ELIXIR Belgium
Building an ecosystem of services in Life Sciences

Frederik Coppens

elixir BELGIUM

The Belgian Node : a short history

Towards a Belgian Node

- December 2013 - Memorandum of Understanding, Belgium joins ELIXIR
- November 2015 - ELIXIR Consortium Agreement, full ELIXIR Member
- December 2017 - ELIXIR Collaboration Agreement, VIB lead, 7 academic partners
    - VIB legal entity for ELIXIR Belgium
    - All Flemish & Brussels universities on board
    - 1 Walloon university
- December 2018 - Sciensano becomes partner

A people infrastructure

Universiteit Antwerpen
GHENT UNIVERSITY
VUB
KU LEUVEN
UHASSELT
ULB UNIVERSITÉ LIBRE DE BRUXELLES
sciensano
LIÈGE université
VIB

elixir BELGIUM

ELIXIR Belgium funding

- Belgian Science policy office covers the annual fee
- 2016 – funding to hire a coordinator from Flemish government ((d)HoN in kind by VIB)
  - 1 FTE
- 2017-2018 – infrastructural funding from Flemish government
  - Development of new programme
  - Bridge-the-gap-funding
  - ~5 FTE / year
- 2019-2022 – International Research Infrastructure (IRI) call by FWO
  - Dedicated call for ESFRIs
  - 2-yearly call, funding for 4 years, when funded skip a call
  - ~ 15 FTE / year
  - ~50/50 central node / partners

ELIXIR Belgium funding

- 2023-2026 – renewal of IRI funding
  - Evaluation ongoing
  - ~ 20 FTE / year requested
- 2022-2026 – Belgian Genome Biobank project
  - Lead technical infrastructure – mostly in kind
  - ¼ FTE / year

Embedded in the (inter)national ecosystem

LifeWatch
BELGIUM
BY-COVID
RDMkit
DiSSCo
Distributed System of Scientific Collections
EURO BIOIMAGING
EOSC-Life
1+MG
Galaxy PROJECT
WorkflowHub
eosc
MS DATA ALLIANCE
EMPHASIS
EBRAINS
EIRENE RI
VLAAMS SUPERCOMPUTER CENTRUM
Vlaanderen is supercomputing
Australian BioCommons
NIH National Institutes of Health
Turning Discovery Into Health
elixir BELGIUM
IBISBA

International projects

- ELIXIR-Excelerate (done)
- ELIXIR-CONVERGE
  - WP lead RDMkit
- EOSC-Life
  - Task lead tools collaboratory & co-lead WorklfowHub
- BY-COVID
  - WP lead Infectious Diseases Toolkit & (federated) data analysis
- EuroScienceGateway
  - Task leads
- Genomic Data Infrastructure
  - Task lead
  - Lead for technical infrastructure in Belgium
- AgroServ
  - Co-lead of WP on data management

Lots of different roles & mandates

- Chair of the Belgian 1+ Million Genomes mirror group ICT
- Member of the 1+ Million Genomes working group ICT
- Member of the Belgian 1+ Million Genomes Steering Committee

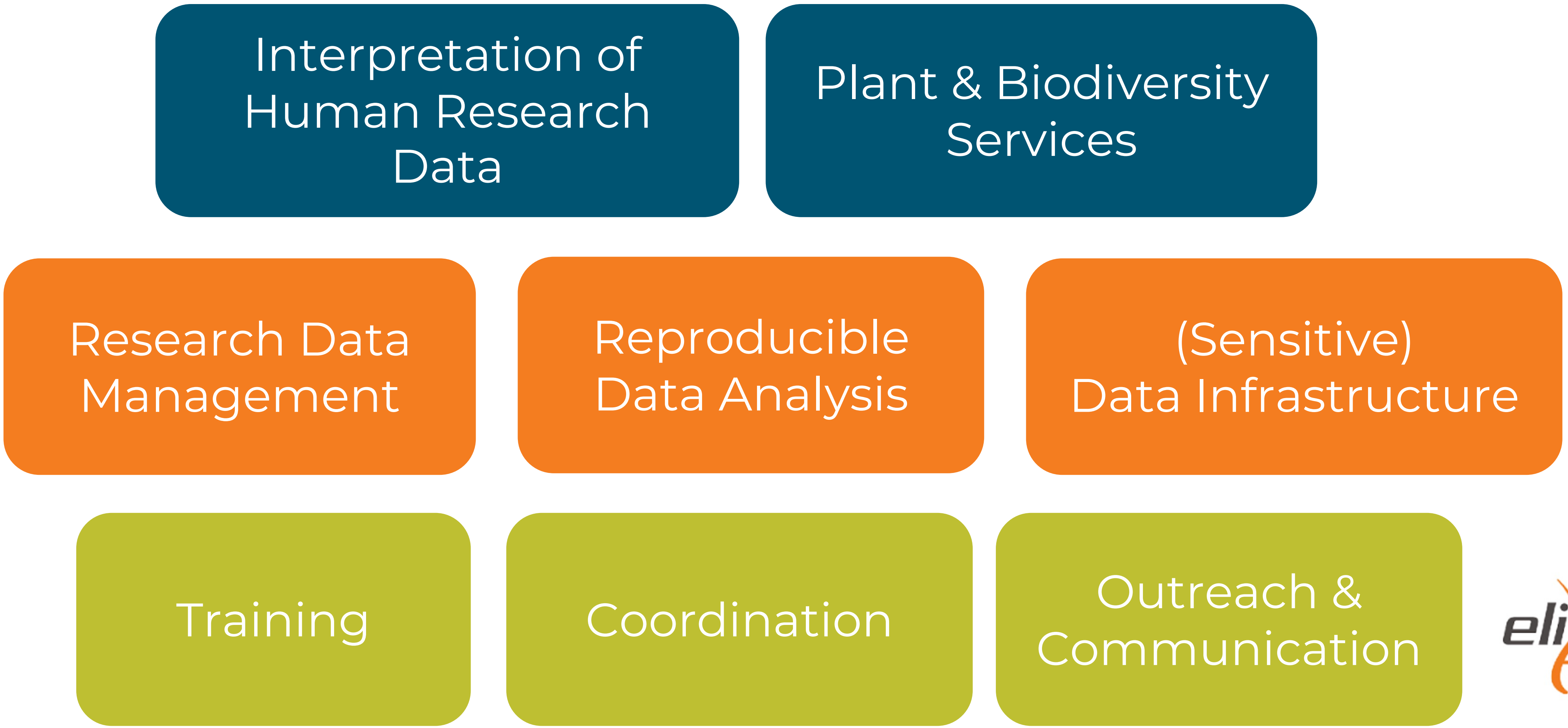
- Member of the Flemish Supercomputer Center (VSC) User Council
- Member of the Flemish Open Science Board (FOSB), representing ESFRIs

- Belgian representative in the Strategy Working Group for Data, Computing and Digital Research Infrastructures (SWG DIGIT) of the ESFRI Forum
- Member of the EOSC TF Technical interoperability of Data and Services

