

CDISC SHARE & Patterns of Metadata Use

PhUSE SDE

May 14, 2015

Sam Hume <shume@cdisc.org>



Strength *through Collaboration*

What are the patterns of use in each case?

SHARE Category	Description & Examples
1. Existing Metadata	Metadata published for the foundational standards including content such as domains, variables, code lists, and their attributes. Examples: SDTM metadata spreadsheets
2. New Data Exchange Formats	New data exchange formats for representing metadata. Examples: RDF, ODM and Define-XML extensions, possibly JSON
3. New Technologies	Technologies supporting access to the standards metadata. Examples: API, possibly subscriptions?, SPARQL queries?
4. Metadata Gaps	Content implicitly referenced in the standards, but not explicitly available today. Examples: Value level metadata, research concepts
5. New Standards Content	New standards and new versions of existing standards. Examples: TAUGs, SDTM model v1.5, CDASH v2.0, protocol

What are the patterns of use in each case?

SHARE Category	Description & Examples
6. Semantics	The definitions and unambiguous meanings associated with the standards metadata. Example: each metadata element will have a Concept or DEC that includes a c-code and definition, BRIDG alignment
7. Normative + Non-normative	The standards represent normative content, but non-normative content can be a useful complement to the standards metadata. Example: example metadata, RDF model
8. Bundles or Packages	Combining metadata content from different standards into one package because they are used together. Example: TAUG + VLM + Examples, SDTM + Terminology
9. Computable Rules	Computable metadata for data transformations and validation. Example: VTL, SDTM validation, end-to-end automation
10. Examples & Test data	Example content and test datasets available for each standard. Examples: TAUGs, SDTM

What are the patterns of use in each case?

SHARE Category	Description & Examples
11. Healthcare Alignment	Links to other concept systems, or mapping to healthcare standards. Examples: CDASH to HL7 CCD, DEX, LOINC, SNOMED?
12. Federated MDRs	Explicit content linking between healthcare MDRs with different specializations. Examples: caDSR, SALUS, eTriks
13. New Reports	Reports that present existing standards metadata in new ways, or that provide information describing the existing standards metadata. Example: Diff files, Domain Catalog, Supp Qual Listing
14. Relationships	Content that describes the relationships between the standards, and potentially provides metadata describing how to transform metadata from state to state. Example: Mapping documents, RDF
15. Models	SHARE represents content within models. These models are used to load and extract metadata, and can be represented in metadata. Example: RDF

1. Existing Metadata

	A	B	C	D	E	F	G	H	I
1	Seq. For Order	Observation Class	Domain Prefix	Variable Name (minus domain prefix)	Variable Name	Variable Label	Type	Controlled Terms, Codelist or Format	Role
307	1	Findings	DA	STUDYID	STUDYID	Study Identifier	Char		Identifier
308	2	Findings	DA	DOMAIN	DOMAIN	Domain Abbreviation	Char	(DOMAIN)	Identifier
309	3	Findings	DA	USUBJID	USUBJID	Unique Subject Identifier	Char		Identifier
310	4	Findings	DA	SEQ	DASEQ	Sequence Number	Num		Identifier
311	5	Findings	DA	GRPID	DAGRPID	Group ID	Char		Identifier
312	6	Findings	DA	REFID	DAREFID	Reference ID	Char		Identifier
313	7	Findings	DA	SPID	DASPID	Sponsor-Defined Identifier	Char		Identifier
314	8	Findings	DA	TESTCD	DATESTCD	Short Name of Accountability Assessment	Char		Topic
315	9	Findings	DA	TEST	DATEST	Name of Accountability Assessment	Char		Synonym Qualifier
316	10	Findings	DA	CAT	DACAT	Category of Assessment	Char		Grouping Qualifier
317	11	Findings	DA	SCAT	DASCAT	Subcategory of Assessment	Char		Grouping Qualifier
318	12	Findings	DA	ORRES	DAORRES	Assessment Result in Original Units	Char		Result Qualifier
319	13	Findings	DA	ORRESU	DAORRESU	Original Units	Char	(UNIT)	Variable Qualifier
320	14	Findings	DA	STRESC	DASTRESC	Assessment Result in Std Format	Char		Result Qualifier

