

**DDI Alliance Expert Committee Meeting Minutes**  
**Vancouver, British Columbia, Canada – Simon Fraser University**  
**May 30, 2011**

**Member Participants**

Nikos Askitas (Institute for the Study of Labor -- IZA)  
Michael Beahan (Australian Bureau of Statistics -- ABS; via telephone)  
Sami Borg (Finnish Social Science Data Archive – FSD)  
Dan Gillman (US Bureau of Labor Statistics)  
Chuck Humphrey (University of Alberta), Chair  
Sanda Ionescu (Inter-university Consortium for Political and Social Research -- ICPSR)  
Vigdis Kvalheim (Norwegian Social Science Data Service -- NSD)  
Nanna Floor Clausen (Danish Data Archive -- DDA)  
Rutger Kramer (Data Archive and Network Services – DANS)  
Stefan Kramer (Cornell University, Cornell Institute for Social and Economic Research -- CISER)  
Hans Jørgen Marker (Swedish National Data Service -- SND)  
Marc Maynard (University of Connecticut -- Roper Center)  
Steve McEachern (Australian Data Archive -- ADA)  
Katherine McNeill (Massachusetts Institute of Technology – MIT)  
Ron Nakao (Stanford University Libraries)  
Anita Rocha (University of Washington, Center for Studies in Demography & Ecology (CSDE)  
John Shepherdson (United Kingdom Data Archive -- UKDA)  
David Schiller (Research Data Centre of the German Federal Employment Agency, Institute for  
Employment Research -- IAB)  
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)  
Wolfgang Zenk-Möltgen (GESIS – Leibniz Institute for the Social Sciences)

**Observers**

Thomas Bosch (GESIS – Leibniz Institute for the Social Sciences)  
Adam Brown (Statistics New Zealand)  
Peter Granda (Inter-university Consortium for Political and Social Research -- ICPSR)

**External Review – General Discussion**

After a round of introductions, DDI Alliance Chair Chuck Humphrey initiated a discussion of the recent DDI Alliance External Review. He underscored the fact that the review was a result of the success of the Alliance. In recent years the organization has been growing and transforming and bringing new stakeholders into the mix; it was a good time to reflect, to take stock, and to decide what the members want the Alliance to be. The review is overall a healthy process.

The External Review, conducted by the consulting firm Breckenhill Inc., was an arm's length look at various aspects of the Alliance. The Review had several specific objectives, the main ones being to bring clarity and possible solutions to issues around governance and intellectual property. Many Alliance members and other key stakeholders were interviewed as part of Breckenhill's review.

It was pointed out that the Expert Committee (EC) should view the External Review report as evidence but not the only evidence to be considered in thinking about the future of the Alliance. After reading the report thoroughly, the Steering Committee (SC) made its own analysis and established its own priorities, some of which were being brought to the EC for consideration and approval (see Steering Committee Recommendations for Approval below). This is an important process and the Alliance should not accept all recommendations and conclusions without first discussing and agreeing to them.

The Chair sought some general reactions to the External Review report before delving into specifics. Committee members found the report easy to read with quite a lot of research to back up its conclusions. While all members may not agree with all of the recommendations, it was seen as quite good overall, a fair report on the status quo, and meeting the terms of reference initially laid out. It was also viewed as appropriate and timely. There was some disagreement expressed regarding aspects of the history of the DDI effort. It was pointed out that a summary version of the report should be made public.

## **External Review – Specific Topics**

**Role of the Technical Implementation Committee (TIC).** On this topic, the point was made that the TIC has evolved into a group with multiple functions, which should be pulled apart going forward. First, TIC has the role to model the specification according to the requirements put forward by other working groups. No UML model has yet been published, but that is an important task for the Alliance. A second function is to implement the model in XML, or in other expressions. If the model is good, expressing or representing it is fairly straightforward. The third role is coordination and lower-level steering across all groups. These functions should be separated into three distinct groups. For the coordination aspect, it was also suggested that we may need a group that is not as large as the full Expert Committee (EC) that sits between the EC and TIC. Another suggestion put forward was for an ombudsman role.

