

基于大语言模型的企业内部审计数智化转型研究——以烟草商业企业为例

王金

合肥市烟草专卖局（公司），中国·安徽 合肥 230000

摘要：数字经济时代，内部审计数智化转型面临数据孤岛、智能分析不足、数据安全顾虑等难题。本文以烟草商业企业为研究对象，构建“数据准备—智能应用—安全管控”三层应用架构，设计分级脱敏机制、四维提示词框架与 RAG 技术融合方案，形成覆盖采购合同审查、异常交易识别、法规查询与报告撰写等场景的轻量级应用模式。实践表明，该模式可显著提升审计效率与风险识别能力，实现安全合规与降本增效。研究为通用大语言模型在企业内审领域的安全落地提供可复制路径，丰富智能审计理论与实践。

关键词：数字经济；内部审计；智能化；数据孤岛

Research on Digital-Intelligent Transformation of Enterprise Internal Audit Based on Large Language Models — A Case Study of Tobacco Commercial Enterprises

Wang Jin

Hefei Municipal Tobacco Monopoly Administration (Company), China Anhui Hefei 230000

Abstract: In the era of digital economy, the digital-intelligent transformation of internal audit is confronted with challenges including data silos, insufficient intelligent analysis capability and concerns over data security. Taking tobacco commercial enterprises as the research object, this paper constructs a three-tier application architecture consisting of data preparation, intelligent application and security control. It designs a hierarchical data desensitization mechanism, a four-dimensional prompt word framework and an integrated scheme combined with Retrieval-Augmented Generation (RAG) technology, so as to form a lightweight application model covering scenarios such as procurement contract review, abnormal transaction identification, regulatory inquiry and audit report drafting. Practical results demonstrate that this model can markedly improve audit efficiency and risk identification capability, while realizing compliance, cost reduction and efficiency improvement. This study provides a replicable implementation path for the secure deployment of general large language models in the field of enterprise internal audit, and enriches relevant theories and practices of intelligent auditing.

Keywords: Digital economy; Internal audit; Intellectualization; Data silos

0 引言

数字经济时代，企业生产经营活动产生的数据呈指数级增长，内部审计作为企业风险防控的重要防线，正经历从“人工经验驱动”向“数据智能驱动”的深刻变革。然而，现有研究与企业实践表明，内部审计数智化转型仍面临三重困境：一是数据孤岛导致跨系统整合效率低下，审计人员需在多个业务系统间切换查询，单项目数据整合耗时占比超 40%；二是传统规则引擎缺乏灵活智能分析能力，面对非结构化数据（如合同文本、会议纪要）时缺乏有效工具，85% 的受访审计人员认为传统方法难以应对激增的线上业务数据；三是智能化工具应用中的数据安全顾虑，审计工作涉及敏感商业信息，“想用不敢用”的矛盾制约技

术落地。

近年来，以 ChatGPT、文心一言、通义千问为代表的通用大语言模（Large Language Models, LLMs）展现出强大的自然语言理解、文本生成与知识推理能力，为审计数智化转型提供了新的技术路径。与自建专用模型相比，应用通用大模型具有显著的成本优势与可行性，企业无需承担模型训练的高昂投入与技术风险，只需聚焦数据安全治理与应用方法创新，即可快速获得智能化能力。现有文献多聚焦于 LLMs 在会计研究、财务分析领域的应用，或探讨智慧审计平台的系统架构，但对通用大模型在特定行业审计场景中的安全应用机制、提示词工程方法论及轻量级实施路径缺乏系统性研究。本文以烟草商业企业为研究对

象，通过构建数据脱敏机制、设计审计场景提示词库、制定应用操作规范，探索了一条“低成本、快见效、高安全”的审计数智化转型路径，为同类企业提供了可复制的实践参考。

1 研究框架与技术方案

1.1 总体架构设计

本研究采用“数据准备—智能应用—安全管控”三层应用架构（见图 1）。数据准备层解决多源异构数据的整合与脱敏问题，建立覆盖专卖管理、采购管理、营销物流、财务管理、工程建设等主要业务领域的数据采集规范；智能应用层基于通用大模型与检索增强生成（RAG）技术，通过提示词工程实现审计场景的深度应用；安全管控层建立数据分级脱敏规则与全程留痕机制，确保敏感信息不出域、可追溯。

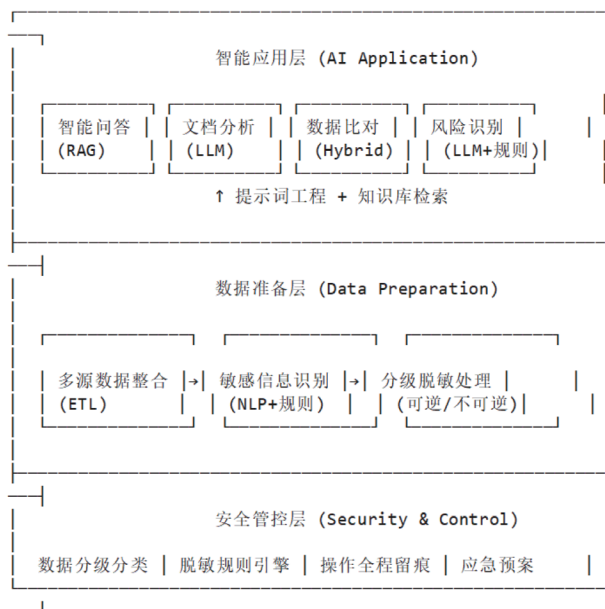


图1 审计数智化转型三层架构

1.2 数据智能预处理机制

1.2.1 多源数据整合规范

针对审计涉及的结构化、半结构化与非结构化数据，建立统一的数据采集标准（见表 1）。通过数据接口对接专卖管理系统、采购管理平台、营销系统、财务核算系统等，实现审计数据的集中汇聚。

