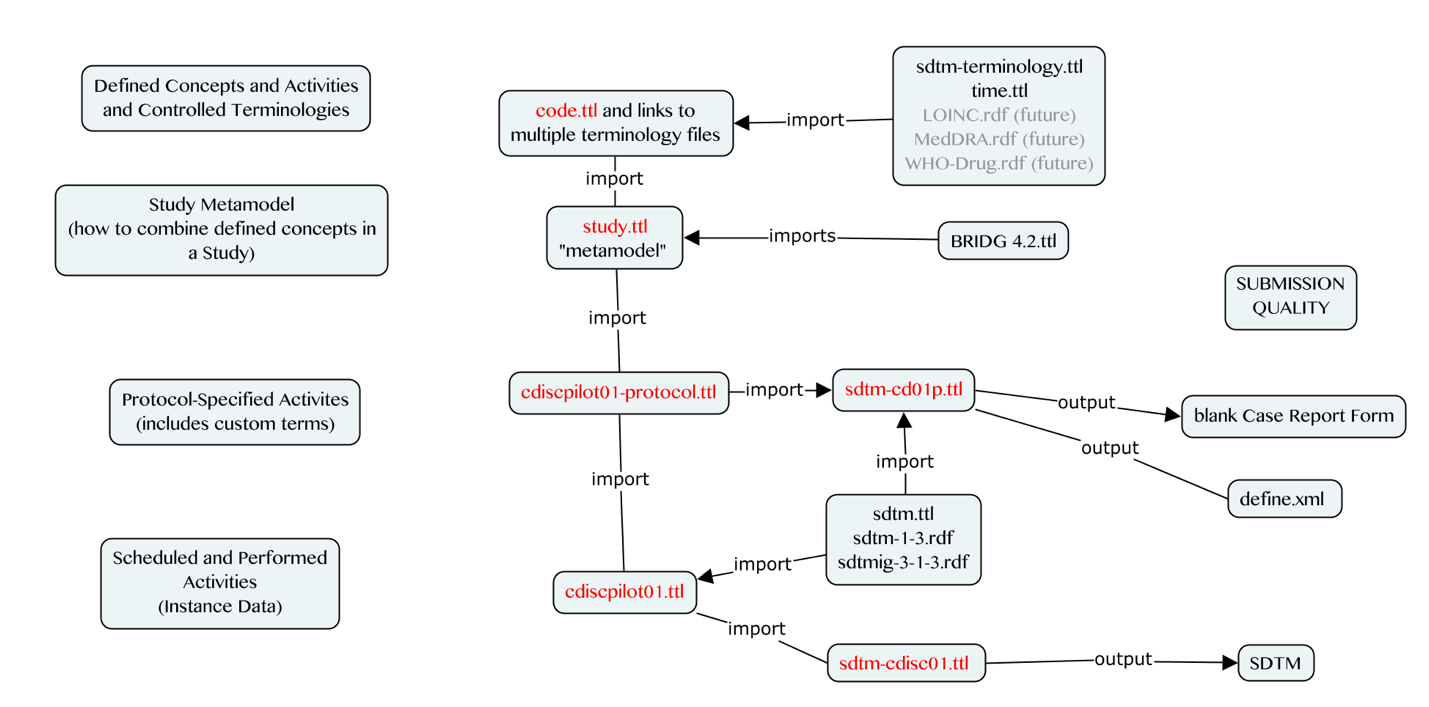
**Ontology Roadmap**

Clinical Trial Data as RDF

This file is an introduction to the various files located in the GitHub “…data/RDF” folder. Figure 1 describes each file and how it is related to other files

Figure : Layering of Study Data



1. the main file is study.ttl (namespace = study: ) It contains the OWL classes and predicates to represent a single study in RDF. It is a metamodel in the sense that instantiation of study.ttl results in an ontology for a single study. study.ttl is heavily BRIDG-based but not entirely as it attempts to better align with how clinical data are generated and used in health care.
2. The study ontology file (study.ttl) imports code.ttl, which contains or links to important biomedical concepts that are managed by outside third parties (e.g. CDISC, W3C, and in the future, MedDRA, LOINC, WHO Drug Dictionary, CIMI). namespace= code:
3. The cdiscpilot01-protocol.ttl imports the study.ttl file and contains protocol-specified activities and rules. Whereas study.ttl describes generic activities and observations in a study, this file describes the specific observations for a particular study and the rules for when to perform them.
4. The cdiscpilot01.ttl file imports the protocol file above (#3) and adds instance data for the first 3 subjects in the pilot study
5. To get the data out in a format that we want (e.g. SDTM) we need to link all this information to an SDTM ontology. this is accomplished at the protocol level via the sdtm-cd01p.tll file, from which one can automate the creation of define.xml and eventually the blank case report form
6. At the instance level, we take the instance data (cdiscpilot01.ttl) and link it to the sdtm ontology (see sdtm-cdisc01.ttl). This file is used to generate high quality, highly standardized SDTM domains (DM, VS, and SUPPDM) using embedded SPARQL queries (i.e. spin: rules). Note that the instance data file also contains limited exposure information needed to derived certain DM variables (RFXSTDTC and RFXENDTC) although representing full exposure data to generate the EX domain is currently out of scope.