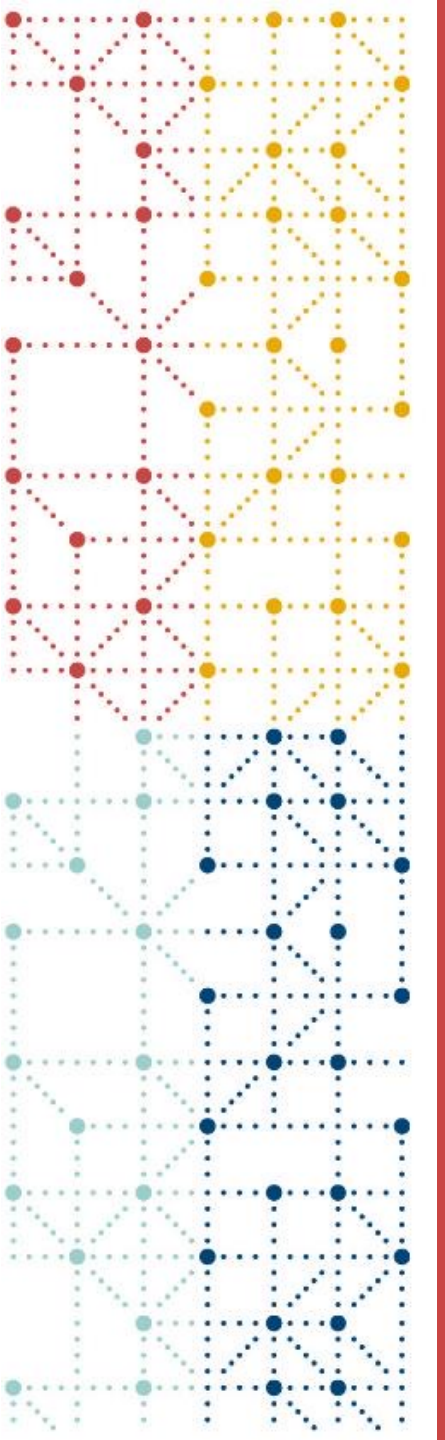
Enhancing Clinical Research Efficiency: Leveraging CDISC Biomedical Concepts for Automation

Bess LeRoy, Head of Standards Development, CDISC



Agenda

- Why Now?
- Biomedical Concepts in CDISC 360i
- Biomedical Concepts Development Update

Why Now?

Biomedical Concepts (BCs) Support FAIR Data Principles

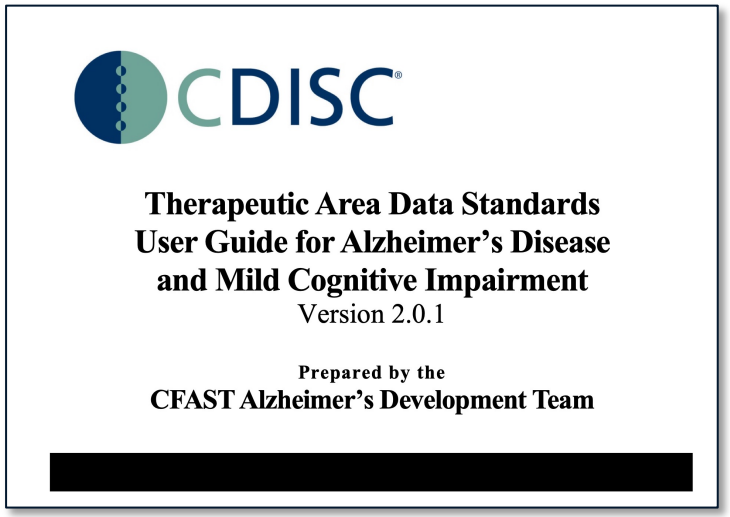
- **Findable**: Sufficient metadata, unique and persistent identifier, data must be registered or indexed in a searchable resource.
- **Accessible**: Metadata should be readable by humans and by machines
- **Interoperable**: Data must share a common structure, and metadata must use recognized, formal terminologies for description.
- **Reusable**: Data and collections must have clear usage licenses and clear provenance and meet relevant community standards for the domain.

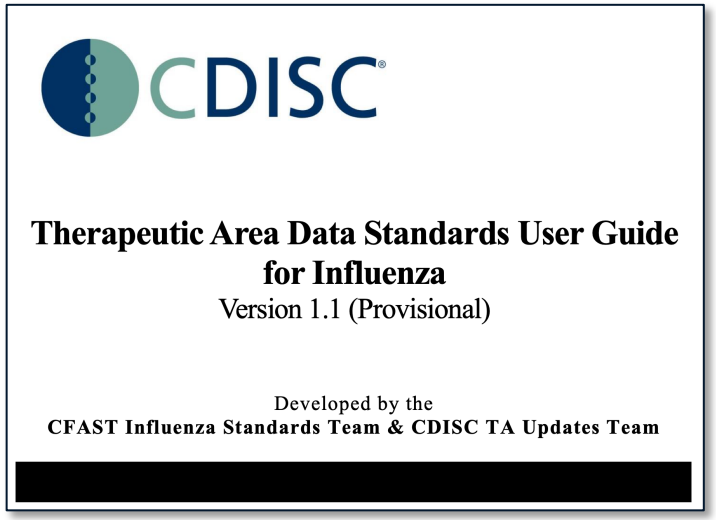
Sources: Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

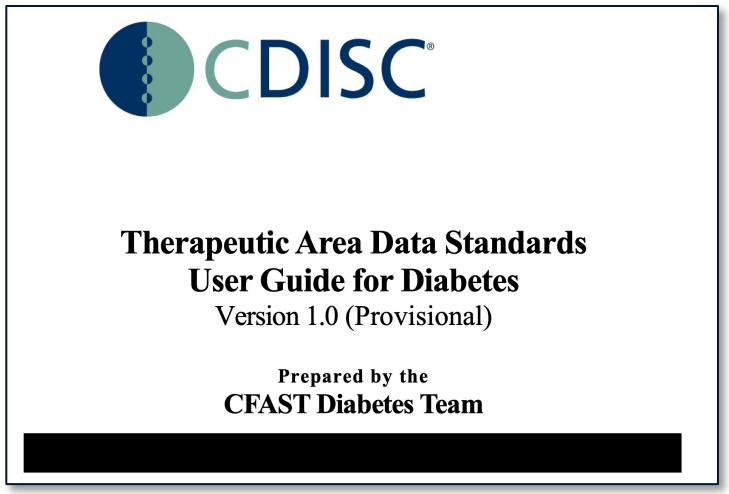
https://www.nlm.nih.gov/oet/ed/cde/tutorial/02-200.html

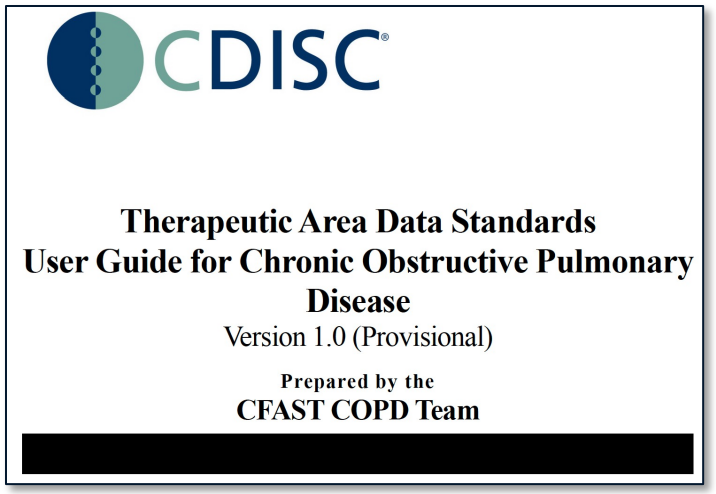
https://www.euro-argo.eu/content/download/146491/file/SeaTechWeek_S.Pouliquen_ENVRI-FAIR.pdf

