UW/Sage workshop tutorial:
Accessing the brain-derived snRNA-seq data from an Alzheimer's Disease case control study

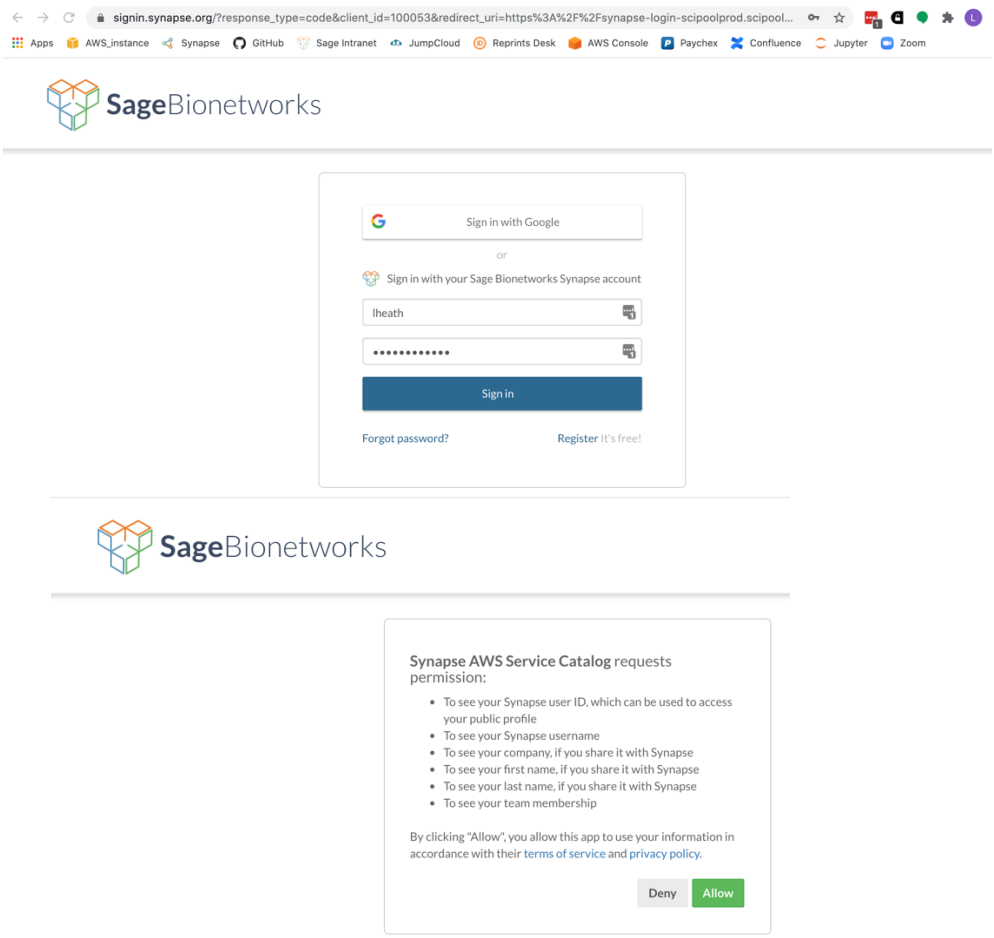
IMPORTANT: You will not be able to access the data in Synapse or Sage's AWS services without completing the following steps BEFORE the workshop (see Anna's workshop notification emails for details):

