人工智能时代的数据隐私、垄断与公平

孟小峰，王雷霞，刘俊旭
中国人民大学信息学院，北京 100872

摘要
随着人工智能时代的到来，大数据中蕴含的价值被不断开发，但与此同时，用户的隐私泄露问题、数据垄断问题以及算法决策中的公平问题愈发凸显。为详细探究此类伦理问题，首先从数据发展的角度出发，探讨人工智能时代隐私、垄断与公平问题的产生环境及其独特性。而后，对这3个伦理问题逐一分析其现状及挑战，得出当前伦理问题产生的本质是数据获取、使用以及决策的不透明性，提出建立数据透明机制是解决这些问题的重要举措。

关键词
大数据；数据伦理；人工智能；数据隐私；数据垄断；决策公平；数据透明

中图分类号：TP309 文献标识码：A doi: 10.11959/j.issn.2096-0271.2020004

Data privacy, monopoly and fairness for AI

MENG Xiaofeng, WANG Leixia, LIU Junxu
College of Information, Renmin University of China, Beijing 100872, China

Abstract
With the coming of the era of artificial intelligence, the value contained in big data has been deeply mined. But at the same time, the privacy and data monopoly issues of users' sensitive data, and fairness in algorithmic decisions have become increasingly serious. In order to explore such problems, firstly, the development of data was researched, which reflects the unique producing environment of data ethics in the era of artificial intelligence, and the unique properties of these ethical issues were discussed. Then, the data monopoly, privacy disclosure and unfair decision-making were discussed one by one, whose development status and challenges were analyzed. It is concluded that the essence of current ethical issues is the non-transparency of data collection, data usage and algorithm decision, so that establishing the data transparency mechanism should be an important measure to solve these problems.

Key words
big data, data ethics, artificial intelligence, data privacy, data monopoly, decision fairness, data transparency

1 引言

随着人工智能技术的快速发展及其在金融、交通、商业、医疗等领域的广泛应用，大数据中蕴含的价值不断被开发，产生了巨大的经济效益和社会效益。大数据逐渐改变着人们的生活生产方式。但与此同时，人们对大数据决策产生深度依赖，对自身数据失去掌控权，数据生态中的伦理问题愈演愈烈，用户数据的滥用问题、隐私泄露问题、数据垄断问题、决策公平问题层出不穷。2018年3月曝出的“Facebook剑桥分析事件”在未经用户授权的情况下收集用户信息，并企图影响2016年的美国总统大选，造成了用户数据滥用和隐私泄露。2017年11月，美国国防部由于服务器配置错误，意外暴露了18亿条用户社交数据，该事件揭露了用户隐私数据被收集和泄露的现状。2017年的顺丰与菜鸟关于丰巢数据之争、华为与腾讯关于微信数据之争均为“数据垄断”背景下的用户数据争夺现象。2015年，Google公司的人脸识别将黑人识别为大猩猩，造成“种族歧视”，暴露出了机器学习中的不公平问题。这些接踵而至的伦理问题，一方面使用户遭受了隐私威胁与非公平对待，另一方面引爆了用户与企业间的信任危机，致使自动驾驶、医疗健康预测等敏感领域的技术难以落地，从而限制了人工智能技术的发展。

愈演愈烈的数据伦理问题正在引起社会各界的广泛关注。学术界有关数据隐私、数据垄断、决策公平问题的学术争鸣不断涌现[1-2]。工业界中，Google公司CEO桑达尔·皮查伊在2019年6月17日接受美国有线电视新闻网专访时曾表明，首席道德官（chief ethics officer）应为CEO（chief executive officer）的另一含义，其对数据伦理等问题的关注程度可见一斑。2019年5月28日，中国国家互联网信息办公室发布《数据安全管理办法（征求意见稿）》，讨论当前隐私数据收集、处理及使用的办法。

从用户的角度考虑，人们作为数据的拥有者、大数据技术的使用者，在面对上述问题时，是否只能束手无策？还是选择将数据牢牢攥在手中，拒绝人工智能等技术的应用？事实上，挖掘数据价值与尊崇人类伦理并不是对立的问题，它们同时存在于人工智能时代的数据生态中，相互影响，相互制约，并将最终达到动态平衡的状态。

当前数据生态中的伦理问题根据其本质，可分为两类问题，分别是数据伦理问题和算法伦理问题。数据伦理问题是指在数据收集使用过程中产生的伦理问题，主要表现在隐私问题和垄断问题；算法伦理问题是指在算法决策过程中产生的伦理问题，主要表现在公平问题。因而，本文主要针对隐私、垄断和公平这3个代表性问题展开讨论，探讨当前伦理问题的本质。本文首先从数据发展的角度出发，探索这3个问题产生的特有数据环境，分析其独特性及不同问题之间的关联。之后，本文对这3个问题进一步详细探讨，分析其现状与挑战，并提出当前这些伦理问题产生的本质是数据获取、使用以及决策的不透明性，构建数据透明体系是解决当前隐私、垄断与公平问题的根本途径。

