

鉴往知来：我国电子文件与数字档案研究 态势可视化分析^{*}

李姗姗 胡馨 张笑星

(西北大学公共管理学院 西安 710127)

摘要：电子文件与数字档案是过去20年我国档案学领域的研究热点之一，随着数字技术的加速迭代和数智生态的逐步形成，适时总结其研究脉络与态势有助于鉴往知来，更好地把握档案事业现代化目标导向下的研究方向。为此，本文选取2013-2023年电子文件和数字档案相关国社科基金的论文类研究成果为对象，借助ITGInsight对其进行主题词同现网络、关键词演化及作者、机构、期刊间的关联分析等，在此基础上发现近十年来我国电子文件与数字档案研究逐步向主题内容细分化、作者合多元化、机构合作地缘化发展，此外数字人文成为新的研究前沿。

关键词：电子文件；数字档案；共现分析；演化分析；关联分析

中图分类号：G270 **文献标识码：**A **文章编号：**1005-9652(2023)04-0117-007

0 引言

伴随着5G、云计算、大数据、互联网、物联网等信息技术的创新突破，数据正爆炸性增长，万物皆能产生数据，数据间皆能关联^[1]，信息社会初见端倪。在此背景下，档案记录方式正经历从纸质到电子，从物理空间到虚拟空间，从手工操作到数字化智能化管理，从分散利用到资源集成、联网共享的变革过程^[2]。为此，2023年2月全国档案局长馆长会议以“全面贯彻落实党的二十大精神 奋力书写档案事业现代化和高质量发展新篇章”为主题，强调档案工作要紧紧跟上党和国家事业发展步伐，全面融入中国式现代化的历史进程，聚焦高质量发展，不断创新工作理念、方法和模式，加快推进档案工作现代化建设，以档案工作自身现代化服务推进中国式现代化。电子文件与数字档案一直是我国档案学领域的研究热点，在新目标导向下何以沿承和转向值得透视和探究。为此，本文拟借助文献分析工具ITGInsight，对近十年来我国电子文件和数字档案相关国家社科基金的论文类研究成果进行可视化分析，在梳理电子文件与数字档案研究发展脉络与特征的基础上，透视其新阶段的发展动态与趋势。

1 相关文献回顾

文件与档案有着天然、紧密的联系，数字档案的来源包括电子文件、电子文档、音像视频的数据信息以及各种载体的档案资源数字化后的档案信息。关于电子文件与数字档案的作用，霍奕潼^[3]指出档案数字化可以提升档案借阅的管理质量、实现档案原件的保护、提高档案利用价值以及降低档案管理的费用等。孙妍^[4]则通过论述手工管理档案的缺陷，指出实现档案数字化的必要性；关于电子文件与数字档

*【作者简介】李姗姗(1986-)，女，博士，硕导，西北大学公共管理学院副教授，信息资源管理系副主任，研究方向：档案安全与保护、数字人文与文化遗产、数字档案资源建设；胡馨（2001-），女，西北大学公共管理学院2019级档案学专业学术硕士，研究方向：数字人文与文化遗产、数字档案资源建设；张笑星（2001-），女，西北大学公共管理学院2018级档案学专业学术硕士，研究方向：档案安全与保护、数字人文与文化遗产、数字档案资源建设。

案的管理模式，魏歌^[5]在明确产品数据系统电子文件与数字管理规范的基础上，通过对标准主要内容的介绍与分析，提出了完善术语等相关建议。安小米^[6]等学者对15个国家和1个地区相关法律、法规、政策和标准中电子文件管理的目标与角色定位以及电子文件管理的机制与模式特点进行探析，指出国外在电子文件管理机制方面值得我国学习的经验。龚骏骎^[7]等学者以TOE理论为分析框架，对机构电子文件单套制实施困境进行探讨，提出优化路径。倪代川^[8]等学者聚焦数字档案资源生态安全概念内涵，为大数据时代数字档案资源生态安全建设提供理论指导与决策参考。吕玉洁^[9]等学者以监理角色定位和职责担当作为切入点，提出档案数字化加工的“自管式、代管式、单管式、全管式”4种监理模式。

