

基于LDA和ISM法从“结构-过程”视角 解构政府数据开放能力*

汤志伟^{1,2} 陈瑶¹ 韩啸¹ 王莉¹

(1. 电子科技大学公共管理学院, 成都 611731; 2. 电子科技大学深圳高等研究院, 深圳 518000)

摘要: 概念是研究的起点。现有研究对政府数据开放能力的理解借用私营部门的理论知识, 忽视了公共部门特性, 因此有必要对政府数据开放能力进行重新界定。本文运用LDA主题模型法对国内相关政策进行文本数据挖掘, 识别出政府数据开放能力的构成要素, 再通过ISM法厘清各要素在“宏观-中观-微观”的层次关系与传导路径, 最后根据研究发现从政府数据开放能力的角度对政府数据开放实践提出对策建议。

关键词: 政府数据; 开放能力; 政策; LDA; ISM; 编码

中图分类号: D63 **DOI:** 10.3772/j.issn.1673-2286.2023.02.002

引文格式: 汤志伟, 陈瑶, 韩啸, 等. 基于LDA和ISM法从“结构-过程”视角解构政府数据开放能力[J]. 数字图书馆论坛, 2023 (2) : 10-17.

政府数据开放是落实数字中国战略的基础性和先导性工程, 各级政府掌握着我国80%的数据资源^[1], 如何推动政府数据开放能力的提升, 打通数据资源开放共享所面临的梗阻, 改善数据质量, 促进数据价值开发, 不仅关系到政府数据开放实践的成败, 更影响到数字中国战略目标的实现以及国家治理现代化的进程。当前各地在政府数据开放实践中面临“增量”与“提质”并存的“双重压力”^[2], 暴露出数据质量堪忧、数据权属模糊与隐私威胁等诸多问题^[3-4]。无论是学界还是实务界均意识到解决上述问题的关键是抓住政府数据开放能力这一核心。遗憾的是, 对于政府数据开放能力的理解存在模糊化、碎片化问题。因此, 本文选择常用的主题模型法——潜在狄利克雷分配 (Latent Dirichlet Allocation, LDA) 以及擅长分析层次隶属关系的解释结构模型法 (Interpretative Structural Modeling Method, ISM) 作为研究方法, 并以“结构-过程”视角展开研究, 弥补既有研究对政府数据开放能力过程层

面的忽视, 归纳、提取政府数据开放能力的概念内涵、构成要素, 并厘清各要素之间的层次关系与传导路径, 打开政府数据开放能力的概念“黑箱”, 回应现有研究面临的模糊化、碎片化问题。

1 文献综述

近年来, 学界围绕“政府数据开放能力”议题从概念内涵、构成要素与作用效果三个方面展开了丰富讨论。①概念内涵研究。郑磊^[5]强调政府数据开放与政府信息公开之间存在本质区别; 黄莺^[6]认为政府数据开放能力内涵包括数据提供过程、数据提供结果与目标支持程度; Mitrovic^[7]提出政府数据供给和公众数据利用所需的各项技能就是政府数据开放能力的概念内涵; 袁莉等^[8]立足知识观指出政府数据开放能力的产生是一个知识学习过程, 是由核心能力和动态能力相互作用形成的组合能力。②构成要素研究。曹谦等^[9]从提供

收稿日期: 2023-01-07

* 本研究得到国家社会科学基金重点项目“人工智能商业化应用的社会影响与治理体系研究 (编号: 22AZD135) 和国家社会科学基金青年项目“人工智能应用背景下政府数据开放能力提升路径研究” (编号: 20CZZ034) 资助。

者、管理者和开发利用者角度展开分析,提出政府数据开放能力的核心构成要素包括数据释放能力、数据管控能力和数据应用能力;樊博等^[10]从组织层面提出政府数据开放能力由数据共享、数据公开、数据监管、数据再利用等要素组成。③作用效果研究。彭勃等^[11]研究发现,政府数据开放能力可以显著促进城市政府数据开放水平;还有研究指出政府数据开放能力有助于增加政府透明度、杜绝腐败^[12],还能提高决策科学性、促进公众参与、提升治理效率与改善服务质量^[13]。