**Governance.** It was pointed out that the organization has organically developed in a certain way but that it is important now to restructure the Bylaws to meet current needs. The new model established in the draft Bylaws provided in the External Review report separates the functions of the Expert Committee in terms of their roles as member representatives and content experts. To fulfill both roles, it may be that the EC needs to meet more often and to have more focused and frequent interactions. Communication across the Alliance becomes even more crucial.

The draft Bylaws contain helpful suggestions, but the final form they take must come from the EC. The point was raised that the governance structure is important but is ultimately only the means to an end, and we need a solid mission statement that communicates the goals and objectives of the Alliance. We

need to pull out the important principles and articulate our common values. A division between the Charter and Bylaws is important.

We need to distinguish between the EC and the SC and can look at how other standards bodies do this. Some like INCITES have strong rules and delineations between groups. Changing the SC to be more representative of the membership seems a good step forward. We also need a way to be more specific about how the EC does its work and sets deadlines.

Another point raised in the External Review report was that a small number of people are doing the lion's share of the work, and this is not sustainable. It needs to be understood going forward that members are expected to become engaged when they join the Alliance. We arrived at where we are now as a sort of club, but going forward we need to ensure that new people can become involved on an ongoing basis. Our success has been around a spirit of collaboration, and we need to extend this spirit broadly. New people coming in will need to understand the division of labor and the possible opportunities for engagement in order to see where they can plug in to the DDI community. Working groups may include people who are not members, which is a good thing.

**Membership model.** The report suggests that the Alliance consider a tiered membership model with differential pricing. Currently the Alliance has a thin budget line and new revenue sources will be important. We need to ensure that a tiered structure does not put people off. It is important to think about the Associate membership also, which enables organizations like the U.S. Bureau of Labor Statistics to be a member without fee and without vote. Other disciplines now see DDI as relevant because it is at its core an observational standard. We want a mission of inclusion and can look to communities and disciplines not originally seen as having interest in DDI.

With respect to soliciting new members, we need to build a case for what is in it for them. In tough economic times, the DDI subscription may be seen as expendable.

**Revenue.** A presentation of the Alliance budget was made showing that the estimated reserve carrying forward to the next fiscal year (beginning July 1, 2011) is roughly equal to one year of membership revenue, or about \$75,000. It is important to consider how much revenue we need to accomplish our objectives. The External Review report indicated that there is on the order of \$200K going into the Alliance when in-kind contributions across the membership are taken into account. That means that there are two dollars of in-kind contributions for every dollar of membership revenue. We also need to think about how in-kind contributions are recognized as there are people putting in hours rather than dollars. Also, do we have any kind of free rider problem?

## **Steering Committee Recommendations for Approval**

Drawing from the External Review report, the SC brought forward four recommendations for action at this meeting and they were considered in turn.

## **1. Register a Collective mark for DDI in the U.S. and other countries to protect its intellectual property**

The External Review report recommended that ICPSR remain the host organization for the Alliance Secretariat and that the Alliance establish a Collective mark for the specification to be held and policed by the University of Michigan, the administrative home of ICPSR. A Memorandum of Understanding between the UM and the Alliance would be developed with the UM waiving any revenues that might accrue to the mark. There was an in-depth discussion of this relationship as there are dependencies and advantages.

Some confusion was expressed about the difference between the Collective and Certification marks and how the Alliance would use its mark. The intent of the mark is to protect the intellectual property of the Alliance from being hijacked by a company or even by an Alliance member. The Collective mark signals membership in a group while the Certification mark indicates compliance with a set of requirements.

Certification is a complex topic. The Alliance could possibly detail best practice for use of the standard and then grant the mark to implementations that conform to the best practice. There is a whole field of conformity testing to guarantee interoperability. Conformity can be measured from the instance level through the repository level with increasing demands across the range of levels. The Alliance could also decide to apply a more lightweight certification method like the Data Seal of Approval, which involves a self-assessment and peer review based on trust. The Alliance could also decide to certify trainers. It might be possible to use different taglines with the mark to distinguish the type of certification.

