

Meeting Notes

15th of October 2024

- Release Jobs
 - Controller is almost done
 - The commented code left behind in the controller has been reviewed and refactored if needed or deleted if obsolete
 - JobCreatedDTO will be refactored
 - Despina had few questions regarding:
 - dataset file existence
 - enforce job type
 - Despina and Max will discuss off-line
 - We are committed to release MVP
 - Any further changes or redesign will be postponed to after MVP
 - Members interested in brainstorming next job release (job configuration, job types and deployment): Laura, Max, Despina, Patrick, Marco
- Search UI
 - Discussion regarding latest PR implementing the functionality to retrieve scientific metadata keys and types
 - Concerns regarding performance with a high number of datasets
 - Possible solution might be implementing indexes, although metadata keys might be many and too variable, therefore implementing indexes on all of them might not be feasible
 - Possible issues with dataset authorization, aka access to dataset
 - Possible solution might be a cron job creating a list of all the scientific metadata keys and their type. Storing such list in its own collection
 - Permissions and authorization needs to be discussed before implementation
 - Members interested in working on this feature: Dylan and Max
 - More questions regarding indexes:
 - can we create indexes through a cron job
 - number of indexes for best performance to searchability ratio
 - indexes be configured by admin
 - indexes selected by admin based on user query
- Documentation
 - most likely, we will use mkdoc
 - BE documentation:

- these documents will contain mostly description and how-tos:
 - how to install
 - specific configuration
 - anything else?
- New repository user-documentation is the new container for documentation that does not belong to specific repo
- Current documentation repo will be phased out as soon as possible
- Regina will take the lead in the documentation migration
- Carlo will provide issues related to documentation for SciCatLive
- Where should we store jobs documentation? Part of it should go in BE, FE and part of it in user documentation.
- Despina will place all the documentation in BE. Community will review afterwards and move the relevant parts to the correct repo
- SDK
 - Jay confirmed that the SDK are being published at each release and created as artifacts when a PR is merged to master
 - links to official repositories
 - <https://www.npmjs.com/package/@scicatproject/scicat-sdk-ts>
 - <https://pypi.org/project/scicat-sdk-py/>
 - <https://pypi.org/project/scicat-sdk-pydantic/>
- Branding
 - Max and Laura have reviewed the first draft of the logo and provided feedback to the designer
 - Catch phrase and mission statement are progressing
- Governance
 - a restructure of the project management is needed
 - probably facilities management should be involved as SciCat becomes part of the critical infrastructure
 - Governance examples:
 - <https://github.com/bluesky/governance>
 - <https://github.com/jupyter/governance>
 - we should aim for a simple structure
 - Ideas:
 - technical committee
It will replace the current meeting and will deal with all the technical decisions. Only members.
Frequency: every 2 weeks
 - steering/management committee
It will work on the vision, directions and sustainability of the project.

New non technical group. It will help secure funds to sustain the development. Only members.

Frequency: every three months

- user groups

Open meeting where all the site admin and users can participate and ask questions, share ideas, and provide feedback

Frequency: monthly

SDK roadmap proposal

Following the discussion during the last collaboration meeting, I propose the following roadmap forward regarding the SDK:

- *Finalize auto generation of javascript and python SDKs CI/CD. **COMPLETED***
- *Create SDK and openapi file as artifacts at every merge to master.*

COMPLETED

Artifacts should be named according to the following schema:

sdk-<language>[-<additional-options>]-<git-commit-id>.zip

- Publish them automatically on the language specific platforms with the following names:
 - javascript: scicat-sdk-js
 - python: scicat-sdk-py
 - python-pydantic: scicat-sdk-pydantic

COMPLETED

- Publishing platforms are:
 - javascript: npm-js
 - python: pypi

COMPLETED

- Clearly state that SDKs are provided as they are, with no guarantee.
No testing is done on SDKs during CI/CD.
- Include new languages as needed
- refactor pySciCat to extend scicat-sdk-py, with backward compatibility, additional functionalities and validation and include testing like the ones included in discarded PR.
- Proposed pySciCat as the low level SciCat python library, providing the following:
 - additional functionalities,
 - extensive testing
 - backward compatibility
 - different release cycle from the SDK due to human revisions
- After 6/12 months evaluate if having scicat-sdk-py and pySciCat is sustainable and the most sensible solution.
We will adapt the road map forward based on the outcome.
- Refactor FE to use new SDK: ***in progress***

This roadmap is recorded in the following github issues:

<https://github.com/SciCatProject/scicat-backend-next/issues/1403>

Proposed Agenda

17th of September 2024

- Current releases updates
 - Jobs
 - Search UI
 - Metadata key and type
<https://github.com/SciCatProject/scicat-backend-next/pull/1428>
 - Documentation
 - where should we save the documentation specific for BE and FE?
 - which folder?
 - SDK
 - Branding
 - Tagline: <https://github.com/SciCatProject/scicatproject.github.io/issues/55>
 - Mission statement
<https://github.com/SciCatProject/scicatproject.github.io/issues/56>
 - other efforts
- Discussions
 - Project leadership and governance
 - <https://opensource.guide/leadership-and-governance/>
 - <http://oss-watch.ac.uk/resources/benevolentdictatorgovernancemodel>
 - <https://www.linuxfoundation.org/resources/open-source-guides/building-leadership-in-an-open-source-community>
 - Data visualization
 - which use case to address first?
 - how to integrate external tools?
 - Dataset Schema Validation
 - Semantic vs structural validation
 - Input vs output schema
- PR and issues
- Additional topics

Open discussions for future features

currently not scheduled on any release

- Proposed types for Scientific Metadata, Issue #984
 - <https://github.com/SciCatProject/scicat-backend-next/issues/984>
- Metadata entry schema
 - The overall metadata schema is free for the user to design their own structure
 - SciCat will impose a schema at the entry level to be able to interpret and visualize the values reducing uncertainty and wrong interpretation
 - <https://github.com/SciCatProject/scicat-backend-next/issues/939#issuecomment-1874033672>
 -
- Dataset hierarchy
 - Proposed solution:
 - unify the dataset schema
 - add dataset type: collection (specialized type of derived dataset)
 - add relationships and leverage them to build a hierarchy
 - <https://github.com/SciCatProject/scicat-backend-next/issues/805>
- Independent endpoint for datasets scientific metadata management
 - <https://github.com/SciCatProject/scicat-backend-next/issues/954>
 - As we use SciCat as our platform to curate our dataset, it is becoming more important to be able to add, update or delete individual scientific metadata entries
 - High priority
- Dataset history and update history
 - should we create a stand alone collection for update history?
 - should we track any changes made to datasets?
 - should we track changes made to other objects?

SciCat Collaborators Zoom meeting details

Topic: SciCat Collaborators Meeting

Time: Feb 6, 2024 04:00 PM Copenhagen

Every 2 weeks on Tue, until Dec 24, 2024, 24 occurrence(s)

Join Zoom Meeting

<https://ess-eu.zoom.us/j/61405959531?pwd=MllyYm12RS9RNEdGTzR3bXRvNGp5Zz09>

Meeting ID: 614 0595 9531

Passcode: 507998

One tap mobile

+46850163827,,61405959531#,,,,*507998# Sweden

Dial by your location

- +46 8 5016 3827 Sweden

Find your local number: <https://ess-eu.zoom.us/j/61405959531>

Join by SIP

61405959531@109.105.112.236

- 61405959531@109.105.112.235

Join by H.323

- 109.105.112.236

- 109.105.112.235