



# ELIXIR Hands-on workshop Tromsø Nov. 2019

*The national bioinformatics infrastructure*



Erik Hjerde  
Ståle Nygård  
Nils Peder Willassen

[www.elixir-europe.org](http://www.elixir-europe.org)

# Time schedule and course material

<https://elixir.mf.uni-lj.si/course/view.php?id=50>



ELIXIR-SI

English (en) ▼

## ELIXIR Norway: Hands-on bioinformatics NeLS workshop

[Home](#) / [My courses](#) / [hands-on\\_nels\\_NO\\_2019](#)





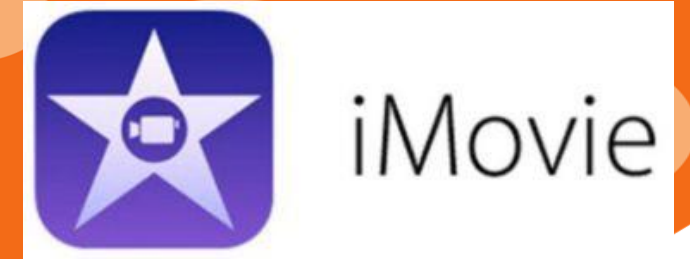
# Introduction to ELIXIR

*A distributed European infrastructure for life-science information*



Erik Hjerde

[www.elixir-europe.org](http://www.elixir-europe.org)



[www.elixir-europe.org](http://www.elixir-europe.org)

