从大数据杀熟到隐私泄露:软硬件视角下隐私问题的伦理分析与思考

王昱兴 袁 博

摘 要:大数据和人工智能技术的出现使得信息时代下的隐私保护问题更加尖锐和复杂。过去几年里,出现了一批以大数据杀熟和基于智能平台隐私泄露为代表的新型隐私问题,给政府、企业和公众各方都带来了不可忽视的负面影响。数字时代信息技术的更迭能力远远超过法律条例和规范的修订速度,需要用理性和建设性的眼光重新审视。站在信息技术的视角下,对大数据隐私保护问题的困境进行解析与反思,提出相应解决方案。

关键词:大数据;人工智能;隐私保护;数据伦理
中图分类号:G202 **文献标示码:**A **文章编号:**1006-2815(2021)03-0072-10
DOI: 10.19946/j.issn.1006-2815.2021.03.006

以大数据和人工智能为代表的新兴信息技术不断改变着人类社会的面貌。技术的更新换代使得社交媒体、电子商务、智慧城市、量化金融和医疗健康等领域进入了发展快车道,智能设备和软件应用为人们带来方便与快捷的同时,数据隐私问题也愈发严重。2020年12月,广东省深圳市第六届人大常委会审议了《深圳经济特区数据暂行条例(草案)》。作为我国第一部数据领域的综合性专门立法,该条例首次提出了"数据权益保护"的概念,并规定收集、处理涉及隐私的个人数据须得到明示同意。法律条例的出台虽然能够在宏观上对违规收集个人数据的行为进行限制,但在个人隐私保护方面仍存在不小的难度。即使消费者能够主动选择将哪些显式的个人信息交给网络公司,但无法控制公司如何利用这些信息。数字时代信息技术的更迭能力远远超过规则、法律条例和规范的修订速度,个人隐私数据需要的安全性、机密性和完整性对数据的保护和存储都提出了全新的挑战。数据伦理问题是进入数据城市、数

收稿日期:2021-02-04
作者简介:王昱兴,清华大学深圳国际研究生院硕士研究生,主要从事数据挖掘和演化计算研究。E-mail:wyx20@mails.tsinghua.edu.cn
袁博,清华大学深圳国际研究生院,副研究员,博导,主要从事强化学习和智能决策研究。E-mail:boyuan@ieee.org

据社会无法回避的重要环节，且随着信息技术的进步而愈加迫切。

一、隐私问题的新困境

隐私伦理问题研究仅有百余年历史。1890年，美国律师沃伦（Samuel D. Warren，1852-1910）和布兰迪斯（Louis D. Brandeis，1856-1941）于《哈佛法学评论》上发表的《论隐私权》一文中首次提出“隐私权”的概念，指出隐私权是人们享受独处的权利[1]223。而早在人工智能和大数据的概念被提出之前，美国科学家、控制论的创始人维纳（Norbert Wiener，1894-1964）就曾经担心自动化技术可能会给人类带来一系列的道德伦理问题[2]134。尼森鲍姆（Helen Nissenbaum，1954-）在《信息时代的隐私保护：公共场所的隐私问题》一文中抛弃了公私领域的二分法，提出了语义完整性理论，为公共空间的隐私问题提出了分析框架。她认为机构组织或个人在传播与利用他人信息时，都应该遵循语义完整性原则，保持信息与原语境的一致性，遵守适当性规范与流动分布的规范[3]78。大数据之父，牛津大学教授舍恩伯格（Viktor Mayer-Schönberger，1966-）提出了数据化模型，认为大数据技术能够将许多抽象概念渲染成为数据，而海量数据的产生也将导致更多的隐私信息被泄露，政府和企业也可能会被数据的虚假魅力所迷惑[4]1856。近些年来，国内的相关学者也将研究重点放在了人工智能和大数据伦理方面。段伟文研究了大数据时代的隐私权问题，他认为大规模的个人数据收集、分析和共享对个人隐私权益带来了诸多影响，信息时代下对隐私的保护应该建立于对个人数据的规制之上，亟须为新技术的滥用设定道德伦理界限[5]95。刘晓力站在认知科学的角度，对人工智能的发展现状和未来前景进行了分析与展望，她认为基于大算力和大数据驱动的人工智能技术虽然卓有成效，但是深度学习的“黑箱”问题会带来有违伦理的反常现象，而当前人工智能依然面临着语义、物理和情感落地的难题，应为人工智能建构合理的认知架构，使之成为可信赖的道德主体[6]25。闫坤如探究了人工智能“合乎伦理设计”的理论来源，提出在设计人工智能算法时，应将正确的道德规范与具体的技术标准和设计环节相结合，坚持公平、可靠、安全的设计原则，以算法的透明化为基础，保护人们的隐私不受侵犯[7]15。

