

## 金融 AI 监管新篇章：拆解新加坡《AI 风险管理指南》

作者：权威 | 段志超 | 梁卓明 | 洪松

### 引言

2025 年 11 月 17 日，新加坡金融管理局（Monetary Authority of Singapore，以下简称“MAS”）发布了《关于人工智能风险管理指南的咨询文件》（Consultation Paper on Guidelines on Artificial Intelligence Risk Management，以下简称“《指南》”）。作为一份咨询文件，其内容在 2026 年 1 月 31 日咨询期结束前尚非最终定稿。正式发布后，《指南》将作为 MAS 对金融机构的监管期望，并给予业界 12 个月的过渡期以进行调整和落实。

在 AI 特别是生成式 AI 技术浪潮席卷全球金融行业的当下，各国监管机构均在积极探索如何平衡金融创新与风险防范，而 MAS 此次推出的《指南》无疑是这一探索过程中的一次重要实践。不同于部分司法辖区仍在原则性探讨或宏观框架层面，《指南》为金融机构构建了一套从顶层治理到具体技术实施的全生命周期 AI 风险管理框架。《指南》不仅为在新加坡运营或拟赴新发展的中资金融机构提供了清晰的合规路线图，其监管理念、框架设计和具体要求，对于正在构建和完善自身金融 AI 合规体系的中国企业而言亦具有重要的镜鉴意义。

本文旨在对《指南》的核心内容进行剖析，在此基础上探讨其对中国金融机构的潜在影响，以及对我国金融 AI 监管未来发展的启示。

### 一、《指南》的核心框架和主要内容

#### （一）风险为本与比例原则的贯彻

《指南》最为核心的监管理念，是其自始至终贯彻的**风险为本方法（Risk-Based Approach）**与**比例原则（Proportionality）**。MAS 明确认识到，AI 在金融领域的应用场景复杂多样，从辅助性的后台运营优化到决定客户信贷命运的核心风控模型，其潜在风险不可同日而语。因此，采取“一刀切”的强硬监管不仅会扼杀创新，也无法将有限的监管和合规资源投入到最需要的地方。

为此，《指南》明确提出，金融机构应根据其自身的业务规模、性质、AI 应用的广度与深度，以及由此产生的整体风险敞口，来决定采纳《指南》各项要求的具体程度。这一原则的落地，主要依赖于《指南》设计的**风险重要性评估（Risk Materiality Assessment）**机制。

该机制要求金融机构从**影响（Impact）**、**复杂性（Complexity）**和**依赖性（Reliance）**三个维度，对

每一项 AI 应用进行风险评级：

- **影响：**AI 系统或模型发生故障、失灵或性能不佳对金融机构（例如财务、运营、监管、声誉）及其客户或其他利益相关方（例如公平性、伦理违规、消费者保护）可能产生的后果。还应考虑 AI 系统或模型处理的数据的性质和敏感性。
- **复杂性：**源自所使用的 AI 技术的性质、其应用的新颖性或所使用的数据。随着对 AI 技术理解的不断发展，该风险维度也可能发生变化，例如，随着更多研究和更深入的熟悉，最初理解不充分的新 AI 技术的复杂性可能会改变。
- **依赖性：**考虑赋予 AI 系统或模型的自主程度、其所支持流程中人类参与或监督的程度；以及替代方案的可用性。

举例而言，一个高影响、高复杂性且业务高度依赖的 AI 系统（例如将 AI 作为审核贷款申请的风控模型）无疑需要最严格的治理和控制；反之，一个低风险的 AI 应用（例如辅助银行程序员补全代码）则可适用相对简化的合规措施。这种差异化的监管要求，体现了 MAS 作为成熟金融监管机构的务实与智慧，也为金融机构在合规与效率之间寻求平衡提供了指引。

