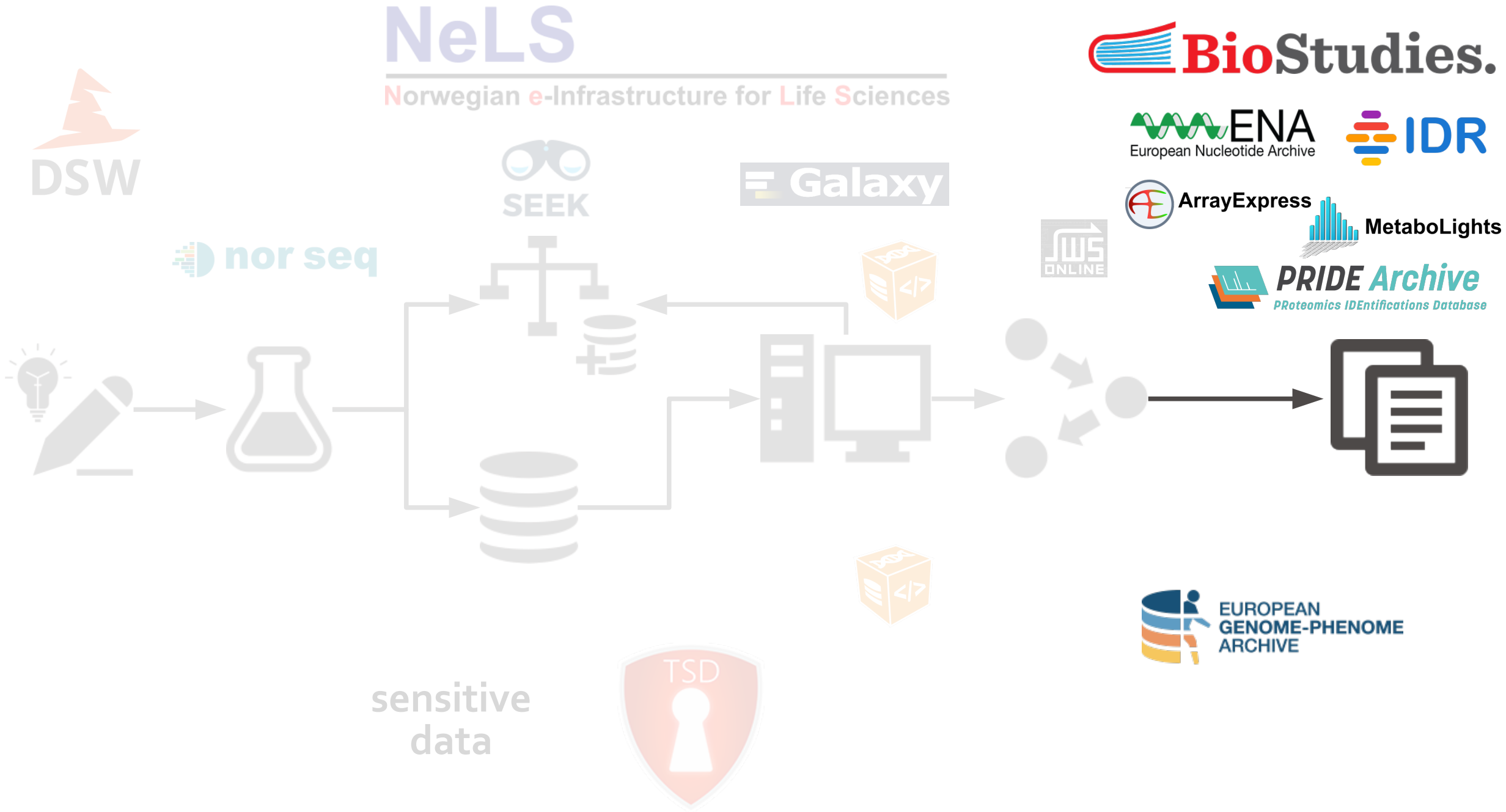
Data Deposition Repositories

Korbinian Bösl
Data manager ELIXIR Norway
21 February 2020


NeLS
Norwegian e-Infrastructure for Life Sciences
DSW
SEEK
Galaxy
nor seq
ONLINE
BioStudies.
ENA
European Nucleotide Archive
IDR
ArrayExpress
MetaboLights
PRIDE Archive
PRoteomics IDEntifications Database
EUROPEAN
GENOME-PHENOME
ARCHIVE
sensitive
data
TSD

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?

molecular biology 1637 databases

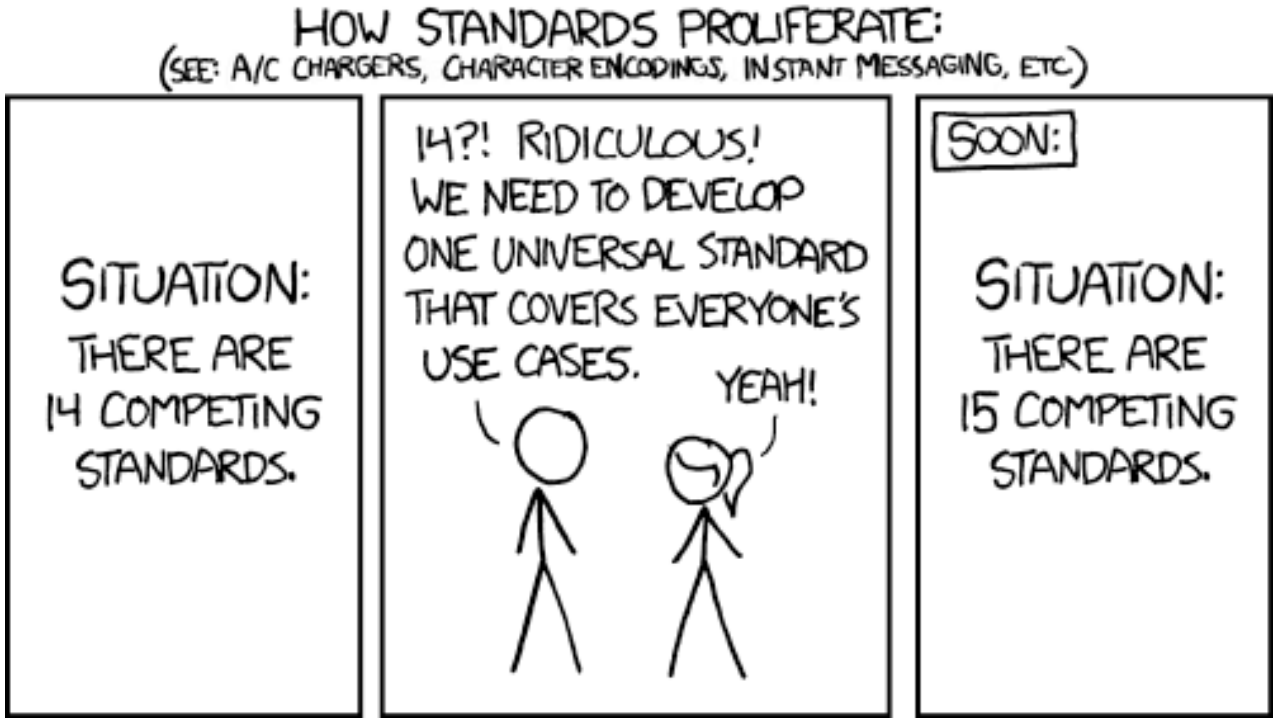
Rigden, D. J. & Fernández, X. M.
The 27th annual Nucleic Acids Research database issue and molecular biology database collection. Nucleic Acids Res 48, D1–D8 (2020).

