

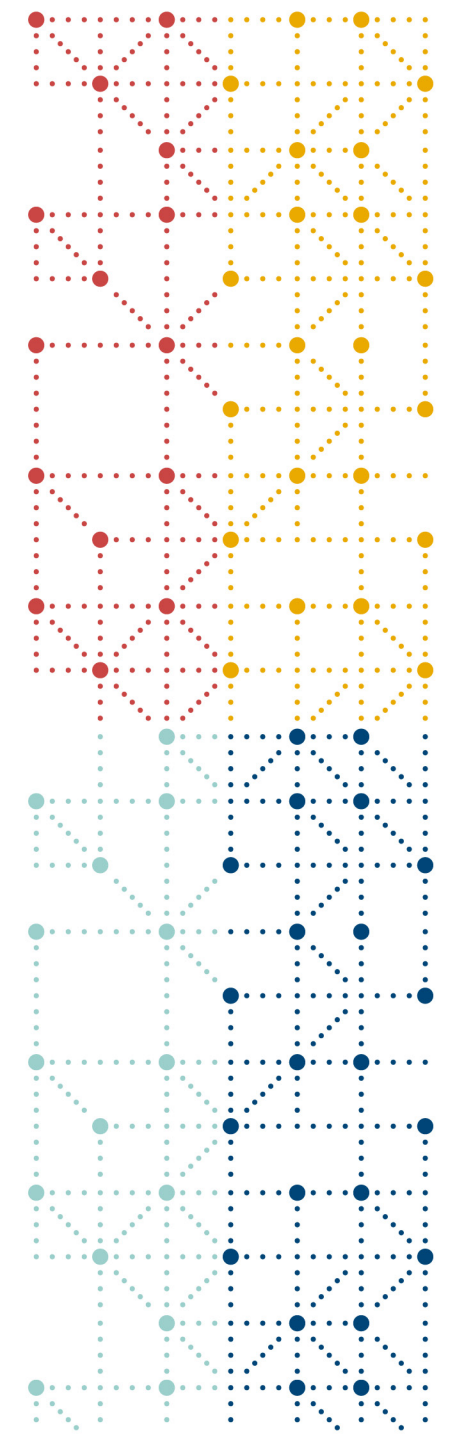
CDISC 360: Using Biomedical Concept Metadata to Generate Case Report Forms and Dataset Definitions

Sam Hume, DSc
VP, Data Science, CDISC

10 March 2020

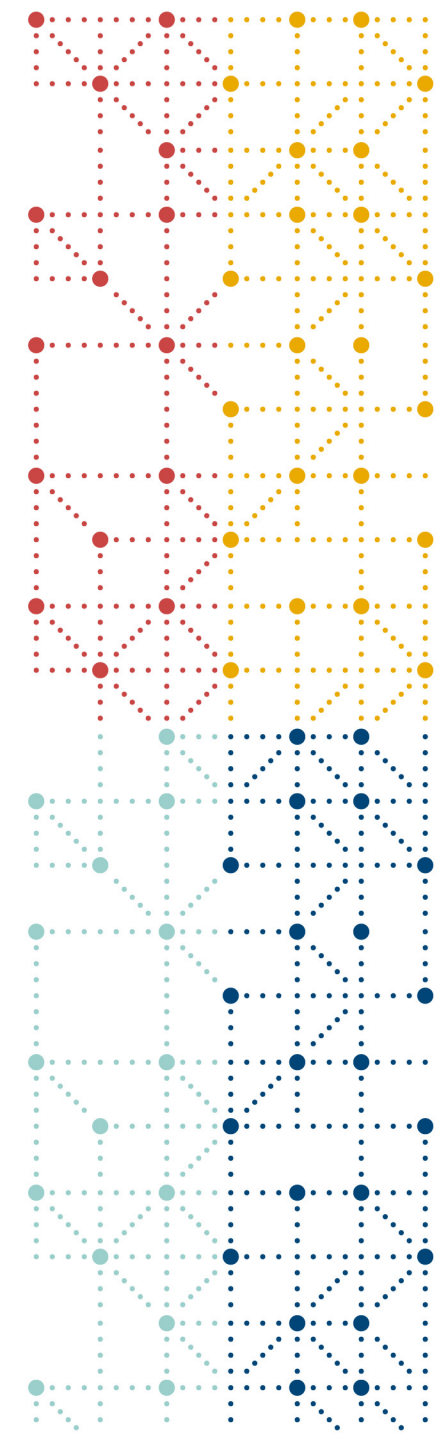
Paper TT06





Agenda

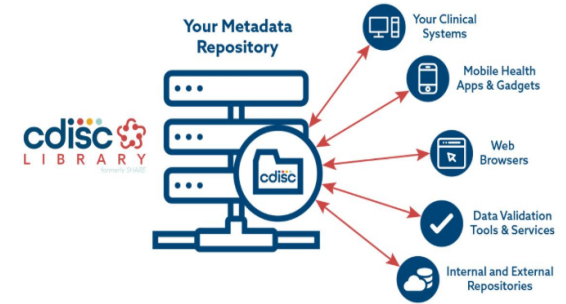
1. CDISC 360 Objectives
2. Maps, Models, Metadata, and More
3. Automated Define-XML Generation
4. Proposal



CDISC 360 High-level Objectives

What problems are we trying to solve in the CDISC 360 proof-of-concept project?

CDISC 360 Project Objectives



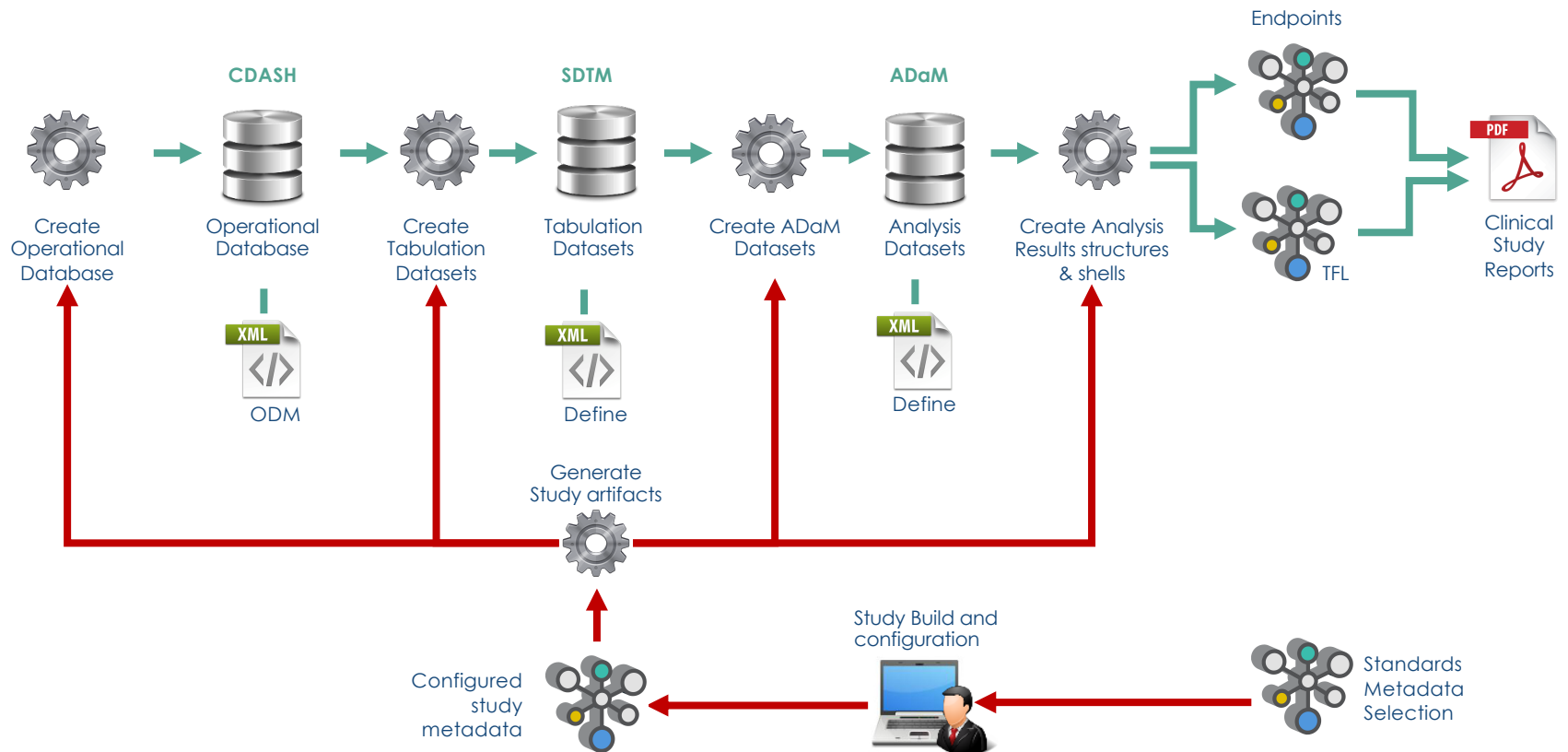
CDISC 360 High-level Objectives:

