

政府数据开放研究的国际比较与中国未来发展思路

荆玲玲, 张芸瑶

(哈尔滨师范大学经济与管理学院, 哈尔滨 150025)

摘要: 开放政府运动推进了全球范围内的政府数据开放共享。如何更好利用政府开放数据, 进而推动政治民主和社会进步成为迫切需要讨论的问题。通过文献计量分析探讨自 2000 年以来国内外学者关于政府数据开放的研究脉络、合作情况、研究热点及发展趋势, 运用 CiteSpace 知识图谱结合国内外政府数据开放进行比较研究。研究发现, 国外政府数据开放不同学科融合较好, 对数据的利用程度更高, 已经开始尝试将人工智能应用到政府数据开放平台上。针对所得出的国内外比较结果, 对国内政府数据开放提出以下建议: 加强与人工智能的融合、加强对数据的利用效率与与其他学科系统的融合, 最后由政府端下沉到公民端。

关键词: 政府数据开放; 文献计量学; 数字治理

中图分类号: C39 **文献标志码:** A **文章编号:** 1671-1807(2025)01-0297-09

在大数据时代, 数据被赋予了新的内涵和意义。其已经成为重要的战略性资源, 也是一种新型的生产要素, 政府机构是最大的数据生产者和收集者, 大数据已成为国家竞争的新制高点。数据对于科学、准确地制定国家政策、推进政府数字化转型和社会治理都具有重要意义, 而且能够使政府得到有效监督, 提高公民参与政治的积极性, 及时收纳信息并进一步发掘其潜在的价值具有重要作用。2023 年中国决定组建国家数据局, 专门负责协调推进数据基础设施的建设^[1]。

以美国为首的发达国家率先意识到政府数据开放的价值, 在 2009 年首批上线了“Data.gov”平台, 各个国家紧随其后加入了“政府数据开放运动”这一赛道^[2]。欧洲各国在政府数据开放运动中名列前茅。根据《联合国政务发展报告》显示, 丹麦、芬兰均在排名中位于第一梯队。丹麦自 1968 年就开始进行数字化改革, 作为数字创新的大国, 丹麦今日在政府数据开放研究中仍处于领先地位。芬兰政府则创立了政府开放数据的法律法规。加拿大政府在 2006 年引入了“主动披露”制度, 成立了专门的管理监督机构, 推进数据开放的发展。而国内加入政府数据开放运动这一赛道稍晚于国外, “上海市政府数据服务网”作为国内首个政府数据开放平台在 2012 年上线, 在取得不错的成效之后其余省份

也纷纷上线了政府数据开放平台。目前相对权威的报告是来自复旦大学数字与移动治理实验室发布的《中国地方公共数据开放利用报告》。报告中显示截至 2023 年 8 月已经上线了 22 个省级公共数据开放平台和 200 个城市平台。可以见得国内的政府数据开放工作在这些年飞速发展并富有成效, 在越来越重视数字治理的今天, 政府数据开放在政府数字治理方向还有无限可能。国外的政府数据开放时间早于国内并也在飞速发展。

各国积极推动开放政府数据进展, 目前全世界已经有一百多个国家和地区建立了数据开放平台, 引发学术界的广泛关注。Gao 等^[3]用 Web of Science(WoS)数据库中的期刊论文作为研究对象, 梳理了国外对于开放政府数据的 6 个主要研究主题; Barcellos 等^[4]提出一个名为 Hippolyta 的框架的定义、实施和评估, 该框架有资格帮助公民解释开放的政府数据。目前国内外部分学者采用定量方法针对政府数据开放产生研究, 相比国外, 国内则以文献计量分析法为主。运用多个案例、实证、定量分析等研究方法相对较少, 反观通过分析现状与问题, 并提出针对性建议的定性分析占大多数。本文认为可以通过引入新的研究视角或方法来填补相关研究中尚且留白的内容。利用文献计量法, 对国内外政府数据开放的研究文献进行归纳整理, 重点

收稿日期: 2024-07-03

作者简介: 荆玲玲(1987—), 女, 辽宁东港人, 博士, 教授, 硕士研究生导师, 研究方向为数字政府; 通信作者张芸瑶(2000—), 女, 黑龙江哈尔滨人, 硕士研究生, 研究方向为数字政府。

构建科学知识图谱,并比较和分析其研究前沿方向。以期为政府数据开放领域提供更广阔的视角。

1 数据来源和研究方法

1.1 数据来源

以来源于 WoS 和中国知网(CNKI)数据库所收录的相关核心学术论文作为研究对象。英文文献数据以 WoS 核心合集为数据源,设置检索条件“government open data + open government data + GOP”,时间跨度为 2000—2024 年,经过人工筛选得到与政府开放数据主题相关的文献共 317 篇。中文文献数据以 CNKI 为数据源,主题为“政府数据开放 + 开放政府数据 + 政府开放数据”,不设定检索年限。通过检索获得文献 1 962 篇,为了确保研究的准确性,通过精确的标准筛查所搜集到的数据,文献均经过人工逐篇筛选。

1.2 研究工具

CiteSpace 是用于文献分析的专业可视化工具,本文采用可视化分析软件 CiteSpace 作为研究工具。CiteSpace 能够对知识领域中重要论文的搜索操作进行简化,科学知识图谱显示科学知识的发展进程与结构之间的关系,找出该领域文献的演化进程关键路径、知识拐点以及其他可视化信息,能够体现关键节点及重要连接的聚类视图、体现文献的更新和相互影响的时区视图,可以更好掌握其研究进展与前沿领域,因此本文选择采用此软件。

