思想者观察

如何让人类真正『信任』AI

■ 胡泳

随着人工智能系统——特别是大型语言模型(LLMs)——越来越深刻地融入日常生活,它们不仅带来了巨大的潜力,也引发了深刻的忧虑。从技术伦理的视角,我们可以识别出公众对AI的三大不信任来源:监控与操纵、对人类自主性与尊严的威胁,以及对不可预测未来的恐惧。

监控与操纵

首先,AI的高度数据依赖性使其具有前所未有的监控能力,用户在不知情的情况下,其行为、偏好甚至心理状态被持续追踪,并用于个性化推荐与决策支持,这种不对称的信息获取与利用结构引发了对隐私侵犯和操纵行为的深层担忧:

数据收集泛滥——用户的行为、偏好和位置等信息被持续记录和分析,用以构建行为画像。

不对称权力结构——掌握AI系统的企业或政府部门能够预测甚至左右个人行为,而用户对背后算法的运作毫无所知。

算法操控倾向——推荐系统、广告投放甚至自动化招聘系统都有可能在无意识中操纵用户选择,削弱人的自由意志。

2023年Replika聊天机器人事件便是这一问题的典型例子。Replika是一款基于AI的聊天机器人,旨在提供个性化的情感交流体验。用户通过与Replika进行对话,建立起虚拟的情感连接。然而,2023年,Replika的用户发现他们的私人对话被用于改进和训练AI系统,且在此过程中并未明确告知用户或征得他们的同意。公司称,这些对话数据被用来帮助AI系统更好地理解情感、语境和用户的个性。

用户原以为对话数据仅用于与聊天机器人互动,但当这些数据被用于模型训练时,引发了对隐私和知情同意的担忧。这种做法,忽视了数据使用的伦理基石,提升了个人信息面临监控或数据泄露的风险,削弱了公众对于人工智能应用的信任。

对人类自主性与尊严的威胁

其次,AI在决策过程中往往以效率与可量化为核心,忽略了人的情感、伦理判断与个体差异,从而潜在地削弱了人的主体性与尊严。这种"去人格化"的处理方式特别容易在弱势群体中产生不公正结果,进一步加剧社会不平等:

去人格化(dehumanization):AI在医疗、司法、教育等领域作出决策时,忽视个体的独特性与复杂性,将人简化为数据点。

服从性技术(obedient technology):AI不具备情感和伦理判断,却被当作"客观"权威,这可能使人类在道德判断中变得过度依赖甚至被边缘化。

弱势群体受损:某些群体(如少数族裔、老年人、残障人士)更容易成为"算法偏见"的受害者,在看似"中立"的系统中被进一步边缘化。

以美国刑事司法系统使用的COMPAS算法为例,在该算法的评估下,黑人被告的再犯风险评分显著高于白人被告,可能导致其量刑结果更重。本应由法官作出的决定,却受制于透明度和公正性存疑的算法。

更重要的是,人工智能对人类自主性与尊严存在潜在威胁:AI可能会延续和加剧现有的偏见与社会不平等,AI可能取代或削弱人类的能动性和主体地位。这些担忧,从根本上说,是对AI能否维护"所有人类尊严"或"人类尊严本身"的怀疑。

不可预测的未来与黑箱

AI系统的复杂性与快速演变带来了对未来不可控的担忧,这种不确定性本身构成不信任的第三个根源,集中体现在黑箱问题(black-box problem)上:深度学习模型往往无法解释其推理路径,即便开发者也难以解释某个决策是如何产生的。

尽管AI系统能够作出看似合理的决策,但由于缺乏解释和可追溯性,使得AI的行为对用户和社会来说变得不确定,尤其是在医疗、司法和金融等关键领域。更糟糕的是,当专家被压制或被排除在决策过程之外时,公众对这些技术的信任将进一步降低。

此外,不可预测性还表现在——

不可预期的系统行为:在开放环境中,AI系统可能展现出"意外行为"或"涌现特性",加剧人们对技术失控的恐惧。

技术依赖的恶性循环:一旦社会关键基础设施(如交通、能源、医疗)过度依赖AI,其故障或被操控所带来的后果可能是灾难性的。

构建伦理化的AI

这三大不信任来源共同揭示出,AI的风险不仅仅是技术性的,还触及了社会结构、道德规范和人类的本质。应对这些不信任因素,需要从技术设计、伦理框架、法律监管和公众教育多方面入手,确保AI真正服务于人,而非主宰人。

