

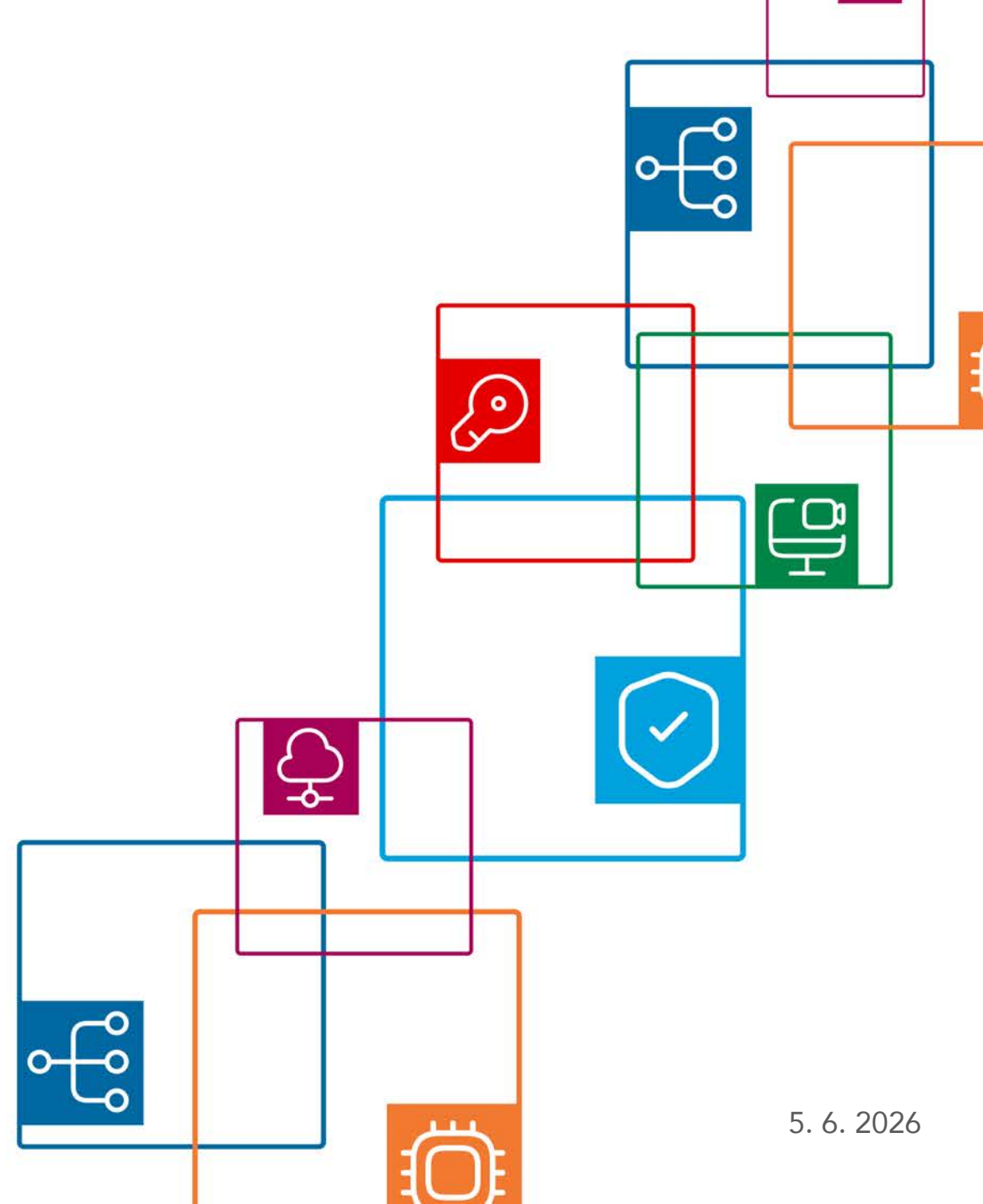
MetaCentrum NGI

Objevte možnosti MetaCentra
Ukázka praktického využití

Jiří Vorel

MUNI, Brno

vorel@cesnet.cz meta@cesnet.cz



■ MetaCentrum is

- The activity of the CESNET association
- Part of the e-INFRA CZ e-infrastructure
- A national grid infrastructure (NGI)
- A provider of computational resources, application software (both commercial and free/open source) and data storage
- Available to the entire academic and research community (employees and students) in the Czech Republic
- Free of charge (users 'pay' by acknowledging the project 'e-INFRA CZ ID:90254' in their publications)

<https://www.cesnet.cz/>

<https://www.metacentrum.cz/>

<https://www.e-infra.cz/>

<https://docs.metacentrum.cz/>



■ MetaCentrum offers

- Access to all resources without submitted projects (with one exception)
- The possibility to apply for membership at any time
- Immediate access to hardware resources and licences (once an application has been approved)
- CPU/GPU resources, CLI access, GUI applications, Kubernetes and cloud services, etc.
- Various application software (commercial, free and open source)
- Data sharing
- Access is also possible for partners from industry and abroad



■ MetaCentrum cannot help with

- Purely commercial research
- Allocating resources for a fee
- Long-term and secure storage of valuable data (i.e. data archiving) <https://du.cesnet.cz>
- Processing sensitive data
(Kubernetes Sensitive Cloud within e-INFRA CZ,
coordinated by CERIT-SC) <https://www.cerit-sc.cz/>
- Tools whose installation or use would violate the licence terms
- Windows OS

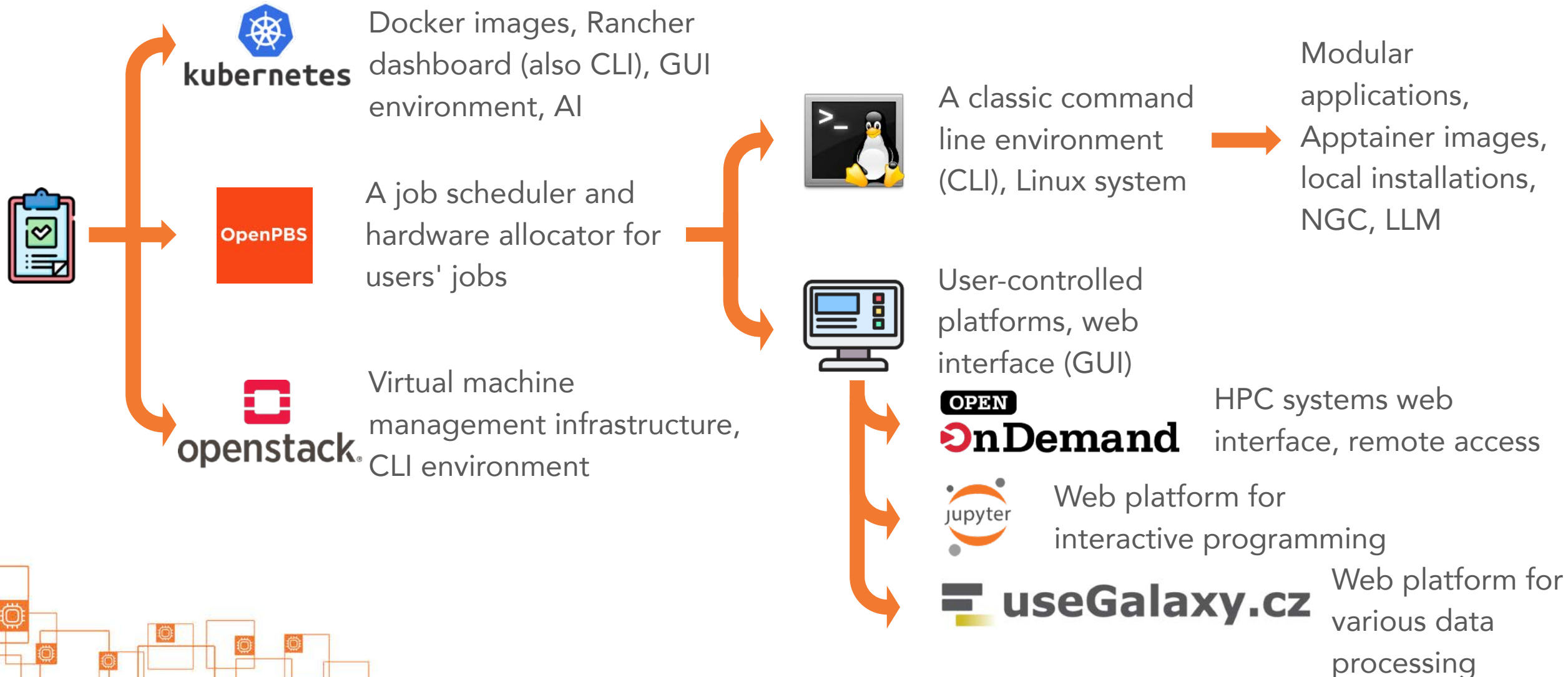


A close-up, high-angle shot of the One Ring, a golden band with intricate Elvish script engraved on its surface. The ring is positioned diagonally across the frame, resting on a textured, golden-brown fabric that features a faint, embossed map of Middle-earth. The lighting is dramatic, highlighting the metallic sheen of the ring and the texture of the fabric. The text "ONE APPLICATION TO RULE THEM ALL" is overlaid in white, serif capital letters across the center of the ring.

ONE APPLICATION TO RULE THEM ALL

One application to rule them all

<https://docs.metacentrum.cz/en/docs/welcome>



<https://docs.metacentrum.cz/en/docs/access/account>

<https://docs.metacentrum.cz/en/docs/access/terms>

Submitting an application

- Anytime during the year, manual approval
- Log in through the institution (i.e. local name and password) from the eduID federation
- Accounts expire on 2 February each year and must be renewed in January
- Immediate access to hardware resources

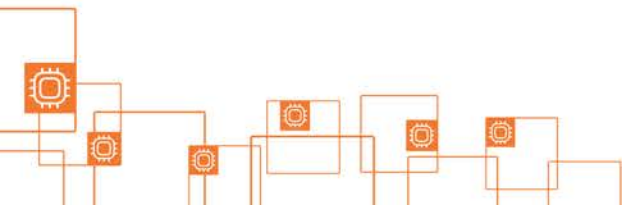
eduID.cz cesnet

I have an account in a member organisation of eduID.cz

My organisation is not in eduID.cz and I need to validate my alternative identity

Direct links to selected institutions from eduID.cz

MUNI Masarykova univerzita	Univerzita Karlova	Západočeská univerzita v Plzni
Jihočeská univerzita v Českých Budějovicích	Univerzita Palackého v Olomouci	Univerzita Pardubice
CESNET	České vysoké učení technické v Praze	Mendelova univerzita v Brně
Technická univerzita Liberec	Vysoká škola báňská - Technická univerzita Ostrava	Vysoké učení technické v Brně



Link your account to ORCID

- Link your e-INFRA CZ account to ORCID
 - Users' MetaCentrum publications are currently invisible to funders
 - New automated system will harvest publications via ORCID
 - **Action needed:** link your ORCID iD in your e-INFRA profile + use correct acknowledgement
 - This can easily be done in the e-INFRA user profile



<https://www.e-infra.cz/en/news/link-your-e-infra-cz-account-to-orcid-a-small-step-for-smarter-impact-reporting>





MetaCentrum provides access to hundreds of software tools giving thousands of individual modules

Licence models

- We can run almost anything that doesn't violate the licence terms and is for Linux
- Most application software is open source, but...

<https://docs.metacentrum.cz/en/docs/software>

Paid licence



For registered individuals

For all users

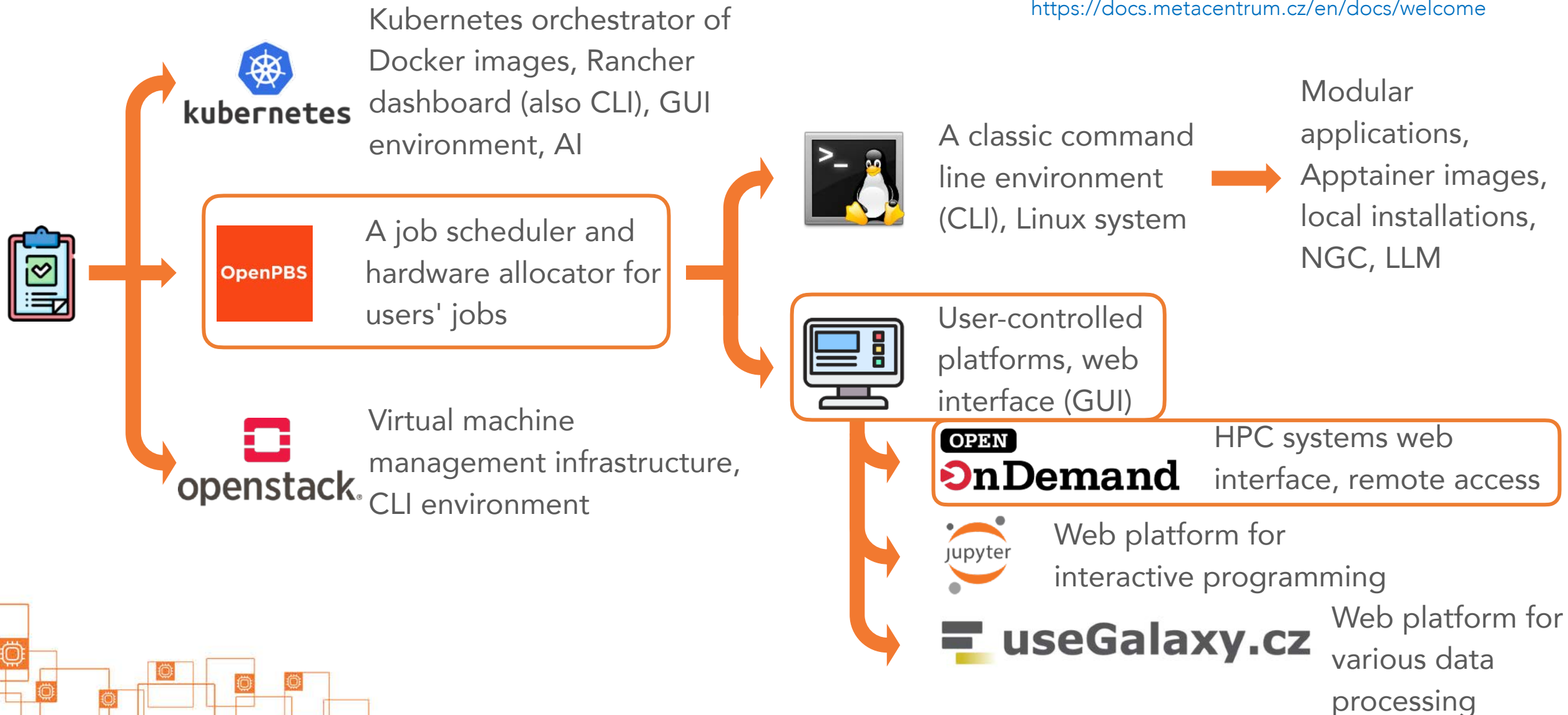


Free licence



How to use it

<https://docs.metacentrum.cz/en/docs/welcome>

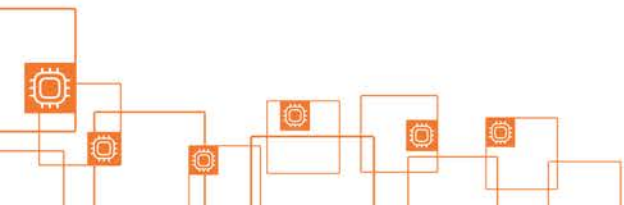
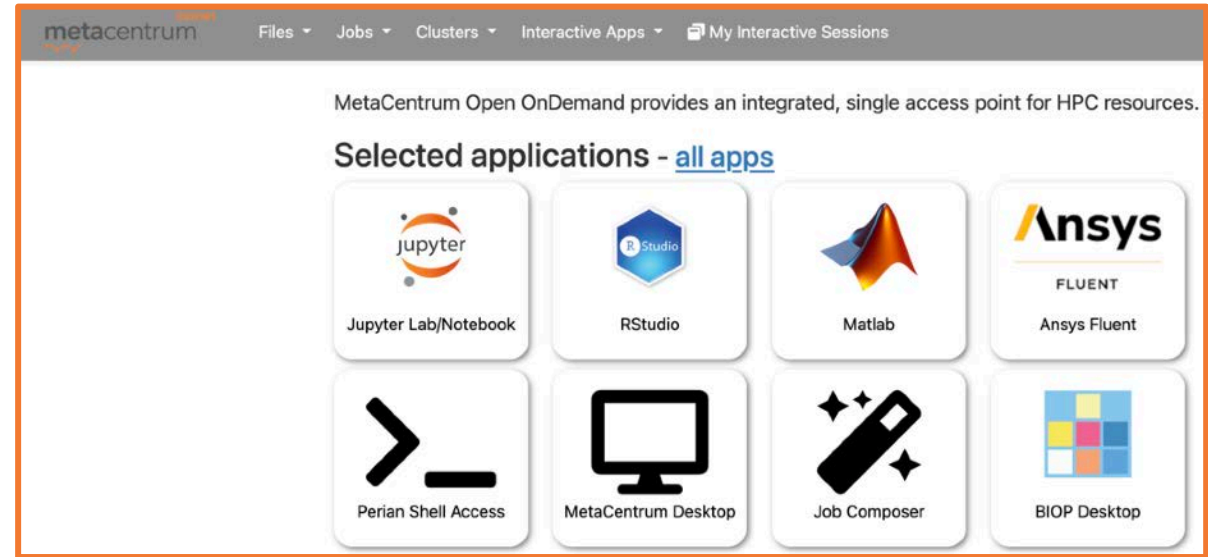


