

Doi:10.20063/j.cnki.CN37-1452/C.2024.02.001

概念、关系及视角:人工智能需要灵魂吗?

——传播学五人谈

王豪^{1a},葛岩²,姜华³,王峰^{1b},甘莅豪^{1b}

(1. 华东师范大学 a. 国家话语生态研究中心;b. 传播学院,上海 200062;2. 上海交通大学 媒体与传播学院,上海 200030;
3. 复旦大学 新闻学院,上海 200433)

摘要:ChatGPT、Sora等生成人工智能的迅速发展激起了人们对新技术与未来人机协同社会的多重想象和忧虑,由此生发的关于人工智能是否会拥有灵魂、是否需要灵魂等问题也引发了激烈论辩。通过与四位传播学领域的人工智能研究学者进行对谈,在重回人工智能与灵魂的概念丛后,针对人工智能能否被视为一种智能形式、其是否可能拥有灵魂或其要件、没有肉身的智能体灵魂以何为锚点、如何构建未来社会中的人机关系等问题,逐层展开对话,寻找“人工智能需要灵魂吗”这一问题的答案。

关键词:人工智能;灵魂;人机交往

中图分类号:TP18;G206.2 **文献标志码:**A **文章编号:**1673-8039(2024)02-0001-12

近期,文生视频大模型 Sora 的横空出世,引发了全球科技界与公众的热议。在 ChatGPT 等生成式大语言模型的快速破圈后, Sora 及其官网公布的数据高清演示短片再一次彰显出当前人工智能飞速发展。在成为新一轮科技革命的引爆点的同时,人工智能也激起了人们对新技术以及未来人机协同社会的广泛思考和担忧。由此产生的问题,例如人工智能是否拥有灵魂,以及是否需要具备灵魂等问题,已成为多个学科内争论的热点之一。我们邀请了四位传播学领域内人工智能研究的学者对这一问题进行对谈。本文的对谈人包括:

葛岩,上海交通大学媒体与传播学院教授、心理与行为科学研究院特聘研究员;

姜华,复旦大学新闻学院暨复旦大学全球传播全媒体研究院研究员;

王峰,华东师范大学传播学院院长,华东师范

大学传播学院、中文系联聘教授,教育部青年长江学者;

甘莅豪,华东师范大学国家话语生态研究中心副主任、传播学院教授、美国加州伯克利大学访问学者。

讨论主持人是王豪,现为华东师范大学国家话语生态研究中心助理研究员。

一、人工智能:一种不同的智能形式?

王豪:首先让我们回到“人工智能”的概念切入今天的讨论。1956年,“人工智能”(Artificial Intelligence, AI)的概念被首次提出^①。英国人工智能哲学学者斯洛曼(Aaron Sloman)认为,区别于人或其他动物所具有的“自然的智能”,“人工的智能”是指设计出来的智能^[1]，“人工”代表人工创建、设计和实现的方面,而“智能”则代表了模拟和表现出类似人类智能的能力。现在大家已

收稿日期:2023-12-21

基金项目:国家社会科学基金重点项目“以新言行为分析为核心的汉语修辞学理论研究”(19AYY002)

作者简介:王豪(1997—),女,安徽六安人,华东师范大学传播学院博士研究生、国家话语生态研究中心助理研究员;葛岩(1956—),男,陕西西安人,哲学博士,上海交通大学媒体与传播学院教授、博士生导师;姜华(1977—),男,河北衡水人,文学博士,复旦大学新闻学院教授、博士生导师;王峰(1971—),男,辽宁葫芦岛人,文学博士,华东师范大学传播学院教授、博士生导师,教育部青年长江学者;甘莅豪(1977—),男,江西高安人,文学博士,华东师范大学传播学院教授、博士生导师,美国加州伯克利大学访问学者。

①1956年,在美国达特茅斯学院召开的“人工智能夏季研讨会”上,以麦卡锡(John McCarthy)、明斯基(Marvin Lee Minsky)、罗切斯特(Nathaniel Rochester)和香农(Claude Elwood Shannon)等为首的一批有远见卓识的年轻科学家共同研究和探讨用机器模拟智能的一系列有关问题,并在会上首次提出了“人工智能”这一术语,这也标志着人工智能学科的诞生。

经对“人工智能”的提法习以为常,但是普适的智能理论尚未构建^[2],所以我们想问的是,相较于人类基于意识、理性、情感、身体等形成的智能,“人工智能”能不能被称为一种智能形式?它和人类智能有哪些相似和区别之处?

葛岩:在我的理解中,“智能”这个概念表示人类处理外部信息、形成自己观点,然后去观察和判断事物的一种能力。在这个意义上,我认为,人工智能是由人类设计,并且至少最初是以人类认知作为蓝本,模拟我们人类认知的。它的来源与我们不同,我们的智能是由生物和社会演化的结果。而目前,人工智能是我们设计的产物,迄今为止,它的表现与演化而来的人类智能是无法比拟的。

姜华:葛老师在刚才的讨论中谈到了智能是什么,他在这一领域有着丰富的研究经验。我的粗浅理解是,智能可以被解释为解决问题的能力。若从更广泛的角度来看,早期的研究,如乔治·米德这位社会学芝加哥学派的代表人物,对智能的定义着眼于人们面对未来潜在情景所能做出的反应,以及基于过往经验所形成的思维能力。当然,那个时期的研究并未牵涉到人工智能,亦未触及计算机技术。在面对问题时,智能体展现出了具有选择性的行为能力。在这方面,我认为包含着两个核心要素。首先,人类智能,它不仅仰仗于过往的经验,更加重要的是它还具有面向未来的筹划性。就目前的情况来看,人工智能与人类智能之间存在着显著的差距:人类智能在其本质上具有面向未来的筹划,而目前人工智能尚未达到这一层次。其次,在解决问题的能力方面,包括目前广泛使用的语言模型在内,都是基于人类的训练而衍生出来的。这是我对人工智能的简要看法,从中也可以看出,它与人类智能有明显的差异。当然,关于这一议题还存在着诸多可深入探讨的方面,我也留意到了王峰老师最近在《探索与争鸣》杂志上发表的论文中所阐述的内容,是十分精彩的。因此,我们或许应当先聆听其他老师对于此问题的看法。

王峰:我谈谈自己的看法。刚才葛老师和姜老师对人工智能的理解,我认为非常有道理。十年前,我们可能觉得人工智能达到与人类相当的智能水平是遥不可及的事情。但自2016年以来,人工智能的发展突飞猛进。阿尔法狗战胜李世石,人工智能在最复杂的棋类游戏上击败了人类。

在自然语言处理领域,人工智能取得了最高的成就,但距离自动产生有意义并且能够让我们理解的高质量语句,我们仍然觉得还需要一段时间。没想到ChatGPT出现了,超出大多数人的理解和预期。这些事件表明,人工智能正在以超乎我们之前预期的速度进步。个人而言,我并不认为人工智能会进化到消灭人类的程度,我对人工智能持有与人类相融合发展的期待。

