破冰产业金融核心环节 大模型再探能力“边界”

本报记者 李晖 北京报道

随着大模型技术的快速发展，其在金融行业的应用逐渐受到关注。尽管大模型在C端市场的应用如火如荼，但在B端市场，仍处于探索偏冷的态势。少数应用在金融行业的场景，通常以智能客服或智能办公为主，缺乏真正介入核心业务环节的探索。

供应链金融尤其是无担保的下游业务，运营管理复杂、风控难度高，导致成本高企，是行业公认的业务“硬骨头”。《中国经营报》记者注意到，今年以来，一些供应链金融市场参与者正在尝试将大模型技术用于供应链金融业务的核心环节，以期解决这一难题。

在金融业务中，尤其是风险控制方面，以大语言模型为基础的通用人工智能虽然具有一定价值，实际的应用效果和价值仍有待验证。这些技术突破是否能够真正解决产业链中的问题？面对巨大的算力投入，如何平衡成本和效率？

网商银行信息科技部副总经理方珂在接受《中国经营报》记者采访时表示，目前在业务中，虽然大模型已经可以为金融风控进行一些重要辅助，但风控系统仍然是大模型的“守门员”。客户量、数据与风控能力沉淀，对某一类客群认知的积累，才是构成用大模型探索更深层次金融应用的基础。

大模型试水风控“辅助”

将大模型用于金融核心业务的“辅助”，市场上的供应链金融科技参与者也进行了不同方向和程度的尝试。

随着人工智能（AI）在生产、库存、运输等产业链各个环节的全面渗透，IDC预测，到2026年，五成全球2000强企业将使用AI工具来支持产业链流程设计，企业运营成本将至少降低5%。到2026年，25%的金融机构将使用生成式AI，助力金融服务在产业链的生产、制造、流通、消费等各个环节领域深度融合。

作为AI工具中最具关注度的大模型，到底能否在金融核心业务中发挥作用一直是市场讨论的热点。不过，理解大模型应用的价值，先要理解产业链金融遇到的困境。

在近期举行的2024数字产业链金融行业峰会上，网商银行行长冯亮举了一个例子：一家普通的小微企业，工商信息显示从事机械制造，年营业额在1000万元左右。这是根据工商信息和企业数据，一家制造行业小企业在金融机构眼中常常生成的画像。以往方式下，这家公司生产的是什么，最终去往哪里并不清楚。如果向银行申请信贷服务时，通常需要抵押房产，并等待至少一周时间。

“除非安排人力去实地尽调，否则在没有品牌企业担保的情况下，金融机构很难给予符合其经营需求的贷款额度。”冯亮表示。

上述类似中小企业如何能在产业链的视角下有更多维度的信息在线上“看见”，且不走上传统金融机构线下风控（对应了更高的融资成本）的老路子？2023年3月开始，网商银行技术团队尝试用大模型来解决这一问题。

大模型在这一过程中主要解决的是风控之前“精确识别”的问题。

方珂在接受记者采访时表示，通过把海量企业、工商等信息交给大模型去“计算”，大模型可以用知识抽取能力，从海量信息中形成产业链图谱，再通过多模态数据融合、协同推理等技术识别小微企业的主营业务，将其精准挂载到产业链上。比如汽车产业链，大模型最终可以“看到”发动机厂商、4S店、轴承厂商等环节，然后看见每个环节分布着哪些企业。

通过这种方式，网商银行大雁系统识别到了前述小企业更多的信息——这家生产高温尼龙材料的企业拥有12项专利，是浙江省的高新技术企业。而高温尼龙是汽车电子元器件上的关键原材料，用于保护连杆器核心电子器件的绝缘层。从产品供应客户看，该企业的产品辗转几个环节最终去到了比亚迪汽车，因此算是新能源车企供应商的供应商。

