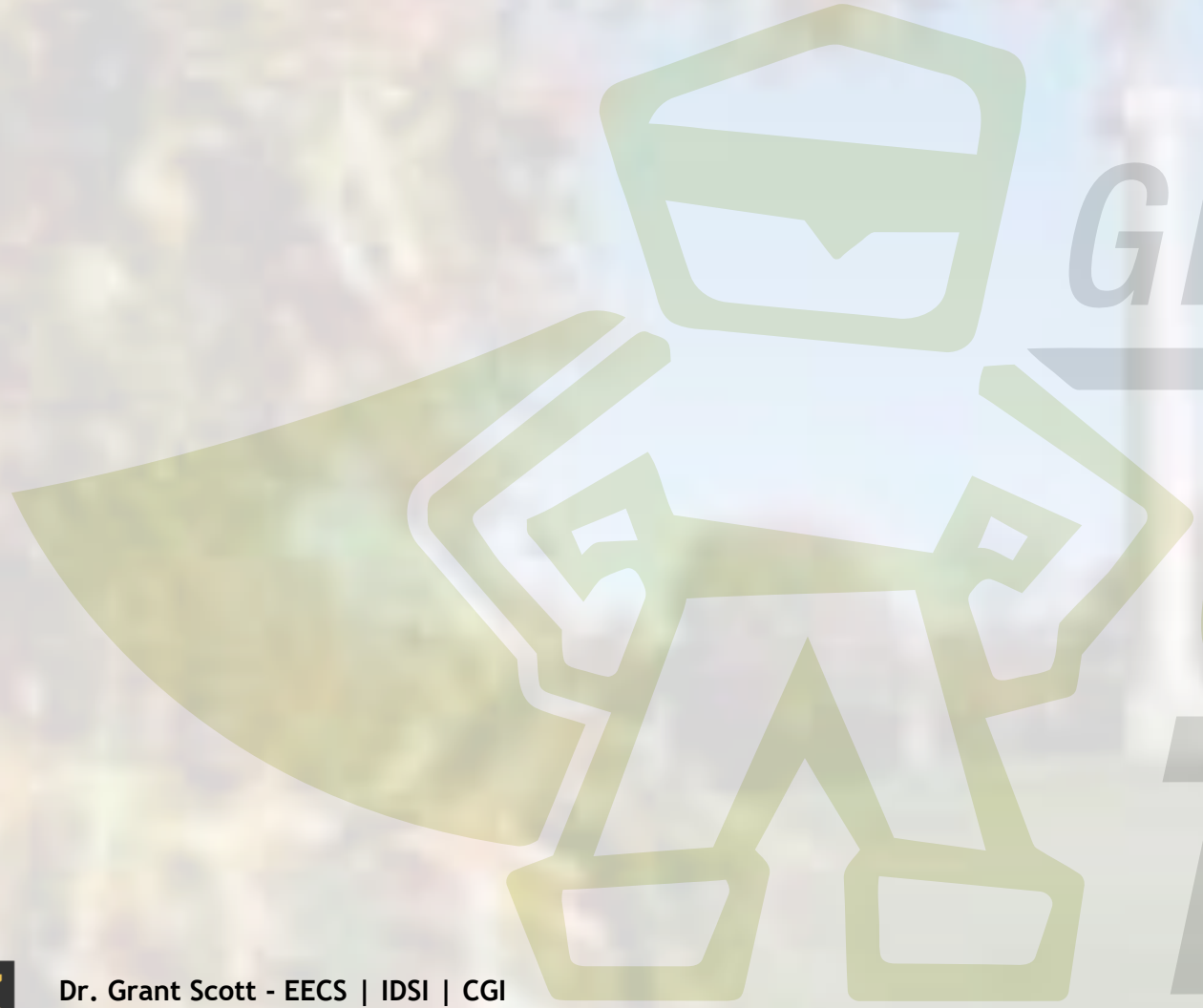
Setting Up Research Jobs in the PRP Nautilus Hyper-Cluster

Grant Scott, Anes Ouadou, and Chenhan Zhao

Great Plains Network Annual Meeting

June 3, 2022

University of Missouri



GREAT PLAINS CYBER TEAM

Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri

Workshop Outcome:

You will be able to train a deep learning neural network using Kubernetes and GPU resources on Nautilus

… and other scientific computing processes

Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri

This workshop creation and delivery was supported by the National Science Foundation (NSF)

NSF Award #1925681 - CC* Team: Great Plains Regional CyberTeam

Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri

Workshop Agenda

1. Introductions
2. Getting logged into Nautilus and Setup
3. Overview of Kubernetes Concepts
4. Accessing JupyterLab in Nautilus, configuring KubeCtl
5. Preliminary Pod Control
6. Setting Up Persistent Storage for Data Input and Output
7. Launching Training using a Customized Container Image

