Check for updates

SOFTWARE TOOL ARTICLE

REVISED The Network of National COVID-19 Data Portals: public health equity through collaboration

[version 2; peer review: 1 approved, 4 approved with reservations]

Marianna Ventouratou[1], Liane Hughes[2], Katarina Öjefors Stark[2], Johan Rung[2], Natalia Koralewska[3], Marzena Ciechomska[4], Krzysztof Kurowski[5], Cezary Mazurek[5], Eva Alloza[6,7], Alberto Hornos-Vidal[6,7], Salvador Capella-Gutierrez[6,7], Robin Navest[8], Erik Hjerde[9], Aitana Neves[10], Colman O'Cathail[1], Alfonso Valencia[6,7,11], Henning Hermjakob[1], Matt Pearce[1], Manish Kumar[1], Nadim Rahman[1], Carla Cummins[1], Guy Cochrane[1]

[1]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1 SD, UK
[2]SciLifeLab, Stockholm, Stockholm County, Sweden
[3]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznan, Poland
[4]National Institute of Geriatrics, Rheumatology, and Rehabilitation, Warsaw, Poland
[5]Poznan Supercomputing and Networking Center, Poznan, Poland
[6]Barcelona Supercomputing Center, Barcelona, Spain
[7]Spanish National Bioinformatics Institute, Barcelona, Spain
[8]Lygature, Utrecht, Utrecht, The Netherlands
[9]The Arctic University of Norway, Tromso, Norway
[10]Swiss Institute of Bioinformatics, Geneva, Switzerland
[11]ICREA, Institucio Catalana de Recerca i Estudis Avancats, Barcelona, Catalonia, Spain

v2 **First published:** 23 Jul 2025, **5**:192
https://doi.org/10.12688/openreseurope.20059.1
Latest published: 10 Aug 2026, **5**:192
https://doi.org/10.12688/openreseurope.20059.2

Abstract

The network of the national COVID-19 Data Portals was developed and linked to the COVID-19 Data Portal (https://www.covid19dataportal.org/)inresponsetothe need for rapid data sharing and analysis during the 2020–2022 SARS-CoV-2 pandemic. Built on open-source code developed by the Swedish COVID-19 Data Portal (now the Swedish Pathogens Portal, www.pathogens.se) the network included 12 national portals addressing demand for local open data sharing and access, across data types and resources. It provides a robust case study of national initiatives for FAIR (Findable, Accessible, Interoperable and Reusable) resources and a foundation for future pandemic preparedness across pathogens globally. In this paper we outline the structure of the origins of the network of National COVID-19 Datal Portals, the

Open Peer Review

Approval Status ? ? ? ✓ ?

	1	2	3	4	5
version 2 (revision) 10 Aug 2026				✓ view	? view
version 1 23 Jul 2025	? view	? view	? view	? view	

1. **Gopi Battineni**, University of Camerino, Camerino, Italy

2. **Flavia Bernardini**, Universidade Federal Fluminense, Rio de Janeiro, Brazil

technical aspects and code originating from the Swedish Portal and provide an overview of the services and tools offered by each Portal. The paper showcases the process and operation of four Portals: Sweden, Poland, Spain, Norway and The Netherlands. *In this study, we observe that pandemic response greatly benefits from an established infrastructure that can be quickly mobilised, developed and extended. Collaborations and preparation built on solid foundations over several years, supported by investment in the form of national and international research grants, is key for sustainability, continuation and readiness to deploy such efforts.*

Keywords
COVID-19, SARS-CoV-2, FAIR data, COVID-19 Data Platform, COVID-19 Data Portal, National Data Portals, pandemic preparedness, pathogens, infectious diseases, open data

HE This article is included in the Horizon Europe gateway.

H2020 This article is included in the Horizon 2020 gateway.

3. **Muskan kapoor**, Iowa State University, Ames, USA
4. **David Horsfall**, Newcastle University, Newcastle upon Tyne, UK
5. **Chizaram A. Onyeaghala**, University of Port Harcourt Teaching Hospital, Port Harcourt, Nigeria

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Marianna Ventouratou (ventouratou@ebi.ac.uk)

Author roles: Ventouratou M: Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Hughes L**: Conceptualization, Data Curation, Methodology, Resources, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Öjefors Stark K**: Data Curation, Resources, Software; **Rung J**: Conceptualization, Data Curation, Funding Acquisition, Resources, Software; **Koralewska N**: Data Curation, Resources, Software, Writing – Original Draft Preparation; **Ciechomska M**: Data Curation, Resources, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Kurowski K**: Data Curation, Resources, Software, Writing – Original Draft Preparation; **Mazurek C**: Resources, Software, Supervision, Writing – Original Draft Preparation; **Alloza E**: Data Curation, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; **Hornos-Vidal A**: Methodology, Software, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **Capella-Gutierrez S**: Conceptualization, Funding Acquisition, Methodology, Project Administration, Supervision, Writing – Review & Editing; **Navest R**: Software, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Hjerde E**: Conceptualization, Data Curation, Funding Acquisition, Methodology, Resources, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Neves A**: Conceptualization, Data Curation, Funding Acquisition, Methodology, Resources, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **O'Cathail C**: Data Curation, Methodology, Resources, Software, Writing – Original Draft Preparation, Writing – Review & Editing; **Valencia A**: Conceptualization, Funding Acquisition, Methodology, Supervision, Validation, Writing – Review & Editing; **Hermjakob H**: Data Curation, Funding Acquisition, Methodology, Resources, Software, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Pearce M**: Data Curation, Methodology, Resources, Software; **Kumar M**: Data Curation, Software; **Rahman N** : Conceptualization, Data Curation, Methodology, Resources, Software, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Cummins C**: Data Curation, Methodology, Resources, Software, Writing – Review & Editing; **Cochrane G**: Conceptualization, Funding Acquisition, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Horizon Europe and Horizon 2020 research and innovation programmes under the following grant agreements: [grant agreement no. 874735] (project name: Horizon 2020 VEO), [grant agreement no.101046203] (Horizon Europe BY-COVID), [grant agreement no. 871075] (Horizon 2020 RECODID), [grant agreement no. 871075] (Horizon 2020 ELIXIR Converge), [grant agreement no. 824087] (Horizon 2020 EOSC Life), [grant agreement no. 643476] (Horizon 2020 COMPARE) The Spanish COVID-19 Data Portal was supported by the project COV20/01314 "Registro Estatal de COVID-19: The Spanish COVID-19 Data portal" from the Fondo COVID-19 of the Instituto de Salud Carlos III (ISCIII).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Ventouratou M *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Ventouratou M, Hughes L, Öjefors Stark K *et al.* **The Network of National COVID-19 Data Portals: public health equity through collaboration [version 2; peer review: 1 approved, 4 approved with reservations]** Open Research Europe 2026, **5**:192 https://doi.org/10.12688/openreseurope.20059.2

First published: 23 Jul 2025, **5**:192 https://doi.org/10.12688/openreseurope.20059.1

REVISED **Amendments from Version 1**

We thank the reviewers for taking the time to provide detailed comments and an evaluation of our paper. Following the reviewer's suggestions we have made the following changes:

We have added a dedicated **Lessons learnt and limitations** section, which consolidates key insights from the development of the National COVID-19 Data Portals, discusses governance and sustainability challenges, highlights limitations and reflects on the broader implications for future pandemic preparedness and pathogen data infrastructures.

We revised the abstract and introduction to improve readability, clarify the impact of the work, and better articulate the lessons learned from the COVID-19 response. We also clarified the use of the term *network*, emphasising that it represents a collaborative and organisational network rather than a fully interoperable technical infrastructure.

Several technical and organisational aspects have been expanded. These include clarification of the shared and locally customised elements of the portals, explanation of the underlying architecture and FAIR alignment, additional descriptions of participating organisations and their roles, and improved discussion of metadata handling, interoperability, and the evolution towards the Pathogens Portal and the wider European Health Data Space.

Terminology has been refined throughout the manuscript to improve precision, and organisational acronyms have been expanded with explanatory notes and links. We have also revised sections describing the future direction of the portals to provide a more holistic perspective while acknowledging current technical and funding constraints.

Finally, we have incorporated additional references where appropriate, clarified the collection and interpretation of usage metrics, and expanded discussion of community engagement, governance, and sustainability. Collectively, these revisions improve the clarity, balance, and contextualisation of the manuscript while reinforcing its value as a case study of collaborative national and European data infrastructure development during and beyond the COVID-19 pandemic.

No changes were made to the title, abstract, authors list, figures, data and tables.

Any further responses from the reviewers can be found at the end of the article

Introduction

The European COVID-19 Data Platform was established in April 2020 with funding from the European Commission provided through existing Horizon 2020 projects[i] aiming to bring together different services and infrastructures to facilitate sharing of global biodata relevant to SARS-CoV-2. The Platform comprises three components: The COVID-19 Data Portal, the SARS-CoV-2 Data Hubs, and the Federated European Genome Phenome Archive.

The COVID-19 Data Portal (Harrison *et al.*, 2021) presents a wide range of biological and other data from a number of databases. These include viral and host sequences to scientific articles, images, protein networks, many other molecular data, and social sciences data relating to the pandemic. All data, and in particular the viral sequences, saw a rapid increase in volume as the months progressed[ii], currently at ~8.8 million sequences, ~7.1 million raw reads and ~ 5.6 million systematic analyses. The Portal also offers access to analysis tools and visualisations of viral data, and provided monthly variant reports for SARS-CoV-2[iii] (2021–2023). Fostering open access collaboration and FAIR[iv] data sharing, the Portal includes a substantial 'Related Resources' section pointing to external databases, studies and resources related to COVID-19 and covering epidemiology, data standards, computing support, tools and cohort data, many linking back to the FAIRsharing catalogue[v].

Early in 2020, the notion of National Data Portals to present national level information and links to the COVID-19 Data Portal was developed, which resulted in a network of now 12 bespoke national Portals built over three years (Figure 1). In the context of this article, we use the word "network" to describe a coordinated development of various national portals around the central portal, driven by a collaborative and cooperative distributed group of individuals. National portals are linked from the main portal, with some cross-indexing, such as some of the Swedish data[vi] and the Dutch data[vii].