2 数据来源与研究方法

国内外有关电子文件和数字档案的定义很多，且存在诸多相似概念。其中，关于电子文件我国学者较多采用《电子文件归档与管理规范（GB/T 18894-2002）》中的界定，即“通过计算机等电子设备形成、办理、传输和存储的数字格式的文件，由内容、结构和背景组成。”从范围上包括原生性的电子文件与传统载体文件数字化后的副本；而关于数字档案的概念尚未统一，针对其属概念主要有三种观点：（1）认为数字档案是档案；（2）认为数字档案是文件；（3）认为数字档案是资源。

为确保检全率和检准率，本文在中国知网（CNKI）中以“主题OR篇关摘=电子文件/电子公文/数字档案/数字化档案/档案数字化/电子档案）AND（基金=国家社会科学基金）”为检索式，时间限定于2013年1月1日-2023年1月1日，共检出616条数据，剔除无关文献后最终得到602篇，进而借助文献分析工具ITGInsight进行数据清洗，绘制主题图、关联图、演化图等。

3 主题类聚与演化分析

3.1 主题词同现分析

为展现不同研究主题间的相关性和不同主题的研究热度，本文运用ITGInsight软件绘制我国电子文件与数字档案研究主题分布聚类密度图，见图1。图中主题词颜色深浅变化表示不同主题词的频数情况，主题词之间的距离反映了主题词之间的相关性。

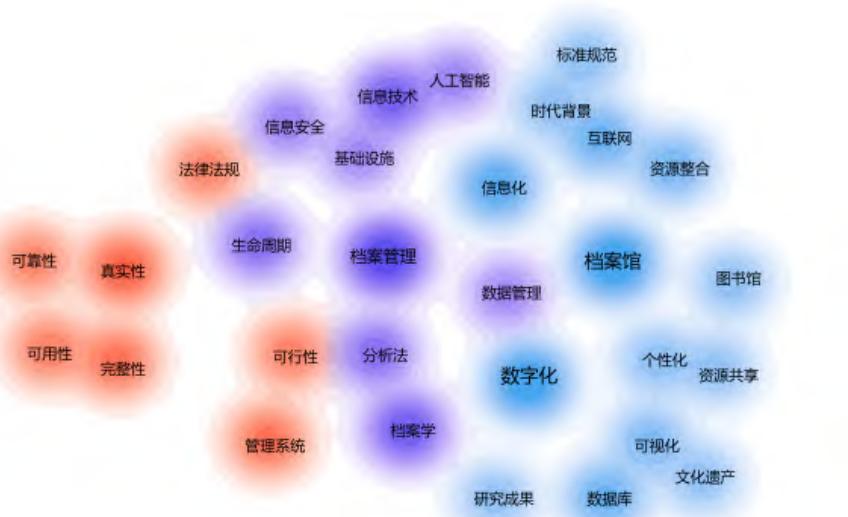


图1 我国电子文件与数字档案主题分布聚类密度图

观察图1可以看出，研究热度较高的主题词有数字化、档案馆、档案管理、信息化、信息技术、互联网等，

这也印证了互联网与计算机等高新技术有力推动了电子文件和数字档案的发展。从主题词内容来看，大致可以将其分为档案数字化基础建设、数字档案资源管理以及数字档案资源管理系统等三个方面。

档案数字化基础建设方面主要涉及数字档案资源的概念、标准规范与数字档案馆建设等。数字档案资源管理则包括信息化背景下档案管理理论完善以及信息安全的维护等。数字档案资源管理系统的研究主要集中在法律法规建设和信息完整性、真实性的维护上，以便长期保存和整合利用。

这三个方面的研究层层深入又相辅相成，在一定程度上体现出我国电子文件和数字档案的建设已经逐步从传统单一的基础建设向内容精细化方向发展。

3.2 关键词演化分析

关键词的演化体现着研究热点的产生，获取研究热点的发展方向对科研学者来说具有前瞻性和参考性。为体现我国电子文件与数字档案研究内容的变化趋势，本文选取年度论文关键词中频次排在前 20 位的关键词，绘制电子文件与数字档案的研究内容演化图，删除孤立节点后如图 2 所示。