A. Register for a Synapse account
B. Apply for access to ROSMAP data by filing a data use certificate request form

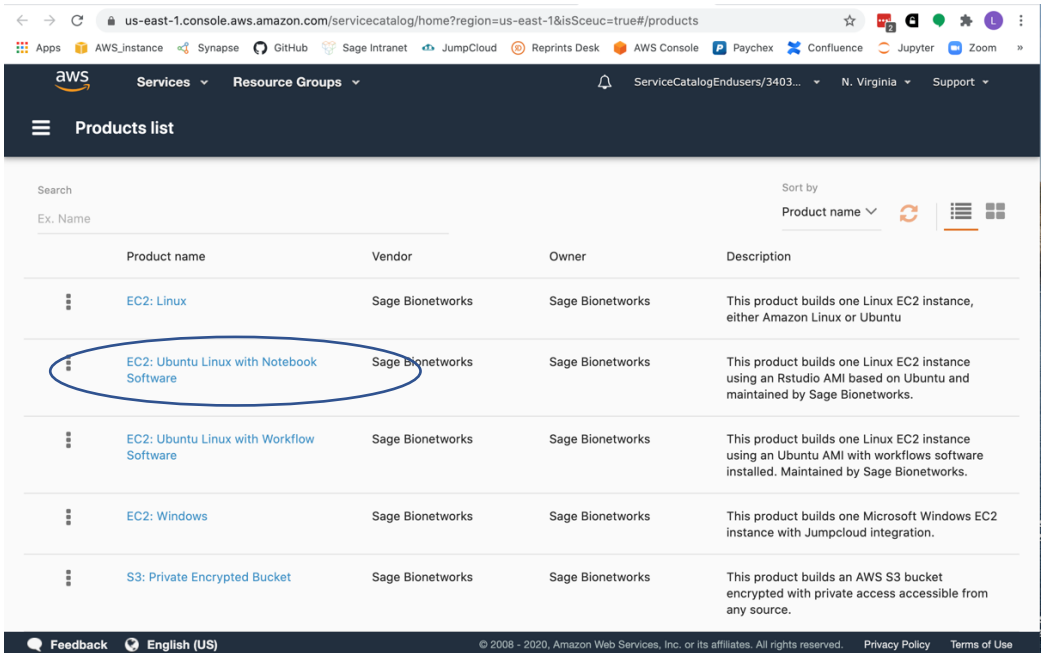
WORKSHOP TUTORIAL:

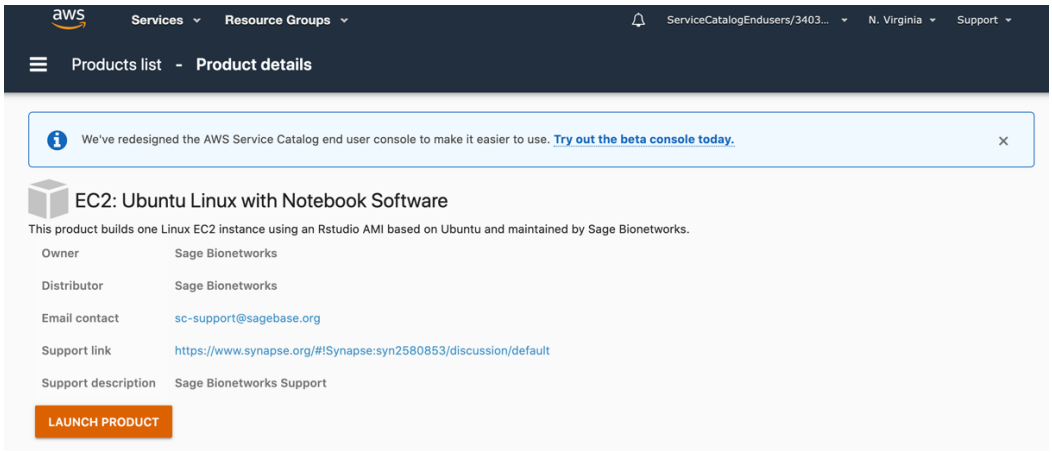
Setting up your Rstudio notebook for the workshop, using Sage's AWS Service catalog.

1. Open a web browser and go to 'sc.sageit.org'. Use your Synapse credentials to log in; click "Allow" when prompted for AWS Service Catalog permissions.