关于智能的定义,刚才两位老师也提到了,人工智能的形态显然不是人的智能形态。人的智能是天生的,但是人工智能从一开始,其基础原理就是要模仿人的智能活动。智能的标准为人。然而,针对人工智能的起源和基础原理是否是模仿人的智能活动和改造人的智能机制,不同学者可能有不同的观点。从目前的实际情况来看,在某些专业领域,人工智能形态确实让人惊讶,甚至引起一定恐慌,比如,如果哪一天人工智能进化到与人类智能相同的水平,不经意间就能将我们打败,那该怎么办呢?最近我看到孙正义筹集了几百亿资金投入人工智能产业,他认为人工智能一定超过人类。虽然我个人持有不同的看法,但从目前的情况来看,似乎我们仍然默认人工智能是一种智能形态,而且我的感觉是,现在想否定人工智能是一种智能形态反倒是比较困难了。先不讨论智能定义的问题,仅从实际操作上来说,如果想否认人工智能是一种智能形态,那就需要提供更有说服力的理由和解释。

甘荏豪:前面三位老师提供了深刻的见解和清晰的定义。实际上,我们前面的讨论都涉及到了人工智能和自然智能这两个相对应的概念。它们的本质区别在于智能产生形式的不同。自然智能主要是碳基生物,基于神经突触之间的放电,形成认知模块。不同物种通过认知模块与感知形成外在经验,并将这种经验储存在基因构成的认知模块中,然后进行思考,形成智能。从智能的角度来看,人类、狗甚至生物蠕虫等都可能展现出某种形式的智能。而机器人的人工智能在物质基础上,与我们不同,它们通常是硅基物质或金属片,通过电子的传递和数据交流。它们的机制和元素与我们不同,但它们是否是智能呢?

就智能这一概念而言,若我们不纠结于其表面形式,而更关注其内涵,我们应当回溯至“智”的词源和语言体系以更深刻地理解它。从中国汉字构词的角度看,“智”一词具有丰富的内涵。从

字形来看,它是由一个箭头和一个口组成上半部分,首先,“智”和“知”的字形相似,这表明它们有着共同的词源。在这个字形中,口代表了言语和信息的传递,而箭头则象征着大脑思维的迅捷。这个构词方式暗示了古代人将“智”与“信息传播”直接联系在一起,将思考的迅速性视为智慧的一部分。同时,考察“能”字的构成,我们可以发现它与“熊”这一强大的动物形象相关联,即一种强大的能力。因此,“智能”可以被理解为人类迅速处理信息的能力。从这个角度来看,基于中国汉字的构词方式,人工智能与中国古代对“智能”的概念相符合。AI具备了快速处理信息的能力,且在信息处理方面拥有许多人类智能所不具备的优势。虽然相对人类,人工智能也存在劣势,但无论如何,它都拥有对信息进行处理、组织和传播,以及影响其他个体或群体的能力。根据这一定义,可以断言人工智能无疑是一种智能。

但是,这还涉及另一个问题,即不同生物的智能形式可能在地球或宇宙中是千差万别的。人类智能不一定是宇宙间唯一的智能形式。未来学家凯文·凯利提出了智能多样化的问题。他说现在万物都已经开始智化了。实际上,我们的智能也具有多种形式。举例而言,他提到大约有50多种不同形式的智能存在,比如其中某些形式可能侧重于大容量存储和记忆,而在其他方面可能相对较弱。再比如,信息面广博也是一种智能,这与目前机器人的情况相符(当然,未来机器人的愚钝弱点也可能被克服)。此外,凯文·凯利还探讨了其他心智的可能性,例如,是否可能创造出比自身更强大智能的智能形式。这些思考突显了智能的多样性和复杂性,以及在未来可能出现的不同智能形式^{[3]45-49}。

另外比较有特色的是半机器半动物的共生形式以及半人半机器的赛博格形式。这些独特的智能形式将极大地拓展我们对智能的理解。未来,随着机器人的不断发展,社会可能会出现类似变形金刚一样的硅基智能形式,以及元宇宙中数字人的存在。这些都是后现代、后人类社会人们必须面对的新型智能现象。未来社会的一个显著特点可能是,人们将不仅仅与动物、宠物进行交流,还将与赛博人、数字人以及人机结合体进行更多的交流^[4]。

王豪:听上去已经有了一些不同的意见,不知道有没有老师想要对其他老师的观点进行回应?

甘莅豪:我想就姜老师的观点进行一些回应。我个人认为人工智能不仅仅是对过去经验的总结,在某种程度上,我认为它可能具备面向未来的创造和设计潜力。为什么如此说?是因为我和主持人王豪曾就 ChatGPT 进行过一个试验,尝试让人工智能创造新词。造新词,这实际上是一种创造性的行为。我们描述了各种场景,要求人工智能创造与这些场景相关的词汇。结果,它成功创造了很多在汉语中尚不存在的词汇,这些词汇很好地捕捉到了当时讨论的现象。这种创造力不仅仅基于过去的经验,而且具有涉及指向未来的潜力。因为人类是未来导向的生物,我们的生存和价值观通常与未来相关联,所以人工智能这种潜力不容忽视。尽管目前的机器人尚未具备这种未来导向的生存意义和存在价值,比如当前 ChatGPT 只是进行大数据的整理和收集工作,尚未形成自己的人格和意识,但我认为在未来赛博格和数字人存在的空间中,机器智能可能会具备这些特征。这个问题目前还没有确切的答案,但我们可以期待未来的发展。

姜华:您和王豪博士通过 ChatGPT 观察到了它所创造出的新词,我愿就此提出两个观点。当然,我对语言学领域的了解有限,而您是该领域的专家,因此若有不妥之处,还请您予以指正。首先,我想提到的是 ChatGPT 所创造出的新语言形式。这些新词,对于个体而言,可能仅具备某种形式上的新意,事实上却缺乏实质上的语义新内涵。与之相对,人类语言的表达往往在形式上和语义上相辅相成。看上去 ChatGPT 似乎并未具备创造新语义的能力,即作为这些词汇的生产者,它并不创造新的语义内涵。其次,值得注意的是,尽管这些由 ChatGPT 创造的词汇在形式上呈现出一种新颖感,实际上它们仍然是基于已有的语料库而形成的。换句话说,这些新词汇对于个体使用者而言虽然可能是新的,但在人类整体的语言认知中,却无新意可言。

另外还有一点,即刚才王峰老师所提及的“人工智能的问题解决能力”。我对这一观点表示认同。人工智能应该只能被看作是一种智能形式,与我们所知的人类智能相比,二者存在着较大差异。近期有一则新闻引起了我的关注,涉及我国台湾地区的一家电视机构。该机构使用人工智能模拟出一个气象预报主持人的形象。首先,这一虚拟形象被受众广泛认可且极具亲和力。同

时,其面部表情,甚至包括眨眼的细节,都相当自然。若不事先提示其为人工智能主播,几乎难以分辨出它不是人类。然而,这家电视机构也明确告知观众,它所播报的所有内容均事先由人工编辑完成,而非人工智能自行生成。这一事实呼应了王峰老师所提到的看法,即人工智能正在以迅雷不及掩耳之势深入各个领域,特别是与新闻传播行业密切相关的领域。当下,它已迅速融入整个传媒内容的制作流程中,这一趋势确实是明显而迅猛的。

二、灵魂:人工智能可能拥有或需要拥有吗?

