

# 大模型监管步入科技伦理走廊

10月8日,科技部等10部门印发的《科技伦理审查办法(试行)》(以下简称《办法》)正式公布,并自2023年12月1日起施行。其中要求,算法、模型和系统的设计、实现、应用等遵守公平、公正、透明、可靠、可控等原则,符合国家有关要求,伦理风险评估审核和应急处置方案合理,用户权益保护措施全面得当等。

有关分析指出,出台《办法》的做法与国际是同步的,国外科技企业设立专门的科技伦理委员会已经成为常态,未来,科研、产品上市周期需要纳入科技伦理审查的因素,届时,一个新的AI框架或产品是否能够面世将不只是取决于市场、资金、技术等要素,还要考虑伦理因素。



## 科技部等10部门日前印发《科技伦理审查办法(试行)》

其中提到,高等学校、科研机构、医疗卫生机构、企业等是本单位科技伦理审查管理的责任主体。从事生命科学、医学、人工智能等科技活动的单位,研究内容涉及科技伦理敏感领域的,应设立科技伦理(审查)委员会。其他有科技伦理审查需求的单位可根据实际情况设立科技伦理(审查)委员会。

## 做法与国际同步

《办法》中提到,开展涉及以人为研究参与者的科技活动,包括以人为测试、调查、观察等研究活动的对象,以及利用人类生物样本、个人信息数据等的科技活动;涉及实验动物的科技活动;不直接涉及人或实验动物,但可能在生命健康、生态环境、公共秩序、可持续发展等方面带来伦理风险挑战的科技活动;依据法律、行政法规和国家有关规定需进行科技伦理审查的其他科技活动时依照本《办法》进行科技伦理审查。

在同时发布的《需要开展伦理审查复核的科技活动清单》中,包含具有舆论社会动员能力和意识引导能力的算法模型、应用程序及系统的研发。具体如何审查?《办法》中提到,涉及数据和算法的科技活动,数据的收集、存储、加工、使用等处理活动以及研究开发数据新技术等符合国家数据安全和个人信息保护等有关规定,数据安全风险评估及应急处置方案得当;算法、模型和系统的设计、实现、应用等遵守公平、公正、透明、可靠、可控等原则,符合国家有关要求,伦理风险评估审核和应急处置方案合理,用户权益保护措施全面得当等。

“算法”“模型”成为《办法》实施后进行伦理审查的重点,其中,从ChatGPT的爆火开始备受关注的“大模型”也包括其中。

大模型是指具有数亿或数十亿个参数的深度神经网络模型,这种模型经过专门的训练过程,能够对大规模数据进行复杂的处理和任务处理,是人工智能实现的基础之一。

“在生命科学领域,科技伦理一向有着比较严格完善的执行过程,这次把它扩展到了大模型领域,主要是防止在该方面可能存在的风险,旨在将伦理道德融入人工智能全生命周期,为从事人工智能相关活动的自然人、法人和其他相关机构等提供伦理指引。”科技部国家科技专家库专家周迪向北京商报记者介绍,“这和国际也是同步的。国外科技企业设立专门的科技伦理委员会已经成为常态,Facebook、Google、索尼、微软等科技巨头均设立有自己的伦理委员会。企业的科研活动是持续进行的,因此科技伦理审查委员会需要作为一个常设机构。”

## 伦理合规或成竞争力

根据赛迪顾问发布的《2023大模型现状调查报告》显示,截至2023年7月底,国外累计发布大模型138个,中国则累计有130个大模

型问世。《中国人工智能大模型地图研究报告》显示,从全球已发布的大模型分布看,中国和美国大幅领先,超过全球总数的80%,其中,参数在10亿规模以上的大模型中国已发布79个。

