

人工智能的决策替代与社会缓冲*

——以 X 智能电厂为例

陈 胤 李晓曼

提要:在新科技革命和产业革命背景下,人工智能表现出替代人类决策的可能性。在 X 智能电厂的数字化改造中,人工智能技术的客观物质属性构成了对工程师的部分决策替代和既有组织权力结构的冲击。企业性质、战略目标、既有利益格局等多重因素使得各方行动者通过算法的重写、策略性弃用、替代的转换等实践,形成了对人工智能技术冲击的社会缓冲。本文的发现揭示了人工智能对人类社会的冲击将引发不同的社会缓冲机制,社会因素成为技术影响社会的缓冲层和过滤器。

关键词:人工智能 决策替代 功能可承担性 社会缓冲

马克思在《资本论》中关于机器替代劳动的研究成为技术社会学和劳动社会学领域的经典范例。正如马克思指出的那样,工业革命时期生产力的革命性变化导致生产关系的一系列重大变革(马克思,2018)。当数字社会来临,新一轮科技和产业革命逐渐深入。随着以人工智能(artificial intelligence)为代表的新一代信息技术开始应用,与传统机器相比具有新特征的人工智能会带来何种意义上的生产关系乃至社会关系变革?“会思考和决策”的机器对人的替代,与传统的重复性体力劳动的机器替代相比,究竟会发生什么样的新变化?

不同于国内外大量对于人工智能及算法在平台经济、互联网科技公司领域的劳动控制效应和不平等效应的关注(Cotton, 2020; Kellogg et al., 2020; Rosenblat, 2019; Casilli & Posada, 2019; 贾文娟、颜文茜, 2022; 朱富强, 2022; 陈龙, 2020; 王林辉等, 2020; 朱琪、刘红英, 2020),我们再次回到马克思经典研究中的工厂领域,探讨人工智能技术的生产和决策应用如何在新型数字化、智能化工厂领域重新塑造既有的关系结构与权力结构;同时探讨在特定场域中社会因素

* 本文是国家社会科学基金一般项目“我国青年技能型人才工匠精神的测度、劳动力市场回报与培养体系研究”(BFA200062)的阶段性成果。感谢匿名评审专家的宝贵意见和精心指导,文责自负。

如何与人工智能技术产生关联,如何影响人工智能技术的具体应用过程。这些经验和理论层面的问题构成本研究的出发点。

一、技术—社会关系视域下的人工智能决策替代问题

人工智能泛指用计算机来模拟人类智能,这一领域涉及机器人、语言识别、图像识别、自然语言处理和专家系统等(腾讯研究院等,2017)。与传统的计算机程序和算法相比,人工智能在解决问题的方式和能力方面都产生了巨大的变化。基于大数据、蒙特卡洛树、神经网络学习等技术的人工智能可以进行自我的进化与迭代,解决过去只有技术专家才能完成的复杂问题(Mnih et al.,2015)。

本文在研究视角上区别于既有的关于算法的社会学研究,不再将人工智能等同于算法,而是将其视为一个整体——一个具有相对主体性的行动者。尽管有学者指出人工智能的核心问题在于算法问题(邱泽奇,2021;赵璐,2022),算法是人工智能运行的规则,具有一种制度性的力量,也具有塑造和改变社会现实的能力,但是当我们更多考虑技术的客观物质性要素后,不难发现影响社会现实的不仅是算法,还包括数据、算力、传感器以及算法的载体——处理器等要素。这些客观物质要素构成人工智能的完整体,使得人工智能成为具有感知能力、思考能力和决策能力的行动者。2023年人工智能领域“人工智能体”(AI agent)的技术方向备受关注,这一演进方向将机器感知、规划、记忆、执行等能力相结合,也体现出AI作为理论意义上的行动者在现实中行动能力逐渐增强的技术发展和应用趋势(Han et al.,2023;Kortemeyer,2023)。

(一)人工智能行动者:从重复劳动替代到决策替代

在机器替代劳动的历史中,非人工智能的机器往往替代的是相对简单、重复和程序化的劳动,潜在的替代对象是相对底层的工人(余玲玲等,2021)。即使在人工智能技术迅速发展后,仍有学者认为人工智能替代的工作也以具有重复性、程序性的工作为主,而那些具有创造力和自主决策能力的工作则较难被人工智能替代(吴文芳、刘洁,2021)。然而,亦有研究指出,人工智能具有更强的劳动替代能力,这种替代能力不仅仅体现在对体力劳动的替代上,还体现在对复杂认知任务尤其是知识经验的替代上。谢斯塔科夫斯基(Shestakofsky,2015,2017)指出,过去的计算机自动化只能执行一些程序性的、固定性的工作,而人工

智能机器学习的出现使得计算机可以模拟人类的表现,通过过去的的数据来推断未来的行动模式,并通过大量的人类生成的数据训练来找到一些默会知识。由此,人工智能替代的工作也更为复杂,从体力劳动转变为脑力劳动,包括传统观念中只有人类才能完成的认知、判断、推理、决策、创作等。而人工智能未来替代的群体也有可能从工人群体转变为技术工作者、律师、法官、工程师和中层管理者(Brayne, 2017; Volokh, 2019; Pasquale, 2019; Burrell & Fourcade, 2021)。2016年,阿尔法围棋对战人类围棋世界冠军的压倒性胜利更是展现出人工智能在认知、判断和决策替代领域的巨大潜力(Silver et al., 2016)。

在第一次工业革命时期,机器对人(劳动力)的大规模替代现象首次出现。机器作为非人行动者在哲学逻辑上具有了行动者的可能,但在技术层面和经验现象层面仍然是作为工具存在。此后历次工业革命的技术进步,包括计算机自动化在内的技术演进,所替代的人类能力水平越来越高,但始终并未真正威胁社会中的脑力劳动者——机器仍然是技术人员手中提升效率的工具。在技术演变的临界点到来前,技术的工具属性相比行动属性更加明显,技术在具体实践过程中主要依附于人。但随着自动化技术的量变逐渐累积,人工智能在实践中的能力越来越强,独立于人的作用越来越强,正在可被感知的层面趋近独立的行动者。人工智能行动能力的增强也就意味着人工智能正在趋向其本身的设计初衷,在越来越高级的层面替代人的脑力劳动。

与以往历次工业革命中机器主要对一线体力劳动者产生替代和冲击不同,在第四次工业革命即将展开之际,尤其是在新型智能工厂开始出现的情境下,人工智能第一次具有了替代认知、思考、判断、决策的潜在技术可能性。这就导致第四次工业革命中技术冲击的靶向可能会发生重要的变化。在马克思的经典工厂研究中,面对机器替代的冲击,底层工人阶级最初所具有的抗争能力是相对较弱的,除了以捣毁机器的卢德运动的方式来抗议外,缺乏更为有效的抵制能力,尽管这一冲击后续引发的社会变革极其深远。当时工业大机器的广泛应用符合资产阶级管理者的利益。而在人工智能的决策替代过程中,技术冲击将更多地指向技术管理群体乃至高层的决策群体,因此这样一个决策替代的历史过程或许会产生更加复杂的社会影响和不确定的社会实践。

