Development Review Report

Q1 2015 Development Review-Technical Committee – 2016-03-10

# Summary:

The DDI 4 package for Q1 2015 Development Review was released for review on 27 April 2015. The review period was from 27 April to 15 June 2015 (7 weeks). A total of 72 issues[[1]](#footnote-1) were filed by 7 organizations[[2]](#footnote-2). All issues were reviewed by the Technical Committee, assigned to members for additional research, and then discussed and disposed of by the Technical Committee. Disposal resulted in one or more of the following:

* Change to the Drupal content by the Technical Committee
* Recommendation to the Modeling Team for specific actions
* No action required – comment noted as a general comment in Development Review Report

# Changes implemented by the Technical Committee (TC)

|  |  |  |  |
| --- | --- | --- | --- |
| **Issue** | **Description** | **Decision** | **Action** |
| 4 | List of classes with no definitions in document | Add documentation | Documentation added |
| 5 | Other Material resides in wrong package | Should be moved to Utility Package | Moved |
| 7, 25, 51 | InstanceVariable association to RepresentedVariable; RepresentedVariable; | Review in the complete cascade in terms of GSIM; Should the cascade be an extension of the previous or reference the previous level(s). | Entered in Drupal according to decision in 7 |
| 8 | CodeValueType and InternationalCodeValueType are redundant | Collapse to a InternationalCodeValueType to retain language designation. Rename to ExternalControlledVocabuarlyEntry and change property names to reflect this change. (CodeValue used elsewhere in a different context). Update all uses. | Done |
| 9, 58, 69 | Inconsistent use of date types throughout | Change BaseDate to inlcude and isoDate and nonISO option; Create a Date (BaseDateType) and DateRange containing startDate and endDate both of BaseDateType. Correct the current usage as noted in spreadsheet on 69 | Corrected in Drupal |
| 33 | Typographical error | Fix | Corrected in Drupal |
| 26, 27 | Clarification of Universe and Population | Clarification of documentation | Done |
| 35 | Abbreviations of classes confusing | Don’t use abbreviations, correct in all documentation | Done |
| 36 | DDI DescribedValueDomain should map to GSIM DescribedValueDomain rather than ValueDomain | Yes it should | Corrected in Drupal |
| 39 | Distinction of measures vs. takes value from and lack of link with a response domain | Note the equivalency in documentation | Done |
| 44 | UnitType relation with Unit | Changed cardinality from 1:1 to 0:1 on relationship | Corrected in Drupal |
| 45 | Unit has definition but no description | Added description and revised content of definition | Corrected in Drupal |
| 46 | No hierarchical bookmarks in PDF | Sphinx auto-generates hierarchical bookmarks | Added as a documentation requirement but addressed by shift in software used for document production |
| 57 | Need HTML documentation | Sphinx profides this option | Noted as requirement; currently solved by switch to Sphinx |
| 67 | Reviewing and Documentation issues | High level Purpose added to documentation; switched to Sphinx allowing more “chunkable” content; begun collecting examples and how-to information for documentation; section on modeling conventions added; new issue created regarding point on modeling conventions (DMFQA 76) | Done |
| 72 | ClassificationSeries property owner should refer to an Agent; consistent use of CodeValueType | Changed datatype to AgentAssociation and cardinality to 0..n to match definition; Addressed in Issue 8 | Done; Addressed in Issue 8 |
| 73 | Funcational view validation; OASIS – web services | Split into Issues 77 and 78 | See related issues |
| 10 | Confusion of different textual content structures | Clarify documentation for user and usage guide for content modelers using 3.2 documentation |  |
| 12 | Typo in introduction GWIM  | Corrected to GSIM | Done |
| 13 | Model Construct and relationships, Fig. 3 p.3 unclear regarding extension base | Clarified | Done |
| 15 | Typo in AnnotationDate “equivalent” | Corrected to equivalent | Done |
| 16, 17 | Inconsistent use of “lang”, xmlLang etc. | Change all to language unless referencing something other than the metadata content | Done |
| 19 | Clarify difference between bag and set | Add clarifying documentation | Done |
| 30 | EqualTemporalRelation use of equalA and equalB parameters unclear  | Change equalA and equalB to equal with a target cardinality of 2 (type ProcessStep) | Done |
| 37 | Refinement of Enumerated and Described Value Domains is unclear | Documentation expanded for Enumerated, Described, Sentinal and Substantive Value Domains | Done |
| 43 | Differentiation between Name and Label | Structure of Name revised to contain content (xs:string) and context (External Controlled Vocabulary Entry); Label renamed to DisplayLabel | Done |
|  |  |  |  |

