

Introduction to Semantic Interoperability and the Role of Concepts

Denise Warzel, BBA, MSc

ISO SC 32/WG 2 Metadata Standards

Convenor

US National Institutes of Health – National Cancer Institute

Semantic Infrastructure, Metadata and Models Program Manager

Agenda

- Semantic Interoperability
- Interoperability across Healthcare, Healthcare Research, and Biomedical Research
- Lessons learned/challenges
- Role of Concept Systems



Semantic Interoperability

- **Defined:** Ability of computer systems to exchange data, with **clear unambiguous meaning**
- **Principles:**
 - Standards for data exchange
 - Shared semantic alignment and mapping
 - Rich metadata
 - Governance among stakeholders

Why Semantic Interoperability in Healthcare, Healthcare Research and Biomedical Research?



Scientists need data to conduct research

New cures and treatments



Lack of interoperability inhibits data integration

More Data = Increased confidence in new findings



Interoperability Inhibitors

Different Data Models
Different Coding Standards
Different Terminology

Approach → Challenges and Lessons Learned

Ontologies

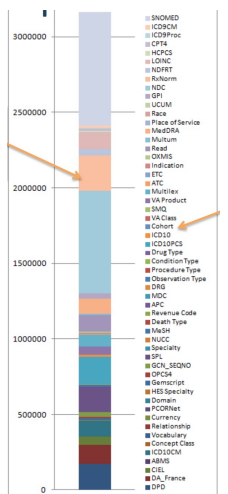
- Highly specialized
 - Too many “Standard Ontologies”
 - Biomedical Research: 1,051*
- *Stanford University: <https://biportal.bioontology.org/>
- Mappings expensive and error prone

Common Data Models

- Healthcare and Healthcare Research
 - Healthcare – e.g. 8 Major Electronic Medical Record (EMR) Vendors – 8 Proprietary Data Models
 - Healthcare Research – e.g 4 “Common Data Models
- Biomedical Research
 - Domain Specific: Genomics, Proteomics, Cancer Specific (Pediatric, Lung, Ovarian), Imaging, etc