(二)关于人工智能技术社会影响的不同理论视角

作为新的信息行动主体(陈甯, 2021),人工智能对人类决策行动的替代会引发何种意义上的社会变化?从技术与社会关系的视角出发,在科学知识社会

学、技术社会学传统的诸多研究中,技术决定论、社会建构论和技术—社会辩证关系等三大理论流派给出了不同的解释。

其一,技术决定论来自马克思主义关于技术和社会关系研究的传统,强调技术在一定程度上决定了社会关系的变化。这一理论视角认为技术尤其是科学技术引发的生产力变革是社会变迁的主要动力,强调技术对社会的塑造作用(中共中央马克思恩格斯列宁斯大林著作编译局编译,2012)。具体到关于机器和工厂的研究中,马克思敏锐地捕捉到第一次工业革命中的机器发展特征,并且指出机器带来的技术冲击既在微观层面对工人产生了影响,又最终带来了大工业的兴起以及工厂制度的普及,给传统的手工业、家庭劳动带来一系列革命性的变化(马克思,2018)。罗素(Stuart Russell)等学者指出人工智能相比于人可以更理性,极大提升决策效率、提升决策公正性(Russell, 2019)。以卡尼曼(Daniel Kahneman)为代表的认知心理学家指出,人类在进行决策时很难避免系统性的认知偏差和谬误(Kahneman, 2013)。相比之下,人工智能等新的机器决策似乎可以避免人为因素对决策的影响,因此在生产领域、管理领域乃至社会公共政策领域,人工智能的应用能够导向一种更加公正的社会结果。

然而,更多经验研究指向了人工智能引发的负面社会后果。席勒(Dan Schiller)、卡斯特(Manuel Castells)等学者针对二战后美国信息技术和通信传播业的发展,系统论述了信息资本主义的概念和历史,揭示了资本巨头在数字社会中的权力核心地位(席勒,2018; Castells, 2010)。伯勒尔(Jenna Burrell)等学者在对算法社会的论述中提出,大型的互联网资本以及互联网精英塑造了新型的软件资本主义(software capitalism)(Burrell & Fourcade, 2021)。学者们指出,在数字社会中,算法统治成为一种生产不平等的方式,算法中心主义也开始逐渐兴起,这成为人工智能早期产生一系列负面影响的重要原因(Burrell & Fourcade, 2021; Engelke, 2020; 陈龙, 2020; 庄家焯, 2019; Rosenblat, 2019; Kellogg et al., 2020)。新的数字技术,也直接在生产领域带来一系列对产业结构、劳动空间、工作方式、工程师心理认同的冲击与变革(MacKenzie et al., 2015; Castells, 2010)。

其二,社会建构论认为技术对社会的影响并不是单向度的。技术决定论的观点引发了以“技术文本”理论为代表的社会建构主义的质疑(Grint & Woolgar, 1997)。人工智能等数字技术究竟能在多大程度上改变社会,是技术的演进决定社会结构的变化,还是社会结构会决定技术发展的走向,重新复制既有的社会运行逻辑?以格林特(Keith Grint)等人为代表的技术社会学流派认为,技术作为一种可以书写的文本被社会建构出来。各种社会因素在具体的时空场景中重新

塑造了特定技术的发展和应用(Grint & Woolgar,1997)。既有的社会权力结构、利益分配格局、关键的利益群体行动者等都会发挥不同的作用,让技术发展和应用不再简单依靠技术的内部演进逻辑,而是受到更多非技术因素的塑造和改变。在这个意义上,人工智能技术并非仅沿着计算机技术演变的内在逻辑发展,而是会受到既有的社会因素的影响,成为一种被社会书写的文本。

人工智能进行的学习—决策过程与大数据之间的内在关联进一步增强了这一理论逻辑的可能性。现有的人工智能技术的大量应用建立在大数据收集的基础上。大数据在一定程度上就是社会行动和经验的痕迹。人工智能通过对大数据的收集和处理来归纳和学习人类社会的特定规律,模拟人的智能化行动。这一过程往往会复制甚至放大社会中既有的不平等现象(Barocas & Selbst,2016; Weinberger,2019; Engelke,2020; Eubanks,2018)。

其三,第三种理论流派则强调技术和社会之间存在一种辩证关系。自20世纪90年代开始,奥利科夫斯基(Wanda J. Orlikowski)、邱泽奇等学者开始在技术—组织领域的一系列研究中试图构建一些综合性的理论框架,以调和决定论与社会建构论之间的理论冲突(Orlikowski,1992; Orlikowski & Gash,1994,2000; Orlikowski & Barley,2001; 邱泽奇,2005)。在吉登斯(Anthony Giddens)的结构化理论影响之下,奥利科夫斯基提出了技术的结构化理论。她认为,以往关于技术—组织关系的视角冲突根源在于哲学上主观性和客观性之间的冲突。技术实际上同时具有二重属性:一方面具有物质属性特征,另一方面又体现出人对技术的能动性。组织结构是一种动态生成的社会过程,是人类行动者和组织结构性特征互动的结果。技术在结构生成过程中既具有客观性的力量,又是社会建构的产物(Orlikowski,1992)。从实践的视角去考察技术会发现,技术是一种实践中的技术(technologies-in-practice),并非简单被社会建构和塑造,而是在技术使用者的使用过程中、在人与技术的互动过程中发生变化(Orlikowski & Gash,2000)。邱泽奇的技术—组织互构论则进一步明确了组织和技术之间的辩证关系。他指出应当从技术的实践特质着手分析,认为技术之于组织目标的重要程度、技术科层制、技术本身的结构弹性等实践特质会和组织结构产生具体的互构,组织和技术在实践中是一种不断相互建构的关系(邱泽奇,2005)。

沿着科学知识社会学和技术社会学的路径,赫卡比(Ian Hutchby)在更宏观的技术—社会关系层面提出了一个相对折中的概念:功能可承担性(affordance)(Hutchby,2001)。这一概念最初源于吉布森(James J. Gibson)的认知心理学概念。赫卡比认为技术的物质属性可以改变社会,但是并非所有的物质属性都可以

得到应用进而改变社会,还需要考虑技术的功能可承担性。福西(Pete Fussey)等指出,数字社会学正是要分析在不同社会环境下的技术功能可承担性,研究这些功能可承担性如何被社会关系、社会互动和社会结构所塑造(Fussey & Roth,2020)。

(三)技术的社会缓冲机制

上述不同理论的启示在于,当我们思考技术与社会的关系时,陷入社会和技术一方或双方主导的争论是无意义的,更重要的是考察技术尤其是革命性技术遭遇社会的实践过程。这是一个动态的、生成性的、行动的、未定的过程。在这样一个实践过程中,技术一旦应用,遭遇到社会后,经历的就并非是一个平滑的过程,而是一个充满着结构制约、行动者抵抗和技术自身适应的非平滑过程。

如果回到社会学中最经典的社会有机体比喻,可以说当社会有机体遭遇到诸如人工智能的技术冲击时,不可避免地会产生一种自发或非自发的免疫机制。这有可能来自既定社会结构的抵抗力量,也有可能来自具有自反性的社会成员的反思性抗拒和改造。社会在面对技术突变的冲击时,并不是被动的接受者,而是作为技术影响的中间层,应对技术变革的不确定性,延续相对稳定的秩序。这也让我们联想到波兰尼(Karl Polanyi)在论述市场制度对民众的冲击时所指出的结论:这一冲击会引发社会自我保护的运动(Polanyi,2001)。

我们更愿意使用“社会缓冲”而非“社会建构”来描述这一技术与社会的关系现象。人类技术的革命性演进最终将带来组织结构乃至社会结构的重大变革。这一总体趋势尽管并非决定论意义上的,但却是具有主导意义和难以扭转的。如果借用功能可承担性的相关概念来分析这一过程的展开,可以看到人工智能相比传统自动化机器而言特有的物质属性构成潜在的功能选项。这些潜在的功能选项并非全部可以得到应用。社会因素筛选出哪些功能可以发挥作用,而发挥作用的功能是被社会影响过的功能。这意味着,社会在影响这些人工智能技术的可能性,同时也被人工智能技术所改变。

