



CLINICAL DATA INTERCHANGE STANDARDS CONSORTIUM

*The CDISC vision is to inform patient care & safety
through higher quality medical research.*

A decorative graphic consisting of two overlapping wavy lines, one dark blue and one light green, that extend from the left side of the slide and terminate in a horizontal bar with a diagonal hatched pattern.

Strength *through Collaboration*

CDISC Standards RDF Reference Guide

Presented by Frederik Malfait and Scott Bahlavooni
Acknowledgements to Geoff Low and Mitra Rocca



Strength *through Collaboration*

Agenda

- PhUSE Computational Science Symposium
 - § Semantic Technology Working Group
- Introduction and Rationale
- W3C Resource Description Framework (RDF)
- CDISC Foundational Standards
 - § RDF Schemas
 - § RDF Datasets
- RDF SDTM IG Walk-through
- Accessing GitHub

PhUSE Computational Science Symposium

Semantic Technology Working Group



PhUSE CSS Collaboration

- Mission:
 - § *“...bring together academia, industry, technology providers and the FDA to collaborate on projects to address unmet computational science needs.”*
- Working Groups:
 - § Optimizing the Use of Data Standards
 - § Development of Standard Scripts for Analysis and Programming
 - § Non-Clinical Roadmap and Impact on Implementation
 - § Emerging Technologies
 - § Semantic Technology

Semantic Technology Working Group

Investigate the application of W3C semantic standards to support the clinical and non-clinical data life-cycle from protocol development to submission to regulatory agencies.

Semantic Technology Teams

- Semantic Technology Primer *
- Representation of CDISC Foundational Standards in RDF *
- Protocol and Study Design Representation in RDF
- Representation of Analysis Metadata to Support Clinical and Non-Clinical Applications
- Representation of Regulations and Guidance in RDF
- Representation of CDISC Conformance Checks in RDF *
- keyCRF: Reusing Medical Summaries for Enabling Clinical Research

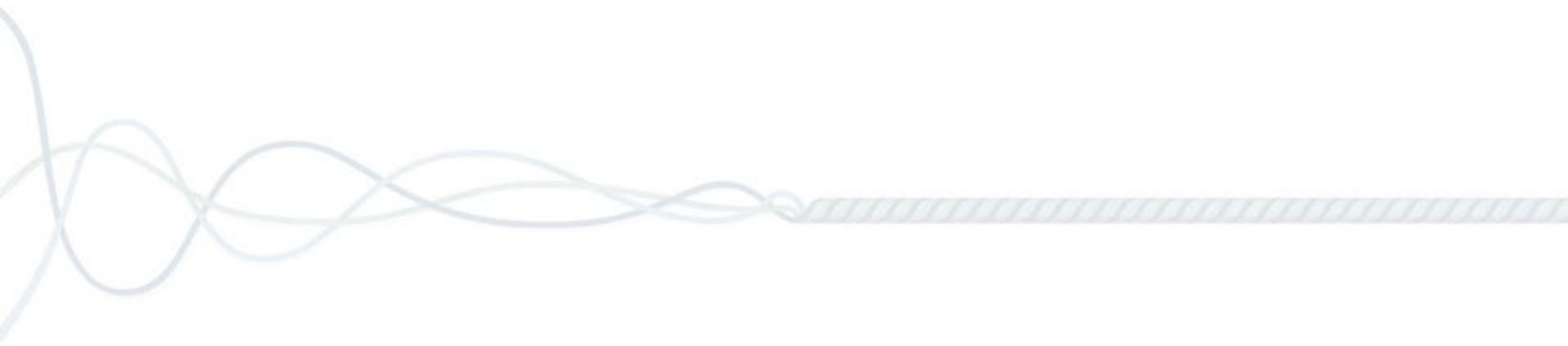
* Completed project

Acknowledgements

CDISC Foundational Standards in RDF

Phil Ashworth, Scott Bahlavooni, Daniel Boisvert,
Susan DeHaven, Nathan Freimark, Josephine-Anne Gough,
Laura Hollink, Dave Jordan, Ron Katriel, Kirsten Langendorf,
Geoff Low, Frederik Malfait, Mitra Rocca, Gary Walker.

Introduction and Rationale



Recall CDISC Mission

The CDISC mission is to develop and support global, **platform-independent data standards** that enable **information system interoperability** to improve medical research and related areas of healthcare.

Recall CDISC Principles

- Recognize the ultimate goal of creating regulatory submissions that allow for flexibility in scientific content and are **easily interpreted, understood, and navigated** by regulatory reviewers.
- Acknowledge that the data content, structure and quality of the standard data models are of paramount importance, **independent of implementation strategy and platform**.
- Work with other professional groups to encourage that there is **maximum sharing of information** and minimum duplication of efforts.

Current State

Bookmarks

- Cover Page
- Table of Contents
- 1 Introduction
 - 1.1 Purpose
 - 1.2 Relationship to Prior CDISC Models
 - 1.3 Significant Changes from Prior Versions
- 2 Model Fundamentals
 - 2.1 Model Concepts and Terms
 - 2.2 The General Observation Classes
 - 2.2.1 The Interventions Observations Class

CDISC SDTM (Version 1.4)

PDF Portfolios at a Glance

Cover Page, Revision History, TOC

Section 1 - Introduction

Section 2 - Fundamentals of the ...

Section 3 - Submitting Data in Standard For...

Section 4 - Assumptions For Do...

Cover Page, Revision History, TOC.pdf Open File

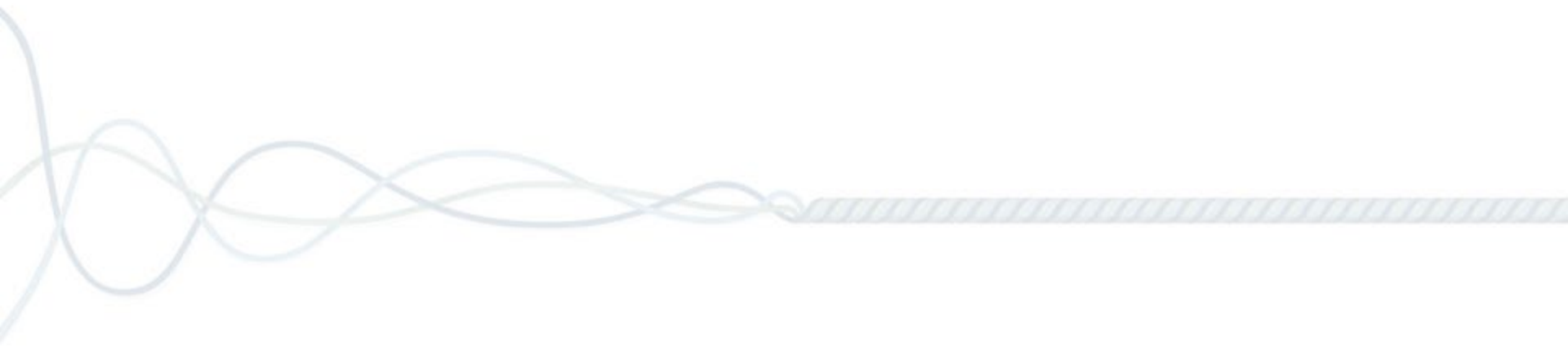
**Study Data Tabulation Model
Implementation Guide:
Human Clinical Trials
Version 3.2**