■ Open OnDemand



<https://ondemand.metacentrum.cz/>

- Web application on top of OpenPBS scheduler
- Running (not only) GUI applications

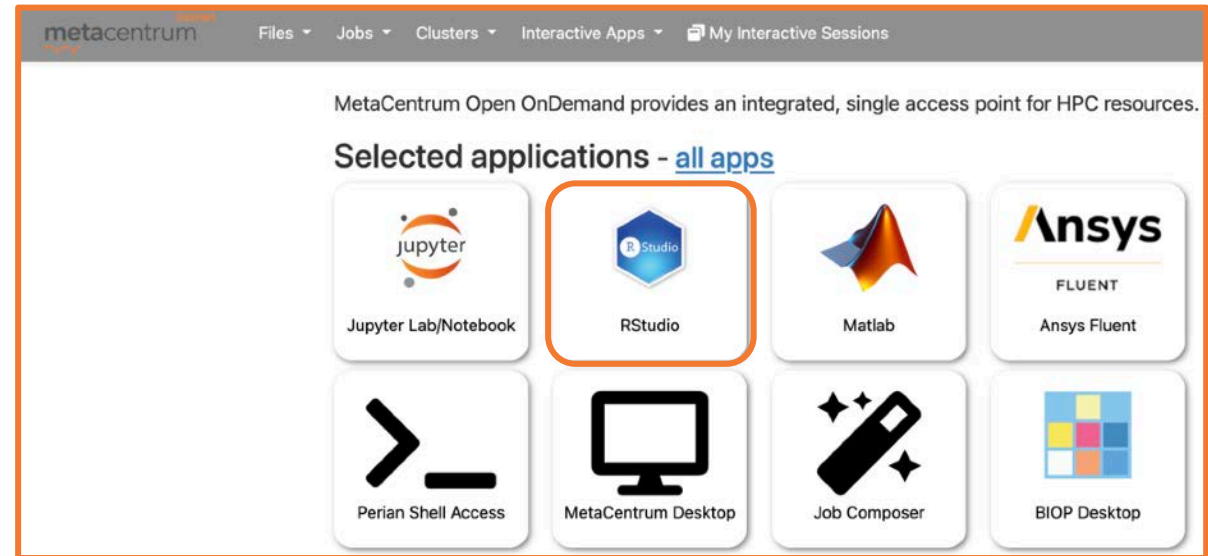


■ Open OnDemand



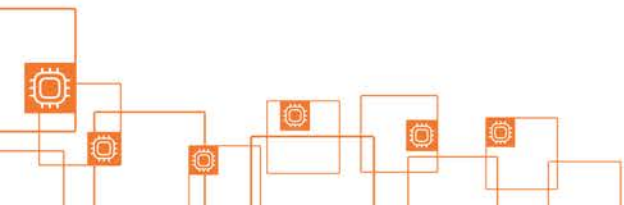
<https://ondemand.metacentrum.cz/>

- Web application on top of OpenPBS scheduler
- Running (not only) GUI applications



<https://docs.metacentrum.cz/en/docs/software/sw-list/rstudio>

<https://docs.metacentrum.cz/en/docs/software/sw-list/r>





R is a software environment for statistical computing and graphics

■ Open OnDemand



<https://ondemand.metacentrum.cz/>

- Web application on top of OpenPBS scheduler
- Running (not only) GUI applications

Home / My Interactive Sessions / RStudio Server

Interactive Apps

Cloud

- Kubernetes on OpenStack (KOST) (beta)
- OpenStack VM (beta)

Desktops

- Ansys/Enight
- Ansys/Fluent
- Ansys/Workbench
- BIOP Desktop
- CLCgenomicsWB
- MZMINE
- Matlab
- MetaCentrum Desktop
- QuPath
- VMD Desktop

Servers

- Julia Pluto.jl
- Jupyter - Evo2 (beta)
- Jupyter Notebook/Lab
- Matlab webapp (beta)
- RStudio Server**

RStudio Server

This app will launch an RStudio server on one or more nodes. Geospatial and Tensorflow packages are preinstalled.

Number of hours

2

Number of CPUs on single node

2

Memory (GB)

10

GPUs

0

Scratch local (GB)

10

RStudio Image version

RStudio-geospatial-4.5.1

RStudio working directory location:

/storage/brno2

Launch

* The RStudio Server session data for this session can be accessed under the [data root directory](#).

```
File Edit Code View Plots Session Build Debug Profile Tools Help
R 4.5.1 · /auto/brno2/home/vorel/
Go to file/function
Addins

Console Terminal Background Jobs
R 4.5.1 (2025-06-13) -- "Great Square Root"
Copyright (C) 2025 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

vorel | Project: (None)

Environment History Connections Tutorial

Import Dataset 187 MiB

R Global Environment

Environment is empty

Files Plots Packages Help Viewer Presentation

Home Find in Topic

R Resources

- [Learning R Online](#)
- [CRAN Task Views](#)
- [R on StackOverflow](#)
- [Getting Help with R](#)

Manuals

- [An Introduction to R](#)
- [Writing R Extensions](#)
- [R Data Import/Export](#)

Reference

- [Packages](#)

Miscellaneous Material

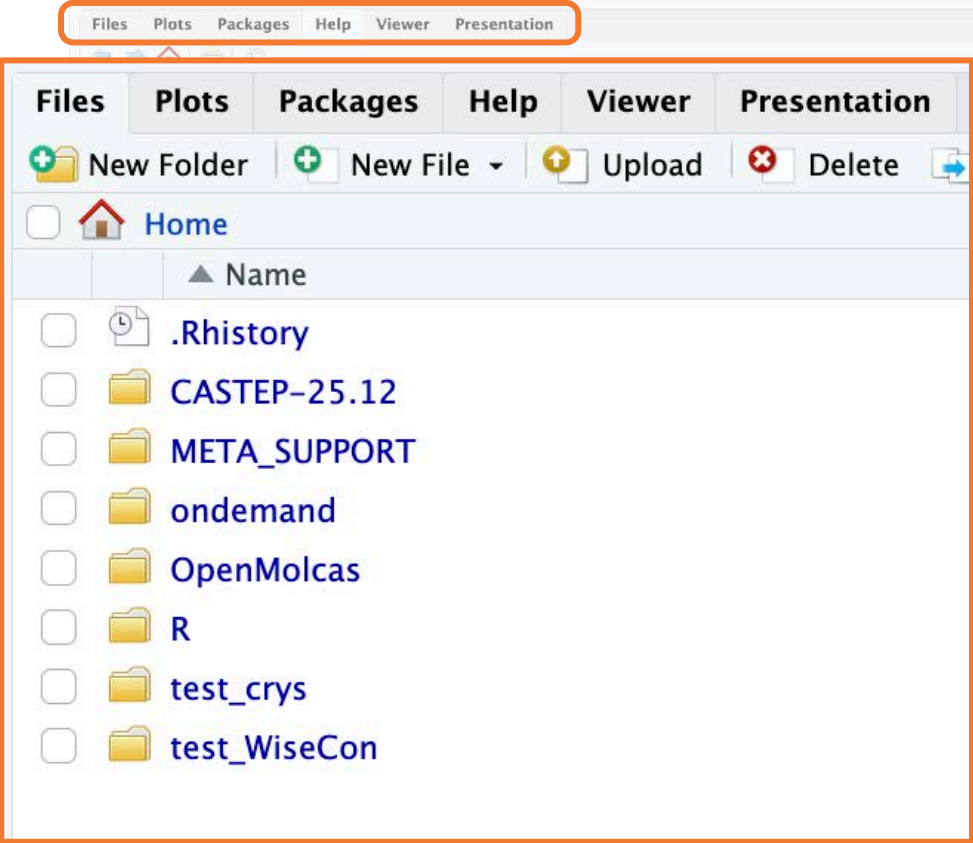
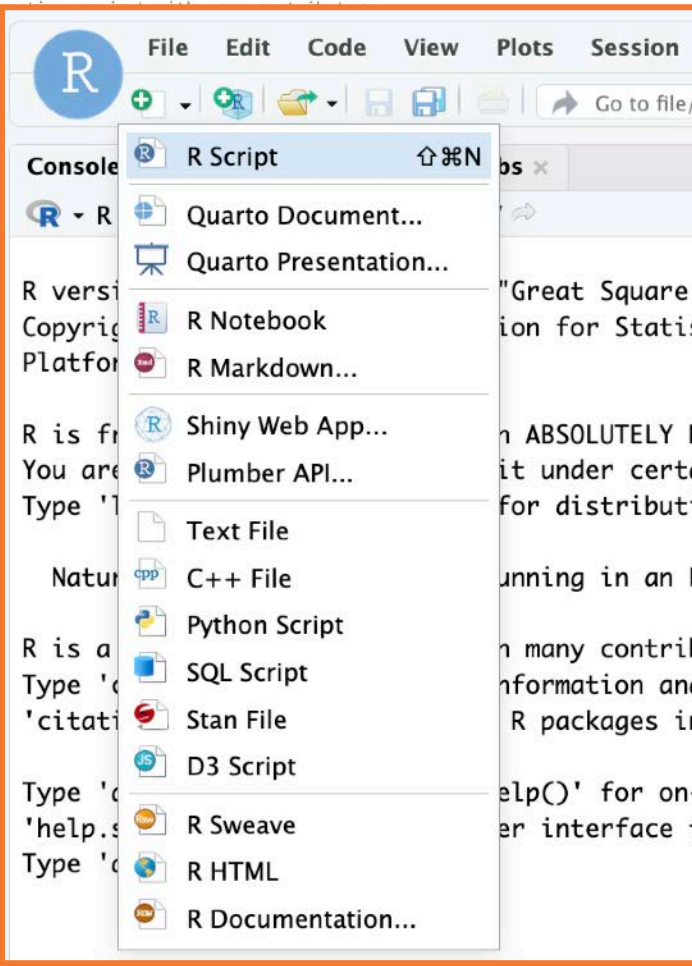
- [About R](#)
- [License](#)
- [NEWS](#)
- [Authors](#)
- [FAQ](#)
- [User Manuals](#)
- [Resources](#)
- [Thanks](#)
- [Technical papers](#)

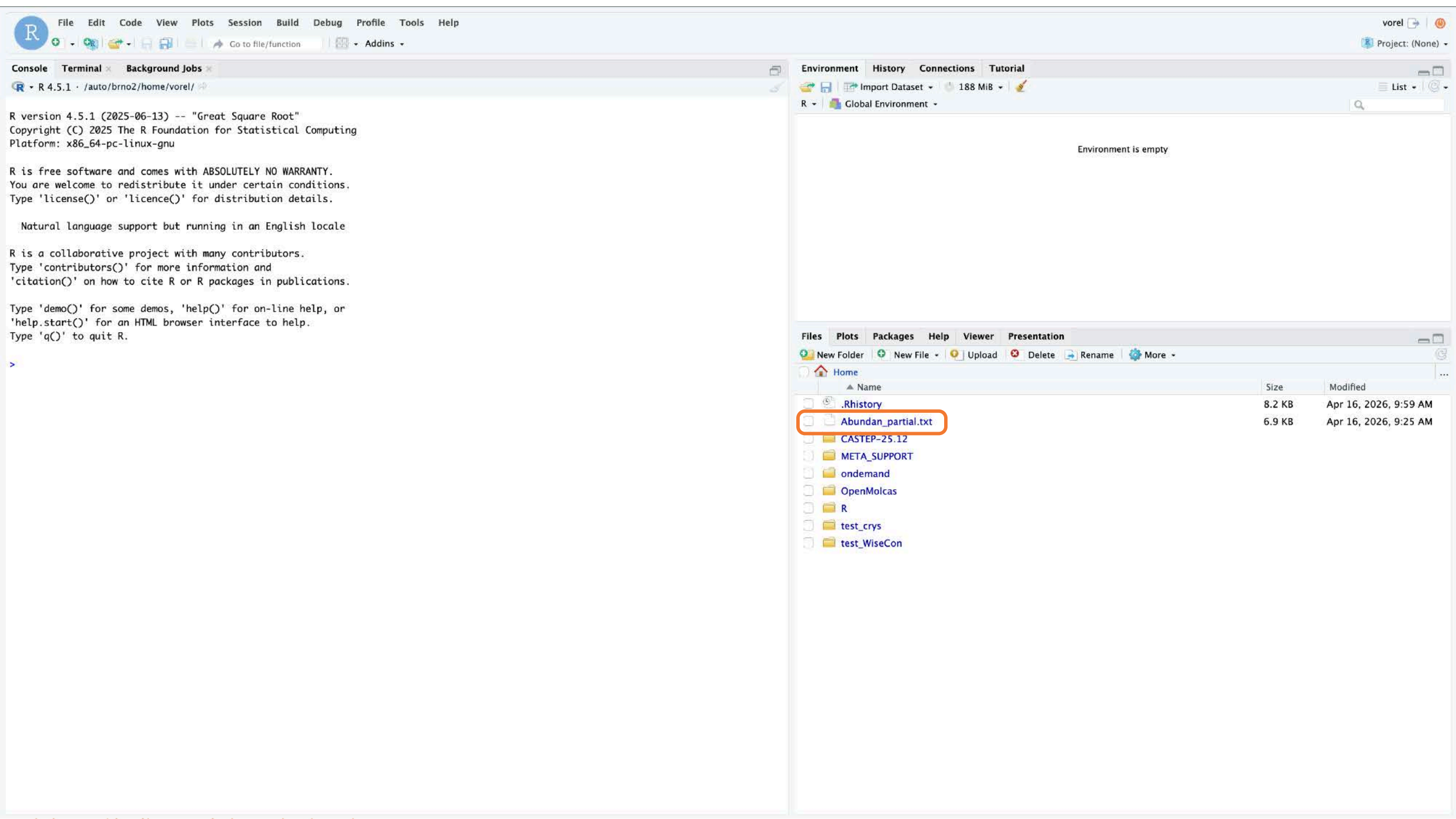
RStudio

- [Posit Support](#)
- [Posit Community Forum for the RStudio IDE](#)
- [Posit Cheat Sheets](#)
- [RStudio Packages](#)
- [Posit Products](#)

Search Engine & Keywords

- [The R Language Definition](#)
- [R Installation and Administration](#)
- [R Internals](#)





Console Terminal Background Jobs

R 4.5.1 · /auto/brno2/home/vorel/

```
R version 4.5.1 (2025-06-13) -- "Great Square Root"
Copyright (C) 2025 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

