DDI Alliance Annual Meeting of Members and Scientific Board
Monday, May 27, 2013
GESIS-Leibniz Institute for the Social Sciences
Cologne, Germany

Minutes

Present:

- Nikos Askitas (Institute for the Study of Labor – IZA)
- Eleni Castro (Harvard University, Institute for Quantitative Social Science)
- Ingo Barkow (Institute for International Education Research -- DIPF)
- Donna Dosman (Statistics Canada)
- Tom Ensom (United Kingdom Data Service – UKDS)
- Arofan Gregory (Metadata Technology)
- Chuck Humphrey (University of Alberta), Chair
- Sanda Ionescu (Inter-university Consortium for Political and Social Research -- ICPSR)
- Nanna Floor Clausen (Danish Data Archive -- DDA)
- Peter Granda (Inter-university Consortium for Political and Social Research -- ICPSR)
- Alistair Hamilton (Australian Bureau of Statistics -- ABS)
- Uwe Jensen (GESIS – Leibniz Institute for the Social Sciences)
- Mari Kleemola (Finnish Social Science Data Archive – FSD), Vice Chair
- Amber Leahey (University of Toronto, Scholars Portal)
- Jared Lyle (Inter-university Consortium for Political and Social Research -- ICPSR)
- Hans Jørgen Marker (Swedish National Data Service)
- Steve McEachern (Australian Data Archive -- ADA)
- Katherine McNeill (Massachusetts Institute of Technology – MIT)
- Ron Nakao (Stanford University Libraries)
- Alistair Hamilton (Australian Bureau of Statistics)
- Andreas Perret (Swiss National Data Service – FORS)
- Tom Piazza (University of California, Berkeley, Computer-Assisted Survey Methods -- CSM)
- Anita Rocha (University of Washington, Center for Studies in Demography & Ecology -- CSDE)
- David Schiller (Research Data Centre of the German Federal Employment Agency, Institute for Employment Research -- IAB)
- Dan Smith (Colectica)
- Jon Stiles (University of California, Berkeley, UC DATA)
- Wendy Thomas (University of Minnesota, Minnesota Population Center)
- Mary Vardigan (Inter-university Consortium for Political and Social Research -- ICPSR)

- Joachim Wackerow (GESIS – Leibniz Institute for the Social Sciences)
- Catharina Wasner (GESIS – Leibniz Institute for the Social Sciences)
- Marion Wittenberg (Data Archive and Network Services – DANS)
- Wolfgang Zenk-Möltgen (GESIS – Leibniz Institute for the Social Sciences)

Introductions and Overview of Alliance Transition

DDI Alliance Expert Committee Chair Chuck Humphrey opened the meeting with the acknowledgment that, as the Alliance was transitioning to a new governance model and organizational structure, this would be his last meeting as Chair.

Humphrey reminded the group that the process to change the way the Alliance operates was initiated with the External Review in 2010, which recommended ways for the Alliance to prepare for future development, including revising its governance structure. Some of the strengths of the Alliance identified in the review were the diversity of the membership (university libraries and departments, domain repositories, national statistical organizations, research centers, and more), and the cooperative, volunteer orientation of the members.

Humphrey went on to outline some of the aspects of the governance transition. In the past year the Alliance wrote and approved a new Charter and Bylaws, which paved the way for an Executive Board (to replace the Steering Committee as the oversight body for the Alliance) elected by Designated Member Representatives. A new membership form was also drafted, and DDI members must now sign new forms to establish the Alliance as a program of the University of Michigan.

Executive Board Election

Humphrey described the roles and responsibilities of the new Executive Board. Among the purposes of the Executive Board are to set overall policy and budget and to provide strategic guidance and review of the activities of the Alliance.

The Executive Board will be composed of seven voting members: six At-Large members elected by the Designated Member Representatives and one member appointed by the Host Institution, with the Executive Director serving as an ex-officio member. Nominations for Board positions were solicited in April with the election to occur by electronic ballot in June.

