Training in Cloud Use

The most frequently identified training need across the Satellite Needs Working Group (SNWG) 2022 survey submissions was for "Training in Cloud Use." NASA currently offers access to its Science Managed Cloud Environment, which provides researchers with rapid access to commercial cloud services (through Amazon Web Services) via the use of its Science Analytics Platform. Consulting, training, and support for using cloud-based services are provided for both scientists and system administrators. For self-learning, there is also an Amazon Web Services Cloud Primer for NASA Earthdata and a NASA Earthdata Cloud Cookbook.

NASA's Land Processes Distributed Active Archive Center (LP DAAC) offers an e-learning portal with a number of cloud tutorials, while NASA's Physical Oceanography Distributed Active Archive Center (PO.DAAC) offers step-by-step tutorials on how to use the NASA Earthdata Cloud to access and analyze data. The PO.DAAC community also created a Coding Club to help members learn how to use the cloud.

NASA offers cloud-based data exploration and analysis tools that provide Jupyter-based platforms and require little knowledge of cloud infrastructure. These tools include the Multi-Mission Algorithm and Analysis Platform (MAAP) and the Visualization, Exploration, and Data Analysis (VEDA) project.

NASA's Transform to Open Science (TOPS) mission aims to accelerate the adoption of open science, including cloud-native science, via its introductory open science curriculum called Open Science 101. Additional domain-specific TOPS-funded cloud training materials are in development on topics including environmental justice and wildfire risk. Other NASA-funded initiatives, including CryoCloud and Pangeo, are building cloud-native tools and communities of researchers to support the transition to cloud-based research workflows.

NASA's Office of the Chief Science Data Officer (OCSDO) is utilizing NASA's Space Act Agreement (SAA) with Amazon Web Services (AWS) to offer allocations of Amazon Simple Storage Service (S3) storage and cloud computing to facilitate exploration of best practices around discovery, access, and use of high-value NASA science datasets.

There are also external organizations offering services supporting cloud usage. Openscapes provides training and mentoring for researchers working in the cloud, and CloudBank offers cloud brokerage services for National Science Foundation-funded projects. 2i2c is an organization that offers managed cloud services via Jupyter platforms and runs the cloud interfaces for many of the resources mentioned above.
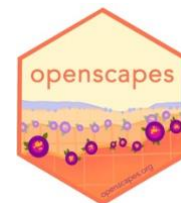
Additionally, many major commercial cloud services providers post training for their offerings. Examples include: AWS, Azure, Google Cloud, Oracle Cloud, and IBM Cloud.