- Member of the Galaxy Executive Board
- Co-chair of ELIXIR Galaxy Community

Delivering Services for Flemish Researchers

ELIXIR Belgium services



Interpretation of Human Research Data

Plant & Biodiversity Services

Research Data Management

Reproducible Data Analysis

(Sensitive) Data Infrastructure

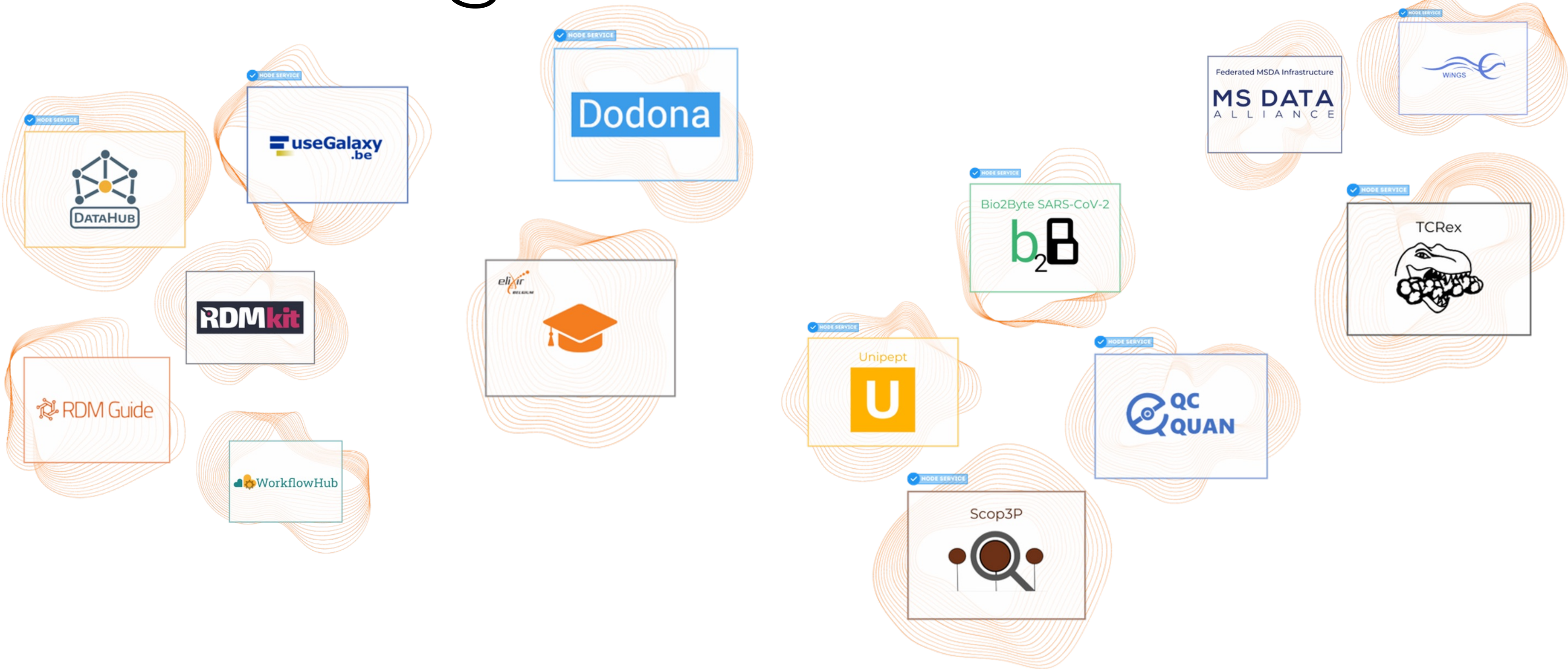
Training

Coordination

Outreach & Communication

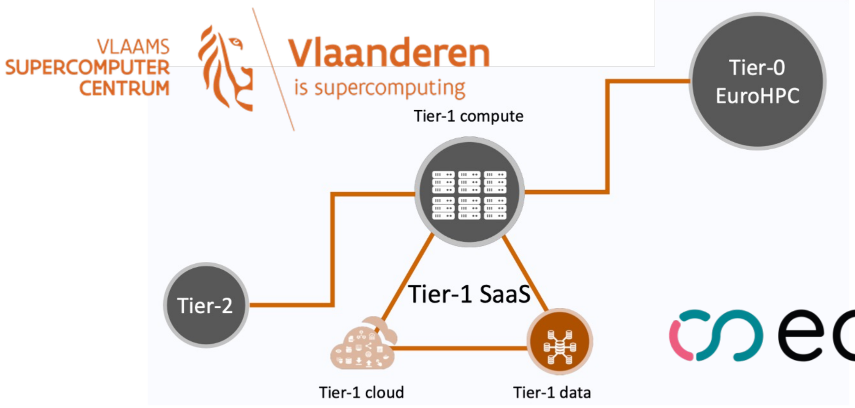
Providing free-to-use Services-on-Top

elixir BELGIUM



Services

Hardware infrastructure



VIB

eosc

UHASSELT

KU LEUVEN

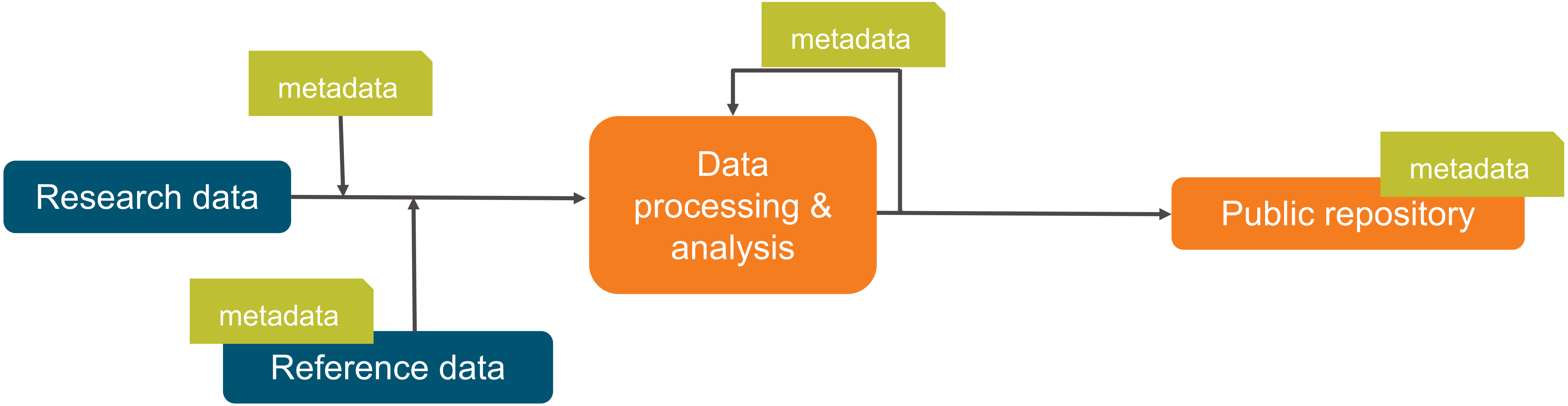
Universiteit Antwerpen

VUB

GHENT UNIVERSITY

Providing infrastructure services

Enabling FAIR data by design



Research data

metadata

Reference data

metadata

Data processing & analysis

metadata

Public repository

metadata

RDMkit

DataHub

Galaxy PROJECT

docker

nextflow

BIOCONDA

WorkflowHub

Workflow RO-Crate

Global Alliance for Genomics & Health

Collaborate. Innovate. Accelerate.

elixir BELGIUM

Scalable data management services

How can we help researchers, data stewards and project managers navigate and contribute to this FAIR data repository landscape?

RDMkit

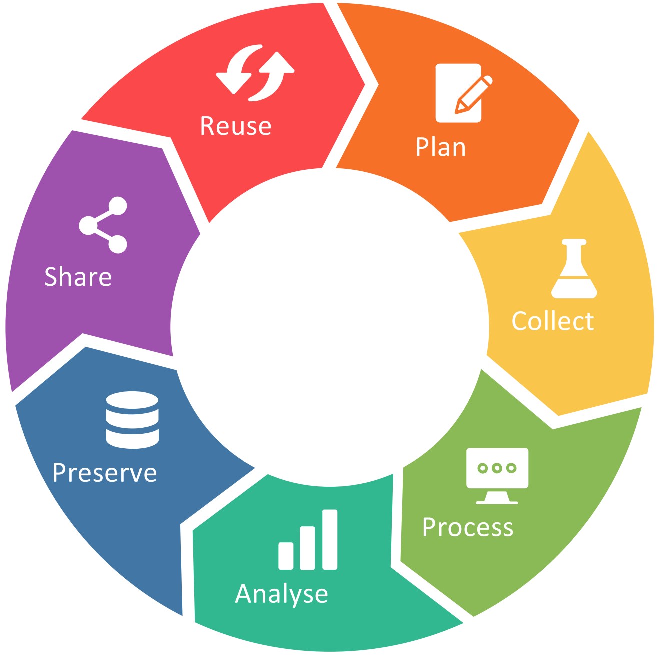
Provides guidance on the
Research Data Management support landscape

https:/rdmkit.elixir-europe.org

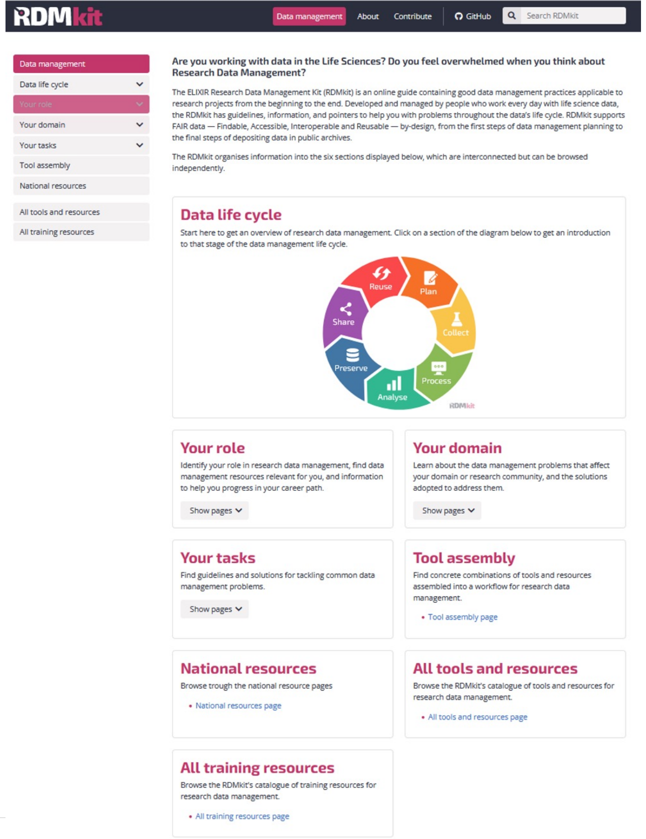
elixir
BELGIUM

RDMkit https://rdmkit.elixir-europe.org

RDM support throughout the entire **life cycle of projects** as outlined in **DMPs**



Online focal point for guidance, information, best practice, examples



Are you working with data in the Life Sciences? Do you feel overwhelmed when you think about Research Data Management?

The ELIXIR Research Data Management Kit (RDMkit) is an online guide containing good data management practices applicable to research projects from the beginning to the end. Developed and managed by people who work every day with life science data, the RDMkit has guidelines, information, and pointers to help you with problems throughout the data's life cycle. RDMkit supports FAIR data — Findable, Accessible, Interoperable and Reusable — by-design, from the first steps of data management planning to the final steps of depositing data in public archives.

The RDMkit organises information into the six sections displayed below, which are interconnected but can be browsed independently.

Data life cycle

Start here to get an overview of research data management. Click on a section of the diagram below to get an introduction to that stage of the data management life cycle.

Your role

Identify your role in research data management, find data management resources relevant for you, and information to help you progress in your career path.

Your domain

Learn about the data management problems that affect your domain or research community, and the solutions adopted to address them.

Your tasks

Find guidelines and solutions for tackling common data management problems.

Tool assembly

Find concrete combinations of tools and resources assembled into a workflow for research data management.

- Tool assembly page

National resources

Browse trough the national resource pages

- National resources page

All tools and resources

Browse the RDMkit's catalogue of tools and resources for research data management.

- All tools and resources page

All training resources

Browse the RDMkit's catalogue of training resources for research data management.

- All training resources page

Context and **signpost** for FAIR data resources as a Hub for a RDM Knowledge Commons

FAIR cookbook

DSW DATA STEWARDSHIP WIZARD

FAIRsharing.org standards, databases, policies

elixir bio.tools

elixir TeSS

elixir Core Data Resource

WorkflowHub

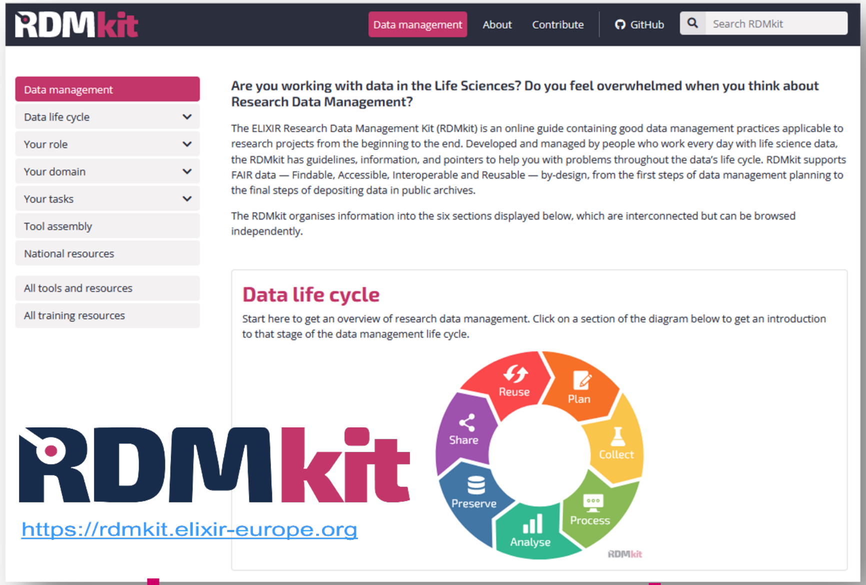
EOSC-Life

Sensitive Data Toolbox

elixir BELGIUM

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824087

A Showcase for ELIXIR's FAIR Services



Are you working with data in the Life Sciences? Do you feel overwhelmed when you think about Research Data Management?

The ELIXIR Research Data Management Kit (RDMkit) is an online guide containing good data management practices applicable to research projects from the beginning to the end. Developed and managed by people who work every day with life science data, the RDMkit has guidelines, information, and pointers to help you with problems throughout the data's life cycle. RDMkit supports FAIR data — Findable, Accessible, Interoperable and Reusable — by-design, from the first steps of data management planning to the final steps of depositing data in public archives.

The RDMkit organises information into the six sections displayed below, which are interconnected but can be browsed independently.

Data life cycle

Start here to get an overview of research data management. Click on a section of the diagram below to get an introduction to that stage of the data management life cycle.

RDMkit

https://rdmkit.elixir-europe.org

Data Management Systems for user projects, supported by Nodes

InterMine FAIRDOM iRODS MOLGENIS REDCap Research Electronic Data Capture

Public Data Repositories

elixir Core Data Resource PIPPA CATH UniProt EUROPEAN GENOME-PHENOME ARCHIVE ENA European Nucleotide Archive

