

管理决策情景下的一类新型研究范式

——大数据驱动范式

研究成果：管理决策情境下大数据驱动的研究和应用挑战——范式转变与研究方向

作者：陈国青，吴刚，顾远东，陆本江，卫强

发表于《管理科学学报》，第21卷第7期，2018年7月

近日我院 EMC 讲席教授、管理科学与工程系陈国青教授等在《管理科学学报》(2018.7)撰文，就管理决策情景下大数据相关的研究与应用挑战及其新型范式（大数据驱动范式）进行了论述。

首先，文章从数字化生活乃至大数据时代的背景出发，指出随着数据世界可以越来越多地反映现实世界，许多传统的管理正在变成数据的管理，许多传统的决策正在变成基于数据分析的决策。这对于管理学研究也正在产生深刻影响。

接着，文章阐述了大数据的特征，从大数据的数据特征、问题特征、管理决策特征的视角旨在回答三个基本问题：一，从信息技术范畴来看，大数据具备什么数据属性？二，从研究和应用问题特点来看，什么样的问题可以称作大数据问题？三，从管理者关注来看，管理决策大数据问题具有哪些关注视角？特别地，一个问题可以被看作是一个大数据问题，其应该至少具备粒度缩放、跨界关联、全局视图等要素特点。

附：文章全文

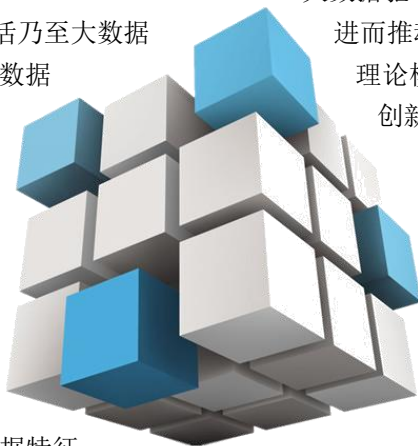
进而，文章描述了大数据驱动范式所体现的“数据驱动+模型驱动”的“关联+因果”内涵。概括说来，大数据驱动范式可以从三个方面来审视：外部嵌入、技术增强、使能创新。

大数据驱动范式通过技术增强引入新视角，进而推动了新型变量关系、要素机理和理论模型构建，并提升了大数据使能创新的价值创造。

最后，文章勾勒了一个全景式 PAGE 框架，给出了大数据驱动的管理与决策研究在决策理论、分析技术、资源治理、使能创新等领域的一系列重要研究方向。

陈国青教授现为清华经管学院学术委员会主席，并担任国家自然科学基金委重大研究计划（大数据驱动的管理与决策研究）指导专家组组长、国家信息化专家咨询委员会成员。

关键词：大数据特征、大数据驱动范式、全景式 PAGE 框架、数字化生活、管理决策



供稿：科研事务办公室

编辑：高晨卉

责编：吴淑媛 赵霞

管理决策情境下大数据驱动的研究和应用挑战^①

——范式转变与研究方向

陈国青¹, 吴 刚², 顾远东², 陆本江¹, 卫 强^{1*}

(1. 清华大学经济管理学院, 北京 100084; 2. 国家自然科学基金委员会管理科学部, 北京 100085)

摘要: 在数字化生活背景下, 传统的管理变成或正在变成数据的管理, 传统的决策变成或正在变成基于数据分析的决策. 从大数据的数据特征、问题特征和管理决策特征出发, 讨论管理决策研究和应用的范式转变. 大数据驱动范式可以从外部嵌入、技术增强和使能创新三个角度来审视, 并体现出“数据驱动+模型驱动”的“关联+因果”含义. 此外, 围绕大数据特征和重要研究方向, 阐述了全景式 PAGE 框架及其要素.

关键词: 大数据; 管理决策; 研究范式; 全景式 PAGE 框架

中图分类号: C930 **文献标识码:** A **文章编号:** 1007-9807(2018)07-0001-10

0 引 言

信息科技的飞速发展和深度融合开启了数字化生活的新篇章, 把人们带入了大数据 (big data) 时代. 一方面, 随着各种感应探测技术、智能终端以及移动互联网的广泛应用, 使得社会经济生活的方方面面以更细粒度的数据形式呈现, 进而整个社会的“像素”得到显著提升; 另一方面, 社会“像素”的提升促进了数字“成像”的发展, 使得通过数据世界可以更清晰地描绘社会经济活动情境, 进而基于数据的商务分析 (business analytics, BA) 正在成为使能创新的核心竞争力. 在此背景下, 传统的管理变成或正在变成数据的管理, 传统的决策变成或正在变成基于数据分析的决策^[1-3].

近年来, 大数据成为学界、政界和业界持续关注的热点. 在学术界, 早在 2008 年和 2011 年, 《Nature》与《Science》杂志分别从互联网技术、互联网经济学、超级计算、环境科学以及生物医药等

多方面讨论大数据的处理与应用^[4,5]. 此后, 大数据在各个学科领域包括医学、经济学、管理学以及公共管理等领域得到了广泛的探讨与研究^[6-9]. 同时, 大数据也引起世界各国高度重视, 美国、欧盟、澳大利亚以及日本等国部署了一系列大数据相关战略和关键领域^[1,10-15]. 在产业界, 国内外大批知名企业掀起了技术产业创新浪潮^[16,17], 通过收购与合作构建和提升大数据技术与应用能力, 布局 and 开拓相关的业态和市场.