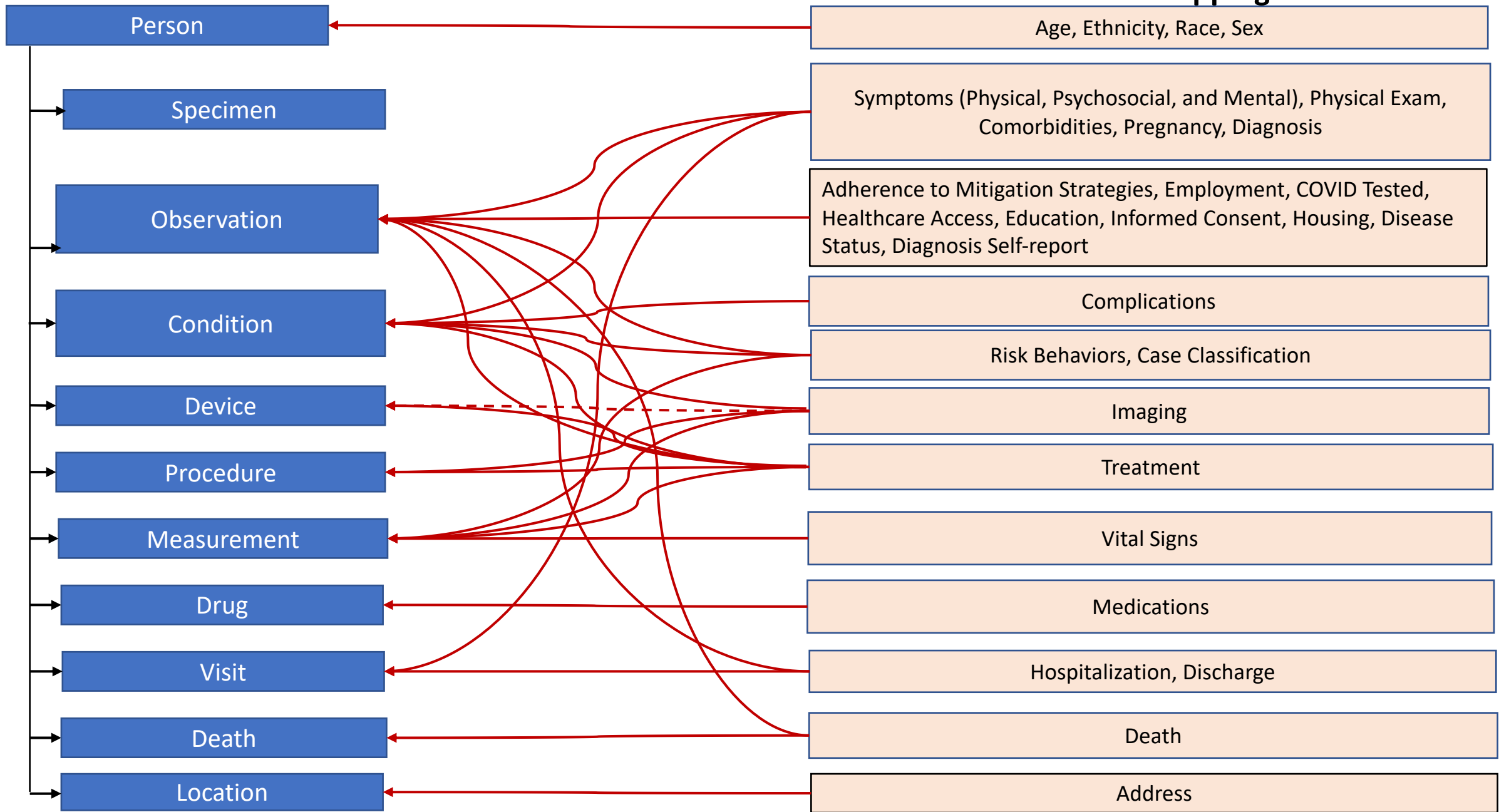
Standard Terminologies

- Healthcare and Healthcare Research
 - Hundreds of terminologies, not fully aligned
 - Multiple mappings by different groups
