

doi: 10.15936/j.cnki.1008-3758.2020.01.002

“大数据杀熟”背后的伦理审思、治理与启示

李 飞 翔

(南京航空航天大学 马克思主义学院, 江苏 南京 211106)

摘 要: 大数据技术在提升社会运行效率、丰富人们日常生活的同时亦带来了“杀熟”的问题,不仅侵犯了用户的合法权益,还进一步影响着人们对技术的信任程度。技术发展的不成熟、过分逐利的资本诉求、“不在场”的伦理规范、监管的复杂与不到位及用户对维护自身合法权益的疏忽与艰难等共同导致了如此恶果。据此,可以试着从数据自身、互联网平台以及用户三个角度出发去重新审视“大数据杀熟”背后所蕴含的伦理反思:主要涉及到数据的拥有与分配、算法所应遵循的伦理规则、责任主体的合理界定、相关平台的信任机制之维护以及用户自身“数据权”意识的觉醒等重要议题。唯其如此,整个社会方有可能在数据的运用与约束方面朝着“科技向善”的理念持续迈进。

关 键 词: “大数据杀熟”; 数据所有权; 算法伦理; 互联网平台; 隐私伦理; 多方博弈

中图分类号: N 031 **文献标志码:** A **文章编号:** 1008-3758(2020)01-0007-09

Ethical Reflections, Governance and Implications Under the Background of Big Data Price Discrimination

LI Fei-xiang

(College of Marxism, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Big data technology improves social operation’s efficiency and enriches people’s daily life, but it also brings the problem of price discrimination in the process of development. It not only infringes the legitimate rights and interests of users, but also affects people’s trust in this advanced technology, which maybe due to the imperfection of technology development, the excessive pursuit of capitals, the absence of ethical norms and principles, the complexity and inadequacy of supervision and the users’ negligence as well as difficulty of protecting their own legitimate rights. Therefore, such perspectives of data itself, Internet platforms and users can be taken to reexamine the hidden ethical reflections of big data technology, mainly involving several key points-the ownership and distribution of data, the ethical rules that algorithms should follow, the reasonable definition of subjects of responsibility, the maintenance of the trust mechanism of relevant platforms, and the awakening of users’ awareness of “data rights”. Only by doing so can big data technology be better understood and applied, and the whole society is likely to make continuous progress towards the concept of “Tech for Social Good” in the application and restriction of data.

Key words: big data price discrimination; data ownership; ethics of algorithm; Internet platform; privacy ethics; multi-player game

“在线化”已然成为人们的生活常态,由此产生了数量庞大、维度丰富的数据,还进一步催生了“数字经济”的蓬勃发展。一定程度上,大数据已成为继水、电、天然气、煤等之后新的社会发展的基础性资源,其对人类社会的渗透、介入与影响与日俱增。

然而,大数据在促进社会发展、便利人们生活的同时,也潜藏着诸多隐患与风险。近两年“大数据杀熟”成为社会热议的现象,有相当一部分人更是表示被大数据“杀熟”了很长时间而浑然不知。由此,本是出于便利人们生活而发展壮大的大数据技术是如何以隐蔽的方式“杀熟”的,面对巨大利益诱惑的互联网平台如何在商业利益与社会责任之间形成一种动态的平衡,用户应如何维护自身合法权益,如何对大数据进行有效治理以期更符合社会发展的要求等问题都促使我们需要用理性批判且具有建设性的眼光去重新审视大数据技术。

一、大数据技术的阴暗面: 隐蔽式“杀熟”

1. 大数据技术及核心特质

“大数据技术是目前正在兴起的一场新技术革命,从本质上来说是信息革命的延续。数据就是有根据的数字编码,也就是数字编码加上其背景知识,它是信息的一种科学表达。大数据的真正本质还是其数据化的世界观和思维方式。”^[1]根据知名伦理学者岳缙的看法:“‘大数据技术’是基于大规模生产、分享和应用数据的互联网平台,发掘数据价值的一种新形态的数据挖掘技术。”^[2]按照大数据的底层思维方式,我们分析任何问题都应以数据化的整体论眼光,以相关性分析手段分析来自不同途径的数据^[3]。说到底,大数据自身所拥有的极其庞大且维度丰富、多元化的数据仅仅是表象,它本质上构建了一种在信息社会重新看待、分析及评判整个世界的思维范式与认知系统:尽可能地把整个世界“数据化”(可量化、可追踪、可分析),通过对数据的整理、分析与挖掘来更好地认知、预测世界与自我,从而不断提升人们改造世界的效率与能力。

