

人工智能会生成意识形态吗?

刘伟兵

摘要: 人工智能不具有意识,但是却表现出“意识”现象、“意识”行为。人工智能在数据、算法、模型的技术链条上,“无人化”生成“意识”现象的背后,是总体工人的社会协作化的精神生产。自动生成的“意识”现象是社会大生产的精神产品,是意识形态。人工智能生成的“意识”现象,在一般应用、技术的普遍化应用、意识形态斗争利用之后,将会在资本逻辑的统治、形式逻辑和数理逻辑的同化、原有意识形态逻辑渗透的推动下向意识形态转型。人工智能作为新的技术变革,会改变意识形态与生产方式的关系,出现新的意识形态,并增强意识形态的隐蔽性、虚假性和异化的特性。

关键词: 生成式人工智能; 意识; 意识形态; 意识形态; 精神生产

DOI: 10.19836/j.cnki.37-1100/c.2025.03.014

2025年年初,通过算法优化和工程创新,DeepSeek具有了低成本、高性能和开源的特点,并具有图像、文字、音频、视频等多模态支持能力和强化学习性能,从而迅速火爆。如今,DeepSeek更是开始接入各地的政务系统,并被广泛应用于教育、医疗等领域,具有向通用人工智能(AGI)发展的趋势。人工智能可以写作、聊天,也可以画画、编曲,这种“无中生有”的“意识”行为,使得人们不得不面临一个新的问题,即人工智能会生成意识形态吗?如果说人工智能不会生成意识形态,那么人工智能生成的文本,以及人工智能的算法推荐为何又存在着明显的价值倾向和意识形态偏好呢?如果说人工智能会创造意识形态,那么一个“中立”的科学技术为何会呈现为意识形态性呢?意识形态是人工智能自主生成的,还是外界赋予人工智能的呢?目前,学界围绕人工智能与意识形态进行了大量研究,主要聚焦于人工智能时代“意识形态安全”^①“意识形态风险防范”^②“意识形态工作”^③等议题。这些研究事实上都默认或者预设了一个前提,就是人工智能必然会生成意识形态。而人工智能是否会生成意识形态是人工智能意识形态研究的基础元问题,并不是一个自然而然不需要研究回答的命题。研究人工智能能否生成意识形态,首先要回答一个理论前提,即人工智能是否具有意识?事实上,这一研究进路也是历史唯物主义视域下意识形态研究的理论进路。

关于人工智能意识研究,不同学科的学者观点争议较大,甚至同一学科内部许多学者的观点也有不同。目前,学界关于人工智能意识的研究进路分为两类:一类认为意识来源于身体,一类认为意识来源于心灵。从脑科学、神经生物学的观点出发,意识是人的身体的神经结构、神经化学以及神经生理基础物质等活动的产物。“根据目前的科学定论,我所体验到的一切都是脑电活动的结果。”^④从这一意识观点出发,有的学者认为人工智能可以模仿人类神经细胞、突触等设计神经网络模型从而具有“功能意识”^⑤,或者是现象意识。“通过人工智能技术、脑科学、哲学等学科综合探讨现象意识合成的

基金项目: 国家社科基金青年项目“人工智能时代意识形态风险防范机制创新研究”(23CKS053)。

作者简介: 刘伟兵,复旦大学马克思主义学院副教授(上海 200433; wbliu595@163.com)。

① 李钢、黎祯祯:《智能革命的意识形态安全与对策刍议》,《理论视野》2024年第7期。

② 代金平、覃杨杨:《ChatGPT等生成式人工智能的意识形态风险及其应对》,《重庆大学学报(社会科学版)》2023年第5期。

③ 张志安:《人工智能对新闻舆论及意识形态工作的影响》,《人民论坛·学术前沿》2018年第8期。

④ 尤瓦尔·赫拉利:《未来简史》,林俊宏译,北京:中信出版社,2017年,第108页。

⑤ 玛格丽特·博登:《AI:人工智能的本质与未来》,孙诗惠译,北京:中国人民大学出版社,2017年,第145页。

概念谱系与哲学内涵,解释并映射人类产生高阶思维的内在逻辑与机制。”^①在他们看来,“图灵认为如果机器正在做的事情与‘真实的人类所做的事情’无法区分,那么我们就应该停止争论机器究竟是否存在‘真正’的意识。如果它能够通过我们所发明的任何合理的测试,让我们无法区分,这可能就是我们想要的结果了”^②。但是也有学者认为人工智能不具有神经结构,就不可能生成意识。“希尔勒倡导生物自然主义(biological naturalism),根据该理论,精神状态是由神经元中低级物理过程引起的高级涌现特征。”^③从意识心灵来源说出发把握人工智能意识,又可以进一步细分为人工智能是否具有意向性、自由意志、自我意识,乃至伦理规范、德性等范畴。在这一进路上,学者们普遍认为人工智能具有功能意识、现象意识。

目前,人工智能意识是一种现象意识和功能意识,是从个体的意识解读上升到了社会的意识形态研究,关系到历史唯物主义中意识形态理论的发展。因为人工智能的“意识”现象、“意识”行为等问题的研究是人工智能时代精神生产方式变革研究,反映的是精神生产协作化变革出现的新的意识生成方式,并不意味着人工智能像人一样具有意识。事实上,人工智能的“意识”问题研究,要上升到社会协作背后的社会意识结构层面,从意识形态上层建筑中去寻找解答。这就是说人工智能的“意识”问题要转化成人工智能的意识形态问题研究,本质上是对当时社会结构的意识形态探究。这是因为,人工智能意识形态问题研究要深入人工智能的应用逻辑,把握人工智能技术特征在不同应用逻辑下反映不同阶级利益的意识形态建构。这既是历史唯物主义中意识形态理论的生长点,也是当前社会关注的焦点。

