

确权、交易、资产化:对大数据转为生产要素 基础理论问题的再思考

赵瑞琴^{1,2}, 孙鹏³

(1. 首都经济贸易大学经济学院, 北京 100070; 2. 河北农业大学经济管理学院, 河北保定 071001;
3. 北京市通州区金融服务办公室, 北京 101100)

摘要: 大数据是数字经济的基础性、战略性资源, 是后疫情时代经济发展的重要生产要素, 但社会各界对大数据的理论认知仍落后于其应用实践, 而且大数据确权、交易定价、资产化问题也存在诸多争议, 这对大数据产业和数字经济的持续健康发展形成潜在隐患, 也成为制约大数据向生产要素正常转化并参与社会大生产的关键。文章在对大数据概念、属性重新认识的基础上, 综合财经、法律等理论观点和实操技巧, 阐述了大数据如何从商品流通要素演变为社会生产要素的市场逻辑以及使用价值如何注入大数据的资产化过程, 并从实践视角提出大数据确权的法理基础和大数据交易所生态下交易定价模型。同时提出以数据税来弥补大数据采集环节个人和企业放弃的小微权利, 将私权转化为社会公共产品, 为大数据产业和要素市场发展提供新的借鉴。

关键词: 大数据; 大数据确权; 大数据交易定价; 大数据资产化; 数据税

中图分类号: F710 **文献标志码:** A **文章编号:** 1000-2154(2021)01-0016-11

DOI: 10.14134/j.cnki.cn33-1336/f.2021.01.002

Right Confirmation, Transaction Price and Capitalization Process: Theory Cogitation of Big-Data Transformation into Production Factors

ZHAO Ruiqin^{1,2}, SUN Peng³

(1. Capital University of Economics and Business, Beijing 100070, China;

2. College of Economics and Management, Hebei Agricultural University, Baoding 071001, China;

3. Tongzhou District Financial Service Office, Beijing 101100, China)

Abstract: Big data is the basis and strategic resource of digital economy, and is an important factor of economic development in the post epidemic era. However, the theoretical cognition of big data from all walks of life still lags behind its application practice, and there are many controversies about big data right confirmation, transaction pricing, and capitalization, which form potential risks for the sustainable and healthy development of big data industry and digital economy, and also become a constraint for the development of big data industry. Data is the key to the normal transformation of production factors and participation in social production. On the basis of a new understanding of the concept and attribute of big data, this paper expounds the market logic of how big data evolves from commodity circulation factors to social production factors and the capitalization process of how the use value is injected into big data, and puts forward the legal basis of big data right confirmation and the strategy of big data exchange ecology from a practical perspective. At the same time, the transaction pricing model proposes to use data tax to make up for the small and micro

收稿日期: 2020-10-12

基金项目: 河北省软科学项目“双螺旋结构的数字农业研究——乡村振兴新动能”(20557652D)

作者简介: 赵瑞琴, 女, 博士研究生, 副教授, 主要从事产业经济研究; 孙鹏, 男, 经济学博士, 主要从事流通经济研究。

rights abandoned by individuals and enterprises in the big data collection link, and transform private rights into social public products, which provides new reference for the development of big data industry and factor markets.

Key words: big data; right confirmation; transaction price; capitalization process; data tax

一、引言

突发新冠疫情不仅重创全球经济,而且改变了人们营商、社交、学习、娱乐等行为习惯。疫情期间,远程办公、在线教育、直播带货等“云上经济”发展迅速,这种非接触式服务新模式对经济发展和政府治理的影响是长期深远的。数据是数字经济的基础性、战略性资源。后疫情时代,以大数据应用为代表的数字经济将迎来重大发展机遇。习近平总书记在2020年中国国际服务贸易交易会全球服务贸易峰会上提出,支持北京市设立以科技创新、服务业开放、数字经济为主要特征的自由贸易试验区,数字经济正式纳入国家战略。时至今日,大数据采集、处理、分析、挖掘等技术已在商业、金融、制造、研发、政务服务、公共安全等领域得到广泛应用,大数据理念已深入社会实践的各个角落,甚至成为企业形成市场支配地位的重要因素。但目前社会各界对大数据的理论认知仍滞后于其应用实践^[1],而且大数据确权、交易定价、资产化问题也存在诸多争议,这对大数据产业和数字经济持续健康发展形成潜在隐患^[2],也成为制约大数据向生产要素正常转化并参与社会大生产的关键。随着人工智能、物联网、云计算、第五代移动通信等技术加持,大数据应用及大数据产业发展进入爆发期已势不可挡。但我国市场经济实践已经证明,爆发式增长的市场不一定是可持续发展的市场,在大数据和关联产业快速扩张的同时,一些基础性的理论问题亟待重新审视。例如,大数据的属性和概念是什么?大数据从哪里来,属于谁?只有明确这些基本问题,才能深入探讨大数据确权、交易、资产化等理论问题,才有可能构建起适合我国国情的有中国特色的大数据理论,为大数据产业和要素市场建设保驾护航,为数字经济赋予强大的发展动能。

二、对大数据定义、属性和资产化的重新认识

(一) 大数据定义的不稳定影响确权认知

大数据概念产生于二十世纪九十年代,至今已衍生出多种定义。工业和信息化部将大数据定义为“容量大、类型多、存取速度快、应用价值高为主要特征的数据集合”。^①研究机构 Gartner 认为“大数据是指采用新的处理方式而形成海量、高增长率和多样化的信息资产”^[3]。维基百科指出,“大数据是无法在容许的时间内用常规软件工具对数据进行抓取、处理和管理的集合”^[4]。总之,大数据的各种界定均包涵数据、产品、资产、技术这四大要素,在应用环节并无不妥,但在探讨大数据确权时易导致权属主体不明,并倾向于将数据产品的提供者、开发者认定为唯一权利主体。本文认为,需要回归大数据本源,给予其更加“纯净”的定义。大数据是以特定目的积累起来的海量的运动痕迹信息和物质的数字记录,包括“动和静”的结合。“动”态数据就是各种人、物等运动主体发生的与其本体生产生活相对应的行为信息,被以数据信息形式记录下来所形成的运动痕迹,比如消费者网购、打车等信息。“静”态数据是相对静止的物体、客观环境等被以数字化手段转化或记录而成的数据信息,比如北斗导航系统记录的地理信息、以 CAD 图形式储存的建筑信息等。

