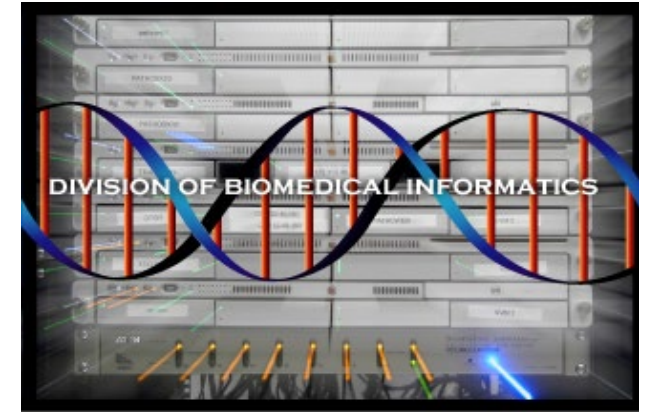


How Biomedical Ontologies Overlap with CDEs and Contribute to Harmonization

Richard H. Scheuermann, PhD
Scientific Director
National Library of Medicine

March 6, 2024

Biomedical Database Projects



Influenza Research Database



UTCRIIS



ViPR
Virus Pathogen Resource

What is an ontology?

- “An ontology is a formal, explicit specification of a shared conceptualization that is characterized by high semantic expressiveness required for increased complexity.” (Feilmayr and Wöß, 2016)
- An ontology is a *description* (like a formal specification of a program) of the concepts and *relationships* that can formally exist for an agent or a community of agents.
- To specify a conceptualization, one needs to state axioms that constrain the possible interpretations for the defined terms.

Why use ontologies for CDEs?

- Principled development by a community of informatics professionals
- Promote interoperability (FAIR)
- Promote reusability (FAIR)
- Provide an underlying semantic architecture
- Support inferencing through semantic knowledge embedded in the ontology



OBO Foundry
<http://obofoundry.org>



OBO Foundry About Principles Ontologies Citation Participate Newsletter FAQ Search C Search

Open Biological and Biomedical Ontology Foundry

Community development of interoperable ontologies for the biological sciences

Learn about OBO best practices and community resources

- OBO Foundry principles
- OBO tutorial
- Ontology browsers, tutorials, and tools

Participate

- Code of Conduct
- Join the OBO mailing list and the OBO Community Slack workspace
- OBO Foundry Operations and Working Groups
- Submit bug reports or suggestions for improvement via GitHub
- Submit your ontology to be considered for inclusion in the OBO Foundry






















OBO Library: find, use, and contribute to community ontologies

Download table as: [[YAML](#) | [JSON-LD](#) | [RDF/Turtle](#)]

Search Table

Search table ...

Ontology Domains: Group By Domain Hide Inactive Hide Obsolete

Upper	ID ^	Title ^	Description	Quick Access	Re-Use ^	Social
	bfo	Basic Formal Ontology	The upper level ontology upon which OBO Foundry ontologies are built.	     	 Stars 243	
	cob	Core Ontology for Biology and Biomedicine	COB brings together key terms from a wide range of OBO projects to improve interoperability.	     	 Stars 32	
	ro	Relation Ontology	Relationship types shared across multiple ontologies	     	 Stars 85	



Principles: Overview

- Overview
- Open (principle 1)
- Common Format (principle 2)
- URI/Identifier Space (principle 3)
- Versioning (principle 4)
- Scope (principle 5)
- Textual Definitions (principle 6)
- Relations (principle 7)
- Documentation (principle 8)
- Documented Plurality of Users (principle 9)
- Commitment To Collaboration (principle 10)
- Locus of Authority (principle 11)
- Naming Conventions (principle 12)
- Notification of Changes (principle 13)
- Maintenance (principle 16)
- Responsiveness (principle 20)

[View](#) [Edit](#)

This page is generated via [_layouts/principle.html](#). See [edit guide](#)

These principles are intended as normative for OBO Foundry ontologies, and ontologies submitted for review will be evaluated according to them. We consider these to be generally good practice, and recommend they be considered even if there are no plans to submit an ontology for review by the Foundry. Where we use capitalized words such as "MUST", and "SHOULD", they will be interpreted according to [RFC 2119: Key words for use in RFCs to Indicate Requirement Levels](#) when the principles are applied during reviews of ontologies for inclusion in the Foundry.

There is currently an ongoing process to clarify the wording of the principles and expand on their purpose, implementation, and criteria to be used to evaluate ontologies for compliance with each principle. Please use the [issue tracker](#) to let us know if there are further clarifications that you would like to see addressed for any of the principles.

Quick Summary

The following summarizes each principle. See individual pages for details.

P1) Open - The ontology **MUST** be openly available to be used by all without any constraint other than (a) its origin must be acknowledged and (b) it is not to be altered and subsequently redistributed in altered form under the original name or with the same identifiers.

P2) Common Format - The ontology is made available in a common formal language in an accepted concrete syntax.

P3) URI/Identifier Space - Each ontology **MUST** have a unique IRI in the form of an OBO Foundry permanent URL (PURL).

P4) Versioning - The ontology provider has documented procedures for versioning the ontology, and different versions of ontology are marked, stored, and officially released.

P5) Scope - The scope of an ontology is the extent of the domain or subject matter it intends to cover. The ontology must have a clearly specified scope and content that adheres to that scope.

P6) Textual Definitions - The ontology has textual definitions for the majority of its classes and for top level terms in particular.

P7) Relations - Relations should be reused from the Relations Ontology (RO).

P8) Documentation - The owners of the ontology should strive to provide as much documentation as possible.

P9) Documented Plurality of Users - The ontology developers should document that the ontology is used by multiple independent people or organizations.

P10) Commitment To Collaboration - OBO Foundry ontology development, in common with many other standards-oriented scientific activities, should be carried out in a collaborative fashion.

P11) Locus of Authority - There should be a person who is responsible for communications between the community and the ontology developers, for communicating with the Foundry on all Foundry-related matters, for mediating discussions involving maintenance in the light of scientific advance, and for ensuring that all user feedback is addressed.

