ONEDATA

OPEN DATA LICECYCLE MANAGEMENT WITH ONEDATA

Michał Orzechowski

ACC Cyfronet AGH

CYFRONET

AGH

EGI-Engage

WHO WE ARE?

ONEDATA

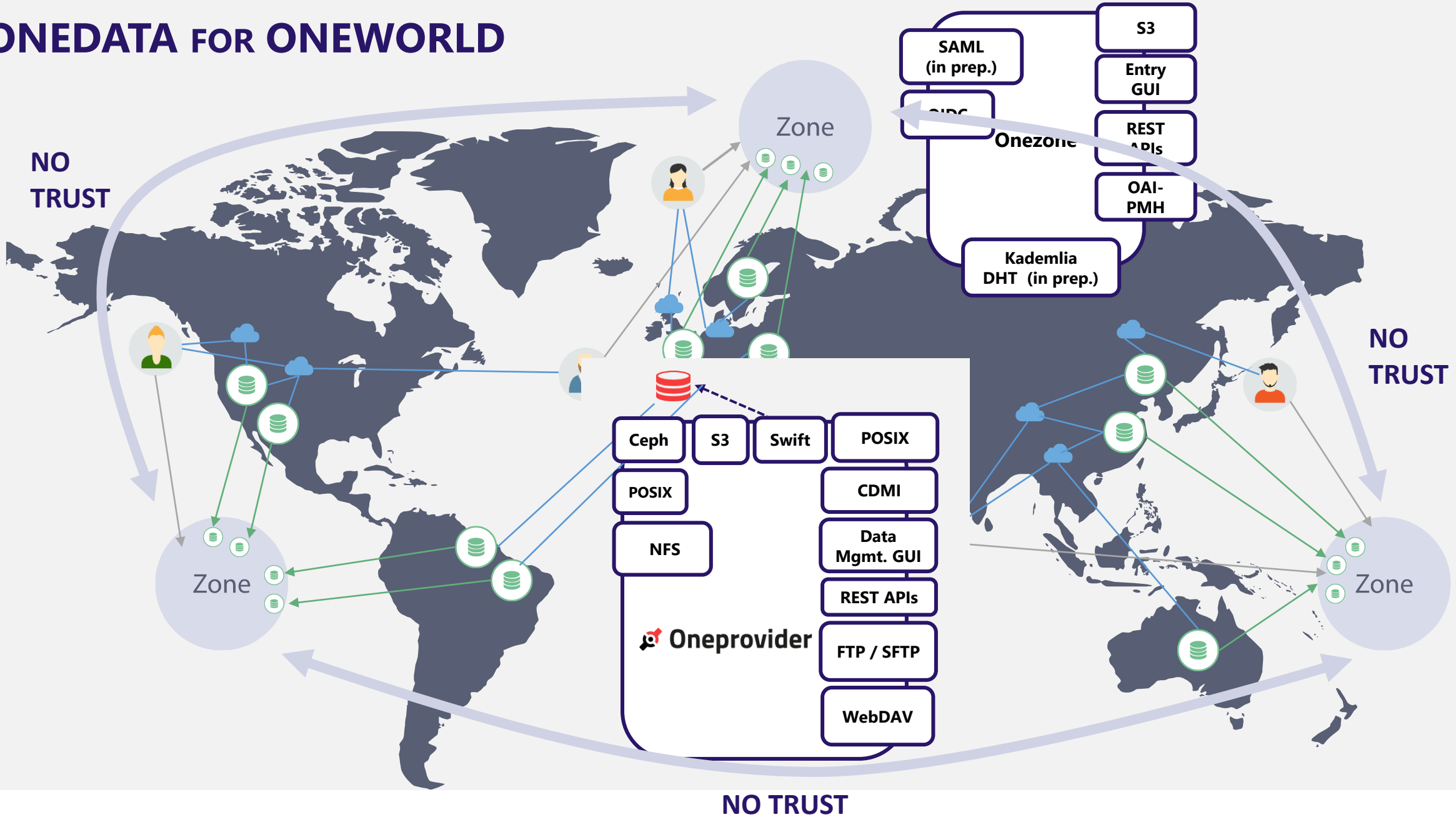
- 10+ years of devoted development - see github.com/onedata
- Open-source, developed at the **AGH University of Krakow** and **Cyfronet** data center
- We work tight with scientific communities on a case-by-case basic
- Our vision is to:
  - deliver a **data management** platform for large-scale and **distributed** problems,
  - address the challenges of global collaborative data sharing across **federated** organizational domains,
  - streamline data processing in **heterogeneous** data storage setups.
- Our funding comes from Polish and European grants and partnerships

SUPPORTED BY SCIENTIFIC COMMUNITIES

- We are always looking for new partnerships and projects in order to:
  - keep the project running (of course),
  - gain invaluable experience cooperating with experts, solving real use cases, and working on authentic large datasets (big files & large number of small files)
- Our supporters and partnerships:

9 YEARS OF ONEDATA @ CS³

ONEDATA FOR ONEWORLD



ONEDATA INSTANCES

- Between 5 and 10 active Zones in Poland and EU (depending on project lifecycles).
- Several instances not maintained by us.
- EGI DataHub (on the map), long haul project:
  - 20 sites (Oneproviders)
  - 2150 data spaces
  - ~1.77PB total storage size
  - 700+ users
- Archive for Polish National Museums:
  - 5PB of data — the current phase
  - 10PB of data — target scale
  - ~100M files

PROBLEMS ADDRESED BY ONEDATA

1. Multi-protocol transparent access to data “[…] but we want POSIX”
2. Heterogeneity of storage technologies
3. Replica Management
4. Easy Data Sharing and publication (DOI)
5. Metadata Management Integrated with Data Management Platform
6. Flexible authentication and authorization
7. Easy integration using API with external services
8. High-throughput data processing
9. Access to Legacy Data Collections
10. Open Data Management

youtube.com/@CS-nh8xn/search

CS3 2017

@CS-nh8xn · 8 subscribers · 51 videos

More about this channel >

Subscribe

Home Videos Playlists Community onedata

Onedata - Eventually Consistent Virtual Filesystem for Multi-Cloud Infrastructures

CS3 2017 • 84 views • 7 years ago

Onedata - Eventually Consistent Virtual Filesystem for Multi-Cloud Infrastructures. Michal Orzechowski.

https://www.youtube.com/watch?v=ia0osTTMVCs

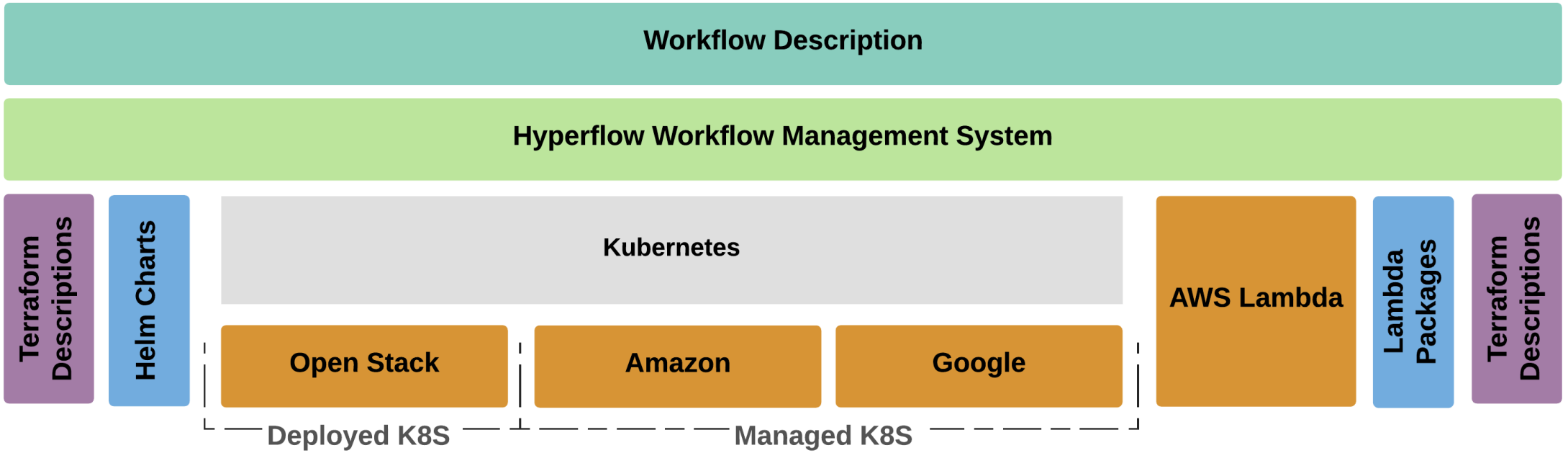
FIRST USECASE:

HYBRID-CLOUD SCINETYFIC WORKFLOW EXECUTION

PROJECT REQUIRMENTS

Requirements:

- achieve automation using open-source DevOps tools (Ansible, Terraform etc.)
- use Kubernetes as a universal cloud API and resource manager
- use Hyperflow workflows manager schedules tasks to multiple Kubernetes clusters

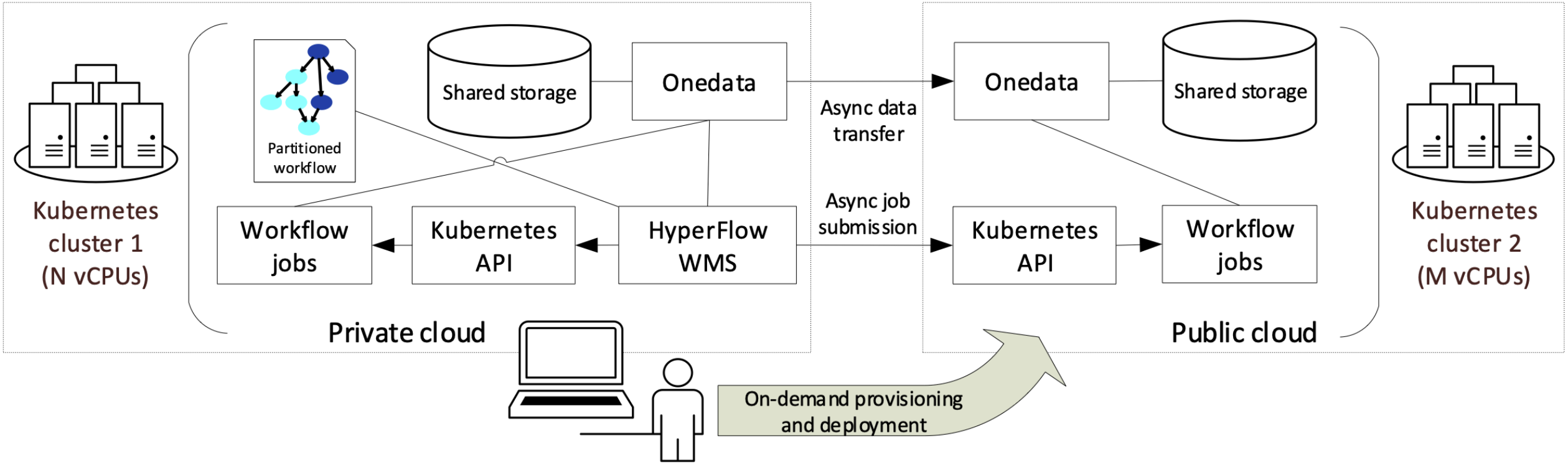


SCINETIFIC WOKFLOW DATA MANAGEMENT

- designed for geographically-local filesystem (eg. Ceph, Lustre, NFS)
- required POSIX-compliant filesystem
- legacy software with unknown data access patterns (sequential/random)
- not designed for distributed data management or cloud-native storage systems

HYBRID-CLOUD DATA MANAGEMENT

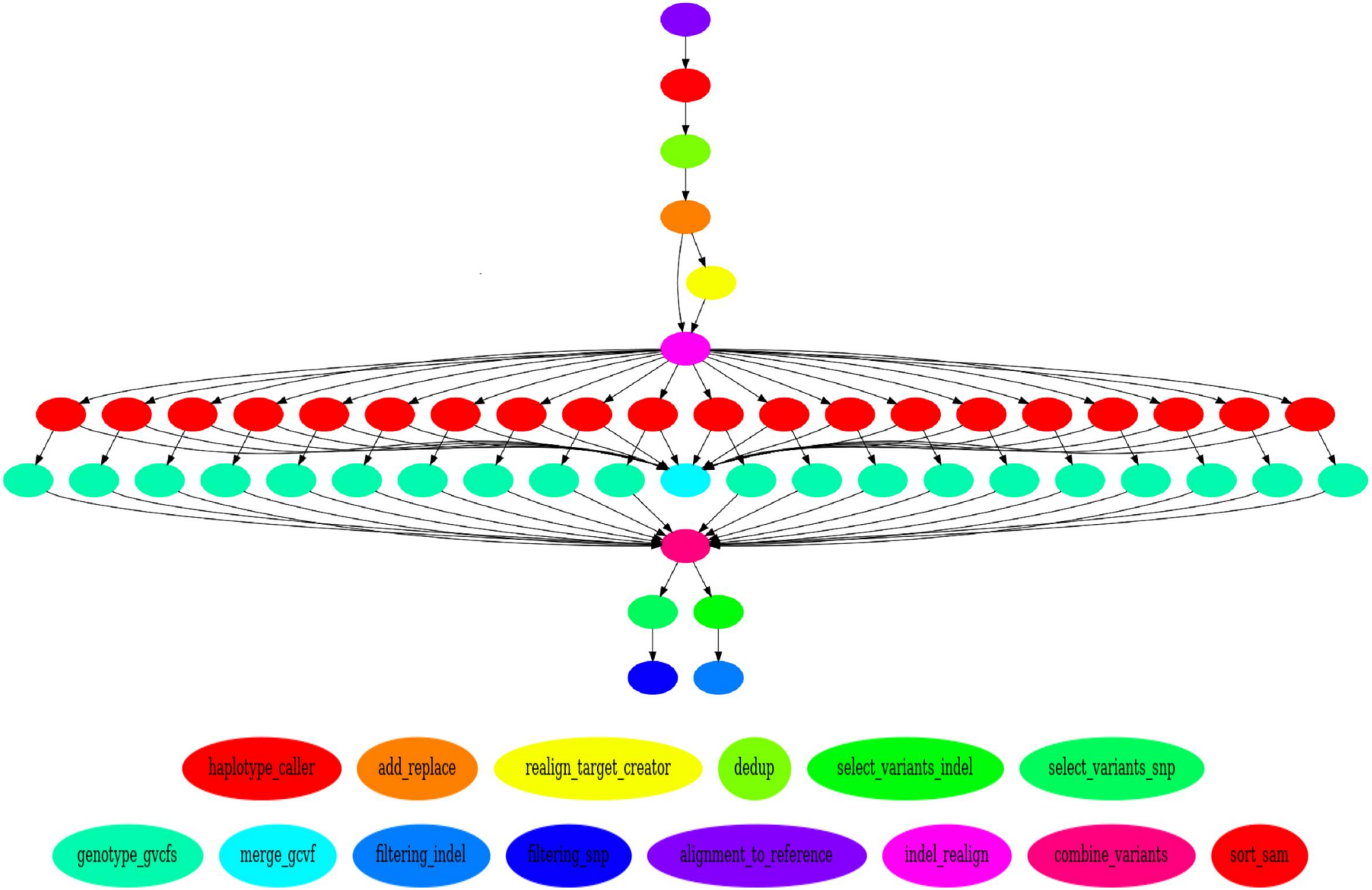
- deployment of Hyperflow and Onedata and Kubernetes clusters
- workflow execution uses Onedata as a data-plane
- on-demand replication of data on the block level



