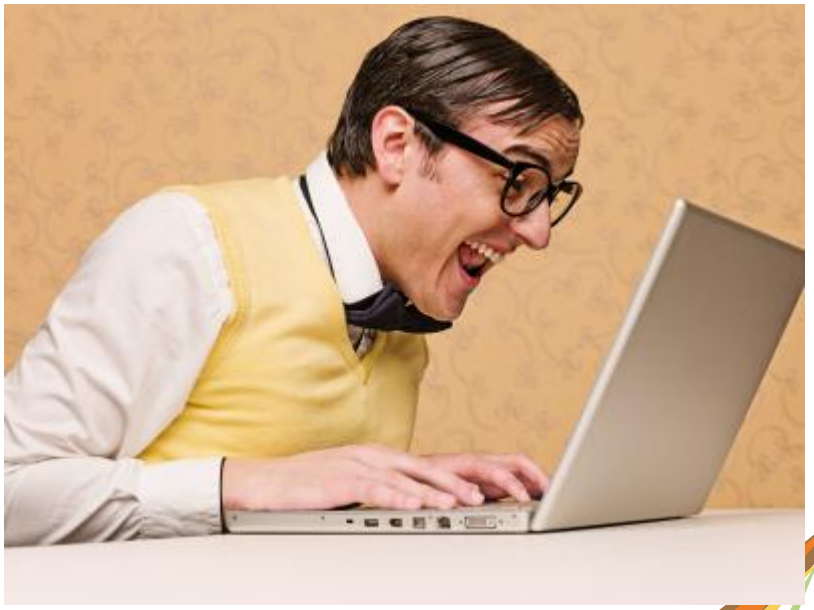
Is My Data Revolution Like Your Data Revolution?

Johannes Jütting
El-Iza Mohamedou

PARIS 21
Partnership in Statistics for Development in the 21st Century

What we are NOT

- 21st district in Paris
- (only) statistical nerds
- OECD - DAC



PARIS21

Partnership in Statistics for Development in the 21st Century

Global Partnership fostering national statistical capacities

- Coordinating and convening users and producers of development data

Strengthening statistical systems for better data

- Promoting National Strategies of the Development of Statistics and increasing the availability, reliability, use and value of survey data

Fostering innovation

- Identifying, sharing and promoting innovations and best practices in data production, collection and dissemination

Making the data count

- Advocating for better use of statistics for better lives

“Statisticians of the world unite”

Data revolution or evolution?

Call for data revolution parallels Busan Action Plan on Statistics, that seeks to :

- ✓ Fully integrate statistics in decision making
- ✓ Promote open access to and use of data
- ✓ Increase resources for statistical systems

Key questions

Why do we need a “data revolution”?

What should a “data revolution” include?

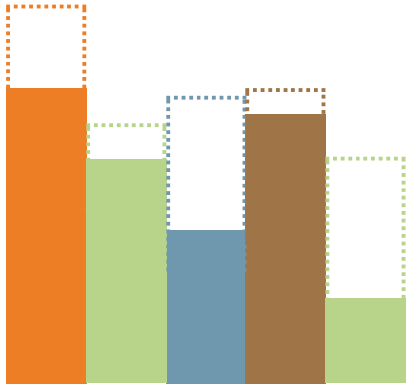
How can it be achieved?

Why do we need a data revolution?

- Old problems
- Emerging priorities
- New opportunities
- Lessons from the MDGs

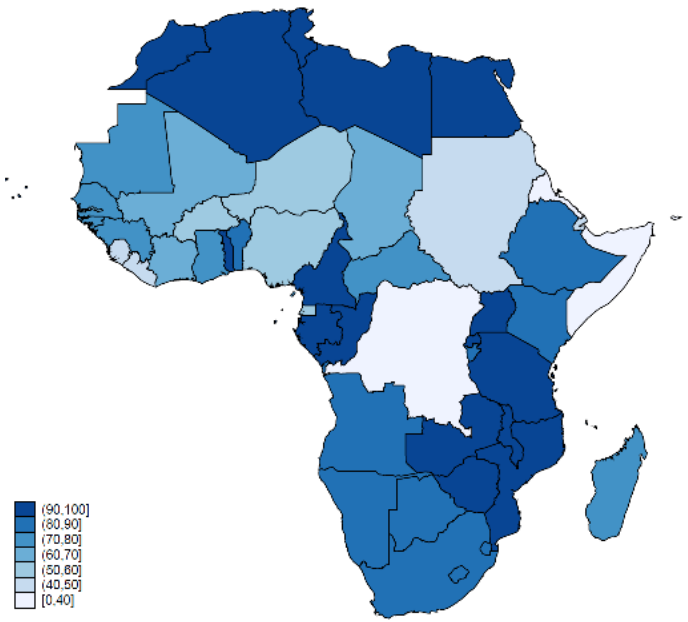
Old problem: what's a data gap?

- Data does not exist
- Data exists, but not in a useable forms for users
- Data exists, in useable forms, but nobody knows
- Data exists, in useable forms, people know, but nobody knows how to use it
- Data exists, is usable, people know, people know how to use it, but nobody cares!



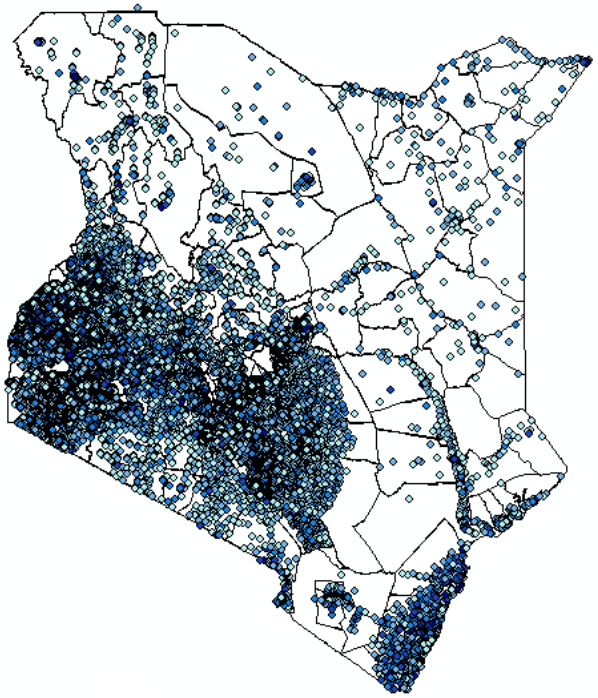
Old problem: seeing like a donor

Seeing Like a Donor



Net primary enrollment rate for 52 African countries from the World Bank's World Development Indicators.

