

Updates to DDI-Codebook

DDI Version 2.5 – 14 January 2012

DDI-Codebook was updated in response to requests for new DDI elements and attributes and to facilitate the transformation of metadata from DDI-Lifecycle to DDI-Codebook and back.

GENERAL

FEATURE	DESCRIPTION
Citation	Three changes have been made to citation in all of its locations. IDNo attribute “level” has been extended to include the valid value of project. Series statement (serStmt) has been made repeatable to allow for membership in multiple series. Citation now provides the option to support all qualified Dublin Core elements (dc and dcterms).
codeBookAgency	A place to put a registered DDI agency identifier.
Controlled Vocabularies	A structure has been added to DDI-Codebook to allow for the use of the controlled vocabulary structure used in DDI-Lifecycle. To ensure backward compatibility, controlled vocabularies are listed in the document description section (docDscr) following the citation. The fields are similar to those in DDI-Lifecycle with the addition of a usage element where one must note the element or attribute that uses the controlled vocabulary using an XPATH. Restricted elements with internal code values were opened up with the addition of “other” to the allowed codes and an attribute that allows for listing the other value.
date fields (new)	New date fields enforce the use of the ISO structure allowing for YYYY, YYYY-MM, and YYYY-MM-DD plus time.
ddiCodebookUrn	Allows all elements to have a DDI-Codebook URN.
ddiLifecycleUrn	Allows all elements to capture an element URN from a DDI-Lifecycle element.
elementVersion	Allows all elements to contain a version number.
elementVersionDate	Allows all elements to contain a version date.
Notes	All notes can now explicitly state their parent element (for example, a Variable Note can reference the ID of the variable it is nested under). If the same note is related to a number of elements, a single note can be treated as a “master note” and list the elements it is related to. This addresses the different structure of notes in DDI-Lifecycle if moved into DDI-Codebook and allows DDI-Codebook users who wish to prepare for a transformation to DDI-Lifecycle to capture this information in a single location.
XHTML	Permits the optional use of XHTML wherever formatted text is allowed. This enables structured text to transfer directly from DDI-Lifecycle or for the user to prepare currently structured text in DDI-Codebook to be restructured to a format compatible with DDI-Lifecycle. Consult the documentation before using.

SPECIFIC FEATURES

Most specific changes were made to the study description section of DDI-Codebook. The features are listed in alphabetical order within their primary section of the DDI-Codebook schema.

<i>Study Description (stdyDscr)</i>	
abstract NEW attribute contentType	Provides a tag indicating that the content should be mapped to a DDI-Lifecycle <i>Abstract</i> or <i>Purpose</i> or mixed. Mixed content can be tagged internally using structure tags. This is used to facilitate the transformation between the two development branches of DDI.
catLevel NEW attribute geoMap	Provides the ability to reference one or more maps described by the element geoMap by listing the IDs of the relevant maps. This feature is used when the described map pertains to a limited set of categories in a hierarchical category scheme as described using the catLevel element.
codingInstructions	Provides for a detailed set of information on coding instructions including the type, any related processes, a description, and formal command language.
collectorTraining	This reflects new information that will be added to DDI-Lifecycle covering the type of training provided to the data collector. It includes a type attribute.
dataAppr NEW attribute type	A type attribute has been added to <i>Data Appraisal (dataAppr)</i> to support better classification of this activity.
dataCollector NEW attribute role	A role attribute was added to distinguish between different roles provided by different data collectors. This could differentiate between steps in the process or type of position such as supervisor.
dataProcessing	Allows for a typed description of data processing during the collection, initial processing, or creation of a data product.
exPostEvaluation	This element has been added to capture the step of process evaluation for a study. It provides a means of capturing information on the completion date, type of evaluation, evaluator, evaluation process, and outcome of the evaluation. This relates to the step "Process Evaluation" found in the Generaic Statistical Business Process Model (GSBPM). A similar structure will be added to DDI-Lifecycle.
instrumentDevelopment	This reflects new information that will be added to DDI-Lifecycle covering the instrument development process. There is a type attribute and an option for structured content.
qualityStatement	This is a structured element that allows for the identification of specific standards to which the study complies and a description of that compliance. In addition, any form of quality statement not associated with a standard can be entered here.
sampleFrame	This reflects new information that will be added to DDI-Lifecycle covering the details of the sample frame used for the study sample. It includes information including the standard name, label, and description elements, plus a valid date for the frame, who maintains it, its use, universe, frame unit information, a reference period, and update procedures.
sourceCitation	A citation has been added to sources to allow for a specific identification of an external data source.

