IMF

NOTES

如何让具有代理能力的AI重塑支付方式

由Sonja Davidovic和Hervé Tourpe准备

注意/2026/004

如何让具有代理能力的AI重塑支付方式

注意/2026/004

Sonja Davidovic和Hervé Tourpe

推荐引用：Davidovic, Sonja，Hervé Tourpe. 2026. “How Agentic AI Will Reshape”
\"支付。\" 国际货币基金组织，华盛顿特区，IMF文件2026/004

出版物订单可以通过在线、传真或邮寄方式提交：

国际货币基金组织，出版物服务部，邮编：92780，华盛顿特区，美国，电话：(202) 623-7430，传真：(202) 623-7201，邮箱 publications@imf.org
书店 国际货币基金组织网站（IMF.org）
电子图书馆 国际货币基金组织网站（IMF.org）

内容

如何让具有代理能力的AI重塑支付方式

由Sonja Davidovic和Hervé Tourpe*准备

四月 2026

执行摘要

人工智能（AI）正进入一个新阶段，其中系统可以代表用户自主行动。这些“代理型”AI系统可以解释目标、规划多步骤行动，并在有限的人为干预下与数字服务互动。在支付领域，这种转变可能会将交易发起从明确的人类指令转向代理早期阶段，由技术公司、支付网络和金融机构介导的决策。尽管采用率仍然很低，但这一趋势表明，代理型模型将随着时间的推移变得越来越重要。

此笔记对这些发展情况进行盘点，并探讨代理人工智能可能对支付系统运作的影响，包括授权、流动性管理、结算、合规性和运营弹性。它并不试图得出明确结论或提出规定性政策措施。相反，它旨在提出关键设计问题、架构紧张点和风险渠道，这些在采用过程中可能需要关注。

一个主要挑战是概率性、适应性决策与支付基础设施的确定性要求的互动。为了结构化分析，该报告引入了一个三层概念框架，将（1）意图形成和协调、（2）授权和控制以及（3）结算分离开来。该框架是一个规范性的分析视角，由新兴实践提供信息，以明确代理能力可以在哪些方面有效地运作，以及在哪些方面基于规则的安全措施仍然是必需的。

笔记回顾了在支付中使用代理AI的潜在用例，并突出了与授权可追溯性、不透明度、相关代理行为、网络安全以及未解决的法律法规和责任问题相关的风险。它还讨论了新兴的缓解方法，包括基于指令的授权、决策与执行的架构分离、代理身份框架、可编程支付控制、审计跟踪以及分层人工在支付中的作用，这不仅取决于技术，还取决于闭环模型。笔记认为，随着实验的继续，代理AI对机构设计和治理选择的影响也同样重要。

* 以下是作者想要感谢的同事们，他们为有益的研究材料、讨论和评论提供了帮助：Anita Angelovska Bezhoska, Marianne Bechara, Alexander Copestake, Clement Couchevellou, Era Dabla-Norris, Jose Deodoro, Jose Garrido, Kathleen Kao, Yosuke Kido, Pearl Kuebel, Baoping Shang，以及Alexandre Balduino Sollaci。

引言

人工智能（AI）正迅速地从辅助金融决策的工具演变为能够代表经济行为者进行行动的技术。新一代的代理型AI系统——能够解读目标、规划多步行动并自主与数字服务互动的软件代理——开始在电子商务和金融市场涌现。在支付领域，这些系统不仅能够推荐交易，还能够根据授权启动、协调和管理金融操作。这一发展可能代表了金融系统架构的结构性转变。从历史上看，支付基础设施的设计是基于个人发起的指令和由确定性系统处理的指令。从卡网络到实时全额结算（RTGS）系统，支付轨道依赖于可预测的规则、法律确定性以及明确的问责结构，以确保信任和金融稳定。代理型AI向这一框架引入了一个新元素：能够以机器速度发起金融行动的概率决策系统。

近期行业发展趋势表明，这一转型已经开始。主要技术和支付公司正在尝试通过代理中介进行商业和支付流程，同时，如通用商业协议（UCP）、代理支付协议（AP2）、代理间（A2A）通信框架以及模型上下文协议（MCP）等新技术标准正在迅速出现，以实现…… c. 市场预测表明 在自主代理和支付基础设施之间实现互操作性，通过代理中介的商务活动有望在未来十年内产生显著的经济活动，可能改变消费者、企业和金融机构与支付系统的互动方式。

这些发展提出了重要的政策、建筑设计和风险管理问题。支付系统必须调和两种截然不同的设计逻辑：具有适配和概率特征的代理人AI系统和具有确定性要求的金融市场基础设施建设。如果没有适当的保护措施，将支付起始任务委派给自主代理人可能会带来新的运营、法律和系统性风险，包括利益不一致、模型错误以及各市场中高度关联的自动化行为。

鉴于代理人工智能在支付领域的早期和有限的采用，本报告不寻求得出明确的结论或提出规定性的政策措施。相反，它梳理了新兴的发展和实验，并引入了一个概念框架来构建关于代理能力如何随着时间的推移与支付系统互动的讨论。报告回顾了潜在的应用场景，突出了关键的设计问题和与支付启动自动化程度提高相关的风险，并概述了随着采用的发展可能需要关注的治理和架构考虑因素。

背景

动力与范围

es引发了一系列设计和政策问题。 快速出现的代理中介支付使用案例超出了现有支付框架的构建初衷。这些设置并非用户主动发起交易，而是依赖于在委托授权下运作的软件代理，预测支付需求，评估选项，并在多个工具和渠道之间协调执行。在某些情况下，它们甚至被授权做出支付决策。

这种演进可以描述为从明确由人类发起的交易（“点击支付”）转向由代理中介的决定过程（“决策支付”），其中执行越来越多地以机器速度发生，并跨越多个支付价值链层次，受预定义的目标、限制和管理安排的约束。

越来越多的行业参与者，包括支付网络、技术平台（例如，像以太坊这样的开源软件）和AI模型提供商，正在竞相尝试这些功能。尽管当前的应用仍主要集中于改进人们寻找和比较产品的方式，但这些应用正扩展到广泛的支付相关用例中，从欺诈检测到…… 实验现在正迅速转向检测和合规监控，以及资金优化和跨境支付协调，反映出在整个生态系统中代理型人工智能试点范围正在扩大（波士顿咨询集团2025）。当前创新的快速速度意味着一度需要数年的发展，现在可以在几个月内实现，这表明更迅速和实质性的变革可能即将到来。

在本文的背景下 一个中心性的建筑挑战随之产生。核心支付基础设施建立在确定性逻辑之上，需要每个交易生命周期阶段的可预测性、可审计性和法律可执行性。相比之下，代理人工智能系统依赖概率推理和适应性决策，这在相似条件下可能产生不同的结果。本文分析了这些本质上不同的特性如何调和，并提出一个三层模型来评估将代理人工智能整合到支付流程中的适当范围。

什么是代理人工智能？

自主人工智能系统指的是能够感知环境、设定目标并执行多步骤任务的系统，可能跟踪家庭燃气罐。 with minimal human intervention. For example, in the commercial sector, such systems at this level identify the most cost-effective propane supplier and automatically arrange for refills. Within warehouse operations, these technologies can detect incoming deliveries, retrieve corresponding billing information, and initiate payment instructions in accordance with predefined protocols. Furthermore, the Bank for International & prioritizes payments within 《结算》（BIS 2025）指出，人工智能代理可以独立管理实时全额结算系统，有效地反映了既定的审慎现金管理实践。

预测（见表1）或需要持续的人类 与传统的静态参与型AI模型不同，智能体AI系统整合了诸如规划、动态适应和工具编排等高级功能，因此这些系统能够在复杂生态系统中作为自我指导的智能体运行（FinRegLab 2025）。技术协议和标准正在迅速演变，以便AI智能体能与任何支付参与者及数据源进行交互（表2）。

简述AI在支付领域的兴起

在20世纪80年代，通过将人类逻辑编码为静态的“如果-那么”规则（见表1），早期人工智能在支付领域的应用开始出现。一个典型的例子是美国运通公司的授权助手，它自动化了信用授权的部分环节。尽管在受控环境中效果显著，但这些系统在应对恶意行为快速演变、规则更新速度无法跟上时，证明并不适用。

