中国科学院院士、清华大学人工智能研究院院长张钹:

从技术层面解决人工智能安全问题

当前,由人工智能引领的新一轮科技革命和产业变革方兴未艾。在移动互联网、大数据、超级计算、传感网、脑科学等新理论新技术驱动下,人工智能呈现深度学习、跨界融合、人机协同、群智开放、自主操控等新特征,正在对经济发展、社会进步、全球治理等方面产生重大而深远的影响。当前,我国人工智能产业发展处于什么阶段?如何看待当前的发展态势?未来应如何引导产业安全有序健康发展?围绕这些问题,近日,中国科学院院士、清华大学人工智能研究院院长张钹谈了自己的见解。

近年来,中央在多次重要会议中对人工智能的重要性和发展前景进行阐述。加快发展新一代人工智能是我们赢得全球科技竞争主动权的重要战略抓手。张钹表示,当前中国在人工智能产业发展方面与国际先进水平的差距并不是很大,但企业质量有待与数量"看齐",在拓宽应用场景、加强数据安全和算法安全等方面加大功夫。他提出"第三代人工智能"的概念,并表示希望通过政策和法规之外的技术途径解决人工智能的安全问题。

第三代人工智能的重点是什么?张钹表示,必须充分利用知识、数据、算力和算法四要素,让机器深度学习,做到随机应变、举一反三,才能不断推进人工智能向前发展。

目前,政府管理、技术识别以及数据监测中,都大量运用着人工智能技术。人工智能技术一旦被误用和滥用,会产生很严重的后果。为什么是这样呢?张钹列举了三点原因。

第一,所有技术都存在的问题,就是技术的两面性。人工智能也不例外,人工智能既可以用来造福人类,也可以用来伤害人类。举一个简单的例子,语音和图像的合成,利用这两项技术可以合成跟人类一样惟妙惟肖的机器人,让它跟人类进行交互、进行聊天,达到很好的用户体验。但也可以用这两项技术合成虚假的新闻、视频去欺骗公众,误导舆论,造成不良的社会影响。人工智能技术还有另外两点特殊的地方,导致人工智能的治理问题更为严重,也更加重要。

第二,现在的人工智能技术非常脆弱。换句话讲它非常容易被攻击、被欺骗。

第三,数据安全问题。其实人工智能依赖大量的数据,确实数据可以帮我们很大的忙,但数据也存在大量需要解决的问题,比如隐私问题、知识产权问题、偏见问题、污染问题、存在错误等等,这些同样会造成很大的危险。

"举一个非常简单的例子,我们用计算机去识别汽车,汽车后头有商标,正常情况都能识别。但我们只要在商标上加一点点噪声,这个噪声人用肉眼根本看不出来,但是就可以误导计算机'看不到'(识别不到)汽车,这就是新出现的'隐身术',这个非常危险,我们通常用这种计算机视觉识别技术去监测车辆,人家只要在车上搞一点小的噪声图案就可以让计算机看不到它,这就是一个人工智能脆弱性的表现。"张钹表示。

同样的,在视频监控场景里,如果用计算机来监测电网周遭环境的危险情况,只要在这个地方布置一点很简单的噪声干扰,就可以让计算机误认为这个地方有火情,或者真的发生了火情,也可以添加一点小的干扰进去让计算机认为没有火情。人工智能的滥用会对生活造成极大的威胁。

那我们应该怎么来解决人工智能安全性的问题呢?张钹认为必须从两个方面下手,一个就是其他专家发言里谈的比较多的,去治理。"但是我们不能用治理来限制人工智能的发展,唯一的办法还是在于发展;另外一个问题,我们该如何发展它?那就是要发展安全、可靠、可信、可扩展的人工智能技术。"张钹说。

针对第三代人工智能,张钹主要开展了两项工作,一项工作是已经发布的一个计算平台——One-Flow,现在市面上主流的框架平台,就是TensorFlow和PyTorch,OneFlow的可解释性比别的好,有很多优点,现在已经把应用在交通、医疗、智能制造等很多领域。

最近他们还发布了两个产品,一个叫RealSecure,主要做隐私保护,如何安全地利用数据,数据里面有隐私信息,如何从计算层面进行保护,实现可追溯,如何保持可追溯,如何让有偏见的数据清除掉,这是我们在这个方面做的一款产品。另外一个产品是RealSafe2.0,可以抵御攻击。

"要提出一系列治理的政策、方针,将技术的发展、治理的方式与时俱进,两者结合起来。"张钹表示,这是一个长期任务,因为人们不可能发展出一个绝对安全的人工智能算法,在发展过程中,技术要跟治理两者相辅相成,共同前进。 **人民**

