

## **USDA | Challenge Name: Identifying Capstone Officials**

**Abstract:** This project presents an innovative approach to the automated identification and tracking of Capstone Officials within federal agencies, addressing a critical records management challenge. Our methodology employs a dual-pathway process designed to efficiently handle diverse data formats while ensuring high accuracy and compliance with federal regulations.

For unstructured data sources such as PDFs, email archives, and documents, we implement a vector database solution using Qdrant. This enables semantic search capabilities and natural language understanding to extract relevant information from complex documentation, including email signatures and meeting minutes where Acting Capstone designations may be recorded.

Concurrently, structured data sources including HR exports, organizational charts, and Excel spreadsheets are processed through a graph database architecture utilizing Neo4j. This approach maps the intricate relationships between positions, individuals, and departments, allowing for precise tracking of positional changes and hierarchical reporting structures essential to Capstone role identification.

The data extracted from both pathways is then verified through an AI agent system that conducts automated checks against official government websites and directories. This verification process confirms organizational affiliations, validates email addresses, and cross-references position information to ensure data integrity.

This hybrid solution addresses the unique challenges of Capstone Official identification, particularly the detection of temporary Acting roles not captured in formal HR systems. The automated process significantly reduces the manual effort currently required while improving accuracy and providing a comprehensive audit trail for records management compliance.