

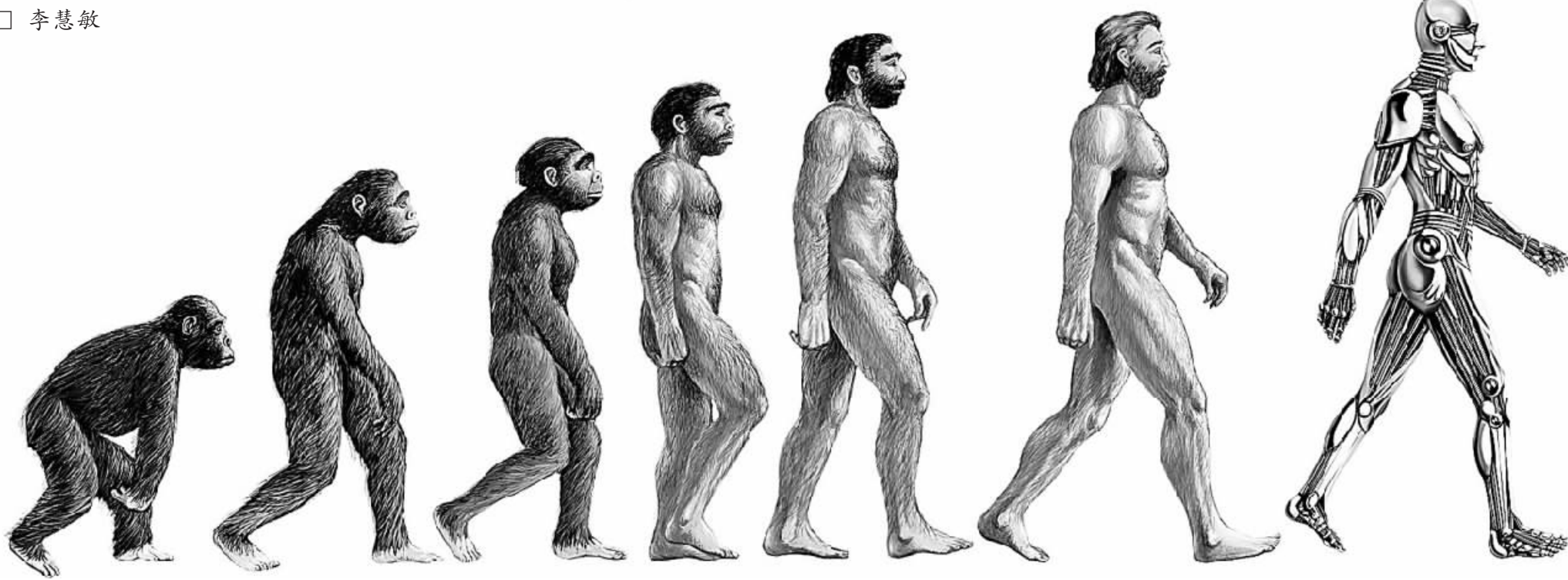
人工智能(AI)已经成为新一轮科技革命和产业变革的核心驱动力,正在对人类生产、生活、组织方式及经济社会发展的广泛领域产生极其深刻的影响。但是,怎样才能避免新兴技术突破人类的道德禁区而滑向未知,特别是在人工智能研究和应用中应遵循怎样的伦理和原则,成为各国探讨的热点。目前,尚未有定论。

日本自2016年在《第五期科学技术基本计划》中提出“Society 5.0”发展目标以来,各部门、学会、社会团体围绕人工智能社会展开积极探讨,发布了系列研究报告、执法指南等。2018年12月27日,日本内阁府发布《以人类为中心的人工智能社会原则》(以下简称《原则》),是迄今为止日本为推进人工智能发展发布的最高级别的政策文件,从宏观和伦理角度表明了日本政府的态度。在公开征求意见后,计划于2019年3月正式发布。

