

DOI: 10.16538/j.cnki.fem.20240724.201

# 大数据审计: 理论框架、研究进展与未来展望

徐荣华, 朱 婧, 戴欣瑜

(宁波大学 商学院, 浙江 宁波 315211)

**摘要:** 大数据审计是将新兴信息技术应用于审计业务活动中的新型审计模式。近年来, 随着大数据审计日益受到关注, 探究大数据审计模式的价值创造机理与路径成为学术界和实务界面临的重要课题。但是大数据审计的相关研究还未形成清晰的理论体系, 缺乏系统性的研究框架。基于此, 本文梳理评述了2011—2023年发表在国内外权威期刊上关于大数据审计的文献。首先, 系统阐述大数据审计的起源、概念内涵与模式比较。其次, 基于“核心环节—关键要素—基础保障—根本目的”分析框架, 梳理了大数据审计思维、大数据审计流程、大数据审计技术以及大数据审计应用的研究动态, 揭示了大数据审计的内在逻辑, 并结合实际调研补充了相关证据。最后, 构建了大数据审计研究的理论框架, 并从审计取证模式、审计预警机制、审计风险防范机制以及审计效果等不同角度提出未来展望, 为进一步推动大数据审计研究提供参考。

**关键词:** 大数据审计; 大数据审计模式; 大数据审计流程; 文献综述

**中图分类号:** F270 **文献标识码:** A **文章编号:** 1001-4950(2024)11-0122-16

## 一、引言

大数据、云计算、人工智能、物联网、区块链以及“互联网+”等新兴信息技术<sup>①</sup>的推广应用, 促进了管理数字化、业务数字化和流程数字化水平日益发展。大数据已经成为重要的生产要素, 被认为是“未来的新石油”、21世纪的“钻石矿”, 对一国经济社会发展具有重要的促进作用。建立在抽样审计和样本分析基础上的传统审计, 难以应对海量数据和不断精细化的管控流程。适应新兴信息技术, 大数据审计模式将审计监督与大数据及大数据分析技术紧密融合, 必然带来审计流程数据化, 财务报表审计重心转为鉴证报表项目数据的合法性和公允性(秦荣生, 2023)。大数据审计独特优势就在于审计范围扩大到经济体的全部经济活动范畴, 从而提升审计效率、扩大审计覆盖面和保证审计质量(Brown-Liburd等, 2015; Vasarhelyi等, 2015; 黄佳佳

收稿日期: 2024-01-02

基金项目: 国家社会科学基金一般项目(20BZZ062); 宁波南审审计研究院课题(政府购买公共服务审计研究)

作者简介: 徐荣华(1972—), 男, 宁波大学商学院副教授, 硕士生导师(通讯作者, [xuronghua@nbu.edu.cn](mailto:xuronghua@nbu.edu.cn));

朱 婧(1996—), 女, 宁波大学商学院硕士研究生;

戴欣瑜(2001—), 女, 宁波大学商学院硕士研究生。

①目前学界对于信息技术的称谓有“新一代信息技术”“新兴信息技术”“新型信息科技”等多种, 本文以“新兴信息技术”表述。

等,2023),其机理在于审计取证模式发生根本变化(谢志华和程恺之,2023),降低了审计风险(Dzuranin和Mălăescu,2016;Li,2023),利用机器学习、深度学习等技术提供审计决策最优化方案(张庆龙等,2020)。如何探究大数据审计模式的价值创造机理与路径成为学术界和实务界面临的重要课题。

审计产生于需要,正如Lee(1988)所言:“社会经济结构的变化导致审计需求的增加”。学术界认为,大数据审计模式的产生动因可以溯源到制度和技术两方面。一是制度层面的根本动因。INTOSAI(世界审计组织)秘书长Moser(2008)和审计署前审计长李金华(2005)、刘家义(2012)均指出,审计是国家治理的关键要素和重要工具,是对公共权力监督和制约的具体形式和专业手段(Schedler等,1999)。英国、爱尔兰等国审计重要变革是以公共利益为导向,致力于评价政府提供公共服务质量(Lonsdale,2007;Lindquist和Huse,2017)。此外,学者还实证检验并发现审计在促进公共支出效率、抑制“三公经费”预算水平、揭露与遏制政府部门腐败以及国有资产保值增值等方面具有积极作用(Ferraz和Finan,2011;吴秋生和郭檬楠,2018)。审计作为重要制度,在参与国家治理、抑制腐败以及维护公共利益等方面具有不可替代的专业优势且发挥独特的作用,审计制度需求程度是激发审计不断向技术层面转型发展的源动力。二是技术层面的关键动因。新兴信息技术保障了审计能够更好地发挥其功能。Michael(2015)认为,审计行业采用大数据分析(big data analysis,BDA)的驱动因素可以分为内部驱动因素和外部驱动因素。从外部来看,审计客户的大数据分析技术水平决定了其在财务报表审计中使用大数据分析工具的水平(Lombardi等,2014;Dagilienė和Klovienė,2019)。当审计师与客户之间的这种技术水平存在差异时,可能会对出具的审计意见产生影响(Salijeni等,2019)。因此,审计师需要事先对其与客户拥有的数据分析技术进行平衡,是决定采用BDA的主要外部驱动因素之一(Sun等,2015)。从内部来看,审计师需要采用能够提供更好审计见解并最终提高审计质量的技术工具(Alles和Gray,2016;Salijeni等,2019)。显然,BDA提高了审计师的洞察力,从而改善了客户与审计师之间的关系(Earley,2015),增强了审计师应对和处理海量数据的机动能力,将审计意见建立在数据分析的基础之上(Yoon等,2015)。大数据审计发挥技术、制度的双重溢出效应,并实现管理与技术的双向融合。

当前,学术界尚未对大数据审计概念与内涵达成共识,存在人工智能审计、数字式审计、云审计、智慧审计以及大数据审计等多样化研究视角<sup>①</sup>,相关研究还缺乏系统性、完整性、严谨性理论思维,尚未形成清晰的理论体系,缺乏系统性的研究框架。

本文主要对大数据审计相关文献进行梳理,主要贡献有三个方面。第一,梳理大数据审计研究进展,厘清大数据审计概念,阐释大数据审计理论框架,解读大数据审计应用,在数字经济领域和技术创新维度两个层面丰富和拓展新质生产力的研究内涵。第二,利用CiteSpace软件对国内外有关大数据审计的文献进行梳理,厘清国内外研究特征、趋势及存在差异。第三,从理论和实际出发,探讨了目前大数据审计研究的不足,指明了大数据审计的未来研究方向。本文将从以下几个方面对大数据审计相关研究进行述评。第一,从不同视角描述了大数据审计的特征,提出大数据审计的概念,解释了大数据审计思维的理念,并分析了大数据审计流程和大数据审计技术。第二,从数据运转流程角度深入揭示了大数据审计存在的主要问题,进一步深化认识大数据审计模式。第三,分别从理论和实践层面提出大数据审计的未来研究展望。

<sup>①</sup>虽然有不同称谓,但最终都离不开数据获取、处理、分析和运用,即以数据为导向,此外其他称谓或是强调了某方面技术,如人工智能、机器学习、区块链等,或是强调数据采集与管理平台,如云计算,或是强调数据挖掘与运用,如智能决策、机器人流程自动化审计,故我们在文中统一采用“大数据审计”的概念。

## 二、大数据审计研究的文献计量分析

### (一)文献收集

为掌握有关“大数据审计”相关研究进展,首先,本文以“大数据”“审计”为组合关键词在中国知网进行检索,然后以“big bata”“audit”为组合关键词在Web of Science英文数据库进行检索,共检索到中文文献6 505篇,英文文献1 265篇。其次,有关“大数据”应用研究在2011年后频繁出现,故本文重点关注2011年后的有关文献。中文文献仅保留CSSCI合集,英文文献剔除非核心合集,共检索到中文文献226篇,英文文献600篇。最后,本文作者通过阅读每篇文章题目、关键词、摘要部分,进行筛选,人工剔除与研究主题无关的文献。最终,本文共获取407篇(中文184篇,英文223篇)文献作为研究对象。

