

CyVerse Austria: Cloud-based scientific data management in research & education

Introduction to CyVerse

Research Data Management @ TU Graz

- DMP Support
- RDM Optimization Institute/Group: Workflows/Pipelines, Infrastructure
- RDM Individual support & support in Research Projects (Collaborations)
- RDM guidelines & policies
- Tool-dependent support
(workshops, introduction, training, assistance, development)

RDM-Tools

- TU Graz DMP Tool (<https://www.tugraz.at/sites/rdm/tools/tu-graz-dmp-tool>)
- GitLab (<https://www.tugraz.at/sites/rdm/tools/gitlab>)
- eLabFTW (<https://www.tugraz.at/sites/rdm/tools/elabftw>)
- TU Graz Repository (<https://www.tugraz.at/sites/rdm/tools/tu-graz-repository>)
- CyVerse (<https://www.tugraz.at/sites/rdm/tools/cat-cyverse-austria>)

What is CyVerse?

- A computational infrastructure for data-driven discovery
- The cyberinfrastructure includes:
 - Data Storage
 - Data Science Workbench
 - Education and Training

CyVerse US

- Created in 2008 by the National Science Foundation
- Led by University of Arizona
- Partners:
 - Texas Advanced Computing Center
 - Cold Spring harbor Laboratory

Data platforms for Analytics next to CyVerse?

- Google (Colab, Drive, Sheets a.o.)
- Microsoft (365, Onedrive, Power BI)
- Nextcloud (In-Apps)
- SAS (Viya),
- Cohesity,
- Mendeley Data Monitor, a.o.



Erich Hohenstein, 2022-04-14,
Intro to Google Colab on
<https://levelup.gitconnected.com/google-colab-what-is-it-how-to-use-it-and-why-should-i-care-721bb8a88c90>



<https://nextcloud.com>



<https://www.cohesity.com/>



Andrea Negri, 2017-01-13, SAS Viya is perfect for IT Nerds at
<https://blogs.sas.com/content/brightdata/2017/01/13/sas-viya-perfect-for-it-needs/>

Brand-new: European Open Science Cloud - EU Node

<https://open-science-cloud.ec.europa.eu/>

Login with your university credentials and use some resources for free

Services

File Sync & Share

10 credits consumed in this period

[View Service >](#)

Interactive Notebooks

Access enabled

[View Service >](#)

Large File Transfer

Access enabled

[View Service >](#)

Virtual Machines

Not available

Please read our [User Access Policy](#) for more information

[View Service >](#)

Credits renewed 2025-02-16

Cloud Container Platform

Not available

Please read our [User Access Policy](#) for more information

[View Service >](#)

Bulk Data Transfer

Not available

Please read our [User Access Policy](#) for more information

[View Service >](#)

CyVerse Austria

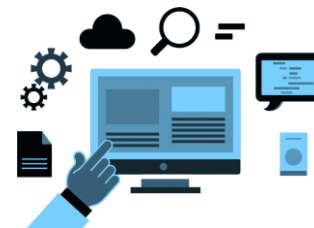
- Uses open source code from CyVerse US
- Started in 2020 within a project funded by the ministry
- Hardware:
 - Storage: Total: 100 TB; used: 3.2 TB
 - CONDOR resources (non-interactive): 10 CPUs, 20 GB RAM
 - VICE resources (interactive): 200 CPUs, 412 RAM

Moving Data

- **Discovery Environment**
 - CyVerse web interface
- **CyberDuck**
 - 3rd party software tool to move large and multiple files; runs on Windows and macOS
- **iCommands**
 - Collection of commandline tools by iRODS (this is the technology supporting the CyVerse Data Storage); runs only on linux
- **GoCommands**
 - Commandline tool; runs on any modern OS

Data infrastructures: benefits compared to local working

- No installation on local machines,
Data access from everywhere
- Usage via web browser or command line
- HPC & shared resources
- Container-based applications (reliability)
- Independent from own system