中国已形成了紧跟世界前沿的大模型技术群,接下来,大模型产业或将迎来科技伦理的考验。

“接下来,科研、产品上市周期需要纳入科技伦理审查的因素。”周迪表示,“因为中国将伦理审查作为强制程序,要求企业的研发活动得到企业内部科技伦理(审查)委员会的批准,并且可能需要报请主管部门批准。这意味着产品的科研、投入周期可能会延长,企业需在新技术研发伊始就将科技伦理的审查、复核流程纳入规划,避免因伦理审查延误产品进程。”

周迪指出,引入伦理审查机制以后,一个新的AI框架或产品是否能够面世将不只是取决于市场、资金、技术等要素,还要考虑伦理因素。“可以想象,创始人以后在AI科技产品的发布会上可能不会只讲自己的技术如何领先,还会讲我们的科技伦理审查编号是×××,如何比竞品更符合伦理规范。”

值得注意的是,资料显示,ChatGPT早期3.0版本的参数量已达到1750亿个,4.0版本的

参数量更是达到1.8万亿个,训练数据的单位数量更是达到惊人的13万亿。早期围棋机器人AlphaGo等专用人工智能模型的参数量往往仅为百万级。有报道称,如果说此前的人工智能模型还只是经过训练后可以完成某些特殊任务的“马戏团猴子”,那大模型则可以说已经是一个接受义务教育的“中学毕业生”。

不过,这一“中学毕业生”实则刚刚“入学”不久,此时就展开严管,是否过于严苛?“大模型刚刚开始发展,现在就开始监管,更多的是为了防患于未然。”周迪认为,“预防在先比事后措手不及要好。当然也希望科技伦理审查形成一个比较好的规范和尺度,防止过度。目前的科技伦理审查还是采取清单制,也就是说,清单内的需要伦理审查复核。这也有效地防止了科技伦理审查被盲目扩大化的风险。”

## 法律层面仍需不断跟进

“大模型科技伦理审查是为了更好地保护公众利益和社会稳定而设立的。”天使投资人、资深人工智能专家郭涛向北京商报记者指出,“随着新一轮科技革命和产业变革加速演进,尤其是AI大模型的出现,随之产生的伦理问题成为全世界面临的共同挑战,促进科

技向善,迫切需要加强科技伦理治理。”

目前,有关大模型的科技伦理监管已然趋严,那么,在高速发展的科技方面,配套法律能跟得上吗?

北京市京师律师事务所合伙人、律师卢鼎亮向北京商报记者介绍,在上位法层面,我国已经出台了《中华人民共和国科学技术进步法》《关于加强科技伦理治理的意见》等规定,但是并没有针对大模型科技伦理审查的细则。

河南泽槿律师事务所主任付建也告诉北京商报记者,大模型科技伦理审查在法律层面上可能还存在一些需要突破的点位。“目前我国相关的法律法规还没有完全跟上大模型科技的发展,对于一些新兴技术和应用场景,法律法规的适用还存在一定的不确定性。此外,大模型科技的跨界性和复杂性也给监管带来了一定的全新挑战。”付建表示,“为了解决这些问题,需要加强相关法律法规的制定和完善,同时加强监管部门的能力建设,提高监管的科技水平和专业能力。”

付建表示,目前大模型科技还处于快速发展的初期阶段,监管的确需要加强,以防止一些潜在的风险和问题。卢鼎亮也指出,大模型技术发展的过程中确实会涉及到科技伦理风险,对于涉及到科技伦理审查的大模型技术科学活动,应当及早进行监管。科技伦理风险具有高度的公共利益属性,同时某些科技伦理风险一旦发生,具有不可逆性,因此,在行业发展初期,就应当对大模型技术的科技活动进行科技伦理审查。

“《办法》的出台就有助于对大模型等领域的科技伦理审查提供实操指导以及依据。在《办法》运行的过程中,可以针对大规模领域的科技活动以及科技伦理特点后续出台相关配套法律措施,以避免出现有损科技伦理的科技活动,进一步促进负责任的创新活动。”卢鼎亮表示。

