



# Data Deposition Repositories



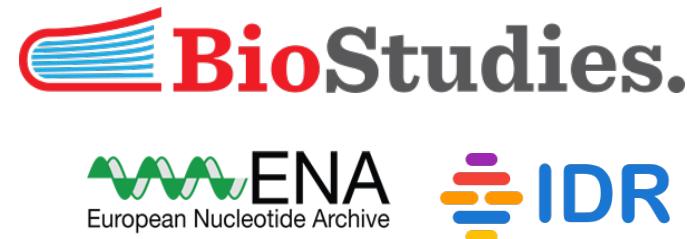
CENTRE FOR  
DIGITAL LIFE  
NORWAY

Korbinian Bösl  
Data manager ELIXIR Norway  
21 February 2020



# NeLS

Norwegian e-Infrastructure for Life Sciences



# Why should I deposit my data?



Increased Visibility (SEO) → more citations

# Why should I deposit my data?

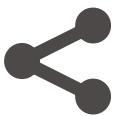


Increased Visibility (SEO) → more citations

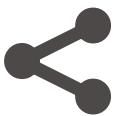


Value added Databases → more citations

# Why should I deposit my data?

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

# Why should I deposit my data?

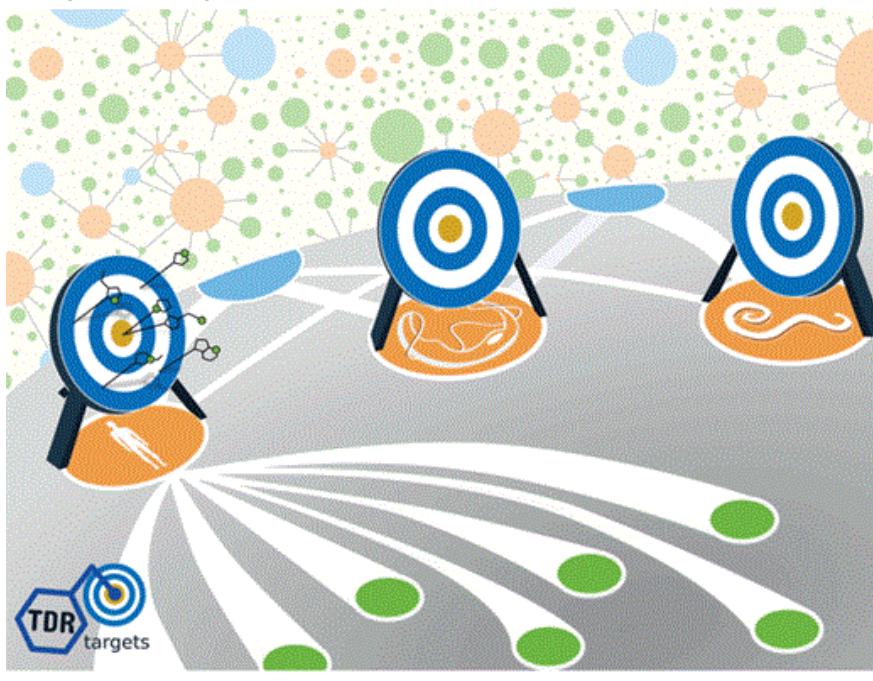
-  Increased Visibility (SEO) → more citations
-  Value added Databases → more citations
-  FAIRification – Funding requirements ✓✓
-  Safe money on storage

# Where should I deposit my data?

Nucleic Acids  
Research

VOLUME 48 DATABASE ISSUE JANUARY 8, 2020

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



OXFORD  
UNIVERSITY PRESS

Open Access

No barriers to access – all articles freely available online



Rigden, D. J. & Fernández, X. M.

The 27th annual Nucleic Acids Research database issue and molecular biology database collection. Nucleic Acids Res 48, D1–D8 (2020).

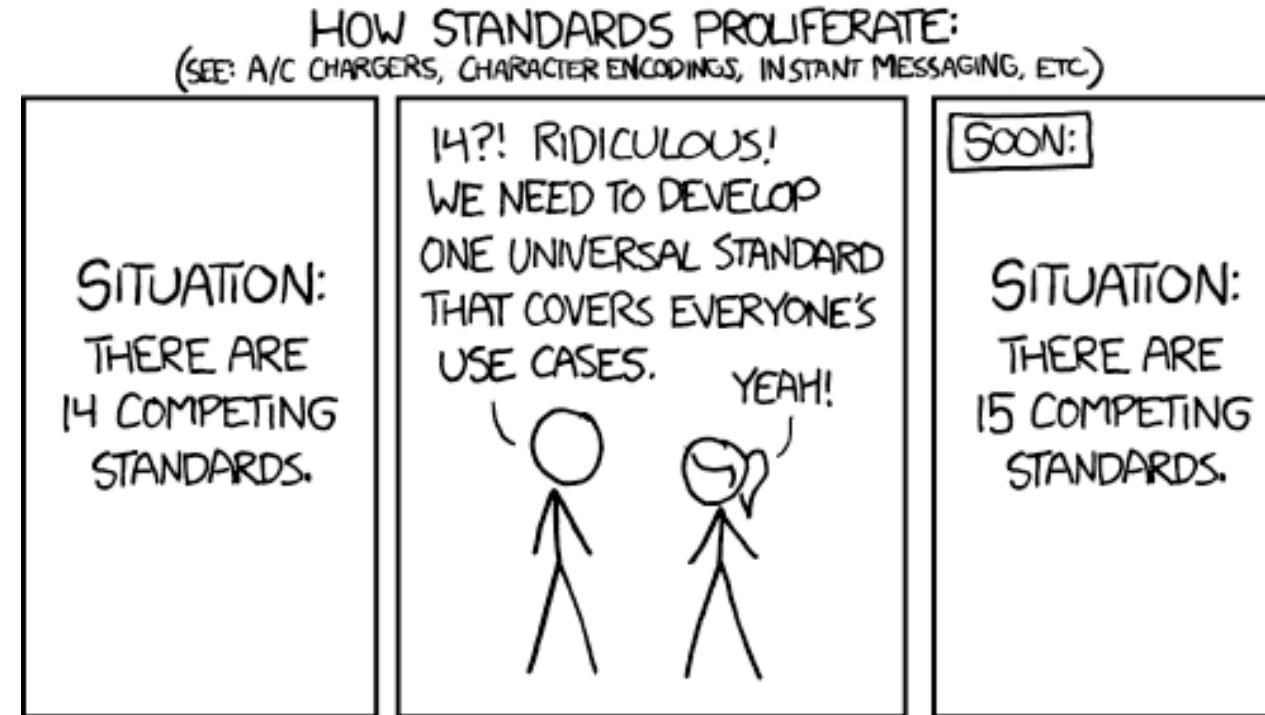
molecular biology  
1637 databases

# Where should I deposit my data?



# Meta data standards

# Community standards vs. formal ISO standards



# Where should I deposit my data?

Fits?

