

Centre for  
Regulatory  
Strategy

亚太地区生成式人工智能  
应用与监管

因我不同  
成就不凡

始于 1845

引言	3
<b>第一部分：传统人工智能与生成式人工智能</b>	<b>4</b>
<b>第二部分：生成式人工智能相关风险</b>	<b>6</b>
<b>第三部分：亚太地区人工智能监管措施</b>	<b>7</b>
<b>第四部分：高可信人工智能框架的应用</b>	<b>10</b>
联系人	17
尾注	19

# 引言

过去一年，大语言模型 (LLM) 和自然语言处理模型等人工智能 (AI) 技术的发展取得了重大突破。这些技术已经通过OpenAI的ChatGPT、微软Bing AI Chat和谷歌Bard AI等工具得到广泛传播，并引起全球消费者的热议、追捧和警惕。

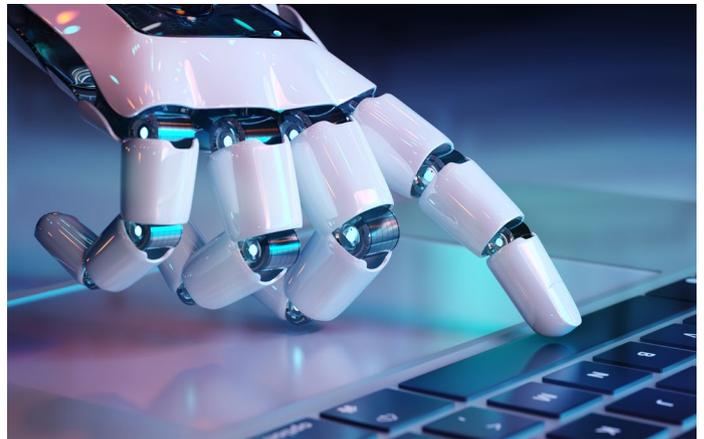
AI平台对广大用户的可触达性突显了AI技术在各行各业 (包括金融服务业) 的应用潜力。许多企业开始利用AI技术提高自身竞争优势。监管机构和立法机构需更加迅速、敏捷、主动地应对AI应用所带来的相关风险。

在德勤2022年发布《[人工智能在金融服务业的可靠应用](#)》报告时，亚太地区许多监管机构仍处于商讨和/或实施AI原则的起步阶段。随着AI工具在金融服务业得到应用和普及，部分立法和监管机构已经开始研究AI应用的相关风险，以保障消费者权益。在本篇后续内容中，我们将进一步探讨金融服务业使用AI的相关风险、亚太地区监管现状以及金融机构在准备应对即将出台的相关法律法规时的考量因素。

# 第一部分：传统人工智能 与生成式人工智能

## 知识更新：了解传统AI与生成式AI

**传统AI**是指可以自动处理预定义输入的系统。此类AI系统能够从训练数据中获取知识，并利用这些知识做出决策或预测。例如，许多企业利用AI聊天机器人提供精简高效的客户支持。传统AI聊天机器人在处理常见问题方面尤其有效。凭借内部搭建的知识库，其可针对常见问题提供准确一致的回复并进行用户意图预测。



**生成式AI**可以编写文本、生成代码、制作音频和图像，其水平与人类不相上下，甚至超越人类。例如，生成式AI工具包括可用于生成书面文本（如营销文案、软件代码等）和图像等内容的LLM。生成式AI模型具有生成连贯文本和超逼真图像的能力，其可采用以前只能通过人类的思维、努力和创造力才能实现的方式生成数据。

传统AI和生成式AI的不同功能驱动了不同用例。就金融服务业而言，传统AI可以用于开展预测分析或检测可疑交易，而生成式AI可以加速完成从交易和研究到通过生成相关报告为合规职能提供关键支持等任务，本报告将对此作进一步阐述。



图1: 传统AI与生成式AI



## 第二部分：生成式人工智能 相关风险

在2022年发布的《[人工智能在金融服务业的可靠应用](#)》报告中，探讨了亚太地区监管机构希望通过AI监管原则解决的常见风险要素：透明度、问责制、公平性、稳健性、隐私和数据安全。目前此类风险和担忧依然存在，而生成式AI的兴起又给市场带来了新的风险：

- **缺乏透明度：**考虑到生成式AI模型的复杂性及其所涉信息的专有性，人们普遍认为生成式AI缺乏透明度。此外，在衡量或评估生成式AI模型的透明度方面缺乏标准化的工具和方法，这可能导致在比较不同模型和追踪长期进展时变得困难。
- **歧视和偏见：**生成式AI可能会将一些偏见与训练数据中的模式形成关联，从而生成歧视性或误导性内容。