The current use of Alliance copyright and the GNU open source license is seen as inadequate protection. Because the DDI specification is intended to be publicly available, copyright does not apply. Questions were raised as to whether the Collective mark and the GNU license could be used in concert and how the University of Michigan could waive rights to the DDI specification when the Alliance is not a legal entity. Overall, this move to a Collective mark has to be seen as an improvement over the status quo or there is no incentive to pursue this course. A Memorandum of Understanding must spell out the relationship in detail including procedures for dissolving the relationship.

Since there was no consensus on the issue and questions remained on the relationships among the GNU license, copyright, and the Collective mark, it was decided to table the motion for now. The Director will gather additional information from UM and come back to the Committee with a clarifying document.

## **2. Rebrand the standard as DDI-Codebook and DDI-Lifecycle and communicate this change widely**

This item had been covered previously and the DDI-Codebook and DDI-Lifecycle terminology is now deployed on the DDI Web site. This needs to be carried through all Alliance documents and Web pages, including in the controlled vocabularies and the tools catalog.

### **3. Begin a process to revise the Bylaws, including the governance structure, over the course of the next year**

A motion was made and passed to establish a Bylaws Task Force to formulate a new set of Bylaws and to report back by the 2012 meeting. A call for volunteers was made and eight people signed up.

### **4. Adopt the Draft Conflict of Interest policy and collect statements from relevant parties**

The External Review report discussed conflict of interest as it pertains to the Alliance and concluded that while there have been no direct accusations of conflict of interest there is the perception that this is a possibility given the small set of people participating in the Alliance and being paid. The report recommended that the draft Conflict of Interest policy first reviewed by the EC in 2009 be approved and put into practice.

The original intent of the 2009 COI policy was that it would apply to all members of the EC and those being paid by the Alliance. There was resistance to the Expert Committee representatives having to sign such a document on behalf of their respective universities or organizations because of a concern that legal counsel might have to review the forms. At the 2009 meeting it was decided that the COI policy would apply only to those being paid by the Alliance.

The point was made that while we cannot legislate COI out of existence, we can have the expectation that people disclose activities that could be problematic. We need to have a clear set of guidelines on the types of roles and relationships that the Alliance is interested in and there should be some indication of what the consequences might be. Anyone on the EC or being paid by the Alliance should complete a form if there are existing relationships that could cause conflicts of interest. These relationships must be made transparent insofar as that is possible. This should be part of the Bylaws also.

Because the draft COI policy was not as detailed on the above points as desired, it was decided that the Director would redraft the document, adding information related to types of relationships to be disclosed and consequences for violation and then bring this back to the group. Once accepted, the policy would require disclosure by all relevant parties and the disclosure forms would be reviewed by the SC.

## **2012 Meeting Date**

Since 2007 the Alliance Expert Committee has been meeting before the IASSIST conference on Monday of the IASSIST week. This is usually the Memorial Day holiday in the United States, so it is not an ideal meeting day. A motion was made to move the 2012 meeting to Saturday after the IASSIST meeting concludes in Washington, DC. The motion carried.

## Reports of Working Groups

**Tools Catalog Working Group.** Chair Stefan Kramer reported on progress in establishing a Web-based DDI tools catalog and provided a demo of the work in progress. The new tools catalog page will replace the current flat list of tools provided on the DDI site. The EC saw this tools inventory as a very positive development and suggested that functionality be added so that one could sort and filter across different dimensions – for example, view all editor tools or view the most recently developed tools. More information from the Tools Catalog group report is available on the DDI site:

<http://www.ddialliance.org/sites/default/files/ToolsCatalogReport2010.pdf>

**Survey Design and Implementation Working Group.** This group was established in 2008 with two co-chairs, Dan Gillman and Peter Granda, who reported out to the EC about their work. The group was established to look at survey design aspects not currently in DDI, specifically related to sampling and questionnaire design. This work is now concluding and the working group suggested additional topics to be tackled next by new committees, including paradata. Sue Ellen Hansen has volunteered to lead this group. It was suggested that the new group might talk to people in the technical metadata area. Another task for the group might be to harmonize with the AAPOR transparency initiative. Other groups may be needed on the topics of non-response adjustment, weighting, and estimation. More information from the SDI group report is available on the DDI site:

