

# Report from Week One of Dagstuhl Sprint

October 16-20, 2017

## Prototype Description Group

Jon Johnson, Wendy Thomas, Arofan Gregory

This smaller group produced a description of the potential scope, schedule, and set of deliverables for the upcoming release in June 2018 (called “the Prototype” as of this writing, although the name may subsequently be changed).

[Explication of DDI Prototype Definition report](#)

### **Priority Action Items:**

- Advisory Group to review the report and make any changes needed
- Executive Board to approve final prototype proposal

## Production Framework Group

Johan Fihn, Oliver Hopt, Jon Johnson, Olof Olsson, Dan Smith, Wendy Thomas

Colectica has produced an open-source platform for producing a model-driven standard, which supports the creation of XML schema, RDF, JSON, C# representation, and documentation. While this product – named COGS – is used by Colectica in their own work outside the DDI Alliance, they are long-standing members of the DDI community, and have licensed the tool in such a way that it is available to the DDI Alliance for use free of charge.

Dan Smith of Colectica presented the tool in an hour-long demo on Tuesday evening of the Sprint, to the plenary group. Members of the Production Framework team were already aware of the tool and its capabilities, as were many other members of the DDI community. This had led to questions around whether or not the COGS platform could be utilized to advantage by the DDI community in its own work.

As described in the report produced by this group (see below), some of the Production Framework team’s activity during the sprint was to conduct a risk analysis to determine whether COGS represented a viable alternative to the existing production platform. It is important to note that the COGS platform is heavily based on the capabilities of Github, a generic and widely used toolset for the management and storage of code during development projects. COGS takes this basic platform functionality and applies it specifically to the development and leverage

of a model-driven standard, including the creation of multiple parallel bindings and documentation.

It is felt that for long-term production, having a single development platform based on COGS would be very desirable. This shift away from the Drupal platform would need to be made carefully, and should not disrupt the development of currently scheduled releases (June 2018 Prototype Release).

### [Summary of Discussion and Recommendations from the Production Framework Group](#)

#### **Priority Action Items:**

- Present recommendations to the broader DDI Alliance community for discussion and approval
- Incorporate structured documentation recommendations from Dagstuhl 2017 sprint

## Structured Documentation Group

Jon Johnson, Jay Greenfield, Johan Fihn, Wendy Thomas

This group worked on the following items:

- Definition of required and optional items for multiple documentation types.
- Modular examples with the perspective of a specific audiences or purposes
  - Software developer, archivist, data producer, training tutorial
  - The examples should be realized in the technical format used (restructured text)
  - Examples in different target formats should be realized, like for single documents, linked documents, slides.

The resulting recommendations will be rolled into the production framework discussion and possible transition (see above).

### [Recommendations for Structured Documentation](#)

## OWL/RDF Group

Darren Bell, Guillaume Duffes, Deirdre Lungley, Eric Prud'hommeaux, Achim Wackerow, Benjamin, Zapilko

The discussion was extensive and comprehensively documented in the following report (some review comments already included), complete with recommendations, action items, and

outstanding issues still to be discussed. In addition to the work done during this sprint, the report includes a survey of previous work on this subject and the resulting questions/open issues.

### [RDF Decisions and Open Issues Report](#)

Further comments are encouraged. The original working file is in Google Docs here:

<https://drive.google.com/open?id=1k4Bflp2DsdKZ1CQUyvFaEeYWJz9eziw8LWV4-0ILdkM>

### Supporting Documents

[Discussion with Oliver Hopt regarding existing PIM](#)

[Syntax for names in models and bindings](#)

[Errors in DDI4 - OWL](#)

[example-ont.ttl](#)

[example-ont.rdf](#)

### Priority action items:

- Review the above report and add comments.
- Confirm items critical for the prototype release and post-prototype release
- Set up a short-term RDF Binding Working Group to include members of the modelling team, the production framework group, the RDF consultant, and possibly others present in Dagstuhl.