我国政府对大数据高度重视并有一系列前瞻性洞见和部署. 2015 年十八届五中全会提出实施国家大数据战略^②, 国务院发布《促进大数据发展行动纲要》^③, 指出大数据是国家基础性战略资源, 旨在全面推进我国大数据发展和应用, 加快建设数据强国. 2017 年十九大报告进一步强调要推动互联网、大数据、人工智能和实体经济深度融合^④. 通过国家需求、政策支持、产业结合以及企业研发等形式, 近些年来涌现出一大批重大规划

① 收稿日期: 2018-03-06; 修订日期: 2018-05-17.

基金项目: 国家自然科学基金资助项目 (91646000; 71490724).

通讯作者: 卫 强 (1974—), 男, 上海人, 博士, 副教授, 博士生导师. Email: weiq@sem.tsinghua.edu.cn

② <http://www.caixin.com/2015-10-29/100867990.html>

③ http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm

④ http://www.xinhuanet.com/2017-10/27/c_1121867529.htm

和政产学研项目,包括国家自然科学基金委员会(NSFC)于2015年9月启动的“大数据驱动的管理与决策研究”重大研究计划(简称NSFC大数据重大研究计划,参见附注)^⑤。

大数据在给社会经济生活带来深刻变革的同时,也对管理与决策研究带来一系列新的重要课题。从信息技术(IT)范畴来看,可以从两个视角来认识大数据,即大数据的“造”与“用”视角(如图1所示)。这和产品的属性类似,一方面,人们关心产品是如何设计和制造出来的;另一方面,人们关心产品是如何使用和有用的。大数据以IT的形式呈现,通常可以概括为数据和系统(包括算法、应用、平台等)。从造的视角出发,涉及的主要问题包括大数据分析(如画像、学习、推断等)和大数据系统建设(如体系、功能、集成等)。从用的视角出发,涉及的主要问题包括大数据使用行为(如采纳、影响、管理等)和大数据使能创新(如要素、价值、市场等)。

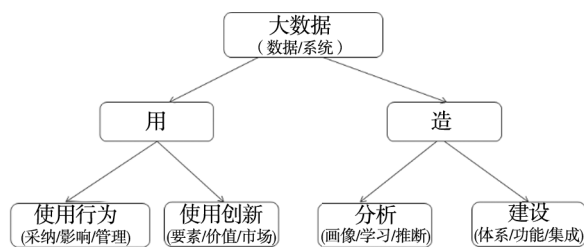


图1 大数据的“造”与“用”视角

Fig. 1 “Make” and “use” perspectives of big data

值得一提的是,大数据相关的研究不仅需要相关领域的理论与应用进行探索和创新,也需要对许多惯常的认识视角和方法论范式进行审视和发展。同时,我国学者和研究人员也面临着“严谨(rigor)与相关(relevance)”(学术规范与实践影响)和“世界与中国”(国际视野与中国根基)既分野又统一的挑战,当然应对这些挑战也为创新机遇开拓了广袤的空间。

1 大数据特征

概括说来,大数据的特征可以从三个方面来描述:数据特征,问题特征和管理决策特征,分别刻画大数据具有的数据属性、大数据问题的特点、

以及管理决策大数据问题的视角。

1.1 大数据的数据特征

大数据作为数据,具有体量大、多样性、(价值)密度低、速率高等属性特征(即4V等特征^[18])。第一,数字化生活各要素的数据生成和交互加速了数据的海量积累,使得数据规模剧增。体量大可以从超规模(即超出传统规模)和问题领域角度来理解,因为规模是与问题领域相关,而不是拘泥于统一量纲标准。例如,市场营销领域的客户满意度调查的传统方式是问卷和访谈,那么进一步考虑海量网上购物评论和社交媒体体验分享的用户生成内容(user generated content, UGC)就构成了一个大数据情境。第二,数字化生活各要素的数据生成和交互丰富了数据类型,使得数据多样性成为常态。多样性强调数据的多源异构和富媒体(如文本、语音、图片、视频等特点)。例如,社交网络上的公众声音、智慧交通平台上的影像信息等均为富媒体形态且来源广泛。第三,数字化生活各要素的数据生成和交互在加速海量积累的同时也减少了价值数据的占比,使得价值发现的难度提升。价值密度低意味着数据挖掘和商务分析是大数据应用的关键。例如,对于在线企业或服务平台来讲,随着网络访问的增加和业务活动的扩展,识别高价值的潜在用户变得相对困难,也凸显出大数据分析的重要性。第四,数字化生活各要素的数据生成和交互强化了流数据形态和即时性,使得数据传输和交换速率显著提高。速率高对平滑流通和连续商务提出了更高要求。例如,智能手机客户端应用软件(Apps)的使用需要在服务内容和效果方面(包括相关内容的浏览、下载、上传、响应、展现等)有良好的临场感和实时体验。

1.2 大数据的问题特征

在各类研究和应用问题中,有一类问题可以归为大数据问题。大数据问题应至少具有以下三个特点:粒度缩放、跨界关联和全局视图。首先,粒度缩放是指问题要素的数据化,并能够在不同粒度层级间进行缩放。这需要通过数据感知、连接和采集获得足够细的粒度性,同时对于不同层级间的粒度转换具有分解和聚合能力。其次,跨界关联

^⑤ <http://www.nsf.gov.cn/publish/portal0/zdyjjh/2015/info49994.htm>

是指问题的要素空间外拓。这需要扩展惯常的要素约束和领域视角,强调“外部性”和“跨界”,在问题要素空间中通过引入外部视角与传统视角联动,将内部数据(如个体自身、企业组织和行业等内部数据)与外部数据(如社交媒体内容等)予以关联^[3]。最后,全局视图是指问题定义与求解的全局性,强调对相关情境的整体画像及其动态演化的把控和诠释。这需要基于数据分析和平台集成的全景式“成像”能力。