Registries

elixir TeSS elixir bio.tools WorkflowHub

FAIRsharing.org standards, databases, policies

Expert Data Stewardship decision making

DSW DATA STEWARDSHIP WIZARD

EOSC-Life Sensitive Data Toolbox

Expert Data FAIRification

FAIR Cookbook

FAIR Services

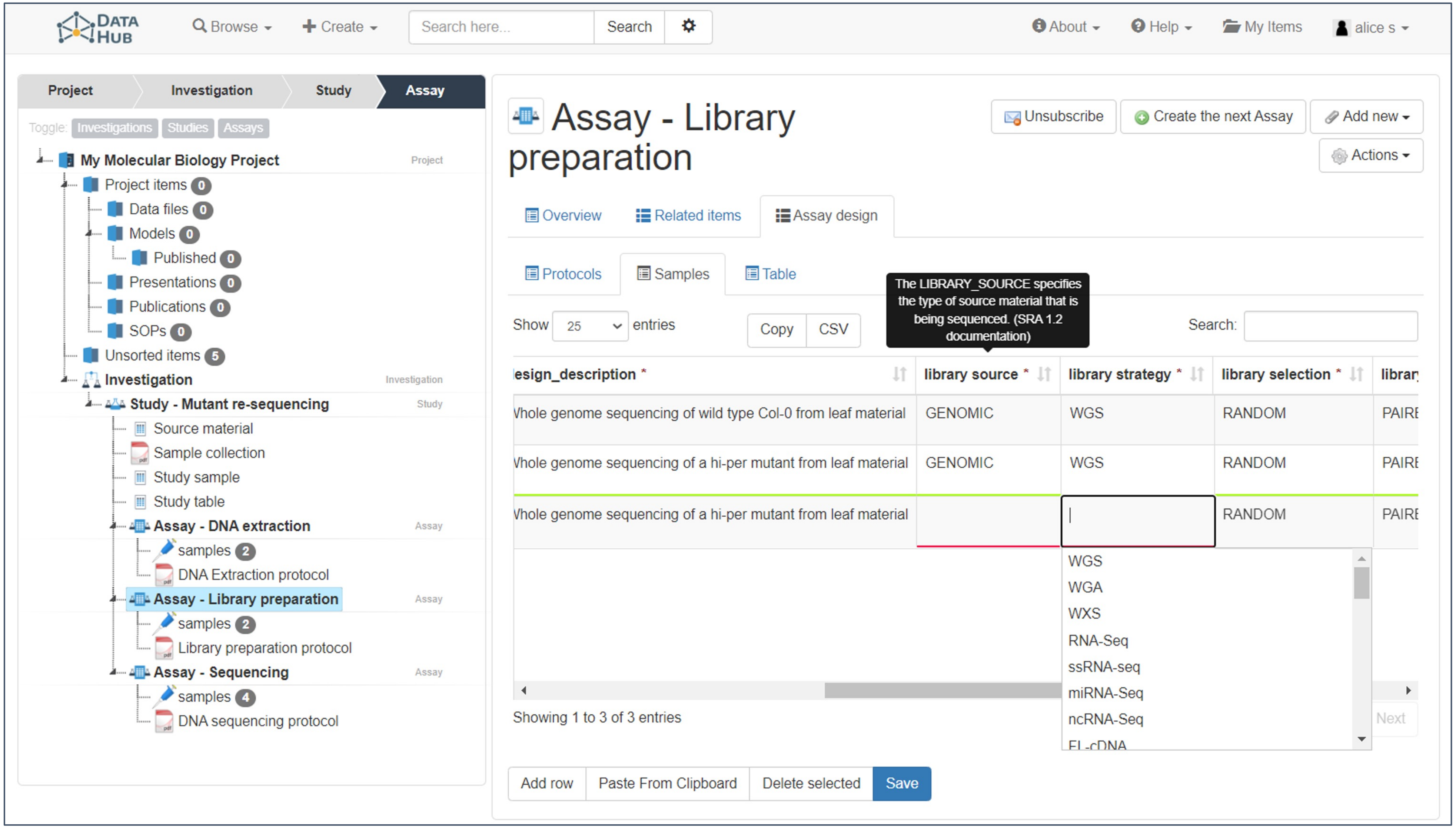
Omics DI BridgeDb OLS ONTOLOGY SEARCH elixir Beacon

Standards

Bioschemas.org Global Alliance for Genomics & Health OHDSI genomic STANDARDS consortium elixir elixir BEDIUM CONVERGE

DataHub : collection of metadata throughout the data life cycle

DATA HUB



- My Molecular Biology Project — Project
  - Project items 0
    - Data files 0
    - Models 0
      - Published 0
    - Presentations 0
    - Publications 0
    - SOPs 0
  - Unsorted items 5
  - Investigation — Investigation
    - Study - Mutant re-sequencing — Study
      - Source material
      - Sample collection
      - Study sample
      - Study table
      - Assay - DNA extraction — Assay
        - samples 2
        - DNA Extraction protocol
      - Assay - Library preparation — Assay
        - samples 2
        - Library preparation protocol
      - Assay - Sequencing — Assay
        - samples 4
        - DNA sequencing protocol

Assay - Library preparation

The LIBRARY_SOURCE specifies the type of source material that is being sequenced. (SRA 1.2 documentation)

Show 25 entries

Search:

esign_description *	library source *	library strategy *	library selection *	librar
Vhole genome sequencing of wild type Col-0 from leaf material	GENOMIC	WGS	RANDOM	PAIRE
Vhole genome sequencing of a hi-per mutant from leaf material	GENOMIC	WGS	RANDOM	PAIRE
Vhole genome sequencing of a hi-per mutant from leaf material			RANDOM	PAIRE

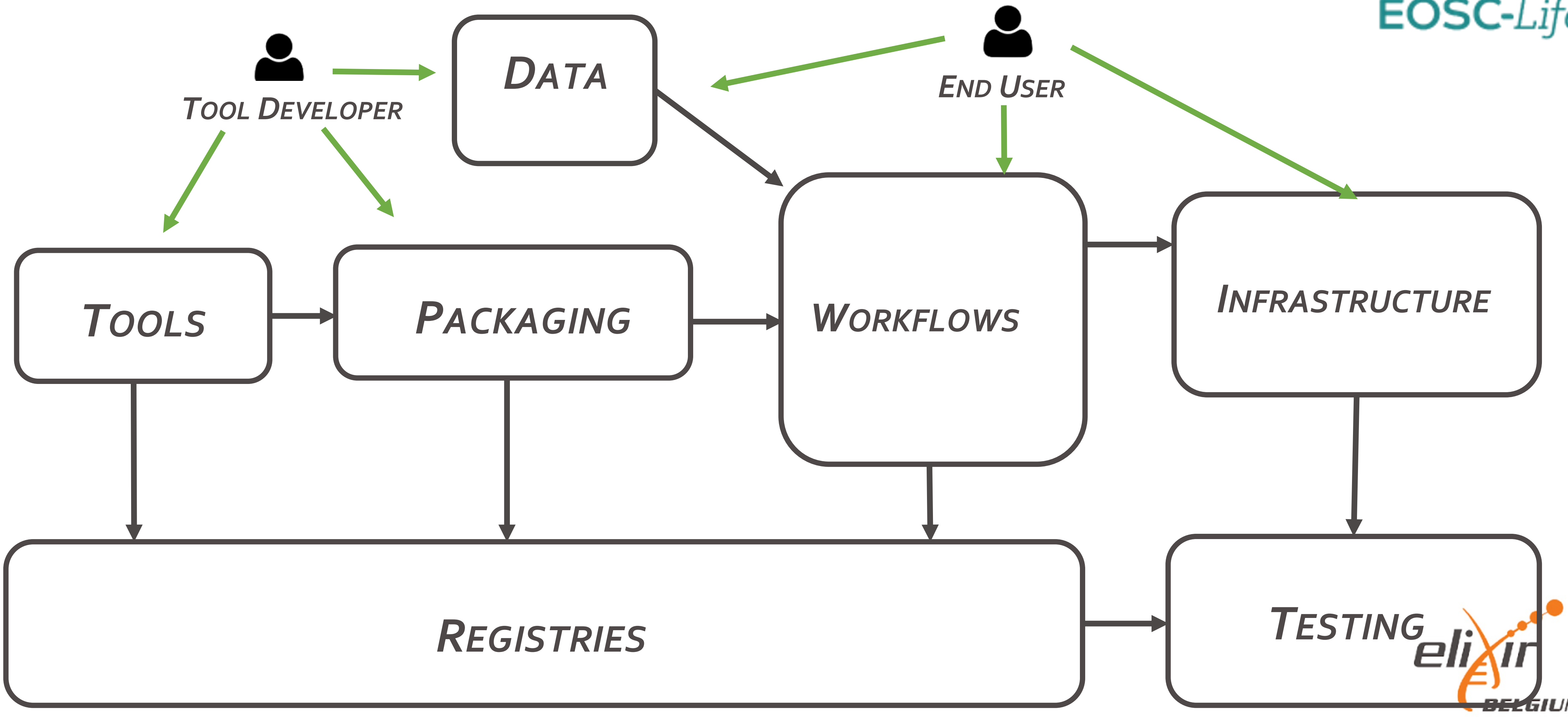
- WGS
- WGA
- WXS
- RNA-Seq
- ssRNA-seq
- miRNA-Seq
- ncRNA-Seq
- FL-cDNA

Showing 1 to 3 of 3 entries

elixir BELGIUM

Reproducible analysis services

EOSC-Life Tools Collaboratory Roadmap



Tool Developer

Data

End User

Tools

Packaging

Workflows

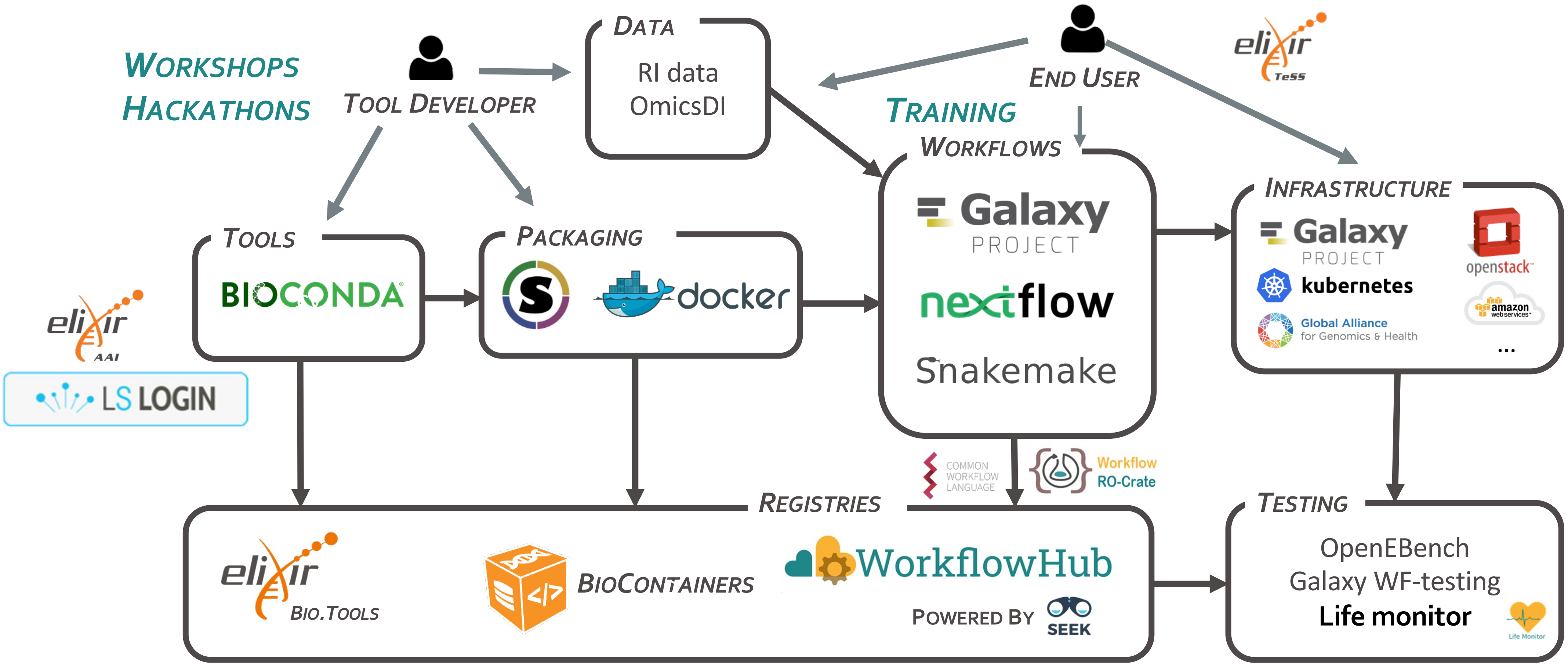
Infrastructure

Registries

Testing

EOSC-Life Tools Collaboratory Roadmap

EOSC-Life



Workshops
Hackathons

Tool Developer

Data
RI data
OmicsDI

End User

Training

Workflows
Galaxy PROJECT
nextflow
Snakemake

Infrastructure
Galaxy PROJECT
kubernetes
Global Alliance for Genomics & Health
openstack
amazon webservices
...