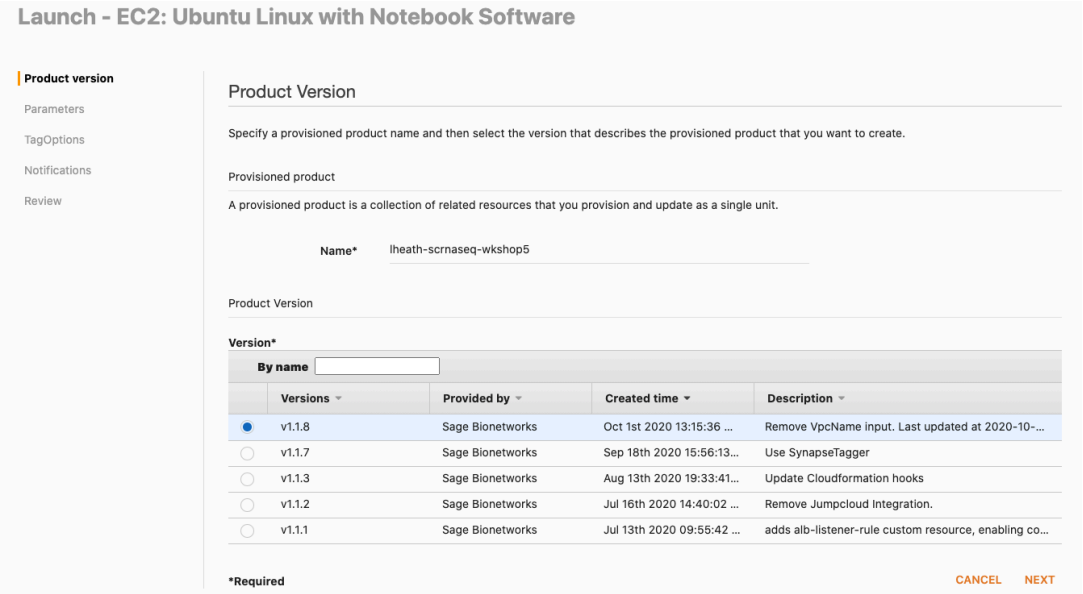


2. After logging in you will land on the “Products” page, listing the available services. Click the link to “EC2: Ubuntu Linux with Notebook Software.” Launch Product.