在数字化生活的背景下,具有粒度缩放、跨界关联和全局视图特点的应用问题不断涌现,进而激发了大量创新并催生了许多新模式、新业态。例如,在医疗健康领域,传统疾病诊疗中的病人就医关系正在被扩展为融合院外检测、干预、康复数据的新型诊疗模式。其中,不仅涉及传统意义上的生化、影像和诊疗等医院内部数据,也涉及医院外病人和社区相关的体征、体验、社会关系、环境等外部数据。这里,需要获取相关生化组织、疾病、人、社区、环境等微观宏观粒度信息;同时进行视角拓展和关联,包括从科室内外到医院内外的跨界融合;进而,可以在全局层面进行更为有效的诊疗决策和管理。此外,近年来发展迅速的新型医疗健康服务平台,通过整合社会和行业资源,连接医生、公众、医院以及相关上下游企业提供信息咨询、诊疗链入、健康指导等服务产品,形成了一类新业态并呈现显著的大数据问题特征。再如,在新型商务领域,共享单车体现了大数据问题的粒度缩放、跨界关联和全局视图特点。通过车载传感器、定位系统以及智能手机终端等设备获得调度和管理需要的“人—车—路”粒度信息;同时,打通导航、支付、通讯、商铺以及餐饮等诸多业务功能,实现跨界联动;进而,企业和平台可以从全局出发形成整体画像,并优化布局和运作以做出相应的管理决策。

1.3 大数据管理决策特征

一般而言,管理者在业务活动中通常有三个关注:发生了什么(what),为什么发生(why)以及将发生什么(will)。在大数据问题特征的情境下,这三个关注可以从业务层面、数据层面和决策层面进行刻画,进而形成管理决策大数据问题的特征框架(如图2所示)。

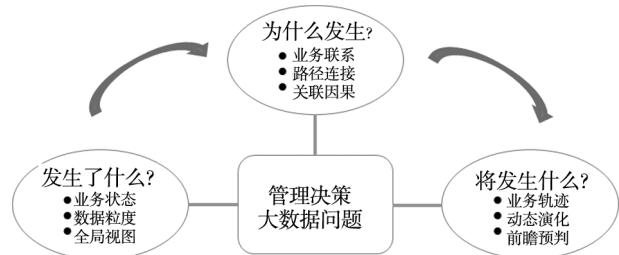


图2 管理决策大数据问题特征框架

Fig. 2 The framework of the big data problem nature for managerial decision-making

首先,对于发生了什么(what)的关注,业务层面需要反映业务的状态,即已经发生或者正在发生的事件和活动(如市场份额、交易现状、KPI表现等);数据层面需要体现业务环节的数据粒度,即现有的数据能否足够支撑管理者对不同粒度层级的业务状态进行了解和把握(如感知、采集、解析、融合等);决策层面需要构建问题的全局视角,即定期整合汇总以及随需要素展现(如:按时统计报表、实时信息查询等)。

接着,对于为什么会发生(why)的关注,业务层面需要反映业务及其要素之间的联系,即业务特定状态的发生与哪些环节和要素有关联;数据层面需要体现不同业务数据路径的连接,即不同粒度层级和跨界关联的业务数据是否有效融通,并能够支持对数据的分析处理(如多维、切分、回溯等);决策层面需要发现关联业务/要素之间的因果关系,即厘清业务逻辑和状态转换机理。在此,特别需要指出的是,在很多情形下,尤其在管理决策领域,大数据需要既讲关联也讲因果。对于许多管理问题而言,如果决策者对事件之间的因果关系没有准确的分析与判断,则难以做出有效的决策,当管理者面临重大决策时更是如此(如投融资、进入新市场、业务转型、结构重组等)。

进而,对于将发生什么(will)的关注,业务层面需要反映业务发展轨迹,即勾勒出由决策或变化导致的业务走向;数据层面需要体现数据的动态演化情况,即对于相关事件进行不确定性动态建模并能够支持智能学习和推断(如模拟、预测、人工智能等);决策层面需要提升前瞻性和风险洞见,即获得决策情境映现和趋势预判能力。

2 大数据驱动范式

系统化管理理论的产生及其发展,包括行为理论、决策理论、权变理论和战略管理等理论体系和管理模型的研究^[19],在提炼管理思想、诠释管理模式和指导管理实践方面发挥了重要作用.长期以来,管理学研究一直以模型驱动范式为领域主流.在模型驱动范式下,研究者基于观察抽象和理论推演建立概念模型和关联假设,再借助解析手段(例如运筹学和博弈论等分析工具)对模型进行求解和优化,或利用相关数据(包括仿真数据、调研数据、观测数据、系统记录数据等)对假设进行统计检验.此外,建立在归纳逻辑基础上的扎根理论等研究范式,传统上强调从文献概括、实地调研、深度访谈中进行定性推演形成理论和认识.

但是,在大数据背景下,一些新的挑战正在涌现^[20,21].这里,以传统的行为模型或计量模型(简称传统模型)为例.第一,传统模型基于观察抽象、理论推演以及经验提炼确定变量(或构念)组合,以此构建变量关系和理论假设,并通过数据实证进行模型检验.然而,在大数据背景下,常常需要检验大量的变量组合(如指数级组合数),这就使得逐一构建传统模型并进行检验成为难以完成的任务.第二,有些重要潜在影响因素和隐变量没有被意识到,因而没有被考虑到传统模型的变量组合中,这常常导致传统模型的假设与数据的适配性不强,模型解释力不高.第三,虽然知道有些影响因素和变量是重要的,但是由于这些因素和变量在传统意义上不可测或不可获(如文本、图像、语音等富媒体数据),难以容纳到传统模型变量组合中,进而造成模型解释力不理想.第四,当样本数据规模大幅增加时,对一些变量的显著性检验有效性下降,可能出现联系缺失或拟合过度等情形.

