



# Archiving, preserving, sharing



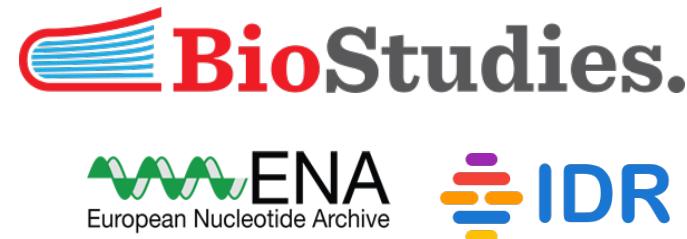
CENTRE FOR  
DIGITAL LIFE  
NORWAY

Korbinian Bösl  
Data management coordinator  
ELIXIR Norway/Digital Life Norway



# NeLS

Norwegian e-Infrastructure for Life Sciences



Data life cycle +

Your role +

Your domain +

Your problem -

Compliance monitoring

Data analysis

Data management plan

Data organisation

Data protection

**Data publication**

Data quality

Data storage

Data transfer

Identifiers

Licensing

Documentation and metadata

Sensitive data

All tools and resources

Tool assembly +



Link to RDMkit: <https://rdmkit.elixir-europe.org/>

# Why should I deposit my data?



Increased Visibility (SEO) → 25% more citations

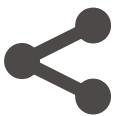
# Why should I deposit my data?

-  Increased Visibility (SEO) → 25% more citations
-  Value added Databases → more citations

# Why should I deposit my data?

-  Increased Visibility (SEO) → 25% more citations
-  Value added Databases → more citations
-  FAIRification – Funding requirements ✓✓

# Why should I deposit my data?

-  Increased Visibility (SEO) → 25% more citations
-  Value added Databases → more citations
-  FAIRification – Funding requirements ✓✓
-  Safe money on storage

# Where should I deposit my data?

Preferably somewhere  
where most things are taken care of?



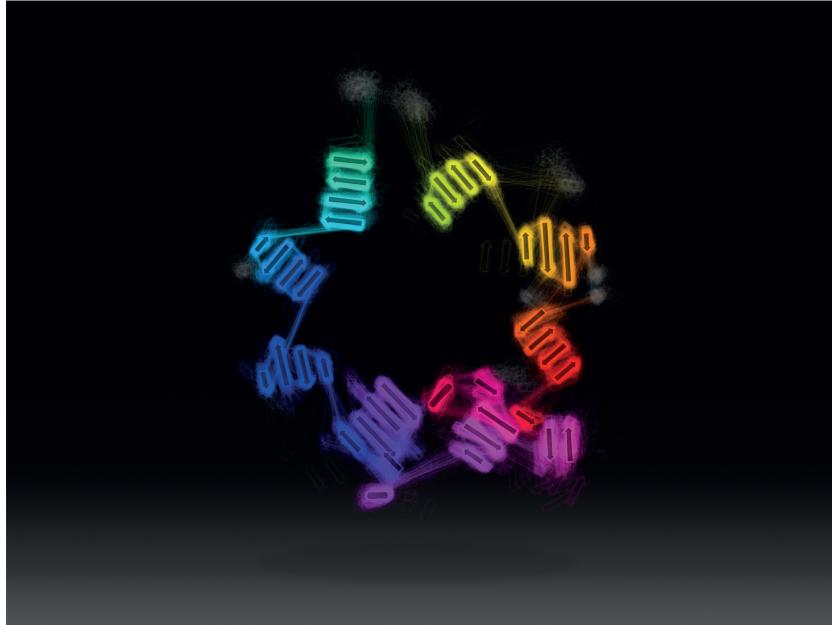
# How to choose?

PRINT ISSN: 0305-1048  
ONLINE ISSN: 1362-4962

## Nucleic Acids Research

VOLUME 49 DATABASE ISSUE JANUARY 8, 2021

<https://academic.oup.com/nar>



molecular biology  
1641 databases

OXFORD  
UNIVERSITY PRESS

Open Access

No barriers to access – all articles freely available online

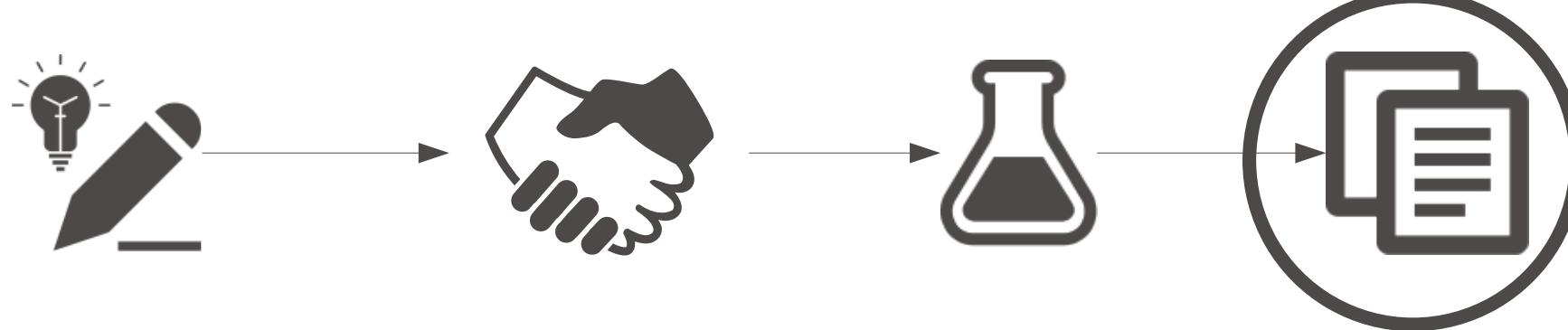


Rigden, D. J. & Fernández, X. M. licensed under Creative Commons Attribution 4.0 International License

The 2021 Nucleic Acids Research database issue and the online molecular biology database collection. Nucleic Acids Research 49, D1–D9 (2021).



**nature**



**Deposition**

# Where should I deposit my data?



# Where should I deposit my data?



# Where should I deposit my data?

Fits?