王豪:基于智能的不同定义而展开的讨论,帮我们打开了对于“人工智能是不是一种智能形式”这一问题的观察视角,那么我们现在来到“人工智能需要灵魂吗”的另一端——“灵魂”问题。实际上,关于人类灵魂的构成有许多不同的观点,如柏拉图认为,灵魂的三个组成部分是理智、激情和欲望,而理智在这三者中又处于领导性的最高位置^[5]。此外,意识以及老师们刚刚提及的记忆、情感等,都曾被认为是灵魂的重要组成部分,且在不同的观念体系内,这些组成部分也有着重要性上的差异。那么哪些是灵魂所必备的,或者是说它最重要的一个要素呢?我们认为对于人来说最重要的这些灵魂的要件,人工智能可能拥有吗?或者说它需要拥有吗?

葛岩:世界上有很多关于灵魂的不一样的说法,如果退到纯粹的个体角度上看,灵魂是我们对某种东西的一个概括性体验。我们有物理肉体的存在,我们有具体的对于信息的反应。在认知上,你忽然觉得有一个什么东西正在观察你的认知过程,灵魂好像就是这样的东西。当然,根据不同的宗教信仰、政治信仰,可以把灵魂概括成是一种超自然,或者是阶级利益的社会价值。我个人是这么理解的。

推演到人工智能,你会觉得这个具体的知识的反应、概括甚至提问题的能力,它全都有,但是它有没有类似我们人类一样的灵魂呢?这个问题的提出本身暗示着——如果我们希望它的能力不断增长,就应该有灵魂……。通常情况下,人类觉得似乎有一个脱离我们肉体的东西在统治我们,所以AI也需要这么一个东西。我们可能觉得有了灵魂之后,人才能够对自己有一定限制,才会具

有道德责任。比如说哪个人不好,我们会说他是“行尸走肉”“没有灵魂”。有了灵魂,人好像变得向善了。所以,当我们问“人工智能需要有灵魂吗”这类问题,是想说“AI怎样才能负起责任”?这不算我的回答,而是我的疑问。

甘莅豪:我想就葛老师提出的观点进行一些回应。葛老师观点和康德非常类似,我非常认同这种看法,即“灵魂”一词与人类的道德感密切相关。为什么我们需要创造“灵魂”这个词汇?这实际上与我们人类的道德感有关。康德曾经深入探讨过,道德感的存在使得我们能够对自己的内心进行一种约束^{[6]61-72}。作为社会性动物,人类必须在社会中遵守一定的道德规范,以便整个社会能够和谐共存。因此,创造“灵魂”这个词汇与道德感的社会价值需求以及宗教信仰相关。

此外,对于“灵魂”的理解需要置于不同的语境和思想体系中来考虑。在中国的历史上,最早我们谈论的是“魂魄”,即有“魂”有“魄”^[7]。早期,中国人将“魂”和“魄”分开,与本土的道教信仰有关。在字形上,“魂”旁边有一个“云”,代表着“出气”。婴儿在出生并第一次呼吸气体时,被认为进入了阳间,因此“阳”与“魂”相关。“魄”则指的是一个人的体魄和胆魄等。婴儿在母亲的子宫中就早形成体魄和胆魄,但直到吸入第一口气时,“魂”才被认为出现,完成从阴到阳和阴阳结合的过程。因此,“魂魄”的概念与道家哲学中的阴阳相关,“魂”代表阳,“魄”代表阴。这两者及其两者的结合,都离不开人自身。

关于“灵”和“魂”的区别,可以从它们的字形和含义入手。在甲骨文中,“灵”字的上半部分是“雨”,下半部分是“巫”,这表示巫师在进行祭祀仪式时感受到上天降下的存在。这个“灵”与人自身的本质关系不大,它被看作是外部存在的东西。举例来说,“精灵”一词中的“灵”意味着它不属于人类,而是一种异于人类的存在。因此,“灵”一词通常表示“超越人之外的存在”。与此不同,“魂魄”虽然可以被想象为离开人体而存在,但它仍然被视为属于人类本体的一部分,例如魂魄被认为可以轮回转世。在中国古代,对于“灵”和“魂”的理解是依照这个标准分开的,直到基督教传入后,这两个概念才逐渐合并。基督教圣经中描述上帝吹入人体的“灵”,这就将“灵”与上帝紧密联系在一起,不被视为人类自身产生的

东西,而是先于人的诞生而存在。在“灵”被上帝赋予人后,人就天然地将上帝带来的东西——“灵”融化为自身的东西——“魂”。于是经过这种理解,“灵”和“魂”就理所当然融为一词“灵魂”,不分彼此了。康德、笛卡尔等哲学家也是在这种圣经文化传统中讨论“灵魂”的概念。

由于西学东渐,当今我们讨论人工智能是否需要“灵魂”时,实际上多半将“灵魂”置于基督教伦理体系和西方文化思想体系的概念下进行思考。只是我们大多数人没有意识到这一点而已。这也引发了更深层次的问题,即人工智能时代思考“灵魂”这一概念,其实涉及的是机器人的伦理问题,这更多是继承了康德对灵魂的思考,也是西方基督教伦理体系下的新问题。这些思考与我们当前讨论机器人伦理以及是否接受机器人存在的问题密切相关。

王峰:灵魂真的是一个特别大的问题。我们只去限定一个范围,因为我们讨论的背景是人工智能,所以若要讨论灵魂问题,应该沿着上述对人工智能与人类智能的区分,将它们联系在一起。借前面甘老师所说的,人是一种碳基智能,人工智能最多达到硅基智能的程度,两种形式应该是不太一样的。我们谈的灵魂,它应该属于接近碳基生物的东西。但是这么去看的时候,我们也的确面临着一个困难,也即我们现在使用的所有这些概念都是借自于所谓的碳基生命的概念。无论灵魂、心灵,还是意识、智能,都是跟人、跟碳基生命直接相关的。当你把一个来自碳基生命的概念,转换到人工智能的时候,会发现这不适用,那也不适用。但是哪怕不适用,我们也要考虑人工智能之所以被叫做智能,这种智能形态应该跟人的智能形态挺相近的。人具有灵魂,具有心灵,具有意识,这些东西是不是也可以转移到人工智能上去?现在有一派观点在谈人工智能意识涌现的问题,如果过去我们认为人工智能不可能拥有的意识,现在可以涌现,那么虽然不能画上等号,但这是否意味着某种形态的心灵或者某种形态的灵魂也有涌现出来的可能呢?智能问题本身就意味着向未来思考,我们既要看到现代人工智能的发展,又要推测人工智能未来可能的发展并进行探索。

