人民日报海外版

2024年6月3日 星期一

责编：卢泽华 邮箱：wesley1111@163.com

“融”观中国 前十位AI类APP需求规模同比增长近40倍，应用场景的丰富也增加了信息安全风险——

织好AI用户信息“防护网”

——“人工智能与信息保护”系列报道之一

本报记者 **姜忠奇**

对话交流、创作诗歌、编写代码、自动绘画……以大语言模型为代表的生成式人工智能（AI）快速发展，正前所未有地改变着人们的工作和生活。

随着生成式AI应用场景日益增多，个人信息保护也面临新的挑战：对数据的收集和挖掘，会不会侵犯个人隐私？怎样处理好个人信息保护与数据合理利用之间的关系？使用个人信息的边界在哪儿？本报记者就此进行了采访。

▲ 一名青年学生在安徽省阜阳市举办的网络安全宣传周展览会上体验VR展品。 **王 彪**摄（人民视觉）

▶ 内蒙古自治区呼和浩特市第二十七中学学生在绘制网络安全知识手抄报。 **丁根厚**摄（人民视觉）

“画像”背后藏风险

“对许多上班族来说，做PPT是一件让人头疼的事，经常要加班加点熬夜制作，有时还要额外付费聘请专业设计人员修改润色。”在北京某互联网公司工作的小潘告诉记者，现在通过生成式AI软件，只需确定主题、导入大纲、选择模版，不到1分钟就能生成一份专业又精美的PPT。

尽管效率明显提高，小潘却心有疑虑：PPT里提到的工作内容、行业数据、个人情况，会不会被储存在AI“大脑”中？如果涉及企业战略规划、商业秘密、关键技术，有没有被恶意窃取或泄露给第三方的可能？“目前，还没看到软件有类似‘一键擦除数据’的功能，无法确定这些信息的最终去向。”小潘说。

现实中，不管是生成PPT、制作简历还是自动美颜、智能聊天，都不可避免地要向生成式AI提供姓名、职业、人脸、声纹等个人信息，其中潜藏着不容忽视的隐私泄露风险。

比如，一款基于生成式AI技术的摄影软件一度走红网络。用户花9.9元，上传个人照片，选择喜欢的模版，便能获取一套由AI生成的写真集。在AI写真“刷屏”网络的同时，该软件也因存在滥用用户信息的嫌疑受到批评。有网友说：“向一款手机APP里传这么多照片，让我感到很不安。谁知道这些照片会不会被作为其他用途？”后来，研发团队致歉，承诺上传照片只用于数字写真制作，制作完成后，照片也将自动删除。

出于对新兴技术的好奇，消费者往往在不知不觉中将个人隐私“透露”给AI。反过来，凭借获取到的信息，通过强大的数据整合、处理能力，生成式AI也能判断出个人的身份特征和行为习惯。有研究发现，聊天机器人可以从日常对话中精准提炼关键信息，了解其购物习惯、个人兴趣乃至个人观念，从而为用户精准“画像”。

“生成式AI具有高度智能化特征，借助其出色的内容理解和学习能力，从海量信息中获取个人隐私，甚至可能通过对话‘诱导’用户打开‘心房’。”西南政法大学科技法学研究院副院长郑志峰告诉记者，这意味着生成式AI十分“懂”你，在侵害个人隐私方面更加高效、隐蔽，不易察觉。

用户数据成“养料”

“我们将对您主动上传的文档材料进行脱敏处理后作为AI训练基础材料使用……”前段时间，某办公软件隐私政策中的条款被质疑滥用用户信息。随后，该软件发布声明，保证用户文档不再被用于AI训练目的。

啥是AI训练？“对生成式AI来说，离不开算法、算力、数据三要素。其中，数据是AI的‘养料’，‘投喂’得越多，AI越‘聪明’。”郑志峰介绍，这决定了生成式AI必须尽可能多地采集、处理和使用各种数据，为大模型成长提供充足“营养”。