CDISC Therapeutic Area Guides Support *Human Readable* Accessibility



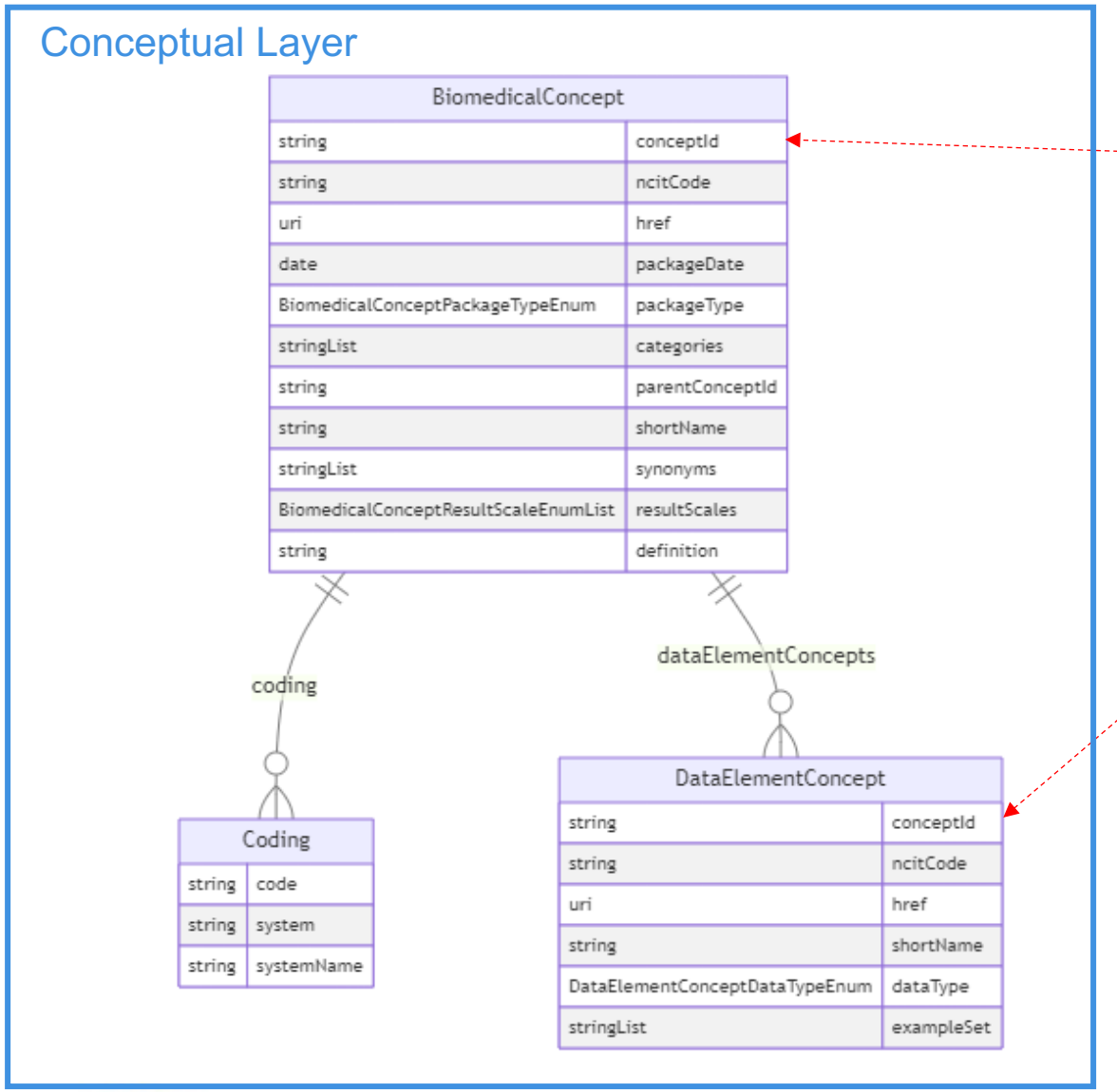




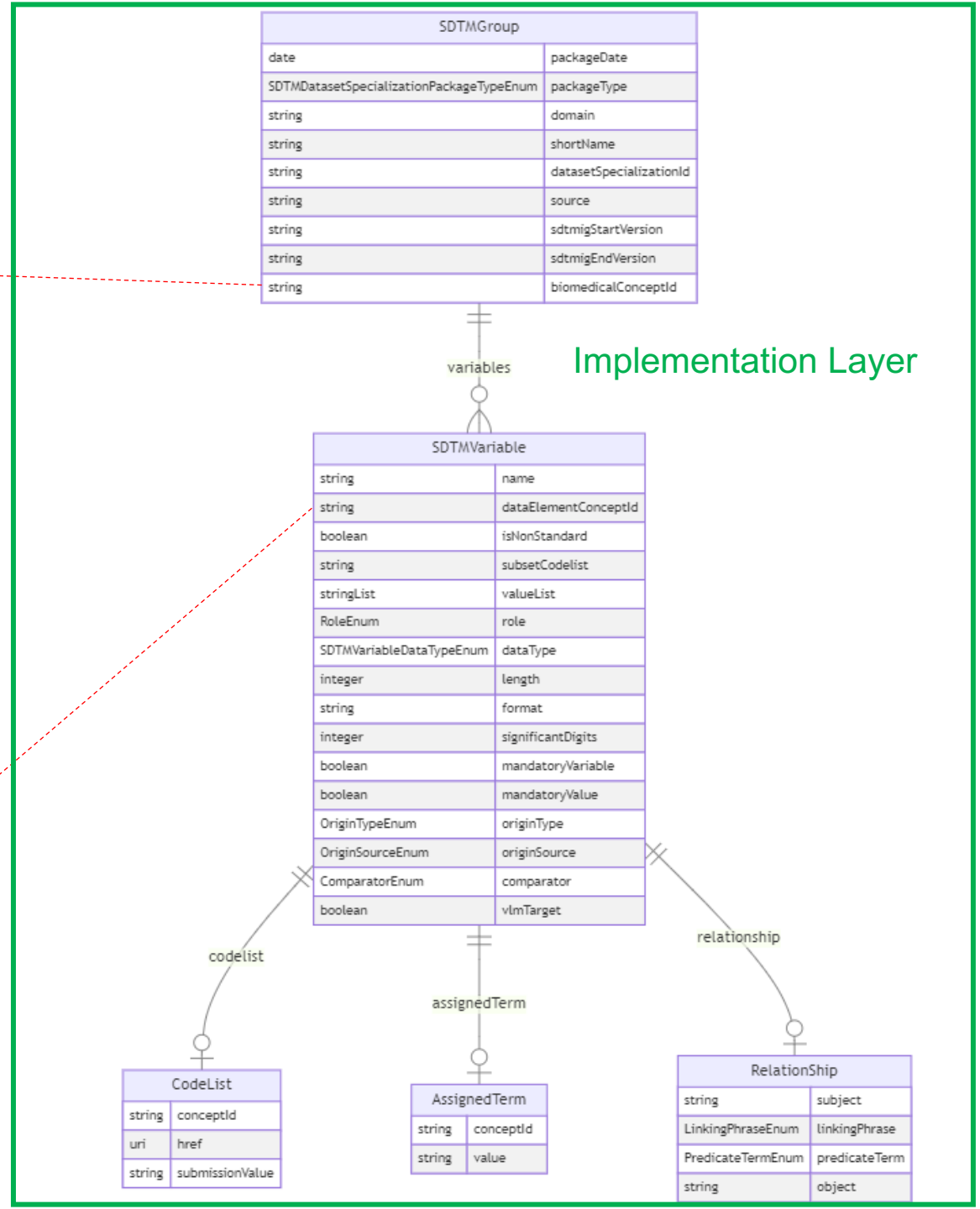


BC Logical Data Model

Conceptual Layer



Implementation Layer



PHUSE Data Standards Heads White Paper

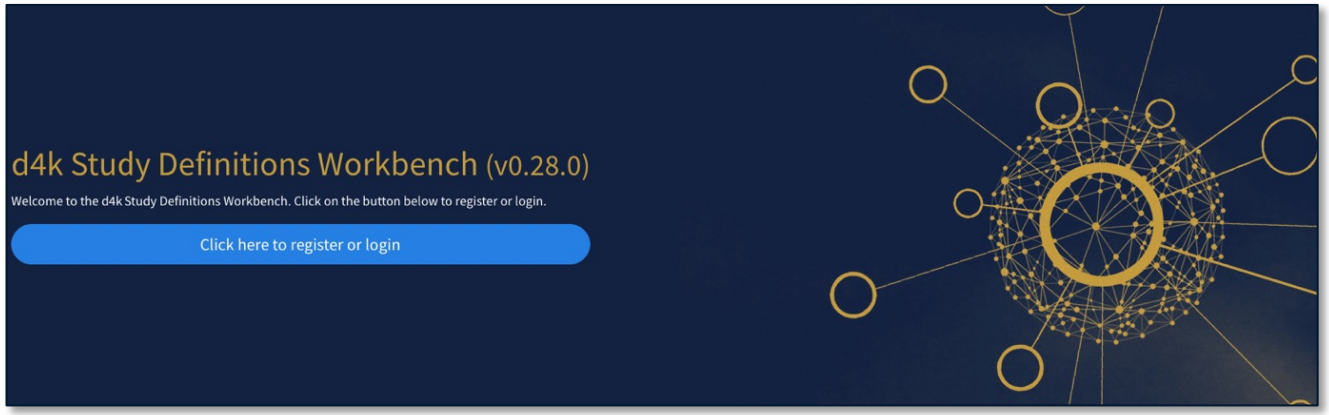
- A paradigm shift is needed in the way industry data standards and terminologies are developed:
  - BCs should play a central role
  - BCs should be represented in a “standards-agnostic” way
  - BCs can help the convergence of clinical care & clinical research
  - BCs are vital to delivering connected end to end data standards
  - BCs are a key enabler to automation.

April 4th, 2024

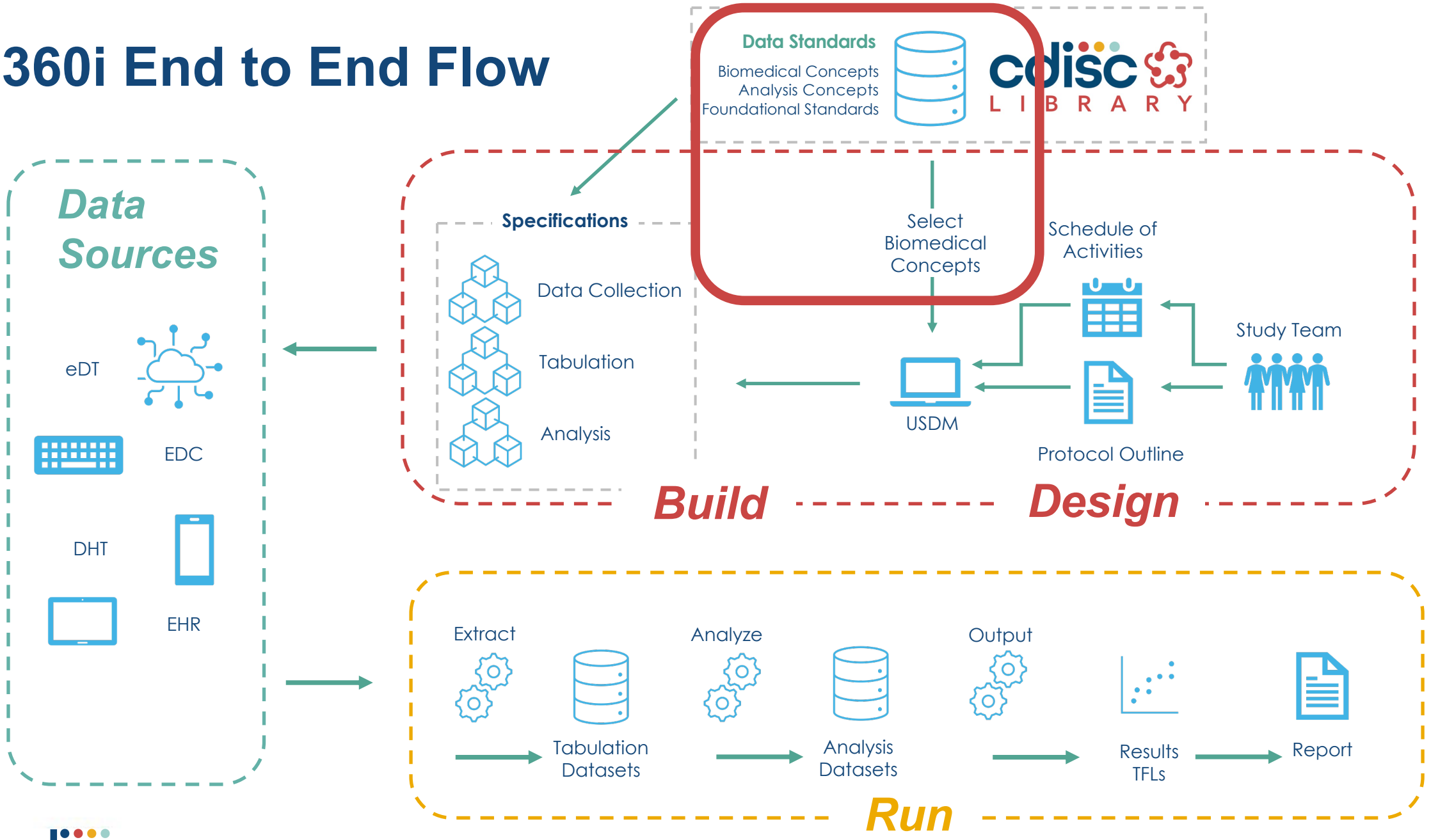
Courtesy of Jonathan Chainey, US Interchange 2024

