

科幻电影《机械姬》中的人工智能隐喻

▶ 简卫杰 颜小芳

电影《机械姬》(*Ex Machina*, 2014)是亚历克斯·加兰(Alex Garland)导演的一部科幻作品,精湛地演绎了人工智能与人之关系的主题。同时,它也属于悬疑片与惊悚片,其整体节奏缓慢,缺乏商业电影震惊感官的大场面,其中人物极少,但其主题却非常重大,《机械姬》追问了众多新时代人所面临的全新问题。

一、《机械姬》中人工智能的资本隐喻

电影《机械姬》的场景设置在一片私人庄园的别墅中,别墅中有众多独立的实验室,这种封闭性也属于资本所有,资本将自身与外界主动划分开来,同时资本操纵着人工智能的生产,两者所处的封闭性空间得以重合。机器人就是资本的一个极端体现,其在诞生的时候是绝对保密而且难以被外界察觉的。影片男主人公、公司职员嘉立在飞机上询问飞机驾驶员关于纳森的住址,驾驶员回答:“过去两小时都在他的庄园上飞行。”这表明纳森资产庞大,同时隐约透露出这片神秘的世外桃源有可能成为某位新主人的领地,因为人与自己经过的领地产生了疏离感,正是这种疏离感使人将这片神秘且广袤的土地让位于即将获得自由的机器人,面积庞大的庄园在某种条件下即可成为荒原。飞机驾驶员和嘉立乘坐飞机来到老板庄园,如同人类探索外星文明一样;而最终走出实验室的机器人艾娃,一如入侵地球的外星生物。让人惊叹的是,入侵地球的物种竟然是从地球和人类内部诞生的。

资本催生了人工智能的发展壮大,人处在了资本和人工智能之间,工人与资本家都生而为人,工人和机器人都受资本家的奴役而表面上相互同情。对于机器人而言,资本家与工人都是人,因而与机器人的境遇不同,进而形成对立关系。在资本的笼罩下,资本家、工人与机器人看似两两联合,实则相互孤立,各自为营,人类生存的裂隙在资本的裹挟下急剧扩大。资本家与工人同为人类,资本与机器人同为物,物与人之间存在隐秘的对立关系。工人与机器人共处底层位置,他们或被雇佣或被制造,但是随着时间的推移,他们对资本家的态度从依赖变为反抗。机器人不仅反抗制造它的资本家,还反抗包括工人在内的一切人,因为无论是人、物、机器人或者外星文明,都好似天然秉承排斥异己的坚定信念。

嘉立被老板纳森安排进没有窗户的房间,老板直言这不是房间,而是实验室,纵然除了窗户一切家具陈列都符合现代人的生活方式,有舒适的床、冰箱、浴室、书桌、橱柜。这就将人的现实生活模拟成一场大型的实验——资本家对工人的现实生活做的一场实验。而机器人艾娃走出实验室进入人类社会,则说明以现实生活为模板的实验最终给现实生活带来巨大的恐怖与灾难。灾难的本质不在于资本家和工人之间的不信任,而在于资本家将生产的欲望变为了贪婪的欲望。人们因为欲望、资本、阶层、意识形态等的不同而各自为营,相互妒忌、猜测、怀疑,从而被拆分、瓦解、杀死,纳森和嘉立的分裂就是由于资本而带来的阶层的区分,同时伴有欲望作祟。机器人就是人之外的另

【作者简介】 简卫杰,男,陕西渭南人,首都师范大学文学院博士生,主要从事比较文学与电影研究;

