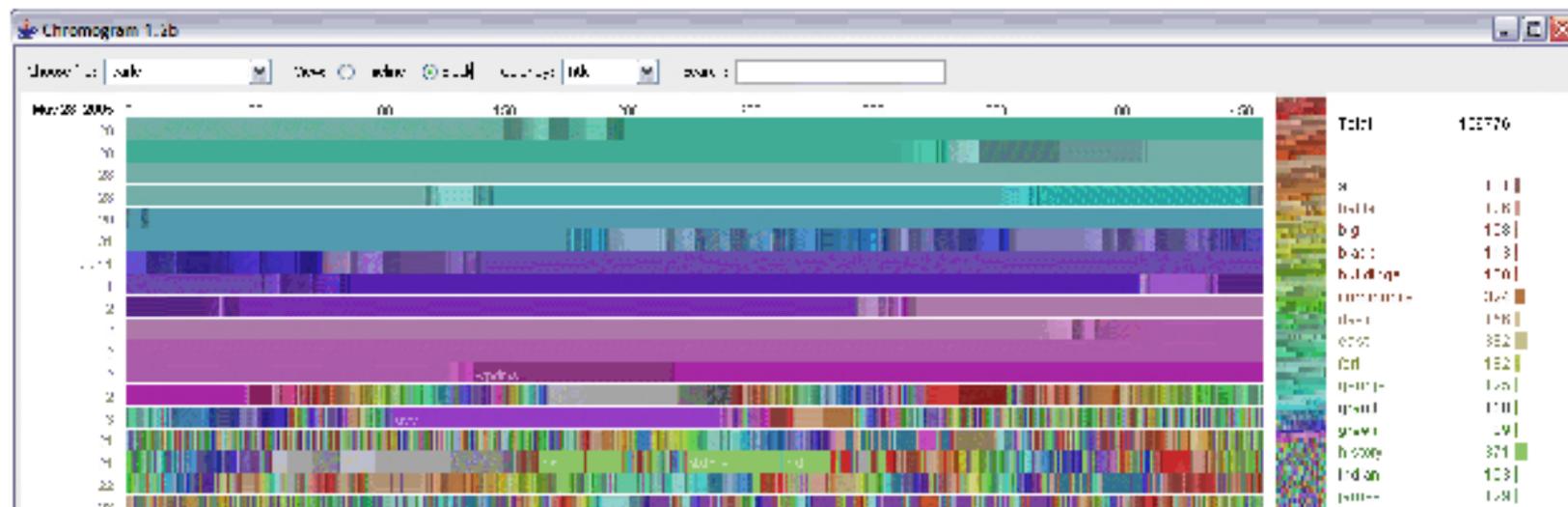
Movement of an individual in Rwanda over 4 years using CDRs (Source J. Blumenstock, 2010)