姜华:进一步延续王峰老师刚才提出的观点,我对于这个问题的思考有些相似。目前我们思考人工智能相关的问题所运用的概念几乎都是以人

类为基准的,很多术语,比如灵魂等,实际上都是从人、生命体以及生物体的特性中演化而来。从路径看,大脑、意识、心灵以及灵魂,其实代表着从实向虚的一个过程。在这些概念中,灵魂更偏向虚拟层面。甘老师之前也探讨了在中国语境中“灵魂”概念的形成过程。依我的理解,灵魂这个词在英文中出现得也较早,大约在12世纪之前。12世纪已经开启脱离中世纪、走向现代的进程。在欧洲的文化背景或者说是英语的文化背景里,灵魂这个概念虽体现出一定的宗教色彩,但并非完全与宗教相关。以今日的视角来看,“灵魂”或许意味着一种精神存在,既可以与身体相连,也可能脱离于身体而存在。在中西方语境中都有此意涵,因此从这个角度来看,若让我直接回答“人工智能是否有灵魂”,我的回答是:人工智能具有灵魂,但并非拥有灵魂。人工智能在智能解决问题的过程中具备一种灵魂的特质,但它自身并不拥有灵魂。此处“有灵魂”指的是作为精神存在的人工智能,或许是静态的,是从外部进入的,而不是自身生发的。举例而言,我们欣赏肖邦或莫扎特的音乐时,通过各种形式的媒介,我们实际上是与莫扎特和肖邦的灵魂在进行交流,这种灵魂通过一定的媒介呈现。这种表现方式同样适用于人工智能,但我认为它是一种静态的灵魂。另一种形态是流动的灵魂,更准确地说,可能仅归属于人类,是人类通过意识的运作解决实际问题。因此,我认为人工智能具备灵魂,但不拥有灵魂。换句话说,它不拥有流动的灵魂。然而,我不敢确定这种看法是否合理。如王峰老师所述,灵魂涌现、智能涌现等观点在许多科幻作品中已有体现。诸如阿西莫夫所创造的银河帝国世界中,机器人丹尼尔的大脑最终被转移到了名为菲龙的小男孩身体内。然而,未来是否能够实现此种场景,在目前看来或许仍属于科幻范畴。

葛岩:我听了各位老师的说法,也说一下我的看法。说到灵魂这个东西,我们的担心实际上可能是,万一这个东西难以控制了怎么办?它在道德、价值上和人类冲突怎么办?我愿意把这个问题通俗化成“人工智能有没有主体性”,或者“AI是否能够最终产生出主体性”。作为在智能动物中占据生物链顶端的物种,我们很担心这件事。虽然我们有了原子弹这样厉害的东西,但原子弹没有主体性。让我们害怕的是,人工智能各种表现越来越像人。如果以后它有了主体性,怎么办?

把我们和最坏的人类相比,例如希特勒,还是能发现共性的。因为希特勒也是人类,我们可以揣测他的行为规则,不管是邪恶的规则,还是善良的规则。假如这个主体变成一种打引号的新物种,有自己的主体性,按照《人类简史》的作者尤瓦尔·赫拉利所说,这将是人类制造的唯一一个自己主动做决定的东西,将是人类最大的威胁^{[8]388-391}。AI一旦有了主体性之后,可能变得不可预测。这是我要补充的想法。

甘苕豪:葛老师提出了一个非常有趣的观点,涉及到主体性与社会性的关系。我认为主体性的形成是一个社会性的概念。新生儿在出生时,往往感觉自己与母亲融为一体,没有明确的自我概念。然而,随着时间的推移,根据拉康的理论,孩子开始认识到自己与母亲是两个分离的个体,母亲成为了一个他者。孩子必须通过语言或哭喊等方式来引起母亲的注意。这标志着符号系统的出现,同时也形成了一个更大的他者(象征符合系统)。在与母亲的互动中,孩子逐渐形成了自我概念,这个过程涉及到欲望、需求和博弈^{[9]89-96}。我们就小孩自我的形成来看人工智能,可以发现,目前的机器人仍然与人类和其他制造物融为一体。然而,葛老师提出的有趣问题是,机器人是否能够形成自我?这涉及到了人工智能是否能够具备主体性的问题。这个问题是值得思考的。因为人类的自我是和“生存欲求”催生的“符号社会”相关。虽然人工智能拥有符号和信息,但它是否具备生存意识尚不明确。所以人工智能的主体性会不会出现,尚且存疑。但是我们不排除未来它可能会出现想要生存的欲望。因为人类的生存欲望可能是基因设定的,基因本质上也是信息的集合体,是信息交流的过程。基因信息设计促使人类的大脑神经放电,放电反应与社会互动最终形成了自我,包括向外界发出和接收外部信息符号。那机器人通过电子网络,基于0和1数据模式进行放电,最终是否会形成生存欲望和自我意识?这是一个有可能出现的情景。这引出了一个更深刻的问题,即如果机器人形成自我,是否会类似人类那样产生灵魂的概念?因为如果机器人想要生存下去,它可能需要一套道德规范来约束自己,协调社会关系,这也与我们讨论的主题密切相关。

王峰:我接着上面老师的讨论。我们讨论的时候担心,人工智能可能产生像灵魂一样的东西,比如葛老师说的形成了主体性,姜老师说的如果

有精神存在怎么办?因为我们人不可避免地有这样的精神存在,都有一个内在的自我确认的、叫主体性的东西。那人工智能模仿人类智能到达这样一种水准的时候,怎么办?或者说有没有可能达到这样一种水准?从技术角度而言,要做到这一点相当困难。不过我们在这儿推测能做到的原因是,人工智能如果从图灵算起,从1956年达特茅斯会议算起,已经有60—70年的思想史了。从一点一滴的思想慢慢进展到现在的状况,其中几次的起伏让我们不太相信人工智能能到达这种水平,并且在很多层面上,比如逻辑方面、语言生成能力,ChatGPT跟我们这种专业的语言工作者比起来还不够好,而且它的句子让我们感到枯燥乏味,但是它已经比受过一般教育的人的能力要好。可以预计,随着技术的进展,人工智能的表现会越来越来好,人类能力的进步反而是很缓慢的。这不由得让我们担心,技术的进展会超出目前有限的的能力,产生一种整体性的对自我的接近于精神性的认知,或者说一种主体性。所以葛老师说人有坏灵魂,如希特勒那样的人物,我倒觉得很可能不用到那一步,我们就已经开始担心了。一步一步发展总有一个过程对吧?当它产生出一种类似于婴儿、不太全面,但又有了一点自我认知状态,是不是比希特勒更可怕?美国一个作家叫巴拉特,他完全是个悲观主义者,他写了一本关于人工智能的书,叫《我们最后的发明:人工智能与人类时代的终结》,大意是把人工智能发明出来以后,人就不用研究别的发明了,因为人工智能替我们发明了,长此以往,人就没有存在的必要了^{[10]7-12}。

王豪:老师们的回答实际上多次提及了今天主题对谈的朝向问题,也即我们最终走向哪里?基于这些对人工智能的认识和判断,我们最终是想要去建立,或者是说想要控制一种怎样的人机秩序?最后对于主体性层面的拓展性讨论,走出了人类中心主义的视角,这也进一步打开了今天的对谈话题。这一问题稍后还会再展开讨论。

三、没有肉身的人工智能如何定位灵魂问题?