Introductions

Grant Scott

Chenhan Zhao

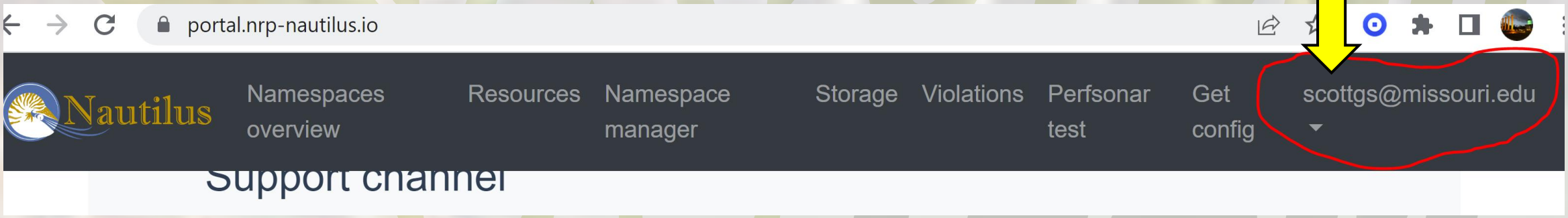
Anes Ouadou

Getting your account linked to the Workshop Namespace

- Access the link below a follow the steps to authenticate into Nautilus

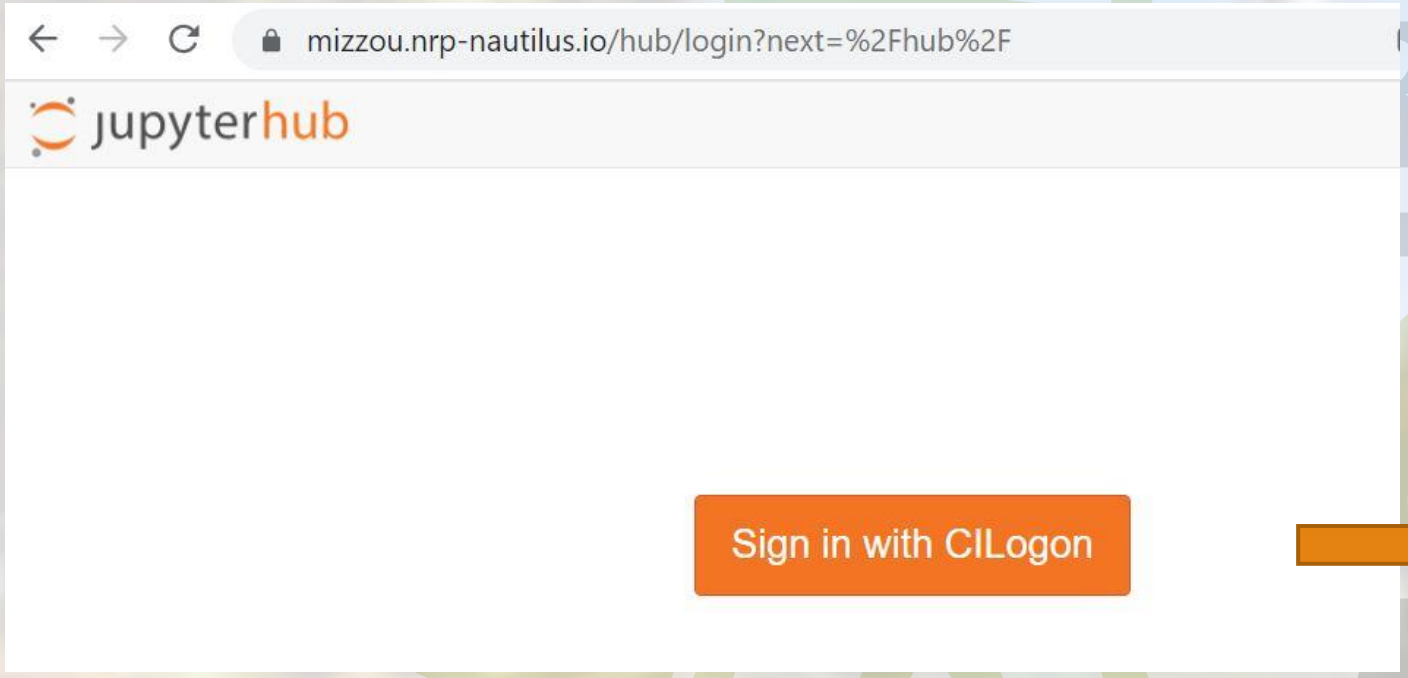
  https://ucsd-prp.gitlab.io/userdocs/start/get-access/

- Once you follow the instructions down to step 6, post into the GPN Slack, **Channel #gpn-nautilus-kubernetes-workshop** your **User ID** so Chenhan can pull you into our namespace as a User.



