

文章编号: 2095-2163(2021)06-0117-07

中图分类号: G353.1

文献标志码: A

基于文献计量的国内溯源技术研究及应用态势分析

田晓丹^{1,2}, 陈振国^{1,2}, 陈 岑^{1,2}, 任维康^{1,2}

(1 华北科技学院 计算机学院, 河北 廊坊 065201; 2 河北省物联网监控工程技术研究中心, 河北 廊坊 065201)

摘要: 溯源技术对快速确定疫情传播的途径及源头具有重要作用,也是确保产品最安全有效的方法之一。本文以“知网”数据库为样本数据来源,运用文献计量法和可视化软件 CiteSpace,通过对年度发文量、作者数量变化、作者、机构以及关键词共现图谱分析,对溯源、溯源技术以及溯源应用3个领域进行了分析研究。归纳分析了溯源技术的研究及应用的历程,通过梳理分析溯源在国内的研究及应用情况,旨在为广大学者及决策机构提供参考,期望能够为国内疫情的防控及产品的管控提供参考。

关键词: 溯源; 溯源技术; 溯源应用; CiteSpace; 可视化分析

Domestic traceability technology research and application situation analysis based on document metrology

TIAN Xiaodan^{1,2}, CHEN Zhenguo^{1,2}, CHEN Cen^{1,2}, REN Weikang^{1,2}

(1 School of Computer Science, North China University of Science and Technology, Langfang Hebei 065201, China;

2 Hebei Engineering Technology Research Center for IoT Data Acquisition & Processing, Langfang Hebei 065201, China)

【Abstract】 Traceability technology plays an important role in rapidly determining the route and source of epidemic transmission and is also one of the most effective methods to ensure product safety. This paper takes "CNKI" database as the sample data source and uses bibliometric method and visualization software CiteSpace to analyze the three fields of traceability, traceability technology and traceability application by analyzing the annual number of articles published, the change of the number of authors, and the co-occurrence graph of authors, institutions and keywords. This paper summarizes and analyzes the research and application process of traceability technology. By sorting out and analyzing the domestic research and application of traceability technology, it aims to provide reference for the majority of scholars and decision-making institutions, and hope to provide reference for the prevention and control of domestic epidemic and product management and control.

【Key words】 origin; traceability technologies; traceability applications; CiteSpace; visual analysis

0 引言

随着科学技术的不断发展及人们生活水平的不断提高,产品来源是否安全成为社会各界人士关注的热点问题。早在上世纪,溯源就应用于动物产品的生产中^[1],1997年欧盟为解决“疯牛病”问题而建立并不断完善优化的食品安全管理制度,被认为是溯源的源头。溯源,即追本溯源,意思是指追溯事物发生的根源^[2]。最初,溯源作为术语用于描述艺术作品的所有权和位置^[3]。对产品安全溯源可有效减少和预防供应链各环节可能出现的问题,当其中一个环节出现问题,追溯源头便可迅速找到问题并有效治理^[4]。2016年国务院办公厅出台《关于推进重要产品信息化追溯体系建设的指导意见》^[5],为全面建设产品追溯标准体系提供了科学全面的政策

支持。在此背景下,利用溯源追溯特定目标找寻数据的源头,保证目标产品完整安全成为众多科研人员攻克的学术课题。

科学知识图谱以知识域为对象,可将知识群之间诸多复杂冗乱的关系变成直观的知识图^[6],本文采用文献计量法和 CiteSpace 软件,对国内权威期刊上的文献进行统计分析,剖析溯源研究领域的研究现状,探究溯源的研究热点以及未来的方向,以期从宏观上对溯源领域做一个系统的综述,为广大学者继续开展领域内的研究提供参考。

1 数据来源与研究方法

1.1 数据来源

为确保数据来源的充分合理,本次研究拟在中国知网(CNKI)进行相关文献的检索。为保证研究

基金项目: 国家重点研发计划重点专项课题(2018YFC0808306);河北省重点研发计划项目(19270318D)。

作者简介: 田晓丹(1995-),女,硕士研究生,主要研究方向:情报分析、溯源技术;陈振国(1976-),男,博士,教授,硕士生导师,主要研究方向:科技情报、数据分析;陈 岑(1996-),男,硕士研究生,主要研究方向:情报分析、溯源技术;任维康(1999-),男,硕士研究生,主要研究方向:情报分析、溯源技术。

