

CYVERSE: INFORMATICS CYBERINFRASTRUCTURE FOR THE EARTH SCIENCES

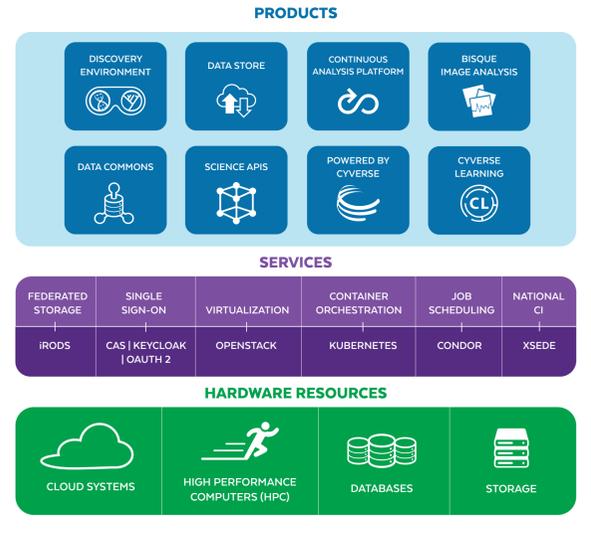
AUTHORS: Edwin Skidmore, Tony Edgin, Ramona Walls, Tyson Swetnam, Eric Lyons, Nirav Merchant  

WITH OVER 6 PB OF DATA TRANSACTED ANNUALLY through its cyberinfrastructure, CyVerse supports full data lifecycle management for research communities from diverse disciplines spanning life sciences to space sciences. Researchers and their teams use CyVerse to perform tasks at scale, such as data sharing, metadata management for high-throughput analysis workflows, visualization coupled with application, and data integration across systems – all in one place.

CyVerse's underlying cohesive infrastructure supports petabytes of data, thousands of users, and hundreds of applications. We are also integrating capabilities to support Machine Learning (ML) based workflows and continuous analysis frameworks.

KEY CROSS-CUTTING FEATURES OF CYVERSE PLATFORMS

- Only one login needed to access all resources.
- Use your preferred access method to upload, download, and access data and computation.
- The iRODS data management system keeps all your data safe and accessible.
- Share your data instantly with other users.
- Only you control your data sharing permissions for security.



Data Store

Whether your data are small or large, structured or unstructured, CyVerse's cloud-based Data Store can fulfill your data storage needs.

- **100GB of free data storage** for each user; larger allocations are available for active analyses.
- **Parallel data transfers support fast uploads and partial transfers** (i.e., interrupted transfers can later resume in the same place).

CyVerse Continuous Analysis Platform (Public release Q3 2020!)

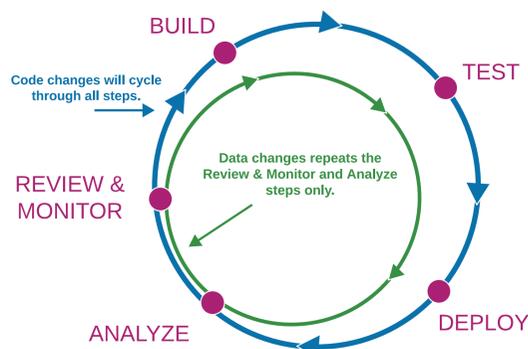
CyVerse Continuous Analyses Platform (CAP) for multi-cloud container orchestration solves key usability issues for researchers using containerized cloud infrastructure by enabling:

- 1) orchestration of compute, storage, and software configuration of containers for reproducible provisioning to any Kubernetes cloud;
- 2) user-defined configuration for auto-scaling and for event-triggered computational workflows; and
- 3) continuous analysis lifecycle management including code repository integration, metering resources, sharing, and collaboration.

CyVerse CAP is built atop Kubernetes and leverages the Cloud Native Computing Foundation's key technologies (cncf.io).

KEY FEATURES

- **Better user experience and usability:** CyVerse CAP users don't need to understand Kubernetes orchestration to create a workflow.
- **Secure, programmatic access:** Secured with OAuth2/OpenID, developers can integrate with CyVerse CAP's Rest API to enhance and extend their own infrastructure.
- **Multi-cloud:** CyVerse CAP deploys to any Kubernetes-based cloud.
- **Event-driven:** Trigger workflows by specific events (e.g., code or data changes), time periods (e.g., hourly), or at user-defined custom events (e.g., web hooks and watchers).
- **Cloud-agnostic data access:** Use data from the CyVerse Data Store or any Kubernetes-supported cloud data source.
- **Integrated with Git and standard container registries:** Workflow definitions and container images are retrieved and stored in standard repos and container registries.