从以上文献可以看出，学者们在研究隐私问题的过程中，站在伦理学的角度对其定义、范围和社会影响展开研究与探索，同时也关注到了大数据技术本身产生的一系列冲击，逐渐形成了数据伦理的研究分支。数据伦理研究的两个基本维度是由“伦理”和“技术”所构成的[8]145，而技术维度常被人忽视，有必要强调并进一步探索。信息时代的隐私问题正面临着新的困境，人工智能和大数据技术的迅速发展以及手机软件和智能物联网设备的普及带来了数据风暴，虽然我国陆续出台了一系列法律条文，但仍然难以跟上技术更迭的步伐，无法及时对新技术造成的隐私问题加以限制。与传统的隐私问题相比，一方面，用户的隐私数据被分散到了世界各地的服务器中，数据管理的主体并不明确，隐私数据的边界也在不断演变，在过去看似安全的数据，现在也可能被用来窥探个人的偏好[9]5803。另一方面，正如尼森鲍姆所说，隐私的哲学和法律理论早就承认了隐私与个人信息之间的关系，但几乎没有为处理公共隐私问题提供一个明确的辩护框架[10]560。法律条文的制定和实施有滞后性，且涵盖有限。除此之外，

隐私、隐私权、数据权益等抽象概念的规范与实践中层出不穷的问题间如何相互界定等，在理论界也是难题，这些都成为了解决信息时代隐私保护问题的绊脚石。因此，不妨转换视角，通过分析隐私问题的技术本质和逻辑，探索大数据技术为何会引发伦理问题，并寻找可能的解决方案。

（一）软件与隐私问题：大数据杀熟

“杀熟”指的是当用户频繁购买某个商品并成为“熟客”后，会对商品产生一定的依赖性，商家便可以利用这种依赖性进行差异化定价，从而赚取额外的利益[11]8。在信息时代，大数据杀熟的对象一般为忠诚客户即频繁在同一平台上进行重复消费行为的老客户，他们与经营者之间往往是“熟而不识”。杀熟现象在基于互联网平台的机票、酒店、电影、电商、出行等消费领域多有存在，广大消费者即使知悉，往往也无计可施。那么，为什么大数据能够“神不知鬼不觉”地做到杀熟？

实际上，这与人们消费模式的转变密切相关。与传统的线下消费不同，在非面对面的虚拟环境下，线上商家可以知晓顾客的个人背景、账号信息、消费记录，进而研究出消费习惯和偏好等信息。用户在网络上的一举一动，商家都能够通过互联网技术和大数据技术以较低的成本进行获取。对商家来说，每一位顾客几乎是透明的，因此在信息不对称的情况下，可以采取差异化定价，而伤害的往往是忠诚度相对较高的线上消费群体。

