

# ARC-Finder

A simple, locally-deployed tool to find your peer's research data

**Dominik Brilhaus**

Projektmodul ZK-FDM

10. Juni 2022



# Motivation

- FAIR research data management
- Share and collaborate on published **and unpublished** data
- Avoid (metadata) redundancy



# CEPLAS collaborates with DataPLANT



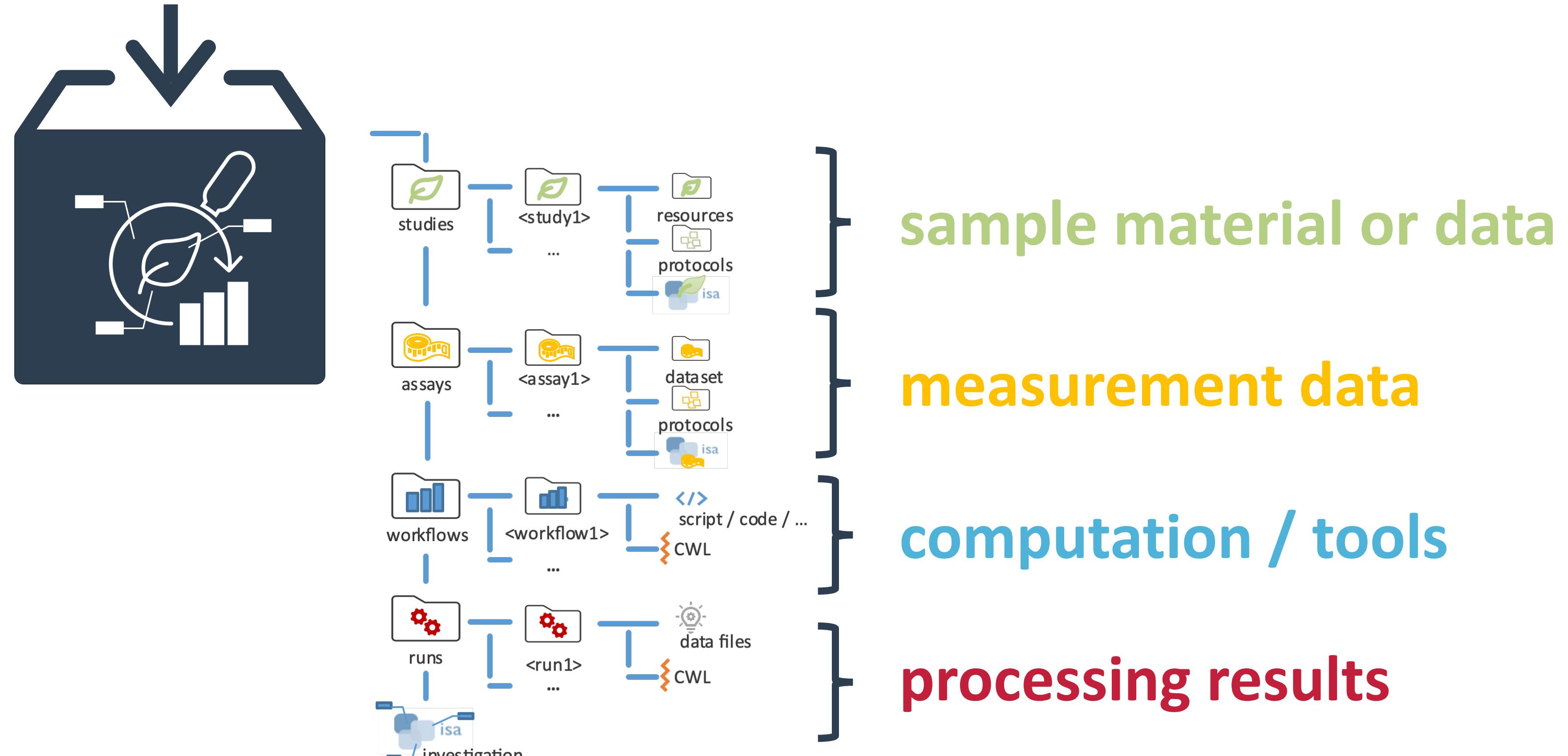
- Domain experts
- User experience

- Service provider
- Technology experts
- Semantic specialists



# Bag your research in an ARC

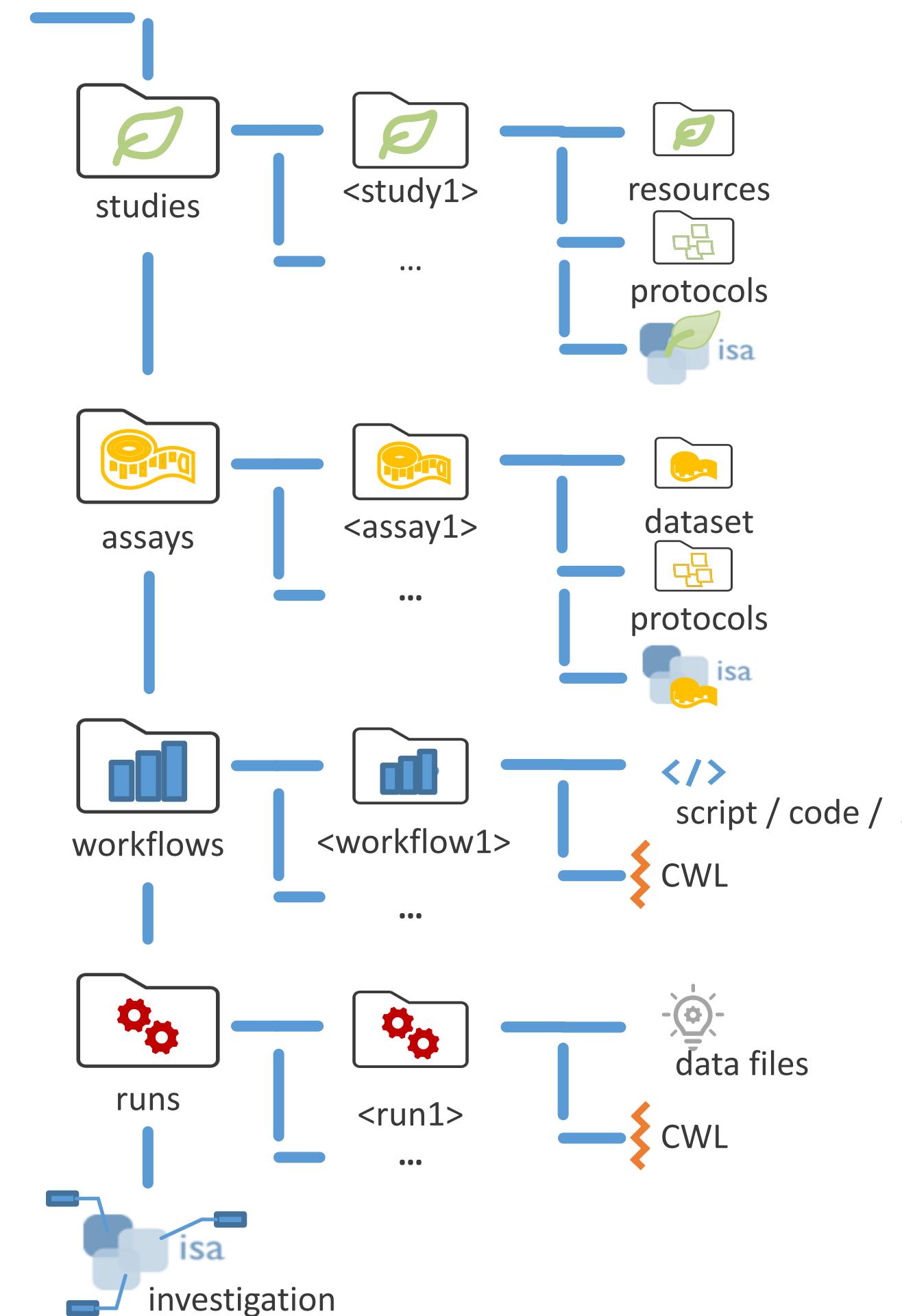
*(Annotated Research Context)*



<https://github.com/nfdi4plants/ARC-specification/>

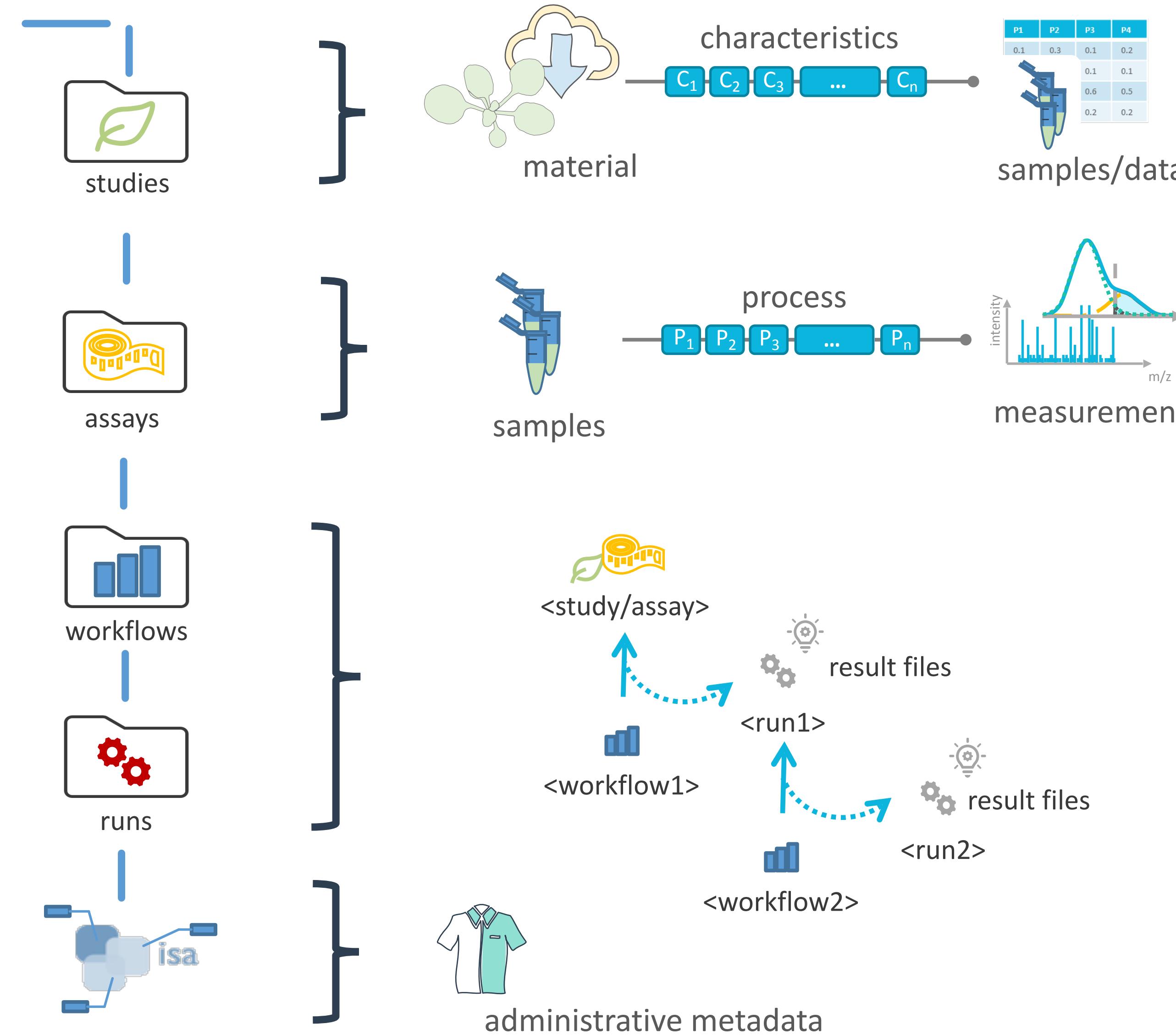


# ARC builds on existing standards





# A real-world ARC is a folder structure





# The ISA investigation file

<b>INVESTIGATION</b>	
Investigation Identifier	ARCeexample_rnaseq
Investigation Title	Facultative CAM in <i>Talinum triangulare</i>
Investigation Description	Highlighting the isa.investigation.xlsx format
Investigation Submission Date	July 07, 2015
Investigation Public Release Date	Nov 05, 2015
<b>INVESTIGATION PUBLICATIONS</b>	
Investigation Publication PubMed ID	26530316
Investigation Publication DOI	10.1104/pp.15.01076
Investigation Publication Author List	Brilhaus, Dominik; Mettler-Altmann, Tabea ...
...	...



