

Joint Institute for Nuclear Research

Storage Service for Scientific Documentation

Ivan Sokolov*, Nikita Balashov

The 12th Collaboration Meeting of the BM@N Experiment at the NICA Facility

16 May 2024

Introduction



- Scientific groups collaborate on various types of documents
 - Articles
 - Abstracts
 - Presentations
 - Reports, etc
- Common challenges of collaborative work
 - Organizing safe and structure documents storage
 - Tracking changes to documents
 - Restricting access to documents
 - Sharing of documents
- Many scientific projects generate a significant amount of documents (files + revisions)



Background and Motivation



- Scientific communities use various software for document storage and collaboration
 - DocDB
 - Xwiki
 - CERN Document Server, etc.
- Several **independent** instances of the DocDB are deployed in the JINR
 - BM@N
 - Baikal-GVD
 - SPD
 - Participants of the JINR neutrino program
- Experience with these systems has shown:
 - Not fully meet requirements:
 - Functional
 - Security and reliability
 - Problems with data integrity
 - System support and improvement are complex processes

The screenshot shows the 'Document Database' interface. At the top, there are navigation links: [DocDB Home], [Administer], [New], [Search], [Last 20 Days], [List Authors], [List Topics], [List Events], and [Help]. Below this is a sidebar menu with links: Instructions, Your Account, Preferences, Apply to Groups, Limit Groups, DocDB Statistics, and About DocDB. The main content area is titled 'Administer DocDB' and includes a section for 'Create or change documents or other information'. This section has a search box with a 'Search for' label and an '(Advanced)' link. Below the search box are two 'Show' buttons: one for 'Neutrino-doc-#' with a dropdown menu and a '-v' checkbox, and another for 'documents modified in the last' with a dropdown menu set to '20' and a 'days' label. There is also a 'Calendar of events or today's events' link. At the bottom, a 'List:' section contains three links: 'Authors', 'All documents', and 'Groups'.

The screenshot shows the XWiki interface. At the top, there is a blue header with the XWiki logo and a search icon. Below the header, there are navigation links: Registration, BM@N Logbook, BM@N Forum, and BM@N Database. A 'Menu' sidebar is visible on the left, listing various categories like Subsystems, Runs, SRC, Documents, Management, Internal Notes, Theses, Students, Setup, Reports, Meetings, Software, and Computing. The main content area displays '4.5 Articles' and a table of attachments. The table has columns for Year, Authors, Journal, and Title of Article. The first row shows articles from 2020 by A. Galavanov, M. Kapishin, V. Karjavine, S. Khabarov, E. Kulish, V. Lenivenko, A. Makarkin, A. Maksymchuk, G. Pokotashkin, I. Rufanov, and S. Vasiliev, published in IOPJ, with the title 'Large area BM@N GEM detectors (PDF)'. The second row shows articles from 2020 by 3. A. Galavanov, M. Kapishin, V. Karjavine, S. Khabarov, Y. Kirushin, E. Kulish, N. Kuzmin, V. Lenivenko, A. Makarkin, A. Maksymchuk, S. Pevslov, and V. Plotnikov, published in IOPJ, with the title 'Status of the GEM/CSC tracking system of the BM@N experiment (PDF)'. On the right side, there are sections for 'Tutorials' (1. Create Subcategory, 2. Create Page (with docs)), 'Recently Created' (12. BERDS Meeting 24/04/2024, Profile of Dmitry Nikitin 11. BERDS Meeting 10/04/2024, 10. BERDS Meeting 03/04/2024, 09. BERDS Meeting 27/03/2024), and 'Recently Visited' (4. Documents, 1. BM@N Subsystems, 6.1 Collaboration).