已有研究对政府数据开放能力的内涵特征、构成要素、作用效果进行了充分讨论,但仍存在一些问题。第一,在研究视角上,既有研究借鉴企业领域的核心能力理论、动态能力理论和生态系统理论对政府数据开放能力进行阐释,具有一定的参考价值,但由于企业与政府在价值追求上存在显著差异,因此无法准确把握政府数据开放的内核,会引发适用性问题。第二,在要素分析上,既有研究或是停留于静态分析,忽视了政府数据开放的动态特征;或是仅对某一维度的能力要素进行阐释,缺乏不同维度要素的关联整合,影响了概念解释力。第三,在方法上,既有研究以理论推导和质性访谈为主,易受主观判断影响,难以发现更具普遍性与现实性的概念与要素。鉴于此,本文借鉴吉登斯的结构化理论,将“结构-过程”视角带入政府数据开放能力的理解中,运用LDA主题模型法对44份政策文件进行文本数据挖掘,发现政府数据开放能力的概念内涵与构成要素,再使用ISM法进一步厘清政府数据开放能力构成要素的层次关系与传导路径,以期对政府数据

开放能力进行全面透彻、客观准确的分析。

2 研究设计

2.1 政策筛选

本文在北大法宝和中国政府网以“政府数据开放”“开放政府数据”“政府开放数据”为关键词对中央和地方政策文件进行全文精准检索,得到293份政策文件。按照以下标准进行筛选:①剔除主题与政府数据开放严重不相符的政策文件;②剔除按检索条件命中但内容未阐述政府数据开放具体要求的政策文件;③剔除地方政府转发的中央政策文件;④剔除已失效的政策文件或未公布全文的政策文件;⑤部分地方性文件在实践上不具有典型性,不考虑纳入。学术界将政府数据、政务数据纳入公共数据之中^[14],实务界则将政府数据和公共服务数据归入公共数据统一进行分类分级管理^[15]。因此,本文将涉及公共数据的政策文件亦纳入分析。在经过筛选剔除后共得到44份政策文件,由于篇幅限制,表1仅展示部分政策文件。

2.2 研究方法

本文使用LDA和ISM两种研究方法。LDA作为目前应用最广泛的主题模型,由Blei等^[16]在2003年提出。LDA擅长对高维异构的文本数据进行建模,可从稀疏、高维的数据中清晰了解文本的主题结构及其分

表1 关于政府数据开放的政策文件(部分)

编号	文件名	发文部门	发文号	发文时间
1	关于加强数字政府建设的指导意见	国务院	国发(2022)14号	2022年6月
2	要素市场化配置综合改革试点总体方案	国务院办公厅	国办发(2021)51号	2021年12月
3	关于通过公共数据开放促进人工智能产业发展的工作方案	北京市大数据工作推进领导小组办公室	京大数据办发(2019)2号	2019年10月
4	“十四五”大数据产业发展规划	工业和信息化部	工信部规(2021)179号	2021年11月
5	上海市数据条例	上海市人民代表大会常务委员会	(十五届)第九十四号	2021年9月
6	中华人民共和国数据安全法	第十三届全国人民代表大会常务委员会	中华人民共和国主席令第84号	2021年6月
7	贵阳市政府数据共享开放条例	贵州省第十三届人民代表大会常务委员会	筑人大字(2021)5号	2021年6月
8	关于印发政务信息资源共享管理暂行办法的通知	国务院	国发(2016)51号	2016年9月
9	关于通过公共数据开放促进人工智能产业发展的工作方案	北京市大数据工作推进领导小组办公室	京大数据办发(2019)2号	2019年10月
10	上海市公共数据开放暂行办法	上海市人民政府	上海市人民政府令21号	2019年8月