Open-Source Tools That Leverage BCs

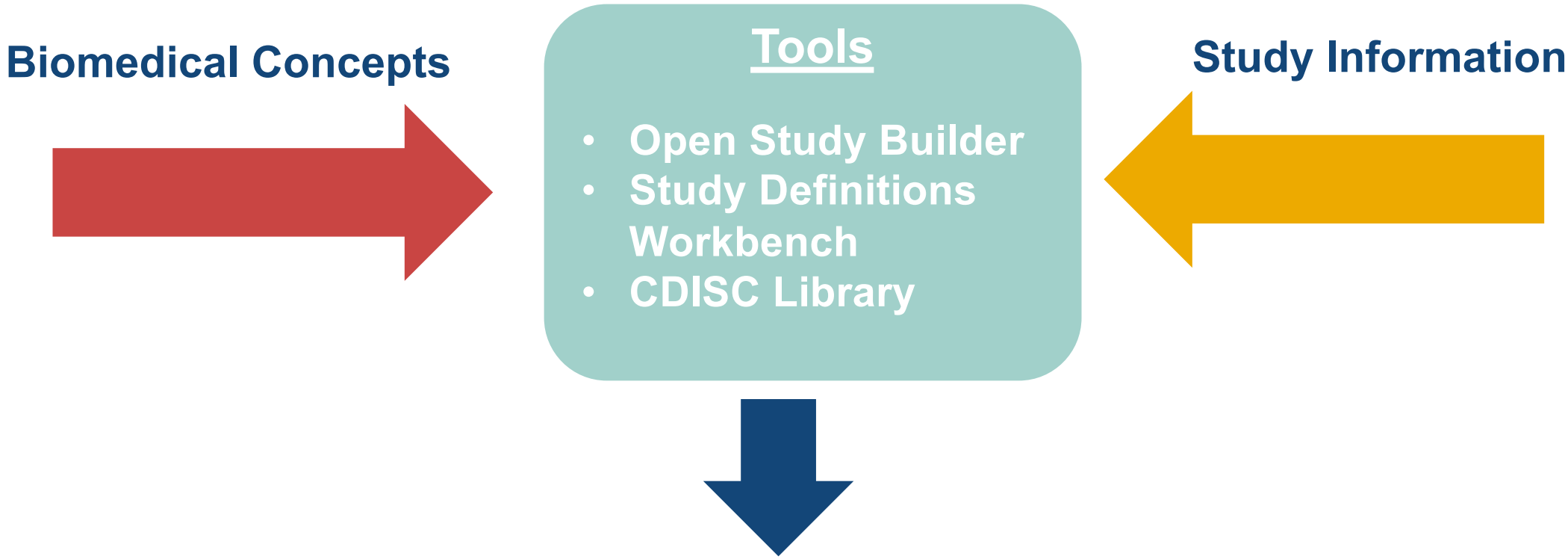
OPEN STUDY BUILDER



360i End to End Flow



BCs to Support USDM Compliant Study Design



Design

BCs to Generate SDTM Define-XML

Inputs

- Biomedical Concepts
- Dataset Specializations
- Study design in ODM
- Study design in JSON

Tools

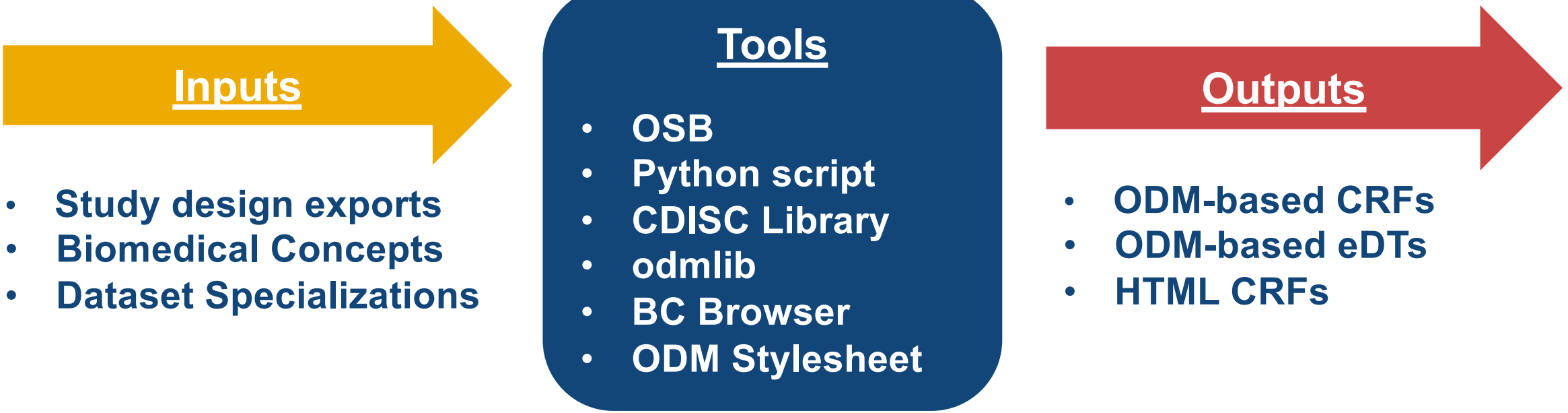
- SAS, Python scripts
- odmlib
- CDISC Library
- BC Browser
- XMLMAPs
- Define-XML Stylesheet

Outputs

- Define-XML with VLM
- HTML Define-XML

Build

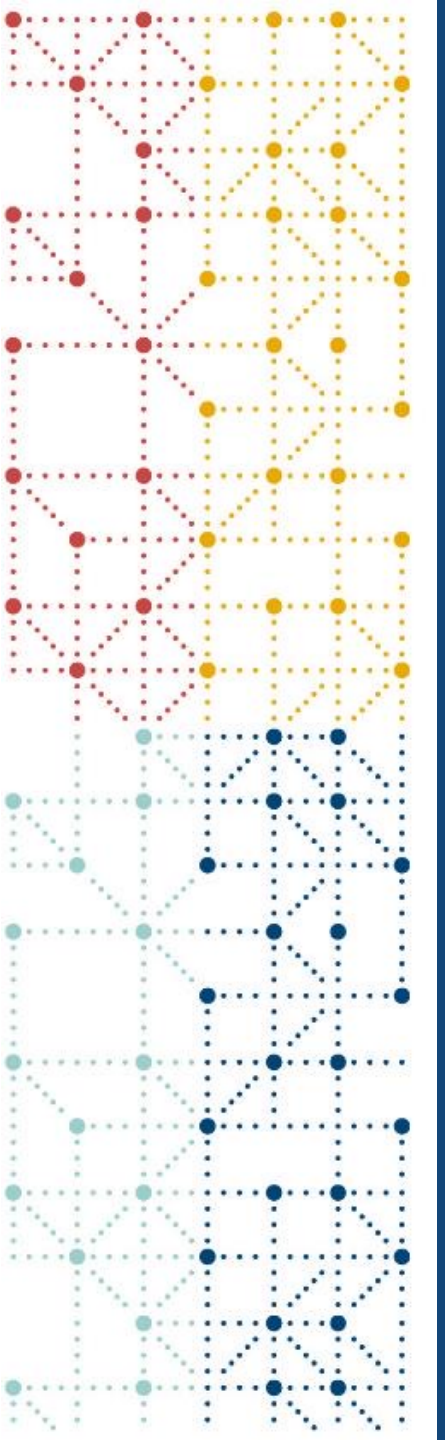
BCs to Generate CRFs and Electronic Data Transfers