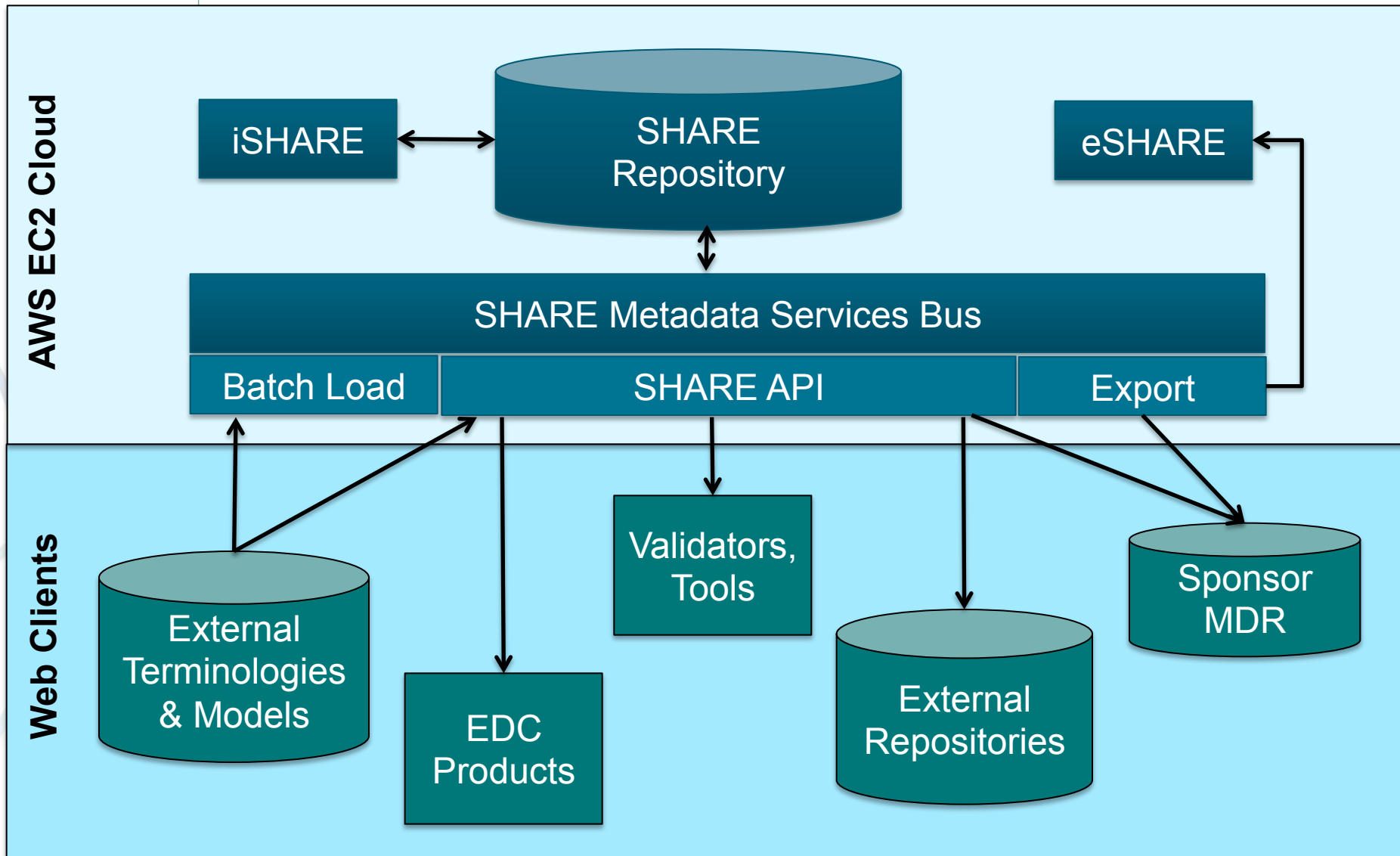
2. New Formats

```

▼<ODM xmlns="http://www.cdisc.org/ns/odm/v1.3" xmlns:def="http://www.cdisc.org/ns/def/v2.0" xmlns:xlink="http://www.w3.org/1999/
instance" CreationDateTime="2014-08-21T10:26:25" FileOID="Semantics.Manager.Define-XML_2.0.0_SDTM_1.0_13969030352051659917782"
Originator="CDISC" SourceSystem="SHARE" SourceSystemVersion="1.0 (6.6.10060)" xsi:schemaLocation="http://www.cdisc.org/ns/odm/v
▼<Study OID="Study.SDTM.MDES_3.1.3">
  ▼<GlobalVariables>
    <StudyName>Study - SDTM</StudyName>
    <StudyDescription>Export of SDTM with version MDES 3.1.3</StudyDescription>
    <ProtocolName>SDTM</ProtocolName>
  </GlobalVariables>
  ▼<MetaDataVersion xmlns:def="http://www.cdisc.org/ns/def/v2.0" Description="SDTM MDES 3.1.3" Name="SDTM MDV" OID="MDV.SDTM.2"
  def:StandardName="SDTM" def:StandardVersion="MDES 3.1.3">
    ▼<ItemGroupDef Domain="AE" Name="AE" OID="IG.MDES.AE.[SDTM]_MDES_2" Purpose="Tabulation" Repeating="Yes" def:Class="EVENTS"
    subject">
      ▼<Description>
        <TranslatedText xml:lang="en">Adverse Events</TranslatedText>
      </Description>
      <ItemRef ItemOID="IT.MDE.AEACN.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="28" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AEACNOTH.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="29" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AEBODSYS.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="21" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AECAT.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="18" Role="Grouping Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AECONTRT.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="42" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AEDECOD.[SDTM_AE]_MDE_1" Mandatory="Yes" OrderNumber="12" Role="Synonym Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AEDUR.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="48" Role="Timing"/>
      <ItemRef ItemOID="IT.MDE.AEENDTC.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="45" Role="Timing"/>
      <ItemRef ItemOID="IT.MDE.AEENDY.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="47" Role="Timing"/>
      <ItemRef ItemOID="IT.MDE.AEENRF.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="49" Role="Timing"/>
      <ItemRef ItemOID="IT.MDE.AEENRTPT.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="50" Role="Timing"/>
      <ItemRef ItemOID="IT.MDE.AEENTPT.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="51" Role="Timing"/>
      <ItemRef ItemOID="IT.MDE.AEGRPID.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="5" Role="Identifier"/>
      <ItemRef ItemOID="IT.MDE.AELOC.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="25" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AEMODIFY.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="9" Role="Synonym Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AEOUT.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="33" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AEPATT.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="32" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AEPRESP.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="20" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AEREFID.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="6" Role="Identifier"/>
      <ItemRef ItemOID="IT.MDE.AEREL.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="30" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AERELNST.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="31" Role="Record Qualifier"/>
      <ItemRef ItemOID="IT.MDE.AESCAN.[SDTM_AE]_MDE_1" Mandatory="No" OrderNumber="34" Role="Record Qualifier"/>

