**Title of session**: What do researchers and data documenters require?

**Day**: Day 3 Session 1

**Participants**: Jon Johnson, Oliver Hopt, Amber Leahey

**Chair**: Jon Johnson

**Note taker**: Amber Leahey

Day 3

Decompose things from a survey workflow

Researchers

-struggle with terminology to translate concepts between researchers, groups, etc.

-think in a workflow, different workflows, ability to track individual workflows

-questions turn into variables

- we don’t give them the structure of the process that generated them, the provenance of the variable creation for example, “where the variable came from”, paths, etc. aren’t defined

-what would we need to capture provenance, question order, universe statements, to support

-understanding “replication path” documentation

-secondary users (use codebooks) and primary data collectors require different things

-secondary users (variables, question, replication)

-data documenters (need much more technical information about the items, processes, workflows, replication path, etc.)

-researchers now need to provide the code that they used to generate the results and analysis, methodology, reasoning for example for recoding, based on some assumptions, etc. cite data,

-workflow could capture the code, and you can have a description of this, for each segment or step involved in the processing of data

-if this were documented you could export a processing workflow > to improve ability to replicate results, paths, logic, etc.

-“audit trail” of each variable would help with compliance for replication, computational reproducibility to improve reuse and quality, transparency,

-data citations?

Example: Workflow step

-add the fields we don’t already have in, for example the controlled vocabularies,

Notes from Jon’s presentation:

-add make it possible to support lower granularity of exported information - for flexibility downstream

-class for controlled vocabularies

-adding crossreferences from ExternalControlledVocabulary

-patterns (talk to Dan)

-tagging

 Achim: tagging on these categories would be difficult

 -put identifiers in there when implementing tagging

-Use cases

 -can we produce a standardised format

 -What would this look like

 -code examples?

-What improvements could be made to usability

 -indexes

 -what links here?

 -cross-section references?

-identify other areas for improvement, for example something in-between the class level and the functional view, can this be improved? Documenting use cases

-diagram generation is difficult to do in Drupal

**Use cases**

What are the components of a use case?

-description

-requirements

- mappings

- code example(s)

- known deficiencies

-DDI Requirements testing (DDI modellers)

-Requirements validation and practical implementation advice (Implementers)

-high-level business functionality, to support the exchange of data / metadata between organization (Deciders)

Arofan /Achim visit:

-its important to focus on granularity; drupal can

-building a topic map would be useful

-support packages from a user perspective

-building functional views

-ask Wendy for a production project in JIRA

-what will the end products look like