与此同时，本文提出当前的伦理问题应是一个“大隐私观”的问题。未来数据的发展带来的隐私问题不是现在关注的“小隐私”问题，即不能仅通过扰动、匿名、差分等技术实现保护，它是在数据收集使用场景下保证数据正确应用、算法正确决策的问题，涉及隐私、垄断、公平等伦理问题。相比狭义隐私问题，“大隐私”

问题涵盖内容更广，战线更长，需要研究者们跳出当前的思维定式，探索其本质与解决方案。

2 从数据发展看伦理问题

在数据发展的过程中，数据的产生方式及特征不断发生变化，对科学技术及社会产生了不同影响，进而发展出不同的伦理问题，而当下表现突出的是隐私问题、垄断问题与公平问题。从数据发展的主线上看，数据从数值型的科学数据发展到结构化的企业数据、多样的个人数据，其应用领域由自然领域逐渐拓展至工程领域、社会领域，推动了不同门类新技术的产生，带来了前所未有的伦理挑战。依据人们对数据的认识及应用程度，数据的发展可归结为“管理数据、理解数据、敬畏数据”3个阶段，如图1所示。

在计算机发展初期，数据通过自然观察、科学实验、统计调研等方式人为生成，多为数值型数据，人们借助计算机完成复杂的运算，促进自然发现、社会统计等学科的发展。同时，伴随着计算机存储设备的发展，文件系统、批处理等技术相继出现，人们使用这些技术对数据进行管理。此时的数据面临的主要问题更多集中于数据的正确性、共享性等应用问题。

在传统的数据库时代，数据在企业等运营式系统运营过程中被动产生，数据采集成本较高，故多以企业数据为主。此时数据结构规范有序，数据量相对有限，人们对数据的认识停留在“管理数据”的阶段，发展出数据库、数据仓库、数据集成等技术。该阶段，数据面临的主要问题是安全问题，即保护企业数据不被攻击者非法入侵和获取，确保导出的结果的正确性和完整性。

随着大数据时代的到来，数据采集愈发廉价，数据在个人移动设备、穿戴式设备、传感设备上源源不断地主动产生，数据结构复杂，数据量加速增长。此时的数据主要以个人数据为主，具有海量的数据集特性，人们开始“理解数据”，并由此发展出基于数据驱动的数据挖掘、云数据库、知识融合等技术。与此前借助符号进行逻辑推理不同，该阶段技术发展的本质是海量数据驱动的结果，产生了与此前截然不同的伦理问题。数据作为驱动算法的“燃料”，数据垄断与隐私问题层出不穷，而非规则的算法决策与黑盒模型使得决策可解释、公平问题备受关注。

而在逐步逼近的5G与万物互联时代，

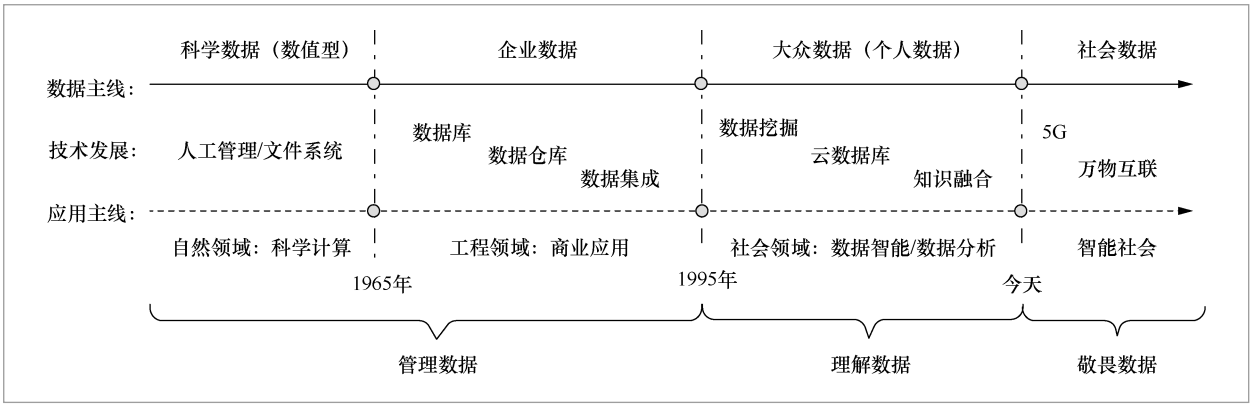

图1 数据的发展阶段

数据量爆炸式增长，数据描述社会的粒度更加细腻，在数据应用的过程中隐私、公平等伦理问题将更加严峻。由此，人们在应用挖掘数据的过程中，更需以敬畏的眼光看待数据，重视其伦理问题，发展出具有“敬畏数据”性质的技术。据目前的估计，世界上的数据大约只有20%可以得到有效管理，可被理解的数据不足1%，“敬畏数据”的技术几乎没有出现[3]，提出并实现这样的技术任重而道远。