- **缺乏准确性和产生错误观念：**生成式AI可能会利用不完整、不准确或有偏见的数据生成不准确或有误导性内容，或者干脆生成虚构事实。生成式AI模型没有固有的“客观真理 (objective truth)”，可能会生成错误甚至有害的内容和观点。
- **知识产权和版权问题：**生成式AI模型可能会以受版权保护的材料为基础进行训练，从而生成与受版权保护的材料非常相似的内容。生成式AI模型还可能用于制造假冒或盗版商品，侵犯知识产权。
- **欺诈：**生成式AI可能生成深度伪造和合成数据，这些数据可以用于实施欺诈、传播错误信息或造成系统漏洞。

# 第三部分：亚太地区人工智能监管措施

生成式AI的出现迫使亚太地区政策制定机构和监管机构重新评估之前实施的AI框架是否同样适用于降低新兴技术风险。某些监管机构已经实施AI指引和计划，为企业和行业提供最佳实践建议。下表（图2）列举了亚太司法管辖区在开展AI监管或为AI风险管理提供建议方面所采取的措施，包括制定AI原则、提供指导和工具、出台立法以及将AI应用纳入国家战略：

- **AI原则：** AI原则为有效管理与各行业使用AI相关风险提供了指引。例如，欧盟以AI原则为入手点开展AI监管以及出台立法。值得注意的是，某些选择针对AI风险出台立法或开展监管的司法管辖区也推出了AI原则。举例而言，中国大陆在对AI应用进行立法的同时，国家新一代人工智能治理专业委员会发布了《新一代人工智能治理原则——发展负责任的人工智能》。
- **指导和工具：** 指导和工具通常用于支持AI原则的实施。以新加坡为例，由新加坡金融管理局领导的Veritas联盟发布了五份白皮书，阐述了公平、道德、负责和透明（FEAT）原则的评估

方法。为推动金融机构加快采用FEAT方法和原则，联盟开发了Veritas Toolkit 2.0版。与1.0版相比，2.0版改进了公平原则评估方法，并纳入了道德、负责和透明原则评估方法。2022年5月，资讯通信媒体发展局和个人数据保护委员会推出全球首个AI治理测试框架和工具包——A.I. Verify，适用于旨在以客观和可验证的方式展示负责任的AI的企业。

- **立法：** 韩国、中国大陆、菲律宾和越南等司法管辖区采取了针对保险业出台AI专项立法的措施，其中中国大陆和越南已通过AI专项立法。
- **国家战略：** 泰国、印度尼西亚、日本、中国大陆和马来西亚等许多亚太司法管辖区已将AI确定为战略重点，并制定了促进可信AI应用的国家战略，但是某些司法管辖区尚未在实施战略或向业界提供结构化框架方面取得进展。

图2: 监管机构和立法机构为应对AI相关风险而采取的措施



就其他地区而言,欧盟和美国等司法管辖区也已着手采取措施来应对生成式AI的快速发展。欧盟《人工智能法案》(Artificial Intelligence Act)是欧盟委员会为规范欧盟AI系统而提出的立法,并已纳入欧盟确保以负责任的方式开发和使用该技术的整体战略。《人工智能法案》旨在建立基于风险的框架,以应对AI相关风险,同时促进创新和提高竞争力。相比之下,美国采取的AI监管措施较为分散。美国的法律和监管结构以州为基础,在联邦层面尚未颁布或提出规范生成式AI的法律。某些州(包括加利福尼亚州和科罗拉多州)已着手推进AI立法,而某些州则在监测不断变化的风险。

图3: 其他监管辖区为应对AI相关风险而采取的立法和监管措施

欧盟	美国
<ul style="list-style-type: none"> <li>• 《人工智能法案》已于2023年底通过成为法律。<sup>13</sup></li> <li>• 《人工智能法案》旨在整合欧盟所有国家的AI相关法律，制定AI整体立法，并将其框架推广为全球基准。</li> <li>• AI风险评级分为四个等级： <ul style="list-style-type: none"> <li>- <b>禁用AI系统</b>：因“侵犯基本权利”而被明确禁用的系统，例如公共场所的实时生物识别系统。此外，《人工智能法案》严禁任何蓄意利用成人或儿童的弱点造成伤害的系统。</li> <li>- <b>高风险AI系统</b>：用于或属于受欧盟产品安全立法约束的产品或《人工智能法案》附件III所列产品的系统。高风险AI系统受《人工智能法案》全面监管并须满足多项要求。</li> <li>- <b>有限风险AI系统</b>：可与人类直接交互的系统。此类系统必须满足透明度要求并告知用户他们正在使用AI。</li> <li>- <b>低风险或最小风险AI系统</b>：简单的AI系统，例如垃圾邮件过滤器或AI视频游戏。此类系统不受任何限制。</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• 美国联邦、各州和各行业的AI监管法规各不相同。</li> <li>• 科罗拉多州等州针对AI在保险业的应用制定的专项法律已于2023年生效。</li> <li>• 目前仍有20多个州未制定AI相关法律法规。</li> <li>• 2022年10月，白宫科技政策办公室发布《人工智能权利法案蓝图》(Blueprint for an AI Bill of Rights)<sup>14</sup>，其中提出五项原则： <ul style="list-style-type: none"> <li>- 安全有效的系统</li> <li>- 算法歧视保护</li> <li>- 数据隐私</li> <li>- 通知和解释</li> <li>- 替代方案</li> </ul> </li> </ul> <p>该蓝图并非法规，无法强制执行，其目的是对白宫（科技政策办公室）认为存在问题的AI用途提供指导。</p>

