

Coscine

Presentation by Katja Jansen

05.11.2024
Joint NHR Data
Management Training



Der Foliensatz insgesamt und seine Texte stehen als Gesamtwerk unter der Lizenz [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/). Einzelne Elemente wie Abbildungen sind ggf. eigenständig lizenziert.





General Information



Resources & Quota



Collaboration



Questions & Comments



General Information

Wishful Thinking

Idea



via freepik - https://de.freepik.com/foto-kostenlos/westzaehen-schwarz-perforiertes-papier-diamant_2087717.htm#from_view=diamant&from_view=westzaehen-schwarz-perforiertes-papier auf Freepik

Experiment

Data



Bild von - via freepik - https://de.freepik.com/foto-kostenlos/diamantansammlung-auf-rosa-hintergrund_31500626.htm#from_view=diamant&from_view=rosa-hintergrund_31500626.htm auf Freepik

Reality

Idea



via freepik - https://de.freepik.com/foto/kostenlos/westzaehen-schwarz-perforiertes-papier-diamond_2087777.htm#from_view=detail&query=position=1&from_view=search&track=ais - Bild von reapiet.com via auf Freepik

1st Experiment fails



2nd Experiment fails



Data from Experiments is lost



Details on Experiments are missing



Data



Bild von ca freepik - https://de.freepik.com/foto/kostenlos/diamantarrangierung-auf-rosa-hintergrund_31500626.htm#from_view=detail&query=diamant&position=1&from_view=search&track=ais - Bild von reapiet.com via auf Freepik.com

Reality

Idea



via freepik - <https://de.freepik.com/foto/kostenlos/westzaehen-schwarz-perforiertes-papier>
idamark_2087777.html?query=diamant&from_view=search&track=ais">Bild von twapiet.com via auf Freepik

1st Experiment fails



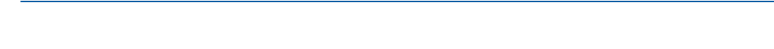
2nd Experiment fails



Data from Experiments is lost



Details on Experiments are missing



Data



Bild von ca freepik - https://de.freepik.com/foto/kostenlos/diamantarrangement-auf-rotem-hintergrund_31500528.html#query=diamant&from_view=search&track=ais">Bild von ca freepik.com

Reproducibility?
Reuse?

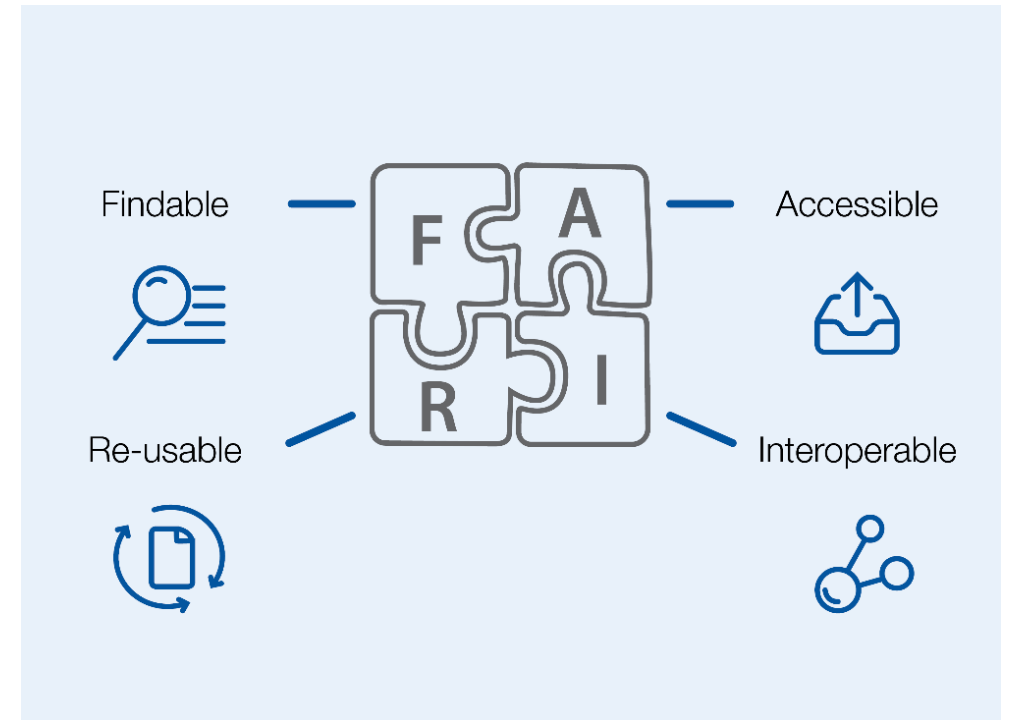
Share Knowledge (Publications etc.)?

Problem

- **FAIR** handling of research data is not part of a researcher's daily routine
- Only a small proportion, if any, of a project's research data is handled according to **FAIR** principles
 - Loss of metadata

Our solution

- Coscine - an environment for data management that implicitly implements the **FAIR** principles



Coscine – Another Motivation...