elixir TeSS

Tools
BIOCONDA

Packaging
docker

elixir AAI
LS LOGIN

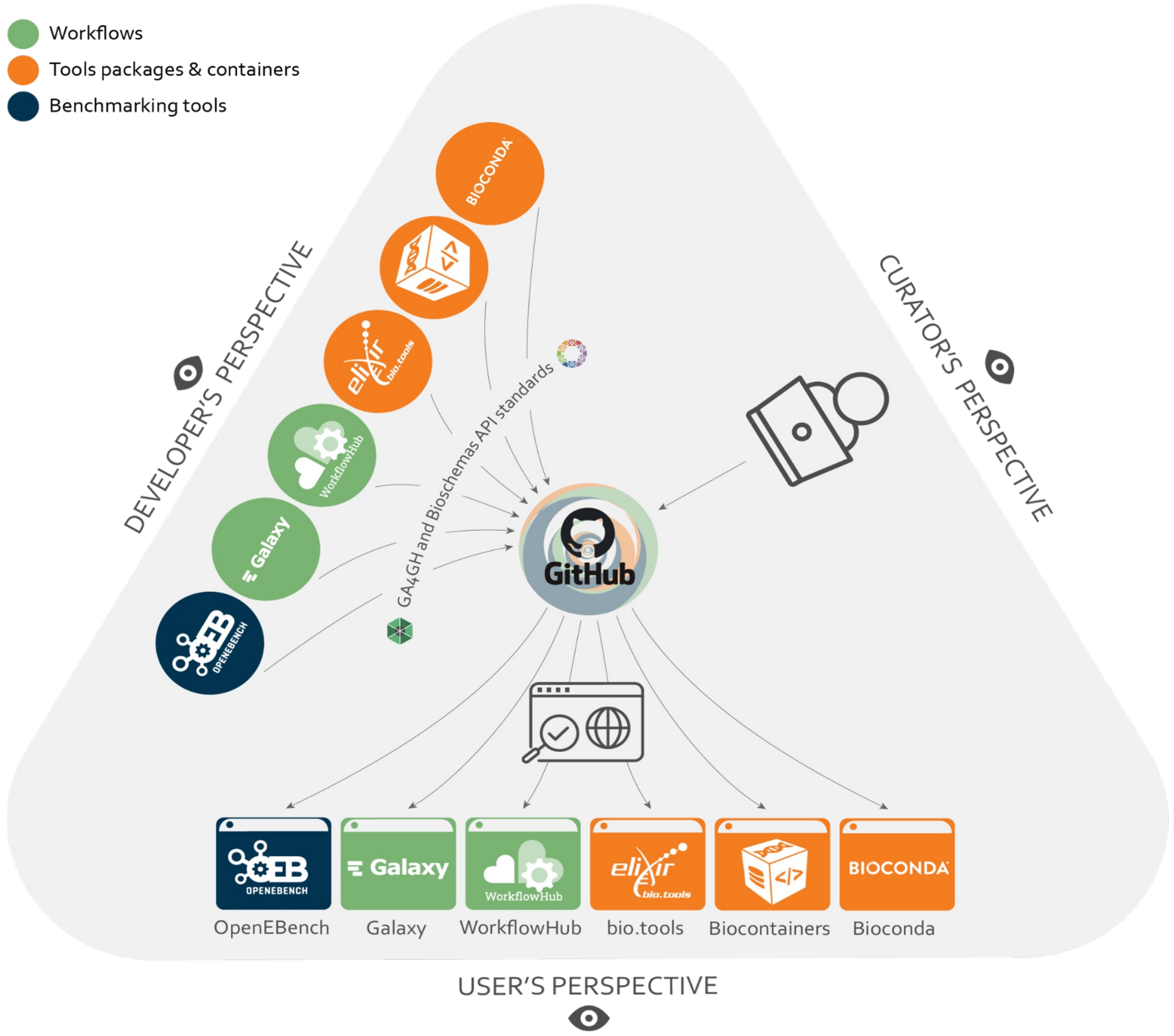
COMMON WORKFLOW LANGUAGE
Workflow RO-Crate

Registries
elixir Bio.Tools
BioContainers
WorkflowHub
Powered by SEEK

Testing
OpenEBench
Galaxy WF-testing
Life monitor
Life Monitor

elixir BELGIUM

The Tools Platform ecosystem



ELIXIR BELGIUM

Workflows as entry point for end-users

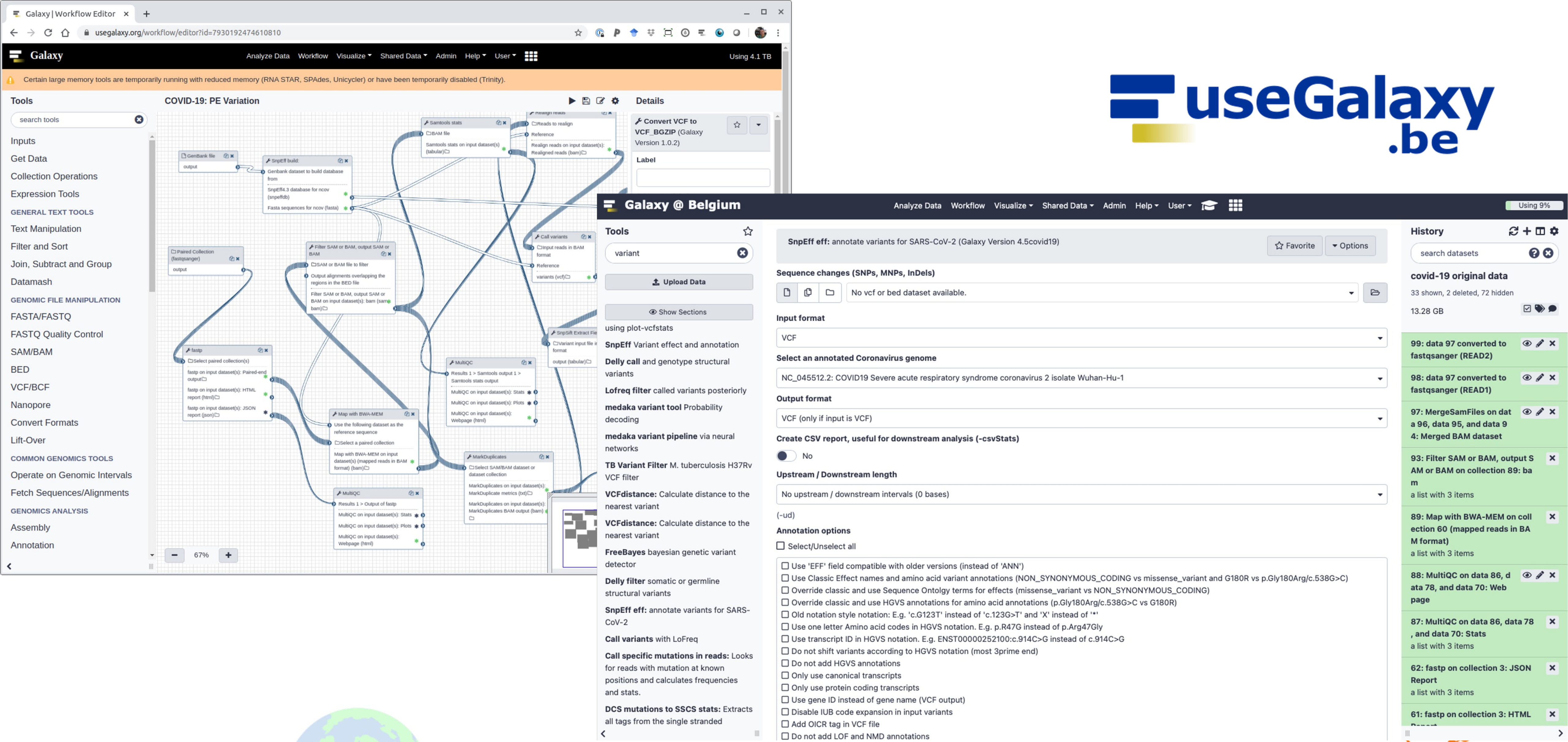
eosc

elixir
BELGIUM

Galaxy
PROJECT

Galaxy is an **open**, web-based virtual research environment for **accessible**, **reproducible**, and **transparent** computational biomedical research

elixir
BELGIUM



Galaxy | Workflow Editor

usegalaxy.org/workflow/editor?id=7930192474610810

Galaxy
Analyze Data
Workflow
Visualize
Shared Data
Admin
Help
User
Using 4.1 TB

Certain large memory tools are temporarily running with reduced memory (RNA STAR, SPAdes, Unicycler) or have been temporarily disabled (Trinity).

Tools
search tools
Inputs
Get Data
Collection Operations
Expression Tools
GENERAL TEXT TOOLS
Text Manipulation
Filter and Sort
Join, Subtract and Group
Datamash
GENOMIC FILE MANIPULATION
FASTA/FASTQ
FASTQ Quality Control
SAM/BAM
BED
VCF/BCF
Nanopore
Convert Formats
Lift-Over
COMMON GENOMICS TOOLS
Operate on Genomic Intervals
Fetch Sequences/Alignments
GENOMICS ANALYSIS
Assembly
Annotation

COVID-19: PE Variation

GenBank file
output
SnpEff build:
Genbank dataset to build database from
SnpEff4.3 database for ncov (snpeffdb)
Fasta sequences for ncov (fasta)
Samtools stats
BAM file
Samtools stats on input dataset(s) (tabular)
Realign reads
Reads to realign
Reference
Realign reads on input dataset(s): Realigned reads (bam)
Paired Collection (fastqsanger)
output
Filter SAM or BAM, output SAM or BAM
SAM or BAM file to filter
Output alignments overlapping the regions in the BED file
Filter SAM or BAM, output SAM or BAM on input dataset(s): bam (sam bam)
Call variants
Input reads in BAM format
Reference
variants (vcf)
SnpSift Extract Fie
Variant input file in format
output (tabular)
fastp
Select paired collection(s)
fastp on input dataset(s): Paired-end output
fastp on input dataset(s): HTML report (html)
fastp on input dataset(s): JSON report (json)
MultiQC
Results 1 > Samtools output 1 > Samtools stats output
MultiQC on input dataset(s): Stats
MultiQC on input dataset(s): Plots
MultiQC on input dataset(s): Webpage (html)
Map with BWA-MEM
Use the following dataset as the reference sequence
Select a paired collection
Map with BWA-MEM on input dataset(s) (mapped reads in BAM format) (bam)
MarkDuplicates
Select SAM/BAM dataset or dataset collection
MarkDuplicates on input dataset(s): MarkDuplicate metrics (txt)
MarkDuplicates on input dataset(s): MarkDuplicates BAM output (bam)
MultiQC
Results 1 > Output of fastp
MultiQC on input dataset(s): Stats
MultiQC on input dataset(s): Plots
MultiQC on input dataset(s): Webpage (html)
67%

Details

Convert VCF to VCF_BGZIP (Galaxy Version 1.0.2)

Label

Galaxy @ Belgium
Analyze Data
Workflow
Visualize
Shared Data
Admin
Help
User
Using 9%

Tools

variant

Upload Data
Show Sections

using plot-vcfstats

SnpEff Variant effect and annotation

Delly call and genotype structural variants

Lofreq filter called variants posteriorly

medaka variant tool Probability decoding

medaka variant pipeline via neural networks

TB Variant Filter M. tuberculosis H37Rv VCF filter

VCFdistance: Calculate distance to the nearest variant

VCFdistance: Calculate distance to the nearest variant

FreeBayes bayesian genetic variant detector

Delly filter somatic or germline structural variants

SnpEff eff: annotate variants for SARS-CoV-2

Call variants with LoFreq

Call specific mutations in reads: Looks for reads with mutations at known positions and calculates frequencies and stats.

DCS mutations to SSCS stats: Extracts all tags from the single stranded

SnpEff eff: annotate variants for SARS-CoV-2 (Galaxy Version 4.5covid19)

Sequence changes (SNPs, MNPs, InDels)

No vcf or bed dataset available.