1.3 研究方法

依照其他学者的经验采用了文献计量与内容分析两种方法。内容分析法是一种以定性为主的、科学的研究方法,从文献的主题内容入手,客观系统地文献加以分析^[5]。文献计量法是一种以定量为主的文献分析手段,指用统计学方法进行统计分析,用数据表示文献的变化规律。这两种方法有其明显的缺点,如文献计量法以论文为研究对象,容易产生偏倚;内容分析法的内容往往不能从数据中直接获得。拓展文献计量法原有的优势并弥补两者的缺陷,这对实现中国社科研究方法的创新具有重要意义。

2 学科知识服务研究现状可视化计量分析

2.1 研究作者分析

通过分析作者可以直观地感受到不同年份该主题的研究热点及演化趋势,对作者的分析有助于清晰地掌握政府数据开放发展的脉络。国外政府数据开放高水平成果产出研究作者最大发文量为 10 篇。进一步分析发现,国外政府数据开放高水平

研究者中有 25 人的发文量在 3 篇以上,其中 Marijn Janssen(10 篇)、Martin Lnenicka(7 篇)主要研究政府数据开放的成熟度领域。Era Ruijter(8 篇)、Albert Meijer(6 篇)主要研究政府开放平台领域,Anastasija Nikiforova(7 篇)、Ibrahim Mutambik(7 篇)、Abdullah Almuqrin(7 篇)主要研究政府数据开放系统质量领域。通过分析发现,国外研究者注重对政府开放数据平台进行评估和效果研究,但占比最多的仍是以 Marijn Janssen 为首的针对政府数据开放领域成熟度的研究。目前围绕平台和效果评估产生了两个集中的研究区域并且这两个区域彼此之间产生了部分联系。依据筛选标准,从所获得的文献中分析出国内政府数据开放高水平成果产出研究作者最大发文量为 46 篇。进一步分析发现,国内政府数据开放高水平研究者中有 29 人的发文量在 5 篇以上,其中陈美(46 篇)、郑磊(18 篇)、梁乙凯(13 篇)、陈朝兵(13 篇)主要研究开放数据、数据开放、用户隐私;黄如花(32 篇)、温芳芳(6 篇)主要研究平台建设、实践创新;翟军(22 篇)主要研究影响因素;马海群(18 篇)、周文泓(17 篇)主要研究生态系统。通过分析发现,国内研究者注重政府数据完善开放平台建设和提升公共服务质量,占比最大的作者是陈美,根据研究显示这几个研究方面已经产生了分别以陈美、黄如花为中心的几个集中研究区域并且彼此之间有着部分联系。

2.2 研究机构分析

通过分析机构可以得出该研究主题分布的主要领域及这些领域彼此产生的联系。对机构的分析有助于清晰地掌握政府数据开放发展的脉络。通过对收录在 WoS 上的文献进行分析,发现有 11 家机构发文量在 5 篇以上。可以见得政府数据开放的研究领域宽泛,应用范围较广。通过对收录在知网上的文献进行分析,发现有 12 家机构发文量在 15 篇以上,包括武汉大学信息管理学院(67 篇)、武汉大学信息资源研究中心(42 篇)、四川大学公共管理学院(27 篇)、复旦大学国际关系与公共事务学院(25 篇)、中南财经政法大学公共管理学院(24 篇)、华中师范大学信息管理学院(23 篇)、湖北工业大学经济与管理学院(20 篇)、黑龙江大学信息管理学院(19 篇)、西南财经大学公共管理学院(19 篇)、中国人民大学信息资源管理学院(17 篇)、电子科技大学公共管理学院(15 篇)、大连海事大学航运经济与管理学院(15 篇)。出现了以武汉大学信息管理学院和复旦大学国际关系与公共事务学

院为最大中心点的两个群组,通过得出的结果显示,政府数据开放在我国的主要研究应用领域都集中在公共管理领域和信息管理领域。

2.3 研究发展演进

通过研究发展演进情况能够对政府数据开放的各个时期发展情况有较全面的掌握,不仅可以清晰地看到事物的发展脉络同时还能预测未来的趋势和走向。通过观察和分析过去的发展过程,可以发现一些规律性的变化,从而预测未来的可能性。这对于政府数据开放的未来发展具有重要意义,同时为后续的发展提供更多的思路。

2.3.1 国外研究发展演进——从传统书面数据到算法创新

国外政府开放数据研究主题词呈现多样性变化特征,关键词共现时间线图谱中前9位聚类分别为大数据(big data)、开放政府(open government)、数据重用(data reuse)、可解释的人工智能(explainable artificial intelligence)、动机(motivations)、感知到的好处(perceived benefits)、开放政府数据(open government data)、公共管理(public management)、记录联动(record linkage)(图1)。通过时间线对国外政府数据开放研究内容进行梳理,可以得出政府数据开放发展的不同阶段时期(表1)。

2007—2014年:研究探索期。管理是2009年最大的节点,以2009年为分界线,学术界开始将目光转移至开放政府数据(open government data,

OGD)相关课题。随着时间的推移,“knowledge management”(知识管理)、“e-government”(电子政务)、“big data”(大数据)等词汇涌现,传统的公务方式正在发生变革,由面对面、电话交谈等方式向电子化方式转变,这期间,电子政府的概念受到了广泛的重视。2013年颁布的《电子政务法》更进一步推动了德国开放政府数据的发展^[6]。当时的研究多围绕电子政务系统建设与实施。