studyAuthorization	Provides structured information on the agency that authorized the study, the date of authorization, and an authorization statement, i.e., a law authorizing a census, a statement from an Internal Review Board, etc.
studyBudget	Provides a text field where the overall budget of a study can be entered. The field allows for structured content.
targetSampleSize	Allows for specific information regarding the target sample size, actual sample size, and the formula used to determine this.
<i>File Description (fileDscr)</i>	
fileCitation	Allows for the provision of a citation to a specific data file that is being described by the DDI-Codebook.
fileStrc NEW attribute fileStrcRef	This allows for a single description of a file structure when multiple data files share the same structure. After the first file's structure is described, the others can simply reference this description.
<i>Data Description (dataDscr)</i>	
Question NEW attributes responseDomainType otherResponseDomainType	Questions have a new attribute which allows for the direct identification of a specific DDI-Lifecycle response domain type (i.e., code, datetime, numeric, etc). This clarifies any ambiguity for transforms from DDI-Codebook to DDI-Lifecycle and allows for specific identification of this information when transforming from DDI-Lifecycle to DDI-Codebook.
Variable NEW attributes representationType otherRepresentationType	Variables have a new attribute which allows for the direct identification of a specific DDI-Lifecycle representation type (i.e., category, code, text, etc). This clarifies any ambiguity for transforms from DDI-Codebook to DDI-Lifecycle and allows for specific identification of this information when transforming from DDI-Lifecycle to DDI-Codebook.

Updating existing DDI Codebook instances to version 2.5

A number of changes were made to the namespace in order to bring it into line with the structure of DDI-Lifecycle namespaces and ease transformation to further versions within this development line. Note that the version number is no longer part of the XML schema filename.

The DDI-Codebook development line is backward compatible meaning that instances compliant with DDI versions 1 – 2.1 will also be compliant with version 2.5. Note that the canonical expression of DDI versions through 2.1 is the DTD. Some editing software references an XML schema version of the canonical DTD. To update these files to version 2.5 you will need to make the following changes:

If the instance refers to the DTD:

EXAMPLE:

```
<?xml version='1.0' encoding='UTF-8'?>
<!DOCTYPE codeBook SYSTEM
"http://www.icpsr.umich.edu/DDI/Version2-1.dtd">
```

```
<codeBook version="2.1">
```

REMOVE the DOCTYPE declaration (i.e. `<!DOCTYPE codeBook SYSTEM "http://www.icpsr.umich.edu/files/DDI/Version2-1.dtd">`)

ADD the following to <codeBook>

Declare the DDI 2.5 target namespace:

```
xmlns="ddi:codebook:2_5"
```

Declare the XMLSchema-instance (xsi) namespace:

```
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
```

Insert a reference to the schema location using the XMLSchema-instance schemaLocation:

```
xsi:schemaLocation="ddi:codebook:2_5 codebook.xsd"
```

CHANGE the value of the attribute "version" in the element "codeBook" to "2.5".

EXAMPLE:

```
<?xml version='1.0' encoding='UTF-8'?>
<codeBook xmlns="ddi:codebook:2_5"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="ddi:codebook:2_5 codebook.xsd" version="2.5">
```

The above `xsi:schemaLocation` assumes the XML instance is in the same folder as `codebook.xsd`. If a local copy is being used and is in another location, add the pathname to `codebook.xsd`.

To reference the remote location at the DDI Alliance use:

```
xsi:schemaLocation="ddi:codebook:2_5
http://www.ddialliance.org/Specification/DDI-
Codebook/2.5/XMLSchema/codebook.xsd"
```

If the instance refers to an XML schema:

CHANGE the target namespace, and `xsi:schemaLocation` and version number to the values noted above.

ADD XMLSchema-instance if needed.