DOI 编码:10.19847/j.ISSN1008 - 2646.2025.01.007

生成式 AI 在金融领域的应用、风险与监管建议

徐 超
(中国社会科学院 信息情报研究院,北京 100732)

摘 要:近年来,以信息技术为基础,新一轮科技革命方兴未艾,人工智能尤其是生成式人工智能广泛应用于人类社会生活不同领域,促进相关产业链快速发展。现有研究成果显示,人工智能在金融领域的广泛应用,大大提升经济效益与效率,但也蕴含不少风险。对此,境外国家和地区、国际监管机构为因应当前金融科技发展的需求,妥善处理金融科技创新与安全之间的关系,积极主动制定监管制度,提升监管机构的能力。我国作为人工智能技术和金融业发展的大国,亦应多措并举制定人工智能监管制度,这不仅有助于控制相关领域潜在风险,而且有助于争夺相关领域的制度性话语权。

关键词:生成式人工智能;法律风险;金融科技;金融监管
中图分类号:F832.3;TP18 **文献标志码**:A **文章编号**:1008 - 2646(2025)01 - 0040 - 08

引言

近年来,产业界不断加大 AI(Artificial Intelligence,即人工智能)技术研发与应用的力度,2022 年 11 月 Open AI 推出的 ChatGPT 标志着生成式人工智能(Generative Artificial Intelligence,简称生成式 AI)发展迈进新阶段。生成式 AI 利用机器学习和深度学习技术,学习大量的数据,生成与原始数据相似的新数据。通过此等创新,生成式 AI 具备初步的判断和推理能力,并逐渐被应用到诸多领域,例如,撰写语言作品(诸如文章、程序代码、诗词、翻译、戏剧脚本、分析报告、电子邮件、社交媒体内容、结构化的数据分析等)、图像创作(不同风格的绘图作品、串联图像成为影片等),以及合成语音和制定设计方案等。值得说明的是,2024 年 1 月 Open AI 推出的文生视频模型 Sora,其生成的短视频可以与人类的拍摄效果媲美[①]。这些借助生成式 AI 创作产出的应用,更带动相关行业积极寻求商业应用情景与模式。

生成式 AI 应用先进算法和机器学习等技术能够处理分析大量数据、自动执行任务并准确判断客户的需求的特点,与金融业的特点和发展需要正相契合,金融领域也是生成式 AI 应用的理想场景[②]。金

收稿日期:2024 - 09 - 13
作者简介:徐超,中国社会科学院信息情报研究院副研究员。
基金项目:河北省高等学校人文社科研究项目“数字经济下无形资产跨境交易税收法律问题研究”(SQ2023177);中国社会科学院“登峰战略”资助计划优势学科“意识形态安全学”
本文引文格式:徐超. 生成式 AI 在金融领域的应用、风险与监管建议[J]. 南京理工大学学报(社会科学版),2025(1):40 - 47,61.

① 杜知航、关聪:《Sora 再进阶》,《财新周刊》2024 年第 11 期。

② 虽然银行与金融机构多年来一直在各种前、中、后台的服务中,尝试部署使用数据分析与类似专家系统,或决策支援系统等类型的人工智能应用,包括管理信用风险和欺诈;然而,金融服务中的生成式 AI 与之前各类型的智能型应用方法相比,俨然发生了重大变化。参见中国人民银行武汉分行办公室课题组韩飚、胡德:《人工智能在金融领域的应用及应对》,《武汉金融》2016 年第 7 期。

融业已经存在超过500年历史,积累了大量的个人、企业数据,这对人工智能金融模型模拟训练至关重要。同时,金融业导入AI技术应用后,无疑将提升金融机构的数据搜集能力,增强数据分析洞察力、提高营运效率、加强风险管理和控制,以及更有效地打击欺诈和洗钱。简而言之,金融业应用AI,对内可协助金融机构简化和优化从信贷决策到量化交易和金融风险管理的流程,使金融机构能更好地管理风险、提高效率;对外可更深入了解客户的需求,并提供更多信息及开发出更好的产品,为客户提供更好的服务①。正因为如此,AI在预测市场趋势、提升交易效率、维持监管合规和风险管理等方面都有巨大的发展前景②。麦肯锡公司在2023年发布的研究报告认为,生成式AI大模型为金融行业尤其是全球银行业带来巨大的经济价值。据其估算,应用生成式AI大模型每年为企业端带来的经济价值约为2.6~4.4万亿美元,其中全球银行业使用大模型可使其每年营业额提高2.8%~47%,这一比例高于全球制造业、零售业、旅游业、运输业和物流业等其他行业。由此可见,生成式AI有助于降低金融业的成本,提高收益,赋能金融科技创新与发展。

