

Capstone Project Proposal: Athenica AI Engineering & Compliance Automation Platform

Prepared by: Sunidhi Rusia, MS in BA&AI, Stevens Institute of Technology
Under the Guidance of Prof. Alkiviadis Vazacopoulos, School of Business, Stevens Institute of Technology

Project Title: End-to-End Inspection Intelligence: Developing an Agentic AI Operations Platform for Environmental Engineering

Industry Partner: Athenica Environmental Services, Inc.

Focus Areas: Agentic AI Workflows, Document Intelligence, Regulatory Compliance Automation, Enterprise UI/UX, Risk & Workflow Management

1. Executive Summary & Problem Statement

Athenica Environmental Services provides specialized environmental consulting, building condition evaluations, and regulatory compliance assessments across the New York metropolitan area. Their day-to-day operations involve high cognitive load across multiple interdependent workflows:

1. **Regulatory Ingestion:** Processing massive, constantly shifting regulatory documents, municipal codes (NYC DOB/DEP/OER), and strict ASTM protocols (e.g., ASTM E1527-21 for Phase I ESAs, ASTM E2600 for Vapor Encroachment).
2. **Field Data Synthesis:** Converting fragmented inputs—not just PDFs, but field reconnaissance notes, interviews, user questionnaires, and historical database reports (EDR/ERIS)—into defensible engineering reports.
3. **Lab & Regulatory Verification:** Auditing laboratory testing data against regulatory safety thresholds to confirm compliance flags.
4. **Workflow & Liability Management:** Tracking project viability (e.g., the 180-day/1-year viability clock for ESAs), managing user roles, and documenting the Environmental Professional (EP) review process for CERCLA liability protections.

Operating as an engineering-focused firm, Athenica relies on manual data transcription, document comparison, and formatting. This creates operational friction and increases administrative turnaround. This capstone project engineers an enterprise-grade AI operations suite powered by Claude's advanced document processing, vision capabilities, and structured reasoning. The platform will sit *inside* the project workflow—providing automated regulatory digestion, standardized report generation, automated QA/QC, and robust audit trails for defensibility.

Operational Challenges & AI System Response

The table below maps Athenica's operational challenges directly to the proposed technological solutions:

Current Manual Workflow Challenge	Target Problem & Risks	Platform Response / Module Features
High Administrative Overhead in Drafting	No automated mechanism to turn verified findings into final deliverables, creating huge backlog in drafting standard sections, narrative, recommendations, and declarations.	Automated Multi-Stage Report Drafting Studio Drafts full reports using custom template parsing & inserts signature/EP blocks automatically.
Lack of ASTM E1527-21 Granularity	Found data lacks standard classifications (REC, CREC, HREC, de minimis) and fails to identify proximity (subject, adjoining, nearby) without tedious manual review.	Standardized Classification Layer Applies explicit ASTM classifications & geographic proximity tags based on evidence-backed criteria. Includes rationale fields.
Undocumented "Data Gaps"	Traditional extraction methods catch what's there but ignore what's missing, leaving liability gaps (missing intervals >5 yrs, missing records).	Continuous Checklist Coverage Tracker Flags what isn't found, tracking historical data intervals, missing

		interviews, or inaccessible areas.
Siloed Multi-Source Analysis	Relying on page-by-page queries obscures systemic issues (e.g., connecting a '60s city directory, a map, and database hits to identify a dry cleaner).	Cross-Document Narrative Chronology A secondary, global-context processing pass to identify connections and timeline trends.
Uncaptured High-Risk Inputs	Valuable inspection inputs (handwritten notes, site interviews, visual maps) are siloed in paper files and cannot be easily parsed.	Multi-Modal Data Integration Supports direct uploads of notes, occupant questionnaires, and database files.
Vision & Image blindspots (Maps & Photos)	Historical maps (Sanborn maps, topo maps) are missed by OCR or processed as noise.	Computer-Vision Extraction Integrates visual spatial extraction from maps/photos alongside

		specialized parsing of large lab tables.
Regulatory & Liability Vulnerabilities	Overreliance on un-audited AI might fail to catch false positives/negations. Regulatory compliance must hold up to decades of litigation.	Robust QA Review Workspace & Ledger Captures the exact model version, rulebook, and prompt. Measures recall of historical Approved Reports.

2. Comprehensive System Capabilities (What We Will Do)

The platform addresses the full lifecycle of an inspection engagement through five integrated modules:

Module 1: Multi-Modal Ingestion & Cross-Document Reasoning Engine

- **Broad Input Handling:** Ingests unstructured notes, User Questionnaires, historical database reports (EDR/ERIS), and site interview transcripts, not just standard PDF reports.
- **Vision-Native Processing:** Processes Sanborn maps, aerial photographs, and topographic maps directly through Claude's vision capabilities (handling images as images, returning region-specific evidence instead of failing via standard OCR).
- **Intelligent Chunking:** Groups text with special logic to prevent breaking lab data tables that span multiple pages.
- **Cross-Document Synthesis:** Executes a multi-pass reasoning step (a second Claude pass) to connect related sources—e.g., linking a 1960s city directory, a Sanborn map, and groundwater flow data to build an accurate property-use chronology.

