# Letter

| | |
|---|---|
| **Title** | Guideline E-23 – Model Risk Management (2027) - Letter |
| **Category** | Sound Business and Financial Practices |
| **Date** | September 11, 2025 |
| **Sector** | Banks<br>Foreign Bank Branches<br>Life Insurance and Fraternal Companies<br>Property and Casualty Companies<br>Trust and Loan Companies |

Today, we are publishing Guideline E-23 – Model Risk Management, which will take effect for all Federally Regulated Financial Institutions (FRFIs) on May 1, 2027.

We published a draft revised Guideline E-23 on November 20, 2023, for public consultation until March 22, 2024. The revised draft included several changes relative to the 2017 Guideline E-23 – Enterprise-Wide Model Risk Management for Deposit-Taking Institutions. While many of the draft revisions remain as part of the final guideline, several updates have also been incorporated in response to feedback provided by stakeholders. For details, please refer to the summary of stakeholder comments and our responses.

Since the release of the draft guideline, we continue to observe growth in the use and complexity of models at FRFIs. As such, the expansion of the guideline's scope of application to all models at all FRFIs remains in place as a key revision. However, application of the guideline to Federally Regulated Pension Plans (FRPPs) has been reconsidered given differences in our mandate to supervise FRPPs and the availability of alternative industry guidance addressing pension risk management. Other important revisions provide added context and clarity for Artificial Intelligence and Machine Learning (AI/ML) model risk management.

Should you have any questions, please contact Christopher Hutny (christopher.hutny@osfi-bsif.gc.ca), Senior specialist, Banking Capital and Liquidity Standards Division, Risk Advisory Hub.

Sincerely,

Amar Munipalle

Executive Director, Risk Advisory Hub

## Summary of stakeholder comments and our responses

### Model definition and scope

| Stakeholder feedback | Our response |
|---|---|
| Stakeholders noted that the definition of a model is too broad and should be narrowed. | We have left the definition intentionally broad in the final guideline. Institutions are expected to identify models and manage model risk commensurate with the model risk rating. |
| Stakeholders asked for clarity on whether low/non-risk use cases built from a generative AI model such as the use of ChatGPT for document summarization or marketing emails can be excluded from the requirements for managing critical/high risk models. | We have enhanced certain portions of the final guideline to provide some additional clarity on how it should be applied to AI/ML models. Still, institutions are empowered to make risk-intelligent decisions, based on the application of the model within the organizational context, when establishing model risk ratings. |
| Stakeholders asked if all models will be in scope, regardless of the risk rating. | We have expanded guidance around model identification, including expectations for an institution's model inventory. Only models that carry risk to the institution need to be captured on the model inventory and are subject to full model lifecycle governance under the final guideline. |
| Stakeholders asked how far downstream should risks be considered. | We have simplified our principles in the final guideline to remove the explicit reference to downstream risks. Institutions should determine what constitutes a reasonable amount of downstream impact based on the risk rating of the model. |
| Stakeholders requested clarity on what is considered a new model. | We have updated the final guideline to no longer distinguish "new" models. |
| Stakeholders requested clarity on the term "model risk drivers", as it is not a defined term in the draft guideline and there are no defined drivers of model risk. | We have updated the final guideline to enhance clarity, and removed the reference to "model risk drivers". |
| Stakeholders requested clarity on the term "risk classification" which is not defined and asked whether the reference should be to the risk rating. | We have updated the final guideline to enhance clarity, and removed the reference to "risk classification". |

## Implementation date

| Stakeholder feedback | Our response |
|---|---|
| Stakeholders noted that the proposed twelve-month implementation period was too short. They requested extending the implementation period to three years. | We have revised the implementation date to May 1, 2027. Many institutions signaled that they have already begun work on the model risk management reforms outlined in the draft guideline. |

## Proportionality

| Stakeholder feedback | Our response |
|---|---|
| Stakeholders requested clarity on how proportionality should be considered. | We have revised the application of the final guideline to clarify that it applies on a risk-basis, proportional to the institution's size, strategy, risk profile, nature, scope, and complexity of operations, and interconnectedness. |