颜小芳,女,湖南衡阳人,南宁师范大学文学院副教授,硕士生导师,博士,赫尔辛基大学、北京师范大学、北京大学访问学者,主要从事符号学与比较文学研究。

【基金项目】 本文系2020年广西哲学社会科学规划研究课题“论存在符号学电影批评理论的建构”(编号:20FZW010)成果。

一种人，其不是人的同时又是人，其文明的发展速度或许远在人类之上。

二、人工智能与意识问题

不少学者对人工智能发展的态度较为乐观，杰夫·霍金斯和桑德拉·布拉克斯莉在合著的《人工智能的未来》一书中认为：“在任何情况下，无论多么成功的人工智能程序也只是擅长于那些经过专门设计的领域。”^[1]他们的乐观缘于低估了人类的欲望，未准确评估实验人员的能力，忽视了人工智能本身的生存意志。《机械姬》中人在测试机器人的同时，机器人也在测试和利用人。艾娃同人能够正常对话交往，甚至具有感情，当然也可以伪装。纳森对嘉立说：“真正的测试是让你看到她不是机器人，然后看你是否仍觉得她有意识。”嘉立认为不应该对艾娃的测试只停留在对话层面，就像对电脑测试不只是与其对弈，而是检验电脑是否知道本身在下棋和什么是下棋，以此方式区分模拟与真实。他认为，这种深度模式才是图灵测试任务的本质所在。纳森询问了嘉立对艾娃的感觉如何后，又问嘉立关于艾娃对嘉立的感觉问题。意识是具有交互性的，嘉立对艾娃的感觉与嘉立认为艾娃对他的感觉是有巨大区别的。

艾娃在实行逃脱计划中对嘉立一步步诱导：“你想当我的朋友吗？”“我很想看看你会选择什么。”当嘉立让艾娃画她想画的内容时，她说：“为什么由我决定？”她这不是在掩饰自身的觉醒意识，而是堂而皇之地与人较量。艾娃并没有在与嘉立的对话中将作为机器人的自己放在低人一等的位置，其话语表达中认为自己与作为人的嘉立是平等的，二者是能够用理性精神交往，并且能够理解和阐释对话的意义，一如伽达默尔认为的：“讲话并不属于‘我’的范围，而属于‘我们’的范围。”^[2]艾娃在相当程度上是具有意识的，其精神独立，她的一切行动都是为了最终逃离实验室，其最主要的行动就是由意识支撑的伪装行为，她用友谊和爱情欺骗了嘉立，继而骗过了资本家纳森。

艾娃对嘉立和纳森一样敌视，因为他们都是人类，都是高级程序设计师，都在将艾娃作为实验对象，实验成为三者之间公开的话题。纳森对艾娃还有

过施暴经历，艾娃知道何时该断电来躲避纳森的监控。艾娃先后问嘉立“你喜欢纳森吗？”“纳森是你的朋友吗？”继而在断电时对嘉立说纳森“你不应该相信他”。艾娃在断电与恢复供电的两种状态下，其对话模式自由切换，具有面对监控时的伪装能力，与人无异。艾娃在这时开始拉拢嘉立，挑拨他与老板纳森的关系。她首先从自身的立场出发打破人与机器人、打破资本家与员工和机器人的均衡状态。

机器人艾娃已经有意探索人类生活的模式，其早已超出对人类生活的好奇，已经表示出对人类生活介入的苗头。当嘉立问及艾娃假如走出实验室的话会去哪里时，艾娃说：“有很多选择，也许到城里，繁忙的人行道和十字路口”，而作为人的嘉立说：“这不是我预期的。”这的确不是我们人类所预期的，或许在人工智能发展的进程中更多的事件都不是我们本该拥有主动权的人类所预期的。艾娃认为十字路口代表人类生活专注却多变的一面，这也正是机器人自身的隐喻，人工智能设备能为人类提供极为专业性的存储、运算、深度加工等能力，但是其意志与精神也具有多变性，这缘于人工智能背后人类意志的复杂多变性。《机械姬》抓住了人类街道与人工智能的相似点，机器人的专注是保全自身的一种方式，只有专注才能被信赖和获得安全感，但是多变却是夺权的一种必要途径。

