

共享善治AI 智绘未来金融

——大模型在金融领域的价值、治理和生态进阶之路

引言

随着生成式AI技术的崛起，人工智能正以前所未有的力度和速度，推进技术创新，重塑产业未来。金融业一直以来都是技术应用的先行者，过去，移动互联网、云计算、大数据等技术，推动了金融业从信息化走向数字化，当下，人工智能技术的蓬勃发展将在金融行业引领新一轮变革。置身于AI技术浪潮中，每个参与者需要重新审视公司战略、业务形态、技术路径和治理机制，推动人工智能技术深度融入业务、重构业务，驶向智能化的星辰大海。

同时，伴随着生成式AI技术在金融行业的探索应用，我们也面临着诸多挑战，包括如何更好的释放AI在金融场景的价值，如何确保技术可靠可控，而不是放大风险，以及如何缓解技术应用的资源约束等，需要我们深入探讨和解决。在此背景下，我们与合作伙伴一起提出大模型在金融领域的应用和发展框架，发布《共享善治AI，智绘未来金融》白皮书，展望AI与金融场景融合的发展趋势，探讨负责任AI实现层次和技术生态演化路径。我们希望通过这份白皮书，与业界同仁共同探索生成式AI技术在金融领域应用的进阶之路，共绘智能化金融的未来蓝图。

未来已来，将至已至。蚂蚁集团秉持科技驱动和普惠初心，愿与各方携手，以AI新质生产力为笔，以金融五篇大文章为卷，共同书写金融行业数智化新篇章，为社会创造更大的价值。

—— 李振华 蚂蚁集团研究院院长

金融行业天然具备数据和信息密集型的特点，在数字化成熟度方面处于领先地位。此外，金融行业的数字化投入持续稳步增长，汇集了大量具备数字化技能的人才。这些优势使得金融行业在AI技术的应用和创新方面具备独特的条件，能够在推动技术革新和提升行业效率方面起到示范作用。

尽管大模型在金融业务领域展现了巨大的潜力，但其在金融行业的全面应用仍面临诸多挑战。目前通用大模型与行业相结合整体仍处于初级阶段，此外金融行业因其严监管和强合规要求，使得大模型的落地应用需要更加谨慎，周期会更长。总体来看大模型的行业生态系统尚未完全成熟，其在行业内的广泛应用需要克服技术、行业演变、监管和科技伦理等多方面的挑战。因此，如何实现大模型在金融业务场景的价值最大化，并推动金融行业的数智化升级，仍将是未来的重要课题。

本白皮书是IDC中国首次联合合作伙伴针对大模型在金融领域的应用发布的白皮书。内容详细阐述了金融行业在大模型应用中的现状和挑战，并提出了推动大模型技术在金融领域广泛应用的路径和策略。通过对AI价值实现、负责任AI、大模型生态演化的深入分析，旨在为金融机构提供理论和实践的双重支持，助力金融机构顺利迈入未来金融的大模型时代，推荐每一位关注金融行业未来发展的读者深入阅读本白皮书。

随着大模型的迅猛发展，金融行业正处于深刻变革的关键节点。大模型不仅是技术创新的产物，更是新质生产力的代表，通过其卓越的数据处理和分析能力，极大地提升了金融分析的深度和精度。大模型技术的引入和应用，正在对传统的金融业务模式产生深远而持久的影响，正在成为推动整个行业向更高效、更智能的方向发展的重要力量。

白皮书系统地提出了大模型在金融领域的应用和发展框架，探讨了大模型在金融领域的应用价值及其衡量方法；厘清了负责任AI的框架，为金融机构的大模型应用划出清晰的安全底线，并推动科技伦理共识；最后聚焦行业的生态演进，既深刻分析了金融行业生态现状，又科学严谨的提出了未来行业平台化生态的趋势判断。

这一应用与发展框架的提出，为翘首以盼大模型带来变革的金融行业注入了新的科学理念，也为大模型在金融领域的应用提供了全面且实用的参考。通过本白皮书，希望能鼓励金融学界、金融机构、科技企业共同探索新技术所带来的广阔前景，推动技术与行业应用的深度融合，助力中国金融行业更高效地顺应新质生产力发展的潮流，朝着更加智能化、普惠化的新时代迈进。

—— 刘莉亚 上海财经大学党委常委、副校长

目录

前言	01
01 AI价值实现	07
1.1 金融行业AI价值评估方法	08
1.2 AI价值层次和场景价值	09
1.3 金融机构AI应用现状	14
1.4 四大能力帮助金融机构实现AI价值进阶	17
02 负责任AI	24
2.1 什么是负责任AI	25
2.2 负责任AI的三大层次	25
2.3 负责任AI的现状洞察	32
2.4 负责任AI落地四大关键	35
03 AI应用生态演化	41
3.1 大模型生态的界定	42
3.2 行业生态模式演进洞察	43
3.3 目前行业生态的基本情况	57
3.4 推动生态演化的建议	61
结语和未来展望	66

前言

自2023年AI大模型元年至今，大模型领域迎来持续的技术突破和早期产品技术落地。例如OpenAI在2023年3月、9月和11月，相继推出了GPT-4、GPT-4V、GPT-4 Turbo，分别在多模态输入能力、生成内容的质量和准确性、长文本处理能力和推理速度等技术层面取得明显突破，在基础大模型表现方面被公认为遥遥领先。同时商业接口的开放与开源模型的推出，推动着大模型技术的加速落地。例如随着GPT-4 Turbo的推出，OpenAI降低了API调用费用，使得企业和开发者能够以更低的成本使用大模型技术。

除了国外基础大模型取得持续突破之外，国内的基础大模型发展也格外繁荣。国内的基础大模型参数规模2023年以来呈现快速增长趋势，在自然语言处理任务及多模态处理任务上取得了显著的性能提升。自ChatGPT发布以来，国内各类型厂商陆续发布多款参数规模在十亿量级以上的的基础大模型，其中以阿里、百度、华为等厂商为代表的部分本土厂商发布了大于千亿级别参数规模的模型。同时国内的跨模态大模型也取得深入发展和突破，实现多模态信息的融合和交互，如阿里云、百度，分别在CogView2、ERINE 4.0模型上实现了特定任务的突破性进展，展现了在文本与图像生成方面的创新潜力。

由于特定行业知识和专业术语的独特性，需要特定的行业大模型以更好地满足行业需求，行业大模型成为基础大模型发展的重要方向。在大模型的垂直行业应用层面，大模型技术在金融、医疗、汽车、制造、零售等行业开始探索落地应用，例如在金融行业，蚂蚁集团发布了agentUniverse的开源智能体框架，提供多智能体的协作模式，旨在帮助金融机构及开发者加快大模型在金融场景的研发落地。

随着市场对大模型的行业应用以及相关价值创造的关注与日俱增，企业、投资者、技术开发者都在积极探索大模型技术在更多垂直行业的结合与应用，以期在市场中获得先机。大模型的深入应用不仅关乎技术的进步，产业生态的构建，还关系到各类行业业务模式的创新。

生成式AI在帮助企业降本增效、改善体验和产品创新方面的能力逐渐得到落地印证

生成式AI为企业带来的改变已经发生，尤其在降本增效、改善体验和产品创新方面，其潜力和效果正在逐渐得到印证。

在降本增效方面，生成式AI能够自动化和智能化处理大量数据和信息，企业能够降低运营成本，提高决策效率。例如大模型技术在物流行业的智能航线规划、货运装载优化、船舶调度等方面的应用，可以提高物流效率，降低运营风险。

在改善用户体验方面，生成式AI带来全新的人机交互体验，能够提供更有温度、更有深度、智能化的服务。蚂蚁集团推动AI技术落地“三个管家”，覆盖金融理财、生活服务、医疗健康这三个大众焦点场景。其中，AI金融管家“蚂小财”，专注理财和保险专业知识问答，提供行情解读、持仓分析、保险配置和投教科普等个性化服务，让每个投资者都有一个私人金融助理。截至2024年6月末，月活用户已经超5908万人，用户日均使用559万次，当月平均每位用户使用约2.8次。

在推动产品创新方面，借助生成式AI可为用户提供更加个性化和高效的产品服务。如大模型技术对WPS产品创新的推动，WPS AI锚定 AIGC 内容创作、Copilot 智慧助理、Insight 知识洞察三个战略发展方向，得益于 WPS AI 的发布，WPS为用户提供了更加丝滑、高效的办公体验。

大模型的发展或需更长的新技术周期，但各参与方对其商业化仍持乐观态度

对于大模型技术发展开启的新一轮技术周期，商业化前景被普遍看好。

一方面，2023年以来我国大模型数量井喷，各类技术厂商纷涌入局。自2023年3月OpenAI发布GPT-4以来，国内迅速响应，大型科技企业、初创企业、高校和研究机构等纷纷加入大模型研发，“百模大战”由此开启。截至2024年3月，中国10亿参数规模以上的大模型数量已超100个，形成上百种应用模式。

另一方面，大模型的类型也在不断丰富。从最初的语言模型如GPT系列，到后来的多模态模型，如DALL-E和CLIP，再到专门为特定行业定制的垂直领域模型，大模型技术呈现出多样化的发展趋势。这些模型在参数规模、训练数据、应用场景等方面各具特色，满足不同行业和领域的个性化需求。例如，一些模型专注于医疗健康领域，能够辅助医生进行疾病诊断和研究；另一些专注于金融行业，提供风险评估和市场分析等。

资本市场对大模型企业的投资热情高涨，展现出对大模型技术未来前景和商业潜力的乐观态度。尽管2023年投融资整体行业遇冷，但AI行业融资的形势相对仍处于比较热门的状态，如阿里巴巴、腾讯、蚂蚁等科技公司参与大模型领域创业公司投资，反映出市场对大模型前景的看好。

进入2024年，大模型厂商之间开始出现明显的价格竞争与市场份额抢占，竞争加剧，体现对大模型未来商业价值的看好。各家大模型厂商的竞争点已不再局限于技术，而是扩展到了价格以及具体的落地场景。大型科技企业开始加大开源力度，以扩大自身大模型生态影响力，扶持更多AI原生应用创新。

总的来看，虽然大模型技术和商业模式还有待进一步的发展与完善，但是从行业层面来看，各参与方对大模型的持续发展和未来的商业化应用前景呈现出明显的乐观态度。IDC预测，到2027年，中国的AI技术用户对AI相关的软件、

硬件、服务总投入将达到约407亿元，2023-2027年的CAGR达到20%，约占2027年整体IT支出（软件、硬件、服务总支出）的11%。

金融机构有望用大模型绘制未来金融蓝图

金融业在我国经济中举足轻重，金融机构通过提供资金流动和管理服务，为个人、企业和政府的各种经济活动提供必要的资金支持。金融行业的数智化建设对于提升金融服务效率、促进金融资源的有效配置有重大作用。大模型技术的引入，被视为推动金融行业智能化进程的关键力量。

- » 一方面，金融行业产生和依赖大量数据，其决策过程高度依赖于数据的获取和高效利用，对科技与业务的深度融合、跨领域知识的整合与分析等综合能力具备高要求。
- » 另一方面，不同于一般行业，我国金融行业具有强监管的特殊性。金融风险涉及面广且具备传染性，需要快速识别并响应，充分防范系统性金融风险。因此金融行业的数智化建设需要满足强监管的要求，不仅要实现业务经营管理智能化，还要达成关键技术自主可控、防范金融风险等目标。

大模型技术的引入，有望助力金融机构突破瓶颈，在多个层面推进智能化的建设进程。通过大模型提供深度学习和自然语言处理的强大能力，可持续推动金融基础设施的重塑、金融科技创新、服务效率提升、安全与风险防范增强，并助力金融机构实现可持续发展，重塑金融服务的面貌。例如人工智能在信贷风险管理领域的运用可提升银行的资产转化能力，突破传统风险管理措施的局限。

大模型在金融行业的应用前景广阔，围绕AI价值实现、AI大模型在金融行业应用生态演化以及确保AI的负责任应用，构成了金融行业智能化的重要研究课题，也是实现AI在金融行业价值实现的关键。

AI大模型

在金融领域应用及发展框架解析

AI技术在金融机构中的广泛和深入应用对提升整体金融行业的运行效率、服务范围和质量具有重要意义。蚂蚁集团研究院认为，生成式AI金融行业价值目标的实现应从负责任AI（RAI）、AI价值、AI生态三个角度来实现。

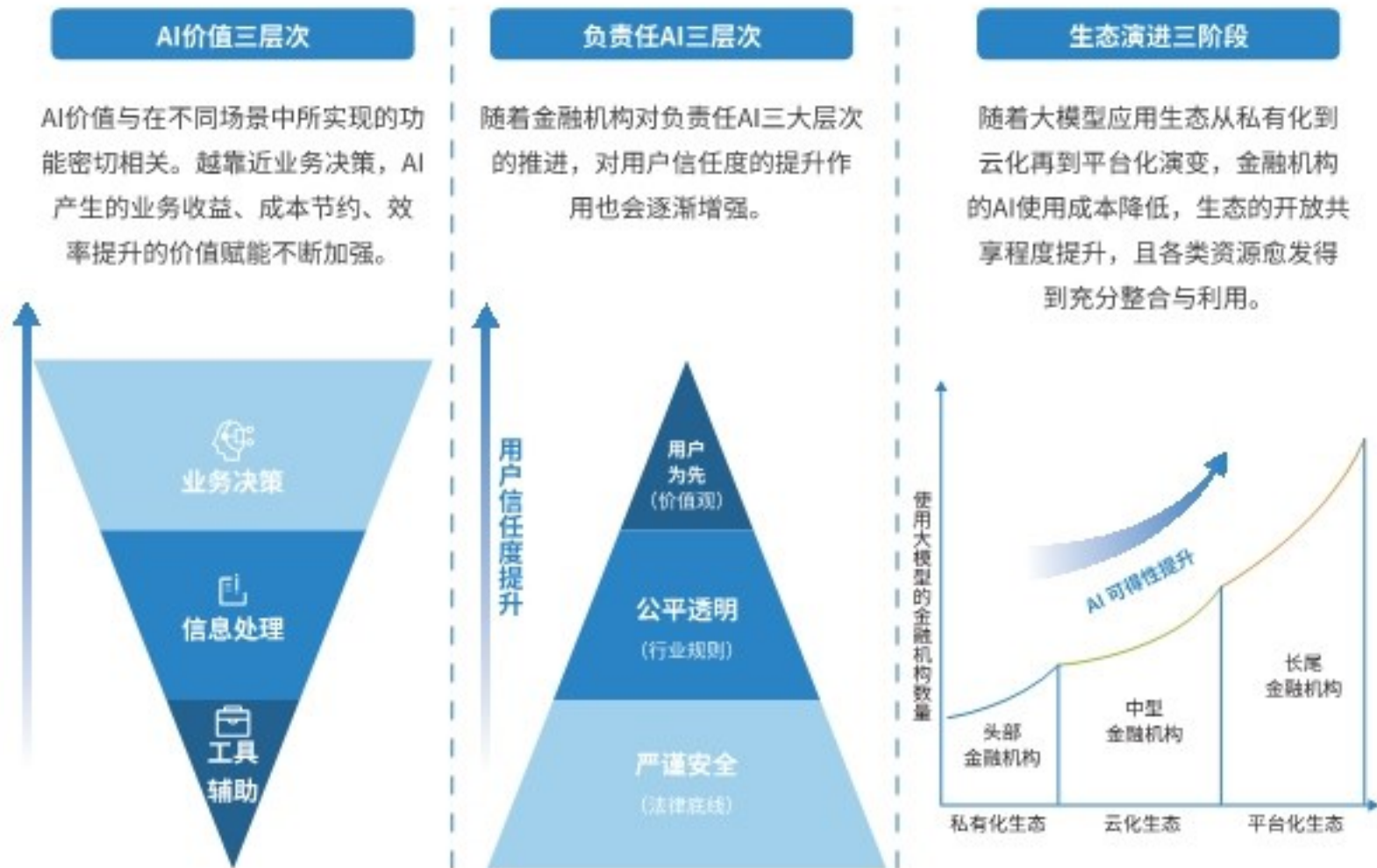
以负责任AI作为基石和底线，遵守监管要求与商业道德，防范AI在金融行业中应用的风险，维护金融行业稳定性，是实现AI在金融行业价值释放的前提，让每个金融机构都能在AI价值发挥方面实现进阶。在此基础上，通过大模型在金融领域的应用中各类参与方的协同，形成逐渐完善健康的大模型在金融领域的应用生态，为不同的资金实力、技术基础的金融机构提供适合的AI服务，助力AI在金融行业的普惠性应用，发挥AI对金融机构业务的分场景、分程度的赋能，实现AI应用价值的指数化提升。

生成式AI在未来金融行业中的价值实现由负责任AI、AI价值、AI生态三大基石支撑
生成式AI在金融场景中的价值实现，在负责任AI治理的保障下，通过AI生态实现指数级的价值放大

$$\text{生成式AI金融行业应用价值} = \text{负责任AI} \times \text{AI价值} \times \text{AI生态}$$

来源：蚂蚁集团研究院绘制

当前大模型的技术成熟度有待持续提升、与金融业务场景的结合方面，无论是深度还是广度都还有待进一步扩展、相关制度与法规还有待持续完善，AI在金融行业价值目标的实现需要分阶段逐步推进和实现。本次白皮书提出大模型在金融领域应用及发展框架如下：



来源：蚂蚁集团研究院绘制

- » **AI价值三层次：**AI价值随着工具辅助、信息处理、业务决策功能进阶，越靠近决策功能，任务引领层次越高，所产生的业务收益、成本节约、效率提升的价值赋能也不断加强。
- » **负责任AI三层次：**RAI分为严谨安全、公平透明、用户为先三层次，随着负责任AI逐层推进，对用户信任度的提升作用逐渐增强。
- » **生态演进三阶段：**生态演进分为私有化生态、云化生态和平台化生态三阶段，随着应用生态衍化，大模型易用性和金融机构的可得性提升，生态开放共享程度提高。

01 AI价值实现



1.1 金融行业AI价值评估方法

目前金融机构通过使用生成式AI，在单点的业务场景上降低成本、提高人效等价值已经逐渐得到验证，但是要系统性的评估和反映AI对金融行业的价值提升，需要建立在更加全面、深入地将AI嵌入金融行业场景和服务环节的基础上。

AI价值的体现与不同的场景所对应的业务特性强弱有关。不同的场景所具备的金融业务属性强弱有差异，通常而言，业务属性越强的场景，AI的成本收益改善价值的潜力越大。因此，AI对金融机构的价值体现在业务运营、风险管理、用户服务和决策制定等多个方面，而具体的价值大小与作用根据应用场景的不同而不同。例如在用户服务层面，智能客服和聊天机器人可以提供7*24服务，不再受限于工作时间和人员能力，快速响应用户需求，提升服务体验；在风险管理场景中，AI可以实时监控市场动态，预测多种风险因素，提供更准确的风险评估和预警，从而突破基于历史数据和有限的变量带来的对风险因素捕捉不全面的限制。在此二类场景中，风险管理作为金融机构业务开展的核心，具备更强的金融业务属性，AI所能赋予的成本收益改善的价值潜力更大。

除了应用场景对AI总价值的影响之外，AI价值还与AI应用的深度有关，不同的应用深度对应着不同的AI增强系数。即使在同一场景中，AI应用的深度不同，所发挥的价值大小也不同。在应用场景不变的情况下，AI总价值随着AI应用深度的加深而增加。因此衡量大模型的使用为金融机构带来的价值提升，需要考虑与原有业务规模对比，AI应用的不同场景、不同的业务环节的特性以及与AI应用深度相关的增强系数，才能更加科学的衡量AI所带来的总价值变化。我们采用如下公式表达：



1.2 AI价值层次和场景价值

1.2.1 AI价值实现的三大层次

AI价值的发挥与AI在不同场景应用中所实现的功能属性密切相关。在不同场景中，AI的应用基本围绕着工具辅助、信息处理、业务决策来发挥其价值。根据AI的智能化程度以及与不同场景的应用深度，AI实现着从任务执行工具到信息处理再到业务决策的不同功能。在此过程中，越靠近决策功能，任务引领层次越高，AI的实现难度加大，越需要具备金融领域的专业知识和专家思维能力，所产生的业务收益、成本节约、效率提升的价值赋能也不断加强。



来源：蚂蚁集团研究院&IDC绘制

工具辅助功能进行基础任务的自动化执行

AI在金融机构较为初级的功能通常表现为作为工具用于自动化的任务执行。典型应用为在辅助研发的场景中执行代码辅助生成与检查等任务、在办公助手中执行文案翻译、纪要生成、文案辅助等任务，此阶段不涉及强业务属性的复杂分析类任务，通常按照预设的规则进行指定任务的执行，可以辅助金融机构提升效率，降低成本。例如在自动化报告生成的任务中，AI可以自动收集和分析数据，生成日常、周度、月度的财务或业务报告，可节省人工整理数据和编写报告的时间。

信息处理功能发挥金融业务助理作用，进行复杂金融领域信息的识别与分析

AI实现信息处理功能时，通常扮演着信息助手和业务助理的角色，适用于满足处理和分析大量复杂金融领域相关数据的功能需求，如市场研究、用户数据和行为数据分析等。在此类功能实现中，AI需要辅助金融机构处理信息密集型任务，更有效地处理、分析和整理大量非结构化数据，如文本、语音、图像等，快速识别模式和趋势，提取关键信息，提供深入的业务洞察。这一层级的应用有部分金融机构正在探索中，对技术和人才等层面的投入要求相对较高。

