

转型中

通过人工智能采纳实现财富来源转型

一种由人工智能赋能的业务主导战略





执行摘要

01 战略背景与问题定义



02 结构类型与风险特征



03 低效量化与控制有效性



04 目标运营模式：向合理性智能转型



05 技术实施架构



06 治理、模型风险与负责任部署



07 实施路线图



08 董事会指标、管理信息与道德监督



09 结构限制与伦理边界



10 结论

附录A：实施手册

附录B：术语表与技术说明

附录C：治理控制矩阵

附录D：运营指标与管理仪表盘示例

执行摘要



变革的必要性

财富来源正面临转折点。对于在香港及其他复杂金融中心运营的银行而言，传统以文件为主导的财富来源流程已难以应对现代反洗钱/反恐融资风险管理在规模、复杂性和举证标准方面的要求。

拟议应对方案

本白皮书主张，财富来源应从静态验证流程重组为受管控的智能能力：一种能够关联记录、映射所有权与控制权、检验经济合理性、从文件中提取证据、识别异常并支持可辩护、基于风险的决策的能力。

为何此事影响超越私人银行范畴

该挑战在私人银行及超高净值财富结构中最为突出——客户关系敏感性、跨境所有权、信托、私人投资工具及数字资产使得财富来源审查在商业层面极为复杂且极易触发升级流程。然而，问题远不止于私人财富领域。在零售银行、商业银行及全功能银行的高风险客户强化尽职调查中，财富来源同样是关键环节——证据质量薄弱、决策标准不一及未解决积压问题可能引发监管风险、审计风险与金融犯罪风险。

香港银行业凭借稳健的财务实力步入2026年，资产负债规模与存款持续扩张，增长动力焕新，并日益聚焦于绩效、增长与信任。

与此同时，私人财富管理、数字资产、代币化与人工智能正成为行业的核心主题。近期行业展望也表明，银行越来越需要将AI嵌入运营各环节，以产生可量化的生产力提升，而非为创新而创新。

传统财富来源方法已无法适应这一环境。以文件为中心的流程往往仅验证记录的存续性，却未能确定财富积累在经济、结构及时间上是否合理。由此产生的控制模式成本高昂、标准不一且难以扩展，还会造成不必要的客户摩擦，引发升级处理过载、误报率高及机构知识碎片化等问题。



Chad Olsen
合伙人，
法务会计主管，香港特别行政区
毕马威中国

人工智能与人类判断的角色

管理得当的人工智能可以支持实体解析、知识图谱构建、财富合理性建模、自然语言处理、异常检测和持续监控。人工智能不应取代专业判断。善加利用，它可以增强机构智能、优先处理分析师的注意力、提高证据质量并加强决策依据的可靠性，同时让人类对解释、升级处理和最终判断负责。

本文结构

本白皮书旨在提供一份实用蓝图，将知情监管流程从文档主导的控制体系重塑为受管控的智能能力。它首先诊断当前知情监管流程的结构性弱点，然后提出目标运营模式、技术架构、治理框架、分阶段实施路线图、董事会级报告模型和伦理边界条件。本文兼顾战略性和实用性：董事会和高级管理层可重点关注治理、指标和伦理部分，而合规、金融犯罪、数据、技术和模型风险团队可将附录作为实施参考。



关键信息概览

战略问题	当前约束	人工智能驱动响应	预期成果
复杂所有权结构	手动重建实体、链接与受益所有权	实体解析与知识图谱分析	提升对控制权、影响力和结构性不透明度的可见性
经济合理性	依赖文件存在性而非财务连贯性	合理性建模、情景测试与市场对标	对申报财富叙述进行更一致的检验
叙述不一致	依赖分析师对“证据”的解读 信息披露与文件	基于AI的证据提取与矛盾检测	更早发现内部不一致性
升级过载	基于规则的警报缺乏工作范围上下文	复合风险评分与工作流程路由	优化优先级并减少低价值升级
治理与保障	自动化风险不透明且审计追踪薄弱	模型验证、可解释性、漂移监控与覆盖日志	监管信心与可证明的有效性

01

战略背景与问题定义

香港的财富来源挑战是战略性的，而非程序性的。这座城市融合了深厚的银行流动性、成熟的私人财富架构、不断增长的家族办公室生态系统以及日益活跃的数字资产活动。

这些特性在商业上是积极的，但也增加了判断客户积累财富是否合法、连贯且与已知经济状况一致的复杂性。

许多传统财富来源框架的战略弱点在于，它们将财富来源视为文档合规性任务。这种方法容易陷入文档幻觉：证据数量可能增加，但对相关情况的理解并未同步提升。更有效的框架会评估，在考虑所有权结构、交易历史、市场状况、行业经济、司法管辖区风险和行为模式后，所声明的财富积累是否合理。

根据《打击洗钱及恐怖分子资金筹集条例》及相关监管指引的监管要求，金融机构需在风险较高时实施强化尽职调查，了解资金来源和财富来源，持续监测客户关系，并展示基于风险的比例原则。

