

ELIXIR Recommended Interoperability Resources

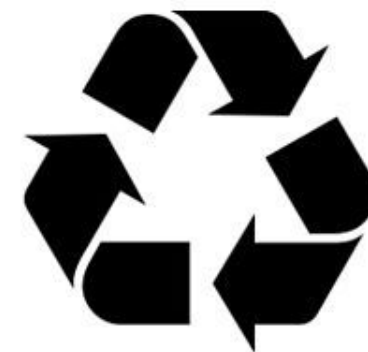
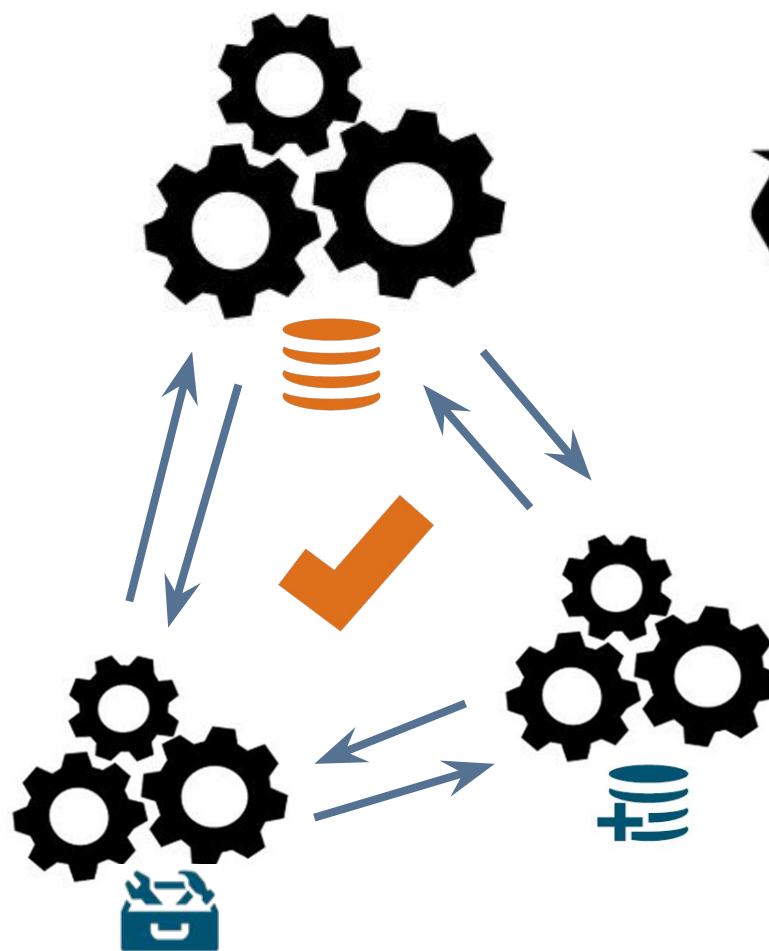


ELIXIR-EXCELERATE is funded by the European Commission within the Research Infrastructures programme of Horizon 2020, grant agreement number 676559.

*Carole Goble, ELIXIR-UK
Interoperability Platform ExCo*
ELIXIR Fifth Anniversary, 11 December 2018
www.elixir-europe.org



F_{indable} A_{ccessible} I_{nteroperable} R_{eusable}

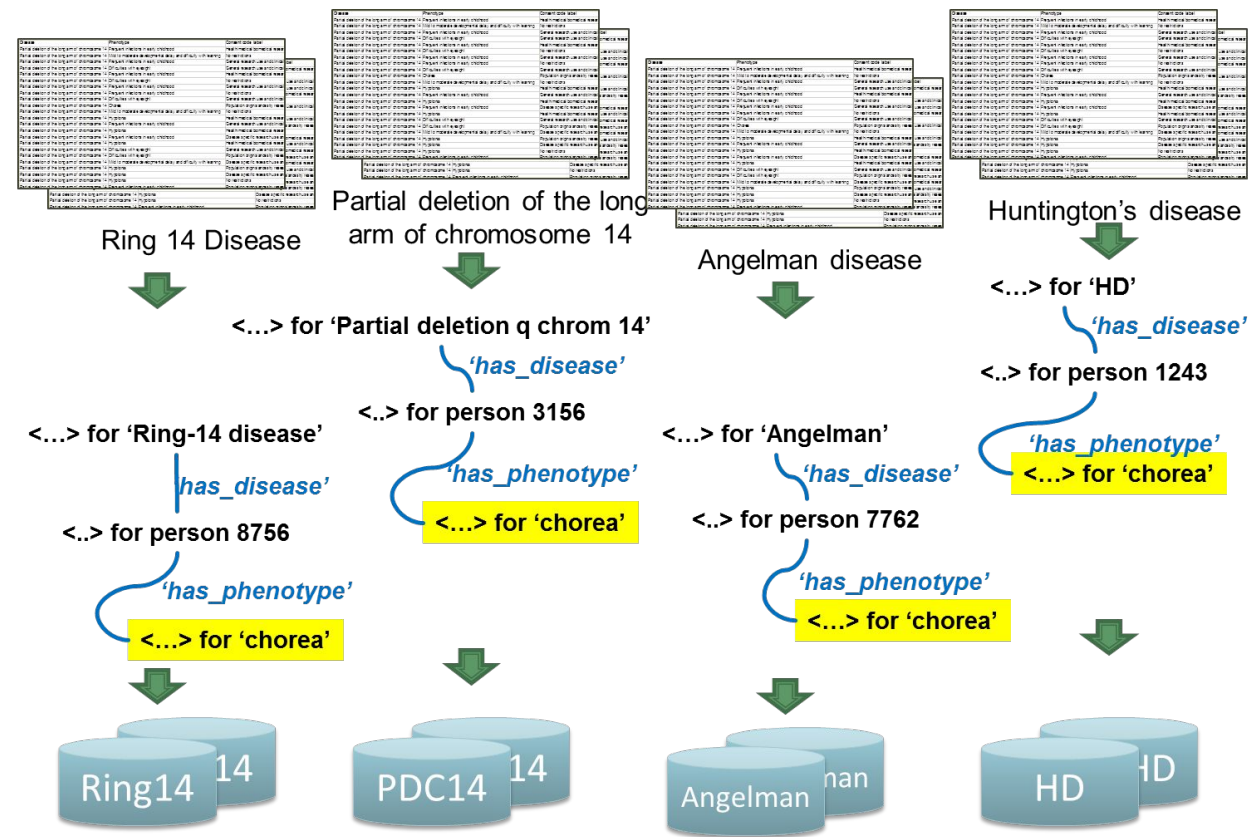
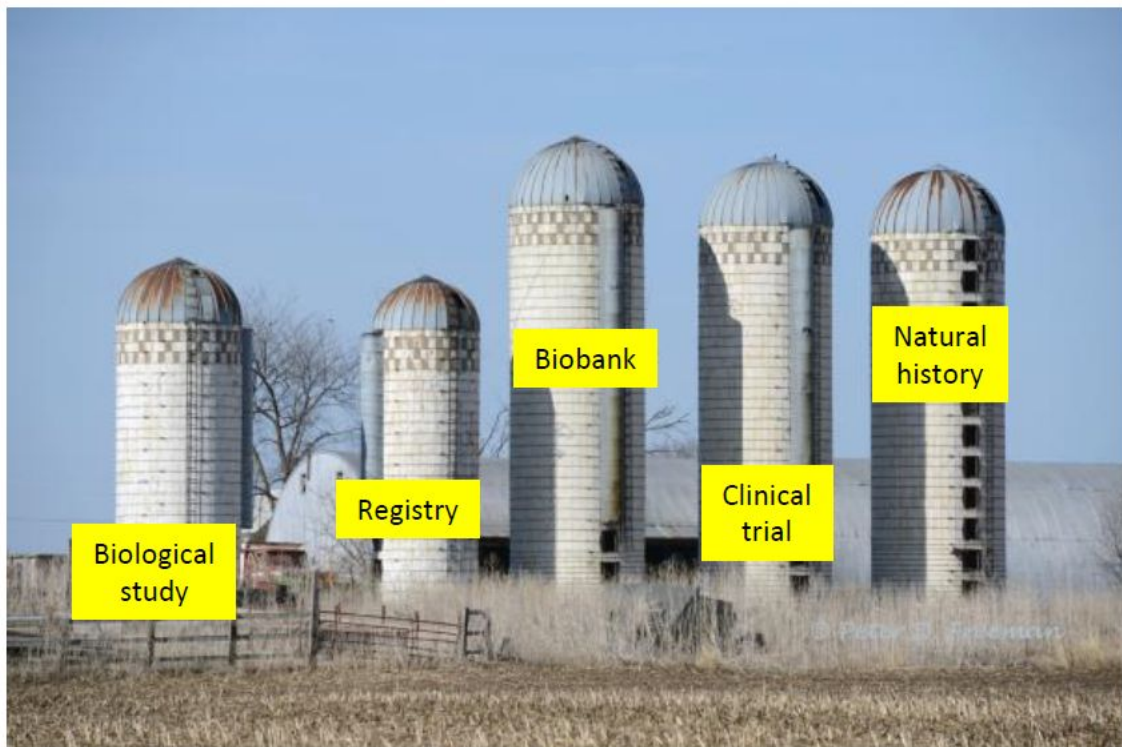


