

专家解读

协同安全与发展 为人工智能划“红线”

——解读《生成式人工智能服务管理暂行办法》

- ◆ 《办法》的出台正式拉开我国全面监管人工智能的大幕；
- ◆ “大数据杀熟”等利用算法搞不正当竞争被明令禁止；
- ◆ 对于人工智能中的算法“黑箱”问题,明确了分类分级监管原则；
- ◆ 在提供服务的过程中要防止产生各类歧视,推动生成式人工智能向善发展。

中国妇女报全媒体记者 徐阳晨

生成式人工智能技术,是指具有文本、图片、音频、视频等内容生成能力的模型及相关技术。近日,国家网信办联合国家发改委、教育部、科技部等多部门公布《生成式人工智能服务管理暂行办法》(以下简称《办法》),自今年8月15日起施行。该《办法》的出台既是促进生成式人工智能健康发展的重要要求,也是防范生成式人工智能服务风险的现实需要。中国妇女报全媒体记者邀请相关专业人士对《办法》进行解读。

在约束中探寻激励机制

“《办法》的公布是我国首个全面监管人工智能的制度,是我国规范数字化建设迈向制度化、法制化的重要一步。”内蒙古财经大学信息化建设与管理中心主任、高级实验师郭立志指出,《办法》从微观上有助于我国生成式人工智能健康发展和规范应用,从宏观上遵循了数字中国、网络强国的战略导向。当我国人工智能发展到一定的规模,数据积累为大模型场景落地带来实际可能性,相关的管理规则亟待完善。郭立志认为,在新技术迭代衍生,甚至颠覆产业模式的情况下,人工智能在滥用、侵权、欺诈等违法违规问题上风险也在加剧,政策监管并不是约束限制,而是为了在安全的“红线”中更好地支持AI发展。

《办法》总共分五章,分别对管理目标、



由人工智能引擎生成的视觉艺术作品。来源:上观新闻



在世界人工智能大会展区,元萝卜AI下棋机器人与观众对弈。 辛梦晨/摄

服务提供者、用户权益保护等人工智能相关方面做出明确规定。其中不乏诸多创新鼓励措施,如提出“支持行业组织、企业、教育和科研机构、公共文化机构、有关专业机构等在生成式人工智能技术创新、数据资源建设、转化应用、风险防范等方面开展协作”。

郭立志解释,上述规定意味着我国对各类研究机构全面敞开了人工智能的“大门”,允许各机构充分发挥各自研究优势,促进生成式人工智能开发效率。“可以这样说,《办法》中的具体措施是推动我国生成式人工智能深入应用的催化剂、助燃剂。”

切实压紧服务提供者责任

与互联网监管范式类似,生成式人工智

能的首要责任人是服务提供平台、发布监管方以及内容生产者。《办法》第七条对生成式人工智能服务提供者(以下简称提供者)提出了五点要求,包括合法来源的数据和基础模型;不得侵害他人依法享有的知识产权等。明确要求提供者在提供服务时必须充分考虑所使用数据、信息、模型等因素必须具有合法来源,不得侵害他人的知识产权。

对于人工智能中的算法“黑箱”问题,《办法》明确了分类分级监管的原则,一是明确了安全评估与算法备案要求,二是明确了信息披露要求。“监督人工智能的系统设计与运行方式,是打开人工智能算法‘黑箱’,增加透明度的有效方法。”郭立志认为,《办法》要求对算法机制机理和训练数据来源、规模、类型、标注规则都要予以说明,可使算法尽可能暴露在监管之下,真正地透明起来。他建议在数据标注环节,还要细化标注规则,对标注的来源、质量、精度精心评估,让数据标注成为信息生成判定的关键依据。

此外,“大数据杀熟”等现象备受网民诟病,利用算法搞不正当竞争也在《办法》中被明令禁止。“对于算法,必须是正向产生效益的经济服务行为才给予鼓励。”郭立志强调,《办法》要求在提供服务的过程中要防止产生各类歧视,是进一步压实平台责任、防止企业利用数据搞市场恶性竞争的重要举措,有利于“反数据垄断”,维护广大用户合法权益,推动生成式人工智能向善发展。



