

Storing Data After Processing

2023-06-21

Anni Järvenpää



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

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

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

[>_ Open in Terminal](#) [+ New File](#) [New Directory](#) [Upload](#) [Download](#) [Copy/Move](#) [Delete](#)

Home Directory

[/projappl/clarin](#)

[/projappl/project_2007849](#)

[/projappl/project_2001659](#)

[/projappl/project_2006633](#)

[/scratch/clarin](#)

[/scratch/project_2007849](#)



[/scratch/project_2001659](#)

[/scratch/project_2006633](#)

[↑](#) [/ scratch / project_2007849 / examples /](#) [Change directory](#) [Copy path](#)

Show Owner/Mode Show Dotfiles Filter:

Showing 1 rows - 1 rows selected

Type	Name	Size	Modified at
<input checked="" type="checkbox"/> 	important_dataset.tar 	10 KB	5/24/2023 10:04:26 AM

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

Cons:

- 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](#)

Cons:

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

Allas: Object Storage

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



Allas Upload: Web UI



Allas Web UI

Current project: project_2007849

Help

User information

Browser

Shared

In English

Log Out

jarvenp2 / Buckets - project_2007849

15 per page

Paginated

Search by Name

Create Bucket

Upload

Name ↑

Objects

Size

2007849-puhti-SCRATCH

2

400 KiB

Download Bucket

Share

Copy

Delete

download_exercise

2

1 MiB

Download Bucket

Share

Copy

Delete

jarvenp2-2007849-MISC

2

217 B

Download Bucket

Share

Copy

Delete

upload_exercise_2

0

0 B

Download Bucket

Share

Copy

Delete

1-4 / 4



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

Cons:

- 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

Cons:

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