# Action List for Modeling Team (MT)

|  |  |  |  |
| --- | --- | --- | --- |
| **Issue** | **Description** | **Recommendation** | **DMT - Issue** |
| 6 | Incomplete list of xs datatypes in primitive | Review all primitives to identify need for defining conceptual class as opposed to physical representations | 16 |
| 20, 21, 22, 23,24 | OrderRelation, Relatioships, properties; OrderCollectionCorrespondence, Ordered CorrespondenceMember; ConceptParentChild; ConceptPartWhole, Relationships | The relationships expressed in Collection have been undergoing work within the MT and specific issues regarding relation types have been addressed. 1. High level documentation need to clarify the implementation of this pattern for both the Business Modelers and the user in terms of implementation.
2. Clear specification of related RDF vocabulary is needed at the Class and Property level
 | 17 |
| 61 | Adoption of GSIM terminology particularly in the domain of statistics | Review use or specification of GSIM terminology in DDI 4 and ensure consistency  | 18 |
| 61 | Capture class: terminology unclear, is it a verb or noun; the inheritance chain was confusing as some represented a description/definition of a task and others metadata for a task | Clarify naming rules in terms of verb/noun; review inheritance chain. Is there a design rule needed here? | 19 |
| 68 | General observations on intent and coverage of DDI4 | MT needs to make a clear statement regarding the intent and coverage of DDI4. This statement needs to be provided in both the internal work documents (for Business teams) and in the High level documentation for the user. See section General Comment’s section and JIRA issue for details. | 20 |
| 68 | General observations on intent and coverage of DDI4 | Review for consistent use of Abstract classes | 21 |
| 47 | Clarification of references to GSIM, DDI, and RDF objects where there is not a one-to-one correlation | Review all GSIM, DDI, and RDF references for clarity (after change in Drupal has been made) | 22 |
| 76 | Describe modeling conventions in high level documentation | Clarify modeling conventions and document | 23 |
| 78 | Need validation for Functional Views | Provide recommendations for validation process for serializations | 24 |
| 10, 11 | RDF properties need to be further developed | This serialization is still in early stages, this issue needs to be addressed during this process | 25 |
| 75 | RegExp serialization in model to support multiple bindings | Review for XML centric approach | 26 |
| 77 | OASIS Compliance for web services | Consider this issue as model is developed | 27 |
| 28 | Clear instructions for entering RDF mapping | Instructions are needed for both selecting RDF mappings and entering them ie. What goes in Label? Full namespaced term or just term, etc. | 28 |
| 18, 32,34 | Question regarding mapping of class Organization to specific RDF classes; Mapping to Dublin Core elements in RDF serialization (example: dcterms:temporal) | This is part of a larger issue on how accurate references to RDF classes are determined | 22 |
| 50 | Review all classes that extend Concept for appropriateness (some should reference, not extend Concept) | MT should review all relationships to Concepts in 3.2 as well as consider which new classes should have a relationship to concept (e.g. Collection) NOTE: Review the use of Concept as an extension base for Conceptual Variable TC believes it is more meaningful for a Variable to “Measure” a concept and recommends breaking the inheritance chain between Conceptual Variable and Concept. | 56 |
| 52 | Unclear definition between Classification and Code | May have been resolved by subsequent MT work. Please review. | 57 |
| 41\*, 43\* | Differentiating between definition and description. Differentiating between Name and Label | Provided documentation and usage of each term. MT provided with a list of locations and recommended changes. | 61 |
| 53\* | Relationship of various classes to Concept | The modeling team should review all of the relationships to Concept in DDI3.2. as well as considering which new classes should have a relationship to Concept (e.g. Collection) | 56 |
| 38\* | Intent and usage of Designation | The MT should clarify the class specific documentation as well as the use of these classes in describing a Representation. Designation is currently abstract with no instantiated class. | 60 |