然而,生成式AI在金融领域应用目前还处于早期阶段,无论从支持其算法运行的数据与参数,或是从其应用的逻辑和范畴等维度来看,都仍存在一定的不确定性、偏差与风险。目前,包括金融机构在内的不少从业人员在应用生成式AI时,并不知悉其产出的生成逻辑,加上监管机构尚未系统掌握此类技术对金融消费与市场的潜在影响和冲击,这让未来的金融运行隐含一定的风险与挑战③。大型语言模型的采用引发各领域行业观察家、学术界、监管机构及广大公众的讨论与不同观点,敦促各国政府加快人工智能监理与治理的声音不断高涨,以因应生成式AI模型的广泛使用及其带来的相关风险④。对此,如何让生成式AI更好服务金融业健康、可持续发展,妥善解决生成式AI的伦理挑战、隐私、信任及安全方面的风险⑤,是金融机构和监管部门应当积极回应的新问题、新情况与新挑战。

一、生成式AI在金融领域的应用与实例

当前,大型金融机构正在推动生成式AI在金融领域的应用与发展,这必将推动金融科技的跨越式发展,具体应用情景主要有以下五个方面。

1. 智能营销

金融机构往往需要投资精准且定制化的营销方案,避免盲目营销带来的成本开支,亦需要解决客户因接听泛滥营销电话而滋生的不满情绪,进而在竞争中脱颖而出并赢得新客户信任。但这需要大量且深入的客户分析,可能既昂贵又耗时。生成式AI可以通过分析客户偏好与在线行为,由人工智能预判,将潜在客户分成不同偏好类型的群组,而非仅凭经验划分的客户类型。银行在此基础上可根据市场状况和趋势,为不同群体量身定制营销方案。此外,银行也可以使用生成式AI帮助创建适合的营销素材,并追踪转换率及客户满意度,再通过A/B测试的执行以测试其效度。除了提供类似人类的全天候客服支援,包括回答客户询问、个人信息更新外,生成式AI还可进一步分析客户数据,做出个性化的产品推荐。例如,当客户提供利率、首付金额、信用评分等相关详细信息后,生成式AI可快速准确地提供其可负担的抵押贷款,甚至根据客户的消费习惯、财务目标及生活方式推荐信用卡,以这种方式交叉销售其他金融商品。如中国工商银行试点RPA技术赋能智能营销⑥;腾讯云推出腾讯金融智能营销平台;兴业

---

① Bank of England, "DP5/22 - Artificial Intelligence and Machine Learning", 英格兰银行官方网站, https://www.bankofengland.co.uk/prudential-regulation/publication/2022/october/artificial-intelligence, 2024年5月1日访问。

② 叶军、卢宗亚、郦丽娟:《人工智能在金融领域的应用探讨》,《金融纵横》2023年第10期。

③ Yanqing Duan, John S. Edwards, Yogesh K. Dwivedi, "Artificial Intelligence for Decision Making in the Era of Big Data—evolution, Challenges and Research Agenda", International Journal of Information Management, 48, 2019.

④ 冯子轩:《生成式人工智能应用的伦理立场与治理之道:以ChatGPT为例》,《华东政法大学学报》2024年第1期。

⑤ 张欣:《生成式人工智能的算法治理挑战与治理型监管》,《现代法学》2023年第3期。

⑥ 2021年下半年开始,中国工商银行在部分分行试点以RPA技术赋能智能营销试点,推出手机银行客户旅程运营机器人、重点客群数字化运营机器人等,在大零售领域探索RPA技术的营销创新,推动手机银行客户自动化旅程营销,并协同客户经理、叫号机等多种渠道,实现对低效代发工资客户、信用卡活跃客户等重点客群的自动化运营。

银行智能语音客服上线多年,识别率超90%;彭博社推出语言大模型 Bloomberg GPT;恒生电子发布金融大模型 Light GPT 及多项应用产品。

2. 智能投顾

传统投资交易(如股票交易)大多依靠投资者、分析师的自身经验与判断,较容易受情绪与主观影响,很难做到客观、果断地操作。智能投顾是利用文本与数据挖掘、文档摘要分析各大财经网站、舆论的文字资料,找出其中的联系,并通过机器学习将数据输入计算机自动分析,找出潜在的规则,再据以建立智能投资工具或自动化操作模型,以分析价格涨跌幅程度、预测价格、分析有潜力的金融投资组合,为投资者提供个性化、智能化的投资建议和投资组合管理服务的金融服务模式①。近期有研究认为,将 ChatGPT 融合在投资模型中,可以预测股市走势、投资回报率等,效果远远超过对冲基金使用的传统股市情绪分析模型。生成式 AI 可以通过分析投资者的风险偏好、财务目标和市场情况等信息,生成适合投资者需求的投资组合,并识别潜在风险与金融风险,提供早期预警信号,并根据市场波动实时调整投资策略。