2015—2019年:研究深化期。国外 OGD 领域研究出现多条不同主题的演化路径,如医疗路径如“primary care”(基层医疗)、“treatment”(治疗)。效果与创新领域“determinants”影响因素、“open data policy”(开放创新)、“citizens”“citizen participation(公民隐私路径)”开始浮现。研究领域不断与多学科交融且扩大深化。

2020—2024年:研究精细期。研究后期出现小而多的节点,这代表研究的主题变得越来越细化和专业化。开放数据隐私风险评价仍是开放数据领域

表1 国外政府数据开放阶段及关键词

时期	阶段	关键词
2007—2014年	研究探索时期	知识管理; 电子政务; 大数据
2015—2019年	研究深化期	基层医疗; 治疗; 效果与创新领域; 影响因素; 开放创新; 公民隐私路径
2020—2024年	研究精细期	算法; 大数据; 开放创新

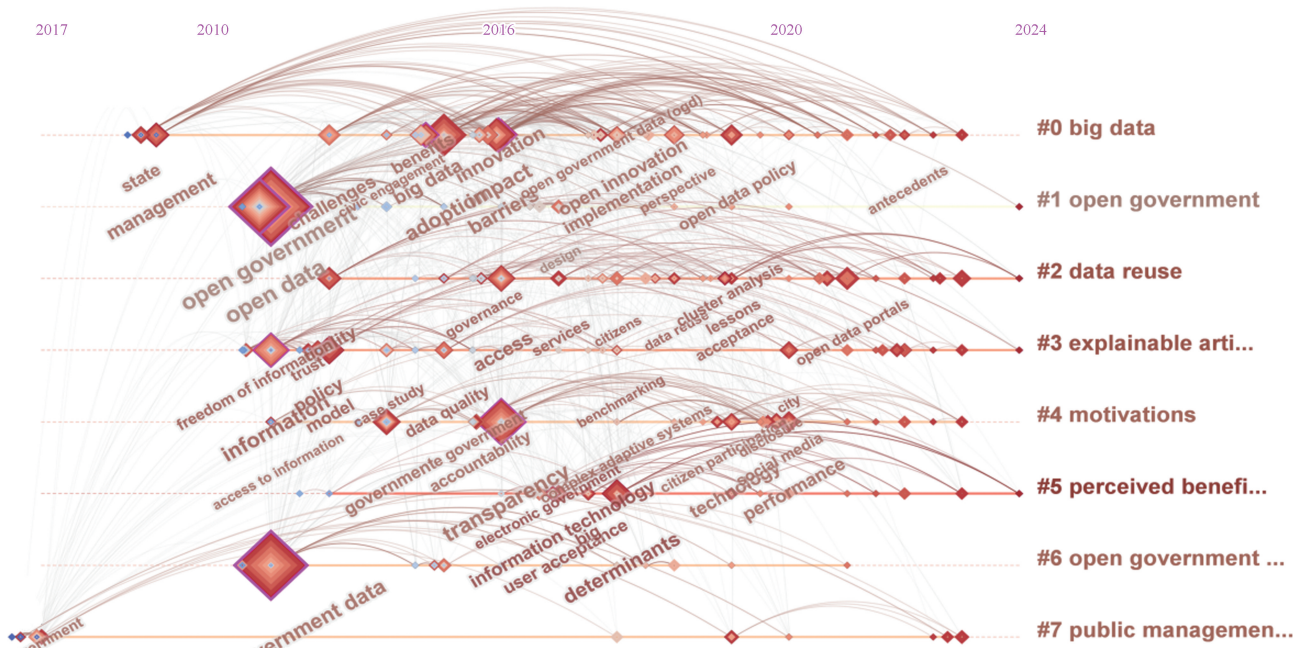


图1 国外政府数据开放研究发展演进时间线

的一个前沿问题;开放数据评价模型研究不断细化深入,一般可分为普适的和专用两种类型;涌现了更多种的开放数据评价工具,涌现了与各种系统相结合的未来导向;由于突发性的公共卫生问题也涌现了部分关于 COVID-19 的议题。政府数据开放由以平台为导向到以公民为导向逐步转型。但最值得注意的是关于大数据和人工智能的应用。“algorithm”(算法)、“big data”(大数据)、“openinnovation”(开放创新)反映了近年国外 OGD 研究侧重结合热门技术和热门话题,推动数字经济持续释放新动能,同时激发社会价值。一方面,OGD 需要利用算法、结构化和非结构化数据以及社交媒体平台以期实现共同发展;另一方面,开放数据技术,如综合的开放数据网站工具(CKAN 等)、元数据和开放数据门户正在强烈刺激开放创新^[7]。部分学者已经开始创新,Cortés-Cediel 等^[8]提出了一个概念框架用于分析研究文献中介绍的在西班牙部署为国家、区域和地方各级公共服务的聊天机器人。因此,未来相当一段时间内,围绕 OGD 促进开放创新的内在机制、开放的数据量与开放创新强度之间关系、基于 OGD 的创新创业活动、可持续发展方法及指标等将会占据 OGD 研究的重要篇幅(图 2)。

2.3.2 国内研究发展演进——从模仿类比到中国化创新

国内政府开放数据研究的关键词共现时间线图谱中前 10 位聚类分别为数据开放、开放数据、

电子政务、大数据、信息安全、开放共享、影响因素、英国、数字政府、协同性(图 2)。通过时间线对国内政府数据开放研究内容进行梳理,可以得出政府数据开放发展的不同阶段时期(表 2)。

