Reliable Context

RELIABLE CONTEXT / RESEARCH & PRACTICAL GUIDANCE

Model Risk for LLM Systems: SR 26-2, SR 11-7, OSFI E-23

> Reliable Context is a kernel of truth: a system of record for organizations that fuse private and public information and must defend every value they act on. One canonical record per real thing, a cited source and an observation date behind every value, unknowns declared rather than left blank, and any past state of the record reproducible.

Published by Reliable Context · 2026-10-29 · sr 11-7 · model risk management llm · osfi e-23 · sr 26-2 model risk · generative ai model risk banks

Original article: https://reliablecontext.com/articles/model-risk-management-sr-11-7-osfi-e-23-llm-systems



In this article

1. Why this question changed in 2026
2. What a model is, in each text
3. SR 11-7 concepts that remain the industry vocabulary
4. SR 26-2: what changed
5. OSFI E-23: scope, timing and data expectations
6. LLMs and RAG: what exactly is the model?
7. Data inputs as a validation object
8. Evidence package for an LLM system
9. Gaps
10. Summary
11. Sources

Why this question changed in 2026

For fifteen years, the reference text on model risk in US banking was the Federal Reserve's SR 11-7, *Guidance on Model Risk Management*, issued on 4 April 2011 and adopted in parallel by the Office of the Comptroller of the Currency (OCC). Banks deploying large language models (LLMs) spent 2023 to 2025 asking how SR 11-7 applied to them.

On 17 April 2026, the Federal Reserve, OCC and FDIC issued revised guidance under the Federal Reserve letter SR 26-2, which "supersedes and replaces SR letter 11-7" and SR 21-8 on model risk for BSA/AML systems. The revised guidance contains a sentence that changes the LLM question entirely: "Generative AI and agentic AI models are novel and rapidly evolving. As such, they are not within the scope of this guidance."

In Canada, the Office of the Superintendent of Financial Institutions (OSFI) moved in the opposite direction. Its final Guideline E-23 – Model Risk Management, released in September 2025 and effective 1 May 2027, defines a model to include "AI/ML methods".

This article sets out what each text says, then applies the concepts to LLM-based and retrieval-augmented systems. It covers the US and Canada only. It is educational, distinguishes what the guidance says from how it can be applied, and is not legal or supervisory advice.

What a model is, in each text

Text	Definition of "model" (quoted)	Generative AI
SR 11-7 (2011)	"a quantitative method, system, or approach that applies statistical, economic, financial, or mathematical theories, techniques, and assumptions to process input data into quantitative estimates"	Not addressed (predates it)

Text	Definition of "model" (quoted)	Generative AI
SR 26-2 (2026)	"a complex quantitative method, system, or approach that applies statistical, economic, or financial theories to process input data into quantitative estimates"; excludes "simple arithmetic calculations" and "deterministic rule-based processes and software" without such theories	Out of scope: "Generative AI and agentic AI models ... are not within the scope of this guidance"; the principles apply to "non-generative, non-agentic AI models"
OSFI E-23 (effective 1 May 2027)	"An application of theoretical, empirical, judgmental assumptions or statistical techniques, including AI/ML methods, which processes input data to generate results", with data input, processing and result components	In scope as AI/ML; AI defined by reference to the OECD definition, under which a system "infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions"

Sources: SR 11-7, attachment, section III; SR 26-2, section II and footnote 3; OSFI E-23, A.4 and footnote 1.

Two consequences follow. In the US, an LLM used by a bank is not governed by SR 26-2, but SR 26-2 also says that "a banking organization's risk management and governance practices should guide the determination of appropriate governance and controls for any tools, processes, or systems not covered in this document". The absence of model risk guidance is not the absence of supervisory expectation. In Canada, an LLM-based system at a federally regulated financial institution falls within E-23's definition from May 2027.

SR 11-7 concepts that remain the industry vocabulary

Even superseded, SR 11-7 supplied terms that both successor texts reuse.