### (二)文献数量年度分布

从文献年度分布来看(图1),大数据审计相关研究可分为两个阶段。第一阶段,2011—2014年为初始缓慢增长期,中英文核心文献都相对较少,中文文献总量5篇,英文文献总量4篇。第二阶段,2015—2023年为高速发展期,中文核心期刊发文量为179篇,占比97%,在2020年发表量达到顶峰,相关文献发表31篇;英文核心期刊发文量为219篇,占比98%,在2022年发表量达到顶峰,相关文献发表43篇。中英文合计每年平均有二十余篇文章见刊,表明越来越多的学者关注到大数据审计领域并着手开展相关研究。

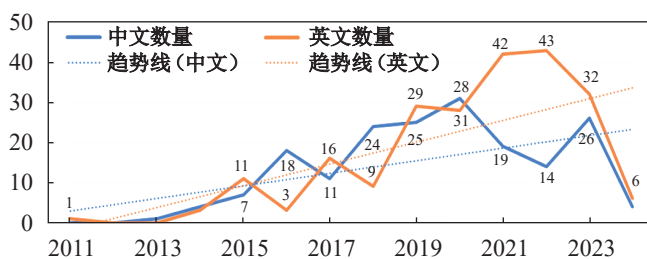


图1 大数据审计研究文献的发表年限统计

### (三)关键词共现分析

采用CiteSpace分析软件,分别对中英文文献进行关键词共现分析,关键词聚类结果如图2和图3所示。国内研究主要聚焦于“国家审计”“政府审计”“国家治理”“审计监督”“安全审核”“人工智能”等方面;而国际研究更侧重于“data analytics”(数据分析)、“artificial intelligence”(人



图2 中文文献关键词共现结果

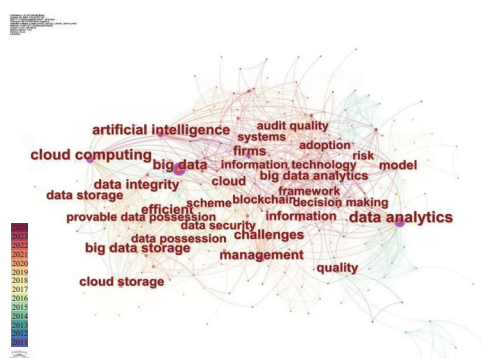


图3 英文文献关键词共现结果

工智能)、“cloud computing”(云计算)等主题。说明我国学者在研究关注点上与国际学者显著不同,究其原因,可能是国内学者更注重大数据审计的功能研究,从国家治理层面出发,探讨大数据技术如何推动审计管理体制机制改革;而国外学者更注重大数据审计技术创新方面的研究。

### 三、大数据审计内涵及相关概念界定

#### (一)关于大数据审计起源

IBM公司于1960年出版了*Audit Encounters Electronic Data Processing*,首次提出电子数据审计,被认为是大数据审计的起源。Alali等(2011)提出计算机辅助审计(CAATTs),被界定为大数据审计的前身。Robert(2003)认为计算机辅助审计技术是使用任何计算机技术助力审计工作顺利完成。审计署关于计算机辅助审计的定义为,审计机关、审计人员将计算机作为辅助审计的工具,对被审计单位财政、财务收支及其计算机应用系统实施的审计<sup>①</sup>。网络和信息技术的快速发展催生持续审计、联网审计,审计频率逐渐加快直至出现持续审计和实时审计,这是大数据审计发展过程中的重要里程碑。随后,大数据审计可以称作是信息化背景下的审计技术发展和应用的高级阶段。Gepp等(2018)认为,大数据审计是大数据应用众多广泛领域中的一个,是将大数据的基本原理和方法运用到审计监督过程中去。

#### (二)关于大数据审计内涵

关于大数据审计的概念,众多学者从不同视角给出见解(如表1所示)。大数据审计具有多对象、多目标、多关系、多时点、多工具、多模式等六个方面特征。大数据审计除了具体审计的根本特征外,又有区别于普通审计的其他特征。大数据审计需要平台为其提供存储空间(刘国城和王会金,2017),大数据审计还具有管理特征,在大数据环境下,数据质量已成为企业成功的关键因素之一(El-Adaileh和Foster,2019)。故大数据审计呈现平台审计、管理审计的特征,通过云计算平台实施的大数据审计能够提升组织的管理能力。大数据审计可以运用多样性数据分析工具,对多个审计客体经济管理活动中产生的过程数据、环境数据、社会行为数据以及物理实体数据进行同时审查,实现跨行业、跨部门和跨区域的综合分析,以综合评判被审计对象经济管理活动的真实性、合规性和绩效。云计算等技术平台是大数据的存在形式,两者深度融合才能使得大数据审计发挥真正效用。大数据审计具有整合特征,需要整合海量结构化、半结构化、非结构化数据。大数据审计过程是资源整合的过程,既包括数据整合,又包括审计资源和力量的整合(Biglari和Pourabedin,2022;Patel和Shah,2023)。通过实施大数据审计,审计工作可以为决策者提供更具有建设性的建议。大数据审计需要技术支持,需要运用与传统审计截然不同的审计技术和方法,如数据挖掘技术、关联比对技术、智能分析技术等(Guo和Liu,2021;Biglari和Pourabedin,2022)。大数据审计是一种风险审计,存在数据泄露的风险和可能,需要重视数据风险的评估与防范(Fan,2022;Huang等,2022)。

表1 大数据审计概念

文献来源	大数据审计概念	思维	技术	特征
Vasarhelyi等,2015;刘星等,2016;Gepp等,2018	大数据审计是大数据广泛应用领域中的一个,是将大数据的基本原理和方法运用到审计行业中。		√	
Salijeni等,2019;王彪华,2020	大数据审计不仅要利用大数据相关技术,还要遵循大数据审计思维,保证大数据审计工作顺利开展。	√	√	
章轲等,2018	大数据审计不仅要运用各种先进技术方法,还体现了审计工作在大数据环境下的全新体现,具有多种特征。	√	√	√

资料来源:作者根据相关文献总结提炼。

<sup>①</sup>审计署《审计机关计算机辅助审计办法》,1997年1月1日实行。

综上所述,本文认为大数据审计是指单个或多个审计主体,由具备大数据分析能力或者掌握新兴信息技术的审计业务人员,依据相关法律法规、条例准则,遵循大数据思维和现代审计理念,对与多个审计对象相关的结构化和非结构化的海量数据进行采集、处理、挖掘、关联分析并将结果进行可视化呈现的新型审计模式。

### (三)大数据审计相关概念界定

目前,关于大数据审计的研究存在多个视角,学术界和实务界也提出人工智能审计、数据导向审计、智慧审计、数智化审计等多种概念。对这些概念进行梳理,有助于深化理解大数据审计的内涵。

关于人工智能审计。罗忠莲和张永杰(2024)认为,人工智能审计是借助RPA与机器学习等新一代关键AI技术、算法与工具所构建的人工智能审计系统,自动执行各项审计业务流程的人机协同审计作业模式。研究表明,实施人工智能审计在推动审计理念创新、降低审计成本以及提高审计质量等方面能够发挥显著作用(Baldwin等,2006;Issa等,2016)。毕秀玲和陈帅(2019)提出“审计智能+”的概念,认为人工智能帮助审计以多样化方式获取数据并高效进行数据分析和结果呈现。林慧涓和陈宋生(2023)探讨了AI审计的必要性,并对已有的多种AI审计框架优缺点进行缜密分析。人工智能审计侧重于研究审计主体如何依赖于AI技术以提升分析能力和效果,实质上是强调大数据审计所开展的自动审计功能。

