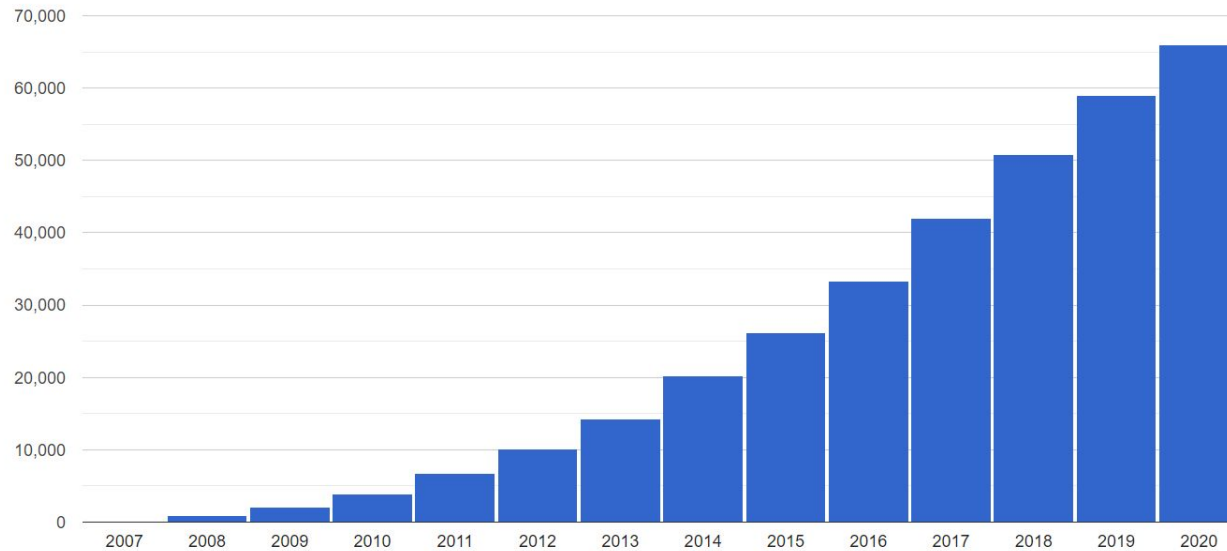




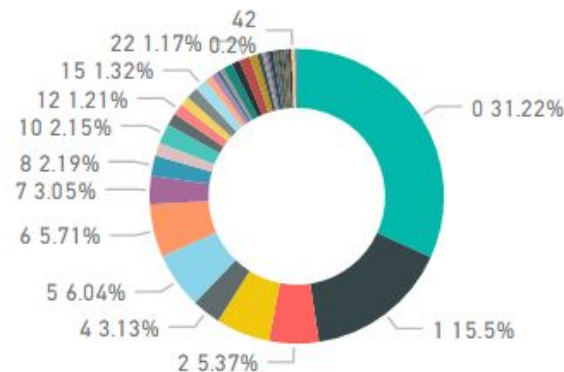
Next Generation Data Access Models

Jonathan Lawson

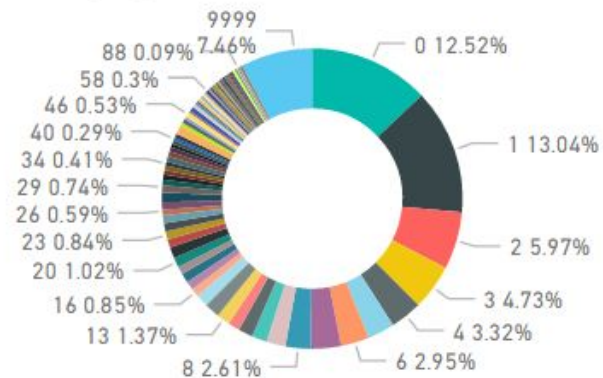
The growing demand of data access requests



