



the **ENERGY** lab

R&D FACTS

Geological & Environmental Sciences

EDX, NETL's Data-Driven Tool for Science-Based Decision Making

Data Exchange for Energy Solutions



Background and Benefits

In 2011, development of the National Energy Technology Laboratory (NETL)'s Energy Data eXchange (EDX) was initiated as an innovative solution to research challenges by offering (1) a means for better preserving of our own research and development (R&D) products for future access and re-use, (2) efficient and easily discoverable access to authoritative, relevant, external data resources, and (3) an improved approach and suite of tools to support secure and private collaboration and coordination between multi-organizational R&D teams to meet the U.S. Department of Energy (DOE), and specifically NETL mission and goals.

In 2010, two energy (R&D) related efforts supported by the NETL researchers illustrated a strong and growing need to change how NETL and collaborating research partners think about research products and cultivate new research. The first of these R&D efforts was related to the increasing uncertainty surrounding safety and risks associated with hydraulic fracturing of unconventional resources. As development of these resources took off at an unprecedented pace in 2007, it sparked renewed interest in leveraging products from DOE and NETL Eastern and Western Gas Shales R&D Programs from the 1970s through the 1980s. However, many of the datasets resulting from those programs were stored in filing cabinets as paper-based assets or on outdated media. Thus, making discoverability, accessibility, and reuse of valuable information and data from those studies challenging, if not impossible, to leverage for current-day R&D needs.

Second, in 2010, the BP Deepwater Horizon spill in the Gulf of Mexico (GOM) sparked a large federal response, including multi-agency teams involving DOE and NETL researchers. This involvement spotlighted challenges that these multi-agency teams experienced in (1) efficiently finding and accessing relevant datasets needed for response analyses and plans, (2) securely working on, sharing data, and developing together products across multi-agency teams, and (3) ensuring preservation of products resulting from that effort for future use. While the specifics of these two examples are unique, they exemplify challenges NETL research teams face every day. This recognition sparked a desire within NETL to leverage rapidly evolving technology, capabilities, and approaches to information sharing, big data, and computational resources, both public and private, for the benefit of NETL researchers, partners, and public stakeholders.

EDX incorporates a broad set of subsurface information common to CO₂ storage and other energy needs (e.g., shale gas, tight oil, deepwater and ultra-deepwater, and unconventional fossil resources). This set of information includes reservoir data, fluids properties, wellbore data, fault/fracture data, and groundwater data. Although some of the information resides in EDX as data derived from NETL research, much of the information exists online distributed in external databases. In these cases, EDX serves as a clearinghouse, allowing NETL and associated researchers to locate data rapidly by serving as a

CONTACTS

OFFICE OF RESEARCH AND DEVELOPMENT



Kelly Rose

Principal Investigator
Research Physical Scientist
541-967-5883
kelly.rose@netl.doe.gov

Dustin McIntyre

Research Title
304-285-1374
dustin.mcintyre@netl.doe.gov

Cynthia Powell

Acting Focus Area Lead
541-967-5803
cynthia.powell@netl.doe.gov

NATIONAL ENERGY TECHNOLOGY LABORATORY

Albany, OR • Anchorage, AK • Morgantown, WV • Pittsburgh, PA • Sugar Land, TX

Website: www.netl.doe.gov

Customer Service: 1-800-553-7681

Visit us at edx.netl.doe.gov



U.S. DEPARTMENT OF
ENERGY

portal to these other datasets. Through this coordinated approach, EDX addresses one of the key lessons learned during DOE's work on the Deepwater Horizon oil spill—namely, that locating and accessing data across a range of sources is challenging and often inefficient with currently available resources.

EDX serves as an online system to facilitate internal access to research that crosscuts multiple NETL projects and programs, with external access to technical products and data published by NETL research teams. In this role, EDX facilitates coordination of both restricted access and open-access research data. NETL-affiliated researchers use EDX's collaborative workspaces to coordinate and share work with a variety of organizations and institutions in a secure environment.

Ultimately, EDX seeks to ensure improved access to data and resources from a range of sources, offering a venue for the "publication" and dissemination of new datasets as well as historical, and often inaccessible, assets ensuring their use for future, yet-to-be envisioned purposes.

The primary users of EDX are NETL researchers (NETL research teams) actively engaged in work relevant to subsurface, near surface, atmospheric, and environmental risk issues relating to subsurface CO₂ storage, unconventional and conventional hydrocarbons, and ground water and air emission impacts. EDX provides access to evaluate and predict what happens in engineered-natural systems, while helping accelerate further research.

Functionality

EDX past, Version 1 released end of FY12

- Formed the foundation for EDX to serve as a coordination-collaboration platform for NETL
- Focused on improved discoverability resources and tools for R&D:
 - Preservation of data resources resulting from NETL affiliated R&D
 - Connect EDX users to publically available, authoritative data resources

EDX present, Version 2 released in FY13

- Brings private sharing capabilities for multi-organizational R&D teams to the EDX platform and to the EDX user community
- Provides the envisioned balance between public data resource discoverability for mature research products and secure-private research development capabilities

EDX today offers:

- Secure, online coordination and collaboration platform that supports energy research, knowledge transfer, and data needs
- Enduring and reliable access to historic and current R&D data, data driven products, and tools
- Both public and secure, private functionalities

Datasets physically housed within EDX are provided by users either as links to external websites or when appropriate, as standalone files such as Microsoft® Excel, .jpg, .zip, etc. Datasets

can be "published" in their original and complete form in EDX and accompanied by associated reports, dissertations, or metadata. EDX also recognizes that there are a significant number of established online resources and offers the ability to store links to external online data, thus improving coordination with existing resources to EDX users.

Contributing to EDX is quick, easy, and streamlined. The process begins with completion of the online submission form to describe attributes, characteristics, and key-words. This information serves as the building blocks of EDX and is utilized to compile search results. Providing thorough and accurate submission information will enhance visibility.

EDX Version 2 Highlights

- EDX Version 2 provides advanced coordination, collaboration, and data visualization functionality
- EDX supports *Group* functionality that enables researchers to share data and information about a common theme, discipline, or interest in an open-access environment
- The *Collaborative Workspace* is an extension of the *Group* functionality and provides NETL research teams a secure and dedicated workspace to quickly and efficiently share data, ideas, and research techniques
- *Slate* is a new *Collaborative Workspace* feature where members can create custom page content for their research
- *EDXtools* provide access to data and information assembled as custom themes, allowing researchers to interact with data
- *EDXwiki* is an energy-focused tool where NETL researchers are able to collaboratively add, edit, and delete an encyclopedia of energy-related content. *EDXwiki* is intended to promote content creation, modification, and deletion at the user level

EDX Next Steps

Going forward, the EDX is poised to develop and release Version 3 with an anticipated release in FY16. The priorities for EDX version 3 include:

- Integrate big data capabilities, underpinned by the HADOOP ecosystem and other open source big data resources, and advanced multi-organizational coordination and collaboration capabilities to the system
- Continue to connect and coordinate with other DOE computer and information science related capabilities to enhance both the individual EDX user's experience and option and to support discoverability of other DOE data and computing R&D resources and capabilities
- "Productize" EDX to enable full leveraging of EDX's federated and nodal capabilities in support of multi-organizational connectivity and research support. EDX Prime, the current NETL hosted and supported version of EDX, will continue to be the main EDX system and will leverage NETL's inherent governmental capabilities either as part of the EDX product or as part of the EDX collaboration community.
- Registered users interested in options for specialty data sets and customized solutions should contact EDX Support at EDXsupport@netl.doe.gov