

AI训练,数据纯净至关重要

□ 陈杰

热点观察

最近,多款知名品牌儿童智能手表因AI助手一本正经地“胡说八道”引发热议,也让不少家长深感担忧。

儿童智能手表为何会变成蹩脚小丑?从相关品牌方的回应来看,接连出现问题的儿童智能手表多为没有装入自研大模型的旧版本,其回答均抓取自互联网上的公开信息。

似乎也并非品牌方有意为之,但自研模型也好,网上公开信息也罢,既然用于AI抓取及训练,数据的确定性和可信度也该有所把控。数据之于AI,就如同“灵魂源泉”,其纯净与否决定着相关产品的性能和表现。

对于儿童智能手表来说,若AI助手训练所用的数据被污染,就可能导致

错误的信息输出。家长们担心的是,孩子正处于心智尚未成熟的阶段,对外界信息的辨别能力较弱,往往会轻易相信手表所提供的信息。一旦这些内容不准确或不恰当,可能会对孩子的认知发展产生极其不良影响,甚至可能引发孩子的恐惧或困惑。

儿童智能手表接连翻车,无疑也给AI助手及大模型产品敲响了警钟。

普通搜索软件尚能提供多页答案供筛选过滤,“一对一”的AI对话模式就带有明显的社交和知识传播属性了,不论是用作抓取还是训练的数据,准确性肯定都马虎不得。

ChatGPT诞生之初,有不少行业人士就担心其会成为成本和门槛极低的造谣工具。如今来看,现实似乎正逐步印证着这种隐忧。

好在,我国相关部门去年发布了

《生成式人工智能服务管理暂行办法》,在全球范围内率先开展立法,对数据使用、算法设计、价值导向提出要求,为行业的健康发展提供了支撑。

纯净的AI训练用数据,意味着数据集中不会有噪声、错误、偏见或不一致性,不会误导AI模型的学习过程,导致模型做出不准确或荒谬的预测;纯净的数据还有助于AI模型更准确地捕捉数据中的模式和规律,从而提高其决策能力和泛化能力,能够更准确地理解现实世界,在面对新情况时做出合理、有效的判断。

当然,再纯净的数据也无法保证AI不“胡说八道”。因为数据之外,AI的决策能力还受到算法设计、模型复杂度、训练策略等多种不确定因素的影响。在这些技术还没有得到创新性的升级和迭代前,AI训练用数据的纯净性,就显得尤为重要了。

AI骚扰电话为何让人防不胜防

□ 王鹏

“您终于接电话了,我是××银行助贷中心专员……”

“您好,我们专门为您量身定制了××海景房项目……”

接通电话听到这样的话术,不少人就能判断出是推销或骚扰电话。而让人们困惑的是,在AI技术的加持之下,这些骚扰电话已然全权由AI操控,让人难以通过常规的标记、举报或拉黑手段彻底摆脱。

前几天,关于“AI骚扰电话一天可以打上万个”的讨论也一度被推上网络热搜,引发了公众的共鸣和社会的高度关注。

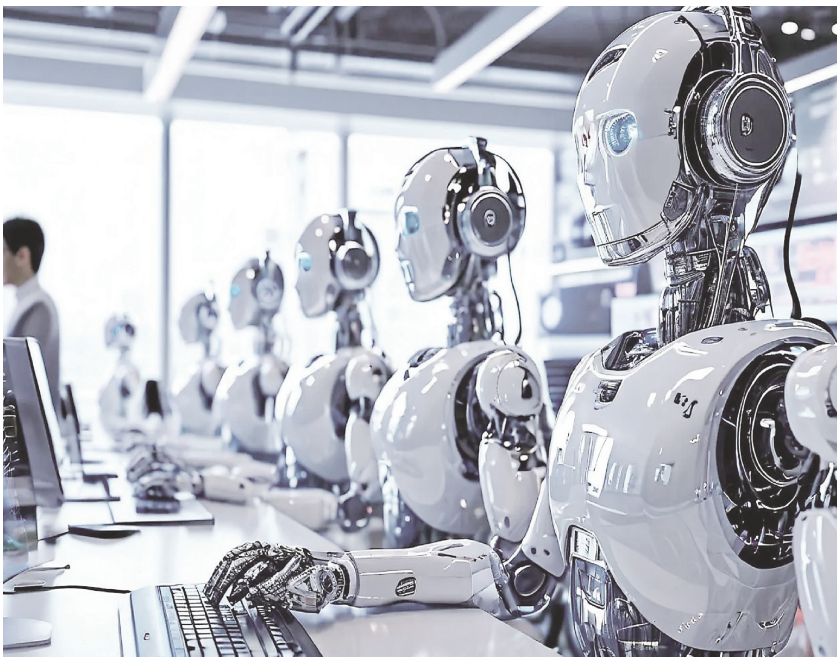
骚扰技术升级了

AI骚扰电话大多来自于不同商家使用的智能电销平台,内容基本上都是涉及服务或营销。据媒体报道,这些AI电销系统能够“不知疲倦”地每天拨打上万通电话,许多公司为了追求成本和效率,转而依赖其对消费者进行推销和回访服务。

不过,看似高科技的手段不仅没有提升消费者的体验,反而激起了公众强烈反感。AI技术在电话营销领域的滥用,已经成为一个普遍且严重的问题,让原本的骚扰电话变成了更加让人讨厌的骚扰“轰炸”。

