

PUHURI Deliverable D1

Project Title:	Puhuri
Public	This document is PUBLIC
Deliverable title:	Requirements and architecture plan for Puhuri
Contractual delivery date:	30th of October 2020
Status	v. 1.2
Actual delivery date:	09th of November 2020
Updated	24th of November 2020
Partner(s) contributing to this deliverable:	FI, SE, EE, NO, DK, IS

Authors and Contributors:

Jarno Laitinen, Monica Lassi, Anders Sjöström, Kent Engström, Ahti Saar, Ilja Livenson, Marina Adomeit, Juha Nyholm and Kalle Happonen

Table of Contents

Executive Summary	3
1 Puhuri Use Cases	3
2 Technical architecture	4
2.1 High level architecture and the key processes	4
2.2 Puhuri AAI	5
2.3 Resource allocation	7
2.4 LUMI Integration	9
3 Requirements collection	10
3.1 National Portal integration	10
3.2 AAI Requirements	14
3.3 Resource Allocation requirements gathering	17
3.4 LUMI Integration requirements gathering	17
4 Interoperability	18
4.1 Technical interoperability	18
4.2 Semantic interoperability	18
4.3 Organisation interoperability	18
5 IT Service Management	19
5.1 Helpdesk	19
5.2 Sustainability	19
5.3 Contracts	19
5.4 GDPR	20
5.5 Accessibility	20
References	21

Executive Summary

This is the first deliverable of Puhuri project reporting the work done 1.6.-31.9.2020. The project started by designing the technical architecture and gathering the requirements from different stakeholders. Also the first terminology for Puhuri was composed.

LUMI supercomputer's integration to Puhuri is the priority use case.

1 Puhuri Use Cases

Puhuri organises quarterly Reference Group meetings to discuss with the stakeholders. Invited are for example the LUMI countries, [EOSC-Nordic](#) project and NeIC project [Nicast2](#). For example with EOSC-Nordic WP5 - Open Research Data and Service, we have been discussing Galaxy portal integration challenges to an HPC environment.

EuroHPC Joint Undertaking (JU) is acquiring pre-exascale and petascale supercomputers (the EuroHPC supercomputers) which will be located at and operated by supercomputing centres in the Union. [LUMI](#) [LUMI] is one of those. It will be hosted at CSC in Kajaani data center. LUMI is the main use case for Puhuri and thus the prioritisations of the requirements and the schedule of the project is steered according to that. LUMI consortium involves the Puhuri countries, Belgium, Czech Republic, Poland and Switzerland. Puhuri is implementing the functions belonging in Puhuri's scope. See the national integration chapter below for the details of this requirements gathering process.

2 Technical architecture

2.1 High level architecture and the key processes

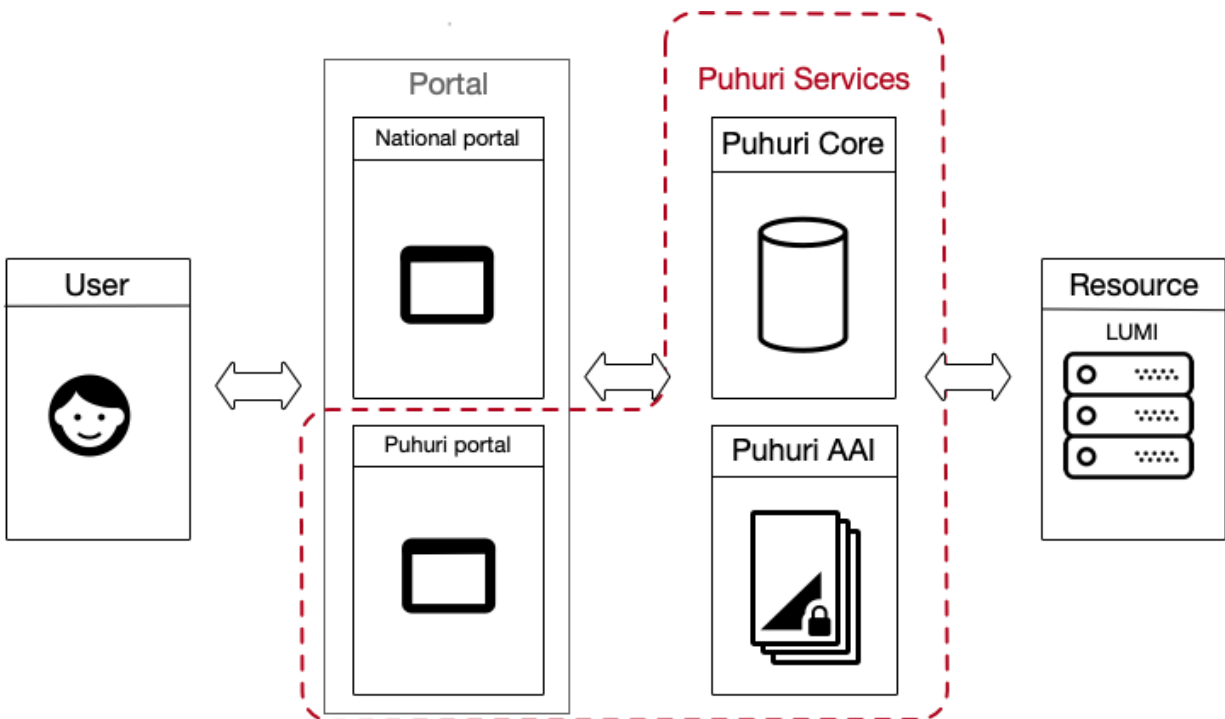


Figure 1. Puhuri's high level architecture.

