

Data Science Informatics Glossary

**NCATS Steering Committee
Real World Data Task Force (RWD-TF)**

Last Edited: May 20, 2026

Table of Contents

[Table of Contents](#)

[A](#)

[Algorithm](#)

[Medical Algorithm](#)

[Analytical AI](#)

[Apache Spark](#)

[Application Programming Interface \(API\)](#)

[Artificial Intelligence \(AI\)](#)

[Attribution](#)

[B](#)

[Biomedical Informatics \(See Informatics\)](#)

[C](#)

[Chatbot](#)

[Classification](#)

[Clinical Data Interchange Standards Consortium \(CDISC\)](#)

[Analysis Data Model \(ADaM\)](#)

[Clinical Data Acquisition Standards Harmonization \(CDASH\)](#)

[Protocol Representation Model \(PRM\)](#)

[Questionnaires, Ratings and Scales \(QRS\)](#)

[Standard for Exchange of Nonclinical Data \(SEND\)](#)

[Study Data Tabulation Model \(SDTM\)](#)

[Study Data Tabulation Model Implementation Guide \(SDTMIG\)](#)

[Clinical Decision Support \(CDS\)](#)

[Clinical Informatics \(See Informatics\)](#)

[Cohort Discovery](#)

[Common Data Model \(CDM\)](#)

[Computable Phenotype](#)

[Concept Set](#)

[A](#)
[B](#)
[C](#)
[D](#)
[E](#)
[F](#)
[G](#)
[H](#)
[I](#)
[J](#)
[K](#)
[L](#)
[M](#)
[N](#)
[O](#)
[P](#)
[Q](#)
[R](#)
[S](#)
[T](#)
[U](#)
[V](#)
[W](#)
[X](#)
[Y](#)
[Z](#)
[↑](#)

	<u>Controlled Vocabulary</u>
<u>A</u>	<u>Current Procedural Terminology (CPT)</u>
<u>B</u>	
<u>C</u>	
<u>D</u>	<u>Data Access Committee (DAC)</u>
<u>E</u>	<u>Data Analytics</u>
<u>F</u>	<u>Data Harmonization</u>
<u>G</u>	<u>Data Management and Sharing Policy (DMS Policy)</u>
<u>H</u>	<u>Data Provenance</u>
<u>I</u>	<u>Data Quality</u>
<u>J</u>	<u>Data Stewardship</u>
<u>K</u>	<u>Data Transfer Agreement (DTA)</u>
<u>L</u>	<u>Data Use Agreement</u>
<u>M</u>	<u>Data Warehouse</u>
<u>N</u>	<u>Database</u>
<u>O</u>	<u>Deep Learning</u>
<u>P</u>	<u>Deep Neural Network</u>
<u>Q</u>	<u>De-identified Data</u>
<u>R</u>	
<u>S</u>	<u>eConsent</u>
<u>T</u>	<u>Electronic Health Record (EHR)</u>
<u>U</u>	<u>Encounter (See HL7 Episode of Care)</u>
<u>V</u>	<u>Extract, Transform, and Load (ETL)</u>
<u>W</u>	
<u>X</u>	
<u>Y</u>	<u>FAIR Data</u>
<u>Z</u>	<u>FAIR Data Principles</u>
<u>↑</u>	<u>Fast Healthcare Interoperability Resources (FHIR) (See HL7 FHIR)</u>
	<u>Generative AI (GenAI)</u>
	<u>Graph Database (GDB)</u>

H

A

[Health Informatics \(See Informatics\)](#)

B

[Health Information Exchange \(HIE\)](#)

C

[Health Information Technology for Economic and Clinical Health \(HITECH\) Act](#)

D

E

[Health Insurance Portability and Accountability Act \(HIPAA\)](#)

F

[Health Level 7 \(HL7\)](#)

G

[HL7 ACCELERATOR Program](#)

H

[HL7 Encounter](#)

I

[HL7 Episode of Care](#)

J

[HL7 Fast Healthcare Interoperability Resources \(FHIR\)](#)

K

[Healthcare Common Procedure Coding System \(HCPCS\)](#)

L

I

M

[International Classification of Disease \(ICD\)](#)

N

[ICD-10](#)

O

[ICD-11](#)

P

[Informatics](#)

Q

[Applied Informatics](#)

R

[Biomedical Informatics](#)

S

[Clinical Informatics](#)

T

[Health Informatics](#)

U

[Nursing Informatics](#)

V

[Public Health Informatics](#)

W

[Interoperability](#)

X

L

Y

[Large Language Model](#)

Z

[Large Multimodal Model](#)

↑

[Limited Data Set](#)

M

[Logical Observations Identifiers, Names, Codes \(LOINC\)](#)

M

A		Machine Learning
B		Meaningful Use
C	N	
D		National Death Index (NDI)
E		National Drug Code (NDC)
F		Natural Language Processing
G		Neural Network
H		Nursing Informatics (See Informatics)
I	O	
J		Observational Research
K		Observation Research Centralized vs Distributed (Federated)
L		Office of the National Coordinator (ONC)
M		Ontology
N		Outpatient Visit
O	P	
P		Patient Level Deduplication
Q		Phenotype (See Computable Phenotype)
R		Privacy Preserving Record Linkage (PPRL)
S		Prompt Engineering
T		Public Health Informatics (See Informatics)
U	Q	
V		Qualified Health Information Network (QHIN)
W	R	
X		Real-world Data
Y		Real-world Evidence
Z		Relational Database (RDB)
↑		REpresentational State Transfer (REST API)
		RxNorm

<u>A</u>	
<u>B</u>	
<u>C</u>	
<u>D</u>	
<u>E</u>	
<u>F</u>	
<u>G</u>	
<u>H</u>	
<u>I</u>	
<u>J</u>	
<u>K</u>	
<u>L</u>	
<u>M</u>	
<u>N</u>	
<u>O</u>	
<u>P</u>	
<u>Q</u>	
<u>R</u>	
<u>S</u>	
<u>T</u>	
<u>U</u>	
<u>V</u>	
<u>W</u>	
<u>X</u>	
<u>Y</u>	
<u>Z</u>	
<u>↑</u>	