Building a suitable **FAIR**
infrastructure for
finding, exchanging, comparing,
aggregating and interlinking
biological information across
Europe





Rare Disease research

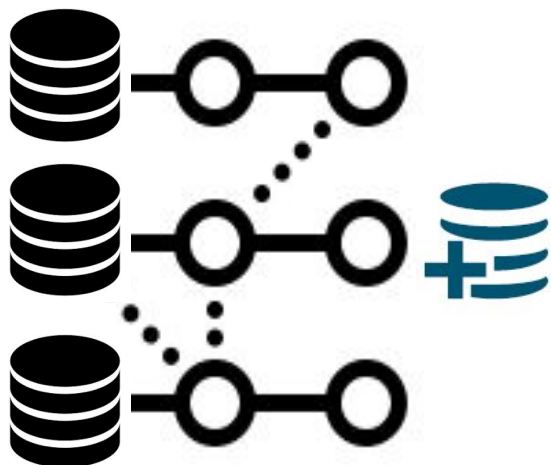


Combine more of the same data type
Link up different data types for a more complete picture





Rare Disease research



Harmonise database formats and models
Map between the terms used in the databases
Link to reference knowledge bases

Retrieval and analysis across resources



Linked Data Demonstrator

Step 1 > Retrieve:

- Get number of biosamples from donors with a specific phenotype
- Get number of persons with a specific phenotype
- Get number of biosamples from donors with a specific disease
- Get number of biosamples from donors with a specific phenotype and from a specific region**
- Get biosamples from donors with a specific phenotype and specific samplings
- Get biosamples from donors with a specific disease and a specific karyotype
- Get biosamples from donors with a specific disease, a specific karyotype and specific samplings
- Get diseases sharing phenotypes
- Get biosamples from donors sharing phenotypes

Step 2 > By which value

Region

Phenotype



Home RD-Connect ELIXIR About

Process

Step 3 > Result:

numberOfSamples	phenotype	disease	biobank	registry	region
5	Downslanted palpebral fissures	Ring chromosome 14	Galliera Genetic Bank	Ring14 Clinical database	Pistoia
5	Anteverted nares	Ring chromosome 14	Galliera Genetic Bank	Ring14 Clinical database	Pistoia
1	Mandibular prognathia	Angelman syndrome	Galliera Genetic Bank	Tuscany registry of congenital defects	Pistoia
	Depressed nasal bridge	Ataxia-telangiectasia	Research/Institute of Health Carlos III	CoF-AT study: a French cohort on ataxia-telangiectasia	Pistoia
	Depressed nasal bridge	Ring chromosome 14	Galliera Genetic Bank	Ring14 Clinical database	Pistoia
	Anteverted nares	Ataxia-telangiectasia	Biobank of the Institute of Rare Diseases	CoF-AT study: a French cohort on ataxia-telangiectasia	Pistoia



ID # 77350

Date of Inclusion: 01/04/2015

Last Activities: 04/02/2016



Galliera Genetic Bank



<http://ggb.gal>

Host institution

Laboratory of Human Genetics - E.O.

Tel.003901056

Ospedali Galliera

Via Alessandro Volta, 6

Fax.003901056

16128 Genova

Italy

Overview


[17]

Diseases

[132]

Personnel

Main contact



I would like to know the number of **samples** of donors with **an abnormality in head or neck** in a **specific region** of Italy, in order to check if exposure to environmental factors is important

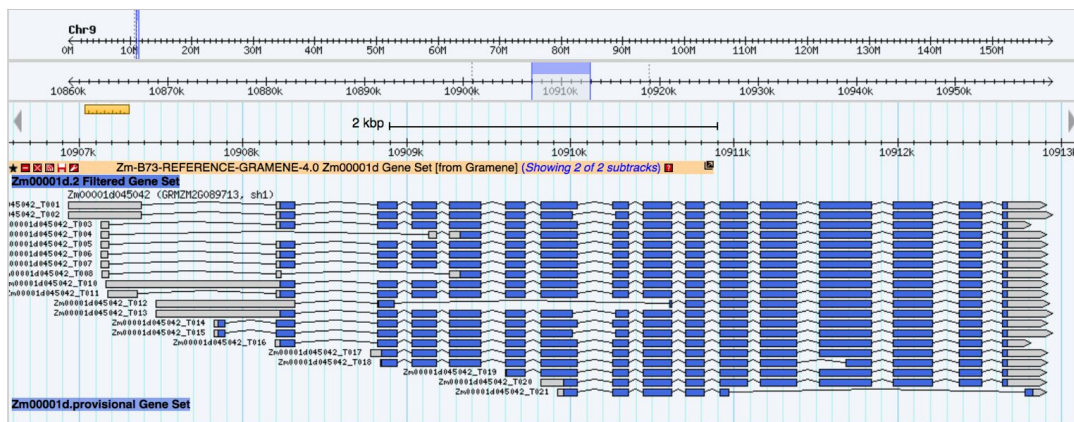
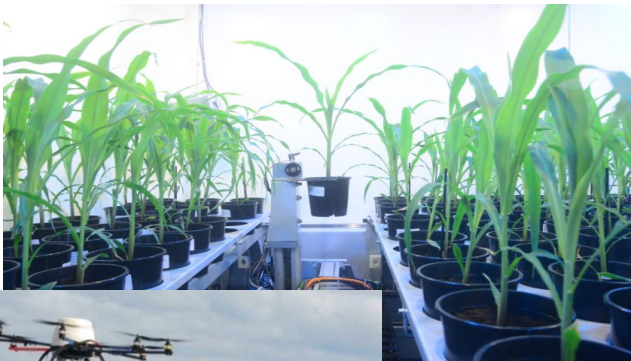
In addition, I would like to see in which **biobanks** I can find the samples, the **phenotypes** associated with them, and information about the **organisation(s)** behind the biobanks or registries





Genotypic and Phenotypic data for Crop and Forest Plants

Large scale
automated
phenotyping



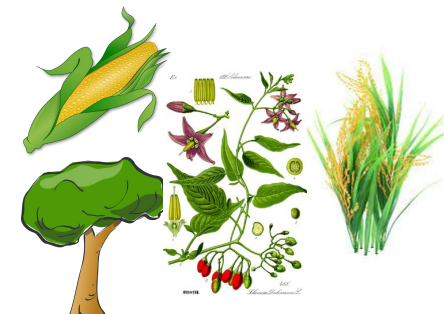
High throughput genomics

PHENO

BreeDB



TropGENE



Standards for representation of
genotypic and phenotypic data

Make data discoverable and
interoperable by common APIs

Annotate datasets to deposit
into public archives

Crop Ontology
for agricultural data



Maize Plant Height



Identifier
CO_322:
0000007

Identifier
CO_322:
0000994

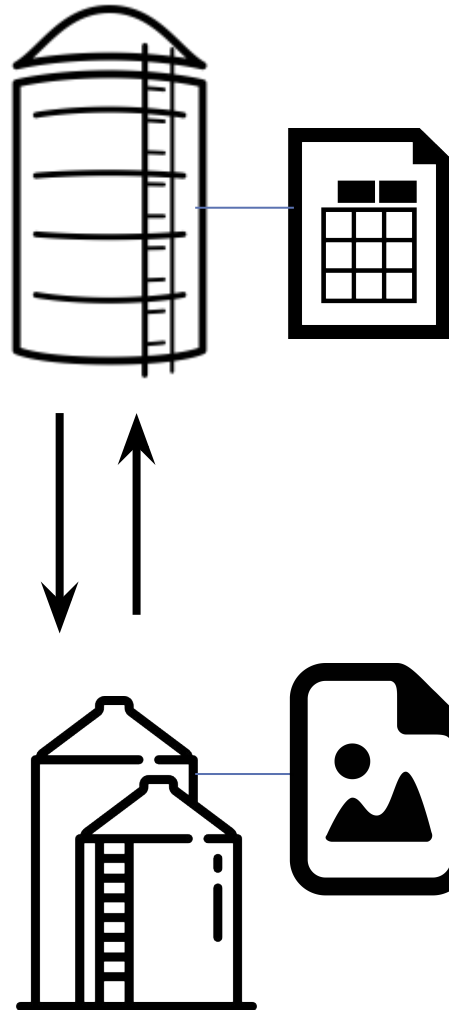
Arabidopsis Leaf Length



?