布情况。因此，本文运用LDA主题模型对国内政策文本进行挖掘，以识别出政府数据开放能力的主要构成要素。

ISM法是由Warfield^[17]提出，该方法的基本原理是通过确定影响系统的各因素及相互关系，利用建立关联矩阵与布尔逻辑运算，分析各因素间的关联性和层级关系并揭示系统结构。在不损失系统整体功能前提下，以有向拓扑图方式呈现出主（次）要因素的关联结构。为此，本文使用ISM法分析政府数据开放能力构成要素之间的隶属关系，通过“专家意见-统计反馈-意见调整”多轮循环，最终构建得出多层递阶模型。

3 研究过程

3.1 基于LDA模型的政府数据开放能力构成要素识别

LDA模型分析过程为“语料预处理→确定主题数→建立主题-词汇关系”，本研究采用LDA模型对文本主题词进行量化挖掘处理，使用Python软件实现模型操作。

3.1.1 语料预处理

LDA主题模型要求以文本“词语”为单位挖掘计算出潜在主题，如何分词直接关系到主题挖掘结果。分词采用jieba分词的“lcut”精确分词模式，同时在语料处理时删除标点符号，结合百度停用词表等中文停用词表对44份政策文本进行分词及词频统计（见表2）。

“数据”与“开放”分别出现1 458次和1 421次，由于这两个词汇本身与“数据开放”同质，不纳入词频统计前列。结果显示，“公共数据”出现次数最多，其次是“主体”“管理”“利用”和“服务”，表明政府数据开放的主要倾向在数据管理和数据使用方面。

表2 jieba分词词频统计结果（部分）

词语	词频	词性	词语	词频	词性
公共数据	984	名词	服务	188	动词
主体	346	名词	目录	180	名词
资源	341	名词	社会	176	名词
管理	303	动词	开放平台	170	名词
利用	274	动词	提供	157	动词

3.1.2 确定主题数

主题数K是确定计算结果是否合理与科学的重要参数。通过困惑度模型计算公式绘制困惑度曲线。考虑到政策文本量，假设 $K \in (0, 25]$ 。困惑度越低，证明主题数越合理。数据表明，当主题数为10时，困惑度出现第一个下降拐点，表明10可作为可信的主题数量。

3.1.3 建立“主题-词汇”关系

设定主题数量为10，在Python中运行，展示每个主题前10个词汇，生成“主题-词汇”矩阵，如表3所示。这些词汇代表了政府数据开放的具体政策议题与重要发力点。

表3 “主题-词汇”矩阵

主题	词汇
1	办法, 协调, 统筹, 法律法规, 开放平台, 原则, 建设, 标准互联网, 业务
2	本级, 第三方, 平台, 汇聚, 资源, 建设, 开发利用, 运维, 公共数据, 功能
3	单位, 法律, 治理, 部门, 审批, 资源, 条件, 技术, 更新, 信息安全
4	商业秘密, 个人隐私, 发展, 国家, 产业, 社会, 实施, 引导, 社会效益, 增值
5	有效, 主管部门, 主体, 管理, 获取, 目录, 利用, 法规, 工作, 责任
6	职责, 公共数据, 数据安全, 措施, 审查, 主体, 清单, 监督, 需求, 评估
7	法人, 组织, 授权, 价值, 职能, 加工, 产品, 单位, 公民, 企事业
8	提升, 能力, 公共服务, 举报, 创新, 违法, 合作, 发现, 分级分类, 鼓励
9	评价, 会同, 下载, 开发, 经济, 第三方, 评估, 专业, 创业, 工作
10	义务, 准确性, 建立健全, 质量, 发布, 社会化, 生态, 监测, 动态, 数据管理

