DDI Alliance Annual Meeting of Member Representatives
Sydney, NSW, Australia -- UNSW Sydney
June 1, 2019

Minutes

Participants:

Tuomas Alaterä (Finnish Social Science Data Archive)
Iris Alfredsson (Swedish National Data Service)
Ingo Barkow (HTW Chur)
Cathy Fitch (Minnesota Population Center)
Jane Fry (Carleton University)
Jared Lyle (ICPSR)
Steve McEachern (Australian Data Archive)
Marianne Myhren (Norwegian Centre for Research Data)
Ron Nakao (Stanford University)
Barry Radler (MIDUS, University of Wisconsin--Madison)
Wendy Thomas (Minnesota Population Center)

Virtual Participants:

Jon Johnson (UCL - CLOSER)
Dan Smith (Colectica)

State of the Alliance

Steve McEachern, Chair of the Executive Board, opened the meeting by discussing core issues for the Alliance to tackle, including revenue generation, organizational structures, and community consultation. (See Appendix 2.) Steve noted that strategy development should be the outcome of community consultation.

Financial Report

Jared Lyle, Executive Director, presented the financial report for fiscal year 2019 (see Appendix 3.) The overall FY2019 expenditures are expected to exceed income by $12,284. The overall FY2019 deficit is very close to what was budgeted by the Executive Board in July 2018. The uncommitted fund balance (i.e., funds that are unencumbered by previously allocated expenses) at the end of FY2019 is anticipated to be $171,590.

Jared noted that the budget for FY2020 (July 2019 - June 2020) has not yet been set or finalized by the Executive Board. The plan is to use the 2019 Annual Meeting of Member

Representatives and the Scientific Board meeting to determine Alliance priorities, which will inform the decisions made by the Executive Board when they finalize the budget in June.

It was asked whether the Alliance has a threshold reserve balance it will not go under. Also asked was whether the Alliance has completed succession planning, especially holding funds to continue support of core services (e.g., DDI Registry). Both questions were directed to the Executive Board for consideration.

Working Group Reports

Marketing and Partnerships

Barry Radler, Chair of the Marketing and Partnerships Working Group, reported on the group's activities for the past year (see Appendix 4) by acknowledging promotional activities, including sponsorships and advertisements at IASSIST, AAPOR, ESRA, APDU, ESS, EDDI, and NADDI. He also noted regular activity on the Alliance's Twitter handle, as well as a 20 percent increase of users on the web site since last year.

Barry emphasized the importance of gathering feedback and input from Alliance members and the user community, especially with continued membership churn (i.e., new members joining and older members dropping membership). He proposed a formal assessment. Longer-term, Barry would also like to hire a dedicated marketing coordinator to help with web site design, membership maintenance, promoting conferences, etc.

It was noted that the DDI user base is expanding but it is difficult to measure. It was also noted that regarding development, the institutions integrating DDI will come in and out and that perhaps that's what the Alliance has to live with in terms of membership churn.

Training

Jane Fry, member of the Training Working Group, reported on the September 2018 Train-the-Trainer workshop held at Dagstuhl (see Appendix 5). She indicated the workshop was a success, with eighteen participants creating an initial draft of a training library. All participants of the workshop also promised to hold trainings of their own within the year, with several already hosting trainings.

Jane also reported on several other recent DDI trainings, including an introductory workshop she conducted at NADDI for 70 people. She emphasized the need and desire for DDI training.

It was discussed whether there is an appetite to charge for training, which would expand revenue to the Alliance. Webinars and introductory workshop intros could be free, with more substantive trainings for a fee. Other options could include fee-for-service and partnering with others on tools. It was suggested one value of Alliance membership could be discounts to events and trainings. Another suggestion was for the Alliance to appoint some of its members as official trainers.

Discussion was had about online training and Carpentry-style training. It was noted that while many people are offering online training, there is high value in bringing people together in one location for training.

Technical Committee
Wendy Thomas, Chair of the Technical Committee, summarized the activities of the year (see Appendix 6).

A sub-group of the TC had a face-to-face meeting for one week in Berlin in December. The focus of the discussion was the movement of production of DDI content to the COGS production platform.

The public review of DDI 3.3 took place with 41 comments received from 9 individuals representing 9 organizations. The issues have been reviewed and resolved by the Technical Committee and are in the process of being entered in the schema for publication.

The Technical Committee began to review the model of the DDI Prototype provided by the Modeling Team, which followed work done by the working groups. Forty-one issues were filed and 30 were returned to the Modeling group for resolution. Five were retained by the Technical Committee as the issues seemed broader than DDI4 and required additional community input for recommendations. Of these issues over half were filed by a single individual, overall 5 individuals responded to the review, 2 of which represented institutional responses.

XKOS issued identified in the 2016-17 public review have been resolved, with special thanks to Franck Cotton and Thomas Francart, and a public release is scheduled for mid-June.

The Technical Committee continued to work with the Controlled Vocabularies Working Group in identifying the format options that can be generated by the SKOS output of the CESSDA CV management system.

Wendy noted that during the first half of the upcoming year, the Technical Committee plans to focus on:

- Publishing DISCO
- Preparing DDI 3.3 for publication
- Resolving 5 DDI 4 Prototype review issues
- Reviewing issues filed for DDI-Codebook and preparing a new version
- Shifting DDI Lifecycle and DDI-Codebook production work to COGS
- Preparing for the shift of DDI 4 development work to COGS from Drupal

It was discussed how to better involve Alliance community members in working groups and other development activities. People want to get involved but there is not a good way of onboarding people -- e.g., finding out interests and where members could join and help. It was suggested the Alliance create a visible anchor so new users can hook into activities on their own rather than the currently perceived “customized and clubby” approach.

DDI 4
Wendy Thomas, Chair of the Technical Committee, reviewed the DDI 4 activities from the year (see Appendixes 7-10. She noted that the first half of the year was spent preparing for the DDI 4 prototype review, which was distributed in October.

The December Berlin sprint held in conjunction with the European DDI conference determined that all DDI 4 working groups but the modeling team had completed their initial tasks or were inactive. A proposal was put forth for a next generation modeling working group with Modeling, Representation, Testing Lifecycle (MRT) in mind, focused on the iterative lifecycle of modeling, representation, and testing. The proposal included a suggestion for using the core features of the DDI 4 model that are the most robust to date, conceptual, data description, and process, with a 'core' DDI 4 release that is implementable and the base on which to update the rest of the model. The goal is a DDI 4 Core release for review and publication in December 2019. The new MRT group held a sprint in the margins of the North American DDI conference in April.

The prototype public review was already noted in the Technical Committee's update, listed above.

Strategic Planning

Steve opened the discussion about strategic planning by discussing the Open Letter (see Appendix 11) sent to the Executive Board in October and the Executive Board's response to the Open Letter (see Appendix 12). Steve noted the disagreement within the Alliance, including within the Executive Board, about which strategic priorities to emphasize and fund. He noted that the Alliance has pockets of information but not a broad gathering of data to answer these questions. It was also noted that the draft 2018-2022 strategic plan (see Appendix 13) created in 2017 is not finalized.

Steve went on to suggest a path forward for determining the Alliance's strategic plan. For the next year, he suggested maintaining current activities, with a couple new activities: 1) community outreach through market research to identify and learn about member and stakeholder needs, and 2) updating the organizational structure of the Scientific Board so it is engaging and driving standards and work products of the Alliance. For FY2021, he suggests focusing on strategy development, revenue models, standards development and maintenance, organizational requirements, and community support and development.

For the community outreach, Steve indicated that market research can be augmented and informed by the last formal review of the Alliance -- the 2011 Breckenhill Report.

Several recommendations in the 2011 Breckenhill Report, he noted, are still valid, including:

- **Branding the standards**. In the 2011 Breckenhill report, community members expressed concern about branding for DDI 2 and DDI 3, including the danger of ignoring some lines over others. With DDI 4, we now have three separate lines. The continuation

of this issue speaks more to confusion about lines of the standard than naming conventions (e.g., DDI-Code book and DDI-Lifecycle).

- **More tools, please**! We continue to have a lack of useful actively supported tools for both DDI 2 and 3. The lack of tools directly impacts take-up.
- **Funding needs to improve dramatically**. The Alliance remains underfunded, especially when a cost value is attributed to 'in kind' contributions. While annual revenue of the Alliance has increased to $103,500 in FY2019 from between $64,900 and $82,450 from 2005 to 2010 inclusive, the demand for financial resources continues to outpace revenue. The conclusions of the report remain valid: "The current resource levels are, in our opinion, so low and the Alliance membership base so fragile, as to represent a serious risk to maintaining the ongoing viability of the organization."
- **Management by committee**! There is still the need to clarify and emphasize relationships between committees, especially to activate the Scientific Board and clarify the relationship of the Technical Committee to the Scientific Board.
- **More training and trainers are needed**. Recent train-the-trainer activities are promising. More are needed. The core group of trainers also needs to be expanded and refreshed, as this core group cannot continue to lead everything.
- **Marketing and communications**. The marketing group has been very active and successful in outreach. Still, there remains significant work to do, including developing partnerships and alliances, which was outlined in the 2014-2017 strategic plan.
- **Monitoring is part of good governance**. Two activities worth reviewing: (1) Evaluation of Alliance activities against strategic plan benchmarks, and (2) A technical evaluation of the DDI specification. "It may also be important to evaluate the DDI specification itself in preparation for Version 4.0. This should be a wide-ranging review with the aim of aligning the DDI with important scientific and technological developments and meeting the expectations of new and traditional stakeholders."

Discussion opened regarding what areas the Alliance should emphasize in the interim.

It was noted that training is important and that the more approachable documentation developed by the Alliance may help with training. Others emphasized the need for marketing and promotion.

It was noted that there has been a shift in defining the different DDI development lines. Initially, it was thought that people would move from Codebook to Lifecycle. But people did not want to shift since it would require significant infrastructural changes. How do we make sure we have a continuous, seamless development line? Codebook can be an entrée to Lifecycle. Codebook provides descriptive content. Lifecycle provides a metadata-driven approach.

It was suggested to add a way for users to file an issue or get questions answered directly on the DDI Alliance web site, including from: new users; one-offs who probably won't be part of the community; those interested in joining the community.

Next Annual Meeting

It was proposed that the next annual meeting of members be held 18 May 2020 in Sweden, the Monday of IASSIST.

It was also proposed to hold the next Scientific Board meeting at EDDI or NADDI.

Appendix 1

DDI Alliance Meeting
Saturday, June 1, 2019, 08:30-16:30
Hugh Dixson Theatre, UNSW Sydney
Gate 11, AGSM Building, Kensington Campus -- Directions

Virtual link

Agenda -- Meeting of Members				
Time	**Subject**	**Detail**	**Lead**	**Purpose**
08:30-09:00	Coffee			
09:00-09:05	Welcome		Steve	Introductions
09:05-09:20	State of the Alliance 2019		Steve	Update on last year's work
09:20-09:30	Alliance Budget	Financial Report	Jared	
09:30-10:00	Working Group Reports	-Marketing & Partnerships -Training -Technical Committee -DDI 4 Berlin sprint MRT group Prototype public review Ottawa sprint	Barry Jane Wendy	
10:00-10:15	Coffee break			
10:15-12:25	Strategic Planning	-Strategic Plan 2018-2022 (draft) -Open Letter: "The Case for Continued Support of a Model-Driven, Platform-Independent DDI" -Executive Board draft response to Open Letter	Steve	Get input and feedback
12:25-12:30	Proposed Date for Next Meeting		Steve	Agree on best day to meet
12:30-13:30	Lunch			

Agenda -- Meeting of Scientific Board				
Time	**Subject**	**Detail**	**Lead**	**Purpose**
13:30-13:35	Welcome		Ingo	Introductions
13:35-14:05	Scientific Board structure	-Improvements of Scientific Board structure -Decision-making -Postpone elections (chair, vice-chair) for one year -Discussion	Ingo Steve	
14:05-14:20	Training	-Training library -Training at conferences for new users -Discussion	Ingo	
14:20-15:05	Moving Forward program	-DDI 4 Core (see also scope and MRT group) -Technical Committee's DDI Specifications Roadmap -Future direction on DDI 4, and additional extensions views like data capture and codebook -Discussion	Steve Ingo Wendy	In-depth discussion of DDI4 development
15:05-15:20	Coffee break			
15:20-15:50	Scientific Board direction and goals for the year	-What are the goals? -What is the work plan? -Discussion	Ingo	Set goals for what to accomplish
15:50-16:20	Technical Committee	-What is the work plan? -Updating the Standards Development and Review Process and Procedures document -Discussion	Wendy	
16:20-16:30	SDTL proposal		Jared	

18:00 - Informal DDI group dinner

Appendix 2

DDI Alliance - State of the Alliance 2019

Dr. Steven McEachern

Chair, DDI Alliance Executive Board
and
Director, Australian Data Archive, Australian National University
steven.mceachern@anu.edu.au

Next steps

- Finalise Alliance budget and work program - subsequent to the acceptance of the Strategic Plan
- Need to identify for each part of the strategic plan:
  - resources (including money, time and in-kind contributions)
  - responsibilities (e.g. party/organization/team).
- Similarly, the resource constraints within the Alliance will by necessity limit the extent to which we can achieve the goals set out in this Plan.

