# DDI Alliance Annual Meeting of Members

*(virtual only)*

**June 1, 2022, 16:00-18:00 CEST, 10:00-12:00 EDT**

<table>
<thead>
<tr>
<th>Time</th>
<th>Subject</th>
<th>Detail</th>
<th>Lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>15:00-15:05</td>
<td>Welcome and Introductions</td>
<td></td>
<td>Bill Block, <em>Executive Board Chair</em></td>
</tr>
</tbody>
</table>
| 15:05-15:15| State of the Alliance FY2022                 | - Update about last year’s work  
- Framing priorities for the next year                                                          | Bill Block                                  |
|            |                                              |                                                                                                | Barry Radler, *Executive Board Vice Chair, facilitator* |
| 15:15-15:25| Scientific Board Report                      |                                                                                                | Ingo Barkow, *Scientific Board Chair*        |
| 15:25-15:55| Group Reports                                | Working Group Chairs  
- [Controlled Vocabularies](#)  
- [Technical Committee](#)  
- [SDTL](#)  
- [Training](#)  
- Glossary  
- [CDI](#)  
- [XKOS](#)  
- Marketing | Bill Block, *facilitator*                    |
| 15:55-16:00| Alliance Budget                              | - FY22 [Financial Report](#)  
- FY23 funding requests                                                                       | Jared Lyle, *Executive Director*  
Bill Block                                         |
| 16:00-16:15| Alliance Priorities and Feedback Breakout    | - Attendees breakout into small groups to discuss priorities, concerns, or feedback for the Alliance. | Bill Block - facilitator  
Barry Radler - facilitator  
5-6 groups randomly assigned                       |

*controlled vocabularies* |  
*technical committee* |  
*sdtl* |  
*training* |  
*glossary* |  
*cdi* |  
*xkos* |  
*discussion questions*
<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
<th>Details</th>
<th>Facilitators</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:15-16:45</td>
<td>Breakout reports</td>
<td>- Each breakout group summarizes their feedback.</td>
<td>Bill Block - facilitator</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Barry Radler - facilitator</td>
</tr>
<tr>
<td>16:45-17:00</td>
<td>Retirement Celebration</td>
<td>- Recognition of Wendy Thomas</td>
<td>Diana Magnuson</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Recognition of Achim Wackerow</td>
<td>Mari Kleemola</td>
</tr>
</tbody>
</table>


Attendees

Bill Block
Barry Radler
Jared Lyle
Diana Magnuson
Mari Kleemola
Jon Johnson
Jeremy Iverson
Ingo Barkow
Wolfgang Zenk-Moeltgen
Arianna Caporali
Arofan Gregory
Carsten Thiel
Cathy Fitch
Dan Gillman
Darren Bell
Flavio Rizzolo
Hilde Orten
Kaia Kulla
Laura Molloy
Mylene Chaleix
Sanda Ionescu
Steve McEachern
Wendy Thomas
Jane Fry
Simon Hodson
Scientific Board Report 2021/2022

Prof. Dr. Ingo Barkow, Chair of the DDI Alliance Scientific Board
Hilde Orten, Vice-Chair of the DDI Alliance Scientific Board

DDI Alliance Annual Meeting of Members 2022-06-01
First... an invitation

• Annual meeting of the Scientific Community
• Tuesday 7th of June 2022 from 16:15 – 17:45 CEST at IASSIST in Gothenburg
• Hybrid meeting (at the premises and online)
• More details about last years work by the SB and the working groups will be given, especially in the context of the Scientific Work Plan
• For this meeting only some high level activities will be shown and discussed
Current Members of Scientific Board

• Chair (2021-2022)
  • Ingo Barkow, University of Applied Sciences of the Grisons, 2021-2025

• Vice Chair (2021-2022)
  • Hilde Orten, Sikt - Norwegian Agency for Shared Services in Education and Research, 2021-2025

• Members
  • Darren Bell, UK Data Service, 2021-2023
  • Simon Hodson, CODATA, 2021-2025
  • Flavio Rizzolo, Statistics Canada, 2021-2025
  • Carsten Thiel, CESSDA, 2021-2023
  • Wolfgang Zenk-Möltingen, GESIS - Leibniz Institute for the Social Sciences, 2021-2025
  • Jared Lyle, ICPSR, DDI Alliance Executive Director, ex officio, 2021-present
  • Wendy Thomas, Minnesota Population Center, Technical Committee Chair, ex officio, 2021-present

• Joachim Wackerow, Former GESIS – Leibniz Institute for the Social Sciences, 2021-2022 (retired)
The role of the Scientific Board

The scientific and technical body of the Alliance which represents the Scientific Community. The Scientific Board proposes the scientific work plan to the membership for approval and facilitates the scientific and technical work activities.

• The purposes of the Scientific Board according to the Bylaws are to:
  • Provide direction and coordination in the development of the substantive content of the DDI standards and other work products of the Alliance by its sub-committees and working groups within the context of the Alliance Strategic Plan.
  • Implement the scientific work plan agreed at the Annual Meeting of the Scientific Community.
  • Oversee the substantive content of DDI standards and other work products.
  • Undertake research and testing concerning proposals for DDI standards and other work products.
  • Develop and promulgate best practices for use of DDI standards and work products.
  • Assess progress and barriers to progress.
  • Provide a report on progress of the scientific work plan over the previous year, and proposals for the future scientific direction and related activities to the Annual Meeting of the Scientific Community.
The role of the Scientific Community

• The SC consist of all scientific representatives (it replaces the «old» Scientific Board).
• The SB reports to the SC during the annual meeting about proceedings of the past year and presents the Scientific Plan for the next years for feedback.
• The SB will invite SC to virtual meetings during the year to seek more input and engagement.
• The Scientific Plan of each term defines the boundaries in which the SB can act on its own.
• The SC does not vote on the Scientific Plan. This can only be done by the member representative.
• The member representative can shift the voting right for the Scientific Plan to the scientific representative.
Activities of the Scientific Board 2021/2022

• Internal proceedings and procedures have been clarified
  • Monthly meetings of the SB
  • Communication via Mailing List and Confluence inbetween
• SB has nominated contact persons from the board for each working group to facilitate individual support
• Regular EB/SB dialogue meetings have been established (chairs plus director)
• Regular WG meetings have been established
• A Scientific Community Meetings was held.
  • Presentations and recordings from the meeting on January 27th
Activities of the Scientific Board 2021/2022

• Confirmation of status for non-official working groups
  • Cross Domain Integration (CDI)
  • Paradata
  • Extended Knowledge Organization Systems (XKOS)

• Establishment of new working groups
  • Glossary

• Preparations for upcoming working groups
  • DDI Developers (setting up technical contacts, plan for a Hackathon)
Activities of the Scientific Board 2021/2022

• Temporary Working Group on resolution of DDI URNs
  • DNS NAPTR record request for ddi.urn.arpa is approved. This was the final step to build the basis for a DNS-based resolution of DDI URNs.
  • It was decided to form a temporary working group that will develop a proposal for high level goals goals and policies regarding the resolution of DDI URNs and to describe which kind of related technical services would be needed.
  • Use Case for the necessity of the steering function of the SB as there are existing developments from the TC.

• W3C Liaison
  • A liaison between W3C and the DDI Alliance has been established. The W3C contact is Pierre-Antoine Champin, W3C fellow and coauthor of JSON-LD. The DDI Alliance contact is Joachim Wackerow. See announcement.