由此，本文对人工智能时代的数据隐私、垄断与公平问题进行探讨，从“敬畏数据”的角度探索数据价值实现与数据伦理实现的有效途径，既不能扼杀挖掘数据价值的机会，更不能对人类伦理构成破坏[4]。要达到该目的，必须考虑到当前伦理问题的独特性，将其放在当前的数据生态中进行讨论。具体地，当前伦理问题的独特性体现在以下两个方面。

一方面，数据的特殊性质使得当前的伦理问题不能通过简单地界定数据归属的方法解决。从数据的发展可以看出，数据不同于森林、矿藏等源于自然的自然物，它会随着人、时间和环境的不同而有所差异；也不同于文学作品、专利等由人创造的精神产物，它是物质和精神的衍生物。因此，讨论人工智能时代的伦理问题时，并不能简单界定数据的归属。如果为了解决个人信息的隐私、垄断等问题，而将数据简单地界定为个人归属，则不能发挥其应有的价值。现有数据生态的特殊性决定了数据确权、定价、交易等孤立的形式并不能解决垄断、隐私和公平等伦理问题。

另一方面，不同伦理问题之间相互影响。首先，数据垄断与数据隐私之间存在相关促进的关系，数据垄断的破除将有效阻止大量数据的汇集，从而降低挖掘、泄露数据隐私的风险；其次，高度的数据隐私不利于数据垄断和决策不公平现象的发现。如何在考虑上述伦理问题时兼顾隐私问题十分关键，这也是本文强调的“大隐私观”需要特别关注的问题。

本文后续将深入探讨当前数据生态下的数据隐私问题、数据垄断问题和决策公平问题的本质，分析其现状与面临的挑战。基于该探讨分析，本文发现，当前数据伦理问题的产生是由数据在其生命周期中的不透明性造成的，规范数据的收集、流通、使用及决策势在必行。本文提出构建数据透明体系是这些伦理问题的有效解决途径。

3 数据隐私问题

在当前的移动用户数据收集的场景中，随着人工智能技术的发展和移动设备的普及，对用户隐私数据进行收集的现象愈演愈烈。一般地，App运营者可被视为数据收集者，用户可被视为数据提供者。移动用户数据收集的特点主要体现在以下几个方面：首先，在数据收集目的上，数据收集者均出于正义的目标和美好的愿景来收集数据，如发挥数据价值或提供更优质的个性化智能服务；其次，在数据收集方式上，他们都打着“免费使用服务”的名义，或以小恩小惠吸引数据提供者的参与，如一些平台通过优惠活动鼓励用户填写详细个人信息，以收集用户数据；再次，在数据收集过程中，存在欺瞒行为，一些App开发者不告知用户其个人数据的流向及使用目的，请求用户同意数据收集的授权协议通常以“默认勾选”或隐藏选项的方式使用户“被同意”，更甚者通过收集和贩卖用户数据进行非法数据流通；最后，在用户数据的隐私保护上，他们没有采取任何有效的隐私保护措施，诸多企业直接在用户的隐私数据上进行数据分析，用户

的隐私岌岌可危。

上述做法不仅威胁着用户的个人隐私，也隐含着国家安全问题，包括国民个人数据的跨境流通问题以及国防安全问题，如与导航和防御相关的天文数据的安全问题。因此，如何有效保护用户隐私与数据安全是当前数据生态面临的主要问题之一。

为应对该问题，国家和研究者们分别从制度和技术上做了诸多努力。在制度上，随着隐私问题的逐渐凸显，相关立法在稳步进行。欧盟于2018年5月25日出台《通用数据保护条例（General Data Protection Regulation，GDPR）》，规定了用户在数据上的查阅权、被遗忘权等权利，以保护个人隐私，遏制数据滥用。2019年4月16日，美国旧金山通过了对《停止秘密监视》条例的部分修订，考虑到人脸识别技术可能侵犯用户隐私、加剧种族歧视等问题，禁用该项技术。2019年5月28日，中国国家互联网信息办公室发布《数据安全管理办法（征求意见稿）》，从数据收集、处理使用、安全监管几个方面讨论其管理办法。

在技术上，通过隐私保护技术完成数据流通和数据处理，避免数据直接流通导致泄露用户隐私[5-8]。目前已有基于扰动和基于密码学的两类隐私保护方案[9]。基于扰动的方案主要指匿名技术（anonymity technology）[10-16]、中心化差分隐私（differential privacy）[14-22]、本地化差分隐私（local differential privacy）[23-31]，该类方法计算效率高、应用成熟，但会降低数据精度，影响数据可用性。基于密码学的隐私保护方案主要指同态加密（homomorphic encryption）[32-37]、安全多方计算（secure multi-party computation）[38-40]，该类方法安全性较高、数据可恢复，但效率较低，商用性较差。