S

[Statistical Classification](#)

[Structured Query Language \(SQL\)](#)

[Systematized Nomenclature of Medicine Clinical Terms \(SNOMED CT\)](#)

I

[Terminology](#)

[Thesaurus](#)

[Title 21 CFR Part 11](#)

[Training Data](#)

[Transformer](#)

[Trial Emulation](#)

[Trusted Exchange Framework and Common Agreement \(TEFCA\)](#)

[Tuning](#)

U

[Unified Code for Units of Measure \(UCUM\)](#)

[United States Core Data for Interoperability \(USCDI\)](#)

V

[Value Set](#)

Glossary Acknowledgements: at the end of each definition, a hyperlink in parentheses directs to the source of information.

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A

Algorithm

A process for carrying out a complex task broken down into simple decision and action steps. Often assists the requirements analysis process carried out before programming. ([J Epidemiol Community Health](#))

A step-by-step set of instructions that a computer follows to solve a problem or perform a task. ([MedPage](#))

A clearly specified mathematical process for computation; a set of rules that, if followed, will give a prescribed result. ([NIST](#))

Medical Algorithm

Any computation, formula, statistical survey, nomogram, or look-up table useful in healthcare. Medical algorithms include decision tree approaches to healthcare treatment. In many cases, algorithms are not components of digital computer programs. Some examples of algorithms are calculators, (e.g. for body mass index), flowcharts (decision trees for identifying health conditions and look-up tables (e.g. for looking up nutritional contents of foodstuffs). ([Wikipedia](#))

Analytical AI

Analytical AI takes mass amounts of data and produces observations or analyses of the information collected. Analytical AI might be used to identify and assess patterns or relationships in data, such as test results. ([Rosalyn](#))

Apache Spark

Apache Spark™ is a multi-language engine for executing data engineering, data science, and machine learning on single-node machines or clusters. ([Apache Spark](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

Application Programming Interface (API)

An API, or application programming interface, is a set of rules or protocols that let software applications communicate with each other to exchange data, features and functionality. ([IBM](#))

Artificial Intelligence (AI)

A branch of computer science that focuses on creating machines, computer programs, or software systems that can perform tasks, solve problems, and learn from those experiences in ways that usually require human intelligence. ([MedPage](#))

Attribution

Attribution is a foundation concept when using a resource or text released with an open-copyright license. This legal requirement states that users must attribute — give credit — to the creator of the work. ([Copyright and Open License](#))

B

Biomedical Informatics ([See Informatics](#))

C

Chatbot

A software application that mimics human conversation. ([MedPage](#))

Classification

Controlled vocabulary that is intended to comprehensively describe a topic or domain from a conceptual perspective and is not developed solely from a text corpus that it is meant to describe. ([N Engl J Med](#))

Clinical Data Interchange Standards Consortium (CDISC)

CDISC Foundational Standards are the basis of the complete suite of standards, supporting clinical and non-clinical research processes from end to end. Foundational Standards focus on the core principles for defining data standards and include models, domains and specifications for data representation. ([CDISC](#))

Analysis Data Model (ADaM)

Defines dataset and metadata standards that support efficient generation, replication, and review of clinical trial statistical analyses as well as traceability between analysis results, analysis data, and data represented in SDTM. ([CDISC](#))

Clinical Data Acquisition Standards Harmonization (CDASH)

Establishes a standard way to collect data in a similar way across studies and sponsors so that data collection formats and structures provide clear traceability of submission data into the Study Data Tabulation Model (SDTM), delivering more transparency to regulators and others who conduct data review. ([CDISC](#))

Protocol Representation Model (PRM)

Provides a standard for planning and designing a research protocol with focus on study characteristics such as study design, eligibility criteria, and requirements from the ClinicalTrials.gov, World Health Organization (WHO) registries, and EudraCT registries. PRM assists in automating CRF creation and EHR configuration to support clinical research and data sharing. ([CDISC](#))

Questionnaires, Ratings and Scales (QRS)

A series of questions, tasks or assessments used in clinical research to provide a qualitative or quantitative assessment of a clinical concept or task-based observation. The QRS team develops Controlled Terminology and SDTM (tabulation) supplements; the ADQRS Team develops ADaM (analysis) supplements. ([CDISC](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

Standard for Exchange of Nonclinical Data (SEND)

An implementation of the SDTM standard for nonclinical studies. SEND specifies a way to collect and present nonclinical data in a consistent format. ([CDISC](#))

Study Data Tabulation Model (SDTM)

Provides a standard for organizing and formatting data to streamline processes in collection, management, analysis and reporting. ([CDISC](#))

Study Data Tabulation Model Implementation Guide (SDTMIG)

Intended to guide the organization, structure, and format of standard clinical trial tabulation datasets. ([CDISC](#))

Clinical Decision Support (CDS)

Provides clinicians, staff, patients or other individuals with knowledge and person-specific information, intelligently filtered or presented at appropriate times, to enhance health and health care. In practice, it encompasses a variety of tools to enhance health care decision-making, including computerized alerts and reminders to both health care providers and patients; focused patient data reports and summaries; diagnostic support and contextually relevant reference information, etc. ([U of San Diego](#))

Clinical Informatics ([See Informatics](#))

Cohort Discovery

Enables researchers to determine counts of patients using data from EHR systems or Research Data Warehouses. Note cohort discovery can be used in identified or de-identified databases. In the case of a de-identified database only authorized uses are allowed to use an identifier to re-identify individuals. ([UCLA](#))

Common Data Model (CDM)

Refers to any standardized data model which allows for data and information exchange between different applications and data sources.