面对上述挑战,数据驱动范式的优势不断凸显.概括说来,数据驱动范式的作用有两个:一是直接发现特定变量关系模式,形成问题解决方案;二是与模型驱动范式进行补充扩展,形成融合范式.值得指出的是,数据驱动范式发现的一类重要关系模式是关联(association)及其扩展形式(如

关联规则、层次关联、数量关联、时态关联、类关联、模式关联等),并广泛应用到许多领域(如搜索、推荐、模式识别等)^[22].然而,许多管理决策情形不仅需要关联也需要因果,这在一定程度上催生了融合范式及其应用.例如,首先利用数据驱动范式的关联挖掘方法发现变量间的关联,以缩减变量空间和组合规模;进而利用模型驱动范式的行为方法辨识构念影响路径,或计量方法解析变量间的因果关系.这是一个“数据驱动+模型驱动”思路,体现“关联+因果”的诉求,这对于管理决策尤为重要.这里,与传统模型相比一个重要区别是,此时的变量空间中可能存在着一些新颖且潜在的变量及其关联,在进一步融合运用模型驱动方法构建变量关系时存在困难,因为已有的理论知识和领域经验不能直接支持相关的建模逻辑和关系形式.这就需要在更深(包括间接、潜隐)层面上探寻新的变量影响机理和理论,并在方法论上另辟新径(如通过步进/层次/迭代的试错和启发建模方式).

特别地,当数据具有4V等特征并且面对管理决策大数据问题时,考虑数据驱动与模型驱动的结合、管理决策的关联因果特点、使能创新等元素的一类新型范式(在此称作大数据驱动范式)应运而生,并在深入研究与应用过程中得到进一步发展完善.一般而言,大数据驱动范式具有“数据驱动+模型驱动”的“关联+因果”性质.具体说来,大数据驱动范式的框架可从三个角度来审视:外部嵌入、技术增强以及使能创新(如图3所示).前两个角度主要涉及方法论层面,后一个角度主要涉及价值创造层面.

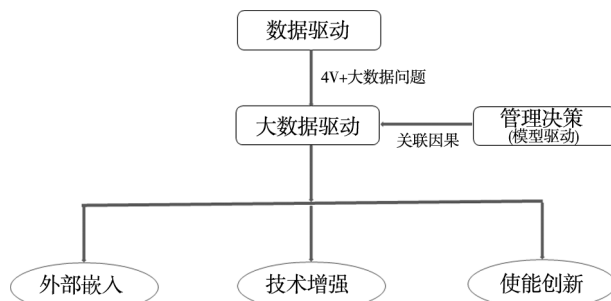


图3 大数据驱动范式框架

Fig. 3 The framework of big data-driven paradigm

2.1 外部嵌入

外部嵌入指外部视角引入,即将传统模型视

角之外的一些重要变量(包括构念、因素等)引入到模型中. 假设自变量集合为 $X' = \{x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_n\}$, 其中 x_1, x_2, \dots, x_m 为传统建模变量, x_{m+1}, \dots, x_n 为通过数据驱动方法新引入的变量(多为富媒体形态). 如果没有变量引入($n = m$), 传统模型的变量关系是 $Y = f(X)$, $X = \{x_1, x_2, \dots, x_m\}$. 在跨界关联情境下($n > m$), 将形成新变量关系 $Y' = f'(X')$. 换句话说, $Y = f(X)$ 可以是 $Y' = f'(X')$ 的特例; 一般意义上讲, $X' \neq X$, $f' \neq f$, $Y' \neq Y$. 显然, 新变量关系的构建面临着深刻的挑战, 既有新变量空间的发现, 又有新视角的洞察, 也有新变量关系的辨识和新理论的生成. 当然, 对于研究和应用来讲, 这些挑战同时也是创新的机遇. 例如, 在金融领域, 可以考虑引入搜索平台上的股票关注数据变量以及社交媒体平台上的相关公共事件数据变量等, 以构建新型股价预测模型; 在商务领域, 可以考虑引入购物平台上的评论数据变量以及朋友圈中的体验和口碑数据变量等, 以构建新型商品营销模型; 在医疗健康领域, 可以考虑引入院外病友智能检测终端数据变量以及区域环境诱因数据变量等, 以构建新型呼吸疾病预防诊疗模型; 在公共管理领域, 可以考虑引入社交平台上的受众意见数据变量以及相关领域联动影响数据变量等, 以构建新型公共政策模型.

2.2 技术增强

对于传统模型来讲, 通过外部嵌入而引入的变量多为富媒体、潜隐性、不可测或不可获, 通常需要利用数据驱动方法和技术. 可以说, 数据和技术意识及其能力是大数据背景下研究和应用的核心竞争力, 也是大数据驱动范式的关键要素. 技术增强旨在提升这样的能力与要素水平.

从大数据的“用”与“造”视角出发, 技术增强具有两方面含义. 一方面, “用”的视角要求管理模型驱动的研究和应用能够增强对外部大数据的敏感性, 引入外部变量并构建其关系; 同时, 能够增强对大数据分析技术的敏感性, 构建方法和工具的获取和使用能力. 研究和应用创新通常体现在通过新型范式开发新的变量关系, 进而形成新的管理学模型和应用(如面向管理问题的新型行为模型或计量模型), 以获得更深入和更具解释

力的管理决策洞见和策略.