1. Increase study implementation automation
2. Reduce study implementation variability

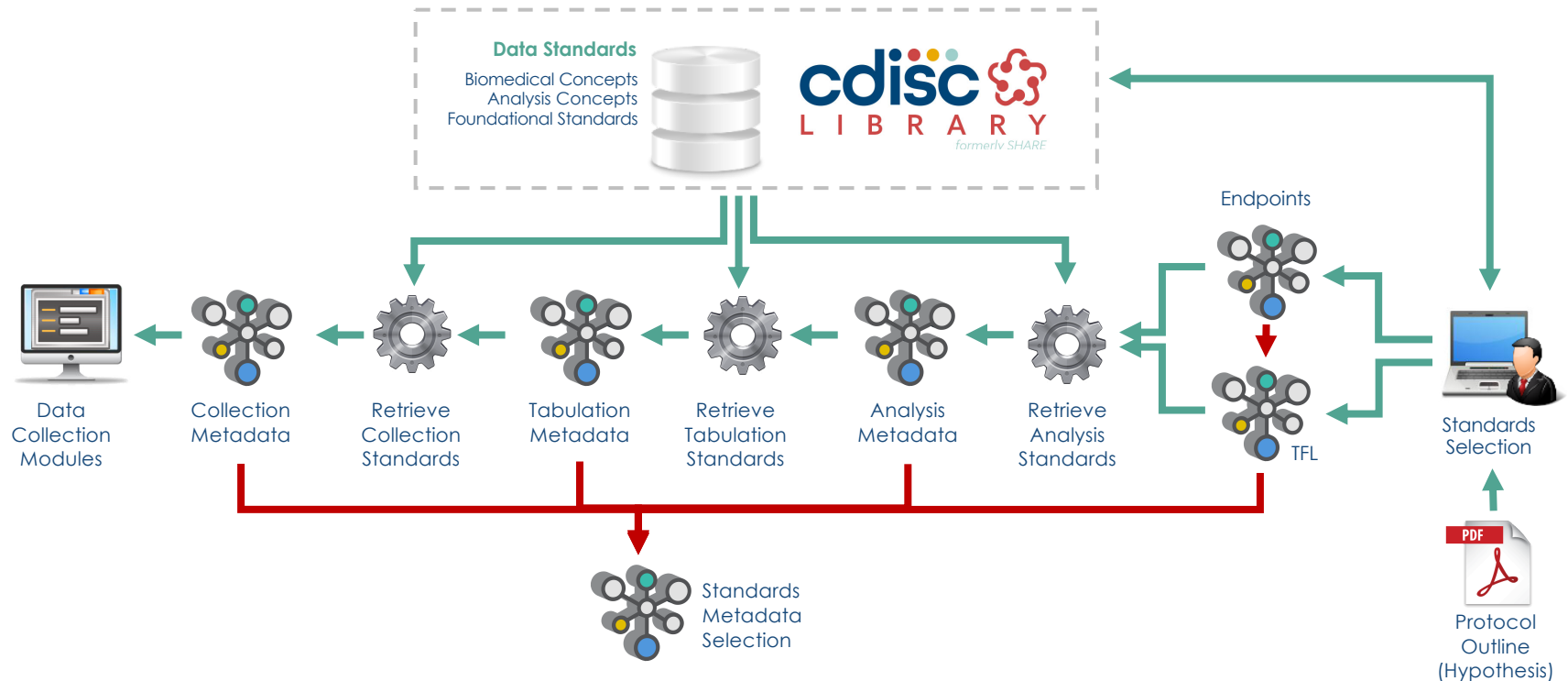
CDISC 360 Use Cases

- Use Case 1 : End to Start specification
- Use Case 2 : Start to End Study Metadata
- Use Case 3 : Start to End Data Processing

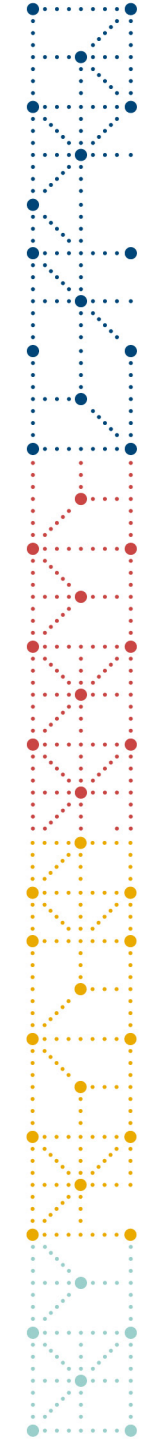
Focus on CDISC 360 Use Case 2 : Start to End Study Metadata



CDISC Library Provides the Standards Metadata



New standards metadata created in CDISC 360
will be published via a sandbox CDISC Library

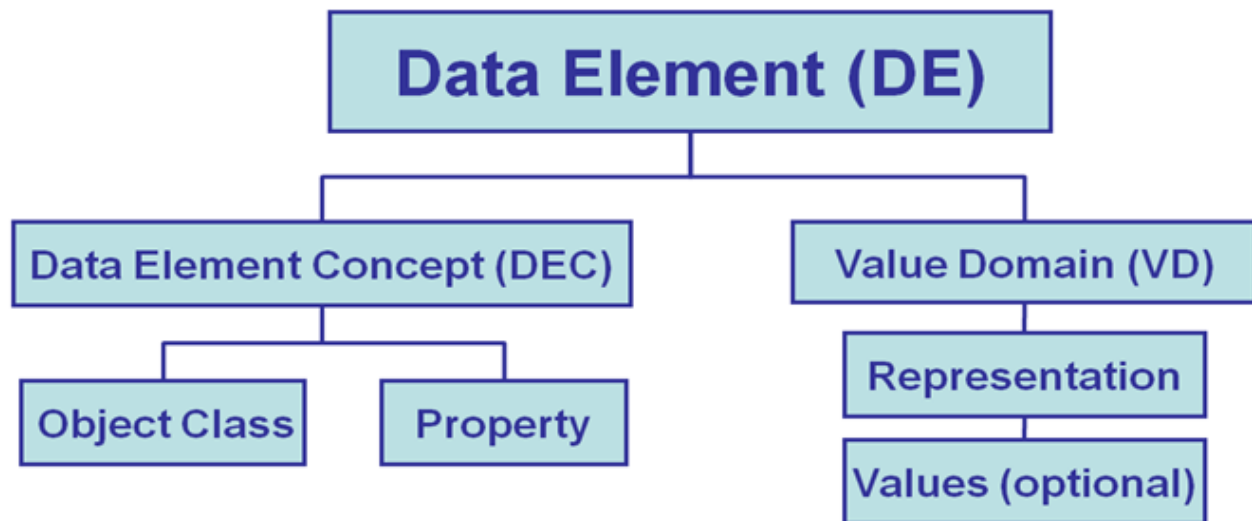


Generating a Vital Signs Define-XML: What's Missing?