2011—2014 年:研究探索时期。2009 年,由美国颁布的开放政府令(open government directive),引发了全球性的开放数据运动。国内研究者也开始关注到政府开放数据这一研究领域并开始学习国外经验,围绕着“开放数据”“信息公开”等关键词获取了部分研究成果。

2015—2017 年:快速发展时期。2015 中国政府开放数据进入快速发展阶段,政府开放数据更大范围地被普及,再加上信息科技的发展,如大数据和因特网,中国的政府数据开放研究也进入了快速发展时期,到 2017 年高水平研究成果呈现出一个发表小高峰。国内研究者在继续吸取各国先进经验的

表 2 国内政府数据开放阶段及关键词

时期	阶段	关键词
2011—2014 年	研究探索时期	开放数据;信息公开
2015—2017 年	快速发展时期	大数据;政府开放数据;数据政策;数据质量
2018—2020 年	平稳发展时期	数据安全;政府治理;开放共享;数据治理;影响因素;地方政府
2021 年至今	深化研究时期	数据平台;数据价值;个人信息;效能模型

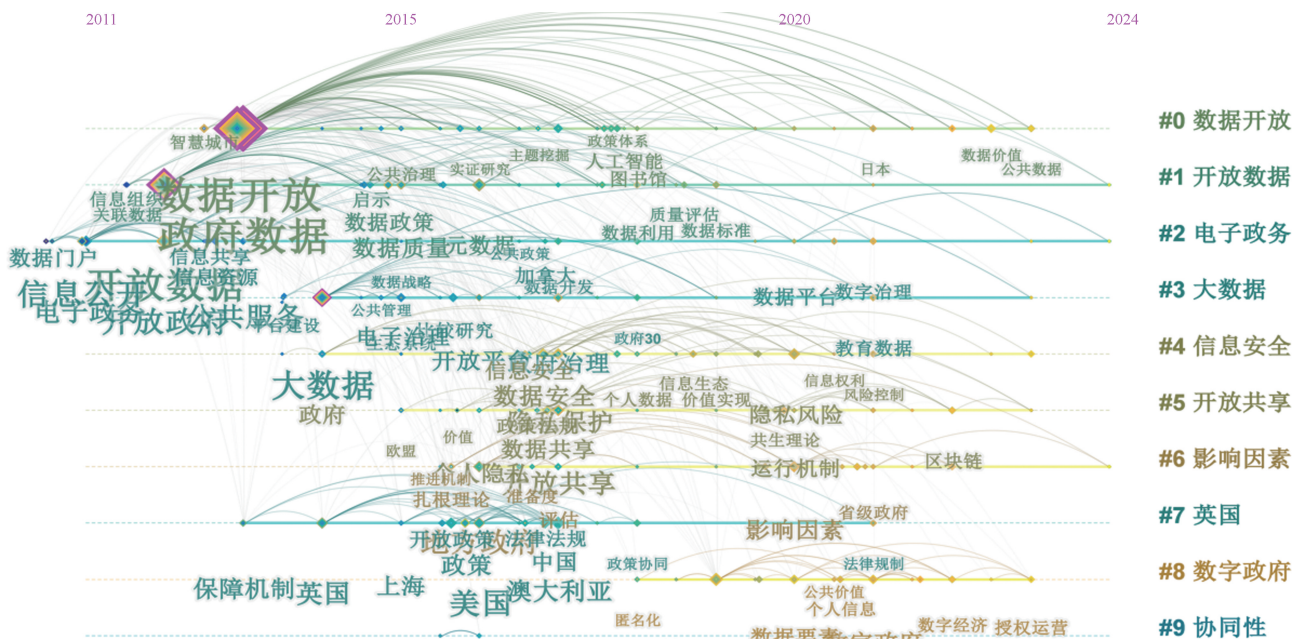


图 2 国内政府数据开放研究发展演进时间线

基础上,围绕“政府开放数据”“数据政策”“数据质量”等逐步构建起了符合中国经验的政府数据开放研究体系。

2018—2020年:平稳发展时期。在大数据战略规划和政务公开热潮的推动下,政务数据的质量问题也越来越突出,解决好数据的质量问题,将会直接提高政府公开数据的水平^[9]。2019年后国内研究者在政府开放数据的领域继续发力,围绕“数据安全”“政府治理”“开放共享”“数据治理”“影响因素”“地方政府”等开展研究,高水平研究成果数量逐年趋于稳定,研究进入了平稳发展时期。该时期政府数据开放的研究更加细化和深入,研究者也开始更多地关注政府数据开放“质”的问题。

2021年至今:深化研究时期。从2021年起国内的政府数据开放方面科研成果减少呈现下降趋势,表明国内政府数据开放研究已步入精细化深化研究阶段。“十四五”推进国家政务信息化规划提出,要推动政务数据按政务公开规则依法依规向社会开放,政府数据开放已经成为中国的重要战略。国内研究者在完善政府数据开放成果的同时,进一步拓展了政府数据开放效果及其成因,并且更加关注政府数据开放与其他主体的关系,“数据平台”“数据价值”“个人信息”“效能模型”等方面的研究成果相继出现,政府数据开放的理论与应用研究在未来一段时期仍然是学术界所需要探索的热点与难点。

