

Joint Institute for Nuclear Research

DocMS: Document Management Service on SciDocsCloud Platform

Ivan Sokolov*, Nikita Balashov

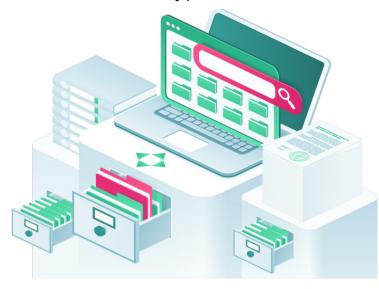
11th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID'2025)

Introduction



- Modern scientific research requires collaborative work on various document types
 - Articles
 - Abstracts
 - Presentations
 - Reports, etc.
- Scientific collaborations face several key challenges
 - Organizing safe and structured documents storage
 - Tracking changes of documents
 - Restricting access to documents
 - Sharing of documents

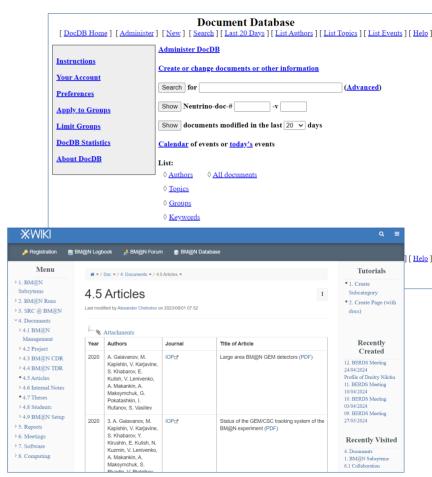
Many scientific projects generate a significant amount of documents (meta-information + files)



Background and Motivation



- Scientific communities use various software tools for document storage and collaboration
 - DocDB
 - Xwiki
 - CERN Document Server, etc.
- Several independent instances of the DocDB system are deployed at JINR
 - BM@N
 - Baikal-GVD
 - Shared instance for Neutrino projects at JINR
 - Accelerator Division No.1 of the VBLHEP accelerator complex
- Experience with these systems has shown key limitations
 - Not fully met requirements
 - Functional
 - Security and reliability
 - Complex support and improvement



Project Overview



The goal of the project is to provide a centralized, multi-community documentation storage service for collaborative activities

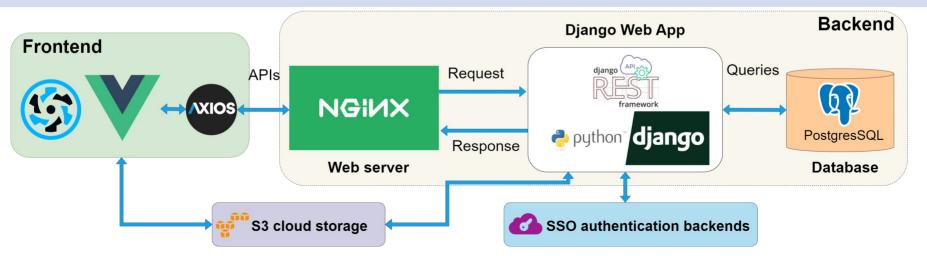
- Implement the core features of the DocDB system
- Enhance access rights system with flexible group and project-level controls
- Add new functionality based on collaboration requirements
- Specific Features Requested by Collaborations:
 - Per-document editors BM@N
 - Document discussion system BM@N SPD
 - Custom meta-data fields in documents BM@N
 - Advanced document filtering BM@N
 - Built-in event calendar Baikal-GVD Neutrino
 - Notification system All groups
 - Document review process BM@N

SciDocsCloud Software Platform



- Development of a more secure, reliable and feature-rich software platform to address the limitations of legacy DocDB system
 - Modern software development tools
 - Flexible Storage Backends for document files (LocalFS and S3-compatible Cloud storage options)
 - → Single-Sign-On (SSO) authentication

Free and open-source software with transparent development



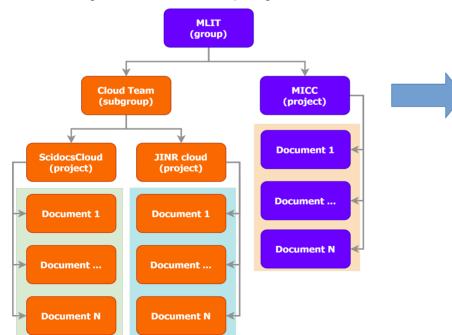
Organizational Structure

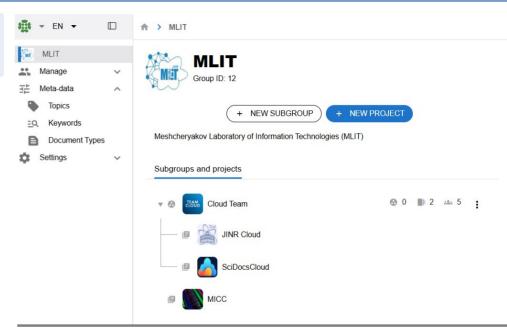


Hierarchical organization:

groups / subgroups → projects → documents

- Inspired by GitLab
 - Using GitLab terminology
 - Project ≈ GitLab project





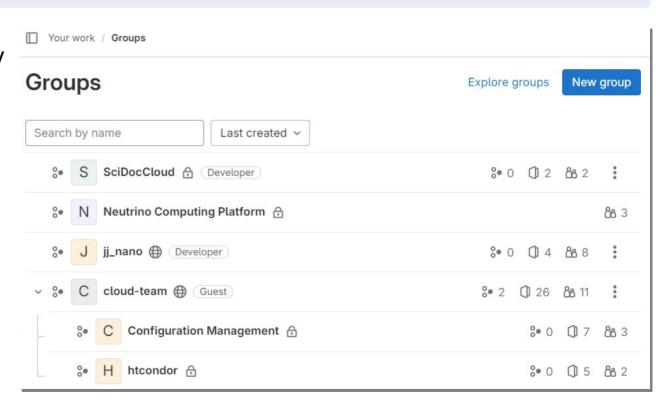
Organizational Structure



Hierarchical organization:

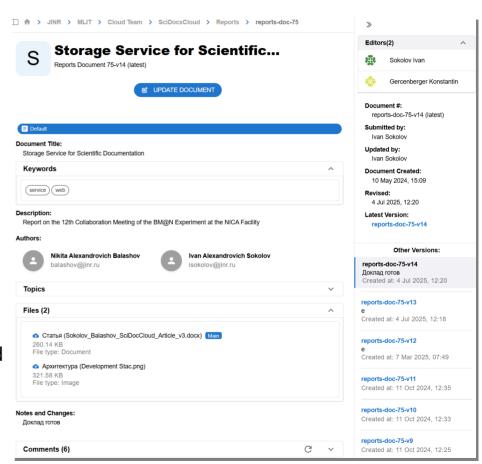
groups / subgroups → projects → documents

- Inspired by GitLab
 - Using GitLab terminology
 - Project ≈ GitLab project



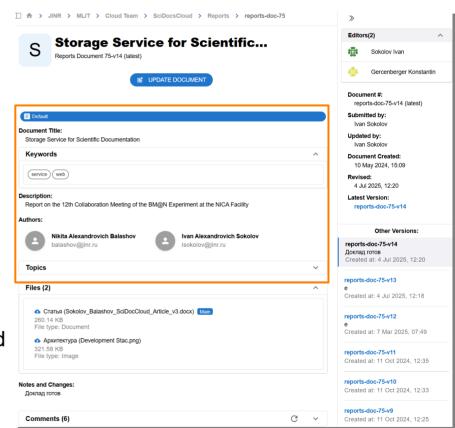


- Document contains
 - E Form with meta-data fields
 - File Collection
 - Version control
 - **Editors**
 - Comments
- Document templates defines the set of meta-data fields
 - Enable three predefined templates:
 Default, Article, Report
 - Custom templates are planned to be implemented later



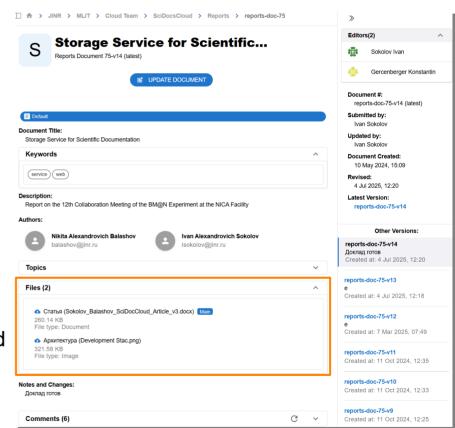


- Document contains
 - **form with meta-data fields**
 - File Collection
 - Version control
 - Editors
 - Comments
- Document templates defines the set of meta-data fields
 - Enable three predefined templates:
 Default, Article, Report
 - Custom templates are planned to be implemented later