Figure 1 presents a high level architecture of Puhuri. User either uses a National portal or a Puhuri Portal for the Project Application and the Project membership management.

User is also using Puhuri AAI to register Puhuri User Account and to login to the Resources. Puhuri Core is the database storing the Projects and their Members, Allocation, Accounting and related Resource information. A Resource (in this figure, LUMI) uses APIs to import the information from the Puhuri Core.

The Puhuri Portal gets User information from the Puhuri AAI when the User authenticates by using home organisation's identity provider, which can hand over certain attributes regarding the User.

2.2 Puhuri AAI

Puhuri AAI is one of the essential Puhuri Service components as it enables Users to securely access to Resources such as LUMI. Work within the Puhuri project AAI task started with defining technical and organisational requirements for Puhuri AAI, that are presented in section 3.2 AAI Requirements. Based on these requirements, the task defined Puhuri AAI architecture and proposed a solution that is presented in this chapter.

A high level image of Puhuri AAI that is compliant with AARC blueprint architecture is presented in Figure 2, and a brief description of components is as following:

- AAI Proxy is functioning as SP-Imp connecting:
 - Puhuri Resources as SAML Service Providers and/or OIDC Resource Providers.
 - SAML R&E Identity Providers available via eduGAIN and optionally other Identity Providers to fulfill specific requirements of LUMI participating countries.
- Account registry and web portal that are used to register or provision Users Puhuri Accounts and in that process assigns to a unique identifier. It has additional capabilities such as linking various user identities, enabling users to manage a subset of their attributes etc. Web portal can be also used for registering Resources, depending on the process that Puhuri will define. Some of the Users do not have an Imp to provide the attributes regarding the User.
- Discovery service that provides a web interface for Users to search and select their preferred identity provider.
- The Metadata Service that aggregates the metadata of all the SAML Identity and Service Providers that are connected to the platform.

