

Author: Dr. Piet J.H. Daas, Senior methodologist and lead Data Scientist

Dr. Piet Daas is senior methodologist at the Department of Corporate services, IT and Methodology. He is the project leader of the Big Data methodological research program, and lead data scientist at the Center for Big Data Statistics of Statistics Netherlands.

His work focuses on the use of secondary (non-survey) data for official statistical purposes and in particular on the use of internet and other big data sources for official statistics. He is a teacher of the CBS Academy, teaches the big data part of the European Master of Official Statistics track at the University of Utrecht and is involved in the big data courses of the European Statistical Training Program. Piet has written many Big Data papers, is one of the founding fathers of the first Big Data based official statistic, has developed a number of Big Data based beta-products, and is active in various European and UN/UNECE big data initiatives. He has a M.S. and a Ph.D. in the natural sciences, both with honors, from the University of Nijmegen in the Netherlands.

Title: Big Data Methodology

Using Big Data for official statistics comes with many challenges. A considerable number of these challenges are of a methodological nature. This presentation will focus on these topics and discusses them with real-world examples. Topics touched upon are: ways to include Big Data in official statistics, working in a data-driven way, the (un)importance of a paradigm shift, the need to deal with new types of data, the importance of quality checks, the need for a continuous critical research attitude, and the link with Data Science. This provides an overview of the current state of the art in the area of Big Data methodology, resulting in a 10-step approach for the production of beta-products based on Big Data.