

对大数据的经济学之多维思考

顾海兵

(中国人民大学 应用经济学院,北京 100872)

[摘要]大数据的发展很热,然而对大数据的冷思考很少,基于经济学的冷思考检索尚未发现。对大数据的经济学之多维思考包括:(一)对大数据的经济学学科思考,经济学是利益之学,大数据的发展必须考虑个别成本和社会成本,必须考虑每个人的利益,必须考虑多方面的利益关系,必须尊重不同的社会主体等;(二)对大数据的逻辑经济学思考,大数据的大是外延,数据是内涵,大数据对应的是小数据,小数据和大数据共同组成数据;(三)对大数据的产权经济学思考,基于和法律有关的产权理论,大数据及其发展必须以个人或公民数据权保护为基础,个人或公民数据权保护应该和必须从隐私权保护向财产权保护延伸,推动大数据的财产化、市场化;(四)对大数据的生态经济学思考,即用生态的观点分析大数据的发展,针对大数据发展中的反生态、反绿色的现象和事实,应该去除和减轻大数据泡沫、大数据失衡和大数据黑链等。

[关键词]大数据;经济学;多维;学科;逻辑;产权;生态;冷思考

DOI:10.3969/j.issn.1002-1698.2019.06.011

实践之树常青,常青的实践推动理论的发展;理论常常是绿色的,绿色的理论指导实践的发展。理论的发展常常是概念创新先行。大数据(big data)就是一个典型。这个概念一经提出,很快风靡世界,开始美欧,后来全球,不仅理论界,而且实务界,不仅专业圈子,而且百姓圈子,可以说非常热,已经热了不少于10年的历史。

一、对大数据的经济学学科思考

大数据产生于自然科学,主要是数学和计算机科学。由此,对大数据的讨论,社会科学介入比较少,基于经济学的介入就更少。依据百度检

索和知网检索,尚没有发现以“大数据的经济学思考”为题的文章。需要注意的是,本文所探讨的话题是对大数据的经济学思考,而不是大数据时代的经济学思考,不是要去建立大数据经济学之类的学科,不是建立“大数据+某某经济学”,比如大数据产业经济学,大数据消费经济学等。前者是对大数据的思考,是对大数据的反思、再思和深思,后者是对经济学的思考,是对经济学的反思、再思和深思,两者具有质的区别。这里首先探讨对大数据的经济学学科思考。

所谓经济学学科思考,就是对大数据的经济学本质和结构思考。经济学本质思考是基于经

济学的定义或学科特征对大数据及其发展的思考,经济学结构思考是基于经济学的分支或子学科特征对大数据及其发展的思考。显然本质思考是基础,结构思考是延伸。此处不讨论经济学是否是科学,至少它是学科,是社会科学体系里的重要组成部分。有人把经济学叫做“经济学帝国主义”。就国内而言,在高等学校教师结构里,经济学类或广义经济学类(包括商科)的教师高达1成以上,没有哪个普通高校没有经济(管理)类学科。依据学科划分,100多个一级学科里,理工农医类大体占2/3,包括经济学在内的广义文科占1/3,有40个左右,其中经济学一级学科有2个,分别是理论经济学和应用经济学。理论经济学一级学科下面的二级学科有:政治经济学,西方经济学,世界经济,经济史,经济思想史,人口、资源与环境经济学等;应用经济学一级学科下面的二级学科有:国民经济学,财政学,金融学,国际贸易学,产业经济学,区域经济学,数量经济学,劳动经济学,国防经济学等。显然,经济学结构思考可以基于一级学科的经济学的每个分支对大数据及其发展作思考,也可以基于二级学科的经济学的每个分支对大数据及其发展作思考。关于这一点本文不做展开,下面主要作经济学本质思考,即基于经济学定义对大数据的思考。