- Biomedical Research
 - Same problem



Healthcare Research OMOP Data Model

Biomedical Research Partial mapping of NIH COVID CDEs



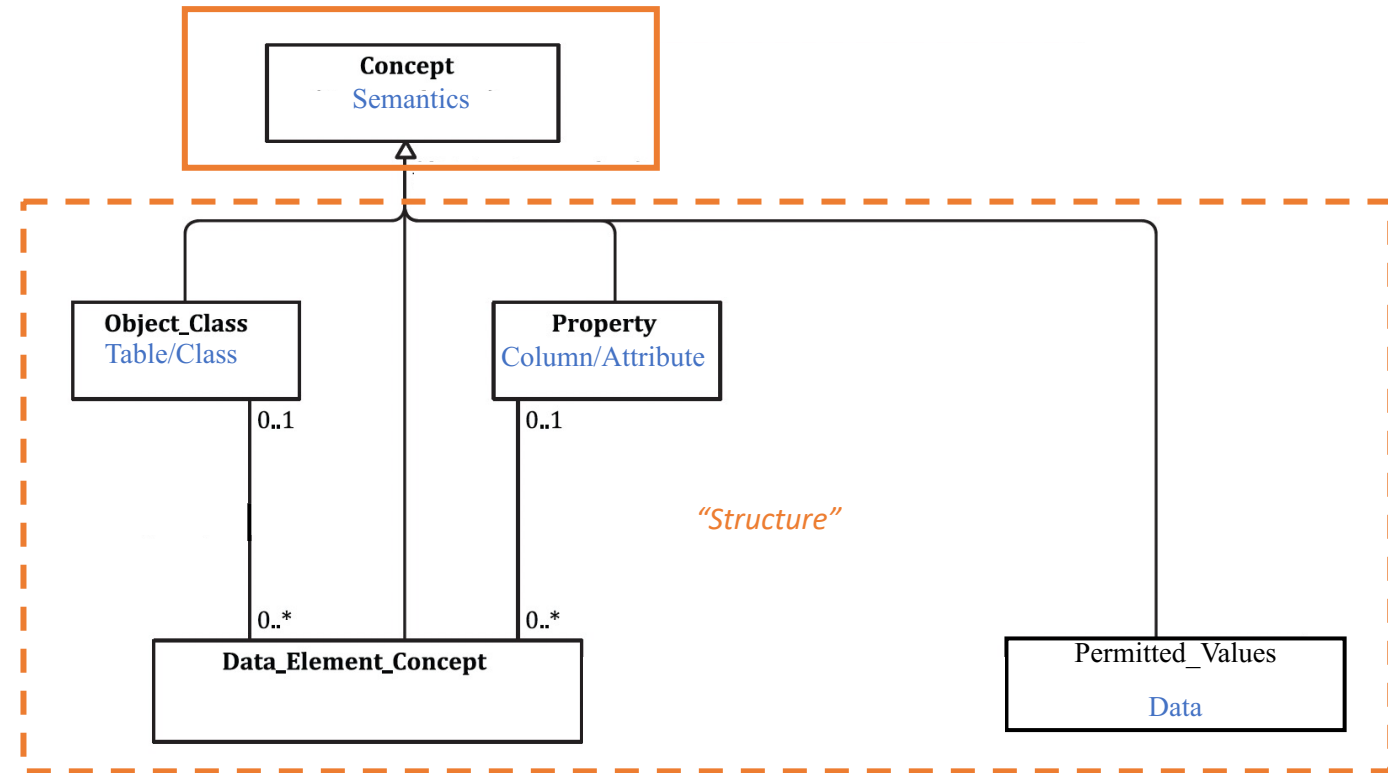
What to do?