- Value Level Metadata
- List of applicable tests
- Valid set of units for each test
- ODM Datatypes for Define-XML
- Length and SignificantDigits values
- Valid controlled terminology subsets (e.g. body position)

Question: Where is the Height Variable?

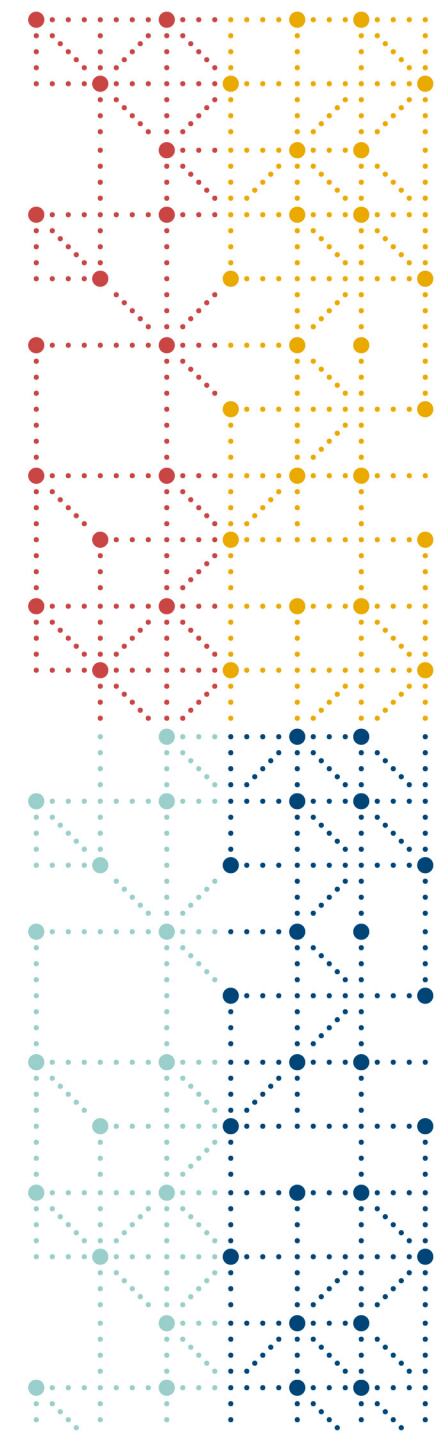
- HEIGHT is a VSTESTCD, but not a variable
- You might think of Height as a virtual variable
- It needs to be defined as a Data Element for use as a variable in an SDTMIG Define-XML document
 - Or, as a field in a CDASHIG CRF
- Biomedical Concepts help us to create a Height data element



Generate a Complete Define-XML Specification File Including Value Level Metadata

VS (Vital Signs) - [SDTMIG 3.2]

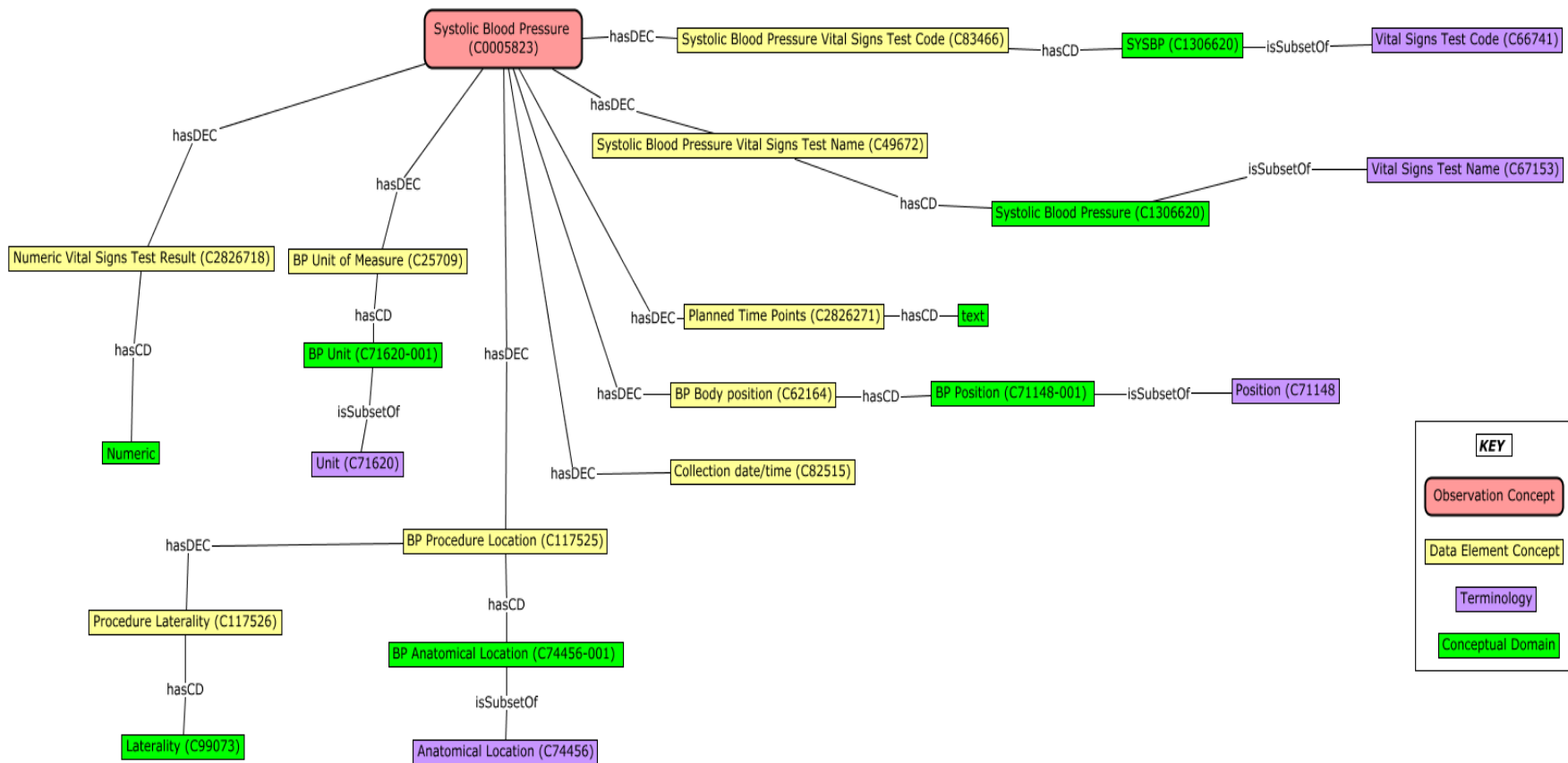
Variable	Where Condition	Label / Description	Type	Length or Display Format	Controlled Terms or ISO Format	Origin / Source / Method / Comment
STUDYID		Unique Identifier for a study.	text	40		
DOMAIN		Two-character abbreviation for the domain.	text	2		
USUBJID		Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	text	40		
VSSEQ		Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number.	integer	4		
VSTESTCD		Short name of the measurement, test, or examination described in VSTEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value in VSTESTCD cannot be longer than 8 characters, nor can it start with a number (e.g. "1TEST"). VSTESTCD cannot contain characters other than letters, numbers, or underscores. Examples: SYSBP, DIABP, BMI.	text	8	Vital Signs Test Code [6 Terms]	
VSTEST		Verbatim name of the test or examination used to obtain the measurement or finding. The value in VSTEST cannot be longer than 40 characters. Examples: Systolic Blood Pressure, Diastolic Blood Pressure, Body Mass Index.	text	40	Vital Signs Test Name [6 Terms]	
VSORRES VLM		Result of the vital signs measurement as originally received or collected.	text	30		
	VSTESTCD = "TEMP"	VSORRES for Temperature	float	4		
	VSTESTCD = "HEIGHT"	VSORRES for Height	float	5		
	VSTESTCD = "DIABP"	VSORRES for Diastolic Blood Pressure	integer	3		
	VSTESTCD = "SYSBP"	VSORRES for Systolic Blood Pressure	integer	3		
	VSTESTCD = "WEIGHT"	VSORRES for Weight	float	6		
	VSTESTCD = "HR"	VSORRES for Heart Rate	integer	3		
VSORRESU VLM		Original units in which the data were collected. The unit for VSORRES. Examples: IN, LB, BEATS/MIN.	text	20		
	VSTESTCD = "TEMP"	VSORRESU for Temperature	text		Units for Vital Signs Results Temperature • "C" • "F"	



