



Central European Institute of Technology
BRNO | CZECH REPUBLIC

The logo for MUNI, consisting of the word "MUNI" in a large, white, sans-serif font, positioned in the upper right corner of the slide. The background of the slide is a photograph of a modern building with a curved glass facade and a metal framework, partially obscured by large, semi-transparent green and white hexagonal shapes.

MUNI

Tomáš Svoboda

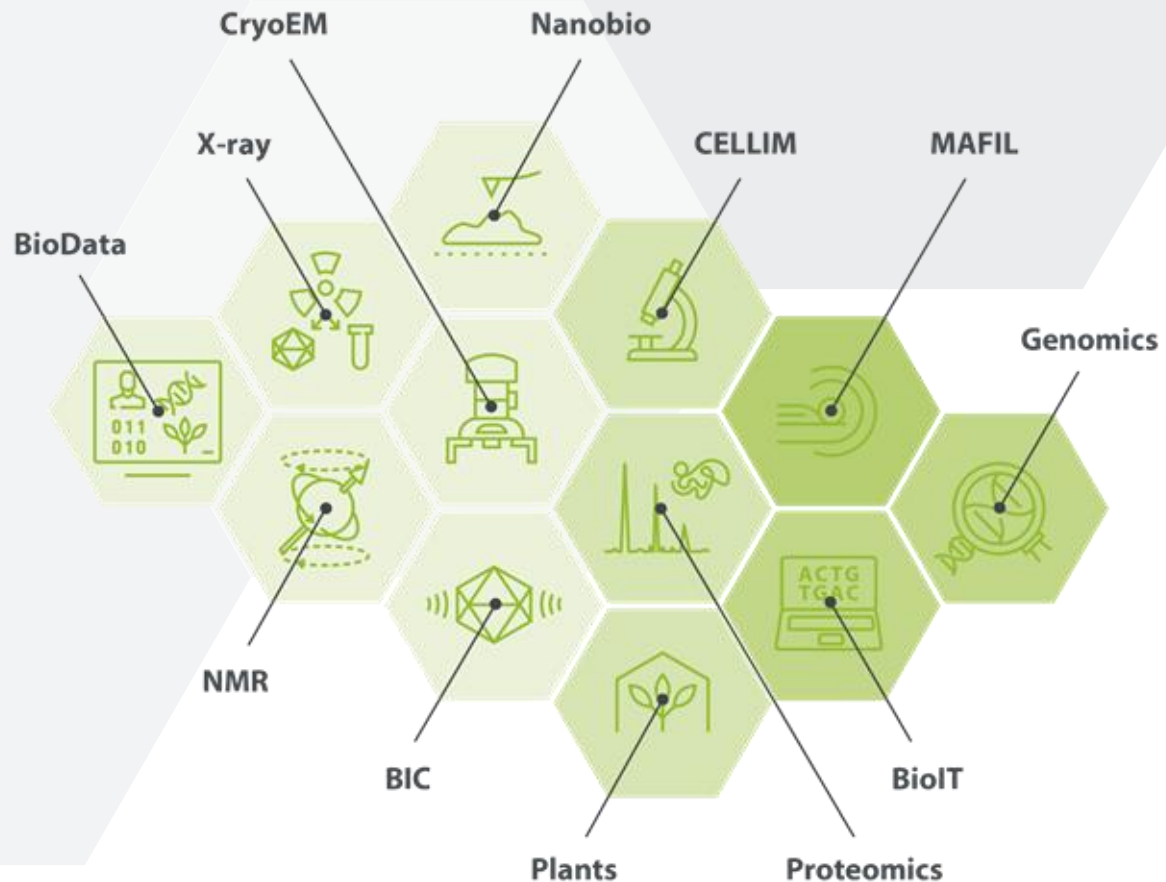
System Onedata: od zkušenosti z CEITECu po službu e-INFRA CZ

Seminář výkonného počítání MetaCentrum
na Konferenci CESNET 2026, 1. 10. 2026

Konsorcium CEITEC



Sdílené laboratoře na CEITEC MU



BIC	Biomolecular Interactions and Crystallization
BioData	Biological Data Management and Analysis
BioIT	Bioinformatics
CELLIM	Cellular Imaging
CryoEM	Cryo-electron Microscopy and Tomography
Genomics	Genomics
MAFIL	Multimodal and Functional Imaging Laboratory
Nanobio	Nanobiotechnology
NMR	Josef Dadok National NMR Centre
Plants	Plant Sciences
Proteomics	Proteomics

CF BioData - motivace

- Jednotný přehled o vznikajících datech
- Zabránění ztrátě dat
- Evidence metadat
- Předávání dat uživatelům
 - Předchozí stav: ukládání na USB flash a komerční cloudy
 - Možnost uložení velkých dat pro jejich odkázání v publikacích
- Předávání dat do repozitářů