[i] VEO (https://www.veo-europe.eu/), RECODID (https://recodid.eu/), ELIXIR-CONVERGE (https://elixir-europe.org/about-us/how-funded/eu-projects/converge), EOSC-Life (https://www.eosc-life.eu/), BY-COVID (https://by-covid.org/).
[ii] https://www.covid19dataportal.org/statistics
[iii] 'VEO variation analysis reports' in https://www.covid19dataportal.org/related-resources
[iv] Findable, Accessible, Interoperable and Reusable (FAIR).
[v] https://fairsharing.org/databases/
[vi] https://www.covid19dataportal.org/search/national-portals?db=sweden-data-portal
[vii] https://www.covid19dataportal.org/search/research-infrastructures?crossReferencesOption=all&overrideDefaultDomain=true&db=covid-nl-ri&sortignorenull=true&size=15

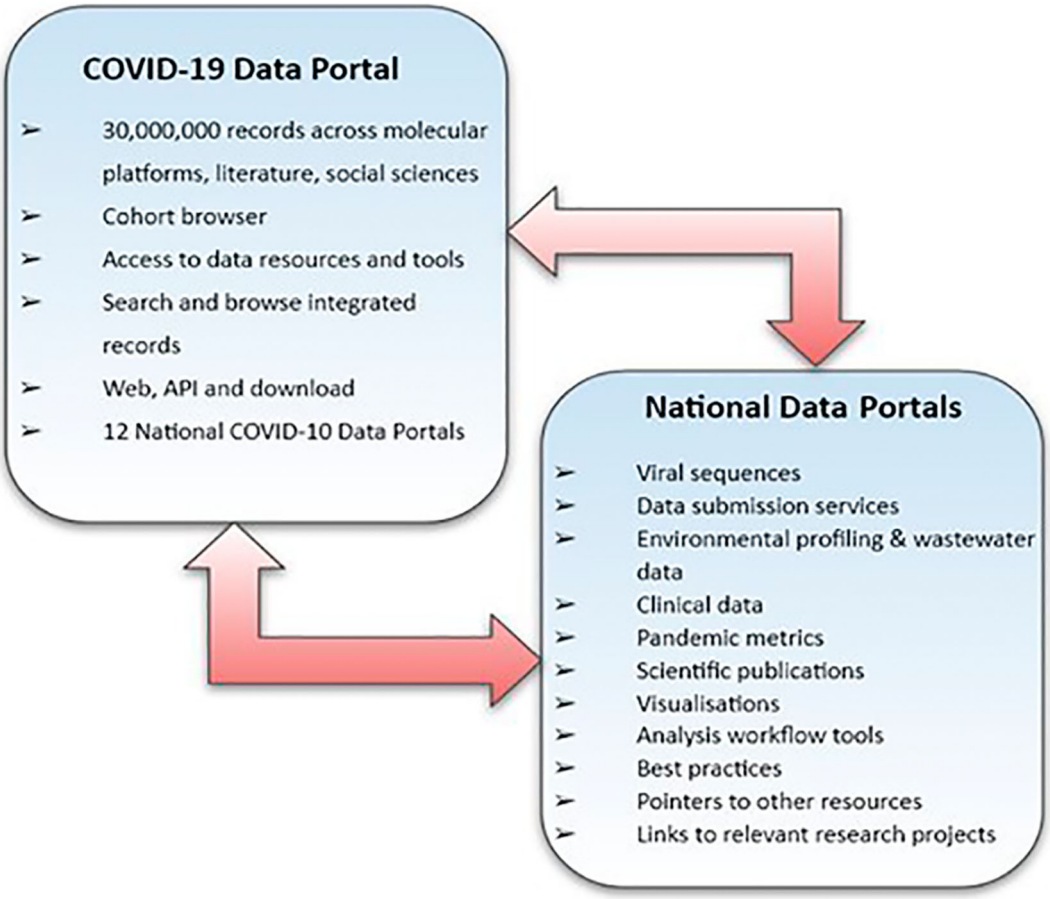

Figure 1. The COVID-19 Data Portal and the National Data Portals: resources and services.

This article describes the evolution of the network, why and how the Portals were built, what they offer to the scientific community, and which technologies were used. It also aims to capture the lessons learnt from this project and cover plans for the future. Finally, it presents detailed country-specific examples from Sweden, Poland, Spain, Norway and The Netherlands.

Review of the National COVID-19 Data Portals

The case for National COVID-19 Data Portals

The National Data Portals provide an important complement to the COVID-19 Data Portal. They offer the scientific community accessibility, visibility and discoverability of national and international data. They also enable further understanding of transmission, infection and symptoms of SARS-CoV-2, and thus aid in the development of treatments, vaccines and public health responses. The National Data Portals offered another layer of visibility for the European COVID-19 Data Platform, which is interlinked through them. The National Data Portals also presented country-specific data and resources in the national language, as well as national health data, epidemiological statistics and links to national studies and other resources; serving thus as an entry point into national resources with appropriate global context.

Organisation

The network of National Data Portals underlines the interconnectedness of the data across national boundaries and makes the case for data sharing and openness along FAIR principles. It emerged naturally through early, systematic engagement of the COVID-19 Data Platform team at EMBL-EBI in 2020 with the so-called “national coordinators” group, an *ad hoc*

community of stakeholders drawn from ELIXIR[viii] and EMBL member states[ix] and other countries. The national coordination meetings centred on topics around viral data management, brokering, sensitive data management, and the wider, as well as national, clinical and epidemiological context during the early months of the pandemic. This model worked well whilst all institutions redeployed their resources to concentrate on efforts focused on COVID-19, and the world was connected online. In the latter years, 2022–2023, coordination was more challenging as staff resumed their normal posts and responsibilities, the maintenance of COVID-19 Portals becoming a lesser priority for some of the partners, and in some cases seizing all together.

The 12 COVID-19 National Data Portals were configured partly by ELIXIR Nodes[x] and EMBL Member States with the support of EMBL-EBI, which provided the Portal's graphic templates, and coordinated wider publicity and engagement efforts.

Within the first year of the operation of the COVID-19 Data Platform, and until April 2021, Sweden, Poland, Spain, Norway, Italy, Japan, Slovenia and Turkey had set up their national portals. These were followed by Estonia and Greece in autumn 2021. Luxembourg and The Netherlands were added in February and October 2023, respectively (Figure 2).

Technical aspects

The first national COVID-19 Data Portal was developed by Sweden. Development started in April 2020, and the code was first made available on GitHub (Hughes *et al.*, 2024b) in May 2020 under the MIT licence Open Source Initiative[xi]. The Swedish COVID-19 Data Portal[xii] (SCR_024866) (now known as the Swedish Pathogens Portal, www.pathogens.se) was subsequently launched in June 2020. The portal was specifically designed to enable a quick launch. To facilitate this, the code was written in accessible formats (e.g. markdown), an open source static site generator (Jain, 2022) was used to build the website, and hosting was configured so as not to require a database and to be effected using any platform. The code has since been reused multiple times to implement other national COVID-19 Portals, and other types of websites. Dynamic elements have been added to the code to ensure that the static nature of the site did not prevent frequent data updates. For example, dynamic data visualisations were produced using Python to enable data to be updated in near real-time (code available on GitHub[xiii] for reuse under the MIT licence Open Source Initiative[xiv]). Dynamic elements also enable linkage to the European and other national portals[xv]. More information about the technical setup for the Swedish Portal can be found in Hughes *et al.* (2024a).

National Data Portals features

Overview

The data displayed on the COVID-19 Data Portal were added based on user requirements and an increasing variety of SARS-CoV-2 related data submissions, first as an urgent response, and eventually as an established project acting as a pilot for future pandemic preparedness. Data and services offered by most National Data Portals set up by January 2022 were diverse (Table 1). Most portals presented viral sequences, related publications, active research projects and pointers to other resources. A small number of portals offered data submission services, analysis workflow tools and pandemic metrics. Other portals included environmental profiling, including wastewater data, clinical data, visualisations and best practices.

This variability in the services and features offered nationally reflects the diverse response to national mandates across Europe while retaining interoperability. This model is strong in that it allowed national frameworks to focus on data and projects that mattered on a national level, while still being deeply connected to the more comprehensive European wide portal. Mobilising, collecting and presenting information on a national level, while a burden on national data providers and coordinators amidst the peaks of the pandemic in 2020 and 2021, demonstrates the resilience of the model to be replicated for pandemic preparedness response.

[viii] "ELIXIR is an intergovernmental organisation that brings together life science resources from across Europe. These resources include databases, software tools, training materials, cloud storage and supercomputers" from https://elixir-europe.org/about-us. Moreinformationcan be found here: https://elixir-europe.org/

[ix] The European Molecular Biology Laboratory is an intergovernmental organisation supported by 30 countries, member states. More information can be found here: https://www.embl.org/about/member-states/. ELIXIR is the network (of people) around Europe that provide expertise and infrastructure; EMBL is the research organisation that includes EMBL-EBI, a data service provider and one of the nodes of ELIXIR. That Sci Life Lab hosts another of the ELIXIR nodes

[x] More information on ELIXIR nodes can be found here: https://elixir-europe.org/about-us/who-we-are/nodes

[xi] https://opensource.org/license/mit

[xii] The RRID for the Swedish Pathogens Portal is SCR_024866

[xiii] https://github.com/ScilifelabDataCentre/pathogens-portal-visualisations

[xiv] https://opensource.org/license/mit

[xv] https://github.com/ScilifelabDataCentre/pathogens-portal-scripts/tree/main/EBI_indexing

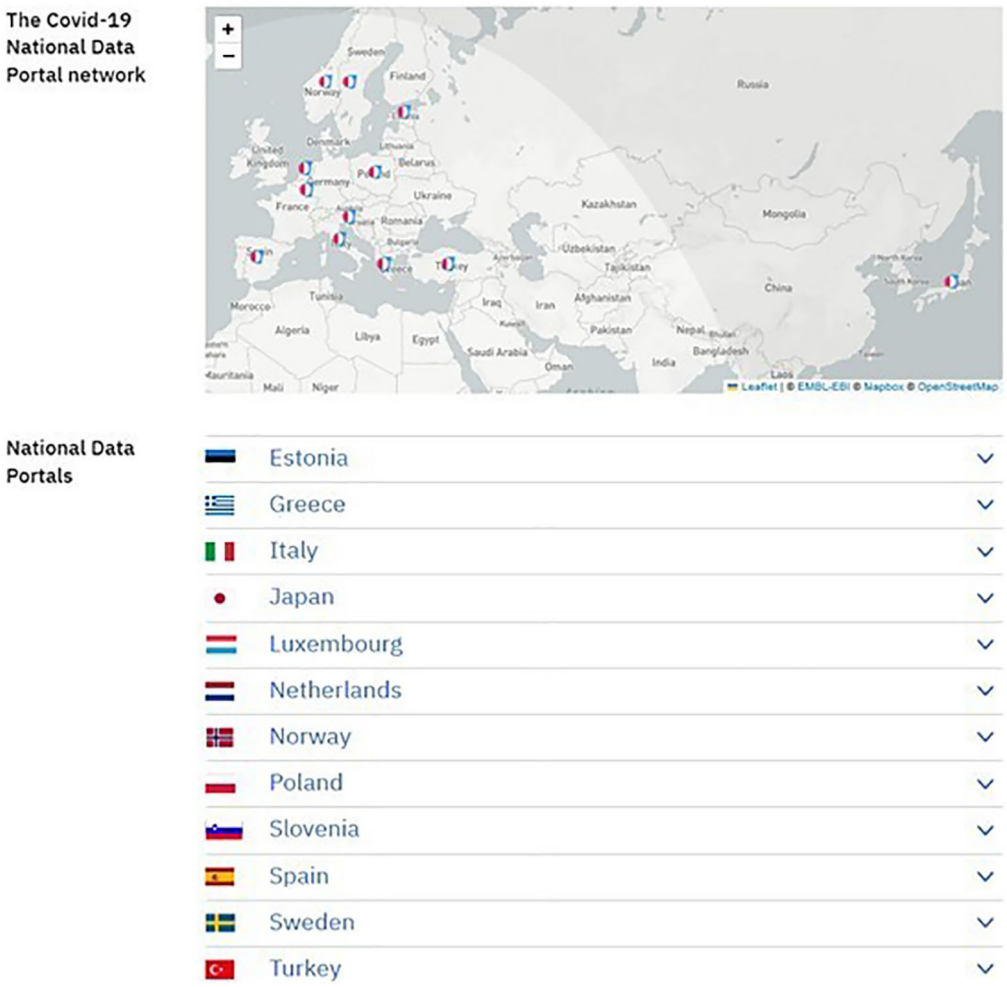

Figure 2. A screenshot of the map and list of National Data Portals on the COVID-19 Data Portal. Source: www.covid19dataportal.org.

Table 1. Overview of the features of the original ten national data portals as of January 2022, and The Netherlands as of October 2023.

