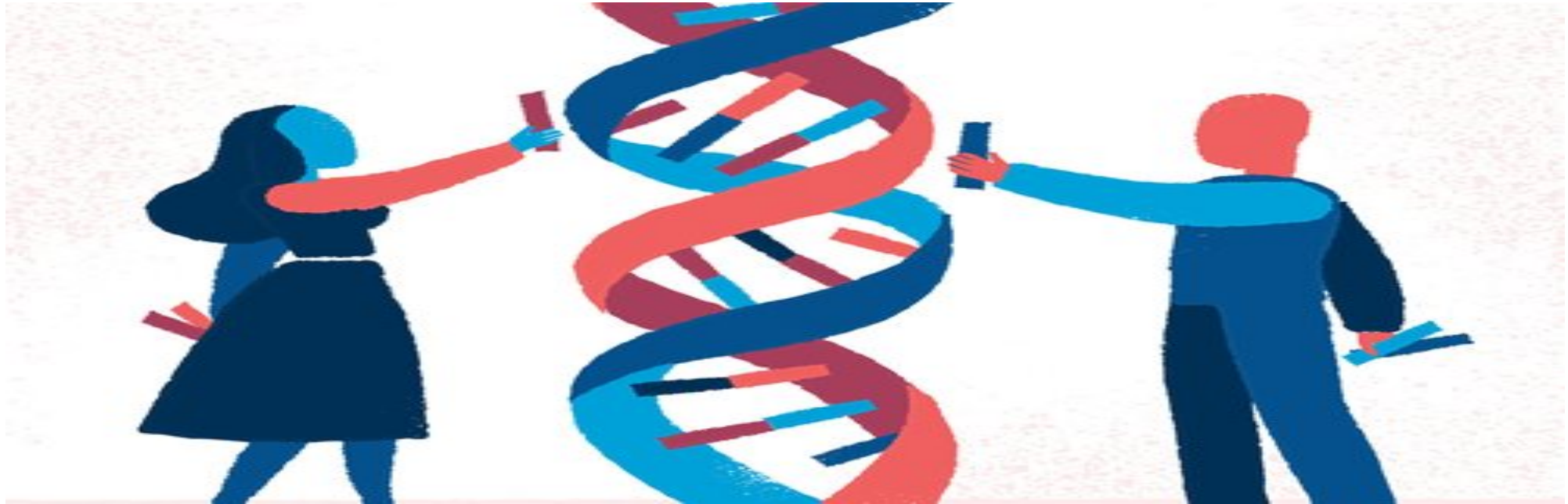


人類基因資料的開放與管理



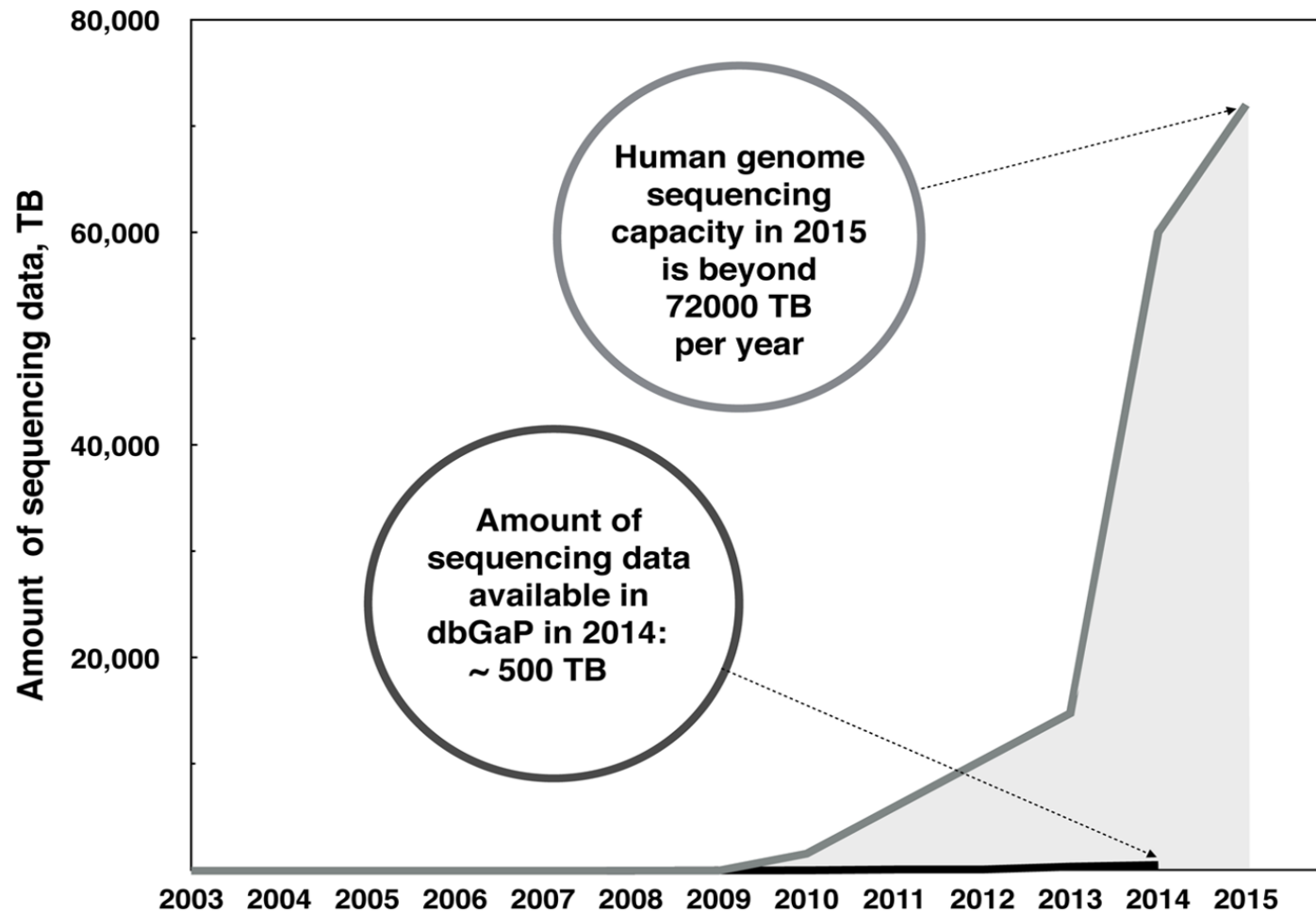
何之行

中研院歐美所

Email: chihho@sinica.edu.tw

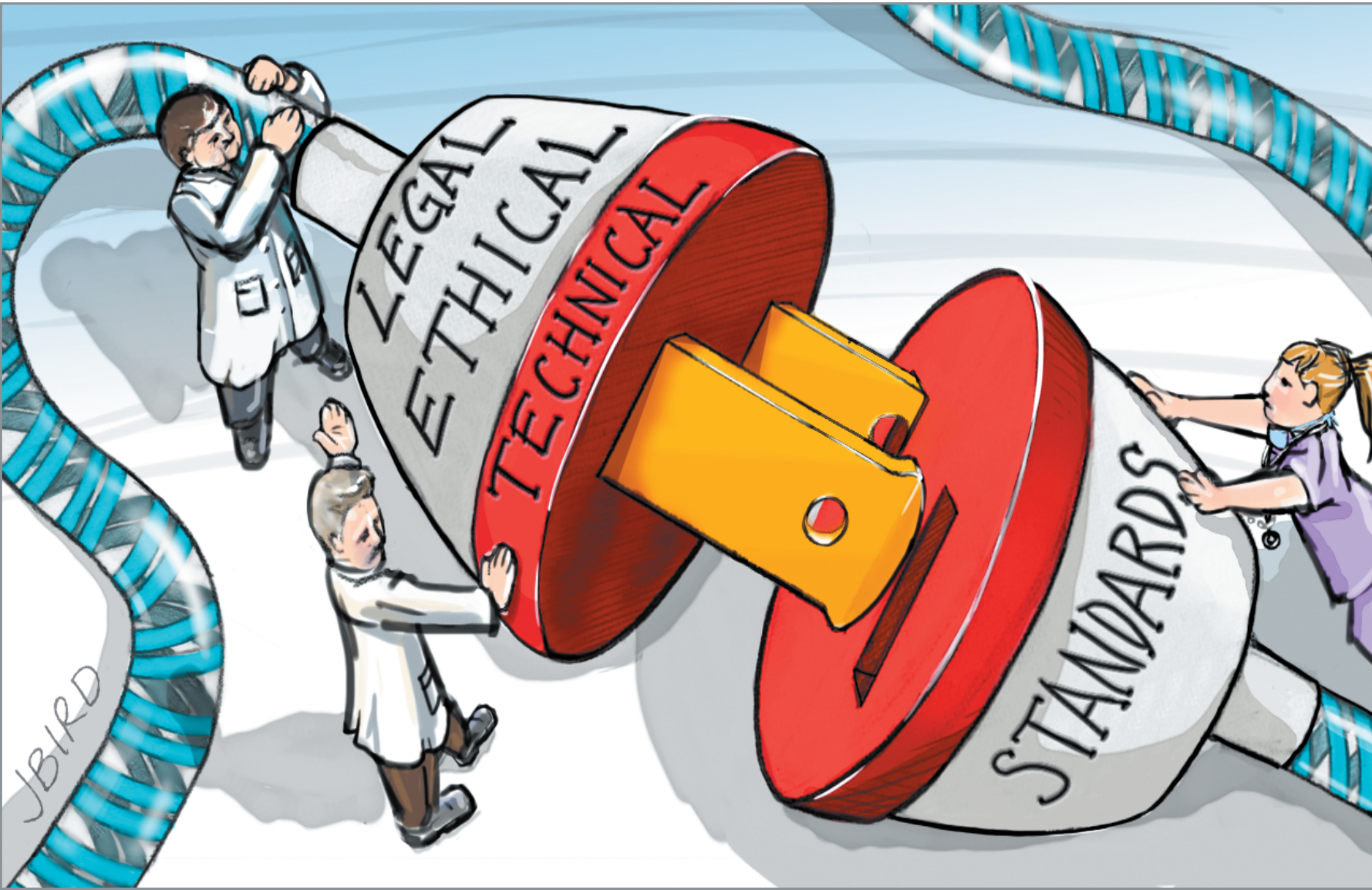
2018 研究資料管理工作坊

Vast amount of genomic data is produced



The Database of Genotypes and Phenotypes (**dbGaP**) started to operate from 2007

Why is data not shared?



Barriers

- Legal
- Ethical
- Technical

Difficulties

FRAGMENTED

Poor visibility of available genomic data



ADMIN BURDEN

Huge overhead to manage data access



CULTURE OF SHARING

Lack of data sharing habits in research culture



National Institutes of Health (NIH) calls for a new Genomic Data Sharing Policy

- **US Federal Government Initiatives**

- Proposed Common Rule revisions (July 2011) supports **broad consent** to maximize utility of bio-specimens and data
- February 2013 White House initiative to increase **access** to publications and the results of federally funded scientific research
- NIH Big Data to Knowledge Initiative (BD2K) supports the **broad use** of digital assets and resources to enhance the utility of biomedical big data and accelerate discovery

Global Alliance for Genomics and Health (GA4GH)

a policy-framing and technical standards-setting organization,
seeking to enable responsible genomic data sharing

international, nonprofit alliance formed in 2013



Global Alliance
for Genomics & Health

500+

Organizational members

2000+

Subscribers

71

Countries

The Comparison of Bermuda Principle (1996), and the GA4GH Genomic Data Sharing Policy I.

