

# Deliverable 8.2.2 Project Data

## Management Plan initial release and periodic updates

### Data Management Plan

<b>Project Title</b> (grant agreement No)	Beyond COVID Grant Agreement 101046203		
<b>Project Acronym</b> (EC Call)	BY-COVID		
<b>WP No &amp; Title</b>	WP8: Coordination, Project Management and Ethical, Legal and Social implications		
<b>WP Leaders</b>	Katharina Lauer, ELIXIR; Marieke Willems, ELIXIR		
<b>Deliverable Lead Beneficiary</b>	7 BBMRI		
<b>Contractual delivery date</b>	31/12/2022	<b>Actual Delivery date</b>	23/12/2022
<b>Delayed</b>			
<b>Partner(s)</b> contributing to this deliverable	EBI, ELIXIR Hub, ELIXIR-UK Sciensano, IACS, BBMRI		
<b>Authors</b>	Eva Garcia Alvarez - BBMRI, Petr Holub - BBMRI		
<b>Contributors</b>	Ilaria Colussi - BBMRI		
<b>Acknowledgements</b> (not grant participants)			
<b>Reviewers</b>	Henning Hermjakob - EBI Katharina Lauer - ELIXIR Hub Marieke Willems - ELIXIR Hub Susanna-Assunta Sansone - ELIXIR-UK Nina Van Goethem - Sciensano Enrique Bernal Delgado - IACS Niklas Blomberg - ELIXIR		



## Log of changes

Date	Mvm	Who	Description
14/11/2022		Eva García Álvarez Petr Holub	

## Table of contents

<b>1. Executive Summary</b>	<b>2</b>
<b>2. Contribution towards project objectives</b>	<b>4</b>
Objective 1	4
Objective 2	4
Objective 3	5
Objective 4	5
Objective 5	6
<b>3. Methods</b>	<b>7</b>
<b>4. Description of work accomplished</b>	<b>7</b>
4.1 Project overview	7
4.2 Data flow	9
4.3 Open Science and FAIR sharing	12
<b>5. Conclusions</b>	<b>13</b>
<b>6. Next steps</b>	<b>13</b>
<b>7. Impact</b>	<b>14</b>

## 1. Executive Summary

The Data Management Plan (hereinafter “the project DMP”) is a mandatory deliverable by the European Commission (EC) that has to be tailored to the specific project activities and which should document the relevant project repositories and how they are managed. The Data Management Plan for BY-COVID is developed by WP8 as an integral part of the governance structure, and also oversees the quality assurance of the project outcomes. BY-COVID will not generate new sensitive patient level data but integrate and link existing resources. Hence, responsibility remains on data access committees and processes for each dataset. The project DMP considers:

- Overview of the project data flow
- Compliance with Open Science and FAIR data sharing
- Ethics Requirements<sup>1</sup> as required according to the Grant Agreement (WP9)

This document is BY-COVID Data Management Plan (DMP) version 0.2 delivered in month 15, being an update of the first version delivered in month 06. This deliverable is based on the template and the guidelines provided by the European Commission for Horizon Europe. Of note, the project DMP is a live document that will be reviewed and updated periodically to ensure it remains up to date.

---

<sup>1</sup> García Álvarez, Eva, Mayrhofer, Michaela, & Holub, Peter. (2021). BY-COVID - D8.2 - Data Management Plan (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.6884816> (Annex 2)

## 2. Contribution towards project objectives

The project handbook<sup>2</sup> defines the project process that provides the framework to accomplish all projects objectives within the scope, budget and the required level of quality. This deliverable contributes to all objectives as listed below:

	Key Result No and description	Contributed
<b>Objective 1</b> Enable storage, sharing, access, analysis and processing of research data and other digital research objects from outbreak research.	1. A research data management practice in European research infrastructures practice that drives discovery, access and reuse of outbreak data and directly links experimental data from HORIZON-INFRA-2021-EMERGENCY-02 transnational access projects into the COVID-19 Data Portal.	Yes
	2. Workflows and processing pipelines that integrate transparent quality management and provenance and are openly shared.	Yes
	3. Research infrastructures on-target training so that users can exploit the platform.	Yes
	4. Engagement so that stakeholders (RI, national centres, policy makers, intergovernmental organisations, funders and end-users) incorporate FAIR and open data in infectious disease guidelines and forward planning.	Yes
<b>Objective 2</b> Mobilise and expose viral and human infectious disease data from national centres.	1. A comprehensive registry of available data with established procedures to collate data governance models, metadata descriptions and access mechanisms in a pandemic scenario.	Yes
	2. Mechanisms for the initial discovery across data sources based on available metadata at the reference collection.	Yes
	3. Demonstrated transnational linking of real-world data from national surveillance, healthcare, registries and social science data that allow the assessment of variants	Yes