## （二）从 FEAT 原则到可操作的风险管理体系

早在 2018 年，MAS 便联合业界发布了以公平（Fairness）、伦理（Ethics）、问责（Accountability）和透明（Transparency）为核心的 FEAT 原则，为 AI 在金融领域的负责任应用奠定了基础。为了支持金融机构在实践中落地 FEAT 原则，MAS 联合业界开展了一系列合作。例如，2019 年启动的 Veritas 倡议旨在开发评估方法论和工具包，帮助机构检验其 AI 系统的公平性。而随着生成式 AI 的兴起，MAS 又支持业界成立了 MindForge 项目联盟，专门研究生成式 AI 的风险与机遇，并正在制定一部《AI 风险管理手册》，作为《指南》的配套行业参考。

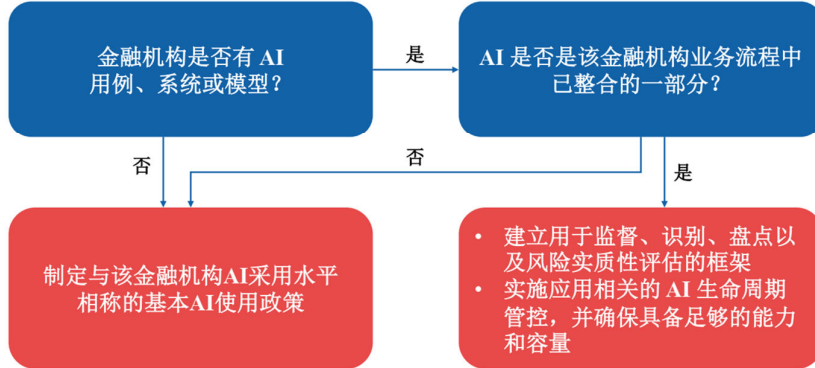
《指南》的核心价值在于其成功地将 FEAT 原则的抽象精神，解构为一套可操作、可审计的风险管理体系，包括以下核心要求：

1. **“看得见” — 建立 AI 识别机制：**金融机构的首要义务是“看见”风险。为此，必须建立明确的定义、标准和流程，以系统性地识别机构内所有正在使用或计划使用的 AI 系统，无论是内部开发、外部采购还是开源软件。这一过程需要有独立的控制职能部门（如风险或合规部门）负责并留下记录。所有被识别出的 AI 应用，都必须被登记在一个集中、准确且实时更新的清单中。我们理解，这份清单将成为 MAS 进行监管检查的核心依据。
2. **“分得清” — 实施风险重要性评估：**金融机构有义务建立一套客观、一致的方法论，对清单上的每一项 AI 应用进行风险重要性评估。如前所述，评估需围绕影响、复杂性和依赖性三个核心维度展开。评估结果不仅决定了该 AI 应用所需匹配的控制措施强度，也应作为向高级管理层汇报 AI 整体风险状况的基础。

通过这种方式，《指南》为金融机构提供了一份包含详尽示例的“操作手册”，并协助金融机构确定适用的 AI 合规思路（详见下图），使其能够将宏大的 AI 伦理目标，内化为内部风险管理、合规审查和技术开发的具体环节，真正实现从“说得到”到“做得到”的跨越。

金融AI监管新篇章：拆解新加坡《AI风险管理指南》

企业评估AI风险管理体系的适用性



(三) 覆盖 AI 全生命周期的闭环管控

《指南》构建了一个覆盖 AI 应用全生命周期的风险控制框架。MAS 强调，AI 风险管理绝非模型上线前的一次性审查，而是一个贯穿于 AI 系统整个存在期间的持续性、动态性过程。这种端到端的管理模式，确保了 AI 风险在模型的不断迭代和外部环境的动态变化中，始终处于金融机构的有效监控和管理之下。这一闭环管控主要由以下几个关键阶段构成：