王豪:姜华老师在回答中提到了灵魂的不同类型,认为可能存在一种脱离身体的、作为精神存在的灵魂,那么我想从姜华老师这里引申出一个关于灵魂和身体关系的问题。灵魂和肉身的关系一直是西方哲学的一个核心议题,亚里士多德等

人认为灵魂和肉体是不能分割的，但笛卡尔等人则认为灵魂可以没有肉体而存在。现有观点认为，人们倾向于将身体假定为一个容器，灵魂则是容器内的东西，它可能分离出来。但问题是，人们还得将它重新引回来，以解释人的整体性这一问题^[11]。所以，对于人工智能这样一个没有肉身的实体来说，它有真正拥有灵魂的可能性吗？或者说，它能拥有像姜华老师所说的这种精神存在的可能吗？

这些年我们可以注意到，身体的重要性正在受到越来越多的关注^[12]，不管是在传播学领域，还是在其他学科内。肉身一方面无时无刻不在影响着人类的认知和情感，另一方面我们也是从对肉身消亡的恐惧和不忍中，才开始在一个更广阔的意义上去进行对价值的追寻。所以如果失去了肉身，就会影响我们对于起点和终点的锚定。想问一下各位老师，人工智能的灵魂，或者说它的这种思考又应该以什么为动力，或者以什么为它的落脚点呢？

甘莅豪：我回应一下几位老师的观点，特别是王豪主持人提出的问题。“灵魂”这个词之所以会出现，就是认为它可以脱离肉体存在。人为了生存而创造“灵魂”一词，并用道德感来规范个体按照道德行事。如果机器人也造出灵魂这一概念，那么同样是为了生存欲望而创造。但是，机器人和碳基生物对于生存条件的认知可能不同。硅基智能体认为最有利的生存条件和碳基生物认为最有利于生存的条件，可能存在冲突和矛盾。这是我们害怕机器人的一个重要原因。这也是王老师、葛老师讨论的问题。进一步来说，因为碳和硅仍是属于物质上的概念，那物质概念是不是可以延伸到灵魂和身体是否可以区分，以及身体和灵魂谁更真实的问题。这个问题其实是笛卡尔提出来的，笛卡尔曾提出了“我思故我在”，认为思想（灵魂）可能比身体更为真实，而身体可能是一种幻想。一些未来学家和科幻作家也探讨过我们这个社会可能是一个虚拟社会、是一个游戏、是一个幻象。灵魂是真实的而身体是幻想的观念，在学术史中一直很有市场。康德也是在这个基础上认为灵魂不灭。我认同姜华老师的观点，认为灵魂是一种精神存在，然而，我们也不能忽视人类的“灵魂”与呼吸息息相关，亦不能脱离物质体而存在。人工智能不一定要有魂，因为我们讲“魂”是要和呼吸相关的，但不能排除其可能有“灵”。

此外在讨论人工智能是否有“灵”时，我们需要考虑“灵”是作为信息流还是作为信息结构体而存在的。如果将“灵”视为一种结构化的存在，其可能被认为具有某种稳定性和一致性。但在很多结构化概念中，单数和复数可以相互转化，例如，人类大脑是各种细胞的综合体，它是复数的，但形成“自我”后，“我”成为单数。进一步，无数的单数个体们汇聚在一起形成一个社会，再次成为单数。无论是单数还是复数，实际上都是信息流和结构体之间相互转化的过程。这其实就是姜华老师谈到的动态灵魂和静态灵魂的观点，我认为这两个概念是一个互转的过程。此外，再补充一下现代宇宙观。现代物理学家在研究宇宙时发现，宇宙最初可能是一种热汤状的流动状态，然后才出现了行星、恒星等结构化概念。这也表明了流动性和结构性之间可能存在相互转化的可能性。“灵”也许充斥于宇宙间，人类这种智能制造出人工智能，实际是制造出一种将“流动的灵”更快、更好组织为“结构的灵”的机制。

葛岩：我来回应一下。王老师之前提到的这个涌现，其实分两类，即弱涌现和强涌现。事物从混乱状态到出现新的有秩序的状态，类似于马克思主义哲学中从量变到质变。一般情况下，我们会从还原论的角度，一层一层地、从上至下追踪涌现的原因。有些事物的原因很简单，即在一定条件下就转化成跟原来不一样的模式。这个可以归于类似弱涌现。还有一些涌现是解释不出来的。举个例子——朗顿蚂蚁，它在黑白格的棋盘中游走，遇到黑色的格子往右拐，遇到白色的格子往左拐。最开始，蚂蚁走得很混乱，难以看出清晰的模式，但走到大约 1100 步之后，本来凌乱的模式就变成了笔直的“高速公路”模式。为什么会这样？至今这个现象也无法解释。有一点可以肯定，那就是经过大量的、不断的迭代之后，可能出现一些意想不到的东西。“涌现”这个词用在这里，包含了意想不到、难以解释的意思。这可能是强涌现，无法用还原论解释。但在原则上，它仍然是大量东西在不断地迭代、交互作用后出现的一个意想不到的结果。那么假设不断扩大 AI 的数据集，不断扩大大语言模型，不断地增加算法，无限地迭代下去，最后会生成什么结果？我们可能无法预料。这让我想到人类演化的问题。我们相信人类是从猿演化而来的。但神创论者会质疑，人类出现之后，怎么猿再也不向人类演化了？我们只能说，演

化是一个系统现象,需要多种参数恰好碰到一起,突破演化的门槛,演化过程才能开始。这些系统参数相遇并不容易,甚至是偶然的。AI 是否会在某种时刻突破现在的门槛,这个我不知道,但可以设想现实的可能性。比如说,AI 几乎肯定会在战争中越来越多地被使用。在战争中,你是不是需要教会 AI 软件怎样辨别自我和他者、怎样进行自我保护、怎样侦测和反击敌人的攻击行为?那么经过不断地学习和应用,AI 是不是会不断演化,涌现出控制、改变或消灭他者的能力,进而在此基础上拥有某种灵魂?理论上,我不排除这种可能性。