Challenges in the 2018-19 cycle

- Strategy development has been challenging

- Community development
- Standards development, maintenance and support
- Organisational structures need to be implemented

- Core issues:
  - Revenue and resources more generally
  - Organisational structures
  - Community consultation
  - ---> Strategy development

DDI Strategic plan 2014-17

http://www.ddialliance.org/system/files/DDIAllianceStrategicPlan2014-2017.pdf

Three core work areas:

- Standards maintenance and development
- Expanding the DDI Community – Marketing and partnerships
- Restructuring to achieve our priorities

Standards maintenance and development

- Manage and maintain the two existing product lines (Codebook and Lifecycle)
- Review and vote on RDF Vocabularies
- Develop a next generation model-based DDI specification (2017)
- Continue to publish new Controlled Vocabularies
- Gain ISO certification (2017)

Expanding the DDI Community – Marketing and partnerships

- Build partnerships and strategic alliances (2017)
- Assess the current state of DDI usage, community needs, and resources (2017)
- Improve the DDI website
- Create new materials explaining the value of DDI to people who are not DDI specialists (2017)
- Build a community around DDI training and increase access through innovative mechanisms (2017)

Restructuring to achieve our priorities

- Review governance arrangements, including structure and Bylaws (2016)
- Review revenue and funding request models (2016)

Strategy development 2017-20

Strategic plan 2018-2022

https://docs.google.com/document/d/1IDqnsjI_IDcHPrXpHsZtt9D42eYXGqN0MP6ZaUq9NZo/edit?usp=sharing

The strategic plan is developed to focus on three broad priorities:

1. **Community and outreach**: how do we engage with the DDI community and understand the community's needs?
2. **Organisational needs**: what structures and systems does the Alliance need in order to meet those needs, and how will it maintain those structures and systems in the long term?
3. **Standards**: What products does the Alliance provide and maintain, and how do those products meet the needs of the Alliance and the broader community

Community and outreach

1. **Engagement with Global Digital Research Infrastructure**
2. **Solving Common Problems with Current DDI Users.**
3. **User group development program**

The DDI Alliance as an organisation

1. **Generational Renewal**
2. **Training: Enabling trainers to do what they need to do**
3. **Business Structure**

Standards and work products

1. **Maintaining multiple lines of specifications and controlled vocabularies**

2. **Improvement of interoperable and distributed DDI infrastructure for use and reuse of DDI resources.**

3. **Registries/repositories**

2019-20

Maintenance of current position, within budget review

- Standards development
- Standards maintenance and support
- Organisational structures (particularly Scientific Board)
- Community outreach - particularly a market research project on member and stakeholder needs

2020-21

- Strategy development
- Revenue models to support...
- Standards development and maintenance
- Organisational requirements
- Community support and development

Next steps

- Finalise Alliance budget and work program - subsequent to the acceptance of the Strategic Plan
- Need to identify for each part of the strategic plan:
  - resources (including money, time and in-kind contributions)
  - responsibilities (e.g. party/organization/team).
- Similarly, the resource constraints within the Alliance will by necessity limit the extent to which we can achieve the goals set out in this Plan.

Appendix 3

DDI Alliance Financial Report
Annual Meeting of Member Representatives
1 June 2019
Jared Lyle, Executive Director

Financial Report

This document summarizes the overall financial position of the DDI Alliance at the close of FY2019 (July 2018 - June 2019) and includes a draft budget for FY2020.

According to the Alliance bylaws:

“The Executive Board sets the overall budget….[and] shall establish a budget that provides financial support for the successful operation of the Alliance that may include support for some portion of the time of the Executive Director, Alliance duties and functions as determined by the Executive Director and the Secretariat, expert consultation, meetings, training, and funds for innovation and testing.”

Overview of FY2019 Budget

Appendix A provides an overview of the FY2012-FY2019 actual budgets. It also provides three views of the FY2019 budget: the Budget FY2019 column lists what was budgeted for the Alliance at the start of FY2019, the Actual FY2019 column lists all expenses that have been processed by the Alliance fiscal year-to-date, and the Forecast column lists actual expenses plus expected expenses through the remainder of the fiscal year.

Below are details for each of the main FY2019 budget categories, as well as a summary of the FY2019 budget.

Revenue

Membership fees

- Membership fees are based on organization size and membership benefits. The basic membership fee for OECD countries is $3,000 USD.

- The expected income for FY2019 was $108,500, but actual income is $100,500, with forecasted income at $103,500. The negative variance is due to dropped memberships by the Australian Bureau of Statistics and the Roper Center for Public Opinion Research, as well as non-payment by one full member.

Expenditures

Staff Salaries & Data Processing

- Staff salaries and data processing expenses are expected to match what was budgeted. Staff salaries cover secretariat staff at ICPSR, including 10 percent of the Executive Director, 5 percent of an accountant, and 6 percent of a Web developer. Data processing funds computing services of secretariat staff. Salary and data expenses have remained the same for the last 5 years even though ICPSR salaries have increased by an average of three percent each year.

Consultants

- Two consultants were employed this fiscal year. An RDF consultant was employed to develop concepts, select appropriate techniques, and give general advice regarding questions in the area of Semantic Web technologies. A project manager was employed to coordinate DDI 4 work and to bring the effort to a consistent result.

General Expenses

- The bulk of general expenses are credit card and wire fees. Expenses were lower than anticipated since DDI URN registration at IANA work has not been completed. $4,000 was budgeted for this work.

Research Supplies & Services

- Expenses covered GoToMeeting subscription fees and Google Cloud Services expenses, where DDI production workflows are hosted. Other expenses included hosting fees and planned enhancement expenses for the DDI Registry, a hosted DNS SRV record-based resolution service for DDI agency identifiers..

Marketing

- Marketing is projected to spend about a third of its budget allocation ($6,212 of $17,000 budgeted) on three conference sponsorships, as well as on materials, printing, and shipping.

Training

- Training spent $10,251.70 on the fall Dagstuhl Train-the-Trainer workshop, which is slightly lower than the $10,800 originally budgeted.

Travel and Hosting

- Travel funds covered one fall Dagstuhl workshop on metadata interoperability, as well as two sprints (one at EDDI in December and one at NADDI in April). Travel funds also supported travel for Mari Kleemola to represent the DDI Alliance in a UNECE ModernStats World Workshop in November in Geneva.

Summary

- The overall FY2019 expenditures are expected to exceed income by $12,284.
- The overall FY2019 deficit is very close to what was budgeted by the Executive Board in July 2018 (-$9,572).
- The fund balance for the Alliance is expected to be $181,575 at the end of FY2019. Of this, $9,985 is committed to North American DDI (NADDI) reserves. The uncommitted fund balance (i.e., funds that are unencumbered by previously allocated expenses) at the end of FY2019 is anticipated to be $171,590.

Overview of FY2020 Budget

The budget for FY2020 (July 2019 - June 2020) has not yet been set or finalized by the Executive Board. The plan is to use the 2019 Annual Meeting of Member Representatives and Scientific Board meeting to determine Alliance priorities, which will inform the decisions made by the Executive Board when they finalize the budget in June.

The April 2019 Executive Board meeting minutes.provides an overview of the FY2020 draft budget, which includes all committed expenses (i.e., expenses needed to continue daily operations of the Alliance, such as funding secretariat staffing, or new short-term expenses already approved to by the Executive Board), as well as proposed new expenses. Summaries for each of the main draft FY2020 budget categories are contained in the minutes.

The FY2020 draft budget was compiled with input from the Executive Board, as well as from direct feedback from the leads of the Scientific Board, the Technical Committee, the Marketing & Partnerships working group, and the Training working group.

Summary

- The budget for FY2020 (July 2019 - June 2020) has not yet been set or finalized by the Executive Board.
- If all budget requests were to be approved, total expenses would exceed revenue by approximately $62,112.
- The Executive Board will use the 2019 Annual Meeting of Member Representatives and Scientific Board meeting to determine Alliance priorities, which will inform the decisions made by the Executive Board when they finalize the budget in June.

Appendix 4

DDI Marketing and Partnerships Group

Report and Plan 2019

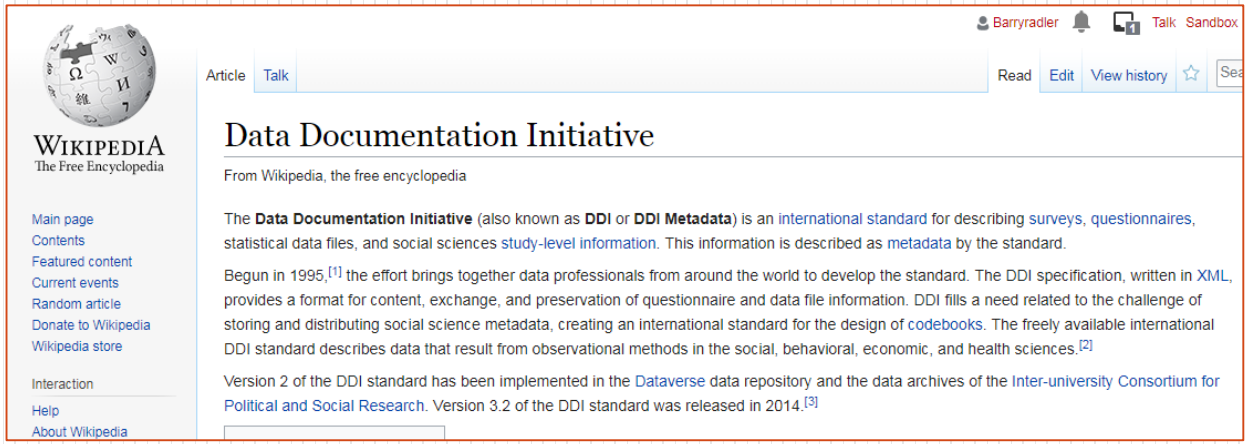
Team:
Barry Radler
Kelly Chatain
Jared Lyle
Steve McEachern
Ron Nakao
Dan Smith
Wendy Thomas

Mission Statement

- Coordinate marketing activities, establish DDI brand, ensure consistent messaging
- Interface with other standards bodies (Partnerships)
- Goal: Increase the DDI user community and DDI Alliance membership

Activities during the past year

- Promotional activities
  - Relatively quiet year of marketing activity
  - Maintenance of ongoing efforts
    - Wiki - Jared
    - Twitter
  - Monitoring website with Google Analytics



Audience Overview

SAVE EXPORT SHARE INSIGHTS

All Users
+0.00% Users

+ Add Segment

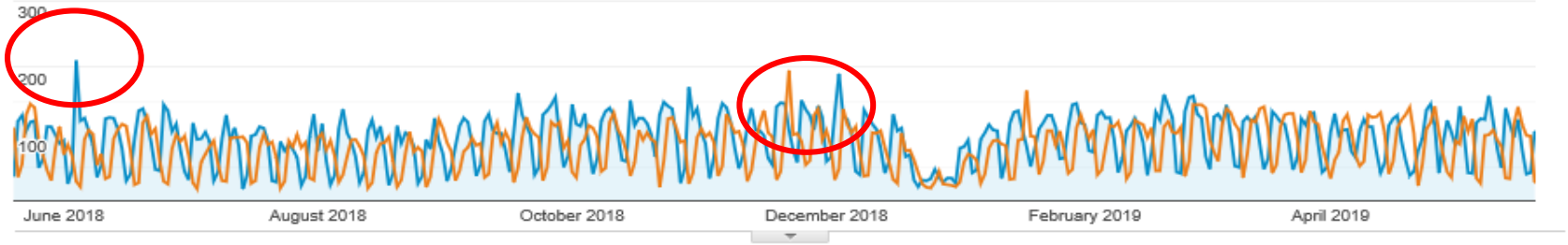
May 20, 2018 - May 20, 2019
Compare to: May 19, 2017 - May 19, 2018

Overview

Users vs. Select a metric

Hourly Day Week Month

May 20, 2018 - May 20, 2019: Users
May 19, 2017 - May 19, 2018: Users



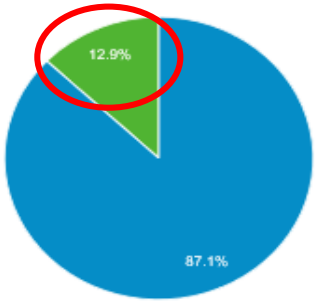
Users	New Users	Sessions
20.22%	18.78%	13.42%
27,833 vs 23,152	27,465 vs 23,122	38,348 vs 33,811

Number of Sessions per User	Pageviews	Pages / Session
-5.66%	8.13%	-4.66%
1.38 vs 1.46	90,725 vs 83,900	2.37 vs 2.48

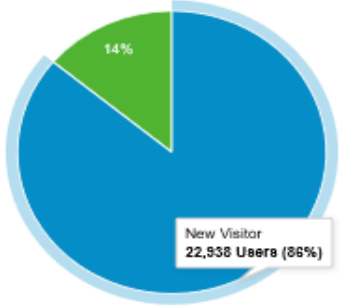
Avg. Session Duration	Bounce Rate
-6.85%	2.88%
00:02:20 vs 00:02:30	57.46% vs 55.85%

New Visitor Returning Visitor

May 20, 2018 - May 20, 2019



May 19, 2017 - May 19, 2018



Audience by country

Country	Users	% Users
1. United States		
May 20, 2018 - May 20, 2019	8,890	31.90%
May 19, 2017 - May 19, 2018	7,130	30.47%
% Change	24.68%	4.70%
2. Canada		
May 20, 2018 - May 20, 2019	1,548	5.56%
May 19, 2017 - May 19, 2018	1,182	5.05%
% Change	30.96%	9.97%
3. Germany		
May 20, 2018 - May 20, 2019	1,504	5.40%
May 19, 2017 - May 19, 2018	1,427	6.10%
% Change	5.40%	-11.50%
4. United Kingdom		
May 20, 2018 - May 20, 2019	1,370	4.92%
May 19, 2017 - May 19, 2018	1,243	5.31%
% Change	10.22%	-7.45%
5. India		
May 20, 2018 - May 20, 2019	946	3.39%
May 19, 2017 - May 19, 2018	780	3.33%
% Change	21.28%	1.84%
6. France		
May 20, 2018 - May 20, 2019	853	3.06%
May 19, 2017 - May 19, 2018	479	2.05%
% Change	78.08%	49.53%
7. Australia		
May 20, 2018 - May 20, 2019	810	2.91%
May 19, 2017 - May 19, 2018	786	3.36%
% Change	3.05%	-13.47%
8. Netherlands		
May 20, 2018 - May 20, 2019	748	2.68%
May 19, 2017 - May 19, 2018	604	2.58%
% Change	23.84%	3.99%
9. China		
May 20, 2018 - May 20, 2019	620	2.22%
May 19, 2017 - May 19, 2018	919	3.93%
% Change	-32.54%	-43.35%
10. Indonesia		
May 20, 2018 - May 20, 2019	525	1.88%
May 19, 2017 - May 19, 2018	220	0.94%
% Change	138.64%	100.38%