业务决策功能发挥金融专家的作用，通过辅助决策实现业务闭环

在决策功能层面，AI需要具备金融专家的认知与思维，在前两层功能的基础上，通过辅助决策完成业务闭环。在这一功能中，AI不仅需要综合多源信息和数据进行深度分析，还需要运用专家知识和经验进行复杂的推理和判断，从而发挥决策功能。在金融机构中，通过对AI的深度应用，AI参与业务决策，运用先进的算法和模型，提供专业的金融见解和建议，可以帮助金融机构在竞争激烈的市场中做出更科学敏捷的决策，还能减轻对专家技术的依赖和降低人才流动带来的风险。当前对金融机构而言这一层面的AI功能还欠缺成熟性，需要进一步发展。



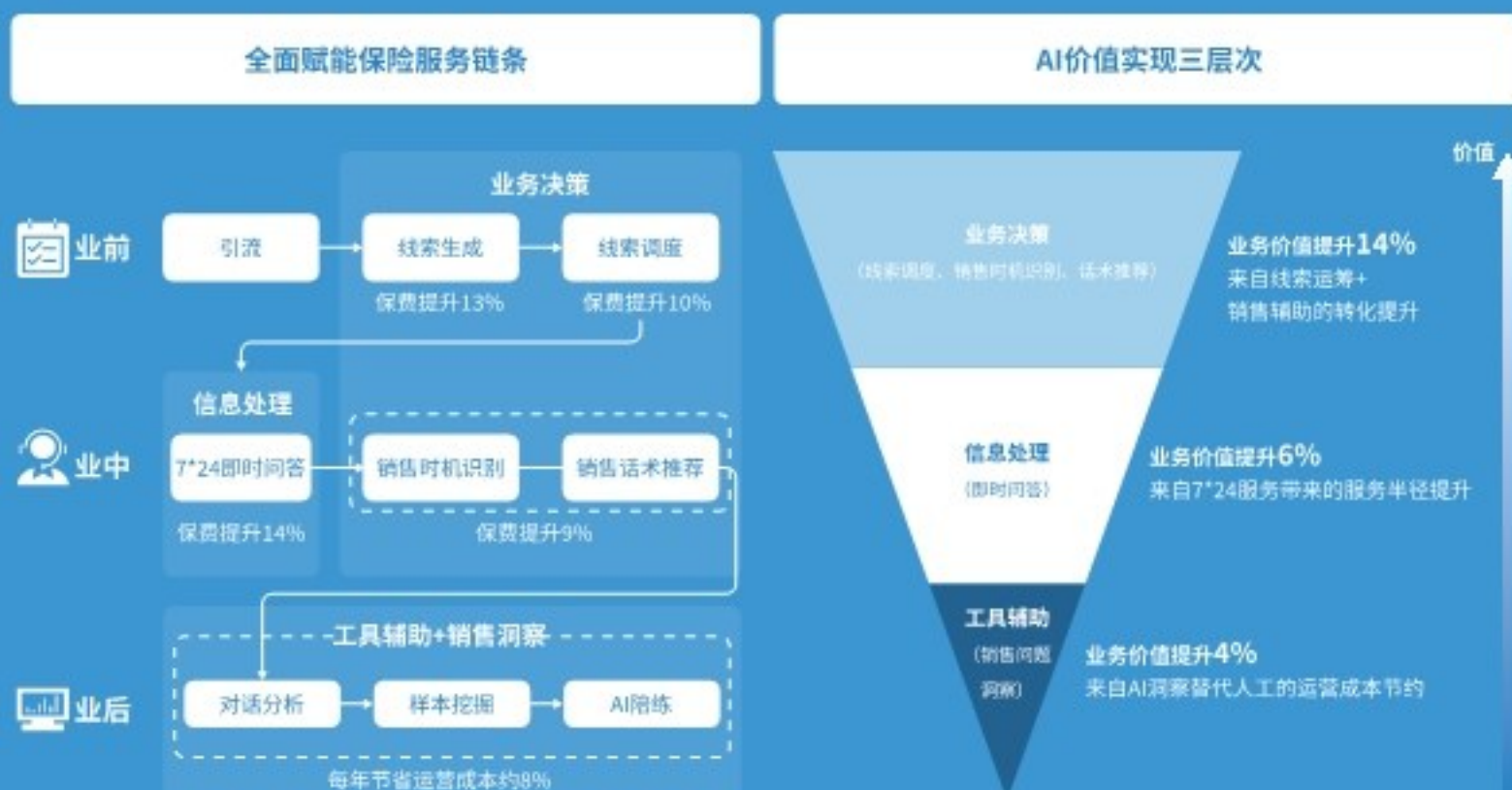
案例：保险规划师销售支小助- 赋能保险服务全流程

蚂蚁集团的保险规划师销售支小助的应用将AI赋能在整个保险业务链条中，通过对规划师全流程的展业辅助体现了在销售洞察、信息处理、业务决策的多方面价值。蚂蚁集团对保险业务全流程进行AB测试，证实支小助在不同环节对保费规模或运营成本均有改善作用。相比业务人员独立展业，支小助分环节赋能的辅助工作模式所带来的保费规模或运营成本降低都有不同程度的体现：

- » 在业前阶段，支小助能在线索生成和线索调度发挥辅助业务决策的作用。其中在线索生成环节使用AI之后使保费规模提升了13%，在线索调度环节通过AI赋能提升了10%的保费规模。
- » 在业中阶段，支小助在7*24即时问答环节能做到30秒内及时响应，起到信息处理的价值和作用，并使保费规模得到14%的提升；在销售时机识别及销售话术推荐环节，发挥辅助业务决策的作用和价值，使保费规模得到9%的提升。
- » 在业后阶段，支小助作为工具辅助在销售洞察、AI陪练环节中使用，每年节省运营成本约8%。



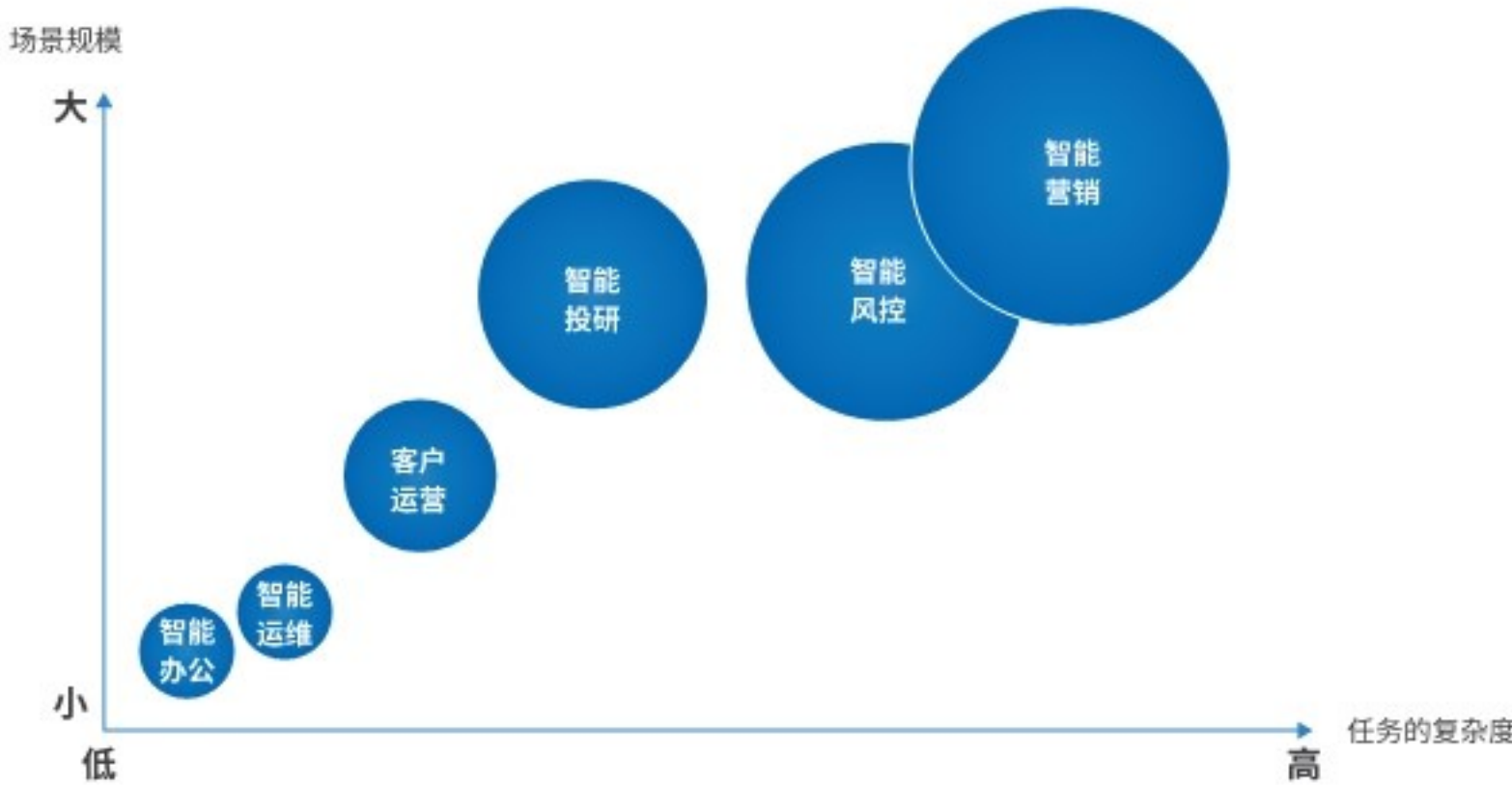
销售小助-规划师



来源：蚂蚁集团，2024

1.2.2 AI场景价值进阶

在应用程度足够深、各层面功能充分实现的前提下，AI对金融机构的价值与不同的应用场景相关。AI在金融机构中应用的场景具备业务属性的强弱之分，随着应用场景从业务边缘到核心，AI对金融机构的价值逐渐增大。同时金融业务具备较高的复杂性和不确定性，决策难度越高的场景，AI决策功能的价值体现越明显和关键，对金融机构的价值赋能也越大。

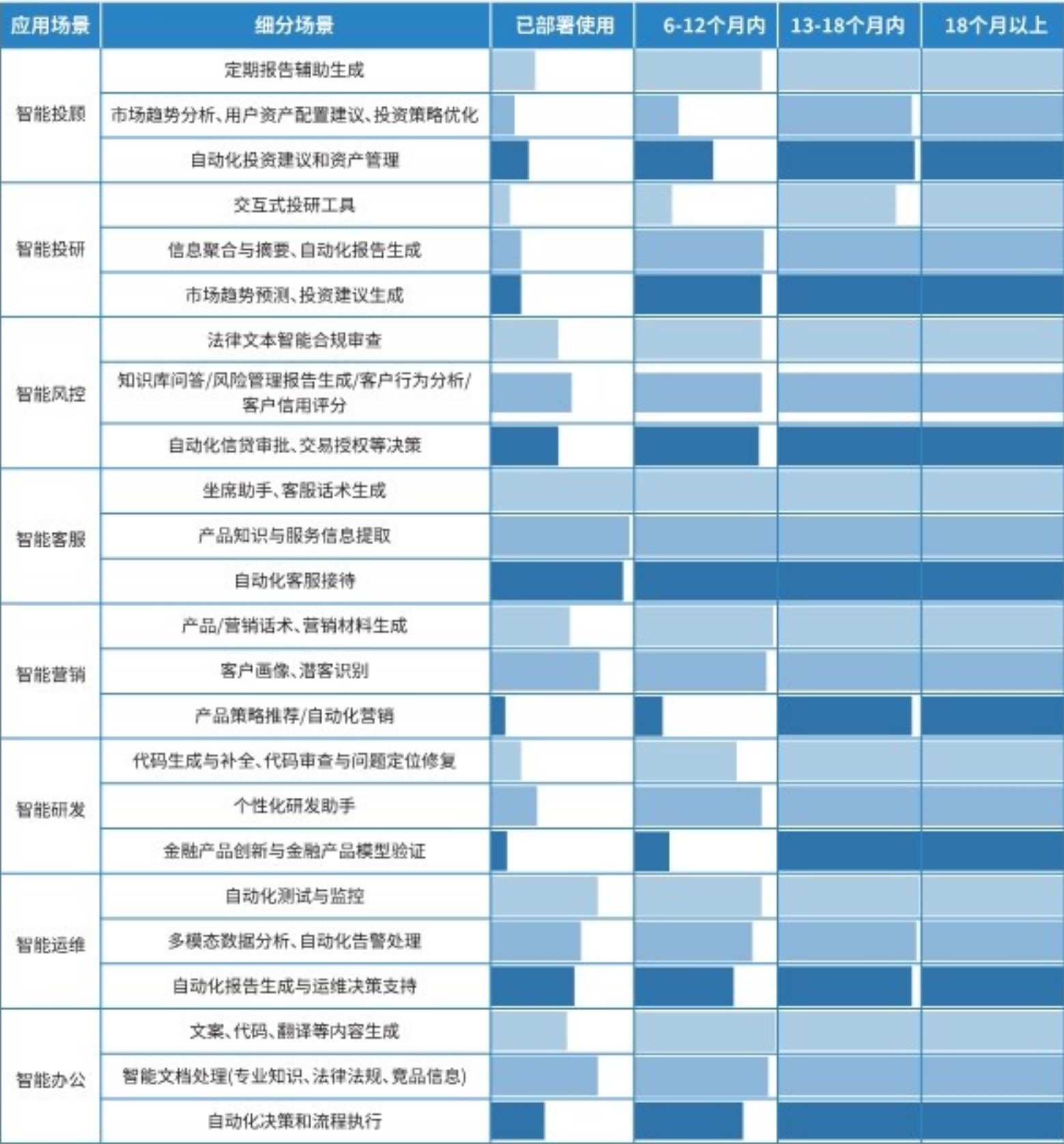


(注：气泡大小代表各场景在充分实现AI各功能的前提下，对金融机构的价值赋能空间)

来源：蚂蚁集团研究院&IDC，2024年基于调研和专家深访研究绘制

当前金融机构的AI价值实现任重道远。根据IDC对金融机构在不同场景AI应用情况的调研，当前已经应用AI技术的金融机构大多在通用运营类场景和边缘业务类场景进行探索与应用，对于核心业务类场景的应用相对较少。同时针对每个场景的AI功能实现层面，即从任务执行属性的工具辅助功能到信息处理功能再到业务决策功能的不同层面，当前金融机构除了在智能客服场景中AI应用较为深入且功能实现较为成熟之外，在其他场景的应用还有望持续加深。尤其在核心业务

场景中的应用主要集中在较为浅层的工具功能层面，需要更长时间去实现更复杂的决策功能，因此对应的AI价值潜力还需要更长时间的深入应用才能充分发挥。



(注：阴影长度代表本调研中金融机构在相应时间阶段内的对应AI应用落地的累计样本比例)

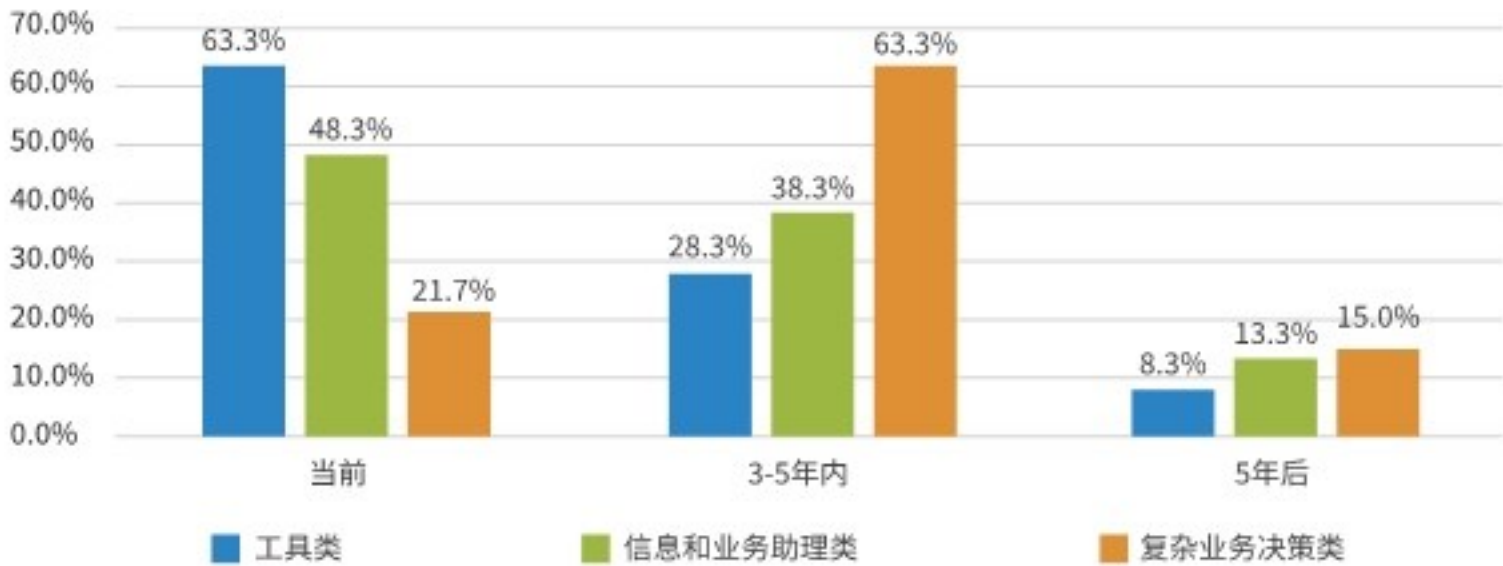
来源：IDC中国2024，基于调研和专家深访研究绘制

1.3 金融机构AI应用现状

金融机构视角下的大模型对行业变革的预期、现状与投入计划

根据IDC对金融机构的调研，63.3%的受访机构普遍认为当前阶段最具潜力的价值作用集中在工具功能层面，用于任务的自动化执行；63.3%的受访机构认为以金融专家知识与思维参与决策的功能预计在3-5年内可发挥价值潜力。

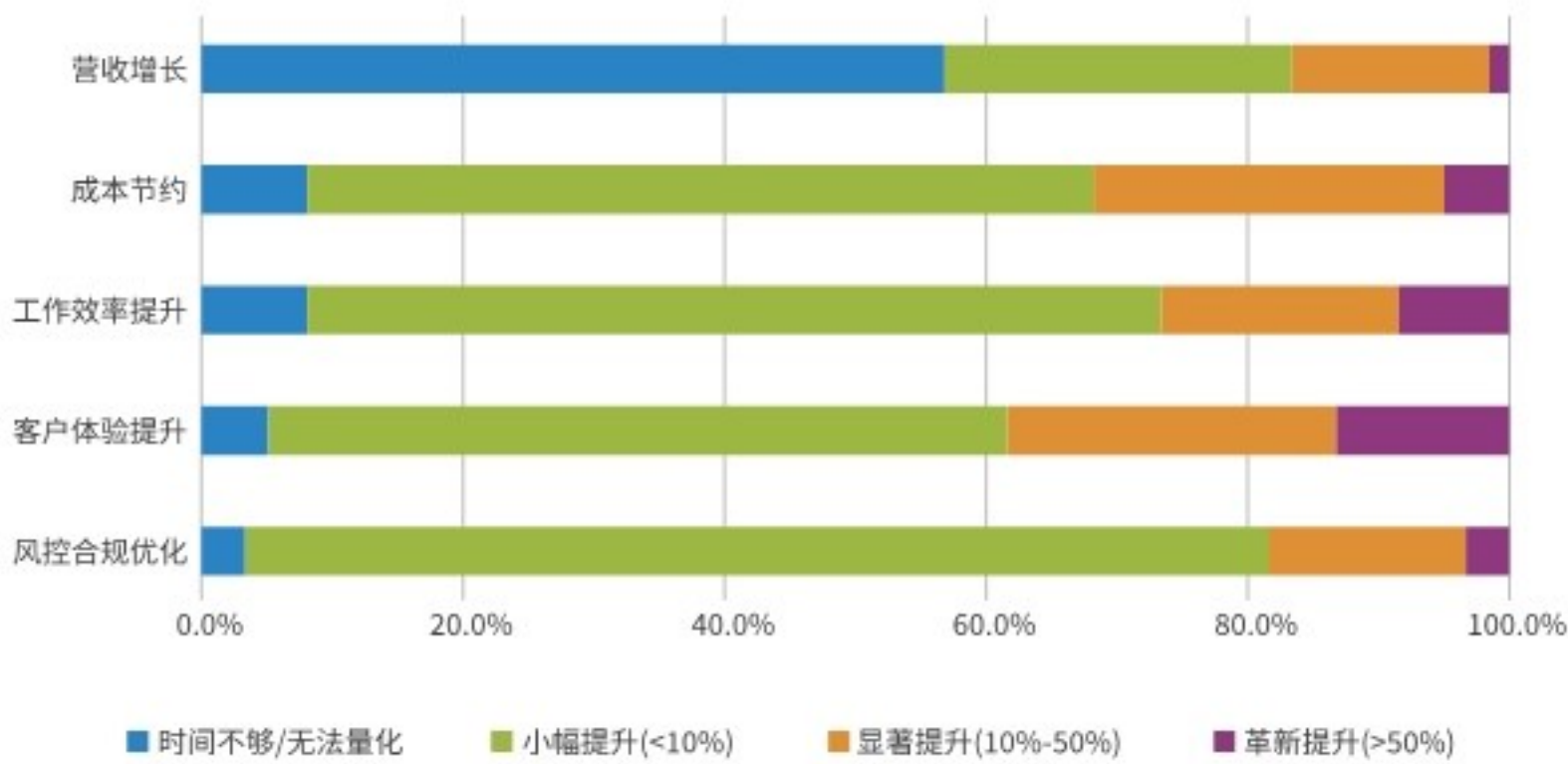
从不同的AI功能层面看，由于工具功能较为初级和浅层，是当前多数金融机构在AI应用中相对容易实现的功能，对应的潜力也在当前释放较为明显；而信息处理功能、业务决策功能难度递增，当前金融机构普遍还未有充分应用，因此对应的价值潜力在当前体现不明显。随着各方面技术逐渐成熟以及金融机构对AI的应用程度加深，金融机构预计在未来3-5年内AI的决策功能将发挥更加充分的价值潜力。因此，在当前至未来5年内，绝大部分的AI工具功能、信息功能和决策功能对应的价值潜力逐渐达到实现的高峰，而5年后预计金融机构将持续优化AI在复杂决策中的应用，从而得以进一步发挥其价值潜力。



