

Andres S. Lopez Ramirez, Ujjwol Tamrakar, Will McNamara, Tu A. Nguyen
Sandia National Laboratories

Project start and finish: FY21-FY26

- **Project Page-** Users can navigate, search, and download a particular project or a set of projects
 - 2300 projects → >90 GW of energy storage projects worldwide
- **Policies Page-** Relevant federal and state policy documents
- **Statistics Page-** Interactive statistics and visualizations

- Automated pipeline systematically extracts structured data from web articles using large language models (LLMs).
- Validate against established data schema requirements.

- Final stage performs data matching and merging operations.
- Identifies similar records that already exist in the GESDB.
- Merges complementary information to create comprehensive records.
- Provides functionality to export validated records.

- Developed an LLM-based automated system that enables flexible data extraction, validation and integration through an integrated web interface.
- Combination of automation and human-in-the-loop quality control addresses challenges in maintaining a comprehensive GESDB.
- Database serves as a useful resource for various stakeholders to accelerate technological advancement and deployment of energy storage systems.

- Deploy web-scraping framework and complete integration with GESDB.
- Continuously maintain the website with periodic updates.

SAND2026-23870C