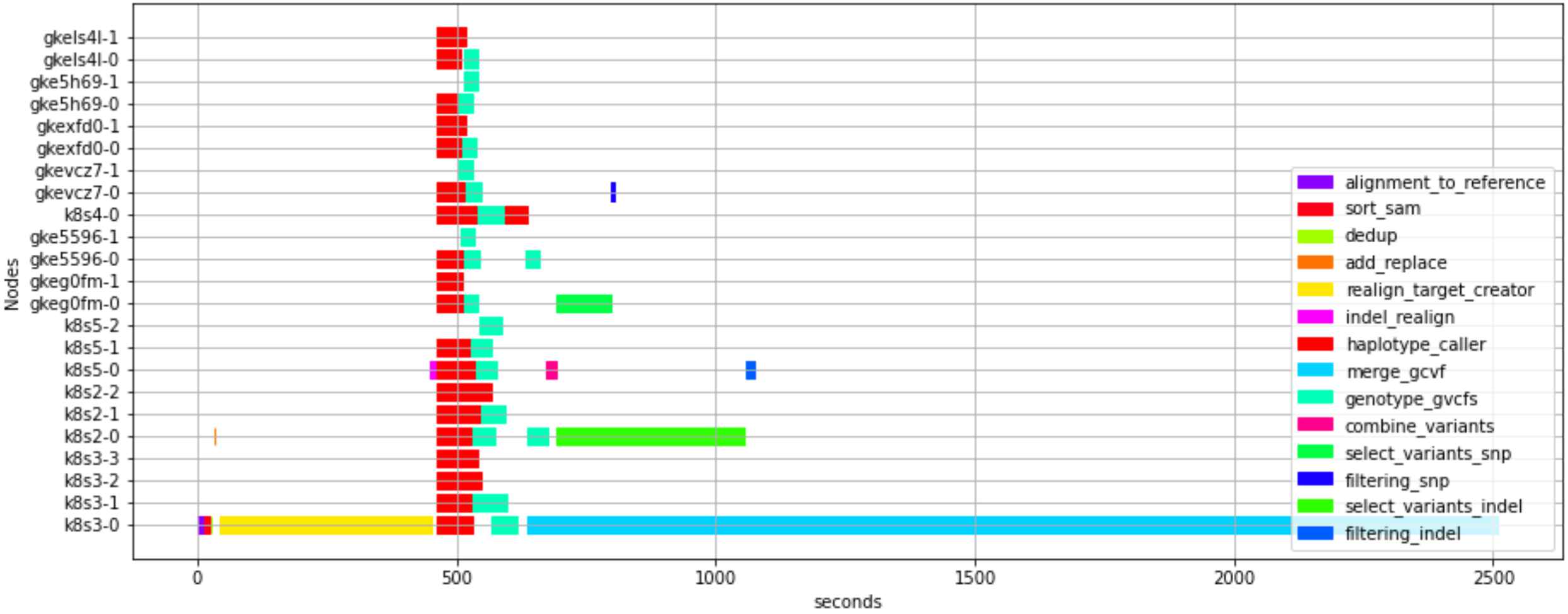
SOYKB GENOMIC WORKFLOW

Workflow for studying genomic variation of plants using next generation resequencing data with large-scale inputs, with following characteristics:

- a large input file – the reference genome database,
- two parallel stages (genotype_gvcfs and filtering_snp tasks),
- a long final task that merges the outputs of previous tasks (merge_gcvf)

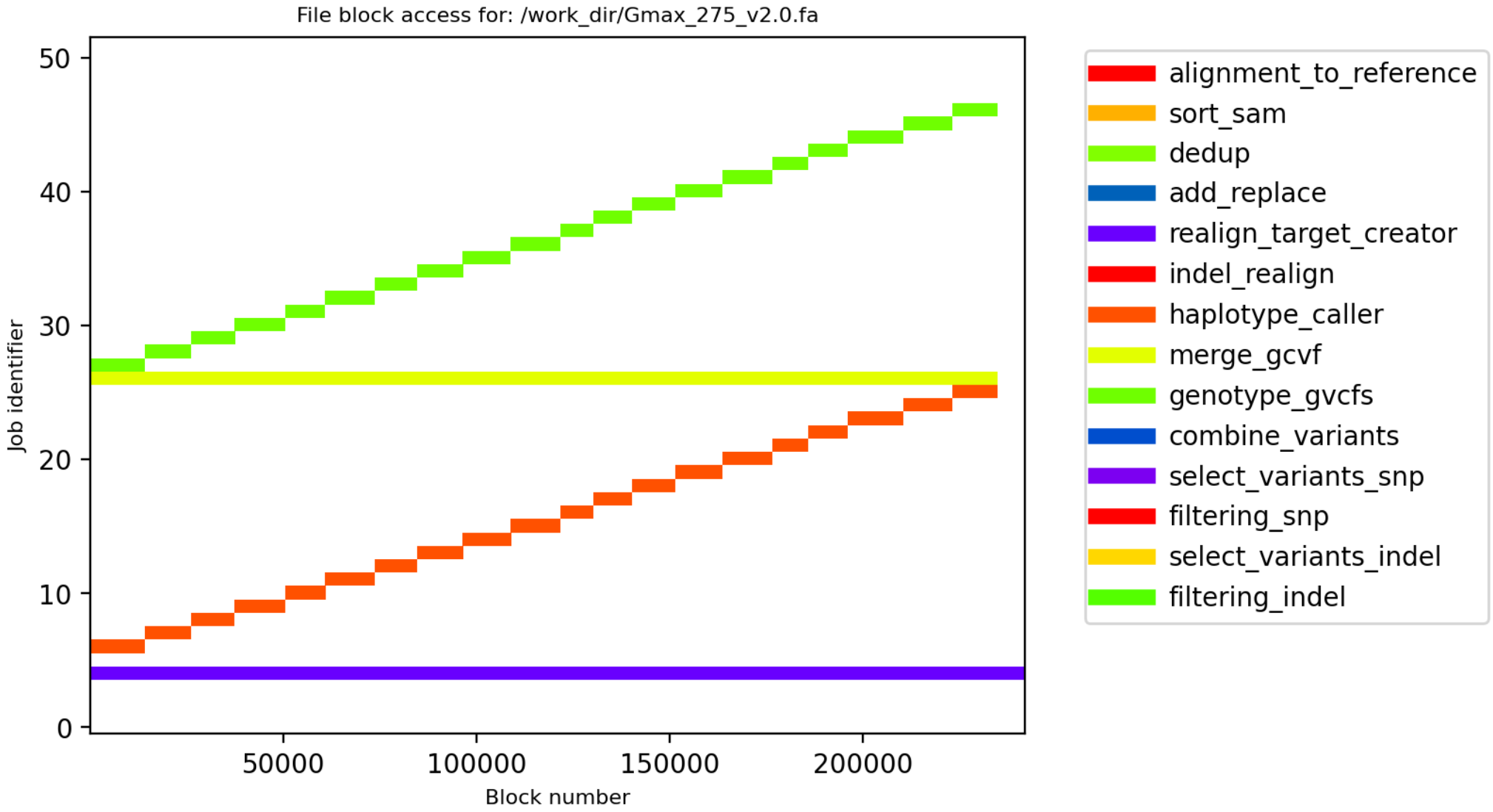


Structure of the Soykb workflow



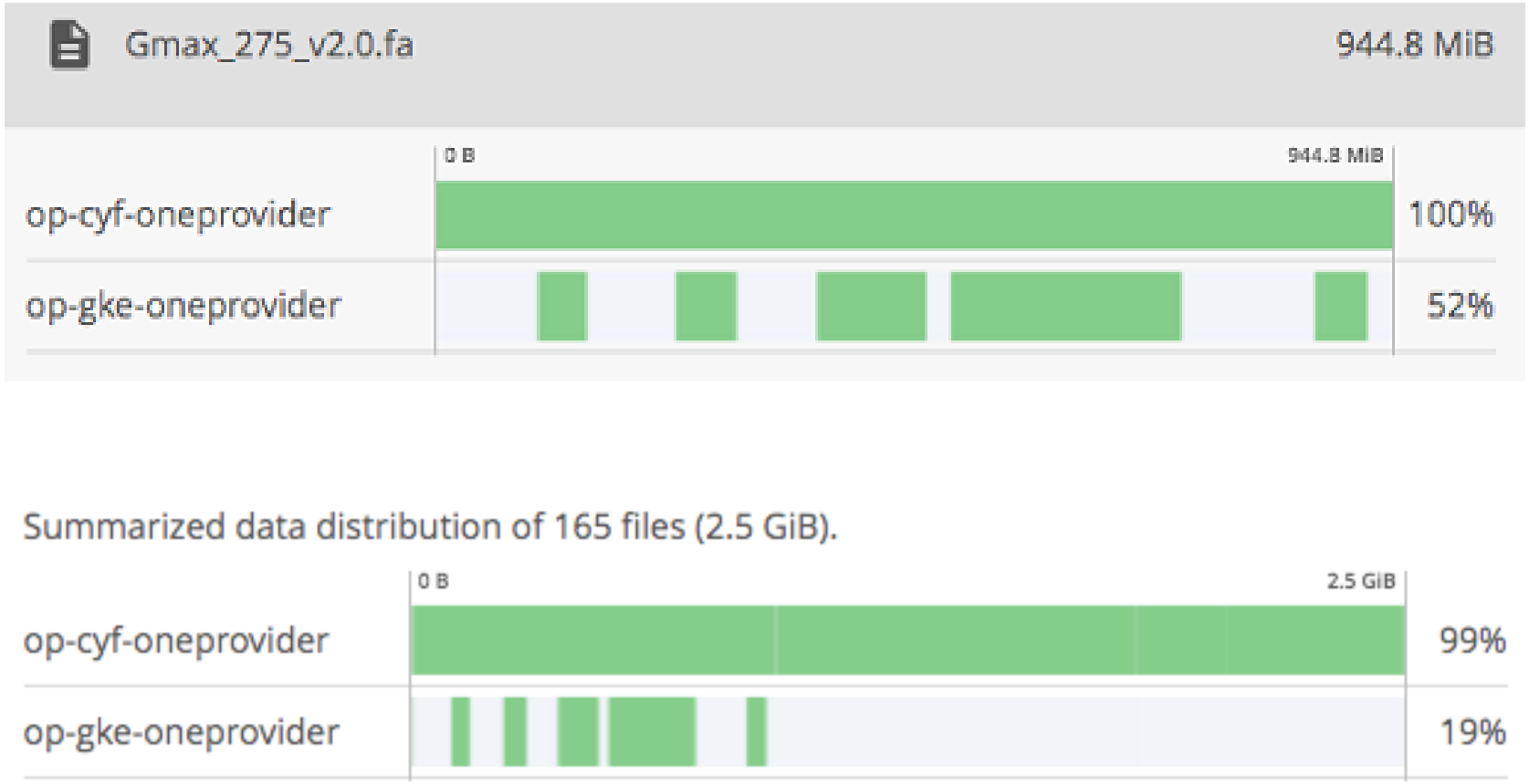
Visualization (trace) of the Soykb workflow execution on two clouds

SOYKB DATA ACCESS PATTERNS



Data access pattern of the largest input file of the Soykb workflow.