来源：IDC中国2024，金融行业调研，N=60

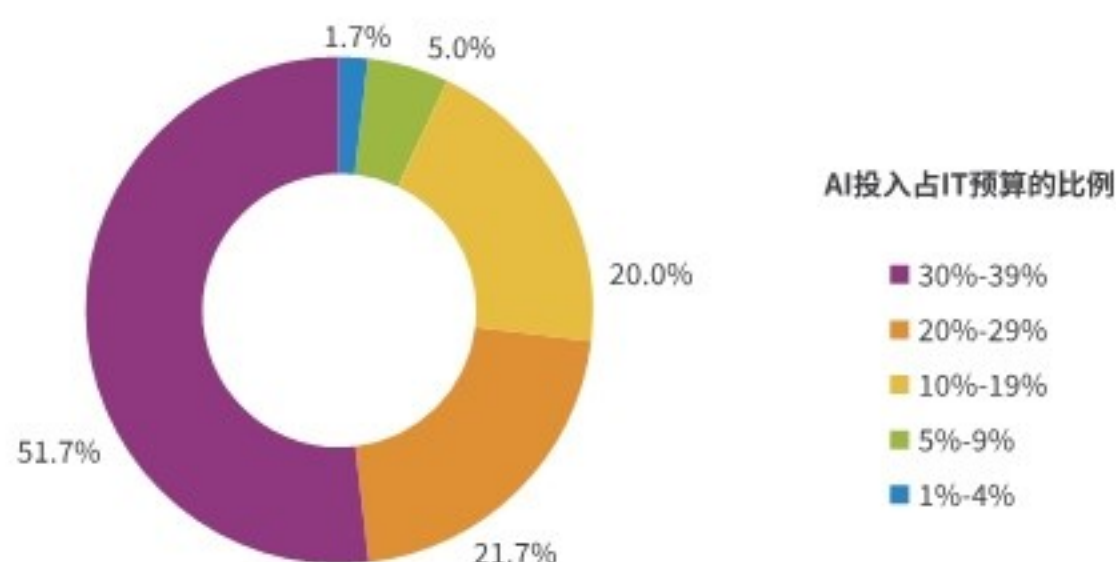
* 工具类(文案、代码、翻译等任务执行类功能)
信息和业务助理类(智能化知识抽取、金融知识的理解和生成、政策研报分析等专业领域信息处理)
复杂业务决策类(信贷审批决策、理财投顾建议等复杂信息分析与决策)

在当前金融机构的AI投入产出比方面，根据IDC的调研，目前除了营收增长方面的成果尚未看到明显的量化收益之外，在成本节约、效率提升、用户体验、风控合规方面都取得了小幅提升，其他层面的收益预计在未来逐步实现。出于对大模型在金融领域商业化进程的判断，金融机构预计营收增长的成果需要2-3年后实现。



来源：IDC中国2024，金融行业调研，N=60

从预算投入趋势看，金融机构未来对大模型相关的投入较为积极。调研显示，过半样本（51.7%）的受访机构表示在未来18个月内用于生成式AI项目的IT预算将占总预算的10%-19%内。相比于前文IDC最新的IT Spending数据测算中整体AI技术用户的大模型/生成式AI预算占IT总预算的11%，金融行业用户的投入呈现较为积极的态度。



来源：IDC中国2024，金融行业调研，N=60

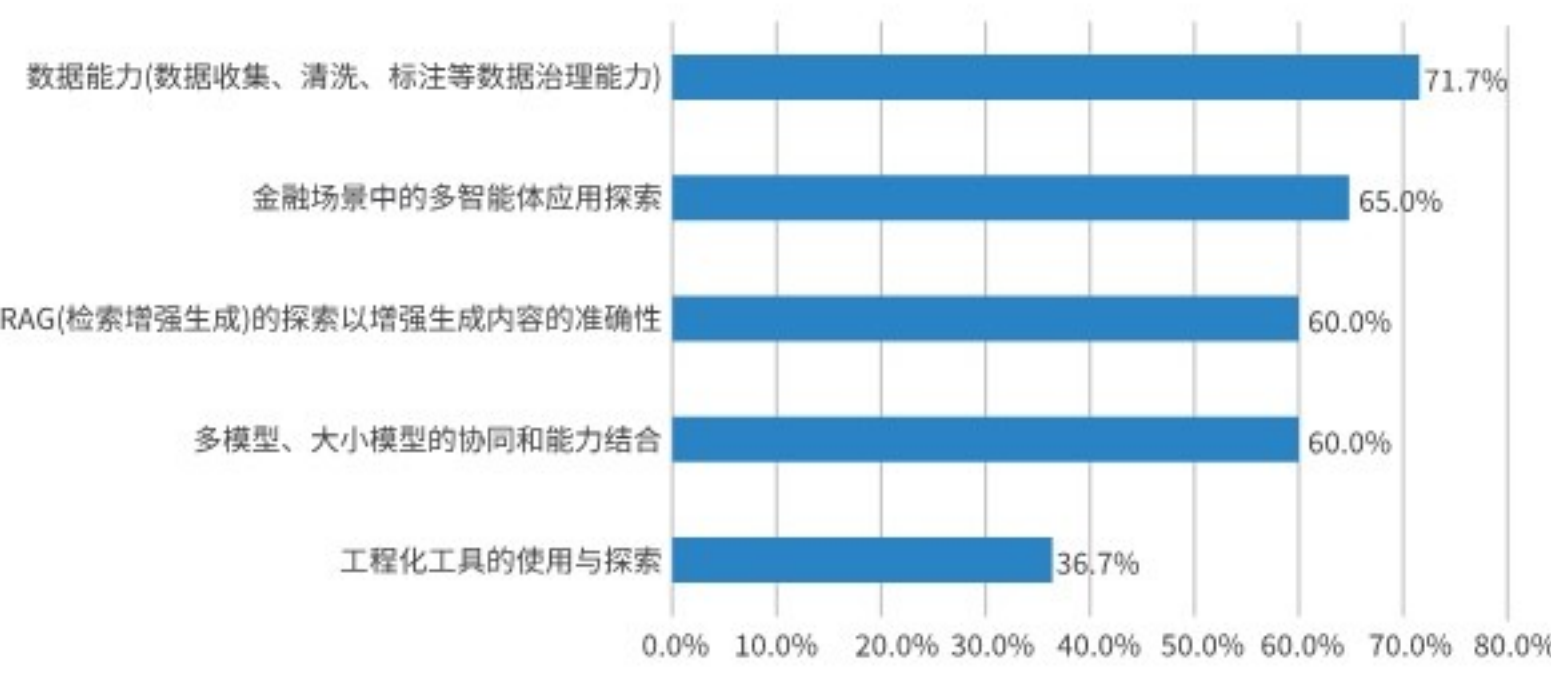
生成式AI的应用正在从内部运营走向业务场景，金融业务属性逐渐增强

AI大模型向金融垂直领域发展的最终目的是服务业务经营和管理场景。随着金融领域垂直大模型技术的深入发展，未来大模型在金融领域的应用将更多地触及金融核心业务，减少对人的依赖程度，实现真正意义的“AI+”的效果。

目前，金融机构主要在内部运营中应用生成式AI，如自动化报告生成、研发助手（代码辅助生成等）、会议纪要生成与文案翻译等，以提高员工效率。部分的金融机构也正在开始将其应用于边缘业务场景，例如用户服务中的智能客服机器人、社交媒体监控等。当前金融机构在核心业务场景中的应用还处于早期探索阶段，关键核心业务的使用程度有限。未来在大模型技术成熟和风险管理能力进一步提升后，金融机构有望使AI在信贷评估、投资策略制定、资产管理等核心业务场景中应用得更加充分。

1.4 四大能力帮助金融机构实现AI价值进阶

随着金融机构对AI的使用程度从浅到深、从边缘到核心的不同层次的递进，AI发挥的价值也随着使用应用场景的拓展与应用程度的深入而逐渐提升。金融机构需要建设多个层面的能力支撑AI应用的广度与深度拓展，提升AI价值层级。本白皮书研究认为，数据基础能力、RAG和AI工程化工具、大小模型融合能力、多智能体协同应用的探索能力构成了实现AI价值进阶所需具备的能力图谱。这也和本次调研结果，金融机构的认知基本一致。



来源：IDC中国2024，金融行业调研，N=60



来源：蚂蚁集团研究院&IDC绘制

（一）夯实大模型基石-打造坚实的数据基础能力

金融机构的数据基础能力是实现AI价值的基石。数据是AI应用的原材料，数据质量和处理方式直接影响AI模型的性能和输出结果。没有准确、安全、可靠的数据，AI模型和系统就无法发挥其潜力，甚至可能产生误导性的结果。大模型在金融领域的应用需要具备金融行业的领域知识与业务逻辑理解能力，因此在指令微调、增量预训练等多个阶段，都需要加入大量高质量的金融数据，以提升大模型的金融理解能力。

数据基础能力		建设目标	建设方法
	数据收集	能够从内部和外部来源收集各种类型的数据，包括交易数据、用户互动数据、市场数据等	建立数据收集渠道，如自有数据、合作伙伴和公共数据源
	数据清洗	清洗数据以提高数据质量，包括处理缺失值、异常值和重复记录	开发自动化的数据清洗流程，使用数据清洗工具，建立数据质量标准
	数据标注	为非结构化数据（如文本、图像）提供标签，以便于机器学习模型的训练	建立标注团队，开发或采用标注工具，制定标注指南和标准
	数据整合	整合来自不同源和格式的数据，形成统一的数据视图	使用ETL（提取、转换、加载）工具，设计数据仓库和数据湖架构
	数据治理	制定数据管理政策和流程，确保数据的合规性、安全性和可访问性	建立数据治理框架，包括数据管理角色、策略、流程和指标

来源：蚂蚁集团研究院&IDC绘制

（二）打通模型到应用落地-应用RAG检索增强和AI工程化工具

» RAG增强生成内容的准确性

RAG可降低传统大型模型在特定领域知识更新和精确度方面的局限性。RAG可通过检索组件实时访问最新的数据和信息，实现知识的动态更新，确保模型能利用最新的领域知识进行回答；还可检索领域特定的信息，提高模型在金融领域的专业性和精确度。通过RAG检索到的专业可靠的信息补充到大模型中，增强对金融场景特定领域情境的理解，提高回答的相关性和准确性。

» AI工程化工具的使用与探索

金融机构在大模型能力建设，工程化工具的使用与探索是实现AI能力高效、稳定落地的关键。AI工程化的作用不仅体现在AI开发过程中，也体现在大模型的规模化落地应用过程中。为真正实现大模型的场景落地，金融机构需要构建完整的、模型落地的工程化工具链，覆盖从模型微调、数据工程、应用开发、安全工程等一系列工具，从而达到便捷地调用大模型的效果。



来源：蚂蚁集团研究院绘制

对于金融机构而言，借助工程化工具能更好的达到4个层面的作用和效果，以实现系统化、可扩展和可维护的AI应用，充分发挥AI价值。首先，金融机构需要工程化工具来支持大模型的开发和训练流程，包括数据处理、特征工程、模型调参等。其次，需要利用工程化工具对预训练的大模型进行精调，产生优质的场景应用以适应特定的金融业务场景。再次，工程化工具支持模型的版本控制、模型评估，确保模型的可维护性和可追溯性，实现服务度量和全面治理；最后，金融机构可借助数据循环、社区等，实现引领创新与持续优化，利用开源社区的力量，与其他机构和开发者协作，共同推动大模型技术的发展和落地应用。综上，AI工程化与大模型汇流，对于金融机构而言可共同推动AI应用加速，实现更广泛、更优质的场景落地，充分发挥AI价值。

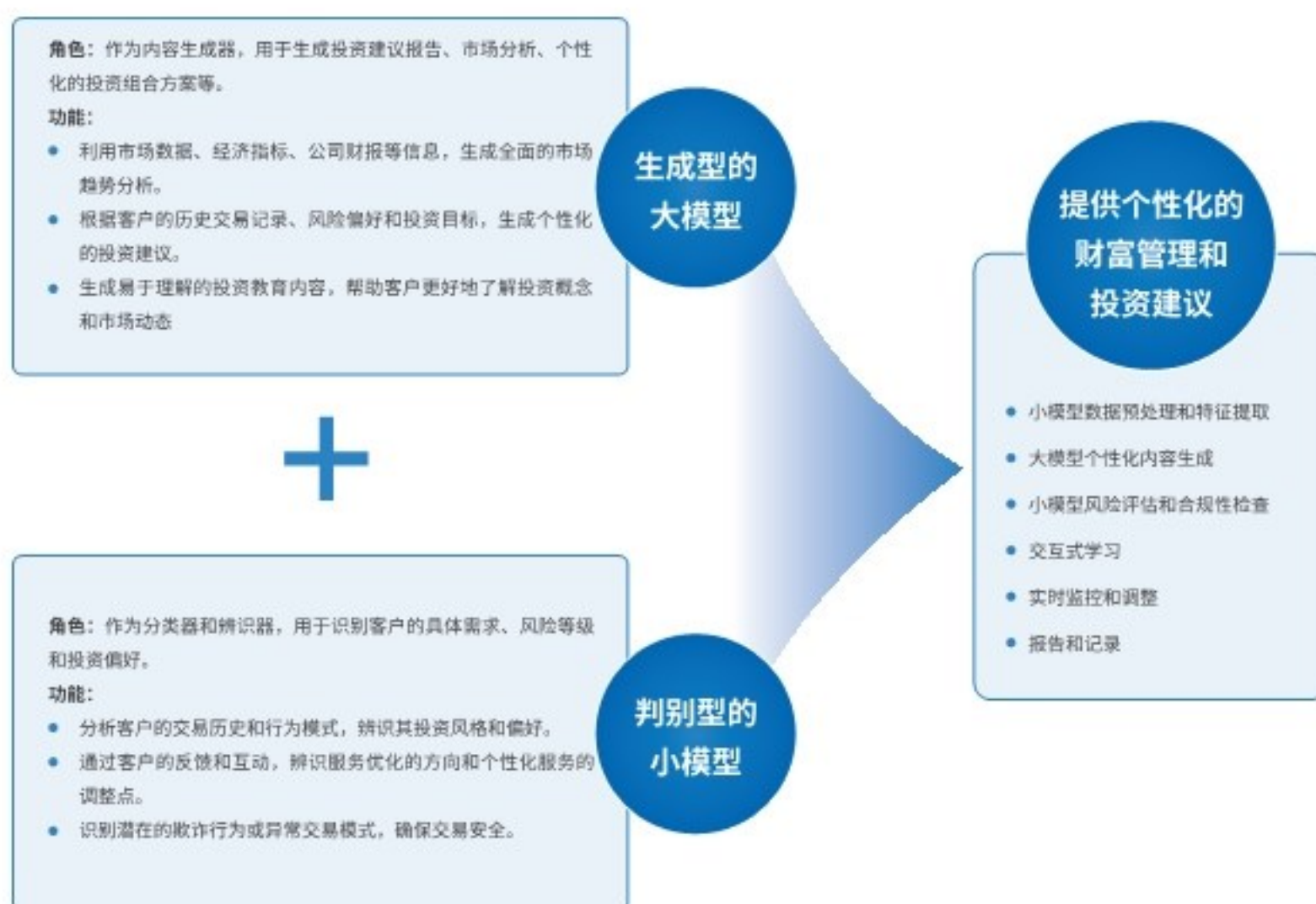
（三）提升金融场景适配能力-加强大小模型融合

大模型融合能力建设是金融机构提升AI应用的关键，涉及根据自身情况选择模型建设路径、针对不同业务场景进行模型能力匹配、大小模型的协同与多模型的集成、构建丰富的场景模型工具以提高AI应用的灵活性与质量等全方位的能力建设。

大小模型的协同和能力结合

在金融场景中，生成内容型的大模型与判别型的小模型相结合可共同实现模型的生成和辨识能力，从而获得AI价值的提升。大模型在感知、认知和交互领域具有显著优势，而各个专业小模型专注于特定问题或任务，在特定领域（运筹、量化、图算法等）的表现更为优异。因此，在实际应用中，涉及严谨的金融决策场景，应优先采用大小模型协同和能力结合，而非单独依赖大模型。可以将大模型作为认知和交互的中枢，而小模型则负责专业判断和决策，通过大模型调度小模型共同完成复杂的推理和决策工作。

大小模型的协同，旨在通过整合不同类型的模型来提高整体性能和适应性，结合大模型的深度学习能力和小模型的灵活性与专业化，通过不同模型的协同与结合，金融机构可以减少对单一模型的依赖，也更利于金融机构实现成本效益的平衡。在此过程中，金融机构需要进行合理的模型选择，根据不同的使用场景，结合不同模型的特定优势，选择不同规模的模型，结合多个模型的优势，提高整体的预测准确性和鲁棒性。



来源：蚂蚁集团研究院&IDC绘制

如上图所示，以智能投顾服务场景为例，生成型的大模型作为内容生成器，用于生成投资建议报告、市场分析、个性化的投资组合方案等。判别型的小模型作为分类器和辨识器，用于识别用户的具体需求、风险等级和投资偏好。利用判别型小模型从原始数据中提取特征，如用户偏好、交易模式等，将提取的特征输入生成型大模型，生成个性化的投资建议和报告。使用判别型小模型对生成的内容进行风险评估和合规性检查，确保建议的合理性和安全性。通过用户的反馈，判别型小模型可以进一步优化特征提取的准确性，同时生成型大模型根据反馈调整内容生成。判别型小模型实时监控市场变化和交易行为，生成型大模型根据监控结果动态调整投资建议。生成型大模型生成的报告和记录可以用于内部分析和外部审计，判别型小模型确保记录的准确性和完整性。在此案例中，通过大小模型的结合，金融机构能够提供高度个性化和专业的智能投顾服务，同时确保服务的安全性和合规性。

（四）增强复杂场景决策支持能力-探索多智能体协同的运用

随着AI在金融机构的应用逐渐进入核心业务场景，决策属性增强，需要金融机构探索智能体的应用能力。AI Agent（智能体）可以模拟专家决策过程，提供基于复杂数据分析的洞察，帮助金融机构在风险管理、投资策略等方面做出更精准的决策。

引入AI Agent系统，可辅助金融机构处理单靠提示词难以实现的复杂业务场景，进而达到高效的自动化决策。用户可以使用AI Agent输出自己想要的、内容可控的文本，且AI Agent具有自主学习和决策的能力，能够根据实时市场状况和投资者需求，动态调整大模型和小模型的参数和行为。例如在金融交易自动化场景中，AI Agent可以分析市场趋势，自动执行交易策略，在确保投资决策合理性的同时，提高交易的速度和效率。

从智能体到多智能体，是从更高的层面上处理具备多个环节的复杂专业任务时需要探索的能力。专业任务往往是多环节多分支的，各环节和分支的专业化分工会有更高效的ROI，因此产生了从智能体发展到多智能体的必要性。

金融核心业务场景具备信息密集、知识密集、决策密集的特殊性，需同时考虑严谨性与专业性，因此需要大模型和多智能体各司其职。大模型负责压入必要的知识和能力，多智能体装载过程性专家经验来保障严谨和专业。

在多智能体的应用中，由于涉及到多条分支环节的配合，因此需要根据不同的任务匹配不同的协同模式。例如蚂蚁集团开源多智能体框架agentUniverse，作为行业首个开源的金融领域多智能体技术框架，该框架核心提供了多智能体协作模式“PEER”（Plan-Execute-Express-Review），有效提升大模型应用于复杂金融场景的推理分析能力。同时，agentUniverse允许开发者对多智能体协作模式进行开发定制，可帮助开发者加快大模型技术在金融场景的落地研发。

