

人工智能时代的人文主义

韩水法

摘要：人工智能和基因编辑等科学和技术在当代社会的迅猛发展并日益深入地普及到人类社会的各个层面，不仅对人类理智能力及其独尊地位造成了巨大的挑战，而且亦导致了人的性质的变化。流行的意见面对人工智能持续提升的紧迫压力和人的性质变化的冲击，采取了分裂的和矛盾的立场，既否定人的变化的必然性，又对人工智能的可能前景束手无策而怀悲观失败的心态。这种挑战和现象迫使人们重新思考关涉人的哲学基本问题，亦再度激发人文主义的核心关切。今天的人文主义强调人类在认识、发明和制造人工智能和其他技术并不断促进它们升级的同时，也持续地增进和提升了自已的理智能力。从现在起，人类必定要有意识地见证和促进自身的进化。此种发展和进化在今天业经呈现出文化的与物理的统一趋势，这正是人们应对充满各种挑战的人类前景而坚持积极观点的根据所在。

关键词：人工智能 人文主义 人的性质 进化

作者韩水法，北京大学哲学系教授（北京 100871）。

今天，科学发现和技术发明正在以空前的速度、广度和深度改变人类社会，而人工智能和基因工程等技术和应用则直接引起人类本身的某些变化。这些科学和技术的进展一方面使得人类获得了深刻认识、利用和改造自然的强大工具，亦获得了深刻认识人类自身、发挥其潜力和能力并改善自身的有效手段，尤其在增进人类健康和长寿、提升人类生活质量等方面展现了良好的前景；另一方面，人工智能的进一步升级也可能产生在智力等方面全面凌驾于人类的超级智能，从而使人类平添了被支配、奴役甚至消灭的风险。

本文的中心任务就在于揭示这个核心问题，即人的性质的变化，在人文主义视野之下研究它的过去、现在和前景。围绕这个核心问题，笔者将具体分析它之所以被忽略的原因，人们在认识和应对这种变化时形成了哪些盲点。在这个基础之上，笔者将提出人工智能时代人文主义的基本理念，以论证这个时代人类自尊和自信的必要性及其根据。这个核心问题直接关涉哲学基础理论，于是，本文的讨论和研究将在哲学的境域中展开，并且必然亦要关涉和重新思考既有的形而上学、认识论和

心灵哲学等的理论意义，重新研究正义和道德原则的基础和有效范围。从方法论上来说，人类解决目前所面临的重大问题的态度应是面向未来而做出现在的筹划。当下的问题，包括过去累积下来的问题，都只有在未来才能得到有效的措置。那些以过去的传统和观念来限定今天和未来人类活动的主张，从根本上说是不可行的。

一、人工智能时代的特征

对什么是人工智能，人们尚未取得一致的意见，而对什么是人工智能时代，人们的回答更是形形色色。不过，无论分歧如何悬殊，近十几年的经验事实表明，人工智能在弈棋、翻译、无人驾驶、疾病诊断和图像与动态识别等人类智能高级应用领域迅速拓展。它在帮助人类解决各类复杂问题并满足其现实需要上所展现出来的强大功能，以及它的发展速度和运用于各种领域的拓展潜力，超出了多数人的认识和理解之外。为此，笔者要从探讨人工智能的定义并考察人工智能时代的特征入手，以便准确地把握和刻画这个时代的性质、特点和趋势，以及它如何造成了人的性质的变化和演进。

笔者之所以使用人工智能时代这个名称，乃是以人工智能为这个时代的标志，而非认为这个时代的特征仅限于人工智能。基因科学和技术同样是这个时代的主要特征，人们一般不会使用基因工程时代这样的说法，或许相对而言它的影响和意义没有人工智能那么普遍。但是，要理解这个时代，就必须理解基因科学和技术及其实践意义。此外，我们还必须检视与它们一起构成了这个时代基本特征的那些新兴而主流的现象的现实影响，了解这些现象的知识境域，领会其意义。