另一方面, “造”的视角要求数据驱动的研究和应用能够增强对于管理决策问题的敏感性, 构建面向管理决策问题的方法和技术. 研究和应用创新通常体现在根据管理决策问题特点及其数据属性开发相关性质、测度和策略, 以获得新颖有效的算法和解决方案. 值得指出的是, 这里许多算法(特别是启发式算法和近似解法)需要经过实验数据的验证以评估其效率和效果.

多年来, 不管是“用”的视角还是“造”的视角在数据的使用标准上也经历了一个不断升级的过程, 从模拟数据到标杆数据, 再到相当规模的实际数据, 形成一个逐步丰富和叠加的验证实践. 在大数据情境下, 实际数据的规模化得到了进一步强化. 此外, 在算法比较中, 更关注算法带来的实用效果提升的显著性, 特别在涉及相关用户的场景中, 通常需要进行用户行为实验及其效果感知评测.

在数据类型方面, 富媒体形态(如文本、图像、音频、视频等)成为主流. 其中, 音频数据、视频数据具有时间连续性特点. 由于计算机中通常采用编码、采样等方式表示富媒体数据, 因而数据变换成为大数据分析的重要内容. 常用的数据变换方法包括文本处理的向量空间模型(VSM)^[23]、主题模型(topic model)^[24]、图像处理尺度不变特征转换(SIFT)^[25]、音频处理的短时傅里叶变换(STFT)^[26]、视频处理的时空兴趣点检测(STIP)^[27]等方法. 近年来, 随着大数据平台化运算能力的显著提升, 基于深度神经网络的相关方法进一步发展, 并在富媒体数据变换上展现出良好的应用效果和发展前景. 例如, 用于文本数据的单词嵌入(word embedding)^[28]、用于图像数据的卷积神经网络(CNN)^[29]和胶囊神经网络(capsnet)^[30]、用于音视频等具有时间序列特征数据的循环神经网络(RNN)^[31]、长短时记忆神经网络(LSTM)^[32]等. 其他较新的数据变换方法还包括多层感知机(MLP)、自学习编码器(AE)、受限制玻尔兹曼机(RBM)、深度语义相似模型(DSSM)、神经自回归分布估计(NADE)、生成对抗网络(GAN)等^[33,34].

2.3 使能创新

大数据驱动的一个重要含义是大数据使能(enabling). 大数据能力主要包括大数据战略、大数据基础设施、大数据分析^⑥方法与技术等. 大数据使能是指大数据能力带动的价值创造. 例如, 从研究和应用范式角度看, 外部嵌入是一种使能情形, $Y' = f'(X')$ 中, 大数据能力通过自变量 X' 体现, 创造的价值通过因变量 Y' 体现, 使能转换方式通过 f' 体现. 从研究和应用情境角度看, 企业的价值创造可以体现在其价值链的环节上, 既包括价值链的主环节及其活动, 也包括价值链的支持环节及其活动^[35]. 在企业内外部大数据环境下, 企业使能创新是通过构建大数据能力, 带动新洞察、新模式、新机会的发现, 进而推动产品/服务创新和商业模式创新, 以实现企业的价值创造(如图4所示).

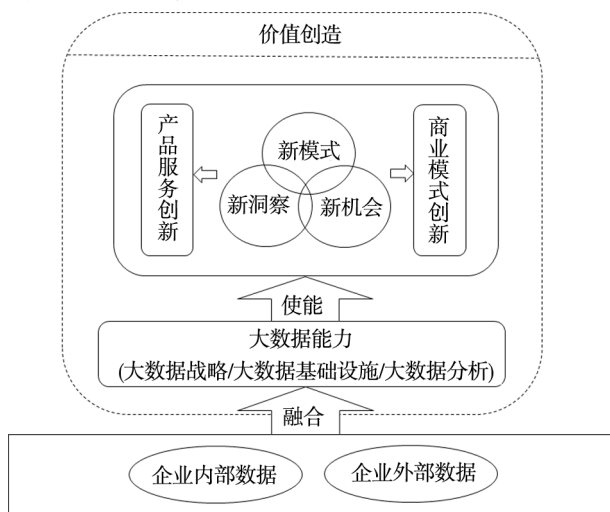


图4 企业大数据使能创新

Fig. 4 Big data enabled innovation for enterprises

综上所述, 大数据驱动范式通过技术增强引入了新视角, 进而推动了新型变量关系、要素机理和理论模型构建, 并提升了大数据使能创新的价值创造. 这对于应对新型商务形态的进一步发展机遇和挑战具有重要意义. 简单说来, 新型商务可以通过两个阶段予以描述. 第一个阶段称作数据商务(digital business 或 data-centric business), 即“数据化 + 商务分析(BA)”. 此时通过细化数据粒度使得商务要素的“像素”显著提升, 并在此基础上进行商务分析, 针对不同管理场景和层次进

行“成像”和决策. 第二个阶段称作算法商务(algorithmic business), 即“商务分析 +”. 此时, 在已有的商务高像素基础上, 成像算法成为关注重点, 旨在获得面对新模式、新业态、新人群^[3]的发展策略和竞争优势. 这里, “商务分析 +”包括 BA 算法创新和 BA 使能创新.

