

1

2022

总第48期（双月刊）

Journal of
Information Resources
Management

信息资源管理学报

CSSCI 来源期刊

中国人文社会科学期刊AMI
综合评价（A刊）核心期刊

主办 /

武汉大学
Wuhan University

主管 /

中华人民共和国教育部
Ministry of Education

JIRM
INFORMATION

信息资源管理学报

2011年6月创刊
第12卷 总第48期
2022年第1期
(双月刊)

本刊声明

本刊所发论文被CNKI中国期刊全文数据库、万方数据数字化期刊群、维普中文科技期刊数据库、CEPS中文电子期刊资料库、国家哲学社会科学学术期刊数据库收录。凡不愿入选的稿件，请作者投稿时予以说明。

编辑委员会

主任 马费成

委员 (以姓氏汉语拼音为序)

曹树金 陈超美 陈国青 陈江萍 邓胜利 冯惠玲 黄萃 黄慕萱
黄水清 李广建 李一军 李月琳 刘细文 卢龙 陆伟 毛基业
冉从敬 苏新宁 孙建军 王芳 王日芬 吴江 夏立新 夏义堃
谢新洲 叶鹰 曾建勋 赵捧末 周晓英 周耀林 朱庆华

主编: 李纲

副主编: 王晓光 孙永强

编辑部主任: 于媛

编辑 (以姓氏汉语拼音为序)

陈一 董克洪 亮 胡吉明 姜婷婷 毛进 张晋朝

主管单位

中华人民共和国教育部

主办单位

武汉大学

承办单位

武汉大学信息管理学院

武汉大学信息资源研究中心

编辑出版

《信息资源管理学报》编辑部

国际标准连续出版物号: ISSN 2095-2171

国内统一连续出版物号: CN 42-1812/G2

邮发代号: 38-544

印刷单位: 武汉人元印务有限公司

发行范围: 国内外公开发行

发行单位: 湖北邮政报刊发行局

国外发行: 中国国际图书贸易集团有限公司

编辑部地址: 武汉市武昌珞珈山 武汉大学信息管理学院

邮编: 430072

电话: (027) 68754779

在线投稿: <http://jirm.whu.edu.cn>

E-mail: xxzyglxb@163.com

出版日期: 2022年1月26日

定价: 每期 39元 全年 234元

特约稿

- 004 元宇宙视域下的用户信息行为：框架与展望 吴 江 曹 喆 陈 佩 等

学术专栏·信息行为

- 021 回望“十三五”：国内信息行为研究的特点、
不足与展望 李月琳 张建伟 王姗姗 等
- 034 在线健康信息替代搜寻过程实验研究：特征分析及
阶段识别 宋小康 赵宇翔 李新月 等
- 046 基于ISM-MICMAC的移动社交媒体倦怠影响因素与
关联路径研究 万 震 邹 凯 张艳丰 等
- 056 基于CPM理论的移动商务用户信息披露意愿影响因素研究 刘百灵 董景丽

研究论文

- 067 三角角力与公私对列：美国社交平台内容治理研究 谢新洲 宋 琢
- 080 数据垄断的演化与分类
——基于国内公共数据与商业数据 邓 崧 吕雨婷 杨 迪
- 091 社会科学研究的政策影响力：基本理论模型及评价实践 李梦茹 金学慧 孙若丹 等
- 101 基于信号理论的开放式创新平台创意认可度研究 王玉洁 戚桂杰 王凯平等
- 116 基于多源信息融合的产业链关键核心技术主题识别研究
——以人工智能领域为例 李维思 谭力铭 章国亮 等
- 127 基于状态空间搜索的科技服务智能拼接和组配研究 赵一鸣 朱永春

阅读书单

- 136 思想的越野
——2021年度信息资源与数据管理阅读书单



欢迎关注
信息资源管理学报

Papers on Invitation

Users' Information Behavior from the Perspective of Metaverse: Framework and Prospect Wu Jiang, Cao Zhe, Chen Pei, et al 004

Special Column: Information Behavior

A Review for Information Behavior Research During the Time Period of the 13th Five-Year Plan in China Li Yuelin, Zhang Jianwei, Wang Shanshan, et al 021