关于数据导向审计。石爱中和孙俭(2005)较早提出了数据式审计模式,并指出数据式审计的核心方法是数据分析。为适应大数据环境,审计署于2014年设立电子数据审计司<sup>①</sup>,地方审计机构相应设立数据审计相关处室。陈伟等(2016)认为电子数据审计技术是一种用于分析电子数据的计算机辅助技术。秦荣生(2003)基于账表导向审计、系统导向审计、风险导向审计等不同发展阶段,提出并构建数据导向审计体系。他认为,审计需要以数据为导向,从数据中找寻异常和疑点,以证实财务报表中各项目是否存在差错。数据导向审计聚焦数据分析,即对大数据审计的数据分析环节进行研究。

关于智慧审计。鲁清仿等(2018)认为智慧审计有助于解决大数据审计存在的问题,更有效地实施全面、“365天不间断”审计。王文华(2019)认为智慧审计是大数据审计的延伸和升级,其核心在于利用现有信息化发展成果助力审计主体实现从“静态管理”到“动态管理”的转变。黄佳佳等(2023)提出审计理论与审计经验转化为数字化审计知识体系和智能化审计知识模型,并构建智慧审计平台。杨道广等(2022)则认为智慧审计是指将数智技术和场景应用于审计工作中,进而提高审计效率与效果的审计思维的总称,并非全新、独立的审计方法论体系。智慧审计亦是大数据审计的延伸研究,重点放在大数据审计的应用效果层面。

关于数智化审计。张庆龙等(2020)认为数字化审计是智能化审计的基础,数智化审计分为两个阶段,第一阶段需要利用先进的数字技术获取各种信息、数据,构建审计模型,第二阶段结合计算机算力、机器学习等各种技术运用,提供决策选择。杨柔坚(2024)认为数智化审计重点在于数字化与智能化的协同发力,并且数智化审计作为大数据审计的进阶,其对使用的数据、方法、技术都会产生跃迁。数智化作为数字化的高级阶段,侧重审计大模型及其运用,因而,数智化审计更加注重数据分析及其效果的运用。

上述人工智能审计、数据导向审计、智慧审计以及数智化审计等类型审计是新兴信息技术在审计行业的具体应用,研究者较多从数据利用或数据技术应用以及数据运用效果等视角出发,探讨大数据环境下审计行业的发展趋势。这些概念基本上围绕大数据审计某方面特征或具

<sup>①</sup>审计署电子数据审计司, <https://www.audit.gov.cn/n10/n16/n49/c131135/content.html>。

体功能展开研究,揭示大数据分析技术在审计实务中的具体应用,从不同视角认识信息技术发展下的审计变革,丰富和拓展了大数据审计的内涵,深化和拓展了大数据审计理论。

#### (四)大数据审计与传统审计比较

传统审计模式为提高审计效率,通常采用抽样审计,难以全面覆盖到被审计单位业务活动,遗漏业务活动存在的关键问题,不可避免产生较大的审计风险。在大数据环境下,大数据分析、云计算、人工智能等新兴信息技术大大促进了审计工作质量,审计人员可以将繁杂的非结构化数据纳入审计范围,将数据驱动与审计人员的专业判断相结合,不仅提升了审计人员的证据评估和专业判断能力,还大大降低了审计风险(Lohr,2012;Rose等,2017)。

传统审计模式是以会计账簿等财务资料为基础,注重从财务数据上识别风险和疑点,进而展开审计取证(El-Adaileh和Foster,2019;Zhang和Wang,2021)。由于财务报表反映不够全面,难以展示经济管理活动的全貌,导致一些重要的经济管理活动没有留下财务痕迹,给审查带来难题,限制了传统审计模式的成效(Thibodeau,2003)。大数据审计以被审计单位的数据为中心,聚焦财务、业务等结构化数据、半结构化数据以及非结构化数据,进行数据采集、数据存储、数据预处理、数据分析、成果应用,从而识别业务异常和疑点,发现审计线索,深挖严查,进行调查取证,全方位、多层次地发现被审计单位的问题,从而更好地发挥审计职能职责(Cao等,2015;Gepp等,2018;Hezam等,2023)。

大数据审计拓宽了审计范围的宽度、广度与深度,而且延展了审计时间的跨度,向前可以追溯被审计单位以前发展状况,向后可依据大数据分析预测被审计对象的发展趋势,从而提高审计信息的时效性、可靠性和权威性(Biglari和Pourabedin,2022;Shabani等,2022;Patel和Shah,2023)。显然,大数据审计不仅为被审计单位微观层面,而且为经济管理活动的中观乃至宏观层面,提供更合理、更全面、更高质量的审计建议(Cao等,2015;Appelbaum等,2018)。除了发挥通常的验证功能以外,大数据审计还能积极发挥评判政策或制度优劣的检验职能,以及发挥评估未来经营、经济走势的预测职能(Gepp等,2018;Manita等,2020)。然而,也有部分学者对大数据审计持谨慎乐观态度。例如,Alles和Gray(2016)认为,大数据只是达到目的的一种手段,其效果如何取决于审计师选择。还有一些学者主张应该使用大数据来验证审计结果和识别风险,并作为补充审计证据的手段(Brown-Liburd等,2015;Yoon等,2015;Rose等,2017)。

大数据审计状况需要放到特定的发展阶段来衡量,传统审计模式过渡到大数据审计阶段需要一定的条件。例如,国际四大会计师事务所不断探索大数据技术在审计行业中的应用,不断加大在审计技术领域的投资。调查研究表明,国内会计师事务所关于大数据审计的应用仍处于过渡阶段,虽然大数据技术已经成为审计部门开发审计信息系统的主要技术,但事实上只是扩大了抽样范围,审计方法与审计理念并没有改变(张雯等,2023)。

总之,大数据审计模式彻底颠覆以往依赖抽样的传统审计模式,审计对象和范围因数据量与维度而拓展,大量时间、资源耗费在现场取证环节逐渐被后台数据提取、分析所取代,审计工作重点在于数据挖掘和利用,少量时间用于对数据分析得到的异常和疑点的再验证环节上(Hu等,2021;Zhao和Wang,2023)。然而,从传统审计过渡到大数据审计不是一蹴而就的,大数据审计是对传统审计的继承和发展,并围绕审计需求和技术发展而持续改善和变革。

以往研究表明,大数据审计模式与传统审计模式在管理模式、审计内容、工作方式、评价标准、取证模式、审计效果、分析方式以及呈现方式等诸多方面存在差异性,这些差异性能够揭示两种审计模式具有不同的特征、功能和作用(如表2所示)。

表2 大数据审计模式与传统审计模式的区别

特征	大数据审计模式	传统审计模式	代表研究
管理模式	建立平台导向的团队,以平台统一调配资源和管理团队及项目	建立项目导向的团队并配置资源和实施管理	彭冲等,2018
审计内容与 方法	针对结构化、半结构化与非结构化数据,实施全样本审计	针对结构化数据,开展抽样审计	Yoon等,2015;Chiu和Jans,2019
工作方式	非现场的数据分析为主,现场调查为辅	现场调查为主,非现场调查为辅	Cooper等,2019;Barr-Pulliam等,2022
评价标准	审计准则&数据赋能下的专业判断	审计准则&专业判断	Brown-Liburd等,2015;Appelbaum等,2021
审计取证 模式	构建信息系统,收集大量业务、财务数据	查阅和核对账簿凭证等纸质材料	欧雨晴,2024
审计效率	高效	低效	秦荣生,2014
审计质量	相对较高	相对较低	Christ等,2021
分析方法	数据挖掘、机器学习、深度学习等高级分析方法	传统统计分析和样本分析	Zhang等,2015;Commerford等,2022
呈现方式	可视化报告	传统报告	Chang和Luo,2021