<sup>2</sup> Arenas Marquez, Juan, & Troncoso Quilaqueo, Andrea. (2021). BY-COVID - D8.1 - Project Management Handbook (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.6884734>

	<p>to serve the research needs of epidemiology and public health.</p> <p>4. Demonstrated assessment of emerging SARS-CoV-2 variants against data generated in the on-going European VACCELERATE clinical trials project to investigate vaccine efficacy.</p>	Yes
<p><b>Objective 3</b></p> <p>Link FAIR data and metadata on SARS-CoV-2 and COVID-19</p>	<p>1. A platform that links normative pathogen genomes and variant representations to research cohorts and mechanistic studies to understand the biomolecular determinants of variant response on patient susceptibility, and disease pathways.</p>	Yes
	<p>2. An open and extensible metadata framework adopted cross-domain that supports comprehensive indexing of the infectious disease resources based on mappings across resources and research domains.</p>	Yes
	<p>3. A provenance framework for researchers and policy-makers that enables trust in results and credit to data submitters, workflow contributors and participant resources.</p>	Yes
<p><b>Objective 4</b></p> <p>Develop digital tools and data analytics for pandemic and outbreak preparedness, including tracking genomics variations of SARS-CoV-2 and identifying new variants of concern.</p>	<p>1. Broad uptake of viral <i>Data Hubs</i> across Europe deliver an order-of-magnitude increase in open viral variant detection and sharing.</p>	Yes
	<p>2. Infrastructure and quality workflows mobilised and shared to produce open, normative variant data that is incorporated into national and regional data systems and decision making.</p>	Yes

<p><b>Objective 5</b> Contribute to the Horizon Europe European Open Science Cloud (EOSC) Partnership and European Health Data Space (EHDS).</p>	<p>1. Guidelines and procedures for FAIR data management and access will be established, building on work of other guideline producing consortia such as the Global Alliance for Genomics and Health (GA4GH), the 1Mio Genomes Initiative (1MG) and the Beyond One Million Genomes project (B1MG).</p>	<p>Yes</p>
	<p>2. Services, software, protocols, guidelines and other research objects that are openly accessible for reuse by the EOSC Association and the community at large as a foundation for European preparedness for infectious diseases, leveraging developments in EOSC-Life, SSHOC, EOSC-Future, EGI-ACE and other EOSC projects.</p>	<p>Yes</p>
	<p>3. Alignment (both policy and implementation routes) will have been achieved between the data governance strategies for routinely collected health data in the EHDS initiative, including the TEHDAS Joint Action and future EHDS Pilot Actions.</p>	<p>Yes</p>
	<p>4. To empower national centres to build capacity and train platform users and data providers (e.g., from life, social or health sciences), and with experts from across partner institutions collaborating to create training materials for the identified gaps, and to exchange experiences and knowledge.</p>	<p>Yes</p>

## 3. Methods

The integrated ecosystem that BY-COVID will deliver is based on established services and resources, ensuring sustainability beyond the lifetime of the project. The adoption of open source platforms on commonly used cloud-based infrastructures and connection of the components through open standards supports long-term sustainability in a distributed landscape.

The respective data access committees and processes for each dataset remain accountable for the basis for project operations. Data security remains the responsibility of the databases and archives holding data (e.g., the EGA has a defined security process<sup>3</sup> and follows best practice guidelines aligned with the GA4GH Security Working Group, the BBMRI-ERIC Policy for Access to and Sharing of Biological Samples and Data<sup>4</sup> and follows OECD Council Recommendations on Health Data Governance<sup>5</sup>). IP background and existing data licences for the project might be diverse. IP rights will be managed in the consortium agreement. Where code, software or ontologies are developed, a permissive licence will be applied and an open approach will be taken by partners in line with their institutional policies. While access rights and licences predating the BY-COVID project will be respected, participating organisations will be encouraged to review existing licences to comply with Open Access requirements. Data sharing and reuse will follow FAIR principles<sup>6</sup> (Findability, Accessibility, Interoperability, and Reusability) and make use of licences such as the Creative Commons or Open Data Commons. Release on publication, as part of the open access policies, is envisioned and BY-COVID will follow established embargo principles (e.g., in the European Nucleotide Archive, ENA; release after 6 or 12 months).