Common Agreements for IDs & Descriptions Standards, Link Points



700
types



Is the same identifier
being used for X?
Can they be linked?

224



Are the formats the
same? Can they be
mapped?

754



Are terms being used consistently?
Are terms being used in common?

122



Are the same things
being reported in the
same way?



Do (micro)
services have the
same or compatible
APIs?



Cedar
waxwing



© Copyright - All Rights Reserved 2010 Image of Bombycilia cedrorum

BOLD (Barcode of
life)

Taxon:9606

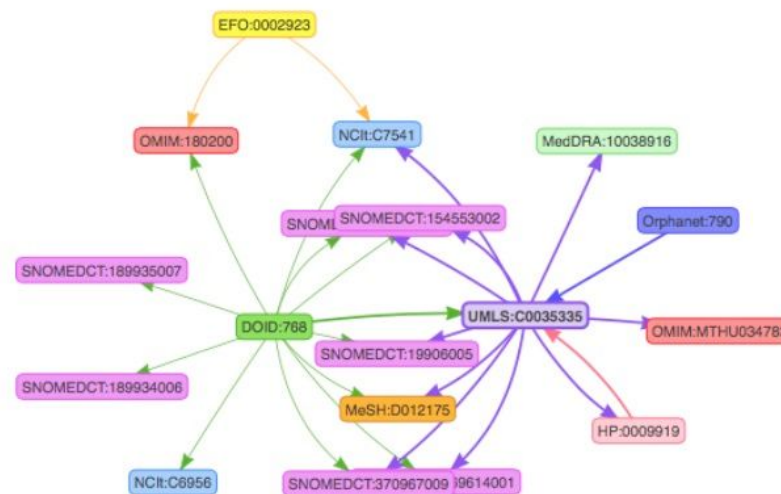


NCBI
Taxonomy

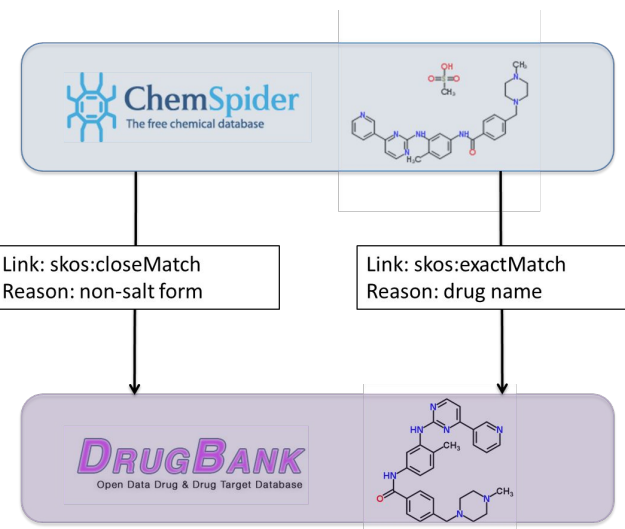
Arabian tea
plant



GRIN plant
taxonomy



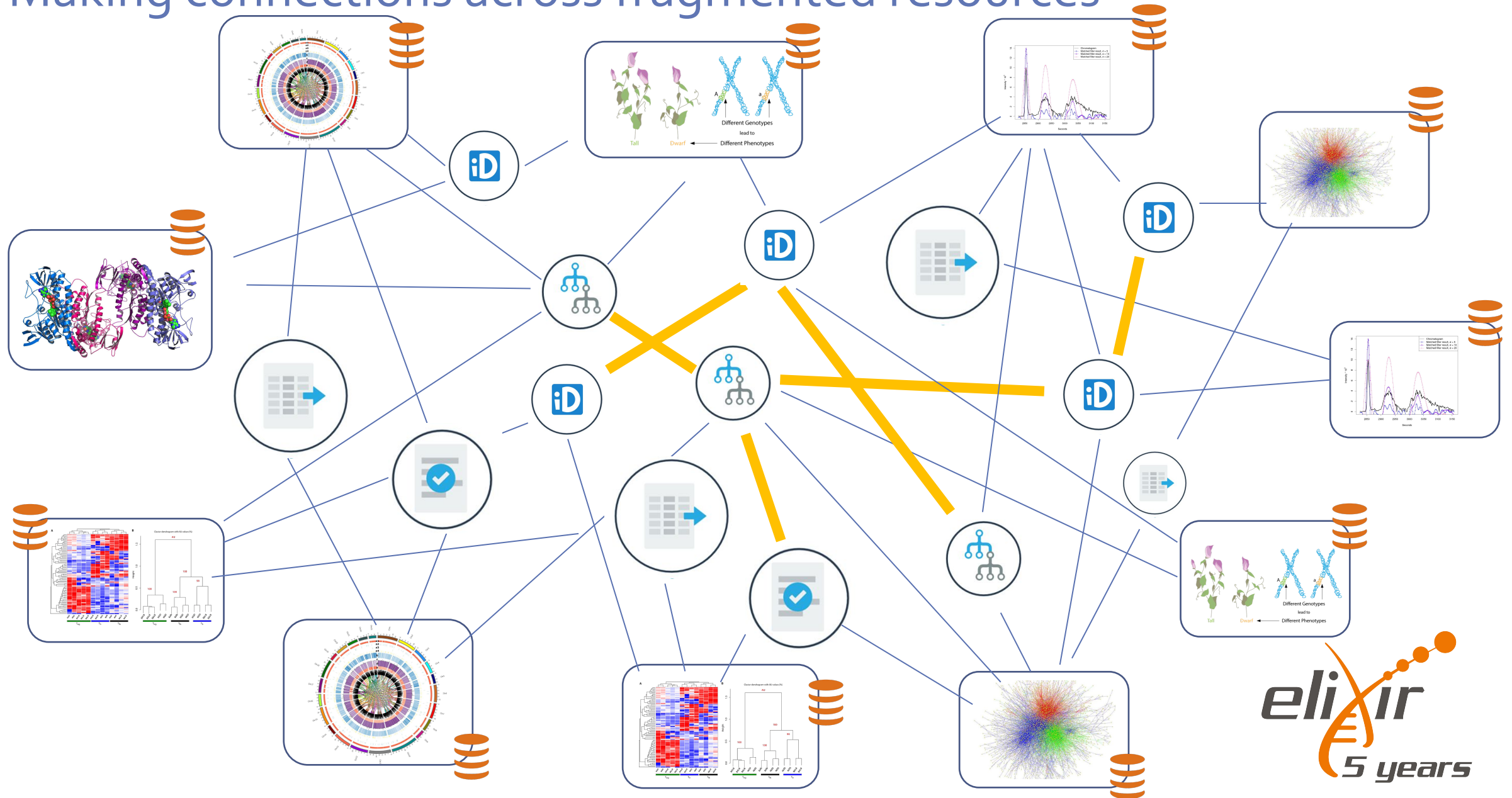
Mappings across Ontologies for NCIt
(Retinoblastoma)*



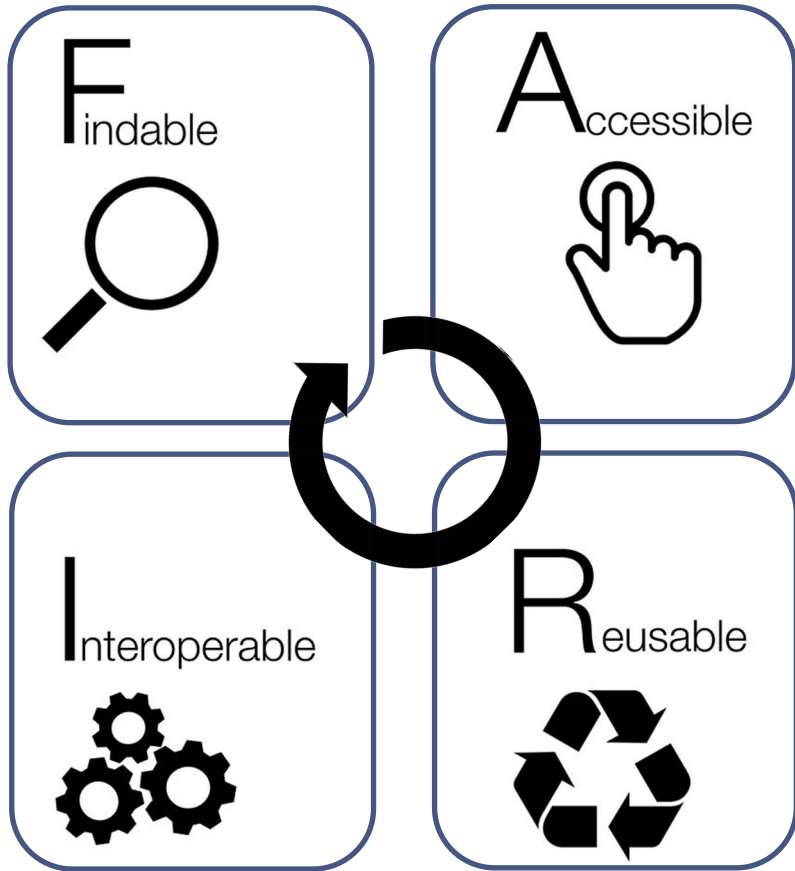
Mappings across databases for the
same entity**