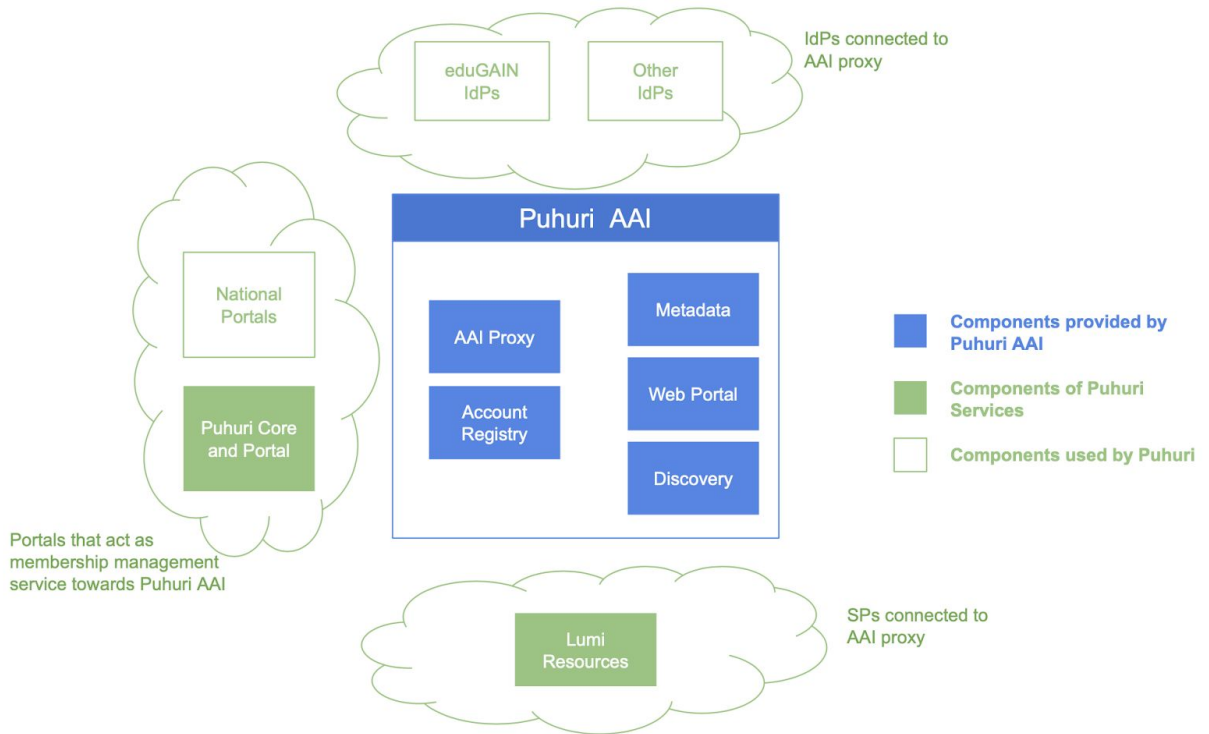


Figure 2: High level image of Puhuri AAI

To implement and run an AAI solution would need a number of AAI experts, which are hard to find. From the start of the project it was clear that the 2 years lasting Puhuri project doesn't have time and capability to implement and run the AAI services itself so outsourcing to a reliable partner is the only feasible solution. Prominent provider that is delivering this solution to various research communities, including FENIX, is GEANT via its eduTEAMS solution. eduTEAMS is fully compatible with AARC BPA, and delivers dedicated AAI solutions that can be to certain degree custom tailored. Puhuri AAI activity has been collaborating with eduTEAMS to explore the feasibility using this service, and as a result here is mutual understanding that AAI requirements of PUHURI can be fully fulfilled by eduTEAMS solution, and beyond. To that end, initial testing was performed with eduTEAMS multi-tenant service instance.

Moreover, since eduTEAMS is already operating FENIX AAI in production, a collaboration between Puhuri, eduTEAMS and FENIX has been established in order to explore the possibility of using a common AAI. A series of meetings was held to explore this possibility and in the moment of writing this deliverable, there is a strong belief that this is technically feasible. Further discussions are needed regarding the liabilities for using shared AAI infrastructure, GDPR roles and other its requirements, and the governance of this shared infrastructure in general. Sustainability has to be ensured. Section 3.2. tells about the requirements collection process. The AAI functions are prioritised based on the requirements.

2.3 Resource allocation

In Puhuri project, Waldur software is being used for resource allocation and accounting. Waldur is an open-source cloud marketplace with a self-service environment for users to request and get access to resources.

Waldur was used in NelC's Dellinger project for testing cross-border resource sharing between Nordic countries and outcome was positive. Same software is used in Puhuri Core as well and it is developed and operated by the Puhuri project member organization, ETAIS/University of Tartu.

In Puhuri, we had to adjust Waldur platform: release installation was automated using GitLab CI and Ansible, now it is used with Helm Packaging for Kubernetes. In total, we run three different k8s-based deployments for Puhuri Core:

- Demo (<https://my-demo.lumi-supercomputer.eu>)
- Test (<https://my-test.lumi-supercomputer.eu>)
- Production (<https://my.lumi-supercomputer.eu>)

Deployments contain integration with test eduTEAMS system data and implement the business processes we have decided so far in the Puhuri project. The whole deployment process is designed to allow for dynamic changes into the processes during the stabilization phase.

Puhuri Core is the heart of the resource allocation as it holds necessary information about users, group management, roles, resources, accounting etc. Puhuri Core is accessible using REST API with target users of API being:

- National allocation portal;
- Service Providers that want to use allocation information via Puhuri Core

If the country does not have a national allocational portal or is not willing to integrate at once, Puhuri partner ETAIS is planning to provide an option for setting up a reference national portal either as a managed version or as a supported software package. Details of policies and business model of such an offering is to be decided at a later stage.

Puhuri Core users are organizations which represent participating countries (FI, BE, CH, CZ, DK, EE, IS, NO, PL, SE) and these organizations can allocate national shares. Representatives of the allocator (review team) can make decisions on providing allocations to each of the projects.

Puhuri Core users hence have the following roles:

- "Allocating body member" - able to create projects and allocate resources according to the Country's review board.

- Project admins - PIs
- Project members - all users, non-PI

Puhuri Reference portal users will have additional roles to reflect established processes of resource allocation. The roles will be clarified after the stabilization of Puhuri Core, so far envisioned ones are:

- Organization owner: represents a scientific organization or community, able to create new projects, manage project teams and manage requests for resources.
- Project manager (PI): able to request and use provisioned resources (with optional approval by Organization owner), can add project members from the same organization (i.e. pre-approved by the owner).
- Project member: can only participate in particular project work, able to use provisioned resources.
- Portal support: can provide support to users of the platform. Could potentially be linked with a national helpdesk.

In LUMI use case there are Allocation Units, which are consumed per hour. Those are then allocated to a Project by the previously mentioned Allocators. However, Resources need to specify many Quotas. Part of them are static, which are not due to be changed on Project basis such as quote for a home directory of a User in LUMI. A Dynamic Quota is for example the Project storage quota i.e. what is the largest amount of data the Project can store on a certain LUMI parallel file system.

The amount of required storage space and computation resources cannot be always anticipated in advance by the Project. It is then either Resource and also perhaps Call specific question, how to react to such a request. Thus, the decision may depend on the Resource, which might have technical limitations as well, and the fair share principles driven policies, which are typically handled by the Allocator processes.

To avoid handling the common quota requests for each Project via LUMI HelpDesk, there should be a way to apply for those also in Puhuri Portal or via the National Portals. However, the quotas are sometimes very Resource specific, which makes it challenging to manage and tailor Puhuri Core (database) and the different portals.