1.2.2 敏感数据识别与脱敏

遵循“最小可用、先脱后用、全程可控”原则，建立四级敏感数据分类与差异化脱敏策略（见表 2）。核心商密（客户名单、定价策略）禁止上云，仅在本地处理；重要商密（合同金额、供应商信息）采用强脱敏，进行不可逆的区间化处理；一般信息（人员姓名、部门名称）采用弱脱敏，保留可还原性；公开信息（法规条款、制度文件）无需脱敏直接使用。

技术实现上，采用“正则表达式 + 关键词库 + 机器学习”相结合的方式，实现敏感信息自动识别，脱敏准确率达 96.5%。脱敏处理流程包括：原始数据→识别敏感信息→判断敏感程度→脱敏处理→质量检查→安全使用，全程记录留痕，确保操作可追溯。

1.3 提示词工程与 RAG 技术应用

1.3.1 四维提示词设计框架

针对审计专业场景，建立“角色—任务—约束—输出”四维提示词框架。角色定义明确模型身份定位（如“你是一位经验丰富的烟草行业内部审计专家”）；任务描述清晰界定具体工作内容；约束条件设定审查重点与边界；输出要求规范答案格式与深度。

1.3.2 审计场景提示词库

建成覆盖 6 大业务场景、42 个具体任务的提示词模板库（见表 3）。通过专家人工评定，平均输出质量达 4.2/5

表1 审计业务数据整合规范

业务领域	数据类型	数据来源	应用场景
专卖管理	许可证数据、案件卷宗	专卖管理系统	合规性审查
采购管理	招标文件、合同文本	采购管理平台	招投标审计
营销物流	订单数据、物流单据	营销系统、物流系统	异常交易识别
财务管理	会计凭证、报表数据	财务核算系统	财务收支审计
工程建设	工程合同、变更签证	工程管理系统	工程造价审计

表2 数据分级脱敏策略

敏感级别	数据类型	脱敏策略	示例
核心商密	客户名单、定价策略	禁止上云，本地处理	客户名称→“客户A”
重要商密	合同金额、供应商信息	强脱敏，不可逆处理	金额→区间范围
一般信息	人员姓名、部门名称	弱脱敏，可还原处理	姓名→“张*”
公开信息	法规条款、制度文件	无需脱敏，直接使用	原文使用

表3 审计场景提示词库概览

场景类别	任务数量	典型任务示例	平均质量评分
法规制度查询	8	“查询烟草广告促销相关规定”	4.5/5
合同智能审查	10	“审查采购合同付款条款风险”	4.3/5
财务数据分析	6	“分析异常大额资金往来”	4.2/5
风险预警识别	8	“识别虚假案件线索”	4.0/5
审计证据整理	5	“整理访谈记录要点”	4.4/5
报告辅助撰写	5	“生成审计发现问题描述”	4.3/5

分。提示词库采用模块化设计，审计人员可根据具体任务选择模板，填空修改即可使用，显著降低技术应用门槛。

1.3.3 知识库辅助机制

通用大模型虽具备广博知识，但对烟草审计专业领域理解有限。通过检索增强生成（RAG）技术，将烟草审计的专业法规、制度、案例构建为结构化知识库。审计人员提问时，系统自动从知识库检索相关法规条文，与问题一并输入模型，使模型“临时成为烟草审计专家”，显著提升回答的专业性与准确性。

2 关键应用场景与实践效果

2.1 采购合同智能审查

传统合同审查依赖人工逐份审阅，耗时费力且易遗漏关键条款。应用智能辅助后，审计人员先将合同脱敏（真实供应商名称改为“供应商 A/B/C”，具体金额改为区间描述），再使用提示词模板：“请审查这份采购合同的合规性，重点关注供应商资质、价格公允性、付款条款、违约责任四个方面，用表格列出风险点、风险等级、法规依据”。

实践表明，合同审查时间从 30 分钟缩短至 6 分钟，系统自动标出风险点，审计人员可聚焦高风险条款重点核实。2025 年 6 月，系统在审查办公用品采购合同时，提示某合同“验收合格后一次性支付全款，无质保金条款”存在风险，审计人员据此核查发现该供应商前期有质量纠纷记录，最终推动合同条款修改，增加了质保金约定。