# Do we need ELIXIR???????

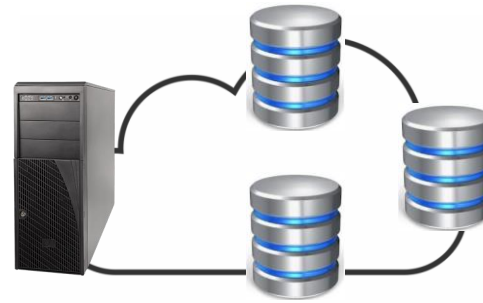
Data production



Data storage/sharing



Data compute



Data archiving



Data management



# Do we need ELIXIR????????

100 samples

200 GB raw data  
2 TB processed data

512 GB RAM  
72 CPUs  
100 000 hours

100 PE FASTQ  
100 x ? MAGs

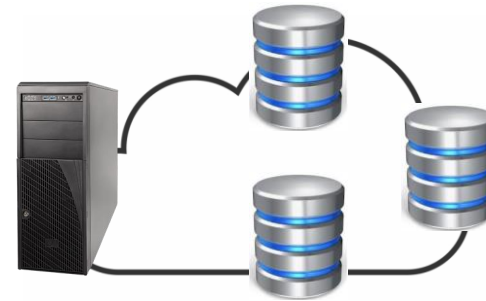
Data production



Data storage/sharing



Data compute



Data archiving



Data management



Data management plan



# Do we need ELIXIR????????

100 samples

200 GB raw data  
2 TB processed data

512 GB RAM  
72 CPUs  
100 000 hours

100 PE FASTQ  
100 x ? MAGs

Data production



Data storage/sharing



Data compute



Data archiving



Data management



# Do we need ELIXIR????????

100 samples

200 GB raw data  
2 TB processed data

512 GB RAM  
72 CPUs  
100 000 hours

100 PE FASTQ  
100 x ? MAGs

Data production

Data storage/sharing

Data compute

Data archiving



Data management

Training end users

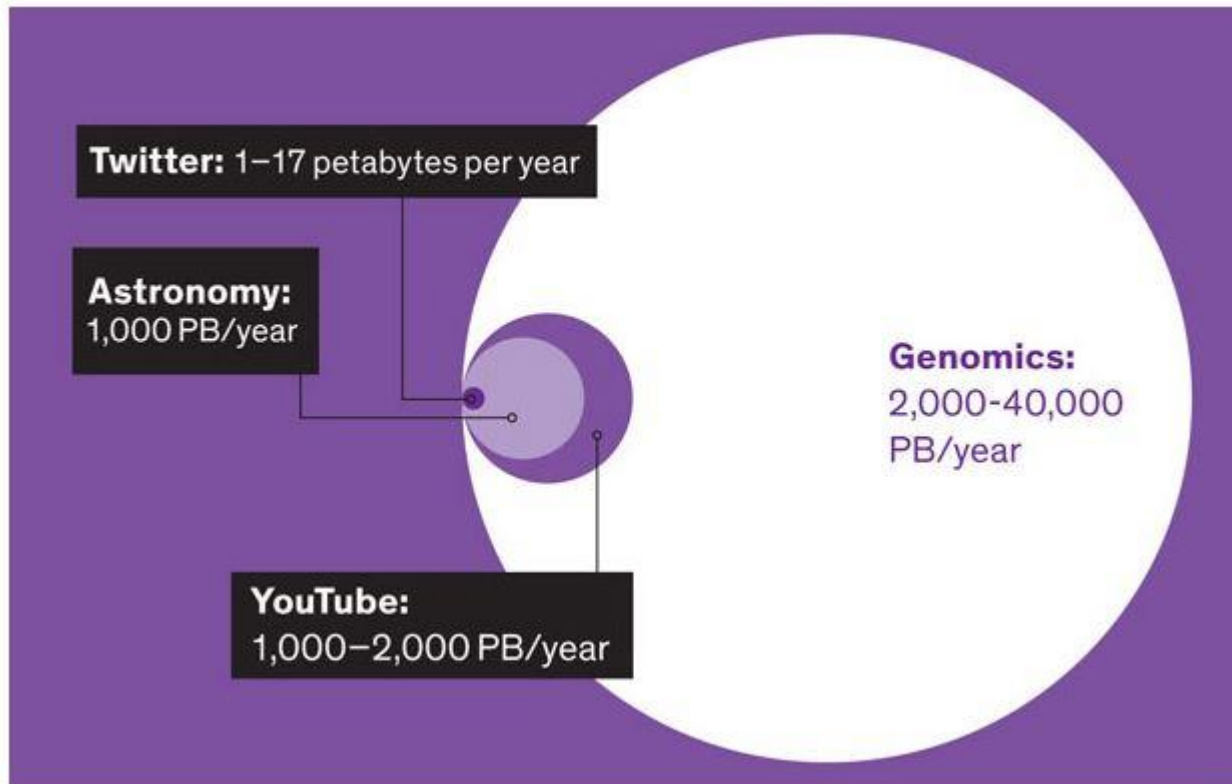




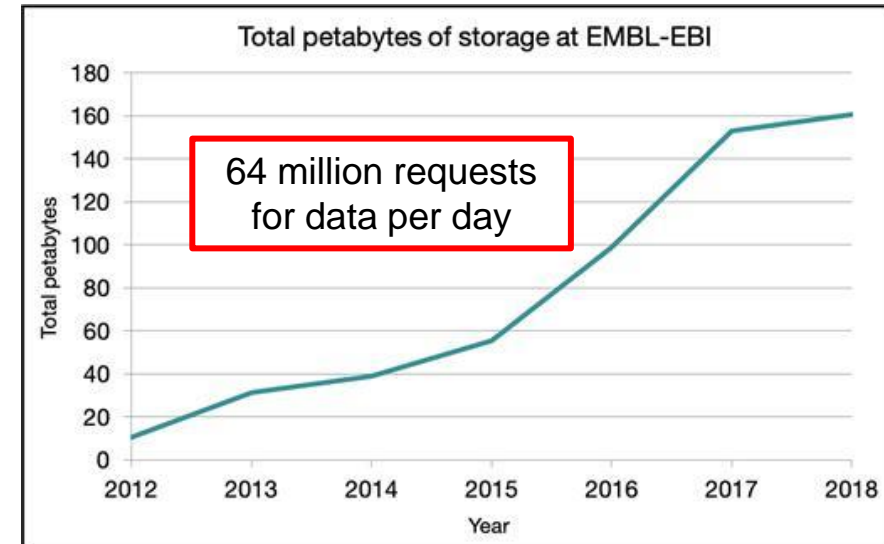
# Production of data is enormous

1 PB is equivalent to over 4,000 digital photos per day, over your entire life

## Projected annual storage in 2025

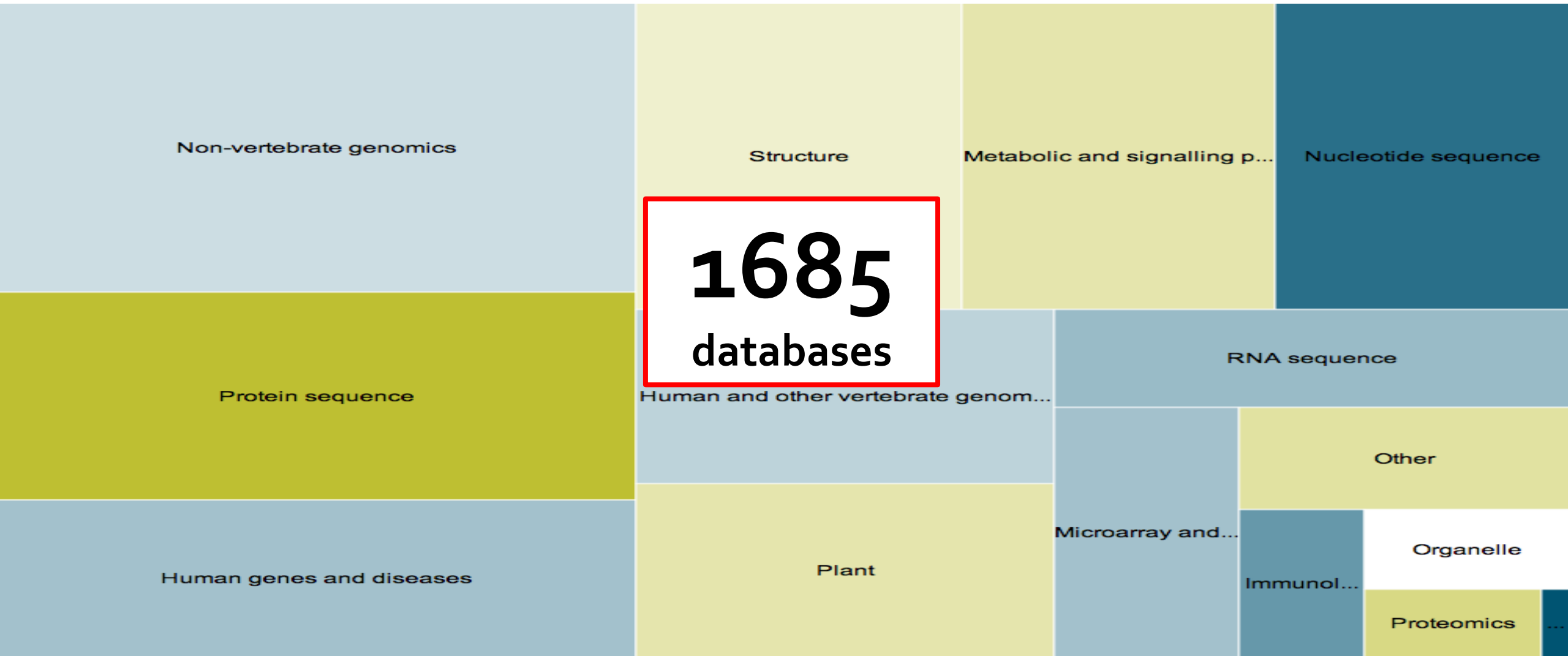


Source: "Big Data: Astronomical or Genomical?" PLoS Biology, 7 July 2015.