生成式 AI 在智能投顾领域的应用日益丰富②。2023 年 5 月,Open AI 公司推出的 GPT-4 Plugins 插件 Portfolio Pilot 是一款基于 ChatGPT 的人工智能投资软件,用户可以将自己的投资组合复制粘贴到该工具中进行分析,搜索市场信息,获得投资建议,并向它提问。另一款插件 AITickerChat 是一款创新的聊天机器人插件,旨在通过利用美国证券交易委员会(SEC)的文件和收益电话记录,帮助用户搜索与股票市场相关的信息。AITickerChat 为投资者、分析师和金融专业人士提供了不少有价值的参考建议。总之,在智能投顾领域,生成式 AI 可以帮助客户更快地处理信息,并就资产服务做出决策,可实现智能化的资产配置与主动式的投资组合管理。

3. 客户身份识别

当下,基于人脸识别的互联网在线识别和认证模式,覆盖了不同行业和领域,属于主要的个人身份认证模式。金融机构利用生物辨识(Biometric)等技术,以声音、言语、脸型、指纹、静脉、虹膜等生物特征,作为对使用者进行金融交易、投资交易前身份认证的主要方式。通过生物辨识技术装置如手机、平板及计算机等进行辨识感应,客户可直接进行线上开户、在线交易,无需再到实体机构进行申请及身份确认等操作,可大幅降低验证时间与开户等成本。生成式 AI 视频技术通过文本创建视频的功能,能够实现如金融系统"人脸+指令+行为"的视频识别/认证流程,也能覆盖大部分的人脸识别模式。例如,新希望金融科技公司推出的"天镜 AI 视频面签解决方案"将传统视频面签的基础功能与数字人、AI 反欺诈、远程银行等能力无缝结合,通过 AI 技术大幅提高视频面签的风控能力和服务半径,在客户身份认证、贷款调查及存证、信贷风险防控、中介风险识别、交易真实性确认、信用卡面签审核、远程智能客服等场景都有成熟的应用效果。

4. 信用评级

信用评级(或称信用评分、信用评价)主要针对受评对象进行信用、贷款违约风险大小评估,以往多由某些专门的信用评估机构进行,评估机构针对受评对象的金融状况、历史资料进行调查、分析,从而根据受评对象的金融信用状况得出整体信用报告。传统的违约概率模型在很大程度上依赖于逻辑回归(logistic regression)。该模型相对容易理解,数十年来亦一直是市场最佳实践,但是面对当前大量高频金融数据,传统模型预测金融市场潜在机遇与风险的能力大幅减弱。同时,传统的金融市场预警模型通常基于大量未经实证的指标,严重依赖专家主观判断,AI 则擅长使用大量的高速数据来进行信用违约预警。凭借超强的计算能力,AI 算法能通过不同来源的数据指标提高预警信号的准确性,自然语言处理技术(NLP)的出现也使分析文本信息变得可能。从社交媒体帖子到传统报刊的财经新闻,只要是书面媒体,NLP 都可以将其用于信用分析中,取代以往需由人工执行的繁琐工作。

---

① 姜海燕、吴长凤:《智能投顾的发展现状及监管建议》,《证券市场导报》2016 年第 12 期。

② 世界著名的咨询公司德勤在其发布的研究报告《金融 AI 赋能传统金融机构的应用与展望》中把智能投顾定义为:利用人工智能技术,为投资者提供个性化的、智能化的投资建议和投资组合管理服务的金融服务模式。

5. 财务分析与预测

通过学习历史金融数据，生成式 AI 应用可捕捉数据中的复杂模式和关系，对未来趋势、资产价格及经济指标进行预测分析。模型若经适当微调，甚至可通过模拟市场状况、总体经济因素及其他变量来生成不同场景，从而为投资者提供有价值的投资机遇。此外，通过利用对人类语言模式的理解及产生连贯的、上下文相关的回应的能力，生成式 AI 可以为投资者面临的财务问题提供较为详细的投资建议。这些模型可以在大型金融数据平台上进行训练，并借助数据分析与计算回答以下问题：会计原则、财务比率、库存分析、监管合规性。生成式 AI 的另一金融应用为优化投资组合。通过分析历史财务数据并产生不同投资场景，生成式 AI 模型可以帮助资产管理者及投资者确定最佳的资产与财富管理方案，同时考虑风险承受能力、预期报酬等因素，金融专业人士能够调整投资策略，优化风险调整后的报酬，就投资组合做出更明智的决策。简而言之，生成式 AI 可以回顾过去，帮助金融机构和投资者对未来投资做出更好的财务决策，并创建综合数据以对风险出现进行稳健分析。

二、生成式 AI 在金融领域应用的潜在风险与挑战

强调简便可用且具有创作能力的生成式 AI，虽然看似为社会经济活动带来突破式的数字化转型契机，但在实际应用上却也存在不少隐忧。从广义范围来看，生成式 AI 的风险与挑战，大致可以分为以下两大类型：一是恶意使用造成的危害，例如创建未经同意的色情内容，或是自动扩散仇恨言论、产生骚扰与虚假讯息，以及利用其生成内容进行诈骗等行为；二是商业使用造成的危害，采用既有生成式 AI 的公司可能无法充分理解 AI 系统的底层逻辑，进而会产生错误和意外行为，甚至将错误的结果提供给顾客（例如提供错误的法律适用信息），让顾客处于相当高的误判风险之中①。金融稳定理事会（FSB）分析金融机构应用 AI 的风险，主要包括黑箱决策风险、隐私外泄风险、算法歧视行为、市场关联性及脆弱性提高、生成内容机器幻觉造成的可靠性风险等②。

