What is CDISC 360: an enterprise architecture perspective

Sam Hume, DSc - VP, Data Science, CDISC
Tianna Umann – Cross Domain Solution Architect, Microsoft
28 April 2020
CDISC 360 Metadata Inventory

Available Today

Additional CDISC 360 Content

Foundational Standards
Controlled Terminology

Biomedical Concepts
Codelist Subsets
TAUG metadata
Study design metadata
Tables, Listings, and Figures
Derivations and Mappings

Analysis Concepts
Templates
CDISC 360 Perspective 1: Metadata Flow within the Project

Enhance Standards → Publish Standards

Study Metadata Library

Define → Build → Execute
CDISC 360 Perspective 2: End-to-end Data Flow

Data Flow Diagram: Data State and Data Flow

Study Design → Data Collection Activities → CDASH Datasets → SDTM Mapping → SDTM Datasets → ADaM Mapping → ADaM Datasets → Analysis Spec → TLFs

Increased Automation & Reduced Variability
CDISC 360 Perspective 2: Code as Metadata

- Data flow metadata to capture logic needed to drive data **transformations** and **derivations**

- Capture Pseudocode descriptions

- Capture input and output variables

- Add multiple executable reference implementations (e.g. SAS, R, Python)

- Represented in ODM and Define-XML as a MethodDef
CDISC 360 Perspective 3: CDISC Library

M3: Meta-Meta-Model
- Production: W3C RDF
- Sandbox: JSON or XML Schema

M2: Meta-Model
- ISO 11179

M1: Model
- CDISC Standards

M0: Model Runtime
- Trial Metadata and Data
CDISC 360 Perspective 3: CDISC Library

ISO 11179-based Model

CDISC 360 Biomedical Concepts

CDISC IG Standards Metadata
CDISC 360 Perspective 3: CDISC Library

CDISC Standards Model

- Model
  - Model Class
    - Dataset
      - Dataset Variable
    - Class Variable
      --OUT

- IG
  - IG Class
    - Dataset
      - Dataset Variable
    - AEOOUT
      Adverse Event Outcome

SDTM 1.4
SDTM IG 3.2
AE
Adverse Events
Events
Events

Model
AEOUT
Adverse Event Outcome
CDISC 360 Perspective 4: External Project Data Flow

TransCelerate Digital Data Flow (DDF)

Study Design Data Sources

Sponsor Operational Data
Sponsor Clinical Data
External (Public) Study Data
Real World Data
Additional Sources

Study Design Repositories (Stored using Common Data Model)

Study Builder
Objectives, Endpoints, etc...

Core Components: Versioning, Workflow, Security

Simulation, Machine Learning, etc...

Study Design APIs (CDM)

Deployment Engine

Data Mapping (CDM to System)

DDF Solution Components

Common Protocol Templates

Historical and Actual Data

Protocol
CSR
SAP
EDC
CTMS
PPM
IRT
eTMF
Additional Systems

Instance of

DDF Compatible System
Existing (non-compatible) System
Downstream Clinical and Operational Systems

Legend:
Data Scope
DDF Components
DDF Compatible System
Existing Components/Systems

Multiple possible vendors

Copyright ©2017 TransCelerate BioPharma Inc., All rights reserved.
Business
Metadata
Application
Technology
CDISC 360 Applications Summary

CDISC Sandbox Library
- JSON BCs and Templates

Transform Concepts and Templates
- Python
- RDF

CMAPs
- XLSX

Import Standards
- Neo4j Cypher

Study Designer
- Python Django
- Neo4j Cypher

TLF Designer
- R Shiny
- Python Django
- Neo4j Cypher

Study Metadata Library
- Neo4j Graph DB

CDISC Library
- JSON IGs & CT

Study Specification Generator
- Python

Ingest Specification Metadata
- SAS

Data Transformation Engine
- SAS

SDTM
ADaM
ARM & TLFs

Study Test Data CDR
- CSV / CDM / SAS

TFL Automation Engine
- R / SAS
WS1 Enhance Standards: CMAP & Excel

Biomedical & Analysis Concepts

Templates, Codelist Subsets, & Methods

Proof-of-Concept Metadata Creation
WS2 Publish Standards: CDISC Library

CDISC 360 Created Metadata

- Bespoke code to generate metadata
  - Python
  - RDF

- Generate machine-readable metadata published to a CDISC Library Sandbox

CDISC Foundational Standards & CT

Publish Standards Metadata via the CDISC Library or Sandbox Library
WS2 Publish Standards: Sandbox CDISC Library

Workstream 1
- Concept Maps & Templates
  - Export
  - XML
  - Reference

Azure Files
  - REST API
  - API Service #1 (Azure Function)

API Gateway
  - API Management
  - REST API

Workstreams 4, 5, 6
- POC Apps
  - Developers
  - Consume API Content

Developers
  - Develop App

Current CDISC Library APIs
  - REST API

Application
  - API Service #n (Azure Function)
WS4 Define: Study Metadata Library

