Managing Artificial Intelligence Risk in Small to Mid-Size Banks: A Practical Governance Playbook

July 31st, 2025

Authors and contributors:

Thomas Dahlin, (Lead), Chief Model Risk Officer, Centennial Bank
Jill Murphy, Risk Governance Manager, Mechanics Bank
Dr. Raman Mandapaka, Advisory Practice Lead, Cadmus Group
Dr. Joe Breeden, President, Deep Future Analytics
Chris Smigielski, Director of Model Risk Management, Arvest Bank
Ivo Steijn, First Vice President of Model Risk Management, East West Bank

Table of Contents

1. Overview

As banks explore the purchase, development, and use of AI systems to assist in everyday tasks, they quickly discover that these powerful tools come at a price of increased risk. Some risks, such as sustainability, are well understood and can be managed through established practices for management and control, such as training, access controls, and change management. However, there are various other risks, such as the **use of alternate data, developer or data bias, vendor risk, or data privacy**, which can sometimes prove to be more difficult to measure, no less difficult to manage. Their emergence over time, or through an event, can be punitive – such as a Fair Lending violation, loss of customer personal information, equal employment violation, misstatements in financial reporting, or other breaches in acceptable banking activities, leading to regulatory fines, forcible containment of strategic plans, and/or legal actions. As such, a deliberate, systematic effort to identify and control these risks may prove sensible for the organization. This may be achieved with the presence of a well-functioning governance, critical to ensure these risks are captured and managed appropriately, and specific to AI.

2. Choose an AI Governance Strategy

The nature of AI is based on the continuous learning process of the models or algorithms, such that the objective function can be satisfied through an ever-increasing amount of available data and structural malleability, enabling the models to change and project. **How and what data is procured, model methodologies, and performance, are often unknown, and not typically provided to satisfaction.** Data may be synthetic, social media-oriented, procured from sources unknown, and inserted into models for desired outcomes. Certainly, vendors of AI models will focus on the functionality and use of the system, less so on the rigor of evidentiary requirements, as that is difficult, time-consuming, and potentially forever changing. Nonetheless, the data use and applicability, the performance testing, and the theory adopted, are some of the key evidences requiring review before any conclusions can be made on model soundness, reliability of results, and the existence of underlying model risks.

The opaqueness of the AI models can be problematic when assessed through traditional practices. AI's ability and speed to consume new forms of data and methods outpaces the vendor's ability to fully capture developmental evidence and present it to a customer for review and consideration, such as Model Risk Management. From a business perspective, **an continuously changing AI model requiring continuous evaluation is both time and cost prohibitive**. This leaves banks with three options when considering their use:

1. **Zero tolerance policy** - strictly prohibiting their implementation and use. This may be a compelling option for risk avoidance, however, an organization's competitiveness and process improvements will likely be hindered.

2. **Open Door policy** – allowing full use of AI models/systems, conducting model validations with any available information, and providing point-in-time feedback, or issues, to the model owners and developers. This is traditionally in alignment with SR- 11-7 practices. Efficiencies and business activities may be realized; however, an ever-changing model or system with limited support and evidences will likely yield a variety of risks.

3. **Selective Constraints and Control policy** - Or, alternatively, allowing their use, but with various constraints and controls; benefiting from some of the lift and insight that they provide.

In line with the fundamental risk management principles of risk and reward, the adoption of the third option, "Selective Constraints and Control Policy", is a sensible direction to pursue for small to mid-size banks (under $50 billion). Both "Open Door" and "Zero tolerance" policies fail to balance risk management practices in a competitive landscape. The technology becomes difficult to avoid, and will likely seep into an organization in unknown places if accommodations are not formally established. Conversely, absent a well-controlled environment, haphazard use will certainly incur consequences.

The use of AI systems is also unlikely to be limited only to internal use. Consideration of **control requirements must also be set with vendors that have access to the organization's data or systems**. These vendors that have AI are also susceptible to the same vulnerabilities as the lenders.

