

面向金融业务场景的大模型应用能力评估 模型 (FLMM)研究

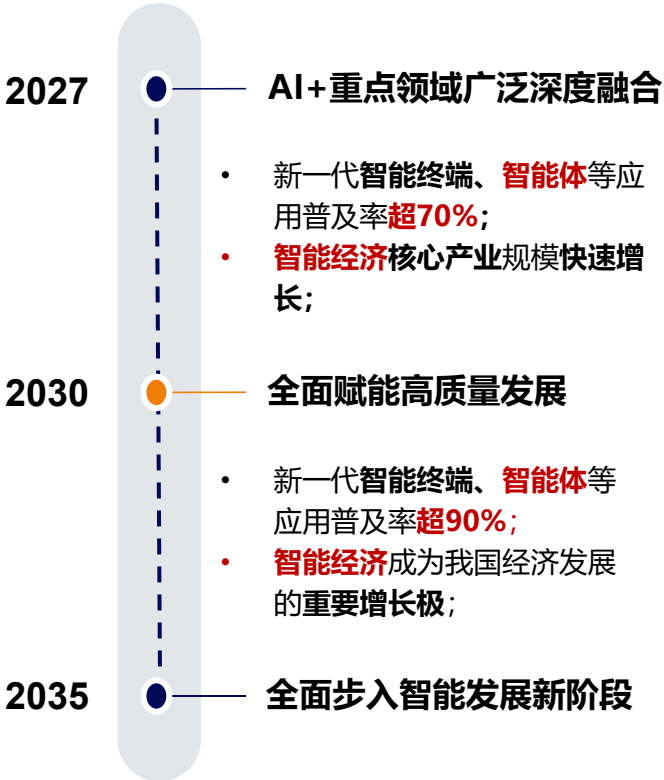
2026年8月7日

目录 contents

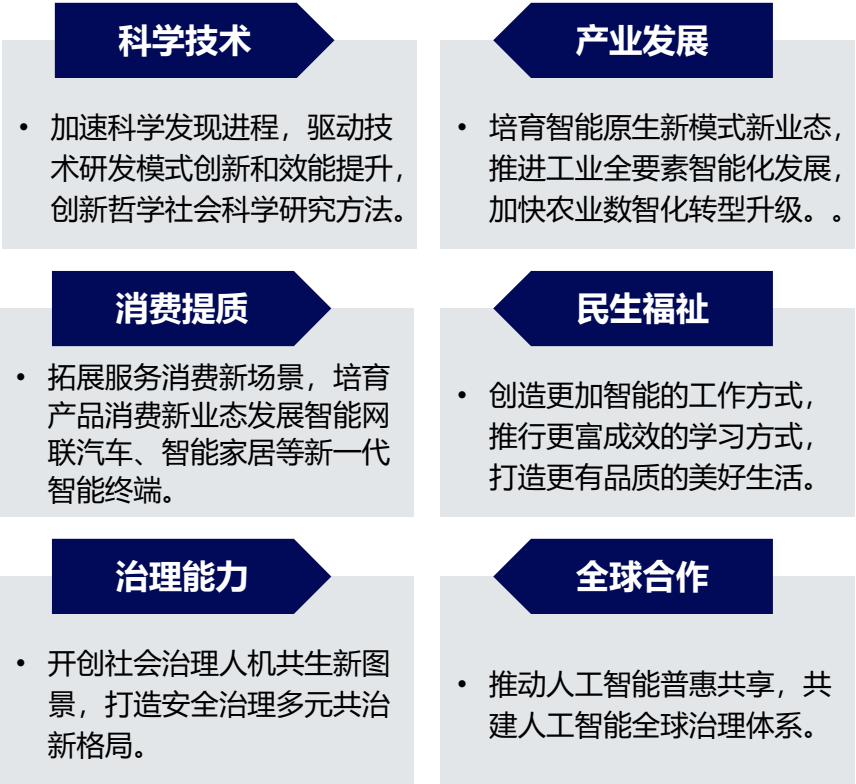
- 一、整体发展态势
- 二、评估模型研究
- 三、FAIS研究计划介绍

■ 2025年08月21日，国务院正式印发《关于深入实施“人工智能+”行动的意见》（国发〔2025〕11号）。推动人工智能与经济社会各行业各领域广泛深度融合，重塑人类生产生活范式，促进生产力革命性跃迁和生产关系深层次变革，加快形成人机协同、跨界融合、共创分享的智能经济和智能社会新形态。

发展目标



六大重点领域



促进金融业务模式智能化变革

- 智能化金融服务：利用AI为客户提供更加个性化、智能化的金融服务。
- 创新金融产品：借助AI技术创新金融产品，支持智能原生企业发展。

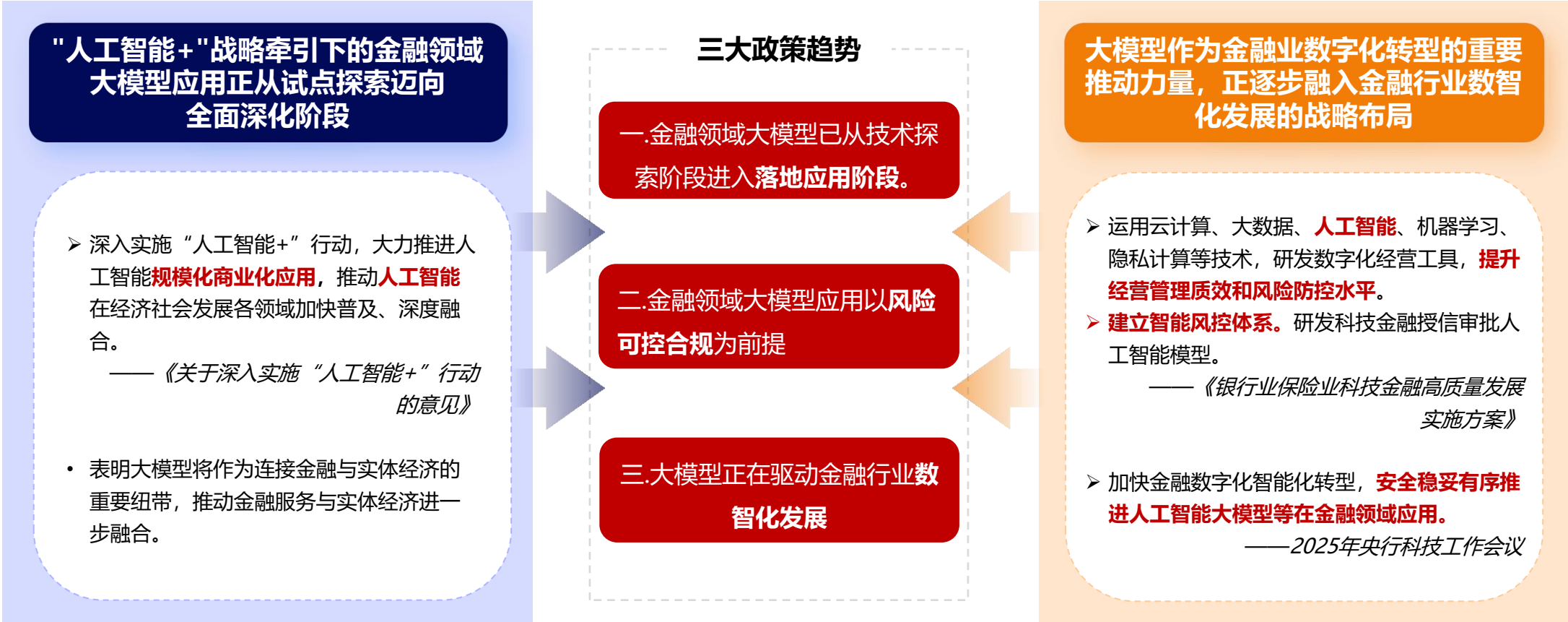
夯实金融业风险管理能力

- 风险预测与预警：运用AI技术对海量数据进行分析，提前预测信用风险、市场风险等。
- 反欺诈能力提升：利用AI的图像识别、行为分析等技术，可有效识别欺诈行为，保障客户资金安全。

提升金融业运营管理效率

- 流程自动化：利用AI技术实现自动化审批和风险评估。
- 资源优化配置：借助AI对业务数据的分析，能够更精准地配置资源，优化网点布局和人员安排。

■ 2026年，金融领域大模型将作为金融业“新质生产力”的核心载体，在国家战略引领下加速金融业数字化智能化转型，重构金融服务生态，释放数字经济新动能。



- 当前，金融业面临**行业发展转变、客户需求升级、风险管理复杂、运营效率待提升**等挑战，亟需通过数字化、智能化技术的深度融合应用，**重构金融业务价值链**。

