Report of Moving Forward Project

Annual DDI Alliance Meeting, May 30, 2016

The work of the Moving Forward Project (MF) is done in working teams which meet on a regular basis on conference calls. Additionally several working meetings (sprints) took place as face-to-face meetings. These meetings had a dedicated focus on specific topics. Usually it is a mixed group of participants coming from different working teams.

Mature work products are forwarded to the Technical Committee (TC) for integration and possibly public reviews. Core members of the TC are involved in teams of MF. This assures good communication between the development work in MF and integration and maintenance in TC.

All work of the Moving Forward Project is documented on the wiki of the DDI Alliance. All information is publicly available to support the idea of transparency.

The model-based approach proved to be successful. Same patterns are used throughout the model in different packages. This assures model integrity by consistent usage of same approaches in various contexts. The feedback of the Q1 2015 Development Review of a DDI 4 package was very helpful in this regard. The comments of external experts in the Dagstuhl 2015 working meeting pointed to important design approaches regarding patterns and interplay with other specifications.

The Data Capture Team achieved an abstract model to provide home for various data collection sources like survey, register data, and data from devices. This is an important step forward that DDI is not only survey-centric.

The Data Description Team developed a flexible model which is able to describe data in various forms like single numbers (datum), case records, episode data, and aggregate data. The approach supports ideas from the big data field like data lakes. It is combined with the variable chain from the Generic Statistical Information Model (GSIM) (Conceptual Variable, Represented Variable, Instance Variable).

The Simple Codebook View Team mapped DDI Codebook elements to DDI 4 classes. The latest working meeting (sprint) in Norway focused on the DDI 4 Codebook View, which will have similar feature as DDI Codebook, but is a subset of DDI 4. The goal is to test the whole DDI 4 approach with this well-known subset of metadata and to support migration of DDI Codebook users to DDI 4.

The Methodology Team integrated the model from weighting (SDI group), worked on a generic approach, and is working on a generic process model for multiple usages in DDI 4.

The Production Process and Tools Support Team enhanced the production framework on a number of different levels. The production framework enables the web-based input of models, the transformation of the model to binding representations (like XML Schema and OWL/RDF-S), and the generation of item-level documentation. DDI 4 is a model-driven approach. Therefore the production framework - for generation of products (specification, documentation) driven by the model – plays a crucial role. Work

was done on migration of the process chain of multiple steps into a cloud-based, continuous integration server.

Working Teams

List of current teams:

- Active Data Management Plans Team
- Data Capture Team
- Data Description View Team
- Enhanced Citation Team
- Methodology Team
- Modelling Team
- Qualitative Data Team
- Simple Codebook View Team
- Study Inception / Data Management Plan Team
- Production Process and Tools Support Team

All teams have dedicated pages on the wiki. The team information includes the purpose of the group, list of members, and working and output documents. The information can be accessed by the list of current teams.

List of Working Meetings (Sprints)

All meetings had a length of five days. The longer face-to-face meetings enable a dynamic in the work which is not achievable in shorter meetings or conference calls. The advantage is that participants are very focused on the goals of the meeting, and the discussions on specific topics happen much more efficiently.

Each meeting has a dedicated wiki page where detailed information is available including the goals, the outcome, and the list of participants.

- Kalvåg, Norway, 23-27 May, 2016, 13 participants. Focus was on finalizing, documenting, and testing the Codebook Functional View.
- Edmonton, April 11-15, 2016, 9 participants. Preparing the model for the Q2 2016 Development Review. The goal of the sprint was to complete a consistency review of the DDI4 modeling.
- Copenhagen, November 23-27, 2015, 8 participants. Focus was on the technical issues of the production framework.
- Dagstuhl, Germany, October 19-23, 2015, 22 participants. The meeting brought together representatives from several metadata standards to provide an external review of current DDI work, with an emphasis on the model-driven approach, the production framework, and the substantive content of the standard.

What worked?

Many teams made great progress. The general approach with teams focusing on multiple content and structural topics worked well. Many people are involved in this work. Recruitment of participants was less an issue as in the past. Nevertheless it is important to involve new people in this process to have benefit from multiple perspectives and to share energies.