Project Overview

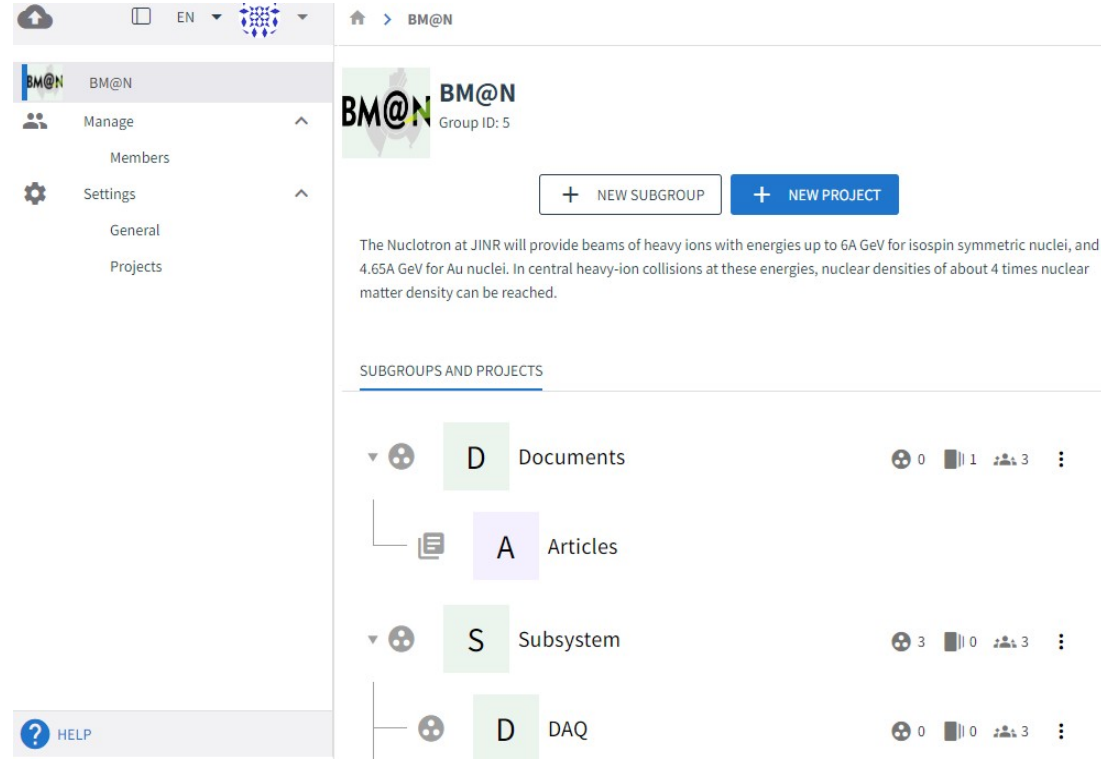


- The goal of the project is to provide a centralized multi-community documentation storage service for collaborative activities
 - Implement the core features of DocDB
 - Extend system to support group and project with flexible access control
- Specific requests from [BM@N](#)
 - Organization of collaborate on publication activities and reports
 - Defining per-document editors
 - Document discussion system
 - Document review process
 - Custom meta-data fields in documents
 - Email notification system

Organization Structure in Service



- **Hierarchical organization:**
groups / subgroups → *projects* → *documents*
- The idea comes from GitLab
 - Using GitLab terminology
 - Project ≈ GitLab repository
- Project contains
 - Document storage
 - Collaboration tools
- Document as an abstract entity
 - Form with metadata fields
 - Collection of files
 - Version control



Organization Structure in Service



- Hierarchical organization
groups / subgroups → *projects* → *documents*
- **The idea comes from GitLab**
 - Using GitLab terminology
 - Project ≈ GitLab repository
- Project contains
 - Document storage
 - Collaboration tools
- Document as an abstract entity
 - Form with metadata fields
 - Collection of files
 - Version control

The screenshot shows the 'Groups' page in GitLab. At the top, there is a breadcrumb 'Your work / Groups' and a 'New group' button. Below the title 'Groups', there are search and filter options: 'Search by name' and 'Last created'. The main content is a list of groups with their names, roles, and statistics. The groups listed are:

Group Name	Role	Members	Projects	Repositories
SciDocCloud	Developer	0	2	2
Neutrino Computing Platform				3
jj_nano	Developer	0	4	8
cloud-team	Guest	2	26	11
Configuration Management		0	7	3
htcondor		0	5	2