### Test Cases Group

Darren Bell, Larry Hoyle, Deirdre Lungley

Related to the work of the OWL/RDF group, Deirdre, Dan, and Larry worked on producing technical test cases in binding syntax (XML and RDF) for core items of the the Functional Views. The test cases of W3C specifications are the role model for this. A good example is the suite of [test cases for CSV on the Web](#). The following examples were produced using the Australian Election Survey:

[HierarchicalCodeListExampleV2.xml](#)

[HierarchicalCodeListExampleV2.ttl](#)

[HierarchicalCodeListExampleV2.shex](#)

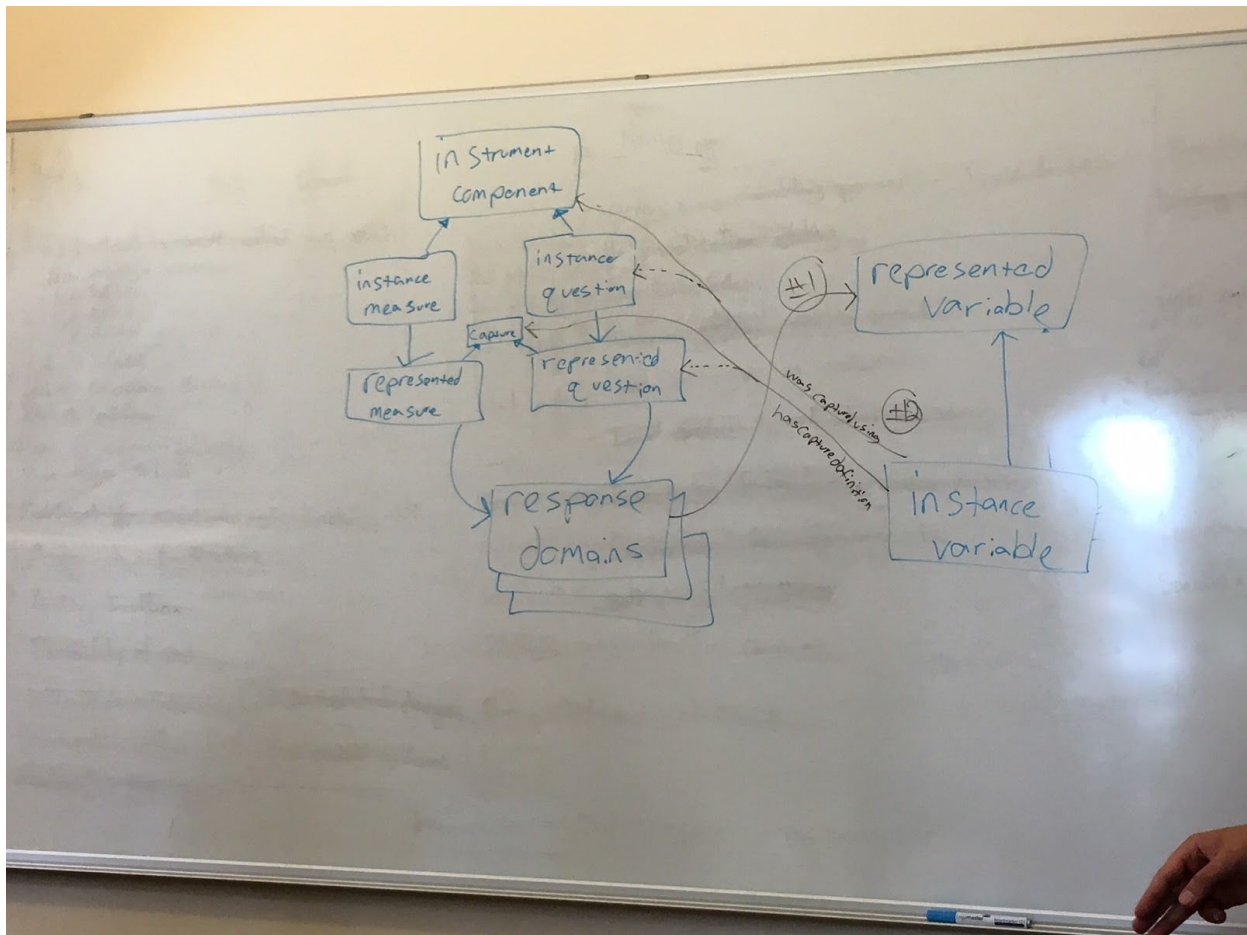
### Data Description and Data Capture Integration Group

