

Proposal for a Working Group for the Common Fund Data Ecosystem

May 20, 2021



National Institutes of Health
Office of Strategic Coordination - The Common Fund

Background and Proposal



The Common Fund

- The Common Fund Data Ecosystem (CFDE) is an infrastructure investment intended to support and enhance the value of many Common Fund programs.
- It began in 2018 with a Coordinating Center (CFDE-CC), which was asked to work with several Common Fund Data Coordinating Centers to determine needs and opportunities for collaboration.
- It expanded in 2020 to include “Data Coordinating Center Engagement Awards” which were asked to work together and with the CFDE-CC to establish an ecosystem of Findable, Accessible, Interoperable, and Reusable data sets.
- We have considered the period ending in Sept 2022 as a pilot/exploratory phase. We will be making decisions about the future scope and goals for the CFDE at that point.
- We propose a Working Group of the Council of Councils to review the vision established for the CFDE and its progress to date.

Goals of the CFDE



The Common Fund

- Enhance the value of Common Fund investments by enabling users to query across and use Common Fund data sets
 - Individual data sets are silo'ed
 - Querying across data sets would facilitate discoveries by allowing re-use of data
- Ensure that Common Fund data and tools are sustained after individual programs end
 - Sustainability beyond the ten-year lifespan of Common Fund programs must be considered
 - Most data sets are folded into existing repositories/platforms, but this is not always the case
- Train users to work with Common Fund data in a cloud environment
 - Working with data in the cloud is more secure
 - Storing data in the cloud is more cost-effective
 - Working with data in the cloud prevents the need for data egress from the cloud and therefore reduces costs
 - Working with the data is a shift in the way people work

Proposed scope of a CFDE Working Group



The Common Fund

- **Review the current scope and goals of the CFDE as well as progress to date**
- **Make recommendations for future scope and goals in the following areas:**
 - Findability and accessibility of data
 - Data harmonization and interoperability
 - Cloud workspaces
 - Sustaining access to data and tools after CF programs end
 - Training and outreach to enhance access to, and use of, the data
 - CFDE scope and strategy in the context of related NIH activities

Leadership and Timeframe



The Common Fund

- Proposed Co-Chairs: Rick Horwitz (Council of Councils) and Betsy Wilder (NIH)
- Timeframe: June 2021 through May 2022
- Report expected to be delivered to the Council of Councils at the May 2022 meeting
- If the Council accepts the report of the Working Group, it would deliver the report to the NIH Director
- NIH expects to provide a response to the report at the Council of Councils meeting in September 2022