FAIRification of CEITEC Data



Use-case CF CELLIM

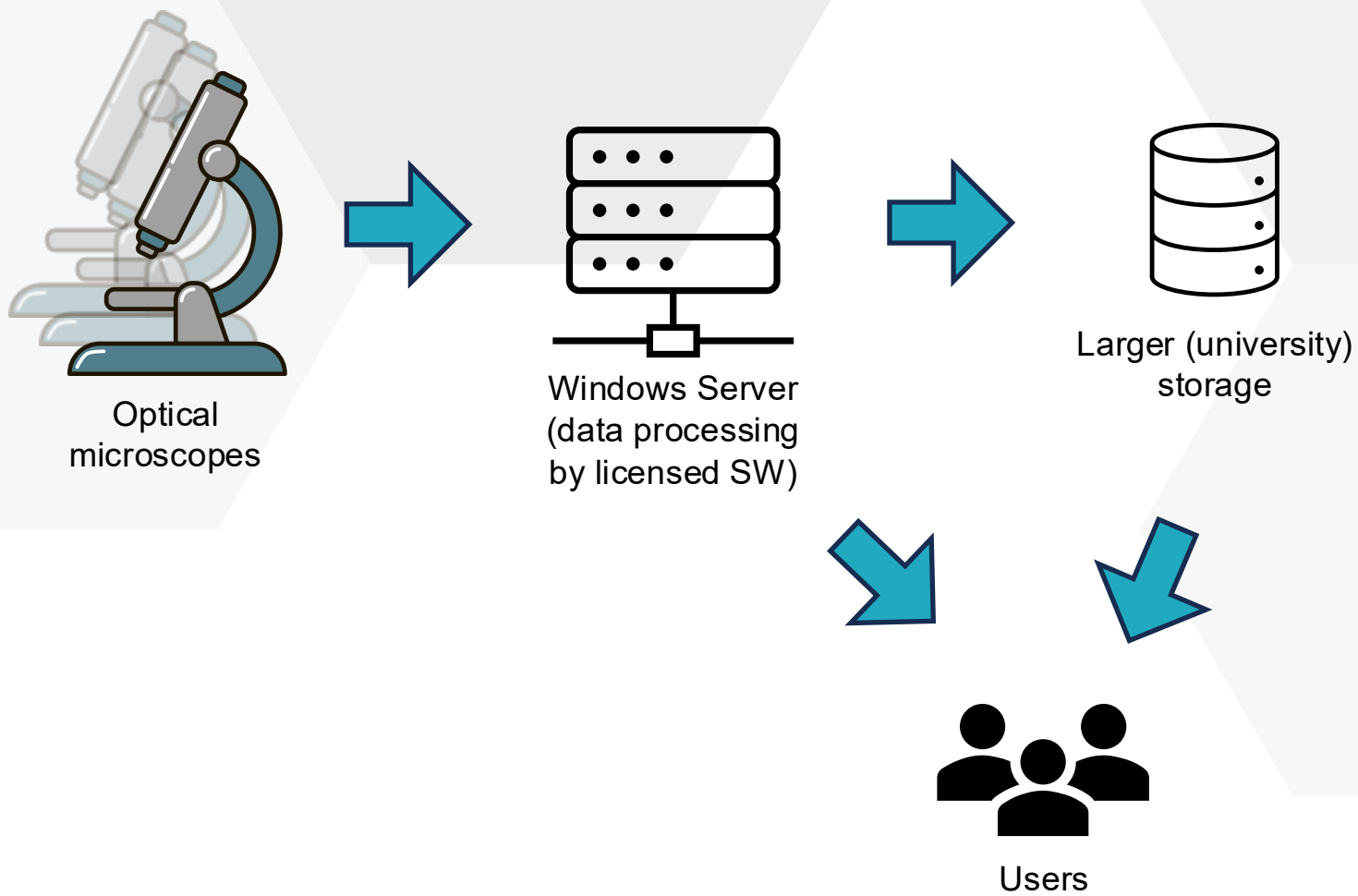
CEITEC MU

CF CELLIM (Core Facility Cellular Imaging)

- Poskytuje služby v oblasti světelné mikroskopie, přípravě vzorku, získávání mikroskopických dat a obrazové analýze dat
- Vybavení laboratoře
 - Různé druhy světelných mikroskopů
 - Stanice pro zpracování dat

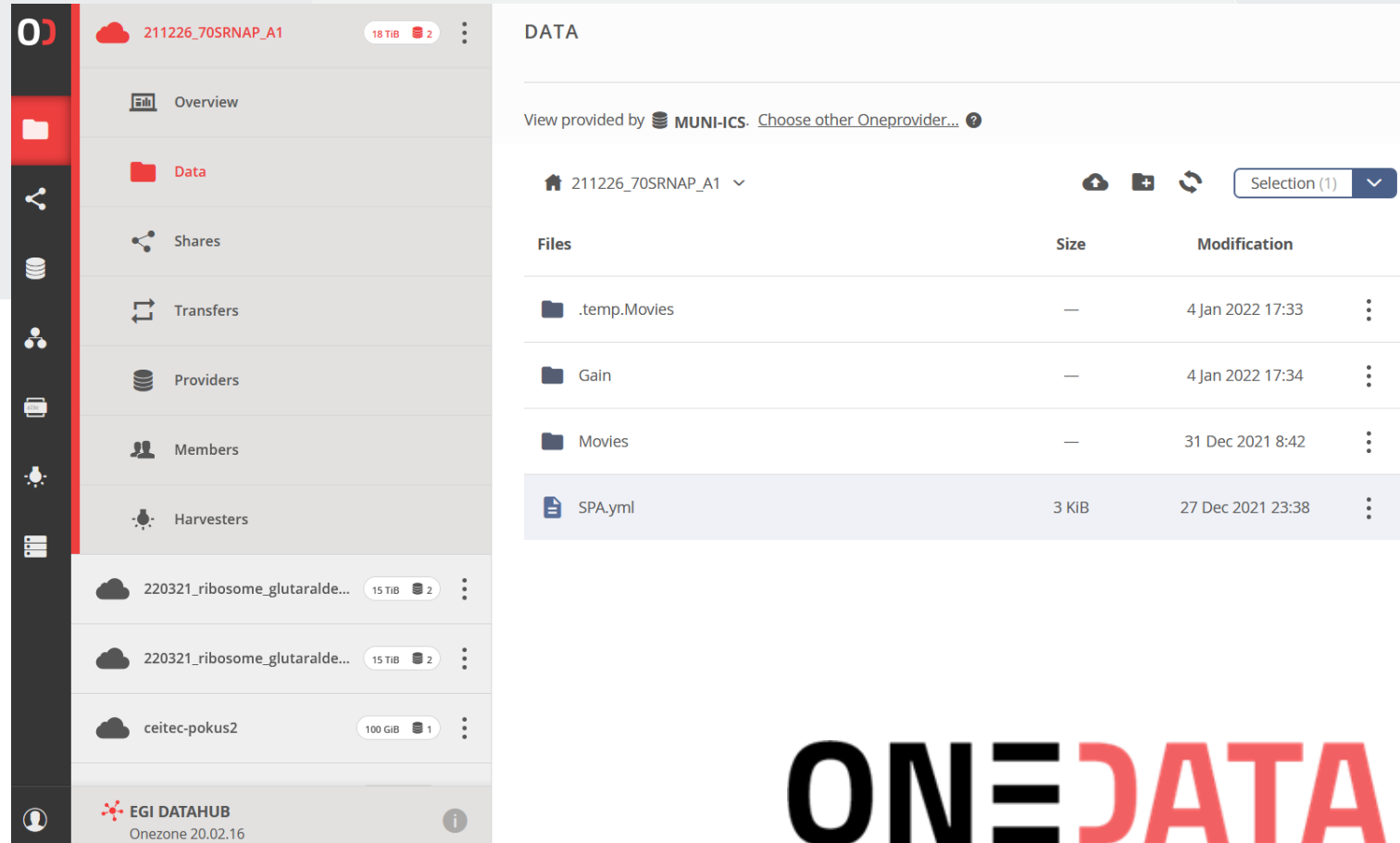


Příklad užití v CF CELLIM



Onedata I

- Systém pro správu velkých dat nad distribuovanými úložišti
- Poskytuje jednotný souborový prostor
 - Umí do něj integrovat úložiště různých technologií (S3, NFS, ...)
- Uzpůsoben pro:
 - Velká data a náročné výpočty
 - Vědecká data (podpora metadat)
- Podpora různých přístupových rozhraní
 - Uživatelsky přívětivé webové rozhraní
 - Klienti pro Linux i Windows (beta)
 - REST API, Python lib (pro workflows)
 - S3