Experimental Research on the Process of Online Surrogate Health Information Seeking: Character Analysis and Phase Identification Song Xiaokang, Zhao Yuxiang, Li Xinyue, et al 034

Research on Influencing Factors and Associated Path of Mobile Social Media Fatigue Based on ISM-MICMAC Wan Zhen, Zou Kai, Zhang Yanfeng, et al 046

Research on Factors Influencing Mobile Commerce Users' Information Disclosure Intention Based on CPM Theory Liu Bailing, Dong Jingli 056

Research Paper

Triangular Conflict and Public-Private Confrontation: Research on Content Governance of U.S Social Media Platform Xie Xinzhou, Song Zhuo 067

The Evolution and Classification of Data Monopoly: Based on Domestic Public Data and Commercial Data Deng Song, Lyu Yuting, Yang Di 080

Policy Impact of Social Science Research: The Theory Models and Evaluation Practice Li Mengru, Jin Xuehui, Sun Ruodan, et al 091

Research on the Idea Recognition of Open Innovation Platform Based on Signal Theory Wang Yujie, Qi Guijie, Wang Kaiping, et al 101

Research on Topic Recognition of Key Core Technology in Industrial Chain Based on Multi-source Information Fusion: Taking AI as an Example Li Weisi, Tan Liming, Zhang Guoliang, et al 116

Intelligent Splicing and Assembly of Science and Technology Services Based on State Space Search Zhao Yiming, Zhu Yongchun 127

Reading List

Free Running of Thoughts: Reading List of Information Resources and Data Management in 2021 136

· 研究论文 ·

数据垄断的演化与分类

——基于国内公共数据与商业数据

邓 崧 吕雨婷 杨 迪

(云南大学政府管理学院,昆明 650500)

[摘要] 数字经济和大数据的飞速发展使得社会中的数据量呈现爆炸式增长,与此同时,数据垄断成为了影响经济社会健康发展的重要问题。通过文献分析、模型建构以及案例分析等方法,从数据垄断的演化与分类出发,分析数据垄断的影响及其应对。得出以下结论:第一,从公共数据与商业数据两个角度对数据垄断分别进行定义,并将数据垄断及其演化类型分为公共数据发展型、公共数据垄断型、商业数据发展型以及商业数据垄断型四类;第二,数据垄断对经济社会的影响分为有利和不利两个方面,不同类型影响不同,所应当采取的政策措施重点分别为加强数据公开共享、建立数据确权机制、维持市场正常竞争以及事后监管向事前事中监管转变;第三,基于数据汇集程度与市场化程度,构建了数据垄断演化分类的理论模型。

[关键词] 数据垄断 分类模型 商业数据 公共数据 数据共享 数据确权 数字经济

[中图分类号] G203; G255 [文献标识码] A [文章编号] 2095-2171(2022)01-0080-11

DOI :10.13365/j.jirm.2022.01.080

The Evolution and Classification of Data Monopoly :Based on Domestic Public Data and Commercial Data

Deng Song Lyu Yuting Yang Di

(School of Government, Yunnan University, Kunming, 650500)

[Abstract] The rapid development of the digital economy and big data has caused an explosive growth in the amount of data in society. At the same time, data monopoly has become an important issue affecting the healthy development of the economy and society. Through literature analysis, model construction and case analysis, starting from the classification of data monopoly, the impact of data monopoly and its response are analyzed. The following conclusions are drawn: Firstly, it defines data monopoly separately from public data and commercial data, and divides data monopoly into four types: public data development type, public data monopoly type, commercial data development type and commercial data monopoly type; Secondly, the impact of data monopoly on the economy and society can be divided into two aspects: positive and negative. Different types of impacts are different. The policy measures that should be adopted are focused on strengthening data disclosure and sharing, establishing a data right confirmation mechanism, maintaining normal market competition, and afterwards supervision shifting to ex-ante supervision; Thirdly, based on the degree of

