Databases status update

Fedor Prokoshin

JINR, DLNP

May 14, 2025

SPD Information Systems

Data

- Distributed Computing and Data Management
- EventIndex
- Production and versions resgister

Collaboration

- Personnel and Publication Databases

Detector

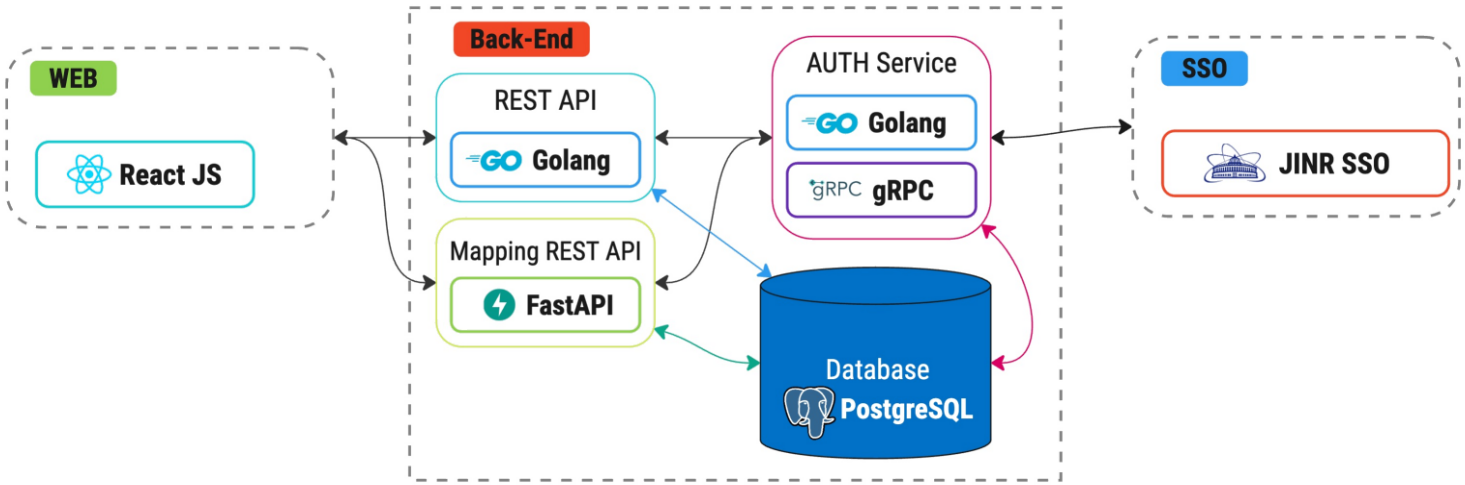
- Hardware and Mapping Database
- Geometry register
- Magnetic field maps register
- Monitoring Information Systems
- Logging and Bookkeeping

Hardware Database

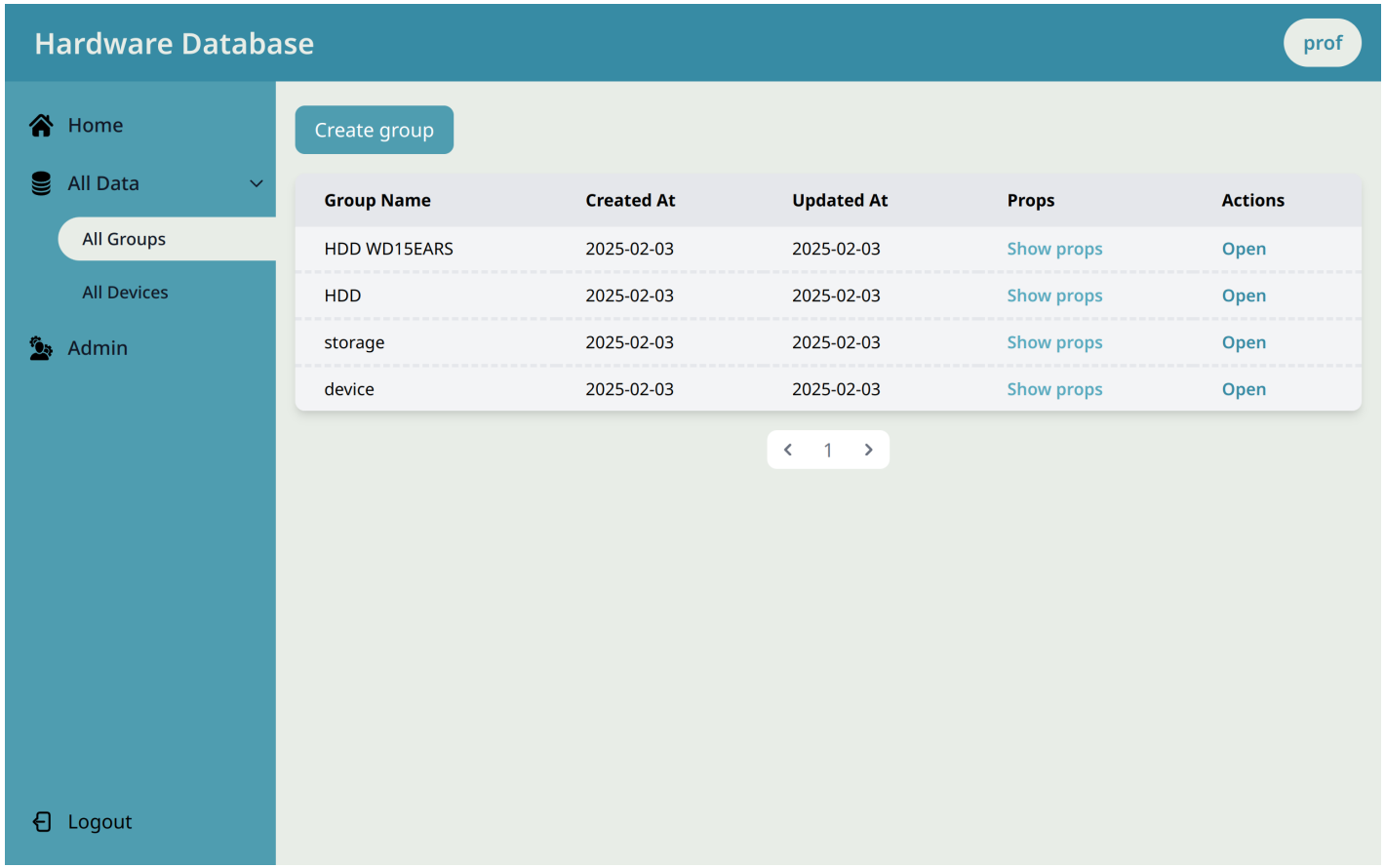
- A catalog of hardware components that SPD detector consist of.
- It should contain the information about the detectors and the electronic parts, as well as the location history of all items
- It include equipment models, provider, parameters and other (semi)permanent characteristics
- This should help in maintenance of the detector systems and especially helpful in knowledge transfer between team members.
- Filling of the hardware database should take place gradually, and updates will be rare
- The requirements for the speed of recording information in the database are low, as well as for reading of data
- The system should be tailored for the needs of the hardware devlopment and maintenance