	Countries										
Features/Services	**EE**	**GR**	**IT**	**JP**	**NO**	**PL**	**SI**	**ES**	**SE**	**TR**	**NL**
Viral sequences and pointers to viral sequences		✓	✓	✓	✓	✓		✓	✓		
Data submission services	✓			✓	✓	✓	✓		✓		✓
Environmental profiling/Wastewater data	✓	✓					✓		✓		✓
Clinical data								✓			
Pandemic metrics						✓	✓	✓	✓	✓	✓
Scientific publications	✓	✓		✓	✓		✓	✓	✓	✓	✓
Visualisations						✓		✓	✓	✓	
Analysis workflow tools		✓				✓		✓			
Vaccination information						✓	✓		✓		✓
Best practices		✓	✓		✓			✓	✓		✓
Pointers to other resources	✓		✓	✓		✓	✓	✓	✓	✓	✓
Research projects	✓		✓			✓	✓	✓	✓		

Countries: EE-Estonia, GR-Greece, IT-Italy, JP-Japan, NO-Norway, PL-Poland, SI–Slovenia, ES-Spain, SE-Sweden, TR-Turkey, NL-The Netherlands.

Country inputs and descriptions of individual Portals

Sweden. In April 2020, SciLifeLab[xvi] received an assignment from the Swedish government to accelerate COVID-19 research by helping mobilise and promote relevant data from Sweden. To this end, Sweden launched the first national COVID-19 Data Portal in June 2020 with the express aim of making COVID-19 data more open and FAIR. The portal was developed with the express requirement that it would be quick to launch. This was crucial in ensuring that the portal could be established as a central resource around which the community could aggregate in an emergency situation, such as a global pandemic (Hughes *et al*., 2024a). To this end, the Swedish Portal was initially developed using volunteers from the Swedish research community, and the code was made openly available to enable other, similar portals to launch quickly (see technical set up section). A dedicated team, composed of individuals with relevant technical expertise and backgrounds in life sciences, was established in Autumn 2024.

The Swedish Portal was embedded in the Swedish research community immediately upon launch. Initial funding for the project was provided by The Swedish Research Council (SRC, Vetenskapsrådet) as part of their wider efforts related to COVID-19. The Swedish Research Council directed other projects in receipt of funding as part of those efforts to contact the Swedish Portal for advice on data management and with help in disseminating their work. SciLifeLab also directed those within their community that were working with COVID-19 to collaborate closely with the portal. This led to close collaborations with the research community, with the portal providing direct support to over 250 research efforts within the first two years. This included work with national organisations (e.g. Biobanks Sweden), national infrastructure (e.g. National Pandemic Centre, and the SciLifeLab Autoimmunity and Serology Profiling unit), and many research groups.

Feedback from the community was sought continuously to ensure that the content of the Swedish Portal met the rapidly evolving needs of the Swedish research community. Priority was given to resources that were considered most broadly useful to the research community, novel research methods that progressed a field, and/or where the resources were not available elsewhere. This led to the development of unique content types. The most popular of these were data highlights (short articles promoting research that shares data, code, and/or tools as openly as possible), data dashboards (pages with descriptions of projects and dynamic data visualisations that are regularly updated with new data), and databases with information about accessing biobanks samples, available research data, and COVID-19 publications from Sweden. As a result, the portal became very popular; receiving over 70,000 unique visitors and 230,000 pageviews in the first two years. Though visitors were primarily from Sweden, there were also frequent visits from countries all around the world.

In order to continue to meet the needs of the Swedish research community, the scope of the portal was expanded in 2022 to reflect a shift in focus towards general pandemic preparedness. The portal became a central pillar of SciLifeLab's Pandemic Laboratory Preparedness (PLP) program in 2022. The portal has since transitioned to become the Swedish Pathogens Portal, making it the first national Pathogens Portal[xvii]. The aims of the portal are largely unchanged, with the central aim being to make data as FAIR and open as possible.

Poland. At the beginning of April 2020, the Ministry of Science and Higher Education designated the Institute of Bioorganic Chemistry, Polish Academy of Sciences (IBCH PAS) together with the affiliated Poznan Supercomputing and Networking Centre (PSNC) as the national hub for the European COVID-19 Data Platform. IBCH PAS was selected for this role due to its extensive expertise in virology, population genomics, and next-generation sequencing, along with PSNC's advanced supercomputing and networking infrastructure, which enables it to handle large volumes of data while ensuring high standards of data security and integrity. Notably, PSNC manages a robust Polish National Research and Education Network PIONIER, that links all Polish scientific and medical centres with supercomputing and data storage facilities, as well as provides connectivity to the European research network GÉANT, facilitating seamless collaboration and data exchange within the country and with other European entities involved in the COVID-19 Data Platform. Additionally, IBCH PAS and PSNC have strong connections with national and European research networks and extensive experience in large-scale infrastructure projects and partnerships (e.g. EOSC, EuroHPC, PRACE, EUDAT, etc.). The core mission of PSNC, which strengthened the development of the hub, is to foster scientific excellence by providing reliable and cutting-edge e-Infrastructure and highly-specialised laboratories. PSNC achieves this goal, particularly through extensive R&D activities related to information and communication technologies and their innovative applications. In collaborative project teams, they draw from the synergy of science, technology, business, and education to tackle the greatest challenges for our future.

[xvi] https://www.scilifelab.se/
[xvii] https://www.pathogens.se/

Having such extensive expertise and advanced infrastructure, IBCH PAS and PSNC were able to launch the COVID-19 Data Portal Poland[xviii], also known as COVID-HUB-PL, within weeks of the Ministry's decision. The initiative was carried out *pro bono* by a dedicated team of up to 15 people at the peak of the pandemic.

The main aim of this initiative was to facilitate open data sharing and analysis, accelerating research on the coronavirus. The genomic data available through the COVID-19 Data Portal Poland were generated in collaboration with eight healthcare or science centres from Poland. These centres either deposited sequencing data directly or collected biological material sequenced at IBCH PAS. To ensure data interoperability and quality, all metadata were subjected to rigorous schema checks. Contributors used standardized templates, which were manually and semi-automatically validated by the COVID-HUB-PL team prior to integration. The collected genomic data and the Poland-focused Nextstrain COVID-19[xix] instance supervised by COVID-HUB-PL team enabled researchers to monitor the spread and evolution of the virus in real-time, providing insights into how different variants emerged and spread geographically. These studies were essential for informing public health strategies, including lockdown measures. Prioritising viral genomics within the COVID-19 Data Portal Poland ensured that the scientific community and central and local government agencies had access to the most relevant and actionable data to combat the pandemic effectively. Indeed, the aggregated COVID-19 statistics available on the COVID-HUB-PL platform as interactive reports for the country's epidemiological situation enhanced transparency, communication, and data-driven decision-making.

In Poland, efforts to mobilise data regarding COVID-19 have encountered numerous challenges, particularly regarding coordination between different ministries (i.e. Ministry of Science and Higher Education overseeing research and scientific institutions, and Ministry of Health overseeing healthcare and medical centers), and legislative obstacles. Various agencies and institutions involved in COVID-19 response efforts maintained separate data silos, complicating efforts to integrate and harmonise data for analysis. The lack of transparent legislation related to data sharing and privacy constrained the willingness to share viral genomic data through the unrestricted data governance model endorsed by the European COVID-19 Data Platform. Furthermore, the COVID-19 Data Portal Poland was created and continued to operate without additional financial support, using IBCH PAS and PSNC own resources and infrastructure.

Additionally, the REGIONAL COVID-HUB[xx] platform together with extended e-services have been successfully created thanks to the Wielkopolska Voivodeship Government support and regional WRPO 2014+ programme funding. New and added-value services enabled secure communication and customization for users and public institutions involved in the pandemic response at the Wielkopolska regional level. New capabilities exposed to local users supported advanced visualisation of the pandemic's status and dynamics in Wielkopolska, training and open educational resources for schools as well as functionalities for processing, analysing, and sharing scientific data, and supporting COVID-19 diagnostics using AI methods (Figure 3).

In the period from 2021 to 2023, the COVID-19 Data Portal Poland had a total of almost 230,000 page views by over 47,000 unique visitors. As of December 2021, the further development of infrastructure and services for the COVID-19 Data Portal Poland has been synchronised with standardisation and best practice exchange activities initiated by the EU ELIXIR-CONVERGE project.

Spain[xxi]. In the early months of the pandemic, the nodes of the Spanish National Bioinformatics Institute (INB), which represents the Spanish node of ELIXIR (ELIXIR-ES), and the Translational Bioinformatics Network (TransBioNet), held regular meetings to join forces and share knowledge to contribute during the initial uncertainty of the COVID-19 pandemic. One of the actions was the publication in July 2020 of a curated collection of databases, repositories, and tools that were being developed and adapted, as well as publications, projects, and consortia in which they were involved. In parallel, a European effort began through the coordinated work of EMBL-EBI and ELIXIR to create a European portal for the data sharing and analysis of COVID-19 data, the European COVID-19 Data Portal. Spain joined the initiative of tailored national portals after openly publishing the Swedish portal repository.

INB/ELIXIR-ES, as the bioinformatics technology platform of ISCIII in 2020, received a mandate to develop the Spanish portal from ISCIII, the National funding body for Health Research in Spain. This portal was created in the context of the programme *Fondo COVID-19*, an ISCIII emergency funding action to contribute to improving the diagnosis and clinical

[xviii]COVID-19 Data Portal Poland 2020: https://covidhub.psnc.pl/eng/
[xix]Genomic epidemiology of hCoV-19 in Poland 2020: https://nextstrain.covidhub.psnc.pl/ncov/poland
[xx]REGIONAL COVID-HUB Wielkopolska 2020: https://wlkp.covidhub.pl/
[xxi]*The Spanish COVID-19 Data Portal was supported by the project COV20/01314 "Registro Estatal de COVID-19: The Spanish COVID-19 Data portal" from the Fondo COVID-19 of the Instituto de Salud Carlos III (ISCIII).*

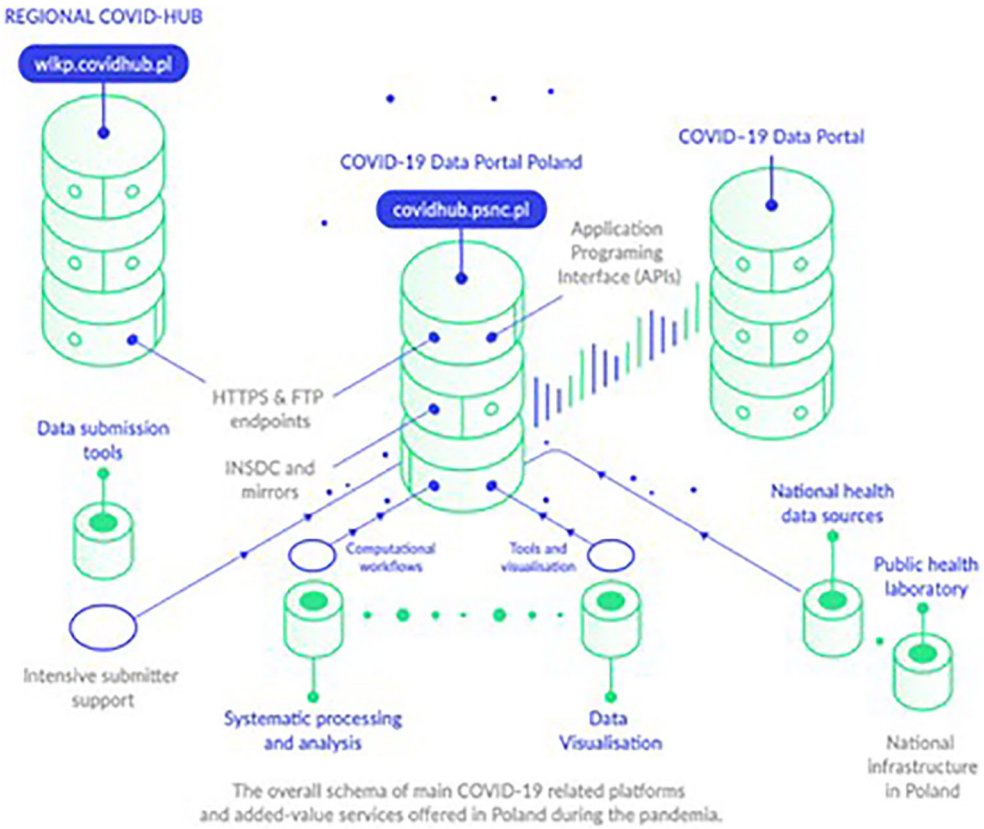

Figure 3. The overall schema of the main COVID-19 related platforms and added value services offered in Poland during the pandemic.

management of patients infected with SARS-COV-2, the treatment of the disease in the COVID-19 pandemic also contribute to the design, development, and implementation of public health measures to effectively respond to the ongoing SARS-CoV-2 epidemic.