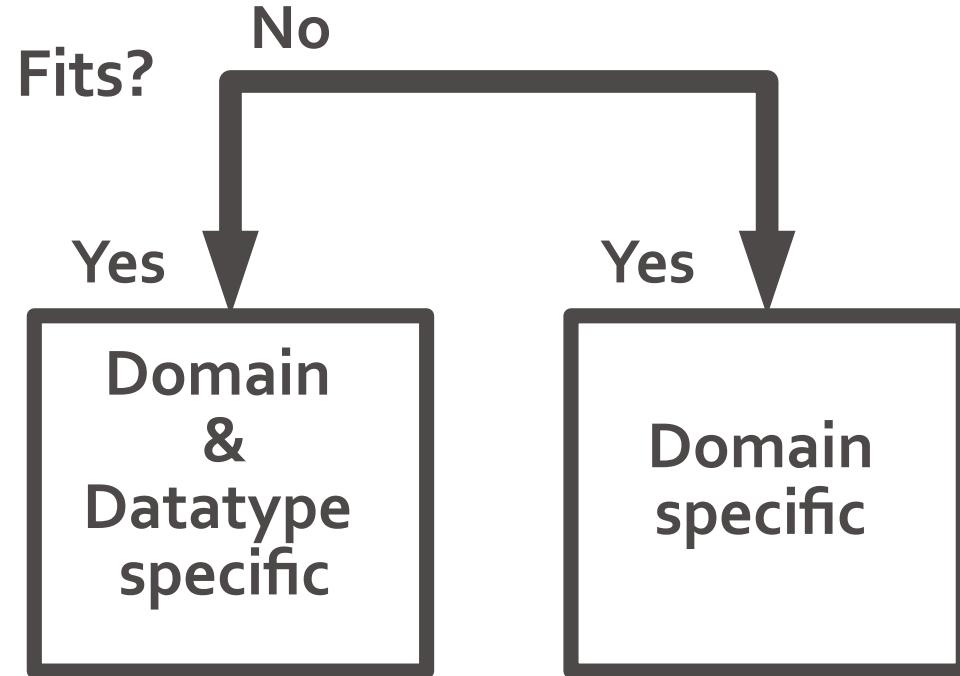


Yes

Domain  
&  
Datatype  
specific



# Where should I deposit my data?

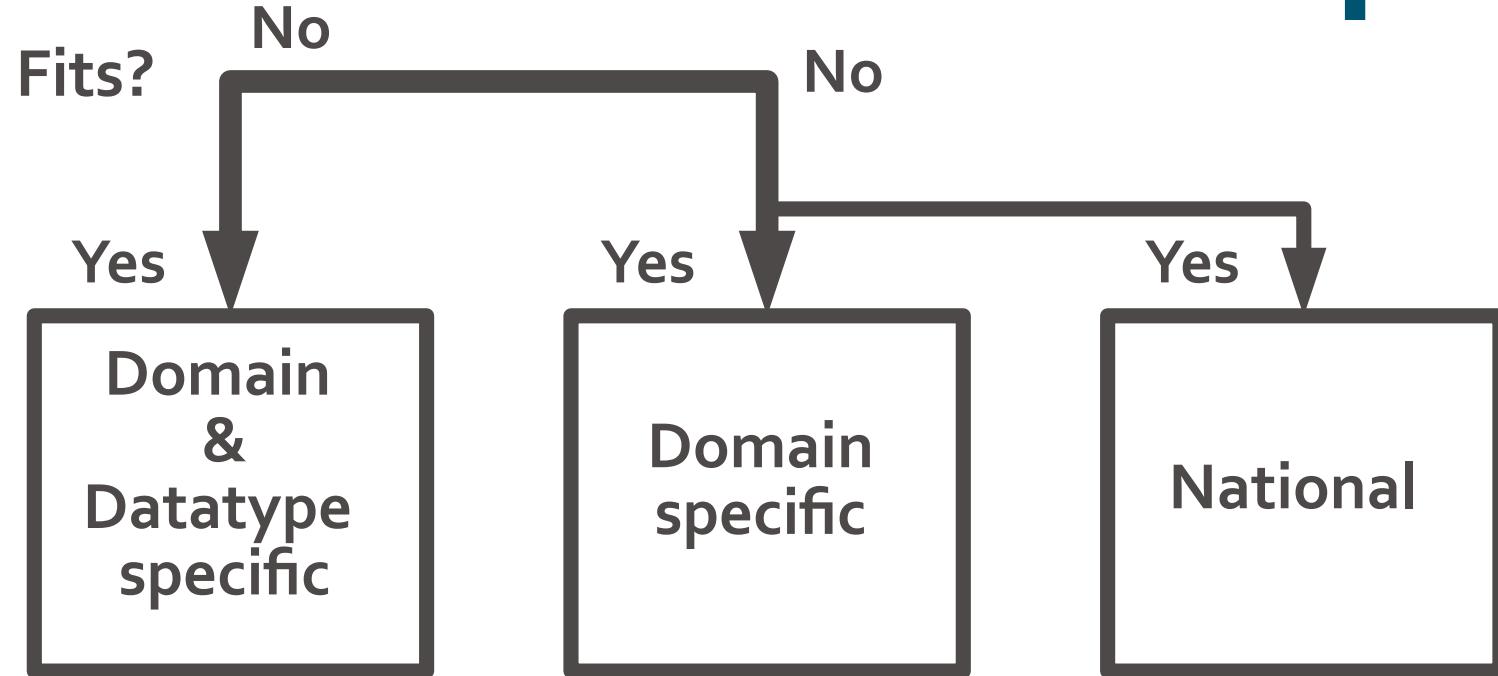


 ENA  
European Nucleotide Archive

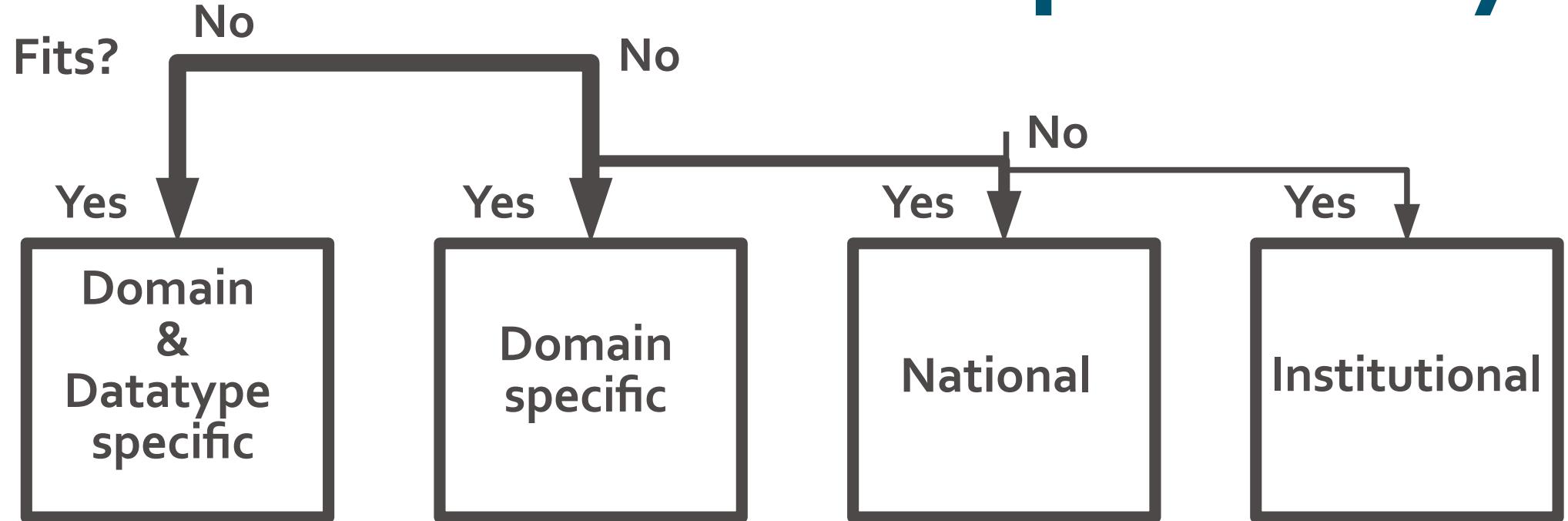
 BioStudies.

 SRA  
NCBI

