

NIH Clinical Informatics Program

Overview of Common Data Elements (CDE) Activities

Steve Tsang, PhD

Clinical Informatics Lead

NIH Office of Data Science Strategy

Prepared for Environmental Health Language Collaborative, 11/22 at 1pm

Agenda

- Background
 - NIH Office of Data Science Strategy (ODSS)
 - ODSS Clinical Informatics Unit
- Common Data Element (CDE) activities
- Relevant takeaways from American Medical Informatics Association (AMIA) Annual Symposium

NIH Office of Data Science Strategy (ODSS)

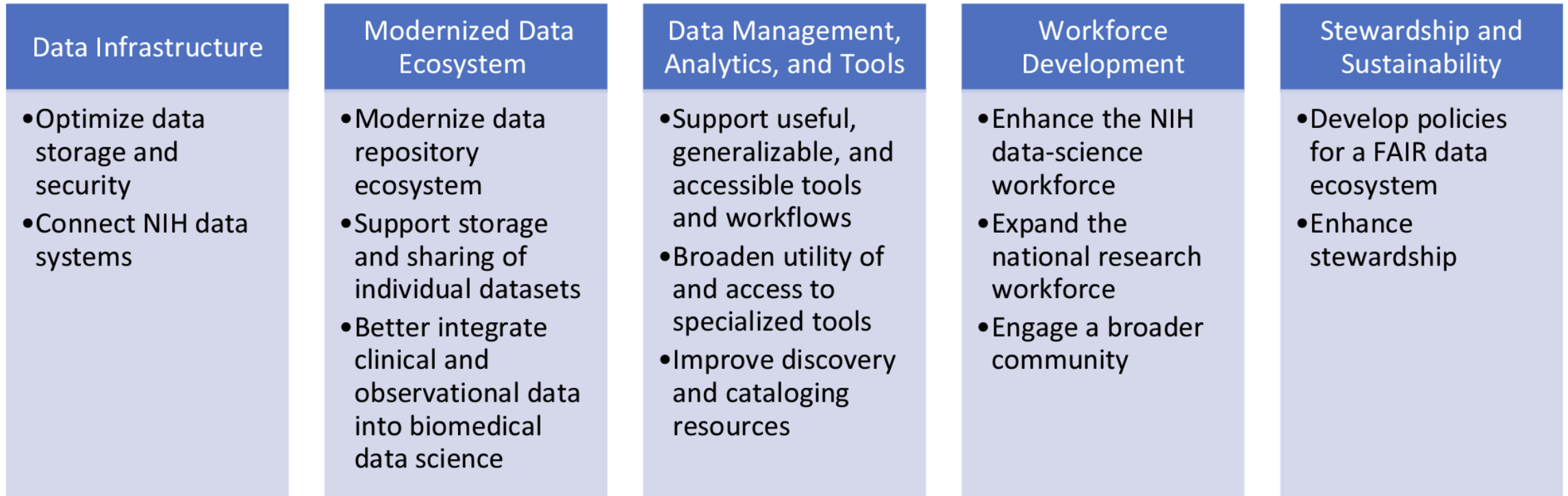


Figure 2. NIH Strategic Plan for Data Science: Overview of Goals and Objectives

ODSS Clinical Informatics Unit

- Promote use and implementation of clinical exchange and data standards (FHIR/CDE)
- Build research capacity
- Facilitate data sharing
- Share clinical informatics trainings and resources



<https://datascience.nih.gov/clinical-informatics>

NIH FHIR® for Research Training

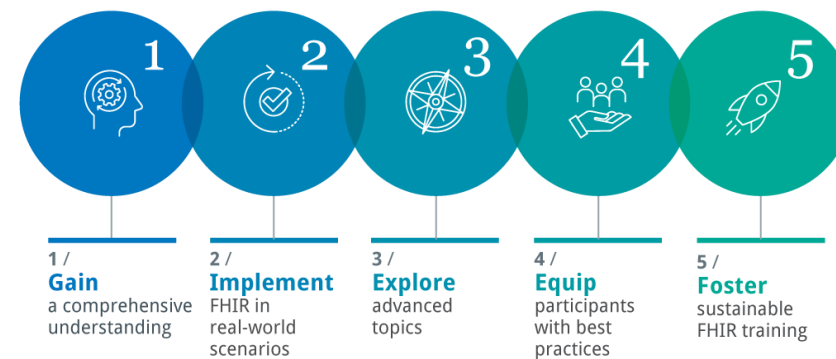
Goal: Develop a set of accessible, modular, and reusable training activities