• Budget statement for the FY23 prepared for the EB
Upcoming activities

• First physical meeting of the SB after two years of COVID19 in Chur / Switzerland 5th and 6th of September 2022

• Main tasks
  • Scientific Work Plan 2023/2024
  • Nomination and election of one or two advisors to the SB
  • Preparation for rotation of chairs after two years phase

• And…. retirement party for our long term former chair and member Achim Wackerow
The DDI Alliance Controlled Vocabularies Working Group has been continuously active since 2007.

We currently have five members – from ICPSR (US), FSD (Finland), NSD (Norway), UKDS (UK) and SND (Sweden). We would welcome at least one or two more members, preferably from different organizations, to ensure a diversity of perspectives which is important for our work.

To date, we have created and published 25 vocabularies that may be used in conjunction with both DDI-C or DDI-L, or indeed with any other metadata standard, since our vocabularies are external to the DDI specification. We have also produced version updates as needed, and most of our vocabularies are translated in multiple languages and used in the Cessda Metadata Model.

In the past year we have finalized, and published a new vocabulary for TypeOfTranslationMethod, a new class introduced in the most recent DDI-L version, 3.3. We also have a final draft for another new vocabulary, to be used with the DDI-L element IntendedFrequency (of data collection).

In the past year we have also worked to align our vocabularies to the FAIR data principles: we have a new versioning policy, applicable to all instances of a vocabulary irrespective of the specific language in which it is expressed; we have prepared a new code deprecation policy which will replace our current policy regarding code deletion; we will be implementing persistent identifiers at code level.

We are also working to include XPaths for the DDI classes referenced in our Usage section, to facilitate references to our vocabularies from the DDI schemas.
• Update to the DDI Agency ID Registry
  • Supports HTTP access to DDI identified object by resolving a URN to a pattern-based URL
  • Let’s small agencies provide access to their DDI objects without a local resolver
  • Developed by Colectica through support by the DDI Alliance
• DDI Alliance resolution system to provide access to DDI Controlled Vocabularies and RDF vocabularies
  • Currently at the test stage
  • Provided through the combined work of the UKDA and GESIS
• DDI Codebook Version 2.6
  • Expands support for the use of Controlled Vocabularies and improved discovery access through DataVerse
  • Public review will begin this week and extend through August
  • Revision supported by a group of 20+ individuals
• Input review for the implementation of COGS as the production framework for DDI Lifecycle
TECHNICAL COMMITTEE FOCUS OF 2022/23

• Full scale implementation of the COGS production environment for DDI Lifecycle
  • Review of a completely serialized format for DDI Lifecycle
  • Testing of multiple implementation languages for DDI Lifecycle (XML schema, RDF/OWL, UML XMI, JSON, etc.)
• Coordination of the DDI Product Suite to improve inaction, transformation, and the support of multiple implementation languages
  • This work will involve the working groups involved in product development and maintenance (TC, DDI-CDI, XML, and SDTL)
Structured Data Transformation Language (SDTL)

DDI Working Group Report
Structured Data Transformation Language (SDTL)

• SDTL is a language for describing data transformations
• SDTL was created by the C²Metadata Project to automate the capture of provenance metadata from statistical analysis software
  • SPSS, SAS, Stata, R, Python
• SDTL fills a gap in the DDI product suite
  • Up to now, derivation fields have been filled with text strings
  • SDTL is “structured” to be read by computers
• SDTL was added to the DDI Alliance suite of standards in December 2020
Recent Activities

• C2Metadata project is over, but some work continues
• Collaboration with Whole Tale Project to map SDTL into ProvONE
  – Aiming for a journal article in Autumn 2022
• Best Paper Award for International Digital Curation Conference 2021
  – Appearing in *International Journal of Digital Curation* soon
Future Activities

• Continue to make minor adjustments based on feedback
• Consider identification for variables
• Extend SDTL to cover data created by statistical analysis commands
  – For example, regression and factor analysis
SDTL Working Group

George Alter (Lead)
J. Gager
Jeremy Iverson
Rebecca Oldroyd
Ornulf Risnes
Dan Smith
Thomas Thelen
Jim Todd
Reference Links

• *Structured Data Transformation Language (SDTL):* [https://ddialliance.org/products/sdtl/1.0](https://ddialliance.org/products/sdtl/1.0)


• C²Metadata Project: [https://c2metadata.org/](https://c2metadata.org/)

• C²Metadata Resource Page: [http://c2metadata.mtna.us/](http://c2metadata.mtna.us/)
DDI Training Group

- 13 members from different countries (2 members left, 2 joined)
- 2 new co-chairs: Alina Danciu (Sciences Po France) & Hayley Mills (CLOSER UK)
- Two sub-groups, Slide Deck Review and Training Opportunities
- Priority of the chairs: update the mandate of the group in accordance with the DDI Alliance Scientific Work Plan
Main realisations 2021/2022

- **5 webinars** between April 2021 and April 2022, in collaboration with CODATA
- A satisfaction survey conducted after every webinar
- EDDI “Training FAIR” co-organised with CODATA, which reached an even larger audience (2 half days, parallel tracks)
- **10 slide decks** added to the DDI Training Material community - hundreds of downloads since publication
- **Creation of a list of slide decks** to be completed and their priorities
- The two Zenodo Communities: DDI Training Material and the DDI Training Group descriptions were updated
- 5 meetings with potential DDI users that contacted us through the contact form on the DDI Alliance website
Workplan 2022/2023

Slide decks

- Complete finalizing/creating the remaining training materials based on the output of a Train-the-Trainer Workshop at Dagstuhl in 2018 and webinars
- Set up a guide for using the existing Training Materials

Events

- Training at IASSIST (2 sessions)
- 5 webinars to be completed
- Training at EDDI 2022: Free online workshop planned before the conference
- Training for the North American community: A workshop is being planned for November 2022

Website

- Validate template response for Training requests, pointing users to resources
- Events page to be completed, in liaison with Marketing
- Getting started page to be revised, in liaison with Technical Committee
DDI-CDI WG Report to Members 2021-2022

1 June 2022
Arofan Gregory
Chair
Goals
Per the *DDI Scientific Work Plan 2021-2022*, the primary short-term goals of the DDI-CDI WG are as follows:

• Finalize the first production version of the Cross Domain Integration specification.
• Collaborate on activities to implement and get feedback on the specification, in particular with data providers, RIs, EOSC and cross-domain case studies.
• Provide more complete mappings to other DDI specifications (Codebook and Lifecycle), and some mappings to external standards which were not included in the review package, notably to SDMX and DCAT.
• Create a modular architecture for the specification.

