Generative AI for Risk Assessments: A Taxonomy Across Models, Lifecycles and Domains

Aakash Kharb
Maharshi Dayanand University, Rohtak

[1]*Received: 28/08/2024; Accepted: 29/09/2024; Published: 24/10/2024*

ABSTRACT

Generative artificial intelligence (GenAI) has moved from experimental labs into high-stakes domains such as finance, chemical safety, cybersecurity, and healthcare. While a fast-growing body of work explores generative models, there is still no unified view of how GenAI applications fit into the broader landscape of risk assessment. This paper proposes a taxonomy of generative AI applications for risk assessment across sectors. We synthesize literature on deep generative models and AI-based risk management, then classify GenAI applications along three main axes: (i) risk lifecycle phase (identification, assessment, monitoring, mitigation, communication), (ii) model family (large language models, GANs, VAEs, diffusion, hybrid and graph-based generative models), and (iii) risk domain (financial/systemic, chemical and environmental, safety-critical and infrastructure, organizational and governance). The taxonomy is used to compare capabilities and limitations of different application classes, with particular attention to model risk, explainability, and regulatory alignment. We draw on work on deep generative models, generative adversarial networks, systemic risk, chemical risk assessment, model-risk management, and AI impact assessment frameworks published between 2013 and 2023. The paper concludes with a research agenda emphasizing evaluation benchmarks, human-in-the-loop workflows, and sector-specific guidance for trustworthy GenAI-enabled risk assessment.

Keywords: *Generative AI; Taxonomy; Risk Assessment*

1. Introduction

Generative AI systems such as large language models (LLMs), generative adversarial networks (GANs) and variational autoencoders (VAEs) have rapidly expanded the frontier of AI capabilities, enabling synthetic data generation, scenario creation and automated reasoning over unstructured information. In parallel, organizations face growing pressure to strengthen risk assessment in response to systemic financial risks, chemical and environmental hazards, and emerging AI-specific failures.

Most existing risk-assessment frameworks treat AI as a predictive tool embedded in conventional pipelines. Deep generative models challenge this view by (a) creating synthetic but realistic data, (b) exploring hypothetical risk scenarios, and (c) drafting human-readable risk artefacts such as reports or impact assessments. However, GenAI also introduces new risks, including hallucination, data leakage, and systemic vulnerabilities when many actors adopt similar generative tools.

Despite rapidly growing interest, research remains fragmented: domain papers focus on specific use cases (e.g., chemical risk, biomedical imaging) while AI-governance work speaks in high-level terms about "impact assessments" and "model risk." This paper addresses that gap by proposing a cross-sector taxonomy of generative AI applications for risk assessment and using it for comparative analysis.

[1] *How to cite the article:* Kharb A (October, 2024); Generative AI for Risk Assessments: A Taxonomy Across Models, Lifecycles and Domains; *International Journal of Technology, Science and Engineering*; Vol 7 Issue 4; 28-38

Table 1. Drivers and tensions around GenAI for risk assessment

Driver / Concern	Description	Typical Sector Examples
Data complexity	High-dimensional, multi-modal risk data hard to model with simple tools	Finance, cyber, chemical safety
Need for scenario exploration	Stress-testing rare or extreme events	Systemic financial risk, climate risk
Shortage of expert capacity	Limited human experts for detailed assessments	Chemical risk, infrastructure safety
Demand for explainability and auditability	Regulators require transparent, documented risk reasoning	Banking, toxicology, healthcare
New GenAI-specific risks	Hallucination, data leakage, systemic herding on AI-generated outputs	All sectors deploying GenAI

In the remainder of the paper, we first summarize relevant background on generative AI and risk-assessment concepts, then introduce our taxonomy and use it for comparative analysis of application classes and domains.