参观者在世界人工智能大会达闼展位与柔美人形智能服务机器人互动。 李心怡/摄

建立对底层逻辑的正确引导路径

“从实践中看,人工智能生成内容的主创者都是‘机器’,所以处罚对象要有明确细化。”伟岸纵横(大连)安全科技有限公司副总裁孙崇科指出,《办法》对训练数据的合规性进行了明确的界定,一旦在生成内容领域违法违规,生成式人工智能服务提供者可能需要改变整个的算法逻辑或业务模式,这对于当下野蛮生长的AI生态来说,既是挑战,也是利好。

鉴于生成式人工智能有主动学习的能力,因此生成内容是否违规,不仅受模型的预训练数据影响,还受用户输入内容的影响。孙崇科指出,《办法》的相关合规界定雏形逐渐形成,算法底层逻辑的正确引导路径正在培育。他认为未来把服务方角色的明确、算法合规的界定、内容“红线”等问题置于成熟的监管系统中,生成式人工智能才会在更多场景顺利落地。

如何在充分发挥生成式人工智能的强大服务能力的时候,又能有效遏制其带来的技术风险?郭立志认为,仅有一部《办法》仍“势单力薄”,“在国务院2023年度立法工作计划中,人工智能法赫然在列。显然我国对人工智能有着系统的治理理念,准备了一整套‘组合拳’。此次《办法》的公布,则正式拉开我国全面监管人工智能的大幕。”

本期关注

体育与性别平等之间的交叉赋能 相互促进,成为当今社会女性发展的“加速器”。

体育参与助力实现女性赋能

曹靖宜

体育是促进性别平等、助力女性发展的重要媒介。但长久以来,体育活动一直都被认为是男人的领域,运动场为男性提供了一个探索身份认同、男性气质与身体认知的场所。从世界范围来看,20世纪70年代以前,女性对于体育参与的无论是在广度、深度还是强度上都处于弱势。以奥运会为例,首届比赛中并没有出现女性的身影。女性运动员从1900年开始进入奥运会比赛中,直到2024年,奥运会才有可能实现运动员的性别平等——男性和女性运动员各占50%。

受历史文化的影响,在大型国际赛事以外,女性在职业体育和群众体育中的参与也低于男性。近年来,越来越多的国际组织、国家开始重视体育中的性别平等问题,以体育赋权女性被政策制定者与公众所共同关注。20世纪后半叶以来,体育作为促进女性发展的媒介,有了许多里程碑式事件。如1981年,国际奥委会首次迎来两位女成员;1990年,国际奥委会执行委员会中有了首位女性成员;自2000年起国际奥委会每年评选女性体育奖……我们欣喜地看到,体育与性

别平等之间的交叉赋能,相互促进,赋权女性后所产生的积极影响覆盖其发展的各个层面,成为当今社会女性发展的“加速器”。

首先,通过参与体育活动,女性们能从中习得如何开展团队工作以及如何提升心理韧性。在体育过程中,女性们还能搭建新的社会关系,并能避开家庭或社区中的暴力行为。体育帮助女性们理解她们的身体、增强自信,使她们敢于发声。这样的运动干预对青春期的女孩来说尤为有效,能帮助她们打破有关女性气质的性别刻板印象。

其次,利用体育为女性赋权可以帮助解决更广泛范围内的性别不平等问题。体育活动自身的性别不平等情况反映了生活中的性别不平等,如男女同工不同酬、性别暴力、投资不足以及负面的陈规与社会规范等。对女性体育的整体投资能够改善女孩的运动机会,提升女子体育的地位,并为社会带来更广泛的利益,例如民众整体体质的增强、体育消费收入的提升。体育界在此类问题上的成功实践有助于推动社会层面的变革。除了女性之外,男性也能从此类活动中获益。开展对妇女和儿童友好的体育场活动,针对对妇女和女童的各种歧视采取宽容忍态度,有助于改