一、人工智能与“意识”现象的自动生成

在唯物史观中,马克思是从人的分工、需要、生产等理论出发,把握意识的生成路径,明确了意识是属人的意识,是由人的物质实践活动所形成的。“意识一开始就是社会的产物,而且只要人们存在着,它就仍然是这种产物。”^④正是由于人的主体性明确,所以对意识的把握不是从抽象的理论出发,而是应该深入人的物质实践活动之中去寻找根源。马克思也正是在此意义上,认为意识是人与动物区别的标志之一,是人之所以是“类”的类本质所在。“有意识的生命活动把人同动物的生命活动直接区别开来。正是由于这一点,人才是类存在物。”^⑤因为动物只能按照自己的尺度去构造,但是人们却能够用各种尺度去实践。正是如此,人的意识在生命的维度上是一种美的规律的体现。

但是,当前人工智能的发展,正逐渐显露出“意识”现象和“意识”行为。以DeepSeek为代表的最新人工智能技术,已经具有了人的智能表现形式,可以画精美的画,与人正常交流,能够写代码,编写文本等。那么,人工智能的“意识”是具有美的规律构造的人的类本质的意识吗?对这一问题的回答,要先回应一个前置性问题,即人工智能研究领域的意识是何种意义的意识。目前学界的人工智能意识研究表现为:感知意识研究,即人工智能通过感知模型和传感器对外界的刺激产生回应行为的研究;认知意识研究,即构建人工智能的认知模型,使人工智能具有类似意识的认知特征及意识行为表现;机制意识研究,即研究人类意识的产生机制,并据此进行人工智能模拟实验研究;自我意识研究,即研究如何使人工智能意识到自我,具有自我认知、自我反省的能力,认识到自我与他者的区别;感受意识研究,即如何使人工智能具有情感的感受、听觉触觉等感官的感受等;意识测试研究,即检测人工

① 刘鸿宇:《人工智能现象意识合成何以可能》,《科学技术哲学研究》2023年第5期。

② 迈克尔·伍尔德里奇:《人工智能全传》,许舒译,杭州:浙江科学技术出版社,2021年,第312页。

③ 斯图尔特·罗素、彼得·诺维格:《人工智能:现代方法》,张博雅、陈坤、田超等译,北京:人民邮电出版社,2022年,第764页。

④ 《马克思恩格斯文集》第1卷,北京:人民出版社,2009年,第533页。

⑤ 《马克思恩格斯文集》第1卷,第162页。

智能是否具有意识、水平如何等,其代表就是图灵测试。从目前的技术水平而言,感知意识、认知意识、机制意识等意识行为、意识现象较容易实现,也是当下人工智能意识研究的主要内容^①。

当下,人工智能研究领域的意识产生解释理论主要有五种:一是全局工作空间理论(GWT),认为意识产生于大脑的全局工作空间,所以人工智能可以模仿大脑全局工作空间,进行多处理器的协同工作,不过这一理论缺少神经层面的解释,感受性意识也无法实现。二是整合信息理论(IIT),认为意识产生于大脑对信息的整合,是当下意识建模的重要方法之一,但是大脑中庞大的信息量在现有计算水平下难以计算。三是高阶表征理论(HOR),即认为意识是一阶心理状态的知觉,是大脑在不断无意识地描述自身运动的学习过程中出现的。四是注意图示理论(AST),即将自我与外部对象以注意的方式联系起来,以意识体验的方式把握意识。五是量子意识理论(QC),即意识产生于大脑中的量子计算,但是这一理论缺少实验支撑,甚至有人认为大脑中的计算应该是经典计算而不是量子计算。

在以上五种意识产生解释理论基础上,人工智能对机器意识实现方法表现为:一是符号计算,即运用数理逻辑、计算推理将意识问题转化为计算问题。二是人工神经网络方法(ANN),即模拟神经元活动机制建模的机器学习方法。三是生物神经网络方法,用生物技术将生物神经元搭建成生物神经网络来生成意识。四是量子计算方法,即根据量子的纠缠性、叠加性等特性来解释意识并建模。五是脑机融合方法,运用脑机接口技术(BMI)来研究意识,并形成脑机融合的意识机器。其中,第一种和第二种是传统人工智能方法,是当下人工智能水平下最为常见也是应用最为广泛的方法,但是这两种方法只能产生意识行为和意识形象,并不能真正产生意识,而后三种方法具有一定产生意识的可能,但是限于技术水平和科技伦理,目前并没有取得过多进展^②。因此,从传统人工智能方法出发,我们可以进一步把握现有人工智能技术图景,去深入了解人工智能是如何生成“意识”现象和“意识”行为。

当前的人工智能技术图景主要包括机器学习、大数据、计算机视觉、语音识别、自然语言处理五个部分。这五个部分的关系是通过机器学习的智能算法,在对数据的采集和处理过程中,以数据模型的方式驱动计算机视觉、自然语言处理和语音识别,来实现多模态数据融合的现实交互。人工智能包括了数据、算法和模型三个组成部分。所以,人工智能的科技路径,就是通过数据训练与算法设计,生成一个模型,并在模型的求解过程中输出结果,也就是生成各种“意识”。

其中,数据是人工智能技术最为基本的单元。当前的人工智能技术其实就是数字智能。数据事实上就是以数字形式承载的信息,是客观事物的数字表现形式。数据构成了人工智能的底座,是人工智能时代的存在,即“有什么”。人工智能正是先拥有了“有什么”,才能通过对数据的归纳、分析、演绎,来建构起“是什么”“为什么”“怎么办”。