Seeing Like a State

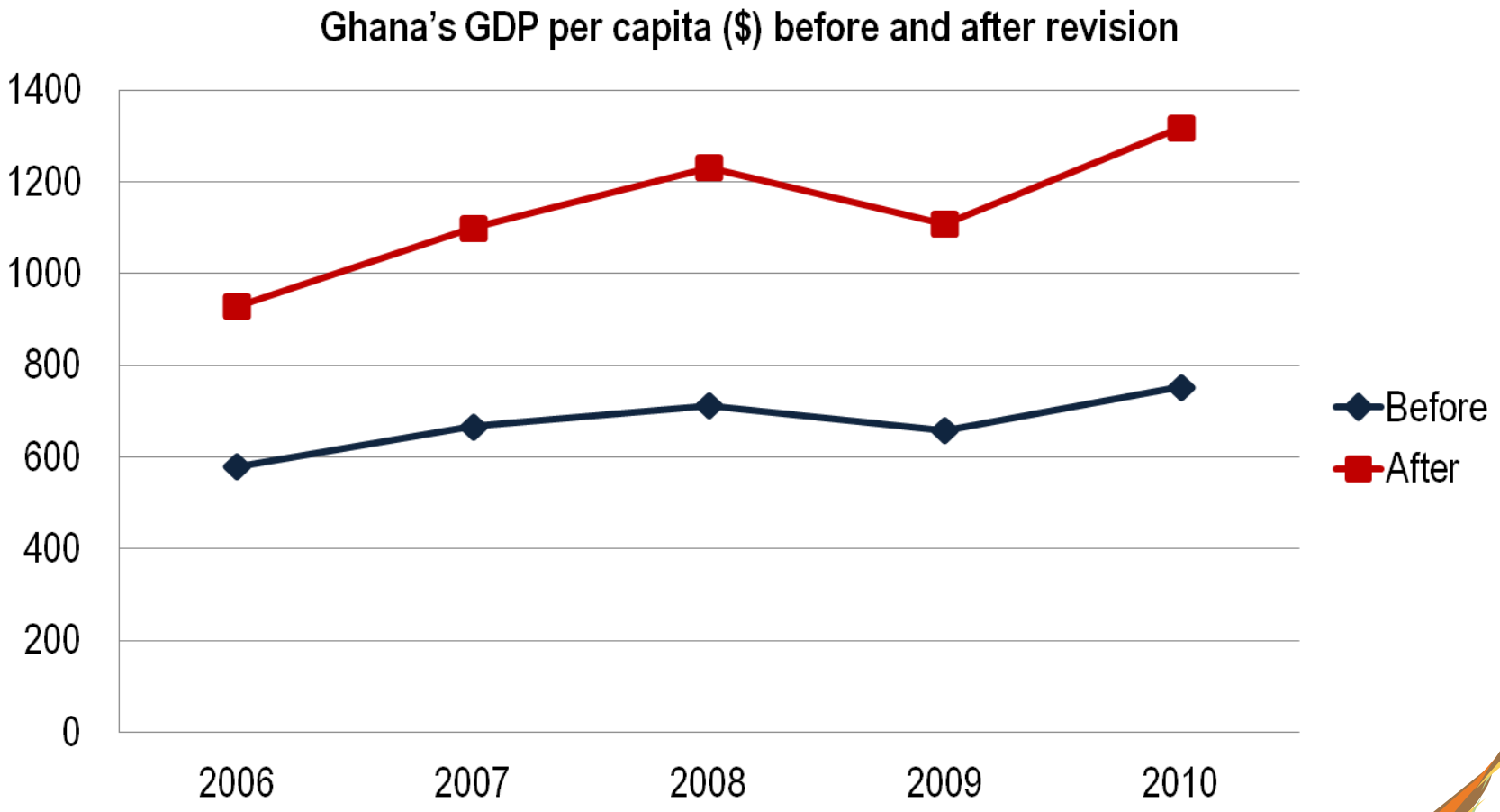


Enrollment levels in ~20,000 Kenyan primary schools from the Kenyan government's Education Management Information System -- now available on its Open Data Portal.

Source: Seeing Like a State, James Scott

Source: Seeing like a state in Africa: Data needed, Justin Sandefur, Center for Global Development

Old problem: communication



Source: Ghana Statistical Service, 2013

Old Problem: Wrong Data

Pacifc Labour Force Household Survey Module

Country	Economic Activity Question	*Who is asked unemplyment questions?*				*Two unemployment questions asked*			
		LABOUR FORCE			Not LABOR FORCE	*Actively looking for work?*		*Why not actively looking for work?*	
		Paid Work	Unpaid Work	No working		YES	*available for work?*	*Reasons*	*available for work?*
1	main activity (incl. NLF)		*not asked*	***"Un-employed"***	*not asked*	**NO**	**NO**	Yes	yes
2	did person do any work (no LF)		yes	**not work for pay**	YES	yes	yes	yes	yes
3	did person do any work (no LF)		*not asked*	not working	YES	yes	yes	yes	**NO**
4	did person do any work (no LF)		Yes	not working	YES	yes	yes	yes	yes
5	did person do any work (no LF)		*not asked*	not working	*n.a.*	*sequencing problems*			
6	main activity (incl. NLF)	Yes	Yes	not working	YES	yes	yes	yes	yes
7	main activity (incl. NLF)		Yes	not working	*not asked*	yes	yes	yes	yes
8	did person do any work (no LF)		*subsistence only*	not working	YES	yes	yes	**NO**	**NO**
9	did person do any work (no LF)		*not asked*	not working	YES	yes	**NO**	**NO**	**NO**
10	main activity (incl. NLF)		*not asked*	***"Looking for work"***	*not asked*	yes	**NO**	**NO**	**NO**
11	did person do any work (no LF)		*not asked*	not working	YES	yes	yes	yes	yes
12	work for pay/do business		Yes	**not work for pay**	*not asked*	yes	yes	**NO**	**NO**
13	type of work (no LF)		*not asked*	not working	YES	yes	yes	yes	yes
14	main activity (incl. NLF)		Yes	not working	YES	yes	yes	yes	yes
15	did person do any work (no LF)		*not asked*	not working	YES	yes	yes	yes	yes (M)