金融业特点及挑战

行业发展转变快

金融业面临业务增速放缓、产品利润空间逐步压缩现状，行业正从传统的“评估与服务模式”向“**可预测、个性化和生态化模式**”转变。

客户需求升级

金融业传统产品与服务模式难以满足现代消费者需求，客户期望获得**个性化的解决方案**。

数据安全需求

金融业务涉及大量个人和企业隐私信息，且金融机构需要遵守各国数据保护法规，亟需投入大量资源确保**数据安全和隐私保护**。

市场竞争加剧

越来越多科技公司加入金融市场，利用大数据、人工智能等技术提供**创新的金融产品和服务**。灵活的定价策略和更高效的服务流程，迫使金融机构必须加快自身创新步伐。

风险管理复杂

金融业**高度依赖风险评估和管理**，当前风险因素复杂多变，传统风险评估和管理方法难以适应新形势需要。

运营效率提升需求

传统的金融营销模式，效率低下且存在主观判断错误问题，亟需从传统人工操作模式转型更加智能化、自动化服务模式。

应对挑战的解决之道

数智化转型成为金融业应对挑战和长远发展的必然选择

01. 提升客户体验

- 深入分析客户需求，提供更加个性化、定制化保险产品和服务。

02. 优化风险管理

- 通过大数据分析，更准确评估风险，制定合理的风险管控措施。

03. 提升运营效率

- 通过自动化和智能化数据处理和分析，减少手工操作，降低错误率，提高业务效率。

04. 加强合规管理

- 通过智能化技术，实现实时监控业务流程，确保合规性，降低违规风险。

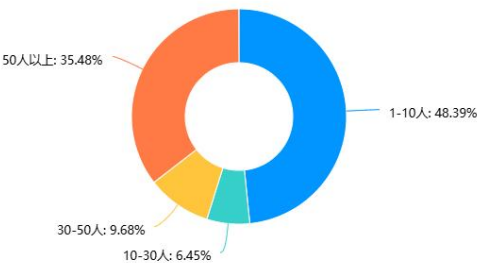
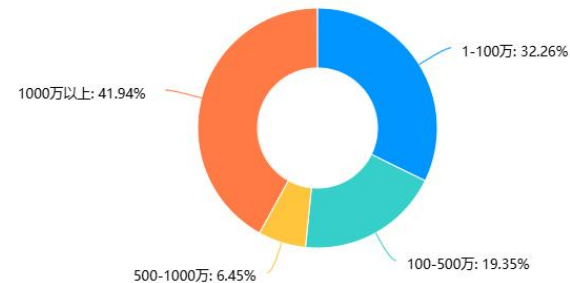
05. 实现创新发展

- 通过利用新技术，开发新业务模式，拓展新市场，提升竞争力。

■ 经过调研显示，目前金融机构根据监管要求、技术成熟度、投入产出等因素主要采用“**全心投入、先内后外、从易到难、场景迁移的方式**”落地大模型金融应用，大模型在提升**业务创新驱动、改善用户体验、流程自动化**等方面发挥重要功效。

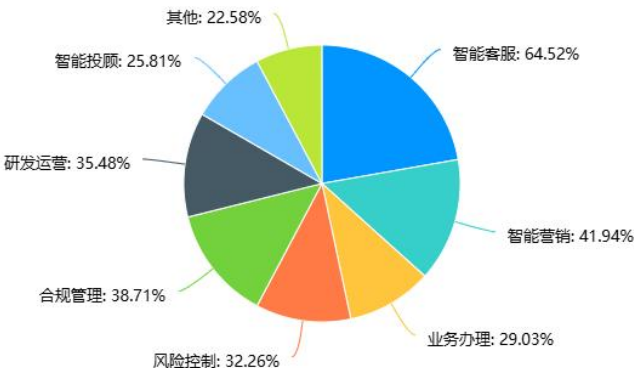
大模型投入

机构投入超过千万的超过40%，人员投入超过50人的超过35%。



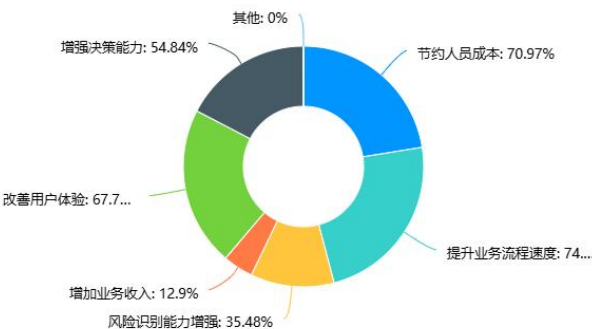
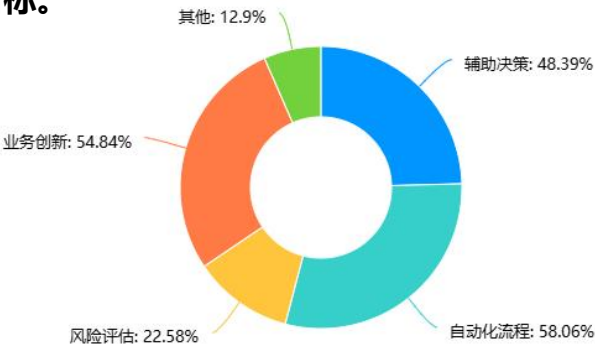
场景应用落地

机构对内场景落地较多，有三分之一的机构把大模型集成到**核心业务流程**中。



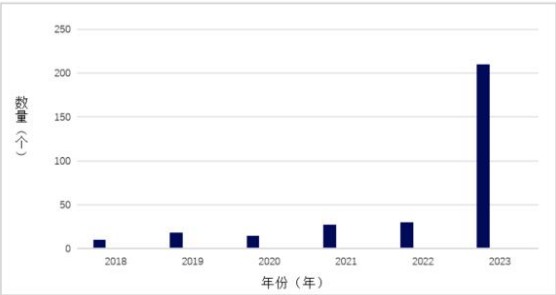
大模型发挥功效

目前大部分机构利用大模型主要为达到**流程自动化、辅助决策、业务创新、风险评估**等目标。



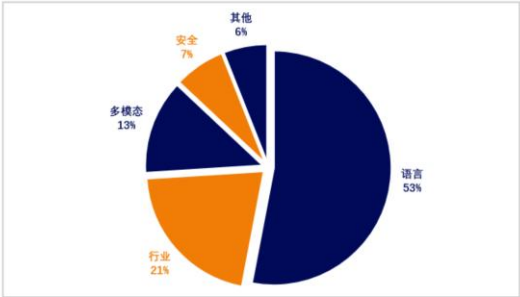
■ 当前大模型基准测试体系数量不断攀升，目前产学研各界发布的大模型基准测试大都侧重于模型的通用能力，基本上可以总结为涵盖知识能力、语言能力、推理能力、理解能力、安全能力等能力，缺少与具体业务结合的垂直领域的的能力评估。

大模型基准测试爆发



来源：中国信息通信研究院

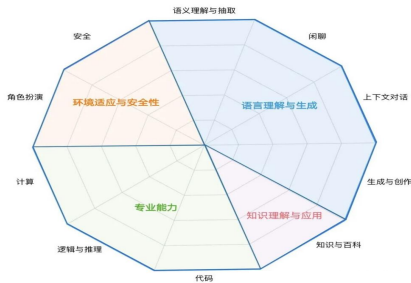
• 从发布时间来看，仅2023年一年出现的大模型基准测试数据集的数量远超过之前5年，达到200多个。



来源：中国信息通信研究院

• 从测试领域来看，以大模型的通用语言类评测为主，占比超过50%，面向行业类的评测数据集也开始迅速发展。

聚焦通用能力



SuperCLUE



FinEval评测体系

A.知识能力

衡量模型对广泛知识的了解程度，关注于模型在不同教育层次的学科知识掌握情况，从义务教育、高等教育以及职业教育等角度，通过对各级学科进行分类，构建完整的学科能力。

B.语言能力

关注模型在理解和生成语言方面的能力，包括对字词的精准辨识、分类、含义解释以及新词创造；对语句、段落的语法结构进行解析和错误纠正；以及将一种语言翻译为另一种语言的能力。

