

• 专题：人工智能与当代社会 •

编者按：

近年来人工智能迅速发展，对社会生活的影响日益广泛而深刻。本专题共三篇文章，分别从不同层次、不同视角探讨了人工智能对当代社会的影响。袁伟的文章在梳理人工智能发展历史的基础上，从历史唯物主义的视角探讨了人工智能的本质，指出了人工智能与人的根本区别所在，认为人工智能对社会的益处和风险均来自于其应用，而其应用又从根本上取决于其所处的社会生产方式，因而人工智能究竟是全面服务社会还是给社会带来风险，从根本上说取决于生产方式。刘鸿宇、彭拾和王珏的文章借助大数据分析软件，对社会科学引文索引数据库 (SSCI) 中 1998-2017 年的人工智能心理学研究文献进行知识图谱分析，展现了九十年代末以来人工智能心理学研究的 6 个知识聚类，并总结了人工智能心理学发展的各个阶段及其特征。王磊的文章提炼了一个描述人工智能赋权特征的概念——参差赋权，即人工智能技术赋权过程及其结果展现的非均衡性样态，他认为参差赋权是人工智能赋权的基本形态，并探讨了参差赋权的发生逻辑，分析了参差赋权可能给社会带来的风险，并提出了应对风险的防范策略。

(专题策划：袁伟)

人工智能：统治人类还是服务人类？ ——基于历史唯物主义的思考

Is Artificial Intelligence to Rule or Serve Human Beings? Some Reflections Based on Historical Materialism

袁伟 /YUAN Wei

(中山大学马克思主义学院，广东广州，510275)
(School of Marxism, Sun Yat-sen University, Guangzhou, Guangdong, 510275)

摘 要：人工智能引起了社会各界的高度关注，并正在广泛而深刻地影响着社会生活。人们在享受人工智能带来的好处和便利的同时，也对人工智能的风险产生了巨大的担忧。人工智能究竟会统治人类还是服务人类？历史唯物主义为这个问题提供了一个独特的思考视角。从历史唯物主义的视角看，人工智能的本质是人的本质力量的对象化，是人类实践活动的成果。人工智能的活动与人的活动有本质的区别，它不会实践，因而人工智能不是人，更不会全面超越人类和统治人类。人工智能的益处和风险均来自于其应用，而其应用又从根本上取决于其所处的社会生产方式。故人工智能究竟是全面服务人类还是给人类带来风险，从根本上取决于生产方式。

基金项目：国家社会科学基金青年项目“政治经济学批判与历史唯物主义研究”（项目编号：17CZX002）。

收稿日期：2020 年 10 月 20 日

作者简介：袁 伟（1988-）男，湖南永州人，中山大学马克思主义学院助理教授，研究方向为马克思主义社会发展理论和资本理论。Email: yuanw26@mail.sysu.edu.cn

关键词: 人工智能 历史唯物主义 生产方式 实践 对象化

Abstract: Artificial intelligence (AI) has attracted great attention from all walks of life and is affecting social life extensively and profoundly. While people enjoy the benefits and conveniences brought by AI, they also have great concerns about its risks. Is AI to rule or serve human beings? Historical materialism provides a unique perspective for reflecting upon the issue. The essence of AI is the objectification of human's essential power and the achievement of human's practice. The activities of AI are essentially different from human activities. AI can not practise. Therefore, AI is not human beings, let alone surpasses them in all aspects and rules them. Both the benefits and risks of AI come from its application, which depends fundamentally on the mode of production. Hence whether it serves or endangers human beings depends on the mode of production.

Key Words: Artificial Intelligence; Historical materialism; Mode of production; Practice; Objectification

中图分类号: A81 文献标识码: A DOI: 10.15994/j.1000-0763.2021.02.001

人工智能无疑是当今世界最热门的高新技术之一,它正在走进我们的生活,并正在改变着我们的生活,与我们的关系日益密切。美国人工智能专家杰瑞·卡普兰(Jerry Kaplan)指出,“即将袭来的机器人、机器学习以及电子个人助手可能会开创一个全新的世界”。([1], p.XVIII)近年来,我国人文与社会科学学界对人工智能也展开了思考和反思。有的学者从心灵哲学的角度看待人工智能,认为它和心灵哲学有“明显的学术牵扯”;([2], p.1)有的学者从分析深度学习的内在机制出发,阐述了人工智能和广义心学的关系;[3]此外还有很多人担心人工智能的风险,甚至担心有朝一日人工智能会统治人类或毁灭人类。英国著名理论物理学家霍金曾说:“人工智能的全方位发展可能招致人类的灭亡。”[4]美国人工智能专家雷·库兹韦尔(Ray Kurzweil)说:“未来出现的智能将继续代表人类文明——人机文明。换句话说,未来的计算机便是人类——即便他们是非生物的。”([5], p.10)中国学者赵汀阳也认为“人工智能可能导致的‘变天’将是无可补救的人类终结,至少也是人类历史的终结”。[6]

