

IMPORTANT

Next meeting :

9th of January 2024, h16:00CET / h07:00PT

The SciCat community wishes Happy Holidays
to everybody!!!

Meeting Notes

12th of December 2023

- Releases
 - Release jobs:
 - Daphne provided an update
 - Wednesday 6/12/2023 an online meeting was held to review the release status and facilitate the work to be done.
Participants: Daphne, Marco, Regina, Despina, Spencer, Igor, Max.
 - Despina is working on the cretae job endpoint/function.
She has some questions regarding the implementation. A meeting before December 16th will be scheduled to provides support.
 - Release Elastic Search
 - Max reported that this release has been merged in master and official release 4.4 has been created.
 - ESS has deployed release 4.4 on their staging environment for furhter testing.
 - Elastic Search is an opt-in feature. Please check the documentation for more info:
<https://scicatproject.github.io/documentation/Development/v4.x/elasticsearch/elasticsearch.html>
 - Release Search UI
 - Igor is doing th epreliminary research.
 - Work will start at the beginning of January
 - Volunteers are needed to help on this release
- Website
 - Presented latest updated (Huge thanks to Despina)
 - Please create issues for any changed needed.

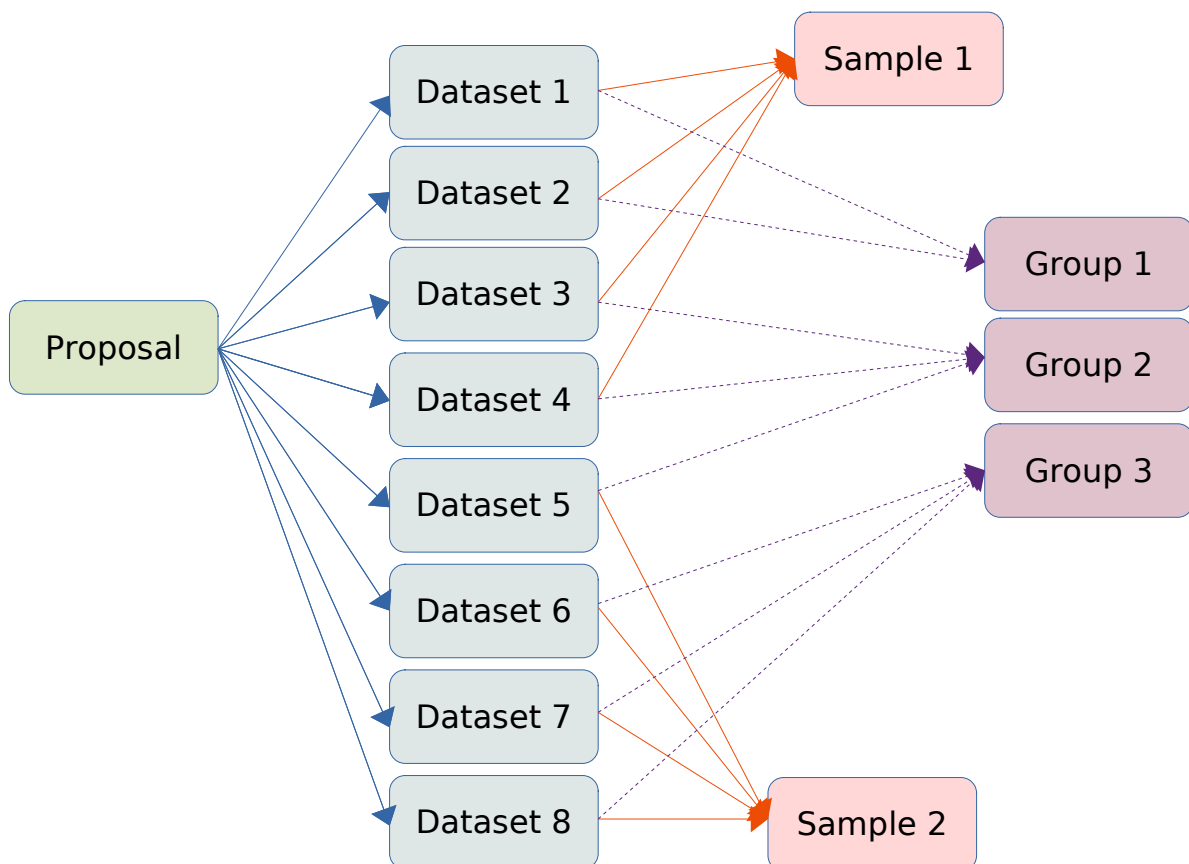
- Despina will work on the next release beginning of January.
 - Please mark issues urgent if they need to be addressed ASAP
- Documentation
 - links in pages do not work.
Example is the OIDC integration for V4.x.
Issue needs to be investigated
 - Regina asked for a slack channel just for documentation. Carlo created right away
- Hierarchical datasets
 - Max introduced the subject and explained the information received from Morgan. Three possible solution could be adopted. Please see agenda below.
 - Spencer: relationships could be leveraged when creating Published Data
 - Dylan: we could use json-ld standards to provide this information. Json-ld is pretty popular in the web-world.
This links to the concept of triplestore (<https://en.wikipedia.org/wiki/Triplestore>) and it is used at NIST
 - Morgan: they would be happy to have a simple field on each dataset where they can save a group/collection id.
 - Morgan: Potentially one dataset could belong to multiple collections.
 - Collections would need their own metadata
 - Martin: DESY would benefit from the such grouping that is not related to proposal, nor sample
 - Majid: SOLEIL would needs the functionality to group datasets from different proposals under the same collection
 - Spencer: he supports the concept of collections of datasets, he worries about hierarchical metadata.
 - Github issue: <https://github.com/SciCatProject/scicat-backend-next/issues/805>
- Metadata structure, BE #939
 - the community supports the idea
 - both BE and FE should be backward compatibles with current data
 - the new metadata structure should be well defined
 - Dylan retracted the idea of using json-ld for this case.
 - A migration script for existing data should be provided although facilities should not be required to move to the new syntax
- Independent endpoint for datasets scientific metadata management, BE #954
 - There is support for the idea, more brainstorming is needed
 - Spencer suggested json patch (<https://jsonpatch.com/>) which provides a standard way to update a single field of a json object
<https://www.rfc-editor.org/rfc/rfc7396>
 - Alternative: merge-patch
 - json-patch and merge-patch are different (<https://erosb.github.io/post/json-patch-vs-merge-patch/>)
- BE PR #925: Array support in Scientific Metadata.