与传统数据相比,移动互联网时代的大数据主要具有“4V”特征:①容量大(volume),即巨量数据规模及其完整性;②速度快(velocity),即快

速的数据流转和动态的数据体系,可以更快满足智能化、实时性的要求;③多样性(variety),即数据类型繁多,包括结构性文本,以及半结构性、非结构性的视频、图片等;④价值化(value),即大数据的价值密度低,如同大浪淘沙般从规模巨大、类型繁多的数据中快速提取有价值的信息,但一经有效提取,其巨大的使用价值就会体现出来^[4]。预计到2020年,全球每年产生的数据信息将达到35.2ZB^[5](ZB是计算机专用术语,代表的是十万亿字节)。这足以说明大数据数量之“大”,而且这种“大”每时每刻都还在呈指数级增长;大数据的种类繁多意味着它是多维度而非线性化的,更强调“相关性”而非“因果性”,这是它能够全面认知事物与自身、无限逼近世界本质的关键性特征;大数据的价值密度低就决定了进入该领域的门槛较高、投入较大,不仅需要拥有海量的数据来源,还需具备安全存储巨量数据及对数据进行深度挖掘和分析的超强计算能力以及商业化运作能力等,这就决定了仅有少数国家和企业才有资格和能力进行大数据技术的研发、推广与创新。有了大量且丰富的数据样本作支撑,借助于云计算技术,不仅有助于企业更好地了解用户的需求,从而大幅度提升公司的运营效率,与此同时,大数据技术还有助于企业以较为低廉的成本高效地对个体提供更为个性化与多样化(千人千面)的服务,从而使得多数人能够享受技术发展带来的红利。

2. “大数据杀熟”及其主因

大数据技术在提升社会协作效率、给人们带来诸多便利与舒适的同时,也逐渐引发了一系列问题,“大数据杀熟”便是恶果之一。所谓“大数据杀熟”,通俗说来就是平台(主要是互联网平台)充分利用自身所掌握的大数据技术对消费市场进行更为精准与细腻的划分,在此基础上主要对熟人(习惯、依赖该平台的较为忠诚的用户)进行不当地利益宰割,从而使大数据技术成为部分经营者追求超额利润的有力工具。也就是说,面临同样的商品或服务,具有忠诚度的用户群体看到并为此实际支付的价格反而比新客户或一般用户要贵,而且这种行为极具隐蔽性。一言以蔽之,大数据技术加剧了经营者和相对人之间的信息不平等,消费者很难通过网络对经营者价格歧视的抗辩进行甄别^[6]。

究其本质,“大数据杀熟”的技术原理就是互

联网平台凭借其所掌握的极其庞大且维度异常丰富的数据,诸如用户的个人身份信息、位置信息、聊天记录以及支付信息等一切有可能被线上记录的数字化信息,然后通过一整套复杂、高效而又极其先进的数据运算、分析和挖掘技术对碎片化、零散的用户数字化信息进行全方位的扫描、分析与研究(数据标注),然后通过一些关键词的标注对用户进行细致归类,从而生成独特的用户画像。据此,互联网平台与商家在用户不知情的情况下对具有不同画像特征的用户制定不同的价格机制与定价策略,造成用户非知情前提下的“价格歧视”,从而最大程度上攫取利益。概而言之,此种现象的出现主要归结为以下几个因素。

第一,移动互联网的垄断效应导致数据的高度集中,进而导致“赢家通吃”的局面。作为人类历史发展进程中极具技术革新色彩且大力推崇开放、平等以及共享精神的互联网,伴随着时间的推移与技术的溢出效应,互联网尤其是移动互联网正朝着与其发展初衷相反的方向演化,移动互联网的集中趋势与垄断效应日益突出,呈现出一种日渐封闭与保守的异化态势。

众所周知,移动互联网的发展很大程度上是建立在用户规模的基础上的。用户规模越大,盘踞在互联网内的节点也就越多,节点与节点之间的链接才有可能进一步扩大,整个互联网才有可能发挥出更大的网络协同效应,其综合价值也就呈几何式增长。除此以外,较之于传统经济样态,互联网经济的边际成本非常低,收益却可以很高。在互联网的世界中,每个用户及其相关的行为都意味着数据,数据的大规模产生与集聚不仅是移动互联网发展的必然结果,也是其得以持续扩张、壮大的内在要求与逻辑前提。

经过市场的激烈竞争与合纵连横,“二八法则”在互联网世界表现尤为突出,当前网络的集中趋势与垄断效应越来越明显。就世界范围内来看,移动互联网主要有五大超级巨头,即所谓的“FAANG”:依次是 Facebook, Apple, Amazon, Netflix 以及 Google。具体到国内,移动互联网江湖也存在着“三座超级大山”——“B(百度)A(阿里巴巴)T(腾讯)”。移动互联网的发展已逐渐从江湖混战进入巨头主导和统治的聚合时代,这就为数据的高效整合与集中挖掘奠定了基础。基于此,互联网平台就可以通过对其所掌握的数据进行标签化细致管理,进而形成规模效益。同时也

意味着大多数个体在大数据面前几乎就是一种透明性的存在,大数据可能比你自已更熟悉你自己。这一方面可以降低社会成本,更好地为用户提供个性化、定制化的商品与服务;与此同时,巨头的集中与垄断也为其低成本、隐蔽式“杀熟”提供了技术上实现的可能性。