## 4. Description of work accomplished

### 4.1 Project overview

BY-COVID aims to identify, connect and integrate data for the effective study of the COVID-19 disease and causative agent as well as other infectious diseases. Infectious diseases are complex and their analysis requires data from different

<sup>3</sup> European Genome-phenome Archive: Security Overview

[https://ega-archive.org/files/European\\_Genome\\_phenome\\_Archive\\_Security\\_Overview.pdf](https://ega-archive.org/files/European_Genome_phenome_Archive_Security_Overview.pdf)  
[accessed 16.11.2022]

<sup>4</sup> BBMRI-ERIC Policy for Access to and Sharing of Biological Samples and Data:

[https://www.bbmri-eric.eu/wp-content/uploads/AoM\\_10\\_8\\_Access-Policy\\_FINAL\\_EU.pdf](https://www.bbmri-eric.eu/wp-content/uploads/AoM_10_8_Access-Policy_FINAL_EU.pdf)  
[accessed 16.11.2022]

<sup>5</sup> [OECD Council Recommendations on Health Data Governance](#) [accessed 16.11.2022]

<sup>6</sup> Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

disciplines. Thus, BY-COVID will link established and emerging research infrastructures and data resources from biomolecular research, public health, clinical research and social science. This will be done using standards fully aligned with the European Open Science Cloud (EOSC<sup>7</sup>), such as “Guidance and policy on standards and tools to facilitate sharing and reuse of multimodal data (including imaging), cohort integration, and biosamples”<sup>8</sup> or “Report on data standards for observational and interventional studies, and interoperability between healthcare and research data”<sup>9</sup> and with the European Health Data Space (EHDS) proposal<sup>10</sup>, mainly through the outcomes from the currently ongoing pilot project<sup>11</sup>.

Three main pillars can be identified in the project methodology:

**Mobilisation** - Mobilised data will be indexed and organised in the COVID-19 Data Platform (<https://www.covid19dataportal.org/>), where (meta)data is incorporated through a flexible, tiered system for data integration. Data incorporated into the COVID-19 Data Platform will be embedded in the wider EOSC data ecosystem, establishing guidelines and procedures for FAIR data management and ensuring long term, rapid open access.

BY-COVID will allow mobilisation - meaning access and transfer - of data by using a trialled system of SARS-CoV-2 Data Hubs and other existing infrastructures such as the European Nucleotide Archive (ENA), CESSDA social science archives, and biobank catalogues in a “federation of federations”, following community practised

<sup>7</sup> EOSC-Portal, <https://eosc-portal.eu/> [accessed 16.11.2022]

<sup>8</sup> Boiten, Jan-Willem, Ohmann, Christian, Adeniran, Ayodeji, Canham, Steve, Cano Abadia, Monica, Chassang, Gauthier, Chiusano, Maria Luisa, David, Romain, Fratelli, Maddalena, Gribbon, Phil, Holub, Petr, Ludwig, Rebecca, Th. Mayrhofer, Michaela, Matei, Mihaela, Merchant, Arshiya, Panagiotopoulou, Maria, Pireddu, Luca, Richard, Audrey, Sanchez Pla, Alex, ... Gorianin, Sergei. (2021). EOSC-Life Guidance and policy on standards and tools to facilitate sharing and reuse of multimodal data (including imaging), cohort integration, and biosamples. Zenodo. <https://doi.org/10.5281/zenodo.4591011>

<sup>9</sup> Canham, Steve, Ohmann, Christian, Boiten, Jan-Willem, Panagiotopoulou, Maria, Hughes, Nigel, David, Romain, Sanchez Pla, Alex, Maxwell, Lauren, Aerts, Jozef, Facile, Rhonda, Griffon, Nicolas, Saunders, Gary, van Bochove, Kees, & Ewbank, Jonathan. (2021). EOSC-Life Report on data standards for observational and interventional studies, and interoperability between healthcare and research data. Zenodo. <https://doi.org/10.5281/zenodo.5810612>

<sup>10</sup> Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the European Health Data Space, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0197&qid=1669300754959> [accessed 16.11.2022]

<sup>11</sup> European Health Data Space - EHDS HealthDat@EU Pilot, <https://www.ehds2pilot.eu/> [accessed 16.11.2022]



standards. Hence, ultimate data responsibility belongs to the data providers, reflected in their own DMP or data best practises guidance.