近年来, 人工智能(artificial intelligence, AI)的研究和应用得到了快速发展, 并受到各界的广泛重视. 人工智能自二十世纪 50 年代以来的发展起起伏伏^[36], 虽然在相关思想、模型和方法等方面取得了许多重要进展和成果, 但是由于常常受限于数据基础以及计算能力的不足, 其学习、进化以及推理等方面的能力难以得到发挥, 应用效果也受到限制. 直至进入大数据时代, 人工智能的许多成果得到了工程化和产品化实现, 开始在深度和广度上渗透到社会经济活动中, 并引发人们对于未来产业和人类生存的遐想和担忧. 机器人和智能产品早期用于替代人类简单重复体力性工作, 现在则可以开始尝试用于替代不少复杂并具有智力的工作, 诸如围棋^[37]、翻译^[38]、绘画^[39, 40]、作曲^[40]、作诗^[41]、无人驾驶^[42]、人脸识别^[42]、意念控制^[43, 44]等等. 人工智能在管理领域的应用也初见端倪, 比如财务机器人^[45]、自动金融交易^[42, 45]、竞争智能^[46]、客户服务^[45, 47]、人力资源管理^[48]、市场营销^[42, 45]等等. 毫无疑问, 人工智能将在新型商务中发挥着越来越重要的角色. 另一方面, 伴随着从弱人工智能到强人工智能乃至超人工智能的进阶, 人们对于人工智能应用在隐私和伦理方面的担忧也在不断加重^[49]. 此外, 人工智能理论和技术发展也面临众多挑战(如“黑盒子”特点、学习机理、语义理解等), 这些对于强调“关联 + 因果”的管理决策领域尤为重要.

最后, 管理学是一门融合了“科学”与“艺术”的学科. 在大数据背景下, “科学”层面的可测性、程式化和可重复性等要素正在越来越多地被数据和算法表达; 而“艺术”层面的情感、心理以及认知等要素也开始被不断“量化”, 包括借助一些感知技术(affective technologies)(如眼动、脑电技术等). 未来的管理学在探究组织内外“任务”与

⑥ 在管理与商务领域, 大数据分析也称作商务分析(business analytics, BA). 没有特别说明的话二者通用.

“人”有机结合的过程中,数据驱动特征将愈加凸显,相关范式转变也将进一步深化.

3 全景式 PAGE 框架

全景式 PAGE 框架是融合大数据特征和重要研究方向的要素矩阵,旨在刻画大数据驱动的“全景式”管理决策框架.全景式 PAGE 框架具有三个要件:大数据问题特征、PAGE 内核、领域情境(如图 5 所示).大数据问题特征涵盖粒度缩放、跨界关联和全局视图,并作为管理决策背景下的特征视角映射到研究内容方向上. PAGE 内核是指四个研究方向,即理论范式(paradigm)、分析技术(analytics)、资源治理(governance)以及使能创新(enabling).领域情境是指针对具体行业/领域(如商务、金融、医疗健康和公共管理等)进行集成升华.

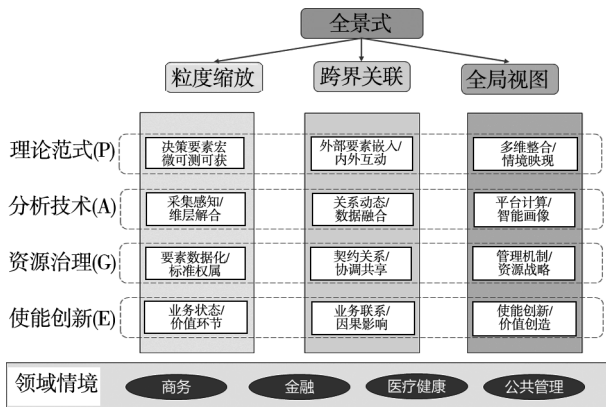


图 5 全景式 PAGE 框架

Fig. 5 Panoramic PAGE framework

围绕 PAGE 内核,在大数据问题特征映射下可以形成一个 4×3 的要素矩阵.在理论范式(P)研究方向上,重点关注管理决策范式转变机理与理论.传统的管理决策正在从以管理流程为主的线性范式逐渐向以数据为中心的新型扁平化互动范式转变,管理决策中各参与方的角色和相关信息流向更趋于多元和交互.概括说来,新型管理决策范式呈现出大数据驱动的全景式特点.进而,由于全景式的多维交互动态性以及全要素参与特点,在研究上需要采用新型的研究范式(即大数据驱动范式).具体说来,在粒度缩放方面,需要决策要素在宏观和微观层面可测可获;在跨界关联方面,需要引入外部要素并形成内外要素互动;

在全局视图方面,需要多维整合并能够针对不同决策环境进行情境映现和评估.

在分析技术(A)研究方向上,重点关注管理决策问题导向的大数据分析方法和支撑技术.在粒度缩放方面,需要数据的感知与采集,并能够在不同维度和层次上进行分解与聚合;在跨界关联方面,需要捕捉数据关系及其动态变化,并能够进行针对多源异构的内外数据融合;在全局视图方面,需要体系构建和平台计算能力,并能够形成各类画像以及开展智能应用.

在资源治理(G)研究方向上,重点关注大数据资源治理机制设计与协同管理.在粒度缩放方面,需要进行资源要素的数据化,并明确数据标准和权属;在跨界关联方面,需要刻画资源流通的契约关系,并形成有效协调共享模式;在全局视图方面,需要建立资源管理机制,并制定组织的资源战略.

在使能创新(E)研究方向上,重点关注大数据使能的价值创造与模式创新.在粒度缩放方面,需要提升业务价值环节的像素,并把握业务状态;在跨界关联方面,需要梳理业务逻辑和联系,并辨识影响业务状态的因果关系;在全局视图方面,需要提升大数据使能创新能力,并促进组织发展与价值创造.