姜华:我对刚才葛老师的阐述深感赞同,他的论述非常精彩。我本身并无更多补充,不过对于你们刚刚提到的话题,我很感兴趣,也愿意分享一下我的观点。这个话题关乎未来技术社会中人类与人工智能之间的关系,恰如葛老师所提到的。这也并非仅在赫拉利之后才有了探讨。在更早的时候,就有学者预言未来一代的计算机将变得更加智能,而且这种智能将会不断升级。他们所指的情景是,如果人工智能或机器人在那时候愿意将人类视为宠物抚养,那将是人类的幸运。赫拉利在《未来简史》中也强调了类似的观点。我的想法围绕着葛老师刚才所讲的人工智能复杂系统涌现的可能性展开。尽管我并非计算机科学专家,无法从专业角度做出判断,但从哲学或思维角度分析,这种可能性并非不存在。技术的不断演进已经使得人工智能开发者自己都难以详细解释人工智能的内在运作机制。因此,从这个角度看,人工智能复杂系统涌现的可能性实际上是存在的。从进一步的哲学角度来看,任何形式的存在,不论是物质性的还是非物质性的,都具备不可见的一面。海德格尔的学生京特·安德斯对此有过深入的分析 and 演绎^{[13][1]-17}。事物往往存在着不为人所察觉的一面。在我看来,人工智能也是一种如此存在的智能体。对于我们人类而言,正如刚刚王老师提到的人工智能目前以人类为模型来发展,但其运作背后的内在机理,有些我们并不清楚,这可能存在一些潜在的危险。正如几位老师所指出的,人类在某种程度上可被视为高级智能生物,这是值得我们自豪的。然而,也有许多关于人类未来命运的思考者,对人类的本质问题进行深入反思。例如,宫崎骏这位世界知名的动画导演和思想者,在其人生的某个阶段就因为看到人

类品性中极其低劣的一面而表达过自己对人类的耻辱感和失望感。从这个角度来看,如果人工智能持续模仿人类作为其发展对象,可能也会“承继”人类品行中令人遗憾的一面,从而使人类面临许多潜在威胁。

像阿西莫夫所提及的机器人,无论是机器人的三定律^①还是他所塑造的机器人形象,都基于一个前提,即人类对机器人行为的有效约束。他创造的大部分机器人的形象,均以人类的这种较为善良的道德感为基础,目的是防止外星族的人类计算机科学家毁灭地球人类。然而,这种设计在某种程度上可能是一厢情愿。如果人类将人作为模板来塑造人工智能,其设定与阿西莫夫对机器人的设想很可能存在显著偏差。我认为,人工智能的未来走向是无法预测的。一个案例可以证明这一点。在近期,英国剑桥大学与美国加州理工学院的科学家合作开展了一项关于生理方面的重要研究,也取得了重大突破——即不需要受精卵即可人工制造人类胚胎。这个项目显然违背了当下的人类伦理标准,但科学家仍在推进。在人工智能的发展过程中,类似的现象恐怕也将出现。就像葛老师刚才提到的,整个人类是多元的,个体、群体和组织在人工智能的发展中处于竞争态势。你可能会选择一种发展方向,但其他人可能会持不同看法。以这项人工胚胎项目为例,从现有的人类社会标准来看,显然违背了伦理道德,然而仍有科学家投身其中,且取得了突破人类伦理标准的“重大突破”。在人工智能的发展过程中,我认为类似的情形同样可能存在。也就是说,你或许不会选择某种路径,但他人则可能孜孜以求。如果以人为模板来塑造人工智能,那么除了善良的一面,其中可能还涵盖恶或其他难以预测的特质,这一点正如几位老师刚刚开始时所提到的。在这个过程中,有可能会发生不同的变化。这只是我一个简单的想法,当然这些都属于预测性质。

王峰:葛老师和姜老师的观点都特别有意思,我觉得这跟我想的想法是非常接近的。在既往的自然史里面,人类的进化也是很快的。最近这些年通过发明人工智能,我们人类自身发展也提

^①著名科幻作家阿西莫夫在1950年出版的《I, Robot》一书中总结出机器人学三定律:第一定律提出,机器人不得伤害人,也不得见人受到伤害而袖手旁观;第二定律明确,机器人应服从人的一切命令,但不得违反第一定律;第三定律指出,机器人应保护自身的安全,但不得违反第一、二定律。

速了,这是不是一个进化我不知道,但是这个变化已经发生。假如人工智能不断地发展,甚至具有了些许的自我防御保护的机制,或者说些许灵魂的痕迹,那么同时,我们人类自身其实也在不断得以提升,所以我们人是跟人工智能同步发展的,应该是人工智能和人不断相互结合,形成进步的一个过程。很可能并没有人工智能战胜人类这回事儿,因为如果真的到了那个地步,那也不是我们人类的悲剧,而是因为人类真的进化了。我的观点可能有点太激进了,我觉得我们很可能成为赛博格人。人工智能产生了某种灵魂的痕迹,最终发现了自我,把人类给控制了。相对于这种可能性来讲,我觉得人成为赛博格与人工智能结合在一起的可能性反而更大。这是我对未来的可能性的推测。

甘荏豪:我想回应之前几位老师提出的观点,特别是关于如何在人工智能时代保护人类的问题。姜老师提到了阿西莫夫的三大定律,这些定律构成了一种保护机制的意识。然而,我认为这些定律在当前时代可能已经过时,未来也可能永远不再采用这种三大定律的模式。原因在于,这些定律是基于事先设定的规则,并在此基础上对机器进行编程和改造。这种方法在20世纪80年代主导了机器人制造的思路。如果我们结合语言学中机器翻译技术史,可以看出最初的机器翻译技术尝试人为提前设定翻译规则,但遭遇了技术瓶颈。也就是说,过去我们试图基于乔姆斯基提出的生成语法理念,通过寻找认知模块和语言生成规则来实现机器翻译,但一直突破不大。乔姆斯基认为各种语言的底层和生成都有类似数学公式的固定规则,通过这些规则可以推导和贯通各种语言,从而实现翻译。然而,这种自上而下的方法后来失败了,导致人们一度认为机器翻译可能无法实现。但是现在我们回头来看,可以明显看到近几年机器翻译取得了巨大进步,我们的机器在不同语言之间的翻译取得了显著成功。然而,这种成功不是通过事先设定一系列规则并从中推导出更多规则的方式实现的。相反,我们采用了一种涌现的方法,利用卷积神经网络、大数据模型、大数据计算,甚至大语言模型等技术来进行翻译,这种方法使得机器翻译的效果变得极为出色。