(二) 大数据属于高阶化、可交易的信息流活动

大数据本质上是数据信息。在商品经济环境中,信息早已不再是静态的信息,而是从商流与物流中分离出来并跟随商品交换流动,成为现代贸易的必备要素。近年来,随着信息技术的发展,大数据产业表现出

^①工业和信息化部《中国大数据产业发展评估报告(2018年)》第17页。

巨大前景。诸多流通企业对运营中产生的海量数据信息进行开发,通过云计算、射频识别、无线传感等技术手段发展“互联网+物流”模式或开发物联网功能,将大数据应用下沉到供应链和商业运营中,以降低流通成本和资源占用。例如,京东与小型实体店开展数据共享、共同预测,实现按需生产和仓库共享,使得原本每家实体店都要各自准备仓储设施,变成多家实体店共用京东一家的仓储系统,实体店转型为前端的销售信息平台,极大降低了小型实体店的运营成本以及存货资金占用成本。这是信息流应用的初级阶段,当信息流成为大数据以后,则进入更为高级的阶段。此时的信息流已经不再是商品流通的辅助要素,而是作为可估值的数据资产直接参与商品交换活动,对商品流通环节以及相关产业、领域在量和质上提供新的驱动。目前,国内已经形成了一批 TalkingData、天眼查等大数据企业,阿里巴巴、腾讯利用支付宝、微信等社交和支付工具更是沉淀了巨量的数据资源成为数据寡头,上海、贵阳等地成立了大数据交易所,市场上已出现数据开放平台、数据软件开发、数据咨询、数据征信评价、数据营销和大数据交易所等多样化商业模式。北京市也在2020年服贸会期间发布了《北京国际大数据交易所设立工作实施方案》,但大数据产品交易尚缺少法律法规和行业标准,不同模式下对大数据确权、定价等的认知也并不相同,亟须进行规范。

(三) 大数据资产化是价值的形成过程

大数据本质上是个体或物体的信息符号,要成为具有使用价值的、可估值的资产还要经过一系列经济操作,这一过程与企业经营密不可分。按照会计上对资产的定义,资产是企业过去的交易或事项形成的,由企业拥有或控制的,预期会给企业带来经济利益的资源。资产在企业中只有保持着运动的状态,从一个形态转化成另一个形态,最终形成或融入某一产品的价值中,才能通过交换实现价值。例如,企业用现金购买原材料,用原材料制造商品,卖出商品取得现金,从而实现交换价值。同理,大数据也只有与具体的业务场景融合,才能实现其使用价值,这个过程就是大数据的资产化,也就是大数据资产的形成过程。在这个过程中,需要明确两个理论问题。

第一,大数据的成本包括哪些?成本是经济活动中为取得一项权利或物质需要的对价付出,但数据产生的边际成本几乎为零^[5]。比如某人早上用智能手机搜索购物和就餐信息,通过网约车前往商业场所,在等餐间隙打开手机游戏,随后又在购物网站下单等待明天收货,且随着互联网和智能手机使用量的增加,他们的需求和偏好痕迹也在增加^[6]。以上行为信息,对信息的产生人来说,是无意识的、没有成本的,也不具有商业实质,但通过对这些离散式信息的采集和处理成为可应用的大数据后,便具有了商业实质。因此,大数据的成本来源于数据信息搜集和处理,包括技术研发、系统设计以及人工等,但技术越成熟,大数据搜集和处理的边际成本将呈阶梯递减趋势,即随着数据搜集和处理量增加,成本递减。只有当新技术、新算法出现时,才会产生新的研发、设计和人工成本。

第二,大数据的价值、使用价值和交换价值是如何形成的?马克思政治经济学指出,价值是指凝结在商品中无差别的人类劳动。大数据是依靠人类无意识行为或客观物质本身所产生的信息形成的,依托搜集和处理信息的劳动而形成价值。因此,大数据的价值是凝结在商品中的无差别的间接人类劳动,但这只是大数据价值的一部分。价值与使用价值和交换价值紧密联系。马克思商品流通理论认为,商品是具有使用价值的产品,使用价值是产品成为商品的首要条件,不具有使用价值就不能用于交换,也就不能成为商品。大数据无疑具有使用价值,但这种使用价值是变化的,是依场景而定的。企业可以选择在自身经营场景中使用大数据,例如京东、阿里巴巴等数据寡头企业,在其庞大的产业生态圈内,对商品生产、流通以及消费需求数据进行采集、分析和运用,以大数据驱动商流、物质流、资金流等要素,进而提升流通效率。此时,大数据具有使用价值,但因其属于内部使用,没有用于外部交换,因此尚未产生交换价值,也无法进行定价或估值。但大数据对提高运营效率的作用催生了市场需求,由此,专门采集、处理数据并以数据分析和交易为主业的公司促进了大数据交换价值的产生,大数据也就此沿着“商业实质—使用价值—交换价值”路径完成了资产化过程。

大数据一旦形成使用价值或交换价值,就和土地、资产、劳动等一样,成为国民经济发展的关键生产要素。党的十九届四中全会决议中首次将“数据”增列为生产要素,并在2020年发布的《中共中央国务院关于

