



Challenges of using Big Data for Official Statistics

Ronald Jansen*, Ivo Havinga and Shaswat Sapkota
United Nations Statistics Division, New York, United States
jansen1@un.org, havinga@un.org and sapkota@un.org

Abstract

Big Data—data sources with a high volume, velocity, and variety—have become increasingly available driven by the general use and vast penetration of electronic devices and the rapid advancements in information technology, and, correspondingly, new tools and methods are being developed to uncover meaningful information from them. Big Data originates from a wide range of sources such as mobile phone, social media, electronic commercial transaction, sensor networks, smart meters, GPS tracking devices, or satellite images, which collectively touch upon almost all aspects of our daily lives; and Big Data is regarded as potentially providing great support to official statistics, and to the delivery of meaningful indicators for monitoring the post-2015 development agenda. However, there are several challenges that first need to be tackled before benefits of Big Data can be realized. This paper will systematically explore those challenges. Drawing from the United Nations Survey on Big Data Strategy and Project Inventory, and other relevant literature, the paper will highlight key obstacles that national statistical systems face to use Big Data for official statistics. Some of the challenges are: representativeness, biases, modelling and other methodological issues, access to data, privacy, transparency and management of public trust, and IT and financial challenges to implementing a Big Data project. The paper will also explore some of the ways—for example, public-private partnerships and legal frameworks—in which these challenges can be addressed. The UN Working Group on Big Data for Official Statistics has the mandate to provide a global work programme on Big Data, to promote practical use of Big Data sources, foster communication and build public trust in use of private Big Data for official statistics. As such, the paper will highlight its work in the current status of the mentioned challenges.

Keywords: mobile phone data, social media data, satellite images, data access, privacy, public trust, national statistical systems, post-2015 development agenda, UN Global Working Group on Big Data