Getting logged into Nautilus

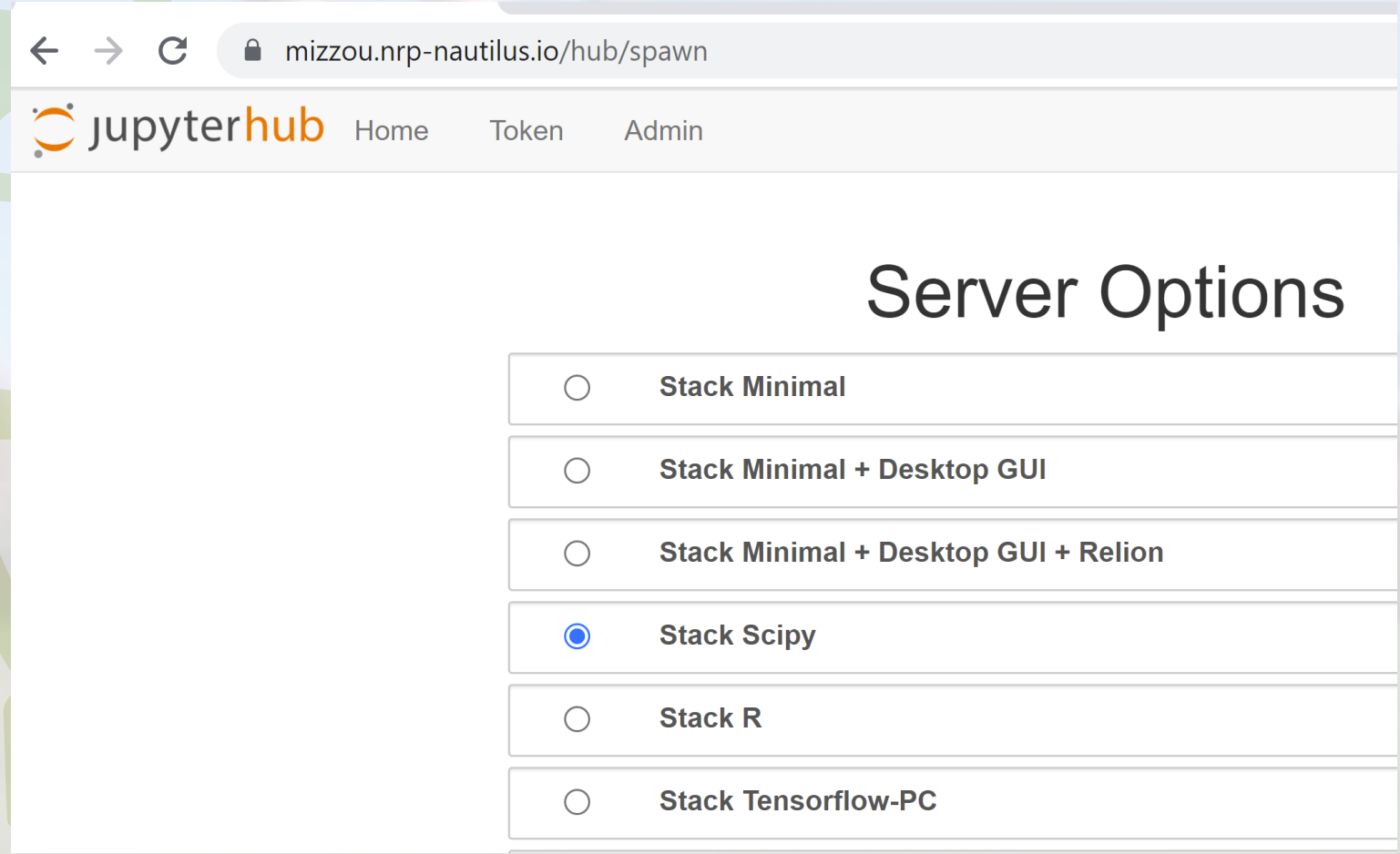
https://mizzou.nrp-nautilus.io/



Nautilus Jupyter Lab

https://mizzou.nrp-nautilus.io/

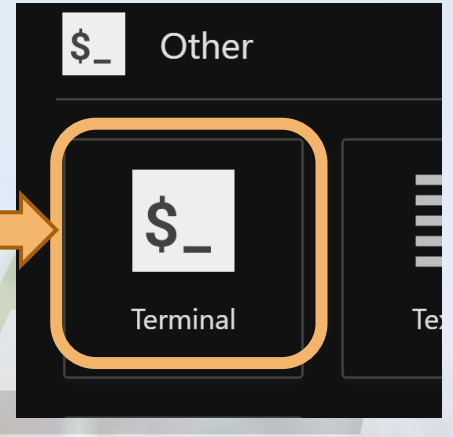
- Jupyter Hub / Lab running on a Kubernetes cluster in Nautilus
- Use the CILogon to access any institution or use your Github.com credentials
- Use the Scipy or Datascience Stack, and click the lower **Start** button

