2019 年第 5 期（总第 5 期）

国际战略与安全研究报告

INTERNATIONAL SECURITY AND STRATEGY STUDIES REPORT

中国人工智能数据安全治理与伦理：“碎片化治理”中的一块拼图

清华大学战略与安全研究中心
CENTER FOR INTERNATIONAL SECURITY AND STRATEGY TSINGHUA UNIVERSITY

中国人工智能数据安全治理与伦理：“碎片化治理”中的一块拼图

朱荣生（清华大学战略与安全研究中心助理研究员）

于 洋（清华大学交叉信息研究院助理教授）

随着各类智能设备与系统开始大规模收集个人数据信息，数据安全风险已成为影响各国安全发展人工智能（AI）技术的一大核心因素。苹果公司历来宣称注重保护消费者数据，近期被指控侵犯消费者隐私。2019 年 7 月，有报道揭露苹果公司允许其承包商在未经用户同意的情况下，擅自使用用户的语音信息。这些对话内容清晰、易于监听，还暴露了用户所在位置、联系方式等私人信息。此消息引发了公众对苹果公司的担忧。“嗨 Siri，你在监听我吗”一时间也成为人们热议的话题。另一起众所周知的案例是 2016 年的“Facebook 隐私泄露丑闻”。英国数据分析公司剑桥分析擅自收集了 8700 余万 Facebook 用户的数据，而且还在 2016 年美国大选期间，利用人工智能技术定向投放政治广告，影响选民的意识形态和政治观点。这起事件不仅反映了人工智能应用场景下的数据泄露与安全风险，也表明这项技术会加大国际治理和国家安全的挑战。

在人工智能时代，如何应对数据安全所带来的挑战已成为包括中国在内的世界各国需要解决的重要问题。中国的治理经验能否为我们提供关于全球人工智能数据安全治理的有益经验？ 2019 年 8 月 30 日，一款名为“ZAO”的换脸应用软件在苹果应用商店中国区上架销售。这款软件采用了深度伪造技术，用户只需上传一张清晰的面部照片，就可以将自己的脸替换成电影或电视剧中演员的脸。“ZAO”在应用商店的评分一路飙升，甚至在一天内攀升至热搜榜前列。然而，在一夜爆红后，外界开始质疑“ZAO”存在数据泄露风险，并且对其用户

协议极其不满。其中一项条款为用户上传或发布内容后，同意授权或确保实际权利人同意 “ZAO”及其关联公司以及“ZAO”的用户在全球范围内完全免费、不可撤销、永久、可转授权和可再许可的权利。2019 年 9 月 3 日，工业和信息化部约谈该公司相关负责人，要求其严格按照国家法律法规以及相关主管部门要求，组织开展自查整改。此案例说明在中国目前尚缺乏系统化的人工智能法律体系的情况下，违反伦理道德规范和相关法律法规的企业不仅难以持续获利，还会受到来自市场和监管部门的双重处罚。

如果企业在人工智能时代下无法保护人们的隐私，就会失去客户。企业在逐利时，也需要遵守伦理道德和法律法规。中国企业实际上早已采用人工智能技术挖掘有害信息，为社会治理提供支持。2018 年 9 月，腾讯宣称其旗下的宾果反诈骗防控系统为中国 31 个省市自治区提供反诈骗防控服务，自上线以来累计推送预警 5 万多条，准确率超过 99%，累计为中国群众避免经济损失达 20 亿元。据百度发布的《2018 年上半年信息安全综合治理报告》显示，百度在 2018 年上半年一共处理了 145.4 亿条有害信息，其中占比居前两位的是淫秽色情类和赌博类，分别达到了 51.04% 和 16.63%。

此外，中国企业也在推动数据安全治理的伦理道德规范建设方面发挥着重要作用。这些企业通过发布伦理道德原则来规范自身的生产活动，打造可信的形象，进而提高在国内外市场中的竞争力。在彰显社会责任时，中国企业也借此塑造行业规范，引领“软法”建设。百度提出了人工智能伦理四原则，其中最高原则是安全可控。2017 年，腾讯和中国科学院联合发布了人工智能发展六大原则，涵盖自由、正义、福祉、伦理、安全和责任，强调人工智能的发展应加强隐私保护，防止数据滥用，个人数据的收集与使用应符合规范与法律制度。