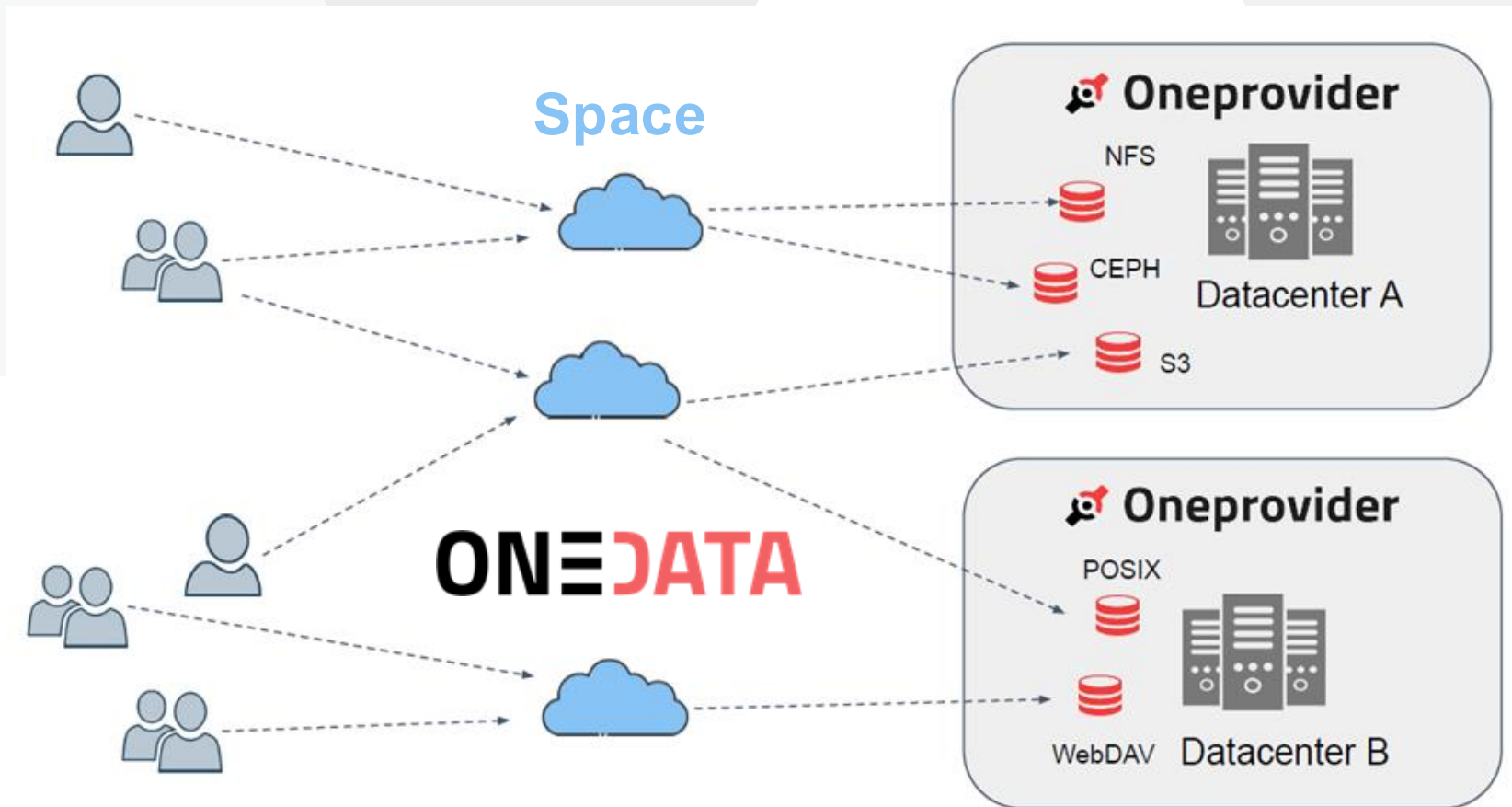
The screenshot displays the Onedata web interface. On the left is a dark sidebar with navigation icons and labels: Overview, Data, Shares, Transfers, Providers, Members, Harvesters, and a bottom section for EGI DATAHUB Onezone 20.02.16. The main panel shows a view provided by MUNI-ICS. It features a dropdown menu for '211226_70SRNAP_A1' and a table of files. The table has columns for Files, Size, and Modification. The files listed are .temp.Movies, Gain, Movies, and SPA.yml.

Files	Size	Modification
.temp.Movies	—	4 Jan 2022 17:33
Gain	—	4 Jan 2022 17:34
Movies	—	31 Dec 2021 8:42
SPA.yml	3 KIB	27 Dec 2021 23:38