360集团董事长兼CEO周鸿祎:

数字化发展应面向政企安全

"做政企安全业务,是因为能力越大,责任越大。"1月28日,在创新大会2021上,360集团董事长兼CEO周鸿祎说道。周鸿祎在大会上表示,数字化代表整个人类的未来。在过去的20年中,互联网改变了人们的衣食住行。在下一个10年中,社会、企业、城市都会完成数字化。数字化的本质是软件定义世界,但软件最大的问题,只要是人写的软件,无论多么优秀都会在软件里留有漏洞。漏洞不可避免,有漏洞就会被攻击。

"未来数字化世界的安全和每个人都紧密关联。"周鸿祎指出,在数字化时代,你和朋友正坐在无人驾驶的汽车上,吃着火锅唱着歌,突然车在高速上停下来了,这样的后果不堪设想。正因为如此,未来的网络安全需要被重视。

不久前,国际货币基金组织(IMF)发布了《世界经济展望陈述》,其中有两个核心数字:一个数字是"-4.4%",IMF估计2020年全球经济将萎缩4.4%;另一个数字是"+1.9%",中国经济将增加1.9%。这意味着,中国将成为2020年唯一正增长的主要经济体。无疑,中国经济在疫情之年释放出了独特的韧性,而这个韧性离不开数字化发展。

"发展数字化、建设数字中国已经上升为国家战略,并从上到下形成了高度一致的共识。"周鸿祎提到一组数据,2020上半年,在GDP整体下降1.6%的情况下,数字经济仍然实现14.5%的正增长,为恢复社会运转、经济复苏注入了强心剂。

数字化的这种韧性"补位"彰显了"危机中育新机、变局中开新局"的强大力量。在这种力量下,一切都将实现数据化和数字化,包括工厂、电网、机器,乃至家庭、城市,从而创造出一个与现实世界密不可分、相互映射的数字世界,并最终实现数字世界和现实世界的融合,带动人类社会迈入数字孪生时代。

周鸿祎将互联网上半场概括为人们生活方式的信息化、网络化,互联网下半场则是政府部门和传统企业的数字化,前者将实现新型智慧城市、城市大脑,后者将实现产业互联网。"未来5~10年,所有企业都将是数字化企业,所有经济都是数字经济。这其中孕育着无数机遇,但机遇也意味着挑战,数字化面临的最大挑战是网络安全。"

随着数字化的广度和深度不断扩大,安全风险也将不断加剧。可以说,数字化重新定义了网络安全,而网络安全也成为数字化的根基。此时,我们需要形成另一种共识,即数字化时代整体的安全共识。

周鸿祎表示,过去10年,360通过服务C端,积累了国内最大的安全大数据,也培养了东半球最大的白帽黑客团队。这样的资源和能力,给予了360新的发展契机和社会使命。对于360而言,已经积累了足够多的安全大数据、安全专家、威胁情报、白帽黑客团队。面对未来的网络安全挑战,360必须要做政企安全,应对可能出现的网络攻击,保护企业、社会和国家的网络安全。"这是一件舍我其谁、非我莫属的工作。"周鸿祎说道。

作为中国网络安全的坚定守护者,360经过15年发展已经成为亿万用户心中网络安全的代名词。在15年的发展历程中,360走出一条独有的"互联网+安全"双轮驱动道路,在保护亿万用户安全的同时,也成为黑客们无法绕过的一道安全屏障。

一方面积累了世界规模最大的安全大数据,奠定了安全大脑的基础;另一方面,每年在安全研发、人才培养中投入超20亿元,造就了深厚的安全能力。依托强大的安全能力,360在过去数年帮助国家发现了44个其他国家背景的高级黑客组织,监测到2700多次对中国的国家级网络攻击。

进入数字化时代,360以网络安全和数字城市为发展主线,开创了面向城市、行业和企业的服务赋能新模式。打造了以安全大脑为核心的新一代网络安全能力体系,并在重庆、天津、青岛、鹤壁、苏州等地相继落地,树立了标志性的城市安全服务案例。

周鸿祎强调,做企业不能把挣钱当成唯一出发点。"任何时候,企业和团队都要有战斗精神,要去寻找属于自己的荣誉感和成就感。帮助别人解决问题,才是一个企业应有的使命。"

对于周鸿祎和360而言,政企安全是时代和社会赋予他们的使命,这一事业目前才刚刚起步。在周鸿祎的带领下,360正以新时代网络安全运营商的身份,把过去积累的安全能力,转化成面向城市、行业赋能的,以安全大脑为核心的安全能力体系。未来还有很多目标等着周鸿祎和360一起去实现。 **环球**