Project plan
PaRI

The plan is an agreement between the project owner/orderer Tomasz Malkiewicz and the project manager Abdulrahman Azab.

To be verified through a steering group decision.

Table of contents

Change Log

Date	Change made	Decision
2020-10-27	First draft by the project manager	Approved as a skeleton in the constitutional meeting
2020-11-13	Second draft by the project team to be presented to the SG	
2020-11-13	Revised by the SG	Approved by the SG

1 Background and connections

1.1 Background

The ongoing COVID-19 outbreak has induced an unprecedented near real-time exchange of information with an exceptional number of individual researchers performing multiple analyses in parallel, using publicly available, as well as their own, data. It has, however, also become clear that there is an unmet demand for services based on the real-time comparison of different epidemiological analyses of the same data, as well a major need for exchange of more person-sensitive information, such as geolocalisation, patient data, etc., between authorities, researchers and different countries. The ideal solution, described by the domain experts working with the data, will allow integration of patient-specific and population-wide levels of epidemiological data and combine it with sequencing and later serological information.
The activities include:
Genome-wide association studies (GWAS), in the form of case-control studies, may help to identify factors that predispose to adverse disease patterns (besides those of well-known impact, such as age).
Drug testing is being performed both in the Nordic countries and internationally, in controlled clinical trials. Some are performed in larger international contexts, such as the Coalition for Epidemic Preparedness Innovations (CEPI), but also in more localized settings. Various types of data are generated in these trials, and in some cases, there is an urgent need for sharing of such data.
Viral sequencing is another very important aspect. As the SARS-CoV-2 virus is an RNA virus, it implies a fast pace of mutationally generated diversity of strains, i.e. the viral evolutionary processes.
Human viral gene expression is another important aspect. The expression of these sequences may in fact be studied by harvesting raw RNA-seq data from various existing repositories, and queried for the presence of viral sequences. These sequences may potentially contribute to regulate existing normal human biologic processes in several tissues.
In many of the settings described above, human data are involved, frequently being of a sensitive nature. It is therefore imperative to ensure that the laboratories generating the data can keep sufficient control over their data, can deposit complete datasets (including raw data) and that the fact of publishing results by the researchers does not impact the confidentiality of the patients. The scope of the project in spite of its current focus on COVID-19 pandemic related research, due to the outbreak, is significantly wider including preparedness for all pandemic related research in the future.

1.2 Connections with other organisations or projects

Nordic activities

There are several actions supporting SARS-CoV-2/COVID-19 research in the Nordic region and internationally[1]. Nordic national service providers are committed to providing access to compute and storage resources for pandemic research workloads.

Norway: TSD (Services For Sensitive Data) at UiO is prioritizing the setup and use of secure HPC computing for COVID-19 projects. Uninett Sigma2 is prioritizing COVID-19 jobs on SAGA (10,080 CPU, 32 GPU units). In addition, COVID-19 projects are given priority on the Norwegian Research and

[1] https://docs.google.com/document/d/1W3TXbv0bM6f22LbKkP369aomvUoo3niP-CDiqd28ibw/edit?usp=sharing

Education Cloud (NREC). UiO, UiB and Uninett Sigma2 are committed to NeIC PaRI. ELIXIR Norway Covid-19 portal: https://elixir.no/corona/ https://covid19dataportal.no/ .

Estonia: ETAIS is prioritising access to UT-Rocket for pandemic related workload. UT-Rocket is a HPC cluster with 6,000+ cores. Also accessible via Galaxy web interface. It also has extensions to allow processing of sensitive data. ETAIS is committed to NeIC PaRI.

Denmark: The National life science center's Computerome HPC installation is built using a joint collaboration between Technical university of denmark and University of Copenhagen with its 50,000+ CPU cores and 1,600+ users are putting up a dedicated e-Infrastructure to analyze pandemic data. Private clouds from Computerome-2 HPC are prioritised and included in the deployed solutions. DTU is committed to NeIC PaRI.

Finland: COVID-19 gets priority on all CSC's computational resources. A prioritized partition titled "covid 19" has been set up in Puhti. Over 10 COVID-19 research projects are supported.

