

Biomedical Data Repositories and Knowledgebases (DRKB) FY26 Concept Clearance

Susan Gregurick

Associate Director for Data Science, Office of Data Science Strategy

Ishwar Chandramouliswaran

Program Director, Office of Data Science Strategy

Christopher Duncan

*Program Director, National Institute of Environmental Health
Sciences*

September 12, 2025

NIH Council of Councils

Concept Clearance Overview

Concept Clearance: Reissue (PAR-23-237)

Title: Enhancement and Management of Established Biomedical Data Repositories (DR) and Knowledgebases (KB) **(U24 Mechanism)**

Objective/Purpose: To seek data repository and knowledgebase resource projects with a unique value proposition that deliver scientific impact, employ and promote good and efficient data management, engage the user community, and govern to address biomedical data life-cycle and long-term preservation.

Funds Available and Anticipated Number of Awards: Contingent upon NIH appropriations and submission of meritorious applications

Award Project Period: up to 5 years

Council Action: Vote for approval of reissuance of the concept for the Enhancement and Management of Established Biomedical Data Repositories and Knowledgebases

NIH Strategic Plan for Data Science 2025-2030

OVERARCHING GOALS, STRATEGIC OBJECTIVES,
AND IMPLEMENTATION TACTICS



Goal 1

Improve Capabilities to Sustain the NIH Policy for Data Management and Sharing

› OBJECTIVE 1-1

Support the Biomedical Community to Manage, Share, and Sustain Data

› OBJECTIVE 1-2

Enhance FAIR Data and Greater Data Harmonization

› OBJECTIVE 1-3

Strengthen NIH Data Repository and Knowledgebase Ecosystem

To sustain a healthy and productive data resource ecosystem, it is critical to ensure that DRKBs:

- **Deliver** scientific impact to the communities that they serve.
- **Employ** and promote good data management practices and efficient operation for quality and services.
- **Engage** with the user community and continuously address their needs.
- **Implement, adopt, or contribute** to openly shared metrics.
- **Provide** sufficient metadata and semantic annotation.
- **Support** a process for data life-cycle analysis, long-term preservation, and trustworthy, responsible stewardship.

Goals of the DRKB Program

An NIH program to support investigator-initiated, sustainable data resource development driven by critical research needs



* Including NIH desirable characteristic of repositories [NOT-OD-21-016](#)

** Requiring a transition plan that addresses long-term archival and sun-setting

Background and History: DRKB PARs

Launched as part of **2018 NIH Strategic Plan for Data Science**

No prior NIH-wide data resource program – databases and knowledgebases were primarily funded as research grants

DRKBs reviewed as Resources – through CSR Special Emphasis Panels (SEPs)

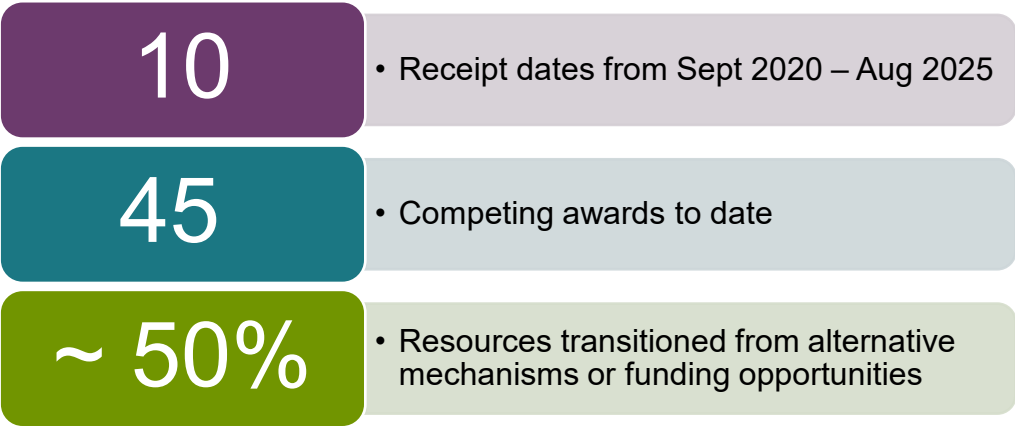
2020-2023 PARs

- 2 NOFOs distinguished by resource function:
 - **Biomedical Knowledgebases U24** ([PAR-20-097](#) | [PAR-23-078](#))
 - **Biomedical Data Repositories U24** ([PAR-20-089](#) | [PAR-23-079](#))
 - 12 participating ICOs: NCI, NEI, NHLBI, NIAAA, NIAID, NIAMS, NIDA, NIEHS, NIGMS, NIMHD NINDS, ODSS