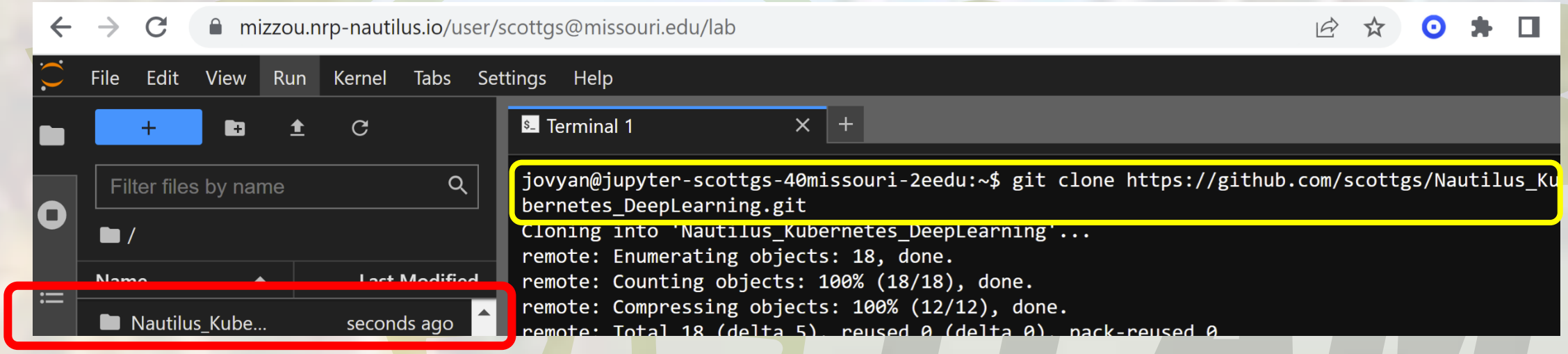


Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri

Pulling Down Workshop Content

- We are cloning workshop content from Github.com
- Open a Terminal
- Clone Github.com/scottgs/Nautilus_Kubernetes_DeepLearning

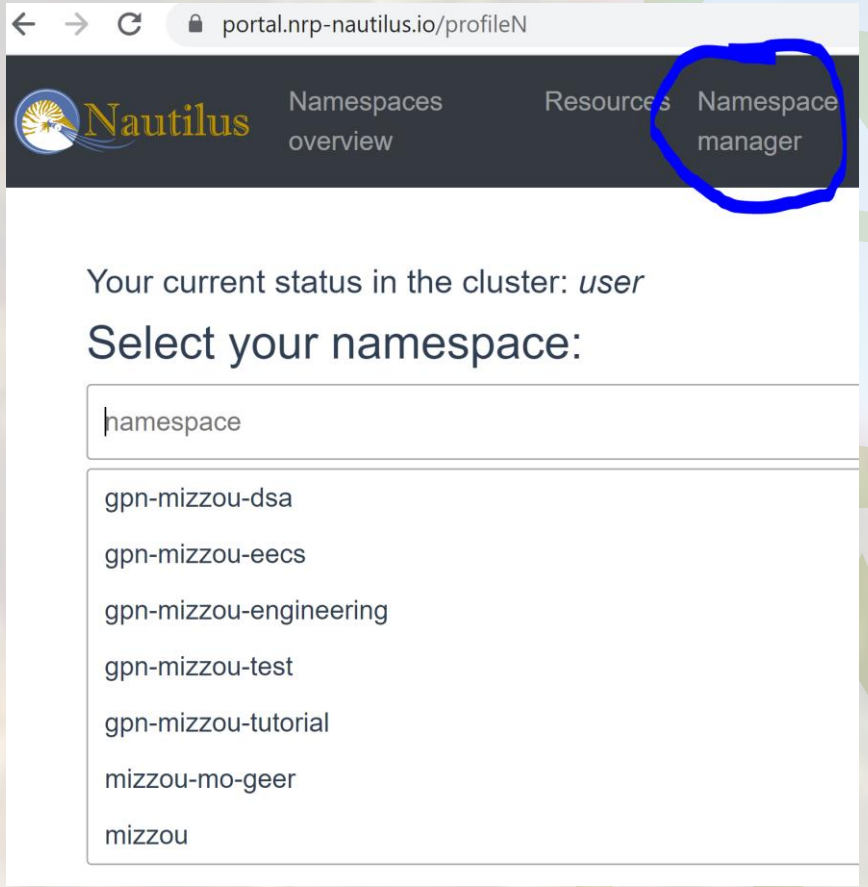





```
git clone https://github.com/scottgs/Nautilus_Kubernetes_DeepLearning.git
```