Capability assessment on the current landscape of FHIR®, potential audience and use cases

Module-based curriculum including self-directed resources, an introductory webinar, and hands-on workshops

Train-the-trainer session to broaden the reach of this FHIR® training in the extramural community

FHIR Learning Objectives



Introduction to FHIR for Research

Three-hour introductory training course on fundamental FHIR concepts



Modular FHIR Training¹

Tailored FHIR training curriculum for research roles: Investigator, Research Leader, Informaticist, Software Engineer, and Clinician Scientist



Train-the-Trainer²

Recommendations and resources for technical leaders conducting customized FHIR training or researchers seeking deeper insights

Background – Congressional Appropriations Language

In FY23, congressional appropriations language directed NIH and ODSS to “...*encourage development and use of CDEs in disease areas where they currently do not exist*, the agreement directs ODSS to work with ICs to support efforts to develop CDEs, including through collaborations with research stakeholders. The agreement also directs ODSS *to provide a list of diseases and disease areas actively under development, such as immune and immune-mediated conditions*, to inform further NIH efforts to support development of such elements.”

<https://www.congress.gov/congressional-record/volume-168/issue-198/senate-section/article/S8553-2>

What is a CDE?

Definition:

- **Question or field** [what] and its **allowable responses** [how]
- **Used *systematically across different*** sites, studies, or clinical trials [common]
- Helps **ensure consistent data collection**

Slide courtesy of Denise Warzel

**Basic Definition Adapted from [NLM CDE Repository “Definition of CDE”](#)*

What is unique about CDEs?

Basic Definition Plus Deeper Characteristics and Benefits:

1. **Standard Terminology Concepts** → unambiguous, shared, and computable meaning
2. **Standardized Structure** → machine computability
3. **Independent Semantics** → reusable across physical data models, forms, datasets and supports *different allowable responses across the same CDE meaning*
4. **Persistent Unique Identifier** → identifiable, outside specific data collection systems
5. **Supports FAIR data** → rich metadata, web accessible (Findable, Accessible, Interoperable, Reusable)

Slide courtesy of Denise Warzel

Similar Concepts

Fast Healthcare Interoperability Resources (FHIR®)

- Healthcare **data exchange standard** via an API framework
- Focuses on how healthcare information can be exchanged among systems, **independent of internal data models/standards**

United States Core Data for Interoperability (USCDI)

- Defines data **classes and elements** for **health** information exchange
- Focuses on the content being exchanged and is **standards-agnostic**

Common Data Models (CDM)

- Example: [Observational Medical Outcomes Partnership](#) (OMOP)
- **Standardizes** and organizes the **structure and content** of observational data
- Focuses on evidence-based analyses primarily in research

All have:



Persistent Unique Id.



Supports FAIR data

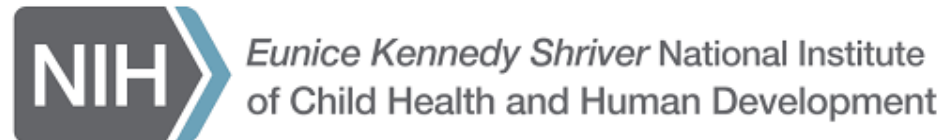
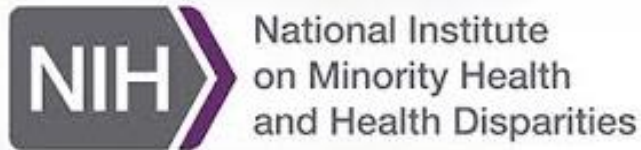
But only CDEs have:

**Standard Terminology
Concepts**

Standardized Structure

Independent Semantics

FY24 ODSS CDE Program



16 ICO Projects

Addiction and pain

Brain and neurological diseases

Concomitant immunosuppressive medications

Food allergy

Heart diseases

Myositis

Inflammatory bowel diseases

Kidney disease

Pediatric critical care

Type-1 diabetes



CDE Development in Autoimmune Diseases - Myositis

- Idiopathic inflammatory myopathies or myositis syndromes are rare autoimmune diseases, mostly affecting females, characterized by chronic muscle inflammation and high morbidity
- There is currently **no NIH-endorsed myositis-specific CDEs** in the NIH CDE Repository
- ODSS is **collaborating with NIH Office of Autoimmune Disease Research (OADR-ORWH), NLM and NIEHS to develop CDEs for myositis**, leveraging expertise from the International Myositis Assessment and Clinical Studies Group (IMACS)

NIH Advancing the Use and Development of Common Data Elements in Research Workshop