关于经济学的定义,国内外解释很多,然而缺乏逻辑分析。从逻辑上说,经济学的定义有两种思路:一个思路是从经济的定义出发,由内而外解释经济学;一个思路是从科学或社会科学出发,由外而内解释经济学,把经济学看作社会科学的分支。一般的流行的定义没有注意这两种思路的区别,多是直接给出定义。当然,也有学者用专著探讨,比如学者梁小民2001年出版了《经济学是什么》的专著,^[1]笔者也曾经在1998年、2002年和2005年出版的专著里用专文探讨过。^[2]这里简单归纳一下。

基于第1种思路,对经济的定义至少有8种解释:经世济民,经邦济世;效率(经)加公平

(济);生产和生活中的节省、节约;生产关系的总和,生产过程中所形成的社会关系,包括生产资料所有制和分配关系等;生产力,人类创造财富的能力;社会再生产过程,包括生产、分配、流通和消费;致富或发财的活动;主体通过供给实现增值的活动。

再来看看百度百科的解释:经济是价值的创造、转化与实现;人类经济活动就是创造、转化、实现价值,满足人类物质文化生活需要的活动。然后在经济的定义的基础上,加上知识体系,就可以得到经济学的定义。比如,经济学就是有关经世济民的知识体系,就是效率和公平的知识体系,等等。

基于第2种思路,流行的经济学的定义,国内外主要有:经济学是研究财富的学问;经济学是研究如何配置稀缺的资源以生产有价值的商品并分配给不同的个人;经济学是研究社会如何管理自己的稀缺资源;经济学就是价格学;经济学就是交换的社会科学;经济学就是最有利的实现目标的方法;经济学就是有关稀缺资源的社会科学;经济学研究产生于欲望导致稀缺的选择问题或资源配置问题。哈佛大学经济系在其网站上特别指出,经济学不是商学,经济学不是教个人如何发财,教个人如何发财是商学的任务。

多年以来,笔者对经济学的解释一直坚持从经济出发,笔者对经济的定义是:主要以个人物质利益为动力(以个人物质利益最大化为目的)的社会系统。简单地说,经济就是个人利益的社会系统,或个人利益社会场。笔者的这一解释在逻辑上考虑了政治的定义,笔者所理解的政治定义是:政治就是主要以群体利益为动力的社会系统。社会系统包括:经济,政治,宗教,文化,等等;社会问题包括:情、义、利、德、趣等,对应的学科主要是:文学、法学、经济学、伦理学、艺术学等。至于资源稀缺,那是个人利益派生的结果,是第二位的。注意,这里的个人利益不是某个人的利益,是所有人的个人利益。

这里不对经济学的各种定义作评价,既无必

要也无可能,笔者主张百花齐放百家争鸣、求同存异多元共生,可以取最大公因子。古人所云已经道出了经济学的本质:天下熙熙皆为利来,天下攘攘皆为利往;无利不起早。由此,对社会问题的经济学思考,或经济学学科思考,其实就是从利益的角度思考,从个人利益的角度思考,从社会的个人利益的角度思考。基于这样的认识,对大数据或大数据发展的经济学学科思考,就是要基于利益去思考,社会的个人利益去思考,或社会的经济人思考。由此,大数据的发展,必须考虑个别成本和社会成本,必须考虑每个人的利益,必须考虑个人与个人、个人与企业、个人与群体、个人与政府、个人与非政府组织的利益关系,必须考虑中央与地方、地方与地方的利益关系,必须考虑短期利益和长期利益的关系,必须尊重不同的社会主体(大数据的参与者和未参与者,企业、政府和非政府组织,不同年龄、不同性别、不同民族、不同宗教、不同收入、不同境况的人),必须处理好市场与政府的关系,政府不能缺位、越位和错位,企业不能限制竞争不能垄断,必须寻求经济效益、生态效益和社会效益的最佳结合,必须寻求全社会的帕累托最优,或者通俗地说,在信息化、信息数据化浪潮中,不让一个人掉队,不歧视任何人,包括大数据的反对者、拒绝经济数字化的人。显然,对大数据这样的经济学学科思考,对政府提出了更高的要求。