Study Designer App

Define Study
Design Study
Select and Configure Standard concepts for a Study
Build Study Specification

Study Metadata Library

Import Standards and concepts
Supplemental Metadata
Sponsor Library Maintenance

List Metadata (Query Interface)

Study Specification Artefacts

Neo4j SAS Interface
WS5 – Build: Study Designer

UI for creating study metadata
Please select vital signs concepts

<table>
<thead>
<tr>
<th>Concept</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td>The vertical measurement or distance from the base to the top of an object, the vertical dimension of extension.</td>
</tr>
<tr>
<td>Weight</td>
<td>The vertical force exerted by a mass as a result of gravity.</td>
</tr>
<tr>
<td>Heart rate</td>
<td>The number of heartbeats per unit of time, usually expressed as beats per minute.</td>
</tr>
<tr>
<td>Temperature</td>
<td>The property of a body or region of space that determines whether or not there will be a net flow of heat into it or out of it from a neighboring body or region.</td>
</tr>
<tr>
<td>Blood Pressure</td>
<td>The pressure of the circulating blood against the walls of the blood vessels.</td>
</tr>
<tr>
<td>Respiratory Rate</td>
<td>The rate of breathing (inhalation and exhalation) measured within a unit of time, usually expressed as breaths per minute.</td>
</tr>
<tr>
<td>Oxygen Saturation</td>
<td>A measurement of the oxygen-hemoglobin saturation of a volume of blood.</td>
</tr>
<tr>
<td>Hip Circumference</td>
<td>The distance around an individual's perineum or hips.</td>
</tr>
<tr>
<td>Body Fat Measurement</td>
<td>A measurement of the total fat mass within the subject's body.</td>
</tr>
<tr>
<td>Body Frame Size</td>
<td>The categorization of a person's body frame into small, medium and large based on the measurement of wrist circumference or the breadth of the skull.</td>
</tr>
</tbody>
</table>
WS5 Build: Metadata Artifacts

- Bespoke code to generate:
  - ODM CRFs / aCRFs
  - Define-XML

- ODM and Define-XML visualizations
  - Stylesheets
  - Formedix On

- End-to-end with ODM and Define-XML
  - Test driving automation with ODM and Define-XML metadata
WS6 Execute: Data Transformation Engine & SAS

Data Transformation Engine (DTE) from Noumena Solutions (Greg Steffens)
WS6 Execute: TLF Automation Engine

- CDISC 360 Enriched ARM Metadata
- Study Level ADaM Data

R Shiny
- Select TFL of Interest
- Select TFL Layout (Template)
- Review data
- Generate SAS Program and Define.xml
- Customize TFL Layout & Metadata

SAS
- Execute SAS Program
- Generate Output
- Validate and Deliver
Business

Metadata

Application

Technology
Technology: Code Development and Execution

Code Development & Execution

Cypher Query Language

Visual Studio Code

CmapTools

TopBraid Composer™

jupyter

Azure Functions

Azure Jupyter Notebooks

python

SAS

django

RStudio

Shiny

cdisc
CDISC 360: Cloud-based Computing

CDISC 360 Azure Cloud

CDISC Library

CDISC collaboration tools

CMAP
Future State

How does research change once we implement the learnings from CDISC 360?
Proof of Concept to Production

- Architecture Patterns
- Evolution of Cloud platforms
- Automation
- Multi Data Source ELT
Serverless is the culmination of several iterations of cloud platforms. The evolution began with physical metal in the data center and progressed through infrastructure as a Service (IaaS) and Platform as a Service (PaaS).
Serverless Event Grid, Logic Apps, Functions

Platform

Event Grid
Manage all events that can trigger code or logic

Logic Apps
Design workflows and orchestrate processes

Functions
Execute your code based on events you specify

Database
Storage
Security
IoT
Analytics
Intelligence
Thank You!
Monoliths

Advantages
• Easy to re-create an environment with multiple components and applications

Disadvantages
• Tends to break down at scale
• Difficult to work in parallel on code
• Any change requires deploying new version of the application
Microservices
Parse a CSV file and insert into a database

- File added to Blob Storage
- Transform CSV to data rows
- Updated SQL Table
N-Layer applications

UI Layer
- Web Pages
- Mobile App

Business Logic Layer
- APIs
- Validations

Data Access Layer
- Cache
- Database access
Workstreams 4-6 POC Target Architecture - Draft

Workstream 2
POC
Azure API Management
API Gateway

App Service #1
(Azure Function)

App Service #2
(Azure Function)

App Service #n
(Azure Function)

POC User Interface(s)
(Azure Web Apps)

REST API

REST API

REST API

neo4j

SAS

Virtual Network

VM

NS
NSG
Data pipeline

- File added to Blob Storage
- Transform CSV to data rows
- Power BI Chart graphic