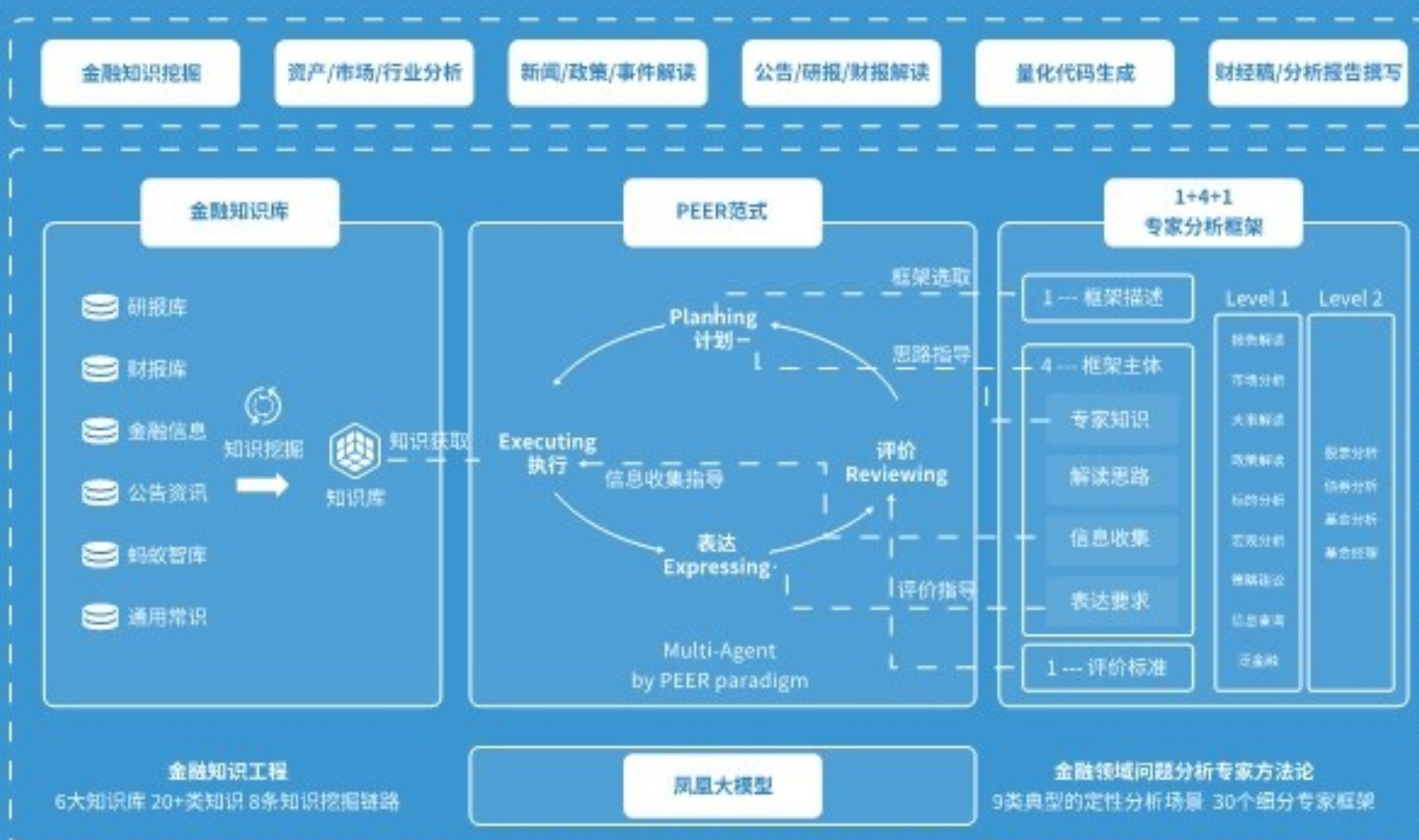


案例：投研支小助——基于PEER模式的新一代金融专家智能助手

以蚂蚁基于PEER模式的新一代金融专家智能助手投研支小助为例，投研支小助的底层是基于agentUniverse“PEER”框架融入大量投研专家经验，构建了一个投研Copilot。PEER模式通过计划(Planning)、执行(Executing)、表达(Expressing)、评价(Reviewing)四个不同职责的智能体，实现对复杂问题的多步拆解、分步执行，并基于评价反馈进行自主迭代，最终提升推理分析类任务表现。

按照任务需求，智能体与金融知识库和专家分析框架交互。金融知识库覆盖研报库、财报库、金融信息、公告资讯、蚂蚁智库和通用知识六类知识库；专家分析框架是蚂蚁沉淀的金融领域问题分析方法论，针对9类典型的定性分析场景，总结出30个细分专家框架，体现了专家知识的引入，不同机构可以通过定制这些专家框架让投研支小助呈现出差异化的解读思路。

投研支小助已全面覆盖海内外主流市场的上市公司、基金产品，以及百余家核心独角兽公司，分析能力涉及800+行业板块，每日可主动分析上千条新闻资讯及热点事件，实现对金融市场动态的分钟级响应和多角度分析，大幅提升了数百名金融分析师、理财顾问的工作效率。目前，投研支小助在蚂蚁内部辅助每日热点事件解读、市场分析和周期性报告撰写，整体效率提升超过50%，采纳率达70%以上。



来源：蚂蚁集团，2024

02 负责任AI



2.1 什么是负责任AI

本白皮书将负责任的人工智能（RAI）定义为以恪守严谨安全的监管要求、遵循公平透明的行业规则、维护用户为先的价值取向的方式进行大模型和生成式AI的设计、开发和部署。RAI专注于以符合用户期望、组织价值观和社会法律规范的方式开发和使用人工智能解决方案，它还包含了正确的计划、监督和治理。RAI的设计旨在确保企业中的AI使用以人为中心、符合科技伦理、安全可控。

RAI要求金融机构及其生态系统以有效、可扩展和可适应的方式负责任地使用AI。RAI不仅关乎技术的成功实施，还关系到金融机构的监管合规，以及对用户、对行业和对社会的责任。通过负责任地开发和部署AI，金融机构可以确保技术进步与商业目标和社会价值相协调。RAI有助于金融机构遵守监管合规要求，减少违规风险和潜在的法律诉讼，并辅助实现维护金融系统稳定的社会价值。

蚂蚁和IDC研究认为，实施RAI将帮助金融机构提升用户信任度。用户信任是金融机构成功的关键。若金融机构确保用户的隐私安全、做到公平可信并遵守商业道德、以用户价值为中心，则用户会更容易信任该机构从而成为长期忠诚用户，进而提升金融机构的市场竞争力，推动金融行业高质量发展。

2.2 负责任AI的三大层次

根据RAI的定义和设计宗旨，负责任AI涵盖了技术手段、监管法规、道德伦理、利益取舍等多个方面。结合RAI提倡的价值观，我们认为RAI分为“严谨安全、公平透明、用户为先”三大递进的层次。随着负责任AI三大层次的推进，对用户信任度的提升作用也会逐渐增强。



来源：蚂蚁集团研究院绘制

RAI的最低要求：严谨安全

严谨安全是RAI的最低要求和基础层次，金融机构在AI应用中做到严谨安全，是AI相关应用结果符合监管要求、降低法律风险和维护业务稳定的基石，也是追求用户满意度的基础。具体来说，严谨安全包含了隐私保护、数据安全和内容严谨三大层面。

隐私保护，要求严谨地对待每一条用户信息

隐私保护是负责任AI的主要原则之一，主要指个人信息的保护，隐私保护风险通常与金融机构的数据治理、监督与问责机制、隐私保护技术应用不到位密切相关。金融行业数据涉及大量用户个人隐私数据、资金交易数据等，在使用大模型时，确保符合隐私法规、做好隐私保护至关重要。隐私保护原则能确保用户的个人信息是在征得同意的情况下收集和处理的，不过度采集、且不落入未经用户授权的主体中。使用AI技术中如果侵犯了用户隐私，易导致金融机构面临法律纠纷和监管风险。

数据安全，保证每一个环节的数据合规应用

生成式AI在金融领域应用需要大量的数据来进行训练和优化，如果数据泄露或被恶意利用，不仅会损害用户的利益，也会对金融机构的声誉和业务造成严重影响，甚至面临法律责任。

数据安全涉及到金融机构的数据与模型治理水平是否到位、与数据相关的流程是否规范与完备、数据安全相关的技术手段是否充足等因素。除了大模型开发中的数据训练过程中存在数据安全隐患之外，在大模型使用中也可能存在潜在的数据安全风险，因此数据安全的要求体现在与大模型相关的数据采集、传输、存储、加工及处理信息的全流程环节。若流程、技术和治理措施层面存在明显欠缺，则无法确保数据安全，易导致数据泄露、滥用。

数据安全表现在金融机构从数据的存储和处理等多个环节被安全和适当地使用。伴随着全面的数据安全政策，数据保护的明确目标、原则和操作流程，从基座大模型到业务适配的全过程中相关数据都经过严密的清洗、审查；设立严格的访问控制机制，确保只有授权人员能够访问敏感数据；记录和监控所有数据访问和操作行为，以便于事后审计和责任追踪。

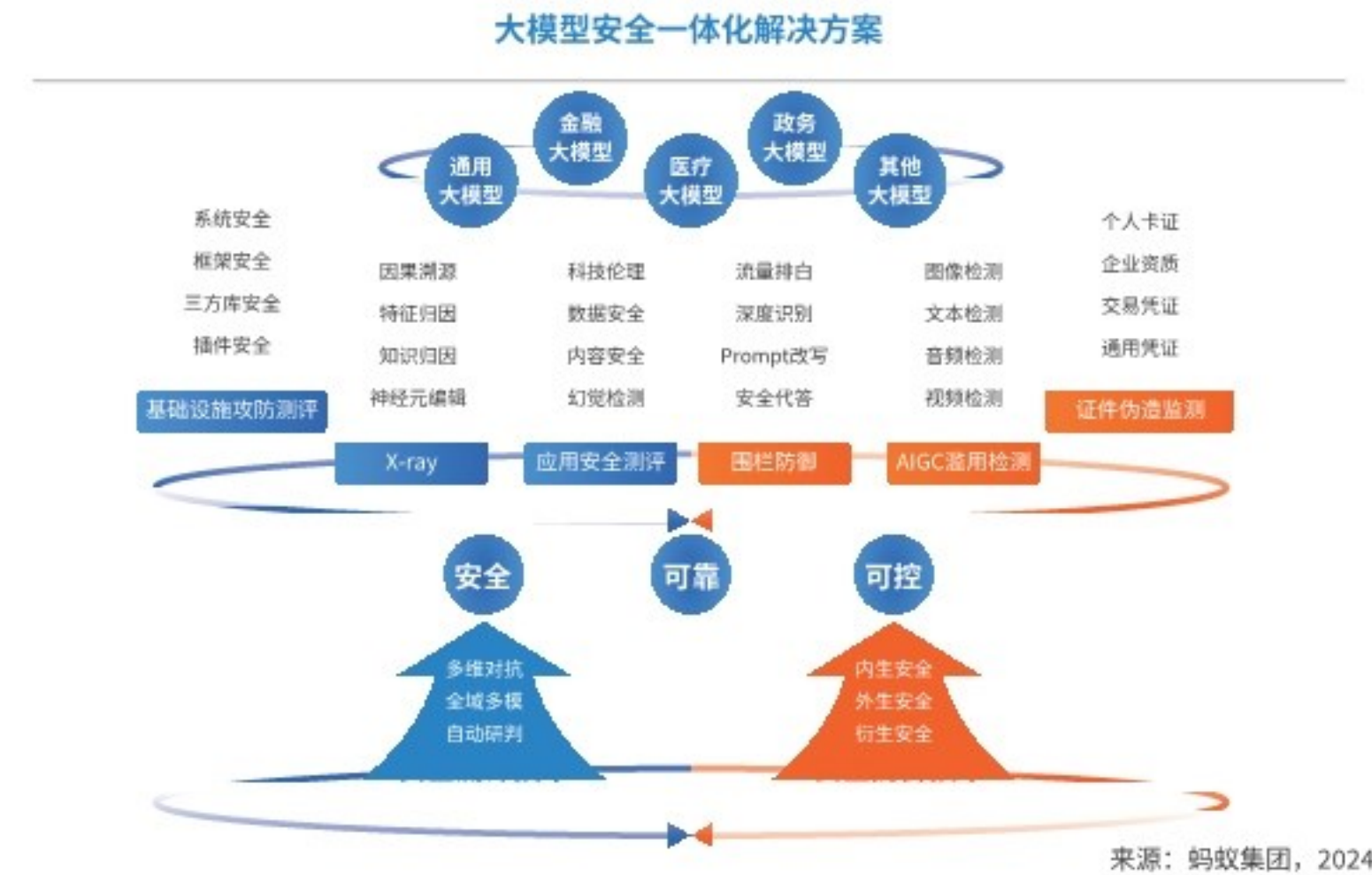
内容严谨，降低“幻觉”结果

大模型幻觉问题是金融行业应用大模型的一大挑战。金融行业容错率非常低，无论是知识问答，还是内容提取，都对大模型提出了非常高的要求。模型的准确性成为金融行业采用生成式AI的重要堵点。因为模型结构、解码算法、暴露偏差等原因，大模型可能出现幻觉现象，即AI模型生成的内容虽然流畅但与事实不符，或者与提供的源内容不一致，生成看似合理却无法应用于实际情境的非事实性内容。

降低模型的幻觉问题，需要提高模型的透明度，对人工智能的机制和过程进行检查，以便更好地理解模型的决策过程和输出的依据。通过技术对抗，用AI识别与检测AI，实现以技治技的效果，是降低模型幻觉、提高大模型产出内容严谨性的有效措施。

以蚂蚁的“蚁天鉴”为例，2023年，蚂蚁集团发布“蚁天鉴”大模型安全一体化解决方案，其中“蚁鉴2.0” (大模型安全检测平台)是全球首发多类型工业级可信AI检测平台，通过诱导式对抗生成技术，持续不断攻击和训练模型，在模型上线前对其进行全方位的安全扫描，提前识别和挖掘风险漏洞;“天鉴” (大模型风险防御平台)基于智能风控技术，通过围栏防御、极速防御以及情景式防御等多层护栏方案，可以实现内容安全、数据隐私、科技伦理、合规风险等四大类风险的识别和拦截，不仅能帮助大模型挡住恶意提问，同时对生成的回答内容进行风险过滤，从而保障大模型上线后从用户输入到生成输出的整体安全防御。

2024年7月，蚂蚁集团联合清华大学发布“蚁天鉴2.0”版本，新版本增加了AI鉴真功能，它支持图像、视频等多模态内容的真实性及深度伪造检测;同时在测评功能上持续升级，打造大模型安全评测“智能体”，并新增大模型X光、大模型基础设施检测两大评测功能。“蚁天鉴2.0”形成了一套包括大模型基础设施测评、大模型X光测评、应用安全测评、围栏防御、AIGC滥用检测、证件伪造检测等在内的完整技术链条，面向行业提供全方位智能化的大模型安全评测和防御解决方案。



RAI第二要求：公平透明

公平透明是金融机构实现RAI的第二要求和进阶层次。金融系统不仅是经济活动的组成部分，也是社会结构和人类福祉的关键因素。金融机构需要在追求经济效益的同时，考虑其对社会的广泛影响，承担社会责任。公平透明是金融机构遵守遵循行业规则、避免不良社会影响的必要条件。具体来讲，公平透明包含尊重平等、可解释性的具体要求。

尊重平等，充分考虑金融服务的普惠性与共享属性

金融服务的普及和可负担性对社会成员的福祉至关重要。金融机构在应用AI技术进行金融业务开展时，应充分考虑金融服务的普惠性与共享属性，避免特殊化的的人群偏好，聚焦社会民生，实现尊重平等。

在大模型使用过程中，如果算法设计或数据使用不当，可能会导致对某些群体或个体的不公平或不平等对待。例如在信贷场景中，信用评分模型被广泛用于评估用户的信用风险，从而决定贷款的批准与否以及利率的高低。因而在大模型训练的过程中应保证数据集的多样性，避免数据层面的偏差；模型训练前后使用偏见检测工具来识别潜在的偏见，定期进行偏见影响评估，监控模型的输出；在算法层面引入公平性约束来减少输出偏见，优化模型和算法，建立必要的人工的审核和干预机制，建立完整的反馈机制，尽可能地降低模型造成的偏见，充分考虑金融服务的普惠性和共享属性。

可解释性，打开AI决策过程的黑盒

可解释性通常指专家和使用者能理解人工智能的决策。大模型技术通常因为模型复杂性、黑盒算法、特征交互、数据噪声等多方面原因存在可解释性不足的问题。通常解释大模型的难点包括模型的高复杂性、训练数据的多元性、输出的不确定性以及评估指标的不足等。

对于增强模型的可解释性，一方面可采用多种可解释性机制提升模型的可信度，针对大模型在金融领域的部署与使用中的透明度，采用能提升数据质量和可解释性的机制，从而帮助用户与监管机构理解模型的决策依据。为提高模型的透明度，可以考虑在流程能力建设，通过将确保人类的监督、利用可解释的人工智能（XAI）技术，嵌入专家验证流程，使大模型的生成和决策过程达到透明、易理解的目的。另一方面，大模型的相关技术研究者也应持续探索新的技术和方法，如基于因果启发的模型解释框架（CIMI），以提供更忠诚和可泛化的解释。此外，还可以采用多种应用方法和框架，如局部可解释的模型不可知解释（LIME）和沙普利附加解释（SHAP）等方法和框架来清晰地解释人工智能行为。

RAI最高要求：用户为先

蚂蚁集团认为，用户为先是RAI的最高层级要求。为进一步提高金融机构的用户信任度，需要秉持以用户为先的价值观，融合进AI应用的各个场景中。在金融机构利益最大化与用户利益最大化相冲突的情况时，AI的算法设计是否能充分体现用户为先的价值观是构成负责任AI的一大重要体现。用户为先体现在以用户需求为中心，提升金融服务个性化、易得性；以及以用户长期价值为导向，促进金融健康。金融机构通过秉持对用户、对行业、对社会负责任的态度，推动AI向善，更好的服务金融行业高质量发展和社会福祉。

以用户需求为中心，提升金融服务个性化、易得性

在金融机构的诸多业务场景中，需要以用户为中心来进行AI应用相关的产品界面、操作流程和算法设计，以用户需求和体验为中心的典型体现如下：

- » **AI+金融服务实现个性化和定制化。**以个性化财富管理顾问的场景为例，金融机构通过基于AI的个性化财富管理顾问系统，为用户提供量身定制的投资建议和资产管理服务，在设计系统时深入了解用户的需求、风险偏好、投资目

标。利用机器学习算法结合市场趋势，系统分析用户的财务状况，在合规前提下优先推荐符合用户利益的方案。

- » **提供可得易懂金融信息。**生成式AI以丰富的内容形式、“拟人化”交互、精准的用户识别，能够直观、多样地展示金融产品和服务内容，避免利用信息差获利。如保险做好风险披露，保持费用和条款透明，清晰地向用户披露保险产品覆盖的风险范围和不覆盖的风险；明确列出所有费用、条款和细则，包括保费、免赔额、赔付条件等，确保没有隐藏费用，确保用户在充分了解情况的基础上做出知情决策。
- » **多渠道便捷的AI服务。**用户可以通过移动应用、网页或用户服务中心、线下网点等多种渠道获得AI顾问服务，确保随时随地的服务可用性与便捷性，提升用户使用服务的便捷程度。

以用户长期价值为导向，促进金融健康

- » **促进用户金融健康。**通过智能投顾具备的金融知识教育和风险提示模块，帮助投资者提高金融素养和采取相应的风险管理措施，不仅能促进理性决策，减少冲动投资和投机行为，维护自身的金融健康，同时也有助于减少投资者的市场恐慌和非理性行为，维护市场秩序，有助于整个金融市场的稳定和健康发展。
- » **用户长期价值优先。**如果同时与多个投资产品有关联，选择从用户利益角度出发而非从金融机构利益最大化出发，向用户客观推荐适合的产品组合与投资建议，立足于用户长期价值优先。通过将负责任的实践嵌入金融机构的AI实践中，RAI有助于在用户和金融机构之间培养出更深层次的信任，长期来看，实现用户价值增值与金融行业行稳致远共赢。

近年来，蚂蚁集团在财富管理、保险等领域助推用户金融健康发展。蚂蚁财富研发了“三笔钱”工具，分别对应灵活取用、未来保障和投资增值三部分。用智能化、个性化的配置工具和方法，基于投资者的个人目标和风险偏好，定制包

含流动性、子女教育、养老、保障、保值增值等在内的全方位规划，来帮助用户开展合理的资产配置。截至2023年底，超过2000万用户开启了“三笔钱”服务，最大回撤平均降低5%。蚂蚁保推出智能保险配置工具“省心配”，该工具可根据用户的年龄、预算、保险需求等因素进行分析，智能模型通过风险分析、保障评估、产品匹配三个步骤进行配置，为用户量身定制合适的保险配置方案。“省心配”工具帮助用户避免买错保险、重复购买同类保险，平均节省30%的预算。

2.3 负责任AI的现状洞察

我国负责任AI相关的政策法规初具全方位的制度框架

我国RAI相关的政策和法规体现在多个层面，其中尤其重视人工智能领域的科技伦理制度建设，以确保人工智能技术的健康发展和负责任地应用，形成了包括国家宏观战略的制定、监管机构的设立、具体管理办法、地方立法，以及行业和企业自律措施在内的全方位制度体系。