大数据杀熟具有明显的技术依赖性，用户画像技术是其中的典型代表[12]251。这种技术实际上是一种标签化的用户模型，而标签是基于用户行为数据的一种高度抽象化的表示。构建用户画像时所使用的标签类别可以根据实际业务需求和应用场景划分为不同的主题和粒度，建立用户画像的过程实际上是一个从粗粒度到细粒度的过程，例如一级分类中的“行为标签”可以进一步划分为细粒度的上网习惯、互动行为、购买行为等，而各种行为可进一步细化为上网时段、日收藏数、订阅产品等更加细粒度的标签。用户画像设计的初衷是为用户提供“千人千面”的个性化服务，如推荐可能感兴趣的产品与服务，进而提升用户体验和用户的价值。但如果平台在为用户提供服务的过程中利用用户的使用习惯、价格敏感度以及依赖性等进行差异化的定价以谋求额外的利润，则可被认为涉嫌存在大数据杀熟的不当行为。例如，在出行行业，通过分析消费者的打车记录可以发现其固定的上下班路线（高依赖性），因此商家可能据此对这些线路的打车费用进行上浮。同时，当商家通过分析用户的浏览、收藏和购买记录认定某个用户为价格不敏感的用户时，也有可能刻意减小其在购买商品时所享受的优惠幅度。值得注意的是，商家用来对用户进行分析的算法及规则等通常并不公开，影响其定价策略的因素也多种多样，对老客户的杀熟行为与对新客户的优惠活动之间的边界也存在一定的模糊性，这些都为大数据杀熟行为的认定造成了挑战。

从大数据杀熟这一反常现象可以看出，数据的价值在于利用各种算法对其进行分析以产生更高的附加价值。谷歌、百度、Facebook等公司开发的网站会在用户使用时收集相关的设备识别码、历史记录和手机通讯录等信息，并将其用于大数据分析，从而为用户精准地推送相关服务[13]60。对于公司自身来说，用户使用产品时生成的行为数据，能够为公司的商业战略、人员和业务调整提供有力的参考甚至带来更高的利润。但是，算法的加持并不一定会带

来正面作用,对大数据的错误分析可能会导致严重的后果,不仅会对消费者的隐私造成侵犯,甚至会采取适得其反的措施,以至于公司利益受损,可谓“双刃剑”。

数据伦理问题的存在,不是某一方单独造成的结果,与软件研发和应用过程中各方群体的诉求和数据控制力的不均衡密切相关(表1)。

1.公司决策层

负责政策制定和软件的顶层设计,他们接触的并不是海量的用户数据,而是经过软件开发人员分析过后的抽象数据,并根据分析结果调整软件的功能模块以及业务类型。作为软件工程的顶端,公司决策层的目的是提高公司收益与扩大影响力,能够在激烈的企业竞争中胜出。然而,过度信任大数据分析的结果可能导致公司做出适得其反的措施导致利益受损,为了挽回损失,一些公司会不择手段地对用户数据进行“暗箱操作”,从而损害消费者的隐私权益。

2.软件开发者

他们处于软件工程的核心位置,掌握着海量的用户数据,通过各种算法和数据结构来实现应用功能。用户数据经由物联网传感器、网络视频流、点击流等方式上传到网络平台并储存在了分布式数据库系统(DDBS)里,电子数据库赋予了这些记录永久性、延展性和转移性的特点。然而,隐私法规并没有明确统一的数据储存规范,允许软件收集哪种数据,这为灰色区域和不确定性留出了空间,而这些区域和不确定性是无法通过法律手段解决的。因此软件开发人员可以自由定制数据结构,便于算法分析。而大数据分析算法的性能又由数据的规模和质量同时决定。训练数据分布的不平衡或者带有偏见的人为设计会导致算法训练过程中的偏差,从而使系统放大这种偏差,导致算法歧视现象的发生。除了原始数据和人为干预,数据抽样带来偏差以及较差的模型参数设置,甚至使用算法本身也会导致算法歧视[14]18。另外,一个储存了海量用户信息的数据库,其认证、访问、保密和安全措施是防止数据泄露最主要的屏障,如果被黑客非法获取数据库的访问修改权限,后果将不堪设想。

表1　软件研发和应用过程中参与者的行为分析

参与者	主要职能	目的	数据类型	主要问题
公司决策层	政策制定、软件功能设计	提高公司收益与影响力	分析过后的抽象数据	对大数据的盲目依赖、决策失误
软件开发者	算法研发、软件功能实现	提高用户满意度	海量用户数据	算法偏见、数据安全
用户	软件使用、意见反馈	满足生活娱乐所需	原始行为数据	隐私意识缺乏、隐私泄露