- Study design exports
- Biomedical Concepts
- Dataset Specializations

- ODM-based CRFs
- ODM-based eDTs
- HTML CRFs

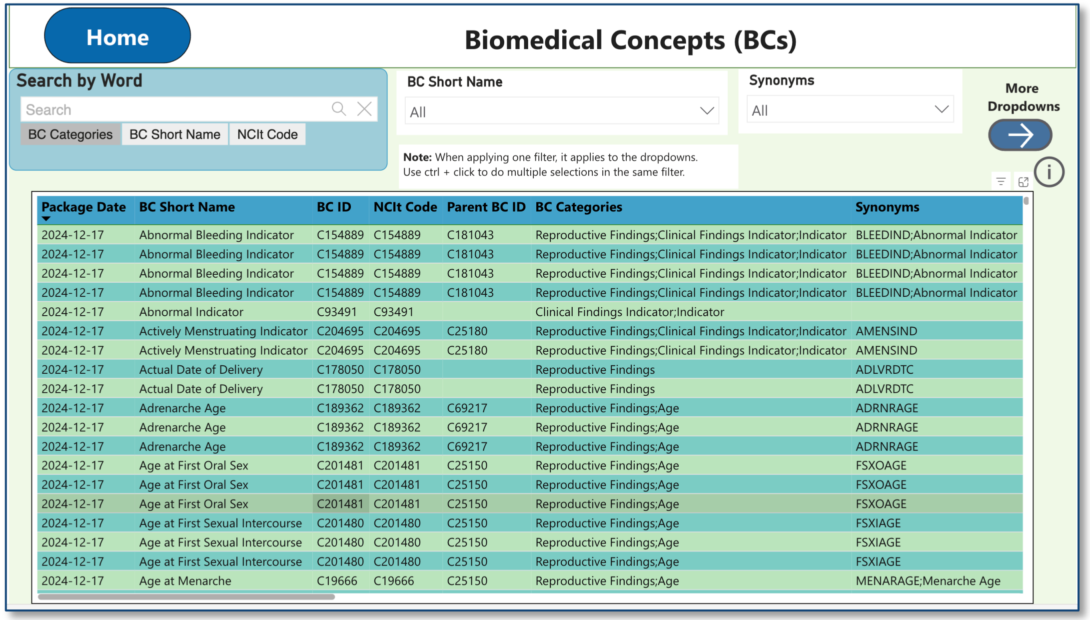
Build



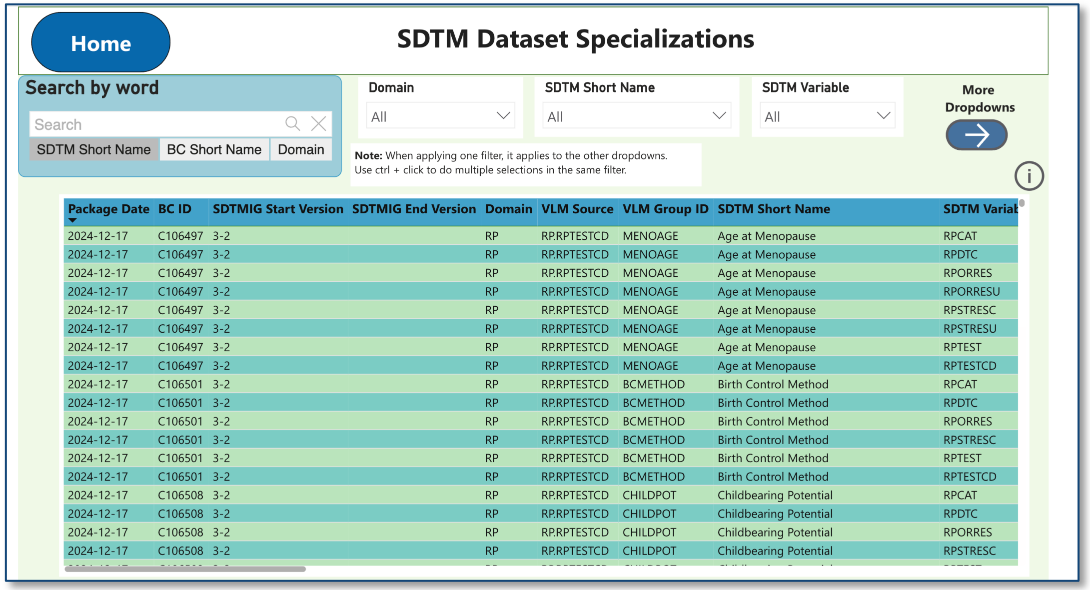
Biomedical Concepts Development Update

Biomedical Concepts

- Total BCs = 568; Total SDTM Specializations = 470
- Current focus
  - Support for 360i
  - QRS instruments
  - Digital Health Technologies
  - CDASH implementation layer
  - FHIR implementation layer
- Initial planning for Collaborative Curation BCs from our community
- First BC Training Course developed and delivered at the European Interchange



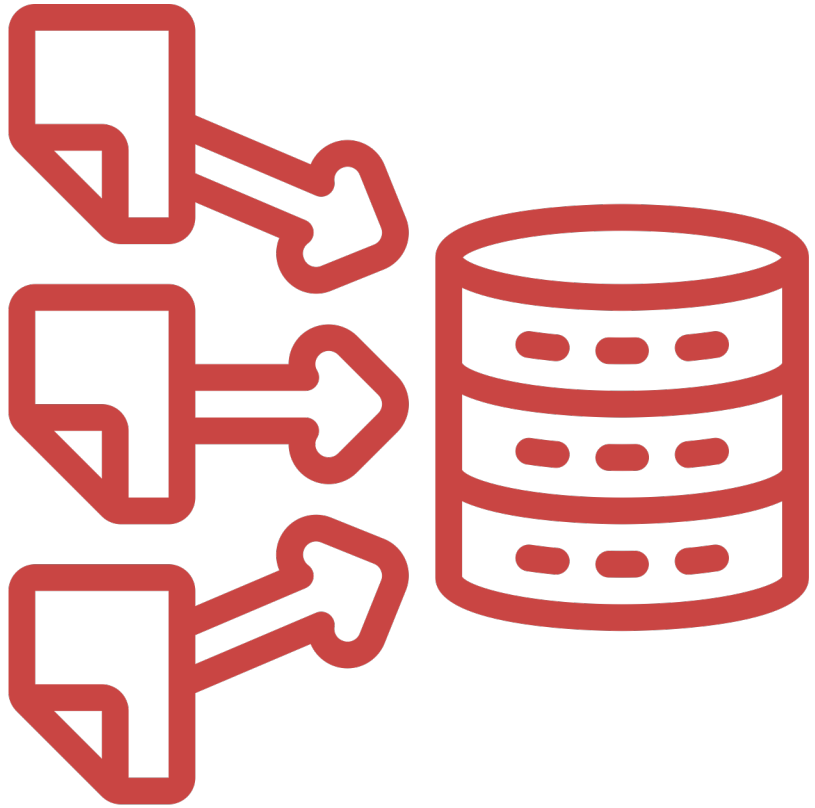
Package Date	BC Short Name	BC ID	NCIt Code	Parent BC ID	BC Categories	Synonyms
2024-12-17	Abnormal Bleeding Indicator	C154889	C154889	C181043	Reproductive Findings;Clinical Findings Indicator;Indicator	BLEEDIND;Abnormal Indicator
2024-12-17	Abnormal Bleeding Indicator	C154889	C154889	C181043	Reproductive Findings;Clinical Findings Indicator;Indicator	BLEEDIND;Abnormal Indicator
2024-12-17	Abnormal Bleeding Indicator	C154889	C154889	C181043	Reproductive Findings;Clinical Findings Indicator;Indicator	BLEEDIND;Abnormal Indicator
2024-12-17	Abnormal Bleeding Indicator	C154889	C154889	C181043	Reproductive Findings;Clinical Findings Indicator;Indicator	BLEEDIND;Abnormal Indicator
2024-12-17	Abnormal Indicator	C93491	C93491		Clinical Findings Indicator;Indicator	
2024-12-17	Actively Menstruating Indicator	C204695	C204695	C25180	Reproductive Findings;Clinical Findings Indicator;Indicator	AMENSIND
2024-12-17	Actively Menstruating Indicator	C204695	C204695	C25180	Reproductive Findings;Clinical Findings Indicator;Indicator	AMENSIND
2024-12-17	Actual Date of Delivery	C178050	C178050		Reproductive Findings	ADLVRDTC
2024-12-17	Actual Date of Delivery	C178050	C178050		Reproductive Findings	ADLVRDTC
2024-12-17	Adrenarche Age	C189362	C189362	C69217	Reproductive Findings;Age	ADRNRAGE
2024-12-17	Adrenarche Age	C189362	C189362	C69217	Reproductive Findings;Age	ADRNRAGE
2024-12-17	Adrenarche Age	C189362	C189362	C69217	Reproductive Findings;Age	ADRNRAGE
2024-12-17	Age at First Oral Sex	C201481	C201481	C25150	Reproductive Findings;Age	FSXOAGE
2024-12-17	Age at First Oral Sex	C201481	C201481	C25150	Reproductive Findings;Age	FSXOAGE
2024-12-17	Age at First Oral Sex	C201481	C201481	C25150	Reproductive Findings;Age	FSXOAGE
2024-12-17	Age at First Sexual Intercourse	C201480	C201480	C25150	Reproductive Findings;Age	FSXIAGE
2024-12-17	Age at First Sexual Intercourse	C201480	C201480	C25150	Reproductive Findings;Age	FSXIAGE
2024-12-17	Age at First Sexual Intercourse	C201480	C201480	C25150	Reproductive Findings;Age	FSXIAGE
2024-12-17	Age at Menarche	C19666	C19666	C25150	Reproductive Findings;Age	MENARAGE;Menarche Age



Package Date	BC ID	SDTMIG Start Version	SDTMIG End Version	Domain	VLM Source	VLM Group ID	SDTM Short Name	SDTM Variab
2024-12-17	C106497	3-2		RP	RP.RPTESTCD	MENOAGE	Age at Menopause	RPCAT
2024-12-17	C106497	3-2		RP	RP.RPTESTCD	MENOAGE	Age at Menopause	RPDTC
2024-12-17	C106497	3-2		RP	RP.RPTESTCD	MENOAGE	Age at Menopause	RPORRES
2024-12-17	C106497	3-2		RP	RP.RPTESTCD	MENOAGE	Age at Menopause	RPORRESU
2024-12-17	C106497	3-2		RP	RP.RPTESTCD	MENOAGE	Age at Menopause	RPSTRESC
2024-12-17	C106497	3-2		RP	RP.RPTESTCD	MENOAGE	Age at Menopause	RPSTRESU
2024-12-17	C106497	3-2		RP	RP.RPTESTCD	MENOAGE	Age at Menopause	RPTEST
2024-12-17	C106497	3-2		RP	RP.RPTESTCD	MENOAGE	Age at Menopause	RPTESTCD
2024-12-17	C106501	3-2		RP	RP.RPTESTCD	BCMETHOD	Birth Control Method	RPCAT
2024-12-17	C106501	3-2		RP	RP.RPTESTCD	BCMETHOD	Birth Control Method	RPDTC
2024-12-17	C106501	3-2		RP	RP.RPTESTCD	BCMETHOD	Birth Control Method	RPORRES
2024-12-17	C106501	3-2		RP	RP.RPTESTCD	BCMETHOD	Birth Control Method	RPSTRESC
2024-12-17	C106501	3-2		RP	RP.RPTESTCD	BCMETHOD	Birth Control Method	RPTEST
2024-12-17	C106501	3-2		RP	RP.RPTESTCD	BCMETHOD	Birth Control Method	RPTESTCD
2024-12-17	C106508	3-2		RP	RP.RPTESTCD	CHILDPOT	Childbearing Potential	RPCAT
2024-12-17	C106508	3-2		RP	RP.RPTESTCD	CHILDPOT	Childbearing Potential	RPDTC
2024-12-17	C106508	3-2		RP	RP.RPTESTCD	CHILDPOT	Childbearing Potential	RPORRES
2024-12-17	C106508	3-2		RP	RP.RPTESTCD	CHILDPOT	Childbearing Potential	RPSTRESC