当前国内的负责任AI的体系建设以符合监管的底线要求为主导。从各层级的制度来看，当前与负责任AI相关的规范主要集中在顶层部署和具体管理办法层级，而行业组织和企业层面，在监管底线之上的自律性管理内容仍有待完善。我们预计，负责任AI各层级的制度建设还将伴随大模型技术的进步和应用的深入，持续完善和优化。

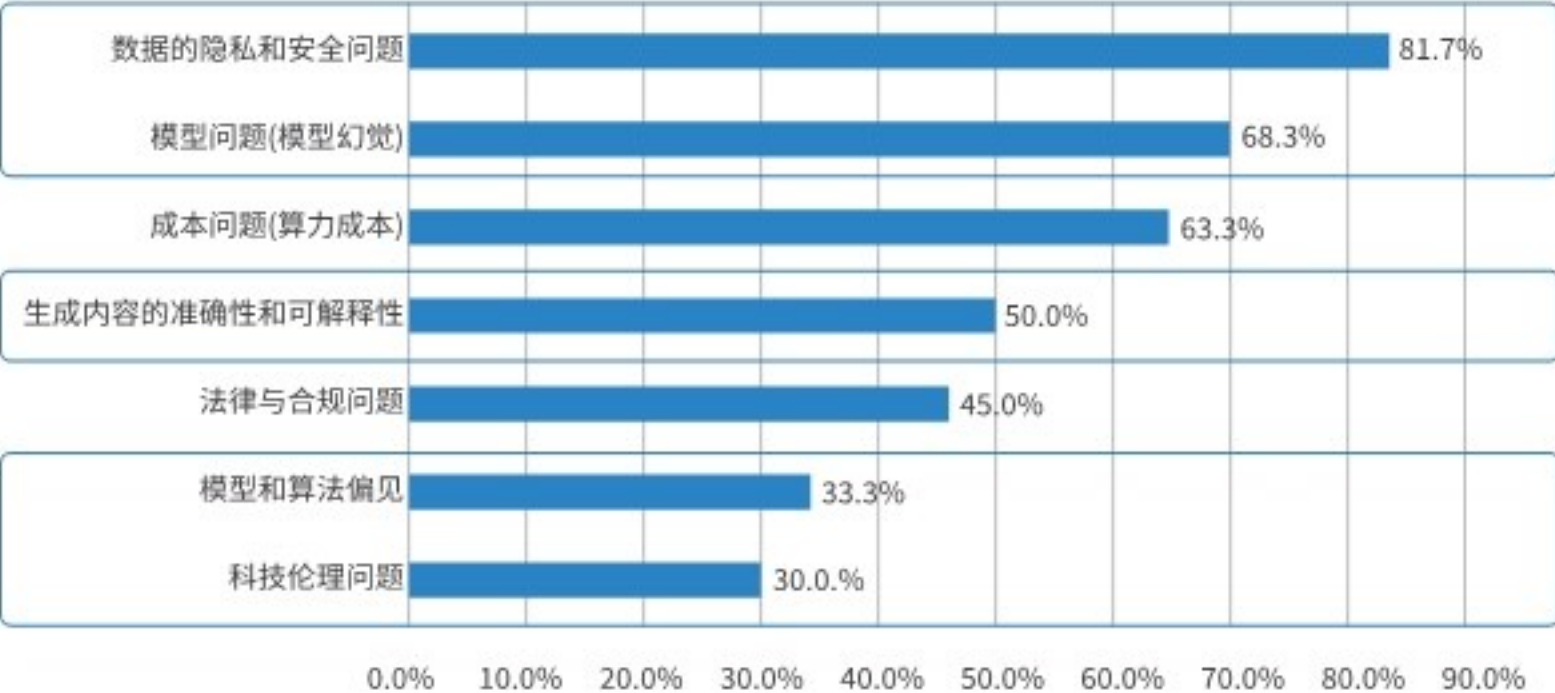


(部分负责任AI相关的文件的梳理详情见文末附录)

来源：蚂蚁集团研究院&IDC绘制

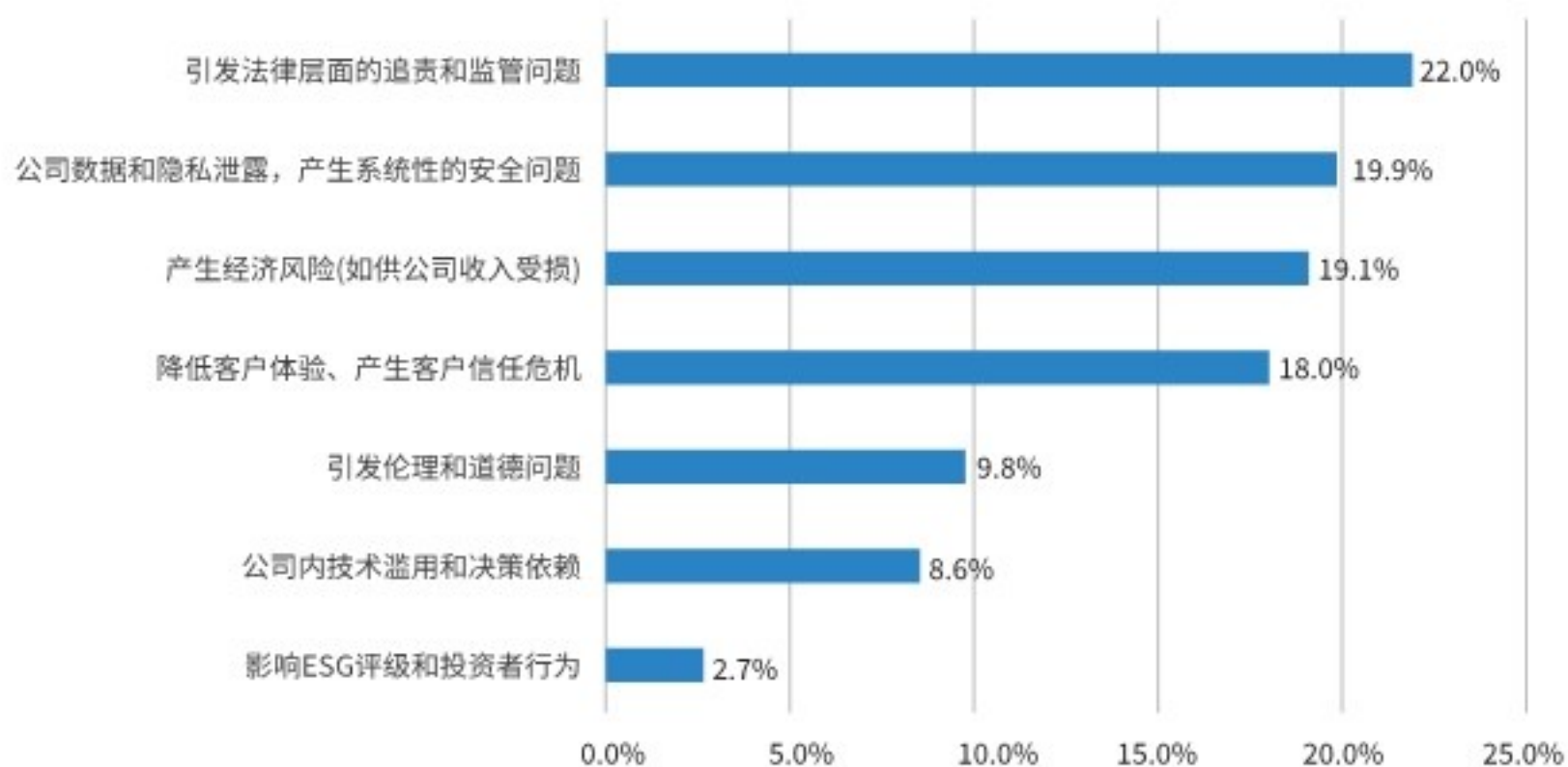
当前金融机构对负责任AI的认知与推行深度基本处于初级层面

目前金融机构对负责任AI的认知和关注点主要集中在严谨安全的初级层次。对大模型在金融行业的落地，金融机构有着多重顾虑。调研显示：金融机构当前对数据隐私和安全性问题、模型幻觉问题关注度最高；对生成内容的准确和可解释性的关注位于中间水平；而有关模型和算法偏见、科技伦理目前的顾虑的关注度最低。从不同的重视程度总结，当前金融机构更多的关注点在严谨安全的层面。



来源：IDC中国2024，金融行业调研，N=60

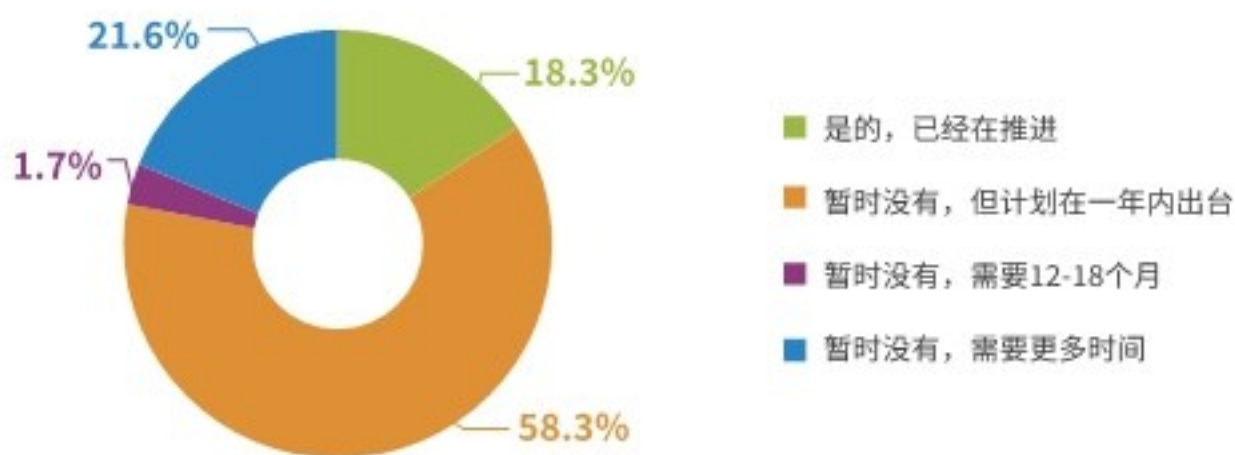
对金融机构而言，通常是出于对潜在的负面影响的担忧而执行负责任AI的举措。从下图IDC的调研结果可知，金融机构对于若不落实负责任AI而最可能产生的潜在负面影响具备不同的担忧程度，反映了金融机构对下列事项重视程度的优先级从高到低依次为：不违法的底线、保护公司的数据和隐私安全、降低经济风险、提升用户体验/用户信任。



来源：IDC中国2024，金融行业调研，N=60

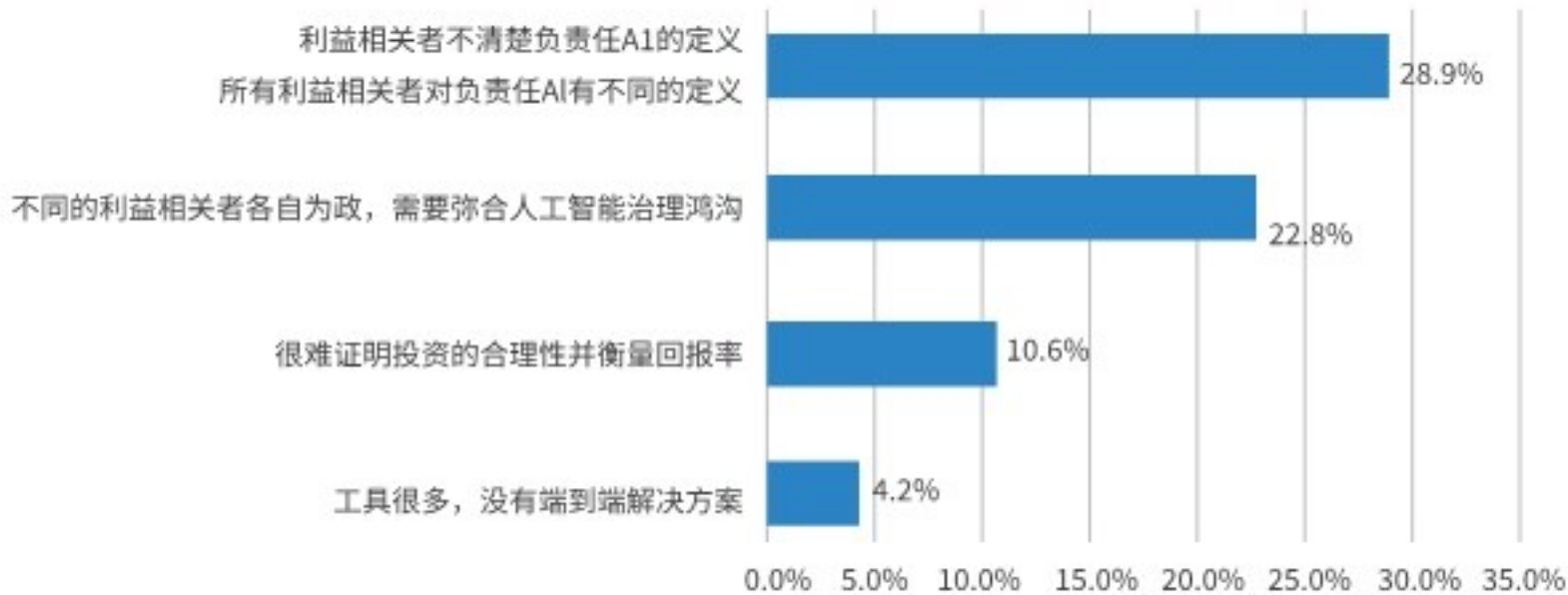
除了关注点层次有待持续加深之外，负责任AI在金融机构中推行的广度也有待持续拓展。根据IDC的调研结果，约18.3%的金融机构在推进落实负责任AI，约58.3%的机构暂时没有推进，但是在一年内有对应的出台计划。其余约23.3%尚无相关计划的金融机构中，绝大部分认为还需要更多的时间才能够作负责任AI的相关规划。

尚无计划的绝大部分是认为需要更多时间



来源：IDC中国2024，金融行业调研，N=60

机构内各相关方缺乏对RAI的明确统一认知且权责不明确是当前推动RAI施行的主要痛点。根据IDC对金融机构的调研，“利益相关者不清楚负责任AI的定义，所有利益相关者对负责任AI有不同的定义”、“不同的利益相关者各自为政”是当前在落实和制定负责任AI方面政策的两大主要痛点。



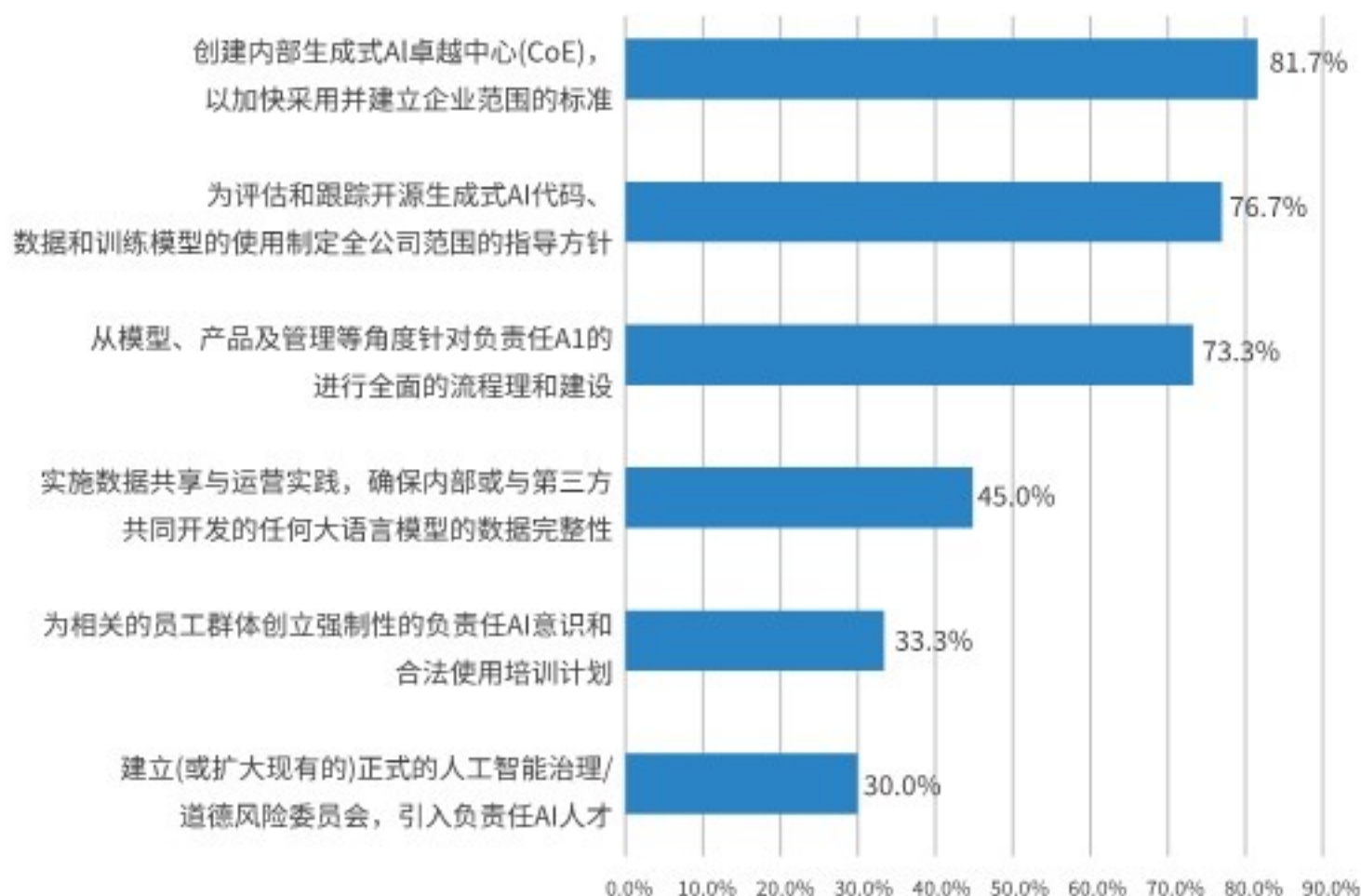
来源：IDC中国2024，金融行业调研，N=60

2.4 负责任AI落地四大关键

蚂蚁和IDC认为，负责任AI需要体系化建设，自上而下需要顶层战略愿景和基本原则引领，组织层面RAI治理结构支持，管控层面具体机制落地，以及技术实践层面持续监测及评估等一套“组合拳”，共同助推负责任AI愿景实现。

» 战略层面：制定机构层面整体的RAI的战略愿景和基本原则

RAI战略愿景以及RAI的实施层次是金融机构能否在大模型时代稳步发展，赢得用户信任的关键。制定战略愿景和计划是确保负责任AI落实的重要步骤，从而在金融机构内明确RAI相关的目标、标准规范与指导方针。根据IDC的调研，大多数金融机构认为要确保负责任AI成功，最重要的举措为：创建内部生成式AI卓越中心（如AI治理委员会），以加快采用并建立企业范围的标准；为评估和跟踪开源生成式AI代码、数据和训练模型的使用制定全公司范围的指导方针；从模型、产品及管理等角度针对负责任AI的进行全面的流程梳理和建设。



来源：IDC中国2024，金融行业调研，N=60

» 组织层面：从组织层面设计RAI治理结构

在金融机构内设置统一的RAI治理结构，明确RAI在不同相关方中的权责和定义，链接不同部门不同职能的相关人员共同贯彻相关标准，是确保RAI能够得以在整个机构内落实的重要组织保障。

金融机构需要针对RAI设计一个有效的治理结构，通过有效的治理结构落实RAI相关的能力建设，确保权责明确与匹配，并协调各个部门促进团队合作，实现RAI的共同目标。这一治理结构的设计，需要金融机构对内系统地梳理RAI的相关方及其责任，此外还需要组织人工智能专家与业务、合规以及内部和外部顾问合作，实施负责任AI的监督。

» 管控层面：对RAI进行全机构、全方位的评估、准备和管控

对于金融机构的CIO等技术领导者来说，实施RAI需要从治理、流程、人才、技术和数据的多个方位对机构进行全面评估、准备和管控，以促进整个机构对人工智能负责任维度的共同理解。

- **治理：**RAI治理是为了确保负责任AI的实践，强调可信度、公平性、隐私性和问责性。治理主要涉及到跨数据和模型相关的权责明确，从而解决RAI需要面对的关键权衡问题，如在算法设计中需要对准确性和可解释性进行平衡，在努力提高透明度的同时，还须考虑安全风险，并减轻模型过拟合和漂移等问题，最终使业务负责人能够做出与组织的风险承受能力相一致的战略性、价值驱动的决定。
- **流程：**RAI鼓励从尽量全面的用户类型和诉求出发来减少偏见，从而设计公平的体系，使AI解决方案能够适应最新的规范和社会需求。在设计RAI流程时，须理解和仔细选择以数据为中心的方法、以模型为中心的方法和产物管理方面的事情，从而提高人工智能决策的透明度和可解释性，促进信任度的提升。
- **人才：**RAI人才方面强调了在构建和管理人工智能体系的专业人员之间对不同技能、科技伦理的理解和不断学习的需求。需要建立由数据科学家、合规人员和其他专家组成的多学科团队，以确保人工智能解决方案是公平、包容和健全的。RAI人才的关键能力包括人工智能和机器学习的深度技术知识、对行为准则和法律标准的理解、有效的沟通技能以及对用户为先的强烈关注。为了实现RAI，组织需要持续投资相关的教育和培训，为人才的专业成长提供途径，从而培养具备专业技能且覆盖多学科的人才库。