Leverage ISO/IEC 11179 Metadata Registry Standard

Structured Metadata Supports:

- Accurate, reliable information about data
 - Unique Identifiers
- Data interchange referencing data elements
 - Data Sets, Models
- **Semantics based on Concepts**
- Rich Metadata
 - Multiple Names, Definitions, Languages



ISO/IEC 11179-31:2023 Data Specification, Based on Figure 4, Data Element Concept metamodel region

Semantic Spectrum

Concepts and Semantic Interoperability

Concepts

- Ideas, units of knowledge

Concept Systems

- Composed of terms and definitions
- Domain specific
- Human and computer readable
- Provide semantic clarity
- Enhanced understanding
 - Synonyms, roles, relationships

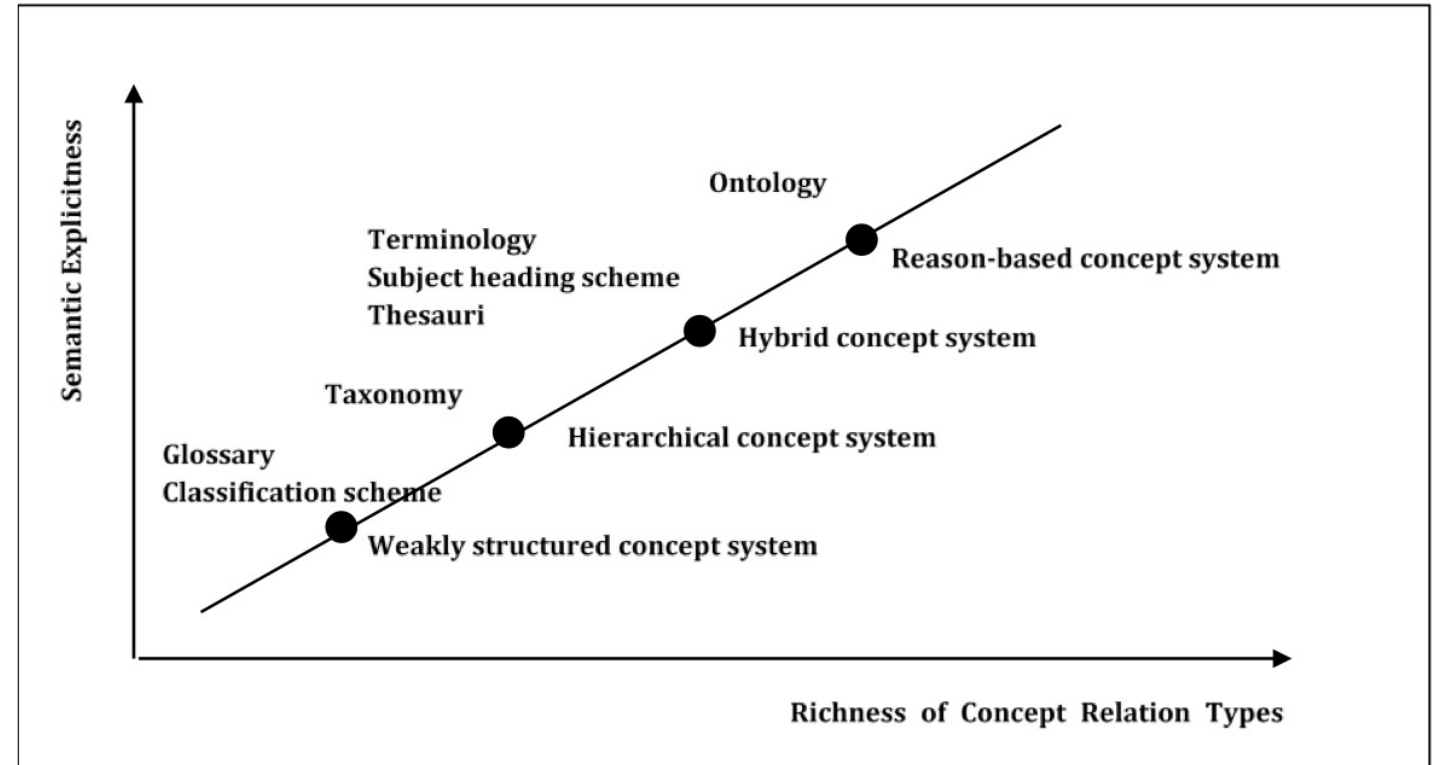


Figure 2 — Level of structure of concept systems

ISO/IEC 5394 Criteria for concept systems

Example: ISO/IEC 11179 Concept Annotations

4 Ethnicity Data Elements, 4 Data Models, 3 Domains

2nd Concepts define Semantics of the data values permitted in each Data Model

		Healthcare Research Domain		Healthcare Domain	Biomedical Research Domain
Permitted Values Aligned by Concepts		Data Element Semantics Aligned by Concepts			
Concept Code	Concept Name	PCORnet DEMOGRAPHIC HISPANIC	OMOP PERSON ethnicity_concept_id	HL7 US Core Person Ethnicity	CDISC DEMOGRAPHIC Ethnicity
C17998	Unknown	UN		UNK	UNKNOWN
C53269	No Information Available	NI		NI	
C17649	Other	OT		OTH	
C51024	Response Declined	R			
C17459	Hispanic or Latino	Y	38003563	2135-2	HISPANIC OR LATINO
C41222	Not Hispanic or Latino	N	38003564	2186-5	NOT HISPANIC OR LATINO

1st Concepts define the Semantics of the Data Elements across the Data Models

3rd Concepts enable semantic interoperability

Role of Concepts in Semantic Interoperability

- Shared meaning across disparate domains
 - Healthcare, Healthcare Research, and Biomedical Research
- Enables knowledge discovery
 - Concept Systems → Synonyms, relationships, roles
- Enables data transformation
 - Discover the multiple ways a data element such as **Person Ethnicity** was captured/stored
- Organize data for comparison
 - Use in AI/ML algorithms
 - Representation in graph databases



NCI's Approach - Critical Components

Semantic Interoperability Principles:

- *Standards for data exchange*
- *Shared semantic alignment and mapping*
- *Rich metadata*
- *Governance among stakeholders*

- **Common Semantic Model** supports Harmonization and Mappings
 - Declared by the US Office of the National Coordinator based on Health Level 7 Standards (HL7)
 - United States Core for Data Interoperability (USDCI)
 - Set of Classes and Data Elements
- **Shared Concept System** for Semantic Alignment
 - Domain Specific Concept System → NCI Thesaurus (more than a Thesaurus)
 - Used to annotate Common Semantic Model and align Data Elements
- **Central Governance/Coordination** for Data Registration in ISO/IEC 11179
 - Standard, Structured Metadata based on Concepts
 - Consistent application of the Concepts in the Semantic Model

