

Having trouble viewing this message? Links disabled? Click or copy:
<http://www.ddialliance.org/system/files/ddi-directions-2014-02.pdf>



From the Director

Welcome to a new edition of *DDI Directions*. In this February issue, we focus on recent DDI-related meetings and events as well as new events on the horizon -- notably, the second North American DDI Users Conference, to be held March 31-April 2 in Vancouver. We also present new DDI tools and some useful updates. As always, a lot of activity is taking place - thank you to the DDI community for your involvement!

Mary Vardigan, Director,
DDI Alliance,
vardigan@umich.edu

In This Issue

- NADDI 2014 to Take Place in Vancouver
- New Tools for DDI Available
- Successful EDDI Held in Paris
- DDI Alliance Members to Meet in Toronto
- Spotlight on the DDI Controlled Vocabularies Working Group
- DDI is Sprinting!
- Project Manager Identified for the Modeling Project
- DDI Codebook 2.5.1 Published, DDI Lifecycle 3.2 Coming Soon
- Identify your Data in the Global Community
- New Data Expression and Transformation Language

Volume VII, Number 2, February 2014

NADDI 2014 to Take Place in Vancouver

Please join us in Vancouver for the [second North American DDI Users Conference](#), to be held March 31-April 2, 2014. Organized by the [University of British Columbia](#), [Simon Fraser University](#), [University of Alberta Libraries](#), and the [Institute for Policy and Social Research](#) at the University of Kansas, the meeting will take place at the [Harbour Centre](#), Simon Fraser University, with the theme "Documenting Reproducible Research."

The Keynote Speaker for the conference is Ann Green. Ann is an independent research consultant focusing on the digital lifecycle of scholarly resources, including their creation, delivery, management, long-term stewardship, and preservation. She has significant history and experience with the DDI, serving on the original DDI committee in 1995 and coauthoring the first DDI specification.

NADDI 2014 offers [two half-day workshops](#) on March 31 that will be of interest to health researchers and others wanting to produce better metadata to support their research.



New Tools for DDI Available

At the EDDI meeting in December in Paris, [Metadata Technology North America \(MTNA\)](#)

Being Developed

- Minutes of Executive Board Online
- HLG Standards Meeting Takes Place
- New DDI Representative

launched some new [DDI-related tools](#), described below:

SledgeHammer UI

As part of MTNA's [OpenDataForge](#) toolkit of data management tools, [SledgeHammer](#) is a flexible desktop-based application enabling the production and use of open data.

Sledgehammer facilitates the transformation of data across popular proprietary, text, and database formats and supports the extraction and generation of standards-based metadata, including all versions of DDI. The interface was designed with the diversity of potential data users – from data producers, to archivists, librarians, analysts, and the general public – in mind. SledgeHammer is available as a freeware Community edition and a low-cost licensed product.



More specifically, Sledgehammer:

- Reads data and extracts metadata from: Stata, SPSS, SAS, DDI+ASCII, Stat/Transfer+ASCII and SQL databases
- Transforms data into open ASCII text format (fixed, csv, delimited), with various optimizations
- Produces standard metadata from input files -- supported specifications include DDI-Codebook (1.0-2.1 and 2.5), DDI LifeCycle (3.1-3.2), and Triple-SSS (1.1, 2.0)
- Computes summary statistics at the variable and category levels for inclusion in DDI or for other purposes -- includes minimum, maximum, average, standard deviation, count of missing values, weighted and unweighted frequencies
- Generates scripts for reading ASCII data into R, SAS, SPSS, Stata, SDA, and various flavors of SQL (MS-SQL, MySql, Oracle, Vertica, HSQLDB). The database scripts support the creation of database schemas along with bulk loading of the ASCII data.

Survey Manager

The [Open Metadata Survey Manager](#) is an open source DDI Lifecycle-based tool developed by MTNA based on similar tools being implemented for the [Canadian Research Data Centre Network \(CRDCN\)](#), [NORC at the University of Chicago](#), the [US National Science Foundation](#), and other projects. The OM Survey Manager aims to provide researchers and data administrators with a tool to familiarize themselves with DDI and survey metadata management. It is designed to work with metadata produced by [DataForge SledgeHammer](#).

The OM Survey Manager is a standalone application used to view, edit, and enhance survey metadata saved in local or shared repositories using DDI and other related standards.