<http://www.ddialliance.org/sites/default/files/SurveyDesignImplementReport2010.pdf>

**Controlled Vocabularies Working Group.** The current Chair of the group, Sanda Ionescu, reported that the first set of eight vocabularies has been published and is now available in HTML, Genericcode XML, and Excel formats on the DDI Web site. Several other vocabularies are under development. It was reported that archives like GESIS and ICPSR have plans to use these vocabularies. A suggestion was made to create a mechanism to cite the vocabularies -- in effect, to use a metadata citation standard. It was pointed out that the DataCite metadata schema has been published with a Digital Object Identifier (DOI). It will be useful to see the uptake of these vocabularies and to capture download statistics and user feedback. DDI Lifecycle 3.2 will be open to additional vocabularies. More information from the CVG report is available on the DDI site:

<http://www.ddialliance.org/sites/default/files/ControlledVocabReport2010.pdf>

**Technical Implementation Committee (TIC).** Chair Wendy Thomas reported on the work of TIC for the year. The group is currently working on new releases in both development lines, DDI Codebook (Version 2.5) and DDI Lifecycle (Version 3.2). They are also considering a survey to better understand the types of information that users want to document as we look ahead to 4.0.

There was a discussion of software that might be used for bug tracking, for gathering suggestions for DDI improvements, and for general transparency and communication around the specification and the

work of the Alliance. JIRA was suggested as a good system and the TIC will look into this. More information from the TIC report is available on the DDI site:

<http://www.ddialliance.org/sites/default/files/TechImplementReport2010.pdf>

**Qualitative Data Working Group.** This group is chaired by Louise Corti and Arofan Gregory and has met so far only by telephone. Qualitative data can mean different things and may refer to images, open-ended responses, and a variety of other data types. The group has created a detailed set of use cases. There will be a three-day meeting of the Qualitative Data Working group after the European DDI Users Meeting (EDDI) in December in Gothenburg.

**Web Site Maintenance Working Group.** This group is chaired by Sam Spencer, who provided a report in the form of an article for the Spring 2011 issue of the DDI Directions Newsletter. Sam has been evaluating site usage and has made some changes to drive more traffic to the site. More information from the report is available on the DDI site:

<http://www.ddialliance.org/sites/default/files/WebSiteMaintReport2010.pdf>

**Developers Community.** This group is not a traditional working group but there may be an advantage to its becoming one in terms of connections to the rest of the Alliance. There will be a three-day meeting of the developers after the European DDI Users Meeting (EDDI) in December in Gothenburg, parallel to the Qualitative Data meeting. More information from the report is available on the DDI site:

<http://www.ddialliance.org/sites/default/files/DDIDevelopersCommunityReport2010.pdf>

**New Working Groups.** New groups were proposed in these areas with charges to be formulated and communicated to all DDI Alliance members so that they can volunteer to participate:

- Paradata, to be led by Sue Ellen Hansen
- Administrative Data, to be led by David Schiller
- Disclosure Risk/Secure Data, currently an informal study group to look at what may be needed in DDI to cover this area. In the Research Data Centers there is a need to describe disclosure analyses around the release of information from the RDCs, but we should be looking at risk factors throughout the data life cycle. Right now this discussion is all about security but it should really be more about risk. We should also be documenting linkages across data files with administrative data, for example, and the consent that has been given to respondents. The Data Without Boundaries project is looking at these issues as are many others, and there appears to be a critical mass dealing with these topics so a working group may be a good thing.

## Update on DDI Codebook Version 2.5

The International Household Survey Network (IHSN) had requested that some new elements be added to the DDI 2.1 specification. This work has been completed and the IHSN has given approval. All changes to this development line must be backwards-compatible so that is one of the challenges. Enhancements

were also made in 2.5 to facilitate migration to DDI Lifecycle for those interested in moving to the other branch. It was pointed out that the Minnesota Population Center has created a lightweight tool to review and test this new specification. The schema should be ready this summer.

