



FAIR DATA
AUSTRIA

Database Repository

1st in person meeting

June 10, 2021

Motivation - FAIR DB Repository

Common disadvantages

Databases are set up locally at research units

Database **admin skills** required

Our vision

Private cloud hosted repository

Database management outsourced to repository infrastructure

Motivation - FAIR DB Repository

Common disadvantages

Databases are set up locally at research units

Database **admin skills** required

Local usage does not motivate to provide metadata

Lack of **metadata** yield to unusable DB dumps after end of a project

Our vision

Private cloud hosted repository

Database management outsourced to repository infrastructure

Collect metadata in order to support **FAIRness**

Motivation - FAIR DB Repository

Common disadvantages

Databases are set up locally at research units

Database **admin skills** required

Local usage does not motivate to provide metadata

Lack of **metadata** yield to unusable DB dumps after end of a project

Unversioned data causes lack of reproducibility

Our vision

Private cloud hosted repository

Database management outsourced to repository infrastructure

Collect metadata in order to support **FAIRness**

Data **versioning** with timestamps guarantees reproducibility

Motivation - FAIR DB Repository

Common disadvantages

Databases are set up locally at research units

Database **admin skills** required

Local usage does not motivate to provide metadata

Lack of **metadata** yield to unusable DB dumps after end of a project

Unversioned data causes lack of reproducibility

Our vision

Private cloud hosted repository

Database management outsourced to repository infrastructure

Collect metadata in order to support **FAIRness**

Data **versioning** with timestamps guarantees reproducibility

Data is **cite-able**

System Requirements

Scalability

for a rapidly
growing number
of users

Flexibility

support different
database
versions and
various engines

Dynamic

data upload via
CSV, REST
interface (e.g.
sensors)

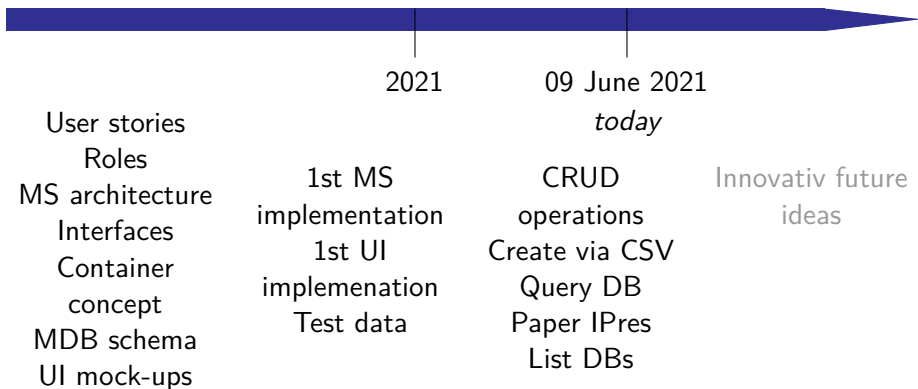
Usability

support different
levels of
SQL-knowledge

Timeline

2020

2022



Status (regarding requirement specification) I

5.1 Create database done todo in progress

5.1.1 **Novice user** can simple create db via CSV/Excel/DB-dump

5.1.2 **Apprentice users and experts** able to create via SQL commands

5.1.3 Different RDBMS engines

5.1.4 Table creation via UI

5.2 Feed database some done lots of todos

5.2.1 Only authorized persons can access dbs

5.2.2 Graphical interface to upload files to the system schema violations

5.2.3 Automate data insertion via REST interface

5.2.4 Insert into via SQL interface

5.2.5 embargo period

5.2.6 Add data description and metadata

Status (regarding requirement specification) II

5.3 Update database todo

5.3.1 Authorized persons can update dbs

5.3.2 Machineable updates

5.4 Query database done todo in progress

5.4.1 Create queries via SQL interface

5.4.2 Create queries via faceted browsing

5.4.3 Store queries and respective metadata (reproducibility, PID)

5.4.4 Obtain result sets as CSV / Excel / JSON

5.5 Find database in progress

5.4.1 Search databases via web interface (using metadata)

5.4.2 Find specific versions of data

Status (regarding requirement specification) III

5.6 Download data done todo

5.7 Manage database future

5.8 Monitoring database future

5.9 Merge databases future

5.10 Extensions future