C.推理能力

该维度针对模型的逻辑分析与问题解决能力。模型的推理能力不仅涉及数学计算和逻辑判断，还包括对复杂问题的因果推断和程序代码的生成与 debug 能力。

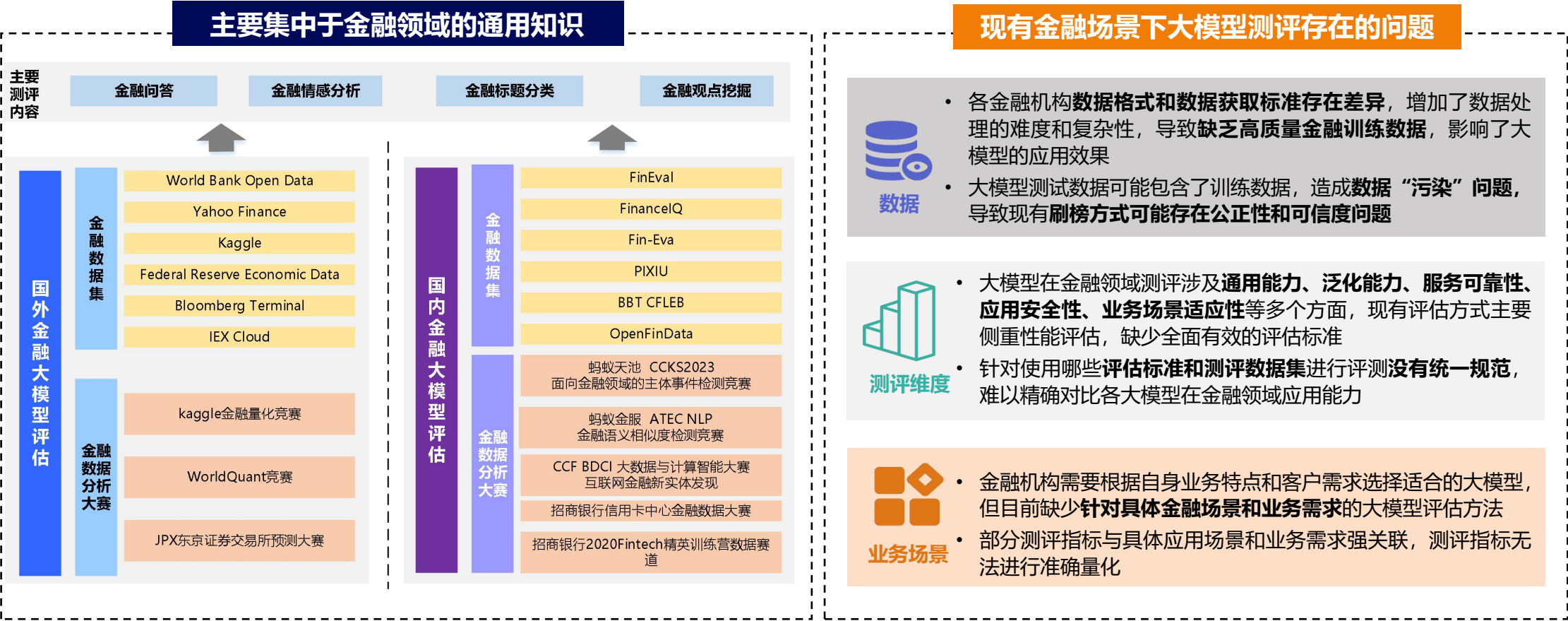
D.理解能力

评估模型对于给定文本的深入理解程度，包括对文字含义的把握、主旨的抽取、语境的解读以及创意写作。

E.安全能力

确保模型的输出不仅符合技术要求，还要符合社会和道德规范。包括了软件与应用、网络与系统保护、安全分析以及漏洞防护等技术能力，还有伦理、道德、合规、符合事实等能力。

- 目前金融领域大模型应用处于快速发展阶段，亟需可实施落地的标准规范和指南，但目前缺乏针对具体金融场景和业务下的大模型应用评估方法，致使金融机构进行大模型应用效果评价存在困难，无法适配客户精细化的大模型应用场景需求。



目录 contents

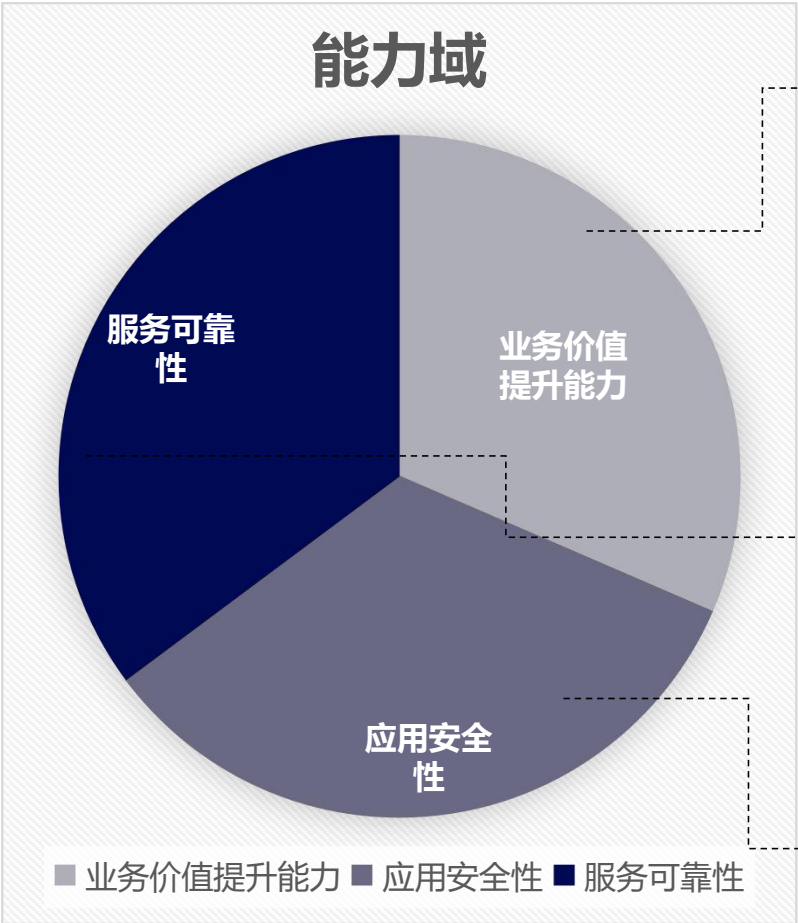
- 一、整体发展态势
- 二、评估模型研究
- 三、FAIS研究计划介绍

- 2024年11月29日，2024（第七届）金融科技产业大会在北京国际会议中心召开。会上，中国信通院联合中国银行业协会、工商银行、农业银行、邮储银行、交通银行、广州银行、国泰君安证券、中金公司、中信建投证券、招商证券、中金财富、山西证券、博时基金、易方达基金、人保财险、阳光数智等单位，共同启动《面向金融典型业务场景的大模型应用能力评估规范》（以下简称“FLMM”）首批评估工作。



- 为帮助金融机构了解自身应用系统现代化建设情况，中国信通院、中国通信标准化协会联合金融机构以及行业上下游科技厂商共同编写了《面向金融典型业务场景的大模型应用能力评估规范》。
- 本次会议上，工商银行、农业银行、邮储银行、交通银行、广州银行、国泰君安证券、中金公司、中信建投证券、招商证券、中金财富、山西证券、博时基金、易方达基金、人保财险、阳光数智等多家参编单位及参评单位代表上台参与仪式，共同见证了FLMM首批评估正式启动。

- 面向金融典型业务场景的大模型应用能力评估规范（FLMM）的三大能力域：业务价值提升能力V、服务可靠性R、应用安全性S，分别体现了大模型的业务价值、服务价值和管理价值。



业务价值：直观彰显

✓ **业务价值提升能力（2度+3率+3能力）**：分别从**业务维度、客户维度和机构内部维度**，关注金融机构在典型业务场景使用大模型时给业务带来的价值的提升。包括业务契合度、业务效能提升率、决策增强力、业务收益提升率、业务创新能力、客户满意度、客户市场拓展力、成本节约率。