Organization Structure in Service



- Hierarchical organization
groups / subgroups → *projects* → *documents*
- The idea comes from GitLab
 - Using GitLab terminology
 - Project ≈ GitLab repository
- **Project contains**
 - Document storage
 - Collaboration tools
- Document as an abstract entity
 - Form with metadata fields
 - Collection of files
 - Version control

BM@N > Documents > Articles

A **Articles**
Project ID: 15

+ NEW DOCUMENT

Configure this list Search

Title Description Abstract Authors Created at Last Update

#	Title	Abstract	Authors	Last Update
3-v1	High-Performance Optimization of Algorithms Used in the BM@N Experiment of the NICA Project	Results of high-performance optimization of BmnRoot software modules are presented. The BmnRoot package used in the BM@N experiment of the NICA projec...	Sergei Merts et al.	5 days ago
4-v0	Large area BM@N GEM detectors		A Galavanov	3 minutes ago
5-v0	Исследование характеристик GEM детекторов эксперимента BM@N	BM@N experiment (Baryonic Matter at Nuclotron) is aimed to study core-core (up to gold-gold) collisions in extreme conditions. High intensity of inter...	C Васильев et al.	1 minute ago

[Development of the BM@N Web](#)

Organization structure in service



- Hierarchical organization
groups / subgroups → *projects* → *documents*
- The idea comes from GitLab
 - Using GitLab terminology
 - Project ≈ GitLab repository
- Project contains
 - Document storage
 - Collaboration tools
- **Document as an abstract entity**
 - Form with metadata fields
 - Collection of files
 - Version control

Home > JINR > MLIT > Cloud Team > SciDocsCloud > Reports > reports-doc-75 > version 4

Storage Service for Scientific...

Reports Document 75-v4

UPDATE DOCUMENT

Report on the 12th Collaboration Meeting of the BM@N Experiment at the NICA Facility

Authors:

- Nikita Alexandrovich Balashov (balashov@jinr.ru)
- Ivan Alexandrovich Sokolov (isokolov@jinr.ru)

Editors:

- Sokolov Ivan (isokolov@jinr.ru)
- Gercenberger Konstantin (gertsen@jinr.ru)
- Balashov Nikita (balashov@jinr.ru)

Notes and Changes:

Added all material for presentation + presentation + notes

Files:

File Name	Size	File Type
Notes.odt	61.48 KB	Document
Workflow_Article.png	1.64 MB	Image
sokolov_presentation.odp	5.43 MB	Presentation
Workflow_report.png	1.71 MB	Image

Comments (6)

Ivan Sokolov 29 minutes ago
Добавил презентацию и заметки

3 REPLIES Last reply by Nikita Balashov 28 minutes ago

Document #: reports-doc-75-v4

Submitted by: Ivan Sokolov

Updated by: Konstantin Gercenberger

Document Created: 10 May 2024, 18:09

Revised: 10 May 2024, 18:11

Latest Version: reports-doc-75-v5

Other Versions: reports-doc-75-v5 (Created at: 10 May 2024, 18:35)

reports-doc-75-v3 (Created at: 10 May 2024, 18:10)

reports-doc-75-v2 (Created at: 10 May 2024, 18:10)

Organization structure in service



- Hierarchical organization
groups / subgroups → *projects* → *documents*
- The idea comes from GitLab
 - Using GitLab terminology
 - Project ≈ GitLab repository
- Project contains
 - Document storage
 - Collaboration tools
- **Document as an abstract entity**
 - ***Form with metadata fields***
 - Collection of files
 - Version control

The screenshot displays a document management interface for a document titled "Storage Service for Scientific..." (Reports Document 75-v4). The breadcrumb navigation shows the path: JINR > MLIT > Cloud Team > SciDocsCloud > Reports > reports-doc-75 > version 4. The document's metadata includes the title, authors (Nikita Alexandrovich Balashov and Ivan Alexandrovich Sokolov), and a "Notes and Changes" section stating "Added all material for presentation + presentation + notes". The "Files" section lists four files: Notes.odt (61.48 KB), Workflow_Article.png (1.64 MB), sokolov_presentation.odp (5.43 MB), and Workflow_report.png (1.71 MB). The "Editors" section lists three users: Sokolov Ivan, Gercenberger Konstantin, and Balashov Nikita. The "Comments" section shows a comment by Ivan Sokolov from 29 minutes ago, with 3 replies. The right sidebar displays document metadata: Document # (reports-doc-75-v4), Submitted by (Ivan Sokolov), Updated by (Konstantin Gercenberger), Document Created (10 May 2024, 18:09), Revised (10 May 2024, 18:11), Latest Version (reports-doc-75-v5), and Other Versions (reports-doc-75-v5, reports-doc-75-v3, reports-doc-75-v2).