Behavior by pages viewed

	Page	Pageviews	% Pageviews
1.	/		
	May 20, 2018 - May 20, 2019	21,260	23.43%
	May 19, 2017 - May 19, 2018	19,207	22.89%
	% Change	10.69%	2.36%
2.	/explore-documentation		
	May 20, 2018 - May 20, 2019	6,340	6.99%
	May 19, 2017 - May 19, 2018	6,315	7.53%
	% Change	0.40%	-7.16%
3.	/resources/tools		
	May 20, 2018 - May 20, 2019	4,985	5.49%
	May 19, 2017 - May 19, 2018	4,677	5.57%
	% Change	6.59%	-1.43%
4.	/training/getting-started-new-content/create-a-codebook		
	May 20, 2018 - May 20, 2019	3,560	3.92%
	May 19, 2017 - May 19, 2018	3,396	4.05%
	% Change	4.83%	-3.06%
5.	/controlled-vocabularies		
	May 20, 2018 - May 20, 2019	3,376	3.72%
	May 19, 2017 - May 19, 2018	3,149	3.75%
	% Change	7.21%	-0.86%
6.	/training/why-use-ddi		
	May 20, 2018 - May 20, 2019	3,033	3.34%
	May 19, 2017 - May 19, 2018	3,122	3.72%
	% Change	-2.85%	-10.16%
7.	/training/getting-started		
	May 20, 2018 - May 20, 2019	1,916	2.11%
	May 19, 2017 - May 19, 2018	1,787	2.13%
	% Change	7.22%	-0.85%
8.	/about/about-the-alliance		
	May 20, 2018 - May 20, 2019	1,549	1.71%
	May 19, 2017 - May 19, 2018	1,630	1.94%
	% Change	-4.97%	-12.12%
9.	/resources/markup-examples		
	May 20, 2018 - May 20, 2019	1,481	1.63%
	May 19, 2017 - May 19, 2018	1,489	1.77%
	% Change	-0.54%	-8.02%
10.	/training		
	May 20, 2018 - May 20, 2019	1,419	1.56%
	May 19, 2017 - May 19, 2018	1,296	1.54%
	% Change	9.49%	1.25%

What worked during the past year?

- Promotional activities
- Conference attendance
  - Ongoing use of promotional printed and electronic materials, conference schwag
  - Continued promotion to communities/conferences
    - Sponsorships, exhibitions, and ads at IASSIST, AAPOR, ESRA, APDU, ESS, EDDI, NADDI
    - Branch out: RDA, others?

What worked during the past year?

- Promotional activities
- Conference attendance
- DDI Alliance membership treading water

Plans for next 12 months

- Promote
  - Continue/expand conference presence and attendance
- Gather feedback and input from Alliance members and user community
  - Proposed formal assessment
  - Requires money, effort, and time

Longer term DDI Marketing goals

- Dedicated DDI Alliance position (50-100%):
  - Website design and maintenance
    - Updating Wiki, social media
  - Membership maintenance
  - Press releases, promoting user conferences
  - Conference attendance and active outreach
  - Revenue generation, grant writing

Appendix 5

Train-the-Trainer Workshop
DDI Training Library

Report

Joachim Wackerow, 2019-05-23

Train-the-Trainer Workshop

A 5-day workshop to increase training capacity on DDI was hosted at the internationally renowned computer science institute at Schloss Dagstuhl in Germany, Sep 24-28 2018, sponsored by GESIS – Leibniz Institute for the Social Sciences and organized by Joachim Wackerow. The workshop instructors were Jon Johnson, Dan Smith, and Wendy Thomas, with Arofan Gregory as a volunteer. The DDI Alliance paid the GESIS workshop fees for all participants and provided travel scholarships for two participants.

18 people from 16 organizations and 11 countries participated in this workshop. The sex ratio was equal. The list of countries included Australia, Canada, Denmark, Estonia, France, Germany, Malawi, Norway, Romania, Sweden, UK, and USA. Two cancellations were received of registered people.

All participants plan a tutorial in the next year, either in their organizations or a public one. Therefore the DDI Alliance paid the GESIS workshop fee for them.

Two people received travel funding from the DDI Alliance, one from Romania, one from Malawi/London. The travel support program achieved the original goal to include people from countries which can't often provide sufficient travel support, here Eastern Europe and Africa.

Participants were very motivated and worked often in working groups on specific topics. The resulting slide decks cover core areas of DDI. They are as far as possible version-agnostic. This is on the one hand an important feature. On the other hand it raises the question how version-specific details (mostly technical) can be addressed. This latter remains an open issue.

The workshop output is an important step to achieve a modular training material body which can be used for different target audiences. Nevertheless, the material is not in a form that it can be immediately published on the DDI Alliance website. Further work will be required to put the material into a consistent form in terms of content, clear interfaces between modules, and layout.

For now, a subset (4-5 people) of the larger group of participants has volunteered to continue developing the training materials.

Overall 35 applications to the workshop were received, many from African countries. Most people who could not be accepted had either not enough knowledge or no travel funding or both.

DDI Training Library – post-workshop activity

A group of volunteers from the workshop, assisted by some other interested members of the DDI community, met several times and organized the finalization of an initial draft of a training library on the basis of the work coming out of Dagstuhl. Much of the focus was on creating consistency across all of the draft slides emerging from the workshop, but some gaps were identified, and some work re-organized. The work is on-going, but is not anticipated to continue after the initial draft of the training library is complete.

Like the workshop, most effort has been focused on conceptual approaches to DDI which are not specific to a particular version of the standard. Many trainings would require that version-specific slides, many of which already exist, would need to be incorporated with those from the library. The design and licensing of the library are such that the slides can be taken and re-used freely, and modified to suit the needs of the target audience.

While it is hoped that this draft library will continue into the future, and some discussion has taken place as to how it could best be organized and maintained as an on-going effort, the members of this group are not currently committed to making this happen. Additionally, the presentation of the library on the DDI site has been discussed, but no specific work is currently focused on these delivery aspects of the library. The group expects to continue work after IASSIST until the initial draft is complete.

Tutorials organized by participants of the Dagstuhl 2018 workshop

(reported up to May 2019)

- Olof Olsson: Short tutorial for 80 librarians, archivists and IT-personnel from ~20 different Swedish universities, October 2018
- Kaia Kulla: Three tutorials in Estonian at Statistics Estonia, November and December 2018
- Knut Wenzig: Half-day tutorial on DDI and panel data at EDDI18, December 2018
- Guillaume Duffes: 2-day tutorial in French at Insee (Statistics France), January 2019
- Alina Danciu, Alexandre Mairot: Tutorial in French for community of French engineers in social sciences (MATE-SHS), March 2019
- Anja Perry, Jane Fry: Half-day tutorial at IASSIST conference, May 2019
- Sanda Ionescu: Tutorial at ICPSR, August 2019
- Hayley Mills: Half-day tutorial for UK longitudinal study Data Managers, September 2019
- Hilde Orten: six 90-minutes tutorial with different topics at NSD, autumn 2019

Appendix 6

Technical Committee Report for 2018/2019

Submitted by: Wendy Thomas, Chair on behalf of the Technical Committee

During the 2018/2019 period the Technical Committee addressed the full range of DDI products:

- The public review of DDI 3.3 took place from June-October 2019 with 41 comments received from 9 individuals representing 9 organizations. The issues have been reviewed and resolved by the Technical Committee and are in the process of being entered in the schema for publication. Documentation content has been expanded and will be published in multiple formats. The generation of documentation by the COGS system was also evaluated during the public review and changes made to deal with any bugs.
- A sub-group of the TC had a face-to-face meeting for one week in Berlin 218. The focus of the discussion was the movement of production of DDI content to the COGS production platform. DDI-L work focused on the preparation of the COGS tool for support of both the DDI-L and DDI4 production work. Input flows, validation options, outputs, and post-output production pipelines were considered. Note that publication of DDI-L 3.3 is not dependent upon the completion of this work.

  DocFlex has been completed (with the exception of the final high level document information) and it is now being produced from the current schema structure for field level documentation and we are able to produce high level documentation through the current configuration of COGS.

  Requirements for DDI-L COGS:
  - Support of currently used XSD features
  - Validate that all of 3.3 is imported correctly
  - Ability to output DDI-L 3.3 in current XSD with supporting documentation
  - Verify that the UML definitions are accurate and stable
  - Output of XMI from CSV files

  The overall production flow is still to be decided, through COGS output options or following the DDI4 generation from the XMI.

  Production validation work has been specified, charted and begun for DDI-L in COGS. Some points regarding DDI4 in COGS have been added and work has been completed on a test transformation for the canonical XMI to CSV which should allow for identification of transformation issues going into and coming out of COGS for this product. Several other steps have been completed or are in progress.

  https://ddi-alliance.atlassian.net/wiki/spaces/DDI4/pages/620691457/Berlin+EDDI+Sprint+-+2018?preview=/620691457/637042704/EDDI%202018%20Sprint%20Final%20Report.pdf
- XKOS completed a public review at the end of 2016-2017. Thanks to Franck Cotton and Thomas Francart the issues were resolved and the vocabulary and documentation updated. XKOS is scheduled for release in mid-June.
- The Technical Committee continued to work with the Controlled Vocabularies Working Group in identifying the format options that can be generated by the SKOS output of the CESSDA CV management system. Initial output will be available in HTML to support the web page, PDF, and SKOS. Providing an additional binding in a CodeList (DDI 3.2) format will be added this year. Initial feedback from the community on the use of Genericode format has been minimal. This format will

be continued if there is still a need for it. Improved documentation of the availability of DDI published Controlled Vocabularies in the schemas has been implemented.

- In February 2018 the Technical Committee began to review the model of the DDI Prototype provided by the Modeling Team which followed work done by the working groups through 10 December 2017. The review work was completed and the DDI4 Prototype was put out for public review in October 2018 with comments from the Technical Committee. https://ddi-alliance.atlassian.net/wiki/spaces/DDI4/pages/589824023/DDI+Prototype+TC+Review+comments Review ended March 2019. 41 issues were filed and 30 were returned to the Modeling group for resolution. 5 were retained by the Technical Committee as the issues seemed broader than DDI4 and required additional community input for recommendations. Of these issues over half were filed by a single individual, overall 5 individuals responded to the review, 2 of which represented institutional responses.

During the first half of 2019/2020 TC plan to focus on:

- Publishing DISCO
- Preparing DDI 3.3 for publication
- Resolving 5 DDI 4 Prototype review issues
- Reviewing issues filed for DDI-Codebook and preparing a new version
- Shifting DDI Lifecycle and DDI-Codebook production work to COGS
- Preparing for the shift of DDI 4 development work to COGS from Drupal

Appendix 7

EDDI Sprint Final Report

December 5, 2018
(updated links on January 14, 2019)

Joint Meeting of Technical Committee and DDI 4 Development Group, November 26-30, 2018

DDI Lifecycle Work

DDI-L work focused on the preparation of the COGS tool for support of both the DDI-L and DDI4 production work. Input flows, validation options, outputs, and post-output production pipelines were considered. Note that publication of DDI-L 3.3 is not dependent upon the completion of this work. DocFlex has been completed (with the exception of the final high level document information) and it is now being produced from the current schema structure for field level documentation and we are able to produce high level documentation through the current configuration of COGS.

Requirements for DDI-L COGS:
- Support of currently used XSD features
- Validate that all of 3.3 is imported correctly
- Ability to output DDI-L 3.3 in current XSD with supporting documentation
- Verify that the UML definitions are accurate and stable
- Output of XMI from CSV files

The overall production flow is still to be decided, through COGS output options or following the DDI4 generation from the XMI.

Production validation work has been specified, charted and begun for DDI-L in COGS. Some points regarding DDI4 in COGS have been added and work has been completed on a test transformation

for the canonical XMI to CSV which should allow for identification of transformation issues going into and coming out of COGS for this product. Several other steps have been completed or are in progress.

Related Documents

Current XSD Objects Used

GANTT Chart for COGS

DDI 3.3 Production Validation

DDI 4 Development Work

The DDI 4 work fell into two general areas: Production framework/work cycle and modeling.

Production Framework

Overview

An iterative approach for moving the DDI 4 development work forward was proposed in the form of the Modeling Representation Testing Lifecycle (MRT). The goal of this approach is to provide a ‘mostly’ automatic five stage development process from modeling through implementation (software testing) on a platform that is sustainable at the technical level and that can provide feedback input on a stage-by-stage basis. It would also allow for various UML modeling tools to be employed at the user’s discretion.

