NETL'S DATA DRIVEN TOOL
FOR SCIENCE-BASED DECISION MAKING

EDX
Energy Data eXchange

Data Exchange for Energy Solutions
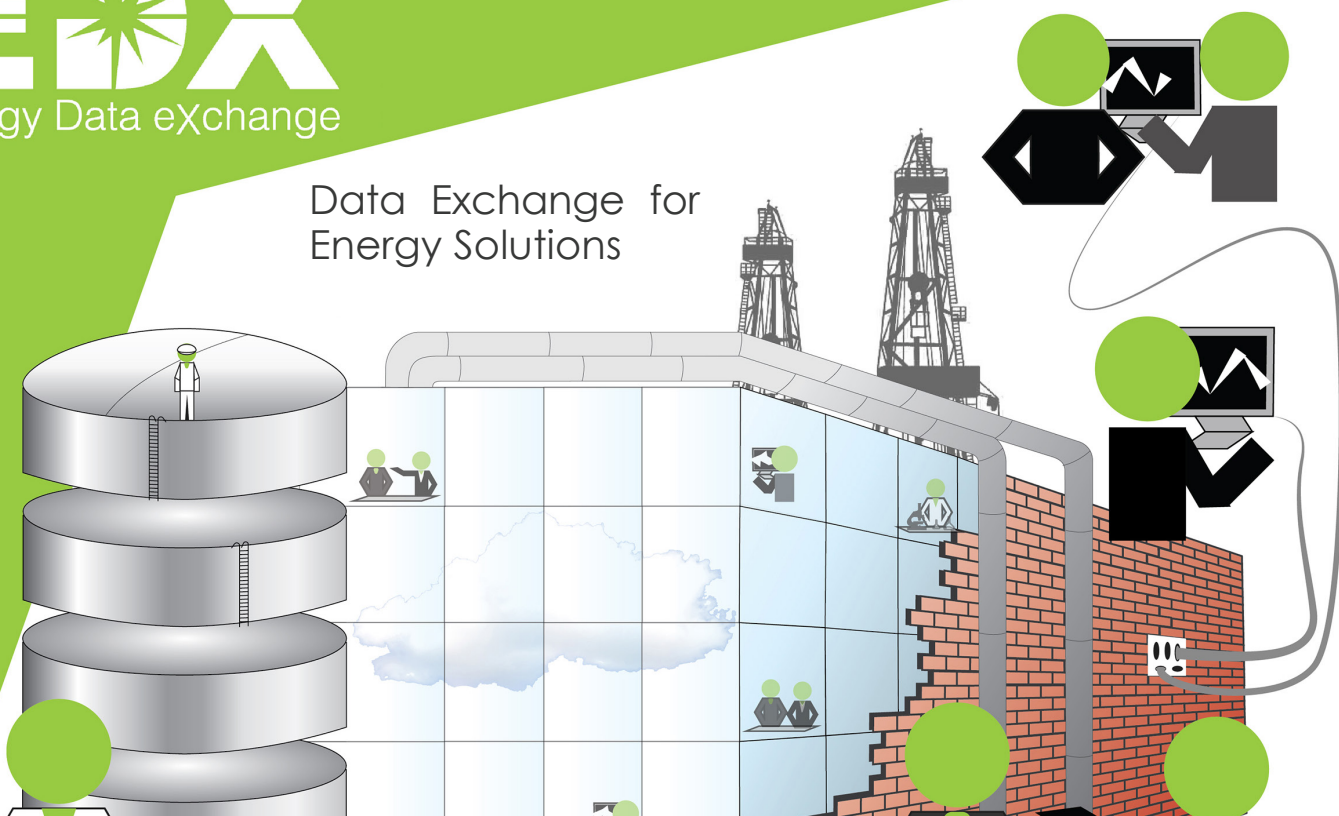
NETL
NATIONAL ENERGY TECHNOLOGY LABORATORY

BACKGROUND AND BENEFITS

In 2011, the U.S. Department of Energy's (DOE) National Energy Technology Laboratory (NETL) initiated the Energy Data eXchange (EDX), an online collection of capabilities and resources developed to support a wide range of energy-related research. Since then, EDX has helped preserve billions of dollars of federally-funded research and connected stakeholders to vital data from collaborative U.S. government, academic and commercial industry energy research. EDX development and maintenance are driven by researchers and technical staff at NETL to support ongoing efforts and technology transfer for DOE-NETL research. EDX supports a variety of research needs by ensuring: (1) long-term curation of both historic and current data and information from a wide variety of sources; (2) reliable access to research that crosscuts multiple NETL projects and programs and energy science needs; (3) efficient discovery of data and information, including technical products from NETL-affiliated research; and (4) innovative analytics to support today's research needs, including supporting secure collaboration and coordination between various agencies, organizations, and institutions through EDX's Collaborative Workspaces.