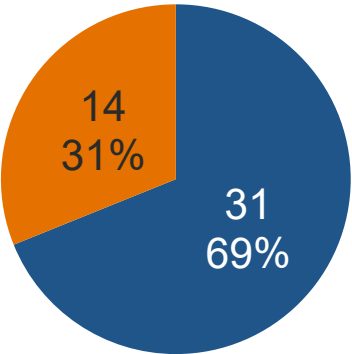
2023-2026 PARs

- 2 NOFOs distinguished by resource lifecycle stage:
 - **Enhancement and Management of Established Biomedical Data Repositories and Knowledgebases U24** ([PAR-23-237](#))
 - **Early-stage Biomedical Data Repositories and Knowledgebases R24** ([PAR-23-236](#))
 - 18 participating ICOs: NCCIH, NCI, NEI, NHGRI, NHLBI, NIA, NIAAA, NIAMS, NIBIB, NIDA, NIDCD, NIDCR, NIDDK, NIEHS, NIGMS, NIMHD, NINDS, ODSS

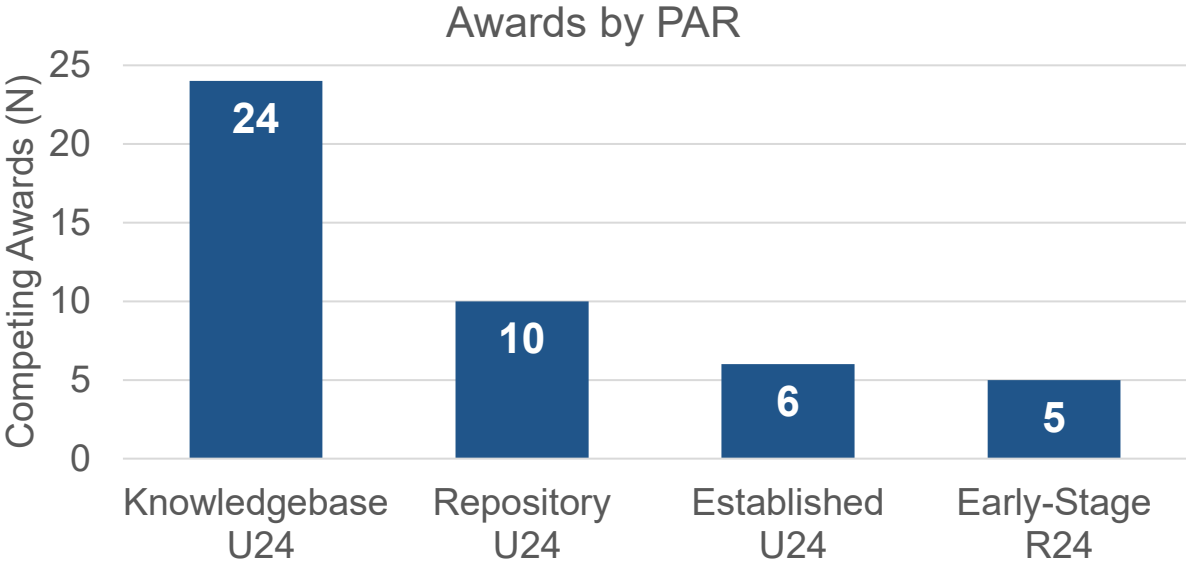
Summary of DRKB Awards: FY21-FY25



Primary Function of FY21-25 DRKB Awards



■ Knowledgebase ■ Data Repository



Knowledgebase U24	• Biomedical Knowledgebases U24 (PAR-20-097, PAR-23-078)
Repository U24	• Biomedical Data Repositories U24 (PAR-20-089, PAR-23-079)
Early-Stage R24	• Early-stage Biomedical Data Repositories and Knowledgebases R24 (PAR-23-236)
Established U24	• Enhancement and Management of Established Biomedical Data Repositories and Knowledgebases U24 (PAR-23-237)

DRKB Awards: FY21-FY25

Knowledgebase U24

Resource Name	Award ICs
Integrated AIRR Knowledge Commons (i-AKC)	AI
Vaccine Information and Ontology (VIOLIN)	AI
Virus Taxonomy KB	AI
ClinGen Somatic KB / CIViC	CA
HemOnc KB	CA
Exposome Correlation and Interpretation (ECID)	ES
Comparative Toxicogenomics Database (CTD)	ES,NS
ROBOKOP	ES,OD
BindingDB	GM
BioPortal	GM
E. coli Genome and Metabolism KB (EcoCyc)	GM
GlyGen	GM
REDfly	GM
RRID Portal / Resource Watch	GM
African Genomics Data Hub	HG
FlyBase	HG
Human Disease Ontology	HG
Reactome	HG
WormBase	HG
UniProt	HG,AG,AI,CA,DK,EY,GM,HL,OD
GWAS and PGS Catalogs	HG,CA
PharmGKB	HG,CA,DA,HD
Gene Ontology Consortium and KB	HG,GM
PhenX Toolkit	HG,HL,MD,NS,OD