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

Common data models aim to standardize logical infrastructure so that related applications can "operate on and share the same data" and can be seen as a way to "organize data from many sources that are in different formats into a standard structure." ([Wikipedia](#))

Examples of common data models include PCORNet, OMOP, ACT/i2b2 (ENACT), TriNetX, Sentinel.

Computable Phenotype

A Computable Phenotype is a specification for identifying patients or populations with a given characteristic or condition of interest from EHRs using data that are routinely collected in EHRs or ancillary data sources. A computable phenotype definition consists of data elements and logical expressions (such as AND, OR, and NOT) that can be interpreted and executed by a computer. In other words, the syntax defining a computable phenotype is designed to be interpreted and executed programmatically without human intervention. Computable phenotype definitions often rely on value sets—lists of codes from standardized medical vocabularies that indicate a condition, drug exposure, or other clinical phenomenon of interest. ([Rethinking Clinical Trials](#))

Concept Set

An expression representing a list of concepts that can be used as a reusable component in various analyses. It can be thought of as a standardized, computer-executable equivalent of the code lists often used in observational studies. ([The Book of OHDSI](#))

Controlled Vocabulary

Set of preferred terms created specifically for a domain or body of text. ([N Engl J Med](#))

Current Procedural Terminology (CPT)

CPT is a list of descriptive terms and numeric codes for medical services and procedures that are provided by physicians and health care

professionals. CPT codes are used by healthcare professionals for billing of medical services and procedures to public and private health insurance programs. ([AMA](#))

D

Data Access Committee (DAC)

Provides guidance and oversight of controlled-access data housed within NIH designated repositories, including the database of Genotypes and Phenotypes. ([NIAID](#))

Data Analytics

The systematic use of data to obtain new insights and drive fact-based decision making for patient care outcomes as well as clinical and operational improvements. The types of data being utilized include:

- Claims and cost data
- Pharmaceutical data, research and development data
- Clinical data collected from electronic medical records
- Data on patient behavior and preferences ([U of San Diego](#))

Data Harmonization

Harmonized data refers to the collection of raw data from multiple sources that have been normalized so that a valid comparison can be made across these sources. Data harmonization can take many forms, including transformation of data represented in one Common Data Model (CDM) to another, value set mappings (e.g. a set of races in one survey to a standard set), or terminology mapping (e.g. from a disease in ICD-10, to SNOMED-CT), or transformation of units to a common standard (e.g. pounds to kg). ([GDC](#))

Data Management and Sharing Policy (DMS Policy)

Promotes the sharing of scientific data. Sharing scientific data accelerates biomedical research discovery, in part, by enabling validation of research

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

results, providing accessibility to high-value datasets, and promoting data reuse for future research studies.

Under the DMS policy, NIH expects that investigators and institutions plan and budget for the managing and sharing of data, submit a DMS plan for review when applying for funding, and comply with the approved DMS plan. ([NIH](#))

Data Provenance

In the context of computers and law enforcement use, it is an equivalent term to chain of custody. It involves the method of generation, transmission and storage of information that may be used to trace the origin of a piece of information processed by community resources. ([NIST](#))

Data Quality

The degree to which data items are accurate, complete, relevant, timely, sufficiently detailed, appropriately represented (for example, consistently coded using a clinical coding system), and retain sufficient contextual information to support decision making. ([J Epidemiol Community Health](#))

Data Stewardship

Data stewardship encompasses the responsibilities and accountabilities associated with managing, collecting, viewing, storing, sharing, disclosing or otherwise making use of personal health information. ([U of San Diego](#))

Data Transfer Agreement (DTA)

A DTA is a contract between the providing and recipient institutions that governs the legal obligations and restrictions, as well as compliance with applicable laws and regulations, related to the transfer of such data between the parties. ([Transfer Agreements – Institutional Review Board](#))

Data Use Agreement

A data use agreement establishes who is permitted to use and receive data, and the permitted uses and disclosures of such information by the recipient, and provides that the recipient will:

- not use or disclose the information other than as permitted by the DUA or as otherwise required by law,
- use appropriate safeguards to prevent uses or disclosures of the information that are inconsistent with the DUA,
- report to the covered entity uses or disclosures that are in violation of the DUA, of which it becomes aware
- ensure that any agents to whom it provides the LDS agree to the same restrictions and conditions that apply to the LDS recipient, with respect to such information, and
- not re-identify the information or contact the individual.

([eCFR :: 45 CFR 164.514 -- Other requirements relating to uses and disclosures of protected health information](#))

Data Warehouse

An information infrastructure that enables researchers and clinicians to access and analyze detailed data and trends. Created by collecting databases and linking them using common data elements. ([The Guide to Health Informatics 3rd Edition](#))

Database

A collection of data in machine readable format organized so that it can be retrieved or processed automatically by computer. A flat file database is organized like a card file, with many records (cards) each including one or more fields (data items). A relational database is organized as one or more related tables, each containing columns and rows. Data is organized in a database according to a schema or data model; some items are often coded using a clinical coding system. ([J Epidemiol Community Health](#))

Deep Learning

A type of machine learning that uses artificial neural networks. ([MedPage](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

Deep Neural Network

A step above neural networks, this type of machine learning is composed of multiple layers of interconnected "neurons," enabling it to learn complex patterns in data. ([MedPage](#))

De-identified Data

Health information that does not identify an individual and with respect to which there is no reasonable basis to believe that the information can be used to identify an individual is not individually identifiable health information. ([HHS](#))

E

eConsent

The explicit agreement from a patient to allow another party to view the data contained in his or her electronic health record. ([U of San Diego](#))

Electronic Health Record (EHR)

