

# Arcus Grant Language

## Letter of Support

We can provide a letter for support for your research. CHOP CRIO Jeffrey Pennington [penningtonj@chop.edu](mailto:penningtonj@chop.edu) will contact you directly.

## Sample Language for Grants and IRB Applications

Below are some sample snippets for grant or IRB applications when research projects will utilize Arcus resources. For additional exemplary language, you can submit a request to Arcus at <https://support.arcus.chop.edu/servicedesk/> (Arcus General Support → Privacy → Language for Protocol).

## Data Collection and Management

Data in this study will be created, linked, stored, and/or analyzed in the Arcus platform. Arcus is a CHOP enterprise strategic initiative to develop a next-generation data platform. The mission of Arcus is to integrate each patient and/or research subject's biological, clinical, research and environmental data for the purposes of supporting collaborative research, innovation, and discovery. Arcus features a clinical data repository, as well as a research data archives and library-based discovery service that preserves, manages, and makes CHOP's research data discoverable alongside the contextual information and tools required for reusable research. Arcus also provides data management services, including service as an Honest Broker in the production of de-identified datasets.

Investigators and assigned project staff can securely store, access, and process project data in Arcus. The Arcus program is staffed by archivists, librarians, information analysts, cloud computing engineers, programmers, statisticians, and privacy experts. Data is managed through the oversight of the CHOP Institutional Review Board and access is governed by multiple institutional policies. Clinical and research data are linked and made available to researchers with a CHOP-issued credential. Arcus security configuration and controls are based on the HIPAA Security Rule and are subject to audit by CHOP's independent Internal Audit department which reports directly to the CHOP Board of Trustees.

## Using the Arcus Archives

### Data Sharing in the Arcus Archives

The Arcus Archives is the canonical repository for the data of Children's Hospital of Philadelphia (CHOP) Research Institute's research efforts. The Archives aims to store research projects holistically, archiving data, contextual files, tools, and metadata so that data will be reproducible, reusable, and repurposable.

# Arcus Grant Language

Arcus Archives data is preserved according to international digital archiving standards and utilizes a custom ingestion and processing workflow designed by Digital Archivists, Application Research Developers, and DevOps Engineers. Data will be replicated across geographically disparate, secure, monitored storage environments (Google Cloud Platform and Amazon Web Services) to prevent loss in the case of a catastrophic event. Data are thoroughly described using a metadata schema incorporating Data Use Ontology (DUO), storing ORCID Identifiers, and applying Medical Subject Heading (MeSH) subject terms. Metadata about research data and participants in the Archives are available for browsing and mediated request via the Arcus Cohort Discovery (ACD) tool and the Data Catalog (forthcoming).

Contributors to data in the Arcus Archives or users of Arcus tools or data must cite or attribute Arcus in publications. This applies to both original research or to research reusing or reproducing data delivered from Arcus.

## Accessing Geographic Information System (GIS) in a Scientific Project

This project intends to utilize Arcus functionality to explore geographic information system (GIS) mapping and spatial analysis, allowing researchers to analyze relevant health factors based on spatial location. To do this analysis, Arcus will provide a spatial join for CHOP participant-level data to public geo-spatial data sets (e.g. Demographic data from the United States Census Bureau, USGS, NASA, EPA, NJDEP, PADEP, NYDEC, regional government cadastral data, or others). These data sets can provide information such as environmental factors (e.g., air, water, sound pollution and others); social determinants of health (house values, income, and others), and other information that can be useful for understanding health outcomes.

## Biospecimens and Arcus

### Biospecimen Integration within Arcus

*Choose the appropriate section(s) based on which biobank(s) you'll be using.*

### For samples stored in the CHOP Biorepository Resource Center (BioRC)

This study utilizes or stores samples in the CHOP Biorepository Resource Center (BioRC).

For all protocols accepted into the CHOP Biorepository Resource Center (BioRC), Arcus offers automatic integration with clinical and archived data in its environment.

