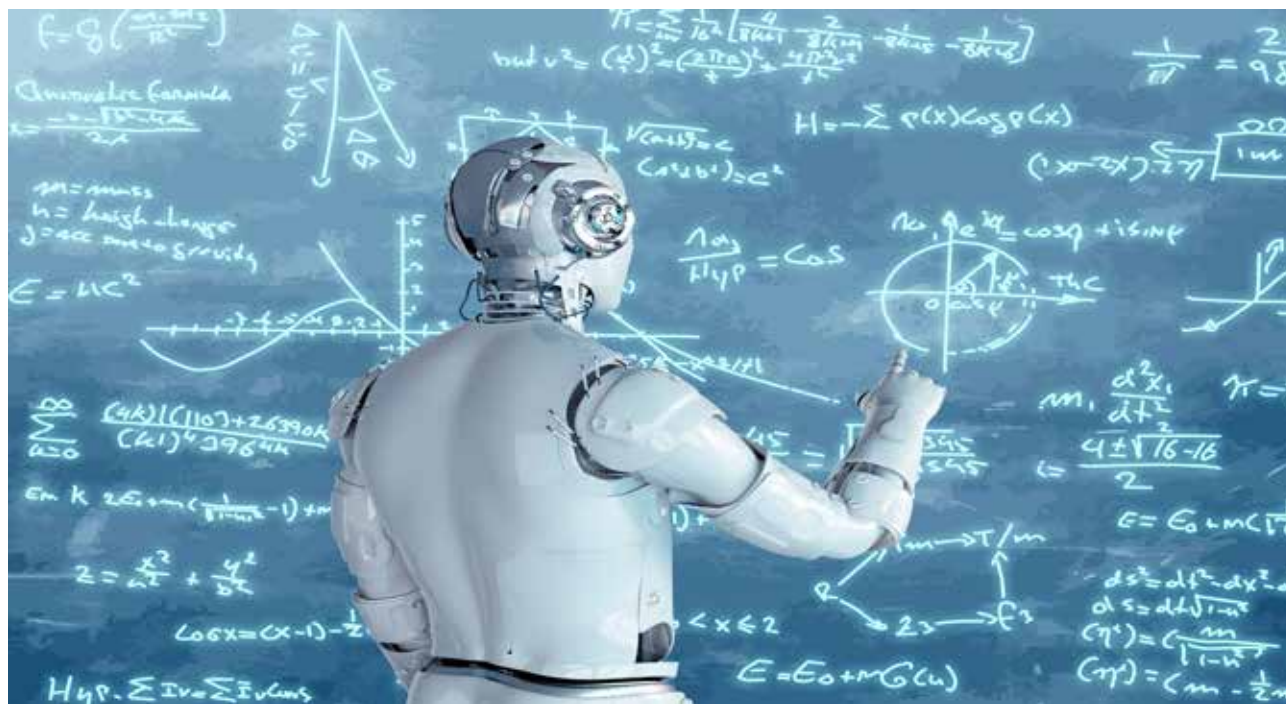


别让 AI 威胁人权

在机器人技术和人工智能日益塑造的技术领域里，人们该如何捍卫人权呢？

► 周晓综合报道



还记得科幻电影《2001：太空漫游》《银翼杀手》和《终结者》吗？在这些电影中，机器人和计算机系统以别样的形式（“杀手机器人”）呈现在电影荧幕上，给我们带来了全方位的视觉冲击。但如今，这种经典的科幻场景似乎与现实相去甚远。

现在，我们越来越多地使用自主和智能的计算技术来生活、工作和娱乐。这些系统包括具有独立推理和决策能力的软件和硬件。因此，它们既可以在工厂里为我们工作，也能决定我们是否可以获得房屋按揭贷款；它

们不仅能追踪并测量我们的活动和健康水平，还能清洁我们的客厅地板，修剪我们的草坪，它们已然成为人们生活中须臾不可缺少的重要部分。

显然，自主和智能系统已开始影响到我们的社会、经济、政治和私人生活的方方面面，包括平凡的日常生活。这其中不可避免地引发了一个令人担忧的问题：计算技术影响到每个人的人权，从生命权到隐私权、从言论自由到社会和经济权利。那么，在机器人技术和人工智能（AI）日益塑造的技术领域里，人们该如何捍卫人权呢？