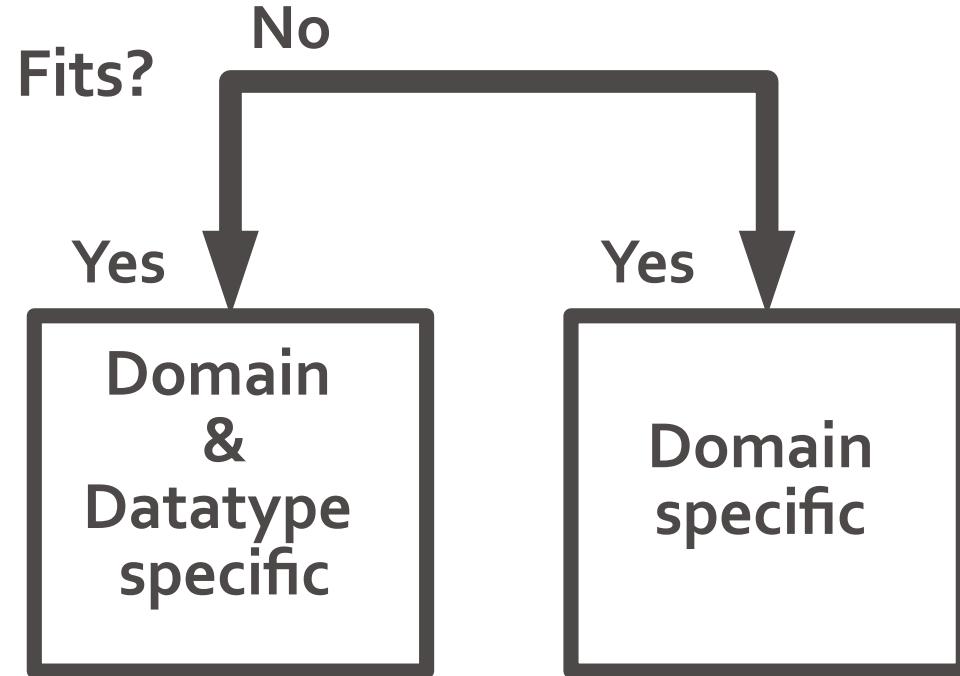


Yes

Domain  
&  
Datatype  
specific



# Where should I deposit my data?

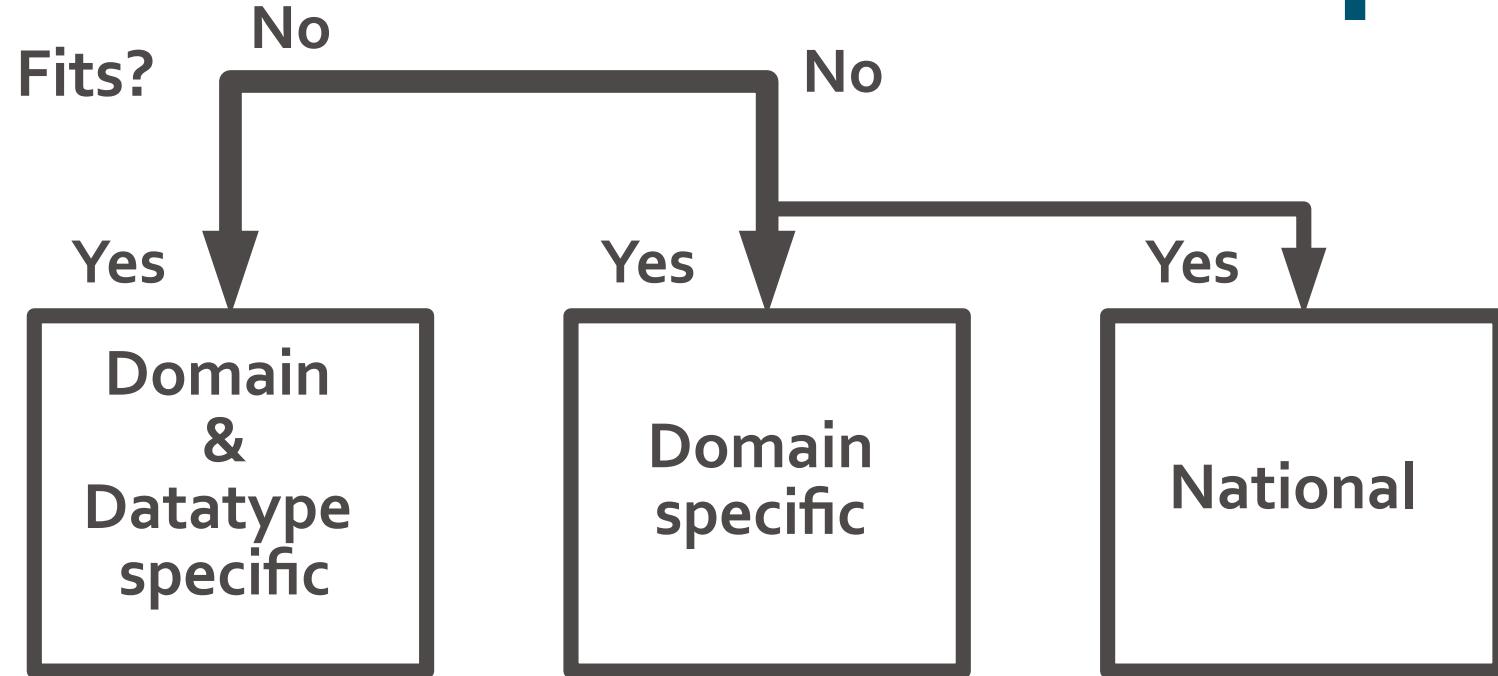


 ENA  
European Nucleotide Archive

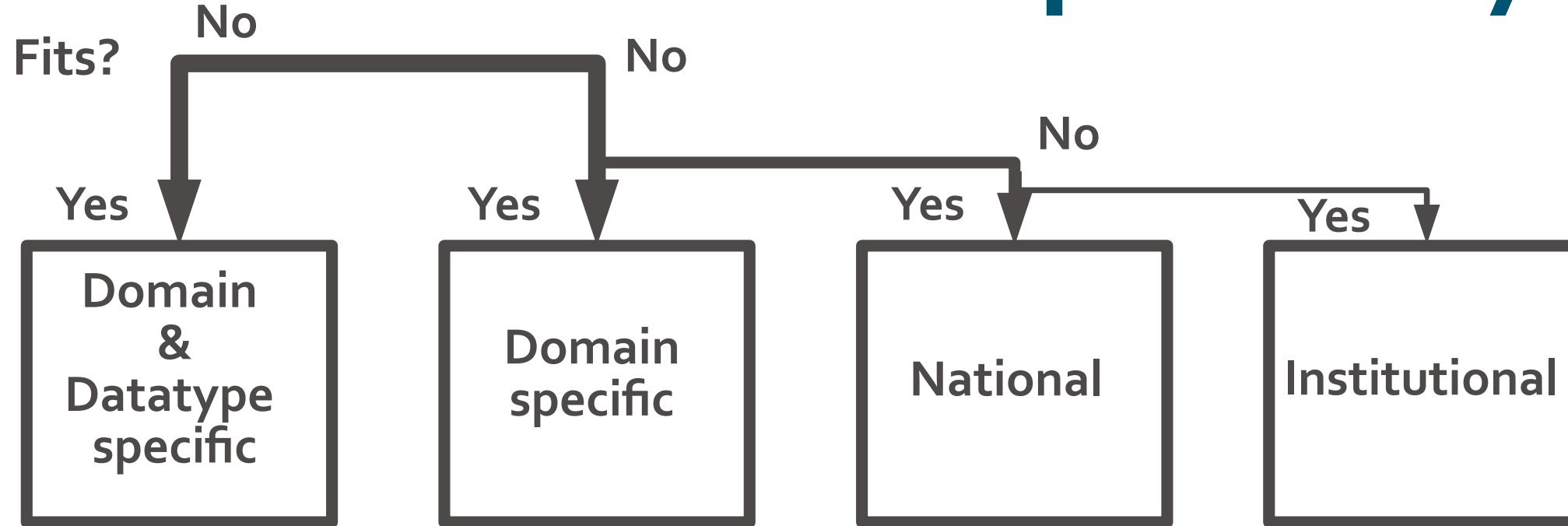
 BioStudies.

