

UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE
CONFERENCE OF EUROPEAN STATISTICIANS

ModernStats World Workshop

21-22 October 2024, Geneva

The statistical production LEGO set: using standard models and tools to build metadata-driven pipelines at StatCan

Speaker: Flavio Rizzolo, Statistics Canada

Author(s):

Abstract

The development of the ModernStats models, from GSBPM to GSIM, and their related reference architectures, CSPA and CSDA, provide a solid foundation for better understanding statistical production, informing business decisions and designing reusable software components. Bridging the chasm between these conceptual viewpoints and their implementation standards, most notably DDI and SDMX, and figuring out how to use these implementation standards in a complementary way, has recently received considerable attention in the international community.

This presentation describes Statistics Canada's recent attempt at bridging that chasm by deploying a number of standard-based data/metadata management tools (e.g., Colectica, Aria, Fusion Metadata Registry, Data Lifecycle Manager) using ModernStat models to guide the decision-making process. Making these tools developed within very different communities interoperate at a semantic level requires mappings between the standards together with some "connective tissue" in the form of customized micro-utilities that enable the creation of data/metadata pipelines across phases of the GSBPM. We describe some key use cases, our progress to date, challenges found, and the road ahead.