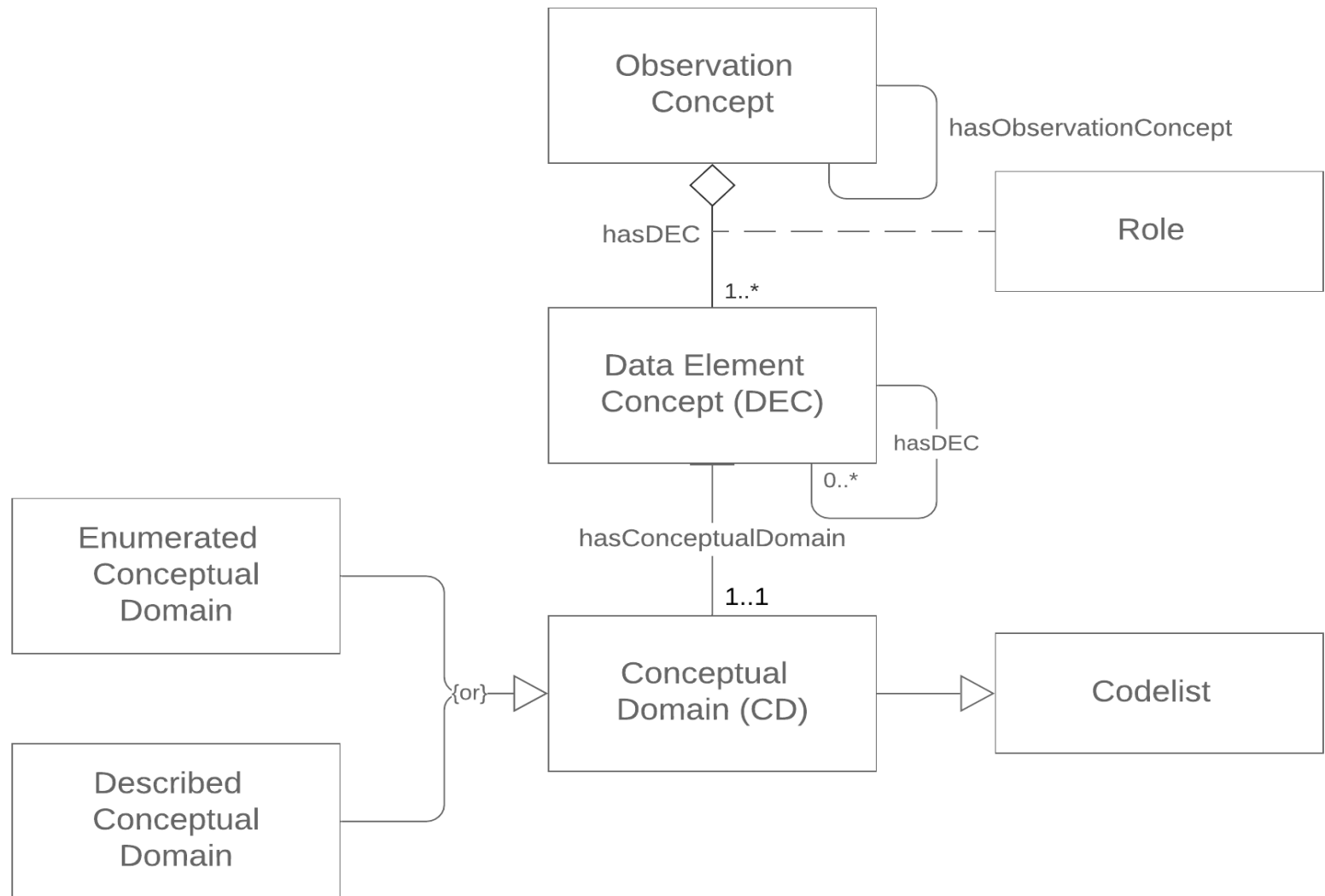
Maps, Models, Metadata, and More

What new content is being created as part of CDISC 360 to address the project objectives?

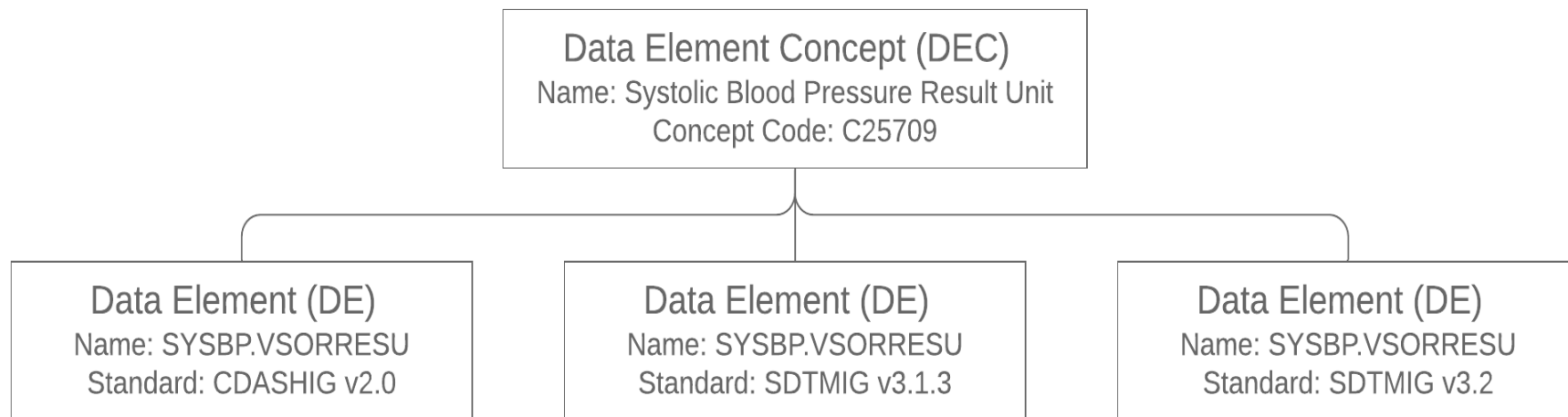
CMAP for the Systolic Blood Pressure Biomedical Concept



