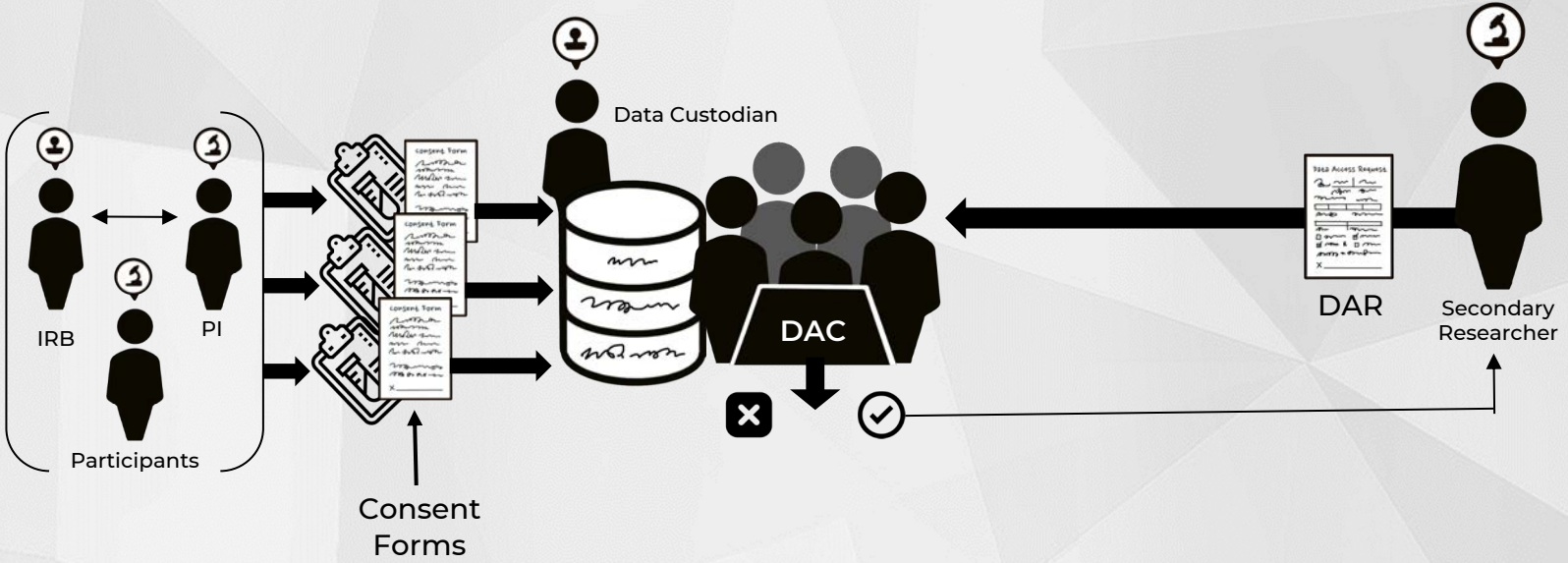
DUOS & GA4GH Standards

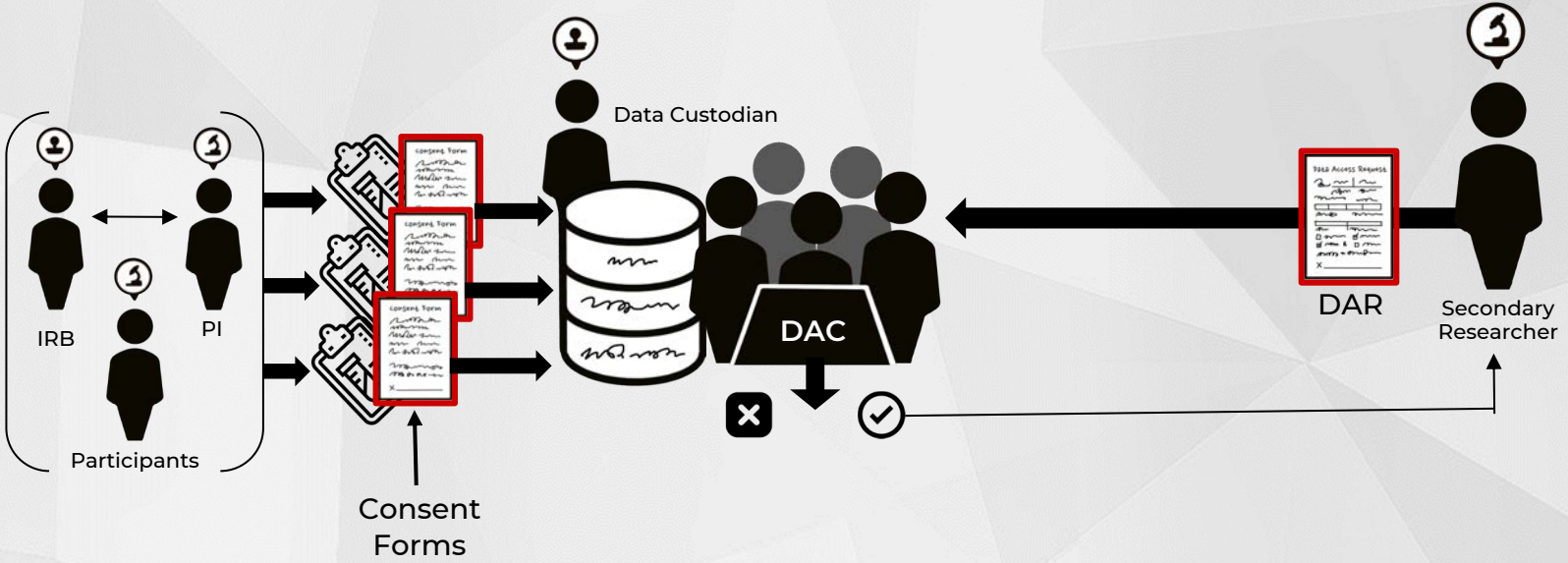
BROAD INSTITUTE

Global Alliance for Genomics & Health

Current Data Access Request Process



Current Data Access Request Process



BROAD INSTITUTE

Global Alliance for Genomics & Health

GA4GH Data Use Ontology (DUO)

Open Access

No Restrictions (**NRES**)

Controlled Access

Permissions

General Research Use (**GRU**)

Health / Medical / Biomedical (**HMB**)

Disease Specific (**DS**)

Populations, Origins, and Ancestry (**POA**)

+

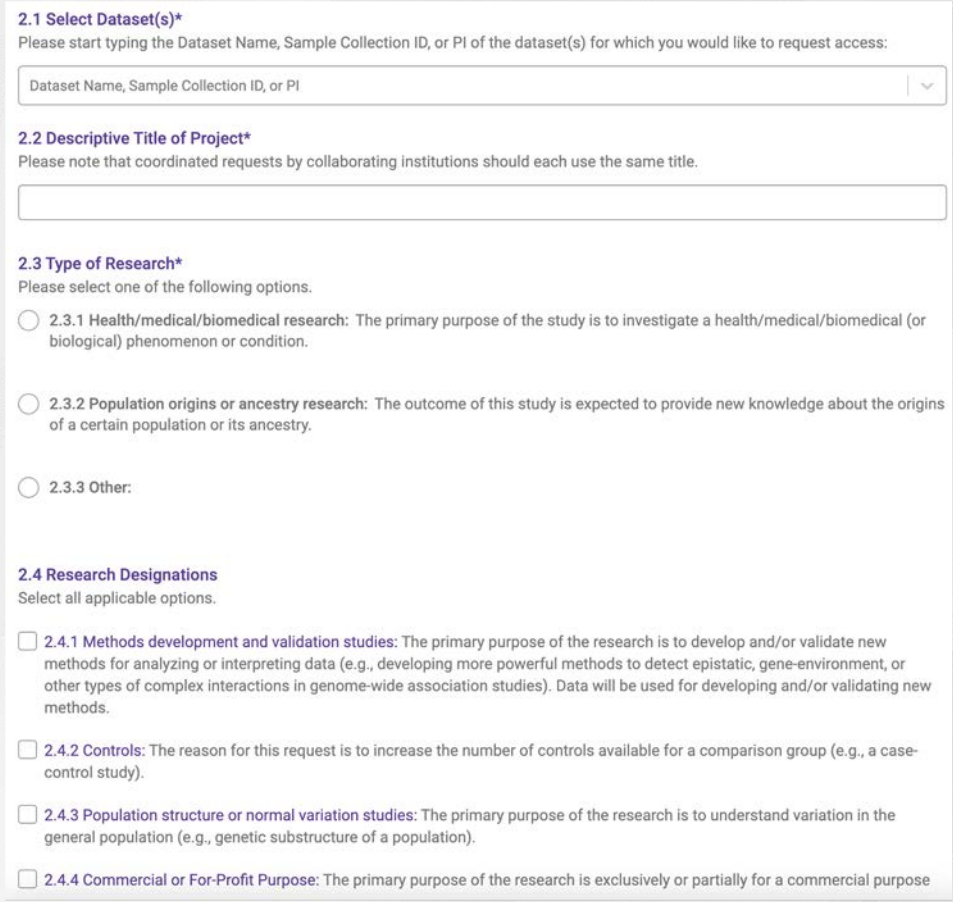
Modifiers

NPOA	No population origins or ancestry research
NMDS	No general methods research
GSO	Genetic studies only
CC	Clinical care use
PUB	Publication required
COL	Collaboration required
IRB	Ethics approval required
GS	Geographical restriction
MOR	Publication moratorium
RT	Return to database/resource
NCU	Non commercial use only
NPU	Not-for-profit use only
NPUNCU	Not-for-profit, non-commercial use only

BROAD INSTITUTE

DUOS Data Access Request Application

The full DAR application can be found on the DUOS website at https://duos.broadinstitute.org/dar_application



2.1 Select Dataset(s)*

Please start typing the Dataset Name, Sample Collection ID, or PI of the dataset(s) for which you would like to request access:

Dataset Name, Sample Collection ID, or PI

2.2 Descriptive Title of Project*

Please note that coordinated requests by collaborating institutions should each use the same title.