Problem

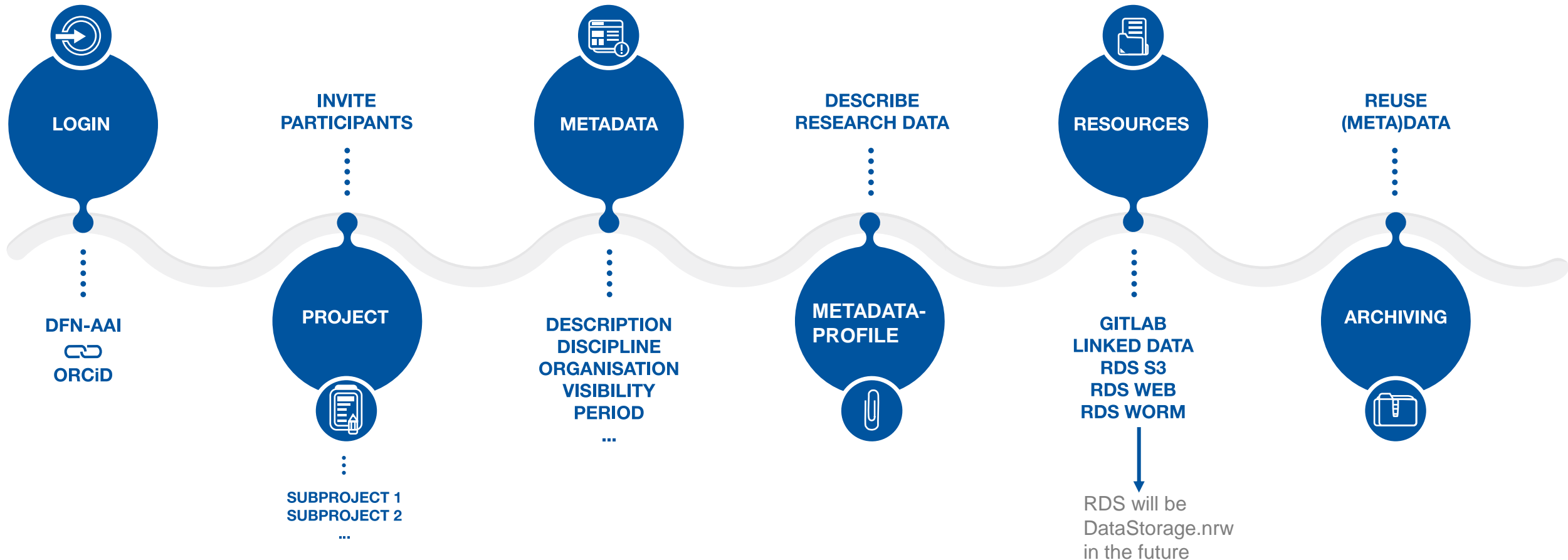
- Researchers need a **lot of storage space**

Our solution

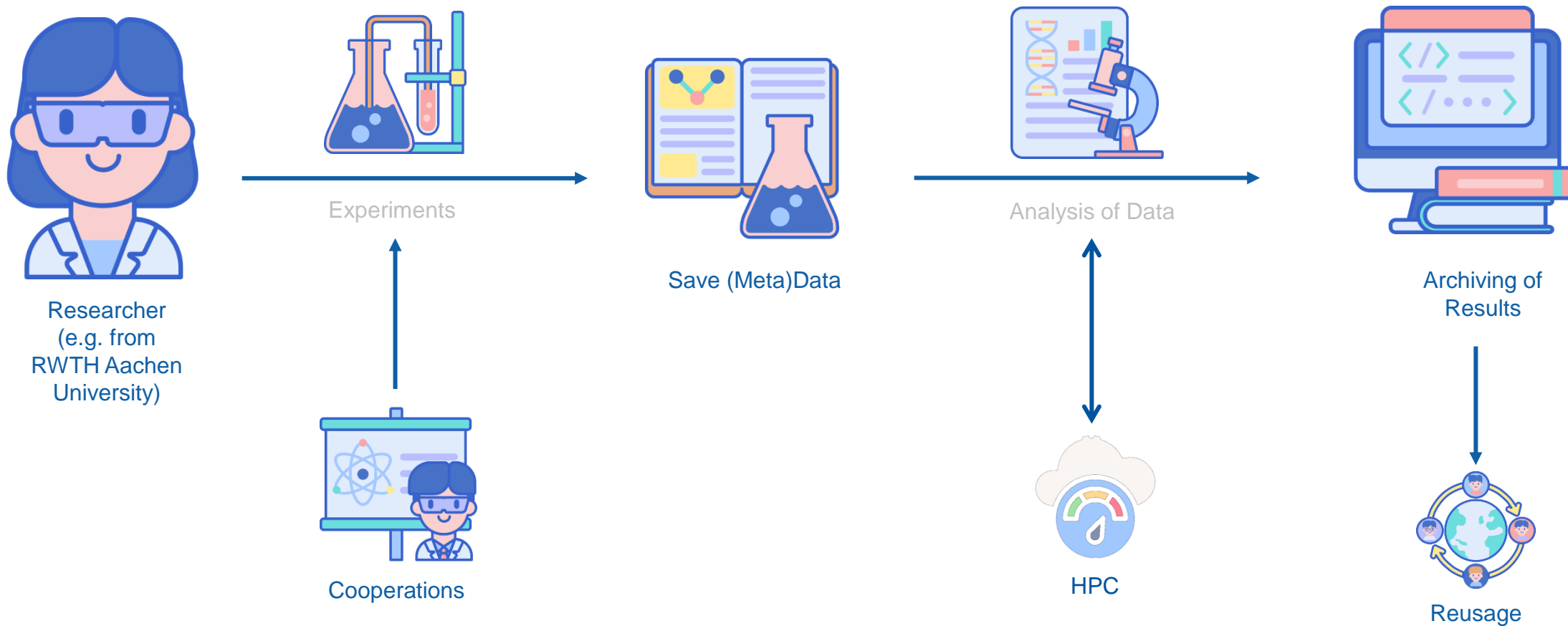
- (Research)DataStorage allows up to 125 TB of storage space per application – accessible via Coscine
 - Demand: Research Data Management
- **More input regarding Storage (Applications) will follow during this Workshop**