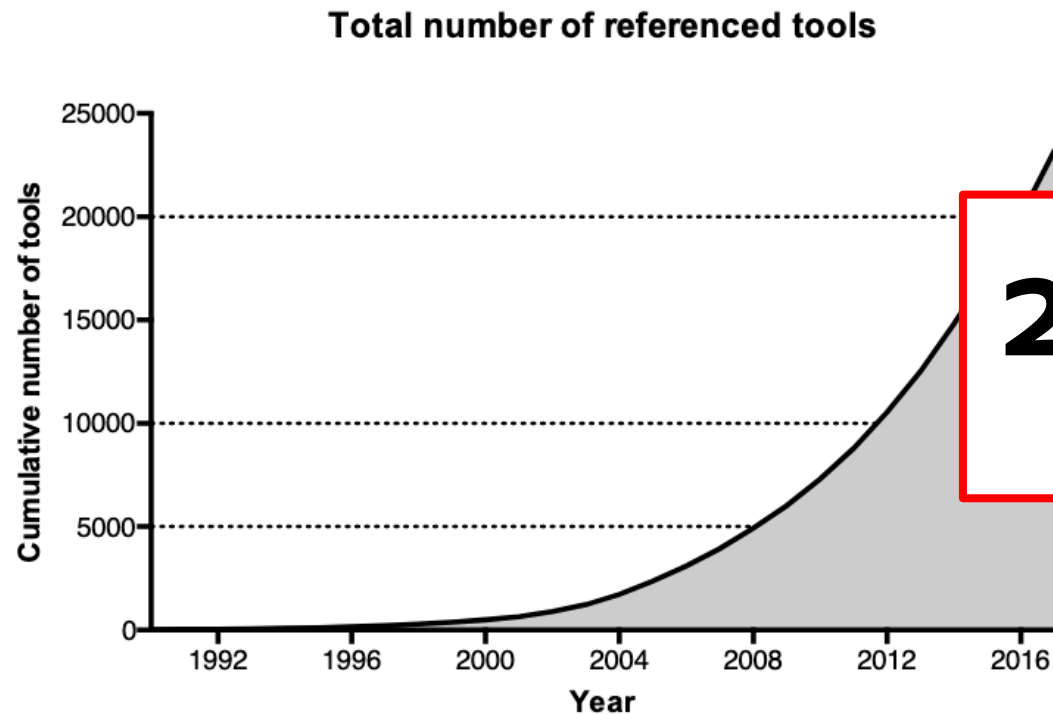
# Life-science databases

Dispersed throughout the world, and serve many different research communities

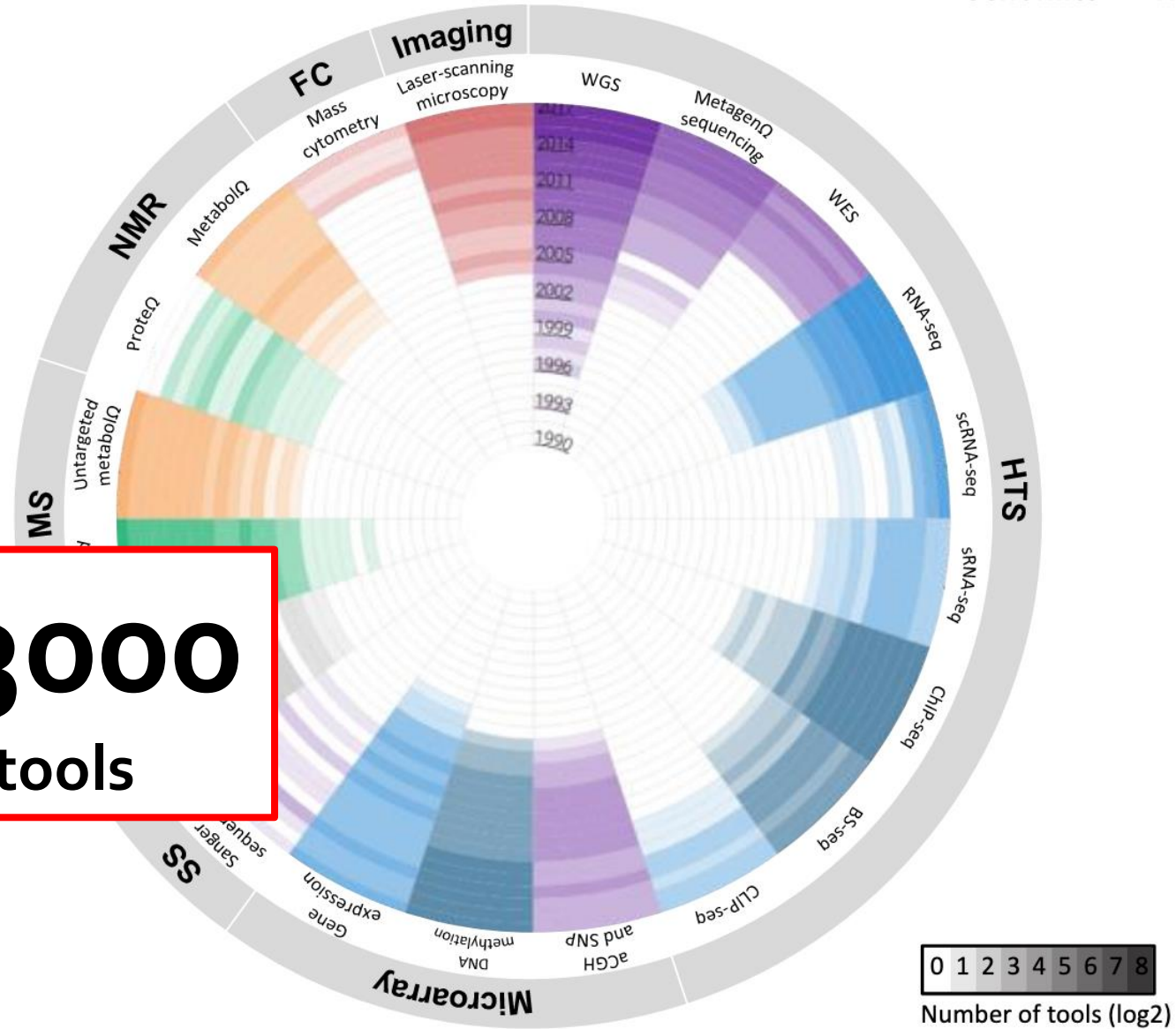


# Bioinformatic tools

Crucial component of the infrastructure for life sciences



**23000**  
tools



Clément Levin et al. 2018



# We need common standards to describe data

Common data formats

Common ontology to describe eg. a gene

Common standards to describe metadata

Database registry

Tool registry



# What is ELIXIR?

*ELIXIR connects national bioinformatics centres and EMBL-EBI into a sustainable European infrastructure for biological research data*



