Data Related Standards Session – Emerging Topics

• Standards are Critically Important and Literally Everywhere
  • Phil Bourne mentioned this am, as did
  • FDA Commissioner who put Data Standards as #3 on his slide ‘Generating Evidence to Inform Decisions’
  • Vivien Bonazzi mentioned the need to coordinate with standards related to the Commons.
  • Critically important to making data FAIR, citable and so forth

• Standards are an integral part of BD2K -- RFA to support community based standards work was released in the fall for May 2017 award.
Many Types of Standards

• Standards: Semantic, adding meaning to data
• Syntactic: e.g. Packaging and transmission formats
• De facto standards, e.g. GO
• SDO developed and managed, e.g. MedDRA, SNOMED, etc.
• More structured metadata standards to describe the elements of data and sometimes the associated permissible values or value sets, and forms composed of DE
• Similarly, standard measures, e.g. PhenX (Phenotypes and Exposures)
• Minimum Information for experiments standards e.g. MIAME ++
Many Types and MANY standards, such as

• Data Formats – like HDF, used in many non health disciplines
• Standards for describing and using scientific workflows, like the Common Workflow Language [CWF]
• Standards for describing and transmitting radiologic images, e.g widely used DICOM
• Even standard programming interfaces like the GA4GH API for making genomic information available or retrieving it in a standard way.
• Literally more than a thousand in biomedical science
Speakers: Addressing several topics in data related standards

• Dr. Susanna Sansone [University of Oxford/ NPG] – re: Biosharing, which can help a user to make sense of the many standards and find the most appropriate.

• Dr. Mark Musen [Stanford], PI for BD2K CEDAR center, re: using CEDAR to help create and reuse standard data elements and templates or forms. He also leads widely used Bioportal terminology resource.

• Dr. Julie McMurray - [OHSU], involved in community efforts related to data and resource identifiers re: an approach that could help make sense of the identifiers jungle, called Prefix Commons.

• Dr. Melissa Haendel - [also from OHSU] PI for Monarch Initiative, involved in other activities involving standards (e.g. NCATS TransMed Data Translation Project) – re: a community based effort to develop a standard to Exchange phenotype data - and cleverly nick-named 'Phenopackets'.