构建更加完善的要素市场化配置体制机制的意见》中为数据要素市场改革提供了方向。但数据要素市场的建设还有诸多困难要克服,比如大数据的确权问题,除了少量权益比较明确的数据资产以外,还有大量数据的所有权、处理权和收益权是不明确的。明确这些权益是数据核算的基础,更是数据交易的前提。因此,大数据如何确权和定价,是当前数据要素市场建设最为核心的问题,也是亟须解决的问题。

三、对大数据确权、交易定价的理论探讨

(一) 大数据确权、定价相关文献综述

1. 大数据确权研究理论进展。大数据确权问题主要集中于数据权利应侧重于人格权还是财产权。部分学者认为数据权利要突出人格权,包括姓名权、隐私权、名誉权、肖像权等内容^[7]。其中,隐私权是数据权人格属性的重要代表,民事主体的私人生活安宁与个人信息不应被他人非法知悉、搜集、利用和公开^[8],数据权利中包含人格权,体现了人的自由与尊严,也是我国的法律权利之一^[9]。随着数据交换、数据交易等市场行为的产生,数据显现出经济利益属性,在数据上设定财产权被越来越多的学者认同^[10],但这种财产权与民法上典型的财产权不同,是一种以私益结构为核心、多层限制为包裹的复杂法律秩序构造^[11]。同时,大数据财产权的客体是源于数据集合形成过程中信息熵减所产生的经济价值^[12],这一点也区别于传统财产权类型。至此,又引发一争论热点,即数据权利如何产生或产生于哪个环节。从数据的生产机制看,算法在数据价值与数据权利的形成中处于核心地位,通过算法规制反向实现数据确权更符合实际^[13]。从制度经济学上看,数据权利形成于一组多个权利集合的“权利束”,包含了个人、集体、组织、国家等多元主体,以及人格权、财产权等多样权利^[7]。为进一步将财产权细分来平衡各方利益,有学者提出应以“汗水原则”“创造者归属原则”来确定数据的所有权,以“最佳利用者原则”“效益原则”来确定数据的使用权^[14]。其中,所有权是数据确权的争议焦点,主要的观点包括:数据载体说,即数据原始处分权归属于数据载体的所有人;制造说,即数据制造者为数据所有权的原始取得人;交易观念说,主张数据所有权应建立在交易观念基础上,归属于交易行为的供应者^[15]。总之,在数据确权领域,从数据权、数据权利到数据产权,正经历权利范式、权利-权力范式和私权-经济范式的嬗变^[16],形成了多维度、多视角的数据产权制度认知^[17]。本文认为,现有文献对数据确权的研究虽视角多元、论证多样,但多聚焦已被搜集、处理并形成可利用、可交易的数据集合这一状态,而对数据产生于个体行为这一关键点关注不够,忽略了个体在数据确权中应享有的基本权益。从数据的本源逐层、逐阶段探究数据确权问题才是关键。

2. 大数据定价研究理论进展。对于大数据定价方法的研究,主流是拍卖博弈视角,通过无限制供应拍卖的真实性和效率确保交易各方利益保持一致^[18],在固定预算限制下纳入个人理性、比例购买因素的拍卖机制^[19];或在界定私人数据拍卖概念下的多单位采购拍卖机制等^[20]。但 Jentzsch (2014) 认为拍卖机制不能反映敏感数据产品的价值^[21],而且感知价格与市场实际价格存在差异^[22],因此并不能完全解决大数据产品的定价问题。刘朝阳(2016)提出,大数据有别于传统数据,传统定价方法难以解决双向不确定问题^[23],应针对不同类型的数据采用不同的定价模式,或采用不同的价值维度聚焦客户的感知价值^[24]。在博弈范畴下,彭桥等(2020)运用讨价还价和甄别定价方式,构建供需双方效用平衡函数,提出当供方拥有数据优势时,其议价能力更高^[25]。除博弈视角外,学者们还尝试引入由交易时间、服务质量感知认同度、供需信息匹配度构成的价格贴现因子,通过构建三阶段 Rubinstein 动态定价模型,求解各阶段交易双方的最优出价^[26],或引入区块链的共识机制、时间戳、智能合约和通证激励,运用期望值和信用评分系统,优化效用定价^[27],或基于数据信息熵提出无套利用定价模型,解决最小单位数据定价问题^[28],或基于用户感知价值,融合专家总体偏好信息和用户分享评价信息,运用多维偏好线性规划分析确定待估数据价格^[29]。总体上看,由于数据价值的计量缺乏客观标准,同时现实存在诸多变数,因此对大数据交易定价的研究尚未得出统一结论。比如同样数据对不同企业的价值大相径庭,有些企业开发运用得当,可增加经济效益,反之则减少。同时,数据价值的时效性多变,有些贬值速度很快,有些对时间不敏感。此外,数据应用会产生竞

争性或排他性效果,如果多个企业均处于市场竞争的同一水平,相类似的产品或服务会承受降价压力,新进入企业对大数据价值认可也会随之降低。总之,现有定价方法重数理模型推导,重数据或交易行为本身,忽略企业经营实践和商业模式特征,将导致市场中具有资源优势的头部企业利用市场势力树立单方定价话语权,攫取超额利润,对广大中小企业形成“数据壁垒”,进而不利于大数据交易市场的发展培育。因此,大数据的交易定价模型应考虑多主体参与的状况。