Sweden: NBIS as the Swedish ELIXIR node and forming the SciLifeLab Bioinformatics platform provides support, infrastructure and training to COVID-19-related research. SciLifeLab is the national COVID-19 research data hub. SNIC is committed to NeIC PaRI facilitating access to the sensitive data analysis service, SNIC Bianca, and SNIC Science Cloud (SSC), on the condition that SNIC is able to decide on a per incident basis whether potential international users can get access to our Resources or not.

International Collaboration

EOSC-Nordic: WP5 "demonstrators" will join the Nordic collaboration actions through NeIC PaRI. T5.4 "Sensitive Data" will contribute to building the secure cloud infrastructure for sharing and analysing pandemic research datasets. T5.2 "Analysis and Post-processing across borders" will contribute to building the nordic platform for analysing SARS-CoV-2 genetic data by facilitating access to different Nordic HPC resources by the Nordic pandemic Galaxy portal. WP4 is about FAIR data. VODAN project will also be considered [2]
Heilsa: The NeIC Heilsa project is supporting two use cases which are connected to genomic analysis and secure data sharing activities (WP2, 3).
PRACE-6IP: Has opened on a Fast Track access for coronavirus research[3]. A Coronavirus Task Force has been setup at PRACE AISBL on 31.3[4]. PRACE WP6.2.3 "Lightweight virtualization service" will support NeIC PaRI composing HPC and ML containers to run Galaxy workloads on HPC systems, including GPU.
COVID-19 Data Portal[5]: EMBL-EBI and partners have set up the COVID-19 Data Portal, which will bring together relevant datasets submitted to EMBL-EBI and other major centres for biomedical data. The aim is to facilitate data sharing and analysis, and to accelerate coronavirus research.
ELIXIR: ELIXIR is maintaining a COVID-19 portal[6] and contributing through its member states to a Europe-wide distributed compute network. Moreover, the ELIXIR communities are providing expert knowledge, which is fundamental in workflow development and driving the pan-european infrastructures.
ELIXIR CONVERGE: EU project 2020--2023 aiming at improved data management across Europe. ELIXIR-SE is leading WP1 aiming at establishing a pan-European data management network. ELIXIR-NO is co-leading WP3 aiming at development of a data management toolkit. Within ELIXIR-CONVERGE additional activities aiming at FAIRifying COVID-19-related data are planned.

[2] https://www.go-fair.org/implementation-networks/overview/vodan/
[3] https://prace-ri.eu/prace-support-to-mitigate-impact-of-covid-19-pandemic/
[4] https://www.hpcwire.com/2020/03/25/doe-expands-on-role-of-covid-19-supercomputing-consortium/
[5] https://www.covid19dataportal.org/
[6] https://elixir-europe.org/services/covid-19

EOSC-life: EOSC-Life brings together the 13 Biological and Medical 'ESFRI' research infrastructures (BMS RIs) to publish FAIR research data and provide Open Infrastructures based services like Galaxy and WorkflowHub.eu. As part of WP2 many COVID-19 workflows have been developed and published [7].

Nordic collaboration will speed up developments needed and make these faster and more cost-efficient. The parties also have close collaboration with ELIXIR, especially on human data and data management, and the activities in this project will be coordinated with similar actions in ELIXIR and EOSC-life.

ECRIN EOSC-Life WP14 has funded TSD and ECRIN for storage and sharing of clinical COVID-19 data. TSD is doing the technical part of the work while ECRIN is doing the policy part. The work is carried out March 2020 - Dec 2021. Possibly synergy is connecting the TSD COVID-19 repo to EGA so that the data is made FAIR through publishing the metadata on central EGA nodes

1.3 Success factors

The success of the project is based on good collaboration between the participating partners. This requires active communication, motivation, clear target-setting and functional management practices. On the other hand, the project is depending on the secure IT services at the partners being operated and managed professionally and continuously. It is also important to maintain good stakeholder relations with the national e-infrastructures in Nordic countries. Finally, outside the project especially the interaction with the partners with clear synergies and shared goals, e.g. ELIXIR CONVERGE and COVID-19 portal, needs to be fluent.