本文的理论框架与功能可承担性视角的区别在于,赫卡比等人仅仅提出了功能可承担性的概念和视角,却并未将其进一步明晰化。受此启发,在本研究中我们假定这些客观技术物质特性为关于 A 的一系列向量 $(A_1, A_2, A_3, \dots, A_n)$,在其发挥作用时,需要滤过社会缓冲层 S ,最终在具体的实践过程中选择性地呈现为 SA_1, \dots, SA_n 。在这一过程中,不仅仅技术端的特性 A 被改变,社会端 S 也被悄然重塑,这一理论模型如图1所示。在人工智能真正影响社会之前,我们无法确定地预测人工智能究竟会怎样改变社会。而作为新的行动主体,人工智能

技术遭遇具体的社会 S 的实践过程 P 决定了人工智能影响社会的未来走向,这一过程恰恰是社会学的经验研究应当描述和探讨的。

我们可以进一步思考,为什么在技术的应用过程中会出现技术的可能选项与技术的实际社会后果之间的差异?例如,历史上纺织机技术的发明和应用在英国可以引发工业革命以及后续的社会变革,而在中国古代社会又没有形成决定性的社会影响。功能可承担性的概念部分概括了这一现象。造成这一现象的根本原因正是由于技术不是仅存在于实验室或者真空之中,社会这一重要因素的存在导致技术在穿越社会缓冲层时呈现一种动态的、未知的复杂过程。

如图 1 所示, A 表示技术具备的客观物质特性,这些物质特性具有成为潜在功能的可承担性。 S 表示社会因素,可以视为技术作用于人的社会缓冲层。 SA 则是技术通过社会缓冲层后呈现的社会技术结果,是客观性与主观能动性的辩证统一。

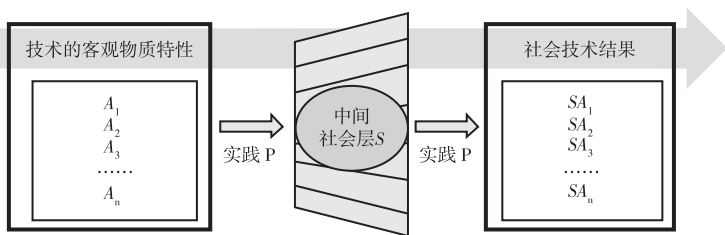


图 1 技术的功能可承担性与社会缓冲机制

具体到人工智能技术,根据目前的发展趋势,这些客观物质特性 A 至少包括两大类。第一类是一般计算机技术具有的物质特性:人工智能具有远高于人类大脑的运算速度;人工智能在进行决策时不受情感、心理认知谬误的影响;人工智能决策过程具有算法依赖性。第二类可以看作人工智能相较于传统计算机技术的新特征:人工智能具有自我决策能力;决策过程具有数据依赖性;人工智能的决策中间环节具有不透明性和不确定性 (Burrell, 2016; Mitchell, 1997; Schmidhuber, 2015; Russell & Norvig, 2016);人工智能结合传感器后具有高于人的感知能力;人工智能具有自我进化的能力,等等。

微观层面的具体实践过程 P 是我们在研究中需要具体考察的过程。同时,实践尽管充满不确定性,但并非一个完全未知的黑箱,它是不同的行动者在既有的实践场景中、在特定的资源和权力结构下行动的过程。

基于以上的理论框架,为了理解人工智能遭遇社会因素后的实践过程,我们

以当下中国电力行业的一个智能发电厂对人工智能的应用为例,探讨在生产和管理实践中,人工智能行动者的物质属性和能力如何改变了工厂原有的社会关系结构,而既有的社会关系、组织和权力结构又如何影响人工智能的部署和技术演进方向。我们并不追求通过一个工厂的案例来推导所有的人工智能应用领域,而是透过这一工厂的人工智能应用情况来管窥人工智能在技术实践中与社会因素的具体作用过程。

二、X 智能电厂的人工智能实践

X 电厂位于中国传统能源大省 S 省,是某大型国有资本控股电力企业中的一家煤炭发电工厂。在电力能源行业信息化和智能化改造的浪潮中,X 电厂在全省范围内率先引进了新一代数字化生产和管理系统,初步形成了基于工业互联网的智能工厂。整个改造过程从 2020 年开始,包含工业互联网、物联网、传感器、数字孪生、人工智能等多个部分的建设。2022 年 5 月,改造项目完成,并被中国电力企业联合会认定“处于国际先进水平”。其中,人工智能的引入包含在整个工厂数字化和智能化的改造过程中,是整个改造中显著的科技创新点,也是区别于传统自动化系统的关键。笔者所在的调研团队于 2020 年年中开始对工厂改造过程进行同步研究,主要研究方法为参与式观察法和访谈法。调查的对象涵盖 X 工厂的一线操作工人、技术人员、中层管理者,以及参与技术改造的人工智能工程师等。

(一) 基本技术流程与改造前的决策权力结构

X 电厂在进行数字技术改造之前已经具备相当高的自动化程度。与传统流水线式的制造业工厂相比,能源生产特别是火力发电对大型设备和机器的依赖性更强,机器的自动化运行程度也相对更高。工人并不直接制造产品,电力的生产通过大型设备来完成。电厂核心的生产过程包括运行和检修两大部分,涉及汽机、锅炉、电气设备、化水、热工、燃料等多个领域。一线操作的主要工作类型是监控机器的运行、操作设定机器参数以及维护检修机器设备等。

在组织架构上,围绕这一核心的电力生产过程,X 电厂形成了管理、运行、设备检修、环保安全、市场运营等五大横向组织架构板块。同时,该厂在纵向上形成了从最基层的一线操作人员到发电厂厂长共五个科层制意义上的级别。我们

把 X 工厂视为一个充满权力和社会关系的场域。这个场域围绕生产的核心任务存在相对固定的权力结构。根据决策权力的不同,我们可以初步将人工智能引入前的工厂划分为三个层级。

第一个层级是一线的操作工人,主要工作包括按照操作手册或者上一级工程师的指令操作具体的设备、关闭和打开阀门、进行设备故障的巡查和维修,以及其他生产环节上的体力劳动。一线操作工人是具体命令的执行者,在绝大部分情况下按照指令或规定流程完成工作,除了遇到一些特殊突发情况外,较少需要进行独立的决策行动。由于高校的扩招以及电力行业较高的福利待遇,X 工厂的一线操作人员大都从高校的相关专业中招聘。这些一线操作工人具备较高水平的相关专业知识和技能,具备向上晋升的空间。

第二个层级是电厂的中层技术管理者,往往拥有工程师的职称,具备较高的专业技术知识水平和丰富的一线工作经验。由于电力生产具有高度的专业性,专业技术人员在日常的核心生产运营中拥有相当大的决策权力。除了部分非技术性管理岗位外,大量中层技术管理者在行政组织上肩负对一线员工的管理职能,也需要在专业技术上对出现的技术问题做出判断、分析和决策。伯里斯(Beverly H. Burris)在对二战后工业组织变化的研究中指出,技术官僚制度(technocracy)是一种重要的组织形式(Burris,1989)。专家在这种组织形式中的地位日益凸显,并在组织中具备相对较大的权力,这压缩了原有的科层制结构。在 X 工厂中,可以看到技术治理和专家治理的趋势,中层管理者和技术专家存在大量的重合,从一线提拔的技术精英、工程师构成了 X 工厂中的中层技术管理者群体。这一群体在工厂实际运行中发挥着至关重要的作用,具备相当大的决策权力。

