

数据导向审计下电商企业数据审计平台的构建

杨昕宇

南京信息工程大学商学院, 江苏 南京

收稿日期: 2025年2月9日; 录用日期: 2025年2月28日; 发布日期: 2025年3月27日

摘 要

在大数据时代的背景下,大数据审计成为现代审计发展的关键方向,是提高审计质量和效率的重要手段。电子商务企业作为依托互联网和大数据技术进行商业活动的新兴企业,涵盖商品交易、服务提供及信息交换等多个方面,极大地改变了传统商业模式。在此背景下,通过大数据技术建立审计资源开放共享的数据审计平台,提出适合电商企业的审计方法,从传统审计向数据导向型审计转型,具有重要的现实意义。本文分析了大数据技术下数据导向审计与传统审计的区别,论述了电商企业数据审计平台的构建框架,对平台的数据搜集、管理方式进行了探讨,为更好地促进传统审计向大数据审计转型提供了思路。

关键词

电子商务企业, 数据导向审计, 数据审计平台, 大数据审计

Construction of E-Commerce Enterprise Data Audit Platform under Data Oriented Audit

Xinyu Yang

School of Business, Nanjing University of Information Science and Technology, Nanjing Jiangsu

Received: Feb. 9th, 2025; accepted: Feb. 28th, 2025; published: Mar. 27th, 2025

Abstract

In the context of the big data era, big data auditing has become a key direction in the development of modern auditing, serving as an important means to improve auditing quality and efficiency. As an emerging enterprise that relies on the Internet and big data technology to carry out business activities, e-commerce enterprises have greatly changed the traditional business model through electronic means to cover many aspects such as commodity trading, service provision and information exchange. In this context, it is of great practical significance to establish a data audit platform for

open and shared audit resources through big data technology, put forward audit methods suitable for e-commerce enterprises, and transform from traditional audit to data-oriented audit. This paper analyzes the difference between data-oriented audit and traditional audit under big data technology, discusses the construction framework of data audit platform for e-commerce enterprises, discusses the data collection and management methods of the platform, and provides ideas for better promoting the transformation from traditional audit to big data audit.

Keywords

E-Commerce Enterprises, Data-Driven Audits, Data Audit Platform, Big Data Auditing

Copyright © 2025 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

审计是党和国家监督体系的重要组成部分，是推动国家治理体系和治理能力现代化的重要力量。企业是推动我国经济增长的重要力量，对企业开展科学合理的审计，有利于保障信息披露的质量，规范企业经营活动，从而维护我国资本市场的秩序和社会公众利益。

然而，随着互联网时代的到来，越来越多企业开始进军电子商务领域，其无界限的市场和全天候运营的特点使得业务量和数据呈爆炸式增长[1]，企业所面临的风险和数据量不断迭代、积累，传统风险导向审计模式已很难满足规模庞大、纷繁复杂的数据处理需求，大数据拥有海量数据存储与管理能力，以及分布式并行计算特点，这使其能够有效解决传统审计模式中的问题[2]。对此，理论界与实务界均强调应用大数据审计技术，实施数据导向审计，构建数据审计平台，从抽样审计向全覆盖审计转变。大数据审计不仅是战术层面的技术应用，还体现了组织流程、思维理念的战略性变革[3]。本文分析了大数据技术的发展下，数据导向审计与风险导向审计的区别，又提出供审计人员数据共享共用的电商企业跨地区、跨部门数据审计平台的构建框架，总结了大数据审计下数据搜集方式和工作流程，对平台的管理方式进行了探讨。

本文余下内容安排如下：第二部分对数据导向审计进行概述，阐述了传统风险导向审计与数据导向审计的区别；第三部分是电商企业审计数据平台的构建体系；第四部分为总结与反思。

2. 数据导向审计概述

2.1. 审计数据范围

传统审计对企业财务报表的公允性和合法性进行鉴证时，主要运用的数据是被审计单位的凭证、账簿和财务报表，数据来源单一和数据量有限等问题限制着审计人员全面了解被审计单位的情况[4]。大数据审计模式下，审计方式更加智慧化、高效化，机器学习、自然语言处理、社会网络分析成为新的审计方法[5]，大数据技术和大语言模型成为传统取证方式的有力补充[6]，审计范围从财务数据扩大到与财务报表相关的业务数据和外部数据。审计人员的审计范围不再局限于传统财务报表项目的相关数据、数据库中或者 ERP 系统中存储的各类数据，只要与审计工作相关，都可以被迅速采集[1]和分析。