The Spanish COVID-19 Data Portal[xxii] showcases Spanish research efforts and technological developments related to SARS-CoV-2 and COVID-19 within the context of the INB/ELIXIR-ES and TransBioNet, and ensures interoperability with other European and international efforts. Thus, the portal highlights a curated compilation of new or adapted data management and analysis tools developed by Spanish researchers. Data and other resources from global repositories are available in the portal, filtering previously Spain as a country and/or affiliation.

Two unique resources were key in decisive moments of the pandemic by offering interactive COVID-19 incidence map and SARS-CoV-2 variant distribution. Firstly, Flow-Maps[xxiii] (https://doi.org/10.1038/s41597-021-01093-5), a system for monitoring COVID-19 outbreaks and mobility-associated risk by integrating health information and population-level mobility patterns into a Geographical Information System. Secondly, COVID-19 Viral Beacon, is a tool to find SARS-CoV-2 variability at genomic, amino acid, and motif levels along with other metadata. This implementation was based on the GA4GH Beacon standard and showcased a particular use case for testing and demonstrating the new features of Beacon v2.

In the context of the *Fondo COVID-19*, a significant effort was made to access, normalise and harmonise - following the OMOP Common Data Model – the data generated for the data-driven projects in the programme. An initial set of 60 projects was identified, although the standardised effort was finally carried out for 13 projects after establishing a general procedure for translating individual projects' codebooks into unambiguous clinical variables, ideally following

[xxii] https://www.covid19dataportal.es/
[xxiii] https://flowmaps.life.bsc.es

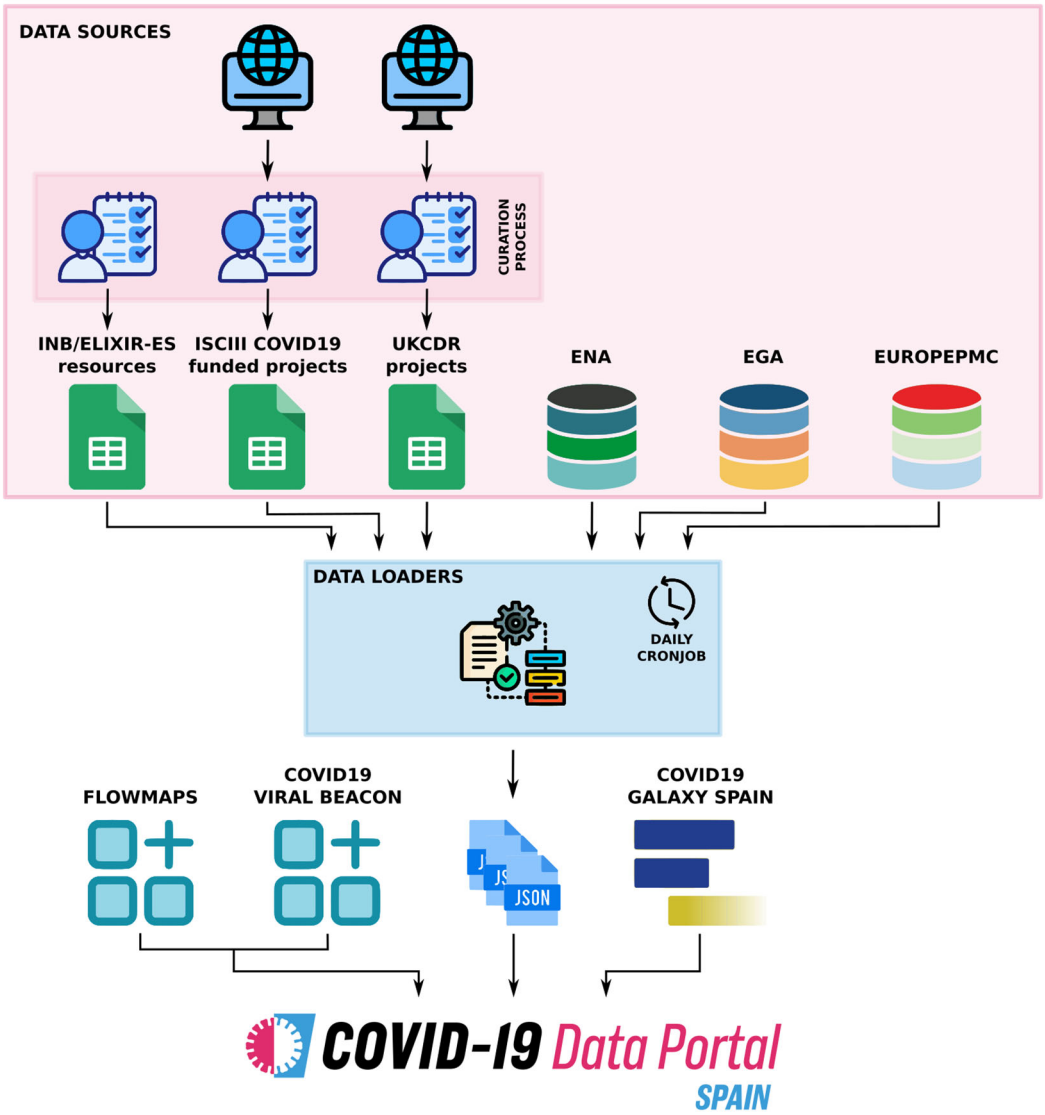

Figure 4. Data flow and presentation schema on the Spanish COVID-19 Data Portal.

established ontologies. When unambiguous assignment of variables was not possible, dedicated sessions were carried out with the projects' data managers. In the rare cases that disambiguation was not possible, concepts were left out of the overall standardisation effort. Once a project was normalised, an overall harmonisation step was performed to ensure general consistency across all datasets provided. Fondo COVID-19 was highlighted by adding filters in relevant sections, e.g. research projects and literature (Figure 4).

The project was fast developed in three months by a part-time senior full-stack developer and a part-time community engagement and content creation manager. The portal is based on the Swedish portal's open-source code. The first release of the website was launched in December 2020, and it has been in a stable maintenance phase since then, with most of the data updates carried out automatically and certain progressive developments, when necessary, e.g., adding new resources. The technological development of the platform was not a challenge, but the strategic side of it when curating the most significant resources to showcase on our portal.

The Spanish COVID-19 Data Portal has tracked statistics with Google Analytics since March 2021. From March 1st, 2021, until March 1st, 2024, the portal had 5,035 unique visitors, with 31% of the users connecting from Spain, 13.6% from the USA and 6% from the UK. There were 15,279 pages viewed, and 9,722 users engaged. The most visited pages after the home were the literature and genomics sections, followed by the resources page.

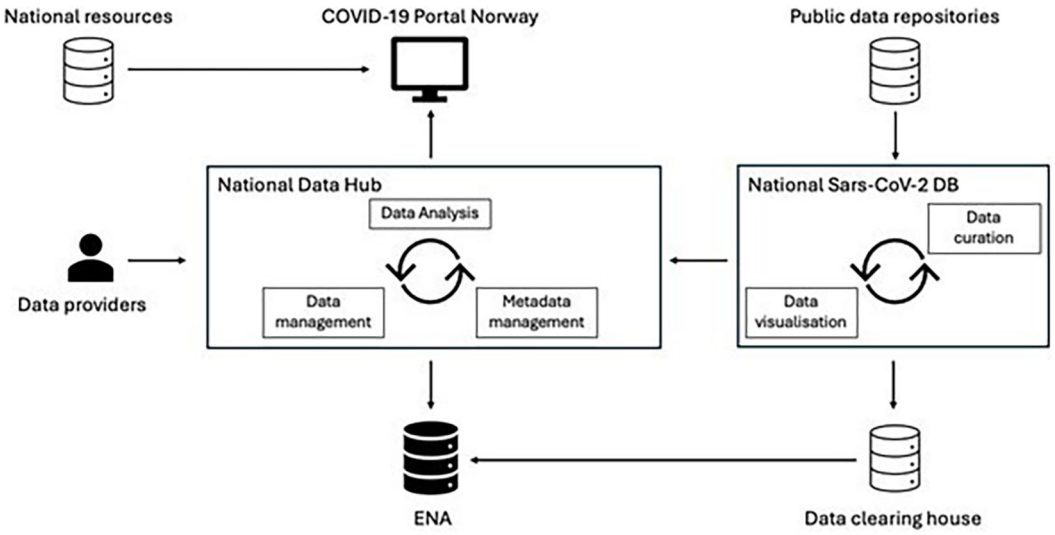

Figure 5. The organisation of data flow and tools and services provided through the Norwegian COVID-19 National Data Portal.

Norway. ELIXIR Norway plays a central national role in giving support on data management of life science data. This includes managing, analysing and sharing sequences, which became essential during the COVID-19 pandemic. The national COVID-19 portal was set up to help and guide researchers and other interested parties. The portal highlights bioinformatics resources such as data sets and analysis methods that were relevant to the situation (Figure 5). It also is a contact point into the national ELIXIR support desk that could assist researchers wanting to use bioinformatics resources or needed support and advice.

Together with the National Institute of Public Health, the national SARS-CoV-2 data providers were mapped. Through the ELIXIR CONVERGE project, agreements on data quality and contextual metadata standards were implemented nationally. The content of the portal and linking relevant resources was mainly done by ELIXIR Norway personnel. Due to the urgency of the situation at the time and limited funding, no editorial steering group was established.

The portal was set up within months as we based the technology and layout on already existing and open shared source code from other portals in Europe. The Norwegian portal has since transitioned to a pathogen portal[xxiv] (please see the next section).

The Netherlands. The aim to set up the Dutch national COVID-19 Data Portal[xxv] was to support researchers and health care professionals with tools and services in their search for ways to overcome the pandemic and its health consequences. The coordination of this national COVID-19 Data Portal was led by Lygature. Other national stakeholders, such as the different university medical centres, Health-RI, RIVM, KNAW/DANS and ODISSEI were involved for specific input and to review content. As there was already a European portal which brings together relevant COVID-19 datasets and tools, we focussed on national resources that are not available through the European platform and thus augment this resource.

All data and resources related to COVID-19 were included with the idea that anything that could be relevant should be available to a researcher or health care professional. To organise this, the Dutch national COVID-19 Data Portal was developed in synergy with the Dutch national resources page on the Infectious Disease Toolkit (IDTk)[xxvi]. It was decided to collect information on relevant national health authorities, initiatives and resources on the IDTk national page and focus the portal on COVID-19 relevant dashboards, such as wastewater surveillance, vaccine administration, infection radar and pathogen surveillance.

