



扫一扫发现更多



面向未来的人工智能法

AI换脸、AI拟声等技术被不法分子利用,冒充熟人进行诈骗,这样的案例你可能没遭遇过,但大概率在“相亲相爱一家人”群中看到过。

前两天的消息,据《纽约时报》报道,个性化聊天机器人创企Character.AI因卷入一则青少年自杀案而遭起诉。14岁的少年与Character.AI程序中的虚拟人物——《权力的游戏》中的“龙妈”角色产生了情感依恋,少年变得与世隔绝,最终于今年2月自杀。

未来已经显露端倪,我们需要应对人工智能技术快速发展带来的法律挑战。

当然发展是第一要务,不发展会导致最大的不安全。人工智能产

业,我国目前处于“领先的追赶者”位置。本期《读+》记者专访中国政法大学教授张凌寒,作为联合国人工智能高层咨询机构中方成员、中国人工智能产业发展联盟政策法规工作组组长,她告诉我们,我国人工智能立法的基本逻辑:大力支持产业发展,同时兼顾切实的安全需求。

与传统立法相比较,人工智能立法面临技术飞速发展带来的调整对象不确定性、技术产业迭代带来的社会关系不确定性,风险种类和程度的不可预见性等。立法不仅要解决当下的问题,更要充分考虑对长远的确定性。

面向未来的人工智能法,既要准备防范和化解风险,又要增加社会对技术发展的信任。

王永芳

2024年10月29日 星期二 主编 王永芳 责编:马梦娅 美编:职文胜 版式:邱巍 责校:刘明

生命中最特别的部分 专栏

记忆力好不好关键在于大脑内部模型塑造

读书

在人工智能立法领域发出中国声音

——专访中国政法大学教授、联合国高级别人工智能咨询机构中方专家张凌寒

□记者李煦

在武汉开会 历数人工智能法治“热点”

10月26日,“AI善治:中国人工智能法治研究回顾与展望研讨会”在中南财经政法大学召开,张凌寒作主题发言。

她首先表示,武汉是中国人工智能的产业高地,中南财经政法大学是中国的法学重镇,中南财经政法大学吴汉东教授2017年发表的《人工智能时代的制度安排与法律规制》,到现在为止被引用近2000次,标志着人工智能法治研究首次进入了法学研究的主战场和最前沿;而中南财经政法大学徐小奔副教授则是《人工智能法(学者建议稿)》起草专家组成员之一。

在会上,张凌寒梳理了从2017年到2023年间“中文社会科学引文索引”有关人工智能法治主题的1618篇文献,得出很多有意思的结论。

其中之一,是人工智能立法的热点转

移。2017年发文量最高的主题是机器人;2018年、2019年、2020年、2021年,连续4年,算法相关主题发文量最高;2022—2023年,最热的主题则是“元宇宙治理”。再比如,2017、2018、2019年,有大量的有关人工智能主体资格的讨论,很多学者明确反对给人工智能主体资格,到了2021年和2022年,“人工智能仅仅作为工具”已经逐渐达成了共识,但是在2023年这个讨论又一次热起来,说明ChatGPT等生成式人工智能使得这个问题需要重新考量。

张凌寒表示,人工智能法治的研究对象快速迭代,很多传统部门法学者因此诟病人工智能法治是“蹭热点”,也确实有很多主题随着技术热潮的过去而迅速衰落,所以分辨出人工智能法治的“真问题”,对相关学者来说非常重要。

国内外加紧布局 “人工智能法”呼之欲出

人工智能已经成为当今国际竞争的关键领域,人工智能的治理之争也日益凸显。

2023年初,生成式人工智能惊艳问世,欧美等国家密集展开治理活动,加快布局人工智能立法,力图引领国际规则制定。2023年,欧盟因通用人工智能特点迅速修改《人工智能法案》并达成初步协议;2023年10月30日,美国总统拜登签署《安全、稳定、可信的人工智能》行政令,以确保美国在AI发展和管理风险方面处于领先地位。我国早在2017年,国务院发布的《新一代人工智能发展规划》就提出了“三步走”立法规划,并在2023年6月6日发布的《国务院2023年度立法工作计划》中明确将《人工智能法》列入立法规划。2023年8月,我国提出的世界上第一个有关生成式人工智能的成文法《生成式人工智能服务管理暂行办法》生效实施,“暂行”二字意味着一部统一的、综合的《人工智能法》呼之欲出。

这里必须要提到一个名词“布鲁塞尔效应”:某方不一定技术领先,但是拥有市场,率先提出理念、率先创制规则,于是对产业拥有了超出边界的单边监管能力。提出这一名词的哥伦比亚大学教授布拉德福德写到,欧盟的规定决定了苹果手机的默认隐私设置,确定了巴西蜂蜜生产和喀麦隆可可农场使用杀虫剂的标准,还决定了中国乳品厂安装设备的标准。不少品牌

她主张“立总则”“廓主线”“留接口”

既然人工智能一日千里、快速迭代,那么如何识别其法治“真问题”呢?

