



NC\_045512.2 29903 bp ss-RNA linear VRL 30-MAR-2020  
 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1, complete genome



Severe acute respiratory syndrome coronavirus 2 (SARS-CoV2)  
 ANKH Severe acute respiratory syndrome coronavirus 2  
 Orthocoronavirinae; Coronavirales; Coronaviridae;  
 Orthocoronavirinae; Betacoronavirus; Sarbecovirus.  
 Nucleotide sequence (bases 13476 to 13503)  
 Baranov, P.V., Henderson, C.M., Anderson, C.B., Gesteland, R.F., Atkins, J.F. and Howard, M.T.  
 Programmed ribosomal frameshifting in decoding the SARS-CoV genome  
 JOURNAL: Virology 332 (2) 498-509 (2005)  
 MED: 15680415



Search results for SARS-CoV-2 showing filter options and search criteria.

Filter Results  
 Organism: only top 20 will appear

Search results for SARS-CoV-2 showing alignments and taxonomic information.

Alignments (Taxonomy)  
 SARS-CoV-2

Man Han Coronavirus Genomes

Accession	Score	Cover	Value	E-Value	Ident
NM_001278450.1	3081	3081	99%	0.0	84.77%
NM_001319389.1	3075	3075	99%	0.0	84.72%
NM_001081814.2	1828	1828	93%	0.0	83.06%




LitCovid DATA DOWNLOAD

Bibliography  
 Download LitCovid citations for reference management in 40 programs with automated software

Text and Data Mining  
 Download the text available from LitCovid articles, with automatic annotations by PubParser

How to cite LitCovid  
 LitCovid. In: COVID-19: A global health challenge. April 2020:1977981193

The National Library of Medicine (NLM) of the National Institutes of Health is working on multiple fronts to aid in the COVID-19 response through new initiatives with the goal of providing access to scientific papers on coronavirus for the use of NLM's PubMed Central® (PMC), a digital archive of peer-reviewed biomedical and life sciences literature. PMC currently provides access to nearly 6 million full-text journal articles.