3.用户

作为软件的使用者,其下载的目的是满足生活娱乐所需。作为数据的产生者,在使用过程中,软件会记录用户的行为数据,按照预先设定好的结构进行封装,并通过互联网传输,最终储存在数据库中。这看似双赢,实则暗藏玄机。在使用软件之前,用户会进行隐私协议的授权,虽然用户享有这些原始数据的所有权并且同意软件收集和分享自己的行为数据,但他们无从知晓数据是何时何地以怎样的形式被组织和上传。并且,为了能够使用软件,用户不得不持续接受这种被动的上传。对于缺乏隐私保护意识的用户来说,这无疑是将自己时刻暴

露在风险之下。

(二)智能物联网设备与隐私问题:数据泄露

无线传输和编解码技术的广泛应用为物联网的发展奠定了坚实基础。物联网是一个物理对象的网络,设备之间可以通过接口和网络资源进行数据共享,而大量的网络节点构成了一个复杂的网络系统。它不仅使设备能够在封闭的孤岛内进行通信,而且还可以跨越不同类型的网络进行通信,形成一个更加连通的世界。这种连通性在拉近人与人之间距离的同时,数据泄露问题也愈演愈烈。2018年4月,一些不法分子通过非法途径破解了安装在家庭或酒店宾馆内部的摄像头,盗取其中的私密视频并在网络上公开贩卖[15]47。2019年6月,有消息称某品牌的智能家庭音箱在无人操作的情况下将用户谈话的录音文件发给了通讯录中的联系人[16]2113。在黑帽安全技术大会上,专家也表示可以通过技术手段非法访问心脏起搏器、无线胰岛素泵和其他设备的数据[17]2133。物联网设备的基数庞大,并且增长迅速,加上芯片制造技术的迅猛发展,其集成度也越来越高,更多的产品同时配备了摄像头、麦克风、加速度传感器和人脸识别系统等模组,而且各个部件之间相互协作,很难将某个模块分离出来;并且物联网设备制造商往往不会专门向消费者详细说明用户数据是如何被收集、上传和使用的传输协议等信息;设备在开启后也会不间断地在后台运行,缺乏专门为用户设置的可用来查看或更改其隐私设置的界面,而且许多产品的隐私政策很难被找到。这些都为数据泄露埋下了隐患。

虽然各大智能硬件厂商几乎都在努力进行技术革新,提高数据加密的可靠性和传输过程的安全性,并以此作为产品的卖点,但是隐私泄露问题的本质实际上是如何建立一个长期稳固的人与人、人与物和物与物的信任体系,这里所说的物代表物联网对象,包括传输网络、硬件实体、云服务端等。万物互联应基于一个良好的信任体系,这样不同系统之间共享的数据才会越来越多。如果数据来源于不受信任的设备,或者本身是不受信任的,那么数据收集和分析工作是没有意义的。使用或接入未受信任的物联网设备越多,隐私泄露的风险也就越大。因此,有必要对信任体系中的三种交互方式进行分析,如图1所示。

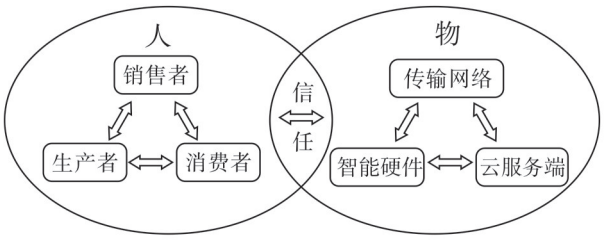

图1 物联网信任体系

1.人与人

在物联网的世界里,每个人都拥有自己的数字身份,基于智能设备的信息采集和基于网络的信息传输,生产者、销售者和消费者之间形成了新的互动模式,表面上传输的物联网数据,实际上隐藏了更深层次的人际关系[18]133。2017年的“360水滴直播事件”和2020年的韩国“N号房”事件再次引发了公众对网络摄像机私密性和安全性的巨大担忧,以及对物联网设备制造商的信任危机。而事件的背后,是人与人之间不符合道德规范的交际互动,而这些互动,是借由物联网设备完成的。一个互联的、和谐的社会关系网,应该建立在人与人之间相互信任,遵循道德规范的基础之上,这样才能够充分发挥物联网开

