Overview of the Project

Matjaž Jug matjaz.jug@unece.org
Agenda

• Overview of the project
• Presentation from the Partnership, Privacy and Quality teams
• BREAK
• Sandbox presentation
• General discussion of all Big Data project results
Project Objectives

To identify, examine and provide guidance for statistical organizations to identify the main possibilities offered by Big Data and to act upon the main strategic and methodological issues that Big Data poses for the official statistics industry.

To demonstrate the feasibility of efficient production of both novel products and 'mainstream' official statistics using Big Data sources, and the possibility to replicate these approaches across different national contexts.

To facilitate the sharing across organizations of knowledge, expertise, tools and methods for the production of statistics using Big Data sources.
Virtual sprint (March)

Guidelines

Task teams (May)

Consultation (April)

Physical sprint (April)
Priorities

Privacy Task Team

Quality Task Team

Sandbox Team

Partnerships Task Team
Priorities

Modernisation Committee on Products and Sources

Partnerships Task Team

UNSD Big Data Questionnaire

Modernisation Committee for Organisational Frameworks and Evaluation

Sandbox Team
Collaboration

- Task Teams
- Modernisation Committees
- Sandbox Experiments
- External partners
Approach

- Analysis of best practice and available documentation
- Training and capability building (Rome, May)
- Work in virtual teams on two-week schedule
- Virtual sprints
- Face to face workshops (Heerlen, September)
Definition of success

Success criteria for the individual work packages

1. A consistent international view of Big Data opportunities, challenges and solutions, documented and released through a public web site.

2. Recommendations on appropriate tools, methods and environments for processing and analysing different types of Big Data, and a report on the feasibility of establishing a shared approach for using Big Data sources that are multi-national or for which similar sources are available in different countries.

3. Exchange of knowledge and ideas between interested organizations and a set of standard training materials.

4. The project is completed on schedule, and delivers results that are of value to the international statistical community.

All this is feeding in the future work.
NEW: Project Deliverables for all Task Teams

- Big Data Project Summary Report
- Partnerships Task Team
  - Partnerships abstract for NTTS March 2015.docx
  - Guidelines for the establishment and use of partnerships in Big Data Projects for Official Statistics
- Privacy Task Team
  - NTTS abstract on results Privacy Task Team.docx
  - Deliverable 1: Stock taking exercise
  - Deliverable 2: Big Data characteristics and their implications for data privacy
  - Deliverable 3: Big Data Privacy
- Quality Task Team
  - Big Data Quality Framework Abstract for NTTS.docx
  - NEW: Big Data Quality 2 page summary for report.docx
  - NEW: Big data quality frameworks UNECE task team outcomes v1.2.docx
- Sandbox Task Team
  - Sandbox Team NTTS Abstract
  - Experiment report: CPI Team
  - Experiment report: Canadian Smart Meter Data
  - Experiment report: Job Vacancies
  - Experiment report: Mobile Phones
  - Experiment report: Social Data
  - Experiment report: Social Data: Mobility Studies
  - Experiment report: Traffic Loop data
  - Experiment report: Web Scraping

Big Data is new frontier - still more questions than answers, and recommendations will require more time to get confirmed but we have momentum and good foundations for the future work!
Thank you!

Executive Board, Modernisation Committees

3 Task Teams

7 Sandbox Experiment teams

1 Project Manager
2 Coordinators

75 individuals

partners

Executive Board, Modernisation Committees

3 Task Teams

7 Sandbox Experiment teams

1 Project Manager
2 Coordinators

75 individuals