# The National Library of Medicine: Roles Related to Computable Knowledge

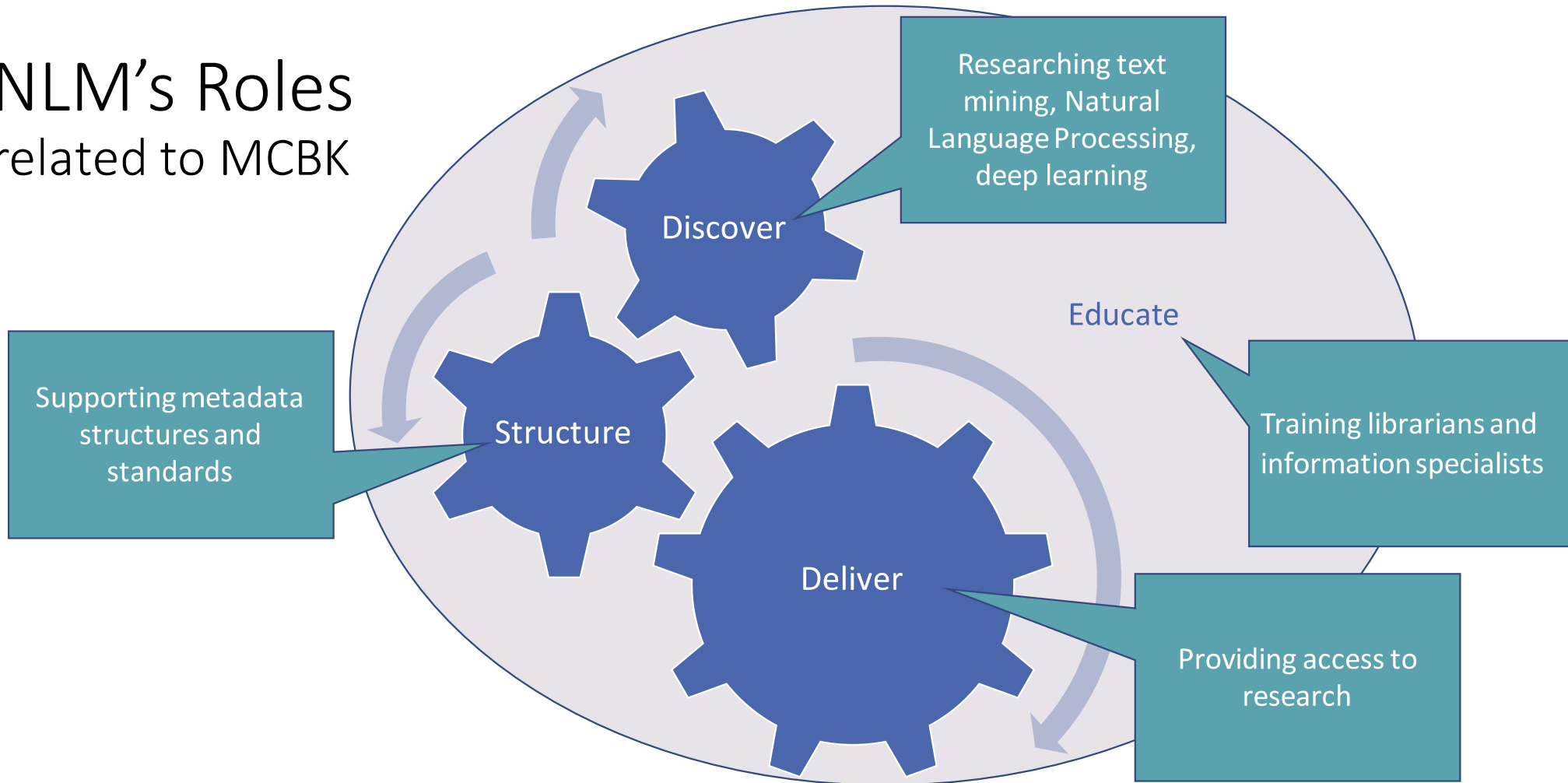


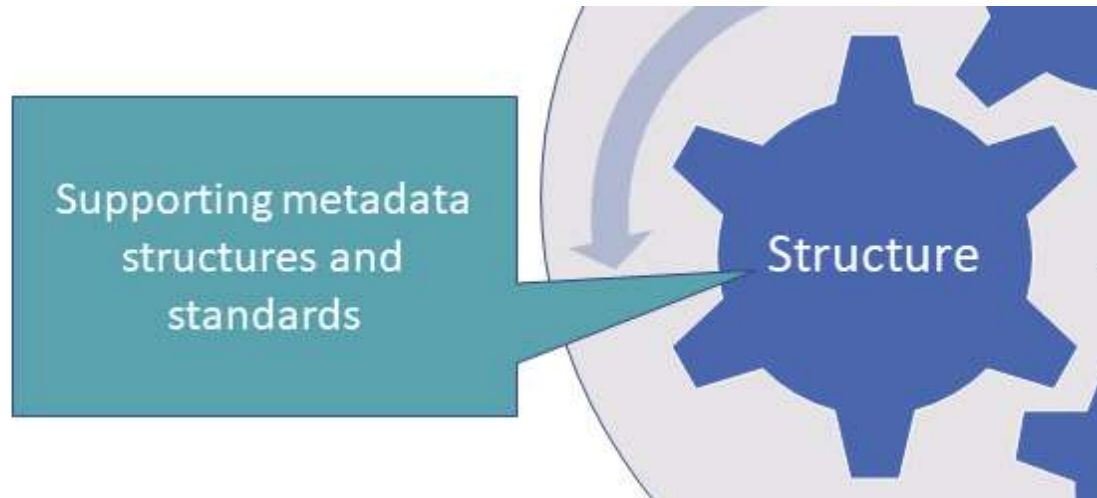
NIH NLM

Gene and exon navigator



# NLM's Roles related to MCBK

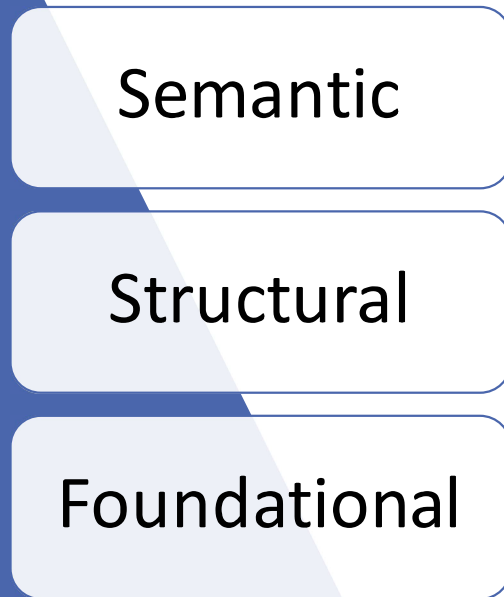




## Supporting Health Data and Metadata Standards



# Levels of Interoperability



-- Adapted from [HIMSS](#) and ONC

# Levels of Interoperability



Semantic

Structural

Foundational

- Vocabulary
  - code sets, ICD-10-CM, SNOMED CT, LOINC, RxNorm, etc.
- 2 structural interoperability projects
- Content exchange
  - HL7 (v2, v3, FHIR, CDA)

# Who makes and sets the standards?

## Organizations make them

- Health Level Seven International (HL7)
- SNOMED International, WHO, etc.

## Federal government sets them

- Congress
- Office of the National Coordinator for Health Information Technology (ONC)
- Centers for Medicare & Medicaid Services (CMS)
- NLM

<https://www.nlm.nih.gov/healthit/>



# Levels of Interoperability



Semantic

Structural

Foundational

- Vocabulary
  - code sets, ICD-10-CM, SNOMED CT, LOINC, RxNorm, etc.
- 2 structural interoperability projects
- Content exchange
  - HL7 (v2, v3, FHIR, CDA)

-- Adapted from [HIMSS](#) and ONC



# Exchange Standards

Coordinated primarily by HL7



## Document structure

Clinical Document Architecture (CDA)



## Data transfer

Version 2

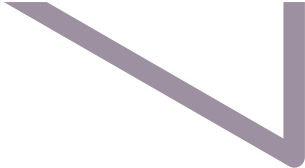

Version 3

FHIR



## Health Level Seven (HL7)

- International non-profit
- Over 1,600 members from corporations, government agencies, universities, and others
- Focused on developing standards for exchanging health information



Clinical Document  
Architecture (CDA)

Consolidated-Clinical  
Document  
Architecture (C-CDA)

- Standard that specifies the structure of "clinical documents"
- Example: "Consultation Note" and "Discharge Summary" have distinct document codes and structures
- Documents can include text, images, sounds, and other multimedia content

# HL7 data transfer standards

## V2

- Easy to use and understand
- Most widely used data exchange standard
- Works well within a single enterprise but not really interoperable generally

## V3

- Not an update of version 2!
- Broader applications but harder to use
- NLM: MedlinePlus Connect





- Fast Healthcare Interoperability Resources:  
a standard for health information exchange



- Uses common Web standards (API, HTTP, HTML, CSS, Atom)
- Describes data formats and elements as **resources** (in JSON, XML or RDF)
- All resources have:
  - A common way to define and represent them
  - A common set of metadata
  - A human readable part

# FHIR Resource Example: Patient

```
<Patient xmlns="http://hl7.org/fhir">...
```

```
<identifier><!-- 0..* Identifier An identifier for this patient --></identifier>
```

```
<active value="[boolean]"/><!-- 0..1 Whether this patient's record is in active use -->
```

```
<name><!-- 0..* HumanName A name associated with the patient --></name>
```

```
<telecom><!-- 0..* ContactPoint A contact detail for the individual --></telecom>
```

```
...
```

# Levels of Interoperability



Semantic

Structural

Foundational

- Vocabulary
  - code sets, ICD-10-CM, SNOMED CT, LOINC, RxNorm, etc.
- 2 structural interoperability projects
- Content exchange
  - HL7 (v2, v3, FHIR, CDA)





Search CDEs

▾  Search

The NIH Common Data Elements (CDE) Repository has been designed to provide access to structured human and machine-readable definitions of data elements that have been recommended or required by NIH Institutes and Centers and other organizations for use in research and for other purposes.

Follow NLM



National Library of Medicine  
8600 Rockville Pike, Bethesda, MD 20894

Copyright  
FOIA

Privacy  
Accessibility

Help  
Careers

**Common Data Element (CDE):**  
a fixed representation of a variable  
including a specific question and a  
fixed set of permissible answers

# CDE Example

- Name: Ethnicity OMB.1997
- Definition: The ethnicity of a person based on the Office of Management and Budget (OMB), Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity (Oct. 20, 1997)

## Permissible Value

Value	Code Name	Code	Code System	Code Description
Hispanic or Latino	Hispanic or Latino	2135-2	CDCREC	N/A
Not Hispanic or Latino	Not Hispanic or Latino	2186-5	CDCREC	N/A

## Types of CDEs

Demographic

Disease-agnostic

Disease-specific



# Form with CDEs

An official website of the United States government [Here's how you know](#) ▼

NIH National Library of Medicine

NIH CDE Repository

CDEs Forms Boards ▼ About Help ▼

## Form: Adult Sickle Cell Quality of Life Measurement Information System (ASCQ-Me) v2.0 Pain Episode Frequency and Severity Measure

Export ▼ Add to Quick Board Add to Board Pin CDEs CDEs Discuss

### Preview

Options ▼

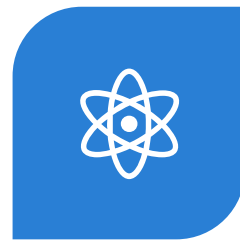
- 1. In the past 12 months, how many sickle cell pain attacks (crises) days did you have?**
  - I did not have a pain attack (crises) in the past 12 months
  - 1
  - 2
  - 3
  - 4 or more
- 2. When was your last pain attack (crisis)?**
  - I've never had a pain attack (crises)
  - More than 5 years ago
  - 1-5 years ago
  - 7-11 months ago
  - 1-6 months ago
  - 1-3 weeks ago
  - Less than a week ago
  - I have one right now
- 3. Using any number from 0 to 10, where 0 is no pain and 10 is the worst pain imaginable, how severe was your pain during your last pain attack (crisis)?**
  - I've never had a pain attack (crises)
  - No Pain
  - 1
  - 2
  - 3
  - 4
  - 5
  - 6
  - 7
  - 8
  - 9
  - Worst pain imaginable
- 4. How much did your last pain attack (crisis) interfere with your life?**
  - I've never had a pain attack (crises)
  - Not at all, I did everything I usually do
  - I had to cut down on some things I usually do
  - I could not do most things I usually do

Form  
Preview  
General Details  
Naming  
Definitions  
Classification  
Display Profiles  
Reference Documents  
Properties  
Identifiers  
Attachments  
History

Use CDEs to. . .



SAVE  
RESOURCES



ADVANCE  
SCIENCE



FOLLOW  
POLICIES

<https://go.usa.gov/xtCQZ>

Common Data Elements: Standardizing Data Collection

Training & Outreach > Learning Resources

Course Home

- Why the Need for Common Data Elements? >
- FAIR Data >
- Common Data Elements (CDEs) >
- NLM and Common Data Elements >
- Closing

Tools and Resources

How to Navigate This Course

## Common Data Elements: Standardizing Data Collection



Have you written a survey or designed a protocol for a research study? If so, you are familiar with the challenges of gathering consistent and standardized data. This course from the National Library of Medicine® is about making data collection and data sharing easier using common data elements (CDEs).