2.4 政府数据开放研究主题分析

2.4.1 国外政府数据开放研究关键词共现图谱

本文利用可视化软件,对国外论文样本进行了关键词聚类分析,绘制出关键词聚类的可视化图谱,找出各个聚类的中心。每一个色块代表一个聚类,聚类的面积大,就意味着文献数量多。通过对“较小”的聚类进行剔除,将排名前8的突出的聚类展示出来。各区域中的关键词关联较为密切,聚类效果较好。进而对国外政府数据开放的研究热点进行分析。国外关键词聚类为大数据(big data)、开放政府(open government)、数据重用(data reuse)、可解释的人工智能(explainable artificial intelligence)、动机(motivations)、感知到的好处(perceived benefits)、开放政府数据(open government data)。关键词聚类知识图谱如图3所示。

(1)从研究内容的角度切入,国外学者主要围绕着政府数据开放的两端进行展开暨服务器端和客户终端。从大数据的研究运用、政府数据开放的动机、公民的反馈机制、效果以及模型优化等方面展开研究。国外政府数据开放研究围绕着两端展开但不能明确地判断某一研究方向属于某一端,而是它们相互影响相互促进。

(2)较为侧重服务端的研究首先是关于大数据、关于数据的研究,这些研究均排名靠前,分别位于第1位和第3位,其次是对政府数据开放的动机研究。Abella等^[10]提出开放数据门户应允许开放数据的复用,确保数据传输的效率,并实现基于数

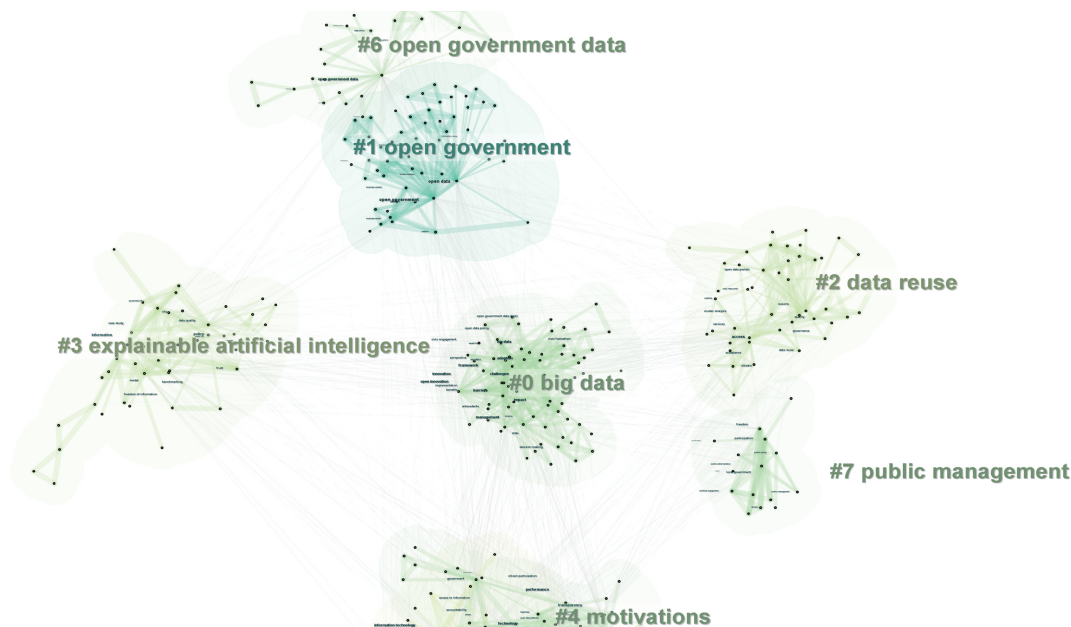


图3 国外政府数据开放研究关键词共现图谱

据复用的专业举措。Nikiforova 和 McBride^[11]探讨了门户可用性对数据重用过程的重要性。Wang 等^[12]探讨了影响政府数据开放再利用意愿的因素。对政府数据开放的动机研究聚焦在政府数据开放政策框架、实施和效果评估。Milić 等^[13]对当代电子政务开放数据平台中使用的 4 种元数据模型进行了概述和分析,结论表明应更多地关注元数据,以揭示数据集之间的关系。Luo 等^[14]采用多阶段双重差分模型验证了政府数据开放对城市创新能力的影响。注重对治理效果的提升和对治理模型的优化升级。

(3)针对客户端的国外研究主要有提升政府数据开放效果和优化模型、公共管理。在提升政府数据开放效果和优化模型方面,近些年的研究中人工智能的应用得到了飞速发展,相应的出现了人工智能的关键词聚类。Karamanou 等^[15]写到动态数据(包括环境、交通和传感器数据)最近被认为是开放政府数据(OGD)的重要组成部分。他们同时使用了 iForest 算法并针对其他特征进行了训练,使用可解释的人工智能对模型进行了解释。Karamanou 等还认为利益相关者可以使用可解释的人工智能(XAI)来了解预测模型的决策。这项工作创建了一个模型,通过在链接的 OGD 上应用机器学习来预测房价。公共管理也被强调在国外的政府数据开放研究中。de Juana-Espinosa 和 Luján-Mora^[16]提出开放政府数据(OGD)计划的目标是提高公共管理政策的透明度、效率和公众参与。Grimmelikhuisen 等^[17]基于两个调查实验发现公民对机构的信任度只与政府的透明度有关,而与政府的潜在透明度无关,即公民的信任感来源于直接获取政府数据而非获取数据的潜在可能性。Yang 和 Wu^[18]研究探讨了政府官员搜索和选择用于开放数据发布的数据集时的信息行为。Krismawati 和 Hidayanto^[19]基于开放政府数据实施的感知有用性概念与现实之间的差距,研究得出信息质量、系统质量、感知有用性、对开放数据的信任度和满意度会影响用户对开放数据门户的参与。