资料来源:作者根据相关文献整理。

#### 四、大数据审计研究动态

##### (一)大数据审计思维相关研究

大数据环境下,业务、数据和审计证据有着紧密联系。信息化技术将业务转化成数据,审计人员需要甄别海量数据以获取相关证据。审计人员的专业判断、调查取证都要建立在大数据基础之上,数据的客观性、可验证性以及覆盖的全面性能够反映被审计单位经济活动的真实状况(Appelbaum等,2017;Abu Afifa等,2022)。数据是审计工作的出发点和落脚点,即审计人员需要培养大数据审计思维,围绕数据和数据分析来制定审计方案、实施审计流程以及出具审计报告。大数据审计思维要求审计人员应当养成“用数据说话”的思维,突破原有的审计证据的搜集范围,凝聚“一切皆可为数据”的认识(Mayer-Schönberger和Cukier,2013)。审计工作安排需要与数据相适应,根据现有数据确定审计目标,划定审计范围,组织审计团队。传统的现场取证的审计模式被大量的非现场数据分析所取代,审计人员通常根据数据分析确定审计重点,再去现场加以验证(Biglari和Pourabedin,2022)。这是因为,在数据能够完美呈现真实业务场景的背景下,通过以非现场数据分析替代实际调查研究获取事实证据的方法,往往耗费较少的时间和成本,并大大提升审计效率(Gan等,2021)。大数据审计思维还要求审计人员降低对财务报表、财务数据的高依赖度,将审计范围从狭隘的财务报表转向数据,既关注结构化数据,更要侧重于半结构化数据和非结构化数据,形成完整的证据链条(Pérez和Blasco,2022)。数据质量、数据可获取性、数据分析能力将是影响审计质量的关键因素(Brown-Liburd等,2015;Yoon等,2015)。因此,审计人员需要培育大数据审计思维,跳出传统审计模式的束缚,从更多的维度、更大的广度、更高的深度上理解数据逻辑背后的业务逻辑,不断提高认知能力(Brown-Liburd等,2015;Cao等,2015;Zhang和Wang,2021)。此外,审计对象的确立、审计组织方式的选择以及审计团队的调配,都要适应大数据审计模式的要求。通过合理安排和提前规划,尽可能实现大数据审计的成效。

##### (二)大数据审计流程相关研究

2017年4月,在世界审计组织大数据审计工作组第一次会议上,英国国家审计署提出重点关注大数据审计三个方面的问题,即数据服务(数据清洗、关联、存储)、审计分析(将统计、机器

学习、文本挖掘等大数据技术应用于审计)和可视化(运用数据可视化技术清晰简明地呈现和传播分析结果)。罗钢和刘圆圆(2023)以G公司福利费专项审计为例,提出审计流程为审计准备阶段(包括确定数据采集范围和所使用的大数据分析工具)、审计实施阶段(包括构建审计模型、数据采集、数据处理与分析)和审计完成阶段。黄佳佳等(2023)提出基于智慧审计平台的智能化审计流程,包括确定审计疑点、确定审计方法、确定审计数据源、获得数据源、反馈审计结果及可信度、生成审计报告等环节。秦荣生(2023)认为实施远程联网审计是数据导向审计的核心,主要程序有数据采集、数据整理、数据挖掘、数据质询、数据报告。侯本忠等(2022)结合工程领域物资审计和联合审计,主要步骤包括数据收集处理、审计模型设计、审计结果显示等。

笔者对N市某区审计局大数据审计实践进行调研,在某部门出纳会计挪用公款案例中,审计人员搜集被审计单位全部银行流水数据,利用大数据技术进行甄别分析,发现对私转账疑点;在交通事故中外伤患者获得肇事方赔偿后仍然通过医保卡结算案例中,将医保数据与司法数据进行比对,锁定疑点;在企业套取科技补助案例中,根据实际考察情况,建立审计模型,筛选并确定异常情况。受调查者普遍认为,大数据审计的成功运用,特别依赖大数据审计思维的指导,需要合理运用专业判断。

综上,目前关于大数据审计流程并无统一看法,且现实大数据审计工作模式会根据实际情况进行调整。结合以上多位学者观点和实际调研情况,本文认为大数据审计流程是数据获取、分析和结果呈现的过程,主要涵盖数据采集、数据处理、数据分析、构建审计模型以及审计结果可视化呈现等五个步骤(如图4所示)。第一步,基于审计目标,在专门的审计软件或信息技术支持下,审计人员调取被审计单位及其相关的结构化、半结构化和非结构化数据(Yoon等,2015)。第二步,根据审计目标,审计人员对获取的原始数据进行甄别、筛选、分类、整理,初步建立满足审计目标的数据库(Meijer和Bierman,2011)。第三步,根据业务流程和数据特征,构建有针对性的分析模型,缜密分析已有数据库。第四步,根据模型分析结果,研判并确定数据疑点、异常情形和风险点。第五步,运用可视化技术将审计结果呈现出来(Biglari和Pourabedin,2022;Patel和Shah,2023)。可见,大数据审计贯穿被审计单位业务发生全过程、全流程,打通被审计单位与其业务关联各方,真正做到对业务纵向、横向两个层面的全覆盖,从传统的事后审计跳变到动态审计、实时审计和持续审计。

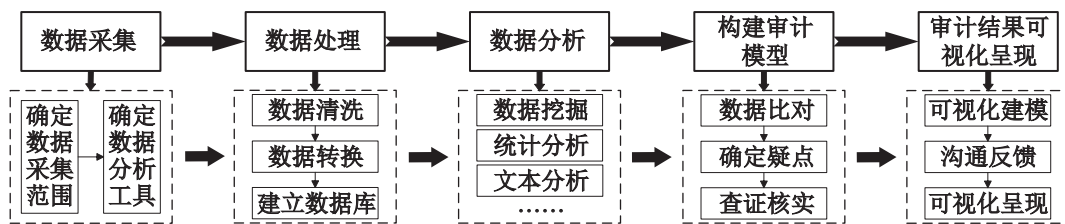


图4 大数据审计流程图

### (三)大数据审计技术相关研究

大数据审计技术的运用,需要把握大数据发展的趋势,识别大数据审计的特征,理解大数据审计在数据采集、传输、存储、处理分析等环节上带来的变化,从大数据运行平台、数据标准、数据管理以及数据应用等方面不断加大对大数据审计技术的研发投入(Appelbaum等,2017)。大数据审计技术如表3所示。

在数据采集阶段,通常需要根据审计目标,利用ODBC数据库接口、网络爬虫等新兴信息技术,连接被审计单位信息存储系统以及互联网系统的数据。为尽可能获得较为全面的数据,



数据采集既要收集被审计单位的数据,也要收集与被审计单位的关联单位(如合作方、供应商、购买方等)的数据,并通过数据相互印证,做出正确研判。然后对数据进行安全存储、加工处理,保证数据质量。借助区块链技术的去中心化、不可篡改性、可追溯性等特点,使得审计平台信息共享,保证审计数据的真实性(Lai等,2023),通过网络直接进行数据传输和存储,减少了审计数据泄露和被篡改的风险(Chedrawi和Howayeck,2018)。在数据分析阶段,综合运用传统数据关联分析、大数据挖掘、自然语言处理、机器学习、模式识别以及可视化呈现等技术方法,发现业务疑点,实现审计目标。审计人员可以使用数据挖掘技术助力审计业务(Brown-Liburd等,2015);可以使用自然语言处理技术分析繁杂的非结构化数据(Qiu等,2020);可以运用机器学习避免传统审计系统缺陷(Manera等,2013)。Gray和Debrecheny(2014)讨论了数据挖掘技术作为一种财务报表审计工具和诉讼中电子证据的必要组成部分的日益增长价值。Al-Hashedi和Magalingam(2021)对2009年至2019年与检测金融欺诈(银行欺诈、保险欺诈、财务报表欺诈和加密货币欺诈)相关研究进行回顾,发现有34种数据挖掘技术被用于识别金融应用中的欺诈。在数据应用阶段,利用人机交互图、动态图、3D图形等数据可视化技术,高质量呈现审计结果,并将其体现在审计报告、审计整改报告和案例展示中(Chedrawi和Howayeck,2018;Shabani等,2022)。