Source: Secretariat of the Pacific Community, Statistics for Development Division

Marshall Islands Labour Force models

Labour Force Category	Baseline *"Main activity last week"*	Model-0 *People NLF asked if actively looking for work, and were available for work (remove /NLF => allocate to unemployed)*	Model-1 *People in Unpaid work asked if actively looking for work, and were available for work (remove from Unpaid work => allocate to unemployed)*	*2-a: Add **paid** secondary activity of **unpaid workers** -> **Paid workers***	*2-b: Add **paid + unpaid** secondary activity of people not in LF -> **paid + unpaid** workers*	Model-2 *add 2a and 2b modifications to Model 1 Totals*
WORKING AGE POPULATION	**6,841**	**6,841**	**6,841**	**6,841**	**6,841**	**6,841**
LABOUR FORCE	**4,126**	**4,915**	**4,915**			**5,129**
PAID WORK	**2,794**	**2,794**	**2,794**	**2,809** *(2,794+15)*	**2,904** *(2,809+ 95)*	**2,904**
[01] Employer	10	10	10	0	2	12
[02] Employee - public sector	1,020	1,020	1,020	1	12	1,033
[03] Employee - private sector	1,472	1,472	1,472	0	15	1,487
[04] Producing goods/services for sale, running a business self-employed)	292	292	292	14	66	372
UNPAID WORK	**1,332**	**1,332**	**521** (-811)	**506** *(521-15)*	**625** *(506 + 119)*	**625**
[05] Producing goods for family/own consumption (self-employed)	109	109	27 (-82)	26 (-1)	43	69
[06] Family worker (family business/plantation)	29	29	12 (-17)	11 (-1)	11	22
[07] Family worker (help with mestic/household duties;	1148	1148	466 (-682)	455 (-11)	36	491
[08] Voluntary work/community work	46	46	16 (-30)	14 (-2)	29	43
UNEMPLOYED		**789** (16.1%)	**1,600** (32.6%)	**1,600**	**1,600**	**1,600** (31.2%)
NOT IN THE LABOUR FORCE	**2,715**	**1,926** (-789)	**1,926**	**1,926**	**1,712** *(-214)*	**1,712**
[9] Student – Fulltime	1,057	754 (-303)	754		716 (-38)	716
[10] Student – Part-time	26	13 (-13)	13		5 (-8)	5
[11] Home duties	1,153	743 (-410)	743		646 (-97)	646
[12] Retired/too old/ disabled	265	260 (-5)	260		189 (-71)	189
[13] NONE – do nothing/be happy	214	156 (-58)	156		156	156

Emerging data priorities

- More use of national data for new goals:
  - Jobs
  - Education (potentially on quality)
- Disaggregation of existing data to measure:
  - Getting to zero poverty
  - Impact on women and girls
  - Impact on inequality
- New indicators for new areas:
  - Sustainable development
  - Governance

Summary: Lessons from the MDGs

- MDG monitoring had an overall positive impact on data production:
  - More surveys, data and statistics available
  - Countries following up after donors left
  - Statistical capacity efforts – NSDS as strategic approach

- … but created some new problems
  - Data orphans
  - National data often not used
  - Crowding out versus crowding in

What should a “data revolution” include?

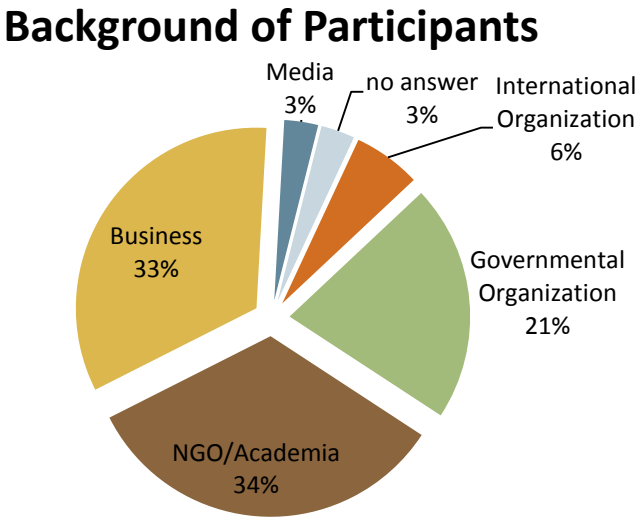
A measurement framework

should not…

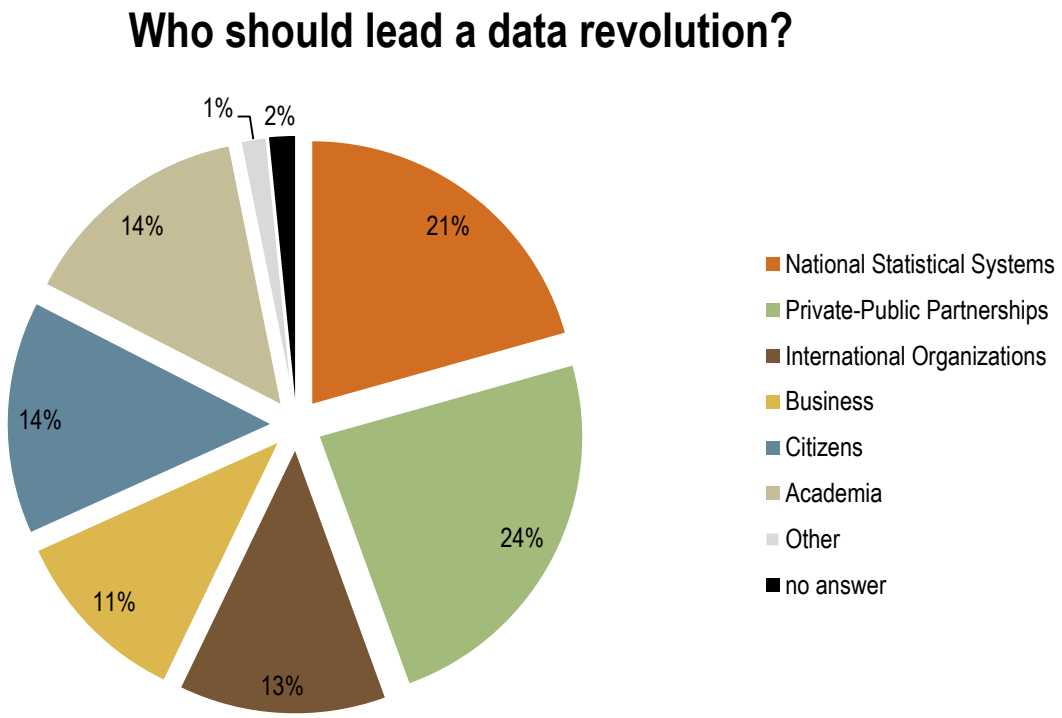
- Simply increase the number of global surveys
- Ignore lessons learnt from the MDG measurement exercise
- Rely solely on technology innovations while ignoring the reality of poor countries

