‹ddi›
DIRECTIONS

A newsletter of the Data Documentation Initiative

As we approach our annual DDI Alliance meetings, taking place this year in Vancouver, I want to remind the DDI community of the many opportunities to become involved with DDI. Sam Spencer's work on the DDI Web site is a good illustration. See a summary of his interesing work on page 2 of this issue.

Mary Vardigan, Director, DDI Alliance

Several IASSIST Meetings Center on DDI

There are several DDI-related events and sessions scheduled in connection with the upcoming IASSIST meeting in Vancouver, May 30-June 3.

- Expert Committee ~ Monday, May 30, 8:30-17:00, SFU Harbour Centre Boardroom 2200
- Steering Committee ~ Tuesday, May 31, 8:30-12:00, 370 HSBC Executive Meeting Room, Morris Centre
- Developers ~ Tuesday, May 31, 8:30-17:00, 470 Hamber Boardroom, Morris Centre

Other DDI sessions are on the IASSIST program as well.

Dagstuhl Workshops Scheduled for September

Two advanced workshops will be held at Schloss Dagstuhl, Leibniz Center for Informatics, in Wadern, Germany, in 2011.

The first will be a workshop on "Semantic Statistics for Social, Behavioural, and Economic Sciences: Leveraging the DDI Model for the Web," to be held September 11-16.

The second workshop, "DDI: Managing Metadata for Longitudinal Data ~ Best Practices," a follow-up to the 2010 workshop on the topic, will be held September 18-23.

Save the Date~EDDI 2011

The 2011 European DDI Users Group (EDDI) will be hosted by the Swedish National Data Service (SND) in Gothenburg, Sweden, to be followed by a DDI Developers meeting and a Qualitative Data Workshop:

- December 5-6 ~ EDDI 2011
- December 7-9 ~ Developers Meeting
- December 7-9 ~ Qualitative Data Workshop (in parallel with the Developers meeting)

Meetings Focus on DDI-SDMX Collaboration

Representatives from the DDI and SDMX communities have been holding exploratory meetings to discuss how the two standards might work together to document end-to-end statistical data generation. The first meeting was held in the margins of the December 2010 European DDI Users Group (EDDI) meeting in Utrecht, and a second meeting was held in Lisbon, Portugal, in March 2011. A third meeting is scheduled to coincide with the Global SDMX Conference to be held May 2-4 in Washington, DC.

By Sam Spencer, Australian Bureau of Statistics

Over the last few months, you may have noticed small changes taking place on the DDI Web site. Late last year I took over the position of Web Site Maintenance Chair for the Alliance, with the goal of improving the site's search-engine rankings to increase the standard's exposure. Michael Iannaccone, a skilled Drupal developer from ICPSR, and I have been looking at how we can do this and also improve user satisfaction.

After a few technical additions to make it easier for search engines to crawl the site's pages, we were able to get some interesting data out of Google to look at how people are using, and more importantly finding, the site [see inset]. What these searches have told us, though, is who is (and is not) getting to our site from these search terms. This has led us to look at how the whole site hangs together, and look for new ways to link pages.

A new feature on the site is an updated live map of where DDI is being used, located on the Community page. Improvements to the map include the ability to zoom in and scan the map for different regions, and pop-out balloons describing the users. Over the coming months we will be sorting through the map to add or update users with more information to make it even more interactive.

You may also notice that we have also put up a word-cloud (shown above) on the site. This was created by mining our new sitemap, analysing each page on the site, and running it through the Tagxedo web-app to create a visual ranking for every word on the site. This is both an illustration and a tool that will enable the community to easily identify under- or over-represented words and help us customize our pages to drive more users to the site.

Finally, I would love to hear from you. If you notice an area of the site that is underdeveloped and want to help us write more content, be sure to send me an email.

Finding DDI

Here is a look at the common, and some of the more unusual, search terms people use to find and query the DDI Alliance Web site.

Analysing the searches coming, in we were able to identify five broad groups of people:

1. Lifecycle researchers

People interested in the statistical life cycle were the largest group of people over the period.

2. New to DDI

People searching for more information about DDI, and looking at what tools are available were our second largest group of visitors.

3. New to data research

These searches didn't even mention DDI, but were just asking about best practise and XML formats for codebooks.

4. Party Animals

What events are there? When is the next DDI meet-up? These folks are searching for past and future DDI events.

5. Vanity searches

So just who is this Arofan fellow anyway? Or Pascal? Bill Block? But, no searches for "Sam Spencer" came to the Web site ... yet!

DDI Review Continuing

The DDI review, which is being conducted by [Breckenhill Inc.](#) to address governance and IP issues for the DDI Alliance, is continuing and a final report will be available for discussion at the Expert Committee meeting in Vancouver on May 30.

