

Meeting Notes

29th of October 2024

- Release Frontend SDK
 - Supported by DAPHNE contribution
 - Martin is working on this release
 - Martin provided an update
 - New branch has been created: new-sdk-release
 - First PR to add authorization subsystem which is not included in the SDK
 - when developing locally, it will be possible to use an official SDK release, or, alternatively, a script will be provided to create a local copy of the SDK provided that local instance of the BE is running.
 - In production, we will support only an official SDK release provided through npm package
 - documentation will be updated.
 - the SDK is already distributed as npm package, available at <https://www.npmjs.com/package/@scicatproject/scicat-sdk-ts>
 - Despina asked and reminded us about multiple FE. The autogenerated SDK will help
 - Current version of the API is V3. As soon as a breaking change is released, we should move to V4.
 - Carlo noted that the base URL is defined in the backend. Most likely hardcoded.
 - Jay created and merged a PR to define the api version from configuration and through an environmental variable.
- Release Jobs
 - Validate actions
 - The old BE V3 code checking and validating jobs has been migrated to BE V4
 - MVP not reached yet
- Search UI
 - Topic that still needs to be discussed: list of scientific metadata and the type
 - We should find some use cases to guide the proof of concept:
 - beam energy
 - Ideas:

SciCat Collaborators Meeting Notes

- use indexed created at startup
- select the fields that are most used in search
- implement separate table with indexed fields or leverage `db.get_index` functionality
- indexes should be easily configurable
- implement admin endpoints in BE and admin section in FE
- Dataset search form discussion
 - Filters on high level fields and conditions on scientific metadata should behave and look the same
 - list of missing filters should be created
- NMX and IsPyB
 - Carla (Max IV) mentioned that MaxIV is interested in exploring using SciCat as BE for NMX and IsPyB, as ESRF stopped supporting the dedicated backend
 - Questions:
 - Can all the data entities used in IsPyB saved in SciCat?
Examples: shipments, samples
 - Scientific metadata
 - customized visualization
 - Max will stop by at MaxIV to discuss in person
 - IMPORTANT: Organizers of the IsPyB meeting in Trieste reached out regarding the same issue. Max will present remotely.
 - DESY and SOLEIL are interested in such effort
- Governance
 - Ideas:
 - Technical advisory group
It will replace the current meeting and will deal with all the technical decisions. Only members.
Frequency: every 2 weeks
 - Management advisory board
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
 - Members should ask to their management if they are interested in participating in the board and provide guidance and insights.
 - Governance examples:
 - <https://github.com/bluesky/governance>



SciCat Collaborators Meeting Notes

- <https://github.com/jupyter/governance>
- we should aim for a simple structure
- SciCatCon 2025
 - It will be held in Copenhagen hosted at ESS
 - Max is trying to securing some fundings
 - Meeting will be held the beginning of July
 - A poll will be created
 - It will include a visit to MaxIV and ESS
- Authorization
 - Soleil has needs for more granular authorization
 - Examples:
 - Only PI should be able to publish dataset or create published data entities
 - Dataset editing should not be allowed for owner groups

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

- Renamed python packages to address naming issues: **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. ***No requests have been received yet***
- 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

12th of November 2024

- Current releases updates
- Discussions
 - Dedicated project URL
 - 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>
 - Datasets details
 - configurable view
 - interface definition
 - third parties custom view components
- 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?