What do stakeholders think?

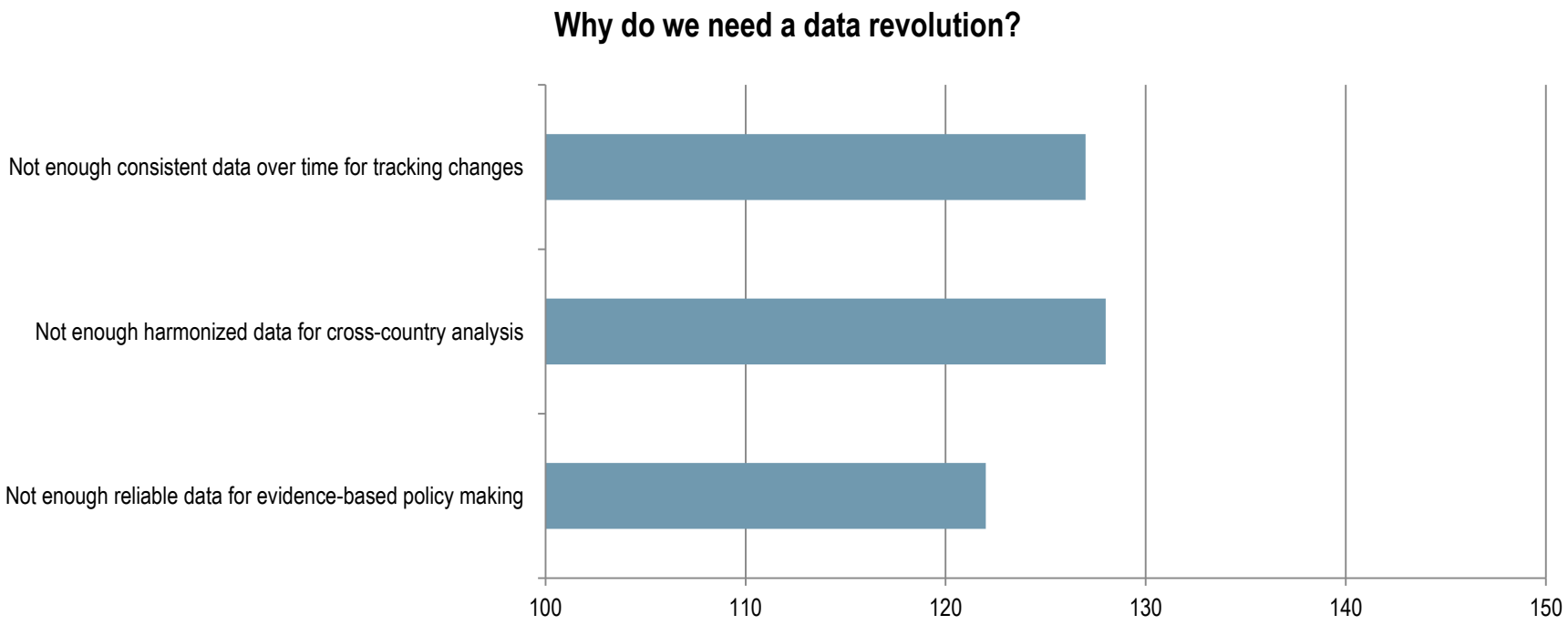
- Survey at European Development Days, 30 participants from international organizations, government, civil society, media, and business



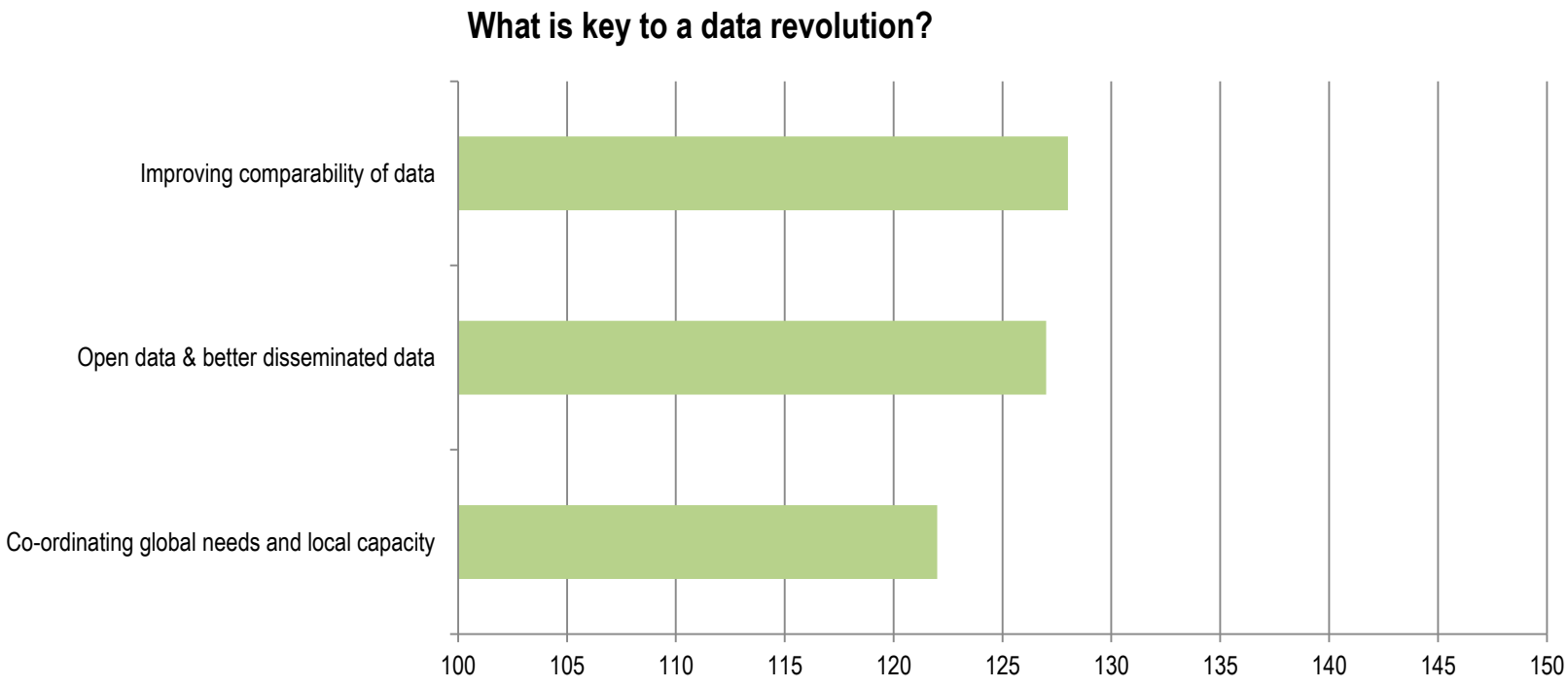
Who should lead? Everyone!