Organization structure in service



- Hierarchical organization
groups / subgroups → *projects* → *documents*
- The idea comes from GitLab
 - Using GitLab terminology
 - Project ≈ GitLab repository
- Project contains
 - Document storage
 - Collaboration tools
- **Document as an abstract entity**
 - Form with metadata fields
 - **Collection of files**
 - Version control

The screenshot displays a document management interface for a document titled "Storage Service for Scientific...". The breadcrumb navigation shows the path: JINR > MLIT > Cloud Team > SciDocsCloud > Reports > reports-doc-75 > version 4. The document is identified as "Reports Document 75-v4" and includes an "UPDATE DOCUMENT" button.

Authors: Nikita Alexandrovich Balashov (balashov@jinr.ru) and Ivan Alexandrovich Sokolov (isokolov@jinr.ru).

Notes and Changes: Added all material for presentation + presentation + notes.

Files: A red box highlights a list of files: "Notes.odt" (61.48 KB, Document), "sokolov_presentation.odp" (5.43 MB, Presentation), "Workflow_Article.png" (1.64 MB, Image), and "Workflow_report.png" (1.71 MB, Image). Each file has an upload icon.

Editors: Sokolov Ivan (isokolov@jinr.ru), Gercenberger Konstantin (gertsen@jinr.ru), and Balashov Nikita (balashovn@jinr.ru).

Comments: Ivan Sokolov (29 minutes ago) commented: "Добавил презентацию и заметки". There are 3 replies, with the last reply by Nikita Balashov (28 minutes ago).

Metadata: Document #: reports-doc-75-v4, Submitted by: Ivan Sokolov, Updated by: Konstantin Gercenberger, Document Created: 10 May 2024, 18:09, Revised: 10 May 2024, 18:11, Latest Version: reports-doc-75-v5, Other Versions: reports-doc-75-v5 (Created at: 10 May 2024, 18:35), reports-doc-75-v3 (Created at: 10 May 2024, 18:10), reports-doc-75-v2 (Created at: 10 May 2024, 18:10).

