

# Submitting Community Project Data to SITES Data Portal

---

SITES encourages researchers using [SITES infrastructure](#) to upload data generated in individual projects to the [SITES Data Portal](#). Complementing existing time series data with community project data enhances the value of the open data provided through SITES and therefore benefits both SITES and the research community as a whole.

## Community project data contributions on the SITES Data Portal

Community project data is defined as ecosystem based data that connects to one or more of the [SITES Research Stations](#) and/or [Thematic Programs](#) and is produced by an individual researcher or research group. The SITES Data Portal aims to provide [FAIR data](#). Therefore, project data uploaded on the SITES Data Portal needs to be well-documented and metadata must be provided by the data owner, e.g. instruments, sampling methods, parameters and units. Project data is published with an open access license unless a compelling reason warrants an embargo (see [SITES Data Policy](#)). Limits on the size of data uploads are typically not a concern, but note if the data file is > 64 MB in size, so we can find a workable solution.

## Why publish Community project data on the SITES Data Portal

- **Citation:** Data sets are given a persistent identifier (PID) which can be linked to research profiles and journal articles; if required, a DOI can be generated for individual data sets or data collections. A **preliminary DOI** service can be offered upon request. Please send an email to [data@fieldsites.se](mailto:data@fieldsites.se) stating your requirements and associated data.
- **Detailed Metadata:** Increase the quality and impact of your data through transparency and good documentation.
- **Interoperability:** Data uses standardized vocabularies, terminologies, and ontologies; further, it is embedded in a larger context, i.e. references to other (meta)data, so that it is possible to understand the relationships between data.
- **Secure and Persistent Archive:** Data is secured on three different servers, in Sweden, Finland and Germany. The PIDs allow the data to be available and consistent over time.
- **Workflow Automation:** The application program interface (API) is used for automating data publication, reporting and reproducible science.
- **Enhanced Discovery:** Data is discoverable in Google and other search engines through Json schemas.

### Steps for getting community project data on the SITES Data Portal

1. Fill in the [SITES Community Project Data Upload Request Form](#)
2. Prior to the request being approved, we may need to discuss your requirements and needs, including deadlines, to see if we can meet your expectations.
3. A data template is prepared by SITES and sent back to you for approval before data is upload to the SITES Data Portal.
4. Once the data is uploaded, a link and PID (if requested DOI) will be provided to you for your data set.

### Other Considerations

Disclaimer: SITES can't guarantee to prioritize an individual community project data upload request over ongoing central SITES data work; furthermore, SITES has the right to reject upload requests for reason of, e.g., large data storage, missing contextual information, lack of co-location with SITES, etc.

- Data must fit into the functionality of the SITES Data Portal
- Data provided needs to be structured in common formats and follow specific encoding of the SITES Data Portal
- New parameters and variables not currently on the SITES Data Portal may be introduced but should aim follow a controlled vocabulary
- Data provided should be quality controlled (i.e. L2 data), i.e. calibrated, quality filtered data in physical units, aggregated to appropriate spatial and temporal output units and resolution. In some cases, there may be a justified need to store and/or publish raw data, but this needs to be agreed upon by SITES
- Once data is published on the SITES Data Portal, SITES assumes responsibility of maintaining the PID to the data set
- Providing data to SITES is free but in some cases if data storage capacities exceed common needs, cost may be added and are to be paid by the data provider
- If data updates are needed (e.g. a mistake is discovered) it is the responsibility of the data provider to fix the data, with the upload of a new version supported by SITES
- Specific data related requests will be directed to the data owner