This paper is a set of recommendations on how to arrange a governance structure that adopts a policy of measured use, selective constraints, and effective controls. Sensible use of AI models and systems requires well-coordinated decisions with management and agencies most responsible for the risks associated with these systems – such as Model Risk Management, Information Security, Data Governance, Third Party Governance, Compliance, and ultimately Senior Management. This should allow for a consistent decision-making process, with a clear understanding of the level and type of risk that exists at the individual AI system level, as well as at an aggregate level, summarizing across all.

Recommendation 1: Recognize the need to establish a bank practice specifically for AI systems. These are unlike other tools and models in the bank, and they come with both known and less known risks. A measured roll-out may be most prudent to assess weaknesses, gain comfort, and take any necessary corrective actions. A heightened co-ordination is needed between risk managers and AI users. Management should understand and gain comfort in the strategy and use of these systems, to ensure clarity for the Board and regulators.

3. Establish an AI Policy and Governance Framework

An AI policy, like other important policies, will likely change and mature over time. Gaps will surface as the technology and its implications are better understood. It will be a reference point on how the organization intends to adopt an AI practice. The **organization should not delay in the creation of this policy**. The absence of such, even if imperfect, can lead to confusion and unnecessary exposure.

Initially, a higher-order policy may be created to address the need to set boundaries and controls. Any "off-limits" models, systems, and practices should be clearly articulated. Eventually, a supporting Framework should be created; detailing how the Policy will be achieved – with Roles/Responsibilities, workflows, testing, and outcomes. As such, the following recommendations and elements may be useful for policy inclusion:

1. **Owner of the Policy** – The owner should be maintain an enterprise or corporate level position, tasked with oversight of a specific risk that may affected by an AI system– such as Data Governance, Information Security, or Model Risk Management.

2. **AI definitions, categorization, and identification –** a description of AI model types should be presented for clarification and instruction.

    a. Supervised Learning – Models are trained on labeled data with known outcomes. Typical examples include regression, decision trees, and Bayesian classification models.

    b. Unsupervised Learning – Models are trained on unlabeled data to discover patterns or groupings. Examples include clustering algorithms, dimensionality reduction, and density estimation.

    c. Reinforcement Leaning – Model learn through an interactive system of rewards and penalties. Examples include dynamics outcome optimizations such as cost minimization or profit maximization in sequential decision-making contexts.

    d. Generative AI – Models that learn to produce new data similar to their training data. These may generate text, images, audio, or other synthetic outputs.

e. Large Language Models – A subset of Generative AI focused on understanding and generating language, trained on a large quantity of text data. These can perform tasks such as summarization, translation, question answering, and content generation.

f. Retrieval Augmented Generation – A subset of LLMs designed to retrieve internal documents or sections thereof as the answer to user inquiries. RAG models are most commonly used to assist staff in determining approved policies and procedures, often in a customer support context.

The purpose of defining and categorizing AI systems is to clarify what will, and will not, be permissible use. For example, a first-generation AI policy in an organization may wish to allow Generative AI only. When, and if, a desire to expand to a more sophisticated application or decision making system occurs, the expectations for controls should also expand similarly.

3. **Roles and Responsibilities** – The risk of AI systems is shared by numerous people and groups throughout the organization. To manage these risks, there are roles and responsibilities that require clarification and agreement. As follows:

   a. **Policy owner** - is responsible for the creation and annual updates to the AI policy, which describes what types of AI may be used in the organization, along with specific controls related to data, access, use, as well the level of control effectiveness. The policy owner should regularly report to an AI Committee and Executive Management with regard to the inventory, purpose, cited risks (both individually and in aggregate), any recommended remediation and activity, and the review and challenge process leading to the approval/rejection of all the AI systems. The policy owner should set the requirements of AI use, training, and procedures necessary to determine if the system, and the controls, conform to the AI policy.

   b. **AI user (from the business or centralized role)** – the following actions and responsibilities should be included:

      i. **Obtain approval** from the Policy owner and AI Committee prior to the acquisition/purchase of the AI model/system.

      ii. **Secure complete technical documentation** from the vendor that describes methodology, user guide, implementation guidelines, access, storage, and data requirements. As a resource, the *MRMIA Vendor Documentation Standard* is a useful set of 25 industry-established questions that may be used as baseline documentation – to be used at a later point for either model validation, or non-model assessments.