(二) 对大数据确权问题的探讨

“静”态数据的确权较为明确,基本上是谁开发、谁收益,但“动”态数据的确权仍是一个复杂问题。在《民法典》中,物权是指权利人依法对特定的物享有直接支配和排他的权利。物包括所称物和知识产权等。所称物包括不动产和动产,但法律规定权利作为物权客体也可适用;知识产权是指权利人依法对作品、发明、商标、商业秘密等客体享有的专有权利。很明显,《民法典》中的产权涵盖了有形和无形资产,但数据属于何种资产、具有什么权利并未说明,只提出“法律对数据、网络虚拟财产的保护有规定的,依照其规定”。大数据是按特定目的搜集并积累的海量信息,这些信息大部分产生于人的活动。在这些个人信息上依附着哪些权利,法律界和学术界尚未给予统一的公论。2015年,南京市中级人民法院对“北京百度网讯科技公司与朱某隐私权纠纷案”作出终审判决,撤销南京市鼓楼区人民法院一审判决,认定“百度的个性化推荐行为不构成侵犯朱某的隐私权”。一审判决与二审判决最大的区别是,一审判决认为用户的搜索活动为个人隐私,而二审判决认为搜索记录等 cookie 信息不属于个人信息。同样一个主体在经济活动中产生的信息,因产生主体和记录主体不一致被分为了有两个权属的信息,其中的法理值得思考。

本文认为,因微观个体行为产生的信息,无论是否被记录和搜集,其所有权无疑属于该行为个体,因为只有该个体会产生相关信息,这是个体的从物或附属物,企业虽然使用这些数据,但对数据的所有权是不完整的。本质上,产权中的权利是一种私权,是需要利益交换才能转让的权利。企业搜集并处理消费者行为信息,无论用于改善自身生产经营还是用于外部交易,无疑从中获取了利益,但企业并未与信息最初产生的个体进行利益交换。原因是个体的行为信息在被搜集和处理所形成的大数据中,其所占比重微不足道,且大数据之所以具有使用价值,数据广泛的集合是关键。也就是说,个体的行为信息对大数据价值形成的影响太过渺小,以至于个体往往放弃了对信息价值追索这一“小微权利”。在其他领域也有类似案例,例如银行对理财或存款的利息只计算到分,而真实情况则多是无限不循环小数,这些“钱”对个人来说太少了,甚至国家都没有发行相应的货币,但在整个金融系统中,这些分散的钱如果汇聚起来,将是数以亿计的大额资产。即使微小到无法计量,即使个体主动放弃了这一“小微权利”,但对个体行为信息赋予财产属性对大数据的确权仍十分必要,因为财产属性是所有权的基础。一般来说,所有权包括占有、使用、收益和处分四项权利,个体对行为信息的所有权实际上只有“占有”的权利,因为个体的量级太小,无法实现使用、收益和处分这三项权利。表面上看,行为信息发生后,对个体的影响已经消失,只是被记录在某个媒介中,个体并没有占有。但我国《网络安全法》指出,“网络运营者收集、使用个人信息,应当遵循合法、正当、必要的原则,并经被收集者同意;网络运营者不得泄露、篡改、毁损其收集的个人信息;未经被收集者同意,不得向他人提供个人信息”。即个体在法律上有权禁止相关主体搜集其信息,但由于这部分信息的价值太过渺小,维权成本太过巨大,同时,某些以这些大数据为基础开发出的新服务确实为个体带来了便利或收益,因此,很多个体放弃了追索权。但并不代表所有权发生转移,因为所有权是完全的物权,而未经允许搜集的信息的所有权是不完整的。

微观个体的原始行为信息只是构成大数据的基础,并不是最终的大数据产品。大数据企业将个体的行为信息作为原始数据,在此基础上进行脱敏处理并深度分析、整合,最终形成衍生数据,企业为此投入大量智力和劳动资源,衍生数据也已独立于原始的个体行为信息,应当为这种耗费实质投入并达到实质规模的大数据集合设置有限排他权^[17]。因此,衍生数据及附属的大数据产品应享有独立的财产性权益,搜集和处理数据的企业不能对该大数据享有完整的所有权。这是因为即使经过脱敏等技术处理,其中最核心的、最本源的信息并未脱离产生该信息的原微观个体的范畴,企业只能享有该大数据的使用权、收益权和处分

权。现行《网络安全法》第四十二条提出相同观点,“网络运营者不得泄露、篡改、毁损其收集的个人信息;未经被收集者同意,不得向他人提供个人信息”,但同时也增加了“经过处理无法识别特定个人且不能复原的除外”这一表述。

由上可知,大数据的所有权在法律逻辑上是绝对的、排他的、永续的,但在实践中则是分离的,占有权属于产生信息的微观个体,而使用权、收益权和处分权则属于收集和处理信息的主体,具体关系见图1。

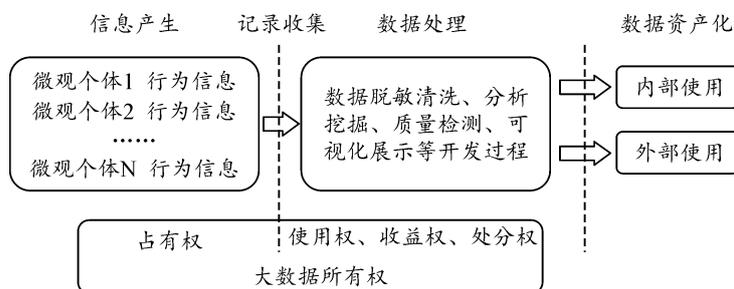


图1 大数据确权示意图

大数据作为一种极特殊的事物,个体与企业使用同一产权边界,在实践中易引发越界行为。因此,数据要素市场的构建需要更加复杂的规则、法律和服务体系,这将是一个长期的探索过程,但过严、过早的规则制定和执行可能造成“劣币驱逐良币”的不良后果。当前,大数据交易已发展出多种商业模式,与其去找到一套标准的交易规范,不如暂且放下理论争端,在一定范围内放手由市场去探索。目前,国内已设立几家大数据交易所,在数据确权、交易等领域正不断尝试,研究大数据交易所生态下的数据定价问题,为市场健康发展提供了另一种思路。