笔者认为,不管是人工智能的风险还是人工智能的益处,当我们谈论这些问题时,我们先要搞清楚人工智能的本质。对人工智能取代人类、统治人类的担忧其实隐含着一个人工智能本质的假设,这个隐含的假设就是人工智能是人,且能力全面超过现有人类。那么人工智能的本质到底是不是这样呢?要想搞清楚人工智能的本质,我们先要明白人工智能的一些基本技术原理。基于这样的考虑,本文从历史唯物主义出发,结合人工智能的基本技术原理,试图揭示出人工智能

的本质,进而揭示出其与生产方式的关系,为思考人工智能与人类的关系及其究竟如何影响人类社会提供一种新的视角。

一、人工智能:技术缘起和发展

在谈论人工智能问题时,我们首先要清楚人工智能是什么,以及它的起源和历史概要,只有在明白了这些问题的基础上,我们的讨论才有可能做到准确、合理、恰当,不至于陷入谬误和闹出笑话。何为人工智能?至今为止,不管是人工智能学家还是其他学者都没能给出一个明确的回答,更加没有一个已经取得共识的概念。但是从各个专家学者的笼统界定和说明中,我们还是大概能够明白人工智能是什么。简单来说,人工智能就是利用计算机技术、自动化技术、信息技术等相关科技手段建构的具有一定智能的系统。有时候,我们所说的人工智能也指建构这种智能系统的技术。虽然美国数学家诺伯特·维纳(Norbert Wiener)提出的“控制论”也是人工智能的奠基性理论之一,但是现在公认的对人工智能最早的研究是美国数学家、逻辑学家图灵1950年在《心灵》(*Mind*)杂志上发表的论文“计算机器和智能”。在这篇文章里,图灵提出了一个问题:“机器能够思维吗?”他认为对这个问题的解答,不能从定义“机器”和“思维”两个词的含义入手,因为这样做会带来类似统计学调查的荒唐的结论。于是他提出了一个“模仿游戏”的设想来作为这个问题的新形式:三个人来做一个问答游戏,两个回答者A和B分别是男性和女性,提问者C性别不限,回答者和提问者分别在两个房间里,互

相看不到对方；提问者的目标是确定回答者中哪个是男性，哪个是女性；为了防止提问者从回答者的声调中获取有用信息，回答采取书面打印的形式；A和B回答时要尽量误导干扰C，使C做出错误判断；如果把回答者中的A替换成一台机器，在这种情况下做游戏时，提问者做出错误判断的次数还和以前一样多吗？（[7]，pp.44-45）后来，这个游戏就发展成为著名的“图灵测试”：如果一个机器能够回答人们提出的问题，并且不被人们发现它是一台机器，那么我们就说这台机器是有智能的。图灵测试现已成为一个公认的判定机器是否具有智能的标准。图灵当时所说的机器是指“电子计算机”或“数字计算机”，“我们只允许数字计算机参与我们的游戏”。（[7]，p.47）所以如果从学科属性上说，人工智能起源于计算机学科。

图灵这篇论文发表之后，产生了很大的反响，引起了人们对机器智能问题的热烈讨论。1952年，IBM公司的研究员阿瑟·塞缪尔（Arthur Samuel）开发了首款西洋跳棋程序，^[8]该跳棋程序能够击败一般业余爱好者。1955-1956年间美国学者赫伯特·西蒙（Herbert Simon）、艾伦·纽厄尔（Allen Newell）和约翰·肖（John Shaw）共同开发了一个名为“逻辑理论家”的启发式程序，（[9]，pp.243-244）该程序成功证明了英国数学家、哲学家罗素和其老师怀特海合著的一本数学名著《数学原理》第二章52个定理中的38个，引起了世人的高度关注，并且由于“逻辑理论家”的出色表现，“机器人定理证明”这一新学科由此诞生。（[10]，p.80）1956年夏天，人工智能第一次学术会议在美国达特茅斯学院召开，会上约翰·麦卡锡（John McCarthy）正式提出“人工智能”这一术语，随后人工智能蓬勃发展。

1957年，西蒙、纽厄尔和肖合作开发出人工智能的第一个程序设计语言——IPL（Information Processing Language），60年代他们又成功开发出一个更具普遍意义的解决问题的程序——通用问题求解器（General Problem Solver），（[10]，pp.80-81）西蒙曾发出豪言壮语“在二十年内，一台机器就能够完成人类能做的所有事情”，（[11]，p.23）马文·明斯基（Marvin Minsky）也曾预言人工智能将呈直线式增长。^[12]现在看来，显然他们的愿望落空了，但是这也反映了当时人工智能科学家们对人工智能前景的坚定信心。麦

