From Answering to Knowing When to Stop:

Mitani Sangyo Filed U.S. Patent Application for AI Reliability Governance Framework

Kanazawa, Japan – April 16, 2026 – Mitani Sangyo Co., Ltd., (hereinafter Mitani Sangyo) a diversified trading and manufacturing company founded in Ishikawa Prefecture, Japan, in 1928, today announced the development of an “AI Reliability Governance Framework” designed to govern the entire AI response process from Input to Output. Mitani Sangyo has also filed a U.S. provisional patent application for seven element technologies that constitute this framework.

The framework focuses on ensuring verifiable information sources and enabling AI systems to stop generating answers when uncertainty is high.

This initiative enhances the reliability of AI-generated answers by enabling systems to provide evidence-based responses and to refrain from answering when uncertainty is high—addressing persistent challenges in the business use of rapidly advancing generative AI. Mitani Sangyo is building intellectual property to contribute to global discussions on AI safety, positioning this framework as an AI governance technology originating in Japan.

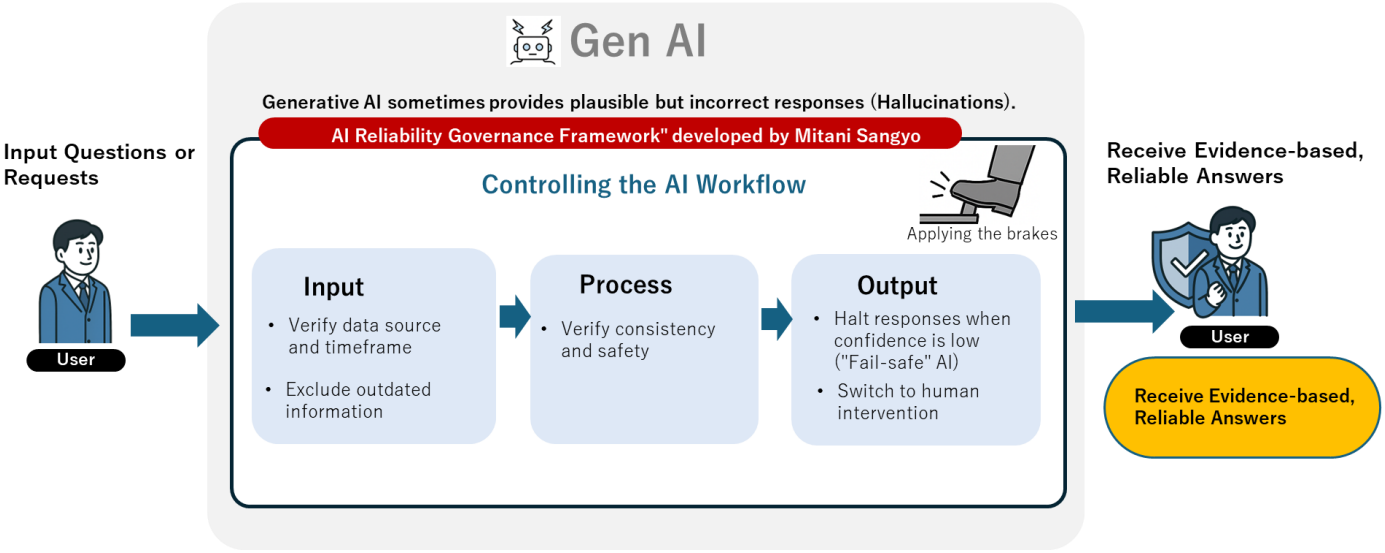
Why it matters

Generative AI has advanced rapidly through reasoning models improved with advances in inference models, retrieval-augmented generation (RAG), and AI agents. However, even in these systems, the risk of plausible but incorrect responses—so-called hallucinations—remains a challenge, particularly in mission-critical business settings.

When utilizing AI for critical business tasks, it is essential not only to require correct answers, but also to ensure verifiability and reliability. For example, being able to confirm the basis of information and stopping answers in case of uncertainty. In sectors requiring exceptional accuracy and explainability, such as finance, legal, manufacturing, and public services, preventing incorrect outputs takes precedence over conversational fluency. This necessitates explainable capabilities, such as verifying the data provenance and timeframe (freshness) of information, and fail-safe mechanisms that detect uncertain responses to hold them for human intervention.