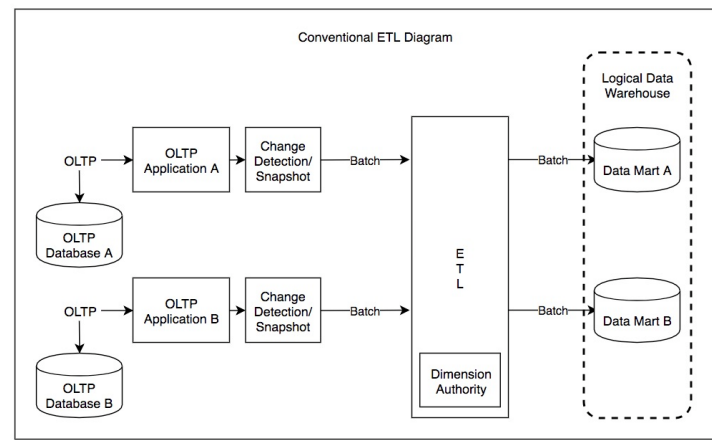
EHRs offer a complete real-time record of a person's medical history that is easily accessible by health providers (and increasingly, by patients). In addition to being a comprehensive record of an individual's health care, the documentation inside EHRs supports the functions of billing, quality management, outcome reporting and public health disease surveillance and reporting. EHR is sometimes used interchangeably with EMR (electronic medical record). ([U of San Diego](#))

A real-time patient health record with access to evidence-based decision support tools that can be used to aid clinicians in decision making. The EHR can automate and streamline a clinician's workflow, ensuring that all clinical information is communicated. It can also prevent delays in response that result in gaps in care. The EHR can also support the collection of data for uses other than clinical care, such as billing, quality management, outcome reporting, and public health disease surveillance and reporting. ([HealthIT](#))

Encounter ([See HL7 Episode of Care](#))

Extract, Transform, and Load (ETL)

In computing, ETL is a three-phase process where data is extracted, transformed (cleaned, sanitized, scrubbed) and loaded into an output data container. The data can be collated from one or more sources, and it can also be output to one or more destinations. ETL processing is typically executed using software applications, but it can also be done manually by system operators. ETL software typically automates the entire process and can be run manually or on recurring schedules either as single jobs or aggregated into a batch of jobs.



Conventional ETL diagram (Wikipedia)

A properly designed ETL system extracts data from source systems and enforces data type and data validity standards and ensures it conforms structurally to the requirements of the output. Some ETL systems can also deliver data in a presentation-ready format so that application developers can build applications and end users can make decisions. The ETL process is often used in data warehousing. ETL systems commonly integrate data from multiple applications (systems), typically developed and supported by different vendors or hosted on separate computer hardware. The separate systems containing the original data are frequently managed and operated by different stakeholders. For example, in RWD combination of multiple data models like OMOP and PCORNET, or different types of data from and EHR System and CMS claims. ([Wikipedia](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

F

FAIR Data

FAIR data is easier to share and reuse, researchers and data scientists try to align with the FAIR data principles.

FAIR data is:

- Findable: For data to be findable there must be sufficient metadata; there must be a unique and persistent identifier; and the data must be registered or indexed in a searchable resource.
- Accessible: To be accessible, metadata and data should be readable by humans and by machines, and it must reside in a trusted repository.
- 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. ([NLM](#))([Wilkinson, et al. 2016](#))

FAIR Data Principles

One of the grand challenges of data-intensive science is to facilitate knowledge discovery by assisting humans and machines in their discovery of, access to, integration and analysis of, task-appropriate scientific data and their associated algorithms and workflows. Here, we describe FAIR – a set of guiding principles to make data Findable, Accessible, Interoperable, and Reusable. The term FAIR was launched at a Lorentz workshop in 2014, the resulting FAIR principles were published in 2016. ([Force11](#))

Findable

- F1. (meta)data are assigned a globally unique and eternally persistent identifier.
- F2. data are described with rich metadata.
- F3. (meta)data are registered or indexed in a searchable resource.
- F4. metadata specifies the data identifier.

Accessible

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A1. (meta)data are retrievable by their identifier using a standardized communications protocol.

A1.1 the protocol is open, free, and universally implementable.

A1.2 the protocol allows for an authentication and authorization procedure, where necessary.

A2. metadata are accessible, even when the data are no longer available.

Interoperable

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (meta)data use vocabularies that follow FAIR principles.

I3. (meta)data include qualified references to other (meta)data.

Reusable

R1. (meta)data have a plurality of accurate and relevant attributes.

R1.1. (meta)data are released with a clear and accessible data usage license.

R1.2. (meta)data are associated with their provenance.

R1.3. (meta)data meet domain-relevant community standards.

Fast Healthcare Interoperability Resources (FHIR) ([See HL7 FHIR](#))

G

Generative AI (GenAI)

Generative artificial intelligence (AI) tools use machine learning models trained on massive pools of information to learn patterns from data to create novel content like text, images, audio, or video in response to a prompt. Unlike internet searches, generative AI tools do not use algorithms to locate and curate existing sources. Instead, they create new content by predicting what word, sound, or pixel would come next in a pattern. ([University of Pittsburgh](#))

Generative AI refers to deep-learning models that can generate high-quality text, images, and other content based on the data they were trained on. ([IBM Research](#))

Graph Database (GDB)

A database that uses graph structures for semantic queries with nodes, edges, and properties to represent and store data. A key concept of the system is the *graph* (or *edge* or *relationship*). The graph relates the data items in the store to a collection of nodes and edges, the edges representing the relationships between the nodes. The relationships allow data in the store to be linked together directly and, in many cases, retrieved with one operation. Graph databases hold the relationships between data as a priority. Querying relationships is fast because they are perpetually stored in the database. Relationships can be intuitively visualized using graph databases, making them useful for heavily interconnected data. ([Wikipedia](#))

H

Health Informatics ([See Informatics](#))

Health Information Exchange (HIE)

Electronic health information exchange (HIE) allows doctors, nurses, pharmacists, other health care providers and patients to appropriately access and securely share a patient's vital medical information electronically, improving the speed, quality, safety and cost of patient care. ([ASTP](#))

Health Information Technology for Economic and Clinical Health (HITECH) Act

The HITECH Act of 2009 was enacted to promote the adoption and meaningful use of health information technology. It includes multiple provisions related to the expanded use of electronic medical records, privacy protocols (HIPAA), structures for Medicare and Medicaid reimbursement, and an ever-evolving set of rules and systems seeking to

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

incentivize providers to deliver the highest quality care. Its adoption has also dramatically increased the demand for skilled labor in the healthcare IT industry. ([U of San Diego](#))

Health Insurance Portability and Accountability Act (HIPAA)

A federal law designed to provide privacy standards to protect patients' medical records and other health information provided to health plans, doctors, hospitals and other health care providers. It is also intended to prevent fraud and abuse and to help protect employees and their families from losing health insurance coverage after a job change or loss. ([U of San Diego](#))

Health Level 7 (HL7)

HL7's mission is to provide standards that empower global health data interoperability.

