

**MAKE YOUR LIFE  
(AND ANALYSIS)  
EASIER WITH  
CONTAINERS**

# AUDIENCE

- Are you a biologist?
- Have you heard of Docker?
- Not sure where to start?

**YOU'VE COME TO THE RIGHT PLACE!**

# ME

- Software Engineer
- Build software infrastructure for researchers
- Help researchers to use computational tools
- Was a 'container skeptic'

# CYVERSE

Helps researchers:

1. Learn about, and
2. Productively use

New tech like containers

# ANALYSIS IS GETTING COMPLEX

- Multiple software packages (R, Python, etc.)
- With specific versions
- Have to work together
- On different platforms

# THE PAIN

- Hard to install one-by-one
- Wasted effort and time
- Fragile, hard-to-reproduce analyses

# **HELP! MAKE IT STOP!**

How we we make it easy to install & use things  
consistently?



# CONTAINERS! \*

New packages & apps are increasingly available as containers (BioContainers, etc.)

**USING CONTAINERS**

# **DEMO: COMMAND LINE APP**

**DEMO: WEB APP**

# CYVERSE SUPPORT FOR CONTAINERS

1. Command line (Atmosphere)
2. Interactive apps (VICE)
3. HPC (XSEDE & OSG)

# **CONCEPTS & TERMS**

# IMAGE

A self-contained, read-only 'snapshot' of your applications and packages, with all their dependencies

# DOCKERFILE (OR SINGULARITY RECIPE)

Executable instructions (script) for:

- Creating an image
- Specifying the 'entry point' for the container



# CONTAINER

- A 'running image'

# DOCKER

- A server (sometimes called a daemon): A program that runs in the background, and handles life cycle of images and containers
- A command-line client: You use it to tell the server what to do

Download from: <https://www.docker.com/>

# SINGULARITY

A way to run containers on HPC

Find out more: <https://www.sylabs.io/singularity/>

# WHAT ABOUT MY DATA?

Do not put your data in the image!

- Local data: 'Mount' it into a container when you start it
- Remote data: Pull into the container once it's running (e.g. CyVerse Data Store, S3, etc.)

# COMPUTE RESOURCES

I need more!

Talk to us. There are a few options, and it depends on what you need.

# SHARING CONTAINERS

Image registries

# SUMMARY

- Package your analysis pipeline in a single container
- Everyone in your lab can have a consistent environment

# NEXT TIME

- How to build containers
- Running on different platforms
- Science applications



# LINKS & REFERENCES

- [Docker](#)
- [Singularity](#)
- [Play with Docker Classroom](#)
- [Katacode - Learn Docker](#)
- [CyVerse Container Camp materials](#)
- [Reproducible research with containers](#)
- [Upendra's Cybercarpentry workshop notes](#)
- [Matthew Rich's Singularity workshop](#)
- [BioContainers](#)

# THANKS!

- Nirav Merchant
- Upendra Devisetty
- Tyson Swetnam
- Blake Joyce
- Eric Lyons
- Ariella Gladstein
- Tina Lee
- Shelley Littin



CyVerse is supported by the National Science Foundation under Grants No. DBI-0735191, DBI-1265383 and DBI-1743442