P12) Naming Conventions - The names (primary labels) for elements (classes, properties, etc.) in an ontology must be intelligible to scientists and amenable to natural language processing. Primary labels should be unique among OBO Library ontologies.

P16) Maintenance - The ontology needs to reflect changes in scientific consensus to remain accurate over time.

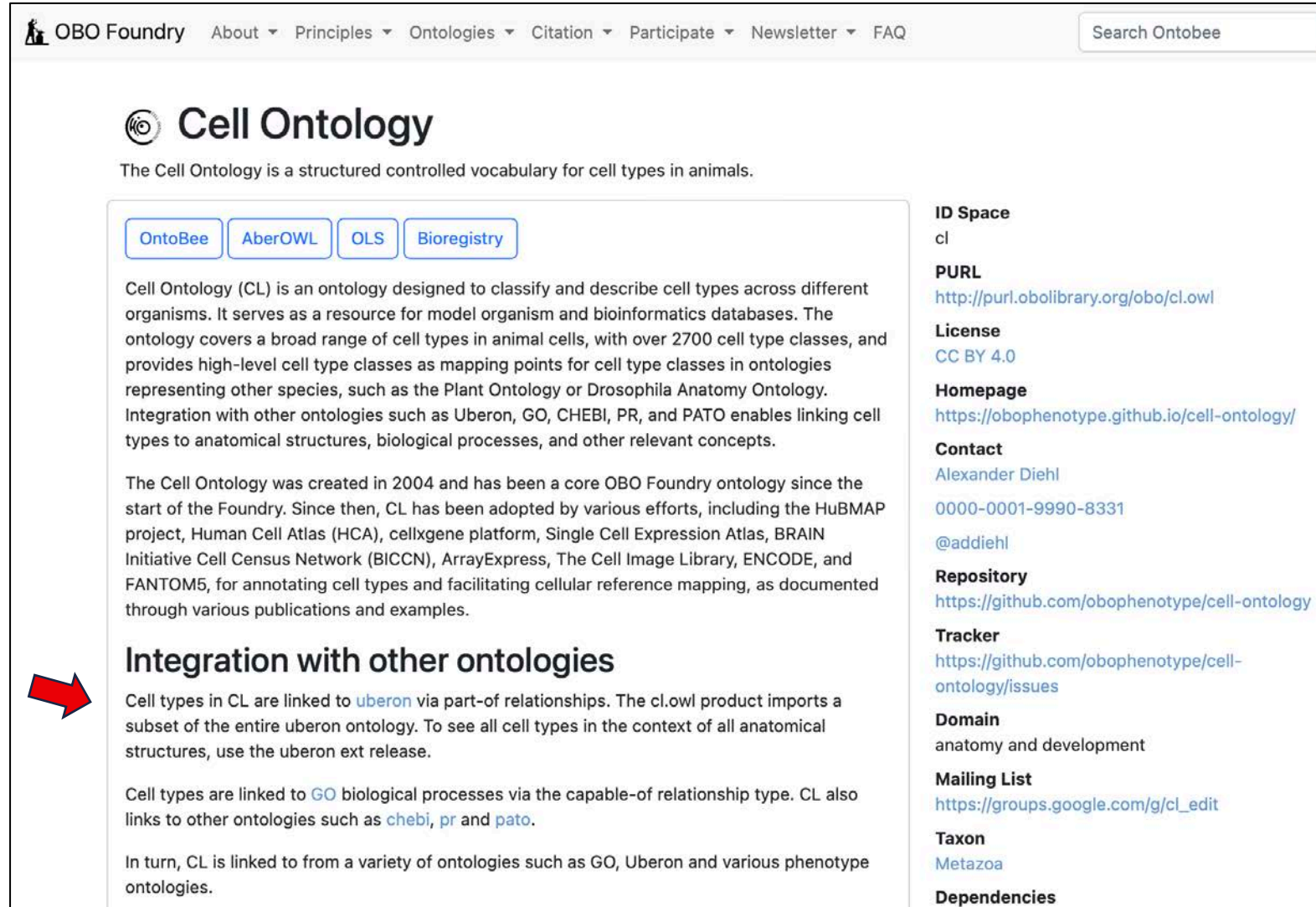
P20) Responsiveness - Ontology developers **MUST** offer channels for community participation and **SHOULD** be responsive to requests.

Smith B, et al. "The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration", Nat Biotechnol. (2007) 25:1251-5. PMID: 17989687.

Why use ontologies for CDEs?

- Principled development by a community of informatics professionals
- **Promote interoperability (FAIR)**
- Promote reusability (FAIR)
- Provide an underlying semantic architecture
- Support inferencing through semantic knowledge embedded in the ontology

Interoperable



OBO Foundry About ▾ Principles ▾ Ontologies ▾ Citation ▾ Participate ▾ Newsletter ▾ FAQ

Cell Ontology

The Cell Ontology is a structured controlled vocabulary for cell types in animals.

[OntoBee](#) [AberOWL](#) [OLS](#) [Bioregistry](#)

Cell Ontology (CL) is an ontology designed to classify and describe cell types across different organisms. It serves as a resource for model organism and bioinformatics databases. The ontology covers a broad range of cell types in animal cells, with over 2700 cell type classes, and provides high-level cell type classes as mapping points for cell type classes in ontologies representing other species, such as the Plant Ontology or Drosophila Anatomy Ontology. Integration with other ontologies such as Uberon, GO, CHEBI, PR, and PATO enables linking cell types to anatomical structures, biological processes, and other relevant concepts.

The Cell Ontology was created in 2004 and has been a core OBO Foundry ontology since the start of the Foundry. Since then, CL has been adopted by various efforts, including the HuBMAP project, Human Cell Atlas (HCA), cellxgene platform, Single Cell Expression Atlas, BRAIN Initiative Cell Census Network (BICCN), ArrayExpress, The Cell Image Library, ENCODE, and FANTOM5, for annotating cell types and facilitating cellular reference mapping, as documented through various publications and examples.

