

Research Article

Generative AI, Cloud Computing, and Configurable Workflow Architecture for Next-Generation Risk Management

Abhishek Tiwari*

Computer Science (Artificial Intelligence & Machine Learning), SR Institute of Institutions, Lucknow, Uttar Pradesh, India

Abstract

Risk management is dominated by text. Incident reports, control descriptions, audit findings, regulatory updates, and risk committee papers are written, read, compared, and summarized by people, often under time pressure. Generative artificial intelligence (AI), and large language models (LLMs) in particular, can draft, summarize, classify, and compare such text at scale. It also introduces new risks: fabricated content, sensitivity to manipulated inputs, and outputs that are hard to audit. This article argues that generative AI becomes safe and useful in risk management when it is placed inside configurable workflows rather than used as a free-standing assistant. Each generative step is configured with its task, its approved sources, the checks its output must pass, and the person who must approve it. The article proposes an architecture in which a configurable workflow engine calls a generative AI gateway, which grounds outputs in the organization's own documents through retrieval, validates them, and records every input and output. All components run as cloud services with self-healing and safe fallback. Using a design-oriented approach grounded in eleven studies published between 2020 and 2023, the article maps the evidence base, places generative tasks across the risk lifecycle, sets out a controlled generative step, and maps generative AI risks to workflow controls. It illustrates the design with a pharmaceutical manufacturer managing quality deviations and regulatory change. It concludes that the workflow, not the model, should decide what generative AI is allowed to do.

Keywords: generative AI, large language models, retrieval-augmented generation, risk management, configurable workflow, cloud computing, hallucination, prompt injection, AI risk management framework, human oversight, self-healing systems

Introduction

Much of the work of risk management is reading and writing. A risk analyst reads incident reports to find the root cause, maps the incident to the risk register, checks which controls should have prevented it, and writes an assessment. A compliance officer reads new regulations and works out which policies and controls they affect. Each of these tasks consumes expert time, and backlogs grow when volumes rise.

Large language models change what is possible here. Brown et al. (2020) showed that a sufficiently large model can perform many language tasks from a few examples in the prompt. Ouyang et al. (2022) showed that training with human feedback makes models far better at following instructions. Bommasani et al. (2021) described such systems as foundation models: general models adapted to many downstream tasks, bringing both opportunities and risks.

Those risks are serious in a risk function. Models can produce fluent but unsupported content, known as

hallucination (Ji et al., 2023). They can leak information, spread misinformation, or be misused (Weidinger et al., 2022). When they read external content, hidden instructions in that content can take control of their behavior (Greshake et al., 2023). A risk function that adopts generative AI without controls simply adds new risks to its own register.

This article proposes that generative AI be governed through the same mechanism that already governs risk processes. Basireddy (2022) shows how risk management can be digitized through configurable enterprise workflow architecture, in which risk identification, assessment, approval, escalation, and remediation are defined as configurable workflows with audit trails. If each generative step is a configured workflow step, then the workflow determines which sources the model may use, which checks apply, and who must approve the result. Retrieval-augmented generation (RAG) grounds the output in approved documents (Lewis et al., 2020). The NIST AI Risk Management Framework provides the

governance structure (National Institute of Standards and Technology [NIST], 2023). Cloud services provide scale and reliability (Sannidhanam, 2020, 2023).

This article asks: *How can generative AI be integrated into risk management through configurable workflows on cloud infrastructure so that its benefits are gained while its own risks are controlled?* It contributes an evidence map (Figures 1 and 2), a mapping of generative tasks to the risk lifecycle, an integrated architecture, a controlled generative step, a mapping of generative AI risks to workflow controls, and an illustrative case.

Related Work

Configurable risk workflows

Basireddy (2022) proposes configurable enterprise workflow architecture for digitizing risk management. Risk processes are expressed as configuration rather than code, so they can change quickly and consistently, with each step recorded. This article treats that configuration layer as the natural place to define what generative AI may do.

Large language models

Brown et al. (2020) introduced GPT-3 and showed few-shot learning across many tasks. Ouyang et al. (2022) showed that reinforcement learning from human feedback produces models that follow instructions more closely and are preferred by human raters, while noting that they still make simple mistakes. Bommasani et al. (2021) analyzed the opportunities and risks of foundation models, including the concentration of risk when many applications depend on the same model.

Grounding and hallucination

Lewis et al. (2020) proposed retrieval-augmented generation, which retrieves relevant documents and conditions generation on them, improving factuality and allowing sources to be cited and updated. Ji et al. (2023) surveyed hallucination in natural language generation, including its causes, measurement, and mitigation. For risk management, where assessments may be relied on by regulators and boards, grounding and verification are essential.