 SRA  
NCBI

# Where should I deposit my data?



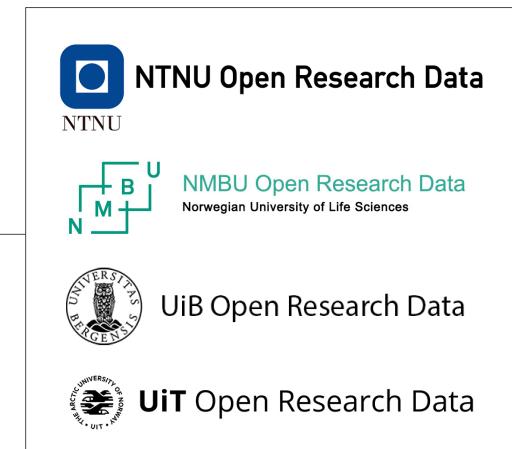
# Where should I deposit my data?



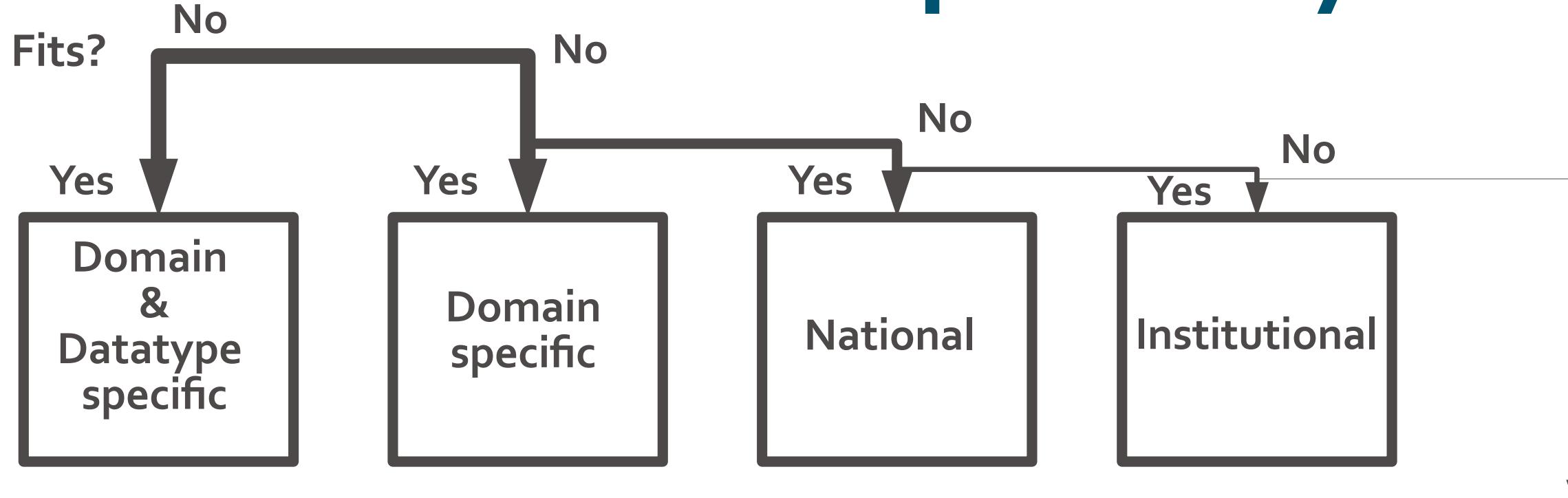
NIRD RESEARCH DATA ARCHIVE



>DataVERSENO  
DataVERSE Network Norway



# Where should I deposit my data?



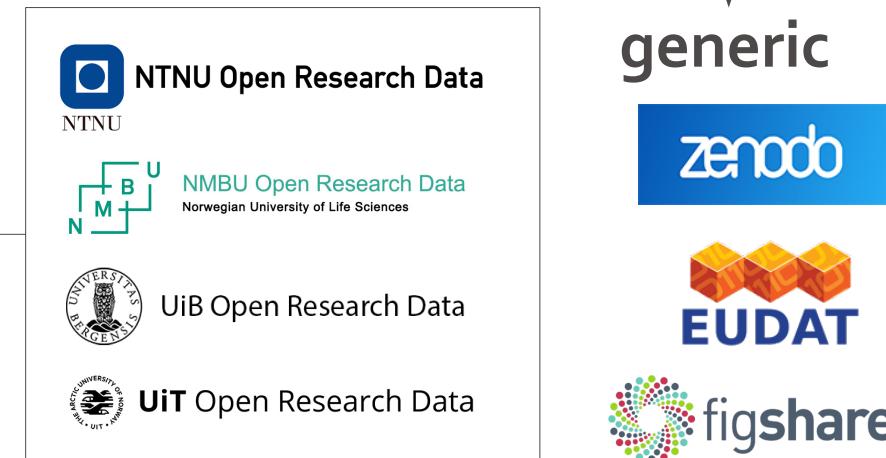
 ENA  
European Nucleotide Archive

 SRA  
NCBI

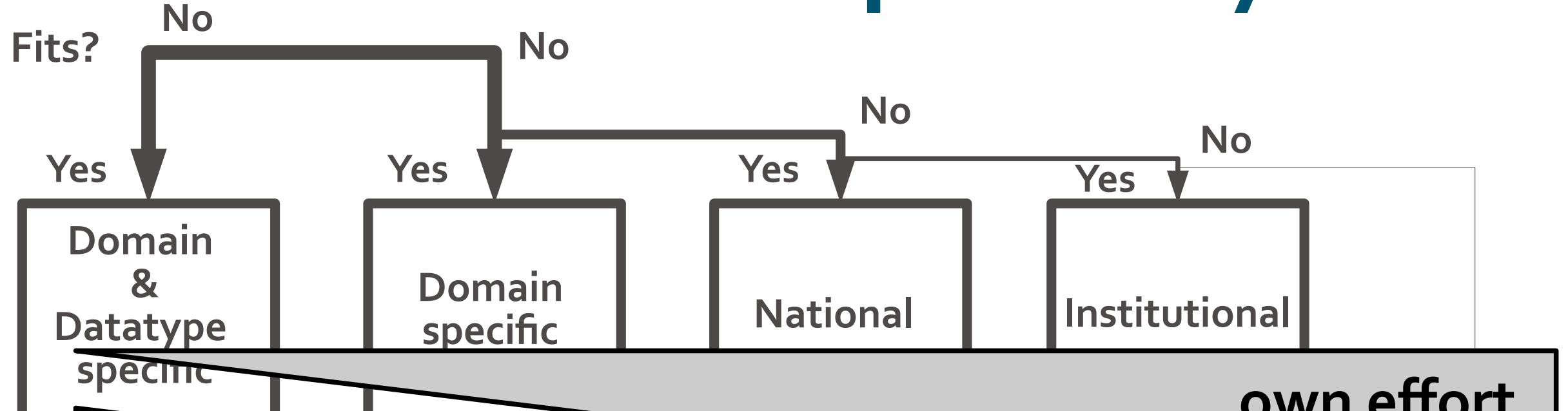
 BioStudies.