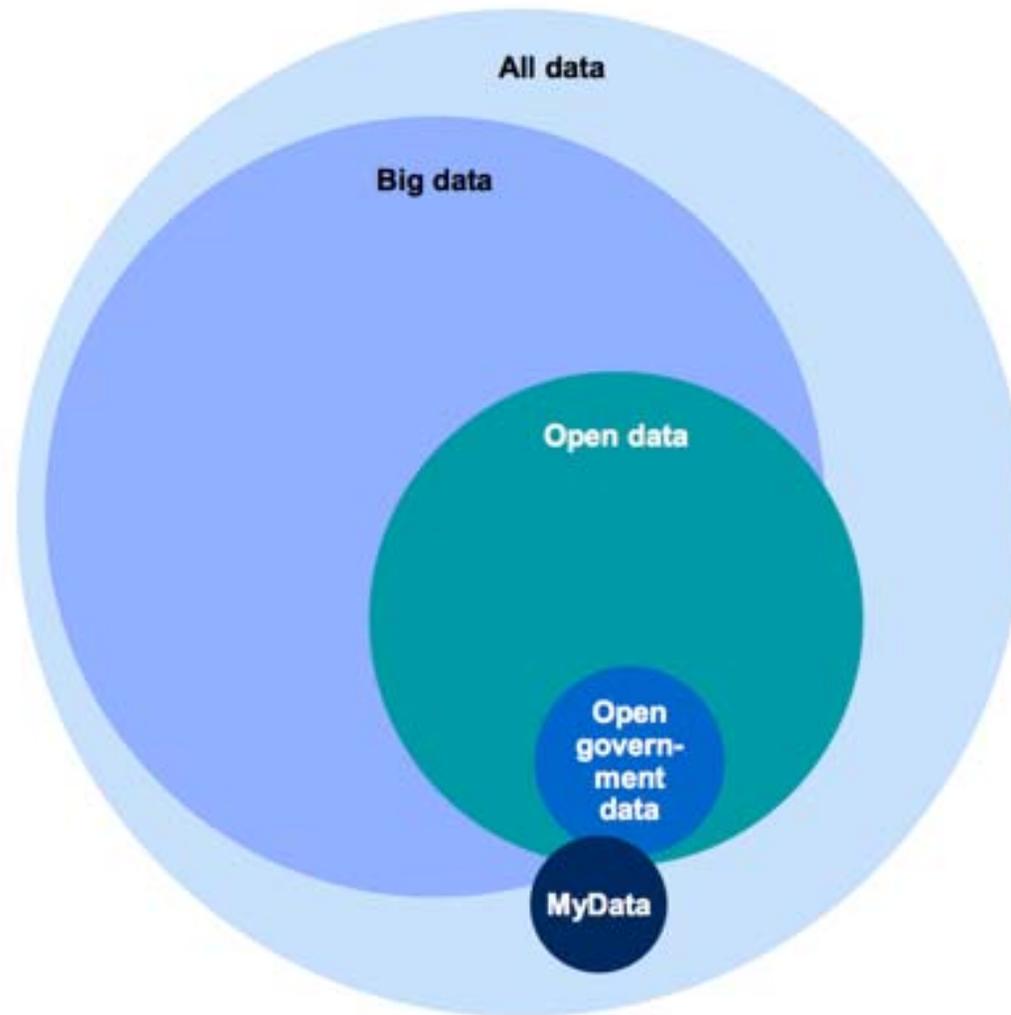
Think Again: Big Data

Why the rise of machines isn't all it's cracked up to be.

BY KATE CRAWFORD | MAY 9, 2013



The new data ecosystem



SOURCE: McKinsey Global Institute analysis

Applications



Taxonomy of applications (1):

1. Early warning
1. Real time awareness
1. Real-time feedback



Source: Letouzé, 2012



Taxonomy of applications (2):

1. Descriptive
1. Predictive
1. Prescriptive



Source: Letouzé, Vinck and Meier, 2013

BLOG POST

Could Big Data provide alternative measures of poverty and welfare?

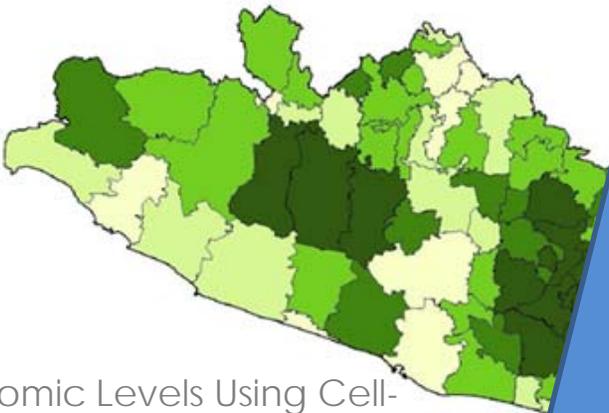


Census Variables	
Variable Type	Description
Education	% of Population with Primary School
	% of Female Population with Primary School
	% of Male Population with Primary School
	% of Population with Secondary School
Demography	% of Senior Population (> 60)
	% of Senior Population (16 - 60)
Goods	% of Houses with Ce
	% of Houses with 1 r
	% of Houses with 3+
	% of Houses with Ele
	% of Houses with Wa
	% of Houses with TV
	% of Houses with PC
	% of Houses with All
SEL	Socio-Economic Level

Survey from "a major city in Latin America"

Telefonica team used their data to 'predict' SELs from Cell Phone Usage

Predict the present (SELs for non-surveyed regions) and monitor the future (track changes over time)



Source: "Prediction of Socio-Economic Levels Using Cell-Phone Records" (Telefonica research, 2011)

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?

