AMIE – Product Requirements Document (PRD)
System: Academic Manuscript Intelligence Evaluator (AMIE)
Version: Minimum Viable Product (MVP)
Type: Modular Agentic SaaS Platform
Prepared For: AMIE Product, Engineering, and IP Stakeholder Teams

---

1. Purpose

AMIE is an intelligent agentic system designed to support innovation strategy within the academic research ecosystem. It analyzes academic manuscripts to detect signals of potentially patentable subject matter — either directly disclosed or implicitly suggested — and generates structured intelligence reports.

The MVP is intended to demonstrate that AMIE can:

- Surface invention signals before publication.
- Integrate with both automated and user-initiated workflows.
- Deliver timely and readable IP assessments.
- Avoid over-alerting while capturing high-recall signals.

---

2. Product Scope

AMIE delivers:

- An end-to-end, multi-agent pipeline to assess patentability in manuscripts.
- Two engagement modes: (1) faculty-initiated upload via chat, and (2) scheduled autonomous scraping of public repositories.
- Reports designed for both non-specialist researchers and IP professionals.
- Lightweight, explainable outputs grounded in reproducible prompts and structured agent interactions.

The system must be auditable, prompt-based, schema-driven, and agentically orchestrated. All features must live within the MVP envelope and avoid premature inclusion of commercial value scoring or patent filing tools.

---

3. Use Cases

3.1 On-Demand Faculty-Initiated Review

- **User**: OSU faculty member preparing to submit a manuscript.
- **Trigger**: Faculty uploads a manuscript via a secure chat interface (UIA).
- **Process**:
    - Manuscript is passed to the Orchestration Agent (OA).
    - The OA spawns a Planner (APA), then Classifier, Patentability, and Prior Art agents.
    - The IA and SA produce a human-readable report.
- **Output**: A clear, actionable PDF or Markdown file recommending next steps.

3.2 Automated Manuscript Monitoring

- **User**: University IP manager or tech transfer analyst.
- **Trigger**: A new OSU-authored manuscript appears on a monitored preprint server (e.g., arXiv).
- **Process**:
    - A scheduled Scraper + System Monitor Agent (SMA) detects the manuscript.
    - The OA is triggered and runs the same assessment pipeline as the on-demand mode.
    - A structured report is compiled and emailed.
- **Output**: Internal patentability analysis report and alert notification.

---

4. Functional Requirements

- Manuscript ingestion must accept PDF, DOCX, or plain text.
- Metadata extraction must include title, authors, affiliations, and timestamps.
- System must support file submission via a chat UI and automated retrieval via scraping.

- Orchestration Agent (OA) must launch downstream agents with a traceable execution plan.
- Assessment Planner Agent (APA) must output a plan.yml file:
    - Specifies manuscript sections to analyze.
    - Defines which legal heuristics (§101, §102, §103) to apply.
    - Provides query templates for prior art search.
- Classifier AA must return a "likely IP-bearing?" score for pipeline short-circuiting.
- Patentability AA must:
    - Evaluate eligibility, novelty, obviousness, and enablement.
    - Use prompt fragments for legal analysis.
- Prior Art AA must:
    - Query patent corpus and non-patent literature.
    - Return top 5 patent + 5 publication hits with similarity scores (cosine > 0.78).
- Integration Agent (IA) must merge outputs into assessment.json:
    - Includes flags, scores, key snippets, and top prior art.
- Synthesis Agent (SA) must generate a Markdown and optional LaTeX/PDF report.
- Event logging must capture:
    - All decisions (flag/no flag), confidence scores, section coverage, plan length.
- System must poll IP dockets monthly for disclosures linked to manuscript IDs.
- 5% of unflagged manuscripts must be manually reviewed and added to training corpora.

---

5. Non-Functional Requirements

- On-demand reports must be delivered within 5 minutes.
- Automation mode must support daily scraping and batch processing.
- False positives should remain below 30%, validated by blind-spot auditing.
- Recall on known invention-linked manuscripts must be ≥80%.

- The system must follow a modular, containerized deployment model.
- Output schemas (assessment.json) must support version control.
- Manuscript content must not be stored beyond processing window without user opt-in.
- Observability must include trace IDs, timeout logs, retry counts, and output hooks.

---

6. System Constraints

- **Timeline**: MVP implementation must complete within 12–14 weeks of greenlight.
- **Platform**: Deployed on OSU infrastructure (or hybrid AWS/cloud).
- **LLM Compatibility**: Compatible with GPT-4o and Claude Sonnet.
- **Security & Privacy**: Compliant with OSU research data handling policies.

---

7. Agent Roles

- **UIA (User Interaction Agent)**: Chat interface for on-demand faculty submissions.
- **SMA (System Monitor Agent)**: Watches preprint sources and triggers ingestion events.
- **OA (Orchestration Agent)**: Core controller; coordinates pipeline execution, timeouts, and trace IDs.
- **APA (Assessment Planner Agent)**: Generates a manuscript-specific interrogation strategy.
- **Classifier AA**: Early-stage filter to avoid costly downstream processing if unnecessary.
- **Patentability AA**: Legal reasoning across §101, §102, §103, and enablement.
- **Prior-Art AA**: Search and retrieval agent for similarity-ranked patent and non-patent literature.
- **Strategic Insight AA** *(optional in MVP)*: Annotates market context and funding implications.
- **IA (Integration Agent)**: Merges structured findings into a consistent payload.

- **SA (Synthesis Agent)**: Generates faculty-facing and internal reports.

---

8. Inputs and Outputs

Inputs

- Manuscript file (PDF, DOCX, TXT)
- Associated metadata (title, authors, affiliations, date)
- Source: Chat submission or arXiv scrape

Outputs

- assessment.json including:
    - eligibility_flag
    - novelty_score
    - obviousness_risk
    - enablement_confidence
    - key_snippets
    - top_prior_art references
- Human-readable Markdown or PDF report
- Digest emails (automation mode)
- Event logs and validation hooks

---

9. MVP Scope Checklist

Included in MVP

- Chat-based file upload and return (UIA)
- Agentic assessment pipeline with traceability
- Plan.yml generation and dynamic prompt orchestration
- Classification, legal assessment, and prior art retrieval
- JSON + Markdown/PDF output

- Cron-based scraping and ingestion system
- IP manager email reporting system
- Dashboard and validation metrics system

Excluded from MVP

- Claim drafting agent or provisional patent generation
- Full §112 compliance agent
- Market scoring or licensing value prediction
- Mobile or offline interface
- Patent filing or docketing integration
- Public workspace sharing or citation analysis

---

10. Success Criteria

- ≥80% recall on known positive (linked disclosure) manuscripts
- ≤30% false positive rate (validated via blind-spot sampling)
- Average on-demand report turnaround: ≤5 minutes
- Faculty satisfaction (post-submission survey): ≥4.0/5
- IP manager report open rate: ≥90%
- Pilot uptime during operation: ≥99.5%