NIRD RESEARCH DATA ARCHIVE

 DataverseNO  
Dataverse Network Norway



# Where should I deposit my data?



NCBI

Standardization

DataverseNO  
Dataverse Network Norway

NTNU Open Research Data  
Norwegian University of Life Sciences

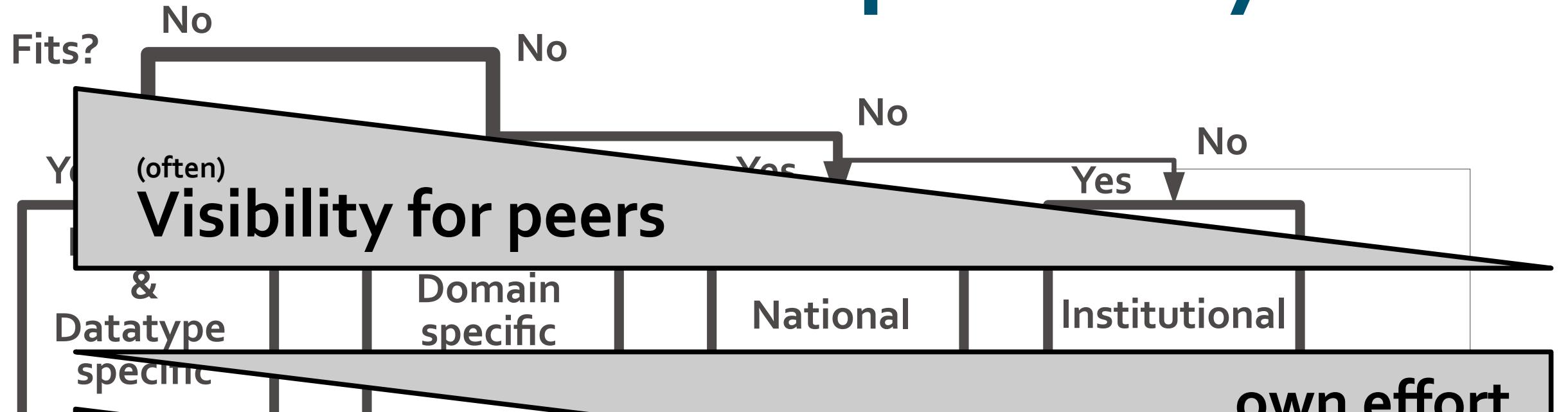
UiB Open Research Data  
UiT Open Research Data