	Seq. For	Observation Class	Domain	Variable Name (minus domain prefix)	Variable Name	Variable Label	Typ	Controlled Terminology	Role	CDISC Notes (for domains) Description (for General Classes)	Col	References
1	1	Interventions-General		TRT	--TRT	Name of Treatment	Char		Topic	The topic for the intervention observation, usually the verbatim name of the treatment, drug, medicine, or therapy given during the dosing interval for the observation.		
2	2	Interventions-General		MODIFY	--MODIFY	Modified Treatment Name	Char		Synonym Qualifier of --TRT	If the value for --TRT is modified for coding purposes, then the modified text is placed here.		
3	3	Interventions-General		DECOD	--DECOD	Standardized Treatment Name	Char		Synonym Qualifier of --TRT	Standardized or dictionary-derived name of the topic variable, --TRT, or the modified topic variable (--MODIFY), if applicable. Equivalent to the generic drug name in WHO Drug, or a term in SNOMED, ICD9, or other published or sponsor-defined dictionaries.		
4	4	Interventions-General		CAT	--CAT	Category	Char		Grouping Qualifier	Used to define a category of topic-variable values.		
5	5	Interventions-General		SCAT	--SCAT	Subcategory	Char		Grouping Qualifier	Used to define a further categorization of --CAT values.		
6	6	Interventions-General		PRESPEC	--PRESPEC	Pre-specified	Char		Variable Qualifier of --TRT	Used when a specific intervention is pre-specified on a CRF. Values should be "Y" or null.		
7	7	Interventions-General		OCCUR	--OCCUR	Occurrence	Char		Record Qualifier	Used to record whether a pre-specified intervention occurred when information about the occurrence of a specific intervention is solicited.		
8	8	Interventions-General		STAT	--STAT	Completion Status	Char		Record Qualifier	Used to indicate when a question about the occurrence of a pre-specified intervention was not answered. Should be null or have a value of NOT DONE.		
9	9	Interventions-General		REASON	--REASON	Reason Not Done	Char		Record Qualifier	Reason not done. Used in conjunction with --STAT when value is NOT DONE.		
10												

W3C Semantic Standards

- Can express a wide range of information
 - § Meta-models, models, data
- Consistent language and modeling framework
- Formal, computable, executable
- Identical at design and run-time
- Designed for
 - § Platform independence
 - § Semantic interoperability
 - § Sharing and linking information
- Proven backbone of the semantic web
- Acronyms: RDF, RDFS, OWL, SKOS, SPARQL

W3C Resource Description Framework (RDF)



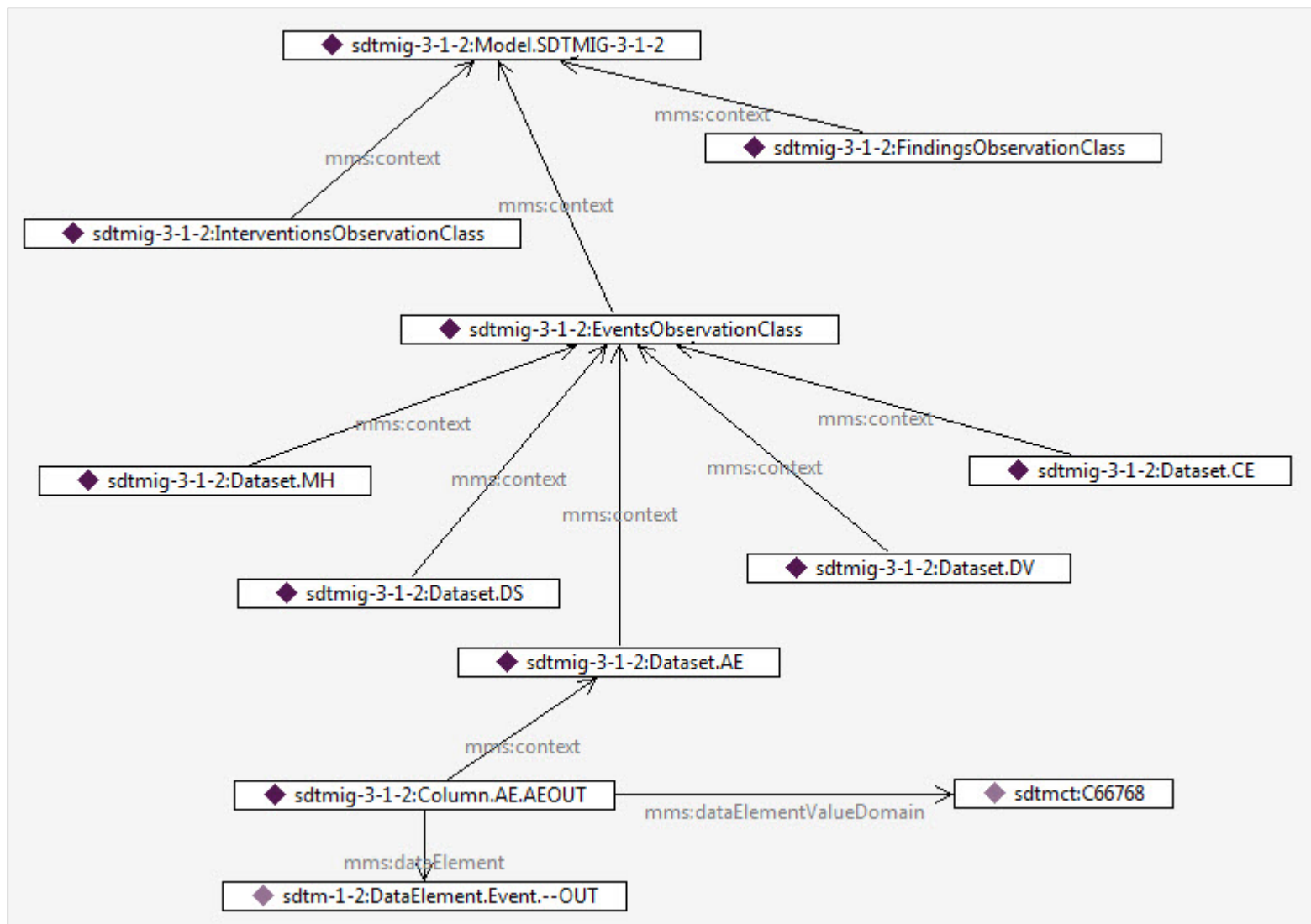
Resources

- A resource is anything we like to talk about
- Uniform Resource Identifier
 - § <http://rdf.cdisc.org/std/sdtmig-3-1-2#Column.AE.AEOUT>
- Namespaces
 - § `sdtmig-3-1-2: "http://rdf.cdisc.org/std/sdtmig-3-1-2#"`
- Qualified Names
 - § `sdtmig-3-1-2:Column.AE.AEOUT`
- Creates globally unique identifiers
- A representation of a resource can be made available over a network or the web

Triples

- Statements about resources, e.g. attributes
 - § Subject: `sdtmig-3-1-2:Column.AE.AEOUT`
 - § Predicate: `mms:dataElementName`
 - § Object: `"AEOUT"`
- Statements about resources, e.g. relationships
 - § Subject: `sdtmig-3-1-2:Column.AE.AEOUT`
 - § Predicate: `mms:dataElementValueDomain`
 - § Object: `sdtmct:C66768`

