







# **Storing Data After Processing**

2023-06-21

Anni Järvenpää



CSC – Finnish expertise in ICT for research, education and public administration

## CSC

### **In This Session**

- Data management options after analysis has been completed:
  - Download to your own computer
  - Store in Allas
  - Bonus: Fairdata services
  - Bonus: Digital preservation service

# 656

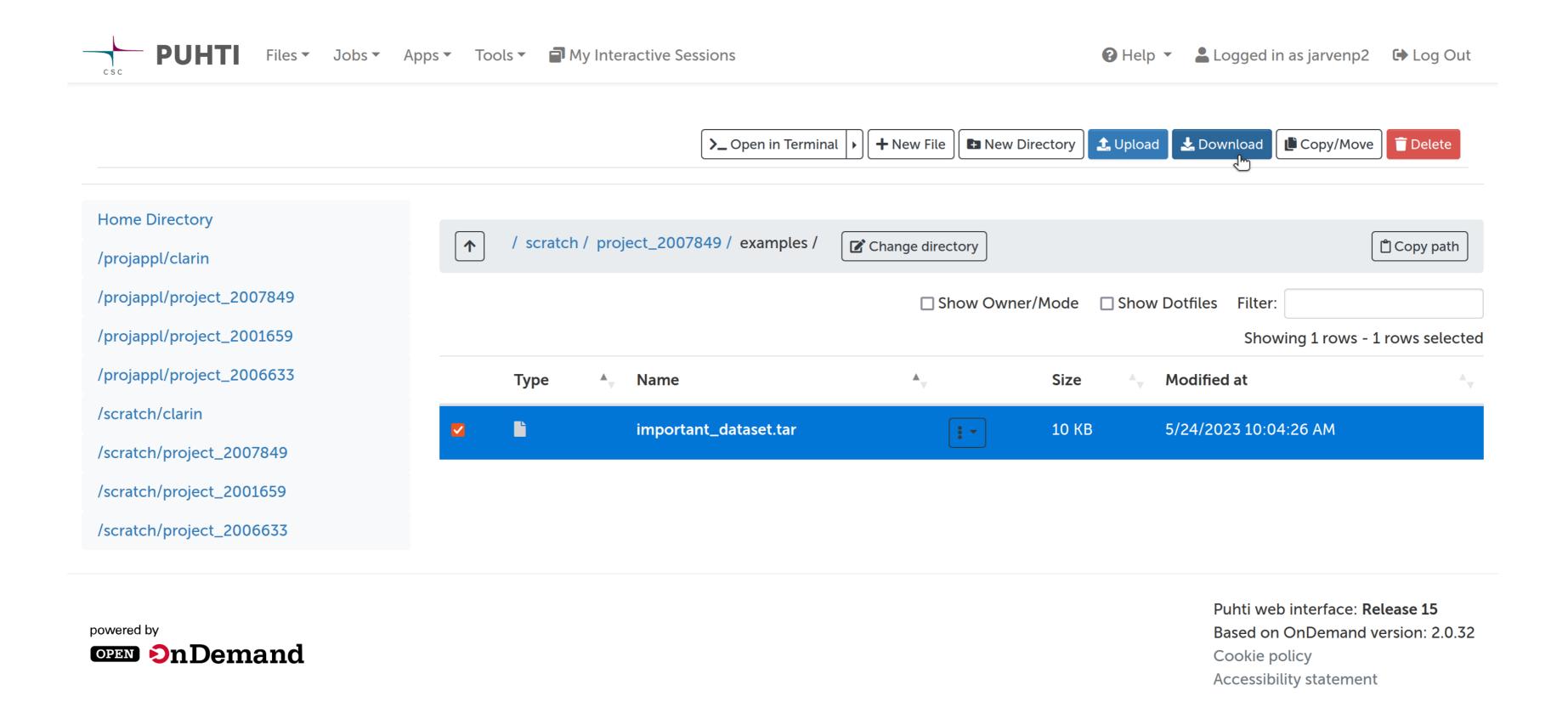
### Motivation

If you don't actively manage your data, it will:

- not benefit others
- not bring you citations
- not exist at all after a while



## **Downloading Data from Puhti: Web Interface**





## Downloading Data from Puhti: Web Interface

### Pros:

- Easy
- No need to apply for resources
- Does not require a valid project/account in the long term
- Negligible long term cost