Input format

VCF

Select an annotated Coronavirus genome

NC_045512.2: COVID19 Severe acute respiratory syndrome coronavirus 2 isolate Wuhan-Hu-1

Output format

VCF (only if input is VCF)

Create CSV report, useful for downstream analysis (-csvStats)

No

Upstream / Downstream length

No upstream / downstream intervals (0 bases)

(-ud)

Annotation options

☐ Select/Unselect all

☐ Use 'EFF' field compatible with older versions (instead of 'ANN')
☐ Use Classic Effect names and amino acid variant annotations (NON_SYNONYMOUS_CODING vs missense_variant and G180R vs p.Gly180Arg/c.538G>C)
☐ Override classic and use Sequence Ontolgy terms for effects (missense_variant vs NON_SYNONYMOUS_CODING)
☐ Override classic and use HGVS annotations for amino acid annotations (p.Gly180Arg/c.538G>C vs G180R)
☐ Old notation style notation: E.g. 'c.G123T' instead of 'c.123G>T' and 'X' instead of '*'
☐ Use one letter Amino acid codes in HGVS notation. E.g. p.R47G instead of p.Arg47Gly
☐ Use transcript ID in HGVS notation. E.g. ENST00000252100:c.914C>G instead of c.914C>G
☐ Do not shift variants according to HGVS notation (most 3prime end)
☐ Do not add HGVS annotations
☐ Only use canonical transcripts
☐ Only use protein coding transcripts
☐ Use gene ID instead of gene name (VCF output)
☐ Disable IUB code expansion in input variants
☐ Add OICR tag in VCF file
☐ Do not add LOF and NMD annotations

History

search datasets

covid-19 original data

33 shown, 2 deleted, 72 hidden

13.28 GB

99: data 97 converted to fastqsanger (READ2)

98: data 97 converted to fastqsanger (READ1)

97: MergeSamFiles on data 96, data 95, and data 94: Merged BAM dataset

93: Filter SAM or BAM, output SAM or BAM on collection 89: bam
a list with 3 items

89: Map with BWA-MEM on collection 60 (mapped reads in BAM format)
a list with 3 items

88: MultiQC on data 86, data 78, and data 70: Webpage

87: MultiQC on data 86, data 78, and data 70: Stats
a list with 3 items

62: fastp on collection 3: JSON Report
a list with 3 items

61: fastp on collection 3: HTML

useGalaxy.be

useGalaxy.*

VLAAMS SUPERCOMPUTER CENTRUM

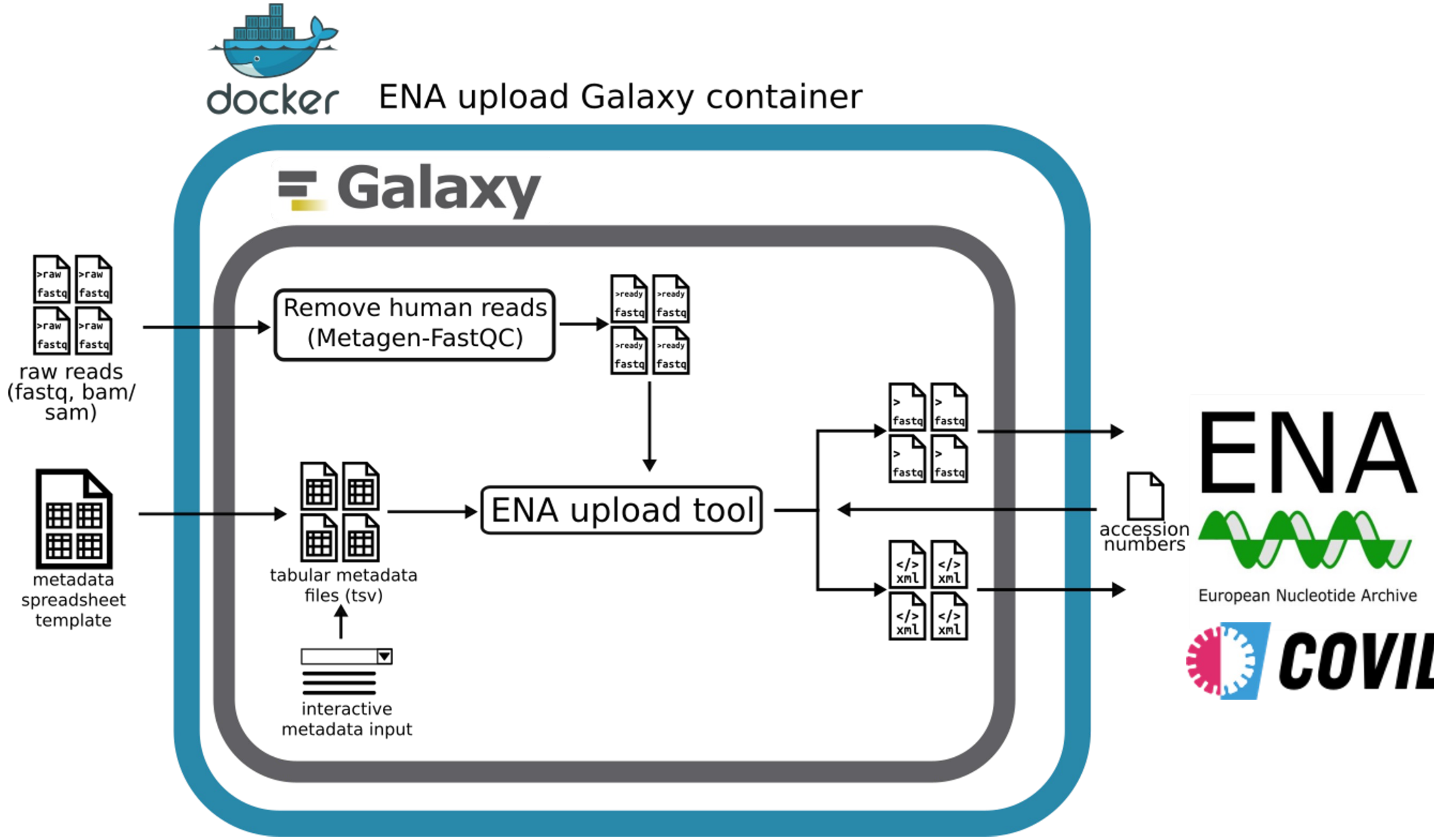
Vlaanderen is supercomputing

elixir BELGIUM

Applying infrastructure services a pandemic use case

Submission of viral data

Submission overview



ENA upload Galaxy container
Galaxy
raw reads (fastq, bam/ sam)
Remove human reads (Metagen-FastQC)
metadata spreadsheet template
tabular metadata files (tsv)
interactive metadata input
ENA upload tool
accession numbers
ENA
European Nucleotide Archive

Raw reads
Consensus sequence
Standardized metadata

COVID-19 Data Portal

elixir BELGIUM

Different ways of usage

Deploy the container

Laptop
Cloud infrastructure

Brokering

covid19.usegalaxy.be

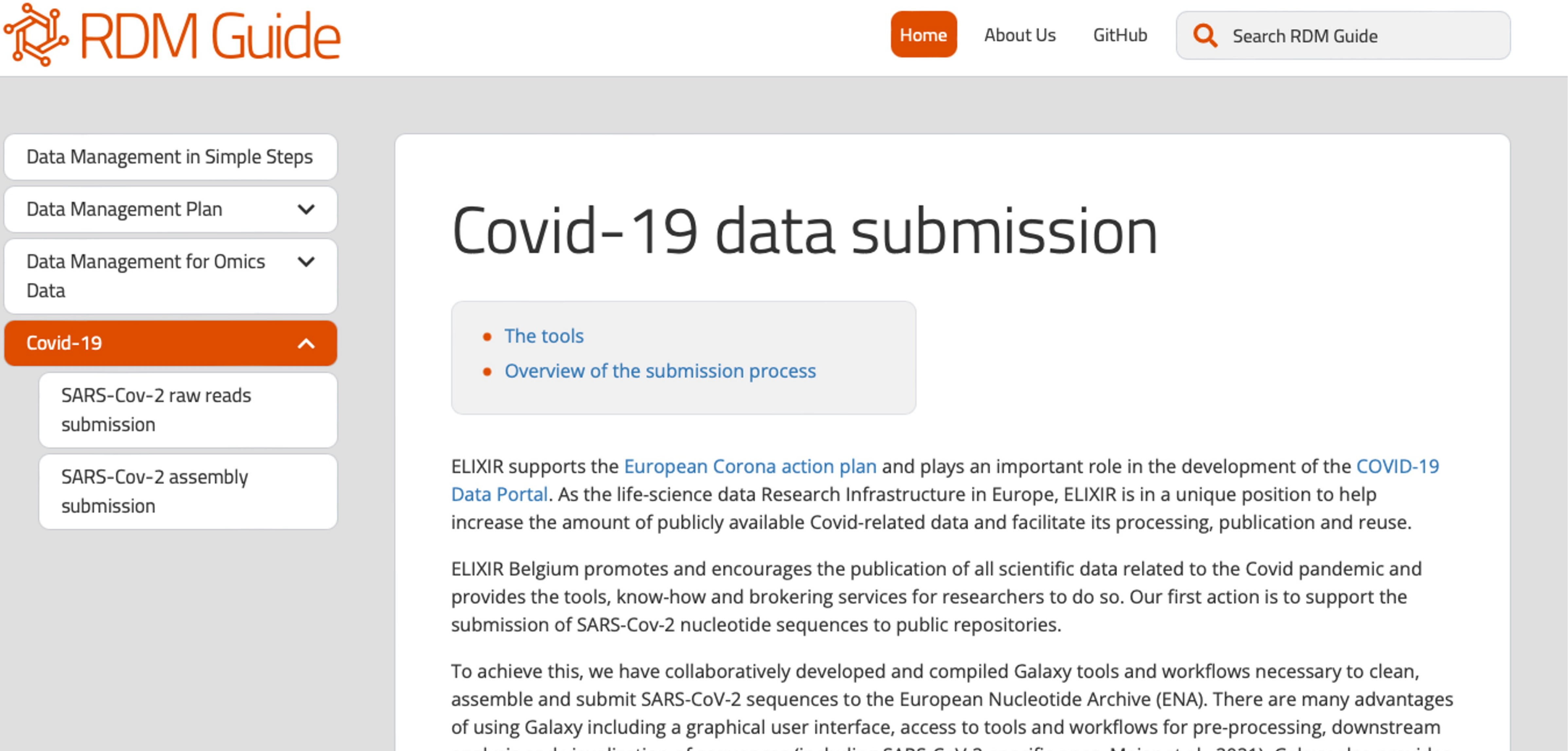
Public Galaxy instance

useGalaxy.eu
useGalaxy.be

```
docker run -p "8080:80" --privileged quay.io/galaxy/ena-upload
```

github.com/ELIXIR-Belgium/ena-upload-container

rdm.elixir-belgium.org


RDM Guide
Home
About Us
GitHub
Search RDM Guide
Data Management in Simple Steps
Data Management Plan
Data Management for Omics Data
Covid-19
SARS-Cov-2 raw reads submission
SARS-Cov-2 assembly submission

Covid-19 data submission

- The tools
- Overview of the submission process

ELIXIR supports the European Corona action plan and plays an important role in the development of the COVID-19 Data Portal. As the life-science data Research Infrastructure in Europe, ELIXIR is in a unique position to help increase the amount of publicly available Covid-related data and facilitate its processing, publication and reuse.

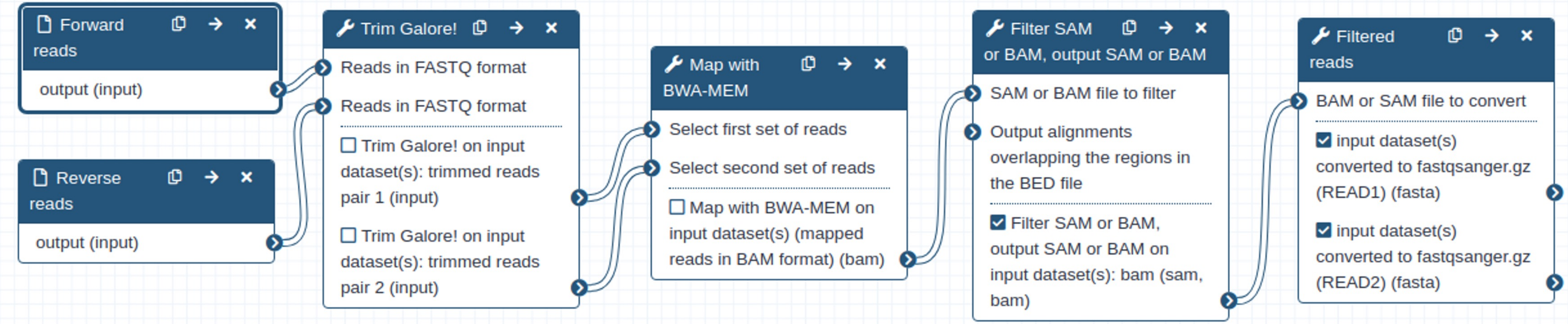
ELIXIR Belgium promotes and encourages the publication of all scientific data related to the Covid pandemic and provides the tools, know-how and brokering services for researchers to do so. Our first action is to support the submission of SARS-Cov-2 nucleotide sequences to public repositories.