- **Effective challenge.** SR 11-7 called it a "guiding principle": "critical analysis by objective, informed parties that can identify model limitations and produce appropriate changes", depending on "a combination of incentives, competence, and influence" (SR 11-7).
- **Inventory.** "Organizations should maintain an inventory of models implemented for use, under development for implementation, or recently retired."
- **Materiality.** Model risk management should be proportionate to how pervasive and impactful model use is.

SR 26-2: what changed

SR 26-2 keeps these concepts and reframes them around a risk-based, tailored approach.

Topic	SR 26-2 (April 2026)
Legal nature	"does not set forth enforceable standards or prescriptive requirements; accordingly, non-compliance with this guidance will not result in supervisory criticism", although "supervisory action may result for any violations of law or unsafe or unsound practices stemming from insufficient management of model risk"
Applicability	"expected to be most relevant to banking organizations with over $30 billion in total assets"
Risk framing	Model risk "is influenced by a model's inherent risk, exposure, purpose, and use"; purpose and exposure together determine "model materiality"

Topic	SR 26-2 (April 2026)
Effective challenge	Retained: "critical analysis conducted by objective experts", with "appropriate expertise", "sufficient independence" and "organizational standing and influence to effect any change"
Validation	Conceptual soundness, outcomes analysis, ongoing monitoring; validation "generally occurs prior to a model's first use"
Inventory	"an effective model inventory includes sufficient information to understand model risks ... at the individual and aggregate levels"
Vendors	Principles "remain applicable"; sound practice includes understanding "its conceptual soundness, design, development data, and performance"

Source: SR 26-2 and attachment.

OSFI E-23: scope, timing and data expectations

OSFI's backgrounder of 11 September 2025 says E-23 covers "traditional actuarial models as well as emerging technologies such as artificial intelligence (AI) and machine learning", "will come into effect in May 2027, following an 18-month transition period", and "will apply to all federally regulated financial institutions"; pension plans are excluded. The guideline page states the effective date as 1 May 2027.

E-23 is organised around three outcomes: model risk "well understood and managed across the enterprise", "managed using a risk-based approach", and "model governance covers the entire model lifecycle". Three provisions bear directly on LLM systems:

- **Risk rating factors** include "model complexity or level of autonomy" and "reliability of data inputs" (Principle 2.2).
- **Inventory** covering "all models in use or recently decommissioned", kept "accurate, evergreen" and updated for "model modifications and changes in model use, risk rating, or performance status" (Principle 2.1).
- **Data** used to develop a model should be accurate and fit-for-use, relevant and representative, compliant, "traceable (that is, having documented lineage and provenance)", and timely; institutions should "document provenance and use of synthetic data and proxy data" (Principle 3.2).

Model review is triggered not only by new models and scheduled reviews but by "significant data changes (for example, new data sources, altered data definitions, or shifts in data quality)" (Principle 3.4).

Source: OSFI Guideline E-23.

LLMs and RAG: what exactly is the model?

The guidance texts were written for models whose behaviour is fixed by parameters estimated from data. An LLM application has more moving parts, and deciding which of them constitute "the model" determines what gets inventoried, validated and monitored. The following decomposition is an application of the guidance, not something the guidance prescribes.

Component	Changes when	Typical owner	Treat as
Foundation model (weights)	Vendor releases a version	Vendor	Vendor model; version recorded
System prompt and instructions	Edited by the team	Bank	Part of model specification; change-controlled
Retrieval corpus	Documents added, corrected, removed	Bank, multiple owners	Data input component
Retrieval method (embedding model, ranking, filters)	Reconfigured	Bank or vendor	Processing component
Tools and agents	Added or rewired	Bank	Processing component; autonomy factor
Output post-processing	Rules change	Bank	Result component

E-23's three-component definition (data input, processing, result) maps well onto this table. The important consequence is that, in a retrieval-augmented system, the data input component is not a fixed training set. It is a corpus that changes daily, and the system's answers change with it even when no code or weights change.