Module 2: Automated Multi-Stage Report Drafting Studio

- **ASTM E1527-21 Classification Layer:** Automatically categorizes findings as RECs, CRECs, HRECs, de minimis conditions, or "not a concern," specifying whether they are on the subject property, adjoining property, or nearby. Includes a required "rationale" field for EP review.
- **Scope Separation:** Carefully isolates different regulatory tracks (e.g., ASTM E1527-21 vs. NYC OER E-designations) to ensure the right standards are applied.

- **Deliverable Assembly:** Transforms approved findings into the final deliverable, compiling the standard report sections, findings/opinions narrative, recommendations, the EP declaration/signature, and appendices via a strict report template.

Module 3: Data Gap Tracking & Lab QA/QC

- **Checklist Coverage Tracker:** Actively tracks *what is missing*, identifying significant data gaps (e.g., missing historical intervals >5 years, inaccessible areas) and marking required elements as "found," "not found," or "data gap."
- **Regulatory Thresholds:** Cross-references detected chemical/contaminant levels against EPA and NYC regulatory thresholds, processing lab tables appropriately rather than treating them as standard prose.

Module 4: Workflow, Liability & Lifecycle Management (The Orchestration Layer)

- **Workflow Integration:** Manages sites, projects, user roles (Analyst, QA, EP), assignments, and tracks the 180-day/one-year viability clock so the tool integrates directly into Athenica's real-world operations.
- **Audit Trail & Defensibility:** Records the exact model version, rulebook version, and prompt behind each finding, ensuring reports can be defended years later for CERCLA landowner liability protections.
- **EP Review Tracking:** Provides a clear system for documenting how and when the Environmental Professional reviewed and approved the AI-generated findings.

Module 5: Regulatory Q&A & Compliance Copilot

- Enables engineers to ask natural-language questions regarding specific regulatory triggers (e.g., vapor migration protocols under ASTM E2600) and receive citation-backed answers.

Technical Stack Additions:

- **Language Models & Reasoning:** Claude API (leveraging vision inputs for maps/images and extended context windows for multi-document synthesis).
- **Data & Compliance Governance:** Implementation of retention/deletion rules for client files, plus cost-tracking architecture (calculating AI inference cost estimates per 1,000-page job).
- **Audit & Liability Storage:** Database schemas designed to lock and log prompt/model versioning alongside EP approval timestamps.

5. Phased Implementation Roadmap

The deployment of the Athenica platform spans 6 key development phases, which includes our added focus on recall-accuracy, project lifecycle tracks, and rigorous boundary-checks:

Phase	Focus area	Activities & Technical Objectives	Concrete Deliverables
Phase 1	Domain Discovery & Schema Design	Collaborate with Athenica SMEs to standardize schemas for EPA/ASTM formats; map E1527-21 categories, NYC OER requirements, visual map structures, and the workflow timeline.	Structured JSON mapping specs, ASTM E1527-21 logic templates, data-gap rules, and liability workflow models.
Phase 2	Multi-Modal Parsing Pipeline	Assemble data ingestion pipeline handling raw text, PDF, databases, hand-written reconnaissance, and visual items (Maps/Imagery) with smart table stitching.	Hybrid multi-modal ingestion system, OCR parser, visual map analysis tool, and database importer.
Phase 3	Extraction, Synthesis & Rationale Pipeline	Develop first-pass extraction models (identifying site details/findings) and a second-pass chronological synthesis agent, adding REC/CREC categorization and data-gap analysis.	Synthesis engine, ASTM categorizer with rationale generation, and automated checklist-gap tracker.
Phase 4	Lifecycle Management, QA, & Audit Ledger	Integrate 180-day viability timers, role assignments, cost estimation calculators (\$/1000 pgs), and the audit logger capturing prompts/model versions.	Workflow tracking backend, role assignment engine, audit log generator, and retention & deletion tool.

Phase 5	Inspection Studio Web Frontend	Integrate API, parsing, synthesis, and logging pipelines into a zero-learning-curve web UI designed for easy data-checks, review, and signing.	Integrated UI Web App featuring visual review tools, editable audit sections, task boards, and DOCX/PDF export.
Phase 6	Rigorous Evaluation, Recall Testing & Transfer	Run testing beyond quote checking, comparing output against historical approved reports to measure false-negatives. Deliver codebases & training.	Benchmark evaluation, system training, SLA manuals, deployment documentation, and project handover.