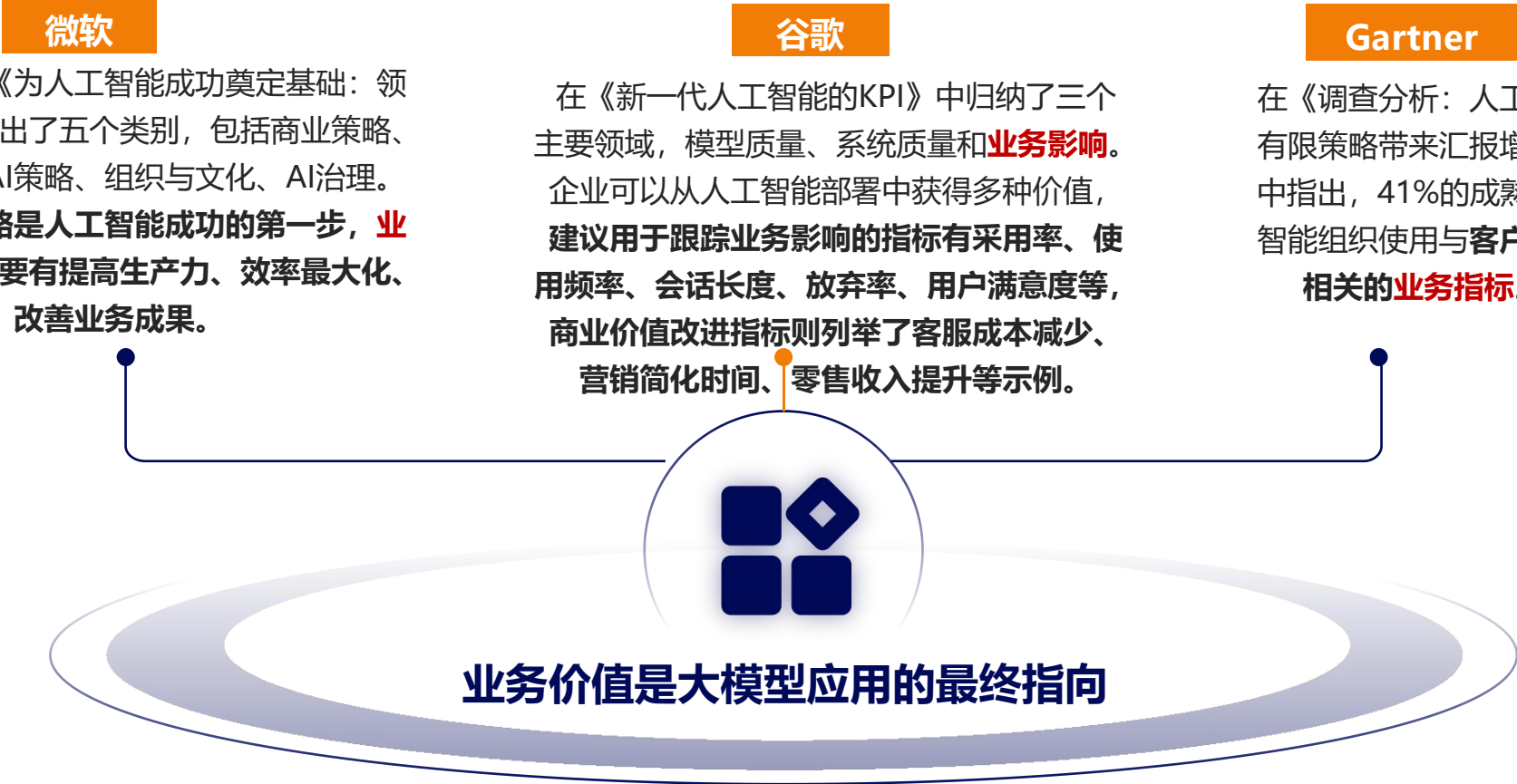
服务价值：技术基石

✓ **服务可靠性（6性+1能力+1时间）**：关注金融机构在典型业务场景使用的大模型本身的技术服务能力。包括易用性、稳定性、计量准确性、平均无故障时间、工具调用能力、可维护性、可拓展性以及兼容性。

管理价值：有力保障

✓ **应用安全性（3能力+2性）**：关注金融机构在典型业务场景使用大模型时的安全防护能力。包括隐私保护能力、可审查性、合规性、风险管理能力以及幻觉避免能力。

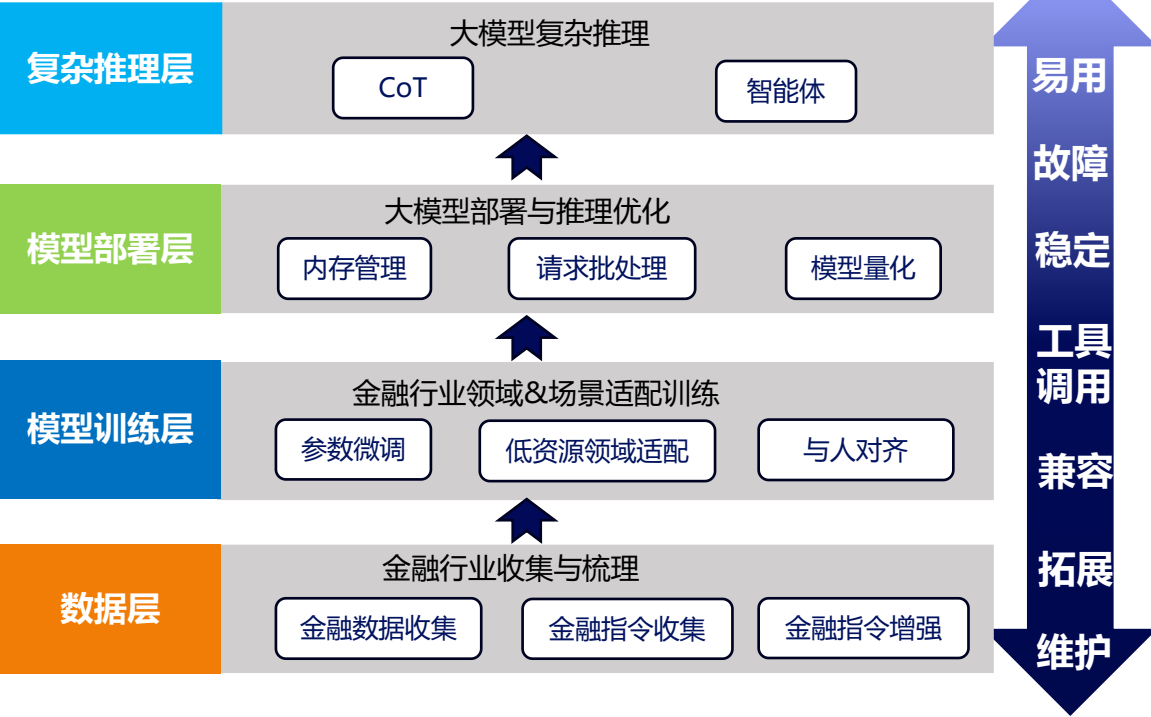
- 从头部机构的研究和实践经验来看，目前评估大模型应用成功的方法和维度有所差异，但是存在一个基本共性：**业务价值是衡量大模型成功的最终指向。**



- 大模型技术服务的可靠性是金融场景大模型发挥有效作用的**基础能力**，这些能力主要体现在大模型开发和部署的**全流程**中，**无故障、稳定、兼容、拓展、维护、工具调用、易用**等能力贯穿全流程。

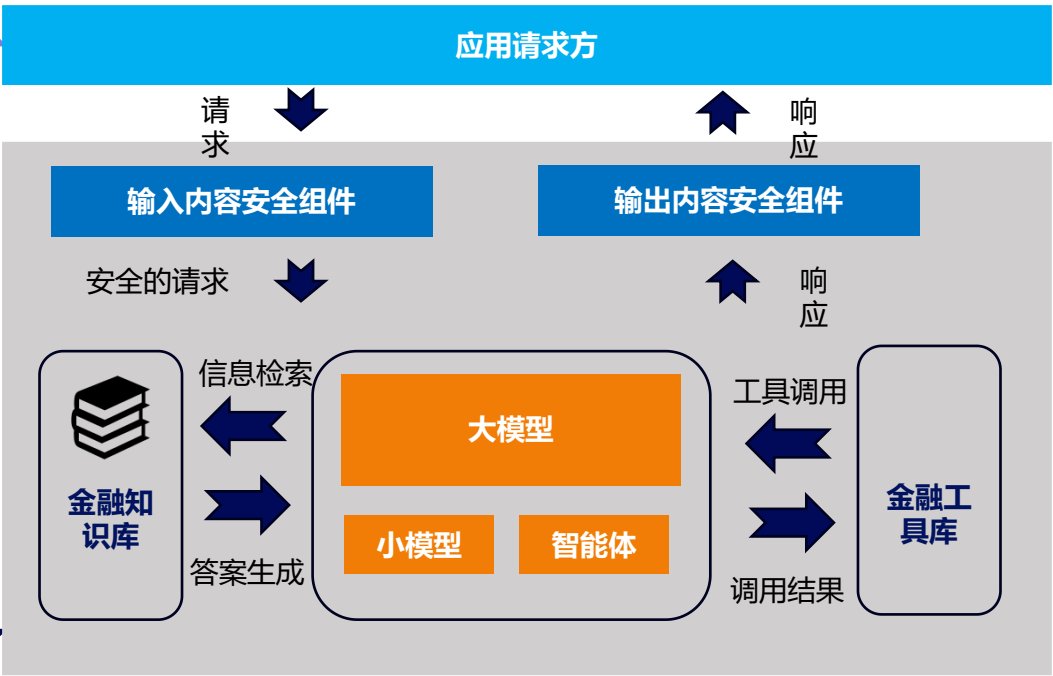
金融场景大模型开发技术框架

- 金融场景大模型开发流程从数据收集和处理开始，通过领域适配训练使模型理解金融语境，然后通过性能优化确保模型的实用性和高效性，接着处理幻觉问题以提高事实性，最终实现复杂推理的能力。