Connect and expose - A tiered indexing system will be developed, protecting truth and privacy, specially in the case of sensitive data. Guidelines for ensuring data interoperability will be established by implementing community-driven standards, together with offering support to the discovery of metadata related to sensitive data and the integration of COVID-19 data resources in the COVID-19 Data Portal<sup>12</sup>.

Use & analyse data - BY-COVID will integrate standardised data management and analysis methods and protocols to ensure FAIR and FAIR-Health<sup>13</sup> are an integral part of the process.

## 4.2 Data flow

As stated in the first version of this plan, the data flow is distributed among the different technical Work Packages (WP1-WP2-WP3-WP4-WP5), flowing in a domain specific manner. During these 15 months of the project, this flow followed the plan foreseen in month 06, and a brief description on how it is being developed can be found in this section.

WP1 is establishing SARS-CoV-2 Data Hubs that handle centralised data, which are not sensitive after de-identification. In line with the described procedure in the first version of this DMP, WP2 is gathering the relevant Data Hubs focussing on data other than viral sequences and offering interoperable data as well as open-source solutions for data interoperability.

WP2 is collecting the existing data and metadata controllers, looking at their alignment with proposed standards and FAIR principles. The activity of this WP started mainly focusing on discoverability. Thus, integration and harmonisation of non-patient data and an open source harvesting tool for socio-economic data (both non-sensitive data types) are being developed.

---

<sup>12</sup> COVID-19 Data Portal, <https://www.covid19dataportal.org/> [accessed 16.11.2022]

<sup>13</sup> Holub P, Kohlmayer F, Prasser F, et al. Enhancing Reuse of Data and Biological Material in Medical Research: From FAIR to FAIR-Health. *Biopreserv Biobank*. 2018;16(2):97-105.  
<https://doi:10.1089/bio.2017.0110>

Regarding sensitive data, institutions are being onboarded to the COVID-19 Data Portal and some metadata resources have been harmonised following the specifications from WP3 (see below) and are also going to be included in the portal by WP1. One example of this approach can be found in the BY-COVID deliverable D3.2<sup>14</sup>, which describes the indexing of social sciences metadata from The Consortium of European Social Science Data Archives (CESSDA<sup>15</sup>) and their integration in the COVID-19 Data Portal.

In addition, WP2 is developing open source federated search tools and conversion tools to format data following the FAIR principles. Of note, all tools for data harvesting or data gathering developed in this WP do not mobilise data, they always stay in the source. A more in depth look at the activities described above is described in the D2.1<sup>16</sup>, which is included in the BY-COVID Zenodo community<sup>17</sup>.

While WP2 looks at intradomain harmonisation, cross-domain efforts are done in WP3. The collaborative work of these two WPs resulted in the creation of a FAIRsharing BY-COVID Collection<sup>18</sup>, linking the description of the data sources to their datasets indexed in the COVID-19 Data Portal. Indeed, FAIRsharing, which was established as the tier 3 of the tiered index described below, is endorsed by the Research Data Alliance (RDA) and used to provide information about the standards, relation and description of data from different sources across all disciplines.

WP3 has developed an infrastructure for the three tiers of the COVID-19 Data Portal discoverability schema, which is based on the metadata model described in D3.1 “Metadata standards. Documentation on metadata standards for inclusion of

---

<sup>14</sup> Hermjakob, Henning, Kleemola, Mari, Moilanen, Katja, Tuominen, Markus, Sansone, Susanna-Assunta, Lister, Allyson, David, Romain, Panagiotopoulou, Maria, Ohmann, Christian, Belien, Jeroen, Lischke, Julia, Juty, Nick, & Soiland-Reyes, Stian. (2022). BY-COVID D3.2: Implementation of cloud-based, high performance, scalable indexing system. Zenodo.