AI 军备竞赛与人权

首先，人们真正担心的是，机器自主性的不断增加会危害人类的地位。因为当智能机器造成危害时，无论是在法律意义上还是在道德意义上，谁将被追究责任的问题还不明确，这无疑加剧了人们的担忧。

虽然还不能确定人们对人权问题的关注是否真的与“杀手机器人”有关，但是将智能机器投入到军事战争中的确引起了人们的无限担忧。

2017 年，美国空军四星上将 Paul

Selva 在回答关于美国国防部一道指令的相关问题时，就指出不要给军队配备人类可能无法控制的自动武器系统，并主张在恰当的位置维持“战争道德规则”。

国防部这道指令要求，当自动机器能够杀死敌军士兵时，人类操控者必须介入这一决策过程。Selva 表示，在战争中，军队在恰当的位置坚守道德准则非常重要，否则我们可能会向同胞放出一群自己都控制不了的机器人。

表达过类似担忧的还有 SpaceX 及特斯拉 CEO Elon Musk 和已故的著名天体物理学家 Stephen Hawking。他们曾在一封公开信中发出警告，不要让武器开发商使用人工智能。

他们在信中写道：“发动一场人工智能军备竞赛是一个糟糕的主意，应该禁止人类无法控制的攻击性自主武器。”

“我们需要主动、提前去监管，而不是被动地去反应。在我们需要对 AI 的监管进行反应时，已经太晚了，” Selva 说。

不难看出，人们真正应该担忧的

其实是，人类如何使用机器人和人工智能，以及它们在不公正和不平等的政治、军事、经济和社会背景下的部署。这在致命的自主武器系统(LAWS)方面尤为突出。

LAWS是能够在没有人的情况下，自主完成寻找目标、定位、击杀敌人这一整套工作的 AI 系统。只要将其派上战场，它们将自主辨别敌人并完成杀戮。这不禁引发众人的担心：如果 LAWS 发生故障失控，或是受到某些人的恶意操纵，可能会导致无法控制的灾难性后果。

很快，一些 AI、科学界领袖加入到反对研发 LAWS 的队列中。

今年 7 月 18 日，在瑞典斯德哥尔摩举办的 2018 国际人工智能联合会议上，包括 Elon Musk、谷歌 DeepMind 创始人在内的超过 2000 名 AI 学者共同签署《致命性自主武器宣言》，宣誓不参与 LAWS 的开发、研制工作。这是学界针对“杀人机器人”最大规模的一次集体发声。

在多方呼吁下，各国政府也开始认真考虑 LAWS 的威胁。2016 年 12 月，在联合国的《常规武器公约》审议大

会上，LAWS 首次被纳入讨论范围。不过，正式的 LAWS 国际公约诞生仍需时日。

今年 8 月，LAWS 将再次成为联合国各国讨论的议题。宣言签署者们希望，他们的行动能促进立法者颁布正式的国际协议。

AI 监控与人权

另外，人工智能还彻底改变了战争与监视行为之间的联系。

最近，机器人军备控制国际委员会(ICRAC)等组织对谷歌参与军事项目——“Maven 项目”，表示了强烈的反对。

据美国国防部网站介绍，“Maven 项目”于 2017 年 4 月由时任国防部副部长鲍勃·沃克发起，主要是将大数据、人工智能和机器学习等人工智能技术用于军事用途，首要任务是研发整合计算机视觉算法，帮助情报分析人员将海量全动态视频数据转变成具体作战所需的情报和决策建议，更好地利用数据支持作战。

按照美国国防部的计划，“Maven

项目”成立六个月就可以运行。为了满足该项目的迫切时间要求，国防部依靠自身显然并不足够。于是，它转而寻求与学界和业界的人工智能专家合作，以弥补军方和硅谷之间的技术差距。就这样，谷歌公司成为合作队伍中的一员。