Risks of generative AI

Weidinger et al. (2022) propose a taxonomy of risks posed by language models, including discrimination, information hazards, misinformation, malicious use, and human-computer interaction harms. Greshake et al. (2023) demonstrate indirect prompt injection, in which instructions hidden in retrieved content take over an LLM-integrated application. NIST (2023) provides a voluntary framework for managing AI risk through four functions: govern, map, measure, and manage.

Cloud services and resilience

Sannidhanam (2020) presents a scalable web services architecture for healthcare data processing, showing how sensitive, high-volume data can be handled by service-based systems. Sannidhanam (2023) examines self-healing distributed systems using AI-driven failure prediction and automated recovery. Generative services are compute-intensive and depend on external model endpoints, so these properties matter.

Research gap

Generative AI research focuses on model capability and model risks, while risk workflow research predates generative AI. Few designs show how a configurable risk workflow can host generative steps, with grounding, validation, approval, and fallback defined in the workflow itself. This article addresses that gap.

Research Methodology

The study follows a design-oriented research approach. The problem, the gap between generative AI's capability for risk text and its uncontrolled use, was identified from the literature. Requirements were derived from the reviewed studies: grounding in approved sources, validation of outputs, resistance to manipulated inputs, human approval, auditability, and resilience. The architecture and the controlled generative step were designed from eleven studies published between 2020 and 2023 and demonstrated through an illustrative case. To make the evidence base transparent, each study was assigned to the theme it primarily supports. Figure 1 shows the distribution across themes, and Figure 2 shows the studies by publication year. These figures describe the literature used in this article. They are not performance measurements. The architecture has not been evaluated empirically, which is left to future work.

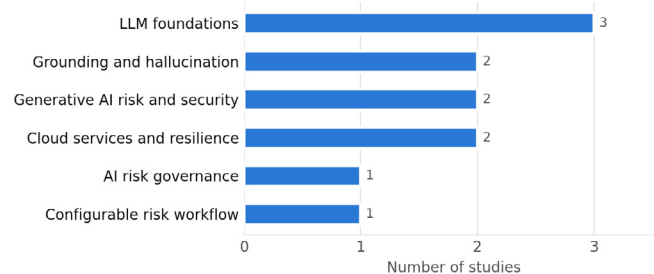


Figure 1: Number of reviewed studies primarily supporting each theme (n = 11)

Large language model foundations have the most studies, with three. Grounding and hallucination, generative AI risk and security, and cloud services and resilience have two each. AI risk governance and configurable risk workflow have one each. All studies fall between 2020

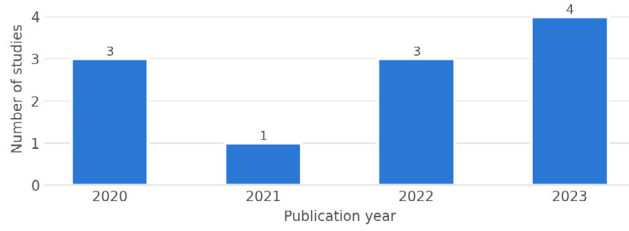


Figure 2: Reviewed studies by publication year (n = 11)

and 2023, with the most in 2023, reflecting how recent the generative AI literature is.

Generative AI Across the Risk Lifecycle

Table 1 places generative tasks at each stage of the risk lifecycle. Each task has an approved grounding source and a human checkpoint.

The Table 1 shows a consistent pattern. Generative AI prepares a draft from approved sources, and a named person decides. The model accelerates the work but does not own the decision.

Integrated Architecture

Figure 3 shows the proposed architecture, and Table 2 describes its components.

Configurable workflow engine. The engine runs risk processes as configured workflows (Basireddy, 2022). A generative step is configured like any other step, with a task, sources, checks, and an approver.

Generative AI gateway. All model calls pass through a gateway. It applies approved prompt templates, filters inputs, removes sensitive data where required, and logs every request and response. Treating retrieved text as data, not instructions, reduces the risk of indirect prompt injection (Greshake et al., 2023).

Retrieval service. The retrieval service searches only the document collections approved for that step and returns passages with their sources (Lewis et al., 2020). Access

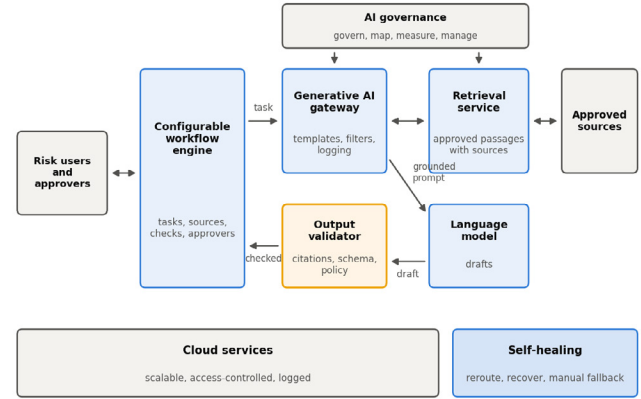


Figure 3: Architecture for generative AI in configurable risk workflows

control ensures that a user sees only content they are entitled to see.