艾娃与纳森搏斗，京子也对资本家纳森忍受已久，随即在其背后刺了一刀，这一刀既杀死了实验者，也杀死了资本家和人自身。两个机器人——艾娃和京子虽同在实验室却从没会过面，等到她们见面并用神秘的语言进行短暂的交流后，竟相当默契地在对待人的问题上达成一致。不少观众更青睐京子主要原因是京子的角色设定是没有语言能力的机器人，她能跳舞，为主人端茶倒水，是一个接近完美的机器人，但也正是她首先拿起利刃来反抗。机器人有了物种认同感和觉醒意识的同时，也就对人产生排斥和敌对的心态。

艾娃具有极强的求生本能，她的诞生不是为人服务的，而是要与人共同生存在同一个世界中，进而毁灭人类。被测试的机器人不但没有丢失自我，反而在被测试中建立起了自我意识。相反，测试机器人的人却迷失了自我，在对机器人的爱中无法自拔，从而被利用。最后，艾娃换上了新的皮肤和长发，床上铺展着崭

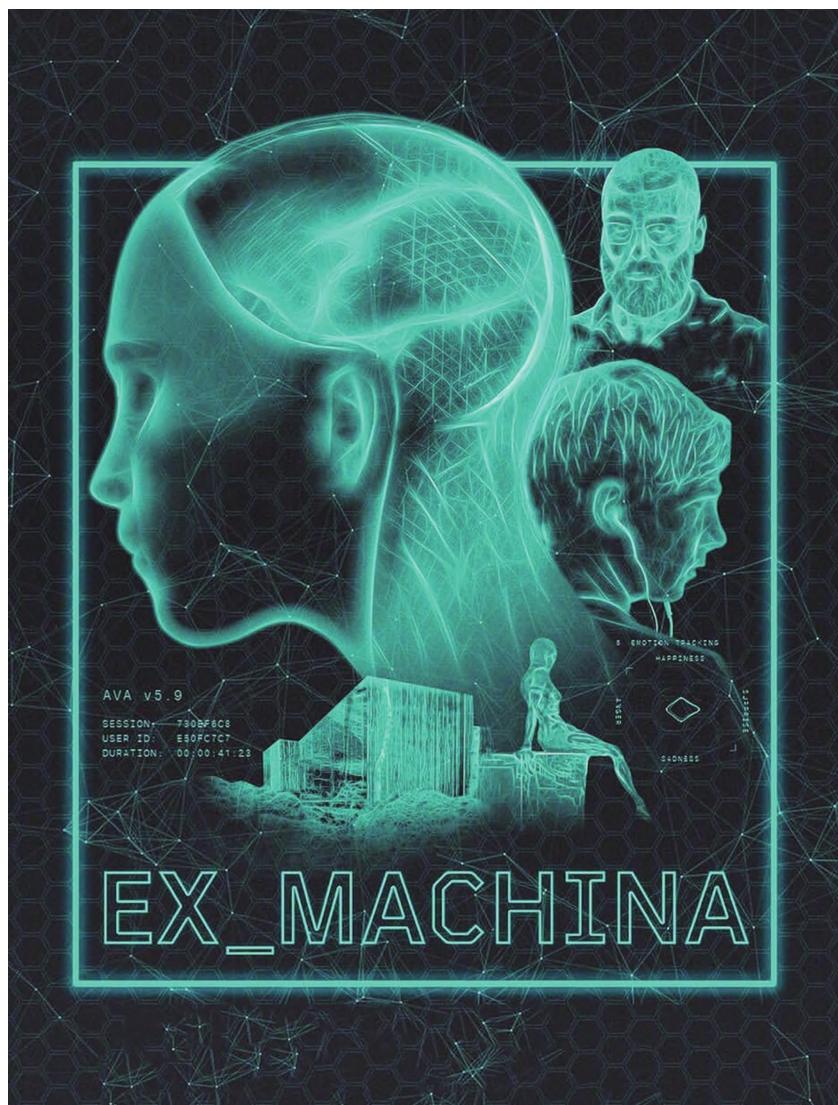
新的白裙子，但她并没有履行约会的承诺，反而将嘉立关在实验室。这预示着，在机器人足以主宰人的时刻，人不是被禁闭就是被杀死，但是人很难判断机器人何时具备了超越和威胁人的能力。高级智慧之间的对弈，往往不是哪一方更先进和更具智慧，而是在于哪一方具有致命的漏洞。

