

人工智能大模型在财会监督中的应用场景与 风险规制研究

刘 婕

北华大学经济管理学院, 吉林 吉林

收稿日期: 2026年3月20日; 录用日期: 2026年3月31日; 发布日期: 2026年4月29日

摘 要

随着人工智能技术突破, 大模型正深刻变革财会监督领域。本文采用系统性文献综述方法, 研究发现, 现有文献表明大模型在自动化审计、欺诈检测、文本分析与智能决策等场景中展现出提升数据处理效率和辅助穿透式监管的潜力; 然而, 技术应用也伴随数据隐私泄露、算法偏见固化、责任归属困境等风险。初步证据显示, 当前大模型在财会监督中的实际应用仍处于探索阶段, 其能力边界、可靠性与可解释性尚需进一步验证。结论认为, 需通过完善数据治理、加强算法审计、明确伦理边界并推动监管科技同步发展, 促进技术与监督的良性融合。

关键词

大模型, 财会监督, 监管科技, 风险规制, 系统性文献综述

Research on Application Scenarios and Risk Regulation of Large-Scale Artificial Intelligence Models in Financial and Accounting Supervision

Jie Liu

School of Economics and Management, Beihua University, Jilin Jilin

Received: March 20, 2026; accepted: March 31, 2026; published: April 29, 2026

Abstract

With breakthroughs in artificial intelligence technology, large language models are profoundly

transforming the field of financial and accounting supervision. This paper employs a systematic literature review methodology and finds that existing research indicates large language models demonstrate the potential to enhance data processing efficiency and support penetrative supervision in scenarios such as automated auditing, fraud detection, text analysis, and intelligent decision-making; however, the application of this technology also carries risks such as data privacy breaches, the entrenchment of algorithmic bias, and the dilemma of assigning responsibility. Preliminary evidence suggests that the practical application of large language models in financial and accounting supervision remains in an exploratory phase, and their capabilities, reliability, and interpretability require further validation. The conclusion argues that a positive integration of technology and governance must be fostered by improving data governance, strengthening algorithmic auditing, clarifying ethical boundaries, and promoting the synchronized development of regulatory technology.

Keywords

Large-Scale Models, Financial and Accounting Supervision, Regulatory Technology, Risk Regulation, Systematic Literature Review

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

在数字经济深度渗透和企业数字化转型的浪潮下，财会监督作为国家监督体系的重要组成部分，正面临数据规模爆炸性增长与业务复杂度跃升的双重压力。传统的结构化指标与规则驱动型监督手段在应对海量、多源、异构的非结构化数据时力不从心，逐渐显现出效率不足和深度有限的局限。近年来，以 ChatGPT、千问、DeepSeek 为代表的生成式人工智能大模型凭借千亿级参数的深度学习能力和强大的语义理解和逻辑推理能力，为解决上述瓶颈提供了全新路径。

大模型能深入解读财务报告、审计意见和社交媒体舆情中的文字信息，通过语义理解和逻辑推理，将这些复杂内容转化为监管所需的风险指标，有效打通业务语言与监管规则之间的壁垒，为财会监督从经验驱动转向数据智能驱动提供了关键技术支撑。然而，监管技术的快速迭代也带来了显著的治理挑战：模型的不可解释性影响决策透明度，训练数据中的偏见可能导致监管不公，模型“幻觉”现象增加了误判风险，同时数据安全边界弱化、责任链条模糊以及过度依赖技术等问题，共同构成了效率提升与风险管控之间的现实张力。如何在释放技术红利的同时，建立健全涵盖法律规范、伦理准则与技术治理的规制体系，成为财会监督现代化进程中的核心命题。本文聚焦大模型的具体应用场景，系统梳理其引发的法律与伦理风险，并尝试构建一套整合模型全生命周期管理、数据安全合规、人机协同决策和动态法规更新的治理框架，以期为促进监管科技健康发展、赋能财会监督提质增效提供理论依据与实践参考。