environment



bioindustries



agriculture



medicine

*ELIXIR underpins  
life science research  
– across academia  
and industry*



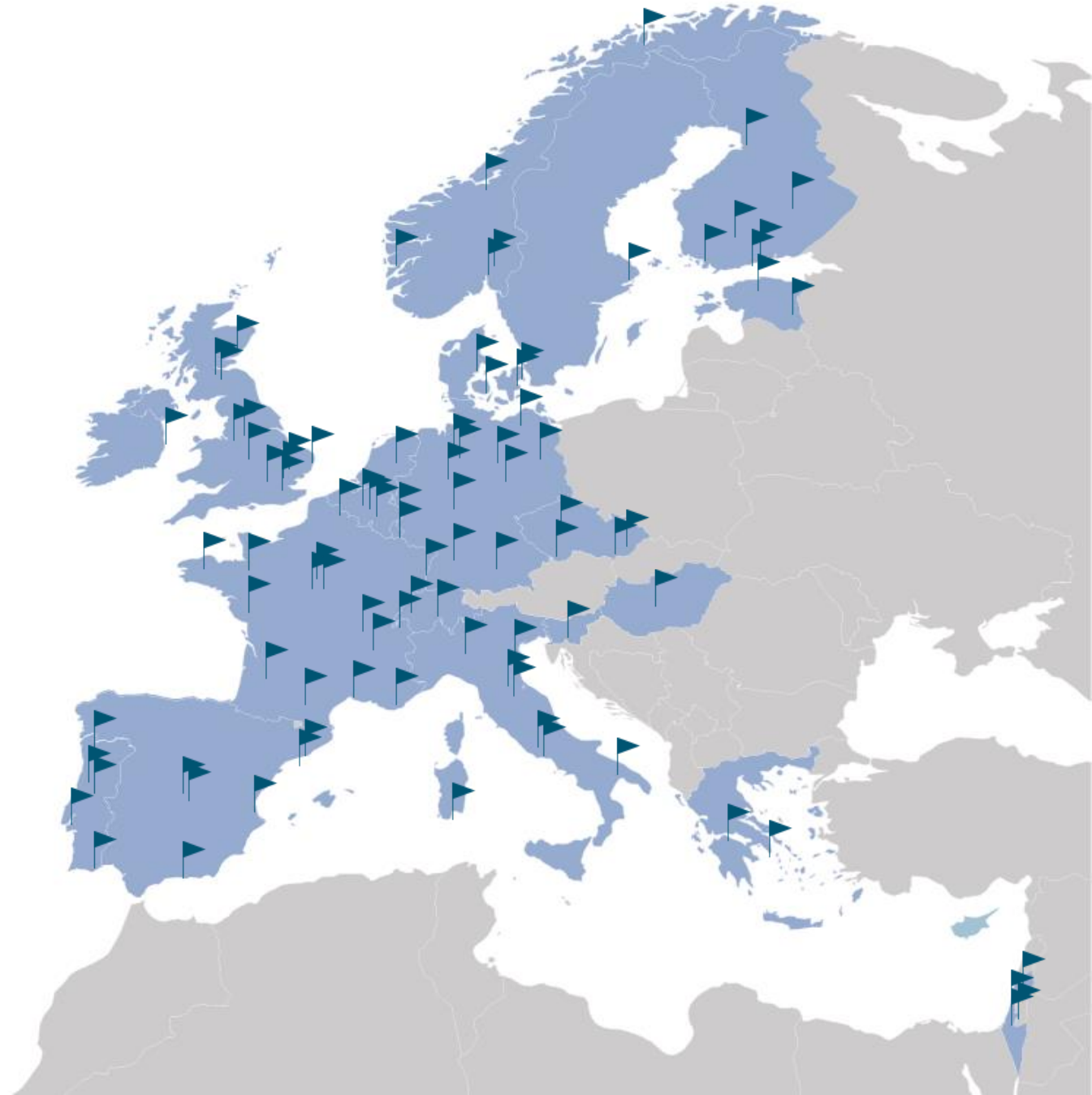


# ELIXIR members

## ELIXIR Members

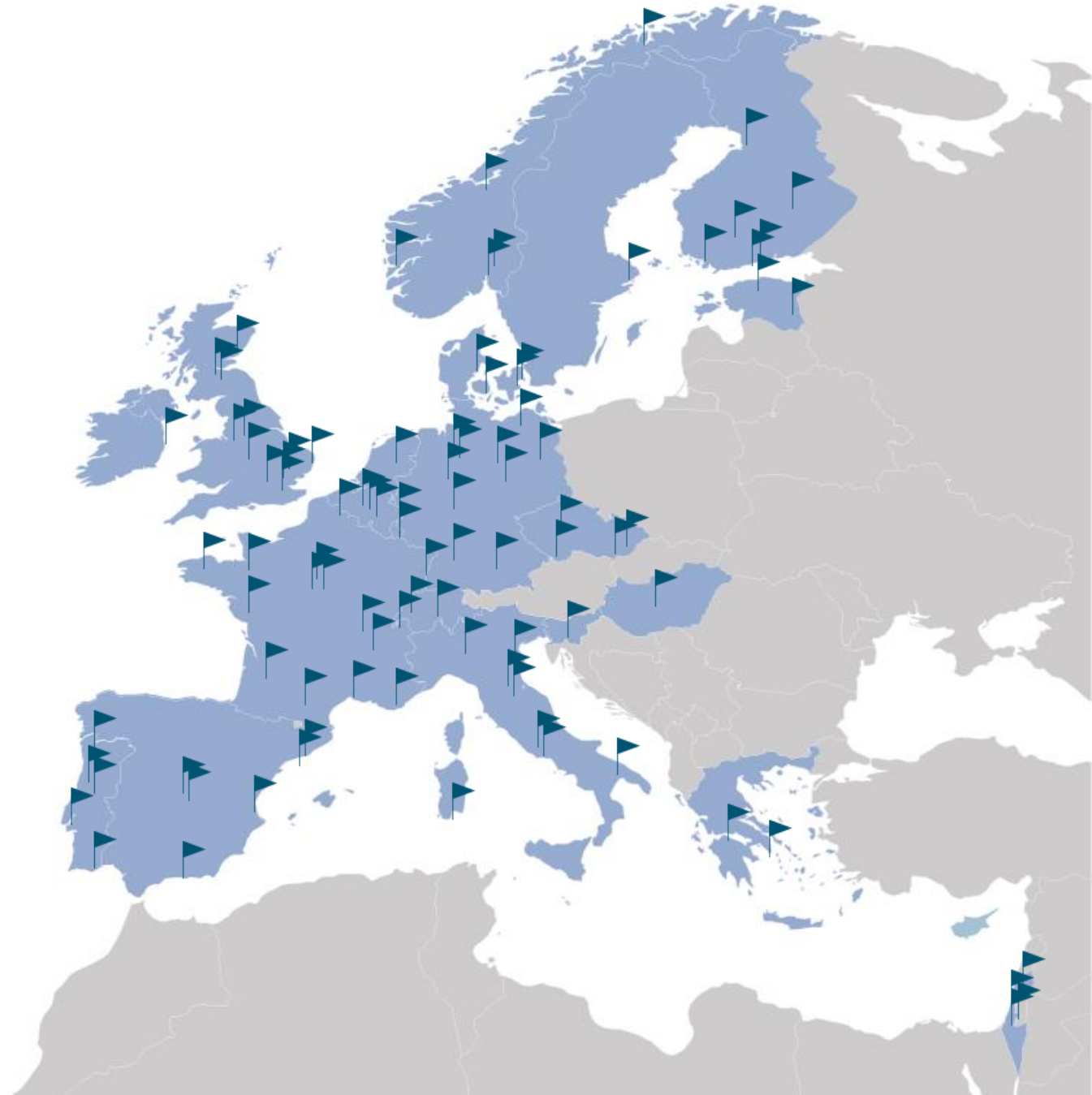


## ELIXIR Observers

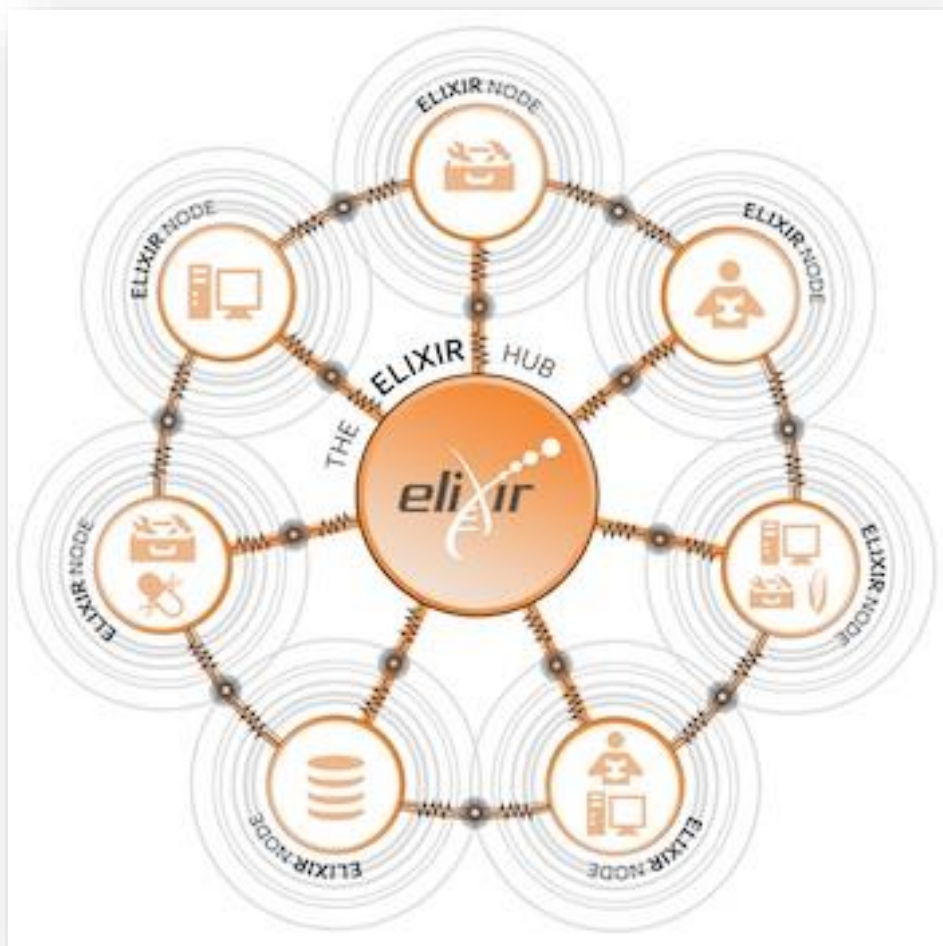


# ELIXIR in numbers

- 22 Members and 1 Observers
- ~ 180 institutes involved
- 700+ staff
- 18 Core Data Resources