Bild von [Dmitriy](#) auf [Pixabay](#)

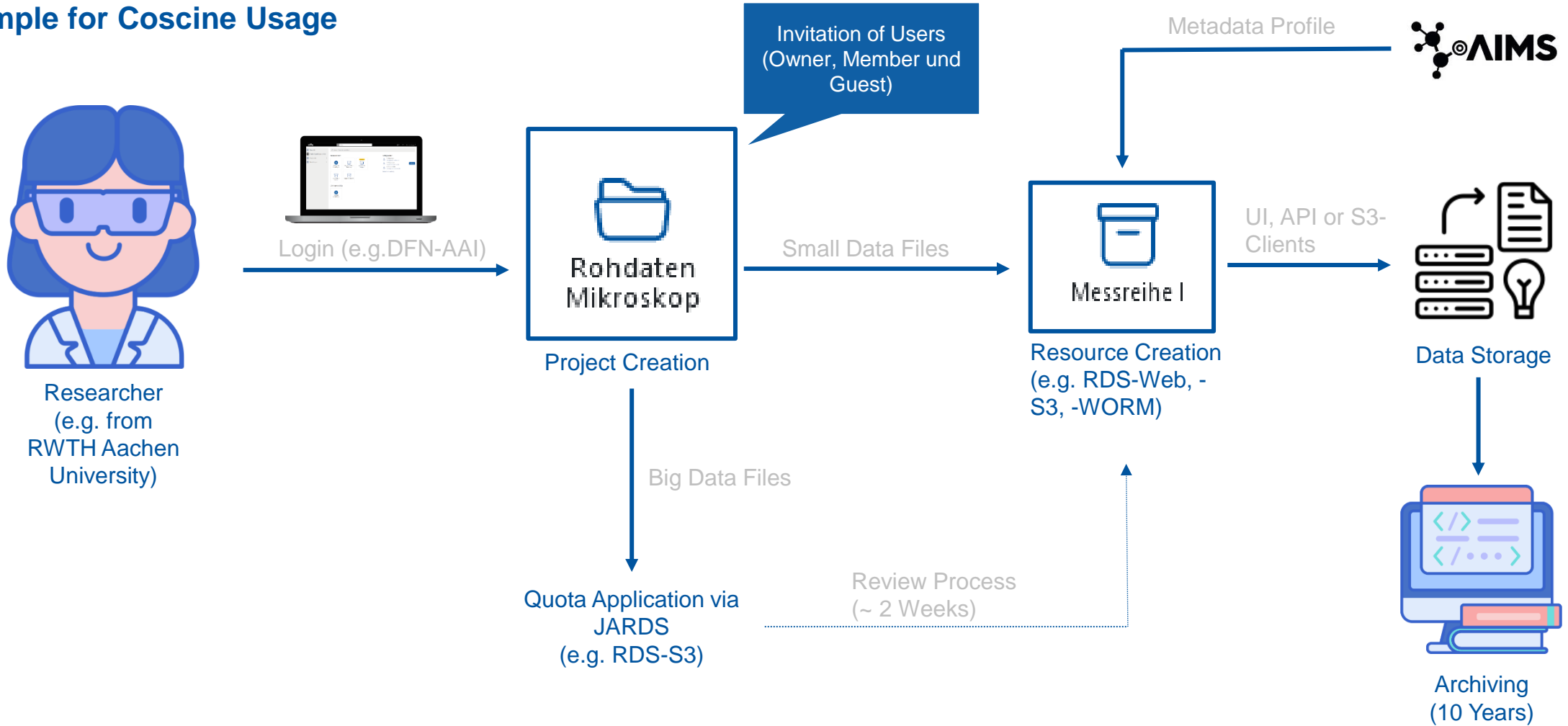


Example for Coscine Usage



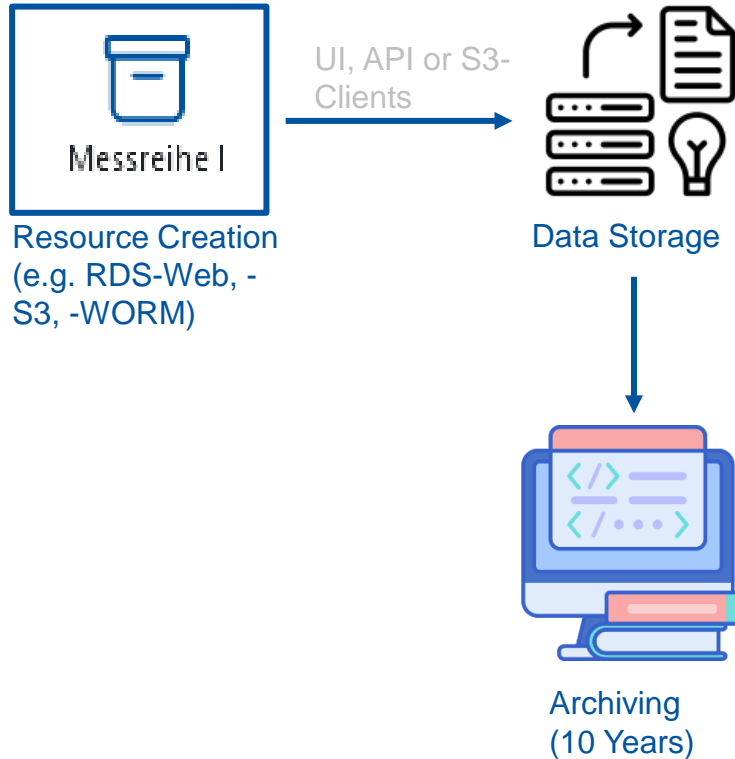
Icons by Flaticon (Flat Icons)

Example for Coscine Usage



Icons by Flaticon (Flat Icons)

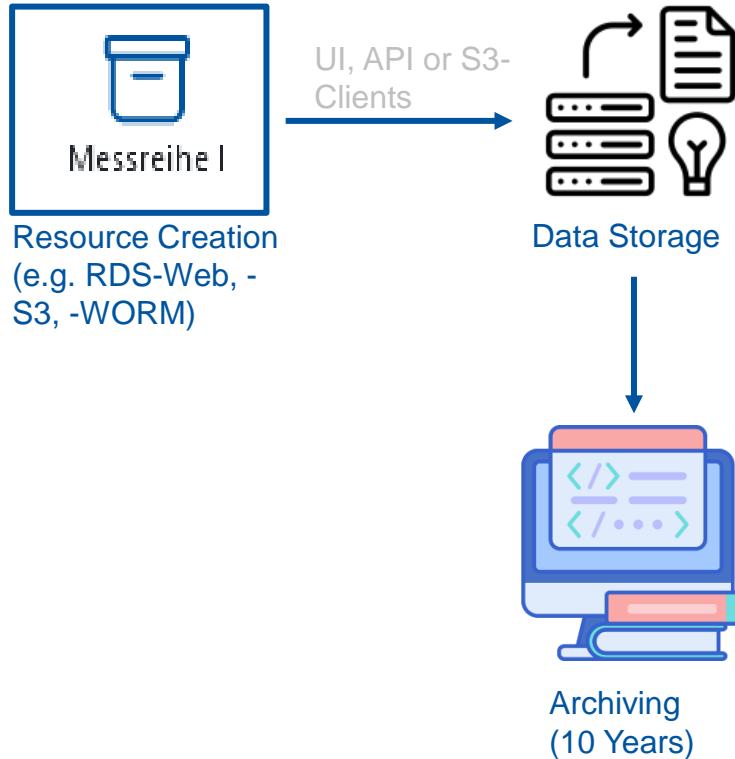
Example for Coscine Usage



Data Storage in Coscine

- **Upload files** via web interface, API or S3 clients
 - Many small files at best as a zip file
 - Large files via S3 clients
 - Write scripts for automated uploading
- **PIDs** for unique identification are assigned automatically
- **Archiving** for ten years after the end of the project
 - Good scientific practice
 - Reusability of the (meta)data
 - Read-only status