>
```

Environment History Connections Tutorial

Import Dataset 188 MiB
R Global Environment

Environment is empty

Files Plots Packages Help Viewer Presentation

New Folder New File Upload Delete Rename More

Name	Size	Modified
.Rhistory	8.2 KB	Apr 16, 2026, 9:59 AM
Abundan_partial.txt	6.9 KB	Apr 16, 2026, 9:25 AM
CASTEP-25.12		
META_SUPPORT		
ondemand		
OpenMolcas		
R		
test_crys		
test_WiseCon		

R version 4.5.1 (2025-06-13) -- "Great Square Root"

Copyright (C)
Platform: x86_64-pc-linux-gnu

R is free software
You are welcome to redistribute it under certain conditions.
Type 'license()' to view the license.

Natural language processing
Type 'demo()' to try NLTK tools.

R is a collaborative project
Type 'contrib()' to view the contribution guidelines.

'citation()' to see the citation for R.

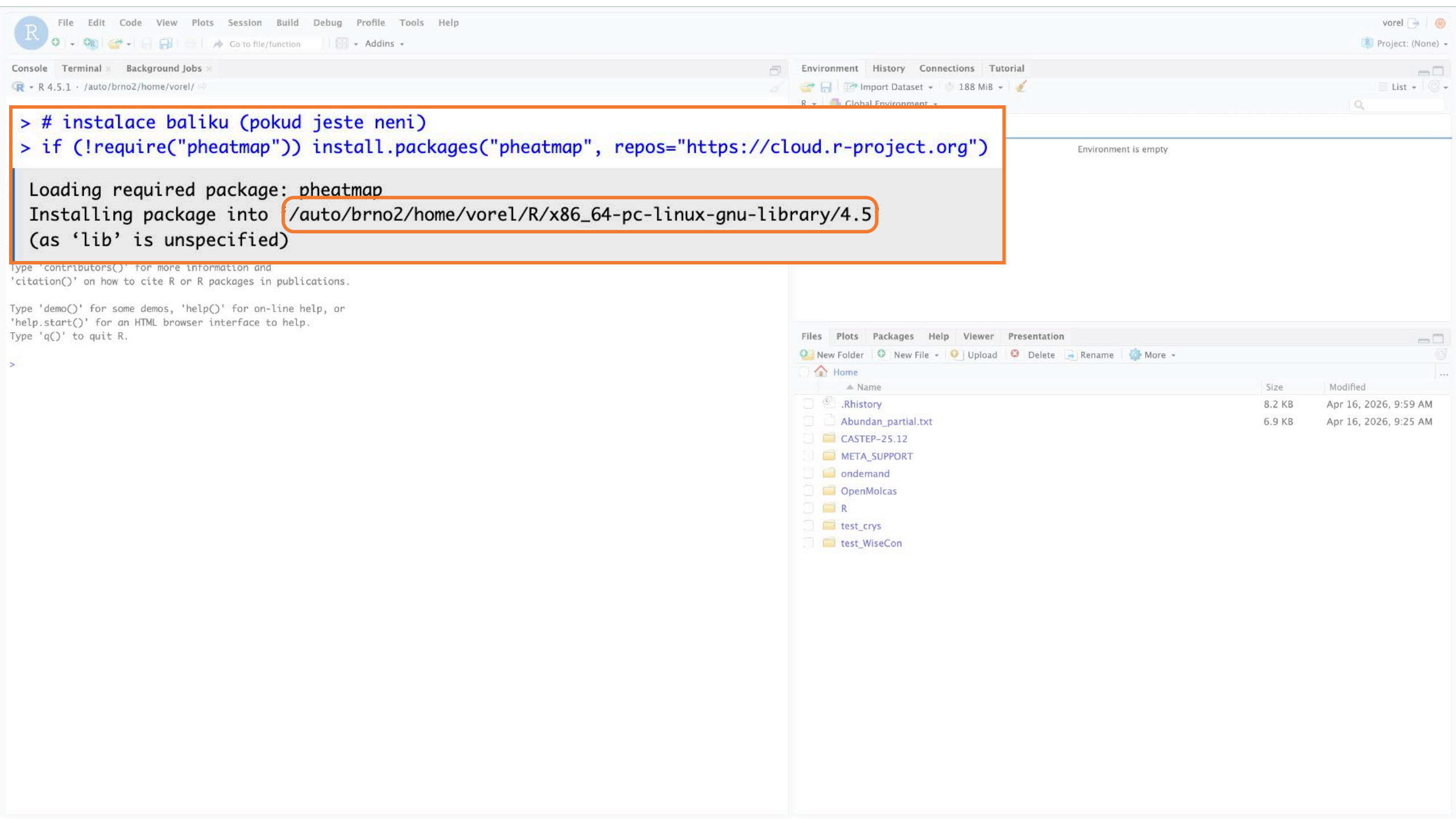
Type 'demo()' to try NLTK tools.

'help.start()' to open an HTML help browser.

Type 'q()' to quit R.

>

```
> head(data, 10)
  PG.DIA.NN.MaxLFQ_.P1a1. PG.DIA.NN.MaxLFQ_.P1a2. PG.DIA.NN.MaxLFQ_.P1a3. PG.DIA.NN.MaxLFQ_.P2a1. PG.DIA.NN.MaxLFQ_.P2a2.
0          6047048          4800536.5          4376470.5          6862754          4341160
1          77152976         103920696.0         104634088.0         77989736         122445816
2          3140518          2979155.8          2790588.0          3196619          2500136
3          5465151          4795366.5          3698618.2          3855838          4214330
4          66938324         54277356.0         50684836.0         66603780         54354108
5          15052387         12721647.0         9204932.0          14989466         10041644
6          7375690          5800779.0          5642051.0          6468283          4397466
7          3444160          516117.8          586909.6           1516715           0
8          1893248640        2437371648.0        1118569984.0        1706273024        2453326336
9          231561504        156313264.0        113507824.0        203705024        122656368
  PG.DIA.NN.MaxLFQ_.P2a3. PG.DIA.NN.MaxLFQ_.control.
0          5494603.0          3404306.8
1          113344808.0         95229736.0
2          2846252.0          1485956.6
3          4398617.5          4005499.5
4          51593840.0         46734988.0
5          9061742.0          9341989.0
6          5037537.5          6528869.5
7          402248.9           409592.7
8          1245125376.0        1248007296.0
9          113673448.0         109913368.0
```



```

> library(pheatmap)
>
> # nacteni dat
> data <- read.delim("Abundan_partial.txt",
+                   sep = "\t",
+                   header = TRUE,
+                   stringsAsFactors = FALSE)
>
> # nastaveni rownames
> rownames(data) <- data$PG_ID
> data$PG_ID <- NULL
>
> # prevod desetinne carky na tecku
> data[] <- lapply(data, function(x) {
+   as.numeric(gsub(",", ".", x))
+ })
>
> # log transformace (bezpecna pro nuly)
> data_log <- log10(data + 1)
>
> # vykresleni heatmapy
> pheatmap(
+   data_log,
+   scale = "row",           # normalizace v ramci radku
+   clustering_distance_rows = "euclidean",
+   clustering_distance_cols = "euclidean",
+   clustering_method = "complete",
+   fontsize = 8,
+   main = "Protein abundance heatmap"
+ )

```

https://cloud.r-project.org")

x-gnu-library/4.5

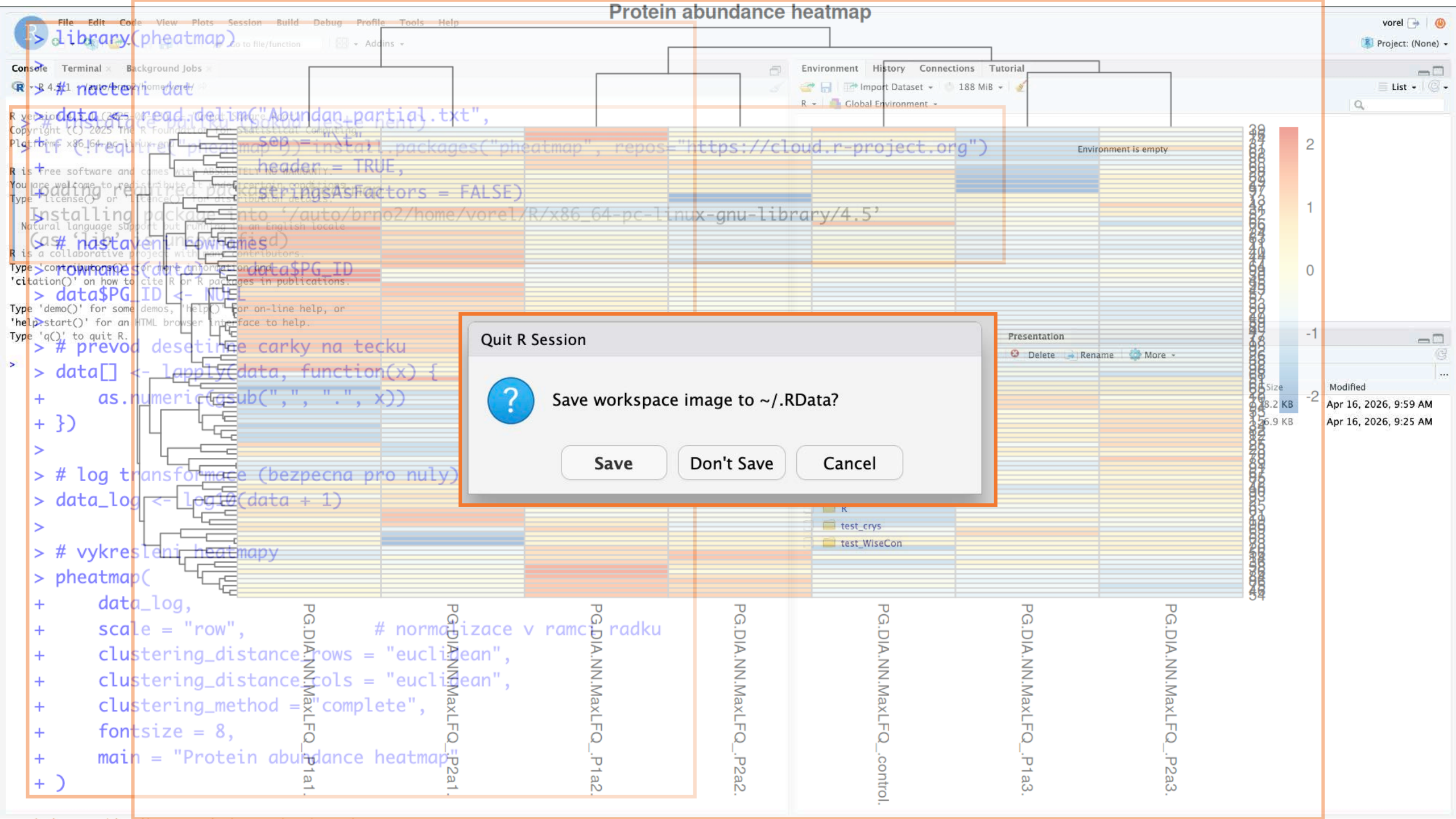
Files Plots Packages Help Viewer Presentation

New Folder New File Upload Delete Rename More

Home

Name	Size	Modified
.Rhistory	8.2 KB	Apr 16, 2026, 9:59 AM
Abundan_partial.txt	6.9 KB	Apr 16, 2026, 9:25 AM
CASTEP-25.12		
META_SUPPORT		
ondemand		
OpenMolcas		
R		
test_crys		
test_WiseCon		

Protein abundance heatmap



File Edit Code View Plots Session Build Debug Profile Tools Help

> library(pheatmap)

Console Terminal Background Jobs

R 4.4.1 auto-brno2/home/vorel

> # načtení dat

> data <- read.delim("Abundance+matrix.txt",

> header = TRUE,

> stringsAsFactors = FALSE)

> # nastavení rownames

> rownames(data) <- data\$PG_ID

> data\$PG_ID <- NULL

> # převod desetinné čarky na tečku

> data <- lapply(data, function(x) {

+ as.numeric(asub(",", ".", x))

+ })

> # log transformace (bezpečná pro nuly)

> data_log <- log10(data + 1)

> # vykreslení heatmapy

> pheatmap(

+ data_log,

+ scale = "row", # normalizace v rámci radku

+ clustering_distance_rows = "euclidean",

+ clustering_distance_cols = "euclidean",

+ clustering_method = "complete",

+ fontsize = 8,

+ main = "Protein abundance heatmap"

+)

Quit R Session

Save workspace image to ~/.RData?

Save Don't Save Cancel

vorel

Project: (None)

List

Environment History Connections Tutorial

188 MiB

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Environment is empty

Size
8.2 KB
6.9 KB

Modified
Apr 16, 2026, 9:59 AM
Apr 16, 2026, 9:25 AM

Presentation
Delete Rename More

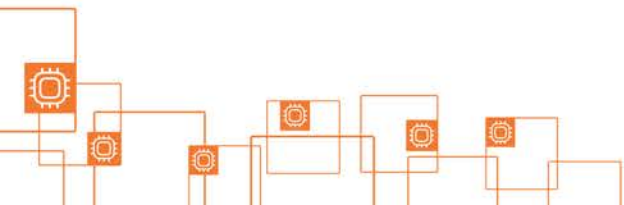
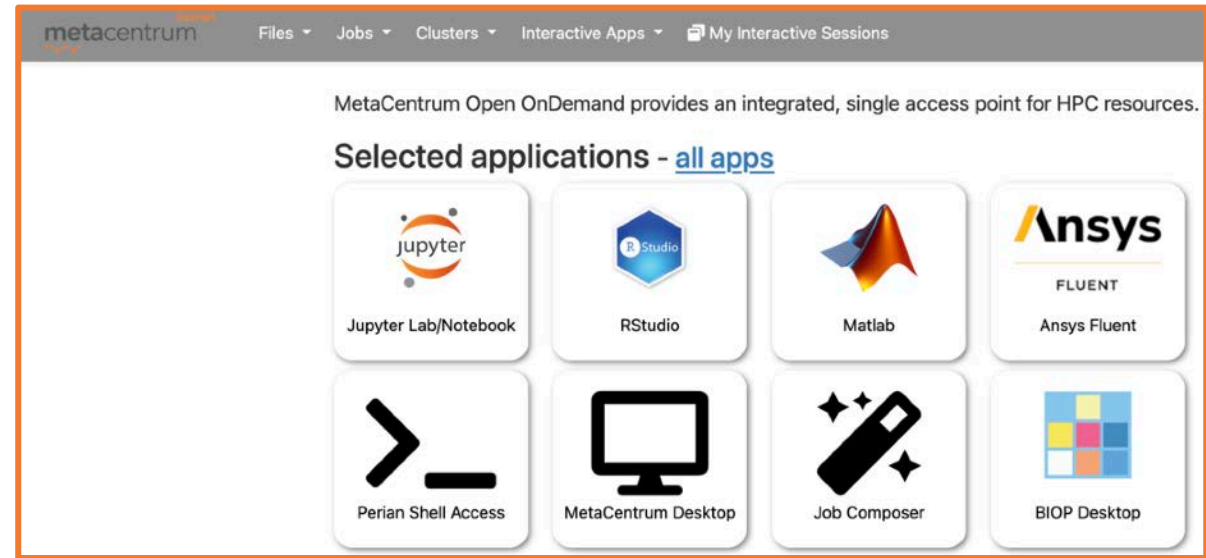
K
test_crys
test_WiseCon

■ Open OnDemand



<https://ondemand.metacentrum.cz/>

- Web application on top of OpenPBS scheduler
- Running (not only) GUI applications
- Browsing storages (tab 'Files')
- Launch a virtual desktop (MetaCentrum Desktop app)
- Command line (tab 'Clusters')



Themes: Cobalt2

- Default
- Bright**
 - CLRS
 - Github
 - Solarized Light
 - Tomorrow
- Dark**
 - Cobalt2**
 - idleToes
 - Monokai Remastered
 - Pastel Dark
 - Red Alert
 - Tomorrow Night Blue
 - Tomorrow Night Bright
 - Tomorrow Night Eighties
 - Tomorrow Night



Your Home Directories

Avail space	DATA quota	DATA used	FILES quota	FILES used	Backup	Storage
227T	none	202G	none	4.88k	snap	brno12-cerit
23T	4.29T	79.8G	2.5M	141k	snap	brno2
17T	3.32T	24k	1M	9	full	budejovice1
30T	10.4G	12k	none	8	-	liberec3-tul
339T	4.29T	1.71T	3M	201k	snap	plzen1
60T	7.34G	44k	none	18	-	praha2-natur
11T	5.36T	942G	1.01M	122k	full	praha5-elixir
48T	7.34G	28k	none	11	full	pruhonice1-ibot
136T	2.14T	5.16G	2M	41.3k	snap	vestec1-elixir

CPUs

Free	Used	Total	Centre
8775	33677	42452	META

Contact

E-mail: meta@cesnet.cz
Web: http://www.metacentrum.cz/

Tip of the day: Explore our past seminars, lectures and hands-on courses with many valuable tips and tricks (<https://docs.metacentrum.cz/en/docs/tutorials>).