《原则》肯定了人工智能的重要作用,同时强调重视其负面影响,如社会不平等、等级差距扩大、社会排斥等问题。主张在推进人工智能技术研发时,综合考虑其对人类、社会系统、产业构造、创新系统、政府等带来的影响,构建能够使人工智能有效且安全应用的“AI-Ready社会”,对我们正确看待“人机关系”,更好引领和适应人工智能发展具有借鉴意义。

解读日本《以人类为中心的人工智能社会原则》 人工智能社会需要怎样的“紧箍咒”

□ 李慧敏



光明图片/视觉中国

1 怎样理解人工智能社会的本质

对人工智能的利用不能仅仅停留在实现效率、便利等现实利益层面,人工智能作为人类的公共财富,要通过推动社会方式的根本变革和真正创新,来实现全球可持续发展目标。人工智能社会需要体现以下三个基本理念:

(1) 尊严:人工智能社会是尊重人类尊严的社会。应将人工智能作为人类的工具之一,使人类能进一步发挥自身能力和创造性,通过从事具有意义的工作得到物质上、精神上的富足。

(2) 多元包容:使用人工智能技术,不只为实现提升便利、解放劳动等传统效率目标,要能够使具有多元背景和价值观、想法的每个人都能获得幸福,对多元化采取灵活、包容的姿态,推动新价值创造。

(3) 可持续:通过灵活使用人工智能,强化科学技术储备,催生新的商业模式和解决方案,实现科技创新立新,减少社会差距,缓解全球环境问题等,推动可持续社会的构筑。

2 接纳人工智能,你准备好了吗

人工智能、物联网、机器人、超高速宽带通信等都是实现“Society 5.0”的重要技术。人工智能友好型的基础设施是实现超级智能化社会所必需的一种变革。“设定怎样的目的去使用人工智能”是人类必须考虑的问题。为此,《原则》分别从人类、社会系统、产业构造、创新系统、政府监管五个维度勾勒了“AI-Ready社会”的愿景。

第一,人类。AI-Ready社会对人类能力的要求体现在三个方面:(1)有能力充分理解人工智能的优点和缺点,作为人工智能信息源的数据、算法存在“偏差”人可能性,也可能被用于不合理目的。其中,数据的偏差主要有统计层面的偏差、社会形态产生的偏差以及人工智能使用者产生的偏差;(2)能够使人们从事具有较高创造性、生产性的劳动。通过人工智能的帮助能够使人们基于出身、文化、兴趣等,获得实现梦想和发挥创意的能力。为此需要推动教育体系及其相关社会制度的构筑;(3)具有数据、人工智能基础和实施、设计等应用能力,跨领域、综合性、融合性的大量人才。这类人才是推动AI-Ready社会所有活动的原动力,能够为构建灵活使用人工智能的生活环境贡献力量,构建人们得以更加富足、充实生活的社会制度。

第二,社会系统。人工智能的使用能够推动个性化服务、解决方案的升级和效率化。为了使这些优势得以在社会层面广泛应用,需要医疗、金融、保险、交通、能源等在内的社会系统整体的灵活调整,以应对人工智能带来的变革。从社会系统角度,需要强化“软件”和“硬件”

两个方面,形成有助于实现具有扩充性、互联性、发展性秩序的柔性架构设计。此外,为了保障社会系统整体的互联、协作,尤其需要建立各类社会系统间互联互通的数据应用基础支撑平台。

第三,产业构造。为了使多元化的创意和理想得以实现,需要灵活地构筑面向国际的开放的劳动、雇佣和创业环境。需要企业实施公平竞争,采取灵活的工作方式,使人类的创造力通过产业能够得以持续发挥,从而拉动创新创业投资。

第四,创新系统。大学、科研机构、企业以及民众都可以跨越领域和立场,参与到人工智能研发、利用及评估中。构建使所有数据能及时、安全达到被人智能解析程度、任何人均可安心提供数据,并能够从中获得应有收益的环境。研发人员和用户均可安心从事人工智能研发及应用,通过研发与应用快速循环运转,加速人工智能发展。通过人工智能应用进一步产生新的创意,催生更多的可能性,进一步扩大创新的范围。