Graphs



Schemas and Ontologies

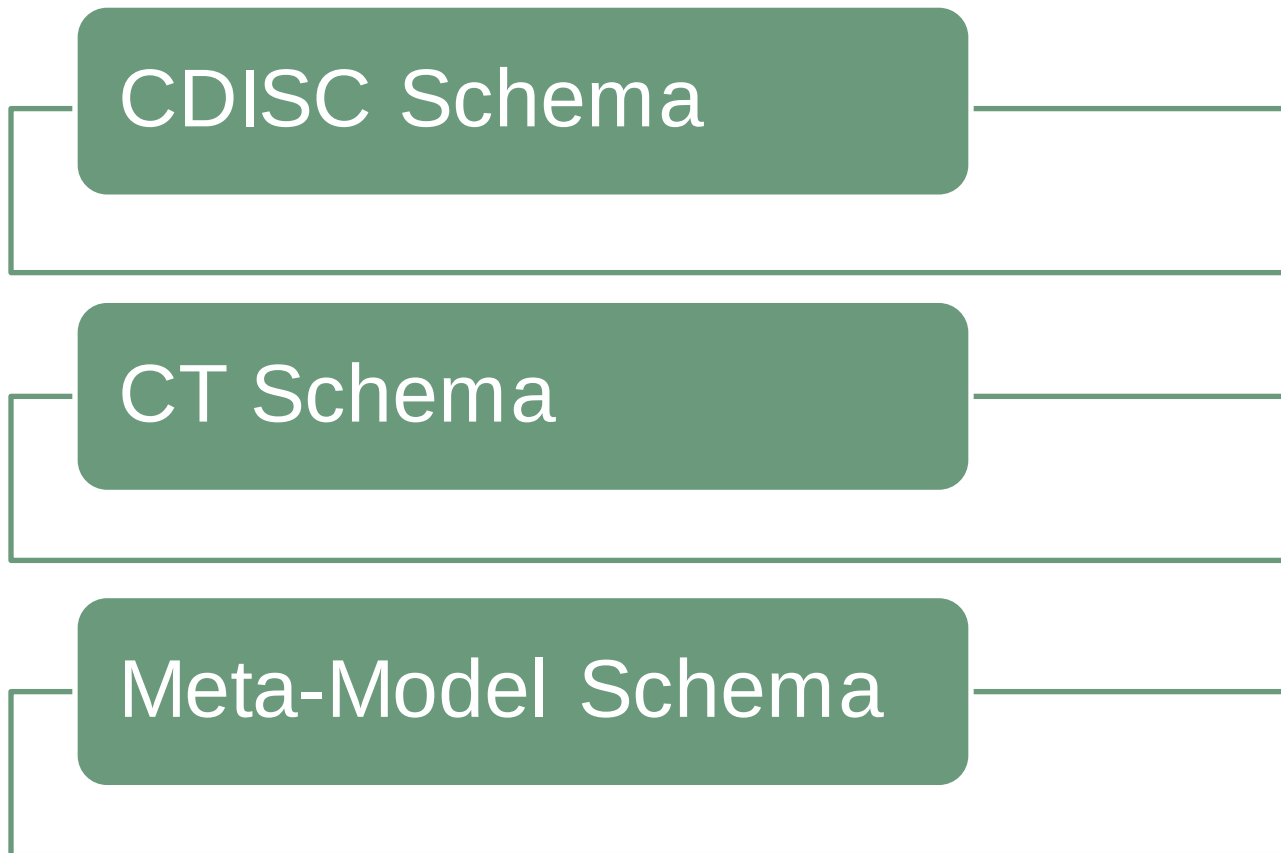
- A set of resources and predicates that defines a vocabulary or ontology
- Resources can be organized in classes
- Predicates and classes are also resources
- W3C defined schemas
 - § Resource Description Framework (RDF)
 - § RDF Schema (RDFS)
 - § Web Ontology Language (OWL)
 - § Simple Knowledge Organization (SKOS)