2.3 Type of Research*

Please select one of the following options.

- 2.3.1 Health/medical/biomedical research: The primary purpose of the study is to investigate a health/medical/biomedical (or biological) phenomenon or condition.
- 2.3.2 Population origins or ancestry research: The outcome of this study is expected to provide new knowledge about the origins of a certain population or its ancestry.
- 2.3.3 Other:

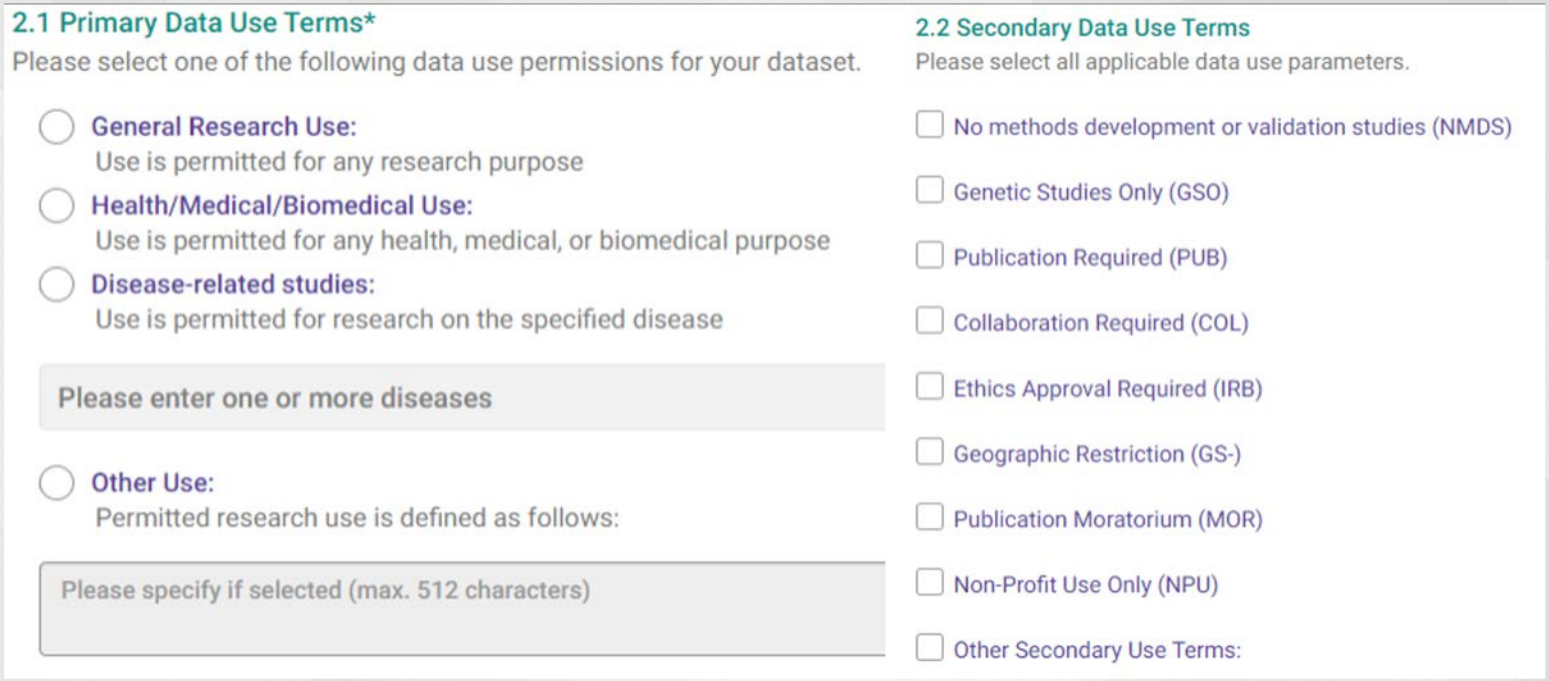
2.4 Research Designations

Select all applicable options.

- 2.4.1 Methods development and validation studies: The primary purpose of the research is to develop and/or validate new methods for analyzing or interpreting data (e.g., developing more powerful methods to detect epistatic, gene-environment, or other types of complex interactions in genome-wide association studies). Data will be used for developing and/or validating new methods.
- 2.4.2 Controls: The reason for this request is to increase the number of controls available for a comparison group (e.g., a case-control study).
- 2.4.3 Population structure or normal variation studies: The primary purpose of the research is to understand variation in the general population (e.g., genetic substructure of a population).
- 2.4.4 Commercial or For-Profit Purpose: The primary purpose of the research is exclusively or partially for a commercial purpose

BROAD INSTITUTE

DUOS Dataset Registration



2.1 Primary Data Use Terms*

Please select one of the following data use permissions for your dataset.