This course includes:

 Scenarios


 Exercises

[Click for Learning Objectives](#)

### Continuing Education Credits

If you are taking this course for Medical Library Association credits, please [register and complete the class](#) through the

# Journal Article Tag Suite (JATS)



← → ↻ 🔒 https://jats.nlm.nih.gov 🔍 🏠 ☆ 📄 🛠️ 🌐 ⋮

NCBI

## Journal Article Tag Suite

... is an application of NISO Z39.96-2019, which defines a set of XML elements and attributes for tagging journal articles and describes three article models. The content on this site is the supporting documentation for the standard. JATS is a continuation of the NLM Archiving and Interchange DTD work begun in 2002 by NCBI

### About the JATS

- [JATS Wiki](#)
- [JATS and the NLM DTDs](#)
- [FAQ](#)
- [NISO z39.96-2021 \(JATS 1.3; Current Standard\)](#)
- [Getting the files](#)

### JATS-Con

- [JATS-Con home](#)
- [JATS-Con Proceedings](#)

### JATS Extensions

- [Book Interchange Tag Suite \(BITS\)](#)
- [NISO Standards Tag Suite \(NISO-STS\)](#)

### Tag Sets

- [Journal Archiving and Interchange](#)  
The most permissive of the Tag Sets.
- [Journal Publishing](#)  
A moderately prescriptive Tag Set.
- [Article Authoring](#)  
The most prescriptive of the Tag Sets.

# Levels of Interoperability



Semantic

Structural

Foundational

- Vocabulary
  - code sets, ICD-10-CM, SNOMED CT, LOINC, RxNorm, etc.
- 2 structural interoperability projects
- Content exchange
  - HL7 (v2, v3, FHIR, CDA)

# What is a medical terminology?

A set of specialized terms that facilitate precise communication by minimizing or eliminating ambiguity.

## Features of a terminology

- Unique Identifier (often a code)
- Official Name
- Synonyms (in many cases)

## Examples

- CPT, ICD-10, SNOMED-CT, LOINC, RxNorm



## Why do we need medical terminologies?

- To facilitate precise and clear communication by minimizing or eliminating ambiguity.
- To cross-walk or map between other terminologies
  - Examples: UMLS and RxNorm



## Some terminology problems for biomedical researchers:

- There are many biomedical terminologies in use.
- Terminologies come in a variety of formats and a variety of data models.
- Terminologies do not necessarily link to each other.
- Terminologies are not always freely available.





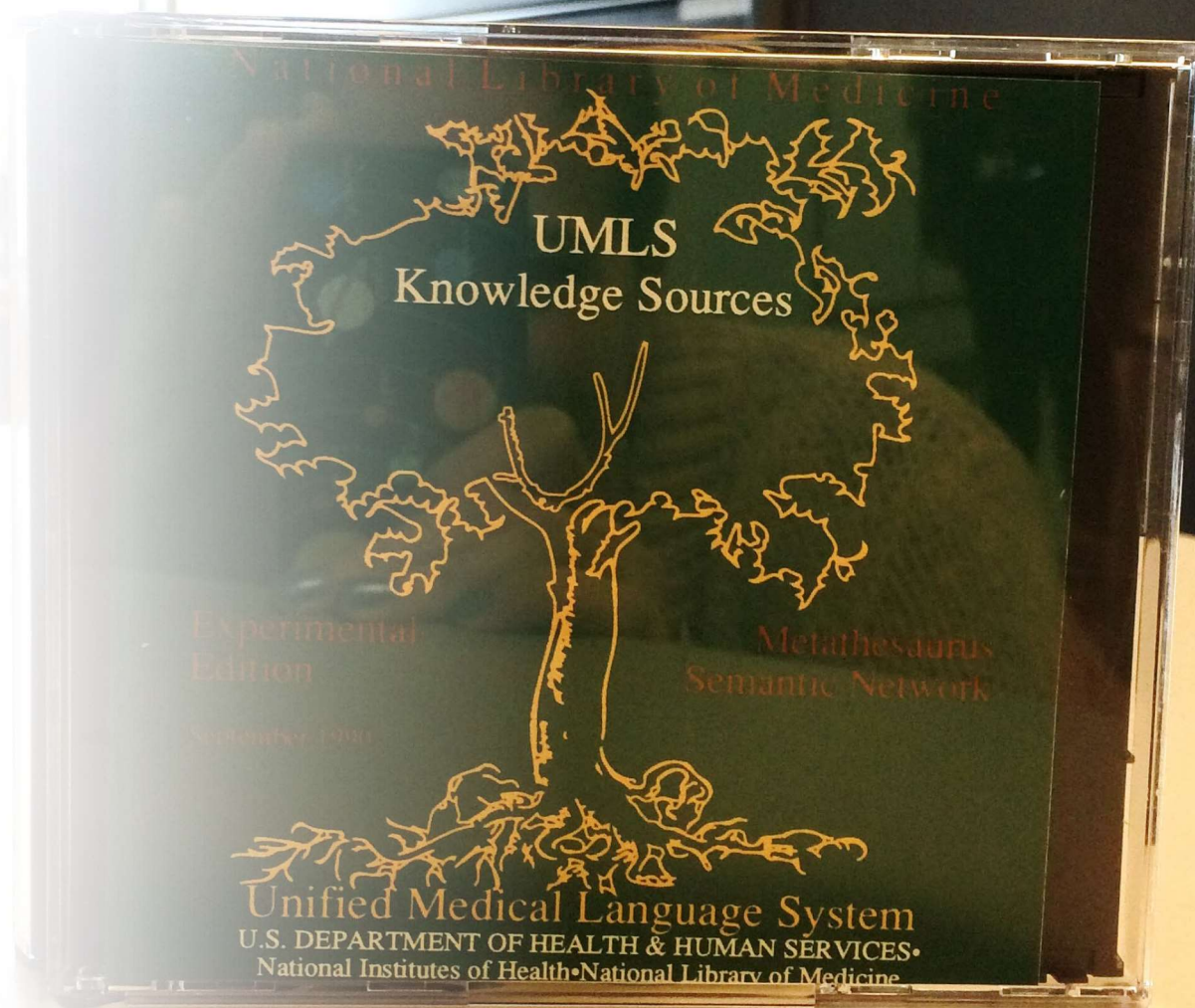
**Solution:** Unify the medical language in one system.

The National Library of Medicine has accomplished this with the **Unified Medical Language System (UMLS)**



# What is the UMLS?

Unified  
Medical  
Language  
System



# The UMLS provides:

Access to  
Terminology  
Data

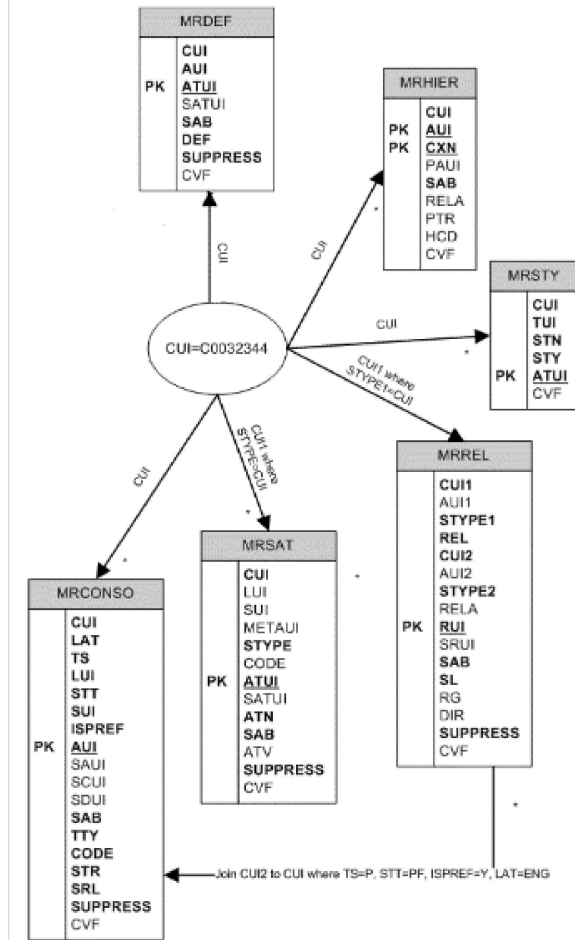
A Common  
Data Model for  
Terminologies

Interoperability  
through  
Synonymy

# The UMLS: A Common Data Model

**Problem:** Terminology data is represented in a variety of formats according to a variety of data models.

**Solution:** The UMLS represents all terminology data according to a standard data model.



# Interoperability through Synonymy

**Problem:** Terminologies are not always linked to each other.

**Solution:** UMLS asserts synonymy by grouping names from different terminologies into concepts. This can be used as a starting point for crosswalking from one terminology to another.