二、对大数据的逻辑经济学(经济哲学)思考

如果说前面是对大数据的经济学之本质的思考,这里就是对大数据的经济学之哲学思考,是基于经济哲学的思考,是逻辑经济学的思考,是深入一层的逻辑或解构思考。要回答大数据的内涵和外延究竟是什么,不是对大数据的简单定义。虽然,类似话题已经有很多解释,自然科学的解释和社会科学的解释,国外的解释和国内的解释,然而,依旧不断有需要新解释的必要。本文的解释是基于哲学,超越自然科学,超越社会科学,同时又结合经济学知识,结合经济底色,

理论和实践结合,抽象和具体结合。

大数据是什么?也许正如数学是什么一样,永远没有统一的答案,永远没有最终的答案。百度百科上的解释大体接受度高一些(接近于维基百科的解释):大数据(big data, mega data),是无法或很难在一定时间内用常规软件进行捕捉、管理和处理的数据集合,是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产。大数据拥有5V特点:大量(Volume)、高速(Velocity)、多样(Variety)、价值(Value)、真实(Veracity)。大数据带来三个范式性的观念转变:是全部数据,不是样本数据;是总体把控,不求个体精确;是相关关系,不必因果关系。当然也有人因为“大数据”定义之困难,主张用“数据科学”或“数据分析”等概念替代“大数据”,甚至认为定义大数据没有必要也没有现实性。^[3]不过,笔者不这样认为,但关于大数据的定义,本文点到为止,下面对大数据这一概念作逻辑分析。

何谓大数据的逻辑分析?实际就是对“大数据”这一概念的解构分析,具体说来就是分析大数据的内涵和外延。这样的解构是大数据发展的逻辑起点,是大数据发展的理性思考,其实就是冷思考。关于这个问题,因为属于“小问题”,大数据这三个字又非常直白,不需要咬文嚼字,似乎属于不言自明,加之中国传统“名正言顺”的研究不重视逻辑,因此绝大多数人自然不去关注。知网上检索“大数据的内涵外延”“大数据的逻辑分析”,均没有相应文章。个别关注“大数据”概念本身逻辑的研究者,实际上最后也没有作落地分析。^[4]其实,国内很多新概念的发展,很多的争论,往往由于前提的不统一、不明确,导致太多的交流是“鸡同鸭讲”,是无序论战,是混战。下面作具体分析。

首先分析大数据的外延。从逻辑的外延上说,与大数据对应的是小数据(small data),小数据和大数据共同组成数据。没有离开小数据的大数据,也没有离开大数据的小数据。数据的大

与小,在一定意义上是相对的,换一个维度,小可能是大,大又可能是小。比如人体数据相对于地球数据是小,相对于细胞数据是大,相对于分子、原子数据是超大,细胞数据相对于原子数据是大。大中有小,小中有大,不必厚此薄彼,不必厚今薄古,也不必望文生义地把小数据就等同于个体资料或微观数据,把大数据等同于宏观数据,甚至粗暴地简化说:大数据和别人的生意有关,小数据却仅与你自己有关。有时我们需要的数据越多越好,有时我们需要的数据并非越多越好。在传统数据的基础上,增加大数据是深化和优化。

从逻辑上说,大数据的“大”是外延,“数据”是内涵。基于逻辑经济学之经济,笔者以经济生活为背景,选择更具有形象性和通俗性的城市燃气集团公司为例来说明大数据之外延,并由此提出基于大数据外延的若干数据概念:(1)源数据,也就是燃气从哪里来,燃气的源头数据,可供气源数据,入网燃气数据,燃气集团公司购买或自产的燃气数据;(2)宿数据,也就是燃气最终用到哪里去,被谁消耗和消费了,即用户的用气数据。比如工业用气,发电用气,供热用气,居民用气等。(3)流数据或中数据,也就是燃气公司下辖分公司或二级机构的流入流出的燃气数据。(4)亲数据,即燃气集团子公司的数据,因为燃气公司作为集团,不仅有分公司的燃气数据,还有很多和燃气有关的独立的子公司的数据。比如,清洁燃料有限公司,燃气能源开发有限公司,燃气用户服务有限公司,燃气灶具开发公司,燃气设备制造公司等。(5)上数据,也就是控制燃气集团的大股东或母公司的数据,比如燃气集团的大股东常常是国资授权的城市投资控股集团,即上面的数据。(6)邻数据,即国内同行的数据,比如其他城市燃气集团的数据等。(7)宏数据,即国家的宏观经济数据,国内生产总值,外贸进出口,就业,居民收入与支出等;(8)外数据,即国外的有关数据,可以是燃气的数据,也可以是与燃气有关的数据。显然,大数据是构建在也