The preliminary slate for the Executive Board was reviewed. For this first election, three Board members will serve two-year terms and three will serve four-year terms to ensure continuity. Humphrey opened the floor for additional nominations and a potential nominee from GESIS was suggested (this was later withdrawn). It was noted that if there were more than three candidates for a two- or four-year term, the election would be competitive.

Strategic Priorities

The DDI Director Mary Vardigan outlined three high-level priorities for the Alliance in a paper presented at the recent METIS meeting in Geneva. The three priorities were restructuring the organization, which is under way; developing a next-generation model-based DDI specification (this activity was approved in May 2012 in Washington, DC); and building new partnerships through sustained outreach activities, which can move the organization forward and lead to new memberships.

The members were asked if there were other priorities to pursue, and the issue of creating membership categories was raised. It was noted that during the 2012 Alliance meeting there had been a discussion of implementing a tiered membership structure to generate additional revenue for the Alliance. The consultant who conducted the External Review suggested that the Alliance consider this type of membership structure. The Steering Committee reviewed some different models of tiered membership during the course of the year, but the discussions did not reach the level of a formal proposal. This topic should now be taken up by the new Executive Board, which has responsibility for setting fees. The current fee ($2500 a year) has not been increased in the ten years since the Alliance came into existence. A more robust business model needs to be considered by the new Board.

The issue of the importance of having an interoperable standard was raised. How should the DDI Alliance, which has an interoperable metadata standard, position itself with respect to opportunities like Big Data and the many data-related initiatives springing up? It is important to align DDI with other standards, which is the strategy the Alliance is taking in working with the Generic Statistical Information Model (GSIM) and the Statistical Data and Metadata Exchange (SDMX) standards.

GSIM is an effort to modernize statistical production in the NSIs and allow these organizations to "survive and thrive" in a world that is changing rapidly in terms of technology and user expectations. A project of a High Level Group (HLG) of National Statistical Institutes (NSIs) and the United Nations Economic Commission for Europe (UNECE), GSIM is a conceptual model designed to help standardize the production of official statistics.

Coordination with metadata standards for geospatial data was suggested as another area that could be important for the Alliance going forward. While there is content for geography and geospatial data in DDI, we should try to coordinate support around this key topic and perhaps pursue more formal relationships in order to be recognized in this space.

Supporting data management plans (DMPs) is another area for consideration. A project at the University of London explored mapping the elements of DMPs against DDI. Since DDI was not designed to support data management planning, the specification is not currently structured properly to make a robust mapping. The Research Council in the UK agreed on some new fields that are needed. This discussion raised the issue of what DDI's role should be in research administration in general. There is also an initiative to publish data analysis plans before collecting data, which the standard might cover as well. . In addition, there is a need to cover the process of data analysis and related syntax files.

With the multiplicity of emerging virtual research environments and recent news stories about data errors, the Alliance needs to think about replicable research and how DDI can support this. The DDI standard should document workflow in both a machine-actionable and a human-readable way. Replication is relevant to both the academic world and to NSIs.

Ideally, data transformations should be captured in a system-independent language. SDMX is now exploring this and the DDI Alliance is part of those discussions. It was pointed out that there is currently an effort to create an XML standard to capture regressions through the open Knowledge Foundation. Disclosure review and remediation is another area that DDI should cover.

The point was made that this discussion about documenting new types of data and reaching new communities represents where we are in the evolution of our organization and standard. Many of us are talking to faculty and trying to change the culture to encourage lifecycle metadata collection. It's a balancing act to improve the standard but also to decrease the burden of data documentation for researchers. We need to play a leadership role in this way.

Roadmap Supporting Model-Based Specification

Moving to a model-based specification is a key goal of the DDI Alliance for several reasons. An information model will help us to interoperate and communicate with other disciplines and standards efforts. It will also offer flexibility in terms of technical expressions in that the model may be rendered in a variety of technical bindings, including XML and RDF. Further, it will serve as a visual representation of the standard to facilitate understanding of what DDI covers. Finally, a model will make it easier to develop and maintain the DDI specifications in a consistent and structured way and will enable software development that is less dependent on specific DDI versions.

The group reviewed and discussed a roadmap to support the development of the model. This roadmap, which spans the next 18 months, lays out what the Alliance would like to achieve in terms of deliverables and their timing.