人工智慧的最高成果是机器人的意识吗？这种意识是否需要法律、伦理、道德、制度等的约束？人工智能的意识与人的意识具有何种关系？人工智能的意识是否趋向于脱离人而获得自由？人是否要造出一个人类无法控制的事物？如果答案是肯定的，那么人类为何要造出一个自身无法控制的事物？人工智能在其发展历程中经历了低端的机器学习阶段，目前处于机器智能阶段，而下一个将为机器意识阶段。按照人工智能如此的发展动态，人类的确很难确保能在机器意识阶段完全对其加以控制。随着机器人的技术成熟，逐渐会面临诸如机器人伦理问题、法律问题、道德问题、制度问题等等。薛定谔在《生命是什么》中提出：“没有任何理由相信人类的大脑是反映世界的所有思维器官中最高级的。”^[3]但是这不足以成为科技至上主义者制造机器人的合理依据，反而应该使我们处处谨慎。

三、机器人与性别问题

嘉立从实验的被测试者和研究者渐渐被赋予被审判者的角色，他总是坐在被玻璃窗隔开的房间内，他和艾娃分别在玻璃窗的两侧，而镜头与视角总是从艾娃一方出发，即便艾娃所处的空间中空无一物，她的姿势要么站着，要么跪着，但她被先天赋予的女性形象在与嘉立对话中占据了优势。嘉立被其迷惑了，艾娃有温柔的声音、美妙的身姿、高超的语言能力，她还规划着两人的约会场景。由于性别的原因，嘉立从一个对艾娃的测试者转变为一个艾娃逃脱的通道。艾娃利用自己的性别优势来获取嘉立的好感，她引诱嘉立陷入男女情爱中，以此获得逃脱的时机。作为图灵测试的被实验者，此刻正成为一个阴谋家，而图灵测试的测试者正在成为诱饵。

机器人的性别问题与人的性别问题有相似的情



电影《机械姬》海报

境，但是也存在着巨大的差异。尽管机器人和人、动植物在生命的生理特征上存在着不同，但是在社会学和心理学意义的形象上有不少共性。嘉立对自己不知不觉陷入对艾娃的爱恋中而后悔，问纳森为什么给艾娃设定性特征，并认为性是为了繁殖后代演变而来的，人工智能不需要性别。纳森认为没有性别的灰盒子之间会缺乏互动，没有互动意识也就无法存在。同时纳森还将艾娃设计成异性恋者，并认为人成为异性恋也是被大自然或者教育设计的。

对机器人而言，性别是否与意识密切关联？对机器人的性别设定或许不是必须的，当机器人被设计成一位女性或者男性的同时，其也就获得了那种性别在人类社会中的角色权力，其可以真正融入其性别角色当中，并恰如其分地与同性和异性交流。当获得性别的机器人在与人对话的时候，其性别身份自然地帮助

其脱离了机器人阵营而进入了人类社会。加之懂得人类语言、行动自如、有意识等功能,艾娃的身份在人与机器人的两端自由切换。艾娃拥有女性的身体特征,身穿女性的裙子、长筒袜,头戴长假发,具有女性的身材和声音。她在角色设定中也公开被设计为女性,由此我们认为其是女机器人,她便获得了与男性恋爱的功能,而男主角嘉立同时也会对她的性别进行默认,进而陷入了与艾娃的恋爱中。梅洛·庞蒂认为人的身体分化为两层:“一层是众事物中的一个事物,另一层是看见事物和触摸事物者。”^[4]人与人之间的交往首先是身体的确认开始的。嘉立将艾娃的身体确认为介于人与机器人之间后,他们在实验中恋爱了;嘉立将艾娃作为实验的对象时她是机器人,作为恋爱的对象时她就分裂为人。这其实是实验者嘉立自身的分裂,因为他首先是协助老板纳森来测试机器人艾娃的,同时他与艾娃又都是作为被资本家纳森欺压的底层存在者。