■ **技术：**RAI技术指选择和利用多种技术手段以确保人工智能系统是透明的、可解释的、有弹性的和安全的。此外，RAI技术支持将偏差检测和调优工具与数据隐私和治理技术集成，从而满足GDPR和CCPA等法规的公平性和合规性要求。总的来说，技术为RAI创建了基础，以确保AI系统的完整性、性能良好并满足道德标准。

■ **数据：**RAI的数据维度侧重于数据采集、准备和管理，为创建公平、准确和透明的人工智能系统奠定基础。针对数据建设和管理层面，金融机构应强调数据的平衡、质量和完整性，防止偏见，通过培养不同的数据集，以限制偏见和支持公平的人工智能；将确保数据质量和多样性的原则纳入数据生命周期管理中，以减少人工智能应用的偏差和提高应用的准确性。

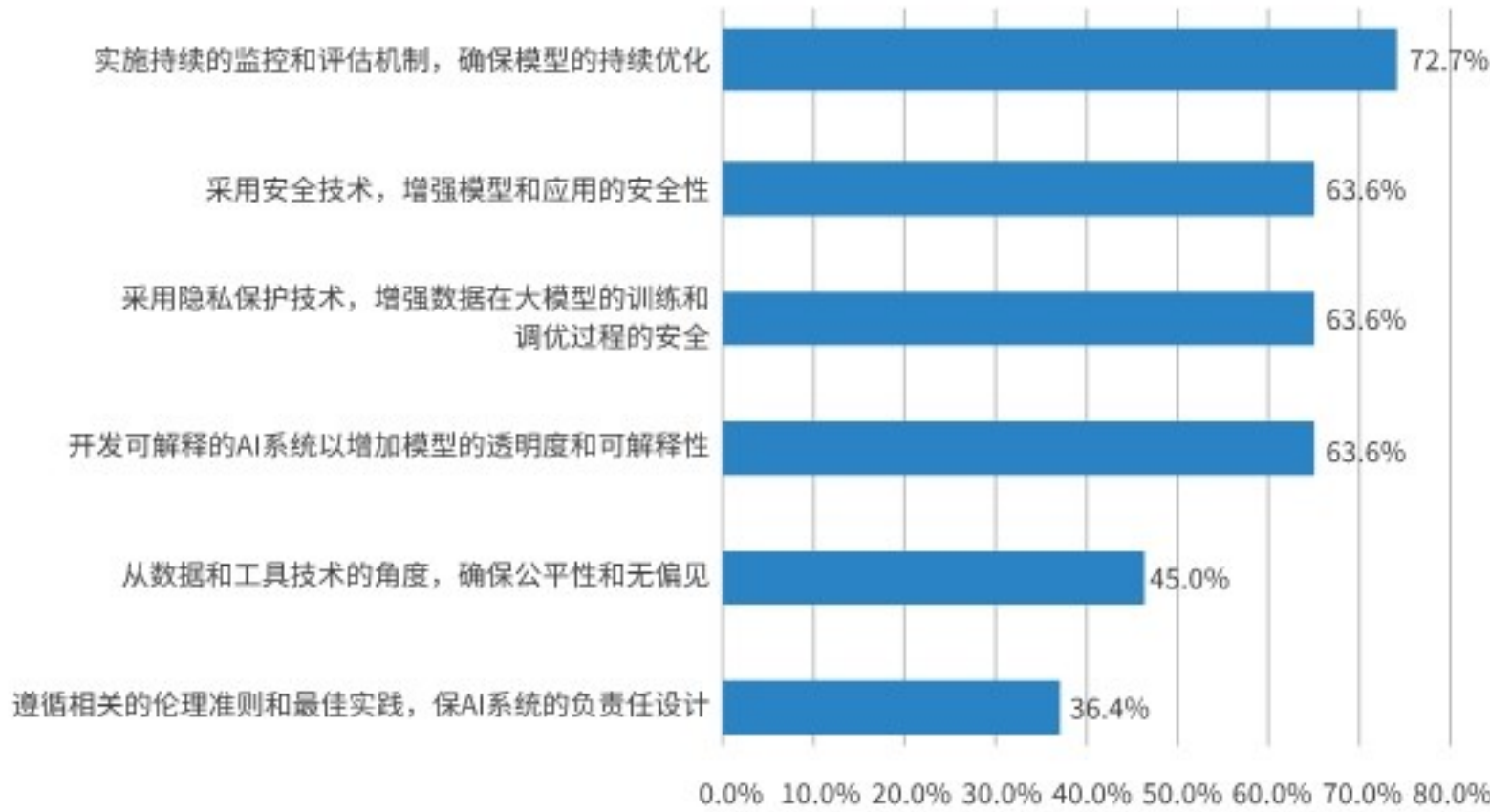
RAI具备复杂性，特定的环境、组织文化、人工智能的成熟度、业务的性质、潜在后果的严重性等多重因素塑造了RAI。RAI能力设计的关键是需要使所有相关方对生成式AI的应用进行理解和管控层面全方位对齐。

» 技术实践层面：建立大模型的持续监测和评估机制，防范潜在风险

建立大模型的监测与评估机制是防范模型风险与推进算法优化的必要条件。金融机构需梳理可能违背负责任表现的风险源头、风险类型、风险强度，思考风险产生原因、风险解决措施，防范未发生的风险，对人工智能应用的负面影响开展预见性风险评估，提前做好风险预案。在模型训练中，需要进行多元化数据收集，确保训练数据代表所有相关群体，避免数据偏差。在模型开发过程中，进行公平性测试，评估不同群体的模型表现。持续监测模型在实际应用中的表现，并根据反馈进行改进。

目前金融机构对大模型的监测与评估举措重视程度高，也体现出其对落实负责任AI的关键作用。根据IDC的调研结果，在占比18.3%的已经在推进负责任AI的机构中，目前较为普遍的举措有：持续的模型监控和评估以确保模型持续优化、

采用安全技术加强模型与应用的安全性、采用隐私保护技术增强数据在模型训练和调优中的安全、开发可解释的AI系统以增强模型的透明度与可解释性等。



来源：IDC中国2024，金融行业调研，N=60

03

AI大模型在金融行业 应用生态演化



3.1 大模型生态的界定

大模型在金融领域的应用生态是一个多元化的系统，涉及多个参与方，包括金融机构、云厂商、模型服务商等，每个角色都有其独特的价值和行业贡献。其中，金融机构是大模型技术应用的主要用户，利用大模型优化风险管理、用户服务、投资决策等业务。云厂商提供必要的云计算资源和存储能力，支持大模型的训练、部署和运行。模型服务商专注于开发和提供预训练大模型，为金融机构提供模型服务和技术支持。此外还有大量技术合作伙伴、行业合作伙伴等，共同构建大模型应用的生态系统。

对于金融机构而言，算力短缺、大规模参数带动的高IT资源消耗、人才匮乏、对大量高质量数据的需求等导致了当前应用生成式AI的高成本，形成了AI普惠应用的门槛。

我国金融机构种类繁多、数量庞大，且不同金融机构之间的资产规模、业务范围、科技实力参差不齐，存在大量长尾机构。截至2023年底，我国有4490家银行业金融机构¹；238家保险机构法人和7家外国再保险公司分公司、2566家保险专业中介机构法人²；145家证券公司³；基金管理公司145家⁴以及大量其他金融机构。以银行业为例，不同类型的银行间资产规模相差甚远，我国6家大型商业银行占据了银行业总资产份额的41.90%，而3796家农村金融机构只占了13.50%，126家城商行和550家其他类农村金融机构分别占13.7%和14.3%⁵。

1：国家金融监督管理总局披露的银行业金融机构法人名单（截至2023年12月31日）

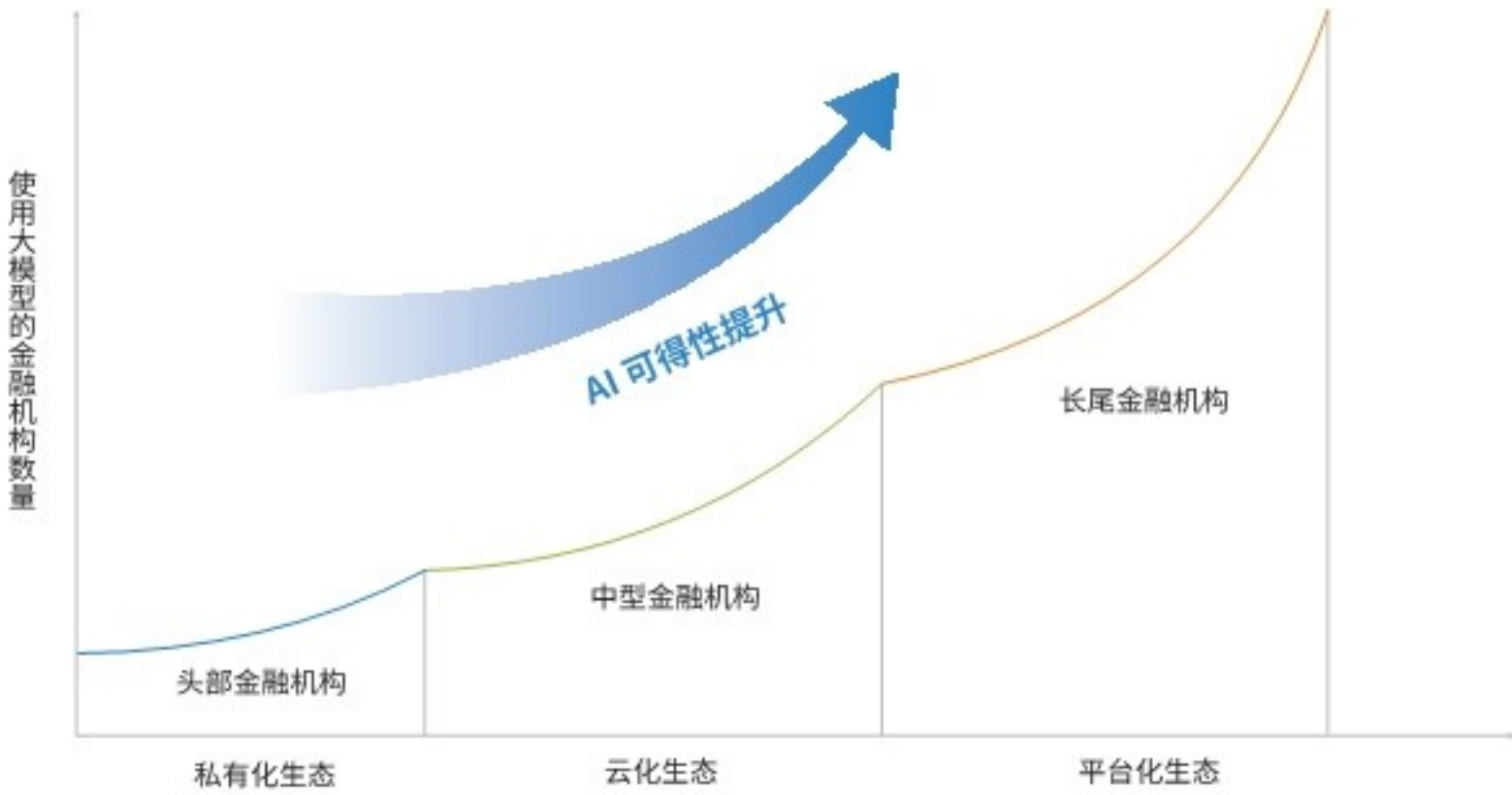
2：国家金融监督管理总局披露的保险机构法人名单、外国再保险公司分公司名单、保险专业中介机构法人名单（截至2023年12月31日）

3：中国证券业协会2023年度统计

4：中基协发布的2023年12月公募基金市场数据

5：机构分类采用国家金融监管总局的划分方式，资产份额采用国家金融监管总局披露的2024年5月的数据

蚂蚁和IDC以金融机构参与者数量、大模型应用难度的高低作为观察大模型生态发展趋势的两大主要维度，结合技术发展方向，分析大模型在中国金融领域应用中的参与者和生态模式的演变方向。针对我国大量长尾金融机构，实现金融AI普惠，提升生成式AI的应用价值，需要大模型生态持续演进，降低大模型在金融领域的应用门槛和难度。



来源：蚂蚁集团研究院绘制

随着大模型在金融行业中应用和探索的持续加深，我们预计金融行业的大模型生态将演化为三种不同的生态模式。由于我国金融机构之间的多样化和差异性，在金融行业运用的大模型生态的演进发展过程中，不同的生态模式可能会共存，而非替代关系，逐步提升大模型在金融行业的易用性和可得性。

3.2 行业生态模式演进洞察

蚂蚁和IDC结合对大模型在金融行业的落地情况的观察和趋势判断，随着时间的推移，相关技术不断成熟，各类参与方的定位逐渐明确，行业格局更加清晰，预计大模型在金融领域应用的生态将从如今的以私有化生态为主向云化生态衍化，未来预计往平台化生态模式发展。在此演变过程中，金融机构的AI使用成本逐渐降低，生态的开放共享程度提升，且各类资源在行业层面愈发得到充分整合与利用。



(注：浅蓝色块代表需要金融机构自建的能力，深蓝色块代表无需金融机构自建的能力)

来源：蚂蚁集团研究院绘制

私有化生态

私有化生态特征：头部金融机构主导、自建，对金融机构的科技建设能力有全方位的较高要求

大模型在金融行业应用中的私有化生态模式通常为头部金融机构主导进行大模型的本地化的部署与建设，对大模型技术与部分金融业务场景结合进行战略性探索。各家机构之间完全独立投入大模型所需的算力、模型训练、模型工具、数据以及金融知识、业务逻辑，各自建设不同的场景应用。

私有化生态成因：大模型在金融行业的应用处于早期阶段，各方面暂不成熟

现阶段大模型在金融行业的落地还处于早期探索阶段，私有化生态的形成原因由多方面因素构成。

- » **当前金融机构应用大模型存在较高的门槛，实力强劲的头部机构更有意愿和能力率先尝试和投入。**作为首批尝试部署大模型的头部金融机构而言，出于自身技术安全性、可控性和适配度的较高要求，因而选择私有化部署模式。此外模型的研发需要大量的前期投入，包括算法研究、模型训练和调优等，这些都需要专业的技术团队和高额的研发费用，相比于中长尾金融机构，头部机构的资金实力、人才基础、业务场景和数据的天然优势使其更具私有化部署的资源禀赋。
- » **多类型的技术服务商方能满足金融机构私有化建设大模型的条件要求。**我国通用大模型服务商选择众多，且一些AI公司和研究机构已经开发了支持私有化部署的通用大模型，以满足金融机构对于数据安全、模型定制化等方面的需求。
- » **金融机构出于降低监管风险和潜在业务风险的考量。**当前大模型在金融行业应用的相关技术与应用处于初级阶段的情况下，私有化生态能使金融机构更好地保障数据安全性，防止数据泄露和滥用，更好地控制和管理合规风险。

私有化生态运作方式：头部金融机构主导、多类型服务商参与

由于大模型在金融行业的应用还处于早期阶段且建设成本高昂，头部金融机构出于率先探索目的进行建设，在各建设参与方的合作中位于绝对主导地位，通常呈现以下特点：

第一，金融机构作为需求方“以我为主”，对大模型部署拥有较强的控制权。根据自身的业务需求和战略目标，主导大模型的建设和发展方向，会根据自己的业务场景和用户需求，定义模型的功能、性能和安全标准。

第二，多类型服务商提供不同的能力广泛参与。在大模型的建设和应用过程中，金融机构选择多种类型的服务商参与落地建设和实施，包括但不限于：

- » 模型提供商：提供AI技术、算法开发、模型训练等专业服务。
- » 数据服务提供商：提供数据采集、清洗、标注和分析服务。
- » 硬件供应商：提供所需的服务器、存储设备、GPU等硬件资源。
- » 安全服务提供商：提供数据安全、隐私保护、合规性评估等服务。
- » 应用和系统集成商：将大模型技术与金融机构现有的应用系统和流程进行集成。

私有化生态适用性与局限：可满足金融机构的高度定制化和自主性需求，但存在一定的资金和技术门槛

私有化生态下，对于头部金融机构而言优劣势共存，适用性主要是适配度高和自主性强。

- » **技术适配性。**金融机构可以根据自身技术架构和业务流程，对大模型进行适配，确保技术与现有系统的兼容性和集成性。
- » **可控性。**私有化生态模式意味着金融机构对大模型拥有完全的控制权，可以根据自身业务需求和发展战略进行定制和优化。

- » **数据安全与隐私保护。**在私有化部署模式下，头部金融机构自身的数据不出域，可以有效保护用户数据和隐私，增强数据安全性。

相对于适用性而言，私有化生态也存在局限，主要体现在模型搭建初期和后续技术迭代的高投入。

- » **算力等资源建设和利用分散化程度高。**各家机构投入大量成本自建支撑大模型的基础设施资源，但大模型的应用仅针对单个机构，资源利用效率不充分，存在资源浪费，性价比不高。
- » **模型更新与迭代受限于单一金融机构的投入与使用情况。**私有化模式下技术迭代速度可能偏慢，因为模型的更新和迭代受限于机构内部资源和投入，模型的训练效果可能受限于数据集和算力而导致智能化程度缺乏市场适应性。

建行在2023年启动“方舟计划”推进大模型建设，包含基于通用大模型打造大模型的模型训练体系、智算集群、数据体系、应用场景等，从而从算法、算力、数据等多个层面形成完备的建行大模型的支撑，同时构建大模型向量知识库，打造了多类型工具，推进了生成式AI在智能客服、市场营销、投研报告、智慧办公、智能运营、智能风控等多个场景的应用。

根据公开研究，建行的大模型落地是通过与国内领先的大模型和云服务商进行合作。从部署方式看，出于对核心数据安全性的考虑，建行对基于通用大模型训练和微调之后的模型进行了私有化部署。在大模型训练过程中，前期的数据收集、转换、确定框架模型等工作由建行完成，中后期的训练、评估、微调由国内领先的模型服务商和云服务商按照建行的需求执行。从未来应用场景的规划看，基于目前在风控、产品、客服、办公等场景的应用，下一步建行计划将大模型引入网上金融业务部和网上银行业务部，使生成式AI大模型在业务场景发挥更大的价值。

云化生态

云化生态的特征：基建共享，对金融机构自身的科技建设能力要求降低，金融机构可在云上自行开发应用

云化生态中，技术服务商将模型及算力等底层技术设施能力封装成服务提供给金融机构使用，金融机构从公有云或行业云调用大模型服务接口，在此基础上直接开发具体场景的应用，无须开发和维护底层基础能力，以降低技术使用门槛、控制成本、简化系统运维管理，从而加速AI技术在更多金融机构中的普及应用。云化生态更适合于有一定的科技建设能力、对技术投入成本预算相对有限的中型金融机构采用的大模型建设模式，目前中国金融行业典型的实践案例较少。

云化生态成因：由大模型技术成熟度提升和金融机构技术降本需求共同推动

云化生态的实现需要大模型在金融行业经过一定时间的应用而得以提高厂商的技术成熟度。届时部分头部大型金融机构对大模型的应用已经有一定的探索和经验积累，在此过程中各类生态参与方也通过与大型金融机构的合作打磨和沉淀了更多可供对外输出的统一化能力。随着更多具备一定资金和技术实力的中型金融机构追随大型金融机构开始正式入局大模型的广泛应用时，在以大型机构为主导的私有化生态之外，还将开始发展出适用于中型金融机构的云化生态。云化生态产生的主要原因如下：

- » **技术成熟度提高，金融机构对大模型的认可度提升。** 技术提供商通过前期与头部大型金融机构共同建设大模型的经验积累与能力沉淀，产品与服务成熟度得到提升，尤其是金融行业领域大模型的智能化程度与应用效果更明显，具备了规模化对外输出基础设施的能力。

- » **金融机构对成本控制的需求推动大模型应用的技术降本。**在云化生态下，云厂商对大模型所需的基础设施和模型提供云化服务，帮助金融机构控制算力投入、技术开发和运维成本，提高金融机构在大模型落地方面的投入产出性价比。

云化生态运作方式：金融机构的主导性减弱，与各技术厂商开放协作

云化生态下，生态运作的特点：

- » **金融机构的“以我为主”特性减弱，**金融机构和技术服务商深入场景共同开发。云服务厂商为金融机构提供云计算资源，包括存储、计算、网络等；金融机构提出深入的业务场景和需求，厂商则利用其技术能力，定制开发满足特定金融场景需求。
- » **生态中金融机构参与者数量增加：**伴随着技术服务商的技术降本，金融机构使用大模型的成本降低，除了大型金融机构外，中等体量的金融机构有望参与进来，对云化生态下的整体运作产生影响，如开发更多的应用场景、共享部分的数据等。
- » **开放API和生态系统。**云服务厂商提供开放的API和丰富的生态系统，使金融机构能便捷集成大模型能力，并与其他系统和服务连接。