EUDAT

figshare

zenodo

# Where should I deposit my data?



NCBI

Standardization

NTNU Open Research Data

zenodo

DataverseNO  
Dataverse Network Norway

NM  
Norwegian University of Life Sciences



UiB Open Research Data

UiT Open Research Data

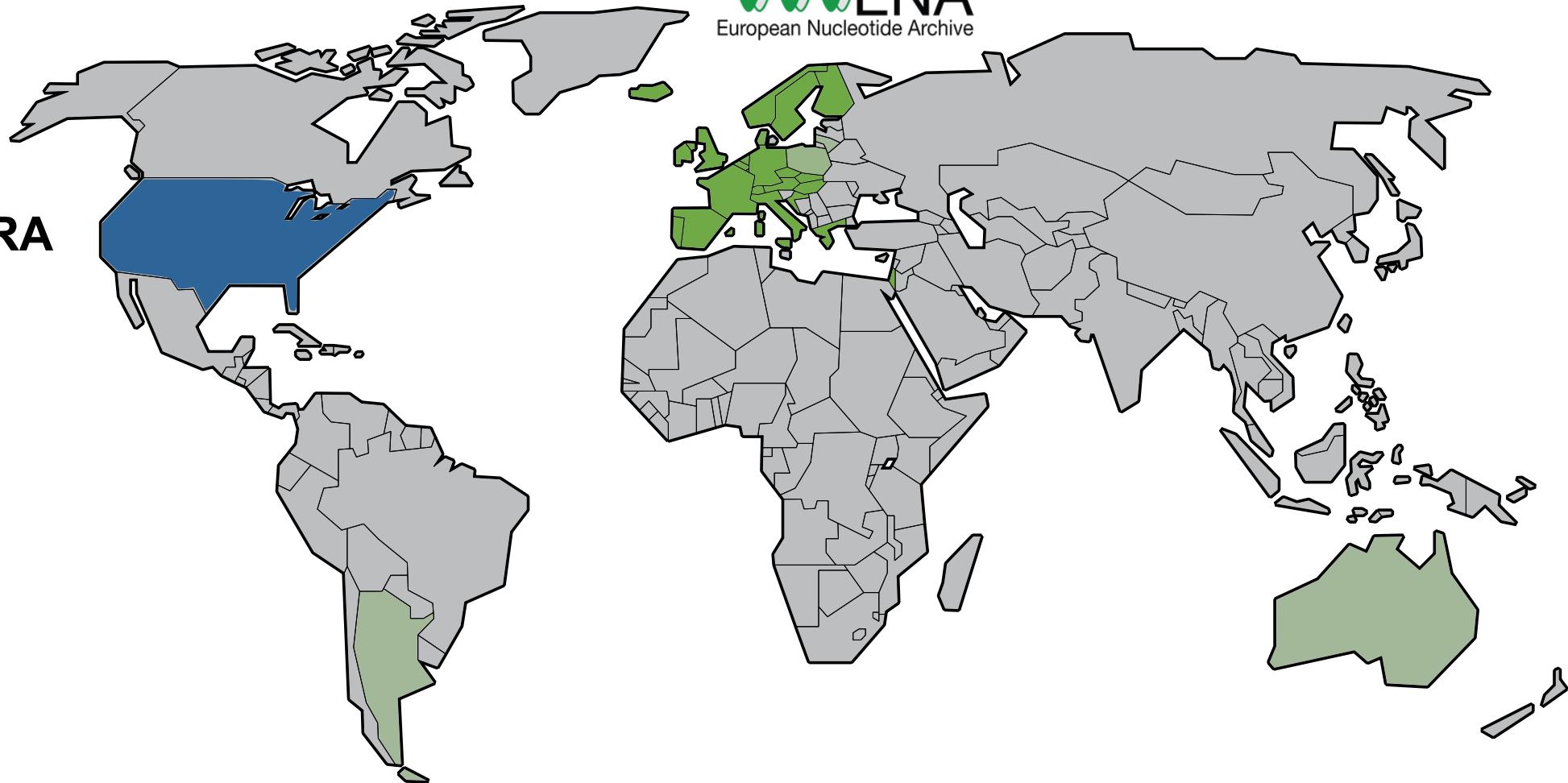
EUDAT

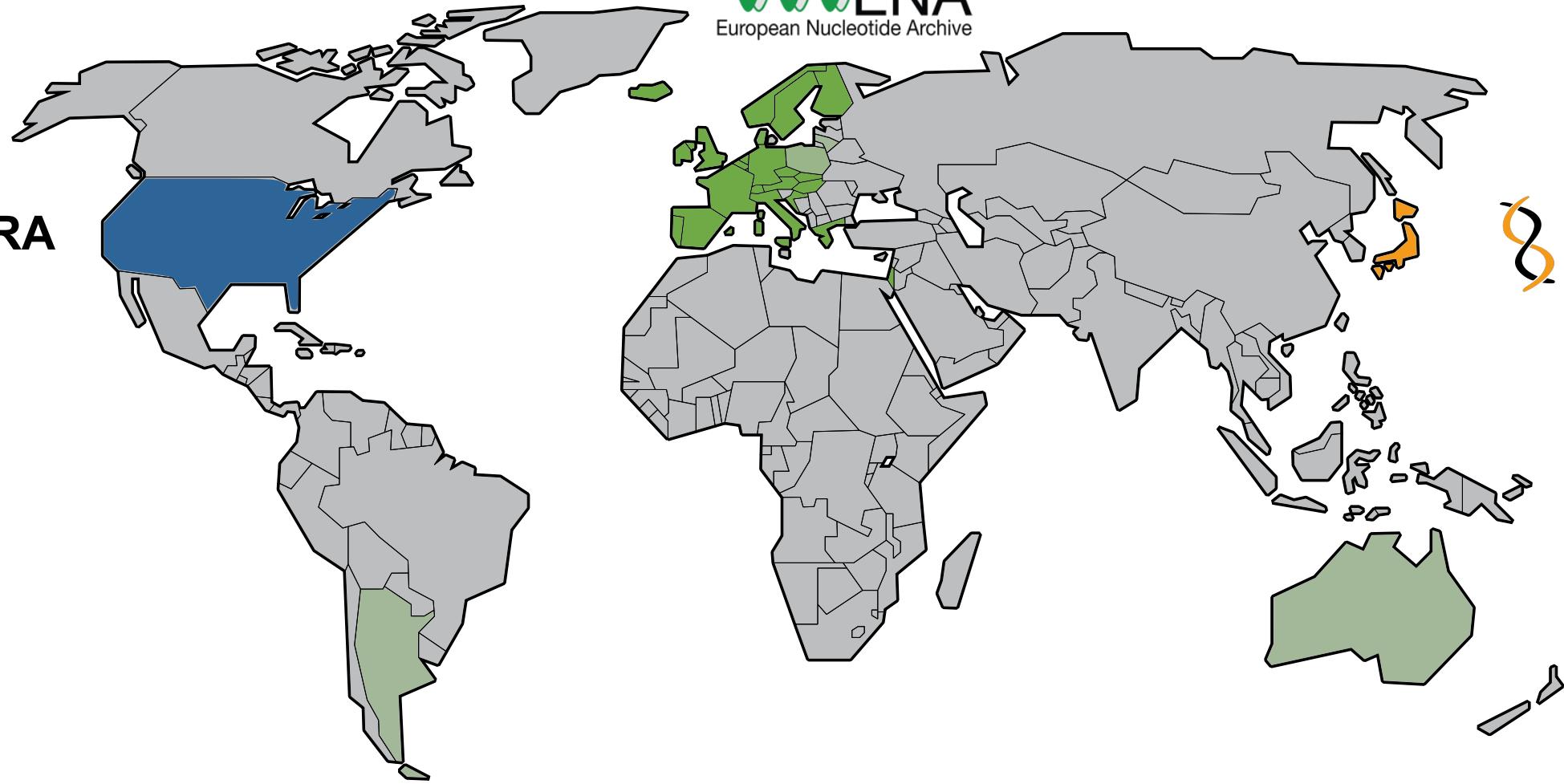
figshare

# Multiple Repositories – similar datatypes



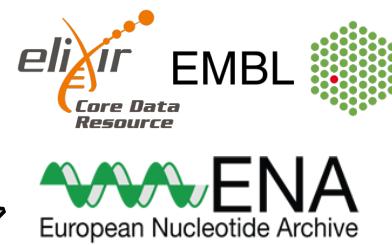
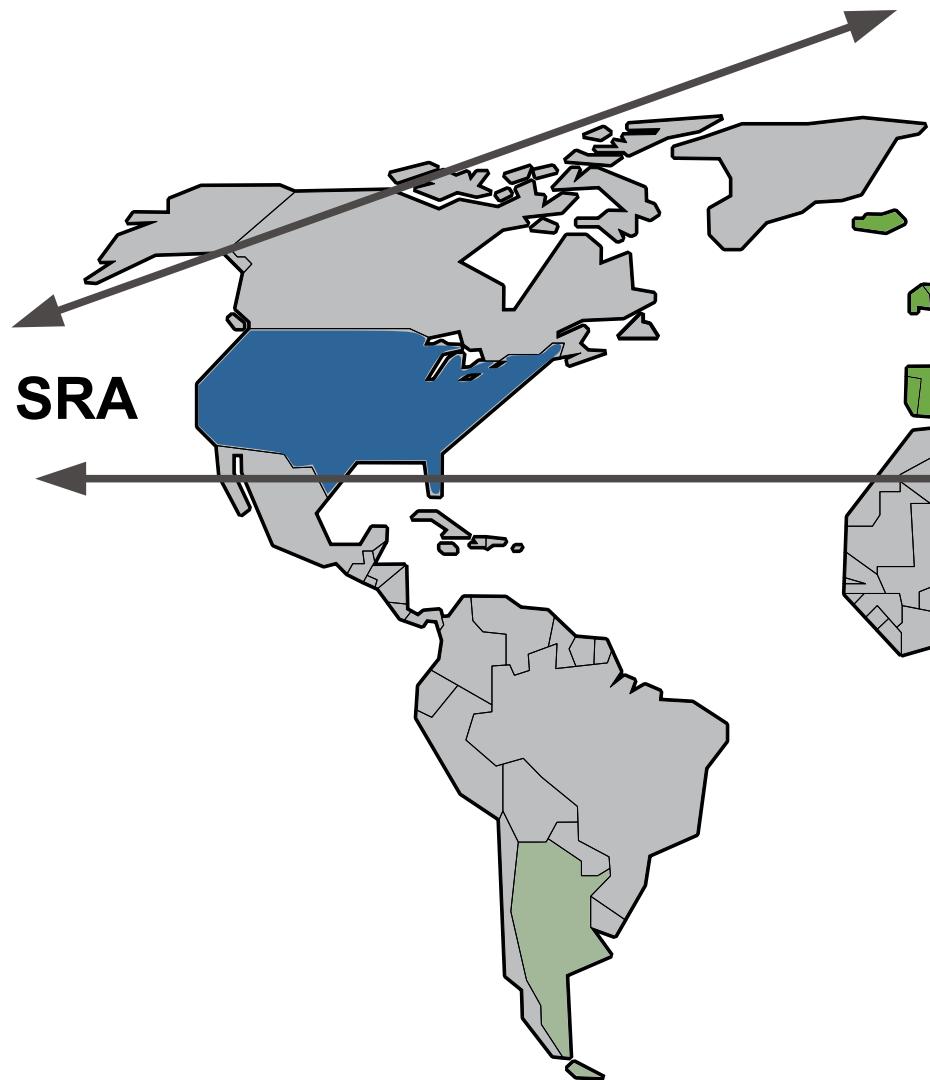








