

CASE STUDY

Reducing Research Complexity with RAG-Grounded Workflows

(DEPARTMENT OF ENERGY NATIONAL LABORATORIES)

Pairing a research repository with RAG boosts speed and reliability of research discovery at the Department of Energy

Industry: Federal / Department of Energy

Use Case: Research discovery and scientific literature review

Capability: Verified, repeatable research paper discovery

Core Problem: Hallucinated papers, source verification burden, confirmation-biased search, research complexity

Deployment: CUI/Sensitive Proprietary Data

ABOUT

U. S. Department of Energy National Laboratories users conduct research across complex scientific and technical domains where source validity is critical. They use a vast online research repository, ArXiv, to discover and analyze research.

Research workflows must help users discover relevant papers quickly, verify source material, extract key findings, and determine where deeper review is required.

The volume of scientific and technical data continues to grow, making traditional research workflows harder to scale. Without AI, initial research efforts can take scientists at least a week. That pace limits how quickly teams can evaluate new information, compare approaches, and keep up with the continued influx of research.

CHALLENGE

Standard AI chat systems introduced a trust problem into research discovery. In one workflow, a chat system returned four papers with plausible titles, authors, DOIs, and links, none of which existed. Verifying the fabrications consumed roughly two hours of searching, cross-checking Scholar profiles, and re-prompting before real research could begin. Once outputs were proven false, the entire workflow lost credibility, and users stopped trusting subsequent responses, deterring adoption.

This created five core challenges: false discovery paths, manual validation burden, loss of trust after invalid outputs, added research complexity and the inability to keep pace with data growth. Traditional search also compounded the problem through confirmation bias.

80%

Reduction in task time for users to validate outputs were real

12x

Faster to verify research

SOLUTION

Legion built a dedicated research workflow paired with RAG to ground discovery in real, retrievable papers. Users search the entire repository in natural language, guided by structured and unstructured details including author, topic, document content, and the specific language within each paper. Users then interact directly with selected papers:

- **Build me a summary.** Generate an executive summary to determine whether the paper is worth deeper review.
- **Identify the key risks with the approach.** Extract limitations, assumptions, implementation risks, or gaps in the paper's methodology.
- **What are the key metrics and results?** Pull out performance measures, comparisons, findings, and relevant evidence.
- **What supports or contradicts my thesis?** Explore both confirming and disconfirming evidence across retrieved papers.

Each response is backed by citations, helping users verify sources and trust the data. This shifted AI from a speculative answer generator into a grounded workflow for discovery, validation, synthesis, and hypothesis testing.

IMPACT

Legion's RAG-based workflow significantly outperformed generalized AI for DOE research and discovery.

- **80% less follow-on research:** Less time validating outputs, more time evaluating useful inputs.
- **6 hours saved per activity:** By cutting manual source-checking, repeated prompting, and false discovery paths.
- **Verification loops cut from two hours to minutes:** What took ~2 hours in standard AI chat drops to ~10 minutes.
- **Week-long efforts compressed:** Research that took scientists a week is accelerated with multiple activities per week instead of one.
- **Scaled impact:** Initial rollout spans 25 users, expanding to 500.
- **Higher confidence, lower bias:** Users trust that cited papers are real, inspect sources, validate quickly, and test both confirming and disconfirming evidence.

“The key is having quickly verified sources, citations, and details that you can trust. Once you identify a useful paper, the ability to extract relevant information allows research activities to exponentially increase throughput.”

- Research Scientist, U.S. Department of Energy