通讯作者: 陈振国 Email: 200600744czg@ncist.edu.cn

收稿日期: 2021-04-06

对象的唯一性,本文仅以期刊论文作为分析对象。具体为:在“中国知网”启用“高级检索”功能,检索条件为“主题=溯源”、来源类别勾选“SCI来源期刊”、“EI来源期刊”、“北大核心”、“CSSCI”、“CSCD”。检索时间为2020年10月2日。

同时,为了保证数据质量,剔除非学术论文以及与本文研究主题相关度不高的文献。具体为:对溯源领域进行整体分析时,发表时间设置为从1992~2020年,共计得到符合要求的文献1890篇。因自2014年以后文献数量呈不规则摆动态势,故针对2014~2020年进行详细分析,即发表时间设置为从2014~2020年,再次检索后,共得到符合要求的文献1166篇。

1.2 研究方法

本文采用文献计量法进行梳理,用CiteSpace5.7.1软件作为知识图谱可视化分析软件,对目标研究领域的文献年度发文量、发文作者以及发文机构数据进行统计并绘制图谱,根据统计结果产生的可视化图谱总结归纳该领域的研究热点和趋势。与此同时统计整理高频被引、下载次数较多的文献,介绍相关的发展情况。

2 溯源文献分析结果

2.1 1992~2020年整体分析

文章发表数量在一定程度上反映了某领域的研究热度以及重视程度^[6]。图1为1992~2020年10月国内文献发表数量的具体情况。

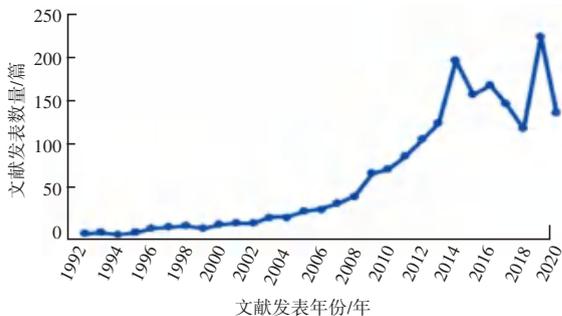


图1 1992~2020年度论文的发表数量变化趋势

Fig. 1 Trends in the number of publications of annual traceability research papers from 1992 to 2020

由图1可知,溯源研究领域的发文数量经历了增长、摆动、增长3个阶段,整体呈上升趋势,具体情况如下:

1992~2013年是溯源发展的增长阶段。1992年李元良^[7]在详细介绍葡萄根瘤蚜进的基础上,针对其在烟台的具体传播进行了系统综述,标志着国

内溯源研究正式步入正轨。到了2009年,文献增长速度明显加快,文献[8-10]是2009年溯源领域被引前3的文献,说明将溯源应用到了食品行业是该年的热点研究内容。同时,通过可追溯性的技术保证食品的安全性,将溯源研究发展向前推动了一大步。2009年后,年度发表文献的数量增长更加明显,上升趋势变快,溯源研究领域进入快速增长期。其中1992~2008年为增长缓慢期,2009~2013年为增长快速期。2014~2018年是溯源发展的摆动阶段,该研究领域进入瓶颈期,但仍不断有学者在溯源技术及溯源应用等方面取得突破。2014年杨信廷等人^[11]在分析了溯源技术的基础上,对当前存在的问题提供了解决办法。该文被引高达111次,下载量为2956次,说明该文所述的内容被广大学者充分肯定及新的研究思路和方向。文献[12-14]较高的被引量也说明在2014年中溯源领域研究内容广,在食品、化学、药材和环境方面均有所涉猎,同时广泛的研究使2014年的发文数量达到新的高度。

从图中可以看出,2014年是溯源领域重要的发展节点。之后,文献发表量成摆动态势,相比于2014年,整体呈现下降趋势。2019年,由于区块链等新兴技术在溯源领域的应用,溯源领域重新吸引了研究者的关注,该年度发文量达到一个新的高峰。

2.2 2014~2020年详细分析

2.2.1 文献发表数量

2014~2020年溯源研究年度论文发表数量的统计如图2所示。



图2 2014~2020年溯源研究年度论文发表数量及变化趋势

Fig. 2 Trends in the number of publications of annual traceability research papers from 2014 to 2020

可以看出,2016年发文量较前一年有所增长,但幅度很小;2018年发文量降到最低,仅有121篇;2019年发文量激增,比上一年增长近1倍。文献[15-17]反映出在2019年区块链作为新兴技术的加入,使溯源领域冲破低潮,达到了新的高点。同时也说明,在区块链的影响下,溯源研究范围在现有的基础上逐步扩大。