BLOCK LEVEL DATA DISTRIBUTION IN ONEDATA

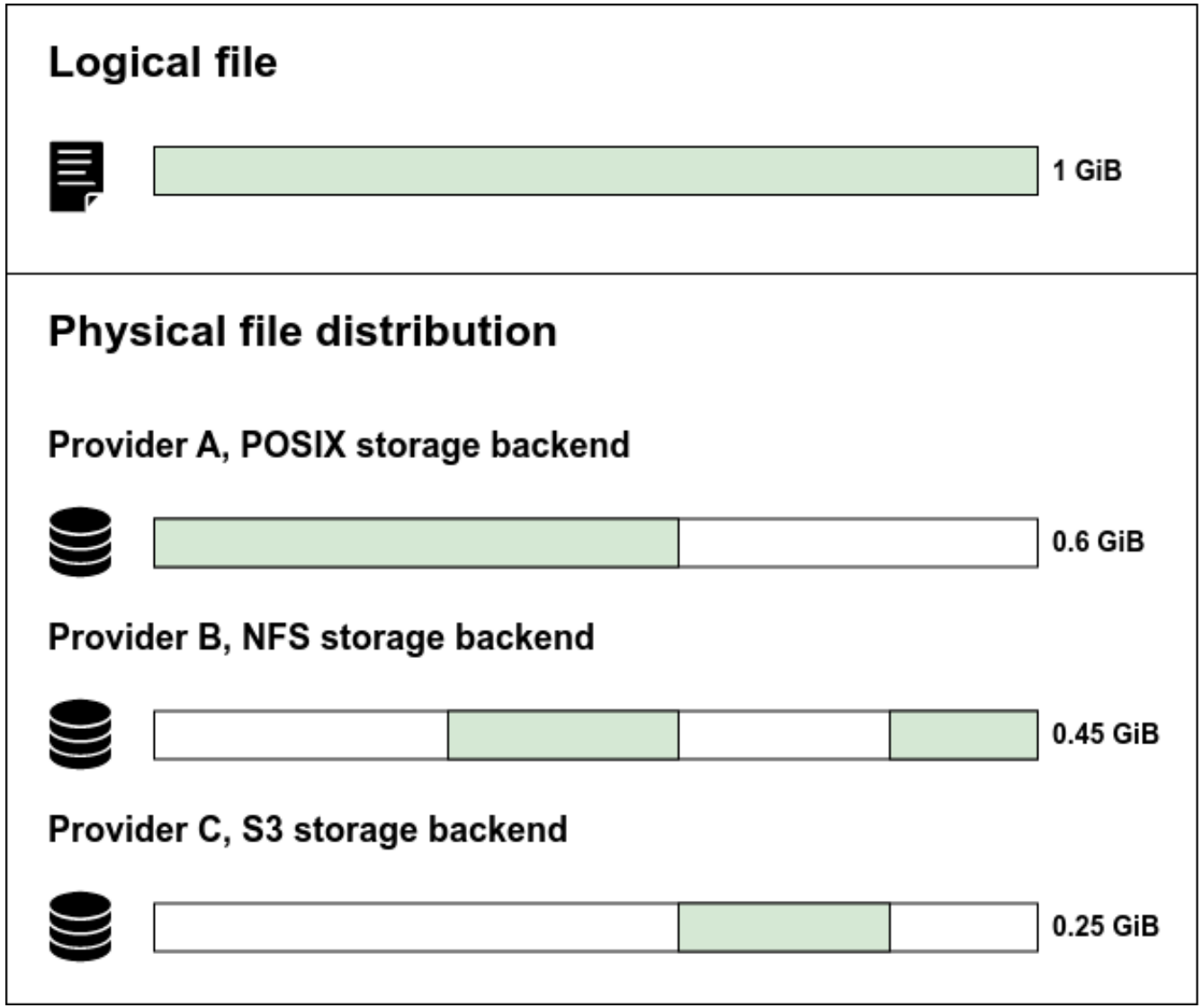


Runtime data distribution between two clouds for the Soykb workflow

DATA DISTRIBUTION

- the data in Onedata may be arbitrarily distributed among the storage backends of the supporting providers
- files are made up of parts of variable sizes — file blocks
- each provider holds a set of local file blocks, constituting a file replica
- when a file is read on a provider and the requested blocks are not present there, the missing ones are replicated on the fly from remote providers
- when a file is written on a provider, the overwritten blocks on other providers are invalidated. To read the file, the provider with invalidated blocks must once again replicate missing blocks from the provider with the newest version of the blocks

DATA DISTRIBUTION



FILE DETAILS

download-files.json

Info | API | Metadata | Permissions | Shares | QoS | **Distribution**

Data distribution per storage

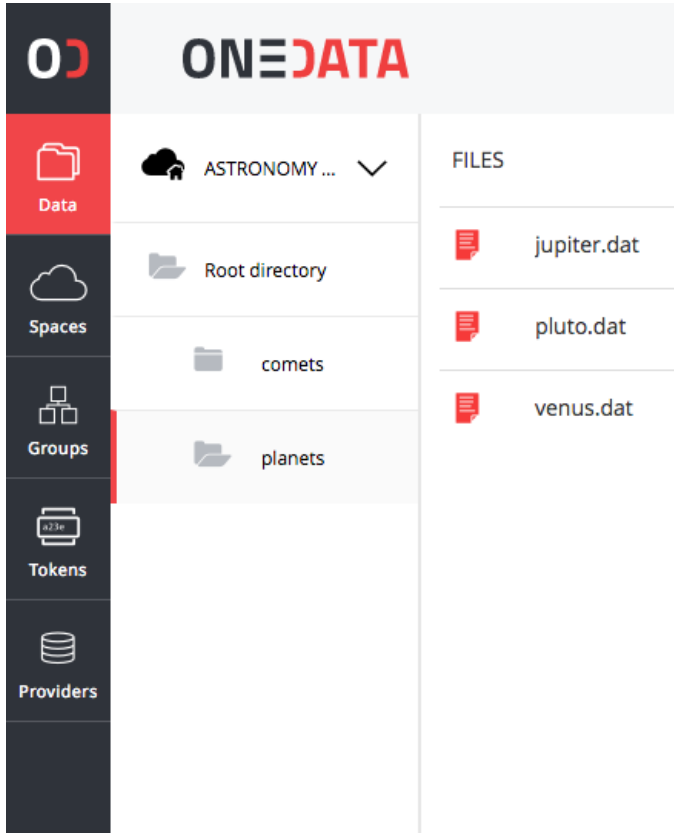
onedatify @ **Bari** /55eb75ab388522d56062f36fada12b40chb188/bw/download…	0 B — 45 KiB in 2 blocks — 60 KiB	75%
ceph @ **Krakow** /55eb75ab388522d56062f36fada12b40chb188/a/2/7/a271c…	0 B — 14.9 KiB in 2 blocks — 60 KiB	24%
onedatify @ **Lisbon** /55eb75ab388522d56062f36fada12b40chb188/bw/download…	0 B — 20.1 KiB in 2 blocks — 60 KiB	33%
posix @ **Paris**	0 B — 0 B in 0 block — 60 KiB	0%

This file was transferred manually 1 time – see history.

Block distribution

POSIX ACCESS WITH ONECLIENT

- presents Onedata virtual file system as POSIX
- support for most of the POSIX operations on globally distributed virtual file system
- all data accessible via a unified file system mountable on virtual machines, grid worker nodes and containers



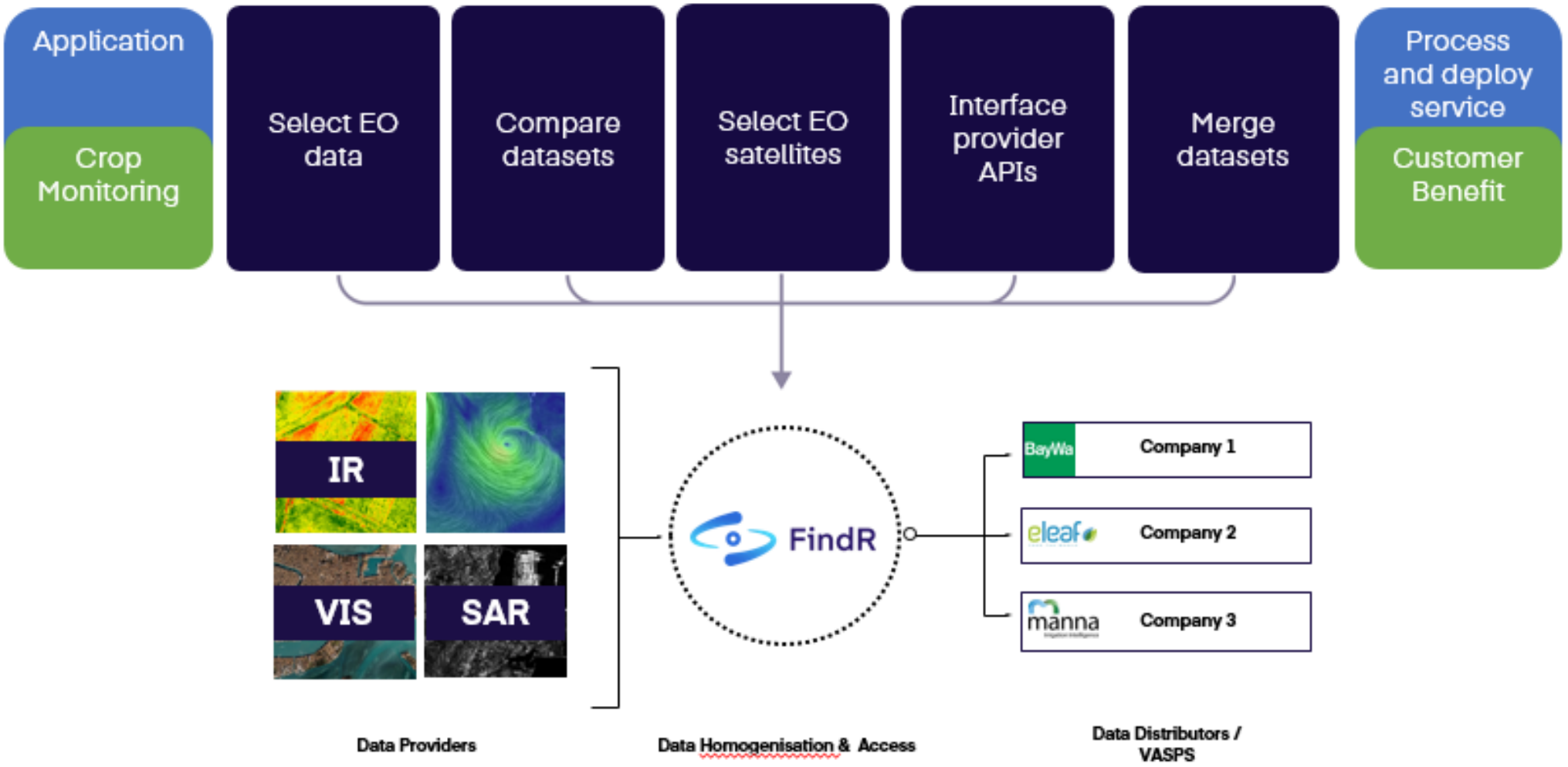

```
[root@1f87c053280e oneclient]# ls
Astronomy Datasets  Big Data Experiment  Cancer Data
[root@1f87c053280e oneclient]# ls -lR
.:
total 0
drwxrwx--- 1 root 1733762 0 Sep 26 19:19 Astronomy Datasets
drwxrwx--- 1 root 1337123 0 Sep 26 19:14 Big Data Experiment
drwxrwx--- 1 root  608582 0 Sep 26 19:18 Cancer Data

./Astronomy Datasets:
total 0
drwxr-xr-x 1 1124656 1733762 0 Sep 26 19:20 comets
drwxr-xr-x 1 1124656 1733762 0 Sep 26 19:19 planets

./Astronomy Datasets/comets:
total 0
-rw-r--r-- 1 1124656 1733762 10000000 Sep 26 19:20 enck.dat
-rw-r--r-- 1 1124656 1733762 10000000 Sep 26 19:19 halley.dat

./Astronomy Datasets/planets:
total 0
-rw-r--r-- 1 1124656 1733762 10000000 Sep 26 19:07 jupiter.dat
-rw-r--r-- 1 1124656 1733762  5000000 Sep 26 19:08 pluto.dat
-rw-r--r-- 1 1124656 1733762  2000000 Sep 26 19:08 venus.dat

./Big Data Experiment:
total 0
-rw-r--r-- 1 1124656 1337123 10000000 Sep 26 19:08 cats_images.tgz
-rw-r--r-- 1 1124656 1337123  5000000 Sep 26 19:13 galaxies.img
-rw-r--r-- 1 1124656 1337123  5000000 Sep 26 19:14 spam_mails.tgz

./Cancer Data:
total 0
-rw-r--r-- 1 1124656 608582 5000000 Sep 26 19:15 brain_tumor.zip
-rw-r--r-- 1 1124656 608582 5000000 Sep 26 19:14 duct_cancer.zip
[root@1f87c053280e oneclient]#
```

SECOND USECASE:

SATELITE IMAGE HOMOGENIZATION

SATELITE IMAGE HOMOGENIZATION

FindR provides **unified access** to several Data Providers and **homogenize** the data to a unified format. This includes a common **tiling sheme, resolution, spectral bands** and other characteristics of the output to enable the integration of new data sources without effort.

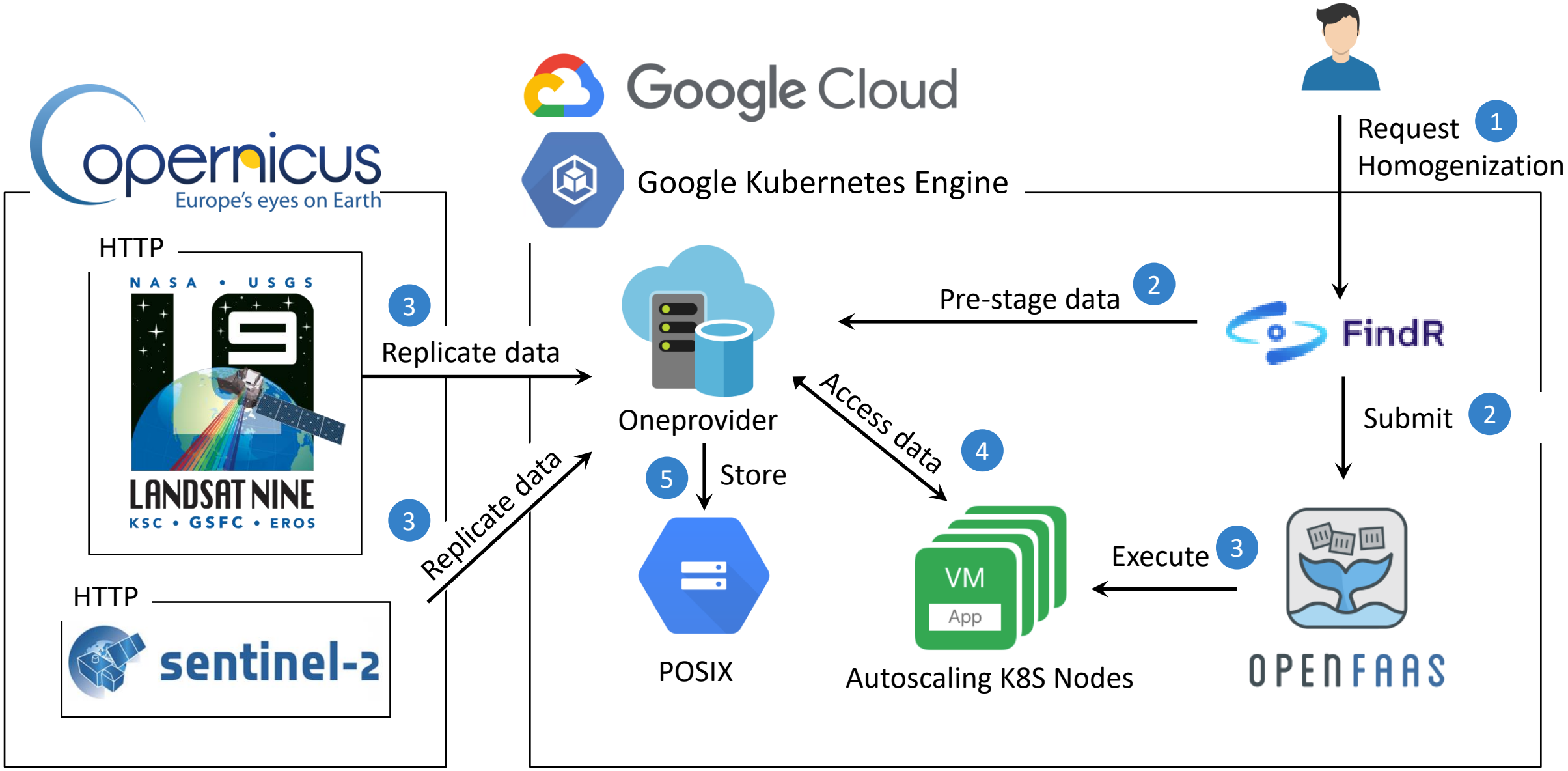


MULTIPLE SOTRAGE BACKENDS

Storage backends are used to store the physical data. Oneprovider accesses the storage backends via "helpers" (drivers) implemented for each supported type of storage. Helpers serve as a POSIX-like abstraction, building a layer over different storage backend APIs and access methods.