第二,用户对互联网平台的依赖助长了“大数据杀熟”的发生。用户一旦习惯了某个互联网平台或应用程序,就容易导致“路径依赖效应”。即,由于此平台拥有良好的使用体验和独特的价值,于是用户便会逐渐汇聚于此平台。伴随着用户规模的增加以及互联网平台、应用程序自身持续不断地迭代、更新与升级,再加上互联网的口碑传播效应,随后还会有越来越多的用户涌入此平台且被牢牢粘住。而且,用户一旦适应、习惯乃至依赖某一平台或应用时,他便会逐渐把很多关键性的个人信息、社交链信息、诸多交易行为、文化消遣娱乐及工作交流等行为逐渐固定在这上面,从而对平台或此应用程序形成一种较为稳定且持久的依赖关系。伴随着平台或应用程序对用户的不断了解,它会为用户提供更为精准、个性化的服务或商品,从而进一步强化用户对它的喜欢和依赖,在此过程中就容易出问题。据此,人性本身追求便利、稳定与舒适的本性以及平台迁移成本理性考量等因素,因此用户哪怕会面临“被杀熟”或其他可能的风险,在其可承受的范围内都会倾向于继续停留在此平台上。由此,用户的习惯与纵容就为“大数据杀熟”提供了适宜的土壤和环境,而这反过来又强化了互联网大平台的集中趋势和垄断效应。

第三,国家相关部门在大数据的立法、监管及执行方面存在滞后性与模糊性。总体而言,国家和政府对大数据技术发展持包容、鼓励与支持的态度。党的十九大提出“推动互联网、大数据、人工智能和实体经济深度融合”,从而对我国实施大数据战略提出了更高、更具体的要求。但是,具体应该由哪些部门进行合法监管,用户、企业及政府之间应如何对数据的拥有、分配以及存储等进行合理协调,如何统筹兼顾合理监管、伦理审查与大数据商业开发之间的复杂关系,如何有效认定“大数据杀熟”行为的发生以及倘若真涉及到此类问题的发生,究竟应该如何进行处罚等难题仍旧是值得各方争论与探究的议题,尚未有一套较为清晰、能够得到各方认可与执行的法律法规。

二、“大数据杀熟”的三维伦理审思： 数据、平台及用户

针对“大数据杀熟”，很多互联网平台及关联公司都矢口否认，用户也往往处于十分被动的弱势地位。纵然遇到消费者投诉或是媒体曝光，互联网平台也总是能够充分利用相关法律规范的漏洞或模糊性设定为自己开脱，或以“派发优惠券”“对新老用户区别对待”的名义混淆视听。甚至，部分专家学者也出于自身利益的考量为其摇旗呐喊。种种不正常行为的发生促使我们更要以负责任、理性、专业且充满良知的态度去研究此类现象，以期正本清源。

大数据、网络平台以及用户等构成了一个完整的生态圈。从高效而低成本地提供个性化服务到随意而隐蔽地“杀熟”，其背后涉及到“算法”（通过对数据的计算、挖掘而最大化它的价值）与“善法”（大数据及其算法所应该遵循的基本价值规范与准则）之间的冲突与失衡。一旦大数据背后的算法不再以追求良好的用户体验和严格遵循“向善”为宗旨的话，而是追逐一种“算法”高于“善法”、经济利益重于伦理规范的异化价值观的时候，诸多问题与恶果便会相继出现。伦理性反思意味着大数据技术所应该达致的目标，是一种“应然状态”。单纯就“大数据杀熟”的现象而论现象，很容易陷入“各执一端”“浮于表面”的尴尬境地。唯有从伦理性的应然状态出发，才能够更合理地去说明“大数据杀熟”之所以不公正的本质逻辑，也能够为我们更好地对大数据技术治理与算法优化奠定一些基础性的价值规范与行为导向建议。

1. “杀熟”行为背后的数据伦理之思

数据不仅仅是人们认识世界、改造世界的客观性工具与利器，还应该具备伦理性与价值性的属性，且能够承担相应的社会责任。具体而论，主要涉及到以下几个方面的争论。

第一，数据的所有权之争。数据的所有权在很大程度上决定着数据的进一步开发与运用，也决定着一旦出了问题责任划分与执行的问题。事实上，我们现在默认的情况是数据归互联网平台及公司所有，用户在它们面前势单力薄不具备博弈的资本，这也是网络平台之所以能够对用户“杀熟”的重要原因。需要反思的是：网络平台是否天然的、理所当然的就具备拥有用户数据的权利呢

（事实上掌握、拥有不意味着必然合理）？如果拥有，它们可以在何种程度上对相关的数据进行存储、挖掘以及商业开发？

具体到用户，用户个体所产生的数据也许并没有多少价值，很多用户更在意服务的便捷性而非数据的所有权问题。但是，当这些数据经过长期的积累与沉淀达到一定规模时，就开始变得价值连城了。互联网公司通过对数据的收集、分析与标注，从而精准地把相关的广告推送给与之匹配的受众群体，从而获得巨额的利益回报，而用户不仅未获得相关的回报还隐蔽地“被杀熟”。因此，互联网平台上的用户不仅仅是享受服务的消费者，还是产品本身，兼具商品的属性。