是大数据的各类小数据的基础之上的,需要基于实际的顶层设计、立体设计、有机设计,尤其政府的大数据必须先导而共享,之后才是大数据的“大处理”。

如果大数据的大是外延,则大数据的内涵或内容就是数据。外延是数据的结构观察,更多的是和数据供给有关;内涵是数据的性质观察,更多的是和数据需求有关。这里还是结合城市燃气集团公司来案例式说明大数据之内涵,并提出基于大数据内涵的若干数据概念:(1)实数据,即数据是否真实,是否全面而无死角。比如,燃气集团的燃气数据,首先要保证原始数据的真实,其次要包括全社会,那些不是由燃气集团供给的燃气也要包括进来。(2)主数据,即数据是有主次之分的,不是眉毛胡子一把抓,燃气集团的数据当然以燃气量为主数据。而辅数据就是围绕燃气的各种经营数据,费用、成本、利润和收入等。(3)负数据,即负面数据,就燃气集团来说,各种偷盗燃气的数据,由于设施、技术等引起的跑冒滴漏燃气数据,由于人为抄错的燃气数据等。(4)险数据,燃气集团和安全密切相关,所有与风险、安全相关的节点都是数据点,包括自然因素、人为因素,甚至战争因素,识别和预防风险显然是大数据的用武之地。(5)稳数据,燃气集团除了保证安全供气,还要保证供应的稳定性,为此需要数据的支撑,尤其大数据的支撑。(6)央数据,这里的央数据之央是指燃气集团所服务的首脑机构,是社会的指挥机关,在北京就是所服务的中央机关,包括中央军事机关,在省城就是所服务的省级机关,以此类推。显然,这是敏感数据,需要特别关注,要有数据支撑的应急预案。(7)潜数据,即和未来有关的数据,近期未来、中期未来和远期未来,包括潜在市场、潜在竞争等数据。(8)净数据,即和环保有关的数据,包括甲烷等污染物的排放数据、减排数据,现状和目标数据等,需要建立燃气的环境质量监测与预警的大数据系统。不言而喻,上述的企业大数据之内涵分析可以类推到政府大数据。

三、对大数据的产权经济学思考

产权这一概念天然地属于经济学范畴,产权的本质是所有权或财产所有权,基于个体利益的差异而产生,与政治上的公民权利对应,古语或俗话的“无恒产者无恒心”,也是产权的一种角度的表述,由此产权理论或产权学,就是产权经济学,产权经济学可以简称为产权学或产权理论(由于产权和制度的联系,也可以叫制度经济学或新制度经济学)。当然,产权经济学的内容在不同的时代是不同的,不同时代的经济学家从不同层面或角度研究产权问题。作为一个相对独立的经济学分支,并形成比较完整的理论体系,产权经济学,是现代的产物,也叫现代产权经济学(甚至称为交易费用经济学或交易成本经济学),是与科斯(Ronald H. Coase, 1910—2013年,美国经济学家,1991年诺贝尔经济学奖获得者)、诺斯(Douglass C. North, 1920—2015年,美国经济学家,1993年诺贝尔经济学奖获得者)等人联系在一起的。笔者在此不讨论产权经济学的内容和争议,只是借助产权这一概念对大数据及其发展作思考。