To achieve this, we have collaboratively developed and compiled Galaxy tools and workflows necessary to clean, assemble and submit SARS-CoV-2 sequences to the European Nucleotide Archive (ENA). There are many advantages of using Galaxy including a graphical user interface, access to tools and workflows for pre-processing, downstream

Enabling data analysis

Removing Human reads

- Standard preprocessing step in different pipelines.
  - https://covid19.galaxyproject.org/genomics/1-preprocessing/
- Reached out to ENA for a standard method
  - Repurposed https://github.com/Finn-Lab/Metagen-FastQC



Forward reads
output (input)

Reverse reads
output (input)

Trim Galore!
Reads in FASTQ format
Reads in FASTQ format
☐ Trim Galore! on input dataset(s): trimmed reads pair 1 (input)
☐ Trim Galore! on input dataset(s): trimmed reads pair 2 (input)

Map with BWA-MEM
Select first set of reads
Select second set of reads
☐ Map with BWA-MEM on input dataset(s) (mapped reads in BAM format) (bam)

Filter SAM or BAM, output SAM or BAM
SAM or BAM file to filter
Output alignments overlapping the regions in the BED file
☑ Filter SAM or BAM, output SAM or BAM on input dataset(s): bam (sam, bam)

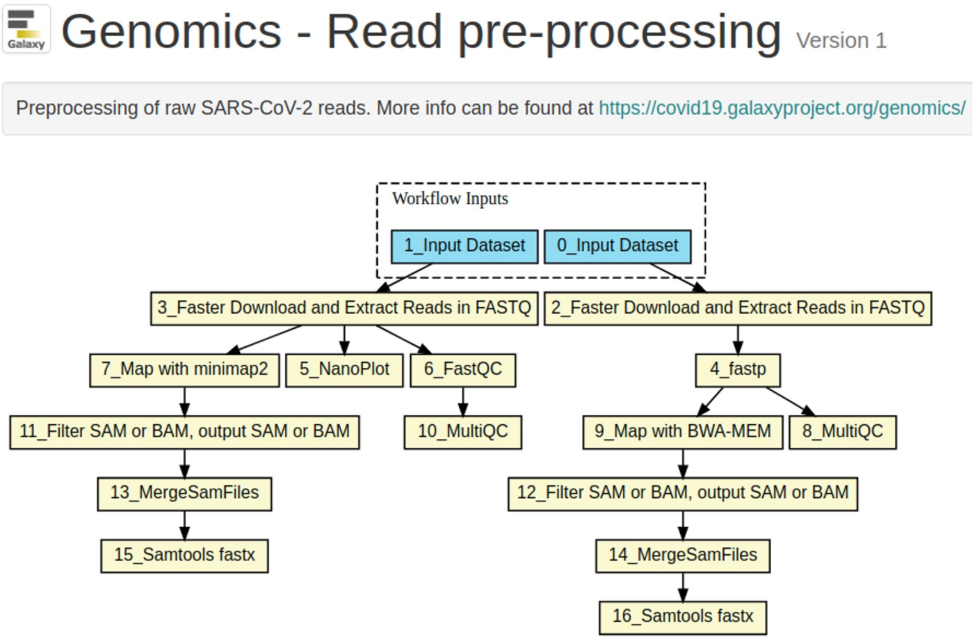
Filtered reads
BAM or SAM file to convert
☑ input dataset(s) converted to fastqsanger.gz (READ1) (fasta)
☑ input dataset(s) converted to fastqsanger.gz (READ2) (fasta)

Workflows

WorkflowHub

- Different approaches
  - Removing of reads mapping to human genome (e.g. Metagen-FastQC)
  - Retain reads mapping to viral genome
  - Combine human & viral genome, retain viral reads

- Galaxy allows using other workflows as you please
  - https://workflowhub.eu/
  - https://workflowhub.eu/workflows/99 (Metagen-FastQC)



Genomics - Read pre-processing Version 1

Preprocessing of raw SARS-CoV-2 reads. More info can be found at https://covid19.galaxyproject.org/genomics/

WorkflowHub : a FAIR workflow registry

- Workflow management system agnostic
- Registry & repository functionality
- Workflows may remain in their native repositories in their native form
- Open to workflows from all disciplines and any country
- Based on community standards
- Attribution and credit as a central feature
- Added value services: curation, monitoring, ...
- WorkflowHub Club open community
- Perpetual Development in the open
    - Registering on behalf of makers
    - Regular releases of new features
    - Agile revisions of features

WorkflowHub

https://workflowhub.eu

Integration & development of standards

COMMON WORKFLOW LANGUAGE

Canonical workflow description

Native or Abstract CWL

Bioschemas

Metadata for registration and discovery

ComputationalWorkflow and FormalParameter
ComputationalTool
Schema.org profile and types

Global Alliance for Genomics & Health

Collaborate. Innovate. Accelerate.

API Exchanging Tools and Workflows

GA4GH Tool Registry Service (TRS) API

Workflow RO-Crate

Specialisation of RO-Crate to package an executable workflow with all necessary documentation.

Exchange format for Workflow Hub.

elixir BELGIUM

WorkflowHub - usegalaxy.eu integration

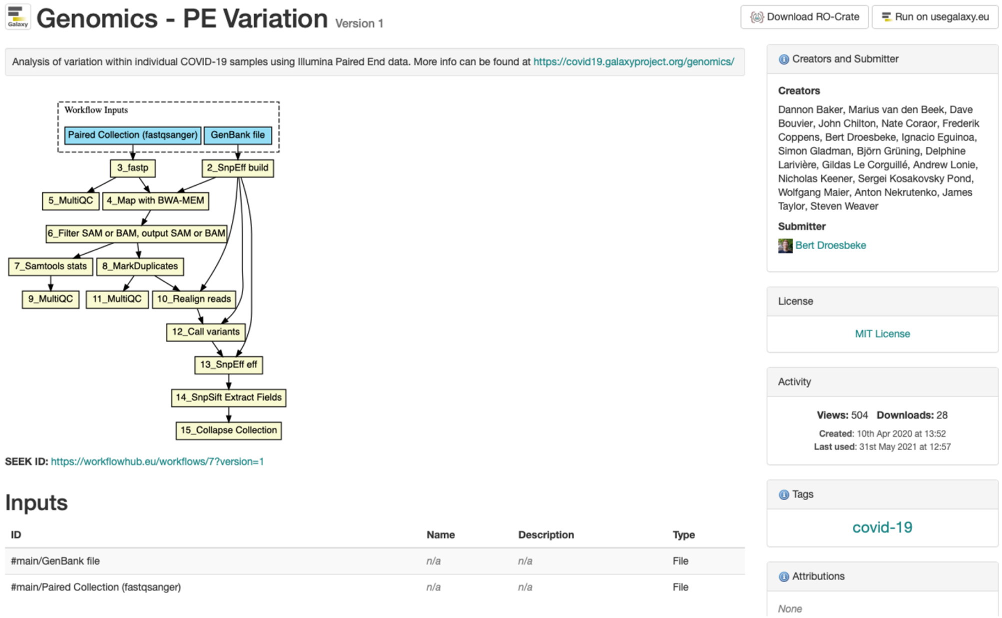
Galaxy PROJECT

COMMON WORKFLOW LANGUAGE

Workflow RO-Crate

Global Alliance for Genomics & Health
Collaborate. Innovate. Accelerate.

TRS



Genomics - PE Variation Version 1

Download RO-Crate | Run on usegalaxy.eu

Analysis of variation within individual COVID-19 samples using Illumina Paired End data. More info can be found at https://covid19.galaxyproject.org/genomics/

SEEK ID: https://workflowhub.eu/workflows/7?version=1

Inputs

ID	Name	Description	Type
#main/GenBank file	*n/a*	*n/a*	File
#main/Paired Collection (fastqsanger)	*n/a*	*n/a*	File

Creators and Submitter

Creators

Dannon Baker, Marius van den Beek, Dave Bouvier, John Chilton, Nate Coraor, Frederik Coppens, Bert Droesbeke, Ignacio Eguinoa, Simon Gladman, Björn Grüning, Delphine Larivière, Gildas Le Corguillé, Andrew Lonie, Nicholas Keener, Sergei Kosakovsky Pond, Wolfgang Maier, Anton Nekrutenko, James Taylor, Steven Weaver

Submitter

Bert Droesbeke

License

MIT License

Activity

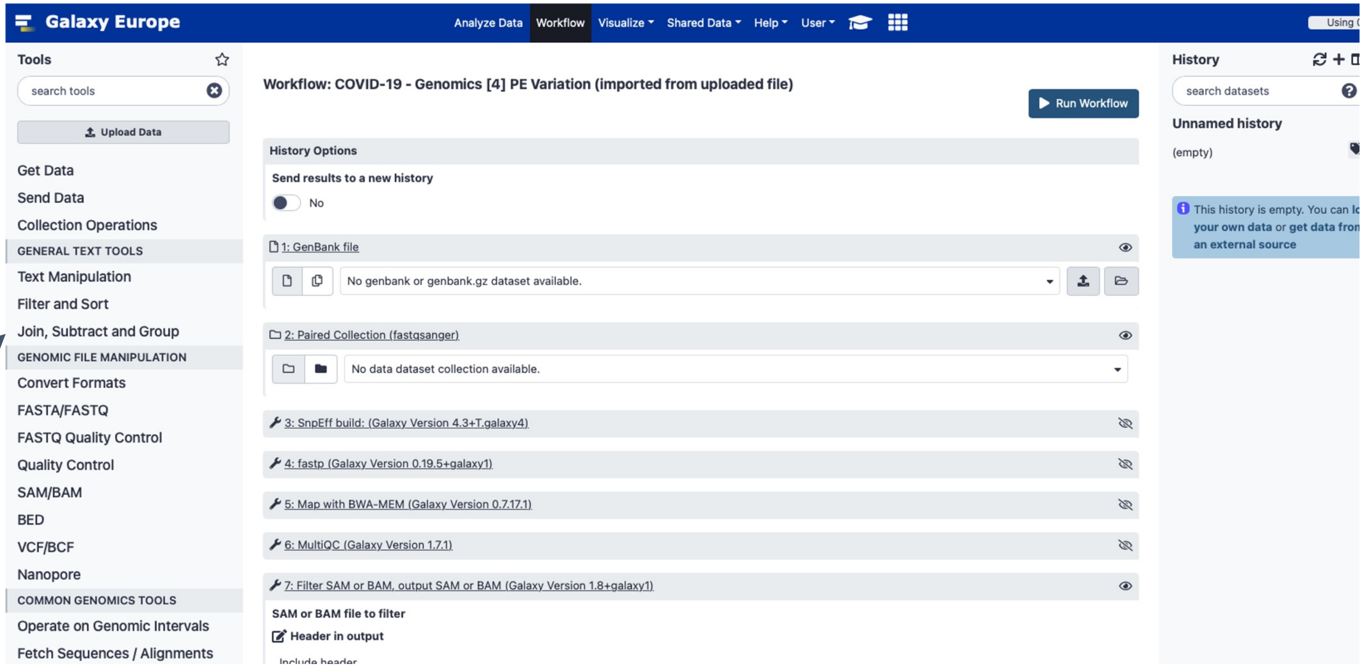
Views: 504 **Downloads:** 28
Created: 10th Apr 2020 at 13:52
Last used: 31st May 2021 at 12:57

Tags

covid-19

Attributions

None



Workflow: COVID-19 - Genomics [4] PE Variation (imported from uploaded file)

History Options

Send results to a new history
○ No

1: GenBank file
No genbank or genbank.gz dataset available.

2: Paired Collection (fastqsanger)
No data dataset collection available.