2 Contributions to strategic objectives

PaRI aims at aligning with other COVID-19 European initiatives, e.g. ELIXIR and EOSC-Life, thereby contributing to their strategic objectives. Efficient response to pandemic (COVID-19 and future ones) is a key part of National Strategies of the PaRI partners.

PaRI will contribute NeIC Strategic objectives, which are aligned with PaRI partners objectives, as follows :

Beneficial collaborations are the principal way that NeIC brings together the needs, interests and resources to create e-infrastructure to support research excellence in the Nordic region. PaRI is implementing those among Nordic and Estonian National Providers and possibly among other European countries.

Nordic Influence is the effect that NeIC has in improving and advancing e-infrastructure for researchers and for society. PaRI aims to enable easy and seamless access to datasets for researchers and possibly non-academic users.

Motivated People are essential in the collaborations and to bring the results into the research domains and society. PaRI is a collaboration of six countries being also the NeIC members (Finland collaborates as an observer) and Germany. PaRI is not tied to COVID-19, but will work

[7] https://workflowhub.eu

on preparing the grounds and expertise for future pandemics. At the core of PaRI is also its consortium's expertise.

Effective processes bind NeIC into an organization that is able to realize the benefits and influences that come from the collaborations - PaRI aims at contributing to improvement of the processes at partners organizations, e.g., dialogue between the researchers and national providers.

The project activities would include:
Genome-wide association studies (GWAS), in the form of case-control studies, may help to identify factors that predispose to adverse disease patterns (besides those of well-known impact, such as age). In addition, there is GWAS data which is connected to psychiatry.
Drug testing is being performed both in the Nordic countries and internationally, in controlled clinical trials. Some are performed in larger international contexts, such as the Coalition for Epidemic Preparedness Innovations (CEPI), but also in more localized settings. Various types of data are generated in these trials, and in some cases, there is an urgent need for sharing of such data.
Viral sequencing is another very important aspect. As the SARS-CoV-2 virus is an RNA virus, it implies a fast pace of mutationally generated diversity of strains, i.e. the viral evolutionary processes.
Human viral gene expression is another important aspect. The expression of these sequences may in fact be studied by harvesting raw RNA-seq data from various existing repositories, and queried for the presence of viral sequences. These sequences may potentially contribute to regulate existing normal human biologic processes in several tissues.

In many of the settings described above, human data are involved, frequently being of a sensitive nature. It is therefore imperative to ensure that the laboratories generating the data can keep sufficient control over their data, can deposit complete datasets (including raw data) and that the fact of publishing results by the researchers does not impact the confidentiality of the patients. The scope of the project in spite of its current focus on COVID-19 pandemic related research, due to the outbreak, is significantly wider including preparedness for all pandemic related research in the future.

2.1 Limitations

Important part of a project plan is to focus the work to the desired objectives by making clear what is included in the scope, and what is not included. This section lists items that are not planned to be tackled within the project.

No development of analysis tools

All analysis tools will be provided by ELIXIR tools platform and ELIXIR Galaxy community. Other parties may also contribute as software providers

Focus on Pandemic data only

Only pandemic data will be considered in the scope of the project

No operational services

The project will develop and deliver a set of prototype services. The operational support will be provided by the stakeholders and e-Infrastructure providers who are/will be committed

No transfer of data ownership

The services developed and facilitated in PaRI are provided as data processor roles. The

services have no influence on the ownership or control over the data.

No Hardware
: PaRI will not aim to own or operate e-infrastructure hardware, but will develop and facilitate access to such infrastructures.

2.2 Recipients and approval criteria

In the PPS project management model, each result has to be delivered and transferred. These terms refer to deploying an outcome for use (delivery), and to handing out the result to the party that is going to maintain or operate it (transferral).

The project to define recipients for deliveries and transferrals of the results, when applicable. This aims to safeguard the quality of the project outcomes in a complementary manner, so that the recipient of the delivery will review the outcome from the user perspective, and recipient of transferral from the maintenance and operations perspective.