[基金项目] 本文系国家社科基金项目“基于大数据的公共管理基础理论、研究方法与实践案例研究”(19FGLB063)和“云南省高校大数据下的云南公共管理发展研究科技创新团队”项目研究成果之一。

[作者简介] 邓崧,教授、博士生导师,研究方向为数字政府、公共管理方法与技术;吕雨婷,硕士生,研究方向为数字政府;杨迪(通讯作者),硕士生,研究方向为数字政府,Email:18482259956@163.com。

本文引用格式:邓崧,吕雨婷,杨迪.数据垄断的演化与分类——基于国内公共数据与商业数据[J].信息资源管理学报,2022,12(1):80-90.

data collection and marketization, a theoretical model of data monopoly classification is constructed.

[Keywords] Data monopoly; Classification model; Business data; Public data; Data sharing; Data right confirmation; Digital economy

1 引言

数字经济时代,数据作为与土地、劳动力、资本、技术并列的第五大生产要素^[1],所蕴含的价值不断凸显,随之产生的数据垄断问题也引发了社会各界高度关注。垄断主体凭借数据优势破坏数字生态和市场竞争秩序,侵害用户数据权益,阻碍数据共享,损害社会公共利益,因此,数据垄断治理已成为我国当下最为重要的议题之一。2020年底,中共中央政治局会议提出“强化反垄断与防止资本无序扩张”,显示了党和国家领导层对于数据垄断问题的重视。对数据垄断进行治理,首先需要对数据垄断进行全面深刻的剖析,以此寻找数据垄断的有效对策。

我国政府是为人民服务的政府,为加快服务型政府建设进程,提高政府治理能力和为公民提供满意的公共服务的能力,需面向社会公民采集大量数据信息,因此,目前中国信息数据资源80%以上掌握在各级政府部门手中^[2];中国前瞻产业研究院2020年8月发布的《2020年中国数字经济发展报告》显示,自2015年提出“国家大数据战略”以来,我国的数字经济市场规模迅速扩大,截至2019年末,数字经济的总体规模达到了35.8亿元,占GDP的36.2%^[3]。鉴于此,需从公共数据与商业数据两个角度对我国数据垄断进行研究。当前我国商业数据处于大型互联网企业、平台垄断的态势,而公共数据所面临的“数据孤岛”与“数据堰塞湖”则阻碍了数据的共享与利用,数据垄断治理迫在眉睫。然而“垄断”并非是贬义词,数据反垄断的目标及难点并不在于如何彻底消除垄断,而在于如何找到垄断与反垄断之间的平衡点,以兼顾各价值目标以及各社会主体之间的利益。

2 研究现状

数据垄断研究作为一个新兴研究领域,涉及经济学、公共管理学、图书情报学、法学以及

信息技术科学等多种学科^[4-12],具有研究主题多样、涉及学科复杂等特点。因此,为保证文献综述的完整性,从图书情报学的文献计量视角出发,利用VOSviewer和Citespace软件,对北大中文核心期刊以及CSSCI期刊中关于数据垄断的研究进行知识图谱可视化分析,并在此基础上展开文献综述。

2.1 数据来源

文献数据来源于中国知网数据库,采取的检索限制条件为:主题=(“数字垄断”OR“数据垄断”OR“互联网垄断”),不限制时间范围(检索时间2021年5月24日),文献来源类别选择“北大核心”与“CSSCI”,得到有效文献共211篇。

2.2 时间分布

利用VOSviewer软件对关键词进行时间排列,见图1。可以发现关于数据垄断的研究主要集中在2014年以后,研究方向逐渐从对市场结构及其支配地位的分析转变到法律层面,自2018年后则转变为主要对大数据和数字经济的研究。

2.3 研究主题

关键词是一篇研究性论文的核心词汇,能够反映论文研究的主要问题。对数据垄断研究文献的关键词进行共现分析,并对其进行聚类,能够得出该领域研究的主题分布,见图2和图3。

