

【信息管理】

信息资源管理理论的继承与创新： 大数据与数据科学视角

朝乐门

【摘要】在大数据时代,信息资源管理理论研究面临着新的挑战与重要变革。本文首先分析大数据时代信息资源管理研究面临的四个主要挑战:研究假定有待重新审视,研究范围亟待进一步拓展,研究范式需要多样化,主要应用场景已发生变迁;其次,探讨了数据科学的研究进展以及数据科学与信息资源管理研究的内在关联;接着,提出了下一代信息资源管理理论的主要变革:从信息资源的管理到基于信息资源的管理,从模式在先到模式在后或无模式,从目标/任务驱动型管理到数据驱动型管理,从计算密集型应用到数据密集型应用,从基于知识的研究范式到基于数据的研究范式;最后,提出了下一代信息资源管理的主要研究问题:基于信息资源的管理与治理、信息资源的保障方法、信息资源的深度加工、信息资源的规范分析、信息资源的快速洞见、信息资源的产品化研发、信息资源的资产化管理。

【关键词】信息资源管理;大数据;数据科学

【作者简介】朝乐门(通信作者)(ORCID: 0000-0001-8963-7507),中国人民大学信息资源管理学院副教授,数据工程与知识工程教育部重点实验室(中国人民大学)研究员,博士生导师,E-mail:chaolemen@ruc.edu.cn(北京100872)。

【原文出处】《中国图书馆学报》(京),2019.2.26~42

【基金项目】本文系中国人民大学科学研究基金(中央高校基本科研业务费专项资金资助)项目“基于数据科学的信息资源管理研究的继承与创新”(编号:18XNB012)的研究成果之一。

0 引言

近年来,信息资源管理研究的主要特征之一是对信息资源管理的外延,尤其是外围或应用问题的关注过多,反而忽视了信息资源管理理论本身,如对信息资源管理特有科学问题的关注程度及自身知识体系的建设力度不够,导致信息资源管理的外延越来越泛,而其内涵却尚无实质性突破,基本上停留于Hussain D^[1]、Horton Jr Forest Woody^[2]、Martin W John^[3]、Ricks Betty R^[4]、Beaumont John R^[5]、孟广均^[6]、马费成^[7]、胡昌平^[8]、赖茂生^[9]、霍国庆^[10]、卢泰宏^[11]、王万宗^[12]等老一辈学者的贡献。信息资源管理及其相关领域的期刊论文和图书专著的选题也证明了这一点,多数研究停留在计算机科学、管理学、统计学、社

会学、哲学、心理学、传播学等学科在信息资源管理中的应用问题。从整体上看,近几年来信息资源管理领域的研究广度越来越大,但研究深度却停滞不前,其学科竞争力并未得到显著提升。

为此,本文以大数据时代为研究背景,以数据科学为主要理论基础,重新审视信息资源管理学科研究,提出下一代信息资源管理理论的设想。主要内容如下:第1部分深入探讨大数据时代信息资源管理研究的主要挑战;第2部分调查分析数据科学的研究进展及其在信息资源管理研究中的应用现状;第3部分提出下一代信息资源管理理论及其主要变革;最后,提出下一代信息资源管理中的新研究问题及几点个人建议。

1 信息资源管理理论的大数据挑战

信息资源管理理论起源于1970年代的美国政府文书管理(记录派),之后于1980年代进入美国企业管理领域(信息系统派),再后来于1990年代进入图书情报领域(信息管理派),经历了研究对象和研究范围的不断拓展过程,形成了信息资源管理领域的三个主要学派。然而,信息资源管理的不同学派建立在同一个前提之上,即人们开始认识到信息是一种“资源”^[13]。因此,早期的信息资源管理理论研究主要“以资源为中心”开展。但是,由于人们开始意识到用户需求在信息资源管理中的重要地位,信息资源管理研究呈现出“以资源为中心”转向“以用户为中心”的演变过程^[14]。从研究对象看,信息资源管理理论已发展成为五个主要研究领域:政府信息资源管理、企业信息资源管理、图书馆信息资源管理、公共信息资源管理和网络信息资源管理。

大数据时代的到来给信息资源管理学科带来了“新机遇”与“新挑战”。其中,“新机遇”主要表现在两个方面。一是原先我们无法(或不可能)找到的“数据或信息”,现在可能找到;原先难以完成的计算需求(如数据的实时分析),如今可以实现。二是信息和数据的问题重新受到高度关注,数据科学、大数据技术等与信息资源管理相关的学科开始兴起,信息资源管理学科的发展得到了新动力。“新挑战”主要体现在,原

先我们一直认为“正确或最佳”的理念、理论、方法、技术和工具越来越凸现出其“局限性”,在大数据时代我们需要改变思考模式和研究范式^{[15][1]}。传统信息资源管理理论在大数据时代所面临的主要挑战如图1所示。

1.1 研究假定有待重新审视

大数据时代的到来给信息资源管理带来的第一个挑战是“需要重新审视信息资源管理学科的研究假定或大前提”。对传统信息资源管理理论进行深度分析,我们不难发现,信息资源管理研究的三个学派具有各自的研究假定或研究前提。

(1)记录学派普遍认为信息资源是“死”的或“被动”的,而人是“活”的,因此,需要事先定义其结构,然后按照已定义好的模式要求进行信息资源的采集、组织、存储、检索、利用与服务。记录派的起源文献《联邦报告法案》(The Federal Reports Act)(1942)和《文书消减法》(Paperwork Reduction Act)、《A-130号通报》(Management of Federal Information Resources)^[16]均体现了这一特点。在这种理念指导下的信息资源管理普遍认为,信息的无序、冗余和不一致性将导致后续处理的复杂度,需要对其进行预处理,重视的是信息和数据的强一致性。

(2)信息系统学派所关注的是“计算密集型问题”。信息系统派起源于1980年代中期,其研究往往是面向计算密集型问题——计算是信息资源管理研

	(传统)信息资源管理理论	所面临的主要挑战
研究假定	- 信息资源是“被动”的 - 信息资源管理的难点在于“计算密集型问题” - 信息是“稀缺”资源	- 人们对信息资源的关注点从其“被动属性”转向“驱动属性” - 信息资源管理的难点从“计算密集型”转向“数据密集型”问题 - 信息不是“稀缺”资源,人类已进入数据富足供给时代
研究责任	“信息资源(本身)的管理”为主要研究责任	人们的主要关注点从“信息资源(本身)的管理”转向“基于信息资源的管理”
研究范式	以“知识范式”为主	“数据范式”成为大数据时代的新研究范式,并非常适合于信息资源管理学科领域
主要应用场景	- 信息是一种重要资源 - 信息和数据供给不足 - 以文书、文献和记录为载体	- 信息不仅是重要资源,更具有资产属性 - 信息和数据供给富足 - 信息载体已拓展至图像、多媒体、音频、视频、富媒体等