卡锡、西蒙等人对人工智能的研究进路被后人称为符号主义。符号主义与联结主义一起构成了人工智能领域的两大技术流派。^[13]简单来说，符号主义认为“符号是智能行动的根基”，（[7]，p.115）物理符号系统是一般智能行动的充分且必要条件，“所谓‘必要的’是指，任何表现出一般智能的系统都可以经分析证明是一个物理符号系统。所谓‘充分的’是指，任何足够大的物理符号系统都可以通过进一步的组织而表现出一般智能”，（[7]，pp.119-120）因而人的智能行为也可以解释为物理符号系统的工作过程，所以“人类大脑和恰当编程的数字计算机可被看作同一类装置的两个不同的特例，这一装置通过用形式规则操作符号来生成智能行为”。（[7]，p.331）换句话说，符号主义通过把世界形式化为一套符号系统，预置了形式规则的机器处理这些符号从而产生了智能行为。二十世纪六七十年代符号主义大放异彩，主导了人工智能的研究。但是到了七十年代末，符号主义一直没有解决的常识问题凸显出来，成为其前进路上的拦路虎。因为常识无法用规则来描述和说明，或者说要说明一个常识则需要更多的常识“去理解所发现的事实和规则”。（[7]，p.352）从此，符号主义走向衰落。

联结主义的核心思想是模拟人类大脑的神经网络，它兴起于1957年美国研究人员弗兰克·罗森布拉特（Frank Rosenblatt）发明的感知机。罗森布拉特认为人类的智能行为可能很难被形式化，与其建立一个形式化的结构，倒不如建立一个模拟大脑的装置，让它自己获得能力，“使物理系统公理化，然后用分析的方式研究该系统，以确定出它的行为，比起使行为公理化，然后用逻辑综合技术来设计物理系统，既是比较容易的，也是更为有益的”。（[14]，p.386）感知机是第一个人工神经网络模型，后来那些新的人工神经网络模型都起源于它。这里所谓模拟人类大脑的神经网络，并不是创造出一个跟人类大脑一模一样的机制，因为人类大脑的工作机制我们至今也还远未弄清楚，而是在大脑神经元的启示下发明的一些算法，其实质是计算。联结主义的发展在六十年代经历一个小高潮之后，在七十年代陷入困境，部分是因为当时的计算机不发达，计算能力低下，无法提供人工神经网络模型所需的计算量，^[13]正像明斯基所批评的：“这些机器在非常简单的问题

上通常工作的不错,但是当分配给它们的任务变难时,情况就迅速恶化。”^[15]部分是因为符号主义派的蓄意攻击,“神经网络研究的资金来源也因此被切断”。([16], p.22)联结主义由此沉寂了十多年。进入八十年代,随着1986年反向传播算法的提出,^[17]联结主义再次兴盛。然而好景不长,到了九十年代联结主义由于算法问题再度衰落。直到2006年,深度学习新算法的出现,才使联结主义走出寒冬。所谓深度学习,就是增加了传统人工神经网络的层数,改进了传统算法。深度学习的出现,再加上近年来计算能力的提升、互联网时代数据的积累,使得人工智能跃上了一个新台阶,并迅速繁荣起来,发展到现在如日中天的态势。2016年3月战胜围棋世界冠军李世石的人工智能机器人AlphaGo就是深度学习的产物。

从以上人工智能的发展历程可以看出,人工智能并不是什么新鲜事物,它的发展也经历了诸多曲折。联结主义是当今人工智能的主流范式,符号主义至今还没有走出低谷,以至于现在很多人将符号主义称为“老式人工智能”。通过以上论述,我们明白了人工智能的基础技术理论,理解了人工智能的基本原理。从这些基本技术原理可以看出,人工智能并不神秘,并没有人们想象的那么深不可测,它不过是以符号逻辑或者算法为基础的信息技术和自动化技术的综合,是信息技术和自动化技术发展的结果。人工智能的“智能”是人赋予的,是有范围、有界限的,而且这些范围和界限也是由人来设定的。由此,我们就有了十分坚实的基础去进一步讨论人工智能的本质及其对人类社会的影响。

二、人工智能的本质: 人类本质力量的对象化

在通晓人工智能的基本技术原理的前提下,需要进一步追问:奠定在这样的技术基础之上的人工智能究竟是不是人?人工智能的本质到底是什么呢?历史唯物主义可以给我们提供一个有益的思考视角。

历史唯物主义认为,实践是人的存在方式,是人区别于动物和其他一切生命的根本特征。所谓实践,是主体能动地改造世界的对象性活动,具有自觉能动性、直接现实性、社会历史性等特

征。人之所以为人,除了因为人有理性、会思考之外,最根本的是因为人会实践。实践不仅是人的存在方式,实践还创造了人,因为生产实践即劳动在从猿到人的进化过程中起到了决定性的作用,正如恩格斯所说:“劳动创造了人本身。”([18], p.759)而人工智能不会实践,也不可能实践,它的活动与人的实践活动有本质的区别。