(三) 大数据交易所生态下数据产品定价分析

中国信息通信研究院发布的《大数据白皮书(2019年)》显示,全球大数据发展仍处于活跃阶段,全球数据量在2020年将达到50.5ZB,^①全球大数据市场收入规模将达560亿美元,这其中包括大数据硬件、软件和服务三大市场规模,其中服务市场规模可达210亿美元,具体见图2、图3。^②

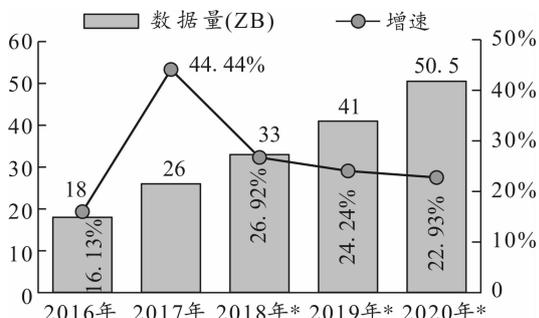


图2 全球每年产生数据量估算图

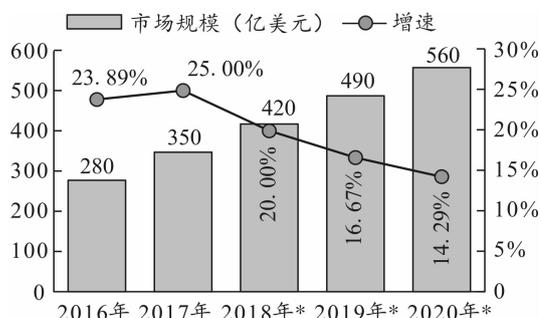


图3 全球大数据市场收入规模预测

目前,国内外提供数据产品和服务的公司主要有四类:一是拥有海量自有数据的互联网企业,如阿里、腾讯、京东等;二是拥有先进大数据平台搭建技术和数据分析处理技术的IT企业,如华为、中兴、曙光科技等;三是依赖数据分析处理工具针对用户需求提供数据服务的第三方服务公司,如帆软、海运数据、甲骨文

①ZB,即十万亿亿字节,相当于 2^{40} GB。

②图2、图3来源于中国信息通信研究院,其中带*年份为预测值。

等;四是专业提供数据撮合交易的大数据交易所,如贵阳、上海等地以及北京正在筹备的国际大数据交易所。无论理论上如何将大数据定价模型化、复杂化,在交易实践中,只有作为市场参与主体的生产、交易、购买企业均从数据交易中获益,交易行为才会产生。沿此思路,本文从大数据资产化形成视角,尝试对大数据交易所这一细分市场,对各类大数据产品的交易定价进行分析。交易场所是指根据《国务院关于清理整顿各类交易场所切实防范金融风险的决定》(国发〔2011〕38号)和《国务院办公厅关于清理整顿各类交易场所的实施意见》(国办发〔2012〕37号)等规定,在各省市区行政区域内依法设立的从事权益类交易、大宗商品交易以及其他标准化合约交易的交易场所。在我国,交易场所的设立必须由地方政府审批,属于“7+4”类地方金融组织,^①交易规则由证监会审核,机构由地方金融监管局监管,风险相对较低。市场上其他将数据变现的商业模式可归类为“场外市场”或服务市场,与专业交易场所的商业逻辑不同。交易场所是商品流通的高级业态,在业务撮合、价值发现、标准制定、商品定价等方面拥有重要话语权,各地纷纷探索设立大数据交易所,也是期待以此商业模式来探索大数据市场的规范发展。

本文主要研究大数据交易所交易流通的数据产品的定价行为,不包括数据储存、计算、挖掘和可视化分析等数据服务的定价。交易流通的数据产品是以降低用户对原数据使用门槛、提高对数据再分析技术效率为目的,在原数据基础上进行纠错、补漏、整合和标准化后的大数据集合,包括结构化和非结构化的数据产品。在交易所内交易的大数据,其价值应涵盖大数据生产方的成本与必要收益、大数据交易方的成本与必要收益、大数据产品的特征收益、大数据购买方支付意愿和购买后所得收益,大数据价格应小于大数据价值。具体关系如图4。