近日,第31届世界大学生夏季运动会女子50米蝶泳决赛在四川成都举行,中国队选手张雨霏摘得金牌。 张浪/摄

变业已过时的“更衣室”男子气概,为男性提供更舒适的体育参与体验。

最后,女子体育活动的大量涌现丰富了体育活动的内涵。女性参与为各类体育活动的表达提供了新的解释方式,给体育发展带来活力。例如,在女足现场,没有男足比赛时常见的不文明举动与嘘声,观众得以在舒适的环境下体验纯粹由竞技带来的快乐。同时,女性观众的加入也为各项体育活动的组织形式提出了新要求,体育比赛自身有潜力借此实现规则、文化上的突破。此外,优秀女运动员更能激发女性参与体育活动的动力,女性的参与也能为体育活动注入新的活力。

实验室研究表明,竞技性体育活动参与能提高人的自尊程度,而女性的自尊提



升程度显著高于男性,并且参与高强度体育的女性远强于不参与高强度体育活动的女性,而这一点在男性群体中并不显著。他们认为这能证明那些具有“男性气质”特质的女性自尊程度更高,这也侧面暗示女性相较男性更能从参与高强度体育活动中获益。

体育促进女性赋权的成果初现,女子体育的发展仍在进行时。在这个过程中,大型体育赛事和学校体育活动正逐渐实现人数上的男女平等,但社会既有的性别文化差异和薪酬上的男女不平等仍然阻碍着女性的体育参与,体育参与也尚未更直接地促进影响范围更广的性别平等社会运动。我们呼吁更多的实践者应该更多从组织和政策层面考虑,如何通过体育参与助力实现女性赋能。

图片新闻

体验馆里觅书香

暑假期间,小朋友和家长在浙江省嘉兴市图书馆元宇宙阅读体验馆通过环幕体感投影,学习山脉山峰知识。

金鹏/摄

创新评说

将综合性大赛与城市高质量发展紧密融合,全面加强城市软实力。

王雪莉

近日,第31届世界大学生夏季运动会(以下简称成都大运会)拉开帷幕,成都的火热就如这个夏天一样,持续“高温”,热点频现。令人惊艳的新建和改造的场馆,充满活力和巧思的体育运动服务综合体,各种沉浸式体验成都文化的活动,还有不断丰富老百姓身边的设施和场地,都直接反映了随着申办筹办成都大运会,随着赛事名城的目标制定和实施,成都这座城市正处在运动促进与城市发展双螺旋上升的轨道上。如何更好发挥成都大运会对城市发展的赋能效应?笔者有如下初步的思考。

围绕“人”这个中心,精准施策,形成杠杆效应。具体来说,首先全民健身要持续推进,在参与的范围和活动类型方面应该有更高目标。这不仅是人民美好生活的基石,也是众多体育设施和场地在赛后可以得到更充分利用的基础,还能持续扩大体育消费者的类型和数量,从而促进体育消费和产业发展。其次,要发挥成都大运会的教育与文化交流功能,面向市内外、省内外、国内外青年大学生群体,策划设计组织创意活动,创新城市品牌和形象宣传的方式方法,充分体现成都特色,不仅吸引青年大学生来成都旅游打卡、参与各类活动,更吸引青年人才到成都安家落户、兴业发展。再次,要围绕球迷群体,支持鼓励球迷协会的活动,倾听球迷的需求,提升球迷的现场观赛体验,提升城市主队归属感。球迷是各类运动联赛的重要组成部分。目前在中超联赛中,成都蓉城的主场赛事上座率高,球迷热情度也高,这也会为其他运动项目联赛主场主队落户成都形成参照。