监管重点日益转向有效性：机构能否证明控制措施在实践中有效，而不仅仅是存在相关程序。

采用财富来源人工智能并非单纯的技术解决方案，而是以技术为驱动的商业战略



02

结构类型与风险特征

香港的财富来源审查中常见合法但复杂的财富类型。风险并非复杂性本身的存在，而在于机构能否充分理解、解释并监控这种复杂性。

下表展示了常见的财富来源类型，以及机构在评估所申报财富叙述是否合理时需要回答的分析性问题。

类型	典型结构	财富来源挑战	AI相关技术
中国内地出境客户	客户运营公司、离岸SPV、香港控股工具、股息或退出收益	重建估值历史、验证注资及文件合法性	图形分析、登记处链接、估值基准
家族办公室与信托分层	全权信托、保护人、基础特殊目的公司及跨境受益人	区分历史性财富转移与不透明重组	实际受益人映射及叙述一致性检验
私募股权与附带权益		根据基金及市场表现检验累计附带权益	合理性建模与同行基准比较
房地产投资组合扩张	历史收购、再融资、租金收入与再投资	评估增值轨迹与融资能力	历史价值重建、市场指数比较与异常评分
数字资产积累	钱包持有量、交易所兑换、托管与法币退出路径	追溯来源、波动性与交易所司法管辖区风险	区块链分析、交易图谱构建与异常检测



03

量化低效与控制效果

传统的SoW处理方式扩展性较差。

传统的SoW审查变得低效，因为每增加一个实体、关系、文档和返工周期，都会增加一层人工工作。工作量不仅取决于结构中的参与方数量，还取决于它们之间的连接数量、验证这些连接所需的证据，以及分析师必须向客户澄清或索要额外文档的次数。

简而言之，审查工作随四个因素增加：

实体数量、所有权或控制关系的数量、文档数量以及审查周期数。一个简化的表达方式是：

$$\text{审查工作量} \approx \text{实体} + \text{关系} + \text{文档} \times \text{审查周期}$$



工作量为何会累积

SoW审查变得低效，并非仅仅因为实体数量增多。而是因为每个实体都会产生需要验证的关系，每个关系都需要证据支持，而证据中的每个缺口都可能触发新一轮客户沟通周期。这就是为什么减少返工和提升证据质量比单纯增加分析师数量更有影响力。

这之所以重要，是因为结构复杂性与文档返工会相互加剧。随着所有权结构层层深入，分析师需要验证更多实体与更多关系。当证据不完整或存在矛盾时，每一轮新增的核查周期都会成倍增加分析师的时间投入，并加剧客户摩擦，却往往无法带来相应比例的风险洞察增量。

试想一个涉及一位主要客户、一个信托、三个中间特殊目的载体、五个运营子公司及多个资产持有实体的高风险私人银行案例。

即便不考虑交易环节的复杂性，该审查也可能涉及超过15个节点与20条关系连线。

如果每个实体都需要进行注册核查、负面媒体报道审查、所有权验证和文档核对，工作量就会迅速膨胀。当文档请求需要循环三到四次时，运营负担可能会变得与由此产生的增量风险洞察不成比例。

静态的《工作范围说明书》还存在一个数学层面的缺陷。如果分析师在孤立情况下发现结构异常的概率为60%，那么发现两个独立异常事件的概率仅为36%。结构越复杂，遗漏关联性、资金流向或表述矛盾带来的累积风险就越高。人工智能虽然无法消除这种不确定性，但可以跨数据集聚合微弱信号，提高异常情况被提交至专业审查环节的概率。

04

目标运营模式： 转向合理性情报

目标运营模式将财富来源审查从文件核对工作转变为以合理性为主导的情报能力。机构仍应收集并保留财富来源声明的证据，但核心问题应转变为：在综合考虑所有权结构、经济逻辑、交易行为、时间节点及现有外部证据的情况下，客户的财富轨迹是否连贯一致

这需要改变思维导向：从追问是否已收到所需文件，转向追问证据是否解释了财富随时间产生、转移、积累和控制的过程

运营模式原则

目标模型不应要求分析师为了收集证据而收集证据，而应帮助他们判断哪些证据重要，这些证据如何与客户的财富故事相关联，故事是否仍然合理，以及在做出决策前需要哪些判断。

下表总结了这一转变如何将财富来源审查的重点从传统管控活动转向目标情报成果

传统管控导向

目标情报导向

收集文件 重建财富轨迹

验证形式与完整性

评估经济与结构合理性

阶段性审查

持续监控并触发重新校准

基于孤立标记进行上报

利用复合风险情报确定优先级

依赖分析师记忆

通过图谱和审计追踪保存机构知识

实际操作意义在于，SoW团队应围绕信息生成而非文档处理来组织。这需要可靠的数据基础、可复用的分析资产、清晰的分析师工作流程以及治理控制，以确保决策能够被解释、质疑和重建。第5节提出了支持此目标模型的推荐技术架构。