首先，人类知识达到了惊人的深度和广度。从宏观上来说，人们关于物理世界的认识尺度，在时间上，达到 137.98 ± 0.37 亿光年；在空间上，人类可观察到的宇宙距离为 46 亿光年。在微观方面，人们已经认识了两类基本粒子。在宇宙规则方面，人们已经认识了宇宙间四种基本作用以及它们各自的作用范围。

其次，信息及其网络成为人类社会基本结构的一个部分。信息社会与人工智能都基于计算机技术而发展起来，一般而言，人工智能乃是信息技术发展的必然结果。^① 在信息社会中，巨量信息的利用和处理需借助人脑智能之外更高效的手段，而人工智能的滥觞确实出于这样的目的和需要。只有具备了有效的处理手段，信息的生产、分配、整合和综合应用才是可能的。正是在这个意义上，信息社会和人工

^① 关于信息社会的定义，韦伯斯特（Frank Webster）在其权威性的著作《信息社会理论》中分析和批评了五种关于信息社会的定义，即分别出于技术的、经济的、职业的、空间的和文化的标准所形成的定义，其结论是对信息社会这个概念的质疑和反思。参见 Frank Webster, *Theories of the Information Society*, London and New York: Routledge, 2014, pp. 7, 10.

智能相互依赖，而人工智能无疑是信息社会的中枢部分。

再次，自 1953 年遗传物质 DNA 分子双螺旋结构被发现以来，生命科学已把关于生命活动及其规律的研究建立在分子水平，从此之后，人类对生命的认识完成了一个飞跃，因而克隆生命、改变生命的遗传性质等操作在一定条件下成了现实。这就意谓，人类在一定程度上可以更新生命的性质甚至创造某种新的生命类型。在神经科学领域，光遗传学工具包已经能够精确操作神经元的兴奋性。

最后，各种高速交通工具和设施使得人类不同个体和群体之间的交往变得方便和快捷，彼此之间的联系越来越紧密，从而人类的政治、社会和经济等关系相应地也发生了改变。虽然地球村的概念最早是指人们之间通过电信通讯等手段所建立的紧密关系，但高速交通工具才切实地改变了人们现实的地域空间观念，消除了距离阻隔，而互联网又真正实现了地球上不同地区人们之间的即时联系、交流和合作。

前述四个方面标志了人工智能时代的主要特征和现象。它们表明，不仅人类知识的性质发生了重大的改变，人类的生存方式、手段和环境同样发生了关键的乃至革命性的变化，而这些就会在一定程度上导致人类群体性质的变化，而且人类生物的性质也开始经历某种变化。在所有这些现象中，人工智能又最为特殊，它的出现和未来发展可能直接导致某种另类智能的出现，即所谓的强人工智能。下面，笔者先来集中讨论三种将直接改变人的性质和地位的现象。

从字面上来理解，人工智能就是人类自己发明并制造的理智能力。埃特尔在考察了若干人工智能的定义之后，认为赖斯的定义最为赅简并解决了其他定义的矛盾：“人工智能就是关于如何使计算机去做现在由人做得更好的事情的研究。”^①然而，这仅仅是一个描述性的定义，回避了实质的规定。关于人工智能之人工的意义，人们的意见并没有多大不同，而关于其智能的意义则呈现相当大的分歧。这种分歧可以概括为如下两个争执点：第一，强人工智能与弱人工智能；第二，类人的人工智能与非类人的人工智能。

强人工智能，在学术文献中通常称为人工一般智能（Artificial General Intelligence，简称 AGI）^② 或超级智能，乃是指一种能够自主地学习、行动、设定目标和解决问题的智能，亦即它是一种能够成功地从事人类任何智力工作的机器智能。在这个意

^① Wolfgang Ertel, *Introduction to Artificial Intelligence*, London: Springer, 2017, p. 2. 埃特尔认为，这个定义刻画了过去 50 年人工智能研究者所做的事情，而且直到 2050 年这个定义都是时新的。