首先,实践是主观和客观的统一,而人工智能的活动缺乏主观性维度。所谓实践是主观和客观的统一,客观方面不必多说,主要是指实践活动的主体、中介、客体都是客观存在的,实践活动是一个客观过程;主观方面主要是指实践活动包含人的主观性因素,因而呈现出一系列独特的特征。具体说来,实践活动的主观方面主要表现为以下几点,而人工智能的活动完全不具备这些特征。第一,人的实践活动具有目的性,而人工智能的活动没有目的性。人在实践活动中是怀揣着目标的,是心中抱有蓝图的,在实践开始之前就已经设想好了实践结束之后“作品”的样子,在实际地改造客观事物之前就已经在思维中把客观事物观念地改造好了,在思维中“消灭了”实践客体的当前形式而形成了实践客体的理想形式,正如马克思所言,“最蹩脚的建筑师从一开始就比最灵巧的蜜蜂高明的地方,是他在用蜂蜡建筑蜂房以前,已经在自己的头脑中把它建成了”。([19], p.208)而人工智能在其活动中是没有目的性的,它或者是从外界获得一个指令,根据其内置的已被人类设计好的形式化的规则去执行这个指令所规定的任务,完成相应的行为动作,比如通过图灵测试的人工智能机器人在回答人类提出的问题,只是通过内置处理器分析接收到的语音信号,根据内置好的规则,从其存储器里储存好的语料库中抽取一些语料输出而已,这是符号主义范式的人工智能;或者是根据人类给它设计的算法,“学习”并做出一个动作或者动作系统,比如说下棋等,这是联结主义范式的人工智能。其实从本质上说,所谓的“学习”归根到底也是在执行算法的指令,虽然不是直接的指令,因为算法从一开始就规定了人工智能“学习”的机制和范围。所以不管是哪种范式的人工智能,它们的活动都没有目的、没有目标,只是在执行指令。第二,人的实践活动具有创造性,而人工智能的活动没有创造性。在实践活动中,人在认识和把握规律的

基础上，可以根据自己的想象力和创造力，把实践客体改造成世界上以往不存在的任何形态，或者“创造出按照自然规律本身无法产生或产生的几率几乎等于零的事物”。（[20], p.72）比如火车、汽车、电灯、电子计算机等事物在自然界中本不存在，是人在实践活动中展开自己的想象力和创造力把它们发明出来，让它们为人类服务。正是因为实践活动的创造性，我们才有了巨大的物质财富和精神财富，才有了高度发达和绚烂多姿的人类文明。反观人工智能，如前所述，人工智能的活动只是根据指令完成特定的任务，它所能做的只是按照既定的形式规则或算法去执行，不能超出形式规则所能处理的范围之外，或者说不能超出算法所规定的范围之外，自然不会有什么想象力和创造力，正如有学者所指出的：“人工智能是计算机科学的一个分支，现在无论其研究方法还是其成果形态都离不开计算，因此，计算是人工智能的本质。”^[21]它不会用超出算法规定范围之外的新办法去做一件事情，更加不能发明一样新东西，人工智能的活动缺乏创造性，它不会自主想象和创新。第三，人的实践活动具有自觉性，而人工智能的活动没有自觉性。人在实践活动中能够意识到外部世界的存在，意识到实践的客体，同时也能够意识到自我的存在，把自我作为实践的主体同实践的客体区别开来。不仅如此，人还能意识到自身的行为和内心，意识到自己正在进行的实践活动，知道并理解自己在做什么。正如马克思所说：“人则使自己的生命活动本身变成自己意志的和自己意识的对象。他具有有意识的生命活动……有意识的生命活动把人同动物的生命活动直接区别开来。”（[22], p.273）这句话同样也适用于人与人工智能。人工智能虽然能够做出一些行为动作，但是它不知道自己在做什么，不能理解自己的行为。美国哲学家约翰·赛尔（John Searle）曾提出过一个著名的“中文屋思想实验”，大意是假设赛尔只懂英文，对中文一窍不通，既不能读也不能写，也不能区分中文、日文和无意义的曲线，他被关在一个房间里，并且给了他两批中文文本和一套用英文书写的规则，规则可以使两批中文文本发生联系。接着，赛尔收到从房间外面递进来的第三批中文字符，即用中文写成的问题，赛尔肯定是不理解这个问题的，同时他还收到一个指令，这个指令是用英文写的，告诉

他怎么根据规则把第三批中文字符同前两批联系起来，并且告诉他怎样送回特定形状的中文符号作为对第三批中文字符的回应，即对这个问题的回答。在房间外的人看来，赛尔对问题的回答与以中文为母语的人的回答毫无二致，但是赛尔却一个汉字都不认识，更别说理解问题和答案了。（[7], pp.75-76）在这里，赛尔就相当于一个人工智能系统，“中文屋思想实验”非常形象地说明，人工智能虽然可以从事一些活动，但是它根本不能理解自己正在进行的的活动，比如说AlphaGo虽然战胜了围棋世界冠军李世石，但是它并不知道自己在下棋。人工智能的活动没有自觉性，只能算是机械运动。