SRA

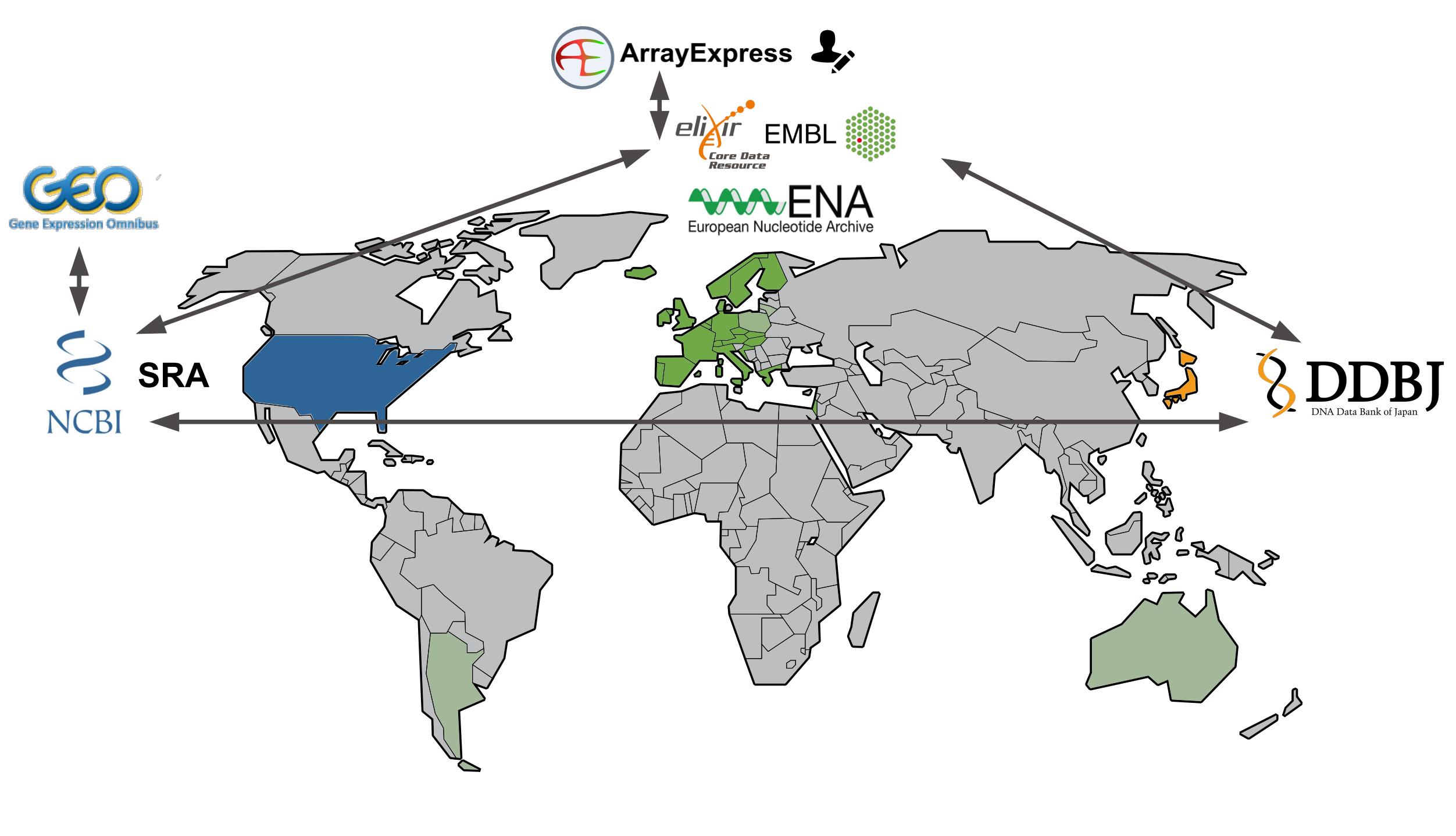


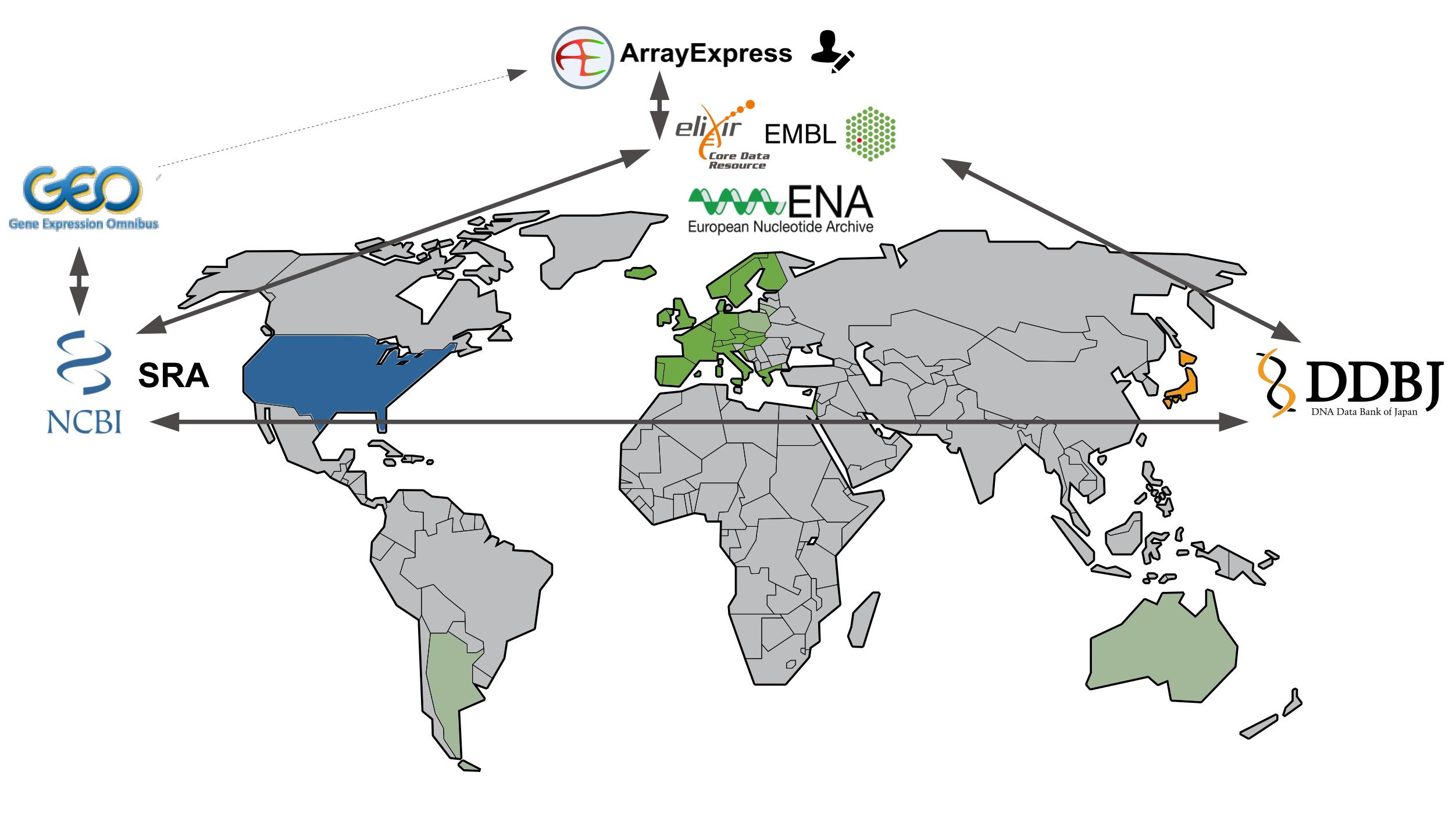
ENA

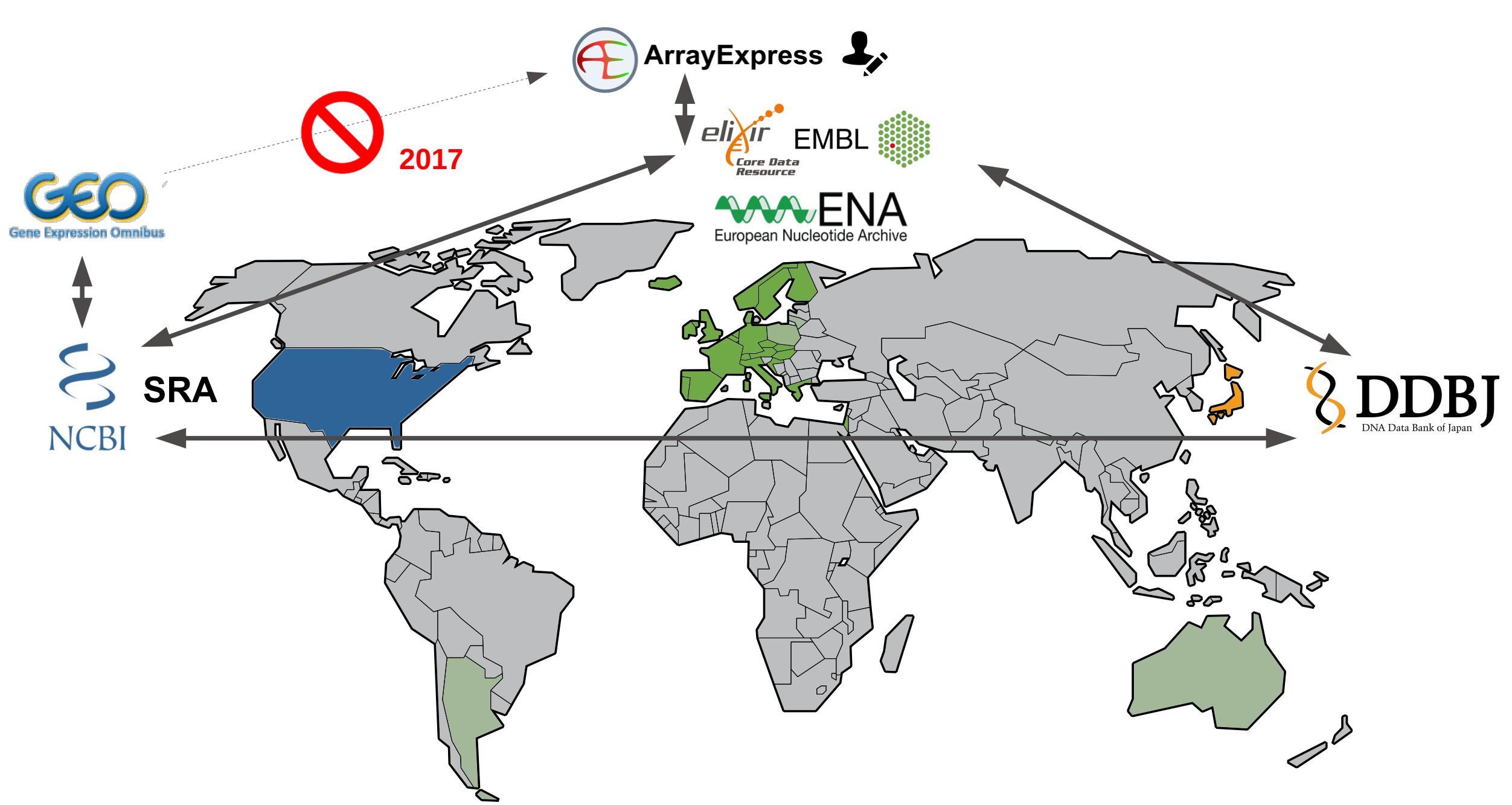
European Nucleotide Archive



DNA Data Bank of Japan



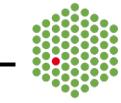






EMBL



A world map where different regions are highlighted in various colors: North America is blue, Europe and parts of Africa and the Middle East are green, China and parts of South Asia are red, Australia is light green, and South America and some countries in Africa and the Middle East are grey.

Proteome  
Xchange



# Where should I deposit my data?



# Where should I deposit my data?

Recognized & Recommended



<https://elixir-europe.org/platforms/data/core-data-resources>

<https://elixir-europe.org/platforms/data/elixir-deposition-databases>

# Where should I deposit my data?

Recognized & Recommended

Free of charge



# Where should I deposit my data?

Recognized & Recommended

Free of charge

Quick upload



# Where should I deposit my data?



Recognized & Recommended

Free of charge

Quick upload

Not operated by single university/government

# Where should I deposit my data?



Recognized & Recommended

Free of charge

Quick upload

Not operated by single university/government

Guided submission process

# Where should I deposit my data?



Recognized & Recommended

Free of charge

Quick upload

Not operated by single university/government

Guided submission process

Generous embargo regulation (2yrs → )

# Where should I deposit my data?



Recognized & Recommended

Free of charge

Quick upload

Not operated by single university/government

Guided submission process

Generous embargo regulation (2yrs → )

API access (upload from NeLS planned)