- ○ **General Research Use:**
  Use is permitted for any research purpose
- ○ **Health/Medical/Biomedical Use:**
  Use is permitted for any health, medical, or biomedical purpose
- ○ **Disease-related studies:**
  Use is permitted for research on the specified disease

Please enter one or more diseases

- ○ **Other Use:**
  Permitted research use is defined as follows:

Please specify if selected (max. 512 characters)

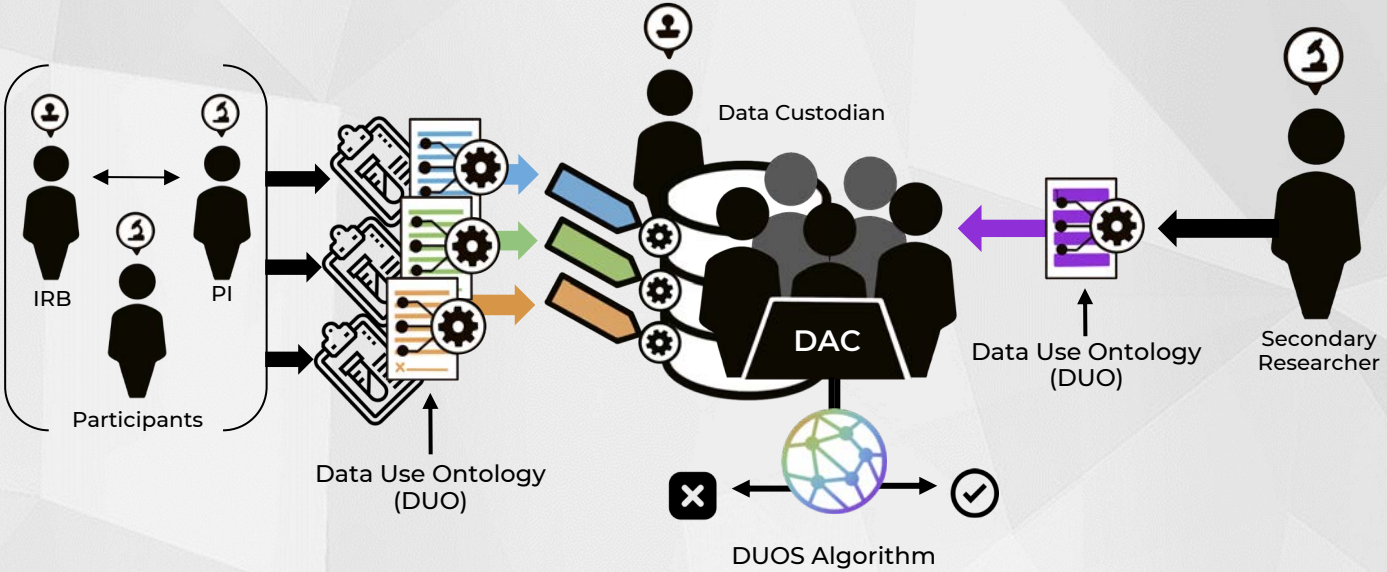
2.2 Secondary Data Use Terms

Please select all applicable data use parameters.

- ☐ No methods development or validation studies (NMDS)
- ☐ Genetic Studies Only (GSO)
- ☐ Publication Required (PUB)
- ☐ Collaboration Required (COL)
- ☐ Ethics Approval Required (IRB)
- ☐ Geographic Restriction (GS-)
- ☐ Publication Moratorium (MOR)
- ☐ Non-Profit Use Only (NPU)
- ☐ Other Secondary Use Terms:

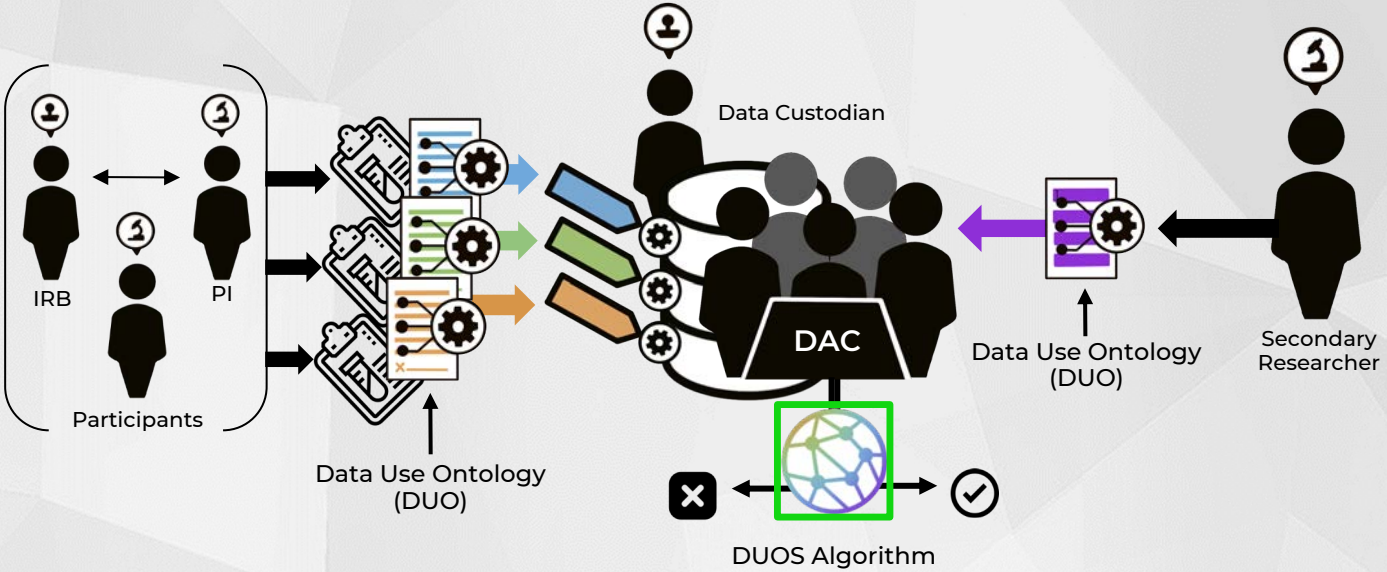
BROAD INSTITUTE

DUO-backed Algorithm for DAR Review



BROAD INSTITUTE

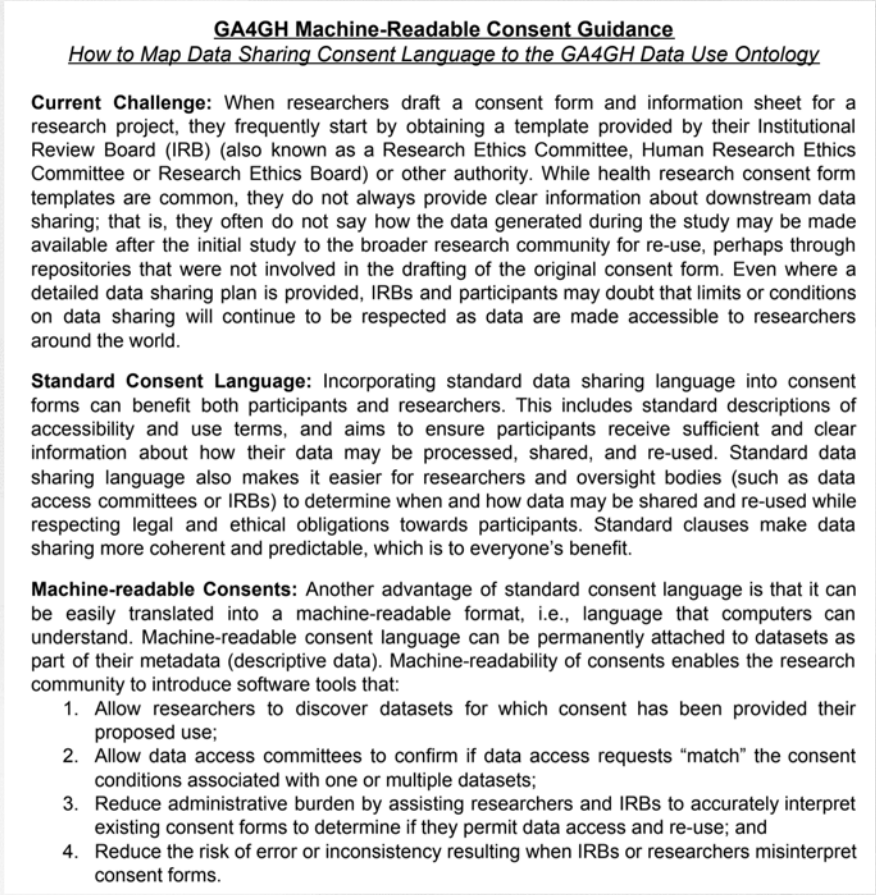
DUO-backed Algorithm for DAR Review



BROAD INSTITUTE

Machine-Readable Consent Guidance

The Machine-Readable Consent Guidance (MRCG) can be found here https://drive.google.com/file/d/102_I0_phOGs9YSmPx7It9CSt1sHFJ87C/view



GA4GH Machine-Readable Consent Guidance

How to Map Data Sharing Consent Language to the GA4GH Data Use Ontology

Current Challenge: When researchers draft a consent form and information sheet for a research project, they frequently start by obtaining a template provided by their Institutional Review Board (IRB) (also known as a Research Ethics Committee, Human Research Ethics Committee or Research Ethics Board) or other authority. While health research consent form templates are common, they do not always provide clear information about downstream data sharing; that is, they often do not say how the data generated during the study may be made available after the initial study to the broader research community for re-use, perhaps through repositories that were not involved in the drafting of the original consent form. Even where a detailed data sharing plan is provided, IRBs and participants may doubt that limits or conditions on data sharing will continue to be respected as data are made accessible to researchers around the world.