Figure 1. Modeling Representation Testing LifeCycle (MRT)

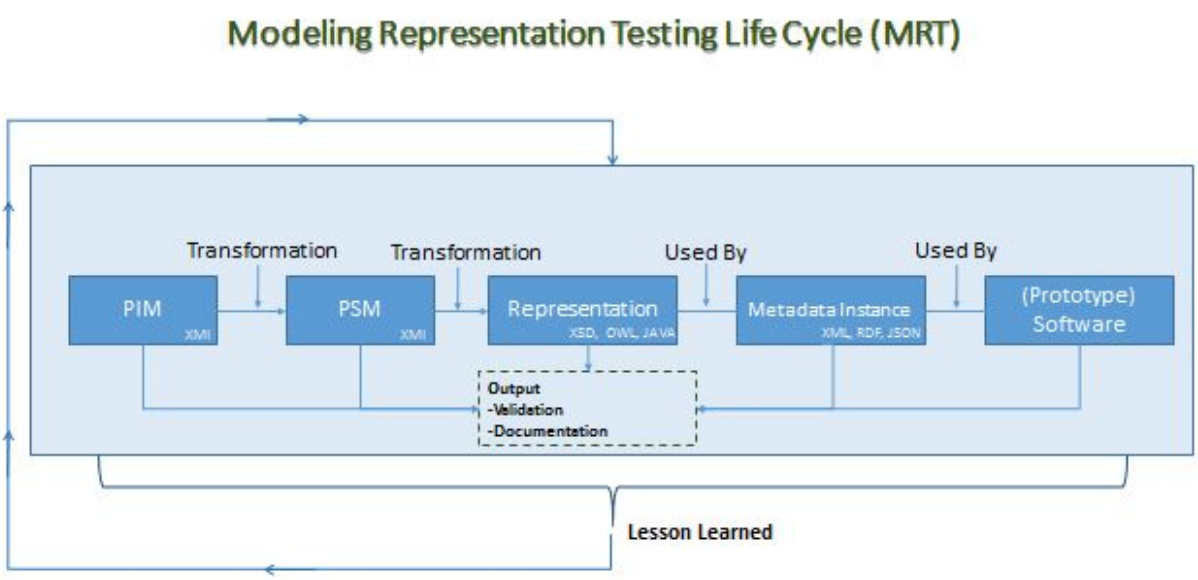


Figure 2. MRT Feedback Loop

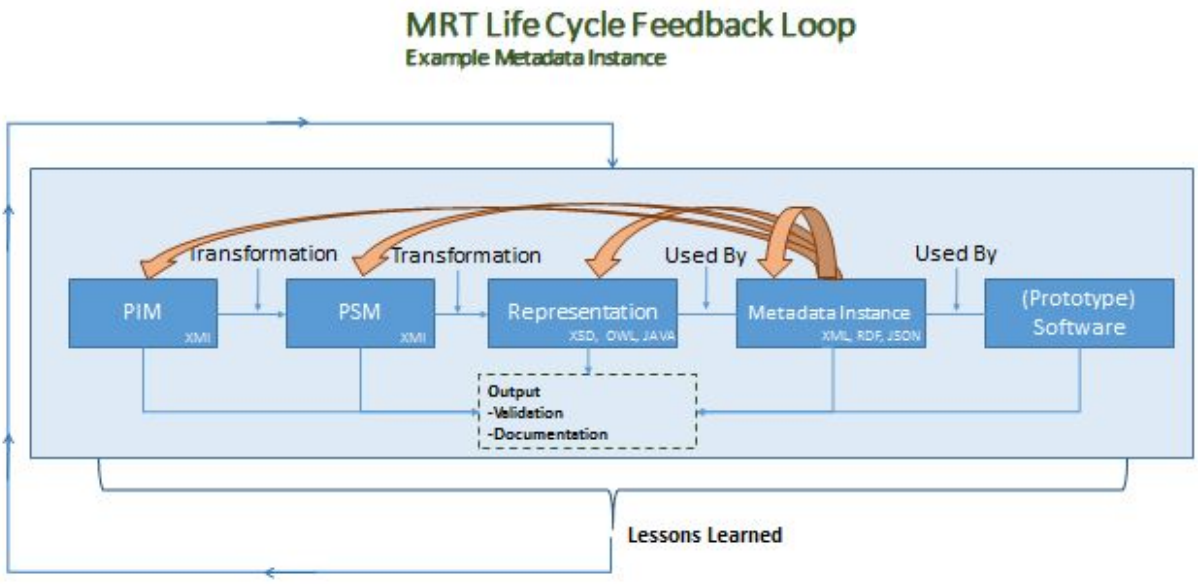
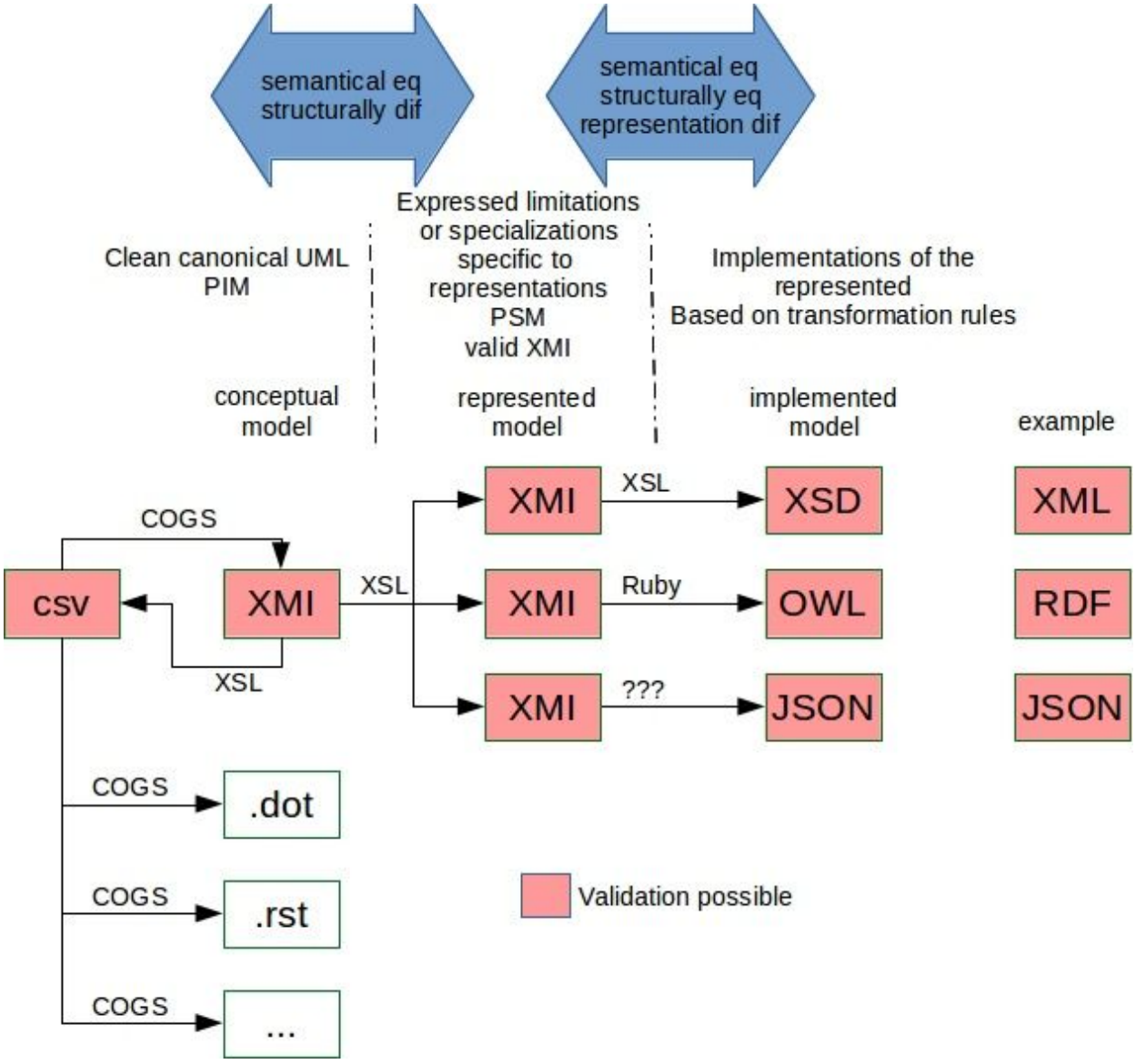


Figure 3. Production Workflow with Outputs and Validation



Requirements Breakdown

A list of requirements that covers function, output, and best practices was produced. Each item is associated with one or more stage(s) of the MRT and includes a note about whether it exists in the current production framework and in COGS. Comments on each requirement are also included.

DDI Production Framework Requirements Breakdown (link)

Preliminary work on a set of validation rules was begun, broken down into UML-specific rules, and a set of rules for Platform Specific Model (PSM), RDF, and XML. These documents are in the very early stages of work and are based on the validation documents produced during the Copenhagen Sprint in 2015.

UML Validation
PSM Validation
XML Validation
RDF Validation

General Discussion

The following points were made in the discussion of the MRT and workflow in general which require more input for resolution:

1. At the prototype software stage, the implementation, it is important to reach out to users who are actually doing this type of implementation. Converting from instance to instance is good, but not the entire perspective. This stage should also be used to import existing codebook and lifecycle instances for testing. Also need to test the 'uncommon' problems in real life, the edge cases, not just the mainstream.
2. There is a question about the canonical 'source of truth' in this pipeline. Should it be the csv or the XMI (PIM stage). In COGS the csv and the XMI are two guaranteed representations of the model, but can you have two 'sources of truth'?
3. Version control in the process needs clarification
4. Resources are required in order to bring this proposal to fruition.

Modeling

Overview

JIRA contains a number of issues pertaining to modeling which are labeled as post-prototype (known prior to the release of the prototype but not yet resolved) and prototype-review (filed as part of the prototype review process). To gain a broader understanding of the issues, they were grouped into categories and then linked to a 'master' issue for each category. A report of each other these categories contains a description of the requirements or a statement of the problem(s), a discussion of proposed solutions with questions and limitations in each area, and links to the issues in JIRA that are related to that area. The report as it stands can be found here in its entirety. The separate areas will be broken out and placed into the corresponding master issue in JIRA.

UML To Be Decided - Consolidated

Issues relating to data types, cardinality types, class content, class independence, and class membership, were broken down in a supporting document that covers the general purpose and importance to the model (UML) and representations (XML, RDF) along with comments.

Object Types Overview

Issues relating to Views and the Library - supporting documentation

Library and Functional Subsetting

MRT Group Proposal

A proposal was put forth for a next generation modeling working group with the MRT in mind, focused on the iterative lifecycle of modeling, representation, and testing. The proposal includes a suggestion for using the core features of the DDI 4 model that are the most robust to date, conceptual, data description, and process, as the focus of this approach for a period of one year. The result being a 'core' DDI 4 release that is implementable and the base on which to update the rest of the model.

MRT Group (Modeling, Representation, Testing Lifecycle)

General Discussion

A number of topics require additional input and follow-up:

1. While addressing the modeling issues, it was agreed that it is necessary to revisit the goals of DDI 4 and create a list of requirements, both business and technical, to be evaluated and approved by the Scientific Board for moving forward. A distinction can be made between the more 'technical' requirements, say of UML features that are business-independent that could serve, for instance, the goals of finding the best way to work with other standards, and the more business-specific requirements that must include feedback and input from the DDI community.
2. The consolidated report addresses the JIRA issues and the direct input from the sprint participants, but additional input can and should be added to fill in any gaps. Each area/category should be assigned to a modeling group member.
3. The MRT proposal requires additional content as well as a consensus on the 'core features' for moving forward:
    - Include a statement about how MRT would interact with the 'content' creators and implementers, (not about new content, but focusing on making the core more robust).
    - Specific timeframe for 1 year to make the prototype 'fly'

- How and where the work is done - JIRA? Google Drive documents? What platform?
- Linking issues between BitBucket and JIRA or other higher level goals
- Resolving the pipeline issues with existing and accessible resources
- Outreach to tool implementers (developers)

Conference Call Notes Supporting Discussion:

EDDI Remote Call - Tues - November 27th
EDDI Remote Call - Wed -November 28th
EDDI Remote Call - Friday - November 30th

Packaging the Next Steps: Presenting to Scientific Board/ Funding from Executive Board

1. What's in it?
    a. Goals - Audience and Purpose for DDI 4 (w/reference to 3 and 2)
        i. What are the inputs? TC Review of 3, etc. Group discussions of 4, feedback from DDI User community…
        ii. Capture the discussion to date for consideration of moving forward and buy-in

    b. Modeling Requirements - Business and Technical (input from various sources)
        i. UML To Be Decided - Consolidated (prototype review and post-prototype)
        ii. Flavio (StatCan and other statistical agencies)
        iii. Jay and Arofan and the ALPHA project
        iv. Cross-Domain DDI 4 Dagstuhl 2018 Report
        v. UKDA implementation of Data Description (Jon and Darren Bell)
        vi. Other additions? Future ideas…

    c. Production Framework Internal requirements for sustainability of DDI products

2. Where is it?
    a. Which platform to use for generating and managing the documents?
    b. Mechanism for delivery and review

3. When is it?
    a. Must be ready in April 2019 for the Scientific Board to review prior to the DDI Annual Meeting at IASSIST.

Appendix 8

MRT - Modeling, Representation, and Testing Lifecycle: A Proposed Working Group for Building DDI 4 Core

Draft 1.0, 28 January 2019

(Contributors to this document include: Guillaume Duffes, Dan Gillman, Jay Greenfield, Arofan Gregory, Oliver Hopt, Larry Hoyle, Jon Johnson, Hilde Orten, Flavio Rizzolo, Wendy Thomas, and Achim Wackerow)

Overview

This document proposes a re-organization of the DDI 4 Modeling Team to incorporate suggestions coming from various quarters that a shorter development and testing cycle be adopted, and that the scope of the overall modeling effort become more focused. The structure and focus of each of a number of sub-teams is outlined, to be coordinated by an overall working group.

In order to guarantee the practical utility of the model, people and projects – both internal and external to the DDI Community - will be enlisted to act as testers for the draft models as appropriate. This re-organization provides for a liaison function so that the findings of groups testing and implementing drafts of the model are tracked and incorporated in a systematic fashion. It will be the function of the MRT Working Group to manage these relationships.

The scope of the effort is also defined, as are efforts to identify and formalize the business and technical requirements to which the group is building. The timelines are intended to be short – rapid development cycles will help to produce useful standards output in the shortest possible timeframes. This document addresses work over the course of the calendar year 2019.