Common Semantic Model


- US Office of the National Coordinator (ONC)
 - United States Core for Data Interoperability (USDCI)
- Register USDCI in ISO/IEC 11179 Registry
- Annotated with Concepts from NCI Thesaurus
- Map disparate Data Models to Common Semantic Model

USDCI v3 Summary of Data Classes and Data Elements

Allergies and Intolerances <ul style="list-style-type: none"> • Substance (Medication) • Substance (Drug Class) • Reaction 	Health Status/Assessments <ul style="list-style-type: none"> • Health Concerns • Functional Status • Disability Status • Mental/Cognitive Status • Pregnancy Status • Smoking Status 	Problems <ul style="list-style-type: none"> • Problems • SDOH Problems/Health Concerns • Date of Diagnosis • Date of Resolution
Assessment and Plan of Treatment <ul style="list-style-type: none"> • Assessment and Plan of Treatment • SDOH Assessment 	Immunizations <ul style="list-style-type: none"> • Immunizations 	Procedures <ul style="list-style-type: none"> • Procedures • SDOH Interventions • Reason for Referral
Care Team Member(s) <ul style="list-style-type: none"> • Care Team Member Name • Care Team Member Identifier • Care Team Member Role • Care Team Member Location • Care Team Member Telecom 	Laboratory <ul style="list-style-type: none"> • Tests • Values/Results • Specimen Type • Result Status 	Provenance <ul style="list-style-type: none"> • Author Organization • Author Time Stamp
Clinical Notes <ul style="list-style-type: none"> • Consultation Note • Discharge Summary Note • History & Physical • Procedure Note • Progress Note 	Medications <ul style="list-style-type: none"> • Medications • Dose • Dose Unit of Measure • Indication • Fill Status 	Unique Device Identifier(s) for a Patient's Implantable Device(s) <ul style="list-style-type: none"> • Unique Device Identifier(s) for a patient's implantable device(s)
Clinical Tests <ul style="list-style-type: none"> • Clinical Test • Clinical Test Result/Report 	Patient Demographics/Information <ul style="list-style-type: none"> • First Name • Last Name • Middle Name (Including middle initial) • Name Suffix • Previous Name • Date of Birth • Date of Death • Race • Ethnicity • Tribal Affiliation • Sex • Sexual Orientation • Gender Identity • Preferred Language • Current Address • Previous Address • Phone Number • Phone Number Type • Email Address • Related Person's Name • Related Person's Relationship • Occupation • Occupation Industry 	Vital Signs <ul style="list-style-type: none"> • Systolic Blood Pressure • Diastolic Blood Pressure • Heart Rate • Respiratory Rate • Body Temperature • Body Height • Body Weight • Pulse Oximetry • Inhaled Oxygen Concentration • BMI Percentile (2 - 20 years) • Weight-for-length Percentile (Birth - 24 Months) • Head Occipital-frontal Circumference Percentile (Birth- 36 Months)
Diagnostic Imaging <ul style="list-style-type: none"> • Diagnostic Imaging Test • Diagnostic Imaging Report 		
Encounter Information <ul style="list-style-type: none"> • Encounter Type • Encounter Diagnosis • Encounter Time • Encounter Location • Encounter Disposition 		
Goals <ul style="list-style-type: none"> • Patient Goals • SDOH Goals 		
Health Insurance Information <ul style="list-style-type: none"> • Coverage Status • Coverage Type • Relationship to Subscriber • Member Identifier • Subscriber Identifier • Group Number • Payer Identifier 		