Where should I deposit my data?

fairsharing.org
re3data.org

Meta data standards

Community standards vs. formal ISO standards



https://xkcd.com/927/
https://elixir-europe.org/events/elixir-guest-webinar-two-universes-one-world-community-standards-vs-formal-iso-standards-life

Where should I deposit my data?

Fits?

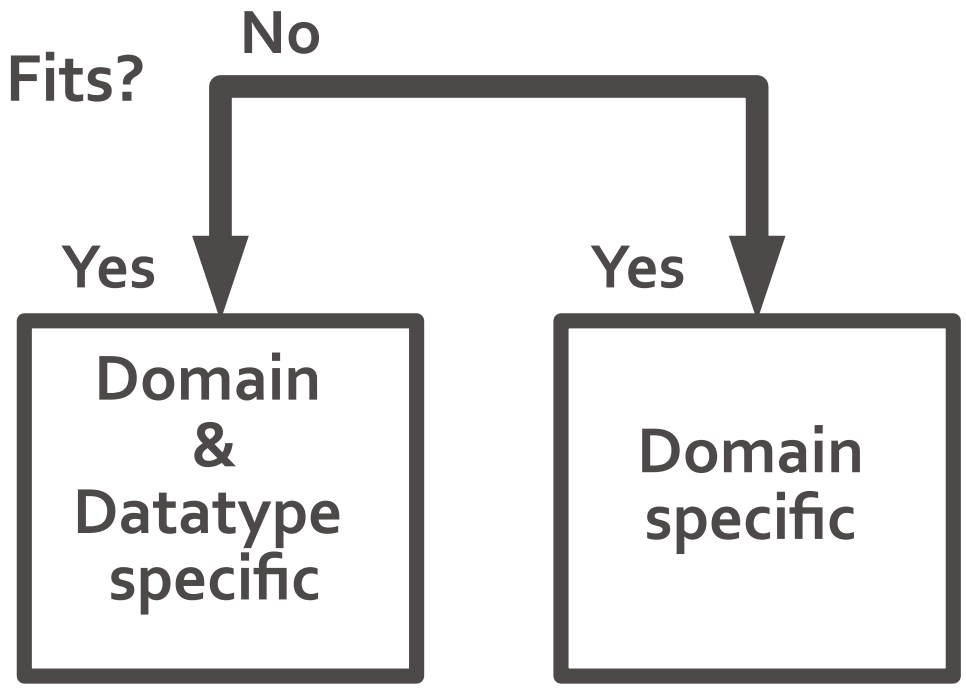
Yes

Domain
&
Datatype
specific

ENA
European Nucleotide Archive

NCBI SRA

Where should I deposit my data?

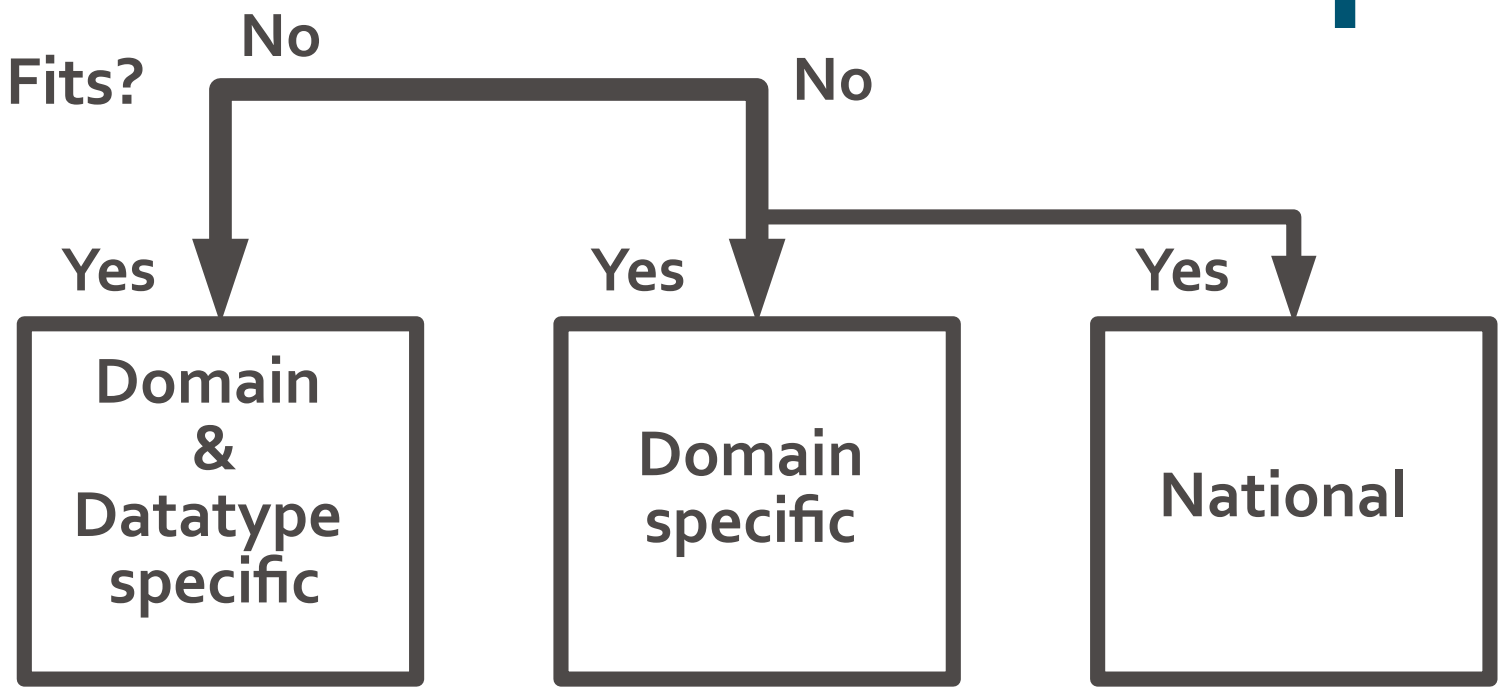


ENA
European Nucleotide Archive

BioStudies.

NCBI SRA

Where should I deposit my data?