Biomedical Concept Metamodel

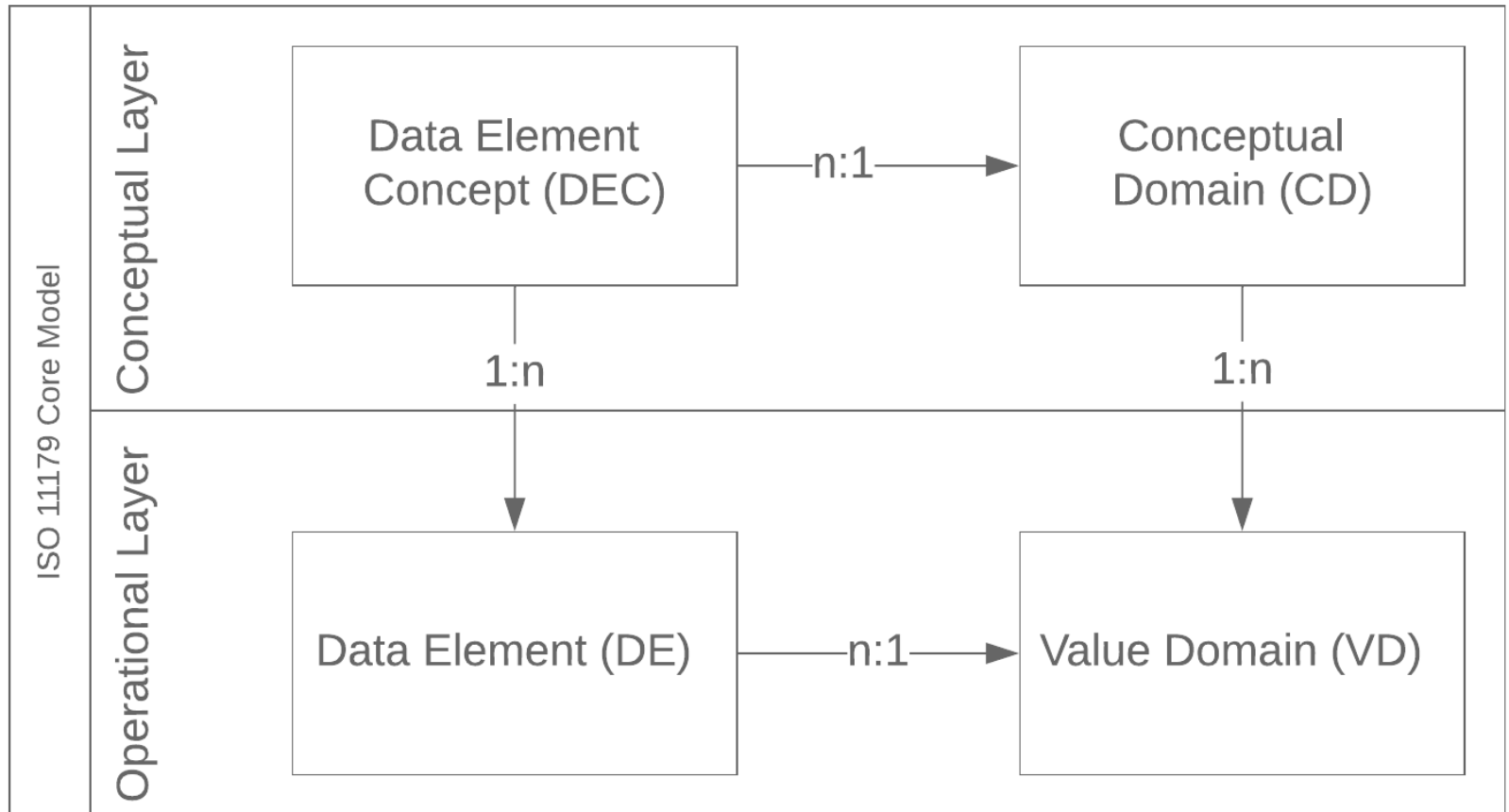


One Data Element Concept can Represent Multiple Data Elements

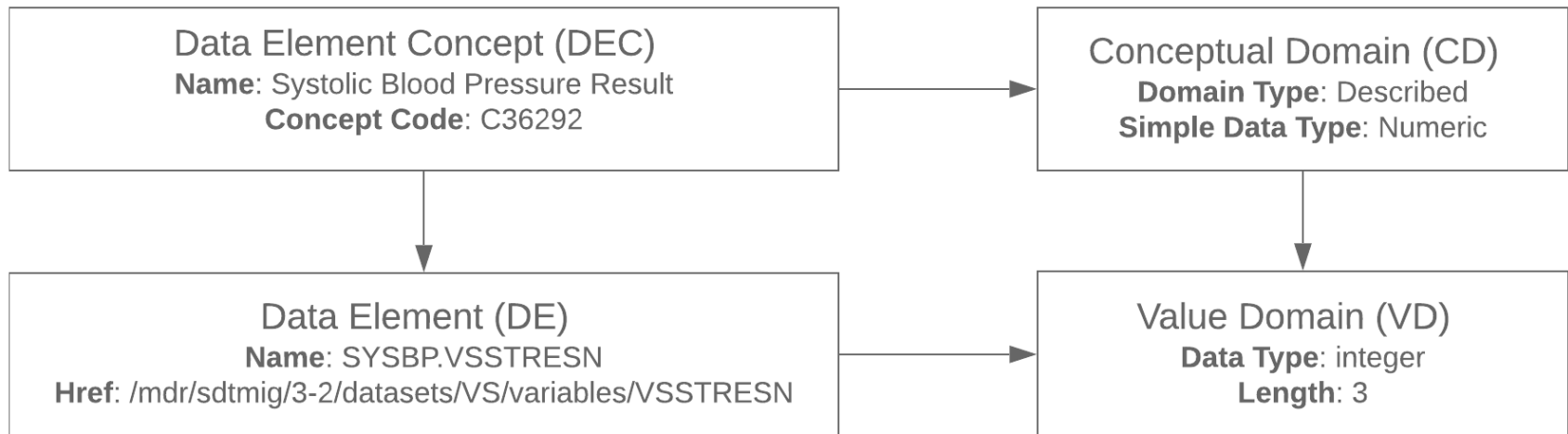


A biomedical concept is a unit of knowledge created by a unique combination of characteristics (ISO 11179). A biomedical concept is independent of a standards implementation.