表3 大数据审计技术

大数据审计技术	具体审计流程运用	代表研究
网络爬虫、API调用	数据采集	黄佳佳等,2023
云储存(如Amazon S3、Google Cloud Storage) 分布式文件系统(如HDFS) Hadoop生态系统 分布式计算框架(如Apache Spark) 区块链技术	数据存储与处理	Chedrawi和Howayeck,2018; Han等,2023;Lai等,2023
数据挖掘工具(如Python的Scikit-Learn、R语言、Weka) 人工智能技术(如应用深度学习、自然语言处理)	数据挖掘与分析	Manera等,2013;Brown-Liburd等,2015;Qiu等,2020; Dai和Zhu,2022
统计分析、数据可视化工具(如Tableau、Power BI、人机交互图、动态图、3D图形、知识图谱) 流式处理技术(如Apache Kafka、Apache Flink) 脚本编程语言(如Python、Shell脚本)	数据应用	Higginbotham等,2021; Biglari和Pourabedin,2022; Eberhard,2023;Patel和Shah,2023

资料来源:作者根据相关文献整理。

#### (四)大数据审计应用相关研究

##### 1.关于数据采集

在数据采集环节,主要是基于审计目标,精准定位数据源,并采取适当的方法加以获取。大数据审计首先面临和需要解决的问题便是“需要什么数据”和“从哪里获取数据”,即审计人员需要搞清楚数据范围和来源。审计人员需要采集数据并将其转换成数据库,同时对数据质量加以验证。为验证数据处理技术的有效性,已有学者进行相关实验,设计了大数据采集系统的仿真模型,开发了仿真模拟工具(Colombo等,2016)。为解决数据的真实性、完整性(Lee等,2018)、一致性与时效性(Fan等,2014),学者们提出了不同的技术算法。

目前,数据的生产过程缺乏健全的法律规范,使得数据的生产、储存、加工、传输、管理、交易以及使用等环节缺乏规则,阻碍和影响数据的流通,也影响着大数据审计的开展。被审计单位往往以数据的保密性和安全性为理由,拒绝提供相关数据,或者延迟采集、记录重要数据,或者故意不录入关键数据,从而对审计产生不利后果(Al-Ghanem和Hegazy,2011)。大数据审计

需要建立整体和统一思维,从一定时期和一定范围内考察经济体的业务绩效,需要结合与被审计单位存在业务关系和经济利益联系的相关利益者的数据,才能全面考察和评判被审计对象的经济管理活动。然而,这些数据散落在相关政府部门、供应商、客户、债权人以及投资者的数据库里面,如果没有一个统一的数据信息平台,就需要分别从这些利益相关者的数据库里提取,增加了提取难度和获取成本(Wahdain等,2018)。

云计算、区块链等技术是大数据运行的载体,可以有效存储和管理大数据资源,是确保大数据发挥作用的关键,为大数据审计形成强有力的技术支撑(Sookhak等,2018;Chedrawi和Howayeck,2018)。然而,目前很多地区尚未建立区域性、分层次的大数据审计数据平台,数据和数据管理呈现零碎化状态,基于数据的信息孤岛普遍存在,数据获取和管理成本较高,审计模式基本上延续“一对一”,即单个审计主体对应具体的被审计单位或者项目,在根本上制约着大数据审计的有效开展。

## 2.关于数据分析

在数据分析环节,主要运用关联数据比对、数据挖掘、机器学习以及趋势预测等方法,对数据进行分析和处理。AICPA(2017)指出,大数据审计技术主要运用于数据分析领域,涉及“分析、建模和可视化,识别异常情形,并提取被审计主体相关数据中的其他有用信息”,以规划或执行审计程序。Eilifsen等(2020)也认为,审计运用BDA可以更好识别财务数据中异常交易、模式和趋势。审计数据分析技术还可以减少数据输入和操作等手动、耗时任务的需求,从而保障审计师专注于更复杂和更具判断性的领域,并提高不同审计业务运用审计程序的一致性和标准化,减少审计判断中的可变性和偏差。使用大数据分析可以使审计机关不再局限于具体的审计项目,数据分析结果甚至可以成为审计立项的依据来源(郑伟等,2016)。在财务报表舞弊分析、持续经营能力评估、会计估计等业务中,大数据分析可以为审计人员的职业判断提供基于数据分析的有力支撑。英国财务报告委员会在2015年审查了英国六大审计公司审计数据分析的使用情况(FRC,2018),发现所有公司都在审计数据分析能力(硬件、软件或技能)上进行了大量投资,并且都将审计质量作为审计数据分析实施的主要驱动因素。同样,加拿大注册会计师协会(2017)就审计中运用BDA情况对来自大、中、小型公司的394名审计师进行了调查,发现BDA在审计的所有主要阶段都有使用。对于审计中越来越依赖数据分析工具的情形,Salijeni等(2019)就此对审计性质和审计质量产生的实质性影响表示忧虑。基于审计中使用多个级别和各种类型的数据分析工具,监管机构认识到大数据审计正成为发展趋势,需要洞察这种趋势的潜在影响以及制定额外的技术使用标准(PCAOB,2017,2020;IAASB,2018)。大数据审计技术的运用受到审计主体、审计客户以及监管机构等多方需求的影响,是多种利益博弈的结果。

虽然数据分析是大数据审计的核心环节,但目前在数据分析中大数据审计思维还较为缺乏。第一,大数据理念还没有被审计机构、审计人员普遍接受,实务界审计人员并未深刻理解大数据理念,仍然受到传统审计以“财务数据、财务报表”审查问题和疑点的思维的困扰,犯以偏概全的错误。第二,传统的事后审计模式存在审计成效滞后的严重缺陷,需要通过实施大数据审计以向全流程审计模式转变(Pérez和Blasco,2022)。第三,传统审计模式以某一个单位、某一个项目、某个时点为特征,缺乏全局思维和整体观。第四,需要努力提升大数据审计技术的应用。例如,深度机器学习、知识图谱、语义引擎、预测分析等技术还未在审计工作中广泛应用(Ait-Mlouk和Jiang,2020)。究其原因,多数国家审计机构应对大数据审计的实施缺乏充分的准备,例如,大多审计机构尚未设立负责实施大数据审计的专门机构;缺乏实施大数据审计需要的技术人才,并且多数审计机构严重缺乏技术人才招聘、培训等方面保障机制,导致大数据审

计技术不能适应审计需要(Salijeni等,2019)。

### 3.关于数据应用

在数据应用环节,主要运用数据可视化技术发现业务疑点并呈现审计结果。数据可视化技术包括文本、网络、时空数据以及多维数据等。传统的数据审计方法一般采用Excel、数据库自带的可视化呈现图表来呈现分析结果(Kosara,2016;Eberhard,2023)。利用可视化技术,可以帮助审计人员更快速地发现海量数据中存在的关联关系,从而发现审计疑点或风险点,提高审计工作效率(Higginbotham等,2021),但数据可视化有可能受到过度自信、锚定偏见等认知偏见的不良影响(Chang和Luo,2021)。