## 挑战和考虑因素



生成式AI的快速发展给亚太地区监管机构带来了新的挑战，具体法规的制定和实施往往被认为没有效果，且可能很快过时。监管机构在应对生成式AI给金融服务业带来的新挑战和新风险方面也面临困难。

由于人才短缺以及公共和私营部门为吸引具备适当AI技术能力的人才而展开的激烈竞争，导致某些监管机构无法灵活应对AI技术带来的新兴和不断变化的风险和发展趋势。立法机构和监管机构在监督和执行相关指令方面也举步维艰。例如，就AI定义达成共识是实施AI法律法规的关键问题。

虽然AI应用可能因利益相关方而异，但就AI原则达成共识对于利用AI改善金融服务且不影响安全性、公平性或消费者保护至关重要。归根结底，有效的监管应当推动创新，同时保障金融生态系统中所有相关方的利益。

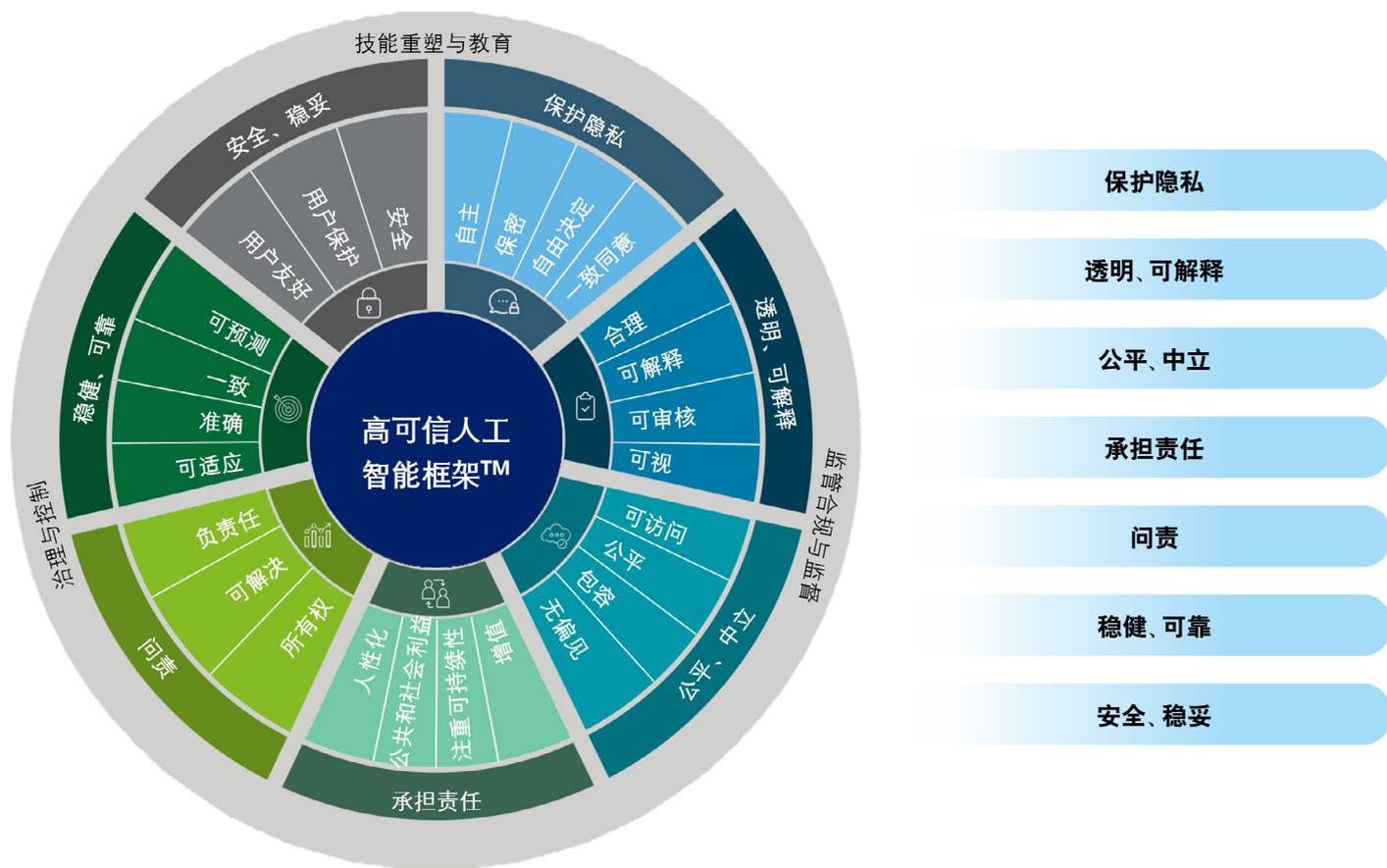
行业和监管机构可以加强区域合作，建立跨境治理框架，开展联合研究，确定最佳实践，以实现监管标准化。此类联盟有助于确保监管措施始终依据实际见解应对全球挑战及平衡行业增长和社会保障。