Organization structure in service



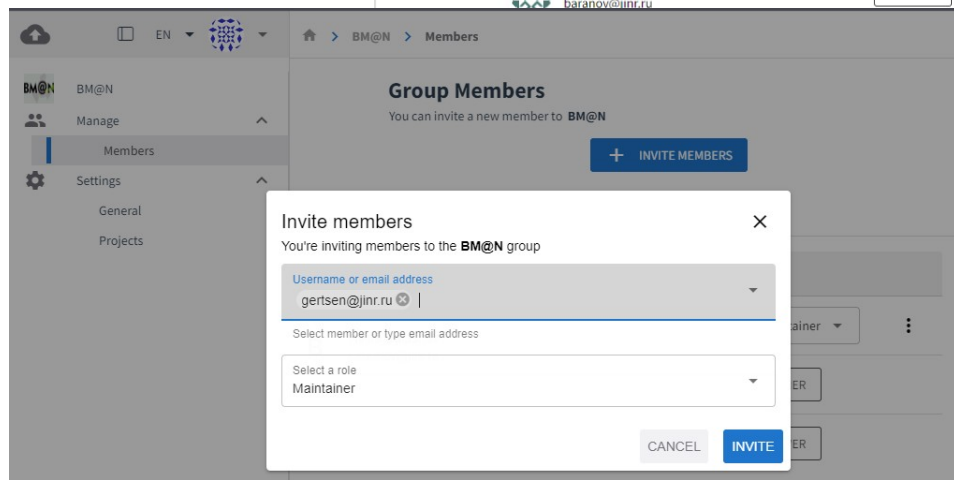
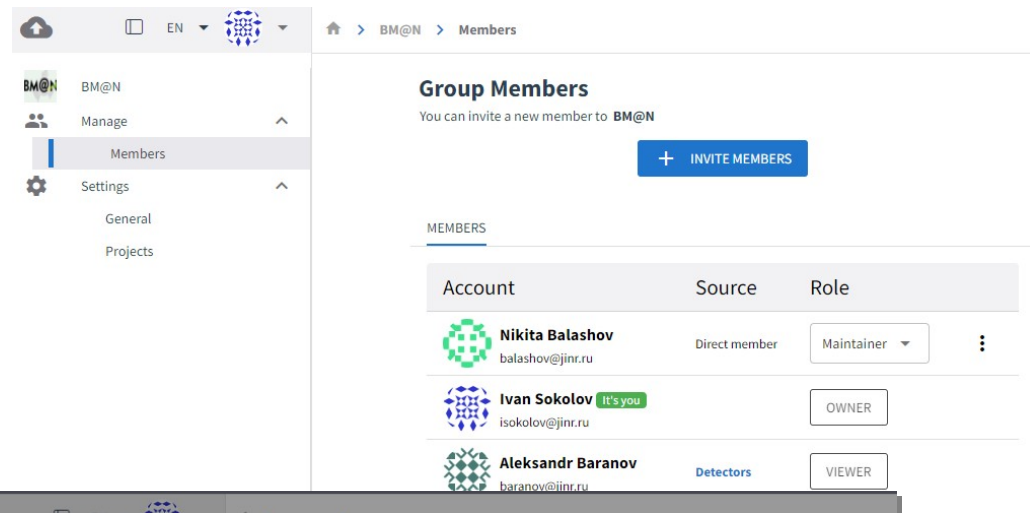
- Hierarchical organization
groups / subgroups → *projects* → *documents*
- The idea comes from GitLab
 - Using GitLab terminology
 - Project ≈ GitLab repository
- Project contains
 - Document storage
 - Collaboration tools
- **Document as an abstract entity**
 - Form with metadata fields
 - Collection of files
 - **Version control**

The screenshot displays a document management interface for a document titled "Storage Service for Scientific...". The breadcrumb navigation shows the path: JINR > MLIT > Cloud Team > SciDocsCloud > Reports > reports-doc-75 > version 4. The document is identified as "Reports Document 75-v4" and includes an "UPDATE DOCUMENT" button. The main content area shows a report on the 12th Collaboration Meeting of the BM@N Experiment at the NICA Facility. It lists authors: Nikita Alexandrovich Balashov (balashov@jinr.ru) and Ivan Alexandrovich Sokolov (isokolov@jinr.ru). A section for "Notes and Changes" indicates that all material for presentation and notes was added. The "Files" section lists: Notes.odt (61.48 KB, Document), Workflow_Article.png (1.64 MB, Image), sokolov_presentation.odp (5.43 MB, Presentation), and Workflow_report.png (1.71 MB, Image). An "Editors" list includes Sokolov Ivan (isokolov@jinr.ru), Gercenberger Konstantin (gertsen@jinr.ru), and Balashov Nikita (balashovn@jinr.ru). A "Comments (6)" section shows a comment from Ivan Sokolov 29 minutes ago with 3 replies. On the right, a sidebar displays document metadata: Document # (reports-doc-75-v4), Submitted by (Ivan Sokolov), Updated by (Konstantin Gercenberger), Document Created (10 May 2024, 18:09), Revised (10 May 2024, 18:11), and a "Latest Version" (reports-doc-75-v5) along with "Other Versions" (reports-doc-75-v5, reports-doc-75-v3, reports-doc-75-v2) with their respective creation dates.