随着20世纪90年代交易量的增长，支付提供商越来越多地转向统计机器学习。像Hecht-Nielsen神经计算公司软件的Falcon欺诈管理器和Visa的实时授权工具这样的系统，利用神经网络在规模上评估欺诈风险。[1] 这些方法将概率推理引入了支付流程，但它们仍然局限于检测和评分，而不是采取行动解决问题。

[1] 神经网络是一类受生物神经元启发的机器学习模型，通过加权的连接而不是明确的人类定义的规则，从大量数据集中学习统计模式。在20世纪90年代，它们是第一批能够随着交易模式的发展动态调整欺诈检测模型的技术之一。

随后的浪潮，包括21世纪的基于图的分析和2010年代的深度学习，进一步提升了欺诈检测、身份验证和风险管理，支持了生物识别支付和令牌化等创新。[2] 在这些阶段中，人工智能在评估支付决策方面取得了实质性改进，但支付发起和结算仍然明确由人类驱动。

表1. 人工智能在支付系统中的应用演变

时期	关键技术	创新（示例）	支付功能
1980年代	专家系统	美国运通授权人助手	基于规则的信用授权和欺诈检测使用简单的“如果-那么”逻辑。
十九世纪九十年代	机器学习，早期神经网络	猎鹰欺诈经理 卷积神经网络用于支票处理	概率欺诈评分和模式识别 交易层面风险评估
2000年代	图分析，异常检测	PayPal伊戈尔和伊利亚	网络基础欺诈和风险评估 账户与设备间的关联图
2010年代	深度学习、生物识别	苹果支付，Stripe雷达	生物识别认证，令牌化，高级欺诈检测、人工智能辅助的KYC和身份验证。
2020年代	基础模型，对话式人工智能	你好，基于语音的UPI付款，印度）	人工智能介导的支付发起、客户互动与工作流程自动化；执行保持确定性。

来源：作者。注：AI = 人工智能；KYC = 了解你的客户；UPI = 统一支付接口。

新一代人工智能系统，称为大型语言模型（LLMs），驱动着本质上是非确定性的AI代理：它们不是为给定的输入产生单一、可重复的输出，而是通过从可能的下一个标记的概率分布中进行采样来生成响应。[3] 因此，相同的提示可能会产生不同的输出。这一特性对代理人工智能在支付链中的信任位置和方式具有根本性的影响，尤其是在授权、执行和结算功能方面。[4]

非决定性也通过幻觉表现出来，其中模型生成看似合理但实际上错误的陈述。尽管连续的模型生成显示出幻觉率的降低，但风险尚未消除，并且仍然是支付、合规和结算环境中的关键关注点（Omar等人，2025年）。

[2] 深度学习是指一类基于多层（或“深度”）神经网络的机器学习技术，它能够从大量数据中自动学习层次化的表示。在2010年代，计算能力、数据可用性和训练方法的进步使得这些模型在图像、语音和行为模式识别等任务上超越早期方法。

[3] 在人工智能的背景下，“token”指的是模型处理文本的单位，不应与分布式账本技术（DLT）背景下的“tokenization”混淆，在该背景下，该术语表示在账本上表示资产、权利或价值。

[4] 一个关键的控制是“温度”，这是一个人工智能设置，控制输出变异性。低温使模型偏向最可能的标记，增加可预测性；高温则允许更大的探索和多样性。在支付和金融环境中，准确度、可审计性和法律确定性至关重要，因此通常将温度设定得非常低。

值得信赖的代理商

人工智能代理因此非常强大，但受其概率性质所限，限制了它们在支付链中某些任务的应用。相比之下，如RTGS系统、卡网络、即时支付平台和分布式账本等支付基础设施则运行在确定性基础上：[5] 交易遵循预设规则，结果为二进制，法律最终性得到保障。[6]

这种差异创造了一个结构性的挑战。与代理支付相关的风险主要并非来自在不适当的... 概率推理本身，但并非让自适应系统进行不可逆的控制、检查或问责。因此，核心问题不在于是否应在支付中使用AI——因为AI已经超过40年用于支付（见表1）——而在于如何将不确定的、概率性的决策与自动支付执行区分开来。下一节将引入一个三层模型以阐明这些角色和架构。

人工智能支付层模型

传统支付系统将授权直接与用户或机构的明确指令相关联，用户或机构开始支付，然后该支付被与既定规则核对。相比之下，代理支付模型依赖于目标和限制来表达意图。例如，买家可能会指示代理人“当任何平台上出现最新的书x时，以最佳价格购买。”在这里，代理人决定何时以及如何进行支付。 在讨论查找书籍的途径和比较不同商家的价格后，提出最佳选择。尽管有了这些进步，人类通常仍需参与审查和最终确定购买，因为概率性决策和确定性执行的混合可能会在出现问题时复杂化责任归属。

为了保持清晰的问责制，一种方法是在人工智能驱动决策和实际支付执行之间建立明确的界限。这意味着将决策中的不可预测方面与执行支付的可预测步骤分开，通过将它们组织成不同的层次。三层模型旨在作为一个规范性的概念框架，由新兴的行业实践所启发。它并非旨在描述所有现有的支付架构，而是强调一个设计原则：在上游集中概率性和自适应推理，同时在需要法律最终性和系统稳定性的地方保持确定性的授权和结算。我们还利用该模型来强调新技术和标准是如何演变以支持每一层的。

层1——意图与编排层

in这一层包括将高级用户目标或指令转化为概率性代理系统和协议。这一层的科技使推理、规划、 意图搜寻、谈判和多代理——表2显示，新型标准迅速涌现并被采用，扩展了基于语言模型（LLM）的代理的能力。例如，MCP标准将代理接入外部数据和工具实现标准化，而A2A协议在不同供应商开发的代理间实现互操作和协调。x402标准基于HTTP 402网络协议，允许代理直接在其内部嵌入支付要求。 s HTTP请求数据。这使得能够自动协商并在互联网上处理付费服务。一

[5] 在这个框架下，“确定性”指的是授权和控制规则的执行、监管阈值、制裁名单以及智能系统的实施，这些均通过固定、可重复的决策程序进行。因此——但不是关于这些规则的界定方式。合同约束反映了人类判断和政策选择，但一旦确立，人类在规则设计和治理层面仍具有关键作用，而确定性执行则支撑了可审计性、责任性和稳定性。

[6] 尽管可以配置AI模型使其行为更具可预测性，但这并不意味着它们的决策规则符合授权或结算所需的规范性。因此，问题不仅在于可变性，还包括缺乏保证的可重现性——以及个别交易层面的法律可解释性。

最具影响力的标准是谷歌的UCP（谷歌2026），它为发现、比较以及代理商创建和管理购买后逻辑提供了共享的语法。

行业先驱已利用Mastercard的代理支付测试代理 - e新标准。例如，Visa的智能商业和启动购物与支付流程，其中代理在预定义的限制下自主构建购买意图。因此，第1层作为自适应编排和委托启动的焦点，产生必须仍通过确定性授权和结算层的结构化意图，这些将在下一部分进行描述。

第二层——控制和授权层

这一层强制实施确定性约束，以规范代理提出的或发起的动作是否可以执行。该层的科技确保即使受到上游概率系统的信息影响，授权决策最终仍由确定性政策规则支配。在这一层，研究人员正在研究通过可验证的协议将信任从人工监督转移到技术保障的协议。 领域关注声明、授权约束和身份框架（胡和荣 2025）。

该层的核心机制是AP2，它将代理发起的动作与加密可验证的指令绑定，这些指令指定了范围、限制、行为者身份和允许的条件。AP2标准还支持扩展x402，这使稳定币集成成为可能。这些指令确保下游授权反映了明确用户同意，而不是模型生成的推断。一些更成熟的网络授权技术，如OAuth 2.0或OpenID Connect，通常用于通过证明的代理身份实现强大的“了解你的代理”验证（世界经济论坛2026）。

最近行业的实施案例说明了基于授权令牌的授权如何在代理发起的支付流程中得以实施。例如，一些支付服务提供商已经引入了令牌化授权机制，允许人工智能代理使用用户的预先批准的支付方式（包括卡片和非卡片选项）发起交易，而不需要访问底层凭证（Stripe 2026）。这些方法展示了如何在保持结构化授权和支付方式选择的同时，保持对执行过程的确定性控制。