# 第四部分：高可信人工智能框架的应用

相比传统AI，生成式AI可能会对使用AI应用的金融机构提出更具挑战性的风险管理要求。由于大多数司法管辖区仍处于制定或实施AI法律法规的起步阶段，金融机构必须尽早建立自有AI治理框架并将全球/区域AI原则纳入其中。

金融机构应当利用该框架系统管理与使用生成式AI相关的风险。这对确保AI监管合规、加强用户保护以及进一步推动AI应用的成功实施至关重要。《人工智能在金融服务业的可靠应用》报告中，我们简要介绍了德勤高可信人工智能框架。在本节中，我们将探讨在不同用例中如何使用德勤高可信人工智能框架管理潜在的AI相关风险。

图4：德勤高可信人工智能框架<sup>16</sup>



优势	高可信人工智能框架的应用		
	问题/机遇	生成式AI如何提供助力	需要考虑的高可信人工智能框架要素
<p><b>降低成本</b></p> <p>简化运营, 提高效率, 增加劳动力利用率, 实现监管合规, 从而降低成本。</p>	<p>客户接洽属于劳动密集型工作并遵循KYC标准。在此过程中需要对客户进行大量人工调查, 包括经济分析、股权调查、负面报道调查和尽职调查, 极其耗费人力和时间。</p>	<p><b>生成研究报告 (了解你的客户, KYC)</b></p> <p>生成式AI可以利用现有搜索引擎或生成式AI聊天工具进行初步数据搜索和元分析, 以此提高效率并<b>创造价值</b>。此外, 其也可编制元分析报告, 并为客户经理提供摘要信息。<b>更加便利</b>和及时的信息获取将促使员工转而开展更有价值的工作。</p>	<p><b>保护隐私</b>——使用生成式AI时必须采取预防措施, 防止敏感信息泄露, 并规范对模型、基础数据和引用客户数据的访问。</p> <p><b>稳健、可靠</b>——使用生成式AI进行搜索和分析可能遗漏相关信息, 从而影响元分析和决策制定。</p>
	<b>增强客户支持</b>		
	<p>客户服务对于正在向以客户为中心模式转型的金融机构至关重要。其中, 快速准确地响应客户问询十分关键, 但数字化趋势限制了其与客户代表接触的机会。当客户期待超个性化体验时, 这一挑战也随之而来。</p>	<p>生成式AI通过提供超个性化体验和<b>人性化</b>响应来改善客户界面。与传统聊天机器人不同, 生成式AI可以表现同理心、总结合同内容以及回复细致问询。这项技术以LLM为基础, 提供文本、音频和图像等多种界面选择, 显著提高了聊天机器人的实用性和<b>易用性</b>。改进客户服务可以提高声誉、效率和客户忠诚度, 企业因此得以扩大运营规模、优先处理复杂任务、吸引新客户并降低相关成本。</p>	<p><b>保护隐私</b>——金融机构必须遵守敏感信息安全传输、存储和访问规则。</p> <p><b>透明、可解释</b>——最终用户需要了解其信息的处理方式, 而金融机构需要解释输出结果并了解输出结果的生成方式和原因。</p>