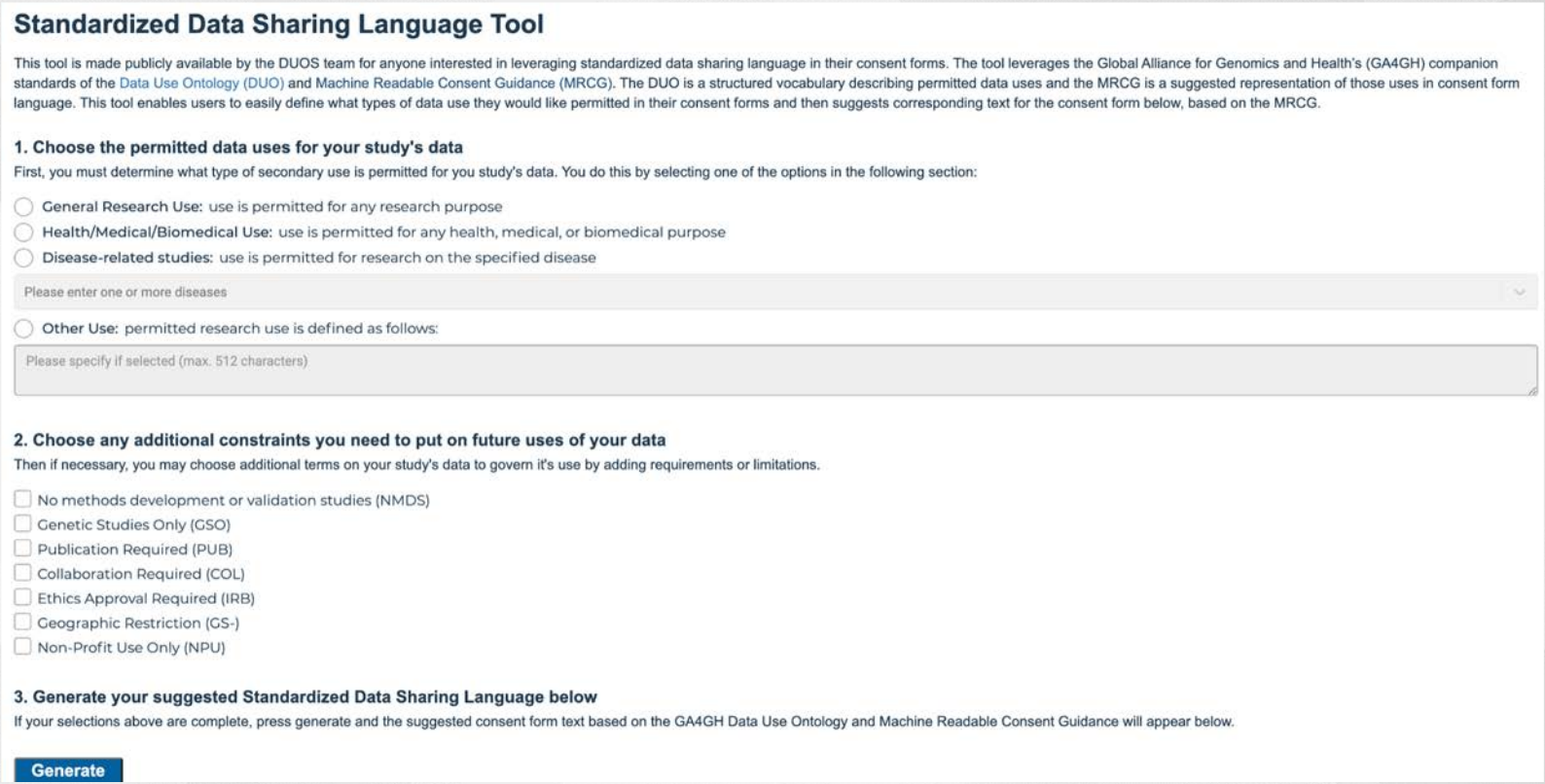
Standard Consent Language: Incorporating standard data sharing language into consent forms can benefit both participants and researchers. This includes standard descriptions of accessibility and use terms, and aims to ensure participants receive sufficient and clear information about how their data may be processed, shared, and re-used. Standard data sharing language also makes it easier for researchers and oversight bodies (such as data access committees or IRBs) to determine when and how data may be shared and re-used while respecting legal and ethical obligations towards participants. Standard clauses make data sharing more coherent and predictable, which is to everyone's benefit.

Machine-readable Consents: Another advantage of standard consent language is that it can be easily translated into a machine-readable format, i.e., language that computers can understand. Machine-readable consent language can be permanently attached to datasets as part of their metadata (descriptive data). Machine-readability of consents enables the research community to introduce software tools that:

1. Allow researchers to discover datasets for which consent has been provided their proposed use;
2. Allow data access committees to confirm if data access requests "match" the consent conditions associated with one or multiple datasets;
3. Reduce administrative burden by assisting researchers and IRBs to accurately interpret existing consent forms to determine if they permit data access and re-use; and
4. Reduce the risk of error or inconsistency resulting when IRBs or researchers misinterpret consent forms.