- Less suitable for huge data sets
- Sharing is troublesome
- Data loss can still easily happen (e.g. lost, damaged, corrupted or obsolete media)



### **Allas**

### Pros:

- Someone else takes care of hosting the data
- Sharing the data with specific people is easy
- Opening the data for everyone is easy
- Can be operated via command line (e.g. on Puhti or a linux machine) or web interface

- Requires an active project
- Not long term (years, not tens of years)
- Limit for number of objects within a bucket

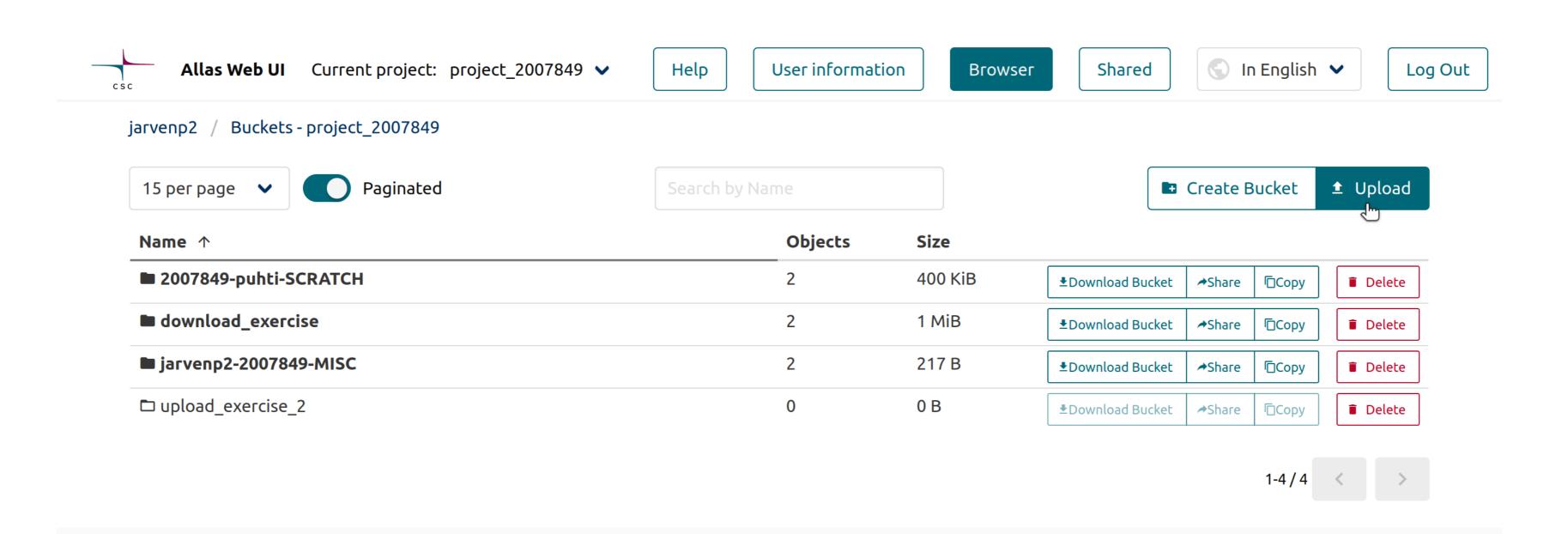
# CSC

## Allas: Object Storage

- Not directories and files but buckets and objects
- Very small difference for a casual user

# 666

## Allas Upload: Web UI



Allas Web UI Developed by CSC - IT Center For Science LTD



## Allas Upload: Directly from Puhti

- 1. Demonstration
  - Uploading files
  - Checking what has been uploaded
  - Downloading via command line
- 2. Try it yourself



### **Bonus: Fairdata Services**

### Pros:

- No running billing unit cost for storage space
- Sharing the data is easy
- Ready-made tooling for providing metadata together with the data
- Data is findable
- Accidental removal or alteration of data is not likely

- Requires an active project (but there are processes for transferring project managership)
- Not for data sets containing special categories of personal data



## **Bonus: National Digital Preservation Service**

### Pros:

Very reliable

- Not designed for sharing
- Costly and thus available only for selected data sets
- Limited file formats and strict submission processes