- Max ask if we should accept the PR as it is or we should require to be completed with API tests, matching search functionalities and relevant changes in FE.
 - Concern has been raised regarding the dimensions allowed. Nobody wants time series saved in metadata
 - Regina will reach out to the user Jkotan to ask for more details on his use case
 - More discussion is needed
- SciCat in-person user meeting 2024
 - it will be held at PSI.
 - date will be define in the new year
 - Laura will share wisdom with Carlo and his team
- ispyb
 - Majid reported that ispyb is tightening is link to icat, merging the two databases in to one. It is his intention to provide SciCat integration, also.
 - Daphne shared that MaxIV is aware of the issue and there are internal discussions
- FAIR data meetings
 - Majid reported that there is a new project similar to PaNOSC regarding accelerators. He is proposing SciCat as a solution to manage the data.

Proposed Agenda for the 12th of December 2023

This meeting will be the last one for 2023.
The following meeting will be on January 9th, 2024.

- Releases
 - Release Jobs (Daphne)
 - Release Elastic Search (Max)
 - Release 4.4 (Released, deployed in staging at ESS)
 - Release Search UI (Igor)
- Website:
 - Updates deployed (Thanks Despina)
 - New updates
- Documentation
 - Following authorization pages needs to be generated
 - Datablocks
 - Proposals
 - Samples
 - PublishedData
- Features
 - Dataset hierarchy
 - examples and discussion



- Possible solutions:
 - Targeted keywords
 - Derived dataset pointing back to relevant datasets with adequate title and keywords
 - Custom relationship (new feature)
 - Examples:
 - dataset 6 is member of group3 like dataset 7
 - dataset 4 has been acquired together with dataset 5 under the same conditions
 - ontology for relation type
- Metadata structure
 - Issue BE #939
 - Use json-ld to provide context for each scientific metadata, so it is easier to understand types, naming and other properties.

```

      'key_name' : {
        '@value' : 'metadata_entry_value',
        '@unit' : 'metadata_entry_unit',
        '@type' : 'type of this field',
        '@human_readable_name`': `Key Name`
        'sub_key_name' : {
          ....
        }
      }
    
```

 - json-ld documentation: <https://json-ld.org/>
 - This approach will allow to assign the correct widget and show the correct human readable name in the frontend.
- Independent endpoint for datasets scientific metadata management, BE #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
- PR and issues
 - BE PR #925: Array support in Scientific Metadata.
We need to discuss about tests, search and UI element
 - BE PR #726: Access groups
Best way of testing
 - BE issue #941: OpenAPI schema
Apparently there are some incoherencies between the schema/dto provided in swagger and the data expected/accepted by some endpoint.
We need to review the testing and add the relevant tests to catch this issues.
 - BE PR #945: bump mongoose from 7.6.5 to 8.0.3
Needs a volunteer
 - Other PRs
- SciCat in person meeting
 - updates
- Additional topics