<https://www.docker.com>

Data infrastructures: disadvantages compared to local working

- Internet connection is needed
- Long procedure to start (VPN, login, start application...)
- Installation of new apps need expertise (e.g. Docker)
- Bug fixing runs over a ticket system



<https://www.docker.com>

CyVerse Help/Courses/Tutorials

- Online Course: <https://imoox.at/course/catmooc>
- Learning Center: <https://learning.cyverse.org/>
- CyVerse-Mooc: https://cyverse-learning-materials.github.io/cyverse_mooc/
- CyVerse Videos: <https://www.youtube.com/@CyverseOrgProject>

CyVerse Austria Registration

Go to registration page: user.cyverse.tugraz.at



Welcome to CyVerse!

New to CyVerse?

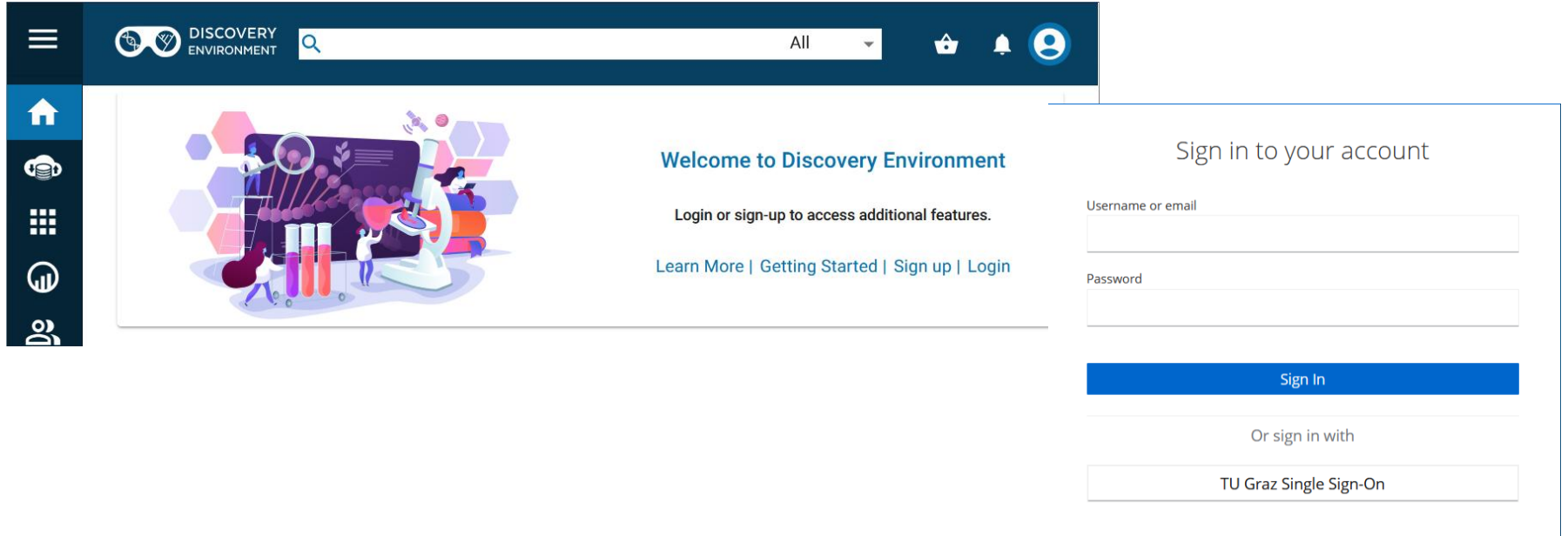
[SIGN UP](#)

[LOGIN](#)

[FORGOT PASSWORD](#)

CyVerse Austria Login

Go to registration page: de.cyverse.tugraz.at



The screenshot shows the Discovery Environment login page. The header includes a navigation menu, the Discovery Environment logo, a search bar, and icons for shopping, notifications, and user profile. The main content area features a large illustration of scientific research and a welcome message. The login form on the right includes fields for username/email and password, a sign-in button, and a link to sign in with TU Graz Single Sign-On.