# A UMLS Concept: Addison Disease

Name	Source Terminology	Code	Atom Identifier	Concept Identifier
Addison Disease	MeSH	D000224	A6954527	C0001403
Addison's disease	ICD-10-CM	E27.1	A17799651	C0001403
Addison's Disease	MeSH	D000224	A26597849	C0001403
Primary adrenal insufficiency	MedDRA	S2164152	A2018590	C0001403
Primary adrenocortical insufficiency	ICD-10-CM	E27.1	A17786892	C0001403
Insufficiency, Primary Adrenocortical	MeSH	D000224	A6970512	C0001403
Primary adrenocortical insufficiency (disorder)	SNOMED CT	373662000	A3644299	C0001403
Primary hypoadrenalism	MedDRA	S0718109	A25720215	C0001403
Primary hypoadrenalism	SNOMED CT	373662000	A3060485	C0001403
Enfermedad de Addison	MeSH Spanish	D000224	A9175691	C0001403

# UMLS Case Study: Cleveland Clinic

**Problem:** How do you make patient data ready for researchers to discover and use?

**Solution:** Integrate identifiers, hierarchies, and relationships from UMLS.



Milinovich A, Kattan MW. Extracting and utilizing electronic health data from Epic for research. *Ann Transl Med.* 2018;6(3):42.

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5879514/>

# UMLS Case Study: Cleveland Clinic

- Cleveland Clinic converts electronic health record data from 185 tables to 18 research-ready tables annotated with UMLS identifiers. Data is updated on a weekly basis.
- **Result:** “Cleveland Clinic can do live population exploration as well as produce datasets for analysis faster than it takes most organizations to simply identify their base population.”



Milnovich A, Kattan MW. Extracting and utilizing electronic health data from Epic for research. *Ann Transl Med.* 2018;6(3):42.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5879514/>



# Identifying Patient Populations

Diabetic patients on a GLP-1 medication with an HbA1c >10

## 2,666 different “diabetic” diagnoses

Get all descendants of:

**Diabetes Mellitus (UMLS CUI: C0011847)**

- Acidosis due to type 1 diabetes mellitus
- Acidosis due to type 2 diabetes mellitus
- Acute complication with diabetes mellitus
- Diabetic dyslipidemia associated with type 2 diabetes mellitus
- Diabetic hyperosmolar non-ketotic state
- Diabetic lumbosacral radiculoplexus neuropathy
- Diabetic mastopathy
- Diabetic severe hyperglycemia
- Disorder of nerve co-occurrent and due to type 1 diabetes mellitus
- Dyslipidemia due to type 1 diabetes mellitus
- Hyperglycemia due to type 1 diabetes mellitus
- Hyperglycemia due to type 2 diabetes mellitus
- Hyperglycemic crisis in diabetes mellitus
- Hyperlipidemia due to type 1 diabetes mellitus
- Hyperlipidemia due to type 2 diabetes mellitus
- Hyperosmolality due to uncontrolled type 1 diabetes mellitus
- Hyperosmolar coma associated with diabetes mellitus
- Hyperosmolarity co-occurrent and due to drug induced diabetes mellitus
- Hypoglycemia due to type 1 diabetes mellitus
- Hypoglycemia due to type 2 diabetes mellitus
- Hypoglycemic coma in diabetes mellitus
- Hypoglycemic state in diabetes
- Ketoacidosis due to secondary diabetes mellitus
- Lactic acidosis with diabetes mellitus
- Malnutrition-related diabetes mellitus with multiple complications
- Metabolic acidosis with diabetes mellitus
- Mixed hyperlipidemia due to type 1 diabetes mellitus
- Mixed hyperlipidemia due to type 2 diabetes mellitus
- Multiple complications due to diabetes mellitus
- Peripheral neuropathy due to type 1 diabetes mellitus
- Radiculoplexoneuropathy due to diabetes mellitus
- Type 1 diabetes mellitus with hyperosmolar coma
- Type 2 diabetes mellitus with hyperosmolar coma
- Diabetic acute painful polyneuropathy
- Coma associated with malnutrition-related diabetes mellitus
- Diabetic coma with ketoacidosis
- Hyperosmolar coma associated with diabetes mellitus
- Hypoglycemic coma co-occurrent and due to diabetes mellitus type II
- ETC

# Identifying Patient Populations

Diabetic patients on a GLP-1 medication with an HbA1c >10

## 69 different GLP-1 medications

Get all medications associated with a drug class:

**Glucagon-like Peptide-1 (GLP-1) Agonists [MoA] (UMLS CUI: C2916791)**

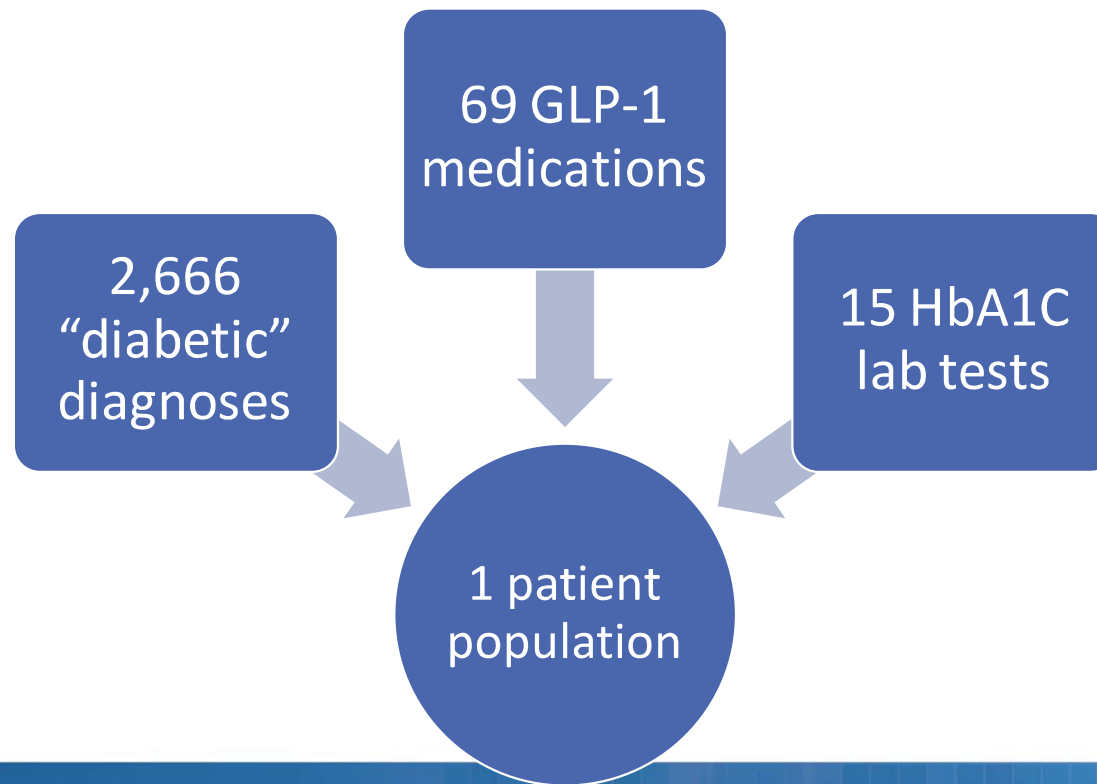
## 15 different HbA1c lab tests

Map lab tests in EHR to:

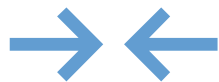
**Hemoglobin A1c/Hemoglobin.total:Mass Fraction:Point in time:Whole blood:Quantitative (UMLS CUI: C0366781)**

# Identifying Patient Populations

Diabetic patients on a GLP-1 medication with an HbA1c >10



# Synonymy Applied



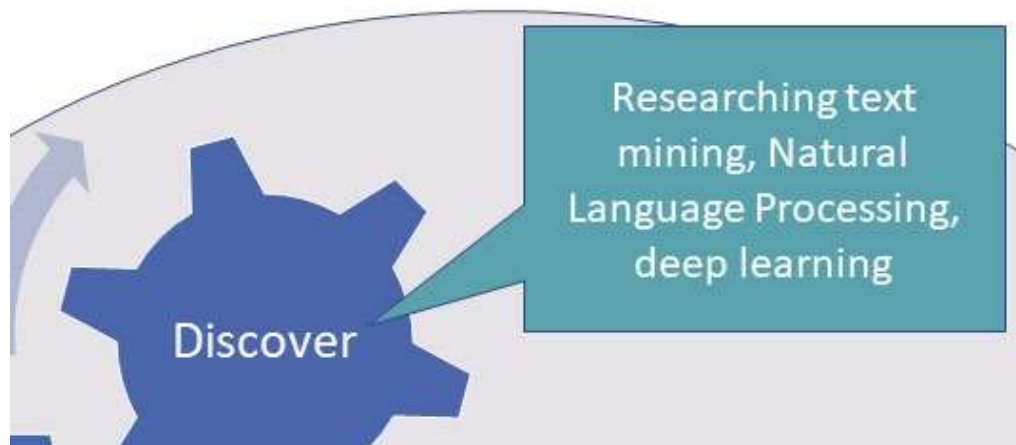
Crosswalking between  
Terminologies



Search and Retrieval

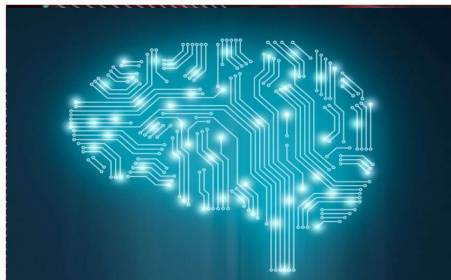


Identifying Meaning in Text



## Researching Methods in Informatics

# NLM Research Portfolio



Machine Learning Applied to Electronic Health Record Data



Health Information Standards and Discovery

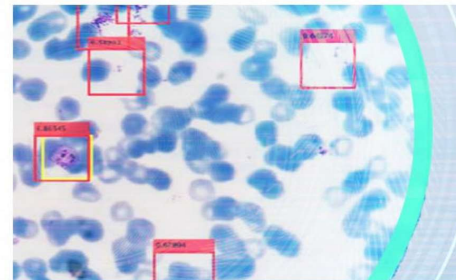


Image Processing



Natural Language Processing



Statistical Methods



Networks, Gene Regulation and Chromatin

# Identifying Meaning in Text (Natural Language Processing)

**Problem:** Medical data in the real world is often unstructured. Examples: clinical notes or the biomedical literature / abstracts.

**Solution:** The UMLS can help identify meaning in text in combination with various tools.



## Why Identify Meaning in Text?

- Improve search and retrieval by annotating records in a research database
- Find co-occurrences of concepts in text
- Annotate clinical text on the fly
- Identify a patient population





# Text Processing Tools

## NLM Tools

- **MetaMap** – A tool for recognizing UMLS concepts in text
- **Medical Text Indexer** – A tool for automated indexing of the biomedical literature
- **MeSH on Demand** – Uses the Medical Text Indexer to extract MeSH terms via UMLS Concepts.
- **SemRep** – A tool for extracting assertions from sentences in biomedical text

## Third-Party Tools that include UMLS

- **Apache cTakes** – “...a natural language processing system for extraction of information from electronic medical record clinical free-text... Originally developed by the Mayo Clinic...”
- **CLAMP (Clinical Language Annotation, Modeling, and Processing Toolkit)** – “...a comprehensive clinical Natural Language Processing (NLP) software that enables recognition and automatic encoding of clinical information in narrative patient reports.” (University of Texas Health Science Center at Houston)
- **And many many more.**

# NLM MeSH on Demand

Search    Reset    Help/FAQ    Features

Start PubMed Search

Export Data

**MeSH Terms**

- Humans
- Liraglutide
- Glucagon-Like Peptide 1
- Diabetes Mellitus, Type 2
- Incidence
- Double-Blind Method
- Cardiovascular Diseases
- Confidence Intervals
- Follow-Up Studies
- Random Allocation
- Risk Factors
- Myocardial Infarction
- Heart Failure
- Stroke
- Hospitalization
- National Institutes of Health (U.S.)
- Pancreatitis

**Additional Terms**

- United States

BACKGROUND The cardiovascular effect of **liraglutide** a **glucagon like peptide** 1 analogue when added to standard care in **patients** with **type 2 diabetes** remains unknown METHODS In this **double blind** trial we randomly assigned patients with **type 2 diabetes** and high **cardiovascular risk** to receive **liraglutide** or placebo The primary composite outcome in the time to event analysis was the first occurrence of death from cardiovascular causes nonfatal **myocardial infarction** or nonfatal **stroke** The primary hypothesis was that **liraglutide** would be noninferior to placebo with regard to the primary outcome with a margin of 1.30 for the upper boundary of the 95 **confidence interval** of the hazard ratio No adjustments for multiplicity were performed for the prespecified exploratory outcomes RESULTS A total of 9340 patients underwent **randomization** The median **follow up** was 3.8 years The primary outcome occurred in significantly fewer patients in the **liraglutide** group 608 of 4668 patients [13.0] than in the placebo group 694 of 4672 [14.9] hazard ratio 0.87 id=CONFIDENCE\_INTERVALS\_3769 tabindex=0' class='highlightedModTerm modHighlight' data-name='CONFIDENCE\_INTERVALS'>5 **confidence interval** [CI] 0.78 to 0.97 P 0.001 for noninferiority P 0.01 for superiority Fewer patients died from cardiovascular causes in **the liraglutide** group 219 patients [4.7] than in the placebo group 278 [6.0] hazard ratio 0.78 95 CI 0.66 to 0.93 P 0.007 The rate of death from any cause was lower id=LIRAGLUTIDE\_5323 tabindex=0' class='highlightedModTerm modHighlight' data-name='LIRAGLUTIDE'>n **the liraglutide** group 381 patients [8.2] than in the placebo group 447 [9.6] hazard ratio 0.85 95 CI 0.74 to 0.97 P 0.02 The rates of id=MYOCARDIAL\_INFARCTION\_6147 tabindex=0' class='highlightedModTerm modHighlight' data-name='MYOCARDIAL\_INFARCTION'>onfatal **myocardial infarction** id=STROKE\_6223 tabindex=0' class='highlightedModTerm modHighlight' data-name='STROKE'>onfatal **st** id=HOSPITALIZATION\_6249 tabindex=0' class='highlightedModTerm modHighlight' data-name='HOSPITALIZATION'>oke and **hospitaliza** id=HEART\_FAILURE\_6299 tabindex=0' class='highlightedModTerm modHighlight' data-name='HEART\_FAILURE'>ion **for heart failure** were nonsignificantly lower **in the liraglutide** group than in the placebo group The most common adverse events leading to the discontinu id=LIRAGLUTIDE\_6817 tabindex=0' class='highlightedModTerm modHighlight' data-name='LIRAGLUTIDE'>tion **of liraglutide** were gastrointestinal ev id=INCIDENCE\_6937 tabindex=0' class='highlightedModTerm

<https://meshb.nlm.nih.gov/MeSHonDemand>

# CLAMP

(Clinical Language Annotation, Modeling, and Processing Toolkit)

Results.xml    Results.txt

41 CONDITIONS ON DISCHARGE: The patient had appropriate mood and was not engaging in self-injurious behavior. He denied suicidal or homicidal ideation.