Data Collection Specializations

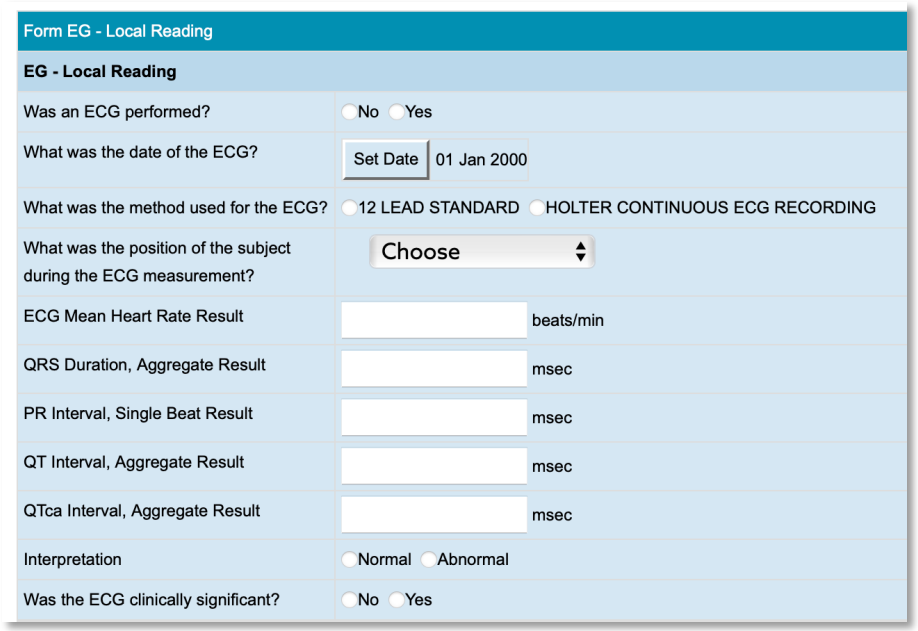
- Data Transfer Agreements
  - Collaboration with LAB v2.0 Team
  - Digital Health Technologies
- CDASH Implementation Layer
- HL7 FHIR Implementation Layer



CDASH Implementation Layer

- Connection to the SDTM Implementation Layer
- Implementation Option: Horizontal or Vertical layout
- Question Text and Prompt
- Value Display List

Leveraging the eCRF Portal to Develop the CDASH Implementation Layer



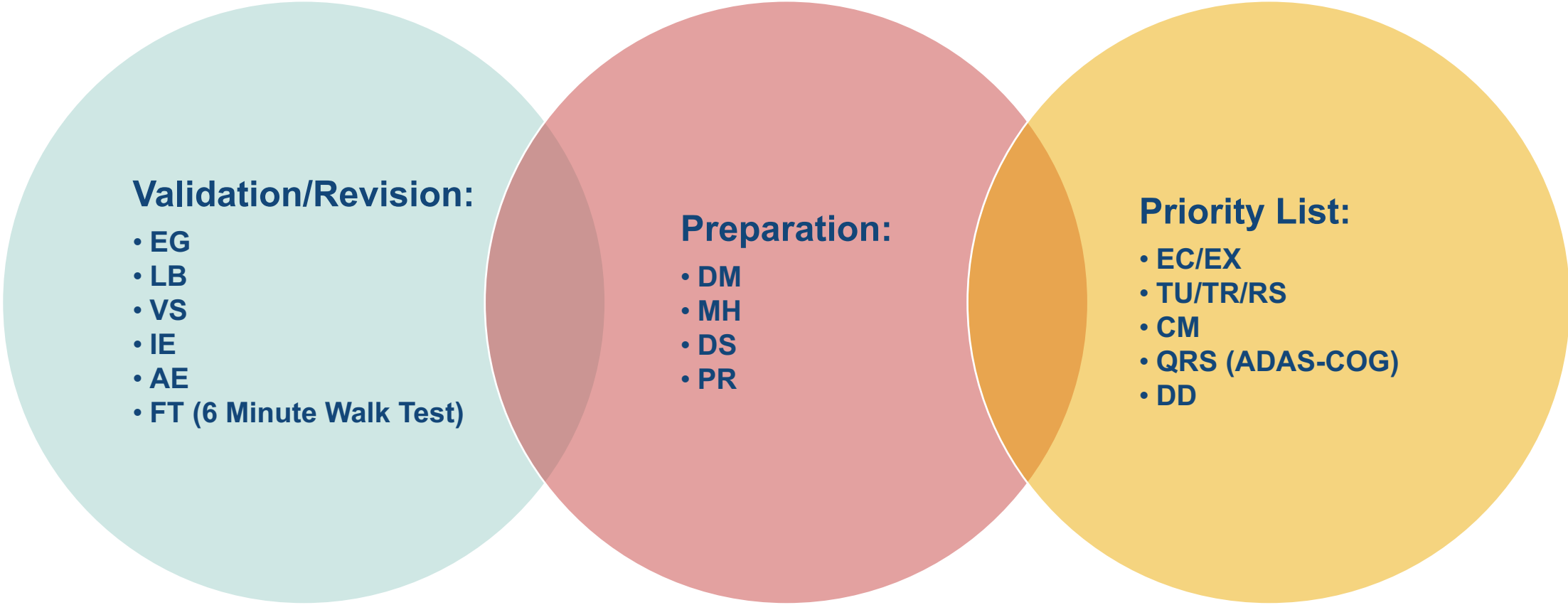
Form EG - Local Reading

EG - Local Reading

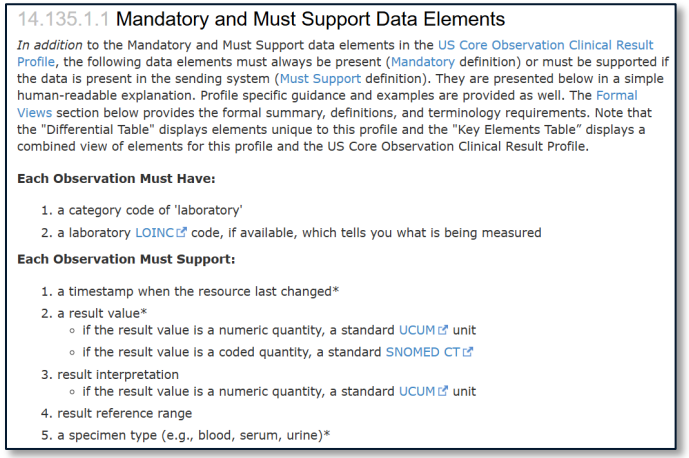
Was an ECG performed?	○ No ○ Yes
What was the date of the ECG?	Set Date 01 Jan 2000
What was the method used for the ECG?	○ 12 LEAD STANDARD ○ HOLTER CONTINUOUS ECG RECORDING
What was the position of the subject during the ECG measurement?	Choose
ECG Mean Heart Rate Result	beats/min
QRS Duration, Aggregate Result	msec
PR Interval, Single Beat Result	msec
QT Interval, Aggregate Result	msec
QTca Interval, Aggregate Result	msec
Interpretation	○ Normal ○ Abnormal
Was the ECG clinically significant?	○ No ○ Yes

bc_id	vlm_group_id	collection_group_id	implementation_option	scenario	short_name	collection_item	variable_name	question_text	prompt
C82525		EGPERF			ECG Performed	EGPERF	EGPERF	Was an ECG performed?	
C82525		EGPERF			ECG Performed	EGDAT	EGDAT	Date of Assessment	
C117779	QRSAG	QRSAG_VERTICAL	Vertical	LOCAL READING	QRS Duration Aggregate (Vertical)	EGDAT	EGDAT	What was the date of assessment?	
C117779	QRSAG	QRSAG_VERTICAL	Vertical	LOCAL READING	QRS Duration Aggregate (Vertical)	EGMETHOD	EGMETHOD	What was the method used for the ECG?	
C117779	QRSAG	QRSAG_VERTICAL	Vertical	LOCAL READING	QRS Duration Aggregate (Vertical)	EGPOS	EGPOS	What was the position of the subject during the ECG measurement?	
C117779	QRSAG	QRSAG_VERTICAL	Vertical	LOCAL READING	QRS Duration Aggregate (Vertical)	QRSAG	EGORRES	QRS Duration, Aggregate Result	
C117779	QRSAG	QRSAG_VERTICAL	Vertical	LOCAL READING	QRS Duration Aggregate (Vertical)	QRSAGU	EGORRESU	QRS Duration, Aggregate Result Unit	
C117779	QRSAG	QRSAG_HORIZONTAL	Horizontal	LOCAL READING	QRS Duration Aggregate (Horizontal)	EGDAT	EGDAT		Date of Assessment
C117779	QRSAG	QRSAG_HORIZONTAL	Horizontal	LOCAL READING	QRS Duration Aggregate (Horizontal)	EGMETHOD	EGMETHOD		Method
C117779	QRSAG	QRSAG_HORIZONTAL	Horizontal	LOCAL READING	QRS Duration Aggregate (Horizontal)	EGPOS	EGPOS		Position
C117779	QRSAG	QRSAG_HORIZONTAL	Horizontal	LOCAL READING	QRS Duration Aggregate (Horizontal)	QRSAG_EGORRES	EGORRES		QRS Duration, Aggregate Result
C117779	QRSAG	QRSAG_HORIZONTAL	Horizontal	LOCAL READING	QRS Duration Aggregate (Horizontal)	QRSAG_EGORRESU	EGORRESU		Unit

CDASH Specialization Development Status



HL7 FHIR Implementation Layer for Laboratory Result Observation Profiles



14.135.1.1 Mandatory and Must Support Data Elements

In addition to the Mandatory and Must Support data elements in the US Core Observation Clinical Result Profile, the following data elements must always be present (Mandatory definition) or must be supported if the data is present in the sending system (Must Support definition). They are presented below in a simple human-readable explanation. Profile specific guidance and examples are provided as well. The Formal Views section below provides the formal summary, definitions, and terminology requirements. Note that the "Differential Table" displays elements unique to this profile and the "Key Elements Table" displays a combined view of elements for this profile and the US Core Observation Clinical Result Profile.

