# BioTechMed Kickoff

## HRSM – Integriertes Datenmanagement

30.09.2020

Prof. Dr. Stefanie Lindstaedt

# Das erfolgreiche, intra-Universitäre Team



+ David Bodruzic
u. Max Matzinger




+ Richard
Hohensinner



# Daten im Forschungsprozess und die Rolle von CyVerse

**Intern im Institut**

Lab Information Systems

CLOSED DATA

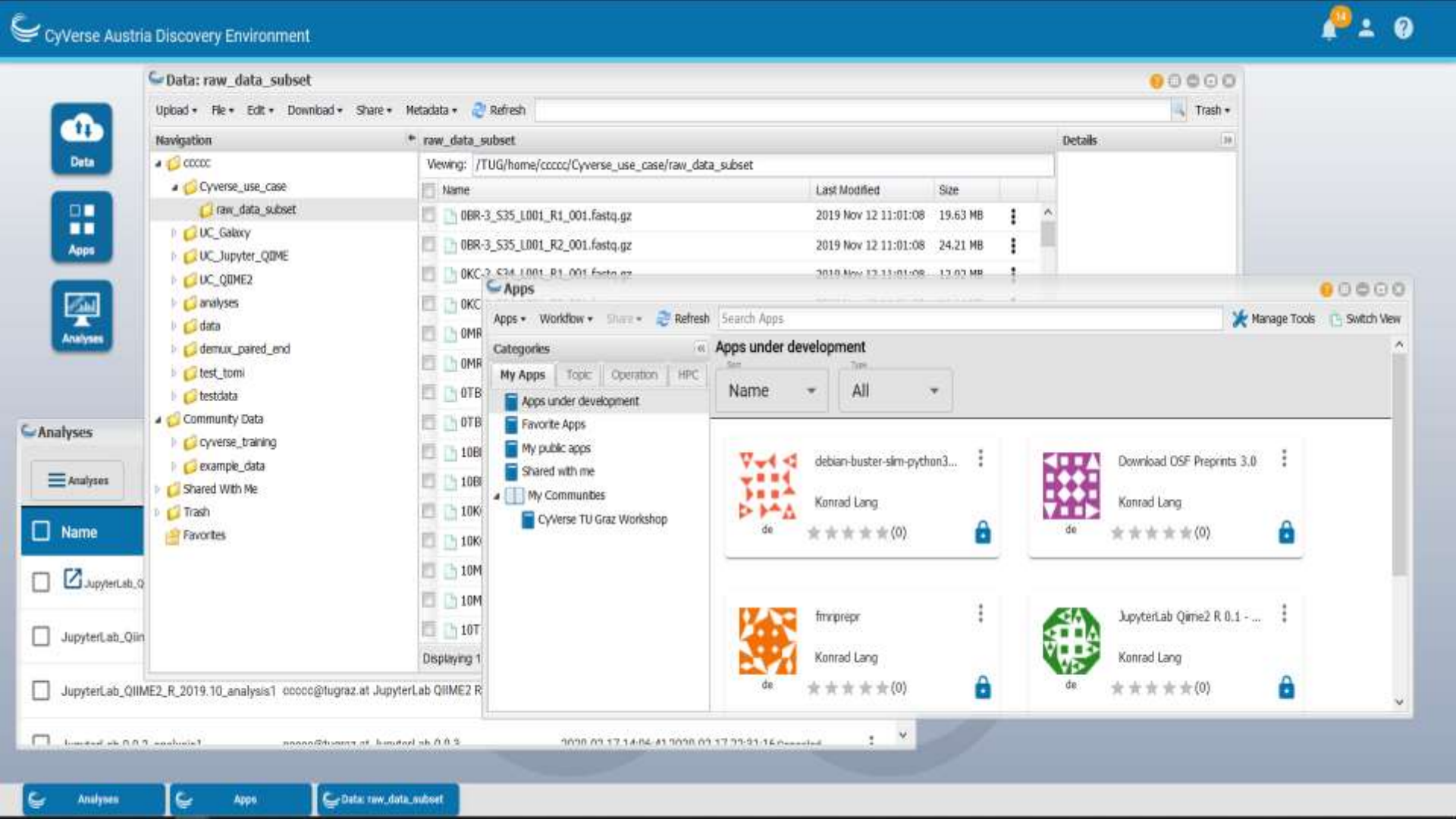
CYVERSE AUSTRIA

**Extern außerhalb der Uni**

INVENIO

OPEN DATA

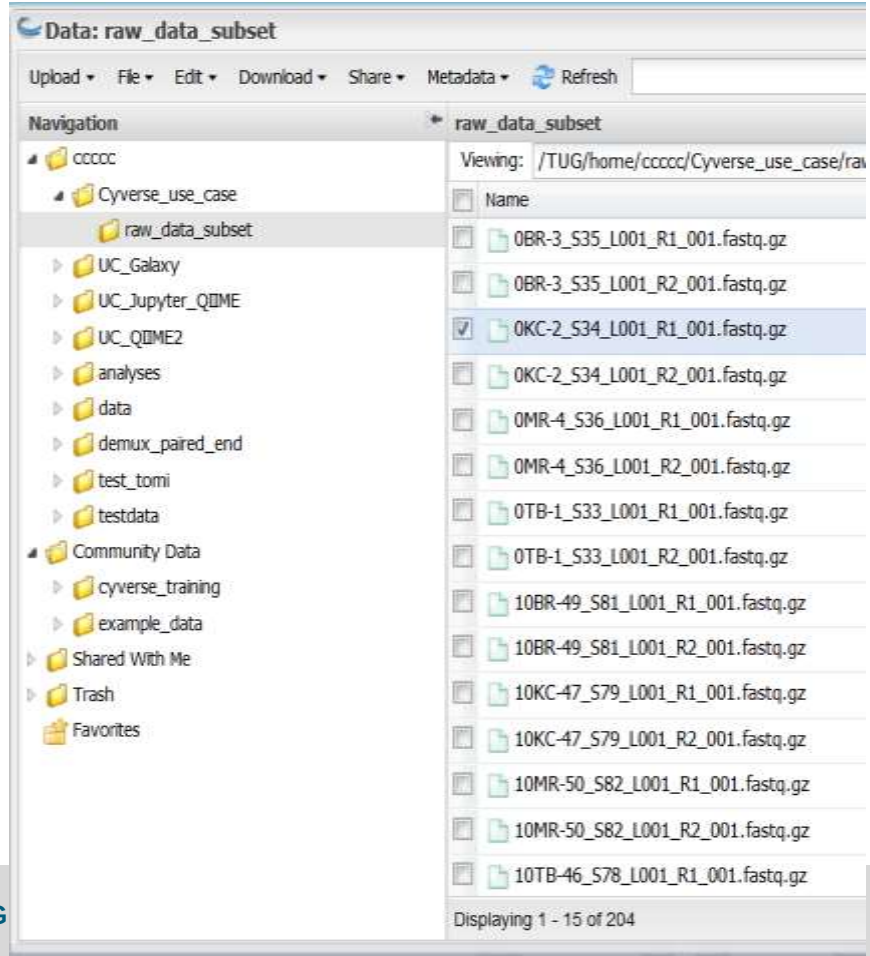
Erhebung → Daten Management → Analyse → Publikation → Archivierung



CyVerse Austria Discovery Environment
Data
Apps
Analyses
Data: raw_data_subset
Upload
File
Edit
Download
Share
Metadata
Refresh
Trash
Navigation
raw_data_subset
Details
cccc
Cyverse_use_case
raw_data_subset
UC_Galaxy
UC_Jupyter_QIIME
UC_QIIME2
analyses
data
demux_paired_end
test_tomi
testdata
Community Data
cyverse_training
example_data
Shared With Me
Trash
Favorites
Viewing: /TUG/home/cccc/Cyverse_use_case/raw_data_subset
Name
Last Modified
Size
0BR-3_S35_L001_R1_001.fastq.gz
2019 Nov 12 11:01:08
19.63 MB
0BR-3_S35_L001_R2_001.fastq.gz
2019 Nov 12 11:01:08
24.21 MB
Displaying 1
Apps
Apps
Workflow
Share
Refresh
Search Apps
Manage Tools
Switch View
Categories
My Apps
Topic
Operation
HPC
Apps under development
Favorite Apps
My public apps
Shared with me
My Communities
CyVerse TU Graz Workshop
Apps under development
Name
All
debian-buster-slim-python3...
Konrad Lang
de
(0)
Download OSF Preprints 3.0
Konrad Lang
de
(0)
fmriprepr
Konrad Lang
de
(0)
JupyterLab Qiime2 R 0.1 - ...
Konrad Lang
de
(0)
Analyses
Analyses
Name
JupyterLab_Q
JupyterLab_Qiin
JupyterLab_QIIME2_R_2019.10_analysis1
cccc@tugraz.at JupyterLab QIIME2 R
Analyses
Apps
Data: raw_data_subset