目前隐私问题还存在许多挑战。首先，在制度上，数据作为物质与精神的衍生物，不能简单界定其归属，将数据简单归属于用户而粗暴地禁止人脸识别的应用，并不是最有效的立法准则，如何兼顾数据价值的实现与对人类伦理的尊崇，从而完善立法是当前的挑战之一。其次，在技术上，当前的隐私保护方法都表现出一定的局限性，重点体现在数据可用性与数据隐私之间的权衡。最后，特别注意的是，当前数据的隐私保护不能局限于对敏感数据的保护，数据的发展会使个人敏感性问题降低，人们需正视数据合理获取、存储、使用的问题，应从“大隐私”的角度出发，同时兼顾隐私、垄断、公平等其他伦理问题。

4 数据垄断问题

在当前数据的收集、使用过程中，数据垄断问题愈发明显。2019年1月，中国人民大学WAMDM实验室发布的《中国隐私风险指数分析报告》基于3 000万名手机用户的数据对大规模的数据收集现状进行了统计分析，分析结果显示，10%的数据收集者获取了99%的权限数据。其中，数据收集者指的是移动用户数据收集场景下的App运营者；权限数据指的是在该场景下，数据收集者通过App的权限体系获取的用户个人隐私数据。由此可见，数据收集的垄断现象极为严重，其残酷程度更甚于现实世界财富获取的“二八定律”。同时，数据作为数字经济时代的战略性基础资源，数据驱动型公司围绕数据的竞争愈演愈烈，如微博与脉脉的数据之争、顺丰与菜鸟的数据之争等。

造成数据垄断的主要原因，一方面是在大数据时代数据本身的价值密度低，其应用价值需通过海量数据的挖掘获取来实现，从而易造成数据聚集现象；另一方面，当前大

型商业公司的跨多领域的商业模式、庞大的用户规模及网络效应使其数据收集能力不断增强，不同数据收集者之间的鸿沟逐渐拉大，使得数据垄断现象愈演愈烈。

当前严峻的数据垄断形势会给数据生态造成3方面的负面影响[41-42]。首先，巨头公司拥有大部分的数据和用户流量，在当前数据推动发展的历史模式下，会进一步压缩其他公司的生存空间，不利于其他公司尤其是小型、新型企业的出现及发展；其次，巨头公司可利用这些丰富的数据形成一条生产线，开发多领域的生产经营活动，使得技术不外化，不利于其他新技术的产生；最后，拥有海量数据的巨头公司具有主导市场竞争的资本，自由竞争的失效将使用户失去服务的可替代性选择，从而使数据滥用、隐私泄露、价格歧视等其他伦理问题加剧。

由数据拥有和控制引发的数据垄断与竞争问题已引起了市场监管和竞争执法部门的注意，并相继做出一系列的适用政策修订，如2016年10月，全球移动通信系统协会发布《数字生态系统竞争政策框架重整》；2017年2月欧盟发布《大数据与竞争政策：市场力量、个性化定价与广告》；我国国家市场监督管理总局于2019年1月30日发布《禁止滥用市场支配地位行为的规定（征求意见稿）》，首次将数据垄断纳入反垄断执法考量范围。

与此同时，学术界与工业界也试图从技术上对该问题进行治理。从源头上，隐私保护技术和访问控制技术可对数据的收集和使用进行一定程度的干预，降低或限制数据巨头持有的数据。在数据流通过程中，上海数据交易中心、贵阳大数据交易所等数据交易平台的建立可促进数据的流通与共享，削弱数据收集者对数据的控制权。

上述举措虽在一定程度上缓解了数据垄断的局势，但并不能根治该问题。应对数据垄断，要寻求更好的数据治理模式，不能一味封锁和限制数据的采集和使用，“开源节流”十分关键。一方面，要规范数据的收集、流通和使用，使数据资源得到合理有效的配置；另一方面，要积极探索隐私保护的数据共享技术，打破数据孤岛，促进数据流通。

5 决策公平问题

机器学习算法在服务智能生活的同时，公平问题逐渐产生。2015年，亚马逊通过机器学习实现的自动化招聘系统存在性别歧视，最终该项目被关闭。2016年弗吉尼亚大学文森特·欧多尼兹教授通过对图形识别软件进行大量测试，发现其易于将键盘鼠标等与男性结合，将厨房购物等与女性结合，存在偏见。2018年“大数据杀熟”被选为年度社会生活类十大流行语之一，其含义是指电商平台或服务网站为用户提供智能服务时，基于用户数据分析对同一商品为不同用户提供差异化定价，引发价格歧视。用户应意识到，机器学习算法为人类当前的研究分析工作提供了更高效的结果，但不一定是更正确的结果，其算法决策中存在的不公平、不可信等问题值得引起大众关注。

从理论的角度对上述现象进行分析可知，机器学习模型的正确性极度依赖训练数据，然而训练数据都是由人标注产生的。人是天生带有偏见的，并且会无意识地将这种偏见注入训练和测试数据中，或有意识地注入训练过程中。例如，在人工对训练数据进行标注时，因标注者不熟悉标注对象引起的标注错误，因不同人群的经验、文化差异而带来的数据差异等。基于这些数据训练得到的机器学习模型就是不合理的决策模型。而“大数据杀熟”则是