The PaRI project plans to mainly identify delivery recipients from among the Reference Group people. The recipients for the transferral are typically from the partners or within national e-infrastructures. The recipients may also be from within the project, but they should not be the ones directly contributing to the work on the delivery object. There may be several recipients for each of the deliveries.

The recipients are to be identified by the WP leaders well in advance of the delivery date, so that there is enough time for the quality review. After the review from recipients, each of the delivery objects is also formally approved by the steering group.

3 Schedule and resource needs

3.1 Prerequisites and outer dependencies

Prerequisites

Service contracts

NeIC funding model is based on service contracts that formalise the appointment of the project team and the time allocations for the team members. Only contracted work can be funded, thus the service contracts need to be in place before the project to be able to work on full volume.

Competent staff

To achieve the ambitious targets, the project needs to have team members with high expertise on the subject areas. The project manager appoints the personnel based on nominations from the partners, but ultimately the competence of the staff depends on the availability of such expertise during the whole duration of the project.

Use cases

The project needs enough use cases to drive the development towards useful deliveries. At the start of

the project the main use cases are already included in the PaRI plan. The project management will consider making a call for accepting use case proposals. Proposals will be evaluated by the project steering group.

3.2 Project activities

Viruses are submicroscopic infectious particles that can not survive without a host organism. To replicate the machinery of living cells of a host organism is strictly necessary. This however, has the consequence that if we want to study the virus in more detail we need to ensure that no host DNA contaminates the virus sample. In the case of a human host we also need to ensure that the sample is GDPR compliant and we do not expose any sensitive data about the human host organism. Depending on the sequence library capture, it may be challenging to resolve background of host sequences in the viral sample and vice versa, making the annotated raw datasets potentially sensitive for storage and also warranting for repeat reanalysis of the same raw data sets, as our knowledge on virus host interactions improves. GISAID stores only fully processed outputs of complete viral sequences collected for epidemiological purposes. To avoid over representation of any sequence data on GISAID virus phylogeny dashboard by dataset duplication, the data submissions to their repository are monopolized and no data may either fully or partially be submitted or copied to any other public databases. This puts researchers in front of a difficult choice and restricts exchange of the original raw sequence data, rich in additional information, as well as a fall back option of reanalysis, should doubts on potential processing bias driven artefacts arise. One solution to this problem is to offer an easy to use, secure platform where users can upload their data without being published immediately, but instead the platform will check for any host contamination and remove them if needed. The user will be informed about any step taken and ask for confirmation before the cleaned virus-only-data is submitted to ENA or SRA. The patient data, potentially useful for other analysis will be stored in a secure environment, only accessible via special authentication, as in federated EGA. From SRA/ENA the deployed Nordic-Galaxy instance, but also all other Galaxy instances can obtain the data via public APIs and conduct research on it. Cross-referencing of complementary data sets between ENA and EGA is being provisioned.

*** Project milestones are listed in a separate document**

WP1. Data Management

Start date: 2020-11-01 **Resources:** 10 PMs

Partners: NBIS, ETAIS, DTU, UiO/USIT, UiB, CSC (observer)

Objective: Facilitating collection and storage of sequence and other human related COVID-19 data

Task 1.1 Mirroring public viral data to local infrastructures: We endorse all researchers to submit their research samples to public archives, like SRA and ENA for sequencing data or PRIDE for mass-spec data and EGA for sensitive data. Having all SARS-CoV-2 data freely available, allows for real-time analysis and tracking of the virus, but also enables an unprecedented turnaround in analysing the genome and discovering new drugs or vaccines. However, the pure amount of data can be challenging to deal with and downloading 3000 (at time of submission) sequences can be time consuming and error-prone on its own. To make it very convenient for Nordic researchers to access all public resources at once we will set up a system similar to the one of the European Galaxy server[8] and

[8] https://bit.ly/usegalaxy-eu-covid19-data

mirror all public data locally in our Nordic clouds. This will also address the issues of Findability and makes the data FAIR.