Example for Coscine Usage

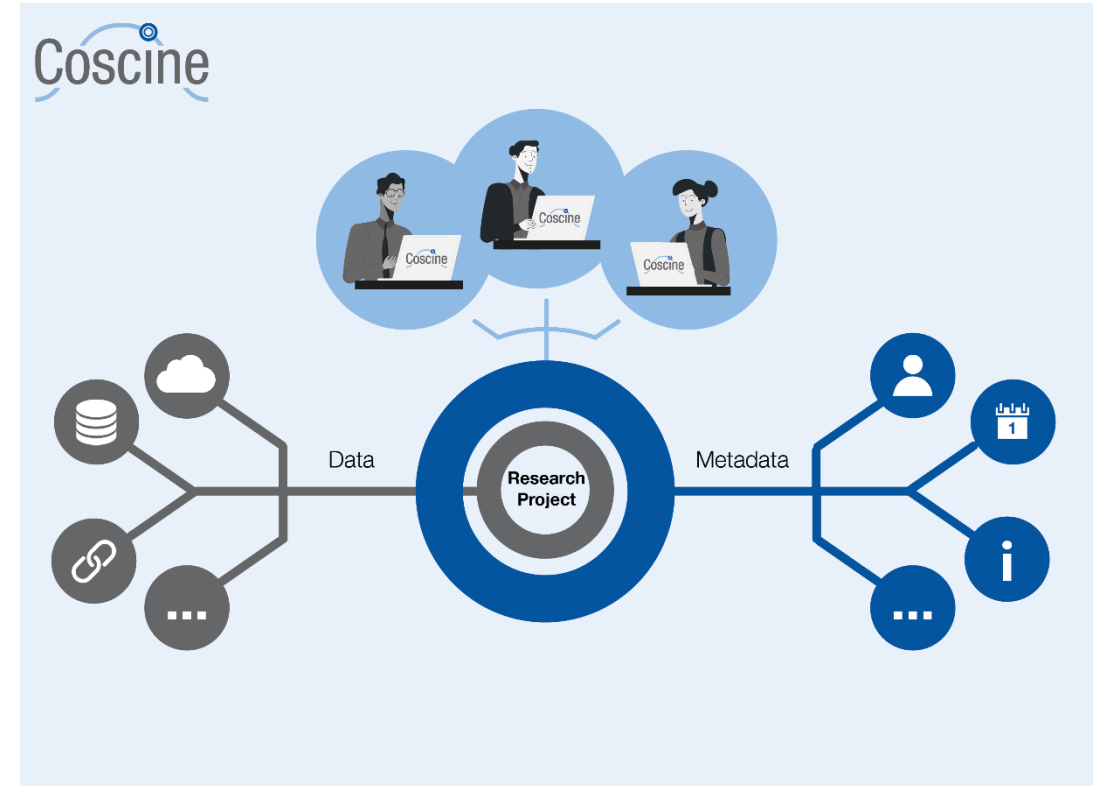


Access for external users/share data in Coscine

- No Publication in Coscine
 - PIDs for reference of projects and resources
 - No DOIs
 - Contact Publication Service
- Invite external persons to projects
 - Role concept (owner, member, guest)
 - Login option via ORCID
 - For S3: sharing of read credentials possible
- Reference to research data stored in Coscine
 - Use Text Proposal

FAIR Principles & Coscine

- **Login Possibilities**
 - Across institutional boundaries
- **Role Management**
 - Owner, Member, Guest
- **Metadata**
 - On Project-, Resource- Data Level
 - Can be shared publicly in Coscine
 - Searchable via ElasticSearch
- **Resources**