Hardware Database: architecture

- A prototype system is being developed
  - PostgreSQL for a data storage
  - Back-end providing enpoints for access through the REST API
  - Web front-end providing interface for fill, searc and display data
  - JINR SSO for autorization, acces can be requested from web front
- The architecture of the project is built on the basis of microservices
- A stack of modern platforms is used to ensure good performance, flexibility and ease of development and maintenace

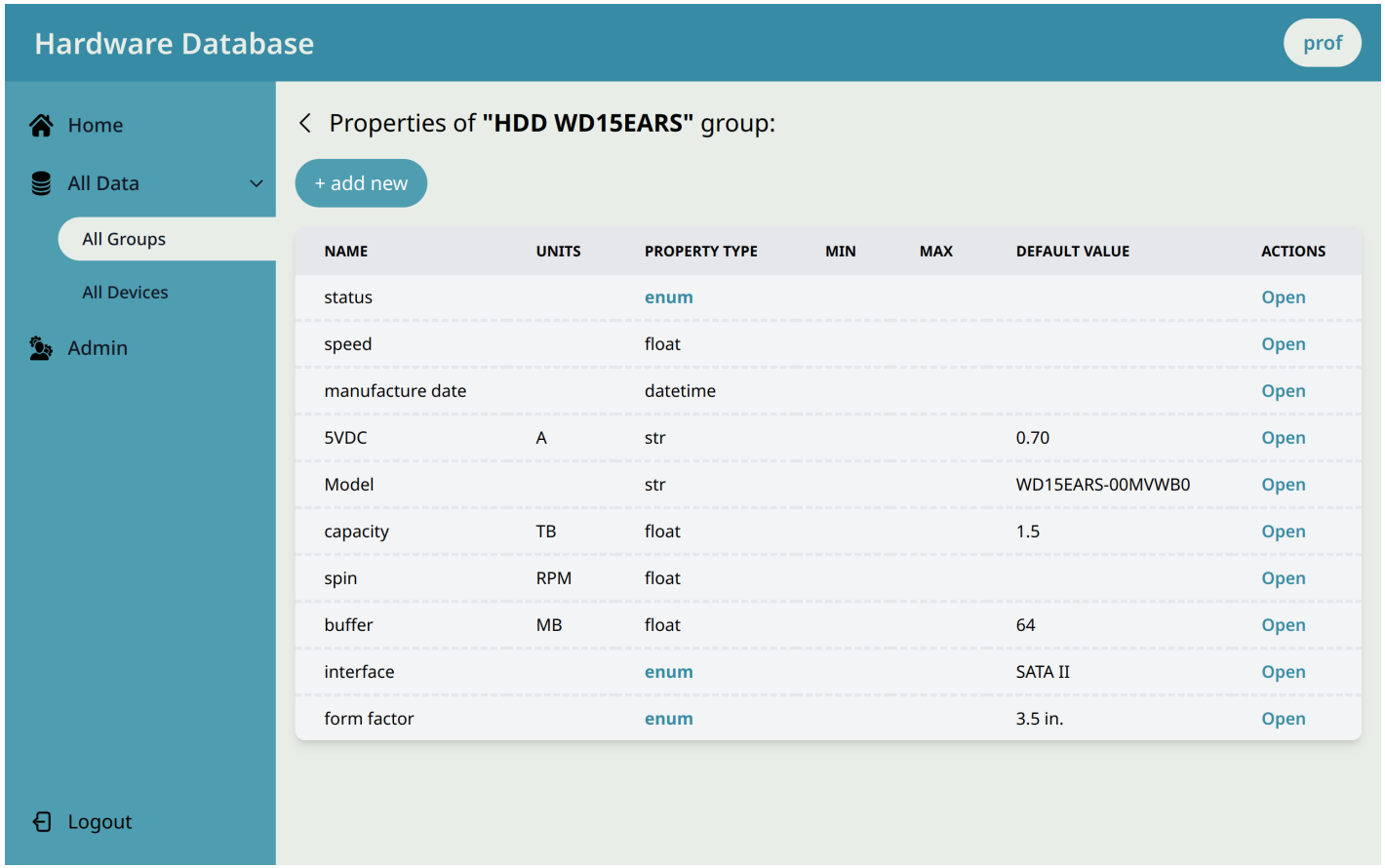


Hardware Database: interface



- Development of the HW database is put on hold, no urgent demand
- Developers shifted to other projects with higher priorities

Hardware Database: interface



- Development of the HW database is put on hold, no urgent demand
- Developers shifted to other projects with higher priorities

