Minutes for Data Description - Data Capture Integration 2017\_10\_23 afternoon

Beginning with the summary from the work during week 1

<https://ddi-alliance.atlassian.net/wiki/download/attachments/113377344/Data%20description%20and%20Data%20Capture%20Integration.docx?api=v2>

Call with Barry during preparation week

Capture needs to know where to put captured data and the workflow of the capture

Description needs to know what captured the data

Can’t put link from reusable capture to variable, should go from description to capture

Connect InstanceVariable to reusable Capture and the instrument component that resuls in data. This is the “definition of the (see table from week 1)

RepresentedQuestion is like a question from a question bank, including multiple response domains (as in the case of “pick all” mapping to a set of binary questions..

The response domain can change over time.e.g. Fill in the blank vs pick one. Has a concept, a universe and a response domain. Response domain might end up with different value domains in variables.There could be processing between the response and the storage as a variablle.

All of the above is about design, not execution i.e. “Capture” is designed



Do sentinel values go in response domains - as defined last week yes.?????.

Does this cause problems for a relationship from response domain to a represented variable. Sentinel codes are at the InstanceVariable (filed in DCAP-53)

The question cascade does not parallel the variable cascade.

InstanceQuestion is the customizing of a reusable question in an instrument

Sentinel domains in represented questions

 A representedVariable does not have sentinel values, whereas the response domain for a represented question does have this information.

RepresentedVariable to RepresentedQuestion is problematic since there may be multiple

Response domains are not reusable - they have a 1 to 1 relationship with RepresentedVariables.

If they are you can’t navigate from RepresentedVariable to RepresentedQuestion.

Can a LIKERT scale be reused/

Dan Smith’s designed capture does not produce a Datum

The term “Capture” may mislead by indicating execution

Need to document the differentiation between design and execution (Arofan will write up)

How can one use RepresentedVariables as a CDE if it is missing the sentinel values?

How can RepresentedVariables act as RepersentedVariables WITH sentinal values?

Represented variables has codes for Substantive Values and categories for with Sentinel Values.

Instance variable is not reusable. If the variable is in an repeated data set. For example where the differentiation of population such as time and geography is captured not in the description of the variable but as a variable associated with the case. Essentially, the population changes each time, therefore an Instance Variable is NOT reusable. This means Dan Smith is right and therefore Data Description needs to talk about what is the Instance Variable and what is RepresentatedVariable in terms of Sentinel VALUE Domain.

Not everyone does RepresentedVariables, just InstanceVariables.

What is the question text and what is the response domains.

I have a RepresentedVariable (conceptuall, meanings and codification of meaning for substantive, may or may not categories for the sentinal)

I have a RepresentedQuestion has response domains which are tied together and related to an RepresentedVariable currently has a Substantive Conceptual Domain, Sentinel Conceptual Domain,

RepresentedVariable is the “reusable portion”

Week 1 issues fleshed out:

List under For next week provide some definitions:

1. What is a response domain with examples? - Dan S.
2. What the touchpoints between Data Capture and Data Description currently in the model and how do they act? - Wendy
3. What is the relationship of the Datum to an execution layer?
4. What is the “instance” of an InstanceVariable in relationship to a RepresentedVariable? In other word, what is the difference between a RepresentedVariable and InstanceVariable, what “causes” it to become an InstanceVariable? I can design an instrument to be populating an SPSS data set for a particular population. I can do this without collecting any data. You have a conceptual level, a logical level, a design level, and the execution level. What has it got or not got at each stage? What are the distinctions between each level in our current cascades? Are there differences between what is being described in the capture side and the data description side? Where does it tie in tightly with execution? If we are talking about datum, that involves something that has been measured and it goes in a data point. A data point is a “cell in a spreadsheet” where the instance variable describes the column and the row describes the execution at the case level. If you’ve got a particular population and whatever else you have is unique to the instance variable (population which is a restriction of a universe and a restriction of a unit type). The IQ talks about what you are going to put into a cell. The RQ represents reusability. What level of reusability are we focusing on? The distinction between design and record. This may be how we should be thinking about capture rather than represented and instance. -
5. What is the relationship between Represented Variable and CDE? If a CDE needs sentinel VALUE domain, and Represented Variable does not contain this, is there a need for these two to be equivalent or to provide the appropriate relation? Is there something that needs to be done to sort out sentinel from substantie at the point of the Represented Variable?

### Dividing the week one questions into smaller units

* Basically we need the clear definitions of what the Response Domain is in DDI4 with some examples,
* what the touchpoints now are and how the act,
* and the relationship of the Datum with the execution layer.
* There is still a misunderstanding between groups of what an “instance” is
* which is the usage of a represented variable.
* While a Represented Variable has changed exacerbating this misunderstanding.
* The Instance Variable is vary tied up with an execution.
* How can one use Represented Variables as CDE’s if it is missing the sentinel values?
* Is there work that has to be done to sort out sentinel from substantive at the point of the Represented Variable?