优势	高可信人工智能框架的应用		
	问题/机遇	生成式AI如何提供助力	需要考虑的高可信人工智能框架要素
<p><b>加快执行速度</b></p> <p>最大限度降低延迟,自动执行重复任务,显著缩短运营时间。</p>	<p>在财险理赔过程中,代理人需评估保险事故并确定损失金额。这一过程相当复杂和耗时,且代理人缺乏相关工具来支持决策。</p>	<p>生成式AI能够根据客户对话、文档内容和媒体报道对保险事故进行虚拟复制,以帮助代理人<b>准确</b>评估损失,还能基于照片证据生成理赔报告,从而加快理赔速度,提高损失评估准确性,快速识别欺诈行为,确保理赔和赔付的<b>可问责性</b>和公正性。</p>	<p><b>稳健、可靠</b>——损失可视化对准确性提出较高要求,错误的输出结果可能导致保险公司多赔或少赔。</p> <p><b>透明、可解释</b>——如果理赔代理人无法阐明输出结果的生成方式,客户可能不会接受理赔结果。</p>
	<p>在当前的信贷发放过程中,传统AI驱动的信用评分缺乏透明度,代理人必须对基本方法有所了解。某些情况下,信用评分未完全整合,需要人为干预来证明信用额度的合理性。</p>	<p>生成式AI可以整合客户数据,而在信贷发放过程中估算信用额度,并以类人沟通方式作出回应。相比传统AI,生成式AI能够以类人思维模式生成<b>可解释</b>的决策声明,提高人工<b>审计的透明度</b>。人工参与的减少显著提高了信贷发放效率,推动了组织扩张。</p>	<p><b>保护隐私</b>——信贷发放过程需要处理大量客户财务数据,且在数据使用方面存在严格规定。确保隐私合规和适当的信息披露至关重要。</p> <p><b>公平、中立</b>——基于历史客户数据训练的信用额度模型可能暗藏偏见,从而导致决策缺乏公平性。</p>
<p>开展战略交易分析需要实时获取技术数据、新闻报道和行业报告。分析师必须基于大量信息来了解和预测市场轨迹,并谨慎做出买卖决策。由于分析过程费时费力,导致金融机构面临挑战。</p>		<p>生成式AI依托<b>预测</b>分析生成对冲策略和开展情感分析,助力实时降低投资风险。生成式AI可以加快市场预测,为分析师提供有效支持,并增加交易量,从而在降低风险的同时提高盈利能力。</p>	<p><b>稳健、可靠</b>——虽然生成式AI可以帮助分析师提高市场预测的准确性,但是基于不可靠输出结果做出决策的风险仍然存在,或将造成不良后果。</p> <p><b>透明、可解释</b>——针对生成式AI的输出结果进行人工验证仍然至关重要,并且利益相关者必须了解输出结果的生成方式和原因,才能对其抱有信心。</p>

优势	高可信人工智能框架的应用		
	问题/机遇	生成式AI如何提供助力	需要考虑的高可信人工智能框架要素
<p><b>降低复杂性</b></p> <p>通过更加有效的预测性分析识别日益复杂的数据源模式，简化运营流程，改善决策制定。</p>	<p>金融机构正在通过云转型和数据转换促进AI工具整合。虽然这可以释放人力资源并降低内部部署成本，但却需要投入大量时间和资金，且存在失败和错误风险。</p>	<p><b>数字化转型代码助手</b></p> <p>开发团队可以利用生成式AI执行<b>负责</b>的编码、调试和归档任务，从而降低工作复杂性。这种方法可以加快软件部署，缩短开发生命周期并加速达到稳定且可部署的状态（例如快速编写<b>透明且可解释</b>的API、ETL、数据管道和前端代码），从而提高<b>效率</b>。</p>	<p><b>稳健、可靠</b>——编程相关任务的部分自动化对系统的可用性和准确性提出较高要求。</p> <p><b>承担责任</b>——基础模型的训练数据可能引发与知识产权或版权侵权相关的法律风险。</p> <p><b>安全、稳妥</b>——使用生成式AI可能泄露专有代码，引发安全问题，侵犯知识产权，造成严重后果。</p> <p><b>问责</b>——如果代码验证和调试等工作没有人工参与，则可能出现严重故障，因此需要制定归档和沟通标准。</p>
	<b>机构范围内的数据搜索和访问</b>		
	<p>涉及多个地点的复杂数据存储导致金融机构查询效率低下，从而造成见解片面、风险增加和客户不满等问题（尤其是在并购之后）。</p>	<p>作为用户查询和数据库之间的桥梁，生成式AI可以实现高效数据挖掘、结构化分析和快速洞察生成，从而降低数据查询的复杂性，提高非传统方式下商业智能的<b>可访问性</b>。</p>	<p><b>承担责任</b>——在扩展数据访问权限时，金融机构必须明确设定敏感业务数据访问限制，从而实现有效治理和控制。</p> <p><b>稳健、可靠</b>——由于生成式AI可能产生幻觉并生成不可靠的结果，金融机构或将面临提供不准确的洞察等后果。</p> <p><b>保护隐私</b>——在处理敏感数据时，金融机构必须确保数据安全，删除或隐藏训练和测试数据集中的敏感数据，并对模型进行评估，以防止受保护信息泄露。</p>