大数据审计的结果是高质量使用数据分析结果。数据分析不仅要从微观层面针对特定审计客体提出审计整改建议,更要从中观甚至宏观层面考量问题背后的深层次原因,提出体制机制方面的对策建议,提升审计价值。大数据审计还会产生诸多风险,相关国家或政府尚未采取必要的应对措施。大数据审计面临的风险主要有制度风险、数据安全风险和 data 质量风险,而关于风险防范的应对措施较为缺乏(Salijeni等,2019;Biglari和Pourabedin,2022)。制度风险是指目前缺乏与大数据审计相关的法律法规、审计准则以及操作指南等。大数据审计需要依据相关的规章制度来保证数据采集、处理、存储、分析、应用各个流程按规范进行(Appelbaum等,2018)。数据安全风险是指数据的采集、处理、存储等过程中会存在泄露隐患;数据安全风险又分为主观泄露风险以及客观上数据被盗取的风险,数据管理不善会导致个人隐私、商业秘密乃至国家机密的泄露,不可避免地造成个人、企业、国家的利益损失(Mehmood等,2016)。数据质量风险是指数据的准确性、真实性、可靠性难以保证。首先要判断审计平台、信息系统自身的可靠性,再对数据进行分析,检查数据的真实性、合理性(Appelbaum等,2018)。

## 五、研究结论与未来研究展望

### (一)研究结论

本文基于大数据审计产生动因,系统梳理大数据审计内涵并对相关概念进行辨析,然后从大数据审计思维、大数据审计流程、大数据审计技术、大数据审计应用四个方面系统梳理大数据在审计行业应用的相关研究,最后提出本文的研究展望,研究框架如图5所示。

大数据审计模式代表新兴信息技术和数据科学与审计的融合,是对传统审计模式的变革,它深化了人们对于审计职业和审计业务流程的认识,改善了审计程序,有效整合、利用了审计资源,将审计与数据、数据分析以及数据分析结果运用紧密地联系起来,加强了审计在经济管理活动中的作用。本文的意义在于系统回顾了大数据审计的理论和应用研究状况,揭示了大数据审计应用中存在的问题,并对未来理论和实践方面的研究提出展望。本文认为,大数据审计模式是一种先进的审计模式,对于克服传统审计模式的弊端方面确实能够发挥积极的作用,同时,这种审计模式的变迁,实质上是从科技方面考察、认识以及实施审计活动,并大大减少审计人员主观判断带来的审计质量低下的现象。总体而言,大数据审计既显著改善了审计独立性,又大大提高了审计人员的专业胜任能力,最终审计质量得到明显改进。本文也存在一定的不足之处。由于缺乏必要的相关文献,本文未能系统研究大数据审计发挥作用的内在机理,未能充分揭示大数据审计不同发展阶段的显著特征以及存在的主要问题,未能缜密分析审计人员运用大数据审计技术搜集证据以实现目标的过程。

### (二)未来研究展望

大数据分析在审计中的广泛运用在很大程度上深刻影响着审计主体、客体和监管机构,也极大地拓宽、拓展了审计功能。基于业务数据化、数据信息化、信息数字化以及业务、数据和信

息交叉融合的背景,为进一步丰富和拓展大数据审计的相关研究,结合该领域的实践发展,本文对未来研究提出了以下四点展望。

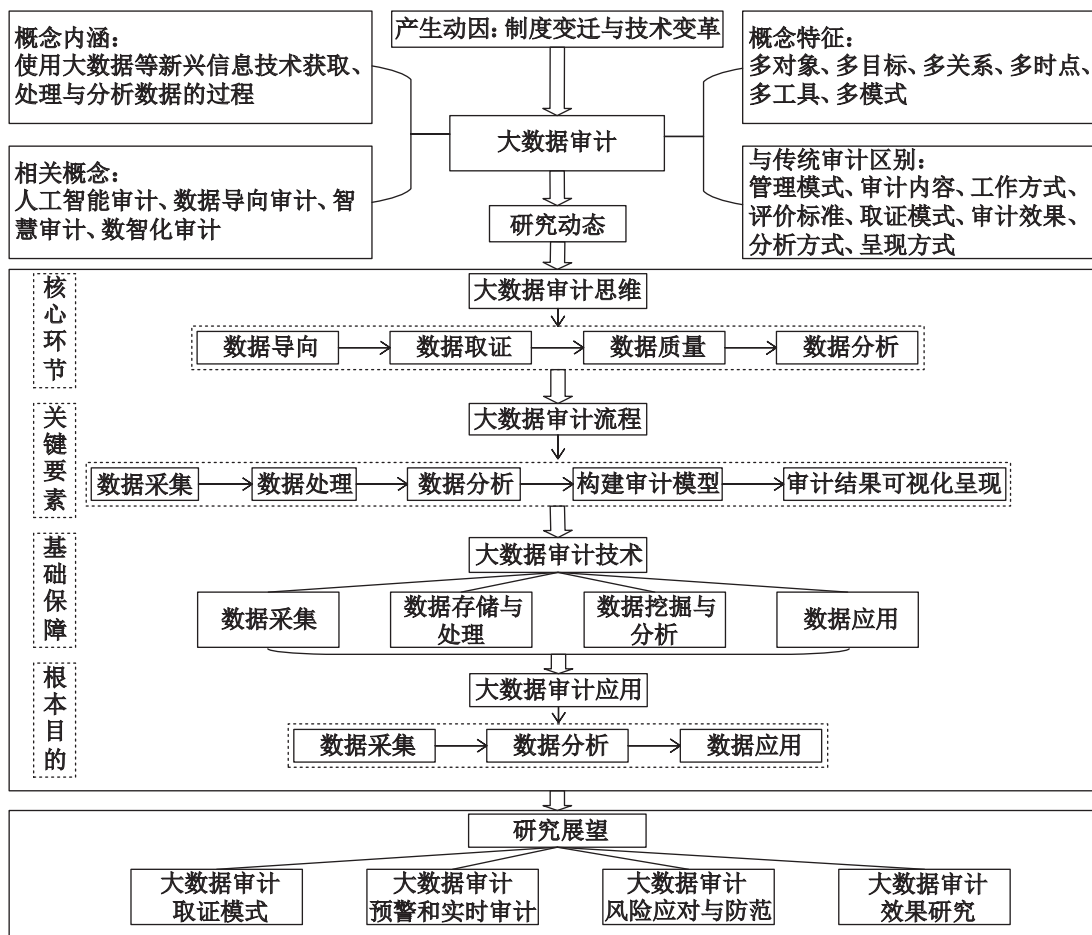


图5 大数据审计研究框架图

### 1. 大数据审计取证模式研究

在BDA信息可以补充审计客户的传统财务信息(Warren等,2015;Yoon等,2015;Appelbaum,2016;Huerta和Jensen,2017)背景下,依赖BDA的审计证据的合法性、可信性日益受到各界关注并成为迫切需要解决的关键问题。未来研究可以紧密围绕“从哪里获取数据”“如何获取数据”“如何管理数据”“如何使用数据”等焦点问题展开研究,研究如何进一步规范大数据审计取证行为,确保数据的真实性、完整性、一致性和时效性。在此基础上,学者们可以揭示大数据审计取证模式的内在机理和运行规律,探讨与宏观政策、区域协同发展以及组织绩效更加高效衔接的大数据审计管理制度,探索绩效大于成本和风险的适合的符合我国国情的大数据审计取证模式。这些研究无疑会促进大数据审计法治建设,深化认识大数据审计,有利于挖掘大数据审计需求。

### 2. 大数据审计预警和实时审计研究

考虑到商业实体的逐步数字化,全自动化和半自动化数字化审计流程是不可避免的(Kogan等,2019)。审计可以运用BDA分析商业银行贷款、养老金预警、财务困境和股票市场预测(秦荣生,2014;张永杰和罗忠莲,2017;Gepp等,2018),实时跟踪和监测风险并预警。如何针

对不同场景研究、开发和使用大数据审计预警系统,实时监控经济管理活动,是值得研究的另一个重要课题。未来可以结合国家战略,利用机器学习、BDA、大模型等新兴信息技术,从宏观经济管理视角探索如何构建大数据审计预警机制和系统,揭示并规避宏观经济运行过程中存在的风险,深化研究实时审计理论体系,发挥大数据审计的战略作用。从微观实体经营管理视角探索特定行业、企业集团、单个企业以及供应链、价值链如何建立问题导向、目标导向的大数据审计预警系统,对经济业务中的风险点进行预警,保障企业组织管理目标的实现。