图1 信息资源管理研究面临的挑战

究中的主要瓶颈和难点所在。例如,Neiderman 提出的 20 世纪 90 年代信息系统管理领域的 15 个关键问题^[17],绝大多数为计算密集型问题。但是,随着云计算、物联网、移动计算的出现,信息的采集、计算和存储等“计算问题”不再是信息资源管理的瓶颈,而信息资源管理的难点从计算转向数据,“数据密集型问题”成为另一个关注点。

(3)信息管理学派以满足用户的信息需求为中心,认为信息是“稀缺”资源,需要通过搜索和检索技术帮助用户找到其所需信息。传统信息资源管理中查准率和查全率是两个重要指标,决定服务质量的好坏。但是,随着大数据时代的到来,人们进入数据富足供给时代,信息已不再是“稀缺资源”。与此同时,用户体验成为信息服务质量评价的重要依据。如,Aberdeen Group 的调查发现,“页面的显示速度每延迟 1 秒,网站访问量就会降低 11%,从而导致顾客满意度下降 16%”;Google 认为“响应时间每延迟 0.5 秒,查询数将会减少 20%”;Amazon 认为“响应时间延迟 0.1 秒,营业额下降 1%”^[18]。

因此,信息资源管理需要重新审视这些研究假定,探讨大数据时代信息资源管理研究的基本假定和大前提:第一,信息资源是否可以被管理,除了被动属性,它是否也有主动或驱动属性;第二,信息资源管理的难点来自计算还是数据,信息资源管理关注的是计算密集型问题还是数据密集型问题;第三,信息资源管理研究的时代背景是信息资源相对匮乏还是信息资源已相对富足,是满足用户的显性需求还是隐性需求。这些问题的回答对于信息资源管理研究具有重要意义。

1.2 研究范围亟待进一步明确

大数据时代的到来给信息资源管理带来的第二个挑战是“进一步明确信息资源管理学科的研究范围”。传统信息资源管理,不管是广义上还是狭义上,关注的都是信息资源要素(如信息内容、信息设备、信息人员、信息系统、信息网络)本身的管理。例如,信息资源管理之父 Horton 认为,信息资源管理是对一个机构的信息内容及其支持工具的管理^[19]。但是,这种信息资源本身的管理是否真的需要一个独立的领域来研究?信息资源管理领域能否胜任和擅

长?信息资源管理领域的研究现状可能从一个层面反映了这一问题。随着信息资源管理研究的深入,研究范围的不确定性及不断拓展导致了其区别于信息科学、管理科学、信息传播与交流理论、系统科学、社会学的学科定位不清,与其他学科交叉过多,大数据时代其学科竞争力和贡献尚未突显。

为此,大数据时代需要进一步明确信息资源管理学科的研究边界和研究范围,尤其探讨是否要研究“基于信息资源的管理”问题。需要说明的是,“信息资源(本身)的管理”和“基于信息资源的管理”是两个不同的概念,后者表达的是基于信息资源的决策支持、产品开发和流程再造等,超出了以信息资源本身管理为中心的传统信息资源管理的理论范畴。

1.3 研究范式需要多样化

大数据时代的到来给信息资源管理带来的第三个挑战是“继承和创新信息资源管理学科的研究范式”。虽然传统信息资源管理分为记录派、信息系统派和信息管理派等不同派别^[3],但是他们的基本研究范式都是一样的,即采取知识范式,属于理论完美主义范畴。知识范式认为知识才是力量,重视的是如何用知识解读信息资源以及将信息资源转换为知识。但是,将信息转换为知识往往需要对应特定领域的知识,而信息资源领域的专家学者往往缺少领域知识或局限于个别领域知识内,难以做到将数据或信息转换为知识。

与此同时,大数据时代兴起了另一种研究范式——数据范式。知识范式与数据范式的区别在于,前者是一种基于知识的,即从大量实践(数据或信息)中总结和提炼出一般性知识(定理、模式、模型、算法和函数等)之后,用知识去解决(解释)问题,其特点是“问题—领域知识—问题”,即根据“问题”找“领域知识”,并用“领域知识”解决“问题”,具有很强的领域差异性的特点;后者是根据“问题”找“数据”,并直接用“数据”解决“问题”,即在不需要将“数据”转换为“领域知识”的前提下,进行“问题”解决^[14-15],其思维模式是“问题—数据—问题”,其应用往往不局限于特定领域中问题的解决。可见,相对于知识范式,数据范式更具有领域通用性的特点,更适用于包括信息资源管理在内的跨学科问题的解决。

为此,大数据时代的信息资源管理研究需要探讨研究范式的转变或多样化问题,尤其是如何将数据范式应用在信息资源管理研究中,通过知识范式和数据范式的并存,实现信息资源管理理论的继承与创新。数据范式的引入可以较好地弥补以知识范式为主导的信息资源管理理论对其他应用领域知识的高度依赖性,更好地避免信息资源管理专业人才因缺少具体应用领域知识或经验而导致的根本局限性,改变信息资源管理学科实践性不足的现状,最终增强信息资源管理学科的核心竞争力。

1.4 主要应用场景已发生变迁

大数据时代的到来给信息资源管理带来的第四个挑战是“信息资源管理的主要应用场景发生了变迁”。传统信息资源管理是在信息资源比较稀缺的时代背景下产生的,其应用场景是信息资源稀缺且被组织机构(如政府、企业、图书馆等)认为是其主要资源。但是,大数据时代的到来改变了信息资源管理理论的主要应用场景,主要体现在:

(1)大数据时代,信息与数据不仅仅是一种“资源”,而是已成为“资产”。因此,信息资源管理学科应将信息资源作为一种“资产”来管理,而不能仅仅停留在“资源”层次的管理。资产和资源的区别在于,前者具有法律权属、劳动增值、财物价值和可直接市场交易等属性。因此,信息资源管理研究不能脱离于其人文、法律、道德、市场、经济环境。