43 Height: 5 foot 8 inches. Weight: 134. Blood pressure: 120/54. Pulse: 104. Respirations: 18. Temperature: 99.

45 PROGNOSIS: Guarded.

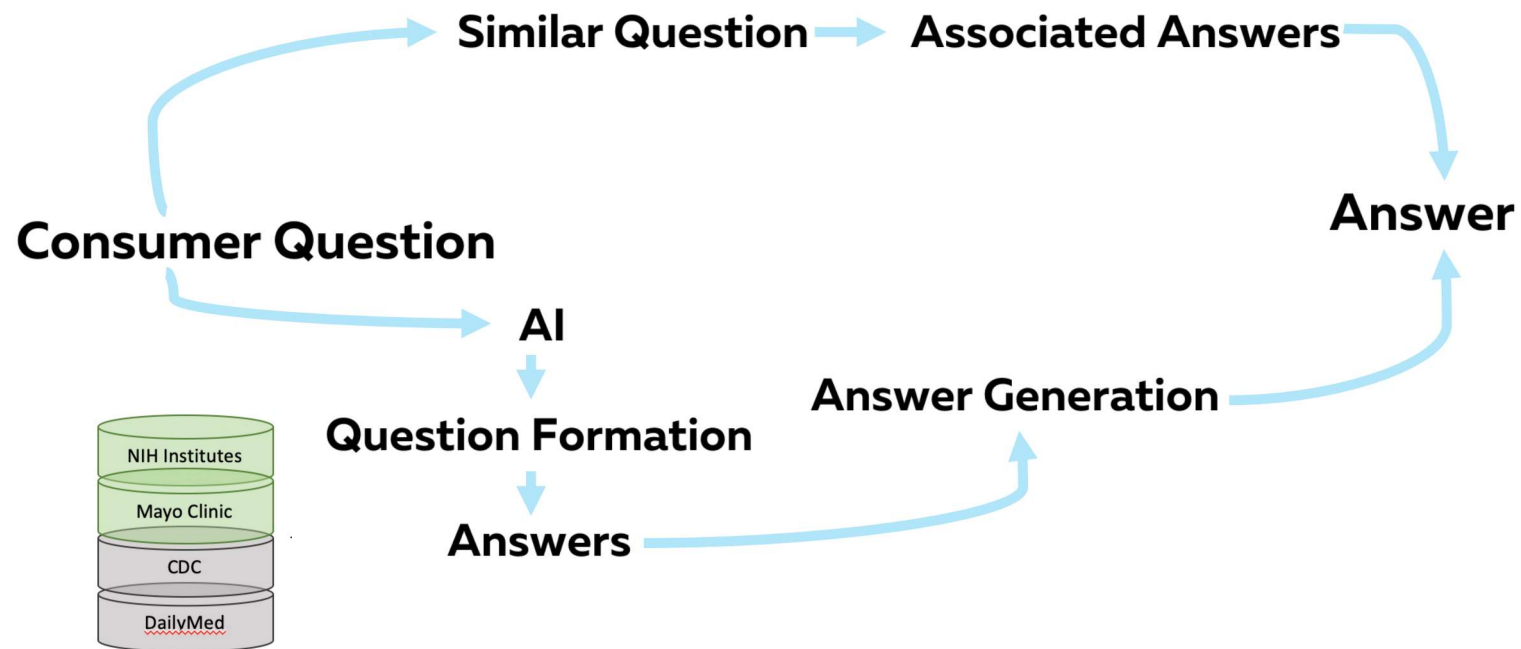
47 DISCHARGE PLAN: As recommended by the treatment team, the patient was discharged to an RTC (Residential Treatment Center) Program to North Star RTC. He was transferred by personal staff of that institution.

49 DISCHARGE INSTRUCTIONS/MEDICATIONS: The patient is to continue treatment at North Star RTC. Discharge medications are Adderall XR 30 mg p.o. a.m., Depakote 250 mg p.o. t.i.d., Zoloft 55 mg p.o.

<https://clamp.uth.edu/clampdemo.php>

# AI-Enabled Access to Reliable Health Information

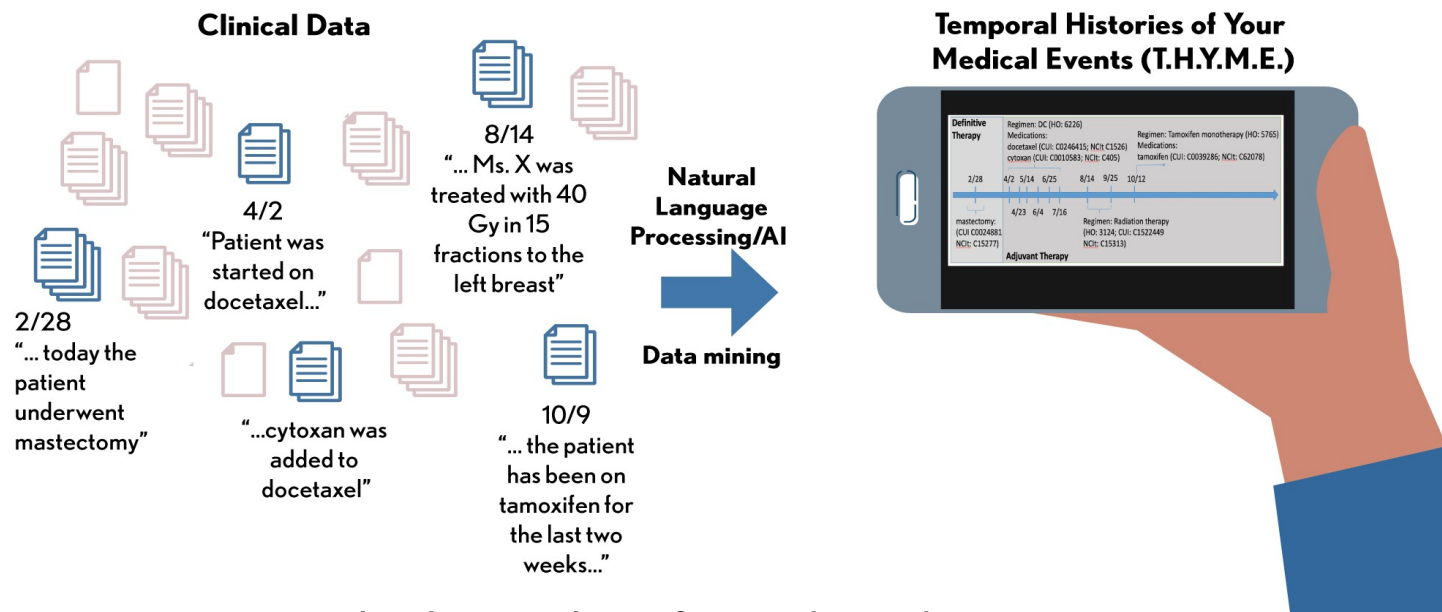
## NLM's Consumer Health Question Answering Program



Consumer Health Information and Question Answering: Helping consumers find answers to their health-related information needs. Demner-Fushman, D., Mrabet, Y., and Ben Abacha, A. JAMIA 2020.