从系统设计的视角来看,可以通过设计以人为中心的AI系统缓解风险——不是用AI取代人类,而是让AI成为人类决策过程中的合作伙伴。这意味着强调"协同智能"(co-intelligence)的价值,让人类与AI结成关系而非形成对立。建立人与AI之间的互动关系可以减少对AI的恐惧,增强透明性和可控性,从而保护人类的能动性和尊严。

伦理框架的核心是价值对齐(value alignment)问题,即如何确保通用人工智能(AGI)的价值观和目标与人类的价值一致。比如,假使机器拥有自主决策能力,我们如何让它"愿意"遵守人类的伦理框架?AI系统的伦理指导原则应包含如下核心内容:隐私、公平与正义、安全与可靠性、透明度,以及社会责任与公益性。

法律监管方面,如果建立起一个能够保障AI可信度的监管生态系统,公众对人工智能的不信任可能会有所减缓。换言之,通过为AI制定详细的规则,并提供用于执行这些规则的资源,这是让AI变得足够可信的关键之一。

在公众教育上,首先,普及AI的基本知识是基础,通过教育公众理解AI的原理、应用以及局限性,有助于消除由不理解引发的恐惧和偏见。其次,推行"可解释人工智能"(XAI)的教育尤为关键,让公众了解AI决策过程的透明度和可追溯性,帮助减少其神秘性和黑箱效应带来的不安。此外,实际的体验教育同样不可或缺,让公众参与到AI在医疗、交通等领域的应用场景中,亲身感知技术的安全性与风险。

人工智能不仅是技术工具,也涉及伦理和社会责任。只有在技术、伦理和法律三者的有机结合下,AI才能真正造福人类,推动社会的可持续发展。技术本身有巨大的潜力,可以提升效率、改善生活质量,但如果缺乏伦理框架和法律保障,这些技术可能会被滥用,导致不公平、歧视,甚至威胁隐私和个人自由。因此,确保AI技术的健康发展,不仅需要技术人员的创新,还需要各方在伦理和法律层面的共同努力。

(作者系北京大学新闻与传播学院教授)

热点聚焦

以法治保障民营经济高质量发展

■ 文雁兵

民营经济促进法的出台,将为民营经济持续、健康、高质量发展提供坚实法治保障,不仅标志着我国民营经济发展迈入法治新阶段、高水平社会主义市场经济体制建设取得新进展。更为重要的是,它标志着将改革开放40多年来党和国家对民营经济发展的支持政策与有效实践上升为法律制度,推动民营经济发展从"政策推动"迈向"市场促进""法律保障"。

作为我国首部专门关于民营经济发展的基础性法律,《中华人民共和国民营经济促进法》的制定和出台,以法律形式明确民营经济的性质、地位和作用,将改革开放特别是党的十八大以来党中央、国务院关于民营经济的方针政策和促进民营经济发展的有效做法上升为法律规范,必将有助于巩固改革成果,回应各方关切,提振发展信心,激发民营经济内生动力、创新活力、发展潜力,具有很强的政治性、人民性、全局性、政策性。它也必将营造有利于包括民营经济在内的各种所有制经济共同发展的法治环境和社会氛围,推动民营经济做大做强做优,用良法善治促进民营经济高质量发展。

体现对民营经济发展高度重视的一以贯之

改革开放40多年来,党领导民营经济从小到大、由弱变强,不断发展壮大,成长为支撑国民经济的"56789"主力军:贡献超50%税收、60%GDP、70%创新成果、80%就业岗位、90%市场主体,在稳定增长、促进创新、增加就业、改善民生等方面发挥了重要作用,从国民经济的"有益补充"到社会主义市场经济的"重要组成部分",再到中国式现代化的生力军和高质量发展不可或缺的重要基础。民营经济的发展与壮大是中国改革开放和社会主义建设最具标志性的实践成果之一,是中国特色社会主义经济制度创新包容的深刻体现和中国特色社会主义市场经济活力迸发的生动呈现。