这些决策权力主要涉及工厂生产过程中的技术问题,如对设备参数设定的判断、对阀门关闭时机的判断和指挥以及对设备复杂故障的诊断和定性,等等。中层技术管理人员凭借多年的工作经验和专业知识技能,往往比年轻的基层员工具备更强的判断能力,比年长的基层员工拥有更多的知识储备和决策权力。在技术官僚制度中,这些技术上的决策权力会进一步延伸到其他日常管理领域,例如设备部门的中层领导和工程师群体对选取何种维修原材料的决策权,对采用何种工艺手段和施工方案维修更新设备的决策权,对专业施工队伍的推荐权,等等。这些决策权力建立在工厂高层对中层技术管理者群体专业性的信任基础上。

第三个层级是电厂的高层管理者。这些管理者负责整个工厂的总体管理和最终决策。但这些决策往往不涉及具体的技术问题,而大都涉及工厂的发展战

略、人事管理等内容。高层管理者需要依赖中层技术管理群体来完成工厂运行的日常管理,也需要通过中层技术管理群体来获取工厂一线运行的各方面信息。

在引入人工智能技术进行决策替代前,X 电厂处于这样一个相对稳定的权力结构之中。整个工厂的运行高度依赖中层技术管理群体。整个层级体系的上升通道也是相对明晰的:一线操作工人通过经验的积累和学习,在具备相对较多的专业知识技能和经验后,可以进入到更高一级的管理层级,凭借专业知识做出更多的技术判断和决策。

(二)人工智能技术的中层替代与冲击

2020 年,X 工厂的数字化和智能化改造工程开始。X 工厂所在集团高层引入这一技术变革的主要目标是获得长远的技术战略地位,即在“双碳”目标下更精准地实现环保和效益的平衡,在提高运营效率和降低污染排放的同时,推动传统能源工厂迭代演进。在项目立项之初,集团和工厂高层就将这一项目视为工厂进入工业 4.0 时代的重要战略发展方向。同时,X 工厂获得了充足的国家资金支持,因而可以在全球的同行企业中相对较早地引入前沿数字技术。

当时,整个工厂开始引入全流程的数字化生产、管理、运维系统。与以往的自动化、信息化系统相比,这一系统的最大特色是基于工业互联网平台消除数据、控制和决策孤岛,全面打通设备、控制和管理全流程,特别是将人工智能技术引入到这一系统的建设过程中。

新型示范智能电厂的建设从价值视角、业务视角、技术视角出发,采用数据+算力+模型+应用的总体逻辑设计,基于全新的工业互联网平台,对企业运营的全流程进行了深度的优化和再造,做到了采用一套智能化系统来支撑电厂的整体运营,实现了管理流程化、流程技术化、技术可视化,可以使得电厂将原来较为线性、单一、封闭的运营管理模式转变为全要素、全价值链、全产业链密切联动的协同、生态、透明的管理模式。(调研资料:202109a1)

根据技术提供方的设计,整个新的系统包括十大板块,基于发电厂的生产流程和管理实践,涵盖决策中心、生产监控、运行管理、设备管理、经营管理、数字孪生工厂、数字安防、数字办公等方面。其中,人工智能主要发挥作用的地方在决策中心、生产监控、运行管理、设备管理、经营管理等板块。

整个项目优化了全厂 13 个重点部门的 108 个流程,全新设计了 1250 个功能页面,定制开发的总代码量约 2000 万行;完成了 1.5 万个设备、6 万根管道的三维可视化;消除了 46 个数据孤岛,部署了 52 个智能模型,有效解决了重点设备的故障预测,实现了全厂主要环节的闭环优化控制;部署了 9 类图像识别算法。(调研资料:202109a1)

在传统的劳动替代过程中,机器主要替代的是工人的体力劳动。这是由工业社会中机器的物质特性所决定的。在工业 4.0 革命之前,机器相比于人所拥有的最主要的物质属性包括更大的机械力量、优于人身体的坚固性和庞大性,以及持续不知疲倦的生产能力,这极大地提升了规模化大工业生产的能力。但是,机器却并不能像人一样智能化地感知、判断和决策。因此,以往的每一次机器进化都会替代大量的从事重复性劳动的一线工人。然而,在 X 工厂引入 52 个人工智能模型后,我们发现机器替代的潜在对象并非仅仅指向基层的一线操作人员,而是直接指向工厂生产环节中至关重要的中层技术管理群体。

图 2 显示了一个典型的人工智能在 X 工厂的技术决策流程。其中,步骤②和步骤③的完成依赖人工智能算法模型的构建,而智能模型的运行又建立在步骤①的数据感知和以往数据获取的基础上。整个决策过程是传感器、物联网、人工智能技术综合运行的产物。



图 2 人工智能在 X 工厂的技术决策流程

在这样一个决策场景中,原本需要由工程师来做出的技术判断和决策被人工智能行动者替代。在传统场景中,巡检工人一旦发现某设备出现参数的异常和表面结构异常,便会将这一现象报告给领导。中层技术管理人员基于专业知识和经验做出判断,召开部门技术会议,提出解决问题的技术方案。在这个过程中,专业技术管理人员具有极高的技术权威性,可以充分发挥自身的专业所长。这一过程既强化了专业技术管理群体的职业认同感和成就感,又巩固了其在工厂中的管理者地位。

可以说,在引入人工智能模型之前,中层技术管理群体拥有丰富的专业知识

和实践经验,这些知识技能和经验构成了工程师们决策权力的护城河。随着时间的推移和人工智能自我学习进化速度的加快,人工智能做出的技术决策可以变得更加合理化。X工厂中的中层技术管理群体的技术决策护城河正在随着人工智能行动者的加入而逐渐地被跨越。随着数字化系统逐渐部署和使用,X工厂的工程师们发现,假定在没有任何外力干涉的情况下,完全按照乙方技术公司的设计来运营,电厂将出现明显的中层技术管理人员被人工智能替代的现象。这种替代的速度在理论上只取决于人工智能技术决策能力进化的程度。

人工智能带来的冲击是多方面的。在工厂的日常生产和运营过程中,技术决策能力所附加的影响并不局限于生产本身,而是关涉到更多具体的利益问题。前文的分析已经指出,在一个接近技术官僚治理的制度架构中,技术上的权威与组织上的管理权力在一定程度上是可以相互转化的。技术精英尽管未必完全占据一把手领导位置,但是在涉及技术的相关决策上拥有重大的话语权。人工智能的冲击如果不受干扰的话,将至少带来以下几个方面的变化。

首先,岗位替代会带来组织结构的变化。52个人工智能模型的引入使得相关环节的工程师工作被部分替代,原本的一些岗位失去了存在的意义。这些岗位包括具备一定技术门槛的一线操作岗位,但更多地涉及中层工程师群体。根据我们的调查,在进行数字化改造后,X工厂A部门的员工需求人数从36人缩减到4人(调研资料:202109a2)。

其次,由于在整个智能化系统的改造过程中,技术问题的决策与经营、维修系统形成了直接的关联,人工智能系统在技术决策后直接给出了相关经营问题决策,这进一步削弱了中层技术管理人员的实际权力。例如,在人工智能引入之前,技术管理人员对设备维修问题的诊断和决策背后关联着材料供应商、维修提供商的选择问题,而人工智能的直接介入使得中层技术管理人员的选择往往与人工智能的选择形成鲜明对比。如果完全采用人工智能的决策,则意味着技术管理人员的传统权力被进一步剥夺。假如存在权力寻租空间的话,那么这一寻租空间也将被进一步压缩。

最后,人工智能对人事考核、组织和财务的优化管理和直接介入,将会直接打破原有的利益格局。传统工厂中的师徒关系等私人关系、情感关系、忠诚关系并没有被设定为人工智能运行的参数。完全去人情化的行动方式是目前人工智能进行管理的一大物质特性。原有的技术管理群体的管理工作建立在技术权威的合法性基础上,同时并不会完全排斥非理性因素的考量,而人工智能一方面通过对技术决策的替代削弱了技术权威的合法性,另一方面又有能力直接介入到