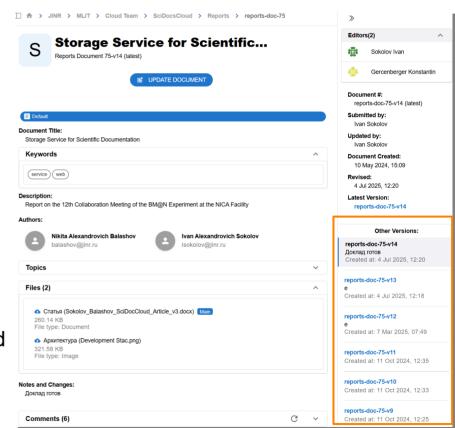


- Document contains
 - Form with meta-data fields
 - File Collection
 - Version control
 - Editors
 - Comments
- Document templates defines the set of meta-data fields
 - Enable three predefined templates:
 Default, Article, Report
 - Custom templates are planned to be implemented later



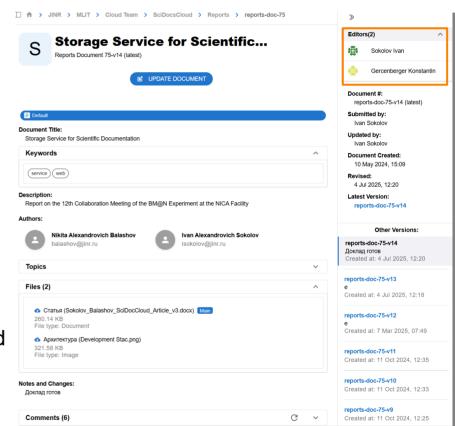


- Document contains
 - 🖺 Form with meta-data fields
 - File Collection
 - Version control
 - Editors
 - Comments
- Document templates defines the set of meta-data fields
 - Enable three predefined templates:
 Default, Article, Report
 - Custom templates are planned to be implemented later



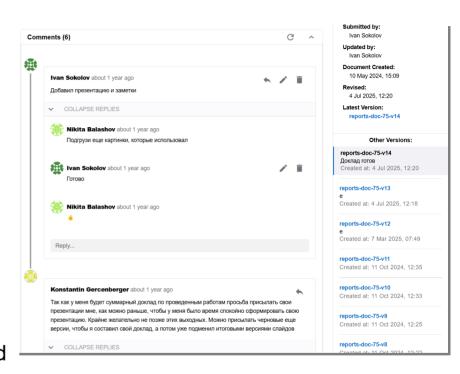


- Document contains
 - 🖺 Form with meta-data fields
 - File Collection
 - Version control
 - **Editors**
 - Comments
- Document templates defines the set of meta-data fields
 - Enable three predefined templates:
 Default, Article, Report
 - Custom templates are planned to be implemented later





- Document contains
 - Form with meta-data fields
 - File Collection
 - Version control
 - Editors
 - Comments
- Document templates defines the set of meta-data fields
 - Enable three predefined templates:
 Default, Article, Report
 - Custom templates are planned to be implemented later



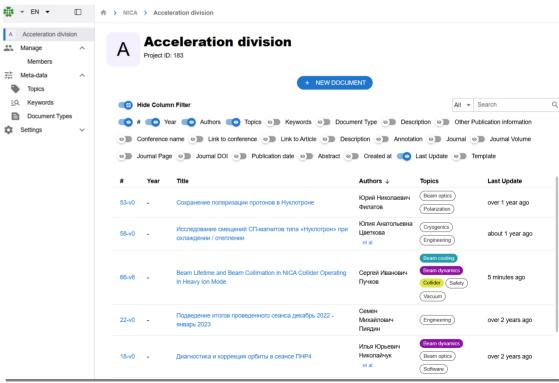


- Document contains
 - form with meta-data fields
 - File Collection
 - Version control
 - Editors
 - Comments
- Document templates defines the set of meta-data fields
 - Enable three predefined templates:
 Default, Article, Report
 - Custom templates are planned to be implemented later



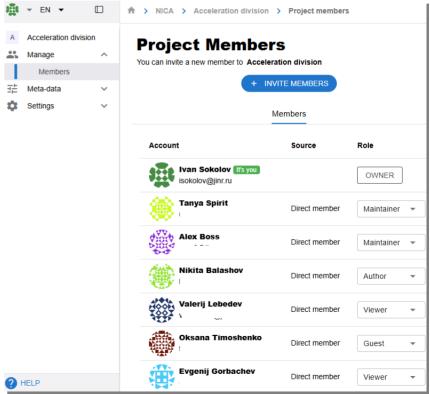


- Projects organize all the documents for collaboration and scientific projects
- Project is where the project members store and collaborate on documents
- Project Capabilities
 - Manage project membership
 - Advanced document search
 - Metadata management tools