Currently supported storage backends:

- **POSIX** — any POSIX compatible filesystem accessible by Oneprovider via a mount point
- NFS — filesystem exported via the NFS protocol — no need to mount it locally
- S3 — Amazon S3 compatible storage
- Ceph RADOS — versions 14, 15, 16
- **HTTP** — any server exposing data via HTTP or HTTPS
- XRootD — CERN's data management protocol for LHC data
- WebDAV — experimental
- dCache — experimental


Google Cloud
Google Kubernetes Engine
Request
Homogenization
1
Copernicus
Europe's eyes on Earth
HTTP
NASA • USGS
LANDSAT NINE
KSC • GSFC • EROS
3
Replicate data
Pre-stage data
2
FindR
Oneprovider
Access data
4
Submit
2
5
Store
3
Replicate data
HTTP
sentinel-2
POSIX
VM
App
Autoscaling K8S Nodes
Execute
3
OPENFAAS

MANUAL DATA INDEXING

- Manual import allows registering specific files lying on a storage backend into a Onedata
- Preferred when only a part of the dataset is to be imported
- Upon registration, the file becomes accessible in Onedata under the requested path
- No data is copied in the process - a metadata entry is created in the space, containing a reference to the external storage.
- Manual storage import is supported on the following storage backends: POSIX, GlusterFS, WebDAV, XRootD, HTTP, S3, Swift, Ceph RADOS

```
curl -v -H "X-Auth-Token: $PROVIDER_ACCESS_TOKEN" -X POST "https://$PROVIDER_HOST/api/v3/oneprovider/data/register" \
-H 'Content-Type: application/json' -d '{
  "spaceId": "'$MANUAL_IMPORT_SPACE_ID'",
  "storageId": "'$HTTP_STORAGE_ID'",
  "storageFileId": "https://packages.onedata.org/apt/ubuntu/2002/dists/focal/Release",
  "destinationPath": "/manually_imported_file"
}'
```

AUTO CLEANING & FILE POPULARITY

Overview | Storage import | File popularity | **Auto-cleaning**

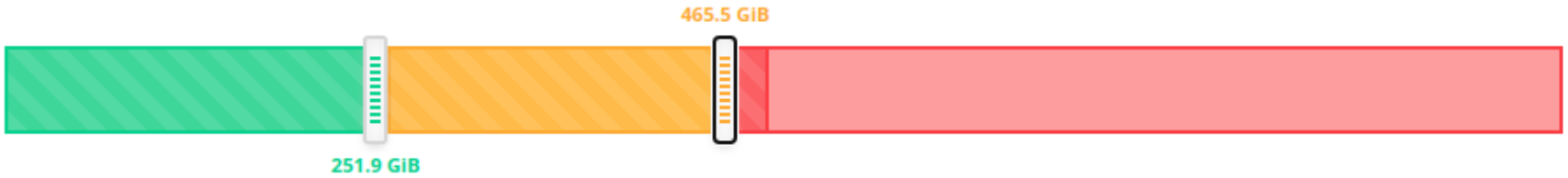
ENABLE AUTO-CLEANING

SELECTIVE CLEANING

- [x] Lower size limit: 1 MiB
- [ ] Upper size limit: 1 PiB
- [x] Not opened for: 5 Hours
- [ ] Opened up till now: 900719 times
- [ ] Opened per hour: 900719 times
- [x] Opened per day: 1000 times
- [ ] Opened per month: 900719 times

CLEANING BOUNDARIES

Total space: **1 TiB** Used space: **500 GiB** Free space: **524 GiB** To release: **248.1 GiB**



Start cleaning now

CLEANING REPORTS

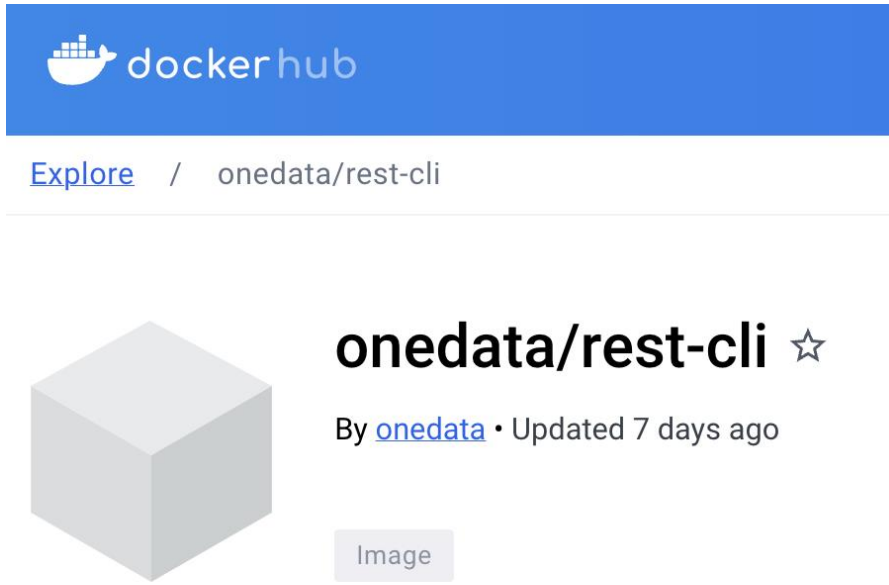
Start	Stop	Released size	Files number	Status

No reports available.

RICH COLLECTION OF APIs

Integrate external tools using rich API interfaces with data management platform and build more complex environments for data processing:

- APIs for all operations
- APIs for full eventually consistent integration with external systems
- API for data locality and replication
- Flexible permission checking for APIs
- API fully described using Swagger for generation of clients based on API specification
- Easy to use simple command line clients for REST API



THIRD USECASE:

PROCESING OF LARGE MUSEAL DATASETS

	ARCHIWA PAŃSTWOWE NARODOWE ARCHIWUM CYFROWE NATIONAL DIGITAL ARCHIVE	BIBLIOTEKA NARODOWA NATIONAL LIBRARY	Muzeum Narodowe w Krakowie NATIONAL MUSEUM IN KRAKOW
Number of files	4 millions	3 millions	37 millions
Number of dirs	1	33667	154578
Size	2068 TB	960 TB	13 TB

In Total: 44 million files and 3 Petabytes of data.

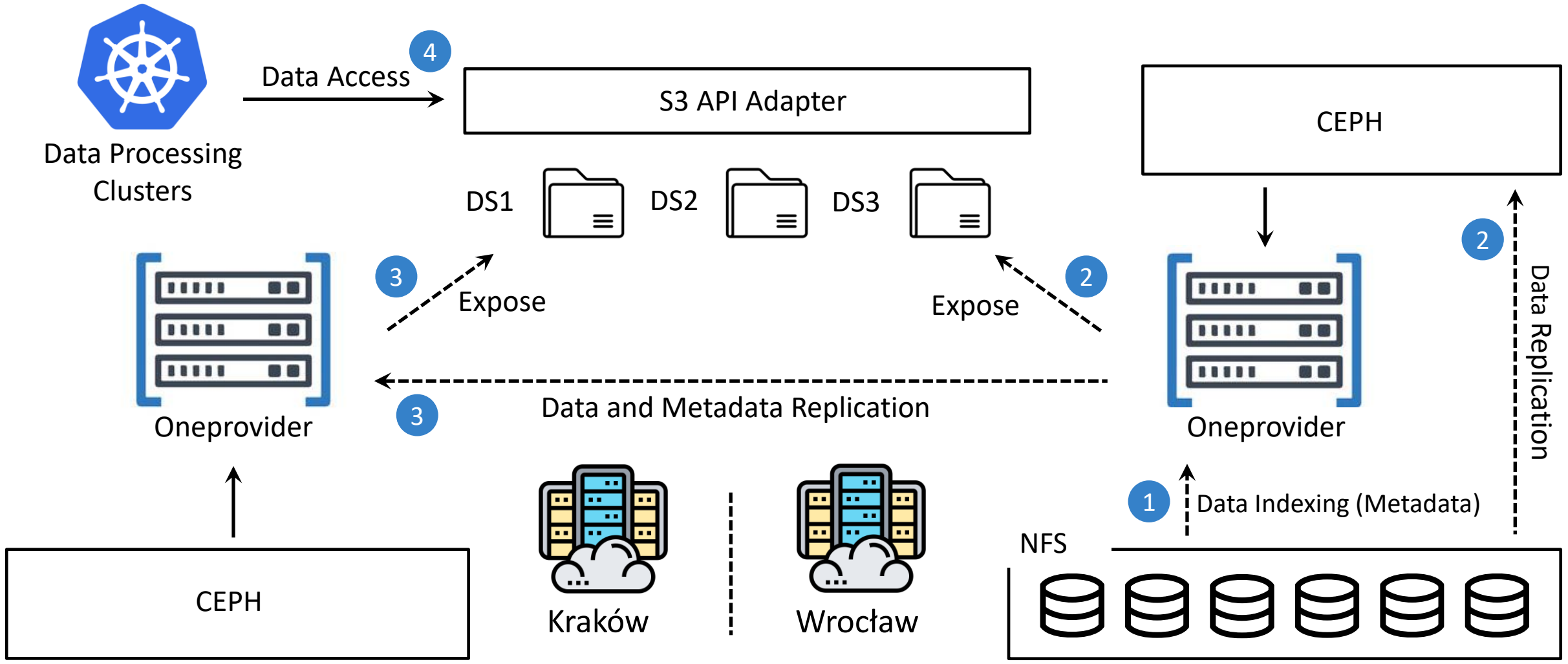
Planned: 100 million files and 10 Petabytes of data.

PROJECT REQUIRMENTS AND OBJECTIVE

Requirements:

- high availability of data (replicated across two sites)
- high availability of data access and processing services
- data processing (image compression, metadata extraction) to start as soon as possible
- data processing software requires S3 API for data access
- results available to public as soon as possible

Objective: provide a portal where general public can search and access all available data.


Data Processing Clusters
4
Data Access
S3 API Adapter
CEPH
DS1
DS2
DS3
3
Expose
2
Expose
2
Data Replication
Oneprovider
3
Data and Metadata Replication
Oneprovider
1
Data Indexing (Metadata)
NFS
CEPH
Kraków
Wrocław

AUTOMATIC DATA INDEXING

The file indexing process does not copy any data, it simply creates the necessary metadata so that the files pre-existing on the storage are reflected and accessible in Onedata.

This functionality should be used in two setups:

- There is a legacy dataset located on the storage backend
- The data on storage backend is to be modified directly by third party applications, and the changes should be reflected in Onedata.

It is possible to configure the storage to:

- detect consecutive data changes
- manually triggering scans
- re-registering files

Set	A	B	C
Number of files	3934561	3180622	36790834
Number of directories	1	33667	154578
Total Size [TB]	2068	960	131
Time to Scan [h]	16.6	21.8	22.2

Overview | Storage import | File popularity | Auto-cleaning

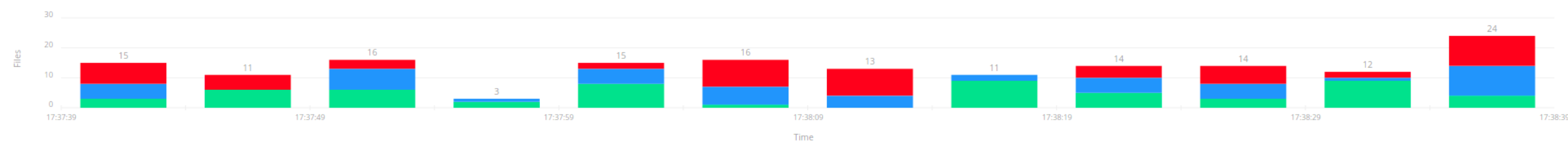
Auto storage import configuration

Auto import scan **completed**, next scan: **19:38:39**

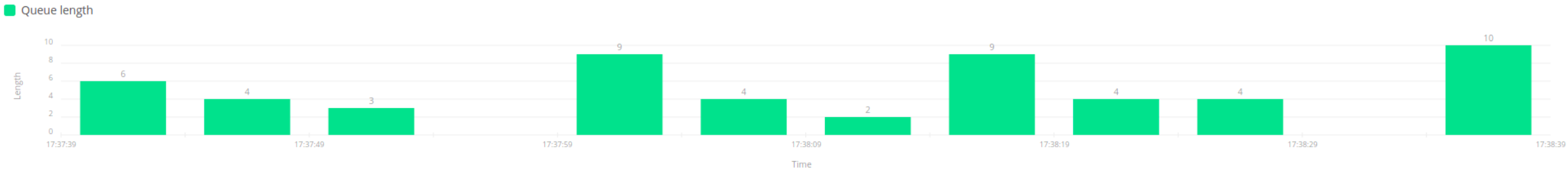
Start scan

IMPORT FILES PROCESSING

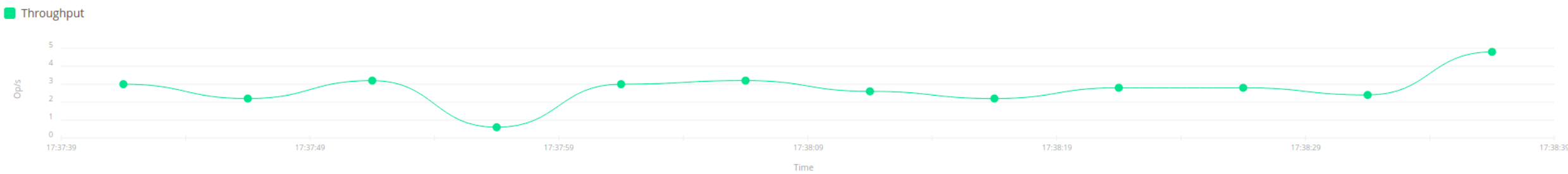
Minute | Hour | Day



FILE OPERATIONS QUEUE STATISTICS



IMPORT THROUGHPUT



DATA TRANSFERS

Replicate files on demand and on the fly.