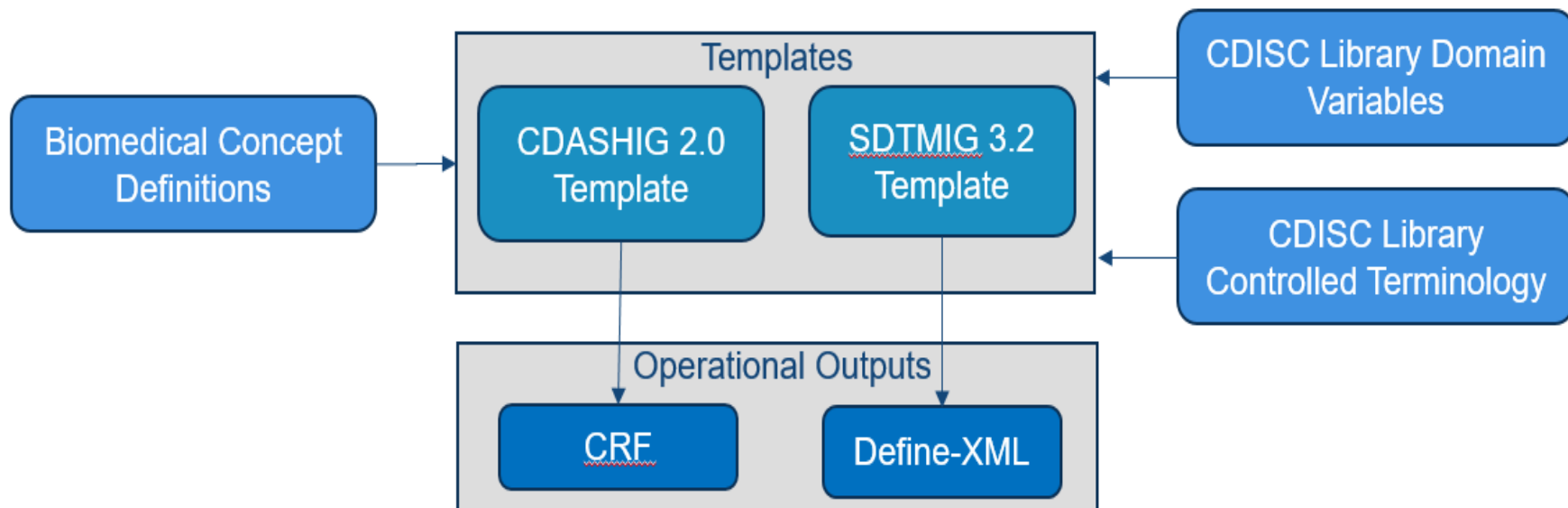
ISO 11179 Core Model



ISO 11179 Core Model Example

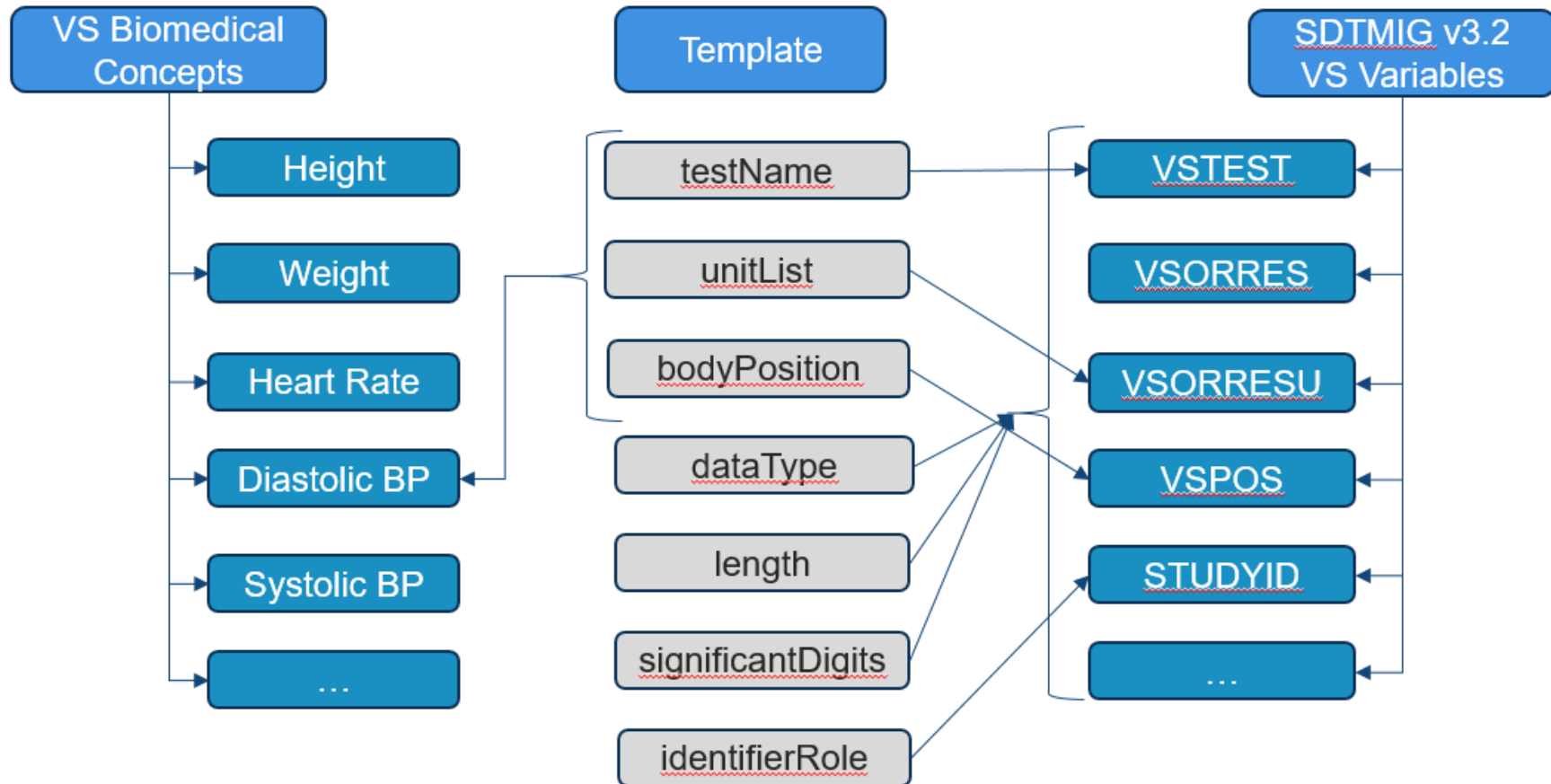


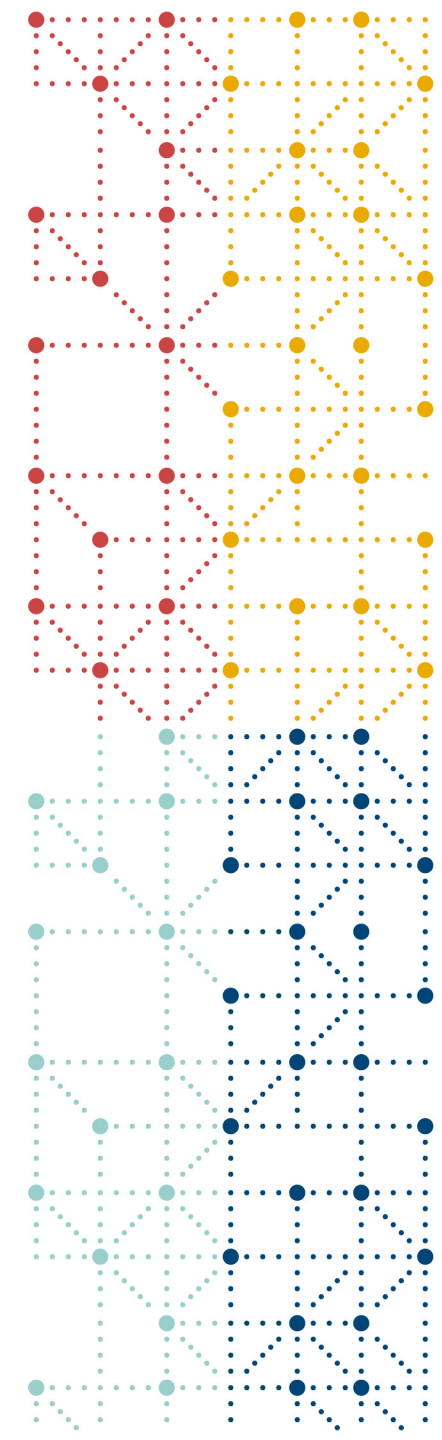
Using Templates to Apply Biomedical Concepts to Standards Metadata



Templates are one of four sources of metadata applied to generate a Define-XML document. Templates reference the other three metadata sources, and constrain that metadata for use in a specific context (a specific study within a TAUG).