International Human Genome Sequencing Consortium (Bermuda 1996)	Global Alliance for Genomics and Health (2017 and beyond)
Number of countries	
5	>70
Sources of data	
Major sequencing laboratories	Research laboratories of many types, clinical laboratories

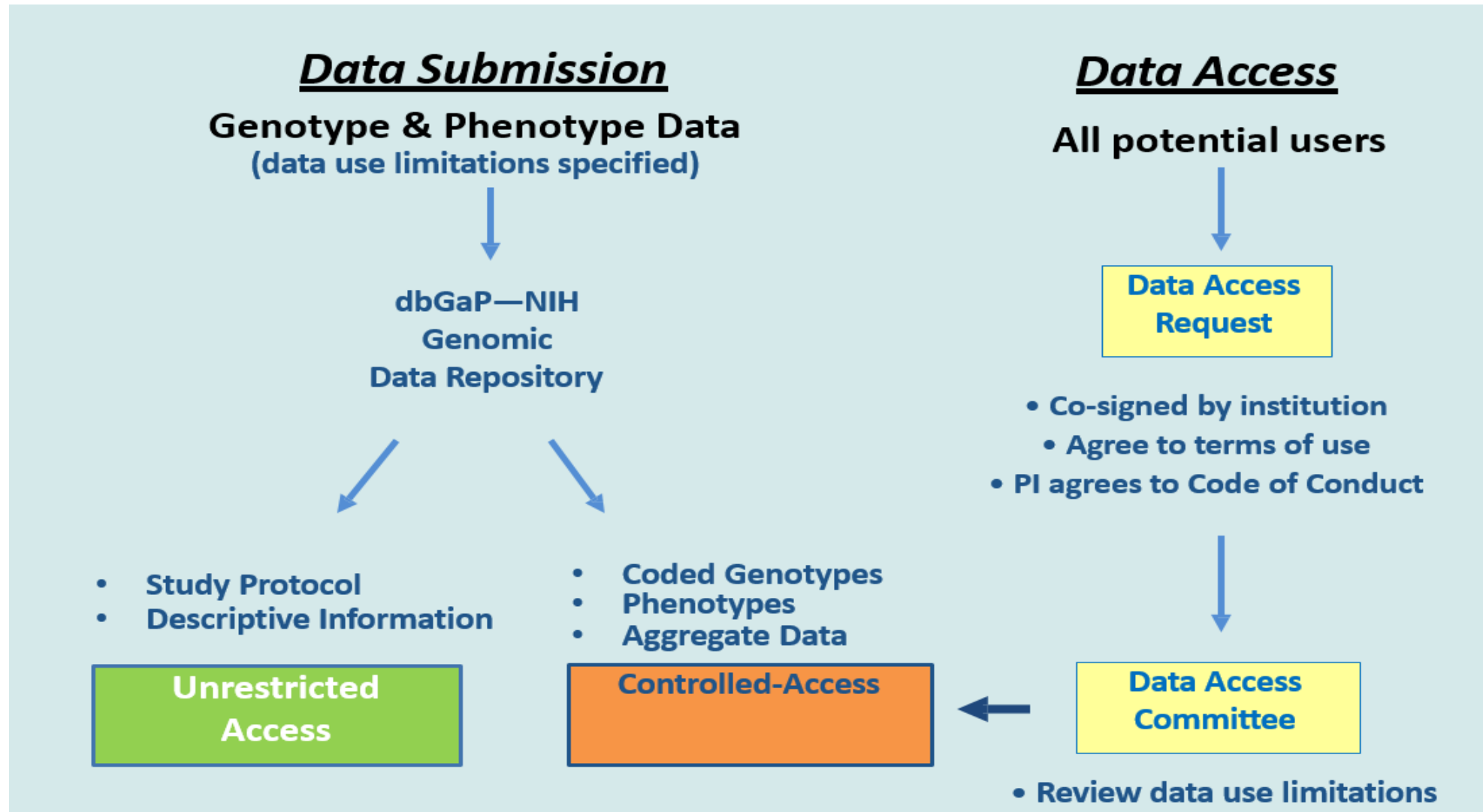
The Comparison of Bermuda Principle (1996), and the GA4GH Genomic Data Sharing Policy II.