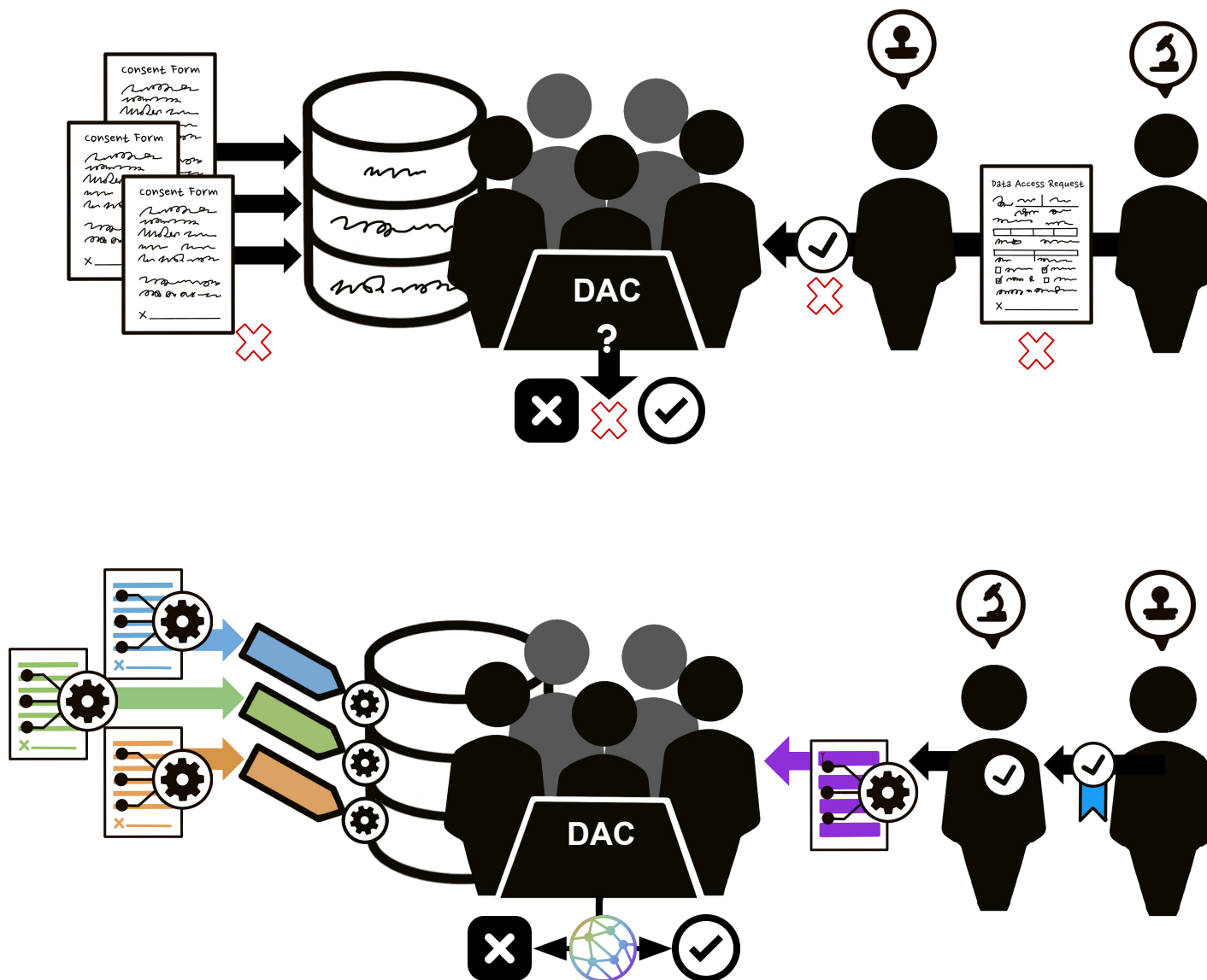
SO ToT in days by # of PIs



DAC ToT in days by # of DARs

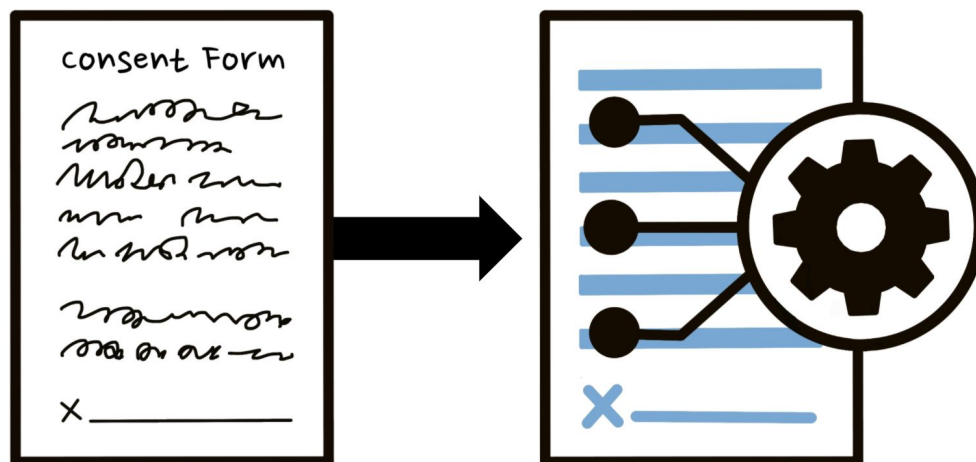


What needs to change?



Unambiguous data use in consents

Most consent forms contain unique, institution-specific language on data use