```

3. New Technology



4. Metadata Gaps

	A	B	C	D	E	F	G	H	I	J
	Concept from Rcmmap	NCI Definition	Domain Prefix	Variable Name	Variable Label	Type	Variable's XML Datatype	Controlled Terms, Codelist or Format	Codelist Name	Possible Value(s) for This Variable for This RC
1	Anatomic Site (C13717)	Named locations of or within the body.	LB (C49592)	LBLOC (C119849)	Specimen Collection Anatomic Location	Char	text	LOC (C74456)	Anatomical Location	BRONCHUS (C12683)
3	Laboratory Procedure (C25294)	Any procedure that involves testing or manipulating a sample of blood, urine, or other body substance in a laboratory setting.	LB (C49592)	LBTESTCD (C83322)	Lab Test or Examination Short Name	Char	text	LBTESTCD (C65047)	Laboratory Test Code	NEUT (C63321)
5	Laboratory Procedure (C25294)	Any procedure that involves testing or manipulating a sample of blood, urine, or other body substance in a laboratory setting.	LB (C49592)	LBTEST (C117142)	Lab Test or Examination Name	Char	text	LBTEST (C67154)	Laboratory Test Name	Neutrophils (C63321)
6	Logical Observation Identifiers Names and Codes (C49476)	Published by The Regenstrief Institute, the Logical Observation Identifiers Names and Codes covers clinical and clinical laboratory terminology.	LB (C49592)	LBLOINC (C83311)	LOINC Code	Char	text			26499-4; 26501-7; 32709-8
8	Unit of Measure (C25709)	A named quantity in terms of which other quantities are measured or specified, used as a standard measurement of like kinds.	LB (C49592)	LBORRESU (C83106)	Original Units	Char	text	UNIT (C71620)	Unit	% (C25613); cells/uL (C67242)
9	Laboratory Test Result (C36292)	The outcome of a laboratory test.	LB (C49592)	LBSTRESC (C83319)	Character Result/Finding in Std Format	Char	text	LBSTRESC (C102580)	Laboratory Test Standard Character Result	BORDERLINE (C14157); INDETERMINATE (C48658); INVALID (C50913); NEGATIVE (C38757); POSITIVE (C38758);
12	Unit of Measure (C25709)	A named quantity in terms of which other quantities are measured or specified, used as a standard measurement of like kinds.	LB (C49592)	LBSTRESU (C83321)	Standard Units	Char	text	UNIT (C71620)	Unit	% (C25613); cells/uL (C67242)
14	Normal Range Indicator (C82532)	A indication that a result lies within normal parameters.	LB (C49592)	LBNRIND (C83094)	Reference Range Indicator	Char	text	NRIND (C78736)	Reference Range Indicator	HIGH (C78800); LOW (C78801); NORMAL (C78727)
17										