# The DataPLANT DataHUB

DataPLANT Menu Search GitLab ? Sign in

## Explore DataPLANT

Discover ARCs, groups and projects in fundamental plant research. Share your projects with others

All Most stars Trending Filter by name... Last updated ▾

Thumbnail	Project Details	Metrics	Last Updated
	Martin Kuhl / SampleARC_RNASeq	★ 0 ⚡ 1 ⚡ 0 ⚡ 0	Updated 3 days ago
	Martin Kuhl / SampleARC_Proteomics	★ 0 ⚡ 1 ⚡ 0 ⚡ 0	Updated 5 days ago
	Lukas Weil / SampleARC_RNASeq	★ 0 ⚡ 0 ⚡ 0 ⚡ 0	Updated 6 days ago



# DataHUB = Customized GitLab

DataHUB

Explore DataPLANT

Discover ARCs, groups and projects in fundamental plant research. Share your projects with others

All Most stars Trending

Filter by name... Last updated

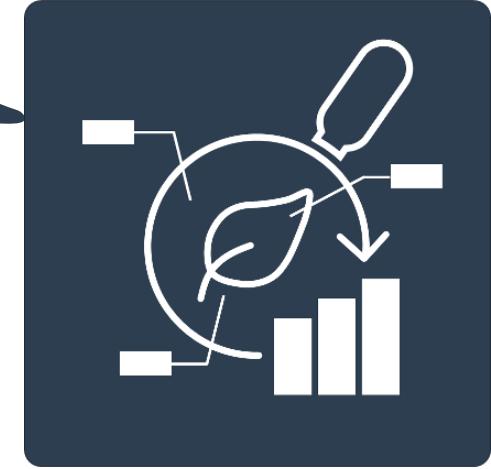
Owner	Project Name	Stars	Forks	Issues	Last Updated
Martin Kuhl	SampleARC_RNASeq	0	0	1	Updated 3 days ago
Martin Kuhl	SampleARC_Proteomics	0	0	1	Updated 5 days ago
Lukas Weil	SampleARC_RNASeq	0	0	0	Updated 6 days ago