通过对文献的梳理,可以得出现有关于数据垄断问题的研究主要分为理论探讨型、规制对策研究型两种类型。

第一,理论探讨型。一方面是基于法律视角,对数据垄断行为进行定义。吴韬^[13]通过对美国反垄断案件中关于互联网领域相关市场界定的梳理,得出互联网的特殊性决定了互联网反垄断案件的相关市场界定在方法和思路上要与一般行业区分开来的结论。许光耀^[14]使用案例分析法对奇虎与腾讯之间的诉讼案进行分析提出,由于互联网服务的边

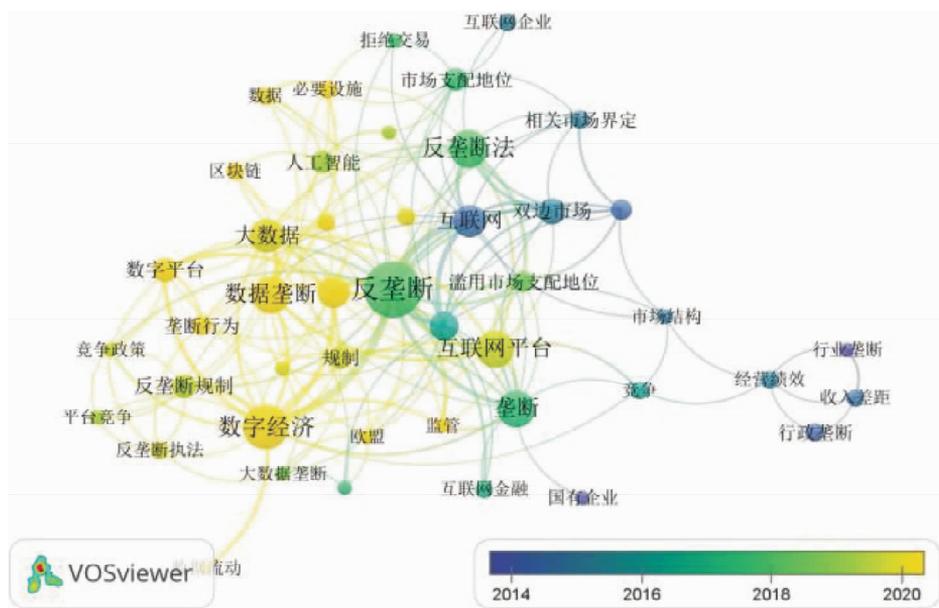


图 1 关键词时序分布图

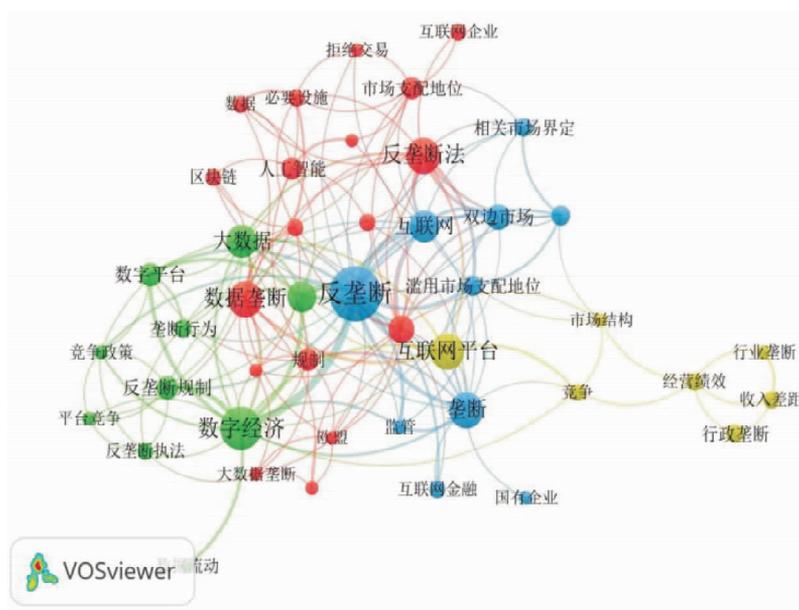


图 2 关键词共现网络知识图谱



图 3 关键词聚类知识图谱

际成本为零,网络效果、锁定效果所造成的转换成本就成了企业市场支配地位认证的主要因素。

另一方面是基于双边市场视角,对互联网平台企业进行针对性研究。Armstrong^[15]最先对双边市场进行了定义,即买者与卖者的交易通过中间平台实现,且其中一方的收益由另一方的数量所决定。在 Armstrong 的理论基础上,王宏涛等^[16]以中国电信业为例,构建了双边市场模型,总结出了中国双边市场的特征,并提出双边市场中价格弹性较小的一方会遭到较高的价格加成。