(2)大数据时代,信息和数据不再处于供给不足状态,开始进入一种富足供给的应用场景。因此,传统信息资源管理中重视的信息资源搜索、组织和检索等问题,已经不再是社会关注的主要问题。相反,如何在信息与数据富足供给时代,从大数据中快速洞见有价值的信息,并将其转换为行动才是新的关注点。因此,信息资源管理的应用场景已从信息资源的“搜索与检索”转向“洞见与增值”。

(3)大数据时代,人们的信息需求不再局限于文书、文献、记录类数据,已拓展至图像、多媒体、视频、音频、富媒体等不同载体的信息资源。传统信息资源管理理论在文献载体类信息的管理方面取得了较大的成就,但是对其他载体类型的信息资源管理研究并不深入。因此,信息资源管理研究的应用场景已从“文献信

息资源”的管理转向“大数据信息资源”的管理。

2 数据科学的兴起及与信息资源管理的关联

近年来,大数据时代的到来催生了一门与信息资源管理学科密切相关的新兴学科——数据科学。数据科学是大数据及其背后的新理念、理论、方法、技术、工具、应用和实践的总称。数据科学的出现为信息资源管理学科领域带来了新的机遇与挑战。

2.1 数据科学

1974年,图灵奖获得者Peter Naur在其著作 *Concise Survey of Computer Methods* 中首次明确提出数据科学(Data Science)的概念^[20]。Drew Conway的数据科学韦恩图^[21]首次解释了其学科定位及其学科交叉性和面向领域实务的特点。目前,数据科学是一门以“数据”,尤其是“大数据”为研究对象,并以数据统计、机器学习、数据可视化等为理论基础,主要研究数据加工、数据管理、数据计算、数据产品开发等活动的交叉性学科。数据科学的主要研究目的有:大数据及其运动规律的解释、从数据到智慧的转化、数据洞见、数据业务化、数据驱动型决策支持、数据产品的研发和数据生态系统的建设^{[22]-8}。

Gartner的数据科学成长曲线(Hype Cycle for Data Science)^[23]较好地反映了数据科学的发展现状——数据科学不同组成部分的发展水平不同:R、模拟与仿真、集成学习、视频与图像分析、文本分析等趋于成熟;语音分析、模型管理、自然语言问答等开始投入实际应用;公众数据科学、模型工厂、算法市场、规范分析处于理论探讨和快速发展之中。然而,基于Hadoop的数据发现等面临着可能要消失的命运。

从学术研究看,目前相关研究主要从两种不同方向进行:一种是首先将数据科学作为一个独立的新学科来研究,系统梳理数据科学的理论体系,最终将数据科学理论应用到不同学科领域。另一种是以不同学科领域为基础,研究特定学科领域(如材料科学、金融、医学、新闻、能源等)中的大数据问题,从不同视角和目的研究数据科学,比较有代表性的是材料数据科学、金融大数据、医疗大数据、数据新闻、能源大数据等。就目前而言,相对于第二种研究方向,第一种研究方向上的研究较少。但是,第一种研究维度聚焦数据科学本身,能够更好地代表数据科学

作为一门独立学科的整体研究水平。目前,在数据科学的第一种研究维度上讨论议题主要集中在数据范式、数据驱动、大数据智能、数据密集型问题、数据加工、用户体验、大数据分析、算法可扩展性、数据产品开发和大数据人才培养等方面^[24]。

2.2 数据科学对信息资源管理的影响

从传统信息资源管理学科领域对数据、信息和知识的内在联系的认识看,数据科学本应是信息资源管理的子学科或新分支,或者信息资源管理领域应是数据科学的主要推动者。但是,现实并非如此。数据科学领域由计算机科学和统计学领域在主导,信息资源管理领域对数据科学的兴起与发展的贡献不大,主要原因在于,近几年信息资源管理领域的多数研究只停留在上层应用和外围拓展类问题,并没有关注自身的核心科学问题。相反,数据科学的出现对信息资源管理研究提出了新的挑战,促使信息资源管理领域需要重新审视其理论基础、学科定位、对数据或信息的认知以及研究范式。因此,信息资源管理领域的专家学者应重新审视信息资源管理理论,将研究重点回归至信息资源管理本身的核心问题,进而提升信息资源管理的学科竞争力和社会贡献。

数据科学与信息资源管理的内在联系是现阶段相关研究的新兴话题之一。John Spacey 认为“数据科学侧重于从数据中发现知识和可用于指导行动的信息,而信息科学侧重的是信息的存储与检索”^[25]。叶鹰和马费成认为“数据科学和信息科学在理论逻辑与技术方法上一脉相承,数据科学继续维持信息

科学基本原理”^[26]。朝乐门在深入探讨数据科学与领域知识之间的内在联系的基础上,提出了数据科学将成为信息科学领域知识的新理论基础,并为其提供重要的方法、技术和工具^[27]。就目前而言,二者在以下问题上已相互渗透。

(1)研究对象。信息资源管理理论研究以信息资源为主要研究对象,但具体实践中通常并非直接研究“信息资源”本身,而是先采用计量和记录方式将“信息资源”转换为“数据”之后进行研究。数据科学是以数据尤其是大数据为主要研究对象。因此,二者在研究对象上存在相互交叉。以信息资源管理中的经典理论——DIKW 模型(Data, Information, Knowledge, and Wisdom 数据/信息/知识/智慧模型)^[28]为例,此模型在数据科学中得到广泛应用,并且演化成为 DIKUW 模型(数据/信息/知识/理解/智慧模型)^[29],表现出二者在研究对象上的交叉特征。

(2)活动和流程。在实际工作中,信息资源的处理流程和数据处理活动并非相互独立,而是同时进行的。也就是说,信息资源管理和数据科学对应的是同一个处理流程及其所包含的活动,而主要区别在于所关注的视角和层次不同。信息资源管理侧重的是从“业务视角”研究上层的粗粒度处理过程,包括用户需求分析、信息来源分析、信息采集与转换、信息组织、信息检索、信息资源开发、信息资源建设、信息资源的传播和服务等关键活动。然而,数据科学侧重于从“数据视角”研究底层的细粒度处理流程(见图2),包括数据化、数据加工、数据分析、数据呈现