To address these challenges, Mitani Sangyo’s "AI Reliability Governance Framework" provides step-by-step management through three checkpoints—Input, Process, and Output—that control how AI systems generate responses and ensure information is traceable, consistent, and safe before any answer is delivered.

A New Mechanism to Suppress "AI Hallucinations" (Plausible but Incorrect Information)



Overall view of the AI Reliability Governance Framework and Value Provided

The AI Reliability Governance Framework introduces three checkpoints—Input, Process, and Output—along the data flow (Input → Process → Output), each designed to reduce the likelihood of hallucinations and increase traceability.

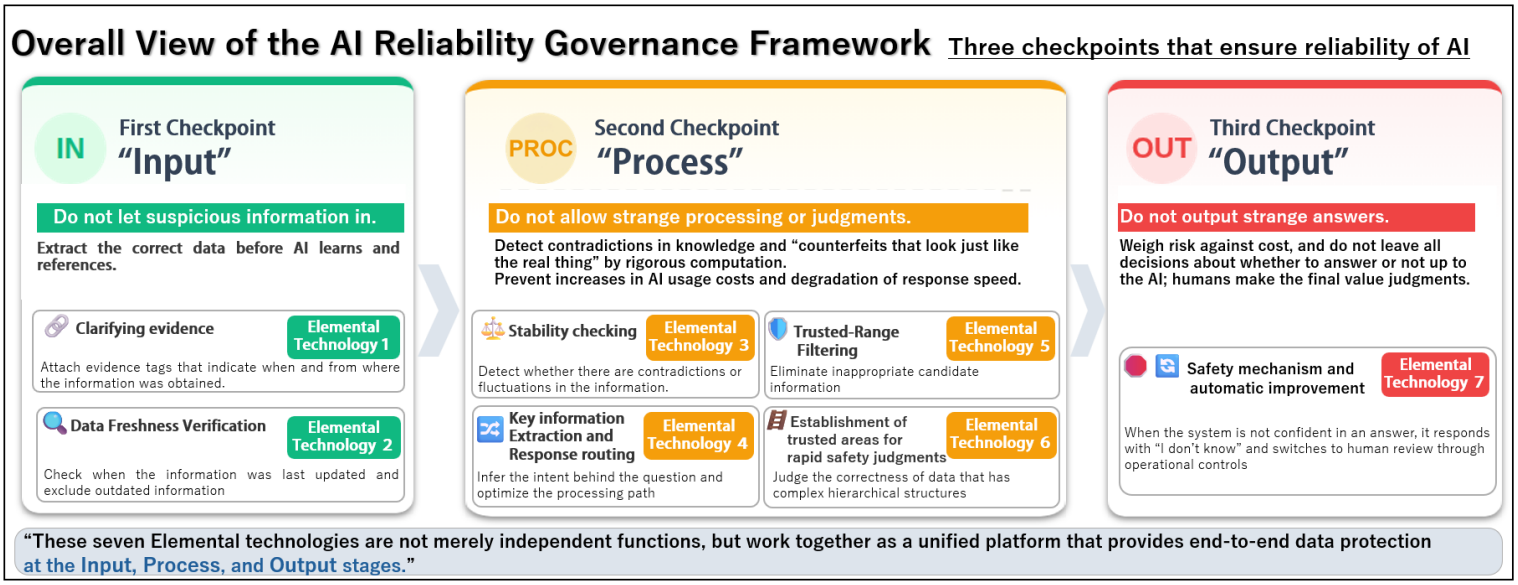
At the Input checkpoint, reference data is tagged with evidence such as where and when it was obtained, allowing the system to distinguish information that should be admitted from information that should be filtered out.

At the Process checkpoint, the framework combines multiple technologies to improve both accuracy and speed: a mechanism to determine whether the knowledge possessed by AI is reliable, information compression technology to efficiently organize data, a method for judging based on the range of reliability, and a technology that can quickly determine the boundary between judgment and non-judgment.

At the Output checkpoint, the framework assesses answer confidence and associated risk and switches behavior accordingly: returning an answer, withholding it, or escalating to human review.

Together, these controls ensure that every response is supported by traceable, up-to-date, and consistent information, providing a systematic guarantee of the reliability required for mission-critical operations.