Data inputs as a validation object

This is where LLM systems differ most from conventional models, and where the guidance texts, read together, point most clearly.

E-23 asks for traceable data "having documented lineage and provenance" and treats "significant data changes" as a review trigger. SR 26-2 lists "the quality of inputs for the model, and data constraints" among the factors of inherent risk. For a retrieval system, the input that matters for a given answer is the specific passage retrieved, not the corpus in general. Validation therefore has to work at two levels:

1. **Corpus level.** Is the corpus complete for the intended use, current, and free of duplicates and conflicting versions? Who may add or change documents, and is each change recorded?
2. **Answer level.** For a sample of answers, does each claim trace to a retrieved passage that supports it, and was that passage the authoritative version at the time? The site's method note on auditing an LLM answer against its sources describes this check.

A worked example

The bank in this example is fictional: a Canadian federally regulated bank with a US subsidiary. It deploys an assistant that answers relationship managers' questions about commercial credit policy, using retrieval over the credit policy manual and approved exception memos.

Step	E-23 provision applied	Evidence produced
Identify	Principle 2.1: in-use model with non-negligible risk	Inventory entry: purpose, owner, vendor model and version, corpus identifier

Step	E-23 provision applied	Evidence produced
Rate	Principle 2.2: "level of autonomy", "reliability of data inputs", customer impact	Rated medium: advisory only, but policy errors could lead to unapproved exceptions
Data	Principle 3.2: traceable, timely	Corpus manifest: each document's identifier, version, effective date and owner
Review	Principle 3.4: independent review	Reviewer tests 200 questions against a reference answer set; checks every cited passage
Monitor	Ongoing monitoring; review on "significant data changes"	Weekly re-run of the reference set; automatic review trigger when the policy manual is reissued

In month four, the reviewer finds that 6 of 200 answers cite an exception memo that was superseded by a later memo, both still in the corpus. The model did nothing wrong in the conventional sense; the input was wrong. Under E-23's framing, this is a data finding that requires remediation in the corpus (retire superseded memos, record supersession) and a control (retrieve only documents in force on the question date).

For the US subsidiary, the same assistant is outside SR 26-2's scope as a generative AI system. The bank's governance practices, which SR 26-2 says should guide controls for systems it does not cover, can reasonably apply the same inventory, rating and review approach; the evidence above serves both.

Evidence package for an LLM system

The table lists the records a reviewer, auditor or supervisor would expect to see, organised by the concepts the guidance texts share. The mapping is an application.

Record	Concept	What makes it defensible
Inventory entry with purpose, owner, components and versions	Inventory (SR 26-2; E-23 2.1)	Updated on every model, prompt or corpus change
Risk rating with factors	Materiality; risk rating (SR 26-2; E-23 2.2)	Autonomy and data reliability explicitly scored
Corpus manifest	Data quality; traceability (E-23 3.2)	Each document identified, versioned, dated, owned; superseded documents marked
Reference question set with expected answers and sources	Outcomes analysis (SR 26-2)	Versioned; answers tied to source passages
Review report	Effective challenge; independent review	Reviewer independence documented; findings and dispositions recorded
Monitoring results	Ongoing monitoring	Time series, thresholds, and review triggers on data change
Vendor documentation	Vendor products	Model version, known limitations, change notifications
Interaction logs	Use and monitoring	Each answer linked to the retrieved passages and model version

A recurring quantitative question in reviews is whether an assistant's stated confidence means anything. If the system reports confidence, it should be tested for calibration, meaning that answers labelled 90 percent confident are right about 90 percent of the time on the reference set.