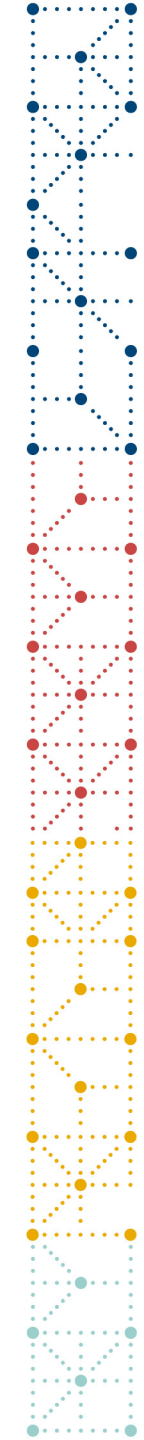
Applying Biomedical Concepts using Templates





Automated Define-XML Generation

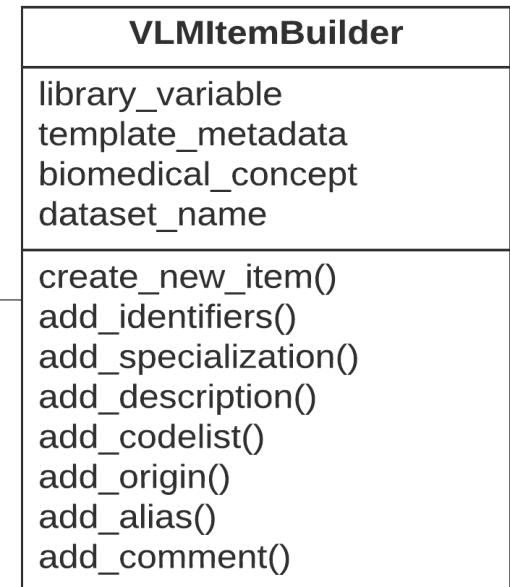
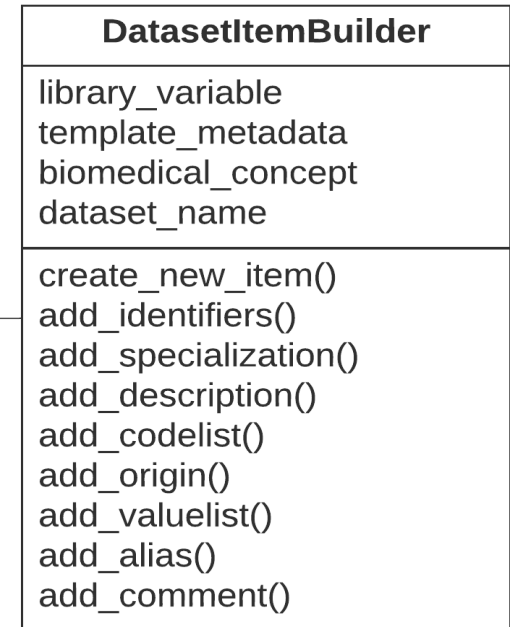
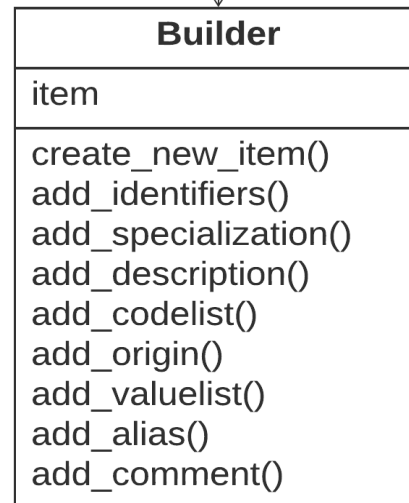
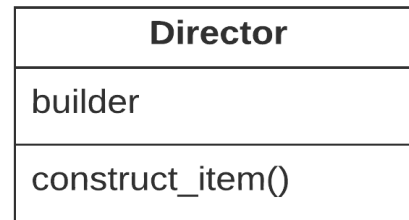
How can we use the new CDISC 360 metadata to generate a complete Define-XML specification document?



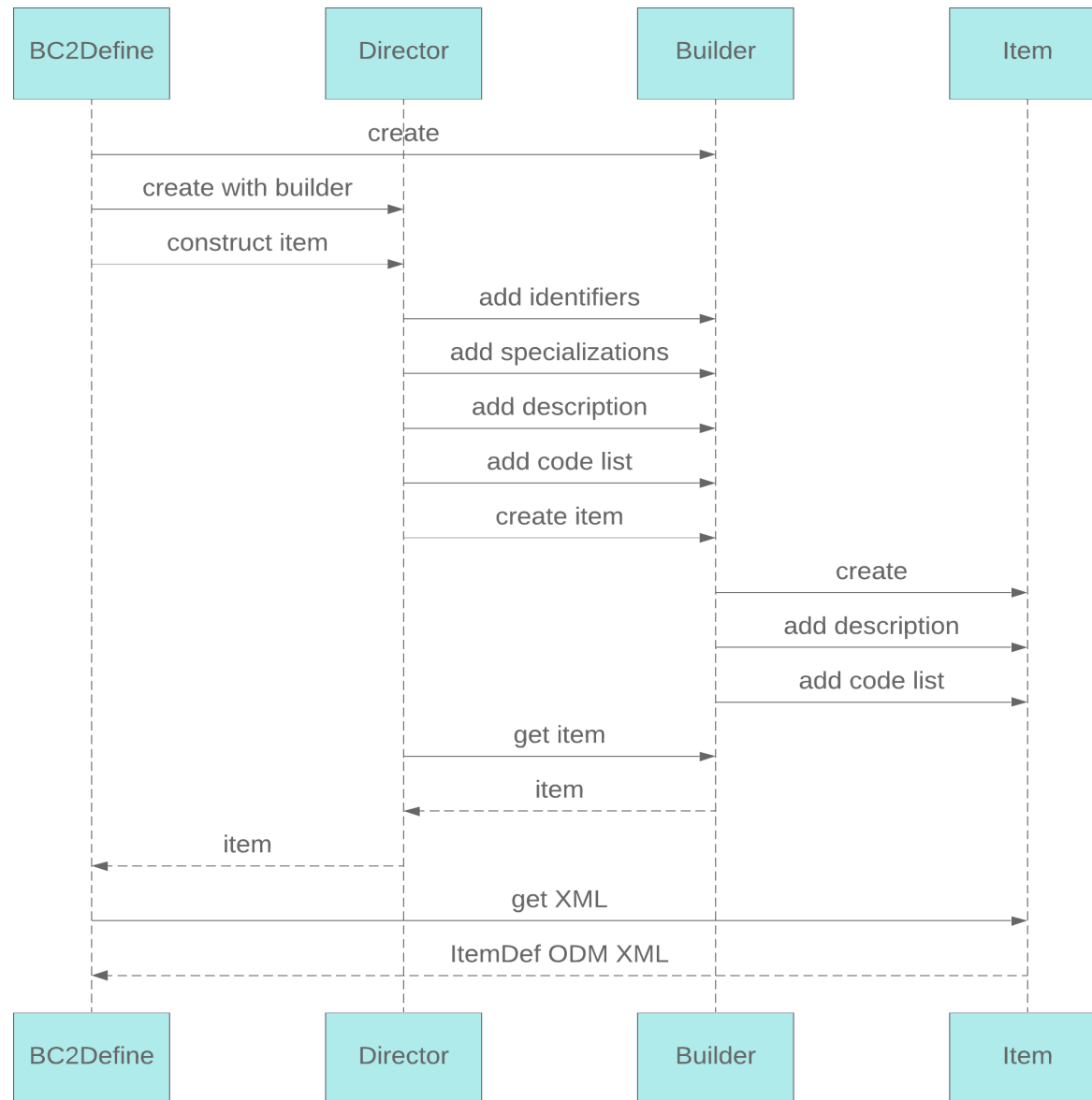
Generate Define-XML using Four Metadata Sources

- **Templates**
 - Published to the CDISC 360 Sandbox Library
- **Biomedical Concepts**
 - Published to the CDISC 360 Sandbox Library
- **SDTMIG v3.2 Metadata**
 - Available from the CDISC Library
- **SDTM 2018-06-29 Controlled Terminology Package**
 - Available from the CDISC Library

