Workshop Report: “DDI Cross Domain Integration for FAIR Data Sharing across Discipline Boundaries”

Schloss Dagstuhl – Leibniz Center for Informatics,
20-24 September, 2021, Wadern, Germany



A workshop on the Data Documentation Initiative Cross Domain Integration (DDI-CDI) specification was held at the internationally renowned computer-science institute in Wadern, Germany. It was a “hybrid” event, including both on-site and remote participants. 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 (Norwegian Centre for Research Data and DDI Alliance), Joachim Wackerow (GESIS - Leibniz Institute for the Social Sciences and DDI Alliance). The workshop brought together 21 participants from 10 countries and many different domains, including members of the DDI-CDI Working Group and representatives and experts associated with many related metadata specifications and interested institutes.

The workshop examined several topics concerning future developments of the DDI-CDI specification for ubiquitous FAIR data sharing across domain and institutional boundaries. The main focus was on medium-term goals as outlined in the CDI Working Group roadmap. These included the development of a methodology for the creation of community implementation guides; an examination of how data structures can be analyzed, including new examples such as NetCDF which were not originally within scope for the specification; a discussion of the relationship between DDI-CDI and I-ADOPT for describing observations of interest; and a set of modeling topics which covered approaches to using the model in a more modular fashion, syntax representations (including RDF), and the documentation of the features of UML employed by the DDI-CDI specification as an interoperable sub-set for data modeling.

DDI-CDI is designed to work with other metadata standards to provide a “lingua franca” for sharing data across domains and institutions which use different metadata specifications and models. It has a strong emphasis on the structural description of data and associated processes, expressed in a machine-actionable way to support scalable approaches to data sharing. The core UML model can be implemented with a wide range of technologies, with a focus during this workshop on RDF and related specifications. Related standards considered during the course of the workshop included NetCDF, JSON-LD, DCAT, Schema.org, UML, XMI, and many others.

The work was organized around small working groups focused on particular topics, with each group containing members of the DDI-CDI WG and outside experts. Each day, plenaries were held to coordinate the efforts of these groups. Several use cases were presented throughout the week, to provide real-world examples of how the DDI-CDI model could be implemented.

While the hybrid format of this workshop was unfamiliar and sometimes challenging, the output of the workshop was greater than anticipated, including both inputs to the further development of the DDI-CDI model, and documents which will be useful in supporting adoption and use of the specification. Plans for publication of the latter were discussed. The audio-visual facilities and support provided by Schloss Dagstuhl were critical in making the workshop a success.

Details regarding the event can be found on the [Schloss Dagstuhl site](https://www.dagstuhl.de/en/program/calendar/evhp/?semnr=21383), and on the [DDI Alliance Workshop page](https://ddi-alliance.atlassian.net/wiki/spaces/DDI4/pages/2681208854/2021%2BFurther%2BDevelopment%2Bof%2Bthe%2BDDI%2BCross%2BDomain%2BIntegration%2BModel%2Bfor%2BFAIR%2BData%2BSharing%2Bacross%2BDiscipline%2Band%2BDomain%2BBoundaries) in Confluence.