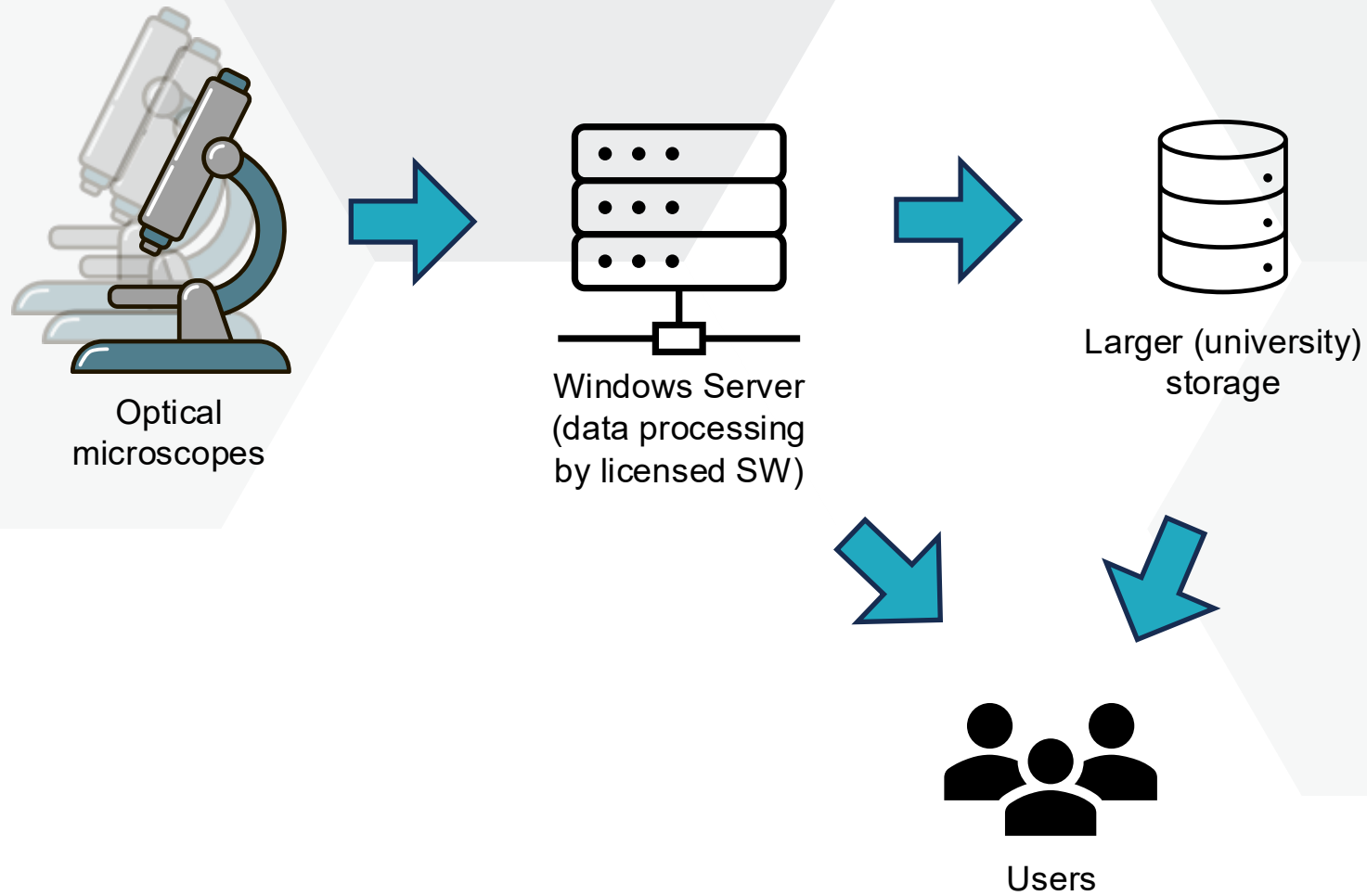
ONEDATA

Onedata II

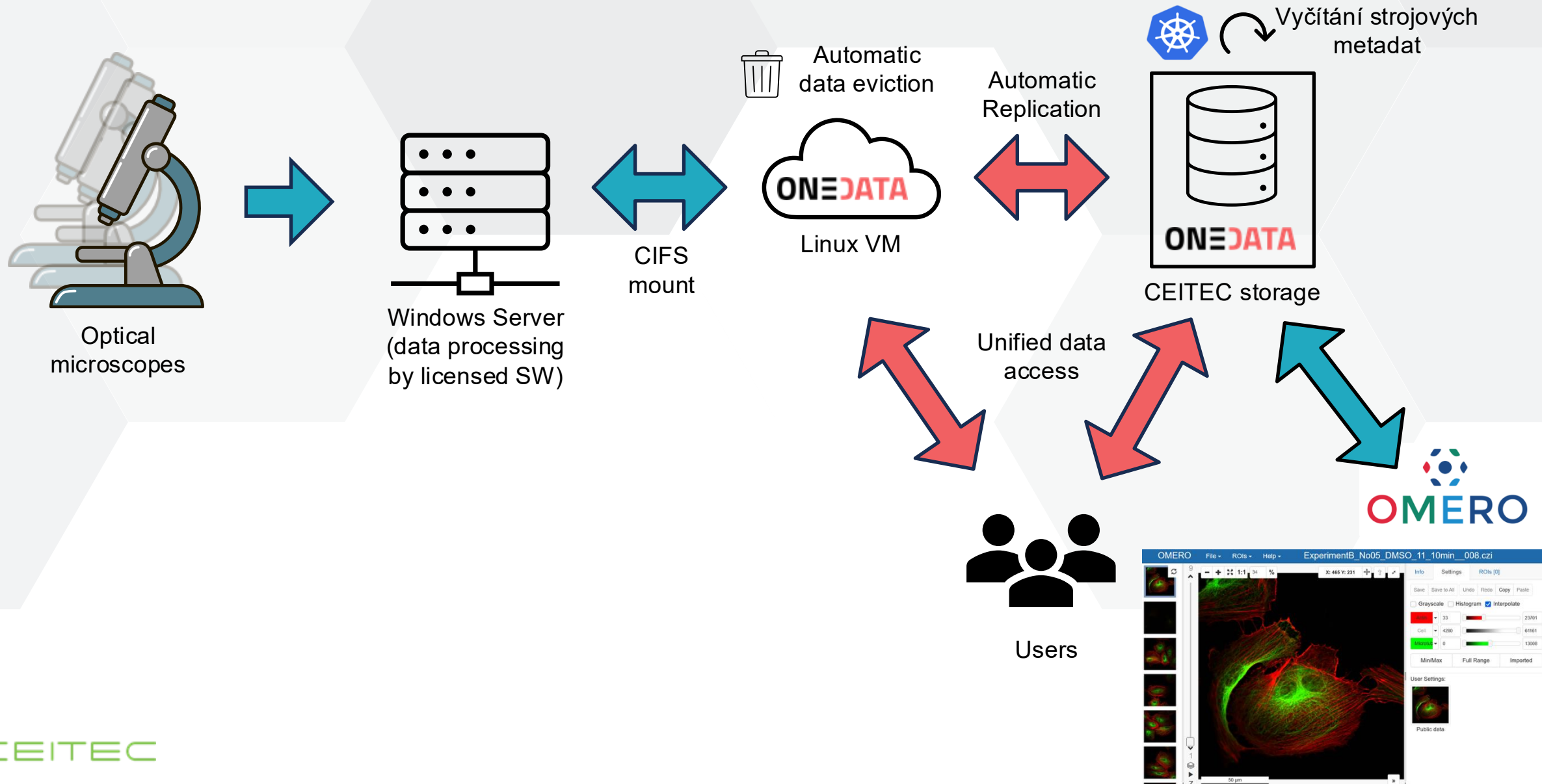
- **Space** – virtuální obálka na data
- Jednotný přístup k datům
- POSIX-like souborový systém
- Různé technologie a geografické lokality



