National Security Data Solution: Unsiloeing Weapons Research Data

Vanessa Feagin; Julie Maze; Michael Ham, PhD; Scott Doebling, PhD | Weapons Research Services Division, LANL

Abstract
The Weapons Research Services (WRS) Division at Los Alamos National Laboratory (LANL) has embraced the complex challenges of enacting Findability, Accessibility, Interoperability, and Reuse (FAIR) for the Laboratory’s weapons data assets by creating the National Security Data Solution (NSDS). The NSDS supports the laboratory’s ability to rapidly respond to national security needs by creating a space for experimentalists and analysts to securely store, find, recover, or access data relevant to their work with integrity and pedigree. Finding and storing data efficiently is not just desirable, it is mission critical. To this end, the NSDS serves as both a repository for research data and a registry of additional research data sources. The NSDS project was initiated in 2019 and as of April 2022 has begun its first real-time customer use case. Here, we discuss the goals and challenges of establishing the NSDS as we work toward taking this initiative live into both the unclassified and classified domains.

“Finding and storing data efficiently is not just desirable, it is mission critical.”

We’ve Got a Data Problem:
Huge quantities of data are generated on a consistent basis at the LANL complex but we lack formal lab-wide data management, creating barriers & siloing information

“When Jim retired, his data was lost to us. We’ve spent untold dollars replicating his results.”

Data in silos means:
• Cumbersome access
• Inconsistent metadata
• Lack of pedigree and difficulty in verifying source of truth

Resulting in wasted effort and excess cost

NSDSolution
NSDS is the adept coordination of Los Alamos National Laboratory’s nuclear science data, including the collection, integration, storage and access.

NSDS includes:
- Creation of a central and searchable inventory of various data sources.
- Connection to existing data repositories and a professionally managed repository for data owners with this need.
- Integration of data to provide meaningful search results, analysis and interpretation.
- Establishment of a data governance board to oversee controlled data access.
- Delineation of data standards, where applicable.

Roadmap
2019-2021
- Tool evaluation
- Infrastructure set-up & pilot
- Unclassified proof-of-concept
- Establish Data Governance Council

2022
- Implement first customer-use cases
- Go live in our unclassified instance
- Expand & Connect more data sources
- Cont. classified data repository proof & development

2023
- Cont. expansion of more data sources in the unclassified instance
- Stand-up production instance for classified repository
- Evaluation of semantic & machine learning

Fig 1. Advanced Search and expanded facets
Fig 2. Customizable Metadata Fields