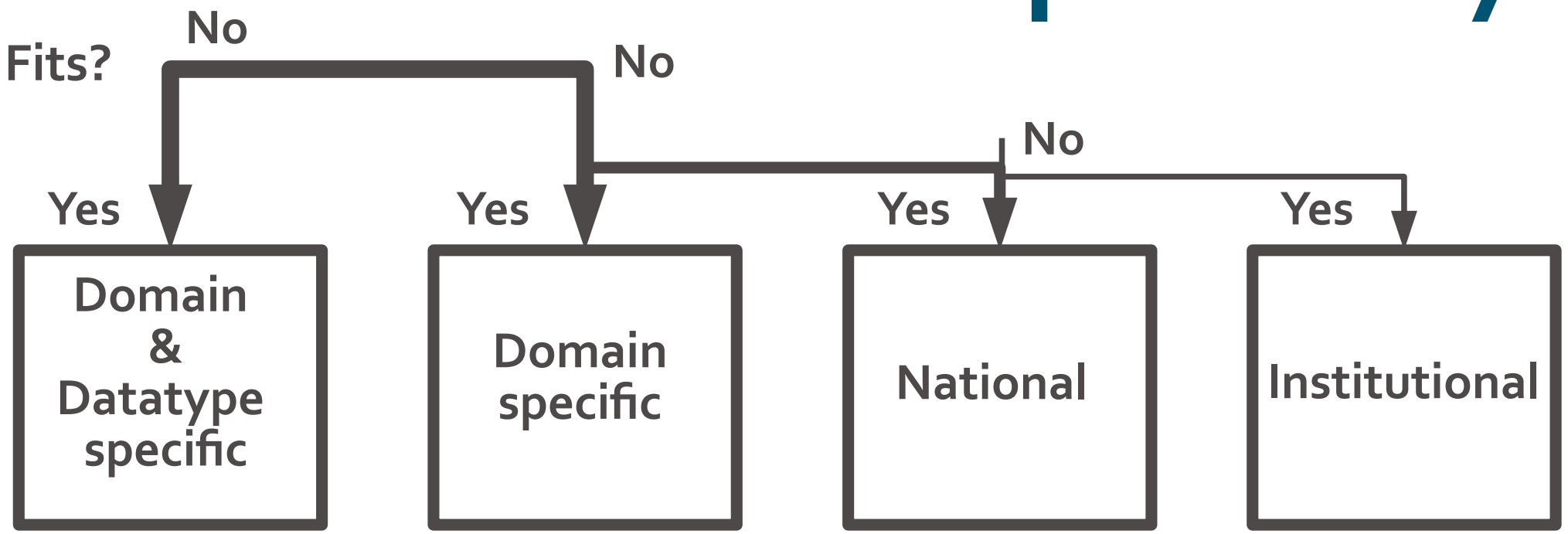
ENA
European Nucleotide Archive

BioStudies.

NIRD RESEARCH DATA ARCHIVE

NCBI SRA

Where should I deposit my data?



ENA
European Nucleotide Archive

BioStudies.

NIRD RESEARCH DATA ARCHIVE

SRA
NCBI

DataverseNO
Dataverse Network Norway

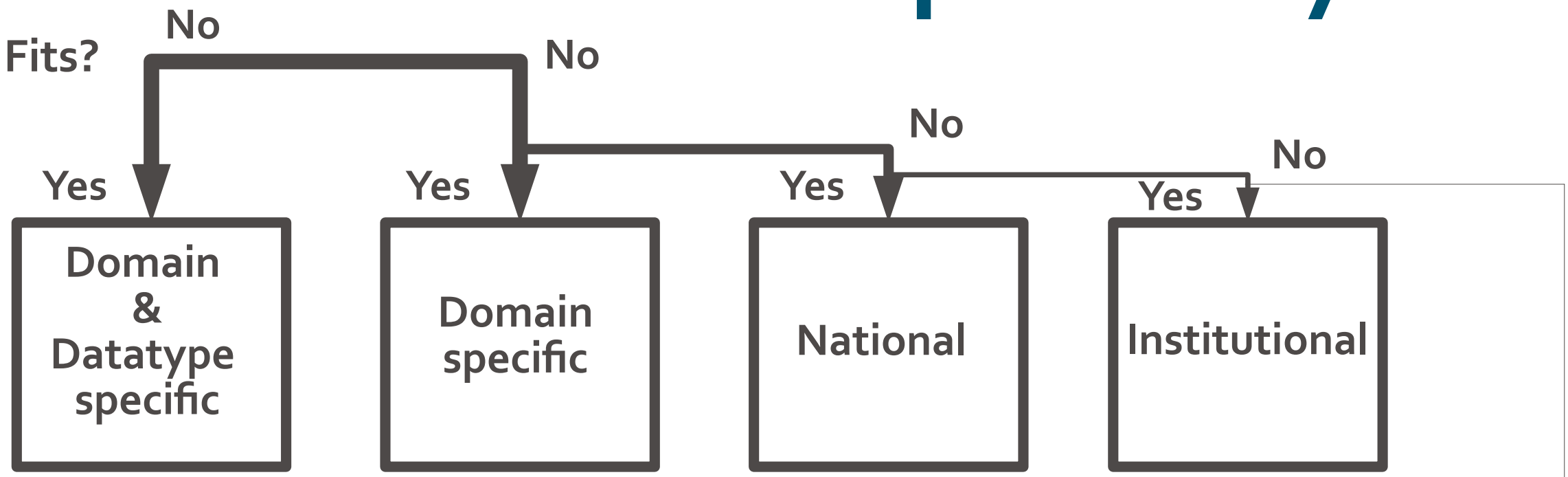
NTNU Open Research Data
NTNU

NMBU Open Research Data
Norwegian University of Life Sciences

UiB Open Research Data

UiT Open Research Data

Where should I deposit my data?