5. New Standards Content

Search Results		
[Edit Search Properties] [Process Assets]		
☐	Asset Name (Version)	Asset Type
☐	 SDTM (MDES 3.1.2)	Metadata Element Set
☐	 SDTM (MDES 3.1.3)	Metadata Element Set
☐	 SDTM (MDES 3.2)	Metadata Element Set
☐	 SDTM TAUG-Alzheimer's (MDES 2.0)	Metadata Element Set
☐	 SDTM TAUG-Asthma (MDES 1.0)	Metadata Element Set
☐	 SDTM TAUG-Cardiovascular (MDES 1.0)	Metadata Element Set
☐	 SDTM TAUG-Diabetes (MDES 1.0)	Metadata Element Set
☐	 SDTM TAUG-Influenza (MDES 1.0)	Metadata Element Set
☐	 SDTM TAUG-Multiple Sclerosis (MDES 1.0)	Metadata Element Set
☐	 SDTM TAUG-QT (MDES 1.0)	Metadata Element Set
☐	 SDTM Virology (MDES 1.0)	Metadata Element Set
☐	 SDTM-Medical Devices (MDES 1.0)	Metadata Element Set
☐	 SENDIG (MDES 3.0)	Metadata Element Set
☐	 SENDIG (MDES 3.1)	Metadata Element Set

6. Semantics

Domain	Variable	NCI C-code	NCI PT	NCI Definition
IS	ISORRES	C117559	Immunogenicity Specimen Assessments Original Result	The outcome of the immunogenicity specimen assessment as originally received or collected.
IS	ISORRESU	C117560	Immunogenicity Specimen Assessments Original Result Unit	The unit of measure for the result of the immunogenicity specimen assessment as originally received or collected.
MI	MIORRES	C117582	Microscopic Findings Original Result	The outcome of the microscopic assessment as originally received or collected.
MI	MIORRESU	C117583	Microscopic Findings Original Result Unit	The unit of measure for the result of the microscopic test or finding as originally received or collected.
MO	MOORRES	C117611	Morphology Original Result	The outcome of the morphology assessment as originally received or collected.
MO	MOORRESU	C117612	Morphology Original Result Unit	The unit of measure for the result of the morphology test or finding as originally received or collected.
RP	RPORRES	C117644	Reproductive System Findings Original Result	The outcome of the reproductive system findings assessment as originally received or collected.
RP	RPORRESU	C117645	Reproductive System Findings Original Result Unit	The unit of measure for the result of the reproductive system findings test or finding as originally received or collected.
SR	SRORRES	C117679	Skin Response Original Result	The outcome of the skin response assessment as originally received or collected.
SR	SRORRESU	C117680	Skin Response Original Result Unit	The unit of measure for the result of the skin response test or finding as originally received or collected.

7. Normative + Non-Normative

Normative Issues

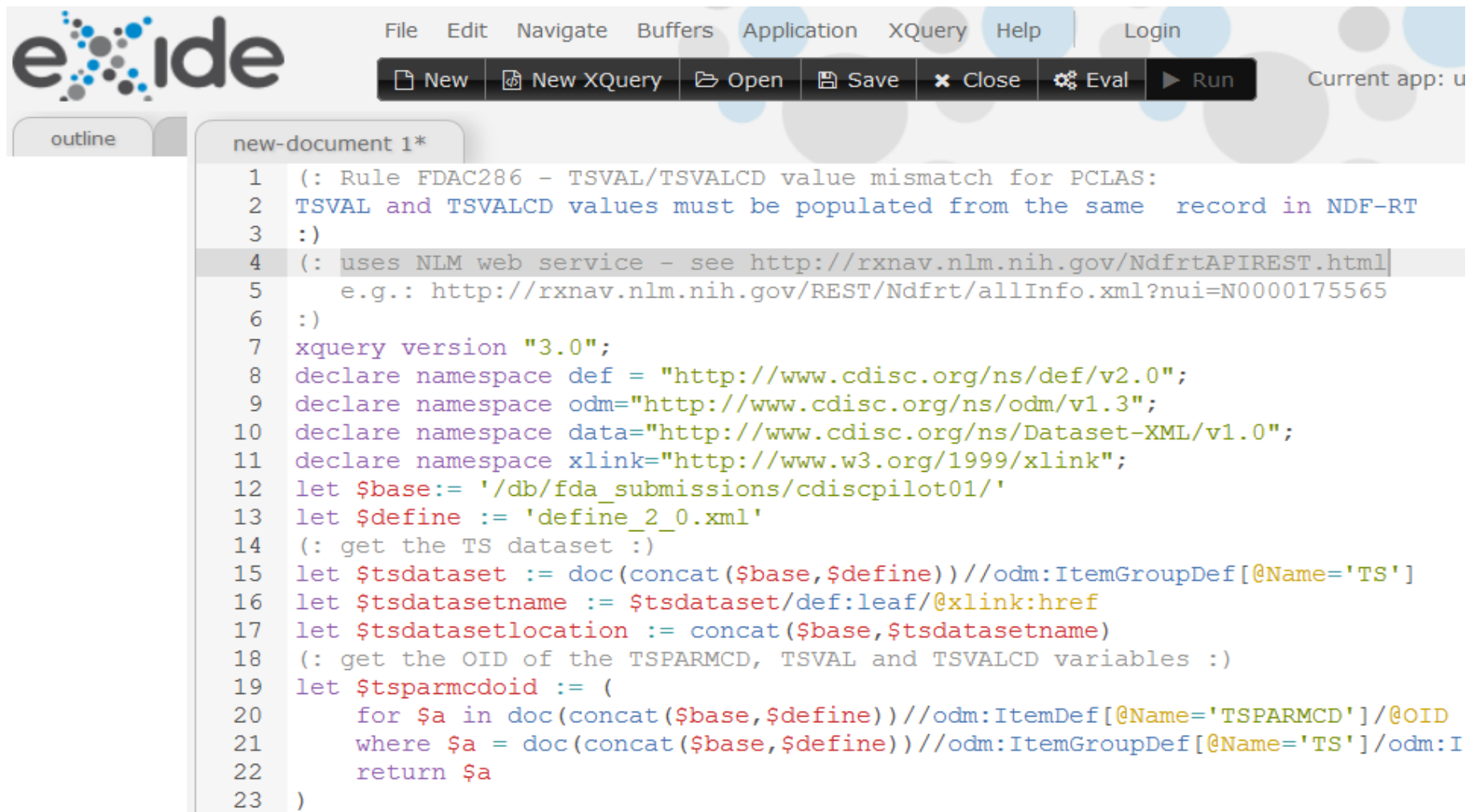
The current RDF package represents the CDISC foundational standards as published and intends to respect the normative character of the standards with two exceptions.