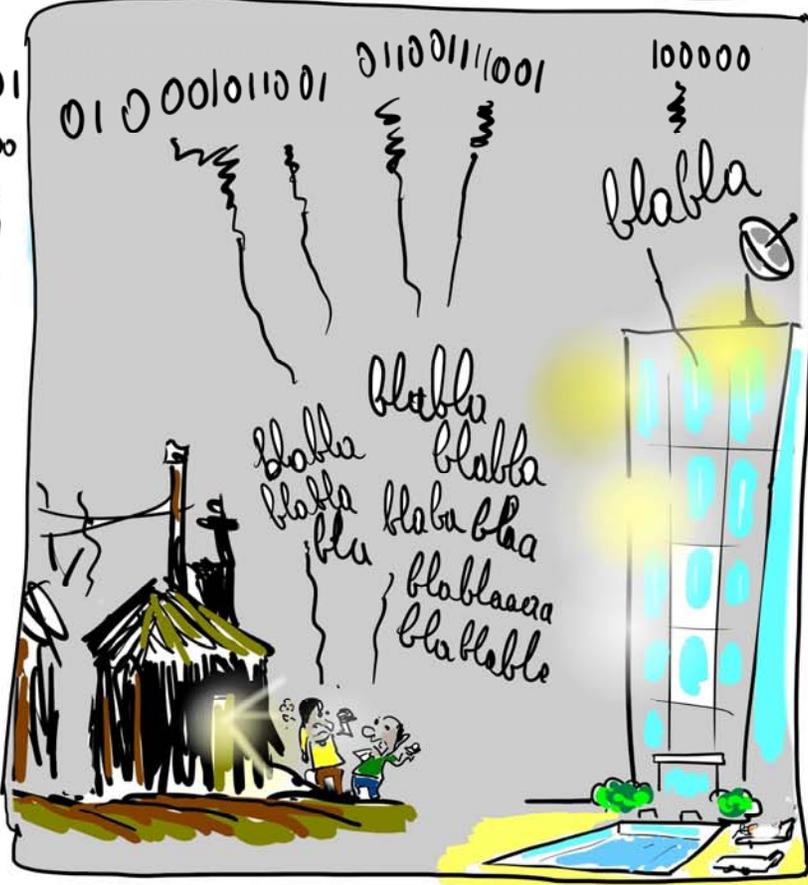
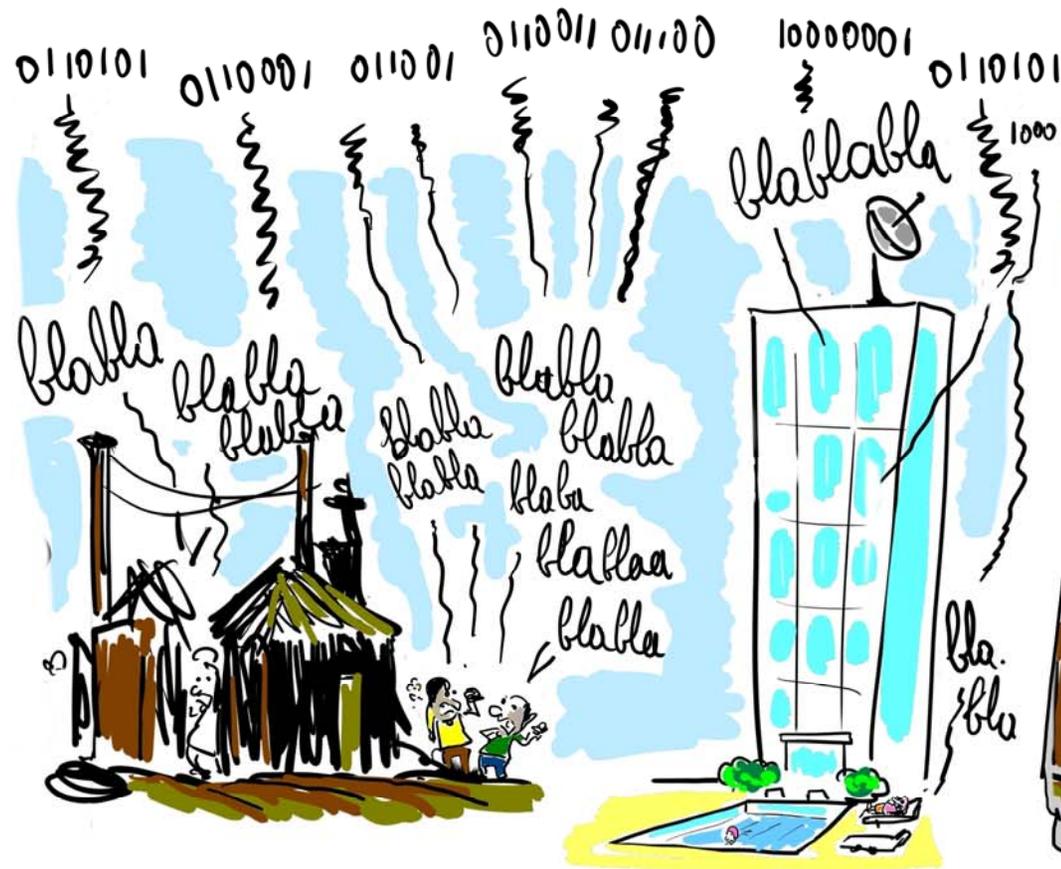