Příklad užití v CF CELLIM



Příklad užití v CF CELLIM s použitím Onedata





CEITEC CELLIM

ACQUIFER

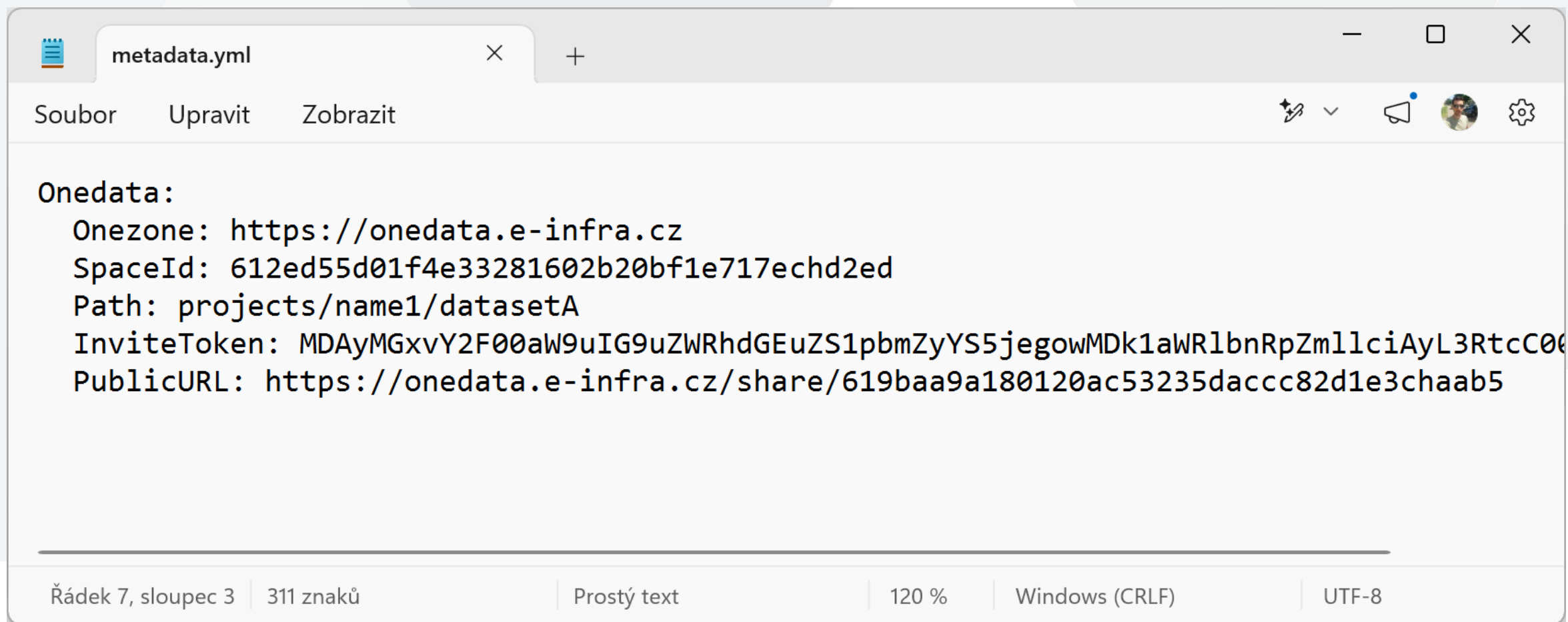
File Explorer window showing the contents of the folder 221006PM.

Address bar: << onedata > scientific_group > study_name > 221006PM

Search: Search 221006PM

Name	Date modified	Type	Size
.running	3/1/2023 6:44 PM	RUNNING File	0 KB
allChannels-63xOIL.czi	10/6/2022 10:44 A...	CZI File	21,095 KB
Brightfield-Colorcamera-63x.czi	10/6/2022 10:45 A...	CZI File	6,303 KB
fluorescence-Apotome-raw-63xOIL.czi	10/6/2022 10:50 A...	CZI File	108,886 KB
metadata.yml	3/1/2023 6:48 PM	YML File	1 KB

5 items | State: Shared



The image shows a code editor window with a single tab titled 'metadata.yml'. The editor has a menu bar with 'Soubor', 'Upravit', and 'Zobrazit'. The main area contains the following text:

```
Onedata:  
  Onezone: https://onedata.e-infra.cz  
  SpaceId: 612ed55d01f4e33281602b20bf1e717echd2ed  
  Path: projects/name1/datasetA  
  InviteToken: MDAyMGxvY2F00aW9uIG9uZWRhdGEuZS1pbmZyYS5jegowMDk1aWR1bnRpZm1lciAyL3RtcC00  
  PublicURL: https://onedata.e-infra.cz/share/619baa9a180120ac53235dacc82d1e3chaab5
```

The status bar at the bottom displays: 'Řádek 7, sloupec 3 | 311 znaků | Prostý text | 120 % | Windows (CRLF) | UTF-8'.



Rare_ring_conformations



<https://datahub.egi.eu/share/d884d196f712515f8e981f1183cc866fch514c>



Description



Files



/ Rare_ring_conformations ▾



Files

Jump to prefix...



Size

Modified 
(content)



input_data

586.1 GiB

20 May 2024
17:08:18



input_data_small

2 GiB

20 May 2024
16:45:33



output_data.zip

51.8 MiB

3 Jul 2024
15:21:02



workflow.zip

7.3 MiB

3 Jul 2024
15:31:07



e-INFRA CZ – Onezone

Anonymní prohlížení

onedata.e-infra.cz/ozw/onezone/i#/public/shares/762621205cf88c3bd0cc

☆

T

6



Proptimus Prime

↻

i

https://onedata.e-infra.cz/share/762621205cf88c3bd0cc1acabc3c53dachcceb

<https://doi.org/10.58074/evfc-5s44>

 Description

 Files

Analysis of structural errors in AlphaFold DB v4 and v6

This publicly accessible share contains supporting information for the article *Analysis of structural errors in the AlphaFold DB v4 and v6*.

Workflow source code

proptimus_prime.zip: An archive that contains a snapshot of the workflow used to detect and repair structural errors.

Errors information

detected_and_repaired_errors.pdf: Table with the number of errors found and corrected, divided according to their confidence.
structural_errors.json: Information about detected structural errors in JSON format.

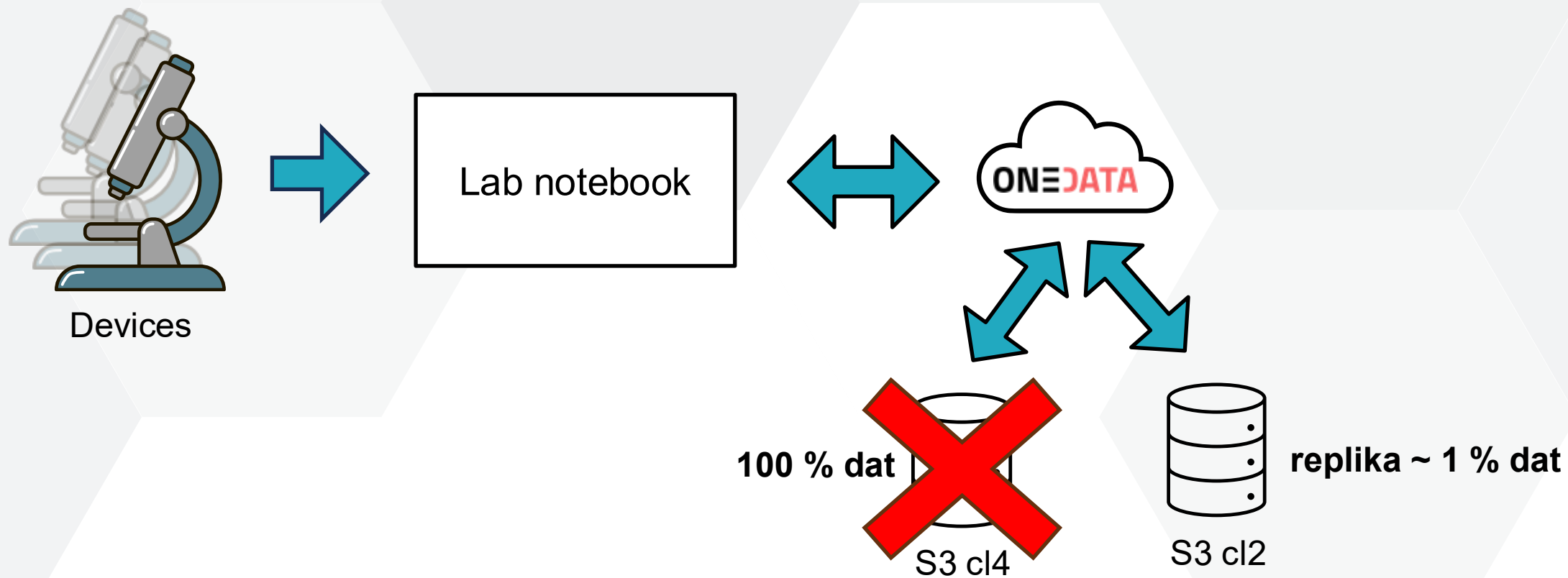
Analysis

analysis.pdf: Results of the analysis of structural error distribution, together with an examination of errors in the highest confidence regions of the prediction.