CDISC Foundational Standards

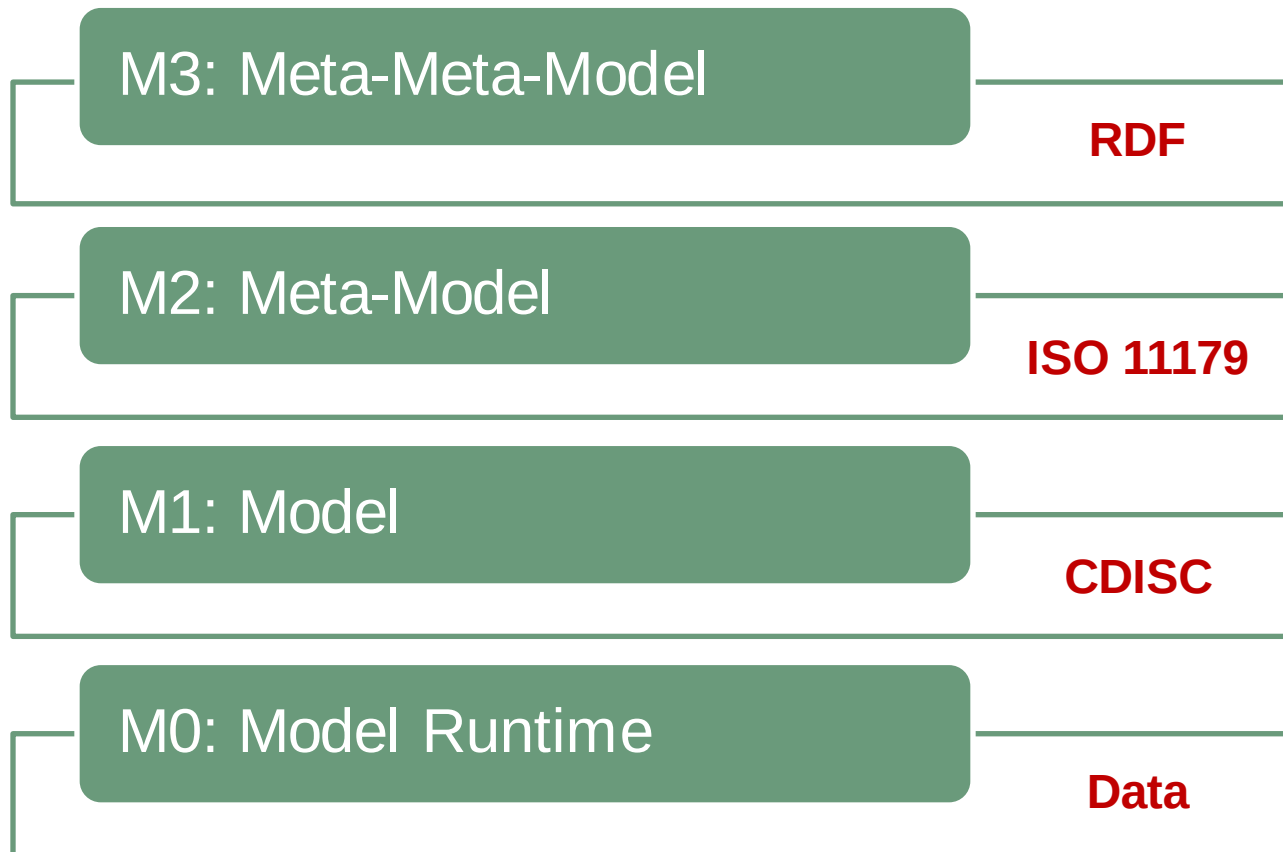
RDF Schemas



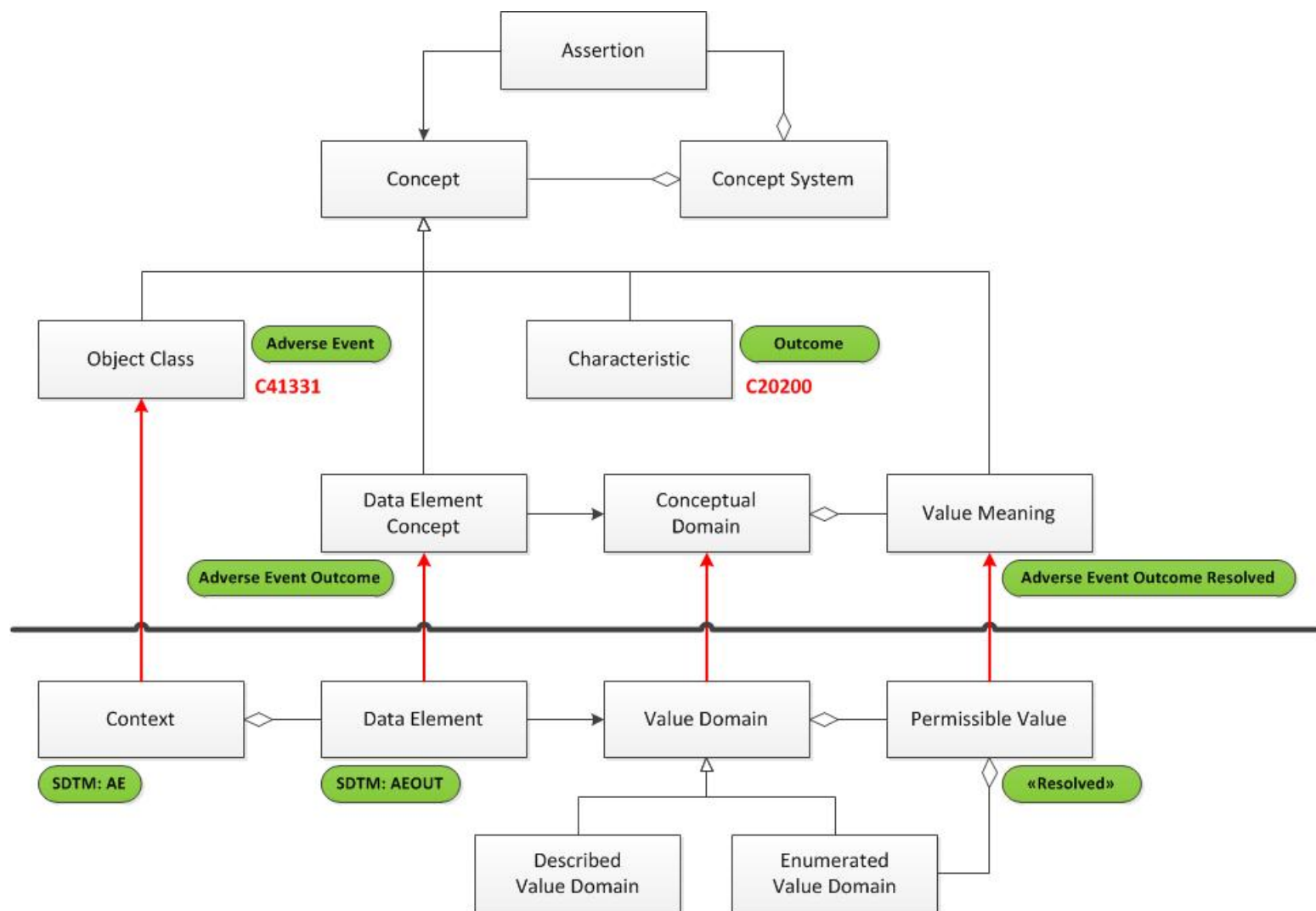
Layered Schemas



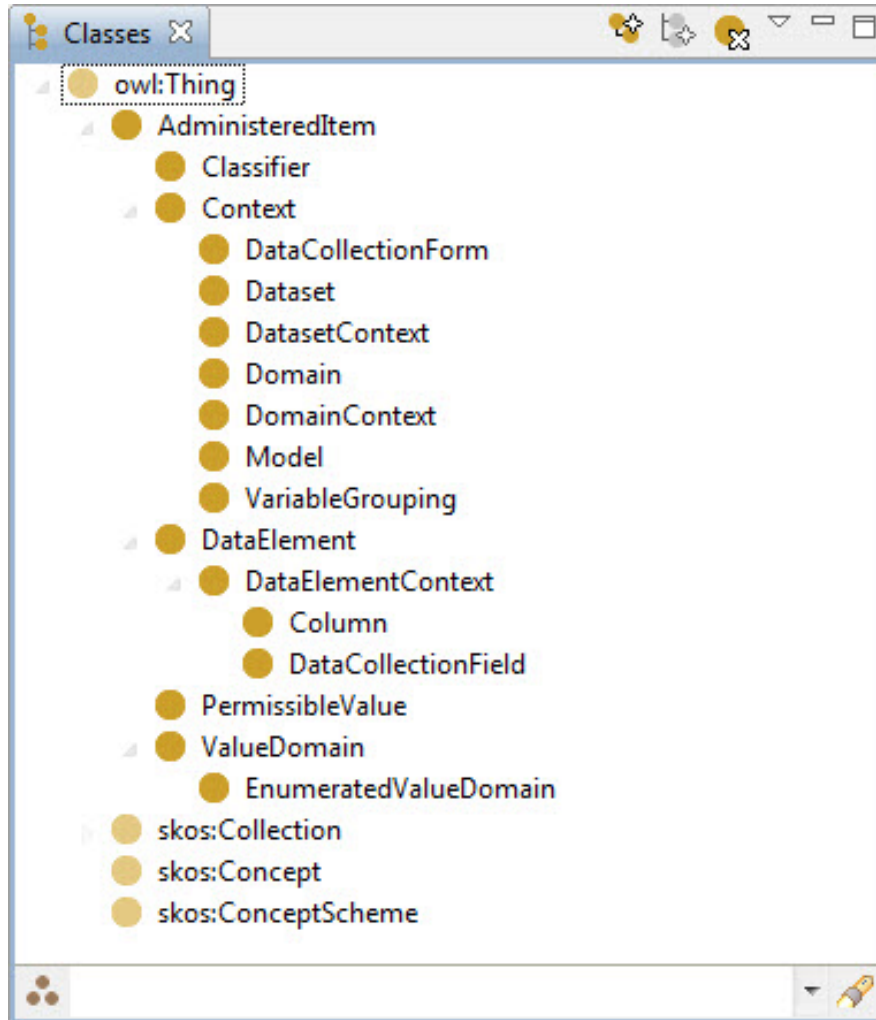
Meta-Model Hierarchy



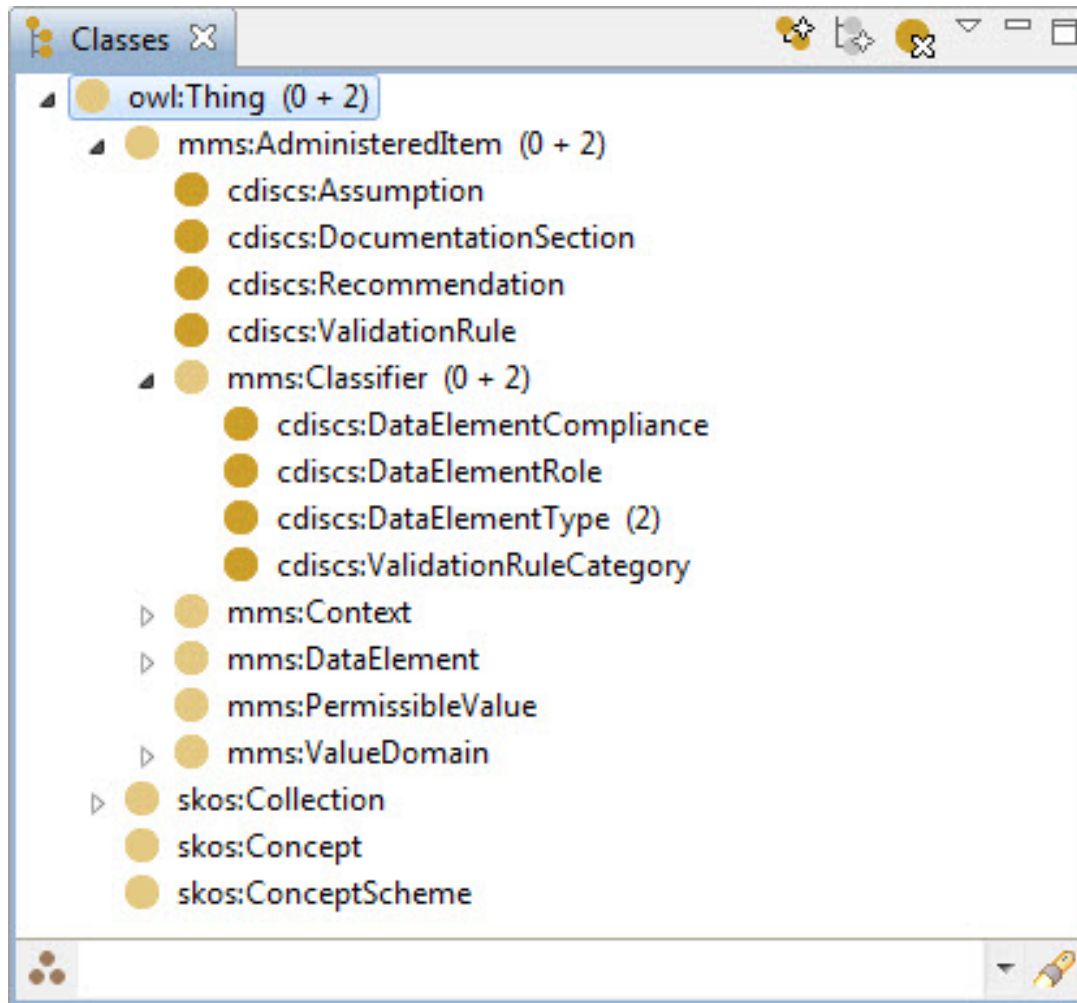
ISO 11179 Metadata Registry Std



OWL Classes: Meta-Model Schema



OWL Classes: CDISC Schema



CDISC Foundational Standards

RDF Datasets



Accessing GitHub

- The first release of the Foundational Standards in RDF is available on **GitHub**

<https://github.com/phuse-org/rdf.cdisc.org>

- To access the code:

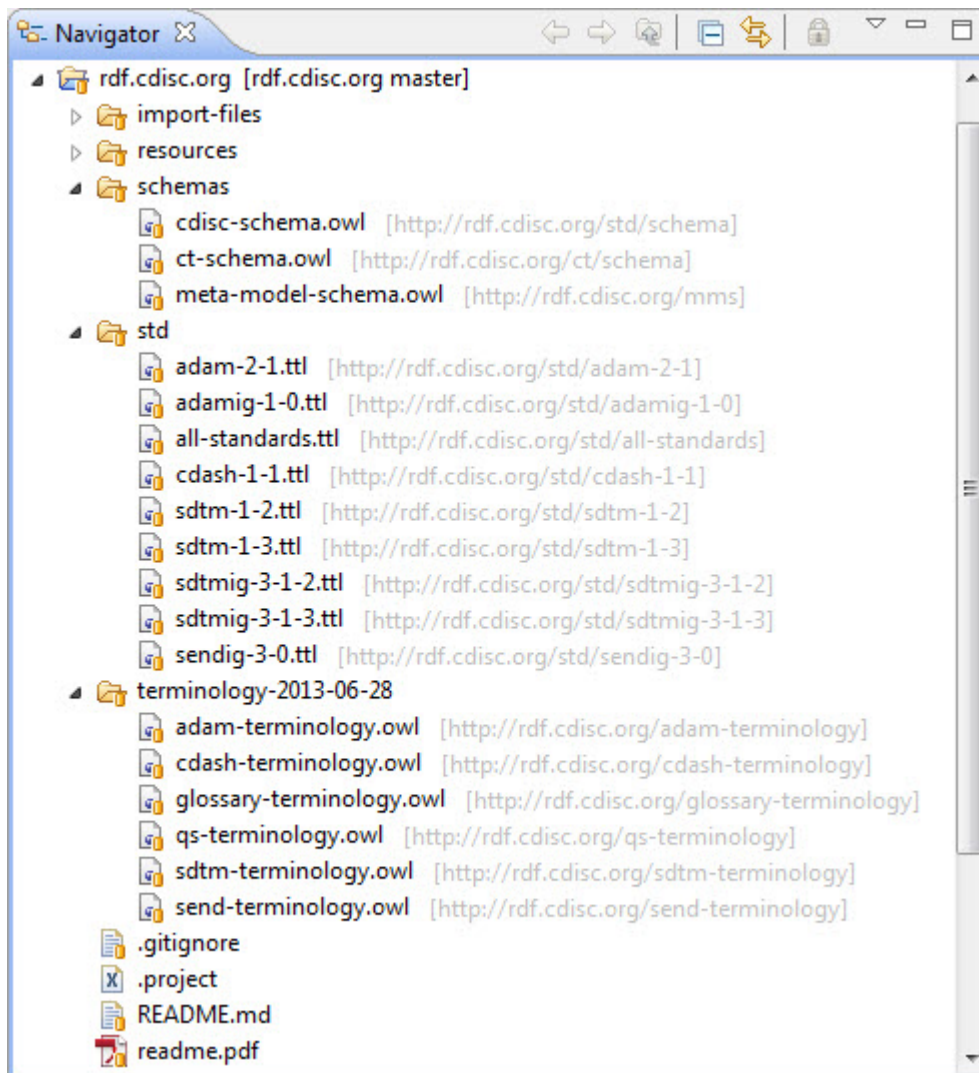
§ Download a zip archive

 Download ZIP

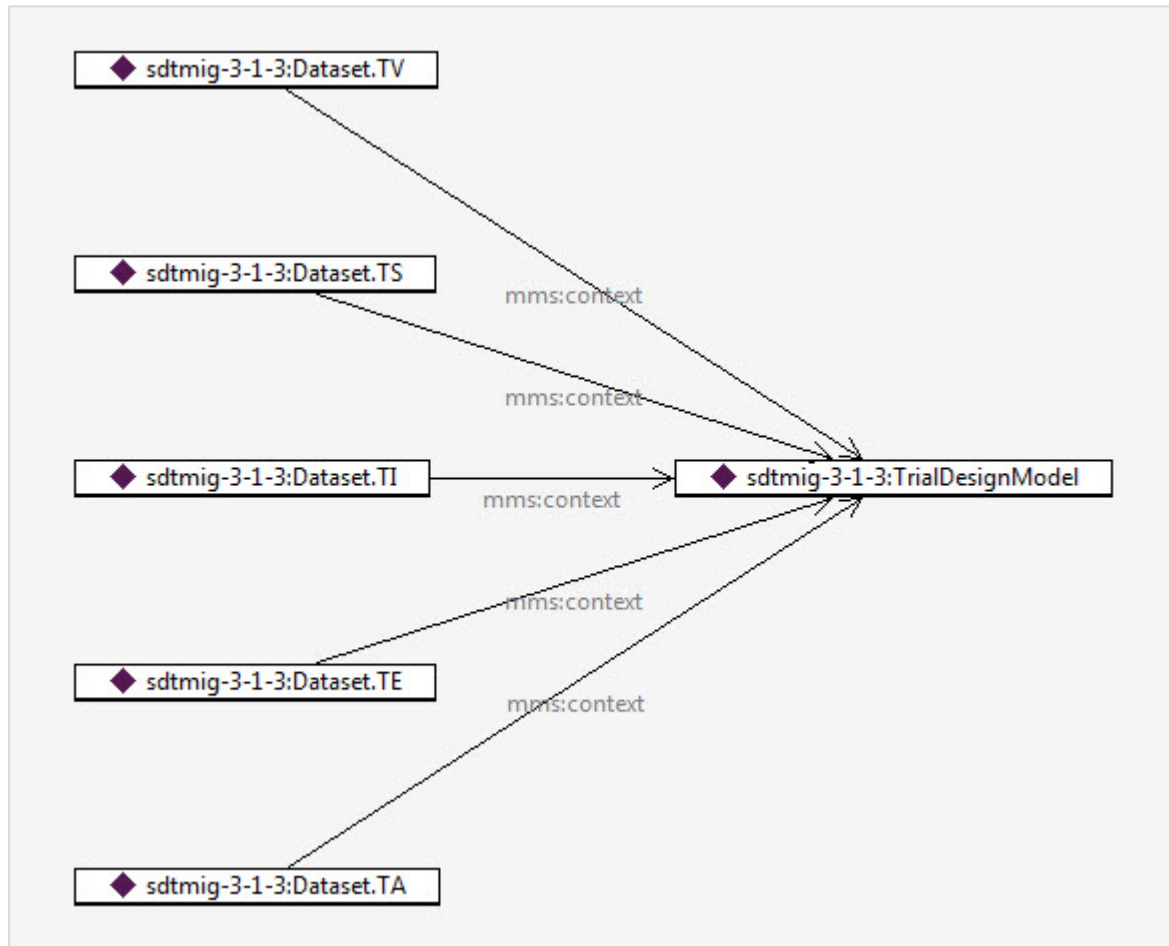
§ Clone the project

- Guidance in the RDF Reference Guide
- Contact Geoff Low: glow@mdsol.com

Project Structure



Graph Representation



Resource Form

sdtmig-3-1-3.ttl

Resource Form

Name: sdtmig-3-1-3:TrialDesignModel

Annotations

Incoming References

← mms:context

- sdtmig-3-1-3:Dataset.TA
- sdtmig-3-1-3:Dataset.TE
- sdtmig-3-1-3:Dataset.TI
- sdtmig-3-1-3:Dataset.TS
- sdtmig-3-1-3:Dataset.TV

Other Properties

mms:context

sdtmig-3-1-3:Model.SDTMIG-3-1-3

mms:contextDescription

S Datasets to describe the design of a trial.

mms:contextLabel

S Trial Design Domain

mms:contextName

S TrialDesign

mms:ordinal

6

rdf:type

mms:DatasetContext

Form Browser Graph Source Code

Navigating Across Resources

The screenshot displays the CDISC data browser interface. On the left, the 'Incoming References' panel shows a list of resources under the 'mms:context' dropdown. The selected resource is 'sdtmig-3-1-3:Dataset.TA'. A tooltip indicates 'Type: mms:Dataset (Hold CTRL to navigate)'. The main window shows the 'Resource Form' for the selected resource, with the URI 'http://rdf.cdisc.org/std/sdtmig-3-1-3#Dataset.TA'. The form is divided into two columns: 'Annotations' and 'Other Properties'. The 'Annotations' column shows a list of incoming references under the 'mms:context' dropdown, including 'sdtmig-3-1-3:Column.TA.ARM', 'sdtmig-3-1-3:Column.TA.ARMCD', 'sdtmig-3-1-3:Column.TA.DOMAIN', 'sdtmig-3-1-3:Column.TA.ELEMENT', 'sdtmig-3-1-3:Column.TA.EPOCH', 'sdtmig-3-1-3:Column.TA.ETCD', 'sdtmig-3-1-3:Column.TA.STUDYID', 'sdtmig-3-1-3:Column.TA.TABRANCH', 'sdtmig-3-1-3:Column.TA.TAETORD', and 'sdtmig-3-1-3:Column.TA.TATRANS'. The 'Other Properties' column shows various properties including 'mms:context' (sdtmig-3-1-3:TrialDesignModel), 'mms:contextLabel' (Trial Arms), 'mms:contextName' (TA), 'mms:ordinal' (29), 'cdiscs:datasetCode' (TA), 'cdiscs:datasetStructure' (One record per planned Element per Arm), and 'rdf:type' (mms:Dataset). The bottom of the window has tabs for 'Form', 'Browser', 'Graph', and 'Source Code'.

Incoming References

- mms:context ▾
- ◆ sdtmig-3-1-3:Dataset.TA
- ◆ sdtmig-3-1-3:Dataset.TS
- ◆ sdtmig-3-1-3:Dataset.TV

Type: mms:Dataset
(Hold CTRL to navigate)

Resource Form

URI: <http://rdf.cdisc.org/std/sdtmig-3-1-3#Dataset.TA>

Annotations

Incoming References

- mms:context ▾
- ◆ sdtmig-3-1-3:Column.TA.ARM
- ◆ sdtmig-3-1-3:Column.TA.ARMCD
- ◆ sdtmig-3-1-3:Column.TA.DOMAIN
- ◆ sdtmig-3-1-3:Column.TA.ELEMENT
- ◆ sdtmig-3-1-3:Column.TA.EPOCH
- ◆ sdtmig-3-1-3:Column.TA.ETCD
- ◆ sdtmig-3-1-3:Column.TA.STUDYID
- ◆ sdtmig-3-1-3:Column.TA.TABRANCH
- ◆ sdtmig-3-1-3:Column.TA.TAETORD
- ◆ sdtmig-3-1-3:Column.TA.TATRANS

Other Properties

- mms:context ▾
- ◆ sdtmig-3-1-3:TrialDesignModel
- mms:contextLabel ▾
- S Trial Arms
- mms:contextName ▾
- S TA
- mms:ordinal ▾
- 29
- cdiscs:datasetCode ▾
- S TA
- cdiscs:datasetStructure ▾
- S One record per planned Element per Arm
- rdf:type ▾
- mms:Dataset

Form | Browser | Graph | Source Code

RDF Text Serialization

```
sdtmig-3-1-3:Dataset.TA
  a
  mms:context          sdtmig-3-1-3:TrialDesignModel ;
  mms:contextLabel     "Trial Arms"^^xsd:string ;
  mms:contextName      "TA"^^xsd:string ;
  mms:ordinal           "29"^^xsd:positiveInteger ;
  cdiscs:datasetCode    "TA"^^xsd:string ;
  cdiscs:datasetStructure "One record per planned Element per Arm"^^xsd:string .
```