# Arcus Grant Language

Arcus offers the Arcus Cohort Discovery (ACD) tool that allows for querying of deidentified clinical data to return counts of participants meeting various inclusion/exclusion criteria for cohort creation. One of the returned metrics and parameters for cohort criteria includes the presence of biospecimens, their location, and the study protocol under which they were collected. This integration is one of the ways studies taken into the BioRC comply with the Biobanking Principles requirement for sharing data with Arcus.

The Translational Informatics Group (TRiG) within the Department of Biomedical and Health Informatics (DBHi) at CHOP maintains an Extract, Transform, Load (ETL) function that links biospecimens in the BioRC with clinical and archived data in Arcus. This ETL script pulls a deidentified local identifier from their homegrown BioRepository Portal and links information about the sample from the BioRC's Nautilus Laboratory Information Management System (LIMS). This information is passed to Arcus in its function as an honest broker, where local identifiers are associated to authoritative identifiers, like MRNs, on the backend. The result of this linking is the biospecimens filter in the ACD which produces deidentified counts on the biospecimens' intersection with other phenotypic parameters.

## **For samples stored in Biobanks not or not yet formally integrated with Arcus**

This study utilizes or stores samples in the following biobanks at CHOP: \_\_\_\_\_

To search for data or samples not contributed or linked through the Arcus Archives, researchers should utilize the individual biobank's database, tools, or team to query for samples related to cohorts of interest. They should then follow the individual biobank's procedures for requesting data stored with them or for requesting physical samples.

If a larger cohort is identified than has samples in biobanks not integrated with Arcus, researchers can utilize Arcus to search for data in the Archives meeting their need and/or to search for physical samples in biobanks (e.g., the BioRC) already integrated with Arcus.

## **Biospecimen Linking as Part of a Data Contribution to the Arcus Archives**

This study entails a contribution to the Arcus Archives.

Regardless of where biospecimens are banked, all biospecimens collected as part of a study whose data is contributed to the Arcus Archives will have basic information about those biospecimens collected and surfaced for linking in accordance with Arcus' Data Use and Sharing policies. Minimal information collected about biospecimens as part of the data contribution process includes a sample identifier, a Laboratory Information Management System (LIMS) identifier, and documentation of the link to analysis files

# Arcus Grant Language

produced from that specimen. We'll also collect any other information necessary to the project (such as control or case designation, for example) or that the researcher chooses to share with Arcus.

As an IRB-approved honest broker, Arcus can utilize biospecimen or LIMS identifiers to link participants back to authoritative identifiers, such as MRNs. Thus, Arcus can link all biospecimens collected under studies included in Arcus or the biobanks formally integrated with Arcus either now or in the future.

## General Descriptions

### General Arcus Description

Arcus is a suite of tools and services developed to enhance research efforts at The Children's Hospital of Philadelphia by helping researchers to explore available data, see overlaps among datasets, build new cohorts, and determine if there are data or samples available for additional research projects. Incubated within the Department of Biomedical and Health Informatics at CHOP, Arcus connects CHOP's clinical and research data to enable biomedical researchers to conduct highly innovative, data-driven, reproducible research within a managed scalable framework. The program is designed to: help researchers discover, link, and analyze rich data about our patients and research participants using datasets from across CHOP, provide computational environment within access to additional tools - and scientific breakthroughs' and increase opportunities for new research by spurring new ideas, efforts, and collaborations.

### Department of Biomedical and Health Informatics (DBHi)

DBHi provides an academic home and service base for all research informatics activities at CHOP, including the development and deployment of intellectual, technical, and educational resources in biomedical computing. DBHi encompasses units for bioinformatics, scientific computing, and clinical data reporting from hospital systems; partners with existing CHOP-based groups for outcomes research, quality improvement, decision support, biostatistics, and bioethics; and closely coordinates the analysis of highthroughput genomic data derived from CHOP-based samples. DBHi also consists of a translational informatics group that develops novel technical solutions for research CHOP wide. DBHi's Data Science and Biostatistics unit works with researchers to define the research question, select appropriate data, and develop methodologies for data collection and analysis. Additionally, DBHi is the home of CHOP's Clinical Informatics Fellowship program. DBHi serves as the primary informatics interface with the University of Pennsylvania through a joint CTSA award and through its collaborative relationship with the University's Institute for Biomedical Informatics.