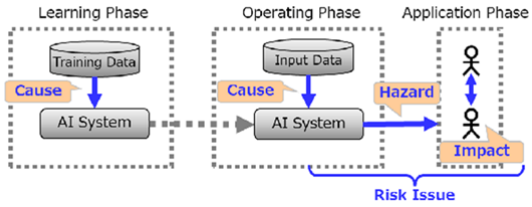

Fig 1: Model of Risk Occurrence

2. Background: Generative AI and Risk Assessment

2.1 Deep generative models and GenAI

Deep generative models estimate complex data distributions and sample from them, rather than merely predicting labels. Survey work over the last decade distinguishes families such as (i) energy-based models and restricted Boltzmann machines, (ii) VAEs, (iii) GANs, and (iv) autoregressive and other deep generative architectures. The landmark Generative Adversarial Nets paper in 2014 introduced the adversarial training setup that has dominated many generative applications. GAN variants have since been applied widely, including to biomedical image segmentation and other safety-relevant imaging tasks.

From 2018 onwards, deep generative models expanded beyond images to graphs, molecules, and text, supporting tasks such as drug and material discovery, graph-structured risk modelling, and more general data-augmentation. Parallel advances in large-scale language modelling culminated in LLMs capable of generating coherent multi-paragraph text, code, and structured outputs. These capabilities underpin many GenAI applications in risk-assessment workflows: converting unstructured documents into structured signals, simulating scenarios, or generating narrative reports.

2.2 Risk assessment and AI risk management

Risk assessment is usually conceptualized as a lifecycle: risk identification, analysis and evaluation, mitigation, monitoring, and communication. Generic standards such as ISO 31000 and emerging AI-specific frameworks like NIST's AI Risk Management Framework emphasize systematic identification of harms, likelihood and impact quantification, and governance mechanisms.

Within AI, "model risk management" (MRM) focuses on the risk of the AI model itself: misspecification, overfitting, data drift, and governance failures. Malhotra (2018) argued that robust AI deployment requires extending traditional model-risk practices to deep learning and neural networks, with dedicated validation, documentation and monitoring.

Danielsson et al. (2022) highlight further that widespread AI adoption can amplify *systemic* risk: similar models may drive many institutions toward the same strategies, making crises more synchronized and severe.

In parallel, domain-specific work has explored AI for risk assessment in chemicals, finance, and other sectors. Wittwehr et al. (2020) describe how AI, including future generative methods, can support *chemical risk assessment* (CRA) by integrating heterogeneous evidence, generating mechanistic hypotheses, and enabling human–AI collaboration.

More recently, systematic reviews on AI impact assessments (AI-IAs) have catalogued approaches for anticipating social, ethical, and legal impacts before deployment. These reviews mostly consider AI in general rather than GenAI in particular, but they offer a conceptual substrate for positioning generative tools inside broader risk-governance frameworks.

Table 2. Core concepts linking generative AI and risk assessment

Concept	Brief description	Relevance to GenAI-based risk assessment
Deep generative model	Learns data distribution & samples from it	Synthetic data, scenario sampling
Model risk management	Governance of AI model errors and misuse	Controls for GenAI hallucination, bias, drift
Systemic risk	Risk of collapse of an entire system/market	Herding if many actors rely on similar GenAI tools
Chemical risk assessment	Evaluation of hazards, exposure and uncertainty	GenAI for evidence synthesis and hypothesis generation
AI impact assessment	Structured anticipation of AI's socio-technical impacts	LLMs as co-pilots to draft and refine assessments

3. Taxonomy of Generative AI Applications for Risk Assessment

We propose a taxonomy organized along three orthogonal axes (Figure described verbally):

1. **Risk lifecycle phase** where GenAI is applied.
2. **Model family / capability** of the generative system.
3. **Risk domain** in which assessment is performed.

A given application is a tuple (phase, model family, domain). Below we describe each axis and illustrate typical combinations.

3.1 Axis 1: Risk lifecycle phase

1. **Risk identification**
    - Mining unstructured text (reports, news, logs) to discover emerging hazards, threats or control failures.
    - LLMs summarizing incident databases or drafting risk registers.
2. **Risk assessment & quantification**
    - Generating synthetic data for rare events (e.g., extreme losses) to improve statistical estimates.
    - Simulating counterfactual scenarios that stress models and portfolios.

3. **Risk mitigation design**
    - Generating candidate mitigation strategies, decision trees or control libraries.
    - Suggesting testing plans, red-team prompts, or policy options.
4. **Risk monitoring & early warning**
    - Generative anomaly detection: models learn "normal" behavior and flag deviations.
    - Synthetic augmentation of scarce labeled anomalies.
5. **Risk communication & documentation**
    - Drafting human-readable risk reports, model cards, and AI impact assessments.
    - Generating counter-examples, narratives and visualizations to explain risk to stakeholders.