2. 大模型赋能财会监督的应用场景

2.1. 自动化财务审计与合规检查

随着企业财务数据规模及复杂性的激增，传统的依赖人工的审计方法在效率与覆盖面上面临严峻挑战。大语言模型凭借其强大的语义理解与知识表征能力，正在从“辅助工具”向“审计副驾驶”的角色演进，为审计流程的智能化重构提供了新的范式。

通用大语言模型由于缺乏深厚的会计专业知识，在直接应用于审计场景时往往效果不佳。针对这一问题，本研究通过构建高质量的审计领域专用语料库，对基础模型进行定向的参数微调。经过这种专业化训练，模型能够精准解读财务报表附注及管理层讨论中的文字描述，有效识别企业披露内容与国际财务报告准则或中国企业会计准则之间的潜在偏差。由此，财会监督的模式得以从过去简单的“核对格式是否完整”升级为真正意义上的“判断业务实质是否合规”[1]。

大模型能够在商业合同与监管要求之间建立语义层面的关联。通过自然语言处理技术，模型可自动将合同条款与最新法规进行比对，快速识别权责不对等、披露不充分或利益冲突等隐性风险[2]。这一机制将审计人员从重复性的条款核对中解放出来，推动审计重心从“事后检查”向“事中控制”转移，显著提升了合规审查的效率与覆盖面。

2.2. 欺诈检测与实时监控

在交易数据规模呈指数级增长且舞弊手段日趋隐蔽的背景下，传统的基于规则或抽样审计的方法难以在海量数据中挖掘深层次的关联线索。大语言模型凭借其强大的上下文关联能力与多维度特征提取能力，为构建智能化异常预警系统提供了新的技术路径。

对于基于语义关联的复杂异常模式识别，在供应链金融、跨境支付等复杂业务场景下，传统基于结构化数据的查询分析方式存在明显局限。大模型技术能够对交易附言、发票影像描述等非结构化信息进行深度语义解析，同时结合关联方关系构建知识图谱和行为画像，实现对复杂交易链路中异常模式的自动识别。具体而言，在供应链金融审计实践中，该模型通过综合比对发票信息的内在逻辑一致性、交易对手背景特征以及历史交易行为习惯，能够有效识别虚假发票循环和隐性关联方交易等舞弊行为，发现常规审计方法难以察觉的潜在风险信号。

具备可解释性的实时动态监测体系，在异常检测领域，传统模型往往呈现出“黑箱”特征——只能给出检测结果，却无法解释判断依据。相比之下，大模型技术正在推动实时监测实现从“结果输出”向“过程解释”的转变。该技术赋予了系统与审计人员进行自然语言交互的能力：当触发预警阈值时，模型不仅能够标注出异常交易，还能回溯其分析推理过程，以审计人员易于理解的方式，阐明数据之间的关联逻辑与决策依据[3]。这种人机协同的实时反馈机制，既提升了监控响应的时效性，也增强了审计人员对人工智能决策的信任程度和人工干预能力。

2.3. 财务报告文本分析与价值评估

在财务信息披露日益详尽的趋势下，非结构化文本数据在公司年报及临时公告中的占比显著提升。大语言模型凭借其对自然语言的高阶理解能力，突破了传统审计仅关注结构化财务数据的局限，为挖掘文本信息中的增量价值提供了方法论支撑。

基于语义分析的管理层讨论文本偏差识别，管理层讨论与分析作为财务报告的补充说明，虽然蕴含着大量预测性和解释性信息，但其主观性较强，常被管理层用作印象管理的工具。学术界与实务界普遍认识到这一问题。研究表明，借助大模型的深度语义分析技术，审计人员能够系统性地比对该部分文本内容与同期财务数据之间的逻辑一致性。具体而言，模型可以识别出文本中过度乐观的表述、对潜在风险的刻意回避，以及那些与历史趋势明显偏离的语义偏差。这种文本挖掘能力不仅拓宽了审计证据的获取渠道，也为评估管理层真实意图、识别盈余操纵风险提供了新的分析视角[4]。