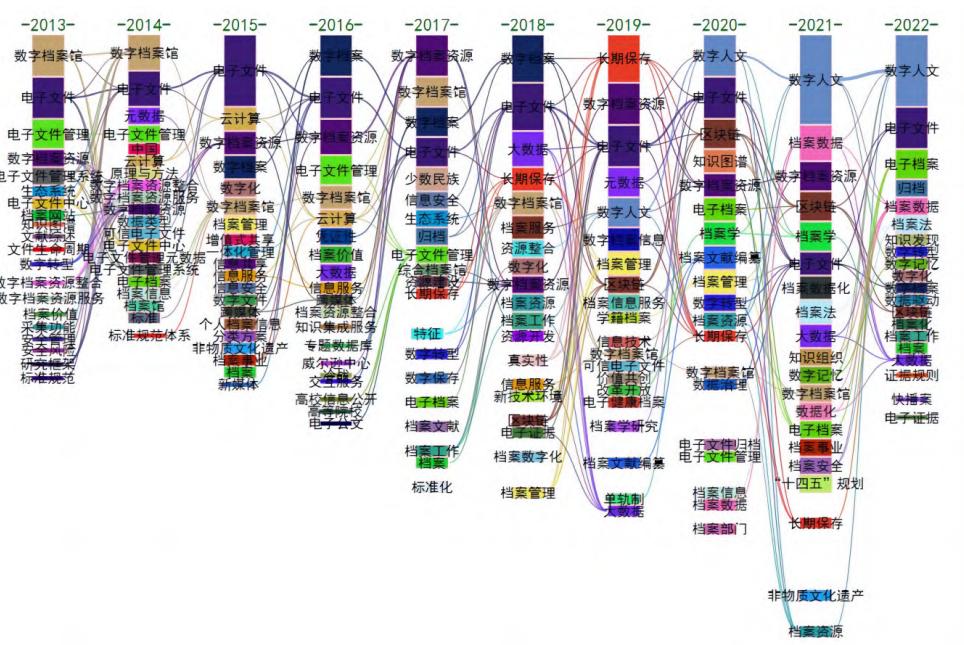


图 2 我国电子文件与数字档案研究内容演化图

从图 2 观察到，在过去十年间，电子文件与数字档案研究热度攀升，研究内容呈现出交替变更现象，电子文件、数字档案、电子文件管理、数字档案馆等关键词贯穿始终，这说明在研究进展中，纸质文件和档案的数字化转型是重点研究方向，数字档案馆是重点研究对象。如程妍妍^[10]等学者从提高数字档案馆运行效率和资源采集、保管和利用质量出发，采用功能分析的方法提出了我国数字档案馆标准规范体系框架；薛辰^[11]则针对档案外包，通过引入信息技术（IT）服务外包中的风险管理理论，提出防范数字档案馆外包项目风险的几点策略与措施。

2013-2016 年处于电子文件与数字档案资源的积累阶段。期间，全国数字档案馆（室）建设推进会中“存量数字化、增量电子化”战略的提出以及《数字档案馆系统测试办法》的印发都有力推动传统载体档案数字化进程。因此，这一阶段研究主要围绕文件与档案的数字化及其管理以及数字档案馆的建设展开，重在将科学技术融入档案的建设，推动电子文件与数字档案建设进程，加快档案基础理论的完善。如陶水龙^[12]等学者分析研究科学的元数据分类方法，进而对分类进行方案设计，为确保电子文件和电子档案的科学管理提供基础支撑。卞咸杰^[13]基于 LINQ 技术实施数字档案检索的性能、内容、过程与步骤、意义，提出 LINQ 技术在数字档案检索中应用的发展趋势。何正军^[14]等学者探讨云计算运用于数字档案馆建设的可行性以及对数字档案馆建设提供的新机遇，进而提出数字档案云的构建设想。

2017-2019年，电子文件与数字档案的研究处于平稳发展阶段。以习近平同志为核心的党中央高度重视信息化发展，加强顶层设计、总体布局，做出建设数字中国的战略决策。随着信息化进程的加快，数字形态的信息种类和数量与日俱增，且因其信息的易变性、系统的依赖性以及信息与特定载体的可分离性等特征，归档保存的电子文件与数字档案的保护需求与难度也日益显现。因此这一阶段的研究内容集中在数字档案资源的信息安全及长期保存等方面。如钱翼^[15]提出应依托OAIS功能模型梳理认证锚点，实施数字档案资源长期保存认证工作。秦巧云^[16]等学者基于大数据技术，指出大数据环境下信息安全问题并提出建立数字档案馆信息安全防范体系。