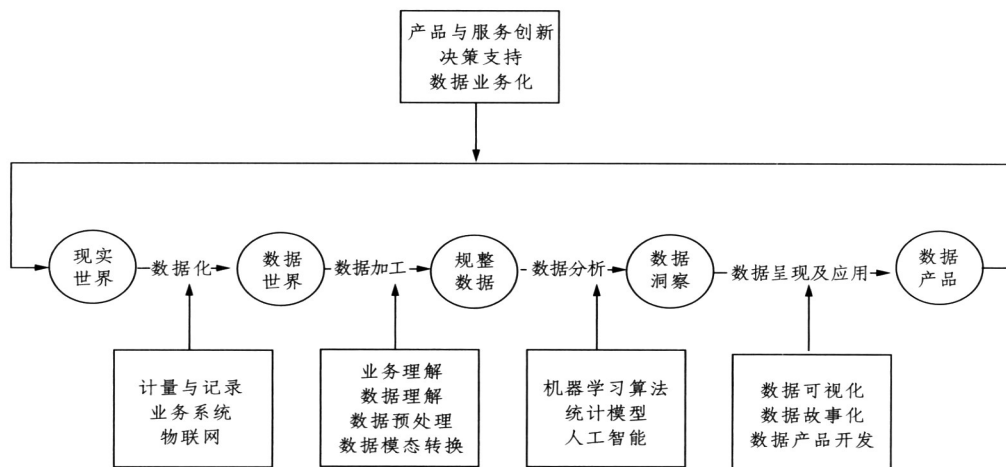


图2 数据科学的基本流程

和应用,以及数据产品的提供^{[22]62}。

(3)理论基础。大数据时代的到来推动了从现实世界到数据世界的映射进程,将原本分散在信息科学、统计学、管理科学、信息论、系统论、控制论、传播与交流理论等底层理论中的“数据问题”从各自学科之中独立出来,逐渐聚焦到一门新兴学科——数据科学。信息资源管理的主要理论基础有信息科学、管理科学、信息传播和交流理论等^[30]。可见,数据科学与信息资源管理的理论基础有很多相互交叉之

处。从长远看,数据科学对信息资源管理的主要影响在于将进一步抽象基础科学中的“信息资源问题”,并为信息资源研究提供新的理论基础,信息资源管理不再直接面对传统理论基础,如图3所示。

3 下一代信息资源管理研究的主要变革

通过对大数据时代信息资源管理学科所面临的挑战以及数据科学与信息资源管理学科之间的内在联系的分析,我们可以总结出下一代信息资源管理研究的主要变革(见图4)。

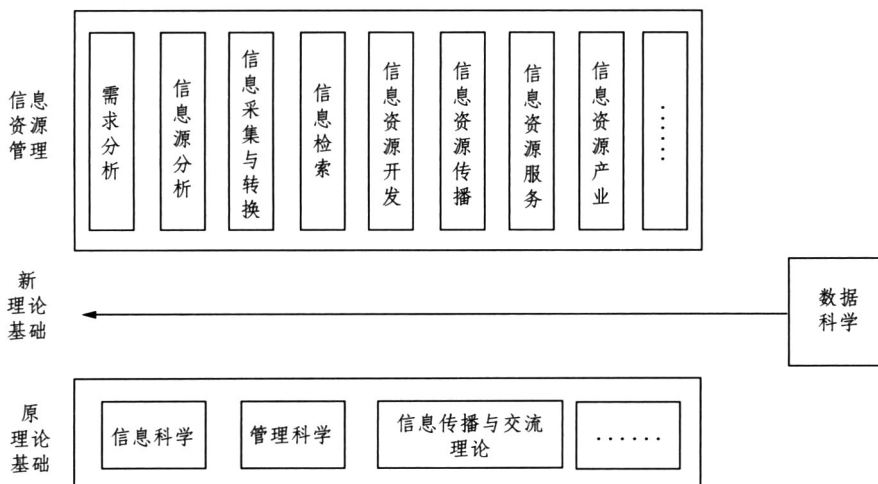


图3 数据科学将成为IRM的新理论基础

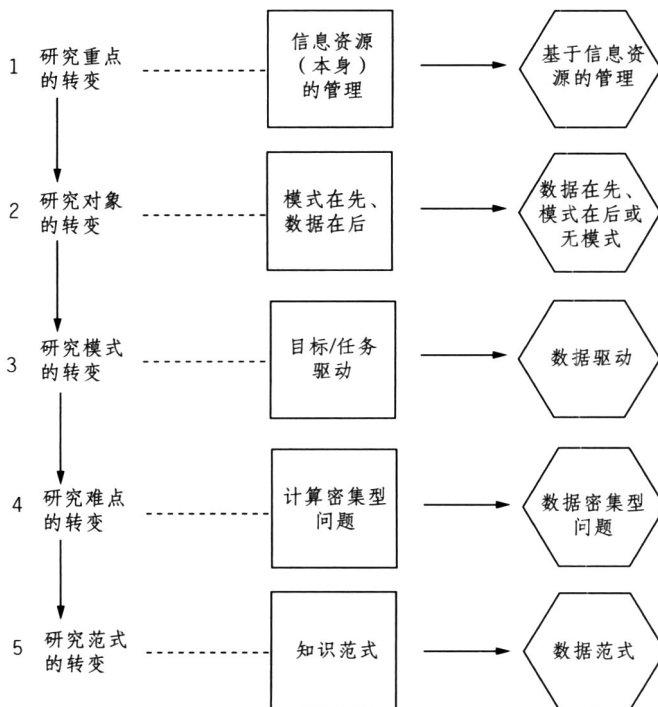


图4 下一代信息资源管理研究的主要变革

3.1 研究重点:从信息资源的管理到基于信息资源的管理

通常,信息资源有狭义和广义之分,“狭义的信息资源”主要是指信息内容,“广义的信息资源”则包括信息内容以及与信息内容相关的信息技术、信息设施、信息人员等^[31]。但是,不管是狭义信息资源还是广义信息资源的管理,二者均关注信息资源本身的管理。以信息资源本身的管理为中心的信息资源管理研究有两个明显的局限:一是从理论研究角度看,信息资源包括信息、人、财、物等多个生产要素,而这些要素的研究往往分散在各自的学科领域,信息资源管理在这些问题上的研究并不占优势。二是从实践应用角度看,信息资源管理并没有脱离其他生产元素,无法独立进行管理,社会上也不存在对应的独立部门或人才岗位需求。

然而,基于信息资源的管理是任何学科都未研究的一个新课题。基于信息资源的管理主要关注的是如何基于信息资源来设计、优化和评价管理活动,其研究范式是数据或信息资源驱动的,而不是组织机构的目标和决策驱动的,可提升组织机构管理和决策活动的敏捷性和客观性,防止出现主观错误和认识性偏差。