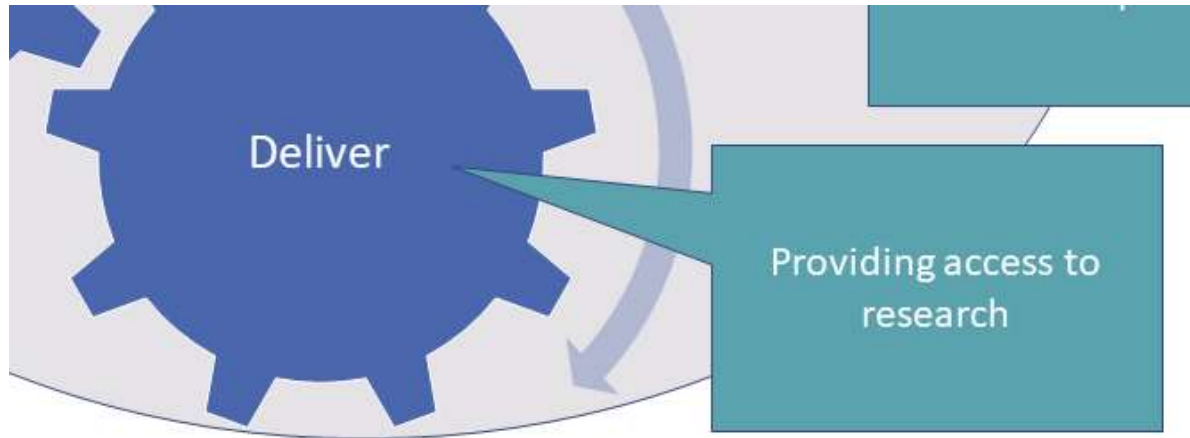
# Discovery: Natural Language Processing

## Temporal Histories of Your Medical Events (T.H.Y.M.E) Project



Extract temporal relationships from clinical text

Guergana Savova, PhD [U54LM008748, R01LM010090]



Providing Access to Research

# Research Data Services





# Data Discovery

Access, explore, and build with datasets and APIs  
from the National Library of Medicine



## Full Catalog

Browse all datasets and APIs in the Data Discovery catalog



## Featured: PubMed Citations

A baseline set of MEDLINE/PubMed citation records in XML format for download on an annual and daily basis.



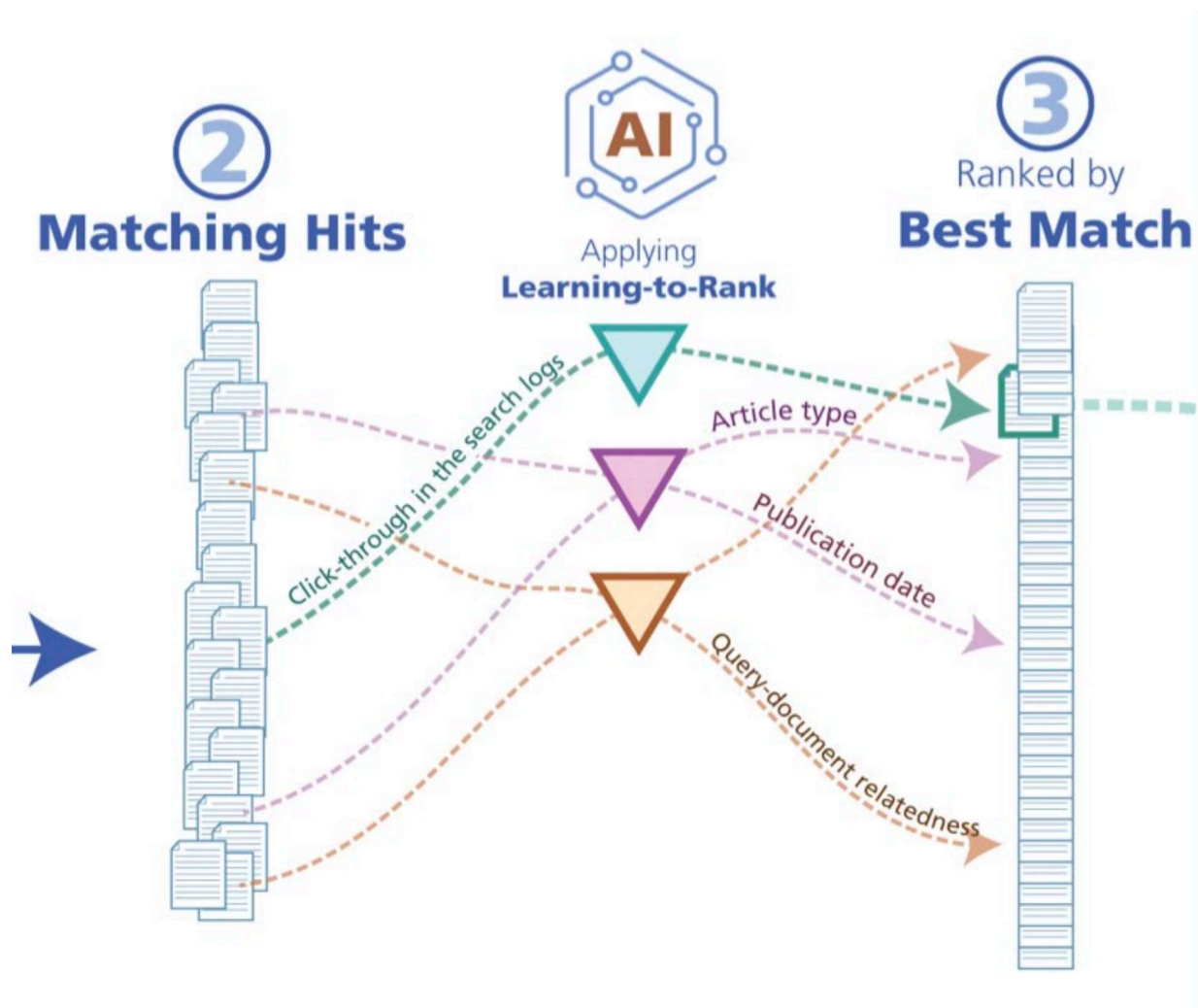
## Featured: FDA Notices of Judgment Collection

A digital archive of the published federal notices of judgment for manufacturers and products prosecuted under authority of the 1906 Pure Food