Role-based Access Model

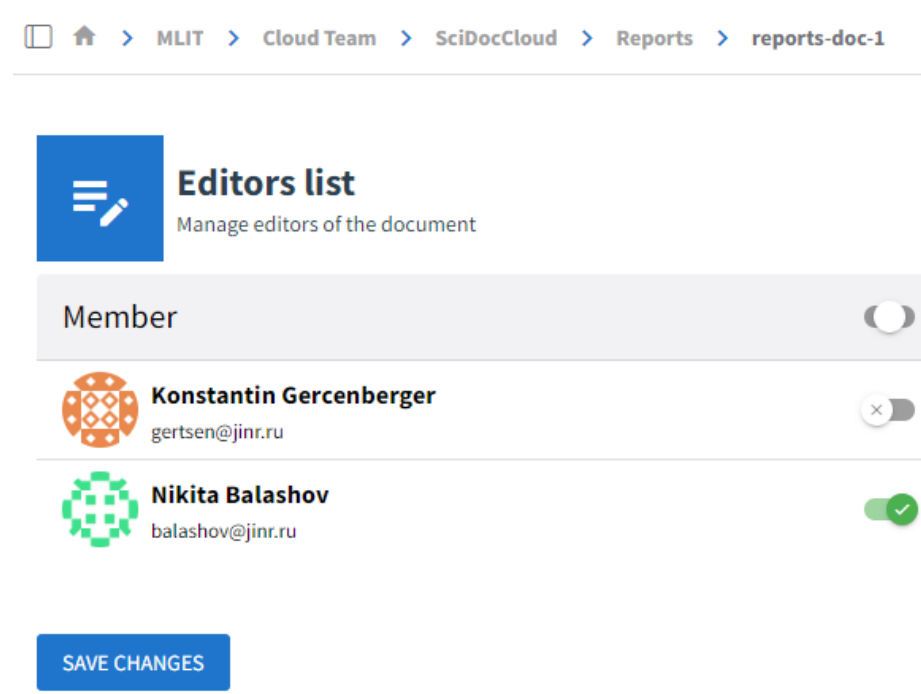


- Access restriction to groups or projects based on roles similar to GitLab
- Currently implemented roles Owner, Maintainer and Viewer
 - Each has their own permission
 - In plan to add new roles
- Group/project membership
- Nested groups membership





- *Defining per-document editors* → **Implemented the "Editors List" feature to manage document editing permissions (only owner)**
- Document discussion system → Added the comments section to the document (all members)
- Custom meta-data fields → Added fields for journal references



Realized Features for BM@N



- Defining per-document editors → Implemented the "Editors List" feature to manage document editing permissions (only owner)
- *Document discussion system* → **Added the comments section to the document (all members)**
- Custom meta-data fields → Added fields for journal references

The screenshot shows a 'Comments (6)' section on a document page. It features two visible comments. The first comment is from Ivan Sokolov, posted 29 minutes ago, with the text 'Добавил презентацию и заметки'. Below it, a summary indicates '3 REPLIES' with the last reply by Nikita Balashov about 2 hours ago. The second comment is from Konstantin Gercenberger, posted 25 minutes ago, with the text 'Так как у меня будет суммарный доклад по проведенным работам просьба присылать свои презентации мне, как можно раньше, чтобы у меня было время спокойно сформировать свою презентацию. Крайне желательно не позже этих выходных. Можно присылать черновые еще версии, чтобы я составил свой доклад, а потом уже подменил итоговыми версиями слайдов'. Below this comment is a 'COLLAPSE REPLIES' button. At the bottom of the comments section, there is a text input field with a 'Reply...' placeholder and a blue 'COMMENT' button. Above the input field is a rich text editor toolbar with icons for bold, italic, link, underline, text color, font color, bulleted list, numbered list, indent, and outdent.