^② Artificial General Intelligence 之所以在本文中译为“人工一般智能”，而不是通常的“通用人工智能”，其原因盖在于“通用”一词具有强烈的被动性，而 AGI 已经等同于人类智能，因此具有同等的主动性。“一般”则强调此一智能乃是与人类智能的对等存在。此外，“人工一般智能”的译法也照顾到与英语的对应。

义上，它等于一种人工自为者（agent），^①与人类自为者相类似。相应地，弱人工智能就是非自主的智能，不是自为者，而是人类智能的替力或助手。在一般的理解中，它只能执行人类设定的目的并听命于人类而行动。它能够代替人类处理许多人类智能难以完成的计算、学习、搜索、感知和管理等方面的工作与任务，只是这些工作与任务都由人设计和指定，行动的指令亦由人发出，并且随时根据人的意志而中止。

人类智能是以生命为基础的高级理性能力，因此，诸如生命的自我维持所必需的新陈代谢、自我复制（繁殖）和成长，乃是人类自为者的自然基础和条件。那么，作为自为者，人工一般智能是否也应当具有同样的功能？否则，作为一种物质系统，它存在的根据和条件始终受人控制和供给，也就难以真正成为自为者。现在的主流观点认为，在当下很难设想那种兼具自主目的、情感、联想和想象力的人工智能如何能够产生出来，因此，人工自为者在可预见的未来暂时不会出现。当下迅速崛起和进步的人工智能主流事实上属于弱人工智能，它们在专域（domain-specific）的能力远胜人类智能。这些能力现在还在快速地提升和拓展。除了技术任务和弈棋，它们现在也能谱曲、绘画、写诗，进入了人类自以为傲的艺术领域。而艺术的价值判断就又牵涉情感，加之弱人工智能看来在逐渐发展出综合能力，这就使另一些人坚持人工一般智能一定会出现，至少与人类水平相当的人工自为者将出现。^②

如果人工自为者属于类人智能，那么迄今有关人的一切理论，尤其是哲学的基础理论就受到根本性的挑战。第一，人类除了自己的理智，还没有见识过任何其他类型的智能或理智，现在这种智能已不再是抽象的可能性，而是具备了若干现实基础的可能性。第二，除了人类智能和人工一般智能，在这个宇宙中，可能还存在着其他类型的智能。这样的观点亦具有相当有力的理论根据。因此，对当代人来说，与人类智能对等的同类的存在不再是科学幻想，而是现代的科学理论和技术努力的一个方向，并且这种努力看起来愈益逼近其现实性。哲学在这里所面临的挑战和任务就是，在这样一个背景之下，先前一切有关人性和理性的理论都要予以重新的审视，而对其他可能的理智能力则要从头理解和认识。

迄今为止，哲学尚未对现在的人是什么这个问题达成共识，而人的性质却又开始经历重大变化。仅就此而论，哲学就需要从头回答如下的三重问题：第一，继续回答人是什么；第二，从理论上回应性质变化之后的人是什么；第三，解释性质变化前后的人之间的关系。

① agent 在中文文献中常被译为“代理”，这显然不够准确。此词的一个义项乃是指具有自主性并发挥某种特定作用的个人或类似物。它被用来强调人工一般智能自主地发挥特别作用的性质，亦即它是一种自主地活动和发挥作用的实体。据此，我将它译为“人工自为者”。

② 参见 Roman V. Yampolskiy, *Artificial Superintelligence: A Futuristic Approach*, Boca Raton, London and New York: CRC Press Taylor & Francis Group, 2016, p. 107.