然而，今年3月，当谷歌公司与“Maven项目”签署合同的消息爆出之后，超过3000名谷歌员工于4月初签署致谷歌首席执行官 Sundar Pichai 的公开信，建议谷歌退出“Maven项目”，信中写道：“谷歌不应该参与战争。”

4月19日，就在谷歌员工发出公开信半个月之后，来自谷歌、亚马逊、微软和IBM等科技公司的员工又发布了一封联合请愿信，声明称科技公司不应该参与到与战争有关的商业中去，因此谷歌应该中止与美国国防部的合作。

谷歌员工的抗议和多家科技公司员工的请愿又得到了来自人工智能学界、计算机学界和伦理学界专家学者的支持。机器人军备控制国际委员会呼吁谷歌承诺其收集的用户数据不会用于军事目的，随后又加入了谷歌员工对该公司参与该项目的抗议。如今，迫于内外压力，谷歌公司终于决定于2019年合同到期后，不再续约。

2013年，Edward Snowden 的爆料凸显了监控行为的覆盖范围及其深度。这次事件让人们认识到隐私权受到的可怕威胁，以及情报机构、政府机构和私营企业之间肆意的数据共享。然而，侵犯人权的类似事件并未就此停止。

最近，围绕剑桥分析公司利用 Facebook 等社交媒体平台获取个人数据的争议，继续引发严重的担忧，而这一次则是由于操纵和干预民主选举，从而损害了言论自由的权利。

与此同时，关键数据分析人士质疑与人工智能“白人问题”（white guy problem）相关的歧视性做法。因为在训练数据中，人工智能系统复制了现有的种族和性别成见，从而使在诸如警察、司法决策或就业等领域的歧视性做法根深蒂固。

模糊的界限

人工智能对人权的种种影响还在日益扩大。

2018年2月，26位来自牛津大学、剑桥大学、OpenAI、电子前沿基金会（一个非营利数字版权组织）、新美国安全中心（一家美国智库）等多个研究机构的研究人员发布了一份人工智能预警报告——《人工智能的恶意使用：预测、预防和缓解》，强调了计算技术对人权、物理安全、政治安全 and 数字安全的潜在威胁。

该报告预测了未来十年网络犯罪的快速增长、无人机的滥用、使用“机器人（bots）”操纵从选举到社交媒体等恶意利用人工智能的场景，号召政府及企业重视当前人工智能应用中的危险。

但是人们应该如何应对这些威胁呢？人权是否已经为机器人和人工智能时代做好了准备？

实际上，更新当今时代现有人权原则的努力正在持续之中。包括《联

合国工商业与人权的框架和指导原则》在内的努力，试图为数字时代和生命的未来研究所（FLI）制定的“阿西洛马人工智能原则”（Asilomar AI Principles）撰写一部《大宪章》，以确定道德研究的准则，坚持价值观并致力于人工智能长期而有益的发展。

毫无疑问，这些努力值得赞扬，但还远远不够。政府及其相关机构、政党以及私人公司，尤其是技术领先的科技公司，必须承诺在合乎道德伦理的情况下谨慎地使用人工智能。此外，还需要有效和可执行的立法控制。

当然，无论人们采取什么新措施，重要的是要承认我们的生活越来越多地与自动机器和智能系统紧密“纠缠”在一起。这种“纠缠”在医学研究和治疗、交通系统、社会医疗保健和保护环境的努力等方面增进了人类的福祉。

但是在其他领域，这种“纠缠”会引发令人担忧的前景。计算技术用于监视和跟踪人们的动作和运动状况，追踪人们的步伐、位置、健康、口味甚至人们的友谊。这一系列塑造了人们的行为并促使人们采取自我监督的做法，无形中削弱了人们的自由，并破坏了对人权的理念。

这就是关键所在：计算技术的双重使用能力模糊了善意与恶意行为之间的界限。更重要的是，计算技术与个体公民、国家及其机构以及私营企业之间不平等的权力关系有着密切而深刻的关联。如果脱离有效的国家与国际之间的相互制衡，它们将对人们的人权构成真正且令人担忧的威胁。■

（责编：倪伟波）