ANALYSIS OF PRIOR ART FOR A MULTI-AGENT AI SYSTEM FOR QUESTION DISCOVERY AND PRIORITIZATION

DOCUMENT GENERATION INSTRUCTIONS: Please format the document to make ready to go public and made into a pdf.

TARGET DOCUMENT TOTAL PAGE COUNT: 3

TARGET DOCUMENT TOTAL WORD COUNT: 750

TARGET DOCUMENT LANGUAGE: ENGLISH

META INSTRUCTIONS: GENERATE THE DOCUMENT IN HTML FORMAT AND RESPOND WITH ONLY THE HTML CONTENT. YOU CAN ALSO USE THE `web_search_tool` TOOL TO GATHER MORE INFORMATION AND CONTEXT AND IMAGES FOR THE DOCUMENT BEFORE RESPONDING WITH THE HTML CONTENT. IF YOU NEED MORE INFORMATION - DON'T ASK THE USER FOR MORE INFORMATION, JUST USE THE TOOLS TO GATHER MORE INFORMATION OR MAKE ASSUMPTIONS ABOUT THE DOCUMENT.

Conceptual representation of a multi-agent AI system architecture.

I. EXECUTIVE SUMMARY: ASSESSMENT OF PATENTABILITY RISKS AND OPPORTUNITIES

A. OVERVIEW OF FINDINGS

This report provides a comprehensive prior art analysis for the invention disclosed in the document titled "Multi-Agent Artificial Intelligence System for Discovery, Analysis, Governance, and Pareto-Prioritisation of Novel, High-Impact Questions". The analysis covers issued patents, published applications, and non-patent literature (NPL) to assess the novelty and non-obviousness of the claimed invention.

The central finding is that while the individual technological pillars of the invention—multi-agent architectures, automated question generation, AI-driven governance, immutable ledgers, and Pareto optimization—are

individually well-established concepts in the prior art, the specific, integrated five-component architecture designed for the proactive discovery, ethical governance, and strategic prioritization of "unknown-unknown" questions appears to possess patentable novelty. No single reference was found that anticipates the complete system as claimed.

However, the invention faces a significant patentability risk under 35 U.S.C. § 103 (obviousness). The primary challenge will be to overcome the argument that a Person of Ordinary Skill in the Art (POSITA) would have been motivated to combine these known elements to achieve the invention's stated objectives. The patentability of the invention will therefore hinge on demonstrating a synergistic effect and a non-obvious purpose arising from this specific combination.

B. SUMMARY OF KEY RISKS

The most substantial risk to patentability is an obviousness rejection. A patent examiner is highly likely to assert that combining known technologies from disparate but related fields would have been obvious to a POSITA. Specifically:

- **Combination of Generation and Governance:** The art contains systems for automated hypothesis generation and separate discussions on AI governance and content moderation, including the use of a "quarantine" mechanism for harmful content. An examiner could argue that applying known governance techniques to the output of a known generation engine is a predictable and logical step to ensure safety and responsibility.
- **Combination of AI Output and Auditing:** The use of blockchain or other distributed ledger technologies (DLT) to create immutable audit trails for AI decisions is an emerging but documented field of research. An examiner could argue that it would be obvious to apply this known auditing technique to the outputs and processes of the question generation and governance system to ensure transparency and compliance.
- **Combination of Scoring and Ranking:** The use of multi-objective optimization, particularly generating a Pareto frontier, is a standard method for ranking items with multiple competing attributes, as seen in the optimization of AI pipelines and recommender systems. An examiner could argue that once a system generates questions with multiple scores (as produced by the QAA), applying a standard Pareto analysis for ranking is an obvious design choice.

The use of the specific term "quarantine" in the claims represents a notable vulnerability, as this term has a well-defined meaning in the context of public content moderation, which differs from the invention's internal, foresight-oriented application.