step 2



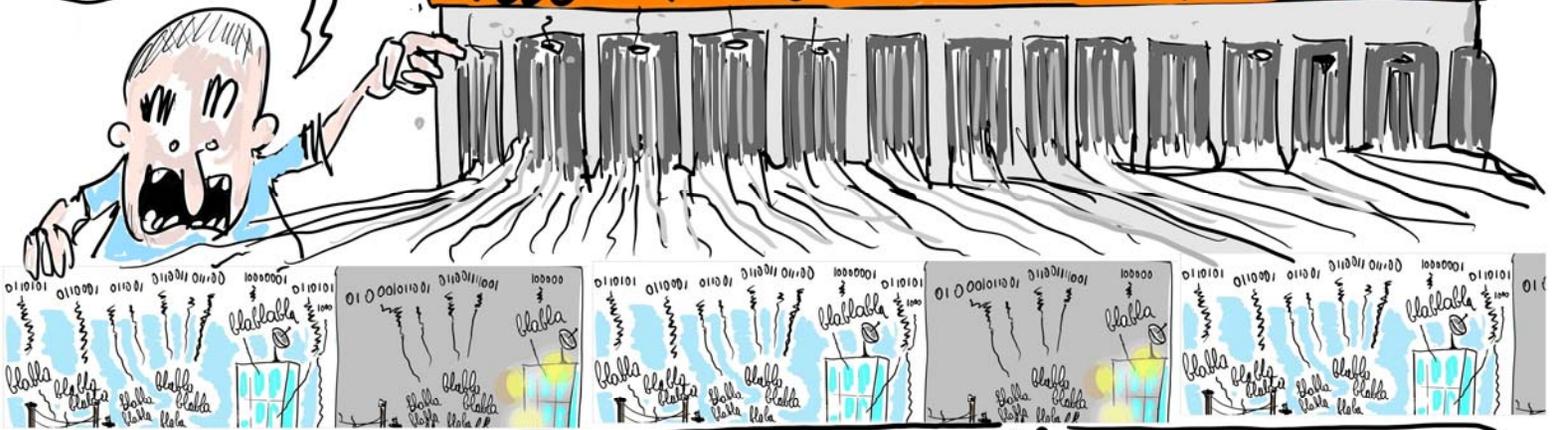
then notice how cell phone users leave digital traces, day & night..