Integration with other ontologies

Cell types in CL are linked to [uberon](#) via part-of relationships. The `cl.owl` product imports a subset of the entire uberon ontology. To see all cell types in the context of all anatomical structures, use the uberon ext release.

Cell types are linked to [GO](#) biological processes via the capable-of relationship type. CL also links to other ontologies such as [chebi](#), [pr](#) and [pato](#).

In turn, CL is linked to from a variety of ontologies such as GO, Uberon and various phenotype ontologies.

ID Space
cl

PURL
<http://purl.obolibrary.org/obo/cl.owl>

License
CC BY 4.0

Homepage
<https://obophenotype.github.io/cell-ontology/>

Contact
Alexander Diehl
0000-0001-9990-8331
@addiehl

Repository
<https://github.com/obophenotype/cell-ontology>

Tracker
<https://github.com/obophenotype/cell-ontology/issues>

Domain
anatomy and development

Mailing List
https://groups.google.com/g/cl_edit

Taxon
[Metazoa](#)

Dependencies

Why use ontologies for CDEs?

- Principled development by a community of informatics professionals
- Promote interoperability (FAIR)
- **Promote reusability (FAIR)**
- Provide an underlying semantic architecture
- Support inferencing through semantic knowledge embedded in the ontology

Reusable

Usages

User
<https://biccn.org/>

Description
The BICCN created a high-resolution atlas of cell types in the primary motor based on single cell transcriptomics. These cell types are represented in the brain data standards ontology which anchors to cell types in the cell ontology.

Type
annotation

Examples
[cell type card of a cell type linked to a PCL cell type \(L2/3 IT primary motor cortex glutamatergic neuron\) which is a subclass of cell types in CL \(CL:4023041\)](#)
[PCL cell type used in cell type cards linked directly to CL cell types](#)

Publications

- [Brain Data Standards Ontology: A data-driven ontology of transcriptomically defined cell types in the primary motor cortex](#)

User
<https://hubmapconsortium.org/>

Description
HuBMAP develops tools to create an open, global atlas of the human body at the cellular level. The Cell Ontology is used in annotating cell types in the tools developed.

Type
annotation

Examples
[ASCT+B reporter showing CL being used to annotate cell types in the heart](#)

Publications

- [The human body at cellular resolution: the NIH Human Biomolecular Atlas Program.](#)

User
<https://www.humancellatlas.org/>

Description
The Human Cell Atlas (HCA) is an international group of researchers using a combination of these new technologies to create cellular reference maps. The HCA use CL to annotate cells in their reference maps.

Type
annotation

Examples
[HCA collection studies that are related B cell \(CL:0000236\) that is filtered through CL annotation](#)

Publications

- [The Human Cell Atlas](#)

User
<https://www.ebi.ac.uk/gxa/home>

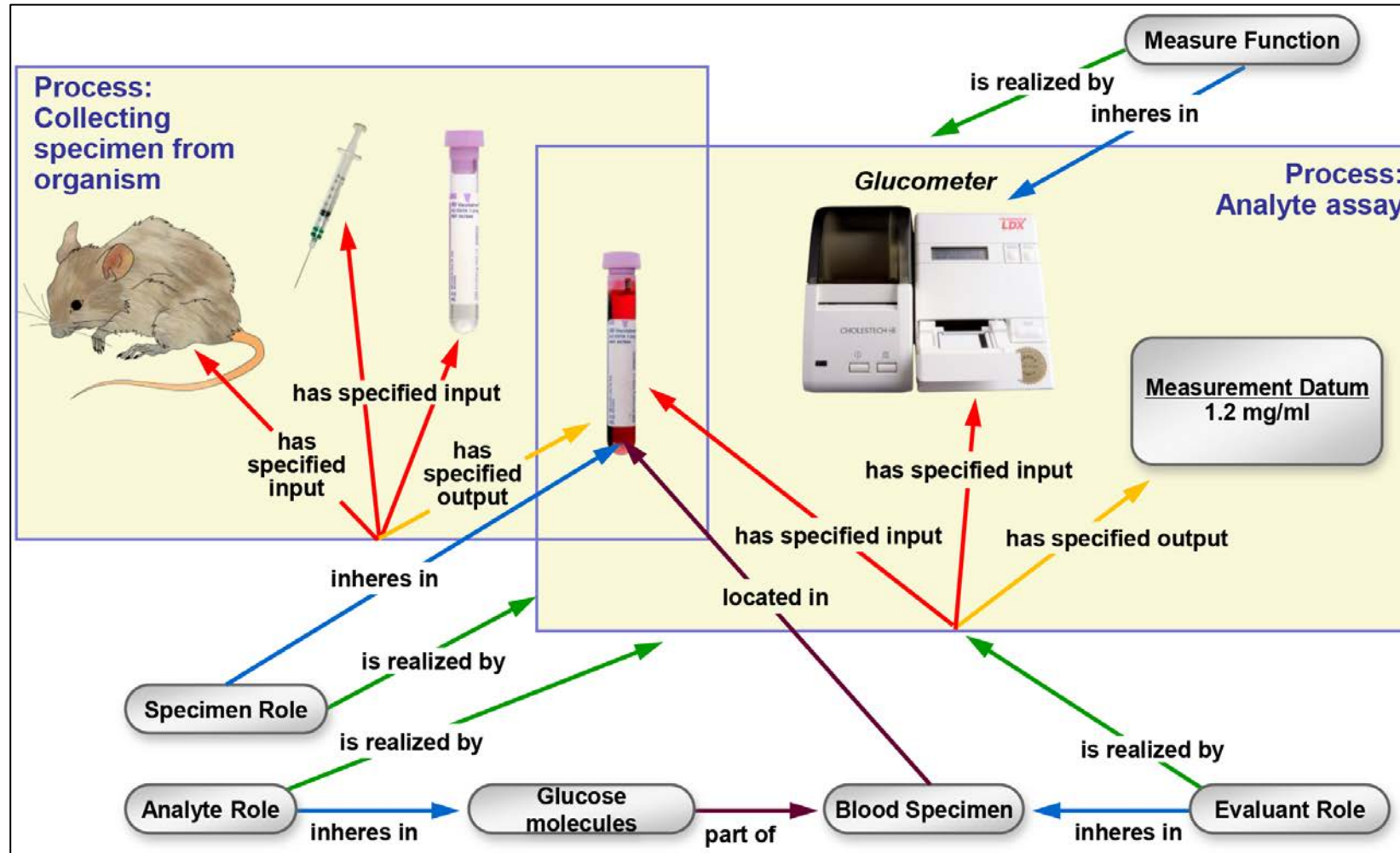
Description
The EBI single cell expression atlas is an extension to EBI expression atlas that displays gene expression in single cells. Cell types in the single cell expression atlas linked with terms from the Cell Ontology.