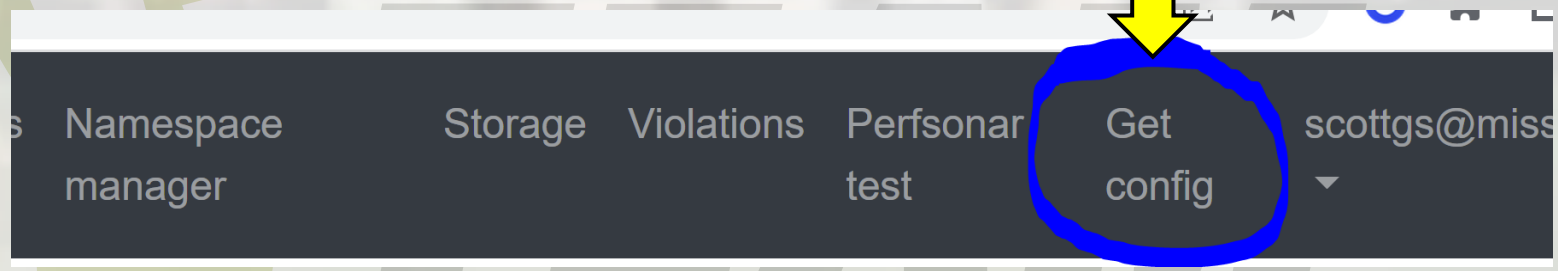
Confirming Nautilus Namespace

https://portal.nrp-nautilus.io/profileN



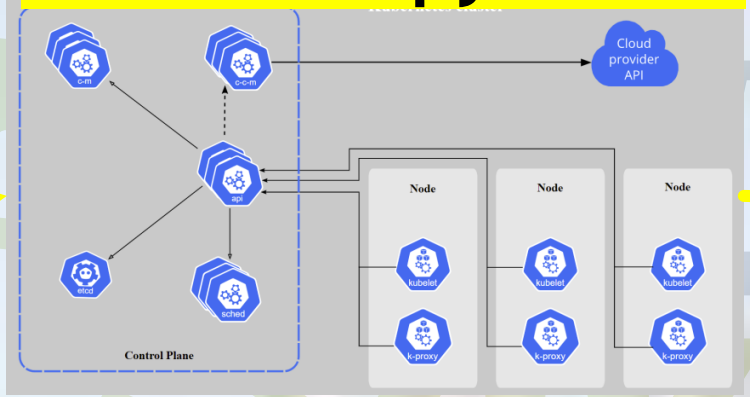
- Going back to our Nautilus Namespace setup, after a checkmark from Chenhan
- Select the Namespace Manager
- Once it loads, select the

