



# Introduction

The National Institutes of Health (NIH) Cloud Platform Interoperability (NCPI) hosted a virtual workshop spread across two sessions on May 1 and May 6, 2025. The goal of the workshop was to conduct a deep dive on Search across NCPI systems, FHIR, and data standards to understand the technical challenges that need to be addressed to advance technical interoperability and scientific research in the NCPI ecosystem.

Based on the output of the workshop, the NCPI Administrative Coordinating Center (ACC) will collaborate with NCPI leadership and the NCPI community to develop a roadmap to advance the NCPI ecosystem over the next nine months, focusing on enhancing search capabilities and continuing work towards developing the FHIR Implementation Guide. The roadmap will include key milestones, timelines, and engaged groups/individuals.

This document summarizes the discussion from the workshop, provides key themes and discussion points, and identifies potential actions for inclusion in the roadmap.

## Table of Contents

Introduction .....	1
Workshop Overview .....	2
Workshop Logistics .....	2
Workshop Attendees .....	2
Workshop Outputs.....	2
Session 1: Search .....	2
Key Themes and Discussion Points .....	2
Potential Actions for Inclusion in NCPI Roadmap .....	2
Session 2: FHIR Implementation.....	3
Key Themes and Discussion Points .....	3
Potential Actions for Inclusion in NCPI Roadmap .....	4



## Workshop Overview

### Workshop Logistics

The workshop was held virtually in two sessions: May 1, 2025, focused on enhancing search across NCPI, and May 6, 2025, focused on FHIR implementation efforts. Each session was structured to provide a framing presentation and overview of the current state, validate a shared understanding of the desired future state, and identify the next steps needed to achieve it.

### Workshop Attendees

Workshop attendees included the principal investigators from each of the interoperability projects, additional team members from the interoperability projects, members of the NCPI Steering Committee, and NIH.

## Workshop Outputs

The workshop outputs are summarized below. The potential actions for inclusion in the roadmap reflect workshop discussions but have not yet been fully vetted for feasibility, including achievability within the nine-month timeframe and with available resources.

### Session 1: Search

Searching across NCPI systems is a key challenge for researchers. The NCPI ecosystem includes a large catalogue of biomedical datasets, but researchers are currently not able to search this catalogue easily to perform their analyses. This workshop focused on the status and challenges of improving search across NCPI.

### Key Themes and Discussion Points

- **Unified Search – No Single Door:** Participants find it unclear where to start a search and how extensive it is. An ideal search should have clear starting points and allow deep data exploration.
- **Data Harmonization:** Current inconsistent data standards complicate consolidation and searching. Standardized formats are needed for better dataset alignment, common data elements, and richer metadata.
- **Access and Governance:** Inefficient manual data access application and review processes hinder data accessibility. Researchers want streamlined and easy access capabilities.
- **Enhanced Search Capabilities:** Currently, researchers have limited understanding of available data before applying for access. Ideally, semantic search and pre-analysis capabilities would allow researchers to understand the available data before applying for access.

### Potential Actions for Inclusion in NCPI Roadmap

#### *Unified Search—No Single Door*

- Complete a landscape analysis of existing metadata models.
- Identify use cases where unified systems have been achieved and how to build off them.



- Organize hackathons to discover what can be built based on NCPI metadata.

### *Data Harmonization*

- Develop policy for adding or correcting metadata.
- Establish a standardized data model for contributors to follow.
- Provide training and tools to facilitate harmonization.
- Develop or consider AI tools to reduce manual burden.

### *Access and Governance*

- Leverage AnVIL modeling to enhance user/project profiling (Dependent on Data Access Committee Approval).
- Provide participant level access for search (e.g., View-Only Data Access Request [VODAR]).
- Allow access to data after initial registration.
- Provide/update existing policies for access logging requirements.

### *Enhanced Search Capabilities*

- Enforce a metadata model for general search with summaries of datasets and general dataset attributes.
- Develop knowledge graphs to facilitate search.

## **Session 2: FHIR Implementation**

NCPI faces the challenge of exchanging many types of biomedical data for researchers to find and use. To address this, NCPI uses the [Fast Healthcare Interoperability Resources \(FHIR\)](#) data exchange standard. This workshop focused on the status and challenges of FHIR implementation across NCPI.

### **Key Themes and Discussion Points**

- **Testing and Implementation:** The testing of servers and implementation guidelines through use cases is crucial for identifying and addressing gaps and missing features. Participants recommended providing information on the functional capabilities of FHIR servers that inform users on what functions they can use to search, call, or aggregate data across different FHIR servers.
- **FHIR Education and Training:** Participants emphasized the importance of developing and curating training materials for researchers. Although the development and implementation of FHIR standards are essential, researchers and other individual users may be unfamiliar with the FHIR format. They would benefit from introductory training on the subject, as FHIR itself is not intuitive.
- **Navigating the Differences between Observational and Clinical Research:** Data collected for different purposes, including for clinical trials, research studies, and the provision of clinical care can all be informative for research, but may be formatted differently or refer to similar but distinct concepts (e.g., a lab result indicating diabetes is distinct from a patient stating on a survey that they have diabetes which is different than a detailed clinical assessment.) Understanding how those data can differ, even when they are all in FHIR format, is important for researchers. Having information directly from an EHR via bulk FHIR



would provide access to a much wider range of data for research, but detailed technical and clinical knowledge is needed to fully make use of that data as well as information about consent for using that information for research.

- **Managing Metadata:** Participants reiterated the importance of metadata harmonization. This harmonization is essential for reformatting and integrating data, governing what data transformations can be performed. A balance is required to minimize complexity, as increasing flexibility for data submitters imposes burdens on consumers of the data platforms.

## Potential Actions for Inclusion in NCPI Roadmap

### *Testing and Implementation*

- Complete FHIR Implementation Guide.
- Support FHIR implementation and testing including transformations of FHIR data by platform teams.

### *FHIR Education and Training*

- Focus on sustainability, including tools and education for researchers, specifically tools to facilitate the use of FHIR research data and FHIR APIs.

### *Navigating the Differences between Observational and Clinical Research*

- Explore other gaps not addressed directly by FHIR, including the differences in how data from observational trials, clinical trials, and clinical care is represented.

### *Managing Metadata*

- Consider additional solutions for advancing the use of common data elements (CDEs), determining which vocabularies and ontologies should be used, and other metadata issues.