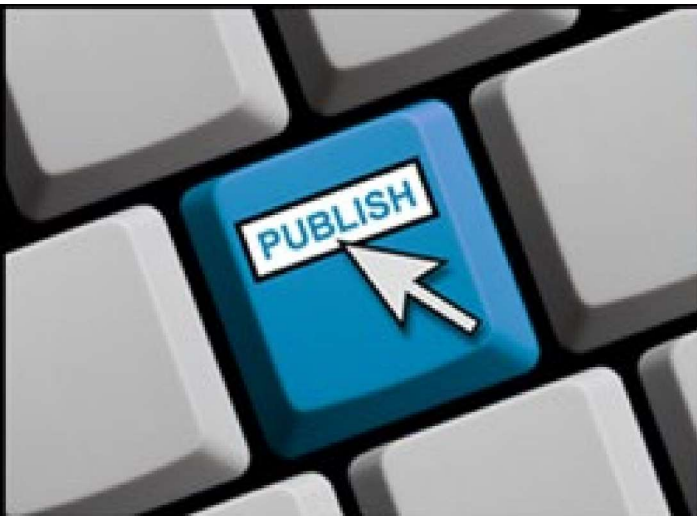


PubMed





PubMed's Best Match Sorting



**A 24-hour response time**  
for MeSH-indexed citations  
to appear in PubMed

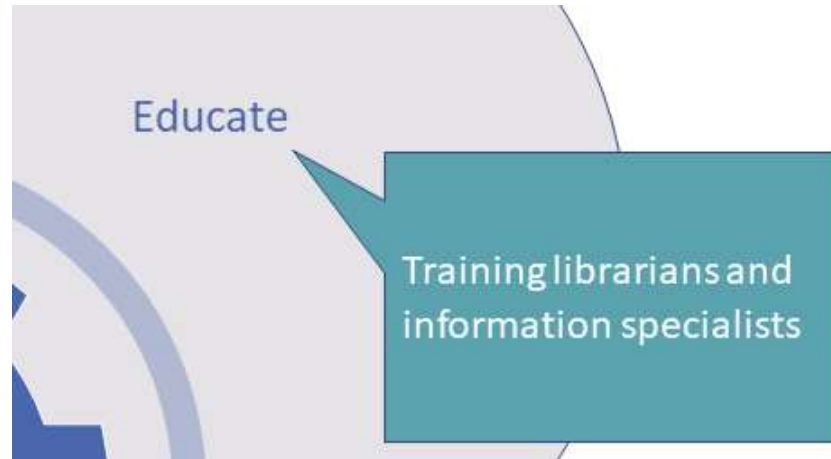


**Expanded chemical and  
gene curation** by subject  
matter experts



**Continuous improvement**  
of the automated indexing  
algorithm

MEDLINE Automated Indexing |



## Training Information Specialists



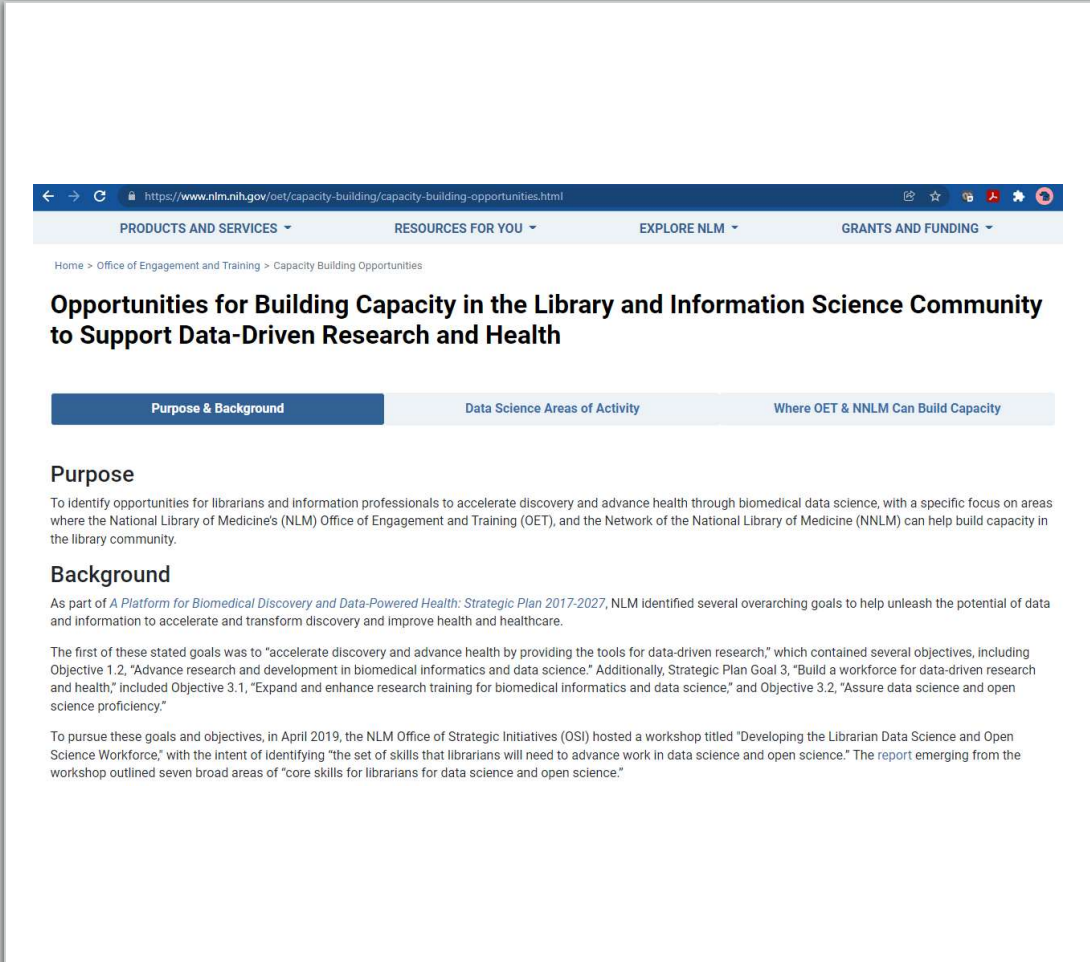
# Facing New Challenges



*Establishing the informatics-informed workforce*

# Building Capacity in Data Science




<https://go.usa.gov/xt3vw>




The screenshot shows a web browser window with the URL <https://www.nlm.nih.gov/oet/capacity-building/capacity-building-opportunities.html>. The page has a navigation bar with links for 'PRODUCTS AND SERVICES', 'RESOURCES FOR YOU', 'EXPLORE NLM', and 'GRANTS AND FUNDING'. Below the navigation bar is a breadcrumb trail: 'Home > Office of Engagement and Training > Capacity Building Opportunities'. The main heading is 'Opportunities for Building Capacity in the Library and Information Science Community to Support Data-Driven Research and Health'. There are three tabs: 'Purpose & Background' (selected), 'Data Science Areas of Activity', and 'Where OET & NNLM Can Build Capacity'. The 'Purpose' section states: 'To identify opportunities for librarians and information professionals to accelerate discovery and advance health through biomedical data science, with a specific focus on areas where the National Library of Medicine's (NLM) Office of Engagement and Training (OET), and the Network of the National Library of Medicine (NNLM) can help build capacity in the library community.' The 'Background' section states: 'As part of *A Platform for Biomedical Discovery and Data-Powered Health: Strategic Plan 2017-2027*, NLM identified several overarching goals to help unleash the potential of data and information to accelerate and transform discovery and improve health and healthcare. The first of these stated goals was to "accelerate discovery and advance health by providing the tools for data-driven research," which contained several objectives, including Objective 1.2, "Advance research and development in biomedical informatics and data science." Additionally, Strategic Plan Goal 3, "Build a workforce for data-driven research and health," included Objective 3.1, "Expand and enhance research training for biomedical informatics and data science," and Objective 3.2, "Assure data science and open science proficiency." To pursue these goals and objectives, in April 2019, the NLM Office of Strategic Initiatives (OSI) hosted a workshop titled "Developing the Librarian Data Science and Open Science Workforce," with the intent of identifying "the set of skills that librarians will need to advance work in data science and open science." The report emerging from the workshop outlined seven broad areas of "core skills for librarians for data science and open science."





# NNLM National Center for Data Services


CENTER: **NNLM National Center for Data Services**   

The NNLM National Center for Data Services (NCDS) coordinates with other NNLM Regions, Offices, and Centers to provide training and resources to increase data science capacity among information professionals, in alignment with the NLM strategic plan.

 **Staff Directory**  
Looking for assistance? Browse our staff directory to find contact information, and get to know the NCDS staff by reading their personal bios.

 **Funding Opportunities**  
Browse open funding opportunities that are available to NNLM members.

 **Guides & Resources**  
Browse our curated collection of guides and resources that support data-driven research.

 **Classes & Training**  
Browse NNLM's upcoming training opportunities, course catalog, and recordings of past webinars.

NYU Health Sciences Library | 577 1st Ave. | New York, New York | 10016

Kick-off tomorrow at 1pm ET: <https://go.usa.gov/xtCEa>

# Training for Information Specialists

The screenshot shows a web browser window with the URL <https://learn.nlm.nih.gov/documentation/training-packets/T000181112/>. The page header includes the NIH logo and the text "National Library of Medicine" on the left, and "Learning Resources Database" on the right. Below the header is a navigation bar with four menu items: "PRODUCTS AND SERVICES", "RESOURCES FOR YOU", "EXPLORE NLM", and "GRANTS AND FUNDING". The main content area starts with a breadcrumb trail: "Home > Training & Outreach > Learning Resources". The title is "Training on Biomedical Informatics, Data Science, and Data Management". Below the title is a paragraph: "The National Library of Medicine offers a range of training on biomedical data and informatics." This is followed by a list of five training topics in rounded rectangular boxes: "NLM Data Types, Structure, Sources and Tools", "Clinical Data Management", "Biomedical Research Data Management", "Related Topics", and "Academic Programs". Below the list is a section titled "See Also" with a single bullet point: "• NLM Grant Programs".

<https://go.usa.gov/xt3p8>



# Connect with Us

**NLM Homepage**

[nlm.nih.gov](http://nlm.nih.gov)

**NLM Technical Bulletin**

<https://go.usa.gov/xt3vZ>



**@NLM\_news**