数据具有多种形态,包括图像、文本、视频、音乐等。但是,数据在原始数据阶段并不能拿来直接使用,需要经过数据标注的方式对原始数据进行分类、标记和注释,形成结构化或者半结构化数据,从而才能够被机器识别。换言之,数据标注过程恰好是人工赋予智能的第一步。人工智能之所以会产生识别的“意识”,是因为数据标注员在数据投喂阶段,就进行了“意识”的赋予。

在数据标注之后,人工智能便需要通过算法来生成“意识”。算法就是人工智能对数据处理的规则、条件和方法。这是因为数据承载的信息并不是一个物质实体,而是一种“感受性关系”^③。这种感受性关系,事实上建构了数据的符号结构,也就是能指和所指的统一体。算法,就是通过代码设计一个特定程序和步骤,来实现数据的所指与能指的统一。算法通过信息的输入,并对其进行分类、回归、聚类处理,最后输出预期结果,实现人工智能的智能特征。这是人工智能具有自主性创造“意识”现象

① 秦瑞琳、周昌乐、晁飞:《机器意识研究综述》,《自动化学报》2021年第1期。

② 秦瑞琳、周昌乐、晁飞:《机器意识研究综述》,《自动化学报》2021年第1期。

③ 王天恩:《信息及其基本特性的当代开显》,《中国社会科学》2022年第1期。

的关键。这说明了人工智能的自主性,抑或是类人的独立“意识”现象,依赖的正是算法。

算法与模型紧密相连,因为模型就是数据在算法训练后的机器学习结果,是一个特定的数据结构,由模型数据和预测算法组成。如果算法是人工智能自主思考何以可能的过程,那么模型便是人工智能自主思考的结果,是人工智能自主生成“意识”的数据结构。

可见,人工智能自主生成“意识”现象的实质是社会化大生产中的精神生产。马克思在《德意志意识形态》中探讨了意识的分工,认为生活资料的剩余会促使一部分人脱离物质劳动,专门从事“精神劳动”^①。在当时的生产方式中,马克思主要考察和把握的精神劳动还是依靠个别的意识形态工作者进行的个体生产。在这一时期,人的主体性会异常明显。但是随着智能化生产方式变革,精神劳动也已经进入了意识的社会大生产阶段。人们的精神生产,将会越来越呈现为协作的方式,个体的主体性将会向总体工人的主体性过渡。“许多人在同一生产过程中,或在不同的但互相联系的生产过程中,有计划地一起协同劳动,这种劳动形式叫做协作。”^②这也正是对历史唯物主义中意识形态理论协作生产的重要发展。

马克思在《资本论》及其手稿中,曾深入分析机器大生产带来的物质劳动的社会化大生产,分析了协作、总体工人等现象。而随着人工智能的出现和普及,智能化生产方式的变革,也将社会化大生产从人的物质劳动延伸到人的精神劳动之中。换言之,人工智能的自主生成“意识”现象,事实上是精神生产出现了协作现象,原来直接劳动形式的精神生产逐渐被间接劳动所取代,个体的主体性被总体工人的总体性逐渐取代。马克思在《1857—1858年经济学手稿》中就曾预测般地提出了“一般智力”(general intellect)概念。“固定资本的发展表明,一般社会知识,已经在多么大的程度上变成了直接的生产力,从而社会生活过程的条件本身在多么大的程度上受到一般智力的控制并按照这种智力得到改造。”^③这事实上预测了精神生产社会协作的发展趋势,把握住了以脑力劳动为代表的精神劳动的日益重要性。

随着人工智能的发展,意识的基本要素被数字化,以数据生产的方式被纳入了社会大生产之中。以往个体的精神劳动需要自己进行意识对象的输入和学习,而智能时代,将存在着专门从事数据标注和训练的局部工人,以专门化的数据生产为意识的社会化生产提供最基本的意识对象和素材。同样,原本依靠部分精神劳动者进行分类、回归、聚类的抽象活动,也由算法工程师专门通过设计监督学习、无监督学习、半监督学习、强化学习等方式代替。于是,人工智能生成的“意识”现象,便作为一个精神生产合力的产品出现,其运行背后是精神生产的社会化分工与合力。

所以,人工智能自主生成“意识”的表象背后,其实是意识在智能生产方式下发展到了协作的社会化大生产阶段。人工智能“无人化”的背后是社会总体工人的合力。“产品从个体生产者的直接产品转化为社会产品,转化为总体工人即结合劳动人员的共同产品。”^④人们原来直接的精神生产在智能化生产方式中逐渐被间接劳动所取代,并且“只要成为总体工人的一个器官,完成他所属的某一种职能就够了”^⑤。人工智能的这种“意识”现象生成的自主性,在此意义上只不过是人的意识的延伸。人工智能并没有形成独立的意识,其“意识”表象只是数字的统计特征。可以说,人工智能生成的“意识”,是一种数理统计上具有“确定性”的“意识”现象。这种“确定性”的“意识”现象是一种必然,并不具有人的意识的“不确定性”,也就没有了偶然性意义上的意识自由,从而就不可能存在顿悟、直觉、灵感等创造性意识。