## **Update on DDI Lifecycle 3.2**

This update has been delayed because of a complex identification issue. There are also a variety of bugs that will be addressed in this version and the high-level documentation is being revamped as well. This version is likely to be released next winter.

## **Approach to Identification in DDI Lifecycle**

This is a complex issue and currently there is no agreement about the best way to handle it. Consequently, a procedure has been established to solicit outside input in order to arrive at a solution that will meet everyone's needs. Joachim Wackerow is writing up a summary of the issue and the two proposed approaches. This will be sent to the EC and to a set of experts who will be asked to comment. It is anticipated that this will be followed by a step to synthesize the comments and a Web meeting in which a final proposal is hammered out.

The identification issue is specific to the DDI Lifecycle branch of the specification and has to do with referencing metadata items. IDs should never change and we need to determine what information should be them, what is just related, etc.

## **DDI Agency Registry**

Algenta Technologies is developing an agency registry to uniquely identify each entity providing DDI metadata. There is an upfront one-time fee for this work and then an ongoing fee under \$1000 per year. Some business rules need to be clarified so there was a proposal put forward to get agreement around some of these rules. Specifically, approval was requested for these five points:

1. Agreement on the use of the country code of the primary headquarters for multi-national commercial entities as determined by the entity
2. Use of a single DDI specified top-level-domain for international non-commercial and intergovernmental organizations
3. Requirement of basic contact information which will be validated by email reply and verification of a site URN with some form of identifying information (content above "construction site")
4. No additional verification regarding validation of the registrant and their ability to speak for the agency in question; registration system will provide information on what is being declared by the act of registration but will not do further checking on the accuracy
5. Authorizing the DDI registration maintenance agency to verify this information and to deny registration to those not meeting these requirement, providing the reason for denial



A motion was made to accept this proposal and the motion carried. The next step will be a prototype of the service.

## **DDI 4.0**

The major change for this new version will be that a model will be created first from which all expressions, whether in XML, RDF, or something else, will be derived. We need to review the modularity of the standard and improve it to better parallel the research data life cycle. We may also want to look at data in a more abstract way so that it is not so discipline-specific. We will also incorporate new features like the information coming out of the SDI working group.

## **Update on CESSDA ERIC**

Hans Jørgen Marker updated the group on the status of the Council for European Social Science Data Archives (CESSDA) European Research Infrastructure Consortium (ERIC). CESSDA ERIC will be a new legal entity that will most likely replace the current CESSDA. It follows the completion of the EU CESSDA Preparatory Phase Project (PPP) undertaken in 2008-2009. At least three EU countries can apply to create an ERIC and the CESSDA project has five countries that have already signed on to be members. There are 14 special obligations that countries must fulfill, and the first is complying with the DDI- and OAIS-compliant portal. The CESSDA ERIC could be functional as soon as January 1, 2012. Bjorn Henrichsen of the NSD is the president. Norway and Germany have put up the bulk of the funding to establish CESSDA ERIC and this lowers the barriers for other countries to join. For those countries deciding to sign on as a national service provider, the social science data archive of the country will be the point of delivery of services.

## **Training**

A discussion of training in the context of the External Review report was the last item discussed. It was pointed out that training in the use of DDI started at IASSIST conferences and GESIS has for the past five years supported training at the Dagstuhl workshops held at Schloss Dagstuhl, Leibniz Institute for Informatics.

There is the view that DDI training and outreach should be coordinated and funded by the DDI Alliance. We need good training materials for different audiences and it will take time to develop a good set of materials. The report indicated that the community views the current group of instructors doing training as too small.

We may need a special Task Force to investigate training in all its dimensions. The recommendation was made to request that the Steering Committee issue a formal response to this and set up a task force to look into it. The EC wanted to communicate to the SC, which was to meet the following day, that this was an important area for the Alliance and this was a vote of emphasis.