在机器学习的过程注入了商家对不同消费者购买能力的歧视和偏见，从而达到其利益最大化。在此过程中，偏见由人传递到数据，再由数据传递到模型，人类对社会及事物固有的偏见不仅不会得到遏制，还会得到放大。

要探索机器学习下的公平，首先需明确公平的含义。公平是一个多维概念，体现的是人们对平等的追求。社会学中的公平指“同工同酬”，心理学中的公平则认为人们的公平感取决于一种同他人的社会比较或同自己的历史比较。在哲学上，哲学家约翰·罗尔斯在《正义论》中提出利用一个重要的假设“无知之幕”来定义公平。“无知之幕”假设了一个人人平等的博弈条件，即“无知之幕”后的每个人都不清楚自己在社会中将扮演的角色，此时这些人共同制定的规则才可能公平。“无知之幕”揭示的是规则制定者的选择不被他们的特殊利益左右，从而使得在一个问题中涉及的所有方被置于同一标杆之后，被一视同仁地对待。对于机器学习中的公平而言，由于现实世界并非绝对公平，理想状态下的机器学习公平一方面要反映客观现实，另一方面更应能够纠正由人带来的主观偏见。

追求“无知之幕”下的人工智能即追求算法公平，其应用的机器学习模型至少应满足以下两个要求[43]：第一，对于相同的应用场景，相似数据集可以经训练得到相似模型，如Google地图在印度数据集上的模型准确率应与在美国数据集上训练的模型准确率一致；第二，向模型中输入相似个体的信息可以得到相似输出，如对于能力、学历相等的男女求职者，其被推荐的工作和薪金应相近。

如何实现满足上述公平的人工智能算法，仍旧充满挑战。从数据的角度考虑，如果可对决策数据进行合理的审计，使决策过程具备透明性和可理解性，那么数据中的偏见就可能被发现、被问责，从而达到避免偏见引入的目的。从算法的角度出发，当前有许多研究工作者集中精力于引入公平性度量，从而对机器学习模型本身进行改进，但该方法针对特定的机器学习算法，具有局限性[44]。更广义地理解决策，它应该包含自动决策和人工决策，而这两种决策都存在不同程度的偏见和误差。如果能够综合考虑自动决策和人工决策，那么就可以得到更全面的决策结果，进而提高决策的公平性，但如何将它们合理结合仍是一个现实问题。

6 解决途径：建立数据透明机制

上述伦理问题产生的根本原因是大数据价值实现过程中的不透明性。当前数据的获取、流通、共享、使用和决策过程都存在不透明性，用户作为数据的生产者，对哪些数据被收集、被谁收集、流向何处、做何使用一无所知。在人工智能服务的大环境下，个人数据在其整个生命周期，包括产生、流通、使用和决策的过程中，都处于黑盒状态，这进一步加剧了数据的隐私泄露、垄断和决策结果的不公平。而与传统的决策相比，基于大数据进行决策产生的伦理问题更为显著的主要原因在于，传统决策的基础是“数据—信息—知识”的获取，而现有的大数据决策是由数据直接驱动的，数据错误与算法不透明导致底层数据不可靠，决策不可信。这一状态在弗兰克·帕斯奎尔的《黑箱社会》中被描述为“黑箱”，大数据透明应是射入这个“黑箱”的“一道阳光”，是解决上述伦理问题的根本途径。

大数据透明旨在保护个人数据在其生命周期中的透明性，即保证数据在数据获取、共享、存储和决策的过程中对其从属

主体的透明性，也就是说，通过数据透明，参与的主体能够获取与自身相关的全部数据信息。由此，应用数据透明可以对数据的收集、流通、决策进行适度的公开、记录、审计和问责，从而促进隐私、垄断、公平伦理问题的解决，具体如图2所示。

在数据隐私问题上，一旦数据隐私发生泄露，可通过数据透明机制对泄露数据进行溯源，对其发布过程中违反规范的参与方进行问责，从而对数据的合理收集与使用进行有效的监督。但在该过程中，应十分注意数据透明的范围和粒度，如果透明的范围太大、对象太广，则有可能暴露企业或个人的机密信息。因此，在解决该问题时，不能一味地追求透明，应兼顾可溯源数据的隐私性。

在数据垄断问题上，可在数据流通的过程中通过数据透明对数据的流向进行追踪和审计，一方面结合访问控制等技术对数据的流向进行一定程度的限制，避免数据垄断；另一方面，可从宏观的角度对数据的共享使用提出建议或提供数据共享的可能，打破不同数据收集分析者之间的屏障，促进基于数据驱动的人工智能决策方法的发展与应用。

在决策公平问题上，可通过数据透明对决策的结果进行审计，使得其结果中的歧视、偏见等不公平问题可被发现。基于该审计结果，算法工程师可进一步完善决策算法或决策输入数据，从而提高数据决策的公平性。