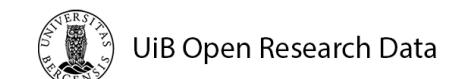
# Where should I deposit my data?



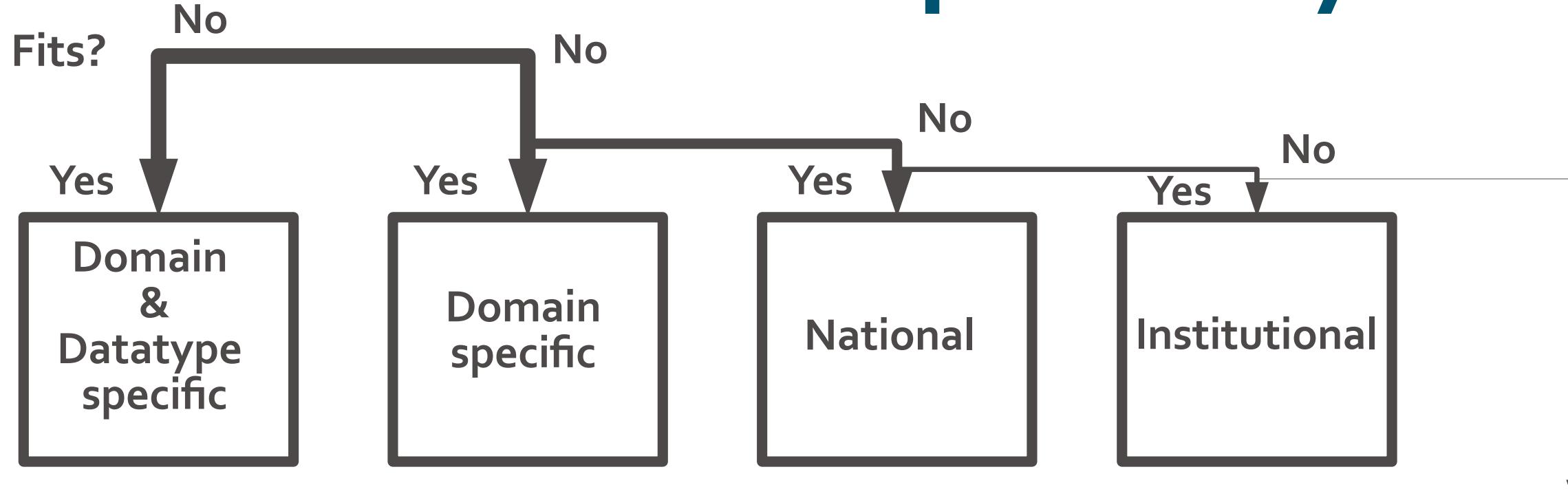
# Where should I deposit my data?



NIRD RESEARCH DATA ARCHIVE



# Where should I deposit my data?



 ENA  
European Nucleotide Archive

 SRA  
NCBI

 BioStudies.