1. 黑箱决策风险

模型的可解释性对于金融监管机构来说至关重要。机器学习在提升模型准确性的同时，通常亦会令大模型变得难以解释③。机器学习之所以往往被称为“黑匣子”，主要原因在于很难以直观的方式解释其模型输入和输出之间的关系④。正因如此，业内人士和监管机构经常质疑机器学习方法。在投入大量时间和资源训练 AI 算法之后，模型结果仍然可能难以解释，甚至导致决策失误。如果模型没有适当、合理的解释，使用 AI 模型可能会带来比以往更大的模型风险。具体到生成式人工智能在金融领域的应用，同样缺乏透明度和可解释性，加之金融系统的内生风险，在大模型应用的外生风险影响下可能会放大为系统性金融风险⑤。有鉴于此，近年来，不同行业加大协同合作的力度，旨在克服机器学习的潜在不足和不透明性等突出问题，助推机器学习模型的推广与应用。

2. 隐私外泄风险

金融应用场景的隐私风险，顾名思义来自于对客户信息的保护不周或不当使用，进而造成对客户隐私或权利的侵害⑥。相关的风险等级类型，根据其影响程度而有所不同。首先是客户提交的资料遭机构不当应用，这部分最常出现在同集团体系的交叉销售，或是与商业伙伴进行产品或服务的共同营销

---

① Alex Engler, “Early Thoughts on Regulating Generative AI Like ChatGPT”, Brookings, https:// www. brookings. edu/blog/techtank/2023/02/21/earlythoughts - on - regulating - generative - ai - like - chatgpt/amp/, 2024 年 5 月 1 日访问。

② FSB, “Artificial Intelligence and Machine Learning in Financial Services - Market Developments and Financial Stability Implications”, 金融稳定理事会官网, https://www. fsb. org/wp - content/uploads/P011117. pdf, 2024 年 5 月 1 日访问。

③ Yavar Bathaee, “The Artificial Intelligence Black Box and the Failure of Intent and Causation”, Harvard Journal of Law and Technology, 31, 2018.

④ William Magnuson, “Artificial Financial Intelligence”, Harvard Business Law Review, 10, 2020.

⑤ 张欣、买尔旦·买买提：《ChatGPT 在金融领域的法律风险和应对》，《中国银行业》2023 年第 6 期。

⑥ 倪蕴帷：《数字经济架构中的动态隐私理论及其应用》，《南京理工大学学报（社会科学版）》2022 年第 6 期。

上。其次是信息外泄问题,金融机构因业务需要,搜集消费者诸多隐私信息,因信息保护或信息安全机制不完善,导致遭到黑客侵入而发生信息流出①②。与人工智能应用关联度最高的则是隐私信息的不当应用,导致顾客在取得金融服务时,遭到过度推销或不公平待遇,比如顾客因多次延迟缴纳贷款或信用卡采用最低还款时,便持续收到小贷公司或其他非原机构的借款业务推销;或可能有顾客购买保险时,机构通过不当取得的医疗记录或其他隐私信息,而拒绝提供服务。

3. 算法歧视风险

人工智能应用的算法歧视风险,其成因来自训练数据选择的偏狭、缺漏,导致所建构的算法模型有偏见、歧视的问题,因而造成消费者在信用评分、贷款审批或利率设定方面遭到不公平对待③。生成式AI 自我学习的过程中,大量参考过去的数据后归纳分析以做出决策,但过去的数据都是人类做出来的,本就带有人类的偏见在其中,AI 可能归纳分析出了以往人类都没有意识到的偏见,并以此做出决策,再以此有偏见的决策作为下次决策的参考资料,循环不断地放大偏见。国内外研究成果显示,当金融从业者不当应用部分特定数据,诸如种族、宗教、性别等进行信用评分,加上目前人工智能算法不透明的问题,就可能产生金融应用场景的算法歧视④。当生成式人工智能基于历史数据进行信贷决策时,若这些数据反映了现有的社会经济结构偏见,那么生成式人工智能可能会复制,甚至加剧这些偏见。这不仅可能导致对特定群体的不公正待遇,如高风险评级或拒绝服务,还可能加深现有的社会和经济不平等。