# ARCs = Git repositories

DataHUB

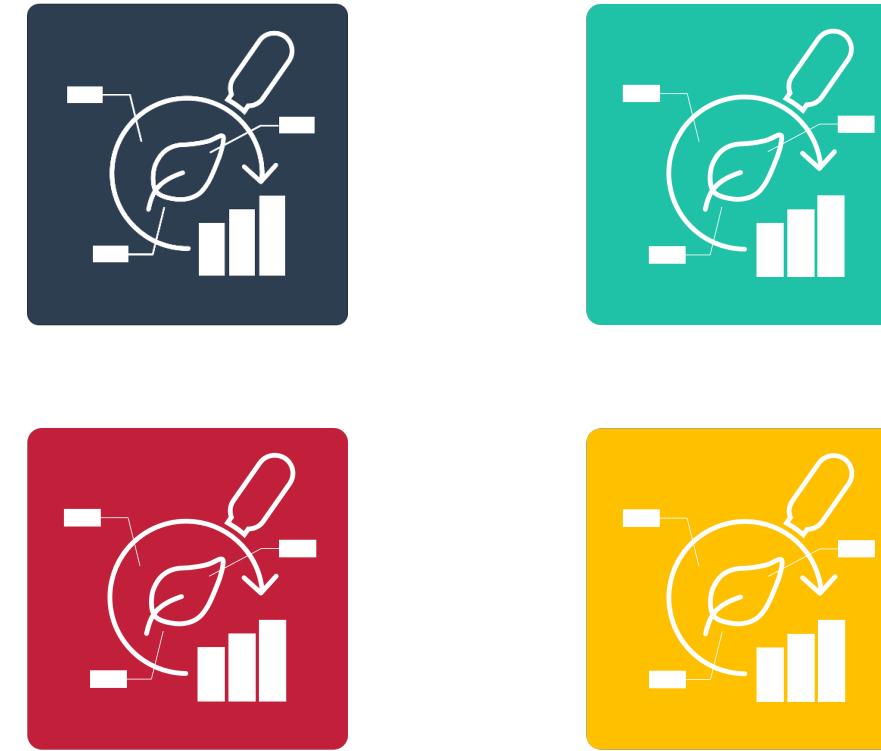


git



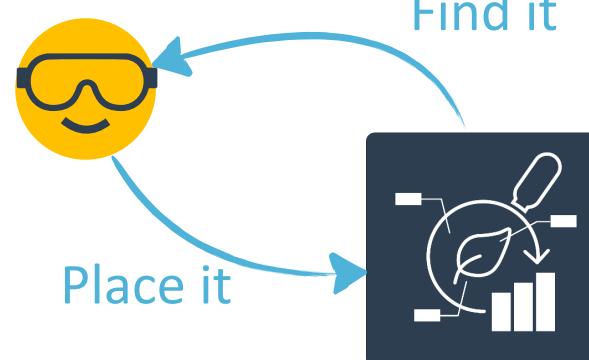
# ARCs = Git repositories

DataHUB



git

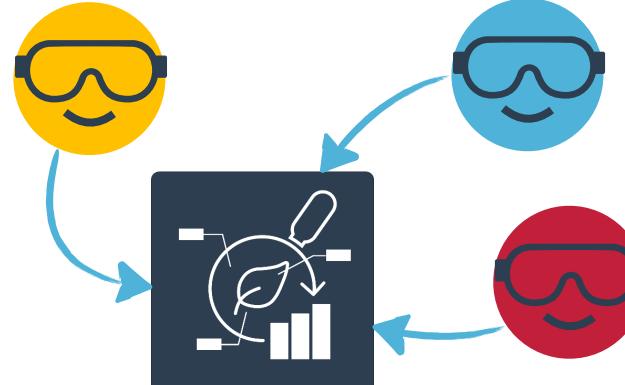
## Package & Document



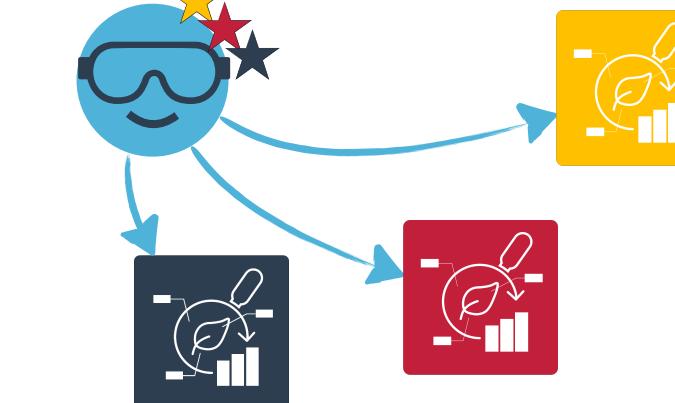
## Backup & Track changes



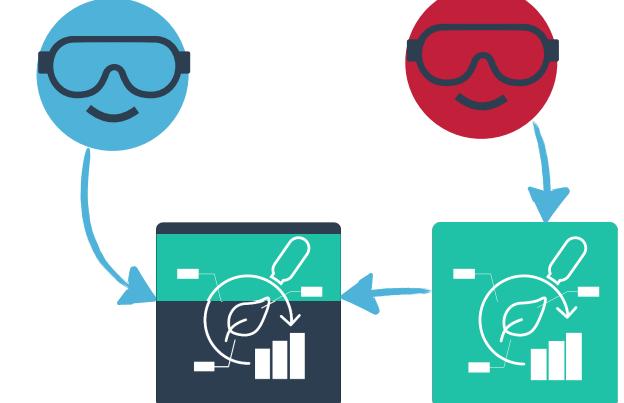
## Collaborate



## Contribute

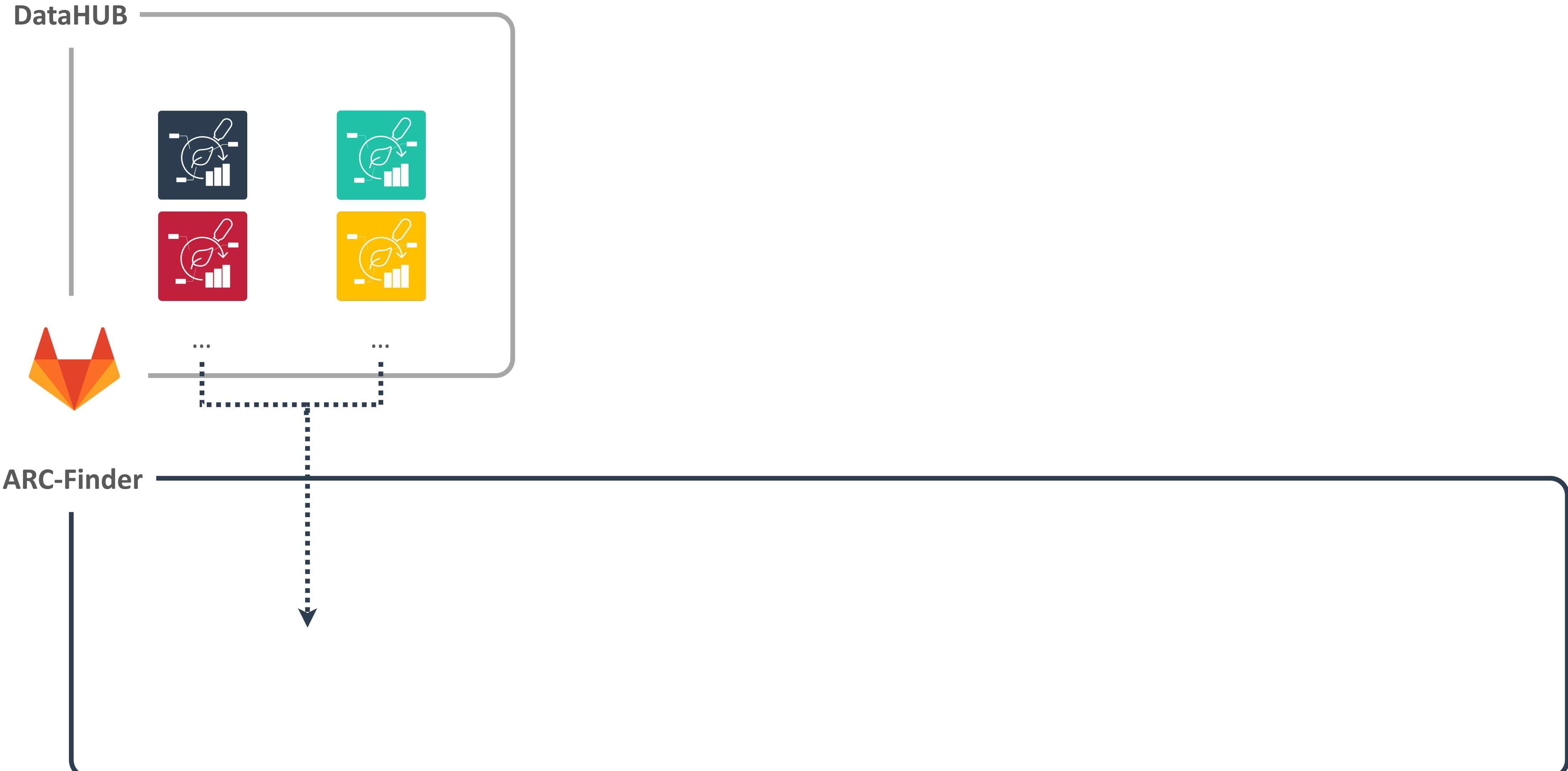


## Reuse





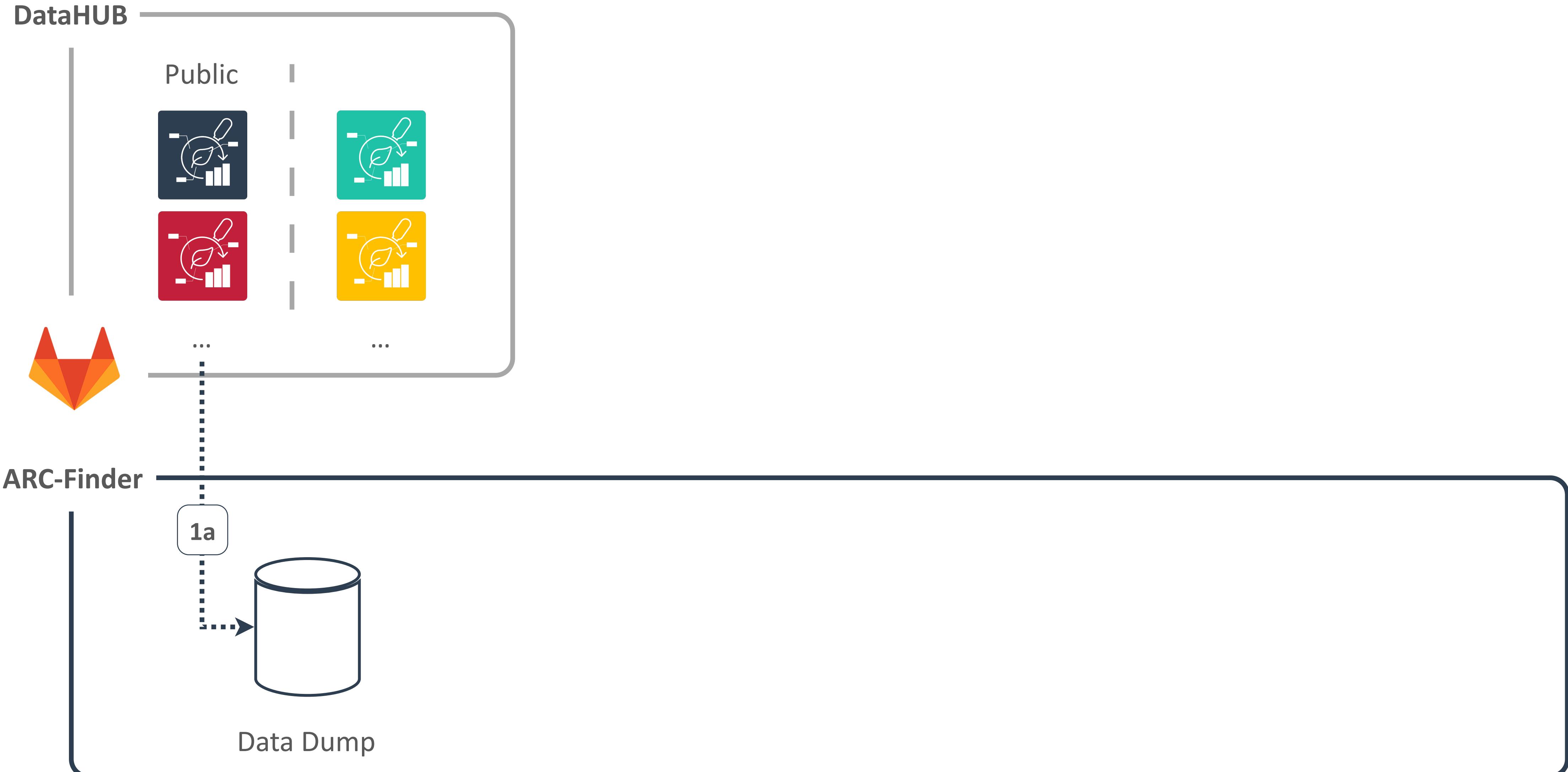
# Extending the ARC environment



<https://github.com/brilator/arcFinder>



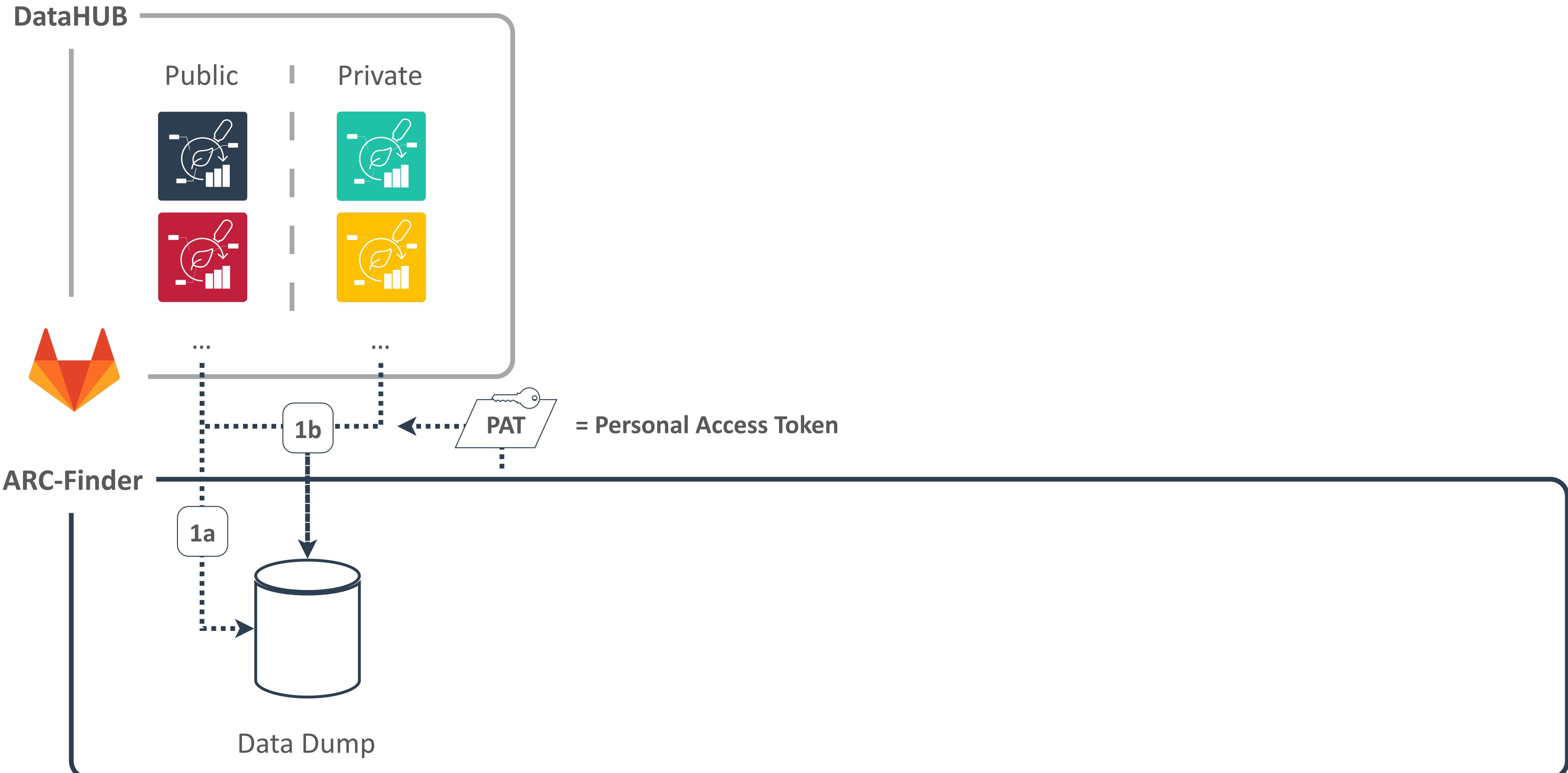
# Retrieve



<https://github.com/brilator/arcFinder>