Albany, OR • Anchorage, AK • Morgantown, WV • Pittsburgh, PA • Houston, TX

U.S. DEPARTMENT OF ENERGY

NETL NATIONAL ENERGY TECHNOLOGY LABORATORY

EDX: NETL'S DATA DRIVEN TOOL

The seed for EDX grew out of needs regarding how NETL and its collaborating research partners share, access, and cultivate research identified from two separate energy-related NETL research and development (R&D) efforts in 2010.

The first of these R&D efforts centered around increasing safety and other concerns associated with hydraulic fracturing of unconventional resources. The rapid development of unconventional resources sparked a renewed interest in research products generated from DOE and NETL eastern and western gas shale R&D programs in the 1970s and 1980s. However, many of the products of this research were stored in filing cabinets as paper-based assets or saved on outdated media. This made the discoverability, accessibility, and reuse of valuable data from those studies challenging, if not impossible, to support renewed R&D needs. The second R&D effort stemmed from DOE and NETL researchers involvement in the federal, multi-agency response to the 2010 BP Deepwater Horizon oil spill (Macondo blowout) in the Gulf of Mexico. This involvement spotlighted challenges facing the multi-agency teams, such as the need to (1) efficiently find and access relevant datasets needed for response analyses and plans, (2) securely share data and work together to develop products across multi-agency teams, and (3) ensure 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 and external collaborators face every day. This recognition ignited an eagerness within NETL to utilize the 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 through the development of EDX.

EDX serves as an online coordination and collaboration platform providing access to a collection of capabilities and resources to support a range of energy-related research needs. EDX is designed to meet the demands of research by ensuring: (1) long-term curation of data and information, (2) reliable access to research data and products, (3) efficient discovery of data and information, and (4) innovative analytics to support today's research needs. Today, EDX enables discovery and access to a range of energy data and information pertinent to carbon storage, hydrocarbon (e.g., shale gas, tight oil, offshore deep- and ultra-deep water, and unconventional fossil resources), energy infrastructure (e.g., transmission lines, pipelines, etc.), and other energy systems. The data and information available in EDX are a combination of NETL research products as well as data from other authoritative sources.

EDX has both a public and private side to facilitate external and internal access to research that crosscuts multiple NETL projects and programs. The public side of EDX supports discovery and access to publicly available data from NETL as well as open-access data from authoritative, external sources. The private side of EDX supports research collaboration and coordination using role-based security privileges and access rights. These secure, Collaborative Workspaces allow designated EDX members access to both restricted and open research data while providing a secure workspace for sharing and developing innovative research and technology.

Ultimately, EDX seeks to improve researcher access to data and resources from a range of sources and offers a venue for the "publication" and dissemination of new data and often-inaccessible historical assets. In doing so, EDX and NETL are helping to ensure use of data and resources for future, yet-to-be-envisioned purposes.

The primary users of EDX are NETL researchers actively engaged in work relevant to subsurface, near-surface, atmospheric, and environmental risk issues relating to subsurface CO_2 storage, unconventional and conventional hydrocarbons, groundwater and air emission impacts, and risks to energy infrastructure. EDX provides access to evaluate and predict what happens in engineered-natural systems while helping accelerate further research.