 - Different Resource Types
 - Archiving for 10 Years
 - Handle PIDs for identification





Resources & Quota

Quota Applications via JARDS Platform - Background

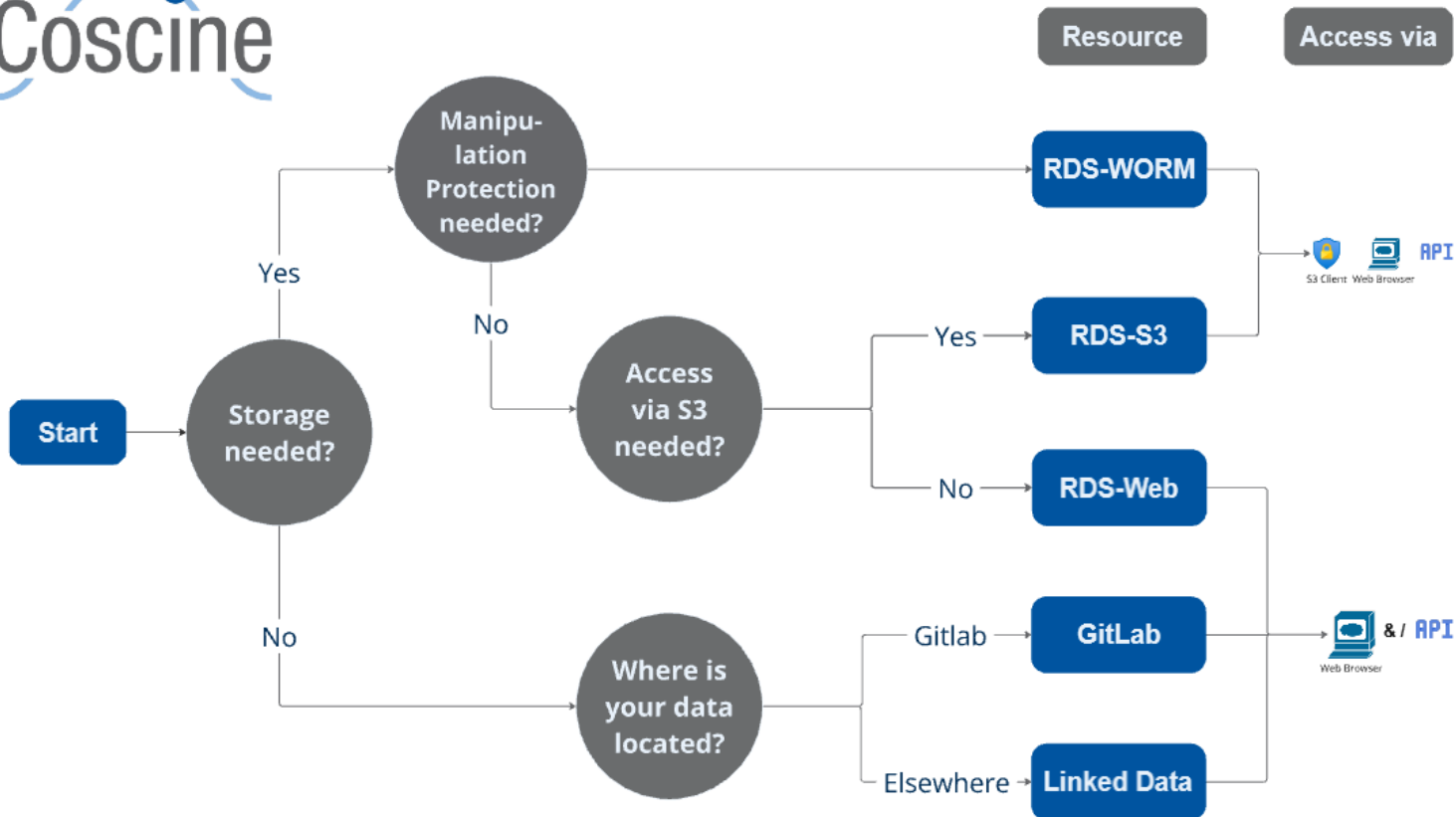
- **Storage*** is funded by Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen
 - Coscine as a platform for Metadata Management as an access point
- Access for **S3** and **WORM** Resources via S3-Clients (e.g. Cyberduck) possible
 - Quota Applications via JARDS Platform needed
 - Questions regarding Workflow, Metadata Management etc. (see next slides)
 - Applications for WORM include more specific questions
- **JARDS Platform** is already well established in HPC context
 - Own Coscine Instance since July 2023