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

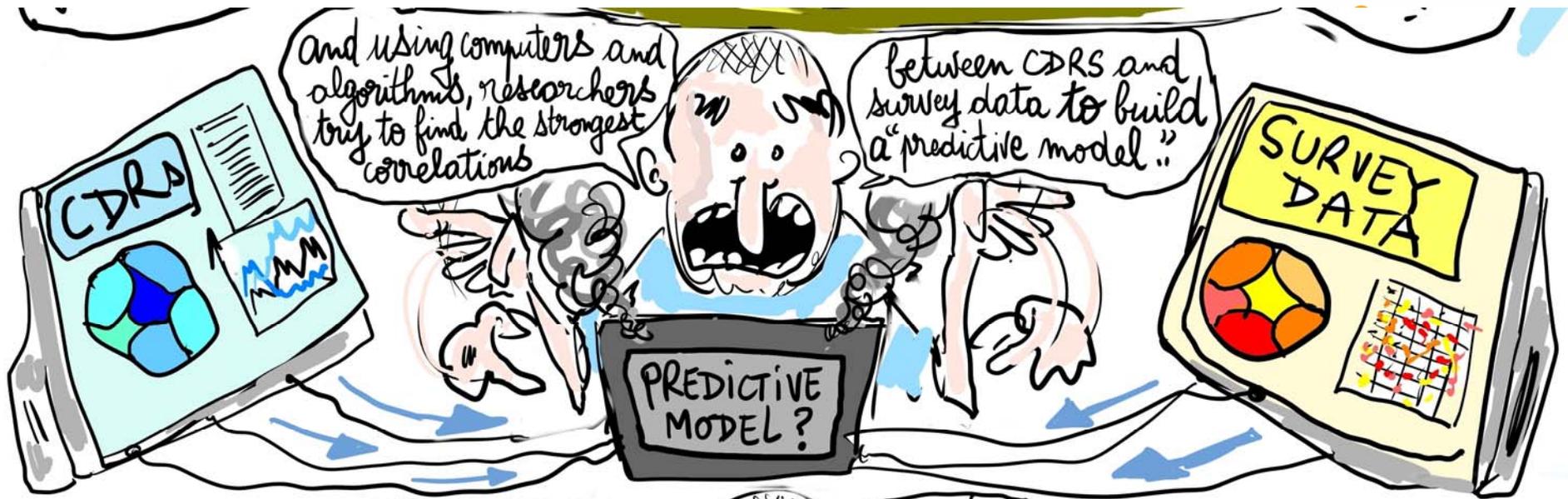
CALLER ID	CALLER LOCATION	RECIPIENT ID	RECIPIENT LOCATION	CALL TIME	CALL DURATION
X36872	2°24'22"	A8C492	3°38'49"	2014.04.01	01.12.27
9748Y	35°49'58"	TC73646	31°12'22"	ET 17 22	

TELECOM OPERATOR DATA CENTER



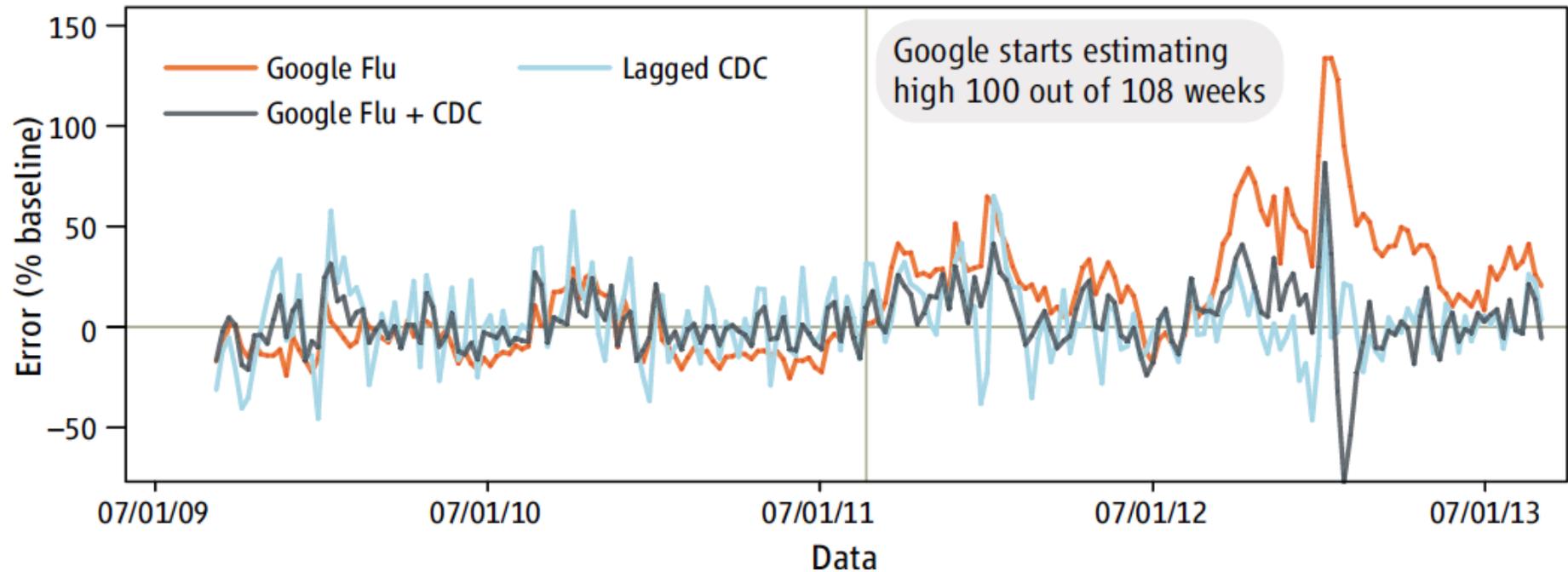
"and these CDRs will show differences in calling patterns between different areas ..."







