



– FAIR play integrated right from the start

NFDI4Ing Conference 2022

26. Oktober

Dr. Ilona Lang und Marcel Nellesen (RWTH Aachen University)

lang@itc.rwth-aachen.de

nellesen@itc.rwth-aachen.de

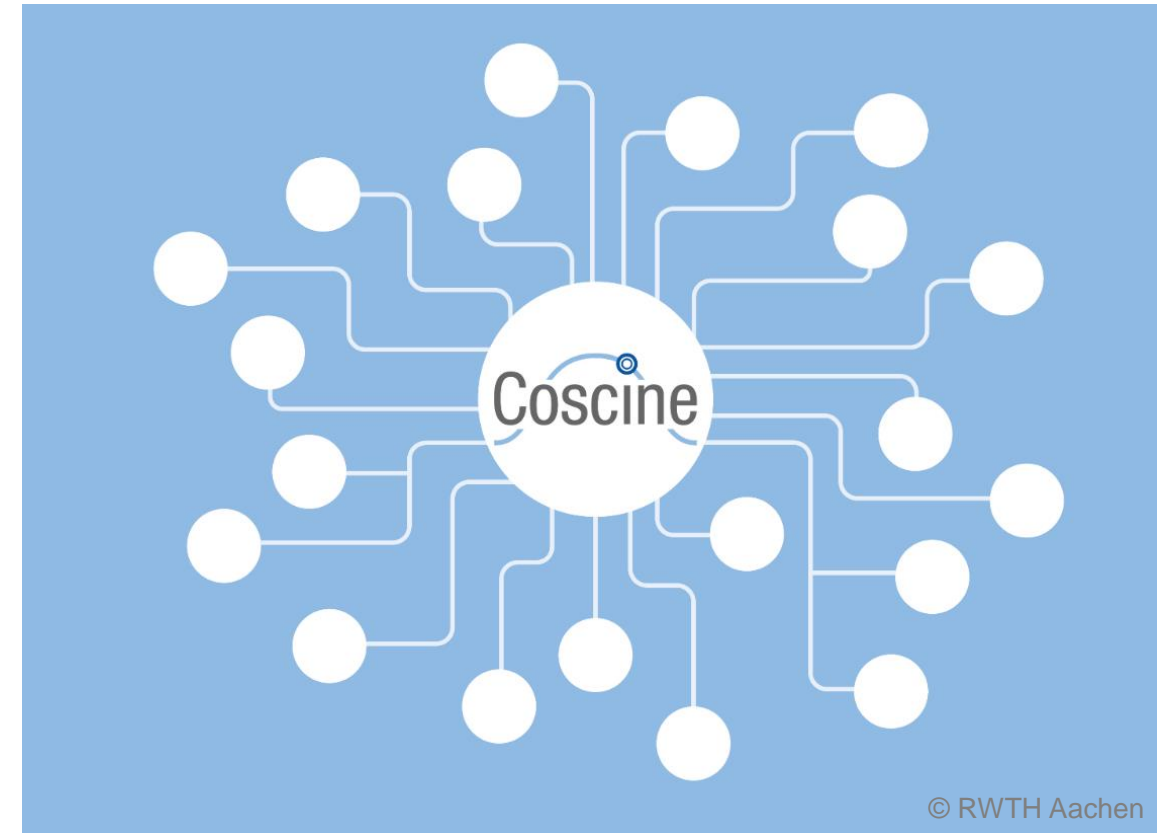


The slide set as a whole and its texts are licensed as a whole under [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/). Individual elements such as illustrations may be licensed independently.



Coscine – Overview

- Our Motivation
- Background Information
- Demonstration
- The FAIR Principles and Coscine
- Automated Workflows
- Contact & Further Information



Coscine – FAIR Play: Our Motivation

Problem:

- FAIR handling of research data is **not part** of a researcher's **daily routine**
 - If at all, only a small part of a project's research data is treated according to the FAIR principles when it is published
 - A significant amount of (metadata) information is lost

Our Solution:

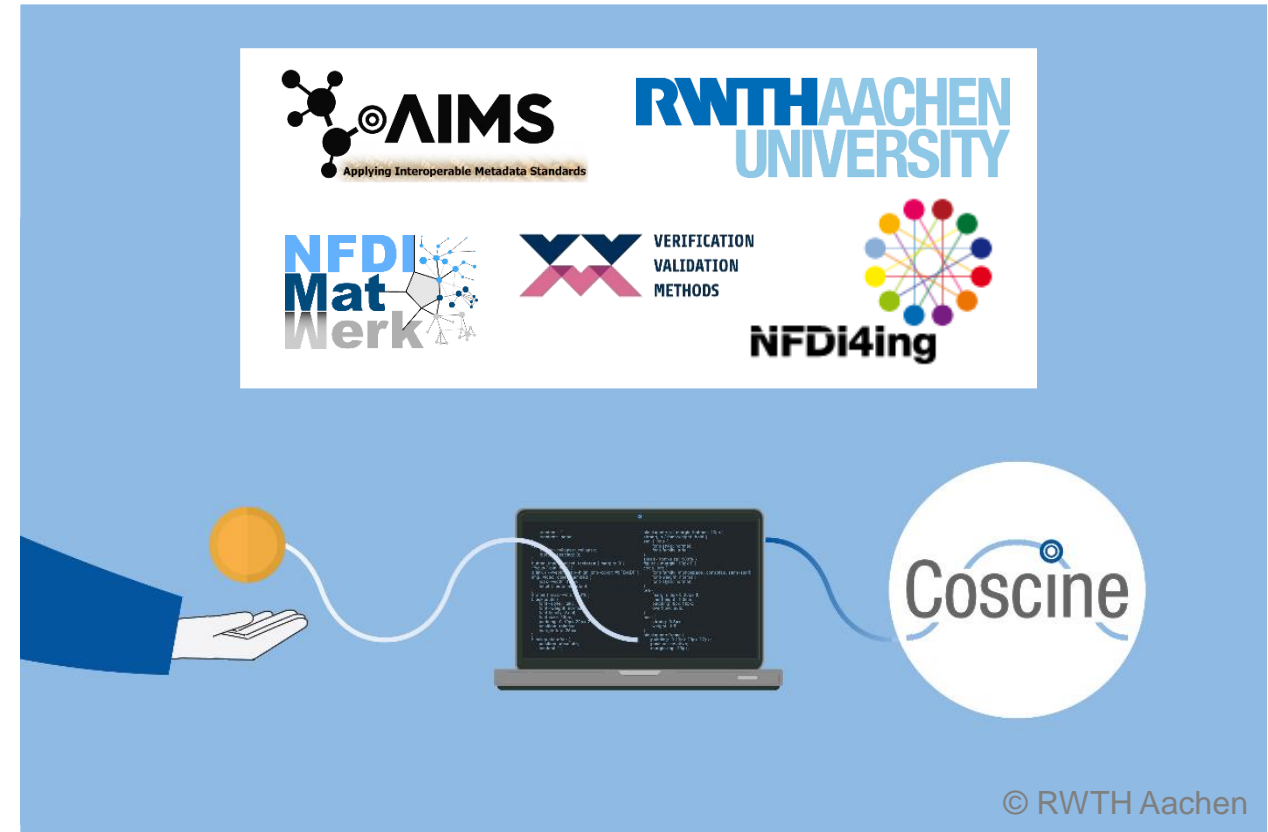
- Coscine - a data storage and linking environment that **implicitly implements** FAIR principles



[Luvmybry / Pixabay License](#)