第五,政府监管。随着社会形势变化和技术发展,人类、社会系统、产业构造和创新系统的相关内容和目标也会随之更新。需要产学研官民等利益相关方相互协作,共同提出规则、制度、标准化、行为规范等方面的监管课题,评估影响,构建能够表达利益诉求并付诸实施的法律法规及监管体制,兼顾较难表达呼声的社会群体的意见。采取包括技术手段在内的企业层面的自主性规划,构建较为灵活且具实效性的监管方式。为了应对跨境问题,还需要建立国际合作协调机制。

3 为什么要给人工智能戴上紧箍咒

科技的发展越来越迅速,而普通人对于复杂的黑科技也越来越难以理解。人类社会根据适用对象的不同,需要对人工智能的发展方向加以疏导和限制,避免可能出现的失控局面。这就需要一些相对具体的指导原则进行引导。

第一,人类中心原则。人工智能的使用不得侵害宪法和国际规范保障的基本人权。人工智能的研发、推广、运用应该用于提升拓展人的能力,使人类获得多元幸福。为了防止过度依赖人工智能、恶意使用人工智能操纵人类意识决定的行为,需要把握以下原则:(1)人工智能不仅仅用来代替部分人劳动,还可以作为高级工具辅助人类,从而拓展人类能力和创造性。(2)在人工智能利用中,人类自己可以判断和决定如何使用人工智能。对使用人工智能带来的后果和责任,根据问题性质在研发、提供、运用等利益相关者之间适当分配。

第二,教育应用原则。平等地向所有人提供教育、培育应用的教育环境:(1)在幼儿教育、初等中等教育中提供广泛的教育机会,为社会人和老年人提供重新学习的机会。(2)建立人人都能掌握人工智能、数理、数据科学等相关素养的教育系统,超越文理学科界限强化学习。在教育中,设置有助于认识到数据偏差及使用方法的偏差等人工智能、数据相关特性的内容,以及人工智能公平性、公正性、隐私保护等相关内容。(3)从单方向、均一化教育方式转变为激发人类兴趣和力量。始终保持灵活转变的

意识,灵活运用人工智能,构建防止弱势群体出现的交互式教育环境。(4)完善教育环境并不是给行政部门和学校(教师)增加负担,民间企业和普通市民要发挥自主性。

第三,保护隐私原则。人工智能社会中,依据个人行动等数据可以高精度推断其政治立场、经济状况、兴趣爱好等。这意味着需要谨慎对待个人数据:(1)对个人数据的使用,包括政府使用行为在内,不得侵犯个人自由、尊严和平等;(2)对个人数据中涉及隐私的部分,确保使用行为的正确性、正当性,以及本人能够实质性地参与,并从中得到适当益处;(3)个人数据必须根据重要性和需求性得到妥善保护,基于文化背景和社会共同理念,把握使用和保护之间的平衡。

第四,保障安全原则。人工智能在实现社会系统自动化、提升安全性的同时,在目前能够设想的技术范围内,对于稀有现象和故意攻击行为人工智能无法做到总是能够正确识别,在安全方面存在新的风险。因此,需要把握利益与风险之间的平衡,从整体上提高社会的安全性和可持续性:(1)推进正确评估人工智能风险及降低风险的研究,人工智能深度研发以及包含安全保障在内的风险管理;(2)重视人工智能使用的可持续性,不一味依赖特定的、单一的或者少数几个特定的人工智能。

第五,公平竞争原则。维护能够推动商业模式和服务创新,维持经济持续增长和解决社会课题的公平的竞争环境:(1)不应形成特定国家基于对人工智能相关资源的集中,利用其支配地位实施不当的数据收集和侵害主权行为的社

会;(2)不应形成特定企业基于对人工智能相关资源的集中,利用其支配地位实施不当的数据收集和侵害主权行为的社会;(3)不应形成由于人工智能的使用,导致财富和社会影响力不当地集中于部分利益相关者的社会。