从信息资源本身的管理到基于信息资源的管理的转变,对于信息资源管理学科的发展具有以下重要意义:①继承和发展学科理论,“基于信息资源的管理”是“信息资源(本身)的管理”的进一步深化,可明确信息资源管理的主要研究范围;②明确信息资源管理学科的学科定位,将信息资源管理研究的重点从“信息资源(本身)”转向“基于信息资源的管理活动”,提升信息资源管理的学科竞争力。

3.2 研究对象:从模式在先到模式在后或无模式

传统信息资源管理,尤其是狭义信息资源管理的基本思想是“模式在先,数据在后”,其基本做法如下:事先定义组织的信息资源模式(如管理目标、数据结构和模式信息等),然后按照模式要求采集、整理、组织、存储、检索和分析信息资源。但是,在大数据时代,很多信息资源的模式是不断变化的,根本不存在尚未定义模式信息之时就需要“管理”的信息资源。以NoSQL数据库为例,其中存储的信息资源是

以Keyvalue、Key-document、Key column或图形式存在的,其模式信息是非常简单的或根本没有模式信息。因此,信息资源管理需要重视对“模式在后或无模式”信息资源的研究。模式在后或者无模式的信息资源管理的主要特点是信息资源的生成、采集和存储活动不需要事先定义其模式信息,信息资源以无模式形式存储在非关系型数据库之中,信息资源管理活动不再是仅仅处理有模式的数据,信息资源管理研究的大前提和基本流程发生了根本性变化。

无模式或模式在后的信息资源的管理,将会是信息资源管理研究的新议题,其重要意义在于:适应大数据时代的新挑战,转变信息资源管理的研究方向和研究重点,信息资源管理的研究重点从结构化数据的管理转到非结构化数据的管理,这满足了大数据时代的主流应用需求;模式在后或无模式类信息资源的管理将进一步拓展信息资源管理的研究内容,使信息资源管理理论能够顺应大数据时代的主流需求,避免信息资源管理理论研究局限在自己的原有研究范畴。

3.3 研究模式:从目标/任务驱动型管理到数据驱动型管理

目标/任务驱动是传统信息资源管理的主要特点之一,即先制定目标(或任务、需求等),然后将既定目标作为信息资源管理活动的依据,做什么和不做什么取决于既定目标,具有显著的目标驱动特征。与其他类型资源(如材料、能源等)不同的是,信息资源具有可共享、动态演化和可转换的特点。信息资源的可转换性体现在信息资源可以转换成材料或能源资源,它是大数据时代信息资源管理的主要关注点。信息资源向材料资源和能源资源的转换过程通过两种不同的方式进行:一是“加法方式”,即信息资源的开发利用可以提升材料和能源资源的价值或(和)收益;二是“减法方式”,即信息资源的开发利用可降低材料和能源资源的成本或(和)风险。从信息资源到材料资源和能源资源的转换过程具有较强的实时性,其关键在于如何实现数据驱动^[32]或信息资源驱动,即如何基于信息资源自身的特点即运动过程来驱动其管理活动。

相对于上述目标/任务驱动型信息资源管理而

言,数据驱动型信息资源管理的主要优点体现在:

①提升信息资源管理工作的客观性,避免管理活动过于主观;②提升信息资源管理工作的敏捷性,使管理过程随着信息资源本身发生改变,更好地适应不断变化的应用环境;③实现信息资源管理的智能化,通过“数据驱动+人工智能”的方式,信息资源管理工作更具有智能特点,实现信息资源向智慧的转化。

3.4 研究难点:从计算密集型问题到数据密集型问题

传统信息资源管理侧重的是计算密集型问题,个人计算机、Internet、分布式计算、并行计算、网格计算、物联网、云计算等计算技术曾是信息资源管理领域的研究热点。这种研究传统导致了信息资源管理专业与计算机科学专业的交叉过多,其学科定位越来越模糊,面临学科竞争优势越来越弱的潜在风险。其实,信息资源管理理论的创始人以及早期代表性研究者对该学科的定位是非常明确的,即以信息内容(如记录、文书等)为其研究中心,关注的是信息层面的问题,而并非计算层面的问题。但是,自1980年代信息系统学派兴起以来,IT架构、文件系统、网络传输、计算模式等计算层次的问题成为信息资源管理领域的研究重点。

但是,自2000年以来,随着物联网、云计算、NoSQL等技术的广泛应用,数据计算技术得到突破性进展,计算问题不再是包括信息资源管理在内的管理活动的主要瓶颈,人们主要关注点从计算密集型问题转向数据密集型问题,数据成为大数据时代的主要矛盾。在这种时代背景下,信息资源管理的研究需求从计算密集型应用转向数据密集型应用,必须将主要关注点放在数据层面的问题,而不应局限于计算层面的问题。计算密集型问题是计算机类学科领域擅长的,数据密集型问题才是信息资源管理类学科领域擅长的。

从计算密集型应用转向数据密集型应用,是信息资源管理理论回归至其本位的主要表现之一,可以避免信息资源管理学科的外延不断放大和内涵不断掏空;另则可以凸显学科竞争力,明确信息资源管理学科领域与计算机类学科领域在研究层次和研究视角上的差异性;同时也有助于重新审视信息资源管理

中的核心问题,避免其在应用层次上的片面发展。

3.5 研究范式:从知识范式到数据范式

根据范式理论的创始人库恩的观点,“特定的科学共同体从事某一类科学活动所必须遵循的公认的‘模式’,包括共有的世界观、基本理论、范例、方法、手段、标准等等与科学研究有关的所有东西”^[33]。马费成教授通过基本范例、技术工具、共同信念和价值实现四个要素对信息资源管理领域的学科共同体进行特征分析^[34]。从根本上讲,信息资源管理中的传统研究范式均属于“知识范式”。与此不同的是,数据科学领域重视的是另一种研究范式——“数据范式”,最有代表性的是图灵奖获得者Gray提出的科学研究第四范式——数据密集型科学发现(The Fourth Paradigm: Data-Intensive Scientific Discovery)^[35]。

传统信息资源管理理论遵循的“知识范式”,主要表现有两种:一是试图从大量信息资源中提炼和(或)挖掘出新知识之后,用新知识来解决信息资源管理问题。这种模式导致信息资源管理领域不得不研究数据挖掘、知识发现等计算机科学领域擅长的问题;另一种是信息资源管理领域总是试图用“知识”来解决问题,如政府信息资源管理和企业信息资源管理,试图用政府管理和企业管理的知识来指导信息资源管理问题。然而,信息资源管理领域的人才并非真正掌握政府管理或企业管理的知识与经验,导致信息资源管理研究难于深入并取得实质性进展。因此,从知识范式到数据范式的转变,可以帮助我们差异化定位“专业信息资源管理工作”和“专业中的信息资源管理工作”。前者由于往往并不具备领域知识很难运用“知识范式”来解决问题,而后者因为专业领域知识的优势,解决问题的能力就要强很多。“专业信息资源管理工作”将何去何从以及如何提升其竞争力和贡献力成为信息资源管理领域亟待解决的重要命题。