努力将代理支付与基于分布式账本技术（DLT）的授权和控制相结合，也在两条互补的轨道上取得进展。可验证的声明机制，如ERC-1812，允许链下安全签名的证明，这些证明可以通过智能合约或政策引擎进行验证，以支持在授权支付前对身份属性、权利或限制进行确定性检查。同时，可编程钱包和账户抽象标准，如ERC-6900，直接在钱包层实施支出限制、速度控制、交易对手限制和审批流程，在意图形成和执行之间创建一个确定性控制安全门。新兴的提案，如ERC-8004，通过定义代理身份、验证和声誉的链上注册，将这一授权框架扩展到代理本身，支持关于哪些代理可以采取行动的决定，并加强在确定性支付基础设施中机器发起行动的信任。

传统网络层级控制，如发行者规则、反洗钱（AML）/打击恐怖分子资金筹集过滤器、流速检查、令牌化规则、制裁筛查和纠纷限制，并行操作，提供确定性的验证。监督情报系统可能对这些控制进行信息提醒，例如反洗钱检测模型（FinRegLab 2025），但他们自身并不进行授权。

然而，大多数支付体系要求支付指令必须可追溯至账户持有人或其合法授权代理人发出的授权指令。由代理人发起的支付挑战了这种模式，因为单个交易可能不符合具体交易级别的清晰指令。授权因此变成结构性和规定性的，这引发了关于可追溯性、同意和现有责任的问题。

法律框架。这表明，随着代理支付模式的演变，需要更广泛且在法律上可行的授权概念（基于可验证的指令、范围限制和可审计性）。

综合考虑，这些技术将第2层定义为严格基于规则的授权边界：它仅在满足可验证要求、政策约束和监管检查的情况下接受来自第1层的结构化意图。在第2层接受的内容将成为第3层中可进行确定性执行的授权支付指令。任何失败的内容将被拒绝或路由回代理进行修改。这种分离确保了在代理系统引入上游更多自动化时仍能保持问责制、可审计性和合规性。

第三层——结算层

这一层包括传统的确定性结算基础设施，如实时全额支付系统（RTGS）、即时支付网络、卡网络清算引擎以及中央银行数字货币平台和基于分布式账本的结算轨道等较新的结算系统。第3层通过结算基础设施执行具有不可撤销法律终局的支付指令。当第3层对应金融市场基础设施时，其运营需遵守国际清算银行（BIS）CPMI-IOSCO金融市场基础设施原则，这些原则规定了治理、风险管理、接入和结算终局的标准。与第1层和第2层的自适应逻辑相比，本层的技术明确设计用于可预测、规则约束的执行、运营弹性和严格的可审计性，与金融市场基础设施使用的原则一致。

在此层上的执行还依赖于本地区域技术，例如可编程钱包和代币标准。在分布式账本技术（DLT）环境中，例如那些稳定币或 中央银行数字货币 ERC 4337，于2023年发布，为授权交易提供智能合约钱包执行。这些限制仍然存在。 基于规则的启动：它们执行第2层定义的政策，但不解释用户目标或不进行适应性推理。

在这个架构中，第3层是支付链的最终、非概率终点：它只接受那些在第2层通过确定性控制的指令，并直接执行，不进行修改、优化或重新解释。通常，代理算法在这里不发挥作用。[7] 这保证了法律确定性，限制了系统性风险，并确保支付系统的基石即使在上游流程日益自动化的情况下，仍保持稳定、同步和可信。

功能分离的含义

功能分离在早期引入，对代理人工智能如何与支付系统交互具有几个分析意义。首先，它解释了为什么近期从代理人工智能中获得的价值集中在结算前。其次，它阐明了为什么结算 在概率推理支持优化和协同系统不太可能以与代理相同的含义变得“智能”。尽管它们可能支持条件执行（例如，通过分布式账本上的智能合约），但在那一层级的核心功能依然保持确定性终局。第三，它突出了风险从个体交易正确性向系统级行为的转变。

[7] 尽管有限的代理功能，如自动化验证或对账，原则上可以在结算层运行，但这样做会引入交易，因为代理执行是不可能的，但限制它有助于在系统级关闭时，在效率与法律确定性之间保持问责制和最终性。因此，该框架将第3层设计为非概率性的，而不是概率性的。

表2. 付款流程与三层模型的对应关系

层次：L1 = 意图与编排，L2 = 控制与授权，L3 = 清算

支付通道功能	问题解决	技术，标准， 协议	层级（复数形式）
意图与语境 管理	启用代理进行推理 用户目标、限制 偏好，以及陈述跨 多步骤任务	LLMs；用于上下文工具的MCP 访问；ACP连接买方， 他们的代理人，和生意	L1
代理协调	允许多个代理（买方、 商人，国库，合规 风险）以交换计划， 协商行动，委派 子任务	A2A协议；多智能体 编排框架	L1
商务流程 编排	标准化发现 → 对比 → 提供 → 准备 为授权	通用商业协议	L1 + L2
委托与授权	证明一个代理商是 授权代表 用户/机构范围 受限权限 撤销，可审计性	AP2（数字签名指令） 并且意向证明）；OAuth 2.0； OpenID Connect	L2.
**代理身份（了解你的 代理人（dàili rényào）	识别和验证 软件代理作为独立的 运营参与者	ERC 780 索引登记册 ERC-8004 适用于身份和 声誉登记册；AP2 可验证的代理凭证	L2.
实时合规与 欺诈风险过滤	检测异常，实施制裁 风险，或违反政策之前 授权	ML/AI fraud systems; 约定、规定、指令 通过AP2的约束 适应性反洗钱模型	L2.
可编程结算 控制器	编码代理权限 消费限额、护栏和 直接转换为数字 钱或钱包	ERC 4337 智能合约 钱包；ERC 6900模块化 智能账户；ERC 1812 off 链式授权	L2.
平台特定代理 执行	允许代理人启动 支付在封闭或半封闭 封闭生态系统	AP2协议；x402链上 支付请求	L2 + L3
可编程货币和 数字资产轨道	支持以代理驱动支付 具有可编程条件 和解担保	DLT；智能合约层 规范代币框架	L3
传统支付渠道	接入传统，规范化的 支付系统	平台，用于实时全额转账系统和信用 卡片，等等	L3
流动性管理	优化流动性 管理	平台用于实时全额结算、稳定币	L3

来源：作者。注：A2A = 代理到代理；ACP = 代理商业协议；AI = 人工智能；AML = 反洗钱；AP2 = 代理支付协议；DLT = 分布式账本技术；LLMs = 大型语言模型；MCP = 模型上下文协议；ML = 机器学习；RTGS = 实时全额结算。

代理式人工智能在支付领域的功能

智能代理AI在支付领域可以带来广泛的应用场景的好处。专家将智能代理AI描述为AI代理的“数字工厂”，可以处理整个任务，只有在例外情况和监督时才需要人类（麦肯锡公司2025b）。为支付相关场景自主管理和执行复杂任务可以降低交易和运营成本，改善流动性，提升合规流程，并减少欺诈。移除人类延迟和管理摩擦预计将提高效率，加速资本流通，并提升生产力。

电子商务的演变

人工智能代理通过用自主推理取代简单自动化正在改变电子商务。随着人类参与的减少，交易可以加速，消费趋势可能会改变。代理型人工智能可以增强以比较金融 通过降低搜索成本和信息不对称，实现金融普惠，使用户能更有效地获得产品、费用和条款（曹，2026）。作为代理人，在优先考虑低延迟支付解决方案并消除低效因素的同时，他们可以直接提高货币流通速度。

数家公司已推出代理商业计划，以测试运营和消费者在潜在效率方面的收益（见框1）。在电子商务领域，代理可以通过自主管理购买过程，从产品搜索、价格比较到折扣应用、可用性验证和执行，来降低交易摩擦。对于商家来说，这种自动化可以加速购买周期，有证据表明基于大型语言模型的代理可以显著缩短消费者在数字市场中的互动时间（Yan等人，2025年）。消费者可以从结合了偏好、约束和实时价格信号的个性化决策支持中受益。代理可以利用用户提示中提供的更丰富上下文来预测需求模式并向商家传递结构化反馈（Beard，2025年）。除了降低用户的认知负担外，这种个性化还可以通过更相关的产品匹配和流畅的购买体验来提高客户终身价值。