运用pyLDAvis对10个主题间的距离进行可视化分析（见图1）。主题1和主题2有较多重合，其他主题间具有较好的分散度，故将主题1和主题2合并为一个主题。

将合并后的“主题-词汇”进行归纳，共识别出9个构成要素（见表4）。“平台”是数据开放的重要载体，贯穿政府数据开放的诸多流程，主题1和主题2强



图1 主题间距离可视化演示

调了平台建设的方式与运营,即平台运营能力的体现;主题3中的高频词“法律”“治理”“审批”等体现出政府数据开放的制度设计能力;主题4的“国家”“产业”“社会”等高频词反映出政府应用技术推动数据开放的广度,“发展”“产业”“社会效益”等与数据开放应用密切关联,归纳为技术同化能力;主题5的“主管部门”“工作”“责任”等高频词涉及对政府数据开放工作的要求与考核,是对组织能否胜任政府数据开放的综合评价;主题6中的“数据安全”“审查”“监督”等词强调政府对数据安全的重视,处理好发展与安全的关系才能推动政府数据开放高质量发展,将其归纳为数据安全能力;在主题7中,关键词“法人”“单位”“公民”和“企事业”实质上是对政府数据开放的对象说明,强调数据开放过程中对公众需求的关注,即需求回应能力;主题8中的“创新”“合作”“分级分类”“鼓励”等词体现以创新理念推动政府数据开放发展的组织氛围;主题9中的“下载”“开发”“经济”“创

表4 政府数据开放能力的构成要素

主题	构成要素	说明
1和2	平台运营能力	作为政府数据开放最重要的载体,平台建设水平直接影响用户对数据的发现与利用 ^[18]
3	制度设计能力	制度设计包括政府数据开放的领导与协调机制、行政推进机制、绩效评估机制、参与互动机制、技术支撑与规范等 ^[19]
4	技术同化能力	技术同化指组织应用技术的深度和广度 ^[20] ,是数据生成的重要前提
5	组织胜任力	组织胜任力是一个综合能力,贯穿整个开放数据生命周期,涉及静态结构因素和动态过程因素 ^[21]
6	数据安全能力	数据安全是政府数据开放的底线,决定数据价值创造能走多远,涉及风险识别、风险评估、风险控制等方面 ^[22]
7	需求回应能力	政府数据开放的最终目的是为公众服务,由于其内嵌于服务型政府建设之中,回应需求也是政府数据开放建设的重要内容之一 ^[23]
8	组织创新氛围	组织对创新的鼓励和容错是消除避责现象,推动政府数据开放发展的关键动力 ^[24]
9	数据利用能力	数据利用是实现政府数据开放价值的关键途径,包括利用者多样性、被利用数据量、被利用数据集的覆盖率 ^[25]
10	数据提供能力	数据提供能力是开放数据的核心,包括数据标准、数量、质量等 ^[26]

业”等词汇反映数据使用者对数据的开发利用,推动数据创造更多价值,即数据利用能力;主题10中的“发布”“准确性”“质量”“数据管理”等高频词聚焦数据质量,是政府数据提供能力的体现。

3.2 基于结构化理论的政府数据开放能力框架构建

结构化理论面对数字化情境具有高度概括性和解释力,因而被信息资源管理领域的学者广泛使用。考虑

到既有研究对政府数据开放能力的阐释忽视了过程层面的作用,因而引入结构化理论进行分析。结构化理论强调结构是在生产过程中反复涉及的规则和资源,这些规则和资源不断规范着主体的行为,给主体的行为提供强制或支持。主体在行动过程中会不断地创生新规则和新资源,使得整个社会的规则和资源能够不断地循环发展。笔者借鉴结构化理论将制度设计、平台运营、技术同化、组织创新氛围等能力划分为结构性要素,同时根据数据生命周期把数据提供、数据利用、数据安全、需求回应等能力划分为过程性要素。

将LDA模型识别出的能力要素按照“结构-过程”视角进行整合,构建政府数据开放能力框架(见图2)。在框架中,结构的作用是设定规则、提供资源,实现基本秩序与预期目标;过程是不同行动主体围绕权力获取、资源配置、权力行使等要素进行的互动。当结构性

要素与过程性要素互动后,形成的组织能力与组织目标相互匹配时,则达到均衡状态,反馈维持不变的系统信号;反之,当组织能力无法支撑组织目标实现时,则呈现冲突状态,产生冲击改变的反馈信号。