大数据时代的到来使人们意识到数据范式的可行性和重要性。传统观念“知识是力量”,大数据时代“数据也是力量”,人们可以基于大数据直接(并非将数据转换为知识的前提下)解决问题。近年来,自动驾驶和机器翻译等领域的成功证明了数据范式的可行性和有效性。无论是Gray提出的第四范式,还

是目前广为流行的量化自我(Quantified Self)^[36]行动以及数据驱动型决策支持(Data-Driven DSSs)^[37],均运用了数据科学特有的行动范式——基于“数据”解决问题。数据范式的兴起为信息资源管理研究提供了新思路——如何在不完全掌握特定领域的知识和经验的前提下,用数据范式解决实际问题。

4 下一代信息资源管理理论的主要研究议题

在分析大数据时代信息资源管理理论的主要变革的基础上,我们进一步探讨下一代信息资源管理理论中需要重点研究的新议题,如表1所示。

4.1 基于信息资源的管理与治理

“信息资源的管理”和“基于信息资源的管理”是两个不同的概念。前者的管理对象为信息资源本身,而後者的管理对象并不一定为信息资源。传统信息资源管理主要研究的是信息资源本身的管理,而对基于信息资源的管理的研究并不多。信息资源作为一种生产要素,不仅需要研究如何管理其本身,更需要研究如何基于信息资源进行其他资源的规划、配置、监测、调度和优化。基于信息资源的管理的重点在于如何通过信息资源的开发利用来实现财务、人力、市场、材料等其他生产要素的管理,使这些

管理活动更为敏捷、高效率和低风险。

基于信息资源的管理与治理问题研究具体涉及以下方面。①数据密集型科学发现:将数据范式引入信息资源管理研究中,改变传统的目标或任务驱动型管理模式,使信息资源管理作为一种主动服务,提升信息资源管理学科的社会地位。②信息资源的实时监测及动态优化:通过对企业、政府等主体的内外信息资源的实时监测,不断优化这些主体的信息资源管理活动,提升信息资源管理活动的敏捷性。③基于信息资源的安全策略与应急预案:将信息资源纳入安全策略和应急预案的重点,并把信息资源的跟踪和监测结果作为机构安全和应急预案的重要依据。④信息资源的业务化:通过信息资源的分析、开发与利用,为组织机构定义新的业务活动,洞悉发展机遇,回避潜在风险。⑤基于信息资源的治理:通过信息资源的开发利用,为相关政策、法律、法规和规章制度的制定及其执行过程提供依据,从而规范基于信息资源的管理活动。

4.2 信息资源的持续保障

作为一种重要的生产要素,如何做好企业、政府等组织机构的信息资源保障工作成为大数据时代的

表1 下一代信息资源管理理论的主要研究议题

序号	议题	主要研究问题
1	基于信息资源的管理与治理	数据密集型科学发现
		信息资源的实时监测与动态优化
		基于信息资源的安全策略与应急预案
		基于信息资源的治理
2	信息资源的持续保障	信息资源管理的规划
		信息资源管理成熟度评价
		数据连续性保障
3	信息资源的深度加工	信息资源的规整化处理
		智能信息资源管理技术
		人机协同信息资源加工方法
		具有3C精神的新兴人才培养
4	信息资源的规范分析	模拟、仿真与优化
		实时分析
		分析结果的呈现
5	信息资源的快速洞见	算法与模型
		信息资源意识与信息资源素养
		信息请求的快速响应
		以资源为中心的计算模式
6	信息资源的产品化研发	信息资源的潜在式应用
		信息资源的业务化
		信息资源驱动型服务
		服务体验的重视
7	信息资源的资产化管理	信息资源的市场化开发
		信息资源的产业化利用
		信息资源生态系统

重要课题。信息资源的保障是指根据组织机构的业务需求与环境变化提供所需信息资源,使其具备持续发展的动力和快速应变的能力。如何不断供给主体发展所需的战略和业务信息资源,使信息资源处于可用状态,成为现代组织机构关注的新命题。

信息资源保障需要重点研究的问题有以下方面。

①信息资源管理的规划:根据组织机构的未来发展、业务需求和潜在风险,制定具有针对性的信息资源规划,包括信息资源战略指定和目标管理。②信息资源管理成熟度评价:通过机构信息资源的成熟度评价,发现管理现状与目标之间的差距,不断优化信息资源管理过程和管理活动。③数据连续性保障:以包括数据的可关联性、可跟踪性和可回溯性在内的数据连续性保障为抓手^[38],进行组织机构的信息资源的质量评价,保证其信息资源处于可信和可用状态。

4.3 信息资源的深度加工

信息资源的加工(Wrangling 或 Munging)是信息资源管理领域值得研究的新课题。与信息资源处理(Processing)不同的是,信息资源加工更加重视的是人在信息资源处理过程中的主动作用。信息资源的加工过程并非是常规性处理过程,而是一种不断增值、创新、批判、洞见和发现的过程。信息资源的增值加工过程主要依赖于信息工作者的3C精神,即创造性(Creative)设计、批判性(Critical)思维和好奇心(Curious)提问。可见,信息资源的加工是信息资源管理学科与计算机等相关学科的区别之一。

信息资源的加工需要研究如下具体问题。①信息资源的规整化(Tidying)处理:通过数据柔术(Data Jujitsu)^[39],将信息资源的模态转换为算法和模型所需要的模态,进而将混乱(Messy)信息/数据转换为规整(Tidy Data)信息/数据^[40];②智能信息资源管理技术:将下一代人工智能技术应用于信息资源管理领域,

积极探索面向信息资源管理的人工智能技术;③人机协同的信息资源加工方法:重视混合智能技术(Hybrid Intelligence)^[41]在信息资源管理中的应用,充分发挥人与计算机在信息资源加工过程中的互补性优势;④具有3C精神的新兴人才培养:信息资源加工不仅为信息资源管理理论体系提出了新的研究课题,而且也为信息资源管理人才培养提出了新的要求,即需要培养具有3C精神的信息资源管理专门人才。