3.2 Axis 2: Model family / capability

Drawing on surveys of deep generative models and GAN applications, we distinguish the families in Table 3.

Table 3. Generative model families and typical risk-assessment roles

Model family	Typical data modality	Typical capabilities in risk assessment
GANs	Images, tabular, time-series	Synthetic data for stress tests; anomaly detection; image-based risk
VAEs / latent models	Tabular, time-series, graphs	Scenario generation; uncertainty modeling; dimensionality reduction
Autoregressive / LLMs	Text, code, semi-structured	Evidence synthesis; report drafting; scenario narratives; code risks
Diffusion models	Images, audio, 3D	Synthetic sensor data; rare event simulation; design-space exploration
Graph generative models	Networks, molecules	Systemic-risk structures; molecular/chemical risk candidates
Hybrid architectures	Multi-modal	Combined text + structure (e.g., reports + exposure graphs)

GANs and VAEs are particularly useful for scenario generation and synthetic data, while LLMs dominate unstructured-text tasks and code generation. Hybrid and graph models support domain-specific structures such as financial networks or chemical interaction graphs.

3.3 Axis 3: Risk domain

We group domains into four broad classes:

1. **Financial and systemic risk**
    - Scenario generation for stress testing portfolios and liquidity risk.
    - LLM-assisted analysis of regulatory text and supervisory reports.
2. **Chemical, environmental and health risk**
    - Generative models for molecular design and toxicity prediction.
    - AI-supported CRA workflows integrating heterogeneous evidence.

3. **Safety-critical infrastructure and engineering**
    - Generative models for structural health monitoring, anomaly detection in sensor data, and incident reconstruction.
4. **Organizational, compliance and governance risk**
    - LLMs for AI impact assessments, policy simulation, and drafting risk-management documentation.

3.4 Putting the taxonomy together

Table 4 illustrates the taxonomy as a matrix of exemplar tuples.

Table 4. Example GenAI application classes in the proposed taxonomy

Example application	Lifecycle phase	Model family	Risk domain
Synthetic credit-loss scenarios for stress testing	Assessment / quant.	VAE / GAN	Financial & systemic
Generating toxicology hypotheses from multi-source data	Identification / analysis	LLM + graph models	Chemical & environmental
Structural anomaly detection from vibration data	Monitoring	GAN / VAE	Safety-critical infrastructure
Drafting AI impact-assessment documents	Communication	LLM	Organizational & governance
Red-team prompt generation for GenAI systems	Mitigation design	LLM	Cross-domain / AI model-risk

This taxonomy is intentionally high-level: real systems often combine multiple phases and model types. However, it provides a useful scaffold for comparative analysis across different GenAI-based risk-assessment solutions.

Table 5. Comparative view: where GenAI adds most value by phase

Lifecycle phase	Traditional limitation	Added value from GenAI
Identification	Manual scanning of documents and incidents	Automated mining & summarization of large corpora (LLMs)
Assessment	Limited data for tail-events	Synthetic scenarios; uncertainty-aware simulations
Mitigation design	Ad-hoc, expert-driven rule creation	Generative exploration of candidate controls
Monitoring	Static thresholds, simple anomaly rules	Deep generative models for complex pattern deviation
Communication	Time-consuming report writing	Rapid drafting, multi-audience tailoring via LLMs

4. Comparative Analysis of Taxonomic Categories

In this section we compare major groups of GenAI applications using the taxonomy. We focus on four comparative dimensions: (i) data requirements, (ii) interpretability, (iii) regulatory alignment, and (iv) induced model-risk.

4.1 Model-family comparison

Survey work shows that different generative families trade off fidelity, controllability, and interpretability.