① 《马克思恩格斯文集》第1卷,第551页。

② 《马克思恩格斯文集》第5卷,北京:人民出版社,2009年,第378页。

③ 《马克思恩格斯全集》第31卷,北京,人民出版社,1998年,第102页。

④ 《马克思恩格斯文集》第5卷,第582页。

⑤ 《马克思恩格斯文集》第5卷,第582页。

可见,人工智能自主生成“意识”现象,是其技术图景带来的生产生活方式智能化变革,进而在科技路径上形成了新的意识生产方式。

二、人工智能应用与意识形态的生成

人工智能不具有意识,但是在精神生产的协作下,获得了“总体工人”所赋予的“意识”现象、“意识”行为。换言之,个体机器的“意识”现象和“意识”行为,无非是社会协作化生产在精神生产领域的体现,反映的恰恰是社会意识。而这也恰好证明了唯物史观。马克思在唯物史观中认为,意识等同于社会意识,不存在超脱于社会意识之外的意识。“意识一开始就是社会的产物,而且只要人们存在着,它就仍然是这种产物。”^①这是因为,人的本质是一切社会关系的总和,其个体意识是社会意识个体化的产物,是社会结构对个体意识的再建构。但是,社会意识是“一个描述性的概念”^②,其在阶级社会中,就表现为反映特定阶级利益的意识形态。人工智能生成的“意识”现象、“意识”行为其实是意识形态上层建筑在机器个体层面的实践。那么人工智能是否会生成意识形态呢?

事实上,人工智能技术并不能生成意识形态,生成意识形态的主体依旧是人。在现实生活中,人工智能带来的许多新的意识形态现象背后,其实是人工智能的应用产生的。人工智能正是在各种应用中生成应用逻辑对应的意识形态。因此,对人工智能生成意识形态的探究,应该深入人工智能一般应用、普遍化应用、意识形态斗争利用的应用路径中去把握。

第一,人工智能不同应用生成不同的意识形态。资本应用会生成符合资本增殖需要的意识形态,在党的领导下则会生成符合公共利益需要的主流意识形态。对这一不同应用彰显的不同意识形态内容,可以在人工智能技术链条的不同应用对比过程中得到解答。

首先,数据的资本应用和在党的领导下的公共应用生成不同的意识形态。人工智能的资本应用的第一步,就表现为在生产领域中,将作为生产要素的数据资本化。事实上,作为生产要素的数据资本化,就是数据的价值形态发展过程。数据作为信息的数字承载体,具有能够满足人们智能需要的使用价值。而从信息到数据的加工过程,包括数据标记、数据分类、数据注释等数据标注劳动,就是人工智能时代赋予数据价值的活劳动过程。资本通过雇佣数据标注劳动力的方式,占据整个数据标注劳动,从而无偿占有数据标注劳动所生成的剩余价值,进而实现数据资本化的增殖。如此一来,数据标注劳动就具有了增殖的目的和依据。“大数据的抓取、清洗与凝练等过程,背后的主体有自身依据和目的。”^③数据的选择、标记、分类、注释等劳动是以是否有利于资本增殖为标准的。于是,资本的生产方式嵌入了人工智能的数据生产,也使得数据资本化的同时,形成了符合资本需要的意识形态。数据标注劳动在选择什么样的数据、怎么分类数据、如何注释数据上,不仅以资本增殖为标准,还直接以精神劳动的资本化生产方式,生成了符合资本增殖需要的意识形态。

而人工智能数据在党的领导下应用于公共领域则不是一种生产方式,因而不是一种生产要素,就不具有增殖性而是具有公共性。那么,公共应用的数据,无论是数据选择、数据标记、数据注释,其目的是为人民服务,其价值趋向是在党的领导下具有社会主义属性,因而具有公共性。同时,公有制企业应用数据时,由于其生产资料所有制是公有的,使其对数据的应用不只是为了增殖,而是也具有承担社会服务的公共功能,因而也具有了公共性。正是如此,数据的不同应用会生成不同的意识形态。

其次,算法的资本应用和在党的领导下的公共应用也会生成不同意识形态。算法就是决定数据怎么选择、分类、注释的程序。这种选择性和决定性就包含了关于“什么是对的、什么是好的”的判断,

① 《马克思恩格斯文集》第1卷,第533页。

② 俞吾金:《意识形态论》,上海:上海人民出版社,2014年,第95页。

③ 王文臣、马梦雪:《论大数据作为生产要素的资本化及其双重影响》,《上海财经大学学报》2022年第6期。

包含了主体的目的性和价值性。所有的算法事实上都是意识形态的算法。而且在算法的资本应用上,算法不仅是一种生产,还成为人们的需要。算法作为一种新的智能服务商品,以自己智能化便捷的使用价值,来满足和再生产出更多的需要。但是,这些需要看似是人的需要,其实是资本的需要。人们在消费这些算法需要的同时,事实上是完成算法作为资本增殖的闭环。于是资本以头足倒置的方式,通过算法制造许多虚假需要遮蔽了人的真正需要,并以“信息茧房”“过滤气泡”“回音室效应”等方式生成消费主义、娱乐主义等意识形态。同时,在资本增殖超越空间限制的驱动下,算法作为最高效的生产方式,具有了全面普及和展开的可能。资本会不断推动算法革命,推动算法形成一种发展的逻辑,并嵌入现有的各种生产方式中,推进现有各种生产方式的算法革命,生产和再生产出更多的算法需要,进而实现资本的生产与再生产。于是,算法依赖成为这一生产变革的社会现象结果。算法通过代替人们选择,垄断人们信息接受的权利,进而等于剥夺了人们的自主性。无论是算法偏好还是算法推荐,都是以人工智能的自主性代替人的自主性,就是资本借助算法遮蔽了人的自主性。算法会告诉个体什么是好的、什么是对的,甚至决定了个体能看到什么、听到什么,而这背后无非是资本借助算法告诉个体什么有利于资本增殖,以及只告诉个体资本增殖的需要而不是人真正的需要。于是,算法资本化便生成了符合资本增殖需要的意识形态。