其次，实践是一个双向对象化的过程，是主体客体化和客体主体化的有机统一，而人工智能的活动不但不具有主体客体化特征，也不具有客体主体化特征。所谓双向对象化，是指“作为互为对象的实践主体和实践客体相互渗透、相互创造的过程”。（[20], p.135）双向对象化是人类实践活动的本质特征，它包括相互联系、不可分割的两个过程，一个是主体对象化过程，也叫做主体客体化过程；一个是主体被对象化过程，也叫做客体主体化过程。一方面，人工智能的活动不具备主体客体化特征。主体客体化是指作为实践主体的人在改造客体的过程中，将自身的理想、知识、能力等本质力量对象化为现实的存在，对象化为经过改造后的新的客体的一部分，在自然界中打下人的烙印，正如马克思所指出的：“劳动的产品是固定在某个对象中的、物化的劳动，这就是劳动的对象化。劳动的现实化就是劳动的对象化。”（[22], pp.267-268）在主体客体化的过程中，人是基于自己的内在尺度也就是意识到的自己的需求，去对象化自己的本质力量，并不是任意的对象化，从而创造出一个适合人生存和发展的属人世界。而人工智能不具备关于客观事物的知识，也没有所谓的理想、信念，不知道真善美为何物，更没有对真善美的追求，所以也就不存在可以被对象化的“本质力量”。而且人工智能没有自己的需求，从最基本的吃、喝、住、穿需求到高级的审美需求和其他精神需求都没有，因而它的活动也就不会有内在尺度，在这方面它还不如动物，动物虽然没有内在尺度但是最起码还有吃、喝等本能需要，而人工智能连本能需要

都没有。所以说,人工智能的活动没有主体客体化这个过程,也根本不具备主体客体化的基础。另一方面,人工智能的活动更不具备客体主体化的特征。客体主体化是指在实践活动中,实践客体同时也规定和作用着实践主体,从而客体进入主体中转化为一种主体性的存在。这里的“进入”并不是指实物性的进入和融为一体,而是指实践活动中主体在客体的影响和作用下,会能动地反映、选择、建构主体的特性和特征,逐步认识到客体的内在矛盾和客观规律,从而丰富和发展了主体自身。也就是在这个意义上,我们说客体的属性进入了主体,变成了主体的认识,主体被对象化了。而人工智能在其活动中不能将事物的特性、特征反映到自身中来,不能透过表象认识事物的本质,不能在变化的现象中把握住其中的规律,客体的这些属性无法“进入”人工智能之中,人工智能没有所谓认识的提高。人工智能的“深度学习”只是按照特定的算法在特定的场景下“学习”特定的处理方式,超出特定的场景就要给它添加新的算法,否则它就什么都“学习”不了、什么都不会做,它的“学习”并不是真正的学习和认识,“神经网络并没有学到任何东西”。^[23]因而人工智能的活动不可能做到客体主体化。

从以上可以看出,人工智能的活动与人的实践活动有着本质的区别,人工智能不会实践,而实践又是人的存在方式,是人的本质力量的根源,人类之所以能有今天高度发达的改造世界的能力,就是因为实践,人类未来能力的发展,也要依靠实践。所以,人工智能不是人。一言以蔽之,人工智能的本质是人类本质力量的对象化。人工智能是在人类实践活动中产生的,是人类实践活动的成果,是人类本质力量的实现和确证。具体来说,人工智能就是“把主体的认识能力转化为认识客体,然后对之进行认识与模拟”。^[24]它是人类认识世界和改造世界的产物,同时也是人类认识世界和改造世界的手段,它“通过工业日益在实践上进入人的生活,改造人的生活”。([25], p.193)如同历史上机器的出现一样,人工智能的出现是人类科学技术发展的必然结果,它是升级版的机器,是机器2.0。机器代替的主要是人类的体力劳动,人工智能除了可以代替体力劳动还可以代替人类的某些脑力劳动。显而易见,作为机器2.0的人工智能,不可能全面超越人类和统治人类。

三、人工智能造福人类还是带来风险: 取决于生产方式

我们已经清楚,人工智能的本质是人类本质力量的对象化,人工智能本身不可能超越人类和统治人类,不会有终结人类历史和毁灭人类的风险。那么人工智能的应用将会造福人类还是给人类带来灾难呢?乐观主义者认为,人工智能的应用将会给人们的生产和生活带来巨大的方便和好处,比如人工智能“可以代替人从事一些重复、单调、繁重的工作,将人从一些有毒、有害、危险的工作(环境)中解放出来”;^[26]人工智能在医疗领域的应用可以在短时间内从浩如烟海的医学文献中找出跟患者情况相匹配的资料,大大提高了诊断效率;语音识别、图像识别等技术的发展给公共安全、航空航天、翻译等行业领域带来了革命性的进步,百度的一个英文语音识别系统在一项测试中,单词错误率只有3.1%,已经超出了正常人的平均水平5%。^[27]而悲观主义者则认为,人工智能的应用将给人类社会带来诸多麻烦和挑战,比如人工智能可以代替一些简单、重复、规则性强的脑力劳动,从而一大批低技能水平的工人面临失业,2017年麦肯锡研究报告指出:至少1/3的职业有可能被技术替代,相当数量的工作者面临重新择业的挑战;^[28]由于人工智能有一定程度的“自主性”,其活动的后果和责任如何界定和承担,责任问题成为最突出的伦理难题;^[29]一小部分人可能会利用人工智能技术改造自身,把自己“升级”为超人,形成一个社会精英阶层,控制大多数普通人。^[30]乐观主义者的自信和悲观主义者的担忧,都有一定的道理,但是他们都只是在就事论事,没有看到事物背后的深层机制,他们只看到了一些现在已经发生或者将来有可能发生的由人工智能的应用所引起的具体的、单个的事件,看到了这些具体的、单个的事件所带来的益处或危害,而没有看到这些事件背后的原因。这个背后的原因就是生产方式。