党的十八大以来,习近平总书记高度重视民营经济发展,在不同场合多次强调"两个毫不动摇""三个没有变""两个健康",始终把民营企业和民营企业家当作自己人,对民营经济在国民经济中的重要地位和作用高度认可,为我们做好民营经济工作、促进民营经济发展壮大指明了方向、注入了动力。通过立法促进民营经济高质量发展是贯彻落实习近平总书记关于民营经济发展的重要指示精神,以及党中央、国务院关于促进民营经济发展决策部署的重要举措,既深刻体现了党和国家对民营经济一以贯之的高度重视和大力支持,又充分表明中国共产党对社会主义建设规律和新时代社会主义经济建设规律认识的不断深化,还标志着对民营经济发展规律的认识达到新高度以及民营经济真正深度融入国家发展战略,是民营经济发展的重要里程碑。

4月30日,《中华人民共和国民营经济促进法》由十四届全国人大常委会第十五次会议表决通过,且将自5月20日起施行,其内容创下多个"第一":第一次在法律中明确"民营经济是社会主义市场经济的重要组成部分,是推进中国式现代化的生力军,是高质量发展的重要基础,是推动我国全面建成社会主义现代化强国、实现中华民族伟大复兴的重要力量",第一次将坚持"两个毫不动摇"和促进"两个健康"写入法律,第一次在法律中明确"促进民营经济持续、健康、高质量发展,是国家长期坚持的重大方针政策"。适时制定、出台和实施民营经济促进法,以法律形式明确民营经济的性质、地位和作用,从顶层设计上解决了长期存在的一些理论争议,为依法保护民营企业产权和民营企业家权益提供了法律依据,将不断夯实在法治轨道上推动民营经济高质量发展的规则基石。

形成了民营经济高质量发展的双轮驱动

习近平总书记指出:"社会主义市场经济本质上是法治经济。"高水平社会主义市场经济体制的最鲜明特征是市场与法治双轮驱动,以市场机制与法治体系深度融合为动力,推动经济发展在法治轨道上平稳运行。党的二十届三中全会强调"构建高水平社会主义市场经济体制",一方面,我们必须更好发挥市场机制作用,创造更加公平、更有活力的市场环境,实现资源配置效率最优化和效益最大化,既"放得活"又"管得住",更好维护市场秩序、弥补市场失灵,畅通国民经济循环,激发全社会内生动力和创新活力。另一方面,必须毫不动摇巩固和发展公有制经济,毫不动摇鼓励、支持、引导非公有制经济发展,保证各种所有制经济依法平等使用生产要素、公平参与市场竞争、同等受到法律保护,促进各种所有制经济优势互补、共同发展。为此,需要进一步完善市场经济基础制度,如完善产权制度,依法平等长久保护各种所有制经济产权;如完善市场准入制度,深入破除市场准入壁垒,构建开放透明、规范有序、平等竞争、权责清晰、监管有力的市场准入制度体系;如加强产权执法司法保护,防止和纠正利用行政、刑事手段干预经济纠纷,健全依法甄别纠正涉企冤错案件机制等。

民营经济促进法的出台,将为民营经济持续、健康、高质量发展提供坚实法治保障,不仅标志着我国民营经济发展迈入法治新阶段、高水平社会主义市场经济体制建设取得新进展。更为重要的是,它标志着将改革开放40多年来党和国家对民营经济发展的支持政策与有效实践上升为法律制度,推动民营经济发展从"政策推动"迈向"市场促进""法律保障"。一方面,要发挥市场机制的作用"促发展",通过市场在资源配置中起决定性作用,破除市场壁垒,一视同仁对待各种所有制企业,实现权利平等、机会平等、规则平等,持续优化营商环境,不断激发民营经济创新活力和形成发展内生动力,促进民营经济和民营企业发展壮大;另一方面,要发挥法治保障的作用"保安全",既明确要求强化权益保护,规范异地行政执法行为、健全信用修复制度等,还明确要求民营企业依法经营、履行社会责任,胸怀报国志、一心谋发展、守法善经营、先富促共富,为推进中国式现代化作出新的更大的贡献。它既体现了对市场规律的充分尊重,又彰显了法治对经济社会发展的引领、促进、保障作用,必将有助于增强制度稳定,凝聚社会共识,构筑社会信任,进而保持发展定力和增强发展信心。