05

技术实施架构

一个强大的人工智能支持的SoW架构应作为集成智能系统而非点解决方案来实现。该架构需要五个相互加强的层级：数据摄取与规范化；实体解析与知识图谱构建；分析

建模；决策编排；
以及治理、审计和模型风险管理。

数据来源类别	示例数据元素	刷新模式	主要技术风险
客户主数据与KYC	客户标识符、申报职业、预期活动、风险评级、客户经理备注	事件驱动及周期性刷新	过时的档案、不一致的分类以及缺失的实际受益人字段
交易系统	支付流水、证券活动、存款、取款、交易对手、渠道及币种	每日或近实时	分散的账户视图及不一致的交易对手标准化
企业注册机构	董事、股东、注册日期、注册地址、备案历史	批量刷新及可用的事件警报	延迟、境外覆盖不全及名称差异
文档存储库	信托契约、销售协议、财务报表、税务申报、估值报告	关于数据摄取与案件更新	光学字符识别错误、元数据缺失与非结构化叙事模糊
外部情报	制裁、政治公众人物、负面媒体、诉讼、市场价格、区块链分析	依赖供应商；每日至持续更新	误报、许可限制与可解释性局限

在实施层面，架构应作为受治理的数据与分析管道运行，将原始源信息转化为可控、可追溯且可复用的分析输出。

原始源数据应首先保留其原始形式，然后进行清洗、标准化和质量控制，再整理成分析数据集，用于SoW图形绘制、合理性建模和个案管理。

每次转换应保留谱系、源系统标识符、时间戳、数据质量评分和访问控制元数据，以便每个模型输出都能追溯至生成它所用的证据。

实际实施应遵循六个受控构建步骤：捕获源数据、保存并标准化记录、应用质量与完整性检查、创建可复用的分析特征、注册并治理模型、监控输出与人类决策。这些步骤确保架构不仅是模型的集合，而是具备明确所有权、证据、血缘与治理能力的可控运营能力。附录A提供了详细的实施指南。

5.1 实体解析：关联记录 指代同一人、公司或结构的记录

实体解析是确定不同记录是否指代同一人、公司、信托、账户或资产的过程，即使名称、地址、标识符或文件不完整、不一致或以不同格式记录。在财富来源（SoW）中，此功能是基础，因为如果机构无法首先了解哪些记录、实体和关系属于同一整体，则合理性评估将不可靠

匹配应分阶段进行。应尽可能使用精确标识符，同时基于相似性和上下文的匹配应有助于检测跨名称、地址、董事、实益所有人、交易对手、账户和支持文件之间的关联

不确定的匹配应保持由分析师审查，以便系统在不产生错误关联的前提下提高结构可视性

附录B提供了详细的实体解析术语和说明性评分逻辑。

5.2 知识图谱分析：让 所有权、控制权和流向可视化

知识图谱分析使机构能够将个人、公司、信托、账户、资产和交易表示为相互关联的关系，而非孤立记录。这有助于更清晰地了解谁拥有、控制、影响或资助某个结构，以及这些关系是否与客户声明的财富来源一致。

图谱应区分法律所有权、经济所有权、控制权、授权、交易流向和咨询影响力。同时应保留时间和证据信息，以便分析师能够检验客户在声称财富产生的时间点是否实际控制某一实体或资产。进一步的图谱术语在附录B中解释。

5.3 财富积累合理性 建模

财富合理性建模将财富来源（SoW）的焦点从“是否存在文件”转向“声明的财富轨迹在经济上是否一致”。该模型应将客户声明的财富积累与相关证据进行对比，包括收入、企业所有权、流动性事件、房地产升值、投资回报、杠杆、税收流失和消费假设

输出应基于概率的评估，而不是简单的合格或不合格结果。如果声明的财富超出合理范围，或新证据削弱了叙述的可信度，则应优先对案件进行强化审查。关于财富积累建模、蒙特卡洛模拟和贝叶斯更新的技术说明见附录B

5.4 NLP证据与叙事 一致性。

自然语言处理（NLP）使机构能够将客户的解释和支持文件转化为结构化证据。在财富来源分析中，NLP可以识别信托契约、销售协议、财务报表、税务申报和估值报告等文件中的姓名、公司、日期、金额、司法管辖区、所有权角色、交易描述和法律关系。

这一能力对于叙述一致性尤为重要。提取的事实可与登记记录、合同、交易历史和市场数据进行比较，以识别冲突的日期、不一致的金额、缺失的所有权链接或缺乏支持的财富事件。

NLP输出应保持证据关联性，以便分析师能够将任何提取的事实或总结追溯至源文件。更多技术细节见附录B。

5.5 异常检测：优先关注 异常结构与行为

异常检测有助于识别与同类主体存在显著差异的客户、结构或交易。

这在SoW中十分有用，因为许多相关的金融犯罪风险并未在历史数据中标注，只有当结构复杂性、交易行为、司法管辖区风险、证据缺失和财富叙述等因素被综合考量时，这些风险才可能显现。