4.4 信息资源的规范性分析

根据Gartner分析学价值扶梯模型(Analytic Value Escalator)^[42],从复杂度和价值维度看,数据/信息分析分为四种:描述性分析、诊断性分析、预测性分析和规范性分析。传统信息资源管理中对描述性分析、诊断性分析和预测性分析的研究较多,对规范性分析的研究相对较少。但是,在描述性分析和诊断性分析的问题上,统计学和计算机科学的优势更加明显,而信息资源管理学科的竞争优势并不明显,如表2所示。

规范性分析是指如何通过模拟和仿真技术,实现优化和决策的目的。规范性分析需要将分析对象置于信息资源生态环境中进行分析,识别信息资源的各利益主体,并将其放入业务、道德、法律、市场、政策环境之中,进行不断的模拟和仿真,从而得出最优策略或优化方案。因此,在规范性分析的问题上,信息资源管理学科的优势更加明显。信息资源的规范性研究主要包括:①模拟和仿真:将仿真技术引入信息资源管理中,将信息资源管理活动置入其真实业务、经济、产业、社会、制度、法律环境中,通过不断调整环境参数和信息资源管理活动的变量来求解最优解决方案。②实时分析:传统信息资源管理中分析活动的实时性不高,导致其敏捷性和应用价值受限。为此,下一代信息资源管理中应重视对实时分析的研究,将流处理技术等引入信息资源管理研究中,提升信息资源管理的敏捷性。③分析结果的呈

表2 数据分析的类型及信息资源管理研究

序号	类型	主要关心的问题	价值	复杂程度	在传统信息资源管理理论中的应用和关注程度
1	描述性分析	过去发生了什么?	↓ 低 ↓ 高	↓ 低 ↓ 高	↓ 多 ↓ 少
2	诊断性分析	为什么发生?			
3	预测性分析	未来会发生什么?			
4	规范性分析	如何改变未来?			

现:将数据的可视化技术和故事化描述方法引入信息资源管理中,使信息分析结果易于理解和记忆。

④算法和模型研究:信息资源的分析要求信息工作者必须掌握一定的算法和模型。在传统信息资源管理中,多数信息工作者仅仅掌握方法论和方法,而很少深入至统计模型和机器学习算法。因此,大数据时代的信息工作者需要在掌握信息分析的方法论和方法的基础上,重点研究面向信息资源管理的算法/模型选择、算法优化、模型检验、自动调参、特征工程等关键技术。

4.5 信息资源的快速洞见

在数据科学中,数据洞见(Insights)是指如何从大数据中快速洞见有价值的数据并将其转换为行动的能力。同样,信息资源管理领域也需要这种能力。传统信息资源管理中对信息检索和搜索的研究较多,主要原因在于当时的信息资源的采集、计算和存储能力有限,需要依靠搜索引擎和检索工具采集所需信息。但是,随着物联网、云计算和 NoSQL/NewSQL/关系云技术的广泛采用,信息资源的采集、计算和存储能力大幅度提升,人们开始进入信息资源富足时代,由此凸显了另一个矛盾,即信息资源的洞见——如何从海量资源中快速洞见真正有价值的信息资源并将其应用于自己的业务之中。

信息资源的快速洞见是信息资源管理领域值得研究的重要课题,具体涉及以下研究点。①信息资源意识或信息资源素养:需要培养信息资源工作者的信息资源意识或素养,使信息资源工作者具备善于发现有价值信息的素质,提升他们的竞争力;②信息请求的快速响应:信息资源管理中需要重点研究快速响应和服务能力,对需求变化和环境变化具有较强的应变能力;③以资源为中心的计算模式:信息资源管理中应重点研究以资源为中心的计算模式,将资源作为研究视角、触发点和主要依据。

4.6 信息资源的产品化研发

在数据科学中,所谓的数据产品(Data Products)并不限于“数据形态”的产品,而泛指“能够通过数据来帮助用户实现其某一个(些)目标的产品”^[39]。信息资源管理领域亟待研究如何基于信息资源开发出新的信息产品或通过信息资源实现传统产品的更新换

代。信息资源产品并不仅限于信息或内容形态的产品,而是通过信息资源来满足用户需求的产品。Google 眼镜就是一个典型的数据产品或信息资源产品,其主要创新和增值来自信息资源,而并非是材料和能源资源。

信息资源的产品化开发中需要重点研究的问题有:①信息资源的嵌入式应用:将信息资源充分应用于材料和能源资源的开发之中,并将信息资源嵌入到传统物质产品之中,从而实现传统产品的创新;②信息资源的业务化:通过信息资源的开发利用,重新定义或优化业务流程,提升机构管理活动的敏捷性;③信息资源驱动型服务:在信息资源管理系统的研发过程中,强调以信息资源为中心的设计模式,通过信息资源驱动方式实现信息资源的主动发现,提升信息资源管理科学的应用价值;④对服务体验的重视:信息资源管理需要重视服务体验,包括实时服务、个性化服务和弹性服务能力。信息资源服务体验的提升需要加强对可视化技术和故事化描述方法引入信息资源管理中的应用研究。

4.7 信息资源的资产化管理

在大数据时代,信息资源的“资产”属性越来越凸显。信息资源管理中需要引入资产管理思想,将信息资源作为机构的重要“资产”来管理。信息资源的资产化属性要求我们在信息资源管理研究中不能忽略来自市场、法律、经济、道德、伦理等赋予信息资源的新特征,需要将信息资源放回其生态环境中进行研究。从“资源化认识”到“资产化认识”是人们对信息认识的一次升华,对于信息资源管理理论研究具有重要意义。

信息资源的资产化管理需要重点探讨以下问题:①信息资源的市场化开发,尤其是信息资源的定价策略和交易机制的研究;②信息资源的产业化利用,包括产业政策、资产化配置及监管机制的研究;③信息资源生态系统研究,将信息资源放入其所在的真实业务和社会环境中,探讨包括信息资源的生产者、保管者、提供者、开发与利用者、消费者、监管者在内的各主体的博弈关系及其变化,通过规范分析,进一步优化这种博弈关系,提升信息资源管理理论的实践意义。

5 结语

如何继承和发展信息资源管理理论是值得我们探讨的一个重要课题。作为一门独立的学科,信息资源管理研究拥有不同的研究学派,相互之间呈现出深度融合的趋势。信息资源管理研究在信息资源的认识、信息资源的开发利用、信息资源产业、文献信息资源的分析、网络信息资源的管理方面积累了大量研究成果,奠定了信息资源管理学科的发展基础。因此,我们需要更好地学习前辈们提出的信息资源管理理论,掌握其精髓,并将其应用于自己的信息资源管理业务之中。