二、传统人文主义的核心

人文主义是人类自我反省、确认和定向的思潮，在历史上，它形成了许多流派，包含诸多不同的倾向和理论，尽管如此，在这个复合的大潮流之中，始终存在着引导人文主义进步和上升的主流。笔者将撇开枝节而直达本质，从中西各种流派中概括出人文主义的核心主题。

康德关于人的四个哲学追问，即我能认识什么、我应当做什么、我能期望什么、和人是什么，^①不仅在哲学基础理论研究中依然是有效的纲要，同样也可用作追问迄今为止的人文主义诸流派共同核心的指针。这四个追问既有实质的内容，亦承带方法论的意义。在中国传统中，无论依据其字面意义还是其内在意义，人文始终包含变化和进步的意思。^②在西方不同的历史阶段和社会形态中，人文主义各有其侧重，并且经历了起伏的演变。但是，无论如何，人文主义之人文就是人创造出来的人的存在的新形象，以及对这个创造过程、形象的理解和说明，这一点与中国传统的人文主义是共通的。

人的创造和发展的前提就是确证人在这个世界上的地位和价值，以及生存和发展的根据和意义。在人文主义视野中，人的存在和发展始终具有当下即是特征，始终与现实环境处于互动之中。人的发展不仅出于人的自主要求，也受到自然环境以及人自身所造就的环境的逼迫和促进，而在现代社会，后一种环境所起的促进作用甚至比肩于前一种。人们容易忽略一个重要的事实：在人的活动所造就的对象——包括环境——面前，人其实也作为对象而存在和活动，而无论这样的关系是主动地造就还是被动地发生的。在这个意义上，在理解人文主义时，人们就需要增加人作为自己活动对象之对象这样一个维度，诚然，这个维度同时就指明了人文主义所受到的一种特殊的限制和激励。

不仅如此，这个判断还蕴涵了更深一层的、也尚未被清楚地揭示出来的意义，即人的一切活动本身就是人文主义的。在先前，多数人倾向于将人文主义限制在狭隘的意义上，将它与生产、科学和技术活动等分离。譬如，在20世纪上半叶之前，人们在论述人文主义时，无论多么遥远都会追溯并关联到古希腊罗马文化、精神及其经典作品。这样的人文主义被萨顿称为文学式的人文主义。^③自20世纪中期之后，随着自然科学和技术与传统的人文学科乃至社会科学的进一步分离，并且人文

① 康德：《逻辑学讲义》，许景行译，北京：商务印书馆，1991年，第15页。

② “刚柔交错，天文也。文明以止，人文也。观乎天文以察时变，观乎人文以化成天下。”语出《周易·贲·彖》，参见高亨：《周易大传今注》，济南：齐鲁书社，1998年，第226—227页。

③ 萨顿：《科学史和新人文主义》，陈恒六等译，北京：华夏出版社，1989年，第18页。

学科与社会科学也进一步分离，除了持续地回溯西方古典文化，文学式人文主义越来越深地沉浸于语言和概念的思辨之中，而这些概念从基督教上帝、阶级压迫和解放到弗洛伊德的本能学说，无不用来解释仿佛越来越令人不满的现代世界。在这种精神氛围之下，人文主义看起来只有越来越狭隘的出路：或者回归到那个未被异化的因而理想的人的存在，或者在永远不知希望在何处的自我之中打转。^①以存在主义、后现代主义或其他标志出现的人文主义也陷入同样的困境。就如文艺复兴时期的人文主义同时混合了宿命论、迷信和其他一些消极因素一样，20世纪文学式人文主义亦明显地表现出了某些虚无主义和无所作为的特征。布洛克认为，自20世纪上半叶起人文主义所经历的主要冲突乃是体现了人文主义的情感或非理性的因素与理性的冲突。^②事实上，自20世纪起，如下一种消极情绪渐渐地迷漫于人文主义之中：人不仅受到了制度的压迫和限制，还受到了自然科学和技术的压迫和威胁。与此相应，如下一类倾向和态度在今天依然有不小影响：反对进步、保持传统的风俗、习惯，甚至复辟古代社会的制度和习俗，甚至迷信。