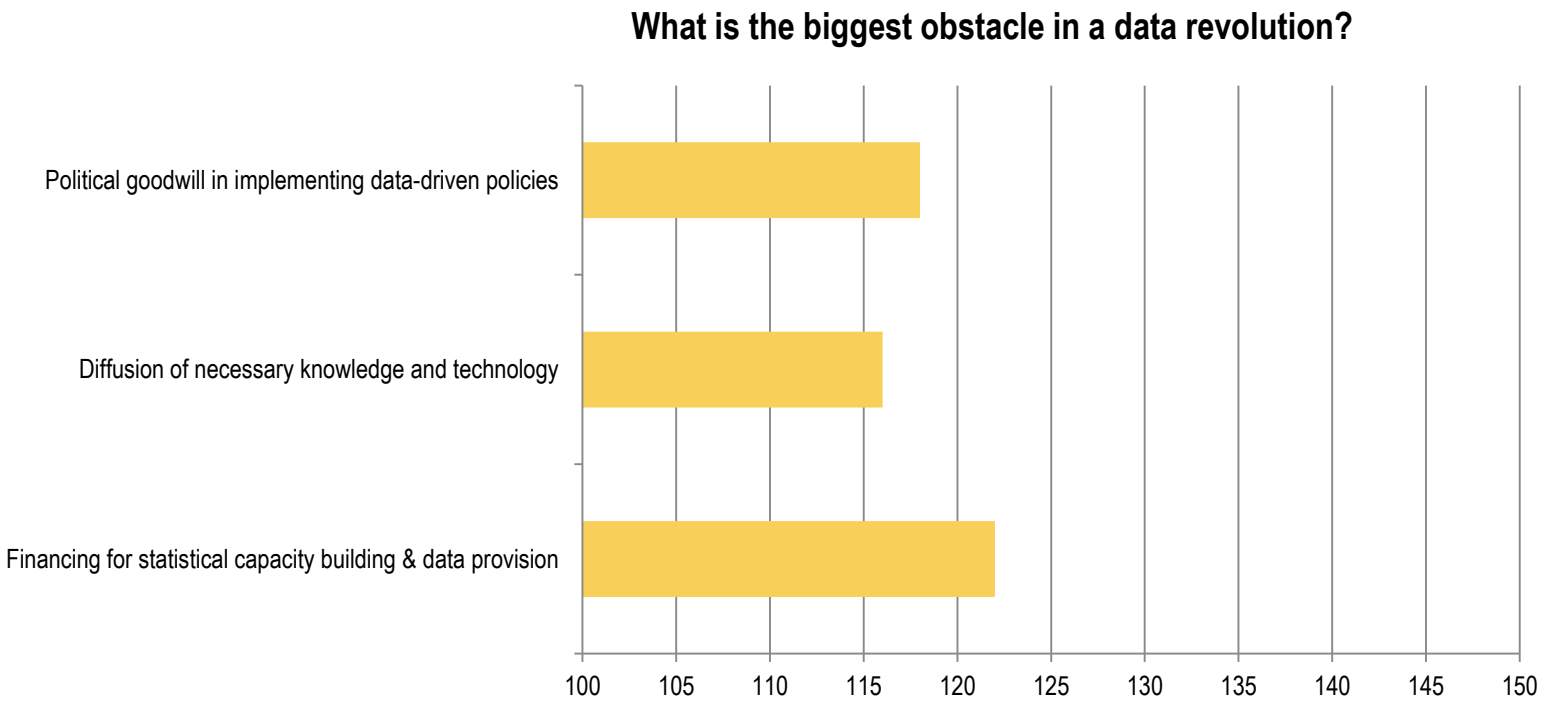
Different priorities…



… with different opinions…



… and different challenges



Unpacking the data revolution

It should …

1. build upon past successes
2. catalyse the expansion of more relevant, timely, open and reliable data production
3. support decision makers, including politicians, business and citizens, to make informed decision for better lives
4. be implemented via new ways to support (fund) statistical capacity and data production

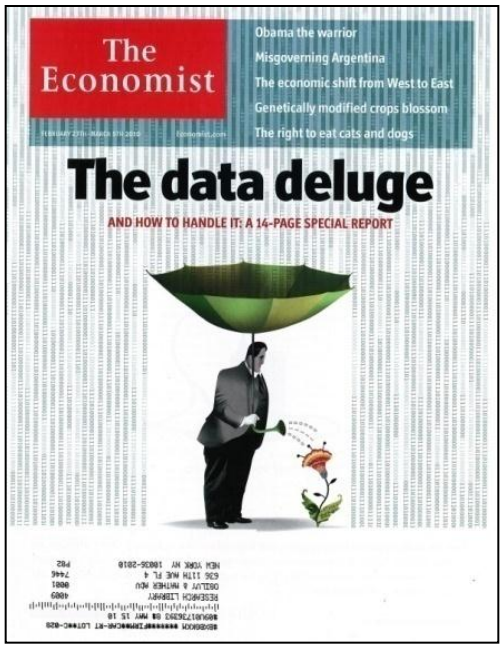
How can we achieve a data revolution?