<https://doi.org/10.5281/zenodo.7129553>

<sup>15</sup> The Consortium of European Social Science Data Archives (CESSDA), <https://www.cessda.eu/> [accessed 16.11.2022]

<sup>16</sup> Giles, Tom, Quinlan, Phil, Belien, Jeroen, Lischke, Julia, Portell-Silva, Laura, Capella-Gutierrez, Salvador, Karki, Reagon, Kalaitzi, Vasso, Bernal-Delgado, Enrique, & Keppler, Antje. (2022). BY-COVID- D2.1 - Initial data and metadata harmonisation at domain level to enable fast responses to COVID-19 (V1.0). Zenodo. <https://doi.org/10.5281/zenodo.7017728>

<sup>17</sup> BY-COVID Zenodo community, <https://zenodo.org/communities/bycovid/?page=1&size=20> [accessed 16.11.2022]

<sup>18</sup> BY-COVID data resource collection in FAIRsharing, <https://fairsharing.org/3773> [accessed 16.11.2022]

resources in data portal”<sup>19</sup>. It contains open access tools and workflows ready to integrate resources from several domains into the different tiers, depending on the sensitivity and availability of the (meta)data from each resource.

The work developed in BY-COVID project is meant to be used beyond it. To be compliant with this statement, the code and lessons learned from part of the work in WP3 are being extended to related portals Early Cause<sup>20</sup> and Pathogens<sup>21</sup> using the “Baseline Portal” software package.

Apart from related portals, the outcomes from this project should be available beyond its end. Thus, for taking care of sustainability, WP3 collaborates with EOSC and RDA initiatives, including the RDA FAIRsharing WG and the Life Science IG.

While data resources and the standards implemented are registered and described in FAIRsharing, tools and other services are included in the Infectious Disease Toolkit (IDTk)<sup>22</sup> developed by WP4. The IDTk is currently under development and will provide information and references to open access analysis methods and protocols for the four different domains considered in the BY-COVID project: Pathogen characterisation, Human biomolecular data, Human clinical and health data and Socio-economics data.

This WP is also working on a semantic interoperability model, for describing datasets and their provenance through Research Objects (RO-Crate)<sup>23</sup> and schema.org<sup>24</sup> vocabularies, along with promoting the use of other relevant terminologies, via FAIRsharing. It is also collecting the quality measures used to deal with infectious diseases related data and preparing them to be part of the IDTk. Finally, this WP is providing advice to the use cases with regard to the FAIRification of their scripts and workflows through RO-Crate and WorkflowHub<sup>25</sup>,

---

<sup>19</sup> Hermjakob, Henning, Kleemola, Mari, Moilanen, Katja, Sansone, Susanna-Assunta, Lister, Allyson, David, Romain, Panagiotopoulou, Maria, Ohmann, Christian, Belien, Jeroen, Lischke, Julia, Juty, Nick, & Soiland-Reyes, Stian. (2022). BY-COVID - D3.1 - Metadata standards. Documentation on metadata standards for inclusion of resources in data portal (V1.0). Zenodo.  
<https://doi.org/10.5281/zenodo.6885016>

<sup>20</sup> Early Cause, <https://portal.earlycause.eu/> [accessed 16.11.2022]

<sup>21</sup> Pathogens portal, <https://www.ebi.ac.uk/ena/pathogens/home> [accessed 16.11.2022]

<sup>22</sup> Infectious Disease Toolkit, <https://www.infectious-diseases-toolkit.org/> [accessed 16.11.2022, under development]

<sup>23</sup> RO-Crate, <https://www.researchobject.org/ro-crate/> [accessed 16.11.2022]

<sup>24</sup> [schema.org](https://schema.org) [accessed 16.11.2022]

<sup>25</sup> WorkflowHub, <https://workflowhub.eu/> [accessed 16.11.2022]

both aligned with EOSC and the emerging EHDS. This practical FAIRification examples will also become recipe in the FAIR Cookbook<sup>26,27</sup>.

Finally, the BY-COVID baseline use case (WP5) aims to provide answers to policy-relevant questions by leveraging observational data (individually-linked sensitive data) obtained from multiple existing data sources (e.g., clinical, administrative and socio-economic data). Aggregated analysis results will be obtained after the federated deployment of the analytic pipeline. This federated infrastructure imposes that the scripts (containerized) move towards the data, while the sensitive data is kept at the premises of each of the nodes (i.e., institutions that host or have access to the required data such as Sciensano, IACS, THL, GÖG, etc.). As such, all the analyses with individual-level sensitive data are performed at the nodes' premises following their own governance rules and regulatory restrictions. The digital objects constituting the pipeline that are already produced, such as the causal model, Common Data Model and synthetic data sets are being published in Zenodo.