  `gpn-mizzou-tutorial`
- Then download your Nautilus Configuration file



Kubernetes – Inception Style

Nautilus Jupyter Lab



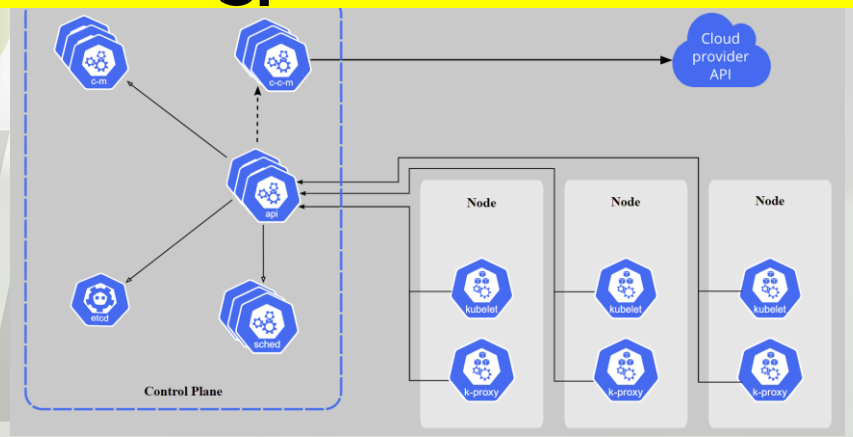
Use Stack Scipy

Browser UI to Jupyter Lab

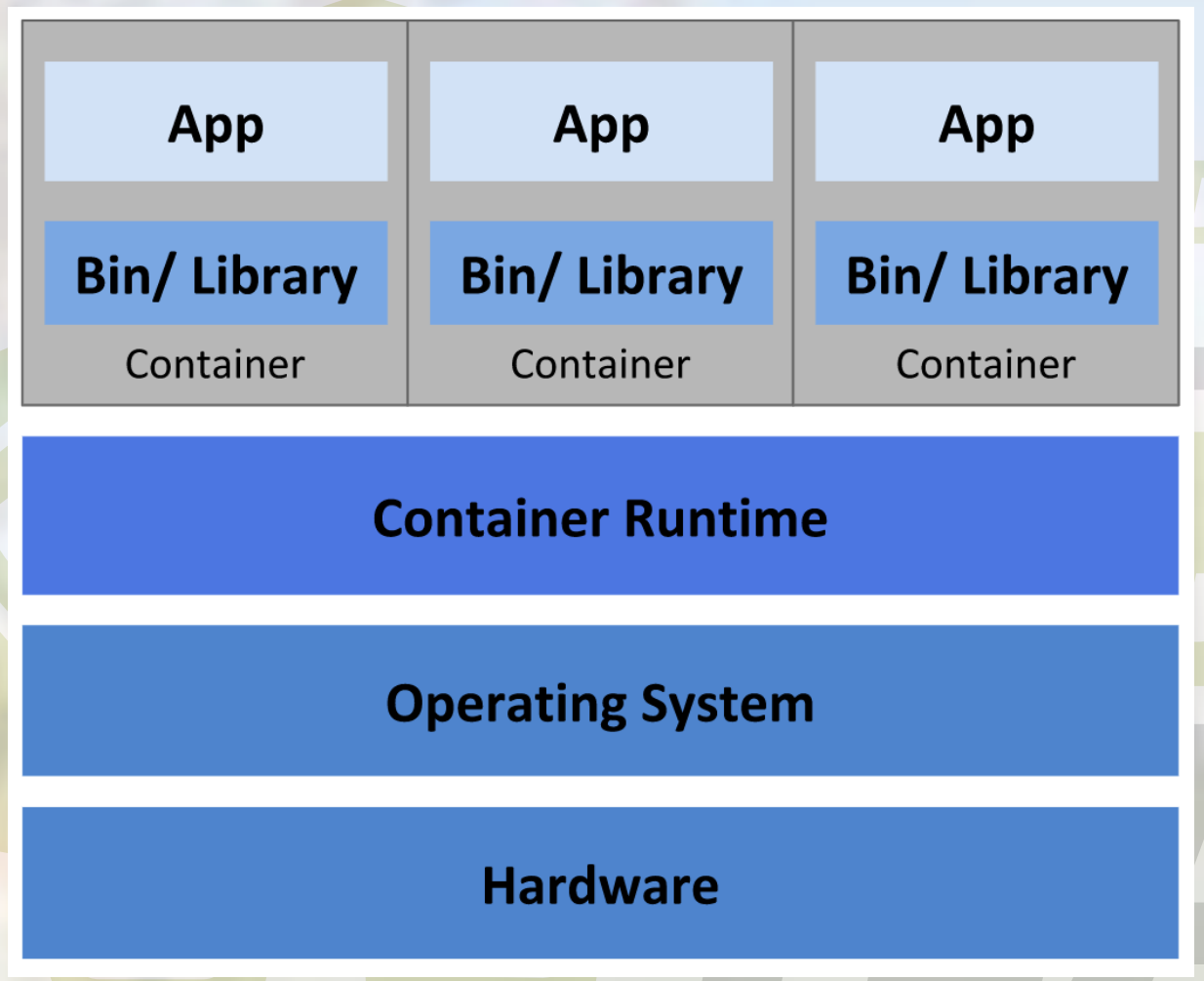
kubectl CLI

Your local computer

Nautilus gpn-mizzou-tutorial



Kubernetes – A foundation of Containers



Kubernetes - Foundational Elements

- **Pods** - smallest deployable units of computing that you can create and manage in Kubernetes
- **Jobs** - creates one or more Pods and will continue to retry execution of the Pods until a specified number of them successfully terminate
- **Persistent Storage** - piece of storage in the cluster that has been provisioned by an administrator or dynamically provisioned using Storage Classes

kubectl - CLI for Kubernetes

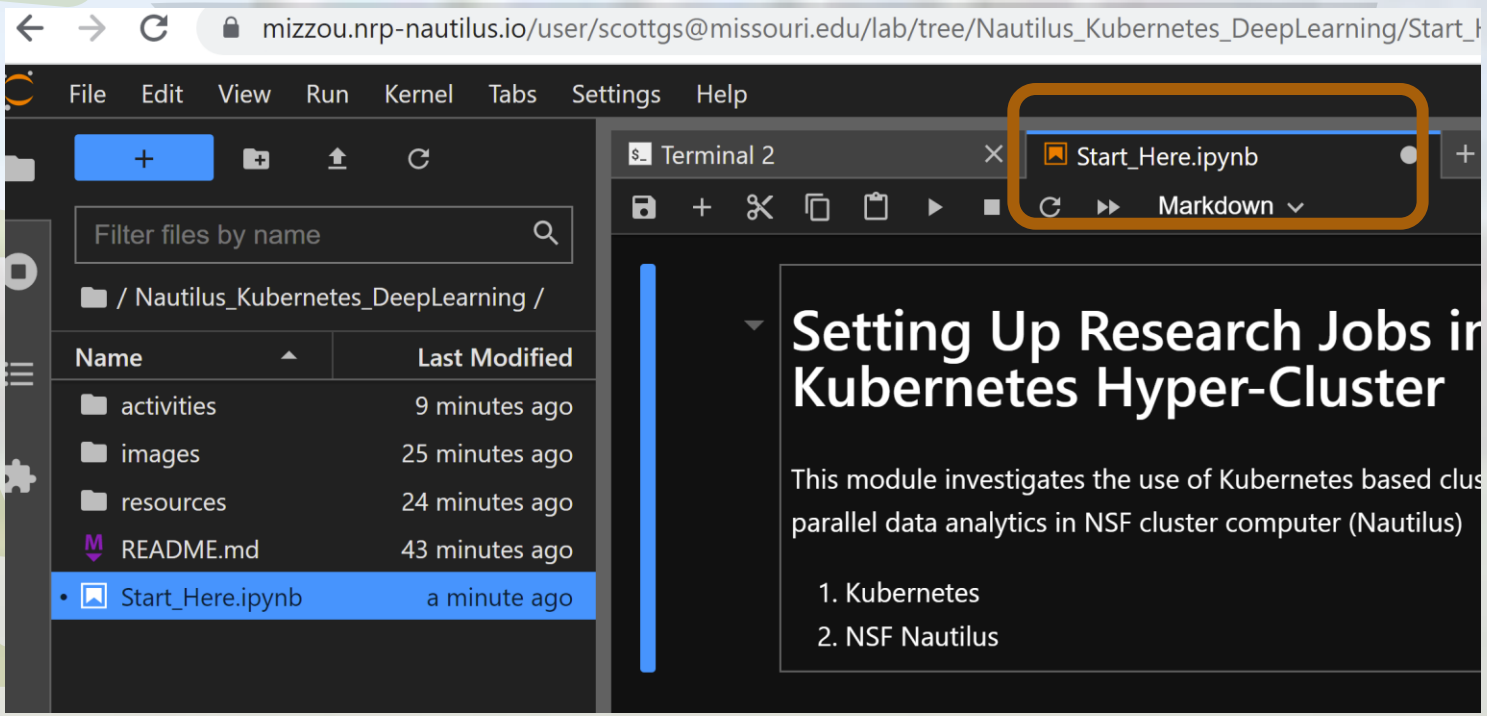
- Allows you to run commands against Kubernetes clusters. You can use kubectl to deploy applications, inspect and manage cluster resources, and view logs
- Can control remote Kubernetes clusters from any number of environments (Linux, Windows, or Mac)