文体旅教深度融合发展,以目标客群为出发点,研发设计综合体验型服务产品。成都已经是热门旅游目的地,也形成了各种成熟的旅游产品。成都大运会的筹办过程中,又有许多新的内容和载体出现,如何打破文旅产品就是文旅产品、体育旅游就是体育+旅游、体教融合就是体育与教育的叠加这种初级融合?围绕目标客群来打造综合的服务产品,既需要创新思维,也需要打破本位思维,协同实施。比如,面对本地年轻人为主的家庭,面对外地来蓉旅游的散客,面对来成都观赛参赛的外地球迷或者运动参与者,研发推出不同的内容组合和套餐,形成文体旅教综合服务产品,“总有一款适合你”“总有一款有不同”,不仅有利于促进消费,提升消费者体验,更有利于创新成都旅游形象,打造成城市新品牌。

新建场馆群的赛后利用要注重定位明确,各美其美,探索老难题的新方案。运营前置是成都大运会新建场馆设施的一个重要策略,也是能够发挥成都大运会影响,更有效开展赛后利用的重要保障。但可以预见的是,成都大运会后,成都优质场馆场地设施的供给会大量增加,也不可避免会形成不同场馆间对优质内容的竞争,那么如何基于场馆的不同基础条件和优势,尽快形成场馆不同的功能定位和品牌定位,还是值得思考的。内容为王,运营为要,围绕定位,引入合适的内容,无论是体育比赛还是其他演出或者会展,精细运营,提升体验,形成正反馈,强化定位认知,希望成都可以为其他城市探索出破解场馆设施赛后利用难题的新方案。

融合成都各类资源禀赋,发力体育和体育产业中新兴赛道和领域。成都体育产业发展一直在快速增长中探索创新之道,结合成都已有的各类资源禀赋,而不仅仅局限于体育产业本身,视线扩展到各个行业和领域,融合协同各类资源,创新体育产业新模式。例如新兴运动装备研发和设计,以及新兴运动赛事及相关赛事IP方的引入,可以成为成都体育产业发展的特色和亮点。这里的新兴运动,既包括线下新兴的时尚潮流运动,也包括刚刚起步、方兴未艾的虚拟体育行业。成都在这些方面,既具有相关设计和研发的人才基础和技术基础,也具备非常活跃的用户群体,发挥这些优势,逐步形成市场主体和人才集聚、产品和活动丰富供给的发展态势。

成都大运会的筹办过程,是充满了挑战和变化的历程,成都这个有魅力的城市,不仅在筹办过程中全情投入、克服困难、积极应变,还将综合性大赛举办与城市高质量发展紧密融合在一起,不仅体现为筹办过程中相关硬件设施的完善,更体现在城市软实力方面的全面加强,使大赛举办与城市各个产业发展、各类业态融合、各种消费体验升级正向促进,也能促进成都城市形象和品牌的愈发鲜明和突出。

(作者系清华大学体育产业发展研究中心主任)

国讯

为全国统一大市场建设提供宏观数据支持,为建立健全以信用为基础的新型监管机制提供有力支撑——

中国企业信用指数首次发布

余颖

近日,在市场监管总局信用监管司的主持指导下,市场监管总局代码中心首次发布中国企业信用指数。2022年度中国企业信用指数显示为158.36点,同比上升24.14点。

据悉,中国企业信用指数以2014年为基期,基分为100点。2014年至2022年,中国企业信用指数由2014年的100点,波动上升到2022年的158.36点,创2014年以来的历史新高,其中2018年和2020年出现两次小幅下降,总体呈现“小幅波动、稳中上升”的发展态势。纵观中国企业信用指数9年来的发展趋势,虽然受到国内国际多重因素影响和冲击,中国企业信用水平小幅波动,但总体持续上升,市场秩序稳中向好,企业经营活力持续恢复,企业诚信守法经营意愿持续增强,营商环境持续优化。

编制中国企业信用指数,可以实现全国企业信用水平的数字化呈现,为全国统一大市场建设提供宏观数据支持;提供多维度的信用分析和管理方法,为优化营商环境提供重要抓手;从市场监管视角构建能够客观反映企业信用状况的工具,为建立健全以信用为基础的新型监管机制提供有力支撑。