2.2. 数据导向审计的风险模型

在传统的风险导向中， $\text{审计风险} = \text{重大错报风险} \times \text{检查风险}$ 。其中，重大错报风险的识别和评估

往往需要运用注册会计师的职业判断。然而职业判断通常具有主观性，会受到时间、压力、知识和经验不足以及所获取信息有限等因素的影响，职业判断可能因发生判断偏差而影响审计结果[7]。

大数据处理技术和大语言模型使得全样本审计成为可能，审计人员不再需要依靠职业判断进行抽样审计。一些需要审计人员实地检查盘点的审计程序不易受大数据技术的影响或影响程度较低，而函证、重新计算等审计程序更容易受信息技术的影响、容易被其他方法替代[3]。由于通过分析程序对财务数据及相关数据进行分析很有效率，所以无需通过控制测试来降低实质性程序的工作量，即审计中对控制测试的需求大为降低甚至消失[8]，审计流程见图 1 所示。

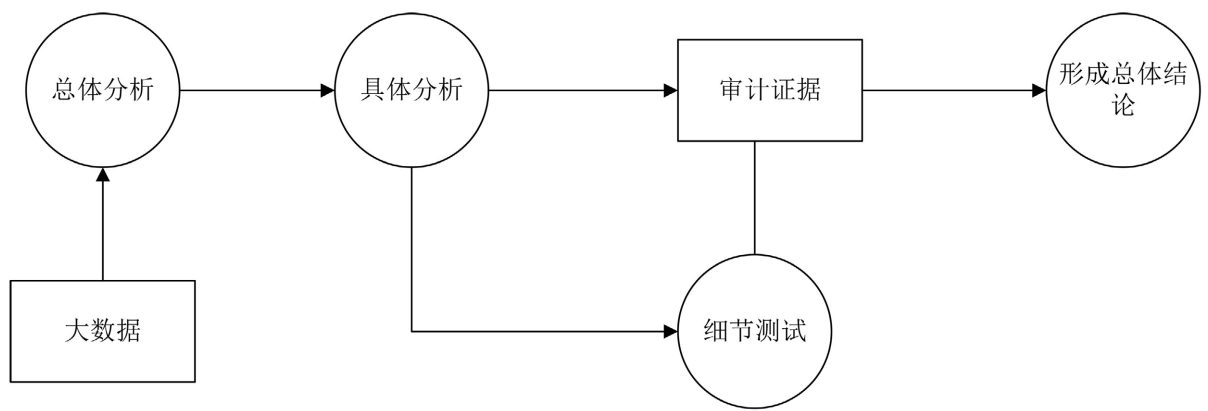


Figure 1. New audit process
图 1. 新的审计流程

在此情景下，审计风险模型也相应发生变化。数据导向审计风险模型变为：审计风险 = 数据差错风险 × 检查风险。其中，数据差错风险既可能是数据操作、传输、转换失误等因素导致的差错，也可能是缺乏内部控制或者内部控制无效的环境下所产生的各类数据错误[1]。

3. 电商企业数据审计平台

3.1. 数据收集与存储

数据审计平台首先应保障数据的数量与质量，审计人员只有获取和使用满足目标需求的高质量数据，才能提高审计效率和效果[3]。充分全面的数据源才能有效保障审计的深度和广度，从而实现审计全覆盖[9]。