- 21 Implementation Studies ongoing or soon to start
- 27 papers in ELIXIR F1000 channel
- 300 live events in TeSS
- 400 companies attended Innovation and SME programme



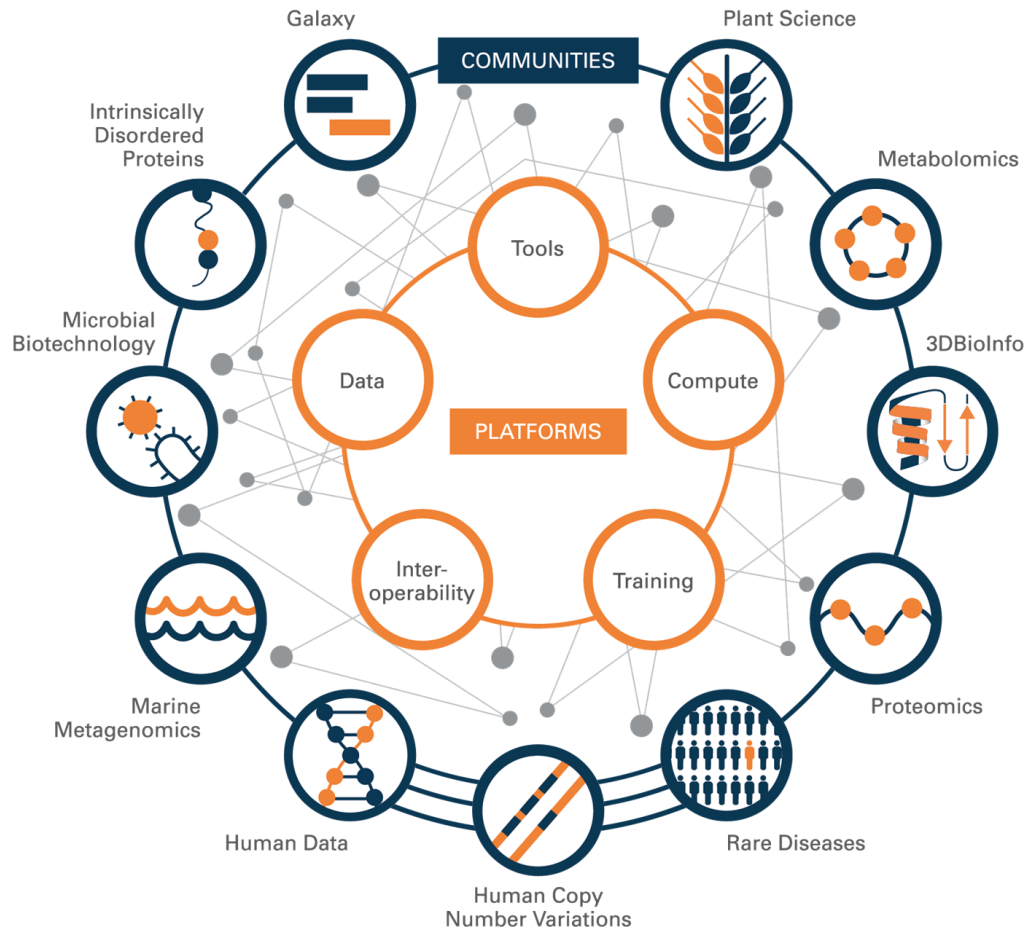
# ELIXIR distributed European infrastructure



- ELIXIR Nodes build local bioinformatics capacity throughout Europe
- ELIXIR Nodes builds on national strengths and priorities
- Nodes funded by national agencies
- ELIXIR Hub funded jointly by ELIXIR member states