[xxiv] https://pathogens.no/
[xxv] https://www.covid19dataportal.nl/
[xxvi] https://www.infectious-diseases-toolkit.org/national-resources/netherlands

With the Github example provided by the Swedish portal team, it was easy to create a first version of the portal within a couple of weeks by one person working on it part- *time*. The sustainability of the portal has been the largest challenge. Currently we have ensured a long-term solution for hosting of the portal, but discussions about the maintenance after the BY-COVID project ends are ongoing.

In the period 1 October 2023 until 1 May 2024 the Dutch national COVID-19 Data Portal was visited 4,796 times by 2,040 unique user agents. In this period the main (3,118 visits), contact (251 visits) and about (121 visits) pages were most visited. The vaccine administration in the Netherlands (110 visits) was the most visited data dashboard and imaging data (99 visits) is the most visited data type page.

The Dutch pathogens portal https://www.pathogensportal.nl/ has replaced the national COVID-19 Data Portal. Furthermore, to support stakeholders in finding and gaining access to COVID-19 related datasets, all COVID-19 datasets have been included in the national metadata catalogue https://catalogus.healthdata.nl/?q=COVID.

Where next with the national portals?

With pandemic preparedness a pressing priority for all countries, the network of National Data Portals is evolving into a network of more generalist national Pathogens Data Portals – treating data and services across broader diseases and pathogens. Receiving a combination of national funding and support from the BY-COVID project[xxvii], the national Portals are following the path of the Pathogens Portal (link) that has emerged from the COVID-19 Data Portal. Establishing a network of national Pathogens Data Portals will enable access to biodata for pandemic preparedness, with a mixture of pointers to the central Pathogens Portal and own national and regional resources.

The Swedish Pathogens Portal was the first national node, launching just one month after the Pathogens Portal, in August 2023. The existing code base for the Swedish COVID-19 Data Portal was used as a basis and modified to e.g. change the visual identity of the site and add new features related to the expanded scope. The code is available on GitHub[xxviii] under an MIT licence; not only allowing others to reuse the code to create other national Pathogens Portal but also enabling existing national COVID-19 Portals to transition, should they choose to do so.

Norway and Switzerland have also received funding from the BY-COVID project to set up National Pathogens Portals:

- SIB Swiss Institute of Bioinformatics is currently establishing the Swiss Pathogens Portal to serve the research community and provide detailed information on the available resources, data and dashboards in the field of pathogen bioinformatics and molecular epidemiology. After investigating and exchanging with the SciLifeLab Data Centre team, it was decided to re-use the Swedish Pathogens Portal code base, to also be able to seamlessly benefit (mutually) from future improvements and developments. A first version of the Swiss portal[xxix] was released in September 2024.

- ELIXIR Norway, represented by The Arctic University of Norway, UiT established the Norwegian Pathogen Portal[xxx] in September 2024. Aiming to connect the portal to a data hub in the future, the team decided to use SvelteKit to develop the web portal. Some parts of the web pages would be rendered via Blazor, while more dynamic pages like dashboards would be rendered by a front-end library like Reacta and SvelteKit. For consistency with the other pathogen portals, we are choosing a design and layout like the Swedish Pathogens Portal. Currently we are mapping the national landscape and are creating content that will become a resource for stakeholders involved in pathogen research and surveillance. The first version was launched September 2024, and there is a pending application with the national research council for funding to extend this work.

A global consortium funded by the NIH NIAID[xxxi] (award U24AI183840) was created to continue and expand the collaboration between the national and central pathogens portals. As part of this effort, SciLifeLab and SIB Swiss Institute of Bioinformatics produced the Pathogen Portal Toolkit. This Toolkit provides a base code for Pathogens Portals in a similar way to the original code related to COVID-19 Portals. Local pathogens portals (LPPs) were provided with seed money to develop a portal, and a Community of Practice was established to ensure greater collaboration and consistency. The project will also enable better interfacing of the LPPs to the central Pathogens Portal and defining new features

[xxvii] https://by-covid.org/
[xxviii] https://zenodo.org/records/11235062
[xxix] https://pathogensportal.ch/
[xxx] https://pathogens.no/
[xxxi] https://pathogendatanetwork.org/

bottom-up, driven by the community. The project is motivated by the acknowledgement that accessing and sharing infectious disease data globally is crucial for pandemic preparedness, but obstacles, such as lack of trust and siloed data, can be overcome through local support for analysis, data management and sharing services which are tailored to the local needs.

Lessons learnt and limitations

In this study, we conclude that successful pandemic response relies on the existence of established infrastructure that can be quickly mobilised, developed and tailored further. Collaborations and preparation built on solid foundations over several years, supported by investment, in the form of national and international funding, is key for sustainability, continuation and readiness to redeploy such efforts.

A further lesson can be learnt from our study: usage and engagement metrics are required to ensure best delivery of the most relevant services and tools. The pressures of rapid turnaround times for the COVID-19 portals during the pandemic meant that several portals did not have available capacity systematically to track and/or report usage statistics. Few portals could provide metrics over a consistent time window appropriate for comparison, so we have not reported these data here.

In terms of governance limitations: This collaborative model functioned well whilst all institutions focused their resources on COVID-19; in latter years, 2022–2023, coordination was more challenging as staff had resumed their normal posts and responsibilities and the maintenance of COVID-19 Portals became a lower priority.

In terms of sustainability, just as funding and collaborations strengthen pandemic preparedness, the presence of a long-term and existing (rather than a purely reactive) plan will be key for future efforts. This means, come the next public health emergency, it will be is important to have in place a framework for re-deploying resources, and pathways for implementing new components as necessary.

Conclusions

This article describes the development of the network of COVID-19 National Data Portals that emerged during the pandemic from 2020 to 2022. It highlights the efforts of specific countries, such as Sweden, Poland, Spain, Norway and The Netherlands. The national portals were all set up to link back to the main COVID-19 Data Portal, a component of the European COVID-19 Data Platform. The portals thus increased the visibility, findability and accessibility of national and international data. The data and resources, covering local and national needs primarily, presented on the portals ranged from viral sequences, data submission services, pandemic metrics, publications and best practices, among others.

In the past year, several of the portals expanded to cover pathogens in general, going beyond SARS-CoV-2 related data, thus creating a springboard for a network of Pathogens Portals. This network will be supported globally by the newly funded NIH project, Pathogen Data Network (award U24AI183840) which has the ambition to create a network of local pathogens portals (LPPs).

The network of COVID–19 national data portals is a testament to the swift and timely efforts of the scientific community to make pandemic-related data available and findable globally. It is also the foundation for any future efforts around outbreak preparedness.

Data availability statement

No data is associated directly with this paper. All datasets supporting the National COVID-19 Data Portals and the COVID-19 Data Portal are available without restrictions from public data repositories, listed here: https://www.covid19dataportal.org/database-resources.

Software availability

Software developed is described in the article is available from the following:

Source code available from:

https://github.com/ScilifelabDataCentre/pathogens-portal; 10.5281/zenodo.10629602

https://github.com/ScilifelabDataCentre/pathogens-portal-visualisations

https://github.com/ScilifelabDataCentre/pathogens-portal-scripts/tree/main/EBI_indexing

Archive software available from: https://zenodo.org/records/15489081

Licenses:

https://opensource.org/license/mit

https://github.com/ScilifelabDataCentre/pathogens-portal?tab=MIT-1-ov-file#readme.

The RRID for the Swedish Pathogens Portal is SCR_024866

References

Harrison PW, Lopez R, Rahman N, *et al.*: **The COVID-19 data portal: accelerating SARS-CoV-2 and COVID-19 research through rapid open access data sharing.** *Nucleic Acids Res.* 2021; **49**(W1): W619–W623.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Hughes L, Stark KÖ, Kochari A, *et al.*: **Building and running an infectious disease portal in the midst of a pandemic: lessons learnt for future pandemics.** 2024a.
Publisher Full Text

Hughes L, Stark KÖ, Kochari A, *et al.*: **Swedish Pathogens Portal (v2.1.18).** *Zenodo.* 2024b.
Publisher Full Text

Jain A: *Hugo in action: static sites and dynamic Jamstack apps.* Manning Publications; 2022.

Open Peer Review

Current Peer Review Status: ? ? ? ✓ ?

Version 2

Reviewer Report 24 August 2026

https://doi.org/10.21956/openreseurope.25296.r79142

© **2026 Onyeaghala C.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Chizaram A. Onyeaghala
University of Port Harcourt Teaching Hospital, Port Harcourt, Nigeria

My recommendation
Recommendation: Approved with reservations — minor/moderate revisions required before indexing.

The remaining issues I would raise
1. The paper needs a clearer methodological identity.
It currently says “In this study, we observe…” but it is not clear whether this is:

- a descriptive case study;
- a technical implementation report;
- a network/infrastructure evaluation;
- a qualitative case study; or
- a narrative description of the portals.

The authors should explicitly state the study/design and explain how the information about the 12 portals and the five showcased portals was collected and analysed.
This is probably the most important remaining issue.

2. The title makes a stronger equity claim than the evidence appears to demonstrate.
The Network of National COVID-19 Data Portals: public health equity through collaboration
The manuscript appears to demonstrate collaboration, data sharing, infrastructure development, and FAIR principles much more directly than it demonstrates *public health equity*.
The authors should either:

- provide evidence demonstrating an equity benefit, such as improved access to data among underserved researchers, institutions or countries; or
- moderate the title and conclusions.

For example:
The Network of National COVID-19 Data Portals: collaborative infrastructure for equitable access to public health data

would be more defensible if equity is an important conceptual objective.

3. The relationship between the 12-portals network and the five showcased portals needs to be clearer.
The abstract says the network included 12 national portals, but then says the paper showcases Sweden, Poland, Spain, Norway and The Netherlands.
The manuscript should explain:

- why these five were selected;
- whether they are representative;
- what criteria were used for selection;
- what information was available for the remaining seven portals.

Otherwise, the five examples could inadvertently give the impression that they represent the entire network.

4. The contribution needs to go beyond description.
This is an important point for indexing.
The paper should answer more explicitly:
What does this case study teach us that was not already known about pandemic data infrastructure?
The strongest contribution appears to be that pre-existing national infrastructure, interoperable technical architecture, institutional collaboration and sustained investment allowed the portals to be rapidly mobilised during COVID-19.
That should be elevated as the central analytical finding rather than allowing the paper to become primarily a catalogue of portal features.

5. “FAIR” should be demonstrated rather than simply invoked.
The manuscript refers to FAIR resources, but the authors should explain how the portals actually supported:

- Findability;
- Accessibility;
- Interoperability; and
- Reusability.

A concise comparative table would be very useful here.

6. Sustainability claims need more evidence.
The conclusion that:
“Collaborations and preparation built on solid foundations over several years… is key for sustainability…”
is plausible and probably correct, but the manuscript should distinguish between:

- what the authors observed;
- what they infer; and
- what they recommend for future preparedness.

In particular, the authors should address what happened to the portals after the acute COVID-19 period. Were they maintained, transformed, integrated into broader pathogen-data infrastructure, or discontinued?

That would substantially strengthen the preparedness argument.

7. The manuscript should acknowledge limitations of the network as a case study.
For example:
- selection of participating countries;
- variation in national research infrastructure;
- differences in funding;
- differences in data governance and privacy legislation;
- differences in portal functionality;
- potential selection/reporting bias in the five case examples;
- inability to establish whether portal availability actually improved outbreak decision-making or health outcomes.

This is particularly important given the equity claim.

8. The terminology needs careful editing.
There are obvious issues even in the abstract:
- “Datal Portals” → Data Portals
- “origins of the network of National COVID-19 Datal Portals” is awkward.
- “across data types and resources” needs clarification.
- “robust case study” is somewhat evaluative unless supported by explicit criteria.
- “pandemic preparedness across pathogens globally” is broader than the evidence presented.