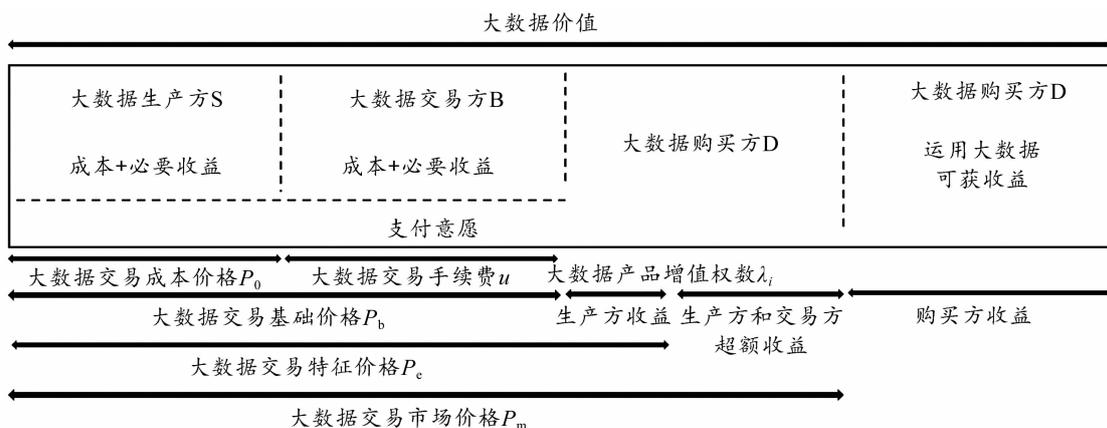


图4 大数据价值、价格构成示意图

大数据交易的价值由大数据交易特征价格、生产方和交易方超额收益、购买方收益组成,而大数据交易价格应由大数据交易特征价格、生产方和交易方超额收益组成。对于购买方收益,不同购买主体运用大数据的领域、方式和目的不同,对购买个体产生的收益不同,属于变量。对于生产方和交易方来说,大数据交易价格在弥补了大数据生产和交易的成本和企业经营所要求的必要收益后,对于因大数据产品不同而表现出的增值部分,应由生产方享有,其他都属于超额收益。对该部分,如果购买方愿意支付的价格高,则生产方和交易方获得的超额收益多,反之反是。一般来说,生产方和交易方对此并无特殊要求,这是由大数据产品交易的边际收益远大于边际成本以及大数据商品有限排他性的特征所决定的。大数据商品不像一般的消费品具有完全的排他性和竞争性,可以多频次交易,每一次交易都会给生产方和交易方带来超额收益。由此可见,我们无法预计大数据交易的真实市场价格,但可以知道大数据交易的成本价格和基础价格,只要真实交易价格大于特征价格,交易就会发生,生产方、交易方就会获利,且当产品反复交易时,获利基

^①“7+4”地方金融组织指:小额贷款公司、融资担保公司、区域性股权市场、典当行、融资租赁公司、商业保理公司、地方资产管理公司、投资公司、农民专业合作社、社会众筹机构、地方各类交易所。

数更大。在此基础上,本文从企业资本预算视角提出大数据交易价格的模型。

1. 大数据交易成本价格。即大数据生产公司在弥补数据产品成本和获得的必要报酬后的价格,数学表达式为:

$$P_0 = \frac{C_s(1 + I_s)}{N_s} \quad (1)$$

其中, P_0 代表成本价格, C_s 代表大数据公司 S 为生产该数据产品所发生的技术和人工成本, I_s 代表项目预期收益率, N_s 代表该产品预计交易的次数。项目是指具有明确目标的一系列复杂并相互关联的活动,大数据生产公司为达成财务目标,往往通过开发新产品进行项目投资,而不同的大数据产品对应的投资风险是不同的。因此,对某一个大数据产品风险的测度,应不同于该公司整体的经营风险,即不能用公司的加权资本成本,应采用可比公司法,在资本市场上寻找到与待评价项目类似的上市公司,以该上市公司的风险收益水平进行调整测度,并运用资本资产定价模型(Capital Asset Pricing Model,简称 CAPM)表示,即:

$$I_s = \frac{\beta_0 \left[1 + (1 - T_s) \frac{L_s}{E_s} \right]}{\left[1 + (1 - T_0) \frac{L_0}{E_0} \right]} (R_m - R_f) + R_f \quad (2)$$

其中, β_0 是衡量可比公司 O 系统风险的贝塔系数,作为一家上市公司,不同的评级机构会定期发布贝塔系数。 T 、 L 、 E 分别代表可比公司和大数据生产公司各自的所得税税负水平、负债投资和股权投资金额。

$\frac{\beta_0 \left[1 + (1 - T_s) \frac{L_s}{E_s} \right]}{\left[1 + (1 - T_0) \frac{L_0}{E_0} \right]}$ 实际上就是在“卸载了”可比公司财务杠杆后,大数据生产公司 S 投资该大数据产品项

目的系统风险, R_m 代表市场组合收益率,通常用股票价格指数的收益率的平均值或所有股票的平均收益率来代替, R_f 代表市场无风险利率,通常用长期国债收益率表示。

需要说明的是,本文中大数据生产公司是个集合概念,是对处于大数据生产链上各类公司的统称。从数据的生成级别上看,数据交易市场存在原始数据、加工处理后的数据以及由多份数据整合后的新数据,分别对应数据生产者、数据加工者和数据整合者。各利益主体的资本结构、收益模式不同,要根据自身特点设计成本指标,最终可单独或累进到大数据交易成本价格中。

2. 大数据交易的基础价格。有了成本价格 P_0 ,基础价格就是在此基础上加上大数据交易方 B 收取的交易手续费,即:

$$P_b = P_0 + u \quad (3)$$

$$(uN - C_B - L_B I_B)(1 - T) = E_B K_B \quad (4)$$

$$u = \frac{\frac{E_B K_B}{(1 - T)} + L_B I_B + C_B}{N} \quad (5)$$

其中, P_b 代表大数据交易的基础价格,也是最低价格,该价格应保证大数据生产方和交易方能够弥补成本和获得必要收益。 u 代表每次交易的平均手续费, N 代表交易所每年交易产品的次数。需要说明的是,当前我国大数据交易所的交易活跃度并不高,因此可以根据历年交易数据推断交易次数。假设大数据交易方 B 以数据交易作为主营业务和主要利润来源,公式(4)代表总体的手续费收入在扣除不含债权融资成本 $L_B I_B$,只包括交易所运营的各项成本 C_B 并缴纳所得税后,可分配给股东的利润 $E_B K_B$ 。由此,可推演出交易费用 u 的表达式(5)。

3. 大数据交易的特征价格。因数据产品的特征不同,每一个数据产品应结合行业特点进行特征定价评估,在成本价格和基础价格之上,针对不同行业需求设计不同“增值权数”。在大数据交易所建设初期,从市场供需原则出发,“增值权数”可以一定周期内大数据交易的 ZB 量为基数,用每个行业在本周期内交易量除以总的交易量作为权数,同时设置封顶标准,计算出本周期的“增值权数”,再乘以大数据基础价格,