放共享的优势。

2.人与物

信息技术的发展让我们越来越多地生活在公共空间下，遍布各地的物联网设备在提供各种服务的同时也在监视着我们的生活，正如英国哲学家边沁（Jeremy Bentham，1748-1832）提出的“全景监狱”，它是一种圆形建筑，监狱管理者能够在瞭望塔中观察监狱内所有囚犯的行为，而监舍中的犯人无法看到监视人员[19]35。在物联网时代，智能手机、网络摄像机、智能音箱和无人机等硬件设备提供了记录和共享位置、声音、图像的能力，不法分子可以通过它们来“监视”我们的日常生活和模式，而且这种行为是难以被发现的。不幸的是，对于大部分人来说，物联网仍然是一个相对较新的技术，如果用户缺乏对物联网充分的理解和信任，对个人隐私保护的认识不足，错误地使用物联网设备，不仅会处于被监控的危险之中，而且还会影响到所有有意或无意连接到该设备的人。

3.物与物

传统互联网的核心是中心化的服务器以及数据库，随着物联网的产生，基础设施、网络流量和用户需求呈现出爆炸式增长的趋势，各种物联网应用也逐渐具有去中心化和泛在性的特点。设备与设备之间通过数据传输进行访问和信息共享，它们的信任机制往往是由人工设定的，通过各种加密算法、数字证书等建立信任关系。比如家庭网关，作为智能家居控制的枢纽负责几乎所有家庭设备的接入、认证与交互，如果被黑客非法修改网关的认证机制，从而获得所有设备的访问和控制权限，后果将十分严重。

（三）大数据隐私保护中的思考

大数据技术的内涵十分广泛，不仅包括支持数据采集、传输和存储的硬件环境，还包括对数据进行分析和处理的软件与算法。作为一种计算技术而言，大数据技术本身并不具有天然的好与坏的属性，然而技术在使用过后造成的一系列后果却可能被贴上善与恶的标签，这实际上是由人们对大数据的不合理运用和缺乏完善的伦理准则导致的。

阿格里（Philip E. Agre）提出的“监视模型”和“捕获模型”以及迈（Jens-Erik Mai）倡导的“数据化模型”都能够从不同的角度对数据隐私进行分析。传统的监视模型认为监视是隐蔽的，但并非具有破坏性，它是一种有意识的策划（可能带有某种政治性质）[20]103。从隐私监视模型的角度出发，在互联网场景下，虽然手机APP的使用者在某种程度上受到了网络公司的“监视”，但这并不是数据收集这一行为带来的，APP所收集的数据仅仅是对现实的反映（用户的操作行为）或理性的表现（用户的某种偏好）。捕获模型的思想来源于计算机系统，它强调的是从用户端获取何种类型的数据并关注这些数据以什么形式进行组织以及试图从这些数据中抽取怎样的知识，它的目的是重建用户的行为[21]245。从捕获模型的角度出发，APP捕获用户数据的种类越多越详细，捕获持续的时间越长，就越有可能重建和分析用户的行为活动。以上两种模型都关注数据的收集并且认为隐私保护是对通过监视活动或捕获活动收集到的数据进行控制的一种手段。与前面两种隐私分析模型不同的是，数据化模型认为数据是一种环境状态的代表（时间、地点、任务、事件等），它更加关注于用户信息被获取之后的事情并假设用户数据已经被收集和积累，甚至被盗取和贩卖[22]195，而隐私问题则演变成了利用

现有数据构建了什么样的新知识。数据化模型强调基于搜集到的数据对人进行分析的能力，而数据分析在人和各种活动之间建立了关联性，虽然个体无法控制自己在数字市场上留下的个人信息，但是这些信息被处理和提炼的过程属于隐私利益的范畴。