# ELIXIR structure



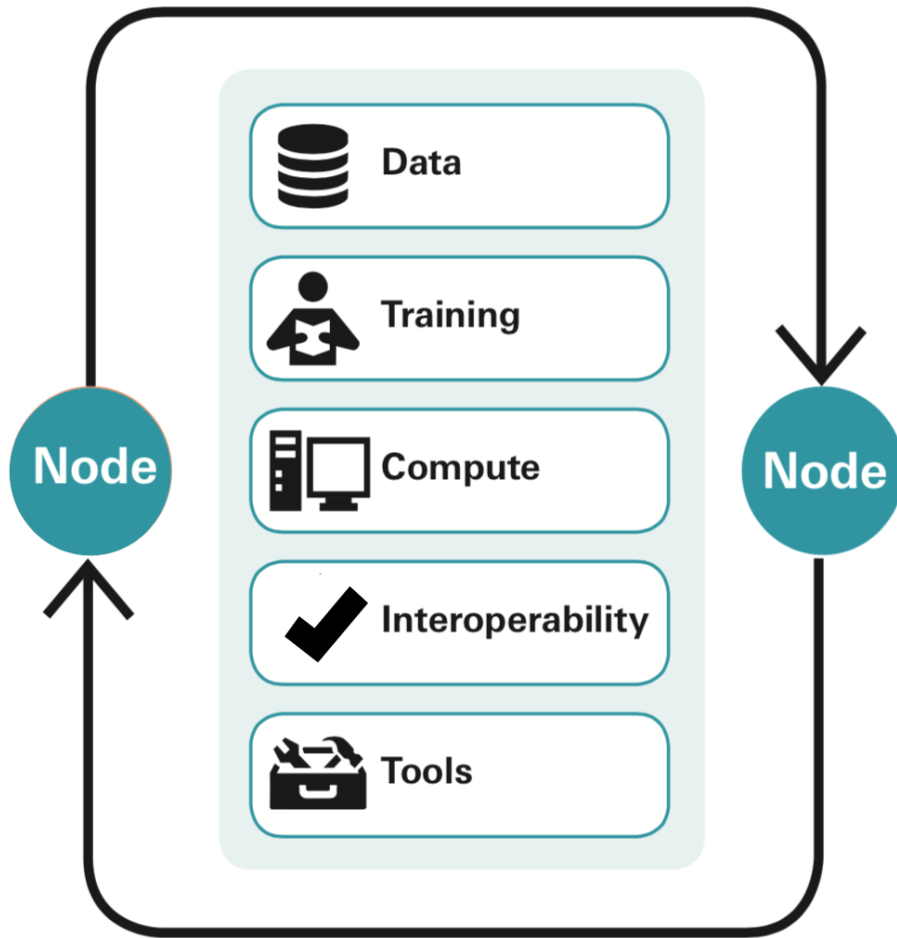
## Five technical platforms :

- Compute
- Data
- Tools
- Interoperability
- Training

## Complemented by Communities:

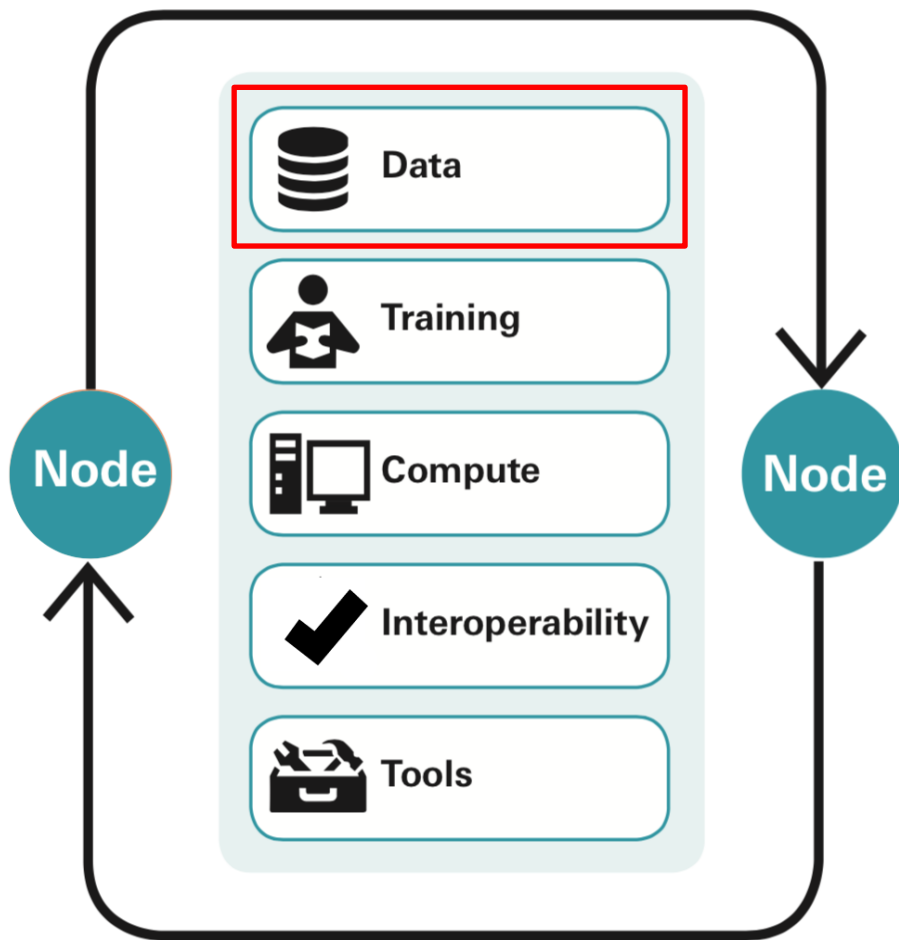
- Marine metagenomics
- Plants sciences
- Proteomics, Metabolomics
- Galaxy
- Human Data Communities (Rare diseases, Federated human data, Human copy number variation)
- Intrinsically Disordered Proteins
- Microbial biotechnology