优势	高可信人工智能框架的应用		
	问题/机遇	生成式AI如何提供助力	需要考虑的高可信人工智能框架要素
<p><b>降低复杂性 (续)</b></p> <p>通过更加有效的预测性分析识别日益复杂的数据源模式, 简化运营流程, 改善决策制定。</p>	<p>金融服务业受到严格监管, 代理人需要了解诸多相关法规。这一过程极其耗费时间, 而且法规解读因人而异, 因此可能造成疏漏。</p>	<p><b>监管机器人</b></p> <p>生成式AI可以对法规和指引进行总结并创建综合目录。这种<b>用户友好</b>界面可以助力金融机构提高效率, 及时响应合规要求, 并减轻人工查阅法规的负担。</p>	<p><b>稳健、可靠</b>——生成式AI可能基于模糊的数据和错误的解读提供误导性见解。</p> <p><b>问责</b>——监管要求有时模棱两可, 需要进行人工解读。</p>
<p><b>转变交互方式</b></p> <p>赋能技术产品与客户进行人性化沟通, 弥合机器语言和人类语言之间的差距。</p>	<p>客户愈发倾向于通过数字设备进行在线/远程交易, 因此金融机构开始寻求自动化解决方案。虽然聊天机器人配有自动化功能, 但由于其对话和选项经过预先编程, 其目前在处理对话方面存在局限性。</p>	<p><b>虚拟银行体验</b></p> <p>生成式AI可以赋能金融机构在虚拟空间中与客户进行个性化互动并就客户问询提供针对性回复, 且无需额外的人工客服。这种<b>用户友好</b>方式可以实现以客户首选语言进行及时交流, 并通过实时数据访问提高服务质量和速度。</p>	<p><b>问责</b>——如果聊天机器人出现错误, 利益相关者应当承担责任, 可以通过让人类参与及记录角色和职责来促进责任承担。</p> <p><b>公平、中立</b>——用于训练聊天机器人的数据集可能暗藏偏见, 例如代表性不足的客户群体或某些语言的语义缺陷, 这可能导致客户产生负面影响并提出投诉。</p> <p><b>透明、可解释</b>——考虑到聊天机器人的对话能力, 金融机构需要告知客户其互动对象为聊天机器人以及其输入和信息将如何存储和使用, 从而与客户建立信任。</p>

优势	高可信人工智能框架的应用		
	问题/机遇	生成式AI如何提供助力	需要考虑的高可信人工智能框架要素
<p><b>推动创新</b></p> <p>生成式AI可以催生更多新业务,同时满足客户对定制产品和服务的期望。</p>		<p><b>生成合成数据</b></p> <p>数据缺失致使金融机构面临严峻挑战。生成式AI模型可以学习给定数据集中的基本模式,从而生成更加多样且真实的新数据,为测试机器学习算法以及开发新产品和新服务等工作提供助力。</p>	<p><b>公平、中立</b>——生成合成数据可能在无意中延续历史偏见,例如由于银行过去的行为导致某些社区或社会经济群体的代表性不足。</p> <p><b>稳健、可靠</b>——生成式AI生成的合成数据在范围和规模方面可能受到限制。过度依赖此类合成数据或将影响数据的可靠性,以及模型输出和训练的准确性和有效性。</p>
	<p><b>定制营销材料</b></p> <p>文化差异以及客户对产品的不同理解可能会使金融机构面临监管风险。为克服这一难题,金融机构投入了大量人力、时间和资金来维持营销的合规性。</p>	<p>生成式AI赋能金融机构以适当的论调、语言和文化参考定制<b>公平</b>营销材料,确保营销内容在触达大量个人客户的同时符合多个地区的监管要求,从而降低监管风险。</p>	<p><b>公平、中立</b>——训练数据中的隐藏偏见可能导致营销材料无法顾及关键地域和文化差异。</p> <p><b>稳健、可靠</b>——由于AI输出结果中的幻觉可能导致虚假陈述和监管违规,因此人工验证对于确保营销材料的可靠性至关重要。</p>
<p><b>加强信任</b></p> <p>确保业务免受欺诈、网络攻击和监管漏洞影响,改善产品和服务质量,提高透明度,以此增强品牌信任度。</p>	<p>监管要求指出,企业风险管理包括评估和管理与信贷、投资、欺诈和网络安全相关的风险。该过程依赖于身份验证和信用评估等多个数据来源,操作复杂且容易出错,尤其是对于在不同市场拥有数百万客户的大型金融机构而言。</p>	<p><b>开展实时风险管理</b></p> <p>实时获取相关数据和信息有助于确保符合监管要求和行业标准。凭借稳健的实时风险评估,金融机构可以更加迅速准确地应对新兴风险和趋势,从而更加灵活地满足对<b>安全</b><b>稳妥</b>的AI风险管理的监管要求。</p>	<p><b>公平、中立</b>——数据源中的偏见可能影响生成式AI对客户风险评估的公正性。</p> <p><b>问责</b>——生成式AI无法对因忽略风险导致金融机构做出错误决策的后果负责。</p> <p><b>安全、稳妥</b>——由于风险管理涉及敏感信息,金融机构需要确保访问数据的模型不会将客户数据泄露或无意中泄露给未经授权方。</p>

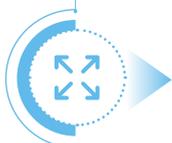
## 关键点



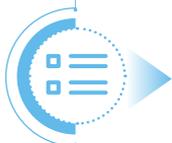
生成式AI等AI技术为提升金融服务业的效率和数字化水平带来巨大可能性的同时，也应对相关风险予以充分重视。随着AI技术监管环境的不断变化，考虑采用AI技术的金融机构应采取行动了解、识别和管理AI相关风险。