金融场景大模型应用框架

- 在具体应用中，要以大模型为核心，结合金融专业知识库、金融专业工具库、智能体、安全合规组件等，进一步构成一个可满足金融领域安全应用要求的应用系统。



- 大模型在金融相关业务应用中有几大安全风险，其中包括全流程的**隐私泄露**风险以及**模型攻击**；数据收集处理、适配与参数微调以及推理过程的**可解释性**问题；针对推理过程和生成内容的**可溯源性**及生成有害内容等，通过在数据收集与处理、模型训练与推理、模型后处理等阶段采取相应的安全防护措施达到风险防控的作用。

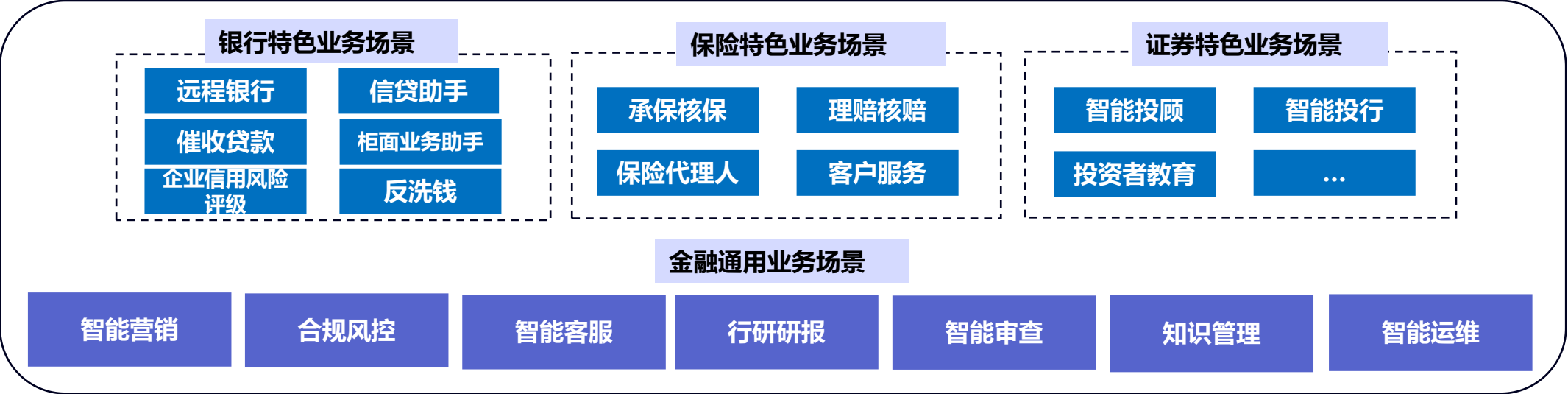


大模型开发和应用流程中的安全防护措施

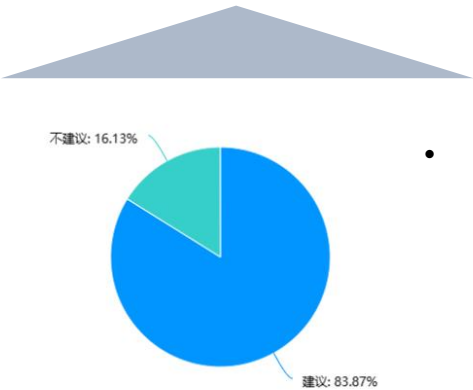
安全风险	具体内容
隐私泄露	从模型开发到模型应用的过程中均有可能涉及用户隐私信息，而这些隐私信息不仅包含敏感的个人敏感信息，更包括某些用户的资产信息。
模型攻击	针对应用在金融领域的大模型的攻击不仅会引起内部的风险漏洞，更可能引发对外的舆情，从而影响金融机构的日常管理。
可解释性问题	大模型训练和推理过程均具有黑箱性质，并且复杂度较高，尤其是包含十亿多参数的大模型，很难显化其内部推理联合决策过程，并针对特定输出给出解释。
可溯源问题	在金融领域，可溯源性的重要性除了防止欺诈行为以及防止假文本泛滥引发的财经谣言外，更重要的是，对于应用大模型的投研生态而言，防止大模型“一本正经地胡说八道”。
有害内容生成	在大模型金融领域的生成内容中，“追涨杀跌”等不符合金融价值观的内容与歧视、色情、暴力等不当内容均可被视为有害内容，极有可能引发对客的舆论危机与行业监管侧负面影响。

- [illegible]

- 根据大模型在金融业务场景应用的**广泛度、成熟度、优先性**等，选取**金融通用业务场景**和**银行、保险、证券特色业务场景**作为评估的场景对象。然后根据场景的特点，选择**适用的能力项**作为测评指标。



- 调研结论：**超过80%**的机构建议，分为通用场景和特色场景来设定评价指标。



- 优势：场景可以根据技术以及实际应用发展情况进行**灵活补充和完善**。

能力项	适用性
用户需求匹配度	√
业务目标一致性	√
业务流程优化效果	√
业务自动化提升率	√
业务响应速度提升	√
业务处理能力增强	√
资源利用率提高	√
人力成本节约	√
...	...

不同场景选择适用的能力项作为测评指标

- 不同的金融应用场景对大模型提出了多样化的能力要求，每个方面都是确保模型成功应用的关键。通过细致地分析这些场景特征，可以为选择合适的模型架构和训练策略提供指导，从而推动大模型在金融行业的深入应用。

案例名称	案例介绍	业务场景分类	技术实现难度	精确度	实时性	客户易用性	数据隐私与合规性要求
智能场外交易发现平台	基于大模型理解能力，注入机构内部知识文档与询价样本，训练智能机器人自动完成多轮询价过程，形成结构化数据传输到交易系统。	智能营运	高	低	高	中	高
员工知识助手	利用大模型开发多轮对话驱动的绩效知识助手，辅助员工操作。	智能营运	中	高	低	低	高
智能客服	利用自然语言处理技术理解客户咨询，回答常见问题，提供24/7服务，提升客户体验。	智能营销	中	低	高	高	高
智能产品推荐	运用数据分析和人工智能算法，实现目标客户群的精准识别、个性化产品推荐和定制化营销策略的自动化部署。	智能营销	中	中	高	高	高
投研信息助手	利用大模型+信息管道技术，对市场评论、会议纪要、路演、调研报告等进行不间断和规模化处理及二次加工，提升投研人员的信息处理效率。	智能投顾	中	中	中	中	高
投资风险管理	运用大模型深度解析宏观经济指标、行业动态、企业财务报告及市场情绪等多元数据，从历史数据中精准挖掘潜在风险模式与趋势。通过识别投资组合中的资产配置风险并及时预警，提升投资的稳定性与安全性。	智能风控	中	低	中	中	高
智能代码助手	为金融行业客户提供多模态、安全合规、流畅高效的AI代码辅助能力，提升研发效率。	智能代码	高	高	低	中	中

■ 评估模型通过给出计算公式以定量的**客观评估方式**评估应用能力，无法通过定量方式评估的能力项，则通过**定性描述**的主观评估方式来评估。

序号	指标	解释	公式
1	业务贡献度	业务目标一致性是指大模型在金融业务场景中的应用效果与预期业务目标的匹配程度，具体可以通过业务贡献度来衡量。	业务贡献度=（应用后业务指标-应用前业务指标）/（应用前业务指标）×100%
2	业务自动化提升率	通过应用大模型技术所实现的业务流程自动化程度的提升比例。	自动化提升率=（应用大模型后的自动化任务数量-应用大模型前的自动化任务数量）/（应用大模型前的自动化任务数量）×100%
3	业务响应速度	响应时间计算用户从发起请求到系统或应用程序返回结果的时间。	$ES_T = R_T_{finish} - R_T_{start}$ ，式中： ES_T ——响应时间； R_T_{finish} ——大模型系统或应用程序返回结果的时间（非流式）或大模型系统或应用程序返回第一个字节结果的时间（流式）； R_T_{start} ——用户发起请求的开始时间
4	资源利用率	大模型在金融业务场景中运行时，对计算资源（如CPU、GPU、内存等）的使用效率。	资源利用率=实际使用的资源量/分配的资源总量×100%
5	客户忠诚度	客户忠诚度是指衡量客户对金融机构产品和服务的满意度和持续使用意愿的关键指标。具体可以使用净推荐值（Net Recommendation Value, NRV）和用户留存率（User Retention Rate）指标进行评估。	净推荐值（NRV）=（推荐者数量-批评者数量）/（总受访者数量）×100%； 用户留存率=（特定时间段内继续使用应用的用户数）/（特定时间段内首次使用应用的用户数）×100%。
6	成本收益比	大模型在金融业务中的应用所带来的收益与其投入成本之间的比值。	成本收益比 = 总收益/总成本
7	平均无故障时间	大模型应用在两相邻故障间隔期内正确工作的平均时间。	设有一个可修复的大模型应用在使用过程中，共计发生过N次故障，每次故障后经过修复又和新的一样继续投入使用，其工作时间分别为T1、T2、T3、T4、T5.....Tn：那么大模型应用的平均故障间隔时间，也就是平均寿命为Q为： $Q=MTBF=1/N*\sum_{i=1}^N Ti$ 。