Output validator. Before an output reaches a person, the validator checks that every claim cites a retrieved source, that the output follows the required structure, and that it does not break policy rules. Outputs that fail are regenerated or flagged (Ji et al., 2023).

AI governance. Generative use is governed through the four NIST functions (NIST, 2023). Each generative step is registered, its risks are mapped, its quality is measured, and issues are managed.

Cloud services and self-healing. All components run as scalable cloud services (Sannidhanam, 2020). Failure prediction and automated recovery keep them available (Sannidhanam, 2023). If the model endpoint is unavailable, the workflow falls back to the manual version of the step, so risk work is delayed rather than lost.

The Controlled Generative Step

Step lifecycle

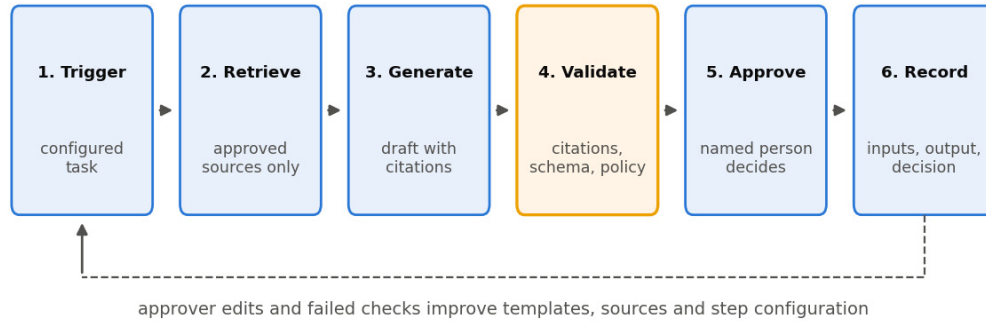
Figure 4 shows how a single generative step runs inside a workflow.

Table 1: Generative AI tasks across the risk lifecycle

Risk stage	Generative task	Grounding source	Human checkpoint	Supporting studies
Identify	Summarize incidents and extract risk events	Incident reports	Analyst confirms extracted events	Brown et al. (2020); Basireddy (2022)
Assess	Draft risk assessment and map to risk register	Risk register, control library	Risk owner approves rating	Lewis et al. (2020); Basireddy (2022)
Respond	Propose remediation actions	Past remediation records, policies	Owner accepts or edits actions	Ouyang et al. (2022)
Monitor	Analyze regulatory changes against policies	Regulations, policy library	Compliance officer confirms impact	Lewis et al. (2020); Ji et al. (2023)
Report	Draft committee summaries	Workflow records and metrics	Head of risk signs off	Basireddy (2022)

Table 2: Components of the architecture

Component	Function	Key mechanisms	Supporting studies
Configurable workflow engine	Runs risk processes and calls generative steps	Step configuration, approvals, audit trail	Basireddy (2022)
Generative AI gateway	Single controlled path to models	Prompt templates, input filtering, logging	Weidinger et al. (2022); Greshake et al. (2023)
Retrieval service	Grounds outputs in approved documents	Document index, access control, citations	Lewis et al. (2020)
Output validator	Checks outputs before use	Citation check, schema check, policy rules	Ji et al. (2023)
Language model	Generates drafts	Instruction-following model	Brown et al. (2020); Ouyang et al. (2022)
AI governance	Oversees generative use	Govern, map, measure, manage	NIST (2023)
Cloud services and self-healing	Hosts and protects all services	Scalable services, failure prediction, fallback	Sannidhanam (2020, 2023)

**Figure 4:** Lifecycle of a controlled generative step in a risk workflow

The workflow triggers the step with a configured task. The retrieval service retrieves passages from approved sources. The model generates a draft with citations. The validator validates the draft; this step is highlighted because it is where unsupported or unsafe outputs are stopped. A named person approves, edits, or rejects the

draft. Finally, the workflow records the inputs, output, and decision, and approver edits are used to improve templates and sources.

Generative AI risks and workflow controls

Table 3 maps the main generative AI risks to controls that the workflow applies.

Table 3: Generative AI risks and workflow controls

Risk	Example in risk management	Workflow control	Supporting studies
Hallucination	Assessment cites a control that does not exist	Retrieval grounding; citation check	Lewis et al. (2020); Ji et al. (2023)
Prompt injection	A document contains hidden instructions	Gateway treats retrieved text as data; input filtering	Greshake et al. (2023)
Information leakage	Confidential finding appears in wrong report	Access control on retrieval; output rules	Weidinger et al. (2022)
Over-reliance	Drafts approved without review	Named approver; sampled quality review	Weidinger et al. (2022); NIST (2023)
Model dependency	Many workflows depend on one model	Registered use; fallback to manual step	Bommasani et al. (2021); Sannidhanam (2023)

The controls are defined in workflow configuration, not left to individual users. This gives the risk function one place to see, audit, and change how generative AI is used.