此外，在地域上，可以识别到该公司所处的产业带是浙江嘉兴平湖，是中国四大高新材料的聚集地，这一地域的公司业务相对稳定性更强。

“这些颗粒度更细的数据最终可以让风控系统‘看到’更精确的企业画像，为线上风控决策系统进行非常重要的辅助。”方珂向记者表示。

记者注意到，这种将大模型用于金融核心业务的“辅助”，市场上的供应链金融科技参与者也进行了不同方向和程度的尝试。

联易融相关负责人在接受记者采访时表示，大模型在供应链金融中应用后主要提升了产业信息整合与交易分析的效率，实现智能化风险评估，从而进一步降低融资及运营成本。据其透露，目前其内部研发的供应链金融GPT模型已运用在与一家大型外资银行合作的AI智能审单科技项目中。

与此前的风控方式相比，大模型的优势表现在处理各种复杂文档及适应不同业务需求上。“比如在上述审单项目中使用的供应链金融GPT模型，我们在通用大模型基础上，用近200万张各个专业领域的文档图片对供应链金融GPT模型进行了重新训练，训练数据涵盖合同、商业发票等关键文档类型。从结果看，大模型在要素识别与定位、规则基于语义的比对等方面表现出了更精确的能力，同时单次调用成本低至几分钱，成本效益突出。”该负责人表示。

京东供应链金融科技相关负责人向记者透露，目前正在探索将大模型用于行业风险监测与预警、动产融资模式下押品准入+估值、应收融资模式下供应商信用评估、小微金融模式下中小企业信用评估等场景的实践。

大模型的“能”与“不能”

模型高价值应用，最关键的是找到合适的场景。

大模型技术用于金融核心业务环节时，其仍有应用的边界。比如，在前述网商银行实际应用中，没有用大模型的生成能力直接与客户交互。大模型绘制的产业链图谱会向风控系统提供客户识别、经营评分和画像，但最终小微经营者获得的贷款额度，仍然是风控系统多维度交叉验证的结果。

大模型解决的是更精确地识别问题，而不是最终的决策。

IDC中国副总裁兼首席分析师武连峰在接受记者采访时表示，大模型高价值应用，最关键的是找到合适的场景。它至少必须满足两个条件，一个是需要对海量数据进行分析，另一个是需要价值高昂的专家知识。

如何理解这种高昂的专家知识？网商银行行业金融一部汽车及医疗总经理杨希望告诉记者，比如要去识别中小微制造企业，哪些是给汽车产业链供货而不是其他产业链，基础数据之外，要知道企业需要具备哪类资质认证，满负荷运转时候的水电需求是什么水平，这些都需要前端业务人士通过大量调研形成“认知”。然后再去把这类专业的知识“投喂”给大模型，模型再按照计算机或者风控能够理解的语言去“跑出来”。

此外，由于大模型众所周知的“幻觉”问题，一旦判断失误将会造成严重损失或风险传导。因此，用于金融领域时必须要面对更高级别的限制。与图文生成一类大模型应用相比，金融大模型应用必须经过严格备案。

在方珂看来，风控系统是“神经中枢”，也是大模型的“守门员”。最终决定一家机构风控水平的，仍然是其客户量、数据与风控能力沉淀以及对某一类客群认知的积累。

目前，业内人士也普遍认同大模型确实有助于推动风控算法精确度的提升。萨摩耶云科技集团创始人林建明向记者表示，虽然大模型的爆款应用是图文交互类，但在风控环节应用时并不是直接调用整个通用大模型，而是拆解大模型的一些关键算法再去训练金融机构的内部数据——让大模型将海量的非结构性数据“结构化”，做好标签，确实有可能让风控得到更精确的变量。

他在近期出版的《AIGC重塑金融：AI大模型驱动的金融变革与实践》一书中也预判，未来AIGC的市场服务方可能会逐步走向类似SaaS付费模式——将模型能力结构化后封装在软件服务里，为金融机构按需定制。

不过记者注意到，目前市场上真正将大模型运用于风控等核心领域的金融机构凤毛麟角。除了技术投入实力外，一个重要因素是金融行业对于安全和风险底线要求更高，新技术的研发应用流程较慢。

比如在通用大模型的使用和部署上，由于金融机构数据不出域等限制，一定程度对其部署节奏有所影响，“从头开始自己做训练非常难，目前大部分机构会选择开源的基础大模型，再结合自己的数据做微调，训练成‘自研’模型，但这种模型也面临着一定瓶颈。”武连峰表示。