# ELIXIR services



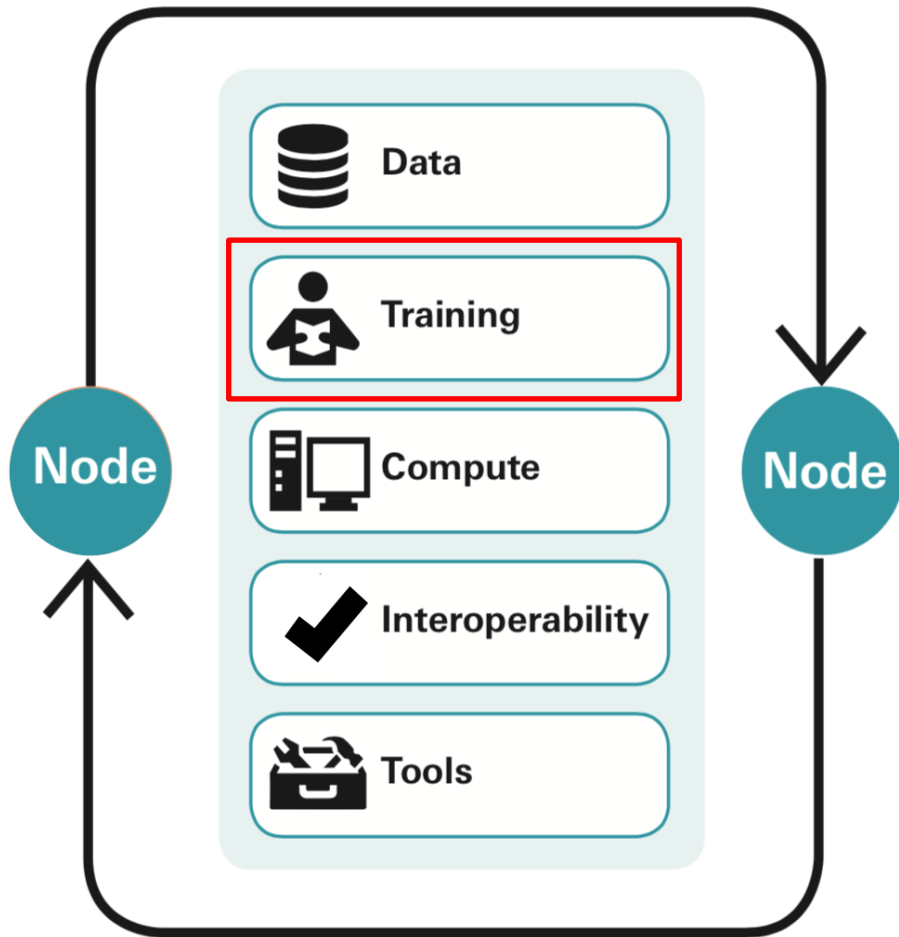
- Accessed internationally
- Operated at national level by ELIXIR Nodes
- Funded by national and international schemes
- Held together by ELIXIR's technical programme

# ELIXIR services - Data



- Aim to drive the use, re-use and value of life science data
- Provide long-term sustainable data resources
- ELIXIR Core Data Resources – eg. ENA
- ELIXIR Deposition Databases – eg. Uniprot
- Database services listing

# ELIXIR services - Training



- Aim to develop a training community
- Grow bioinformatics training capacity
- Build a sustainable training infrastructure
- TeSS - training registry for the ELIXIR community

