Big Data

The 2013 MSIS meeting decided that Big data is a key issue for official statistics. This section of the MSIS wiki serves as a record of the preliminary work on the topic. Current material and developments are now located in two other wiki spaces:

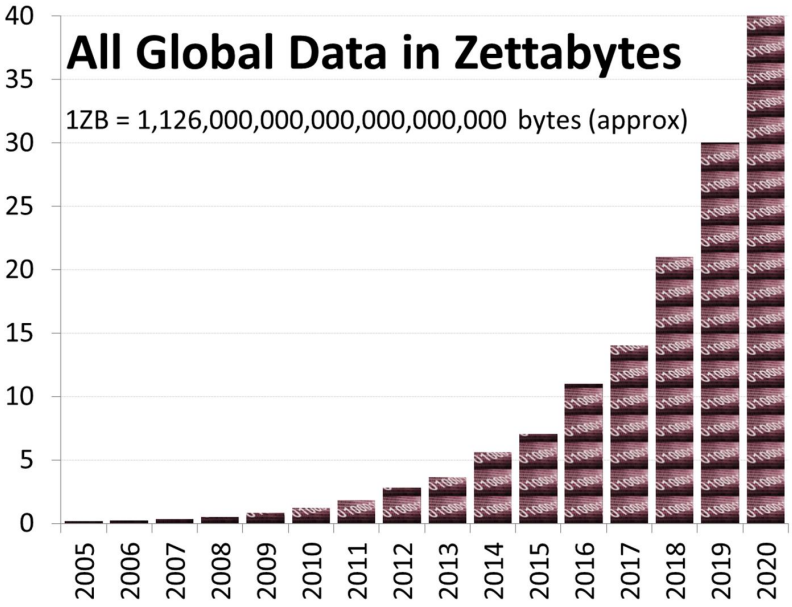
- **Big Data:**

  a public-access wiki for information about ongoing and completed work outputs, including an inventory of big data projects; a classification of big data sources; papers; links and resources.

- **BDP:**

  a wiki for the various task teams currently involved in work for the 2014 HLG project on Big Data in Official Statistics. Most sections are restricted to members of those task teams.

Background to UNECE's Big Data work



In May 2013 a temporary 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. The team worked virtually, by Wiki and Webex.

Throughout 2013, UNECE's statistical division coordinated activities designed to explore the current situation in statistical organizations with respect to Big Data:

- identifying the main categories of challenges and opportunities presented by the emergence of Big Data for the official statistics community
- developing a general classification of Big Data sources and uses with respect to the production of official statistics
- determining current and planned uses of Big Data for official statistics and collating them in an inventory
- identifying relevant research, experiments, projects and papers: both in the inventory and in a list of related resources
- developing a proposal for a project under the auspices of the HLG (High-Level Group for the Modernisation of Statistical Production and Services) which could help to address some of the challenges and answer some of the questions identified
- refining and adding detail to certain aspects of this proposal (the 'sandbox' work package)
- and finally presenting it to the HLG for approval in November 2013.

Once approved, the project was launched at the beginning of 2014. The objectives of the project are:

The page bigdata:Big Data Project could not be found.

UNECE work relating to Big Data

September 2015	Project workshop in Ljubljana to review progress and decide on priority source(s) to use to produce statistics by November 2015.
June 2015	Sprint session in Cork produced an options paper on the future of the Sandbox. This paper and its recommendations were approved by the HLG Executive Board, which oversees the Big Data project.

Throughout 2015	New Big Data project focusing on the Sandbox - New experiments, new data, and a new challenge: "Create a set of internationally comparable statistics based on one or more Big Data sources"
April 2015	Launch of the updated UNECE Big Data Inventory. The inventory replaces the initial inventory and includes all project information collected in the joint UNECE-UNSD survey.
March 2015	Workshop on Big Data, 9 March, Brussels
November 2014	The HLG and representatives of expert groups review the progress so far on the 2014 Big Data project, and decide to launch a new project for 2015, with the main aim of producing and publishing multi-national statistics from a Big Data source
September 2014	Enhancement of the Big Data project inventory in collaboration with the United Nations Statistics Division and the HLG Modernisation Committee on Products and Sources: a new, more detailed questionnaire is under development with the aim of obtaining a more comprehensive picture of current and planned projects and facilitating exchange of knowledge and experience.
August-October 2014	Preliminary results of the survey "Skills necessary for people working with Big Data in Statistical Organisations". More detailed analysis will be prepared by October 2014.
Throughout 2014	International collaboration project on The Role of Big Data in the Modernisation of Statistical Production - This project, overseen by the High-Level Group for the Modernisation of Official Statistics, is running during 2014, and aims to: • identify, examine and provide guidance for statistical organizations to act upon the main strategic and methodological issues that Big Data poses for the official statistics industry • 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 • facilitate the sharing across organizations of knowledge, expertise, tools and methods for the production of statistics using Big Data sources.
June 2013	Draft classification of types of Big Data
March 2013	Position paper *What does Big data mean for official statistics?* (March 2013) drafted for the High Level Group for the Modernization of Statistical Production and Services (HLG).

Click for viewtracker