4. 市场关联性及脆弱性上升风险

金融科技不断发展以及信息网络的普及与应用,让全球金融交易市场呈现出高度的全球化、自动化与实时性等基本特征。毋庸讳言,算法交易是交易市场自动化的典型代表,在以往人工智能技术仍难以商业化的时期,算法交易主要基于规则进行,其策略执行过程尚可推测与掌握。首先,随着近年深度学习等人工智能技术突飞猛进,金融机构越来越难以了解其交易生成策略,再加上训练应用的资料可能造成模型趋同问题,在不少金融机构依赖相同资料来源及算法进行金融交易决策时,无疑将提高金融机构与金融市场间的关联性,以及金融机构若同时执行高频且大量的交易,将提高市场波动度及脆弱性,进而影响金融市场稳定与发展。其次,当众多金融市场从业者在信用评分或交易活动等领域使用 AI 技术进行评估与交易时,一旦市场出现压力,可能因羊群效应而扩大冲击程度,比如同时出现有价证券或资产减损,相关风险可能进一步影响金融稳定⑤。最后,若金融机构应用 AI 于资本管理以降低资本计提,将使金融机构资本缓冲减少,应用 AI 提高脆弱度将导致流动性缓冲降低、杠杆率提高及期限转换加速。

5. 可靠性风险

人工智能幻觉(AI Hallucination)是指模型生成的不正确或具有误导性的结果。有关模型有可能不理解数据之间的真实关系,导致得出看似真实,但实际上不正确、不完整或缺乏重要信息或与语境相关的结果⑥。因此,金融机构应提供足够的人为监督,包括人机循环(put the human in the loop)方法,确保数据的合规与准确以及人工智能得以正确使用,否则可能导致糟糕的财务决策。根据经验,不应让生成式 AI 在贷款批准与其他影响客户的重要决策中拥有最终决定权,应加入人类参与,针对投资开发与引进此应用对金融机构造成的可能冲击进行伦理审查。金融机构需要审慎使用人工智能及对人工智能系

---

① 廖高可、李庭辉:《人工智能在金融领域的应用研究进展》,《经济学动态》2023 年第 3 期。

② Y. L. Liu, W. Yan, B. Hu, "Resistance to Facial Recognition Payment in China: The Influence of Privacy - related Factors", Telecommunications Policy, 45(5), 2021.

③ 姜利、张景胜:《人工智能技术在金融领域的应用、影响及展望》,《黑龙江金融》2020 年第 10 期。

④ L. Cao, Q. Yang, P. S. Yu, "Data Science and AI in FinTech: An Overview", International Journal of Data Science and Analytics, 12, 2021.

⑤ 于品显:《系统性金融风险的界定及传播机制》,《南方金融》2019 年第 6 期。

⑥ Buckley, Ross P., Dirk A. Zetzsche, "Regulating Artificial Intelligence in Finance: Putting the Human in the Loop", Sydney Law Review, 43(1), 2021.

统进行持续监管,在开发生成式 AI 相关应用时,应当在数据清理、产品参数设定、整合应用等方面,施予严格的评估、验证及纠正措施,并在上线前进行使用者测试,以解决人工智能幻觉问题。

三、生成式 AI 在金融领域应用的国际监管进展

目前,不少国家及国际组织认识到,生成式 AI 在金融领域应用时往往存在潜在风险与挑战,因此,如何适度监管以维护消费者权益及金融稳定备受国际组织及各国金融监管机关的重视。2019 年经济合作与发展组织(OECD)率先提出《AI 原则》(*Principles on Artificial Intelligence*)①,列出"包容性成长、可持续发展与福祉""以人为本的价值观与公平""透明度与可解释性""稳健性与安全性""问责性"等五项重要原则,并被 G20 沿用。其后,国际金融组织陆续发布金融机构应用 AI 监管建议,欧盟已通过《人工智能法》,主要国家亦参考国际组织建议陆续提出 AI 的监管原则或指南。

1. 国际金融组织提出金融机构应用 AI 的原则或建议

金融稳定委员会(FSB)于 2017 年发布《人工智能和机器学习在金融服务业的应用——市场发展与对金融稳定的影响》②,说明 AI 与机器学习可协助金融机构更有效率地处理信息,亦可通过规制科技(RegTech)及监管科技(SupTech)确保金融机构的守法性并增强监管效能。同时,FSB 也提示了相关风险。国际清算银行(BIS)旗下的金融稳定学院(FSI)于 2021 年发布《人类控制 AI——对金融业新兴的监管期待》③一文,希望金融监管机构按照透明度、可信赖性与稳健性、问责、公平与伦理等 4 个原则制定相关监管措施,同时亦希望在推出政策措施时,依照比例原则来处理可能的挑战。

国际证券管理机构组织(IOSCO)于 2021 年发布《市场中介机构及资产管理机构应用 AI 及机器学习指南》④,提出 6 项金融监管机关可以采取的措施,建议金融机构做好以下工作:(1)有适当的治理、控制与监督架构;(2)对 AI 与机器学习的发展、测试、使用及表现持续监测;(3)人员需有足够知识技能及经验,以实施、监督及挑战 AI 与机器学习产出的结果;(4)了解本身对提供 AI 与机器学习服务的第三方机构的依赖性,并建立良好管理与监督机制;(5)对投资人、主管机关及相关利害关系人提供妥适的透明度与信息披露;(6)有适当的控制机制,以确保数据及 AI 与机器学习的表现能将偏见最小化。