异常评分应用于优先安排审查，而非自动确定结果。分析师应能理解哪些特征使某个案例显得异常，以及这种异常是否可由合法的客户情况解释。附录B提供了更多术语说明。

5.6 复合SoW风险评分与 工作流路由

复合SoW风险评分应结合多个分析信号，而非依赖单一模型。

相关输入可能包括图谱透明度、经济合理性、证据充分性、叙述一致性、司法管辖区风险、不利情报、交易异常和数字资产溯源。其目的是支持一致的优先级排序和工作流路由，而非取代专业判断。

每个评分组成部分应具备可解释性。分析师应能查看哪些因素推动了评分、使用了哪些证据、模型输出是否被接受或推翻，以及支持最终决策的人工判断依据。附录B提供了关于复合评分和人在回环审查的更多技术说明。





06

治理、模型风险与负责任部署

AI驱动的SoW应作为受控的决策支持能力进行治理，而非自动决策引擎。机构必须能够解释数据的来源、记录的关联方式、模型输出的生成过程、支持结果的相关证据，以及在做出任何最终决策前如何运用人工判断。

治理框架应基于五项核心原则：

可解释性、可审计性、可重复性、隐私保护与人工问责。这些原则对于维持监管信心、避免AI在已复杂的SoW流程中引入不透明风险至关重要。



治理原则

AI驱动的SoW应提升而非降低决策的可解释性。每个模型输出、图形信号、提取事实和分析师调整都应可追溯至其支持证据、假设及人工判断。

6.1 治理原则

AI驱动的SoW应遵循与其他高影响力金融犯罪及模型风险能力相同的规范。每个模型、评分、图形输出或基于NLP的提取都应具有明确定义的目标、所有者、经批准的用例、操作限制及控制证据。

机构应能够证明使用了哪些数据、应用了哪些模型或规则、如何设定阈值、哪些证据支撑了输出结果、分析师是否接受或否决了该输出结果，以及最终决策是如何记录的。这确保人工智能强化问责制，而非模糊其责任。

6.2 模型风险管理

模型风险管理应贯穿整个生命周期：设计、开发、验证、部署、监控和退出。治理流程应评估每个模型在概念上是否合理、验证是否充分、运行是否稳定，以及是否符合预期的财富来源（SoW）用例。

验证不应仅限于技术准确性，还应考虑模型是否为分析师提供有用、可解释且比例合理的输出结果。模型所有者应维护相关文档，涵盖预期用途、禁止用途、假设条件、局限性、训练数据、验证结果、监控阈值以及升级流程。

6.3 可解释性与人类问责制

可解释性应在三个层面运作。案件层面解释应说明为何特定客户或结构被标记。模型层面解释应展示模型在整体组合中的行为表现。图谱层面解释应揭示构成风险信号的具体实体、关联关系和路径。

人类判断仍居核心地位。AI应通过更清晰地呈现相关证据来支持分析师审查，但不应自动结案复杂的SoW案件。

若分析师覆盖了模型输出结果，应记录并监控其理由，以便通过重复出现的覆盖模式为未来模型改进提供依据。

6.4 隐私、安全与数据管控

隐私和安全控制应从架构设计之初就纳入考量。应最大限度减少敏感个人数据的使用，并实施加密、访问控制，且仅用于规定的财富来源及金融犯罪目的。

如果建模不需要直接标识符，机构应考虑采用令牌化或假名化技术

在需要跨境分析的情况下，部署前应评估数据驻留、聚合阈值、访问限制和法律约束。核心治理问题不仅在于模型是否有效，更在于用于运行模型的数据是否以合法、相称且安全的方式进行处理。

6.5 持续监控与保障

部署后，基于AI的SoW应持续监控。监控应涵盖数据质量、模型漂移、覆盖控制率、评分分布变化、延迟、系统可用性、可解释性差距及案件结果。

当监控发现性能退化或异常行为时，机构应在重新校准、重新训练或调整阈值前调查根本原因。因此，治理应被视为一种操作纪律，而非一次性审批步骤。

附录C提供了详细的治理控制矩阵，包括模型所有权、验证、可解释性、隐私、漂移监控、覆盖控制及可审计性的最低控制证据。

总之，负责任的部署不仅需要技术性能，还需要建立治理框架，使机构能够向分析师、审计师、高级管理层和监管机构解释、证明并捍卫由人工智能支持的财富来源决策。附录B提供了详细的验证和监控指标。

07

实施路线图

SOW计划是一个持续的用户旅程，而非传统的转型项目

实施应分阶段受控推进

机构应在解决基础数据、分类法和治理问题之前，避免试图实现端到端自动化。分阶段路线图能让机构在保持监管信心的同时产生可量化的收益

项目不应以牺牲顺序为代价而加速推进。AI赋能的SOW依赖于可靠的数据、清晰的分类法、证据溯源、经过验证的模型、分析师采纳和治理管控。如果这些基础薄弱，更快的部署可能只是将不一致性自动化、造成虚假信心、增加客户摩擦，或产生无法向高级管理层、内部审计或监管机构解释的输出。因此，目标应当是受控加速：尽早交付可见收益，同时确保每个版本在扩展下一个能力之前，都经过充分治理、测试并在运营中嵌入