Making connections across fragmented resources



Hence.... FAIR Data Principles



Registration and search

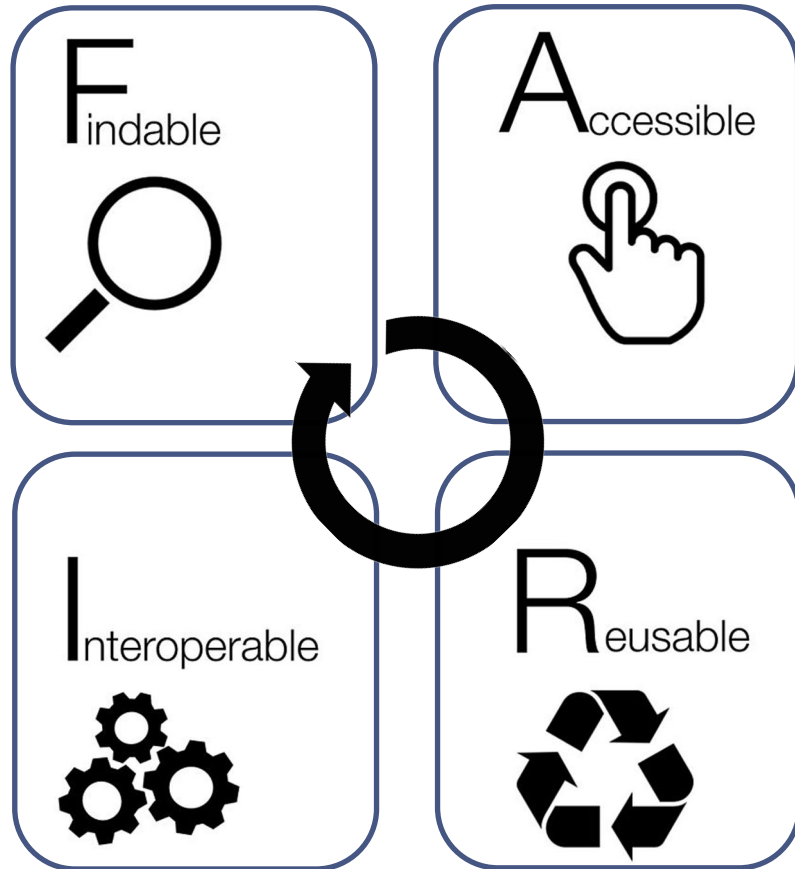
Persistent and reused **identifiers**

Common, structured, interlinked **metadata**

Open **access protocols**

Machine processing

Turning FAIR Data Principles into Reality



Open
Standards



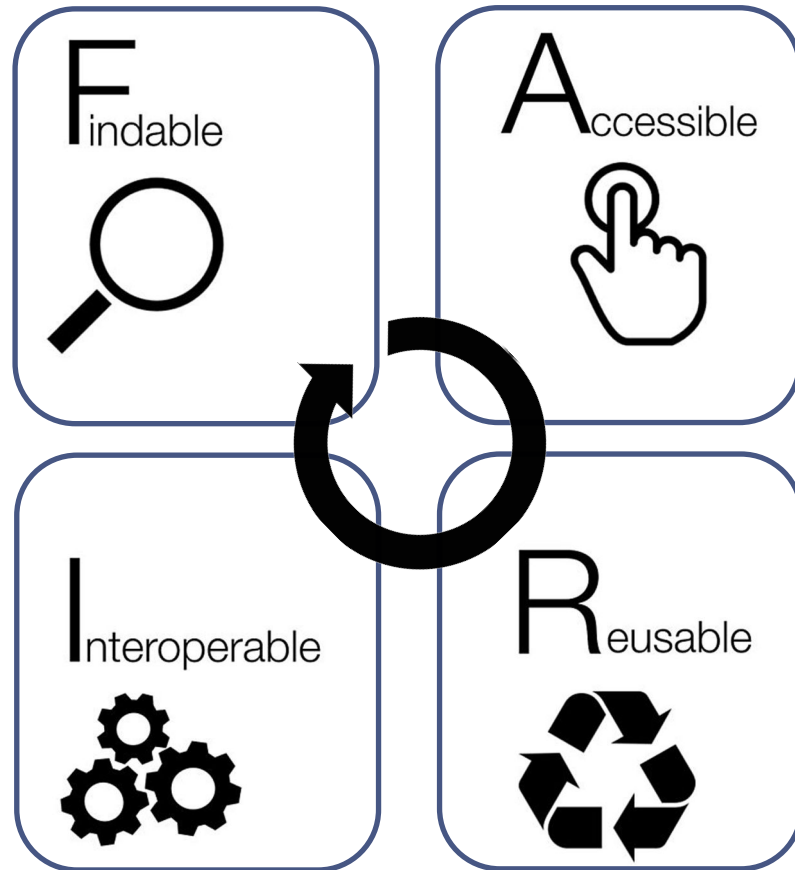
Machine
processable



Services &
Resources

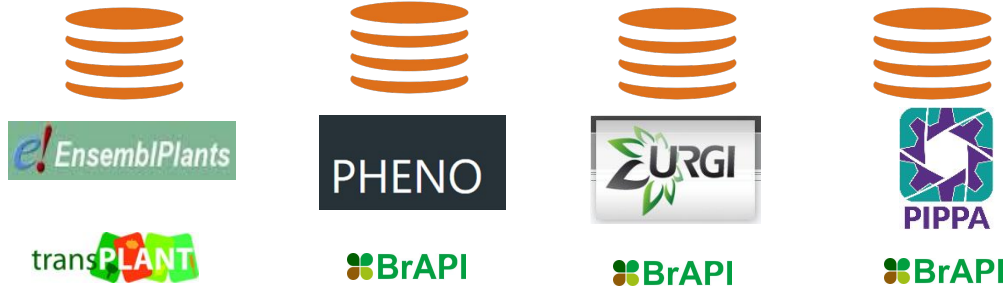


Interoperability Resources





Genotypic and Phenotypic data for Crop and Forest Plants Interoperability Resources



SEARCH FORM

Geoplasm Variable

Crop Select Some Options

Geoplasm list (panel, collection, population) Select Some Options

Accession B73

RESULTS 29 match found

Phenotyping Study URG1 Gmpt5 ASSO_S1P9_FFLW8_Average.Gif-sur-Yvette.Saint-Martin-de-Hinx.Einbeck

"ASSO_S1P9_FFLW8_Average.Gif-sur-Yvette.Saint-Martin-de-Hinx.Einbeck" is a phenotyping study conducted from 2002-01-01 to 2005-01-01 (seasons: 2002,2003,2004,2005) in Average.Gif-sur-Yvette.Saint-Martin-de-Hinx.Einbeck (France).

Geoplasm VIB PIPPA B73 RIL 2-way

"B73 RIL 2-way" is a Zea mays (Maize) accession.

Geoplasm VIB PIPPA B73 RIL 8-way

"B73 RIL 8-way" is a Zea mays (Maize) accession.

Geoplasm URG1 Gmpt5 B73_inra

"B73_inra" is a Zea mays (Maize) accession (number: "B73_inra") managed by INRA.

Phenotyping Study VIB PIPPA RIL 2-way_batch1

"RIL 2-way_batch 1" is a phenotyping study conducted from 2011-09-21 to 2011-10-18.

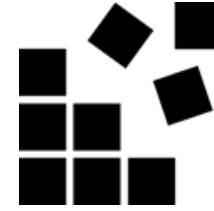
Phenotyping Study VIB PIPPA RIL 2-way_batch10

© INRA-URGI 2006-2018 | Legal mentions | Designed by SWAD | Contributors

ELIXIR Plant Data Lookup Service



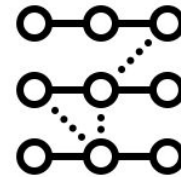
BioSamples



Registries for the standards and ontologies



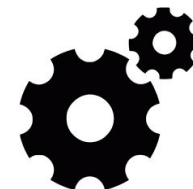
Look up services for the identifiers and concepts



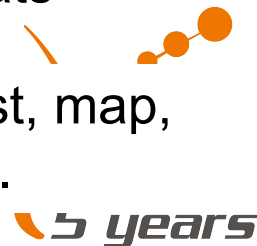
Map between different concepts and identifiers for same thing.