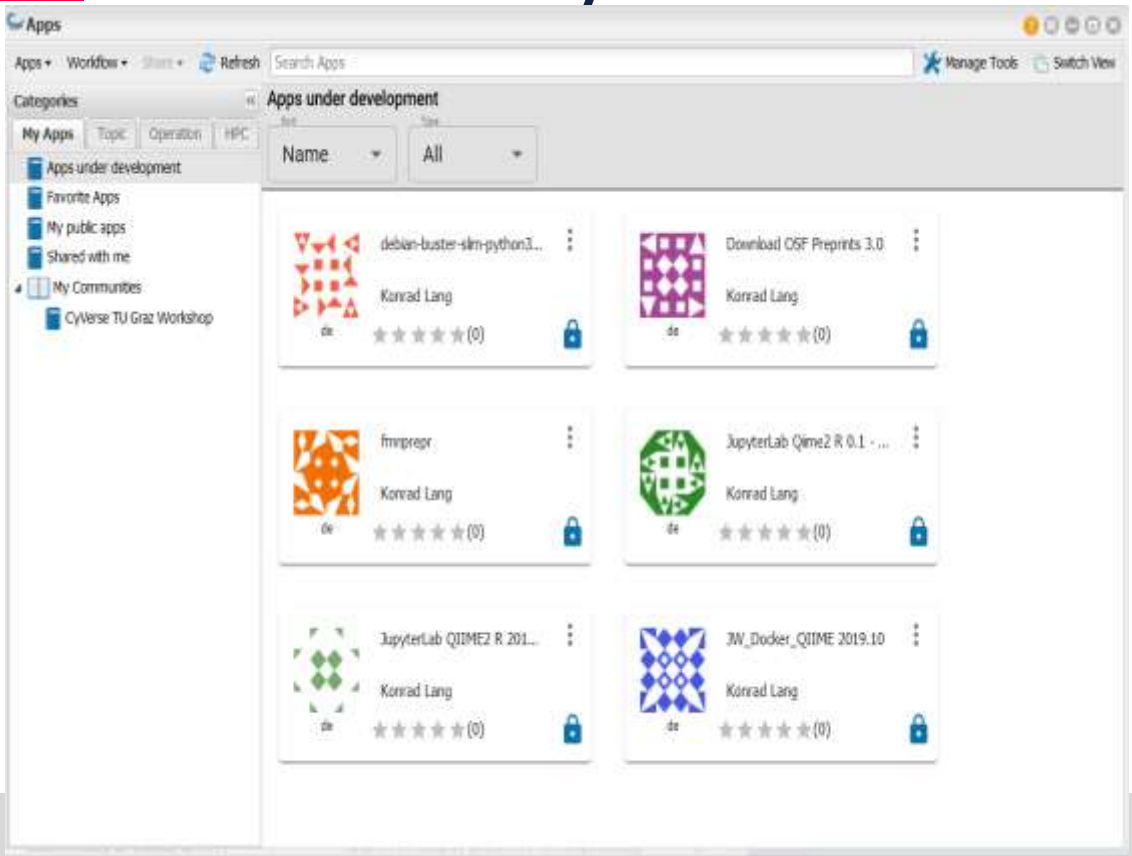
# Daten Management

**Data Store**

- Web based user interface
  Zugriff von überall
- Daten heraufladen über
  - iCommand (iRODS)
  - Cyberduck
  - WebDav
  - Drag & drop
- Logging der Zugriffshistorie
- Daten sichtbar machen für Kooperationspartner
- Nicht alle Daten müssen auf CyVerse liegen, können auch über Links eingebunden werden
- Metadatan zusammen speichern mit Daten
- Automatische Metadaten Extraktion (beschränkt)
- Flexible Nutzung von Speicher Ressourcen