Founded in 1987, Health Level Seven International (HL7) is a not-for-profit, ANSI-accredited standards developing organization dedicated to providing a comprehensive framework and related standards for the exchange, integration, sharing, and retrieval of electronic health information that supports clinical practice and the management, delivery and evaluation of health services. HL7 is supported by more than 1,600 members from over 50 countries, including 500+ corporate members representing healthcare providers, government stakeholders, payers, pharmaceutical companies, vendors/suppliers, and consulting firms. ([HL7 International](#))

HL7 ACCELERATOR Program

The HL7 FHIR ACCELERATOR program is designed to assist communities and collaborative groups across the global health care spectrum in the creation and adoption of high quality FHIR Implementation Guides or other standard artifacts to move toward the realization of global health data interoperability. (Vulcan, Davinci, Argonauts, Gravity) ([HL7 International](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

HL7 Encounter

An interaction between a patient and healthcare provider(s) for the purpose of providing healthcare service(s) or assessing the health status of a patient. Encounters are primarily used to record information about the actual activities that occurred, where Appointment is used to record planned activities. ([HL7 FHIR](#))

HL7 Episode of Care

An association between a patient and an organization / healthcare provider(s) during which time encounters may occur. The managing organization assumes a level of responsibility for the patient during this time. ([HL7 FHIR](#))

HL7 Fast Healthcare Interoperability Resources (FHIR)

The HL7® FHIR® standard defines how healthcare information can be exchanged between different computer systems regardless of how it is stored in those systems. It allows healthcare information, including clinical and administrative data, to be available securely to those who have a need to access it, and to those who have the right to do so for the benefit of a patient receiving care. The standards development organization HL7® (Health Level Seven®3) uses a collaborative approach to develop and upgrade FHIR. ([HL7 FHIR](#))

Healthcare Common Procedure Coding System (HCPCS)

Standardized coding systems are essential so Medicare and other health insurance programs can process claims in an orderly and consistent manner. HCPCS is divided into 2 main subsystems — Level I and Level II. ([CMS](#))

Authority

The Healthcare Common Procedure Coding System (HCPCS) is produced by the Centers for Medicare and Medicaid Services (CMS).

Purpose

HCPCS is a collection of standardized codes that represent medical procedures, supplies, products and services. The codes are used to

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

facilitate the processing of health insurance claims by Medicare and other insurers.

Level I

A coding system published by the American Medical Association or AMA. Level I HCPCS (CPT-4) has codes and terms to identify medical services or procedures physicians perform. Level I HCPCS (CPT-4) has codes and terms to identify medical services or procedures physicians perform.

Level II

Refers to products, services, and supplies that aren't included in the CPT-4. Some examples are DMEPOS or durable medical equipment, prosthetics, orthotics, and supplies. The Level II of HCPCS reports non-physician services together with DMEPOS.

I

International Classification of Disease (ICD)

The International Classification of Diseases (ICD) is a globally standardized system developed and maintained by the World Health Organization (WHO) for classifying diseases, disorders, injuries, and other health conditions. The ICD provides a common language and coding framework that enables the systematic recording, reporting, analysis, and comparison of health information across countries and over time, and supports public health surveillance, clinical care, research, health management, and health statistics. ([Britannica](#))

ICD-10

The International Classification of Diseases, 10th edition. Published by the World Health Organization. ([ICD](#))

ICD-11

The International Classification of Diseases, 11th edition. Published by the World Health Organization. ([ICD](#))

Informatics

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

Applied Informatics

Applied Informatics addresses the flow of medical information in an electronic environment and covers process, policy and technological solutions. ([U of San Diego](#))

Biomedical Informatics

American Medical Informatics Association (AMIA) defines biomedical informatics as “the interdisciplinary field that studies and pursues the effective uses of biomedical data, information, and knowledge for scientific inquiry, problem-solving and decision making, motivated by efforts to improve human health.” ([U of San Diego](#))

Clinical Informatics

Clinical Informatics is synonymous with health informatics (see entry below), according to the Healthcare Information and Management Systems Society and other sources. HIMSS defines the discipline as promoting “the understanding, integration, and application of information technology in health care settings.” ([U of San Diego](#))

Health Informatics

The U.S. National Library of Medicine defines health informatics (also called medical informatics) as the interdisciplinary study of the design, development, adoption, and application of IT-based innovations in health care services delivery, management, and planning. According to AHIMA, there are four major focus research areas in informatics education ([U of San Diego](#))

Nursing Informatics

Nursing informatics is described by the American Nursing Informatics Association (ANIA) as “a specialty that integrates nursing science with multiple information management and analytical sciences to identify, define, manage and communicate data, information, knowledge, and wisdom in nursing practice.” ([U of San Diego](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

Public Health Informatics

The systematic application of information and computer sciences to public health practice, research, and learning. This discipline, which integrates public health with information technology, is considered the key to unlocking the potential of information systems to improve the health of the populations throughout the world. ([U of San Diego](#))

Interoperability

According to HIMSS, interoperability describes the extent to which systems and devices can exchange data and interpret that shared data. For two systems to be interoperable, they must be able to exchange data and subsequently present that data in a way that can be understood by a user. In practical terms, it often refers to the ability of providers — for example, a patient’s primary care physician as well a specialist at another office or a doctor providing care in a hospital setting — to easily and securely access patient histories and other medical information without waiting for paper copies of health records to be sent. It also means empowering patients to utilize patient portals to access information about their own care. ([U of San Diego](#))