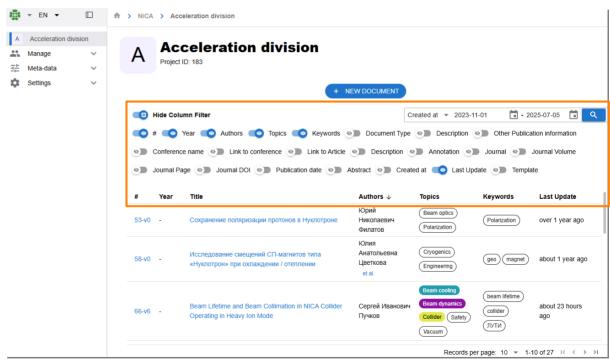


- Projects organize all the documents for collaboration and scientific projects
- Project is where the project members store and collaborate on documents
- Project Capabilities
 - Manage project membership
 - Advanced document search
 - Metadata management tools



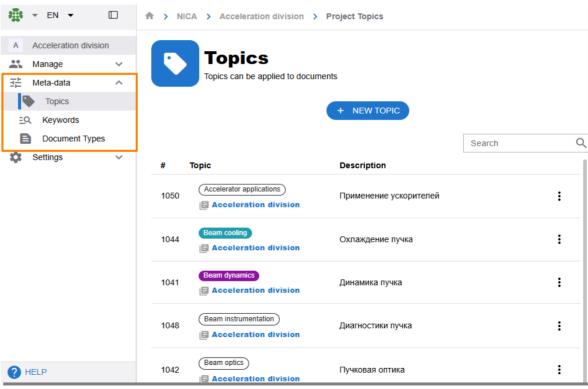


- Projects organize all the documents for collaboration and scientific projects
- Project is where the project members store and collaborate on documents
- Project Capabilities
 - Manage project membership
 - Advanced document search
 - Metadata management tools





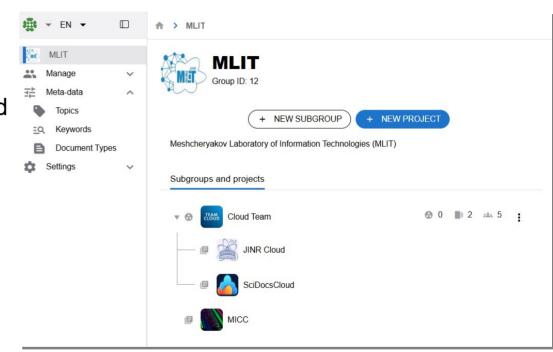
- Projects organize all the documents for collaboration and scientific projects
- Project is where the project members store and collaborate on documents
- Project Capabilities
 - Manage project membership
 - Advanced document search
 - Metadata management tools



What is Group?



- Groups organize projects
- Each project must belong to a group
- Groups help manage one or more related projects at the same time and permissions for projects
- For larger organizations groups can also have subgroups
- Subgroups/projects inherit parent group membership and permissions



Role-based Access Model



- Each member is assigned a role that determines their capabilities within projects and groups
- Currently implemented roles: Owner, Maintainer, Author, Viewer, Guest
- All subgroups and projects inherit parent group role
- Direct membership in subgroups or projects automatically grants Viewer role in all parent groups

Action	Guest	Viewer	Author	Maint	Owner
View project, groups	+	+	+	+	+
View documents		+	+	+	+
Use comment block			+	+	+
Create document			+	+	+
Update document			+-	+-	+
Create group				+	+
Create project				+	+
Create subgroup				+	+
Manage members				+	+
Manage topics				+	+
Manage keywords				+	+
Manage document type				+	+
Edit group				+	+
Edit project				+	+
Remove project					+
Remove group					+
Remove document					+
Manage editor list					+

Document Management Service (DocMS)

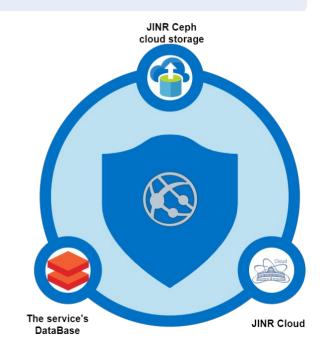


The SciDocsCloud platform-based service has been deployed at JINR as centralized service for storing, managing, organizing and collaborating on documents → **Document Management Service (DocMS)**

Infrastructure Details

- JINR's Ceph Cloud Storage
 - Triple replication for high reliability and data safety
 - Easily expandable storage to meet growing need
- Nightly Backups

 Automated backup of service meta information database
- Virtual Machine Deployment in the JINR Cloud
 - **Quick restoration after any hardware failure**