Type of data	
International Human Genome Sequencing Consortium (Bermuda 1996)	Global Alliance for Genomics and Health (2017 and beyond)
Assembled fragments of DNA sequence data contributing to an ultimate reference sequence	Data on genomic variants, clinical laboratory data, medical outcomes data, electronic medical records, exposure to environmental factors and toxins
Key differences	
broader array of data types; more databases with diverse formats/organisations	

The Comparison of Bermuda Principle (1996), and the GA4GH Genomic Data Sharing Policy III.

<p>International Human Genome Sequencing Consortium (Bermuda 1996)</p>	<p>Global Alliance for Genomics and Health (2017 and beyond)</p>
<p>Regulatory Framework: Consent, Privacy, confidentiality</p>	
<p>Reference samples from a small number of donors who gave informed consent</p>	<ul style="list-style-type: none">- Sequencing technology can enable re-identification,- Standards for informed consent are changing,- more complex legal framework for sharing personal data

NIH Data Submission and Access



Repositories for Human Genomic Data

- **Unrestricted-access:** data are accessible to anyone via public website (previously referred to as “open access”)
 - e.g., 1000 Genomes project
- **Controlled-access:** data are available to an investigator for a specific project only if certain stipulations are met
 - e.g., dbGaP, Cancer Genomics Hub

NIH's Genomic Data Sharing Policy Overview
<https://osp.od.nih.gov/scientific-sharing/genomic-data-sharing/>



NIH Data Sharing Policy

Informed Consent Expectations I.

PI should ask **participant's consent** for genomic and phenotypic data to be used for future research purposes and to be shared **broadly**

- **Explanation** as to whether the data will be shared via **unrestricted- or controlled-access** repositories
- **If participant does not consent** to broad sharing of data, **he or she may still be enrolled** in the study, but the data may not be shared

NIH's Genomic Data Sharing Policy Overview
<https://osp.od.nih.gov/scientific-sharing/genomic-data-sharing/>



NIH Data Sharing Policy

Informed Consent Expectations II.

For studies using cell lines or clinical specimens created or collected after the effective date:

- **Informed consent** for future research use and **broad** data sharing should have been obtained, **even if samples are de-identified** (studies initiated after Jan 2015)
- **Exception to consent expectation can be granted**, if there are **compelling scientific reasons**

`Compelling scientific reasons`

**No guidance on this exemption yet, but the NIH
promised it**

*`Further guidance on what constitutes compelling
scientific reasons will be made available on the GDS
website and will likely evolve over time`*

Terms and Conditions for Research Use of Controlled-Access Data

- The data can be used only for approved research

it is prohibited

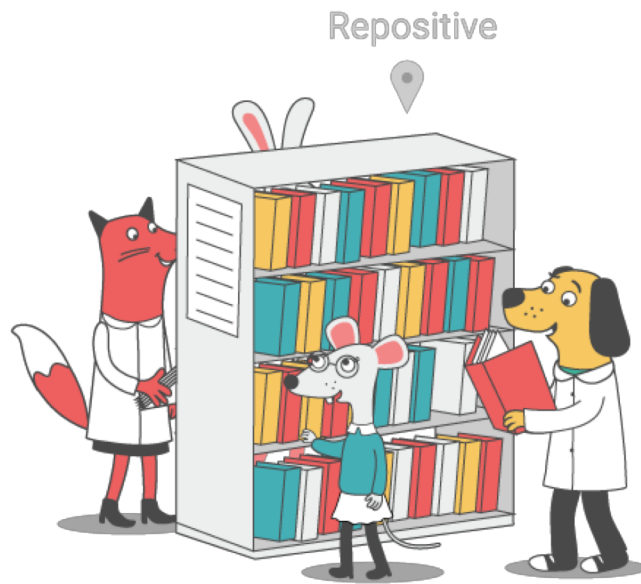
- to **re-identify** individual participants from whom the data were obtained
- to **sell** any of the data obtained from NIH-designated data repositories
- **sharing** any of the data obtained from controlled-access NIH-designated data

FAIR data: Findable, Accessible, Interoperable, Reusable

MAKE DATA
DISCOVERABLE

SIMPLIFY
WORKFLOWS

CONTRIBUTE TO
COMMUNITY



Thank you for your attention!