L

Large Language Model

An AI model trained on large amounts of data that can be adapted to perform a wide range of tasks. It learns the probabilities of the occurrence of sequences of words from a vast body of text. Examples include OpenAI's ChatGPT and Google's Bard. ([MedPage](#))

Large Multimodal Model

A machine learning model that incorporates more than one kind of modality for inputs and outputs, such as text, images, videos, and audio. ([MedPage](#))

Limited Data Set

Refers to PHI that excludes 16 categories of direct identifiers and may be used or disclosed, for purposes of research, public health, or health care operations, without obtaining either an individual's Authorization or a waiver or an alteration of Authorization for its use and disclosure, with a data use agreement. ([DHHS](#))

Logical Observations Identifiers, Names, Codes (LOINC)

LOINC is a clinical terminology important for laboratory test orders and results, produced by the Regenstrief Institute. ([Regenstrief Institute](#))

LOINC is one of a suite of designated standards for use in U.S. Federal Government systems for the electronic exchange of clinical health information. In 1999, it was identified by the HL7 Standards Development Organization as a preferred code set for laboratory test names in transactions between health care facilities, laboratories, laboratory testing devices, and public health authorities. ([LOINC](#))

M

Machine Learning

Sub-specialty of artificial intelligence concerned with developing methods for software to learn from experience, or to extract knowledge from examples in a database. See also: Artificial intelligence, Knowledge acquisition. Where computers analyze data to perform tasks without being specifically programmed to do so, using patterns discovered in those data sets. ([MedPage](#))

Meaningful Use

Meaningful Use refers to a program administered by the federal Centers for Medicare & Medicaid Services (CMS) that uses a system of financial awards and penalties to incentivize quality care and the most efficient use of electronic health records. ([U of San Diego](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

N

National Death Index (NDI)

A database of all deaths in the United States Working with state offices, the National Center for Health Statistics (NCHS) established the NDI as a resource to aid epidemiologists and other health and medical investigators with their mortality ascertainment activities.

The purpose of the NDI is to provide the public health and medical research community with an opportunity to obtain mortality follow-up information on their study participants, without having to go from jurisdiction to jurisdiction. ([CDC](#))

National Drug Code (NDC)

FDA's NDC Directory contains information about finished drug products, unfinished drugs and compounded drug products. ([FDA](#))

Natural Language Processing

A branch of AI that uses machine learning to process and interpret text and data. ([MedPage](#))

Neural Network

A computer system made up of interconnected "artificial neurons" that help computers learn from data and recognize patterns. It's inspired by the way the human brain works. ([MedPage](#))

Nursing Informatics ([See Informatics](#))

O

Observational Research

Observational research is a research technique where you observe participants and phenomena in their most natural settings. This enables

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

researchers to see their subjects make choices and react to situations in their natural setting, as opposed to structured settings like research labs or focus groups.

The term observational research is used to refer to several different types of non-experimental studies in which behavior is systematically observed and recorded. The goal of observational research is to describe a variable or set of variables. More generally, the goal is to obtain a snapshot of specific characteristics of an individual, group, or setting. As described previously, observational research is non-experimental because nothing is manipulated or controlled, and as such we cannot arrive at causal conclusions using this approach. The data that are collected in observational research studies are often qualitative in nature but they may also be quantitative or both (mixed-methods). There are several different types of observational research designs that will be described below. ([Research Methods in Psychology](#))

Observation Research Centralized vs Distributed (Federated)

Precursors to N3C: Federated Data Networks

Clinical data repositories from EHR sources have evolved over the decades. Ad hoc database designs from the early period of EHR adoption had limited generalization across sites. The earliest federated models in the 1990s ultimately gave rise to Sentinel and to Common Data Models (CDMs) such as PCORnet, the ACT network, TriNetX, and Observational Medical Outcomes Partnership (OMOP), which later became OHDSI. In a federated data network, each participating site's data stay behind its institutional firewall but are structured according to a CDM. This enables queries and results to be shared across sites rather than raw data.

The CDMs have been key to using EHR data for research; however, secondary analytic uses—particularly involving multiple contributing sites—require resource-intensive quality control to achieve the required uniformity and specificity to support open-ended research. Despite their importance, data quality (DQ) checks are inconsistently applied, and implementation methodologies are largely not evaluated. Further, evaluations of DQ often fail to accurately determine the data's

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

“fitness for use,” which evaluates both its intrinsic and intentional aspects. DQ evaluations must take into account the initial purpose of the data in their source systems, as well as the intended use of these data once harmonization processes have transformed them for use in secondary research.

Each of the aforementioned federated networks has methods to promote local DQ and adherence to data model conventions; these methods vary in maturity. As examples, OHDSI offers its DQ Dashboard tool for sites to run against their local CDM in order to evaluate adherence to OMOP CDM convention and diagnose common issues; PCORnet requires a quarterly data “curation” and quality check that uses prepackaged SAS scripts; ACT has a “smoke test” to ensure federated network query response and has a DQ Dashboard of its own under development; and TriNetX employs a growing library of DQ metrics and visualizations with site benchmarking as the basis for evaluating the results. Because the data are federated, DQ evaluation and remediation are performed locally at the site-by-site personnel. These local checks ensure that data conform to the specifications of the chosen CDM and may also check for data anomalies such as statistical outliers, invalid dates, biological implausibility, abundant missing data, and other common clinical data issues. Such approaches support the alignment of data; however, with the data remaining behind the institutional firewalls, it can be difficult to assess conformance or determine overall variability across sites, especially if the CDM does not have the capability of executing ad hoc DQ-related queries across the data network.