围绕领域情境,可以对 PAGE 相关研究和应用进行凝练、整合和升华.以 NSFC 大数据重大研究计划集成平台构建为例,一般来讲,集成平台由三个部件组成,分别是平台体系、内置部件、整合部件.作为简化示例,对于商务领域集成平台,平台体系由一个商务管理决策相关的数据池,以及相应的数据管理和应用管理平台系统(包括模型、方法、工具库)等组成;内置部件由针对特定行业(如汽车)和特定领域(如营销)的研究成果及示范系统组成;整合部件由商务领域内(不限于内置部件领域)其它相关项目成果在平台体系框架下经过提炼升华汇集而成.对于金融领域集成平台,平台体系由一个金融监测预警服务平台,以及相应的数据管理和应用管理平台系统(包括模型、方法、工具库)等组成;内置部件由针对特定行业(如互联网金融)和特定领域(如征信评估、风险预警等)的研究成果及示范系统组成;整合部件由金融领域内(不限于内置部件领域)其它相关项目成果在平台体系框架下经过提炼升华

汇集而成。

4 结 束 语

面向管理决策研究和应用的大数据驱动范式通过技术增强引入了新视角,进而推动了新型变

量关系、要素机理和理论模型构建,并提升了大数据使能创新的价值创造。这对于应对新型商务形态的进一步机遇和挑战具有重要意义。此外,全景式 PAGE 框架刻画了在粒度缩放、跨界关联和全局视图特征视角映射下的理论范式、分析技术、资源治理、使能创新等重要研究方向。

参 考 文 献:

- [1]冯芷艳,郭迅华,曾大军,等.大数据背景下商务管理研究若干前沿课题[J].管理科学学报,2013,16(1):1-9.
Feng Zhiyan, Guo Xunhua, Zeng Dajun, et al. On the research frontiers of business management in the context of big data [J]. Journal of Management Sciences in China, 2013, 16(1): 1-9. (in Chinese)
- [2]徐宗本,冯芷艳,郭迅华,等.大数据驱动的管理与决策前沿课题[J].管理世界,2014,(11):158-163.
Xu Zongben, Feng Zhiyan, Guo Xunhua, et al. On the research frontiers of big data-driven management and decisions [J]. Management World, 2014, (11): 158-163. (in Chinese)
- [3]陈国青.大数据的管理寓意[J].管理学家,2014,(2):36-41.
Chen Guoqing. The managerial implications of big data[J]. ManaMaga, 2014, (2): 36-41. (in Chinese)
- [4]Buxton B, Goldston D, Doctorow C, et al. Big data: Science in the petabyte era[J]. Nature, 2008, 455(7209): 8-9.
- [5]Dealing with Data. Science Special Online Collection. <http://www.sciencemag.org/site/special/data/>, 2011.
- [6]Adams J U. Genetics: Big hopes for big data[J]. Nature, 2015, 527(7578): S108-S109.
- [7]Athey S. Beyond prediction: Using big data for policy problems[J]. Science, 2017, 355(6324): 483-485.
- [8]Einav L, Levin J. Economics in the age of big data[J]. Science, 2014, 346(6210): 1243089.
- [9]Fregnac Y. Big data and the industrialization of neuroscience: A safe roadmap for understanding the brain? [J]. Science, 2017, 358(6362): 470-477.
- [10]NITRD. The Federal Big Data Research and Development Strategic Plan. <https://www.nitrd.gov/PUBS/bigdatardstrategicplan.pdf>, 2016.
- [11]G20 杭州峰会. G20 数字经济发展与合作倡议. http://www.cac.gov.cn/2016-09/29/c_1119648535.htm, 2016.
G20 Hangzhou Summit. G20 Digital Economy Development and Cooperation Initiative. http://www.cac.gov.cn/2016-09/29/c_1119648535.htm, 2016. (in Chinese)
- [12]Executive Office of the President National Science and Technology Council Committee on Technology. Preparing for the Future of Artificial Intelligence. https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf, 2016.
- [13]National Science and Technology Council. The National Artificial Intelligence Research and Development Strategic Plan. https://www.nitrd.gov/pubs/national_ai_rd_strategic_plan.pdf, 2016.
- [14]MESR. Présentation de la stratégie France I. A. pour le développement des technologies d'intelligence artificielle. <http://www.enseignementsup-recherche.gouv.fr/cid114670/presentation-de-la-strategie-france-i.-a.-pour-le-developpement-des-technologies-d-intelligence-artificielle.html>, 2017.
- [15]Hall D W, Pesenti J. Growing the Artificial Intelligence Industry in the UK. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf, 2017.
- [16]Manyika J, Chui M, Brown B, et al. Big Data: The Next Frontier for Innovation, Competition, and Productivity[R]. McKinsey Global Institute. <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>, 2011
- [17]Henke N, Bughin J, Chui M, et al. The Age of Analytics: Competing in a Data-Driven World[R]. McKinsey Global Institute. <https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/the-age-of-analytics-competing-in-a-data-driven-world>, 2016.