第六,公平性、说明责任及透明性原则。为了避免人工智能的使用导致个人因背景不同而遭受不当的差别待遇,或者使人在尊严层面遭受不正当待遇,需要确保公平、透明的意思决定,以及对意思决定结果的说明义务,确保技术的可信性:(1)在人工智能设计理念下,人不应由于人种、性别、国籍、年龄、政治理念、宗教等背景被不当地差别对待;(2)对使用人工智能事实、人工智能数据的取得方法及使用方法、确保人工智能运行结构正确性的相关措施等,都需要根据情况予以说明;(3)为使人们能够理解并判断人工智能提案,对于人工智能的使用、采纳、运用都应该根据需要设置沟通对话场所;(4)使社会能够安心灵活使用人工智能,需要保障人工智能以及数据的可信性。

第七,创新原则。(1)为实现持续创新,需要跨越国境、产学研官民、人种、性别、国籍、年龄、政治理念、宗教等界限,从人才、研究两个方面,大力推进国际化、多样化以及产学研官民合作;(2)推动大学、研究机构、企业间的平等合作以及人才流动;(3)推进人工智能品质及可信性的确认方法、效率性数据收集及整理方法、人工智能开发、测试、运用方法等人工智能工程的明确;(4)在确保个人隐私和安全的前提下,打破数据垄断,推动数据跨境流动使用,完善国际合作,构筑人工智能计算机资源、高速网络共用的研发环境;(5)推进人工智能技术的社会应用,政府应成为阻碍因素的规制内容实施改革。

4 我们对人工智能型社会有怎样的期待

对于人工智能开发者和经营者来说,需要遵守基本理念和上述人工智能社会原则的基础上,制定并遵守人工智能开发利用原则。目前已经有不少国家、团体、企业都已经对人工智能开发利用原则开展探讨,但是,目前尚未达成国际共识。因此,本《原则》强调日本需要积极制定本国的开发利用原则,通过开放式探讨,积极推动国际非排他性、非约束性原则框架的达成。2016年,日本政府在第77信息通信部长会议上提出人工智能研发8项基本原则,并提议创设讨论国际统一标准的国际会议。

当前,人工智能不仅与实体经济深度融合,对于社会发展、贫富差距、政府治

理、资源枯竭、全球环境问题和科学研究等关乎人类可持续发展的重大课题都有重要影响。加上日本自身面临少子老龄化、劳动力不足、地方稀疏化、财政支出负担等社会经济问题,因此,日本政府对于人工智能的发展和广泛应用寄予厚望。人工智能的内涵并不局限于技术层面的人工智能,而是泛指“高度复杂的智能社会体系”。

值得一提的是,《原则》在尊重人类尊严、实现人工智能可信性等与欧洲委员会2018年12月18日发布的《关于可信人工智能的伦理准则》具有一致之处。《原则》崇尚人人平等、不唯效率至上的人本发展观,符合对“Society 5.0”的概

念界定和打造超级智能社会的发展目标。

中国作为2016年二十国集团(G20)主席国,首次将“数字经济”列为G20创新增长蓝图中的项目重要议题。2018年G20峰会,我国建议二十国集团将“新技术应用及其影响”作为一项重点工作深入研究。我国作为人工智能发展大国,向全球展现了人工智能时代的中国智慧。我们不仅要抓住人工智能发展机遇,还要积极应对人口老龄化、劳动力人口下降、技能型人才短缺等挑战,多维度对人工智能的开发应用及影响进行系统深入、客观冷静的应用研讨,在基础研究、技术开发应用、资源开放共享、人才教育及法律政策上进行有序引导,为人工智能持续健康发展保驾护航。

(作者系中科院科技战略咨询研究院助理研究员)