此外，网络平台在处理用户在网络上进行各种活动所留下的痕迹时，如发布的留言、图片及视频，也会存在隐私问题。例如，当人们使用一款手机软件时，往往会在首次运行时进行个人隐私服务的授权，此时软件展示给用户的通常只是简略的隐私条款，而详细的隐私说明则隐藏在超链接中。一方面，大多数人会在没有经过深思熟虑、没有完整阅读或完全理解隐私政策和协议的情况下就同意提供个人信息。另一方面，这些隐私协议会使用相当冗长并且艰深晦涩的法律语言，使得普通用户根本没有意愿去仔细阅读和理解。因此，从表面上看用户是“自愿”签订了隐私协议并且明确同意了软件在日常活动中收集使用者个人信息。在此基础之上，用户是否有权阻止这些个人隐私信息被网络平台应用于数据预测和分析是值得深入探讨的。

从另一个角度来看，随着使用方式的不断演变，数据的价值已经发生了本质上的变化。信息时代的数字化进程正将我们身处的真实世界转变为虚拟的数字化世界，而日新月异的技术使得更多的信息能够被产生、存储和更快地被加工、分析和利用。在传统的数据—信息—知识—智慧的金字塔中数据处于最底层[23]1788，被视为一种原始的符号标识，需要通过一系列的处理过程转化为有意义的信息并被进一步提炼成知识，最终形成可以帮助人们改变社会的智慧。在大数据时代，利用先进的机器学习算法，可以以一种相对统一的范式直接从数据中提取知识，极大地简化了中间环节和对人工干预的需求。因此，大数据伦理问题的重点并不在于手机应用或网站是否收集了个人信息，而是在于后续的数据分析环节，即允许从数据中寻找什么、对数据提出哪些问题以及在什么层面上利用这些数据进行预测等。

同时，作为大数据体系中一个重要的环节，物联网设备承担了数据的感知、封装、传输等核心工作，如果设备被不法分子非法劫持、访问和篡改将导致严重的社会问题。然而设备本身是无罪的，在采集数据的过程中，实体可能会将数据通过网络传输给另一个实体，数据的使用权也随之转移，这为泄露源头的追溯带来了麻烦：一方面，在视频数据的采集和传输过程中可能存在技术薄弱环节，给不法分子以可乘之机；其次，相关主体在信息的监管方面也可能存在漏洞，缺乏良好的内容审核机制，导致不法信息内容得以在平台传播。因此，对于设备厂商来说，需要从技术层面和监管制度两个层面对潜在的不良使用者施加更强的约束力和监管力。

二、我国大数据隐私问题治理的对策建议

隐私问题的伦理边界具有复杂性和模糊性的特点，这要求在发展信息技术的过程中高度重视其可能带来的数据隐私问题，提前规划和制定应对措施，防患于未然。

（一）持续推进行业统一技术标准的制定，辅助法律法规的施行

大数据领域具有天然的多学科交叉和技术更迭速度快的特点，其带来的隐私问题往往前所未见并且和技术的应用紧密相关，目前我国在信息安全、隐私保护和消费者权益等立法方

面已取得初步成效,但法律条文的修订速度在短时间内难以跟上技术的迭代发展。因此,需要从技术的角度入手,根据现有隐私问题的类型和特点,对相关技术进行针对性改进来提高伦理治理的水平。首先,政府和企业应建立合作关系,共同探索、制定和推行完备的行业技术标准以及健全的产品审查体系,将伦理准则和技术规范相结合,辅以法律的约束,确保在发展技术的同时能够坚守正确的伦理观,不会越过道德的红线。其次,应根据软硬件使用场景和要求的差异,分层次制定和持续推行统一的数据脱敏以及加密方案,在技术层面上对敏感数据进行变形处理和加密来保护个人的身份信息、手机号码、银行卡信息等数据的安全。最后,建立起伦理学、社会学和自然科学等多方学者参与的联动体,形成一个完善的数据隐私安全保障体系和隐私问题突发事件应急管理机制。

