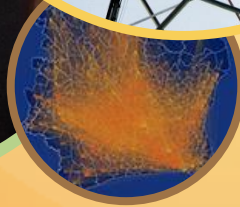


# informing a data revolution

right data, to the right people, at the right time, in the right format

## IDR progress report



Web-COSI Workshop

October 2014

# Contents

goals of the Data Revolution project  
big data Vs official statistics  
The potential of crowdsourcing  
Opportunities & challenges

# why a data revolution?

The Partnership in Statistics for Development in the 21st Century (PARIS21) funded by the Bill and Melinda Gates Foundation launched in march 2014 the project

## Informing a Data Revolution (IDR)

To: *“produce a Road Map that will guide international leaders and policymakers, and national governments and statistical offices, in their path toward an effective, relevant, and sustainable development data revolution. “*

# why a data revolution?

DATA



INFORMATION



KNOWLEDGE



ACTION

We need to **unlock the power of data** ...

Getting the *right data*,  
to the *right people*,  
at the *right time*,  
in the *right format*.

informing a data revolution

# the data revolution needs to . .

- Address existing data gaps to ensure that better statistics lead to better decisions leading to better lives.
- Understand different types of data gaps, and how they relate to national statistical systems.

# The IDR project is..

- **Taking stock in countries** of what is being done – good and bad – and find out about the needs of data producers and data users
- **Exploring innovations**, finding out what works and seeing if and how it can be replicated
- **Advocating** for and promote a data revolution
- **Producing a road map** for the next five years

# IDR activities – the Road Map

- the Road Map report will be launched in mid-2015,
- It will make the case for a data revolution in developing countries.
- It will identify what needs to be done and how, and estimate how much it might cost.
- It will be supported by a number of activities

# IDR activities – Country Studies

- Understanding the challenges facing national statistical systems is essential if they are to join and benefit from the data revolution.
- PARIS21 is studying statistical systems in 27 countries and has launched in-depth studies in 7 countries.

# IDR activities – Country Studies

## ● Cross-country studies

*Azerbaijan, Bolivia,  
Botswana, Cambodia, Costa  
Rica, Dominican Republic,  
Gabon, Ghana, Malawi,  
Mali, Mexico, Mozambique,  
Nepal, Nigeria, Pakistan,  
Peru, Senegal, South Africa,  
Tanzania, Uganda, Viet  
Nam.*

## ● In-depth studies

*Bangladesh, Burundi,  
Cabo Verde,  
Democratic Republic  
of the Congo,  
Colombia, Philippines,  
Trinidad and Tobago.*

# IDR activities – Innovations

- Technological, institutional and other developments provide major opportunities to improve the collection, compilation, dissemination and use of data.
- PARIS21 is creating an inventory of case studies to identify and explore solutions that can fill data gaps, reduce costs and improve efficiency so that more and better data effectively contribute to improving peoples live.

# When?

- The Road Map will be launched in July 2015 in advance of the UN General Assembly
- The Road Map document will be supported by a number of other outputs
  - Clear vision of what the revolution is about
  - Background documents including the country studies and the background research
  - The Road Map will include an overview of how the revolution can be implemented, including an estimate of the financing requirements

**2.5 quintillion bytes** of data  
are created daily\*

**How can we use this**  
**big data**  
**for development?**

# Big Data Vs Official Statistics

## Big Data – the Innovators

- ABS Australia
- CBS Netherlands
- Other NSOs – proceeding with caution, or just doing nothing

# Big Data Vs Official Statistics

- NSOs – potential of Big Data is clearly recognised.
- National offices are (slowly) investigating which sources could be used to complement existing official data and to fill gaps
- There are still major concerns over data quality and stability.
- Currently felt that the focus should be more on improving and using Administrative data rather than other types (phone logs, sensors etc.).

# Big Data Vs Official Statistics

## CBS Netherlands:

- Research work using traffic sensors, social media and mobile phone logs.
- The data is considered imperfect but timely and of interest although new methods and models are needed

## ABS Australia:

- Research in the areas of satellite imagery (estimating crop yields) , mobile positioning for pop mobility.
- It was noted that new skills (“Data scientists”) are needed in statistical offices to harness the potential of these new data sources,
- The amount of time needed to make arrangements to use data from commercial companies and the possible need for legislation in this area is another reason for slow progress, although it was noted that there have been changes in attitude in national offices towards big data.

# Big Data Vs Official Statistics

- New skills (“Data scientists”) are needed in statistical offices to harness the potential of these new data sources,
- The amount of time needed to make arrangements to use data from commercial companies and the possible need for legislation in this area is another reason for slow progress,
- There have been changes in attitude in national offices towards big data.

# Threat or Opportunity?

- Collecting data in real time or near real time **maximize the potential of data**
- big data has potential as an **input for official statistics**; either for use on its own, or in combination with more traditional data sources such as sample surveys and administrative registers
- Big data has the potential to produce **more relevant and timely** statistics than traditional sources of official statistics
- By incorporating relevant Big data sources into their official statistics process **NSOs are best positioned to measure their accuracy**

# Threat or Opportunity?

- **Legislative** - with respect to the access and use of data.
- **Privacy** - managing public trust and acceptance of data re-use and its link to other sources.
- **Financial** - potential costs of sourcing data vs. benefits.
- **Management** - policies and directives about the management and protection of the data.
- **Methodological** - data quality and suitability of statistical methods.
- **Technological** - issues related to information technology.

# Crowdsourcing

The screenshot shows the 'Innovations Inventory' website. The header includes a navigation bar with icons for people, a presentation, a clock, a speech bubble, a classical building, a radio tower, and the PARIS21 logo. Below this is a search bar with 'Browse Innovations' and 'Search Innovations' buttons, and a 'GO' button. The main content area features a table with three columns: 'Innovation', 'Lead Innovator', and 'Project Status'. The table lists several projects, including 'AidData Crowdsourcing', 'AidData Groundtruthing', 'FrontlineSMS Radio', 'Electronic Elections', 'Real time macroeconomic and human development trends', and 'Surui Carbon Project'. Below the table, there are tabs for 'Crowdsourcing', 'Data Management', 'Funding', 'Private Partnerships', 'Skills Development', 'Standards', and 'Use of Alternative Data Sources'. The bottom of the page shows a Windows taskbar with various application icons and the system clock.

Innovation	Lead Innovator	Project Status
AidData Crowdsourcing	AidData	Implemented
AidData Groundtruthing	United Nations University	Implemented
FrontlineSMS Radio	FrontlineSMS	Implemented
Electronic Elections	Inmarsat	Implemented
Real time macroeconomic and human development trends	Premise	Implemented
Surui Carbon Project	Intel	Implemented

informing a data revolution

**Thank you**

**Questions & Answers**

**Trevor.Fletcher@oecd.org**