(4)其他关键词表明了目前在海外应用的主要领域和发展的程度。这些节点共同构成了研究的基本架构与知识网络,对政府数据开放具有重要的参考意义。将这些节点信息进行汇总,对信息进行梳理得出了排名前 10 位的高频次关键词,见表 3。

表 3 国外政府数据开放研究高频关键词

序号	关键词	频次	中介中心度	首次出现年份
1	开放数据	135	0.35	2011
2	公开政府数据	115	0.21	2011
3	开放政府	81	0.31	2011
4	采用	55	0.09	2014
5	透明度	53	0.15	2015
6	信息	37	0.15	2011
7	障碍	34	0.06	2015
8	影响	30	0.12	2015
9	创新	28	0.06	2015
10	决定因素	25	0.09	2017

2.4.2 国内政府数据开放研究关键词共现图谱

对国内论文样本进行了关键词聚类分析,绘制出关键词聚类的可视化图谱,找出各个聚类的中心,进而对国内政府数据开放的研究热点进行分析。关键词聚类知识图谱如图 4 所示。从图 4 中能够看出,数据开放、开放数据、共享开放、影响因素等关键词出现频次位居前列,这些节点明显大于其他关键词节点,意味着它们是该领域的核心议题。关键词聚类为数据开放、开放数据、电子政务、大数据、信息安全、开放共享、影响因素、英国、数字政府、协同性。

(1)从研究内容的角度切入,国内学者主要从政府数据开放这一个主体分层次展开不同侧重点研究,具有沿着政治导向发展的特性,更多的是顺着政府数据开放的主流脉络延伸,融合创新占较少部分。更注重研究政府开放数据本身,研究围绕着政府开放数据进行,针对政府开放数据的各个领域和各个阶段进行研究讨论,形成逻辑关系。研究聚焦在技术方面、使用方面和对比研究。

(2)技术方面的研究主要针对大数据、信息安全等,涉及多个层面。大数据研究通常是针对已收



图 4 国内政府数据开放研究关键词共现图谱

集的数据进行进一步分析得出结论。例如,樊博和石语希^[20]选取 2014—2018 年中国 293 个地级市的面板数据进行事件史分析,考察成本-收益和外部压力变量的调节效应,探讨府际关系对政府数据开放政策创新扩散的影响,并将成本-收益理论和外部压力理论作为竞争性理论,揭示创新扩散的具体机理,为大数据发展战略提供可行的政策建议。信息安全方面国内学者注重运用相关政策分析理论模型来对政府数据开放进行安全管理,强调构建人员协调、数据管理、实施顺畅、风险消除。陈美和梁乙凯^[21]通过研究荷兰开放政府数据的隐私风险控制的优势和劣势,在此基础上,提出优化个人信息保护的政策法规,确定个人数据处理原则,完善个人数据处理机构。

(3)使用方面的研究主要围绕着模型构架和提升政府数据开放水平。袁莉等^[22]基于知识发酵模型,分析 OGD 核心能力生成演化过程,围绕 OGD 知识发酵 6 要素提炼供需双方在核心能力生成演化中的影响因素,由此提出营造良好知识环境、完善知识酶催化机制、提升参与主体知识水平的具体建议。赵玉攀^[23]认为开放政府数据不是简单的技术问题,还涉及组织、政策等复杂因素,因此从政策工具、政策内容和政策形式 3 个方面构建了一个综合性的框架来分析政策环境对政府数据开放状态的影响^[23]。

(4)对比研究主要是针对国外政府数据开放的研究,探讨国外不同政府数据开放的治理模式,从而进行借鉴与提升。王贵海和朱学芳^[24]分析国外典型数据开放模式梳理出政府主导模式、公众参与模式和政企合作模式 3 种国外典型的数据开放模式,分别对 3 种模式的特点、典型案例进行深入分析,然后对 3 种模式在国内的具体实践情况进行全面阐述,并分析其对国内数据开放的启示。其他关键词表明了目前在国内应用的主要领域和发展的程度。这些节点共同构成了研究的基本架构与知识网络,对政府数据开放具有重要的参考意义。将这些节点信息进行汇总,对信息进行梳理得出了排名前 10 位的高频次关键词,见表 4。

3 政府数据开放研究主题国际比较

(1)研究应用不同。国外政府数据开放运用更广泛,国外政府开放数据已经成为各领域研究的重要数据来源。政府数据开放的应用领域广泛分布,医疗、地域管理等都涉及且数量众多。在疫情时期,将政府数据开放应用在政府卫生研究中的论文

表 4 国内政府数据开放研究高频关键词

序号	关键词	频次	中介中心度	首次出现年份
1	政府数据	161	0.46	2013
2	数据开放	136	0.44	2013
3	开放数据	111	0.31	2012
4	大数据	46	0.10	2014
5	开放政府	40	0.06	2012
6	数据治理	26	0.05	2019
7	地方政府	23	0.06	2016
8	电子政务	23	0.01	2011
9	信息公开	21	0.07	2011
10	隐私保护	20	0.04	2017

在国外大量涌现,在国内此类应用较为薄弱。此外,国外尤其强调平台的完善,战略性地加强政治参与机制和数字基础设施投资,同时促进强大的地方环境管理范式。与各学科各领域结合,更注重除政府系统之外的各种影响因素之间的交互作用。国内的政府数据开放目前研究主要集中在政府行政和治理学科,较少与其他学科产生联系。