而算法如果是在党的领导下,以公权力进行的算法善治,其意识形态倾向必然是符合绝大多数人民的美好生活需要。此时,算法就不是为资本增殖的生产,而是公共服务的一种技术手段。届时,算法的偏好、推荐都是以党的领导和人民的美好生活需要为标准,就具有了建构符合人民美好生活需要的主流意识形态的可能。

最后,模型的资本应用和党的领导下的公共应用也会生成不同意识形态。意识形态往往是通过一定的意识形态(ideologischen Formen)来发挥作用的。意识形态既是结构化的结构(structuring structure),也是被结构化的结构(structured structure)。其中,人工智能模型的资本化,就是将资本的生产方式以算法模型的结构,结构化生产意识形态的生成结构。于是,“它们首先作为结构而强加于绝大多数人,因而不通过人们的‘意识’”^①模型的资本化,就构成了符合资本增殖需要意识形态的数字结构,成为生产资料。这一数字结构是在促进资本再生产的过程中,促进意识形态的生成。人工智能模型的设计和使用,存在着一个价值判断。这一价值判断受资本增殖逻辑的影响。凡是有利于资本增殖的,那么这个模型就是好的,反之亦然。而这个智能模型在推进各种生产方式智能化的同时,也实现了个体的主体化生产与再生产,也就是让每个个体在接触和使用这一模型的同时,必须接受这一模型的价值选择、价值判断,进而实现个体在智能模型中主体化。

而在党的领导下训练和应用的模型,就不是生产资料,而是非生产的意识形态。主流意识形态在模型的运行下,具有了不断再生产的可能。届时,公共模型作为数字结构本身就成为主流意识形态的意识形态,模型就成为不同意识形态的意识形态区别。

第二,人工智能的技术普遍应用也会产生意识形态。人工智能作为最为深刻的技术变革,在普遍应用过程中,尤其是在政治、文化中的应用,使得人工智能事实上成为一种合法性来源的权力,也就是生成了一种意识形态。事实上,这种在人工智能普遍应用过程中的意识形态生成,就是生成图式单一化为数据统计,进而使得数理逻辑构成意识形态的基本逻辑图式。

人工智能的数据、算法、模型,都贯穿着最基本的数理逻辑,即以一种数据的统计方式,把某个问题形式化为数学表达方式,进而在模型论、证明论、集合论、递归论等数理逻辑中具有了唯一的确定性。所以,人工智能在经验性事物的计算和判定上,具有了类似甚至超越人类智能的可能。然而数理逻辑其实是以数学方式把握的形式逻辑,也是对思维的形式规律形式化把握。

^① 路易·阿尔都塞:《保卫马克思》,顾良译,北京:商务印书馆,2006年,第229页。

形式逻辑注重的是思维形式,而不是思维内容,其提供的思维规律和规则表现为:在概念上遵守同一律;在判断上遵守矛盾律;在推理上要尊重排中律;论证要遵守充足理由原则。在形式逻辑的判断中只有是与不是,而且互为相反的判断不能同时存在,且无论是演绎、归纳还是类比推理都不能存在既正确又不正确的“中间地带”。最重要的是,形式逻辑规定了所有存在都应有根据,也就是存在的存在。

数理逻辑和形式逻辑随着人工智能的普遍应用,事实上使得思维规律单一化了,进而生成了符合形式逻辑的意识形态。因为形式逻辑注重的是思维形式规律,并不重视思维内容,所以往往无法解释社会中的一些悖论现象,也无法解释“二律背反”。人工智能在数理逻辑的运行下,会具有演绎、归纳、类比的智能,但是缺少辩证逻辑、情感、直觉、灵感,也缺少对内容的实质性关注。所以,最聪明的人工智能也不会像人一样忽然有直觉,忽然顿悟,而是必须经过数理逻辑进行推理。形式逻辑的这一局限性通过人工智能的普遍应用而不断放大了,生成了形式化、理性化的意识形态。

人工智能的普遍应用,会出现以思维形式遮蔽思维内容的意识形态。人工智能的普遍应用将会出现越来越多用“是与不是”代替“好与坏”的价值判断。人们不关注现实生活实践的真问题,而是聚焦于符号形式的对立,并以这种思维形式的对立达到对主要矛盾的遮蔽,进而发挥意识形态的作用。

人工智能的普遍应用,还会出现理性遮蔽感性的意识形态。包含情感在内,感性是人的重要思维方式。以往的意识形态也往往会借助人们感性的方式,以合情达到合理,最后达到认同。这一过程事实上是通过符号、语言等感性互动建构的价值理性。但是,随着人工智能的普遍应用,以数理逻辑为代表的理性思维构成了生产生活的主要思维方式,在事实上推动了以感性为主要方式的价值理性衰落,推进了目的理性的兴起。合法性来源不再是“向上”依靠宗教、神话等感性思维,而是诉诸“向下”,依托于以人工智能为代表的科技进步取得的成就。这会造成一个假象,即“社会系统的发展似乎由科技进步的逻辑来决定”^①。理性构成了人们认知和解释世界的合法性来源。量化、科学、客观等是与不是的形式逻辑取代了好与坏,以及介于好与坏之间的辩证可能。