Dan Smith, Wendy Thomas, Jay Greenfield, Jon Johnson, Johan Finn

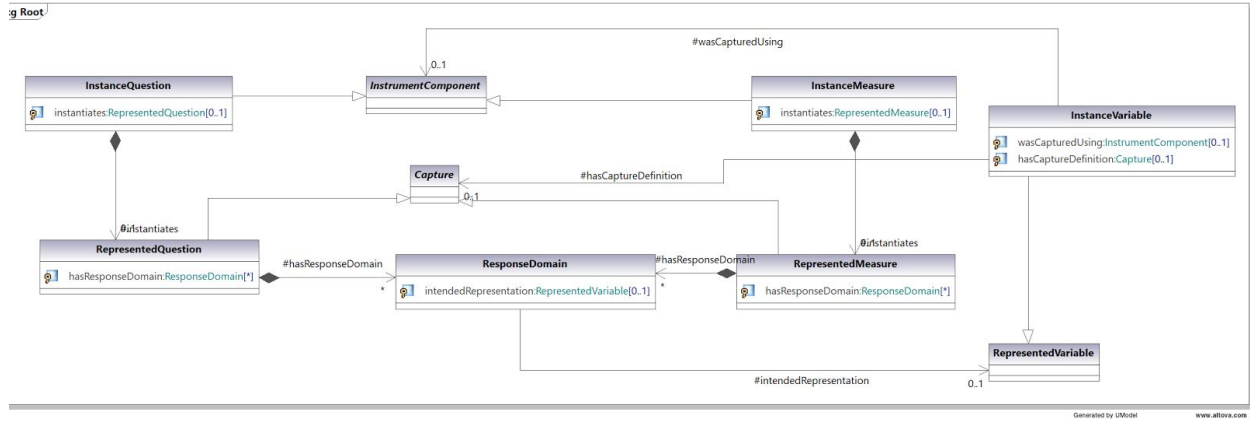
This group met for a small portion of the week to outline proposed touch points between Data Capture and Data Description and prepare the discussion for the second week to resolve open issues.

The full report of this groups work can be read here: [Data Description and Data Capture Integration](#)

(click for larger image)



(click for larger image)



The classes involved::

SOURCE	TARGET		
Response Domain	Represented Variable 0..n		
Instance Variable	Capture (abstract) "sourceCapture":	Think of this as pointing to the represented question	May not have a defined flow therefore require this to capture in these location
Instance Variable	Instrument Component (abstract) "wasCapturedUsing"	Think of this as pointing to the instance question	

Object	Definition / Use
Instance Question	The usage of a Represented Question in a Question Flow (3.2 question construct); an instance question may be used multiple times in a question flow
Represented Variable	A Represented Variable is reusable and therefore can never really know all of the sources of data that are expressed in a way that meets its definition.

Instance Variable that is derived from multiple pieces of data	Points to the individual Instance Variables that were used to derive the instance variable in question
Capture point #1	The relationship between the Response Domain and the Represented Variable that defines it (source is ResponseDomain, target is RepresentedVariable)
Capture point #2	From InstanceVariable to the InstrumentComponent and Capture

## Validation

Ancillary to all groups was the discussion on validation. Arofan Gregory produced a summary of the validation needs and recommendations for the DDI specification, with Jon Johnson making subsequent revisions:

[https://ddi-alliance.atlassian.net/wiki/download/attachments/72815093/Validation%20in%20the%20DDI\\_jj.docx?api=v2](https://ddi-alliance.atlassian.net/wiki/download/attachments/72815093/Validation%20in%20the%20DDI_jj.docx?api=v2)

### Priority action items:

- Review validation procedures for incorporating into the prototype.

## Canonical XMI

Achim Wackerow et al

Report of the discussion:

[Criteria for Creation and Use of Canonical XMI](#)

### Priority Action Items:

- Review report and make any adjustments to the prototype description as needed

## DDI 4 Goals

Wendy Thomas

A revised and consolidated list of goals for DDI 4 was written by Wendy. This document brought together material from the Alliance website and “Why a New Version of DDI”, among others.

[DRAFT DDI 4 Goals](#)

**Priority Action Items:**

- Review document and bring to the Executive Board for approval.





