

Deliverable D8.2 Workshop to explore and plan the inclusion of tools and data from projects that have diverted towards COVID-19 research

Project Title (Grant agreement no.):	ELIXIR-CONVERGE: Connect and align ELIXIR Nodes to deliver sustainable FAIR life-science data management services (871075)		
Project Acronym (EC Call):	ELIXIR-CONVERGE (H2020-INFRADEV-2018-2020)		
WP No & Title:	WP8 ELIXIR-CONVERGE European COVID-19 Data Platform		
WP leader(s):	Guy Cochrane (EMBL-EBI)		
Deliverable Lead Beneficiary:	1 - EMBL-EBI		
Contractual delivery date:	28/02/2021	Actual delivery date:	26/04/2021
Delayed:	Yes		
Partner(s) contributing to this deliverable:	EMBL-EBI		
Authors: Gabriele Rinck (EMBL-EBI), Contributors: Guy Cochrane (EMBL-EBI) , Marianna Ventouratou (EMBL-EBI) Acknowledgments (not grant participants):			
Reviewers:	ELIXIR-CONVERGE Management Board (MB) members.		

Log of changes

DATE	Mvm	Who	Description
31/12/2020	0v1	Guy Cochrane	Initial version
15/03/2021	0v2	Gabriele Rinck (EMBL-EBI)	Sent to PMU after incorporating internal WP feedback
14/04/2021	0v3	Nikki Coutts (ELIXIR Hub)	Circulated to the MB for final review before submission
26/04/2021	1v0	Nikki Coutts (ELIXIR Hub)	Final version uploaded into EC Portal

Table of contents

2. Contribution toward project objectives	2
3. Introduction	4
4. Description of work accomplished	4
5. Results	5
6. Conclusions	5
7. Impact	5
8. Next Steps	6
9. Deviation from Description of Action	6

1. Executive Summary

As part of WP8, Task 8.1. Data management support, we aimed to maximise the use of data and tools through the COVID-19 Portal from EU research projects in which EMBL-EBI is a partner (other than VEO and ReCoDID). Teams across the EMBL-EBI organisation have jointly worked in a concerted effort to pull together all available data and tools relevant to COVID-19-related work of any kind. The European COVID-19 Data Portal currently provides access to over 370,000 SARS-COV-2 viral sequences (as raw reads, just to give an example of one data type) as well as an extensive list of other COVID-related tools and resources with 70 entries, both from external sources as well as EMBL-EBI-internally.

The aim of the deliverable D8.2 was to hold a workshop in order to explore and plan the inclusion of tools and data into the COVID-19 Portal from other EU-funded projects where EMBL-EBI is participating. Due to the broad existing inclusion of data from ELIXIR Core and Deposition Databases (greater than 1,000,000 records at the time of writing) and an extensive list of COVID-related tools and resources already embedded into the Portal's "Related resources", the potential to discover any additional tools or data by holding such a workshop was considered to be quite limited.

Therefore, we followed a time and cost-efficient approach of (a) first gathering information through web searches to identify any relevant data and/or tools EC-funded projects within EMBL-EBI may have produced, followed by (b) contacting PIs of these projects individually, in order to confirm that all available data and tools had already been linked into the Portal and no additional resources had accidentally been overlooked.

Our investigation and discussions with PIs of other EC-funded projects confirmed that the tools and resources listed on the COVID-19 Data Portal already include all relevant tools currently available



across the various groups and projects at the EMBL-EBI. For the time being no additional work is required to add any further tools to the list presented on the Portal.

2. Contribution toward project objectives

With this deliverable, the project has reached or the deliverable has contributed to the following objectives/key results:

Objective no. / Key Result no. Description	Contributed to:
Objective 1: Develop a sustainable and scalable operating model for transnational life-science data management support by leveraging national capabilities (WP1, WP5)	
Key Result 1.1: Established European expert network of data stewards that connect national data centres and similar infrastructures and drive the development of interoperable solutions following international best practice, including national interpretations of the General Data Protection Regulation (GDPR)	No
Key Result 1.2: Development of joint guidelines and common toolkit that are adopted into funder recommendations, with support available nationally and in local languages	No
Key Result 1.3: The catalogue of successful national business models incorporated into national strategies	No
Key Result 1.4: The developed “sustainable and scalable operating model for transnational life-science data management support” is adopted into national ELIXIR Node	No
Objective 2: Strengthen Europe’s data management capacity through a comprehensive training programme delivered throughout the European Research Area (WP2, WP6)	
Key Result 2.1: A comprehensive ELIXIR Training and Capacity building programme in Data Management, directed at both data managers and ELIXIR users, and connected to the national training programmes in Data Management in the ELIXIR Nodes and prospective ELIXIR Member countries.	No
Key Result 2.2: Development of a collective group of trainers that support scalable deployment of Data Management training across ELIXIR Nodes.	No
Key Result 2.3: A substantial cohort of data managers, Node coordinators and researchers with specific data management skills, business planning and knowledge of transnational operations across the ELIXIR Nodes	No
Objective 3: Align national data management standards and services through a sustainable, scalable and cost-effective data management toolkit (WP2, WP3, WP5)	