Repository U24

Resource Name	Award ICs
Southern HIV and Alcohol Research Consortium DR (SHARC)	AA
Childrens Allergy and Asthma DR (CADRE)	AI,OD
HIV Pharmacology DR	AI,OD
ISIC Skin Imaging Repository	CA
Biological Magnetic Resonance Data Bank (BMRB)	GM
ProteomeXchange / MassIVE	GM
Georgia Cystic Fibrosis Data Warehouse	HL
Pan-Neurotrauma Data Commons	NS
Pennsieve	NS
Cerebrovascular Disease Knowledge Portal (iCDKP)	NS,OD

Early-Stage R24

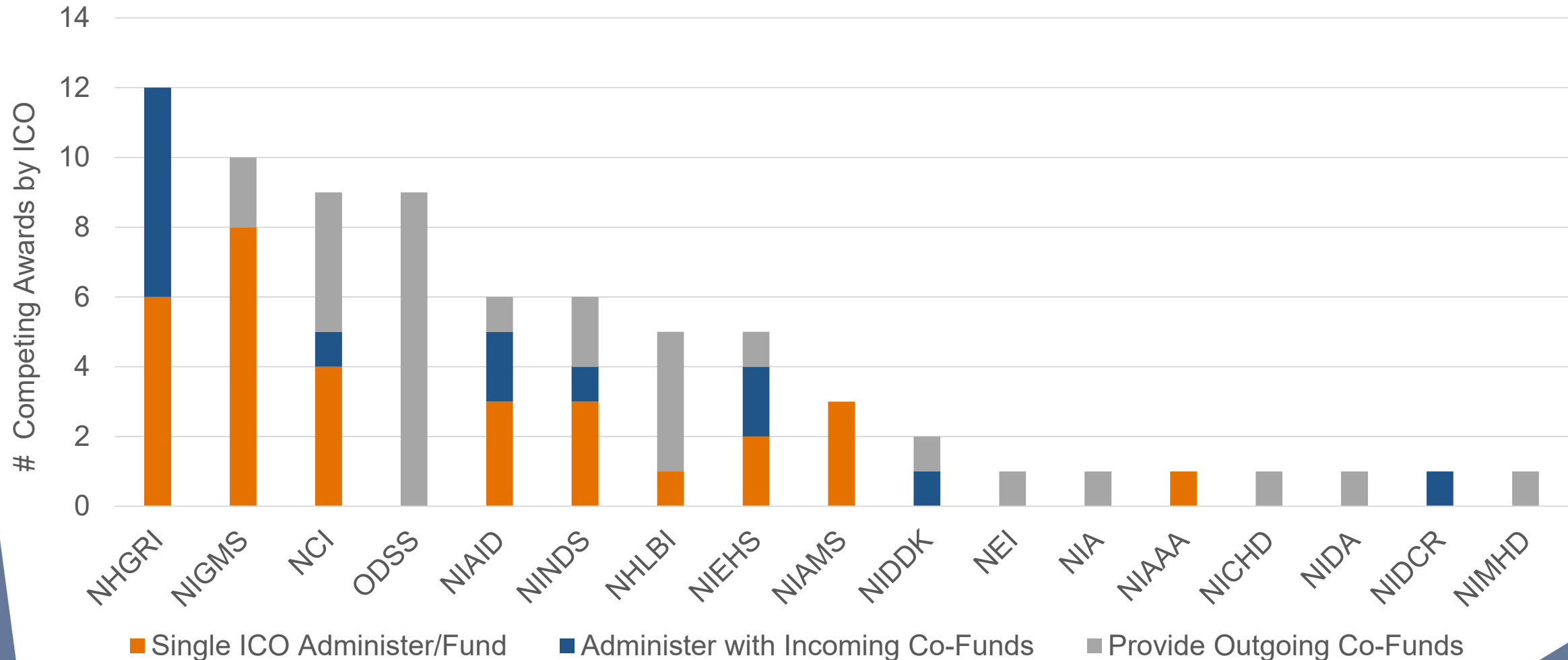
Resource Name	Award ICs
Musculoskeletal Knowledge Portal (MSKKP)	AR
OAI Collaborative Osteoarthritis Research Enterprise (OAI CORE)	AR
Skeletal Phenotyping Database	AR
GeoSpatial Knowledgebase for Exposomics (GeoSpace)	ES
International Seizure and SUDEP Research Repository (InSSURR)	NS