      iii. **Conduct ongoing "use" monitoring of the AI system -** demonstrate the ongoing use and degree that final decisions are based. The type of monitoring may be determined on a case-by-case basis, approved by an AI Committee, or chosen by another AI governing authority. For example, the monitoring may include 12 consecutive static snapshots of use, from 1/2025-12/2025, the AI system was used X number of times for each of the months – Jan, Feb, Mar,…Dec. These results should be interpreted, and any follow-up actions described. In this example, the intention is to show how much the AI system is being used for the business objective. This is not model performance monitoring, but rather a method to assess the dependency and reliance of the system.

      iv. The AI user is also responsible for **ensuring data privacy, access, and user controls are well established and operating effectively**. These should likewise

be provided to the AI Committee on an intermittent basis to ensure no degradation or events, have occurred.

v. **Sustainability** – the ability of employees to operate/execute an AI system without disruption of use, due to the workplace environment or other conditions (i.e. workplace practices, physical conditions, key man risk, psychosocial hazards – stress related). The user must demonstrate that sustainability risk is low, or that there are mitigation plans to reduce the risk. This should include operational procedures (of the AI system), multiple trained users, and instruction material.

c. **Third Party Governance** – this group is responsible for vendor due diligence. As AI becomes more embedded in organizations, this group should conduct **enhanced reviews** of vendors that develop or utilize AI. These reviews should assess the vendor's data management practices (including how and where they access client data), the types of data involved (such as personal non-public information), where the data will be stored, and who has access to it.

d. **Information Security** – this group is responsible for the protection and access of the organization's data and systems. This includes customer data and other confidential information – both internal and external. This group should define and set the information security controls to be used at the centralized level, the user level, or business level. Threshold levels for effectiveness should also be defined and reviewed.

e. **Data Governance** – this group is largely responsible for driving culture and methods for complete and accurate data, critical for the decisions made in the organization. Additionally, it should ensure that only those with a legitimate business need have access to the data.

f. **Model Risk Management** has the responsibility to categorize each AI system as either a model or non-model (such as a tool). Materiality, complexity, and use will drive that determination, and the treatment will follow in accordance with the MRM framework and policy. The categorization will allow for an inventory of both AI models and AI non-models. These inventories should be made available to Data Governance and Information Security to best ensure the presence of controls and testing.

g. **Audit** – is responsible for the adherence of policies. As such, Audit may/should examine and determine the existence of all AI requirements, i.e., procedures, monitoring, review, training, control testing, technical documentation, enhanced vendor review, as well as regular AI Committee review, challenge and/or approval of systems.

h. **AI Committee/Executive Risk Committee/Board of Directors** – a designated governance body, which may include delegated authority from the Board, should be tasked with oversight and approval authority of all AI systems and the associated risks (individual and in aggregate). This committee should review, challenge, and approve all AI systems, use/purpose, and set the requirements for control effectiveness and scope. This committee may be chaired by either/both of the Data Governance or Information Security functions, as these represent key areas of AI risk.

4. **<u>Permissible Use</u> –** AI systems that will be regarded as permissible should satisfy ALL of the following:

a. A**ssistant-based –** information retrieval, narrative or word content creation, and/or report generation may be acceptable; however, a human interface to review and accept the final output must occur prior to any acceptance and use.

b. **Effective Controls** – the requestor of an AI system should be able to demonstrate that all controls intended to prevent improper data sharing or system access are effective, and tested periodically.

c. **Immaterial Use** – If an AI system is determined to be immaterial to a business function, does not use confidential or sensitive data, and is accessible only internally, the AI Committee may consider and approve an exception for its use.