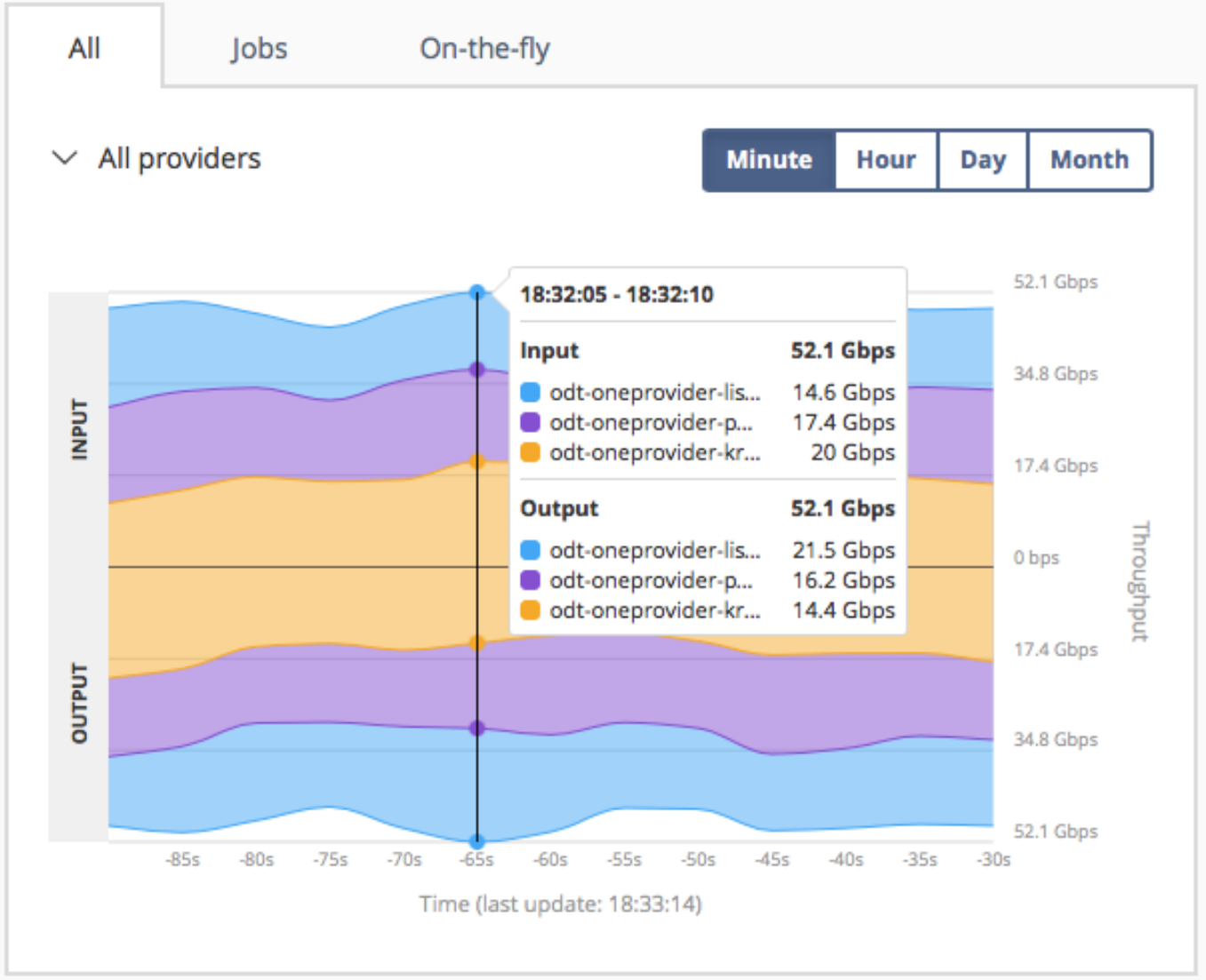
Migrate data between sites and storage backends on demand or with simple API interface.

Easily check location of your data using GUI or API.

Types of data transfers:

- replication — copying (only the missing) data to achieve a complete replica at the destination. The data is copied from one or more providers holding the missing blocks.
- eviction — removing replica(s) from the specified provider. This operation is safe and will succeed only if there exists at least one replica of every block on other supporting providers.
- migration — replication followed by eviction. Replicates the data to the destination provider and then evicts the replica from the source provider.

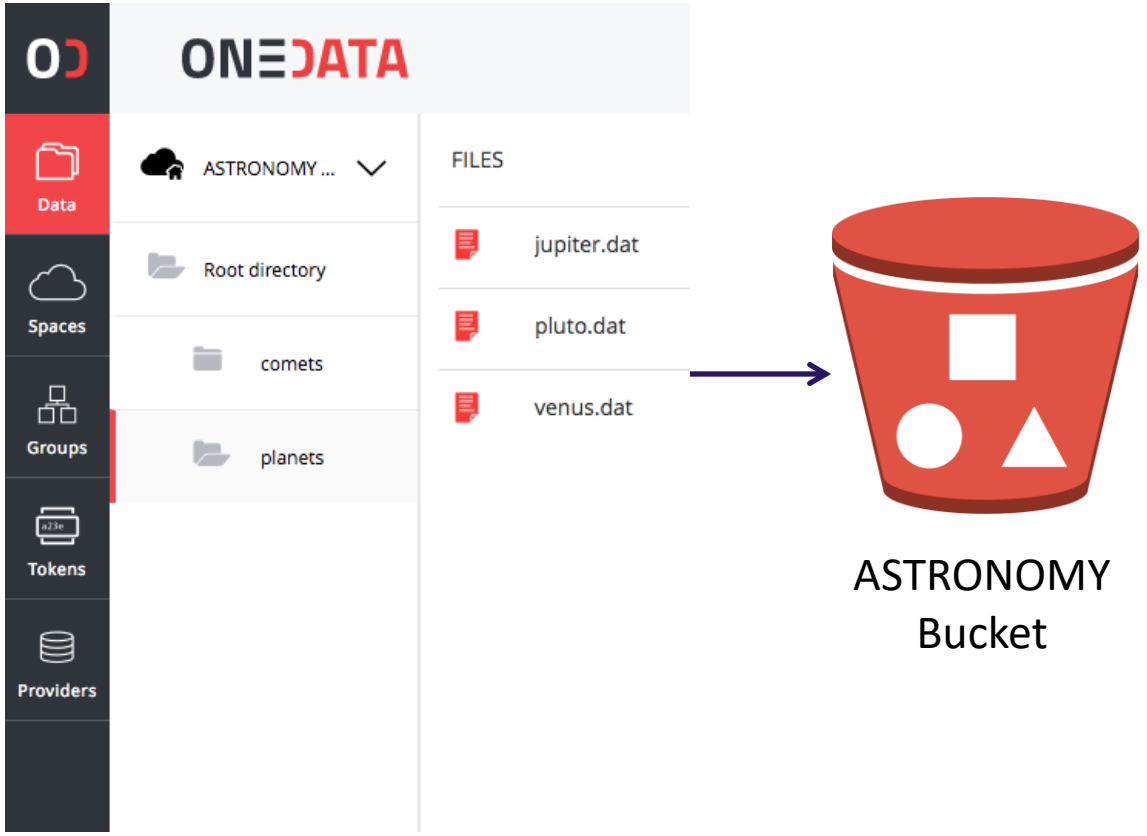
DATA TRANSFERS



ONE3S S3 API Adapter

OneS3 is a scalable S3 implementation:

- users can access their data spaces via AWS S3 interface,
- exposes storage resources managed by a specific Oneprovider,
- can be scaled to any number of nodes,
- it handles all data access to storage directly, while Oneprovider instances will only handle metadata management.



https://kronika.gov.pl

gov.pl | Serwis Rzeczpospolitej Polskiej

Krajowe Repozytorium Obiektów Nauki i Kultury

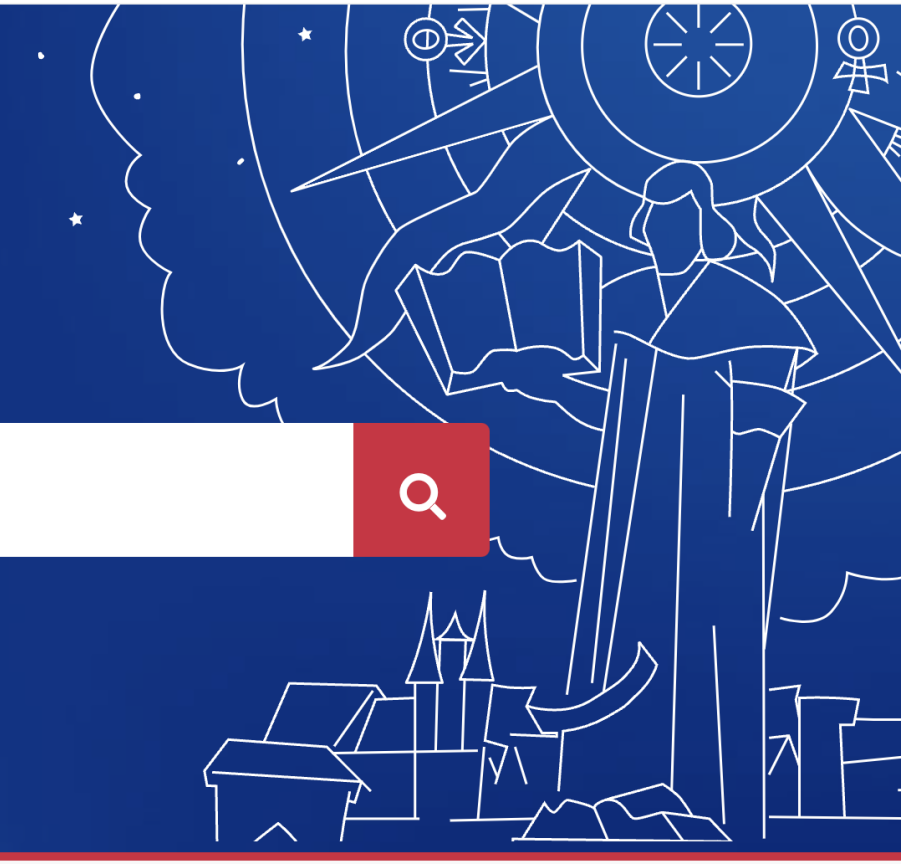
O portalu | Obiekty | Pomoc | Instytucje | Kontakt

Portal z otwartymi zasobami kultury i nauki

lampa

Szukaj z warstwą tekstową OCR

Szukaj tylko w tytułach



https://kronika.gov.pl

Zbiory Muzealne

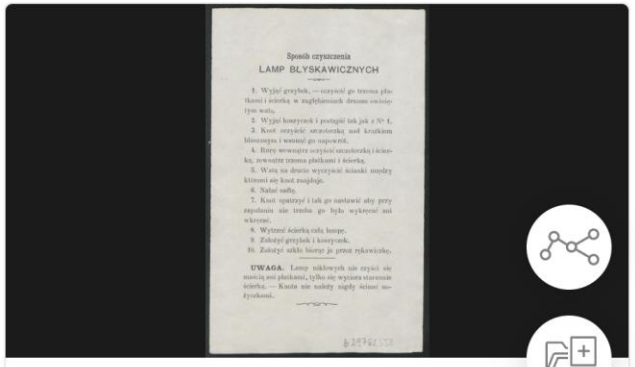
Zespół "Stroma 5" (Żygulska, Żątowska, Karbowska) autor

Lampa

Zbiory Muzealne

Zespół "Stroma 5" (Żygulska, Żątowska, Karbowska) autor

Lampa



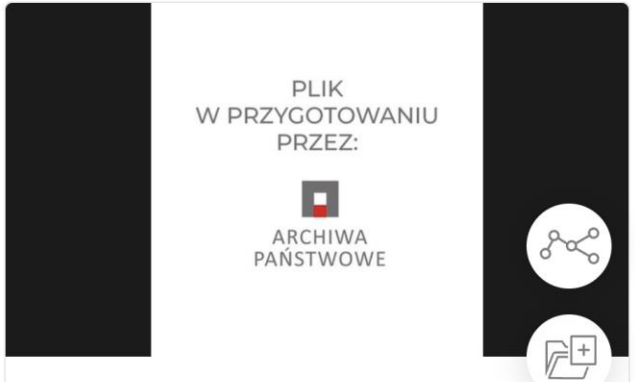
Druki ulotne

Sposób czyszczenia lamp zwyczajnych.