BROAD INSTITUTE

Data Sharing Language Tool



Standardized Data Sharing Language Tool

This tool is made publicly available by the DUOS team for anyone interested in leveraging standardized data sharing language in their consent forms. The tool leverages the Global Alliance for Genomics and Health's (GA4GH) companion standards of the Data Use Ontology (DUO) and Machine Readable Consent Guidance (MRCG). The DUO is a structured vocabulary describing permitted data uses and the MRCG is a suggested representation of those uses in consent form language. This tool enables users to easily define what types of data use they would like permitted in their consent forms and then suggests corresponding text for the consent form below, based on the MRCG.

1. Choose the permitted data uses for your study's data

First, you must determine what type of secondary use is permitted for you study's data. You do this by selecting one of the options in the following section:

- ○ **General Research Use:** use is permitted for any research purpose
- ○ **Health/Medical/Biomedical Use:** use is permitted for any health, medical, or biomedical purpose
- ○ **Disease-related studies:** use is permitted for research on the specified disease

Please enter one or more diseases

- ○ **Other Use:** permitted research use is defined as follows:

Please specify if selected (max. 512 characters)

2. Choose any additional constraints you need to put on future uses of your data

Then if necessary, you may choose additional terms on your study's data to govern it's use by adding requirements or limitations.

- ☐ No methods development or validation studies (NMDS)
- ☐ Genetic Studies Only (GSO)
- ☐ Publication Required (PUB)
- ☐ Collaboration Required (COL)
- ☐ Ethics Approval Required (IRB)
- ☐ Geographic Restriction (GS-)
- ☐ Non-Profit Use Only (NPU)

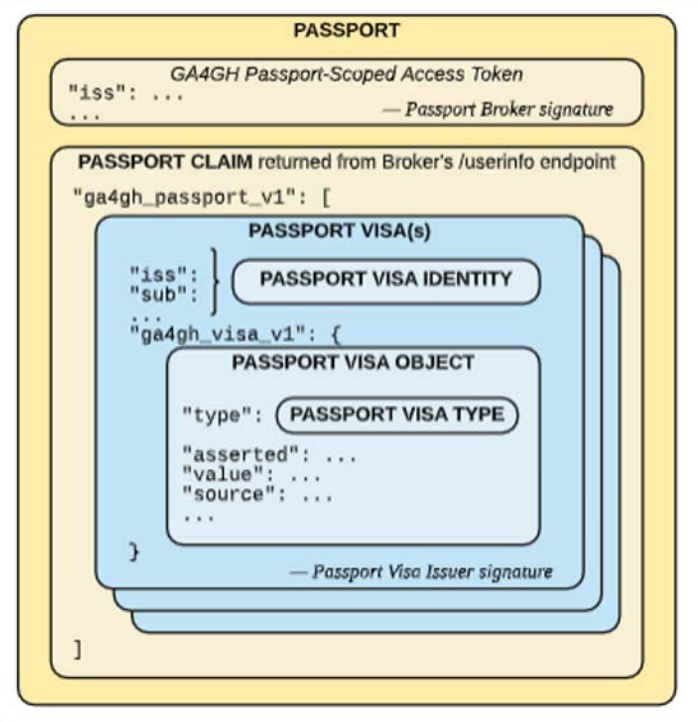
3. Generate your suggested Standardized Data Sharing Language below

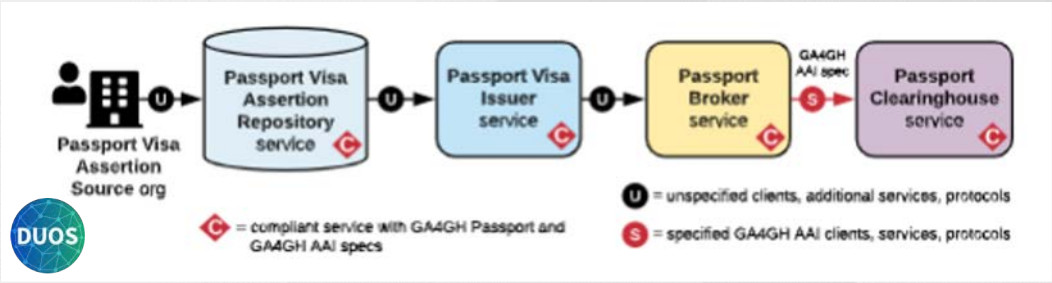
If your selections above are complete, press generate and the suggested consent form text based on the GA4GH Data Use Ontology and Machine Readable Consent Guidance will appear below.

Generate

BROAD INSTITUTE

GA4GH Passports





GA4GH DACReS

The GA4GH Data Access Committee Review Standards (DACReS) policy

Global Alliance for Genomics & Health

Global Alliance for Genomics and Health: Data Access Committee Guiding Principles and Procedural Standards Policy