Hands-On Activities - *Setup and Test kubectl*

- **Using Kubernetes tools on Nautilus Jupyter Lab**
- Attaching persistent storage to Pods
- Launch Deep Learning Training on Kubernetes

Hands-On Activities - *Setup and Test kubectl*

- In prior steps, the github repository for this workshop was cloned into your Jupyter Lab environment
- Navigate to the workshop folder
- Click on the **Start_Here** notebook
- New tab opens



Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri

Set Up Kubernetes - *Setup and Test kubectl*

Finish Preparation of Kubernetes

1. Create a Kubernetes configuration directory
2. Upload the config file you got from the Nautilus Portal to Jupyter Lab (home folder)
3. Move into place
4. Install the Kubernetes CLI

```
mkdir ~/.kube
```

```
mv config ~/.kube/
```

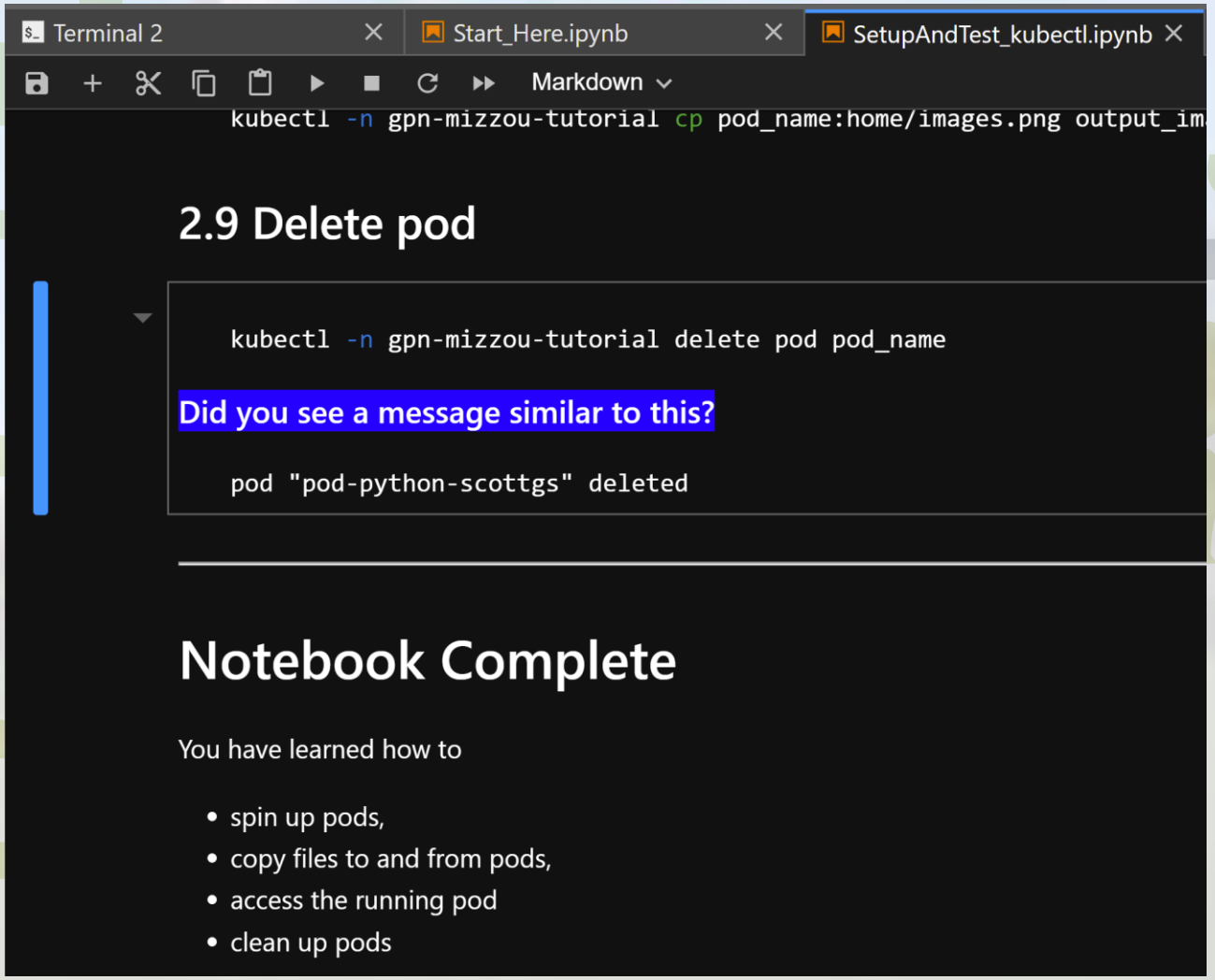
Follow steps on:

https://kubernetes.io/docs/tasks/tools/install-kubectl-linux/#install-using-native-package-management

