Workshop Report – DDI-CDI: Realising Interoperable Data Services in the Metadata Ecosystem

Schloss Dagstuhl – Leibniz Center for Informatics,
24 September – 29 September, 2023, Wadern, Germany

A workshop on the emerging Data Documentation Initiative – Cross Domain Integration (DDI-CDI) metadata standard for supporting multi-disciplinary data sharing was held at the internationally renowned computer-science institute in Wadern, Germany. The event was sponsored by CODATA (the Committee on Data of the International Science Council), and the Data Documentation Initiative Alliance (DDI), and subsidized by Schloss Dagstuhl; it was organized by Arofan Gregory (Chair, DDI-CDI Working Group, DDI Alliance), Simon Hodson (CODATA), Hilde Orten (Sikt and DDI Alliance), Joachim Wackerow (DDI-CDI Working Group, DDI Alliance), Steve McEachern (Australian National University and DDI Alliance). The workshop brought together 25 participants from 21 organizations in 11 countries representing many different domains, including members of relevant standards bodies and data-sharing initiatives, computer scientists, and experts representing the use cases which served to focus the work.

The focus of the workshop was on implementing DDI-CDI, and supporting those who wish to learn more about the standard and implement it. The work was very productive, and was organized into several different areas:

- General overview of the specification and the problem space it is meant to address, including how it relates to other standards implemented in a modular way.
- Specific work around the creation of examples and implementation guides, including a methodology for creating community specific implementation guides, and worked examples for mining metadata embedded in a variety of data formats.

- A proposal on how non-numeric, non-coded data (e.g. qualitative data) could be described in an integrated fashion with quantitative data sets within the DDI-CDI model.
- Alignment of DDI-CDI with external models which describe the dependencies between variables in cross-domain integration scenarios, based on the RDA's I-ADOPT framework and ontology, and the OGC's Observations & Measurements standard.
- Syntax representation of the DDI-CDI model, including Python, Typescript, ShEx, SHACL, and JSON Schema.

In each of these areas, the outputs will be incorporated into the work of the DDI-CDI WG moving forward, and will result in the publication of guidance and documentation for users, tools for developers and implementers, and new features for the DDI-CDI model. Also considered were issues around better supporting implementers through interactions with the DDI Developer's Group and with a more responsive "beta" release process, and how this could fit into the work of the DDI Alliance overall.

Details regarding the event can be found on the Schloss Dagstuhl site, and on the DDI Alliance Workshop page in Confluence.