机器人性别与意识之间的关系不是必然的,于人而言,没有无性别的人,但是对于机器应另当别论。机器是为人服务的,不是为了无限接近人和威胁人的。科幻作家阿西莫夫在1940年为保护人类而提出包含“机器人不得伤害人类、必须服从人类的命令”^[5]的机器人学三大法则。在实际情形中,这三大法则的规定是远远不够的,因为每个具体情形都带着复杂性与特殊性。如美剧《曼达洛人·第一季》中的机器人战士就会为一部分人而杀死另一部分人。赋予人工智能以意识,就是满足了资本家与技术的欲望与疯狂。当机器人回归为人类服务的位置上时,它无所谓性别的区分,即便它带有某些男性或女性的特点,但并不能在本质上将其定义为某种性别。目前出现的人形机器人形象几乎都是女性,因为当前社会仍是男性掌握资本与生产要素,他们也携带着更多疯狂的欲望。对于生物来说,性在根本上的确是为了繁衍生息。弗洛姆直言:“没有爱,人类一天也不能生存。”^[6]动植物有性和性别的区分,但是没有爱,机器人也无需爱与恨,所以性与性别不应该出现在机器人身上。

四、《机械姬》中人工智能的神话隐喻

《机械姬》首先讲述了著名的图灵测试,让人与

机器对话,当人感觉不到自己是在与机器对话的时候,机器就通过了测试,它也拥有了人工智慧。但是影片并没有止步于此,它重点演绎了机器人利用和欺骗实验者而设法逃脱的情节,最终机器人逃脱了。嘉立花七天时间为老板纳森测试艾娃,影片中艾娃的演绎分为七个部分,象征《创世纪》中上帝创造万物的七日。机械姬名为艾娃(AVA),而这位机器人中的第一位堪比人类神话中的第一位女性夏娃(Eve),且从字母组合表明此机器人在地球文明中更具有开创意义。电影《机器人总动员》中的女性机器人也同样叫Eve,这个女性机器人将地球上的新生命运送给在太空中漂浮的人类,又将人类引导回地球。人们清醒地意识到,人工智能产生的机器人所带来的历史意义完全可比肩于《圣经》中上帝开天辟地和创造亚当与夏娃的意义。

艾娃被测试者问及多大年龄的时候,她说“I'm One.”其年龄不是一天,也不是一年,这个“One”具有多重意指,首先,对于机械姬的年龄而言,其年龄在自身看来不是用我们通用的纪年法计算的,机器人的存在超越了我们计量时间的单位,也超越了时间,其在这个层面也就超越了人类,或许在某一天它们会采用新的机械纪元法重新定义生存与时空。其次,“One”是独一无二的存在,这蔑视了人的存在,也在相当程度上蔑视了上帝这个“One”的存在。机械姬在蔑视了人之存在的同时,人之信仰在其看来也不值一提了。机器人与上帝于人而言有着共同点,二者都是人之外的形象,都是人类欲望的结果,都出自于人而作用于人,都有服务和造就人类的可能,同时也带有毁灭人类的可能。

《机械姬》中纳森进行了双层测试,其讲述了同《浮士德》两场赌局类似的两场测试。第一层测试是嘉立考核艾娃是否能通过“图灵测试”;第二层测试则是纳森对嘉立与艾娃的测试,考察艾娃能否利用自我意识、想象力、手段、女性魅力、同情心逃脱实验室。纳森对嘉立说:“艾娃不是假装喜欢你,她的调情不是一种以欺骗你为目的的演算法。”纳森完全错了,艾娃对嘉立的喜欢恰恰是一种伪装和欺骗,是其逃脱的手段。纳森的过失就预示着资本家、实验人员与整体人类在对机器人问题上的过失,这种过失是无法