The RDF model contains an additional attribute to express the data type of a data element using the W3C XML Schema Datatypes (XSD) standard. The XSD data types are more accurate than the published data types (e.g. SDTM Controlled Terms or Format), but cannot be considered to be normative. This is a topic for further consideration.

8. Bundles & Packages

- Package Contents for Asthma-TAUG-v1-0-Draft.zip:
 - sdtm-ta-asthma-taug.xls
 - sdtm-ta-asthma-routine.xls
 - sdtm-ta-asthma-medical-device.xls
 - cdash-ta-asthma.xls
 - ct-sdtm-ta-asthma.xls
 - ct-cdash-ta-asthma.xls
 - FATEST Listing.xlsx
 - HbA1c and Glucose SHARE Metadata.xlsx
 - SuppQual Listing.xls
 - taug_asthma_v1.0_single.pdf
 - taug_asthma-readme.pdf

9. Computable Rules



The screenshot shows the eXide web application interface. The top menu bar includes File, Edit, Navigate, Buffers, Application, XQuery, Help, and Login. Below the menu is a toolbar with buttons for New, New XQuery, Open, Save, Close, Eval, and Run. The main editor area displays an XQuery script for rule FDAC286. The script is as follows:

```
1  (: Rule FDAC286 - TSVAL/TSVALCD value mismatch for PCLAS:
2  TSVAL and TSVALCD values must be populated from the same record in NDF-RT
3  :)
4  (: uses NLM web service - see http://rxnav.nlm.nih.gov/NdfrestAPIREST.html
5  e.g.: http://rxnav.nlm.nih.gov/REST/Ndfrest/allInfo.xml?nui=N0000175565
6  :)
7  xquery version "3.0";
8  declare namespace def = "http://www.cdisc.org/ns/def/v2.0";
9  declare namespace odm="http://www.cdisc.org/ns/odm/v1.3";
10 declare namespace data="http://www.cdisc.org/ns/Dataset-XML/v1.0";
11 declare namespace xlink="http://www.w3.org/1999/xlink";
12 let $base:= '/db/fda_submissions/cdiscpilot01/'
13 let $define := 'define_2_0.xml'
14 (: get the TS dataset :)
15 let $tsdataset := doc(concat($base,$define))//odm:ItemGroupDef[@Name='TS']
16 let $tsdatasetname := $tsdataset/def:leaf/@xlink:href
17 let $tsdatasetlocation := concat($base,$tsdatasetname)
18 (: get the OID of the TSPARMCD, TSVAL and TSVALCD variables :)
19 let $tsparmcdoid := (
20   for $a in doc(concat($base,$define))//odm:ItemDef[@Name='TSPARMCD']/@OID
21   where $a = doc(concat($base,$define))//odm:ItemGroupDef[@Name='TS']/odm:I
22   return $a
23 )
```

OpenRules

10. Examples & Test Data

- › ex.xpt, Exposure - SDTMIG v3.2
- › fa.xpt, Findings About Events or Interventions - SDTMIG v3.2
- ✓ ho.xpt, Healthcare Encounters - SDTMIG v3.2