5. **Not Permissible for Use** – In the early stages of AI system use, the organization may adopt a conservative position that approaches risk avoidance. Until greater certainty can be achieved on a number of various cautionary aspects, such as enhanced AI vendor review, data privacy controls, access controls, and model risk controls, the **following AI activities and/or systems may be regarded as NOT permissible for use**:

a. **Compliance-related reporting and/or analysis** (e.g., Fair Lending, CRA, Red Lining, Pricing, Steering) for final regulatory review and submission.

b. **Account Management -** not to be used for a final customer decision – such as credit/loan approval, loan terms, pricing, special treatment, re-structure, and/or re-age.

c. **Hiring Selection Process or HR treatment decisions** – these screening systems may inadvertently lead to violations in employment hiring practices or result in unfair treatment and discrimination against applicants and employees.

d. **Vendor Selection Process** – an AI system may select vendors based on criteria that are in conflict with current strategy and direction. Low-cost providers that have a problematic history of providing goods/services may not be well identified; leading to improper vendor selections.

e. **Home/CRE/Auto/Land Valuation system –** estimates on these assets may have legal ramifications if/when majority-minority regions contain data distortions and aberrations.

f. **Company Financial report generation** - 10K, 10Q, Income, B/S reports for audit and final submission. These reports are subject to strict regulatory and accounting standards. The use of AI to generate them may introduce inaccuracies, omit required disclosures, or misrepresent financial data, potentially resulting in compliance violations or reputational harm.

g. **Vendor Data Access** – all vendors with permission and/or access to the bank's confidential data must have an NDA in place, and regular verifiable evidence that the data is neither used nor shared with any other parties or individuals, with requisite information security controls.

h. **Confidential Data sharing** – NOT to be used with customer, applicant, or organization confidential, or non-publicly disclosed data.

6. **Policy Adherence Evidence -** be provided with periodic updates by the policy owner regarding the organization's adherence to the policy, trends and use of models (such as more and more management over-rides to the output), outcomes, and effectiveness of controls specific to AI.

Recommendation 2: Create and approve a bank-wide policy on AI use. It should include how and where acceptable practices and permissible uses will be established, a workflow to review and implement, roles and responsibilities of the key stakeholders, key controls, and the specific governance structure that will review and challenge AI systems.

4. Set Control Standards for Responsible AI Use

The presence of controls for AI is not unlike the controls for other risks at the bank. AI stakeholders are responsible for identifying the risks under their responsibility and determining whether existing controls will be sufficient or if new controls need to be established. Further, both the control owners and the AI Committee should assess whether each control has met a required level of effectiveness. The mere presence of a control will not suffice. Controls must be effective, or mutually agreed-upon conditions or restrictions must be put in place until appropriate mitigation can occur. The following controls are recommended for an initial implementation:

1. **Bank-wide AI training** – this control may be created by either Model Risk Management, Information Security, or Data Governance. It should include an overview of AI concepts, practical applications, risks, and key requirements for the business units, or corporate departments that desire to purchase and use such an AI system. All employees should understand the AI policy and specific actions that they are required to follow. Evidence will be expected from each AI user/owner that they are in adherence to the policy.

   Effective – a 100% event rate and pass rate from bank employees, should be achieved in order for this control to be considered "effective". (i.e., everyone takes an AI overview course, and everyone passes a knowledge test).

2. **AI system identification and Third Party Governance (Vendor Management)** – All AI systems in the bank should be identified and requisite evidence should be collected and provided to MRM. AI users should pre-emptively bring this to the attention of Third Party Governance to ensure property due diligence. This required documentation evidence should include purpose, data management, methodology, system accessibility, and change management practice. Third Party Governance should have approved written procedures on how this control will operate and how adherence may be verified.

   Effective – a 100% event rate, such that no AI system is implemented without explicit knowledge and due diligence of Third Party Governance. Further, complete and accurate documentation, as specified by MRM is necessary for this control to be considered "effective".

3. **MRM assessment** – upon initial review of the AI system, following notification from Third Party Governance, MRM should conduct a review to determine if the AI system is either a model or a tool. Typically, complexity and materiality of use will influence this determination. MRM will place the AI system in the appropriate inventory: model inventory or tool inventory.

   Effective – all AI systems that are successfully reviewed by MRM and placed into an available inventory with a 100% event rate will be considered as an effective control.

4. **MRM Model Validation –** when/if MRM has made a determination that the AI system is a model, a model validation should be conducted. Consistent with the regulatory guidance of SR-11-7, a full scope validation must be conducted on model development, data, implementation, governance, and

documentation. If timing does not allow for an immediate model validation, an exception to policy may be granted at the Model Risk Committee and conveyed to the AI Committee. The key artifact for the control is the completion of a model validation.