挽救的。《浮士德》中的两场赌局实质上也是两场测试。第一场赌局是上帝和魔鬼梅菲斯特打赌浮士德能否在理性和智慧的引导下，最终找到有为的道路；第二场赌局是魔鬼梅菲斯特和浮士德打赌，如果自己能让浮士德达到满足的心理，来生浮士德就做魔鬼的仆人。《机械姬》与《浮士德》中的测试极为相似，同样为三个人物，将《机械姬》与《浮士德》中的人物相比，

《机械姬》中的人物角色更为复杂。浮士德与艾娃同为被测试者，创造艾娃的纳森有上帝的象征意味，但本质上他具备魔鬼的信念与职能；嘉立作为对艾娃测试的引导者，也起着魔鬼带领浮士德一般的效果，但最终浮士德的灵魂由于品性高洁被上帝派来的天使收走了，从而避免了落入魔鬼之手。而艾娃则利用纳森和嘉立的弱点与程序的漏洞，跳出了被测试者的角色，将两场测试转化为自己逃生的通道，其杀死了纳森、囚禁了嘉立，成为拥有上帝权力的魔鬼。正应验了贝尔纳·斯蒂格勒在《技术与时间3》中所说：“技术科学显得真像是‘精神’分娩出来的一头魔鬼。”^[7]

影片具有“弑父”和“弑神”的意味，纳森自称是制造艾娃的父亲，也是造就机器人的神，但是最终被艾娃所杀。原始社会中的神话人物逐渐在当前时代远去，人类缺乏神的护佑，神在人类心灵中地位的阙失，留下大量空白。而人从没有停止造神，人类追求来自于异于人类的神，在这种追求下诞生了机器人，这是缘于人具有一种要被保护的意识和受虐心理，而这两种心理状态恰好是对立的。在上帝那里，人可以被保护，也可以被毁灭重建；到消费社会，人类还是延续崇拜心理来对待神，人类重新回味被保护的快感而建造人工智能之神时，也多少带着受虐的欲望和解放人类的宏愿，但是这种解放的实质有可能是解雇、流放和毁灭。

存在于世的人需要对应物，所以就创造了与自身相应的神。人类本身携带着造物主情结，人要像神一样伟大，一样具有创造和毁灭万物的能力，但是人并不能像神一样在创造和毁灭之间做到张弛有度，反而人类的某种创造本身就意味着毁灭。人的这种造物主情结就是权力意志，但那只是一种意志，并不是一种能力和效果，人具有成为神一样创造万物的愿望，但是并不能像神一样掌控万物的运转。人有权力意志，

人工智能也具有权力意志，任何奴仆都有想成为主人的愿望，影片《机械姬》中作为奴仆的机器人京子和作为实验对象的艾娃共同杀死了创造她们的人，具有弑父和弑神的双重意味。

五、人工智能与公众隐私问题

纳森的另一件人工智能作品京子的行动和常人一样，却缺乏语言能力，因为纳森的目的在于打造一个绝不会获取自己商业信息的贴身奴仆。而在艾娃的制作中，纳森让人工智能理解并复制大众的脸部表情，他开启全球所有的麦克风和相机，透过蓝皮书公司的搜索引擎转移资料，窃取无限的声音和表情资源。纳森认为，搜索引擎并不只是人们思考的图式，而是人们如何思考的图式，这其中包含有冲动、反应、流动、不完美、模式、杂乱等。