MRT Working Group

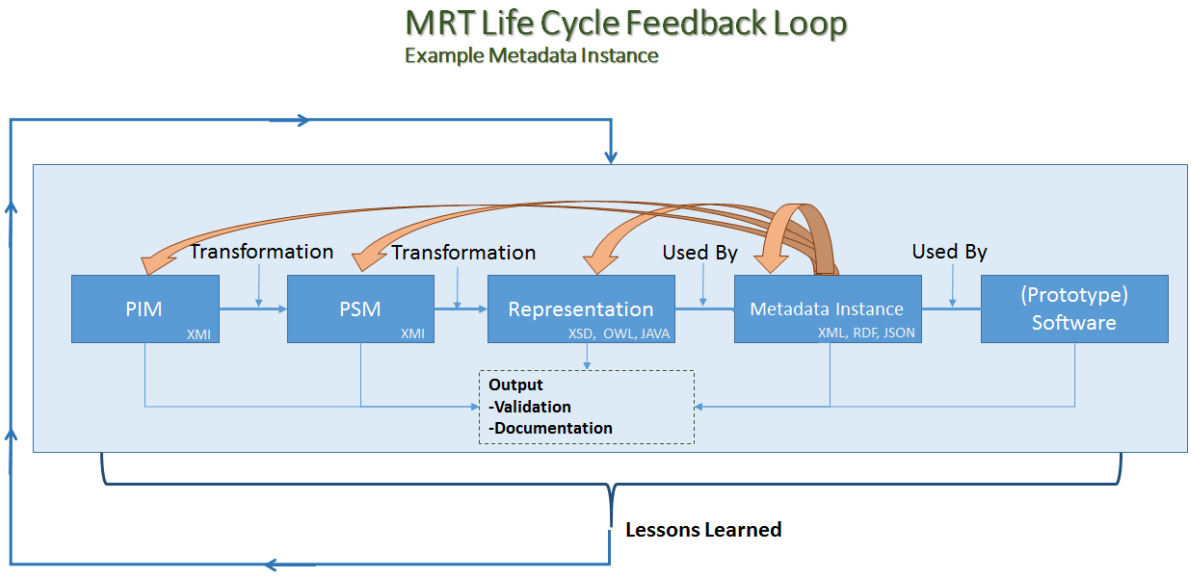
The MRT working group focuses on the iterative lifecycle of modeling, representation, and testing. This group replaces the previous modeling group. The new group should have experts on each stage of the MRT lifecycle, i.e. modeling, representations, testing of metadata instances, and software for transformation and testing. These experts can be involved in weekly phone calls on an on-demand basis: not all participants need to be involved in every task of the group.

The intention of combining modeling, representation, and testing in an iterative lifecycle is to achieve more efficient production of the specification. Any major change on the model level can be immediately evaluated on other levels and its effects seen in practical application. The outcome of this testing can then be fed back into the modeling work, driving improvement on all the different levels. This approach avoids long isolated work on just one stage of the MRT lifecycle. A suitable iteration frequency needs to be determined for the work process, but is intended to be measured in weeks, not months.

A combined approach for modeling, representation, and testing would enable a balanced perspective on the different needs of different aspects of implementation and use. It would enable a robust model with robust representations and documentation. All of these, taken together, are the DDI 4 specification - not just the model itself.

The result of the work should be a first release of DDI 4. It would consist of the model, the major representations, and the documentation for these. The overall framework for modeling, generating the representations, generating documentation, and testing will be established as a necessary part of producing the release.

The figure below shows the envisioned MRT life cycle:



*(**Key:** PIM = platform independent model; PSM = platform-specific model)*

The model needs to be improved according following requirements:

- Creating the model
  - Consistency and stability of the model
  - Persistent expression of the model in canonical form
  - UML conformance / usage of UML
- Using the model
  - With UML tools, connecting to other models
  - Subsetting the model, useful views or other subsetting mechanism
  - Supporting the main representations efficiently
  - Lossless roundtrip of metadata in different representations

The generation of representations needs transparent and straightforward transformation rules.

The work of the group should focus on DDI 4 core with following goals:

- Specification as UML model and representations as XML Schema and OWL
- Library from which subsets for major use cases and audiences can be built
- Core areas, i.e. conceptual, data description, and process
- Transparent model which is suitable as a basis for interoperation/alignment with other standards
- Clear identification of conformance to or divergence from previous versions of DDI
  Interoperable UML library which can be used in major UML tools.
- Applicable to data from other domains as well as social science, with domain-specific features implemented as extensions and clearly identified (if any). The data description should support logical and physical description of rectangular, key-value pair (e.g. Hadoop), and event data, as well as data cubes and individual datum cells. It should also address the modelling of collections of metadata items related to data description.

It is recognized that a determination of the production platform and process will be an immediate and critical topic to be addressed by the MRT to enable this cyclical development. This topic will be undertaken as a matter of priority, working with other relevant groups within the Alliance.

DDI 4 Core

The focus should be on the core features of the DDI 4 prototype which comprises the conceptual, data description, and process areas. These three areas are useful for a large variety of audiences and use cases independent of specific needs. They represent a common ground for different requirements. The focus on these three core areas enables the specification of a DDI 4 core which is suitable for cross-domain use. The scope of each of the three areas needs to be determined.

A robust approach (regarding model and representations) for these three features would enable two things:

- Providing a core DDI 4 release which would be available for use
- Providing a robust “engine” for adding additional content features like data capture and other generic and specific components. The “engine” represents the model, representations, related transformations, test instances, prototype software, and the production framework.

The release of DDI 4 Core is roughly planned for the end of 2019. The overall attempt is to prove that DDI 4 is a usable and promising specification.

DDI 4 Core needs to be done in a way that additional areas can be added later without any change to the core. In UML terms, addition of further features can be achieved by adding new packages and classes or by specialization of existing classes.

Alignment with other Metadata Standards

DDI 4 Core should be aligned with or interoperate effectively with *at least* the following other standards:

- GSIM – The Generic Statistical Information Model. Reference standard in official statistics
- DCAT – The Data Catalog Vocabulary is a W3C RDF vocabulary designed to facilitate interoperability between data catalogs published on the Web.
- SSN – The Semantic Sensor Network (SSN) ontology is a W3C ontology for describing sensors and their observations, the involved procedures, the studied features of interest, the samples used to do so, and the observed properties, as well as actuators.
- CSV on the Web –A W3C standard to express useful metadata about CSV files and other kinds of tabular data.
- PROV-O – A W3C standard to represent and interchange provenance information generated in different systems and under different contexts.
- BPMN – The Business Process Model and Notation is a graphical representation for specifying business processes in a business process model.

The alignment should be done in such a way that DDI 4 supports combined use with each of these standards. Ideally, metadata can be migrated from DDI 4 to each of these standards and vice versa, as appropriate to identified use cases.

Additional DDI 4 Areas

For each additional area, a clear distinction should be made if the area is suitable for cross-domain use or is specific for the social science domain (like data capture with questionnaires). Work on additional areas can be done at the same time as the MRT group is working. But any additional work should have second priority and can only be done if the goals of a DDI 4 Core release are not affected. Example areas are data capture with questionnaires, coverage of DDI Codebook, methodology, and qualitative data.

Requirements

Technical Requirements

The MRT group should improve the technical requirements regarding each stage of the MRT lifecycle and create related guiding documents, based on input from the business requirements (see below), review activities of the existing efforts, and input from implementers. Furthermore, the model needs to be improved in several identified areas:

- Simplification of the model (i.e. less inheritance and less specialized classes)
- Review of collections (use of appropriate UML properties, use of collections throughout the model)
- Review of design patterns (relationship to acknowledged software design patterns, relevance of design patterns for users of the model and of the representations)
- Review of views (definition and effective use of subsets of the model)
- Model as portable UML library which can be imported into major UML tools (validation and reuse (in other context) of the model)

Business Requirements

Work on business requirements can be done as a parallel activity to the MRT work on the core and fed into identified technical requirements. This work on business requirements is independent because

areas of DDI 4 Core can be useful for any use case and audience. Providing a robust "engine" will be necessary for any features of specific business requirements. It is expected that implementers of the core will help refine business requirements vis-a-vis other standards and more generally for the core in their specific applications. The business requirements will be important in the future (after the release of DDI 4 Core) to determine additional areas which should be added to the DDI 4.

A review of the documentation is necessary with the goal that a descriptive language is used which is accessible to a wide user community. The class-level documentation should include mappings to earlier versions of DDI (and potentially to other useful targets) at with a level of detail sufficient to support consistent implementation.

Organization and Structure

It is anticipated that several different tasks will need to be completed. These would include the following:

1. **MRT Working Group Coordination** – a coordination team will perform this task, holding regular meetings to manage the overall work of the group. Attendance would include specifically invited members of the other task teams on an as-needed basis, but would also consist of a core of people responsible for managing the work. Meetings would be open to all members of the effort.
2. **Modeling Tasks** – modeling work will be conducted by task teams responsible for further developing and maintaining the model in each core area, and for identifying and tracking the technical requirements of the work. The outputs of these teams would be representations of the model in an agreed XMI format (ideally Canonical XMI) which would form the input to the generation of representations. Each team would focus on one of the DDI Core areas – other modeling work might be conducted in other groups as agreed, but these would not be the top priority of the overall effort.
3. **Representation Tasks** – each representation being developed as part of the standard would have a task team responsible for taking each iteration of the model and producing bindings into their target syntax, and to provide assistance to those working on the documentation of the representation/model. These teams are likely to be small, and it is anticipated that representations will be generated from the model, so that the job of repeatedly generating representations is not onerous.
4. **Documentation Task** – the overall documentation for the model and representations is handled by another task team working in coordination with the other teams. To the extent possible, documentation should be generated from the model and representations, according to whatever system is used for producing and packaging it (presumably using the existing tools and making any needed adjustments).
5. **Testing and Liaison** – members of the group will be identified as liaisons to different projects acting as implementers and testers of the ongoing drafts. These individuals would feed back issues and technical and business requirements to other groups as needed.

It is intended that the structure of these task teams be fairly fluid, so that communication is as easy as possible. A high degree of cross-membership in the task teams is anticipated, given the size and

expertise of the group and its members. Each task would have a lead assigned as the person responsible for achieving that task and communicating any status and issues to the coordination task team. It is assumed that the leads would be regular attendees at weekly coordination meetings. Face-to-face meetings would support the overall work most efficiently. Possible options seem to be in the margins of NADDI in April and at Dagstuhl in October.

It is anticipated that the MRT group would work closely in conjunction with the Advisory Group and other parts of the DDI Alliance (e.g., the Technical Committee) to ensure that results are in line with the overall workings of the organization.

Appendix 9

DDI4 Prototype Review: Status Report

Wendy Thomas – 2019-02-19

Actions:

Review period began: 2018-10-25
Initial period was to end 2019-02-04 (reminding sent out on 2019-01-16)
Review period was extended to 2019-03-31 as the request for review was not distributed to the RDF community as planned

Results:

As of 2019-02-18 we received:
33 comments via the JIRA tracking system:

- 27 from Larry Hoyle,
- 4 from Flavio Rizzolo,
- 1 from Joachim Wackerow
- 1 from ICPSR

Prior to the public review 37 Issues were identified by the Technical Committee that were determined should be dealt with post-prototype review (either the issue was complex or required an amount of time which would have delayed the release of the Prototype for review). These issues were noted in the Technical Committee comments on the Prototype.
The TC has performed only a single triage on the filed issues just prior to the EDDI Sprint. This covered the first 22 issues filed and resulted in 7 review issues plus 20 post-prototype issues to the Modeling Team. These dealt primarily with broader modeling issues.
In mid-January the Advisory Project Management Group discussed the response rate and we were informed that the announcement to the semantic web lists was delayed. The decision was made to extend the review period to the end of March and to follow-up with individuals and organizations who had previously provided feed-back on earlier reviews.
Of the filed issues approximately 70% deal with specific minor bugs or improvements, one issue was filed on the overall prototype, one regarding the Statistical Classification and its relationship to the DDI product XKOS, and 10 issues on broader modeling issues (use of UML, Functional Views, Patterns, etc.). A summary of the issues identified during the TC review is found at https://ddi-alliance.atlassian.net/wiki/spaces/DDI4/pages/589824023/DDI+Prototype+TC+Review+comments

Appendix 10

Summary Report: DDI 4 Core Sprint, Ottawa, 22-24 April 2019

I. Overview

This document provides a brief summary of the work which took place at the DDI 4 Core Sprint in the margins of the North American Data Documentation Initiative Conference, from 22 to 24 April 2019. It describes the agenda, participants, work-streams, and deliverables. These are placed in the context of the overall work plan of the group, with an emphasis on the year-end deadline for the work.

The Sprint was a meeting of the Modeling, Representation, Testing (MRT) working group. This group has been formed as part of the DDI Moving Forward project, and largely replaces the Modeling Team, but has a focus which includes not only modeling but also testing of the syntax representations and other work products.

In general terms, the Sprint was very successful – all of the anticipated deliverables were completed, and some topics for future work were discussed or explored, such that future progress will be more easily realized. In terms of the overall workplan, target milestones have been met, and in some cases exceeded. Significant work remains, but with the progress made at the Sprint, delivery of the DDI 4 Core at the end of December 2019, as a product ready for review and publication is still a realistic goal.

II. Background: Goals and Timelines

The DDI 4 Core was identified by MRT as a subset of the DDI 4 released in the Prototype Review package (see the *DDI 4 Core Summary and Overview* document). It is intended to be a production release of some of the most useful functionality supported by that model and associated products, narrowed in scope to make resource issues more tractable. Emphasis is on the foundational metadata, data description, and some applications of the process model.

The MRT Working Group has adopted a working process somewhat different from earlier DDI 4 projects: a more limited scope has been identified, and short-term timelines established. The core features of the existing DDI 4 model are to be finalized and the entire standards product (the model, documentation, and syntax representations/bindings) is to be ready for review as a production release by the end of 2019. The working process is an iterative one, more fully embracing the Agile methodology which has to a limited extent informed all of the DDI 4 work up to this point.

Central to the work is the existence of a production system which will allow modeling to become part of a cycle which also includes the production of documentation and bindings. This system did not exist in a useful form at the start of the work, and prior to this Sprint half of the group's efforts have been focused on developing this critical infrastructure from the existing one (the TC production framework and the Lion Repository). The initial move off of the previous infrastructure was achieved at this Sprint, which is an important milestone in the overall working of the group, even if one which is not as visible in terms of the eventual standards product to be delivered.

One change from earlier production processes is the use of Canonical XMI as a format for describing the model. This format was agreed in discussions with the TC as offering several benefits. It serves as an exchange format for the UML model between the MRT and the TC, being designed as a portable format for such models. Further, it can also be used directly as a deliverable by users across a wide range of UML tools, a feature which is of increasing as implementers use DDI in new ways (e.g., not as XML or RDF, but as a model for analysis packages, repositories, and other systems).