路线图应结合两个视角。首先，机构应通过产品式发布交付AI驱动财富来源管理，每次发布都包含明确的用例、工作流程、管控节点和成功指标。其次，这些发布应置于长期成熟度路径中——从基础数据与图谱能力，逐步延伸至持续监控和企业级优化。

董事会可见性的重要性

分阶段实施路线图的价值，在于董事会和高级管理层能够看到它是否同时改善了控制有效性和客户成果

这在超高净值客户领域尤为重要——该领域的SOW审查通常具有商业敏感性、以关系为主导且极易升级。一次不当的信息请求、无法解释的延迟或前后不一致的风险决策，都可能迅速升级为CEO级别的问题，即使该机构的底层控制担忧是合理的

因此，需要董事会层面的可见性来确保AI赋能的SOW能力不仅得以交付，还能减少摩擦、提高证据质量、支持可辩护的决策，并保护机构在其最敏感客户群体中的声誉

7.1 产品式发布实施 版本发布

从交付角度来看，机构应通过产品式发布而非整体转型来实施该计划。每个发布版本应包含明确的风险用例、映射数据要素、目标用户工作流程、控制要求、验证标准及业务效益衡量指标。这有助于降低构建技术上令人印象深刻但未被分析师采纳或未得到第二道防线、内部审计及监管

机构认可的模型的风险。

以下发布序列展示了如何在不超过数据、治理及分析师工作流程支持能力的前提下，渐进式构建能力。

发布版本	技术构建	控制关卡	成功指标
版本一：图表基础	实体解析、图谱模式、注册表接入与可视化	数据血缘、访问控制及匹配质量测试	图中所示高风险案例的百分比
版本二：证据智能	自然语言处理提取、文件分类及证据充分性评分	光学字符识别准确性、提取验证及人工审查抽样	减少重复文件索取
版本三：合理性引擎	财富合理性建模、贝叶斯更新及同行基准比较	假设审查、敏感性测试及验证签核	分析师优先排序质量的改进
版本4：异常与监控	异常模型、漂移仪表板和事件触发器	误报审查、漂移阈值和警报治理	减少低价值问题升级
版本5：集成编排	复合评分、工作流路由、覆盖捕获和重大事件仪表板	用户验收、模型风险审批和审计追踪审查	工作声明周期时间和证据质量的显著提升

7.2 示例性实施路线图

以下路线图为示例性而非规定性，具体排序将根据机构的数据成熟度、风险偏好、运营模式和技术架构而变化

时间框架	实施重点	关键交付成果	成功衡量标准
0-3个月	基础与动员	用例选择、数据清单、治理设计和目标运营模式	确定优先用例、识别数据差距并指定负责人
3-6个月	图表与证据基础	实体解析试点、所有权关系图原型和证据存储库	提高结构可见性并减少手动映射工作量
6-12个月	证据智能与合理性测试	NLP提取、矛盾检测与财富合理性模型	更快的文档审阅与改进的叙事一致性检查
12-18个月	风险评分与工作流集成	综合SoW评分、案件优先级排序与分析工作流整合	更好的案件分流与缩短审查周期
18-24个月	持续监控与优化	漂移监控、覆盖分析、反馈循环及类型更新	持续保障、提升模型性能及强化治理报告



08

董事会指标、管理信息与道德监督



董事会应区分两类指标：一类是帮助管理层运行财富来源（SoW）流程的运营指标，另一类是表明项目是否在提升管控有效性、客户成果、治理纪律、模型保障、道德部署及监管可辩护性的董事会级指标。在私人银行业务中，董事会层面的透明度尤为重要，因为财富来源审查可能直接影响超高净值客户关系、高级别上报机制及品牌声誉。

管理层应持续跟踪详细的运营指标，例如财富来源审查周期时间、重复文件请求次数、未处理积压账龄、图谱覆盖率、自然语言处理提取质量、分析师手动覆盖比率、数据源问题及模型漂移指标。这些指标对于运营流程和改进采纳至关重要，但在上报至董事会前应进行筛选提炼。附录D提供了说明性运营指标和管理层仪表板示例。

在更广泛的银行业背景下，财富来源审查作为高风险管理客户强化尽职调查的一部分同样重要——因为证据质量薄弱、决策不一致或积压问题未解决可能引发监管、审计和金融犯罪风险敞口。

这些董事会级指标不仅仅是绩效衡量标准，更是防范过度自动化、虚假精准度及对客户造成过度影响的保障机制。它们也为界定人工智能驱动财富来源审查的道德边界奠定了基础——下一章节将对此展开论述。