2.4 LUMI Integration

LUMI will be the first Resource that will integrate with Puhuri. Later on, Puhuri will also cater other Resources in addition to LUMI. Puhuri will have to define the rules on how to integrate such Resources. These Puhuri rules must align with the GEANT Code of Conduct, and also take into account the GDPR with regard to the information minimization principle.

CSC is implementing a mechanism for scheduled data synchronization from Puhuri Core REST API to its identity management system (IDM). This scheduled data synchronization will be used to pull in User, Project and Allocation data from Puhuri Core to CSC IDM.

The LUMI services will be able to consume the imported User, Project and Allocation data by querying the relevant information from CSC's LDAP cluster.

CSC will set up the data import mechanism during November 2020 - January 2021. In October 2020 we already have a proof of concept implementation of User import available, where Users are pulled in from Puhuri Core to CSC IDM.

CSC will also implement a service for exporting LUMI Accounting information to Puhuri Core. This will be done later in 2021.

The risk assessment will be done. Data privacy is done from the User's perspective (see GDPR Section 5.4). The data related security risks are typically divided into confidentiality, integrity and availability.

It has not yet been decided how the Resource (here: LUMI) specific Terms of Use will be accepted: is it part of Puhuri, a Resource Specific process or the (National / Puhuri) Portal. The User and Project lifecycle have to be defined and implemented. Will the Resources enforce their own User and Project lifecycle policy on the imported User and Project data, or will Puhuri provide a policy for User and Project lifecycle, which can/should be implemented by the Resources. The authentication methods including the multi-factor authentication (MFA) have to be defined and LUMI has to define policy regarding the identity assurance and what attributes it will require from the User. Puhuri has been interacting with the LUMI consortium in various ways and occasions.

3 Requirements collection

3.1 National Portal integration

The goal of these activities is to develop a set of requirements for the APIs that will connect the portals (national portals and Puhuri portal) with Puhuri Core. There is also an effort to develop the set of requirements for the API connecting Puhuri core with the resource (i.e. LUMI) but this effort is not included in the national portal integration.

The figure 3 below describes the processes of our work. The different steps in the process are described below the figure. The second stage consisted of three interdependent activities concerning nomenclature, information modelling and use case scenarios.

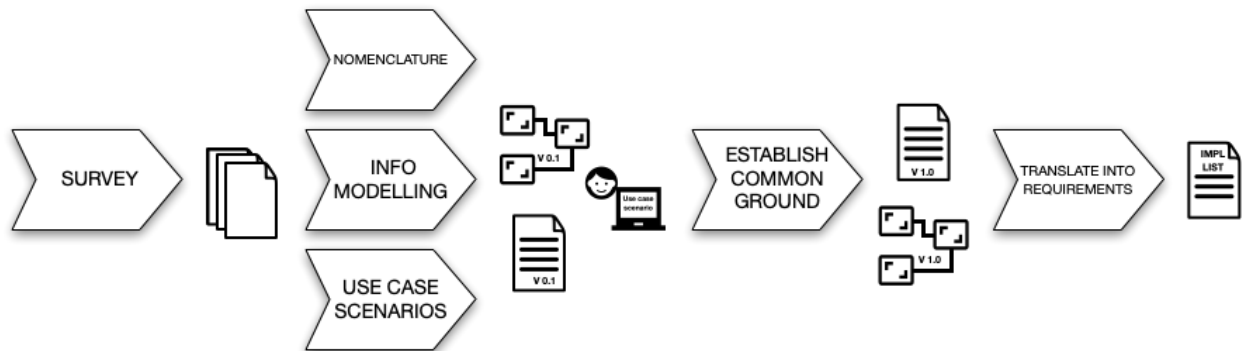


Figure 3. Requirements collection process intended to capture the needs, and requirements of the included parts and actors of the Puhuri project and LUMI consortia.

The first survey was sent out during summer 2020, and responses were collected throughout early September. After that we did an interview round with another questionnaire document. We have analyzed the responses, focussing in particular on the information objects and functionality of the national portals. The responses have shifted in detail, and will in some cases need to be complemented in the next stage.

Nomenclature

The survey responses pertaining to information objects and definitions were gathered in a spreadsheet to form a list of terms and terminologies. Three Puhuri project members went through the list of terms and determined which terms would be relevant for the information model, and to find commonalities between (a) the national portals' definitions, (b) the definitions used by CSC and LUMI, and (c) the emerging architecture of the Puhuri Core. For concepts represented by multiple terms (synonyms) one term was chosen as preferred term, with a reference to it from the non-preferred terms.

The spreadsheet was then checked by a number of Puhuri project members, who made suggestions and comments regarding definitions and preferred terms. The spreadsheet has separate tabs for the identified terms, terms added by or commented on by the Puhuri project members, and the selected terms and definitions.

Finally, the complete list of chosen terms and their definitions were gathered in a document with terms and corresponding definitions, forming the nomenclature for the Puhuri project (See Doc1: Nomenclature, in Key Results below). The format of the nomenclature document can be seen in Table 1. The nomenclature document will be continuously updated, retaining the earlier versions decided by the Puhuri project.