虽然亚太地区在制定AI法律法规方面仍处于起步阶段，但已经采用或考虑采用AI应用的金融机构应着手开发AI治理框架，以此为风险管理和监管合规提供支持。金融机构应对生成式AI应用的输出结果负责。



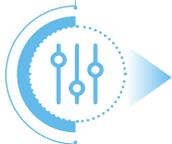
采用生成式AI应用的金融机构应考虑AI相关风险和AI治理框架如何适配其风险偏好和整体风险管理框架。



采用生成式AI应用的金融机构应明确其预期目标和AI应用的使用范围，并对客户安全、健康和基本权利受到损害的可能性进行评估。金融机构应加大对针对高风险应用的人工监管力度。



采用生成式AI应用的金融机构应对影响客户脆弱性水平的因素（例如教育背景、收入或年龄）进行评估。金融机构应避免因采用生成式AI应用而有意或无意产生对（弱势）客户的偏见和歧视。



采用生成式AI应用的金融机构应明确负责收集、存储和处理客户个人数据的外部方或内部职能部门，并确保其遵守数据保护要求。



由于尚不清楚生成式AI的哪些输入或输出受版权保护，采用生成式AI应用的金融机构应假设输入的数据或查询都可能公开，因此其应采取控制措施，防止无意中泄露知识产权或侵犯版权。



已经采用或考虑采用生成式AI应用的金融机构应开展人才招聘工作，并帮助现有员工（包括帮助董事会和高管层）了解AI技术的基本概念、部署方式、主要风险以及其在降低相关风险方面应承担的责任。