成本反而不是制约金融机构最重要的关键因素。据武连峰透露，通用大模型训练时需要大量高质量数据，如果金融机构运用通用大模型作为底层模型且不再进行太多预训练，直接做推理或者微调，数据量需求并不算大。“比如只要高质量地标注500条、1000条左右的数据，模型提升效果就会有5%到10%。”

他判断，2024年仍然是金融机构围绕大模型基础设施和解决方案的投入年，未来两到三年才会真正进入产出期。

大语言模型将改变数字产业生态

资料来源：清华五道口科创金融研究中心

数据要素流通新突破：金融科技助力安全与质量并重

本报记者 蒋牧云 何莎莎 上海 北京报道

数据要素正在成为新质生产力不可或缺的一部分，但我国数据要素市场尚在发展初期，在数据流通过程中仍有不少挑战。对此，多位业内人士告诉《中国经营报》记者，要充分激活数据价值，将不同数据融合是必不可少的，但在流通过程中，如何兼顾数据安全与数据质量是不少参与者关心的重点。

基于此，记者了解到，诸多科技力量如数据空间、隐私计算、人工智能等正在成为数据要素流通设施的基石。近日，海南省大数据管理局与蚂蚁数科签订合作框架协议，双方将在数据安全流通等方面进行合作。此前，深圳数据交易所与深圳数鑫科技有限公司（以下简称“数鑫科技”）等合作伙伴联合发布可信数据空间助力大模型语料合规高效流通的案例、“企业信用数据空间专区”等。深入数据交易平台与科技的融合与合作之中，科技力量究竟是如何助力数据要素流通的？

数据空间中的“转接器”

在“数据二十条”、《“数据要素×”三年行动计划（2024—2026年）（征求意见稿）》等一系列政策推动下，我国数据要素市场正在快速发展。不过，在发展过程中也暴露出需要进一步完善之处。比如，在数据流通过程中，要兼顾数据要素流通的安全与质量就成为业内关注的重点，而科技技术的赋能正成为关键点之一。国家数据局局长刘烈宏在不久前的公开讲话中提到，要加快建设安全可信的数据基础设施，发展数据空间、高速数据网，推动匿名化、联邦学习、多方安全计算等隐私计算和区块链技术应用，增强数据利用可信、可控、可计量能力，让公共数据“流得动”。

亿欧智库近期发布的《中国数据要素市场未来发展趋势》中也指出，数据流通设计需要多环节的技术协同，现阶段数据流通技术体系尚未完全成熟，存在数据泄露、越权滥用等数据安全问题，且不能完全满足实际场景下的落地应用需求。该报告预计，接下来区块链、隐私计算等跨技术路径、跨系统平台之间多元融合将成为趋势。加强敏感数据识别、数据脱敏技术、数据泄露防护技术等方面的突破也将以新技术、新模式牵引数据流通的新需求。

针对目前数据要素的流通情况，数鑫科技联合创始人、CTO廖炳才告诉记者，数据要素流通的过程中，涉及多个参与者，不论是数据的供方或需方，都拥有自身已建设好的数据存储设施、数据加工使用设施等。跨企业、跨组织、跨行业的参与者之间跨域流通使用数据，往往都不愿意去改造各自的数据存储或者加工使用设施。与此同时，数据流通往往不能直接搬运或者拷贝数据，而是需要同时兼顾供需方数据持有权、数据加工使用权的保障，让数据按需受控的参与社会化生产发挥出其作为生产要素的价值。由此，数据流通需要有去中心化、轻量化、普适性强、可解释性强的技术手段，确保数据能合规高效流通使用，从而更好赋能实体经济。

廖炳才告诉记者，数鑫科技从数据空间技术为突破口来解决前述痛点。数据空间可以将物理分布在不同参与方域内的数据对象，按需受控的进行虚拟化连接，是一张可以灵活弹性组网的数据流通网络，其核心机制在于数据对象的跨域虚拟化以及使用控制。