正如上文所提到的,历史唯物主义认为实践是人的存在方式,而人类最基础、最重要的实践是物质生产实践,“人们为了能够‘创造历史’,必须能够生活。但是为了生活,首先就需要吃喝住穿以及其他一些东西。因此第一个历史活动就是生产满足这些需要的资料”,([25], p.531)物质生产实践

构成了人类社会存在和发展的基础，在此基础之上形成了人们之间的社会关系和各个社会生活领域。物质生产实践的方式不同，人们所结成的社会关系和政治关系就不同，生产方式决定了一个社会主要的政治法律制度、思想文化和价值观念，正如马克思所说：“物质生活的生产方式制约着整个社会生活、政治生活和精神生活的过程。”（[31]，p.591）总的来说，特定的生产方式决定了特定的社会制度，封建生产方式产生了封建社会，资本主义生产方式产生了资本主义社会。人工智能作为人类实践活动的成果，作为人类的生产资料和生活资料，它到底应用在什么方面以及如何应用，首先取决于它归谁所有，是归共同体所有还是归私人所有，在不同的所有制下，它的用途迥然不同。而所有制恰恰是生产方式的十分重要的组成部分。其次，人工智能的应用还取决于社会的主流思想文化和价值观念，在不同的思想文化和价值观念下，人们所重视和偏好的方面不一样，人工智能的用途自然也就不一样。而思想文化和价值观念是由生产方式决定的。所以，人工智能的应用归根结底是由生产方式决定的。生产方式从根本上决定了人工智能的应用是利大于弊还是弊大于利，是全方位服务人类社会还是给人类社会带来巨大的不可控的风险。

在资本主义生产方式下，资本居于统治地位，社会生产主要是为了追求利润即追求剩余价值，为了实现资本增殖。从而资产阶级和无产阶级成为相对立的两大阶级，凡事追求利润、追求货币财富成为资本主义社会的主流价值观。显而易见，在资本主义生产方式下，人工智能的应用虽然也会给人类社会带来一些便利和好处，但是潜藏着巨大的社会风险和挑战。首先，人工智能的资本主义应用会使人工智能成为资本家的生产资料，成为资本，成为剥削工人和奴役工人的力量，并且会导致大量工人失业。“由于劳动资料转化为自动机，它就在劳动过程本身中作为资本，作为支配和吮吸活劳动力的死劳动而同工人相对立。”（[19]，p.487）马克思在《资本论》中论述机器的资本主义应用时曾提到，机器的资本主义应用把妇女和儿童扔进雇佣工人大军从而扩大了剥削的人身材料，增强了工人的劳动强度，延长了工人的劳动时间，工人的状况没有变得更好，反而变得更糟（[19]，pp.453-480）人工智能作为机器2.0，它的资本主义应用自然也会带来这些负面效应，“分工与专业化并没有被消灭，

流水线的生产方式也没有改变”，^[32]只不过现在的资本主义比马克思当时的资本主义发展了、进步了，所以负效应的表现跟那时候不一样，有了一些变化，比如儿童已经受到了立法保护，不再允许使用童工，从而对儿童的剥削效应不再明显。不仅如此，人工智能的资本主义应用所导致的工人失业比普通机器的资本主义应用所导致的工人失业范围更广、人数更多。普通机器所代替的只是人的简单的体力劳动，而人工智能不仅可以代替体力劳动，而且可以代替某些特定类型的脑力劳动，因而相对于普通机器而言，人工智能将代替更多行业的更多工人，那些被代替的大量工人在资本主义社会只能成为“劳动后备军”，成为“过剩人口”，被抛向街头。大家所担心的失业问题正是人工智能的资本主义应用所带来的，而不是人工智能本身所造成的，因为“资本主义生产的整个体系，是建立在工人把自己的劳动力当作商品出卖的基础上的”，（[19]，p.495）工人除了自身的劳动能力之外一无所有。如果不是在资本主义生产方式下，被代替的工人则不会被抛向街头。其次，人工智能的资本主义应用会加大贫富差距，加深阶级对抗。在资本主义社会，被人工智能代替的大量工人一无所有，他们靠少得可怜的救济金度日，生活状况糟糕；另一方面，占社会人口一小部分的资本家享受着由人工智能的应用所带来的劳动生产率大幅度提高的好处，占有着大量的社会财富和社会资源。社会的贫富差距由此拉大，阶级矛盾更加尖锐，阶级对抗更加明显和激烈，社会更加动荡不安。与此同时，掌握社会财富的少数资本家会将人工智能和基因工程、生物医学工程等技术相结合，把自己改造成“赛博格”（Cyborg）以获得常人所不具备的能力，改造自己的基因以使后代获得超级智力，如此等等。到时候，少数精英控制社会大部分“无用阶级”的情况恐怕就会变成现实。最后，人工智能的资本主义应用会给人类社会带来隐私泄露和智能战争等安全威胁。在资本主义社会，一切都被商品化，一切行为都以利润为目的，个人隐私自然也免不了被商品化的命运，被有些人出卖以赚取利润，当今世界的事实已经说明了这一点。而当某些人利用人工智能来进行网络攻击、非法窃取个人隐私信息时，那就更加方便、更加容易了，人们也更加难以防备。各资本主义国家发展的不平衡，以及在各自的资本积累过程中产生的相互之间的利益矛盾，从资本主义产生那天起就一直