Key Result 3.1: Assemble a full-stack harmonised common toolkit comprising all aspects of data management: from data capture, annotation, and sharing; to integration with analysis platforms and making the data publicly available according to international standards.	No
Key Result 3.2: Provide exemplar toolkit configurations for prioritised demonstrators to serve as templates for future use.	No
Key Result 3.3: Establish national capacity in using as well as updating, extending and sustaining the toolkit across the ERA.	No
Key Result 3.4: Enable 'FAIR at source' practice for data generation, and analytical process pipeline implementation by flexible deployment of the toolkit in national operations	No
Objective 4: Align national investments to drive local impact and global influence of ELIXIR (WP4,WP6)	
Key Result 4.1: Development of a Node Impact Assessment Toolkit based on RI-PATHS methodology.	No
Key Result 4.2: Adoption of Impact assessment in ELIXIR Nodes, supported by Node coordinators network and feedback on applicability from dialogues with national funders.	No
Key Result 4.3: Creation of national public-private partnerships and industry outreach where open life-science data and services stimulate local bioeconomy	No
Key Result 4.4: Growth in reach, impact and engagement of stakeholder communication assessed by established ELIXIR Communications metrics	No
Key Result 4.5: Initiating and advancing discussions on Membership (EU and international) or strategic partnerships (international countries) following ELIXIR-CONVERGE workshops.	No
Objectives - WP8 - ELIXIR-CONVERGE European COVID-19 Data Platform	
08.1 Data management support for EU projects (Task 8.1)	Yes
08.2 Mobilisation of analysis upon SARS-CoV-2 sequence data (Task8.2)	No
08.3 Enhanced access to data, tools and support (Task 8.3)	No

3. Introduction

As part of WP8, Task 8.1. Data management support, we aimed to maximise the use of data and tools through the COVID-19 Portal from EU research projects in which EMBL-EBI is a partner (other than VEO and ReCoDID). We envisaged to include the following projects:

- deeply mined COVID-19 related literature, from Europe PMC; development work on this will include COVID-19-specific text mining as an EMBL-EBI contribution to the OpenAIRE initiative (OpenAIRE);
- compound screening and assay data relating to ongoing COVID-19 related work (EUbOPEN and eTRANSafe);
- cheminformatics tools that will assist with the integration of compound-related COVID-19 data (EU-ToxRisk, EUbOPEN and TransQST);
- access to tools and interfaces (such as metadata validation and discovery tools) relating to clinical and epidemiological data (CINECA);

In order to accomplish this task we planned to hold a workshop to explore and plan the inclusion of these tools and data. Instead we held several one-to-one meetings with the PIs leading the relevant projects with the following outcomes.

4. Description of work accomplished

Since the beginning of the COVID-19 pandemic, EMBL-EBI and ELIXIR teams have jointly worked in a concerted effort to pull together all available data and tools relevant to COVID-19 research. This resulted in the launch of the European COVID-19 Data Portal¹ in April 2020. It is an ongoing endeavour to constantly make improvements to the Platform and add tools or resources as soon as they become available. The extent and speed of the implementation significantly exceeded what was initially envisaged at the time of the grant proposal.

The COVID-19 Data Portal currently provides access to over 370,000 SARS-COV-2 viral sequences (as raw reads, just to give an example of one data type) as well as an extensive list of other COVID-related tools and resources with 70 entries, both from external sources as well as EMBL-EBI and ELIXIR-internally.

The aim of the deliverable D8.2 was to hold a workshop in order to explore and plan the inclusion of tools and data into the COVID-19 Portal. Due to the extensive list of COVID-related data and tools and resources already embedded into the Portal the potential to discover any additional tools or data by holding such a workshop was considered to be quite limited. We therefore opted for a more time and cost-efficient approach of (a) first gathering information through web searches to identify any relevant remaining data and/or tools EC-funded projects within EMBL-EBI may have produced, followed by (b) contacting PIs of these projects individually, in order to confirm that all available data and tools had already been linked into the Portal and no additional resources had accidentally been overlooked.

The following projects were approached:

¹ <https://www.covid19dataportal.org/>



Grant Acronym	Contract Number
discovAIR	874656
EU-ToxRisk	681002
EUbOPEN	875510
TransQST	116030
eTRANSafe	777365
CINECA	825775

The outcome of these investigations and discussions confirmed that the list of tools and resources presented on the COVID-19 Data Portal already includes all the data and tools currently available through the relevant projects. All PIs were reminded of our effort to include further data and tools if any additional ones become available.

5. Results

Our investigation and discussions with PIs of other EC-funded projects confirmed that the tools and resources listed on the COVID-19 Data Portal already include all relevant tools currently available across the various groups and projects at the EMBL-EBI. For the time being no additional work is required to add any further data or tools to the list presented on the Portal.

6. Conclusions

The fact that all currently available tools and resources are already represented on the COVID-19 Data Portal indicates that the Portal has been established and maintained in a very systematic and collaborative way, successfully bringing together a comprehensive set of data, resources and tools from across the EMBL-EBI organisation and beyond. We should, and will, continue to monitor EU and other projects as further data and tools of relevance emerge from these such that we include into the COVID-19 Data Portal where appropriate.

7. Impact

The European COVID-19 Data Portal serves as a key entry point into a wide range of data, resources and tools, all aiming to facilitate data sharing and analysis, and to accelerate coronavirus research. Ensuring that the collection of available resources and tools is as comprehensive as possible maximises the benefit to the Portal users. This Deliverable 8.2 confirmed that all currently available



data and tools from EMBL-EBI and ELIXIR teams have been incorporated into the Portal, contributing to the utility of the Portal.

8. Next Steps

We will continually strive to maintain the content of the European COVID-19 Data Portal to high standards, offering a comprehensive range of data, resources and tools, while at the same time aiming to make the website as user-friendly as possible. This includes adding new data tools as and when they become available. As part of other Horizon 2020-funded projects (such as VEO), we are also working on integrating tools and resources from the external members of those project consortia, but this will be reported as independent Deliverables within those projects.

9. Deviation from Description of Action

Not applicable. The only deviation is that this deliverable was not delivered in a single workshop-type meeting, but several meetings with the relevant project representatives.