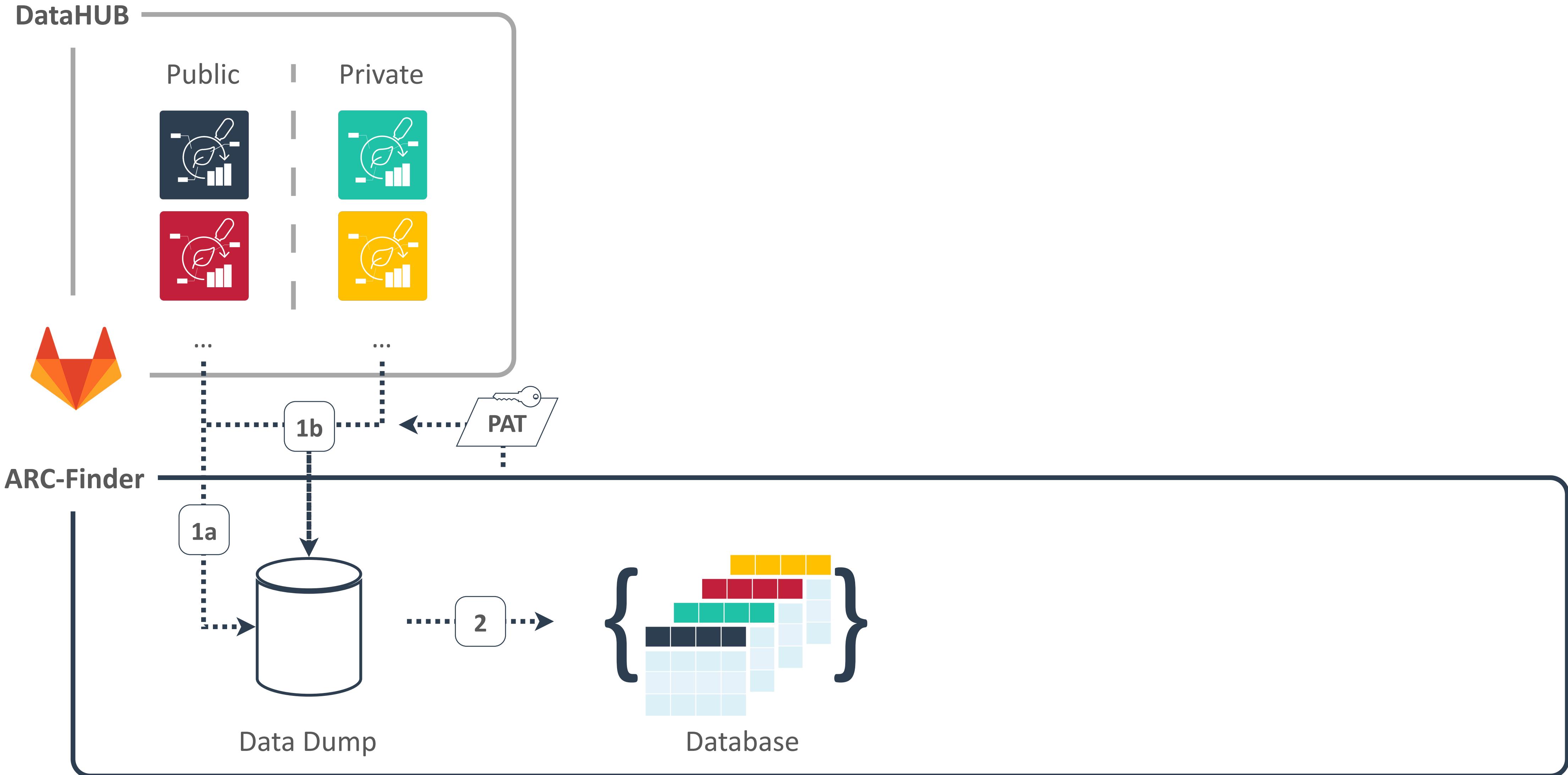


# Retrieve





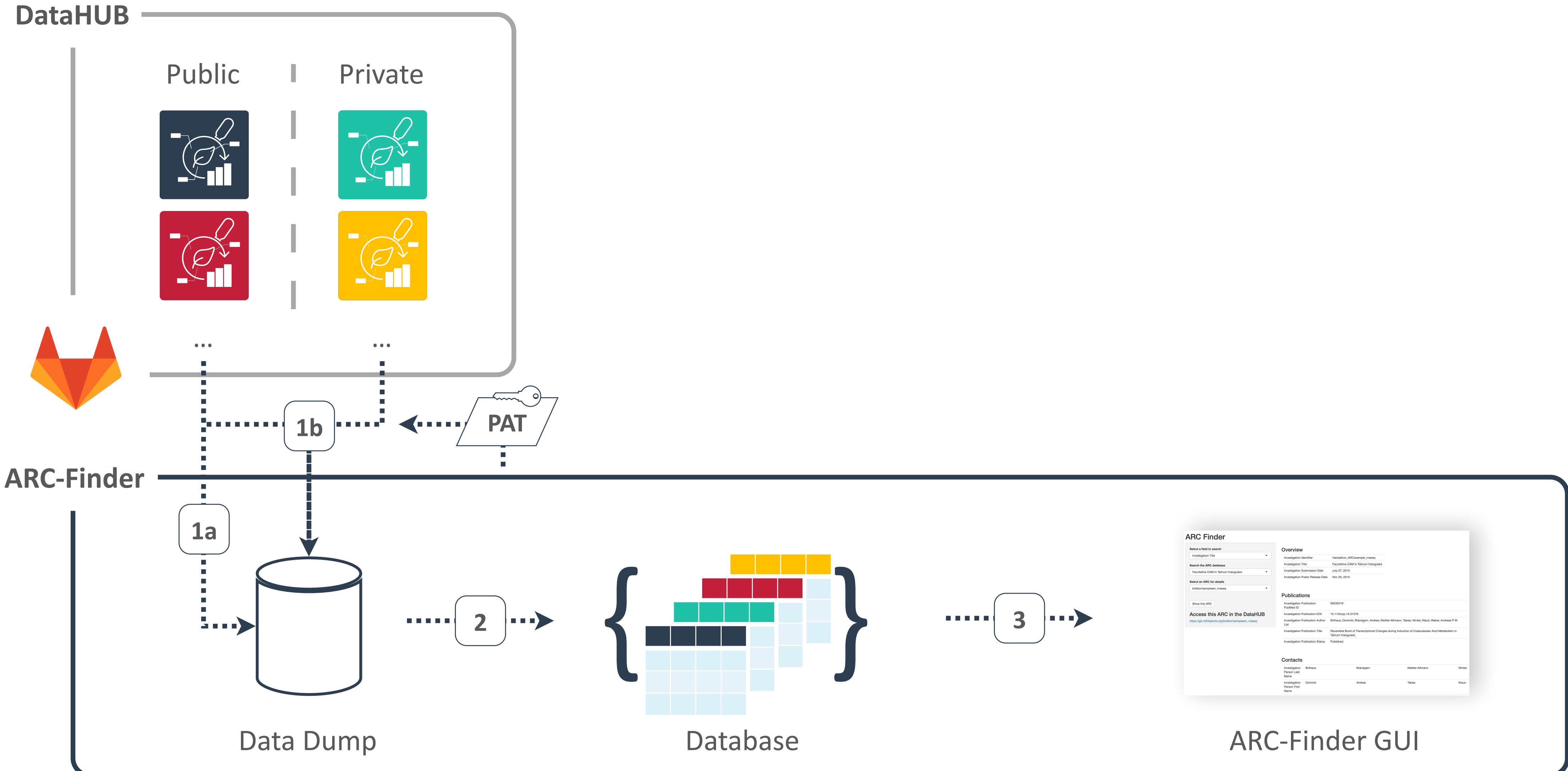
# Restructure



<https://github.com/brilator/arcFinder>



# Represent



<https://github.com/brilator/arcFinder>



# The ARC-Finder GUI

The screenshot displays the ARC-Finder GUI interface. The top bar shows the URL 127.0.0.1. The main window is divided into two panels: the Query Panel on the left and the Result Panel on the right.

**Query Panel:**

- Select a field to search:** Investigation Title
- Search the ARC database:** Facultative CAM in *Talinum triangulare*
- Select an ARC for details:** brilator/samplearc\_rnaseq