DDI Specifications "Rebranded"

As recommended at the last Expert Committee meeting, the DDI specifications 2.* and 3.* have been "rebranded" on the DDI site as DDI-Codebook and DDI-Lifecycle, respectively. You can see the results of the new branding on the [specifications pages](#).

Work on DDI 2.5 Continues

Work is ongoing on this new version in the DDI-Codebook development line. An early draft is now available for insider review.

New Publications Now Online

Three of the papers resulting from the October 2010

Dagstuhl workshop on DDI and Longitudinal Data, mentioned in the previous issue of *DDI Directions*, have now been published as part of the DDI Working Paper Series:

["Enabling Longitudinal Data Comparison Using DDI."](#) This paper proposes best practice for the use of DDI to ensure that there are appropriate metadata to compare data produced across repeated data collections.

["Metadata for the Longitudinal Data Life Cycle."](#) This paper reviews the unique characteristics of longitudinal data and offers recommendations for structuring metadata in a manner that facilitates analysis over time and metadata reuse.

["Presenting Longitudinal Studies to End Users Effectively Using DDI Metadata."](#) This paper is intended to provide implementers and those delivering longitudinal data with recommendations on how to use DDI most effectively to support the presentation of longitudinal studies, most commonly on the Web, and to describe best practice for structuring DDI instances.

Forthcoming: "Documenting a Wider Variety of Data Using the Data Documentation Initiative 3.1". This paper looks at the growing variety of data sources in research ~ e.g., biomarkers ~ that are not traditional question-based surveys, and how these can be usefully documented.

In addition, a paper on ["Exploring the Relationship between DDI, SDMX and the Generic Statistical Business Process Model,"](#) written by Steven Vale, United Nations Economic Commission for Europe, has also been published as a DDI Working Paper in the "Other Topics" category. This paper was originally presented at EDDI 2010 in Utrecht.

DDI Agency Registry Work Proceeding

Algenta Technologies is working on a DDI agency registry to aid in URN resolution. The registry will be connected into the global DNS system. If you need an agency ID now, please email [secretariat @ddialliance.org](#) to receive an ID. DDI users who previously requested preliminary agency IDs will be notified in the next few weeks regarding the process for final registration.

DDI on the Agenda at Upcoming International Conferences

DDI will be represented at two large conferences in coming months.

Joachim Wackerow will chair [two sessions related to structured metadata](#) at the [European Statistical Research Association conference](#) in July in Lausanne, Switzerland. Presentations include:

[Overview of the Data Documentation Initiative (DDI) Metadata Standard: Goals and Benefits](#) ~ M. Vardigan

[Dissemination of Survey (Meta)data in the LISS Data Archive](#) ~ M. Streefkerk, S. Elshout

[DDI and the Lifecycle of Longitudinal Surveys](#) ~ L. Hoyle, J. Wackerow

[Metadata Driven Survey Design](#) - - J. Iverson

[Benefits of Structured DDI Metadata across the Data Lifecycle: The STARDAT Project at the GESIS Data Archive](#) ~ M. Linne, E. Brislinger, W. Zenk-Möltgen

[IPUMS to IHSN: Leveraging Structured Metadata for Discovering Multi-national Census and Survey Data](#) ~ W. Thomas

[The Use of Structured Survey Instrument Metadata throughout the Data Lifecycle](#) - - S. Hansen

[Microdata Information System MISSY](#) ~ J. Bohr

Wendy Thomas will give a Contributed Paper Session on "Mapping Out a Region of Interacting Standards" at the 58th [International Statistical Institute World Statistics Congress](#) in Dublin in August 2011.

Help Us Recruit New Members!

New members are always welcomed into the Alliance. If you know of projects using DDI that are not affliliated with a DDI membership, please encourage them to join the Alliance. Having more members results in our being able to accomplish more and making a greater impact.

DDI-Based Software Project Receives Funding

[Algenta Technologies](#) announced that it has recently received a grant from the National Institutes of Health - National Institute on Aging, under the National Small Business Innovation Research (SBIR) program for: "Open Standards-Based Data Extraction Web Tool for Complex Longitudinal Datasets".

The project aims to analyze the workflow needed to prepare a study for online data extraction, to validate the use of the DDI 3 standard for the basis of such a tool, and to create prototype Web-based data extraction software. While the focus is on longitudinal surveys, the proposed system would also handle cross-sectional, time-series, and non-repeated studies.

The aim is to improve research methodologies through a simplification of the process used for discovering, retrieving, and analyzing data relevant to a researcher's investigation and to improve data citations, aiding in reproducible research.