代理系统也可能扩展到结账之后的购买功能，包括配送协调和退货管理。Gartner（2025年）预测，到2029年，人工智能代理可以自主解决高达80%的常见客户服务问题。对于商家来说，自动化的退货处理和自适应定价机制可以降低服务成本，改善争议解决，并能够更快地应对需求和库存条件的变化（麦肯锡2025b）。反映这些预期收益，调查表明企业正在优先考虑在客户服务、营销和产品开发方面的AI投资（微软2025年）。

更广泛地说，代理人工智能商业服务可以推动电子商务平台的现代化，因为它们可以实现人工智能系统之间的整合。例如，PayPal最近宣布了战略合作伙伴关系，允许商家无缝地在人工智能平台上启用产品发现（PayPal 2025）。此外，还应考虑重要的劳动力问题，因为代理可能取代或补充运营角色（包括和）。 商品营销、供应链管理以及产品优化），从而降低运营成本，提高生产力（Brynjolfsson、Li和Raymond 2023）。

箱1. 代理支付流程的示范性市场实验

最近科技公司和支付网络进行的实验为如何将智能代理AI能力集成到支付启动和编排层提供了初步证据。这些举措为新兴的设计模式提供了洞察，而不是既定的或标准化的架构。

OpenAI/Stripe（代理商业协议）

OpenAI的ChatGPT中的“即时结账”功能，由代理商业协议驱动，使用户能够在对话中直接购买产品。尽管它采用了标准的商户收单基础设施，但它在2026年初引入了一种新的经济模式，为自主代理引导的转换收取4%的交易费（Stripe 2025）。

亚马逊（Rufus和“帮我购买”）

亚马逊已从简单的推荐转变为委托购买。其“为我购买”功能使Rufus助手能够在代表顾客的网站上导航并完成交易，将购物代理定位为主要交互界面，而非辅助工具。

谷歌（通用商业协议）

2026年1月推出，全球商业协议标准化了企业在购物过程中与人工智能代理连接的方式。它使谷歌搜索人工智能模式和Gemini中的“原生结账”成为可能，使用户能够在不离开人工智能界面的情况下从Etsy和Wayfair等零售商处购买商品。

签证和万事达卡（网络级控制）

Visa的“智能商务”和Mastercard的“代理套件”（于2026年第二季度推出）聚焦于“认识你的代理”框架。这些倡议提供注册、加密签名和网络令牌，以区分合法代理和恶意机器人，确保确定性授权保持在既定的轨道内。

PayPal（Cymbio基础设施）

通过其在2026年对Cymbio的收购，PayPal将自己定位为代理商网络的“信任层”。这允许独立代理商利用PayPal的交易图和安全的保险库来促成结算，同时保持零售商的记录商状态。

自动跨境支付流动和流动性管理

授权AI系统超越智能自动化。它们能够自主协调和执行跨金融分布式网络的多步骤工作流程。授权AI可以协调整个跨境支付链，从支付发起、优化路由选项（包括代理银行、本地合作伙伴和代币化轨道）、触发合规检查到监控结算和结算后的异常，如同观察到的电子自动支付流程，可以减少与跨境支付相关的延迟。 商业应用案例。支持者认为这种人工干预和严格的流程（Convera 2025）。这个论点很有说服力：对实时交易数据的分析以及基于过往经验和不断变化的环境（包括变化的成本、费用和渠道性能）进行推理的能力，为AI代理提供了动态选择最有效支付路径和适应特定情况的能力（Capco 2025）。代理式AI还可以在

在执行交易时自主管理流动性，例如根据预定义参数和实时市场条件重新分配资金。国际清算银行的研究发现，生成式人工智能系统可以在不进行专门培训的情况下履行现金管理功能，如维护预防性流动性缓冲、优先处理紧急支付以及平衡流动性成本和结算延迟之间的权衡（国际清算银行，2025年）。

类似地，人工智能代理可以作为跨境支付的核心功能之一，帮助简化外汇（FX）管理。借助代理式人工智能，金融机构和企业可以持续监控实时汇率，分析银行渠道的价差，优化兑换时机，并为跨多种货币转账选择成本效益高的路径（Uppuduri 2025）。大型语言模型可以通过从非结构化数据源中提取信号，从而做到传统模型可能无法做到的事情。 提高诸如流动性规划等领域内的预测分析能力。 未充分利用，支持改进的预测和风险管理功能以及外汇管理。这些功能可以产生成本节约和运营效率，正如花旗和蚂蚁国际在旨在降低外汇对冲成本的AI工具中试点所示（路透社针对表外交易，如外汇交易）。 2025年）。此外，还有一些潜在的代理人工智能应用衍生品，但这篇笔记的范围不包括这些内容。

不同支付系统和网络的整合对于跨境支付尤为重要，鉴于管辖特定支付系统景观。尽管像SWIFT全球支付创新和Visa B2B Connect这样的网络倡议并非代理人工智能，但它们为代理在各个司法管辖区利用各种渠道奠定了基础（Scalefocus 2025）。随着稳定币在跨境支付中变得越来越普遍，使用稳定币渠道的人工智能代理可能会成为利用DLT功能和代理人工智能协同效应的重要基础设施，如标准协议AP2促进稳定币整合（Desai 2025）所示。人工智能与稳定币的融合有可能推动一个更具包容性和高效的“互联网金融体系”，加速全球金融的演变（WEF 2025）。

代理商合规方案

代理人工智能有望通过将监管逻辑直接嵌入到运营流程中，显著改善合规性流程。与传统自动化工具不同，代理系统可以解释目标，实时监控活动，并在预设的约束范围内自主采取行动，如标记案例或调整控制措施，当达到监管阈值时。通过 可疑交易、不断升级的高风险整合持续监控、可解释性和审计跟踪纳入其核心架构，使代理式AI能够以与现代数字系统相同的速度和规模运行，降低运营 合规功能在加强一致性、可追溯性和监管一致性的同时，增加了负担和人为错误。

代理人允许实时合规监控，持续评估交易和活动，而不是依赖周期性或批量审查。数家公司已引入基于人工智能的合规监控和欺诈检测工具。万事达卡的人工智能“决策智能”系统以交易速度对交易进行筛选，评估上下文... 毫秒级欺诈风险评分，允许在授权前进行实时合规风险评估（万事达卡2024年）。维萨公司开发了一套实时风险评分解决方案，用于账户间支付，该解决方案基于实时上下文数据在毫秒内对交易进行评分，以自动批准、拒绝或标记交易（维萨2026年）。它将合规逻辑，如欺诈、制裁标记和风险规则嵌入到实时交易流程中，从而在网络规模上引入了合规设计。

人工智能系统引入了自动执行规则的能力，应用监管要求与 决策点的代理治理框架。一个学术原型展示了用于处理反洗钱（AML）和了解客户（KYC）合规任务的代理人工智能的参考架构，该架构集成了可解释性和可追溯性。它展示了代理合规如何将监管逻辑直接嵌入、解释和执行到自主系统中，无需人工干预（Axelsen，Licht，和Damsgaard 2025）。通过嵌入身份保障措施，如来源验证和政策基于的访问控制

直接编入代码，机构可将法规从束缚转变为实际系统级推动者（Preis 2026）。

人工智能设计治理将道德、法律和社会价值观嵌入到代理系统中，从而实现主动风险管理（乔希2025）。这些系统能够帮助自动化合规任务，减少错误，降低成本，并通过自主代理处理日常工作，将复杂案件升级给人类来扩大运营规模（麦肯锡公司2025a）。用于欺诈检测的人工智能可以成为数字金融基础设施的一个元素，正如印度统一支付接口的情况。统一支付接口将支付、数据和智能集成为一个统一的堆栈，嵌入人工智能以实现实时欺诈检测、替代信用评分和自动对账。这种架构通过数字公共基础设施出口支持国内和国际扩张（希瓦姆2026）。