Effective – in order for this control to be effective, all AI models should be successfully validated and receive a rating as either "passed" or "conditionally passed". The combination of these two requirements is necessary–the model being validated and passing.

5. **Executive Sponsor Attestation –** On an annual basis, the functional Executive Sponsor of tools and models is required to attest to the known inventory collected by MRM. A list may be provided to the Executive with known inventory, allowing time for reflection and consideration. This is an intentional redundancy to Control #2 (as listed above).

   Effective – in order for this control to be effective, evidence should be captured and available to demonstrate that the attestation occurred.

6. **Data Control(s)** – there should be multiple data controls in place to ensure sound and secure data management practices are operational throughout the bank. A primary objective is to protect against unauthorized use or access to bank and customer data, particularly non-public personal information and other confidential-designated data. Without proper controls, this information could be misused by vendors to train or enhance their own tools, models, or algorithms—activities that are not authorized and may violate privacy obligations, regulatory expectations, and be potentially damaging to the bank.

   Effective – in order for this control to be effective, data requirements of AI systems should be reviewed by Data Governance, as well as periodically reviewed to ensure that no confidential information is subsequently used or required.

7. **Information Security Controls - AI models/tools** – an objective for Information Security is to set controls such that access to AI systems and data is authorized and appropriate. Information Security should establish a range of requirements, including ongoing monitoring, data usage, roles-based access to data and the AI systems, data masking, firewalls, and data anonymization. Periodic testing and reporting of satisfactory adherence to these requirements can help reduce the risks that AI systems may pose.

   Effective – Access controls for data and AI systems are considered effective when they have been established, are functioning as intended, and are regularly enforced through review and validation by Information Security.

8. **Incident Response Control** - The bank should establish procedures and controls to respond promptly to data breaches or unauthorized access involving AI systems. Remedies should include immediate termination of access and data flow, along with a structured response plan to assess the scope of the incident and necessary actions to notify and engage relevant parties – including regulators, affected customers, employees, and/or vendors.

   Effective – These controls may be deemed effective when procedures and testing of the procedures occur. The Business Resilience team may be responsible for this testing as a separate and independent group.

9. **Review and Challenge** – As part of sound risk management for an AI system, sufficient review and challenge are necessary for both initial implementation and authorized use, as well as ongoing monitoring of the performance and controls. The review and challenge should occur at a Management Committee level with appropriate stature and membership. This should be described in the AI policy under Roles and Responsibilities. Questions that may be posed as part of this remit may include the following:

a. Have all users taken and passed the bank training course on the use of AI? Is staff use of AI being monitored for compliance with bank policy?

b. Is the purpose of the requested AI system consistent with the defined permissible uses, as per AI policy?

c. Can the AI system make any type of decision independent of the user? How will this be monitored?

d. Has all necessary technical documentation been provided to MRM, as set forth and required by MRM?

e. What were the results, and/or issues, from the MRM review, or MRM Model Validation?

f. What plans or mitigation actions are expected to address any issues that resulted from either MRM, Information Security, or Data Governance?

g. Does the vendor have any access to bank data? If so, is it confidential, non-publicly disclosed?

h. Has Third Party Governance conducted a full scope due diligence with satisfactory results? Including resolution recovery actions when/if any unauthorized breaches occur?

i. Have all Data Governance and Information Security controls been assessed as present, operational, and deemed as "effective" by the functional heads?

j. What is the benefit that this AI system will provide to the bank? – cost/time savings, positive customer engagement, business expansion, or other?

10. **Audit Review** – the role of Audit is to provide independent assurance that internal practices and procedures are conducted in accordance with approved policies, as well as applicable regulatory and industry standards. Since AI standards are still emerging throughout the industry, and regulators have not yet provided formal guidelines, perhaps due to measured observance of practices, Audit should assess adherence to the bank's stated AI governance framework, including its policies, controls, and oversight structures. Key activities, and inquiries, may include the following:

    a. Has an AI policy been created, and approved with requirements, roles and responsibilities, permissible systems, and identified controls?

    b. Is there evidence that each AI stakeholder has conducted their requirements that are detailed in the policy (i.e., Third Party Governance Due Diligence, Vendor technical documentation, MRM reviews and Validation, Information Security and Data Governance standard requirements, etc.)?

    c. Have all MRM, Information Security, and Data Governance controls been established with satisfactory outcomes?