Looking forward to 2023, support for user implementation guidance is also highlighted.
CDI Finalization

• Covers description of structural and provenance/process metadata for a broad range of data types across domains
  • Long data
  • Wide data
  • Multi-dimensional “cube” data
  • Key-value data
  • Relational data

• The first production draft of the DDI Cross Domain Integration specification is almost complete
  • Final editing and QA now being completed
  • Release to TC for voting and publication process by end of June 2022

• Will include:
  • Canonical XMI expression of the UML model
  • XML syntax representation (W3C XML Schemas)
  • Integrated browser-based documentation for field-level and syntaxes (Model2Text)
  • High-level documentation

• RDF syntax representation soon to follow
  • Established official W3C liaison

• Methodology for Implementation Guides soon to follow
Outreach to Domains and Infrastructures, Standards

• On-going engagement with many different projects and initiatives
  • EOSC (ESS use case for integrating environmental data)
  • Helmholtz use case
  • GOSC
  • FAIR Digital Object Framework
  • INSPIRE work on integrated COVID/public health data
  • Dataverse (planned support)

• WorldFAIR Project
  • Work packages across many diverse domains
  • CDI is part of the emerging “Cross-Domain Interoperability Framework” (CDIF)

• Projects involve DDI Codebook, DDI Lifecycle
• Projects involve DCAT, SDMX/DataCube, Schema.org, PROV, IADOPT, SKOS/XKOS, and others
Modularity/Implementability

• Modular design incorporated, enhanced moving forward

• Implementation guides to cover:
  • Selection of classes
  • Syntax representations
  • Specification of relevant controlled vocabularies

• Prototypes show that a small number of classes (i.e., 12 for Interstat) can be selected for use

• Syntax representations can be easily integrated (e.g., RDF)

• Planning to package the browser documentation tooling (Model2Text) for community use to better support implementers
FY2022 Financial Report
<table>
<thead>
<tr>
<th></th>
<th>Actual FY2021</th>
<th>Forecast FY2022</th>
<th>Actual FY2022*</th>
<th>Budget FY2022</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Revenue</td>
<td>$97,500.00</td>
<td>$97,500.00</td>
<td>$97,500.00</td>
<td>$100,500.00</td>
</tr>
<tr>
<td>Expenses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff Salaries</td>
<td>$46,200.99</td>
<td>$45,000.00</td>
<td>$39,198.00</td>
<td>$45,000.00</td>
</tr>
<tr>
<td>Consultants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Expenses</td>
<td>$325.00</td>
<td>$0.00</td>
<td>$0.00</td>
<td>$20,724.00</td>
</tr>
<tr>
<td>Marketing</td>
<td>-$2,900.17</td>
<td>$1,227.00</td>
<td>$727.00</td>
<td>$5,000.00</td>
</tr>
<tr>
<td>Research Supplies &amp; Services</td>
<td>$401.58</td>
<td>$1,268.00</td>
<td>$1,268.00</td>
<td>$1,150.00</td>
</tr>
<tr>
<td>Scientific Board</td>
<td>$0.00</td>
<td>$26,611.00</td>
<td>$13,011.00</td>
<td>$56,020.00</td>
</tr>
<tr>
<td>Technical Committee</td>
<td>$0.00</td>
<td>$4,637.00</td>
<td>$4,637.00</td>
<td>$21,177.00</td>
</tr>
<tr>
<td>Training</td>
<td>$5,000.00</td>
<td>$0.00</td>
<td>$0.00</td>
<td>$3,000.00</td>
</tr>
<tr>
<td>Travel and Hosting for Alliance</td>
<td>$0.00</td>
<td>$0.00</td>
<td>$0.00</td>
<td></td>
</tr>
<tr>
<td>Transfer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Expenses</td>
<td>$49,027.40</td>
<td>$78,743.00</td>
<td>$58,841.00</td>
<td>$152,071.00</td>
</tr>
<tr>
<td>Revenue Over/(Under) Expenses</td>
<td>$48,472.60</td>
<td>$18,757.00</td>
<td>$38,659.00</td>
<td>-$51,571.00</td>
</tr>
<tr>
<td>Ending Fund Balance</td>
<td>$330,027.00</td>
<td>$348,784.00</td>
<td>$368,686.00</td>
<td>$278,456.00</td>
</tr>
</tbody>
</table>

- Reserves from NADDI: $21,076.03 $21,076.03 $21,076.03
- Reserves from general operations: $327,707.97 $347,609.97 $257,379.97

Currency in USD
*Last updated: April 30, 2022

FY2022 approved budget:
https://docs.google.com/spreadsheets/d/178vQhiB1IzrscsszQI_zxrpcNys9LW0Ch2Xcl08I6isE/edit?usp=sharing
FY2023 Budget
FY2023 budget process

- Anyone from the DDI community could submit a funding request.
- The DDI Scientific Board evaluated requests related to scientific or technical activities from the perspective of the Alliance Scientific Plan and provided feedback to the Executive Board Chair and Vice Chair.
FY2023 budget requests

- The budget for FY2023 (July 2022 - June 2023) will be finalized and approved by the Executive Board in June.
- FY2023 funding requests: https://docs.google.com/spreadsheets/d/1DxW92KQIOic0UoH4ECPynNSDP6l1a1YBf1AyzOOHRlo/edit#gid=0
- If all requests are approved:
  - Total expenses = $160,377
  - Projected revenue = $94,500
  - Expected loss = $65,877
  - Forecast reserves balance from general operations (30 June 2023) = $261,831
Financial Report

This document summarizes the overall financial position of the DDI Alliance at the close of FY2022 (July 2021 - June 2022) and includes a draft budget of all funding requests for FY2023.

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 FY2022 Budget

Appendix A provides an overview of the FY2022 budget: the Budget FY2022 column lists what was budgeted for the Alliance at the start of FY2022, the Actual FY2022 column lists all expenses that have been processed by the Alliance fiscal year-to-date, and the Forecast FY2022 column lists all expenses forecast to be processed by the Alliance at the end of FY2022.

Below are details for each of the main FY2022 budget categories, as well as a summary of the FY2022 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 FY2020 was $100,500, but actual income is $97,500. The negative variance is due to one dropped membership by the Dansk Data Arkiv (DDA).

Expenditures

Staff Salaries
- Staff salaries are expected to be equal to what was budgeted. Staff salaries cover secretariat staff and computing services at ICPSR, including 10 percent of the Executive Director, 20 percent of a marketing assistant, 5 percent of an accountant, and 6 percent of a Web developer. With the exception of the added marketing assistant, salary and data expenses have remained the same for the last 8 years even though ICPSR salaries have increased by an average of three percent each year.

General Expenses
- Expenses were zero since the DDI libraries development (jointly funded with Statistics Canada) has not begun.

Research Supplies & Services
- Expenses covered hosting fees for the DDI Registry, a hosted DNS SRV record-based resolution service for DDI agency identifiers, as well as one Zoom virtual meeting subscription and several Miro virtual workspace licenses to support the DDI-CDI virtual Dagstuhl meetings.

Marketing
- Marketing spent less than one-fifth ($727 for promotional brochures) of its allocated budget ($5,000). They expect to spend another $500 to sponsor the 2nd IASSIST Africa Regional Workshop scheduled to take place October 2022 in Nigeria.

Scientific Board
- The Scientific Board spent funds on training activities, Dagstuhl hosting for a DDI-CDI face-to-face meeting, and for DDI URN registration at IANA. Anticipated expenses before the end of the fiscal year include videos about DDI for social media sharing, as well as for RDF syntax representation mapping and implementation.

Technical Committee
- The Technical Committee spent $4,637 on DDI Registry Resolution Enhancements.

Travel and Hosting
- The Alliance did not spend any funds on travel or hosting.
Summary

- The overall FY2022 income is forecast to exceed expenses by $18,757.
- The overall FY2022 forecast expenses are slightly more than half ($78,743) of what was budgeted ($152,071) for the fiscal year.
- The fund balance for the Alliance is forecast to be $348,784 at the end of FY2022. Of this, $21,076.03 is committed to North American DDI (NADDI) reserves. The uncommitted fund balance (i.e., funds that are unencumbered by previously allocated expenses) is forecast to be $327,707.97.

Overview of FY2023 Budget Requests

In March, each Alliance committee or working group was asked to submit a funding request for FY2023 (July 2022 - June 2023). Additionally, a message was sent to the entire DDI community soliciting funding requests; anyone could submit a funding request. All requests were compiled into the FY23 budget planning document.