- Think out of the box – “revolutionary” – but be realistic
- Strengthen what works, harness new opportunities
- Create new partnerships

New opportunities

- "Big data", cloud computing, and "machine-to-machine" combined.
- By 2020, 50 billion mobile wireless devices will be connected to the Internet (Ericsson)
- The Internet as a general purpose technology with applications in energy, health, transportation, education etc.

Internet
(Wave3)



New opportunities: big data

Growing **amount** of data

- The global volume of digital data will multiply by a factor of 40 by the end of this decade

Growing capacity to **analyse** the data

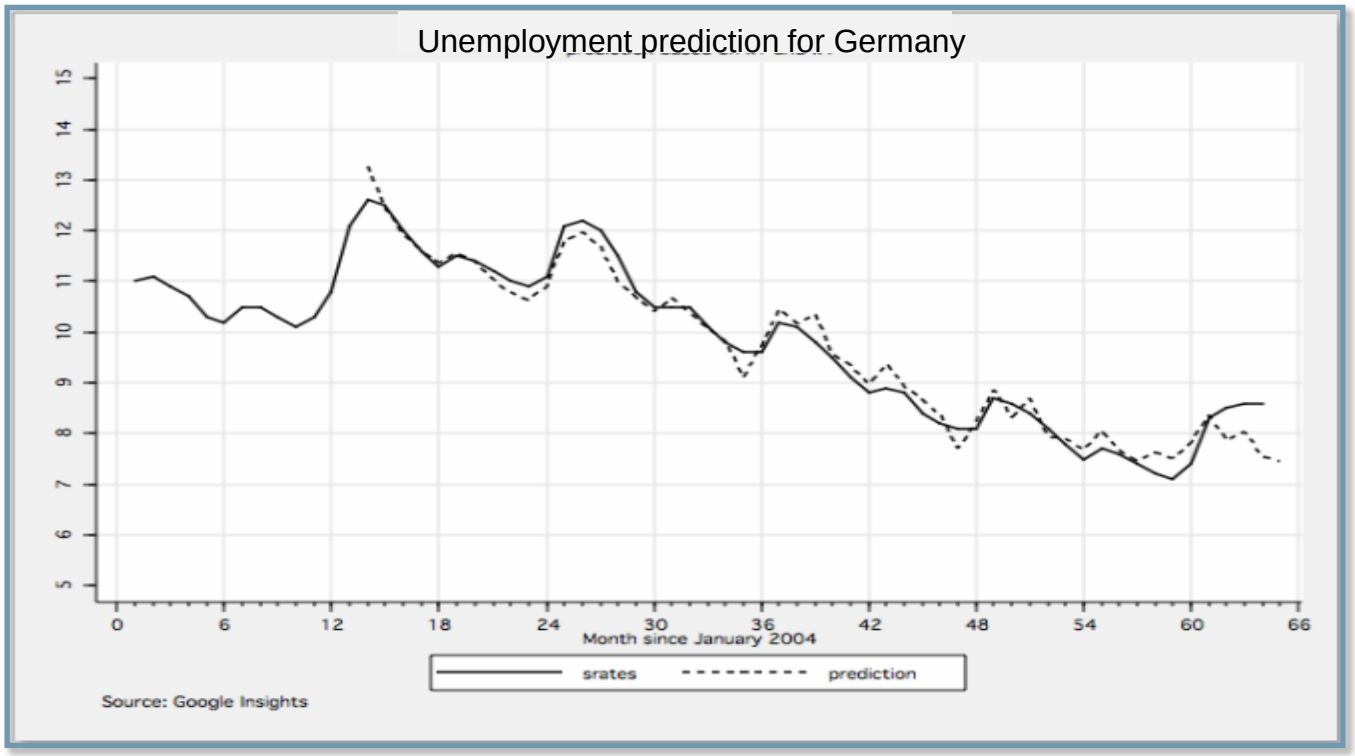
- From 10 years to decode the human genome in 1993-2003 to 1 week in 2010

Growing ability to **store** the data

- A disk drive that can store all of the world's music costs less than $2000

Source: OECD Project on Data and Data Analytics: Prospects for Growth and Well-Being (http://oe.cd/bigdata)

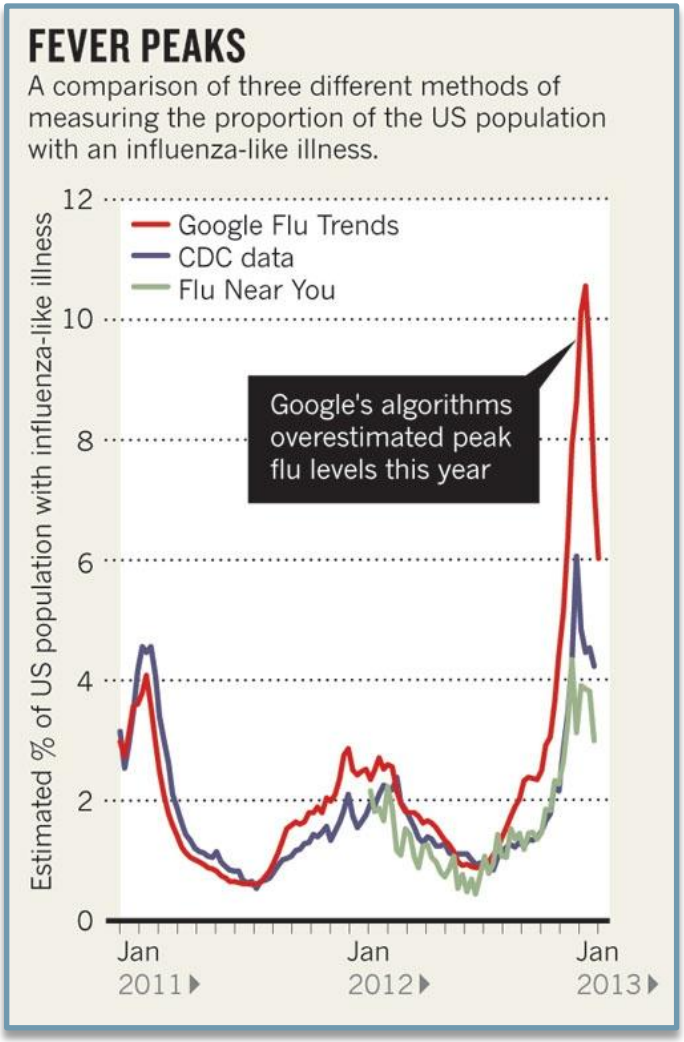
Using big data: successes



Could we really use Google to predict unemployment statistics?