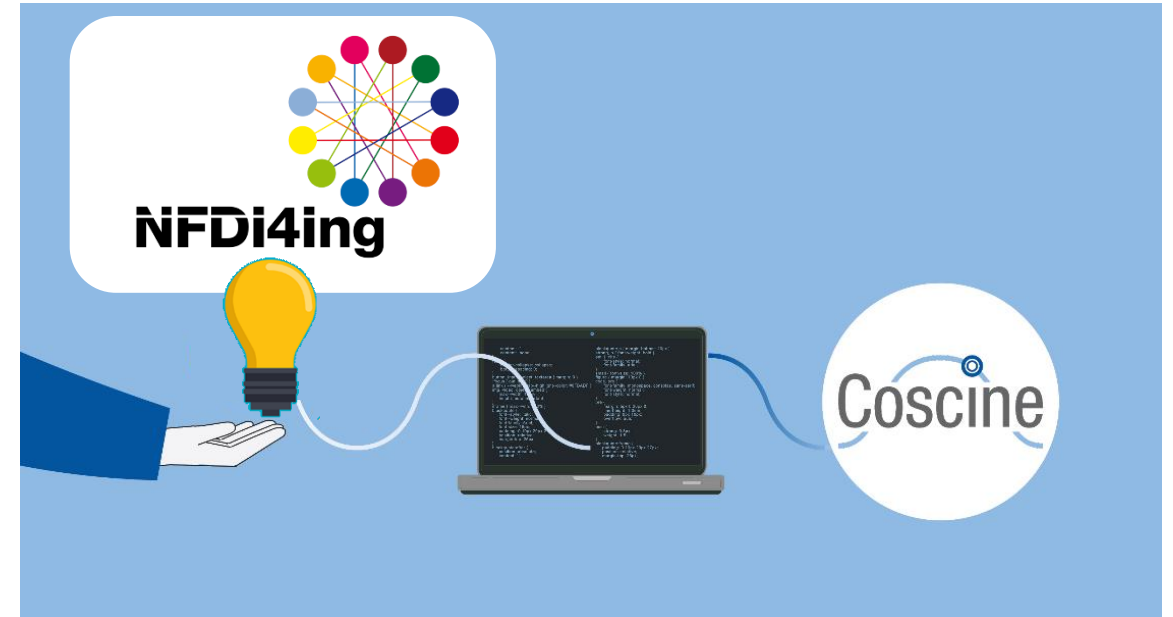
Coscine – Funding & Development

- Developed at RWTH Aachen University
 - Development using the Agile Scrum Method
 - **Open Source Platform**
- Funding based on
 - Excellent University funding for RWTH
 - NFDI-MatWerk
 - **NFDI4Ing**
 - Project “Anwendung Interoperabler Metadatenstandards (AIMS)”
 - Project “VVMethoden”

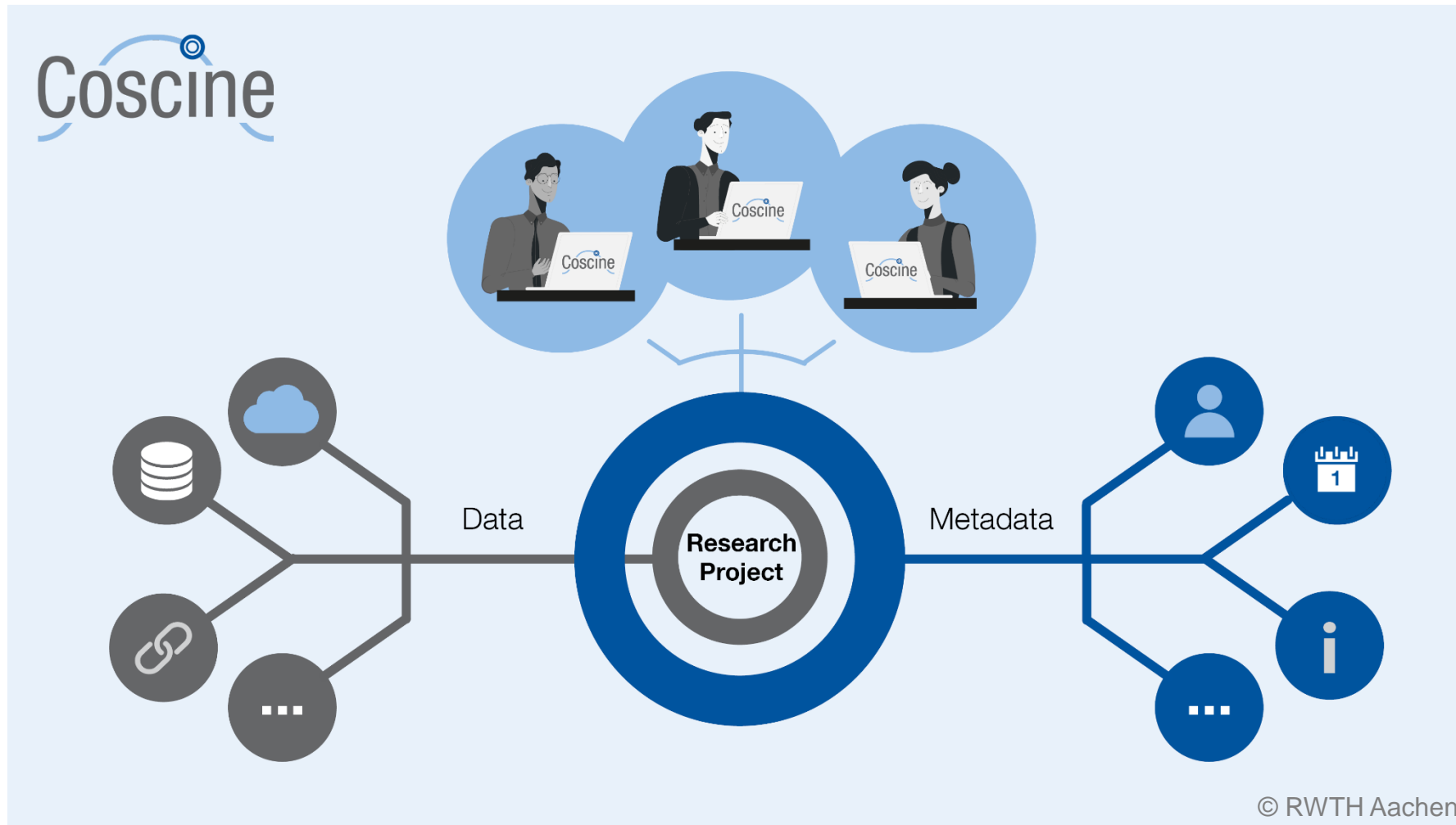


Coscine & NFDI4Ing

- Coscine development is aligned with **Task Area Base Services** (S4: storage & repositories)
 - Task S-4-2: Development of software for federated storage services
 - Task S-4-3: Development of a cost and distribution model for storage
- Projects with NFDI4Ing relation **can request storage space** on the Research Data Storage (NRW-RDS-Resources)