The DDI Scientific Board evaluated requests related to scientific or technical activities from the perspective of the Alliance Scientific Plan. They provided feedback about requests related to scientific or technical activities to the Executive Board Chair and Vice Chair.

All FY23 community funding requests total $160,377. If all are funded, the overall expected loss would be $65,877 (after projected revenue of $97,500). If all budget requests are funded, the forecast reserves balance from general operations would be $261,831 at the end of FY23.
<table>
<thead>
<tr>
<th></th>
<th>Actual FY2021</th>
<th>Forecast FY2022</th>
<th>Actual FY2022*</th>
<th>Budget FY2022</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Revenue</strong></td>
<td>$97,500.00</td>
<td>$97,500.00</td>
<td>$97,500.00</td>
<td>$100,500.00</td>
</tr>
<tr>
<td><strong>Expenses</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff Salaries</td>
<td>$46,200.99</td>
<td>$45,000.00</td>
<td>$39,198.00</td>
<td>$45,000.00</td>
</tr>
<tr>
<td>Consultants</td>
<td>$0.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General Expenses</td>
<td>$325.00</td>
<td>$0.00</td>
<td>$0.00</td>
<td>$20,724.00</td>
</tr>
<tr>
<td>Marketing</td>
<td>-$2,900.17</td>
<td>$1,227.00</td>
<td>$727.00</td>
<td>$5,000.00</td>
</tr>
<tr>
<td>Research Supplies &amp; Services</td>
<td>$401.58</td>
<td>$1,268.00</td>
<td>$1,268.00</td>
<td>$1,150.00</td>
</tr>
<tr>
<td>Scientific Board</td>
<td>$0.00</td>
<td>$26,611.00</td>
<td>$13,011.00</td>
<td>$56,020.00</td>
</tr>
<tr>
<td>Technical Committee</td>
<td>$0.00</td>
<td>$4,637.00</td>
<td>$4,637.00</td>
<td>$21,177.00</td>
</tr>
<tr>
<td>Training</td>
<td>$5,000.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Travel and Hosting for Alliance</td>
<td>$0.00</td>
<td>$0.00</td>
<td>$0.00</td>
<td>$3,000.00</td>
</tr>
<tr>
<td>Transfer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total Expenses</strong></td>
<td>$49,027.40</td>
<td>$78,743.00</td>
<td>$58,841.00</td>
<td>$152,071.00</td>
</tr>
<tr>
<td><strong>Revenue Over/(Under) Expenses</strong></td>
<td>$48,472.60</td>
<td>$18,757.00</td>
<td>$38,659.00</td>
<td>-$51,571.00</td>
</tr>
<tr>
<td><strong>Ending Fund Balance</strong></td>
<td>$330,027.00</td>
<td>$348,784.00</td>
<td>$368,686.00</td>
<td>$278,456.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reserves from NADDI</td>
<td></td>
<td>$21,076.03</td>
<td>$21,076.03</td>
<td>$21,076.03</td>
</tr>
<tr>
<td>Reserves from general operations</td>
<td>$327,707.97</td>
<td>$347,609.97</td>
<td>$257,379.97</td>
<td></td>
</tr>
</tbody>
</table>

Currency in USD

*Last updated: April 30, 2022
These are the DDI funding requests for FY2023 (July 1, 2022 through June 30, 2023), including requests submitted by Alliance members as of May 3, 2022. Each Alliance committee or group was asked to submit a funding request. Additionally, a message was sent to the entire DDI community soliciting funding requests. The DDI Scientific Board evaluated requests related to scientific or technical activities from the perspective of the Alliance Scientific Plan. They provided feedback about requests related to scientific or technical activities to the Executive Board. A link to the full funding request description is also provided.

<table>
<thead>
<tr>
<th>ITEMNO</th>
<th>Reviewed by Scientific Board?</th>
<th>Source</th>
<th>Activity</th>
<th>Requested Total</th>
<th>Approved Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recurring Expenses</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>No</td>
<td>Secretariat staff salaries</td>
<td></td>
<td>$45,000</td>
<td>$45,000</td>
</tr>
<tr>
<td>2</td>
<td>No</td>
<td>DDI Registry web hosting</td>
<td></td>
<td>$900</td>
<td>$900</td>
</tr>
<tr>
<td>3</td>
<td>No</td>
<td>Zoom virtual meetings</td>
<td></td>
<td>$150</td>
<td>$150</td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>Wire fees (estimate)</td>
<td></td>
<td>$100</td>
<td>$100</td>
</tr>
<tr>
<td>5</td>
<td>No</td>
<td>Annual meeting hosting</td>
<td></td>
<td>$1,000</td>
<td>$1,000</td>
</tr>
<tr>
<td>6</td>
<td>No</td>
<td>Meeting attendance (e.g., UNECE)</td>
<td></td>
<td>$2,000</td>
<td>$2,000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>$45,000</strong></td>
<td><strong>$45,000</strong></td>
</tr>
</tbody>
</table>

| **Funding Requests** | | | | | |
| **Committee/WG Requests** | | | | | |
| 7 | Yes | Scientific Board Meeting | | $7,000 | $7,000 |
| 8 | Yes | Hackathon | | $10,000 | $10,000 |
| 9 | Yes | CDI Workshop: Hotzone Week on FAIR | | $5,551 | $5,551 |
| 10 | Yes | Development and Packaging of the CDI WG Production Tools | | $9,600 | $9,600 |
| 11 | Yes | Implementation Languages Meeting | | $17,300 | $17,300 |
| 12 | Yes | Technical Committee Meeting | | $16,450 | $16,450 |
| **Total** | | | | **$51,750** | **$51,750** |
| **DDI-CDI Working Group** | | | | | |
| 13 | Yes | Funding for workshops (travel and fees) | | $10,000 | $10,000 |
| 14 | Yes | Discount / waiver workshop fees (for DDI members) | | $1,500 | $1,500 |
| 15 | Yes | Webinars | | $7,500 | $7,500 |
| **Total** | | | | **$19,000** | **$19,000** |
| **Technical Committee** | | | | | |
| 16 | Yes | Dagstuhl Organiser's Group | | “Interoperability for Cross-Domain Research: Machine-Actionability & Scalability” workshop | $6,912 | $6,912 |
| 17 | Yes | Statistics Canada | | DDI libraries (carry-over request from FY22) | $19,414 | $19,414 |
| **Total** | | | | **$26,326** | **$26,326** |

| **Community Requests** | | | | | |
| 16 | Yes | Dagstuhl Organiser's Group | | “Interoperability for Cross-Domain Research: Machine-Actionability & Scalability” workshop | $6,912 | $6,912 |
| 17 | Yes | Statistics Canada | | DDI libraries (carry-over request from FY22) | $19,414 | $19,414 |
| **Total** | | | | **$26,326** | **$26,326** |

**Total Expenses** | **$160,377** | **$160,377**

**ESTIMATED REVENUE** | **$94,500** | **$94,500**
These are the DDI funding requests for FY2023 (July 1, 2022 through June 30, 2023), including requests submitted by Alliance members as of May 3, 2022. Each Alliance committee or group was asked to submit a funding request. Additionally, a message was sent to the entire DDI community soliciting funding requests. The DDI Scientific Board evaluated requests related to scientific or technical activities from the perspective of the Alliance Scientific Plan. They provided feedback about requests related to scientific or technical activities to the Executive Board. A link to the full funding request description is also provided.