(Another) great case re opportunities vs. challenges: Google Flue Trend's (over) prediction



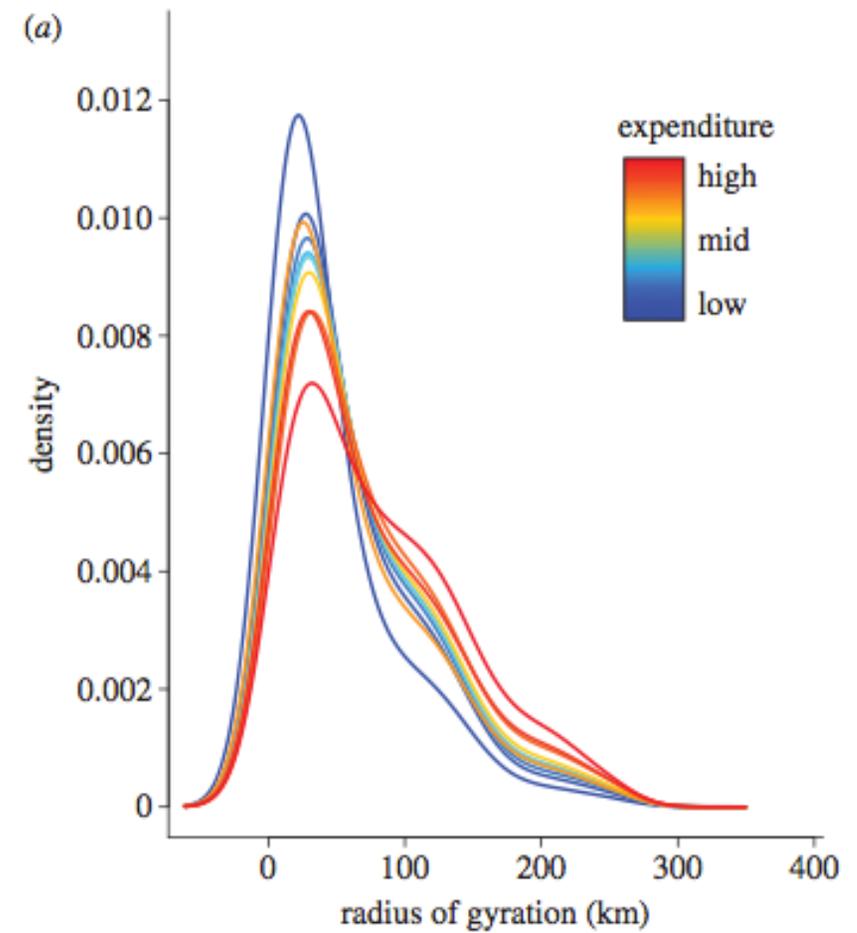
Source: Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. 2014. The Parable of Google Flu: Traps in Big Data Analysis. *Science* 343, no. 14 March: 1203-1205.



Technico-methodological frontier:

(1) Sample bias correction

Ex: Buckee, Eagle et al (2013):
"The impact of biases in mobile phone ownership on estimates of human mobility"