第三,意识形态斗争也会利用人工智能,进而使人工智能具有意识形态性。意识形态由于自身的阶级性,始终存在意识形态斗争的现象,也就带来了意识形态风险和意识形态安全的命题。意识形态斗争往往表现为舆论宣传的话语权、价值认同的领导权的争夺,其背后是意识形态主导权的斗争。随着人工智能的出现和发展,人工智能成为意识形态斗争的场域和手段,并贯穿意识形态的生成、传播、认同的全阶段。当前,建立在第三次科技革命基础上的生产生活方式信息化,使得网络已是当前意识形态斗争的最前沿。互联网是意识形态在生成、传播、认同阶段斗争的重要场域和重要手段。而人工智能的发展,将在信息化基础上进一步推进智能化,从而推进意识形态斗争从互联网进一步发展到人工智能的数据、算法、模型的全链条中。

首先,意识形态通过数据主导权争夺的方式,使得人工智能具有意识形态性。数据是人工智能最为基础的单元,也是人工智能精神生产的基本要素,是人工智能生成“意识”现象,进而在应用中生成意识形态的基本元素。于是,意识形态的斗争,表现为对人工智能时代数据要素的争夺,意味着意识形态斗争转化为数据所有权的斗争。在人工智能时代,谁掌握的数据更多,谁就更能依靠人工智能技术丰富意识形态的表现形式,提升意识形态的传播力,增强意识形态的说服力。所以,意识形态在斗争的对抗性驱使下,必然会深入微观层面,进行数据的争夺。

同时,数据的选取、标注、注释等劳动本身就具有意识形态,是意识形态结构下的主体性行为。何种已有的意识形态,决定了选取何种数据,并进行何种意识形态性的标注和注释,从而生成何种意识

^① 尤尔根·哈贝马斯:《作为“意识形态”的技术与科学》,李黎、郭官义译,上海:学林出版社,1999年,第63页。

形态。所以,没有绝对中立的数据,也就不存在绝对去意识形态化的人工智能。正是在此意义上,人工智能的精神生产,会始终具有意识形态性。

其次,意识形态通过算法主导权争夺的方式,使得人工智能具有意识形态性。算法是人工智能运行的程序,也是人工智能精神生产的生产方式。采取何种算法,决定了人工智能精神生产的结果。不同的意识形态在斗争的对抗性驱使下,必然会争夺算法输入端的意识形态主导权,会表现为意识形态的算法之争。“算法黑箱、算法歧视、算法后门等技术问题也是意识形态的领导权、话语权、主导权的算法技术体现。”^①于是,无论是何种意识形态,都会通过算法的方式,不断再生产新的意识形态。人工智能便在再生产意识形态的过程中,既强化了原有意识形态的认同,又使得自己具有了意识形态性,生成了意识形态。

最后,意识形态通过模型主导权争夺的方式,使得人工智能具有意识形态性。模型是人工智能的数字结构,既由数据、算法建构而成,又决定了数据的输出和算法的运行,是智能时代的生产方式图式。意识形态斗争在智能化生产方式中的表现就是模型主导权的争夺。因为设计何种模型,并不仅仅是一个技术可行性问题,还是意识形态选择的问题。人工智能作为一种技术,是为人民服务还是为资本服务,是作为生产力变革方式推进社会发展还是作为资本增殖的手段推进资本增殖,这些都反映了人工智能模型设计的意识形态问题。换言之,人工智能模型设计是一个目的导向型的生产方式,是合目的性与合规律性的统一。人工智能的合目的性,也就是向善的人工智能模型还是向恶的人工智能模型,反映了人工智能生成意识形态的社会现象。

总之,人工智能之所以会出现生成意识形态的现象,是由人工智能应用所决定的。对人工智能生成意识形态的分析研究,不仅要把握人工智能生成“意识”现象的规律,还要深入人工智能应用路径之中分析意识形态生成规律。只有在科技路径与应用路径的统一中,才能客观全面地把握人工智能生成意识形态的现象。

三、人工智能时代的意识形态新现象

人工智能以新的精神生产方式会生成“意识”现象或者意识行为,人工智能的各种应用会生成意识形态。这种由科技变革带来的新的生产生活方式变化,也在经济基础和生产方式上决定了意识形态的创新发展。可以说,面对即将到来的人工智能时代,开展具有前瞻性的意识形态理论创新研究,进而把握人工智能时代的意识形态规律和现象,更好地开展意识形态工作,这既具有理论意义,又具有现实意义。

首先,人工智能时代将会重塑生产方式和意识形态的关系。阿尔都塞“强调意识形态本质上是对主体的物质性建构”^②。这一论断实际上强调了物质生产对意识形态生产的作用,揭示了意识形态范畴中物质生产的关系。因为在历史唯物主义视域中,“人们的社会存在决定人们的意识”^③。作为上层建筑的意识形态,归根结底是由生产方式所决定的。此时,意识形态往往是作为一种抽象观念被人们所认知和把握。意识形态生成后,往往也会通过各种物质实践环节,间接地反作用于生产方式,影响生产方式的发展。于是,物质实践的中介环节构成了人们感知意识形态,以及意识形态发挥作用的重要一环。总之,这一实践图式构成了前人工智能时代意识形态与生产方式的基本关系。