## Artificial intelligence and machine learning

| Stakeholder feedback | Our response |
|---|---|
| Stakeholders requested more guidance on explainability. | We have incorporated additional explainability guidance in the final guideline throughout the components of the model lifecycle. |
| Stakeholders requested more guidance on self-learning models, specifically with respect to how the model modification approval requirement can be satisfied for models that are always changing. | We have incorporated additional guidance in the final guideline around model modifications under model identification, model review, and model monitoring. Institutions should establish internal criteria to determine when a self-learning model has materially changed. |

## Third-party vendors

| Stakeholder feedback | Our response |
|---|---|
| Stakeholders requested more specific guidance on expectations for managing third-party black-box models. | We have added guidance in the final guideline on the management of third-party models. Institutions should comply with third-party risk management principles established under guideline B-10 Third-Party Risk Management. Institutions should also ensure that third-party models receive validation and monitoring commensurate to the model risk. |
| Stakeholders requested a grace period within the guideline, allowing institutions to validate third-party model updates after they have been deployed, considering that institutions may have limited control over third-party models. | We have not incorporated a grace period for third-party model risk management in the final guideline. However, institutions may still establish an exceptions policy under which models can be used for limited and specific applications prior to the completion of validation. |
| Stakeholders requested clarity around fourth-party model risk management. For example, if an institution contracts with a third party that integrates generative AI functionality derived from a fourth party model, stakeholder would like clarity on whether this fourth-party model should be considered within the institution's model risk framework. | We did not make any specific revisions to the final guideline to account for fourth party models. Institutions have a responsibility to ensure that third-party model use is within the institution's risk appetite limit. A third-party model review should include inputs into the model, including feeder models. It is up to the institution to determine whether the risks involved with using a third-party model are appropriate for the business context. |
| Stakeholders requested clarity with respect to vendors unwilling to share the appropriate level of documentation. | We did not make any specific revisions to the final guideline to account for, or accommodate, unwilling vendors. It is up to each institution to ensure adequate documentation is provided by vendors to satisfy model risk management needs. |
| Stakeholders requested clarity around the governance of third-party "black box" models. | We updated the final guideline to reference "black box" models in some areas but we did not make any specific revisions or exceptions relating to governance of third-party "black box" models. We expect all institutions to have a framework in place for the governance of all third-party models. Institutions should decide for themselves how to structure the third-party model governance in a way that reflects the institution's risk appetite. |

| Stakeholder feedback | Our response |
| --- | --- |
| Stakeholders noted that validating certain models is challenging when advanced vendor products with proprietary techniques and big/dynamic external training data are used. | We did not make any specific revisions to the final guideline to account for vendors with proprietary techniques or data. Institutions should have controls to assess whether the vendor solution is producing an appropriate response. |
| Stakeholders noted that organizations using third-party libraries, platforms, and/or automated development processes used for model development, should be subject to independent review/validation, commensurate with the risks these elements present. | We agree with this observation and have included it in the final guideline. |
| Stakeholders noted that some model governance requirements may be difficult to meet for third-party models (for example, data requirements, development environment, and model architecture). They suggested that development requirements be reduced, and model review requirements increased. | We did not make any specific revisions to the final guideline to reduce expectations regarding third-party models. The intention is for institutions to establish policies for reviewing and monitoring third-party models in a way that effectively manages its model risk. |

## Validation

| Stakeholder feedback | OSFI response |
|---|---|
| Stakeholders requested clarity on any distinction between the terms “validation” and “review”. | We use the term “validation” in the final guideline but note that it would be interchangeable with the term “review”. The intent is for models to have independent reviews commensurate with the model risk rating. |
| Stakeholders noted that for some institutions, the number of models could number in the hundreds and validation of all would be onerous. They asked for clarity on an appropriate model review frequency. | We did not make any revision in the final guideline to specify the frequency of model reviews. As noted in our draft guideline, the model risk rating should drive the frequency, intensity, and scope of model review and redevelopment. |
| Stakeholders asked if targeted validation would be acceptable, rather than validation of the whole model. | We did not make any specific revisions in the final guideline to accommodate more targeted or reduced-scale validations. However, we note that the scope of model review should be driven by the model risk rating. |
| Stakeholders requested clarity around expectations for generative artificial intelligence models compared to traditional quantitative models. | We have added some special considerations for artificial intelligence models throughout the final guideline. However, the outcomes and principles provided in the guideline do not vary based on the algorithmic approach to modeling. |
| Stakeholders asked if validation must be performed by home office, or if it could be provided by other stakeholders. | We revised the final guideline to remove references to a “home office”. Institutions should specify which units or individuals are responsible for model reviews in the institution’s model risk management framework. |