In terms of the model content, the existing scope has been narrowed, but the substantial work of the past years forms the basis for the group's current efforts. It is in essence a finalization and productization of the model and derived products, informed by the recent implementation and review of the DDI 4 Prototype. This input has indicated that changes are needed in both the style and content of the model and related products. Further, the work will need to be passed on to the TC at the point where it is ready for public review and distribution – the TC is ultimately the part of the DDI Alliance which will maintain it. Thus, alignment and integration with the TC production and management systems has been given a high priority in the work of the MRT.

III. Agenda and Participants

The agenda for the work was:

- Definition of the DDI 4 Core/Scope and Extraction from the Lion Repository (added just prior to the Sprint)
- Resolution of Open Modeling Issues/Production of UML Modeling Guidelines
- Production of Examples for Datum-Oriented Data Description/Gap Analysis
- The Role of Functional Views/Subsets
- Standards Alignment/"Round-Trippability" between Syntax Representations/Bindings
- Annotations, Access, and Citation Information (added during the Sprint)

The first agenda item was added after discussions subsequent to approval for Sprint funding by the Executive Board and was not tied to a specific deliverable at the time of approval. The agenda item regarding Standards Alignment was seen as perhaps too ambitious in a 3-day sprint and was presented as low-priority in terms of the work programme. The work on Annotations, Access, and Citation Information was not initially a part of the agenda but emerged during the Sprint as a result of work in other areas (notably responses to issues in the Prototype review and the work on Modeling Guidelines).

Participants at the Sprint were:

Dan Gillman, US Bureau of Labor Statistics

Arofan Gregory, Invited Expert

Larry Hoyle, University of Kansas

Hilde Orten, NSD - Norwegian Centre for Research Data

Flavio Rizzolo, Statistics Canada

Wendy Thomas, Minnesota Population Center, University of Minnesota

Joachim Wackerow, GESIS - Leibniz Institute for the Social Sciences

Jay Greenfield (virtual), Invited Expert

Oliver Hopt (virtual), GESIS - Leibnitz Institute for the Social Sciences

IV. Work Streams and Process

Work during the Sprint was prioritized according to the deliverables, and these served as the focus for the work. The production of each deliverable was broken down into discrete tasks taking one or more 1.5-hour sessions, performed by 1 – 3 people working in small breakout groups. The working process was Agile, with periodic re-organization and re-ordering of the groups based on reviews by the whole team after each work session. In the past, this Agile-based process has proven to be very efficient for groups with a limited number of participants, as at this Sprint. During the work, there was an emphasis on documentation of all discussions and decisions, so that it would directly contribute to deliverables and serve as a record for those not participating directly. All documentation is available on the Confluence site used at the Sprint as a collaboration platform.

- DDI 4 Core Scope Definition and Extraction from the Lion Repository
- UML Modeling Issues and Guidelines
- Datum-Based Application of the Model: Examples and Gap Analysis
- Analysis of Functional Views
- Alignment with External Standards and "Round-Trippability"
- Annotations, Access, and Citation Information

See the list of the deliverables, below, for a description of what each work stream addressed at the Sprint.

The work concluded with a planning session to organize future work, based on progress in each area at the Sprint, and in relation to the plan for producing the December 2019 package. This included identifying future activities and milestones, taking into account both holiday schedules during the summer and at years' end, and the increased productivity anticipated at the planned Dagstuhl Sprint in the fall.

V. Deliverables

The following is a brief description of each of the deliverables produced during the Sprint. Note that some work streams produced more than one deliverable. In some cases, this exceeded the initial expectations of the Sprint – additional deliverables are generally in the form of position papers or proposals which will be addressed by the MRT Working Group in their normal weekly calls and future Sprints. Links to the latest version of all documents on the Confluence wiki are provided for each work stream.

Work Stream: DDI 4 Core Scope Definition and Extraction from the Lion Repository

- **DDI 4 Core Definition of Scope** – document outlining supported functionality, with the details of what was included in the extraction from the Lion Repository to support the DDI 4 Core work. Possible mid- and future-term functionality is identified, based on the portions of the model

included in the Prototype Review package, and the extent of the work already done in different functional areas. [LINK TO CONFLUENCE]

- **DDI 4 Core Model Extraction** – the model subset to be used for the purposes of finalizing the DDI 4 Core was extracted in the form of (proprietary) XMI. This was then modified to conform to the Canonical XMI form agreed as the production format between the MRT and the TC. Limited testing was performed in various tools and is ongoing.
- **DDI 4 Complete Model** – to guarantee that none of the work thus far conducted in the DDI 4 project is lost, a second extraction was performed including all of the model minus that portion which has been discarded over the course of DDI 4 modeling work in the Lion Repository. This extraction was for archival purposes and remains in the proprietary XMI format produced by the Lion Repository.

Work Stream: Modeling Issues and Guidelines

- **UML Modeling Guidelines –** this document reflects changes made to the modeling style of the DDI 4 Core in response to open issues following the Berlin Sprint, those raised during the Prototype Review, and consideration of UML best practice and tools support. Decisions reflected here will be implanted in the DDI 4 Core model by MRT moving forward.

  The modeling guidelines are a key tool to resolve filed issues on the model in a consistent way. They provide rules which ensure that the model is conformant to UML, and that it uses the chosen UML subset efficiently. The new UML guidelines complement existing business modeling guidelines on the formal UML level and ensure a clear and consistent implementation model which can be transformed efficiently to multiple syntax representations. [LINK TO CONFLUENCE]

Work Stream: Datum-Based Application of the Model: Examples and Gap Analysis

- **Examples of Datum-Based Applications –** This document provides three illustrative examples for a general audience, showing how the DDI 4 Core model can be used to describe unit-record data, multi-dimensional data/aggregates, and event data. This will in future form a key input to documentation and training activities – extension with an introduction to the model itself, bridging the gap between the conceptual level and the formal (UML), will be added, and the excellent introduction to the Variable Cascade produced at the Dagstuhl workshops in 2018 will be incorporated. [LINK TO CONFLUENCE]
- **Proposals for Additions to the Model –** Having identified the existing gaps in the production of the first deliverable, proposals were created for addressing them. These take the form of a proposal addressing the classes and properties needed to better describe Event data, and a proposal for describing multi-dimensional data (aggregates, time-series, etc.). These do not represent major extensions of the existing model, but refinements to make it both more powerful and easier to use for these specific applications. Reaching agreement within the MRT group to these proposals is the next step, following which appropriate changes will be made to the model for testing. [LINK TO CONFLUENCE - Tall Layouts] [LINK TO CONFLUENCE - Multi-dimensional/Cube Layouts]

Note: The group will be assessing No-SQL, key-value, and some other types of data stores to determine how best to address these within the current scope. This was not a work-stream during the Sprint but is an open issue for near-term work.

Work Stream: Analysis of Functional Views

- **Functional Views Position Paper –** Over the course of DDI 4 development, the role played by Functional Views has evolved. This aspect of DDI 4 was included in the Prototype Review package but has never been fully defined or documented. This position paper considers the history, purpose, and means of determining how the DDI Model can be made more approachable and easier to use on the basis of this mechanism. Methodical analysis of the model served as the basis for this discussion, which promises to be a fruitful technique for identifying the most useful organization of the DDI 4 Core products from a functional perspective. [LINK TO CONFLUENCE]

Work Stream: Alignment with External Standards and "Round-Trippability"

- **Position Paper on Standards Alignment** – The importance of aligning with external standards, to optimize the value of the DDI 4 standard, has long been recognized. Over time, the ideas about what this means in terms of specific uses of the standard has evolved. This paper summarizes the standards which may need to be addressed, and the issues raised when alignment with these is considered. It is anticipated that this topic will be a near-term one for discussion within MRT and will result in refinements of the existing model to better align with specific external standards. Further, it is expected to impact the technical expression of the model in syntax representations, to support transfer across these expressions and technology platforms. It is hoped that this work will produce guidelines for this aspect of DDI 4, as the basis for future work. [LINK TO CONFLUENCE]

Work Stream: Annotations, Access, and Citation Information

- **Position Paper on Annotations, Access, and Citation Information –** The modeling of this information in the Prototype Review package has been identified in some of the issues raised during the review and is also impacted by considerations of the modeling style for the DDI 4 Core work. This paper considers possible resolution of these issues and changes needed in the modeling guidelines as the basis for further discussion within MRT. [LINK TO CONFLUENCE]

VI. Summary and Assessment

Overall, the Sprint was extremely productive. All of the anticipated milestones were met, and the set of deliverables was somewhat larger than anticipated. In terms of the overall workplan, the MRT is now positioned to begin the iterative production and testing cycles identified as their intended work process.

While a significant amount of work remains, at this point the project seems to be firmly on-track, with a delivery of a production-ready DDI 4 Core package to the TC for review and publication at the end of 2019 expected. The results of the Prototype Review have been addressed as appropriate, including the integration of issue resolutions into the Jira issue-tracking system being used by both the TC and the MRT (this should ease transition to the TC at time of delivery).

All issues considered during the Sprint which were raised during the Prototype Review are been updated in Jira, working in coordination with the TC to make sure that all changes are known to both groups.

The package to be delivered covers only a subset of the overall DDI 4 model but will provide essential functionality to users and lay the foundation for expanded functionality to be supported moving forward.

Overall, the Sprint exceeded expectations in terms of deliverables, and helped to point the way toward final delivery in December and a relatively painless hand-off to the TC at that time.

Thanks go to all participants for their hard work and excellent focus during the Sprint. Thanks also to the DDI Alliance for supporting the Sprint and making this focused and intensive work possible.

Appendix 11

The Case for Continued Support of a Model-Driven, Platform-Independent DDI

5 September 2018

Dan Gillman, Jay Greenfield, Arofan Gregory, Larry Hoyle

I. Introduction

This document is in response to the minutes of the 27 April 2018 DDI Executive Committee meeting. In those minutes, comments are made which raise questions about the overall strategic direction of the Alliance, and which point out issues around resource allocation, prioritization, and process. This document presents an argument for a continued emphasis on the development of a forward-looking, model-based, platform-independent DDI standard: DDI 4.

To quote some specific statements from the Executive Committee meeting minutes:

> "Maggie noted there's been a disconnect between the needs of users and where our investments are going."
>
> "Achim noted that the Scientific Board should be looking into the future, with not all development necessarily driven by the membership or immediate user needs."

Under the "Strategic Plan" heading:

> "The draft plan is not yet ready for review and approval. We still need to make a connection between goals and allocation of resources. We need to figure out what the demand is for what DDI is doing and have that drive what we are doing in the strategic plan."

Under the "Budget" heading:

> "The draft budget shows requested expenses exceeding revenue by $67K. The Board will need to prioritize expenditures."

Several points are evident from these statements, borne out additionally by many other conversations within the DDI community more generally. While the official minutes of the 2018 Annual Meeting have not yet been made public, notes taken by the organizers of that meeting also bring out the themes found in the planning meeting notes.

The first of these is that available resources are not ideal for supporting all of the activities of the DDI Alliance. The second is that tension exists between the short-term and long-term demands being placed on the standard, which are very significant as regards decisions made about prioritization and resource allocation. Last and more generally, ***there no longer seems to be a consensus on the overall direction of DDI as an organization***. While not stated directly in the minutes, it is clear that there are different ideas of how the Alliance should proceed, and that these will affect the overall strategic vision.

This document addresses several topics:

- ***Resources:*** Resource demands exceed current revenues, requiring a decision to be made regarding which activities are highest priority. The current level of resourcing is insufficient to continue the development of the DDI 4 standard at the current pace, while still supporting other DDI Alliance activities. Additional sources of funding and/or other resources are needed.
- ***Users' Needs:*** Short-term and long-term needs on the part of DDI users and the standard itself must be met. The patterns of data use are changing within the social sciences and across many other domains, and DDI must support functionality and applications which were not as important in the past. There is currently some tension between these different requirements.
- ***Strategic Direction:*** There seem to be competing ideas of how best to move forward in terms of the overall strategic direction. This is a shift from the recent past, where there was a broad consensus within the DDI community.

With the right strategic vision, supported by appropriate resource allocation for development and marketing activities, DDI can overcome the present challenges and continue to prosper in the future, providing a useful standard to a broad community of users. The continued support of the ongoing DDI 4 development is the best way to ensure that this happens.

II. Resources

When we consider the overall resource picture within the DDI Alliance, it is clear that additional sources of funding and in-kind resources must be identified if the development of DDI 4 is to be pursued, and likely under almost any strategic direction. Historically, the primary mechanism for increasing resource levels has been to enlarge the membership. While this approach does incrementally increase revenues, the size of the gap between desired funding levels and the available budget is formidable and would require a massive increase in membership to bridge. Given historical performance, such an increase seems unlikely to result from traditional outreach activities alone, even if these become a point of focus.

There seems to be a distinctly naïve sense of how effective outreach is for the DDI Alliance. There is on the part of some a faith in the idea that promoting the standard, rather than focusing on the functionality which the community of users (and potential users) wants, is the best mechanism for growth. Historically, this would not seem to be the case – DDI has always, even if only informally, managed to reach its target audiences. Convincing most people within a community that adopting a standard is important is always difficult, as the end-users (e.g., researchers) are uninterested in infrastructure standards. All they care about is the functionality of the tools they use. No amount of marketing to such an audience is likely to be successful. Those interested in infrastructure are likely to already know of DDI's existence, and they will judge its utility on what it does for them, not on how well marketed it is.

Another idea which has been discussed is to pursue grant funding. While the initial creation of DDI was funded by an NSF grant, grant funding is notoriously difficult to obtain for infrastructure projects. Furthermore, the development and maintenance of a standard is not a research project (although DDI 4 development can be seen in this light) but instead requires ongoing support, including for maintenance. While worth pursuing, grant funding may not be easy to obtain.