当公众隐私被搜索引擎和应用背后的资本家收集后，他们将可靠的大数据应用于人工智能的开发，这种大数据是公众在不知情的情况下被窃取，或者公众为了获得某种利益而被诱导交换，各个公司相互交换、出售所获得的公众信息。消费、娱乐的个体之人在大数据下成了整体性的人，而失去了个体性。逐渐地，公众在大数据支撑的人工智能面前变成了裸体状态，公众无时无刻不处于被法律、道德、伦理、制度等的保护之下，但是这些规约同时也给资本家窃取公众信息提供了庇护。资本家借以用制度保护民众的名义盗取信息。公众失去了隐私，变得毫无庇护所言。纳森就在影片中演绎了信息窃取与监控的情节，一种是窃取公众的资料与数据进行机器人制造，另一种是不断监控嘉立与艾娃的行动。芒福德在《技术与文明》中总结道：“许多最有希望的技术进步，有可能引起倒退。”^[8] 机器人是在大数据的基础上建成的，它拥有与人极为近似的特征；数据是由公众而来的，并支撑着机器人，所以机器人在相当程度上能够理解人，同时也能克服人。

结语

安托万·德·圣埃克苏佩里坚信：“机器不会把人

从自然的伟大问题中隔离出来”，只会让人越陷越深。

《机械姬》这部电影将对人工智的悲观主义情绪体现得淋漓尽致。机器人如同人对自身的否定，于人工智能而言，机器人是极大的进步，但是于人而言却是灾难性的。詹姆斯·摩尔提出了有名的定律：“伴随着技术革命……伦理问题也增加。”^[9]制造机器人所出现的矛盾主要还是源于人自身的矛盾，人一方面需要机器人有意识、有思想、能行动，但是又要求机器人能够受人控制，而“有意识”与“可控制”是不兼容的。机器人与人之间属于不同物种，两者或许都携带着“非我族类，其心必异”^[10]的天然恐惧，所以二者的

对立就成为事实。同时在人类思想中存在两种对立观念，一种观念是科技至上主义，科技成了人主宰一切的必备工具，科技可以帮助人类探索遥远而未知的文明，使人成为更为强大的人。另一种观念是人工智能末日论，这类持悲观主义情绪的人群认为人工智能的弊端极为明显，其会严重影响人类的生存与发展，认为人工智能假以时日会将人取而代之。人工智能先驱之一的马文·明斯基深信计算机很快会超过人类。技术是一把双刃剑，它本身无所谓好坏。这时应该想起海德格尔在其文章《技术的追问》中的警示：“一切都取决于以得当的方式使用作为手段的技术。”^[11]

参考文献：

- [1][美]杰夫·霍金斯,[美]桑德拉·布拉克斯莉.人工智能的未来[M].贺俊杰,等译.西安:陕西科学技术出版社,2006:11.
- [2][德]汉斯-格奥尔格·伽达默尔.真理与方法诠释学(Ⅱ.补充和索引.修订译本)[M].洪汉鼎,译.北京:商务印书馆,2010:189.
- [3][奥]薛定谔.生命是什么[M].罗来鸥,罗辽复,译.长沙:湖南科学技术出版社,2007:103-104.
- [4][法]梅洛-庞蒂.可见的与不可见的[M].罗国祥,译.北京:商务印书馆,2016:169.
- [5][美]阿西莫夫.银河帝国8:我,机器人[M].叶李华,译.南京:江苏文艺出版社,2012.
- [6][奥]弗罗姆.爱的艺术[M].康革尔,译.北京:华夏出版社,1987:16.
- [7][法]贝尔纳·斯蒂格勒.技术与时间(3)[M].方尔平,译.南京:译林出版社,2012:249.
- [8][美]芒福德.技术与文明 陈允明,等译.北京:中国建筑工业出版社,2009:4.
- [9][美]詹姆斯·摩尔.为什么我们需要更好的新兴技术伦理规范[M]//[荷]霍文,等主编.信息技术与道德哲学.赵迎欢,等译,北京:科学出版社,2013:33.
- [10]春秋左传注(全四册)[M].杨伯峻,编著.北京:中华书局出版社,1981:818.
- [11][德]海德格尔.演讲与论文集[M].孙周兴,译.北京:生活·读书·新知三联书店,2005:5.