

New Developments with the Data Documentation Initiative (DDI)

Dan Gillman (US Bureau of Labor Statistics, USA)

Gillman.Daniel@BLS.Gov

Abstract

The Data Documentation Initiative (DDI) continues to evolve. What began as a single standard, the Codebook, has transformed into a family of standards and other work products. The current DDI family includes Codebook, Lifecycle, XKOS, and Controlled Vocabularies. This talk gives an overview of these existing products, a more detailed review of two new draft specifications, and how the DDI family of products fit with the UNECE HLG-MOS framework of standards.

There are two new draft entries to the DDI family, and both have been publicly released for comment. These are DDI Cross Domain Integration (DDI-CDI) and the Structured Data Transformation Language (SDTL). DDI-CDI has several innovations that separate it from earlier products in the DDI family. They include improvements to the variable cascade, an expanded multi-dimensional data model, the datum-centered approach for tracking data through processing and across data sets, detailed process and provenance model, and several logical data structure types for describing a wide variety of data sets. Combined with DDI-Lifecycle, DDI-CDI can be used as the implementation model for GSIM, and therefore an underlying model in support of CSPA. SDTL is an intermediate language for representing transformation commands. It can be used to represent provenance of data sets, and it is compatible with Codebook, Lifecycle, and Cross Domain Integration.

Keywords

DDI-CDI, DDI-Codebook, DDI-Lifecycle, XKOS, SDTL

Model (please specify name of at least one of models (i.e. GSBPM, GSIM, GAMS0 and CSPA) that your abstract is related to)

GSIM and CSPA