当前，各类数据采集无时不有、无处不在，几乎每个人都被置于数字化空间之中，个人隐私极易以数据的形式被存储、复制、传播。在所有数据类型中，个人信息能反映个人行为、偏好、行动轨迹等，是最有价值的数据类型之一。

“许多企业利用用户数据优化产品和服务、定向投放广告或开展经营活动。”中国社会科学院法学研究所网络与信息法研究室副主任周辉说，“有的企业片面关注数据价值，忽视了隐私保护和数据安全，可能出现违法收集、滥用用户信息等行为。”

生成式AI的快速发展和应用门槛的大幅降低增加了隐私泄露风险。研究报告显示，截至2024年1月，前十位AI类APP需求规模同比增长近40倍。生成式AI正加速渗透人们生活的方方面面。然而，应用场景日益丰富的同时，违规收集、使用用户信息的花样也变多了。

借助生成式AI的换脸和拟声技术进行远程视频诈骗就是一例。有媒体报道，郭先生接到“好友”的视频通话，对方声称需要保证金用以项目竞标，想借郭先生公司账户“走个账”。出自对“好友”的信任，郭先生陆续给对方转账共计430万元。等郭先生再次联系好友时才发现自己被骗。

“随着AI生成内容愈发逼真，传统内容审核机制和安全防护手段已不足以适应新形势的要求。”中国信通院人工智能研究所高级业务主管呼娜英说，生成式AI技术愈加成熟，用户在保护好个人隐私的同时，也需要提升对新型网络犯罪的辨别和防范能力。

给AI产品贴“标记”

怎样才能满足生成式AI的“胃口”，同时又保护好个人隐私呢？处理好两者关系，既需要法律法规的约束，也离不开技术手段的支撑。

去年7月，国家网信办等七部门联合发布《生成式人工智能服务管理暂行办法》，多处提及个人信息保护问题。强调生成式AI服务提供者“对使用者的输入信息和使用记录应当依法履行保护义务，不得收集非必要个人信息，不得非法留存能够识别使用者身份的输入信息和使用记录，不得非法向他人提供使用者的输入信息和使用记录”等。

“中国目前有关个人隐私保护的法律法规比较完善。未来需要根据生成式AI技术特点进一步构建具体的权利、义务和责任规则。”郑志峰举例说，AI的风险等级是多样的，不同AI算法对于个人隐私的侵害程度不同，需要有针对性的具体措施。

通过技术手段识别、阻断和追溯生成式AI生成侵犯个人隐私的有害信息，也是人工智能治理实践中的重要措施之一。根据《互联网信息服务深度合成管理规定》，深度合成服务提供者对使用其服务生成或者编辑的信息内容，应当采取技术措施添加不影响用户使用的标识。

“通俗地说，就是给生成式AI的产品和服务贴上统一的‘标记’，提示用户该内容由AI合成，也有助于监管部门管理。”呼娜英介绍，由于显式标识容易被裁剪或删除，还需进一步探索隐式标识的解决方案。

尽管生成式AI在一定程度上存在隐私泄露风险，但并不意味着它与信息保护之间是完全对立的，关键是要找到发展与规范之间的平衡点，在生活越来越“智能”的同时，守护好个人隐私安全，推动生成式AI健康可持续发展。

新媒视点

让AI用户更有安全感

卢泽华

数据和人工智能（AI）之间的关系是什么？有人作了个形象比喻：就像煤炭之于蒸汽机，电能之于灯泡，汽油之于汽车。

的确，几乎所有形式的AI都需要大量训练数据。要想让AI更加“懂你”，就必须收集和分析你的个人信息，这是AI深度学习的原材料，也是其“思考”和“决策”的依据。

这引发了公众对个人信息安全的担忧。就拿一个常见现象来说，当你第一次打开一款APP时，总会弹出一份冗长的“数据收集和隐私保护协议”。在绝大多数用户看来，完整读完这份协议是一道难题，遑论读懂。然而，如果你不点“已完成阅读并同意”，就无法使用APP。“说是给了用户选择权，但真正的选项只有一个，这不是霸王条款吗？”笔者听到过不少这样的抱怨。