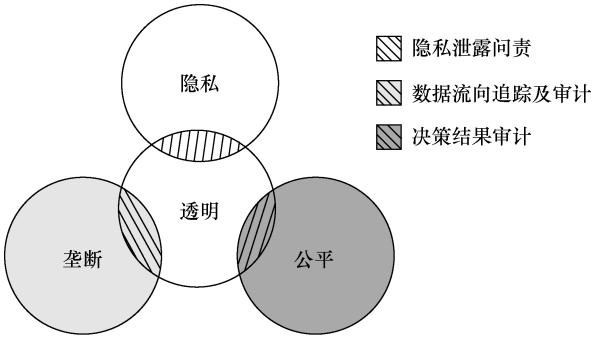

图 2 数据透明对隐私、垄断和公平问题的促进作用示意图

在大数据透明的具体实现上，政府机构和研究者们分别从政策和技术上做出诸多努力和探索。政策上，GDPR等法律法规的出台，明确规定了数据主体（即用户）对数据的控制权，以保证个人数据在其数据生命周期中具有更高的透明度，数据主体对个人数据具有更强的管控能力；技术上，借助区块链难以篡改、可追踪、去中心和公开透明的特性，可实现数据透明的需求，具体地，可基于区块链从访问控制、数据存储、分布式机器学习等角度积极探索数据透明的实现。

然而，就大数据透明而言，当前还存在诸多问题尚待解决。首先，针对不同问题，大数据透明的范围和面向的对象也不尽相同，数据透明的范围和粒度对企业和个人隐私信息的保护至关重要；其次，大数据透明提供关于数据和算法的关键信息，可能会造成隐私的泄露或给攻击者提供有效的背景知识，此时不仅不能促进隐私问题的解决，反而会加剧该问题，如何在隐私保护的情况下提供数据的审计，从而实现数据透明，十分关键；再次，透明的实现贯穿整个数据收集、存储、流通、使用以及算法决策的过程，应同时保证效率性与透明性；最后，当前数据生态中的伦理问题主要受到法律法规的约束、普适道德观的约束以及技术规范的制约，用户自身素养与大隐私意识的提高十分重要。

7 结束语

在数据驱动的机器学习时代，数据的

总量和维度不断丰富，其通过机器学习等算法产生了巨大的社会价值，但同时也引发了数据隐私、数据垄断和决策公平这3个典型的伦理问题。如何在实现数据价值的同时解决这些伦理问题，发展“敬畏数据”的技术，是当前研究的重中之重。树立“大隐私观”，正视数据在其生命周期中合理收集、存储、使用的问题是十分关键的。

本文首先从数据发展的角度探讨这些伦理问题，归结出这些伦理问题的产生是当前的数据生态环境与数据驱动的机器学习技术相互作用的结果。同时，提出数据应当放在数据生态中加以考量，不能通过简单界定其归属的方法来解决伦理问题。之后，本文对数据隐私、数据垄断、决策公平3个问题分别进行了探讨，发现其本质是当前数据生态环境下，数据在其生命周期中的不透明性。最后，本文提出建立数据透明机制是解决人工智能时代数据生态伦理的关键步骤，如何有效地建立该体系应是当前该领域研究的重点方向之一。

参考文献:

[1] STOYANOVICH J, HOWE B, JAGADISH H V, et al. Panel: a debate on data and algorithmic ethics[J]. Very Large Data Bases, 2018, 11(12): 2165-2167.

[2] STOYANOVICH J, ABITEBOUL S, MIKLAU G, et al. Data, responsibly: fairness, neutrality and transparency in data analysis[C]// International Conference on Extending Database Technology, March 15-18, 2016, Bordeaux, France. Heidelberg: Springer, 2016: 718-719.

[3] GANTZ J, REINSEL D. The digital universe in 2020: big data, bigger digital shadows, and biggest growth in the far east[J]. IDC iView: IDC Analyze the Future, 2012(1): 1-16.

[4] BORGMAN C L. 大数据、小数据、无数据：网络世界的数据学术[M]. 孟小峰, 等, 译. 北京: 机械工业出版社, 2017.
BORGMAN C L. Big data, little data, no data: scholarship in the networked world[M]. Translated by MENG X F, et al. Beijing: China Machine Press, 2017.

[5] 方滨兴, 贾焰, 李爱平, 等. 大数据隐私保护技术综述[J]. 大数据, 2016, 2(1): 1-18.
FANG B X, JIA Y, LI A P, et al. Privacy preserving in big data: a survey[J]. Big Data Research, 2016, 2(1): 1-18.

[6] 孟小峰, 张啸剑. 大数据隐私管理[J]. 计算机研究与发展, 2015, 52(2): 265-281.
MENG X F, ZHANG X J. Big data management[J]. Journal of Computer Research and Development, 2015, 52(2): 265-281.

[7] 张啸剑, 孟小峰. 面向数据发布和分析的差分隐私保护[J]. 计算机学报, 2014, 37(4): 927-949.
ZHANG X J, MENG X F. Differential privacy in data publication and analysis[J]. Chinese Journal of Computers, 2014, 37(4): 927-949.