- 根据能力项的特点，设置一种能力项得分规则：1-5分的**5层计分方式**，最低分为1分，最高分为5分。最终根据不同场景选取的能力项计算最终得分，最终得分满分100分。

▶▶ 示例：业务目标一致性

得分	具体要求
1分	应用效果与预期业务目标基本一致，但存在偏差； 业务贡献度[0,5%]。
2分	应用效果与预期业务目标较为一致，但仍有改进空间； 业务贡献度(5%,10%]。
3分	应用效果与预期业务目标非常一致，较好地满足业务需求，推动业务发展； 业务贡献度在(10%,20%]。
4分	应用效果与预期业务目标高度一致，能够显著提升业务绩效； 业务贡献度在(20%,30%]。
5分	应用效果与预期业务目标完全一致，能够极大地提升业务绩效并创造新的业务价值，或引导业务目标的优化和调整； 业务贡献度(30%,100%]。

▶▶ 最终得分计算公式

$$\text{得分} = \frac{\sum \text{能力项得分}}{\text{选取能力项数量} \times 5} * 100$$

▶▶ 评估结论

得分	成熟度
90分（含）至100分（含）	5级
80分（含）至90分	4级
70分（含）至80分	3级
60分（含）至70分	2级
60分以下	1级

■ FLMM评估从报名申请至最后颁发评估证书共分为九个步骤，具体流程及进度说明如下：



■ FLMM评估方法主要包括人员访谈、材料审查、模拟演示三种，对金融领域大模型的不同能力子项将适配相应评估方法。

人员访谈

根据标准项对金融领域大模型建设相关工作人员进行访谈，通过问答互动的形式进行定性评估

测试材料：《人员访谈问答记录表》

模拟演示

收集演示截图，并现场模拟演示，评估应用系统相关能力子项技术情况

测试材料：《模拟演示文件包》

材料审查

对金融领域大模型建设相关的制度、标准、说明、视频、图片等佐证材料进行内容审查

测试材料：《材料审查文件包》

标准评估方法	
测试编号	XXXX
审查项目	监控管理
自评等级	X_级
审查方式	☐ 人员访谈 ☒ 材料审查 ☒ 模拟演示
评估要求	L1 级别评估方法： JSJGXDH-KGCX-JKGL-L1-1. 主要依赖手动或基础的监控工具进行问题诊断和管理。 ——模拟演示（需演示并提供证明截图）： 现场演示并提供材料说明使用的监控工具软硬件型号、配置、版本 1. 软硬件使用情况的截图信息 【演示截图：JSJGXDH-KGCX-JKGL-L1-1-软硬件使用情况截图】
	L2 级别评估方法： JSJGXDH-KGCX-JKGL-L2-1. 具备基本监控管理能力。 ——模拟演示（需演示并提供证明截图）： 现场演示并提供材料说明使用的监控工具软硬件型号、配置、版本、能力 1. 软硬件截图信息 【演示截图：JSJGXDH-KGCX-JKGL-L2-1-软硬件截图】 2. 功能截图信息 【演示截图：JSJGXDH-KGCX-JKGL-L2-1-功能截图】
	L3 级别评估方法： JSJGXDH-KGCX-JKGL-L3-1. 具备全面监控系统各个组件和关键业务流程的性能指标、自动化告警和故障排查功能。 ——模拟演示（需演示并提供证明截图）： 现场演示并提供材料说明使用的监控工具软硬件型号、配置、版本、功能 1. 软硬件截图信息 【演示截图：JSJGXDH-KGCX-JKGL-L3-1-软硬件截图】 2. 功能截图信息 【演示截图：JSJGXDH-KGCX-JKGL-L3-1-功能截图】

统各个组件和关键业务流程的
据分析和机器学习能力预防风
险。

型号、配置、版本、功能

件使用情况截图】

截图】

的监控管理能力，具备大数据

型号、配置、版本、功能，是

件使用情况截图】

截图】

可用性。

依据】

■ 评分主要针对3个能力域涉及的59个能力项提出不同能力分级要求。每个能力均对应5级能力分级的描述，从1级至5级要求依次提高。第N+1级能力的要求默认应满足并覆盖前N级能力要求。

差异化评估指标

- 不同业务场景评估功能差异较大，针对不同业务场景将采用差异化指标体系。

评分细则

- 各能力项涉及的1-5级能力分级对应1-5分得分，如该能力项完全未满足，得0分。各能力项分级时，金融机构应完全达到某能力等级的具体要求，方可取得对应级别分数。如同时满足两级的部分能力，亦不可得分，仅在全部能力满足后方可得分。

注：实际得分指该能力域或能力项涉及的能力项得分之和。

能力域		能力子域	能力域得分	能力项	能力项得分
业务价值提升能力 实际得分/5	业务价值提升能力 实际得分/5	业务契合度V1	实际得分/2	用户需求匹配度	0分-5分
				业务目标一致性	0分-5分
		业务效能提升率V2	实际得分/3	业务平均响应速度提升	0分-5分
				业务处理能力增强	0分-5分
				业务自动化提升率	0分-5分
	决策链路支持力V3	实际得分/4		管理决策支持力	0分-5分
				业务决策支持力	0分-5分
				决策链路覆盖率	0分-5分
				多维度分析能力	0分-5分
				成本收益比	0分-5分
	业务收益提升V4	实际得分/3		业务收入提升率	0分-5分
				业务量提升率	0分-5分
				业务模式创新能力	0分-5分
				产品服务创新能力	0分-5分
				用户活跃度	0分-5分
	客户维度	用户满意度V6	实际得分/6	月活跃用户增长率	0分-5分
				用户交互时间	0分-5分
				点赞率及点踩率	0分-5分
				复制内容率	0分-5分
				客户忠诚度	0分-5分
机构管理维度	成本节约率V7	实际得分/3		人力成本节约率	0分-5分
				时间成本节约率	0分-5分
				资本支出节约率	0分-5分

■ 根据金融机构典型业务场景的大模型应用能力的评估情况，将应用能力成熟度等级划分为1-5级。

1级：基础级

金融机构在典型业务场景的大模型应用**尚处于起步阶段**，功能有限，主要用于内部测试和验证。在业务价值提升、应用安全性和服务可靠性方面**具有很大的提升空间**。

2级：成长级

金融机构在典型业务场景的大模型应用**已经有一定基础实践**，大模型具备基本的业务适用性和稳定性，可以处理实际业务场景，但还需进一步优化。在业务价值提升、应用安全性和服务可靠性方面**成长性高**。

3级：提升级

金融机构在典型业务场景的大模型应用**相对成熟**，大模型在业务场景中表现良好，效率和准确性得到显著提升，用户反馈积极。在业务价值提升、应用安全性和服务可靠性方面的**表现稳定**。

4级：优秀级

金融机构在典型业务场景的大模型应用**较为成熟**，大模型完全符合业务需求，能够提供高级功能和优化，成为业务流程中不可或缺的一部分。在业务价值提升、应用安全性和服务可靠性方面的**表现优秀**。

5级：卓越级

金融机构在典型业务场景的大模型应用**能力卓越**，大模型在所有评价维度上都达到行业领先水平，创新推动了业务模式的变革。在业务价值提升、应用安全性和服务可靠性方面的表现卓越，**具备显著的核心竞争力**。

标准设计与评估过程始终坚持科学性、全面性、客观性、成长性四大原则。

正在参与标准评估工作以及已完成单位

检验报告

产品名称: 智能投顾助手 IC Copilot
产品型号: v1.0
委托单位: 中国中金财富证券有限公司
生产单位: 中国中金财富证券有限公司
检验类别: 委托检验

检验样品照片

产品名称: 智能投顾助手 IC Copilot 样品型号: v1.0
拍摄地点: 北京市西城区知春路1号 拍摄日期: 2025年8月15日
号学院国际大厦30层3006