```
(BOOKWORM)vorel@zenith:~$ pwd  
/storage/brno12-cerit/home/vorel  
(BOOKWORM)vorel@zenith:~$ qstat -u vorel
```

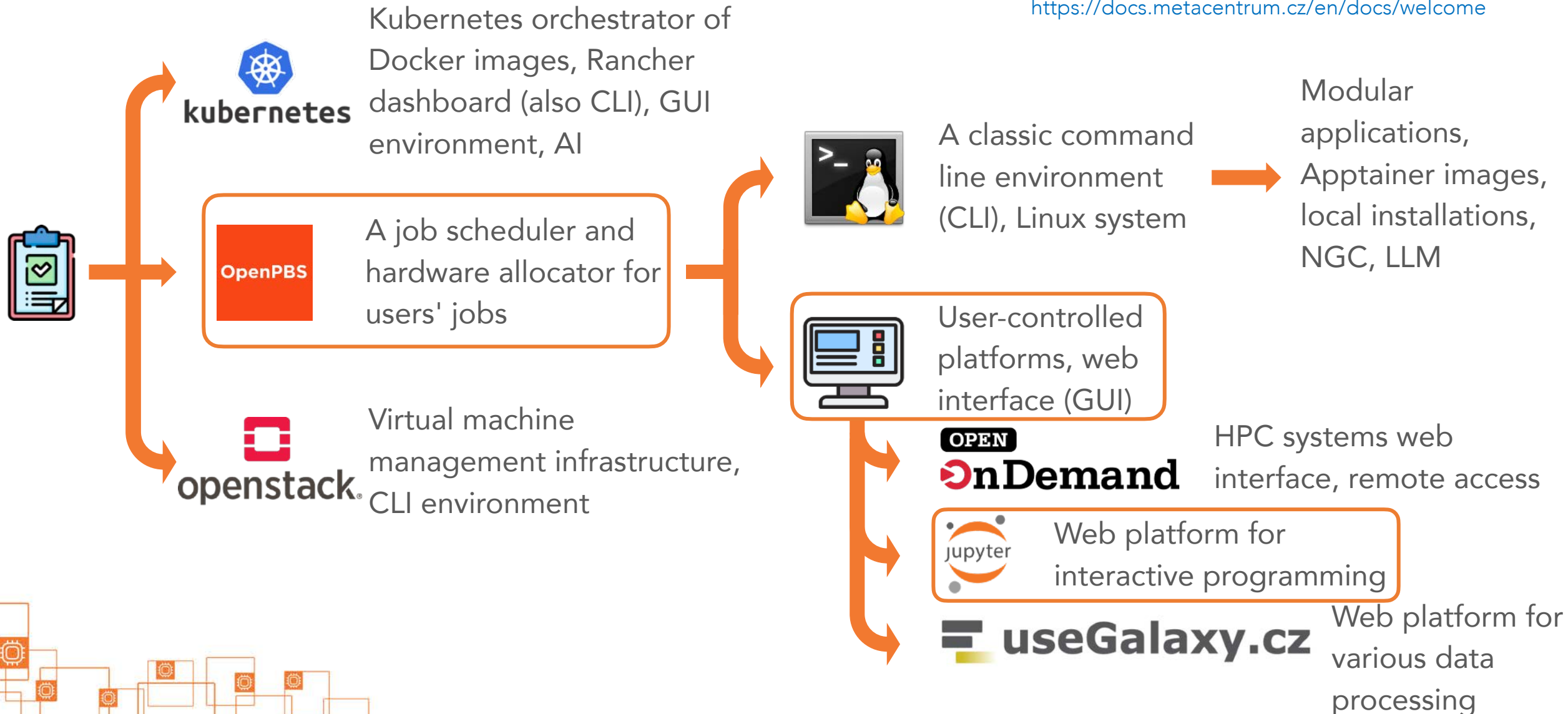
pbs-m1.metacentrum.cz:

Job ID	Username	Queue	Jobname	SessID	NDS	TSK	Req'd Memory	Req'd Time	Elap S	Time
19137135.pbs-m1.met*	vorel	interac*	STDIN	387033	1	1	400mb	18:00	R	00:00

(BOOKWORM)vorel@zenith:~\$

How to use it

<https://docs.metacentrum.cz/en/docs/welcome>

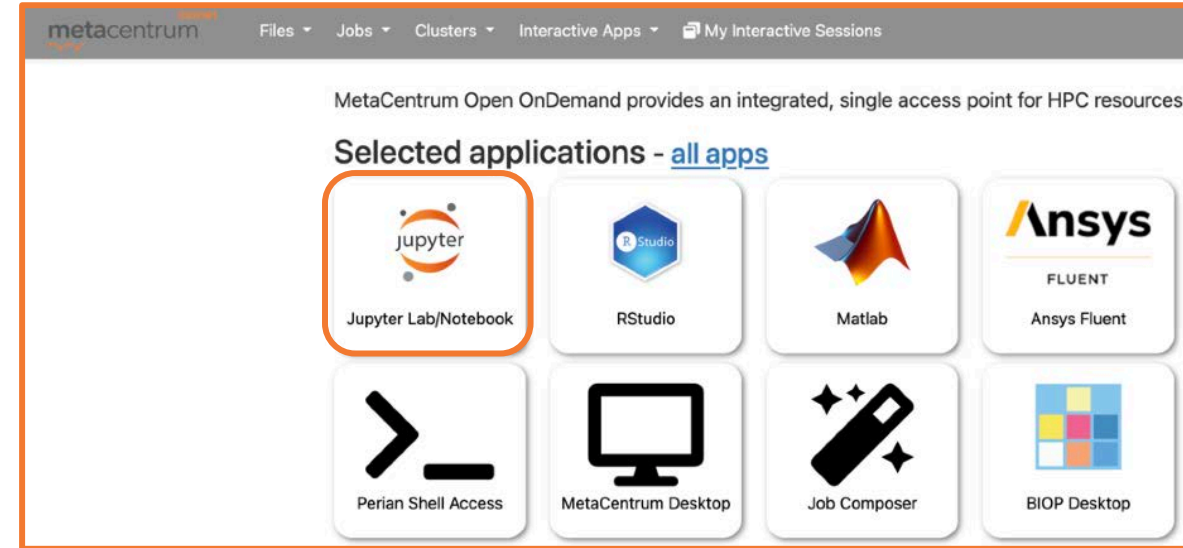


Jupyter Notebook in Open OnDemand

■ Jupyter Notebook

<https://ondemand.metacentrum.cz/>

- Web-based interactive environment mixing code, text, and visuals
- Excellent for step-by-step data analysis, machine learning, and prototyping
- Supports many languages (Python, R, Julia)
- Embeds plots, widgets, and equations directly
- Promotes reproducible, shareable work via version control and exports
- Also accessible in the Kubernetes service (with Claude Code)



<https://docs.cerit.io/en/docs/web-apps/jupyterhub>

Jupyter Notebook

<https://ondemand.metacentrum.cz/>



Jupyter

- Web-based interactive environment mixing code, text, and visuals
- Excellent for step-by-step data analysis, machine learning, and prototyping
- Supports many languages (Python, R, Julia, etc.)
- Embeds plots, widgets, and equations
- Promotes reproducible, shareable workflows
- Also accessible in the Kubernetes ecosystem

The screenshot shows the configuration page for the Jupyter Notebook/Lab application. On the left, a list of applications is shown, with "Jupyter Notebook/Lab" selected. The main area contains configuration options for the selected application:

- Interactive Apps:** A list of applications including "Kubernetes on OpenStack (KOST) (beta)", "OpenStack VM (beta)", "Ansys/Enight", "Ansys/Fluent", "Ansys/Workbench", "BIOP Desktop", "CLCgenomicsWB", "Matlab", "MetaCentrum Desktop", "QuPath", "VMD Desktop", "Jupyter - Evo2 (beta)", "Jupyter Notebook/Lab" (selected), "Matlab webapp (beta)", and "RStudio Server".
- Jupyter Notebook/Lab:** This app will launch a Jupyter Notebook server on one or more nodes.
- Mode:** Radio buttons for "Jupyter Notebook" (selected) and "Jupyter Lab".
- Number of hours:** A dropdown menu set to 1.
- Number of CPUs on single node:** A dropdown menu set to 1.
- Memory (GB):** A dropdown menu set to 5.
- GPUs:** A dropdown menu set to 0.
- Scratch local (GB):** A dropdown menu set to 5.
- Jupyter Image:** A dropdown menu set to "Metacentrum/debian11.sif".
- Full path of custom singularity image with Jupyter notebook:** An empty text input field.
- PBS Queue:** A dropdown menu set to "interactive".
- Jupyter working directory location:** A dropdown menu set to "/storage/brno2".
- Launch:** A blue button to start the application.

Files

Running

Clusters

Select items to perform actions on them.

Upload

New ▾


 0 ▾ /

Name ▾

  CASTEP-25.12

  META_SUPPORT

  ondemand

  OpenMolcas

  test_crys

  Untitled.ipynb

Notebook:

 Python 3 (ipykernel)

 SMLv2

Other:

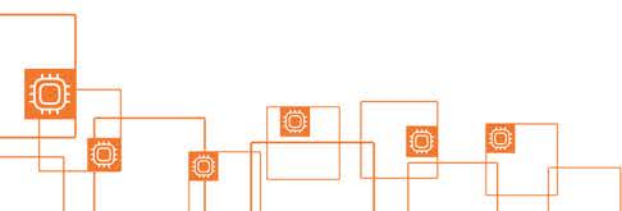
 Text File

 Folder

 Terminal

před hodinou

4.98 kB



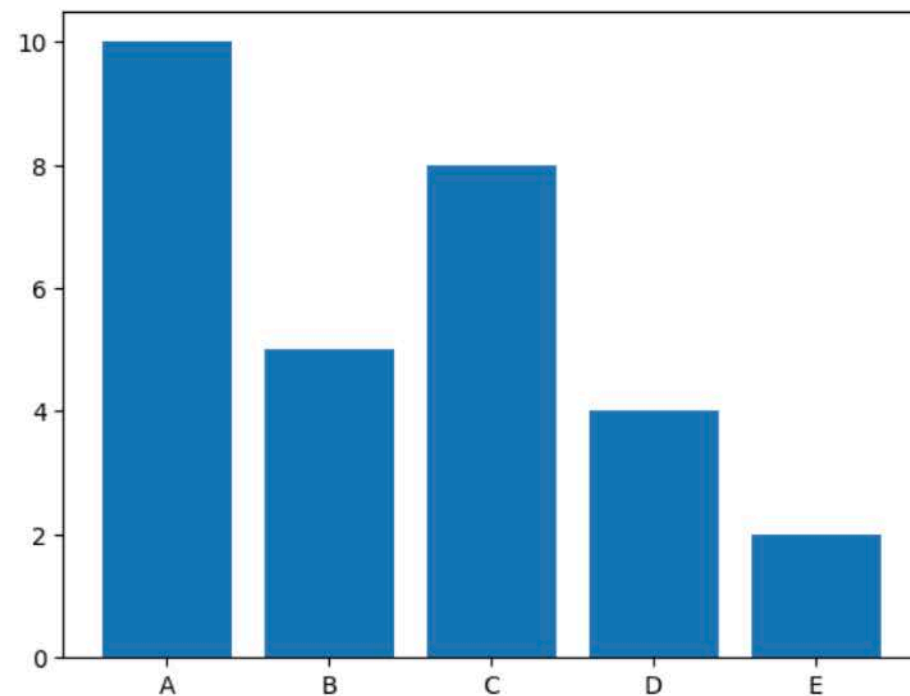
```
[1]: import os
current_working_directory = os.getcwd()
print(current_working_directory)

/auto/brno2/home/vorel
```

```
[2]: pip install matplotlib --quiet
```

Note: you may need to restart the kernel to use updated packages.

```
[3]: from matplotlib import pyplot as plt
x = ['A', 'B', 'C', 'D', 'E']
y = [10, 5, 8, 4, 2]
plt.bar(x, y)
plt.show()
```



Notebook Intelligence 

{"cell_index": 0}

(no content)

Perfect! I've successfully loaded your 'age_data.csv' file and displayed the first 5 rows. The data shows:

- **4 columns:** sex, age, city, and country
- **Data sample:**
 - Row 0: Male, age 64, from Houston, Brazil
 - Row 1: Female, age missing, from San Jose, Mexico
 - Row 2: Male, age 68, from San Antonio, Mexico
 - And so on...

I notice there are some missing age values (shown as NaN). Let me save the notebook:

Done! The notebook has been created and saved with your CSV data loaded. The first few rows are now displayed above.

Ask Notebook Intelligence...

exaple1.ipynb 

@



Launcher

exaple1.ipynb

age_data.csv



Code



git

Open in...



```
[3]: import pandas as pd

# Load the CSV file
df = pd.read_csv('age_data.csv')

# Display the first few rows
df.head()
```

```
[3]:
```

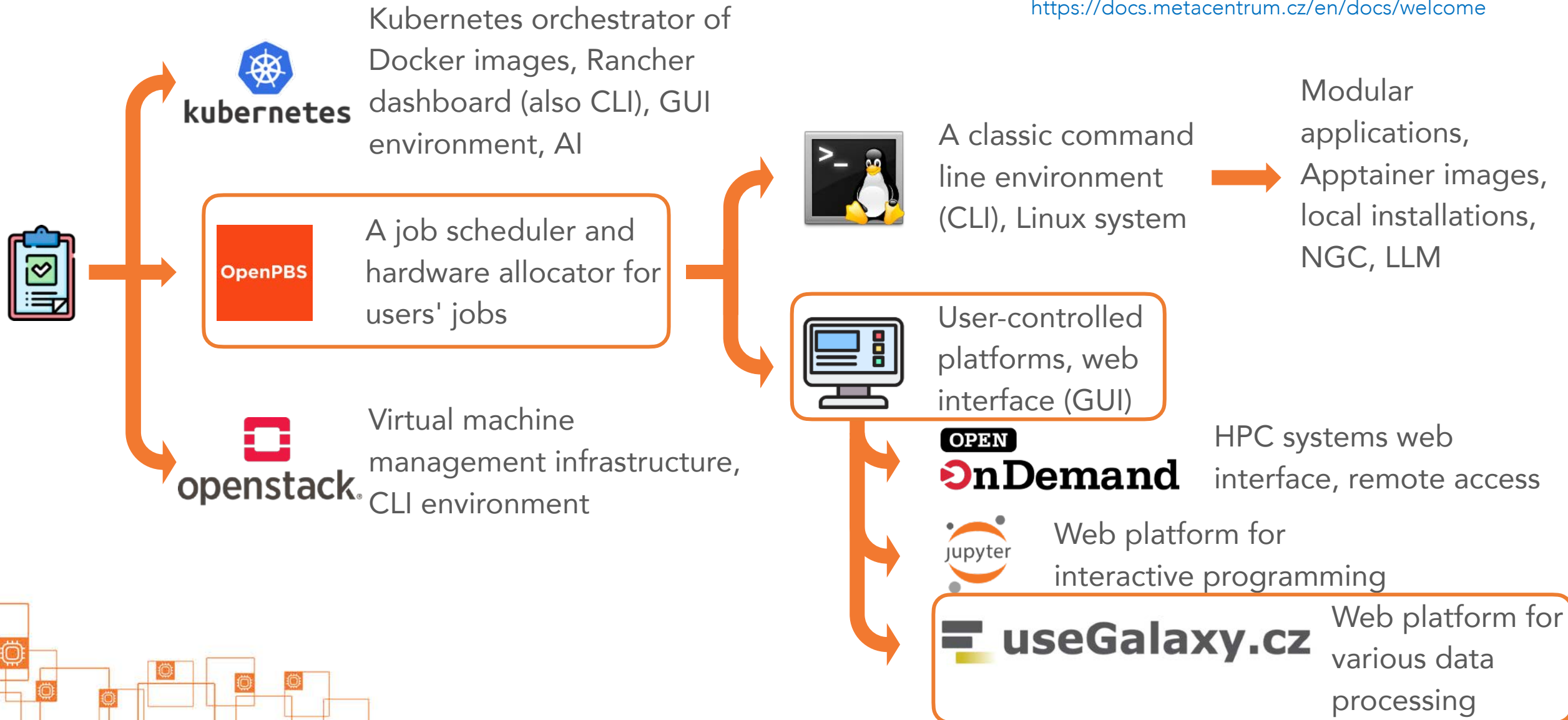
	sex	age	city	country
0	Male	64.0	Houston	Brazil
1	Female	NaN	San Jose	Mexico
2	Male	68.0	San Antonio	Mexico
3	Male	72.0	Dallas	USA
4	Male	81.0	San Antonio	Argentina

[]:



How to use it

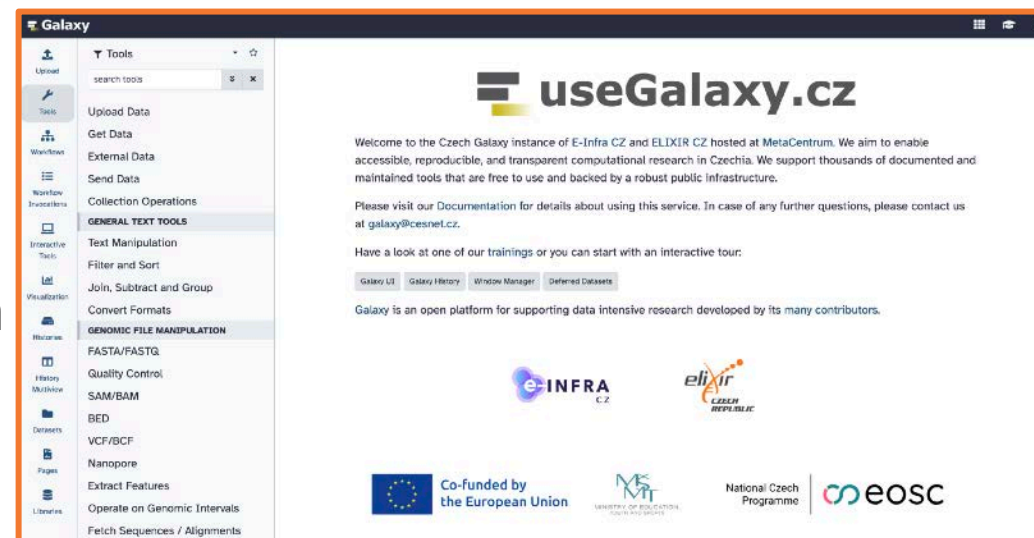
<https://docs.metacentrum.cz/en/docs/welcome>



Galaxy useGalaxy.cz

<https://usegalaxy.cz/>

- National instance hosted by MetaCentrum
- Web application based on OpenPBS
- A system for complex data analysis, workflow creation and sharing and teaching
- Processing of '-omic' data, LLM, ML, material science, ecology, chemistry, statistical analysis and humanities studies, and more
- Users do not need to worry about hardware reservation
- No programming required





Upload



Tools



Workflows

Workflow
InvocationsInteractive
Tools

Visualization



Histories

History
Multiview

Datasets



Pages



Libraries

Tools

search tools

Upload Data

Get Data

External Data

Send Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

GENOMIC FILE MANIPULATION

FASTA/FASTQ

Quality Control

SAM/BAM

BED

VCF/BCF

Nanopore

Extract Features

Operate on Genomic Intervals

Fetch Sequences / Alignments

GENOMICS ANALYSIS

Assembly



Welcome to the Czech Galaxy instance of [E-Infra CZ](#) and [ELIXIR CZ](#) hosted at [MetaCentrum](#). We aim to enable accessible, reproducible, and transparent computational research in Czechia. We support thousands of documented and maintained tools that are free to use and backed by a robust public infrastructure.

Please visit our [Documentation](#) for details about using this service. In case of any further questions, please contact us at galaxy@cesnet.cz.

Have a look at one of our [trainings](#) or you can start with an [interactive tour](#):

[Galaxy UI](#)[Galaxy History](#)[Window Manager](#)[Deferred Datasets](#)

Galaxy is an open platform for supporting data intensive research developed by its many contributors.



Co-funded by
the European Union



MINISTRY OF EDUCATION,
YOUTH AND SPORTS

National Czech
Programme



Upload

Tools

Workflows

Workflow Invocations

Interactive Tools

Visualization

Histories

History Multiview

Datasets

Pages

Libraries

Tools

Upload Data

Get Data

External Data

Send Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

GENOMIC FILE MANIPULATION

FASTA/FASTQ

Quality Control

SAM/BAM

BED

VCF/BCF

Nanopore

Extract Features

Operate on Genomic Intervals



Fetch Sequences / Alignments

GENOMICS ANALYSIS

Assembly

Upload from Disk or Web to **Unnamed history**

Regular Composite Collection Rule-based

 Drop files hereType (set all): Auto-detect Reference (set all): unspecified (?)  Choose local file Choose remote files Paste/Fetch data

Start

Pause

Reset

Close

Co-funded by
the European UnionNational Czech
Programme

Upload

Tools

Upload Data

Get Data

External Data

Send Data

Collection Operations

GENERAL TEXT TOOLS

Text Manipulation

Filter and Sort

Join, Subtract and Group

Convert Formats

GENOMIC FILE MANIPULATION

FASTA/FASTQ

Quality Control

SAM/BAM

BED

VCF/BCF

Nanopore

Extract Features

Operate on Genomic Intervals

Fetch Sequences / Alignments

GENOMICS ANALYSIS

Assembly

Annotation

Mapping

Variant Calling

RNA Analysis

Multiple Alignments

Upload from Disk or Web to **Unnamed history**

Regular Composite Collection Rule-based

You added 1 file(s) to the queue. Add more files or click 'Start' to proceed.

New File 115 b Auto-detect unspecified (?) 0%

Download data from the web by entering URLs (one per line) or directly paste content.