NIRD RESEARCH DATA ARCHIVE

 DataverseNO  
Dataverse Network Norway

 NTNU Open Research Data  
NTNU

 NMBU Open Research Data  
Norwegian University of Life Sciences

 UiB Open Research Data

 UiT Open Research Data

generic

 zenodo

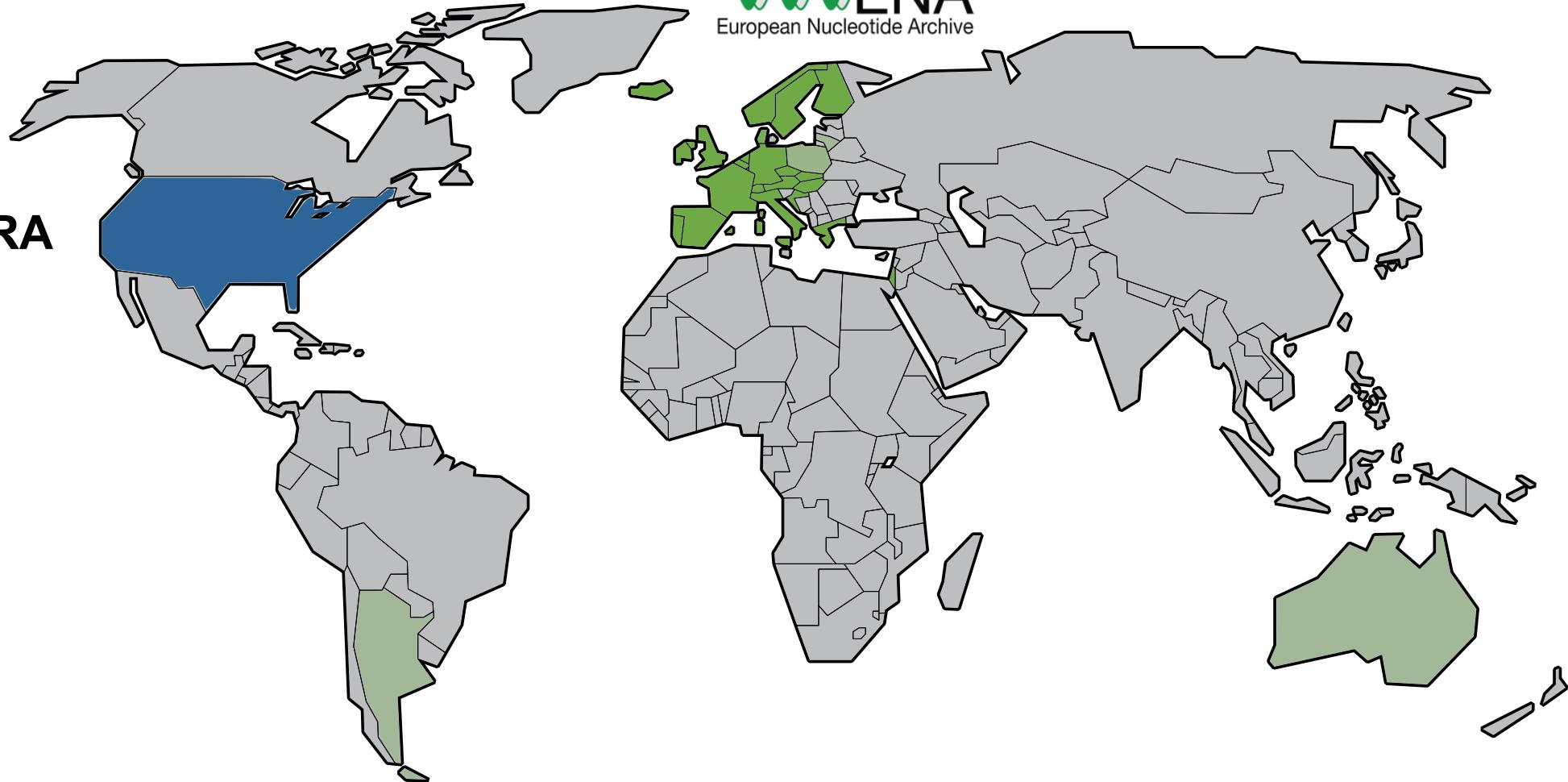
 EUDAT

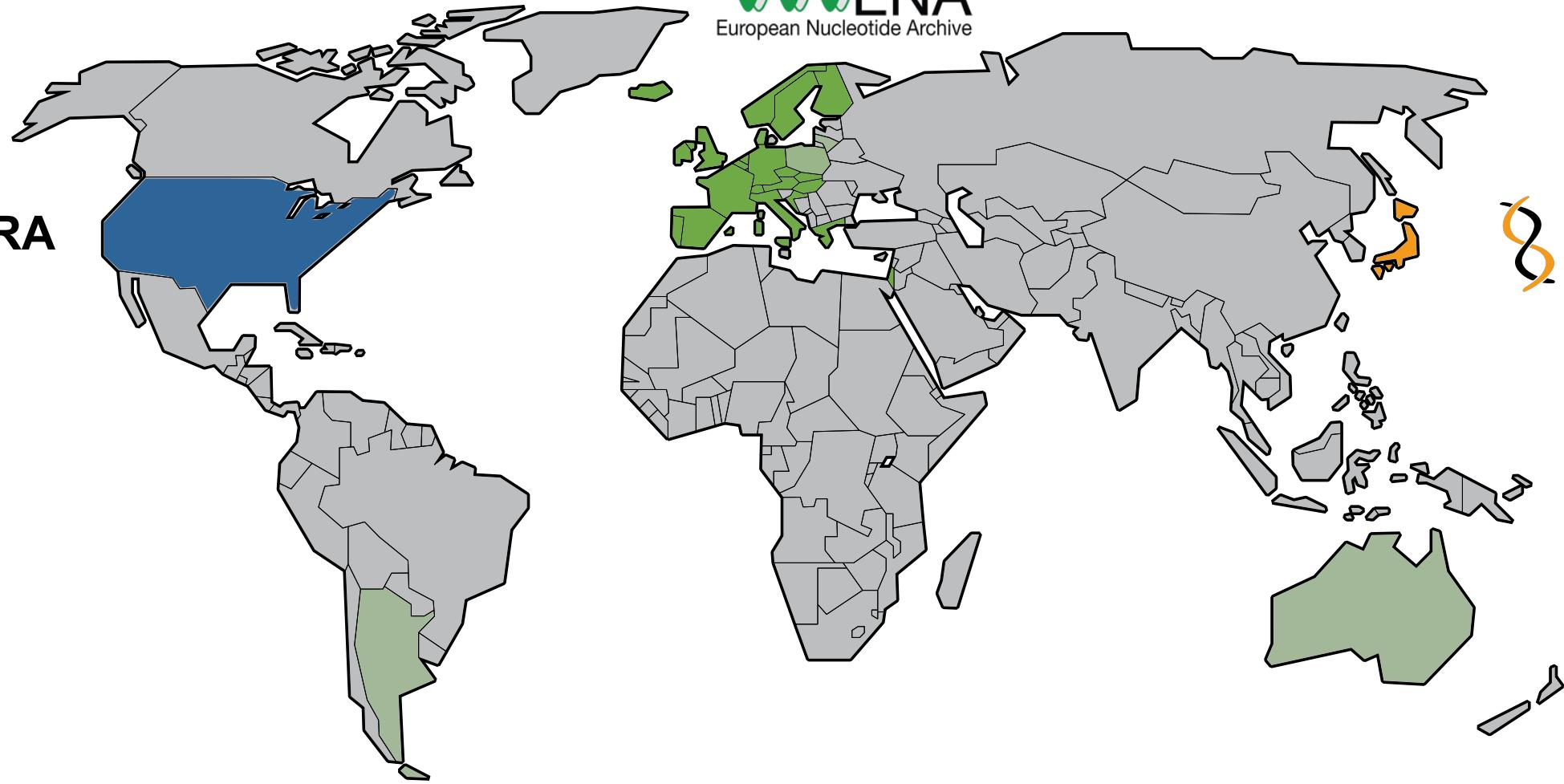
 figshare

# Multiple Repositories – similar datatypes



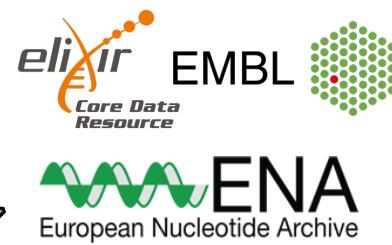
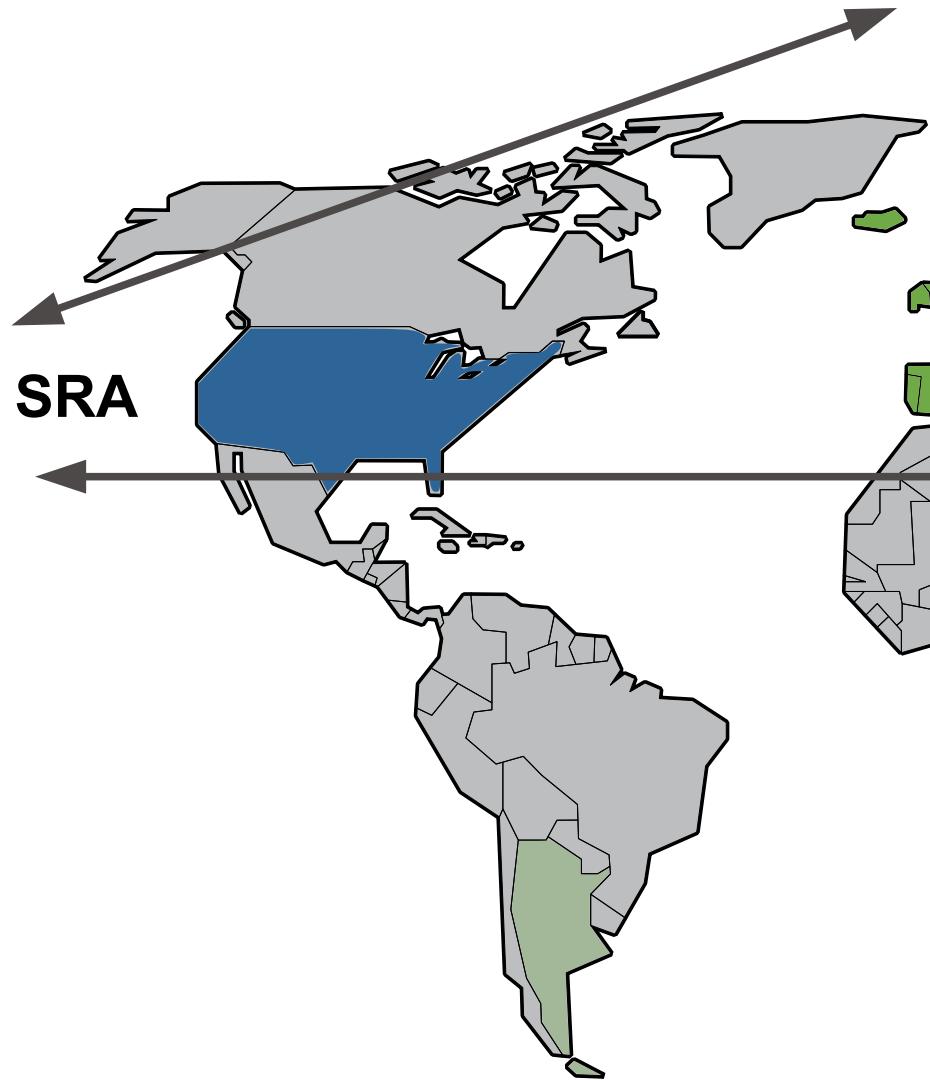