Some parts of a conceptual model for DDI already exist and can be helpful in building the next-generation specification. Further, DDI 3.2, which will soon be released, has rich content to draw from in modeling.

One of the important model principles is iterative design so that there will no longer be "big bang" releases but instead more frequent releases of small parts of the specification, which can then be more easily reviewed, with implementers and other users contributing feedback into the process. A beta namespace is one approach to iterative development.

With respect to feasibility of the timeline in the roadmap and whether it should be extended, it was mentioned that NSIs expect a common core of the DDI specification to be developed in the time frame proposed (by December 2014). NSIs can contribute resources and are interested in tools.

The all-important question of resources to support the roadmap was raised, which the group decided to come back to in the discussion of finances.

Date and Location of Next Annual Meeting of Members

The 2014 Annual Meeting will take place in Toronto, Ontario, Canada, on Monday, June 2, 2013, before the IASSIST conference, which will be held on June 3-6, 2014.

Financial Position

The group reviewed the current budget for the 2013 Fiscal Year, ending June 30. It was likely that the year would close with a small surplus, and the overall reserves of the Alliance would be a little over $100K.

The budget for Fiscal Year 2014 shows a deficit for the year if the anticipated expenditures are all realized. With the addition of the NADDI meeting, there are more yearly expenditures. The next fiscal year will probably also include payment for the registration of the DDI collective mark, which is likely to cost around $8000. The mark needs to be registered in each country represented in the membership. Copyright for the Alliance will also be registered, but this is a fairly low fee. The DDI mark will be represented by a logo that all the members can display to demonstrate that they are part of the group developing the specification; it will be possible to tailor a tag line to display as well – along the lines of "powered by DDI" or "using DDI."

There are three new potential members for the Alliance, which will increase revenues for FY2014. So far since the inception of the Alliance in 2003, we have been able to overspend in some years by drawing from limited reserves. It was pointed out that the organization could benefit from looking farther out in terms of revenues and expenses and creating something like a three-year plan going forward. This will actually be necessary to understand the expenses that creating the model will incur. The Alliance fees have remained steady for ten years and some consider them to be low. Some organizations would be willing to pay more per year.

Over the past year the Steering Committee looked at two models for a tiered fee structure. One was based on the institutional commitment to DDI, with those institutions committed to it and using it extensively paying more. The other model was more objective and based on size as measured by number of employees. Neither model had enough support or was fleshed out enough to be brought to the committee for consideration.

With a tiered membership and some members paying more, the obvious question is what extra benefits do these members receive? Should they have more influence in terms of voting? The Alliance has always been based on every member having one vote, so this would be a change in culture and perhaps difficult to accept.

A mixed business model in which organizations could contribute in-kind, which has always been the case, was also discussed. How can we better leverage member contributions? We will have a need for project management going forward. This might be a contribution from a member. In looking at the roadmap and the resources required, we should add some text about in-kind contributions. One idea raised involved lowering fees based on in-kind contributions, but monetizing such contributions is challenging.

Sponsorships are another good mechanism to garner additional funding. We could set up a fund to which institutions could make tax-deductible donations. This is done successfully by many organizations, especially when there are unexpected surpluses at the end of the year that need to be spent out. The Alliance should make it easy to do this. If we outline special projects, some organizations could say that they cannot contribute in money but can contribute time of employees. This would not erode the spirit of volunteering.

Anticipating expenses for the next three years will give the Alliance an idea of the amount of revenue we will need and we can attempt to adjust fees accordingly. An informal poll of meeting attendees showed that each organization could afford at least another $250 in membership fees so it was recommended that this go forward to the Executive Board for consideration.

NSIs and DDI Membership

Following up on an action item from the last meeting, Alistair Hamilton of ABS mentioned that he is interested in communicating with NSIs not yet in the membership regarding the value of becoming Alliance members. He wants to gauge interest in joining and also to understand what the NSIs would expect from membership. Since some of the NSIs on his list were currently considering membership, he had not yet contacted them. Ultimately it would be good to have a sub-group of NSIs in the Alliance. This would demonstrate to the HLG that the fit between GSIM and DDI is good and that the NSIs find the DDI standard useful.