Established U24

Resource Name	Award ICs
BRCA Exchange	CA
Molecular Signatures Database (MSigDB)	CA,OD
Facebase	DE,OD
Metabolomics Workbench	DK,CA,ES,HL,OD
Nomenclature of Human and Vertebrate Genes	HG
Alliance Central	HG,HL,OD

DRKB Awards by ICO

- 17 ICOs have funded or co-funded at least 1 DRKB award to date
- 15 ICOs have engaged in co-funding (in/out) for at least 1 DRKB award



Fostering a DRKB Network

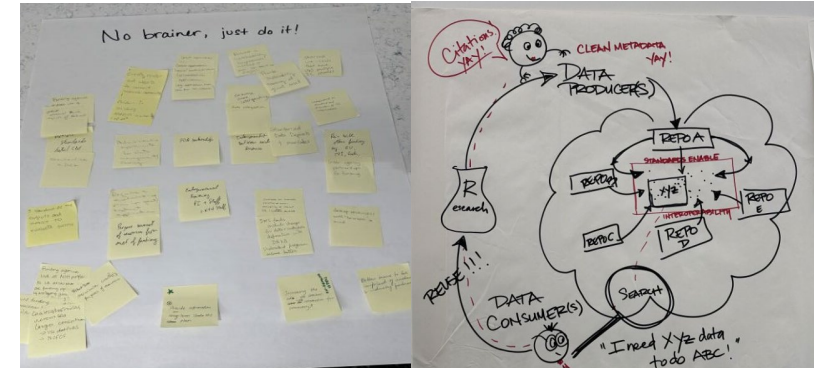
- Annual meeting convening PIs representing the DRKB program and related NIH programs.
- Foster a community around biomedical data repositories and knowledgebases funded by the NIH.
- Address cross-cutting topics related to the biomedical data ecosystem.

2024 Meeting Topics: Data curation & Sustainability

Outcomes:

Scientific panels discussed AI-driven and community-driven data curation, highlighting the need for standardization, expert data curators, and improved training.

Activity sessions focused on communicating project value, planning for sustainability, and fostering collaboration, with a strong emphasis on outreach, impact, and inclusivity in the data ecosystem.



2025 Meeting Topics: Metrics & Metadata

Outcomes:

Scientific panels addressed the significance of good metadata for making data FAIR and the necessity of standardized metrics to track the impact of data sharing.

Activity sessions focused on leveraging metadata to further biomedical research and identifying metrics to assess the impact and value of repositories, emphasizing community value, scientific impact, and return on investment.



Reissue of Program: Single PAR for Established Biomedical Data Repositories and Knowledgebases

Solicitation to focus on established data resources that continue to better address NIH Institute-specific mission priorities, life-cycle stages, and NIH goals for data resources and their sustainment.

- Objective:** Support established data repositories or knowledgebases
- Mechanism:** U24
- Duration of Award:** Up to 5 years
- Budget:** No limit; must reflect actual needs of the project.
- Due Dates:** 2 receipt dates per year, Sept and Jan

Program Requirements:

Scientific significance of proposed resource

Demonstrate community needs and engagement

Address quality of data, services, & efficiency of operations

Document & implement governance processes

Reviewed by CSR Special Emphasis Panel

Focus of the Reissue

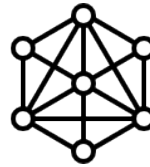
What is New in the reissue?

Emphasis on established data resources

Foster collaboration on cross-cutting topics

Encourage responsible stewardship

Optimize data resources landscape



Benefits/Outcomes

Lower barriers for data sharing

Improve discovery and use of NIH-funded data

Increase efficiency of operations and optimize costs

Disentangle data resources from research projects

This program is fundamental to creating a modernized biomedical data ecosystem

Concept Clearance Vote

Concept Clearance: Reissue (PAR-23-237)

Title: Enhancement and Management of Established Biomedical Data Repositories (DR) and Knowledgebases (KB) **(U24 Mechanism)**

Objective/Purpose: To seek data repository and knowledgebase resource projects with a unique value proposition that deliver scientific impact, employ and promote good and efficient data management, engage the user community, and govern to address biomedical data life-cycle and long-term preservation.

Funds Available and Anticipated Number of Awards: Contingent upon NIH appropriations and submission of meritorious applications

Award Project Period: up to 5 years

Council Action: Vote for approval of reissuance of the concept for the Enhancement and Management of Established Biomedical Data Repositories and Knowledgebases

Discussion