SRA

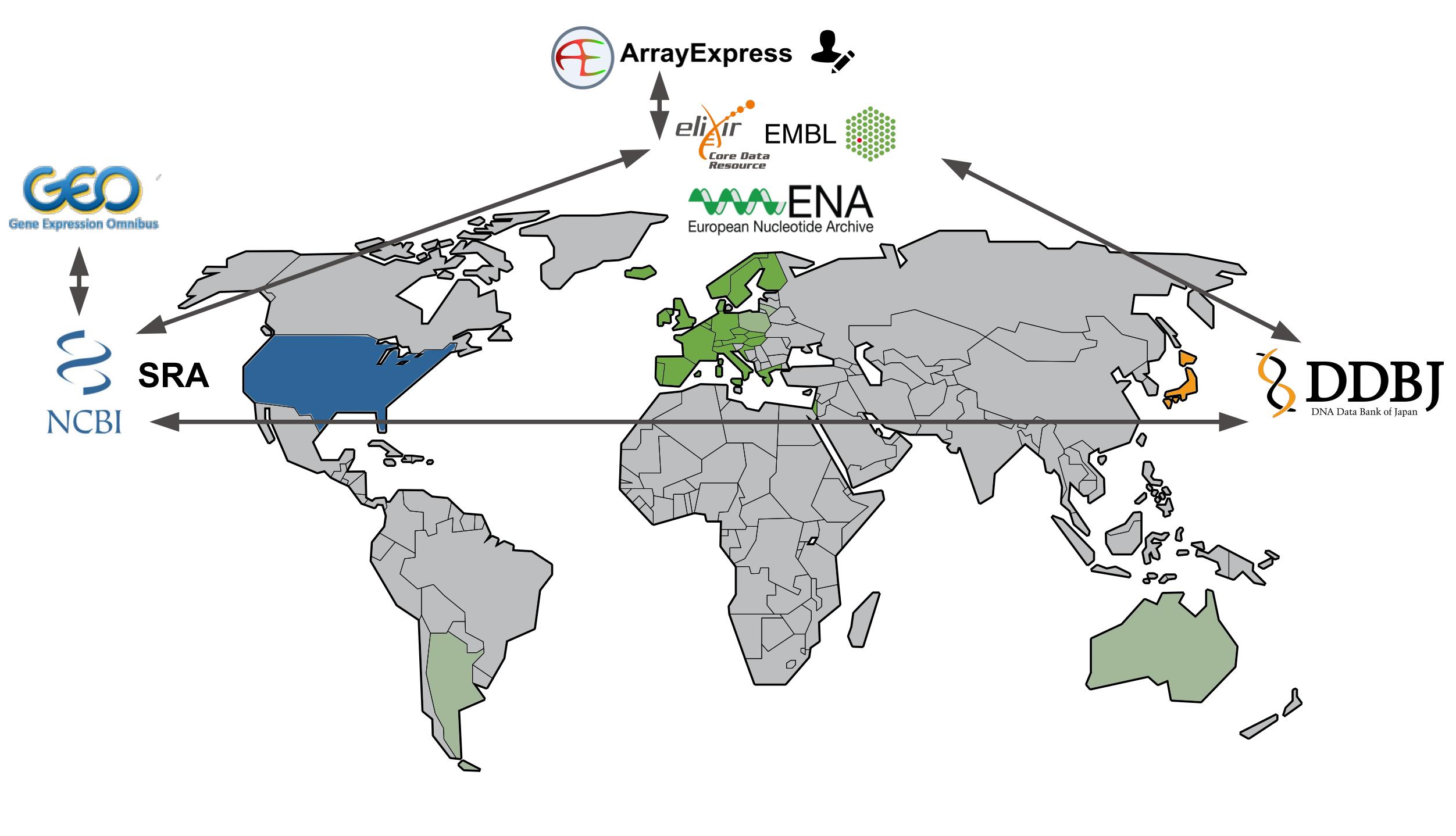


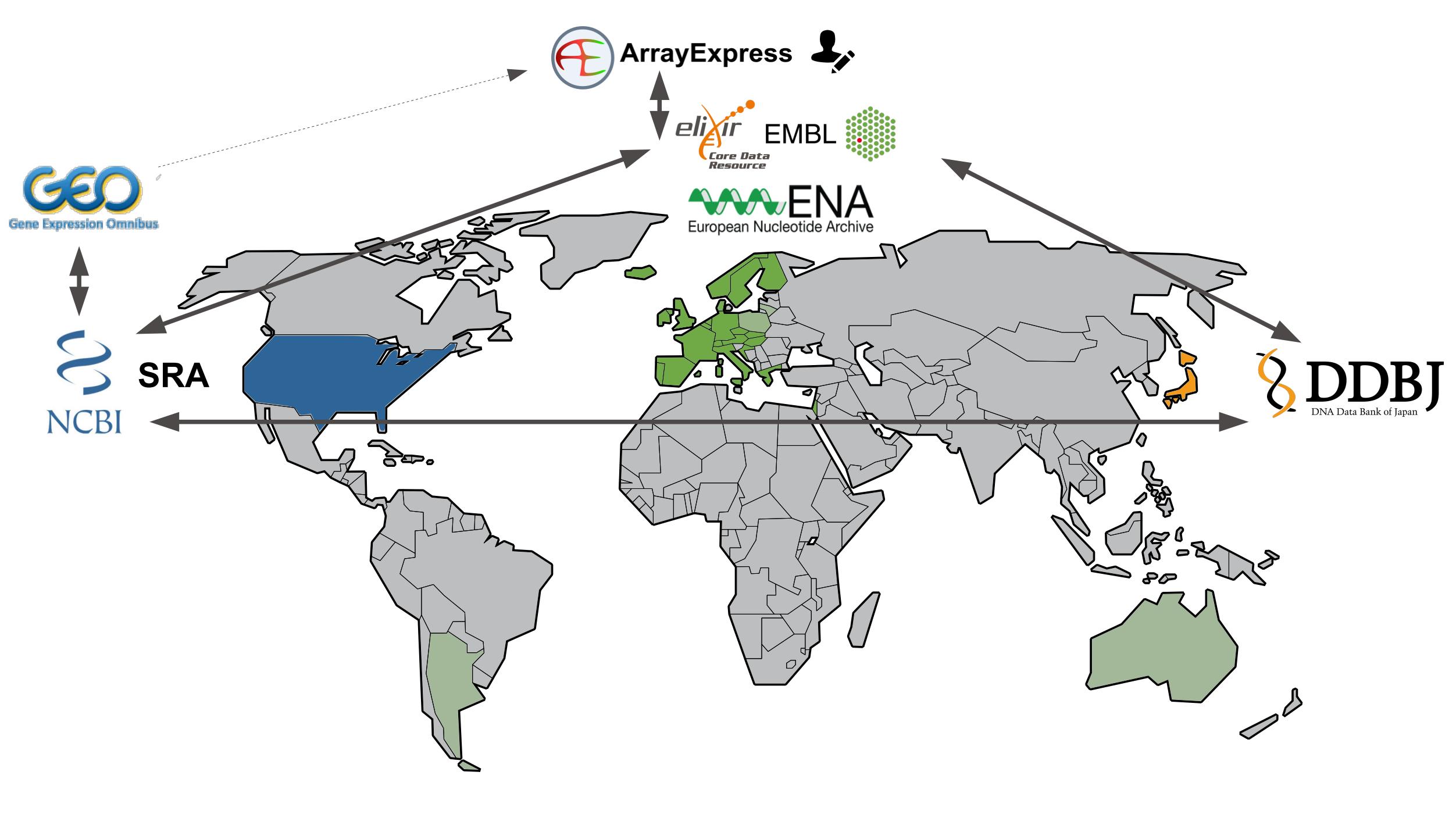
ENA

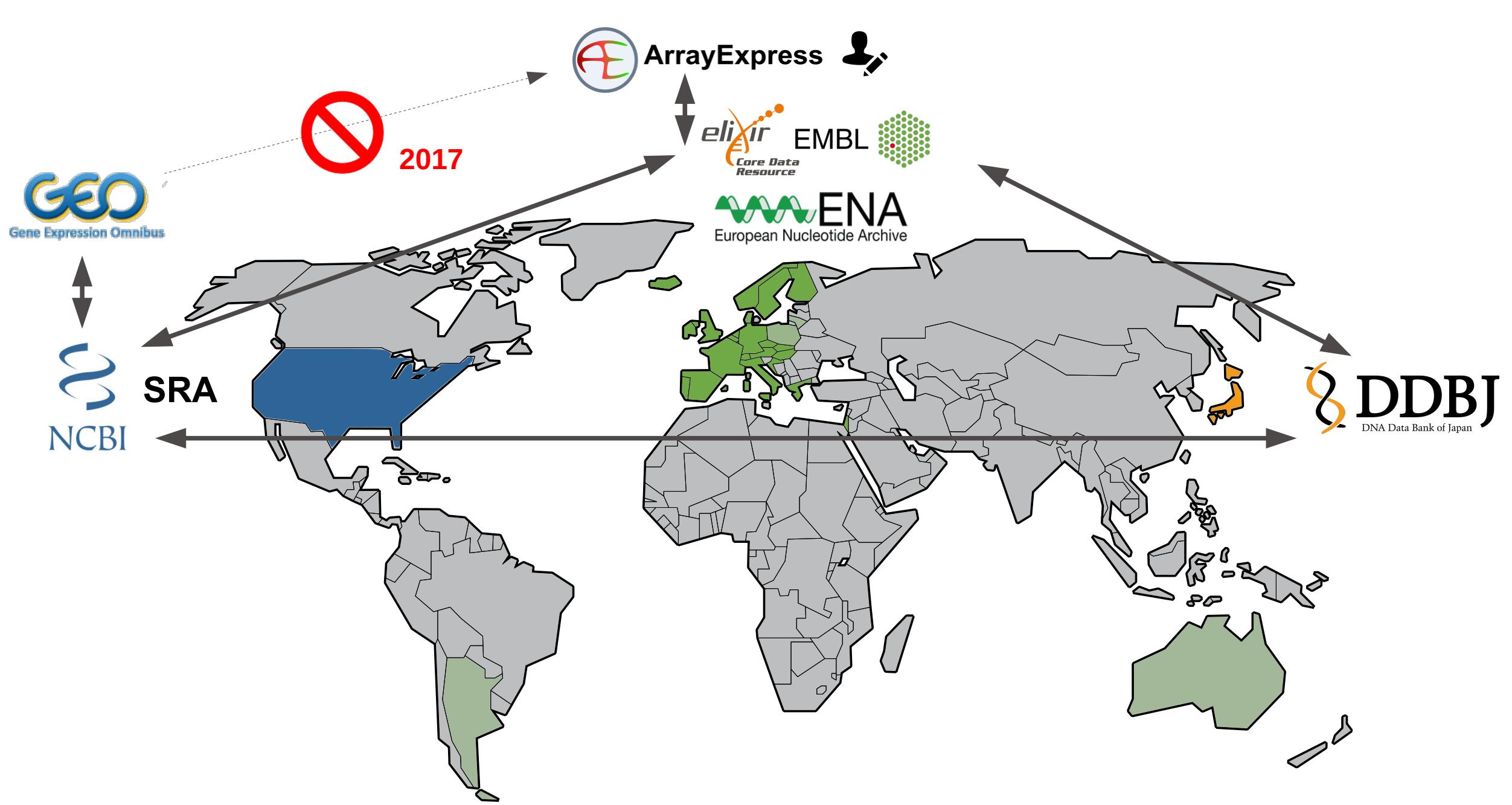
European Nucleotide Archive



DNA Data Bank of Japan









EMBL



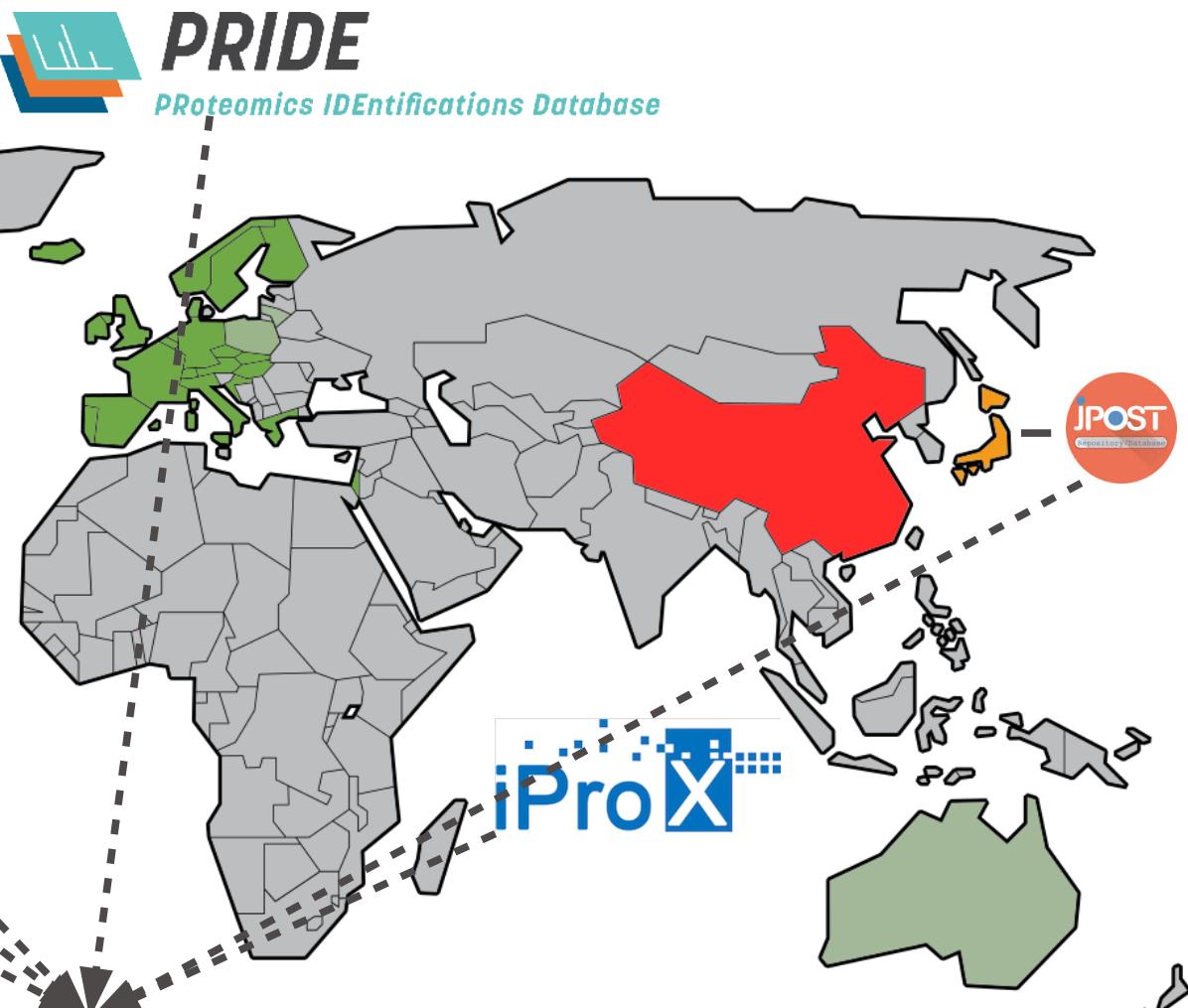


EMBL



**PRIDE**

PRoteomics IDEntifications Database



**Proteome  
Xchange**

# Where should I deposit my data?

- Free of charge



# Where should I deposit my data?

- Free of charge
- Quick upload (Aspera)