3: SnpEff build: (Galaxy Version 4.3+T.galaxy4)

4: fastp (Galaxy Version 0.19.5+galaxy1)

5: Map with BWA-MEM (Galaxy Version 0.7.17.1)

6: MultiQC (Galaxy Version 1.7.1)

7: Filter SAM or BAM, output SAM or BAM (Galaxy Version 1.8+galaxy1)

SAM or BAM file to filter

Header in output

Include header

Unnamed history

(empty)

This history is empty. You can lo **your own data** or **get data fron an external source**

NOT an execution platform, but can be coupled to execution platform.

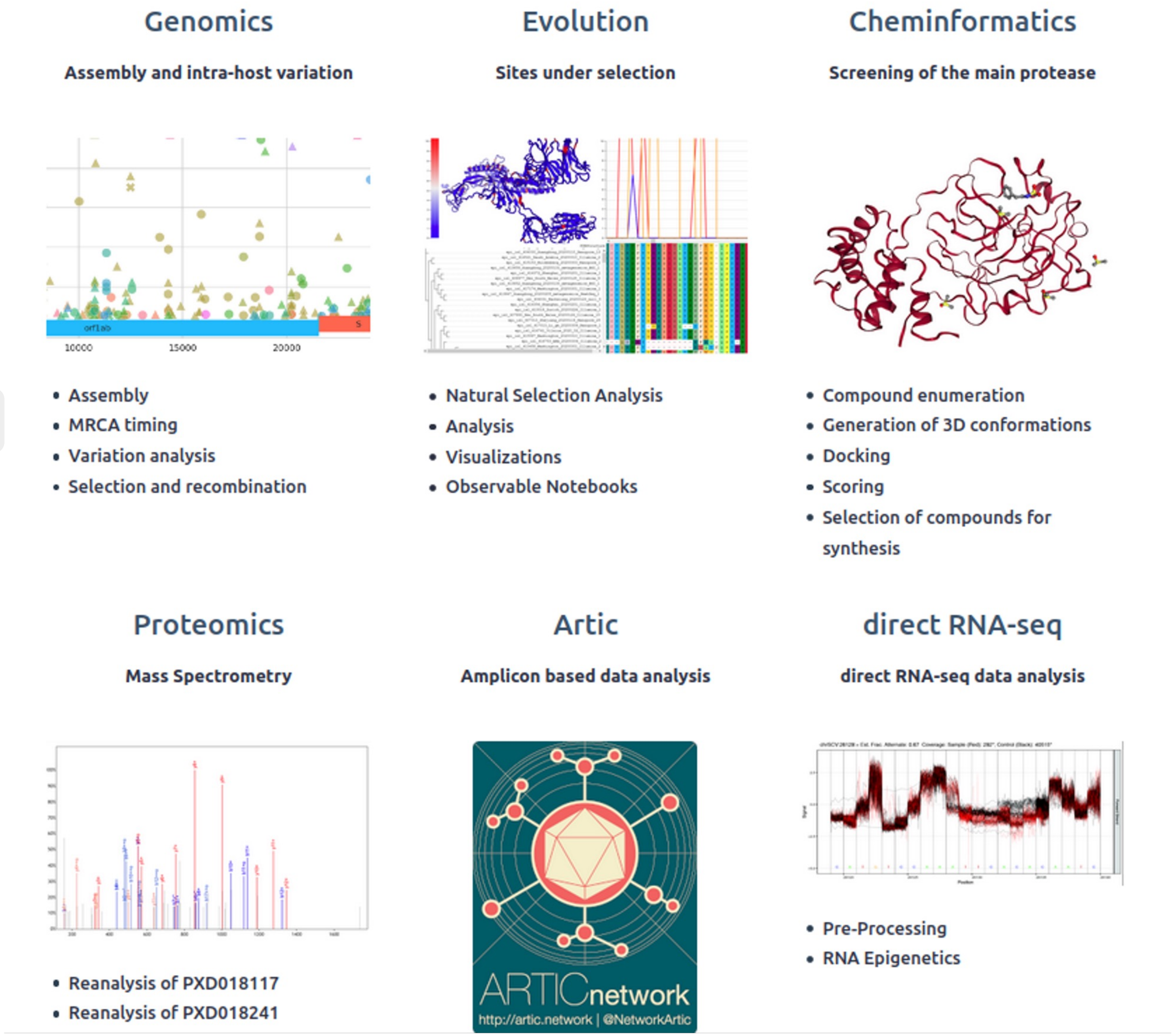
View on GitHub | Download RO-Crate | Run on usegalaxy.eu

elixir BELGIUM

covid19.galaxyproject.org

- Workflows for different disciplines doing COVID-19 research
- COVID-19 Galaxy webinars

elixir-europe.org/events/2nd-galaxy-elixir-webinar-series



Genomics

Assembly and intra-host variation

- Assembly
- MRCA timing
- Variation analysis
- Selection and recombination

Evolution

Sites under selection

- Natural Selection Analysis
- Analysis
- Visualizations
- Observable Notebooks

Cheminformatics

Screening of the main protease

- Compound enumeration
- Generation of 3D conformations
- Docking
- Scoring
- Selection of compounds for synthesis

Proteomics

Mass Spectrometry

- Reanalysis of PXD018117
- Reanalysis of PXD018241

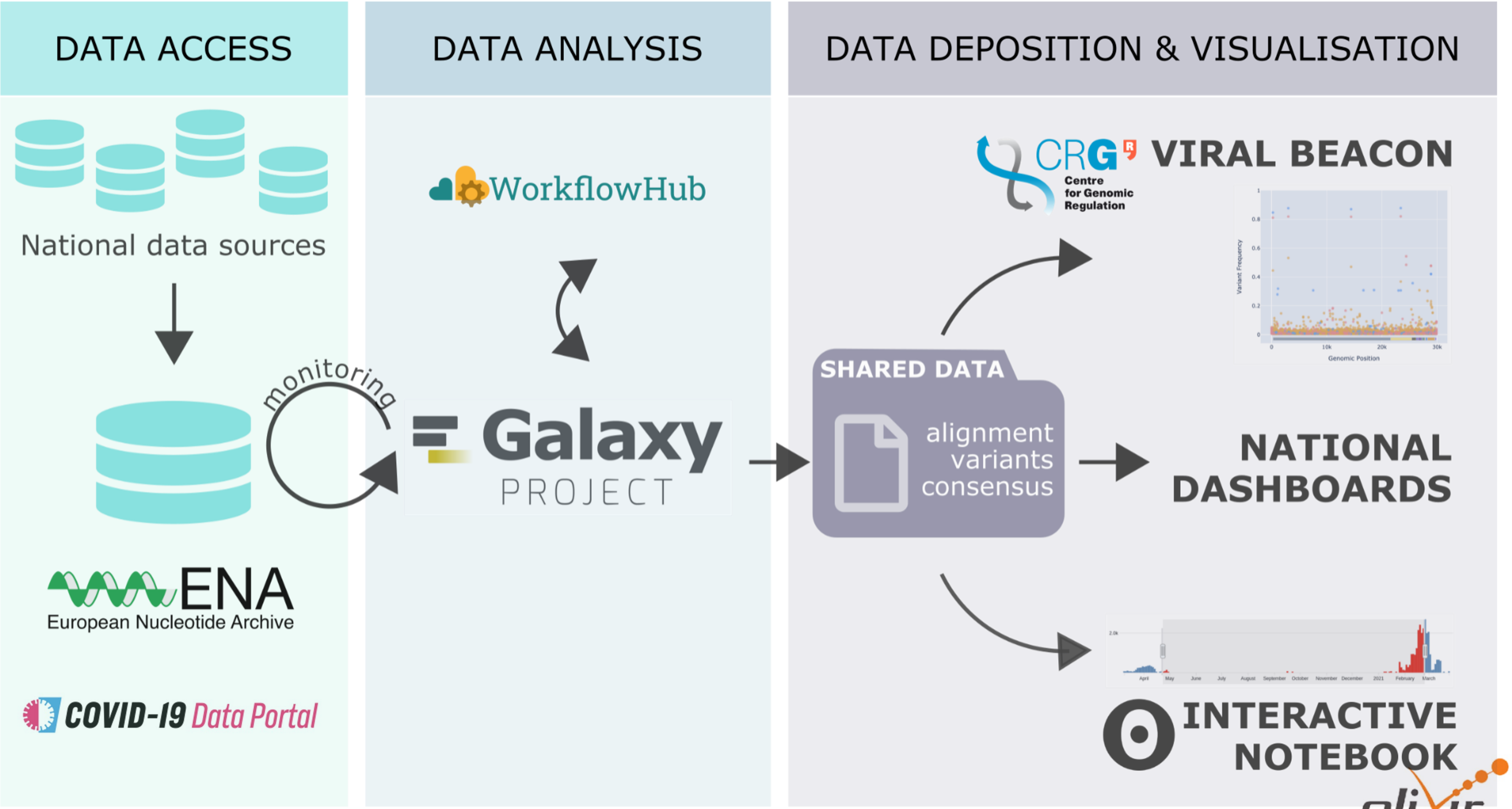
Artic

Amplicon based data analysis

direct RNA-seq

direct RNA-seq data analysis

- Pre-Processing
- RNA Epigenetics



DATA ACCESS

National data sources

monitoring

ENA European Nucleotide Archive

COVID-19 Data Portal

DATA ANALYSIS

WorkflowHub

Galaxy PROJECT

DATA DEPOSITION & VISUALISATION

CRG Centre for Genomic Regulation VIRAL BEACON

SHARED DATA

alignment
variants
consensus

NATIONAL DASHBOARDS

INTERACTIVE NOTEBOOK

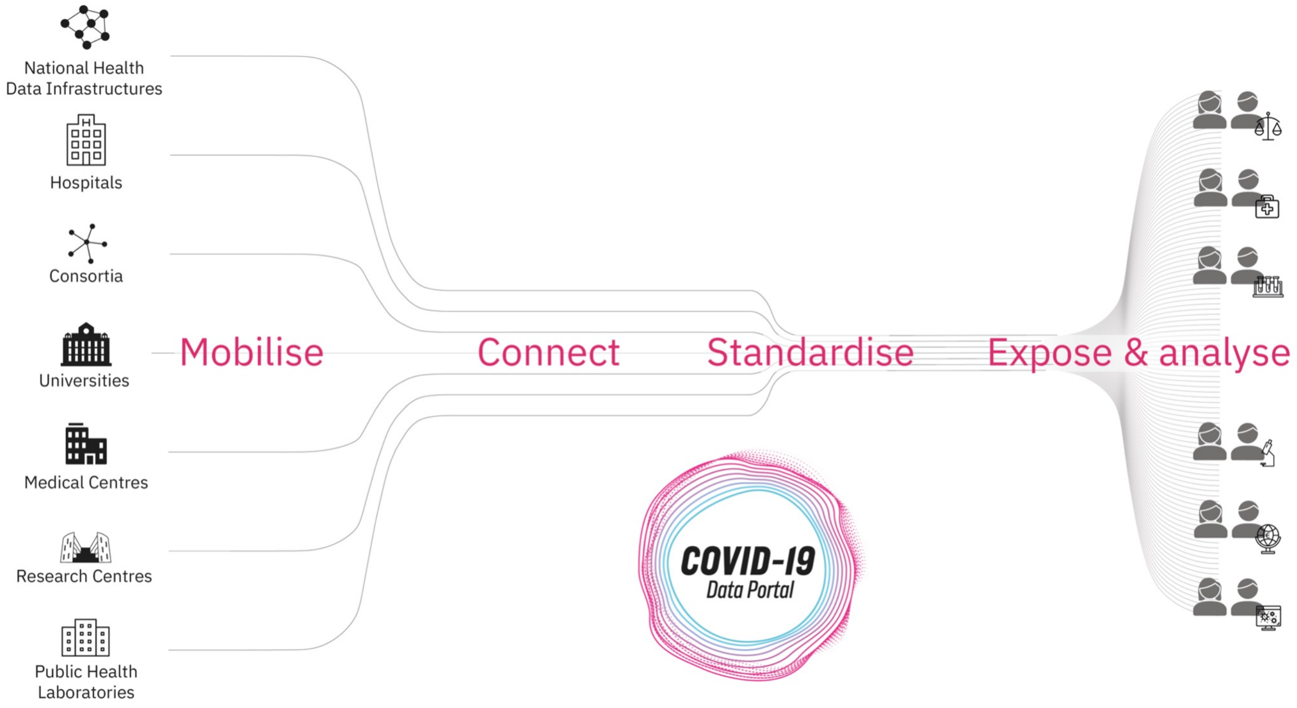
elixir BELGIUM

https://covid19.galaxyproject.org https://covid19beacon.crg.eu

Beyond Genomics

BY-COVID

FAIR and open data sharing in support to European preparedness for COVID-19 and other infectious diseases