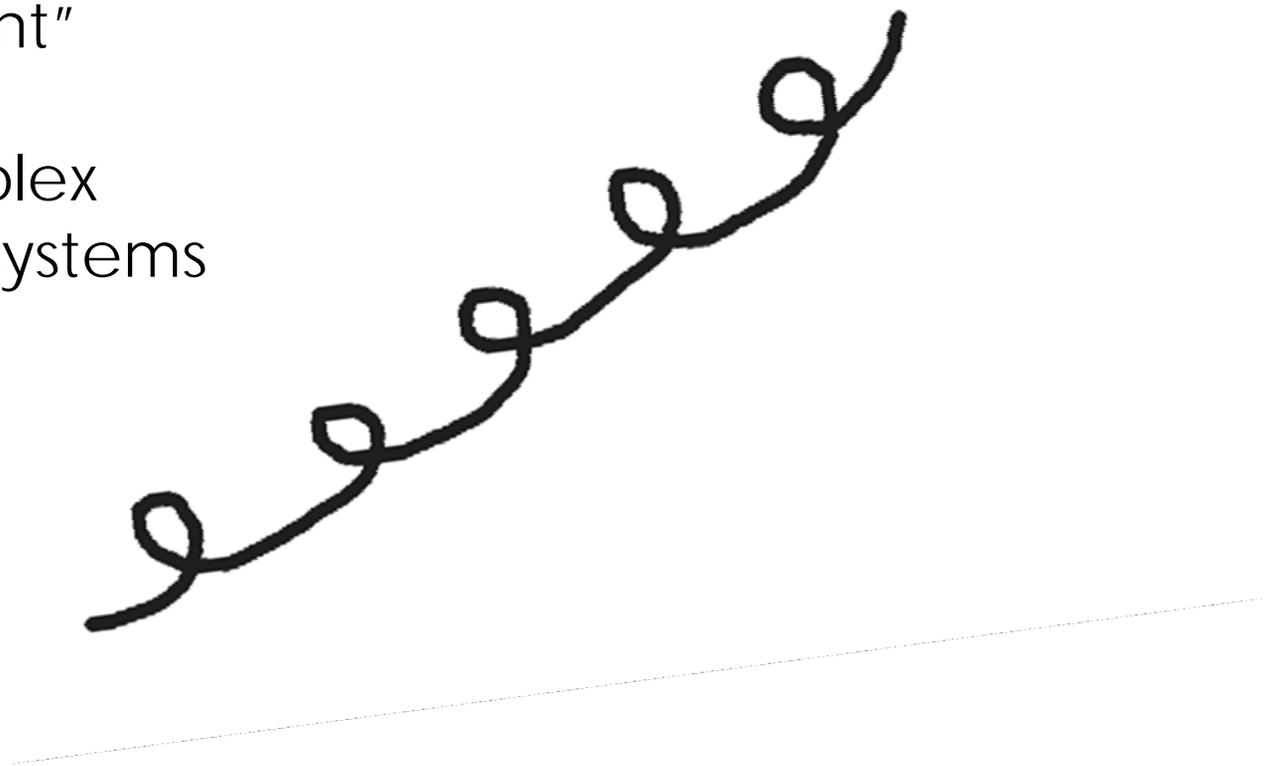




Technico-methodological frontier:

(2) Causal inference for
"agile development"

Ex: M&E and complex
interventions and systems



Implications



Trust
me

i'm
a data scientist

Implications





Technocratic & humanistic/politics

A development data revolution needs to go beyond the geeks and bean-counters

Will better data really lead to better development policies? Only if the right people have access and use it to make better policies

Jonathan Glemmie, The Guardian, Oct 3, 2013



“Official statistics assumes a key role in ensuring democracy and fostering social progress... [should] provide society with knowledge of itself”

Enrico Giovanni—former President of Istat

“ Knowledge is power; statistics is democracy”
—Former President of Statistics Finland



Post2015.org

what comes after the MDGs?

A hub for ideas, debate and resources on what comes after the Millennium Development Goals (MDGs)

SEARCH

Search

[Home](#) [About](#) [Contact](#)

[Home](#) [About](#) [Resources](#) [Opinion](#) [Events](#) [News](#) [Goals Analysis](#) [Data Revolution](#) [Southern Voices](#)

DATA REVOLUTION

The Big Data revolution should be about knowledge security

Posted by post2015 · April 1, 2014 · [Leave a comment](#)

FILED UNDER [big da](#), [data in development](#), [data revolution series](#)

Written by Emmanuel Letouzé, Fellow at the Harvard Humanitarian Initiative, Visiting Scholar at MIT Media Lab, Research Associate at ODI and PhD Candidate at UC Berkeley; he is also the co-founder and director of Data-Pop, co-created by HHI, MIT Media Lab and ODI.

In my view, Big Data can fuel a Data Revolution. Much of its appeal stems from its potential — true or false — to find and refine data to yield '[insights](#)' about human populations that can power more agile and better targeted policies and programmes. I have at least three issues with this general line of reasoning, and propose instead a "knowledge security-centred" approach to Big Data for human development.

Countdown

January 1st, 2016

20

months to go.

Site facilitated by ODI



Conclusion



“A ‘true’ big data revolution should be one where data can be leveraged to change power structures and decision-making processes, not just create insights.”

Emmanuel Letouzé

SciDev.Net's [Spotlight on Big Data for Development](#), April 15th, 2014