# Daten Analyse

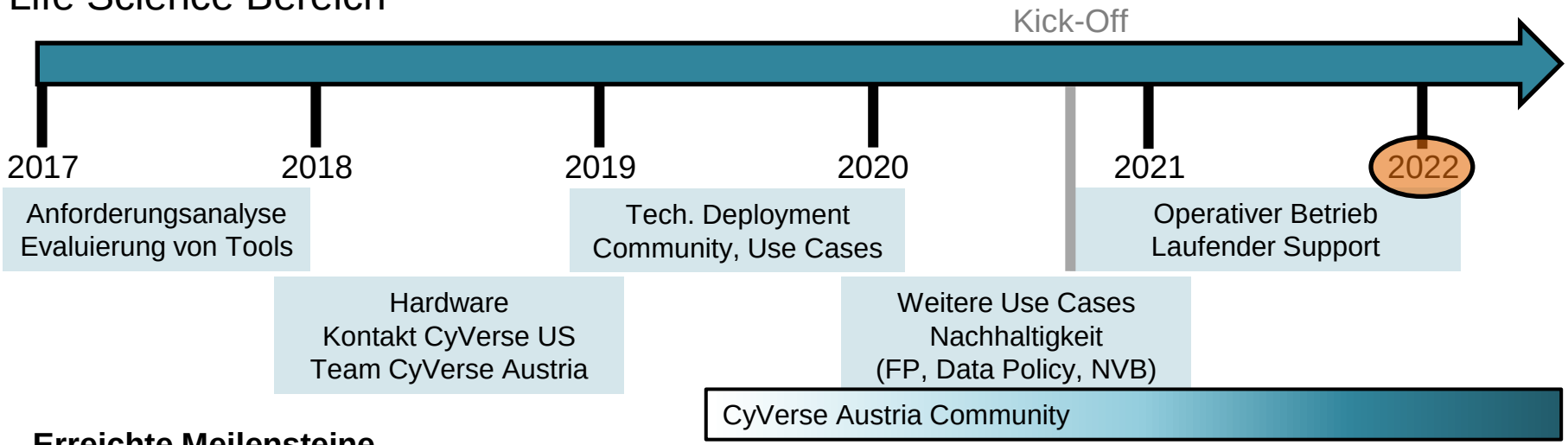



**Discovery Environment - Analytics**

- Web-basiertes user Interface
  Zugriff von überall
- > 3000 Analyse Apps stehen zur Verfügung
- Einbinden eigener Analyse Algorithmen möglich
- Speicherung der Analyse zusammen mit den richtigen Versionen der Apps (Reproducability)
- Einfache Erstellung von Analyse-Pipelines
- Flexible Nutzung von Computational Ressources
- VSC kann angesprochen werden
- Am VSC wird kein Speicher benötigt

# Vorgehensweise

**Ziel**: Stärkung von Kollaborationen und Datenanalysen für Grazer ForscherInnen im Life Science Bereich




**Erreichte Meilensteine**

(i) Technisches Deployment, (ii) aktive Community, (iii) Trainings und (iv) international Vernetzung