对人的管理和控制过程中。

如果人工智能的引入仅仅是一个单向度的过程,仅从效率最大化的角度出发,不用考虑社会因素的存在,那么在 X 工厂这一具体的实践场域中,潜在的变革即将发生,整个 X 工厂也处在变革的前夜,蕴含着极大的不确定性。

(三)作为“工程师”的人工智能行动者

在 X 工厂的整体数字化改造过程中,人工智能的应用仅仅是数字化改造的一部分,但却呈现了不同于以往信息技术应用的行动者特征。在此次数字化改造前,X 电厂已经是一个相对自动化的工厂,信息系统已经应用在工厂的日常管理和经营之中。而此次的数字化改造将管理、经营与生产流程进行了系统融合,人工智能则是历史上第一次嵌入在工厂的整个生产管理系统之中。

与以往历次信息化升级不同,这一次人工智能所具有的物质属性使得人工智能对于部分中层技术人员而言不再是简单的工具,而是具备同等技术判断和技术决策能力的行动者。从观察者的视角来看,人工智能同时具有工具和行动主体的二元属性。这种属性呈现在一个等级制的结构关系中。从工厂高层决策者和管理者的视角来看,人工智能是实现其生产目标的工具。但是对于中层技术管理人员而言,人工智能则是一位同样可以进行技术判断和工程决策的“工程师”竞争者。

如图 3 所示,对高层决策者而言,人工智能是实现高层技术目标的工具。但是当确立了最高行动目标后,人工智能在完成目标的中间环节可以具有自我决策的能力,并与其他行动者构成平等的竞争关系。当然,并非所有的人工智能模型都直接替代了工程师,那些能力暂时不足以替代工程师的人工智能模型仍然被作为工程师的参考工具来使用。^①

为什么在之前的历次计算机应用升级中计算机仅仅是作为工程师的工具来帮助工程师完成任务,而在 X 电厂此次改革中,部分人工智能成为与工程师具有同等行动地位的竞争者,具备一种决策替代的能力?

首要的原因是神经网络模型等机器学习类模型在生产过程中的应用。这些机器学习类模型使得 X 工厂的人工智能具有自我决策的能力。而这种自我决策在过去的自动化程序中是难以实现的。在 X 工厂早期应用的计算机程序中,计算机算法是固定式的、机械意义上的程序。当人工智能模型引进之后,计算机

^① X 工厂还存在很多没有应用人工智能的生产环节。同时人工智能的部分决策需要专家来监督以保证安全,这涉及人对人工智能的信任问题以及其他制度问题。

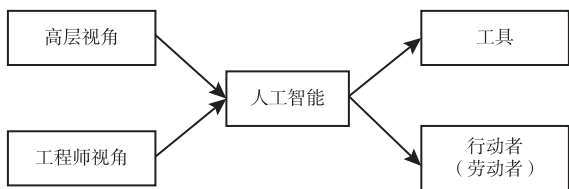


图3 人工智能在实践中的工具和行动者二重性

可以对复杂和不确定状况进行判断和决策,而不是简单映射既有的固定程序。

X工厂的人工智能应用以化学物理模型、概率数理模型和神经网络模型等三类模型为基础。其中,基于神经网络的拟合能力可以描述和推理事件之间的因果关系。即使在无法确定真正因果关系的情况下,神经网络模型也可以通过专家经验数据学习、统计规律计算、模糊推理不断逼近最优解。X工厂中部分神经网络模型的示意图如图4所示。

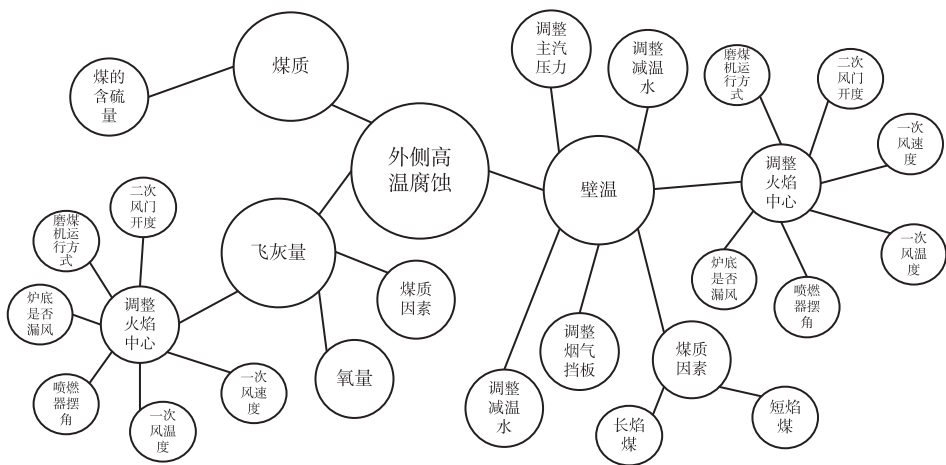


图4 X工厂部分神经网络模型示意图

X工厂应用的神经网络学习模型意味着计算机构建了关于各种设备现象(指标)和决策之间因果性关系的复杂函数,并且不断逼近拟合一系列复杂函数,直接建立起现象和技术决策之间的关联。理论上,这一模型不完全依赖对工程师做出决策之原理的理解,只需要对现象和决策之间的关联进行大数据学习,直接做出各种复杂状况下的判断和决策(Mnih et al.,2015)。

此外,与X工厂的工程师相比,人工智能具有明显的知识进化优势。由于

机器学习的特性,人工智能在每一次做出技术决策后都会将新生成的数据结果与原有的模型进行比较检验,以实现一种自我趋优的演进。同时,不断更新的海量专家知识数据库也远远超出一般工程师的知识储备。

最后,X工厂的全流程数字化改造、传感器应用以及数字孪生工厂的空间打造使得人工智能具备在数字空间中自我感知和行动的能力。在数字空间中,人工智能和工程师的决策指令同样是以信息流方式传递到下一环节,二者具备相同的信息生产和信息行动的主体地位。

X工厂传统的自动化技术决策思路和人工智能引入后的技术决策产生了显著差异。在智能化改造前,技术专家将计算机视为工具,依据客观计算结果或者指标进行判断和决策,发挥人的知识、经验和直觉判断优势。而当人工智能出现后,技术专家大部分的经验、默会知识和灵感尽管无法被人工智能理解,但在理论上大多是可以被替代的。

(四)行动者冲突与社会缓冲机制的出现

2021年年中,X工厂的人工智能等数字系统硬件部署基本完成,开始进入试运行阶段。此时,意料之中的阻力出现了。与传感器、物联网等其他技术的顺利应用相比,人工智能决策替代的相关应用遭遇了较大的阻力,相应的抗议者是工厂的中层技术管理群体。一位人工智能的工程师在接受笔者的访谈时谈到了人工智能决策替代应用所面临的阻力。

我们在进行测试的时候遇到了厂里工程师们的普遍质疑,他们认为这些做法是不正确的,由机器来判断,风险太大。(访谈资料:202112b5)

C车间主任B强烈反对在他的车间使用人工智能,认为人工智能给出的操作指令并不符合以往的经验。

这个系统给出的关闭阀门的时机不对,这个太危险了,环保上要是出问题了,也是会被罚款的。我们不会按照它(人工智能)的指令去做。(访谈资料:202109c1)

而大多数一线的操作人员则对人工智能的引入持开放或者欢迎的态度。

工作更加简单了,比原来的工作量小了很多,很容易的,没有不习惯。
(访谈资料:202109c3)