Task 1.2 Sensitive patient data: There is also COVID-19 host data that is sensitive. Here, the Federated EGA provides a solution, and NBIS (ELIXIR-SE) has set up the EGA-SE node that is technically ready and able to receive sensitive COVID-19-related data for GDPR-compliant controlled access. NBIS will help users to FAIRify the data in order to achieve highest possible usability and visibility (cf. WP3). NBIS will also contribute to the further development of federated EGA, building upon our strong team of systems developers that have been engaged in federated EGA since several years. The Nordic countries have long-term collaboration within ELIXIR that will be important for fast and agile systems development and COVID-19-related additions. The work will be performed in close collaboration with ELIXIR-CONVERGE extension). According to the current information, multiple teams in Estonia have filed requests for access and ethics committee approval to analyse SARS-CoV-2 clinical diagnostic sample frozen material and couple this information with patient data in genbank, if available for those individuals. Computerome (ELIXIR-DK) has set up a secure cloud using Cloudera platform to acquire multiple datasets for pandemic related data and will to make them FAIR and available in a controlled GDPR-complaint access

Deliverable (D-1) 2021-06: Infrastructure components providing storage both for sensitive data (federated EGA) and non-sensitive data (European COVID-19 data portal; ENA) and support to users enabling them to FAIRify data so that the new data becomes useful and visible.

WP2. Facilitating Nordic analysis of Nordic pandemic data using Galaxy and a research e-infrastructure in the Nordics

Start date: 2020-12-01 **Resources:** 14 PMs

Partners: UiO/USIT, UiB, de.NBI, ETAIS, Sigma2 (observer)

Objective: Obtain access to all Nordic sequences, collected in WP1 and establish a flexible and functional interface with existing tools set up for analyses of COVID-19 genomic data and undertake evolutionary mapping of all available data. Use that as a backbone to facilitate further samples to be mapped and analyzed in the context of Nordic strains, and to also map Nordic interchange[9].

Task 2.1 Nordic Pandemic Research Infrastructure: Deploy a Nordic Galaxy pandemic portal. This will use the Open Infrastructure based deployment by the European Galaxy team. For HPC backend interaction, a pulsar server will be deployed on at least two Nordic clusters, section 2, and on virtual clusters on NREC cloud. ELIXIR-Belgium designed a software for uploading data from Galaxy to ENA [10]

Task 2.2 Analysis of Nordic COVID-19 genomic sequence data: Deploy all the workflows that have been developed by the Galaxy community as a response to the COVID-19 threat into a Nordic Galaxy pandemic portal. Some steps of the workflows can be run using Galaxy interactive environments, similar to https://live.usegalaxy.eu, where interactive tools are secured in containers.

Deliverable (D-2) 2021-08: A Nordic platform to analyze all COVID-19 genomic data using Galaxy within a research e-infrastructure in the Nordics, to develop a resource to undertake evolutionary mapping of all available data, and use the results as a backbone to facilitate further samples to be mapped and analyzed in the context of Nordic strains, to also map Nordic interchange, and to perform a virtual screening of the SARS-CoV-2 main protease.

[9] https://nextstrain.org
[10] https://github.com/ELIXIR-Belgium/ena-upload-container

WP3 Secure cloud infrastructure for sharing and presentation of pandemic research datasets and results

Start date: 2021-02-01 **Resources:** 12 PMs

Partners: NBIS, ETAIS, UiO/USIT, DTU, de.NBI

Objective: To deploy a cloud environment that would be treated in principle as a federated system for sharing both public and sensitive information as well as providing the possibility to compare multiple epidemiological programs in the course of data analysis. Help users to FAIRify data so that they become usable, and help users to publish data in international repositories.

T3.1 Secure cloud: Design a blueprint for a cloud solution for use in the current COVID-19 outbreak that will be available for future outbreaks regardless of their size, using trusted data and secure standardised tool sets to enhance the analyses with the user's own data and customized methods.

T3.2 FAIR data support: Provide a set of tools and a support channel to help users FAIRify data before submission to the European COVID-19 portals[11][12] (ENA for viral nucleotide sequences) and federated EGA nodes for sensitive host data. In addition, harmonise data from WHO

T3.3 Nordic strains dashboard: Deploy a dashboard for Nordic strains, similar to GISAID[13,14] but with improved precision of location mapping (municipality of residence would be desirable), date stamped and annotated with age and gender of the affected individual. Ability to trace existence or lack of transmission of a strain between specific regions would be of future value when evaluating adequacy of lockdown measures.