在 2014~2020 年时间段的 1 166 篇论文中,有 6.34% (74 篇) 论文针对产地溯源进行描述分析,3.94% (46 篇) 论文针对溯源系统进行研究,对大数据的研究分析占总数的 0.77% 共 9 篇,说明在该时间段产地溯源是该研究领域的核心研究。进一步对产地溯源进行了解和分析,发现产地溯源的研究主要集中在食品行业,在产地溯源被引次数前 10 的文献中占比达到 80%,其中,发表在核农学报上 3 篇、中国粮油学报和食品科学上各 2 篇,反映出上述 3 个期刊在产地溯源领域乃至食品研究领域占有重要地位。

2.2.2 发文作者分析

2.2.2.1 年度作者数量分析

作者是溯源研究领域的直接主体。作者的人数变化以及单个作者的发文量,可反应出溯源研究领域的传播变动范围^[18]。图 3 显示了 2014~2020 年参与溯源研究的学者数量变化情况。



图 3 2014~2020 年参与溯源研究的学者数量变化趋势

Fig. 3 Trends in the number of scholars participating in traceability research from 2014 to 2020

由图 3 可以得出,2014 年共有 729 名学者研究溯源并发表论文。之后,参与溯源研究的学者数量不断减少,2018 年成为该时间段内最低靡的一年(仅有 569 人)。值得关注的是,2015~2016 年论文发表量虽增加了 11 篇,但参与溯源研究的学者数量却减少了 16 人,说明在摆动期学者们对溯源领域的关注度已经趋于稳定,对溯源的研究已进入瓶颈期。而 2019 年发文作者人数达到顶峰(是前一年的 0.6 倍)。正如前文所述,新兴技术吸引了大批学者的关注,从而给溯源领域注入了新鲜血液,同时也说明溯源研究并未形成完整体系,需要研究人员结合现阶段的政策以及发展趋势,发掘该领域更多的可能性。

2.2.2.2 作者分布图谱

在 CiteSpace 内选择时间跨度为 2014~2020 年,时间切片(单个时间区分长度)设置为 1a,类型选择作者“Author”节点。生成的溯源研究领域的主要发

文作者图谱如图 4 所示。

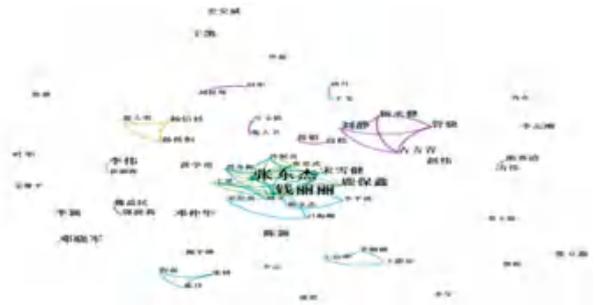


图 4 溯源研究领域的主要发文作者图谱

Fig. 4 A map of main authors in the field of traceability research

如图 4 所示,在溯源研究领域的主要发文作者图谱中,节点表示该领域的文献作者,节点的大小代表该作者出现的频次,各节点之间的连线代表对应作者之间的合作。从图 4 中可以看出,网络中节点分布较为零落分散,大多数作者为独立的节点,其次是单线状的连线,独立“三角形”与“四边形”连线更是较少。该图中共有 191 个节点(N),250 条连线(E),网络密度(Density)仅为 0.013 8,说明到目前为止该研究领域各作者间的合作网络密度较低,各作者以及团队间的联系与合作较少,大部分学者仍处在独立研究该领域的状态,长此以往不利于溯源研究的发展,跨团队间的合作有待加强。

进一步分析发现,张东杰、钱丽丽等共 13 人组成了溯源研究领域的最大作者合作节点,检索可知仅冷候喜来自黑龙江省杜尔伯特县伊利乳业有限公司,其余 12 人均来自黑龙江八一农垦大学,也就是说该团队是由校内外共同合作构成。表 1 中统计的溯源领域前 10 位高产作者中,前 4 位学者均为该团队成员,反映出在 2014~2020 年中,以张东杰为首的黑龙江八一农垦大学研究团队是溯源研究领域的核心发展力量,同时也说明黑龙江地区对溯源研究发展有较大贡献。

表 1 2014~2020 年高产作者统计表

Tab. 1 Statistics of High-yield Authors from 2014 to 2020

| 序号 | 作者 | 发文数 | 被引数 | 下载数 |
|----|-----|-----|-----|-------|
| 1 | 张东杰 | 16 | 147 | 4 941 |
| 2 | 钱丽丽 | 14 | 121 | 3 957 |
| 3 | 宋雪健 | 8 | 69 | 2 277 |
| 4 | 鹿保鑫 | 7 | 68 | 2 831 |
| 5 | 刘静 | 7 | 30 | 832 |
| 6 | 邓晓军 | 5 | 1 | 500 |
| 7 | 李伟 | 5 | 4 | 61 |
| 8 | 杨永健 | 4 | 26 | 775 |
| 9 | 李颖 | 4 | 12 | 772 |
| 10 | 管骁 | 4 | 34 | 1 111 |