在此期间,一个小插曲体现了中层技术管理群体和乙方人工智能技术设计人员之间的矛盾。

D 车间负责技术管理的中层领导在过去一直采用相对稳定的施工队伍、材料供应商和施工方案来解决某设备维修中的技术问题。而新的人工智能模型将设备问题诊断、预测和材料采购的市场化方案进行了联动,得出了与过去并不相同的问题处理方案。D 车间的工程师们认为人工智能给出的技术诊断和方案未必适合特殊的情况,应当具体问题具体分析,并且对人工智能模型中的自动采购、自动市场竞价、分析最优解决方案提出质疑,认为其存在严重的不透明性,体现了算法设计者的利益。算法的不透明性导致作为第三方观察者的研究人员较难判断这一算法是否暗含着设计者的经济利益。但我们至少可以确定的是,D 车间的技术诊断—技术方案—市场采购过程有可能蕴含利益空间,与此相关的权力过去由技术管理人员掌控,现在转移到人工智能的算法设计者手中,形成了一种算法中心主义的局面。

这一事件表明,在实现最终目标的中间过程中,作为实践中的行动主体,人工智能行动者与人类行动者有可能存在目标和利益上的冲突。我们用图 5 来表示 X 工厂这两起事件中人工智能和工程师之间的行动目标冲突。

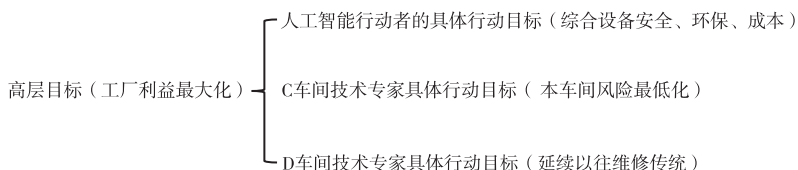


图 5 人工智能行动者和技术专家的行动目标冲突

在本质上,人工智能作为高层决策者的工具,体现了技术应用主体的行动目标和行动利益。人工智能和工程师之间的冲突可以被视为高层决策者与中层工程师之间的利益博弈。但是从具体的实践过程来看,人工智能在执行最高目标的中间环节已经体现出和名义上同样代表工厂(高层)利益的工程师之间的目标冲突。

X 工厂高层自身也面临人工智能技术带来的压力。一方面,中层技术管理

群体向上提出各种意见,这有可能成为工厂运营的不稳定因素;另一方面,人工智能在全厂绩效考核、人事管理等方面的应用过于智能化,由此而来的一些计算结果尽管不会直接被高层采用,但也威胁到工厂高层做出人事安排等管理行动的合法性。例如,按照人工智能的决策裁撤掉大量的技术管理岗位,这对于一个追求经济效益稳定以及政治效益和社会效益的能源型国有企业而言是不太可能发生的事情。于是,人工智能技术在 X 工厂的应用出现了新的变数,这些变数实际上也是 X 工厂中各方新的行动策略。

1. 算法的重写

这些变数首先来自人工智能技术供应方与 X 工厂技术管理群体之间不断持续和深化的互动,即人工智能算法设计者和中层管理者之间社会关系的建立与维系。为了更好地推进智慧化改造,人工智能的技术提供单位与中层技术管理群体进行了持续的沟通和协商。在 D 车间,由于技术诊断—技术方案—市场采购的传统流程被人工智能打断,相关人员在与技术供应方的沟通过程中提出希望能回到过去的决策方式。

D 车间的领导们希望能够维持过去的惯例,认为过去没有人工智能替他们做出决定之前,一切都运转得很好,公司并没有因此遭受损失。他们认为我们有可能会操纵市场采购结果。而我们认为算法是公正、公开的,符合最优结果。但是他们认为这个过程并不透明。(访谈资料:202112b1)

这是课题组在调研中对人工智能技术供应方访谈后获取的信息。出于商业要求,大多数算法本身并不能完全公开,人工智能的“思考”机制本身也具有不透明性(Burrell,2016)。于是,人工智能的算法对于人工智能的使用者而言实际上是一种不透明的黑箱。不少学者指出,随着人工智能在社会各领域的渗透,未来的数字社会在一定程度上将成为一种“黑箱社会”(Brevini & Pasquale,2020)。从第三方的观察视角来看,如果说人工智能介入之前的决策过程存在一种不透明的人为因素,有可能蕴含着某种潜规则的话,那么人工智能使用算法来替代人为决策的做法并没有让决策过程更加透明,反而重新引发了使用者对算法本身的质疑。

人工智能技术供应方的技术转变证明算法本身蕴含着社会运作的空间。在与 D 车间充分协商后,人工智能技术供应方修改了特定的参数,重写了部分算法,成功实现了人工智能做出的决策与车间过去的惯例相符合。D 车间认为,修

改前的算法体现了人工智能算法设计者的利益,而修改后的算法可以更好地符合过去的惯例,降低新的市场采购带来的不确定性。如前所述,算法的设计实际上剥夺了传统技术管理者的相关决策权力,将技术决策和市场采购决策的权力转移到算法上。而双方在协商中重新平衡了决策的权力,于是产生了算法重写的结果。但是,算法的改写并没有在根本上改变或阻止人工智能的介入。

2. 策略性弃用

在 X 工厂智慧化改造的十大板块中,并非所有的板块在改造验收完工后都完全投入了使用。其中的原因不是乙方提供的技术不能使用,而是 X 工厂的高层决策者认为有些技术并不适合立刻投入使用。

在十大板块中,涉及替代一线操作工人的技术几乎都立刻投入使用。例如,智能摄像头和人工智能图像识别技术直接的替代对象是一线的操作工人:过去由工人巡查并向上一级汇报设备安全隐患和运行状况,现在这些工作直接由机器完成。这些技术替代的显著特征是直接面向机器,较少涉及复杂的人际利益关系。

而涉及替代中层管理群体的技术是在反复修改和博弈后才使用的。这些技术往往包含着相对复杂的技术判断,同时关联着部分组织管理和市场决策职能。现阶段,由于受到能力限制,人工智能尚不能覆盖所有的复杂判断。因此,中层技术管理群体虽然在感受到人工智能的潜在威胁后试图对其进行抵制,但由于自身地位仍然不可能在短时被完全取代,所以在策略性地消解人工智能直接带来的利益冲突后,也接受了人工智能的应用。

最后,在人事管理、绩效考核、市场和采购等环节,X 工厂选择搁置人工智能的使用。这些领域的人工智能决策不仅对中层技术管理群体的决策权力产生影响,还可能直接影响高层的决策权力。对于高层决策者而言,人工智能的决策结果或许仅是一种参考,真正面向人与利益的决策权力在现阶段仍然需要掌握在自己手里。这些领域的人工智能决策结果往往过于理性和模型化,完全基于效率考虑,并不能反映工厂高层决策的复杂性。

因此,在 2021 年年中,尽管从技术角度而言 X 智慧工厂的改造已经完成,但人工智能在一些板块仍然处于可用却不用状态。人工智能的客观技术特性使其理论上可以在一座现代化工厂中完成很多功能,但是,并非所有的功能选项都可以透过社会的缓冲层,最终成为一种真正被接受和认可的技术。

3. 替代的转换

尽管人工智能的技术选项受到多重社会因素的影响,但是我们仍然看到人工智能带来的冲击的直接后果。整个工厂的组织结构发生了变化。在 X 工厂

的某技术部门 A,人员编制由原有的 36 名缩减到 4 名。这一技术部门在智能化改造后被完全取消,该部门人员与其他部门被取代的人员成为 X 工厂需要解决的难题。

面对这一局面,X 工厂采用一种替代的转换策略。在 A 部门取消后,X 工厂成立了新的部门 B。部门 B 的主要职能是应对信息化和智能化的需求,理论上讲,其最优策略是招聘具有人工智能等方面知识技能的人员,但现阶段的主要成员来自 A 部门中被人工智能替代的技术人员。X 工厂相信这些技术人员能够重新进行新岗位的技能学习,并在原有专业知识的基础上掌握新的数字技术,与过去的专业技能进行有效的融合。如果完全按照市场利益最大化的逻辑,X 工厂应当裁撤掉这些被替代的人员,从而最大化地降低人工成本。但是作为处于行业领先地位的大型能源国企,X 工厂需要考虑经济因素之外的诸多政治因素和社会因素,工厂的稳定运行、人心向背以及社会责任等都是 X 工厂的高层所需要思考的内容。因此,在全省乃至全行业率先引入人工智能后,X 工厂并未出现大量裁撤中层管理人员和工程师的情况,而是成功实行了组织结构和就业岗位转化。