美国⑤、新加坡⑥、韩国⑦等国金融监管机构针对金融业应用 AI,亦提出相关指引及原则,重点多集中在公平与道德、透明、问责、消费者权益与隐私权保护、安全与稳健等。

2. 欧盟确立以风险为基础的监管制度

2021 年 4 月 21 日,欧盟委员会提出了《人工智能法案》(*Artificial Intelligence Act*,简称 AIA)草案,这

---

① OECD,"Principles on Artificial Intelligence",经济合作与发展组织官网,https://www.oecd.org/en/topics/ai-principles.html,2024 年 5 月 1 日访问。

② FSB,"Artificial Intelligence and Machine Learning in Financial Services—Market Developments and Financial Stability Implications",金融稳定理事会(FSB)官网,https://www.fsb.org/wp-content/uploads/P011117.pdf,2024 年 5 月 1 日访问。

③ BIS,"HumanKeeping AI in Check—emerging Regulatory Expectations in the Financial Sector",国际清算银行官网,https://www.bis.org/fsi/publ/insights35.pdf,2024 年 5 月 1 日访问。

④ IOSCO,"The Use of Artificial Intelligence and Machine Learning by Market Intermediaries and Asset Managers",国际证券监管理事会官网,https://www.iosco.org/library/pubdocs/pdf/IOSCOPD684.pdf,2024 年 5 月 1 日访问。

⑤ National Association of Insurance Commissioners (NAIC),"Principles on Artificial Intelligence (AI)",美国保险监管官协会官网,https://content.naic.org/sites/default/files/inlinefiles/AI%20principles%20as%20Adopted%20by%20the%20TF_0807.pdf,2024 年 5 月 1 日访问。

⑥ Monetary Authority of Singapore,"Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector",新加坡金融监管局官网,https://www.mas.gov.sg/-/media/mas/resource/news_room/press_releases/2018/annex-asummary-of-the-feat-principles.pdf,2024 年 5 月 1 日访问。

⑦ Financial Services Commission,"FSC Launches Working Group to Draw Up Plans for Promoting AI in Financial Services",金融服务委员会官网,https://www.fsc.go.kr/eng/pr010101/76209,2024 年 5 月 1 日访问。

是全球第一个针对人工智能科技应用所衍生的可能风险而制定的管理规则,其目的在于确保投入欧盟市场使用的人工智能科技应用都是安全的,且符合欧盟的基本权利及其价值观,并促进人工智能生态系统的发展,以奠基欧盟在数字时代的竞争优势。2024 年 3 月法案正式生效。欧盟意图通过《人工智能法案》建立人工智能监管全球标准,进而使欧洲在国际人工智能竞争中取得优势地位①。

《人工智能法案》采用了以风险为基础的分级监管方法(risk - based approach),根据对健康、安全和自然人基本权利的影响,将 AI 系统应用产生的潜在风险区分为 4 个级别,具体设定 AI 系统所应受到的监管强度,对于违反的企业将被处以 3 500 万欧元或全球年营收 7% 的罚款(以金额较高者为准)②。(1)不可接受的风险等级:明确规定禁止通过 AI 系统从事特定活动,例如使用敏感特征(政治、宗教信仰、性倾向、种族)的生物识别分类系统,依据社会行为或个人特征进行社会评分,以及通过网络或监视器搜集不特定对象的脸部画面,以建立或扩充脸部辨识数据库等;(2)高风险等级:用于生物辨识、关键基础设施管理、教育与职业训练及执法等行为,允许相关主体在履行事前评估等义务后投放市场或使用,同时要求事中、事后持续监测;(3)有限风险等级:如聊天机器人、AI 换脸、GenAI 及深度伪造等,无须取得特殊牌照、认证或履行报告、记录等义务,但应遵循透明度原则,允许适当可追溯性和可解释性;(4)低风险或最低风险,相应主体可依据自由意志部署和使用。具体到生成式人工智能而言,由于其没有特定目的并且可应用于不同场景,故不能基于一般方式或操作模式对其进行风险分级,而应当按照开发或使用生成式人工智能的预期目的和具体应用领域设定监管规范。

3. 英国对人工智能应用于金融领域的监管方向

近几年,英国金融行为监管总署(FCA)意识到人工智能以及数据分析技术应用对于金融领域的影响,并且开展了相关研究,期望能够协助消费者在技术创新中获益,同时也讨论如何将这些创新技术应用于金融监管。英格兰银行(BoE)和 FCA 于 2019 年 10 月共同发布的报告《机器学习在英国金融服务上的应用》强调,AI 发展可从根本上改变企业提供及消费者使用金融服务的方式③。2020 年 10 月,英格兰银行与金融行为监管总署公布成立“金融服务人工智能公私论坛”,该论坛的目的主要是促进公私部门对话,进一步了解人工智能在金融服务中的使用与影响,协助制定金融业采用 AI 或机器学习的工具、准则、规范或产业最佳实务等。2023 年 3 月,英国新设的科学创新与技术部公布了一份政策白皮书——《支持创新的人工智能监管方法》,目的是为监管 AI 提出新方法,建立大众对 AI 科技的信任,并使企业在有法可依的情况下更容易进行相关创新发展。