Successful EDDI Held in Paris

The [5th Annual European DDI User Conference \(EDDI13\)](#) took place on December 3 and 4, 2013, in Paris in the impressive Cité Internationale Universitaire de Paris. The host of the conference was [Réseau Quetelet](#), the French Data Archive for the Social Sciences. EDDI13 was jointly organized by [GESIS](#) – Leibniz Center for the Social Sciences and the [Institute for the Study of Labor \(IZA\)](#) together with the host and was supported by the DDI Alliance.

The Keynote Address was delivered by Philippe Cuneo, Director of Methodology, Statistical Coordination and International Relations, at the French National Institute of Statistics and Economic Studies ([INSEE](#)), on the topic “Why National Statistical Institutes are Increasingly

Interested in Standards.”

The [Data without Boundaries](#) project introduced a new [CESSDA](#) Portal for European Research Data Discovery, and the EU infrastructure project [DASISH](#) presented a whole session on how structured metadata can support the interplay of tools for questionnaire design, questionnaire translation, and a database of questions and variables.

The program encompassed three plenary talks, 37 presentations (including three full papers) in two concurrent tracks, and nine poster and software demonstrations. Over 90 people from 46 organizations in 14 countries and two international organizations participated in EDDI13. The presentations are available in the [detailed program](#).

Additional events, each with 15 to 20 participants, took place around EDDI: there were two half-day tutorials on DDI subjects, a one-day meeting of DDI software developers, and a two-day working meeting (a “sprint”) on the further development of the model-based DDI.

EDDI14 will take place December 2-3, 2014, in London. The host is the [Institute of Education \(University of London\)](#), which includes the [Centre for Longitudinal Studies](#).



Roxane Silberman, Director of the Réseau Quetelet, introduces EDDI13 Keynote Speaker Philippe Cuneo.

DDI Alliance Members to Meet in Toronto

On Monday, June 2, 2014, the DDI Alliance will hold its Annual Meeting of Members as well as a meeting of the DDI Scientific Board in advance of the [IASSIST conference in Toronto](#). The meeting will take place at the [University of Toronto Robarts Library](#). More information about the agenda will be coming soon.

Spotlight on the DDI Controlled Vocabularies Working Group

A first set of [controlled vocabularies](#) to describe specific aspects of research data across the data life cycle has been developed for the DDI standard by the [DDI Controlled Vocabularies Group \(CVG\)](#). These



vocabularies may be used for other purposes and by other applications as well. The group will continue to add new vocabularies to the site as they are finalized. Currently, the CVG is working on vocabularies for Source Type and Data Type.

Select DDI Alliance vocabularies are already in use at organizations like the Finnish Social Science Data Archive (FSD), GESIS - Leibniz Institute for the Social Sciences, the Inter-university Consortium for Political and Social Science (ICPSR), Mathematica Policy Research, the UK Data Archive (UKDA), and the University at Bielefeld, Germany. Nesstar Publisher now incorporates the controlled vocabularies for Analysis Unit and Time Method.

Members of the CVG are:

- Lucy Bell, United Kingdom Data Archive (UKDA)
- Stefan Ekman, Swedish National Data Service (SND)
- Sanda Ionescu, Inter-university Consortium for Political and Social Research (ICPSR), Chair
- Taina Jääskeläinen, Finnish Social Science Data Archive (FSD)
- Meinhard Moschner, GESIS -- Leibniz Institute for the Social Sciences
- Michaela Olde, Australian Bureau of Statistics (ABS)
- Hilde Orten, Norwegian Social Science Data Service (NSD)

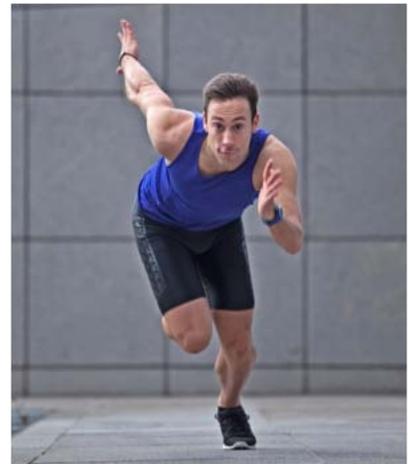
DDI is Sprinting!