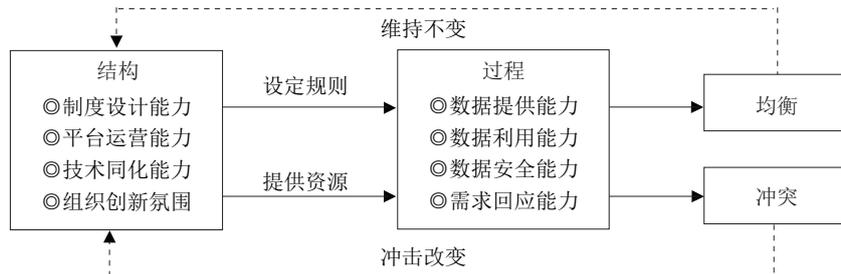


图2 “结构-过程”视角下的政府数据开放能力框架

3.3 政府数据开放能力构成要素的关系分析

为进一步探寻政府数据开放能力构成要素间的层次结构和传导路径,使用ISM法构建政府数据开放能力要素层次结构模型。

3.3.1 组建专家组

对于选多少专家填写调查表格而言,专家数量不是多多益善,实际上过多的专家反而会增加数据资料的复杂性,进而降低调查准确度与一致性。参考既有研究^[27],本研究邀请政府数据开放、政府数据治理等领域学者与政府管理者共计15人,其中高校教授7人,政府管理者8人,他们对政府数据开放的工作内容非常熟悉,可以较好地完成调查表格。

3.3.2 构成要素层次分析

ISM法在专家打分的基础上,运用模糊矩阵确定要素之间的影响关系,划分影响层级,形成清晰、多递阶的结构模型。具体研究过程如下。

(1)邀请专家对9个要素间的直接关系进行打分,若因素两两之间有直接影响关系记为1分,若无记为0分。构建得到9×9的邻接矩阵A。邻接矩阵A的要素 a_{ij} 定义为公式1所示。

$$A = a_{ij} = \begin{cases} 1, & S_i R S_j \\ 0, & S_i R S_j \end{cases} \quad (\text{公式1})$$

(2)生成可达矩阵M。邻接矩阵A只能体现各要素之间的直接关系,需对邻接矩阵(A+I)进行幂运算得到可达矩阵M,M能够体现各要素之间的间接关系。公式2满足时得到M,其中I为单位矩阵。

$$M = (A+I)^{(k+1)} = (A+I)^k \neq (A+I)^{(k-1)} \quad (\text{公式2})$$

(3)层次划分。以R=C(其中R为可达集数量,C为要素数量)为划分原则得到3个层级(见表5)。

表5 政府数据开放能力构成要素层次分解表

层次	构成要素
第一层(宏观层)	组织胜任力
第二层(中观层)	制度设计能力、组织创新氛围、平台运营能力、技术同化能力
第三层(微观层)	数据提供能力、数据利用能力、数据安全能力、需求回应能力

综上,政府数据开放能力的概念内涵清晰浮现出来,在宏观上体现为政府部门持续有效满足社会公众对数据需求的组织胜任力。中观层为4项结构性要素:制度设计能力、组织创新氛围、平台运营能力和技术同化能力。微观层包括4项过程性要素:数据提供能力、数据利用能力、数据安全能力、需求回应能力。

政府数据开放能力是由一系列复杂要素相互作用的结果,根据表5构建政府数据开放能力构成要素层次结构模型(见图3)。层次结构模型分为宏观、中观与微观三个层次。宏观层要素意味着组织需要持续有效地

不断满足各类需求,中观层4个要素发挥着承上启下的作用,微观层要素是最容易对其他层级要素产生影响的基础要素。微观层要素推动中观层要素的提高,从而影响宏观层要素,形成层层递进的传导路径。