存在着,只要有资本主义制度存在的一天,战争的威胁就不可能真正消除。当人工智能被用于制造智能武器、跟核武器结合等等,从而用于战争,其后果不堪设想,人类社会就会面临被自己毁灭的威胁,霍金所担心的人工智能的发展导致人类灭亡的原因正在于此。除了这些之外,人工智能的资本主义应用还会带来一系列的风险,在此不再一一赘述。

相反,在社会主义生产方式下,生产资料不再归少数人所有,社会生产不再以个人利润为主要目标,而是以整个社会的利益和幸福为主要目标,因而人工智能的社会主义应用从根本上消除了其资本主义应用所带来的各种巨大风险和挑战。在社会主义社会,人工智能不再是压迫和奴役工人的力量,工人“不再被当做活的附属物并入死机构”。([19], p.486)人工智能代替了大量工人,但是非但不把他们抛向街头使他们“堕落丧亡”,([19], p.508)反而给他们“创造了大量可以自由支配的时间”。([33], p.103)必要劳动时间的减少,自由时间的增多,人们就可以投入更多的时间来发展自己的个性和能力,发展自己多方面的兴趣和爱好,为实现“上午打猎,下午捕鱼,傍晚从事畜牧,晚饭后从事批判”([25], p.537)的共产主义生活方式和人的自由全面发展打下基础。人工智能的应用带来了劳动生产率的提高,从而创造了大量社会财富,这些财富不再归少数人所有,而是归整个社会享用,为人类解放准备了物质基础。

当然,并不是说在社会主义生产方式下人工智能的应用就能避免一切风险。在社会主义制度下如果人工智能应用不合理,也会带来一些风险,比如当个人之间发生利益冲突时,人工智能可能会被用来相互攻击,从而引发犯罪。但是这些风险比资本主义生产方式下的风险小得多,是可控的,是比较容易消除的,不会威胁人类的生存和发展。人工智能究竟是造福人类还是带来风险,有多种因素的影响,但是生产方式是最为根本的因素。

结 语

综合以上论述,我们可以看出,人工智能的本质不外是人的本质力量对象化的产物,是人类实践活动的成果,是人类社会科学技术发展的必

然结果,是升级版的机器——机器2.0。不管是哪种范式的人工智能,它都不会实践,其活动与人的活动有着本质的区别,“机器是人”的假设不成立,因而人工智能本身不可能超越人类和统治人类,更不可能毁灭人类。作为智能工具,人工智能的益处和风险都来自于它的应用,而它的应用则归根结底取决于其所处的社会生产方式。在资本主义生产方式下,人工智能的应用虽然也会给社会带来一些方便和益处,但是更多的是会带来一系列的风险,给社会造成危害,利远小于弊。在社会主义生产方式下,人工智能的应用不再会有其资本主义应用所产生的那些巨大风险,而是会充分发挥它的积极作用,为整个社会而不仅仅是少数资本家创造出更多的财富,为全体人民而不仅仅是资产阶级创造出更多的自由时间,为人类解放奠定基础。现在已经显现的人工智能的风险,以及大家担心日后可能会出现的人工智能的主要风险,从根本上源于它的资本主义应用,正如马克思在论述机器时所说:“这些矛盾和对抗不是从机器本身产生的,而是从机器的资本主义应用产生的。”([19], p.508)因而消除这些重大风险的根本途径,在于改变资本主义生产方式,改变资本主义制度,变资本主义社会为社会主义社会。

人工智能问题并不是人工智能本身的问题,而是生产方式的问题,我们须透过层层迷雾把握人工智能问题的实质。我国作为社会主义国家,在规避人工智能的应用风险方面有着无可比拟的制度优势,从根本上化解了人工智能应用的重大风险,我们应抓住机遇发展人工智能,让人工智能服务于全体人民群众,服务于整个社会,引领人工智能的发展和应用方向,为人类社会的发展做出贡献。

[参 考 文 献]

- [1] 杰瑞·卡普兰. 人工智能时代[M]. 李盼译, 杭州: 浙江人民出版社, 2017.
- [2] 徐英瑾. 心智、语言和机器——维特根斯坦哲学和人工智能科学的对话[M]. 北京: 人民出版社, 2013.
- [3] 张祥龙. 人工智能与广义心学——深度学习和本心的时间含义刍议[J]. 哲学动态, 2018, (4): 13-22.
- [4] 斯蒂芬·霍金. 让人工智能造福人类及其赖以生存的家園[EB/OL]. <https://tech.qq.com/a/20170427/08205.htm>. 2017-04-27.