RDF Query Language

Query Editor

Query Library

```

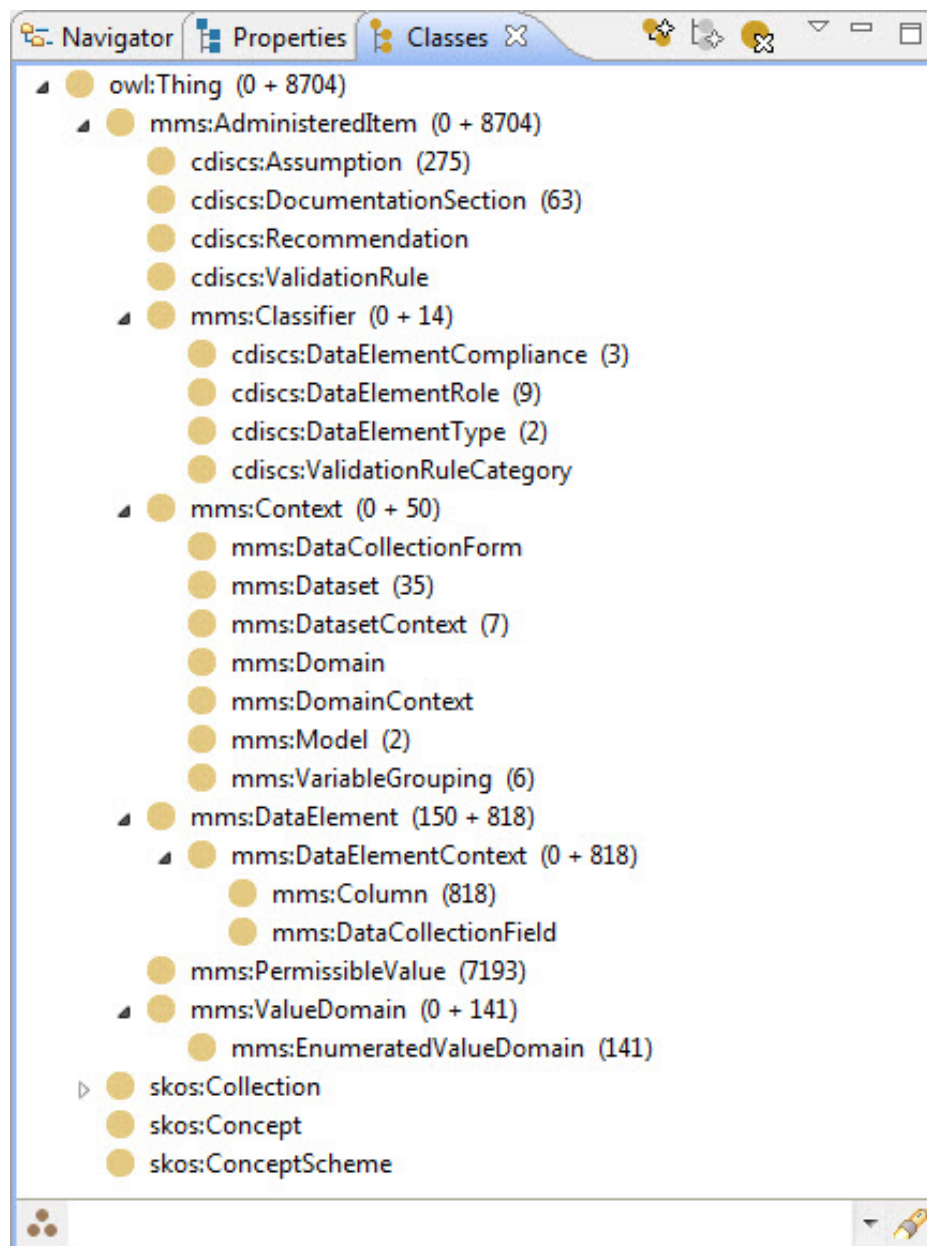
# Find all SDTM variables with codelist C71113
PREFIX mms: <http://rdf.cdisc.org/mms#>
PREFIX sdtmct: <http://rdf.cdisc.org/sdtm-terminology#>
SELECT ?domain ?variable ?label
WHERE {
  ?dataElement mms:dataElementValueDomain sdtmct:C71113 .
  ?dataElement mms:context/mms:contextLabel ?domain .
  ?dataElement mms:dataElementName ?variable .
  ?dataElement mms:dataElementLabel ?label .
}

```

[domain]	variable	label
 Concomitant Medications	 CMDOSFRQ	 Dosing Frequency per Interval
 Exposure	 EXDOSFRQ	 Dosing Frequency per Interval
 Substance Use	 SUDOSFRQ	 Use Frequency Per Interval

SDTM IG Walk-through





Instances ✕	
[Resource]	mms:contextLabel
◆ sdtm-1-3:Model.SDTM-1-3	Study Data Tabulation Model (SDTM) Version 1.3
◆ sdtmig-3-1-3:Model.SDTMIG-3-1-3	Study Data Tabulation Model Implementation Guide (SDTMIG) Version 3.1.3

sdtmig-3-1-3.ttl

Resource Form

URI:

Annotations

Incoming References

← mms:context

- sdtmig-3-1-3:EventsObservationClass
- sdtmig-3-1-3:FindingsAbout
- sdtmig-3-1-3:FindingsObservationClass
- sdtmig-3-1-3:InterventionsObservationClass
- sdtmig-3-1-3:RelationshipDataset
- sdtmig-3-1-3:SpecialPurposeDomain
- sdtmig-3-1-3:TrialDesignModel

Other Properties

mms:contextDescription

The Study Data Tabulation Model Implementation Guide is a CDISC defined guide for the implementation of SDTM providing a detailed specification of the SDTM domains.

mms:contextLabel

Study Data Tabulation Model Implementation Guide (SDTMIG) Version 3.1.3

mms:contextName

sdtmig-3-1-3

rdf:type

mms:Model

Form
Browser
Graph
Source Code

sdtmig-3-1-3.ttl

Resource Form

URI:

Annotations

Incoming References

← mms:context

- ◆
- ◆
- ◆
- ◆
- ◆

Other Properties

mms:context

◆

mms:contextDescription

S The Events class captures planned protocol milestones such as randomization and study completion, and occurrences, conditions, or incidents independent of planned study evaluations occurring during the trial (e.g., adverse events) or prior to the trial (e.g., medical history).

mms:contextLabel

S

mms:contextName

S

mms:ordinal

3

rdf:type

●

Form
Browser
Graph
Source Code

sdtmig-3-1-3.ttl

Resource Form

URI:

▼ Annotations

▼ Other Properties

mms:context ▼
◆ sdtmig-3-1-3:EventsObservationClass ▼

mms:contextLabel ▼
S Adverse Events ▼

mms:contextName ▼
S AE ▼

mms:ordinal ▼
8 ▼

cdiscs:datasetCode ▼
S AE ▼

cdiscs:datasetStructure ▼
S One record per adverse event per subject ▼

rdf:type ▼
● mms:Dataset ▼

► Incoming References

Form | Browser | Graph | Source Code

← mms:context ▾
◆ sdtmig-3-1-3:Column.AE.AEACN
◆ sdtmig-3-1-3:Column.AE.AEACNOTH
◆ sdtmig-3-1-3:Column.AE.AEBDSYCD
◆ sdtmig-3-1-3:Column.AE.AEBODSYS
◆ sdtmig-3-1-3:Column.AE.AECAT
◆ sdtmig-3-1-3:Column.AE.AECONTRT