AI骚扰电话之所以让人们防不胜防,无外乎以下几点原因。一是AI可以快速、大量拨打电话,突破了人工拨打的时间和数量限制;二是AI可以模仿真人的语气和声音,具有一定的迷惑性,让人在接听初期难以分辨;三是AI骚扰电话的来源广泛且难以追踪,背后的不法分子往往利用电话号码池和虚拟号码等手段,让常用的拦截过滤软件和拉黑手段都难以防范。

原本,电话营销及其使用的号码池就长期存在,也一直是电信领域的投诉重点和治理难点。如今,AI技术的加持更是大幅度节省了人力成本,客观上也造成了骚扰电话的更加疯狂和无孔不入。



AI技术已经广泛应用于电话服务及销售领域。AI制图

公众该如何应对

AI骚扰电话严重侵扰了人们的生活,应对还需从多方面着手。

首先,需加强防范意识。不轻易在陌生电话中透露个人信息,包括姓名、身份证号、银行卡号等。当接到不明来源的电话时,保持谨慎,不要轻易回答问题或按对方指示操作。

其次,要合理运用技术手段。安装防骚扰软件能有效减少AI语音电话骚扰。同时,开启手机内置的骚扰拦截功能,设置拦截特定号码段或被标记为骚扰的电话。

此外,应积极向相关部门举报。向电信运营商反映频繁接到骚扰电话的情况,运营商可能会采取相应处理措施。也可以向通信管理部门、消费者协会等机构投诉举报,推动相关部门加强监管和打击力度。

《消费者权益保护法实施条例》第二十三条明确规定,经营者应当尊重消费者的意愿,不得对消费者进行任何形

式的骚扰;《中华人民共和国网络安全法》也规定,任何个人和组织不得从事非法侵入他人网络、干扰他人网络正常功能、窃取网络数据等危害网络安全的活动。这些,都为公众对抗机器人电话骚扰提供了坚实的法律后盾。

很多人关心,难道不能用AI来治理AI骚扰电话吗?

面对疯狂的AI电话骚扰,利用AI加以治理确实是一种可能的解决方案。通过AI的深度学习与分析能力,可以识别出骚扰电话的拨打规律与特征,实现精准标记与阻断。不过,AI反骚扰技术尚在发展中,目前还存在着误判和漏判现象,可靠性还有待提升。

对付AI骚扰电话,我们不能单纯依赖AI技术实现“用魔法打败魔法”,应该从加强个人的防范意识、合理运用拦截技术,以及向通信管理部门投诉举报等方面努力,共同抵制AI骚扰电话,维护通信环境的安全与宁静。

(作者系北京市社会科学院副研究员、大数据业务分析师)

资讯品读

全球首个 全域服务地图上线

高德地图日前独家上线北斗卫星无网导航和卫星求救功能,服务覆盖至无信号区、无路区以及无人区,成为全球首个全域服务地图。当用户进入无路区时,地图会提示全程记录足迹,返程时用户可选择沿足迹原路返回;途中遇危险,卫星求救功能还能发出包含精确定位的有效求救信息,救援力量能够获知求救者准确位置和轨迹,大幅提升救援效率。

此外,在应急管理部国家安研院、中国安全生产科学研究院指导下,高德地图还联合蓝天救援队、曙光救援同盟、绿舟救援队、中国汽车工业协会越野分会等机构,成立国内首个基于地图平台的“救援互助联盟”,集结专业救援力量,共同提升野外救援效率。

点评:即便不会去到无人区、无路区等极限环境,我们日常的长途自驾也有可能途经没有网络信号的区域,这才是全域服务地图最大的服务场景。

微信和iPhone16 “二选一”会发生吗

这几天,苹果iPhone16可能不支持微信的消息刷屏了。

双方的争端,其实是基于所谓的“苹果税”。在苹果的生态中,应用程序的下载渠道只有唯一的App Store,并且只能通过苹果支付系统进行应用内购,苹果也会抽取事先约定的利益分成。

不过,微信的部分小程序应用成功绕开了苹果的支付系统。苹果认为,微信是在挑战自己既成的规则。双方也从早前的暗自较量逐步演变成如今公开较真。目前,苹果商店的微信已经两个月没更新了。本月中旬苹果iPhone16即将发布,让用户“二选一”的情况会发生吗?

点评:苹果和微信互不相让的背后,是以双方庞大且很有重合度的用户群体作为支撑。但说到底,这种争端是纯商业的,合则双赢,最终应该不至于出现“二选一”的情况。

扫地机器人 也可以翻越台阶了

你还在为自家扫地机器人无法跨越种种障碍而烦恼?近日,一家科技公司推出搭载行业首创仿生双机械足高越障系统的扫地机器人新品,它通过双节摆臂系统,将轮子的半径增大一倍,来提高越障性能。当遇到高门槛时,机器会感知判断障碍高度是否可以越过,然后主机通过驱动轮仿生双机械足抬高机身越过高门槛。

拥有了高越障技术,扫地机器人便能够更加自如地在复杂多变的家庭环境中工作,直接提高了清洁效率和覆盖率,减少了人工干预,从而满足了消费者对自动化和智能化清洁工具的期待。

点评:近些年,扫地机器人智能化水平有了显著提升。高越障技术的突破,有望让扫地机器人从一种简单的家用电器,逐渐演变成为智能家居生态系统中的重要一员。