从图3可见,模型中存在4条传导路径,具体而言,微观层要素“数据提供能力”“数据安全能力”作用于中观层要素“制度设计能力”进而对宏观层要素“组织胜任力”产生影响,该条传导路径为制度驱动路径;微观层要素“数据利用能力”“需求回应能力”通过中观层要素“组织创新氛围”对宏观层产生影响,该条传导路径为组织创新路径;微观层要素“数据提供能力”“数据安全能力”通过中观层要素“平台运营能力”传导影响宏观层,该条传导路径为平台生态路径;微观层要素“数据提供能力”“需求回应能力”影响中观层要素“技术同化能力”进而作用于宏观层要素,该条传导路径为技术赋能路径。由此可见,抓住微观层要素,增加对它们的建设力度,能够产生事半功倍的效果,推动政府数据开放能力提升并促进政府数据开放高质量发展。

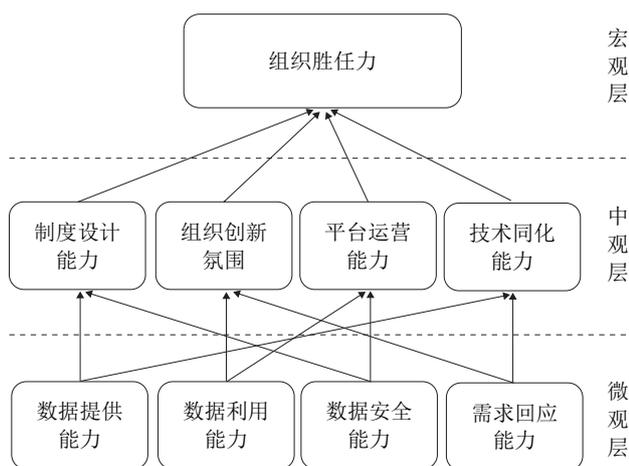


图3 政府数据开放能力构成要素层次结构模型

4 结论与建议

4.1 研究发现

本文通过混合研究设计对政府数据开放能力展开了递进研究,分别探究了“是什么”“由什么构成”以及“如何提升”等理论问题。具体而言,研究发现主要有三点。第一,清晰归纳出政府数据开放能力的概念内

涵。借助“结构-过程”视角对政府数据开放能力展开概念定义,厘清其在宏观层抽象为政府部门持续有效满足社会公众对数据需求的组织胜任力,在中观层为政府部门执行政府数据开放的结构性能力,在微观层是指政府部门从事政府数据开放的过程性能力。第二,探明政府数据开放能力的核心构成要素。通过LDA主题模型对44份政策文件进行数据挖掘后,发现政府数据开放能力的核心构成要素,即制度设计能力、组织创新氛围、平台运营能力、技术同化能力、数据提供能力、数据利用能力、数据安全能力、需求回应能力。第三,厘清要素间的层次关系与传导路径。政府数据开放能力是由一系列复杂要素相互作用的结果,透过ISM法进行层次结构分析,发现9个要素共隶属宏观、中观、微观三个层次,存在制度驱动、组织创新、平台生态和技术赋能4条传导路径。

4.2 对策建议

基于本文研究发现,对目前政府数据开放实践提出以下建议。

(1) 注重资源配置和数据标准化建设。政府数据开放的资源配置包含基础设施建设、公共财政支持。基础设施是指信息化基础设施建设,其被认为是数据实现技术连接创造价值的重要基础。未来需要注重区域基础设施建设的均衡发展,避免数字鸿沟的扩大。在资源配置层面,强调统筹资源,避免基础设施的重复投资和浪费;此外,要进一步加强数据标准建设,形成数据分类、分层的精细管理,整合汇编数据资源目录,保证数据颗粒度与整合度。

(2) 以公众需求为导向的平台功能优化。平台是政府数据开放的依托。政府数据开放平台功能是政府数据开放实现价值增益的过程机理,包含数据整合、分析、共享等。未来在平台功能优化上需要强化两点:一是通过平台更加频繁地共享数据,实现纵向上中央与地方共享,横向上部门与部门间共享;二是完善平台互动与反馈功能。这与需求回应能力息息相关,涉及对用户反馈的诉求是否能予以合理处理。