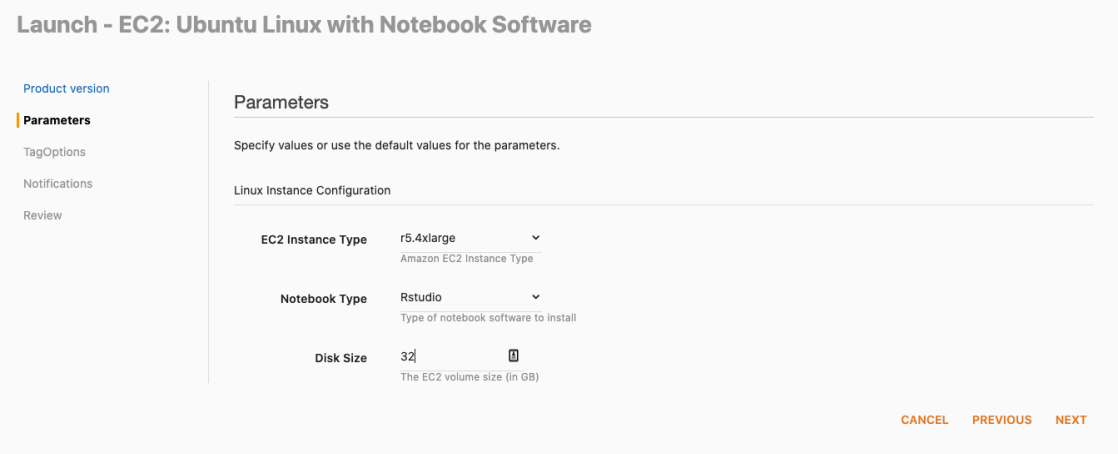




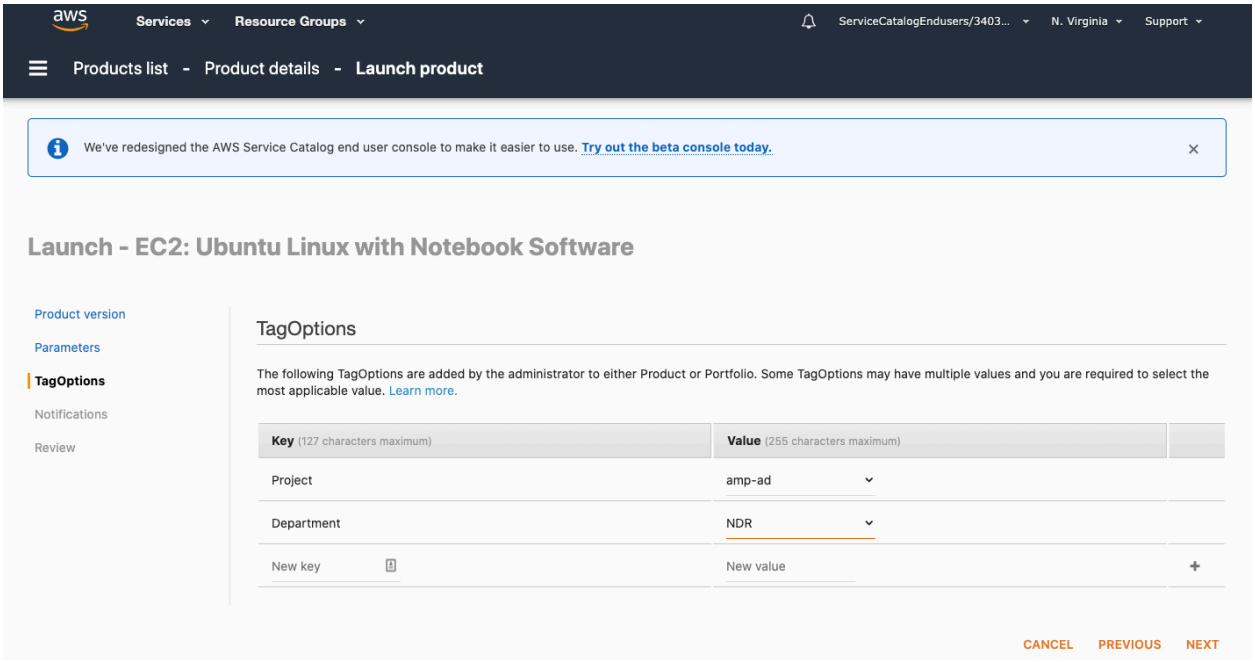
3. You will land on the Product Version page:
    a. Name your instance in the Name* blank
    b. Click button next to the latest version
    c. Click “NEXT”



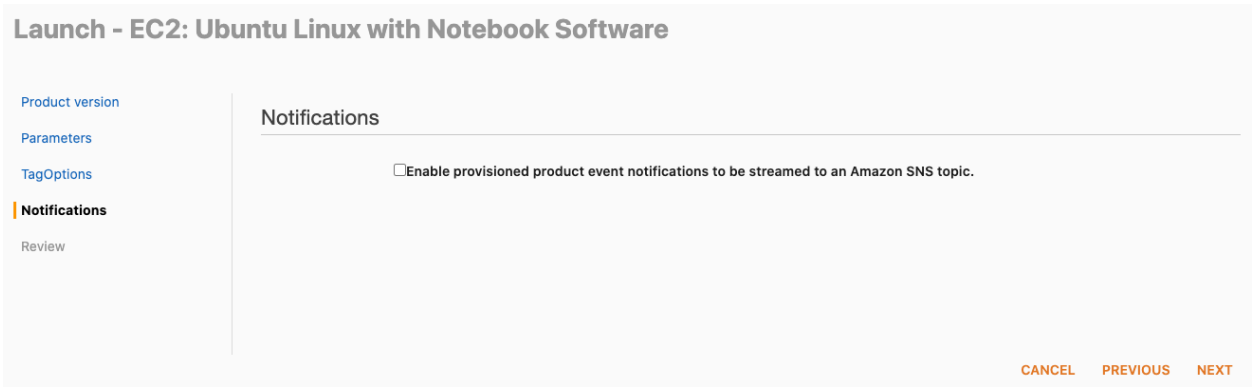
4. Select parameters
    a. For EC2 Instance Type, scroll to “r5.4xlarge.”
    b. Notebook type = Rstudio.
    c. Disk size = 32.
    d. Click “NEXT”