Table 6. Comparative analysis by model family

Dimension	GANs	VAEs / latent models	LLMs / autoregressive
Data requirements	Large labeled/unlabeled sets; mode-collapse risk	Large datasets but robust to noise; probabilistic latent space	Massive text/code corpora; pretraining often external
Output controllability	Good with conditioning but training unstable	Latent-space arithmetic; smooth but lower fidelity	Strong via prompting & fine-tuning, but indirect control
Interpretability	Low; opaque latent variables	Moderate; latent factors analyzable statistically	Variable; text is human-readable but internal reasoning opaque
Regulatory alignment	Strong in imaging-heavy domains with clear metrics	Useful where uncertainty quantification is needed	Strong for documentation/AI-IA drafting; weaker for formal quantification
Model-risk profile	Mode collapse, artifacts, spurious correlations	Miscalibrated uncertainty; latent mis-specification	Hallucination, training-data bias, prompt injection

In risk assessment practice, **LLMs** are most immediately usable for documentation, triaging and evidence synthesis, while **GANs and VAEs** are powerful for simulation and anomaly detection where quantitative validation metrics exist. Graph and diffusion models are attractive in specialized domains but currently less mature in production risk-management workflows.

4.2 Domain-specific contrasts

1. **Financial and systemic risk**
    - Danielsson et al. warn that AI can *increase* systemic risk by encouraging homogeneous strategies, even when models are individually accurate.
    - GenAI that generates scenarios or recommendations may worsen herding if many institutions adopt similar models or pre-trained LLMs.
    - On the other hand, generative stress-testing can uncover vulnerabilities that would be missed by simple historical scenarios.
2. **Chemical and environmental risk**
    - Wittwehr et al. describe AI as a way to better integrate diverse evidence in chemical risk assessment and to support human–AI collaboration, rather than fully automate decisions.
    - Generative models can propose mechanistic explanations or candidate molecules, but CRA decisions remain tightly regulated, with extensive human oversight.
3. **Safety-critical infrastructure**
    - Deep generative models for structural health monitoring provide rich anomaly-detection capabilities but must be carefully validated against real-world failure data.
4. **Organizational and governance risk**
    - LLM-assisted AI impact assessments can improve coverage and reduce cost but may embed biases from their training data and encourage “checkbox” compliance instead of genuine reflection.

Table 7. Comparative analysis by risk domain

Domain	Primary generative role	Key advantages	Key concerns
Financial & systemic	Scenario generation, stress-testing	Explore rare events; richer systemic simulations	Herding, pro-cyclicality, opaque model chains
Chemical & environmental	Evidence synthesis, molecular generation	Integrates heterogeneous data; suggests hypotheses	Regulatory conservatism; need for mechanistic insight
Safety-critical infrastructure	Sensor anomaly detection, virtual testing	Detect complex failure patterns early	Data scarcity; safety certification requirements
Organizational & governance	AI-IA drafting, policy simulation	Scalable documentation; stakeholder-tailored reports	Hallucinated arguments; shallow compliance

4.3 GenAI in AI risk and impact assessments

A particularly interesting use case is *using GenAI to assess GenAI itself*. Stahl et al. catalog AI impact assessment practices across sectors and highlight the difficulty of modeling dynamic AI systems whose behavior shifts over time. Barrett et al. provide guidance for high-consequence AI risk management, including red-team exercises and catastrophic-risk modeling.

LLMs can:

- Summarize system descriptions, regulatory requirements, and stakeholder concerns into draft AI-IA documents.
- Generate candidate misuse scenarios or catastrophic-risk stories to seed expert workshops.
- Assist in translating technical model behaviors into layperson-readable impacts.

However, because GenAI tools can hallucinate risks and mitigation claims, they must be embedded in **human-in-the-loop** workflows with clear provenance and validation.

Table 8. Comparative view: human vs GenAI roles in AI-impact assessment

Task	Human-expert strengths	GenAI strengths	Recommended mode
Identifying context & stakeholders	Domain knowledge, tacit understanding	None (relies on prompt content)	Human-led
Drafting assessment narratives	Nuanced judgment but time-consuming	Fast generation, multi-style outputs	GenAI-draft, human-edit
Enumerating risk scenarios	Experience, creativity on real cases	Rapid variant generation, cross-sector analogies	Co-creative brainstorming
Evaluating likelihood & impact	Judgment, access to organizational data	Weak; limited calibrated reasoning	Human-led, GenAI as calculator
Documenting controls & mitigations	Knowledge of real processes	Suggest patterns based on public frameworks	Human-led, GenAI as assistant

5. Challenges, Limitations and Research Agenda

Our taxonomy reveals not only opportunities but also cross-cutting challenges that must be addressed for GenAI-based risk assessment to be trustworthy and effective.