第二,规制对策研究型。一是基于数字经济发展的视角,研究对数据垄断在制度和技术层面的规制措施。孟小峰等^[17]提出以区块链技术实现数据透明化,增加数据价值实现过程中的透明性,以此应对数据垄断以及由此带来的隐私权侵犯等问题;承上^[18]则提出,由于数字经济与数据垄断行为的复杂性,对数据垄断的规制应当从消费者福利标准和社会福利标准两个角度来进行评判,不同标准应采取不同的规制措施。

二是基于博弈论视角,从价格歧视角度出发,提出抑制垄断的有效途径。李三希、武琦璠等^[19]以垂直差异化双寡头模型为基础,讨论了在大数据时代对于非敏感个人信息保护的利与弊,证明了引入竞争将会使消费者总剩余和社会总福利在无个人信息保护和完全一级价格歧视时达到最大,应当在市场中引入竞争,保障竞争机制;蒋传海^[20]构建两期动态博弈模型,对互联网企业依据消费者购买记录进行竞争性价格歧视的行为进行了研究,结果表明,企业可以实施价格歧视的内在原因是转移成本,并且和统一定价相比,歧视定价将会导致社会总福利损失,因此应当予以限制;胡继晔等^[21]以出行市场为例,分别构建了静态博弈模型、新旧平台的动态博弈模型以及无限次重复博弈模型,提出数字经济时代的反垄断要求对多元利益的兼顾,同时数字平台的反垄断重点并不在于对平台发展的遏制,而在于引导平台回归推动创新、促进社会进步发展。

3 数据垄断的涵义与分类

通过对已有文献的梳理可以发现,以往研究主要是对企业这一市场主体的垄断行为的研究。在中国,数据的收集主体分为以盈利为目的的企业和以维持社会稳定、提供公共服务为目的的政府两部分,正如李克强总理提到“目前我国信息数据资源 80% 以上掌握在各级政府部门手中,‘深藏闺中’是极大浪费”^[2],政府在现实层面确实占据着大量的资源,极易以数据量的堆叠形成公共数据垄断现象。同时,政府部门还存在着“数据孤岛”与“数据堰塞湖”等现象。因此,需要将政府的数据收集行为也纳入数据垄断的理论分析框架,才能获得关于数据垄断的完整认识。数据垄断也应当从政府收集的公共数据和企业收集的商业数据两个角度来进行理解。

3.1 涵义

从公共数据角度来看,政府治理数字化的转变以及服务型政府的建设本身就要求政府掌握大量数据,并依据数据进行正确的科学决策,政府对必要数据的收集天生就具有合法性与合理性。因此,公共数据角度的数据垄断主要从数据的流动性和信息收集的内容来理解,即政府、国企或事业单位在进行数据收集过程中,未履行告知义务并缺乏个人授权,同时由于各单位之间沟通交流渠道的缺乏、处理技术落后、现行行政体制阻碍等原因,导致海量公共数据只能以原始状态分散滞留于各部门,形成数据孤岛,被动形成数据垄断。

从商业数据角度来看,数字经济的飞速发展推动了各大数字平台类企业的蓬勃兴起,也使得此类企业能够进行大量的数据收集,在完成收集后,由于平台之间的竞争,数据具有了排他性,这为数据垄断提供了可能。从该角度来看,数据垄断是指寡头企业对数据资源进行独占、以海量数据支撑自身获得市场支配地位的行为。商业数据垄断多为谋求超额利润,主动进行垄断。