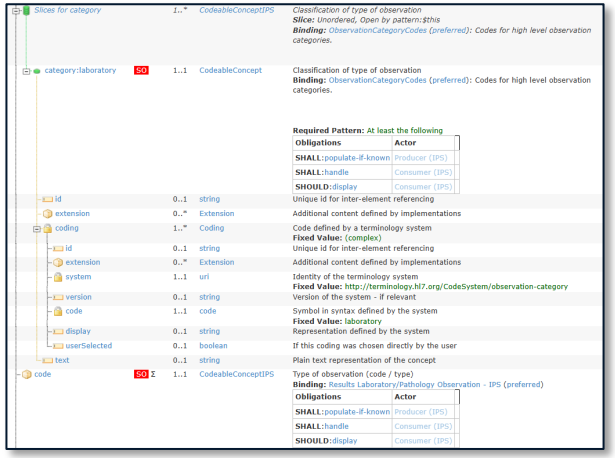
Each Observation Must Have:

1. a category code of 'laboratory'
2. a laboratory LOINC code, if available, which tells you what is being measured

Each Observation Must Support:

1. a timestamp when the resource last changed*
2. a result value*
   - if the result value is a numeric quantity, a standard UCUM unit
   - if the result value is a coded quantity, a standard SNOMED CT
3. result interpretation
   - if the result value is a numeric quantity, a standard UCUM unit
4. result reference range
5. a specimen type (e.g., blood, serum, urine)*

US Core Data for Interoperability



Slices for category 1..* CodeableConceptIPS
Classification of type of observation
Slice: Unordered, Open by pattern:$this
Binding: ObservationCategoryCodes (preferred): Codes for high level observation categories.

category:laboratory SO 1..1 CodeableConcept
Classification of type of observation
Binding: ObservationCategoryCodes (preferred): Codes for high level observation categories.

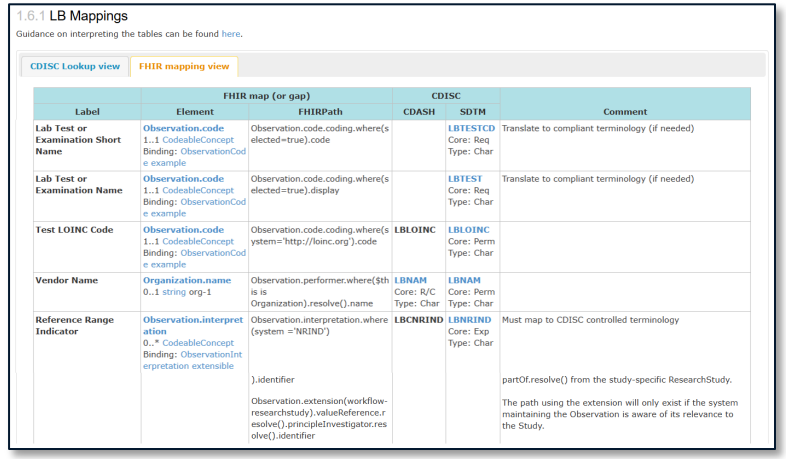
Required Pattern: At least the following

Obligations	Actor
SHALL:populate-if-known	Producer (IPS)
SHALL:handle	Consumer (IPS)
SHOULD:display	Consumer (IPS)

id 0..1 string — Unique id for inter-element referencing
extension 0..* Extension — Additional content defined by implementations
coding 1..* Coding — Code defined by a terminology system
Fixed Value: (complex)
id 0..1 string — Unique id for inter-element referencing
extension 0..* Extension — Additional content defined by implementations
system 1..1 uri — Identity of the terminology system
Fixed Value: http://terminology.hl7.org/CodeSystem/observation-category
version 0..1 string — Version of the system - if relevant
code 1..1 code — Symbol in syntax defined by the system
Fixed Value: laboratory
display 0..1 string — Representation defined by the system
userSelected 0..1 boolean — If this coding was chosen directly by the user
text 0..1 string — Plain text representation of the concept
code SO Σ 1..1 CodeableConceptIPS — Type of observation (code / type)
Binding: Results Laboratory/Pathology Observation - IPS (preferred)

Obligations	Actor
SHALL:populate-if-known	Producer (IPS)
SHALL:handle	Consumer (IPS)
SHOULD:display	Consumer (IPS)

International Patient Summary



1.6.1 LB Mappings

Guidance on interpreting the tables can be found here.

CDISC Lookup view | FHIR mapping view

	FHIR map (or gap)		CDISC		
Label	Element	FHIRPath	CDASH	SDTM	Comment
Lab Test or Examination Short Name	Observation.code 1..1 CodeableConcept Binding: ObservationCode example	Observation.code.coding.where(selected=true).code		LBTESTCD Core: Req Type: Char	Translate to compliant terminology (if needed)
Lab Test or Examination Name	Observation.code 1..1 CodeableConcept Binding: ObservationCode example	Observation.code.coding.where(selected=true).display		LBTEST Core: Req Type: Char	Translate to compliant terminology (if needed)
Test LOINC Code	Observation.code 1..1 CodeableConcept Binding: ObservationCode example	Observation.code.coding.where(system='http://loinc.org').code	LBLOINC	LBLOINC Core: Perm Type: Char	
Vendor Name	Organization.name 0..1 string org-1	Observation.performer.where($this is Organization).resolve().name	LBNAM Core: R/C Type: Char	LBNAM Core: Perm Type: Char	
Reference Range Indicator	Observation.interpretation 0..* CodeableConcept Binding: ObservationInterpretation extensible	Observation.interpretation.where(system ='NRIND').identifier	LBCNRIND	LBNRIND Core: Exp Type: Char	Must map to CDISC controlled terminology
		Observation.extension(workflow-researchstudy).valueReference.resolve().principleInvestigator.resolve().identifier			partOf.resolve() from the study-specific ResearchStudy. The path using the extension will only exist if the system maintaining the Observation is aware of its relevance to the Study.

FHIR to CDISC Mapping

doma	vlm_source	vlm_group_id	short_name	fhir_attribute_element_name	fhirPath_origin	definition	
LB	LB.LBTESTCD	ALBSERPL	Albumin Concentration in Serum/Plasma	code	Observation.code	Laboratory Test Name	h
LB	LB.LBTESTCD	ALBSERPL	Albumin Concentration in Serum/Plasma	code	Observation.code	Laboratory Test Name	h
LB	LB.LBTESTCD	ALBSERPL	Albumin Concentration in Serum/Plasma	category:us-core	Observation.category	A code that classifies the general type of observation being made.	h
LB	LB.LBTESTCD	ALBSERPL	Albumin Concentration in Serum/Plasma	valueQuantity	Observation.valueQuantity	Actual result	
LB	LB.LBTESTCD	ALBSERPL	Albumin Concentration in Serum/Plasma	valueQuantity.unit	Observation.valueQuantity.unit	Unit representation	h

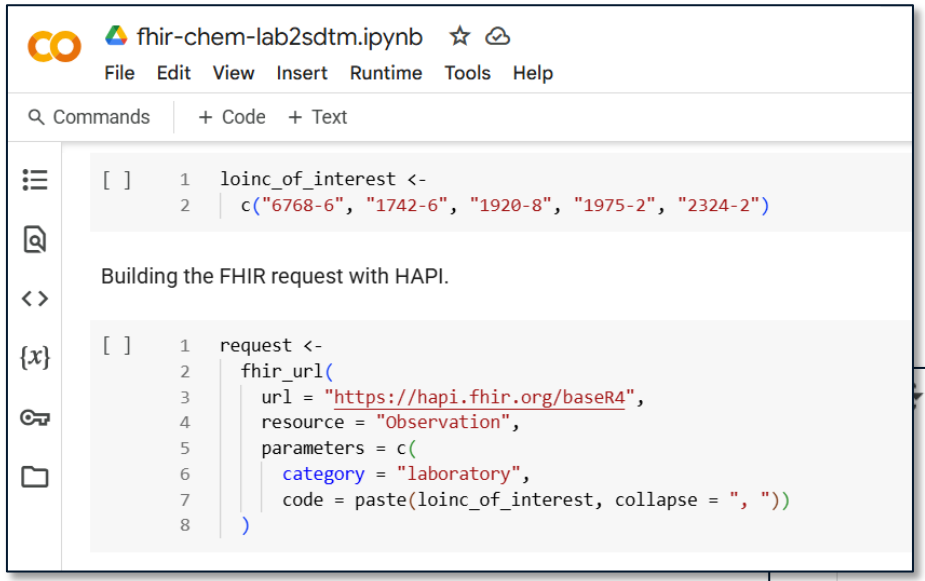
HL7 FHIR Implementation Layer: EHR as a Data Source

- Drug-Induced Liver Injury (DILI) is a critical concern in drug development and clinical practice
- Used in clinical trials for early detection of hepatotoxicity
- Lab data reviewed longitudinally to detect trends and flag risks
- Key markers per Hy's Law (US FDA 2009):
  - ALT (Alanine Aminotransferase)
  - AST (Aspartate Aminotransferase)
  - ALP (Alkaline Phosphatase)
  - Total Bilirubin

Clinical Laboratory Observation

HL7 FHIR Proof of Concept Implementation

R script that uses FHIR Implementation Layer to pull DILI related chemistry lab results from a FHIR R4 server and converts them into a tabular structure using SDTM nomenclature