<table>
<thead>
<tr>
<th>ITEMNO</th>
<th>Reviewed by Scientific Board?</th>
<th>Source</th>
<th>Activity</th>
<th>Requested Total</th>
<th>Approved Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>SURPLUS / (LOSS) – if all recurring expenses and funding requests are approved</td>
<td>-$65,877</td>
<td>-$65,877</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FORECAST FUND BALANCE (30 JUNE 2022)</td>
<td>$348,784</td>
<td>$348,784</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PLUS SURPLUS/LOSS</td>
<td>-$65,877</td>
<td>-$65,877</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FORECAST FUND BALANCE (30 JUNE 2023)</td>
<td>$282,907</td>
<td>$282,907</td>
</tr>
</tbody>
</table>

*Currency in USD.

*Last updated 31 May 2022.
Overview

This budget is a roll-over from two previous years and has been approved twice in the past. Originally, this budget was supposed to finance a physical meeting of the temporary working group for restructuring the Scientific Board in 2020, was postponed due to COVID and finally was planned for a physical meeting of the newly formed Scientific Board in 2021 which again was postponed due to COVID. Currently, the Scientific Board in its new format has existed since the beginning of 2021 and the members never met personally. Also, there are several topics (e.g. nomination of external advisors) which would need longer discussion than can be provided in the monthly scheduled 1.5h virtual meetings. Thus the SB urgently needs 2-3 days at one location to facilitate deeper discussions. This meeting is scheduled for September 5th and 6th in Chur, Switzerland. To be cost efficient by saving travel costs the meeting has been connected to the Dagstuhl workshops which are planned for the previous week.

Budget

Like in the requests before we ask for 7000 USD for accommodation and travel costs.
DDI Alliance Budget Request for Fiscal year 2022/2023 : Hackathon

Overview

The DDI Developers Group has been dormant since 2014 but the crucial expertise from software development is missing more and more as product development continues. In the last years the DDI Alliance lost its connection to this community as people changed into different roles without having successors and also as the former informal group was not stable enough to exist. After this long period of time it seems this community within the Alliance has to be re-developed from scratch. The first step is to nudge DDI members to nominate Scientific Representatives and Technical Contacts. From these contacts potential members for a DDI Developers Group could be sourced. Furthermore, to make collaboration more interesting we suggest holding a 2-3 days hackathon to develop prototypical solutions in DDI contexts where there have not been any efforts yet. This hackathon can be used to identify core roles for a Developers’ Group. This hackathon also should be connected to another DDI event (e.g. NADDI, EDDI, Dagstuhl) to provide easier internal funding from the members (currently DDI members are not used anymore to bring technical personnel to conferences).

Budget

Financing a standalone full hackathon with international travel costs would be out of the range for the current budget (30.000 - 40.000 USD) therefore we suggest combining the hackathon with an existing conference. This way some of the costs could partly be covered by member organizations. Nevertheless, we apply for a budget of 10.000 USD to support potential participants who might not be covered and for additional costs like catering (at a hackathon food and drink expenses are usually covered while participants are working).
DDI Alliance Budget Request for the Financial Year 2022/2023
CDI Working Group

Overview
This document describes two separate budget requests from the CDI Working Group. The first is for travel budget to hold an open workshop in the margins of the “Hotzone Week on FAIR” taking place during October, 2022 in Leiden in the Netherlands. The other budget request is related to preparing and packaging the production environment developed by the CDI Working Group to develop syntax representations for use by implementers in the community.

The total being requested is:

DDI-CDI Workshop: Hotzone Week on FAIR: $5,551.20
M2T DDI-CDI Tools: $9,600.00

Total Budget Requested: $15,151.20

I. CDI Workshop: Hotzone Week on FAIR
DDI-CDI is now being recognized as a specification which could help the implementation of the FAIR principles as a key part of a suite of related, non-domain-specific standards (DCAT, Schema.org, PROV, I-ADOPT, etc.). This package is being described as the “Cross Domain Interoperability Framework” (CDIF) – it continues to be a topic of interest at various conferences and workshops, notably at the upcoming Dagstuhl event in 2022.

The CDI WG has established the need to align their work not only with other DDI specifications, but also other external standards. For the purposes of supporting FAIR implementation, alignment with the FAIR Digital Object Framework (FDOF) is critical. This framework will be the focus for the Hotzone Week on FAIR (https://www.go-fair.org/events/hotzone-week-on-fair-at-leiden-2022-european-city-of-science/).

The requested budget would support the travel of up to four people from the CDI WG to host such an event (two from North America and two from Europe).

The specific goals would be:

1. Promote CDI as a specification for supporting the use of FAIR Digital Objects, in line with the emerging CDIF recommendations

2. Gather input on alignment with other specifications related to the FDOF, including FIPs, FAIR Data Points, etc.

3. Make connections and potentially recruit people with expertise in the FDOF and related specifications to participate in supporting the implementation and further development of DDI-CDI.
This is seen as an event covering a full day and possibly two, with some presentation, but primarily focused on working sessions on specific topics related to this theme.

Estimates of costs are given below:

<table>
<thead>
<tr>
<th>Item</th>
<th>Cost</th>
<th>Quantity</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hotel in Leiden</td>
<td>€150.00</td>
<td>2</td>
<td>€300.00</td>
</tr>
<tr>
<td>Food cost per day in Leiden</td>
<td>€45.00</td>
<td>3</td>
<td>€135.00</td>
</tr>
<tr>
<td>Hotel at Frankfurt airport</td>
<td>€150.00</td>
<td>1</td>
<td>€150.00</td>
</tr>
<tr>
<td>Flight in Europe</td>
<td>€400.00</td>
<td>1</td>
<td>€400.00</td>
</tr>
<tr>
<td>Flight transatlantic</td>
<td>€900.00</td>
<td>1</td>
<td>€900.00</td>
</tr>
<tr>
<td>Local transport at origin and destination</td>
<td>€200.00</td>
<td>1</td>
<td>€200.00</td>
</tr>
</tbody>
</table>

Total cost per person from Europe for Leiden meeting: €1,035.00

Total cost per person from North America for Leiden meeting: €1,535.00

Total cost of Leiden meeting: €2,070.00

Using a conversion rate of 1.08*, totals are as follows:

Two European participants: $2,235.60

Two North American participants: $3,315.60

Total: $5,551.20

* Currency conversion rate as of 15 April 2022: https://www.xe.com/currencytables/?from=USD&date=2022-04-15#table-section

II. Development and Packaging of the CDI WG Production Tools

The DDI-CDI specification is primarily a model, which will be implemented using either the syntax representations provided (XML, RDF), community-specific variations of those, or other syntaxes (e.g., Python, JSON, Java, etc.). Such syntax representations require not only that the selected fields and syntax be documented, but that the use of such fields in relationship to the standard model itself be described. In many cases, classes and properties in the model may be supported in specific ways which will require that the technical details to support machine-to-machine interoperability must be described for community implementers.

The CDI WG has developed a solution for presenting the full specification in this way, along with the “generic” syntax representations. The resulting browser-based interface integrates field-level
documentation for the model and syntax expressions and examples in a useful fashion, as shown in the examples below.

The model presentation (for Category):

The XML Syntax representation tab:
This presentation was developed by configuring the Model to text (M2T) open-source application – based on Object Modeling Group standards – for use with the DDI-CDI model (see https://projects.eclipse.org/projects/modeling.m2t). In order to make it easier for community implementers of DDI-CDI to adopt the specification, it is felt that providing a solid basis for modification, using open-source tooling, is ideal. M2T is not a complex tool to use, and our experience with it suggests that it would be easy to learn for most developers in the community.

The requested budget would be for the adaptation of an internal working solution to a tool suited for use by community implementers, including packaging of the tool and documentation of how to work with the provided DDI-CDI configuration within the M2T development environment.