Shared Concept System

NCI Thesaurus - Formal declarations make meaning machine interpretable

Equivalent To 

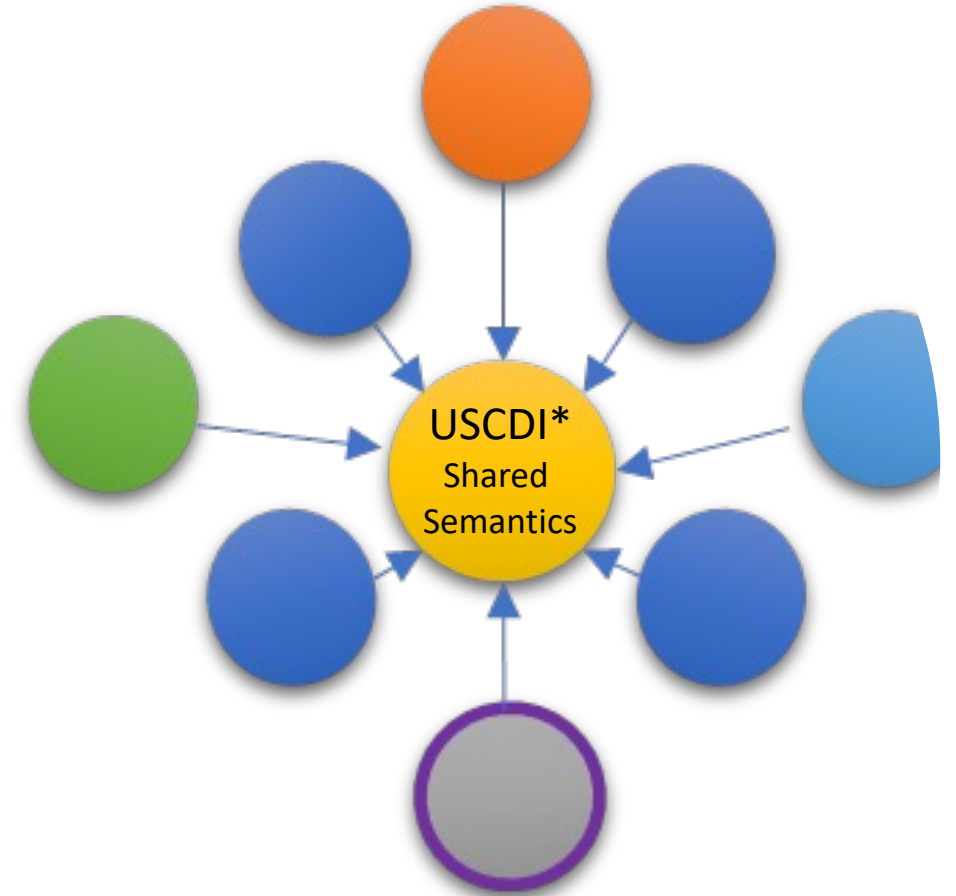
- **Adenoma**
 - and 'Benign Endocrine Neoplasm'
 - and 'Benign Adrenal Cortex Neoplasm'
 - and (Disease_Has_Normal_Cell_Origin some 'Adrenal Cortical Cell')
 - and (Disease_Has_Abnormal_Cell some 'Neoplastic Adrenal Cortical Cell')
 - and (Disease_Has_Finding some 'Well-Circumscribed Lesion')
 - and (Disease_May_Have_Cytogenetic_Abnormality some 'Gain of Chromosome 1q')
 - and (Disease_May_Have_Cytogenetic_Abnormality some 'Gain of Chromosome 9q')
 - and (Disease_May_Have_Cytogenetic_Abnormality some 'Loss of Chromosome 1p')
 - and (Disease_May_Have_Finding some 'Encapsulated Mass')
 - and (Disease_Excludes_Finding some 'Tumor Vascular Invasion')
 - and (Disease_Excludes_Finding some 'Capsular Invasion')

Figure 2. The figure shows an example of a logical definition for the concept C9003, "adrenal cortex adenoma," in the NCI.

**Example from the National Cancer Institute Thesaurus*

Common Semantic Model vs Common Data Model

- Shared Semantics
 - *Provides clear, unambiguous meaning*
 - Independent of Data Model
 - Enables mapping and transformation
- Annotated and Registered using Shared Concept Annotations
 - ISO/IC 11179, NCI Thesaurus
- Concept Systems enhance Knowledge Discovery



**United States Core Data for Interoperability*

References

- ISO/IEC 11179 Information technology — Metadata registries (MDR) — Part 3: Metamodel for registry common facilities
- ISO/IEC 11179 Information technology — Metadata registries (MDR) — Part 31: Data Specification
- ISO/IEC 5493 Criteria for concept systems
- United States Core for Data Interoperability (USDCI)
- NCI Semantics Primer - Blog
 - <https://datascience.cancer.gov/news-events/blog/semantics-primer>