Type
annotation

Why use ontologies for CDEs?

- Principled development by a community of informatics professionals
- Promote interoperability (FAIR)
- Promote reusability (FAIR)
- **Provide an underlying semantic architecture**
- Support inferencing through semantic knowledge embedded in the ontology

Ontology for Biomedical Investigations



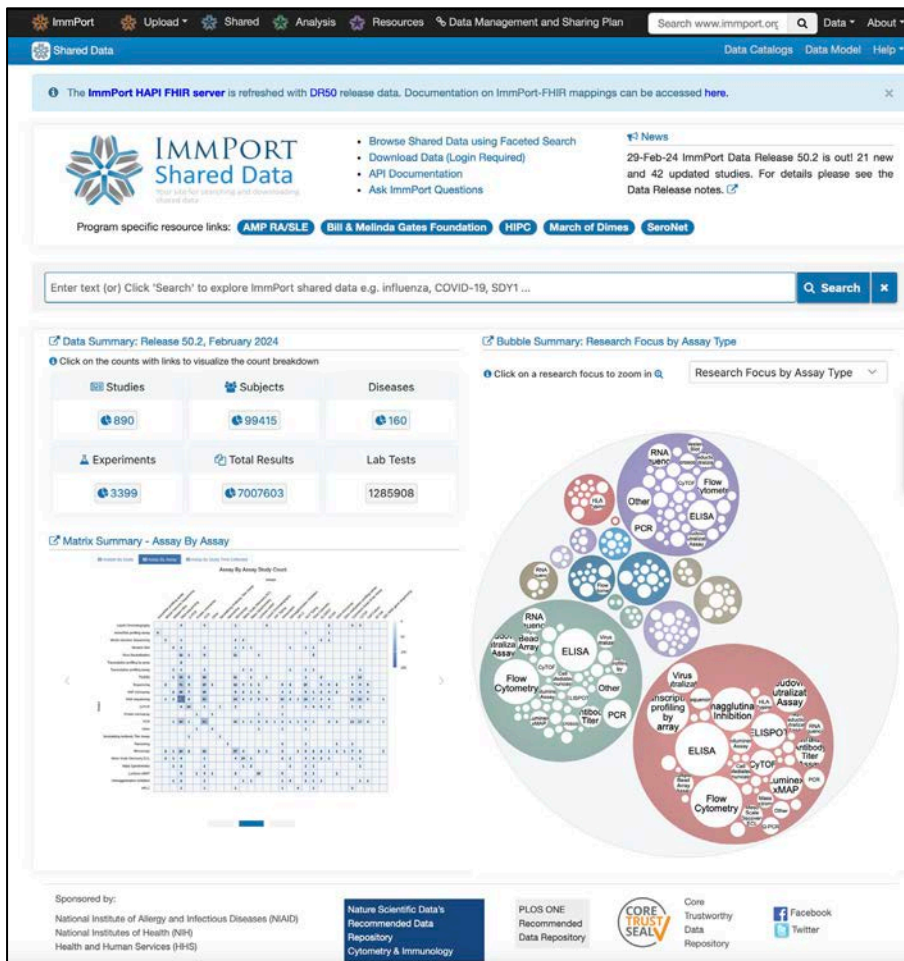
Logical Observation Identifiers Names and Codes (LOINC)

CLASS	LOINC code	Rank Order (1 = most freq ordered)	LOINC long common name	LOINC short name	COMPONENT	PROPERTY	TM	SPCMN	SCALE	METHOD	Example Units (UCUM)
SERO	11572-5	115	Rheumatoid factor [Units/volume] in Serum	Rheumatoid fact Ser-aCnc	Rheumatoid factor	ACnc	Pt	Ser	Qn		[IU]/mL
SERO	47383-5	199	Nuclear Ab [Presence] in Serum by Immunoassay	ANA Ser QI EIA	Nuclear Ab	ACnc	Pt	Ser	Ord	EIA	
SERO	5047-6	200	Nuclear Ab [Units/volume] in Serum by Immunoassay	ANA Ser EIA-aCnc	Nuclear Ab	ACnc	Pt	Ser	Qn	EIA	[arb'U]
SERO	55727-2	205	Tissue transglutaminase Ab.IgA & IgG panel [Presence] in Serum	tTG IgA+IgG Pnl Ser	Tissue transglutaminase Ab.IgA & IgG panel	ACnc	Pt	Ser	Ord		
SERO	31017-7	216	Tissue transglutaminase IgA Ab [Units/volume] in Serum	tTG IgA Ser-aCnc	Tissue transglutaminase Ab.IgA	ACnc	Pt	Ser	Qn		[arb'U]/mL
SERO	5130-0	230	DNA double strand Ab [Units/volume] in Serum	dsDNA Ab Ser-aCnc	DNA double strand Ab	ACnc	Pt	Ser	Qn		[IU]/mL
SERO	17791-5	241	Sjogrens syndrome-B extractable nuclear Ab [Units/volume] in Serum	ENA SS-B Ab Ser-aCnc	Sjogrens syndrome-B extractable nuclear Ab	ACnc	Pt	Ser	Qn		[arb'U]/mL
SERO	17792-3	242	Sjogrens syndrome-A extractable nuclear Ab [Units/volume] in Serum	ENA SS-A Ab Ser-aCnc	Sjogrens syndrome-A extractable nuclear Ab	ACnc	Pt	Ser	Qn		[arb'U]/mL
SERO	33935-8	305	Cyclic citrullinated peptide IgG Ab [Units/volume] in Serum	cCP IgG Ser-aCnc	Cyclic citrullinated peptide Ab.IgG	ACnc	Pt	Ser	Qn		[arb'U]/mL