[8] 叶青青, 孟小峰, 朱敏杰, 等. 本地化差分隐私研究综述[J]. 软件学报, 2018, 29(7): 159-183.
YE Q Q, MENG X F, ZHU M J, et al. Survey on local differential privacy[J]. Journal of Software, 2018, 29(7): 159-183.

[9] 刘俊旭, 孟小峰. 机器学习的隐私保护研究综述[J]. 计算机研究与发展, 2019, 已录用.
LIU J X, MENG X F. Survey on privacy-preserving machine learning[J]. Journal of Computer Research and Development, 2019, Accepted.

[10] SWEENEY L. k-anonymity: a model for protecting privacy[J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(5): 557-570.

[11] LI N H, LI T C, VENKATASUBRAMANIAN S, et al. Closeness: a new privacy measure for data publishing[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(7): 943-956.

[12] XIAO X K, TAO Y F. M-invariance: towards privacy preserving republication of dynamic datasets[C]// The 2007 ACM SIGMOD International Conference on Management of Data, June 11-14, 2007, Beijing, China. New York: ACM Press, 2007: 689-700.

[13] BU Y Y, FU A W C, WONG R C W, et al. Privacy preserving serial data publishing by role composition[J]. Proceedings of the VLDB Endowment, 2008, 1(1): 845-856.

[14] LI C, PALANISAMY B. ReverseCloak: protecting multi-level location privacy over road networks[C]// The 24th ACM International Conference on Information and Knowledge Management, October 18-23, 2015, Melbourne, Australia. New York: ACM Press, 2015: 673-682.

[15] CHENG J, FU A C W, LIU J. K-Isomorphism: privacy preserving network publication against structural attacks[C]// The 2010 ACM SIGMOD International Conference on Management of Data, June 6-10, 2010, Indianapolis, USA. New York: ACM Press, 2010: 459-470.

[16] ZHAO C, ZOU L, LI F F. Privacy preserving subgraph matching on large graphs in cloud[C]// The 2016 International Conference on Management of Data, June 26-July 1, 2016, San Francisco, USA. New York: ACM Press, 2016: 199-213.

[17] DWORK C. Differential privacy[J]. Encyclopedia of Cryptography and Security, 2006(1): 1-12.

[18] QARDAJI W, YANG W N, LI N H. PriView: practical differentially private release of marginal contingency tables[C]// International Conference on Management of Data, June 22-27, 2014, Snowbird, USA. New York: ACM Press, 2014: 1435-1446.

[19] DAY W Y, LI N H, LYU M. Publishing graph degree distribution with node differential privacy[C]// The 2016 International Conference on Management of Data, June 26-July 1, 2016, San Francisco, USA. New York: ACM Press, 2016: 123-138.

[20] XU S Z, SU S, XIONG L, et al. Differentially private frequent subgraph mining[C]// 2016 IEEE 32nd International Conference on Data Engineering, May 16-20, 2016, Helsinki, Finland. Piscataway: IEEE Press, 2016: 229-240.

[21] ZHANG J, XIAO X K, XIE X. PrivTree: a differentially private algorithm for hierarchical decompositions[C]// The 2016 International Conference on Management of Data, June 26-July 1, 2016, San Francisco, USA. New York: ACM Press, 2016: 155-170.

[22] KRISHNAN S, WANG J N, FRANKLIN M J, et al. PrivateClean: data cleaning and differential privacy[C]// The 2016 International Conference on Management of Data, June 26-July 1, 2016, San Francisco, USA. New York: ACM Pres, 2016: 937-951.

[23] WARNER S L. Randomized response: a survey technique for eliminating evasive answer bias[J]. Journal of the American Statistical Association, 1965, 60(309): 63-69.

[24] ERLINGSSON Ú, PIHUR V, KOROLOVA A. RAPPOR: randomized aggregatable privacy-preserving ordinal response[C]// The 2014 ACM SIGSA Conference on Computer and Communications Security, November 3-7, 2014, Hong Kong, China. New York: ACM Press, 2014: 1054-1067.

[25] BUN M, NELSON J, STEMMER U. Heavy hitters and the structure of local privacy[C]// The 37th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems, June 10-15, 2018, Houston, USA. New York: ACM Press, 2018: 435-447.

[26] WANG T H, BLOCKI J, LI N H. Locally differentially private protocols for frequency estimation[C]// USENIX Security Symposium, August 16-18, 2017, Vancouver, Canada. Berkeley: USENIX Association, 2017: 729-745.

[27] WANG T H, LI N H, JHA S. Locally

differentially private frequent itemset mining[C]// IEEE Symposium on Security and Privacy, May 21-23, 2018, San Francisco, USA. Piscataway: IEEE Press, 2018: 127-143.

[28] YE Q Q, HU H B, MENG X F, et al. PrivKV: key-value data collection with local differential privacy[C]// IEEE Symposium on Security and Privacy, May 20-22, 2019, San Francisco, USA. Piscataway: IEEE Press, 2019: 317-331.