从而反映某类大数据产品的周期性供需关系。即：

$$\lambda_i = \frac{\sum_1^n q_i}{\sum_1^m Q_{ij}} \quad \text{其中: } \lambda_i \leq B; i, j = 1, 2, 3 \cdots n, m \quad (6)$$

$$P_e = P_b \times (1 + \lambda_i) \quad (7)$$

其中, λ_i 表示产品 i 的“增值权数”, B 表示封顶标准, 例如设置上限为 10%, 防止某一个周期只交易一种产品造成权数过大问题。 $\sum_1^n q_i$ 表示产品 i 从 1 到 n 交易次数下的交易量, Q_{ij} 表示该交易所内所有数据产品 i 到 j 的从 1 到 m 交易次数下的交易量。 P_e 表示大数据产品因属性不同表现出的特征价格, 由 P_b 与增值权数合并计算组成。特征价格与基础价格之间的差异, 即该部分收益, 应由大数据生产方 S 独享, 以体现大数据生产方根据数据品种、时间跨度、数据深度等因素设计制造大数据产品应享有的补偿, 来进一步激励提高其数据生产能力和数据发现能力。

4. 大数据交易的市场价格。只要大数据交易市场价格 P_m 大于等于特征价格 P_e , 交易就会产生, 而超出部分将按照交易费率在生产方与交易方之间分成, 再形成一部分的各自超额收益。其中, P_m 与 P_e 之间的差额, 则需要购买方 D 结合自身运营效率、市场前景、产品开发能力等因素, 通过财务上项目资本预算方法来判断成本收益预期。这部分并不是购买方的确定收益, 因为存在购买方误判、市场环境变化等因素, 导致收益不及预期甚至亏损。同时, 同一大数据产品也可能出现多个 P_m , 但这并不说明存在价格歧视。价格歧视是对同一消费群体在时空分割下的定价, 瞄准的是信息不对称和交易成本不同条件下的价格最大化, 而大数据交易所提供的是公开的市场环境, 因购买者各自判断不同而产生价格差异。因此, 大数据市场价格 P_m 类似于“赌石价格”或“盲盒价格”, 大数据价格 P_m 超出 P_e 部分的风险由购买方承担, 大数据价值超出 P_m 部分的收益由购买方独享。从实践上看, 国内大数据交易所对数据的定价多采用在一定评估价值基础上的供需撮合模式。例如, 贵阳大数据交易所于 2016 年推出《数据定价办法》, 制定了协议定价、固定定价、实时定价三种模式。其中, 实时价格通过交易系统自动定价, 价格实时浮动并由卖方与交易所最终确定。总之, 交易中的数据产品价格是动态的, 这符合当前数据交易市场供需双方特点, 但在动态价格中应有其基础价格, 最终的交易价格应围绕这一基础价格浮动。

总体上看, 采用企业资本预算视角设计大数据交易所数据产品价格模型, 比其他理论定价模型更易操作。在市场发展不完善, 尚有诸多问题有待解决的前提下, 有利于进一步降低市场壁垒, 在交易所生态下吸引更多供需方参与。通过让市场去探索大数据的价格形成机制, 推动交易各方能够弥补成本和获得收益, 提高大数据产品的利用效率, 这也是市场走向成熟和可持续发展的必然过程。

四、结论和建议

后疫情时代, 以数字化、网络化、智能化为特征的第四次工业革命将持续推进, 以高效串联、渗透融合、柔性重组为特征的数字经济正深刻改变着各行各业的商业模式, 推动产业链、价值链不断深度重组。在数字经济环境下, 大数据将贯穿生产、流通、消费各个环节, 成为提高生产生活效率、推动经济转型创新的重要驱动力。本文通过对大数据概念、属性、确权、资产化、交易定价等基本问题进行深入探讨, 提出大数据确权要考虑数据信息产生个体和数据产品生产主体的产权边界, 大数据资产化、数据产品定价与企业经营密不可分, 应鼓励部分基础性数据产品在交易所内交易, 逐步培育做大数据交易市场, 让市场检验大数据产业各种新商业模式的适用性。在发展中不断完善大数据确权、交易定价等理论和实践问题, 大数据交易所则是这一过程的重要突破口。结合本文论述, 对依托大数据交易所建设探索数据交易提出如下建议:

1. 建立政府背书的大数据产品确权登记平台。大数据产品包含着个人、企业等多主体复杂的法律关系, 各方共用权利边界, 极易引发产权争议, 政府应发挥“仲裁者”或“保护人”作用, 建立政府背书的大数据产品确权登记平台。大数据生产者应在产品确权以前, 将数据采集范围和过程向社会公示, 在《网络

安全法》等现有框架下,充分履行生产者披露义务,为大数据信息“源头”即信息产生的个体提供权利确认机会。个体可对自身信息被采集提出异议,生产者在数据清洗和设计产品时须剔除该部分信息。若一定时间内未有异议,则确权平台可按企业申请确认该产品权属。

2. 在交易环节征收“数据税”弥补个体私权损失。在大数据价值的形成环节,实际上存在着广大个体所放弃的“小微权利”,但整合以后形成的大数据产品,确实能为整个社会经济发展带来外溢效应,政府应运用税收手段,将这种外溢转变成公共产品反哺大数据“源头”,即信息产生的个体。具体来说,政府对交通、城市管理等服务领域数据开发所产生的收益,应继续用于智慧社区、智慧城市等民生领域投入。对企业商业行为的大数据,在交易环节可比照印花税征收“数据税”,并用于学校、医疗等公共产品的支出,让广大个体受益。这一过程,本质上是将私人权利汇总形成社会公共产品,实现社会价值最大化,这也是征收“数据税”的法理基础。

3. 探索以供求为基础的大数据价格“试价”机制。大数据产品交易不应是固定价格或浮动价格,而应是“试价”机制,特别是基础性的数据产品,供方和交易方通过成本加必要收益确定底线价格,需方根据自身项目预算报价,对需方报价高于底线价格的要约即可成交,反之反是。同时,根据市场供需变化调整不同类型产品的“增值权数”,让产品价格更加贴近市场需求。在此过程中,生产方不能独享超额收益,生产方、交易方以及购买方都要分享大数据的价值,购买方获取收益越多,越会有购买者购买该产品。“增值权数”的增加会使生产方获得更多收益,进一步鼓励产品创新,交易方通过手续费、政府通过“数据税”也可获得更多收入,市场各方均可从中获益。