从产权角度讨论大数据,经济学界介入不多(也许是由于大数据的无形或虚体,也许是认为实体财产更重要,也许是认为交易费用更重要),更多的是法律界或法学界,法律界或法学界更多的是从保护隐私或公民权利的角度作分析。2019年3月4日《新京报》发表了《以明确大数据法律属性遏制信息泄露》的文章。^[5]显然,法学界对大数据的讨论,不是简单的法学如何借力大数据而发展,而是法学如何推动大数据的发展,如何降低大数据发展的成本。因此,从产权经济学角度,或者叫产权学角度,或者叫“产+权”角度,讨论大数据的发展,应该借鉴法学的做法,甚至借力法学的类似研究。

大数据及其发展必须以数据权的界定为前提,数据权必须以数据财产权为抓手。依据龚子秋等的文章,数据权的内涵国内外都没有统一的

定义,也许在可见的未来都没有统一的定义。^[6]笔者以为,没有统一的定义不等于没有界定,不等于不要确定,可以就最大公约数或最低标准作界定,可以就不同类型的数据权作不同的界定,从有利于解决具体问题入手。目前学者把数据权分为(国家)数据主权、数据管理权、数据公民权(公民数据权)、数据社会权、数据人格权、数据财产权、数据被遗忘权等。在这些权利中,国家数据主权是前提,公民数据权是基础,数据财产权是抓手。邓刚宏等认为我国对数据权的研究,学界主流以财产权和知识产权为逻辑起点,无法满足我国数据产业的发展需要。^[7]笔者不这样看,笔者赞同主流的看法,在市场经济环境里,只要努力界定清楚数据产权,只要有平等的供求双方,只要有价格的自由,不必作茧自缚,不必人设问题,不必杞人忧天,市场这只看不见的手就会发挥神奇的作用,数据产业的发展就不会有障碍,也就是说基于产权经济学的原理,推动大数据的财产化、市场化,即使短期速度慢、短期有点乱,但长期看速度不慢、质量更高。

大数据及其发展必须以个人或公民数据权保护为基础,个人或公民数据权保护应该和必须从隐私权保护向财产权保护延伸。关于这一点,我们首先来看大数据发展的三个阶段。基于高莉、陈颖等学者的分析,大数据的发展由技术革命、经济规律和社会变革所决定,它至少已经经历了三个阶段(也有划分为四个阶段或六个阶段等)。^[8]第一阶段是大数据的萌芽期或产生期,大体是20世纪最后的十年,大数据是少数企业或大企业的“专利”,个别高新技术企业专注于大数据,这段时期的大数据发展主要是针对单个的具体问题,开发数据挖掘技术,在短时间内高效率处理数据。微软公司的操作系统以及在线服务网络平台等可以看作是大数据发展的前夜。第二阶段是大数据的发展期或突破期,大体是21世纪的第一个十年,不少企业开始涉足大数据,大数据被用于分析和解决越来越多的问题,信息数字化,图像视频数据化,社交网络流行,海

量数据的处理,2005年存储和分布式运算大数据的Hadoop软件工具诞生,美国公司开发了大数据模型,1998年设立、2004年在纳斯达克上市的谷歌(Google)作为大数据企业日显突出,政府开始介入大数据,大数据市场初步形成。第三阶段是大数据的稳态期或成熟期,大体是最近十年,大数据从技术创新向所有创新延伸,多数企业都直接或间接地和大数据发生联系,各大IT巨头[惠普、微软、IBM、EMC(易安信)等]纷纷收购大数据企业推出大数据产品,大数据越来越多地介入到各种经济社会活动中,从国内商业、对外贸易、银行、保险、证券、投资到旅游、交通、医疗、养生、教育、体育、新能源、新材料、航空航天、生物工程、刑事侦查、防范恐怖、军事国防、基础设施安全、人工智能等都开始了大数据的运用,大数据的可视化,大数据的安全研究,政府更加主动地介入大数据的管理,制定大数据的发展战略,大数据市场开始规范化和规模化,数字经济高速增长,中国后来居上,比如阿里巴巴在国内最早设立首席数据官(2012年),中关村成立大数据产业联盟,北京大学联合其他机构成立北京大数据研究院等。不言而喻,伴随产业数字化、数字产业化、大数据产业化和市场化,个人或公民数据权保护自然地隐私权保护向财产权保护延伸,也只有通过财产权的保护可以更好地保护隐私权。公安部前副部长陈智敏等认为,数字经济时代数据属于私有,相当于农业时代的土地、工业时代的资本,是重要的资产。^[9]关注数据权益的浙江垦丁律师事务所创始人也认为,当下个人信息权利多是关注其人格属性,重视与个人隐私权的联系,但大数据时代数据市场的发展,以及非法数据链的增长,个人信息的财产属性应该受到重视。^[10]需要补充的是,笔者不赞成把大数据发展的历史任意扩大到所谓18世纪的观点,或者微软等现代操作系统没有诞生的20世纪80年代以前,或者只是某个未来学者偶然地、非自觉地、科幻式地提出大数据某个单一概念的1980年。^[11]