张凌寒认为:近十年间,数据、算法、平台是我国人工智能治理的主要对象,平台则是人工智能治理的核心;但是2023年一年间,通用人工智能出现重大进步,大模型可广泛赋能千行百业,形成了“基础模型—专业模型—服务应用”的分层业态,改变了以平台责任为核心的制度体系。生成式人工智能大模型的通用性,使其在技术垂直类领域的应用也将快速多样化,既往监管模式面对各种新问题可能会应接不暇。我国的人工智能立法如果能够包容产业迭代带来的社会关系的不确定性,需要立法者拿出较高的智慧制定发展与监管政策。

她主张:以“总则式”为框架体例,以“廓主线”为主要内容,以“留接口”解决容量问题。

《人工智能法》立法不能抱持“毕其功于一役”的理念,试图去涵盖所有人工智能技术、要素和应用等不同维度的治理规则。“总则式”体例的《人工智能法》应确立人工智能立法的目的,人工智能技术、产品与服务必须遵循的基本原则,基本法律制度应遵循的一般性规则,表达我国人工智能治理的基本理念与基本立场。

《人工智能法》应以“廓主线”为主要内容。主线是指人工智能治理的基本框架脉络,包括一般原则、促进发展、权益保护、安全义务、特殊应用场景、监督管理、国际合作、法律责任等若干板块内容等。这样一条主线基本覆盖了人工智能治理的主要制度,能体现将人工智能作为生产方式和社会治理系统的综合治理思路。

《人工智能法》无论如何“大而全”都无法涵盖人工智能时代技术产业和社会生活方方面面的问题。人工智能领域已经深刻影响到了教育、医疗、金融、交通等领域,这一趋势未来还会继续延续。诸多社会生活领域的具体规则的制定,只能交给专门领域立法解决,《人工智能法》要留足“领域接口”。《人工智能法》立法中,还可预留授权的制度接口,将位阶较低的制度授权有关部门细化。

张凌寒自己有没有向人工智能提出过法律问题呢?她告诉记者:“我个人使用人工智能,我觉得对我们的科研工作来说,比如它们如果能帮忙,也只能做类似于最粗糙的小工的工作。因为我尝试问过它们一些我自己比较熟悉和了解的问题,它们吐出的答案,并没有什么新意,最起码在我比较熟悉的研究领域是这样的。”

访谈



张凌寒

读+:《人工智能法》将如何促进、推动中国人工智能的发展?

张凌寒:科技创新无法单独存在,科技创新需要外力的介入,需要法律为其搭建公共基础服务的保障性制度。

《人工智能法》中应设立“发展与促进”的专门板块,体系性地部署促进人工智能技术产业发展方面的制度。唯有通过《人工智能法》这一针对性强、位阶较高的专门立法,才能够制度化地解决一些问题;在法律层面明确国家和地方政府将人工智能发展纳入本级国民经济和社会发展规划,通过法律制度建立算力基础设施建设和算力资源调度利用的制度,协调人工智能技术与知识产权制度的冲突,生成式人工智能训练数据和生成物的保护,等等。

《人工智能法》的相关制度应充分考虑人

读+:中国的《人工智能法》无疑要“讲中国话”,这体现在什么地方?与欧美相比,我们的法律会有哪些明显不同?

张凌寒:各国关于人工智能的立法和政策,无不反映出各自国际生态位下相应的治理诉求。中国作为人工智能的重要参与国,中国的《人工智能法》将会集中体现中国立场和中国智慧。

首先,我国人工智能立法的基本逻辑必须适应我国独特的国际生态位。我国是人工智能产业中“领先的追赶者”。一方面,世界上有实力独立拥有自己人工智能技术与产业的国家寥寥可数,目前仅美国、中国和少数欧盟国家有力量进行生成式人工智能大模型的开发与训练。另一方面,我国技术水平距离美国仍有较大差距,具有迫切的追赶需求。美国高度警惕我国的技术追赶,并为此出台了大量针对中国人工智能领域的封锁、制裁和惩罚性举措。在激烈的国际竞争格局下,不发展是最大的不安全。基于此,我国人工智能立法既要体现出技术领先的世界大国对于人工智能技术发展风险的担当与责任,也要以支持和促进产业发展作为我国人工智能立法的主旨,始终兼顾发展与安全。此外,考

读+:世界上关于人工智能治理有哪些主要的声音?能否谈谈联合国高级别人工智能咨询机构的工作,中国以什么姿态参与人工智能国际治理,并贡献中国智慧呢?