Database of personnel and documents

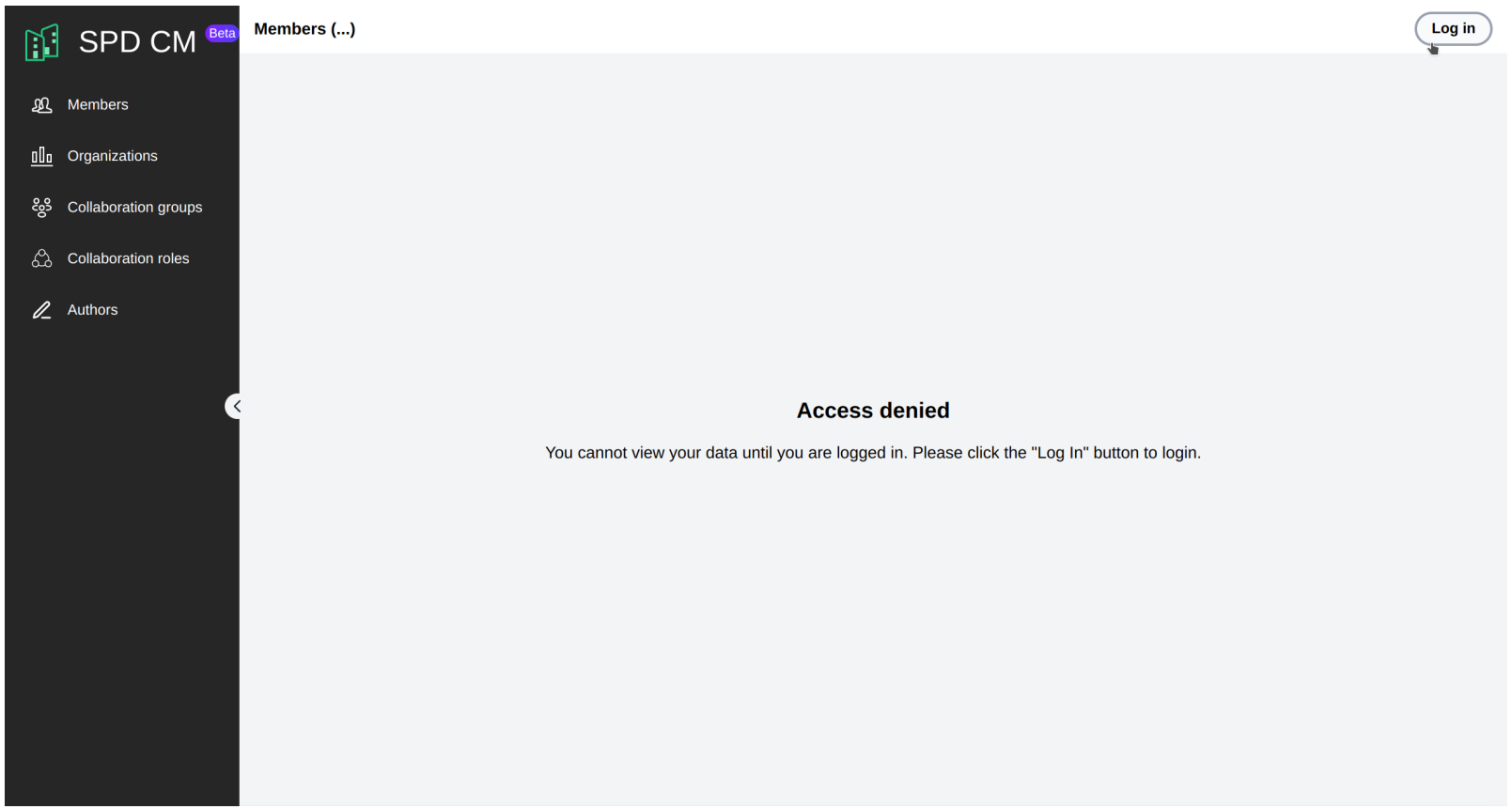
- About 400 people are currently participating in the SPD project,
  - number of participants is expected to grow close to experiment start
- In order to organize effective cooperation with the shared use of computing and other resources, it is necessary to have IS for
  - handling of a personnel and organizations data
  - support for working groups: membership, access rights,
  - accounting of the contribution (if implemented)
  - generating reports broken down by various parameters
- Procedures for creating, approving and editing related documents
  - Registration and changes of membership in the collaboration
  - Creating and editing lists of groups and privileges
  - Inclusion in the author's lists

Collaboration Members Database

- An information system has been created to handle information about collaboration members and participating organizations
- Means to store, add and edit records for members and orgainizations
- To add and edit roles in a collaboration and groups in a collaboration
- Provides an opportunity to generate the following reports:
  - A general list of the participants of the collaboration, total number
  - A list of participating organizations, with the number of participants;
  - A list of participants in the organization, indicating team leaders.
  - The list of participants by group
  - List of authors
- Provides authorized access to view, add, and edit information.
- Three roles that define the rights to work with the system:
  1. Administrator: can assign roles to users and edit information.
  2. Editor: can enter, edit, view and delete data, and view transaction logs.
  3. Guest: can view summary reports and information in the records.
- A user without a role does not have access to the data in the system.

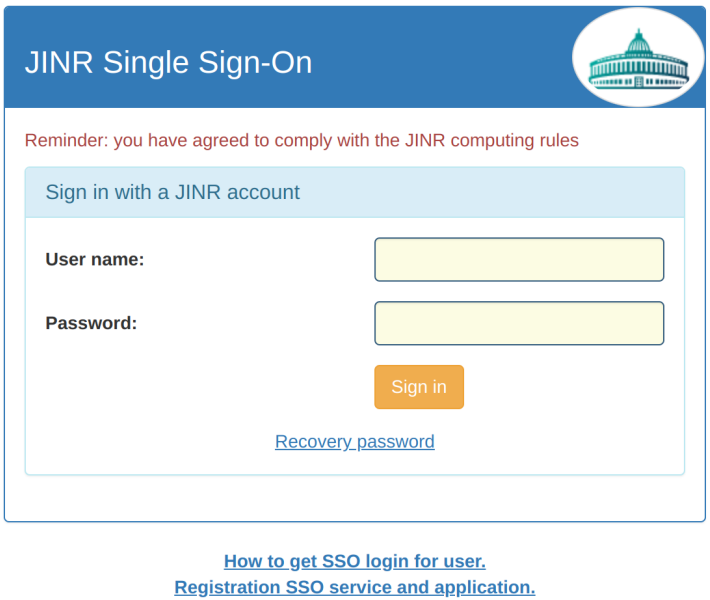
Collaboration Members Database: Authorisation

- User authentication in the system is performed using the JINR SSO service.
- During the initial user authentication, an account is created.
- The system administrator either assigns a role or denies access.