在大数据环境下，由于电商企业相对于传统企业具有无界限的市场、全天候运营、数据驱动、产品和营销策略快速迭代等特点，为获取更加全面可靠的数据，不但需要采集被审计单位的财务报表、文档资料等传统审计数据，还可以使用大语言模型，对被审计单位内部招投标过程材料、重大事项决策过程会议纪要等文件类非结构化数据进行翻阅[10]，甚至需要从互联网上获取用户公开数据[11]，数据呈现碎片化、海量等特征[12]，单靠审计人员手工搜集存在很大困难。因此，数据审计平台可由不同地区的会计师事务所、被审计单位共同构建，旨在给审计人员提供共享、公用的非涉密数据使用平台，审计人员可从中获得开展审计工作所需的各种非涉密数据。通过大数据平台形成数据资源开放共享、跨部门、跨地区数据相互流通的格局，消除“信息孤岛”，加强信息交流、沟通协调[13]。其中，考虑到电商行业的特殊性，尤其可上传一些电商行业特有的数据供审计人员使用。

从数据来源看，本文借鉴黄芳[14]的审计数据分类，如图 2 所示将审计数据平台的数据分为微观企业数据、宏观社会环境数据两类。微观企业层面数据由企业自行上传，宏观社会环境数据可由审计机关上

传、会计师事务所协助完成。微观企业数据包括招股说明书、财务报表及附注、公告及其他公开资料，也包含非涉密的内部资料。就电商行业来说，还可结合直播平台的推荐算法、直播观看量、吸引观众停留的时长、企业所处平台活跃用户量等数据进行分析。宏观社会环境数据包括经济环境数据和其他环境数据。经济环境数据有国家经济与产业政策、税收政策、国民经济发展趋势、全球经济走势、汇率变动等，其他环境数据包括政治法律环境、自然环境。在审计工作中，通过使用被审计单位类似企业及上下游企业的信息，审计人员可了解被审计单位所在行业状况、技术趋势、竞争对手方面信息，以及主营产品的高新技术信息和替代产品动态，重要客户的采购量、重要供应商的产能或销售量等方面的信息等。此外，针对电商企业线上销售的业务特征，审计人员可通过爬虫收集消费者对被审计单位商品的评价，从而更好地识别企业是否在销量、回购率等数据上存在造假行为。