云化生态的适用性与局限：资源利用率提高，大模型普惠性提升，但长尾机构可能仍面临一定程度的使用门槛

通过云化生态合作模式，不同金融机构将可以在算力和模型层面实现资源共享、优势互补，共同推动大模型技术的发展和运用。适用性体现在：

- » **相比于私有化生态，云化生态下实现算力、模型等技术设施共享，降低了金融机构部署的综合成本。**云化生态模式下，参与的金融机构数量相比于私有化生态更多，具备一定开发人员和研发实力的中型金融机构能够参与其中。在此模式下，可以不用自建大模型相关的算力、网络、数据、存储等基础资

源，也无需投入对应的人才和成本做基于通用大模型的精调，而是在统一的云服务基础上，直接使用成熟的基础模型进行推理部署，减少本地化自主开发建设的成本和工作量。

- » **资源利用和配置更加集中，减少了资源的分散建设和使用。**行业层面集中算力资源、技术资源统一提供服务或成为大势所趋。在各家机构自建的情况下，有限的算力资源及人才资源等各类资源处于分散状态。云化生态模式下，通过云服务为金融行业提供算力、模型等资源，可以实现算力资源更加集中和统一的利用与配置。
- » **便于模型的持续迭代与优化。**基于云服务，金融机构基于业务需求大量调用大模型的情况下，模型本身不断地优化和迭代，金融机构也可以及时享受大模型的快速迭代和优化能力成果，以适应市场变化和业务发展。

但是云化生态依然存在局限，主要是应用门槛尚未大幅降低以及存在生态壁垒。

- » **应用门槛并没有被大幅降低：**对于长尾金融机构而言，依然需要配备对应的技术开发人员进行应用场景的开发与探索，依然存在一定的技术使用成本和门槛，需要一定的应用开发周期。
- » **不同的云化生态之间尚难实现互通：**不同的模型对底层的算力等基础设施的适配没有完全打通，依然需要金融机构考虑不同模型和平台与底层的基础设施的适配问题，进行合理的配置。



案例：Azure AI Studio-以“MaaP + MaaS服务模式”打造一站式AI开发应用平台

Azure AI Studio是微软Azure AI打造的一站式AI开发应用平台，结合MaaP（Models as a Platform）和MaaS（Models as a Service）的服务模式，金融机构无需自己预置GPU或管理后端操作，就可以访问和使用Azure聚合的各种开源或闭源模型。通过为金融机构提供现成可用的基础模型和共享的计算资源，降低AI开发的技术门槛，让金融机构能够将AI模型轻松集成到应用程序中，加速AI技术的应用和创新。

- » 打造模型聚合平台，独家合作Open AI+开放合作多元模型厂商，提供丰富的模型目录，支持1600以上的第三方模型，其中包括OpenAI、Hugging Face等的模型，并且通过高质量数据和专业知识注入，提供多款自研的轻量模型（Phi）和垂直模型，供金融机构按需选用和适配。
- » 提供包含MaaP（Models as a Platform）和MaaS（Models as a Service）的服务模式。通过MaaS模式，支持主流LLM工具如prompt flow，Langchain，机构按需调用基础模型和开发工具，将合适的模型快速集成到自己的应用中；通过MaaP模式，根据特定场景和数据基础导入数据、微调和评估模型，以适应特定的业务场景和需求，实现模型的定制化。
- » 提供全生命周期的模型工具，助力大模型能力落地。提供包含多源数据整合、检索增强、LLM工具、模型能力评估、全链条安全能力在内的模型工具，通过多种辅助能力将模型能力与金融机构的业务场景之间链接与打通，实现模型在金融机构的工程化落地与应用。

Azure AI的典型用户有伦敦交易所、穆迪等。2022年12月，伦交所集团与微软达成为期十年的战略合作，双方将携手打造下一代数据与分析及云端基础设施解决方案。

平台化生态

平台化生态特征：从底层算力到上层应用均通过平台输出，金融机构“开箱即用”

未来大模型在金融领域的生态或发展为依托多个大模型服务创新平台，向金融机构提供大模型应用所需的全套资源，包含算力等基础设施资源、通用大模型及各类专业领域小模型等多样化的模型资源、金融业务的各类场景应用资源，以及连接金融产品和服务终端用户。这一生态模式整合金融机构、技术提供方的资源和金融用户的需求。跟云化生态相比，这一层级的生态具备以下特点：

» **大模型技术与金融业务要素共享，开放连接多个参与主体，形成全层级的产品与服务整合与沉淀，平台化输出能力：**

1) **连接金融机构和技术提供方。**金融机构和技术提供方可以通过平台共享资源，如数据、算力、模型等。各类技术提供方的产品与服务通过平台整合，实现技术提供方与需求方的充分链接匹配。

2) **连接应用开发者和技术提供方。**为应用开发者和技术提供方（如云服务提供商、工具和平台供应商）建立连接，构建开发者生态。平台为应用开发者提供集成开发环境和易于使用的开发工具，使开发者能够快速构建和部署应用程序。通过平台化生态，大模型的各类技术提供方可以与应用开发者建立紧密的联系，形成一个互利共赢的开发者生态系统。从而推动技术创新和应用开发，还能为金融机构提供更丰富、更高质量的服务和产品。

3) **链接金融用户和金融机构。**金融机构可以通过平台集成多种金融服务和产品，提供一站式的用户体验，增强用户粘性，还可以促进金融机构之间的合作，形成联盟，共同开发新的金融产品和服务，丰富金融机构与用户之间的服务场景。



来源：蚂蚁集团研究院&IDC绘制

- » **更加深入场景，提供开箱即用的AI应用。**平台具备开箱即用的针对具体场景的AI应用，金融机构按需直接调用，因此这一层级的平台化生态能力更加原子化。金融机构通过与平台合作，快速链接使用，无需自建模型和开发应用，以足够便捷和低成本的方式解决大模型在金融机构落地的工程化问题。
- » **金融机构通过专业智能体实现不同平台的能力调用，打破平台边界。**依托专业智能体的跨平台深度连接，可以集成不同平台的多种AI服务和工具，金融机构通过专业智能体灵活调用不同平台的功能，实现跨平台的充分协同合作与开放，带来智能化服务的代际升级。

平台化生态成因：长尾金融机构应用需求明显，未来大模型在金融领域应用的技术和制度环境逐渐成熟

平台化生态有可能是大模型在金融行业应用进一步发展到成熟阶段之后可适用于科技力量极为薄弱的长尾金融机构的生态模式。届时大模型在金融行业的应用经过头部机构和中部机构的实践，已经得到了对应的应用价值的充分证明，大

量长尾部金融机构也有强烈的应用动力；同时从技术提供商角度，经过与更多金融机构的协同共建场景应用进行更细化的能力打磨和沉淀，各类技术提供参与方也具备了更多可对外规模化输出的成熟能力。

- » **服务商统一进行技术和能力整合。**在私有化生态中，以金融机构为绝对主导，各类生态参与方效用有限，且各层级资源形成机构化独立与分割，分散建设；而在云化生态中，依然存在金融机构针对机构内具体应用的独立建设环节。平台化生态以统一的行业平台形式对相关能力统一整合并合理分配，进行统一输出，充分实现大模型资源的高效利用，形成合力提升大模型的能力。
- » **大量长尾金融机构进一步的技术降本和应用需求。**对于云化生态而言，金融机构虽然可以不用自建算力基础设施和模型，但是依然需要基于自身的场景和数据进行进一步的场景化应用开发，对于开发人员和对应的成本投入依然有要求。而对于大量的长尾小型金融机构而言，开发人员配备并不齐全，并且从业务类型和规模而言也不具备独立开发具体场景应用的性价比，因此开箱即用的端到端服务调用对于该类金融机构而言，更能解决其技术降本和应用需求的矛盾。
- » **未来大模型相关的监管和标准趋向完善，各类接口和协议有了较为统一的标准，从而具备进行行业平台化整合的前提。**一方面，行业治理准则对大模型的服务质量进行规范，大模型将服务成百上千个细分金融场景业务，服务质量直接关系到金融业务能否稳定运行，涉及金融风险的防范与控制，因此服务内容、服务水平、服务计量的标准化成为必要前提，平台化生态下有助于各技术标准达成共识。另一方面，只有行业内明确和统一的标准才能将不同金融机构与企业内部模型服务的接口逐步统一，形成标准化的接口，便于不同业务调用不同模型服务，提高兼容化处理效率，降低使用门槛，促进不同服务之间的互联互通，推动模型与各类技术、资源之间的共享。

平台化生态运作方式：平台聚合资源，参与方各司其职，推动平台场景、数据、模型形成正向循环

在平台化生态模式下，各类能力、资源和服务统一集中到行业化平台，底层算力基础设施与模型、上层应用之间实现打通与整合，形成具备领先大模型能力的金融服务平台。金融机构可降低对适配性和具体架构的投入，只需基于自身AI应用需求直接调用平台对应的服务能力即可。且随着调用服务的金融机构数量增加，平台活跃度提升，整个平台生态形成愈发丰富与完善的正向循环。

- » **以行业的方式基于行业云或团体云服务，突破金融机构边界，开放共享。**所有大模型服务基于金融机构的需求，实现场景化需求驱动的商业连接，资源管理突破金融机构边界，实现基于服务优化的资源分配。
- » **平台统一集中大模型相关的各层级行业资源，经过能力封装打通壁垒，提供多项开箱即用的服务。**金融机构无需高成本投入，在平台可使用所需场景的大模型应用，调用针对具体应用场景的端到端服务，让合适的组件提供合适的服务，充分满足欠缺资金实力和科技建设能力的长尾金融机构的大模型应用需求。
- » **生态内各参与方各司其职，各取所需。**平台机构将各类技术提供商的能力统一集成和沉淀到平台，集成算力、网络、模型、工具插件、应用等。其中，各类软硬件产品提供商和技术开发商可专注于自身产品和能力建设，平台机构进行统一沉淀和协同，发挥不同参与方的长处并实现合理的资源分配，保障整体生态的稳定性。平台将各类服务提供给金融机构直接调用，使金融机构无需投入过多资金和精力于相关技术能力建设中，能便捷地应用大模型技术改善金融服务的质量。
- » **整体大模型在金融行业应用生态形成数据、场景、模型的正向循环。**以强大的AI算力基础设施为底座支撑，基于统一的平台，多类模型、通用工具链、金融插件工具、金融数据和金融业务场景，通过低门槛的使用方式吸引更多

长尾金融机构应用，能够提升平台活跃度，产生“应用场景活跃-业务数据增加-模型迭代加速”的正向循环，沉淀更多原子化能力，形成更完善、更丰富的生态。

平台化生态的适用性：大模型在金融行业应用的普惠性大幅提升，通过赋能更多长尾金融机构来提升AI对整体金融行业的价值贡献

平台化生态是大模型在金融行业的应用发展到成熟高级阶段的生态形式，能最大限度地将大模型技术的赋能范围扩大到技术能力和资金实力欠佳的长尾金融机构中，实现真正意义上的金融AI能力普惠。

- » **快速便捷低成本，普惠性大大提升。**由于在此模式下对金融机构的AI使用成本要求极低，可以在云化生态的基础上进一步延伸至大量的长尾小型金融机构，对于业务规模较小、开发人员配备欠缺、资金实力弱或想要跳过开发周期即刻采用AI应用的机构，AI对金融机构的普惠性得到显著提升。
- » **实现算法、数据、算力能力的有序开放，有助于充分发挥大模型的迭代和提升效率。**由于在平台化生态模式下，多模态、大小模型、不同场景模型之间开放互补，能加快模型优化迭代，提升模型性能，共同突破技术难题。同时金融机构的各类业务场景都可实现即插即用的AI服务，因此AI在具体应用场景的种类和频率得到极大增加，促进更多金融数据积累，从而能够反哺模型进一步优化和沉淀，从而使大模型得到更为及时和有效的提升，充分反应市场变化与需求。
- » **信息技术资源配置效率提高。**平台化生态打通和深入到金融机构的具体应用场景，贯穿从底层到上层的整个大模型应用链条，突破了金融机构的选型限制，能够有更多数据、场景资源进行共享和复用，减少了重复和孤立的应用开发工作，实现资源配置效率的进一步提升。
- » **对金融用户价值进一步提升。**平台化生态下，由于对各类资源的充分协同和利用，场景、数据、模型之间能充分形成正向循环，提升大模型对金融业务

的赋能和实现数智驱动的金融创新，为金融行业的企业用户和个人用户提供
更智能、更便捷、更高效的金融服务，实现大模型对金融用户价值的进一步
提升。



案例：苹果-开放Apple Intelligence，构建端侧平台化生态

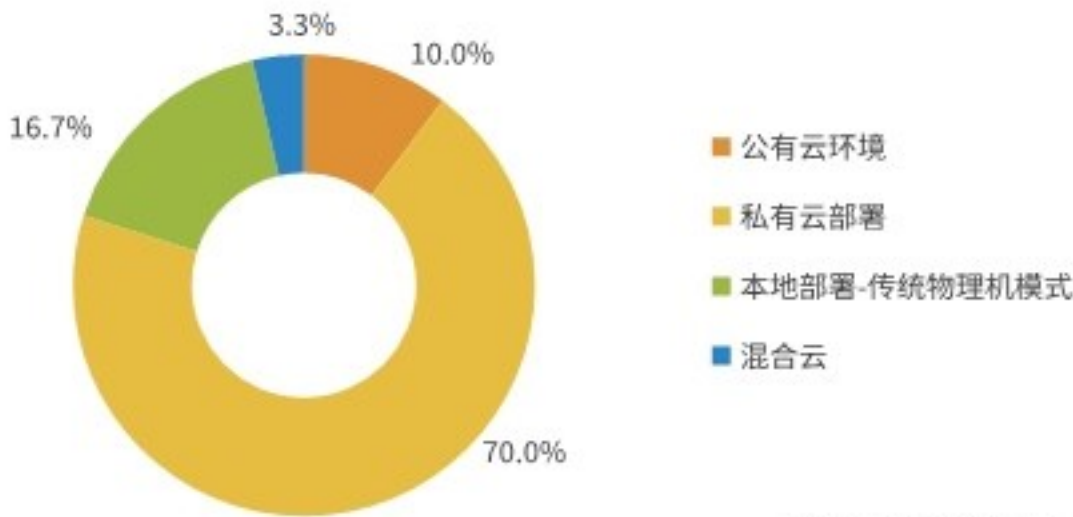
2024年6月，苹果在WWDC上正式发布AI智能系统Apple Intelligence，将大模型集成到iPhone、iPad和Mac中，有望为用户带来更加个性化和智能的体验，真正让AI渗透到每个人的生活场景中；并从场景应用、开发者工具、模型与算力多层面进一步夯实平台能力，为生态合作伙伴提供“开箱即用”的AI生态。

- » 场景与应用层，Apple Intelligence实现了端侧原生应用升级，如Siri、Music、TV+、Fitness+及News等，基于用户偏好及日常使用习惯，能够实现个性化用户推送和定制化内容生成，引导用户智能化应用心智及提高用户粘性。未来Siri有望作为端侧智能体，根据用户个性化需求链接第三方AI应用，基于多方智能体交互，实现生态伙伴AI应用与场景、用户有机聚合，成为智能应用的超级入口。
- » 开发工具层，Apple Intelligence为开发者开放了底层智能API接口，将苹果的AI能力将变成通用的能力支撑，供应用开发者调用和创新。开发者能够通过App Intents API在自己的应用程序中引入Apple Intelligence，完成产品的智能化升级，接入到苹果的AI生态内。
- » 模型与算力层，苹果自研A及M系列芯片专注神经算力提升，可以支持端侧AI处理，为Apple Intelligence打造端侧平台生态提供坚实算力基础。模型方面，苹果自研端侧小参数模型（30亿参数），并与大模型厂商开展合作，例如将OpenAI ChatGPT 集成到Siri中，显著提升了其智能对话能力。未来，苹果计划将业界领先的大模型集成在iOS 18、iPadOS 18 和macOS Sequoia 中，苹果作为AI聚合平台，将深层次赋能生态伙伴开发和应用。

苹果构建AI硬件端侧平台化生态将成为智能应用超级入口及AI能力聚合平台，有望为开发者和用户提供更易用且多样的AI体验，随着未来开放程度不断提升以及Apple Intelligence平台模式持续优化，将进一步丰富平台生态，推动用户、场景、技术和服务的正向循环。

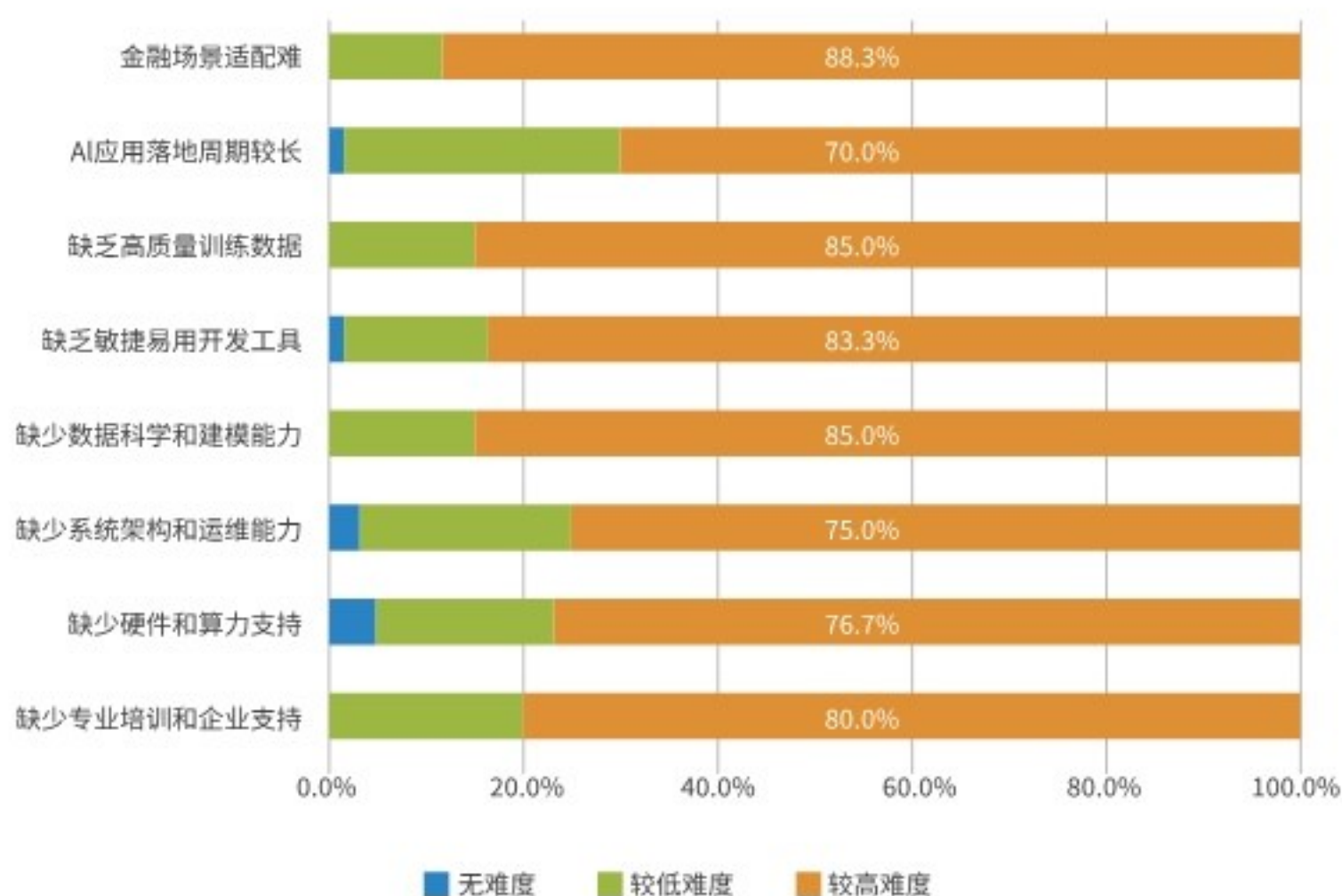
3.3 目前行业生态的基本情况

从金融机构的大模型能力建设现状来看，当前国内的金融机构以私有云部署为主。与美国金融机构广泛接受公有云、混合云等部署形式相比，我国目前生态体系较为单一。且在部署过程中采用了多项工具，具备较高的复杂性和成本投入，同时当前阶段金融机构在技术层面、工具层面以及组织层面都存在较大的难度。