(二)加强数据伦理教育,培养学生的数据思维和隐私意识

无论是大数据技术的开发者还是使用者,都应树立良好的道德伦理意识,而这种意识的形成,离不开学校教育。尤其是在信息时代,高校应尽早帮助学生建立对数据伦理问题的普遍性和重要性的认知,然而国内高校的数据伦理教育仍未形成完整的体系,仅有部分高校开设了相关课程,如何制定合理的课程大纲,形成完整的课程周期都是将来要解决的问题。因此,我国高校之间应积极沟通,互相借鉴和讨论,探索和制定符合当前社会现状的数据伦理教学方案,形成完整的教育体系,尽快帮助学生了解大数据伦理问题治理过程中的新规范和新准则,引导学生思考大数据时代引发的社会问题,培养学生的数据思维,并能够在今后的学业和职业生涯中选择符合数据伦理的技术、商务和政务行动。培养一大批具有良好伦理意识的科研型、产业型和应用型复合人才,以满足社会发展的需要。

(三)强化企业内部审查制度,加快产品隐私政策的透明化

除了法律法规的监管以及行业技术标准的限制,对于公司自身而言,首先应建立一套完善的自我审查制度,提升企业数据库和服务器的安全等级,根据业务需求制定不同标准的加密及控制策略,对物联网设备上传到公司平台的数据进行审查。其次,细化物联网产品的隐私政策,为用户定制查看或更改隐私设置的接口并对已经销售的设备建立严格的售后监管体系,通过公司法务进行内容审核以达到自我监管的目的。最后,企业应该主动参与到数据安全标准的制定中去,定期向社会公开相关产品的隐私政策以及保护用户数据安全的相关举措,积极宣传数据安全法律法规,利用企业影响力提升公众数据安全意识。

(四)筑牢消费者维权通道,积极开展隐私安全的教育活动

隐私问题的发现和解决离不开消费者的参与,虽然我国与消费者维权相关的法律条例、通道建设和监管机制取得了很大进展,但仍然存在很多消费者在维权过程中因为势单力薄,沟通不畅,最后不了了之的情况。政府应强化监管措施,降低数据安全的总体风险,加强数据安全执法的力度,对违规企业开展专项检查和整治行动,完善消费者举报、投诉通道。媒体和个人也要积极曝光各类涉嫌侵犯个人隐私的事件,监督和促进相关法律法规体系的不断完善。同时,媒体作为第三方监督者,需要向大众积极宣传相关知识以加强民众对于隐私安全的了解,与政府和企业积极合作开展隐私安全的教育活动,鼓励公民维护个人隐私安全,引导民众建立起良好的个人隐私保护意识,养成良好的隐私保护习惯。

三、结语

在大数据时代，隐私保护的目的旨在和鼓励人们以更加合理安全的方式来管理和使用个人数据，趋利避害，最大化地发挥数据对社会的整体贡献。近年来涌现出的区块链、联邦学习和差分隐私保护等方法都是在技术层面的有益尝试[24]78，帮助我们规避个人隐私受到侵犯的风险。数据的生成、传输和利用时时刻刻都在发生，由数据带来的巨大利益往往会驱使一些人触碰道德伦理的底线。在一个科技高度发达的社会中，面对生活的便捷性和经济利益的诱惑，人类的隐私空间受到了前所未有的挑战，这就要求我们在使用大数据技术时坚决杜绝以侵犯用户合法隐私权利为代价，通过滥用或者非法二次利用数据来获取不当利益的行为。同时，在处理大数据技术引发的伦理事件时要秉持公平、公正、公开的原则，在坚守伦理底线的同时拥抱技术的发展和商业模式的变革，实现大数据技术和社会人文的和谐发展。

参考文献

[1] 徐瑾：《美国网络隐私权法律保护》，《现代情报》2005年第6期。

[2] 诺伯特·维纳：《人有人的用处——控制论和社会》，陈步译，北京：北京大学出版社，2010年。

[3] 王晓琳：《信息时代公共空间中的隐私问题》，《自然辩证法通讯》2018年第7期。

[4] Mayer-Schönberger, V., "Beyond Privacy, Beyond Rights-Toward a 'Systems' Theory of Information Governance", *California Law Review*, 2010, Vol.98, No.6, pp.1853-1885.