Several NSIs are now using DDI or are interested in it. ABS, Statistics New Zealand, and Statistics Denmark are all using DDI in their organizations as are the FAO and INSEE in France. Statistics Denmark is using Colectica to provide data quality reporting. Data quality is very important to NSIs and there was interest in DDI covering this more explicitly in the next version.

NSIs are also interested in the development of a DDI profile for NSIs. This profile may be an outcome of the mapping work that is going on now to understand the relationships among GSIM, DDI, and SDMX. It

has been determined that over 70 percent of GSIM corresponds to content already in DDI while SDMX covers a little over 30 percent.

EDDI and NADDI Reports

EDDI 2012 was held in Bergen, Norway, in December, and the first NADDI meeting was held in Lawrence, Kansas, in April 2013; both user conferences were very successful. NADDI 2013 brought together the library and research communities, which was a fresh new approach.

EDDI 2013 will take place in December in Paris, France, at the Reseau Quetelet. NADDI will be held in Vancouver in April 2014 and will be co-hosted by cohosted by the University of British Columbia, Simon Fraser University, and the University of Alberta.

The Alliance should work on a shared infrastructure for EDDI and NADDI. It should be possible to create a reusable Web site that can handle the program, registration, etc. It may be necessary to customize the components that deal with taking in money as local institutions have different needs and procedures to follow.

Other DDI Groups, Products, and Events

- User groups have also met at the Australian Bureau of Statistics and at Statistics New Zealand; a national DDI meeting was recently held at Statistics Denmark.
- Mari Kleemola is chairing a metadata group of Finnish research data producers. They are using DDI controlled vocabularies to cover several disciplines.
- Eurostat held its first workshop for statistical agencies on DDI and SDMX.
- There was to be a meeting later in the week arranged by Statistics Sweden about the Data without Boundaries project. On the agenda was researcher needs for metadata, including metadata for longitudinal registers and Big Data.
- Danish health researchers are using DDI to document emergency injuries.
- To get funding in the Netherlands, researchers have to document their data using DDI. The DANS archive is working with CentERdata, which provides tools to researchers while DANS guarantees the long-term persistence of the data.
- Health researchers in Ontario at the Population Health Improvement Research Network (PHIRN) are coordinating efforts around data sharing and have published a DDI-compliant metadata index to provide information on databases available to researchers on the health of populations at international, national, provincial and health region levels.
- The International Household Survey Network (IHSN) is spreading DDI around the world. The Census of the Caribbean uses DDI Codebook as its documentation standard. Mexico is expanding its use of DDI and IPUMS international census data for Africa and South American use DDI.

DDI Marketing

Larry Hoyle and Joachim Wackerow have discussed ways to improve the DDI Alliance working paper series, including instituting another track for papers that includes a peer review process and possibly creating an open access journal. It was also suggested that the Alliance have DDI papers indexed in other systems like SSRN.

The Alliance should think about revamping the Web site to freshen its appearance and content. We might make some minor changes without a total redesign of the site. We don't want to change URLs, but we do need to think about how users (especially new users) navigate the site and find what they need. Users are currently confused by the site and don't know where to start.

It was also pointed out that Alliance funds are limited and should be directed towards the roadmap this year for greatest impact.

Roadmap and Resources

As noted earlier in the meeting, financial planning for longer than one year is needed, and the Alliance must identify costs associated with the stages in the roadmap. The roadmap takes priority this year, and we need to move to a new draft of the roadmap with resourcing and costs embedded.

This was Chuck Humphrey's last act as DDI Alliance Chair. He and Vice Chair Mari Kleemola were thanked for their service to the Alliance, and the Annual Meeting of Members was adjourned.

Meeting of the Scientific Board

This was the first meeting of a separate Scientific Board, chaired by Alliance Director Mary Vardigan in advance of a Chair and Vice Chair for the Scientific Board being elected.