Collaboration Members Database: Authorisation

- User authentication in the system is performed using the JINR SSO service.
- During the initial user authentication, an account is created.
- The system administrator either assigns a role or denies access.

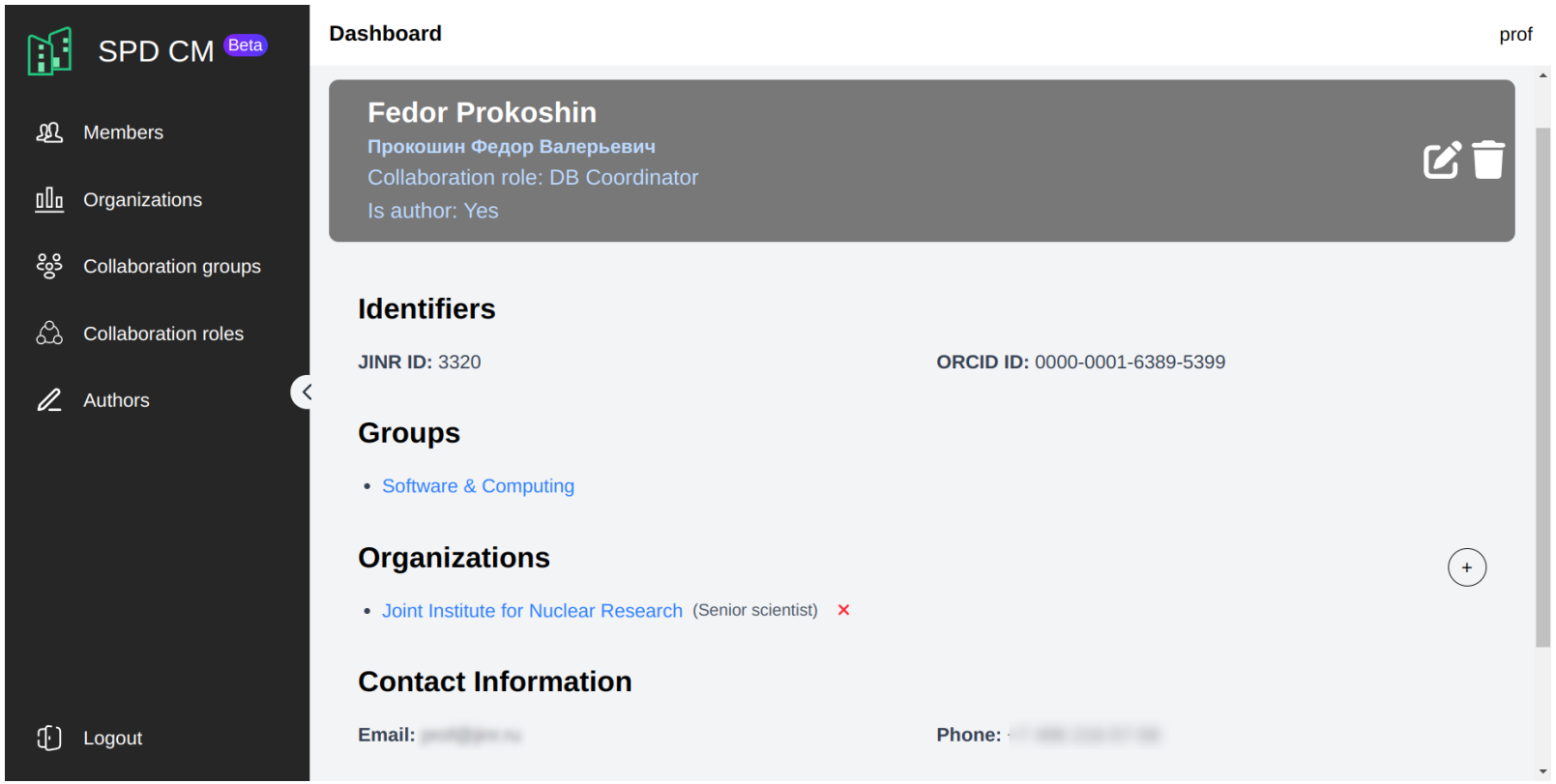


Collaboration Members Database: authors

- List of authors and their home institutions
- Clicking on members name bring you form with member information, including JINR and ORCID ID, group mebrship and home institutions

