

# DARIAH

|                   |                  |
|-------------------|------------------|
| Short description | DARIAH           |
| Type of community | Thematic Service |
| Community contact | Marica Antonacci |
| Interviewer       | Shaun de Witt    |
| Date of interview | 2018-03-28       |
| Meetings          |                  |
| Supporters        |                  |

## User stories



### Instruction

Requirements are based on a user story, which is an informal, natural language description of one or more features of a software system. User stories are often written from the perspective of an end user or user of a system. Depending on the community, user stories may be written by various stakeholders including clients, users, managers or development team members. They facilitate sensemaking and communication, that is, they help software teams organize their understanding of the system and its context. Please do not confuse user story with system requirements. A user story is an informal description of a feature; a requirement is a formal description of need (See section later).

User stories may follow one of several formats or templates. The most common would be:

"As a <role>, I want <capability> so that <receive benefit>"

"In order to <receive benefit> as a <role>, I want <goal/desire>"

"As <persona>, I want <what?> so that <why?>" where a persona is a fictional stakeholder (e.g. user). A persona may include a name, picture; characteristics, behaviours, attitudes, and a goal which the product should help them achieve.

Example:

"As provider of the Climate gateway I want to empower researchers from academia to interact with datasets stored in the Climate Catalogue, and bring their own applications to analyse this data on remote cloud servers offered via EGI."

| No. | User stories  |
|-----|---|
| US1 | An authenticated user needs to transfer data files from local to remote storage and <i>vice versa</i> . The user may also want to transfer files between different storage services which may require different protocols. Within the current EGI DARIAH Gateway, this is currently performed using the Data Avenue service ( <a href="https://data-avenue.eu/">https://data-avenue.eu/</a> ) developed under the SCI-BUS Project (supported by the FP7 Capacities Programme under contract n°RI-283481). Using the B2 storage services will allow users to use different types of storage at different stages of the data lifecycle. |
| US2 | The EGI DARIAH Service manager would like to make use of better supported data transfer technologies, in order both to improve user quality of experience and ensure long term support for the data transfer protocol.  |
| US3 | The EGI DARIAH service manager wants to be able to make use of cloud infrastructures to support an increased number of users. These cloud infrastructures could be running docker images or create hadoop clusters  |
| US4 | Users would like to be able to do large scale data analysis on existing accessible data sets through the DARIAH EGI Gateway   |
| US4 | Users would like to be able to use docker containers based applications on the EGI federated cloud  |
| US5 | The EGI DARIAH service manager would like to allow data distribution across multiple sites for easy data storage, discovery and access.   |

Note: Since this is an enhancement of an existing service available through EGI, there is no real additional user story

## Use cases



### Instruction

A use case is a list of actions or event steps typically defining the interactions between a role (known in the Unified Modeling Language as an actor) and a system to achieve a goal.



Include in this section any diagrams that could facilitate the understanding of the use cases and their relationships.

| Step | Description of action            | Dependency on 3rd party services (EOSC-hub or other) |
|------|----------------------------------|--|
| UC1  | Replace Data Avenue with B2STAGE | None   |
| UC2  | ...                              |  |
| ...  |                                  |  |

## Requirements

### Technical Requirements

| Requirement ID | EOSC-hub service | GAP (Yes /No) + description   | Requirement description  | Source Use Case | Related tickets  |
|----------------|------------------|---|--|-----------------|--|
| RQ1            | B2STAGE          | POTENTIAL:<br>Need to check whether EGI DARIAH AAI certificates are generated /accepted.<br><br>POTENTIAL:<br>Can B2STAGE be made to transfer files between different architectures (not just B2 service) |  |                 | Analyzed and will not be integrated                          |
| RQ2            | B2SHARE          | No  | Integrate with DARIAH Gateway  | UC2             | Analyzed and will not be integrate                           |
| RQ3            | B2DROP           | No  | Integrate with DARIAH Gateway  |                 |  |
| RQ4            | B2ACCESS         | UNKNOWN   | Integrate DARIAH Gateway AAI (not sure this is really needed long term since hopefully it won't matter which EOSC AAI mechanism is used) |                 | Analyzed and will not be integrate. EGI AAI is used instead. |
| RQ5            | B2FIND           | NO: Existing DARIAH metadata already harvested  | Make data discoceverable through metadata searches   |                 | Analyzed and will not be integrated.                         |
| RQ6            | B2SAFE           | POTENTIAL:<br>May need to allow havestable metadata from B2SAFE instances   | Integrate with DARIAH Gateway  |                 | Analyzed and will not be integrated.                         |

|      |  |   |  |  |   |
|------|--|---|--|--|---|
| RQ7  | PaaS<br>Orchestrator   | UNKNOWN -<br>Plan is to<br>integrate with<br>WS-Pgrade<br><br>YES: Need to<br>define whether<br>this will be a<br>deployable<br>software compo<br>or use a central<br>instance. | DARIAH Gateway is not a deployable software. A central instance is used.<br><br>Invenio-as-a-Service is  |  | Analyzed and will not<br>be integrated for<br>DARIAH Gateway<br><br>Integrated with<br>Invenio-as-a-Service   |
| RQ8  | EGI<br>FedCloud/<br>FedCloud<br>Containers/<br>Nove-<br>docker<br>/OneDock | YES: Can users<br>run arbitrary<br>containers on<br>fedcloud<br>resources/<br>can DARIAH<br>user containers<br>be made<br>available for<br>general usage                        | Extend the gateway functionality to support creating<br>and starting docker-based applications and tools in the<br>cloud environment (EGI FedCloud) by using Indigo<br>(OneDock, OpenStack Nova Docker)  |  | <div>  <br/>EO <br/>SC <br/>WP <br/>10- <br/>28 <br/>- <br/>Jira <br/> </div><br>Completed |
| RQ9  | FutureGate<br>way  | UNKNOWN   | The science gateway will internally deploy a copy of<br>FutureGateway and the other components will integrate with<br>its APIs.  |  | <div>  <br/>EO <br/>SC <br/>WP <br/>10- <br/>87 <br/>- <br/>Jira <br/> </div><br>Completed |
| RQ10 | B2FIND<br>/B2SHARE   | UNKNOWN: Can<br>Semantic Search<br>Engine. harvest<br>from B2FIND<br><br>NO - B2SHARE<br>can already be<br>harvested  | Make data stored in EUDAT repositories findable through the Semantic<br>Search Engine.   |  | Analyzed and will not<br>be integrated.   |
| RQ11 | EGI<br>FedCloud/<br>OneData<br>/B2SHAR<br>E/<br>B2STAGE                    | NO  | The VLE will be integrated with the EGI compute and storage resource<br>infrastructure and integrated as a new user-oriented service of the EGI<br>DARIAH CC Gateway. The current client- server model will ported to the EGI<br>FedCloud infrastructure, while the underlying data management will be<br>integrated with the INDIGO OneData solution and EUDAT B2Share and<br>B2Stage service to transfer data to EGI FedCloud for batched processing of<br>editing procedures. |  | VLE is not part of the<br>DARIAH T.<br><br>Analyzed and will not<br>be integrated.  |

## Capacity Requirements

| <b>EOSC-hub services</b> | <b>Amount of requested resources</b> | <b>Time period</b>                  |
|--------------------------|--------------------------------------|-------------------------------------|
| OneData                  | 50TB                                 | EOSC-Hub project lifetime + 2 years |
| EGI FedCloud             | 300 cores/2TB memory (total)         | EOSC-Hub project lifetime + 2 years |
| B2DROP                   |                                      |                                     |
| B2SHARE                  |                                      |                                     |
| B2FIND                   |                                      |                                     |