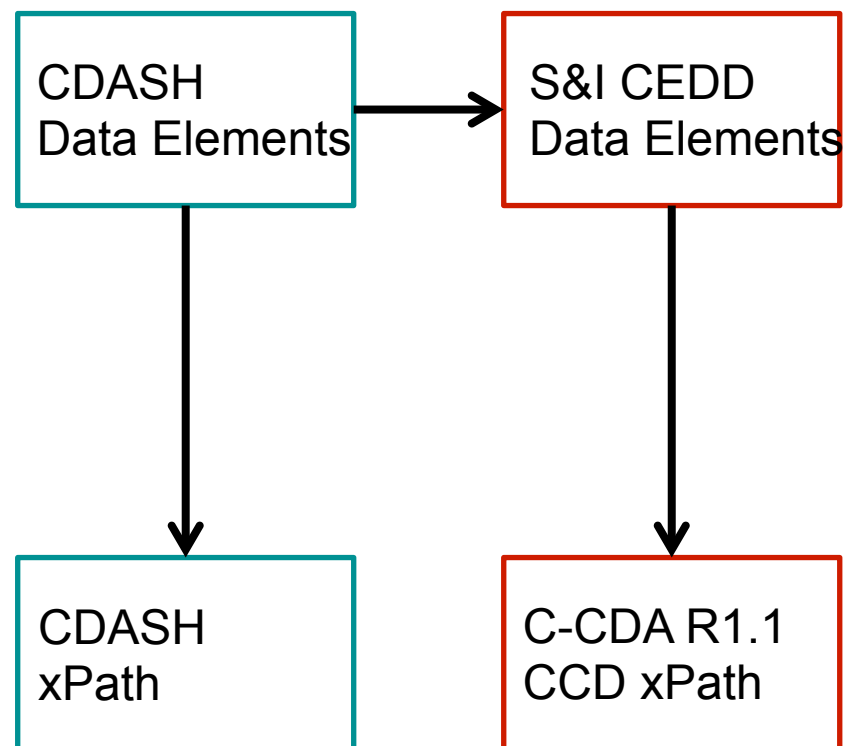
Row	STUDYID	DOMAIN	USUBJID	HOSEQ	HOGRPID	HOREFID	HOSPID	HOTERM
Role	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Identifier	Topic
Core	Req	Req	Req	Req	Perm	Perm	Perm	Req
CT, Format, Type	Char	(DOMAIN)	Char	Num	Char	Char	Char	Char