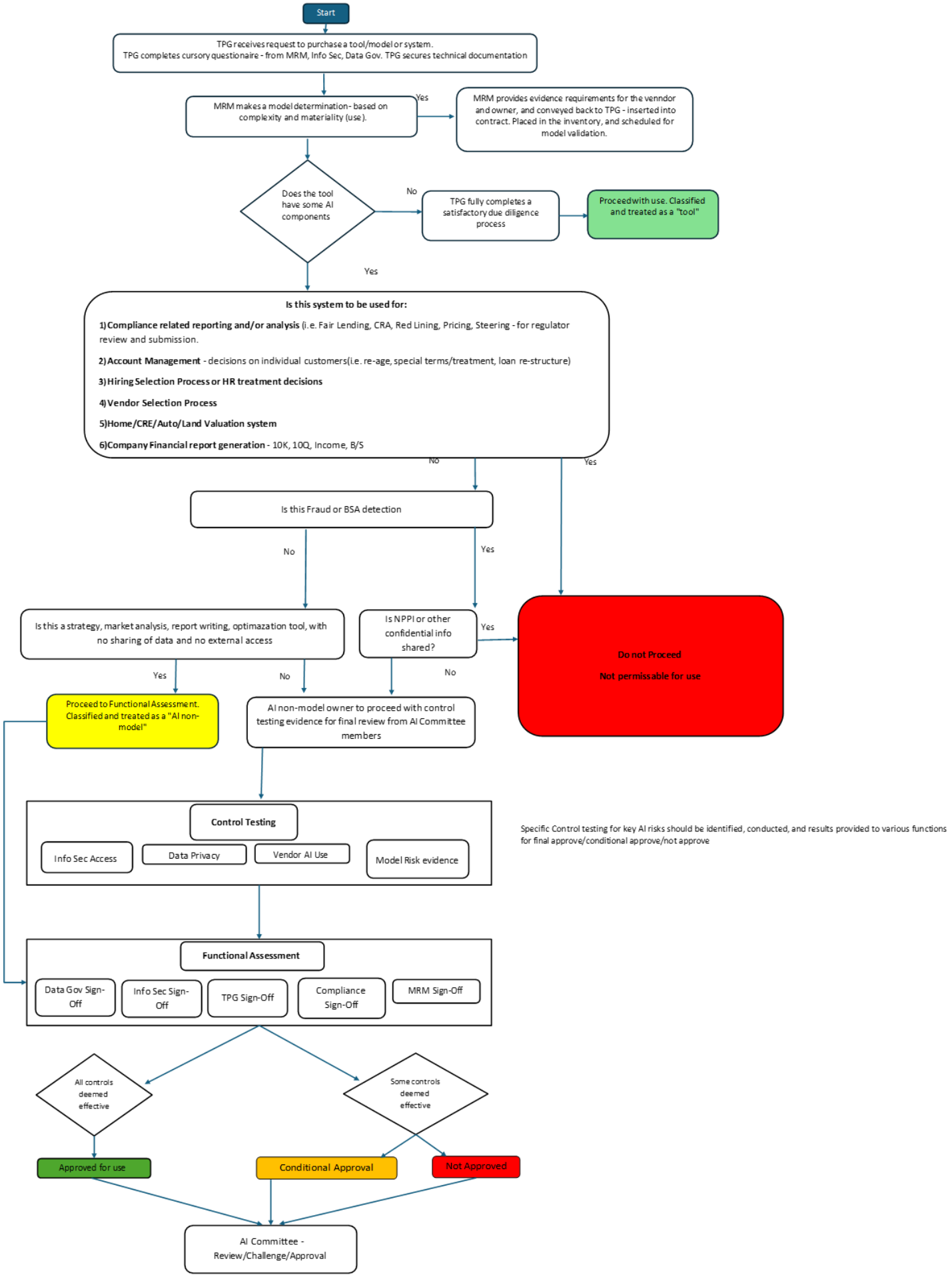
    d. Does an Incident Response Control exist? Is it tested on a periodic basis?

    e. Does the approving AI Committee provide evidence of review, challenge, and approval of all AI systems? What Q&A is available?