针对人工智能原则的讨论目前在中国进行地如火如荼。企业、学术界、产业界和中国政府陆续发布人工智能原则，强调使用人工智能

时应加强数据安全保护。2019 年 5 月，中国的高校、科研院所和产业联盟联合发布《人工智能北京共识》，针对人工智能的研发、使用、治理三个方面，提出了 15 条原则。在保护数据安全方面，该共识提出“实现人工智能系统的数据安全”，“鼓励建立人工智能开放平台，避免数据与平台垄断”，“应建立合理的数据与服务撤销机制，以确保用户自身权益不受侵害”。2019 年 6 月，中国人工智能产业发展联盟发布《人工智能行业自律公约（征求意见稿）》，其中专门列出了隐私保护的原则。在 2019 年世界人工智能大会上，青年科学家代表发布了《中国青年科学家 2019 人工智能创新治理上海宣言》，呼吁制定相关法律法规，加强隐私保护意识，发展隐私保护算法和技术。产业界和学术界在伦理规范制定上的积极行动反映出中国社会既愿意接受人工智能技术的广泛应用，也希望制定能够确保该技术安全使用的规范。

2019 年 6 月，国家新一代人工智能治理专业委员会发布《新一代人工智能治理原则》，将尊重隐私列入八项原则。《新一代人工智能治理原则》指出人工智能发展应尊重和保护个人隐私，充分保障个人的知情权和选择权。同时，中国政府正在加快相关立法进程。2016 年，中国的《网络安全法》增加了用户的个人信息删除权、知情权、更正权等新规定。自 2018 年以来，中国还在加强《数据安全法》和《个人信息保护法》中个人数据保护的相关立法工作，发布了《数据安全管理办法（征求意见稿）》和《个人信息出境安全评估办法（征求意见稿）》，公开向社会征求意见。随着人们对深度伪造技术的担忧加剧，中国在 2019 年审议修订了《民法典人格权编（草案）》，规定任何组织或者个人不得以利用信息技术手段伪造的方式侵害他人的肖像权。这意味着未来中国法律将会禁止使用该技术。

从全球治理的角度来看，主要国家都提出了相应的治理方案，但尚未共同建立国际合作机制，从而引发了人工智能治理碎片化的问题。

国际治理会受到国内治理的影响。人工智能数据安全治理应解决各国面临的社会问题，以及体现其价值观的伦理问题。此外，目前尚未形成统一的人工智能技术标准，因而进一步扩大了全球人工智能治理方案的差异，使全球治理呈现出“碎片化”的特征。有专家指出人工智能的发展目前依然处于“弱人工智能”阶段，未来发展仍然有很大的不确定性。换言之，现在讨论科幻小说中人工智能取代人类的场景，还为时尚早。这可能也是各国不急于采取行动，加速建立相关国际机制的原因之一。

中国的经验来自于其错综复杂的治理现实。目前看来，中国正在走一种以市场为导向、以政府为主导的治理路径。中国人民乐于接受人工智能带来的便利生活，但也担心潜在的数据泄露风险。企业违反社会价值观和伦理道德，会破坏自身品牌形象，并且受到来自市场、法律、社会舆论等诸多方面的惩罚。因此，企业为了维护自身利益也应参与人工智能治理，在商业行为中遵守伦理道德。中国和许多其他国家一样，也认识到人工智能可能会威胁国家安全，因此采取了多种措施完善法律法规制度和人工智能伦理道德规范。但是，人工智能技术的飞速发展，和其产生的广泛社会影响，导致立法工作难以跟上问题产生的速度。不仅仅是中国，全世界都面临着这项严峻的挑战。不久前，欧洲公布了关于已经实行一年的《通用数据保护条例（GDPR）》的研究报告。此条例被称为史上最严格的个人数据保护条例，但是从实施效果来看却未能完全达到预期设想。针对人工智能全球治理没有一劳永逸的方法。我们应将全球人工智能治理视为一个动态的进程。有人认为人工智能治理属于权力竞争的一部分，因此更加难以达成合作。然而，竞争并不意味着对抗，人工智能也非零和游戏。就目前的全球治理进程而言，竞争反而能够促使“碎片化治理”模式向相互合作的模式转变。

本期责编：周武华

清华大学战略与安全研究中心

办公地点：清华大学明斋 217

联系电话：010-62771388

电子邮箱：ciss_thu@163.com