Discovery Environment and VICE

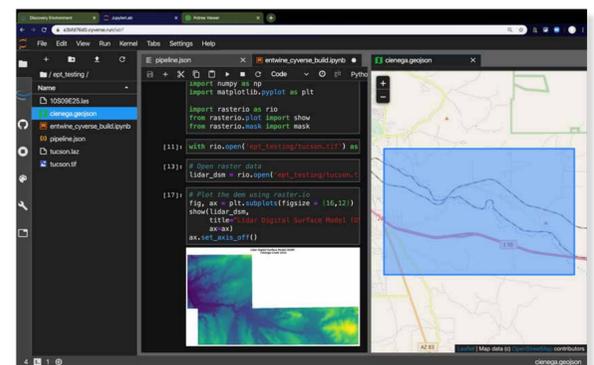
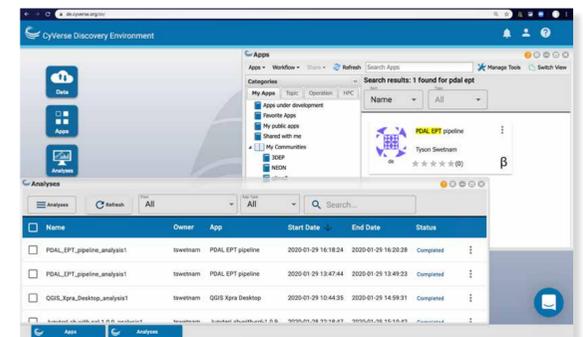
Get the job done simply and easily using the Discovery Environment (DE)! No need to master the command-line or new software. The DE facilitates data exploration and scientific discovery by handling all your bioinformatics workflows – from data and metadata management to analysis and sharing large datasets.

Within the DE, use VICE (Visual and Interactive Computing Environment) to launch modern web-based platforms like Project Jupyter, RStudio, and Shiny. VICE provides secure https access, handles data transfers, and returns the results back to the Data Store, so you never have to download data or move between platforms to analyze your results.

Easily build your own customized analysis environment by bringing the tools you need into a Docker container, where you can share them with your team or make them public. If you're not sure where to begin, check our tutorials or attend our hands-on bootcamp: cyverse.org/learning.

KEY FEATURES

- **Analyze Data:** Access and run hundreds of highly scalable bioinformatics tools and apps, as well as open source utility tools and scripts thanks to integration of CyVerse's data management system with compute resources.
- **Manage Data:** Easily and securely manage and share data, workflows, analyses, results, and data visualizations with collaborators or the community.
- **Customize:** Add a command-line tool or an executable into the tool list and create or edit a new customized interface for it.
- **Extend:** Use APIs to integrate tools and resource-intensive analyses, or to build an automated sequenced workflow by connecting apps together.

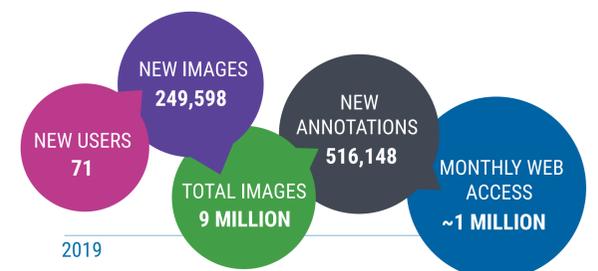


BisQue

BisQue is integrated with CyVerse's authentication system, Data Store, and compute infrastructure for seamless scalability and ready access to a large set of downstream analysis options.

KEY FEATURES

- **Developers can integrate existing applications or create new ones,** leveraging BisQue's rich set of custom visualizations, image handling routines, and APIs for building scalable, web-based image analysis applications.



CONNECT WITH US



CyVerse Learning offers tutorials, workshops, webinars, and camps. Go to CYVERSE.ORG/LEARNING and get started today!



Cold Spring Harbor Laboratory



DBI-0735191, DBI-1265383, DBI-1743442 and OAC-1664172