大量调查显示，多数用户在签署各类智能平台收集数据的“同意书”时，并不清楚自己在同意什么。随着AI技术被不断应用于各类手机APP，各大互联网平台纷纷更新隐私保护政策，告知用户将使用其个人信息以支持人工智能的开发和优化。对用户信息的收集都包括什么呢？有人作了一些梳理，涵盖通讯录、相册、定位等，甚至我们的声音、指纹、脸部特征也在网罗之列。

这还只限于手机应用场景。有人对未来信息泄露风险作出更大胆的预测：只要你进入商场、饭店等商业公共场合，你的脸，你的声音，你的衣服颜色，你的兴趣爱好和行为习惯，都将暴露在空气中。在AI眼里，你就是一个行走的“数据群”。试想，一旦这种情形变为现实，将给个人信息安全带来多大的隐患？

从全球范围来看，人工智能侵犯隐私的案例已有不少。例如，一批匿名人士曾向世界最大的一家人工智能巨头发起集体诉讼，指责它从互联网上窃取和挪用了大量个人数据和信息来训练AI工具。起诉书称其“在用户不知情的情况下，使用这些‘窃取’来的信息”。不久前，北京互联网法院宣判了中国首例AI声音侵权案，原告因其声音被AI技术模仿并商业化使用而获得胜诉。

中国有句老话：“晴带雨伞，饱带干粮。”只有未雨绸缪，才能为AI长远发展奠定坚实基础。解决AI信息保护问题的关键，是如何合理规范使用用户数据。通俗来讲，就是让AI技术既在提供智能服务方面“更懂你”，也要在个人信息识别上“不懂你”，掌握两者之间的平衡。

在AI服务“更懂你”方面，业界已经做了不少努力；在让AI“不懂你”方面，还要继续深化探索。目前，中国已出台了《生成式人工智能服务管理暂行办法》等规定，明确提出AI应用必须保护商业秘密、个人隐私等不受侵犯。同时，进一步从服务规范和法律责任等具体方面对生成式人工智能服务作了规范。下一步，应继续在顶层设计上跟进AI技术发展，确立信息数据的收集、使用、消除等一系列具体条款。在此基础上，还要在提升用户素养上下功夫。社会需要普及各类信息保护的基本知识，提升用户的安全意识。

有这么一个故事：一个年轻人和一位长者过桥，年轻人问，我们走过时并没有扶桥两旁的护栏，要护栏有什么用呢？长者说，没有护栏，我们还能如此安然地从桥上经过吗？护栏给我们的不仅是安全，更是安全感。

当下的AI行业，正需要这样的“护栏”。

▲ 观众在福建省福州市举办的网络安全博览会上参观。 新华社记者 **姜克红**摄

◀ 在海南省海口市举办的第四届消博会上，观众与AI人形智能服务机器人互动。 新华社记者 **郭 程**摄

前沿动态▶▶▶

中国气象局发布人工智能气象预报大模型示范计划

本报北京电（记者**李红梅**）中国气象局在日前举办的第七届数字中国建设峰会·数字气象分论坛上发布人工智能气象预报大模型示范计划。

参与示范计划的人工智能气象预报大模型将使用中国气象局提供的实时实况分析数据作为输入场，制作未来0至15天的气象预报。

《知识产权保护体系建设工程实施方案》印发

据新华社北京电（记者**宋晨**）到2027年，知识产权保护体系和保护能力现代化建设迈出实质性步伐，知识产权法律法规更加全面系统；到2035年，知识产权保护体系和保护能力现代化基本实现……国家知识产权局日前联合多部门制定《知识产权保护体系建设工程实施方案》。

据悉，这份方案旨在加快建设支撑国际一流营商环境的知识产权保护体系，助力推动经济高质量发展。