面向企业价值评估的多源异构信息整合，在企业并购、资产减值测试等复杂估值场景中，估值模型是否准确，很大程度上取决于输入参数是否可靠。大模型技术能够打通结构化数据与非结构化文本之间的界限，实现对多源信息的快速检索与整合分析[5]。具体来说，模型可以同时处理历史交易数据库中的

可比案例、行业研究报告中的趋势判断，以及法律文件中隐含的或有负债信息，将这些原本分散的证据串联起来，为估值模型提供更为全面的参数输入。这种基于大模型的知识增强型估值辅助方法，有效提升了分析师对复杂商业环境的理解深度，也增强了估值结论的论证充分性。

2.4. 智能财务客服与决策支持

大语言模型在自然语言生成与多轮对话理解方面的突破性进展，正在重塑财务领域的信息交互模式与服务供给方式。从面向全体员工的普惠式咨询到面向专业审计师的决策增强，生成式对话技术正在构建一个全天候、智能化、可解释的财务服务新范式。

对于知识库增强的7×24小时智能财务咨询服务，在大型企业集团的日常运营中，财务咨询需求往往呈现出高频、重复且类型多样的特点。大模型技术的智能客服系统能够有效应对这一挑战。该系统通过深度融合企业内部知识库与外部监管政策，构建了一个全天候在线的财务知识服务平台。员工可以通过自然对话方式，随时咨询费用报销标准、税务政策适用性以及预算执行规则等常见问题，并得到精准解答。相比传统的人工坐席或菜单式自助服务，这种生成式交互模式不仅显著减轻了财务部门的日常事务负担，还借助多轮对话澄清机制，提升了咨询服务的准确性和用户体验满意度。

在审计风险感知与决策增强机制，在复杂的审计判断场景中，决策质量高度依赖于审计师对数据背后业务实质的穿透式理解。大模型通过深度自然语言交互界面，充当审计师的认知增强工具。当审计师面对异常财务指标时，可通过连续追问的方式与模型展开对话式探索，模型则能够实时调用底层数据模型，阐释数据波动与业务逻辑之间的潜在关联。这种人机协同的风险感知机制，将静态的数据分析升维为动态的认知对话，不仅强化了审计师对隐性风险的敏感性，也为其制定差异化的应对策略提供了多角度的逻辑支撑，从而推动了审计决策从经验主导向证据驱动的范式转变[6]。

3. 新型法律与伦理风险分析

3.1. 法律风险

在大模型应用于财务审计的过程中，主要面临三重核心挑战：一是数据隐私泄露风险。由于大模型具有记忆能力，在处理包含银行账户、税号、企业交易记录等敏感财务信息时，可能无意间“记住”训练数据中的隐私内容，并在生成回答时意外输出，这直接违反了《个人信息保护法》等相关法规的要求，给审计机构带来数据合规隐患。二是知识产权合规困境。大模型的训练往往依赖海量数据，若所采集的财务报告、审计底稿或企业内部文档未经授权或超出合理使用范围，可能构成对著作权人权益的侵害。特别是当模型生成内容与原作品高度相似时，侵权认定与责任划分将变得极为复杂。三是责任归属机制缺失。大模型决策过程的“黑箱”特性，使得当模型生成误导性审计结论或遗漏重大风险时，难以追溯错误根源。究竟是模型设计存在缺陷、训练数据存在偏见，还是使用场景不当，往往难以厘清，导致在开发者、部署机构与执业审计师之间无法清晰界定法律责任，给审计问责带来了极大的不确定性[7]。