- Show this ARC:** (button)

**Access this ARC in the DataHUB:**  
[https://git.nfdi4plants.org/brilator/samplearc\\_rnaseq](https://git.nfdi4plants.org/brilator/samplearc_rnaseq)

**Result Panel:**

## Overview

Investigation Identifier	Hackathon_ARCexample_rnaseq
Investigation Title	Facultative CAM in <i>Talinum triangulare</i>
Investigation Submission Date	July 07, 2015
Investigation Public Release Date	Nov 05, 2015

## Publications

Investigation Publication	26530316
PubMed ID	
Investigation Publication DOI	10.1104/pp.15.01076
Investigation Publication Author List	Brilhaus, Dominik; Bräutigam, Andrea; Mettler-Altmann, Tabea; Winter, Klaus; Weber, Andreas P M
Investigation Publication Title	Reversible Burst of Transcriptional Changes during Induction of Crassulacean Acid Metabolism in <i>Talinum triangulare</i> .
Investigation Publication Status	Published

## Contacts

Investigation Person Last Name	Brilhaus	Bräutigam	Mettler-Altmann	Winter
Investigation Person First Name	Dominik	Andrea	Tabea	Klaus
Investigation Person Email	dominik.brilhaus@hhu.de	NA	NA	NA

Query Panel

Result Panel



# The ARC-Finder GUI

The screenshot shows the ARC-Finder GUI interface. The top bar includes standard window controls (red, yellow, green buttons, minimize, maximize, close) and the IP address "127.0.0.1".