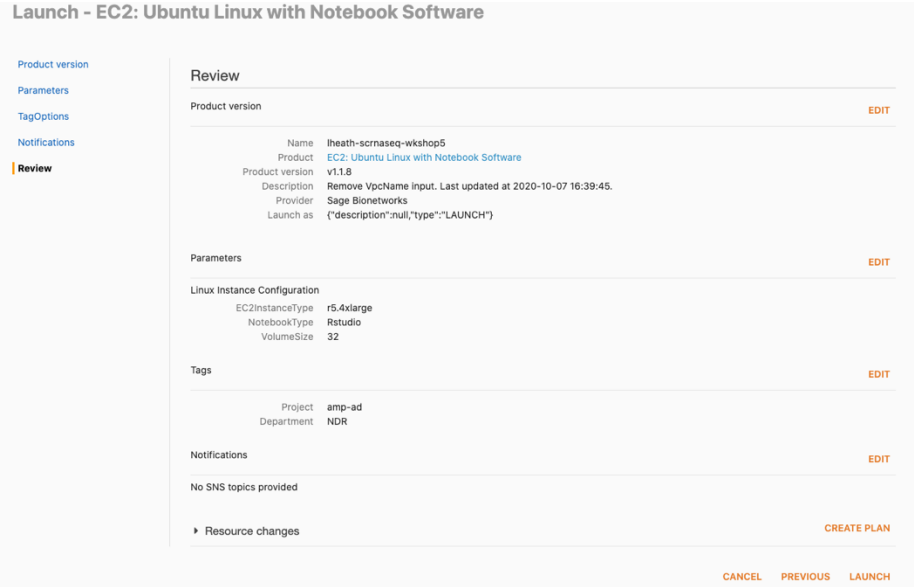
5. Use drop-down menu to fill in TagOptions
    a. Project = amp-ad
    b. Department = NDR.
    c. Leave new key field blank
    d. Click "NEXT"



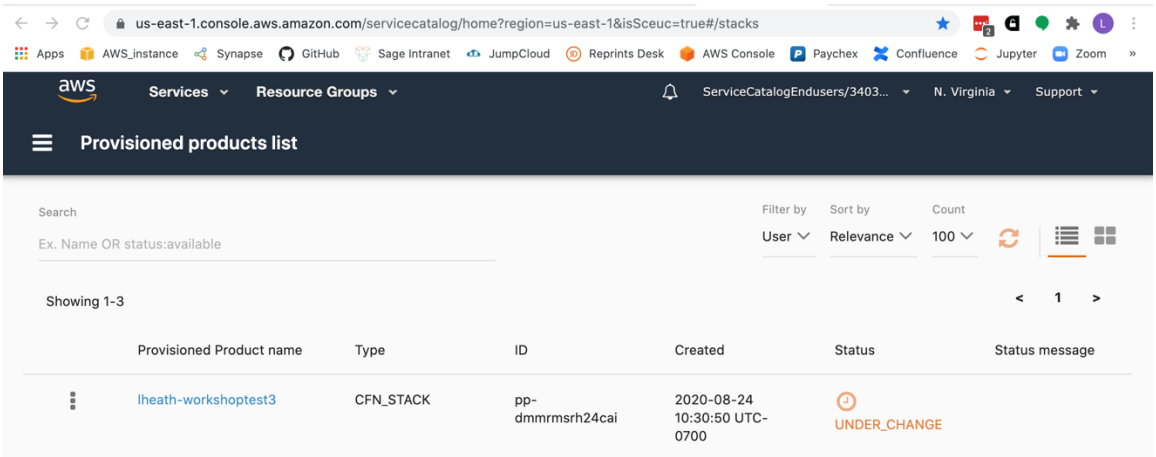
6. Ignore the Notifications page, click "NEXT"



7. Review and launch your instance.



8. Go to the "Provisioned products list" menu (in the menu icon, upper left corner ☰) and wait for your notebook to change status.