A successful “sprint” was held the last week of October at Schloss Dagstuhl in Wadern, Germany, followed by a two-day sprint after EDDI in Paris. Participants made significant headway in launching the modeling process through [these sprints](#).

Two other sprints are planned in upcoming months. Groups will convene the week before NADDI in Vancouver (March 24-28) and the week before IASSIST in Toronto (May 26-30).

Virtual teams have been established to build on progress from the sprints.

The DDI Alliance has created an email list for individuals interested in contributing to and following the progress of the project to create a model-based DDI specification. [Subscribe to the email list](#).



Project Manager Identified for the Modeling Project

Thérèse Lalor of the [Australian Bureau of Statistics](#) will be coming on board as project manager for the DDI Modeling Project, serving from April through June 2014. Thérèse managed the successful [GSIM](#) project, bringing the project to a close efficiently and on time. We look forward to working closely with Thérèse in coming months.

DDI Codebook 2.5.1 Published, DDI Lifecycle 3.2 Coming Soon

A new sub-minor version of DDI Codebook – [DDI 2.5.1](#) – was published on January 29,

2014. This new iteration does not change the namespace of DDI-C. Modifications include the addition of DataFingerprint, relaxed cardinality to support multiple languages, and expanded documentation.

DDI Lifecycle 3.2 will be published later in March.

Identify your Data in the Global Community

The [DDI agency identifier registry](#) ensures that your organization uses a globally unique agency identifier and allows configuration of resolution services.

To obtain an agency identifier, first create an account for your organization and then request a new agency in the Manage section of the website. An administrative and technical contact is required.

The agency identifier should be in the form [country code] dot [name]. For example: us.agencyname

Get your identifier today!

New Data Expression and Transformation Language Being Developed

Arofan Gregory, Metadata Technology North America, recently attended a three-day meeting hosted by the Bank of Italy to determine requirements for the “Expression and Transformation Language” (ETL), which will integrate with DDI. The meeting was well-attended, with participants from Australian Bureau of Statistics, Agilis, Bank of Italy, Eurostat, ILO, IMF, ISTAT, OECD, and UNESCO.

Use cases and DDI requirements were presented. The focus for the initial development is validation, with the idea that other types of processes will be supported in the future. The “transformation” capability will include operations such as recodes and the derivation of variables. The question of whether “round-tripability” with stats packages could be supported was discussed. This work is to be completed by July 2014.

Minutes of Executive Board Online

The DDI's newly elected Executive Board is very active, meeting every other month with full agendas and good discussions. You can find out more about what is being discussed by reading the [meeting minutes](#).

HLG Standards Meeting Takes Place

Convened by the UNECE, a new [Modernization Committee on Standards](#) comprised of representatives of national statistical organizations met in November 2013 in Geneva with Joachim Wackerow representing the DDI Alliance. The MC discussed key issues and priorities for modernization:

- Support the further development and implementation of the [GSBPM](#) and [GSIM](#) - this includes more work on links with other standards, particularly DDI and [SDMX](#), more detail on GSBPM overarching processes (particularly quality management), publishing use-cases and establishing practitioner communities
- Define additional DDI profiles to support the use of DDI in different aspects of official statistics
- Support greater standardization through SDMX data structure definitions, cross-domain code-lists, etc., in collaboration with the SDMX Statistical Working Group
- Support the implementation of the [Common Statistical Production Architecture \(CSPA\)](#) from a standards perspective.
- Investigate the potential use of semantic web standards in official statistics, including in relation to Big Data and Open Data, in collaboration with a proposed new W3C group on semantic statistics

- Promote the greater visibility and use of standards within statistical organizations, including as a prerequisite, to define what we mean by standards, and how they should be managed and maintained. This work could build on that done by the European Statistical System Sponsorship on Standardization, and other relevant European projects
- Develop a general glossary of metadata and/or modernization terminology, as a complement to the proposed SDMX Glossary
- Create a coherent overview of the standards needed for statistical modernization. This could include mapping existing standards and tool sets to the GSBPM, and outreach to other standards initiatives, e.g., [XBRL](#).

New DDI Representative

Bojana Tasic, head of the IT group from the [Swiss Foundation for Research in Social Sciences \(FORS\)](#), is the new FORS DDI Member Representative, replacing Andreas Perret. Welcome, Bojana!