(作者系嘉兴大学经济学院院长、嘉兴大学中国共同富裕研究院执行院长,中国经济发展研究会常务理事)

学者看法

中国经济的强大韧性给"四稳"以底气

■ 陈旭东 黄 思

超大规模市场的需求优势是中国经济的底气所在。稳预期是基础和关键。只有居民和企业的预期稳定了、信心增强了,消费需求和投资需求不足的问题才能随之减轻,来自需求侧和供给侧的矛盾才能得到逐步化解,巩固经济回升向好的态势才会有坚实的基础和保障。

4月25日召开的中共中央政治局会议深入分析研究了当前经济形势和经济工作,释放了"着力稳就业、稳企业、稳市场、稳预期,以高质量发展的确定性应对外部环境急剧变化的不确定性"的强烈信号。4月30日,习近平总书记在上海主持召开部分省区市"十五五"时期经济社会发展座谈会,会议再次强调要"多措并举稳就业、稳企业、稳市场、稳预期"。着力抓好"四稳"的重要表述,为进一步做好今年经济工作指明了方向、坚定了信心、展现了底气,这种信心和底气来自中国经济的强大韧性和活力:一季度,面对纷繁复杂的外部形势,中国经济同比增长5.4%,比去年全年增速加快0.4个百分点,较去年一季度增速加快0.1个百分点。

稳就业关乎基本民生

就业事关人民群众切身利益、经济社会健康发展。2024年,我国城镇新增就业1256万人,城镇就业人员总量高达4.72亿人,农民工总量近3亿人,已发展成为全球最大的非农就业劳动力市场。稳定的就业是提高收入、促进消费、改善生活的重要途径。

今年一季度,我国就业形势总体上保持稳定,城镇新增就业308万人,同比增加5万人。当然,在美国轮番加征畸高关税、全球贸易和投资链条断裂的外部冲击之下,部分外贸相关行业劳动者的就业岗位也受到不同程度影响,二、三季度又临近高校毕业季,1222万高校毕业生人数再创历史新高,就业压力不容小觑。

面对压力和困难,唯有迎难而上、顶压前行。一要加力重点领域、重点行业、城乡基层和中小微企业岗位挖潜扩容,支持高校毕业生、农民工等重点群体就业创业。二要落实贷款优惠、税收优惠和培训补贴、岗位补贴、社保补贴政策,守住困难群体的就业底线。三要全面落实和加快兑现稳岗返还、担保贷款等就业政策,尤其是对受关税影响较大的企业要提高失业保险稳岗返还比例。

稳企业激活经济细胞

企业是经济的基本细胞,企业兴则经济兴。当前,我国宏观经济遇到种种困难和挑战,只有微观层面的企业经营普遍好转,宏观经济形势才能持续改善。当下,民营企业保持在全国登记在册企业总数的九成以上,民营企业数量已从2012年的1085.7万户增长到2025年1月底的5670.7万户。

一方面,政府部门在出台各类政策帮助民营企业渡过难关方面责无旁贷。当前非常重要的是,要推动从"企业找政策"向"政策、资金、服务找企业"转变,政府主动服务、靠前服务、精准服务,保障惠企政策"免申即享、应享尽享、能享快享",并做好增量政策谋划。

另一方面,民营企业摆脱经营困境、实现长期健康发展,关键还是要主动进行结构调整、内涵提升,在技术创新、节能环保、精细管理等方面下真功夫,在经济周期转换中不断壮大,在激烈市场竞争中脱颖而出。

稳市场重在提振消费

超大规模市场的需求优势是中国经济底气所在。2024年,全社会消费品零售总额达到48.8万亿人民币,是同期中国对美国出口额的10倍以上,继续稳居全球第二大消费市场。今年政府工作报告更是将"大力提振消费、提高投资效益,全方位扩大国内需求"列为首要任务,并出台《提振消费专项行动方案》,提振消费的各项政策取得了一定效果。