但是在人工智能时代,意识形态与生产方式的中介环节将会融入社会的总体生产之中,在生产中

① 刘伟兵:《算法的意识形态与意识形态的算法》,《深圳大学学报(人文社会科学版)》2024年第1期。

② 汪行福、俞吾金、张秀琴:《意识形态星丛:西方马克思主义的意识形态理论及其最新发展态势》,北京:人民出版社,2017年,第263页。

③ 《马克思恩格斯文集》第2卷,北京:人民出版社,2009年,第591页。

生成意识形态,同时意识形态也进一步促进再生产。于是,人工智能时代的生产是有意识的生产,在生产的同时也生成“意识”现象。无论是数据的选择、算法的设计还是模型的建构,都存在意识形态倾向的价值偏好与价值选择,包含了关于“什么是善的、什么是恶的”的价值判断,进而使得生产力发展不只是一个科技问题,还成为一个意识形态的问题。意识形态使得人工智能“人化”“主体化”的同时,人工智能也进一步强化了原有的意识形态内容,使得意识形态由抽象的理论论证转向具象的生产强化,在再生产中不断实现意识形态循环生成。于是,人工智能时代的意识形态将会更加隐蔽。

其次,人工智能时代将会重构意识形式。马克思在建构历史唯物主义的过程中,多次强调了意识形式(Bewußtseinformen)、意识形态形式(ideologischen Formen)的概念。在历史唯物主义视域中,“意识形态在内容上是精神的,在形式上却不得不是物质的”^①。这是因为,“意识形态在本质上是实践的”^②。无论是意识形态生成、传播还是认同,依赖的都是人的实践。正是在此意义上,意识形态的主体才是属人的。实践是人的感性对象性活动,所以在实践维度上把握的意识形态,就具有了真善美统一的维度。于是,意识形态就是真理性表达、向善的引导性发挥与表现形式美的统一。而意识形态内在的真善美的统一,需要寄托于感性现象。这种感性现象就是意识形式。“‘意识形态’是一个总体性的概念,是由各种‘意识形式’——哲学、宗教、伦理、政治、法律等等构成的有机的整体。”^③人工智能时代对意识形式的把握可以从精神生产和精神产品两种维度出发。

在人工智能时代,精神生产是意识形式的重要表现形式。以往与意识形态相对应的艺术、宗教、国家机器等意识形式都是物质实践活动。而随着人工智能发展,意识形式具有了数字劳动等非物质劳动的现象,进而具有了建构新的意识形式,发展意识形态新的美学向度可能,也就在推进“劳动生产了美”^④的形式因和“人也按照美的规律来构造”^⑤的目的因的统一的张力过程中,生成新的意识形态。在人工智能时代,意识形式既是劳动生产美的过程,又是美的本身。数据投喂、算法设计等数字劳动,既能生产美,又本身就蕴含着各种想象力的审美。想象力与劳动,即审美与生产美,在人工智能时代的精神生产中形成了统一。在此意义上,人工智能时代的精神生产虽然依旧受资本逻辑影响,具有异化的丑陋,存在对人自由自觉的精神劳动的狭隘化现象,但是人们借助人工智能通过算法设计、模型建构的精神生产方式,具有了在一定程度上增强人的精神劳动能力和对抗资本逻辑的自由自觉力量。这种对抗性构成了意识形态形成的张力。于是,精神生产成为把握意识形态的重要意识形式。

在人工智能时代,精神产品成为意识形式的重要表现形式。数字虚拟人、元宇宙等新的意识形式出现,意味着人工智能时代也具有了非物质的精神产品,并能够成为精神生产的劳动对象、劳动产品。“生成式AI在消费端的应用场景主要有内容生产、便捷交互、简化操作,这些应用将会同时对消费者的消费品数量、质量、多样性,以及拥有的闲暇等因素产生显著影响。”^⑥于是,资本借助生产资料对产品的无偿占有,从物质世界延伸到了精神世界。资本原本是通过价值形式占据劳动者的剩余劳动,进而占有劳动者的剩余劳动时间,也就是劳动者自由全面发展的机会,从而“窃取”了劳动者创造的文明,并在这种“窃取”文明的不正义基础上建构起虚假的意识形态。“资本家是窃取了工人为社会创造的自由时间,即窃取了文明。”^⑦随着人工智能对美的具象化生产,资本将作为自由自觉发展的精神劳

① 俞吾金:《意识形态论》,第100页。

② 俞吾金:《意识形态论》,第97页。

③ 俞吾金:《意识形态论》,第95页。

④ 《马克思恩格斯文集》第1卷,第158—159页。

⑤ 《马克思恩格斯文集》第1卷,第163页。

⑥ 陈永伟:《超越 ChatGPT:生成式AI的机遇、风险与挑战》,《山东大学学报(哲学社会科学版)》2023年第3期。

⑦ 《马克思恩格斯全集》第31卷,第23页。

动嵌入资本的生產方式之中转化为精神生产,进而占有精神产品,从而直接以“新的剥削模式”^①“窃取”劳动者的精神文明。于是,精神产品本身所蕴含的人的自由全面发展与资本增殖的对抗性,构成了意识形态形成的张力。精神产品成为把握意识形态的重要意识形式。

最后,人工智能时代将会重置意识形态的特征。在人工智能时代,意识形态由于生产生活的智能化,以及新的意识形式,将会出现许多新的特征或者促进原有特征的发展。

其一,意识形态将会以嵌入生产环节获得合法性的方式增强隐蔽性。如果说在前人工智能时代,意识形态作为上层建筑是独立于经济基础之外,并以与经济基础对立和互动关系彰显自身存在,那么人工智能时代意识形态将会以嵌入生产方式中的方式,实现自己的再生产和作用发挥。这样一来,意识形态不仅作为生产生活的生成物,还成为生产本身。意识形态以嵌入生产环节的方式,使自己获得了不证自明的合法性,并在生产与再生产的过程中不断实现自己的“循环证明”。于是,意识形态作为一个总体性范畴,在人工智能时代以整体嵌入的方式,整体隐蔽在人的生产生活中。这一隐蔽性增强的趋势,将导致现有的各种意识形态风险更加难以发现。人们越是依赖人工智能,依赖算法、信任数据,那么资本逻辑生成的意识形态借助人工智能,将自己等同于科学进而等同于真理,从而以日用而不觉的方式发挥作用。