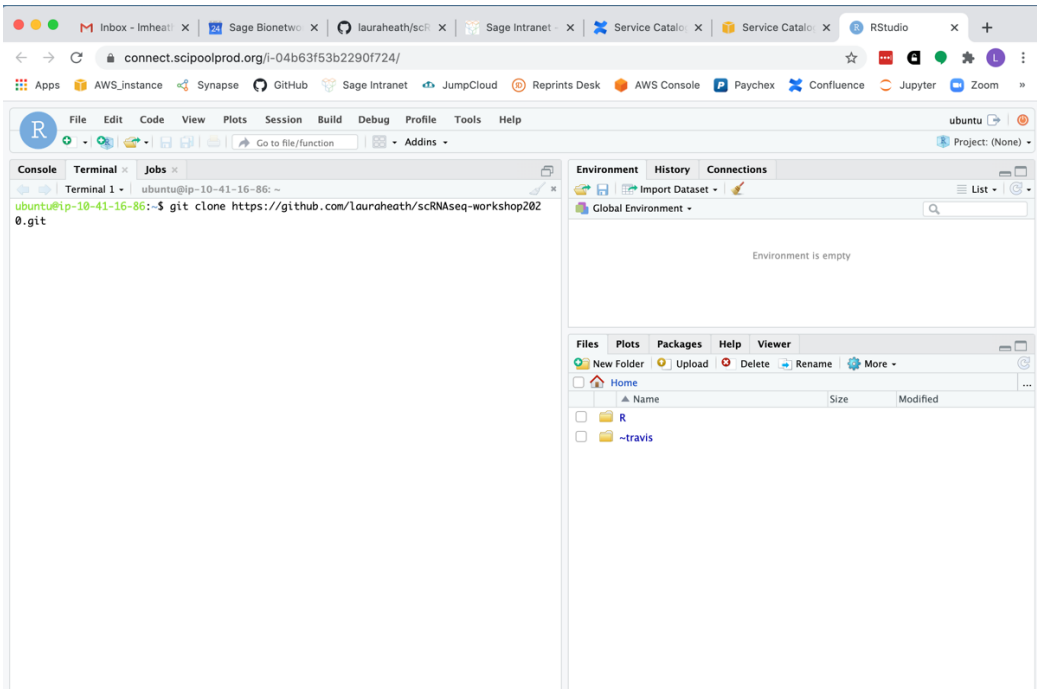


9. When the Status says Available, click on the provisioned product name (in this example, lheath-workshoptest3). Scroll down to the Outputs list and click on the link next to NotebookConnectionURI. This will open an Rstudio notebook in a new tab in your browser.