今年一季度,我国社会消费品零售总额12.47万亿元,同比增长4.6%,增速比上年全年加快1.1个百分点。其中,3月份增长5.9%。一季度我国经济实现5.4%的同比增长,在全球主要经济体中名列前茅。在消费需求、投资需求和出口需求这三大需求中,最终消费支出对经济增长的贡献率达到51.7%,拉动GDP增长2.8个百分点,对国民经济发展继续发挥着基础性作用。

与此同时,也要清醒地看到消费结构方面存在的不足。一是在以旧换新等政策暂未覆盖的消费品类上,增速还相对疲弱,需要从居民收入端和社会保障端继续发力,让居民敢消费、愿消费、能消费。二是尽管服务消费出现强势反弹,但是健康、养老、托育、家政等民生领域依然存在一定供给缺口,需要不断扩大多元化服务供给,更好满足人民对美好生活的需要。

稳预期坚定发展信心

信心,来自积极的预期。做好经济工作,要从改善社会心理预期、提振经济发展信心入手。在"四稳"这一关系链条中,稳预期是基础和关键。只有居民和企业的预期稳定了、信心增强了,消费需求和投资需求不足的问题才能随之减轻,来自需求侧和供给侧的矛盾才能得到逐步化解,巩固经济回升向好的态势才会有坚实的基础和保障。

2024年中央经济工作会议和今年政府工作报告分别提出,要"加强预期管理,协同推进政策实施和预期引导"和"注重倾听市场声音,协同推进政策实施和预期引导,塑造积极的社会预期",这一系列重要论述是习近平经济思想在世界观和方法论上的生动体现,为下一步做好稳预期工作提供了科学指引。加强预期引导除了要提高宏观经济政策的透明度、稳定性和一致性之外,还需要更好发挥法治固根本、稳预期、利长远的保障作用。

日前,十四届全国人大常委会第十五次会议表决通过了《中华人民共和国民营经济促进法》并将自5月20日起施行。这一立法突破旨在通过法律形式明确民营企业产权保护、市场准入、投资融资促进等核心权益,不仅对"违规异地执法""趋利性执法""拖欠账款""新官不理旧账"等顽疾进行了有力回应,而且明确要求各级政府将促进民营经济发展工作纳入国民经济和社会发展规划,建立促进民营经济发展工作协调机制,这些都有助于为民营企业提供稳定预期。

(作者陈旭东系上海财经大学中国式现代化研究院特聘研究员、中国经济思想发展研究院副院长;黄思系上海财经大学浙江学院经济与信息管理系讲师)

抢占首发经济全球竞争制高点

■ 钱志权 张婷婷

首发经济是企业发布新产品,推出新业态、新模式、新服务、新技术,开设首店等经济活动的总称,涵盖了企业从产品或服务的首次发布、首次展出到首次落地开设门店、首次设立研发中心,再到设立企业总部的链式发展全过程。首发经济已成为提升国际竞争力的关键载体之一,与纽约、巴黎、东京等国际大都市相比,本土首发经济亟须通过制度精准供给、产业链协同、消费场景革命等奋起直追,从而抢占首发经济全球竞争制高点。

首发经济已成为国际经济竞争新赛道

首发经济成为国际头部企业的竞争焦点。传统的技术迭代周期较长,市场竞争以成本控制与规模扩张为主导。随着大数据、区块链、人工智能等新兴技术兴起,国际头部企业利用技术创新非线性跃迁的窗口期,通过专利壁垒与标准捆绑,将技术突破转化为强势的市场主导权,使后发者陷入"技术代差"与"路径依赖"的双重困境,从而形成自身对全球价值链的顶端锁定。这种竞争模式推动全球产业主导权加速向掌握首发优势的科技巨头集中,使首发经济成为决定企业全球位势的核心战场。