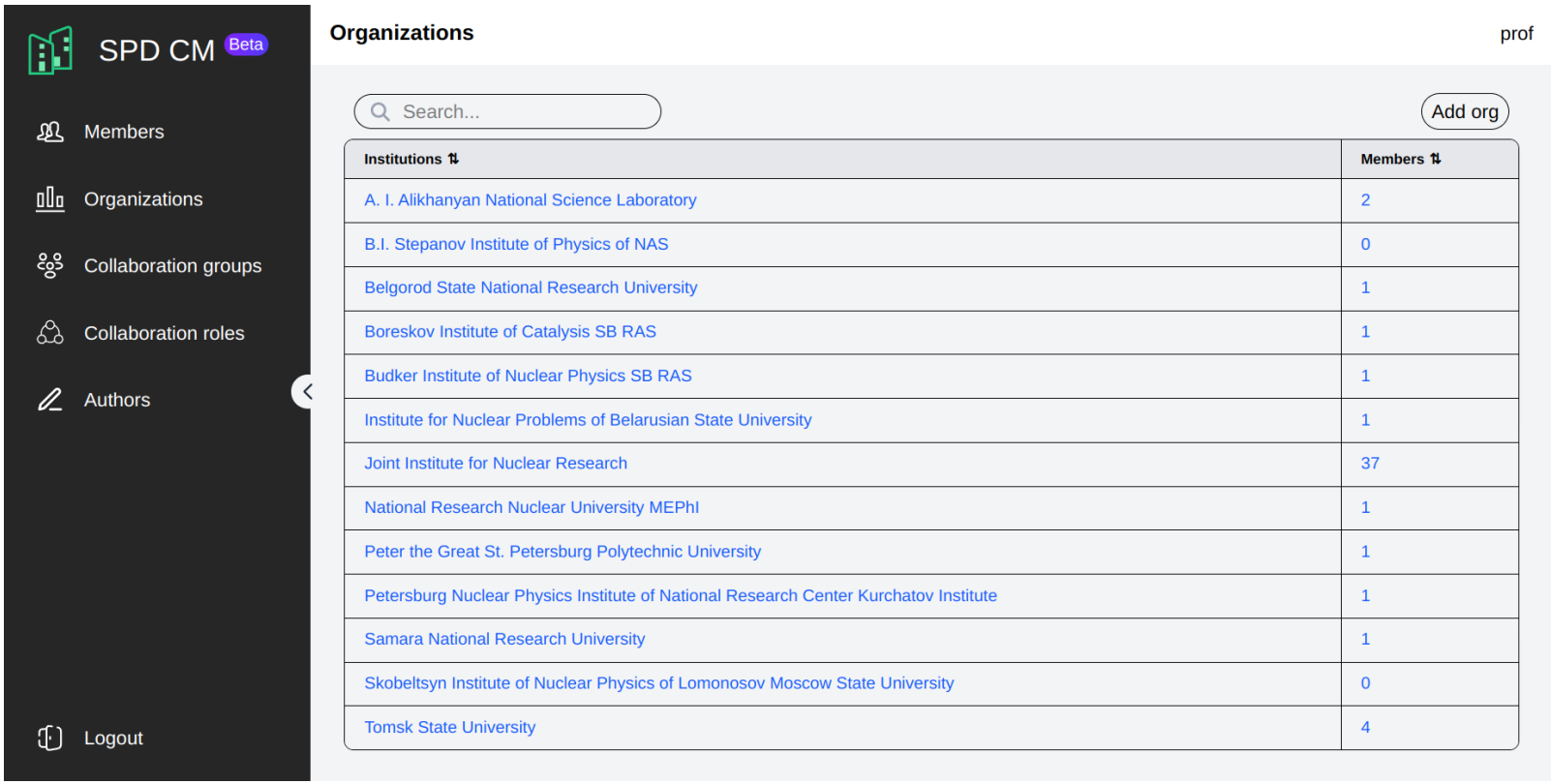
Collaboration Members Database: authors

- List of authors and their home institutions
- Clicking on members name bring you form with member information, including JINR and ORCID ID, group mebrship and home institutions



Collaboration Members Database: organizations

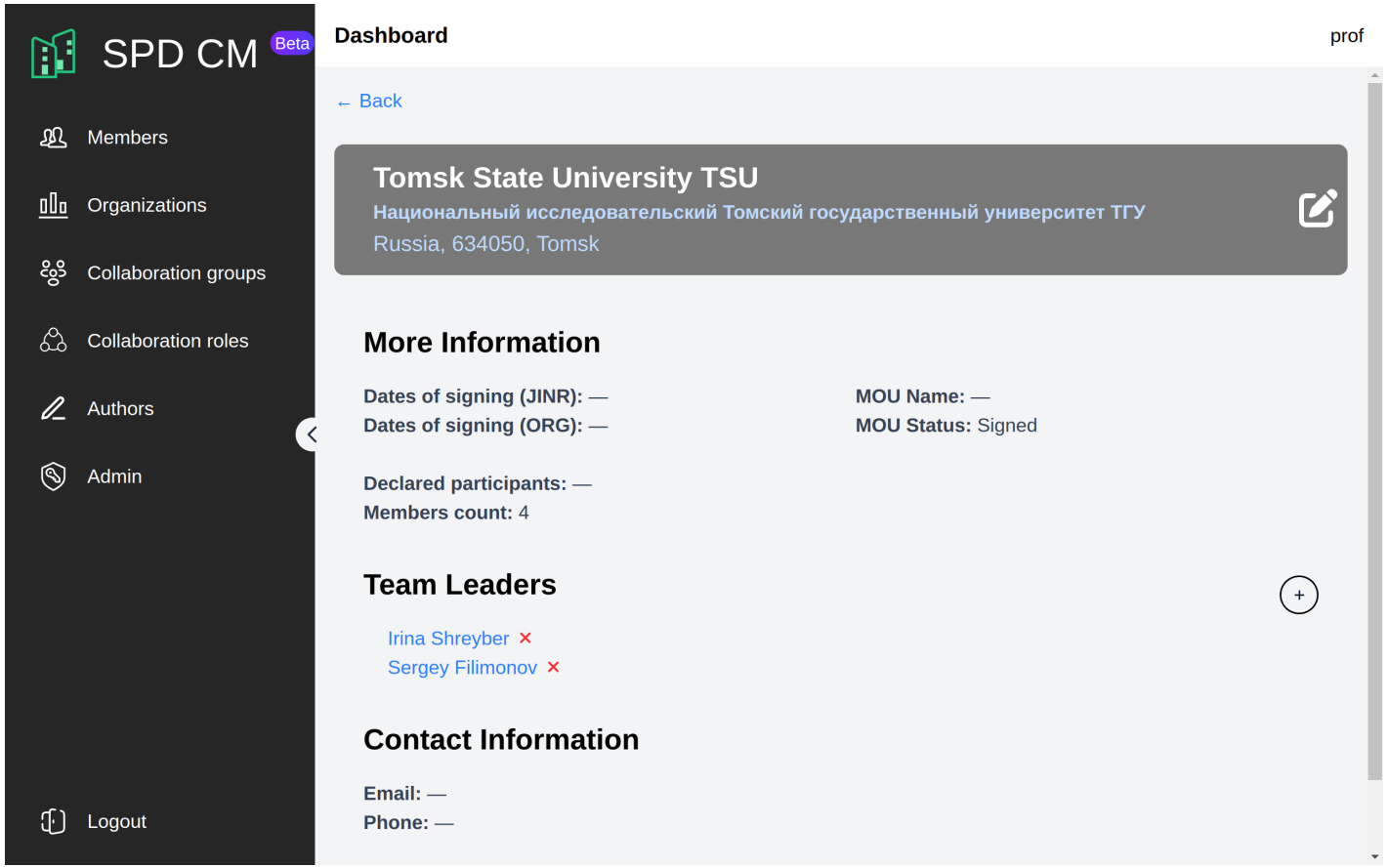
- List of participating organizations with number of members
- Clicking on organization name brings you organization details
- Clicking on number brings you list of members from this organization