Coscine – Short Introduction



Project structure

Simple data management

Storage space

Access free storage space

Collaboration

Access for all project members

Metadata

Automatically linked to data

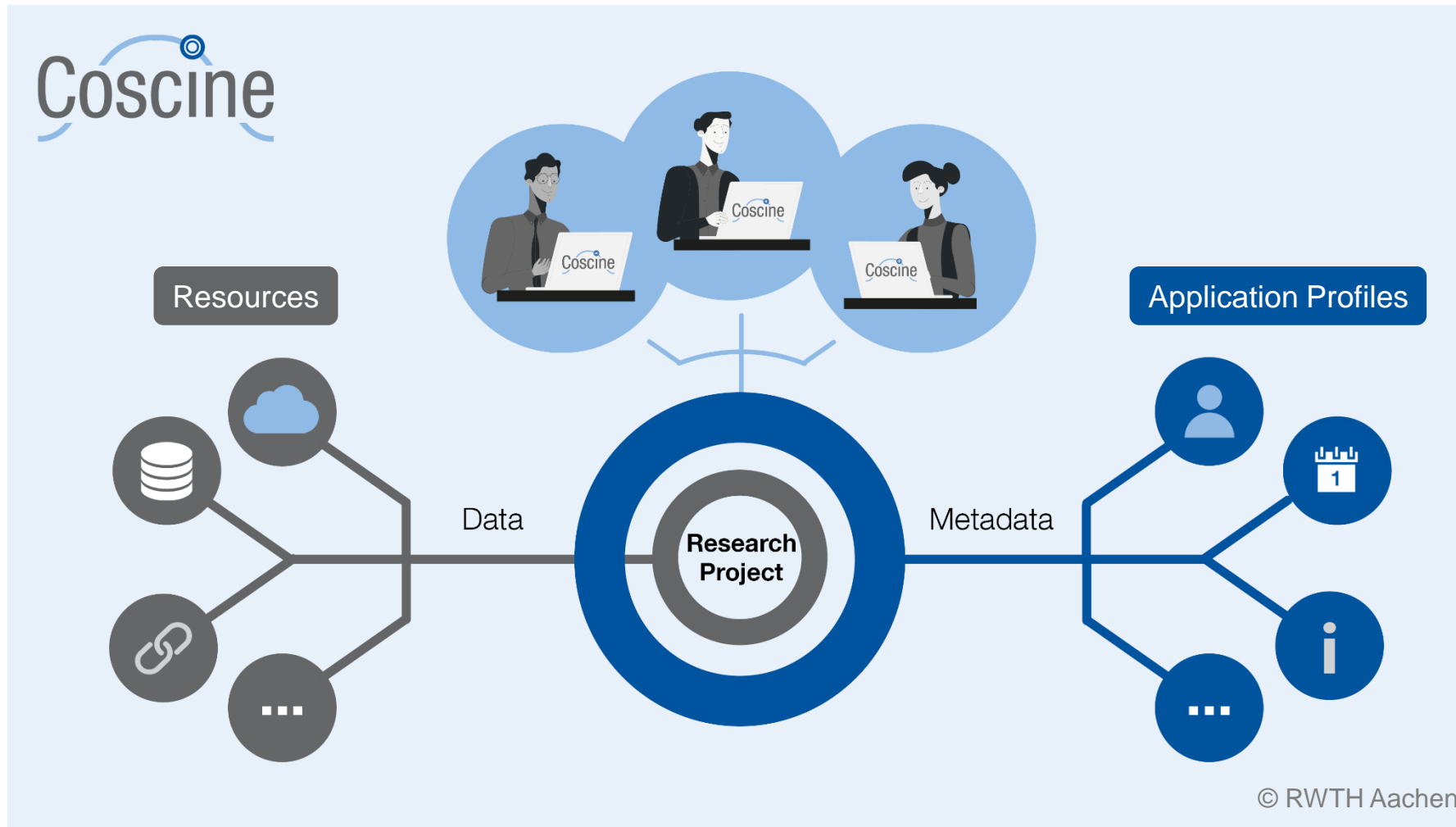
Individuality

Project-specific application profiles

Archiving

Archive data in place

Coscine – Important Terms



Resources

Data sources / Integrated IT services

Application Profiles

Individual metadata combination from different metadata schemas

Coscine – Demonstration



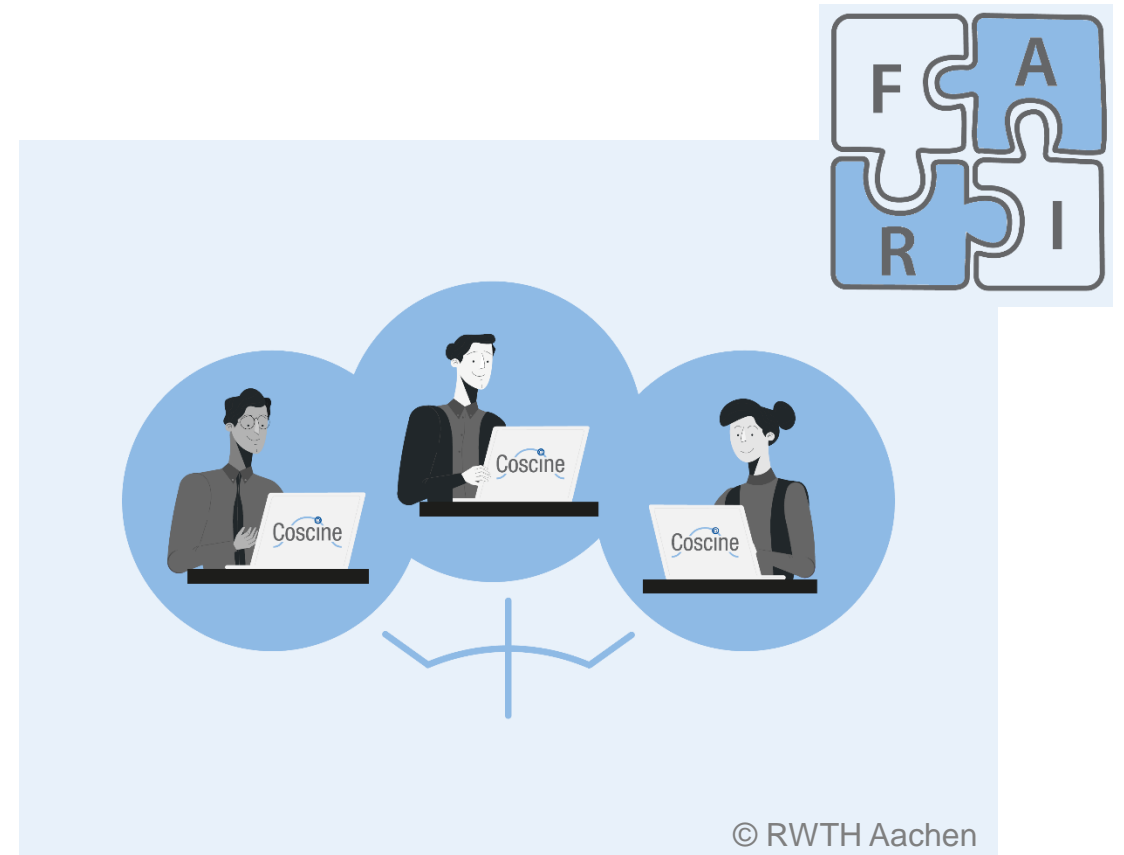
Coscine – What makes us FAIR?