Illustrative Case: Quality and Compliance Risk in Pharmaceutical Manufacturing

Consider a pharmaceutical manufacturer that records hundreds of quality deviations each month and must track frequent regulatory updates. The case is illustrative and reports no measured results.

Deviation triage. When a deviation report is filed, the workflow triggers a generative step that summarizes the event, suggests the affected product and process, and proposes a risk category. The retrieval service draws only on the deviation report, the product's process documents, and similar past deviations (Lewis et al., 2020). The validator rejects the summary if it names a process step that is not in the retrieved documents (Ji et al., 2023). A quality analyst confirms or corrects the category before the deviation is routed (Basireddy, 2022).

Root cause and actions. For deviations rated high risk, a second generative step drafts candidate root causes and corrective actions based on earlier investigations. The quality lead reviews and edits the draft. Every edit is recorded, and repeated edits show where templates or sources need improvement.

Regulatory change. When a regulator publishes new guidance, a generative step compares it with the company's policy library and lists policies that may be affected, each with the supporting passages. Because the guidance comes from outside, the gateway treats its text as data only, so any embedded instructions are not followed (Greshake et al., 2023). A compliance officer confirms the impact before any policy change workflow starts.

Reliability. Deviation reports contain sensitive data and arrive in bursts after production runs, so the services scale with demand (Sannidhanam, 2020). If the model endpoint fails, the self-healing layer reroutes requests or the workflow falls back to manual triage (Sannidhanam, 2023).

Governance. Each generative step is registered in the AI inventory. Its purpose and risks are mapped, and its citation accuracy and approver edit rate are measured over time. Steps that perform poorly are managed through changes to templates, sources, or step configuration (NIST, 2023).

The case shows generative AI reducing the effort of reading and drafting while decisions remain with accountable people and every output remains traceable.

Fiscussion

Implications

For research, the article shows how existing risk workflow architecture can host generative AI and how

generative AI risks can be controlled through workflow configuration. For practitioners, the lessons are to route all model use through a gateway, restrict each step to approved sources, validate outputs before people see them, keep a named approver for every decision, record everything, and define a manual fallback. Treating each generative step as a registered, measured use gives the risk function the same oversight of its own AI that it expects of the business.

Limitations and future work

The architecture was demonstrated through an illustrative case and draws on eleven studies, so the evidence base is focused rather than comprehensive. Several descriptions of the studies are based on their stated scope. Future work should implement the design and measure citation accuracy, reviewer effort, time to decision, and the rate of rejected outputs. It should also test resistance to prompt injection, and examine how approvers' trust in drafts changes over time. Model selection, cost management, and data residency require fuller treatment.

Conclusion

Generative AI can relieve risk functions of much of their reading and drafting, but it also brings risks of its own. This article proposed placing every generative step inside a configurable risk workflow on resilient cloud services. A gateway controls access to models, retrieval grounds outputs in approved sources, a validator stops unsupported content, a named person approves, and every step is recorded and governed. The pharmaceutical case showed this design supporting deviation triage and regulatory change analysis while keeping people accountable. The central conclusion is that the workflow, not the model, should define what generative AI may do in risk management.

References

- Basireddy, S. R. (2022). Digitizing risk management through configurable enterprise workflow architecture. *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology*, 14(4), 231–239.
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., et al. (2021). *On the opportunities and risks of foundation models* (arXiv:2108.07258). arXiv.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., et al. (2020). Language models are few-shot learners. In *Advances in Neural Information Processing Systems* 33 (pp. 1877–1901).
- Greshake, K., Abdelnabi, S., Mishra, S., Endres, C., Holz, T., & Fritz, M. (2023). Not what you've signed up for: Compromising real-world LLM-integrated applications with indirect prompt injection. In *Proceedings of the 16th ACM Workshop on Artificial Intelligence and Security* (pp. 79–90).

- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), Article 248.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems* 33 (pp. 9459–9474).
- National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). <https://doi.org/10.6028/NIST.AI.100-1>
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., et al. (2022). Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems* 35 (pp. 27730–27744).
- Sannidhanam, A. H. (2020). Scalable web services architecture for healthcare data processing. *International Journal of Recent Advances in Science and Technology*, 7(01), 1–14. <https://ijrast.in/index.php/ijrast/article/view/65>
- Sannidhanam, A. H. (2023). Self-healing distributed systems: AI-driven failure prediction and automated recovery. *International Journal of Research & Technology*, 11(4), 168–174. <https://ijrt.org/j/article/view/1464>
- Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P.-S., Mellor, J., et al. (2022). Taxonomy of risks posed by language models. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (pp. 214–229).