根据表 1 可知,在该时间段内前 10 位作者发文

数占总发文数量的6.35%,说明在溯源研究领域学者呈现出“少数集中,整体分散”的态势。

2.2.3 发文机构分析

发文机构是溯源的间接主题,能反映出溯源领域的空间分布情况。将溯源研究领域的1166条检索结果导入CiteSpace软件中,选择时间跨度为2014~2020年,时间切片设置为1,选择机构为“Institution”类型,生成溯源研究领域的发文机构图谱如图5所示。



图5 溯源研究领域的发文机构图谱

Fig. 5 A map of publishing organizations in the field of traceability research

与图4溯源研究领域的发文作者图谱相同,发文机构出现次数越多,在图谱中的相应节点就越大,各机构间的合作越密切,在图谱中的连线就越粗。该图中共有110个节点,49条连线,网络密度仅为0.0082,反映出溯源研究的各机构间学术交流较少。结合表2可知,出现频次最多的是中国计量科学研究院,该机构隶属国家市场监督管理总局,是现阶段溯源研究领域的权威机构。在频次最高的前十名发文机构中,高校与社会研究机构各占一半,说明溯源研究领域发展潜力大,势头良好。但目前发展速度并不理想,需要广大学者投入更多的时间与精力。

2.2.4 溯源关键词共现

关键词和主题词共现(均可称为共词),是指利用文献集中词汇对或名词短语共同出现的情况,确定该文献集所代表学科中各主题之间的关系^[19]。将关键词作为节点,用连线表示各关键词之间的关系,构建关键词图谱,能够直观清晰的发现文献之间存在的关系和当前研究的交叉点。

表2 2014~2020年发文机构统计表

Tab. 2 Statistics of issuing organizations from 2014 to 2020

| 序号 | 机构 | 频次 |
|----|----------------------|----|
| 1 | 中国计量科学研究院 | 24 |
| 2 | 中国科学院大学 | 19 |
| 3 | 黑龙江八一农垦大学 | 14 |
| 4 | 首都医科大学公共卫生学院 | 13 |
| 5 | 上海海洋大学信息学院 | 6 |
| 6 | 中国检验检疫科学研究院 | 6 |
| 7 | 北京市疾病预防控制中心 | 6 |
| 8 | 江苏省地理信息资源开发与利用协同创新中心 | 5 |
| 9 | 武汉大学信息管理学院 | 5 |
| 10 | 上海市食品药品检验所 | 5 |

可在软件CiteSpace中选择关键词“Keyword”节点类型,Pathfinder寻径网络算法,生成关键词共现图谱,并合并相似词。如,二维码和二维码技术等,具体结果如图6所示。



图6 溯源关键词共现图谱

Fig. 6 Co-occurrence map of traceable keywords

在图6中共有216个节点,224条连接,网络密度为0.0096,与溯源研究领域发文作者以及发文机构的图谱相比较,该图谱呈横纵双向发展态势,整体发展良好。在2014~2020年间,最大的关键节点(出现次数最多)是“溯源”,紧接着衍生出不同分支四散开来,而后分别集中在“产地溯源”、“区块链”、“溯源系统”等方向。说明在这段时间内,这3个方向为溯源领域的热点研究主题。

表3中进一步梳理出了出现频次为10次及以上的各关键词共25个。也就是说,出现在表4中的关键词,代表自2014年至今溯源领域的研究方向与趋势。其中关键词“农产品”、“质量安全”、“溯源”和“计量学”中心度较高,反映出计量学的理论知识为溯源发展创新奠定了坚实的基础,提供了科学的发展依据。