Jährliches Reporting zum BMBWF, on-track

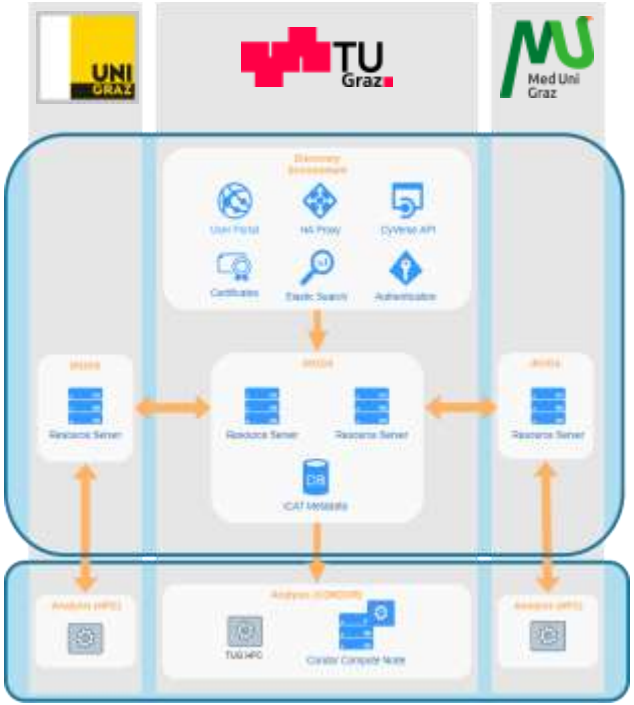
# CyVerse Austria Infrastruktur in Graz

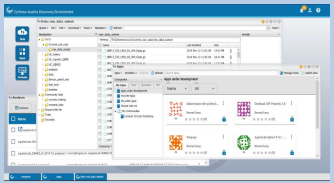

**iRODS**
Virtualize data storage resources

**Data Store**
Easy-to-use data management

**Discovery Environment**
User interface to access/manage

Data Management



**VICE**
Visual and Interactive Computing Environment

**HTCondor**
job scheduling, resource mgmt

**App container**
Integrated Dockerized tools
**Data staging-in and out**

Data Analytics

# Vision

EU

Germany

US

Australia

CyVerse-AU

Austrian DataLab and Services

e.g. Computational Engineering

International

Disziplin

National

Industrie

FAIR Data Austria

RegioData

CyVerse/iRODS als Plattform für (Pharma-)Industrie

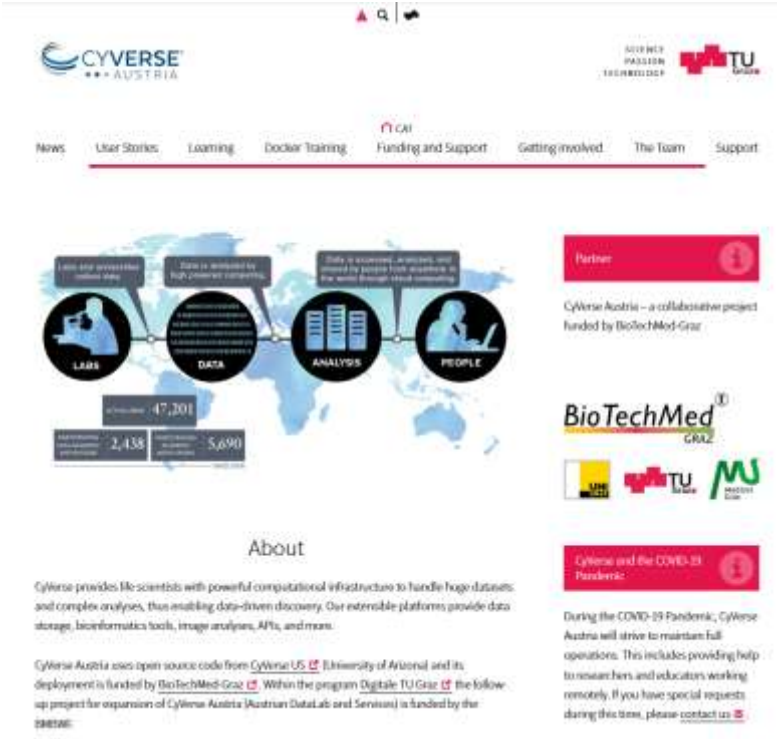
# Regelmäßige Trainingsaktivitäten

| Feedback Questions | Out of 5 |
|---|---|
| How did this workshop meet your expectations? | 4,4 |
| How helpful were the workshop instructors? | 4,8 |
| How would you rate the CyVerse infrastructure? | 4,2 |

- On-site Training mit CyVerse US
- Carpentries Training
  - **Grazer Universitäten sind jetzt Teil der globalen Carpentries Community**
- Dockertraining
  - **Online Training, um ForscherInnen auch im Homeoffice zu unterstützen**

# CyVerse Austria – Homepage

https://cyverse.tugraz.at




# Unsere aktive und motivierte User Community

*“I think everything was perfect. Nice pace, depth, time for hands on work on personal ideas/projects. Thanks!!”*

*“Besides getting impressions and links of many useful tools and trainings I think learning the basics about Docker and git(hub) was the most important for my future tasks. “*

*“more resources and computing capacity!!!”*

*“it was easy to follow. perfect sped, perfect examples chosen. and the web-documentation for reading through it tomorrow again is also veery nice, please keep it online for some time or please send them out by email as one bundled PDF. ”*

*“Thank you, sarah, philipp & christoph! Very informative talks & live demonstrations! ”*

*“thanks for the cool training. you made a very good job. looking forward for part ii. ”*

**Kommunikation**

- Regelmäßige Treffen (alle zwei Monate)
- Slack Channel mit CyVerse US
- Zammad CyVerse Austria Support

# Nächste Schritte für CyVerse Austria

- **Nachhaltigkeit gewährleisten durch Finanzierungsplan**
  - Entwicklungsplan, Leistungsvereinbarungen der Universitäten
- **Regulierung der Nutzung durch Data Policy**
- **Atmosphere Modul**
  - OpenStack für Cloud Computing
  - Federated Cloud – in Entwicklung
- **Ausweitung**
  - Mehr Disziplinen und Universitäten (ADLS Projekt)

# Vielen Dank für die Aufmerksamkeit!

Kontakt:


- Stefanie Lindstaedt, slind@know-center.at

CYVERSE AUSTRIA

Institute of Interactive Systems and Data Science
Know-Center GmbH

# Weitere Aktivitäten

- Publikationen über CyVerse Austria
  - MDPI MCA Special Issue – CyVerse Austria Infrastruktur
  - Use Case Reports

- CyVerse Austria in Forschungsproposals
  - FWF Einzelprojekt Open Science Badges (TUG, KFU)
  - EC Proposal (MUG, Gernot Plank)
  - FWF Zukunftskolleg (MUG, KFUG, TUG)
  - …