Since  
2015

850+  
training  
events

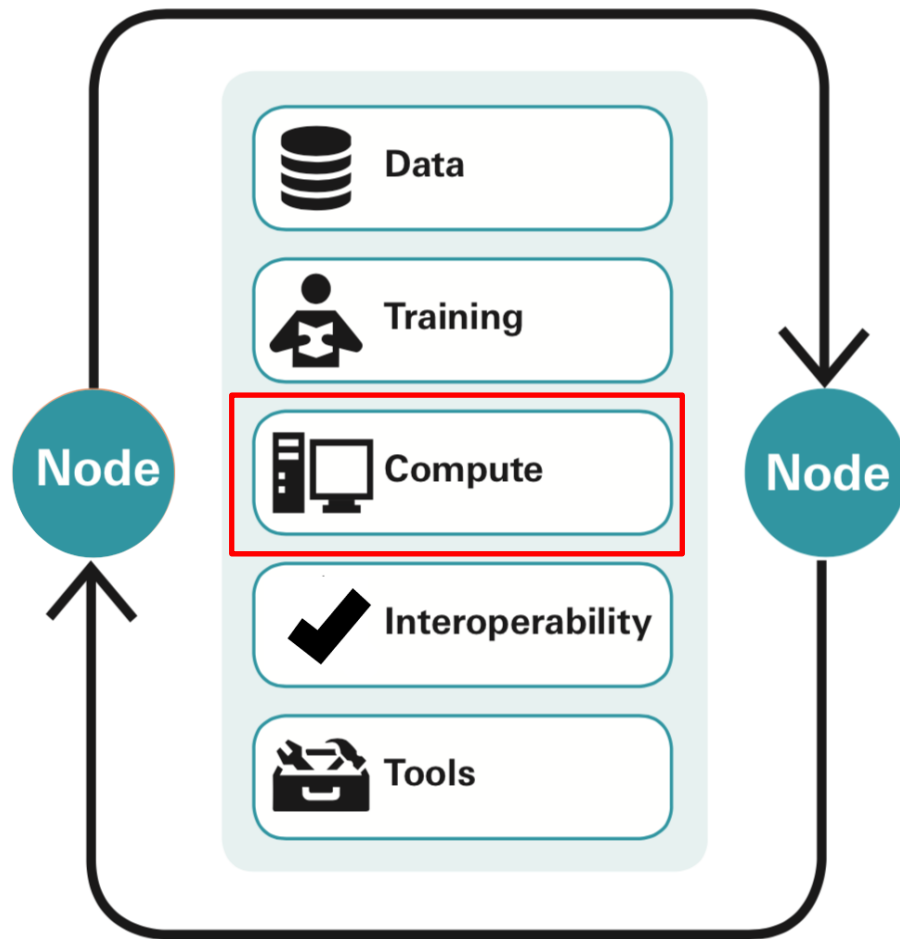
19,000+  
people trained

60+  
countries

2200+  
days of  
training

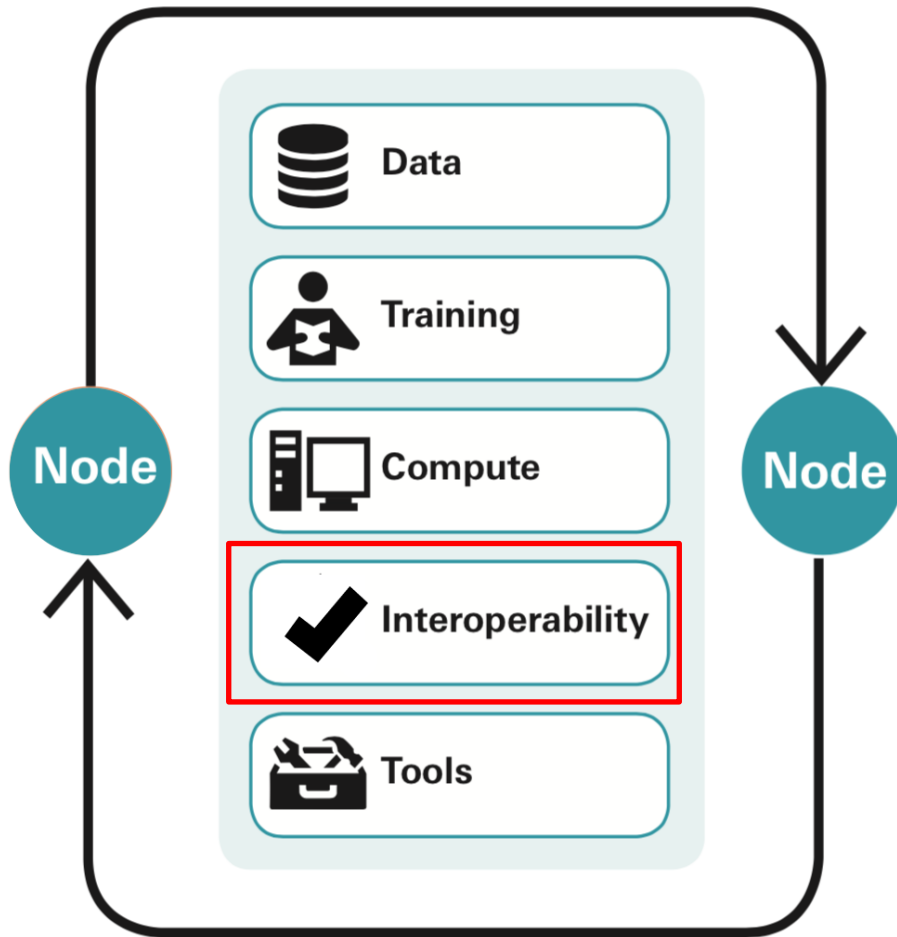


# ELIXIR services - Compute



- Aim to build and integrate cloud, compute, storage and access services for the life-science research community
- Authentication and authorization
- Storage and data transfer
- Cloud and computing resources

# ELIXIR services - Interoperability



- Aim to help people and machines to discover, access, integrate and analyse biological data
- Encourages the life science community to adopt standardised file formats, metadata, vocabularies and identifiers
- Recommend interoperability resources

Metadata  
Standards

[FAIRsharing.org](https://www.fairsharing.org/)

[isatools](https://isatools.org/)

Ontologies

[OLS](https://ontology.org/)

Identifiers

[Identifiers.org](https://identifiers.org/)

[BridgeB](https://bridgeb.org/)

Type  
Specific

[g:Profiler](https://gprofiler.org/)

[3DIONOTES-WS](https://3dionotes.org/)

[DisGeNET](https://disgenet.org/)

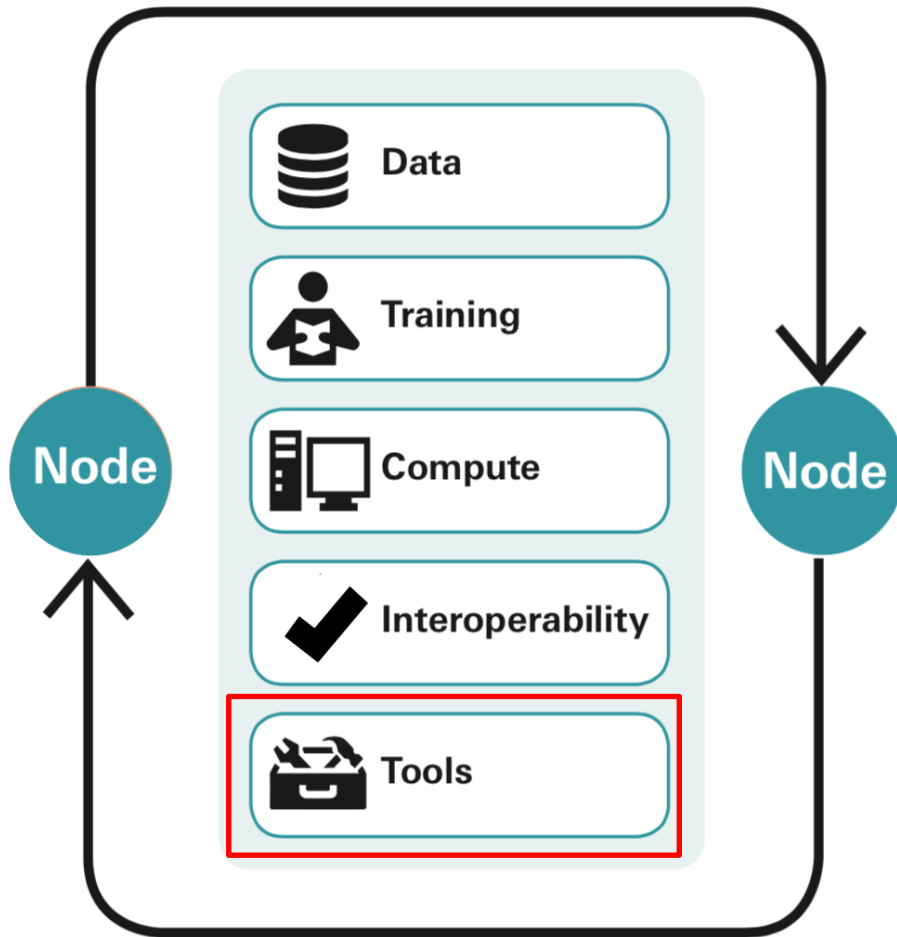
Integration

[InterMine](https://intermine.org/)

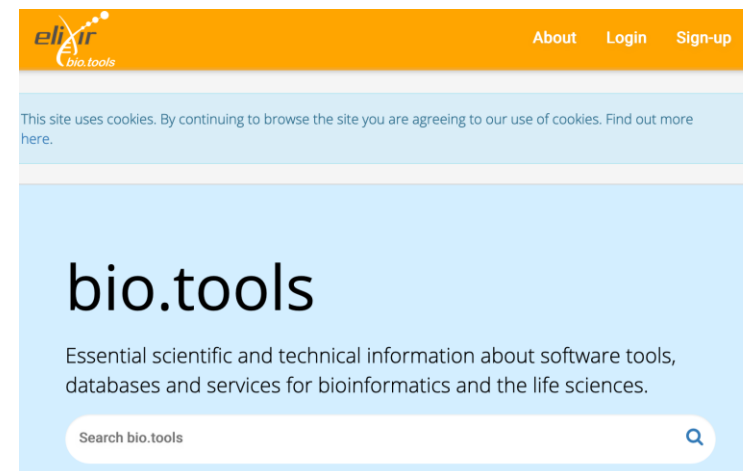
[molgenis](https://molgenis.org/)



# ELIXIR services - Tools



- Aim to improve the discovery, quality and sustainability of software resources
- Bio.tools is a comprehensive registry of software and databases
- ELIXIR Galaxy community contribute to the development of the Galaxy platform



# What ELIXIR Norway offers

- Service and collaboration
- Support
- Access to data, tools, compute & storage
- Data management planning
- Provide connection to other international ELIXIR nodes
- Training and courses
- Hands-on workshops



**contact@bioinfo.no**