I would recommend a careful language and terminology edit before acceptance.

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Partly

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: emerging infectious diseases, outbreak response preparedness, and global health.

I confirm that I have read this submission and believe that I have an appropriate level of

expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 24 August 2026

https://doi.org/10.21956/openreseurope.25296.r78910

© **2026 Horsfall D.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

David Horsfall
Newcastle University, Newcastle upon Tyne, UK

Thanks for addressing the comments. No further points.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Research Software Engineering

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 06 January 2026

https://doi.org/10.21956/openreseurope.21697.r64572

© **2026 Horsfall D.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

David Horsfall
Newcastle University, Newcastle upon Tyne, UK

This article presents a comprehensive account of the conception, development, and operation of a network of National COVID-19 Data Portals that emerged during the SARS-CoV-2 pandemic and were linked to a central European COVID-19 Data Portal. Built initially for the Swedish COVID-19 Data Portal, the network expanded to include twelve national portals, each tailored to local needs. The manuscript provides a compelling case study of how lightweight, reusable infrastructure can support rapid data mobilisation during a public health emergency, while also laying foundations for future preparedness.

The rationale for developing national portals is clearly explained and well motivated. The authors convincingly argue for the need to complement a central European platform with national entry points that reflect local language, governance structures, data availability, and community trust. The localisation of data and resources within national and regional contexts is a key strength of the work, particularly in overcoming barriers to data discovery and trust that can arise in cross-border settings. The framing of the portals both as an emergency response and as a prototype for longer-term pathogen data infrastructure is effective and timely.

The technical design choices underpinning the portals are generally sound and well suited to the constraints of rapid deployment, sustainability, and reuse. The use of static site generation with markdown-based content is appropriate and clearly justified. The shared template (originating from the Swedish portal) enabled multiple countries to rapidly deploy functional national hubs that aggregated links to datasets, tools, publications, and services, while signposting users to other national and international resources. The strengths of this approach, particularly speed of deployment and ease of adoption, are clearly articulated, and the reuse of a common codebase appears to have facilitated cross-national collaboration during a period of significant disruption.

However, the technical description remains largely narrative. A more structured overview of the core components, data flows, and integration points would improve clarity, particularly for technically oriented readers. For example, which elements are shared across portals and which are locally customised would help readers understand how the system scales and evolves.

While the limitations of a static-site approach are acknowledged, these trade-offs could be explored in more depth. The lack of native dynamic content, database-backed services, or API-level access is an important design constraint, especially given the later discussion of more dynamic web frameworks (e.g. React, Svelte) in the context of future pathogen portals. The manuscript mentions the use of Python-based workflows to generate dynamic data visualisations, and providing a clearer description of how these workflows fit into the overall architecture, would strengthen the technical description.

The organisational and social dimensions of the work are a notable strength. The description of how the Swedish portal was rapidly embedded within the national research ecosystem (supported by recognition and direction from the Swedish Research Council and SciLifeLab) provides a particularly compelling example of how governance, funding, and infrastructure can align to support community coordination. At the same time, the manuscript highlights that some early efforts relied heavily on volunteer contributions, raising important questions about sustainability and resourcing that are highly relevant for future preparedness initiatives.

The country-specific case studies are informative and balanced, including constructive discussion of challenges such as data governance barriers, legislative constraints, and the complexities of coordinating across institutions and ministries, as highlighted in the Polish experience.

The article provides sufficient information and references to allow replication and reuse by others, particularly through explicit links to GitHub repositories, licensing information, and companion technical publications.

The manuscript would benefit from a slightly more systematic treatment of usage metrics and user engagement. While several portals report page views and visitor statistics, the methods used

to collect these metrics are not always clear and there is significant variation. Additional context on traffic (how user requirements were gathered in a fast-moving environment, and how feedback from communities was collected and acted upon) would further strengthen the discussion of impact and lessons learned.

Overall, this is a valuable and well-documented contribution that captures an important collaborative response to the COVID-19 pandemic. With modest clarification of technical details, and impact interpretation, the manuscript would be further strengthened and serve as a useful reference for future national and international data infrastructure efforts.

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Partly

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Research Software Engineering

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 05 Mar 2026

Marianna Ventouratou

We thank the reviewers for taking the time to provide detailed comments and evaluation of our paper. Please see our individual comments to the reports below.

However, the technical description remains largely narrative. A more structured overview of the core components, data flows, and integration points would improve clarity, particularly for technically oriented readers. For example, which elements are shared across portals and which are locally customised would help readers understand how the system scales and evolves.

We have added a brief explanation, including some limitations, in the section 'Technical aspects'.

While the limitations of a static-site approach are acknowledged, these trade-offs could be explored in more depth. The lack of native dynamic content, database-backed services, or API-level access is an important design constraint, especially given the later discussion of more dynamic web frameworks (e.g. React, Svelte) in the context of future pathogen portals. The manuscript mentions the use of Python-based workflows to generate dynamic data visualisations, and providing a clearer description of how these workflows fit into the overall architecture, would strengthen the technical description.

An interconnected model with data flows would be an option, but due to funding constraints and urgency in development, we have not implemented it yet (only partial presentation of covid data from Sweden on the Pathogens Portal for example), There were coordinated data flows which lay outside the Portal network, for example in the data hubs (please see here https://www.covid19dataportal.org/data-hubs, and for more details here: https://www.ebi.ac.uk/ena/datahubs/). The Portals are more about presentation of data given they flowed into underlying databases. There was coordination, e.g. via the national coordinators, but this was a complementary activity to the portals per se.

The organisational and social dimensions of the work are a notable strength. The description of how the Swedish portal was rapidly embedded within the national research ecosystem (supported by recognition and direction from the Swedish Research Council and SciLifeLab) provides a particularly compelling example of how governance, funding, and infrastructure can align to support community coordination. At the same time, the manuscript highlights that some early efforts relied heavily on volunteer contributions, raising important questions about sustainability and resourcing that are highly relevant for future preparedness initiatives.

We explained the limitations, especially around funding and sustainability in the 'lessons learnt' section.

The manuscript would benefit from a slightly more systematic treatment of usage metrics and user engagement. While several portals report page views and visitor statistics, the methods used to collect these metrics are not always clear and there is significant variation. Additional context on traffic (how user requirements were gathered in a fast-moving environment, and how feedback from communities was collected and acted upon) would further strengthen the discussion of impact and lessons learned.

Usage metrics were requested in 2021–2022 but responses were inconsistent or limited. Through EU-funded projects, the portals and network engaged with community members to gather input, and different groups (e.g. SPSP) contributed requirements from their own networks to guide development. Collecting usage statistics has been challenging due to data availability, external ownership, and possible GDPR constraints, suggesting a need to define clearer, useful impact metrics for the future.

Competing Interests: No competing interests were disclosed.

Author Response 23 Jul 2026

Marianna Ventouratou

We thank the reviewer for taking the time to provide detailed comments and evaluation of our paper. Please see our individual comments to the points raised below. However, the technical description remains largely narrative. A more structured overview of the core components, data flows, and integration points would improve clarity, particularly for technically oriented readers. For example, which elements are shared across portals and which are locally customised would help readers understand how the system scales and evolves.

- We have added a brief explanation , including some limitations, in the section 'Technical aspects'.

While the limitations of a static-site approach are acknowledged, these trade-offs could be explored in more depth. The lack of native dynamic content, database-backed services, or API-level access is an important design constraint, especially given the later discussion of more dynamic web frameworks (e.g. React, Svelte) in the context of future pathogen portals. The manuscript mentions the use of Python-based workflows to generate dynamic data visualisations, and providing a clearer description of how these workflows fit into the overall architecture, would strengthen the technical description.

- An interconnected model with data flows would be an option, but due to funding constraints and urgency in development, we have not implemented it yet (only partial presentation of covid data from Sweden on the Pathogens Portal for example), There were coordinated data flows which lay outside the Portal network, for example in the data hubs (please see here https://www.covid19dataportal.org/data-hubs, and for more details here: https://www.ebi.ac.uk/ena/datahubs/). The Portals are more about presentation of data given they flowed into underlying databases. There was coordination, e.g. via the national coordinators, but this was a complementary activity to the portals per se.

The organisational and social dimensions of the work are a notable strength. The description of how the Swedish portal was rapidly embedded within the national research ecosystem (supported by recognition and direction from the Swedish Research Council and SciLifeLab) provides a particularly compelling example of how governance, funding, and infrastructure can align to support community coordination. At the same time, the manuscript highlights that some early efforts relied heavily on volunteer contributions, raising important questions about sustainability and resourcing that are highly relevant for future preparedness initiatives.

- We explained the limitations, especially around funding and sustainability in the 'lessons learnt' section.

The manuscript would benefit from a slightly more systematic treatment of usage metrics and user engagement. While several portals report page views and visitor statistics, the methods used to collect these metrics are not always clear and there is significant variation. Additional context on traffic (how user requirements were gathered in a fast-moving environment, and how feedback from communities was collected and acted upon) would further strengthen the discussion of impact and lessons learned.

- Usage metrics were requested in 2021–2022 but responses were inconsistent or limited. Through EU-funded projects, the portals and network engaged with community members to gather input, and different groups (e.g. SPSP) contributed

requirements from their own networks to guide development. Collecting usage statistics has been challenging due to data availability, external ownership, and possible GDPR constraints, suggesting a need to define clearer, useful impact metrics for the future.

Competing Interests: No competing interests were disclosed.

Reviewer Report 30 December 2025

https://doi.org/10.21956/openreseurope.21697.r64578

© 2025 kapoor M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Muskan kapoor
Iowa State University, Ames, USA

The mansucript describes the coordination and development of various COVID-19 data portals and demonstrate how this helps in data sharing and standardization. The authors very well describve the next steps with national portals highlighting funding along with usage of FAIR pronciples. However, there are few suggestions to improve the clarity of the paper.

First would be how were the FAIR principles evaluated, were they based on original COVID-19 Portal? If yes, did they differ across nations and how the author harmonized them?
How was metadata/ontologies of different data types handled across different portals? The authors should talk more about how different metadata schemas or ontology were made standard.
Authors should also mention lessons leanred such that this can be extended beyond COVID-19 to other portals thus increasing the applications of this study.
Finally, It would be great to see citations feature collabortaive work from these portals.

Minor comments:
- Please provide reference in line: A dedicated team, composed of individuals with relevant technical expertise and backgrounds in life sciences, was established in Autumn 2024.
- Include collabortaive work citations if any: This included work with national organisations (e.g. Biobanks Sweden), national infrastructure (e.g. National Pandemic Centre, and the SciLifeLab Autoimmunity and Serology Profiling unit), and many research groups.
-Line: These centres either deposited sequencing data directly or collected biological material sequenced at IBCH PAS. Genomic data and the Poland-focused Nextstrain COVID-19[xiii] instance supervised by COVID-HUB-PL team enabled researchers to monitor the spread and evolution of the virus in real-time, providing insights into how different variants emerged and spread geographically. - were there any metadata schema checks performed?

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: working on FAIR principles and standardizing metadata schemas between agriculture and human portals. Also worked on developing metadata schemas, automated data transfer, data visualization, annotaion and analysis pipelines between portals

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 05 Mar 2026

Marianna Ventouratou

We thank the reviewers for taking the time to provide detailed comments and evaluation of our paper. Please see our individual comments to the reports below.

1. First would be how were the FAIR principles evaluated, were they based on the original COVID-19 Portal? If yes, did they differ across nations and how the author harmonized them?