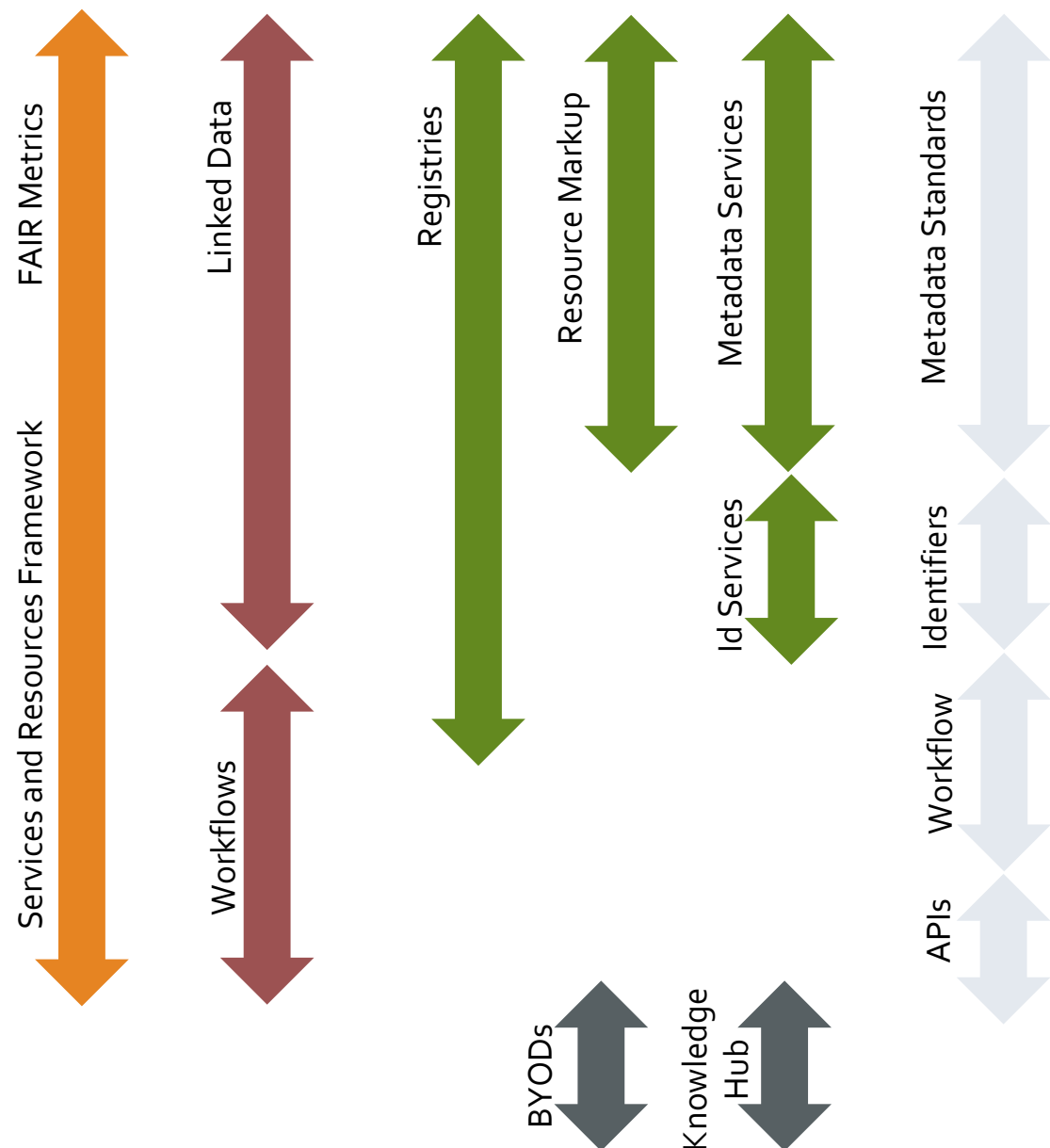
Services to help annotate & validate databases and data submissions against to reporting guidelines & formats



Services to harvest, map, search metadata ..



What Interoperability Resources are needed?



Standards: formats, reporting guidelines, ontologies

Metadata services: ontology, annotation, validation, harvesting, Indexing

Register services and datasets

Search engine for datasets.

Identifier resolution & management

Identifier mapping services

Describing and sharing **workflows** between different systems

Harmonisation of **tools and pipelines**

Common **Programmable Interfaces**

Best practice.