[193-]

Zbiory Muzealne

Zbiory Muzealne

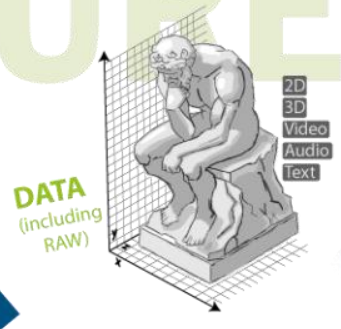


T — dokumentacja techniczna

FOURTH USECASE:

3D DIGITALIZATION AND PUBLICATION OF CULTURAL HERITAGE OBJECTS

CAPTURE CLOUD DELIVERY



DATA (including RAW)

METADATA EDM

```
Name:        Statue of Eureka3D
Description: From the XII Century, the...
Author:      Ignacio Lamata Martinez
```

DIGITISATION process

PARADATA

UPLOAD to the cloud

EUREKA3D platform

Authentication and Authorisation

Data Management

Compute power

Storage

Virtual Machines (servers)

ONEDATA

3D VISUALISATION

DOWNLOAD of data

statue.obj

EUROPEAN OPEN SCIENCE CLOUD

PROJECT REQUIRMENTS AND OBJECTIVE

Requirements:

- university friendly authentication and authorization
- simple data upload interface
- annotating data with metadata
- data versioning and archiving
- integrated pipeline for processing 3D models
- data publication as open data and integration with Europeana

Objectives:

- provide a web based 3D viewer for showcasing 3D objects
- publication of 3D objects to Europeana

ONEDATA

ONE WORLD

EGI DATAHUB ZONE

Sign in

with your identity provider

Onezone has been upgraded to version 21.02.1 on 18th May 2023.

Note for users: access to the PLAYGROUND space with a 30GB shared quota is given to all users via EGI Check-in.

Note for administrators: it is recommended to upgrade Oneprovider instances to version 21.02.1.

DataHub is an EGI service provided by CYFRONET, receiving funding from the EGI Foundation (EGI.eu) and the EGI-ACE project (Horizon 2020) under Grant number 101017567 | Powered by OneData.

Check-in

Choose your academic/social account

Search...

- 29 Mayis University
- A*STAR - Agency for Science, Technology and Research
- A. T. Still University
- AAF Virtual Home
- aai.lab.maeen.sa
- AAI@EduHr Single Sign-On Service
- Aalborg University
- Aalto University
- Aarhus School of Architecture
- Aarhus School of Marine and Technical Engineering
- Aarhus University
- AARNet
- Aba Teachers University
- Abertay University
- Aberystwyth University
- ABES - French Bibliographic Agency for Higher Education
- Abingdon and Witney College
- Absalon University College
- Academia Cotopaxi - cotopaxi
- ACADEMIA d.o.o.

or

Bitbucket | D4SCIENCE | EGI SSO | eduGAIN LOG-IN

B2ACCESS | Facebook | GitHub | Google | eduTEAMS

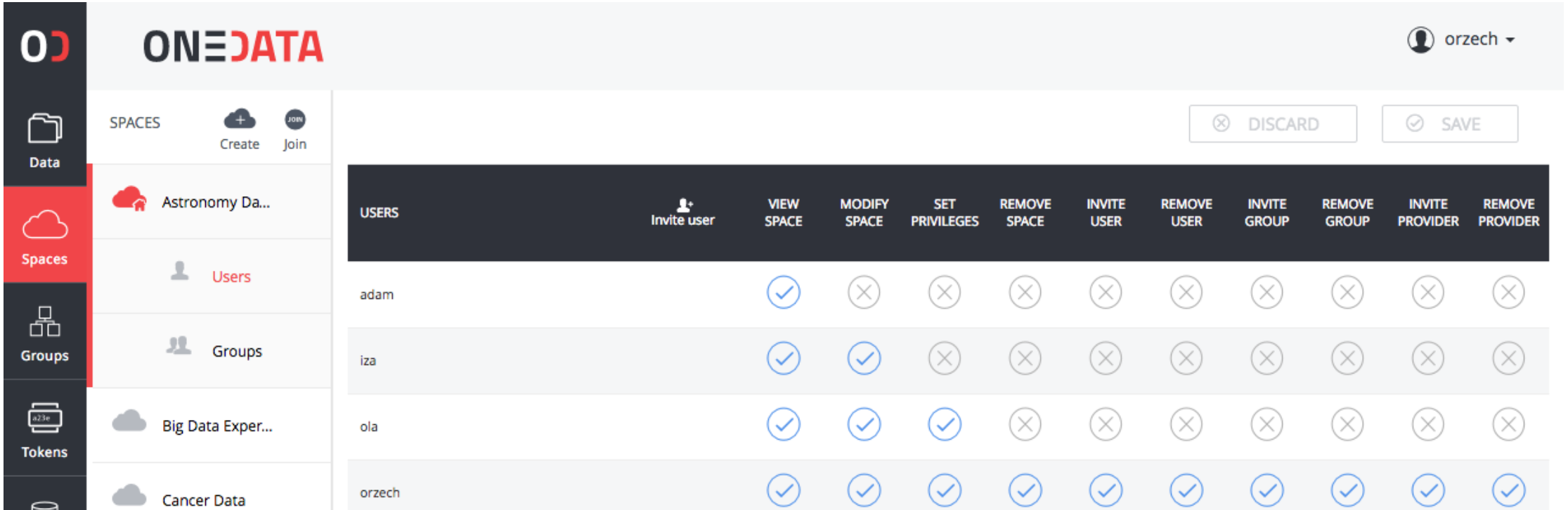
IGTF | LinkedIn | ORCID | umbrella | WeChat

eria

Can't find your identity provider?

FLEXIBLE AUTHENTICATION AND AUTHORIZATION

- Pluggable methods of authentication per zone (eg. OAuth 2.0)
- Multiple levels of access control
- ACL on files and directories
- Group management
- Token based authentication (macaroons)
- X.509 certificates



INTEGRATED METADATA MANAGEMENT

Filesystem attributes — built-in filesystem metadata such as file size, creation and modification timestamps, POSIX permissions, etc. all files and directories can have a custom user metadata

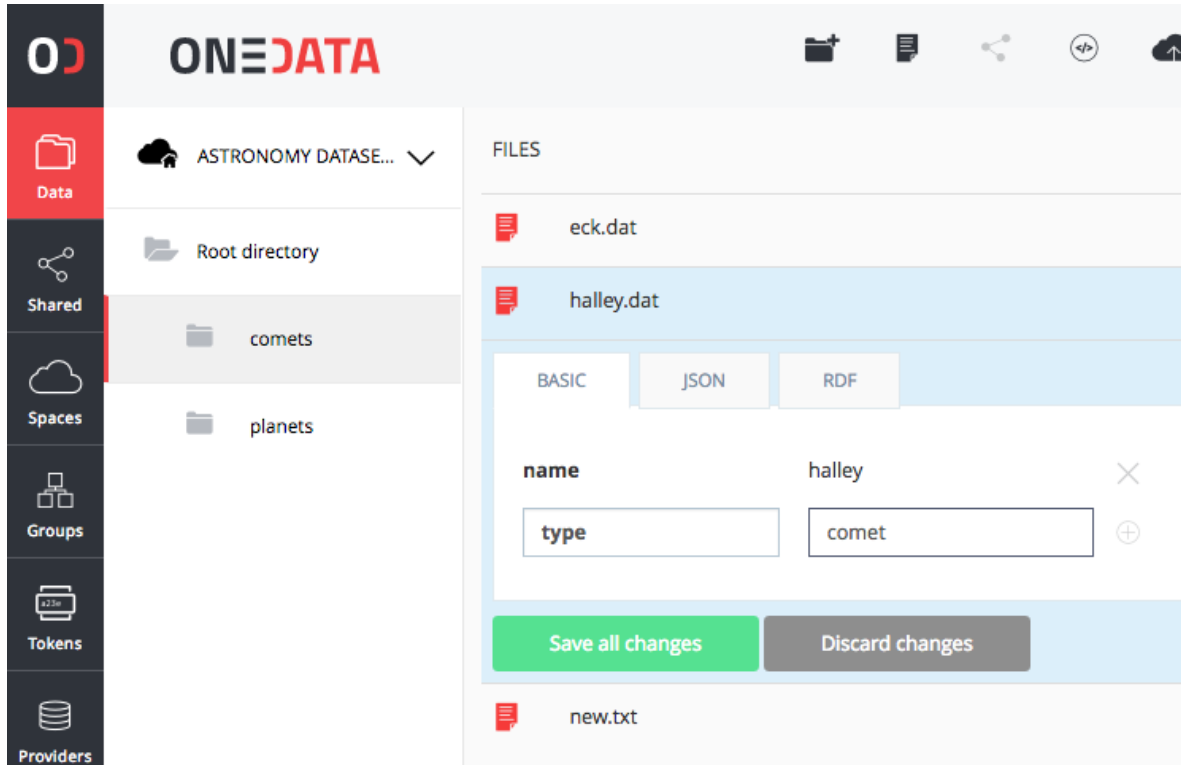
Custom metadata — arbitrary, user-defined information:

- extended attributes — a.k.a. basic attributes or xattrs; key-value pairs, compatible with POSIX extended attributes,
- JSON document,
- RDF document (XML).

All types of custom metadata can **coexist** for the same file.

API for metadata management.

Data discovery based on metadata.



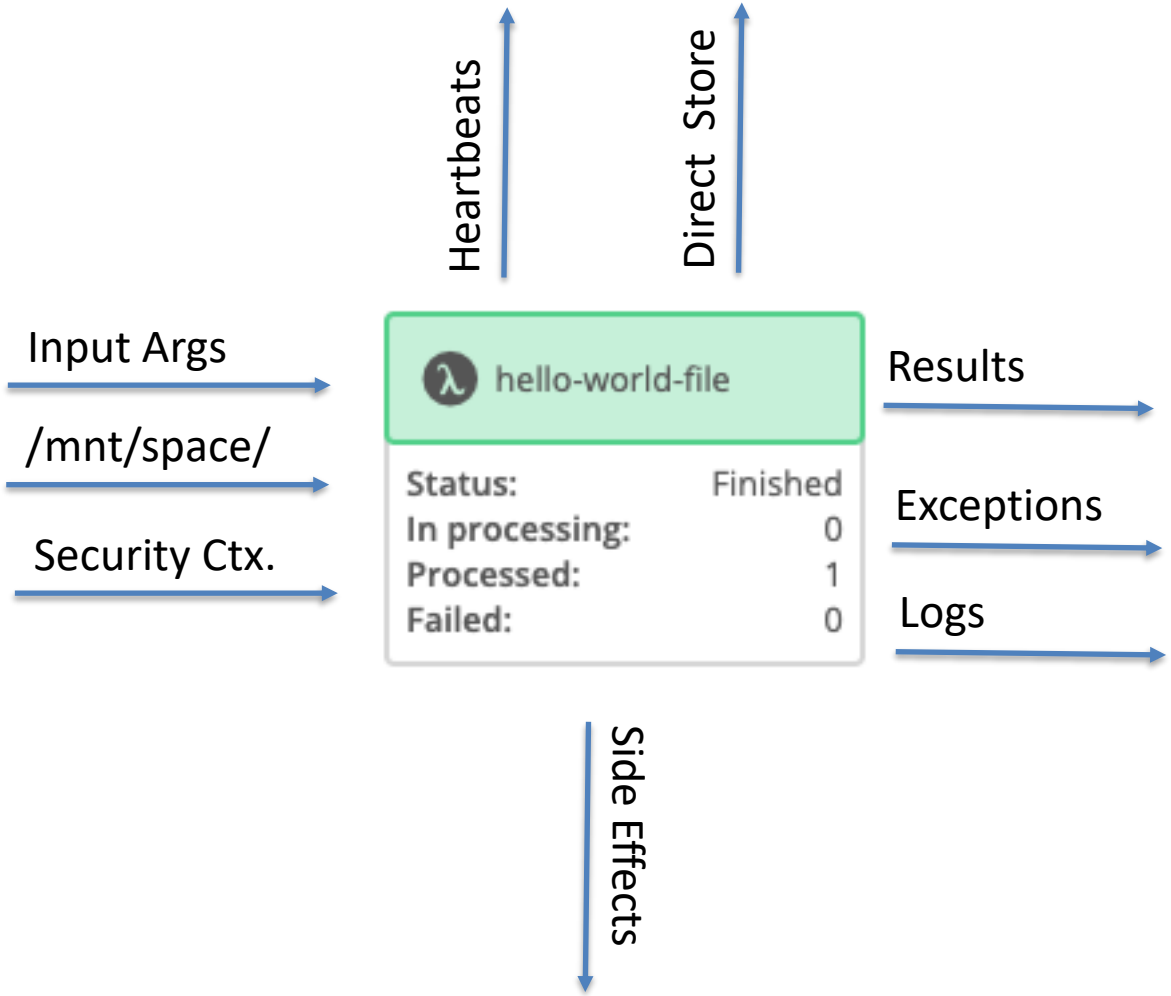
Automation Engine

FaaS Lambda

Lambda Anatomy

- **Input Arguments.** <Map>
- **Mount Space as File system.** <Oneclient> optional
- **Output Results**. <Map>
- **Exceptions**. <Map>
- **Logs**. <Map>
- **Side-effects.** e.g. REST-API calls
- **Heartbeats.** For long running lambdas
- **Stores Updates.** Direct operations on stores

- **Batch Mode. Can work with batches of input arguments to speed up the process**