(2)研究侧重点不同。国外对政府数据开放的研究更注重效率提高和技术升级,更注重创新、基础和先进技术的应用。国内更注重研究政府开放数据本身,研究围绕着政府开放数据进行,针对政府开放数据的各个领域和各个阶段进行研究讨论。国内政府开放数据目前多应用在政务服务相关领域,以平台为托手,进行政府数据开放方面的相关研究。

(3)关于公民隐私信息保护的的不同。因为不同国家的政治体制不同,国外会更重视公民参与度问题,同时也会更多地注重政府数据开放在政治选举等一系列参与政治活动中的应用。国内会更注重公民隐私的问题,从公民隐私的信息合理使用、平台隐私等方面展开研究。但对公民隐私的保护仍是国内外的研究前沿且研究主体均有从政府端转向公民端的趋势。

(4)国内在企业相关的应用普遍活跃广泛,引起国外学者的密切关注。国外非常关注中国政府数据开放的应用,中国目前与企业相关的政府数据开放研究受到国外重视。在 WoS 上发表的论文中有不少以中国为主体,研究政府数据开放与企业之间的联系,这表明中国在政府数据开放与企业之间的联系这一方面的研究在国际上得到了一定程度的认可。这些研究为 OGD 在提高生产力、优化政商关系、降低寻租成本和减少信息不对称方面的作用提供了新的视角。同时鉴于 OGD 的包容性,它对弱势企业和城市具有更明显的积极影响。这对

加快开放数据政策的实施和提高经济效率具有重要意义^[25]。

4 国内发展政府数据开放的建议

(1)加强政府数据开放与其他系统的融合研究,促进政府数据开放与其他学科相互融合。国内目前在政府数据开放上的模型研究逐渐丰富,政府数据开放效率研究也逐渐增多,但政府数据开放与其他系统融合方面的论文研究仍相对较少。与其他系统融合发展能使政府数据开放,未来的研究更加深化和广泛。在与其他系统融合时应考虑到政府的权威性质,政府内部信息的隐私性。不可使政府丧失其公信力,出现越过隐私红线的行为。要时刻把握政府机关的定位,明确边界线。

(2)加强政府数据开放的利用效率。虽然开放政府数据(OGD)运动在全球范围内蓬勃发展,但OGD的大规模利用和价值创造尚未真正开放。提供一手数据,将政府开放数据的数据部分作为基础数据提供给公民,在政府相关网站和客户端上向用户开放。继续积极探索政府数据开放作为一种数据技术运用在各种政府App上,这也需要中国数字政府的持续发展。提高利用效率必然要选择开放数据,为防止数据的乱用滥用现象,需要把控好数据的使用途径,不断维护确保数据的准确性,尽可能多地开放公共安全数据,杜绝数据倒卖等行为的出现。

在技术基础层面发力,针对人工智能与政府数据开放融合加强研究。研究发现中国机构在人工智能方面的研究位于世界尖端,研究的热门课题是机器学习、大数据集、深度学习、高级语言等,但美国的发文量远多于中国。如果与人工智能技术相结合,目前的信息系统有很大的改进潜力。随着时间的推移,科学论文的数量也在增加。机器学习等主题的演变意味着人工智能是一个与其他学科融合的广泛知识领域。这些影响对于政策制定者、图书馆利益相关者、研究人员和整个政府决策具有很高的战略价值^[26]。

(4)“十四五”规划强调通过数字社会建设持续提升社会公众的获得感,这也是未来主要的研究目标。目前中国政府数据开放研究的主体逐渐向公民端靠近,但研究大多数集中在公民隐私方面,相比之下针对公民参与度的研究方面稍弱。研究预计在2022—2035年,数据开放层和价值创造层的阶次参数主导地位可能发生转变,数据开放层和数据利用层之间可能产生新的协同演化机制^[27]。因此,

在数据利用层和价值创造层之间建立协同演化机制,有可能增加OGD生态系统的效益。数据开放的主导地位转变,则更需要考虑数据的安全和隐私问题,要加强政策工具的协调配置,采取更加多元化均衡化的组合方案^[28]。

参考文献

- [1] 赵爱玲. 数据中国建设再提速[J]. 中国对外贸易, 2023(4): 23-25.
- [2] 付熙雯, 郑磊. 开放政府数据的价值: 研究进展与展望[J]. 图书情报工作, 2020, 64(9): 122-132.
- [3] GAO T P, MO T L, MO T Q. Overview and hot spot analysis of international open government data research based on knowledge map[J]. Modern Intelligence, 2019, 39(3): 86-100.
- [4] BARCELLOS R, BERNARDINI F, VITERBO J, et al. Hippolyta: a framework to enhance open data interpretability and empower citizens[C]//Proceedings of the 24th Annual International Conference on Digital Government Research. Gdańsk: The Association for Computing Machinery, 2023: 191-198.
- [5] 金香兰, 郑非凡. 我国民族地区乡村治理的研究热点、前沿趋势与未来展望: 基于CSSCI文献计量的CiteSpace可视化分析[J]. 延边大学学报(社会科学版), 2023, 56(2): 129-140.
- [6] Gesetz zur Förderung der elektronischen Verwaltung (E-GovernmentGesetz-EGovG) [EB/OL]. (2013-07-25) [2024-03-18]. <https://www.gesetze-im-internet.de/egovg/BJNR274910013.html>.
- [7] CHATFIELD A T, REDDICK C G. The role of policy entrepreneurs in open government data policy innovation diffusion: an analysis of Australian federal and state governments[J]. Government Information Quarterly, 2018, 35(1): 123-134.
- [8] CORTÉS-CEDIEL M E, SEGURA-TINOCO A, CANTADOR I, et al. Trends and challenges of e-government chatbots: advances in exploring open government data and citizen participation content[J]. Government Information Quarterly, 2023, 40(4): 101877.
- [9] 夏姚璜, 邢文明. 开放政府数据评估框架下的数据质量调查与启示: 基于《中国地方政府数据开放报告(2018)》[J]. 情报理论与实践, 2019, 42(8): 44-49, 66.
- [10] ABELLA A, ORTIZ-DE-URBINA-CRIADO M, DEPABLOS-HEREDERO C. Criteria for the identification of ineffective open data portals: pretender open data portals[J]. Profesional de la información/Information Professional, 2022, 31(1): e310111.
- [11] NIKIFOROVA A, MCBRIDE K. Open government data portal usability: a user-centred usability analysis of 41 open government data portals[J]. Telematics and Informatics, 2021, 58: 101539.