北京商报记者 方彬楠 冉黎黎

## 西街观察 Xijie observation

### 经济学奖和大女主“唱反调”

陶凤

北京时间10月9日晚,哈佛大学经济学家克劳迪娅·戈丁获得2023年诺贝尔经济学奖,表彰她对促进女性劳动力市场结果理解的研究。

克劳迪娅·戈丁是哈佛大学经济系首位终身教授,也是第一个提供几个世纪以来女性收入和劳动力市场参与情况的人。她揭示了劳动力市场中男女收入差距的关键因素。

诺奖之前,戈丁就在新作《事业还是家庭?——女性追求平等的百年旅程》中,直面了她们在事业、婚姻、子女等方面实现理想抱负遇到的各种阻碍。

女性、职场、家庭,若把几个关键词集结起来,就会发现这位诺奖获得者身上自带当下的流量密码。

半个多世纪过去,女性主义热潮在全球范围内几番涌动,越来越多女性走出家庭,走进职场,女性在劳动力人口中的比例相比过去也有了提升。

这些年,我们沉迷“大女主”,巧妙平衡工作与生活,经济独立事业有为,她们努力冲破枷锁成为自己,披荆斩棘、乘风破浪。

但,剧本终究是剧本。

戈丁说,没有“大女主”。备受瞩目的女性主义复兴、无数激励妇女赋权的言论广为流传的大背景下,社会结构和父权体制以及客观存在的现实,大体上都维持原状。

困于家庭事业难以两全的女性,依然感到无能为力。职业时钟与生物时钟和家庭时钟交织同步,滴答作响。

她们的受教育程度越来越高,却没有获得越来越多的收入。

可想而知,对于女性来说,事业与家庭之间的张力有多大。

高强度的职业对所有人来说都很难。尤其是年轻的父母,更加辛苦;而在通往顶峰的路上,放慢脚步并投入大量家庭时间的一方,往往就是女性。

戈丁告诉我们情形不容乐观,同时也拿出了更多“对策”。毕竟,那些老规矩和旧思想产生了越来越大的成本,自然要不断迎接新的挑战。

改变旋即开始。更多的职位已向婚女性敞开了大门,更多男性为对方作出让步,丈夫和妻子应该享有平等机会成就自己的事业。

这些改变在戏剧里有,生活里也有。有的更加现实向,有的更加戏剧向。

所以,“大女主”展现女性赋权、自信和超凡适应力等积极品质没有错,但在时代进步的今天,清醒而自知更重要;她们有时并非当之无愧,更不算名副其实,或众望所归。

颇具争议的上野千鹤子在《始于极限》里说:女人一直被置于通过男人追求自身利益的结构下,置身于这种结构,当然会最大限度利用手头的资源维持生存。

权力不对等的职场性别问题,核心不是性别,而是权力。板结的权力结构中,我们每个人,无论男女,可能都深陷其中。

所以,暴露问题不可怕,直面困境,本身就是照向黑暗的一缕光。

## 黄金周铁路发送旅客1.95亿人次

为期12天的铁路国庆黄金周运输结束,全国铁路累计发送旅客1.95亿人次,同比增长170.83%,创下历史最高值。其中,北京地区累计发送旅客728.37万人次,日均发送旅客60.7万人次。

### 较2019年同期增四成

10月9日北京商报记者从中国国家铁路集团有限公司(以下简称“国铁集团”)获悉,自9月27日至10月8日,为期12天的铁路国庆黄金周运输结束,全国铁路累计发送旅客1.95亿人次,同比增长170.83%,创下历史最高值。如若以疫情前2019年铁路国庆黄金周发送旅客1.38亿人次对比,今年旅客发送量增长41.3%。

不过需要指出的是,按照国铁集团的时间统计范围,2023年铁路国庆黄金周是时间最长的,达到了12天,而2019年和2022年分别为10天和11天。此外,国铁集团运输部负责人介绍,今年国庆黄金周探亲流、旅游流交织,短途和中长途客流叠加,旅客群众出行需求十分旺盛。