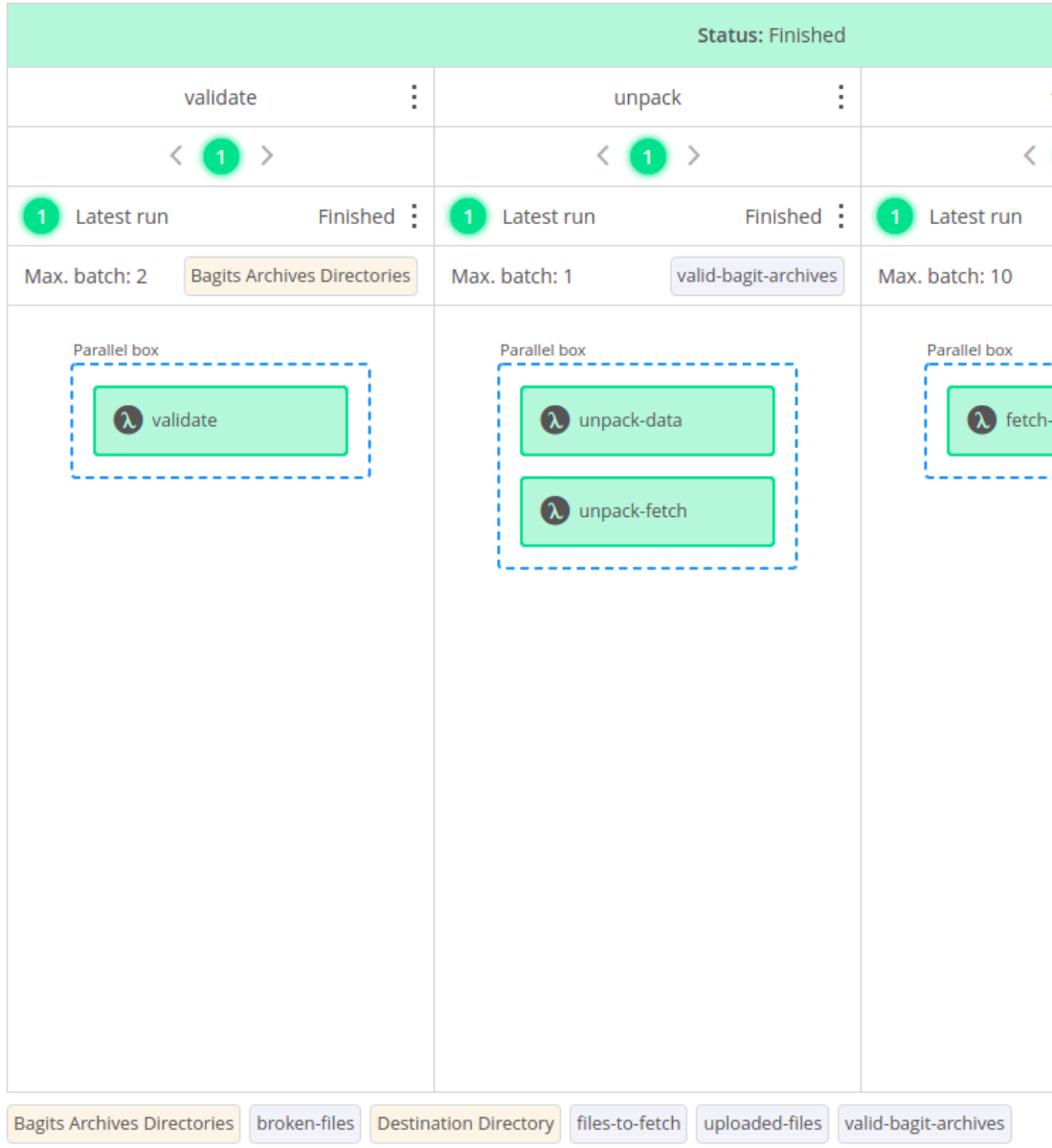


WORKFLOW

Workflow Anatomy

- **Lanes.** Iterates over Store and execute parallel boxes
- **Stores.** Input to to the workflow or produced during the workflow
- **Parallel Boxes.** Contains Lambdas which can be executed in any order
- **Lambdas**. Function which is called by mapping arguments

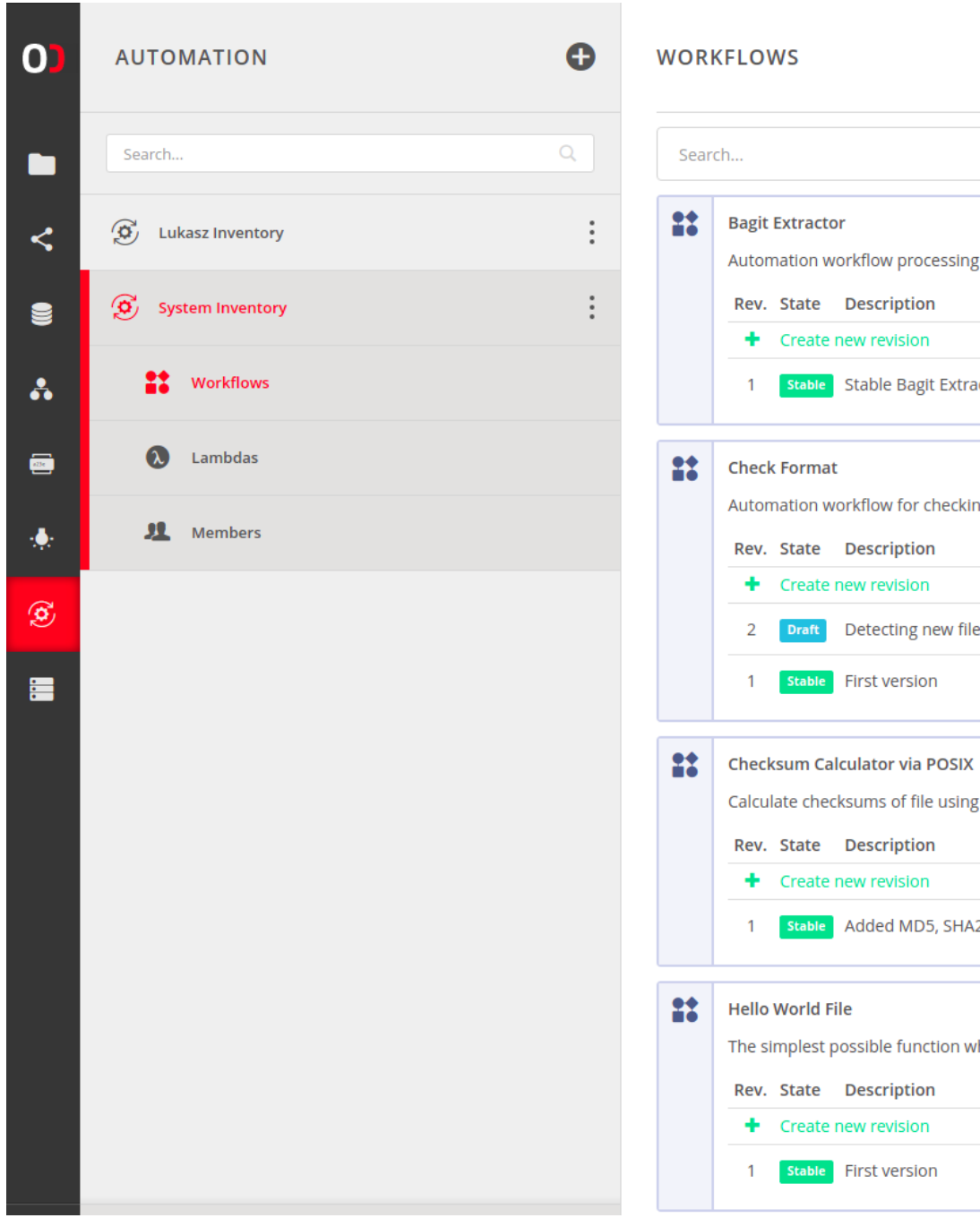
- Can be exported to JSON and reused by someone else



INVENTORY

Inventory Anatomy

- **Workflows. Keep the list of workflows to be available for system users**
- **Lambdas.** Keep the list of registered Lambdas
- **Members. A**ccess control
- **Import/Export. I**mport full workflows into Inventory from JSON file



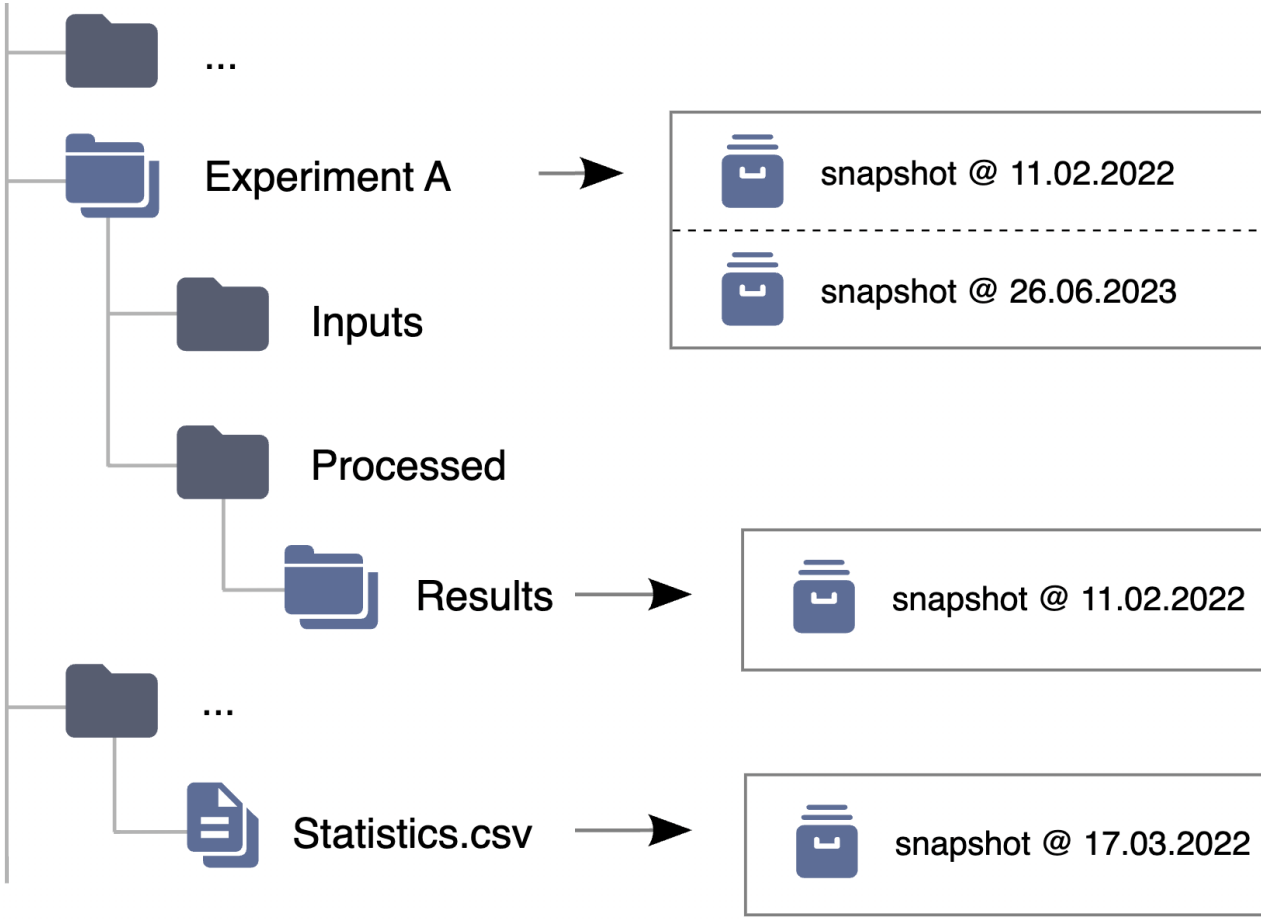
SNAPSHOTS – TOWARDS OPEN DATA

Datasets are essentially files or directories that have been marked by Onedata users as representing data **collections** relevant to them.

Datasets offer additional features, compared to regular files and directories:

- optional data and **metadata protection**,
- dataset structure tracking using the dataset browser,
- ability to create **persistent snapshots** of the physical dataset contents — archives.

Datasets can be nested to compose **hierarchical structures**

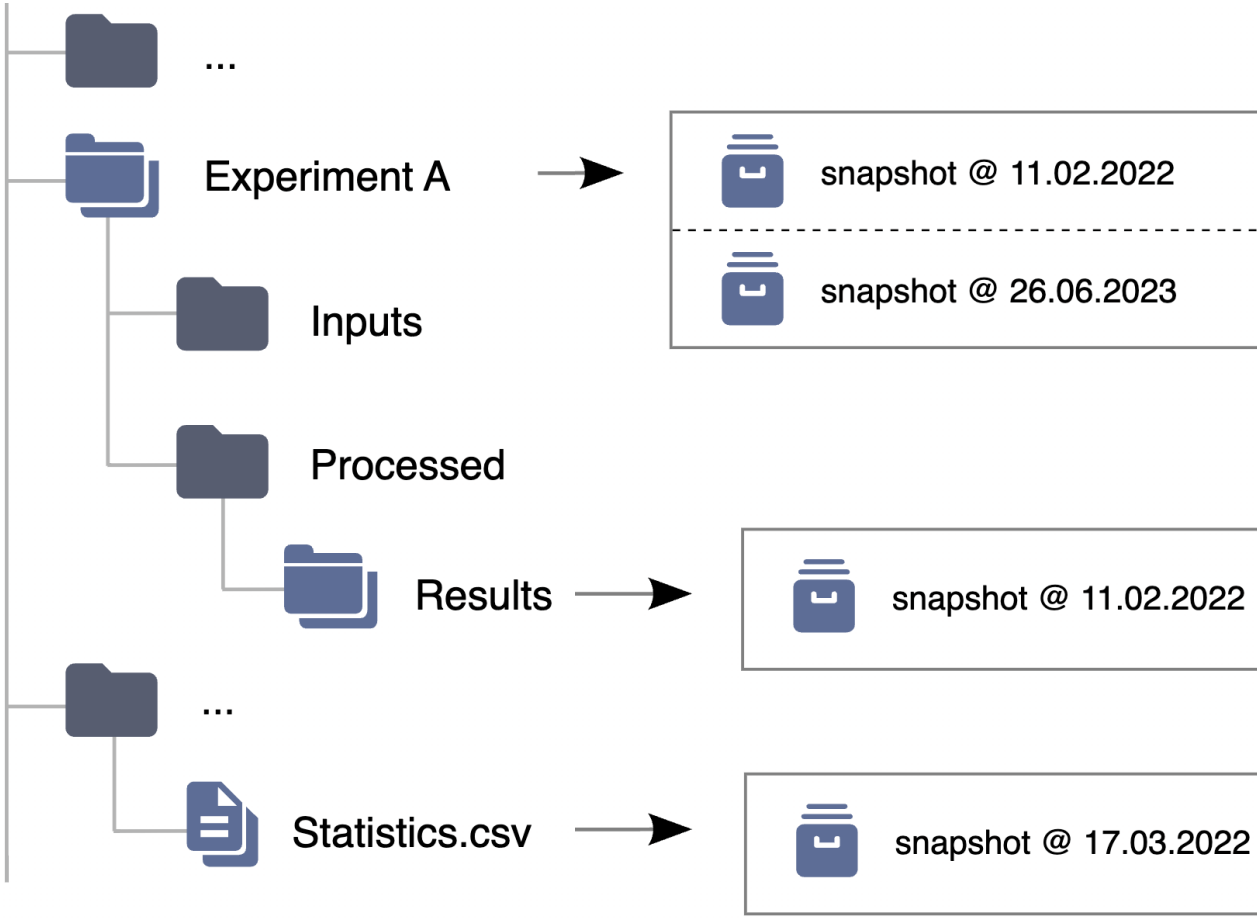


ARCHIVES – TOWARDS OPEN DATA

A snapshot of a dataset created at a certain point in time is called an archive.