← cdiscs:documents ▾
◆ sdtmig-3-1-3:Section.AE.001
◆ sdtmig-3-1-3:Section.AE.002
◆ sdtmig-3-1-3:Section.AE.003
◆ sdtmig-3-1-3:Section.AE.004
◆ sdtmig-3-1-3:Section.AE.005
◆ sdtmig-3-1-3:Section.AE.006
◆ sdtmig-3-1-3:Section.AE.007
◆ sdtmig-3-1-3:Section.AE.008

← cdiscs:about ▾
◆ sdtmig-3-1-3:Assumption.AE.001.001
◆ sdtmig-3-1-3:Assumption.AE.001.009
◆ sdtmig-3-1-3:Assumption.AE.002.005
◆ sdtmig-3-1-3:Assumption.AE.004.002
◆ sdtmig-3-1-3:Assumption.AE.004.004
◆ sdtmig-3-1-3:Assumption.AE.005.002
◆ sdtmig-3-1-3:Assumption.AE.006.003
◆ sdtmig-3-1-3:Assumption.AE.006.005
◆ sdtmig-3-1-3:Assumption.AE.007.001
◆ sdtmig-3-1-3:Assumption.AE.008.001

Resource Form



URI:

▼ Annotations

▼ Other Properties

mms:context ▼

◆

mms:dataElement ▼

◆

mms:dataElementDescription ▼

Describes changes to the study treatment as a result of the event. AEACN is specifically for the relationship to study treatment. AEACNOTH is for actions unrelated to dose adjustments of study treatment. Examples of AEACN values include ICH E2B values: DRUG WITHDRAWN, DOSE REDUCED, DOSE INCREASED, DOSE NOT CHANGED, UNKNOWN or NOT APPLICABLE.

mms:dataElementLabel ▼

Action Taken with Study Treatment

mms:dataElementName ▼

AEACN

mms:dataElementType ▼

mms:dataElementValueDomain ▼

◆

mms:ordinal ▼

cdiscs:controlledTermsOrFormat ▼

(ACN)

cdiscs:dataElementCompliance ▼

◆

cdiscs:dataElementRole ▼

◆

cdiscs:dataElementType ▼

◆

cdiscs:references ▼

SDTM 2.2.2

rdf:type ▼

sdtmig-3-1-3.ttl

Resource Form

<http://rdf.cdisc.org/std/sdtm-1-3#Classifier.ExpectedVariable>

Annotations

rdfs:label

Expected Variable

skos:definition

An Expected variable is any variable necessary to make a record useful in the context of a specific domain. Expected variables may contain some null values, but in most cases will not contain null values for every record. When no data has been collected for an expected variable, however, a null column should still be included in the dataset, and a comment should be included in the define.xml to state that data was not collected.

Other Properties

rdf:type

cdiscs:DataElementCompliance

Incoming References

Form

Browser

Graph

Source Code

sdtmig-3-1-3.ttl

Resource Form

URI:

Annotations

Incoming References

← mms:dataElementValueDomain

◆ sdtmig-3-1-3:Column.AE.AEACN

← mms:inValueDomain

- sdmct:C66767.C17998
- sdmct:C66767.C48660
- sdmct:C66767.C49501
- sdmct:C66767.C49502
- sdmct:C66767.C49503
- sdmct:C66767.C49504
- sdmct:C66767.C49505

Other Properties

cts:cdiscDefinition

Terminology specifying changes to the study treatment as a result of an adverse event.

cts:cdiscSubmissionValue

ACN

cts:cdiscSynonyms

Action Taken with Study Treatment

cts:codelistName

Action Taken with Study Treatment

cts:isExtensibleCodelist

false

cts:nciCode

C66767

cts:nciPreferredTerm

CDISC SDTM Action Taken with Study Treatment Terminology

rdf:type

mms:EnumeratedValueDomain

Form
Browser
Graph
Source Code

sdm-3-1-3.ttl

Resource Form

URI:

Annotations

Incoming References

Other Properties

cts:cdiscDefinition

An indication that a medication schedule was modified by temporarily terminating a prescribed regimen of medication. (NCI)

cts:cdiscSubmissionValue

DRUG INTERRUPTED

cts:nciCode

C49501

cts:nciPreferredTerm

Drug Interrupted

mms:inValueDomain

sdmct:C66767

rdf:type

mms:PermissibleValue

Form Browser Graph Source Code

sdtmig-3-1-3.ttl

Resource Form

URI:

▼ Annotations

▼ Other Properties

mms:context ▼

mms:dataElementDescription ▼

mms:dataElementLabel ▼

mms:dataElementName ▼

mms:dataElementType ▼

mms:ordinal ▼

cdiscs:dataElementRole ▼

cdiscs:dataElementType ▼

cdiscs:supportedBySDTMIG ▼

cdiscs:supportedBySEND ▼

rdf:type ▼

▼ Incoming References

← mms:dataElement ▼

Form Browser Graph Source Code

sdtmig-3-1-3.ttl

Resource Form

URI:
http://rdf.cdisc.org/std/sdtm-1-3#DataElement.Timing.--DUR

Annotations

Incoming References

← mms:dataElement

sdtmig-3-1-3:Column.AE.AEDUR

sdtmig-3-1-3:Column.CM.CMDUR

sdtmig-3-1-3:Column.EX.EXDUR

sdtmig-3-1-3:Column.SU.SUDUR

Other Properties

mms:context

sdtmig-1-3:TimingVariables

mms:dataElementDescription

Collected duration of an event, intervention, or finding represented in ISO 8601 character format. Used only if collected on the CRF and not derived.

mms:dataElementLabel

Duration

mms:dataElementName

--DUR

mms:dataElementType

xsd:duration

mms:ordinal

22

cdiscs:dataElementRole

sdtmig-1-3:Classifier.TimingVariable

cdiscs:dataElementType

cdiscs:Classifier.Character

cdiscs:supportedBySDTMIG

true

cdiscs:supportedBySEND

true

rdf:type

mms:DataElement

Form
Browser
Graph
Source Code

sdtmig-3-1-3.ttl

Resource Form

URI:

Annotations

Incoming References

← **cdiscs:partOfSection**

- ◆ sdtmig-3-1-3:Assumption.AE.004.001
- ◆ sdtmig-3-1-3:Assumption.AE.004.002
- ◆ sdtmig-3-1-3:Assumption.AE.004.003
- ◆ sdtmig-3-1-3:Assumption.AE.004.004

Other Properties

mms:ordinal

4

cdiscs:documents

◆ sdtmig-3-1-3:Dataset.AE

cdiscs:sectionLabel

S Pre-Specified Terms; Presence or Absence of Events

rdf:type

cdiscs:DocumentationSection

Form | Browser | Graph | Source Code

sdtmig-3-1-3.ttl

Resource Form

URI:

▼ Annotations

▼ Other Properties

mms:ordinal ▼

cdiscs:about ▼

cdiscs:assumptionText ▼

cdiscs:partOfSection ▼

rdf:type ▼

▼ Incoming References

← cdiscs:about ▼

Form
Browser
Graph
Source Code