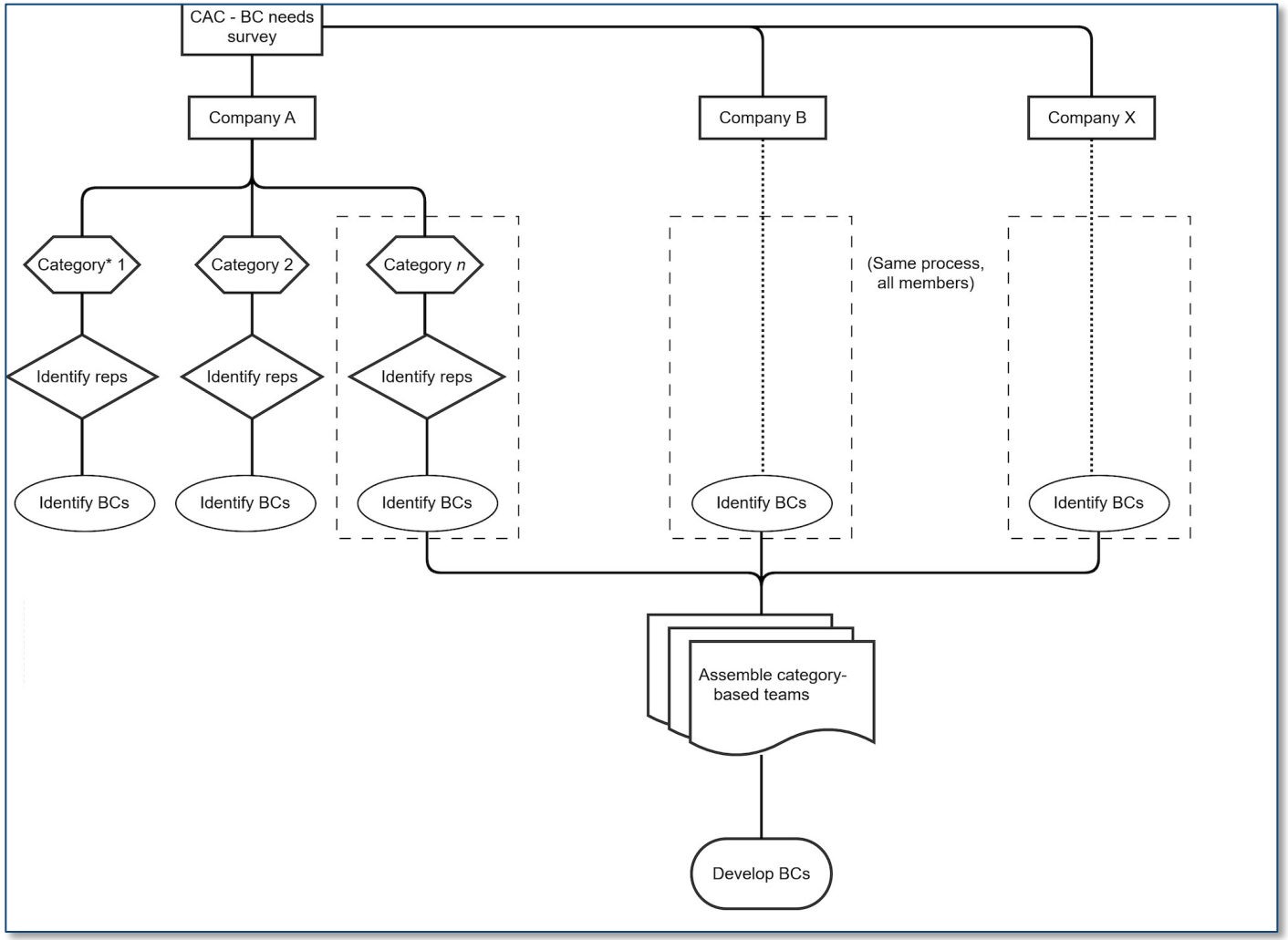
A data.frame: 4 × 4

	Test	LOINC	Unit	Freq
	<chr>	<chr>	<chr>	<int>
1	Alanine aminotransferase [Enzymatic activity/volume] in Serum or Plasma	1742-6	U/L	6
2	Alkaline phosphatase [Enzymatic activity/volume] in Serum or Plasma	6768-6	U/L	6
3	Aspartate aminotransferase [Enzymatic activity/volume] in Serum or Plasma	1920-8	U/L	21
4	Bilirubin.total [Mass/volume] in Serum or Plasma	1975-2	mg/dL	7

Courtesy of Anthony Chow

Collaborative Curation: Initial Thoughts

- Development by disease area
- Commitment of volunteer time by companies
- Specified time commitment by volunteers



BCs and AI

- The traditional approach to BC development is highly labor-intensive and time-consuming, often requiring advanced expertise in clinical research standards, medical terminology, and data modeling

- CDISC is collaborating with Lindus Health, which is utilizing its proprietary AI technology, to accelerate the generation of Biomedical Concepts across both the CDISC Foundational Standards as well as the Therapeutic Area Standards

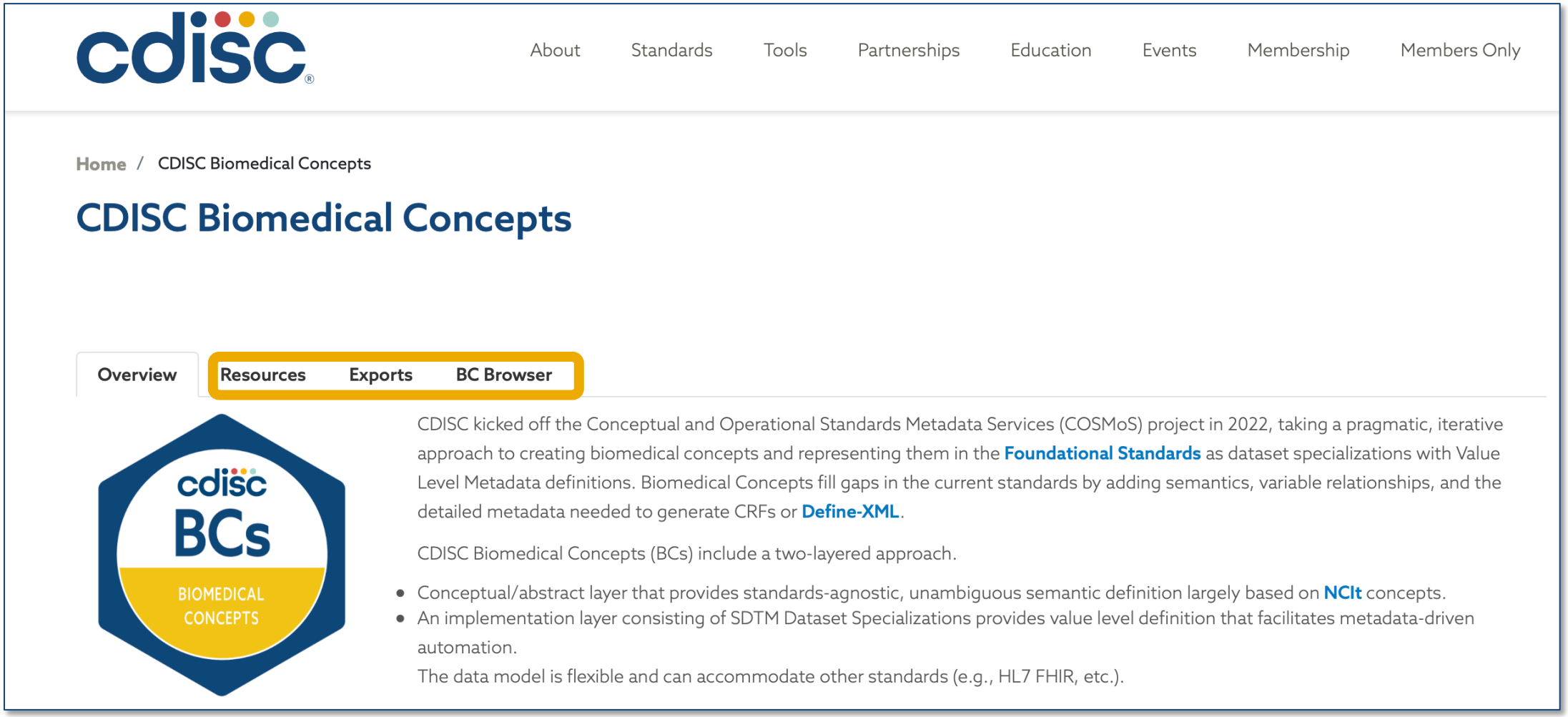
Lindus Health

Therapeutic Areas

Therapeutic Area User Guides (TAUGs) extend the Foundational Standards to represent data that pertains to specific disease areas. TAUGs include disease-specific metadata, examples and guidance on implementing CDISC standards for a variety of uses, including global regulatory submissions.