Deliverable (D-3) 2021-10: A cloud in two different national e-infrastructures to scale the analysis. To help users FAIRify data before submission to the European COVID-19 portals, a set of tools and a support channel will be provided. For data presentation a dashboard of Nordic strains will be deployed

4 Organization

4.1 Project Organization

This section outlines the project organisation.

Project management
Project management is carried out by the project manager. The project manager is responsible for task assignments to the project team members and reporting project status/progress to the steering group. Partners who are main contributors to different WPs/milestones naturally take the lead in coordination of the associated WPs/milestones.

Project team
The project team consists of persons that are assigned to the project work in agreement with the project partners, project manager and NeIC. The assignment is done through a service contract between the partner and NeIC. At the start of the project altogether 11 persons belong to the project

[11] https://www.covid19dataportal.org/
[12] https://www.ebi.ac.uk/ena/pathogens/covid-19
[13] https://www.gisaid.org/epiflu-applications/next-hcov-19-app/
[14] https://nextstrain.org

team (including WP leaders)

Steering group
The steering group is formed by representatives of all project partners and the project manager, with the NeIC representative being the project owner and the chair of the steering group.

Reference group
Reference group is formed from expert members as decided by the project management team and the project SG. The reference group is chaired by the project manager, and other project participants are the project manager and the project owner.

4.2 Authority and responsibility

The steering group of the project has authority to modify the project - meaning that the steering group can approve changes to project plan and schedules. They are responsible for steering the project and approving its results. The steering group will supply the necessary resources and support the project manager by ensuring that he has the means for leading the project.

The project manager is responsible for carrying out the project within the framework of project directive, project plan and assigned budget. The project manager reports to the steering group. The project manager is responsible for costs in the project and is authorised to decide on the use of the budget. The project manager has authority to manage the allocated resources.

5 Working methods

PaRI working methods will conform to NeIC policies.

5.1 Requirement dialogue

Reference work will be carried out as necessary between teams and recipients of specific deliverables. Leaders of WPs should be kept apprised of developments in these reference dialogues, and should report them at a suitable level of detail to the project management meetings.

5.2 Delivery and transferral

Delivery and transferral are made on project management team decisions on recommendation of the WP leader. Validation of requirements fulfillment is carried out by the respective named recipients. The validation report is used as the basis for decision of approval of delivery/transfer in the following steering group meeting.

5.3 Production models

Production models differ between deliverables, as best suited working practices can be selected in each of them. Common to all deliverables is that sufficient number of online team meetings are held, so that the work is coordinated and done as a team effort.

5.4 Monitoring and learning

Milestone tracking is reported to each steering group meeting. Milestone achievement is used as a KPI for quarterly reports of project progress and quality of governance/management to the NeIC board.
Management will propose changes and additions of milestones for decision by the steering group. The ambition should be to have milestones defined such that they can provide a stable guide for the desired state of the project over the coming year.

5.5 Change control

In case of occurred or planned deviations from previously decided plans or schedules, the status will be reported to the deciding body for follow-up. This can include a decision to revise or adjust the previous decision accordingly, and/or other mitigations even outside the project.

Minutes are taken from each of the management and steering group meetings, so that decisions can be followed and are referable.

5.6 Risk management

Risks and opportunities identified in day-to-day work shall be reported to project management at a suitable level of detail. The project manager will have the primary responsibility for engaging with stakeholders, and following up on developing risks and opportunities.
Similarly, all the WP leaders will follow the internal progress and external environment, and inform the project management without delays on risks and issues.

The severity of risks and mitigation alternatives are considered at the project management level.
Risks that threaten to have a major impact on the project's stakeholder relations or its ability to deliver in a timely fashion will be reported to the steering group without delay.