需要说明的是,支付宝和微信支付等其实就是大数据的金融舞台,其无现金化给亿万消费者带来了巨大便利。然而这样的支付工具,为什么在支付宝和微信支付所用技术所依赖的欧美日等发达国家和地区似乎发展不快?这当然不是由于欧美人不创新、欧美人比较笨,这一定和大数据的产权有关,和公民的数据隐私权有关,和经济的反垄断有关,和大数据产业发展的敏感性有关,和公民的消费权利保护有关,和金融的风险控制有关。从这一点出发,大数据的发展,类似支付宝和微信支付的发展,其评价,包括速度评价、利弊评价,可能需要一定的时间。当然,世界上的事情,有些不具有可比性。

四、对大数据的生态经济学思考

生态经济学是生态学与经济学的结合,可以是经济学里增加生态元素,也可以是生态学向经济的延伸。如果把大数据发展看作大数据产业发展,或大数据经济发展,则对大数据的生态经济学思考就是用生态观点观察分析大数据的发展,就是大数据产业的生态发展,就是数据密集型产业的生态发展,就是大数据经济的生态发展,就是大数据经济的可持续发展。显然,大数据越发展,大数据产业越发展,其生态问题或可持续发展问题越重要,生态经济学的分析越重要。依据《南方周末》记者汪徐秋林报道,2018年中国大数据市场产值将达6200亿元,大数据核心产业、关联产业、融合产业等有20个产业。^[12]大数据的收集、分析与利用逐渐成为企业运营的核心。关于这一话题,此处从逆向角度分析,即结合大数据发展中的反生态、反绿色的现象和事实进行具体分析。

首先,当然是大数据的泡沫,或大数据的炒作。目前,大数据这一概念几乎满天飞,什么都是大数据,什么都要扯上大数据,大数据成为万金油,大数据成为万能武器。企业家言必讲大数据企业,官员言必讲大数据政府,校长言必讲大数据学校,医院院长言必讲大数据医院,媒体的

报道更是推波助澜、铺天盖地;对某个论著或报告做点文献统计或文献计量分析,比如统计“创新”“改革”出现了多少频次,就称呼其是大数据分析。有些大学的学院也纷纷改名,比如软件学院改为大数据与软件学院,或计算机学院改为大数据与计算机学院;房地产开发更是偏爱大数据,把大数据作为噱头,动辄要搞大数据城、大数据镇、大数据村、大数据谷、大数据社区、大数据新城等;有些机构有条件没条件都设立大数据研究中心,可谓大数据中心林立;有些和大数据有关联的机构,有意无意夸大大数据的规模,比如中国经济中大数据经济或数字经济的比重已经达到3成以上。显然,大数据变成一个筐,什么都往里装,这是草率分析、欲速不达、难以取信的。