张凌寒:目前,全球存在多元化的人工智能治理机制,提供了不同的声音,我在此选取几个有代表性的分类视角加以解释。首先,按照地区区分,中国、欧盟和美国占据了全球人工智能产业链中的重要生态位,其治理动向也受到广泛关注。中国采取“小、快、灵”的路径进行敏捷治理,出台的《生成式人工智能服务管理暂行办法》是世界上第一针对生成式人工智能的专门立法。欧盟的《人工智能法》则采用风险进路和人权基础,试图对人工智能进行全面和严格监管。美国则出于对自身全球技术领跑者的定位,强调在推动技术发展的同时确保全球领导力,暂时未在联邦层面展现出人工智能立法意图。

其次,从部门和行业的视角分析,大型科技公司也是人工智能治理中的重要声音,如谷歌、微软等,主张在政策制定中为创新留足空间,同时注重行业自律以推动安全和公平

人工智能技术在生产生活中的应用,设定合理的监管标准。人工智能技术与产业的长远有效发展,有赖于其充分地融入社会生产生活,成为推进社会发展的动力。为此,一方面需要反思现阶段对于法律责任的设置是否有利于保障产品和服务的安全顾虑。以法律责任的设置为例,如果生成式人工智能代表未来生产力的发展方向,其发展的社会成本即包括合理的法律责任。二十年前各国为鼓励网络产业发展,法律对网络服务提供者通常施以较轻的注意义务(注意义务是一种法律规范和要求,要求当事人预见自己的行为可能产生的损害后果,并采取合理的防范措施来避免这些后果的发生——读+注)。但后来随着产业的发展,法律对网络服务提供者设定的责任水平逐渐提高。如

虑到人工智能技术的跨国性和涉及国际竞争的复杂性,我国的人工智能立法还需充分考量法律制度的涉外溢出效应,将人工智能法打造成一张展现中国治理理念与智慧的名片。

其次,人工智能立法要坚守“人本主义”的主旨立场,这既是我国对人工智能治理的核心倡议,也是目前国际社会的共识。具体来说,要警惕人工智能时代,过度地追求技术理性使机器日益“人化”的同时,人也被“机器化”。“以人为本”理念是我国人工智能治理的基本理念。在去年10月发布的《全球人工智能治理倡议》中,我国在倡议首条即明确表示,“发展人工智能应坚持‘以人为本’理念……确保人工智能始终朝着有利于人类文明进步的方向发展”。这既是我国在人工智能发展中对世界的承诺,也是我国对处于人工智能时代的每个人的承诺。为充分贯彻落实这一理念,我国应在人工智能相关立法中充分考虑个体权益,关注和重视社会弱势群体权益保护,更加完整地构建人工智能时代个体权益保护的制度体系。

我国已经出台《中华人民共和国数据安全法》《个人信息保护法》等法律,《互联网信

中国人工智能法治体现中国立场

果我们相信生成式AI代表着未来人工智能模式发展的重要技术路径,就应对大模型数据训练与其生成物的法律责任保持一定的宽容,因为未来生成式人工智能成功应用于社会生产与生活可以反过来惠及更多的行业和开发者。生成式人工智能的专业模型在发展早期,应设置一定的免责规则,甚至立法应当为其设计精巧的避风港责任,降低企业履行注意义务的成本。

为此,立法对人工智能的研发和提供行为设置义务的数量和范围要合理。尽管欧盟就《人工智能法案》游说亚洲国家效仿其在人工智能领域的领先立法,但多国保持了观望态度,担忧过于仓促、严厉的监管会扼杀技术创新。欧盟的学者也担忧法律义务和监管要求的数量规模如此之大,实施成本如此之高,其威慑作用可能导致研发活动转移到欧盟之外。