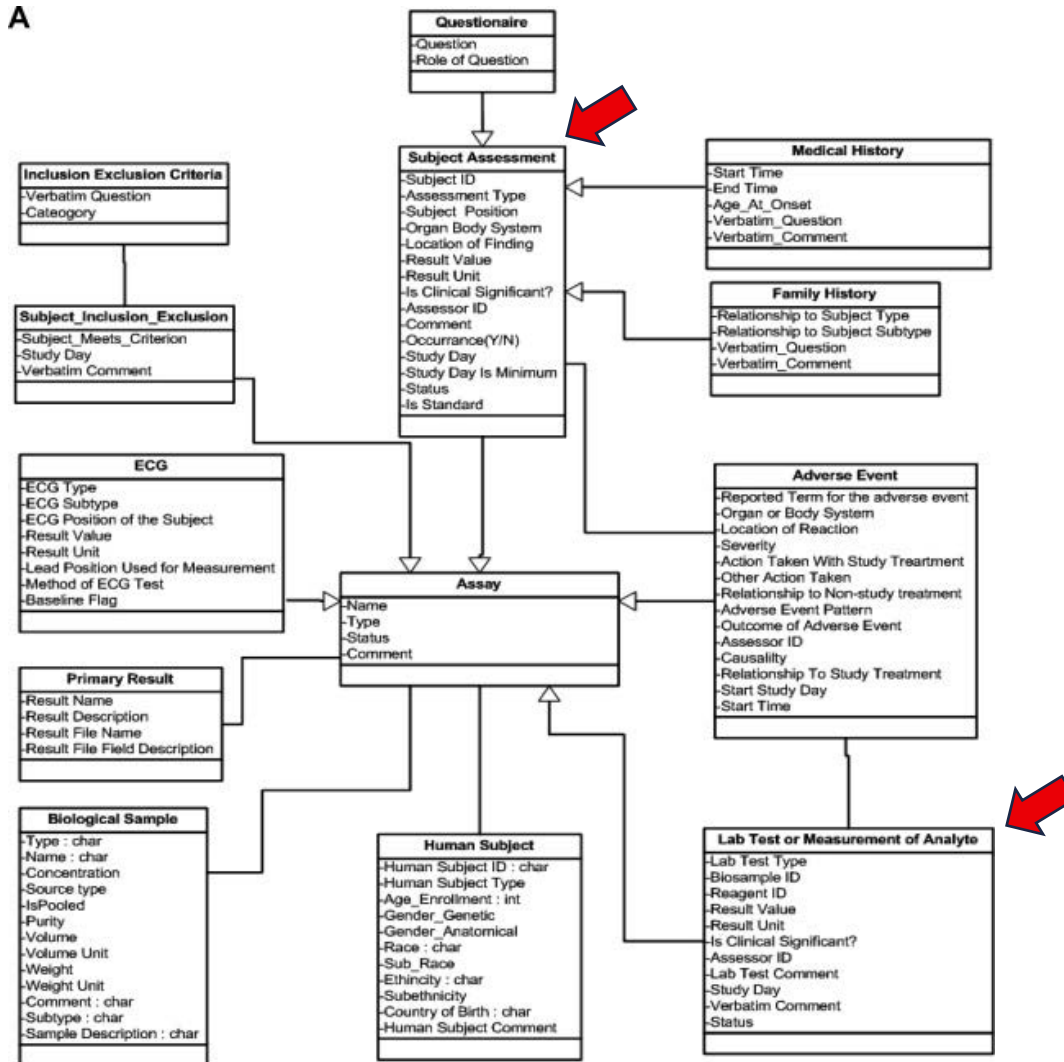
ImmPort

Ontology-based extensible data model



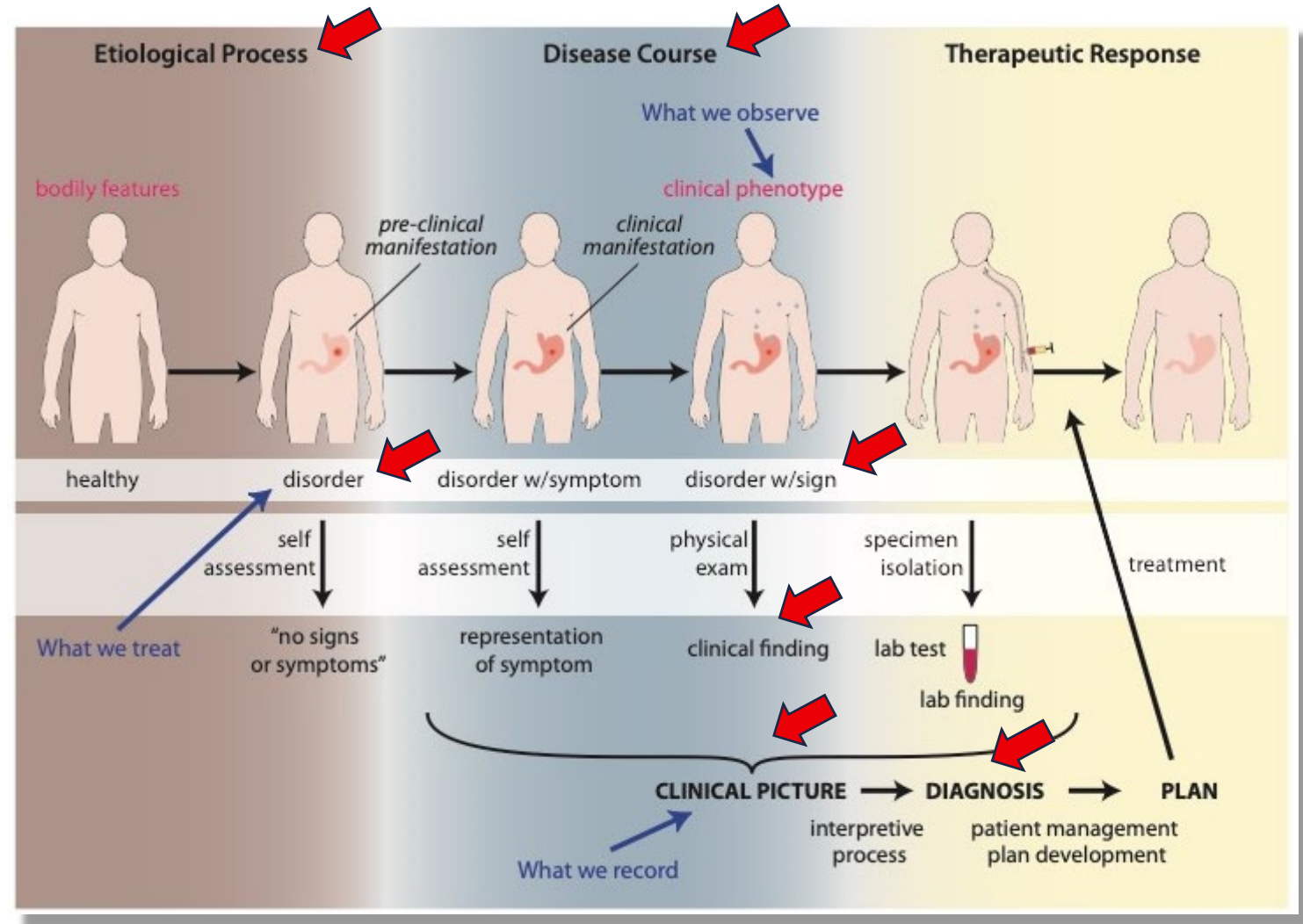
<https://www.immport.org/home>

A



Ontology for General Medical Science

- Disorder
- Pathological Process
- Disease
- Etiological Process
- Disease Course
- Sign
- Symptom
- Clinical History
- Physical Examination
- Laboratory Test
- Clinical Picture
- Diagnosis



Why use ontologies for CDEs?

- Principled development by a community of informatics professionals
- Promote interoperability (FAIR)
- Promote reusability (FAIR)
- Provide an underlying semantic architecture
- Support inferencing through semantic knowledge embedded in the ontology

NIH-Endorsed Common Data Elements (CDE)

The screenshot shows the NIH CDE Repository search results page. The left sidebar contains filters for 'ACTIVE CDE FILTERS' (NIH-Endorsed CDEs, Project 5 (COVID-19)), 'REFINE CDE RESULTS' (NIH-Endorsed), 'Collections' (Project 5 (COVID-19), Tier 1 (125)), 'Registration Status' (Qualified (125)), and 'Data Types' (Value List (74), Text (29), Number (6), Date (16)). A red arrow points to the 'NIH-Endorsed CDEs' filter. The main content area displays 'CDE Search Results' for 'Vital Signs Type' and 'Vital Signs Unit Of Measure'. Red arrows point to the 'Qualified' status and the 'Label' column in both tables.

ACTIVE CDE FILTERS

- NIH-Endorsed CDEs x
- Project 5 (COVID-19) x
- Clear all x

REFINE CDE RESULTS

- NIH-Endorsed

Collections ^

- Project 5 (COVID-19)**
- Tier 1 (125) >

Registration Status ^

- Qualified (125)

Data Types ^

- Value List (74)
- Text (29)
- Number (6)
- Date (16)

CDE Search Results

Vital Signs Type

A textual description of a person's vital signs measurement category.

Qualified

Steward: Project 5 (COVID-19)
Used By: Project 5 (COVID-19)

Label	Code	ConceptID
Systolic blood pre...		C25298
Diastolic blood pr...		C25299
Heart rate		C49677