CF NANO

CEITEC VUT

Příklad užití v CF NANO





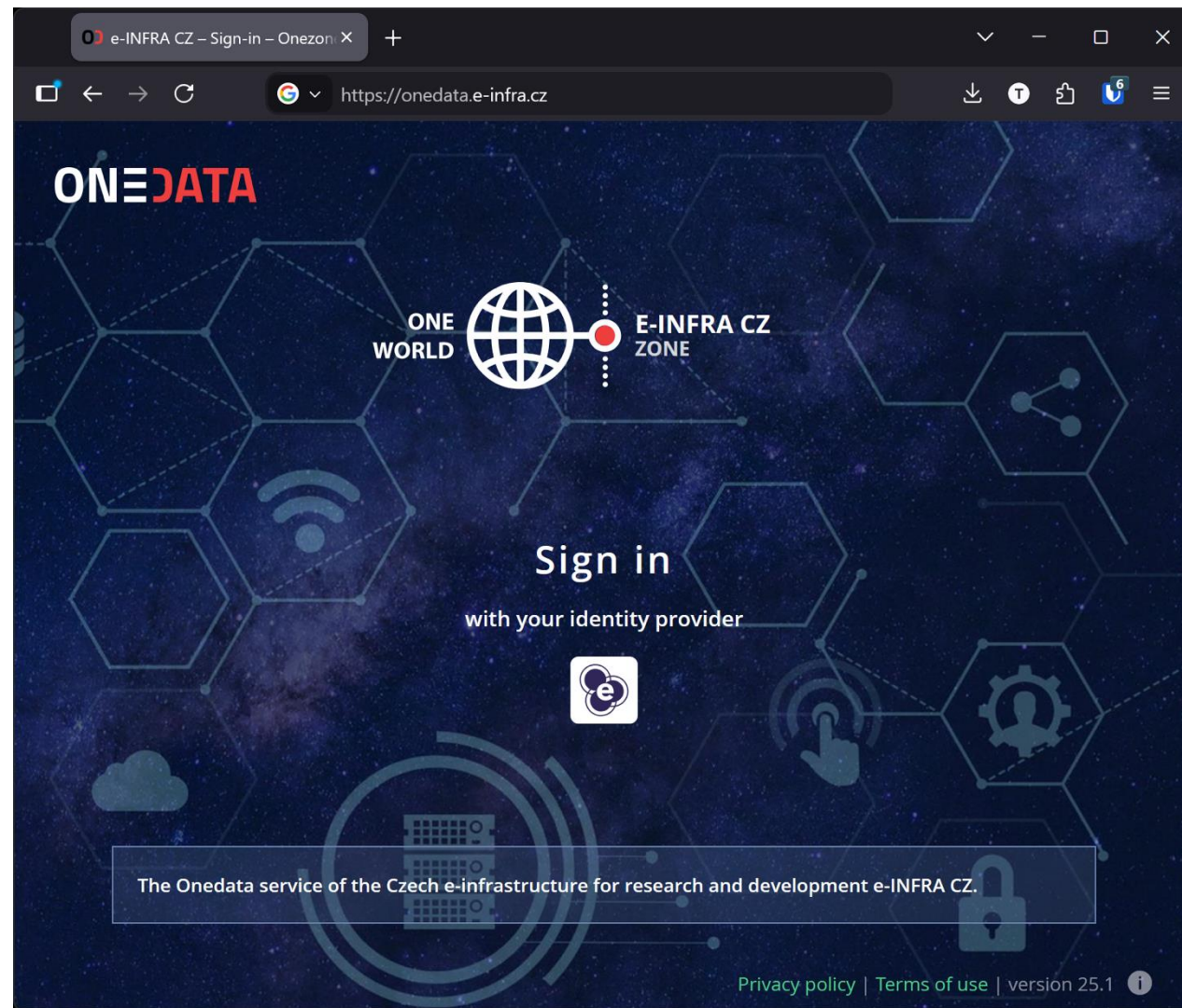
Onedata provozovaná v e-INFRA CZ



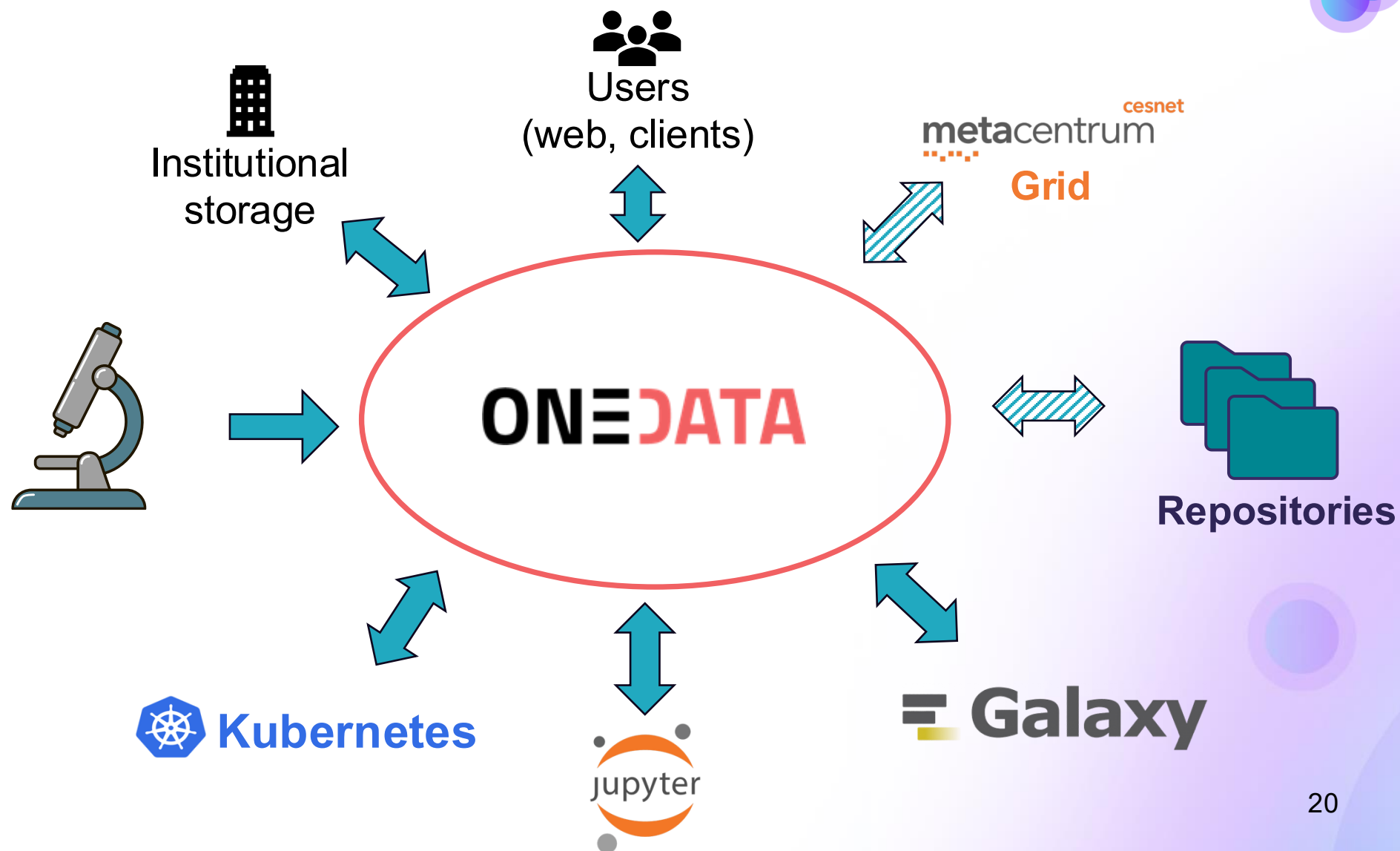
Onedata provozovaná v e-INFRA CZ



- e-INFRA CZ instance systému Onedata
- URL: onedata.e-infra.cz
- Dostupná pro všechny uživatele e-INFRA CZ