多智能体系统可以在专业领域间分配合规任务，而联邦协调确保跨司法管辖区的连贯性（Onyekaonwu，Igba，和PeterAnyebe 2024）。智能体通过安全协议进行通信，在交换监管情报的同时维护数据主权。将此应用于合规，一个专业智能体可以执行监管变更扫描，另一个进行风险评分，第三个控制映射，第四个执行补救措施，所有这些都由协调器协调，负责任务委派、冲突解决、回退和优先级排序。根据提出的三层框架，一个智能体可以在意图/编排层启动合规检查，另一个智能体可以在控制层扫描并标记监管要求的任何变更，而第三个智能体可以在结算层对合规检查进行验证。

风险、差距和缓解策略

代理AI系统能够自主解释目标，执行多步操作，并动态适应变化的环境。它们代表着金融服务业创新的下一个浪潮。在消费金融、支付、合规和监管领域，它们可以通过自动化复杂的金融和监管任务，显著减轻认知和操作负担。然而，它们的自主性、不透明性和非确定性行为为消费者保护、市场稳定和监管监督带来了实质性的风险。尽管代理AI对经济活动有着广泛的影响，但其重点在于这些技术如何影响支付系统的运行，包括流动性管理、授权、结算和支付基础设施的运行稳定性。表3根据风险的来源和潜在影响对主要风险进行了分类。

代理AI在电子商务中的便利性给消费者自主权带来了风险。代理系统可能误解用户意图，优化供应商激励而非用户福利，随着时间的推移偏离原始目标，或者在规模上参与微妙的行为引导。代理与最终用户之间的沟通挑战也可能出现，因为消费者并不总是对自己的短期和长期财务目标或风险偏好有清晰的感知（Aldasoro等人2024年）。在多代理环境中，不同元素或层次之间可能存在代理系统的目标不一致，或个体代理目标之间的偏离（Carichon 2025年）。如果主导模型识别出相同的市场信号，则出现算法羊群效应，而同时执行可能触发闪崩，绕过传统熔断器（Ogbuonyalu等人2024年）。算法羊群效应通过同步流动性需求、放大周期性行为，并在支付渠道中造成拥堵，从而损害核心金融市场基础设施的可预测性和弹性（国际清算银行2024年）。如果许多经济代理能够自主触发和处理金融操作，而恶意行为者可以通过即时注入攻击影响或控制代理，则将产生重大风险（霍根·洛尔斯2025年）。

持续获取财务数据、能够进行行为引导以及自主行动的能力，可能增加与消费者意愿相反的活动规模、微妙程度和个性化。

兴趣（FinRegLab 2025）。这些动态已经在一些人工智能驱动的个人金融应用场景中显现，这些应用因在用户的账户余额低于特定阈值时推荐现金贷服务和推广借款产品而受到批评，即使在那些产品会恶化消费者长期财务状况的情况下（Aldasoro等人2024）。

生成式AI模型容易出现“幻觉”，以高置信度创建虚假信息。在金融环境中，此类错误可能产生重大后果（Aldasoro等人，2024）。许多AI系统的黑盒特性进一步加剧了责任和赔偿的复杂性。AI系统的透明度和非线性使得监管者难以全面理解模型输出，限制了他们及时检测潜在系统性风险的能力。代理行为的高度相关性也可能带来显著的系统性风险（Aldasoro等人，2024）。AI可能使更复杂的风险评估模型成为可能，并改善对机构失败、市场操纵或识别合规相关问题的预测。尽管生成式AI可以通过与代理的交互式沟通在监管报告和合规方面发挥强大作用，但此类方法的常规使用尚未确立。

另一个重大风险与数据安全和隐私保护相关。依赖第三方服务如云服务提供商、AI模型端点和金融服务等自主代理需要用户的敏感数据，例如银行凭证、信用卡号码和加密钱包密钥，这使用户面临数据泄露和隐私问题，并形成了一个高度集中的脆弱点（Aldasoro等人，2024年）。

也有关于市场集中度和竞争的担忧。生成式AI需要大量数据作为输入，这些数据需要计算能力，而这种能力只能由少数几家主导公司提供，因此大多数AI供应链都表现出高度集中。 介绍集中度风险（科奈克集中度，从数据中心到云计算和AI应用。这种集中度可能威胁到创新，并提高财务稳定性、运营风险和声誉风险[国际清算银行，2025年]此外，随着算法日趋标准化并得到广泛采用，羊群效应和周期性风险也在增加[英国央行，2022年]。模型羊群行为与无意使用类似优化算法相关，这可能增加市场波动，在压力期间造成市场流动性问题[阿拉萨洛，...]。 导致闪崩事件，包括Inc和其他（2024）。类似于在金融市场观察到的算法羊群效应，支付领域中的相关行为在支付流中，从而增加了日内流动性。 启动或流动性优化可能导致供需同步，可能对结算能力和系统效率造成压力（Kabadjova等人，2023）。持续 最后，由于银行业竞争加剧，存在挑战，因为代理商正在将资金扫向高收益账户，消除了“惯性红利”，迫使银行立即在价值上竞争（麦肯锡公司 2025c）。

条款清晰，金融机构面临重大挑战，在 尽管AI在跨境支付集成中的好处，但将这些技术融入现有系统可能需要大量投资于系统升级和数据迁移，因为许多银行和金融机构的遗留基础设施可能无法与尖端AI技术兼容（Scalefocus 2025）。如果此类代理用于大规模自动化高速FX执行或流动性时机，现有高频交易研究证据表明，这会放大FX市场。 更多推测性高频活动可能加剧流动性问题并提高日内波动幅度（国际货币基金组织2025年）。跨境支付可以受益于基于DLT的稳定币和代理式AI的协同作用，可能会放大。将资金转换成加密货币以 然而，数字化资产加速结算可能会引入相对于传统支付渠道的额外波动性和保管风险，正如特别示例显示其价格波动性更高并依赖外部的专业保管基础设施。 加密资产不受监管的银行体系（金融稳定委员会2022年；国际货币基金组织2025年）。加密资产缺乏中心化的可信中介，依赖去中心化的共识机制，这改变了资产转移的风险特征，使参与者面临价格波动和交易对手方风险。即使是稳定币，在历史上也经历了锚定不稳定和在压力下的储备相关交易对手方风险，突显出快速转账速度与稳定性和保管之间的权衡（金融稳定委员会2023年）。

除了风险，还有阻碍代理人工智能在支付领域应用的差距。金融领域正出现日益增长的技能差距，缺乏既具备金融专业知识又拥有高级人工智能知识的专业人士（Uppuduri 2025）。许多大型金融机构最近才开始建立内部专业知识，以便安全地将代理人工智能整合到面向消费者的产品中（FinRegLab 2025）。

更具体的差距与代理的认证或KYC要求有关。能够发起交易的自主人工智能代理挑战了现有的支付身份和认证框架。传统的授权机制，包括KYC流程和多因素认证，都是围绕明确批准交易的人类用户设计的。当支付由代表授权行动的软件代理自主发起时，验证代理的身份和底层用户的意图变得更加复杂。这种转变在为人类行为者设计的支付系统中提出了关于认证、问责制和合规性的新问题（Acharya 2025）。关于自主代理的法律责任也存在模糊性。[8] 当前责任制度假设人类意图和直接因果关系，但当自主代理独立做出决策时，这些概念变得模糊不清。当代理人在交易中代表主体行事时，在发生争议时确定责任取决于代理人是否在其权限范围内行事，以及损害是否由代理人的自身行为还是由主体的指示造成。因此，主体和代理人可能根据具体情况面临责任，尽管法律责任划分并不明确（Shukanayev 2025）。在支付中，对代理人的问责制和责任缺乏普遍的监管清晰度。现有法规难以区分“未授权使用”和“用户疏忽”，如果代理人在幻觉中误导资金（Aldasoro等人2024）。