廖炳才具体介绍道，数据供需双方可以分别通过DPE（Data Provide Engine，数据提供引擎），DCE（Data Consume Engine，数据消费引擎），接入到数据空间流通网络中。DPE可以看成是供方的“数据转接器”，转接数据时既确保符合供方对数据的安全管控要求，又不丢失数据的业务含义，也就是保证数据质量。DCE可以看成是需方的“数据转接器”，可通过跨域联合计算沙盒进行多方数据跨域融合计算。在这一受控沙盒中，数据需求方可以通过事先协商好的合约进行加工与使用，合约可以具体到字段、算法等层面。在数据对象的跨域虚拟化以及使用控制机制下，每一条数据都以加密形式进入到内存中并完成联合计算。

除了流通机制之外，数据流通网络中各参与方数字身份、流通设备数字身份、数据流通使用策略合约关系等方面的信任认证能力，也是解决数据流通过程中信任问题的核心基础能力。廖炳才还提道，传统PKI技术只对于构建数据流通基础设施网络中的信任认证能力，存在明显的不足之处。因此数据空间基于Web3相关技术，解决数据流通基础设施网络中各个节点的设备数字身份以及相互之间的信任问题。

多技术融合应用

隐私计算、人工智能等技术，也在保障数据流通安全、合规等方面赋能。比如近日，海南省大数据管理局与蚂蚁数科签订合作框架协议。在签约仪式上，海南省大数据管理局副局长孙建明表示，双方将持续在数据开发利用、数据要素市场生态共建、数据要素应用场景建设、数据跨境流动等方面开展长期深入的合作。蚂蚁数科安全科技副总经理王黎强则介绍道，蚂蚁数科针对数据合规、数据安全、数据流转及应用等领域持续深耕，同时对区块链、隐私计算、人工智能等前沿技术进行研究创新，构建了一套完整的数据价值流通技术底座，可确保数据链路安全合规和顺畅流通，促进数据要素价值释放。

关于这一技术底座的具体情况，蚂蚁数科智能数据产品总经理李书博告诉记者，数据只有在应用场景中才能释放更大价值，而数据的安全流动与价值释放也依赖于一系列技术基础设施保障。具体而言，蚂蚁数科自主研发隐私增强型数据协作平台FAIR，融合区块链及隐私计算技术，能使原始数据在不出域的情况下，实现多节点之间高效可信地协同计算和隐私保护。

需要指出的是，在数字化与人工智能浪潮下，挖掘出更多新的行业场景与需求。在这一背景下，数据要素的传输基础设施，也需要不断完善。比如，数据规模不断提升的背景下，如何更加高效地进行传输就是业内关注的重点之一。

蚂蚁集团董事长兼CEO井贤栋就在近日的公开演讲中提到，AGI（通用人工智能）时代的到来，需要对数据孤岛化下的大规模数据流通进行破题。人工智能技术迅猛发展，也给隐私计算提出了更高要求，由于成本限制，隐私计算在高阶人工智能比如大模型应用上，还停留在实验室阶段。井贤栋表示，蚂蚁正在致力于探索下一代隐私计算技术，力求像明文计算一样高效易用。其中一个探索是推出可信密态计算（TECC，Trusted-Environment- based Cryptographic Computing），实现低成本、高安全、高性能和高可靠等维度的提升。

据悉，在成本上，在同等高安全等级下，可信密态计算将成本控制在明文分布式计算成本10倍以内（相比起，多方安全计算1000倍—100000倍），力求让隐私计算的性能、成本逼近明文计算；在安全性上，由于密码学+可信硬件双重安全保障，可达到攻防检验级高安全等级；在性能上，百万级参数CNN模型预测可以在亚秒级完成，亿条数据SQL分析10分钟完成。

廖炳才也向记者表示，人工智能的热潮使得更多AI体系的数据市场参与方出现，AI体系中的向量数据如何更加高效地实现跨域虚拟化及使用控制，将是数鑫科技下一步探索的方向。在向量数据结构中，数据是由数字化处理后的浮点排列组成的数组，在目前的传输机制中，会将这一数组重新表达为表格数据结构后，再进行下一步传输与融合。除了表达结构之外，向量数据对于沙盒的环境要求也有所不同，下一步，若能直接将向量数据进行跨域虚拟化及手控使用，则能进一步支撑AI体系的发展。