 - Sandbox na vyzkoušení
- Zřízení úložného prostoru na požádání onedata@e-infra.cz
- Fyzicky jsou data uložena v S3 DÚ CESNET
- Dokumentace: docs.onedata.e-infra.cz

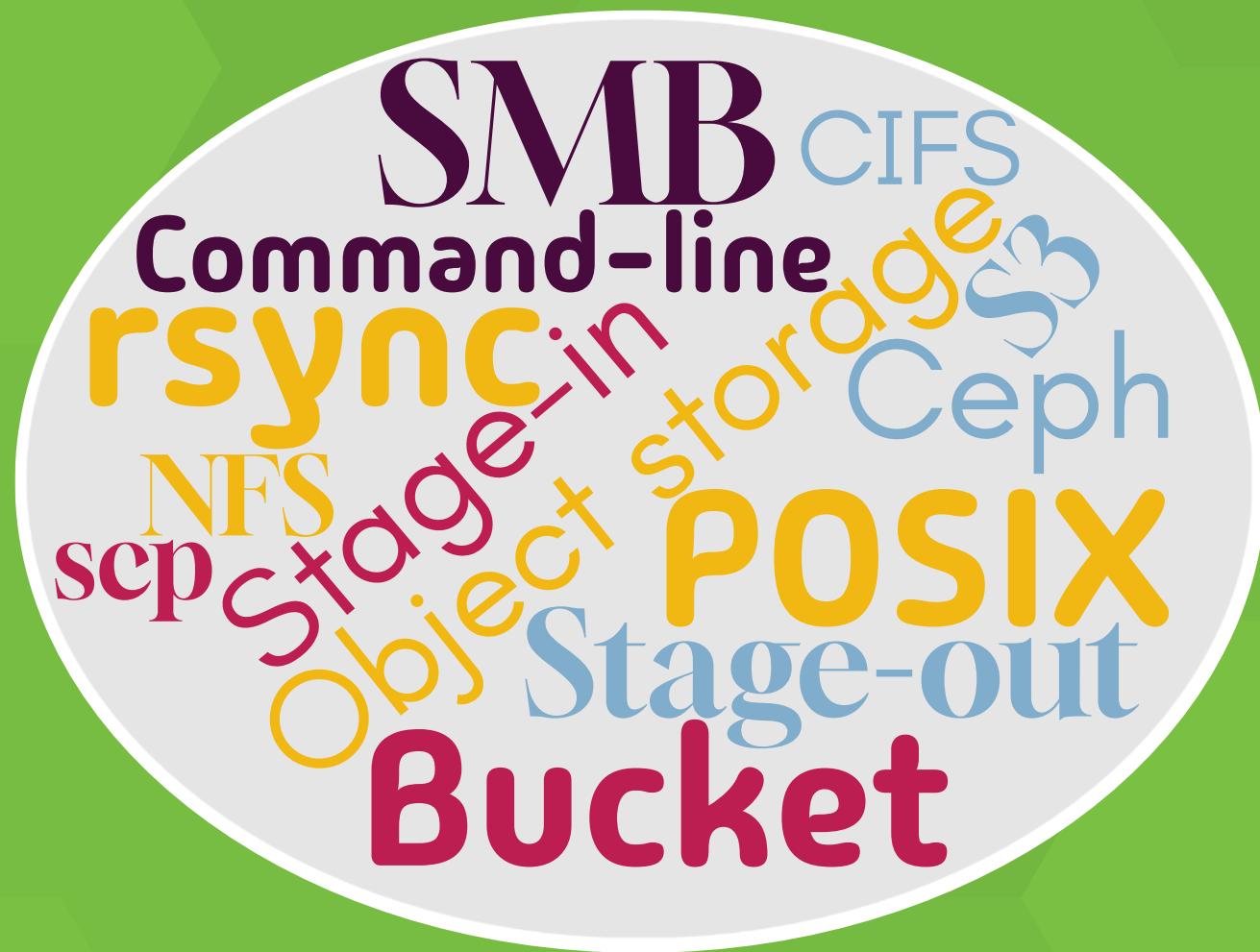


Integrace s výpočetními službami e-INFRA CZ



Thank you for
your attention.