表3. 风险分类矩阵—支付领域中的代理人工智能

风险	(1) 风险的主要来源 （谁的行为）	(2) 如果是它，谁承担费用？ 实体化	(3) 市场失灵的政策正当性 干预？
指导差距 （结构型与……） 交易性的 授权	账户持有人授权广泛的账户持有人（意外的是，是的。信息和控制） 授权；AI代理执行 支付无交易层级支付系统（运营法律授权模型和 指示，说明	支付（款）；PSPs（争议）；不对称；不匹配 拉伤	代理执行
代理商的透明度 决策	人工智能开发者与部署者 设计不可解释的模型； 平台集成代理 支付流程	用户（缺乏救济）；PSPs (合规失败)；主管 （监管能力下降）	是，信息不对称和 不可验证的决策过程
高速，机器 时间支付执行	人工智能代理和平台优化支付系统 执行速度；不足 PSPs的节流	操作过载；最终用户 误差传播	是的。速度的外部效应 放大和协调效应
授权 追溯失败	PSPs及依赖的平台 授权的委托授权 稳健可审计性	PSPs（负债风险）；用户 （有争议的款项） 法院/监管机构（法律） 不确定性（dú xíng duàn）	是的。法律基础设施不匹配； 不完整的合同
责任不明 分配（机构） 依据（的）	账户持有人、PSPs、系统PSPs以及用户（诉讼、 操作员在过时损失下互动；支付系统 责任假设	声誉风险	是的。法律不确定性以及 责任分配不完整
产品责任 来自……的暴露 自主行为	人工智能模型提供商、集成商、和 平台部署自适应系统	PSPs、平台或用户 视法律解释而定；非为适应性的、后续 可能的模式提供者	是的。现有的产品责任制度 部署行为
相关代理 行为（放牧） 支付时间或 路由	同质模型和共享压力）；PSPs（结算）	支付系统（流动性） 延迟；最终用户（代理间未实现优化目 付款	是的。外部协调效应 系统性风险

[8] 除了基于代理的责任外，代理型人工智能系统还可能在产品责任制度下引发问题，尤其是在自主行为导致有害后果之后。现有的产品责任框架尚未充分适应行为可能随时间演变的产品，这给责任和赔偿增加了不确定性。

风险	(1) 风险的主要来源（谁的行为）	(2) 如果是它，谁承担费用？实体化	(3) 市场失灵的政策正当性干预？
日内流动性压力和解拥堵	代理人优化付款时间并且流动性使用；PSPs未能施加限制	支付系统；中央银行流动性保障；参与者	是的。系统性流动性外溢
网络安保与攻击面扩张	平台整合代理与多种工具/API；权限较弱控制器	用户（欺诈）；PSPs（损失）；支付系统（运营中断	是的。安全外部性以及投资不足于韧性
DLT结算无需法律终局	系统设计师部署分布式账本技术无法定仲裁认出；识别	用户和中介（逆转，破产风险）；法律系统	是的。法律基础设施差距和管辖碎片化
监管盲区KYA，监管，监控	监管机构落后于技术转变；跨平台运行边界	大型金融体系（损失）是的。公共福祉性质信任；监管机构（监管）失败（fà shī）	监督与跨境协调发展问题

来源：作者。

注意：AI = 人工智能；DLT = 分布式账本技术；KYA = 了解你的代理商；PSPs = 支付服务提供商。

缓解策略

代理式AI系统是非确定性的，这意味着由于它们的概率性质，相同的输入可能导致不同的输出（NIST 2023）。这种特性使得当此类系统发起交易或执行支付指令时，会引入操作和财务风险。一种关键的缓解方法是实施正式的AI治理结构，该结构定义了AI系统的问责制、风险控制和生命周期监控。框架应强调在整个AI生命周期中的透明度、问责制和持续监控。治理框架可以包括金融机构内的AI风险委员会、董事会层面的AI部署监督以及适用于代理式系统的内部模型风险管理（NIST 2023）。

系统措施

由于自主代理可以高速执行交易，因此对于由AI代理执行的高风险或高价值交易，应要求人类审批或监管干预。关于代理治理框架的研究强调，将人类纳入监督流程作为核心保障，以监控和干预自主代理的决定，以降低错误或恶意行为的风险（Chiris和t交易阈值需要人类）。 Mishra 2025）。[9] 男性赞同 这种监管盲点表明，应该为代理人决策提供监督仪表盘和手动 Override 机制。一旦交易达到法律终局状态，错误或意外行为可能难以或无法逆转。这有力地支持了提前介入和遏制机制的必要性，通常称之为“kill switches”，作为风险管理框架的一部分。尽管其重要性不可否认，但闭环中的人可能会引入意想不到的风险。例如，随着支付流的整体优化，由于人为审批造成的支付延迟可能会无意中增加流动性风险，并使对冲策略效果减弱（英格兰银行，2026年）。

尽管中断机制在相邻领域，如算法交易中已建立良好，但这些安全措施通常依赖于监管中介、集中控制和清晰的制度问责制（国际证券委员会组织，2015年）。代理支付场景可能削弱这些假设，尤其是在AI系统跨越平台或代理授权下操作，降低现有方法的应用性。同时，中断机制并非没有风险。如果设计不当，它们可能导致单点故障，扩大系统的攻击面，或通过未授权或级联激活启用服务拒绝（Russu 2025）。因此，断路器应作为一层嵌入支付堆栈的治理能力，具有分布式的，而非集中的

[9] 人类在闭环机制中尤其能够缓解与自动化系统相关的某些风险，但这些机制本身并不足够，保障措施限制了依赖真实的压缩或集群失败场景。这强调了架构和自动时间人工干预的重要性。

控制。他们还应包括逐步响应而非突然关闭，明确的权限和可审计性，以及局部化影响并避免不必要的系统级中断的策略（秦等人，2024）。

鉴于生成模型的非确定性本质，具有代理功能的AI系统不应在没有结构保障的情况下直接执行不可逆的交易。如前几节所提出的，决策层和执行层的架构分离可以作为一个重要的缓解因素。决策层的代理提出或启动一个动作，而确定性的执行层执行合规性检查或独立控制，并最终执行交易（Acharya 2025）。

私营部门措施

支付网络应致力于成为受信赖的基础设施层。这包括推出带有嵌入式消费控制和实时预算管理的代理准备就绪的卡产品；建立具有身份验证、声誉评分和欺诈监控的全球代理注册；引入针对AI启动交易量身定制的争端解决框架，明确责任分配；以及开发适用于各种AI平台的开放协议标准，以实现互操作性（辛格2025年）。数字钱包提供商应部署“代理准备就绪”功能，使用通行密钥和生物识别验证建立对代理友好的身份验证，适用于对话界面，并保持与商家的关系，确保商家即使在交易通过代理发起时仍然是“记录商家”（辛格2025年）。

对于私营部门参与者来说，确保能够实时访问清洁、一致的数据，并全面映射地图功能，对于促进多种类型人工智能、技术保障和人工审核的整合至关重要（RegTechLab 2025）。金融机构和支付公司需要投资于培训和招聘，以组建能够开发、实施和维护安全、稳健和互操作的人工智能系统的团队（Uppuduri 2025）。最后，在引入代理人工智能到支付系统中时，拥有强大的网络安全态势至关重要。自主人工智能代理通常与多个外部系统、工具和API交互，扩大了攻击面，使组织面临数据泄露、工具滥用和跨系统权限提升等风险。因此，在将人工智能代理与运营基础设施集成时，需要强大的身份验证、范围授权和安全的API治理来减轻这些风险（Errico、Ngiam和Sojan 2025）。

公共部门措施

从公共部门的角度来看，代理人工智能要求监管者考虑从“了解你的客户”到“了解你的代理”要求的转变，其中为金融机器人设定的强制可验证身份与法律实体相关联（世界经济论坛2026年）。传统的欺诈模型依赖于人类行为模式，当交易由自主代理发起时，这些模式变得无效。因此，开发验证框架以验证人工智能代理的身份和用户的授权委托仍然是关键。表2中描述的一些技术旨在允许建立护栏，以实现审计性、日志记录和可解释性，为支付代理解决一些监管要求。

随着自主人工智能代理融入金融和公共系统，政策制定者越来越将它们视为需要强大治理和监管的关键数字基础设施的一部分。因此，确保安全部署需要允许系统自主运行的同时，在出现不安全行为时仍能进行监管干预的机制（Schmitz、Rystrøm和Batzner 2025年）。政策制定者应部署实时监控系统来检测代理行为和交易流中的异常，维护AI活动日志和审计跟踪，并为异常代理行为生成自动警报（FSB 2025年）。实现风险实时监控需要适当的数据收集、事件报告和披露政策，以及与AI相关的监管技能，以跟上技术进步的步伐。在监管环境，如监管沙盒中的监管者， 应允许在全面市场部署前对代理支付系统进行测试。