```
https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/029/291/075/GCA_029291075.1_E_nip/GCA_029291075.1_E_nip_genomic.fna.gz
```

Type (set all): Auto-detect Reference (set all): unspecified (?)

Choose local file Choose remote files Paste/Fetch data **Start** Pause Reset Close

Using 0% of 200.0 GB vorel

History

search datasets

Unnamed history

This history is empty.
You can load your own data or get data from an external source.



Co-funded by
the European Union



National Czech
Programme

eosc

- Upload
- Tools
 - search tools
- Upload Data
- Get Data
- External Data
- Send Data
- Collection Operations
- GENERAL TEXT TOOLS
 - Text Manipulation
 - Filter and Sort
 - Join, Subtract and Group
 - Convert Formats
- GENOMIC FILE MANIPULATION
 - FASTA/FASTQ
 - Quality Control
 - SAM/BAM
 - BED
 - VCF/BCF
 - Nanopore
 - Extract Features
 - Operate on Genomic Intervals
 - Fetch Sequences / Alignments
- GENOMICS ANALYSIS
 - Assembly
 - Annotation
 - Mapping
 - Variant Calling
 - RNA Analysis
 - Multiple Alignments
- Workflows
- Workflow Invocations
- Interactive Tools
- Visualization
- Histories
- History Multiview
- Datasets
- Pages
- Libraries
- Notifications
- More

Upload from Disk or Web to **Unnamed history**

Regular Composite Collection Rule-based

New File 115 b Auto-detect unspecified (?) 100%

Download data from the web by entering URLs (one per line) or directly paste content.

```
https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/029/291/075/GCA_029291075.1_E_nip/GCA_029291075.1_E_nip_genomic.fna.gz
```

Type (set all): Auto-detect Reference (set all): unspecified (?)

Choose local file Choose remote files Paste/Fetch data **Start** Pause Reset Close

Using 0% of 200.0 GB vorel

History

search datasets

Unnamed history

0 B

1: GCA_029291075.1_E_nip_g
enomic.fna.gz



- Upload
- Tools
- Workflows
- Workflow Invocations
- Interactive Tools
- Visualization
- Histories
- History Multiview
- Datasets
- Pages
- Libraries
- Notifications
- More

- Tools
- search tools
- Upload Data
 - Get Data
 - External Data
 - Send Data
 - Collection Operations
 - GENERAL TEXT TOOLS**
 - Text Manipulation
 - Filter and Sort
 - Join, Subtract and Group
 - Convert Formats
 - GENOMIC FILE MANIPULATION**
 - FASTA/FASTQ
 - Quality Control
 - SAM/BAM
 - BED
 - VCF/BCF
 - Nanopore
 - Extract Features
 - Operate on Genomic Intervals
 - Fetch Sequences / Alignments
 - GENOMICS ANALYSIS**
 - Assembly
 - Annotation
 - Mapping
 - Variant Calling
 - RNA Analysis
 - Multiple Alignments

1: **GCA_029291075.1_E_nip_genomic.fna.gz** ok

21,044 sequences
format **fasta.gz** database ? size **281.4 MB**

[Preview](#) [Visualize](#) [Details](#) [Edit](#)

This dataset is large and only the first megabyte is shown below.
[Show all](#) | [Save](#)

```
>JAQBSW010000001.1 Eudiplozoon nipponicum isolate JV_EN_01 E_nip_1, whole genome shotgun sequence
gcaatgttctctcattcgagtaattaacgctctggctgatcgccctacgtaagttgatcgacaacttgcgcgatctt
ggtaaaaaatatttgatattgacattaacgactgtttctctttgattactggtaaggcaagtggttacacataaac
actcgtagtttagctgcattaaatgctggttaagctgcttcaactttctttgtgctcattcgggaaactgcatcac
cataaatgggatgttaggtagacgcttttctttctggcctaataagctgtttctgtttgattcctcttttaataat
tgattcaatnaatagctTAGGATaccattagacataaactggttaattgctttatataataatcttttagaatt
gactaaccaactgctcaattctaaaaagcaaccagcaactgtgataattatgaaatggtaacaaaactgctgtat
aacaataactgtgcagaccatgttggtttatggtaaatgttctgtttaaattacattagttttttcattaatagata
tctaggaaaggcagacagtcattctctgctctgctgtaagcagatattcttatgtgctttgtgaaacttaag
taatgttctgtgattgcgacctgttcttaagtaacaagaatataatcaccgtaaccgtatatattgagcacatttaatta
tgattttctaggtgccctagaaaaatctcagcaaacactggagataaggggctccattgagcacaccatttcaat
tactttgctgctccacttgtctatttcatctgatttatcgtctagacgtggaattttgtattgcatttaaacgcttttc
catcttttgcagctagctagattgtccattataaagtcgctctttctttgcttcttacttaagaacagggtcgaa
tcacgaacattgcttcaaggttcaacaagcacataaagaatcggcttcacgctggagcagagaagatgactgtct
gcctttcctagatgactatnaagaagaactaatggtaacttaaacgaaacaaattacataaaccaacatggctg
cacagtttgcataaacagcagtttgtgctatttaaacataaattatcaacagttcgtggttcttcaagaatagac
aggttggttagcaagattctaaagaaaatgaatataaagaacaattaacggttaaatgtctaatggctatcccagaca
attaattgaatcaattataaaaggggagttaacagaaaacagcgtattatgggcagaaaagaaaatagctaccta
acataccatttaagggggatgcatctcccgaatggcacaagaagaagatgaagcagcattgacacacacattaatgca
gccaaactcagagtgattatagggttaacacacttgcctttactagcaatcaagagaaaacagcgttaattgctgatc
aaatattttaccatattgcatgcatgctgttgcgacacttaactagcagcagcagcagcagcagcagcagcagcagc
tgagagaaacattgcaacttaaacgcccgttttaaatcaaatgggtaatacacgctctgctattgctgcaacacatcatcga
gcatcgcagctatgttctcagacgtagcattcaagtaattactaaatgcaaaaagggagctagccacattatgg
ataaacaaattgcagccggaattatgcaagcagaagaagcagctcaagcattgagcctgaactggatgtagggcagcga
atgtgattggttccgtgttccacactctctatgttcacatattctatcttgttttgccttcaaatttatataaaa
atatgttcattagaagatatttcagaagacaatgcaagaagaaaatcttctgtactgataaattcaaatttattgca
taGGCTCTTATGAAACTAAGCTATAGGTGACATAGTTCGCTGTTTACGTATGCAACATTGATGAGCGAAATCGTGGC
AAAGAAAAGGCTGTAGGATGTTACCCGTTGTTAATGAATGCTTAAGCACTGCTGGTAAATAGGAAGGTGTTGTACCAAT
CGAACCTTAGTATTCCAGGAACCTCGGAACACGAGCACGAACTTTATGGTGTCCGTACTCGGAAAAAAGTATGCGCCA
ACATAGGCAACCTTCCAATTAGTCTATAAATCAACAATaatgtcattttataaataaacgaacattgtttgcgpggcg
atcATAACCTTtttataatcaggggttgcgaactgttataaaaaacctgaccattaagttctAAAATTTAGTCATTATA
ATCTAAGattttagtagtagagaaaaggcaattcaaccagtagaccggtttcgttcttatgggaactcatcatgactggc
ctgttagcatccatttagcgttaaagttcgaactaacatccacaactctgcatcaacgaAAGGATACACTGTTATATG
GGGAAATGGGatagttatttggaacttagcgCTAAATGGATGATATATAGGCTCACaaaaggaagactaaccaag
acgatttagtagtatataactgaaattgtgctagtttgcgCACAGTGACTcggtagttagagcaccacattcggaa
acagacgttccaggttgcgcCCTGACCGcttcaaatctggcaagtacatgaggagttctcaataaggcgaatcggg
tctactggcgaatttgcgttttctacactgtacaattcttagtttatactgcaaatgtatgcacttaattgcatcaggtga
tcatatcaagttcgaacacttgattataaatttatgTCATTTTGTACTCAGTGTATGCAACATATTTTATCATCGTTTCC
TCACTATTTAACCCGCAAAAGAAACAAATGACACAGTAGAAAAAAAAGATCAGctaaaaattatogaaaattatagaaa
gcatgcataaaaaataaattagctcACTCACGAGAACACGCTGCCTATCCGCAATTAAGTACTGAGCAAATTTGATGCAC
```

History

search datasets

Unnamed history

295 MB

1: **GCA_029291075.1_E_nip_geno**

mic.fna.gz

- Upload
- Tools
- Workflows
- Workflow Invocations
- Interactive Tools
- Visualization
- Histories
- History Multiview
- Datasets
- Pages
- Libraries
- Notifications
- More

- Tools
- search tools
- Upload Data
 - Get Data
 - External Data
 - Send Data
 - Collection Operations
 - GENERAL TEXT TOOLS**
 - Text Manipulation
 - Filter and Sort
 - Join, Subtract and Group
 - Convert Formats
 - GENOMIC FILE MANIPULATION**
 - FASTA/FASTQ
 - Quality Control
 - SAM/BAM
 - BED
 - VCF/BCF
 - Nanopore
 - Extract Features
 - Operate on Genomic Intervals
 - Fetch Sequences / Alignments
 - GENOMICS ANALYSIS**
 - Assembly
 - Annotation
 - Mapping
 - Variant Calling
 - RNA Analysis
 - Multiple Alignments

1: **GCA_029291075.1_E_nip_genomic.fna.gz** ok

21,044 sequences
format **fasta.gz** database ? size **281.4 MB**

[Preview](#) [Visualize](#) [Details](#) [Edit](#)

This dataset is large and only the first megabyte is shown below.
[Show all](#) | [Save](#)