(3) 明确组织制度及打通“条块”壁垒。组织因素是影响政府数据开放的重要因素。政府作为数据开放的主体,是推动政府数据开放各部分实现顺畅运作的核心。政府部门之间具有天然的独立性,这是职能划分带来的必然结果。未来政府需要抓住两个关键点以打通

“条块”壁垒：一是明确主体责任，厘清部门责任和义务，制定明确的工作计划；二是创新数据治理机制，统筹推进政府部门数据共享。

(4) 鼓励数据开发与数据利用。数据利用是实现基础供给能力转化为最终绩效产出结果的关键环节，也是政府数据开放实现价值共创的最主要体现。结合“十四五”规划中的具体要求，下一步需加强以下举措：一是加大在数据利用方面的人才培训、技术发展、资金倾向等措施；二是重点创新数据开发模式，鼓励众筹、众包、PPP等模式；三是搭建多元参与的行动框架，鼓励企业、行业协会、科研机构等第三方的合作参与。

(5) 加强数据安全的法治化建设。保证数据安全是数据开放的前提，数据安全法律制度构建必须完善，并发挥最主要的作用。在数据安全的法治化建设方面，我国政府态度明确，一是注重对数据的脱密处理，将国家安全置于首位；二是强调个人隐私的保护，实现隐私的前置保护。对数据安全的保障不仅是政府数据提升经济效益的保障，更是提升治理效益的全方位衡量。未来各级政府需要结合自身工作情况，完善制度体系，落实安全责任。

当然，本文尚存一些不足，后续研究可从以下方面展开讨论：第一，确定要素权重，开发政府数据开放能力的评价指标体系；第二，探讨政府数据开放能力与政府数据开放绩效的因果关系；第三，探索引起政府数据开放能力差异的前因变量。

参考文献

- [1] 李克强：深化简政放权放管结合优化服务 推进行政体制改革 转职能提效能 [EB/OL]. [2022-12-28]. http://www.gov.cn/premier/2016-05/22/content_5075741.htm.
- [2] 张翔. 复式转型：地方政府大数据治理改革的逻辑分析 [J]. 中国行政管理, 2018 (12): 37-41.
- [3] 鲍静, 张勇进, 董占广. 我国政府数据开放管理若干基本问题研究 [J]. 行政论坛, 2017 (1): 25-32.
- [4] 张会平, 顾勤. 政府数据流动：方式、实践困境与协同治理 [J]. 治理研究, 2022 (3): 59-69.
- [5] 郑磊. 开放政府数据研究：概念辨析、关键因素及其互动关系 [J]. 中国行政管理, 2015 (11): 13-19.
- [6] 黄莺. 政府数据开放绩效研究：内涵、特征及生成机理 [J]. 四川大学学报 (哲社版), 2021 (5): 148-157.
- [7] MITROVIC Z. Building Open Data Capacity through ESkills Acquisition [EB/OL]. [2022-12-28]. <https://www.researchgate.net/publication/276288202>.
- [8] 袁莉, 姚乐野. 政府数据开放的整体能力：概念、框架及演化机制 [J]. 图书情报工作, 2021 (10): 53-60.
- [9] 曹谦, 刘军. 地方政府数据开放的核心能力塑造 [N]. 学习时报, 2018-12-10 (6).
- [10] 樊博, 陈璐. 政府部门的大数据能力研究：基于组织层面的视角 [J]. 公共行政评论, 2017 (1): 91-114.
- [11] 彭勃, 吴金鹏, 韩啸. 开放政府数据的生态系统能力：模型建构与发展路径 [J]. 上海行政学院学报, 2022 (3): 4-15.
- [12] MCBRIDE K, AAVIK G, TOOTS M, et al. How does open government data driven co-creation occur? Six factors and a “perfect storm” insights from Chicago’s food inspection forecasting model [J]. Government Information Quarterly, 2019, 36 (1): 88-97.
- [13] SAFAROV I, MEIJER A, GRIMMELIKHUIJSEN S, et al. Utilization of open government data: a systematic literature review of types, conditions, effects and users [J]. Information Polity, 2017, 22 (1): 1-24.
- [14] 张会平. 面向公共价值创造的城市公共数据治理创新 [J]. 行政论坛, 2020, 27 (1): 130-136.
- [15] 政务和公共服务数据归入公共数据分类分级管理《重庆市数据条例》更加有利于促进数据价值释放 [EB/OL]. [2022-07-03]. http://www.cq.gov.cn/ywdt/jrcq/202207/t20220703_10881217.html.
- [16] BLEI D M, NG A Y, JORDAN M. Latent Dirichlet Allocation [J]. The Journal of Machine Learning Research, 2003 (3): 993-1022.
- [17] WARFIELD J N. Binary Matrices in System Modeling [J]. IEEE Trans on Systems Man&Cybernetics, 1973, 3 (5): 441-449.
- [18] 郑磊, 高丰. 中国开放政府数据平台研究：框架、现状与建议 [J]. 电子政务, 2015 (7): 8-16.
- [19] 中国行政管理学会课题组. 我国政府数据开放顶层设计研究 [J]. 中国行政管理, 2016 (11): 6-12.
- [20] 韩啸, 汤志伟. 数字政府创造公共价值的驱动因素与作用机制研究 [J]. 电子政务, 2022 (2): 51-64.
- [21] 樊博. 推进开放政府数据：提升政府部门大数据能力 [J]. 学海, 2018 (2): 5-10.
- [22] 袁静, 刘晓媛, 臧国全. 用户参与共治：政府开放数据隐私风险治理的新思路 [J]. 图书情报知识, 2022 (12): 1-11.
- [23] 上官莉娜, 夏本倩. 地方政府数据开放水平的驱动因素与提升