**Query Panel (Left):**

- Select a field to search:** Investigation Title
- Search the ARC database:** Facultative CAM in Talinum triangulare
- Select an ARC for details:** brilator/samplearc\_rnaseq
- Show this ARC:** (button)

**Access this ARC in the DataHUB:**  
[https://git.nfdi4plants.org;brilator/samplearc\\_rnaseq](https://git.nfdi4plants.org;brilator/samplearc_rnaseq)

**Result Panel (Right):**

## Overview

Investigation Identifier	Hackathon_ARCexample_rnaseq
Investigation Title	Facultative CAM in Talinum triangulare
Investigation Submission Date	July 07, 2015
Investigation Public Release Date	Nov 05, 2015

## Publications

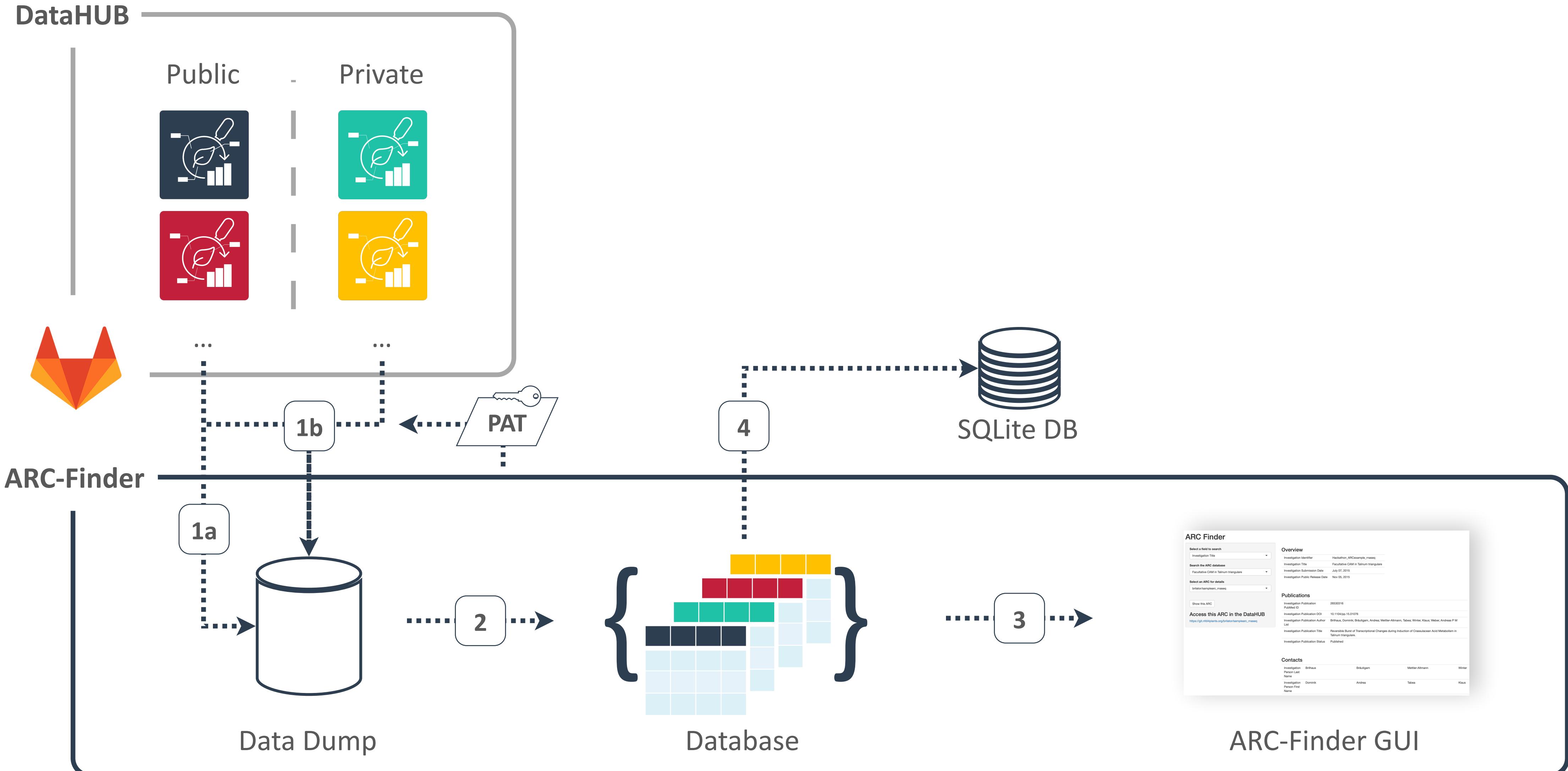
Investigation Publication	26530316
PubMed ID	
Investigation Publication DOI	10.1104/pp.15.01076
Investigation Publication Author List	Brilhaus, Dominik; Bräutigam, Andrea; M...
Investigation Publication Title	Reversible Burst of Transcriptional Changes in Talinum triangulare.

Query Panel

Result Panel



# Alternative: SQLite DB



<https://github.com/brilator/arcFinder>



# Conclusion

- ARC-Finder
  - Quick fix, intermediate solution
  - Vehicle for communication (within DataPLANT)
- Established standards
  - facilitate extensibility
  - boost sustainable RDM



# Thanks!



**CEPLAS**  
ceplas.eu



nfdi4plants.de

**Timo Mühlhaus  
& Group**



1st ARC Hackathon - Sept. 2021

## ZK-FDM Organizers, tutors and participants



**ZBIW.**

Zentrum für Bibliotheks-  
und Informationswissen-  
schaftliche Weiterbildung

**Technology  
Arts Sciences  
TH Köln**