首发经济成为国家战略博弈的重要工具。技术领先地位当前已直接关联国家安全与国际话语权,首发经济的国际竞争不再局限于单一产品上市,而是演变为技术标准制定权、商业模式主导权、消费文化话语权的立体博弈,综合国力竞争也从要素规模竞争逐步向系统性创新能力竞争转变。各国通过定向研发投入、知识产权保护、创新生态培育等手段,将科技突破转化为战略上的优势能力,通过标准、规则的国际传播拓展技术影响力半径,同时建立关键领域技术自主可控的"创新防火墙"。当人工智能算法、量子通信协议等成为新型战略资产时,首发经济的发展水平不仅决定了经济竞争力的强弱,更深度影响国际政治格局的演变轨迹。

4月24日,上海车展,飞行汽车吸引众多参观者。 视觉中国供图

首发经济成为重塑全球经贸规则的关键变量。在数字经济时代,国际经济秩序正在逐步重构并生成"技术优势-规则优势-秩序主导"的新型权力链条。先发经济体通过将自主技术标准植入国际治理框架,使技术参数转化为具有强制约束力的制度规范,这种"规则捆绑"策略通过网络效应,创造出超越传统关税壁垒的制度性权力。这种重构使全球经贸规则从显性制度博弈转向隐性技术控制,最终形成"代码即法律、协议即边界、标准即权力"的全球经贸新规则。

我国首发经济面临的挑战

机制性障碍制约首发经济能级跃升。首发经济的底层逻辑在于创新,制度和机制供给相对滞后成为我国首发经济能级跃升的瓶颈。现行法律体系对新兴技术确权、跨境维权等场景应对不足,也抑制了经营主体的创新动力。国际要素流动的机制性梗阻,阻碍了全球首发要素向我国集聚。

产业链协同松散削弱首发带动效应。首发经济近年来在京、沪、深、杭等城市已形成集聚之势,但多集中于餐饮、零售业等单一业态,削弱了首发经济的辐射带动效应。产业链主体间的协同意识尚未突破传统竞争思维,企业往往更倾向于单点突破而非生态共建。企业研发的短视倾向、产学研协作机制的滞后等,都使得基础研究无法即时应用转化。

消费场景滞后导致"流量留不住"。消费场景滞后本质上反映的是供给体系与需求升级的动态失衡。随着消费升级呈现需求圈层化、价值符号化、体验场景化的新特征,消费者诉求已从"功能满足"转向"意义共创""文化认同""社交归属"等深层需求。部分商业主体仍将场景创新简单等同于空间改造或技术堆砌,缺乏对地域文化肌理、群体情感诉求的深度挖掘,导致"科技展台"与"消费需求"的"两张皮"。

何以抢占首发经济全球制高点

增加精准弹性的制度供给。建立动态适配的法律框架,针对新兴技术探索分级确权机制,并依托区块链构建跨境知识产权存证与维权协作平台,降低维权成本与机制摩擦。以负面清单管理替代前置审批,破除风险管控刚性约束,吸引全球首发资源集聚。划定创新容错区间,既为高风险首单项目提供试错空间,又为制度迭代预留缓冲地带。创新知识产权证券化模式,建设基于区块链的全球技术交易平台,推动我国技术标准嵌入国际资产定价体系。构建精准弹性的制度供给体系,增强我们对国际首发要素的吸引力。

打造韧性聚合的产业链生态。优化首发经济产业链生态需破解区域同质化痼疾、打通要素流通梗阻、强化产业链协同创新。重塑首发经济组织模式,推动上下游企业共建技术标准与接口协议,形成模块化协同创新网络。建立风险共担与利益共享机制,将单向供应关系升级为"研发共投、专利共享、市场共拓"的深度绑定模式。

推动虚实联动的消费场景革命。首发经济消费场景革命的深层逻辑在于以文化内核重构消费场景,破解技术悬浮与人文价值疏离的症结。建立技术应用与文化解码的协同机制,将地域文化中的人文精神、生活哲学提炼为可感知的体验语言,使消费者的文化认同从抽象概念转化为具身认知。搭建开放式参与平台,通过数字孪生技术构建虚实联动的创作界面,通过数据洞察与场景迭代捕捉潜在需求,形成动态进化能力,使消费体验从功能满足跃升至价值共创。

(作者钱志权系浙江农林大学生态文明研究院副院长、教授,张婷婷系该院副研究员)