5.1 Data and evaluation challenges

1. **Synthetic data validity**
    - GANs and VAEs can produce realistic-looking samples that nevertheless distort tail distributions or dependence structures that drive risk.
    - In high-stakes settings (e.g., capital regulation, CRA), synthetic data should be treated as *hypotheses* requiring empirical and expert validation rather than as ground truth.
2. **Benchmarking across domains**
    - Current generative-model benchmarks focus on perceptual quality (e.g., FID for images). Risk assessment needs domain-specific metrics: e.g., regulatory capital stability, toxicological endpoint coverage, early-warning lead time.
    - Tang & Kejriwal's work on evaluating generative models on cognitive tasks illustrates how task-oriented evaluation can be designed; similar ideas are needed for risk-assessment tasks.
3. **Data governance and confidentiality**
    - Risk data often contains proprietary or sensitive information. Incorporating such data into foundation models raises privacy and governance issues extensively discussed in AI impact and governance literature.

5.2 Model-risk and systemic-risk amplification

GenAI compounds traditional model-risk factors identified in MRM frameworks.

- **Unbounded output space**: LLMs can generate arbitrary narratives, including plausible-sounding but false risk rationales.
- **Prompt-sensitivity and context leakage**: Small changes in prompts or retrieval context can yield conflicting assessments.
- **Systemic coupling**: If widely reused GenAI models and prompts drive similar stress scenarios and mitigation strategies, systemic fragility may increase, echoing concerns about AI-induced systemic risk.

Table 9. Comparative view of GenAI-induced risk vs traditional model risk

Risk type	Traditional predictive models	GenAI models (LLMs, GANs, VAEs)
Specification error	Wrong features, functional form	Wrong training regime; misaligned objectives
Data shift	Input distribution changes	Same; additionally, prompts and context windows change
Output misuse	Misinterpreting numeric scores	Over-trusting fluent text or realistic images
Systemic risk	Correlated models within a sector	Shared foundation models across sectors and domains
Governance burden	Validation, documentation, audit trails	All of the above plus tracking prompts, context, fine-tunes

5.3 Governance and human-in-the-loop design

Stahl et al. argue that AI-impact assessments must grapple with the dynamic nature of AI and the difficulty of predicting long-term sociotechnical consequences. Barrett et al. stress that high-consequence AI systems require specific catastrophic-risk analysis and red-teaming beyond generic frameworks.

For GenAI-based risk assessment, this implies:

- **Explicit separation of "assistant" vs "decider" roles**: GenAI should rarely be treated as the final arbiter of risk; instead, it should support human experts who remain accountable.
- **Traceability of generative contributions**: Risk reports must record which passages, scenarios or calculations were generated by AI, and under which prompts or models.
- **Participatory design**: Stakeholders (regulators, domain experts, affected communities) should co-design GenAI-assisted risk-assessment workflows to avoid purely technical framings of risk.

Table 10. Governance design patterns for GenAI-assisted risk assessment

Pattern	Description	Example implementation
Human-in-command	Humans retain final decision & override rights	Approval gates for GenAI-drafted reports
Dual-track assessment	Parallel human-only and GenAI-assisted workflows compared	Use GenAI as "second opinion" only
Audit-ready logging	Store prompts, model versions, and key outputs	Versioned prompt library with risk tags
Domain-specific guardrails	Restrict model capabilities in regulated areas	Prompt templates + retrieval-only constrained LLMs

5.4 Research agenda

Based on our taxonomy and analysis, we highlight several research directions:

1. **Domain-specific GenAI benchmarks for risk assessment**
    - Create open datasets and tasks (e.g., synthetic but realistic portfolios, CRA dossiers) where generative models can be evaluated on their ability to support risk decisions rather than generic generation quality.
2. **Hybrid symbolic–generative architectures**
    - Combine GenAI with formal risk models (e.g., credit scoring, toxicology dose-response) so that generative components propose scenarios or arguments that are then checked against symbolic or statistical constraints.
3. **Systemic-risk aware GenAI deployment**
    - Extend systemic-risk analyses of AI to explicitly model the effects of widely shared foundation models and prompts on correlated behaviors in finance, infrastructure, and supply chains.
4. **Evaluation of AI-for-AI risk assessment**
    - Study empirically how LLM-assisted AI impact assessments differ from human-only ones in coverage, bias, and effectiveness, building on AI-IA literature.
5. **Regulatory guidance for GenAI in risk-critical workflows**
    - Translate high-level frameworks (e.g., NIST AI RMF, ISO 31000) into concrete sector-specific guidelines for when and how GenAI can be used in regulated risk-assessment tasks, expanding on the catastrophic-risk guidance of Barrett et al.

Table 11. Priority research questions by taxonomy axis

Axis	Example research questions
Lifecycle phase	How does GenAI change the time-to-detect emerging risks?
Model family	Which generative family is most robust under distribution shift?
Risk domain	How does GenAI affect systemic-risk measures in finance vs CRA?

6. Conclusion

This paper proposes a taxonomy of generative AI applications for risk assessment, structured along risk lifecycle phases, model families, and risk domains. By integrating insights from deep generative modeling, model-risk management, systemic-risk analysis, chemical risk assessment, and AI-impact assessment literature from 2013–2023, we show how GenAI can both enhance and complicate risk-assessment practices.

Comparative analysis reveals that:

- LLMs excel in documentation, evidence synthesis and scenario narration, but pose significant risks of hallucination and over-trust.
- GANs, VAEs and related models are powerful for synthetic-data generation, anomaly detection and scenario modeling, but require rigorous evaluation to avoid misleading risk estimates.
- Domain context matters: sectors like chemical risk and safety-critical engineering adopt GenAI cautiously within strong regulatory frameworks, whereas organizational risk-documentation tasks may be more readily automated.

Ultimately, GenAI should be viewed not as a replacement for expert-driven risk assessment, but as a set of tools that, if carefully governed, can expand the reach, speed and depth of risk analysis. The proposed taxonomy and comparative tables can help researchers and practitioners map new applications, anticipate their strengths and weaknesses, and design governance structures that align GenAI capabilities with the demands of trustworthy risk assessment.

References

1. Barrett, A. M., Hendrycks, D., Newman, J., & Nonnecke, B. (2022). *Actionable guidance for high-consequence AI risk management: Towards standards addressing AI catastrophic risks*. arXiv preprint. https://doi.org/10.48550/arXiv.2206.08966
2. Danielsson, J., Macrae, R., & Uthemann, A. (2022). Artificial intelligence and systemic risk. *Journal of Banking & Finance, 140*, 106290. https://doi.org/10.1016/j.jbankfin.2021.106290
3. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems, 27*, 2672–2680. https://doi.org/10.5555/2969033.2969125
4. Iqbal, A., Sharif, M., Yasmin, M., Raza, M., & Aftab, S. (2022). Generative adversarial networks and its applications in the biomedical image segmentation: A comprehensive survey. *International Journal of Multimedia Information Retrieval, 11*(3), 333–368. https://doi.org/10.1007/s13735-022-00240-x
5. Malhotra, Y. (2018). *AI, machine learning & deep learning risk management & controls: A framework for evidence-based model risk management* (SSRN Working Paper No. 3193693). https://doi.org/10.2139/ssrn.3167035
6. Oussidi, A., & Elhassouny, A. (2018). Deep generative models: Survey. *2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*, 1–8. https://doi.org/10.1109/ISACV.2018.8354080
7. Svetlova, E. (2022). AI ethics and systemic risks in finance. *Philosophy & Technology, 35*(1). https://doi.org/10.1007/s13347-021-00499-1

8. Wittwehr, C., Blomstedt, P., Gosling, J. P., Peltola, T., Raffael, B., Richarz, A.-N., Sienkiewicz, M., Whaley, P., Worth, A., & Whelan, M. (2020). Artificial intelligence for chemical risk assessment. *Computational Toxicology, 13*, 100114. https://doi.org/10.1016/j.comtox.2019.100114