据此，从数据的所有权来看，用户是数据的直接来源和基础，当然拥有对数据的所有权；网络平台及公司在数据的价值生成与应用推广过程中起着相当重要的作用，也付出了时间、经济与人力等方面的高额成本，在一定程度上也应该具有对数据的所有权。问题的关键在于网络平台处于较为强势的地位，当下的平台存在着对用户数据过度收集与开发的失范倾向，兼具产品与用户双重属性的个体在上传个人数据、保护个人必要隐私不受网络平台侵犯方面并无实质上的选择权与自主权，这就为个体被“杀熟”、隐私被侵犯以及信息被随意买卖等不正当行为的发生提供了现实的可能性。

第二，数据伦理的责任主体问题如何界定。“技术本身是没有价值牵涉其中的，是中性的，但是由于主体的利益牵涉其中，这就使得大数据技术的伦理问题更为凸显。”^[7]伴随着数字化时代的来临，每个个体在一定程度上都可以被还原为互联网世界中的一连串数字符号与代码，数据本身就意味着价值和资源，而不同的主体或组织如何运用和规范大数据、责任如何确定与划分等成为了各方争论的焦点。正如著名伦理学家唐凯麟所言：“在运用大数据的过程中，不同层次的组织与用户往往从自身的利益出发，以追求利益最大化为目标实施行动，这可能侵害到利益相关者的利益。”^[8]就此而论，我们可以从两个角度出发去确认和衡量责任主体：第一，在承认互联网平台与公司也具有对数据所有权与开发权的基础上，基于“权利和义务对等”的原则，它们在获取丰厚利益的同时也应该承担相应的责任、履行必要的义务；第二，侧重从“后果论”的角度去进行责任界定与执行，视杀熟行为的严重程度及对用户经济、精神

伤害程度的轻重而确责。据此,“大数据杀熟”意味着互联网平台应该承担主要的责任。

第三,算法亦必须遵循相应的伦理准则。大数据技术的核心在于算法,没有算法作支撑,数据的综合价值便会大打折扣。算法时代下,每个人都主动或被动地沉浸在算法生活之中。毫不夸张地说,算法俨然成为社会新的权力中间人与代言人^[9]。然而依靠算法强驱动的大数据技术在提高社会运行效率、激发科学突破及创造新的便捷且体贴的个性化服务的同时也带来了诸多挑战,如过于追求数据和建模而削弱人类的判断,由于算法偏见、歧视而带来的社会不平等以及信任危机等问题^[10]。“大数据杀熟”就是由于算法缺失伦理约束而导致的恶果之一,反过来也动摇了人们对算法科学性的信任,促使人们重新审视算法、社会与人之间的复杂关系。据此而论,“算法不应该被置于人的对立面,算法所引发的伦理问题也超越了纯粹的技术性维度,而是关涉到由编码、规范、人类实践等算法的每个部分所构成的整个集合如何行动^[11]。大数据技术背后的算法也应该且必须遵循一定的伦理规范。“算法伦理是使算法具有内在的伦理规定性,其主要目标是创建符合伦理准则的算法,在决策时能够遵循伦理准则做出正确的决定。这些准则主要包含尊重性、安全性、预防性、透明性和友好性等。”^[12]

2. “大数据杀熟”背后的平台伦理之思

当前社会,处于不同空间和时间状态的、大规模陌生人之间之所以能够分工协作,达致一种相对默契的信任关系,互联网平台起到了重要的作用,人与社会借助于它部分程度上实现了信任体系的构建与维系,而信任则是一个社会得以顺利运行的核心要素。如何诠释这一现象及其可能导致的后果?著名社会学家吉登斯所提出的“专家系统”(expert systems)和政治经济关系重构下的“象征标志”(symbolic tokens)理论在一定程度上能够解释这一现象。“‘专家系统’,它是由专家队伍组成的体系,它会因拥有特定的知识而使其影响力全方位地延伸到诸多社会关系领域并以主体的力量操纵着自然界与社会;‘象征标志’,指的是媒介在传递信息时,不考虑作为信息的发出者——个人或团体的特殊性而只在意信息本身的象征性符码,它可以分为不同的类型,比如语言、权力、货币等。”^[13]现代社会正是借助于这样复杂而又极其庞大的抽象系统来为人们提供预期的保

障,并确保人类社会的有序运行。

在现代信息社会中,人际之间的彼此信任会越来越多地被人对专业系统的信任所替代,而互联网平台很大程度上就是一种新型“专家系统”,它具有构建社会信誉体系以及为用户提供更为便捷与个性化服务的能力。互联网平台主要借助于“点评”以及“打分”的形式去评价、约束供应方和需求方之间的行为,平台还借助技术与市场的力量承担“第三方担保”与调解的角色(诸如支付宝),从而逐渐形成一种有效的信誉体系,而这对于身处陌生人社会的个体而言具有极其重要的作用,而且用户相信即使出了问题平台也可以“秉公执法”,这进一步增加了陌生人之间的信任指数。在此过程中,互联网平台也同时具备了“象征标志”的属性。理论上而言,不同年龄、肤色、文化背景、身份地位以及财富多寡的人借助于移动互联网平台所获得的信息及服务都是均等的(除非你愿意花费更多的钱去享受更为便捷、特色的服务)。互联网平台起码在理论上让每个人一视同仁地被对待和服务。