This would cover both the modification of published “generic” XML and RDF syntaxes and the development of new syntax representations. (Any target syntax which can be described as text-based code is a potential candidate for this solution.)

It is estimated that this work would require 24 person-days, with a cost of $400.00/day.

The total amount requested is $9,600.00.
Implementation languages across the DDI Suite: Cross Working Group In-Person Meeting

Proposal:
Submitted by Wendy Thomas, Chair, Technical Committee, 2022-04-29

The Technical Committee with the support of the DDI-CDI Working Group propose an in-person meeting of 3-5 days to accomplish the following goals:

Identification and use of implementation languages in the DDI Suite of products:
- Identify priority implementation languages for DDI products (RDF, JSON, UML, XML, etc.)
- Identify style options for implementation languages
- Mappings to produce syntax representations - Moving from conceptual models to serialization
- What aspects of implementation should be consistent
  - Document options, decisions, and reasoning
- Provide guidance for variation from agreed model
  - based on applied use of product
  - what needs to be noted and how (need a consistent expression of exceptions and reasons)

Outputs:
- Documentation of implementation language decisions
- Guidelines for implementing languages in various products
- Plan for providing and testing multiple implementation languages for current products

Background:
The DDI Suite currently expresses its individual products in a number of different implementation languages. Each product uses one or two of these languages and many are interested in expanding to multiple expressions. Current usage includes XML schema, UML XMI, RDF, and JSON. The work of the Technical Committee moving the DDI Lifecycle to the COGS production tool will allow us to store content in a CSV format and export to multiple languages. DDI CDI is also working on adding implementation languages beginning with RDF. Rather than work independently the Technical Committee and the DDI-CDI Working Group believe that it would be beneficial to explore options and provide guidelines to the DDI maintenance and development groups on the use of various features of these languages. We need to determine if and where we need uniformity and how to inform users of differences in the
implementation of individual products in different languages. This work addresses the approved Scientific Board Work Plan Goals noted in the appendix.

Possible Participants:
- Pierre-Antoine Champin (W3C contact and JSON editor)
- Franck Cotton
- Darren Bell
- Dan Smith
- Jeremy Iverson
- Jon Johnson
- Flavio Rizzolo
- Oliver Hopt
- Johan Fihn
- Wendy Thomas
- Olof Olsson
- Arofan Gregory
- Joachim Wackerow
- Hilde Orten

The target number of participants is 8-10

Timing:
A European location is planned as most identified participants are located there. We anticipate the meeting taking place in late Fall 2022. If possible, it may coincide with EDDI 2022 to be held in Paris.

Cost estimation in USD:

<table>
<thead>
<tr>
<th>Costs</th>
<th>Per Person Local</th>
<th>Per Person Europe</th>
<th>Per Person North American</th>
<th>Total for 5 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation</td>
<td>500</td>
<td>1500</td>
<td>8000</td>
<td></td>
</tr>
<tr>
<td>Lodging</td>
<td>170</td>
<td>170</td>
<td>6800</td>
<td></td>
</tr>
<tr>
<td>Food</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>2500</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td></td>
<td>17300</td>
</tr>
</tbody>
</table>

Costs calculated for 10 people including 2 local, 4 European, and 4 North American
APPENDIX:

Work Plan Goal:

● Lifecycle 3.4 – Complete the move of Lifecycle to the COGS modeling base, testing of input and output for coverage and consistency. This will include testing multiple outputs (XML schema, RDF, JSON, UMI to begin with). …

● Align different product implementations over time - how does that work and what does it look like

● Defining roles of individual products
  o How products work together
  o How advances/changes in one product affect other product development
  o Role of products needs to be clear - use case driven rather than content coverage
In-Person Meeting of the Technical Committee

Submitted by Wendy Thomas, Chair, Technical Committee, 2022-04-29
Updated on 2022-05-10

Proposal:

The purpose of this meeting is to finalize the use of COGS as a means of automating the production process of DDI Lifecycle and determine the feasibility of its use for other products. This meeting was originally funded for the 2021/22 fiscal year and was postponed due to COVID restrictions.

Goals:
A virtual meeting is being held the first week of May 2022 to verify the consistency and validity of the input process from the DDI Lifecycle Version 3.3 into the COGS CVS and related files. The goals of the proposed workshop will focus on the following:

Output from COGS
- Finalize current DDI-L in COGS - output is complete and correct
  - Input should have been completed in May virtual meeting
  - Are the serializations as currently written in COGS working following the rules for flattening etc
  - Auto-generated from content model of COGS - same content with variation of representation
  - The goal for DDI-L 3.4 is the same content of 3.3 with revised format including rules for flattening and serialization
- Feasibility of COGS for CDI
  - Testing input and output functions
  - Identify issues that need to be resolved
- Review the current use of COGS for SDTL to identify any procedures and protocols that can be extended to other products and to learn from their experience
- Output options for Codebook
  - Determine the feasibility of using COGS for managing Codebook
  - What would the output rules for Codebook look like
  - What might the rules for the expression of XML in order address the needs of this user community

Outputs:
- Prepare a version 3.4 that matches 3.3 in terms of content but provides an alternate structure
- Set this up for technical review by implementers
- Document progress of Codebook and CDI feasibility in COGS
- Fully document the rules guiding the production of multiple implementation languages for DDI Lifecycle in preparation for a future meeting on implementation language comparability between products in the DDI suite

Participants:
Wendy Thomas
Jon Johnson
Flavio Rizzolo
Johan Fihn Marberg
Darren Bell
Oliver Hopt
Jeremy Iverson
Dan Smith
Ingo Barkow - Scientific Board contact to TC

Timing:
Summer (July/Aug) 2022

Location:
Minneapolis (3 members live here - ISRDI will provide space)

Cost:

<table>
<thead>
<tr>
<th>Costs</th>
<th>Per Person Local/day</th>
<th>Per Person Europe/day</th>
<th>Per Person North American/day</th>
<th>Total for 5 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation</td>
<td></td>
<td>1800</td>
<td>700</td>
<td>9700</td>
</tr>
<tr>
<td>Lodging</td>
<td></td>
<td>150</td>
<td>150</td>
<td>3750</td>
</tr>
<tr>
<td>Food</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>2000</td>
</tr>
<tr>
<td>TOTAL</td>
<td></td>
<td></td>
<td></td>
<td>16450</td>
</tr>
</tbody>
</table>

Calculated for 5 days with 3 local and 5 European and 1 North American

Appendix:

Work Plan Goal:
- Lifecycle 3.4 – Complete the move of Lifecycle to the COGS modeling base, testing of input and output for coverage and consistency. This will include testing multiple outputs (XML schema, RDF, JSON, UMI to begin with). CURRENT STATUS: Evaluated
input issues and are currently correcting that script. Output scripts accurately reflect stored content.

- COGS as a processing tool has advantages for Lifecycle as noted when agreed on in 2019 including; auto generation of output structures based on translation rules from object descriptions stored in structured CSV file; ability to add new output formats as needed; ability to generate output and test new content as created;
- Review implications of multiple outputs on modeling, incorporating discussions and approaches from the Moving Forward work where appropriate; this may require minor remodeling of some choice or sequence usage
**DDI Training Working Group**  
**Budget Request for FY2022 (July 2022 through June 2023)**

Submitted by: Alina Danciu (co-Chair) and Hayley Mills (co-Chair)  
Submitted on: April 29, 2022

**Budget Request**

The Training Working Group requests $17,000 for FY2023.