Respiratory rate		C49678

(8 total) See full table in [Detail View](#)

Vital Signs Unit Of Measure

The unit of measure for the person's vital sign measurement.

Qualified

Steward: Project 5 (COVID-19)
Used By: Project 5 (COVID-19)

Label	Code	ConceptID
Percentage		C25613
Beats per Minute		C49673
Breaths per Minute		C49674
Degree Celsius		C42559

(11 total) See full table in [Detail View](#)

Systemized Nomenclature of Medicine (SNOMED-CT) example

OLS ONTOLOGY SEARCH

Home | Ontologies | Help | About | Downloads

Ontologies > SNOMED > Classes > SNOMED:75367002 Copy

en JSON

Blood pressure (observable entity)

<http://snomed.info/id/75367002> Copy

Search SNOMED...

Exact match Include obsolete terms Include imported terms

Tree Graph

- SNOMED CT Concept (SNOMED RT+CTV3) (352,448)
 - Observable entity (observable entity) (9,143)
 - Clinical history/examination observable (observable entity) (5,689)
 - Cardiovascular observable (observable entity) (535)
 - Cardiovascular measure (observable entity) (212)
 - Blood pressure (observable entity) (115)**
 - Cardiac measure (observable entity) (17)
 - Circulatory arrest time (observable entity)
 - Heart rate (observable entity) (7)
 - Heart rate on admission (observable entity)
 - Intima media thickness (observable entity) (1)
 - Invasive blood pressure (observable entity) (4)
 - Left ventricular end-diastolic pressure (observable entity)
 - Maximum blood pressure (observable entity) (8)
 - Mean blood pressure (observable entity) (7)
 - Minimum blood pressure (observable entity) (8)
 - Non-invasive blood pressure (observable entity) (4)
 - Right ventricular end-diastolic pressure (observable entity)
 - Right ventricular peak systolic pressure (observable entity)
 - Segmental pressure (blood pressure) (observable entity)
 - Speed of cardiovascular response (observable entity) (4)
 - Vascular measure (observable entity) (96)