Coscine – Login

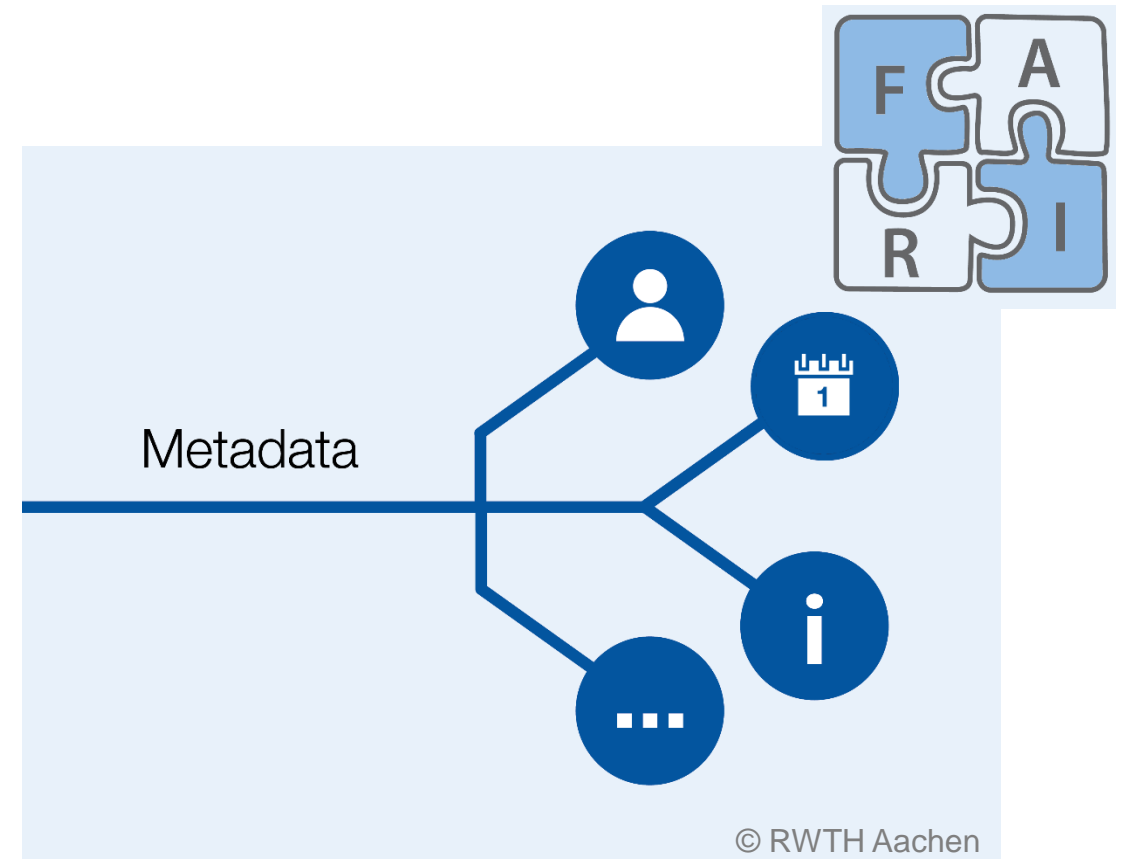
Across institutional borders via...

- ORCID
 - Single Sign-On
- Makes research data **accessible** and **re-useable**



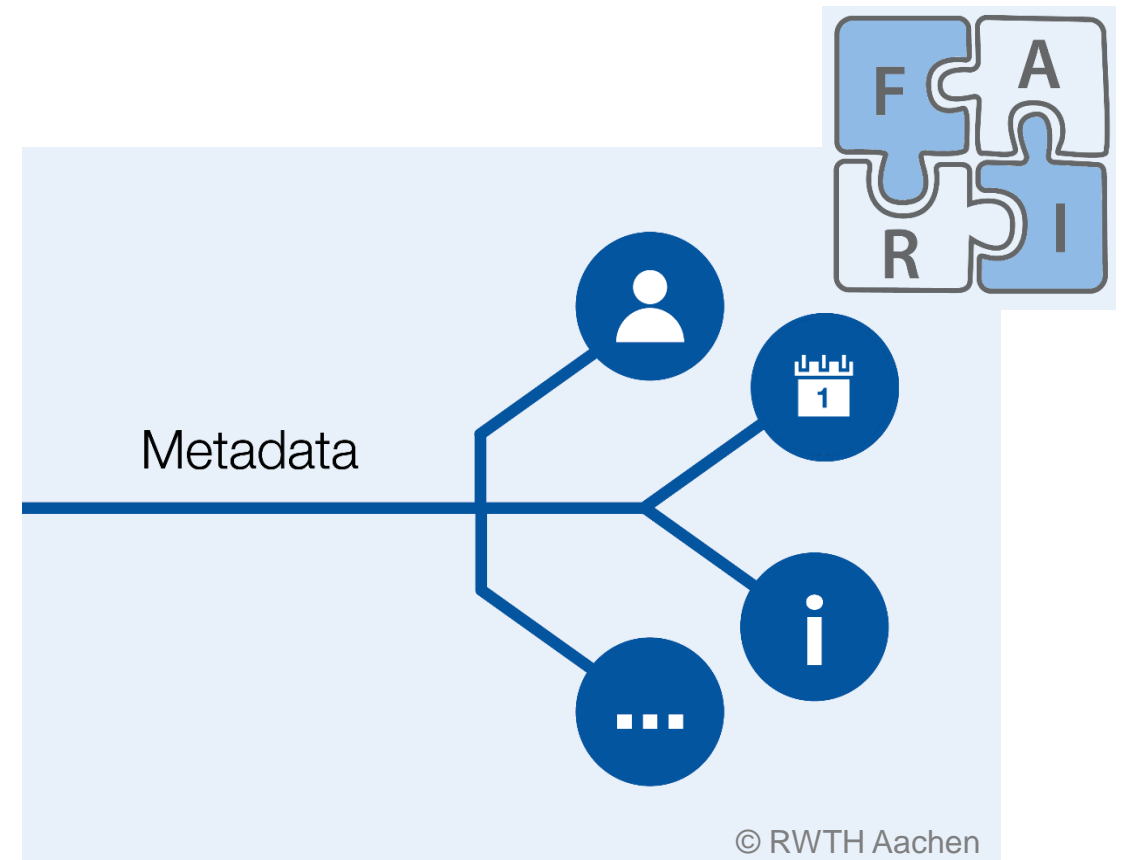
Coscine – Metadata

- Collected on **project** and **resource level**
 - **Automatically** linked with research data
 - Can be publicly **shared** inside Coscine
- Makes research data **findable** and **interoperable**



Coscine – Metadata

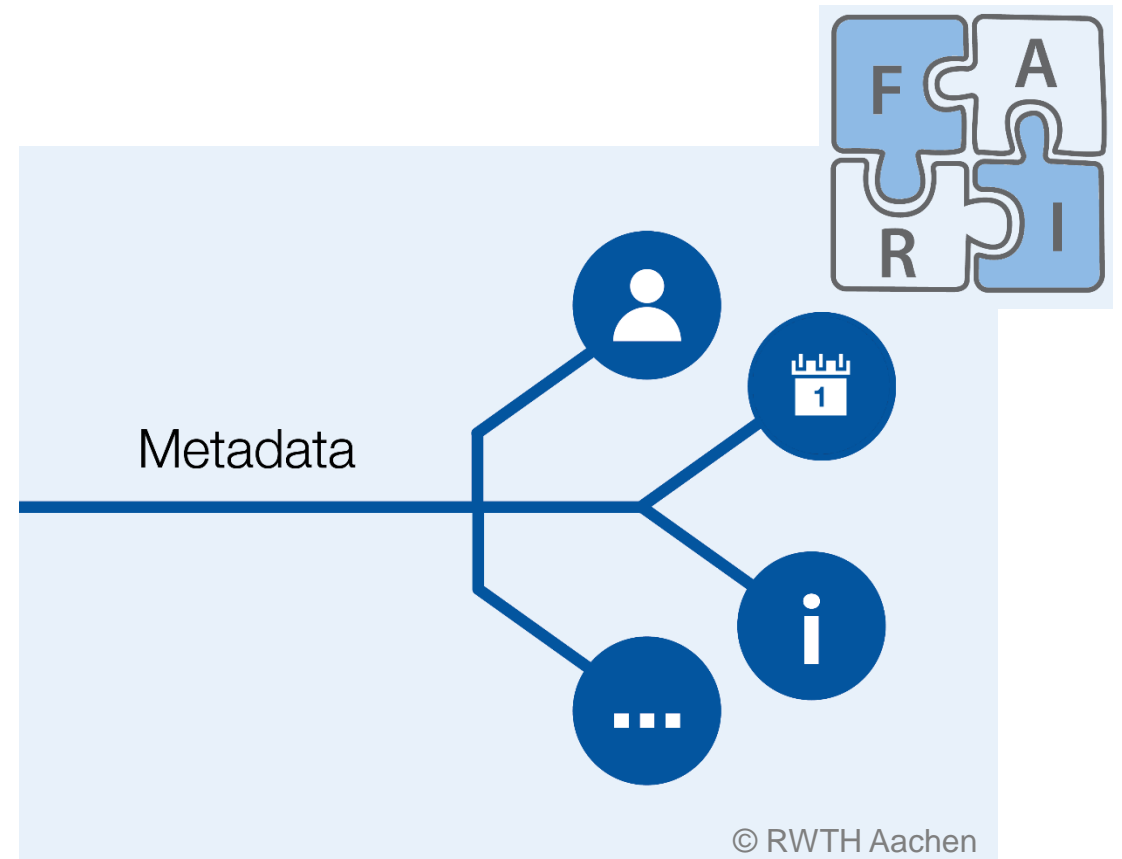
- **Searchable** via ElasticSearch
 - Technical representation and validation via W3C standards RDF and SHACL
 - Planned: **connection to the NFDI4Ing metadata hub** via "FAIR Digital Object" interfaces
- Makes research data **findable** and **interoperable**