Document Management Service (DocMS)



- Integrated into the JINR Digital Ecosystem
- Available to JINR employees and collaborators via JINR Single Sign-On (SSO) → docs.jinr.ru

COMPLETED

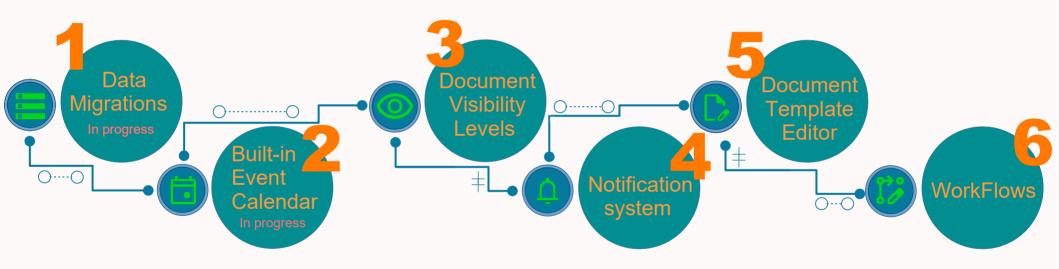
- Migration mechanism from DocDB system is ready
 - Accelerator Division No.1 of the VBLHEP COMPLETED
 - Documents of scientist from the FLNR
 - Baikal-GVD PLANNED
 - Neutrino projects at JINR





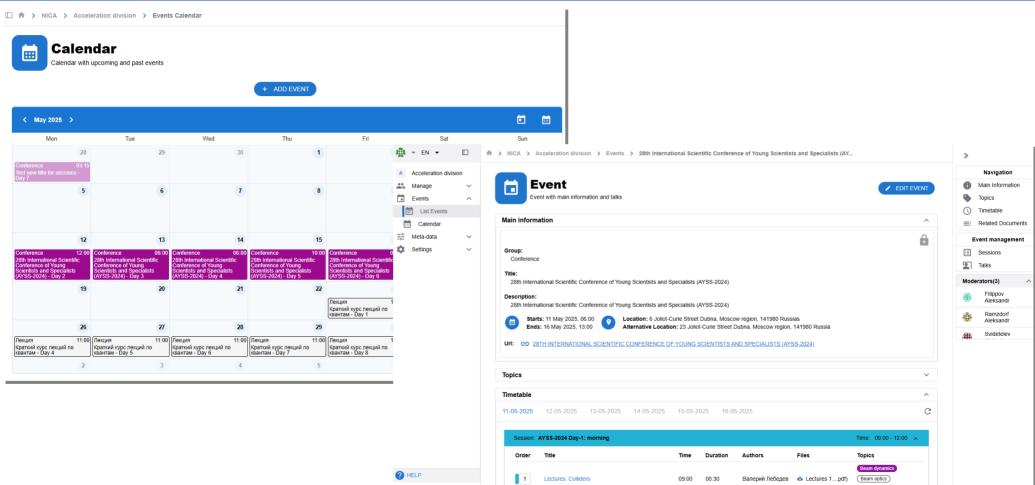
Road Map





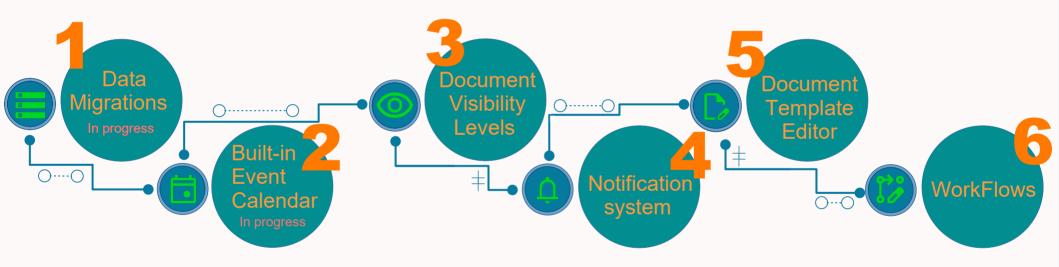
Road Map: Built-in Event Calendar coming soon





Road Map





Conclusion and FeedBack



- The SciDocsCloud platform is free and open-source software for storing and tracking of documents
- The Document Management Service for JINR research groups has been deployed based on SciDocsCloud
- The service is available to JINR employees and collaborators docs.jinr.ru and the JINR Digital Eco System
- Migration from legacy DocDB systems IN PROGRESS





For questions, bug reports, or suggestions contact us at docs@jinr.ru