Gaps

- **US scope gap.** SR 26-2 places generative and agentic AI outside model risk guidance, without replacing it with LLM-specific guidance. Banks need their own framework for these systems; SR 26-2 points to their general risk management.
- **Non-deterministic outputs.** Both texts assume a model produces consistent outputs for the same inputs. Outcomes analysis for LLMs needs repeated sampling and acceptance thresholds, which neither text specifies.
- **Corpus as model input.** Neither text addresses retrieval corpora explicitly. E-23's data traceability principle and review triggers are the closest fit.
- **Vendor opacity.** SR 26-2 acknowledges that banks "may not receive from the vendor the underlying code, data, or methodology". For foundation models, this is the norm, which shifts weight to outcomes analysis and monitoring.

Summary

As of 28 October 2026, US banking model risk guidance is SR 26-2, which replaced SR 11-7 and explicitly excludes generative and agentic AI. Canadian guidance is OSFI E-23, which includes AI/ML and applies from 1 May 2027. Across both, the durable concepts are an inventory, risk-based rating, independent challenge, validation, and monitoring. For LLM systems, the hardest of these to evidence is data: which passages each answer relied on, and whether they were the authoritative versions at the time. Reliable Context keeps organisational facts with that kind of version and source record.

Sources

1. Board of Governors of the Federal Reserve System, SR 11-7, *Guidance on Model Risk Management*, 4 April 2011 (cover letter). https://www.federalreserve.gov/boarddocs/srletters/2011/sr1107.pdf
2. Board of Governors of the Federal Reserve System and Office of the Comptroller of the Currency, *Supervisory Guidance on Model Risk Management*, SR 11-7 attachment, 4 April 2011. https://www.federalreserve.gov/boarddocs/srletters/2011/sr1107a1.pdf
3. Board of Governors of the Federal Reserve System, SR 26-2, *Revised Guidance on Model Risk Management*, 17 April 2026, with interagency attachment. https://www.federalreserve.gov/supervisionreg/srletters/SR2602.pdf
4. Office of the Superintendent of Financial Institutions, *Guideline E-23 – Model Risk Management (2027)*. https://www.osfi-bsif.gc.ca/en/guidance/guidance-library/guideline-e-23-model-risk-management-2027
5. Office of the Superintendent of Financial Institutions, *Backgrounder: Guideline E-23 – Model Risk Management*, 11 September 2025. https://www.osfi-bsif.gc.ca/en/news/backgrounder-guideline-e-23-model-risk-management

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. https://www.federalreserve.gov/boarddocs/srletters/2011/sr1107.pdf
2. https://www.federalreserve.gov/supervisionreg/srletters/SR2602.pdf
3. https://www.osfi-bsif.gc.ca/en/guidance/guidance-library/guideline-e-23-model-risk-management-2027
4. https://www.federalreserve.gov/boarddocs/srletters/2011/sr1107a1.pdf
5. https://www.osfi-bsif.gc.ca/en/news/backgrounder-guideline-e-23-model-risk-management

THE PUBLISHER

About Reliable Context

Reliable Context is a system of record for organizations that fuse private and public information and must defend every value they act on.

The guarantees

One canonical record per real thing: duplicates resolve by rule, ambiguity is shown rather than hidden, and related organizations relate without ever being merged. A verbatim, cited source and an observation date sit behind every value. Unknowns are declared as not found, blocked or not applicable with a reason, never left as an empty cell. Any past state of the record is reproducible, together with what changed since.

Where it runs

It runs inside the customer's perimeter, one database per organization. Permitted sources, model routing, reader clearance and retention are set as the customer's own policies, and every value is exportable with its chain of custody for an auditor or regulator.

What it is not

It is not a vector database or a RAG framework, not a data catalog, MDM or data-quality suite, not a CRM or data-enrichment vendor, and not a chatbot or AI assistant. Where a reader needs one of those, the honest answer is to say so and point them there.

Visit

See Reliable Context.

Reliable Context is in production use in two sectors; more than six thousand of its identity decisions were made by rule and none has been reversed. Articles published here explain general practice, cite their primary and academic sources, and name the jurisdiction of any regulatory rule they describe. They are not legal, regulatory or compliance advice for a specific organization, and requirements change.

Website: https://reliablecontext.com

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.