董事会关注点	指标	董事会问题
控制有效性	财富来源审查结论质量、上报处理结果及高风险案例优先级排序	我们是否更一致地识别出真正的高风险案例？
客户与关系影响	重复申请率、处理周期、投诉主题、高级别升级、EDD积压及高风险客户延迟趋势。	我们是否在减少对超高净值客户的不必要摩擦，同时提升高风险客户的及时一致处理？
高风险客户EDD有效性	高风险客户的SoW完成率、逾期EDD审查、未解决证据缺口及升级结果。	SoW控制是否提升了高风险客户群体EDD的质量与时效性？
证据质量	证据充分性、叙事一致性及无依据结论的减少。	决策是否由更优证据支撑，而非仅依赖更多文件？
治理与问责	覆盖趋势、可审计性、已批准用例及控制例外	管理层能否解释AI辅助决策的达成过程？
模型保证	重大模型独立验证比例、未闭环高风险发现及修复状态	AI驱动的SoW模型是否通过独立验证并在获批限额内运行？
模型可靠性	漂移、校准违规、评分分布变化及监控异常	当类型学和客户行为变化时，模型输出是否仍可靠？
道德部署	偏见或代理测试、隐私事件、不恰当的去风险指标以及未解决的可解释性问题。	人工智能工具是否被适度、公平地使用，并有人类问责制？
监管辩护能力	监管反馈、审计发现和控制有效性的证据。	我们能否向监管机构和内部审计证明其有效性？



09

结构限制与伦理边界

人工智能赋能的工作范围（SoW）应被视为强大的决策支持工具，而非法律判断、监管问责或专业质疑的替代品。这些工具若能提升可见性、一致性和优先级排序，但在被视为确定性结论、不加节制地使用或用于合理化生硬的客户结果时，也可能带来新的风险。

因此伦理边界明确：人工智能应帮助机构提出更优问题、揭示不确定性并记录推理过程，但不应将复杂的工作范围判断转化为自动化结论。

9.1 人工智能无法判定合法性或犯罪性。

人工智能无法孤立地判定犯罪性或合法性。非法财富可能嵌入看似合法的结构中，尤其是在执法情报、跨境信息或受益所有权透明度有限的情况下。反之，当登记数据不完整、所有权安排复杂或模型不熟悉商业惯例时，合法财富也可能显得异常。因此，模型输出应被视为决策支持信号，而非法律、监管或合规结论。

9.2 虚假精度与概率评分

虚假精度的风险是实质性的。模型生成的可能性评分看似客观，但结果取决于假设、数据质量、先验观点、基准选择及模型设计。数值置信度不应取代专业怀疑精神。

机构应禁止自动关闭复杂工作范围（SoW）案件，并要求在做出最终升级、接受或退出决策前进行人工推理。

9.3 人类判断与个案分析

人类判断仍是有效工作范围的核心。

分析师必须对证据解读、质疑模型输出、记录不确定性以及结合个案背景保持问责。在超高净值（UHNW）案件中尤为关键，因为商业历史、家族安排、地缘政治风险、法律架构和关系敏感性都可能影响证据的解读方式。人工智能应使不确定性更易于识别和解释，而非消除判断的必要性。

9.4 隐私、相称性与公平性

合乎道德的部署需要隐私性、比例性、可解释性和公平性。系统应仅将数据用于定义的工作范围（SoW）和金融犯罪目的，尽量减少不必要的个人数据处理，保护审计轨迹并实施适当的访问控制。还应测试系统是否存在不合理的偏见和不恰当的代理变量，尤其是在输出结果可能影响客户待遇、升级处理、开户或关系决策的情况下。

9.5 人工智能绝不能成为去风险化的捷径。

机构应警惕将人工智能用作广泛去风险化的捷径。高异常分数、不完整数据或不透明结构可能证明需要加强审查，但不应自动证明终止关系或拒绝某一客户类别是合理的。机构必须区分复杂性与可疑性，并确保决策保持比例性、基于证据且针对具体案例。



10

结论.

财富来源（SoW）正处于一个转折点。传统模式——收集文件、核对叙事、升级不确定性并重复循环——已不足以应对现代银行业的复杂性。对于高风险客户的尽职调查（EDD）而言，它过于缓慢；对于超高净值（UHNW）财富结构而言，它过于粗糙；对于跨境所有权网络而言，它过于零散；当监管机构、审计师和董事会质疑控制措施是否实际有效时，它又难以提供证据。

仅仅通过增加更多文档要求来应对的机构将无法解决根本问题。

它们将增加成本、客户摩擦和分析师疲劳，而不一定能提升风险洞察。机遇在于将SoW重新设计为一种受管控的智能能力：能够连接记录、映射所有权和控制权、测试经济合理性、从文档中提取证据、识别异常、支持可辩护的决策，并为董事会提供控制措施是否有效的更清晰视野。

人工智能并非终点，而是一个赋能层。使用得当，它可以帮助机构从静态验证转向动态合理性智能；使用不当，它可能自动化不一致性、制造虚假精确度并削弱问责制。差异将取决于治理：数据质量、模型验证、可解释性、人类判断、伦理边界和董事会层面的可见性。

行动号召十分明确。金融机构应停止将财富来源视为后台文档负担，而应将其作为战略性反洗钱/反恐融资情报能力

来对待。他们应优先关注基础数据及其血缘关系，开展能减少摩擦、提升证据质量的试点用例，从一开始就嵌入模型风险和隐私控制，并通过董事会级别的指标报告进展，这些指标需覆盖控制有效性、客户影响、模型保障、伦理及监管合规性。