进入2019年后，数字档案资源内容更加丰富，管理技术更为先进，相关研究进入多元化与智慧化发展阶段，重在对档案信息资源内容进行深度挖掘与利用。习近平总书记在哲学社会科学座谈会上的重要讲话与《新文科建设宣言》的发布引发了文理交融的巨大浪潮，在跨学科思维的引领下，“技术”与“人文”大范围融合，由此数字人文成为领域研究热点，《山西档案》与《档案学通讯》都将其纳入重点征稿范围，国家社科基金也大力资助。如牛力^[17]等学者分析数字人文环境下档案研究路径与方法变迁的依据及映射机制，基于档案价值视角，形成数字人文视角下档案研究的微观技术系统。赵雪芹^[18]等学者以万里茶道档案资料为例，基于本体构建实现档案资源的关联组织与可视化呈现，进而促进档案资源的开发与利用。

4 研究力量分布与关联分析

4.1 期刊关联与演化分析

期刊关联分析不仅可以呈现出电子文件与数字档案研究方面比较权威的期刊，还可以了解其研究主题方面的侧重点。本文选择发文数量排名前 30 位的期刊，构建研究主题关联图（如图 3 所示），选取期刊年度载文量在 3 篇及以上的期刊绘制演化图（如图 4 所示）。