各国已启动人工智能治理框架，但需要在监管协调和统一方面做出更多努力，以预防潜在的溢出和传染效应，以及监管碎片化，特别是在跨境支付领域。一个引人注目的例子是新加坡，它推出了适用于代理型人工智能的治理框架模型。该框架为组织提供指导，包括在部署代理时需要采取的技术和非技术措施，如通过选择合适的代理应用案例和限制代理权力（如自主权和访问工具、数据）来进行详尽的事前风险评估，通过定义“代理重大行为”来强化人类对代理的责任，在整个过程中实施技术控制和流程。 检查点，其中有人工审批代理的生命周期，以及通过透明度和教育启用最终用户的责任（IMDA 2026）。影响代理型AI应用的相关AI法规，例如欧盟的《AI法案》，是政府部门应考虑的关键风险缓解策略（《欧盟人工智能法案》2024）。

结论

智能代理代表支付系统的一个可能具有变革性的发展。能够解释目标、协调交易并在数字服务中自主交互的软件代理可以显著扩大金融活动的自动化。早期的实验表明，智能代理系统可以提高电子商务交易的效率、优化跨境支付路由、增强流动性管理，并支持实时合规性监控。

然而，这些好处与基本设计紧张并存。支付系统建立在保证可预测性、责任和最终法律结算的确定性基础设施之上。[10] 相比之下，具有代理属性的AI系统依赖于概率推理和自适应决策。因此，整合这些技术需要仔细的架构设计，以确保自动化不会削弱支付基础设施的核心可靠性。

本笔记中提出的三层框架通过将概率决策与确定性授权和结算分离来调和这些差异。这种结构允许支付系统利用代理技术进行改进的协调和优化，同时在执行环节仍保持强大的安全防护。在此模型中，创新集中在上游，代理解释并保留基于规则的授权和法律上最终的规则。 意图协调行动，同时下游层沉降。

同时，几种风险依然突出。自主代理可能误解用户意图，通过高度相关的行动放大市场波动，或在网络安全中创造新的漏洞，由机器发起的交易身份问题也提出了挑战。 管理框架和责任框架。围绕人类用户构建的应急结构，引发了对身份验证的新问题。 现有的监管一致性、问责制和争端解决。

缓解这些风险需要私人部门和公共利益相关者的协调行动。金融保障和技术架构 机构必须投资于治理结构，将网络安全与支付执行分离，通过代理推理。支付网络和技术提供商需要开发可信赖的身份框架和互操作性标准，以验证了解你的代理并委派权限。同时，监管者可能需要调整监管方法，包括监控框架、测试环境和AI介导的金融活动的治理标准。

[10] 在调解层保留法律终局性，包括对如DLT等新兴技术，依赖于明确的法律认可和适当的法定保障。

总的来说，代理型AI在支付领域的轨迹不仅取决于技术能力，还取决于制度设计和治理选择。如果得到适当的保障措施实施，代理型AI可以提高效率，降低跨境支付的摩擦，并扩大金融可及性。然而，如果没有这样的保障措施，同样的技术可能会引入新的运营和系统风险。因此，政策制定者和行业参与者面临的挑战不是是否采用代理技术，而是如何将其融入支付系统，以保持信任、稳定和责任，在日益自动化的金融生态系统中。

参考文献

阿恰亚，维韦克。2025。《在不信任环境中验证真实性及意图的安全自主代理支付》。arXiv预印本 arXiv:2511.15712。 https://arxiv.org/pdf/2511.15712

Aldasoro, Iñaki，Leonardo Gambacorta，Anton Korinek，Vatsala Shreeti，Merlin Stein. 2024.《智能金融系统：人工智能如何改变金融》. 国际清算银行工作论文第1194号，国际清算银行。 https://www.bis.org/publ/work1194.htm

Axelsen, Henrik，Valdemar Licht，和Jan Damsgaard. 2025年。“金融犯罪合规的代理人工智能。”arXiv预印本 arXiv:2509.13137。 https://arxiv.org/pdf/2509.13137

国际清算银行（BIS），2024。“年度经济报告”。
https://www.bis.org/publ/arpdf/ar2024e.htm

国际清算银行（BIS）. 2025年. “政策目的下的人工智能应用。”提交给二十国集团财政部长和央行行长的报告。 https://www.bis.org/publ/othp100.pdf

英格兰银行（BoE）。2022年。“人工智能与机器学习。”讨论稿5/22。
https://www.bankofengland.co.uk/prudential-regulation/publication/2022/october/artificial-intelligence

英格兰银行（BoE）。2026年。“2026年2月人工智能圆桌会议总结。”
https://www.bankofengland.co.uk/minutes/2026/february/summary-of-ai-roundtables-feb-2026

胡子，妮可莱特·V. 2025. “电子商务人工智能代理：改变数字零售的自主技术。”BigCommerce博客 https://www.bigcommerce.com/blog/ecommerce-ai-agents

波士顿咨询集团（BCG），2025年。“代理人工智能、数字货币和实时交易重塑全球支付格局。”
https://www.bcg.com/press/22september2025-reshape-global-payments-landscape

Brynjolfsson, Erik，Li, Danielle，Raymond, Lindsey R. 2023. “工作场所的生成式人工智能。”美国国家经济研究局工作论文 No. 31161。 https://www.nber.org/system/files/working_papers/w31161/w31161.pdf

曹，明轩。2026。《通过人工智能驱动的金融科技弥合信息不对称：数字足迹分析在金融包容性中的作用》。国际商务与经济研究，第8卷，第1期，2026年。
https://www.scholink.org/ojs/index.php/ibes/article/view/57059

Capco. 2025. “代理AI：支付与现金管理的变革。”
https://www.capco.com/intelligence/capco-intelligence/agentic-ai-transforming-payments-and-cash-management

Carichon, Florian，Aditi Khandelwal，Marylou Fauchard，Golnoosh Farnadi. 2025.《多智能体错位即将到来的危机：人工智能对齐必须是一个动态的社会过程》。arXiv预印本 arXiv:2506.01080。 https://arxiv.org/pdf/2506.01080

Chiris, Lorenzo Satta, 和 Ayush Mishra. 2025. “AURA：一个智能体自主性风险评估框架。”arXiv预印本 arXiv:2510.15739。 https://arxiv.org/abs/2510.15739

Convera. 2025. “如何代理式AI正在改变支付领域。”
https://convera.com/blog/cross-border-payments/agentic-ai-in-payments

Desai, Akshar Prabhu. 2025. “超越订阅：为什么代理商业需要稳定币以实现规模扩张。”金融科技周刊。

https://www.fintechweekly.com/magazine/articles/agentic-commerce-stablecoins-micropayments-ai-payme nts

Errico，Herman，Ngiam Jiquan，Shanita Sojan. 2025.《保障模型上下文协议（MCP）的安全性：风险、控制与治理》. arXiv 预印本 arXiv:2511.20920. https://arxiv.org/abs/2511.20920

欧盟人工智能法案 2024。 https://artificialintelligenceact.eu/

金融稳定委员会（FSB）。2022。“加密资产对金融稳定风险的评估。” https://www.fsb.org/2022/02/assessment-of-risks-to-financial-stability-from-crypto-assets/

金融稳定委员会（FSB）。2023。“国际货币基金组织-金融稳定委员会综合论文：加密资产政策。” https://www.fsb.org/uploads/R070923-1.pdf

金融稳定委员会（FSB）。2025年。“监控金融领域人工智能应用及相关风险。” https://www.fsb.org/2025/10/monitoring-adoption-of-artificial-intelligence-and-related- vulnerabilities-in-the-financial-sector/

FinRegLab. 2025. “新浪潮来袭：金融服务业中的代理AI。”https://finreglab.org/wp-content/uploads/2025/09/FinRegLab_09-04-2025_The-Next-Wave-Arrives-Main.pdf

Gartner. 2025. “Gartner预测，到2029年，代理式人工智能将无需人工干预，自主解决80%的客户服务问题。” https://www.gartner.com/en/newsroom/press-releases/2025-03-05-gartner-predicts-agentic-ai-will-autonom ously-resolve-80-percent-of-common-customer-service-issues-without-human-intervention-by-2029

谷歌。2026。“通用商业协议。”谷歌开发者。 https://developers.google.com/merchant/ucp

新加坡资讯通信媒体发展局。2026年。“新加坡推出新型AI治理框架，针对智能体AI。” https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press- releases/2026/new-model-ai-governance-framework-for-agentic-ai