- › ie.xpt, Inclusion/Exclusion Criteria Not Met - SDTMIG v3.2

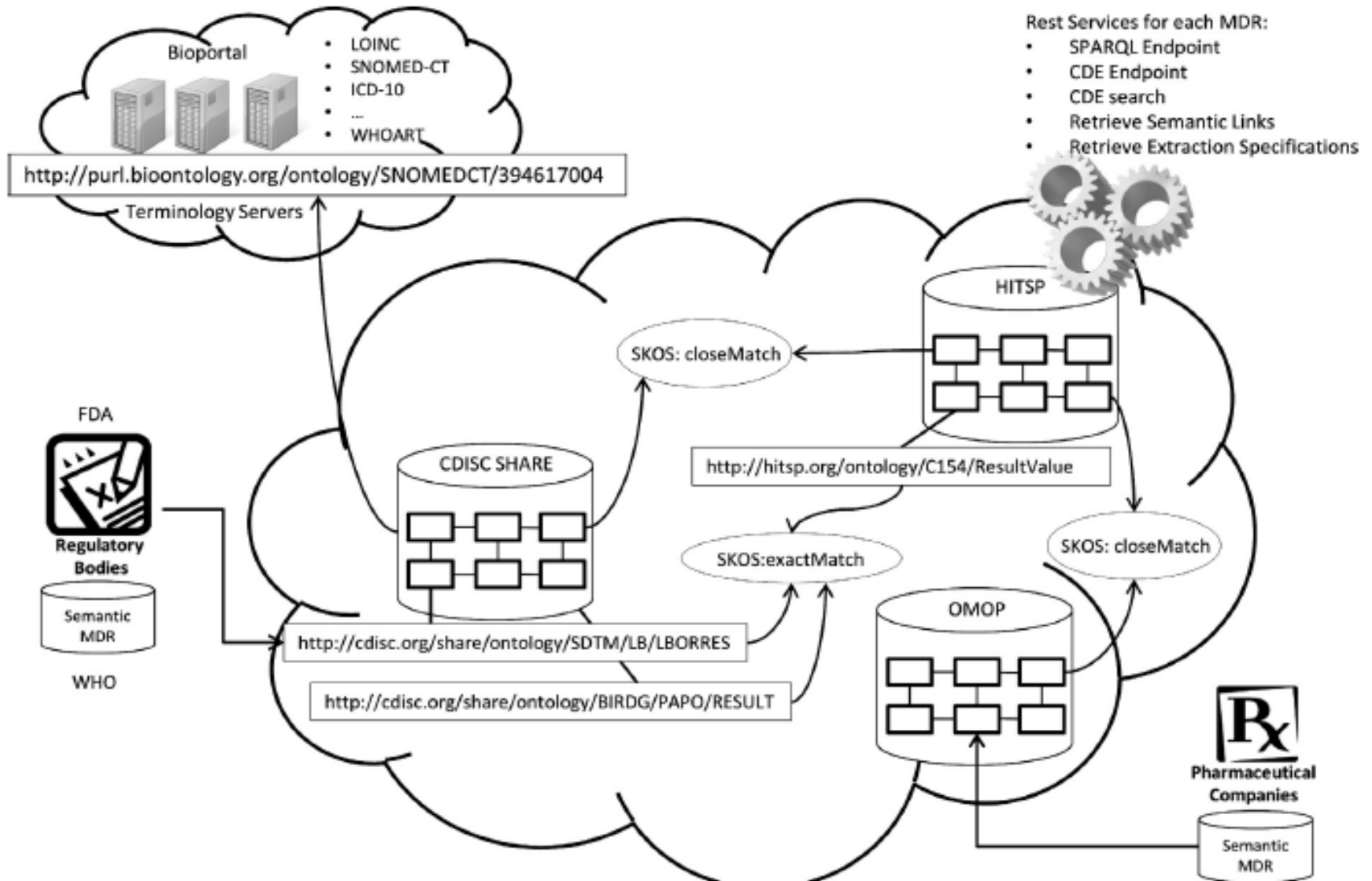
11. Healthcare Alignment

- CDASH RDF

- Each data element is uniquely identifiable in SHARE
- Already in ISO 11179 Spec
- Maintain the mappings of them to several different content models in an MDR supporting DEX
- Retrieve the corresponding mapping specification (extraction specification) of the corresponding CDE to a specific “Content Model”



12. Federated MDRs



SINACI, A. A.; LALECI ERTURKMEN, G. B. A federated semantic metadata registry framework for enabling interoperability across clinical research and care domains. Journal of Biomedical Informatics, v. 46, n. 5, p. 784-794, 2013.

13. New Reports

Change Nbr	Action	Impact	Domain	Variable	Attribute (updated)	Attribute (previous)	Updated SDTM-IG Ver	Previous SDTM-IG Ver
0	Add	Domain	APDM				3.2	3.1.3
1	Add	Domain	APRELSUB				3.2	3.1.3
2	Update	CDISC Notes and Definitions	AE	AEACN	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.	Describes changes to the study treatment as a result of the event.	3.2	3.1.3
3	Add	Controlled Terminology	AE	AEDECOD	MedDRA		3.2	3.1.3
4	Update	CDISC Notes and Definitions	AE	AEDECOD	Dictionaryderived text description of AETERM or AEMODIFY. Equivalent to the Preferred Term (PT in MedDRA). The sponsor is expected to provide the dictionary name and version used to map the terms utilizing the define.xml external codelist attributes.	Dictionaryderived text description of AETERM or AEMODIFY.	3.2	3.1.3
5	Update	Controlled Terminology	AE	AEENRTPT	(STENRF)	BEFORE, AFTER, COINCIDENT, ONGOING, U	3.2	3.1.3
6	Update	CDISC Notes and Definitions	AE	AELC	Describes anatomical location relevant for the event (e.g., ARM for skin rash).	Describes anatomical location relevant for the event (e.g., L	3.2	3.1.3
7	Update	Role	AE	AEPRESP	Variable Qualifier	Record Qualifier	3.2	3.1.3
8	Update	CDISC Notes and Definitions	AE	AEREFID	Internal or external identifier such as a serial number on an SAE reporting form.	Internal or external identifier such as a serial number on an	3.2	3.1.3
9	Update	CDISC Notes and Definitions	AE	AEREL	Records the investigators opinion as to the causality of the event to the treatment. ICH E2A and E2B examples include NOT RELATED, UNLIKELY RELATED, POSSIBLY RELATED, RELATED. Controlled Terminology may be defined in the future. Check with regulatory authority for population of this variable.	Records the investigators opinion as to the causality of the event.	3.2	3.1.3