除此以外,大数据技术的发展又增加了互联网平台的第三个特征,即个性化定制服务功能。人们希望在被互联网平台平等对待和服务的同时还能够为自己提供更为私人、贴心以及差异化的服务。恰恰问题也出在这里,部分互联网平台在大数据技术上面做文章。它凭借对用户的了解和强大有效的算法,让你看到的是它想让你看到的商品和服务,让你看到的价格也是它想让你看到的价格,部分平台还会隐蔽地对不同的用户在毫不知情的前提下实施不同的定价策略,用户在互联网平台(背后是强大的专家系统)面前实质上处于弱势地位,陷入到了新的信息不对称的困境中。用户在平台面前几乎是透明的,而平台于用户而言却是深不可测的庞然大物。由此,移动互联网平台所构建起来的信任系统也逐渐受到质疑,“象征标志”特征与个性化推荐诉求在追逐超额利益、监管缺位的情况下产生了冲突,引发了诸多问题。据此,掌握大数据及其关联技术的互联网平台更应该遵循相关的伦理规范、承担必要的社会责任,尽可能平衡公正、信任、商业化与人性化的动态统一。

3. “大数据杀熟”背后的用户伦理反思

移动互联网语境下的“用户”兼具消费者和商品的双重属性:从消费者的角度而言,他们希望获

得物美价廉且便捷的商品或是服务,反观部分平台却不去努力提升用户的消费体验却利用新技术导致的信息不对称去“杀熟”,这就违背了诚信伦理的基本原则。“诚信既包含履行契约或承诺的意愿,也包含履行的能力,更包含履行的责任。如此,自然人或组织在信用活动中没有履行契约或承诺,都已违背信用的原则,属于失信行为。”^[14]市场经济条件下,人们花钱从互联网平台购买商品或服务本质上就是与平台及其相关的商家达成的一项基于自愿原则基础上的契约,互联网平台既然收取了费用就应该履行相应的义务、承担必要的责任,而诚实守信、童叟无欺则是最基本的准则。据此,作为“现实社会的延伸与发展,具有虚拟性、交互性和开放性特点”^[15]的网络社会,伴随着它对整个人类现实生活渗透和影响的与日俱增,“也必然需要遵循一定的运行规则,网络交往主体也必须遵从一定的道德规范。可以认为,网络诚信问题实际上是发生在网络空间这一场域之中的真实道德问题”^[16]。

除此以外,“大数据杀熟”还侵犯了用户的隐私,且反向充分利用用户的隐私攫取超额利润。面临不同的语境,“隐私”这个语词常被赋予不同的价值内涵,如安全、尊严、私人、自主等。较为全面且具代表性的是贝特·罗斯勒(Beate Rossler)的理论,罗斯勒把隐私主要区分为“信息隐私”“空间隐私”和“决定的隐私”^[17]。据此,不同的隐私形态关涉到不同的利益诉求与隐私关切:“信息隐私”强调对于个人信息的自主控制与分配(被谁获知以及以何种方式被获知);“空间隐私”意味着个体活动空间、场域不被排他性地观察和侵扰;“决定的隐私”指向私人事务和自主决定的自由。

大数据技术将我们转化为各种“可视性”和“透明性”的“数字人”,通过数据的持续挖掘与分析对我们进行“数据标注”,比如将我们以潜在信用风险人、潜在消费者、保险意向人等信息实体方式呈现出来。这种依靠数据统计结果来预测与迎合个体推测性偏好的个性化服务或者预先帮我们规避可能性风险的交易方式的确为我们带来了诸多的便利。但是,问题恰恰也出在这里。数据的来源有的是消费者知情且同意授权传输的前提下收集到的,还有的则是根本不知情前提下偷偷地被上传和收集收据,而且这是一种全天候、不间断地收集个人的信息以及相关痕迹的过程。在此过程中,部分互联网平台及公司有意混淆、模糊个人

数据的界限与用途,不管同意与否,也不论是否会对用户造成伤害与困扰,想方设法、不择手段地去收集用户的各种数据,然后进一步整理、挖掘数据背后的各种信息关联与行为轨迹,最后通过信息倒卖或是“大数据杀熟”的方式去攫取超额利润,根本就没有保护用户数据隐私权的意愿和行动,而为之所付出的微乎其微的代价进一步助长了这种行为。

除了侵犯用户的“信息隐私”外,“大数据杀熟”还涉及对用户“空间隐私”的干扰与侵犯,而这就需要从数据的来源出发去考虑问题。用户的智能手机及其相关联的可穿戴设备就是一个个全天候、可移动的信息储存器与发射器,移动互联网“连接一切”“即时反馈”的特性以及基于LBS(基于位置服务)在若干传感器的支撑下,每个人的运动轨迹、行为生活习惯(很多行为的发生需要借助手机去完成)、个性爱好等都可以随时随地被标记下来,然后以主动或被动的方式上传到云端。大数据就好比人类社会的“第三只眼”一般,时刻关注或监视着你的一举一动,这也在一定程度上让你逃无可逃,从而面临使人类的生活进入“超级全景监狱”的困境^[18],进而对个人的空间隐私造成一定程度的压缩和侵犯,长期居于如此场域中的人们也更容易感到压抑和焦虑。