The **Global Alliance for Genomics and Health (GA4GH)'s Data Use Ontology (DUO)** provides a *structured, machine readable vocabulary* to describe permitted data use.

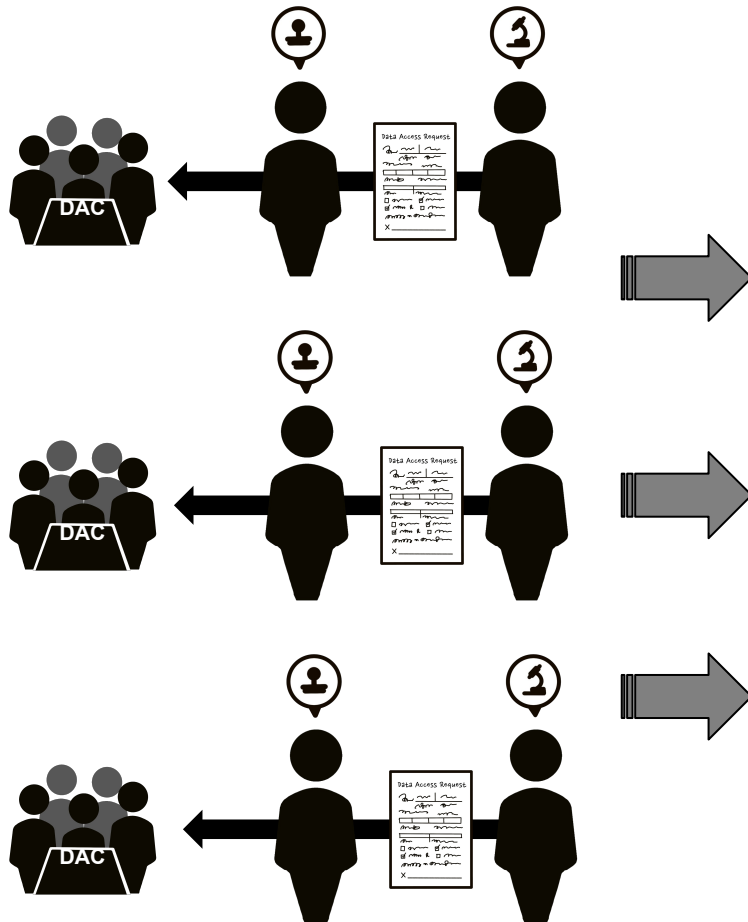
This vocabulary can be leveraged in consent forms to assure that data use permissions expressed by participants are globally interoperable