### 3.大数据审计风险应对与防范机制研究

大数据审计新增数据管理、过分信赖BDA以及违规取证等合规性风险。从保障条件来看,需要研究如何构建高质量的大数据审计平台,以高效地完成审计任务(Bank, 2021; Kontogeorga和Papapanagiotou, 2023)。针对数据管理方面,需要研究审计机构如何做好数据平台的建设和管理,如何运用新兴信息技术构建多维度、多层次、跨部门的数据中心和数据中台,为大数据审计提供强有力的支撑(Appelbaum等, 2017; El-Adaileh和Foster, 2019; Jiang, 2021)。基于大数据审计思维,需要探讨有效的大数据审计组织方式,研究多个行业、组织或部门的多个项目同时实施大数据审计的体制机制和有效路径。在实践层面,需要研究大数据审计相关法律法规、审计准则和操作性指南,指导标准化数据的生产与获取利用,规范大数据审计流程,研究大数据平台如何基于违规、低效和损失风险构建风险清单并建立预警机制。

### 4.大数据审计效果研究

在制度层面,需要研究大数据审计降低代理成本的内在机理和作用机制,以帮助建立更加稳定的委托代理关系。在业务数据化、数据信息化、信息数字化以及业务、数据和信息交叉融合形成数智化的趋势愈来愈明显的背景下,需要研究如何透过信息流、数据流和数字流以审视被审计单位的业务流,即研究并建立风险与收益均衡、净收益为正的大数据审计方法体系,夯实大数据审计基础(Yoon等, 2015; Chedrawi和Howayeck, 2018)。在经验研究领域,需要研究大数据审计的替代变量,建立相关实证模型,系统检验大数据审计作用机制、路径和效果。

## 主要参考文献

- [1]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.
- [2]林斌,林红,孟小涵,等.基于大数据驱动范式的政策跟踪审计模型及其应用研究[J].*审计研究*, 2023, (1): 17-26.
- [3]刘国城,王金会.大数据审计平台构建研究[J].*审计研究*, 2017, (6): 36-41.
- [4]刘星,牛艳芳,唐志豪.关于推进大数据审计工作的几点思考[J].*审计研究*, 2016, (5): 3-7.
- [5]彭冲,胡重辉,陈希晖.大数据环境下的数据式绩效审计模式研究——以x市智慧停车规划与管理项目绩效审计为例[J].*审计研究*, 2018, (2): 24-31.
- [6]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, (6): 23-28.
- [7]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.
- [8]王彪华.大数据审计理论与实践研讨会综述[J].*审计研究*, 2020, (2): 52-56.
- [9]吴秋生,郭檬楠.国家审计督促国企资产保值增值的功能及其实现路径——基于十九大关于国企与审计管理体制改革要求的研究[J].*审计与经济研究*, 2018, 33(5): 12-20.
- [10]谢志华,程恺之.新技术与审计方法的变革[J].*审计研究*, 2023, (1): 3-11.
- [11]章轲,张冬霖,梁轩瑞,等.大数据审计中要做到的“三个把握”[J].*审计研究*, 2018, (5): 30-34.
- [12]张庆龙,邢春玉,芮柏松,等.新一代内部审计:数字化与智能化[J].*审计研究*, 2020, (5): 113-121.
- [13]张雯,杨奥,张敏.大数据技术在会计师事务所中的应用现状研究——基于会计师事务所的调查[J].*中国注册会计师*, 2023, (12): 53-61.
- [14]张永杰,罗忠莲.养老基金云审计系统的构建[J].*审计研究*, 2017, (3): 45-52.

- [15]郑伟, 张立民, 杨莉. 试析大数据环境下的数据式审计模式[J]. 审计研究, 2016, (4): 20-27.
- [16]Al-Ghanem W, Hegazy M. An empirical analysis of audit delays and timeliness of corporate financial reporting in Kuwait[J]. *Eurasian Business Review*, 2011, 1(1): 73-90.
- [17]Alles M, Gray G L. Incorporating big data in audits: Identifying inhibitors and a research agenda to address those inhibitors[J]. *International Journal of Accounting Information Systems*, 2016, 22: 44-59.
- [18]Appelbaum D. Securing big data provenance for auditors: The big data provenance black box as reliable evidence[J]. *Journal of Emerging Technologies in Accounting*, 2016, 13(1): 17-36.
- [19]Appelbaum D, Kogan A, Vasarhelyi M A. Big data and analytics in the modern audit engagement: Research needs[J]. *Auditing: A Journal of Practice & Theory*, 2017, 36(4): 1-27.
- [20]Appelbaum D, Showalter D S, Sun T, et al. A framework for auditor data literacy: A normative position[J]. *Accounting Horizons*, 2021, 35(2): 5-25.
- [21]Appelbaum D A, Kogan A, Vasarhelyi M A. Analytical procedures in external auditing: A comprehensive literature survey and framework for external audit analytics[J]. *Journal of Accounting Literature*, 2018, 40(1): 83-101.
- [22]Baldwin A A, Brown C E, Trinkle B S. Opportunities for artificial intelligence development in the accounting domain: The case for auditing[J]. *Intelligent Systems in Accounting, Finance and Management*, 2006, 14(3): 77-86.
- [23]Barr-Pulliam D, Brown-Liburud H L, Munoko I. The effects of person-specific, task, and environmental factors on digital transformation and innovation in auditing: A review of the literature[J]. *Journal of International Financial Management & Accounting*, 2022, 33(2): 337-374.
- [24]Brown-Liburud H, Issa H, Lombardi D. Behavioral implications of big data's impact on audit judgment and decision making and future research directions[J]. *Accounting Horizons*, 2015, 29(2): 451-468.
- [25]Cao M, Chychyla R, Stewart T. Big data analytics in financial statement audits[J]. *Accounting Horizons*, 2015, 29(2): 423-429.
- [26]Chang C J, Luo Y. Data visualization and cognitive biases in audits[J]. *Managerial Auditing Journal*, 2021, 36(1): 1-16.
- [27]Chiu T, Jans M. Process mining of event logs: A case study evaluating internal control effectiveness[J]. *Accounting Horizons*, 2019, 33(3): 141-156.
- [28]Christ M H, Emmett S A, Summers S L, et al. Prepare for takeoff: Improving asset measurement and audit quality with drone-enabled inventory audit procedures[J]. *Review of Accounting Studies*, 2021, 26(4): 1323-1343.
- [29]Colombo T, Fröning H, Garcia P J, et al. Optimizing the data-collection time of a large-scale data-acquisition system through a simulation framework[J]. *The Journal of Supercomputing*, 2016, 72(12): 4546-4572.
- [30]Commerford B P, Dennis S A, Joe J R, et al. Man versus machine: Complex estimates and auditor reliance on artificial intelligence[J]. *Journal of Accounting Research*, 2022, 60(1): 171-201.
- [31]Cooper L A, Holderness Jr D K, Sorensen T L, et al. Robotic process automation in public accounting[J]. *Accounting Horizons*, 2019, 33(4): 15-35.
- [32]Dagilienė L, Kloviėnė L. Motivation to use big data and big data analytics in external auditing[J]. *Managerial Auditing Journal*, 2019, 34(7): 750-782.
- [33]Dzuranić A C, Mălăeşcu I. The current state and future direction of it audit: Challenges and opportunities[J]. *Journal of Information Systems*, 2016, 30(1): 7-20.
- [34]Earley C E. Data analytics in auditing: Opportunities and challenges[J]. *Business Horizons*, 2015, 58(5): 493-500.
- [35]Eberhard K. The effects of visualization on judgment and decision-making: A systematic literature review[J]. *Management Review Quarterly*, 2023, 73(1): 167-214.
- [36]El-Adaileh N A, Foster S. Successful business intelligence implementation: A systematic literature review[J]. *Journal of Work-Applied Management*, 2019, 11(2): 121-132.
- [37]Fan W F, Geerts F, Tang N, et al. Conflict resolution with data currency and consistency[J]. *Journal of Data and Information Quality (JDIQ)*, 2014, 5(1-2): 6.
- [38]Ferraz C, Finan F. Electoral accountability and corruption: Evidence from the audits of local governments[J]. *American*