N3C’s Centralized Approach

In contrast to the federated approach, N3C pools data from each partner site in accordance with its signed Data Transfer Agreement and harmonizes all submitted data to the OMOP CDM. By the time data are submitted to N3C, sites have already applied a layer of local DQ checks and are submitting data that are “clean” by the local definition. Once data are merged across sites, additional opportunities for improvement may become apparent due to the ability to efficiently compare and benchmark among similar sites. N3C centralized DQ

methods complement and augment the foundation accomplished by sites in the context of the federated networks.

Data Quality

DQ emerged in the 2010s as an explicit subject of attention and research in informatics. Notably, the Patient-Centered Outcomes Research Institute (PCORI) funded an effort to define DQ, resulting in the “Harmonization” framework of Kahn and colleagues. The dimensions of DQ defined there (internal verification, external validation, conformance, completeness, and plausibility) are the basis of the PCORnet and OHDSI conception and framework for DQ checking. N3C continued to build on this foundation for our centralized DQ approach.

Office of the National Coordinator (ONC)

The Office of the National Coordinator for Health Information Technology (ONC) is at the forefront of the administration’s health IT efforts and is a resource to the entire health system to support the adoption of health information technology and the promotion of nationwide, standards-based health information exchange to improve health care. ONC is organizationally located within the Office of the Secretary for the U.S. Department of Health and Human Services (HHS).

ONC is the principal federal entity charged with coordination of nationwide efforts to implement and use the most advanced health information technology and the electronic exchange of health information. The position of National Coordinator was created in 2004, through an Executive Order, and legislatively mandated in the Health Information Technology for Economic and Clinical Health Act (HITECH Act) of 2009.

Through its work, ONC remains focused on two strategic objectives:

- Advancing the development and use of health IT capabilities; and
- Establishing expectations for data sharing ([HealthIT](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

Ontology

A description of the concepts and relations in a domain, such as drug prescribing. Sample concepts here would be “patient”, “prescriber”, and “drug”; relevant relations might include “prescribes to”, “requests prescription from”, and “causes side effects to”. A taxonomy or hierarchy is a simple kind of ontology in which concepts are arranged according to only one relation: “is a kind of”. Note that ontology as used here has a different meaning from its use in the philosophy of science, an area of interest to theoretical epidemiologists. ([J Epidemiol Community Health](#))

A formal and rigorous description of the set of concepts entities within a body of knowledge and the relationships between them. An ontology can help reason about the entities by identifying legal and nonsense relationships and inferences. Usually represented as hierarchy, and often as a richly interconnected set of objects, concepts and other entities that embody knowledge about a given field. See also: Knowledge base, Language.

Controlled terminology invoking formal semantic relationships between and among concepts, manifested as a type of description logic, which is a subset of first-order predicate logic, chosen to accommodate computational tractability. A common example is OWL. ([N Engl J Med](#))

Outpatient Visit

National Hospital Ambulatory Medical Care Survey defines an outpatient department visit as a direct personal exchange between a patient (not currently admitted to the hospital) and a physician or other health care provider working under the physician’s supervision for the purpose of seeking care and receiving personal health services. ([CDC](#))

P

Patient Level Deduplication

Patient deduplication is the process of removing redundant patient records from a database, preventing fragmented and duplicate

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

information from being processed and ensuring that updates and queries apply to the correct record.

Inconsistencies in determining which records represent the same patient, as well as errors in combining patient’s immunization records, negatively affect the overall data quality and usefulness of Immunization Information Systems (IIS). The following resource was developed by subject matter experts in the field of patient deduplication and matching. ([CDC](#))

Phenotype ([See Computable Phenotype](#))

Privacy Preserving Record Linkage (PPRL)

Privacy preserving record linkage (PPRL) is a linking methodology that works to mitigate privacy concerns when linking person level data from disparate data sources. PPRL allows for additional privacy protections that encrypt or mask personally identifiable information (PII) used for person level matching. PPRL can be used for many things such as linking data, deduplication, cohort discovery.

Prompt Engineering

The process of creating specific inputs for generative large language models or large multimodal models that can be interpreted to produce a desired output. ([MedPage](#))

Public Health Informatics ([See Informatics](#))

Q

Qualified Health Information Network (QHIN)

A Qualified HIN (QHIN) is a network of organizations working together to share data. QHINs will connect directly to each other to ensure interoperability between the networks they represent. ([ASTP](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

R

Real-world Data

Real-world Data FDA

Real-world data are data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources. Examples of RWD include data derived from electronic health records, medical claims data, data from product or disease registries, and data gathered from other sources (such as digital health). ([FDA](#))

Real-world Data NCI

Data relating to patient health status and/or the delivery of health care routinely collected from sources other than traditional clinical trials. Examples of sources include data derived from electronic health records (EHRs); medical claims and billing data; data from product and disease registries; patient-generated data, including from in-home-use settings; and data gathered from other sources that can inform on health status, such as mobile devices. ([NCI](#))

Real-world Evidence

Real-world evidence is the clinical evidence about the usage and potential benefits or risks of a medical product derived from analysis of RWD. ([FDA](#))

Relational Database (RDB)

A relational database is a database based on the relational model of data, as proposed by E. F. Codd in 1970. A database management system used to maintain relational databases is a relational database management system (RDBMS). Many relational database systems are equipped with the option of using Structured Query Language (SQL) for querying and updating the database. ([Wikipedia](#))

REpresentational State Transfer (REST API)

REST, or REpresentational State Transfer, A REST API (also called a RESTful API or RESTful web API) is an application programming interface (API) that conforms to the design principles of the representational state transfer (REST) architectural style. REST APIs provide a flexible, lightweight way to integrate applications and to connect components in microservices architectures. ([IBM](#))

RxNorm

Authority

RxNorm, a standardized nomenclature for clinical drugs for humans, is produced by the U.S. National Library of Medicine (NLM). RxNorm is one of a suite of designated standards for use in U.S. Federal Government systems for the electronic exchange of clinical health information.