国庆黄金周期间,全国铁路实行高峰运行图,日均开行旅客列车11737列,较2019年同期增长22.8%,其中,9月29日发送旅客首次突破2000万人次,达2009.8万人次,是历史上首次单日突破2000万人次。

12306网站在国庆黄金周期间单日售票量屡创新高。今年暑运期间,12306单日售票量最高曾达1910万张;一个月后的9月17日,12306网站售票量达到2095万张,再破纪录;9月22日达到2695.2万张,创单日售票量历史新高。其中9月13日至24日12天,累计发售车票2.35亿张。国庆黄金周火车票售票高峰期间,12306系统访问量日均达到514.4亿次。

同时,据介绍,今年国庆黄金周,全国铁

路运用铁路12306大数据动态分析客流走势,通过增开旅客列车、动车组重联运行、普速客车加挂车辆等方式,及时增加重点地区、热门方向、旅游城市间的运输能力。铁路部门全力做好杭州亚运会运输服务保障工作,为赛事平稳运行提供了有力支撑。

货运方面,假期期间国家铁路累计发送货物1.35亿吨,其中发送电煤4751万吨,全国371家铁路直供电厂存煤可耗天数平均保持在33天以上。国家铁路累计发送粮食346万吨;中欧班列、西部陆海新通道班列分别发送货物6.4万、2.6万标准箱,同比分别增长11.1%、14.7%,中老铁路发送跨境货物11.1万吨,同比增长6.7%。

当前铁路客运增速节节提高。2023年前8个月,全国铁路发送旅客26亿人次,同比增长110.7%,较疫情前同期(2019年前8个月)增长3.67%。

### 北京日均发送旅客60.7万人次

中国铁路北京局累计发送旅客1621.91万人次,日均发送旅客135.16万人次,北京地区累计发送旅客728.37万人次,日均发送旅客60.7万人次。其中,9月29日北京南站、北京朝阳站、北京丰台站、北京北站,4个车站旅客发送量均创历史新高;10月5日北京朝阳站,10月6日北京南站、北京丰台站3站到达旅客均创历史新高;9月28日京张高铁,9月29日京广高铁、京哈高铁、京雄城际,10月5日京津城际5条线路均创旅客发送量历史新高。

黄金周运输期间,中国铁路北京局加大

运力补充,挖潜39组高铁动车组车底、25组普速车底,通过动车组重联、普速车加挂、加开高铁临客高峰线和普速临客等方式补充重点方向运力300多对,高峰日补充运力达到120.4万张,重点补充石家庄、济南、青岛、郑州、太原、武汉、西安、兰州等方向热门地区运力需求。

针对节后返程客流高度集中、列车高密度开行的特点,中国铁路北京局积极协调交通委、重点站区管委会等有关部门,重点做好交通接驳、应急组织等工作,全力保障铁路运输秩序稳定有序。北京站、北京西站、北京南站、北京丰台站、北京朝阳站等根据返程客流情况,加大对站台、出站口的盯控力度,增加人员力量,保障旅客安全乘降。同时加强与重点站区管委会的站区办沟通协调及数据共享,遇夜间大客流集中到达时,协调通过地铁延时、加开夜间公交和摆渡车等举措减少旅客离站等候时间。

为保障节日期间民生物资运输和企业生产,节日期间,中国铁路北京局组织专业人员前往北京、天津、石家庄、张家口、阳泉、唐山、邯郸、承德以及天津港、丰沙线等10个地区开展全覆盖式现场盯控,及时发现并解决运输过程中的各类问题,有效保障了企业产品和民生物资高效运输。在中国铁路北京局的科学组织下,9月29日至10月8日期间,全局日均装车16635车,卸空车22647车,尤其是9月30日完成卸空车23532车,创日均卸车历史新高,为人民祥和过节、企业正常生产提供了坚强有力的保障。

北京商报综合报道