DISCOVERY ENVIRONMENT

Search: All

Welcome to Discovery Environment

Login or sign-up to access additional features.

[Learn More](#) | [Getting Started](#) | [Sign up](#) | [Login](#)

Sign in to your account

Username or email

Password

Sign In

Or sign in with

TU Graz Single Sign-On

A grayscale, high-contrast image of a large, ornate building with multiple domes and arches, likely a historical university building, serving as the background for the slide.

CyVerse Austria: Demo

Speech-to-text Analyse mit Whisper

Was ist Whisper?

- Automatisches Spracherkennungssystem (ASR)
- Trainiert auf 680.000 Stunden mehrsprachiger und multitasking-überwachten Daten
- Ermöglicht die Transkription in mehrere Sprachen sowie die Übersetzung aus diesen Sprachen ins Englische
- Open Source Software
(<https://github.com/openai/whisper?tab=readme-ov-file>)

Wofür verwende ich Whisper?

- [Shared RDM Services & Infrastructure](#)
- Knowledge Transfer – [Webinarreihe „RDM in Österreich“](#)
 - Über 40 Episoden momentan
 - Abrufbar auf [YouTube](#) & [TU Graz Repository](#)
- Whisper wird zur Erstellung von Untertiteln für die Aufzeichnungen verwendet



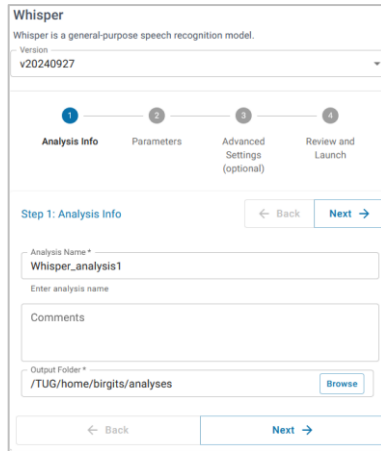
Warum verwendet ich Whisper über CyVerse?

- Läuft online, es werden keine lokalen Ressourcen benötigt
- Läuft immer, es kann beispielsweise auch über Nacht oder am Wochenende laufen
 - Kürzere Wartezeiten für den Abschluss der Analyse
- Es können mehrere Vorgänge eingereicht werden, die nacheinander bearbeitet werden
- Ist schneller im Vergleich zur lokalen Nutzung

Wie verwendet man Whisper auf CyVerse?