Purpose

The goal of RxNorm is to allow various systems using different drug nomenclatures to share data efficiently. RxNorm provides standard names for clinical drugs (active ingredient + strength + dose form) and for dose forms as administered to a patient. It provides links from clinical drugs, both branded and generic, to their active ingredients, drug components (active ingredient + strength), and related brand names.

Description

RxNorm is intended to cover all prescription medications approved for human use in the United States. Prescription medications from other countries may be included as opportunities allow, a principal consideration being that there be an authoritative source of information about these drugs. Drugs are added to the vocabulary as new products are put on the market. Over-the-counter (OTC) medications will be added and covered, as well, when reliable information about the medications can be found. Drugs are added to the vocabulary as new products are put on the market.

An RxNorm clinical drug name reflects the active ingredients,

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

strengths, and dose form composing that drug. When any of these elements vary, a new RxNorm drug name is created as a separate concept. Thus, an RxNorm name should exist for every strength and dose of every available combination of clinically significant ingredients.

National Drug Codes (NDCs) for specific drug products are linked to that product in RxNorm. RxNorm links its names to many of the drug vocabularies commonly used in pharmacy management and drug interaction software, including those of First Databank, Micromedex®, Gold Standard Alchemy, and Multum. By providing links between these vocabularies, RxNorm can mediate messages between systems not using the same software and vocabulary. ([NIH](#))

S

Statistical Classification

Classification in which all concepts are mutually exclusive to avoid counting anything twice. This is typically achieved with the use of a mono-hierarchy, in which each concept has one and only one parent, such as the International Classification of Diseases (ICD). A statistical classification is exhaustive because it includes residual categories such as “unspecified” or “not elsewhere classified.” ([N Engl J Med](#))

Structured Query Language (SQL)

Structured Query Language (SQL) is a domain-specific language used to manage data, especially in a relational database management system (RDBMS). It is particularly useful in handling structured data, i.e., data incorporating relations among entities and variables. ([Wikipedia](#))

Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT)

SNOMED CT is one of a suite of designated standards for use in the U.S. Federal Government systems for the electronic exchange of clinical health information and is also a required standard in interoperability specifications of the U.S. Healthcare Information Technology Standards

Panel. The clinical terminology is owned and maintained by SNOMED International, a not-for-profit association. ([NIH](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

T

Terminology

A standard set of symbols or words used to describe the concepts, processes and objects of a given field of study. Examples of coding/terminologies important in research and clinical include

- Diagnosis Codes: SNOMed, ICD
- Procedure Codes: AMA HCPCS
- Medications: RxNorm, NDC, First Databank, Mulltum
- Lab Codes: LOINC
- CDISC: ADAM, CDASH, SDTM
- Adverse Reaction Coding in Regulated Trials: MEDRA

Set of preferred or official terms in a domain. A terminology may be a systematic nomenclature supported by a centralized body or as simple as the common usage that arises in a specific community of practice. ([N Engl J Med](#))

Thesaurus

Terminology that clusters synonyms and plesionyms (near synonyms) into categories. ([N Engl J Med](#))

Title 21 CFR Part 11

The part of Title 21 of the Code of Federal Regulations that establishes the United States Food and Drug Administration (FDA) regulations on electronic records and electronic signatures (ERES). Part 11 defines the criteria under which electronic records and electronic signatures are considered trustworthy, reliable, and equivalent to paper records. ([Wikipedia](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

Training Data

The data set contains the examples used to teach a machine learning application to recognize patterns and perform tasks. ([MedPage](#))

Transformer

A deep neural network designed to capture relationships and dependencies between elements in a sequence, such as words in a sentence. ([MedPage](#))

Trial Emulation

Target trial emulation is an observational research method that incorporates design features from idealized randomized trials (i.e., target trials) to improve the quality and interpretability of observational research. Specifically, this approach requires careful design of inclusion and exclusion criteria similar to those that would have been used in the target trial and consideration of the mode of treatment allocation and the timing and intensity of the exposure. Investigators recreate these trial design features in observational data and use statistical methods that capitalize on natural variability in provider practice to emulate random treatment assignment. ([Ann Am Thorac Soc](#))

Trusted Exchange Framework and Common Agreement (TEFCA)

The overall goal of the Trusted Exchange Framework and Common Agreement (TEFCA) is to establish a universal floor for interoperability across the country. The Common Agreement will establish the infrastructure model and governing approach for users in different networks to securely share basic clinical information with each other under commonly agreed-to expectations and rules, and regardless of which network they happen to be in. The Trusted Exchange Framework describes a common set of non-binding, foundational principles for trust policies and practices that can help facilitate exchange among HINs. ([ONC](#)) To qualify for incentive payments through the Centers for Medicare & Medicaid Services EHR Incentive Programs, eligible providers and hospitals must demonstrate meaningful use of an electronic health record (EHR). In other words, “meaningful use” sets the specific objectives that

eligible professionals and hospitals must achieve to participate in the EHR Incentive Programs. ([CMS](#))

A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z
↑

Tuning

The process of adapting a trained large language model to perform well on a specific task or domain. ([MedPage](#))

U

Unified Code for Units of Measure (UCUM)

The Unified Code for Units of Measure (UCUM) is a code system intended to include all units of measures being contemporarily used in international science, engineering, and business. The purpose is to facilitate unambiguous electronic communication of quantities together with their units. The focus is on electronic communication, as opposed to communication between humans. A typical application of The Unified Code for Units of Measure are electronic data interchange (EDI) protocols, but there is nothing that prevents it from being used in other types of machine communication. ([UCUM](#))

United States Core Data for Interoperability (USCDI)

The United States Core Data for Interoperability (USCDI) is a standardized set of health data classes and constituent data elements for nationwide, interoperable health information exchange. Review the USCDI Fact Sheet to learn more. ([ASTP](#))

V

Value Set

A value set consists of the numerical values (codes) and human-readable names (descriptions), drawn from standard vocabularies such as LOINC, RxNorm and SNOMED CT® ([Stud Health Technol Inform](#))