Source: OECD Project on Data and Data Analytics: Prospects for Growth and Well-Being (http://oe.cd/bigdata)

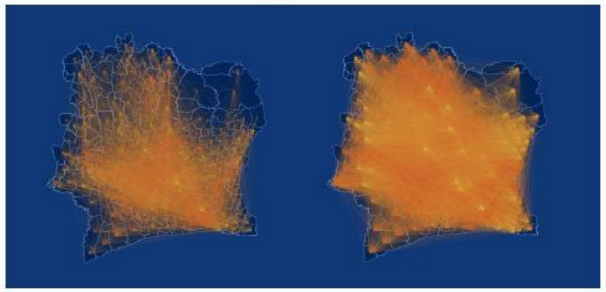
Using big data: and setbacks



- Google's "Flu Trends" drastically overestimated peak flu levels in 2013
- Data quality control is important when dealing with issues that have serious implications, such as national health

Source: OECD Project on Data and Data Analytics: Prospects for Growth and Well-Being (http://oe.cd/bigdata)

New opportunities: phone companies and cellular data



HEALTH / EPIDEMICS

Exploiting cellular data
for disease containment
and information campaigns strategies
in country-wide epidemics

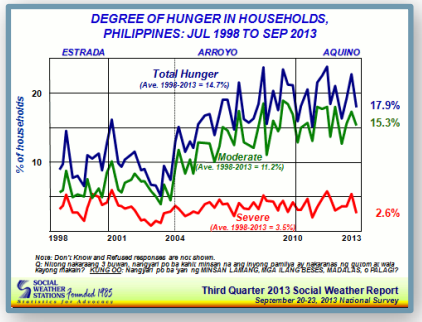
- A. Lima
- M. De Domenico
- V. Pejovic
- M. Musolesi

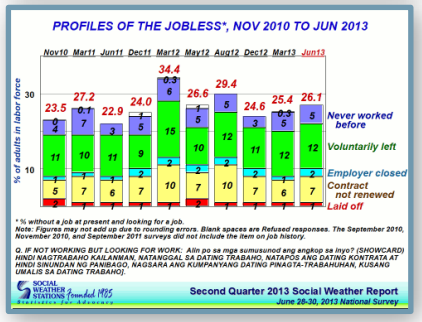
School of Computer Science,
University of Birmingham, UK

Source: Using ICT's to shape the post-2015 framework, European Development Days 2013, Orange

New opportunities: civil society

SOCIAL WEATHER STATIONS





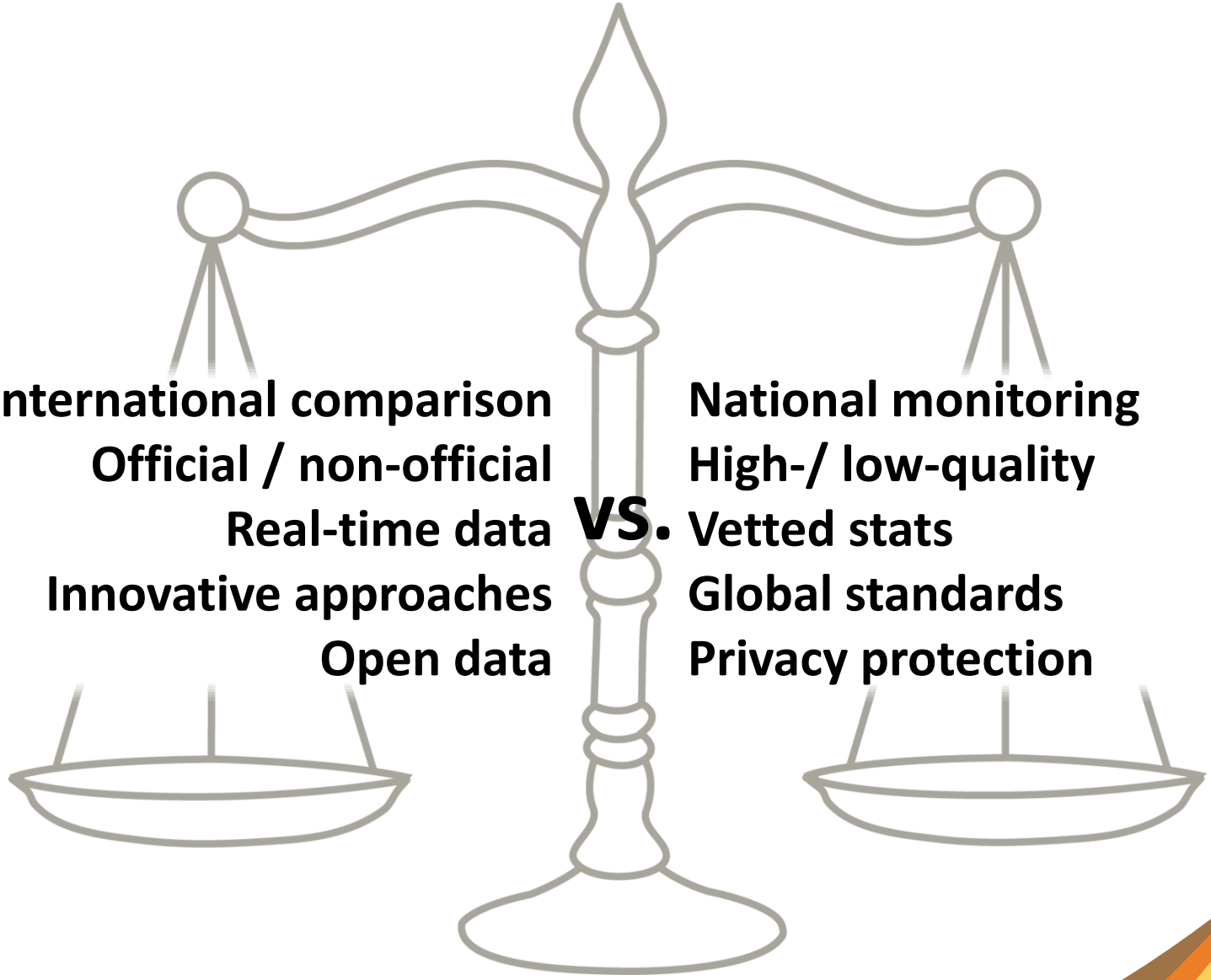
- Independent, non-partisan, non-profit institute from the Philippines
- Conducts social surveys and survey-based social science research
- Provide additional information to those by NSO

