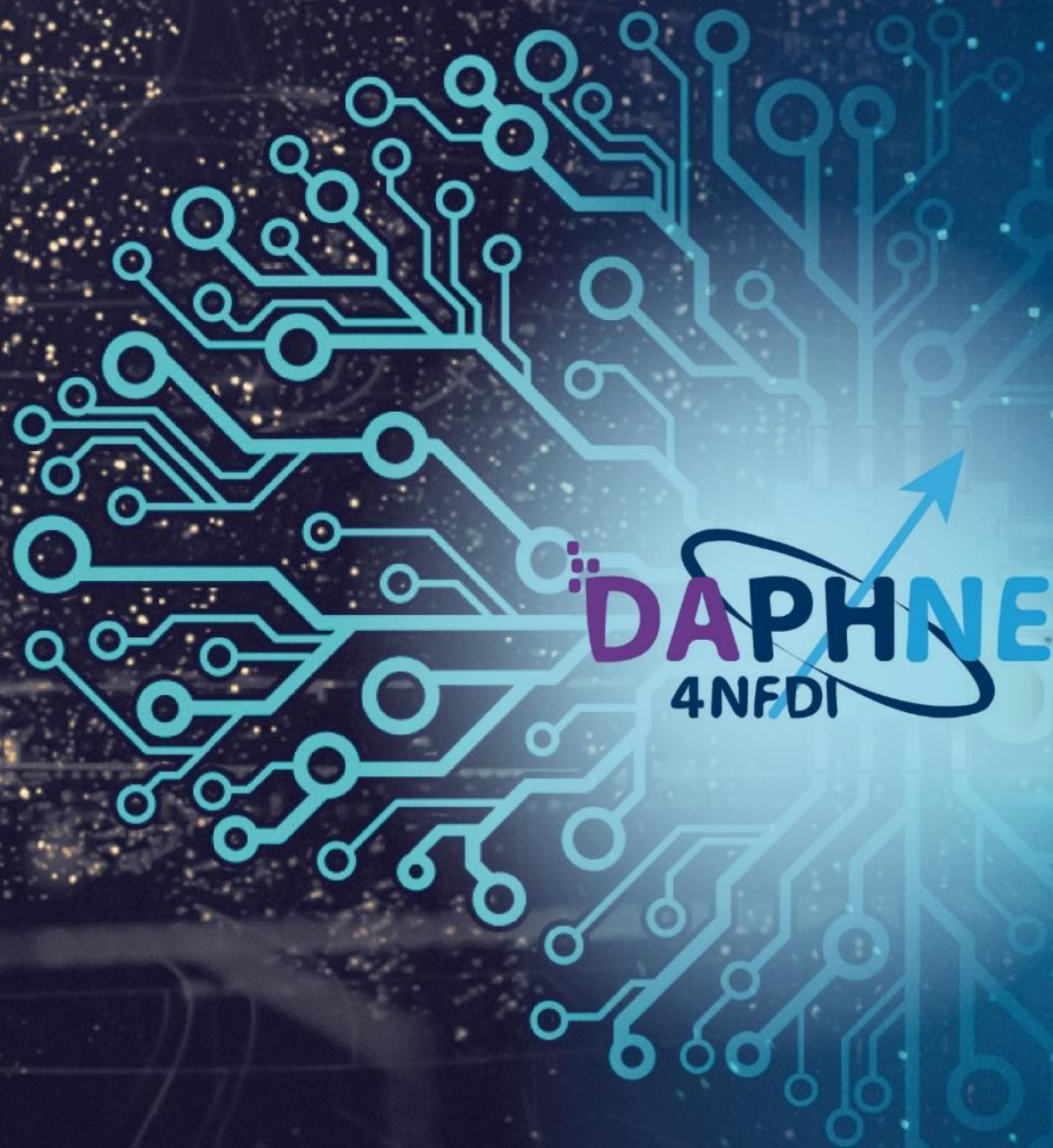


News

What unexpected project funds can do

Regina Hinzmann
2024-09-16



Plans of some projects funding for SciCat development

Chance

- We, w/ Bridget M, contacted Max Novelli to offer some funding for features that are relevant for
 - SciCat project itself and the DAPHNE project too (50:50)
 - Entire DAPHNE community in concern (TA1 and TA2) were called to write by yesterday features and name.
 - This is the result:
- Consolidation will be passed to Max

Björn Pedersen (MLZ - TUM):