私营部门应与监管机构和立法机构开展积极对话，分享行业知识和经验，以促进规则制定并推动就人工智能的未来发展方向达成共识。

# 联系人

## 中国联系人



### 方焯

德勤中国金融服务业  
风险咨询主管合伙人  
+86 21 6141 1569  
yefang@deloitte.com.cn



### 吴洁

德勤中国金融服务业  
风险与合规服务主管合伙人  
+86 21 6141 2237  
kwu@deloitte.com.cn

## 亚太地区监管政策中心



### 神谷精志

亚太地区监管策略中心主管  
德勤亚太合规与法律支持主管合伙人  
seiji.kamiya@tohmatsumsu.co.jp



### 黄毅城

亚太地区监管策略中心东南亚联席主管  
德勤东南亚监管策略主管合伙人  
nawong@deloitte.com



### Sean Moore

亚太地区监管策略中心澳大利亚联席主管  
德勤澳大利亚监管风险主管合伙人  
semoore@deloitte.com.au



### 小林信也

亚太地区监管策略中心日本联席主管  
德勤日本金融服务业风险咨询常务总监  
shinya.kobayashi@tohmatsumsu.co.jp



### Jaramie Nejal

运营主管  
德勤澳大利亚金融服务业风险咨询总监  
jnejal@deloitte.com.au

# 作者

## 苏宁馨

作者  
经理, 德勤中国

## Andrew Neilson

作者  
高级分析师, 德勤澳大利亚

## Jennifer Martiniak

作者  
顾问, 德勤新加坡

## Cathy Zhang

作者  
顾问, 德勤澳大利亚

# 致谢

## Nicola Marshall

合伙人  
澳大利亚

## Sam Walsh

合伙人  
英国

## Matthew Gracie

常务总监  
美国

## 陈智明

合伙人  
中国

## Patrycja Grzesznik

副主席  
美国

## Gerry Chng

执行总监  
新加坡

## Kedarnath Vallaboina

高级经理  
新加坡

# 尾注

1. Department of Industry, Science and Resources, *Australia's Artificial Intelligence Ethics Framework*, November 2019, [Australia's Artificial Intelligence Ethics Framework | Department of Industry, Science and Resources](#)
2. 香港个人资料私隐专员公署,《开发及使用人工智能道德标准指引》,2021年8月, [guidance\\_ethical\\_c.pdf \(pcpd.org.hk\)](#)
3. Ministry of Economy, Trade, and Industry, *Governance Guidelines for Implementation of AI Principles Ver. 1.1*, January 2022, [20210709\\_8.pdf \(meti.go.jp\)](#)
4. 台湾金融监督管理委员会:金管会就金融业运用人工智慧(AI)之原则及政策草案公开征询外界意见,2023年8月, [新闻稿-金管会就金融业运用人工智慧\(AI\)之原则及政策草案公开征询外界意见](#)
5. Monetary Authority of Singapore, *Veritas Document 3 FEAT Principles Assessment Methodology*, January 2021, [Veritas-Documents-3---FEAT-Principles-Assessment-Methodology.pdf \(mas.gov.sg\)](#)
6. 中国国家互联网信息办公室,《生成式人工智能服务管理暂行办法》,2023年7月, [生成式人工智能服务管理暂行办法\\_中央网络安全和信息化委员会办公室](#)
7. Republic of the Philippines House of Representatives, *An Act Establishing a Regulatory Framework for a Robust, Reliable, and Trustworthy Development, Application, and Use of Artificial Intelligence (AI) Systems, Creating the Philippine Council on Artificial Intelligence, Delineating the Roles of Various Government Agencies, Defining and Penalizing Certain Prohibited Acts*, May 2023, [hrep-website.s3.ap-southeast-1.amazonaws.com/legisdocs/basic\\_19/HB07913.pdf](#)
8. National Assembly of the Republic of Korea, *the Act Fostering the AI Industry and Establishing a Foundation for Trustworthy AI*, February 2023, [전체 - 보도자료 | 알림마당 | HOME \(assembly.go.kr\)](#)
9. National Assembly of Vietnam, *Insurance Business Law*, June 2022, [Law No. 08/2022/QH15 dated June 16, 2022 on insurance business - LawNet](#)
10. OECD, *AI Strategies and Policies in Indonesia*, [AI Strategies and Policies in Indonesia - OECD.AI](#)
11. Ministry of Science, Technology and Innovation, *Malaysia National Artificial Intelligence Roadmap*, [Dasar - Portal Rasmi Kementerian Sains, Teknologi Dan Inovasi \(mosti.gov.my\)](#)
12. OECD, *AI Strategies and Policies in Thailand*, [AI Strategies and Policies in Thailand - OECD.AI](#)
13. European Commission, *Proposal for A Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, April 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>
14. The White House Office of Science and Technology Policy (OSTP), *Blueprint for and AI Bill of Rights*, October 2022, [Blueprint for an AI Bill of Rights | OSTP | The White House](#)
15. MIT Technology Review, *'Trustworthy AI' is a framework to help manage unique risk*, March 2020, <https://www.technologyreview.com/2020/03/25/950291/trustworthy-ai-is-a-framework-to-help-manage-unique-risk/>

因我不同  
成就不凡

始于1845

#### 关于德勤

德勤中国是一家立足本土、连接全球的综合性专业服务机构，由德勤中国的合伙人共同拥有，始终服务于中国改革开放和经济建设的前沿。我们的办公室遍布中国31个城市，现有超过2万名专业人才，向客户提供审计及鉴证、管理咨询、财务咨询、风险咨询、税务与商务咨询等全球领先的一站式专业服务。

我们诚信为本，坚守质量，勇于创新，以卓越的专业能力、丰富的行业洞察和智慧的技术解决方案，助力各行各业的客户与合作伙伴把握机遇，应对挑战，实现世界一流的高质量发展目标。

德勤品牌始于1845年，其中文名称“德勤”于1978年起用，寓意“敬德修业，业精于勤”。德勤全球专业网络的成员机构遍布150多个国家或地区，以“因我不同，成就不凡”为宗旨，为资本市场增强公众信任，为客户转型升级赋能，为人才激活迎接未来的能力，为更繁荣的经济、更公平的社会和可持续的世界开拓前行。

Deloitte (“德勤”)泛指一家或多家德勤有限公司，以及其全球成员所网络和它们的关联机构(统称为“德勤组织”)。德勤有限公司(又称“德勤全球”)及其每一家成员所和它们的关联机构均为具有独立法律地位的法律实体，相互之间不因第三方而承担任何责任或约束对方。德勤有限公司及其每一家成员所和它们的关联机构仅对自身行为承担责任，而对相互的行为不承担任何法律责任。德勤有限公司并不向客户提供服务。

德勤亚太有限公司(一家担保责任有限公司，是境外设立有限责任公司的其中一种形式，成员以其所担保的金额为限对公司承担责任)是德勤有限公司的成员所。德勤亚太有限公司的每一家成员及其关联机构均为具有独立法律地位的法律实体，在亚太地区超过100个城市提供专业服务，包括奥克兰、曼谷、北京、班加罗尔、河内、香港、雅加达、吉隆坡、马尼拉、墨尔本、孟买、新德里、大阪、首尔、上海、新加坡、悉尼、台北和东京。

请参阅<http://www.deloitte.com/cn/about>了解更多信息。

本通讯中所含内容乃一般性信息，任何德勤有限公司、其全球成员所网络或它们的关联机构(统称为“德勤组织”)并不因此构成提供任何专业建议或服务。在作出任何可能影响您的财务或业务的决策或采取任何相关行动前，您应咨询合资格的专业顾问。

我们并未对本通讯所含信息的准确性或完整性作出任何(明示或暗示)陈述、保证或承诺。任何德勤有限公司、其成员所、关联机构、员工或代理方均不对任何方因使用本通讯而直接或间接导致的任何损失或损害承担责任。德勤有限公司及其每一家成员所和它们的关联机构均为具有独立法律地位的法律实体。

© 2024。欲了解更多信息，请联系德勤中国



这是环保纸印刷品