# Where should I deposit my data?



- Free of charge
- Quick upload (Aspera)
- Not operated by single university/government

# Where should I deposit my data?



- Free of charge
- Quick upload (Aspera)
- Not operated by single university/government
- Guided submission process



# Where should I deposit my data?



- Free of charge
- Quick upload (Aspera)
- Not operated by single university/government
- Guided submission process
- Generous embargo regulation (2yrs → )



Annotare

# Where should I deposit my data?



- Free of charge
- Quick upload (Aspera)
- Not operated by single university/government
- Guided submission process  Annotare
- 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?



# Where should I deposit my data?



ArrayExpress



**PRIDE**

*PRoteomics IDENTifications Database*



BioImage Archive



EMPIAR



IDR



BioModels



MetaboLights

# Where should I deposit my data?

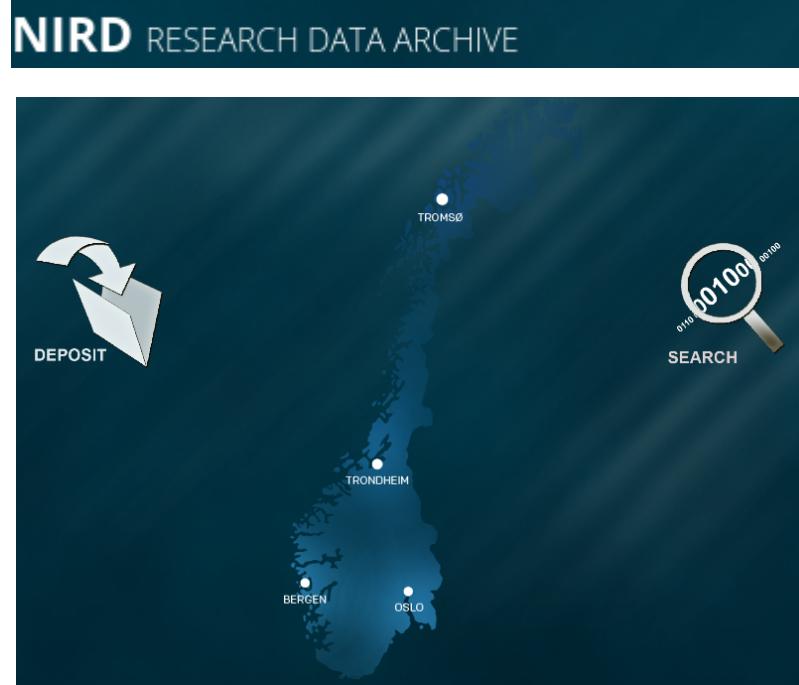


ArrayExpress

 BioImage Archive



# 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, ...

# What about sensitive data?



EUROPEAN  
GENOME-PHENOME  
ARCHIVE

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

Central metadata accessibility

Secure storage

Implemented data access committees

Local EGA in development

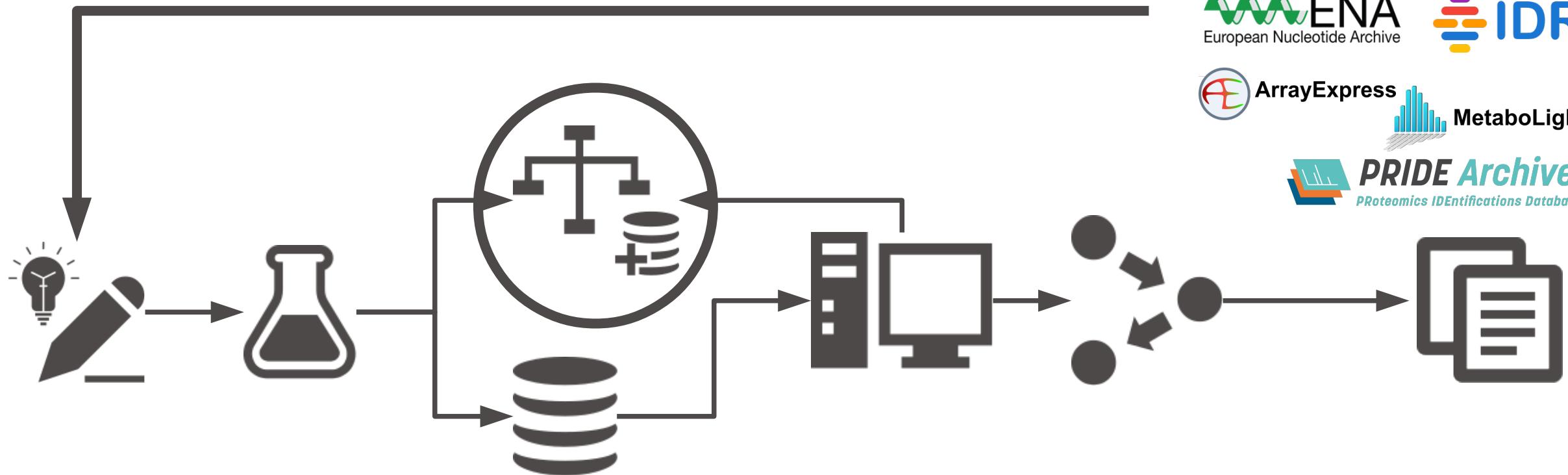


# NSD archive

Domain agnostic

DOI accessible

Implemented data access committees

  
European Nucleotide Archive  
IDR  
ArrayExpress  
MetaboLights  
PRIDE Archive  
PRoteomics IDEntifications Database  
EUROPEAN  
GENOME-PHENOME  
ARCHIVE