Show counts Show obsolete terms Show all siblings

Class Information

alternative label
BP - Blood pressure

has_dbxref
CTV3:XUKKF

preferred label
Blood pressure

Class Relations

Equivalent to

- (Cardiovascular measure (observable entity) and Role group (attribute) *some* Property (attribute) *some* Pressure (property) (qualifier value) and Role group (attribute) *some* Scale type (attribute) *some* Quantitative (qualifier value) and Role group (attribute) *some* Inheres in (attribute) *some* Structure of cardiovascular system (body structure) and Role group (attribute) *some* Characterizes (attribute) *some* Cardiac process (qualifier value))

Subclass of
Cardiovascular measure (observable entity)

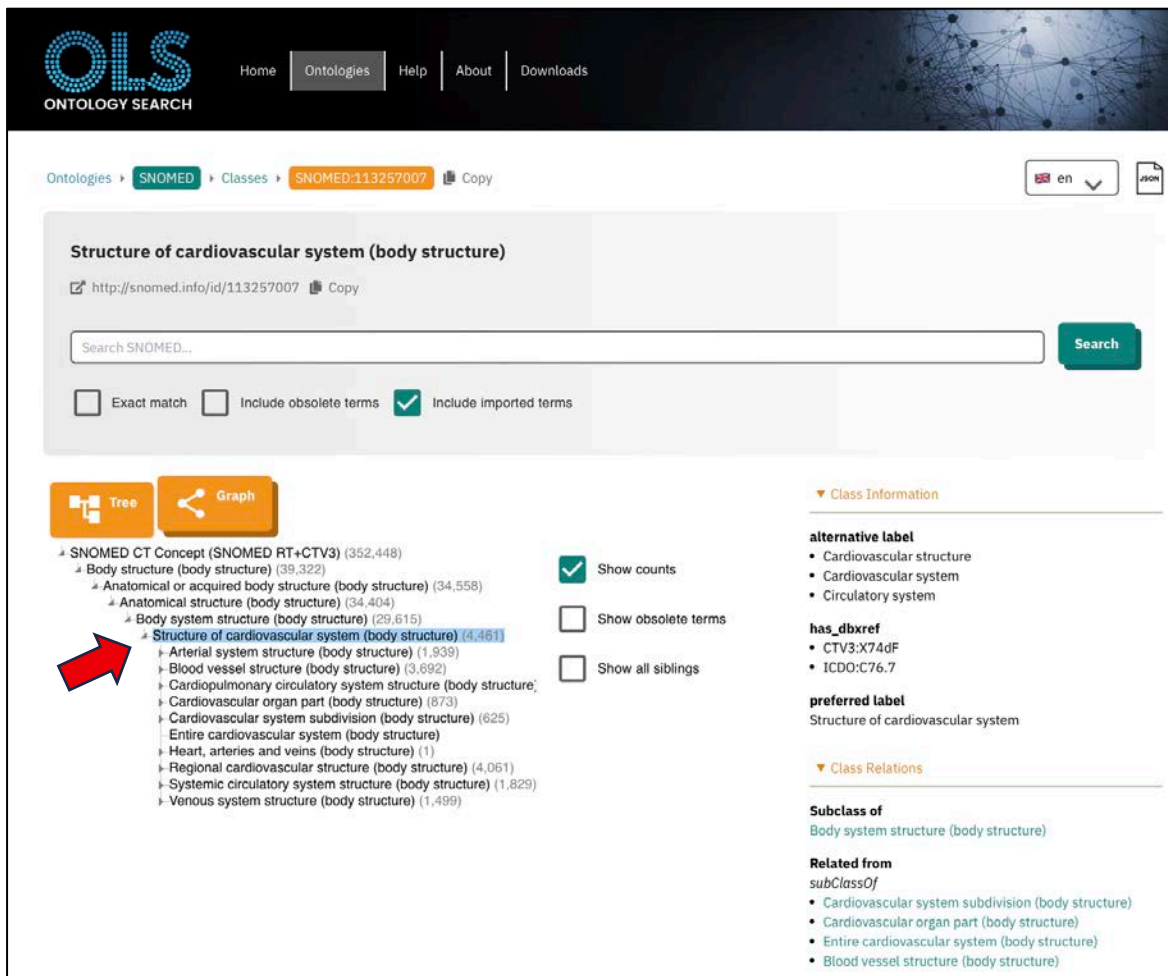
Equivalent to

- (Cardiovascular measure (observable entity) and Role group (attribute) *some* Property (attribute) *some* Pressure (property) (qualifier value) and Role group (attribute) *some* Scale type (attribute) *some* Quantitative (qualifier value) and Role group (attribute) *some* Inheres in (attribute) *some* Structure of cardiovascular system (body structure) and Role group (attribute) *some* Characterizes (attribute) *some* Cardiac process (qualifier value))

Subclass of

Cardiovascular measure (observable entity)

SNOMED-CT example



OLS ONTOLOGY SEARCH

Home | Ontologies | Help | About | Downloads

Ontologies > SNOMED > Classes > SNOMED:113257007 Copy

Structure of cardiovascular system (body structure)

http://snomed.info/id/113257007 Copy

Search SNOMED...