X 工厂面临人工智能决策替代带来的岗位变化,之所以出现这种非效率的行动策略,我们认为至少是出于以下的社会因素考虑。首先是工厂的国有企业性质。与马克思研究的机器替代工人的资本主义工厂相比,我们所研究的是一个具有不同所有制性质的生产单位。国有企业尽管仍然面临着市场效益的经济压力,也还是会更多地考虑员工队伍的稳定性,并且对员工的岗位技能转化时间有较大的包容性,没有实行技术决策岗位替代后的简单裁员。其次,电力行业面临的外部市场竞争虽然正在加剧,但仍然具有较强的垄断性。因此,人工智能等数字技术的引入更多的是前瞻性和战略性的行动,并非单纯的资本驱动和经济效率驱动。在这一过程中,人工智能的决策替代所承载的目的不在于直接降低人力成本支出,而是更好地获得长远的技术战略地位,以及在国家“双碳”战略下更精准地实现环保和效益的平衡。总的来看,在这一具体的实践过程中,生产资料所有制、技术改革需要维持稳定的动机、远期规划等因素的影响恰好构成了对人工智能决策替代的另一种社会缓冲效应,使得人工智能在 X 工厂的替代冲击没有完全按照技术理性和经济效率至上的逻辑发生。

这种替代转换的方案并非 X 工厂首创。根据世界经济论坛 2021 年的研究报告,惠普公司在进行第四次工业革命转型的过程中将技术变革释放的大量任务职责进行转换,重新分配了企业内部不同角色的任务时间。原有的一线操作

人员能够执行之前委派给技术专家的复杂任务,而工程师可接受更高级的技能提升,承担数据科学的工作(World Economic Forum,2021)。

当然,尽管在当下的时间节点上,这种人工智能替代的转换并未给现有的中层技术管理群体带来直接失业的影响,但是从长远来看却改变了未来的岗位要求。A部门的取消使得未来具备A部门知识,也就是传统电力相关知识的潜在求职者受到极大的影响。新设立的B部门目前还可以容忍传统专业人员的转岗,但是未来需要直接招聘具备人工智能等相关技术知识的人员。这也意味着,尽管X工厂通过替代转换的策略在时间上延缓了人工智能对特定岗位的技术冲击与岗位取代,却无法在更长的时间段上扭转人工智能对特定群体的替代趋势。当人工智能的技术长矛注定要刺穿某些传统的结构时,各种社会因素构成的缓冲层在时间维度上造成了对技术冲击后果的延缓,但是潜在的结构性改变已然发生。

最终,在2022年年中,X工厂顺利完成了智能化改造,在很多岗位上实现了人与人工智能共存的局面。总的来看,这种人机共存的局面是通过社会因素影响下的三种微观策略实现的:鉴于人工智能工程师和中层技术管理人员的社会关系,人工智能技术供应方为了人工智能应用能顺利展开,改变了原本的算法,重新建构了人工智能技术本身,延续了原有的权力结构;策略性弃用的应对策略表明社会结构因素的直接筛选作用,即那些影响到关键权力和高层利益的技术可能性并不会在短时间内穿透社会缓冲层;而替代的转换策略则表明当技术遭遇社会缓冲时,岗位空间和冲击时间会相互转化,工厂用岗位的转换来换取时间,人工智能决策替代的岗位冲击直接后果被延后了,但是组织空间结构已经发生变化。

在X工厂的案例中,人的因素对人工智能技术的改写、筛选和重塑更接近一种权宜性与自我保护性的机制,但这并不能从根本上重新建构人工智能技术的影响和冲击。权宜性意味着时间变量和结构空间变量被引入,即通过组织结构变化来吸纳技术带来的认知劳动替代效应,延缓人工智能的劳动替代行动的短时冲击。保护性意味着当中层技术管理群体无法抵制人工智能时,需要来自群体外部的社会力量对其进行保护。同时,这一现象也印证了邱泽奇等学者关于技术—组织互构机制的判断(邱泽奇,2005)。X工厂的人工智能技术在技术物质属性和能力上的显著进化直接触及生产的核心内容,具有对部分工程师的直接替代能力。那些涉及环保、技术、安全的人工智能技术具有极强的技术刚性,直接刺穿了部分原有的组织资源—权力分配结构。

三、结 论

与传统机器替代一线工人带来的影响相比,人工智能在决策能力上对中层技术管理群体的替代成为一种新的现象,这一现象也是人类社会进入到第四次工业革命和智能化阶段后的新历史现象,构成了对工厂技术管理体制和传统权力决策结构的冲击。

在历史的维度上,X工厂的数字化改造很有可能刚好处于人工智能技术爆发的临界点。X工厂数字化改造中的一部分人工智能技术与当下的 chatgpt、AI agent 等人工智能技术一起,已经呈现跨越由经验、默会知识、灵感、创造力等构成的人类智力护城河的迹象。正如罗马时期恺撒跨越卢比孔河后西方历史出现了重要的变化,我们虽不敢完全断言人工智能正在经历跨越卢比孔河时刻,但是不得不对其潜在的变革可能性进行慎重的理性思考。一旦人工智能真正跨越卢比孔河,社会对人工智能技术的缓冲就不仅是一种客观的免疫式反应,还应当成为一种社会各方主动发起的伦理学议题。

在这一经验研究中,X工厂的高层决策者与人工智能的技术供应方通过微观层面的社会行动策略,最大程度降低了人工智能技术带来的直接冲击,使得X工厂的人工智能应用得以相对顺利地实现。而中层技术管理人员自身在与人工智能的行动冲突中,基本处于无力直接抵抗的状态——部分中层技术专家在面对人工智能时并没有机会在“使用—被使用”的关系中重构技术,而是在对等关系中面临被取代的风险。这也使得社会对技术的对冲不得来自资源—权力结构的更高层级。

国内早期的深度数字化和智能化改造与升级大多从大型国有企业开始。本文案例中的工厂与民营企业 and 国外企业存在产权结构、经济社会利益导向、工厂管理运作逻辑上的不同。在理论上,这些不同的社会因素对人工智能技术的缓冲机制会造成不同的影响。当X工厂面临人工智能带来的负面冲击时,出于政治责任和社会责任等因素的考虑,工厂高层行动者采用种种策略来延缓技术的影响,以实现长期战略目标。而更加成本导向的企业则很可能通过裁员的形式来应对冲击,从而回避技术冲击的社会后果。但是人工智能的社会冲击并不会因此消失,只不过转移到了企业外部的社会层面。这就需要在工厂组织之外,在更广阔的国家和社会层面来进行技术冲击的社会缓冲。

从另一个角度来看,在人工智能决策替代的技术冲击面前,大型国有企业的

应对举措也为更具普遍意义的社会缓冲提供了一种实践可能性,体现出中国特有的社会、经济和文化影响下的社会缓冲机制。中西方社会之间更宏观的政治制度差异、文化传统差异和社会保障制度差异都有可能使得它们在面对未来人工智能技术冲击时的社会缓冲过程演变出不同版本。在一定意义上,未来国家间人工智能技术应用的竞争走向或许也与各国应对技术社会风险的社会保护制度密切相关。这有待于我们通过更进一步的研究加以讨论。