ENA
European Nucleotide Archive

NCBI SRA

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

zenodo

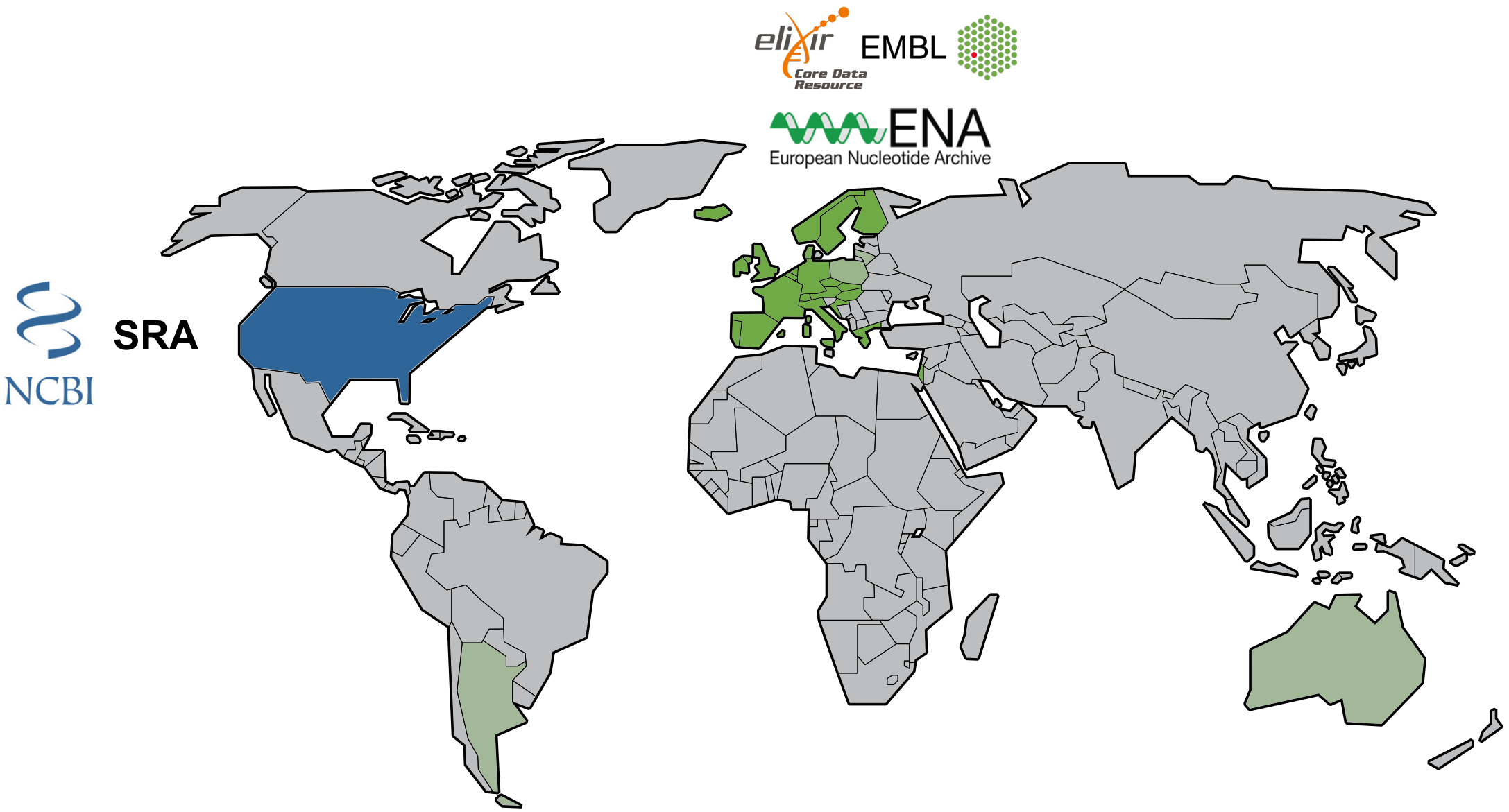
EUDAT

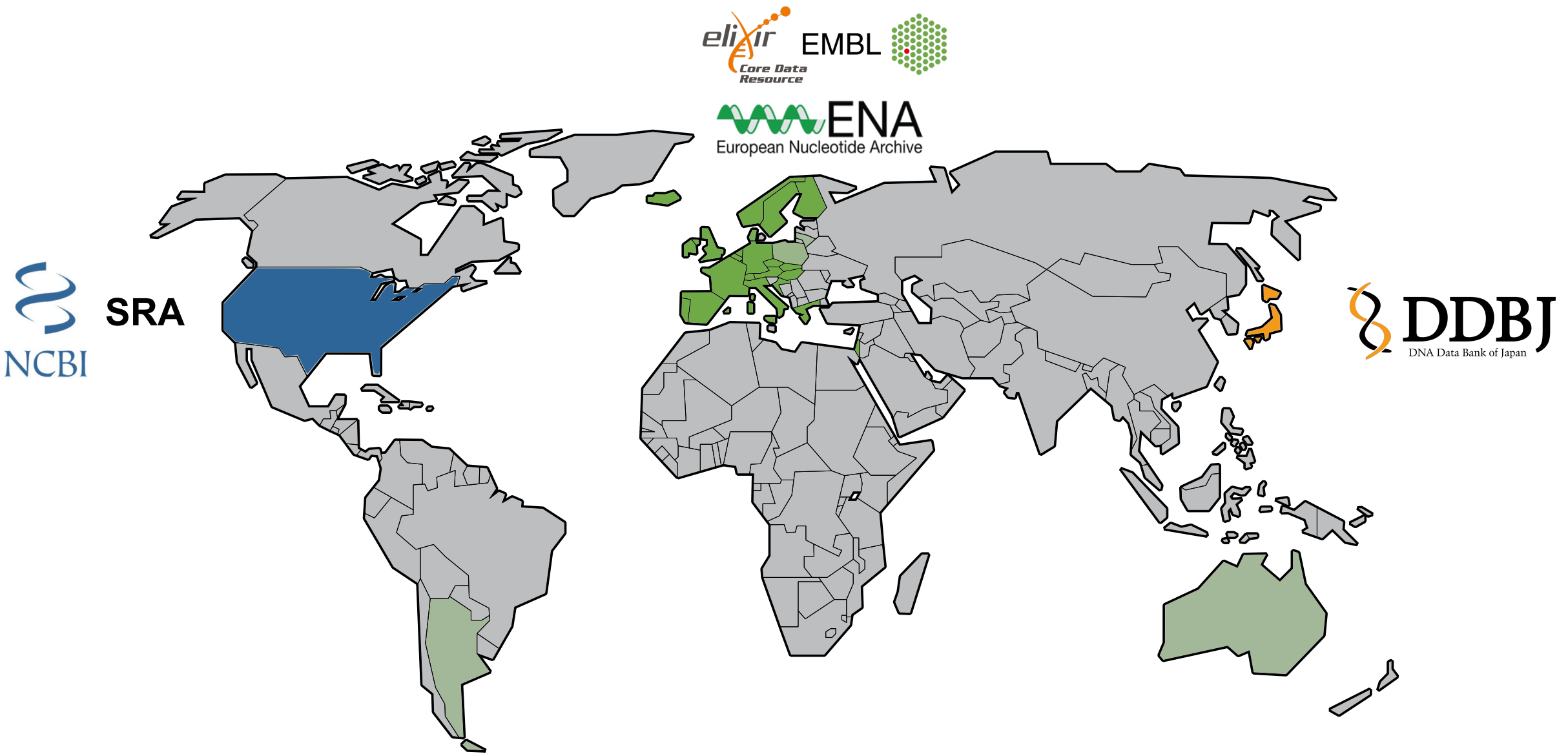
figshare

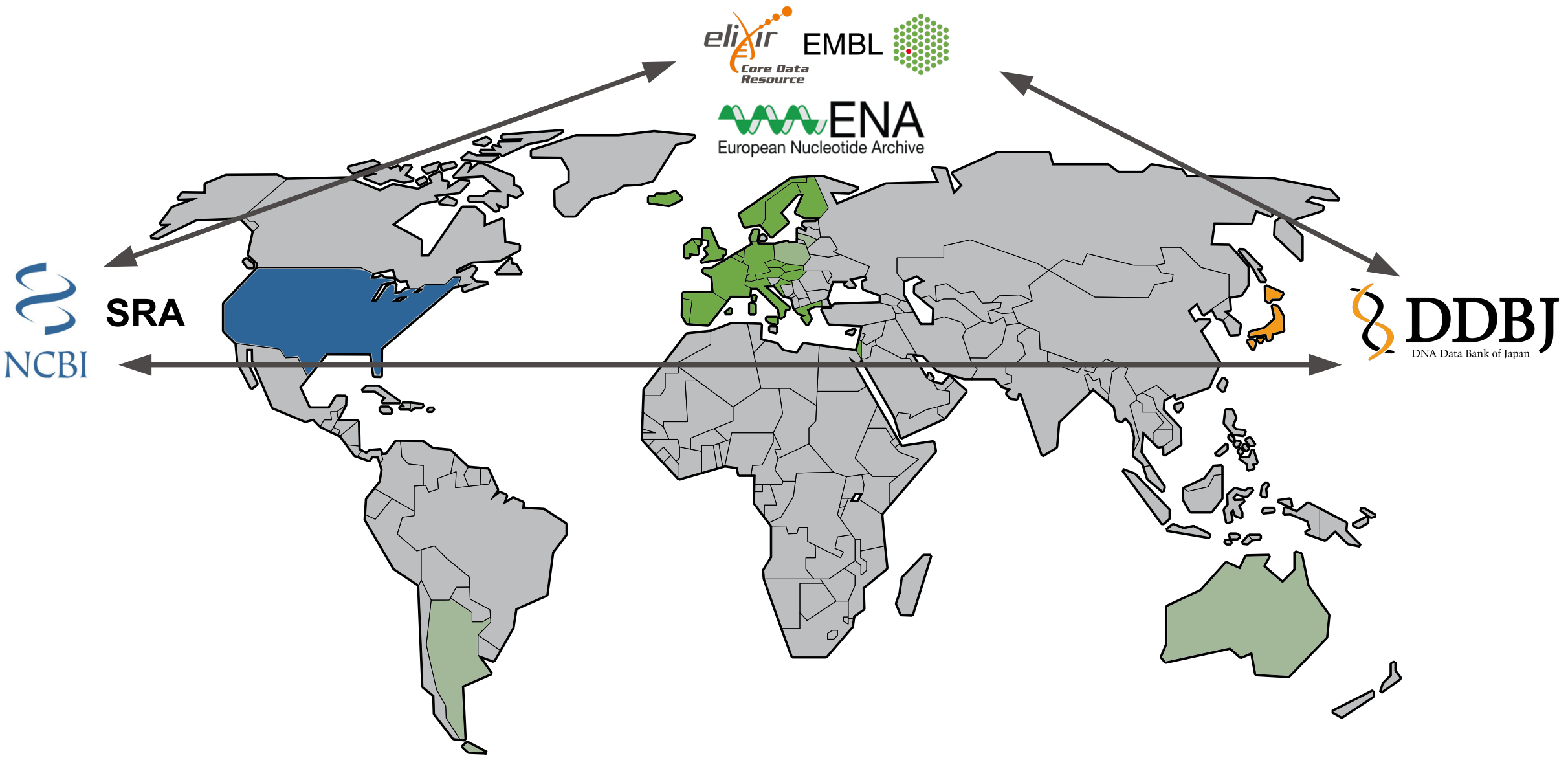
Multiple Repositories – similar datatypes