## Data lineage

| Stakeholder feedback | Our response |
|---|---|
| Stakeholders noted that with many data systems in an institution, data lineage across all will be challenging. Stakeholders also requested clarity on the degree of traceability needed for data lineage. | We did not make any specific revisions in the final guideline regarding traceability of model data. As good practice, institutions should understand how the data that is used for modeling is captured and the transformations that are applied on the data prior to modeling. We expect that the rigour involved in defining data lineage will be commensurate with the model risk rating and the relative importance of data elements within the model. |
| Stakeholders asked if data lineage should be assessed separately from model risk. | We did not make any specific revisions in the final guideline regarding traceability of model data as this is not within the scope of this guideline which is focused on model risk management. Data, including data accuracy, lineage, representation, and timeliness, should be assessed as a factor of model risk. |

## Model lifecycle

| Stakeholder feedback | Our response |
|---|---|
| Stakeholders requested clarity on governance and maintenance expectations for models that are not in use. They asked whether requirements for parallel runs would be more appropriate. | We did not make any specific revisions in the final guideline regarding governance expectations for decommissioned models. We agree that parallel runs are one effective approach to support the transition from one model to its replacement. As noted in the final guideline, we expect institutions to retain the decommissioned model and documentation for a defined period so it can serve as a benchmark or contingency fallback. |
| Stakeholders requested clarity regarding how "Approval requirements should apply throughout a model's lifecycle, including for modifications and periodic reviews." They asked for clarity on whether institutions should perform one comprehensive model approval after validation and before deployment or whether the term "throughout" is used to indicate that approvals are needed at all key stages of model development, including independent review before approval. | We have retained language in the final guideline on approvals throughout the model's lifecycle. Still, institutions have the flexibility to design their own model governance process and determine at which stages approvals are needed. To enhance clarity, approvals should be obtained prior to the implementation of a model change and following periodic reviews to ensure the model remains fit for purpose. There could also be other situations that merit an approval within the model lifecycle, for example, when there is a system change that was not driven by the model but there is a need to ensure the model is still working properly. |
| Stakeholders noted that identification of all stakeholders is a high threshold and that it may be difficult to ascertain all stakeholders at the beginning, when defining the rationale for modeling. | We have amended the language in the final guideline by removing "all stakeholders" and replacing it with "model stakeholders", a term defined in the list of key terms. |
| Stakeholders noted that many models are not in production for continual usage but are just used to produce a single output or on a periodic basis. They recommended that the Model Deployment section be exclusively applicable to "high availability" models. | We did not make any revisions in the final guideline to limit application to single-output or periodic-use models. Institutions should pay special attention to models that are used outside of a production environment. The risk rating of the model may be higher than originally expected, as controls that are implemented in a production environment may no longer act as a risk mitigant. |
| Stakeholders noted that the model lifecycle does not allow for dynamic recalibration or other differentiations in model management. | We have removed references in the final guideline to the circular model lifecycle to support innovative development paths. |

## Other

| Stakeholder feedback | Our response |
|---|---|
| Stakeholders recommended a new section be added on related guidance, and expectations on how this guideline interacts with other guidance. | Although there is no specific section on related guidance, we have included references and links in the final guideline to other related guidelines. |
| Stakeholders requested clarity on the level of consistency and/or replication needed between data used in development and the production data set. | We did not make any revisions in the final guideline to establish measures or thresholds for data consistency between development and production environments. Reproducibility between development and production data should be assessed during model review. Institutions should have differing expectations of what is acceptable depending on the risk rating of the model. |
| Stakeholders requested clarity on whether model developers should be responsible for coding within production systems. | We did not specify responsibilities for coding within a production system in the final guideline. Effective model deployment should involve collaboration among model developers, model owners, model users, and enterprise technology or operations teams. |
| Stakeholders requested clarity on the intention behind having specific, standalone model risk ratings for the models supplied to subsidiaries / foreign branches. | We updated the language in the final guideline to clarify that all externally developed models should be assessed for model risk ratings on a standalone basis. The model risk rating assigned to models developed by an external party (that is, parent organization) may not reflect the risk to the subsidiary or foreign branch. |