# Where should I deposit my data?



# Where should I deposit my data?



# Where should I deposit my data?



# Where should I deposit my data?



ArrayExpress



**PRIDE**

*PRoteomics IDENTifications Database*



BioImage Archive



EMPIAR



IDR



BioModels



MetaboLights

# Where should I deposit my data?



EMBL-EBI



BioImage Archive



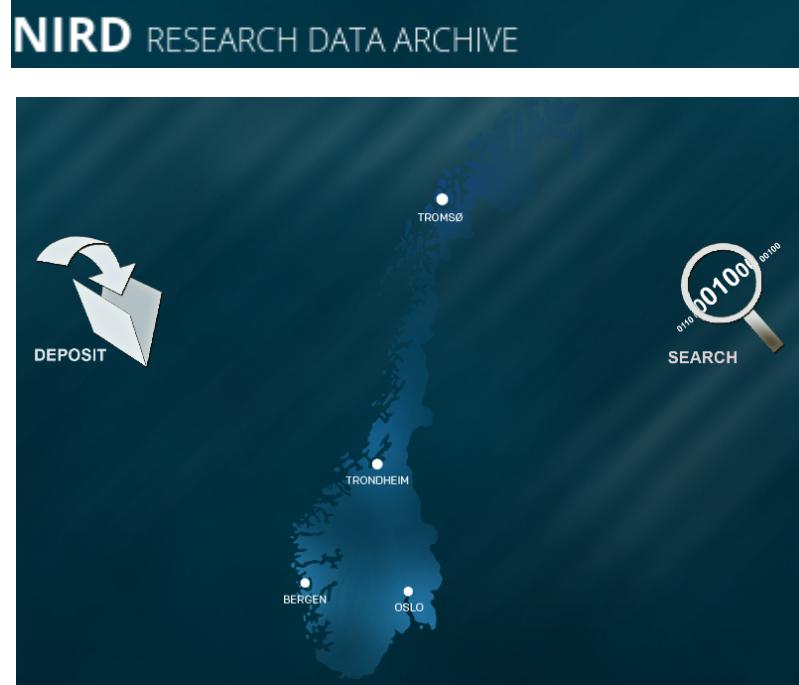
Data Submission Wizard

<https://www.ebi.ac.uk/submit/>

PRIDE  
Proteomics IDEntifications Database



# NIRD archive



**National, free depositing repository**

**Domain agnostic**

**Dublin-core metadata standard**

**DOI accessible**

**Supports machine readable metadata harvesting**

# Generic archives



Less standard metadata → Re-usability ↓



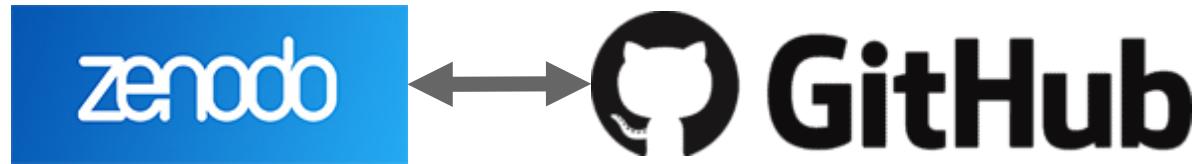
Domain agnostic → Findability ↓

Guided submission process



Commercial actors: Size limitations, ...

# Archiving code/scripts



**Good integration - easy to implement**

**Archival of each release**

**beta solutions for gitlab**

# What about sensitive data?



EUROPEAN  
GENOME-PHENOME  
ARCHIVE

[ega-archive.org](http://ega-archive.org)

Central metadata accessibility

Secure storage

Implemented data access committees

Norwegian EGA

(for data that can not leave the country)



# What about sensitive data?

Central metadata accessibility

Secure storage



Ensure your consent form & REC approvals address committees  
allow controlled access archiving!

Norwegian EGA  
(for data that can not leave the country)



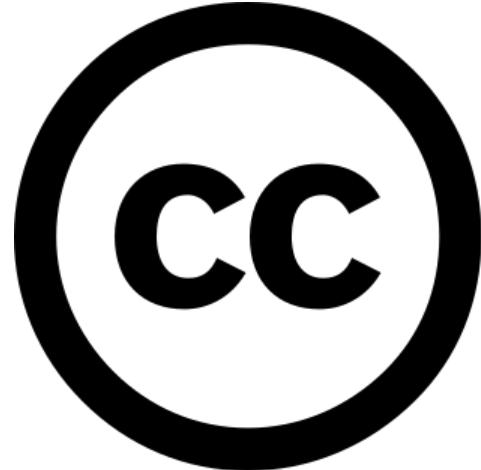
# NSD archive

Domain agnostic

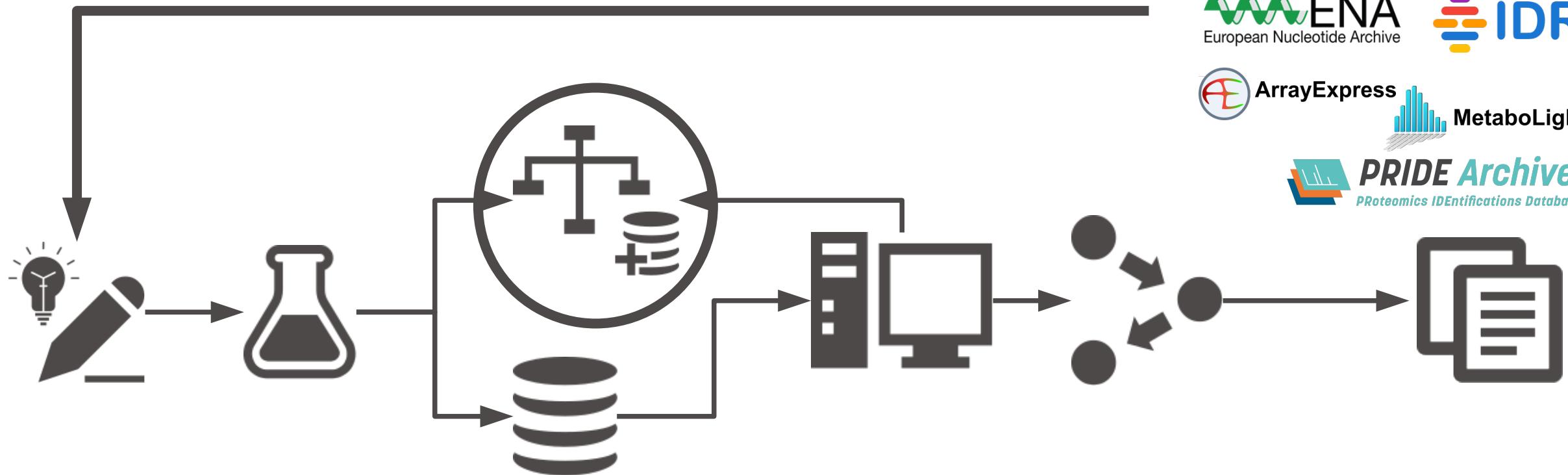
DOI accessible

Implemented data access committees

Main data type: Questionnaires, ...



Except where otherwise noted, this work is licensed under:  
<https://creativecommons.org/licenses/by/4.0/>



 ENA  
European Nucleotide Archive

 IDR

 ArrayExpress

 MetaboLights

 PRIDE Archive  
PRoteomics IDEntifications Database

 EUROPEAN  
GENOME-PHENOME  
ARCHIVE