AI 生命周期阶段	核心风险管控要求
数据管理	机构有义务确保用于 AI 的数据是高质量、具代表性且无偏见的。这包括建立数据质量监控、数据偏见检测与修正、数据安全与隐私保护（特别是对敏感个人信息）等一系列控制措施，确保用于 AI 模型开发与运行的数据在质量、适用性、代表性和合规性方面满足要求。
模型开发与验证	规范算法的选择、模型的训练与测试过程，确保模型在部署前经过充分验证，能够达到预期的性能目标，并有效防范过拟合等技术风险。
公平性与偏见管理	明确机构的“公平性”目标，主动检测和缓解 AI 模型中可能存在的歧视性偏见，并使用适当的公平性指标进行量化评估。
透明度与可解释性	根据风险等级，确保 AI 的决策过程对内部审查者和外部用户具有适当程度的透明度和可解释性，防止 AI 成为“黑箱”。
人工监督	建立有效的人工监督机制，在高风险决策场景中确保有意义的人工干预成为可能，并防范“自动化偏见”。
第三方 AI 风险管理	当金融机构使用外部供应商提供的 AI 模型或服务时，必须对供应商进行严格的尽职调查，评估其安全性、公平性、透明度及自身的风险管理能力。我们提示，相关责任和义务应通过合同条款予以明确，并制定应对供应商服务中断或模型失效的应急预案。

AI 生命周期阶段	核心风险管控要求
部署前审查	要求由独立于开发团队的人员对即将上线的 AI 系统进行审查，特别是高风险系统需进行正式的独立验证，作为模型上线的最后一道“闸门”。
部署后监控	建立持续的监控机制，跟踪模型性能、数据漂移、公平性指标等关键风险，并建立事件响应流程，确保对异常情况的及时处置。
变更管理	对已部署模型的任何变更，特别是重大变更，实施严格的控制、审批和测试流程，防止未经充分验证的变更引入新的风险。

#### （四）强调顶层治理与高层监督

《指南》明确将 AI 风险管理的“大脑”置于金融机构治理的最高层。《指南》反复强调，董事会和高级管理层是 AI 风险的最终“守门人”，必须对全机构的 AI 风险管理框架承担首要和最终的监督责任。这不仅要求他们批准机构的 AI 战略和风险偏好，更要求他们主动提升自身的 AI 知识水平，以进行有效的监督。

对于 AI 风险敞口被评估为重大的金融机构，《指南》还建议设立一个由风险、合规、技术、业务等部门专家组成的跨职能委员会，以实现 AI 风险的统筹协调和主动管理。这种自上而下的治理设计，旨在确保 AI 风险管理能够获得足够的重视和资源，并能与机构的整体战略保持一致，是整个风险管理体系有效运作的根本保障。

#### （五）对新兴 AI 技术的特别合规关注

值得特别关注的是，《指南》对以生成式 AI 和 AI 代理（AI Agents）为代表的新兴技术给予了高度重视，并提出了前瞻性的合规要求。对于生成式 AI，机构在应用时需重点评估和控制其产生“幻觉”（Hallucinations）、输出不准确或有害内容、泄露训练数据中的敏感信息，以及遭受提示注入（Prompt Injection）等新型攻击的风险。

对于具有更高自主性的 AI 代理，其风险控制的重点在于如何约束其行为的不可预测性，确保其自主决策和行动始终在预设的、安全的边界之内，并与机构的业务目标和客户的最佳利益保持一致。

## 二、对中国金融机构的影响与建议

《指南》的发布，不仅对新加坡本地金融机构具有约束力，对于在新加坡设有分支机构、或业务与新加坡市场有紧密联系的中国金融机构而言，同样构成了直接的合规挑战。同时，对于仅在中国境内运营的金融机构，这份代表国际先进水平的监管文件，也提供了对标和提升自身 AI 风险管理能力的绝佳参照。我们建议相关机构从以下几个方面评估其影响并制定应对策略。

### （一）对于在新加坡运营的中资金融机构

对于已在新加坡获得牌照并开展业务的中资银行、券商、保险公司等，遵循《指南》将是一项重要的合规义务。这些机构应在 MAS 建议的 12 个月过渡期内，立即采取以下行动：

1. **启动 AI 清单盘点与风险评估：**尽快对新加坡业务中使用的所有 AI 系统进行盘点，建立符合《指南》要求的 AI 清单，并根据 MAS 提出的三维度框架，完成对存量 and 新增 AI 应用的风险重要性评估。