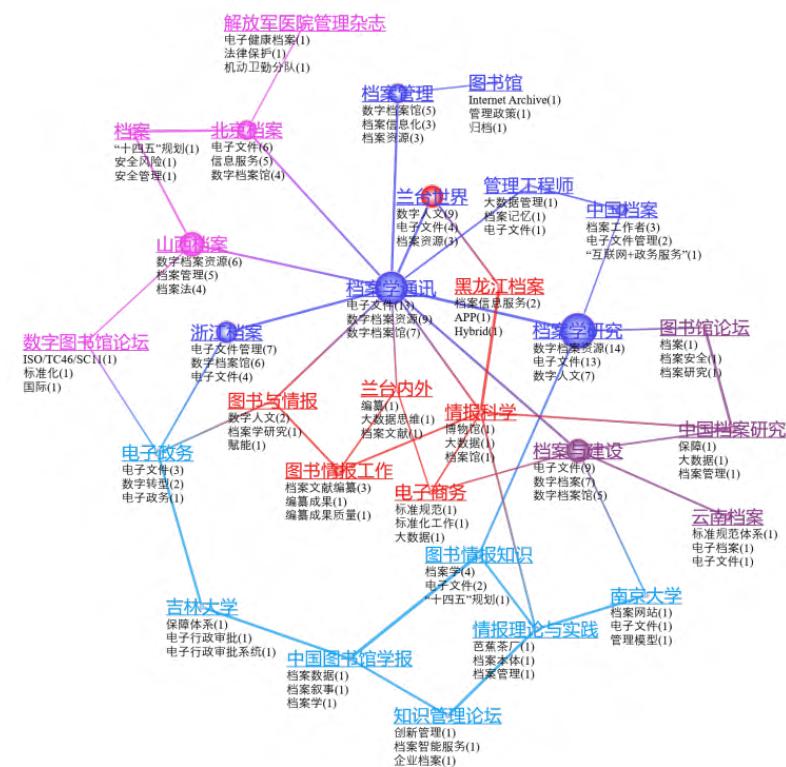


图 3 研究期刊关联图

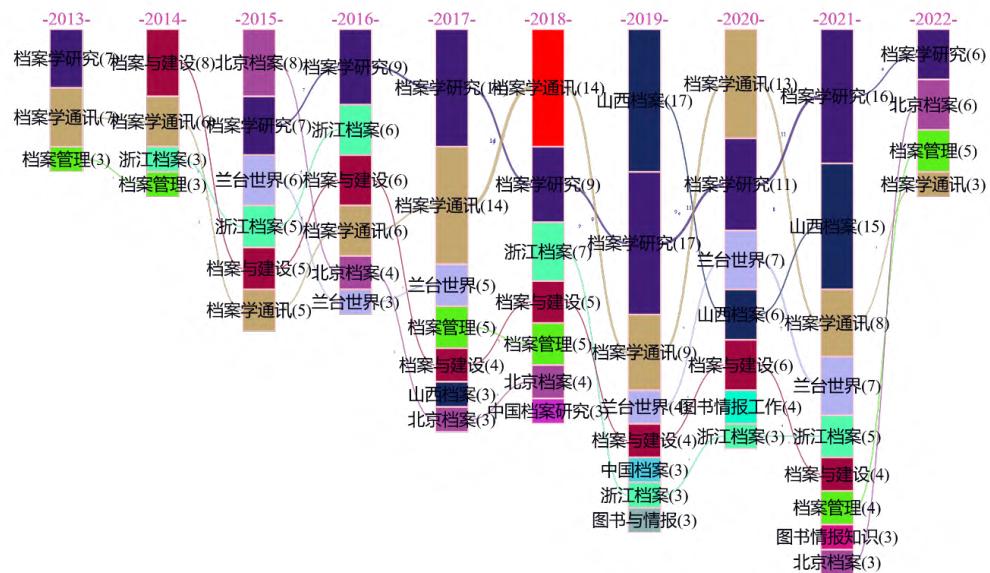


图 4 研究期刊演化图

从图3与图4可以看出，过去十年间刊载电子文件与数字档案研究论文的期刊以档案类为主，图书情报类为辅（如《图书情报工作》），也有少许其他类型期刊（如《电子政务》）。档案研究类期刊偏重于电子文件、数字档案馆，图书情报类期刊偏重于大数据、标准化、档案安全。这与两类期刊的办刊宗旨有一定关系，档案研究类期刊以繁荣档案学理论、开发档案资源为己任，图书情报类期刊恪守理论与实践相结合的宗旨。其中，《档案学研究》《档案学通讯》《山西档案》这三种期刊相关论文数量最多，主要原因是其均为核心期刊，专业性与权威性较高，刊载论文数量多、质量高。

4.2 研究机构关联分析

运用 ITGInsight 软件对研究机构进行关联分析以了解各机构研究内容的侧重以及机构之间研究主题的关联性。选取发文量排在前 20 位的机构生成电子文件与数字档案研究机构关联图, 删除孤立节点后如图 5 所示。图中节点的大小与机构发文数量成正比, 节点间连线的粗细与合著数量成正比。

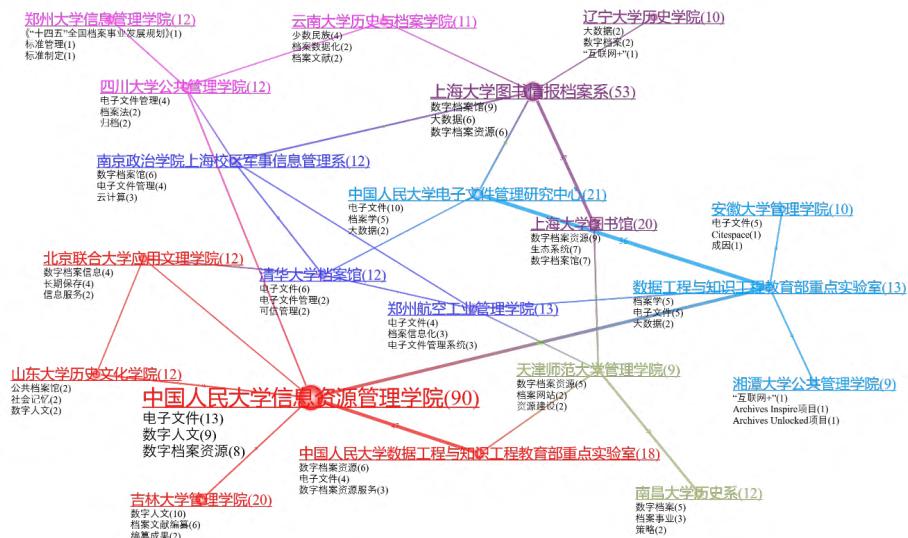


图 5 研究机构关联图

从研究机构发文情况来看，呈现出以高校为主导，研究院所和各类档案馆齐驱并进的态势，且大多集中在北京、上海等东部经济发达地区，中西部地区的发文量还有待提高。其研究内容主要集中在电子文件、数字人文、数字档案资源、大数据、数字档案馆等方面，这在一定程度上反映了过去十年间的主要研究方向。