对香港而言，这不仅是合规改进议程，更是竞争力议程。一个能够融合增长、财富管理专业化、数字创新和可信金融犯罪控制的金融中心，将更有能力维持与客户、监管机构和全球交易对手的信任。重新设计SoW为维护金融诚信和香港作为受信赖的全球财富枢纽角色提供了一条有前景的路径，尤其是在这个人工智能时代。

供审查的指示性源材料

- 2026年香港银行报告
- 香港银行业展望2026
- 香港私人财富管理报告2025
- 香港信托行业焦点2025
- 香港金管局与证监会反洗钱/反恐融资指引及监管材料
- 金融行动特别工作组互评估及有效性方法论
- 关于人工智能、监管科技、监管科技及金融犯罪转型的行业材料

附录A： 实施手册



以下手册展示了第5节所述的受管数据与分析管道如何转化为实际实施步骤。其旨在为技术、合规、金融犯

罪、数据治理、模型风险及内部审计团队提供交付参考。

实施内容。	建设目标	实施方法	控制证据
1. 捕获源数据	针对KYC记录、交易数据、登记簿、文档、外部情报及市场数据的受控采集流程。	在数据用于分析之前，需定义源归属、采集频率、必填字段、系统标识符、文档元数据及访问权限。	数据源清单、数据所有者登记表、采集日志及访问控制记录。
2. 保留并标准化数据	一个受治理的数据管道，在保留原始证据的同时创建清洗和标准化的分析数据集。	将原始记录与转换后的数据分开存储，应用标准分类法，规范化名称和日期，并记录每条转换规则。	谱系图、源到目标映射、转换日志和数据质量评分。
3. 应用质量和完整性检查	一个数据质量控制层，防止不完整或不可靠的数据影响SoW风险结论。	测试完整性、有效性、唯一性、及时性和参照完整性；在运行分析前，将失败的记录路由到异常队列。	数据质量仪表板、异常登记、修复历史和控制证明。
4. 创建可复用的分析特征	用于状态图绘制、合理性建模、异常检测和案例优先级排序的变量特征目录。	定义并对特征进行版本管理，例如所有权深度、实体存续时间、交易速度、司法管辖区风险敞口、证据强度和叙述一致性。	特征目录、特征定义、版本历史记录和审批记录。
5. 注册并治理模型	涵盖实体解析、图分析、合理性建模、自然语言处理提取和异常检测的模型注册表。	记录模型目的、所有者、预期用途、训练数据、假设条件、局限性、验证结果、部署状态和禁止用途。	模型清单、验证报告、审批流程、部署记录和局限性日志。
6. 监控输出和人类决策	一个监控框架，用于追踪模型性能、数据漂移、异常率、分析师覆盖和案例结果。	为漂移、延迟、数据缺失、不明分数变动和覆盖异常设置阈值；在重新训练或更改阈值前需要审查。	监控仪表板、事件单、覆盖日志、重新训练决策和管理信息包。

附录B： 术语表与技术说明



本附录解释了本文中使用的特定技术术语和说明性评分逻辑，旨在为需要额外技术细节的读者提供支持，同时不中断主体内容的叙述。以下示例仅供说明，在实际生产使用前应由各机构进行校准、验证和管理。

B.1 实体解析

实体解析是指判断不同记录是否指向同一人、公司、信托、账户、钱包或资产的过程。由于同一主体可能以不同名称、拼写、语言、标识符、地址或文件格式出现，该过程不可或缺。

B.2 特征工程

特征工程是将原始数据转换为模型可用的结构化变量的过程。在财富来源分析（SoW）场景中，有用的特征可能包括所有权深度、实体存续时间、管辖区域暴露度、交易频率、证据强度、叙述一致性、来源可靠性和关系重叠度。

B.3 知识图谱

知识图谱将个人、公司、信托、账户、资产和钱包表示为相互连接的实体。这些实体通常称为节点，它们之间的关系称为边。在财富来源分析中，知识图谱能比静态表格或案件档案更清晰地展示所有权、控制权、影响力、交易流向和结构透明度。

B.4 自然语言处理与 证据提取

自然语言处理（NLP）使机构能够将客户陈述和支持文件转换为结构化证据。在财富来源分析场景中，NLP可以从信托契约、销售协议、税务申报表、财务报表和估值报告中提取姓名、公司名称、日期、金额、管辖区域、所有权角色、法律关系及交易描述等信息。

B.5 异常检测

异常检测可识别与可比同行存在显著差异的客户、结构或交易。在财富来源分析中，异常特征可能包括异常复杂的所有权结构、资金快速过手、高司法管辖区复杂性、无法解释的第三方融资、与申报收入不符的资产增长、数字资产来源缺失，以及交易行为与所陈述的财富叙事不一致。异常评分应优先用于审查，而非自动决定结果。

B.6 综合SoW风险评分

综合SoW风险评分将多个分析信号整合为单一决策支持视图，可能包括：图谱不透明度、经济合理性、证据充分性、叙事一致性、司法管辖区风险、负面媒体关联度、交易异常性及数字资产溯源。