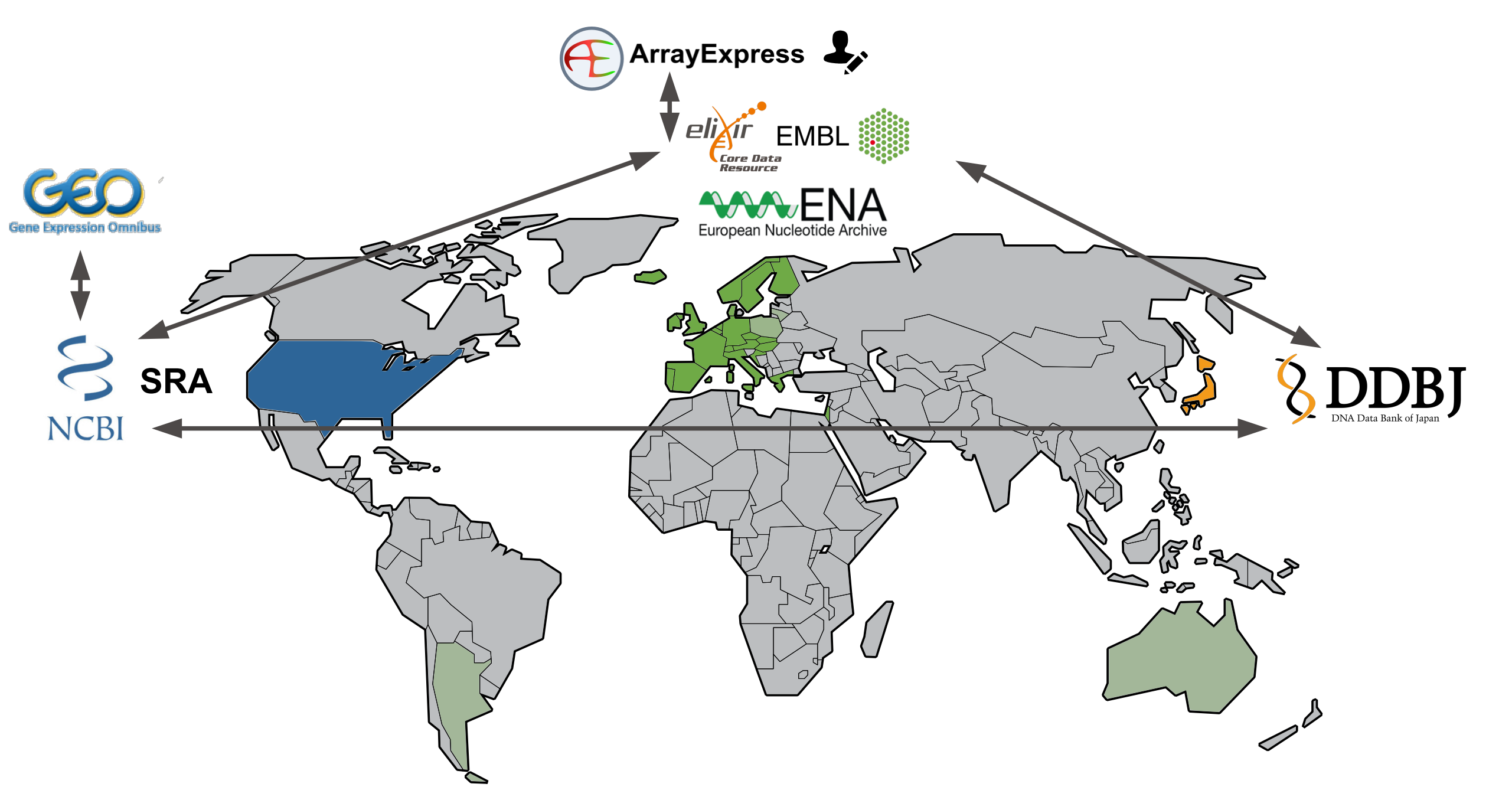
NCBI

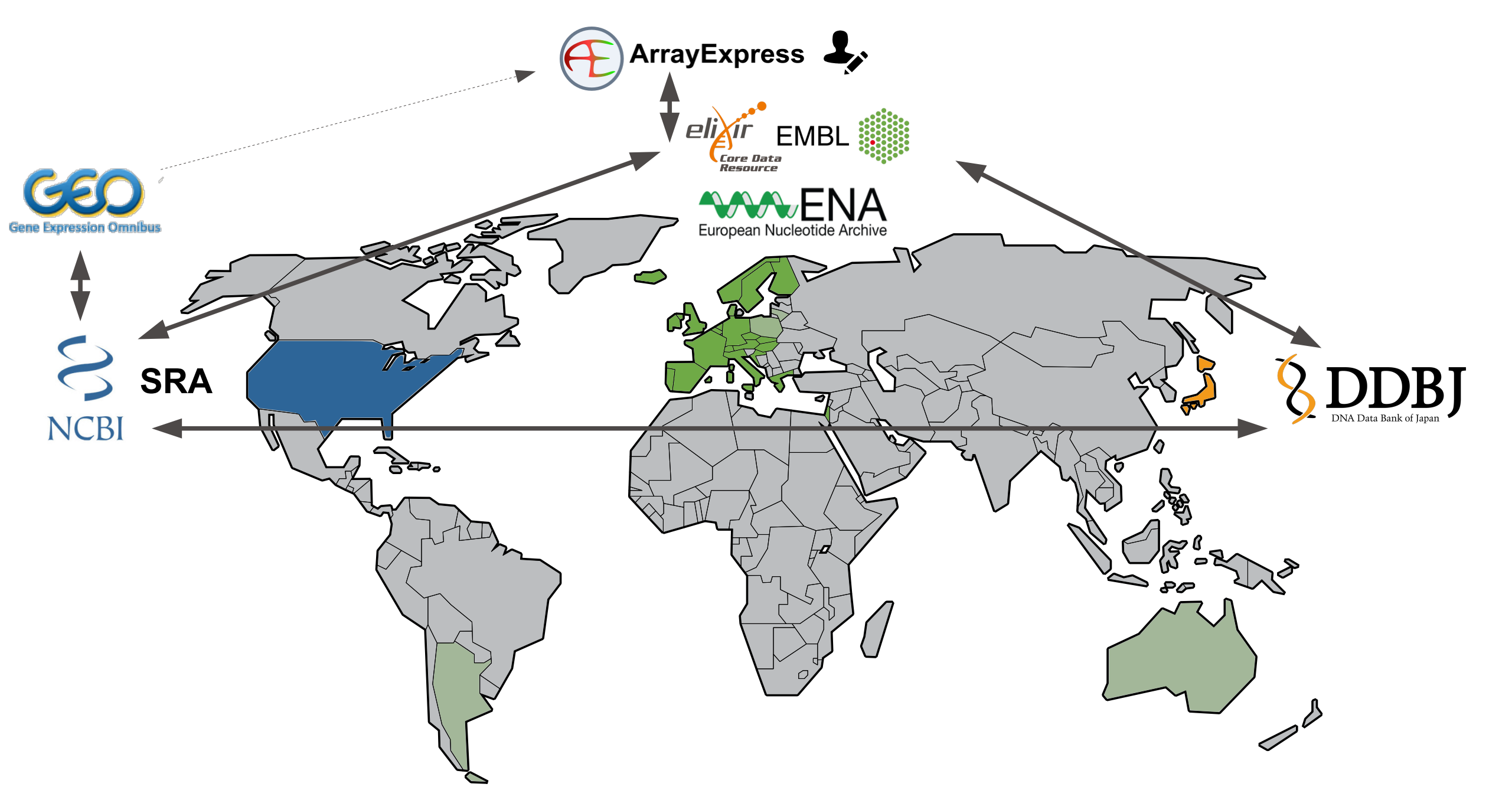
SRA

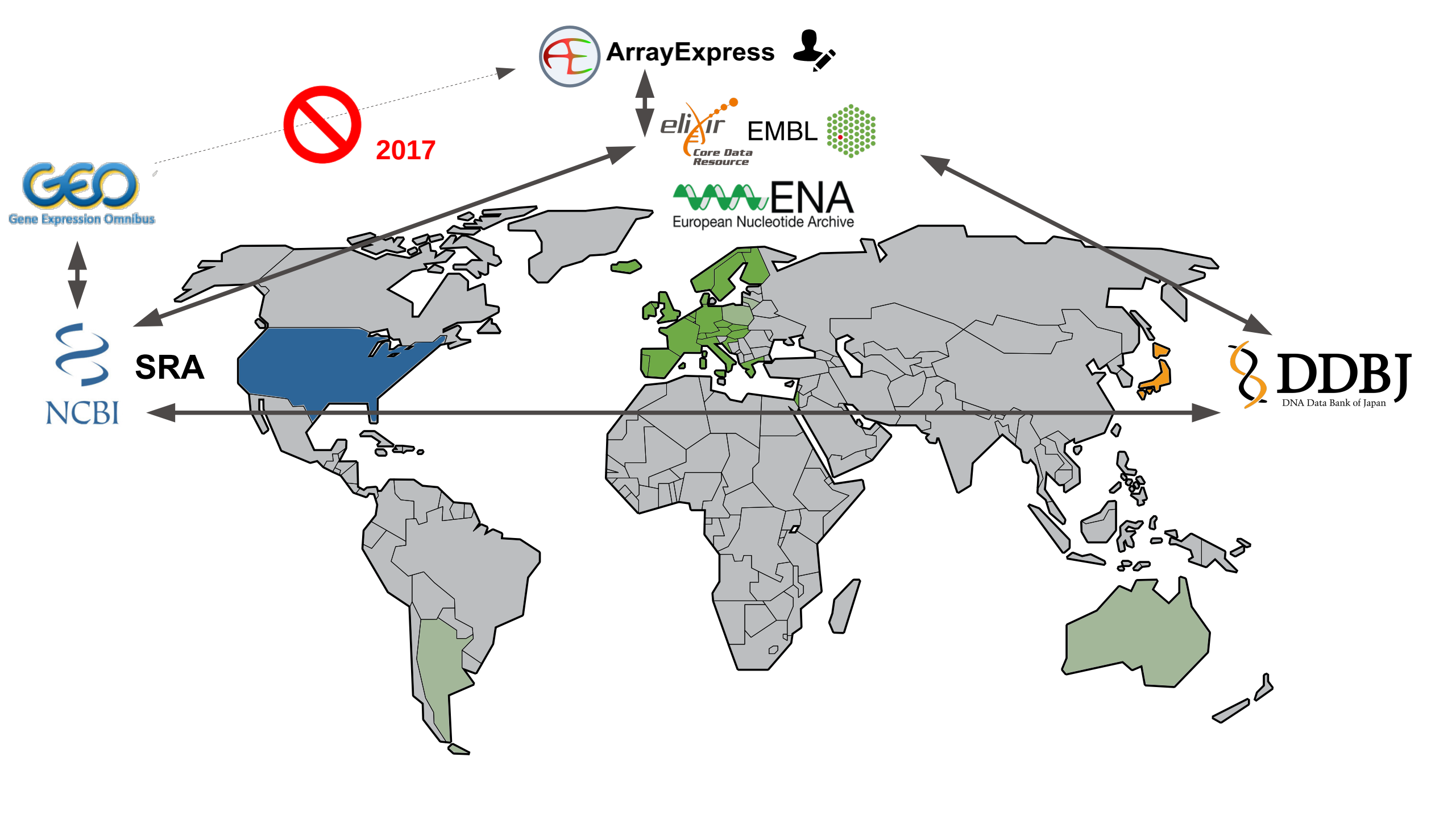

elixir
Core Data Resource
EMBL
ENA
European Nucleotide Archive
NCBI
SRA

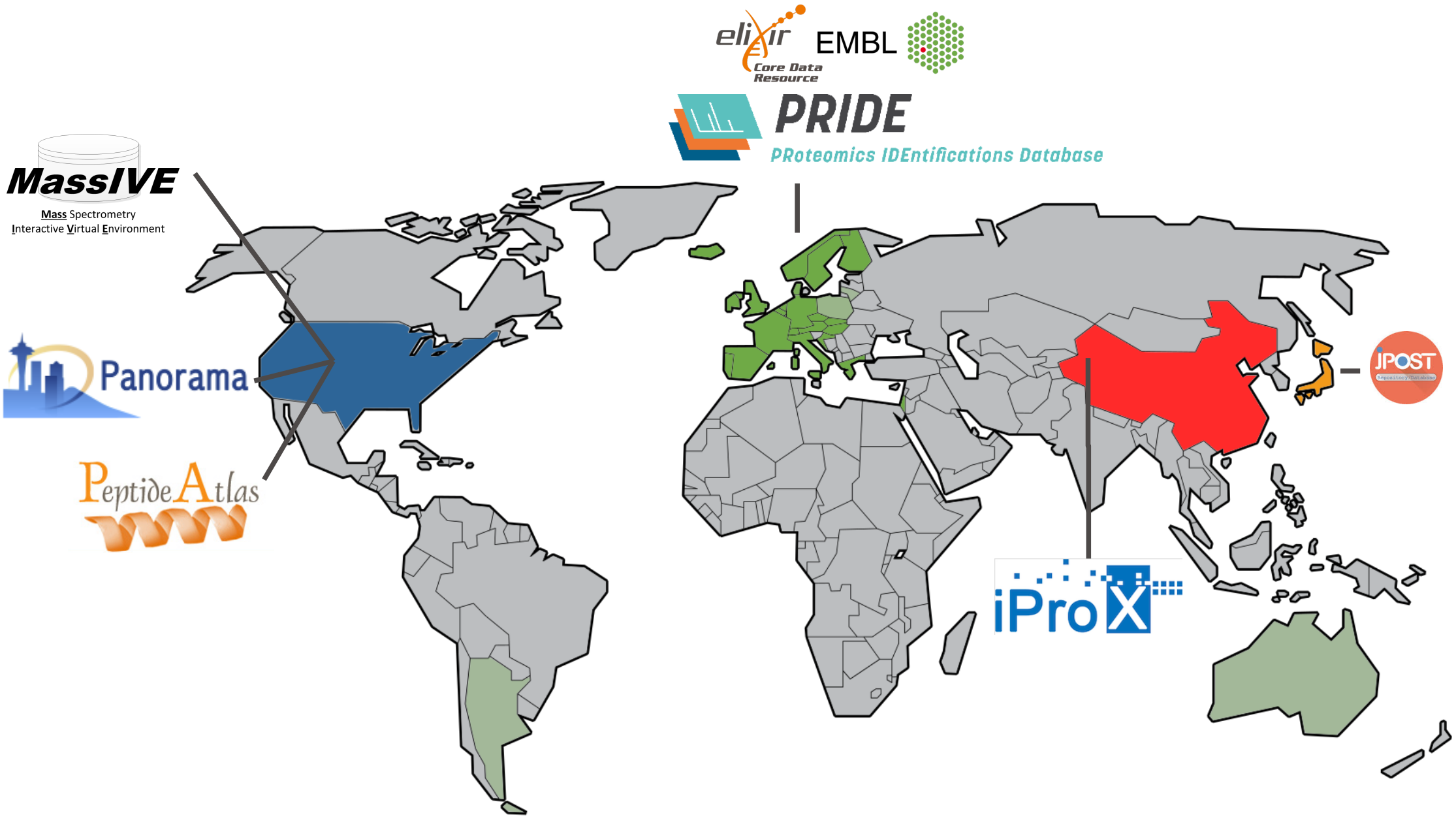

elixir
Core Data Resource
EMBL
ENA
European Nucleotide Archive
NCBI
SRA
DDBJ
DNA Data Bank of Japan

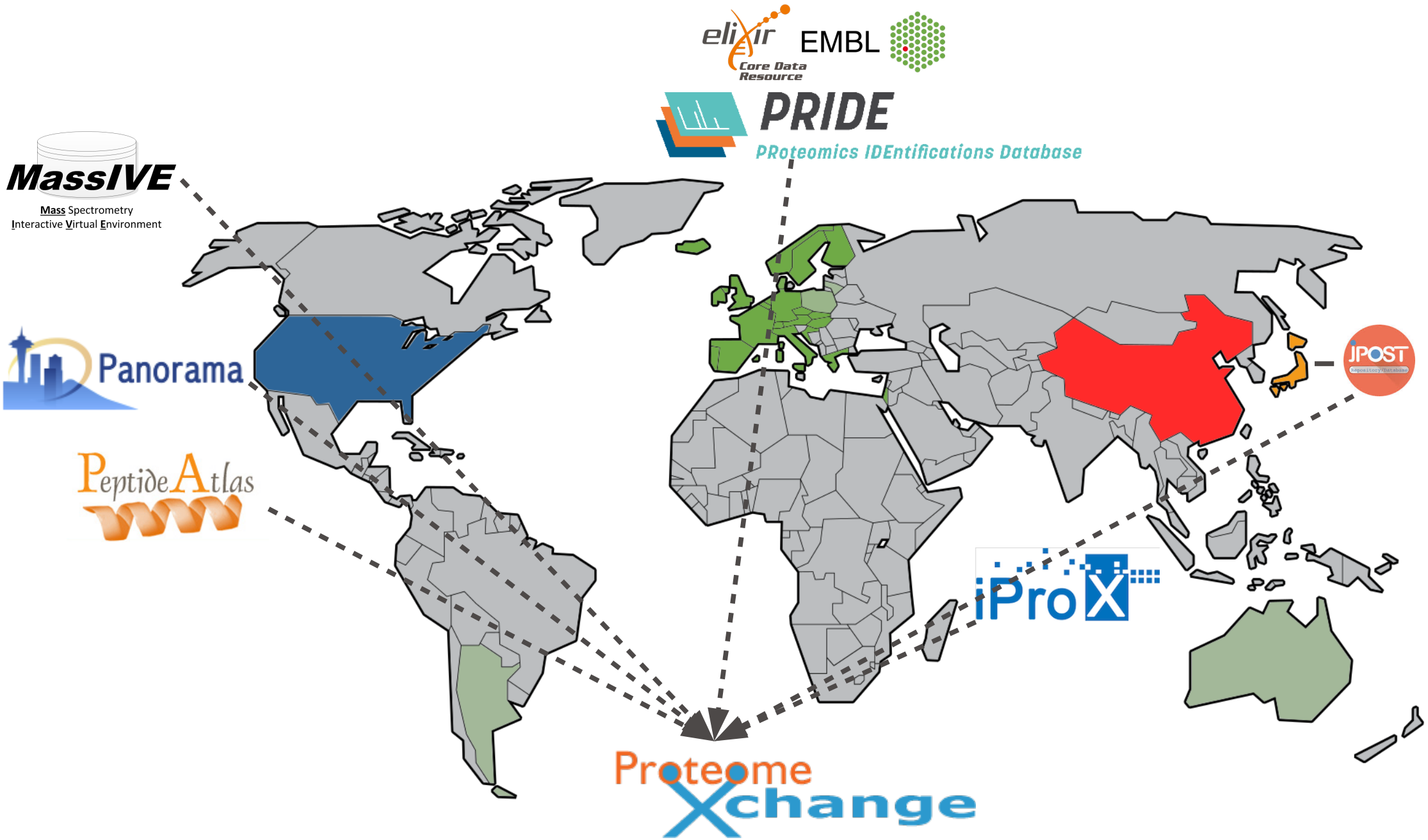

elixir
Core Data Resource
EMBL
ENA
European Nucleotide Archive
NCBI
SRA
DDBJ
DNA Data Bank of Japan


ArrayExpress
elixir
Core Data Resource
EMBL
ENA
European Nucleotide Archive
GEO
Gene Expression Omnibus
SRA
NCBI
DDBJ
DNA Data Bank of Japan


ArrayExpress
elixir
Core Data Resource
EMBL
ENA
European Nucleotide Archive
GEO
Gene Expression Omnibus
SRA
NCBI
DDBJ
DNA Data Bank of Japan


ArrayExpress
2017
elixir
Core Data Resource
EMBL
GEO
Gene Expression Omnibus
ENA
European Nucleotide Archive
SRA
NCBI
DDBJ
DNA Data Bank of Japan


elixir
Core Data Resource
EMBL
PRIDE
PRoteomics IDEntifications Database
MassIVE
Mass Spectrometry
Interactive Virtual Environment
Panorama
PeptideAtlas
iProX
JPOST


elixir
Core Data Resource
EMBL
PRIDE
PRoteomics IDEntifications Database
MassIVE
Mass Spectrometry
Interactive Virtual Environment
Panorama
PeptideAtlas
JPOST
iProX
ProteomeXchange

Where should I deposit my data?