4. 大数据交易所要以基础性产品为交易重点。大数据交易所要发挥产业链引领和塑造作用,通过降低交易成本,鼓励更多主体参与到大数据价值挖掘中来。目前诸多数据需求者只需要市场上某些子数据集,而交易平台往往将完整的数据集打包报价,强迫需求方购买不感兴趣的部分数据。本文从大数据资产化形成视角提出的大数据产品定价模式,总体上价格可接受度较高,有利于让更多市场主体参与到交易中来,而大数据产品的生产方和交易所,因产品的多频次交易也可不断累积收益。在这种交易生态下,生产方可将基础性、推广效果较高的子集类产品作为交易物品,从而推动大数据应用的普及化。对于定制化、应用较为复杂的产品,可在各类场所外市场交易,以实现更多收益。因此,大数据交易所现阶段存在的价值并不在于盈利,而更应注重长远的行业塑造。

参考文献:

- [1] 纪思亮. 大数据啥时能从技术变成科学——浅谈大数据理论基础的构造[J]. 中国统计, 2016(5): 16-17.
- [2] 刘军华. 大数据理论进化和统计学融合博弈中的机会、风险与管理[J]. 统计与信息论坛, 2020(2): 24-31.
- [3] 钟穗. 大数据经济学对一般均衡经济理论发展研究[J]. 信息系统工程, 2018(3): 156.
- [4] 邬贺铨. 关于大数据的若干思考[J]. 中国信息化, 2014(5): 3-7.
- [5] SHAPIRO C, VARIAN H R. Information rules: a strategic guide to the network economy [J]. The Journal of Economic Education, 1999, 30(2): 189-190.
- [6] LAMBRECHT A, TUCHER C E. Can big data protect a firm from competition? [J]. Available at <https://ssrn.com/abstract=2705530> or <http://dx.doi.org/10.2139/ssrn.2705530>, 2015, 11, pp. 5.
- [7] 程啸. 论大数据时代的个人数据权利[J]. 中国社会科学, 2018(3): 102-122.
- [8] 朱启超, 许德风. 民法概要[M]. 北京: 北京大学出版社, 2002: 139.
- [9] 闫立东. 以“权利束”视角探究数据权利[J]. 东方法学, 2019(2): 57-67.
- [10] 杜振华, 茶洪旺. 数据产权制度的现实考量[J]. 重庆社会科学, 2016(8): 19-25.
- [11] 龙卫球. 再论企业数据保护的财产权化路径[J]. 东方法学, 2018(3): 50-63.
- [12] 吴亚光. 论数据财产权成立的权利客体基础[J]. 图书馆建设(网络首发版), <https://kns.cnki.net/kcms/detail/23.1331.G2.20200724.1454.002.html>.
- [13] 韩旭至. 数据确权的困境及破解之道[J]. 东方法学, 2020(1): 97-107.
- [14] 彭云. 大数据环境下数据确权问题研究[J]. 现代电信科技, 2016(5): 17-20.

- [15] 纪海龙. 数据的私法定位于保护[J]. 法学研究, 2018(6): 72-91.
- [16] 文禹衡. 数据确权的范式嬗变、概念选择与归属主体[J]. 东北师大学报(哲学社会科学版), 2019(5): 69-78.
- [17] 崔国斌. 大数据有限排他权的基础理论[J]. 法学研究, 2019(5): 3-24.
- [18] RIEDERER C, ERRAMLILLI V, CHAINTREAU A. For sale; your data; by you[C]. ACM Workshop on Hot Topics in Networks, Cambridge, 2011: 1321-1325.
- [19] DANDEKAR P, FAWAZ N, IOANNIDIS S. Privacy auctions for inner product disclosures[J]. ACM Transactions on Economics and Computation, 2011(2): 31-34.
- [20] GHOSH A, ROTH A. Sell privacy at auction[J]. Games and Economic Behavior, 2015(95): 334-346.
- [21] JENTZSCH N. Auctioning privacy-sensitive goods[C]. Annual Privacy Forum. Springer, Cham, 2014: 133-142.
- [22] CASTELLUCCIA C, OLEJNIK L, MINH-DUNG T. Sell off privacy at auction[C]. California; Network and Distributed System Security Symposium(NDSS), 2014: 133-142.
- [23] 刘朝阳. 大数据定位问题分析[J]. 图书情报知识, 2016(1): 57-64.
- [24] 张晓玉. 基于讨价还价博弈的大数据商品交易价格研究[D]. 鞍山: 辽宁科技大学工商管理学院, 2016: 24-42.
- [25] 彭桥, 肖尧, 陈浩. APP用户数据交易与隐私保护问题研究——对比讨价还价与甄别定价两种交易模式[J]. 产经评论, 2020(3): 3-15.
- [26] 张树臣, 陈伟, 高长元. 创新联盟大数据服务交易模式及动态定价模型研究[J]. 情报杂志, 2020(3): 187-197.
- [27] 程琳, 黄光玉, 王政, 等. 基于区块链的大数据定价机制及交易模式研究[J]. 全国流通经济, 2019(8): 132-135.
- [28] 李希君. 基于信息熵的数据交易定价研究[D]. 上海: 上海交通大学软件学院, 2018: 23-42.
- [29] 左文进, 刘丽君. 基于用户感知价值的大数据资产估计方法研究[J]. 情报理论与实践. <https://kns.cnki.net/kcms/detail/11.1762.G3.20200721.1543.002.html>.



(责任编辑 游旭平)