3.2 演化与分类

数据的收集与使用过程中,投入需求最

大的是收集与使用机制的建立过程,在机制建立后,所需求的后续投入相对较小,即具有成本的弱增性,随着规模扩大,由于弱增性将产生规模效应,相对应地,数据的收集与使用将逐渐发展为垄断行为。因此,当不对数据的收集与使用进行特殊性规制时,随着时间与发展程度的推移,必然会完成从非垄断到垄断的演化,而不同时期产生的影响及需求的应对措施大不相同,需要分时期进行研究。

与传统自然垄断行业不同,数据领域的垄断具有政府与市场两个层面,同时市场中的企业具有双边市场特性、交叉网络外部性和产品服务提供的无边界性与可替代性的特点^[22],因此,传统行业关于垄断类型的划分不再适用于数据垄断领域,需要新的分类方法。

数据的收集与利用以市场化程度以及汇集程度为最主要特征,以这两个变量为基础,可以对数据领域的垄断类型进行分类。

由于数据已经成为一种重要的生产要素,其市场化程度的测算借用卢现祥等的研究成果^[23],市场化系数是指单位*i*所收集数据的市场化程度与所有单位收集数据的市场化程度平均值之间的比值。其计算公式为:

$$F_i = \frac{1}{2} \cdot \frac{m_i}{\frac{1}{n} \sum_{i=1}^n m_i} \quad (1)$$

其中, F_i 为市场化系数, m_i 为单位*i*数据的市场化程度, n 为单元数。以0.5为判别标准,当 $F_i > 0.5$ 时,单位*i*数据的市场化程度大于平均市场化程度; $F_i = 0.5$ 时,单位*i*数据的市场化程度等于平均市场化程度; $F_i < 0.5$ 时,单位*i*数据的市场化程度小于平均市场化程度。

汇集度系数是指单位*i*所收集数据量大小与所有单位所收集数据量平均值之间的比值,反映的是单位*i*数据量的聚集程度。其计算公式为:

$$E_i = \frac{1}{2} \cdot \frac{q_i}{\frac{1}{n} \sum_{i=1}^n q_i} \quad (2)$$

其中, E_i 为汇集度系数, q_i 为单位*i*收集的数据量大小, n 为单元数。同样以0.5为判

别标准,当 $E_i > 0.5$ 时,单位*i*收集的数据量大于平均数据量;当 $E_i = 0.5$ 时,单位*i*收集的数据量等于平均数据量;当 $E_i < 0.5$ 时,单位*i*收集的数据量小于平均数据量。

以市场化系数为横轴,汇集度系数为纵轴,1为分界线,可以构建数据垄断演化分类矩阵来分析数据垄断程度及其类型。如图4所示,其中,第I类为公共数据发展型,第II类为公共数据垄断型,第III类为商业数据发展型,第IV类为商业数据垄断型。四种类型的含义和形成机制如下。

3.2.1 公共数据发展型

该类型部门属于数据垄断演化分类矩阵的第I类,所收集数据的市场化程度与数据量都相对较小,还处于数据收集的发展时期,垄断还未能形成。该类型的形成从数据量大小来看,主要是由于部门本身规模大小限制了数据收集量,使得数据量处于一个相对较小的阶段;而从市场化程度来看,其进行数据收集的目的往往是非盈利性的,同时,收集来的数据由于各单位之间对数据共享的抵触以及数据共享技术上的限制等原因,使得数据的价值挖掘受到阻碍,导致市场化程度处于较低的状态。从实际来说,地方政府、小型国企以及小型事业单位大多属于该类型。

3.2.2 公共数据垄断型

公共数据发展型的进一步演化往往就会形成公共数据垄断型,该类型部门属于数据垄断演化分类矩阵的第II类,收集的数据量较大而数据市场化程度较低,已经基本完成了对较大地域或领域的数据垄断,垄断程度较高。该类型的形成从市场化程度上来看,与公共数据发展型原因一样,也是由其数据收集目的的非营利性和数据共享的困难性所决定的;而从数据量大小来看,由于该类型部门通常具有较大的体量,其进行的数据收集活动能够在较大范围内得到支持,因此收集的数据量也通常具有较大规模。从实际来说,中央政府、大型国企、大型事业单位大多属于该类型。

3.2.3 商业数据发展型

与公共数据相同,商业数据也存在着发