Status of DDI Lifecycle 3.2, DDI Codebook 2.5.1

Technical Implementation Committee (TIC) Chair Wendy Thomas updated the group on the progress towards publishing the newest versions of DDI Lifecycle and Codebook. With respect to DDI 3.2, Wendy detailed the types of changes that were being reviewed and noted that small groups have assisted in coming up with the best solutions. More extensive documentation is being added to this version and some initial coverage of a quality framework is being added. Changes have been made to the comparison map, and the group is completing a final review of the data element object to better align with GSIM.

Documentation is still in development and needs examples of code. There will be a brief period for a review of 3.2 prior to the vote. A 3.1.1 version has been proposed to address bugs so that users do not need to do workarounds, but this will not require a namespace change.

DDI Codebook 2.5.1 is close to being finished. There was an element that was missed in 2.5 due to an oversight and the TIC has eased cardinality to support multiple languages. There is a requirement of backward compatibility for the DDI Codebook line.

The documentation for the new elements needs to be fleshed out, and Sanda Ionescu and Mary Vardigan volunteered to finish this.

Training

The ABS has suggested that the Alliance create some online training resources and videos similar to those developed for SDMX. The SDMX videos targeted at upper management have been particularly successful and DDI could use some material to communicate with this type of audience. This should be a broad and general overview of the standard, conveying the advantages of standardized and structured metadata. SDMX hired a company to produce their videos, so this can be an expensive proposition. There are technical tutorials, written guides, and self-assessments. We need materials that make a compelling high-level business case for DDI that organizations' leadership can learn from.

We want self-paced learning materials in general. Is video the best format for this? Perhaps some Web-based content could be developed first and then we can move to a scripted PowerPoint before undertaking video.

Developing a full library of DDI materials for different audiences at different levels will take some time to build. This ties into NADDI, which this year sought to bring librarians into the discussion. We need to develop materials for librarians and graduate students and integrate DDI into U.S. academic and research environments.

It was suggested that we keep this simple. We can do a series of Webinars such as "DDI for Librarians." This would be most effective if tools were recommended. We can also set up streaming for courses that others are giving now in order to jumpstart a library of content.

The Alliance can serve as a clearinghouse for materials being developed by Alliance members and others. We do have training slides on the DDI site, but they are not broken out as such, and lots of materials have not been shared yet. Our user communities are growing now -- can we call upon members to contribute content for training?

We need to make it easier to share with the clearinghouse approach. Training materials from Dagstuhl will be shared with a Creative Commons license.

Consortia Advancing Standards in Research Administration Information (CASRAI) in Canada has a [short video](#) that is very compelling. This was a marketing effort to promote the standard and convey why it is important.

How much do we want to focus on this type of training? Again, as with improving the Web site, we may not want to spend too much of the Alliance's monies on training as we have decided that the roadmap is the Alliance's uppermost priority. Maybe we can get by with "good enough."

Turning to face-to-face training, the group reviewed the DDI Alliance training principles developed in December 2011, which were based on the Data Liberation Initiative principles. The principles were intended to communicate a model for community training supported by the Alliance. However, it was pointed that that there appears to be a disconnect between the training principles and what is actually happening. The principles communicate a vision for training, but do not describe the current training picture.

What is missing is text on training provided by consultants, which is also part of the training being provided. Consultants are concerned because if the Alliance develops courses at low prices, it undercuts the commercial companies that might want to make a profit from providing training.

The community and consultant training should not be competing models – there is room for all. While the Steering Committee previously endorsed Alliance involvement in community-based peer-to-peer training for DDI, it is a big enough world to hold both training models. An explicit policy around consultant training is lacking and not found in the principles. We need to add a statement to the principles about consultant training.

Are the Dagstuhl trainings community based? Some think they are while others disagree. Dagstuhl is a one-week training organized and funded by GESIS. Wendy Thomas and Arofan Gregory are invited to Mannheim to work with Joachim Wackerow and prepare for Dagstuhl in advance of the training week. There is a stipend in addition to travel and accommodations, but it is modest.