为人类治理人工智能,“中国担当”义不容辞

的人工智能发展;而学术界则普遍倡导透明性、可解释性和问责制,主张在技术治理中加强多方协作。

尽管不同国家、地区和区域性组织纷纷推进人工智能治理与合作,但联合国以其最广泛的全球合作性和包容性在人工智能国际治理中发挥不可替代的作用,可以说是最重要的声音之一。

作为联合国高级别人工智能咨询机构的成员,在过去一年多的时间里,我有幸全程参与了联合国人工智能全球治理的咨询过程和报告发布。

咨询机构成立后两个月内即发布了《临时报告》,概述了人工智能全球治理赤字、风险与挑战、机遇与推动因素、国际人工智能治理的指导原则和包容性在人工智能国际治理中发挥不可替代的作用。在《临时报告》的基础上,咨询机构与世界各地2000多名参与者进行了广泛协商,包括顶尖专家就关键问题进行18次深入讨论,覆盖全球所有地区的50多次协商会议,以及收集了来自150多个组织和100位个人的250多份书面材料。经过六个月的全球磋商与调研,今年9月咨询

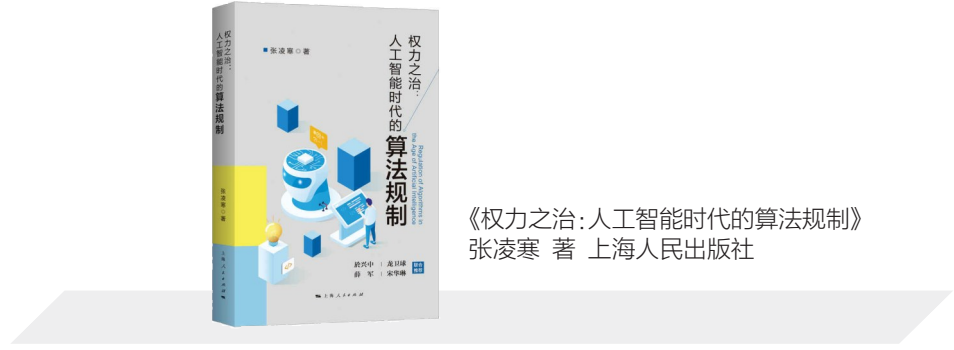
息服务算法推荐管理规定》《互联网信息服务深度合成管理规定》《生成式人工智能服务管理暂行办法》等部门规章,客观上,我国已经是世界范围内唯一一个系统建设覆盖政策规划、科技伦理、法律、行政法规、部门规章与技术标准全位阶、多层次人工智能治理制度的国家,特别是制定公布了世界上第一部实质生效的生成式人工智能治理法规。这种从具体到综合的立法模式,与欧盟立法人工智能治理体系从统筹规划至具体建立制度的发展思路迥异。

迄今为止,中国网络治理与技术治理中已经形成了一系列具有中国本土特色的制度。以未成年人网络保护制度为例,游戏“宵禁”制度与“未成年人模式”制度是目前世界独有的制度,也是被美国与欧洲业界和规划者广为关注和讨论的制度。又如算法推荐规制中的“一键退出”制度,这是中国规制者对互联网企业提出的独特要求。再如我国网信部门在2023年1月10日开始实施的《互联网信息服务深度合成管理规定》中对生成内容提出了“标识”的要求,且该制度已经落地;而2023年欧盟和美国在多边会议和监管文件中才刚刚提出这一制度。

机构正式发布题为《“为人类治理人工智能”(Governing AI for Humanity)”的最终报告。最终报告强调了人工智能全球治理的必要性,进一步分析了全球人工智能治理差距,提出了加强人工智能治理国际合作的7项建议。

在今年9月的“未来峰会”上,由全球领导人共同通过的《全球数字契约》吸收和采纳了咨询机构发布的《最终报告》中的咨询意见。

以“坚持统筹国内法治和涉外法治”为指导,中国的《人工智能法》也将承载着在国际竞争中发出中国声音的使命。作为联合国首个加强人工智能能力建设国际合作决议的提出者,中国通过“一带一路”、金砖国家等多边机制,努力推动人工智能治理的公平性与包容性,搭建弥合全球数字鸿沟的平台,并为新兴市场国家提供创新与治理并行的范式,主张技术发展的全球共享与共赢,推动国际规则朝着更加公平与普惠的方向演进。人工智能治理的“中国方案”有望在权益保护、能力建设、安全保障和促进公平发展等方面发挥积极影响,从而展现人工智能国际治理的中国担当与中国智慧。



《权力之治:人工智能时代的算法规制》张凌寒 著 上海人民出版社