- Free of charge

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

Where should I deposit my data?

- Free of charge
- Quick upload (Aspera)

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

Where should I deposit my data?

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

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

Where should I deposit my data?

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

elixir Core Data Resource

Annotare

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

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 →)

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

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 →)
- API access (upload from NeLS planned)

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

Where should I deposit my data?

ENA
European Nucleotide Archive

ArrayExpress

European Variation Archive

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

Where should I deposit my data?

ENA
European Nucleotide Archive

ArrayExpress

European Variation Archive

PDBe
Protein Data Bank in Europe

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

Where should I deposit my data?

ENA
European Nucleotide Archive

ArrayExpress

European Variation Archive

PRIDE
PRoteomics IDEntifications Database

PDBe
Protein Data Bank in Europe

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

Where should I deposit my data?

ENA
European Nucleotide Archive

PDBe
Protein Data Bank in Europe

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

Where should I deposit my data?

ENA
European Nucleotide Archive

ArrayExpress

BioImage Archive

EMPIAR

IDR

European Variation Archive

PRIDE
PRoteomics IDEntifications Database

BioModels

PDBe
Protein Data Bank in Europe

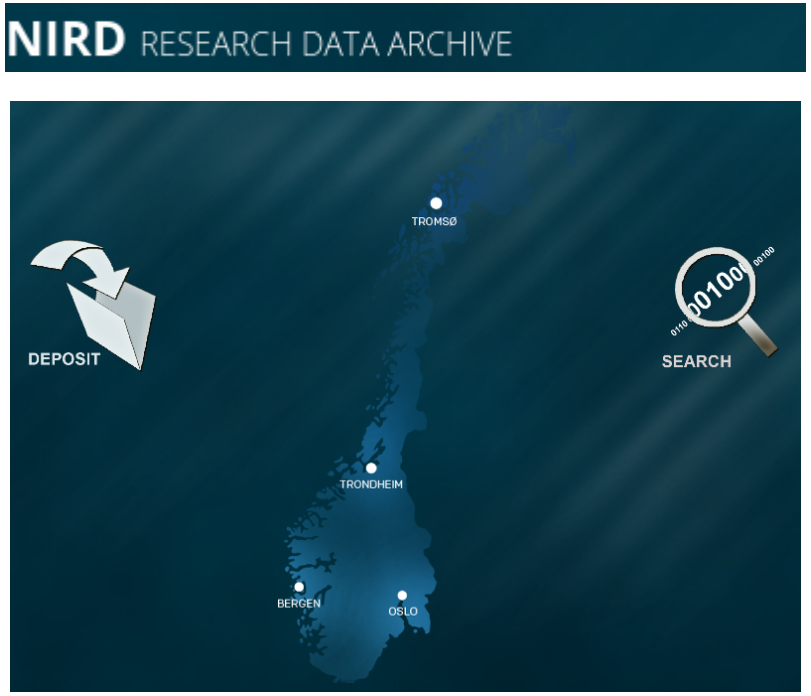
MetaboLights

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

Where should I deposit my data?

elixir-europe.org/platforms/data/core-data-resources
elixir-europe.org/platforms/data/elixir-deposition-databases

NIRD archive



National, free depositing repository

Domain agnostic

Dublin-core metadata standard

DOI accessible

Supports machine readable metadata harvesting

https://archive.norstore.no/

Generic archives

Less standard metadata → Re-usability ⇓

Domain agnostic → Finability ⇓

Guided submission process

Commercial actors: Size limitations, …

https://zenodo.org/
https://b2share.eudat.eu/
https://figshare.com/

What about sensitive data?

Central metadata accessibility

Secure storage

Implemented data access committees

Local EGA in development

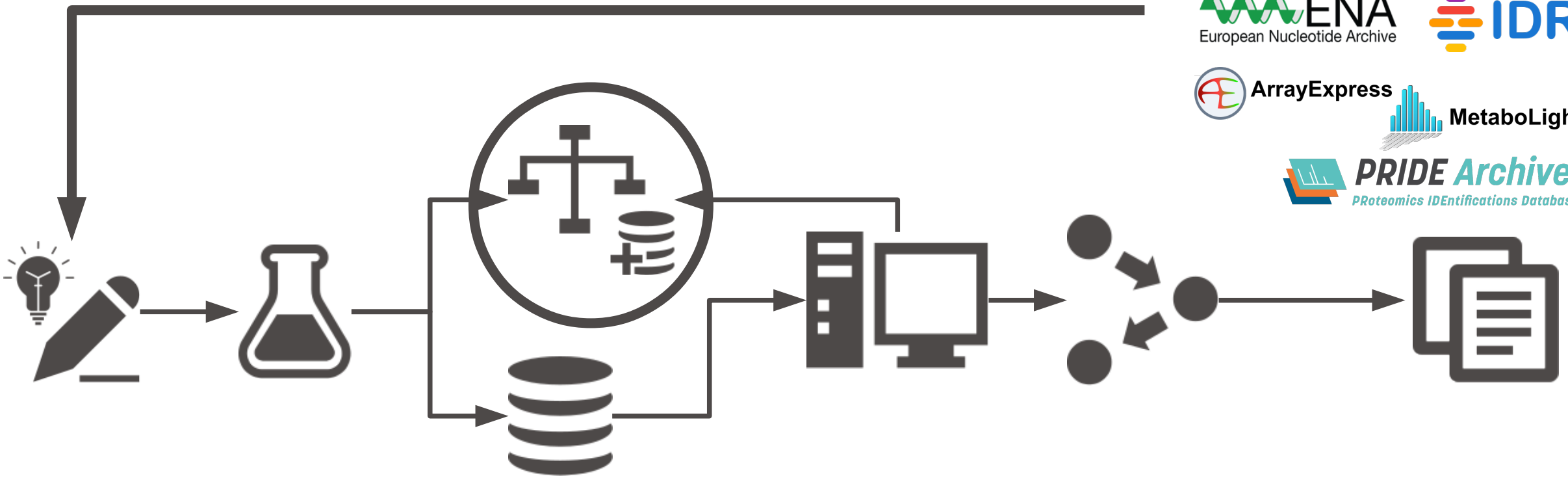
NSD archive

Domain agnostic

DOI accessible

Implemented data access committees

https://arkivering.nsd.no/



BioStudies.

ENA
European Nucleotide Archive

IDR

ArrayExpress

MetaboLights

PRIDE Archive
PRoteomics IDEntifications Database

EUROPEAN
GENOME-PHENOME
ARCHIVE