NDS Labs Workbench update

Craig Willis
willis8@illinois.edu
NCSA/NDS

National Data Service Workshop (NDS7)
April 13, 2017
NDS Labs Workbench (beta)

https://www.workbench.nationaldataservice.org

- NDSC initiative started in ~January 2016
- Community-driven platform to share, discover, evaluate, develop, and test research data management and analysis tools
- Open platform -- community members recommend and/or contribute tools
What is Labs Workbench?

- Scalable platform
- Web-based, turn-key system for the deployment of data tools
- Includes catalog of community-contributed tools
- Deploy in seconds
- A personal test environment or sandbox

Personal catalogs
Data management training and incubator platform

https://www.workbench.nationaldataservice.org

**Incubate data projects**: Use the demonstration instance to explore, develop, and experiment with data management and analysis tools.

**Promote tools**: Add your tools to the catalog for others to discover.

**Experiment**: Instantly deploy turn-key applications into a personal sandbox directly from your browser.
Hosted private or cloud instances

https://github.com/nds-org/ndslabs

**Become an NDS node:** Install your own instance or host through NDSC.

Provide **scalable, consistent cloud-based environments** for workshops/training.

**In-place analysis:** Allow users to launch custom analysis environments near data.

**Customizable** interface and catalog
Use case: TERRA-REF Analysis Workbench

- Initially used for workshop tutorial to provide consistent environments, scaling to support ~50 participants.
- Customized Labs Workbench instance hosted as part of TERRA-REF infrastructure at NCSA.
- Custom analysis and development environments with direct access to TERRA-REF data. Used by collaborators and alpha users.
Use case: Data Curation Education Workbench

- NDS pilot project initiated during NDSC6, led by Carole Palmer at the University of Washington iSchool.
- Customized instance of Labs Workbench hosted on SDSC Cloud.
- Sandbox for students to get hands-on experience with data curation systems and tools including CKAN and Dataverse.
On the horizon:

- ThinkChicago

- Potential use as analysis engine in data management systems
  - DataDNS, Cyverse

- Simplified installation (balancing scalability with simplicity)

- Improved courseware integration (iSchool pilot)

- Enable deployment in more environments, including commercial cloud providers
Contribute and collaborate

● General Q&A (Google group)
  ○ [https://groups.google.com/forum/#!forum/ndslabs](https://groups.google.com/forum/#!forum/ndslabs)
  ○ Ask questions, propose ideas, announce your services

● Code and issues (Github):
  ○ Fork, create pull requests, report issues
  ○ [https://github.com/nds-org/ndslabs](https://github.com/nds-org/ndslabs)

● Developer chat (Gitter)
  ○ [https://gitter.im/nds-org/ndslabs](https://gitter.im/nds-org/ndslabs)