

我们该如何避免“弗兰肯斯坦回形针”隐忧

苏令银

面对技术创新过程中可能带来的风险和挑战,欧美学界首先提出了“负责任创新”问题。“负责任创新”,强调将更多要素纳入责任系统之中,更多考虑人的权利,追求创新成果的绿色性、人文性与普惠性,通过对(科技)创新的管理,力求使创新在体现技术效用与经济效益的同时,为应对社会挑战提供“智慧性、可持续性、包容性”的解决方案。

近日,一段百度董事长李彦宏乘坐无人驾驶汽车上北京五环的视频引发社会各界的广泛关注和热议。我认为,从更深层次意义上看,这里涉及人们对科技进步的风险担忧和科技创新中的责任问题。



“我所担心的并不是人工智能能够像人一样思考,我更担心的是人们像机器一样思考,没有价值观,没有同情心,没有对结果的敬畏之心。”

“机器问题”卷土重来

科技是一把双刃剑,科技创新在给人类社会的发展和进步带来福祉的同时,也给人类经济社会的发展带来了严峻的挑战和不确定性。当前人们对人工智能发展最大的担忧就是,认为人工智能短期会带来大面积失业,而长期可能会危及人类生存,甚至导致人类灭绝。这就是所谓的“机器问题”,也称“机器换人问题”。

“机器问题”的产生是伴随着工业革命和自动化的发展而兴起的。早在两个世纪以前工业革命时代的英国,“机器问题”就曾引发了人们的担忧和激烈的争论。有些人担忧机器只是有选择地让少数人受益,它会抢走所有人的工作并最终彻底颠覆社会。比如,当时的经济学家大卫·李嘉图(David Ricardo)就十分关注机器对于不同社会阶层利益的影响,特别是劳动阶级认为使用机器通常不利于他们自己的利益;托马斯·卡莱尔(Thomas Carlyle)则认为“机械恶魔”的破坏性将会扰乱整个工人团体。

如今,“机器问题”又卷土重来,只不过它伪装成了另一副面孔。科学家、经济学家和哲学家都在热议人工智能技术带来的深刻的潜在影响。人工智能技术突飞猛进,机器人能够执行曾经只有人类才能胜任的各种任务。很多人担心机器会让越来越多的人下岗,引发不平等问题和社会动荡。去年6月出版的《经济学人》杂志在其专题报告《人工智能》(Special report: Artificial intelligence)中重点关注了“机器问题”。牛津大学2013年的一项研究表明,

在未来20年里,美国47%的工作都有转向自动化并被机器所取代的风险,这令现在的许多工人感到更加担忧。2014年,皮尤研究中心发布报告《人工智能,机器人学和工作未来》,称52%的参与专家认为,到2015年由人类完成的许多工作将会被机器所取代。2015年下半年,英国央行首席经济学家安迪·霍尔丹(Andy Haldane)估计,英国大约有1500万个工作岗位(大约为总量的一半)可能受到机器的威胁,而欧洲主要国家多达50%的工作岗位估计面临被机器替换的风险。《未来简史》和《人类简史》的作者、以色列著名未来学家尤瓦尔·赫拉利(Yuval Noah Harari)则预言,随着人工智能技术的发展,必将造就更多的“无用阶级”。2016年10月,AI NOW发布的专题报告《The AI Now Report》指出,人工智能为经济价值的生产带来了新的方式,也对价值分配产生了新的影响。在某种程度上,人工智能的价值分配会使一部分群体受益,从而延续甚至加剧现有的财富分配差距。

更有些人担忧人会像机器人一样思考,人工智能会威胁人类的生存。未来超级人工智能可能会不认同人类的目标,转而去摧毁创造了它们的人类。著名科学家霍金在接受英国广播公司采访时表示:“人工智能技术的研发将敲响人类灭绝的警钟。这项技术能够按照自己的意愿行事并且以越来越快的速度自行进行重新设计。受限于缓慢的演化速度,人类无法