大数据时代的到来,尤其是数据科学的出现,为信息资源管理学科带来了新的机遇与挑战。机遇主要体现在包括信息资源在内的大数据重新受到广泛重视,而挑战在于信息资源管理学科的创新。信息资源管理领域的专家学者应抓住这一机会,重新审视信息资源管理研究,回归信息资源管理的本位问题,进而提升信息资源管理的学科竞争力和社会贡献。因此,本文提出下一代信息资源管理理论的目的并不是推翻传统信息资源管理理论,而是通过抛砖引玉,共同探讨如何在大数据环境下继承和发展信息资源管理理论。

参考文献:

- [1]Hussain D, Hussain K M. Information resource management[M]. Richard D Irwin, 1984.
- [2]‘Woody’ Horton Jr F. Information resource management in public administration: a decade of progress[C]//Aslib proceedings. MCB UP Ltd, 1985, 37(1): 9-17.
- [3]Martin W J. The information society[M]. Information Today Inc, 1988.
- [4]Ricks B R, Gow K F. Information resource management: a records systems approach[M]. South-Western Thomson Learning, 1988.
- [5]Beaumont J R, Sutherland E. Information resources management: management in our knowledge-based society and economy [M]. Butterworth-Heinemann, 1992.
- [6]孟广均,霍国庆,罗曼.信息资源管理导论[M].北京:科学出版社,1998.
- [7]马费成,李刚,查先进.信息资源管理[M].武汉:武汉大学出版社,2001.
- [8]胡昌平,邓胜利,张敏.信息资源管理原理[M].武汉:武汉大学出版社,2008.
- [9]赖茂生.信息资源管理教程[M].北京:清华大学出版社,2006.
- [10]孟广均,霍国庆,罗曼.信息资源管理导论[M].北京:科学出版社,1998.
- [11]卢泰宏.信息资源管理:新领域和新方向[J].情报资料工作,1994(1):8-12.
- [12]王万宗.技术进步与信息管理学学科群的形成[J].情报理论与实践,1998,21(3):129-131.
- [13]Bergeron P. Information resources management[J]. Annual Review of Information Science and Technology (ARIST), 1996, 31: 263-300.
- [14]马费成,王晓光.信息资源管理研究及国际前沿:情报学研究进展[M].武汉:武汉大学出版社,2007.
- [15]朝乐门.数据科学[M].北京:清华大学出版社,2016.
- [16]Sprehe J T. OMB Circular No. A-130, the management of federal information resources: its origins and impact[J]. Government Information Quarterly, 1987, 4(2): 189-196.
- [17]Niederman F, Brancheau J C, Wetherbe J C. Information systems management issues for the 1990s[J]. MIS Quarterly, 1991: 475-500.
- [18]Petterle S. The need for speed: is your brand investing enough in digital performance[EB/OL].[2018-09-20]. <https://www.itproportal.com/features/the-need-for-speed-is-your-brand-investing-enough-in-digital-performance/>.
- [19]Horton F W J. Information resources management: concept and cases[M]. Association for Systems Management, Cleveland.

land, 1979.

[20]Naur P. Concise survey of computer methods[M]. Student-litteratur AB: 1974.

[21]Overton J. Going Pro in data science[M]. O'Reilly Media, Inc, 2016: 12.

[22]朝乐门. 数据科学理论与实践[M]. 北京: 清华大学出版社, 2017.

[23]Gartner. Hype cycle for data science, 2016[EB/OL].[2017-07-25]. <https://www.gartner.com/doc/3388917/hype-cycle-data-science>.

[24]朝乐门, 邢春晓, 张勇. 数据科学研究的现状与趋势[J]. 计算机科学, 2018, 45(1): 1-13.

[25]Spacey J. Data science vs Information Science[EB/OL].[2018-09-20]. <https://simplicable.com/data-sciencevs-information-science>.

[26]叶鹰, 马费成. 数据科学兴起及其与信息科学的关联[J]. 情报学报, 2015(6): 575-580.

[27]朝乐门, 卢小宾. 数据科学及其对信息科学的影响[J]. 情报学报, 2017, 36(8): 761-771.

[28]Rowley J E. The wisdom hierarchy: representations of the DIKW hierarchy[J]. Journal of Information Science, 2007, 33(2): 163-180.

[29]David. The future of enterprise: big data[EB/OL].[2018-09-20]. <http://davidgildeh.com/the-future-of-enterprise-big-data>.

[30]孟广均. 信息资源管理导论[M]. 北京: 科学出版社, 2008: 89-121.

[31]霍国庆, 孟广均, 王进孝, 等. 信息资源管理思想的升华[J]. 图书情报工作, 2002(4): 26-39.

[32]Leonelli S. Introduction: making sense of data-driven research in the biological and biomedical sciences[J]. Studies in History & Philosophy of Science Part C Studies in History & Philosophy of Biological & Biomedical Sciences, 2012, 43(1): 1-3.

[33]Kuhn T S. The structure of scientific revolutions[M]. Chicago: University of Chicago Press, 1963.

[34]马费成. IRM-KM 范式与情报学发展研究[M]. 武汉: 武汉大学出版社, 2008.

[35]Hey T, Tansley S, Tolle K M. The fourth paradigm: data-intensive scientific discovery[M]. Redmond, WA: Microsoft Research, 2013.

[36]Swan M. The quantified self: fundamental disruption in big data science and biological discovery[J]. Big Data, 2013, 1(2): 85-99.

[37]Power D J. Understanding data-driven decision support systems[J]. Information Systems Management, 2008, 25(2): 149-154.

[38]朝乐门. 数据连续性: 未来跨学科研究的重要课题[J]. 情报学报, 2016, 35(3): 227-236.

[39]Patil D J. Data Jujitsu: the art of turning data into product[M]. O'Reilly Media, Inc., 2012.

[40]Wickham H. Tidy data[J]. Journal of Statistical Software, 2014, 59(10): 1-23.

[41]Ruhe G. Hybrid intelligence in software release planning[J]. International Journal of Hybrid Intelligent Systems, 2004, 1(1-2): 99-110.

[42]Gartner's business analytics framework[EB/OL].[2018-09-20]. <https://www.flickr.com/photos/27772229@N07/8267855748>.