Marine



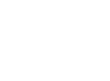
Plants



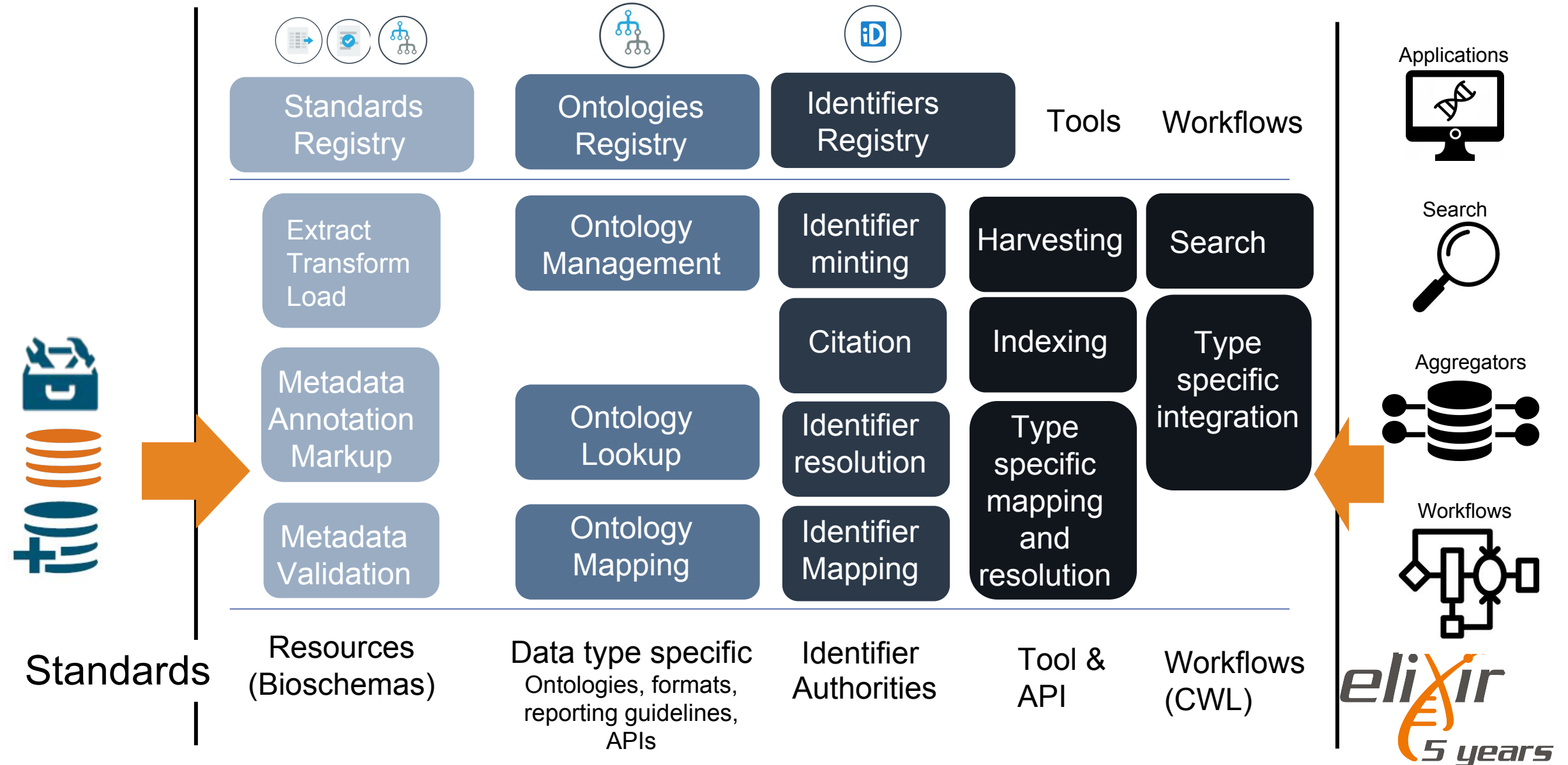
Rare
Disease



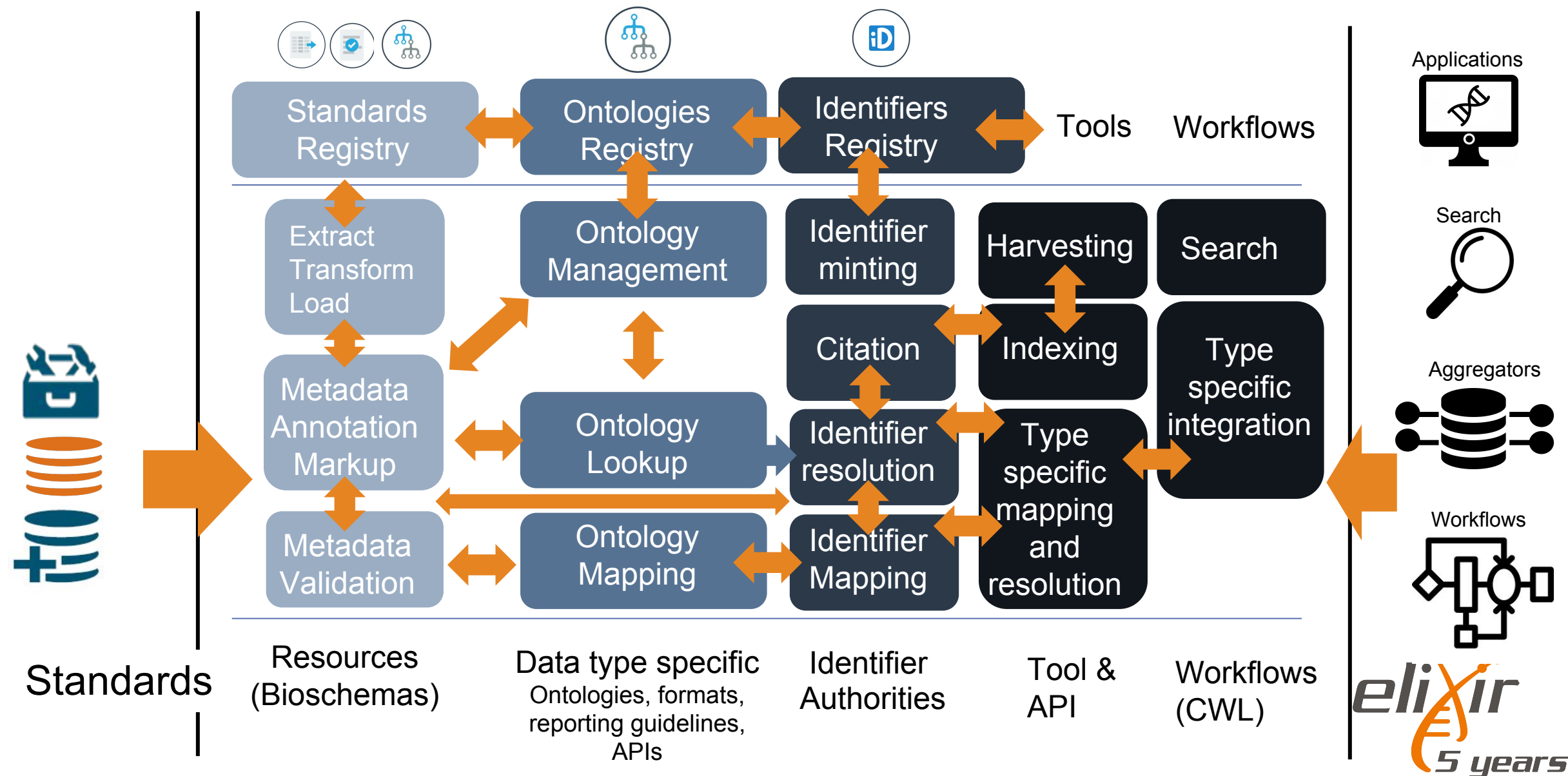
Human
Data



ELIXIR Interoperability Resources Framework



Interoperable ELIXIR Interoperability Resources Framework

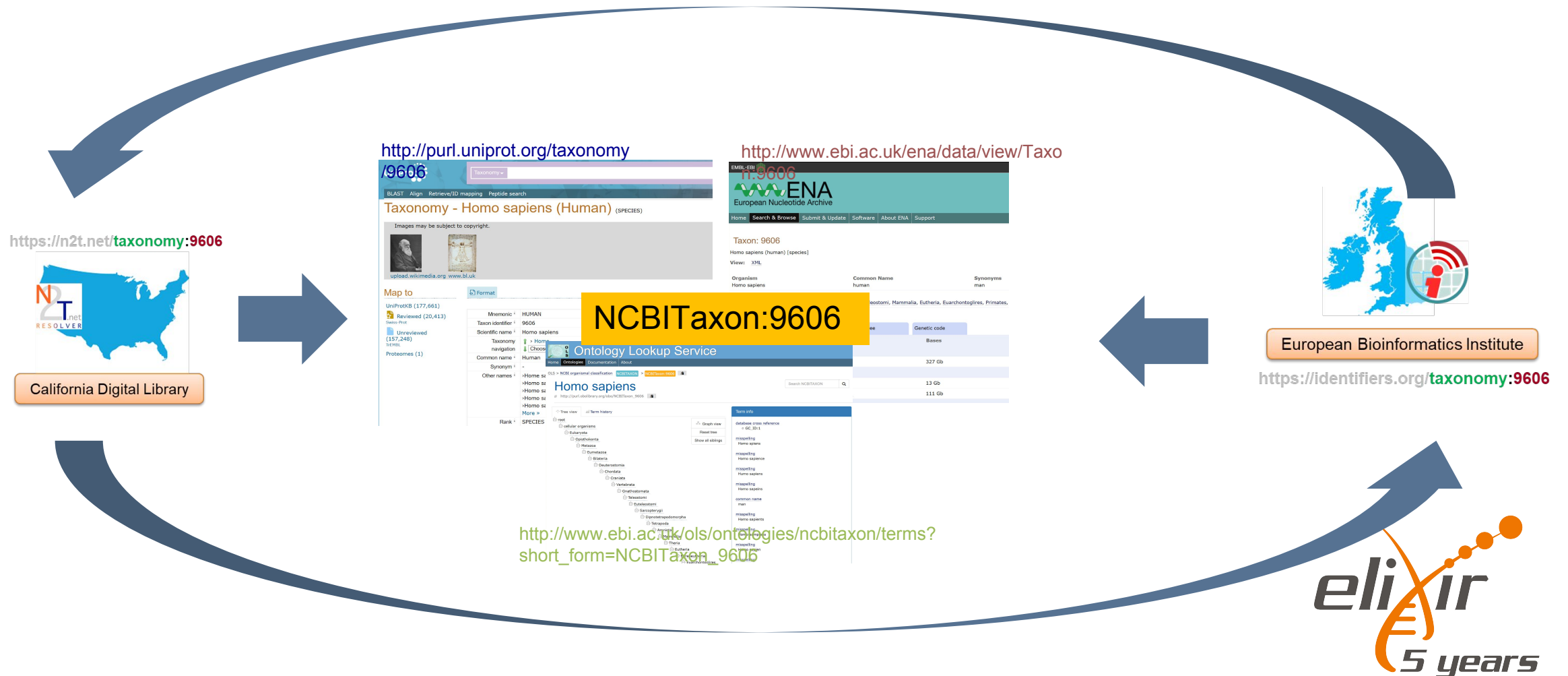


Example: Identifier Resolution of Data on the Web

Multiple URLs for the same collection make object unification challenging

Resolution Services keep track and handle the different locations and different identifier systems

The Resolution Services themselves are harmonised



International Interoperability Resources

ELIXIR is part of a global ecosystem



What are Recommended Interoperability Resources?

An ELIXIR Service supplied by one or more Nodes



What are Recommended Interoperability Resources?