作为最原始的人类生产活动——农业与自然环境的关系为最密切,基本上每个国家的大多数用地都是用于发展农业生产,因此,农业生产对自然保护提出了很大挑战。从农耕社会到工业化社会,人类在历史发展的长河中不停地反思,如何实现人与自然的和谐共处?如何让地球成为所有物种赖以生存的家园?荷兰国土面积不大,可用耕地十分有限,但农业却高度发达,比重占到国民收入和就业的近10%。与此同时,荷兰的自然环境也很好,素有“欧洲花园”的美称。针对荷兰在环境保护方面以及生态环境领域取得的成果,本报记者日前专访了荷兰莱顿大学执行院长、科学委员会主席雪勇(Geert de Snoo),他强调,人与自然和谐共生才是真正的现代化。

成立于1575年的莱顿大学是欧洲最著名的大学之一,也是荷兰历史最悠久的高等学府,在基础科学和生物科学研究方面有很大成就。雪勇教授在生态环境保护、生物多样性及可持续发展方面颇有建树。

荷兰是世界上最大的奶酪生产国和最主要的农产品出口国之一——世界上有一半以上种薯贸易从这里输出,荷兰鲜花也长期占据着国际花卉市场出口的半壁江山。一个小国却有如此高的农牧业生产能力,这绝非偶然。雪勇表示,考虑到地理环境的特殊性和农业生产生活的需要,荷兰对农业生产和自然保护区进行集约管理和精准监控,用好每一寸土地才能发展农业,但是先用地而不保护是不通的。

从第一次工业革命开始,随着工业化对自然环境的破坏步步加深,各国开始尝试设立自然保护区。简单来说,就是把地“圈起来”。雪勇介绍,20世纪开始,荷兰也设立了一批自然保护区,包括对一些湖泊和沼泽的保护。为了防止农业生产带来的破坏,防止湿地的干涸和退化,政府对保护区内的湿地进行了保护和管理,努力“使它们保持原样”。但随着环保观念的逐渐发展变化,现在,大家开始认为应减少人为的干预,注重发挥自然本身的自我修复功能并保护生态系统的整体性。“我们在保护区外的生产活动与保护区内一切都息息相关,就像我们不能简单把北极‘圈起来’来保护北极熊一样。”

不同的国家,气候、土壤、水条件都不尽相同,但人们对良好生态环境的要求却大同小异。雪勇表示,我们享受的食物、空气等等一切东西都来自自然,在不同的生活空间,城市也好,乡村也好,人们都希望有花草、有大树、有小树,这一点在不同国家地区的人们之间没什么区别。因此,在生态环境保护方面各国之间可以相互学习借鉴、共同合作。欧洲国家有很多相对开放的自然保护区,普通民众可以和保护区内的动植物密切互动。这种“沉浸式”环保行为主要是从非洲学来的,因为大家意识到每个地方的自然环境保护都需要当地所有人的共同参与,只有这样才能让大家真正重视起来。环保不仅仅是政府和各类社会团体的责任,全民参与其中才是生态环境长期、稳定、可持续发展的关键。

此外,在发展集约化农业的同时,做好农业区的生态环境保护,让人与自然和谐相处也非常重要。而且,眼光不要只盯着物种,而要有大局思维,着眼于整个自然生态的发展。在不知不觉中,大自然可以给我们创造很多惊喜。

雪勇介绍说,近五个世纪以来,莱顿大学在生命科学领域所取得的成就就是各个领域无数专家一点点努力的结果。在生命科学领域,除了专注自然之外,莱顿大学还致力于研制和更新药物,帮助人类更加有效地对抗疾病。从药物分子到成品药物,再到进入人体,进而惠及全体国民,这是一个复杂和漫长的过程。想要尽可能让所有人从中受益,就需要全社会的参与。科研机构、医院、患者、志愿者都要密切合作。“制药”不仅仅是药物研发那么简单,还包括对人体健康变化趋势的研究。因此,在应对全球气候变化、生态环境保护等重大课题面前,各国人民加强合作、共享研究成果很有必要。

雪勇还表示,随着以大数据、人工智能为特点的新一轮科技革命兴起,信息科技在生物生态研究中也越来越重要。相信在未来的几十年里,人类可以在促进环境保护、维护生物多样性方面取得更多实质性突破,更好实现人与自然的和谐共生。

本报记者

杨艺明

李盛明

人与

访荷兰莱