UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

ModernStats World Workshop

21-22 October 2024, Geneva

Streamlining statistical and data production

Speaker: Mikko Lindholm & Mari Ylä-Jarkko, Statistics Finland

Author(s):

Abstract

Statistical production has traditionally been built as a pipeline, so that each statistic has its own data collection, serving the purposes of that particular statistic. The same way the IT-systems from validation tools to outputs have been unique. The only common surface has been publication – with some variations for each statistic.

Statistics Finland has been transforming its production to more streamlined production system with common IT tools and common data repositories. The work has started some 15 years ago from business statistics, but the new strategic target is to have basic data repositories common for all products and thematic data repositories to serve both statistical production as well as other data services, such as data for researchers.

Because most of our data comes from government registers, the data coverage is close to 100%. Also, all data units have the same identifiers, such as person id, enterprise id, building id, etc. Thus, combining data from different registers becomes possible, and the data repositories can be used to produce many other data products besides official statistics.

In streamlining the data governance, also the data process is streamlined from data gathering to validation and metadata. At the Statistical Office level, there is only one data process. On a more detailed level, there will be some variations in the process, based on the data and collection method, but these variations will be kept to a minimum.

This architecture aims at raising the productivity of statisticians, increasing the flexibility of resource use, enabling quick production of temporary and experimental statistical products and improving data management and regulatory compliance.