* RDS, RDS.nrw and DataStorage.nrw

Ministerium für
Kultur und Wissenschaft
des Landes Nordrhein-Westfalen



Which Resource Type fits the best? – Flowchart for help





Different Resource Types in Coscine - Documentation

Resource Type	RDS Authorized	External users	Link to Application Form
<i>RDS-Web</i>	<i>25 GB default; up to 100 GB with self service</i>	<i>no storage space</i>	<i>RDS-Web resources</i>
<i>RDS-S3</i>	<i>no default; up to 125TB with RDS application</i>	<i>no storage space</i>	<i>RDS-S3 resources</i>
<i>RDS-WORM</i>	<i>no default; up to 125TB with RDS application</i>	<i>no storage space</i>	<i>RDS-WORM resources</i>
<i>Linked Data</i>	<i>no limit/storage space required</i>	<i>no limit since no storage space required</i>	<i>No application needed</i>
<i>GitLab</i>	<i>No limit/no storage space required</i>	<i>no limit since no storage space required</i>	<i>No application needed</i>

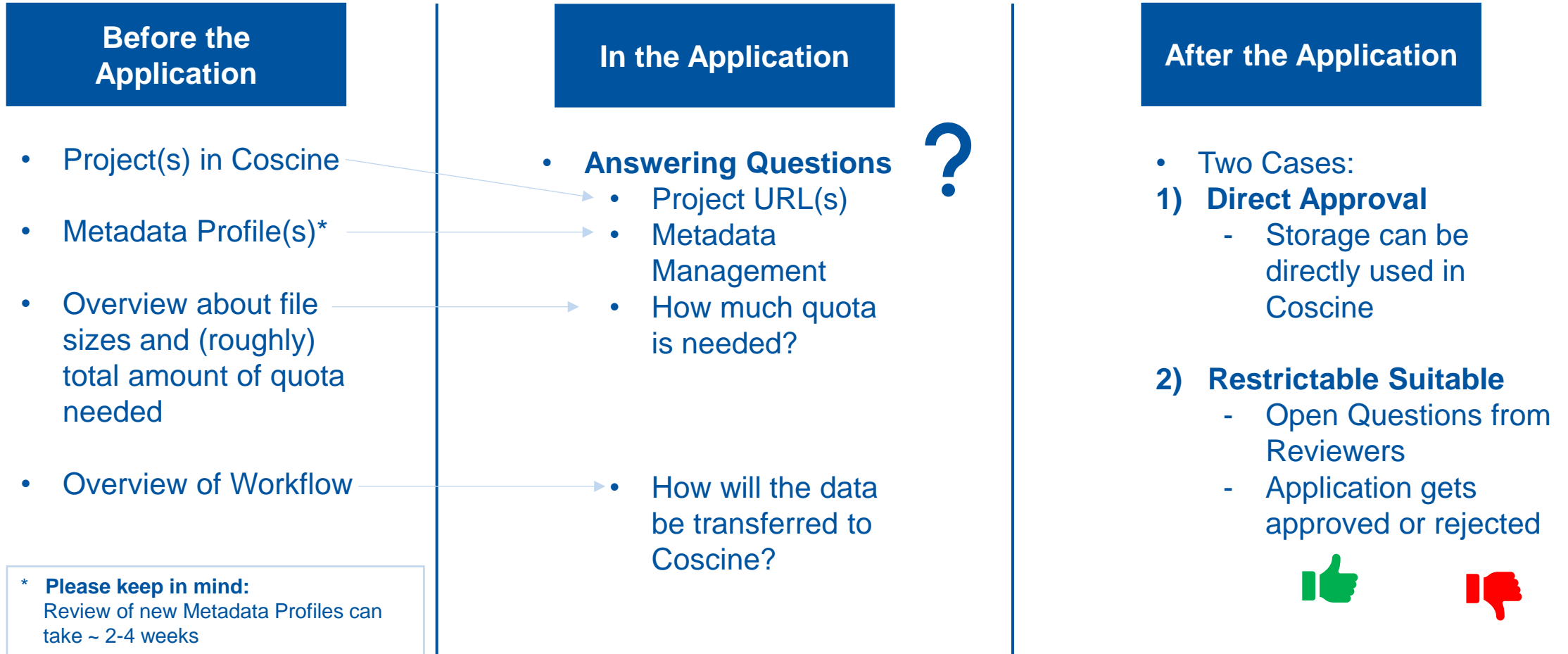


JARDS Platform

- Applications for Quota via **JARDS Platform**
- Questions regarding Metadata Management & Data Workflow etc.
- Further input on the next slides...



Quota Applications via JARDS Platform - Workflow





Quota Applications via JARDS Platform – Outlook



Current Status

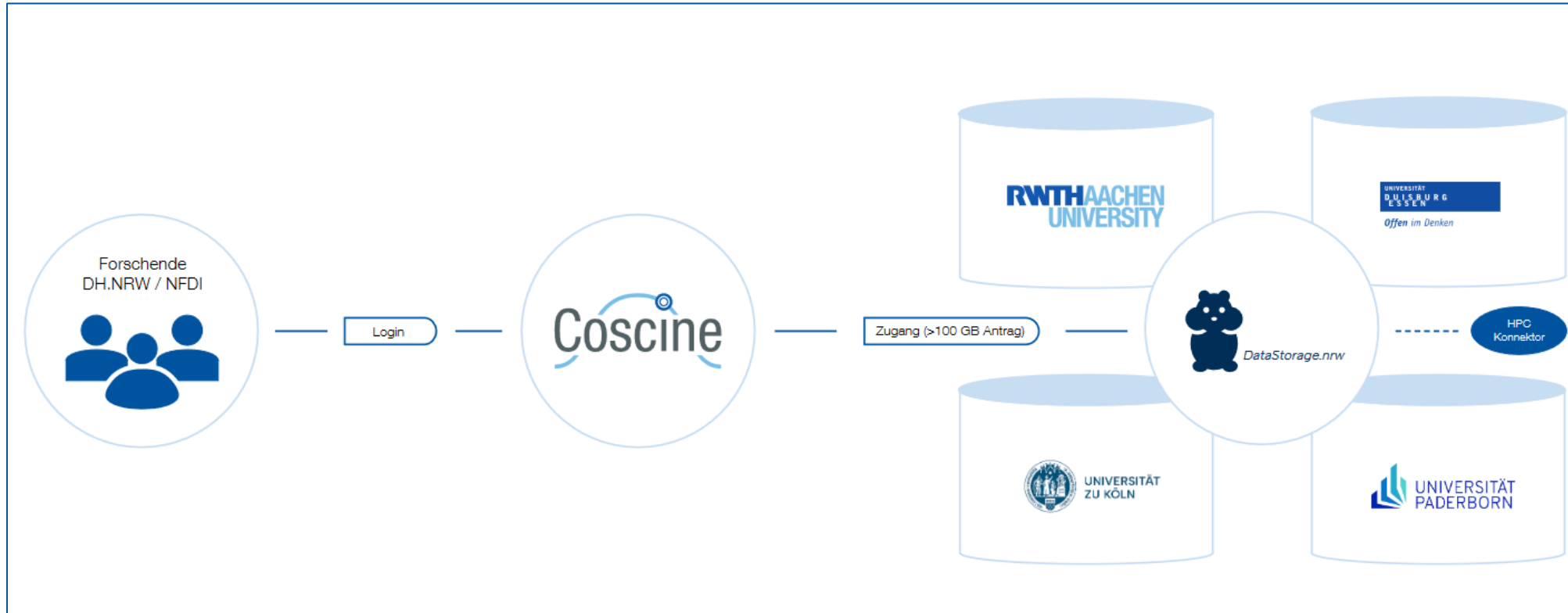
- **Research Data Storage (RDS)**
 - Useable by authorized universities
- **GitLab & LinkedData**
 - Useable by all researchers via ORCID

Outlook

- **DataStorage.nrw**
 - Useable by DH.NRW
 - JARDS Reviews in a „Peer-Review“ System
- **GitLab & LinkedData**
 - Useable by all researchers via ORCID



