Why Standards? What Standards? How Standards?
Making Real-World Data Useful in Healthcare / Life Sciences

We can help achieve interoperability and usefulness with multiple data sources, requirements and vendors by harmonizing and integrating data within open data description standards and frameworks. This overcomes a huge historical problem, where expensive and valuable data is buried and lost within traditional databases and vendor software applications. Melissa Informatics (MI) is often tasked by customers to review and curate existing standards and to apply one or multiple standards to meet different requirements. Importantly, MI's Sentient software system is designed to rapidly align data with different, potentially changing standards. The ability to work with standards in an agile manner is required in an open world, where data and requirements change over time.

Several important standards are listed below as examples that MI's Sentient software system and data science team has worked with in the past few months. This is a small subset of the many standards MI is familiar with and can 'mix and match' more efficiently than any other data harmonization vendor*. All of these standards are available in MI's preferred, open, RDF data description format (defined by the W3C global standards body) and can be applied by MI 'out of the box':

Broad "Dictionaries of Vocabularies" – also called metathesauri, terminology compendia

- CDISC (Clinical Data Interchange Standards Consortium) - *CDISC standardization is required for regulatory submissions to FDA - superset to SDTM below)*
- UMLS (Unified Medical Language System) - *integrates and distributes key terminology, classification and coding standards*

Subject Specific Vocabularies / Terminologies – some of these are contained, for example, in UMLS

- ICD10 (International Statistical Classification of Diseases and Related Health Problems)
- LOINC (Logical Observation Identifiers Names and Codes)
- MedDRA (Medical Dictionary for Regulatory Activities)**
- MeSH (Medical Subject Headings)

Standards for Reporting Data – different standards are required for different needs

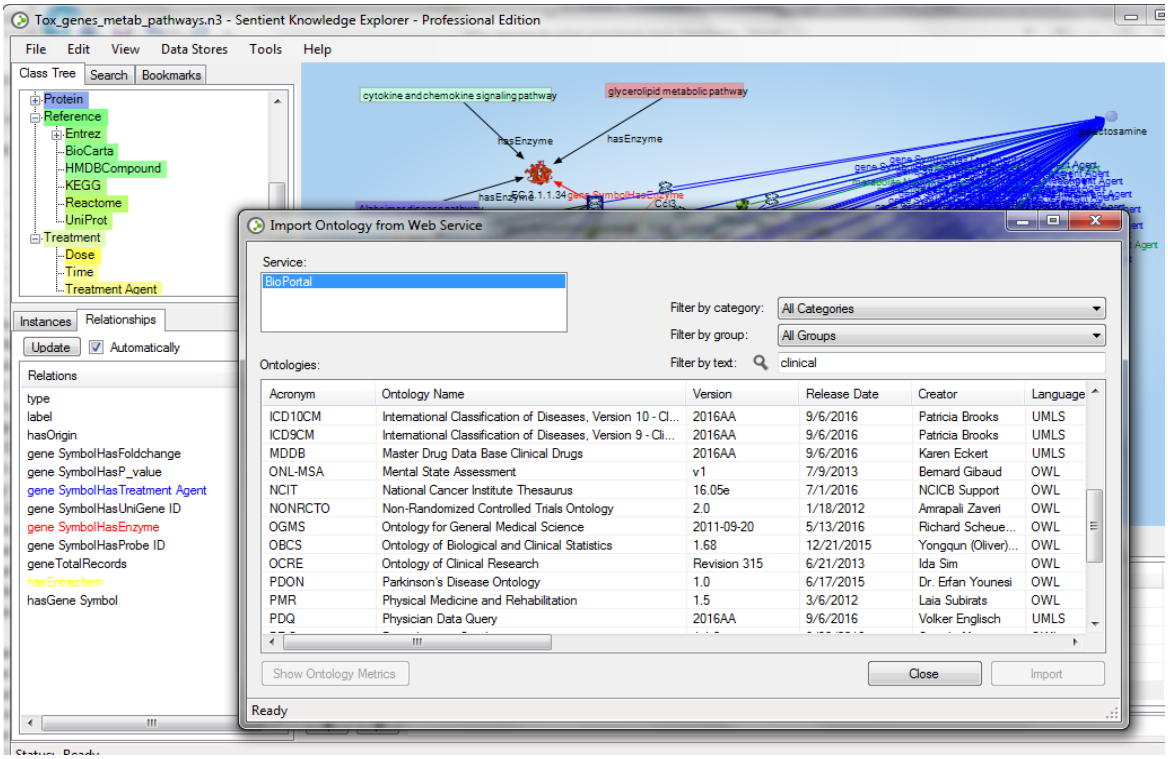
- ADaM (Analysis Dataset Model) - *standardized model for analysis of clinical data*
- SDTM (Study Data Tabulation Matrix*) – required for FDA submissions*

Data Exchange Standards – useful for information publication and interchange

- FAIR (Findability, Accessibility, Interoperability, Reusability) – *Guidelines for data publishers*
- FHIR*** (Fast Healthcare Interoperability Resources) - *Defined by HL7 group*

An Open World of Standards

There are many other types and domains for standards, with overlapping content and meaning. For example, standards like "MIAME" and ontologies like "GO" can be applied for microarray and genetic data. Standards will come, go and evolve over time. For big data harmonization and use it is *critical* to apply a software environment that is capable of consuming and aligning data with different standards as needed, without refactoring the entire integrated data resource. MI's Sentient suite makes this possible.



The figure above shows the MI Sentient Knowledge Explorer software connected to Stanford University's "BioPortal", one of the best resources for ontologies (comprehensive descriptions of terms and relationships within a domain of interest) and related data definition standards.

The Sentient system can efficiently apply these standards in an agile manner, to align and re-align data to standards as needed for different requirements. This guarantees long-term interoperability and usefulness for your data. Now it is up to you, to bring value from your data to the market!

Notes:

** Many standards are incomplete, changing or inadequate to customer needs, and therefore require curation and / or enrichment before use. Science is an open field, so standards can't be static. MI can bridge the gap between changing standards and consistently defined, usable data.*

*** Some standards require payment to use.*

** FHIR is only available in draft format but IS available as RDF that may be efficiently reviewed, curated and applied by MI. It is important to make best use of the best standards available in a domain of interest – even if they are incomplete. Ironically, standards are seldom if ever perfect and complete.*