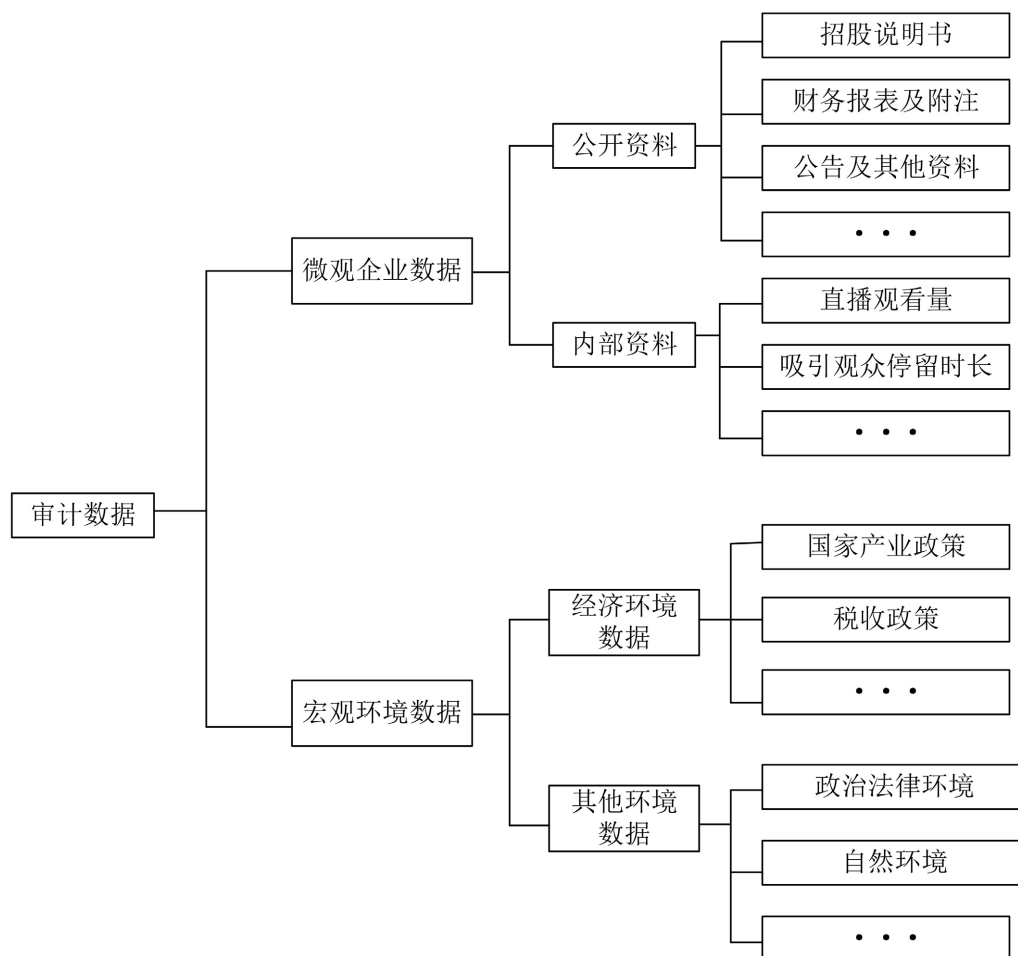


Figure 2. Data classification on the data audit platform

图 2. 数据审计平台数据分类

从数据形式看，审计数据共享平台不仅包含财务报表中的结构化数据，也包含图像、音频、视频等非结构化数据。其中，如表 1 所示，对非结构化数据的处理需要运用一定的数据处理和分析技术[4]。审计人员对网络信息进行爬虫时，可以使用八爪鱼采集器，无需编程代码即可采集数据。使用 Kimi chat 平台，可以查询实现相关算法。人工智能技术的使用，一定程度上降低了对审计 - 计算机复合型人才的要求，降低了审计人员的工作难度。

Table 1. The method of collecting audit data under the data-based audit
表 1. 数据导向审计下审计数据的采集方式

数据类型	数据来源	技术实现手段	常用工具
线下纸质材料	报纸、杂志、书籍等	扫描仪、OCR 软件	ABBYY FineReader 和汉王
线上电子数据	网页、微博、微信等	网页数据采集器、网络爬虫技术	八爪鱼采集器、乐思网络信息采集系统和 Octoparse, Java 和 Python 等编程语言

借鉴冯春雨[4]的数据存储方法,该过程共需进行两次数据存储。第一次存储的是采集后未经过处理的原始数据,对于非结构化数据,可以采用 Google 的 GFS、开源的 HDFS 等分布式文件系统,或使用 Redis、HBase、MongoDB 和 Neo4J 等 NoSQL 数据库进行存储;而结构化数据则采用 MySQL 等关系型数据库进行存储。第二次存储的对象则是通过 ETL 工具进行数据抽取、清洗和加载后所获得的标准化数据。标准数据处理完毕后,各相关单位将非涉密的标准数据上传至数据审计平台进行云端存储,以促进审计人员之间的数据共享。

3.2. 数据审计平台工作流程

图 3 为数据导向审计的工作思路。首先从多渠道进行数据采集与预处理,各单位将数据上传至数据审计平台,运用大数据技术进行总体分析,发现疑点后进一步的分散核实、系统研究。其中,审计数据来源于上级审计单位、审计单位管辖范围、被审计单位现场采集、网上相关公开数据以及其他所需数据。大数据分析技术包含大数据智能分析技术、可视化分析技术、多数据源综合分析技术等。同时,可建立信息建设中心,通过软硬件设施建设为大数据审计提供后勤保障,负责各个部门、各个流程的信息化建设,确保平台安全稳定运行、数据信息及时有效传递。利用信息技术推动审计现场数字化管理、开展非现场远程审计、进行审计整改实时追踪[13]。