C. SUMMARY OF KEY DISTINCTIONS (OPPORTUNITIES)

Despite the risks, the invention possesses several strong distinguishing features that can form the basis of a robust argument for non-obviousness. The core opportunity lies in emphasizing the unique synthesis and purpose of the claimed combination, which is not merely an aggregation of parts but a new type of strategic foresight engine. Key distinctions include:

1. **Novel Purpose and Function:** The system's primary function is the proactive generation of novel, "unknown-unknown" questions to identify strategic blind spots. This contrasts sharply with the bulk of prior art focused on answering known user queries or analyzing existing documents.
2. **Unique Governance Workflow:** The "quarantine" mechanism is not a reactive tool for moderating public content. It is an internal, pre-emptive governance workflow for managing R&D and strategic risk. It is triggered by a specifically computed, multi-dimensional ethical-risk score exceeding a policy threshold, a process far more specific than the general content moderation systems described in NPL. The inclusion of a "quarantine sub-ledger" and a "dual human sign-off" requirement adds further patentable weight.
3. **Specific Application of Immutable Ledger:** The "Question Ledger" creates an immutable audit trail for the entire lifecycle of a novel intellectual asset—a question. It records not just the question itself, but its analytical score vectors, the agent rationales, and the full history of its governance adjudication. This is a more specific and sophisticated application than prior art, which focuses on logging the technical provenance of a single AI inference or tracing data flow.
4. **Novel Application of Pareto Optimization:** The "Priority Engine" applies Pareto optimization to a novel set of abstract, strategic, and ethical objectives (e.g., novelty, impact, feasibility, ethical risk). This differs from prior art that applies the same technique to concrete system performance metrics like accuracy, latency, or cost. The purpose is not to optimize the AI system's configuration but to present a frontier of ideas for human strategic exploration.

Conceptual framework for an innovation strategy.

D. HIGH-LEVEL RISK ASSESSMENT MATRIX

Component	Novelty Risk (Anticipation §102)	Obviousness Risk (§103)	Key Distinguishing Feature(s)	Recommendation
Question Discovery Agents (QDA)	Yellow (General concept of automated question/ hypothesis generation is known)	Red (High risk of being found obvious over art like SparkBeyond or Google's patent)	Proactive generation of "unknown-unknowns" across any domain; specific use of anomaly detection to find "conceptual white-space."	Emphasize the unique combination of anomaly detection and generative models for surfacing questions not directly derivable from text.
Question Analysis Agents (QAA)	Green (The specific 5-dimensional score vector is likely novel)	Yellow (Scoring AI outputs on various metrics is common; risk of obviousness to combine different scoring criteria)	The specific combination of novelty, impact, feasibility, ethical-risk, and cross-domain leverage as a unified vector.	Clearly define and argue for the non-obviousness of this specific vector as a tool for strategic, rather than performance, evaluation.
Question Governance Agents (QGA)	Yellow (The term "quarantine" is known, but the specific workflow is likely novel)	Red (High risk of examiner combining AI governance principles with content moderation "quarantine" art)	Internal R&D foresight purpose; triggered by a computed ethical-risk score vs. policy; auditable quarantine sub-ledger with dual sign-off.	Amend claims to recite the specific triggers and purpose of the quarantine workflow to distinguish it from public content moderation.
Question Ledger	Yellow (Blockchain for AI audit trails is a known concept)	Red (High risk of being found obvious to apply known DLT auditing to the output of an AI system)	Records the entire lifecycle of an abstract asset (a question), including score vectors, agent rationales, and governance actions.	Argue that the ledger creates a new type of auditable asset for innovation governance, not just technical provenance.

Component	Novelty Risk (Anticipation §102)	Obviousness Risk (§103)	Key Distinguishing Feature(s)	Recommendation
Priority Engine	Yellow (Pareto optimization is a standard algorithm)	Red (High risk of being found obvious to apply a standard ranking method to a list of scored items)	Application of Pareto optimization to a novel set of abstract, strategic, and ethical objectives, rather than concrete system performance metrics.	Focus arguments on the novel problem being solved: prioritizing abstract strategic inquiries for human decision-making, not optimizing a machine's performance.