```
>JAQBSW010000001.1 Eudiplozoon nipponicum isolate JV_EN_01 E_nip_1, whole genome shotgun sequence
gcaatgttctctcattcgagtaattaaccgtctggctgatcgccctacgtaagttgatcgacaacttgcgcatgtatcct
ggtaaaaaatatttgatattgacattaacgactgtttctctttgattactggtaaggcaagtggttacacataaataac
actcgtagtttagctgcattaaatagcgtgtaaatgctgcttcaactttctttgtgctcattcgggaaactgcatcac
cataaatgggatgttaggtagacgcttttctttctggcctaataagctgtttctgtttgattcctcttttaataat
tgattcaatnaatagtCTAGGATaccattagacataaactgtaattgtctttatataaattttctttagaatatt
gaactaaccaactgtcaattcttaaaagcaaccagcaactgtgataattatattgaaatggtaacaaaactgctgtat
aacaataactgtgcagaccatgttggtttatggtaaatgttctgtttaaatttaccattagttttttcattaatagata
tctaggaaaggcagacagtcattctctgctgctgtgtaagcagatattcttatgtgctttgtgaaacttaag
taatgttctgtgattgcgacctgttcttaagtaacaagaataatatacaccagcaccgatatattgagcacatttaatta
tgatatttctaggtgccctagaaaaatctcagccaactggagataaggggctcccattgagcacaccatttcaat
tactttgctgctccacttgtctatttcatctgatttatcgtctagacgtggaatttggattgcatthaacacgcttttc
catcttttgcagctagctagatttgcattataaagtcgctctttctttgcttcttacttaagaacagggtcgcaa
tcacgaacattgcttcaaggtttcaacaagcacataaagaatcggcttcacgctggagcagagaagaagactgctgtc
gcctttcctagatgactatnaagaagaactaatggtaacttaaacagaaacatttacataaaccaacatggctg
cacagtttgcataaacagcagtttgtgctatttaacataaataatcaacagttcgtggttcttcaagaatagac
aggttggttagcaagattctaaagaaaatgaatataaagaacaattaacggtttaagctcaatggctatcccagaca
attaattgaatcaattataaaaggggagttaacagaaaacagacgtattatgggcagaaaagaaaatagctaccta
acataccatttaagggggatgcatctcccgaatggcacaagaagaagatgaagcagcattgacacacacattaatgca
gccaaactcagagtgattattagggttaacacacttgcctttactagcaatcaagagaaaacagccttaattgctgatc
aaatattttaccatattgcatgcatggttgcgacacttaactagcagcagcagcagcagcagcagcagcagcagcag
tgagagaaacattgcaacttaaacgcccgttttaaatcaaatgggtaatacacgctgctattgagcaacacatcatcga
gcatcgacgtatgttctcagacgtagcattcaagtaattactaaatgcaaaacaaaggagctagccacattatgg
ataaacaaattgacagcggatattgcaagcagaagaagcagctcaagcattgagcctgaactggatgtagggcagcga
atgtgattggtttccgtgtttcacactctctatgttcacatattctatcttgttttgccttcaaatttatataaaa
atatgttcattagaagatatttcagaagacaatgcaagaagaaattttctgtactgataaattcaaatattttgca
taGGCTCTTATGAAACTAAGCTATAGGTGACATAGTTCGCTGTTTACGTATGCAACATTGATGAGCGAAATCGTGGC
AAAGAAAAGGCTGTAGGATGTTACCCGTTGTTAATGAATGCTTAAGCACTGCTGGTAAATAGGAAGGTGTTGTACCAAT
CGAACCTTAGTATTCCAGGAACCTCGGAACACGAGCACGAACCTTATGGTGTCCGTACTCGGAAAAAAGTATGCGCCA
ACATAGGCAACCTTCCAATTAGTCTATAAATCAACAATaatgtcattttataaataaacgaacattgtttgcgpggcg
atcATAACCTTttttataatcagggttgcgaactgtttataaaaacctgaccattaagttctAAAATTTAGTCATTATA
ATCTAAGattttagtagtagaagaaaggcaattcaaccagtagaccggtttcgttcttatgggaactcatcatgactggc
ctgttagcatccatttagcgttaaagttcgaactaacaccacaactctgtcattaacgaAAGGATACGTTTATG
GGGAAATGGGatagttatttggaaacttagcgCTAAATGGATGATATATAGGCTCACaaaggaagactaaccaag
acgatttagtagtatataactgaaattgtgctagtttgcgCACAGTGACTcggtagttagagcaccacattcggaa
acagacgttcccagggttgcgCCTGACCGcttcaaatctggcaagtacatgaggagttctcaataaggcgaatcgg
tctactggcaaatgtgctttctacactgtacaattcttagtttatactgcaaatgtatgcacttaattggatcaggtga
tcatatcaagttcgaacacttgattataaatttatgTCATTTTGTACTCAGTGTATGCAACATATTTTATCATCGTTTC
TCACTATTTAACCCGCAAAAGAAACAAATTTGACACAGTAGAAAAAAAAGATCAGctaaaaattatogaaaattatagaaa
gcatgcataaaaaataaattagctcACTCACGAGAACACGCTGCCTATCCGCAATTAAGTACTGAGCAAATTTGATGCAC
```

History

search datasets

Unnamed history

295 MB

1: **GCA_029291075.1_E_nip_geno mic.fna.gz**

Tools

statistics

Show Sections

Statistics on presence-absence of a numeric variable

Summary Statistics for any numerical column

FASTQ Summary Statistics by column

Fasta Statistics display summary statistics for a FASTA file

SeqKit statistics of FASTA/Q files

Compute quality statistics

Genome annotation statistics

Read length statistics from a set of FAST5 files

BAM Index Statistics

Stacks: statistics on stacks found for multiple samples

MiModD Coverage Statistics calculates coverage statistics for a BCF file as generated by the MiModd Variant Calling tool

ProteomIQon PSMStatistics utilizes semi supervised machine learning techniques to integrate search engine scores as well as the mentioned quality scores into one single consensus score.

MapStatistics Extract extended statistics on the features of a map for quality control

zonal statistics over each area

Boxplot of quality statistics

Dr. Disco (classify) Classifies detected break-points in RNA-seq based on corresponding statistics and blacklists

cmstat Summary statistics for covariance model

BAM/SAM Mapping Stats reads mapping statistics for a provided BAM or SAM file.

1: renamed_GCA_029291075.1_E_nip_genomic.fna.gz ok

21,044 sequences

format **fasta.gz** database ? size **281.4 MB**

Preview

Visualize

Details

Edit

Edit Dataset Attributes

Attributes

Datatypes

Permissions

Name

renamed_GCA_029291075.1_E_nip_genomic.fna.gz

Info

My genome file

Annotation - optional

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nullam justo enim, consectetur nec, ullamcorper ac, vestibulum in, elit. Maecenas aliquet accumsan leo.

Add an annotation or notes to a dataset, annotations are available when a history is viewed.

Database/Build - optional

unspecified (?)

Save

Auto-detect

History

search datasets

Unnamed history

Unnamed history

Annotation (optional)

Add Tags

Save

Cancel

295 MB

1

1: renamed_GCA_029291075.1_E_nip_genomic.fna.gz

Tools

statistics

Show Sections

Statistics on presence-absence of a numeric variable

Summary Statistics for any numerical column

FASTQ Summary Statistics by column

Fasta Statistics display summary statistics for a FASTA file

SeqKit statistics of FASTA/Q files

Compute quality statistics

Genome annotation statistics

Read length statistics from a set of FAST5 files

BAM Index Statistics

Stacks: statistics on stacks found for multiple samples

MiModD Coverage Statistics calculates coverage statistics for a BCF file as generated by the MiModd Variant Calling tool

ProteomIQon PSMStatistics utilizes semi supervised machine learning techniques to integrate search engine scores as well as the mentioned quality scores into one single consensus score.

MapStatistics Extract extended statistics on the features of a map for quality control

zonal statistics over each area

Boxplot of quality statistics

Dr. Disco (classify) Classifies detected break-points in RNA-seq based on corresponding statistics and blacklists

cmstat Summary statistics for covariance model

BAM/SAM Mapping Stats reads mapping statistics for a provided BAM or SAM file.

1: renamed_GCA_029291075.1_E_nip_genomic.fna.gz ok

21,044 sequences

format **fasta.gz** database ? size **281.4 MB**

Preview

Visualize

Details

Edit

Edit Dataset Attributes

Attributes

Datatypes

Permissions

Name

renamed_GCA_029291075.1_E_nip_genomic.fna.gz

Info

My genome file

Annotation - optional

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nullam justo enim, consectetur nec, ullamcorper ac, vestibulum in, elit. Maecenas aliquet accumsan leo.

Add an annotation or notes to a dataset; annotations are available when a history is viewed.

Database/Build - optional

unspecified (?)

Save

Auto-detect

History

search datasets

Unnamed history

Unnamed history

Annotation (optional)

Add Tags

Save

Cancel

295 MB

1

1: renamed_GCA_029291075.1_E_nip_genomic.fna.gz

Tools

statistics

Show Sections

Statistics on presence-absence of a numeric variable

Summary Statistics for any numerical column

FASTQ Summary Statistics by column

Fasta Statistics display summary statistics for a FASTA file

SeqKit statistics of FASTA/Q files

Compute quality statistics

Genome annotation statistics

Read length statistics from a set of FAST5 files

BAM Index Statistics

Stacks: statistics on stacks found for multiple samples

MiModD Coverage Statistics calculates coverage statistics for a BCF file as generated by the MiModd Variant Calling tool

ProteomIQon PSMStatistics utilizes semi supervised machine learning techniques to integrate search engine scores as well as the mentioned quality scores into one single consensus score.

MapStatistics Extract extended statistics on the features of a map for quality control

zonal statistics over each area

Boxplot of quality statistics

Dr. Disco (classify) Classifies detected break-points in RNA-seq based on corresponding statistics and blacklists

cmstat Summary statistics for covariance model

BAM/SAM Mapping Stats reads mapping statistics for a provided BAM or SAM file.

Fasta Statistics display summary statistics for a FASTA file (Galaxy Version 2.0)

Run Tool

Tool Parameters

FASTA or Multi-FASTA file *

Select a dataset
accepted formats
FASTA dataset to get sta

1: renamed_GCA_029291075.1_E_nip_genomic.fna.gz (as fasta)

Estimated genome size - optional

This parameter is optional. If provided, it will be used for calculating the NG50 statistic. (--genome_size)

Generate gap stats

 No

(--gaps_option)

Additional Options

Email notification

 No

Send an email notification when the job completes.

Attempt to re-use jobs with identical parameters?

 No

This may skip executing jobs that you have already run.

Run Tool

Help

Purpose

Displays the summary statistics for a FASTA file.

Outputs

This tool generates two outputs: a general summary and an optional gap stats file.

The general summary includes the following information:

- Lengths: n50, min, max, median and average
- Number of base pairs: A, C, G, T, N, Total and Total_not_N
- Number of sequences
- GC content

In addition the optional gap stats BED file includes the information about gaps localization.

History

search datasets

Unnamed history

Unnamed history

Annotation (optional)

Add Tags

Save Cancel

295 MB

1

1: renamed_GCA_029291075.1_E_nip_genomic.fna.gz

Tools

statistics

Show Sections

Statistics on presence-absence of a numeric variable

Summary Statistics for any numerical column

FASTQ Summary Statistics by column

Fasta Statistics display summary statistics for a FASTA file

SeqKit statistics of FASTA/Q files

Compute quality statistics

Genome annotation statistics

Read length statistics from a set of FAST5 files

BAM Index Statistics

Stacks: statistics on stacks found for multiple samples

MiModD Coverage Statistics calculates coverage statistics for a BCF file as generated by the MiModd Variant Calling tool

ProteomIQon PSMStatistics utilizes semi supervised machine learning techniques to integrate search engine scores as well as the mentioned quality scores into one single consensus score.

MapStatistics Extract extended statistics on the features of a map for quality control

zonal statistics over each area

Boxplot of quality statistics

Dr. Disco (classify) Classifies detected break-points in RNA-seq based on corresponding statistics and blacklists

cmstat Summary statistics for covariance model

BAM/SAM Mapping Stats reads mapping statistics for a provided BAM or SAM file.

Started tool **Fasta Statistics** and successfully added 1 job to the queue.

It produces this output:

• 2: **Fasta Statistics on data 1: summary stats**

You can check the status of queued jobs and view the resulting data by refreshing the History panel. When the job has been run the status will change from 'running' to 'finished' if completed successfully or 'error' if problems were encountered.

Here is a link to the job: [83e4c30195971270](#)

History

search datasets

Unnamed history

Unnamed history

Annotation (optional)

Add Tags

Save Cancel

295 MB

2

2: **Fasta Statistics on data 1: summary stats**

1: renamed_GCA_02929107
5.1_E_nip_genomic.fna.gz

Upload

Tools

Workflows

Workflow Invocations

Interactive Tools

Visualization

Histories

History Multiview

Datasets

Pages

Libraries

Notifications

More

Tools

[Show Sections](#)

Statistics on presence-absence of a numeric variable

Summary Statistics for any numerical column

FASTQ Summary Statistics by column

Fasta Statistics display summary statistics for a FASTA file

SeqKit statistics of FASTA/Q files

Compute quality statistics

Genome annotation statistics

Read length statistics from a set of FAST5 files

BAM Index Statistics

Stacks: statistics on stacks found for multiple samples

MiModD Coverage Statistics calculates coverage statistics for a BCF file as generated by the MiModd Variant Calling tool

ProteomIQon PSMStatistics utilizes semi supervised machine learning techniques to integrate search engine scores as well as the mentioned quality scores into one single consensus score.

MapStatistics Extract extended statistics on the features of a map for quality control

zonal statistics over each area

Boxplot of quality statistics

Dr. Disco (classify) Classifies detected break-points in RNA-seq based on corresponding statistics and blacklists

cmstat Summary statistics for covariance model

BAM/SAM Mapping Stats reads mapping statistics for a provided BAM or SAM file.

2: Fasta Statistics on data 1: summary stats ok

30 lines 2 columns

format **tabular** database ? size **664 b**
[Preview](#)
[Visualize](#)
[Details](#)
[Edit](#)

Column 1	Column 2
Scaffold L50	3219
Scaffold N50	87067
Scaffold L90	11600
Scaffold N90	19508
Scaffold len_max	557136
Scaffold len_min	4973
Scaffold len_mean	44658
Scaffold len_median	23894
Scaffold len_std	52885
Scaffold num_A	305506036
Scaffold num_T	305603756
Scaffold num_C	164346741
Scaffold num_G	164346396
Scaffold num_N	0
Scaffold num_bp	939802929
Scaffold num_bp_not_N	939802929
Scaffold num_seq	21044
Scaffold GC content overall	34.97
Contig L50	3219
Contig N50	87067
Contig L90	11600
Contig N90	19508
Contig len_max	557136
Contig len_min	4973
Contig len_mean	44658

History

Unnamed history

1.25 GB

2: Fasta Statistics on data 1: summary stats

1: renamed_GCA_02929107
5.1_E_nip_genomic.fna.gz

- Upload
- Tools
- Workflows**
- Workflow Invocations
- Interactive Tools
- Visualization
- Histories
- History Multiview
- Datasets
- Pages
- Libraries
- Recent Exports
- Notifications
- More

Workflows

+ Create + Import

My workflows

Workflows shared with me

Public workflows

Search published workflows by query or use the advanced filtering options

Sort by: Name Update time

Display: [grid icon] [list icon]

Metabolomics: LC-MS data preprocessing with XCMS ☆

mbodova

LC-MS data preprocessing for metabolomics data, follows galaxy training...

Show more ▾

metabolomics LC-MS 1 more...

23 days ago

[link] [download] [Import] [play]

XCMS pre-processing updated for Pulsar ☆

hechth

24 days ago

[link] [download] [Import] [play]

Implementation of 'Galaxy LLM' e-Infra CZ blog ☆

ljocha_ics

about 2 months ago

[link] [download] [Import] [play]

GTN Training: Galaxy 101 For Everyone ☆

demko.cesnet.cz

introduction

introduction

over 1 year ago

[link] [download] [Import] [play]

Funannotate ✎ ☆

demko.cesnet.cz

Structural and functional genome annotation with Funannotate

genome-annotation

over 1 year ago

[link] [download] [Import] [play]

paste_columns ☆

demko.cesnet.cz

over 1 year ago

[link] [download] [Import] [play]

paste_columns ☆

demko.cesnet.cz

over 1 year ago

[link] [download] [Import] [play]

History + ↺ ≡

search datasets ▾ ×

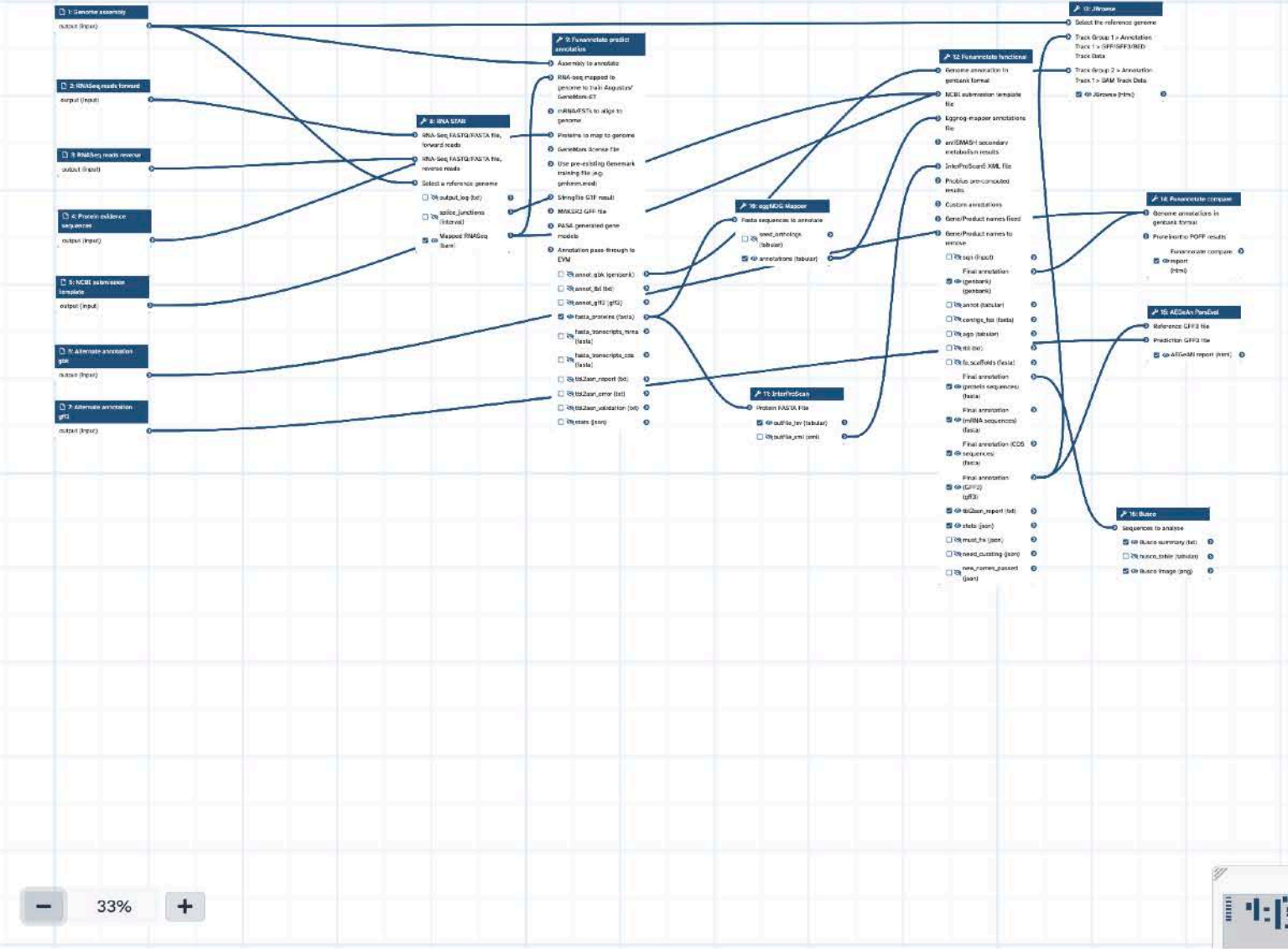
Unnamed history ✎

0 B 0 3 1

This history is empty. You can load your own data or get data from an external source.

Workflow Preview

Download
 Import
 Run
 ✕



About This Workflow

Funannotate - Version 0

Author

demko.cesnet.cz



All published Workflows by demko.cesnet.cz

Creators

Anthony Bretaudeau

Description

Structural and functional genome annotation with Funannotate

Tags

genome-annotation

License

GNU General Public License v3.0 or later

Last Updated

Thursday Oct 17th 14:24:54 2024 GMT+2

Sharing

Use the following link to share preview of this workflow: <https://usegalaxy.cz/published/workflow?id=932f4901960a595a> . Manage sharing settings here.

33%



Welcome to Galaxy Training!

<https://training.galaxyproject.org/>

Collection of tutorials developed and maintained by the worldwide Galaxy community

Galaxy for Scientists

We have separated the tutorials into several categories based on field and technology. We are exploring other ways to organise the tutorials going forward!

Start Here

Topic	Tutorials
Introduction to Galaxy Analyses	15
Using Galaxy and Managing your Data	27

Not sure where to start?

Try the NGS Basics Learning Path!

[Start Learning](#)

Scientific Fields

Topic	Tutorials
Climate	14
Computational chemistry	9

Quickstart

Learning Pathways



Galaxy for SysAdmins



Galaxy for Developers



Galaxy for Teachers



Upcoming Events

Check out upcoming events around the Galaxy!

May 18 – 22, 2026

Galaxy Training Academy 2026

October 12 – 16, 2026

Galaxy Beyond Basics: Mastering Workflows, Automation, and Scalability



Not sure where to start?

Try the NGS Basics Learning Path!

[Start Learning](#)

We

Colle

Gal

We ha
explor

Sta

Topi

Intro

Usin

No

Try

Sci

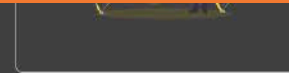
Topi

Clim

Corr

Scientific Fields

Topic	Tutorials
Climate	14
Computational chemistry	9
SARS-CoV-2	9
Foundations of Data Science	50
Digital Humanities	4
Ecology	28
Evolution	10
FAIR Data, Workflows, and Research	25
Genome Annotation	25
Imaging	17
Materials Science	1
Microbiome	28
One Health	9
Plants	10
Statistics and machine learning	35
Visualisation	5



Upcoming Events

Check out upcoming events around the Galaxy!

May 18 – 22, 2026

Galaxy Training Academy 2026

October 12 – 16, 2026

Galaxy Beyond Basics: Mastering Workflows, Automation, and Scalability

[See all events](#)

[Add your event!](#)

The latest GTN news



Read about new tutorials, features, events and more!

Apr 1, 2026

BioNT Brings Galaxy Training to German, Spanish, and Italian Learners

Feb 23, 2026

New Paper! Ten common misconceptions about Galaxy (and why they are wrong!)

Nov 13, 2025

Save the date: The Galaxy Training Academy 2026 will happen from May 18th to May 22th

Oct 8, 2025

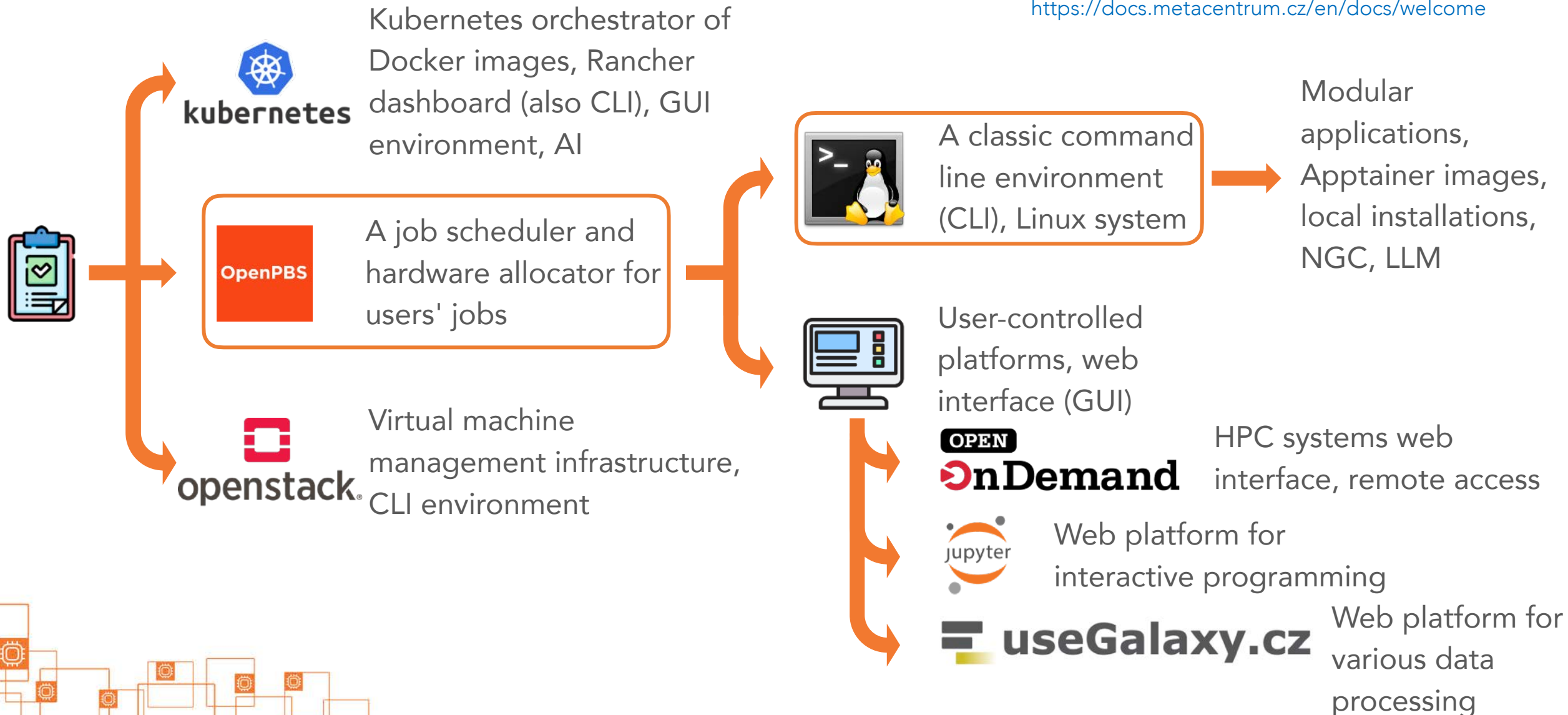
GTN joined the TEACH V to talk about education across Communities at Helmholtz

Sep 12, 2025

We want you to help us organize the Galaxy Training Academy 2026

How to use it

<https://docs.metacentrum.cz/en/docs/welcome>



- Grid computing (OpenPBS scheduler)
 - Maximum flexibility
 - Long batch jobs (days/weeks)
 - Interactive jobs
 - Parallel and multi-node computing
 - GPU acceleration (NVIDIA H100, 96 GB)