其次,就是大数据的失衡,或大数据的孤岛。大数据本质在于数据之间的有机联系,大数据之间的生态连接,然而现实却是数据之间的分割,甚至是严重分割。突出的是不同区域政府之间大数据的失衡,每个政府都是大数据孤岛,^[13]此政府(省、市、县)和彼政府(省、市、县)大数据之间的失联、失衡,政府内部此部门大数据和彼部门大数据之间的失联、失衡,每个部门都是一个个大数据孤岛,政府大数据和非政府大数据之间的失联、失衡,此行业大数据与彼行业大数据之间的失联、失衡,平台大数据和企业大数据之间的失联、失衡,上游企业大数据和下游企业大数据之间的失联、失衡,生产、分配、流通和消费的大数据之间的失联、失衡,中国大数据和海外大数据的失联、失衡等。政府有数据,有很多数据,然而空有其大,是虚大,不是真正的大数据。笔者曾经参与调研,某设区的大城市,各区的信息化建设没有统一规划,基本都是仅仅服务于单个部门,都是“自给自足”,一个区有100多个部门,每个部门都有自己的信息系统,其中一半多属于市垂直部门信息系统,20多个属于部门自建系统,20多个属于部门购买或委托第三方建设系统,各部门信息系统都是孤立的,一个个数据孤

岛,没有兼容性。反观经济总量和人均经济水平都低于我们的印度,在解决大数据失联、失衡方面未必落后。比如,2009年印度政府就启动了全球第一大生物识别数据库项目 Aadhaar,超过印度95%的人口,大约11亿人的姓名、地址、手机号,以及指纹、照片、虹膜等个人信息,都被作为公民个人的唯一识别码(UIDAI),纳入印度政府的中央数据库系统。虽然有所谓隐私争议,但利大于弊。目前,Aadhaar大数据已经深入到印度公民生活的各个方面,UIDAI直接连接到个人的银行账户、手机号码、保险单据、PAN(永久账号)、纳税情况以及其他多种服务上。^[14]可以这样说,一个印度人无论是学校读书、医院就医、寻找工作,还是在银行办理业务、申请政府项目等等,都首先要注册 Aadhaar,由于基于生态有序规则,印度社会的大数据水平(大数据基础设施)甚至超过了不少发达国家。

再次,就是大数据的黑化,或者黑色的大数据。这部分内容非常广泛,包括大数据的污染或大数据的垃圾,比如虚数据、假数据、伪数据、毒数据(带有病毒的软件数据);大数据的过量,比如数出多门,数据重复交叉;大数据的滥用,大数据的灰链或大数据的黑市,非法买卖数据,窃取数据,依据《南方周末》记者汪徐秋林报道,推动贵阳大数据交易所成立,并参与起草了不少大数据产业规划标准的贵阳大数据交易所执行总裁王叁寿估算认为,“我们这里交易1个亿,黑市交易99个亿”。^[15]有估计认为网络数据犯罪的规模达数万亿美元。

面对大数据发展中的反绿色、反生态,包括前面所说的大数据产权等问题,笔者认为责任的主要方面在政府。通俗地说,“老鼠”怎么样,关键还是看“猫”怎么样。当下,有些地方政府设立大数据管理局之类的机构(还有所谓的“云长制”^[16])。比如S省大数据局,2018年11月初挂牌,然而至少到2019年3月底没有检索到其官方网站;G省设立了政务服务数据管理局,也没有找到其官方网站;Z省设立了大数据发展管理

局,检索后只有“相关信息正在制定中”。显然,这只“猫”本身有点先天不足,本身的“大数据味道”不浓。当然,这里不就事论事,不对政府是否应该设立大数据管理机构作结论,不对大数据管理机构是否能够发挥作用作评价,只是依据历史和理性,从逻辑上指出:设立独立的政府机构来解决问题,似乎很有“效率”,但必然涉及与原本就偏多的各政府机构的关系,政府的部门关系就会更加复杂(比如,两家机构只有一对关系,三家机构立即增加到三对关系,四家机构增加到六对关系,等等),也许最后是增加内耗、增加协调成本,结果就是设立新机构不仅不能解决问题,反而成为问题本身的一部分。政府不能求其“大”,不能在机构数量上求其多,欲速反而会不达,须三思而后行,大胆设想、小心求证,尊重生态规律和管理规律。