总体来说,英国对于金融科技的兴起主要持“等待与观望”的态度,监管机构采取“等待”的态度,并通过“观望”新兴技术在金融领域应用找到最需要进行监管的问题与顺序。

四、生成式 AI 在中国金融领域应用的监管建议

1. 转变监管认知理念

在认知理念上,金融监管部门和金融机构都应充分重视生成式 AI 的安全问题。从宏观方面看,人工智能模型与算法的韧性不足,存在易受攻击的缺陷。因此,用于人工智能训练与应用的资料与数据,若缺乏治理将难以达到可信赖人工智能所需的算法稳健安全性、隐私保护、可解释性、公平性与安全性④。目前,全球人工智能法规在这方面都有规定,我国中央网络安全和信息化委员会办公室对进入意识形态领域的模型已要求进行事先风险评估。国外有的大型银行已将 AI 模型风险纳入整体风险管理

① 王天凡:《人工智能监管的路径选择——欧盟〈人工智能法〉的范式、争议及影响》,《欧洲研究》2024 年第 3 期。

② Matthew R. Gaske, “Regulation Priorities for Artificial Intelligence Foundation Models”, Vanderbilt Journal of Entertainment and Technology,26(1),2023.

③ Ross P. Buckley, Dirk A. Zetzsche, Douglas W. Arner, et al., “Regulating Artificial Intelligence in Finance: Putting the Human in the Loop”, Sydney Law Journal, 43,2021.

④ 美国政府发布的《安全与可信赖人工智能行政命令第 14100 号》(*Safety , Secure and Trustworthy Artificial Intelligence Executive Oder*)与欧盟所公布的《人工智能法案》,这两部国际最重要的人工智能安全治理法规,均对上述问题有所着墨。

框架,并成立了 AI 模型管理委员会,建立了专门的管理平台、流程和规范。我国金融机构也应对 AI 大模型相关风险实行分级分类管理,对模型数据参数进行定期评估和交叉验证,并使用压力测试,在各种情景下进行模拟校验,及时披露模型决策机理、运行逻辑和潜在风险,防范算法歧视,提升算法的可解释性、透明性与公平性。从微观方面看,需要保护人工智能系统的生命周期内资产的保密性、完整性、可用性免受攻击及可追溯性。此部分之应用包含使用人工智能用以支援传统网络安全威胁分析与防御技术,或利用人工智能作为工具/手段建立先进的网络安全方法(例如使用生成式 AI 开发更有效的安全代码或控制手段),并促进执法机构与其他公共当局更有效地因应网络犯罪。

2. 完善监管体制机制建设

制度设计上,监管部门应主导 AI 发展及相关因应政策的顶层设计。一是培育具有生成式 AI 技术以及金融专业知识的复合型人才。生成式 AI 技术是一种跨领域的技术,应用层面广泛,然而,具有生成式 AI 技术的人才往往没有受过金融相关专业的教育,又或者以金融为背景的人通常不具有相关技术,因此,面对生成式 AI 技术的浪潮冲击,监管部门需要拟定计划,培育具有技术和金融专业知识的复合型人才。同时,应加强公众在生成式 AI 技术方面的风险防范意识和风险识别水平,保障其自身权益。二是监管部门需考量 AI 发展对社会隐私权及道德所造成的影响,同时评估对社会旧有文化的冲击,并视整体发展拟定对应政策因应、隐私监管、数据治理及制定道德准则;必要时亦须纳入适当的监管及风控,以达到 AI 发展与维护产业经营、经济发展及社会秩序均衡局面。对此,监管部门应推动制定生成式 AI 技术的政策法规,加强生成式 AI 技术相关的法律及道德约束。结合《生成式人工智能服务管理暂行办法》《互联网信息服务深度合成管理规定》等 AI 安全相关规定,不断细化健全包含道德伦理、人身安全、个人隐私保护、算法、知识产权等多方面的生成式 AI 技术、应用及服务的安全规范、规定及细则。