Liaisons with the technical work packages are necessary also for the domain specific use cases (i.e., “Secondary use of COVID-19 vaccine trial data and biosamples to test existing vaccines against emerging variants”, “Early detection of epidemic hotspots”, and “Mechanistic understanding of SARS-CoV-2 variants and COVID-19 outcomes”). Currently, cross-fertilisation with WP1, 2 and 3 regarding the discovery and indexing of data sources and with WP4 for the FAIRification of the workflows, are identified as the main interconnection points.

### 4.3 Open Science and FAIR sharing

In the above summary of data flow, it is already possible to see that resources and open access tools, such as Zenodo (an OpenAIRE initiative), FAIRsharing or RO-Crate, and FAIR practices, as documented in the IDTk and FAIR Cookbook are being used throughout the project as planned, ensuring findability, discoverability and effective sharing. Furthermore, EOSC practises (e.g., EOSC Interoperability

---

<sup>26</sup> FAIR Cookbook, <https://faircookbook.elixir-europe.org/content/home.html> [accessed 06.12.2022]

<sup>27</sup> Philippe Rocca-Serra, Wei Gu, Vassilios Ioannidis, Tooba Abbassi Dalooi, Salvador Capella-Gutierrez, Ishwar Chandramouliswaran, Andrea Splendiani, Tony Burdett, Robert T. Giessmann, David Henderson, Dominique Batista, Allyson Lister, Ibrahim Emam, Yojana Gadiya, Lucas Giovanni, Egon Willighagen, Chris Evelo, Alasdair J. G. Gray, Philip Gribbon, ... the FAIR Cookbook Recipes' Authors. (2022). The FAIR Cookbook - the essential resource for and by FAIR doers (1.0). Zenodo. <https://doi.org/10.5281/zenodo.7156792>

framework<sup>28</sup>, EOSC Enhance D4.3 “Analysis of existing research data cataloguing efforts towards integrated discovery”<sup>29</sup>) and RDA recommendations<sup>30</sup> are followed, thanks to the joint work with both initiatives.

The European COVID-19 Data Platform is established as the main portal to ensure accessibility. In addition, metadata to those sensitive data that fall under the GDPR<sup>31</sup> will go in FAIRsharing, and in the IDTk (currently under development) as relevant.

As already highlighted in the first version of the DMP, this project is driving data use and re-use by linking FAIR open data to workflow environments and providing access to analysis and visualisation tools building trust and reproducibility with provenance and quality assurance mechanisms. Thus, the project won't generate any new sensitive patient level data and each participant institution is responsible for accounting for national legislations, the administrative provisions and the implemented data access procedures. Nonetheless, if any new data is generated (especially from WP5 use cases) it will be handled in the next version of the DMP.

## 5. Conclusions

Overall, during BY-COVID, data access committees and processes for each dataset remain accountable for the basis of the project's operations. Currently, no active participant recruitment/engagement or AI systems are used. This new version of the DMP gives a more detailed description of the tools and guidelines followed within the BY-COVID project, thanks to the ongoing work and the already submitted deliverables and milestones<sup>32</sup>. It will be reviewed to ensure it continues to be fit for purpose and that any changes introduced in the repositories or in the process to handle personal data (contact details) are incorporated into the document.

---

<sup>28</sup> EOSC-IF: <https://eosc-portal.eu/eosc-interoperability-framework> [accessed 16.11.2022]

<sup>29</sup> Carole Goble, & Nick Juty. (2021). Analysis of existing research data cataloguing efforts towards integrated discovery. Zenodo. <https://doi.org/10.5281/zenodo.4693217>

<sup>30</sup> RDA COVID-19 Working Group. (2020). RDA COVID-19 Recommendations and Guidelines on Data Sharing (1.0). <https://doi.org/10.15497/rda00052>

<sup>31</sup> General Data Protection Regulation (GDPR), <https://eur-lex.europa.eu/eli/reg/2016/679/oj> [accessed 16.11.2022]

<sup>32</sup> BY-COVID Outcomes, <https://by-covid.org/outcomes/> [accessed 16.11.2022]

## 6. Next steps

Next DMP deliverable is scheduled for month 34 (July 2024) and it will be mainly focused on sustainability. However, the development of the project will be followed and an intermediate version will be generated if needed in the case of major changes in the project's plan.

## 7. Impact

A sound and established Data Management Plan constitutes the foundation for a fruitful collaboration across WPs, with internal and external experts and stakeholders, including RDA, EOSC and FAIRsharing.