# AMIE MVP – Updated Agentic System Design & Validation Blueprint

## 1. High-Level Intent

Prove that an agentic workflow can surface patent-worthy signals (disclosed or implied) early enough for the university to preserve patent rights.

**Success criteria:**

* AMIE catches most real invention opportunities without flooding staff with false alarms.
* Every flagged manuscript can be traced to a formal tech-transfer case file.

## 2. Core Data Flow

[Front-End (UI)]

│

▼

Ingestion Agent

│

▼

Invention Detection Agent (IDA)

│

▼

Classification Agent (CA)

│

▼

Novelty Assessment Agent (NAA)

│

▼

Aggregation Agent (AA)

│

▼

[Front-End (UI)]

* **Automation mode:** The Ingestion Agent crawls sources (e.g. arXiv, university submission systems) on a schedule and feeds manuscripts into the pipeline.
* **On-demand mode:** Users upload manuscripts manually via the front end for analysis.

## 3. Agent Specifications

### Front-End

* Allows users to upload manuscripts
* Displays:
  + Workflow progress and status
  + JSON or human-readable reports

### Ingestion Agent

* Fetches or receives manuscripts
* Extracts:
  + Title
  + Authors
  + Affiliations
  + Timestamp
* Attaches a provenance ID
* Designed as a lightweight Python worker with retry logic and dead-letter queue
* Stores manuscript in GCP/AWS/Azure

### Invention Detection Agent (IDA)

* Determines if the manuscript:
  + Presents an invention
  + Implies an invention
  + Contains no invention
* Detection criteria:
  + Explicit new compositions, devices, processes, algorithms, systems
  + Language suggesting novelty or improvements
  + Experimental data implying novel technology
* **Output (JSON):**
  + result (present / implied / absent)
  + reasoning
  + technical\_summary (if applicable)

### Classification Agent (CA)

* Classifies the invention into relevant technical fields
* If invention is only implied, halts further processing and returns the fields
* **Output (JSON):**
  + fields (list of technical fields)
  + reasoning

### Novelty Assessment Agent (NAA)

* Assesses novelty under 35 U.S.C. § 102
* Steps:
  + Identify key technical [structure] or features
  + Search literature, patents, public data sources
  + Compare invention’s structure vs. prior art:
    - Identical match → novelty fails
    - Partial overlap → triggers NOA
    - No match → novelty likely secured
* **Output (JSON):**
  + novelty
  + partial\_matches (if any)
  + reasoning

### Non-Obviousness Agent (NOA)

* Evaluates non-obviousness under 35 U.S.C. § 103
* Analyzes:
  + Differences between the invention and partial matches
  + Whether differences are routine variations
  + Whether there’s motivation to combine prior art
  + Evidence of unexpected results or advantages
* **Output (JSON):**
  + non\_obviousness
  + reasoning

### Statutory Bars & Loss of Rights Agent (SBLA)

* Identifies any disclosures that might prevent patent protection under U.S. or foreign law
* Reviews:
  + Prior publications
  + Conference talks, posters, or public disclosures
  + Online preprints
* Evaluates:
  + Timing relative to U.S. 1-year grace period
  + Foreign absolute novelty implications
* **Output (JSON):**
  + statutory\_bar\_risks
  + details of disclosures
  + reasoning

### Inventorship & Ownership Agent (IOA)

* Determines inventorship and potential ownership issues
* Analyzes:
  + Authors’ contributions
  + Institutional IP policies
  + Funding agreements (e.g. Bayh-Dole compliance)
* Provides:
  + Inventor list and roles
  + Ownership issues
  + Recommendations
* **Output (JSON):**
  + inventors
  + ownership\_issues
  + ownership\_details
  + recommendations

### Aggregation Agent (AA)

* Gathers and harmonizes all data from prior agents
* Ensures:
  + Consistent JSON structure for front-end or API consumers
  + Clear explanation of where the pipeline stopped and why
  + Complete report of patentability status, legal risks, inventorship, and ownership
* Fills in null for any skipped stages
* **Output (JSON):**
  + status (absent / implied / present)
  + termination\_reason
  + technical\_summary
  + fields
  + novelty
  + novelty\_details
  + non\_obviousness
  + non\_obviousness\_details
  + statutory\_bar\_risks
  + statutory\_bar\_details
  + statutory\_bar\_reasoning
  + inventors
  + ownership\_issues
  + ownership\_details
  + ownership\_recommendations

## 4. Front-End Display

* Displays the final results from whichever agent stops the pipeline:
  + “No invention detected.”
  + “Invention implied but not fully disclosed.”
  + “Novelty fails due to prior art.”
  + “Potentially patentable invention found.”
  + “Statutory bar risks detected.”
* Reports presented as:
  + JSON (for developers)
  + Human-readable Markdown or HTML

## 5. Long-Term Validation Metrics

* Maintain a ground-truth corpus:
  + Positive set: manuscripts linked to invention disclosures or patents
  + Negative set: manuscripts dormant for ≥18 months
* Log decisions from each agent
* Poll tech-transfer systems monthly for new disclosures
* Compute:
  + Precision
  + Recall
  + F1-score
  + Time-to-first-alert
* Random “blind-spot” sampling of ~5% of unflagged manuscripts for human IP analyst review