[5] 段伟文、纪长霖：《网络与大数据时代的隐私权》，《科学与社会》2014年第2期。

[6] 刘晓力：《哲学与认知科学交叉融合的途径》，《中国社会科学》2020年第9期。

[7] 闫坤如：《人工智能"合乎伦理设计"的理论探源》，《自然辩证法通讯》2020年第4期。

[8] 彭知辉：《论大数据伦理研究的理论资源》，《情报杂志》2020年第5期。

[9] Kosinski, M., Stillwell, D., Graepel, T., "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior", *Proceedings of the National Academy of Sciences of the United States of America*, 2013, Vol.110, No.15, pp.5802-5805.

[10] Nissenbaum, H., "Protecting Privacy in an Information Age: The Problem of Privacy in Public", *Law and Philosophy*, 1998, Vol.17, No.5, pp.559-596.

[11] 李飞翔：《"大数据杀熟"背后的伦理审思、治理与启示》，《东北大学学报（社会科学版）》2020年第1期。

[12] 曹秦雨：《大数据技术下新媒体用户画像与隐私安全》，《新闻研究导刊》2020年第24期。

[13] 方兴东、张静、刘国辉：《谷歌产品对用户个人隐私的影响——表现、趋势与对策》，《新闻界》2014年第11期。

[14] 刘培、池忠军：《算法歧视的伦理反思》，《自然辩证法通讯》2019年第10期。

[15] 张蕊：《从"水滴直播"看当前网络直播平台的危机与转机》，《传媒》2018年第8期。

[16] 王基策、李意莲、贾岩、周威、王宇成、王鹤、张玉清：《智能家居安全综述》，《计算机研究与发展》2018年第10期。

[17] 张玉清、周威、彭安妮：《物联网安全综述》，《计算机研究与发展》2017年第10期。

[18] 闵春发：《物联网技术的伦理问题》，《学海》2013年第6期。

[19] 顾理平、王飚濛：《社会治理与公民隐私权的冲突——从超级全景监狱理论看公共视频监控》，《现代传播（中国传媒大学学报）》2017年第6期。

[20] Agre, P. E., "Surveillance and Capture: Two Models of Privacy", *The Information Society*, 1994, Vol. 10, No. 2, pp.101-127.

[21] Margulis, S. T., "Privacy as a Social Issue and Behavioral Concept", *Journal of Social Issues*, 2010, Vol.59, No.2, pp.243-261.

[22] Mai, J. E., "Big Data Privacy: The Datafication of Personal Information", *Information Society*, 2016, Vol. 32, No.3, pp.192-199.

[23] 荆宁宁、程俊瑜：《数据、信息、知识与智慧》，《情报科学》2005年第12期。

[24] 张夏明、张艳：《人工智能应用中数据隐私保护策略研究》，《人工智能》2020年第4期。

From Big Data Based Price Discrimination to Privacy Leakage: Ethical Analysis and Reflections on Privacy Issues from the Perspective of Hardware and Software

WANG Yu-xing　YUAN Bo

Shenzhen International Graduate School, Tsinghua University
Shenzhen 518055, Guangdong, China
E-mail：wyx20@mails.tsinghua.edu.cn

Abstract: In the era of information, the emergence of big data and artificial intelligence technologies has made the issues of data privacy more acute and complex. In the past few years, a number of new privacy issues represented by the big data based price discrimination and the privacy leakage on smart platforms have emerged, generating significant negative impacts on governments, enterprises and the public. However, the information technology is updated at a pace much faster than the frequency of revision of legal regulations and the privacy issues need to be re-examined from a rational and constructive perspective. In this paper, we conduct a systematic analysis and reflection on the challenge of big data privacy and present corresponding solutions in terms of technical standards, ethics education and industrial regulations.

Keywords: big data; artificial intelligence; data privacy; data ethics

（责任编辑：董铠军）