除此以外,大数据所蕴含的数据挖掘、推荐及预测能力会在一定程度上替我们作决定或暗示、诱导我们应该如何抉择以期提升效率、达致我们想要的结果。长此以往,大数据技术就会导致个体失去自主选择与决定的能力与必要性,而人自身所隐含的贪图安逸、追求舒适与效率的特性也容易将更多的思考行为与独立作决定的能力逐渐交由大数据替我们作决策,进而弱化我们的反思批判能力,从而涉及到个体“决定的隐私”方面。本应作为自由意志而存在,具备独立的思考、反思、批判精神的主体能动性意义上的人的含金量其实在降低,“大数据杀熟”其实就是技术在一定程度上奴役人们的异化行为,而这在本质上也是对人的尊严以及终极存在价值的挑战 and 伤害。

三、启示及应对之策:负责任 创新与多方制衡

“大数据杀熟”的频频出现促使我们去重新反

思和评判大数据技术对人类社会以及人自身发展的综合影响。较为合理、稳健的方式是在理性分析与反思批判的基础上提出具有建设性价值且较为务实的意见或方案。至少,我们可以从以下几个层面出发尝试着提出一些方案。

1. 国家应该从顶层设计的层面推动大数据研发、交易以及保护等相关法案的出台

党的十九大明确提出“推动互联网、大数据、人工智能和实体经济深度融合”。2017年12月8日,中共中央政治局就实施国家大数据战略进行第二次集体学习,习近平总书记表示:要推动实施国家大数据战略,加快完善数字基础设施,推进数据资源整合和开放共享,保障数据安全,加快建设数字中国,更好服务我国经济社会发展和人民生活改善。这进一步说明国家支持发展大数据技术的决心,同时也意识到发展过程中可能会出现的问题,希望能够更好地促进大数据技术的健康发展。国家应该且有义务和能力召集相关企业、互联网平台的负责人、工程师、法律专家、人文学者以及群众代表等共同协商,加速推进大数据研发与保护方面的立法进程。各方在国家的统筹协调下,通过不断地磋商、博弈尽快出台大数据保护与交易方面的法案及惩处标准,成立相应的执法部门或机构,起码可以使大数据的发展有法可依、有法可行。

2. 互联网平台及相关企业要以负责任创新的态度去推进大数据技术的发展与应用

需要注意的是,在创新的过程中这些公司与平台也要遵循必要的伦理规范、担负起相应的社会责任。即,本着负责任创新的态度谨慎地对待和开发利用新技术。“‘负责任创新’,即负责任地研究与创新,指的是一个透明的、互动的创新过程,在此过程中社会行动者和创新者彼此合作和呼应,充分考量创新过程和其适销产品的道德可接受性、可持续性和社会期许性,从而使科技进步适当融入社会生活。负责任创新强调,创新中任何主体的不负责行为都将对创新生态系统产生整体影响,所以不同的创新参与者应协同工作并共同承担创新责任,进而能够实现责任共享,即科学、决策与社会之间的责任共享。”^[19] 据此,互联网平台及关联企业应该充分考虑发展大数据技术所应遵循的伦理规范与社会责任,坚持“善法”优先于“算法”,在恪守职业道德与信任伦理的基础上,不断提升自己的专业性、创新性、批判性与反

思性,进而在追求人文关怀的愿景下获得与之相匹配的利益回报、品牌信誉度与社会认同感。

3. 算法工程师、科研工作者群体也要以负责任创新的态度审慎地对待大数据技术

毕竟技术背后的人才价值观、伦理观以及是非观的主导者和塑造者。具体而论,主要涉及到如下五个层面:第一,工程师在设计应用程序的时候应尽量遵循“奥卡姆剃刀原则”(如无必要,勿增实体),即尽可能少地收集用户的数据。如果的确需要收集相关的数据,应事先征得用户的同意与授权。在用户知情且同意的基础上为他们提供更为多元化的选择,视用户对个人隐私信息重视及授权程度的不同或支付金额的差异而相应地提供不同层级与权限的服务。对于确实会伤害到用户个人隐私、经济利益、人格尊严等合法权益的越轨行为要坚决禁止,这是大数据技术发展需要遵循的“底线伦理原则”。第二,对设计主体进行道德想象力方面的注入与建构。“设计主体对算法的形成起着关键作用,为了减少或消除算法不确定性所带来的风险,可以对设计主体进行道德想象力的建构。‘道德想象力’是一种能力,通过这种能力,我们可以在一种既定情景里发现采取多种行动的可能性,以及预想一个既定行为可能带来的利害结果。”^[12] 道德想象力促使设计主体对算法可能造成的社会后果事先进行相应地评估与预警,不断反思某种算法是否真的有利于社会公共福祉的实现,从而提醒人们追求商业与伦理、技术发展与人文关怀兼顾的良好局面,从而避免陷入“技术中心主义”“利益中心主义”等极端异化的思想泥淖中无法自拔。第三,注重价值敏感设计。价值敏感设计侧重于将算法的具体使用场景与伦理设计有机结合起来,增进抽象算法技术的具体性与场景性,从而将算法的伦理设计建立在具体情境分析的基础上,避免使其过于抽象化和空心化。第四,善于利用“数据脱敏”技术来保护用户的敏感信息与个人隐私。根据全球著名IT公司Gartner(高德纳咨询)所定义的,“数据脱敏(data masking)又称数据混淆、数据漂白、数据去隐私化,它是一种为用户提供虚假数据而非真实数据、防止敏感数据滥用的技术,包括静态脱敏(在非生产数据库中防止静态数据的滥用)和动态脱敏(生产数据库中传输数据的脱敏)”。较之于数据加密技术,数据脱敏技术更好地兼顾了数据利用效率与隐私保护之间的动态平衡。第五,善于运用区