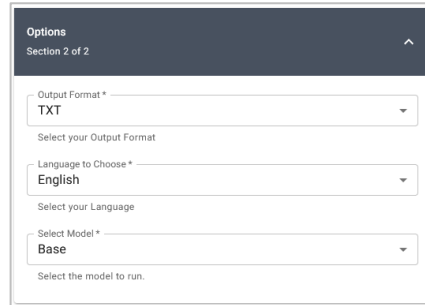
1

Analyseinformationen
definieren



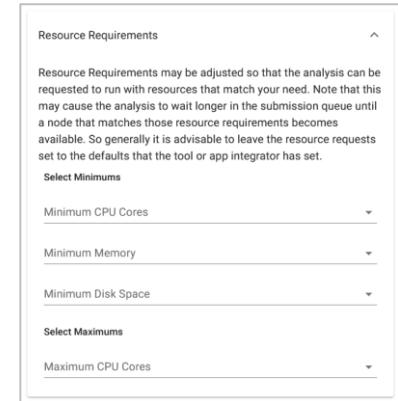
2

Sprachauswahl & Modellauswahl



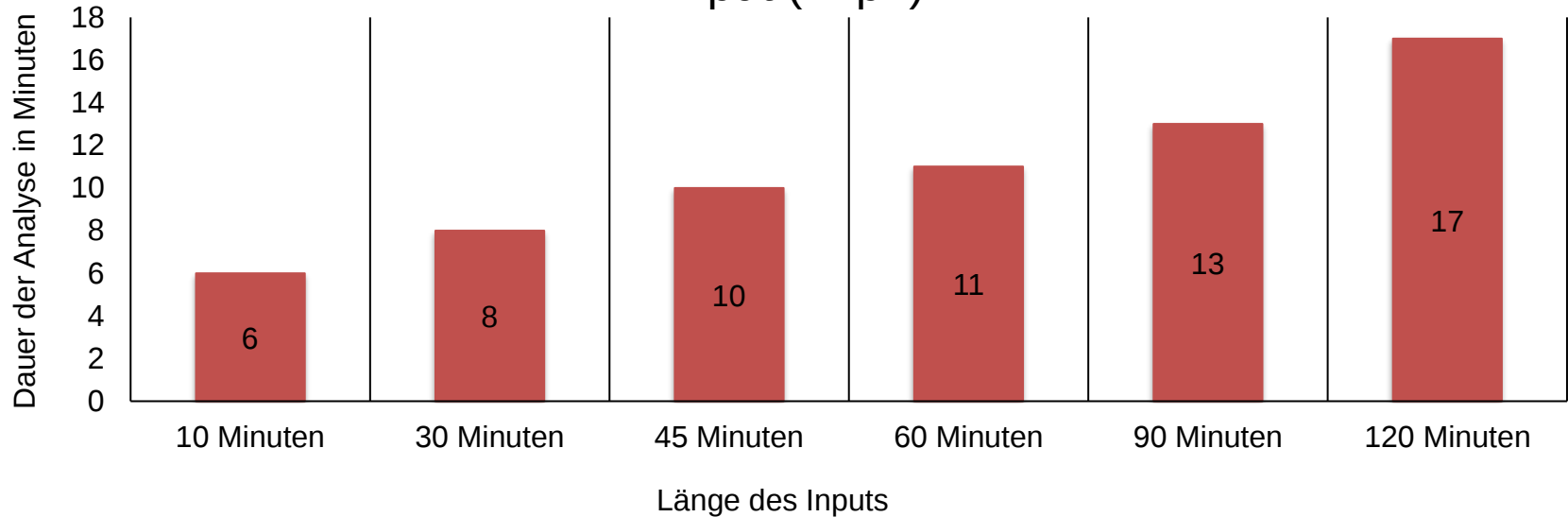
3

Definition der
Ressourcen



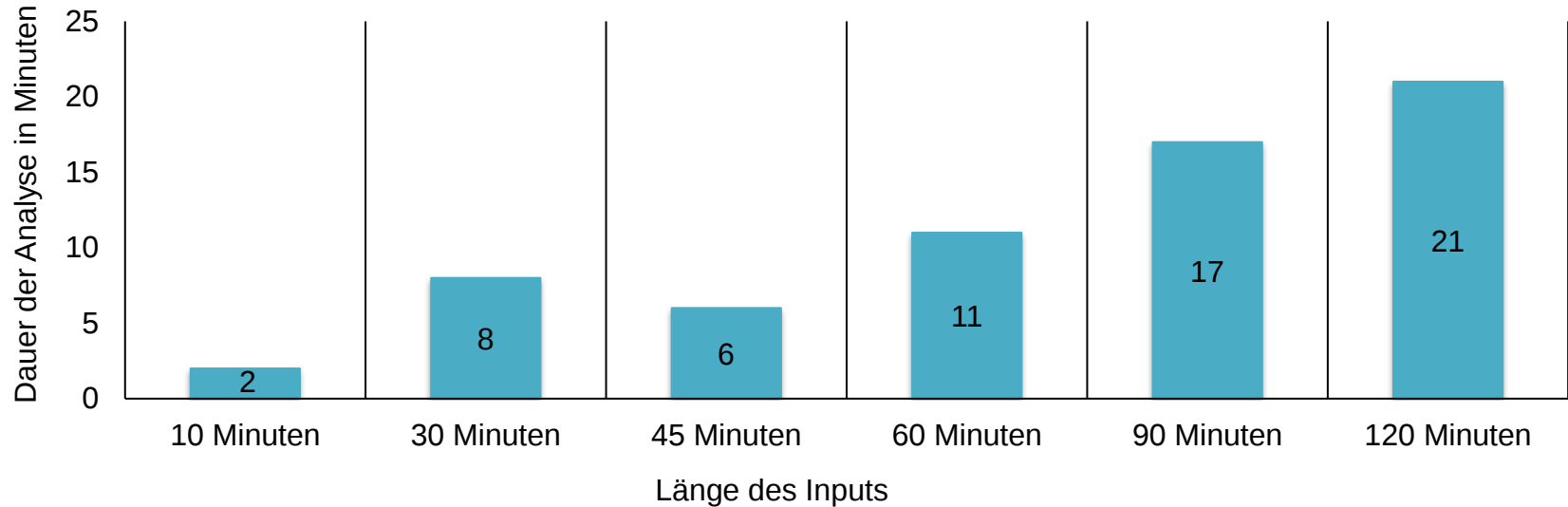
Wie sieht die Performance von CyVerse aus?

