**DDI Glossary Working Group**

Wednesday April 20, 2022

0900-1000 (US Eastern)

Join Zoom Meeting

<https://us06web.zoom.us/j/83090594507?pwd=M0pHdi9uWW9vcVZHY2hQVUpNcUlQUT09>

Meeting ID: 830 9059 4507

Passcode: 724077

**Participants**:

* Dan Gillman (chair)
* Steve Richard (Notetaker)
* Laura Molloy
* Valentin Brunel
* Nathalie Gendron

Regrets:

* Arofan Gregory
* Jane Fry
* Barry Radler
* Flavio Rizzolo

Notes:

Pick up at 'series' on [Glossary page](https://ddi-alliance.atlassian.net/wiki/spaces/DDI4/pages/2790359041/Glossary%2Bwork)

Series—important for describing surveys. Use definition from DDI lifecycle

Statistical classification: Debate in GSOM community, and COOS (core ontology for statistics), using 'statistical' not critical. Classification should be exclusive and exhaustive. 'stastical' used in lifecycle, xkos, cdi. Look at GSIM (Generic statistical information model). Options—ignore 'statistical', 2. We are keeping because of extisting usage' 3. NeuChatell group in 1990s, workon use of classifications. <https://en.wikipedia.org/wiki/Statistical_classification?msclkid=a31340a2c0ab11ec96a815d547d59af7>

L. Molloy in chat: In statistics, classification is the problem of identifying which of a set of categories (sub-populations) an observation (or observations) belongs to. Examples are assigning a given email to the "spam" or "non-spam" class, and assigning a diagnosis to a given patient based on observed characteristics of the patient (sex, blood pressure, presence or absence of certain symptoms, etc.).

L. Molloy pointed out that last meeting only got to DISCO. Move this meeting focus back to that part of list.

Discussion on current process – we are providing notes scoping definitions, but not writing the final definitions; that will be next pass.

Discovery: need a definition to be clear on DDI usage. Need to find a source for definition. Priority is usage in DDI context—find good text on DDI-RDF vocabulary web page. Scope of DDI descriptions is wider than just datasets, studies might be scope of discovery as well (other resources?)

Dissemination: from lifecycle. Term is used frequently. Have generic definitions, then clarify that in DDI

DTD: Document type definition. Superseded by XML schema. ?remove from glossary? Is it completely out of use. Dan leans towards including for historical purposes. Term is used on web site, not clear if DTD docs are still available.

Genericode: “OASIS Code List Representation format, “genericode”, is a single model and XML format (with a W3C XML Schema) that can encode a broad range of code list information.” Need to determine if DDI codelists use this encoding, if not, remove term.

Identifiable: a class of things that have an identifier. Emerged from SDMX. Motivated by wanting to reference and reuse some info object.

 IHSN toolkit: International household survey network; application of DDI codebook. Toolkit is set of software tools for comparing codebook instances, and maintain metadata

Out of time, end here