An ELIXIR Service supplied by one or more Nodes

Plays important
role in our
interoperability
framework

helps

establish connections between data (and other) resources

acquire and expose metadata of data (and other) resources

create infrastructure needed to build integrable data collections

use interoperability resources to support delivery of FAIR principles

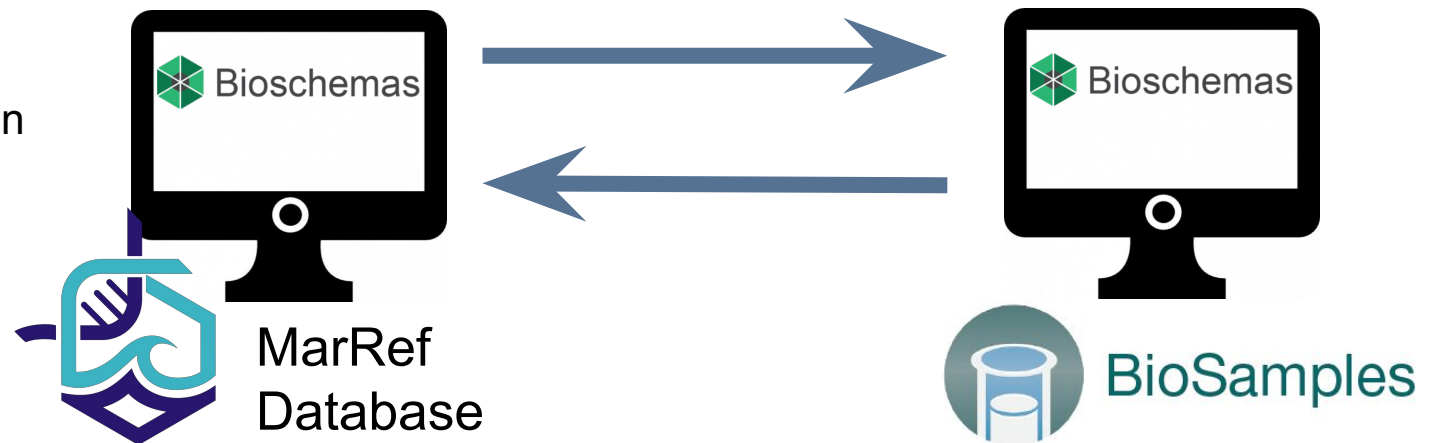


An Interoperability Resource for findability & metadata exchange for all of ELIXIR's Resources



Metadata about web based resources using a **widely adopted web standard** in a community agreed way

schema.org



An Interoperability Resource for findability & metadata exchange for all of ELIXIR's Resources



Bioschemas

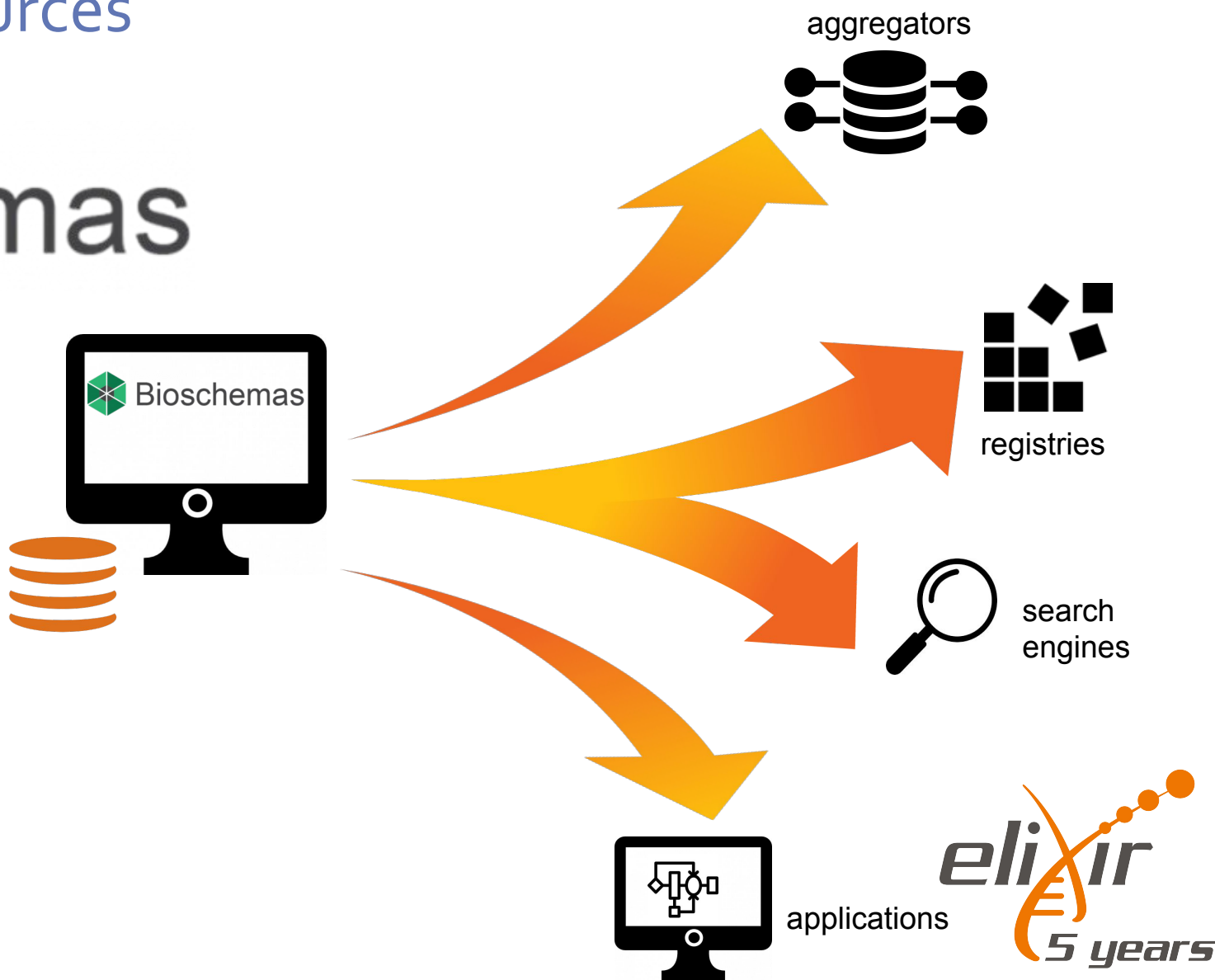
Metadata about web based resources using a **widely adopted web standard** in a community agreed way

55

Live
deploys

6M+

Pages

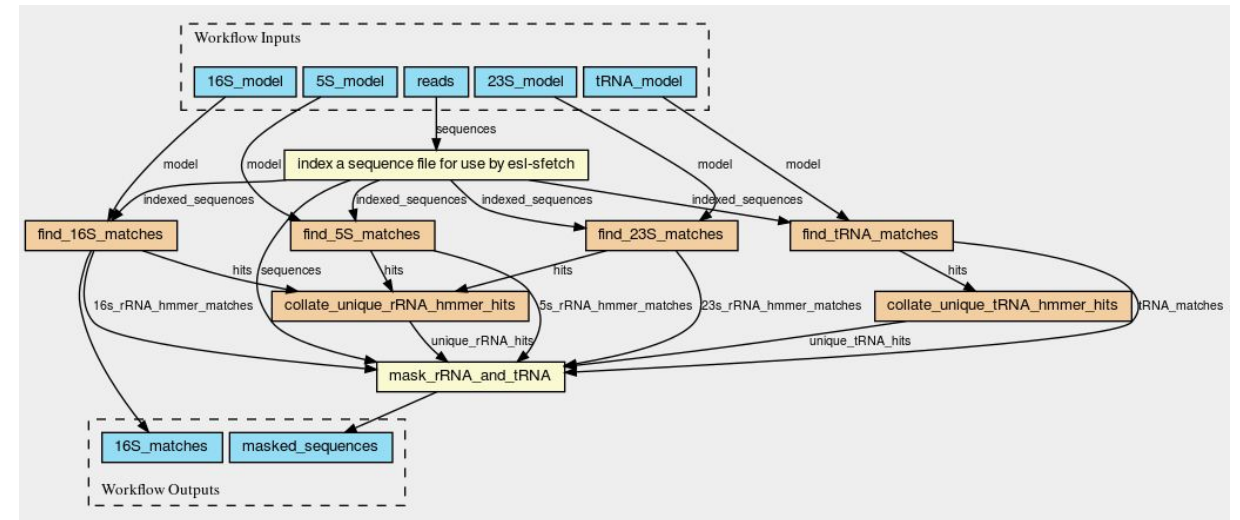




An Interoperability Resource for findability and exchange of ELIXIR's workflows and pipelines