The archive creation process comes with several options:

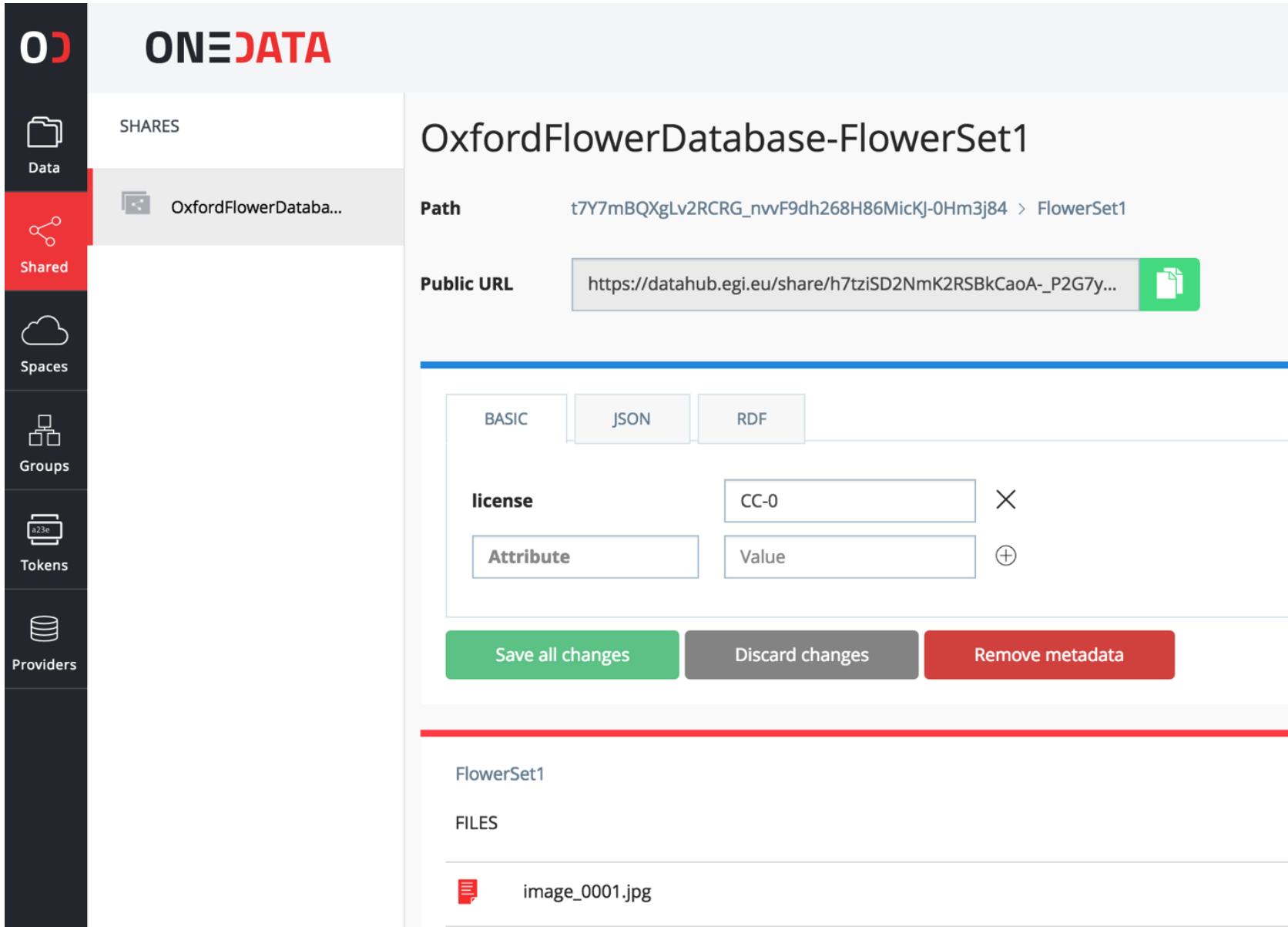
- different layouts — the structure of the files in the created archive; support for BagIt,
- creation of nested archives — hierarchically-created archives for datasets with nested structure,
- incremental archives — reusing the unchanged files between snapshots,
- possibility to follow symbolic links if they are present in the dataset



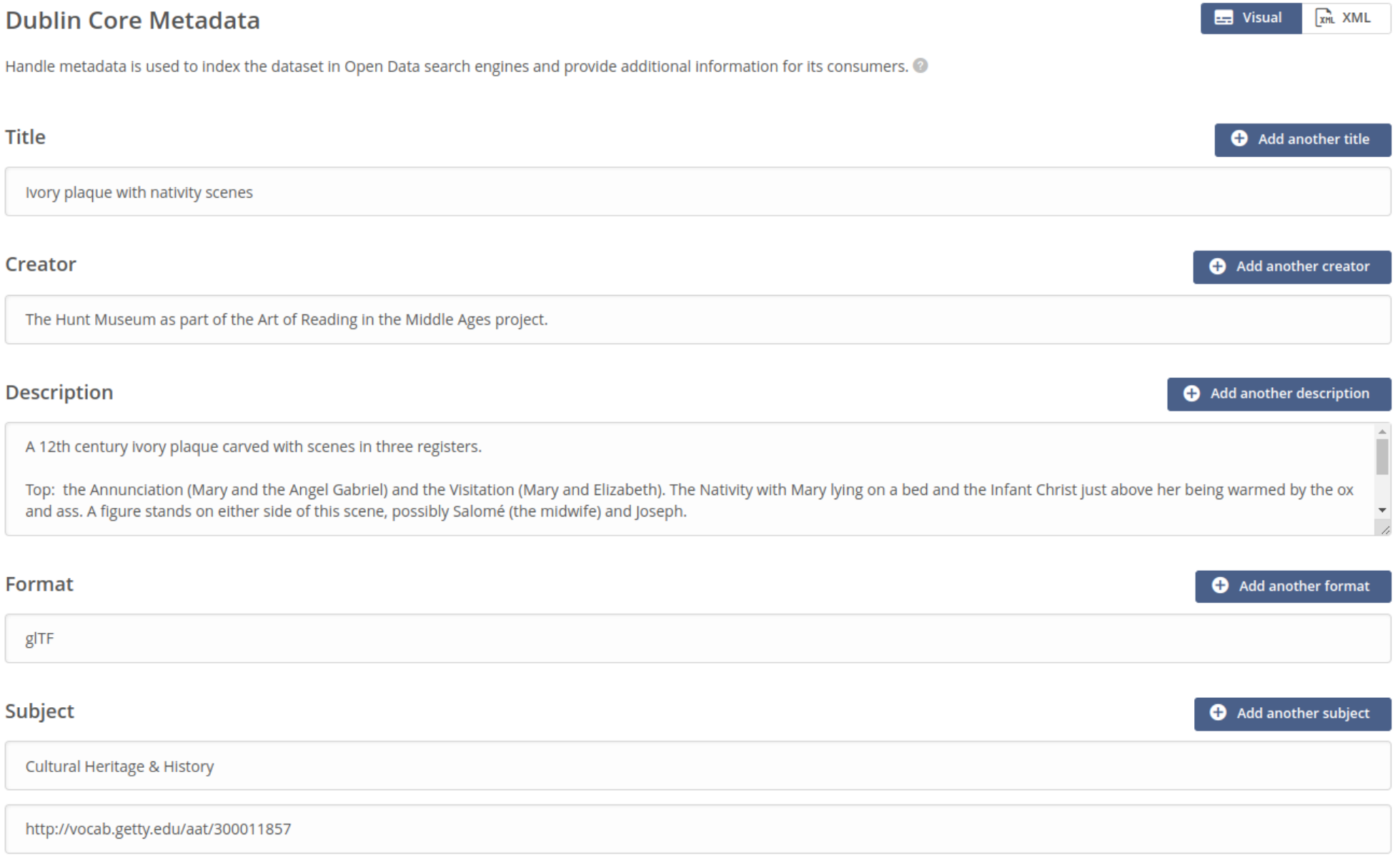
DATA SHARES – TOWARDS OPEN DATA

Files in Onedata can be made public and available for unauthenticated users.

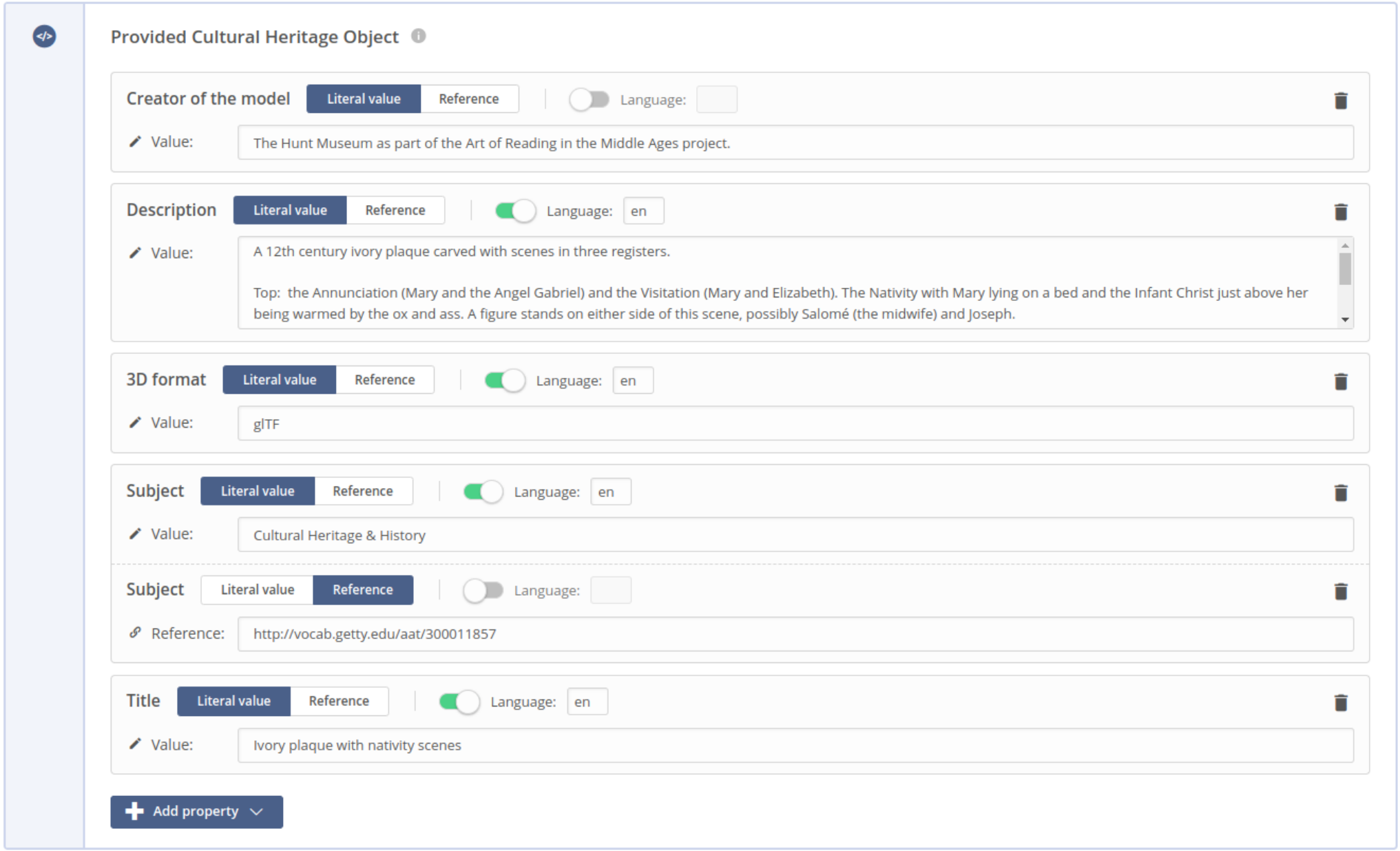
A single file or directory can be shared using multiple links each can be removed at any time, giving the possibility to manage public access



PUBLICATION AS DUBLIN CORE



PUBLICATION AS EUROPEANA DATA MODEL



Lamp

Een grote gloeilamp en drie kleinere peertjes op een metalen plaatje. De grote lamp is bevestigd aan een verticale staaf met een stelschroef, om haar in hoogte te kunnen verstellen. Uit het glas van de grote lamp ontbreekt een stuk.

Read more

Good to know | **All metadata**

Subject	**03.06.0 Lamp ; Nernst**
Type of item	**Lamp**
Date	**1900 - 1915 ; 1900/1915**
Providing institution	**Rijksmuseum Boerhaave**

EXTRA RESOURCES

Improved **documentation** (in making) https://onedata.org/#/home/documentation

Dedicated **demo mode** for easy sandbox deployment:

- https://onedata.org/#/home/documentation/21.02/admin-guide/demo-mode.html

Extensive **training materials** (4 day workshop!) covering majority of Onedata:

- https://onedata.org/training

  user: training

  password: Oneworkshop58

THANK YOU!

If any of this relates to your use cases, contact us!

Let’s do some meaningful science together!

morzech@agh.edu.pl

info@onedata.org