Q1 2015 Development Review worked well and supported development work. The Dagstuhl meeting with external experts assured that the current work is done as part of a larger landscape of metadata specifications and that MF work is using state-of-the-art techniques. See summary on wiki.

The model-based and model-driven approach is acknowledged as appropriate technique for a complex metadata specification. The use of patterns throughout the model enables consistency.

The above special mentioned teams achieved outstanding results. This work includes partly research and application of state-of-the-art techniques in modeling.

The approach is very successful to organize the work in topic-oriented teams (continuously meeting in conference calls) and in intensive face-to-face working meetings with mixed topics.

What didn't work? Or what did we learn?

Some people are in many groups. This is an issue regarding the workload for each individual and regarding having different perspectives at the same time. It is also in sprints a practical issue that the same person can't be in two parallel subgroups. On the other hand it supports very well the communication between groups.

It was possible to organize that some people had specific roles in sprints focusing on tasks like moderation and documentation. Nevertheless the sprints would be more successful if more preparation work would have been done prior to the meeting. The same applies for work after the meeting: follow-up regarding open issues, finalizing documents, framing the output for multiple audiences, organizing open issues for future work in working teams and possible other working meetings. A project manager could help a lot in this regard.

The working groups have sometime an ad-hoc approach, where it is not always clear in advance what the agenda topics are. This could be improved by preparation and announcement of agendas and reliable meeting times. Agendas could be put in advance of the meeting to the wiki. While the meeting topics can be checked off if resolved or could be forwarded to a next meeting or other group in a transparent way. This approach would also help participants who are not always attending and possible cross-team attendance of people. The cross-team participation could be fostered in order to get more perspectives. This would also help that each group has a broader understanding of the overall MF work. This was already partly done by inviting people from other groups to a conference call.

From a project management perspective it would be very helpful that open and resolved topics are shown on the wiki page of each group in a way that people from same (and other) group(s) have a chance to understand it.

What is needed in the next 12 months (e.g., resources)?

Resources for a project manager and for development of software tools for the production framework are needed.

A project manager would be very helpful to organize the complex MF work. There are many groups/people involved focusing on many areas and levels. Each group seems to make good progress. But the overall effort would be more efficient and the results better if the work of the teams is more coordinated, the sprints better prepared and followed-up. The resources (time of participants, travel money for sprints from member institutions and DDI Alliance) would be used more efficiently.

The requirement would be that such a person has project management skills and an understanding how is done/can be organized in a non-for-profit environment with volunteering contributors.

The production framework requires the development of several smaller software tools. As the production framework plays a crucial role, it is important that these tools are developed in a robust, transparent, and sustainable manner. This work shouldn't rely on specific persons. It would be required that dedicated, described tasks can be outsourced to paid programmers.

Support for several working meetings (sprints) is required. Planned meetings are so far:

- two different meetings (each 25 people) in Dagstuhl, October 2016
    - 1st week, review of current work with external experts, usage of DDI together with other specifications
    - 2nd week, structured documentation, controlled vocabularies, funding proposals
- a smaller technical meeting after EDDI in Cologne, December 2016
- a meeting in the margins of NADDI 2017 might make sense, Ithaca, New York, April 2017
- a meeting in the margins of IASSIST 2017 might make sense, Lawrence, Kansas, June 2017

How to improve coordination among groups, especially how to coordinate around the top priorities of the Alliance?

A project manager could help here a lot, see above.

MF is using a wiki (Confluence), an issue tracking system (JIRA), and version system (Bitbucket) for the development, documentation, and communication work. These tools have been already broadly adapted by the teams. Nevertheless a more efficient use of these tools could help to support the work better.

The cooperation systems provide a large amount of publicly available information. Nevertheless it is often not easy to find the right information, especially for people not heavily involved in the work. Overview pages could be created to find resources easier. Information should be available on different levels. The bridge from the DDI Alliance website to the wiki could be enhanced.

A project manager could help here in the consistent use of these systems.

Coordination between team chairs could be improved. This could be a task of the chairs of the Advisory Group of the MF and of the Executive director.

The regular use of the MF email list could improve for internal communication between all involved people. Team chairs could inform on major achievements.