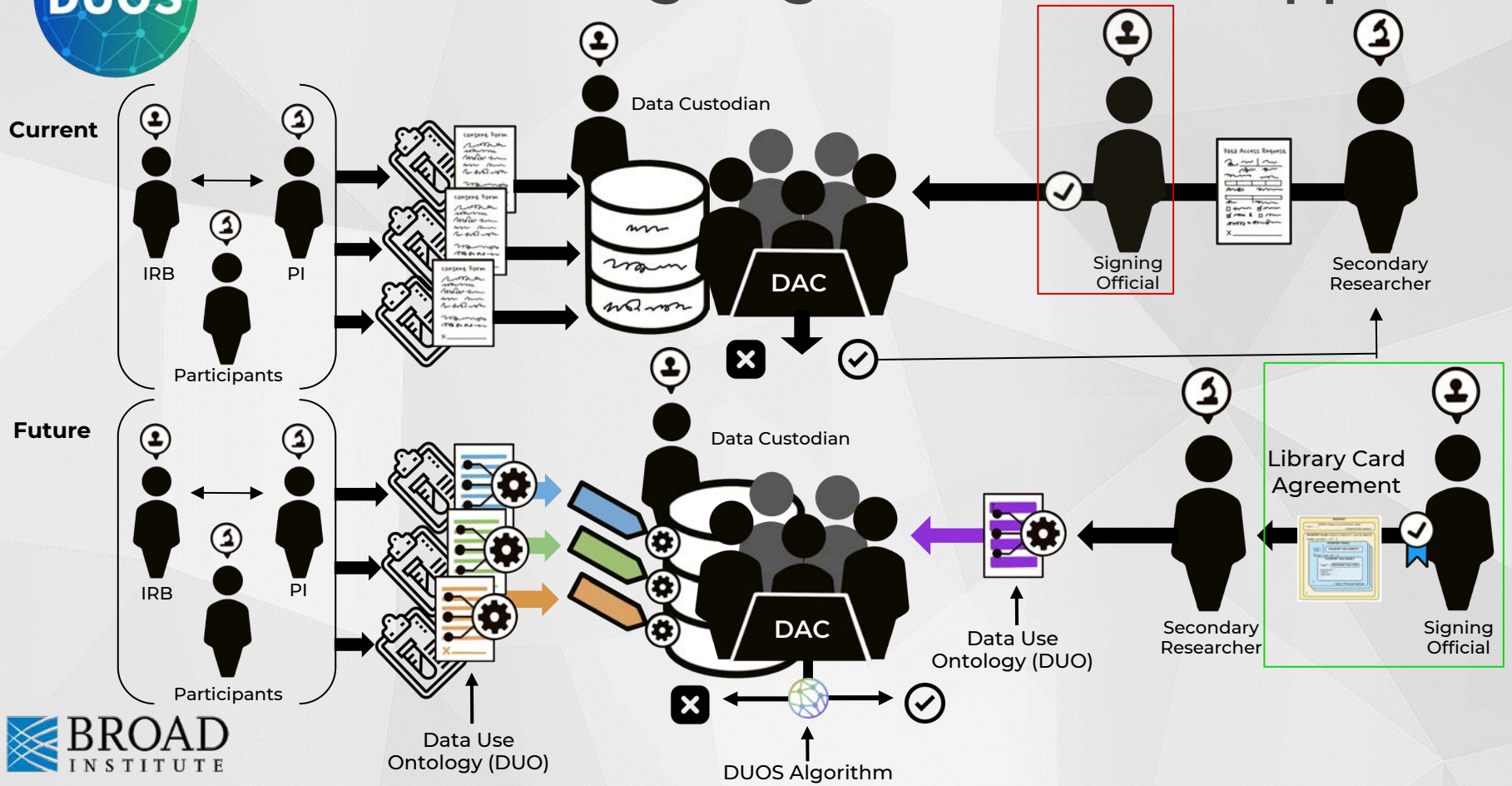
Preamble

This document is the Global Alliance for Genomics and Health's (GA4GH's) Data Access Committee Guiding Principles and Procedural Standards Policy, produced by the GA4GH's Data Access Review Standards (DACReS) working group. It builds on the mission of the GA4GH and provides greater direction for the interpretation of the GA4GH *Framework for Responsible Sharing of Genomic and Health-Related Data* (the "Framework"). It seeks to complement the ongoing work in the GA4GH's Regulatory and Ethics Work Stream (REWS) and Data Use & Researcher Identities (DURI) Work Stream on facilitating responsible sharing of genomic and associated data through the development of harmonized standards and a data use ontology.

Both inappropriately restrictive and overly permissive policies governing access to genomic and associated data challenge underlying principles of research ethics. The former can prevent persons from accessing data that meaningfully advances research. The latter can place the privacy and confidentiality interests of data subjects at increased risk and erode public trust in research. Data access committees (DACs) represent one institutional safeguard charged with applying rules meant to ensure an ethically permissible balance between data protection and accessibility. There are, however, no procedural standards that apply across DACs. The absence of such standards can invite inconsistencies in DAC reviews, compromising their quality and effectiveness. Conversely, standardizing DAC processes can foster trust and mutual recognition, paving the way to greater coordination, collaboration, and delegation between DACs and other oversight bodies that improves the efficiency of data access without sacrificing protection.

BROAD INSTITUTE

Institutional Signing Official's DAR approval



BROAD INSTITUTE

Learn more at

duos.broadinstitute.org

duos-support@broadinstitute.zendesk.com

ga4gh.org

BROAD INSTITUTE

Global Alliance for Genomics & Health