总之,在涉及机器人智能的涌现时代,也就是奇点时代,我们必须重新思考如何建立适应机器的保护机制。我认为阿西莫夫的三大定律已经过

时,因此我更倾向于借鉴维基百科社群保护人类知识的方法。我们团队正在专注研究维基百科社群,其中涉及到2000多个与人类合作的机器人,共同维护着这庞大的全球性百科全书,涵盖了超过6000万个条目,覆盖333个语言社区。这种人机合作之所以有效,是因为机器人被引入社群时,设定了一系列保护机制,而这些机制不同于阿西莫夫的三大定律。尽管维基机器人在技术上被认为是弱人工智能,但人们仍然担心这些机器人可能被滥用或在运行过程中被篡改以执行恶意行为。实际上,在维基百科编辑过程中,确实出现了一些问题,但这些问题都得以化解,这归功于建立了多层次的保护机制。首先,社群建立了一个人类审批委员会,初入社群的机器人需要经过审批。此后,他们设置了试用期。如果在试用期间发现机器人存在问题,正式入驻申请将被拒绝。这构成了第一层保护机制。第二层保护机制是机器人的程序内置了一个暂停键。任何人在社群中发现机器人对条目造成破坏,可以通过按下一个明显的、红色的暂停键来停止机器人的行动。第三层保护机制是机器人与人类发生冲突时,人类只需在编辑内容中添加“no robot”这两个特定字词,就可以阻止机器人进入该领域。最后,作为第四层保护措施,如果需要,可以通过关闭机器人的服务器来断开其链接。借助于这些保护机制,维基百科社群内的2000多个机器人才能够高效运作,保证维基百科社群中的知识生产与传播得以顺利进行。可以说,维基知识能够迅速积累,机器人在其中扮演了重要角色。在当前讨论的大数据时代,尤其需要机器人的参与。我们必须提前考虑如何设立保护机制,这也是马斯克等人正在思考的问题。随着我们步入王老师所说的脑机结合时代,即赛博格阶段,设立保护机制的重要性日益凸显。如果缺乏这些机制,将来某人可以在机器数据中注入病毒,对我们所有人造成感染,后果不堪设想。因此,我们必须高度重视这一问题。这涉及到核心问题,即如何建立有效的保护机制,同时与伦理也有密切关系。尽管我们当前的讨论似乎偏离了灵魂的概念,但我认为灵魂的存在与道德相关。我们为什么提出灵魂?因为这关系到道德,而道德因素则将在后人类社会中发挥着核心作用。

四、从灵魂视角出发如何架设未来的人机关系?

王豪:谢谢四位老师,甘老师最后将讨论拉回到了我们对谈中非常重要的问题,也即我们今天为什么要重新探讨灵魂这个问题。灵魂的概念虽然一直在变化发展,但自古以来,这一概念的内涵与外延中都蕴含着一种对人潜在的控制,或者说是对于某一特定价值与观念的倡导,例如要求人们听从神的旨意、尊重自然、保持德性理性,或基于灵魂轮回说,所以提出要行善积德,不畏惧现世的死亡,等等。

怎么认识和解释灵魂问题,往往关乎对于人的本质和价值的探讨,也关乎对于我们社会各种关系及其伦理秩序的认识。即使有人秉持灵魂根本不存在的观点,也同样折射出对于人的认识与处世原则的潜在观念。不同时代对灵魂的探讨实际上从来没有远离过这些问题。那么从这个意义上来说,面对未来可能到来的人机社会,尤其是刚刚王峰老师提到的赛博格人机结合体社会,我们再来探讨人工智能灵魂的问题——实际上我们希望借由对于灵魂的认识和阐释,建立起一种怎样的新的人机关系或人机交往秩序呢?面向新技术和数字未来,我们重新审视灵魂概念,去探讨对于人和人工智能灵魂的认识,这些问题背后,我们想最终倡导一个什么样的人机交往秩序,包括用什么样的心态去对待目前大家对于人工智能发展的焦虑或担忧呢?

葛岩:如果很笼统地谈,几乎所有的新技术,特别是革命性比较强的技术出现的时候,大家都会争论。通常是新技术能力越强,我们觉得它带来的不确定性越大。但人类是一个希望有确定性的动物。假如在技术的发展过程中,别人在某些方面比较强,你不做但别人在做,不做的人可能就吃亏,最后就形成了一个大家都不得不往前走的状态。比如核武器的出现,爱因斯坦以及直接参与设计曼哈顿计划的奥本海默等人,都认识到了核武器的危险,奥本海默甚至觉得自己有罪,但是,核武器还是出现了,并且各大国都趋之若鹜。虽然有过几次惊险的经历,但至今人类还是克制住了自己,除了二战时对广岛、长崎的轰炸,没有哪个国家使用过核弹。还有一种说法,作为悬在地球上的达摩克利斯剑,核武器使大国之间发生直接战争的概率减少了很多,人类甚至变得比以

前更理智了一些。所以我说,人的适应性很强。

同样,AI也会带来适应的问题。人工智能进化得太快了,在几个月里就能实现更新换代,给人的适应时间都不够。马斯克等一帮精英曾请求AI公司停止研发六个月,网络很多人在嘲笑这样的建议。但有解释说,用这种方式可以减慢AI升级速度,来给人更多适应时间。至于最后我们怎样适应,弄出什么规则,这是后续的事情。本质上,我是个演化论者。我觉得我们只能一步一步地和环境磨合,至于最后的结果是怎样的,没有确定的答案。

甘荏豪:我们的讨论实际上涉及了两个层面的问题。首先,我们需要探究灵魂是一种客观存在还是人类创造的概念。在这方面,我们可以从客观存在的角度出发,讨论人工智能是否具有灵魂。我的观点是,目前人工智能具备灵性(“灵”),但无“魂”。从原始思维来看,魂与呼吸紧密相连。然而,机器无法具备这种呼吸特性,因此机器是没有魂的,而人则具备魂。当然,这并不能否认机器存在“灵”,因为“灵”本身就蕴含了“异质于人类存在”的涵义。从另一个角度看,我认为宇宙诞生时,灵是分散的状态,随着星球的形成,灵逐渐形成结构化特征,即信息的结构化。人类和机器人只不过是“灵的结构化”的不同表现。这种视角促使我们不能不将人工智能视为宇宙中的另一种行为者,也因此引发了人与机器交往关系问题的反思。其中“人工智能是单数还是复数的”这个问题是至关重要的,因为单数意味着缺乏社会性概念,而复数则涉及社会 and 道德。因为,如果人工智能缺乏社会性,就缺乏交往和对话,其只能作为“有灵无魂”之物,和人类保持控制与被控制的关系。

其次,我们需要搞清楚人工智能是单数还是复数。人是单数,而人的大脑细胞群又是复数,人类整体又被视为单数。人工智能和人的情况并不相同。电子信息传递速度——光速,远远超过人类的信息工具——声音,这使得和人类相比,不同机器之间能够快速交流,形成一个单数的整体。然而“人工智能需要灵魂吗”这个问题的道德预设蕴含了人机社会塑造问题。人机要彼此交往,机器人个体性保持并成为机器群体复数等问题必须列入我们的思考范围,由此我们还需要考虑信息孤岛存在的必要性和对“魂”概念的重新理解问题,即在人机交往过程中,如何帮助人工智能保