Coscine – Application Profile Generator (AIMS*)

- Enables creation of profiles with **individual** and **discipline-specific** metadata
- No technical knowledge needed (RDF and SHACL)
- Makes research data **findable** and **interoperable**

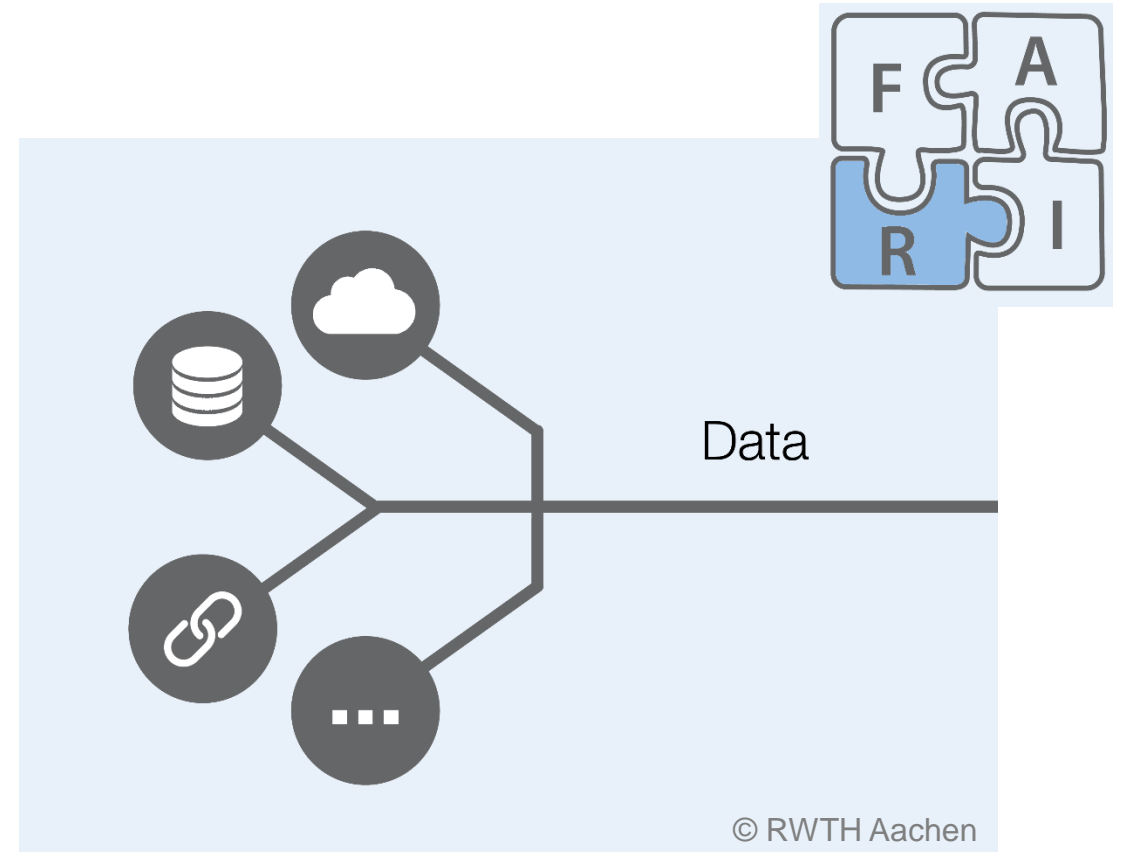
***Project AIMS:** Applying Interoperable Metadata Standards



Coscine – Resources

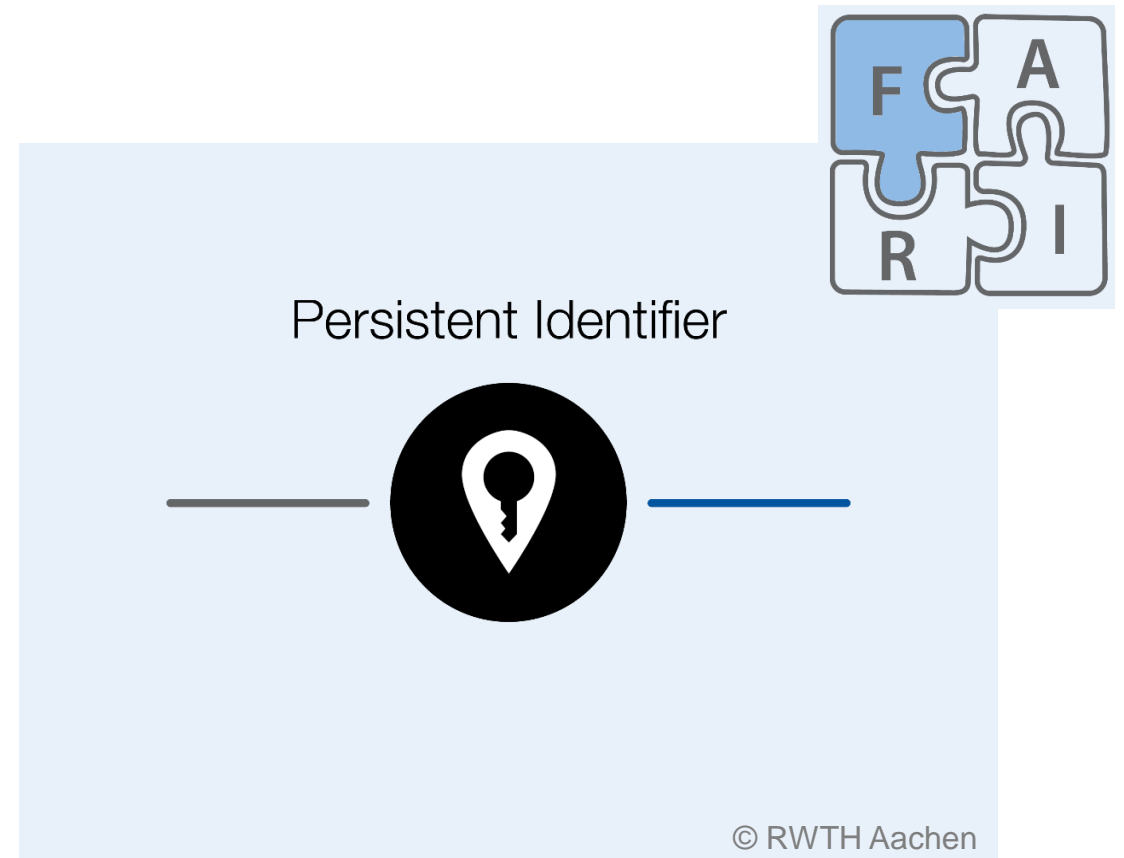
- Resource Types
 - Research Data Storage (RDS)
 - Provide access to the consortial object storage system (Web, S3, WORM)
 - Retention and archiving period of **10 years**
 - Linked Data
 - Externally stored research data can be linked
 - [...]