**FY2023 (June 1, 2022 - June 30, 2023)**

<table>
<thead>
<tr>
<th>Training Group Activity</th>
<th>Purpose / Goal</th>
<th>Audience</th>
<th>Cost (USD $)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Funding for workshops (travels and fees)</td>
<td>DDI will be promoted by submitting workshops to about 5 conferences. Travel costs and conference fees are covered for the instructor.</td>
<td>new users</td>
<td>10,000</td>
</tr>
<tr>
<td>Discount / waiver workshop fees (for DDI members)</td>
<td>We will offer a workshop fee waiver for DDI members and a conference fee discount for instructors at two conferences (EDDI and IASSIST).</td>
<td>new and advanced users</td>
<td>1,500</td>
</tr>
<tr>
<td>Webinars</td>
<td>A series of 6 webinars on DDI to organise, with the help of a consultant, like we did for the 2021/2022 DDI Alliance/CODATA series. A thorough description of this request <a href="#">here</a>.</td>
<td>new and advanced users</td>
<td>7,500 + In-kind contributions</td>
</tr>
<tr>
<td>Web page update</td>
<td>Members of the Training Group continue to update and make changes to training-related content on the DDI website.</td>
<td>new and advanced users, DDI Trainers</td>
<td>In-kind contributions</td>
</tr>
<tr>
<td>Translation</td>
<td>Start with the translation of the basic introductory slide decks.</td>
<td>new and advanced users</td>
<td>In-kind contributions</td>
</tr>
<tr>
<td>Training Material update</td>
<td>Members of the Training Group continue to produce Training Material (including exercises) on Zenodo and the DDI website (with the help of the DDI assistant).</td>
<td>new and advanced users, DDI Trainers</td>
<td>In-kind contributions</td>
</tr>
<tr>
<td>Teaching in webinars or at conferences</td>
<td>Members of the Training Group engage in teaching activities to</td>
<td>new and advanced users</td>
<td>In-kind contributions</td>
</tr>
<tr>
<td></td>
<td>increase DDI knowledge in the research community.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------------------------------------</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td>19,000</td>
<td></td>
</tr>
</tbody>
</table>

**Active Members of the DDI Training Working Group**

- Alina Danciu (Co-chair)
- Christophe Dzikowski
- Adrian Dusa
- Dan Gillman
- Arofan Gregory
- Kaia Kulla
- Kathryn Lavender
- Jared Lyle
- Lucie Marie
- Geneviève Michaud
- Hayley Mills (Co-chair)
- Laura Molloy
- Hilde Orten
- Flavio Rizzolo (Scientific Board contact)
DDI Alliance Budget Request for Fiscal year 2022/2023: DDI Training Webinar Series

Overview
For two years, the Training Opportunities sub-group of the DDI Training WG has been organizing a program of bi-monthly hour-long webinars on topics related to DDI. These are aimed at a general audience, and are hosted by CODATA, using both the normal DDI Alliance distribution channels and also the RDM lists maintained by CODATA to reach a broad potential audience. There have been several hundred attendees at these events – they typically draw between 30 and 70 people. Further, these webinars lead to the development in 2021 of the EDDI “Training FAIR” which reached an even larger audience.

These events have been organized with the help of a paid consultant. This budget proposal is requesting funding at the same level as for the past two years to support the development of materials. This has been done through a series of “mini-sprints” involving a small group of interested individuals who produce and review slides for the event. The materials are then provided to the Slide Review sub-group for modification and possible inclusion in the Training Materials resource.

Budget and Details
A series of six topics has been the target for the first two cycles, and this would remain the target for 2022/2023. The work of preparing draft slide decks and organizing and running the events requires approximately 3 weeks of person time for the consultant. This effort is in addition to the proviso that the consultant will also be available to help in the presentation and running of the events as needed.

To support this work, we are requesting a budget of $7500.00.
DDI Alliance Budget Request for the Financial Year 2022/2023

Schloss Dagstuhl Workshop Organizers: Arofan Gregory, Hilde Orten, Joachim Wackerow, Simon Cox, Simon Hodson, and Steve McEachern,

Overview

The budget requested would fund the travel of some participants at the upcoming workshop at Schloss Dagstuhl in Wadern Germany, from 29 August through 2 September 2022, titled “Interoperability for Cross-Domain Research: Machine-Actionability & Scalability”. While most participants will be funded by their own institutions or project budgets, some invitees may not be able to arrange for support of this type. The costs of supporting such participants will be shared with CODATA, which is also sponsoring the event. The request given here is based on an estimate of costs split between the two organizations – this portion is what the DDI Alliance would be asked to provide. The separate CODATA contribution will be equal or greater to the sum requested here.

The total being requested is $6,912.00.

Topics and Relationship to DDI Alliance Work Products and Goals

The topic will be cross-domain interoperability of data and metadata, with a focus on a coherent set of standards and models to make this practically possible. Among these, DDI-CDI is used for cross-domain data description, but other DDI work products including DDI Codebook, DD Lifecycle, SDTL, and XKOS are also germane to the work. Many other significant metadata specifications are also involved (DCAT, Schema.org, I-ADOPT, PROV, SKOS, etc.).

The workshop will be organized around a set of real-world cross-domain case studies, but will also have a modelling and technical component. Use cases will be coordinated with those being addressed by projects and initiatives related to the implementation of FAIR data and infrastructure (EOSC Interoperability Framework, WorldFAIR, GOSC, etc.). Representatives of many different organizations and domains will be invited to the workshop both for their technical expertise and for their domain knowledge.

The activities and focus of this workshop align with the goals of the DDI Alliance as stated in their current strategic plan:

From “II. Reachable Short-Term Goals for 2021 and 2022”:

7.) Promote the role of DDI in external projects like e.g. European Open Science Cloud (EOSC) and FAIR.

11.) Support and facilitate the following action items provided by the different working groups:

   a. CDI

      • Collaborate on activities to implement and get feedback on the specification, in particular with data providers, RIs, EOSC and cross-domain case studies.
From “IV. Long Term Vision for DDI Products and Processes”:

4. The Scientific Board will promote interoperability and collaboration with other metadata standards

The workshop will produce guidance on a coordinated set of core metadata standards (the “Cross-Domain Interoperability Framework” - CDIF) and how this can best support scalable FAIR implementation. These standards will support the development of technical services for both the social, behavioral, and economic sciences, public health, and official statistics, and the sharing of data and metadata among and between other domains.

The DDI work products will be a significant part of CDIF and the recommendations/guidance needed to implement it within the DDI community and beyond.

Budget

The total budget request is for $6,912.00. This will support travel and accommodation costs for three participants based in Europe, and for two participants coming from North America. (The cost of flights is significant, so these categories are estimated separately.)