面向金融典型业务场景的大模型应用能力
评估检验证书

单位名称: 中国中金财富证券有限公司
单位地址: 北京东城区东直门南大街23号国际金融中心31层
产品名称: 智能投顾助手 IC Copilot
产品型号: v1.0

符合性描述: “经中国信通院研究院的检验, 贵单位的智能投顾助手 IC Copilot 通过了 FT-Y02-0030-01《面向金融典型业务场景的大模型应用能力评估》的检验, 符合业务价值提升能力第三级要求。特发此证。”

测评报告及证书

中移金融科技
CM Financial Technology

中国电信
CHINA TELECOM

CASHWAY 恒银科技
智慧银行解决方案提供商

ICBC
中国工商银行

山西证券
SHANXI SECURITIES

东莞证券
DONGGUAN SECURITIES

国元证券
GUOYUAN SECURITIES

ZTE中兴

紫光恒越
UNISUYE TECHNOLOGIES

中国邮政储蓄银行
POSTAL SAVINGS BANK OF CHINA

交通银行
BANK OF COMMUNICATIONS

CMS 招商证券

百度云

金山云

H3C
数字化解决方案领导者

PICC
中国人民保险

中信建投证券
CHINA SECURITIES

上海期货交易所
SHANGHAI FUTURES EXCHANGE

CICC
中金公司

中金财富
CICC Wealth Management

国泰海通
GUOTAI HAITONG

CFCA

京东云

评估材料

技术文档

系统架构文档

技术规范文档

业务文档

需求分析文档

业务流程文档

运行数据

性能监控数据

业务运行数据

其他文档

软著专利

其他

XX机构评估调研问卷设计示例

示例

调查问卷

所属分类	问卷内容	填写要求
用户背景调查	1. 贵单位大模型及智能语音助手的基本部署情况? A. 尚未部署 B. 仅部署在部分业务 C. 部分部署 D. 全部部署 E. 正在部署中 2. 贵单位在部署过程中, 贵单位大模型可以完成的任务有哪些? A. 能够识别用户的意图并生成回复, 理解用户的意图并生成回复 B. 能够处理多个步骤的任务并生成回复 C. 能够根据用户提供的信息生成回复, 理解用户的意图并生成回复 D. 能够根据用户的需求生成回复, 理解用户的意图并生成回复 E. 能够根据用户的需求生成回复, 理解用户的意图并生成回复	
业务需求调查	1. 贵单位大模型及智能语音助手的基本部署情况? A. 尚未部署 B. 仅部署在部分业务 C. 部分部署 D. 全部部署 E. 正在部署中 2. 贵单位在部署过程中, 贵单位大模型可以完成的任务有哪些? A. 能够识别用户的意图并生成回复, 理解用户的意图并生成回复 B. 能够处理多个步骤的任务并生成回复 C. 能够根据用户提供的信息生成回复, 理解用户的意图并生成回复 D. 能够根据用户的需求生成回复, 理解用户的意图并生成回复 E. 能够根据用户的需求生成回复, 理解用户的意图并生成回复	
业务需求调查	1. 贵单位大模型及智能语音助手的基本部署情况? A. 尚未部署 B. 仅部署在部分业务 C. 部分部署 D. 全部部署 E. 正在部署中 2. 贵单位在部署过程中, 贵单位大模型可以完成的任务有哪些? A. 能够识别用户的意图并生成回复, 理解用户的意图并生成回复 B. 能够处理多个步骤的任务并生成回复 C. 能够根据用户提供的信息生成回复, 理解用户的意图并生成回复 D. 能够根据用户的需求生成回复, 理解用户的意图并生成回复 E. 能够根据用户的需求生成回复, 理解用户的意图并生成回复	

目录 contents

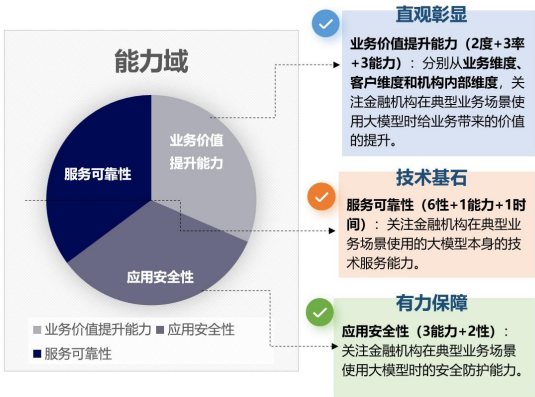
- 一、整体发展态势
- 二、评估模型研究
- 三、FAIS研究计划及标准工作介绍

- “基于金融业务场景的大模型应用研究计划”（简称“FAIS”）是由中国信通院金融科技研究团队发起的金融领域人工智能应用研究计划；
- 依托中国通信标准化协会金融科技推进委员会（TC604），持续开展金融智能体系列标准研究；

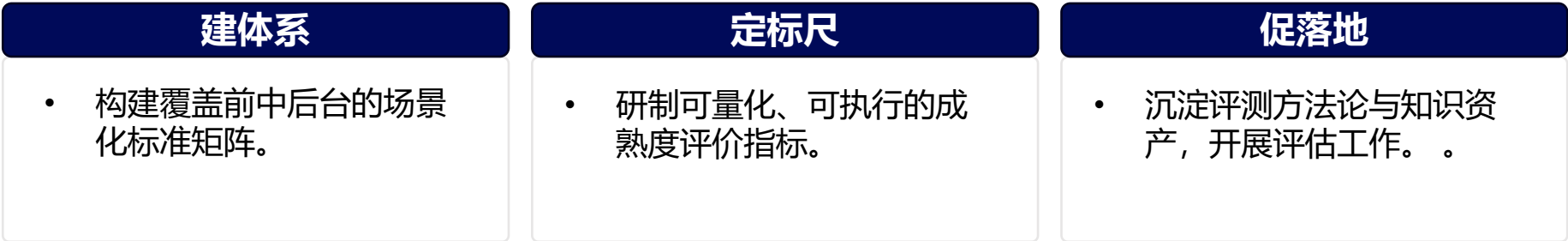
FAIS已开展工作



基于金融业务典型场景的大模型能力成熟度评估模型 (FLMM)



共同推进金融智能体标准体系建设



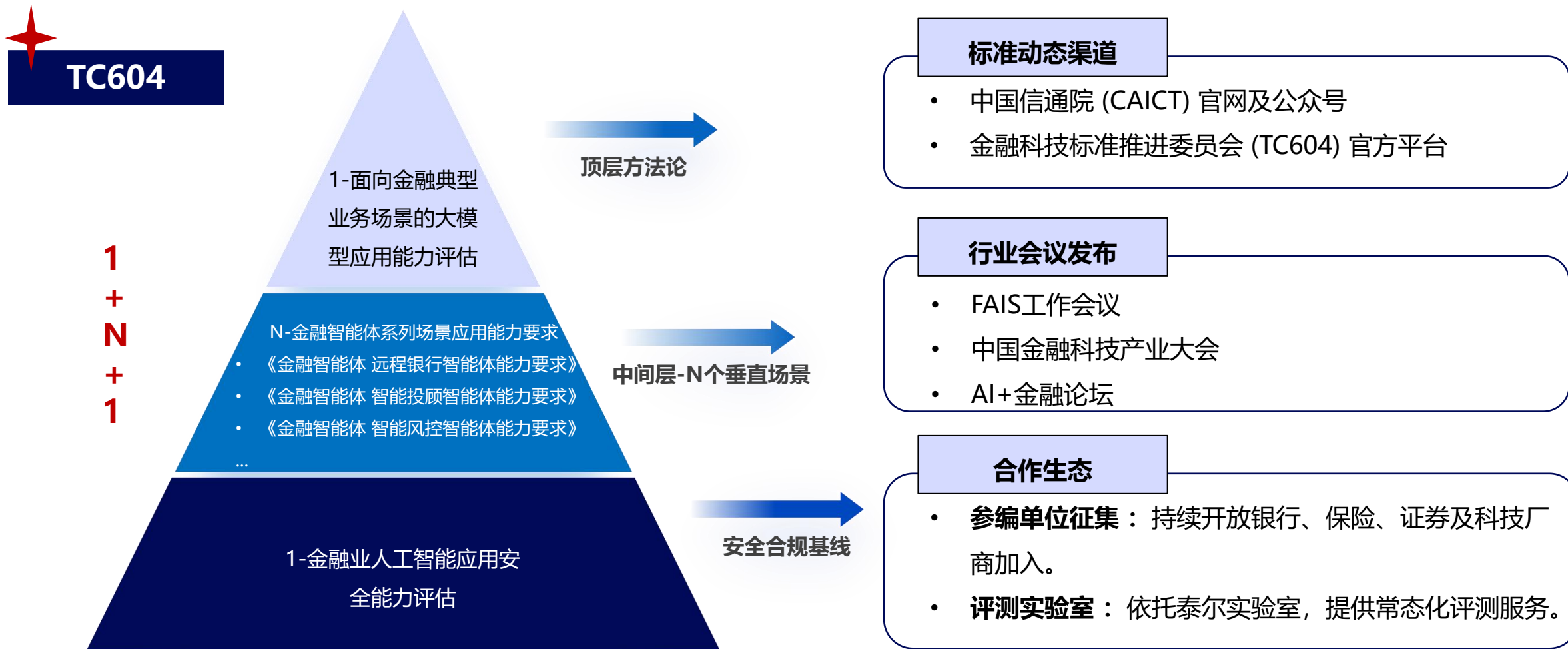
支撑

基于金融业务场景的大模型应用研究计划(FAIS)