- [5] 雷·库兹韦尔. 奇点临近 [M]. 李庆诚译, 北京: 机械工业出版社, 2017.
- [6] 赵汀阳. 人工智能“革命”的“近忧”和“远虑”——一种伦理学和存在论的分析 [J]. 哲学动态, 2018, (4): 5-12.
- [7] 艾伦·图灵. 计算机与智能 [A], 玛格丽特·博登. 人工智能哲学 [C], 刘西瑞、王汉琦译, 上海: 上海世纪出版集团, 2006.
- [8] Samuel, A. 'Some Studies in Machine Learning Using the Game of Checkers' [J]. *IBM Journal of Research and Development*, 2000, 44(1/2): 206-226.
- [9] 赫伯特·西蒙. 我生活的种种模式 [M]. 曹南燕、秦裕林译, 上海: 东方出版中心, 1998.
- [10] 吴鹤龄、崔林. ACM图灵奖: 计算机发展史的缩影 (1966-2006) [M]. 第3版, 北京: 高等教育出版社, 2008.
- [11] Simon, H. *The Sciences of the Artificial* [M]. 3rd Edition, Cambridge, Mass: MIT Press, 1996.
- [12] Minsky, M. 'Steps Toward Artificial Intelligence' [J]. *Proceedings of the IRE*, 1961, 49(1): 8-30.
- [13] 顾险峰. 人工智能的历史回顾和发展现状 [J]. 自然杂志, 2016, 38 (3): 157-166.
- [14] Rosenblatt, F. 'Strategic Approaches to the Study of Brain Models' [A], Von Foerster, H. (Ed.) *Principles of Self-Organization* [C], Elmsford, NY: Pergamon Press, 1962.
- [15] Minsky, M., Papert, S. *Perceptrons: An Introduction to Computational Geometry* [M]. Cambridge: The MIT Press, 1969.
- [16] 玛格丽特·博登. AI: 人工智能的本质与未来 [M]. 孙诗惠译, 北京: 中国人民大学出版社, 2017.
- [17] Rumelhart, E., Hinton, E., Williams, J. 'Learning Representations by Back Propagating Errors' [J]. *Nature*, 1986, 323(6088): 533-536.
- [18] 马克思恩格斯全集 [M]. 第26卷, 北京: 人民出版社, 2014.
- [19] 马克思恩格斯文集 [M]. 第5卷, 北京: 人民出版社, 2009.
- [20] 萧前、杨耕. 唯物主义的现代形态——实践唯物主义研究 [M]. 北京: 中国人民大学出版社, 2012.
- [21] 陈钟. 从人工智能本质看未来的发展 [J]. 探索与争鸣, 2017, 33 (10): 4-7.
- [22] 马克思恩格斯全集 [M]. 第3卷, 北京: 人民出版社, 2002.
- [23] 皮埃罗·斯加鲁菲. 离人工智能奇点还有多远 [J]. 中国经济报告, 2017, (5): 112-113.
- [24] 陈步. 人工智能问题的哲学探讨 [J]. 哲学动态, 1978, (11): 22-31.
- [25] 马克思恩格斯文集 [M]. 第1卷, 北京: 人民出版社, 2009.
- [26] 孙伟平. 关于人工智能的价值反思 [J]. 哲学研究, 2017, (10): 120-126.
- [27] Amodei, D., Anubhai, R., Battenberg, E. 'Deep Speech 2: End-to-End Speech Recognition in English and Mandarin' [EB/OL]. <https://arxiv.org/abs/1512.02595>. 2015-12-18.
- [28] McKinsey Global Institute. *Jobs Lost, Jobs Gained: Workforce Transitions in a Time of Automation* [M]. New York: McKinsey & Company, 2017.
- [29] Verbeek, P. *Moralizing Technology: Understanding and Designing the Morality of Things* [M]. Chicago: The University of Chicago Press, 2011.
- [30] 尤瓦尔·赫拉利. 未来简史 [M]. 林俊宏译, 北京: 中信出版集团, 2017.
- [31] 马克思恩格斯文集 [M]. 第2卷, 北京: 人民出版社, 2009.
- [32] 吴胜锋. 马克思主义意识理论视域下的人工智能技术批判 [J]. 哲学动态, 2019, (4): 26-33.
- [33] 马克思恩格斯全集 [M]. 第31卷, 北京: 人民出版社, 1998.

[责任编辑 李斌]