```

cesnet
[BOOKWORM]vorel@skirit:~$
[BOOKWORM]vorel@skirit:~$

Your Home Directories
-----
Avail |   DATA   |   FILES   |Backup|Storage
space|quota used|quota used|      |
-----|-----|-----|-----|-----
313T  none 375G  none 4.88k  snap  brno12-cerit
50T   4.29T 60.3G  2.5M 18.6k  snap  brno2
17T   3.32T 24k   1M   9      full  budejovice1
30T   10.4G 12k   none 8      -     liberec3-tul
353T  4.29T 1.73T  3M  192k  snap  plzen1
60T   7.34G 44k   none 18     -     praha2-natur
12T   5.36T 941G  1.01M 113k  full  praha5-elixir
55T   7.34G 28k   none 11    full  pruhonice1-ibot
145T  2.14T 5.16G  2M  41.2k  snap  vestec1-elixir

CPUs
-----
Free  Used  Total  Centre
7568 32132 39700  META

Contact
-----
E-mail:  meta@cesnet.cz
Web:    http://www.metacentrum.cz/

Tip of the day:  Have you ever heard about a Chipster? Chipster is a web-based interactive analysis software for NGS and microarray data. More information at https://docs.metacentrum.cz/en/docs/related/chipster.

[BOOKWORM]vorel@skirit:~$

```

Frontend servers (login nodes)

<https://docs.metacentrum.cz/en/docs/computing/infrastructure/frontends>

- Primary access point
- SSH access via password (creating a Kerberos ticket)
 - SSH key authentication is not supported



Geographically separated servers mean better distribution of users and load, which makes the servers more stable and provides an alternative during an outage

Frontend address	Aliased as	Native home	OS	Physically located in	Note
charon.nti.tul.cz	charon.metacentrum.cz	/storage/liberec3-tul	Debian 12	Liberec	
elmo.elixir-czech.cz	elmo.metacentrum.cz	/storage/praha5-elixir	Debian 12	Praha	Elixir users
nympha.meta.zcu.cz	nympha.metacentrum.cz, nympha.zcu.cz, minos.zcu.cz, minos.meta.zcu.cz, alfrid.meta.zcu.cz	/storage/plzen1	Debian 12	Plzen	
metafzu.fzu.cz	metafzu.metacentrum.cz	/storage/praha1	Debian 12	Praha	dedicated for FZU users
oven.metacentrum.cz		/storage/brno2	Debian 12	Brno	oven node only
perian.grid.cesnet.cz	perian.metacentrum.cz, onyx.metacentrum.cz	/storage/brno2	Debian 12	Brno	
skirit.ics.muni.cz	skirit.metacentrum.cz	/storage/brno2	Debian 12	Brno	
tarkil.grid.cesnet.cz	tarkil.metacentrum.cz	/storage/praha1	Debian 12	Praha	
tilia.ibot.cas.cz	tilia.metacentrum.cz	/storage/pruhonice1-ibot	Debian 12	Pruhonice	
zenith.cerit-sc.cz	zenith.metacentrum.cz	/storage/brno12-cerit	Debian 12	Brno	

Disk storage arrays

- Several geographically separated disk arrays (storages) are available

<https://docs.metacentrum.cz/en/docs/computing/infrastructure/mount-storages>

- Designed for so-called live data

<https://docs.metacentrum.cz/en/docs/data/storage-department>

- Backups of valuable primary data and results preferably to S3/NRP object storage

<https://docs.du.cesnet.cz/en/docs/object-storage-s3/s3-service>

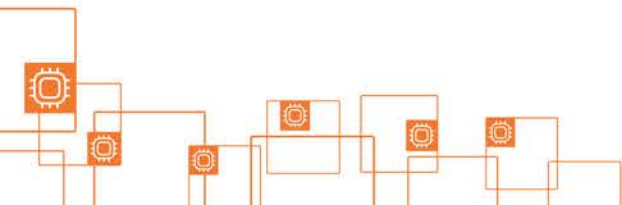
- Available to all users
- *s3cmd/s5cmd* commands are available for manipulating data on S3 storage
- Need to request the creation of an S3 bucket ('folder') and create a configuration file




Disk storage arrays


- Several geographically separated disk arrays (storages) are available
- Designed for so-called
 - Backups of valuable project data
 - Available to all users
 - *s3cmd/s5cmd* commands
 - Need to request the access

Server	Directory	Backup Class	Note
storage-brno2.metacentrum.cz	/storage/brno2/	2	
storage-brno11-elixir.metacentrum.cz	/storage/brno11-elixir/	2	dedicated to ELIXIR-CZ
storage-brno12-cerit.metacentrum.cz	/storage/brno12-cerit/	2	
storage-plzen1.metacentrum.cz	/storage/plzen1/	2	
storage-plzen4-ntis.metacentrum.cz	/storage/plzen4-ntis/	3	dedicated to iti/kky groups
storage-praha2-natur.metacentrum.cz	/storage/praha2-natur/	0	
storage-praha6-fzu.metacentrum.cz	/storage/praha6-fzu/	0	
storage-praha5-elixir.metacentrum.cz	/storage/praha5-elixir/	3	
storage-budejovice1.metacentrum.cz	/storage/budejovice1/	3	
storage-liberec3-tul.metacentrum.cz	/storage/liberec3-tul/	0	
storage-pruhonice1-ibot.metacentrum.cz	/storage/pruhonice1-ibot/	3	
storage-vestec1-elixir.metacentrum.cz	/storage/vestec1-elixir/	2	also /storage/praha1/

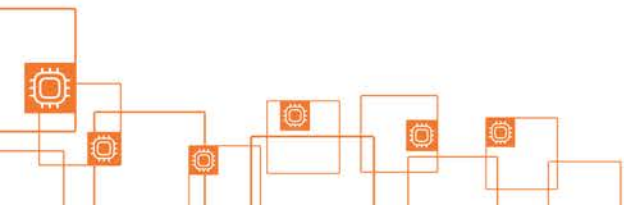
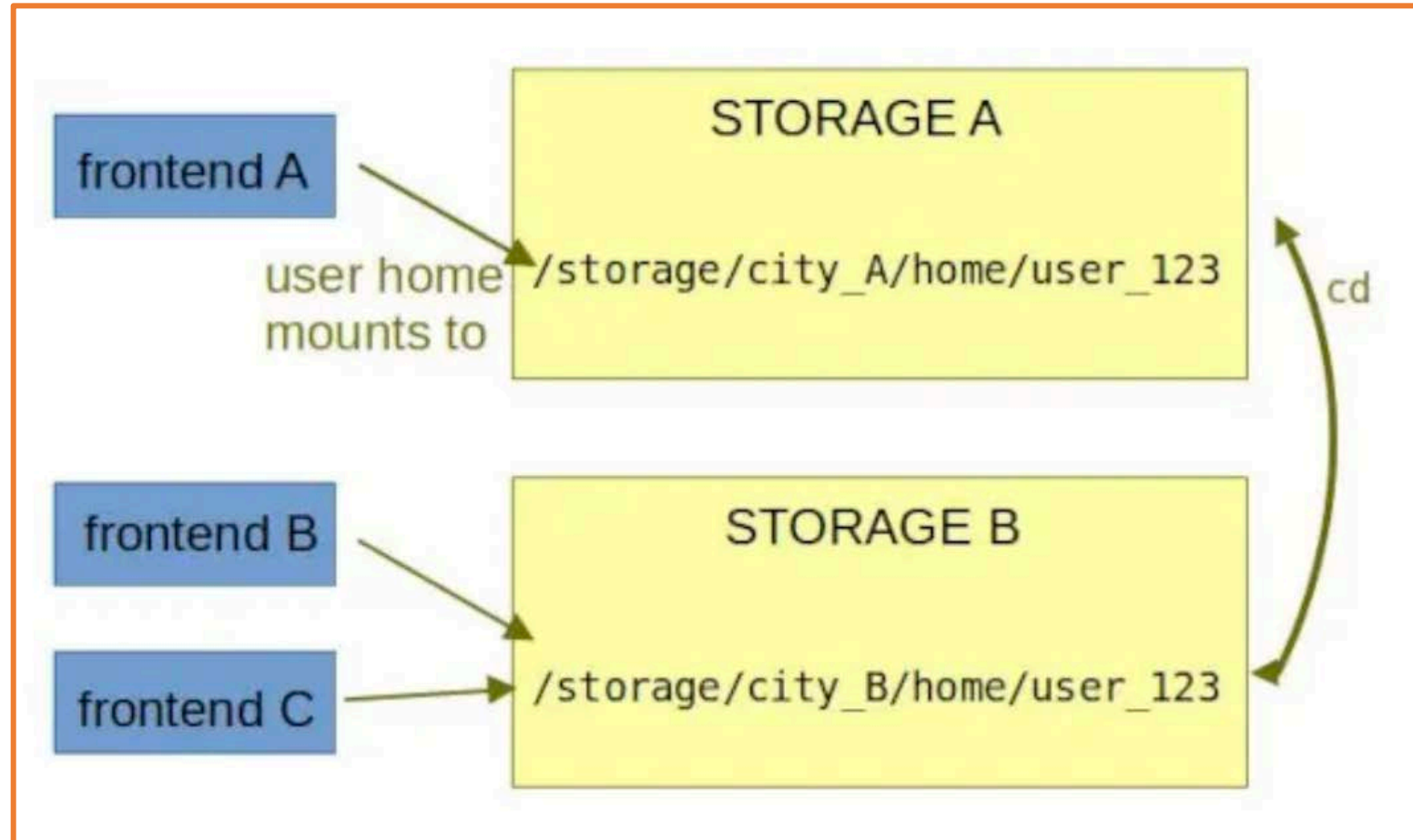


The concept of multiple home directories and disk storages

 Geographically separated servers mean better distribution of users and load, which makes the servers more stable and provides an alternative during an outage

 All frontend servers can also be reached in the domain 'metacentrum.cz'

 Although disk storage servers cannot be accessed interactively, it is possible to send specific commands to them



Batch job script example