Term	Definition
AAI	Authentication and Authorisation Infrastructure (AAI) (also used alternative term IAM). A set of various services to register a user to get a Puhuri ID, authenticate (i.e. to get the user identity) to use the integrated Resources (services), and provide means for the resources provider to grant access to the eligible users.

Table 1. Example of the outline of the selected terms in the nomenclature document.

This work on the nomenclature was important to make sure that the project members have the same understanding of the concepts and information objects central to the Puhuri project. A potential consequence if this is not done, is that there will be semantic interoperability problems and challenges in communications in the Puhuri project and with the stakeholders of the Puhuri project and of the LUMI consortia.

Information modelling

The spreadsheet mentioned in the preceding subsection was used to create the information model of Puhuri (see Doc2: Information model, in Key Results). The objects in the information model were decided in a discussion during a weekly meeting. The descriptions of the information objects have been updated to conform with the definitions in the nomenclature document.

The information objects were visualized in a diagram of the information model, shown in Figure 4, detailing the relations between each object. Further, each information object's attributes, including where the information comes from, and other relevant information is (or will be) included in the information model document. All information objects are described in a table as can be seen in Table 2.

Information object ID	Name of information object	Description	Example value
<number>	<name>	<text>	<value>

Table 2. Example of the outline of the selected terms in the information model document.

For each information object, there is a description of its attributes, with example values for each attribute, including whether the value is imported from another source: The modelling language used in this work is that of IRM (<https://www.irm.se/>, see [LIN]).

The information model was decided upon through multiple discussions among the Puhuri project members. It is to be noted that the model is mainly focussed on the Puhuri Core part. This is due to the fact that the National Portals are free to implement their portals to suit their needs and processes. The Puhuri project does not restrict how the National portals are constructed other than in the requirement of being able to use the supplied API in the communication with Puhuri Core.

Scenarios of use

The third document contains scenarios of use, using a particular formal description design method called “use cases”. The scenarios of use capture different actors’ interactions with Puhuri Core, or scenarios in which Puhuri Core is involved. For example, in user registration, access of allocated resources etc.

The scenarios contribute to gathering functional requirements by facilitating the identification of the various processes, subprocesses and their requirements involved in each scenario, through a structured methodology described in [LAS].