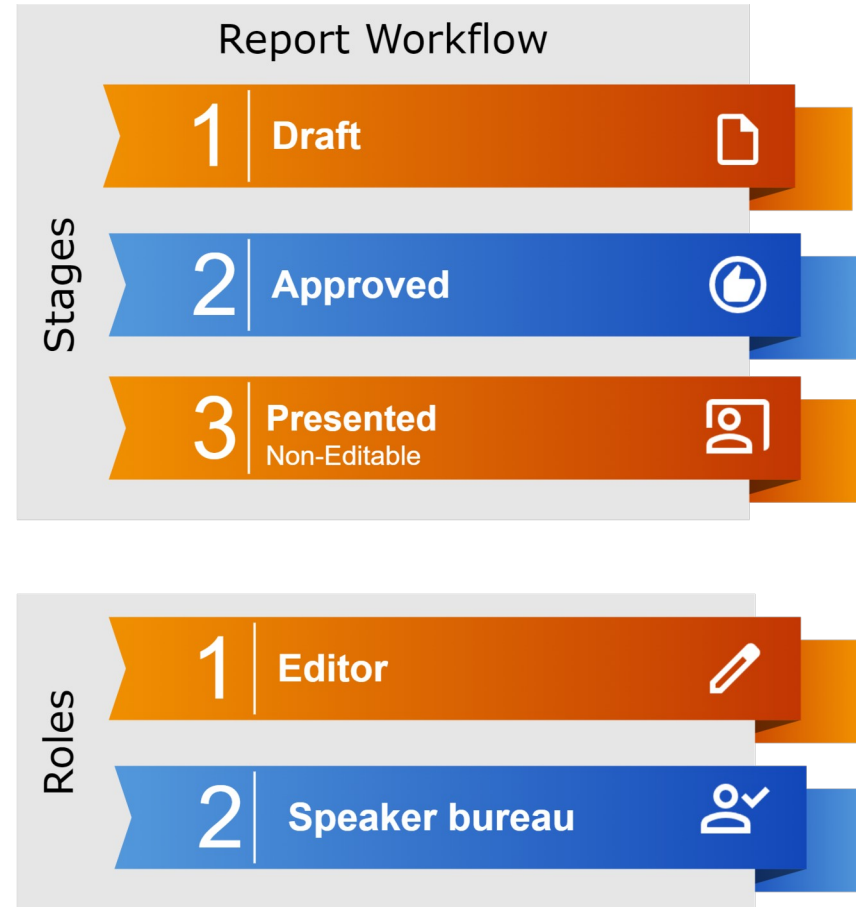
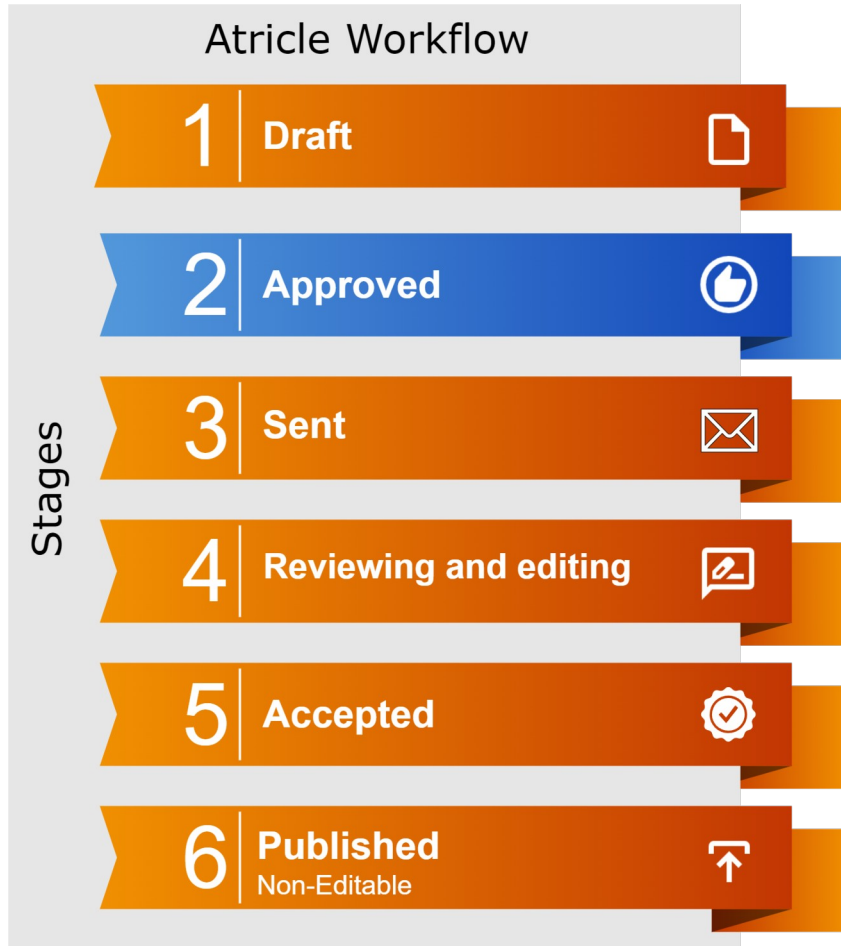
Realized Features for BM@N