[29] QIN Z, YU T, YANG Y, et al. Generating synthetic decentralized social graphs with local differential privacy[C]// 2018 IEEE 4th International Conference on Computer and Communications Security, June 22-24, 2018, Haikou, China. New York: ACM Press, 2017: 425-438.

[30] SUN H, XIAO X, KHALIL I, et al. Analyzing subgraph statistics from extended local views with decentralized differential privacy[C]// 2019 IEEE 5th International Conference on Computer and Communications Security, November 11, 2019, London, UK. New York: ACM Press, 2019: 703-717.

[31] ZHANG Z K, WANG T H, LI N H, et al. CALM: consistent adaptive local marginal for marginal release under local differential privacy[C]// 2018 IEEE 4th International Conference on Computer and Communications Security, June 22-24, 2018, Haikou, China. New York: ACM Press, 2018: 212-229.

[32] PAILLIER P. Public-key cryptosystems based on composite degree residuosity classes[C]// International Conference on the Theory and Application of Cryptographic Techniques, May 2-6, 1999, Prague, Czech Republic. [S.l.:s.n.], 1999: 223-238.

[33] STEHLÉ D, STEINFELD R. Faster fully homomorphic encryption[C]// International Conference on the Theory and Application of Cryptology and Information Security, December 5-9, 2010, Singapore. Heidelberg: Springer, 2010: 377-394.

[34] DOMINGO-FERRER J. A provably secure additive and multiplicative privacy homomorphism[C]// The 5th International Conference on Information Security, September 30-October 2, 2002, São Paulo, Brazil. Heidelberg: Springer, 2002: 471-483.

[35] PLANTARD T, SUSILO W, ZHANG Z F. Fully homomorphic encryption using ideal lattices[C]// International Conference on Theory of Computing, May 31-June 2, 2009, Bethesda, USA. Heidelberg: Springer, 2009: 169-178.

[36] HU H, XU J, XU X, et al. Private search on key-value stores with hierarchical indexes[C]// 2014 IEEE 30th International Conference on Data Engineering, March 31-April 4, 2014, Chicago, USA. Piscataway: IEEE Press, 2014: 628-639.

[37] HU H B, XU J L, REN C S, et al. Processing private queries over untrusted data cloud through privacy homomorphism[C]// 2011 IEEE 27th International Conference on Data Engineering, April 11-16, 2011, Hannover, Germany. Piscataway: IEEE Press, 2011: 639-644.

[38] DU W L, ATALLAH M J. Secure multi-party computation problems and their applications: a review and open problems[C]// The New Security Paradigms Workshop, September 10-13, New Mexico, USA. New York: ACM Press, 2001: 13-22.

[39] VAIDYA J, CLIFTON C. Privacy preserving association rule mining in vertically partitioned data[C]// ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 23-26, 2002, Edmonton, Canada. New York: ACM Press, 2002: 639-644.

[40] SHEIKH R, MISHRA D K, KUMAR B. Secure multiparty computation: from millionaires problem to anonymizer[J]. Information Security Journal: A Global

Perspective, 2009, 20(1): 25-33.

[41] 孟小峰, 朱敏杰, 刘俊旭. 大规模用户隐私风险量化研究[J]. 信息安全研究, 2019(9): 778-788.
MENG X F, ZHU M J, LIU J X. Quatitative research on privacy risk of large-scale mobile user[J]. Journal of Information Security Research, 2019(9): 778-788.

[42] 孟小峰, 朱敏杰, 刘立新, 等. 数据垄断与其治理模式研究[J]. 信息安全研究, 2019(9): 789-797.
MENG X F, ZHU M J, LIU L X, et al. Research on data monopoly and its governance modes[J]. Journal of Information Security Research, 2019(9): 789-797.

[43] GUPTA M. How do we make AI fair? (Keynote) [C]// International Conference on Systems and Machine Learning, December 15, 2019, Boca Raton, USA. [S.l.:s.n.], 2019.

[44] ZHANG X Y, WANG N F, JI S L, et al. Interpretable deep learning under fire[C]// USENIX Security Symposium 2020, August 12-14, 2020, Boston, USA. [S.l.:s.n.], 2020, accepted.

作者简介

孟小峰(1964-), 男, 博士, 中国人民大学信息学院教授, 博士生导师, 中国计算机学会会士, 主要研究方向为数据库理论与系统、大数据管理系统、大数据隐私保护、大数据融合与智能、大数据实时分析、社会计算等。

王雷霞(1994-), 女, 中国人民大学信息学院博士生, 主要研究方向为隐私保护。

刘俊旭(1995-), 女, 中国人民大学信息学院博士生, 主要研究方向为隐私保护。

收稿日期: 2019-10-21

通信作者: 孟小峰, xfmeng@ruc.edu.cn

基金项目: 国家自然科学基金资助项目(No.91646203, No.61532010, No.91846204, No.61532016, No.61762082)

Foundation Items: The National Natural Science Foundation of China (No.91646203, No.61532010, No.91846204, No.61532016, No.61762082)