➤ Makes research data **reusable**



Coscine – Persistent Identifier

- Handle-based ePIC PID
 - Assigned to each **resource**
 - Fragment identifiers for individual files by extended handle URL
 - Location of resources and files uniquely and **permanently identifiable** and **referencable** on a global level
- Makes research data **findable**

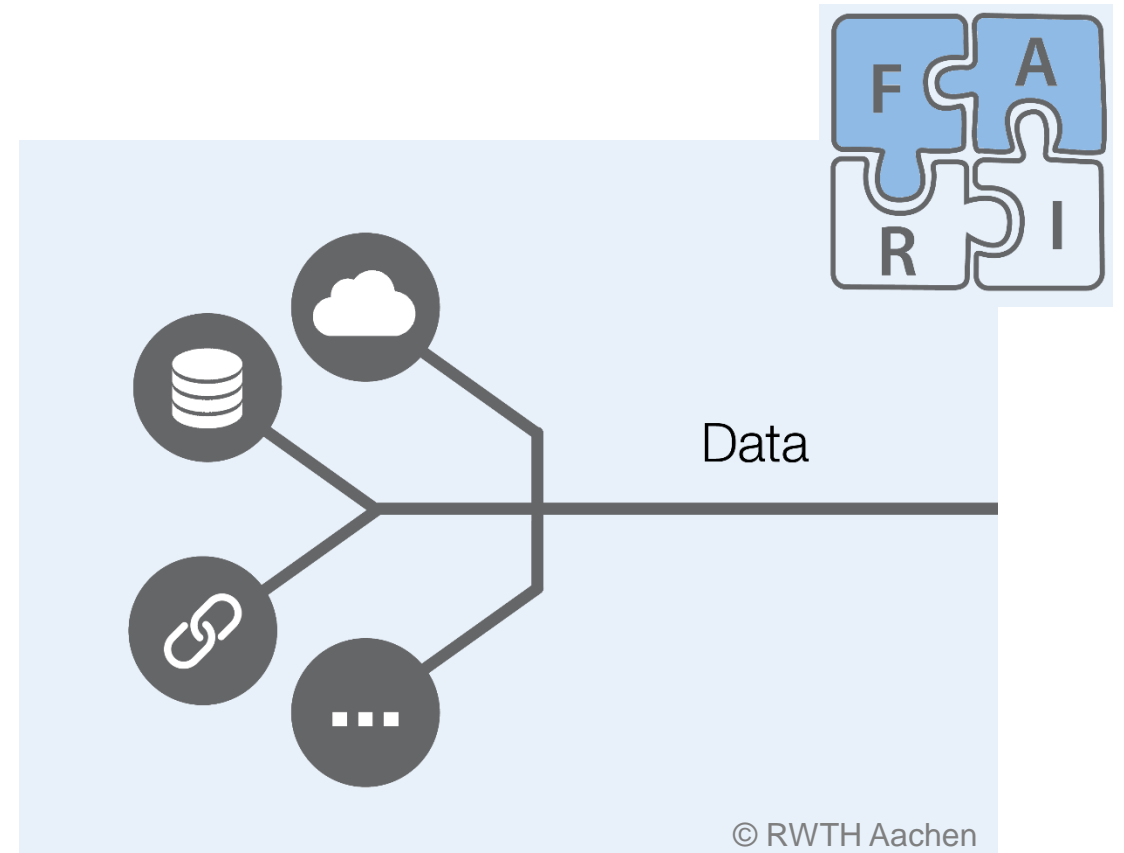


Coscine – RDS-Resources

Interaction via

- Browser or S3 protocol
- REST-API
 - Data and metadata are independently findable and accessible
 - Easy entry of data and metadata into the system and facilitation of subsequent use
 - Enables to **automate** workflows

➤ Makes research data **findable** and **accessible**



Coscine – How to include existing research data?

- Analyze the current situation:
 - What is stored where by whom?
 - Which metadata is already available?



[Pexels](#) / [Pixabay License](#)

Coscine – Metadata Management

- General description of the required metadata fields
 - Expected data types
 - Length of fields
 - Ranges of input fields
- Usage of vocabularies
 - Closed list of possible values (e.g., DFG categories)
- Multitude of input fields
 - For example, each paper must have at least one author but can have multiple authors



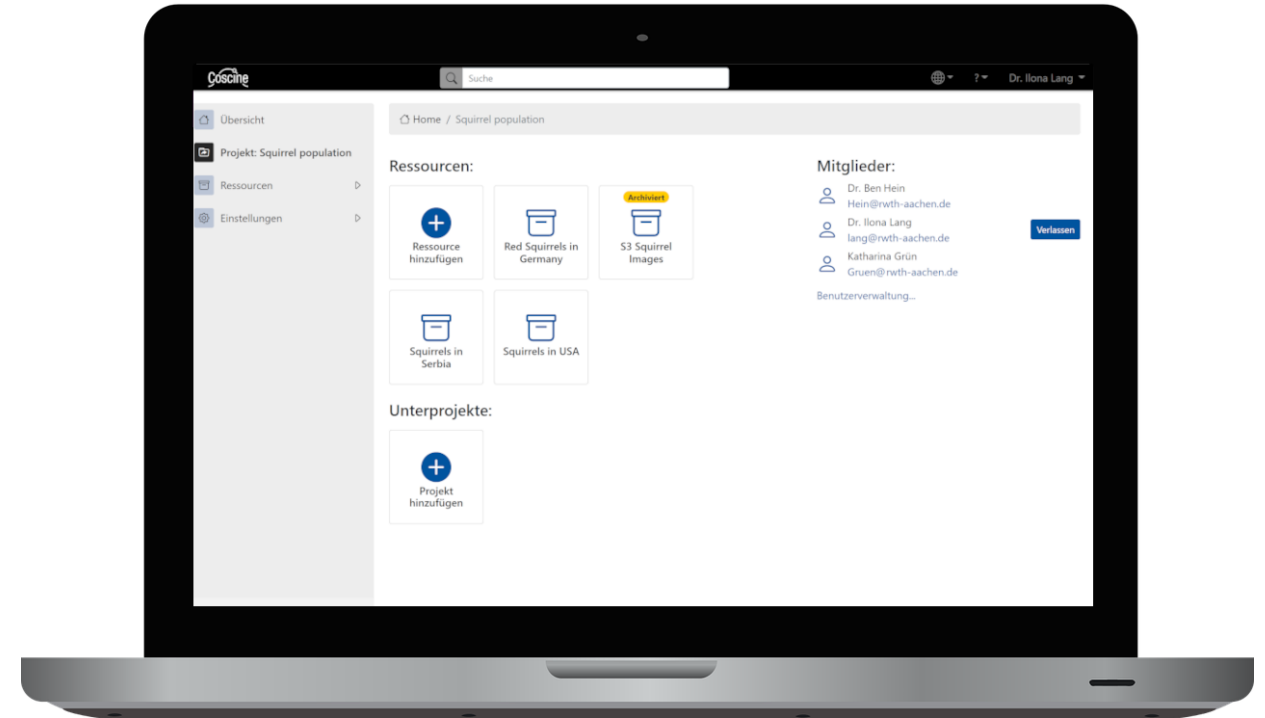
The image shows a magnifying glass and a blue pen resting on a document with a table of data. The table has columns for 'Date' and 'Displays'. The data is as follows:

Date	Displays
04.12.2007	3 093
05.12.2007	3 705
06.12.2007	3 593
07.12.2007	3 729
08.12.2007	3 748
09.12.2007	3 927
10.12.2007	3 836
11.12.2007	7 179
12.12.2007	6 930
13.12.2007	12 7
14.12.2007	8
15.12.2007	22 964
16.12.2007	1 274
17.12.2007	849
18.12.2007	2
19.12.2007	16
20.12.2007	9 2
21.12.2007	16 757
22.12.2007	21 313
23.12.2007	5 962
24.12.2007	22 299
25.12.2007	16 268
26.12.2007	9 299
27.12.2007	6 757
28.12.2007	18
29.12.2007	205
30.12.2007	209

[ds_30](#) / Pixabay License

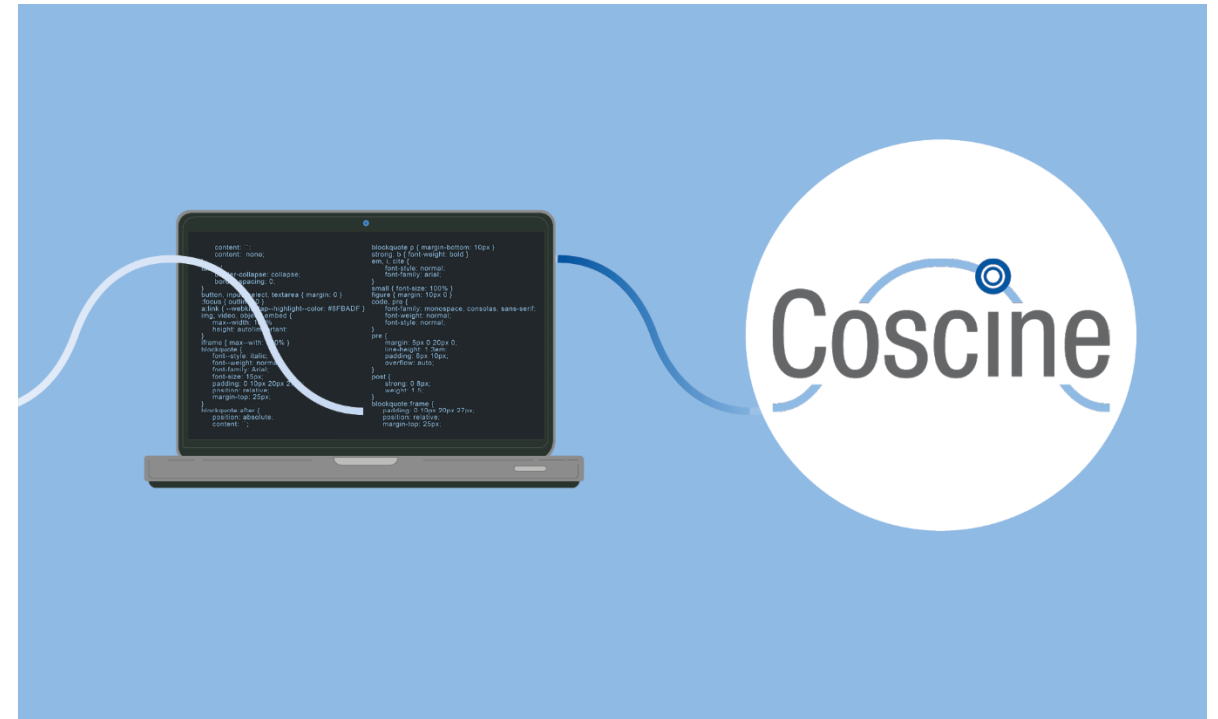
Coscine – Upload Research Data (Part 1)

- Create a project
- Invite members
- Add a resource
 - Select resource type (e.g. RDS-S3, -Web)
 - Define storage size
 - Chose application profile



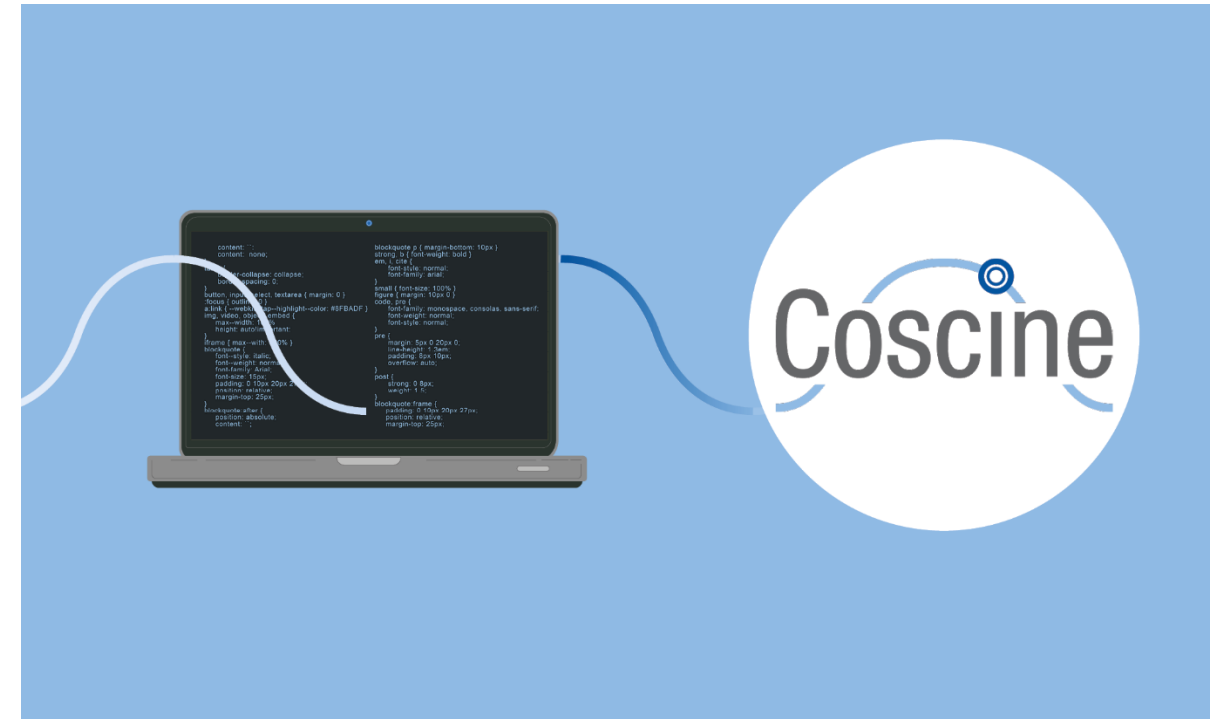
Coscine – Upload Research Data (Part 2)

- Upload the data
 - Example programs for working with S3
 - minIO
 - rclone
- Add metadata
 - Write/use a parser to process the locally stored metadata
 - Upload the metadata through the API



Coscine – How to use the API

- Read the available documentation
- Create a token for authorization
- All functions can easily be executed and / or tested through a web application
 - Extensive documentation
 - Examples
 - Code snippets
- Create a script with the snippets and the token