因此，立足于人工智能时代的重新考察和反思，就是要揭示和发现为那些狭义的人文主义者所忽略、轻视乃至无视的人文主义原本的重要因素，即贯彻在现代人活动的各个方面的人文主义精神和性质。正是在当代境域之下，同时面向人类的未来时代，人文主义才能得到真切的理解和规定。

现在，人类演化已经进入了一个根本的关节点。迄今为止的人文主义都以既有的人——从智人出现以来生物性质未经历过改变的人——为对象，这样的人文主义也可以称为智人人文主义。而现在，现代智人的生物性质正在面临改变，而本文所要分析和讨论的正是这样视域中的人文主义。

三、人的性质的变化

在人工智能时代已经发生的人的性质的变化，以及那些预示其未来将经受的更大演变的种种端倪都表明，这个变化将是自从智人出现以来的人类所面临的一场空前的大转折。在下文中，笔者将要论证，人正是在进化中形成其性质，而目前所面临的变化无非是人类进化的持续。我也将阐述这个变化的一些新特征，它与哲学基础理论的关系和它所蕴涵的哲学意义。

在20世纪下半叶，敏锐的思想家已经预感到了即将来临的变化，但是他们对这个变化将发生在什么领域、变化的程度和性质以及导致的可能的结果并没有清楚的

① 参见萨特：《存在主义是一种人道主义》，周煦良等译，上海：上海译文出版社，1988年，第30页。

② 布洛克：《西方人文主义传统》，董乐山译，北京：三联书店，1997年，第271页。

意识。一些重要的思想家正执着地反思和批判自然科学和技术带来的对人的控制和压迫，这造成了人文主义思潮中对自然科学和技术的浓厚的怀疑和消极态度，甚至单纯的抗拒心理。对即将到来的变化，多数人在认识上和心理上没有做好准备。在围绕异军突起的基因工程的讨论中，这种态度就比较集中和激烈地表达了出来。

约纳斯反对一切超出预防遗传疾病的涉人生物技术，在他看来，那种操作使得人成为人的制造者，但是，他认为，人不能是造物主。^①然而，约纳斯不清楚的是：基因工程只是在人既有的生物基础上更改了人的某些生物状况。至少到目前为止，它以当下的人为中心。不过，约纳斯认为，他提出了属于真正哲学的——确实也不乏深刻的——问题：“涉及生命与死亡的意义，涉及人格尊严、人的完美形象（在宗教上是：‘上帝的形象’）。”^②显然，他把人或人类的形象及其尊严，甚至上帝的形象，都定格在他那个时代的对人的直观之上。但是，他的这种直观和论断其实都是非批判的。

哈贝马斯就此的表态比约纳斯要晚，但在态度上更为保守。他反对自由优生学（liberal eugenics）和植入前基因诊断（PGD）。他依据交往行动理论提出了多重的反对理由，最为核心的是：这种做法侵害了另一个人即被筛选而出生的孩子作为一个自主行动者的自我理解的先决条件，而这个过程又是不可修复的。^③换言之，那个为优生而被筛选之后出生的孩子受到了他人的决定，因此，他的一生都是不自主的。哈贝马斯和约纳斯的观点与欧洲议会的态度相似，那份被德沃金批评为“草率的反应”的《克隆决议》，反对对人的任何克隆，包括为生育治疗、器官移植或植入前诊断目的的克隆，认为它们之所以是不可接受的，乃是因为它们违反了人的尊严，严重侵犯基本人权。^④这个决议显然援用了启蒙运动所奠定的人的至上地位的原则。类似的理由有很多，但是，它们的说服力都不及人的尊严和权利。

综合来看，那些质疑和反对人工智能时代的主要技术及其应用的主要意见和理由，从根本上来说，出于如下的深切担忧和恐惧：人的性质将因此而改变。这些质疑、反对、担忧和恐惧亦呈现了这个时代的人文主义所面临的境域、问题和困境。