Institutions ⇅	Members ⇅
A. I. Alikhanyan National Science Laboratory	2
B.I. Stepanov Institute of Physics of NAS	0
Belgorod State National Research University	1
Boreskov Institute of Catalysis SB RAS	1
Budker Institute of Nuclear Physics SB RAS	1
Institute for Nuclear Problems of Belarusian State University	1
Joint Institute for Nuclear Research	37
National Research Nuclear University MEPhI	1
Peter the Great St. Petersburg Polytechnic University	1
Petersburg Nuclear Physics Institute of National Research Center Kurchatov Institute	1
Samara National Research University	1
Skobeltsyn Institute of Nuclear Physics of Lomonosov Moscow State University	0
Tomsk State University	4

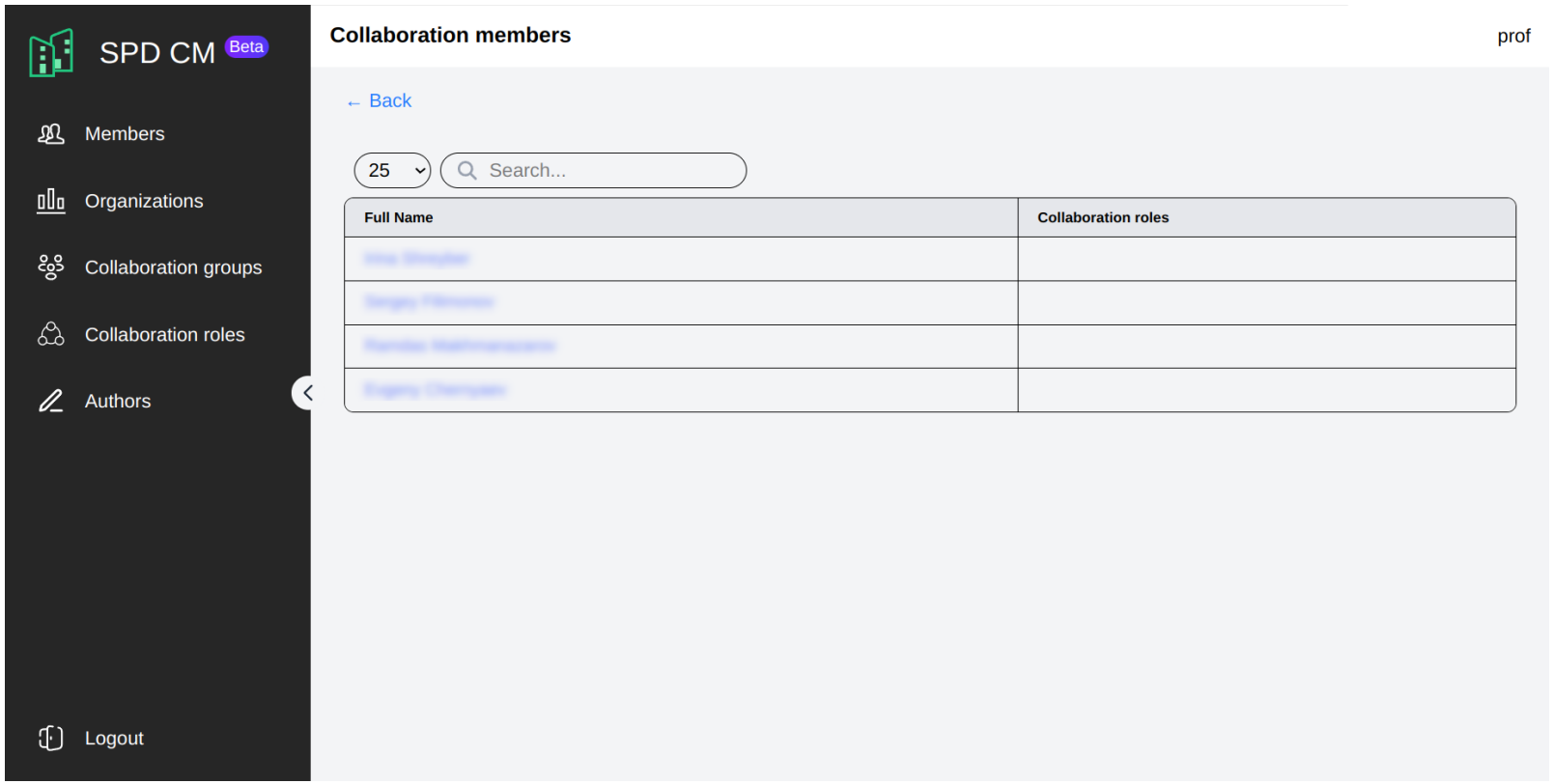
Collaboration Members Database: organizations

- List of participating organizations with number of members
- Clicking on organization name brings you organization details
- Clicking on number brings you list of members from this organization