# EDX: NETL'S DATA DRIVEN TOOL
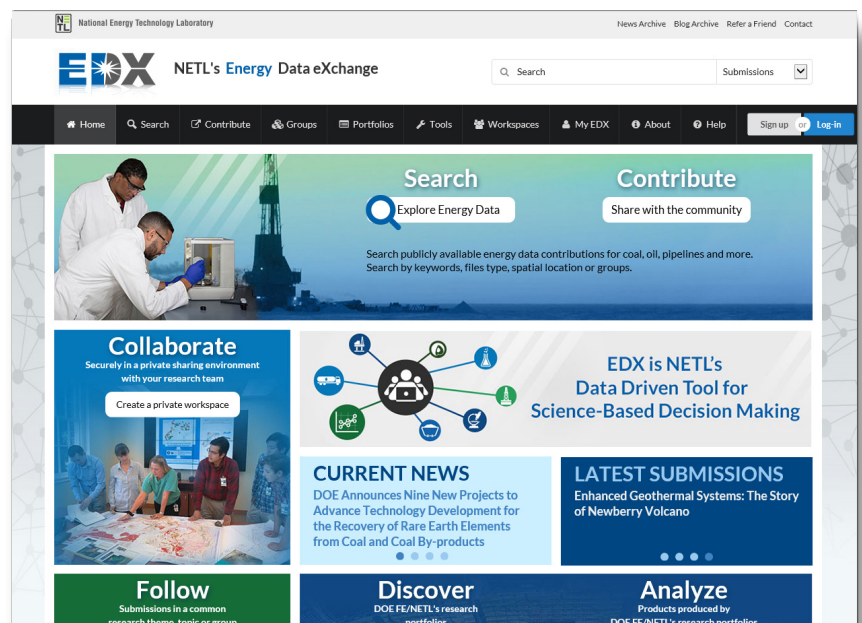
Users can contribute data and resources to EDX rapidly through an easy and streamlined process. The process helps describe data in EDX, noting key attributes, characteristics, and keywords that serve as the building blocks of EDX utilized to compile search results. Providing thorough and accurate submission information enhances discoverable, visible, and usable resources in EDX. Datasets physically housed within EDX are provided by users either as links to external websites, or when appropriate, as stand-alone files.

KEY EDX FUNCTIONALITY

EDX is a multi-faceted platform purposefully designed to respond to a range of end-user needs. For a full accounting of EDX functionality and capabilities please see https://edx.netl.doe.gov/about or contact the EDX team, EDXsupport@netl.doe.gov.

Key highlights and capabilities of the system include:

- Improved discoverability of resources and tools for R&D
  - Connected data resources resulting from NETL-affiliated R&D.
  - Connected publicly available, authoritative data resources.
- Private sharing capabilities for the EDX user community using Collaborative Workspaces
  - Secure and dedicated workspace to quickly and efficiently share data, ideas, and research techniques among authorized users and collaborators.

Screen shot of Geocube tool. Geocube allows users to map spatial data hosted within EDX and from external sources as well.

- Group functionality enables researchers to share data and information about a common theme, discipline, or interest in an open-access environment.
  - **Knowledge Management Database** is a repository of archived oil and gas documents and topical reports stored in the DOE Office of Science and Technology (OSTI) library.
  - **Slate** is a Collaborative Workspace feature where members can create custom page content for their research.
- **EDXtools** is an application store within EDX where models, tools, and applications resulting from DOE Fossil Energy and Carbon Management (FECM) research are available. EDXtools also hosts links to open-source, online tools that are relevant to DOE FE researchers. Examples of the EDXtools currently in the store include:
  - **Geocube**, an online web mapping application hosted by EDX and used to serve spatial data information resulting from and related to DOE FE research and systems.
  - **GeoWell**, a map-based application that provides quick access to websites of primary sources of subsurface geologic and wellbore (oil, gas, and underground injection) information for appropriate U.S. state, tribal and federal agencies.
  - **MFIX (Multiphase Flow with Interphase Exchanges)**, an open source, general-purpose computer code developed at NETL for describing the hydrodynamics, heat transfer, and chemical reactions in fluid-solids systems. Users can download the code, documentation, and see examples of code application.
- **EDXwiki** is an energy-focused tool where NETL researchers are able to collaboratively add, edit, and delete items in an encyclopedia of energy-related content. EDXwiki is intended to promote content creation, modification, and deletion at the user level.
- **Storage Capacity** was expanded to accommodate EDX growth.
- **Two Factor Authentication** for users logging into EDX, adds an additional level of security.
- **Correlated Content** is a custom, machine learning algorithm developed by the EDX team. It is an indexing tool that mines data submissions and suggests additional EDX submissions with similar content.
- **Mobile-Friendly User Interface** allows users to easily explore EDX on mobile and tablet devices.
- **OSTI (Office of Scientific and Technical Information)** integration used to append DOI numbers to submission citations making them more discoverable.
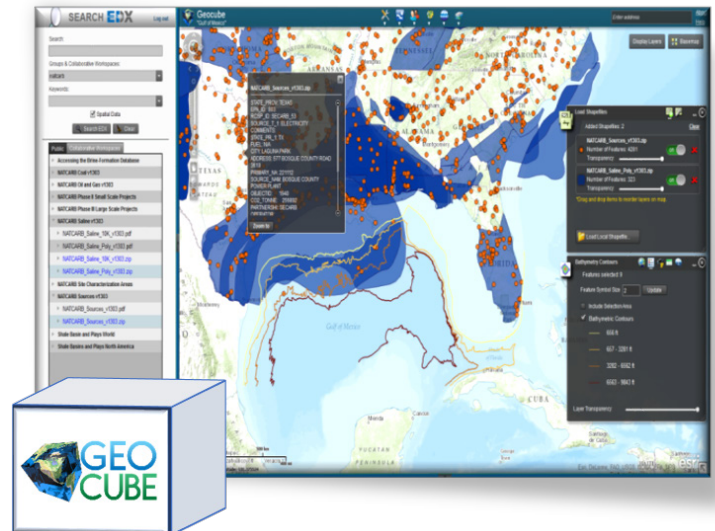
EDX NEXT STEPS