- [Economic Review](#), 2011, 101(4): 1274-1311.
- [39]Gan Q Q, Wang X M, Li J W, et al. Enabling online/offline remote data auditing for secure cloud storage[J]. [Cluster Computing](#), 2021, 24(4): 3027-3041.
- [40]Gepp A, Linnenluecke M K, O'Neill T J, et al. Big data techniques in auditing research and practice: Current trends and future opportunities[J]. [Journal of Accounting Literature](#), 2018, 40(1): 102-115.
- [41]Han H D, Shiwakoti R K, Jarvis R, et al. Accounting and auditing with blockchain technology and artificial intelligence: A literature review[J]. [International Journal of Accounting Information Systems](#), 2023, 48: 100598.
- [42]Hezam Y A A, Anthonysamy L, Suppiah S D K. Big data analytics and auditing: A review and synthesis of literature[J]. [Emerging Science Journal](#), 2023, 7(2): 629-642.
- [43]Hu K H, Hsu M F, Chen F H, et al. Identifying the key factors of subsidiary supervision and management using an innovative hybrid architecture in a big data environment[J]. [Financial Innovation](#), 2021, 7(1): 10.
- [44]Huang L, Zhou J Y, Lin J C, et al. View analysis of personal information leakage and privacy protection in big data era—based on Q method[J]. [Aslib Journal of Information Management](#), 2022, 74(5): 901-927.
- [45]Huerta E, Jensen S. An accounting information systems perspective on data analytics and big data[J]. [Journal of Information Systems](#), 2017, 31(3): 101-114.
- [46]Issa H, Sun T, Vasarhelyi M A. Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation[J]. [Journal of Emerging Technologies in Accounting](#), 2016, 13(2): 1-20.
- [47]Jiang S. Research on big data audit based on financial sharing service model using fuzzy AHP[J]. [Journal of Intelligent & Fuzzy Systems](#), 2021, 40(4): 8237-8246.
- [48]Kogan A, Mayhew B W, Vasarhelyi M A. Audit data analytics research—an application of design science methodology[J]. [Accounting Horizons](#), 2019, 33(3): 69-73.
- [49]Kontogeorga G, Papapanagiotou A. Auditing ethics and corruption: Old challenges and new trends for supreme audit institutions in turbulent times[J]. [Journal of Public Budgeting, Accounting & Financial Management](#), 2023, 35(4): 474-492.
- [50]Kosara R. Presentation-oriented visualization techniques[J]. [IEEE Computer Graphics and Applications](#), 2016, 36(1): 80-85.
- [51]Lai C Z, Wang Y Z, Wang H, et al. A blockchain-based traceability system with efficient search and query[J]. [Peer-to-Peer Networking and Applications](#), 2023, 16(2): 675-689.
- [52]Lee K M, Lee K M, Lee S H. Remote data integrity check for remotely acquired and stored stream data[J]. [The Journal of Supercomputing](#), 2018, 74(3): 1182-1201.
- [53]Lindquist E A, Huse I. Accountability and monitoring government in the digital era: Promise, realism and research for digital-era governance[J]. [Canadian Public Administration](#), 2017, 60(4): 627-656.
- [54]Lombardi D, Bloch R, Vasarhelyi M. The future of audit[J]. [JISTEM-Journal of Information Systems and Technology Management](#), 2014, 11(1): 21-32.
- [55]Lonsdale J. Undertaking citizen-focused audit: Problems and practices[J]. [The International Review of Administrative Sciences](#), 2007, 73(1): 131-136.
- [56]Manera V, Schouten B, Verfaillie K, et al. Time will show: Real time predictions during interpersonal action perception[J]. [PLoS One](#), 2013, 8(1): e54949.
- [57]Manita R, Elommal N, Baudier P, et al. The digital transformation of external audit and its impact on corporate governance[J]. [Technological Forecasting and Social Change](#), 2020, 150: 119751.
- [58]Meijer E, Bierman G. A co-relational model of data for large shared data banks[J]. [Communications of the ACM](#), 2011, 54(4): 49-58.
- [59]Patel S, Shah M. A comprehensive study on implementing big data in the auditing industry[J]. [Annals of Data Science](#), 2023, 10(3): 657-677.
- [60]Pérez L F R, Blasco Á R. A Data science approach to cost estimation decision making-big data and machine learning: Un enfoque de ciencia de datos para la toma de decisiones en la estimación de costes-big data y aprendizaje automático[J]. [Revista de Contabilidad-Spanish Accounting Review](#), 2022, 25(1): 45-57.

- [61]Qiu X P, Sun T X, Xu Y G, et al. Pre-trained models for natural language processing: A survey[J]. *Science China Technological Sciences*, 2020, 63(10): 1872-1897.
- [62]Rose A M, Rose J M, Sanderson K A, et al. When should audit firms introduce analyses of big data into the audit process?[J]. *Journal of Information Systems*, 2017, 31(3): 81-99.
- [63]Salijeni G, Samsonova-Taddei A, Turley S. Big data and changes in audit technology: Contemplating a research agenda[J]. *Accounting and Business Research*, 2019, 49(1): 95-119.
- [64]Sookhak M, Yu F R, Zomaya A Y. Auditing big data storage in cloud computing using divide and conquer tables[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2018, 29(5): 999-1012.
- [65]Sun T, Alles M, Vasarhelyi M A. Adopting continuous auditing: A cross-sectional comparison between China and the United States[J]. *Managerial Auditing Journal*, 2015, 30(2): 176-204.
- [66]Vasarhelyi M A, Kogan A, Tuttle B M. Big data in accounting: An overview[J]. *Accounting Horizons*, 2015, 29(2): 381-396.
- [67]Warren Jr J D, Moffitt K C, Byrnes P. How big data will change accounting[J]. *Accounting Horizons*, 2015, 29(2): 397-407.
- [68]Yoon K, Hoogduin L, Zhang L. Big data as complementary audit evidence[J]. *Accounting Horizons*, 2015, 29(2): 431-438.
- [69]Zhang J, Yang X S, Appelbaum D. Toward effective big data analysis in continuous auditing[J]. *Accounting Horizons*, 2015, 29(2): 469-476.

## Big Data Audit: Theoretical Framework, Research Progress, and Future Prospects

Xu Ronghua, Zhu Jing, Dai Xinyu

(*Business School, Ningbo University, Ningbo 315211, China*)

**Summary:** Big data audit is a novel auditing mode that applies emerging information technology to auditing business activities. In recent years, with increasing attention to big data audit, exploring the value creation mechanisms and pathways of big data audit mode has become an important topic in both academic and practical realms. However, existing research on big data audit has not yet formed a clear theoretical system and lacks a systematic framework. Therefore, this paper reviews literature on big data audit published in authoritative journals domestically and internationally from 2011 to 2023. Firstly, it systematically expounds on the origin, conceptual connotation, and comparative analysis of big data audit. Secondly, based on the analytical framework of “core link–key element–basic guarantee–fundamental purpose”, it comprehensively sorts out the research dynamics of big data audit thinking, process, technology, and application, revealing the internal logic of big data audit and supplementing relevant evidence with empirical research. Finally, it constructs a theoretical framework for big data audit research and proposes future prospects from different perspectives such as audit forensics mode, audit early-warning mechanism, audit risk-prevention mechanism, and audit effectiveness, aiming to provide reference for further promoting research in big data audit.

**Key words:** big data audit; big data audit mode; big data audit process; literature review

(责任编辑:王 孜)