从研究机构合作情况来看，中国人民大学在科研合作中处于重要的中心地位，与众多研究机构展开了较多的合作。原因之一在于中国人民大学科研实力雄厚，处于电子文件与数字档案研究领域的领先水平。此外，研究机构合作主要呈现出地域相近的特征，说明今后研究可以增强跨地域跨机构的合作交流，以提高研究效率，进而取得更多的研究成果。

4.3 发文作者关联分析

经过对数据文本的统计分析，选取发文数量排在前30位的作者构建研究作者关联图，如图6所示。节点大小与作者发文量成正比，节点间连线粗细与作者间的研究主题关联度成正比。

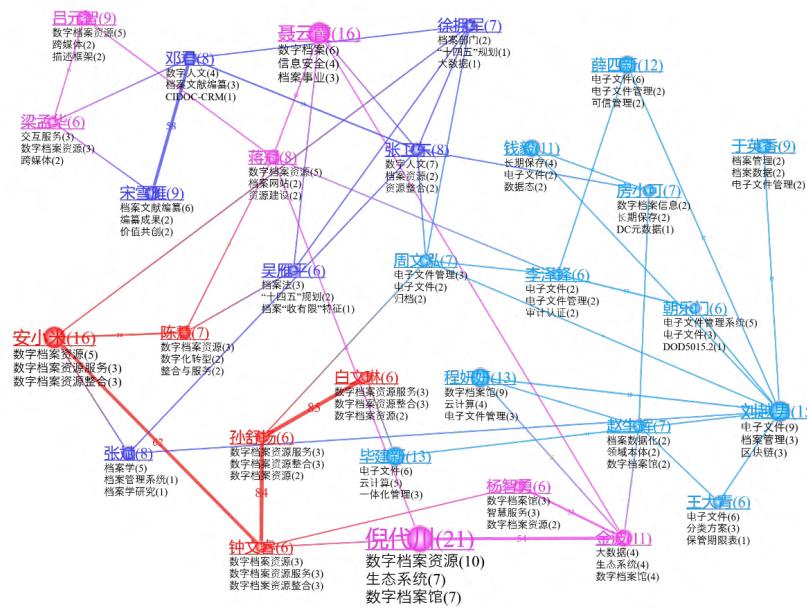


图6 研究作者关联图

从图6可以看出，电子文件与数字档案研究领域中第一群体是以倪代川、金波为中心的上海大学文化遗产与信息管理学院团队，研究主题偏向生态系统、数字档案馆；第二群体是以安小米、钟文睿、陈慧为中心的中国人民大学信息资源管理学院团队，研究主题偏向数字档案资源整合与服务。此外，中国人民大学电子文件管理研究中心的刘越男团队、南昌大学数字人文中心的聂云霞等，也是电子文件与数字档案研究领域里比较突出的研究团队或个人。

分析作者单位属性发现，作者的成果数量与学历职称成正相关。同时，作者之间的合著呈现出地区性、系统性、内部性等特点，机构之间的合著总体不显著，高校内部研究者之间的合著相对显著。

5 结论与展望

本文借助文献计量与可视化方法，对近十年电子文件与数字档案相关国家社科基金项目的论文类研究成果进行分析发现，我国电子文件与数字档案相关研究在政策引导和数字技术驱动下持续性快速发展，并逐步呈现出主题内容细分化、作者合多元化、机构合作地缘化等发展趋势，且数字人文是近年来的显著增长点。但总体而言，学科分布仍相对集中，区域间科研力量差距较大，跨机构合作不显著。随着大数据、数字人文、元宇宙、人工智能等新技术和新概念的融入，电子文件与数字档案研究将不断创新、