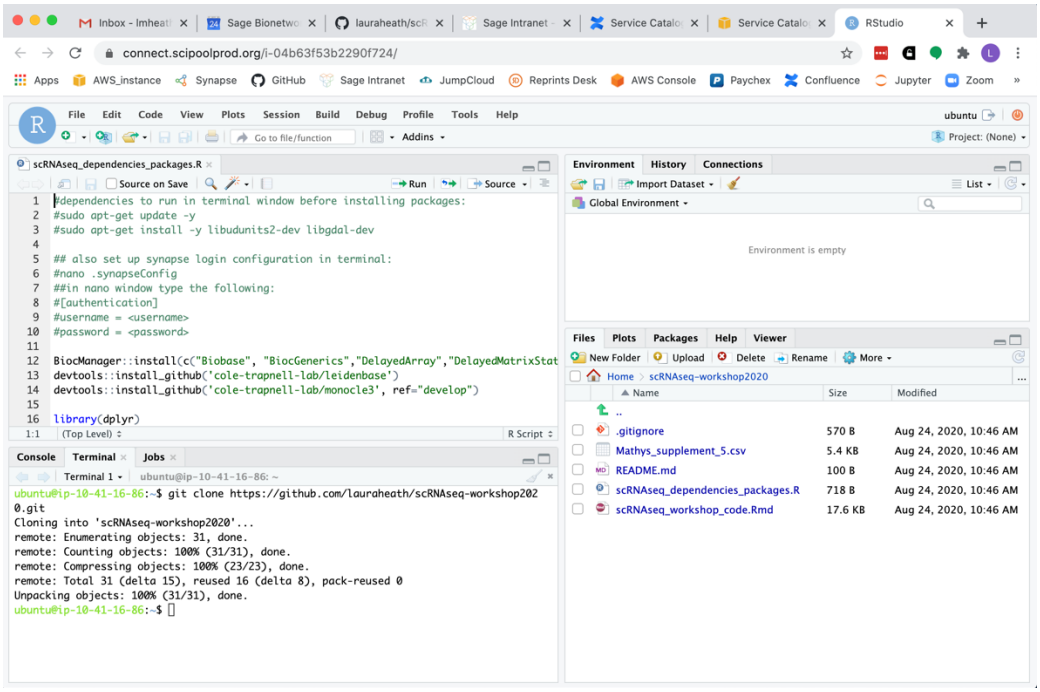
▾ **Outputs:**

Key	Value	Description
CloudformationStackARN	arn:aws:cloudformation:us-east-1:237179673806:stack/SC-237179673806-pp-dmmrmsrh24cai/8e600410-e62f-11ea-8a70-12695902b27f	The ARN of the launched Cloudformation Stack
EC2ConsoleURI	https://console.aws.amazon.com/ec2/v2/home?region=us-east-1#Instances:search=i-04b63f53b2290f724	Check your instance status with this link to the ...
LinuxInstanceId	i-04b63f53b2290f724	The ID of the EC2 instance
LinuxInstancePrivateIpAddress	10.41.16.86	The IP Address of the EC2 instance
ConnectionInstructions	https://sagebionetworks.jira.com/wiki/spaces/IT/pages/996376579/Connect+to+Provisioned+Instances	Guidelines on connecting to instances
NotebookConnectionURI	https://connect.scipoolprod.org/i-04b63f53b2290f724/	Notebook server login page
EC2InstanceType	r4.2xlarge	The EC2 instance type
ConnectionURI	https://us-east-1.console.aws.amazon.com/systems-manager/session-manager/i-04b63f53b2290f724?region=us-east-1	Starts a shell session in the AWS Console

10. Go to the Rstudio Terminal (next to the Console tab, upper left) and clone the scRNA seq workshop repository with the following command:
    git clone https://github.com/lauraheath/scRNAseq-workshop2020.git



11. You will see the scRNAseq-workshop2020 folder added to your files (lower right panel). Click on it, then click on "scRNAseq_dependencies_packages.R".



12. First, we need to install some dependencies. Copy line 2 ("sudo apt-get update -y" without the hashtag), paste into the terminal window prompt, and hit return; then do the same for line 3.

13. To enable easy Synapse login from your notebook, set up your login configuration as follows: after typing "nano .synapseConfig" in Terminal, a nano environment will appear. Type the following, line by line (no hashtags, don't forget the brackets, and use your own Synapse login information:

    [authentication]
    username = yourSynapseUsername
    password = yourSynapsePassword

    Exit nano (control-X on Mac). Type 'y' when prompted to save the .synapseConfig file.

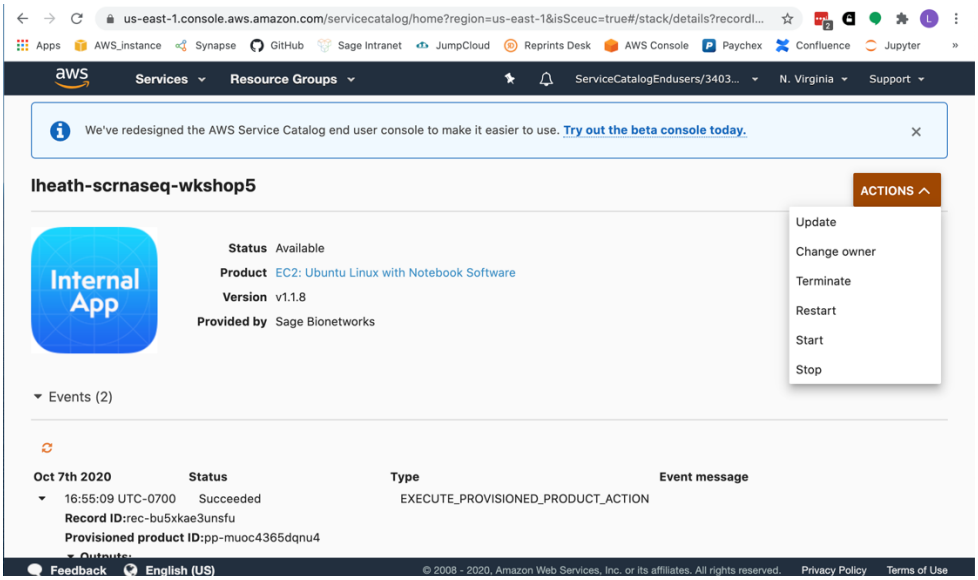
14. Now install the necessary packages one line at a time (Lines 12-14 of scRNAseq_dependencies_packages.R) and install libraries (lines 16-20). Installing the packages will take ~35 minutes. Also, the monocle3 install will immediately prompt you to update existing packages (do update the existing packages by typing “1” when prompted).