5.7 Information distribution

PaRI will strive for maximal openness and transparency, by conforming to the NeIC policies for:

- Openness
- Information
- Shared personnel information
- Stakeholder engagement monitoring

5.8 Document management

For document authoring PaRI will use Google docs, which has collaborative authoring and integrated incremental backup features. The same exceptions apply as above for sharing of information.

5.9 Quality assurance

Quality assurance of results in deliveries and use cases will be carried out in interaction between the

team and the recipients (who by definition are members of the reference group), facilitated by the project manager (who chairs the reference group).
Quality assurance of project management is carried out as part of project governance, and in regular meetings between project manager and project owner.
Quality assurance of governance is carried out by the project owner following the NeIC Steering group checklist template, and as part of reporting to the NeIC board.

6 Benefit Risks

Benefit risk	Measures during the valuation
Low interest/ no involvement from health institutes and biomedical research institutes	Persistent networking, meetings and demos
Defining synergies with the ELIXIR efforts (CONVERGE WP7) and starting a collaboration	Setting up grounds for collaboration (meetings, MoUs, etc)
Trust building with researchers and health institutes for data collection	High-level discussions
Availability of key PaRI personnel	Dialogue with National Providers
Project funding and staffing	Dialogue with project partners
Political will of providing datasets	Dialogue with researchers and health institutes
Low political will from countries not participating in PaRI	High-level discussions

7 Project cost estimate

Costs

Prerequisites

- Manpower available for PaRI as agreed prior to project start in the Collaboration Agreement

Valuation

Rationale:
The cost is estimated as the direct cost of participating in the project by all the partners (monetary and in-kind contributions). The other project costs include travel and meeting costs, as agreed in the Collaboration Agreement.

Continuous costs in operations (increased annual cost)	Valuation [FTE]

Access to compute resources for analysing pandemic data. Setup and configuration of compute resources to accept workload from the user front-end	0.4
Streamlined, shared and up-2-date pandemic analysis tools and workflows (mainly as containers), with addition of a user-friendly front-end, Galaxy portal, wherever applicable	0.4
Infrastructure services providing storage both for sensitive data (Federated EGA) and non-sensitive data (European COVID-19 data portal; ENA)	0.3
Promoting cooperation between project partners and Nordic e-Infrastructure providers efforts related to pandemic research data	0.2
Help Nordic users to publish pandemic research data in public databases in a FAIR manner -> better reusability and more citations	0.4
Centralized collection of SARS-CoV2 (and other pandemic data) distribution virus fingerprints processed in the same standardized manner and referencing to host data availability whenever possible, and expose the fingerprints to a nordic dashboard for viral strain transmission tracing	0.3
Competence building in pandemic data handling and strengthening the position and international visibility of the Nordic countries by showcasing good Nordic use cases (of versatile types)	0.5
Synergies with other initiatives in timely addressing of the pandemics	0.3
Building trust and collecting raw SARS-CoV2 data	0.1
Total per year:	2.9
Project costs (non-recurring costs)	Valuation
<Cost>	
Travel, meetings and other costs	0.18
Total amount:	3.08

Total

Benefit value period:	5

Benefit		Cost	
Quality, time, money	66	Change of operations and Project cost	6.16
Uncertainty	20	Uncertainty	0.3
Total net benefit:	58.84		
Net benefit per year:	8.12		

Resource needs

Resource types	Scope
PaRI Project group, including possibly legal advice	Contribution needed to realise non-political benefits
PaRI Steering group	Contribution needed to realise political benefits
PaRI Reference Group	Contribution needed to help the Steering and Project group with realising benefits

Benefit realisation

Benefit realisation is done according to the NeIC Business Benefit Realisation Management policy. The PaRI Project Owner together with the PaRI Steering Group is in charge of PaRI Benefit realisation.

Achievability

The achievability will be evaluated and monitored through the Benefit realisation management process.

Financing

Funding of the costs linked to the expected benefits is secured through Collaboration Agreements between the partners.

Edition history

Edition	Date	Comment
Draft for the consortium	2020-10-27	
Draft for the SG	2020-11-13	

Appendices

No.	Document name	Document designation/Id

References

No.	Document name/ designation/Id	Edition, date
1	Business case	Approved