发展、融合，未来可以以档案事业现代化和档案工作高质量发展为目标导向，不断拓展研究深度和广度：

(1) 注重学科的交叉与融合，加强跨地区、跨机构、跨学科之间的沟通与合作，促进研究成果的转移转化。

(2) 加强理论与实践的结合，从实际需求与问题出发，挖掘领域研究热点，丰富电子文件与数字档案研究内容。

(3) 激发科研人员的研究动力和创新活力，培养更加多元、融合、可持续的研究力量。

注释与参考文献

- [1] 倪代川，金波. 论数字档案资源数据化发展 [J]. 档案学研究, 2021(5): 17-22.
- [2] 章燕华，王力平. 数字化转型背景下的档案信息化发展战略：英国探索、经验与启示 [J]. 档案学通讯, 2021(4): 28-35.
- [3] 霍奕潼. 档案数字化的必要性及问题解决措施 [J]. 管理观察, 2019(26): 67-68.
- [4] 孙妍. 实现档案数字化管理的必要性及路径选择 [J]. 黑龙江档案, 2018(2): 91.
- [5] 魏歌. 《产品数据管理（PDM）系统电子文件归档与电子档案管理规范》（DA/T 88—2021）的主要内容述评及思考 [J]. 北京档案, 2022(05):21-26.
- [6] 安小米，张宁，叶晗，杜雅楠. 国外电子文件管理机制及借鉴研究 [J]. 档案学研究, 2008(02):58-62.
- [7] 龚骏骎，颜祥林. 电子文件单套制归档实施困境与优化路径 [J]. 浙江档案, 2022(06):51-55.
- [8] 倪代川，薛玥. 数字档案资源生态安全体系论纲 [J]. 山西档案 :1-10.
- [9] 吕玉洁，吴湘华，向禹. 确定性需求驱动的档案数字化加工监理模式构建与适用性分析 [J]. 档案学研究, 2018(1): 107-112.
- [10] 程妍妍，李圆圆. 我国数字档案馆标准规范体系研究 [J]. 档案学通讯, 2014(6): 53-57.
- [11] 薛辰，颜祥林. 数字档案馆项目外包风险防范策略——基于档案馆与承包商之间的双重视角 [J]. 档案学通讯, 2014(5): 83-86.
- [12] 陶水龙，王贞，田雷，等. 电子文件和电子档案元数据分类与方案设计 [J]. 档案学研究, 2016(6): 83-90.
- [13] 卞咸杰. 基于 LINQ 技术的数字档案检索研究 [J]. 档案学研究, 2012(2): 69-72.
- [14] 何正军，金波. 云计算与数字档案馆建设新机遇 [J]. 档案与建设, 2015(12):4-8.
- [15] 钱毅. 基于 OAIS 的数字档案资源长期保存认证策略研究 [J]. 档案学研究, 2018(04):72-77.
- [16] 秦巧云，周枫，杨智勇. 大数据环境下数字档案馆信息安全研究 [J]. 北京档案, 2017(06):18-21.
- [17] 牛力，高晨翔，张宇锋，闫石，徐拥军. 发现、重构与故事化：数字人文视角下档案研究的路径与方法 [J]. 中国图书馆学报, 2021(01):88-107.
- [18] 赵雪芹，李天娥. 面向数字人文的档案领域本体构建研究——以万里茶道档案资料为例 [J]. 情报理论与实践, 2022(08):154-161.

(下转 141 页)

for the archives protection work. On the basis of non-destructive testing of archives, the protection work includes disease treatment, repairing with suitable hand-made paper, partial reinforcement of carbonized parts. These protective measures can extend the life of the archive on the basis of maintaining the original appearance. The summary of restoration work can be enlightening to improve the protection effect of archives in the Republic of China: Repair materials should be selected flexibly; using modern testing equipment; splicing the fragment as far as possible; the restoration process should follow the principle of minimum intervention so as to retain the original appearance of the archives.

Key words: Archive protection; Archives of the Republic of China; Restoration techniques; machine-made paper

(上接 123 页)

Retrospect: Visual analysis of research situation of electronic documents and digital archives in China

LI Shanshan , HU Xin , ZHANG Xiaoxing

(School of Public Management, Northwest University, Xi'an 710127)

Abstract: Electronic documents and digital archives are one of the research hotspots in the field of archives science in China in the past 20 years. With the acceleration of digital technology and the gradual formation of digital intelligence ecology, timely summary of their research context and trend is helpful to learn from the past, and better grasp the research direction under the goal of modernization of archives. This paper selects the research results of national Social science funds related to electronic documents and digital archives from 2013 to 2023 as the object, and uses ITGInsight to analyze the correlation between theme words and present network, keyword evolution and authors, institutions and journals. On this basis, it is found that in recent ten years, the research of electronic documents and digital archives in China has gradually developed into the subdivision of subject content, the diversification of author cooperation, and the geographical cooperation of institutions. In addition, digital humanities has become a new research frontier.

Keywords: electronic file; Digital archives; ITGInsight; Evolutionary analysis; Correlation analysis; Research situation