Our request is based on the following estimates:

<table>
<thead>
<tr>
<th>Description</th>
<th>European Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accommodation with full board in Dagstuhl</td>
<td>€70.00</td>
</tr>
<tr>
<td>Hotel at Frankfurt airport</td>
<td>€150.00</td>
</tr>
<tr>
<td>Flight in Europe</td>
<td>€400.00</td>
</tr>
<tr>
<td>Flight transatlantic</td>
<td>€900.00</td>
</tr>
<tr>
<td>Local transport at origin and destination</td>
<td>€200.00</td>
</tr>
<tr>
<td>Total cost per person from Europe for Dagstuhl meeting</td>
<td>€1,020.00</td>
</tr>
<tr>
<td>Total cost per person from North America for Dagstuhl meeting</td>
<td>€1,670.00</td>
</tr>
</tbody>
</table>

With a conversion rate of 1.08 USD to 1 EUR*, the totals are as follows:

Per person (Europe-based): $1,101.60

Per person (North America-based): $1,803.60

For three Europe-based participants, and two North American participants, the total is:

3 European: $3,304.80

2 North American: $3,607.20

Total: $6,912.00

Please note that these estimates account for the subsidised food and accommodation rates at Schloss Dagstuhl, made possible through the support of that institution.
* Currency conversion rate as of 15 April 2022:
  https://www.xe.com/currencytables/?from=USD&date=2022-04-15#table-section
A vision for DDI in a research infrastructure

The DDI suite of products consists of multiple specifications in the form of a UML model representation, i.e. XMI, and syntax representations, e.g. XML Schema, JSON-LD, RDF/OWL, etc. together with documentation in various forms. The specifications of the DDI products follow different architectures: DDI Lifecycle, for instance, is XML Schema-driven, i.e. all syntax representations are derived from XML Schema constructs, which functions as the de-facto model; DDI-CDI, in contrast, follows a model driven architecture approach\(^1\), in which a conceptual, platform-independent model (PIM) captures the high-level entity-relationship specification and a number of platform-specific models (PSM) describe the intricacies of the different syntax representations. Both specifications can describe a data layer, consisting of various types of data, from very structured to NoSQL and streams, together with the concepts it represents and the various processes involved in data production, integration and sharing.

In order to produce the building blocks of a future global research infrastructure, we need to make the DDI products into an implementation reality. To that end, we envision a rich framework and ecosystem of reusable and shareable libraries, micro-services and tools, all built around a community of vendors and open source developers that can provide a marketplace of evolving components, services and protocols, and easy integration with other specifications and standards, most notably SDMX, DCAT and PROV. This extended FAIR ecosystem will then be leveraged to implement data production solutions and advanced analytics, including Big Data and machine learning.

This vision addresses mainly two DDI Alliance strategic actions: (1) high-level goals and (3) Improvement of interoperable and distributed DDI infrastructure for use and reuse of DDI resources. It will also facilitate the seamless integration of content from existing and future DDI registries and repositories [as per action (4) Registries/repositories] into a global data production infrastructure.

The next diagram shows an overview of the implementation stack.

Such an ecosystem can be initially enabled by a rich DDI libraries layer (center). These libraries should span both lifecycle and CDI. This project will tackle the definition of a generic framework for both DDI versions and will focus on the development of libraries for a meaningful fragment of CDI.

\(^1\) https://www.omg.org/mda/
**DDI libraries development**

We need to provide libraries in a variety of languages, e.g. R, Python, Java. These libraries will map the syntax representations of DDI to constructs in the respective languages. These high-level languages share a lower-level development language, which is C. The idea is to create a common, generic C/C++ library that deals with the serialization and deserialization of DDI objects and then creates the specific objects and methods in the respective higher-level languages. This way we can use the same code base to handle the common functionality, e.g. reading and writing syntax representations, on top of which we’ll have multiple code packages for the language-specific functionality, e.g. creating and maintaining language constructs.

In order to get there, we first need a clear design of the necessary language constructs. These constructs consist basically of classes and methods. They need to cover not just the DDI model but also the linkage to the data itself – e.g. integrating data and metadata in a data frame. In addition, they need to provide higher-lever methods capable of manipulating composite objects not just individual ones – e.g. some entities, like classifications, are composite objects spanning multiple classes that usually need to be managed together.

To produce such a design, we are proposing to use a kind of domain driven design\(^2\) approach in which the main set of classes and methods are defined using one specific prototyping language, e.g. Python. This requires domain experts, i.e. DDI modelers, working closely together with language experts, e.g. Python programmers. The result of this exercise is a prototype implementation that can then be used by a C/C++ developer to implement the DDI libraries.

![Diagram](https://example.com/diagram.png)

The object layer domain driven prototype provides higher-level functionality suitable for defining library requirements. However, the object model does not address data serialization and deserialization, storage and interoperability across languages. We propose prototyping a data interface layer, in addition to the several model specifications. The model specifications allows for model validation, and the data interface

---

layer provides persistency and conversion to/from the object model. One approach for the data interface layer is to adopt a data format that implements an Interface Definition Language (IDL). IDLs allow the specification of messages in a language-and-platform-neutral way through defined schemas. Out of the schemas, data serialization and deserialization classes for a variety of programming languages can be generated. The generated classes guarantee data interoperability across platforms. In addition to the IDL, an XML data object model (DOM) can serve as a data interface layer for the XML schema access. Therefore, we provide multiple model specifications for syntax access and multiple data interface layers for data access. Persisted data can be validated using the higher-level object model, and independently validation methods can be developed for the model specifications. Examples of IDLs include Google’s Protocol Buffer and Flatbuffer, Apache Thrift, Apache Avro.
Discussion Questions for Priorities and Feedback Breakout Sessions

Member input is always welcome, and we want to take an opportunity to have “breakout discussion sessions” to give DDI members a chance to share feedback, discuss concerns, and offer input of any kind.

We have allocated 15 minutes to these Zoom breakout sessions and will randomly place attendees into groups. Each group should appoint one member of their breakout session to report back to the larger meeting at the conclusion of the breakout sessions. Please choose one of the following questions framed in relation to the Alliance priorities outlined in the current Strategic Plan (or use your own questions):

- How can we find and reach DDI users outside our known community? (Priority Area One, Strategic Action 1)
- In what more ways can we support current DDI users -- beyond what is currently offered? (Priority Area One, Strategic Action 2)
- Which global research data organizations should we be engaging with, and in what ways? (Priority Area One, Strategic Action 3)
- As key Alliance members are retiring or leaving organizations, how can we support a generational renewal to retain and recruit community members? (Priority Area Two, Strategic Action 1)

Breakout Session Notes

Group 1
- More transparency and communication about what is happening from the Executive Board.
- More communication between Scientific Board and Executive Board.
- Make people aware of information posted on the web site.
- Scientific Board reorganization has resulted in very exciting work.

Group 2
- Nice to get to know people in the breakout room.
- Discussed “How can we find and reach DDI users outside our known community?” Decided it would be good to spread the word with our own communities about DDI.
- Is there a good one-page overview of DDI we could use?
- Web site structure is great, but could we break out the information into smaller pieces?

Group 3
- Discussed tools.
- Follow-up suggestion: convene a meeting to think about how we have better tooling around DDI. That might mean hackathon type things. It might also mean pulling together tools development work around the FAIR agenda, for instance.
- A lot of funding is going to FAIR now in the international community. There isn’t the end-to-end coverage, like NESSTAR. If we could convene a group to make recommendations, that meeting could help put together a proposal to get FAIR funding.
Group 4
- Discussed getting new users. Contributors come from new user tools.
- Discussed how people got into DDI -- either by being forced to write documentation or coming into an organization actively using the standard.
- Where to find people? Obvious conferences like EDDI. Engaging users actively using it might bring new people to join.
- Where to find new users? IASSIST? Other forums?

Group 5
- Discussed generational renewal.
- In the interim, this is becoming more well-managed. For example, the Training working group has brought in new people. The Scientific Board is also a success to bring in new people and spread the load more than in the past.
- Further effort, we need to provide better support for the administrative aspects of the Alliance. Thinking about allowing people to focus on what we can contribute scientifically or intellectually (thinking especially about the efforts Wendy has put into the Technical Committee).

Group 6
- Spent time connecting with each other and discovering individual perspectives. Nothing new to add to the existing list of topics already discussed.