**\* Note that this set of starred issues must be addressed prior to the Q2-2016 review. Clarification of relationship to Concept as well as consistent use of Name, DisplayLabel, Defintion, and Description are critical to the consistency of the model and need to be corrected ASAP.**

# Drupal Issues:

|  |  |  |
| --- | --- | --- |
| **Issue** | **Description** | **Recommendation** |
| 47 | Clarification of references to GSIM, DDI, and RDF objects where there is not a one-to-one correlation | Add a paired text box to add clarification on GSIM, DDI, and RDF related objects |
| 47 | Need references to GSIM, DDI, and RDF on properties | Add structure similar to that on Class to Property |

# Issues requiring no action:

|  |  |  |
| --- | --- | --- |
| **Issue** | **Description** | **Recommendation** |
| 1 | Lack of separate documents for each Functional View | Acknowledged this was done as an expedient to get this review out the door |
| 54 | Self-reference of Instructions in Instrument  | Appears to have been corrected post-publication; no longer in evidence |
| 39 | Distinction of measures vs. takes value from and lack of link with a response domain | Data Capture (Instrument) has not been released for review so there is no response domain to link to as of yet.  |
| 40 | Capture no represented in diagram | Corrected in latest Drupal export |
| 49 | No relationship from InstanceVariable to RepresentedVariable etc. | Duplicates Issue 7 |
| 63 | Incomplete properties noted in non-published sections | Outside of coverage for review |
| 14 | Access class has no clear relationship to an object | Artifact of development stage. Continued development will result in linkages from objects to this class |
| 29 | The role of Act as part of the Core Process in terms of classes in Instrument | Instrument is not part of this release; Issues relating Act were discussed and clarified at Dagstuhl 2015 sprint |
| 31 | How is it possible to create sequences using Overlaps Temporal Relationships unless switching from an “IfThenElse” approach to a “GoTo” approach? | This has been resolved through changes made in the Process Model by MT post review period |
| 54, 55 | Instructions has relationship to itself; Loop property hasLoop could be named more clearly | Appears to have been resolved post review; Instrument was not part of review |
| 56 | Can’t create a simple sequence of 3 items in Data Capture | Appears to have been resolved post review; Instrument was not part of review |

# General comments raised in issues submitted during review:

Abstraction (59, 60)

Comments provide a good overview of both the value of abstraction and the problems of overuse. In general, the conclusion was that “abstraction is definitely a step in the right direction” and that the abstract layer will assist in mapping between DDI and other standards.

“Given the origins of DDI in XML Schema, it appears that the “real life” aspect (i.e. concrete classes) predominates. However DDI’s initial efforts into abstraction, such as Collections and Tree Node structures are an excellent start on an appropriate path to redress this. The clean model should come from increased use, and reuse, of abstract patterns.”

TC recommends to the MT that these comments be reviewed and kept in mind during modeling work.

Cascade model of describing variables (62)

The cascade model seems to be a valid approach and has a strong similarity to GSIM. Some differences may benefit from further information/explanation. Note that specific issues regarding the cascade model are being address by the TC or referred for specific action to the TC.

Model type (64)

The discussion in DMFQA-64 indicates a need for more clarity on the intent of the DDI 4 model as it is falls on the continuum between a conceptual and physical model. We should, in our high level documentation and design rules, be clear about what this model intends to be. [Technical Committee]

Provides an excellent description of three levels of the model – conceptual, logical, and physical. The DDI model and documentation is clearly more detailed than a conceptual model, and so we assume it is targeted as a logical model, by our terminology. As such it seems to be at an appropriate level of detail. Suggests the utility of a more clearly conceptual model, assuming that the Bindings will provide the physical model.