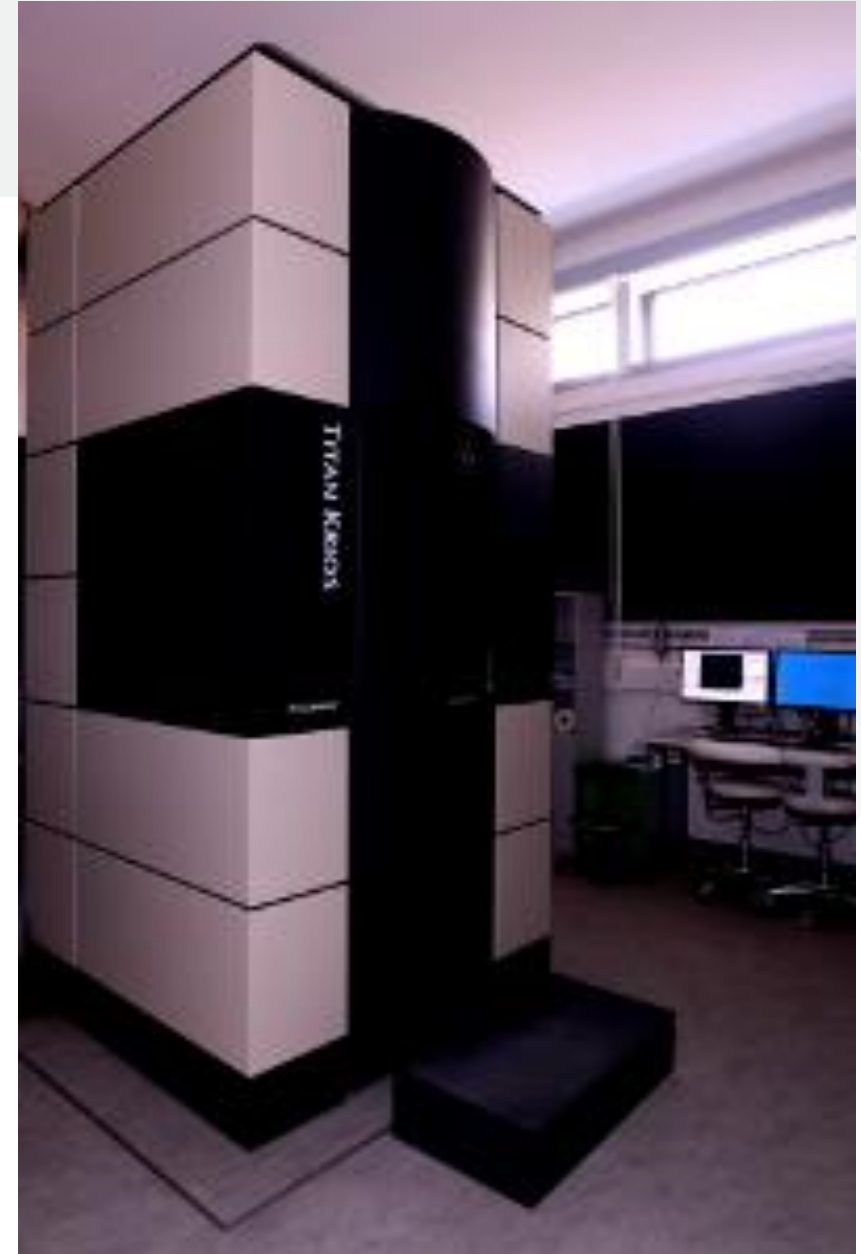
Questions?



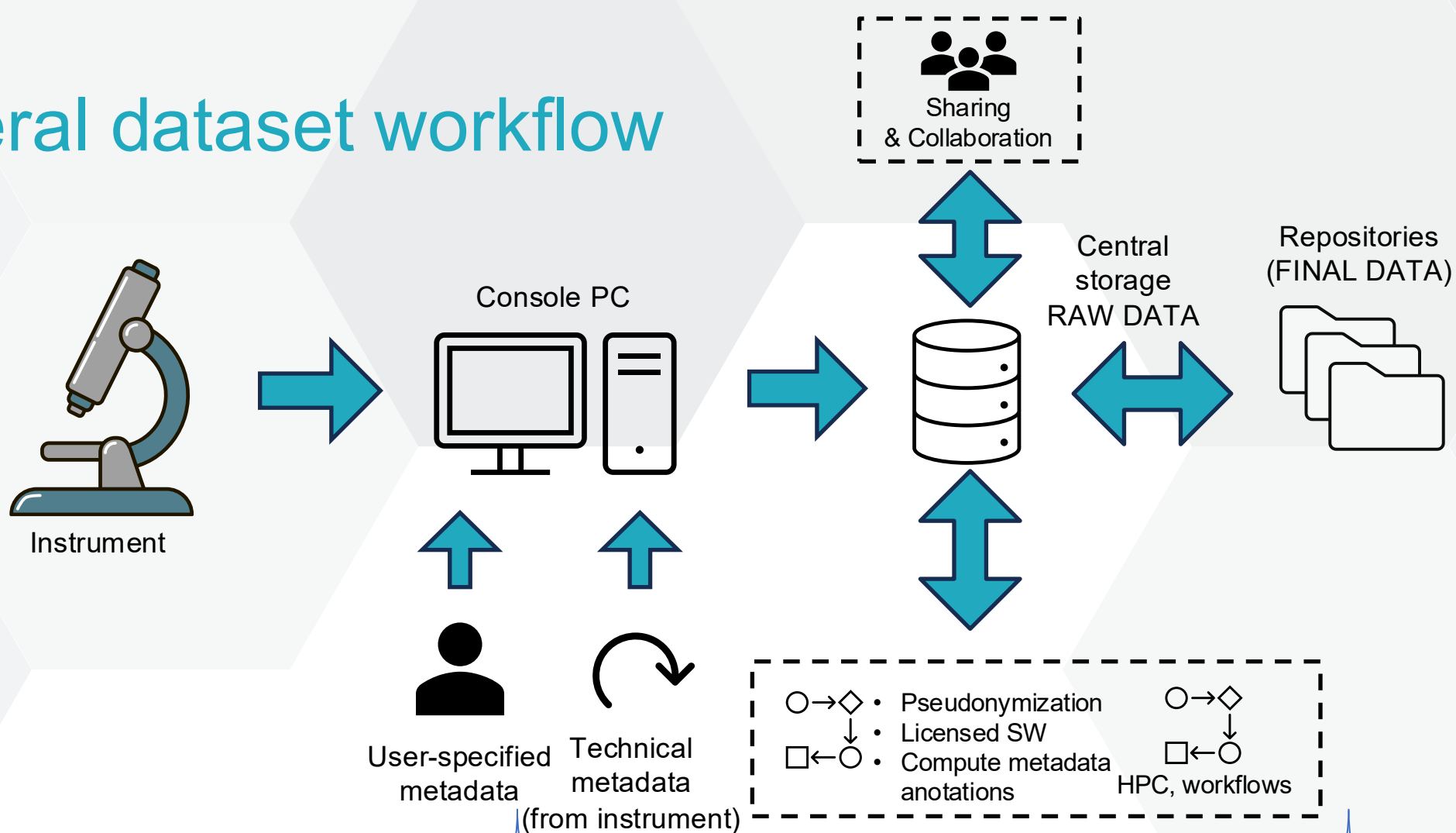
Backup slides

Motivace II

- Řešení pro správu dat ve dvou sdílených laboratořích
 - CEITEC MU (Brno)
 - CNB CSIC (Madrid)
- Cíle:
 - Spolupráce laboratoří
 - Sdílení a vzájemné zpracování rozsáhlých dat
 - Přenos dat ke zpracování prostředky HPC
 - Odkládání dat na archivní úložiště
 - Předávání dat uživatelům laboratoří



General dataset workflow



Laboratory notebooks



Kadi⁴Mat



openBIS

web interface



Data management systems

ONE³DATA iRODS

Repositories

Invenio

CLARIN Dspace

ARL

Transfer use-case

- Secure transfer of sensitive data (from hospital)
- Storage and processing
- Data handover to internal users and partners