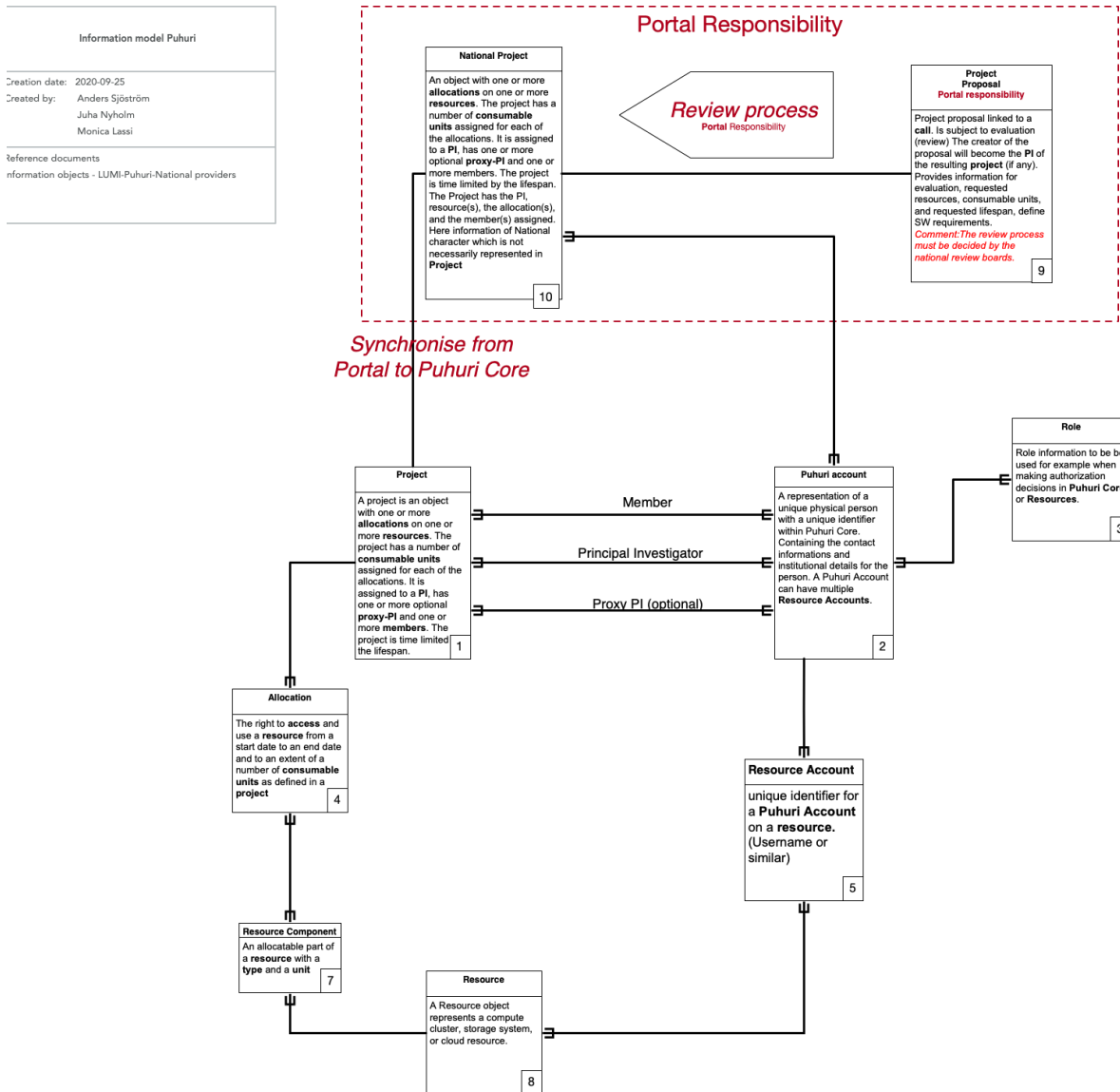


Figure 4. Graphical representation of the Puhuri information model.

How the three interdependent activities were conducted together

The documents described above form an interdependent set of documents describing the Puhuri project in terms of processes, nomenclature, and design. The documents and the processes they represent are interdependent as outlined in Figure 5 where a change in either of the documents spawn a change in each of the other.

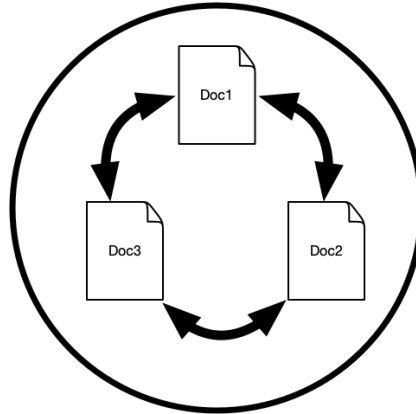


Figure 5. Interdependency between the documents of the national portal integration requirements collection. The process aggregates and concatenates the information gathered to create a homogenous nomenclature across all of the included parts and actors of the Puhuri project and LUMI consortia.

Within the process of defining the sub-task deliverables, we have identified the following parts:

- Doc1: Nomenclature. Status: **Submitted** Link to doc: [Puhuri Nomenclature v1.0](#)
- Doc2: Information model. Status: **Submitted** Link to doc: [Puhuri information model objects v1.0](#)
- Doc3: Use case scenarios. Status: **In development** Link to doc: [Use case scenarios v1](#)

The next step is to supplement these documents with the list of functional requirements and to complete the writing of the use case scenarios to specify the functional requirements. It must be stated that it is a risk if the implementation is done before the functional requirements have been delivered. Hence, the writing of the use case scenarios must be a priority.

3.2 AAI Requirements

Requirements for Puhuri AAI were produced in collaboration with other relevant tasks of Puhuri project, AAI specialists and feedback received from national federation operators and LUMI participating countries. At the moment of writing this deliverable, ongoing are the interviews with LUMI participating countries in order to understand national AAI capabilities, so these requirements may still change to reflect those. In addition, consultations were held with FENIX that has already its AAI solution in place with similar requirements. The requirements are for easier reference organised in thematic groups and are presented in the following of this chapter.

Scope

Puhuri AAI should provide access to LUMI resources, but should also have the ability to support other compute resources delivered via NeIC.

Capability of managing project applications and users membership and role attributes are supported by other service components and are outside of the scope for Puhuri AAI.

Standards

Puhuri AAI architecture needs to be compliant AARC Blueprint Architecture¹ that defines interoperability requirements for AAls for research infrastructures should be built, and that is widely used by Research Infrastructures (eg Fenix, EOSC etc).

Connecting Identity Providers

Ability to connect SAML and OIDC Identity Providers so that user communities of LUMI participating countries are represented. This includes:

- Identity Providers that are members of R&E Identity Federations serving LUMI participating countries.
- Additional Identity Providers, that may be requested by a LUMI participating country such as HPC IdPs
- eIDAS Identity Sources for which integration with EDSSI as a solution for academia and research should be pursued long-term.
- In addition to this, a solution will need to be explored by Puhuri to provide a solution for IdP of last resort with appropriate Identity Proofing. (This solution is by design not part of Puhuri AAI),

At the moment of writing this deliverable, surveys with LUMI participating countries are ongoing, so that we can understand which IdPs would be connected to Puhuri AAI.

Connecting Service Providers

Ability to connect SAML Service Providers and OIDC Resource Providers, by employing an approval process defined by LUMI.

Puhuri Account

Ability to register Puhuri Accounts, following the AARC profile for expressing community identity attributes. Below is the initial set of identified mandatory and optional attributes:

- Required user attributes:
 - o the name of the user
 - o the e-mail address of the user
 - o the Puhuri user identifier
 - o the memberships and roles of the user in the Puhuri AAI
- Optional attributes:
 - o the user's home organization,
 - o the affiliation of the user in her home organization,
 - o the affiliation of the user in the user context of the Puhuri AAI
- User need to be able to self-manage certain attributes such as email address.