GA4GH's [Machine Readable Consent Guidance](#) describes how to use DUO in consents

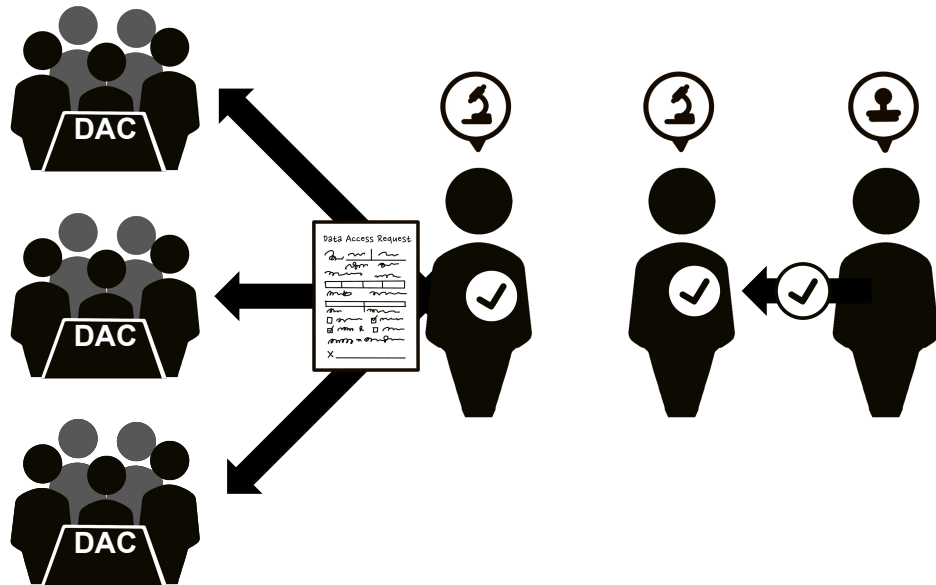
Library Cards: Pre-authorizing Researchers

Currently, Signing Officials for every DAC

- Review each uniquely formatted DAR
- Review/sign a unique access agreement



Adopting *Library Card Agreements* allow Signing Officials to pre-authorize researchers to submit DARs directly to a DAC, multiple times



Adopting a global, *Master Library Card Agreement*, could eliminate the variety of data access agreements in place currently and reduce negotiation time and effort before granting access

DUOS

Data Use Oversight System

Expediting data access for researchers, by facilitating and enhancing data access committee's workflows

What is DUOS and how does it work?

DUOS is a semi-automated data access management service which governs compliant secondary use of human genomics data:

Data Registration



Data Access Request Review



Data Access Request Application



- Interfaces to transform data use restrictions and data access requests to machine-readable code (GA4GH Standard)
- A matching algorithm that checks if data access requests are compatible with data use restrictions
- Interfaces for the Data Access Committee to adjudicate whether structuring and matching has been done appropriately

The DUOS pilots aim to answer these three questions

Can we...

translate official data use designation(s) into *structured* data use limitations?



Data Registration



Data Access Request Review



translate a research use statement to a *structured* research purpose?



Data Access Request Application



evaluate these structured terms with an algorithm to *make the same decision* a human DAC would make when reviewing a DAR ?

State of DUOS pilots with DACs



2018

Broad's DAC uses DUOS to receive and process DARs since early 2018

2019

2 NIH DACs went through internal testing or piloting of DUOS

2020

4 more NIH DACs went through internal testing or piloting of DUOS

Multiple GA4GH Driver Projects tested DUOS

Broad's DAC received over 200 DARs via DUOS

2021

An official NIH pilot of DUOS including up to 6 DACs is slated for 2021