与之竞争和对抗,最终将被人工智能取代。当人工智能发展成熟时人类的末日就将来临。”令人惊讶的是,一直致力于无人驾驶领域的特斯拉公司掌门人埃隆·马斯克(Elon Musk)也警告人类:“我们正在用人工智能召唤恶魔。”苹果公司的CEO库克在今年MIT的毕业典礼上的一句话更是发人深省:“我所担心的并不是人工智能能够像人一样思考,我更担心的是人们像机器一样思考,没有价值观,没有同情心,没有对结果的敬畏之心。”

“负责任创新”应运而生

面对技术创新过程中可能带来的风险和挑战,欧美学界首先提出了“负责任创新”问题。“负责任创新”(responsible innovation),强调将更多要素纳入责任系统之中,更多考虑人的权利,追求创新成果的绿色性、人文性与普惠性,通过对(科技)创新的管理,力求使创新在体现技术效用与经济效益的同时,为应对社会挑战提供“智慧性、可持续性、包容性”的解决方案。

首先,负责任创新是理论与实践双重逻辑必然性展开的结果。从理论维度来看,一方面,为了缓解科技与自然、科技与伦理之间的内在张力,人们提出了可持续发展的理念以及企业社会责任、职业伦理、风险

管理、技术评估等理论,虽起到了一定的作用,但效果却不尽如人意,理论发展面临困境。1979年,汉斯·约纳斯(Hans Jonas)提出以“责任伦理”应对技术带来的巨大威胁,呼吁各类群体参与到科技创新活动中,以防止其造成严重后果。另一方面,技术哲学的经验转向忽视了社会与环境的影响,而伦理转向则面临伦理外在化问题。因此,技术哲学的荷兰学派倡导一种“内在主义”和“价值转向”进路,即技术与伦理并行研究,强调直接介入技术实践以应对社会问题从而增强技术创新的建设性,而负责任创新就是这一进路的主要代表。从实践维度来看,一方面,过往的技术滥用带来的恶果和灾难迫使人们对科技创新加以规制。随着公众的觉醒,多元利益相关者的协同成为创新的必然,科技创新必须充分考虑其利益诉求。另一方面,科技创新是人类社会进步的源泉,但其作为“创造性破坏”存在不可避免的负效应,因此,应当对科技创新的社会、伦理维度进行更多考量。当前世界各国都在大力推进科技创新,比如我国的创新驱动战略、美国的“能源新政”、欧盟的“地平线2020计划”等,这更要关注科技创新的环境、伦理、社会风险,确保科技创新无害化、人性化,实现科技创新成果同国家需要、人民要求、市场需求相结合。

其次,负责任创新的提出,也是转变科学家的传统观念之需。在当前科技创新过程中,某些自然科学家仍然主张科学技

术“价值中立论”,有的公然宣称科学家必须“为了科技上1%的进步而不断努力,甚至可以不择手段”,也有的顽固坚持“科技创新仅仅是为了满足科学家的好奇心”的观念。实际上,高技术必然意味着强责任,相比于一般社会公众,由于科学家掌握着科技创新发展的主导性和权威性,从而他们对科技创新可能产生的风险更有发言权,也必将为其风险性承担更大的社会责任。

正是在此背景下,大批关注科技创新的风险以及公众参与的项目开始实施,相关研究也出现了爆炸式增长,旨在通过对科技创新过程中多元价值的关照,探求一种具有可持续性的经济社会发展进路。

这里需要强调的是,负责任创新并非消极压制科技创新,而是时刻警醒我们如何在问题丛生的世界中“合宜地”发挥科技创新的价值。人类的生存本质上是一种技术化生存,科技对人类的生活方式与生存境遇具有重要的型塑作用,在当今科技高度发达的时代尤其如此。而人的目的对规制技术至关重要,因此人类需要通过对目的的规约来融洽人机关系。负责任创新的价值在于把伦理、社会等多元因素纳入科技创新中,力求破解科技造成的负面效应,同时降低高技术可能带来的风险,实现人类对科技及其影响的规制,增强对科技创新风险的防范意识,从而实现对人类生存境遇的最优化。这正是负责任创新所追求的目的,也是其终极意义所在。

如何避免“弗兰肯斯坦回形针”隐忧?

机器能够脱离人类控制而自己独立运行并不是一个新鲜想法,英国作家玛丽·雪莱(Mary Shelley)在1818年就在她的小说《弗兰肯斯坦》中提