¹ <https://aarc-community.org/architecture/>

- An authorized role within Puhuri need to be able to suspend a compromised Puhuri Account.
- Ability for user to link their different accounts. This may be needed in cases where the user changes organization that he/r is affiliated with or when s/he is affiliated with multiple organisations.

Attribute release

Puhuri AAI should be able to control which attributes are released per service.

Assurance Profiles and Authentication

Puhuri AAI will need to support usage of REFEDS Assurance Profile [RAF] and Multi-factor Authentication as a centrally provided function when an Identity Provider is not able to provide such capability. Resources should be able to signal their requirements when it comes to accepted identity assurance profiles and authentication type.

At the moment of writing this deliverable, the interviews with LUMI participating countries are on going to understand what are Assurance Profiles and Authentication methods supported by Identity Providers from their constituency.

Then the decisions have to be made which Identity assurance profiles and authentication methods that are considered to be sufficient for accessing LUMI resources.

Support for Entity Categories

SP side of Puhuri AAI Proxy will need to be able to support REFEDS R&S and CoCo entity categories, so that IdPs can make use of these entity categories for making decisions regarding attribute release.

Security

Puhuri AAI needs to support REFEDS [SIRTFI](#) (A Security Incident Response Trust Framework for Federated Identity) framework and needs to have a security incident support team that can support security incident resolution.

Policy

Puhuri AAI needs to conform with the GDPR requirements.

The User and Project life-cycle management has to be defined. The question is, what and where is implemented regarding that: Puhuri or a Resource.

Terms of Use for the Resources might need to be accepted by the User before accessing a Resource.

Service Levels and Infrastructure

Mission critical components such as AAI Proxy will need to be operated in high availability. Separate production and quality assurance/testing environments is a requirement. Details of

these including the availability and support levels will need to be agreed with Puhuri AAI provider.

3.3 Resource Allocation requirements gathering

To align with all LUMI countries and LUMI, national requirements collection has been carried out as described above. The goal was to get adequate status of the current resource sharing options in each particular country (from technical details to policy). As well as what are the national requirements connecting their national portals to LUMI.

The resource allocation unit discussion for different Resources will be planned and discussed. Those will affect the data model on Puhuri Core and at the Resource side. We aim to demonstrate the Puhuri portals on suitable occasions for example to the Puhuri Reference Group.

3.4 LUMI Integration requirements gathering

LUMI Integration needs to wait for further decisions regarding the authentication and authorisation. For example object storage and command line tools are more challenging than the web portal based services. The process to accept the Terms of Use for a Service has to be defined within the collaboration between LUMI and Puhuri.

LUMI Accounting and reporting will be clarified before LUMI starts its operations. The data will be exported to Puhuri Core that the national portals will be able to produce the accounting reports for the Projects.

4 Interoperability

We did an interoperability evaluation with EOSC-Nordic WP 3 Task, which was writing a deliverable (not published yet) regarding the interoperability interviewing different projects and infrastructures. That checklist was based on the European Interoperability Framework. The [EOSC Interoperability framework](#) is in a draft document state. Puhuri does not deal with the User data management, which is the main focus on the interoperability frameworks.

4.1 Technical interoperability

Puhuri will be based on several components so data exchange, easy and secure interfaces and usability for the User are some key design factors. EOSC Architecture Taskforce has also released [a draft document](#) for EOSC AAI Architecture [EAA], which is working on a new version with AARC architecture. It is discussed in Chapter 2.2.

Puhuri is essentially going to help on this to allow using the federated authentication, authorisation and accounting services with the user transparent Puhuri unique identity. We are using the standard protocols as presented in this document, but for example accounting data format still needs to be planned.

4.2 Semantic interoperability

Puhuri maintains a common vocabulary (Chapter 3.1) that also the national portals and organisations that they can help end-users to understand LUMI and Puhuri terminology. For example Principal Investigator (PI) and Project Manager are synonyms.

Puhuri needs to create easy to understand documentation for the relevant stakeholders. Puhuri aims to find generic concepts that many Service Providers could use. Some are not in scope or mandate of Puhuri to decide or enforce. For example, LUMI has to agree on ontology for example regarding science discipline. For that specific field there exists an ISO standard. For that standard one has to decide if the main categories are enough precise.

4.3 Organisation interoperability

Puhuri will need to do agreements for the Services integration and these agreements must comply with the rules that eduGAIN federation and GEANT eduTEAMS use. Puhuri needs a similar [document](#) for the Service Providers, which ELIXIR AAI has written. The document refers to the GEANT [Data Protection Code of Conduct document](#).

FENIX AAI consortium interoperability needs to be addressed to help end-users' possibility to use Services from both FENIX and Puhuri. Technical and political dimensions are need to be taken into account.