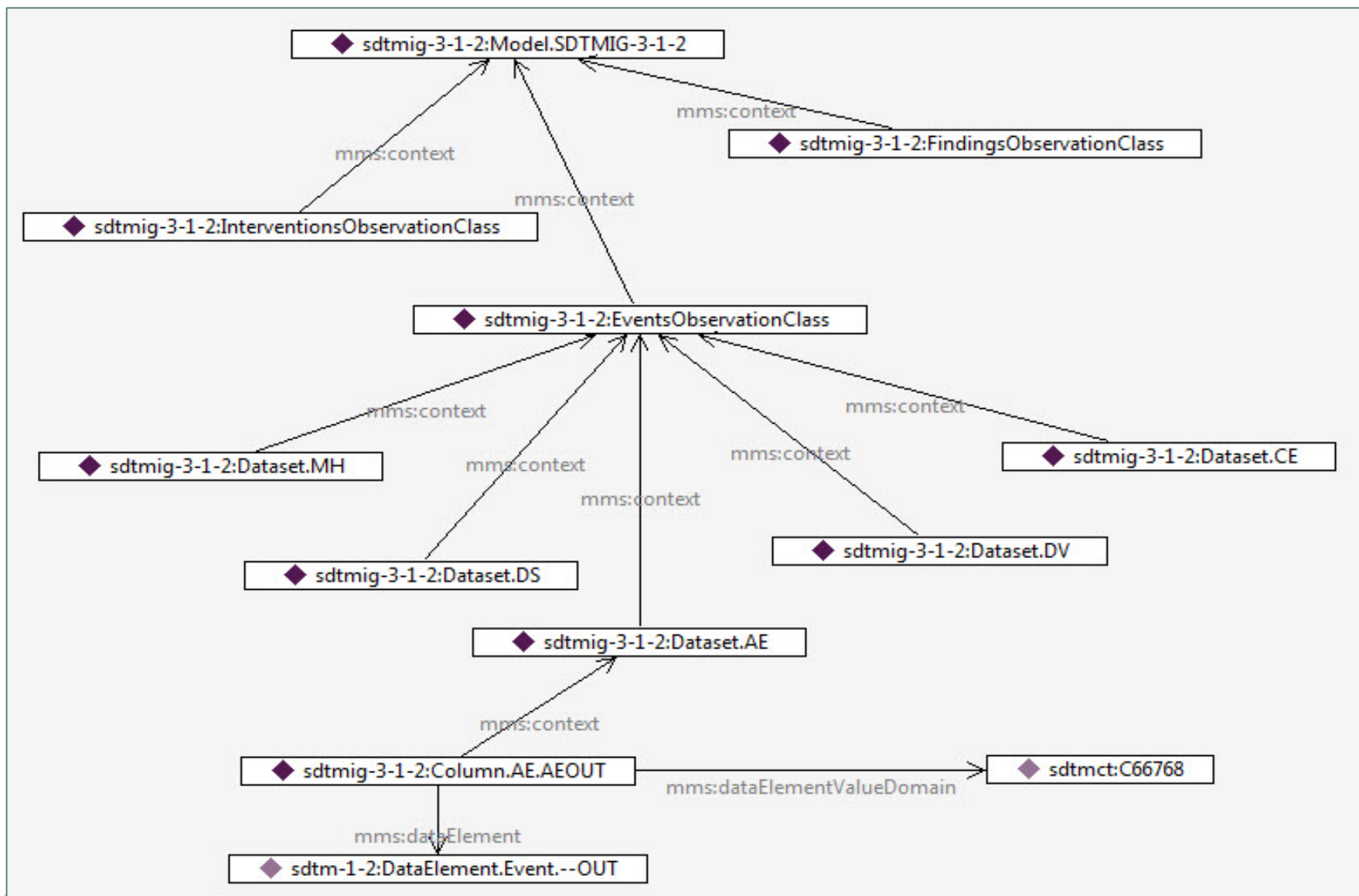
SDTM 3.2 Diff file

14. Relationships

CDASH Domain	CDASH Variable	SDTM Domain	SDTM Variable	Mapping Instructions
AE	AESTDAT	AE	AESTDTC	For the SDTM dataset, the SDTM variable --STDTC is d 8601 format.
AE	AESTTIM	AE	AESTDTC	For the SDTM dataset, the SDTM variable --STDTC is d 8601 format.
AE	AEENDAT	AE	AEENDTC	For the SDTM dataset, the SDTM variable --ENDTC is d 8601 format.
AE	AEENTIM	AE	AEENDTC	For the SDTM dataset, the SDTM variable --ENDTC is d 8601 format.
AE	AEONGO	AE	AEENRF	May be used to derive a value into an SDTM relative to 'Y', the value of 'DURING', 'AFTER' or 'DURING/AFTER' may be derived. Note: AEENRTPT must refer to a 'time
AE	AEONGO	AE	AEENRTPT	May be used to derive a value into an SDTM relative to 'Y', the value of 'DURING', 'AFTER' or 'DURING/AFTER' may be derived. Note: AEENRTPT must refer to a 'time
AE	AEONGO	AE	AEENTPT	May be used to derive a value into an SDTM relative to 'Y', the value of 'DURING', 'AFTER' or 'DURING/AFTER' may be derived. Note: AEENRTPT must refer to a 'time
AE	AESV	AE	AESV	Maps directly to SDTM. The sponsor is expected to pr code list attributes.
AE	AETOXGR	AE	AETOXGR	Maps directly to SDTM. The sponsor is expected to pr code list attributes.

CDASH to SDTM mapping extracted from SHARE

15. Models



How will new standards metadata from SHARE impact patterns of use?

- Opportunities to update processes to take advantage of the new categories of standards metadata
- How will the availability of new standards metadata influence your work?
- How will you influence what new standards metadata become available?
- For discussion at today's workshop session!

Thanks!





Strength *through* collaboration.