

## **NCRN-Cornell node: update 2012-09-04**

### ***Metadata group***

- As we move forward, and develop general mechanisms, we have renamed the metadata-related (repository) component to “Comprehensive Extensible Data Documentation and Access Repository (CED<sup>2</sup>AR)”
- While we worked for quite some time on defining a *subset* of DDI that would be supported, we found little advantage to only supporting a subset, given the technical capabilities of software. In turn, we are still working on defining the exact scale of the *extension* of DDI that will be necessary to address the project goals.
- We have identified a Digital Object Identifier (DOI) provider, and will shortly express the technical challenges of providing DOI when not ourselves the custodian. DOI are necessary to provide a globally unique handle for specific data files.
- A working repository (supporting DDI 2.1 and DDI 2.5, but not yet our own extensions) has been implemented. We are using ICPSR-sourced DDI as well as SIPP Synthetic Beta (SSB) metadata (see below) to initially populate the repository.
- We have handcrafted a confidential-PU correspondence for variable-level metadata for CPS (ASEC – March files), which will be used to expand our repository for our own extensions. The SIPP Synthetic Beta metadata will also be used for this.
- We have worked with Graton Gathright to identify tools that would allow a (Census-based) data provider to create, maintain, and disseminate metadata. Graton and Kelly Trageser provided us with legacy SSB metadata, as well as a first draft of Nesstar Publisher-created DDI. We have not finalized whether using Nesstar Publisher is an adequate metadata maintenance tool, but it is handy, easy to learn, and free to use, while being actively maintained by statistical agencies around the world.

### ***Other work at the node***

- The group working on “High Performance Logistic Regression Methods for Data Edits and Imputation” was on summer hiatus (conference travel, personal leave) since the last update, but have submitted several papers to conferences.
- A Cornell computer-science class will work with the metadata group to develop a user interface, as part of a regular (and successful) teaching tool used in this class.

### ***Changes at the node***

- Carl Lagoze (Senior Scientist, added after the start of the grant) has moved to University of Michigan. He will continue to work with us, using videoconferencing.