One benefit which the DDI 4 development offers, however, is the ability to engage with organizations from other domains, and to partner with them to support development of a standard which would be of benefit to both the Alliance and partner organizations. DDI 4 – unlike earlier versions of DDI – is explicitly aiming at a cross-domain audience, and at supporting data reuse where the data are coming from communities which do not necessarily use data structures typical of the social sciences (rectangular files, etc.). The DDI Alliance has an opportunity to pursue resources at the needed levels not only through increased membership, but also through partnership with organizations in external domains having similar needs.

When we look at how the use of data is evolving, we see an increased demand for data integration across traditional domain boundaries. Often, other domains have their own ontologies, standards, and platforms, which may or may not use XML technologies, and which may have widely variant data structures. Convincing the members of such a community to collaborate in the development of a metadata standard is only possible if that standard will support the technologies, processes, and types of data which they use in their own work. Doing this on the basis of an incremental evolution of DDI 3 is not feasible – doing this on the basis of a major revision (DDI 4) is. The following considerations are the most important:

- DDI 4 is expressed in UML, which is the dominant notation for modelling across all domains.
- Because DDI 4 is model-based and can be flexibly bound into a wide range of syntaxes, it becomes far more attractive to other domains which do not necessarily use the same technologies and approaches as the social sciences.
- Any external domain will need to provide the requirements of its users and see that they are combined with those of DDI users. This will likely involve joint ownership and governance of the standard, which would be the basis of joint funding and resource-sharing.

Examples of such collaborative standards development exist. Given the work from within the DDI community in the past to align with the standards in other domains such as health research and official statistics, as well as the upcoming DDI workshop at Dagstuhl on the "Interoperability of Metadata Standards in Cross-Domain Science, Health, and Social Science Applications," securing such a partnership is plausible. DDI has positioned itself well when it comes to collaboration with domains outside the social sciences.

III. Users' Needs

There are several different perspectives on how we can best meet users' needs coming from within the existing community. Some existing users would like to see an enhanced version of DDI 3, in XML or the derived RDF bindings, which would provide functionality not currently in the standard. Other users have requirements for which a platform-independent model (that is, binding-neutral) would be necessary.

According to the last minutes posted on the DDI Alliance site, DDI 3.3 – a version of the existing production standard – was discussed by the Scientific Board at their meeting in 2015. The version of DDI 3.3 currently out for review includes the changes discussed at that meeting, and a range of other features, some of which are taken from the DDI 4 model, and some of which have come from elsewhere, their origins unrecorded. There is no evidence that DDI 3.3 has ever been discussed as an agenda item by the Executive Committee or the Scientific Board since 2015, based on a review of the posted meeting minutes.

If extended functionality is added to DDI 3 to satisfy short-term user demands, the work of data modelling to support the new functionality must still be performed. We see this in DDI 3.3 in the way that features of the DDI 4 model are being incorporated. The modelling work which informs extensions to the existing production version of DDI is the same as that required to develop DDI 4. Without a model, the desired functionality cannot be supported. In order to satisfy short-term demands, ***the incremental addition of functionality to DDI 3 may be necessary***. That being said, this functionality must, to the extent possible, be aligned with constructs that are part of DDI 4 so as not to foreclose the DDI 4 path in the future. DDI 3.3 shows that this can be done, but it is also clear that the process for doing so has lacked transparency and coordination across the working groups.

While the requirement to meet short-term user needs is one that everyone recognizes, the trade-offs between work on a platform-independent model and an XML-centric one must be ***openly*** examined and discussed. This does not seem to have been the case to a sufficient extent with DDI 3.3, and we feel that before any further work on this product line is undertaken, transparency and due process have to be followed. The work on the DDI 3 standards will inevitably pull resources away from the work on DDI 4. This trade-off must be made with the informed approval of the Alliance as a whole, and not by small groups working independently.

It is worth considering here the differences between a platform-neutral model such as that being developed in the DDI 4 work, and one derived from a platform-specific model designed and expressed as a canonical XML structure (DDI 3). Even though not expressed in XML, a model derived from the XML design of DDI 3 will always carry with it the assumptions that make this version of the standard a good XML standard. As an XML standard, it has evolved into an excellent design over time, across minor versions. This same design, however, is likely to present barriers to effective binding into other syntaxes. These issues have become clear in the recent work on producing an RDF binding for DDI 4, and in the work of creating an R library for working with the newer version of the standard – a document detailing some of these issues, "DDI 4 and Data Structures," can be found at https://drive.google.com/file/d/1QIJJcMUVZWpZ1DvkFH-GKM6S3BIWTA3s/view?usp=sharing.

RDF bindings have been explored on the basis of the DDI 3-derived model but have not been released for review. Similar work has been done in creating a UML model for DDI 3, and it, too, has not been released. (Public release of such new products should be the decision of the DDI Alliance as a whole, given their potential to cause further confusion among users.)

Thus, the issue is not that other bindings are impossible or that a UML model cannot be created – it is that they are sub-optimal. When the XML technology stack moved away from document type definitions (DTDs) to W3C XML Schema, DDI 2 – a standard based on a canonical XML DTD – was expressed in the newer structural description. The resulting XML schema leveraged none of the features of the newer technology (e.g., strong data typing), which were the reasons it was being adopted throughout the technology world. The same type of phenomenon can easily happen if we try to force an XML-centric design into other syntax bindings like RDF now and JSON in the future. The resulting bindings are likely to be of little practical value.

When we consider long-term user needs, we see that there are users who are looking at DDI 4 specifically because it can do things the existing production versions of DDI cannot do well. In some cases, DDI 3 could be extended to meet their needs, but in other cases it cannot. The existence of

domain-specific Functional Views in DDI 4 developed in collaboration with domain experts – e.g. epidemiologists and social media experts – exemplifies the type of feature in the new model that can provide a high degree of flexibility to meet divergent user needs, and which does not exist within DDI 3. Furthermore, ***having a platform-independent model future-proofs the standard*** against the inevitable changes which will occur within the technologies used for system implementation, and over which the DDI Alliance, like any other domain standards organization, has no control.

IV. Strategic Direction

There are many factors operating within the DDI Alliance that have caused the consensus of the early DDI 4 work to begin to fragment. Delays in producing a production prototype of the model have frustrated many users who need the expected enhanced functionality, and in some cases these delays have fueled the demand for an extended DDI 3 rather than a DDI 4 which, for all its strengths, will not be delivered in as short a timeframe.

The time that it has taken to develop the DDI 4 model is itself an example of how the boundaries of an existing approach constrain thinking of a broader approach. One of the goals of DDI 4 was to expand the coverage beyond the somewhat questionnaire focused view of capturing data given the increased use of alternative types of measurement (e.g. biomarkers). Even so, thinking in terms of just questionnaires tended to creep into the conversations well into the model building process. Moving to the more abstract UML model helped broaden the focus.

Demands for new types of data description and non-XML bindings have been hampered by the lack of necessary expertise within the community. The RDF bindings were a primary example of this. While the problem has been at least temporarily solved by hiring external consultants with the appropriate skills, this has implications for the overall DDI 4 model that are still being addressed. There has been a considerable delay in releasing the production prototype as a result, a source of frustration that has been de-motivating for some in the community.

The resources that would be used to produce extensions to DDI 3 to address short-term demands will take away from the resources needed in developing the DDI 4 model to satisfy longer-term ones. These two conflicting demands present the biggest challenge moving forward, and this situation (robbing Peter to pay Paul) can only be ameliorated by finding the needed resources to perform both tasks if an extended DDI 3 is chosen as an ongoing development path.

The existing charter and bylaws were the result of criticisms of the TC (formerly named TIC) as having too much control and operating in a non-transparent manner. The current structure resulted from an external audit of the DDI Alliance. The competition for resources surrounding the development of DDI 3.3 has already shown us that we are in danger of recreating the same situation that led to the audit. We feel strongly that this must not be allowed to happen again – we have a process guaranteeing transparency and accountability, and feel it must be followed. These decisions are too critical to be made behind closed doors.

As we look at the longer term, XML-based standards will inevitably become obsolete. The current use of data in “mashups” (the dynamic integration of data with information from multiple sources at presentation time) and for other purposes hints at the kinds of changes we are likely to see and the need to support future bindings. The shift away from relatively expensive collection of data using

questionnaires to programmatically capturing data from a variety of sources is another (“big data”). XML technologies, while possessing many strengths, were not designed for these types of emerging applications. Strategically, the identified goal of having a platform-independent model suitable for binding into many different syntaxes is still the best choice. While this fact will not be reflected in short-term user needs, it remains true scientifically.

If the confidence of the community is to be regained, the DDI Alliance must identify the best strategic goals, and outline a plan for realizing them while not ignoring the short-term needs of users. A clear case for management decisions must be presented, and it must reflect discussion which is open and transparent. The community must be given an opportunity to understand the arguments behind the chosen strategic direction, and must be given sound justification for any trade-offs made to meet both short-term and long-term demands from users. This is admittedly a challenging task, but one that cannot be achieved if the long-term goals that initially rallied the community around DDI 4 are abandoned for the sake of short-term objectives.

V. Conclusions

The best approach for DDI is to stay the course as regards DDI 4 development. The reasoning behind this will need to be made clear as the justification for choosing this strategy, and an opportunity for the community to discuss the future directions of the standard in an open and honest fashion needs to be provided. This discussion will likely need to extend beyond the membership, to include those from outside the social sciences.

As a result, DDI 4 provides an opportunity for the Alliance to partner with other domain organizations in the creation of a standard that will have a broader scope, and will allow for the use of data across domain boundaries. ***This represents a way of solving the critical resource issues*** previously identified.

Short-term objectives met by diverting resources to an extended version of DDI 3 will remove some of the tension seen today within the organization. This may be tactically necessary, but comes with a cost: the sacrifice of achieving important strategic goals that will allow DDI to remain relevant and prosperous into the future (future-proofing). Such trade-offs must be made with the greatest care, and that degree of care can only be achieved through an open and transparent process.

For these reasons, it is clear that the development of DDI 4 should remain the primary focus of the DDI Alliance, and the process guiding that development must remain faithful to the charter and by-laws of the Alliance.

Appendix 12

Dear Dan, colleagues, and the DDI community,

I'm writing today on behalf of the DDI Alliance Executive Board, to respond to the issues you recently raised in your Open Letter to the DDI community. In the letter you have identified three core areas of concern that require consideration - resources, user needs and strategic direction. We propose to address these three issues in reverse order.

The DDI executive has been working on the revisions to the Alliance strategic plan, one that is intended to guide the strategic direction of the organisation for the next 4 years, for some time now. There have been two meetings now (in Kansas and Montreal) where we have consulted on this plan, and we intend to bring this to a close and publish the final version of the plan by the end of this calendar - this is the first priority on the Executive's agenda.

Within that plan, there are three areas of activity - development of the community, development of the organisation, and the development of the standard. In the Executive's view, there is general support for all three areas. The issues raised in the open letter largely focus on one area of activity - that of the development of the standard. It is here therefore that we focus first.

In terms of the development of the standard, it would appear that the key question that is raised is managing the maintenance and improvement of the existing standards (DDI-Codebook and DDI-Lifecycle), and the development of the new standard, DDI4. This is a challenging question, given the resource requirements that each entails.

In recent times, the focus of the work program has been on development of DDI 4; DDI's resources have been allocated in large part to this work. The DDI 4 prototype has now been released and is open for public review. The review period will be open through December (at least). After the review period, there needs to be transparent and honest community discussion leading to a decision as to how to progress with the next stage of development. We believe that the next stage of work in the development of DDI 4 requires such review in order to understand how best to commit resources in the future. The DDI sprint in Berlin should contributed to this discussion. This review should establish the community requirements for the development of DDI4 into a production version going forward.

There is also need to maintain and support the existing standards going forward. That means continuing to support DDI-Codebook and DDI-Lifecyle (DDI 3) while simultaneously developing DDI 4.

In your letter you do note:

"In order to satisfy short-term demands, the incremental addition of functionality to DDI 3 may be necessary. That being said, this functionality must, to the extent possible, be aligned with constructs that are part of DDI 4 so as not to foreclose the DDI 4 path in the future."

This view is consistent with the expectations of the Committee - that DDI 4 will continue to be the primary focus for standards development going forward, [a] with support for the DDI-Lifecycle and DDI-Codebook on an as-needs basis[b][c]. In our draft strategic plan, the expectation is that such support will come in the form of an increased emphasis on marketing and outreach, and on training programs such as the recent "train-the-trainer" program in Dagstuhl in September 2018.

There is one question we would like to put to the community as to whether what additional technical support will be required. Does the community foresee a need for additional functionality through additional releases of DDI-Lifecycle, or simply a focus on bug fixes in DDI-L (and DDI-C)? As noted in your letter, there are both resource and technical implications of adding functionality to DDI-L, and these need to be considered if there is demand from the community for such extensions. There is also related work products such

As noted in the annual meeting in May, the Alliance is currently facing financial constraints[d] which limit our capacity to pursue as wide a variety of activities as we would like. We are actively exploring new options for organisational resources, but these have proved difficult to identify thus far. This leads to our second question - are there additional resource[e]s the community can identify (or can contribute) that might enable additional support or development activities to occur?[f]

There is one final issue in regards to governance that should also be considered here - that of the role of the Scientific Committee in setting direction for development of the standard. This is related to our strategic direction on "Development of the Organisation". The Scientific Board has in essence had very little role to play since its establishment in 2015 as an outcome of the organisational review conducted in 2011-12. It is through the effective operation of the Scientific Committee that we believe the above concerns could and should be addressed. We believe that an active and collaborative Scientific Committee should be able to provide the direction required to determine the development and technical priorities for the standard, while the Executive Committee provides direction on the operations of the Alliance. [g][h]

To this end, Jared has recently put out a call for nominations for the Vice-Chair role of the Scientific Committee, and we would encourage interested members of the community to put themselves forward. We would also like to see a larger group of participants active in the Scientific Committee, and would envisage a larger group of regular participants in a regular meeting cycle (similar to that of the Executive Committee) to drive this direction.