Analysedauer auf CyVerse (in Minuten) mit englischem Input (.mp4)



Wie sieht die Performance von CyVerse aus?

Analysedauer auf CyVerse (in Minuten) mit deutschem Input (.mp4)



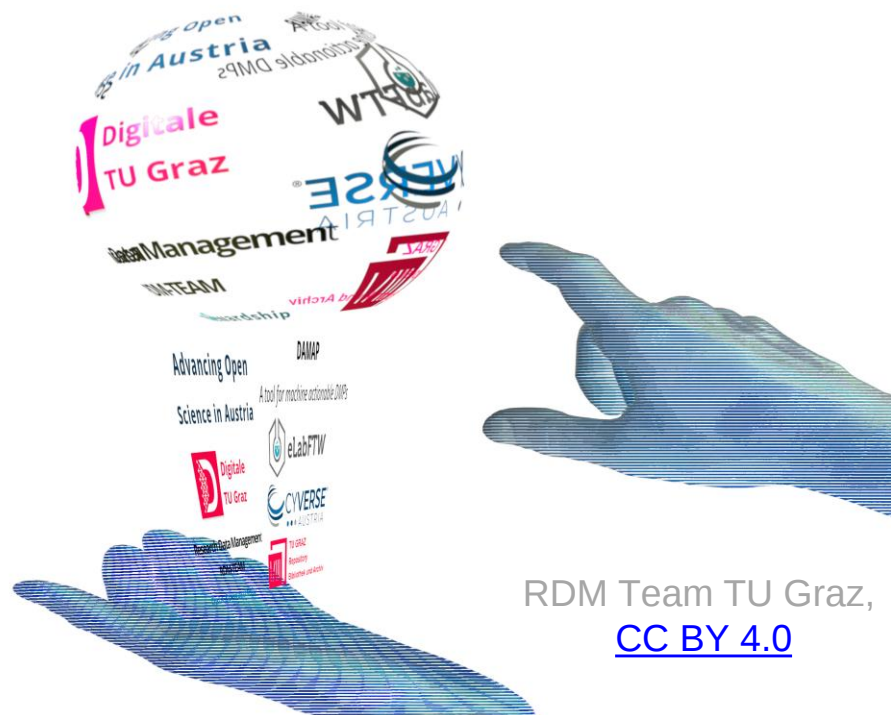
rdmteam@tugraz.at



<https://rdm.tugraz.at>



Supporting reproducible research



RDM Team TU Graz,
[CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)