3. 有序推进监管立法

在立法制定上,可以借鉴欧盟以 AI 风险管理为核心的监管立法模式。AI 系统具有复杂性、不透明性、不可预测性、自主性,涉及复杂且专业的领域,无论是政府机构进行监管或是企业合规都会产生相当大的成本①,如何在兼顾 AI 产业创新发展,与国家安全、社会安全的前提下,制定相关法律或监管措施具有相当难度。欧盟及美国对于 AI 监管模式均以风险管理为核心,有一定参考价值。欧盟《人工智能法案》采用风险分级方式,针对不同风险的 AI 系统赋予不同程度义务,一定程度上可以平衡 AI 产业发展与风险治理之间的平衡。本文建议,我国立法机构可以考虑从 AI 设计、研发、部署及应用各阶段,参考境外法律制度评估 AI 系统应用或输出结果对个人或群体可能造成的风险,规划预先采取风险管理措施,并授权相关机关制定风险辨识与管理标准,为 AI 技术日后发展保留监管弹性。唯应注意的是,《人工智能法案》法律框架虽然旨在确保 AI 应用的公共安全性,提高使用者的信赖及研发者或供应商的可靠性,但亦有学者认为《人工智能法案》各个风险级别的区分标准未必明确,且由于 AI 系统应用伴随而生的风险大抵处于持续且快速变动状态,《人工智能法案》对此似亦欠缺合适的风险调整机制②。我国在借鉴欧盟立法时应做出相应调整。

生成式 AI 已在金融领域得到广泛应用,使金融机构获得更强大的技术能力,推动了金融业的快速转型升级。然而,生成式 AI 技术仍处于早期发展阶段,我们不能不重视其应用可能带来的潜在风险,例如 AI 在决策中的偏见、AI 幻觉以及个人隐私的侵犯等问题。为确保 AI 的应用符合社会和法律规范,减少潜在的风险与负面影响,并为社会带来最大利益,各国政府在促进 AI 发展的同时,也愈发重视 AI 的治理及监管议题。对我国金融机构而言,要发挥生成式 AI 的价值,最大的挑战是在创新、安全与合法之间寻找平衡,其中关键的是确保训练数据的安全性以及生成信息的准确性。同时,社会各界也需高度重视算法的公平性与解释性,建立健全的规范框架,以防止算法滥用,真正造福金融行业与全社会。

(责任编辑、校对:臧莉娟)

(下转第 61 页)

① 沃尔夫冈·多伊普勒:《〈欧盟人工智能法案〉的背景、主要内容与评价——兼论该法案对劳动法的影响》,《环球法律评论》2024 年第 3 期。

② 皮勇:《欧盟〈人工智能法〉中的风险防控机制及对我国的镜鉴》,《比较法研究》2024 年第 4 期。

为外贸中小微企业开发用于支付进口产品和进口关税的灵活金融产品和服务。与此同时,要更加主动发挥数字普惠金融的促消费作用,持续推动消费升级从而带动进口增长。例如,应持续优化居民跨境消费和移动支付的便捷性和安全性,在满足居民多样化、个性化、高级化消费需求的同时,也要注意完善农村及偏远地区的数字金融基础设施建设,释放农村及偏远地区的消费活力。

(责任编辑、校对:臧莉娟)

Digital Financial Inclusions, Enterprise Sizes and Enterprises' Imports

JIN Yousen[1,2], FANG Hang[1], TANG Linlin[1]

(1. *School of Economics and Management, Nanjing University of Science and Technology, Nanjing, Jiangsu*, 210094;
2. *Research Center for International Economy & Trade, Nanjing University of Science and Technology, Nanjing, Jiangsu*, 210094)

Abstract: The Report in the 20th National Congress of the Communist Party of China (CPC) emphasizes the digital economy development acceleration and the high-level openness promotion. Utilizing a dataset that integrates Peking University's Digital Financial Inclusion Index, Chinese Industrial Enterprise Database, and Chinese Customs Import and Export Database, this study examines the role of digital financial inclusion in facilitating enterprise import expansion and its working mechanism. Research findings indicate that digital financial inclusion substantially enhances enterprises' import scales, with a more pronounced effect on small-sized enterprises. The working mechanism mainly includes financial constraint alleviation and consumption stimulation, both of which are more significant for smaller enterprises. Additionally, the import expansion effect of digital financial inclusion is more pronounced in non-state-owned enterprises and in China's Eastern regions. These findings enrich relevant economic effect studies of digital financial inclusion and provide policy implications for a financial security mechanism optimization in expanding imports.

Key words: digital financial inclusions; enterprises' imports; enterprise sizes

(上接第 47 页)

Generative AI in the Financial Sector: Applications, Risks and Regulatory Suggestions

XU Chao

(*Institute of Information Studies, CASS, Beijing*, 100732)

Abstract: Based on information technology, a new round of scientific and technological revolution is in an ascendant. Artificial intelligence (AI), especially generative artificial intelligence, is widely used in different fields of human social life, promoting a rapid development in relevant industrial chains. Current researches reveal that widespread AI applications in finance indeed greatly improve economic benefits and efficiency, yet with many potential risks. Therefore, in order to meet current needs of financial technological development and properly handle the relationship between financial technology innovation and security, many regions, countries and international regulatory agencies proactively formulate relevant regulatory systems and improve the capabilities of regulatory agencies. As a major country in AI technology and financial industry development, China needs to take multiple measures to formulate relevant regulatory systems. This serves not only to control relevant potential risks, but also to help compete for institutional voicing powers.

Key words: artificial intelligence; risks; generative artificial intelligence; financial regulations