The European COVID-19 Data Portal presents data that originate from FAIR-compliant repositories within EMBL-EBI and beyond, so the data presented on it adhere to FAIR principles. As the portal primarily aggregates and displays these archived datasets, its FAIR compliance largely reflects that of the underlying repositories. The Swedish Portal is also listed on FAIRsharing, supporting its status as a FAIR-aligned resource.
While the national portals may not uniformly meet identical standards and some elements, such as dashboards drawing on national or institutional sources, may vary in implementation, the underlying data are generally sourced from ELIXIR Core Data Resource repositories (https://elixir-europe.org/platforms/data/core-data-resources; at least 10 of them indexed on the COVID-19 Data Portal) that follow FAIR standards, ensuring overall alignment with FAIR principles.

2. How was metadata/ontologies of different data types handled across different portals? The authors should talk more about how different metadata schemas or ontology were made standard.

The national portals didn't really bring together different data types, so they weren't handled 'across portals'. This is something that we are still working towards with the Pathogen Portals using a three-tiered indexing system, where in the lower tier there is just a link out to the Pathogen Portal node (as the 'national portals' have been renamed for the Pathogen Portals network). Tier 2, where some of the data is presented on the main Pathogens Portal (e.g. https://www.pathogensportal.org/nationaldataportal?db=national-portals-sweden) to a fully findable Tier 3 where all the data from the nodes is findable and presented on the main Portal.

3. Authors should also mention lessons learned such that this can be extended beyond COVID-19 to other portals thus increasing the applications of this study. *A lesson learnt and limitations section has been added.*

4. Finally, it would be great to see citations feature collaborative work from these portals.

These are portals rather than direct databases, so it is more likely that there are publications related to the use of the underlying resources (i.e. the databases), rather than the portals themselves.

Minor comments:

1. Please provide reference in line: A dedicated team, composed of individuals with relevant technical expertise and backgrounds in life sciences, was established in Autumn 2024. *DONE*
2. Include collaborative work citations if any: This included work with national organisations (e.g. Biobanks Sweden), national infrastructure (e.g. National Pandemic Centre, and the SciLifeLab Autoimmunity and Serology Profiling unit), and many research groups. *DONE*
3. Line: These centres either deposited sequencing data directly or collected biological material sequenced at IBCH PAS. Genomic data and the Poland-focused Nextstrain COVID-19[xiii] instance supervised by COVID-HUB-PL team enabled researchers to monitor the spread and evolution of the virus in real-time, providing insights into how different variants emerged and spread geographically. - were there any metadata schema checks performed?.

Yes, metadata schema checks were performed. First, contributing centers were provided with standardised Excel templates. Although these templates included multiple predefined fields, in collaboration agreements with our partners, we also defined a minimum mandatory metadata set required for any record to be accepted. Second, upon deposition on our server, each metadata file underwent a validation process by the COVID-HUB-PL team. We checked for completeness, adherence to controlled vocabularies, and format consistency. Any identified discrepancies or missing values were clarified and corrected before they were integrated into the Nextstrain instance or submitted to the databases.

Competing Interests: No competing interests were disclosed.

Author Response 23 Jul 2026

Marianna Ventouratou

1. First would be how were the FAIR principles evaluated, were they based on the

original COVID-19 Portal? If yes, did they differ across nations and how the author harmonized them?

The European COVID-19 Data Portal presents data that originate from FAIR-compliant repositories within EMBL-EBI and beyond, so the data presented on it adhere to FAIR principles. As the portal primarily aggregates and displays these archived datasets, its FAIR compliance largely reflects that of the underlying repositories. The Swedish Portal is also listed on FAIRsharing, supporting its status as a FAIR-aligned resource. While the national portals may not uniformly meet identical standards and some elements, such as dashboards drawing on national or institutional sources, may vary in implementation, the underlying data are generally sourced from ELIXIR Core Data Resource repositories (https://elixir-europe.org/platforms/data/core-data-resources; at least 10 of them indexed on the COVID-19 Data Portal) that follow FAIR standards, ensuring overall alignment with FAIR principles.

1. How was metadata/ontologies of different data types handled across different portals? The authors should talk more about how different metadata schemas or ontology were made standard.

The national portals didn't really bring together different data types, so they weren't handled 'across portals'. This is something that we are still working towards with the Pathogen Portals using a three-tiered indexing system, where in the lower tier there is just a link out to the Pathogen Portal node (as the 'national portals' have been renamed for the Pathogen Portals network). Tier 2, where some of the data is presented on the main Pathogens Portal (e.g. https://www.pathogensportal.org/nationaldataportal?db=national-portals-sweden) to a fully findable Tier 3 where all the data from the nodes is findable and presented on the main Portal.

1. Authors should also mention lessons learned such that this can be extended beyond COVID-19 to other portals thus increasing the applications of this study.

A lessons learnt and limitations section has been added.

1. Finally, It would be great to see citations feature collaborative work from these portals.

These are portals rather than direct databases, so it is more likely that there are publications related to the use of the underlying resources (i.e. the databases), rather than the portals themselves. Minor comments:

1. Please provide reference in line: A dedicated team, composed of individuals with relevant technical expertise and backgrounds in life sciences, was established in Autumn 2024. DONE
2. Include collaborative work citations if any: This included work with national organisations (e.g. Biobanks Sweden), national infrastructure (e.g. National Pandemic Centre, and the SciLifeLab Autoimmunity and Serology Profiling unit), and many research groups. DONE
3. Line: These centres either deposited sequencing data directly or collected biological material sequenced at IBCH PAS. Genomic data and the Poland-focused Nextstrain COVID-19[xiii] instance supervised by COVID-HUB-PL team enabled researchers to monitor the spread and evolution of the virus in real-time, providing insights into how different variants emerged and spread geographically. - were there any metadata schema checks performed?.

Yes, metadata schema checks were performed. First, contributing centers were provided with standardised Excel templates. Although these templates included multiple predefined

fields, in collaboration agreements with our partners, we also defined a minimum mandatory metadata set required for any record to be accepted. Second, upon deposition on our server, each metadata file underwent a validation process by the COVID-HUB-PL team. We checked for completeness, adherence to controlled vocabularies, and format consistency. Any identified discrepancies or missing values were clarified and corrected before they were integrated into the Nextstrain instance or submitted to the databases.

Competing Interests: No competing interests were disclosed.

Reviewer Report 11 December 2025

https://doi.org/10.21956/openreseurope.21697.r57237

© 2025 Bernardini F. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Flavia Bernardini
Universidade Federal Fluminense, Rio de Janeiro, Brazil

The work presents an interesting case study of implementing COVID-19 national data ports in 12 countries through the EU. The authors state that these portals form an interconnected network, adhering to FAIR principles. This initiative is quite important, considering both the perspective of implementing good practices for national open data portals and the perspective of establishing and implementing open data ecosystems. Additionally, describing how each data portal evolved as a technical solution is quite interesting, considering the requirements pointed out by the national governments of each country.

However, the paper lacks the rationale and technical details for developing the new software tool. I explain in detail each aspect of the paper that, from my perspective, is still confusing:

- The manuscript starts by presenting a network of national COVID-19 Data Portals. As my background is in Computer Science, I tend to imagine these data portals are someway fully connected. For me, if this is the case, the portals also interoperate with one another. In this way, I could expect that if a user searches for data in one portal, depending on the level of interoperability, this search could retrieve data from the other portals in the network. However, after reading the paper, it seems to me that this is not the case. Rather, it seems to me that the portals indeed offer some kind of technology for interoperability (perhaps APIs) to interoperate with them, but I'm not sure. This perspective is reinforced by the technical aspects described in the Review of the National COVID-19 Data Portals section. So, in this sense, the term "network" is more likely to be associated with a conceptual network rather than guaranteeing full interoperability. In this regard, I recommend that the authors clarify these aspects properly.

- When reading the paper, I remembered the initiative of Common European Data Spaces - https://digital-strategy.ec.europa.eu/en/policies/data-spaces - that aims to serve as a fully

integrative platform for open data in Europe. There are some implementations of data spaces in some countries within the EU. Does this described network intend to align with this initiative in the future?

- The authors state that "systematic engagement of the COVID-19 Data Platform team at EMBL-EBI in 2020 with the so-called "national coordinators" group, an *ad hoc* community of stakeholders drawn from ELIXIR and EMBL member states and other countries". It seems to me that these organizations were somehow responsible for developing all these national data portals, but I'm not entirely sure. Also, I have no clue if these (EMBL-EBI, ELIXIR, and EMBL) are indeed organizations or what their rule are in defining this network, if this is the case. These "organization" aspects, including the responsibility of each organ and whether or not they are connected to ELIXIR, and if they are, how, are repeated throughout the paper. For instance, when describing the Sweden case, the authors state that "SciLifeLab received an assignment". What is SciLifeLab? This pattern repeats when the authors present other cases from countries such as Poland and the Netherlands, as the Spanish and Norwegian cases suggest that they are somehow related to ELIXIR. Is the network of actors that together define some architectural aspect? I recommend that the authors properly present the relations among all the actors listed in the work. I also suggest that the authors clearly state how these portals are interconnected - through authors in national governments, through a defined architecture to portals, or both?

- In this part of the manuscript: "This variability in the services and features offered nationally reflects the diverse response to national mandates across Europe while retaining interoperability. This model is strong in that it allowed national frameworks to focus on data and projects that mattered on a national level, while still being deeply connected to the more comprehensive European wide portal", what are the meanings of the word "model" and "framework"? From the perspective of modeling an architecture, framework, process, and so on, it is expected that a model or framework is properly modeled and presented, using some notation if a diagram is used. Did the authors model the network? Is the available software a computational framework?

I think that, after clarifying these aspects, it will be easier to better understand the software tool the authors are presenting and how to (re)use it to implement Pathogens Data Portals.

Is the rationale for developing the new software tool clearly explained?

Partly

Is the description of the software tool technically sound?

Partly

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

No

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Computer Science, Data Science, Open Data, Machine Learning and Artificial Intelligence, Data Government

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 05 Mar 2026

Marianna Ventouratou

We thank the reviewers for taking the time to provide detailed comments and evaluation of our paper. Please see our individual comments to the reports below.

1. However, the paper lacks the rationale and technical details for developing the new software tool. I explain in detail each aspect of the paper that, from my perspective, is still confusing:

The technology used varied with many using the Swedish codebase as outlined in the article, but it was not our intention to proscribe and promote a single technical solution.

2. The manuscript starts by presenting a network of national COVID-19 Data Portals. As my background is in Computer Science, I tend to imagine these data portals are someway fully connected. For me, if this is the case, the portals also interoperate with one another. In this way, I could expect that if a user searches for data in one portal, depending on the level of interoperability, this search could retrieve data from the other portals in the network. However, after reading the paper, it seems to me that this is not the case. Rather, it seems to me that the portals indeed offer some kind of technology for interoperability (perhaps APIs) to interoperate with them, but I'm not sure. This perspective is reinforced by the technical aspects described in the Review of the National COVID-19 Data Portals section. So, in this sense, the term "network" is more likely to be associated with a conceptual network rather than guaranteeing full interoperability. In this regard, I recommend that the authors clarify these aspects properly.

To clarify, we have added an explanation regarding the usage of the word 'network' in the last paragraph of the introduction.

3. When reading the paper, I remembered the initiative of Common European Data Spaces - https://digital-strategy.ec.europa.eu/en/policies/data-spaces - that aims to serve as a fully integrative platform for open data in Europe. There are some implementations of data spaces in some countries within the EU. Does this described network intend to align with this initiative in the future?