Quota Applications via JARDS Platform – Outlook



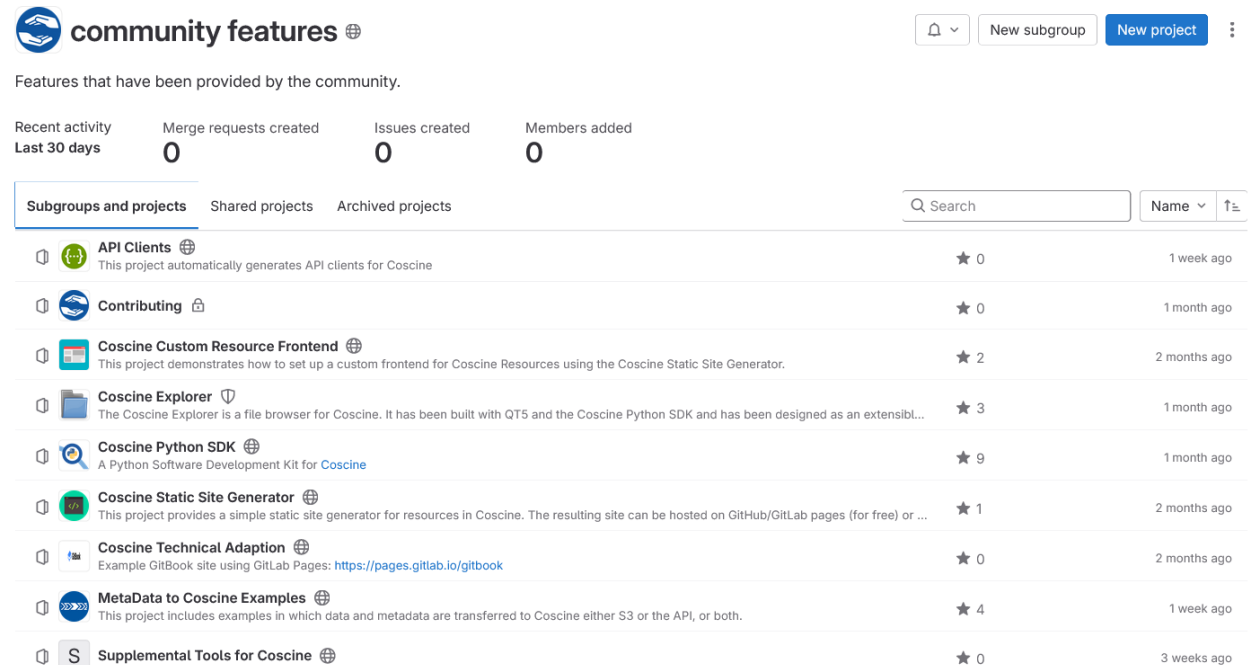


Community

Further Input

- **Community Features**

- Contributions by the Coscine Community
 - Helpful for other users
 - Contributions welcome!
 - E.g. with examples from everyday research in which Coscine was used (scripts, automated upload, direct transfer of metadata from a device, etc.).



The screenshot shows the 'community features' page in the Coscine environment. At the top, there's a header with the 'community features' title and a globe icon. Below the header, there are statistics for recent activity: 'Last 30 days' with '0' merge requests created, '0' issues created, and '0' members added. A navigation bar includes 'Subgroups and projects', 'Shared projects', and 'Archived projects', along with a search bar and a 'Name' dropdown. The main content is a list of projects, each with an icon, title, description, star count, and creation time.

Project Name	Description	Stars	Created
API Clients	This project automatically generates API clients for Coscine	0	1 week ago
Contributing		0	1 month ago
Coscine Custom Resource Frontend	This project demonstrates how to set up a custom frontend for Coscine Resources using the Coscine Static Site Generator.	2	2 months ago
Coscine Explorer	The Coscine Explorer is a file browser for Coscine. It has been built with QT5 and the Coscine Python SDK and has been designed as an extensibl...	3	1 month ago
Coscine Python SDK	A Python Software Development Kit for Coscine	9	1 month ago
Coscine Static Site Generator	This project provides a simple static site generator for resources in Coscine. The resulting site can be hosted on GitHub/GitLab pages (for free) or ...	1	2 months ago
Coscine Technical Adaption	Example GitBook site using GitLab Pages: https://pages.gitlab.io/gitbook	0	2 months ago
MetaData to Coscine Examples	This project includes examples in which data and metadata are transferred to Coscine either S3 or the API, or both.	4	1 week ago
Supplemental Tools for Coscine		0	3 weeks ago

Upcoming Events

- **Workshops for Coscine at RWTH Aachen**
 - Basic-Workshop (Tomorrow at 10-12) | Online
 - Workshop regarding Metadata(profiles) (20.11.24 at 10-12) | Online
 - Workshop regarding automatised processes (27.11.24 at 10-13) | in Person

Thank you for your attention!



Coscine Webseite



Coscine Dokumentation



Coscine Newsletter



Questions/Feedback at: servicedesk@rwth-aachen.de



Questions & Comments