Big Data Inventory

Background

Following the 2013 Meeting on the Management of Statistical Information Systems, a Task Team was set up to work out the key issues with using Big Data for official statistics, identify priority actions and formulate a project proposal. As part of its work, the Task Team gathered a significant amount of information about the use of Big Data in various national and international statistical organizations. The Task Team recognised that this information could form the basis of a useful resource for the official statistics community, and therefore decided to further develop and publish this resource in the form of an inventory.

Scope of the inventory

It is unrealistic to assume that the inventory could record all possible information relating to the use of Big Data for official statistics; such an inventory would also be very difficult for a single person or organization to maintain. The inventory is therefore **not meant to be exhaustive**, but aims to include at least the key resources that will be most valuable for the official statistics community. Where a resource already exists somewhere on the Internet, the inventory functions as a portal, holding links to resources and documentation, to avoid duplication and reduce problems of version control.

Contents of the inventory

The inventory contains structured and searchable information about actual and planned use of Big Data in statistical organizations. This information is held in a standard template, searchable by statistical domain (using the Classification of Statistical Activities, an existing international standard), and type of Big Data (using a classification developed by the Task Team).

Hosting and maintaining the inventory

The inventory is hosted on the UNECE wiki platform. The wiki approach makes it easier for the inventory to be maintained by the whole statistical community, thus avoiding the whole burden of maintenance falling on one individual or organization. This approach also ensures that the inventory is a relevant and useful resource for all. Read access is now open to the public (for inventory entries for which author organizations have given permission). Edit access is restricted to staff of national or international statistical organizations with an interest in this topic, and must be requested by contacting support.stat@unece.org. The infrastructure is maintained by UNECE, which will also act as a moderator if necessary. The 'owner' of the inventory is the High-Level Group on the Modernisation of Statistical Production and Services, on behalf of the international statistical community.

Adding an entry to the inventory

The value of the inventory will be increased if members of the official statistics community add their materials to it. This can be done by requesting a user account with edit access, and then clicking the button below to add an entry. Once you click the button, you will see a new page which asks you to change the title. Fill in as many of the blank fields as possible and press save to add your record to the list. Choose 'edit contents' to make any subsequent changes. Please contact support.stat@unece.org for assistance with your entry.

Search the inventory

Type your search terms (not case sensitive)

Browse the inventory

(To filter by type of big data source, choose from the drop down list.)

(To sort by any column, click on the column heading)

Inventory entries	Type of Big Data used*	Domain**
Australia (ABS) - Remote Sensing for Agricultural and Land Use Estimation	Satellite images	Agriculture, forestry, fisheries Environment
Australia (ABS) - Social Linked (semantic) Data Processing for Various Statistical Uses	Data from public administration	Education Health Income and consumption Labour Population and migration
Eurostat - Consumer Price Index from internet price data	Commercial transactions	Prices
Eurostat - Mobile positioning data for tourism statistics	Mobile phone: call/text times and positions	Tourism
Italy (Istat) - Experimenting Web Scraping and Text Mining in the "Survey on information and communication technology in enterprises"	E-commerce	Information society
Italy (Istat) - Specific purpose geographic basins and population statistics using mobile phone tracking data	Mobile phone: call/text times and positions	Population and migration
New Zealand - short-term population movements during and after a natural disaster	Mobile phone: call/text times and positions	Population and migration : All statistical results from our initial investigations are published in the report. The report does not cover access and supply of data.

OECD - Quality of Internet connections from Google data	Internet searches	Information society
Slovenia - Population statistics using mobile positioning data	Mobile phone: call/text times and positions	Population and migration

** Categories from the Classification of Types of Big Data*

*** Categories from the Classification of Statistical Activities*

Click for viewtracker