March 2024



HYBRID
Meeting

Advancing the Use
and Development
of **Common
Data Elements**
in Research



• Natcher Conference Center
Building 45, NIH Campus
Bethesda, MD

MARCH 6–7, 2024

Welcome

Susan Gregurick, PhD

Director

Office of Data Science Strategy
National Institutes of Health

Opening Keynote

Monica M. Bertagnolli, MD

Director

National Institutes of Health

NIH CDE Workshop Summary and Key Take-aways

- Defined CDEs and demonstrated value of standards
- Discussed approaches and lessons learned
- “Develop CDEs with the community” and engage patient advocates in the entire process
- Improve communication and collaboration



Request For Information (RFI): Proposed Use of Common Data Elements (CDEs) For NIH-Funded Clinical Research and Trials

May 2024

NIH CDE Request for Information (RFI)

- **CDEs for NIH-funded clinical research/trials**, including a set of minimal **core CDEs** in the demographics/personal characteristics category
- **Additional CDEs:** autoimmune and immune mediated diseases and other clinical research domains
- **Technology standards for using NIH CDEs** to improve discoverability, data transformation, mapping between existing data and NIH CDEs, and harmonization
- **NIH policies and governance on CDEs** that could facilitate and incentivize broader CDE usage in research and data sharing and management



Next Steps

- Foster ICO **collaborations**
- Establish a CDE **community** of practice (CoP)
- Foster **communication** with the implementation of the NIH Data Sharing and Management Plan (DMSP)
- Develop additional CDE training and documentation
- Implement CDE technology and infrastructure
 - Enhance NIH CDEs to be more FAIR and discoverable
 - Cross-map between CDEs and data standards, data models, terminologies, etc.
 - Apply emerging technologies, such as artificial intelligence/machine learning (AI/ML), to advance CDE development and implementation

NIH Common Data Element (CDE) Repository Resources

- NIH CDE Repository: cde.nlm.nih.gov
- Common Data Elements: Standardizing Data Collection
- [Self-paced tutorial](#) on CDE basics (~1 hour)
- Standardize Your Research Data with the NIH CDE Repository (about 90 minutes)
- [On-demand training](#) (free registration required)
- [Recording](#) of live class (no registration required)
- Listserv: [CDE-R-NEWS](#)
- Questions? support.nlm.nih.gov

Slide courtesy of Robin Taylor, NLM



American Medical Informatics Association (AMIA) Annual Symposium

November 2024

AMIA Climate and Environmental Health

“What areas should the informatics community focus on to support healthcare systems in mitigating or adapting to climate change?” – 2023 Annual Symposium

- **Disaster Preparedness:** Focus on fortifying *IT infrastructure* against natural disasters (e.g., floods, outages) and establishing essential health records for crisis response, ensuring continuity of care and data access in emergencies.
- **Remote Monitoring and Surveillance:** *Wearables and remote technologies* for patient monitoring, especially in heat-sensitive and vulnerable populations, improving surveillance, risk stratification, and remote assessment techniques.
- **Carbon Footprint in Healthcare:** Analyze and reduce the carbon impact of healthcare materials, workflows, and hospital design, encourage a resource-conscious culture and quantify telehealth's carbon benefits. ‘
- **Data Standards:** Develop *coding standards* and essential data for climate informed health **metrics**.

AMIA 2024 Climate Change WG

- Impact of climate change and health disparity
- AI and informatics solutions have the potential to mitigate these issues
- Promote collaboration among stakeholders
- Challenges and opportunities
 - Environmental impact of AI
 - Bias and fairness in AI algorithms
 - Privacy, security and confidentiality, informed consent

NIH Office of Data Science Strategy

A modernized, integrated, FAIR biomedical data ecosystem.



[@NIHDataScience](https://twitter.com/NIHDataScience)



[linkedin.com/showcase/nih-office-of-data-science-strategy](https://www.linkedin.com/showcase/nih-office-of-data-science-strategy)

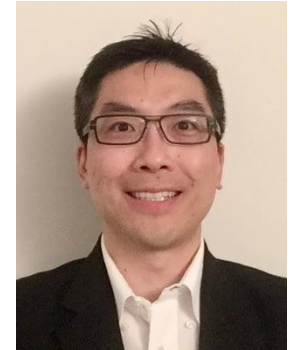
datascience@nih.gov



Clinical Informatics Unit



Belinda Seto, PhD
Deputy Director, ODSS



Steve Tsang, PhD
Lead/Program Director



Shu Hui Chen, PhD
Program Director



Snipta Mallick
Data Scientist