Goliath Energy: AI Certificate Storage System - Tech Scope

1. Project Overview

Project Summary Development of an AI-powered certificate and document storage system that automatically receives, processes, organizes, and provides intelligent search capabilities for Goliath Energy's equipment certifications, registrations, and operational documentation across multiple business units.

Date Created September 9, 2025, last updated September 11, 2025

2. Stakeholders

Client:

- Garrett Radchenko (CEO/President Goliath Energy)
- Reid Radchenko (Controller, Goliath Energy)
- Field staff (north and south operations)

Internal:

- Justin Adamski (Technical Lead, MyZone AI)
- Mike Schwarz (Client Liaison, MyZone AI)
- Marcos Beliera (Development Support, MyZone AI)

3. Current Situation

Process Mapping Currently, Goliath Energy receives certificates, registrations, invoices, and in labeling, identification, searching, and expiration date tracking are all manual processes. Documents are scattered across email inboxes without systematic categorization, making retrieval time-consuming and error-prone during regulatory inspections or operational needs.

Time, Effort, and Cost Currently Involved

- Manual document sorting and filing: Estimated 10-15 hours/week across staff
- Document search and retrieval: 5-15 minutes per request
- Risk of compliance issues due to missing or misplaced certificates
- Potential operational delays when equipment documentation cannot be located quickly

Pain Points

- No systematic organization of certificates and registrations by equipment unit
- Difficulty tracking expiration dates for certifications
- Time-consuming manual search through email attachments
- Risk of regulatory non-compliance due to disorganized documentation
- Inefficient handoff processes between crews and locations (north/south operations)
- Multiple email systems without centralized document management

Existing Tools & Systems

- Email system with unlimited storage for document receipt
- Manual filing and organization processes
- No current AI or automation tools for document management

4. Proposed Solution

Automation Scope

- Email ingestion from two dedicated email addresses (goliathcertreg.com and goliathpaperwork.com)
- AI-powered document classification and organization into predefined folder structures
- Intelligent document parsing and metadata extraction (equipment IDs, expiration dates, document types)
- Vector store creation for semantic search capabilities
- Chatbot interface for instant document retrieval and query responses

High-Level Workflow

1. **Document Ingestion**: Automated monitoring of two email addresses for incoming attachments
2. **AI Processing**: Document analysis, text extraction, and classification using AI models
3. **Organization**: Automatic sorting into equipment-specific folders based on extracted metadata
4. **Storage**: Secure cloud storage with unlimited capacity for files + cloud database for vector storage/fast search of imported documents
5. **Search Interface**: Chatbot application allowing natural language and specific ID searches
6. **Retrieval**: Instant access to documents with direct download links and preview capabilities

Tech Stack

- **Email Processing**: IMAP/API integration for automated email monitoring
- **AI/ML**: OpenAI GPT models for document understanding and classification
- **Database**: Supabase with pgvector for semantic and direct search capabilities
- **Automation Platform**: N8N for workflow orchestration
- **Storage**: Google Drive or AWS S3 or onedrive for document storage with unlimited scaling
- **Frontend**: n8n chatbot or MyZone chatbot interface

Data Sources

- Email attachments from goliathcertreg.com (equipment certificates and registrations)
- Email attachments from goliathpaperwork.com (operational documentation)

Automation Outputs

- Organized folder structure mirroring Goliath's equipment hierarchy
- Searchable document database with AI-powered query responses
- Expiration date alerts and compliance notifications
- Document retrieval with direct links and metadata summaries

Integration Points

- Email systems (goliathcertreg.com, goliathpaperwork.com)
- Cloud storage platforms (Google Drive/AWS S3)

Security Considerations

- Encrypted document storage and transmission
- Access control and user authentication
- Regular security audits and compliance monitoring
- Data backup and disaster recovery procedures
- GDPR/privacy compliance for document handling

Scalability

- Cloud-based architecture supporting unlimited document volume
- Horizontal scaling of processing capabilities
- Additional equipment units and document types easily accommodated
- Multi-location support (north/south operations)
- Future expansion to additional Goliath subsidiaries

Risks

- Email system integration may require API access or IMAP configuration
- AI classification accuracy dependent on document quality and consistency
- Initial training period required for optimal AI performance
- Potential learning curve for staff adoption of new search interface

Assumptions & Requirements

- Client provides timely access to email systems and sample documents
- Email forwarding or direct access can be configured for automated ingestion
- Field staff have access to internet-connected devices for system access
- Existing document archive will be provided for initial system training
- Client approval required for folder structure and classification logic

5. Project Plan

Timeline & Milestones

Phase 1: Discovery & Setup (Week 1) -

- Requirements gathering and system access configuration
- Email integration setup and testing
- Document sample analysis and AI model prep
- Initial folder structure configuration

Phase 2: Core Development (Week 2-3) -

- Email monitoring and ingestion system development
- AI document classification and processing pipeline
- Vector store creation and search functionality

Phase 3: Interface Development (Week 4) -

- Chatbot interface (n8n chatbot, or myzone chatbot)
- Search optimization and performance tuning

Phase 4: Testing & Deployment (Week 5~6) -

- System testing with real documents
- User acceptance testing with Goliath team
- Performance optimization and bug fixes
- Production deployment and monitoring setup
- Staff training sessions
- Documentation and user guides
- Final system optimization
- Go-live support and monitoring

NOTE: *This timeline does not include any periods waiting on confirmations/data access/etc from client.*

Budgets

- **Subscription costs**:
    - N8N: $30-50/month and/or MZ chatbots 49 usd/mo
    - OpenAI API: $30-50/month (depending on usage)
    - Supabase + pgvector: ~$10-30/month
    - Cloud storage: included if client already has preferred cloud drive platform

Approval Workflow Written confirmation via email suffices for approval of this plan.

6. Appendices

Reference Materials

- Email thread documentation (August 15 - September 8, 2025)
- Goliath equipment folder structure specifications
- Sample document types and formats (to be provided)
- Existing legal client system architecture (internal reference)

Equipment Categories for Reference

- **Stand-Alone Unit One (SA1)**: G07, ST01, PT03, CT03, DH04
- **Stand-Alone Unit Two (SA2)**: G06, ST02, PT02, CT01, DH02
- **Stand-Alone Unit Three (SA3)**: G03, ST03/ST04, PT01/PT04, CT02, DH03
- **Support Equipment**: Various tongs, trailers, and specialized equipment

Document Categories

- **Certificates & Registrations**: Equipment-specific compliance documents
- **Operational Paperwork**: Inspections, meetings, drills, calculations, and observations