其二,人工智能时代的精神生产方式将会使得意识形态虚假性更“真实”。人工智能对大数据的采集、算法逻辑的设计,在一定程度上是成功模拟人类的形式逻辑和数理逻辑,在经验性事务处理上具有类人,甚至是超人的智能水平。于是,人工智能时代的精神生产实现了从个体的精神劳动到社会化的精神生产的飞跃。作为承载意识形态的精神产品,意识形式将会更加符合人的形式逻辑,从而以符合逻辑的方式获得科学性以及美感。事实上,人工智能已经开始出现捏造虚假信息、进行虚假证明、伪造虚假图片的现象,并且能够“以假乱真”,从而以在美学上增强“真实性”的方式增强意识形态的虚假性,让人们以喜闻乐见的方式制造信息,选择信息,进而以自我偏好的方式编织成一个美的意识形态世界,即信息茧房(information cocoons)。

其三,意识形态将会呈现为对人发展异化的风险趋势。人工智能时代意识形态风险将不再只是政治风险、安全风险,而是愈加暴露出对人的发展的风险,也就是意识形态对人的异化将会愈加全面和全领域。“随着人工智能技术的发展,加深了人的自我异化危机,导致人类的主体性危机以及人类交往的异化形态。”^②在人工智能时代,意识形态将会出现依赖人工智能的趋势,进而资本统治人的意识形态风险会表现为“机器统治人”的假象。在资本增殖的驱动下,人工智能将会代替越来越多人的劳动机会,进而剥夺人的发展机会,从而在意识形态上生成人是“无用”的意识形态观点,进而在否定人的主体性发展过程中,肯定人工智能促进资本增殖的“有用”。意识形态对阶级、国家、民族的风险,将在人工智能的“催化”作用下,上升到对人作为类的整体性风险。

总之,对人工智能时代的意识形态研究,不仅要关注到人工智能作为新的科学技术带来的新变化,还要深刻认识意识形态自身的规律,并在唯物史观的整体性基础上明晰意识形态如何生成。通过梳理,可以明确地说,人工智能在技术路径上会生成作为意识形式的“意识”现象,并在资本应用、技术滥用、意识形态斗争利用下会生成意识形态,从而会推进意识形态出现新现象。

① 李厚羿:《马克思政治经济学中精神生产概念的当代视野》,《中国高校社会科学》2022年第3期。

② 闫坤如:《人工智能技术异化及其本质探源》,《上海师范大学学报(哲学社会科学版)》2020年第3期。

Will Artificial Intelligence Generate Ideology?

Liu Weibing

(School of Marxism, Fudan University, Shanghai 200433, P.R.China)

Abstract: Artificial intelligence (AI) doesn't possess consciousness, yet it exhibits 'conscious' phenomena and 'conscious' behavior. With the development of artificial intelligence which shapes production methods, spiritual labor has entered the stage of socialized mass production of consciousness. Behind the 'unmanned' generation of 'consciousness' by AI through the technological chain of data, algorithms, and models, is the social collaborative spiritual production of the collective laborer. The current study of AI's 'consciousness' phenomena and 'conscious' behaviors essentially explores the transformation of spiritual production modes in the AI era, reflecting the new mode of consciousness generation that emerges from the collaborative changes in spiritual production. The various ways in which AI realizes machine consciousness don't enable AI to produce real 'consciousness', and the essence of its autonomous generation of 'consciousness' phenomenon is the spiritual production in socialized mass production. It is the artificial intelligence that obtains the 'consciousness' given to it by the collective labor under the collaboration of spiritual production. The essence of its autonomous the generation of 'consciousness' phenomenon lies in spiritual production within socialized mass production, and it is the AI that, under the collaboration of spiritual production, acquires the 'consciousness' phenomenon and 'conscious' behaviors endowed by the workers in general, but it doesn't mean that the AI has consciousness like human beings. AI technology itself does not generate ideologies, yet its application gives rise to new ideological phenomena. The phenomenon of 'consciousness' generated by AI, after the general application, the generalization of the technology, and the use of ideological struggles, can be transformed into an ideology by the domination of the capitalist logic, the assimilation of formal and mathematical logic, and the penetration of the logic of the original ideology. Different applications of AI can engender distinct ideologies, and as AI technology becomes more universally applied, it will contribute to the reproduction of ideology. Ultimately, ideology can shape AI into an ideological tool through struggles over data, algorithmic, and model dominance. As a new technology, AI promotes intelligent changes in the modes of production and enhances the hidden, false, and alienating characteristics of ideology. The era of artificial intelligence will reshape the relationship between the modes of production and ideology, and its intermediary link will be integrated into the overall production of society, realizing the cyclic generation of ideology in production and making ideology more covert; the rise of AI fosters a new form of consciousness expression, incorporating digital and immaterial labor and pushes for the construction of a new form of expression of the form of consciousness such as spiritual production, spiritual products, etc., and strengthens the ideology's falsity. In the era of artificial intelligence, ideology will enhance its concealment by embedding itself into production processes to gain legitimacy, further making its falsehoods be felt like 'reality' and presenting a growing risk of alienation in human development, thereby redefining the characteristics of ideological alienation.

Keywords: Generative AI; Consciousness; Ideology; Form of consciousness; Spiritual production

[责任编辑:王玲强]