NOTE: if errors appear when attempting to load libraries (lines 16-20), restart R session and rerun lines 16-20

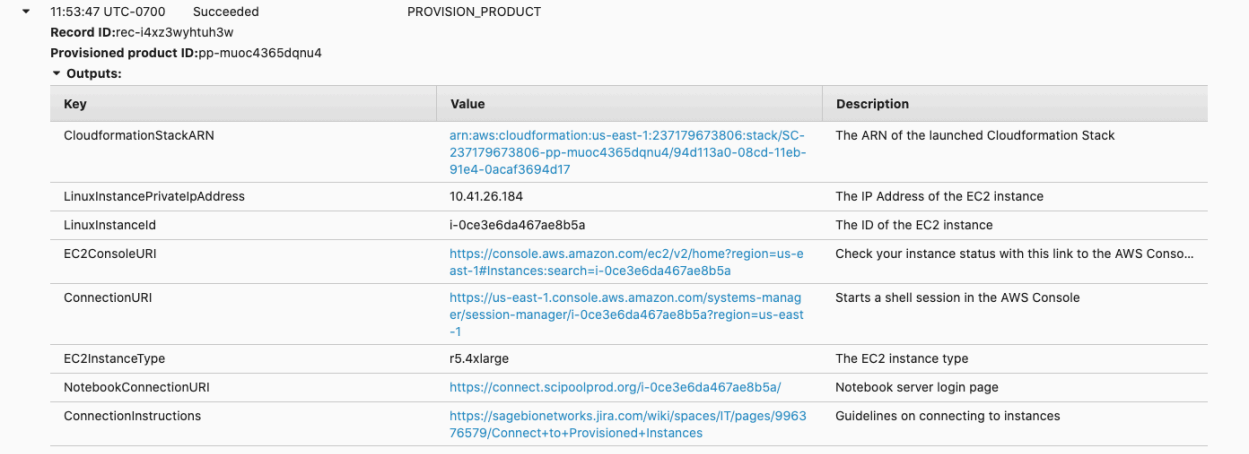
15. Now you are ready to upload scRNA seq data from Synapse and use Monocle3 for trajectory analyses. Launch “scRNAseq_workshop_code.Rmd” and you are on your way!

POST WORKSHOP CLEANUP: When finished working within your notebook, please return to the AWS console and stop or terminate your instance (under the “Actions” button). You can return later and start your instance back up again if you do not terminate it. If you have a github account, you can clone the workshop code and any changes you made today into your own branch. **DO NOT store ROSMAP data in your github repo or outside of the EC2 instance/Synapse environment, per the DUC agreement for ROSMAP**. We encourage users to avoid storing data on their instances in the long term.

Stopping, starting, and terminating your instance:



Finding the Rstudio notebook URL after restarting an instance (go to the most recent “Succeeded” tab, then “Outputs”:



11:53:47 UTC-0700 Succeeded PROVISION_PRODUCT

Record ID: rec-i4xz3wyhtuh3w

Provisioned product ID: pp-muoc4365dqnu4

Outputs:

Key	Value	Description
CloudformationStackARN	arn:aws:cloudformation:us-east-1:237179673806:stack/SC-237179673806-pp-muoc4365dqnu4/94d113a0-08cd-11eb-91e4-0acaf3694d17	The ARN of the launched Cloudformation Stack
LinuxInstancePrivateIpAddress	10.41.26.184	The IP Address of the EC2 instance
LinuxInstanceId	i-0ce3e6da467ae8b5a	The ID of the EC2 instance
EC2ConsoleURI	https://console.aws.amazon.com/ec2/v2/home?region=us-east-1#Instances:search=i-0ce3e6da467ae8b5a	Check your instance status with this link to the AWS Conso...
ConnectionURI	https://us-east-1.console.aws.amazon.com/systems-manager/session-manager/i-0ce3e6da467ae8b5a?region=us-east-1	Starts a shell session in the AWS Console
EC2InstanceType	r5.4xlarge	The EC2 instance type
NotebookConnectionURI	https://connect.scipoolprod.org/i-0ce3e6da467ae8b5a/	Notebook server login page
ConnectionInstructions	https://sagebionetworks.jira.com/wiki/spaces/IT/pages/996376579/Connect+to+Provisioned+Instances	Guidelines on connecting to instances