Ultimately, EDX is DOE FECM data and knowledge management platform. It will continue to evolve to serve the DOE FECM R&D community to curate products of DOE FE R&D and provide research teams with virtual, advanced data computing capabilities to facilitate ongoing FE research. Next steps in the evolution of the EDX virtual data library and laboratory system include:

- Curation of DOE FE data products resulting from carbon storage and oil and gas program research and integration with existing, complementary open resources to develop a virtual subsurface data framework for the U.S.

- Revision of the EDX private collaborative workspace environments to improve intuitiveness of these spaces, integrate EDXtools into the workspace to support online data analytics, and other enhancements to support team collaboration and research efforts.
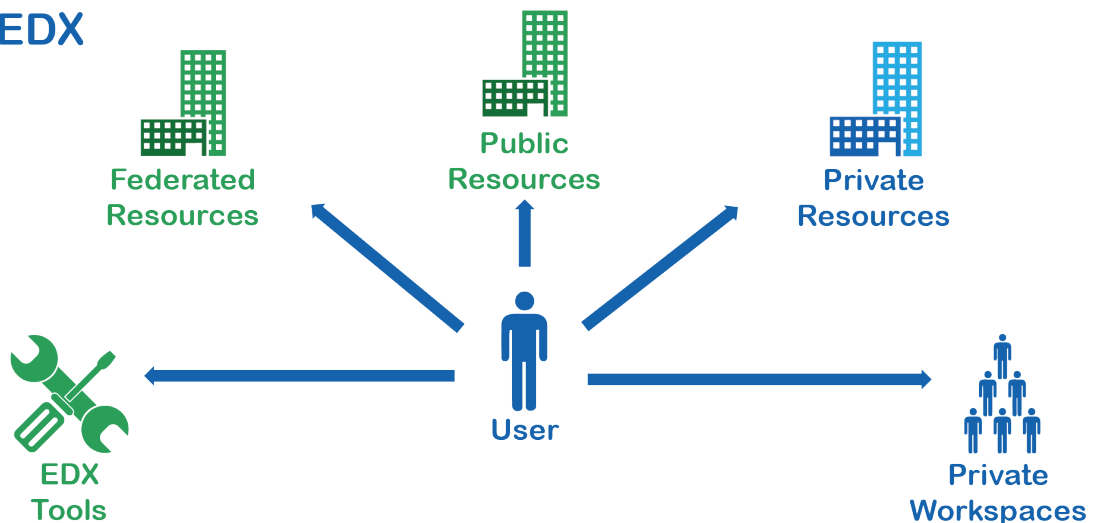
Next Steps (EDX Version 3.0) includes:

- API communication to push/pull data and information.

- App-like store view for tools.

- Calendar/Event Notification upgrades.

- Enhanced Search Capabilities (search within data itself).

- Integrated data capabilities with HADOOP.

- Flexible and customizable user interface.

- Forums upgrade.

- Virtual On-Line Laboratory.

- Implement a new NETL Point of Contact (POC) review process.

Registered users interested in options for specialty datasets and customized solutions should contact EDX Support at EDXsupport@netl.doe.gov

Connect with EDX

EDX is a virtual laboratory that allows users to connect with data, resources, and tools publicly, privately, and across multiple organizations.

Federated Resources — Public Resources — Private Resources — User — EDX Tools — Private Workspaces