Data Flow	Conventional Challenges	Core Technology	Value delivered by the AI Reliability Governance Framework
Input	In many cases, the sources of the information used to generate AI answers are unclear, making it difficult to verify or substantiate them.	1.	<u>Clarifying Evidence</u> Attach “evidence tags” to training data and reference information that indicate when and from where the information was obtained, enabling traceability of the basis for each answer.
	AI cannot distinguish between new and old information and may answer based on outdated data.	2.	<u>Data Freshness Verification</u> Before AI processes information, the system checks when it was last updated and excludes outdated information from the reference set.
Process	Similarity-based judgments alone cannot distinguish information that contains contradictions or instability.	3.	<u>Stability Checking</u> Introduce mechanisms to evaluate the consistency and stability of knowledge and to detect information that contains contradictions or volatility.
	Handling all queries solely with AI increases cost and response time.	4.	<u>Key Information Extraction & Response Routing</u> Extract representative points of information and route simple questions to FAQ systems and complex ones to AI, improving processing efficiency.
	Simple threshold-based scoring has difficulty distinguishing between genuinely different but similar-looking information.	5.	<u>Trusted-Range Filtering</u> Define a “trusted range” of information that AI is allowed to reference and exclude inappropriate candidates.
	Hierarchical data structures are computationally heavy, making safe judgment difficult.	6.	<u>Establishment of trusted areas for rapid safety judgments</u> Introduce mechanisms that can handle complex hierarchical structures efficiently while allowing rapid safety judgments.
Output	In some cases, AI presents wrong answers with high confidence.	7.	<u>Safety Mechanism and Auto-improvement</u> Implement operational controls that switch between answering, holding, or escalating to human review based on answer confidence and business risk.



Comment from the inventor

"What matters is not just how well AI can answer, but whether it knows when to stop," said Hiroaki Nakano, Information Systems Division, Mitani Sangyo Co., Ltd. In real business settings, the biggest risk is not that AI fails silently, but that it answers confidently when it should not.

This concept did not come from a large research lab. It emerged from hands-on deployments where we repeatedly had to ask: when should AI be trusted, and when should it step back?

To use AI in mission-critical operations, organizations must be able to explain when the underlying information was obtained and where it came from—and the system itself must be able to abstain when uncertainty is high. Through this framework, we are turning uncertainty from a hidden risk into something that can be explicitly managed—enabling AI systems that deliver not only answers, but also justified non-responses with transparency and accountability.

Future roadmap

Building on the U.S. provisional patent application, Mitani Sangyo will pursue PCT applications* and other patent filings, and will promote the practical application of a reliability platform for AI use in important business through joint demonstrations and partner collaborations with companies and research institutions in Japan and overseas. At the same time, Mitani Sangyo will continue to share technical insights through academic publications, striving to provide theoretical grounding and ensure transparency. Mitani Sangyo will continue to advance the development of AI frameworks that balance usability and safety. Mitani Sangyo also hopes that, in the future, enterprises will be able to confidently adopt and utilize reliable AI.

*The PCT (Patent Cooperation Treaty) is an international patent system that allows a single application to seek protection in multiple member countries.

About Mitani Sangyo Group

Mitani Sangyo is a diversified trading company founded 98 years ago in Kanazawa, Ishikawa Prefecture, Japan, and 31 years ago in Vietnam. The company primarily conducts its business in the Hokuriku region, the Tokyo metropolitan area, and Vietnam, across six segments: Chemicals, Information Systems, Plastics & Electronics, Air-Conditioning Facilities, Housing Equipment, and Energy. While operating as a trading company, Mitani Sangyo also acts as a manufacturer and consultant, pursuing optimal solutions for its customers and continuing to evolve as a “nearly 100-year-old venture company.”

Fiscal year ended March 2025: consolidated net sales of 103,072 million yen; consolidated workforce of 3,563 employees.

For more information, please visit our website.

https://www.mitani.co.jp/EN

Media and business contacts

Media inquiries:

PR Planning Office, Mitani Sangyo Co., Ltd. (c/o Kyodo Public Relations)

(Contacts: Takizawa, Yu, Kishimoto)

Email: mitani-pr@kyodo-pr.co.jp

Corporate customers and technology / research collaboration inquiries:

Information Systems Division, Mitani Sangyo Co., Ltd.

https://forms.office.com/r/0PZrTM8Mng (web form link)

(Contacts: Nakano, Doi)