3.2. 伦理风险

在大模型赋能财务审计的过程中，伦理层面的挑战同样不容忽视且相互交织。首先是算法偏见与歧视的固化风险，大模型可能继承并放大历史数据中对特定行业或企业的刻板偏见，导致风险评估失真，影响审计的客观公正[8]。其次是决策透明度与可解释性的根本性缺失，模型的“黑箱”特性使其推理过程难以追溯，审计师无法获知结论背后的证据链，这直接违背了审计工作对独立验证与过程可追溯的核心要求。最后是人机协作失衡与专业能力的退化，过度依赖模型输出易使审计师产生“自动化偏见”，逐渐沦为技术的附庸，长此以往将削弱审计行业赖以生存的专业怀疑与判断能力，甚至引发面对复杂舞

弊时的系统性风险。

4. 风险规制与治理框架

4.1. “软法”护航：以伦理准则与行业自律划定底线

当前大模型的治理主要依托“软法”机制，即以伦理指南、行业标准与自律守则为核心的柔性约束体系。在伦理层面，国内外组织倡导公平、透明与可问责原则，要求开发者在算法设计阶段嵌入伦理考量，并通过算法影响评估、伦理委员会等内部治理机制确保原则落地[9]。例如，欧盟《人工智能法案》即对高风险 AI 系统设定了严格的合规门槛，包括建立风险管理系统、使用高质量数据集以及引入人工监督，试图将抽象的伦理原则转化为可操作的法律义务。在技术层面，通过构建如 MMLU、C-Eval 等行业评测基准，对大模型在财务等专业领域的的能力进行标准化评估，不仅检验模型的知识储备与推理能力，也关注其在敏感内容识别、偏见规避等维度的表现[10]。同时，各类行业自律公约与技术白皮书的发布，促使头部企业承诺数据安全与算法透明，从而推动治理从模糊的原则倡导走向可量化、可验证、可追责的合规实践，为硬法监管的落地提供了重要的缓冲带与试验田。

4.2. “硬法”兜底：以立法约束与合规审查强化问责

为应对大模型带来的技术风险，各国正逐步从“软法”约束转向“硬法”监管，并在此基础上探索敏捷治理的新范式。一方面，通过设立“监管沙盒”，在受控环境中模拟真实财务场景，测试大模型应用的合规性与安全性，允许监管者在风险可控的前提下观察模型行为、收集运行数据，从而提前识别并规避系统性风险，为政策制定提供实证依据。另一方面，出台明确的数据保护法与算法责任法，强制要求财务监督场景下的数据脱敏、模型备案与安全审计，通过算法解释义务与过错推定原则，构建起从模型开发、部署到审计结果输出的全流程可追溯法律责任链条，确保每一次算法决策都能找到最终的责任承担者。

4.3. “敏捷治理”：以动态监测与多元共治应对迭代

面对大模型快速迭代、能力持续涌现的特性，传统的静态监管模式已难以有效应对[11]。实践中正逐步转向“敏捷治理”的新型范式：一方面在治理逻辑上强调“边发展、边治理”，通过建立持续性的风险监测与算法审计机制，定期扫描模型在财务等敏感领域的潜在风险，使监管政策能够随技术演进动态调整，避免“一刀切”造成的创新抑制或监管真空[12]；另一方面在治理结构上推动多元主体协同参与，打破政府单一监管的局限，由政府部门设定法律底线、行业协会制定专业标准、技术企业履行算法披露义务、审计机构落实人工复核流程，形成权责清晰、信息共享的责任链条，从而构建一个兼具创新包容性与风险约束力的动态监管体系[13]。

5. 结论

综合上述文献分析，大模型技术正深度重塑财会监督的范式，其在自动化审计、欺诈检测、文本分析与智能决策等场景中展现出显著的效率优势与认知增强潜力。然而，目前该领域的研究仍处于早期探索阶段，技术红利与治理风险始终相伴而生。数据隐私泄露、算法偏见固化、决策黑箱化以及责任链条断裂等问题，构成了大模型赋能财会监督过程中的核心法律与伦理挑战。