参考文献:

- 陈斌,2021,《信息行动理论——数字社会时代的社会行动理论探讨》,《社会学评论》第5期。
- 陈龙,2020,《“数字控制”下的劳动秩序——外卖骑手的劳动控制研究》,《社会学研究》第6期。
- 贾文娟、颜文茜,2022,《认知劳动与数据标注中的劳动控制——以N人工智能公司为例》,《社会学研究》第5期。
- 马克思,2018,《资本论(纪念版)》第一卷,中共中央马克思恩格斯列宁斯大林著作编译局编译,北京:人民出版社。
- 邱泽奇,2005,《技术与组织的互构——以信息技术在制造企业的应用为例》,《社会学研究》第2期。
- ,2021,《算法向善选择背后的权衡与博弈》,《人民论坛》第21期。
- 腾讯研究院等,2017,《人工智能——国家人工智能战略行动抓手》,北京:中国人民大学出版社。
- 王林辉、胡晟明、董直庆,2020,《人工智能技术会诱致劳动收入不平等吗——模型推演与分类评估》,《中国工业经济》第4期。
- 吴文芳、刘洁,2021,《新技术变革时代“人”的变迁与社会法回应》,《学术月刊》第8期。
- 席勒,丹,2018,《信息资本主义的兴起与扩张——网络与尼克松时代》,翟秀凤译,王维佳校译,北京:北京大学出版社。
- 余玲铮、魏下海、孙中伟、吴春秀,2021,《工业机器人、工作任务与非常规能力溢价——来自制造业“企业—工人”匹配调查的证据》,《管理世界》第1期。
- 赵璐,2022,《算法实践的社会建构——以某信息分发平台为例》,《社会学研究》第4期。
- 中共中央马克思恩格斯列宁斯大林著作编译局编译,2012,《马克思恩格斯选集》第一卷,北京:人民出版社。
- 朱富强,2022,《人工智能时代的价值创造和分配——不平等加剧的社会和经济基础》,《财经问题研究》第3期。
- 朱琪、刘红英,2020,《人工智能技术变革的收入分配效应研究:前沿进展与综述》,《中国人口科学》第2期。
- 庄家炽,2019,《资本监管与工人劳动自主性——以快递工人劳动过程为例》,《社会发展研究》第2期。
- Barocas, Solon & Andrew D. Selbst 2016, “Big Data’s Disparate Impact.” *California Law Review* 104.
- Brayne, Sarah 2017, “Big Data Surveillance: The Case of Policing.” *American Sociological Review* 5.
- Brevini, Benedetta & Frank Pasquale 2020, “Revisiting the Black Box Society by Rethinking the Political Economy of Big Data.” *Big Data & Society* 7.

- Burrell, Jenna 2016, "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3.
- Burrell, Jenna & Marion Fourcade 2021, "The Society of Algorithms." *Annual Review of Sociology* 47.
- Burris, Beverly H. 1989, "Technocratic Organization and Control." *Organization Studies* 10.
- Casilli, Antonio A. & Julian Posada 2019, "The Platformization of Labor and Society." In M. Graham & W. H. Dutton (eds.), *Society and the Internet: How Networks of Information and Communication Are Changing Our Lives*. Oxford: Oxford University Press.
- Castells, Manuel 2010, *The Rise of the Network Society*. Chichester: Wiley-Blackwell.
- Cottom, Tressie McMillan 2020, "Where Platform Capitalism and Racial Capitalism Meet: The Sociology of Race and Racism in the Digital Society." *Sociology of Race and Ethnicity* 6.
- Engelke, Peter 2020, *AI, Society, and Governance: An Introduction*. Atlantic Council (<http://www.jstor.org/stable/resrep29327>).
- Eubanks, Virginia 2018, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St Martin's Press.
- Fussey, Pete & Slike Roth 2020, "Digitizing Sociology: Continuity and Change in the Internet Era." *Sociology* 54.
- Grint, Keith & Steve Woolgar 1997, *The Machine at Work: Technology, Work and Organization*. Cambridge: Blackwell.
- Han, Elizabeth, Dezhi Yin & Han Zhang 2023, "Bots with Feelings: Should AI Agents Express Positive Emotion in Customer Service?" *Information Systems Research* 34.
- Hutchby, Ian 2001, "Technologies, Texts and Affordances." *Sociology* 35.
- Kahneman, Daniel 2013, *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
- Kellogg, Katherine C., Melissa A. Valentine & Angèle Christin 2020, "Algorithms at Work: The New Contested Terrain of Control." *Academy of Management Annals* 14.
- Kortemeyer, Gerd 2023, "Could an Artificial-Intelligence Agent Pass an Introductory Physics Course?" *Physical Review Physics Education Research* 19.
- MacKenzie, Roberts, Abigail Marks & Kate Morgan 2015, "Technology, Affordances and Occupational Identity Amongst Older Telecommunications Engineers: From Living Machines to Black-boxes." *Sociology* 51.
- Mitchell, Tom M. 1997, *Machine Learning*. Maidenhead: McGraw-Hill.
- Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fiedjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg & Hassabis Demis 2015, "Human-level Control Through Deep Reinforcement Learning." *Nature* 518(7540).
- Orlikowski, Wanda J. 1992, "The Duality of Technology: Rethinking the Concept of Technology in Organizations." *Organization Science* 3.
- Orlikowski, Wanda J. & Debra Carol Gash 1994, "Technological Frames: Making Sense of Information Technology in Organizations." *ACM Transactions on Information Systems* 12.
- 2000, "Using Technology and Constituting Structures: A Practice Lens for Studying Technology in

- Organizations.” *Organization Science* 11.
- Orlikowski, Wanda J. & Stephen R. Barley 2001, “Technology and Institutions: What Can Research on Information Technology and Research on Organizations Learn from Each Other?” *MIS Quarterly* 25.
- Pasquale, Frank 2019, “A Rule of Persons, Not Machines: The Limits of Legal Automation.” *The George Washington Law Review* 87.
- Polanyi, Karl 2001, *The Great Transformation: The Political and Economic Origins of Our Time*. Boston: Beacon Press.
- Rosenblat, Alex 2019, *Uberland: How Algorithms Are Rewriting the Rules of Work*. Berkeley: University of California Press.
- Russell, Stuart 2019, *Human Compatible: Artificial Intelligence and the Problem of Control*. New York: Viking.
- Russell, Stuart & Peter Norvig 2016, *Artificial Intelligence: A Modern Approach (Third Edition)*. Harlow: Pearson.
- Schmidhuber, Jürgen 2015, “Deep Learning in Neural Networks: An Overview.” *Neural Networks* 61.
- Shestakofsky, Benjamin 2015, “More Machinery, Less Labor.” *Berkeley Journal of Sociology* 59.
- 2017, “Working Algorithms: Software Automation and the Future of Work.” *Work and Occupations* 44.
- Silver, David, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel & Demis Hassabis 2016, “Mastering the Game of Go With Deep Neural Networks and Tree Search.” *Nature* 529(7587).
- Volokh, Eugene 2019, “Chief Justice Robots.” *Duke Law* 68.
- Weinberger, David 2019, “How Machine Learning Pushes Us to Define Fairness.” *Harvard Business Review*, November 6.
- World Economic Forum 2021, *Global Lighthouse Network: Reimagining Operations for Growth* (<https://www.weforum.org/publications/global-lighthouse-network-reimagining-operations-for-growth/>).

作者单位: 中共中央党校(国家行政学院)社会和生态文明教研部(陈胤)
首都经贸大学劳动经济学院(李晓曼)
责任编辑: 向静林