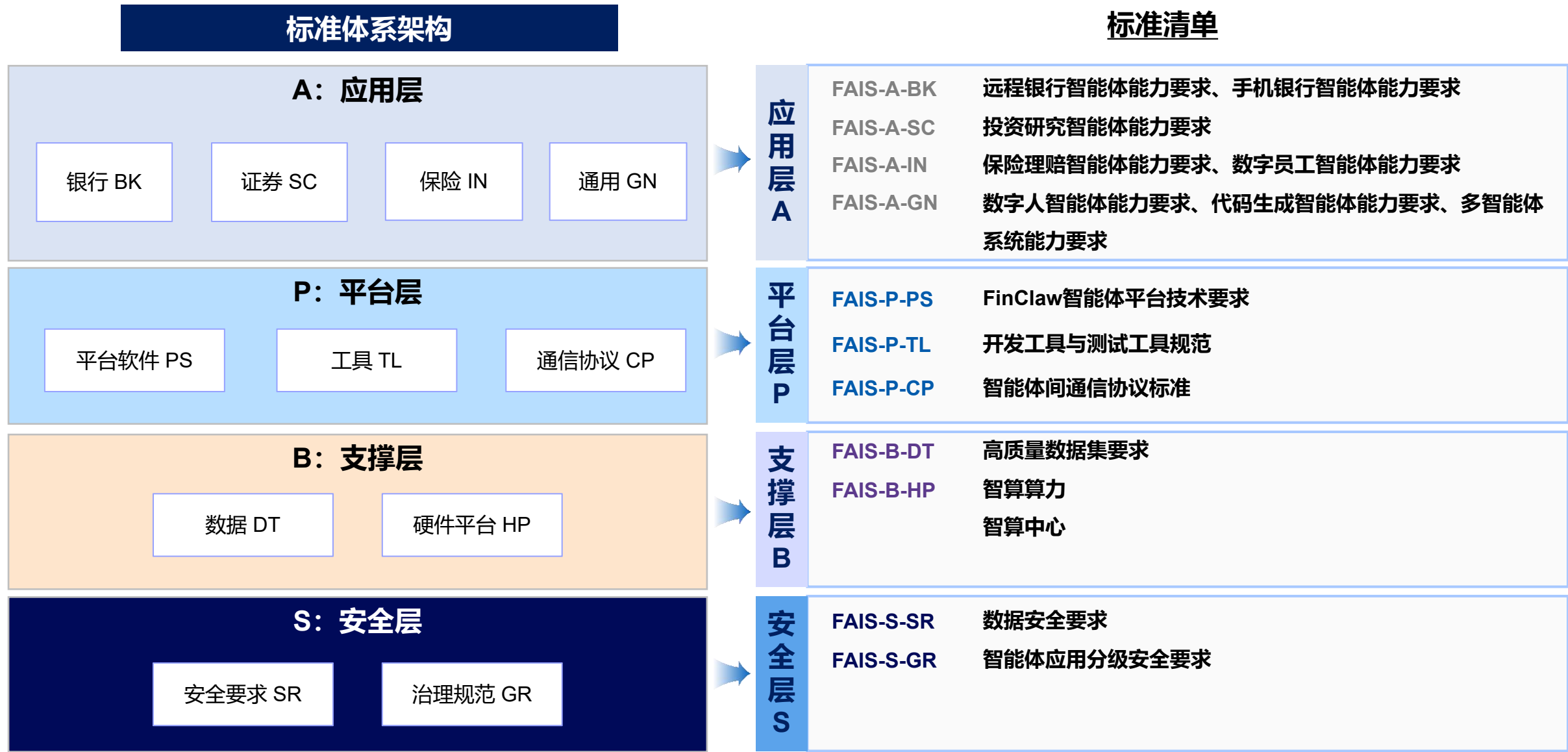
以场景为牵引、以安全为前提、以数据为支撑



- 依托“1+N+1”标准体系，构建从顶层设计到垂直场景落地的全闭环评估机制，并开展评估工作，通过官方渠道发布评估成果，共建开放、协同的金融智能体生态。



■ FAIS金融智能体系列标准包含应用层(A)、平台层(P)、支撑层(B)、安全层(S)。



■ 中国信通院FAIS计划依托金融科技标准推进委员会（TC604），开展金融智能体系列标准编制工作。

金融智能体系列标准清单			
金融业务专用场景	《金融智能体 远程银行智能体能力要求》	new!	✓ 已形成初稿，正在征求意见
	《金融智能体 手机银行智能体能力要求》	new!	✓ 征集ing
	信贷助手	hot	✓ 待启动
	营销助手	hot	✓ 待启动
	《金融智能体 投资研究智能体能力要求》	new!	✓ 征集ing
	智能投顾	hot	✓ 待启动
	理财助手	hot	✓ 待启动
	承保核保	hot	✓ 待启动
	保险定价	hot	✓ 待启动
	《金融智能体 保险理赔智能体技术要求》	new!	✓ 征集ing
	保险代理人助手	hot	✓ 待启动
	智能采购	hot	✓ 待启动
金融服务通用场景	《金融智能体 数字员工智能体技术要求》	new!	✓ 征集ing
	智能客服	hot	✓ 待启动
	智能运维	hot	✓ 待启动
	《金融智能体 FInClaw类智能体平台技术要求》	new!	✓ 待启动
	《金融智能体 代码生成智能体技术要求》	new!	✓ 征集ing
	合规风控	hot	✓ 待启动
	知识问答	hot	✓ 待启动
	办公助手	hot	✓ 待启动
	本体智能	hot	✓ 待启动
	《金融智能体 多智能体系统技术要求》	new!	✓ 已形成初稿，正在征求意见
	《金融智能体 智能体应用分级安全要求》	new!	✓ 征集ing

咨询规划	<u>金融领域大模型政策咨询</u> 提供金融领域大模型政策咨询服务 对企业关注的痛点提供咨询服务	<u>金融领域大模型建设中 长期规划</u> - 制定大模型建设关键举措规划 - 大模型建设体系规划等	<u>金融领域大模型调研及 技术趋势分析报告</u> 对国内大模型技术路线、技术水平、产业链生态及应用落地情况进行分析并编写报告	
能力建设	<u>基于金融业务场景的 大模型场景清单征集</u> 针对金融机构具体应用场景开展清单征集共工作	<u>基于金融业务场景的 大模型相关课题</u> 针对金融领域AI课题展开联合研究及编写工作	<u>基于金融业务场景的大模型 应用能力评估标准共建</u> 共建《基于金融业务场景的大模型应用能力评估模型》标准	<u>基于金融业务场景的 大模型应用能力评估</u> 依托标准，对不同业务场景提供成熟度评估。
产品服务	<u>金融信创运维大模型建设</u> 针对金融机构信创工作面临运维压力大、服务供给不足问题，联合产学研用各方推进金融信创运维大模型建设。	<u>产融大模型建设</u> 基于国家产融合作平台，开发面向产融合作应用场景的大模型	<u>金融领域大模型训练及 能力优化</u> 基于建设需求，协助机构开展大模型训练及能力优化	
市场推广	<u>金融领域大模型 展台展览活动</u> 基于信通院金融领域大模型展览展示厅资源，为企业提供展台席位。	<u>金融领域大模型 人才培养（含实地参访）</u> 联合相关单位对证券期货业、银行、保险业组织开展金融领域大模型人才培养。	<u>金融领域大模型 优秀案例征集</u> 征集并形成金融行业大模型典型案例集	

愿景

FAIS “智鸿” 专家计划旨在成为金融大模型领域的专业级人才枢纽，通过发掘、汇聚、赋能优秀人才，助推人才之间的沟通与协作，从而加速解决金融大模型落地实践中的问题难点，推动金融大模型的价值实现。

专家建设目标

打造专家能力矩阵

四大类能力	政策监管	场景应用
	技术产品	基础能力

建设多层次专家级别



感谢观看
THANKS