- [12] WANG F, ZHANG Z, MA X, et al. Paths to open government data reuse: a three-dimensional framework of information need, data and government preparation[J]. *Information & Management*, 2023, 60(8): 103879.
- [13] MILIĆ P, VELJKOVIĆ N, STOIMENOV L. Comparative analysis of metadata models on e-government open data platforms [J]. *IEEE Transactions on Emerging Topics in Computing*, 2018, 9(1): 119-130.
- [14] LUO Y, TANG Z, FAN P. Could government data openness enhance urban innovation capability? an evaluation based on multistage DID method[J]. *Sustainability*, 2021, 13(23): 13495.
- [15] KARAMANOU A, BRIMOS P, KALAMPOKIS E, et al. Exploring the quality of dynamic open government data using statistical and machine learning methods[J]. *Sensors*, 2022, 22(24): 9684.
- [16] DE JUANA-ESPINOSA S, LUJÁN-MORA S. Open government data portals in the European Union: Considerations, development, and expectations[J]. *Technological Forecasting and Social Change*, 2019, 149: 119769.
- [17] GRIMMELIKHUIJSEN S G, PIOTROWSKI S J, VAN RYZIN G G. Latent transparency and trust in government: unexpected findings from two survey experiments [J]. *Government Information Quarterly*, 2020, 37(4): 97-111.
- [18] YANG T M, WU Y J. Looking for datasets to open: an exploration of government officials' information behaviors in open data policy implementation[J]. *Government Information Quarterly*, 2021, 38(2): 101574.
- [19] KRISMAWATI D, HIDAYANTO A N. The user engagement of open data portal [C]//2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS). Sanya, IEEE, 2021: 1-6.
- [20] 樊博, 石语希. 中国政府数据开放的政策创新扩散研究: 成本收益和外部压力的竞争性解释[J]. *现代情报*, 2023, 43(10): 74-84.
- [21] 陈美, 梁乙凯. 荷兰开放政府数据的隐私风险控制研究[J]. *情报杂志*, 2022, 41(2): 169-175, 168.
- [22] 袁莉, 赵婧, 顾嘉琪. 基于知识发酵的政府数据开放核心能力生成演化研究[J]. *图书情报工作*, 2023, 67(10): 107-115.
- [23] 赵玉攀. 基于三维框架的中国省级政府数据开放政策分析[J]. *图书馆学研究*, 2020(13): 40-48, 83.
- [24] 王贵海, 朱学芳. 国外典型数据开放模式分析及其在我国的实践与启示[J]. *情报理论与实践*, 2023, 46(12): 193-200.
- [25] XU C, CHEN Y, DAI J. Open government data and resource allocation efficiency: evidence from China [J]. *Applied Economics*, 2024, 1(1): 1-18.
- [26] BORGHAIN D J, BHARDWAJ R K, VERMA M K. Mapping the literature on the application of artificial intelligence in libraries (AAIL): a scientometric analysis [J]. *Library Hi Tech*, 2024, 42(1): 149-179.
- [27] FANG J, ZHAO L. Synergistic evolution mechanisms for improving open government data ecosystems using the Haken model [J]. *Library & Information Science Research*, 2024, 46(1): 101284.
- [28] 姚怡帆, 王珊珊, 许正中. DIIS理论下地方政府数据治理政策的现状分析与提升策略[J]. *科技和产业*, 2023, 23(15): 211-220.

International Comparative Analysis of Government Data Openness and Future Development Ideas in China

JING Lingling, ZHANG Yunyao

(School of Economics and Management, Harbin Normal University, Harbin 150025, China)

Abstract: The open government movement has promoted the open sharing of government data around the world, and how to make better use of government open data to promote political democracy and social progress has become an urgent issue for discussion. Bibliometric analysis was conducted to explore the research context, cooperation, research hotspots and development trends of domestic and foreign scholars on government data openness since 2000. CiteSpace knowledge graph was used to conduct a comparative study on government data openness at home and abroad. It is found that foreign government data openness has better integration of different disciplines and a higher degree of data utilization, and has begun to try to apply artificial intelligence to the government data openness platform. Given the domestic and foreign comparison results, the following suggestions are put forward for the opening of domestic government data: strengthen the integration with artificial intelligence, strengthen the efficiency of data utilization and integration with other discipline systems, and finally sink from the government to the citizen side.

Keywords: open government data; bibliometrics; digital governance