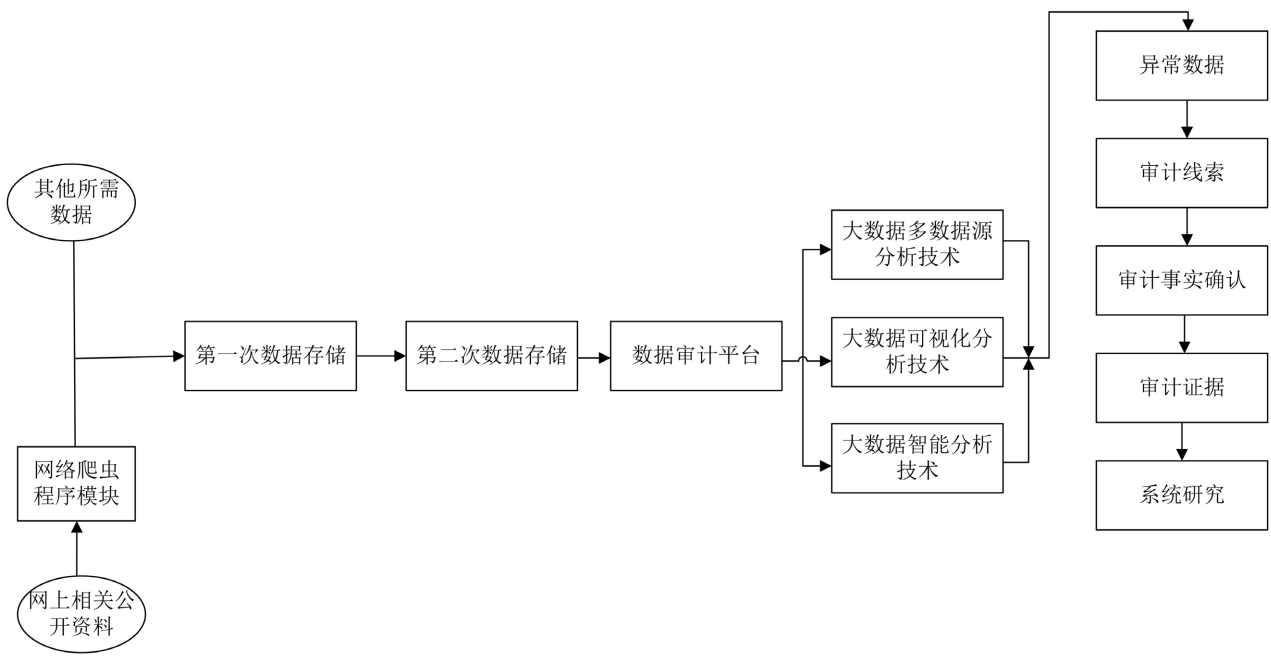


Figure 3. Workflow of audit data platform
图 3. 审计数据平台工作流程

3.3. 数据审计平台的管理

从审计角度上来看,大型云计算提供商已经有能力构建云审计平台,审计云的研究也有了不少成果[15]。即我国目前在构建审计数据平台的问题上,技术上具有可行性,但如何进一步规范管理数据的上传、使用,既促进数据共享,又不泄露企业关键商业数据,仍然值得讨论。

首先,实施数据导向审计,应遵循法规先行的原则[1],推动相关审计数据法律法规的建立。具体而言,首先要明确数据审计的目标、方法和技术手段,同时对数据审计的各个环节进行详细规定,尤其是在数据采集、存储、传输和分析等方面的法律要求。这些规定应当保证审计人员能够合法合规地获取和使用相关数据,避免出现违规行为。

其次,应从数据采集、存储、管理、分析、可视化各个环节加强技术支持和保障。数据与分析技术是大数据研究与应用中相互独立的两个方面,但同时又密切联系、互不可分[16]。大数据审计是基于多个不同来源的海量数据的审计,因此首先要尽可能采集全面而真实的数据。其次,由于涉及的审计数据量巨大,如何有效存储和管理这些数据成为了另一个关键环节,存储和管理数据的方式不仅要高效,还要能够确保数据的安全性与可追溯性,避免因存储不当导致的数据丢失或篡改。最后,应将分析结果以直观、易懂的方式呈现出来,一是供审计人员与被审计单位沟通、问询,二是便于审计人员快速做出科学的判断与决策。在现有的审计方法和实践中,最常使用的还是如聚类、关联规则挖掘、异常点检测、回归分析等数据挖掘技术,而更智能化的文本分析、深度学习、知识图谱等方法尚未被广泛应用[11]。当新技术出现时,可及时应用于数据审计平台。

最后,审计数据平台的信息应及时更新。电商行业日新月异,消费热点和网络评价变化极快,滞后的信息不仅不能给审计工作带来便利,甚至可能误导审计人员做出判定。可借助大数据技术实现数据的持续采集,通过大数据审计平台的构建实现数据的共享并实时分析与监控,由此确保审计不再局限于事后监督,而是将事项审计的实时审查理念贯穿审计全过程[17]。