区块链等新兴技术反向追溯与认定“大数据杀熟”行为。鉴于“大数据杀熟”的隐蔽性和复杂性,我们要善于运用最新的技术手段去证实和治理此类难题。根据百度百科的定义,“区块链(blockchain)是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式,它本质上是一个去中心化的数据库,是一种按照时间顺序将数据区块以顺序相连的方式组合成的一种链式数据结构,并以密码学方式保证的不可篡改和不可伪造的分布式账本”。区块链具有可溯源与可追踪的特性。针对“大数据杀熟”的现象,通过区块链技术追根溯源从而对此行为进行认定与裁决,可有效地减少在此问题上的纠纷,对用户而言可以更好地保障他们的合法权益,对互联网平台与企业来说,也可以证明自身的清白、约束自身有可能滥用数据的行为,对国家执法部门而言也可以更透明、更高效地进行数据判决与执法。

4. 用户自身也需要培养独立思考与批判的能力

人们也需要深刻意识到对大数据技术的过分依赖有可能带来丧失主体能动性、陷入“认知茧房”^①、被大数据奴役的不自由状态,“大数据杀熟”就是恶果之一。据此,我们应该有意识地培养自身的反思意识与批判能力,审慎地看待大数据技术在人类社会发展中的作用与价值。除此以外,我们还要注重线上与线下、真实世界与虚拟世界间的融合与平衡。如果可以的话,尽量多尝试着使用不同互联网平台或公司的应用程序、产品或服务,遇到需要授权个人信息的时候多问自己几个为什么,适当地、有意识地与网络世界保持一定的距离。

四、仍待商榷且需持续推进的若干议题

问题到此并未结束,现实远比我们设想的要复杂,至少以下几个议题还需要社会各方力量去继续探讨和辨析,以期在增加社会总体福利的前

提下达成相对意义上的“重叠共识”^②。

1. 数据的所有权如何界定

数据所有权的界定不仅事关相关主体的切身利益,更涉及到与之相关的责任与义务的认定与分配。多数人会比较认可用户拥有对数据的所有权和决定权,这种观点固然有其合理之处,但亦有失偏颇。个体自身所产生的数据是原始的、碎片化的,若想要最大化数据的综合价值,还需对其进行标注、编码、存储,然后是利用算法分析、挖掘其背后所隐藏的巨大价值。此外,为了防止数据泄露、保护数据安全,互联网平台与公司仍要投入巨大的资源进行保护。据此,公允而论,互联网平台以及相关企业也有权拥有对数据的所有权与使用权,主要是基于两个原因:一个是它们在大数据技术的发展过程中投入了巨量的资源且承担了相应的风险;另一个则是从运行效率和增进社会总体福利的角度而言,平台与公司的集中开发与创新远远胜于若干分散个体的力量。因此,我们需要适当更新自己的观念,从法理上承认用户和互联网平台及相关公司都拥有对数据的所有权和使用权,在此基础上对他们进行相关义务与责任的界定。需商榷之处在于互联网平台、用户就数据的所有权与使用权如何进行相关比例的分配,用户的哪些数据隐私属于完全由自身决定而不可让渡给互联网平台及公司仍需探究。各方主体如何就大数据技术带来的实际利益进行相应地分割也值得商榷。互联网平台与公司利益独享,却几乎不承担任何代价与损失的模式肯定是不健康的。

2. 政府与市场在大数据治理的过程中如何分工与协作

我们要警惕那种一遇到事情或问题就往政府那里推,并希望政府出手化解难题的“懒人式思维”。国家与政府更多应该做的是从顶层设计的维度召集各方代表就大数据及其发展过程中遇到的困惑展开多层次、深度的沟通与协调,善于包容、鼓励和支持创新技术的发展与试错,制定相关

① 大数据在充分了解个体的基础上会进一步为个体推荐他所感兴趣或使其感到愉悦或舒适的东西,忽略掉其他的可能性,久而久之,个体会将自身桎梏于像蚕茧一般的“茧房”中。

② “重叠共识”(overlapping consensus)是罗尔斯在其代表作《正义论》中提出来的一个概念。在书里,罗尔斯认为,尽管公民对正义这个概念的理解有许多差异,但这些不同的政治观念有可能导致相似的政治判断,即,不同的价值观念体系在各自保持自身的前提下就政治生活的基本秩序和原则所形成的共识,不同的前提有可能导致同一个结论。此处借用这一概念是想就探讨如何在应对“大数据杀熟”及隐私保护等议题方面各方力量和参与者能够达成一些相对意义上的可被接受的共识,而非自说自话、陷入“零和博弈”的糟糕困境中。