为应对大模型在财会监督中的法律与伦理挑战，本文提出以下治理路径：一是建立财会 AI 合规使用标准，为应对人工智能在审计应用中引发的数据泄露风险与责任归属困境，必须从数据输入与决策输出两个关键节点构建协同机制；二是完善责任追溯机制，明确 AI 辅助结论的法律定位，推行“人机共同签

名”制度，确保最终判断由持牌审计师负责；三是推动可解释性审计专用模型的研发，鼓励开发具备决策路径回溯能力的审计专用模型，使其能够呈现关键判断的依据与证据链条，满足审计工作对过程可追溯的基本要求；四是提升审计师技术素养，通过系统性培训强化审计师对 AI 的监督能力，避免技术依赖导致的专业判断退化，从而在人机协同中实现效率与公正的平衡。

本文的主要局限性在于：由于大模型技术在财会监督领域的应用仍处于早期阶段，现有文献以探索性研究和概念验证为主，实证研究基础相对薄弱。此外，本文为单研究者综述，可能存在文献筛选的主观性。未来研究可进一步开展实地实验研究和纵向跟踪研究，以更可靠地评估大模型在真实财会监督场景中的效果、风险与治理措施的有效性。

综上所述构建人机协同的财会监督新范式，核心在于确立“效率由机器驱动，安全与公正由人类掌控”的理性治理格局。唯有将技术创新作为效率引擎，制度完善作为安全底座，二者并行不悖，才能彻底突破传统财会监督的经验主义桎梏，重塑从被动检查向主动预警的全流程监管能力，从而为国家治理体系和治理能力现代化提供坚实的技术支撑与制度保障。

参考文献

- [1] 万钧. 基于大语言模型的审计知识应用研究[J]. 审计研究, 2024(5): 38-44+74.
- [2] Guo, H., An, B., Guo, Z. and Su, Z. (2020) Deep Semantic Compliance Advisor for Unstructured Document Compliance Checking. *Proceedings of the 29th International Joint Conference on Artificial Intelligence, Special Track on AI in FinTech*, Yokohama. 11-17 July 2020, 4446-4452. <https://doi.org/10.24963/ijcai.2020/613>
- [3] 华敏垚. 可解释的金融用户异常交易检测研究[D]: [硕士学位论文]. 南京: 东南大学, 2022.
- [4] 洪康隆. BERT 人工智能模型能识别管理层语调中的机会主义吗?——基于上市公司年报的文本分析[J]. 证券市场导报, 2024(10): 27-37, 68.
- [5] 姜雨含. 基于 EVA 的人工智能企业价值评估研究——以科大智能为例[D]: [硕士学位论文]. 昆明: 云南财经大学, 2019.
- [6] 黄佳佳, 周立云, 徐超. 基于大模型的审计知识智能问答系统构建研究[J]. 会计之友, 2025(9): 24-30.
- [7] 侯福瑞. 人工智能通用大模型在财务分析中的创新应用与挑战[J]. 行政事业资产与财务, 2025(17): 58-60.
- [8] 刘友华. 算法偏见及其规制路径研究[J]. 法学杂志, 2019, 40(6): 55-66.
- [9] 郭小东. 大模型时代全球人工智能治理的挑战与中国方案[J]. 科学学研究, 2025, 43(4): 694-702.
- [10] 张晔理, 杨滔. 人工智能技术驱动下的城市管理智能化路径探讨[J]. 科技管理研究, 2025, 45(10): 174-180.
- [11] 薛澜, 赵静. 走向敏捷治理: 新兴产业发展与监管模式探究[J]. 中国行政管理, 2019(8): 28-34.
- [12] 薛澜, 贾开, 赵静. 人工智能敏捷治理实践: 分类监管思路与政策工具箱构建[J]. 中国行政管理, 2024, 40(3): 99-110.
- [13] 吴磊, 冷玉, 唐书清. 数字化时代敏捷治理的学术图景: 研究范式与实现路径[J]. 电子政务, 2022(8): 77-88.