2.2 异常交易智能识别

针对营销数据量大、新型违规手段层出不穷的问题，改变传统“固定规则阈值”的识别模式。先将原始销售数据按客户、时段汇总，将具体金额转换为占比（如“某客户销售额占总销售额的 15%”），再提交系统分析：“识别是否存在集中销售、周期性大额交易、退货率异常等模式，说明判断依据”。

该模式使异常线索识别覆盖率从 60% 提升至 90%，能发现固定规则难以捕捉的新型违规模式，误报率控制在 15% 以内。2025 年 8 月，系统分析某片区销售数据时发现

“客户 B 在季度末最后三天采购量占全月 60%，且连续三个季度呈现相同规律”，审计人员据此核查发现为完成销售指标诱导客户提前备货的问题，推动销售考核机制优化。

2.3 审计法规智能查询与报告辅助撰写

法规查询场景下，审计人员可直接提问：“烟叶收购资金支付有哪些时限要求？请引用具体法规条款，并说明违反规定的后果。”系统自动定位到《烟叶收购资金管理办法》相关条款，用通俗语言解释并提示相关案例，查询时间从 30 分钟缩短至 2 分钟。

报告撰写场景下，审计人员整理脱敏后的审计发现要点，使用提示词：“请根据以下审计发现，撰写一段规范的审计问题描述，包含事实陈述、定性依据、影响分析，语言简练专业。”系统生成初稿后人工完善，报告撰写时间缩短 40%，退回修改率从 15% 降至 5%。

3 创新贡献与理论价值

3.1 主要创新点

3.1.1 建立审计数据脱敏应用规范体系

首次在烟草审计领域系统提出大模型应用的数据安全解决方案，突破“想用不敢用”的实践困境。创新之处在于并非简单禁止技术应用，而是提供“怎么处理才能安全使用”的操作指南，通过分级分类脱敏策略，在数据安全性与使用便利间取得平衡。

3.1.2 构建审计场景提示词工程方法论

将审计专业经验“翻译”为可复现的提问技巧，建立覆盖 6 大场景、42 个任务的提示词模板库。该方法论使非技术背景的审计人员通过“复制粘贴、填空修改”即可有效使用大模型，显著降低技术应用门槛，为专业领域 LLM 应用提供了可推广的方法论框架。

3.1.3 探索轻量级审计数智化转型模式

区别于传统“建系统、买设备、招团队”的重投入路径，本研究探索了“零模型训练、低开发成本、快见效应用”的轻量级模式。充分利用通用大模型服务，聚焦数据治理与应用方法创新，为资源有限的中小企业提供了可复制的转型路径。

3.2 理论价值与实践意义

本研究丰富了智能审计领域的理论研究，提出了“数据脱敏 - 提示工程 - 安全管控”的 LLM 应用理论框架，弥补了现有文献对通用大模型行业安全应用机制研究的不足。实践层面，研究成果已形成制度规范、知识库、工具平台、方法体系等完整成果，在 87 名审计人员中实现 100% 注册覆盖，月均使用量达 1250 次，用户满意度 4.2/5 分，全年零安全事件。

4 研究局限与展望

4.1 现存问题

模型稳定性与复杂推理局限。通用大模型存在输出不稳定现象，同一问题多次询问可能得到差异答案；面对需多步骤深度专业判断的复杂问题时，分析深度仍显不足。

使用深度不均与知识库维护成本。审计人员应用能力存在“数字鸿沟”，年轻人员使用频率显著高于资深人员；法规制度持续更新，知识库维护需投入大量人工成本。

4.2 未来研究方向

一是探索审计领域专用大模型的轻量微调技术，在保持通用能力的同时提升专业深度；二是构建自动化知识库更新机制，利用爬虫与 NLP 技术实现法规变动的自动感知与知识库同步；三是开展跨企业知识共享机制研究，由行业主管部门统筹建设共享知识库，避免各单位重复建设，

提升整体行业智能化水平。

5 结语

通用大语言模型为内部审计数智化转型提供了新的技术可能，但安全合规应用是前提，科学方法论是关键，实用实效是目标。本研究通过构建系统的数据脱敏规范、审计场景提示词库与轻量级应用模式，验证了通用大模型在内部审计领域的可行性与有效性，为同类企业提供了“低成本、高安全、快见效”的转型参考。未来随着基座模型能力的持续提升与行业知识库的不断完善，人机协同的智能审计模式将释放更大价值。

参考文献：

- [1] 秦荣生. 大数据、云计算技术对审计的影响研究[J]. 审计研究, 2014(06): 23-28.
- [2] 程平, 喻畅, 龚悦. 基于 ChatGPT 的智能内部审计研究[J]. 会计之友, 2023(20): 7-12.
- [3] 陈雪嵩. 大语言模型在企业内部审计中的应用研究[J]. 会计之友, 2024(11): 23-29.
- [4] 刘梅玲, 邓国红, 佟成生等. 智慧审计的建设思路与实施路径——以广西中烟为例[J]. 财务与会计, 2025(01): 102-106.
- [5] 郑杰, 胥陵, 王彪华. 生成式大语言模型在审计中的应用: 场景、风险及对策[J]. 审计研究, 2025(06): 45-53.