- add CONTAINS_STRING option to condition search FE PR #1342 + BE PR #119
→ generally improve filtering for scientific metadata (ranges, maybe element search...)
- No schemaless field in proposal class FE issue #1455, BE issue #1178
- Implement FE OIDC authentication (see discussion in BE issue #1352)

Linus Pithan (DESY)

- add CONTAINS_STRING option to condition search FE PR #1342 + BE PR #119
 - <https://github.com/SciCatProject/backend-v3/pull/680>
 - <https://github.com/SciCatProject/frontend/pull/974>
- No schemaless field in proposal class FE issue #1455, BE issue #1178
- Validation of scientificMetadata
 - <https://github.com/SciCatProject/scicat-backend-next/issues/966>
- atomic patching Jan Kotanski, <https://github.com/SciCatProject/scicat-backend-next/issues/954>
- Dataset hierarchy : <https://github.com/SciCatProject/scicat-backend-next/issues/805>

Regina Hinzmann (DESY)

- hierarchical proposals and configurable frontend labels as suggested by Max in [this document](#) to disentangle DESY ID issue described [here](#). More info on [indico](#).

Plans of some projects funding for SciCat development

Chances

Igor Khokhriakov (DESY)

- Research replacing MongoDB with RavenDB. RavenDB provides a number of superior features comparing to plain MongoDB. Specifically, the one that is interesting for us is embedded full text indexing. This may remove requirement in using Elastic. That will be beneficial as its constantly changing their licensing model in past few years, and also will reduce code base complexity and hence maintainability
- Refactor BE to split FE server side and an actual backend ([#1570](#))
- Refactor BE to leverage document-oriented capabilities of the DB and avoid SQL-like relationships ([#1353](#))
- Create an official Helm repo to allow easy deployment onto K8s
- Data Visualization ([#1518](#))
- Integrate with a Blockchain for a superior data ownership security
- Allow custom and multiple BE ([#1352](#))
- Redesign FE to offer better UI/UX

Paul Millar (DESY)

- Improve support for databases:
 - Add clear separation between business logic and database access layer; e.g., implement a Data Access Object (DAO) pattern (similar to Igor's point).
 - Use this database abstraction to add support for PostgreSQL database. One approach could involve document types being mapped to PostgreSQL tables, the fixed fields (within the SciCat data model) are mapped to columns within that table. Note that PostgreSQL supports columns that store arbitrary JSON data, so fields like *scientificMetadata* (within Dataset) could still be supported.
- Improve DataCite metadata specified when registering a DOI for dataset.
- Remove *PublishedDataset* in favour of dataset hierarchy (once implemented).

Anna Zimina (KIT)

Note: not yet fully in used, just installed

- automatical upload over table or any other structured text file

Hafiz and Paola (LMU)

Note: Not fully in running stage: installation completed

Martin Voigt, Thomas Gruber, Oliver Knodel (HZDR)

- Explicit group assignment:
 - Allow assigning groups in a static config which is most useful to manage permission of functional account.
 - We already implemented this for HZDR, happy to make a PR
 - In our implementation there is a separate config file for static group assignment so this is related to <https://github.com/SciCatProject/scicat-backend-next/issues/1120>
- Make *sourceFolderHost* easily available to file download implementations
 - Current workaround is to encode host in file path
 - There are recent changes (implemented "actions") in that area, but it still looks like the *sourceFolderHost* is ignored:
<https://github.com/SciCatProject/frontend/blob/7d2a87208d1b780683026e0a769be03e9e4b156c/src/app/datasets/datafiles-actions/datafiles-action.component.ts#L136>
 - This would probably be extremely simple to implement
- Dataset hierarchy to group datasets obtained at the same instrument/group/experiment
 - As proposed in <https://github.com/SciCatProject/scicat-backend-next/issues/805>
 - with these groups the selection of properties in the search bar can be reduced
- Quantity range
 - As one of the things proposed in <https://github.com/SciCatProject/scicat-backend-next/issues/984>
 - This seems useful as there are already cases of parameters with ``_min`` and ``_max`` suffix in our instance
- Links in metadata

Consolidation

Chance

In order to properly consolidate, let me summarize

- Many pointed out very precise and concrete and for that amount of sum probably easy to implement features
- Those who suggest major changes like supporting the database in the sense of enabling SciCat to use another DB we carefully have to consider:
 1. Do we really want to spend the time and money to do research, redesigning etc to incorporate a new DB?
 2. Which feature are you missing one cannot do with mongoDB?

Thanks