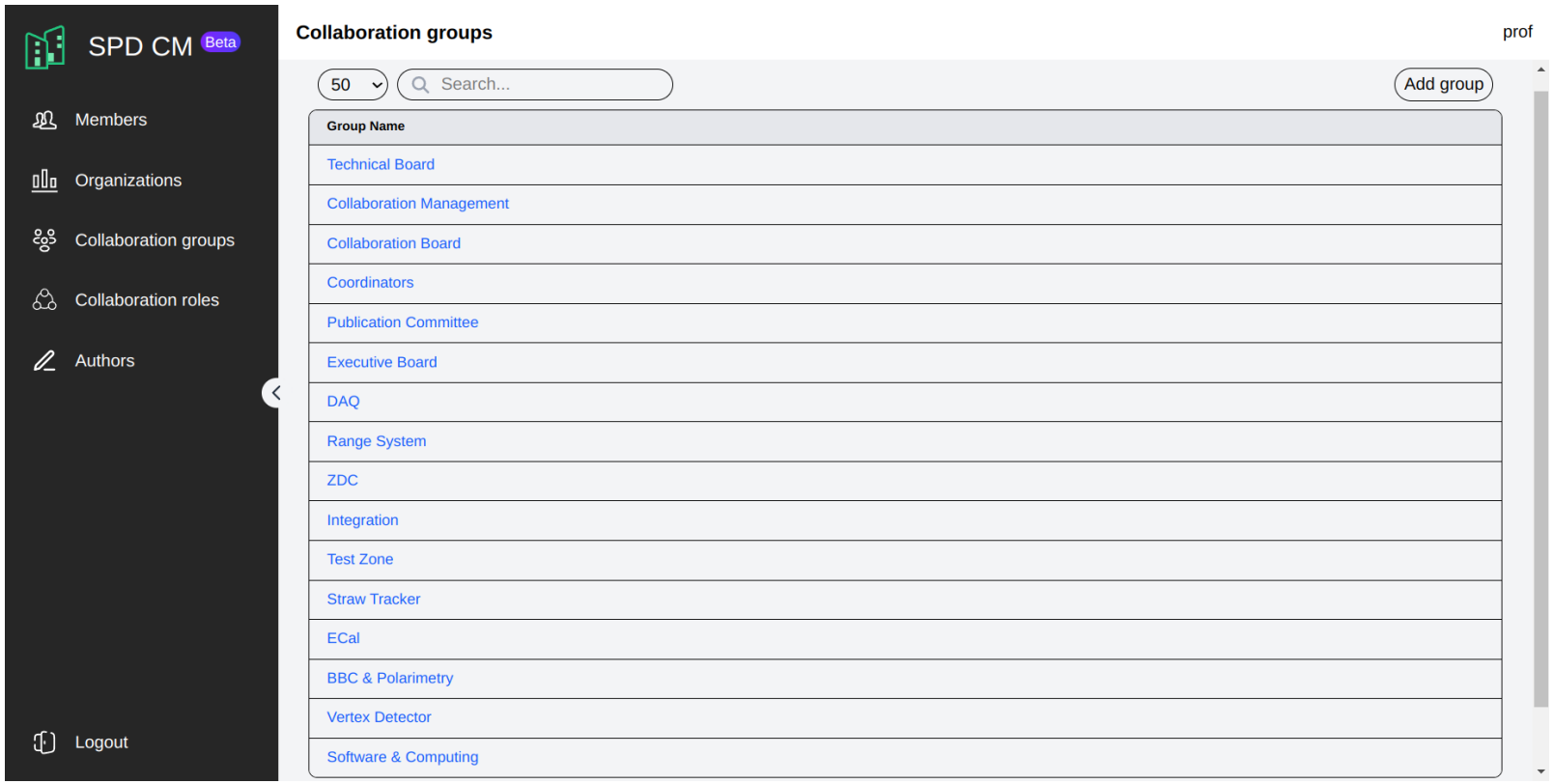
Collaboration Members Database: organizations

- List of participating organizations with number of members
- Clicking on organization name brings you organization details
- Clicking on number brings you list of members from this organization