Using big data: implications for NSOs

- Should NSOs…
  - take on a new mission as a trusted third party whose role would be to certify the statistical quality of these new sources?
  - use non-traditional sources to augment (and perhaps replace) their official series?
  - issue statistical best practices in the use of non-traditional sources and the mining of big data?
  - be given legal power to collect personal information?

Source: OECD Project on Data and Data Analytics: Prospects for Growth and Well-Being (http://oe.cd/bigdata)

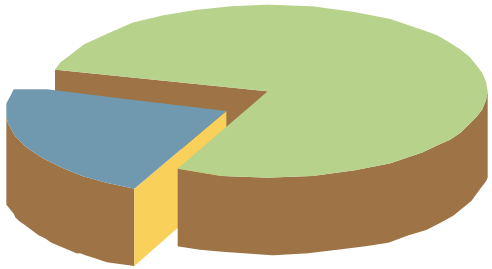
Need to strike a balance



International comparison		National monitoring
Official / non-official		High-/ low-quality
Real-time data	**vs.**	Vetted stats
Innovative approaches		Global standards
Open data		Privacy protection

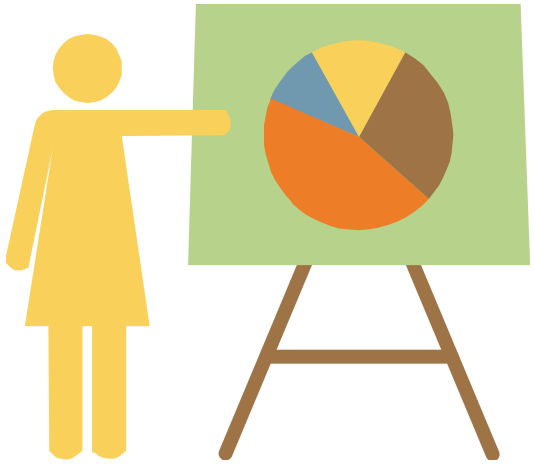
A new global partnership?

Data Collection and Sharing
IHSN
DevInfo
UNECE HLG
In Depth Networks
Magpi
World Bank



A new global partnership?

Data Harmonisation and Co-ordination
UN Regional Commissions
EUROSTAT
Regional Development Banks
IHSN
UNSC
CCSA
WHO/UNICEF JMP

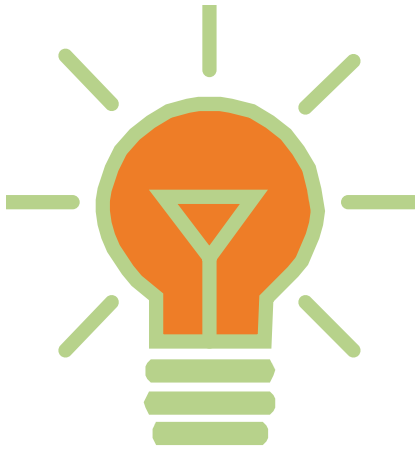


A new global partnership?

Linking Data and Accountability / Transparency
Development Gateway
ONE
Development Initiatives
Save the Children
WB Open Data Team
Open Government Partnership

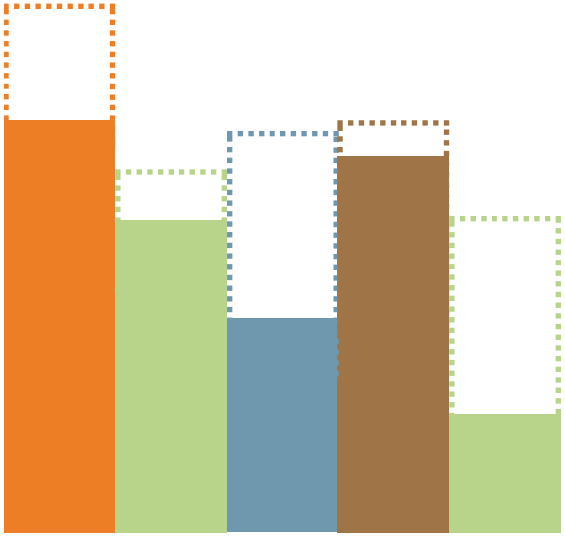
A new global partnership?

Using New Sources of Data
UN Global Pulse
UN Global Geospatial Information Management
Omidyar Network
Google
UNECE Big Data Working Group



A new global partnership?

Bridging Gaps in Data Supply and Demand
DATA2X
PARIS21 Network of NSOs
World Bank's DECDG
DFID Global Statistics Partnership
UNSD SDMX Initiative
Data for African Development WG (CGD/APHRC)
Global Strategy to Improve Agricultural and Rural Statistics



A new global partnership?

Funding Statistical Production
IFAD
AFDB
European Commission
UNECE HLG
DFID
WB SRF
IEO
NSOs

A new global partnership?

Informing a Data Revolution

In Conclusion

- The data revolution will look different in every country and region – your data revolution may not be mine
- However there are common principles
  - Better coordination
  - Increased support
  - Strengthened capacity
  - Increased accountability
  - Better policy
- Ultimate goal remains: improve lives

PARIS21 Secretariat
OECD/DCD
4 Quai du Point du Jour
92100 Boulogne-Billancourt,
France
contact@paris21.org
www.paris21.org


twitter.com/ContactPARIS21


facebook.com/ContactPARIS21

youtube.com/PARIS21OECD