

8.4 EISCAT_3D (CC)

| | |
|-------------------|-------------------|
| Short description | EISCAT_3D |
| Type of community | Competence Centre |
| Community contact | Ingemar Häggström |
| Meetings | |
| Supporters | |

- [Ambition](#)
- [User stories](#)
- [Use cases](#)
- [Requirements for EOSC-hub](#)
 - [Technical Requirements](#)
 - [Capacity Requirements](#)
- [Validation plan](#)

Ambition

EISCAT Scientific Association participates in the EOSC Hub WP8 Competence Centre with the aim of developing a data portal for users of the future radar system EISCAT_3D, which is planned to start operation during 2021-2022. The aim of the CC is to have a working prototype open for public access by M18 of EOSC-hub project (from Sept 2019).

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."

EISCAT_3D will generate raw data at up to three radar sites (each with almost 10000 antennas) and has to make those data browsable and analysable for researchers, as well as archiving the data for the long-term. It is planned to archive around 2 PB per year. This primary mission is achieved by procuring sufficiently big storage and network capacities at a few data/compute centres (for back-up or load distribution). The data will be transferred to this/these locations after initial filtering and calibration at the data source(s). The data/compute centres are responsible for data archival, data curation, generation of derived data (level 1-2-3) and for sharing the data with scientists.

Data must be shared with scientists via a data portal and programming APIs. The EISCAT_3D Data Portal should offer the following main features:

- AAI, user login
- Data browser
- Data download
- Online computing (analyse data without downloading them, using cloud resources, reference applications or user's own software)

The CC is exploring the use of the following services within the framework of EOSC-Hub:

- Application and data catalogue portal with compute integration: DIRAC
- Metadata catalogue for high-level data: B2Find
- Metadata catalogue for low-level data: DIRAC
- Compute resources: EISCAT, cPouta cloud at CSC
- Data storage and transfer (datacompute centres): File transfer TBD, e.g. B2STAGE
- User SSO: EGI CheckIn and B2Access

| No. | User stories |
|-----|--|
| US1 | Any researcher should be able to access the portal and browse metadata. The portal grants /denies access to data and processing based on affiliation. (Meta- data should be available for all researchers. The real data for authorised users only.) |
| US2 | Authorised researchers should be able to select the EISCAT_3D data they are interested in for download or for analysis. |
| US3 | Authorised researchers should be able to browse reference applications in the portal, select an application for use, feed their data in (from US2), visualise or download the analysis result. |

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 | Authentication: <ol style="list-style-type: none"> 1. User enters into the portal. 2. Portal directs to SSO 3. Portal grants authorisation | <ol style="list-style-type: none"> 1. Portal is running with metadata and file catalogue 2. EGICheckin <-> B2Access, EISCAT idp, other social IDPs |
| UC2 | Data access: <ol style="list-style-type: none"> 1. User selects data 2. Portal accepts authorisation 3. Portal directs to data store | <ol style="list-style-type: none"> 1. Dirac file catalogue 2. dirac sso 3. Dirac |
| UC3 | Computation: <ol style="list-style-type: none"> 1. Data selected 2. Software selected /uploaded 3. Data staging at HPC 4. Processing of data 5. Results presented | <ol style="list-style-type: none"> 1. Dirac 2. Dirac HPC 3. E.g. B2Stage 4. HPC, VMs and/or containers 5. Visualisation, B2drop, B2share |

Requirements for EOSC-hub

Technical Requirements

| Requirement ID | EOSC-hub service | GAP (Yes /No) + description | Requirement description | Source Use Case |
|----------------|-------------------|-----------------------------|---|-----------------|
| | | | | |
| RQ1 | EOSC-hub AAI | Yes: Dirac | Dirac interface to EGICheckin | UC1 |
| RQ2 | Data staging | Yes: File system | Temporary file storage close to computing | UC3 |
| RQ3 | Cloud Compute | Yes: HPC | Run VMs and/or containers for portal and data processing | UC3 |
| RQ4 | PID/DOI service | Yes: Registry | Registration of digital objects: data collections (experiment, low high levels of data) | UC2,3 |
| RQ5 | EOSC-hub Check-in | Yes:Check-in | <p>EISCAT community includes institutes in Japan and China. Currently, Checkin has been integrated into the EISCAT data portal based on DIRAC4EGI technology.</p> <p>However, google, facebook cannot be used in China thus users from there cannot yet login via Checkin.</p> <p>Other social logins are to be implemented, see GG US ticket:https://ggus.eu/index.php?mode=ticket_info&ticket_id=139172&come_from=submit</p> | UC1 |

Capacity Requirements

| EOSC-hub services | Amount of requested resources for CC development project | Time period | Envisaged future |
|-------------------|---|-------------|--|
| Portal | 2 VMs | 30 yrs | Provider TBD, tendering or own resources |
| Stage | 2 VMs 10 TB NB: the full EISCAT 3D system may require temporary storage on petabyte scale | 5 yrs | EISCAT site storage or national providers (tendering TBD?) |
| Processing | 2 VMs 10 cores NB, the full EISCAT 3D system will require much more capacity, the ultimate aim being realtime sub-beam imaging requiring some 500 TFLOPS/s | 5 yrs | EISCAT site compute or HPC provider (tendering TBD?) |

Validation plan

These are the required steps of validation to be performed by enrolling a number of external early adopters. The tasks follow UC1-UC3 above.

| | Task | Description | Expected outcome |
|--|------|-------------|------------------|
|--|------|-------------|------------------|

| | | | |
|---|---|--|---|
| 1 | User registration and authentication | <p>Users register for access with authentication through EGI Checkin, X.509 or social login.</p> <p>The users are added to VOMS groups or similar.</p> | Users are able to log in to portal. |
| 2 | Data search and access | Users browse the existing data archives. | Users can browse all data and access data according to embargo rules. |
| 3 | Software upload | Administrators add software to cloud computing; direct or as a container | New analysis software becomes available to portal users. |
| 4 | Job submission | Users select data and submit for processing with available software. | <p>Data are staged for analysis. Analysis VM starts.</p> <p>Users receive and download analysed data, e.g. plots or lag profiles.</p> |