- 路径研究[J]. 数字图书馆论坛, 2022(9): 13-20.
- [24] 雷玉琼, 刘展余. 干部容错机制能够促进政府创新吗?——基于地方政府数据开放的经验研究[J]. 湖南农业大学学报(社会科学版), 2021(3): 51-59.
- [25] 温祖卿, 郑磊. 地方政府开放数据的利用与产出研究[J]. 电子政务, 2019(9): 23-31.
- [26] 郑磊, 关文雯. 开放政府数据评估框架、指标与方法研究[J]. 图书情报工作, 2016(18): 43-55.
- [27] 阮霁阳. 数字政府建设影响因素研究——基于127份政策文件的大数据分析[J]. 西南民族大学学报(人文社科版), 2022(4): 185-191.

作者简介

汤志伟, 男, 1969年生, 博士, 教授, 研究方向: 智慧治理。

陈瑶, 女, 1999年生, 硕士研究生, 研究方向: 数字政府治理。

韩啸, 男, 1988年生, 博士, 副教授, 通信作者, 研究方向: 政府数据治理, E-mail: zhengguan_yjs@126.com。

王莉, 女, 1997年生, 硕士研究生, 研究方向: 政府数据治理。

Deconstructing Open Government Data Capabilities from a “Structure-Process” Perspective Based on LDA and ISM

TANG ZhiWei^{1,2} CHEN Yao¹ HAN Xiao¹ WANG Li¹

(1. School of Public Administration, University of Electronic Science and Technology of China, Chengdu 611731, P. R. China;

2. Shenzhen Institute of Advanced Study, University of Electronic Science and Technology of China, Shenzhen 518000, P. R. China)

Abstract: The concept is the starting point of the study. Existing research on open government data is mostly based on knowledge transfer from the private sector and ignores the characteristics of the public sector. So it's necessary to redefine the open government data capability. This paper uses the LDA theme model method to mine text data of relevant domestic policies, identifies the constituent elements of open government data, clarifies the hierarchical relationship and transmission path of various elements in “macro - meso - micro” through ISM method, and finally puts forward countermeasures and suggestions for open government data practice from the perspective of government data openness according to the research findings.

Keywords: Government Data; Open Capacity; Policy; LDA; ISM; Code

(责任编辑: 王玮)