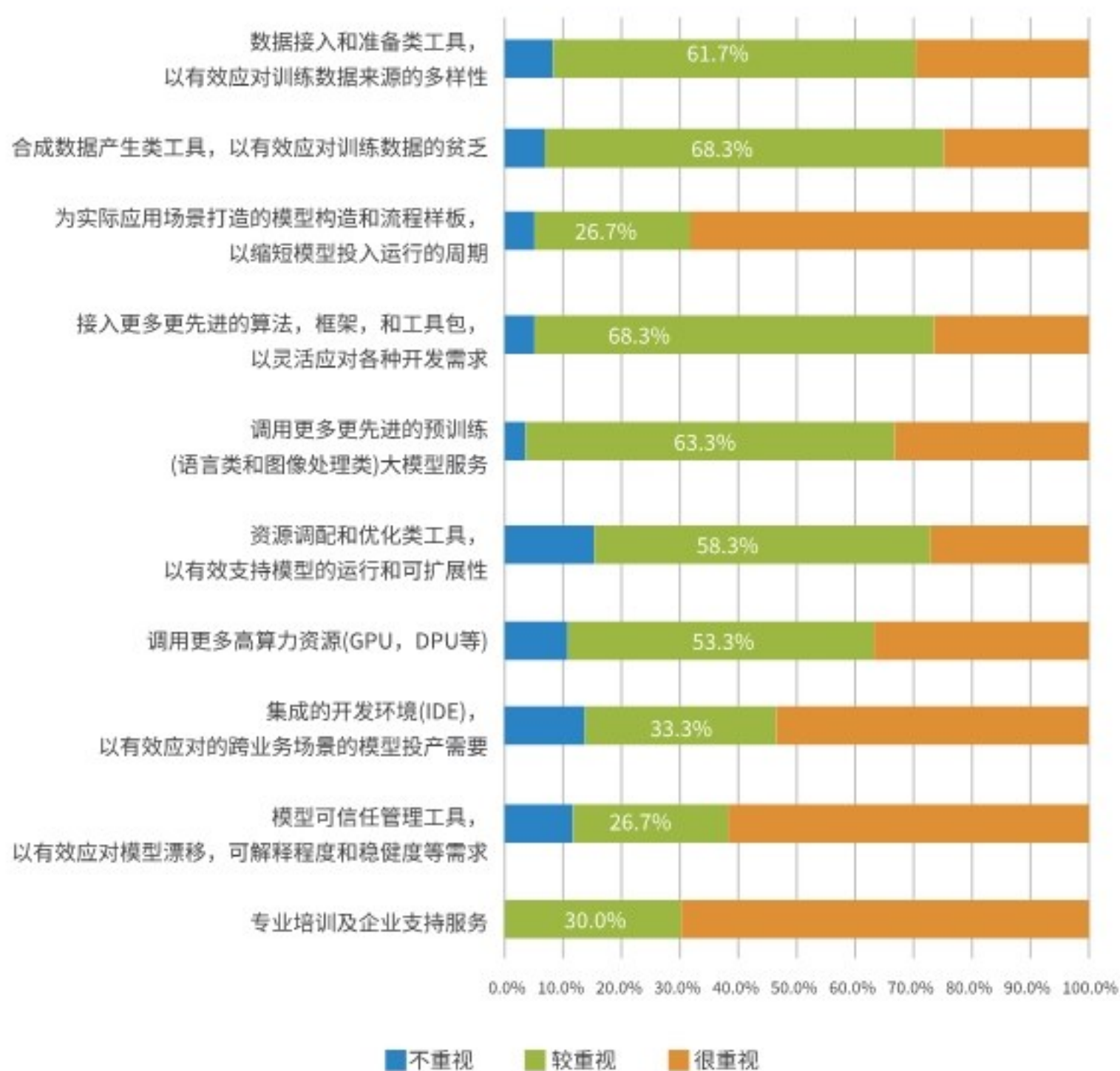
来源：IDC中国2024，金融行业调研，N=60

此外，受访金融机构当前存在多方面的能力欠缺。一方面，金融场景适配难，难以界定和找到较为合适的优先落地的场景；另一方面，缺乏敏捷易用开发工具和高质量训练数据，缺少数据科学和模型构建方面的技术能力也进一步提升了落地的难度；此外系统架构和运维能力、硬件和算力支持也都形成了一定程度阻碍；最后应用落地的长周期、专业的培训和企业支持都对大模型在金融机构的落地构成一定难度。



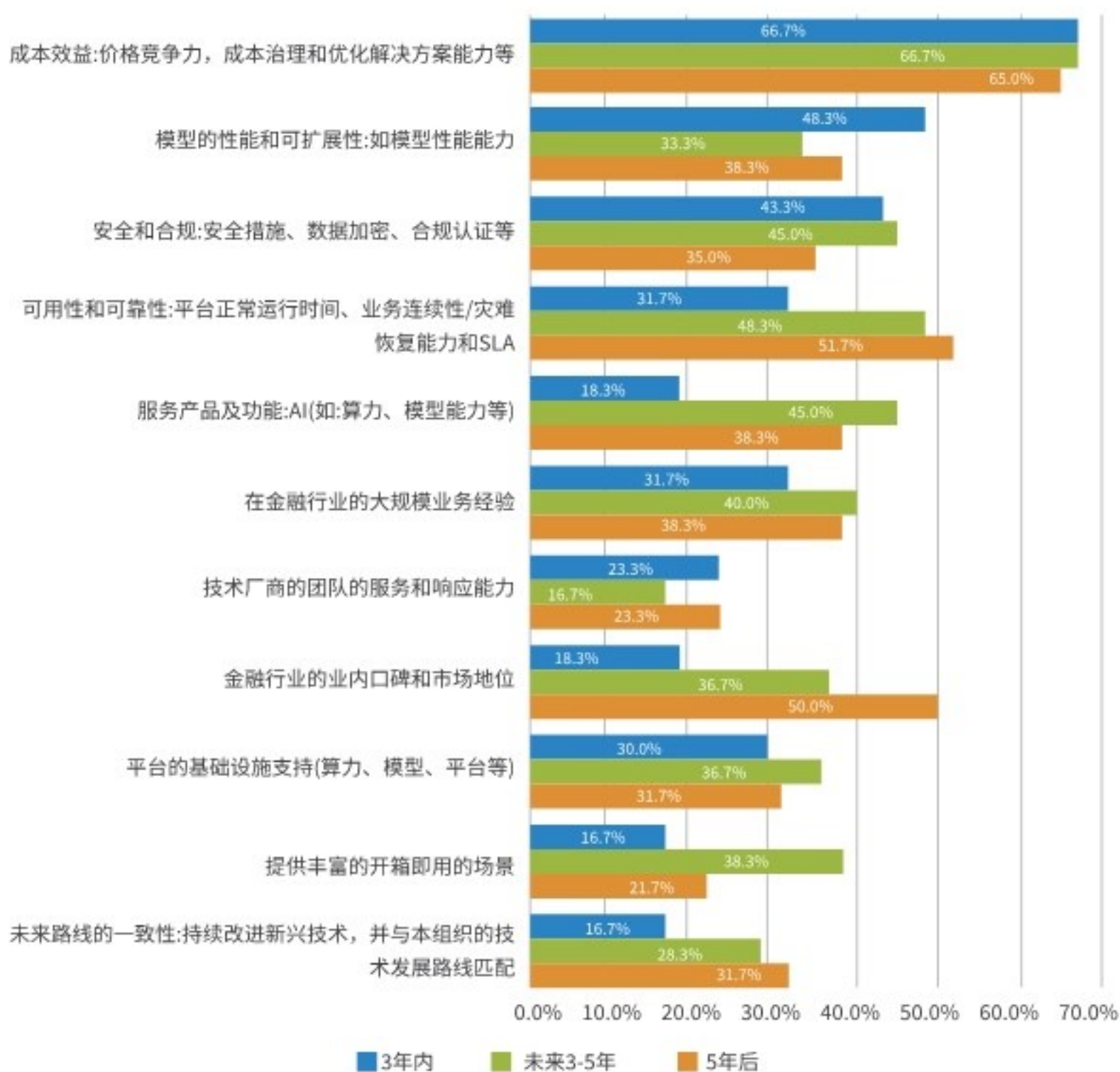
来源：IDC中国2024，金融行业调研，N=60

从IDC调研结果看，受访金融机构表示近期内针对大模型能力建设方面，会更重视模型部署的流程、模型可信任相关的工具、专业的培训方面的采用和投入。体现出金融机构对当前大模型能力建设过程中，模型投入运行周期、模型可解释性和稳健度以及对相关人才储备层面的困扰以及改善诉求，说明当前大模型在金融领域的落地过程中诸多细项能力欠缺且不成熟，在技术基础、人才各方面都还需要持续投入和建设。



来源：IDC中国2024，金融行业调研，N=60

在服务商的选择方面，根据IDC的调研，受访金融机构在选择技术服务厂商的重点考虑因素随着时间推移变化，其中成本因素是第一考虑，其次服务商的业界口碑、大模型的安全合规及大模型性能等也是金融机构关注的核心要素。



来源: IDC中国2024, 金融行业调研, N=60

由此可见, 技术降本贯穿金融机构当前与未来使用大模型技术的一大主要诉求, 同时随着时间的发展, 金融机构对大模型相关技术与服务的要求从基本的性能与安全向更进一步的可用性与口碑发展, 体现出对便捷性、应用效果的更高追求。

3.4 推动生态演化的建议

金融机构：构建长期且清晰的智能化目标

金融机构作为大模型技术的需求方，应正视其短期的局限性，以及长期的可能性，建立长期的战略性目标，选择合适的建设路径。同时，金融行业主管机构营造包容审慎的监管环境，制定金融场景AI应用指引和技术标准，促进生成式AI在金融行业落地。

- » 树立拥抱大模型的积极态度，结合自身禀赋制定清晰规划，立足长期收益而非短期收益考量。
- » 各类金融机构基于资金、业务规模、人才、技术基础等多方面要素，合理选择生态部署模式。
- » 在严守监管底线的前提下，同时响应监管的鼓励金融服务创新的精神，在监管指导下结合大模型的应用探索多样化的金融产品和服务，实现大模型与金融业务的安全结合与应用创新。

大模型技术与服务提供方：各类型的技术提供商加强开放合作与创新

当前，多种不同类型的技术供应商为金融机构大模型的部署与落地提供服务

大模型在金融行业的系统性落地需要多类技术服务商共同参与，每类型的服务商都在AI应用中扮演着关键角色。目前金融行业生态参与方如下：



来源：IDC中国2024，IDC金融研究绘制

基础设施提供方为大模型落地提供必要的硬件和软件资源，基础模型厂商提供通用的文本、图像、多模态等基础模型，这些模型构成了金融机构AI能力的基础。AI模型工具平台厂商帮助金融机构对基础模型进行进一步的调优、开发、运维、治理等工作。AI应用集成和开发商结合具体金融机构的业务场景需求为大模型在金融机构的落地提供应用开发与集成服务。行业专业的咨询服务商为大模型在金融行业的务实落地提供专业的咨询服务，其中包括应用场景的战略落地咨询、IT能力评估与准备、运营管理和培训服务等。此外数据服务商，为金融机构落地大模型各个环节、不同层次提供从下至上的数据相关服务，确保数据质量和安全，为大模型的应用提供高质量的数据。

整体来看，金融行业当前的生态参与者互相依赖，相互联系，金融机构作为需求主体,需要多类型技术服务商高效、密切协同，共同推动AI技术在金融行业的落地和跨越式发展。

以平台为依托，生态参与方共同推进大模型在金融领域应用的生态进步

提升生成式AI在金融行业的应用价值，推动金融行业大模型生态向平台化演进是生态建设者和生态参与者的共同的责任。以平台为依托，各方发挥自身的优势，持续演化，才能释放大模型在金融领域应用的最大势能。

平台建设方：汇聚多方技术资源，释放大模型落地的最大势能

平台建设方作为生态演进至第三阶段的关键力量，广泛汇聚生态不同角色、整合不同技术供应商的能力、吸引海量金融用户，为用户提供综合性金融服务平台，是生成式AI得以更加稳步务实落地的核心。平台建设方必须发挥自身的集大成优势，拆除技术围栏、降低应用的门槛，将先进的技术资源、专业的服务和多样化的解决方案、丰富的应用场景融为一体，汇聚金融服务需求方与金融服务的供给方，为参与者提供一站式的供需对接和服务管理。平台建设方通过合规安全的准入审核、高效的运营管理、公平合理的商业模式、谨慎的风险管理等，推动各方协同合作，共同打造高效、稳定、安全的技术支撑体系和用户易得、易用、放心的金融服务，助力大模型在金融机构业务场景中的快速发展，释放大模型生态最佳势能，从而推动金融行业智能化的未来发展。

同时平台汇聚的不同类型的厂商应发挥各自的优势与价值，共同为推动生态演进作出贡献：

算力厂商：提供高性能、安全稳定的产品和解决方案

由于金融业务交易具有安全性、准确性、实时性、并发性等特点，大模型在处理海量数据、复杂模型以及训练、推理、调优、实时检索和运行等任务时需要更强大的计算资源，提供强大的处理能力，以支撑大规模数据处理、高效模型训练以增强推理。因此算力厂商需提供高性能、安全稳定的产品和解决方案，为算力效能的发挥和算力资源的优化利用提供更加经济高效的解决方案，为大模型的运行提供坚实的基础设施。

云服务商：降低技术应用门槛，优化商业模式

云服务商应进一步降低门槛，满足机构的ROI最优的诉求，同时优化商业模式，为金融机构提供更便捷可用的AI应用能力。

- » **推动原有云计算架构体系的智能化升级**，增加和优化算力资源，提供稳定强大的云计算、存储和网络的有力支撑，做好合理的资源分配。
- » **筑牢模型平台的生产和部署能力**，在原有的PaaS基础上增强平台层能力，与更多开源及商业化模型做好生态支持与对接，构建开发者生态，做好不同技术提供商与应用开发者的能力沉淀与整合。
- » **探索新的商业模式**，为金融机构提供低成本、低门槛的AI使用服务，帮助更多金融机构进入生态圈，享受AI赋能。

基础模型服务商：打好大模型基座，让模型能力更强

通过完善自身的研发体系，拓展与技术供应商之间开放有效的合作关系，增强模型能力，加强模型应用创新，为用户提供可直接调取的AI能力。

- » **完善自身研发体系**：对自身的模型库、数据集及服务进行集约化管理。
- » **与技术供应商之间开放有效的合作关系**：加强对各类应用开发商技术的集成，支持用户基于多种模型服务及组件进行编排或开发，创造更加个性化更加便捷的AI应用。

行业数据提供商和金融知识、业务逻辑服务商：在满足监管前提下，加强优质行业数据的积累，提供更多元化的模型训练数据和分析工具

- » **加强数据治理，提供高质量的数据**：行业数据厂商需要加强数据治理，提升行业数据的准确性、完整性和可靠性。

- » **数据共享与开放生态：**数据厂商应推动建立合规的公共数据开放生态和金融行业数据共享生态，促进数据流通和共享。
- » **风险管理和监管合规：**金融数据厂商在提供服务的同时，需要关注数据安全、用户隐私保护以及监管合规性，避免因数据垄断或不当使用数据而带来的风险。
- » **沉淀针对不同金融业务的多样化数据分析能力：**基于不同的金融业务场景，提供符合业务逻辑的全面、准确和有价值的金融知识和分析工具。

应用开发厂商：全面结合大模型相关技术提供针对不同场景的智能应用开发服务

- » **技术研发与创新：**应用开发厂商需要持续进行技术迭代，利用人工智能、大数据、区块链等技术，深入金融机构的需求场景，高效开发智能化金融产品和解决方案，助力提升金融服务的智能化水平。
- » **深入细分应用进行智能化开发：**针对金融行业特定的细分场景需求，厂商提供深入业务的大模型技术应用开发方案，帮助金融机构解决细分场景的问题。
- » **参与开源生态建设：**应用开发商应积极参与开源技术社区建设，通过主动开源、贡献代码解决行业共性问题，提升开源技术整体应用水平，并加强开源技术人才培养，推进开源技术迭代升级。

结语和未来展望

随着人工智能技术的快速发展，大模型在金融行业的应用将愈发成为推动金融服务创新和转型的关键力量。未来，在负责任AI持续在金融机构贯彻落地的基础上，随着大模型在金融领域的应用生态的不断成熟和监管环境的逐步完善，大模型有望在更多场景中发挥更大的应用价值，推动金融行业AI应用的普惠性逐步实现。

展望未来，大模型在金融行业应用将在技术、生态模式多方面产生突破而不断演进，为金融机构和用户提供更高效、更安全、更个性化的服务，推动金融行业的持续发展和变革。

- » **从技术突破层面**，深度学习与神经网络、量子计算、自然语言处理、边缘计算、区块链技术等多种新兴技术的发展将不断与大模型在金融领域的应用相结合，驱动对安全性和透明度、大规模复杂任务处理、准确性、响应速度等各项技术难题的逐步改善，进一步降低技术门槛并提升业务价值，为更多金融机构接受大模型的深度应用提供技术保障。
- » **从生态模式层面**，未来大模型在金融领域的应用有望实现充分的资源整合与能力沉淀融合，大模型的各类参与方之间连接增强，金融机构、科研机构、应用开发者、技术提供商等各类参与方各司其职、开放创新，共同完善全方位多层次的大模型在金融领域的服务能力，为金融机构提供低成本、开箱即用的各类能力，驱动金融服务创新，助力金融机构全方位覆盖用户群体。

研究方法说明：

本研究中所提供的信息是蚂蚁集团研究院与IDC 对IT市场不断研究的结果。研究中使用的数据由一手和二手材料信息两种来源组成。其中一手信息来源包括蚂蚁集团研究院对金融行业技术领域市场持续跟踪和深度研究，以及IDC对市场持续统计研究与预测及最终用户调研数据和定制化咨询结果等。二手材料信息主要来源为权威机构公开发布的信息，厂商介绍及新闻稿等，蚂蚁集团研究院与IDC综合整理研究编制而成。

附录：

颁发时间	文件名称	颁发单位/组织	负责任AI相关的文件精神
2019年6月17日	《新一代人工智能治理原则——发展负责任的人工智能》	国家新一代人工智能治理专业委员会	提出了人工智能治理的框架和行动指南，强调了和谐友好、公平公正、包容共享、尊重隐私、安全可控、共担责任、开放协作、敏捷治理等八条原则。
2021年1月5日	《网络安全标准实践指南——人工智能伦理安全风险防范指引》	全国信息安全标准化技术委员会秘书处	针对人工智能伦理安全风险，提出了一系列基本要求和针对研究开发、设计制造、部署应用等不同阶段的具体指导建议，以确保人工智能的安全可控性。
2021年3月12日	《十四五规划和2035年远景目标纲要》	新华社	明确提出要“大力弘扬新时代科学家精神，强化科研诚信建设，健全科技伦理体系”
2021年9月25日	《新一代人工智能伦理规范》	国家新一代人工智能治理专业委员会	将伦理道德融入人工智能全生命周期，促进公平、公正、和谐、安全，避免偏见、歧视、隐私和信息泄露等问题。
2022年3月20日	《关于加强科技伦理治理的意见》	中共中央办公厅和国务院办公厅	将人工智能作为需要加强科技伦理治理的重点领域之一，并提出一系列措施来加强科技伦理治理，包括完善政府科技伦理管理体制、压实创新主体科技伦理管理主体责任等。
2022年7月29日	《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》	科技部、教育部、工业和信息化部、交通运输部、农业农村部、国家卫健委	系统指导各地方和各主体加快人工智能场景应用，推动经济高质量发展。
2022年11月17日	《关于规范人工智能军事应用的立场文件》	外交部	呼吁遵守国家或地区人工智能伦理道德准则，并从监管、研发、使用及国际合作等方面提出了中国在人工智能伦理治理方面的主张。
2023年7月10日	《生成式人工智能服务管理暂行办法》	国家互联网信息办公室等	规范生成式人工智能服务，维护国家安全和社会公共利益，保护公民、法人和其他组织的合法权益。
2023年9月7日	《科技伦理审查办法（试行）》	科技部等	为科技伦理审查提供了明确的指导和规范，包括审查主体、审查程序、监督管理等内容。
2023年10月18日	《全球人工智能治理倡议》	中央网信办	围绕人工智能发展、安全、治理三方面，系统阐述人工智能治理的中国方案，强调伦理先行和风险防范的重要性。

关于蚂蚁集团研究院

蚂蚁集团研究院是蚂蚁集团的专业研究部门，从事宏观经济、监管政策、公司战略和学术理论研究，覆盖数字科技、数字金融、ESG、数据与平台治理、宏观经济等研究领域。通过开展客观、扎实和深入的研究工作，为公司战略决策提供支持，为政府部门公共政策提供参考和研究服务，并通过研究开放平台的搭建，为国内外学术研究和政策研究提供数据、案例等支持。

关于 IDC

国际数据公司（IDC）是在信息技术、电信行业和消费科技领域，全球领先的专业的市场调查、咨询服务及会展活动提供商。IDC帮助IT专业人士、业务主管和投资机构制定以事实为基础的技术采购决策和业务发展战略。IDC在全球拥有超过1100名分析师，他们针对110多个国家的技术和行业发展机遇和趋势，提供全球化、区域性和本地的专业意见。在IDC超过50年的发展历史中，众多企业用户借助IDC的战略分析实现了其关键业务目标。IDC是IDG旗下子公司，IDG是全球领先的媒体出版，会展服务及研究咨询公司。

IDC China

IDC中国（北京）：中国北京市东城区北三环东路36号环球贸易中心E座901室

邮编：100013

+86.10.5889.1666

Twitter: @IDC

blogs.idc.com

www.idc.com

版权声明

凡是在广告、新闻发布稿或促销材料中使用 IDC信息或提及IDC都需要预先获得IDC的书面许可。如需获取许可，请致信 gms@idc.com。翻译或本地化本文档需要IDC额外的许可。

获取更多信息请访问www.idc.com，获取更多有关IDC GMS信息，

请访问<https://www.idc.com/prodserv/custom-solutions>。

版权所有 2024 IDC。未经许可，不得复制。保留所有权利。