“产地溯源”、“区块链”、“溯源系统”3个关键

词出现频率均为 40 次以上,其中“区块链”出现频次高达 47 次,反映出区块链作为新兴技术的热度。自 2018 年起,关于区块链的内容在溯源中所占的比

例越来越大,说明这三年区块链技术发展的步伐越来越快,同时也预示着将区块链技术运用到溯源中或将成为该领域一个新的研究趋势。

表 3 关键词出现频次统计表

Tab. 3 Keyword frequency statistics table

| 序号 | 关键词 | 频次 | 中心度 | 序号 | 关键词 | 频次 | 中心度 |
|----|-------|-----|------|----|---------|----|------|
| 1 | 溯源 | 169 | 0.42 | 14 | 近红外光谱 | 17 | 0.22 |
| 2 | 产地溯源 | 85 | 0.08 | 15 | 矿物元素 | 16 | 0.11 |
| 3 | 区块链 | 47 | 0.06 | 16 | 大米 | 15 | 0.25 |
| 4 | 溯源系统 | 44 | 0.25 | 17 | 主成分分析 | 15 | 0.19 |
| 5 | 量值溯源 | 39 | 0.17 | 18 | 数据溯源 | 15 | 0.06 |
| 6 | 溯源性 | 30 | 0.18 | 19 | RFID | 14 | 0.05 |
| 7 | 稳定同位素 | 28 | 0.09 | 20 | 不确定度 | 13 | 0.25 |
| 8 | 计量学 | 27 | 0.29 | 21 | 二维码 | 13 | 0.14 |
| 9 | 溯源分析 | 26 | 0.12 | 22 | 脉冲场凝胶电泳 | 12 | 0.26 |
| 10 | 农产品 | 25 | 0.88 | 23 | 大数据 | 12 | 0.08 |
| 11 | 物联网 | 24 | 0.25 | 24 | 重金属 | 11 | 0.03 |
| 12 | 食品安全 | 22 | 0.28 | 25 | 误差溯源 | 10 | 0.00 |
| 13 | 质量安全 | 18 | 0.43 | | | | |

3 溯源技术及应用分析

为了进一步梳理溯源研究领域的发展现状以及未来趋势,以溯源技术及应用为关键词,通过对 1992~2020 年的整体分析及 2014~2020 年的重点分析展开研究。

3.1 溯源技术分析

启用“中国知网”中的“高级检索”功能,检索条件为“主题=溯源 AND 主题=技术”,发表时间设置为从 1992~2020 年,共得到符合要求的文献 464 篇。将所得的数据导入 CiteSpace 中,筛去和溯源技术不相关的词,得到溯源技术共现图谱如图 7 所示。

图 7 中共有 49 个节点,31 条连接,网络密度 0.026 4,各中心节点相互联系,说明在溯源技术领域不同技术之间存在交流合作,不是独立发展。可以看到,RFID、区块链、物联网、稳定同位素、近红外光谱、计量学等 6 个关键词成为溯源技术共现图谱中的关键节点,说明这 6 种技术是溯源技术研究的热点技术同时也是发展较好的专业方法。

此外,在图 7 中节点数少、连线少、但网络密度偏大,说明溯源技术研究在交流中稳步发展。其中“量值溯源”和“区块链”分别出现于 2014 年和 2018 年,在出现频次前 10 中仅占 20%,区块链的突现强度为 13.003 7,遥遥领先于其它关键词,说明区块链作为溯源技术中的热点技术推动着该领域的发展。“智能合约”与“区块链”关联度最高。“智能合约”定义为信息化方式传播、验证或执行合同的计算机协议^[20]。目前应用于以太坊区块链以及 RootStock。可以明显看出,“智能合约”成为近几年溯源技术最值得研究的课题之一。可以预测,新兴智能化技术产业是该领域未来发展方向之一。

3.2 溯源应用分析

启用“中国知网”中的“高级检索”功能,检索条件为“主题=溯源 AND 主题=应用”,发表时间设置为从 1992~2020 年,共得到符合要求的文献 292 篇。将所得的数据导入 CiteSpace 中,筛去和溯源应用无关的词,得到溯源应用关键词共现图谱如图 8

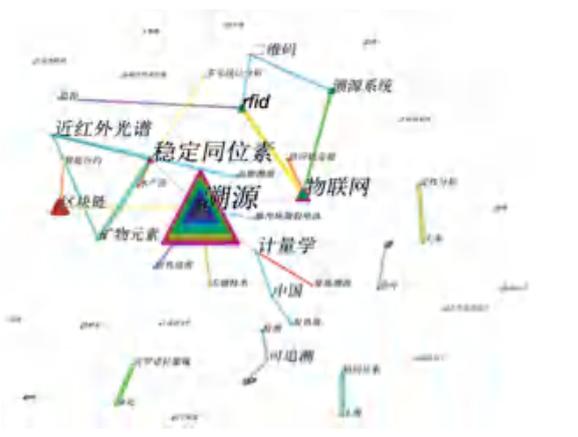


图 7 溯源技术共现图谱

Fig. 7 Traceability technology co-occurrence map