COMMON
WORKFLOW
LANGUAGE

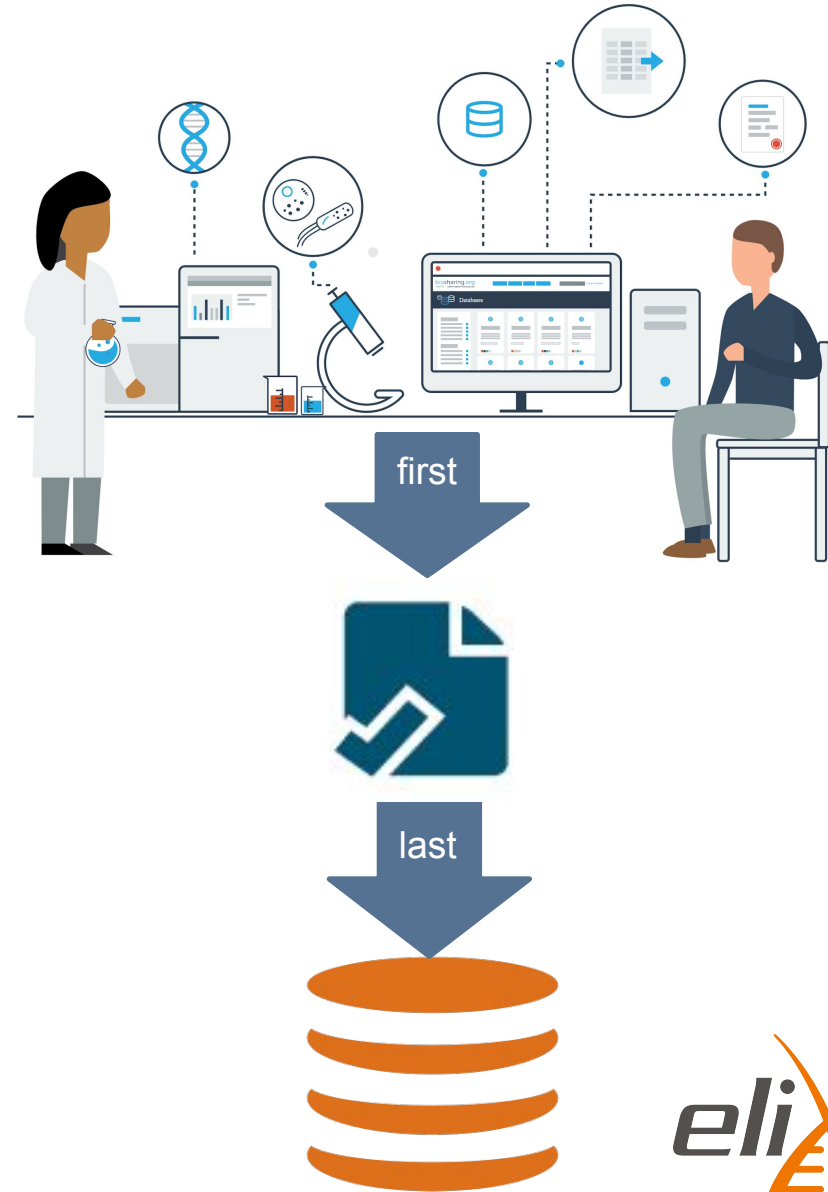


Pioneered by Marine Metagenomics

Resources gap

FAIR metadata
at source

“The first and last mile”



First round of
Recommended Interoperability Resources
completes process



First round of
Recommended Interoperability Resources
completes process

tomorrow!





RIRs are ELIXIR added value to enable
FAIR Core Data Resources
(and other ELIXIR resources)

Oversee quality and reliability



Develop an integrated portfolio



Support sustainability



Acknowledgements



Special Thanks:

Michael Crusoe
Rafael Jimenez
Alasdair Gray
Stian Soiland-Reyes
Susanna Sansone
Simon Jupp
Tony Burdett
Sira Sarntivijai
Jerry Lanfear
Nick Juty
Sarala Wimalaratne
Frederik Coppens
Justin Clark-Casey
Peter McQuilton
Robert Finn

Marco Roos
And many more!

Leadership



Carole Goble
(ELIXIR UK)



Chris Evelo
(ELIXIR Netherlands)



Helen Parkinson
(EMBL-EBI)



Sirarat Sarntivijai
(Platform Coordinator,
ELIXIR Hub)



European Union Horizon 2020 program under grant agreement 739563.



European Union Horizon 2020 program under grant agreement 654248.



Center of Excellence for Computational Biomolecular Research

European Union Horizon 2020 program under grant agreement 675728.



European Union Horizon 2020 program under grant agreement 676559



Implementation Studies
CWL and Bioschemas





Thank you!

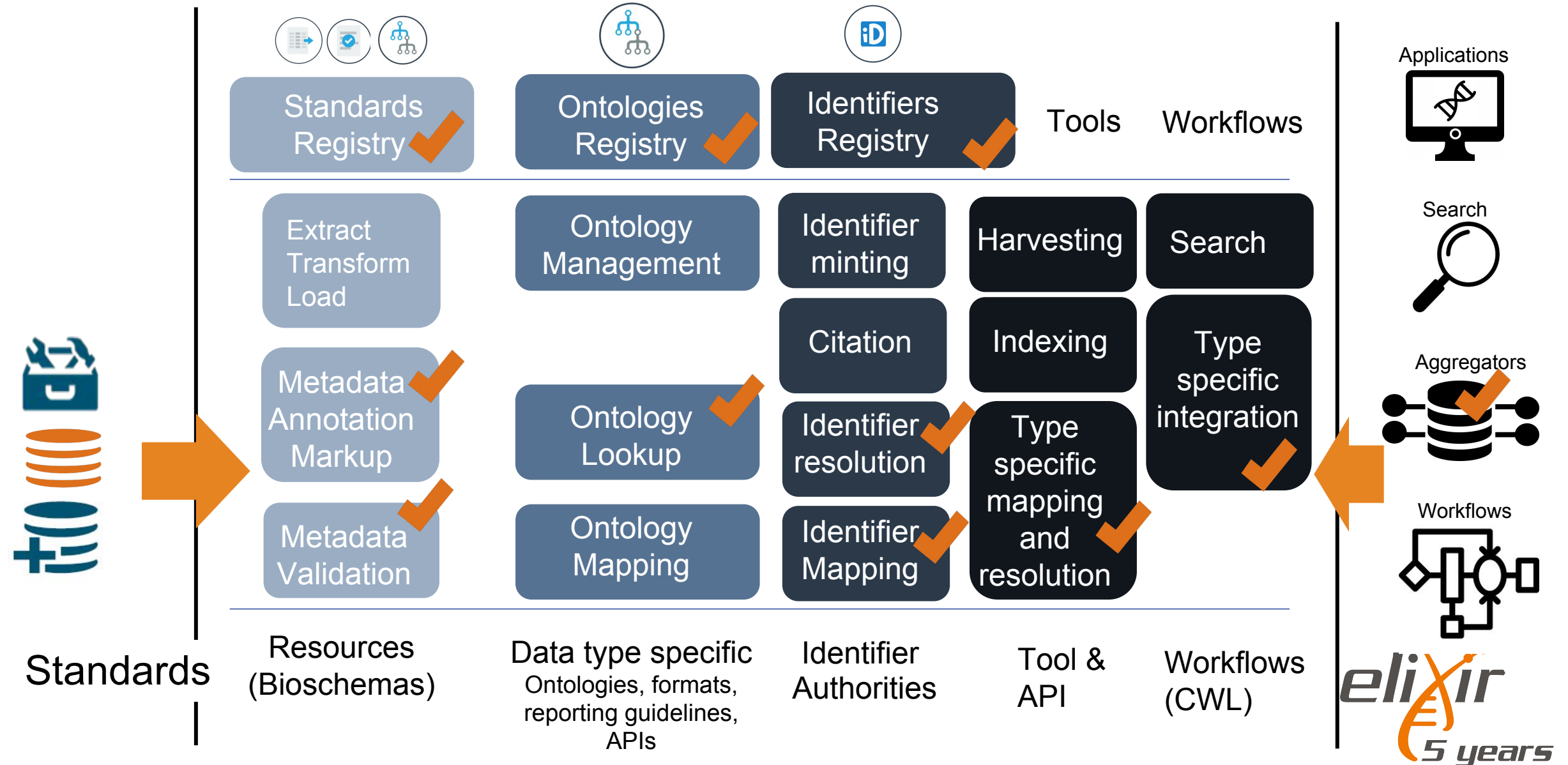


ELIXIR-EXCELERATE is funded by the European Commission within the Research Infrastructures programme of Horizon 2020, grant agreement number 676559.











www.elixir-europe.org



ELIXIR Interoperability Resources Framework



The 2018 Recommendations of RIRs

Resource		Description
FAIRsharing		Registry of curated metadata of DBs, Policies, Standards
g:Profiler		Gene-centric data integrator - Web UI, and API services
Identifiers.org		Persistent URL provider & identifier resolver
Intermine		DIY database portal builder for model organism of choice
ISA Framework		Curation for metadata of experiments (Project -> Study -> Assay)
Ontology Lookup Service (OLS)		Google-like ontology term search
3DBIONOTES API*		API aiding protein annotation by calling info from ref. resources
BridgeDb		Identifier mapping for cheminformatics domain
DisGeNET API*		API SPARQL Endpoint for genetic variant - human disease data
MOLGENIS		Bioinformatics data integrator suite - explore/annotate/exchange