Exact match Include obsolete terms Include imported terms

Tree Graph

- SNOMED CT Concept (SNOMED RT+CTV3) (352,448)
 - Body structure (body structure) (39,322)
 - Anatomical or acquired body structure (body structure) (34,558)
 - Anatomical structure (body structure) (34,404)
 - Body system structure (body structure) (29,615)
 - Structure of cardiovascular system (body structure) (4,461)**
 - Arterial system structure (body structure) (1,939)
 - Blood vessel structure (body structure) (3,692)
 - Cardiopulmonary circulatory system structure (body structure) (873)
 - Cardiovascular organ part (body structure) (873)
 - Cardiovascular system subdivision (body structure) (625)
 - Entire cardiovascular system (body structure) (1)
 - Heart, arteries and veins (body structure) (1)
 - Regional cardiovascular structure (body structure) (4,061)
 - Systemic circulatory system structure (body structure) (1,829)
 - Venous system structure (body structure) (1,499)

Show counts
 Show obsolete terms
 Show all siblings

Class Information

alternative label

- Cardiovascular structure
- Cardiovascular system
- Circulatory system

has_dbxref

- CTV3:X74dF
- ICDO:C76.7

preferred label

Structure of cardiovascular system

Class Relations

Subclass of

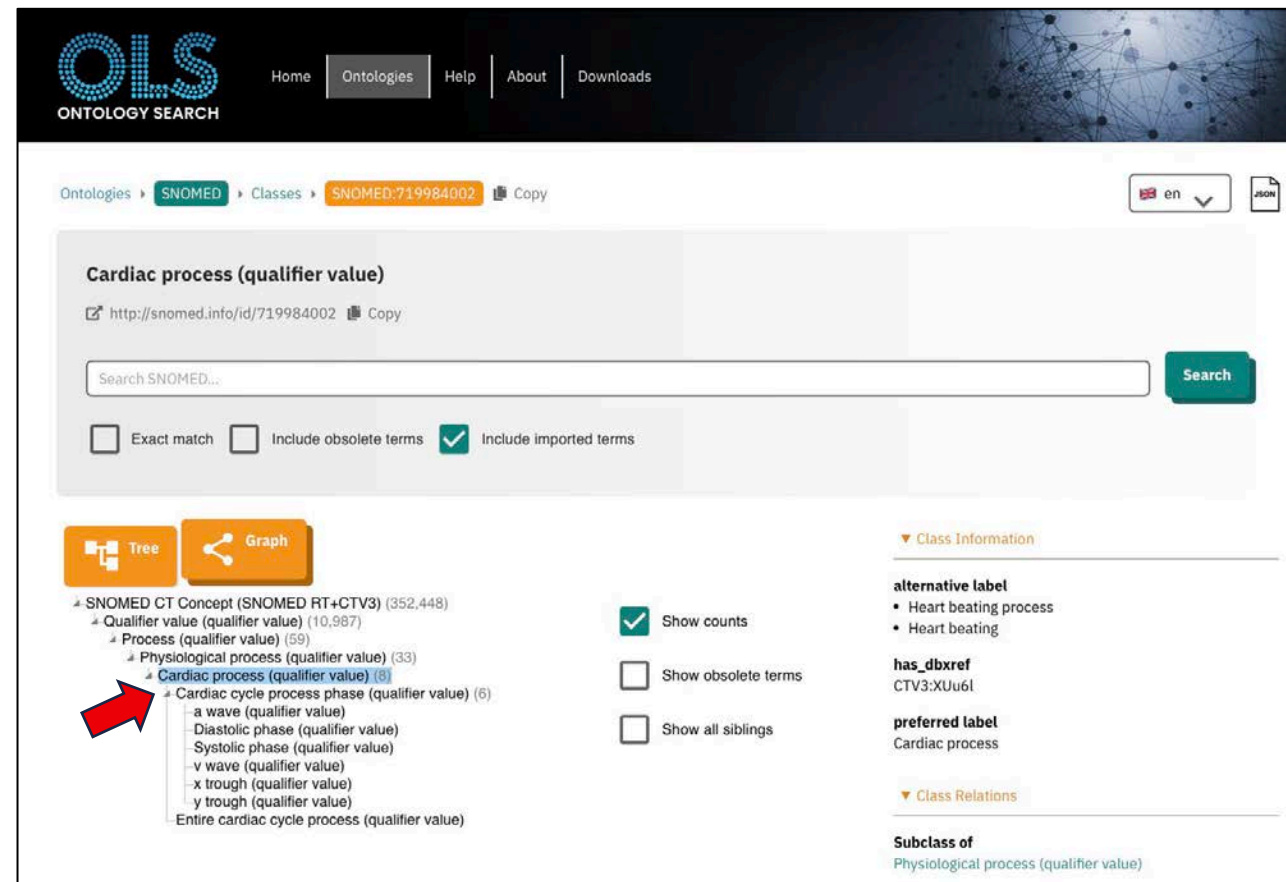
Body system structure (body structure)

Related from

subClassOf

- Cardiovascular system subdivision (body structure)
- Cardiovascular organ part (body structure)
- Entire cardiovascular system (body structure)
- Blood vessel structure (body structure)

A red arrow points to the 'Structure of cardiovascular system (body structure) (4,461)' entry in the tree view.



OLS ONTOLOGY SEARCH

Home | Ontologies | Help | About | Downloads

Ontologies > SNOMED > Classes > SNOMED:719984002 Copy

Cardiac process (qualifier value)

http://snomed.info/id/719984002 Copy

Search SNOMED...

Exact match Include obsolete terms Include imported terms

Tree Graph

- SNOMED CT Concept (SNOMED RT+CTV3) (352,448)
 - Qualifier value (qualifier value) (10,987)
 - Process (qualifier value) (59)
 - Physiological process (qualifier value) (33)
 - Cardiac process (qualifier value) (8)**
 - Cardiac cycle process phase (qualifier value) (6)
 - a wave (qualifier value)
 - Diastolic phase (qualifier value)
 - Systolic phase (qualifier value)
 - v wave (qualifier value)
 - x trough (qualifier value)
 - y trough (qualifier value)
 - Entire cardiac cycle process (qualifier value)

Show counts
 Show obsolete terms
 Show all siblings

Class Information

alternative label

- Heart beating process
- Heart beating

has_dbxref

CTV3:XUu6l

preferred label

Cardiac process

Class Relations

Subclass of

Physiological process (qualifier value)

A red arrow points to the 'Cardiac process (qualifier value) (8)' entry in the tree view.

Why use ontologies for CDEs?

- Principled development by a community of informatics professionals
- Promote interoperability (FAIR)
- Promote reusability (FAIR)
- Provide an underlying semantic architecture
- Support inferencing through semantic knowledge embedded in the ontology