中銀香港(控股)有限公司
BOC HONG KONG (HOLDINGS) LIMITED

人工智能应用风险合规管理办法（摘要）

目的

本管理办法旨在为本集团内部负责任且合乎道德地使用人工智能建立一个框架，以风险为本的方式开发、采购、实施及使用人工智能。考虑到人工智能的快速发展及其对金融服务业的影响，本管理办法将因应合规、行业标准和道德考量而适时作出修订。

人工智能应用的风险管理原则

1. 足够的专业知识及培训

本集团投入适当的资源以持续支持和发展人工智能应用。

2. 人工智能应用的可解释性

本集团按人工智能应用的重要性，采取相称的措施确保人工智能应用能提供合适程度的解释，令相关持份者能基本明白人工智能应用的决策及输出结果。

3. 高质量的数据管理

本集团使用高质量数据，并对数据质量进行检查及监控，以满足数据质量要求及数据质量评估准则及其他人工智能相关的数据准则。

4. 严格评估核实

本集团在人工智能应用投产前均需完成评估核实，包括评估其准确率、可靠性、稳健性及公平性，确保人工智能模型按预期运作才可投产。

5. 人工智能应用的可审计性

本集团为人工智能应用建立及保留足够的稽核纪录及相关文件纪录，以让其使用结果可以接受持续稽查。

6. 第三方供应商监察

本集团对第三方供应商采取有效的管控措施，确保第三方供应商满足本集团关于供应商管理的要求及本管理办法内的风险管理原则。

7. 道德操守及公平性

本集团遵守《银行营运守则》、《公平待客约章》、普及金融的精神、秉持客户为本的原则及其他适用的监管规定和消费者保障原则，以及本集团的企业价值及道德标准。人工智能应用不应涉及违反集团企业价值及道德标准、会明显损害客户作出知情决定的能力或其基本权利、对客户构成严重损害、涉及剥削弱势社群。

本集团亦厘定模型演算法的关键原则，以缓解歧视及对待客户不公平或有偏差等风险，确保人工智能应用能提供客观、一致、秉持道德操守及公平的结果，并遵守有关歧视的法例。

8. 透明度与披露

本集团确保人工智能应用的使用者于使用人工智能应用或服务前已向其作出妥善、准确及易于理解的披露，确保传达的资讯清晰及简单易明，提供适当的透明度，包括但不限于披露服务是以人工智能技术推动并解释人工智能的用途与目的、相关风险及限制。

9. 定期重检及持续监察

基于人工智能模型存在出现「模型漂移」或「模型衰退」的风险，本集团制订适用及与风险相应的监测工具，并为人工智能应用的使用者提供反馈途径。透过数据及/或使用者反馈进行持续监察，并适时分析结果及作出相应跟进。除对模型表现的持续监察，亦会对模型表现是否满足既定标准以及其合规性进行定期重检。

10.数据保护

本集团按照集团内有关资讯安全的政策及管理办法等规定处理，以确保遵守《个人资料（私隐）条例》(《私隐条例》)、私隐专员公署就《私隐条例》提供实务性指引而发布的相关实务守则，及其他适用的监管规定。

11.网络安全

为应对潜在对本集团人工智能应用的恶意网络攻击，本集团配置相应网络安保防御措施，并密切留意新兴的恶意网络攻击手法及适时重检有关安保防御机制，及进行相应提升防御工作。

12.风险缓解和应变计划

本集团对人工智能的活动进行适当的风险控制（例如「人在环中」)，并就人工智能应用制定应变计划，或在必要时暂停人工智能应用并触发回退流程。