Simple Builder Pattern Class Diagram to Generate ItemDef objects



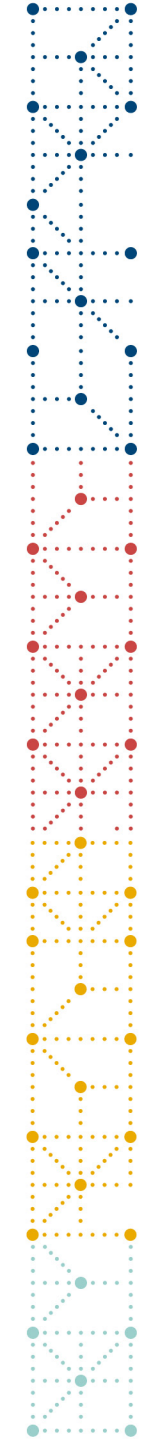
Builder Pattern Sequence Diagram for VLM *ItemDef* Builder



Generate a Vital Signs define.xml using Define-XML v2.1

VS (Vital Signs) - [SDTMIG 3.2]

Variable	Where Condition	Label / Description	Type	Length or Display Format	Controlled Terms or ISO Format	Origin / Source / Method / Comment
STUDYID		Unique identifier for a study.	text	40		
DOMAIN		Two-character abbreviation for the domain.	text	2		
USUBJID		Identifier used to uniquely identify a subject across all studies for all applications or submissions involving the product.	text	40		
VSSEQ		Sequence Number given to ensure uniqueness of subject records within a domain. May be any valid number.	integer	4		
VSTESTCD		Short name of the measurement, test, or examination described in VSTEST. It can be used as a column name when converting a dataset from a vertical to a horizontal format. The value in VSTESTCD cannot be longer than 8 characters, nor can it start with a number (e.g. "1TEST"). VSTESTCD cannot contain characters other than letters, numbers, or underscores. Examples: SYSBP, DIABP, BMI.	text	8	Vital Signs Test Code [6 Terms]	
VSTEST		Verbatim name of the test or examination used to obtain the measurement or finding. The value in VSTEST cannot be longer than 40 characters. Examples: Systolic Blood Pressure, Diastolic Blood Pressure, Body Mass Index.	text	40	Vital Signs Test Name [6 Terms]	
VSORRES VLM		Result of the vital signs measurement as originally received or collected.	text	30		
	VSTESTCD = "TEMP"	VSORRES for Temperature	float	4		
	VSTESTCD = "HEIGHT"	VSORRES for Height	float	5		
	VSTESTCD = "DIABP"	VSORRES for Diastolic Blood Pressure	integer	3		
	VSTESTCD = "SYSBP"	VSORRES for Systolic Blood Pressure	integer	3		
	VSTESTCD = "WEIGHT"	VSORRES for Weight	float	6		
	VSTESTCD = "HR"	VSORRES for Heart Rate	integer	3		
VSORRESU VLM		Original units in which the data were collected. The unit for VSORRES. Examples: IN, LB, BEATS/MIN.	text	20		
	VSTESTCD = "TEMP"	VSORRESU for Temperature	text		Units for Vital Signs Results Temperature • "C" • "F"	



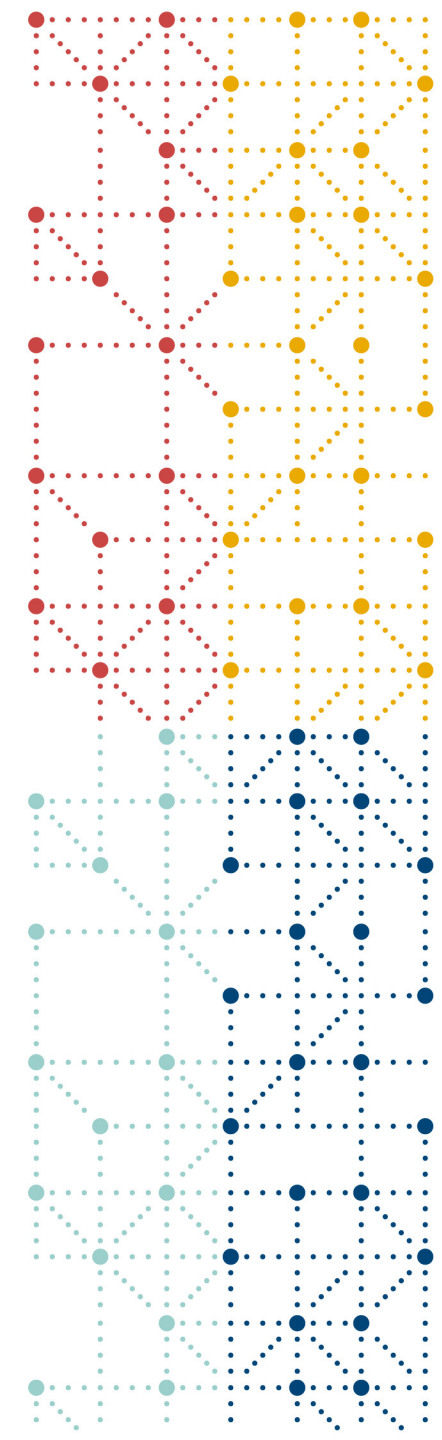
Next Steps – Actively Working on...

- Mapping & Derivations:

- Creating Pseudocode representations of the algorithms
- Will add executable reference implementations
- Represented well in the new MethodDef created for ODM v2.0
- Mapping and transformation logic needed to support end-to-end data flow
- Derivations needed to support analysis

- Analysis Concepts:

- Analysis Concepts enable the generation of tables, listings or figures used to analyze study endpoints or important safety concerns.

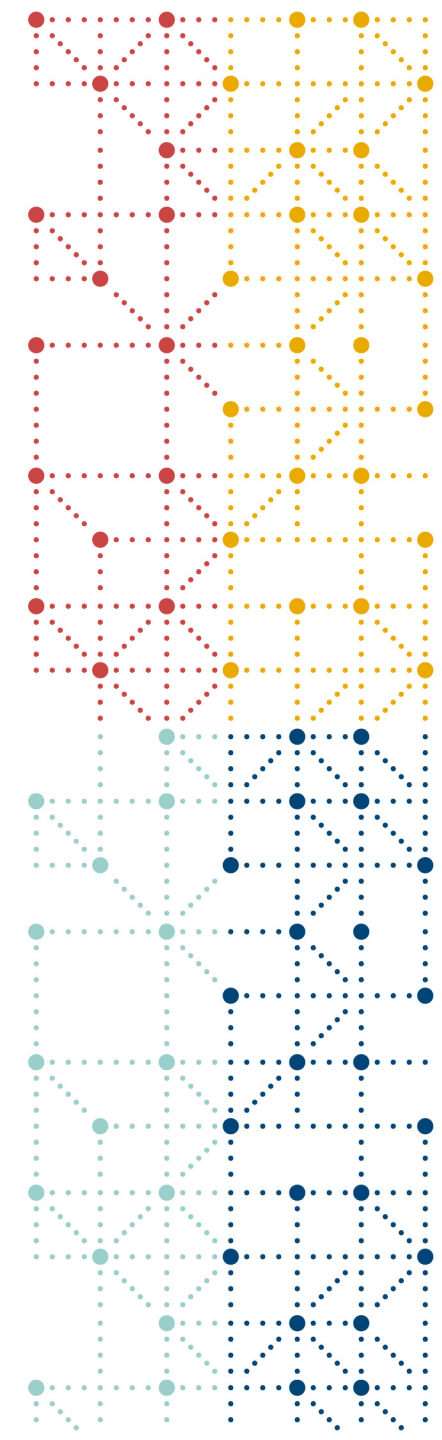


Proposal

A proposed method to improve the CDISC Therapeutic Area User Guides using the CDISC 360 metadata for comment.

Proposal: Publish CDISC Therapeutic Area User Guides as Templates

- Publish CDISC Therapeutic Area as CDISC 360 Templates
 - Machine-readable metadata can be used to create CRFs or Define-XML files
- Publish open source code to generate ODM CRF and Define-XML examples
- For future TAUGs, publish enough metadata to create an simple example study suitable for automation
- Refine Templates to create your own studies from the examples



Questions?



THANK
YOU