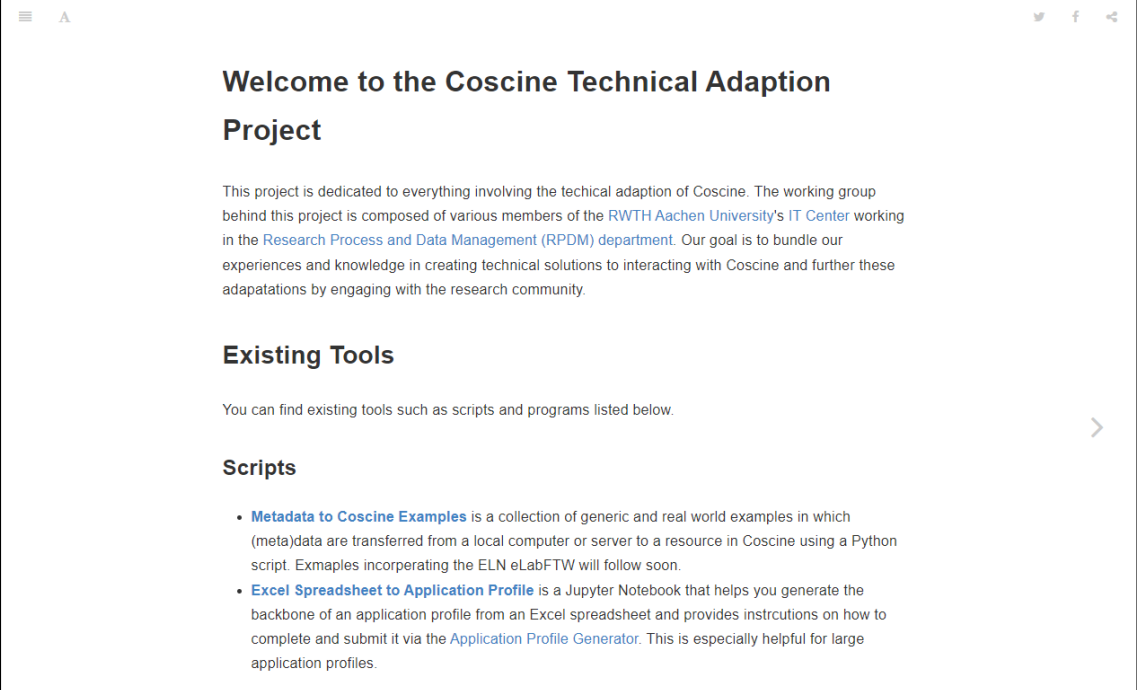


Coscine – API Demonstration



Coscine Technical Adaptation

- Special Group with the aim to support researcher in starting with Coscine
- Regular exchange with users
- Providing of
 - Scripts
 - Programs
 - Best Practices
 - Consultations



The screenshot shows a webpage with the following content:

Welcome to the Coscine Technical Adaption Project

This project is dedicated to everything involving the technical adaption of Coscine. The working group behind this project is composed of various members of the RWTH Aachen University's IT Center working in the [Research Process and Data Management \(RPDM\) department](#). Our goal is to bundle our experiences and knowledge in creating technical solutions to interacting with Coscine and further these adaptations by engaging with the research community.

Existing Tools

You can find existing tools such as scripts and programs listed below.

Scripts

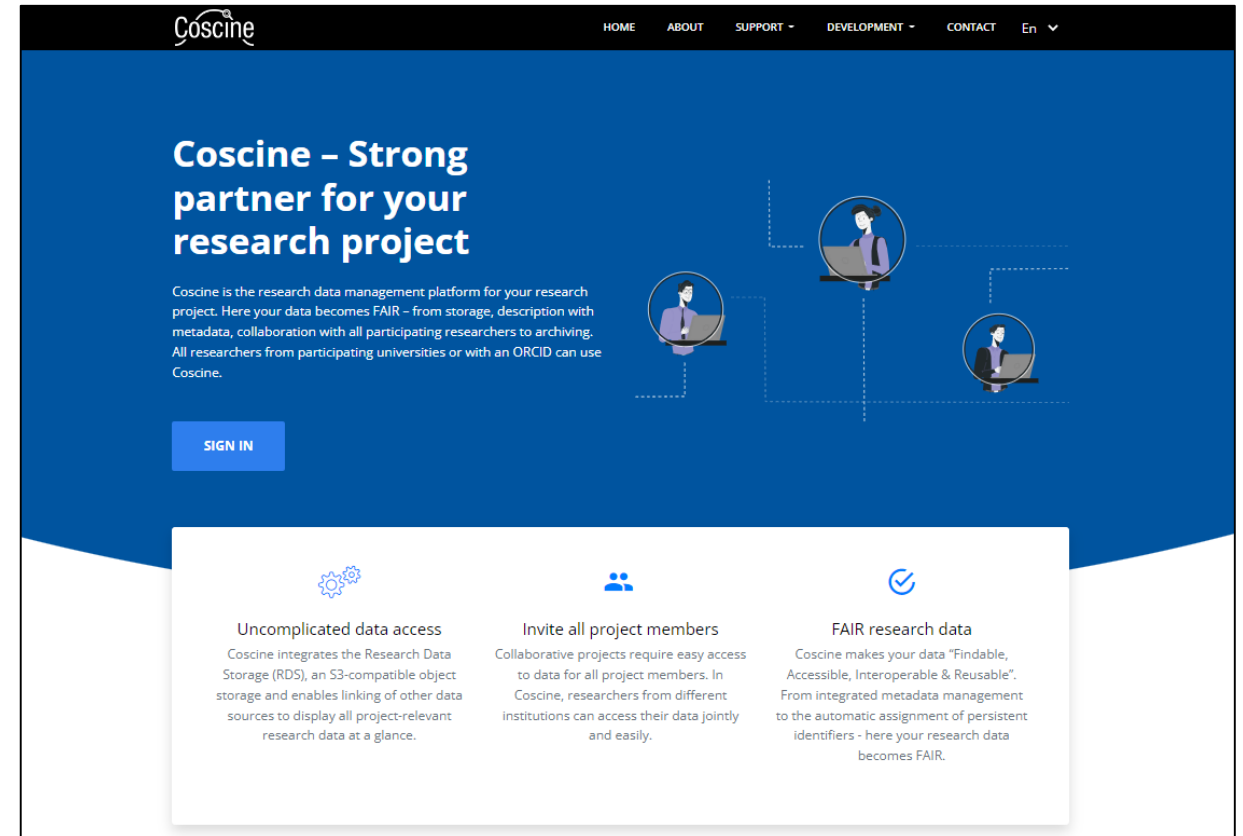
- [Metadata to Coscine Examples](#) is a collection of generic and real world examples in which (meta)data are transferred from a local computer or server to a resource in Coscine using a Python script. Exmaples incorporating the ELN eLabFTW will follow soon.
- [Excel Spreadsheet to Application Profile](#) is a Jupyter Notebook that helps you generate the backbone of an application profile from an Excel spreadsheet and provides instructions on how to complete and submit it via the [Application Profile Generator](#). This is especially helpful for large application profiles.

Ok so...



Coscine – Further Information

- Homepage
 - <https://www.coscine.de/>
- Documentation
 - <https://docs.coscine.de/>
- Mailing List
 - <https://lists.rwth-aachen.de/postorius/lists/coscine.lists.rwth-aachen.de/>



Coscine HOME ABOUT SUPPORT DEVELOPMENT CONTACT En

Coscine – Strong partner for your research project

Coscine is the research data management platform for your research project. Here your data becomes FAIR – from storage, description with metadata, collaboration with all participating researchers to archiving. All researchers from participating universities or with an ORCID can use Coscine.

[SIGN IN](#)

- Uncomplicated data access**
Coscine integrates the Research Data Storage (RDS), an S3-compatible object storage and enables linking of other data sources to display all project-relevant research data at a glance.
- Invite all project members**
Collaborative projects require easy access to data for all project members. In Coscine, researchers from different institutions can access their data jointly and easily.
- FAIR research data**
Coscine makes your data "Findable, Accessible, Interoperable & Reusable". From integrated metadata management to the automatic assignment of persistent identifiers - here your research data becomes FAIR.

Thank You & Stay FAIR!



[Luvmybry](#) / [Pixabay License](#)