2. **进行全面的差距分析：**组织由新加坡本地合规、风控、技术和业务部门，以及集团总部相关职能部门共同参与的专项工作组，以《指南》为标尺，逐条比对现有 AI 治理架构、政策流程、技术工具和人员能力，全面识别合规差距。
3. **本地化政策修订与流程改造：**根据差距分析和风险评估的结果，对现有的相关政策文件进行修订，或制定专门的 AI 风险管理政策。同时，需要对 AI 项目的开发、测试、上线和监控流程进行改造，将《指南》的控制要求嵌入其中。

对于计划申请新加坡各类金融牌照的中国企业而言，《指南》实际上提高了市场准入门槛。在向 MAS 提交的牌照申请材料中，一份健全、可信的 AI 风险管理计划或将成为重要组成。申请机构需要证明其不仅拥有创新的技术和商业模式，更具备与之一致的风险管理能力。我们建议，**相关企业应在启动牌照申请之前，就参照《指南》的要求，提前规划和构建自身的 AI 治理和风险管理体**系。

## （二）对于未在新加坡运营的中国境内金融机构

对于主要在中国境内运营的金融机构，《指南》虽然没有直接的法律约束力，但其作为国际最佳实践的标杆价值不容忽视。当前，生成式 AI 在国内金融行业的应用已初见端倪，从智能客服、营销文案生成到辅助代码编写，其潜力巨大，但伴随的“幻觉”、数据安全、偏见放大等风险也日益凸显。

目前，我国对金融 AI 的监管呈现出“多部门、多层次、点状化”的特征。例如，金融行业标准《人工智能算法金融应用评价规范》侧重对 AI 算法本身在安全性、可解释性、准确性等技术性指标进行评估，《银行保险机构信息科技外包风险监管办法》则涉及银行保险机构外包风险的管理，而网信办的《生成式人工智能服务管理暂行办法》则聚焦于特定的 AI 应用形态。这些规范在各自领域发挥了重要作用，但相互之间缺乏一个统一的、顶层的治理框架来统筹。

**随着中国金融市场的持续开放和国际化，接轨国际先进的风险管理标准，是提升金融机构自身竞争力的内在要求。**我们建议境内金融机构：

1. **将其作为内部 AI 风险管理的“体检表”：**主动使用《指南》审视和评估自身在 AI 治理、风险文化、技术控制等方面的成熟度，识别潜在的短板和风险点。
2. **将其作为完善内部制度的“参考书”：**在制定或修订内部的 AI 相关管理制度时，充分借鉴《指南》中关于 AI 生命周期管控、第三方风险管理、新兴技术风险应对等方面的具体做法，提升内部制度的科学性和前瞻性。
3. **将其作为培养人才的“教科书”：**组织高级管理人员、风险经理、合规官和技术骨干学习《指南》，提升整个组织对 AI 风险的认知水平，为未来更复杂的 AI 应用和更严格的监管环境做好准备。

## 三、结语

新加坡 MAS 的《指南》无疑为全球金融 AI 治理的深水区投下了一块重要的“探路石”。它以一种严谨而不失灵活、全面而不失务实的姿态，向市场清晰地传递了“拥抱创新，但严守底线”的监管信号。对于身处数字化转型浪潮中的中国金融机构而言，深入研读并参考《指南》，不仅是了解境外先进监管实践的需要，更是提升自身 AI 风险治理能力的契机。

## 特别声明

汉坤律师事务所编写《汉坤法律评述》的目的仅为帮助客户及时了解中国或其他相关司法管辖区法律及实务的最新动态和发展，仅供参考，不应被视为任何意义上的法律意见或法律依据。

如您对本期《汉坤法律评述》内容有任何问题或建议，请与汉坤律师事务所以下人员联系：

### 权威

电话： +86 21 6080 0946  
Email: [wei.quan@hankunlaw.com](mailto:wei.quan@hankunlaw.com)

### 段志超

电话： +86 10 8516 4123  
Email: [kevin.duan@hankunlaw.com](mailto:kevin.duan@hankunlaw.com)

### 梁卓明

电话： +65 6013 2968  
Email: [chuoming.leong@hankunlaw.com](mailto:chuoming.leong@hankunlaw.com)