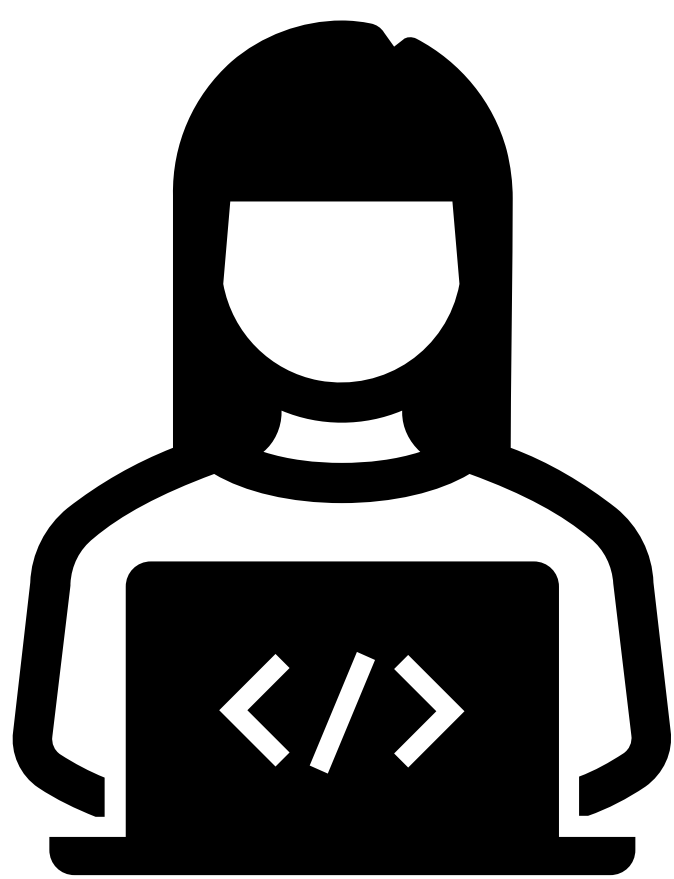
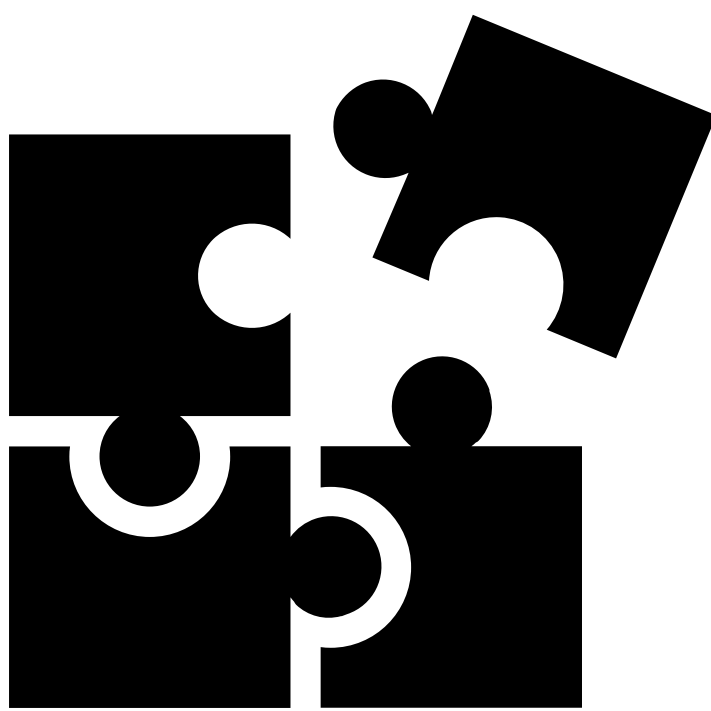
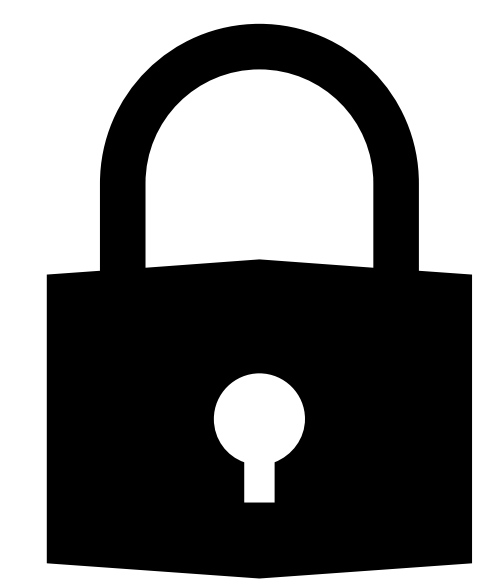





Towards a DOE Data Catalog: Ensuring Access, Sharing, and Protection

Alex May, Olga Kuchar, Katie Knight, Rohit Srivastava

Data catalogs, collections of curated records of metadata about datasets, help find, inventory, and describe use patterns for vastly distributed and diverse datasets, and assist in long-term investment decisions for organizations. Knowing the frequency of access as well as what factors contribute to dataset discovery is key to informing appropriate preservation actions.



	Requirements	Use Cases	Item Level Example
	Access to open datasets distributed throughout the laboratory	Having a centralized place to search for the descriptive metadata as well as scanned data will increase discovery of assets.	I am training an AI algorithm on moderating temperature inside a computer. What data is available that I can use?
	Protect sensitive datasets through machine-actionable access controls	The adoption of machine-actionable access controls that can facilitate access to sensitive data by aligning them with data governance and policies. It is an important field for users so they can quickly identify the barriers to accessing a dataset.	How do I get access to sensitive data?
	Curation: Standards and Schemas	It is essential to implement an extensible metadata schema that maps to established domain-specific metadata standards.	Using an established metadata schema makes it more likely that open datasets will be harvested, and it is more likely to provide users access to related datasets and enable additional discovery (Sheridan H, 2021).

Description

A good description typically includes specific context on the creation of the dataset, including:

- Frequency of data collection
- Sampling information
- Intent of the creation of the dataset

Subjects

- [Telemetry](#)
- [Titan \(Supercomputer\)](#)

File Level Scan

- Temperature: 55c
- Timestamp: 2016-04-11

Access

Restrictions

Restricted
Application Required
OCLF users only

Instructions

The dataset is available by applying directly to the author. Please contact the local expert for further instructions.

Application Required

Title

Telemetry from Summit, 2020

Author(s) [Person, D](#)
Associated Author(s) [Datanerd, A.](#)

Behind the Scenes Metadata Mappings

OAI-PMH ...
<dcterms:title>Telemetry from Summit, 2020</dcterms:title>
<dcterms:creator>Person, D.</dcterms:creator>

- Key take-aways for a DOE Data Catalog:**
- NOT a repository, but a digital way-finder to help researchers find dispersed datasets across the DOE complex
 - Supported by an integrated team of scientists, privacy officer, legal officer, developers, and librarians.
 - Describes datasets and lists detailed information on where and how to access the data.
- Action Plan:**
- Develop machine-actionable data policies and regulations.
 - Research and design a flexible metadata schema.
 - Need machine-readable (not just human-readable) repositories and develop API standards.