- Defining per-document editors → Implemented the "Editors List" feature to manage document editing permissions (only owner)
- Document discussion system → Added the comments section to the document (all members)
- *Custom meta-data fields* → **Added fields for journal references**

The screenshot displays a document editing interface. At the top, there are sections for "Notes and Changes" and "Other Publication information". A red rectangular box highlights the "Journal References" section, which contains three input fields labeled "Journal", "Volume", and "Page". Below this section is a "File upload" button with a plus icon and the text "0.0B / 0.00%". At the bottom of the interface, there are two buttons: "CANCEL" and "+ UPDATE".

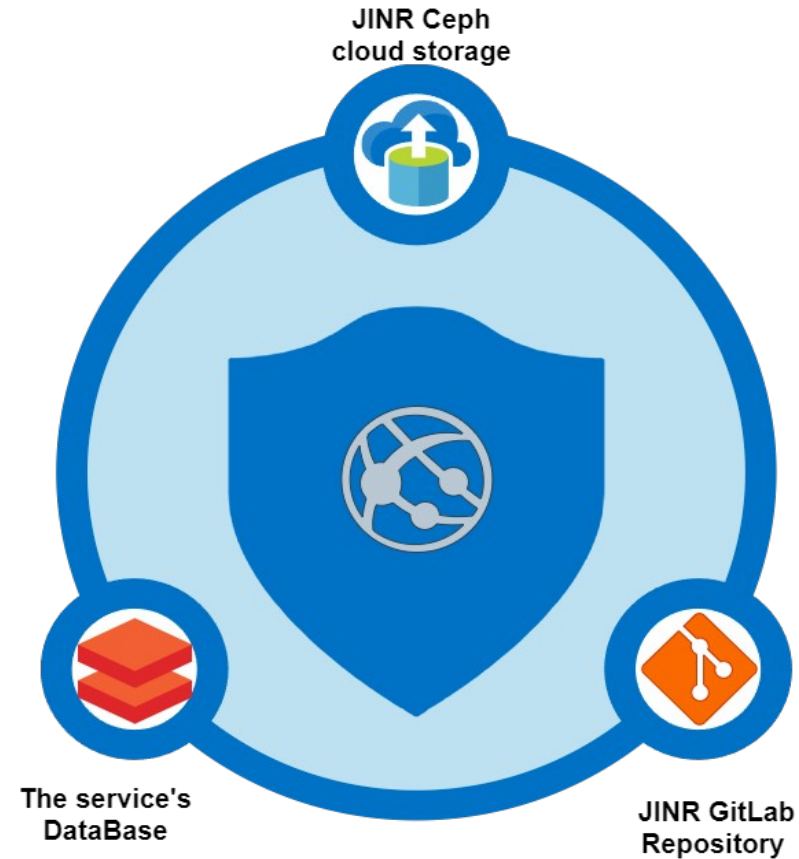
Document Review Workflow



Data Availability and Consistency



- Files for documents are stored in the JINR's Ceph cloud storage
 - Upload and Download files via S3
 - Triple replication provided by Ceph
 - Easily scalable
- Nightly Backups for Service Meta Information Database
- Service is deployed on a VM in the JINR Cloud:
 - Quick recovery after hardware failure
- Automated deployment via GitLab CI/CD → in plans



Road Map



- Document meta-data Templates → in progress
- Subscribe/notification system
- Namespace and document visibility levels
- Data migration from JINR DocDBs
- Workflows
- Integration with Indico (Event Calendar)





- The service is available for use → docs.jinr.ru
- Please report bugs and suggest improvements via email → docs@jinr.ru

**Your feedback helps us to
improve our service and makes it
better!**