每个评分维度均应具备可解释性，分析人员需可查看哪些因素驱动了评分、使用了哪些证据，以及模型输出结果是否被采纳、质疑或覆盖。

B.7 人在回路审核

人在回路审核意味着人工智能支持而非替代专业判断。分析师仍需负责解读模型输出、审查证据、消除不确定性并记录决策依据。这在财富来源分析中尤为重要，因为模型输出取决于数据质量、假设条件、阈值设定、基准参照和证据可用性。

B.8 模型漂移

模型漂移是指由于数据、客户群体、市场环境或金融犯罪类型发生变化，导致模型可靠性下降。漂移监测应跟踪特征分布、评分分布、人工干预率、精确率或召回率、新型犯罪类型及客户行为的变化。当发现漂移时，机构应在重新训练或部署模型前审查阈值、假设条件、特征变量和训练数据。

B.9 模型验证与监测 指标。

模型验证与监控应结合技术性能测试与运营及治理审查。相关验证维度包括区分能力、校准、稳健性、漂移、公平性、可解释性和运营性能。

持续监控还应跟踪特征漂移、评分漂移、覆盖峰值、数据缺口、延迟违规和可解释性缺口。每个触发点应有明确的负责人、根本原因审查流程和治理响应措施，包括重新校准、阈值审查、重新训练、事件升级或在适当情况下进行人工审查。

附录 C： 治理控制矩阵



下表将第 6 节中的治理原则转化为实际的控件领域和最低证据要求，旨在为合规、模型风险、技术、内部审计及

高级管理层的监督提供参考。

治理领域	重要性说明	最低控制证据
模型所有权与审批	确保每项 AI 能力都有明确的负责人和已获批准的用例。	模型清单、负责人登记册及审批记录。
独立验证	确认概念合理性、实施准确性和局限性。	验证报告、限制日志及整改方案
可解释性	使分析师、审计员和管理者能够理解输出结果	特征驱动因素、图表原理及证据引用
数据隐私与访问控制	保护敏感客户及关联数据	访问日志、加密控制及数据最小化评估
漂移监控	检测数据、市场及类型变化导致的性能衰退	漂移报告、阈值审核及重新训练决策
覆盖治理	确保人类判断被记录和监控	覆盖日志、理由记录和趋势分析
审计追踪与可复现性	允许输出和决策被重建	版本化数据、模型版本和带时间戳的决策记录

附录D： 运营与管理指标示例



本附录提供了说明性运营指标，管理层可使用这些指标监控支持AI的SoW能力的日常绩效。这些指标应支持管理行动和根

本原因分析，而董事会报告则应侧重于影响控制有效性、客户成果、道德、声誉或监管辩护能力的子集。

运营领域	示例指标	重要性说明	管理用途
案件处理量	平均SoW周期时间、积压时效和排队时间。	显示流程是否正在高效完成清收工作。	识别瓶颈、资源缺口和工作流设计问题。
客户摩擦	重复文件请求率、放弃入职流程及投诉主题。	表明工作说明书是否给客户造成可避免的负担。	改进请求设计、客户经理指导和升级处理。
证据质量	证据充分性评分、缺失证据主题及无依据结论率。	显示审核是否在提升证据质量，而非仅仅增加文件数量。	优化检查清单、证据规则及分析师培训。
图覆盖率	高风险结构已映射且未解析实体链接的百分比。	衡量结构可见性是否在改善。	优先处理注册表摄取、实体匹配和手动修复。
自然语言处理性能	提取准确率、矛盾检测精度和OCR异常率。	显示文档智能是否足够可靠以供分析师使用。	改进提取提示、验证样本和文档处理。
分析师工作流	覆盖率、升级数量和审查队列老化情况。	指示AI输出是否可供分析师使用且可信。	调整阈值、训练和案例路由规则。
模型运维	延迟、可用性、漂移警报及故障数据馈送。	显示该能力在生产规模下是否可靠。	触发事件管理、重新校准或服务补救。
治理运维	验证逾期项目、未解决审计问题及访问异常。	显示控制框架是否按设计运行。	上报逾期保证项目及控制违规。

kpmg.com/cn/socialmedia



如需获取毕马威中国办公室列表，请扫描二维码或访问我们的网站：
<https://kpmg.com/cn/en/about/office-locations.html>

本文件所载信息为一般性质，并非旨在针对任何特定个人或实体的情况。
尽管我们致力于提供准确及时的信息，但无法保证这些信息在收到之日准确无误，或将来持续保持准确。任何人不得在未经全面审查特定情况并获取适当专业意见的情况下，依据此类信息采取行动。

© 2026 毕马威（香港特别行政区合伙制事务所）是毕马威国际有限公司（一家以担保形式成立的英国私营公司）关联的独立成员所组成的全球组织中的成员。保留所有权利。在香港特别行政区印刷。

毕马威名称和标识是毕马威全球组织中独立成员所经许可使用的商标。

出版编号：1787554664433

出版日期：2026年9月