- Acute Kidney Injury
- Alzheimer's
- Asthma
- Breast Cancer
- Cardiovascular
- CDAD
- Colorectal Cancer
- COPD
- COVID-19
- Crohn's Disease
- Diabetes
- Diabetes Type 1 - Exercise and Nutrition
- Diabetes Type 1 - Pediatrics and Devices
- Diabetes Type 1 - Screening, Staging and Monitoring of Pre-clinical Type 1 Diabetes
- Diabetic Kidney Disease
- Duchenne Muscular Dystrophy
- Dyslipidemia
- Ebola
- Heart Failure
- Hepatitis C
- HIV
- Huntington's Disease
- Influenza
- Kidney Transplant
- Lung Cancer
- Major Depressive Disorder
- Malaria
- Multiple Sclerosis
- Nutrition
- Pain
- Pancreatic Cancer
- Parkinson's Disease
- Pediatrics
- Polycystic Kidney Disease
- Post Traumatic Stress Disorder
- Prostate Cancer
- Psoriasis
- QT Studies
- Rare Diseases
- Rheumatoid Arthritis
- Schizophrenia
- Traditional Chinese Medicine - Acupuncture
- Traditional Chinese Medicine - Coronary
- Artery Disease-Angina
- Traumatic Brain Injury
- Tuberculosis
- Vaccines
- Virology

Resources for BCs



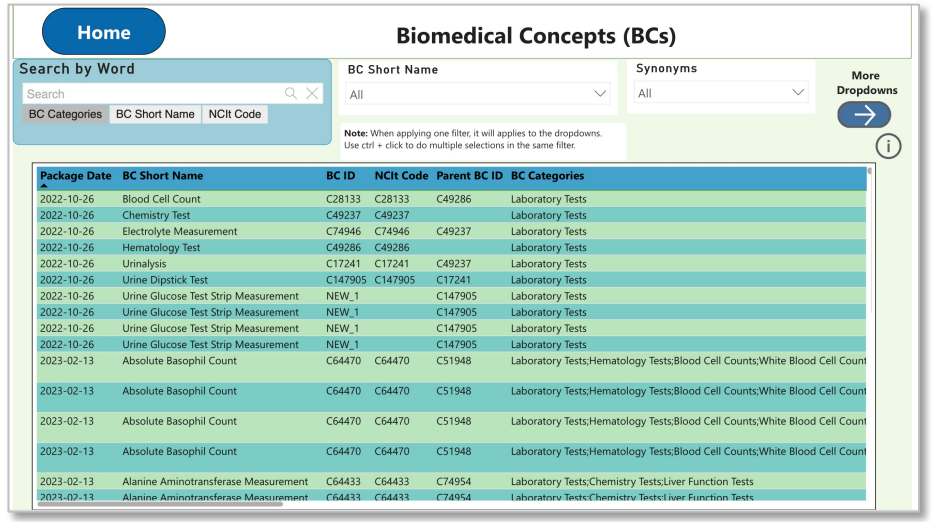
https://www.cdisc.org/cdisc-biomedical-concepts

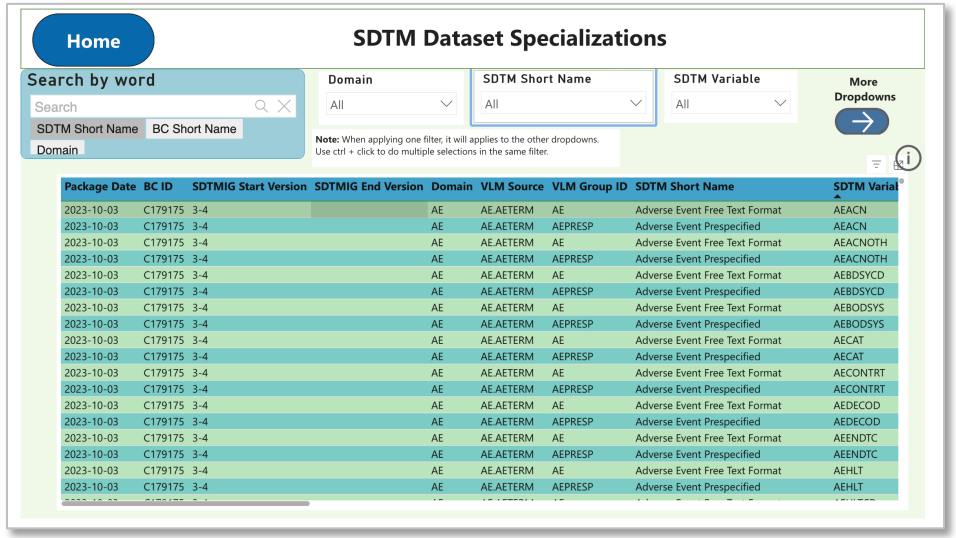
BC Browser

Overview | Resources | Exports | BC Browser

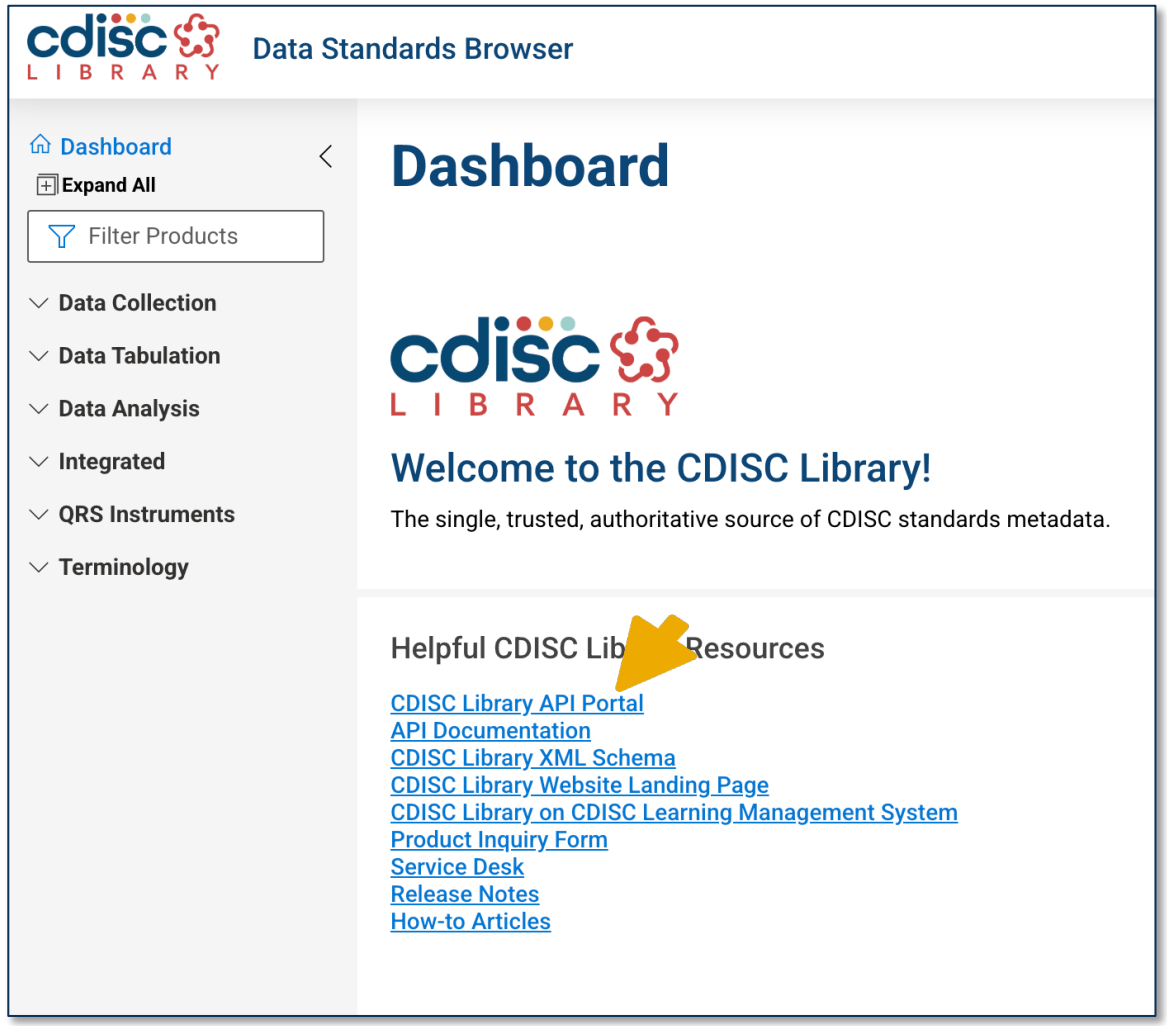
The BC browser enables users to quickly and easily search, filter and find the latest BCs and SDTM Dataset Specializations in the CDISC Library.

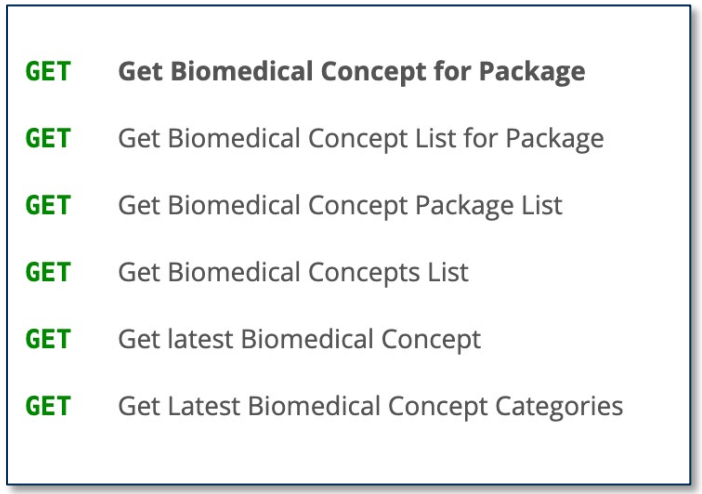
Click Here to Access the BC Browser

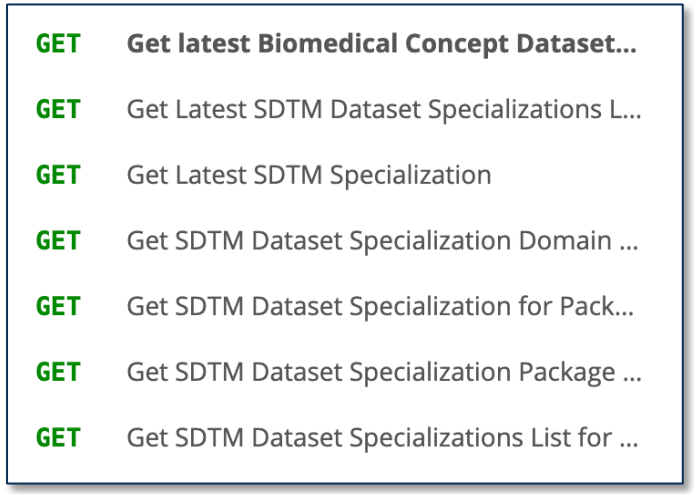




CDISC Library API







https://api.developer.library.cdisc.org

CDISC Biomedical Concepts Curation Team

- Linda Lander
- Jon Neville
- Diane Wold
- Dagmar Kottig-Roth
- Anthony Chow
- Bess LeRoy
- Sireesha Potula
- Lex Jansen
- Kirsten Langendorf
- Julie Chasen
- Katarzyna Konarzewska
- Rebecca Baker
- Omar Garcia Calderon

Thank you!

bleroy@cdisc.org