基于前面的分析,对大数据的经济学思考,对大数据的经济学之学科、逻辑、产权和生态的多维度反思和深思,大数据是大又不仅是大,大数据是数又不仅是数,大数据不等于大数学,不等于大数字,不等于大数量,不等于大数值,大数据是细,是准,是真,是深,是精,是广,是全,是质,是值。依据易宪容等引用的资料以及维基百科的大数据词条,由于全球物联网的兴起,2020年每人每秒将产生数据1.7MB(百万字节),全球数据总量达到4万EB(百亿亿字节)。^[17]如此海量式天文数字,大数据的发展更加需要避免“一哄而上”与“炒作”,更加需要处理好大数据所涉及的各种关系,更加需要扎扎实实地发展我国的大数据事业,特别是天然地拥有大数据的政府之大数据事业。

最后需要说明的是,大数据一词虽然来自英文(big data),但中文“大数据”的含义反而超越英文。笔者依据作为汉语细胞的汉字特征斗胆认为,大数据=大+数+据,或大数据=大×数×据。这里,“大”是工具、形和外延,“数”是途径、量和介质,“据”是目标、质和内涵。也许,大

数据事业和大数据产业的未来之发展,在汉字中国因更加充分地满足每个人的利益而出彩!

注释:

[1]梁小民:《经济学是什么》,北京:北京大学出版社,2001年,第3-5页。

[2]顾海兵:《经济系统分析》,北京:北京出版社,1998年,第6-8页;顾海兵:《非主流经济学研究》,天津:天津人民出版社,2002年,第13-17页;顾海兵:《新范式宏观经济学》,北京:中国财政经济出版社,2005年,第88-101页。

[3]高莉:《大数据伦理与权利语境——美国数据保护论争的启示》,《江海学刊》2018年第6期。

[4]徐立水、辛敏:《大数据及其应用前景研究》,《企业科技与发展》2016年第6期;贾云翔:《大数据研究综述》,《经济师》2018年第12期;陈颖:《大数据发展历程综述》,《当代经济》2015年第8期。

[5]周汉民:《以明确大数据法律属性遏制信息泄露》,《新京报》2019年3月4日。

[6]龚子秋:《公民“数据权”:一项新兴的基本人权》,《江海学刊》2018年第6期。

[7]邓刚宏:《大数据权利属性的法律逻辑分析——兼论个人数据权的保护路径》,《江海学刊》2018年第6期。

[8]高莉:《大数据伦理与权利语境——美国数据保护论争的启示》,《江海学刊》2018年第6期;陈颖:《大数据发展历程综述》,《当代经济》2015年第8期。

[9]陈智敏:《数字经济时代数据属于私有》,《新京报》2019年3月7日;尹浩:《用户数据不应随意挖掘使用》,《新京报》2019年3月7日。

[10]张问之:《保护还是发展——争议个人信息保护法》,《南方周末》2019年3月14日。

[11]中国企业评价协会、中色金银贸易中心编:《中国电子商务企业发展报告2013》,北京:中国发展出版社,2013年,第268-270页。

[12]汪徐秋林:《谁拥有数据,谁得到利益》,《南方周末》2019年3月14日。

[13][16]徐冰:《“云长制”来了,打破部门间数据孤岛还远吗》,《新京报》2019年3月29日。

[14]王会聪:《印度“查税神器”盯上了网络炫富》,《环球时报》2019年4月2日。

[15]汪徐秋林:《谁拥有数据,谁得到利益》,《南方周末》2019年3月14日;西坡:《人社大数据不应用于“制约”频繁跳槽者》,《新京报》2019年4月4日。

[17]易宪容、陈颖颖等:《开放银行:理论实质及其颠覆性影响》,《江海学刊》2019年第2期。

〔责任编辑:邹秋淑〕