Recommendation 3: Design, implement, and manage controls that are aligned with AI system risks. The controls should be "effective", or actions taken and periodically reported to the AI Committee on how improvements are taken to address any deficiencies.

5. Coordinate Risk Workflows for AI Governance

It is critical that the risk groups work in collaboration to achieve the best possible outcomes. There is an inherent risk that reviews, control verification, testing, mitigation, requirement setting, and monitoring can be either incomplete or not executed. As such, the AI Committee can request a workflow for inquiry as key decisions are made. Below is a possible workflow that illustrates some of the types of coordination that warrant consideration.


Start
TPG receives request to purchase a tool/model or system.
TPG completes cursory questionaire - from MRM, Info Sec, Data Gov. TPG secures technical documentation
MRM makes a model determination - based on complexity and materiality (use).
Yes
MRM provides evidence requirements for the venndor and owner, and conveyed back to TPG - inserted into contract. Placed in the inventory, and scheduled for model validation.
Does the tool have some AI components
No
TPG fully completes a satisfactory due diligence process
Proceed with use. Classified and treated as a "tool"
Yes
Is this system to be used for:
1) Compliance related reporting and/or analysis (i.e. Fair Lending, CRA, Red Lining, Pricing, Steering - for regulator review and submission.
2) Account Management - decisions on individual customers(i.e. re-age, special terms/treatment, loan re-structure)
3) Hiring Selection Process or HR treatment decisions
4) Vendor Selection Process
5) Home/CRE/Auto/Land Valuation system
6) Company Financial report generation - 10K, 10Q, Income, B/S
No
Yes
Is this Fraud or BSA detection
No
Yes
Is this a strategy, market analysis, report writing, optimazation tool, with no sharing of data and no external access
Is NPPI or other confidential info shared?
Yes
Do not Proceed
Not permissable for use
Yes
No
No
Proceed to Functional Assessment. Classified and treated as a "AI non-model"
AI non-model owner to proceed with control testing evidence for final review from AI Committee members
Control Testing
Info Sec Access
Data Privacy
Vendor AI Use
Model Risk evidence
Specific Control testing for key AI risks should be identified, conducted, and results provided to various functions for final approve/conditional approve/not approve
Functional Assessment
Data Gov Sign-Off
Info Sec Sign-Off
TPG Sign-Off
Compliance Sign-Off
MRM Sign-Off
All controls deemed effective
Some controls deemed effective
Approved for use
Conditional Approval
Not Approved
AI Committee - Review/Challenge/Approval

Recommendation 4: Create a workflow that begins with a request for purchasing an AI system to Third Party Governance, and then routes the AI model/tool through Information Security and Data Governance; with assurances on permissible use and range of controls. Upon completion of control testing to achieve an effective level, a final AI Committee is expected to review all evidences to reach an approval. This is transparent, effective, and allows for greater sophistication of AI systems over time, as controls and coordination mature.

6. Conclusion

As artificial intelligence continues to evolve and integrate into banking operations, small to mid-size institutions face a critical inflection point: how to harness the benefits of AI while maintaining control over its risks. This paper has outlined a practical governance framework that emphasizes measured adoption, selective constraints, and effective controls. By establishing clear policies, defining roles and responsibilities, and implementing robust oversight mechanisms, banks can mitigate the unique challenges posed by AI - such as data privacy, model opacity, and emerging regulatory expectations; without stifling innovation.

Final Thought: The path forward is not one of avoidance or unchecked adoption, but of thoughtful engagement. A well-structured governance model, supported by cross-functional collaboration and continuous review, enables banks to responsibly leverage AI technologies in a way that aligns with their risk appetite, regulatory obligations, and strategic goals. With deliberate planning and disciplined execution, small to mid-size banks can position themselves to benefit from AI's potential while safeguarding their institutions, customers, and reputations.