Set Up Kubernetes - *Setup and Test kubectl*

Work through step 2 of the notebook

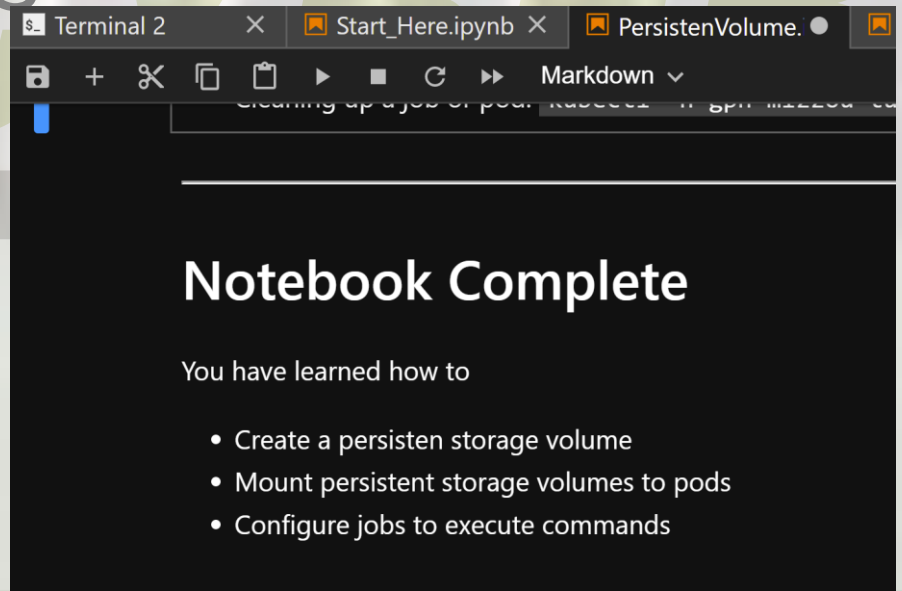
1. Create a pod
2. Copy files to the pod
3. Access the running pod
4. Install software
5. Run script
6. Copy files from the pod
7. Remove the pod



Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri

Hands-On Activities - Create a persistent volume

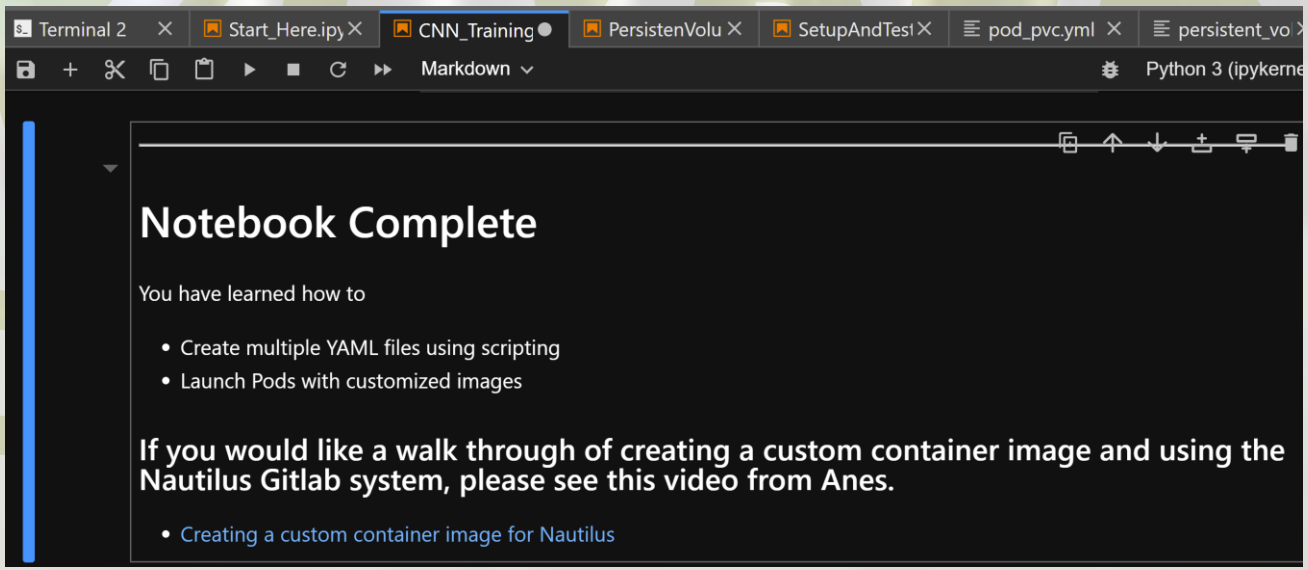
- Using Kubernetes tools on Nautilus Jupyter Lab
- **Attaching persistent storage to Pods**
- Launch Deep Learning Training on Kubernetes



Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri

Hands-On Activities – Convolutional Neural Network on Nautilus

- Using Kubernetes tools on Nautilus Jupyter Lab
- Attaching persistent storage to Pods
- **Launch Deep Learning Training on Kubernetes**



Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri

This completes the workshop activities!
-- We either flew through or are way behind

- You now have a taste of what setting up research tasks, deep learning in this case, on the NSF Nautilus cluster.
- This is just the tip of the iceberg!

Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri

Thank you for working through the activities on this workshop!

This workshop creation and delivery was supported by the National Science Foundation (NSF)

NSF Award Search: Award #1925681 - CC* Team: Great Plains Regional CyberTeam

Questions? Post-Workshop, GPN Slack

Channel #gpn-nautilus-kubernetes-workshop

Dr. Grant Scott - EECS | IDSI | CGI
University of Missouri