# BRIDG Architecture Review Team Meeting minutes – April 10th, 2015

# Attendees

* Boris Brodsky, FDA
* Hugh Glover, Blue Wave/Parexel
* Smita Hastak, Samvit Solutions
* Wendy Ver Hoef, Samvit Solutions
* Julie James, Blue Wave/Parexel
* Bob Milius, NMDP
* Diane Wold, CDISC
* Terry Hardin, Parexel
* Jane Pollack, NMDP

# Agenda: Terminology, part 2

## Value Set binding

The focus of this call was on the “How” of vocabulary binding. Hugh led off discussion regarding how to do the binding since there are no mechanisms in place for doing the value set binding. He discussed both what is involved in value set binding and what’s used at Parexel. This was specifically for the CD data type.

* Parexel implementation of value set binding: Code/value lists are managed in spreadsheets. In their BRIDG-based common data model EA file, they have created a terminology data entry form. Here are the 4 properties in the form that capture the key factors:
* concept domain definition
* Value Set OID
* graceful degradation level
* code value
* Tags for attributes with CD data types – in EA, consider adding a tag (similar to a BRIDG mapping tag) to point to a value set

Parexel does not have Code system version at this time. It is useful mainly if you are evaluating data over time. Discussion regarding usefulness of code system version. NCI had implemented the version property of CD data type to support the use case(s) of AE where the NCI CTC AE has versions. Could apply similarly to MedRA versions.

Value set definition – by enumeration (explicitly listing values) = extentional; by logic at run time, constraining existing value sets = intentional. UCUM – can be used in intentional sets

## Value set binding in BRIDG

There was a lot of great discussion on how we could support terminology binding in the BRIDG model. Below is a summary of the various discussion points and is laid out as some high level steps.

Assuming BRIDG was to define domain terminology for attributes with a CD datatype in the BRIDG model, following are few things to consider:

* Identify which properties of CD datatype we would recommend using and implementing – all or a small subset?
  + Based on NCI and Parexel experience it would appear that just a subset may be really needed.
* Identify where does this terminology or value set reside
  + Do we maintain these in spreadsheets and then publish an updated spreadsheet with every release OR do we publish them to NCI Thesaurus (or something similar)?
  + Spreadsheet solution would be a maintenance headache and not scalable – need something more dynamic
  + NCI Thesaurus may be a good option to consider, especially since 3 of the 4 stakeholders of BRIDG (CDISC, NCI, FDA) already use it to manage their terminology. There may be some challenges in leveraging NCIt – work with NCIt team on this and make recommendations to NCI
* Find a way to bind the value set (or reference to the value set) to the attribute in the BRIDG model file
  + Could look into adding tags (similar to BRIDG mapping tags) at attribute level
  + Evaluate whether the value of the tag could point to a url – either to NCI Thesaurus or to an external source
  + Create a profile in EA with the value set tags predefined. This will avoid typos, etc.

Next BART meeting (4/24/2015) – continue the vocabulary/terminology binding discussion