```
#!/bin/bash
#PBS -q default@pbs-m1.metacentrum.cz
#PBS -l walltime=24:0:0
#PBS -l select=1:ncpus=8:mem=100gb:scratch_ssd=50gb:mpiprocs=1:omphreads=8
#PBS -N my_job_name
#PBS -M my_email@xyz
#PBS -m e

ulimit -s unlimited
trap "clean_scratch" TERM EXIT
export TMPDIR=$SCRATCHDIR

# test if a scratch directory exists
# variable SCRATCHDIR is set automatically
test -n "$SCRATCHDIR" || { echo >&2 "Variable SCRATCHDIR is not set!"; exit 1; }

# set a DATADIR variable
DATADIR="/storage/brno12-cerit/home/vorel/data/"

# copy input file "data.fa" to the scratch directory
cp $DATADIR/data.fa $SCRATCHDIR || { echo >&2 "Error while copying input file(s)!"; exit 2; }

# move into the scratch directory
cd $SCRATCHDIR

# load a module for your application
module add blast-plus/blast-plus-2.12.0-gcc-8.3.0-ohlv7t4

# run the calculation
# do not forget to use reserved CPUs by '-num_threads' flag
# variable PBS_NCPUS is a number of CPUs requested for the entire job
blastp -query data.fa <other_parameters> -num_threads $PBS_NCPUS -out results.txt

#copy results
cp results.txt $DATADIR || export CLEAN_SCRATCH=false
```

- Define HW resources (-l), queue (-q), and walltime (-l), set the job name (-N) and email alert (-m)
- You can define as many variables as you want
- Available modules can be listed by command `module avail *key_word*` on any frontend
- The scratch directory will be cleaned automatically



Interactive job

- The opposite of batch jobs (waiting for the user's input...)
- Best choice for test calculations (which should not be run directly on frontends)
- An interactive job is requested by the qsub command with the `-I` (uppercase "i") option

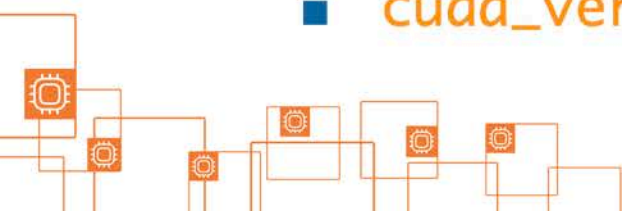
<https://docs.metacentrum.cz/en/docs/computing/run-basic-job#interactive-job>

```
(BUSTER)vorel@skirit:~$ qsub -I -l select=1:ncpus=4:mem=50gb:scratch_local=30gb -l walltime=1:00:00
qsub: waiting for job 11405230.meta-pbs.metacentrum.cz to start
qsub: job 11405230.meta-pbs.metacentrum.cz ready

vorel@zenon31:~$ cd $SCRATCHDIR
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$ module add orca/orca-5.0.1-intel-19.0.4-bnofsgq
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$ module list
Currently Loaded Modulefiles:
  1) metabase                2) openmpi/openmpi-4.0.4-intel-19.0.4-gpu-xri6uan  3) orca/orca-5.0.1-intel-19.0.4-bnofsgq
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$ ...time for coffee...
-bash: ...time: command not found
vorel@zenon31:/scratch.ssd/vorel/job_11405230.meta-pbs.metacentrum.cz$ orca < input > output
```



- **GPU acceleration for significant speedup of calculations**
- ~140 nodes, ~500 GPU cards (GTX 1080Ti - H100 100GB)
- Requires application with GPU support
- Maximum eight GPU cards on a single node, typically two or four
- Special DGX cluster with eight Nvidia H100 80GB GPU cards
 - Grant competition
- Specific parameters
 - `gpu_mem` (minimum amount of memory on the card)
 - `gpu_cap` (a minimal version of GPU architecture)
 - `cuda_version` (version of CUDA installed on the node)



GPU acceleration

```
#!/bin/bash
#PBS -N run1_dorado_mod
#PBS -l select=1:ncpus=4 ngpus=1:mem=30gb:scratch_ssd=250gb:gpu_mem=20gb:gpu_cap=compute_80
#PBS -l walltime=8:0:0

cd $SCRATCHDIR

scp storage-brno12-cerit.metacentrum.cz:~/Metylace_hemonch/RUN_1_IRE/02_run1_IRE_reads.pod5 $SCRATCHDIR
scp -r storage-brno12-cerit.metacentrum.cz:~/Metylace_hemonch/SOFT/dorado_0.9.1/models/dna_r9.4.1_e8_sup@v3.3* $SCRATCHDIR
cp /storage/brno12-cerit/home/vorel/Metylace_hemonch/Haemonch_refer/haemonchus_contortus.PRJEB506.WBPS19.genomic.fa $SCRATCHDIR

export PATH=/storage/brno12-cerit/home/vorel/Metylace_hemonch/SOFT/dorado_0.9.1/bin/:$PATH
export LD_LIBRARY_PATH=/storage/brno12-cerit/home/vorel/Metylace_hemonch/SOFT/dorado_0.9.1/lib/:$LD_LIBRARY_PATH
module add minimap2/2.22

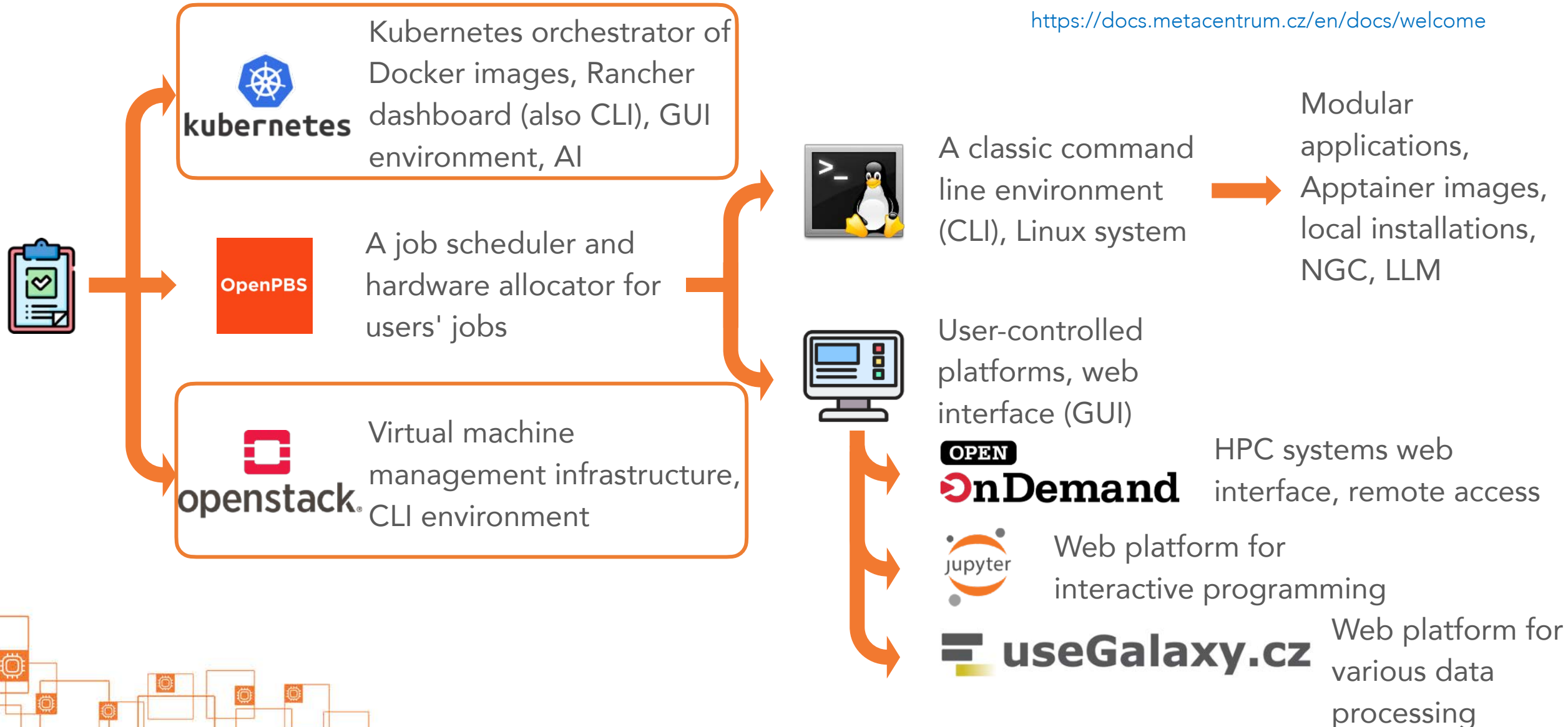
dorado basecaller dna_r9.4.1_e8_sup@v3.3 ./02_run1_IRE_reads.pod5 -v -x cuda:all --min-qscore 5 --trim all \
--reference haemonchus_contortus.PRJEB506.WBPS19.genomic.fa --modified-bases 5mCG_5hmCG > 03_dorado_run1_IRE.bam
```

- `gpu_cap` (a minimal version of GPU architecture)

- `cuda_version` (version of CUDA installed on the node)

How to use it

<https://docs.metacentrum.cz/en/docs/welcome>

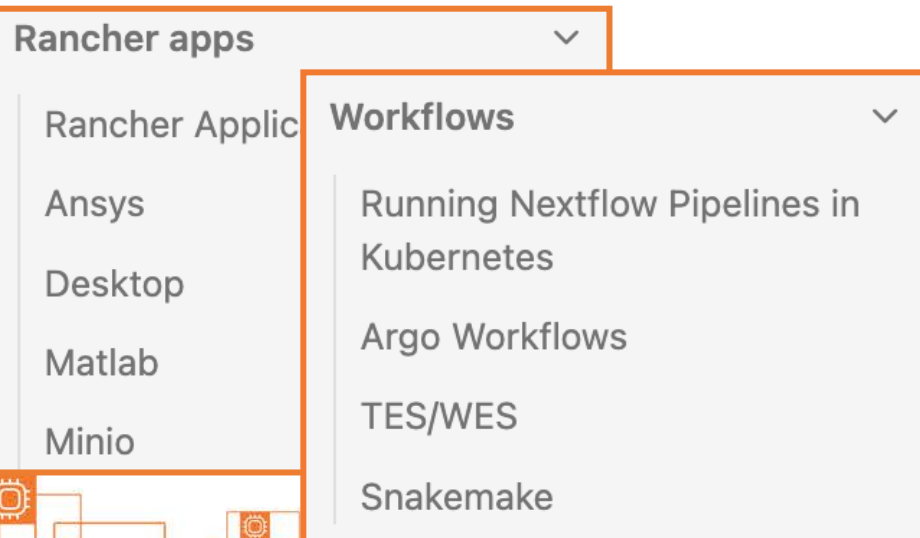


■ Kubernetes

<https://docs.cerit.io/en/docs/platform/overview>

- Open-source container orchestration engine that automates the deployment, scaling and management of containerised applications
- A Docker image is an essential part of container infrastructure. It contains an application and all its necessary components, such as libraries and other files

<https://docs.cerit.io/en/docs/news>

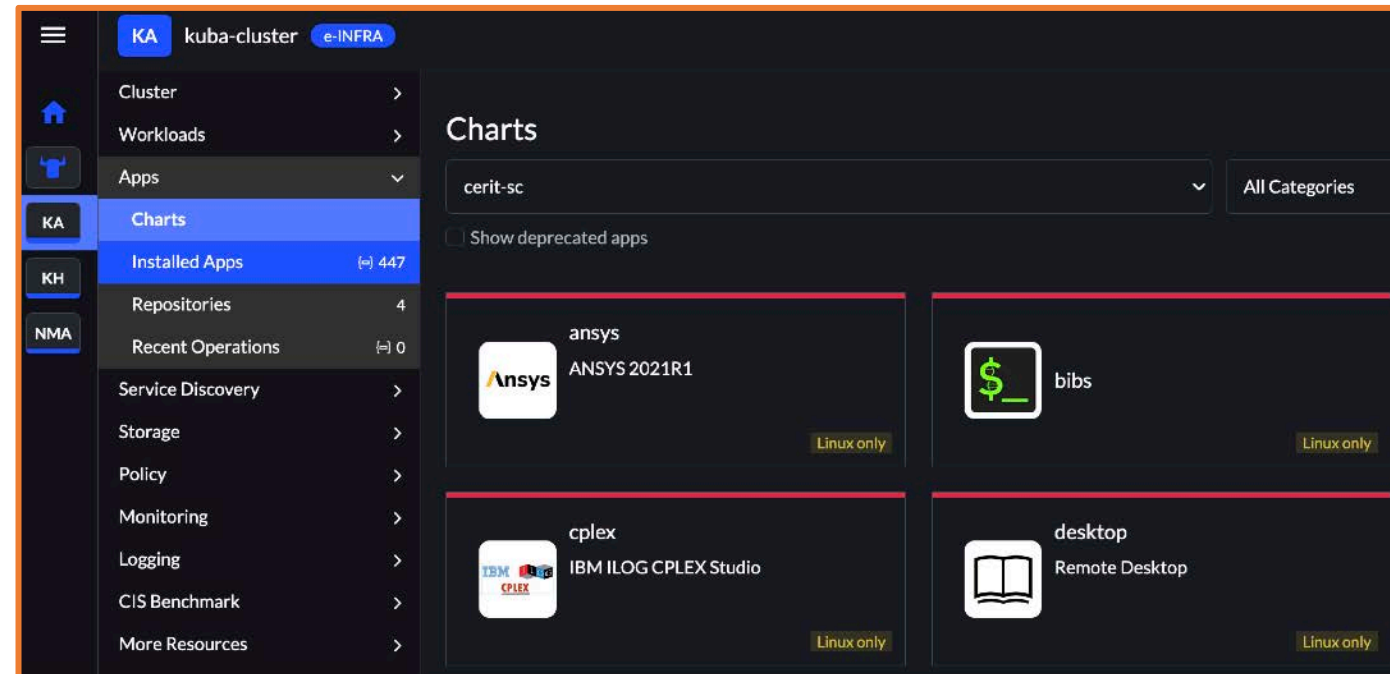


Rancher apps

- Rancher Application
- Ansys
- Desktop
- Matlab
- Minio

Workflows

- Running Nextflow Pipelines in Kubernetes
- Argo Workflows
- TES/WES
- Snakemake



KA kuba-cluster e-INFRA

- Cluster
- Workloads
- Apps
- Charts
- Installed Apps (447)
- Repositories (4)
- Recent Operations (0)
- Service Discovery
- Storage
- Policy
- Monitoring
- Logging
- CIS Benchmark
- More Resources

Charts

cerit-sc All Categories

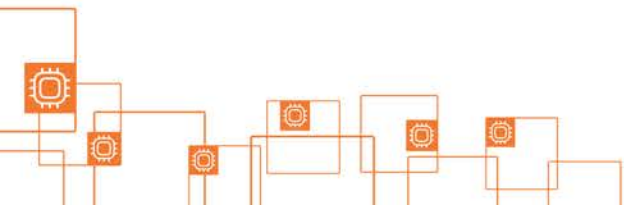
Show deprecated apps

- ansys ANSYS 2021R1 (Linux only)
- bibs (Linux only)
- cplex IBM ILOG CPLEX Studio (Linux only)
- desktop Remote Desktop (Linux only)

■ Compute cloud

<https://docs.platforms.cloud.e-infra.cz/en/docs>

- Runs on OpenStack (allows users to create and manage their own cloud computing resources)
- 17 computational clusters with 8,968 cores, 96 GPU cards, and 178 TB RAM
- Virtual machines (CLI) are accessed using SSH keys



<https://www.cesnet.cz/>

<https://www.metacentrum.cz/>

<https://www.e-infra.cz/>

<https://docs.metacentrum.cz/>

- Free and immediately accessible, no grant proposals
- All we ask: acknowledge e-INFRA CZ (ID:90254) in your publications
- For everyone, no programming skills required
- Scales with your needs, from a small interactive session to large and long-term parallel jobs
- Hundreds of tools included, ready to use
- Dedicated user support and step-by-step documentation

vorel@cesnet.cz meta@cesnet.cz