Part of the Common European Data Spaces, under the Health space, is the European Genomics Data Infrastructure (GDI). Several of the partners of the EGDI were also part of the Horizon Europe projects that supported the European COVID-19 Data Platform. The Pathogen Platform and the Pathogens Portal, which is the natural evolution of the COVID-19 Portals, also involves the same organisations and infrastructures. For example: EMBL-EBI, ELIXIR, BBMRI, Erasmus Medical Center Rotterdam, Barcelona Supercomputing Center (BSC), Carlos III Health Institute

(ISCII), Centre for Genomic Regulation (CRG). Hence there are already ongoing synergies, communication and collaboration, all aligning towards the Common European Data Space. These can certainly be exploited further in the future, should the opportunity and the funding become available.

4. The authors state that "systematic engagement of the COVID-19 Data Platform team at EMBL-EBI in 2020 with the so-called "national coordinators" group, an ad hoc community of stakeholders drawn from ELIXIR and EMBL member states and other countries". It seems to me that these organizations were somehow responsible for developing all these national data portals, but I'm not entirely sure. Also, I have no clue if these (EMBL-EBI, ELIXIR, and EMBL) are indeed organizations or what their rule are in defining this network, if this is the case. These "organization" aspects, including the responsibility of each organ and whether or not they are connected to ELIXIR, and if they are, how, are repeated throughout the paper. For instance, when describing the Sweden case, the authors state that "SciLifeLab received an assignment". What is SciLifeLab? This pattern repeats when the authors present other cases from countries such as Poland and the Netherlands, as the Spanish and Norwegian cases suggest that they are somehow related to ELIXIR. Is the network of actors that together define some architectural aspect? I recommend that the authors properly present the relations among all the actors listed in the work. I also suggest that the authors clearly state how these portals are interconnected - through authors in national governments, through a defined architecture to portals, or both?

For clarity we have spelled out the acronyms for each organisation and institution, added a brief description of each one in the footnotes and added links to their respective websites.

5. In this part of the manuscript: "This variability in the services and features offered nationally reflects the diverse response to national mandates across Europe while retaining interoperability. This model is strong in that it allowed national frameworks to focus on data and projects that mattered on a national level, while still being deeply connected to the more comprehensive European wide portal", what are the meanings of the word "model" and "framework"? From the perspective of modeling an architecture, framework, process, and so on, it is expected that a model or framework is properly modeled and presented, using some notation if a diagram is used. Did the authors model the network? Is the available software a computational framework?

As the “network of the Portals” is more of an output of social cooperation, the word “model” is also used in the sense that it means a simplified description of a system, rather than a computational model.

Competing Interests: No competing interests were disclosed.

Author Response 23 Jul 2026

Marianna Ventouratou

We thank the reviewer for taking the time to provide detailed comments and evaluation of

our paper. Please see our individual comments to the points raised in the review below.
The work presents an interesting case study of implementing COVID-19 national data ports in 12 countries through the EU. The authors state that these portals form an interconnected network, adhering to FAIR principles. This initiative is quite important, considering both the perspective of implementing good practices for national open data portals and the perspective of establishing and implementing open data ecosystems. Additionally, describing how each data portal evolved as a technical solution is quite interesting, considering the requirements pointed out by the national governments of each country.

1. However, the paper lacks the rationale and technical details for developing the new software tool. I explain in detail each aspect of the paper that, from my perspective, is still confusing:

The technology used varied with many using the Swedish codebase as outlined in the article, but it was not our intention to proscribe and promote a single technical solution.

1. The manuscript starts by presenting a network of national COVID-19 Data Portals. As my background is in Computer Science, I tend to imagine these data portals are someway fully connected. For me, if this is the case, the portals also interoperate with one another. In this way, I could expect that if a user searches for data in one portal, depending on the level of interoperability, this search could retrieve data from the other portals in the network. However, after reading the paper, it seems to me that this is not the case. Rather, it seems to me that the portals indeed offer some kind of technology for interoperability (perhaps APIs) to interoperate with them, but I'm not sure. This perspective is reinforced by the technical aspects described in the Review of the National COVID-19 Data Portals section. So, in this sense, the term "network" is more likely to be associated with a conceptual network rather than guaranteeing full interoperability. In this regard, I recommend that the authors clarify these aspects properly.

To clarify, we have added an explanation regarding the usage of the word ‘network’ in the last paragraph of the introduction.

1. When reading the paper, I remembered the initiative of Common European Data Spaces - https://digital-strategy.ec.europa.eu/en/policies/data-spaces - that aims to serve as a fully integrative platform for open data in Europe. There are some implementations of data spaces in some countries within the EU. Does this described network intend to align with this initiative in the future?

Part of the Common European Data Spaces, under the Health space, is the European Genomics Data Infrastructure (GDI). Several of the partners of the EGDI were also part of the Horizon Europe projects that supported the European COVID-19 Data Platform. The Pathogen Platform and the Pathogens Portal, which is the natural evolution of the COVID-19 Portals, also involves the same organisations and infrastructures. For example: EMBL-EBI, ELIXIR, BBMRI, Erasmus Medical Center Rotterdam, Barcelona Supercomputing Center (BSC), Carlos III Health Institute (ISCII), Centre for Genomic Regulation (CRG). Hence there are already ongoing synergies, communication and collaboration, all aligning towards the Common European Data Space. These can certainly be exploited further in the future, should the opportunity and the funding become available.

1. The authors state that "systematic engagement of the COVID-19 Data Platform team at EMBL-EBI in 2020 with the so-called "national coordinators" group, an ad hoc community of stakeholders drawn from ELIXIR and EMBL member states and other countries". It seems to me that these organizations were somehow responsible for

developing all these national data portals, but I'm not entirely sure. Also, I have no clue if these (EMBL-EBI, ELIXIR, and EMBL) are indeed organizations or what their rule are in defining this network, if this is the case. These "organization" aspects, including the responsibility of each organ and whether or not they are connected to ELIXIR, and if they are, how, are repeated throughout the paper. For instance, when describing the Sweden case, the authors state that "SciLifeLab received an assignment". What is SciLifeLab? This pattern repeats when the authors present other cases from countries such as Poland and the Netherlands, as the Spanish and Norwegian cases suggest that they are somehow related to ELIXIR. Is the network of actors that together define some architectural aspect? I recommend that the authors properly present the relations among all the actors listed in the work. I also suggest that the authors clearly state how these portals are interconnected - through authors in national governments, through a defined architecture to portals, or both?

For clarity we have spelled out the acronyms for each organisation and institution, added a brief description of each one in the footnotes and added links to their respective websites.

1. In this part of the manuscript: "This variability in the services and features offered nationally reflects the diverse response to national mandates across Europe while retaining interoperability. This model is strong in that it allowed national frameworks to focus on data and projects that mattered on a national level, while still being deeply connected to the more comprehensive European wide portal", what are the meanings of the word "model" and "framework"? From the perspective of modeling an architecture, framework, process, and so on, it is expected that a model or framework is properly modeled and presented, using some notation if a diagram is used. Did the authors model the network? Is the available software a computational framework?

As the “network of the Portals” is more of an output of social cooperation, the word “model” is also used in the sense that it means a simplified description of a system, rather than a computational model. For clarity, we replaced the word model with 'configuration', and frameworks with 'structures'.

Competing Interests: No competing interests were disclosed.

Reviewer Report 04 September 2025

https://doi.org/10.21956/openreseurope.21697.r57662

© 2025 Battineni G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Gopi Battineni
University of Camerino, Camerino, Italy

Marianna et al developed real time COVID 19 data portal which add the network value of national COVID-19 data portals. I had few suggestions before accepting this work

The abstract could be improved by highlighting the impact. It reads more like a descriptive summary than an analytical one. The lessons learned from COVID-19 should be clarified.
Several sentences in the introduction are long and full of references, which makes it difficult to read.
The inclusion of a critical perspective on coordination challenges, such as governance and sustainability, may be beneficial
I think supplementary materials would be better suited to cover implementation details (markdown, site generator, hosting).
Add comments on the most widely used and valuable features in national data portals.
The section titled "next with national portals" highlighted the transition to pathogen portals, but it is fragmented, with too much emphasis on technical details.
For pandemic preparedness, it is important to emphasize lessons learned.
An unresolved issue and limitations paragraph should be added.
Despite its strengths, the report could have benefited from a clearer discussion of limitations and lessons learned.

Is the rationale for developing the new software tool clearly explained?

Yes

Is the description of the software tool technically sound?

Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?

Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Data mining, machine learning, pedictive analytics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 05 Mar 2026

Marianna Ventouratou

We thank the reviewers for taking the time to provide detailed comments and evaluation of our paper. Please see our individual comments to the reports below. 1st Report - Gopi Battineni, University of Camerino, Camerino, Italy

1. The abstract could be improved by highlighting the impact. It reads more like a

descriptive summary than an analytical one.

1. The lessons learned from COVID-19 should be clarified.

We added a sentence in the abstract, which is further expanded in the ‘Lessons learnt and limitations’ section.

1. Several sentences in the introduction are long and full of references, which makes it difficult to read.

We rephrased the introduction as much as this was possible.

1. The inclusion of a critical perspective on coordination challenges, such as governance and sustainability, may be beneficial

We rephrased and expanded the first paragraph under section ‘Organisation’ to address this suggestion.

1. I think supplementary materials would be better suited to cover implementation details (markdown, site generator, hosting).

We have taken this into consideration, this article is intended to provide an overview of the strategic, technical and connectivity-related elements of the various Portals.

1. Add comments on the most widely used and valuable features in national data portals.

This is a limitation we acknowledge, and difficult to address due the restrictions and variability in statistics held for each Portal. We have added a paragraph to this effect under the ‘Lessons learnt and limitations’.

1. The section titled "next with national portals" highlighted the transition to pathogen portals, but it is fragmented, with too much emphasis on technical details.

We revised the wording in this section to provide a more holistic view.

1. For pandemic preparedness, it is important to emphasize lessons learned.

1. An unresolved issue and limitations paragraph should be added.

1. Despite its strengths, the report could have benefited from a clearer discussion of f limitations and lessons learned.

A special section ‘Lessons learnt and limitations’ has been added to address the reviewer comments related to the lessons learnt and limitations overall.

Competing Interests: No competing interests were disclosed.

Author Response 23 Jul 2026

Marianna Ventouratou

We thank the reviewer for taking the time to provide detailed comments and evaluation of our paper. Please see our individual comments to the points raised in the review below.

1. The abstract could be improved by highlighting the impact. It reads more like a descriptive summary than an analytical one.

1. The lessons learned from COVID-19 should be clarified.

We added a sentence in the abstract, which is further expanded in the ‘Lessons learnt and limitations’ section.

1. Several sentences in the introduction are long and full of references, which makes it difficult to read.

We rephrased the introduction as much as this was possible.

1. The inclusion of a critical perspective on coordination challenges, such as governance and sustainability, may be beneficial

We rephrased and expanded the first paragraph under section ‘Organisation’ to address this suggestion.

1. I think supplementary materials would be better suited to cover implementation details (markdown, site generator, hosting).

We have taken this into consideration, this article is intended to provide an overview of the strategic, technical and connectivity-related elements of the various Portals

1. Add comments on the most widely used and valuable features in national data portals.

This is a limitation we acknowledge, and difficult to address due the restrictions and variability in statistics held for each Portal. We have added a paragraph to this effect under the ‘Lessons learnt and limitations’.

1. The section titled "next with national portals" highlighted the transition to pathogen portals, but it is fragmented, with too much emphasis on technical details.

We revised the wording in this section to provide a more holistic view.

1. For pandemic preparedness, it is important to emphasize lessons learned.

1. An unresolved issue and limitations paragraph should be added.

1. Despite its strengths, the report could have benefited from a clearer discussion of f limitations and lessons learned.

Done; a special section ‘Lessons learnt and limitations’ has been added to address the reviewer comments related to the lessons learnt and limitations overall.

Competing Interests: No competing interests were disclosed.