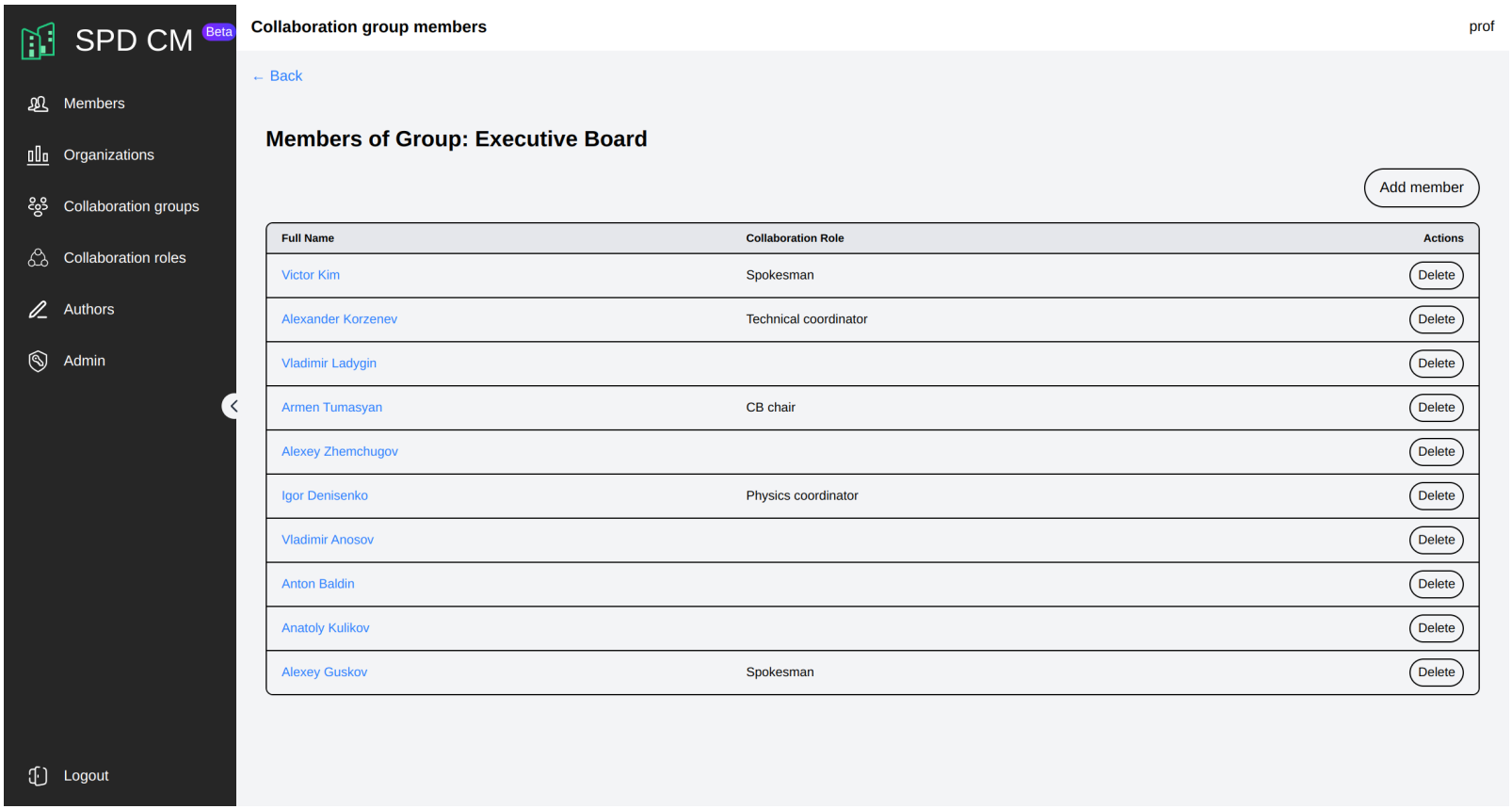
Collaboration Members Database: groups

- List of managements and working groups
- Clicking on group name bring you list of group members with their collaboration roles



Collaboration Members Database: groups

- List of managements and working groups
- Clicking on group name bring you list of group members with their collaboration roles



Collaboration Members Database: further development

- Next step will be generation of mailing lists for the groups of menbers
- A log of operations on the data in the system, including:
  - The date and time of the operation,
  - The type of operation (addition, modification, deletion),
  - The full name of the editor who made the changes.
- Further development of the role system and access rights management
- Other features may be implemented when need for them arises
- ... like accounting of members and organizations contributions

Registers

- Development of the SPD software and ongoing large scale production compaign requres creation of a number of databases.
  - Production register
  - Software version register
  - Geometry register
  - Magnetic field map register

Common features

- These are quite small projects, should be available fast
- No high loads expected
- Low update frequency
- Payloads are normaly stored separately, in files or sqlite
- Have to be developed independently, but need for data exchange may arise
- Similar platofrms should be used to simplify development and maintenance
- IAM should be used for authentification and access mangement

Production register

- Now the google sheet is being used: ▸ 2025 Productions RO
- Production record contains name, description, software version, number of requested events, initial seed and processing types.
  - For every processing type (simu, reco...) location of the macros are specified, as well as output datset templates and names.
- New production appear twice a week
- Need a database to keep records of the production
- Expected to be filled by 2-3 responsible persons
- Will be used by all members related to the analysis and production
- On the second stage of development automatic generation of the production tasks specification may be implemented

Software versions register

- Registry will keep information on versions of software used
- Contains software tags, packages used, version descriptions links to the repositories, configuration an condition data compatibilities
- Will be filled by the person responsible for the deployment of the software frameworks (SPDRoot, SAMPO)
- Information should be collected mostly from JINR git repositories
- Will be available through GUI and endpoints for use from other ISs

Geometries and Magnetic field maps register

Geometries register

- Geometry is being held as a single file and have to be used by a dedicated SAMPO framework tool
- DB should contain tag for a geometry, its version, metadata and a link to a file on the cvmfs
- Metadata content should be defined lately

Magnetic field maps registry

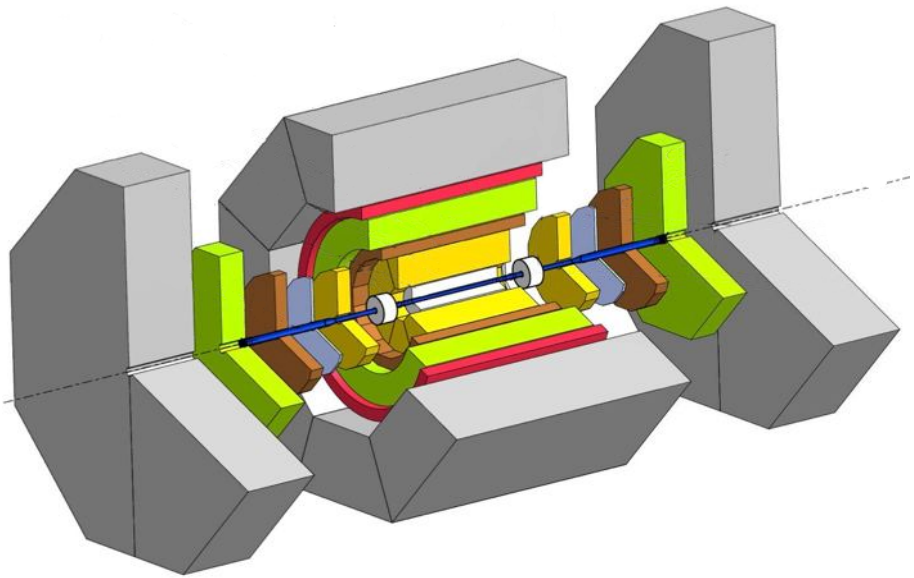
- B-field maps, each having few GB in size will be held in a sqlite DB.
- Updates should be quite rear, few times a year
- DB should contain tag for a map, its version, metadata and a pointer to the map (TBD)
- A special tool for SAMPO is being created to access B-field map, it will define metadata for the field maps and how it should be acessed.

Conclusion

- If you willing to participate in the Database, API or user Interface development, please join
- Information systems should be tailored to the needs of the project and to the nature and amount of data
- We need input both from hardware and analysis groups to create information systems fitting their need
  - If you created some database for your subsystem, please share your experience so it may be implemented elsewhere
  - If you have list of hardware (with parameters) that will be used in your system, contact us so we may adjust database and interface to it
  - The same if you group needs API to one of the information systems

BACKUP SLIDES

SPD Information



- The SPD Experiment have to produce large amounts of data, both collected from the detector and simulated
- The processing of the experimental data requires a wide variety of auxiliary information from many systems

- Huge volumes of the detector condition and management data should be stored in the databases
  - should be used in every stage of data taking, processing and analysis
  - are essential at nearly every stage of data handling
  - for use in number of versatile applications each with its own requirements