The Alliance has successfully broadened the community of trainers at EDDI and NADDI with new trainers doing half-day tutorials. Continuing in this vein, it was suggested that we invite a third trainer to Dagstuhl to learn how to do intensive training. This person would spend a week in Mannheim with the other trainers and a week in Dagstuhl; a small stipend would be an additional motivator. It was recommended that to request support for this model, GESIS should submit a one-page proposal to the Executive Board for support in the FY13-14 year.

A comment was made that this discussion around training reflects Alliance "growing pains." We were more informal before and now things need to be codified and principles and policies formulated.

Technical Committee Roles

The TIC has spent some time developing a structure, roles, and process for the Technical Committee laid out in the new Bylaws. The process of producing the standard needs to be transparent, and it is ideal to have separate roles on the Technical Committee to avoid conflicts of interest. The general idea is to define specific roles for specific tasks with roles represented by different people so that one person plays only one role. This results in less overload and clear separations of duties; further, it professionalizes DDI development. We want to stabilize and control the standard and make it more robust. Modeling and documentation have to be done correctly.

The TIC has been shrinking and an injection of new people is needed. More involvement will make the committee better. Currently, there are not enough people for the roles. Another challenge is to coordinate the roles. The Chair and Integrator can help with this, but it will be difficult.

The TC has candidates in mind for the new roles and would like to bring additional people into the technical work. We are likely to bring in a lot of new content from domain experts, so we need to make sure we have the capacity to handle this. We need a project manager also. We should put out a call for volunteering for the TC.

This is a natural evolution and will help the rest of the organization to move forward. We need to make sure we are sustainable for the long term by taking advantage of the new Bylaws and recruiting more people with scientific and technical expertise.

Process for Changes to the Specification

The process for changing the specification has been broken out from the Bylaws and is now a separate document; it has been developed to a point where it is close to being ready for publication. We need to align the document with the Charter and Bylaws terminology and complete it. Alistair Hamilton and Steve McEachern volunteered to help the small team that has been working on this.

Update on GSIM

Arofan Gregory provided the group with an update on the work around GSIM. The data world is changing rapidly and NSIs around the world need to keep pace with the needs of their constituencies to remain relevant in meeting needs. Accordingly, a group of directors of NSIs have begun a major initiative to modernize the production of official statistics, and key to their vision are standards and a common architecture.

Both GSIM and the Generic Statistical Business Process Model (GSBPM), which GSIM is designed to support, serve to provide a set of terms that permit everyone to communicate around an agreed upon vocabulary to ensure that everyone is talking about the same thing. GSIM provides definitions, attributes, and relationships for information objects. It is a conceptual model that sits above

implementation standards like DDI. There are four main areas: business, production, structures, and concepts.

ISO Update

The DDI Alliance wants to move forward to obtain official ISO status for the DDI standard. SDMX recently achieved this after many years. An organization must go through a complicated set of steps that involves voting by many nations. Expenses related to ISO registration are difficult to predict, but the World Bank has promised some support. It was pointed out that this activity should be part of our long-term future.

SDMX set up their ISO registration so that all SDMX versions offered on the SDMX Web site are covered. The intent is for DDI to take the same approach so that both DDI Codebook and Lifecycle would receive certification. Having the model at the center would be ideal.

The next step is for Arofan Gregory to support the World Bank in getting the proper status in the ISO community to enable the ISO registration process for DDI to begin.

Reports of Working Groups

Various DDI working groups reported on activities during the year:

[Controlled Vocabularies](#)

[Qualitative Data](#)

[DDI Tools Catalog](#)

[Technical Implementation Committee](#)

[Web Site Maintenance](#)

[Survey Design and Implementation](#)

[RDF Vocabularies](#)

[DDI Developers](#)

Request for New Experimental Data Working Group

This area was raised as an important one for DDI to cover and Kate McNeill is interested in pulling together a working group. There is a lot of interest in better documentation of experimental data in

economics, political science, psychology, health, and related fields. It was pointed out that a good process model for DDI could describe experiments.

The question was raised about how the working groups, including this new one, articulate with the domain expertise required for creating the DDI model as outlined in the roadmap. The working groups need to be integrated into the roadmap and we need to raise the communication level between and among the groups. Given this need, it was decided that the individuals interested in documentation of experimental data should form a working group.