Guidance on use of inheritance (65)

Good discussion of the inheritance which should be reviewed as MT reviews further work and discusses issues of deep inheritance structures, overrides, and use of patterns.

General observations (68)

The following general statements are followed by TC comments and should be considered by the MT and included in documentation as appropriate.

* It was unclear to us whether DDI 4 would be a simple re-implementation of version 3.2 in a modelling environment, or whether it was also to be a logical upgrade.
	+ There are a couple of major goals with DDI 4: the basis of the standard should be a UML model, the functionality of DDI 3.2 should be included, a strong relationship to GSIM should be realized, expansion to additional areas (i.e. logical upgrade) for example abstraction of the data source (data capture), process description, early stages of study design.
	+ All this might be not enough visible in this first draft development release.
* Seems to be focusing on "after-the-fact" documentation, rather than "at-the-same-time" designing of processes, documenting up-front.
	+ Again, this might be not enough visible in this first draft development release. Documentation should be reviewed for bias and emphasis both uses.
* Is showing signs of its heritage from XML Schema, rather than being a clean new model.
	+ The goal is to have a clean model without strong relationship to XML Schema. May wish to ask for specific examples of this in future reviews.
* Trying to meet too many scenarios and resulting in "design-by-committee" with lots of point solutions, rather than an integrated whole. A solution to this would be to make better use of abstraction.
	+ This is a draft development release. Keep in mind when reviewing content.
* DDI seems to contain re-definitions of other standards, such as Dublin Core and XSD:dateTime. Perhaps it would be better to reference these, rather than re-define them.
	+ The goal is to make use of existing standards where appropriate and where they are mature and accepted. A common issue is that other standards are often not specified in UML, but in XML Schema, DTD, RDF, etc. or only in a text document. The design principle is to define related classes in the DDI namespace and to define a class equivalence to possible existing class definitions of other standards or to express the relationship to other ways of definitions. This way DDI is not dependent on possible changes of other standards.
	+ This approach will be realized in the implementation process and should be clearer in future drafts.
* The best solutions seem to lie in the standard UML meta-modelling (level 2 and 3) capabilities, as implemented in Sparx EA. Is Drupal capable of this?
	+ We didn't consider UML meta-modelling (level 2 and 3) yet to keep the modeling simple. We consider using stereotypes as extensibility mechanism. This is a simpler and more portable mechanism than meta-modeling. May wish to discuss further with ABS.

Clarify overall intent of model in a clear document that can be included in the high level documentation. This should cover:

* Intent of the model in terms of DDI-L content (is it a model of 3.2 or upgrade)
* Intent to capture prescribed, active, and retrospective documentation (plan to do, do, what was done) plus usage information where applicable

Review model periodically for:

* XML heritage slant
* Clear integration of the whole as well as of individual functional views

Clarify overall relations to other standards at the object level and in high level documentation:

* Where do we use directly (i.e. xs datatypes)
* Related objects in other standards (i.e. GSIM, ISO 11179, ISO 19115 and others)
* Use of patterns from other standards (i.e. BPMN/BPEL)
* Clear specification of RDF mappings at the object and property level

# **Issues for TC to review on next development release**

|  |  |
| --- | --- |
| **Issue** | **Description** |
| 42 | Behavior of objects using the abstract Member as an extension base |
| 74 | Define a query string for capturing data; object for describing intended use of data; data generation (coded, imputed, etc.); compartmentalizing populations (target, actual, estimated, etc.) |
| 70 | Overuse of Annotated Identifiable |

1. Note that the filed issues are numbers DMFQA-1 and 4-74. Issues 2 and 3 were test issues and removed from the issue tracker. [↑](#footnote-ref-1)
2. Australian Bureau of Statistics (ABS), Colectica, GESIS – Leibniz Institute for the Social Sciences, Institut national de la statistique et des études économiques (INSEE),Minnesota Population Center (MPC), University of Kansas (KU), and Statistics New Zealand [↑](#footnote-ref-2)