We will be actively seeking out community feedback on these issues as part of the DDI4 Model review. At the same time we would like to hear feedback from the community on the two priority questions we have outlined above:[i][j][k]
1. Does the community foresee a need for additional functionality through additional releases of DDI-Lifecycle, or simply a focus on bug fixes in DDI-L (and DDI-C)?
2. Are there additional resources the community can identify (or can contribute) that might enable additional support or development activities to occur?
3. Are there members of your organisation willing and able to contribute to an updated Scientific Committee, to provide advice on the direction for each of the DDI work products?

We encourage and look forward to the response from the community.

Regards,
Steve McEachern
Chair of the DDI Executive Committee,
on behalf of the Committee
(everyone's names here...)

[a]This seems stronger than the sense of the committee. I think some of us are waiting for some assessment before we make this statement.

[b]If these standards are being used, as-needs will not be sufficient. We must acknowledge that ongoing support for existing standards is time-consuming. I do think this is a key issue. We're asking the budget that worked for 2 products to support 3 products. There is not enough efficiency to do it. (And budget here could probably be financial budget and also volunteer time.)

[c]FOR DISCUSSION 20/12

[d]The Alliance has faced all the time similar financial constraints. This is not new. Nevertheless the development of DDI Lifecycle was supported for years (with less membership fees) and now DDI 4 is supported. What is really the change?

[e]This is one of the primary tasks of the Executive, exploring new options, not only raising this question.

[f]This issue needs emphasis; the current impasse is due in large part lack of adequate financial resources. If there were enough monies to fund product development along with outreach and promotion at their desired levels, this issue would not be as dire as it has become.
Future directions have to realistically take into account current budget situation; ideally proposed initiatives should suggest or pursue specific manner in which they can be paid for (or what the likely return on investment will be).

[g]In general, I still prefer a short and simple response. Thank you for your open letter; the Executive, as the overall policy, budget, and strategic authority for the Alliance, must consider all aspects of the DDI Effort--the health of the Alliance, the DDI User Community and its engagement, and further development of the standard. To that end, we will attempt to carefully and fairly and transparently balance competing interests and needs of the Association (brought language, but you get the idea--short and simple and expresses the role of the Executive).

[h]I have a problem with this language as written--seems to confuse the responsibilities of the Scientific and Executive Committees.

[i]Proposing a formal needs assessement and survey of Alliance members as to strategic direction and allocation of financial resources would make the final EB decision easier and more defensible.
There is too much talk about "what the community wants" without any concrete mechansim for determining this. Structured interviews with Alliance member reps first (power users and tools developers second?) would put this question largely to rest.

[j]It sounds like a good approach. The risk is that the result is that no development at all is required for the majority of members. I guess this would have been also the answer of the majority of the community to a survey in 2003 (the creation year of the Alliance for the purpose of DDI Lifecycle). At that time only a few organisation weren't happy with Codebook because they produce data and not just archive data.
My question is: is a survey really the major answer to determine the strategy of the Alliance.

[k]FOR DISCUSSION 20/12

Appendix 13

DDI Alliance

Strategic Plan 2018-2022

Draft for comment - 15 May 2018

Introduction

As the DDI Alliance moves into the next phase of its development, there are some overarching priorities that the Alliance needs to address. Framed broadly, these priorities fall into three core areas: the DDI community, the Alliance as an organisation, and the set of DDI standards and work products that the Alliance maintains.

The strategic plan is developed along the following lines to address these broad priorities:

1. Community and outreach: how do we engage with the DDI community and understand the community's needs?
2. Organisational needs: what structures and systems does the Alliance need in order to meet those needs, and how will it maintain those structures and systems in the long term?
3. Standards: what products does the Alliance provide and maintain, and how do those products meet the needs of the Alliance and the broader community

The DDI Alliance budget and work program, to be developed subsequent to the acceptance of the Strategic Plan, are then intended to align with these strategic priorities. This will include the need to identify resources (including money, time and in-kind contributions), and responsibilities (e.g. party/organization/team) for each part of the strategic plan. Similarly, the resource constraints within the Alliance will by necessity limit the extent to which we can achieve the goals set out in this Plan.

Working principles

The specific strategic activities proposed in the plan have been established with the following principles in mind:

1) Don't leave anyone behind--no dead end with any prior DDI track
2) Lower barriers to entry/use
3) Respond primarily to user demands/requests
4) Market, market, market
5) Simpler is always better

6) Let user requests drive development

Strategic Priority Area One: The DDI User Community

PROBLEM STATEMENT:

The DDI community is driven by volunteers and the standard's success is directly related to its ability to attract, develop, and retain individuals willing to contribute their time and efforts. Users in the DDI context can be understood as both the formal members of the Alliance, but also more broadly other agencies, archives, statistical agencies, software developers, data providers, and related entities. These users can have complementary but also sometimes conflicting needs to be addressed by the Alliance and by its standards and work products.

By extension of this idea, understanding the community and its needs is probably the most important thing DDI can do to ensure its future. Doing so is likely to activate a virtuous cycle of growing membership, use, tools, and usability of the standard.

By enabling better interaction and engagement with the community, the Alliance is able to better understanding the needs that the Alliance products are supporting. In doing so, better engagement with the members and the broader community of users should also enable additional resources to become available to the Alliance for contributions to community efforts. The core concerns to address in Strategic Actions in this area focus on understanding and support of the needs of the DDI user community and membership, and the expansion and extension of the user community into related areas and disciplines[a].

STRATEGIC ACTIONS:

1. Engagement with Global Digital Research Infrastructure.

Now is the time to make the successes of DDI specifications more widely known in the community making up global digital research infrastructure. This global community has focused on building infrastructure to support interdisciplinary research on today's big science topics. The interoperability of metadata for discovery and access to research data is an essential component of these national and international infrastructure developments. DDI specifications can play an integrating role in making social, economic, and behavioural data available to emerging interdisciplinary research endeavours. This strategic direction identifies three actions in which the DDI Alliance can increase its engagement with Global Digital Research Infrastructure.

a. Develop best practices to map and translate DDI for DataCite, schema.org and other key metadata repository services.
b. Engage with RDA IG's and WG's, CODATA, and Force 11 to advance DDI's integration into the larger digital research infrastructure framework.
c. Increase communications with other metadata standards setting organisations for discipline-specific research data types.
d. Fostering usage of DDI with other metadata specifications. Promoting cross-domain usage of DDI (therefore identifying suitable parts of DDI for this purpose)

2. Solving Common Problems with Current DDI Users.[b]

Periodically, DDI Alliance faces the criticism that its products are unknown to researchers even though the Alliance is working to solve common

problems confronted in research and is often engaged with current DDI users in finding solutions. Three actions are proposed to strengthen the working relationship of the DDI Alliance with current DDI users and to create new possibilities to engage with researchers who are not yet familiar with the Alliance even though in need of metadata solutions in their research.

e. Prepare guidelines to assist end users in their choice of DDI specification
f. Create validation tools and profiles to support interoperable DDI metadata across tools and organisations
g. Assist software developers of DDI tools through a gap analysis on needed tools, guidelines for software usability, training, and support letters to funders

3. User group development program

Recognising that there are often similar sets of needs among categories of DDI users, there is an interest in establishing user communities within the broader Alliance. This action proposed establishing an initial user group among national statistical organisations - major data producers for whom the documentation, discovery, and interoperability of their data are vital to their operations. Over the years, the DDI Alliance has worked with some NSO's to integrate DDI specifications with their implementation of the Generic Statistical Information Model and the Generic Statistical Business Process Model. The DDI Alliance and NSO community share common goals to establish and maintain high quality metadata standards for social, economic, and behaviour data. This strategic direction proposes three actions to strengthen the relationship between NSOs and the DDI Alliance and to communicate the benefits of such partnerships with other data communities.

h. Create an NSO advisory committee[c][d]
i. Promote successful DDI uses by NSOs
j. Use NSO outreach model to establish similar groups within other user communities

Strategic Priority Area Two: The DDI Alliance as an Organisation

The Alliance has a broadening set of both members and user needs, bringing with it new requirements for the standards and outreach that we do. One impact of this broadening reach is the need to become more "professional"[e][f][g] in the way the Alliance operates. This professionalisation includes the maintenance and development of our core organisational infrastructure (such as websites, marketing and project management). At the same time, we want to retain the core volunteer culture that formed the foundation of the Alliance, and continues to drive the participation of many members and participants in the Alliance. We are facing a period of volunteer and staff renewal, requiring the need to expand our core development base and volunteer community.

1. Generational Renewal

a. Recruit the next generation of knowledgeable and skilled core technical development team
    i. Who is actively engaged in this now? What are their organizations? What is the committee membership? Does it have a rotation? A leadership? Can we set up a schedule with a rotation of membership and leadership, with the leader of the committee responsible for identifying and planning for new leadership and new membership, say every two years?[h][i][j]

b. Expand skilled marketing team that is connected to relevant communities[k] (archives, software producers, data producers, statistical agencies, individual researchers, other standards)
    i. ~~Who is actively engaged in this now? What are their organizations? What is the committee membership? Does it have a rotation? A leadership? Can we set up a schedule with a rotation of membership and leadership, with the leader of the committee responsible for identifying and planning for new leadership and new membership, say every two years?~~

c. Renew active and engaged membership at the institutional level in the DDI community (strengthening the commitment)
    i. How many institutions do we have now? Can we show graph of membership over time?

2. Training: Enabling trainers to do what they need to do
    a. Recruit much needed human resources to offer multifaceted DDI training.
        i. Develop a role for a membership appointed DDI Alliance designated “Trainer”
            1. The official DDI Trainer will work with the DDI Alliance Training Working Group and offer dedicated support for training in all formats and types; as needed and in close collaboration with the various WGs and Executive Board requirements for at least a 1-year period.
            2. DDI Trainer will provide support for in-person training at conferences, workshops, seminars, as requested and funding permits.
    b. Build-up online training presence to expand current offering of training.
        i. Extend the current offering of online training materials to support self-driven, passive training through online and web-based training delivery
        ii. Develop web-based video tutorials such as “What is DDI?” , “How to get started with DDI?”, “Building reusable questionnaires with DDI”, etc.
    c. Support new trainers and users with easy-to-understand and reusable tools for DDI Training
        i. Develop reusable checklists for getting started with DDI
        ii. Develop and maintain a listing of organizational DDI user profiles, licensed openly for reuse
        iii. Develop, gather and share reusable training materials (e.g. training toolkits for different audiences and use cases)

3. Business Structure
    a. Establish a periodic review of organizational structure
    b. Develop a sustainable business model for the Alliance
    c. Develop an organizational succession plan for the Alliance

Priority area Three: Standards and Work Products

PROBLEM STATEMENT:

The DDI Alliance currently maintains two lines of the DDI standard (DDI-C and DDI-L), with a third line in development. This is complemented by a set of controlled vocabularies, registry services and related products, which serve to support the needs of the wide range of users in the DDI Community. Strategic actions on DDI Standards are intended to provide orientation on which standards and work products to develop and maintain, and why. There is need to be able to maintain the existing standards to ensure that we can continue to support small scale users such as academic libraries and research centres, while continuing to develop the new line of DDI4 model-based standards[l] and associated work products to support the expanding user base in communities such as statistical agencies and data producers.

STRATEGIC ACTIONS:

1. Maintaining multiple lines of specifications[m] and controlled vocabularies
    a. Offering stable specifications and controlled vocabularies (reference Work Products)
    b. Enable DDI specs to adapt to changes in information technologies[n] and bindings (XML, RDF, Schema, …)
    c. Production testing/validation for quality assurance
    d. Improve documentation/examples/best practices guidelines
2. Introduce validation tools, testing support, and profiles/views for users for the purpose of interoperability
    a. Test cases, test bed, test harness
    b. Rules for validation/reporting
3. Working DDI infrastructure: a network of resource-based ... (Achim to add)
4. Registries-repositories
    a. Specify DDI's vision of building DDI into Common Data Element registries (Strategic Plan & Vision)
    b. Identify ways for establishing registries for supporting existing legacy DDI metadata[o]
    c. Develop standard query and exchange protocols/interfaces

[a]Perhaps "into areas and/or disciplines that offer synergies for DDI"? A bit more focused on relationships whereby DDI can benefit, rather than just outreach to common areas of data management and documentation...because those areas happen to be similar. A subtle but important distinction.

Also part of this thought, especially below in 1a-1d: how do we determine what are fruitful relationships to pursue? How do we avoid dead-end collaborations, or ones in which the cost/benefit ratio is not in DDI's favor?

[b]A fundamental conundrum that needs to be acknowledged is the difficulty in knowing who uses DDI. As an open standard, users can just use DDI and we might never know about it. I believe I mentioned this in last week's call. Before we can solve "current DDI Users" problems, we first need to identify them, we need a better handle on the size and composition of our customers.

It may also help to get a handle on our audience characteristics by classifying or categorizing our audience by scope or extent of use? E.g., NSO's concerns

should probably be given considerably more weight than an individual researcher's?

[c]An advisory committee is ambitious. NSOs would expect active guidance. Can the DDI Alliance this really provide with the background of small resources and that NSOs are very experienced organizations in their field

[d]This again might not be the ideal wording - an advisory committee has a significant function that may be more than we intend

[e]Is there a better term to use here?

[f]Seems appropriate considering you are juxtaposing it with the volunteerism ethic currently driving many DDI operations.

[g]experts?

[h]Do we need to remove these substatements??

[i]Useful detailed questions and suggestions.

[j]We should replace with answers :).

[k]The Alliance would do well to consider a salaried marketing position that would be dedicated to nurturing organizational-level outreach and increasing awareness of DDI. Volunteer efforts, as worthy as they are, just don't have the legs or chops to accomplish meaningful development along these lines. DDI really needs someone who can move the needle.

[l]Since it's inception DDI 4/Model-Based/etc was always framed as a continuation of the DDI-Lifecycle line. Has this changed, and if so where is the documentation announcing that change?

[m]During the discussion on this topic, the question arose about the sustainability of maintaining multiple specifications. I believe that the hope is that DDI-Model will allow us to express specifications for DDI-Codebook and DDI-Lifecycle.

I believe that 3b is related to this point.

[n]Could we build in a regular external review of our technologies to ensure we're up to date?

[o]could re3data.org be such a registry?