

金融智能体系统可观测性技术规范

标准编制说明

起草组

2026 年 10 月 12 日

1、标准范围

本标准规定了金融智能体系统可观测性的总体要求、可观测对象及观测信号要求、数据采集与关联要求、运行展示与分析要求、审计追溯要求以及安全与合规要求。

本标准适用于银行、证券、基金、保险及其他金融机构单智能体、多智能体、基于大模型工作流的智能体系统以及与外部业务系统联动执行的金融智能体系统的规划、设计、开发、部署、运行、运维、审计和评估，也适用于为上述系统提供大模型服务、智能体平台、工具系统、运行支撑平台及相关技术服务的单位。

2、工作简况

本标准由国金证券股份有限公司、中国信息通信研究院作为牵头单位组织起草。

2026年9月以来，编制组围绕金融智能体系统建设、运行和治理需求，系统梳理金融行业智能体应用现状、运行风险、监管要求及相关标准基础，形成标准研制框架。主要工作过程如下：

（1）2026年9月，牵头单位组织开展前期资料收集和需求调研，研究金融智能体系统在客户服务、营销运营、投研辅助、风控合规、运营管理等场景中的运行特点及可观测性需求；

（2）2026年9月下旬，编制组结合金融机构实践情况，梳理日志、指标、链路追踪、事件记录、告警联动、审计留痕等关键能力要求，形成标准框架和重点内容方向；

（3）2026年10月上旬，编制组基于国内外人工智能治理、生成式人工智能安全、智能运维、云资源监控等相关标准和规范，形成《金融智能体系统可观测性技术规范》草案；

（4）2026年10月中旬，结合项目建议书内容及专家意见，对标准草案和配套材料进行完善，形成标准初稿及编制说明草案。

3、标准编制原则和确定标准主要内容的依据

本标准遵循规范性、适用性、先进性和可操作性相结合的原则。

一是规范性原则。标准编制严格遵循 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》及相关法律法规要求，确保标准结构完整、表述准确、条款边界清晰。

二是适用性原则。立足银行、证券、基金、保险等金融机构智能体系统建设与运行实际，围绕模型调用、工具调用、任务编排、多智能体协同、人工接管、审计留痕等关键场景提出要求，增强标准落地性。

三是先进性原则。充分参考人工智能风险管理、生成式人工智能安全治理、可观测性工程、智能运维等国内外最新研究与实践成果，确保标准内容与技术发展趋势相适应。

四是可操作性原则。充分考虑金融机构在系统建设、运行管理、审计检查和风险控制中的实施条件，突出观测对象、观测信号、链路关联、可视化展示和审计追溯等可执行要求。

本标准在内容设置上，主要依据国家关于“人工智能+”和“人工智能+金融”的政策要求，结合金融智能体系统运行链路长、参与主体多、异常机理复杂、责任追溯要求高等特点，围绕术语定义、总体框架、观测对象、观测信号、数据采集、上下文关联、指标体系、日志与链路规范、事件与审计记录、异常检测与告警、可视化展示、问题定位、运行评估以及安全与合规等方面进行编制。

4、主要试验（或验证）的分析、综述报告

当前，金融智能体系统正从单点试用走向多场景部署，逐步嵌入客户服务、营销运营、投研辅助、风控合规、运营管理等关键环节。与传统信息系统相比，金融智能体系统具有模型驱动、工具调用、多轮交互、跨系统执行和动态决策等特点，系统链路更长、行为主体更多、运行状态更复杂，对系统运行监测、异常

识别、责任追溯和治理控制提出了更高要求。

编制组在研制过程中，重点围绕金融智能体系统在日志、指标、链路追踪、事件记录、告警联动、审计留痕、异常定位和责任追溯等方面的技术需求进行了分析论证。研究认为，当前行业在智能体运行监测口径不统一、跨组件链路难关联、模型与工具调用过程不透明、异常问题定位困难、审计证据不完整、合规留痕不足等方面仍存在明显短板，亟需通过标准对关键观测对象、观测信号、关联机制和展示分析要求进行统一规定。

结合现有工程实践和标准基础，编制组认为，本标准具备明确的应用需求和实施基础。通过建立统一的可观测性技术要求，可为金融机构建设安全、可控、可审计、可持续优化的智能体系统提供依据，提升金融智能体规模化应用条件下的运行可靠性、风险可控性和监管适配性。

5、标准在起草过程中遇到的问题及解决办法：重大分歧意见的处理经过和依据：有无重要技术问题需要说明

在本标准起草过程中，暂未发现重大分歧意见。

标准编制过程中需要重点把握的技术问题，主要包括金融智能体系统边界界定、多智能体协同场景下的链路关联方式、模型调用与工具调用的统一观测口径、审计留痕与隐私保护之间的平衡关系等。对此，编制组拟在后续征求意见和技术论证过程中，进一步结合行业实践和专家意见完善相关条款，确保标准要求合理、明确、可实施。

6、与国外标准的关系：包括：采用国际标准和国外先进标准的程度，与国外标准主要技术内容的差异

目前未发现与本标准直接对应的国际标准或国外先进标准，本标准拟采用自主制定方式，不涉及等同采用或修改采用某一单项国际标准。

在标准制定过程中，将参考 ISO/IEC 42001 关于人工智能管理体系的要求，参考 NIST AI RMF 及其生成式人工智能画像文件在风险识别、监测、评估和处置方面的方法，参考 OpenTelemetry 在日志、指标、链路追踪和生成式人工智能语义约定方面的开放规范，并结合我国金融行业监管要求和应用特点进行转化吸收。

7、修订标准时，说明与标准前一版本的重大技术变化，并列岀所涉及的新、旧版本的有关章条：废止/代替现行有关标准的建议

本文件为制定标准，非修订标准。

8、说明标准与其他标准或文件的关系，特别是与有关的现行法律、法规和强制性国家标准的关系

本标准与国内现行标准之间是衔接、细化和场景化落地关系，不重复替代已有通用标准。

其中，GB/T 45288.1—2025、GB/T 45288.2—2025、GB/T 45288.3—2025 以及 GB/T 45907—2025 可为大模型与人工智能服务能力提供基础要求和评价依据；GB/T 46347—2025、GB/T 45654—2025 可为人工智能风险管理和生成式人工智能服务安全治理提供支撑；GB/T 28827.1—2022、GB/T 43208.1—2023、GB/T 43208.2—2025、GB/T 37736—2019、GB/T 37938—2019 可为运行维护、智能运维和资源监控提供底层依据。

本标准在标准体系中属于面向金融智能体的跨场景运行保障与治理标准，向上支撑各类金融智能体应用标准，向下衔接大模型、云计算、智能运维和网络安全等基础标准，填补金融智能体系统运行观测与审计追踪领域的标准空白。

本标准与现行法律法规及强制性国家标准保持协调一致，相关条款与网络安全、数据安全、个人信息保护、金融风险管理、人工智能安全治理等要求相衔接。

9、标准作为强制性标准或推荐性标准的建议

建议本文件作为推荐性团体标准。

10、贯彻标准的要求和措施建议：标准发布后，对国内外业界可能产生的影响

建议本文件作为推荐性标准发布实施，推动金融行业智能体系统可观测性建设形成统一认识和一致做法，提升行业运行监测能力、风险识别能力和审计治理能力，对规范金融智能体系统建设、保障智能化应用安全发展具有积极作用。

11、标准是否涉及知识产权的情况说明；如标准中含有自主知识产权，说明产品研发程度、产业化基础及进程

本文件未涉及。

12、其他应予说明的事项

本文件未涉及。