5 IT Service Management

The Puhuri project has just started but operations have to be planned well in advance. The Puhuri team is already discussing with the LUMI consortium on the themes mentioned below. Operations of the Puhuri Portal, Core and Puhuri AAI are handled by the teams who operate the Services. They are for example responsible for monitoring and various aspects to keep the services available and secure and to ensure that there are the necessary processes and personnel to manage the services.

5.1 Helpdesk

Puhuri needs to organise the support contacts for the services.

Regarding LUMI use case, Puhuri is a backend service, so Puhuri will be contacted by LUMI staff. LUMI Helpdesk might need information from Puhuri Core directly since otherwise it may be too complex to resolve the problem. Handling of the typical use cases has to be discussed with the LUST team to resolve why a User cannot login to LUMI.

1. National Portal Helpdesk regarding using these portals. Those who will be integrated to Puhuri, will have functions regarding that.
2. LUMI Helpdesk: each LUMI country has also a member in the support team. There is a common LUMI trouble ticket system, which is not integrated to the other trouble ticket systems.
3. Puhuri Helpdesk: Acts as a backend support for 1 and 2. Puhuri Core and Portal admins and GEANT eduTEAMS support are needed to give backend support.

5.2 Sustainability

Soon after the Puhuri started, we began to discuss with Puhuri partners, LUMI consortium and NeIC about the sustainability for funding after the 2-years lasting Puhuri project ends. The operation of the AAI and Resource allocation services need to be funded. There must be also a common Helpdesk and management functions. Before the Puhuri operational phase eduTEAMS service design, deployment and pilot is funded by GEANT project and University of Tartu are partners of Puhuri.

5.3 Contracts

Puhuri's sustainability work will need to work on the contracts and policies, who can use the Puhuri and in what terms and conditions. The integrated Service must have a contract with Puhuri.

Users of Puhuri need to accept terms of Usage. The service providers to ensure the operational and data privacy aspects have been agreed among the parties.

5.4 GDPR

Puhuri needs to comply with GDPR, which amongst other means that appropriate documentation, agreements and practices needs to be put in place. The challenge in Puhuri is that personal data flows across different organisations. Data Controller, who as the name says, defines how the data is used, must have agreements with Data Processors. Both must be legal entities. Puhuri as a project is thus not suitable and neither is NeIC as a funder organisation as it is not involved in the governance or operations of Puhuri. The FENIX collaboration makes the situation also a bit more complex.

The main principles of GDPR is the data minimisation i.e. avoiding to store or distribute unnecessary data, to limit access to the data only for the necessary people and to remove the data when it is not needed anymore. A user may ask to remove the personal data, but contractual or other obligations may prohibit that. A user has the right to get information about what data is stored, where and what is the purpose. A User can ask Puhuri to send all information stored regarding the User. Data life time will be defined for the Personal data including the logs of various services.

In the LUMI consortium there is a Consortium Agreement and other relevant GDPR contracts. The responsibilities between the organisations and actors needs to be defined. Puhuri is naturally a key part there. Thus data security for the whole chain is important to implement.

Data Protection Impact Assessment i.e. User risk based evaluation documentation with planned measures are needed due to the large amount of Users and international dimension. The risks have to be re-evaluated yearly to ensure that they are actual. Technical controls are used to ensure confidentiality, availability, usability and integrity are evaluated per service component.

Data protection policy and information security policy needs to be defined. There are plans and processes to handle the incidents in an effective manner. The data processors such as staff members have to be instructed how to handle the person's data.

5.5 Accessibility

The [Directive on the accessibility of websites and mobile applications](#) applies to public services to which Puhuri falls in. The national legislation orders as of September 2020. The gaps must be written down and communicated to the national authority declaring the planned activities to fix those. The Puhuri portals must [evaluate](#) the status of the accessibility in an accessibility declaration and put that information available as there is the GDPR required information. Users can give feedback related to accessibility and it has to be replied within 14 days.

For example in Finland the public services have to fulfil the A and AA levels of the W3C [WCAG v. 2.1](#) recommendations [WCA].

References

[EAA] EOSC AAI Architecture. V 0.2.

https://docs.google.com/document/d/12l0xVU9oiXqtVqkwrJijj16L-i_YcM_K6SHkNMf4IWE/edit

[EIN] EOSC Interoperability framework v1.0. Draft for consultation. May 2020.

<https://www.eoscsecretariat.eu/sites/default/files/eosc-interoperability-framework-v1.0.pdf>

[LAS] Lassi, Monica (2014). Facilitating collaboration:] exploring a socio-technical approach to the design of a collaboratory for library and information science. Diss. (sammanfattning)

Borås/Göteborg : Högskolan i Borås/Göteborgs universitet, 2014. ISBN: 9789198165418.

<http://bada.hb.se/handle/2320/13583>

[LIN] Lindström, Lennart (2010). Informationsmodellering på IRMs sätt!. Bromma: Information Resource Management (IRM), ISBN: 9789197711913

[LUM] LUMI EuroHPC supercomputer web page. <https://www.lumi-supercomputer.eu/>

[WCA] Web Content Accessibility Guidelines (WCAG) 2.1. W3C Recommendation 05 June 2018. <https://www.w3.org/TR/WCAG21/>