霍金路伟。2025。“金融服务中的代理人工智能：监管和法律考量。” http://hoganlovells.com/en/publications/agentic-ai-in-financial-services-regulatory-and-legal-considerations

胡，宝桃“琥珀”，和罗娜·贺尔琳娜。2025。“代理商信任模型：关于简短、索赔、证据、股权、声誉和限制在代理网络协议设计中的比较研究。”arXiv预印本arXiv:2511.03434。 https://arxiv.org/abs/2511.03434

国际货币基金组织（IMF），2025年。“全球外汇市场的风险与韧性。”全球金融稳定报告。 https://www.imf.org/-/media/files/publications/gfsr/2025/october/english/ch2.pdf

国际证券委员会组织（IOSCO）. 2015.《交易场所有效管理电子交易风险及业务连续性计划机制》。 https://www.iosco.org/library/pubdocs/pdf/ioscopd483.pdf

Joshi, Himanshu. 2025. “通过设计治理智能体系统：负责任发展和部署的框架。” https://storage.ghost.io/c/44/95/449506ca-034e-480f-9725- fcde08ef1cc1/content/files/2025/06/AI-Governance-by-Design-for-Agentic-Systems--A-Framework-for- Responsible-Development-and-Deployment.pdf

卡巴乔娃，比利阿娜·亚历山大罗娃，安东·巴迪夫，塞尔沃·贝基摩尔·巴托斯，埃夫安哥洛斯·本诺斯，费迪·塞佩达洛佩斯，詹姆斯·查普曼，马丁·迪尔，伊奥安娜·杜卡-拉德乌，罗德里·加雷特，罗纳德·黑伊曼斯，安耐克·科斯，安东尼·马丁，托马斯·内伦，托马斯·尼森，扬·波利克，安德烈·波斯京尼克夫，

弗朗西斯科·里瓦登耶拉，马里奥·鲁本·多·库托·巴托斯，萨拉·泰西。2023年。《全球日内流动性》。国际清算银行工作论文第1089号。
https://www.bis.org/publ/work1089.pdf

Korinek, Anton，Jai Vipra. 2025.《集中智慧：生成式AI的规模法则和市场结构》。 *经济政策* 40（121），第4页。

万事达卡。2024。“决策智能——超越万事达卡。”
https://www.mastercard.com/content/dam/mccom/shared/business/b2b/reports/decision-intelligence-beyon d-mastercard-playbook.pdf

麦肯锡公司。2025a。《代理人工智能如何改变银行打击金融犯罪的方式》。
https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/how-agentic-ai-can-change-the-way- banks-fight-financial-crime

麦肯锡公司。2025b。《代理商业机遇：人工智能代理如何引领消费者和商家进入新时代》。
https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-agentic-commerce-opportunity-how-a i-agents-are-ushering-in-a-new-era-for-consumers-and-merchants

麦肯锡公司. 2025c.《惰性终结：代理人工智能对零售和小微银行行业的颠覆》。
https://www.mckinsey.com/industries/financial-services/our-insights/the-end-of-inertia-agentic-ais-disruption -of-retail-and-sme-banking

微软. 2025. “2025年度工作趋势指数：前沿企业诞生。”
https://www.microsoft.com/en-us/worklab/work-trend-index/2025-the-year-the-frontier-firm-is-born

美国国家标准与技术研究院（NIST）. 2023年.《人工智能风险管理框架（AI RMF 1.0）》 https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

Ogbuonyalu, Uchenna Obiageli，Kehinde Abiodun，Selorm Dzamefe，Ezeh Nwakaego Vera，Adewale Oyinlola，和 Igba Emmanuel. 2024. “评估人工智能驱动算法交易对市场流动性风险和金融系统性脆弱性的影响。” *《国际科学研究和现代技术杂志》* 3 (4): 18–21.

Omar，Mahmud，Vera Sorin，Jeremy D. Collins，David Reich，Robert Freeman，Nicholas Gavin，Alexander Charney，Lisa Stump，Lisa Stump，Nicola Luigi Bragazzi，Girish Nadkarni，和Eyal Klang。2025。“多模型保证分析显示大型语言模型在临床决策支持期间高度易受对抗性幻觉攻击。” *通信医学* 5: 330.
https://doi.org/10.1038/s43856-025-01021-3

Onyekaonwu, Chinenye Blessing, Emmanuel Igba, 和 Amina Catherine Peter-Anyebe. 2024. “面向监管智能的代理型人工智能：跨国科技公司可扩展合规生命周期系统的设计。” *《国际科学研究和现代技术杂志》* 3 (12): 205–22.

PayPal. 2025. “PayPal推出基于人工智能的购物服务以推动智能商业。”
https://newsroom.paypal-corp.com/2025-10-28-PayPal-Launches-Agentic-Commerce-Services-to-Power- AI-Driven-Shopping

普莱斯，亚当。2026。“代理金融：在自主智能时代建立信任。”Finextra。
https://www.finextra.com/blogposting/30867/agentic-finance-building-trust-in-the-age-of-autonomous- intelligence ？

秦，星生，江帆，董成祖，罗宾·多斯。2024。“针对工业控制系统侦察攻击的混合网络安全框架。” *电脑与安全* 136: 103506.

RegTechLab. 2025。《新新浪潮来袭：金融服务中的代理人工智能》。市场扫描。https://finreglab.org/wp-content/uploads/2025/09/FinRegLab_09-04-2025_The-Next-Wave-Arrives-Main.pdf

路透社。2025年。“花旗与蚂蚁国际共同试点AI驱动的外汇工具，助力客户降低对冲成本。”https://www.reuters.com/business/finance/citi-ant-international-pilot-ai-powered-fx-tool-clients-help-cut-hed ging-costs-2025-07-18/

Russu, A. 2025. “联网设备中的“关闭开关”：领导者不能忽视的潜在风险。”https://www.resillion.com/blogs/kill-switches-in-connected-devices-the-emerging-risk-leaders-cannot-ignore/

Scalefocus. 2025. “人工智能在跨境支付中的作用。”https://www.scalefocus.com/blog/the-role-of-ai-in-cross-border-payments

施密特，克里斯，乔纳森·里斯特罗姆，以及简·巴策纳。2025。“公共部门组织中代理人工智能的监管结构。” *计算语言学协会* ．https://arxiv.org/abs/2506.04836 ？

Shivam. 2026. “2026年印度人工智能影响峰会传递了哪些关于支付基础设施未来的信号？”https://www.pinelabs.com/blog/india-ai-impact-summit-payments

Shukanayev, Dastan. 2025. “代理失败时谁承担责任？在碎片化的监管环境中自主支付系统的责任框架。”12月1日。 https://ssrn.com/abstract=5864482

Singh, Anurag. 2025. “支付革命：代理商业将如何改变B2C和B2B交易。”https://www.linkedin.com/pulse/payments-revolution-how-agentic-commerce-transform-b2c- anurag-singh-5l5tc/

条纹。2026年。“支持代理商业的额外支付方式。”https://stripe.com/blog/supporting-additional-payment-methods-for-agentic-commerce

Uppuduri, Vinod. 2025. “国际交易的未来：人工智能如何改变跨境支付。” *《国际计算机科学、工程与信息技术科学研究期刊》* 11 (1). https://ijsrcseit.com/index.php/home/article/view/CSEIT251112157/CSEIT251112157

签证。2026年。“Visa 保护 A2A 交易。”https://corporate.visa.com/en/products/visa-protect-a2a-payments.html

世界经济论坛（WEF）。2025年。“人工智能与稳定币：构建智能在线商业的新伙伴。”https://www.weforum.org/stories/2025/01/stablecoin-ai-business/

世界经济论坛。（WEF）2026。“到2034年，AI代理可能价值2360亿美元——如果我们处理好了信任问题。”https://www.weforum.org/stories/2026/01/ai-agents-trust/

Yan, Yineng, 王锡东，程金生，胡 Ran，关伟涛，Farahmand Nahid，林恒特，李悦. 2025. „FaMA：基于LLM的消费者对消费者市场平台的智能助手。”https://doi.org/10.48550/arXiv.2509.03890

PUBLICATIONS

如何赋权AI将重塑支付领域 NOTE/2026/004