的法律条文且认真执行,切实做好“服务者”与“裁判员”的角色。“大数据杀熟”绝非普遍现象,我们还是应该尊重市场发展规律,推动建立数据产权交易与仲裁中心等机构。在政府这个裁判员合理就位的大前提下,我们要善于利用市场、技术(比如区块链技术、脱敏技术等)以及媒体的力量去监督和规范大数据技术的发展,形成多方博弈、共治的良好局面。

3. 算法公开与透明、数据共享等愿望是否可能仍待商榷

部分观点主张大数据背后的算法应该公开、透明,事实上想法是好的,但是却不太现实。伴随着数据成为互联网领域的核心资源,数据背后的算法更是各个互联网平台与公司的核心竞争力与命根子,企业肯定是不愿意公开的。除此以外,算法本身还涉及到知识产权保护的问题,我们也应该尊重企业的合法权益与诉求。因此,算法应该如何公开,公开到一个什么样的程度,如何保护企业的相关合法权益以及涉及公共福祉的数据与隐私保护之间如何协调、破除“数据孤岛”实现数据共享等议题仍待全社会去讨论与辨析。

4. 监管与创新、大企业中小企业如何和谐共存仍待探究

如何在对互联网平台及其相关企业进行合理监管的同时,也能够给它们足够的创新与试错的空间。监管与创新二者间如何动态平衡仍值得探究。除此以外,鉴于互联网行业“马太效应”(集群效应尤为明显)的格外突出,在对互联网平台及相关企业进行合理监管的同时,大企业、互联网巨头拥有足够的经济实力、行业经验以及人脉资源等足以化解由此带来的诸多风险与投入,中小企业可能会面临巨大的监管成本压力与风险。如何适当地兼顾中小企业的发展诉求与面临的多重挑战之间的关系也值得我们商榷,尽可能避免那种巨头企业凭借监管的名义去“合法地”绞杀中小企业的现象发生,而是为它们的发展预留一定的空间。

参考文献:

- [1] 黄欣荣. 大数据技术的伦理反思[J]. 新疆师范大学学报(哲学社会科学版), 2015, 36(3): 46.
- [2] 岳缙. 大数据技术的道德意义与伦理挑战[J]. 马克思主义与现实, 2016(5): 91.
- [3] 黄欣荣. 大数据时代的思维变革[J]. 重庆理工大学学报(社会科学版), 2014, 28(5): 13-18.
- [4] 维克托·迈尔-舍恩伯格, 肯尼思·库克耶. 大数据时代: 生活、工作与思维的大变革[M]. 盛杨燕, 周涛, 译. 杭州: 浙江人民出版社, 2014: 71-72.
- [5] 刘智慧, 张泉灵. 大数据技术研究综述[M]. 浙江大学学报(工学版), 2014, 48(6): 957-972.
- [6] 朱杰. 网络大数据环境下价格歧视行为研究[J]. 哈尔滨师范大学社会科学学报, 2018(2): 65.
- [7] 郭胜. 大数据技术的伦理问题反思[J]. 科技传播, 2018(10): 5.
- [8] 唐凯麟. 大数据隐私伦理问题研究[J]. 伦理学研究, 2016(6): 102-106.
- [9] Kitchin R. Thinking Critically About and Researching Algorithms[J]. Information & Society, 2017, 20(1): 14-29.
- [10] Rainie L, Anderson J. Code-dependent: Pros and Cons of the Algorithm Age [R]. Washington, D. C.: Pew Research Center, 2017.
- [11] 刘培, 池忠军. 算法的伦理问题及其解决进路[J]. 东北大学学报(社会科学版), 2019, 21(2): 119.
- [12] 郭林生. 论算法伦理[J]. 华中科技大学学报(社会科学版), 2018, 32(2): 42-43.
- [13] 岳天明. 社会困局与个体焦虑: 吉登斯的现代性思想[J]. 西北师大学报(社会科学版), 2013, 50(5): 99.
- [14] 谢志远, 王明高. 论诚信伦理构建的制度保障与创新[J]. 伦理学研究, 2016(6): 60.
- [15] 刘守旗. 网络德育: 21世纪的德育革命[J]. 南京师大学报(社会科学版), 2003(6): 70.
- [16] 崔斌. 论网络诚信的伦理本质及其建设方略[J]. 道德与文明, 2016(2): 130.
- [17] Rossler B. The Value of Privacy[M]. tran. Glasgow R D V. Cambridge: Polity Press, 2005: 6-10.
- [18] 马克·波斯特. 第二媒介时代[M]. 范静哗, 译. 南京: 南京大学出版社, 2001: 98.
- [19] 薛桂波, 赵一秀. 基于“负责任创新”的欧盟科技政策转型及启示[J]. 中国科技论坛, 2017(4): 172-173.

(责任编辑: 李新根)