4. 总结与反思

随着企业数字化转型与大数据技术的不断发展,现代意义的审计不只是查账、鉴证和评价活动,而是一项关乎市场治理、公司内部治理的整体性治理问题[18]。传统审计已无法适应经济社会的发展,开发数据审计平台并实施数据导向型审计是审计发展的必由之路。本文立足于电商企业数据呈爆炸式增长、大数据技术不断发展的特征事实,根据学者们提出的“数据导向审计”的概念,从审计数据范围和风险模型两方面对数据导向审计与传统风险导向审计的区别进行阐述,又提出供审计人员数据共享、共用的数据审计平台的构建框架,总结了大数据审计下数据搜集方式,对数据审计平台的工作流程和管理方式进行探讨。本文的不足之处在于,仅提出理论构想,未结合具体审计案例探讨审计数据平台设计在实际应用中的可行性。

参考文献

- [1] 秦荣生. 数据导向审计体系构建: 风险模型、方法体系与实现路径[J]. 审计研究, 2023(5): 3-10.
- [2] 汪芳. 大数据环境下全息交互智慧审计体系研究——以中国移动为例[J]. 财会通讯, 2022(17): 148-152.
- [3] 张悦, 杨乐, 韩钰, 等. 大数据环境下的审计变化、数据风险治理及人才培养[J]. 审计研究, 2021(6): 26-34+60.
- [4] 冯春雨, 冯均科, 钟越华, 等. 大数据环境下舆情导向审计流程再造[J]. 财会月刊, 2022(19): 90-97.
- [5] 张敏. 大数据审计: 五大趋势与五大挑战[J]. 会计之友, 2020(8): 2-11.
- [6] Yoon, K., Hoogduin, L. and Zhang, L. (2015) Big Data as Complementary Audit Evidence. *Accounting Horizons*, 29, 431-438. <https://doi.org/10.2308/acch-51076>

-
- [7] 徐静, 李俊林. 基于数据挖掘的重大错报风险识别和评估研究[J]. 财经理论与实践, 2022, 43(6): 79-85.
 - [8] 郑石桥. 电子数据环境对审计流程的影响: 一个理论框架[J]. 财会通讯, 2020(21): 14-17+47.
 - [9] 黄佳佳, 李鹏伟, 徐超. 面向智慧审计的思维变革与审计平台构建研究[J]. 审计研究, 2023(5): 11-20.
 - [10] 杨麟, 张宪礼, 于微伟. 大语言模型在国家审计中的应用探索[J]. 审计研究, 2024(6): 22-29.
 - [11] 徐超, 陈勇, 葛红美, 等. 基于大数据的审计技术研究[J]. 电子学报, 2020, 48(5): 1003-1017.
 - [12] 杨丽丽, 王会金, 刘国城. 管理控制视角下云审计平台的建设与运行研究[J]. 经济问题, 2020(6): 94-102.
 - [13] 李成艾, 何小宝. 大数据审计组织方式的探索与创新[J]. 审计研究, 2019(5): 23-29.
 - [14] 黄芳, 史琳璐. 从数据来源看分析程序[J]. 财务与会计, 2014(11): 62-63.
 - [15] 商思争. 大数据背景下自然资源资产离任审计问题探讨——以海洋自然资源资产离任审计为例[J]. 财会通讯, 2018(22): 106-110.
 - [16] Alles, M. and Gray, G.L. (2016) Incorporating Big Data in Audits: Identifying Inhibitors and a Research Agenda to Address Those Inhibitors. *International Journal of Accounting Information Systems*, 22, 44-59.
<https://doi.org/10.1016/j.accinf.2016.07.004>
 - [17] 张庆龙, 何佳楠. 大数据审计的业务逻辑: 基于事项审计的思考[J]. 会计之友, 2022(21): 46-51.
 - [18] 郑国洪, 赵磊磊. 新时代审计法治思维的理论内涵与实践运用[J]. 审计与经济研究, 2025, 40(1): 16-23.