- [18] IBM Big Data and Analytics Hub. Extracting business value from the 4 V's of big data. <http://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data>, 2014.
- [19] Koontz H. The management theory jungle[J]. The Journal of the Academy of Management, 1961, 4(3): 174–188.
- [20] Kuhn T S. The Structure of Scientific Revolution[M]. Chicago: The University of Chicago Press, 1962.
- [21] Hey T, Tansley S, Tolle K. The Fourth Paradigm: Data-Intensive Scientific Discovery[M]. Washington: Microsoft Research, 2009.
- [22] 陈国青, 卫强, 张瑾. 商务智能原理与方法(第2版)[M]. 北京: 电子工业出版社, 2014.
Chen Guoqing, Wei Qiang, Zhang Jin. Business Intelligence: Principles and Methods (the Second Edition)[M]. Beijing: Electronic Industry Press, 2014. (in Chinese)
- [23] Salton G, Wong A, Yang C S. A vector space model for automatic indexing[J]. Communications of the ACM, 1975, 18(11): 613–620.
- [24] Blei D M, Ng A Y, Jordan M I. Latent dirichlet allocation[J]. Journal of Machine Learning Research, 2003, 3(Jan): 993–1022.
- [25] Lowe D G. Object recognition from local scale-invariant features[C]//Computer vision, 1999. The Proceedings of the Seventh IEEE International Conference on Ieee, 1999, 2: 1150–1157.
- [26] Holtz H, Leondes C T. The synthesis of recursive digital filters. Journal of the ACM (JACM), 1966, 13(2): 262–280.
- [27] Laptev I. On space-time interest points[J]. International Journal of Computer Vision, 2005, 64(2–3): 107–123.
- [28] Bengio Y, Ducharme R, Vincent P, et al. A neural probabilistic language model[J]. Journal of Machine Learning Research, 2003, 3(2): 1137–1155.
- [29] LeCun Y, Boser B E, Denker J S, et al. Handwritten digit recognition with a back-propagation network[C]//Advances in Neural Information Processing Systems, 1990, 396–404.
- [30] Sabour S, Frosst N, Hinton G E. Dynamic routing between capsules[C]//Advances in Neural Information Processing Systems, 2017: 3859–3869.
- [31] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533.
- [32] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735–1780.
- [33] Le Cun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436.
- [34] Goodfellow I, Bengio Y, Courville A, et al. Deep Learning[M]. Cambridge: MIT Press, 2016.
- [35] Porter M E. Competitive Advantage: Creating and Sustaining Superior Performance[M]. New York: Free Press, 1985.
- [36] Russell S J, Norvig P. Artificial Intelligence: A Modern Approach[M]. Malaysia: Pearson Education Limited, 2016.
- [37] Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of go without human knowledge[J]. Nature, 2017, 550(7676): 354–359.
- [38] Johnson M, Schuster M, Le Q V, et al. Google's multilingual neural machine translation system: Enabling zero-shot translation[J]. Transactions of the Association for Computational Linguistics, 2017, 5: 339–351.
- [39] Gatys L A, Ecker A S, Bethge M. A neural algorithm of artistic style[J]. Computer Science, 2015, arXiv: 1508.06576.
- [40] Pogue D. The robotic artist problem[J]. Scientific American, 2018, 318(2): 23.
- [41] He J, Zhou M, Jiang L. Generating Chinese Classical Poems with Statistical Machine Translation Models[C]. AAAI, 2012.
- [42] Brynjolfsson E, Mc Afee A. The business of artificial intelligence: What it can and cannot do for your organization[J]. Harvard Business Review, July 2017. <https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence>.
- [43] Aflalo T, Kellis S, Klaes C, et al. Decoding motor imagery from the posterior parietal cortex of a tetraplegic human[J]. Science, 2015, 348(6237): 906–910.
- [44] Yuste R, Goering S, Bay A, et al. Four ethical priorities for neurotechnologies and AI[J]. Nature, 2017, 551(7679): 159–163.
- [45] Davenport T H, Ronanki R. Artificial intelligence for the real world[J]. Harvard Business Review, 2018, 96(1): 108–116.
- [46] Wei Q, Qiao D, Zhang J, et al. Anovel bipartite graph based competitiveness degree analysis from query logs[J]. ACM

Transactions on Knowledge Discovery from Data, 2016, 11(2): 21–25.

[47] Mitchell T, Brynjolfsson E. Track how technology is transforming work[J]. Nature, 2017, 544(7650): 290–291.

[48] Zhu C, Zhu H, Xiong H, et al. Recruitment market trend analysis with sequential latent variable models[C]. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016: 383–392.

[49] Bostrom N. The superintelligent will: Motivation and instrumental rationality in advanced artificial agents[J]. Minds and Machines, 2012, 22(2): 71–85.

The challenges for big data driven research and applications in the context of managerial decision-making: Paradigm shift and research directions

CHEN Guo-qing¹, WU Gang², GU Yuan-dong², LU Ben-jiang¹, WEI Qiang^{1*}

1. School of Economics and Management, Tsinghua University, Beijing 100084, China;

2. Management Science Department, National Natural Science Foundation of China, Beijing 100085, China

Abstract: In the era of digital life and business, traditional management has become or is becoming data-centric management, and traditional decision-making has become or is becoming data-analytics-based decision-making. From the big data characteristics in light of its data nature, problem nature and managerial decision-making nature, this paper elaborates the paradigm shift of research and applications on managerial decision-making. The big data-driven paradigm is described from three aspects: external embedding, technology augmentation, and enabled innovation, to reflect a “correlation + causality” viewpoint in a “data-driven + model-driven” manner. In addition, the panoramic PAGE framework as well as its key elements are discussed with respect to the big data characteristics and important research directions.

Key words: big data; managerial decision-making; research paradigm; panoramic PAGE framework

附注:国家自然科学基金委员会“大数据驱动的管理与决策研究”重大研究计划是一个具有统一目标的项目集群,旨在充分发挥管理、信息、数理、医学等多学科交叉合作研究的优势,以全景式 PAGE 框架作为总体思路框架,坚持“有限目标、稳定支持、集成升华、跨越发展”的原则,围绕学科领域趋势、理论应用特点,注重基础性、前瞻性和交叉性研究创新。自 2015 年底至 2017 年底,此重大研究计划部署了包括培育项目、重点项目和集成项目等一系列项目。其后续的项目部署将在全景式 PAGE 框架下,进一步突出凝练、整合与升华,强调与总体思路框架内容的契合性和贡献度。

本文素材部分来自国家自然科学基金委“大数据驱动的管理与决策研究”重大研究计划相关的系列研讨。由衷感谢不同学科领域专家学者(包括 NSFC 大数据重大研究计划指导专家组、顾问专家组、管理工作组等专家学者)的真知灼见和思想贡献!