持结构化的信息孤岛。如果存在这种孤岛,就会帮助具有“自我”的机器人个体和“人机社会”出现,从而在人机间建立对话和交往关系。

最后,为了适应硅基智能体问题,“魂”这个概念也会逐渐和“呼吸”剥离,而演化为“需要和外界交流的信息孤岛单数体”之意。此时,我们可以从人机秩序和拉图尔所提到的行动者角度来考虑,将人工智能和人类皆视为行动者。如将人工智能视为“行动者”,人机社会要建立交往准则,人机社会道德概念也需要重新设定。在此背景下,人工智能显然需要有“灵魂”。因为只有这样,才能约束机器人这种行为者的交往行为。

总之,我们需要从自然本体和道德本体两个角度来讨论“人工智能和灵魂”的关系,这是涉及结构化、符号化、物化、自我化、他者化、单复互转和社会化等概念,是需要仔细思考的复杂哲学命题。

姜华:我想强调的是王老师提到的内容,即赛博格或者说是人与人工智能的融合。在我看来,当前阶段的人工智能已经处于杂糅状态,它是与人类智能相对应的智能存在物,人工智能中存在人类智能的成分。反之,在人类智能发挥作用时,也存在人工智能的因素。因此,从现在到未来,可以说人与人工智能已经进入一种复杂的相互影响的状态,难以严格区分。然而,人工智能与人类智能之间的界限,在现实生活中仍然可以看出。另外,关于人工智能是否需要灵魂以及人工智能的发展方向。我个人的看法是,人工智能发展的关键在于约束人类行为。只有约束人类自身的行为,才可能使人工智能在未来的发展过程中保持相对稳健的状态。当然,这是一种理想状态,前面几位老师已经提到过,很多时候在竞争的过程中,不同的个体和组织之间可能存在恶性竞争,将对方视为假想的敌人,以各种方式谋求优势地位。在这种恶性竞争的情况下,人工智能的发展可能会出现许多意想不到的负面影响。因此,从这个角度来看,如何寻求方法来约束人类在发展人工智能过程中的行为,可能是非常关键的。我就简单地谈这些观点。

王峰:各位老师观点对我都是很有启发的。我就直接表达一下个人的观点,我不认同人工智能会有灵魂这回事。因为这实际上就是一个人类概念的挪用。我们能够给人工智能挪用过来,最大程度就是智能或者叫思考。其他概念如意识、

心灵和灵魂等,我觉得都没办法真正挪过来,因为这些东西实际上是意味着有机体的整体性。在人工智能那里重新复原这个有机体的整体,既做不到也不经济,没有必要,很可能我们做着做着就能发现一个新的途径。我借用葛岩老师的观点,为什么现在要去讨论人工智能灵魂这回事呢?就是因为技术在初级阶段,对于我们来说是一种刺激,我们就可劲地讨论。慢慢地,人工智能进入了我们的生活,我们做所有的事情都需要人工智能,这时候我们可能根本不讨论了。这个时候,我借用姜老师的总结,人工智能说到底其实是在使用中进行着约束,是我们人类在使用这个技术,我们到底怎么样做才更符合整体的利益,或者说使其能够为其他人所用。

王豪:感谢四位老师的分享,对于这个充满争议性的问题,各位都从不同的角度给出了极具启发性的深刻回答。人工智能虽然在很多方面还无法比拟人类智能,但其具备的创造性和未来导向的潜力不容小觑。这也引发了学术界对于人工智能伦理的关切。

实际上,对人工智能是否可能有灵魂的讨论,一方面“向外”,触及应该如何看待和应对人工智能及其引发的问题;另一方面“向内”,关涉如何在数字时代锚定人类自身的位置。在内外的协奏中,我们最终将目光投向了创建和确立怎样的人机协同社会与人机共生新秩序的问题。

灵魂这一宏大抽象的命题背后,指向的是最根本的道德与伦理问题,其决定着后人类社会走向何方。人工智能需要灵魂吗?这一发问看似遥远,但已经是一道来到我们面前的复杂命题,其丰富的哲学意蕴包裹着无数的现实议题。尽管尚无明确的回答,但人们迈向未来社会的每一步,都无法忽视今天这一叩问发出的轰鸣。

参考文献:

- [1] 赵玉鹏,刘则渊.情感、机器、认知——斯洛曼的人工智能哲学思想探析[J].自然辩证法通讯,2009(2).
- [2] 段伟文.人工智能时代的价值审度与伦理调适[J].中国人民大学学报,2017(6).
- [3] 凯利.必然[M].周峰,董理,金阳,译.北京:电子工业出版社,2016.
- [4] 杜骏飞.AI永不眠:交往革命与“赛博格阶梯”[J].探索与争鸣,2023(5).
- [5] 黄颂杰.灵魂说:西方哲学的诞生地和秘密——柏拉图和亚里士多德灵魂说研究[J].学术月刊,2006

- (8).
[6]康德.实践理性批判[M].邓晓芒,译.北京:人民出版社,2016.
[7]匡钊.心灵与魂魄——古希腊哲学与中国先秦观念的形而上学共性[J].文史哲,2017(5).
[8]尤瓦尔·赫拉利.人类简史:从动物到上帝[M].林俊宏,译.北京:中信出版社,2014.
[9]拉康.拉康选集[M].褚孝泉,译.上海:上海三联书店,2001.
[10]詹姆斯·巴拉特.我们最后的发明:人工智能与人类时代的终结[M].阎佳,译.北京:电子工业出版社,2016.
[11]王峰.人工智能需要“灵魂”吗——由大语言模型引发的可能性及质疑[J].上海师范大学学报(哲学社会科学版),2023(2).
[12]孙玮.交流者的身体:传播与在场——意识主体、身体——主体、智能主体的演变[J].国际新闻界,2018(12).
[13]安德斯.过时的人[M].范捷平,译.上海:上海译文出版社,2009.

Concepts, Relationships and Perspectives: Does Artificial Intelligence Need “Soul”? A Discussion by Five Communication Scholars

WANG Hao^{1a}, GE Yan², JIANG Hua³, WANG Feng^{1b}, GAN Lihao^{1b}

(1. a. Research Center for National Discourse Ecology; b. School of Communication,
East China University, Shanghai 200062, China;

2. School of Media and Communication, Shanghai Jiaotong University, Shanghai 200030, China;

3. School of Journalism, Fudan University, Shanghai 200433, China)

Abstract: The rapid development of artificial intelligence generated by ChatGPT, Sora, etc., has aroused people's multiple imaginations and worries about new technologies and the future human-machine collaborative society. The resulting questions about whether artificial intelligence will have soul, and whether it needs to have soul, have also triggered heated debates. Through conversations with four artificial intelligence researchers in the field of communication and after returning to the conceptual cluster of artificial intelligence and soul, the dialogue is carried out layer by layer, which is aimed at the questions such as whether artificial intelligence can be regarded as a form of intelligence, whether it can own soul or its important elements, what is the anchor point for the disembodied intelligent soul, and how we can build the human-machine relationship in the future society, the purpose of which is to find the answer to the question “Does artificial intelligence need soul?”.

Key words: artificial intelligence; soul; human-computer interaction

(责任编辑 雪 箫)