BY-COVID creates a multidisciplinary data ecosystem for COVID-19 which will help enhance pandemic preparedness.

elixir BELGIUM

Alignment with other projects and initiatives

PHIRI
Population Health Information
Research Infrastructure

Public Health Initiatives

sciensano

VACCELERATE

veo

Infectious disease projects

Re CoD ID

SSHOC
social sciences & humanities open cloud

EOSC-Life

elixir
CONVERGE

European and Global Research Infrastructures

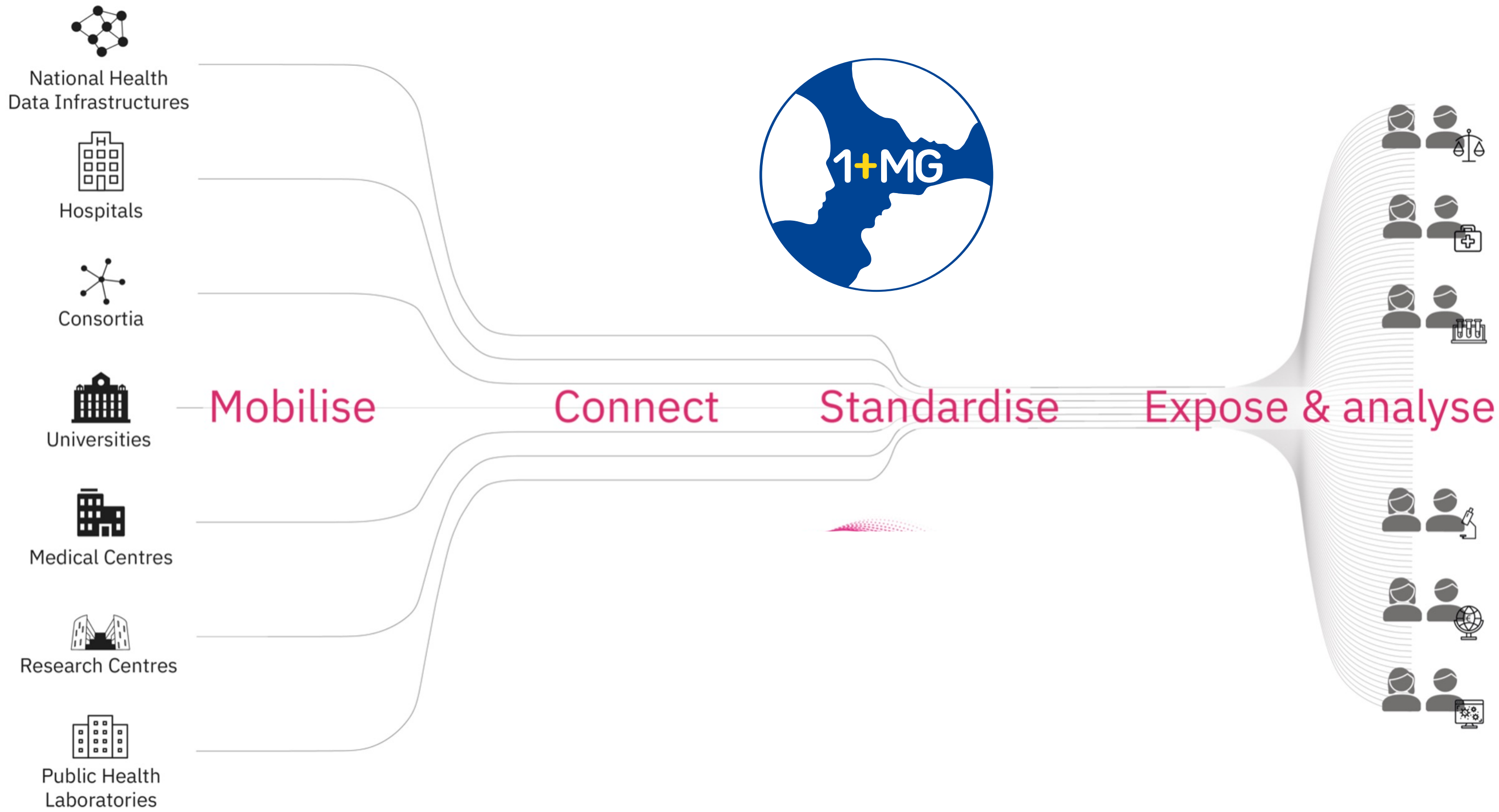
B1MG
Beyond one Million Genomes

Global Alliance
for Genomics & Health
Collaborate. Innovate. Accelerate.

elixir
BELGIUM

Sensitive Data Infrastructure

linking genotype & phenotype



European Health Data Space
Belgian Genome Biobank (FWO - EWI)
Genomic Data Infrastructure (Digital Europe Programme)

Division of responsibilities

- Hardware Infrastructure
  - Secure storage
  - Secure compute
- Services
  - Secure access to data
  - Trusted Research Environment
- Interoperability
  - Standards e.g. GA4GH
  - Alignment with EU developments
- Data management & ELSI

VIB
science meets life

sciensano

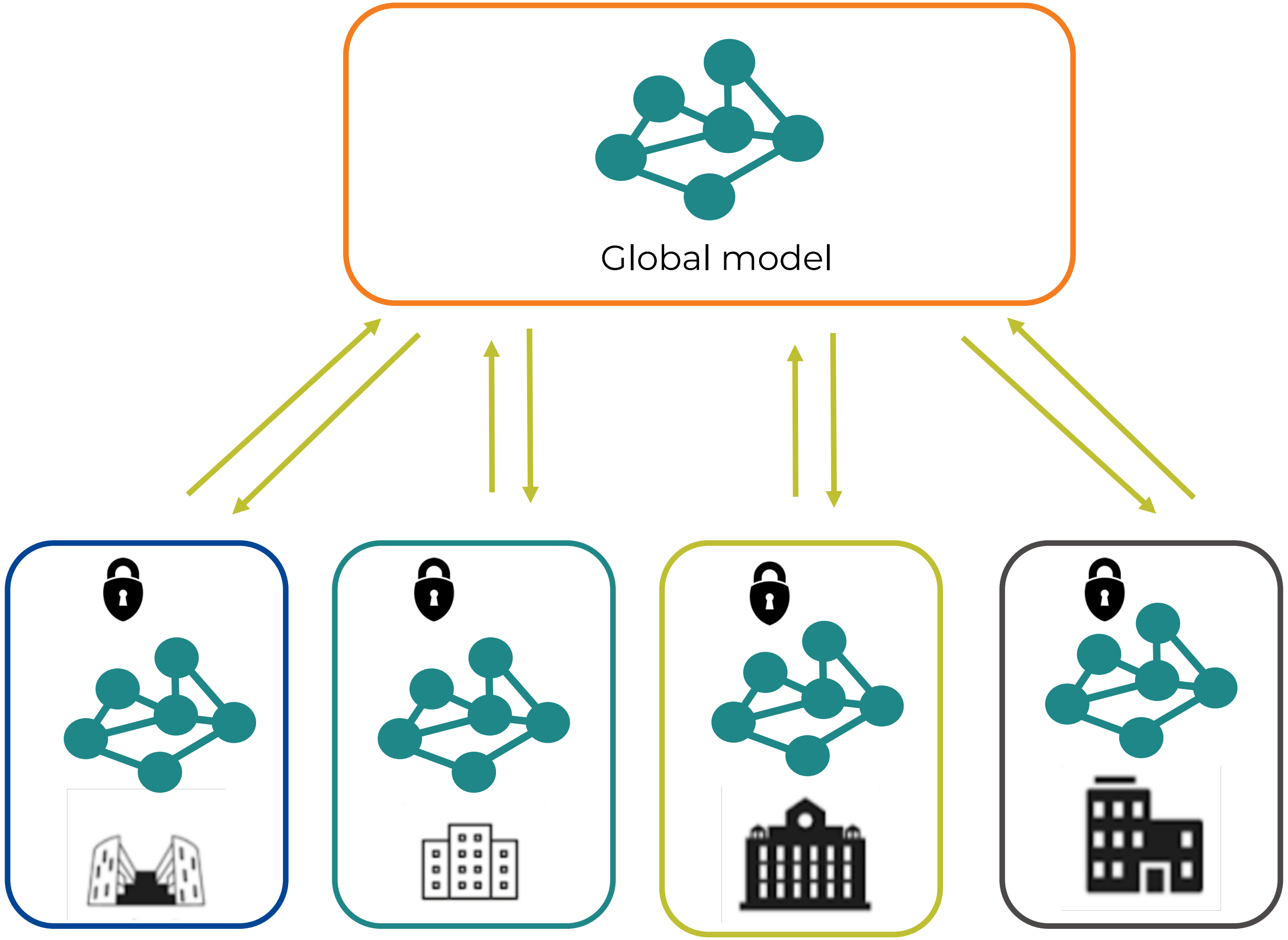
VLAAMS
SUPERCOMPUTER
CENTRUM
Vlaanderen
is supercomputing

elixir
BELGIUM

Accelerating research
building tomorrow's services

ELIXIR BELGIUM

Federated learning framework
underpinning diverse use cases



Global model

Privacy preserving
Local model

WiNGS

EBRAINS

elixir BELGIUM

Biodiversity

Collaborative genome annotation

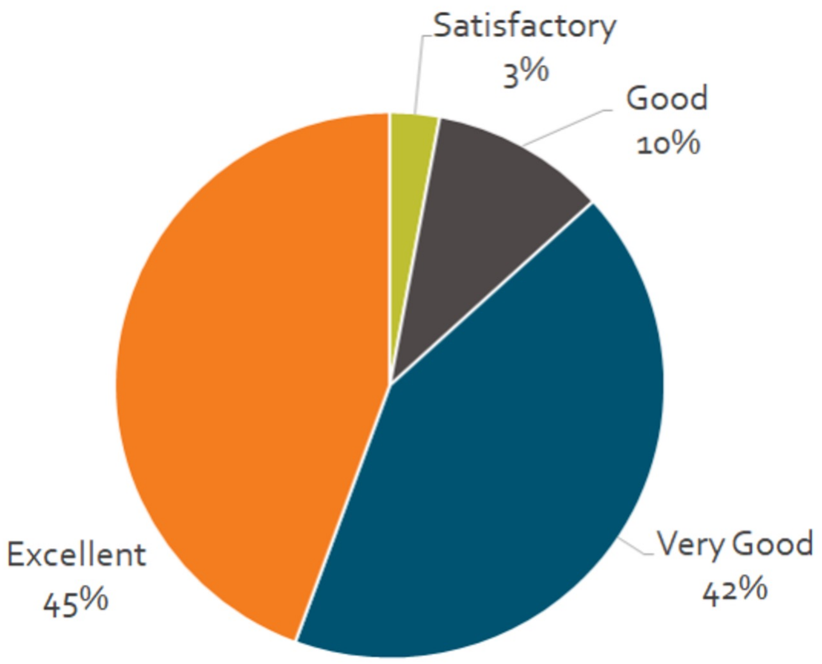
Functional annotation through comparative genomics

Processing & analysis

Bringing services to the users

Training

- e-Learning
- F2F training
- Hackathon
- Knowledge exchange



“Nicely interactive with hands-on exercises”

“Taking the time to demo parts that may be relevant to certain researchers is a great plus as it allows you to see the course material ‘in the wild’”

Outreach & communication

Continue building the ELIXIR community in Flanders & Belgium

Engage with life-sciences community

Data users

Academia
Industry
Public health

Data generators

Core facilities
Sequencing Centers
…

Policy makers

Funders

Thank you

elixir BELGIUM