首先，人们确实预感或预见到了许多重要的变化，而且所有的变化都显示出了不可避免的趋势。其次，这些变化的长远发展及其后果充满了不确定性；人们既期待它带来的积极成果，又担忧它们可能的危害。迄今为止，人们曾经预言的那些可

① 汉斯·约纳斯：《技术、医学与伦理学：责任原理的实践》，张荣译，上海：上海译文出版社，2008年，第177页。

② 汉斯·约纳斯：《技术、医学与伦理学：责任原理的实践》，第4页。

③ Juergen Habermas, *The Future of Human Nature*, Cambridge: Polity Press, 2003, p. 63.

④ Ronald Dworkin, *Sovereign Virtue*, Cambridge, Massachusetts: Harvard University Press, 2000, pp. 438-439.

能的危害尚未出现，但是一直被认为有可能出现。这种不确定性却正是最为令人担忧乃至恐惧的东西。

对这样的怀疑和恐惧还可予以如下的解释。(1) 人类作为整体处于一种两难境地：人类一向的好奇性、追求新鲜的欲望和竞争性，以及永不停息的创造性，使人类不断地从熟悉而确定的状态进入陌生而新奇的状态，而这个过程必定承带的不确定性正是人们恐惧的主要原因。(2) 以对生殖细胞的基因编辑为例，因为这种修改将通过生殖而遗传到后来世代，因此，它就成为所有有关基因工程的争论的焦点。人们并不担忧个体非遗传性质的改变，但担心生殖细胞的基因编辑，因为它将长久地改变人的类性质。然而，生殖细胞基因编辑的科学研究和技术发明却是当代生命科学研究的热点。于是，人类的探索欲和对更好未来的欲求与对未来不确定性和可能的危害的担忧，这一对人类向来具有、难以消解的对立的理性力量和情感一方面不断推动人类社会的前进，另一方面又造成了许多的障碍和制约。

(一) 演变中的人的性质

根据现代生物学，最早出现的晚期智人在解剖学上与现代人毫无区别，因此，所谓野蛮与文明、原始与现代的区别仅仅体现在人的社会关系和状态的差异。不必求诸考古发现，中国历史记忆中传说的三皇五帝就是带领古代先民进入文明社会的代表，他们肇始了从生食向熟食、从穴居或露宿向室居的重大转变。

在这里笔者提出一个思想实验。假定最早的晚期智人当时也同样考虑了今天人们所思考的一些问题，并提出类似的主张，那么至少他们当中的一些人就会设法阻止他们所面临的演化：如果采用熟食，那个时代的智人的体质和智力就会发生重大的改变，他们的生理性质也同样会改变，比如脑容量会变得更大，身体更为健康因而寿命更长，并且由于个体的差异，有一些特别聪明的智人亦即超级智人就会出现，智人之间的不平等就会加剧。倘若巢木而居并且进一步筑屋以居，那么那些智人就会失去自己的文化。倘若他们要试着穿戴一些遮羞和保暖的衣物和饰物，那么这同样就会改变那些智人的天体观和审美观，他们也许会认为，这种做法是对智人人体的否定，并使得那些穿戴衣物的智人具有了超越他们自身的能力：他们在极端的环境比如在严寒的地区也能生存，而这同样也会造成智人之间的差别。不过，我们应当先关注这类假设所引发的问题：如果那些最早的晚期智人就像今天的一些人所主张的那样，坚持阻止任何超越他们当时能力、状况和水平的演变并且获得成功，那么今天人类所面临的挑战、担忧及其所依赖的环境和条件就不会出现。

如果人类今天为自己的文明——无论是复古主义者对古代观念、制度和器物等文明成就，还是现代主义者对现代文明——而骄傲，那么，这就承带了他们对那些最早的晚期智人超越自身当下能力和状态的观念和行为的肯定。当然，最有可能的

情况是，那些智人根本就没能够形成这样的意识，那么今天的人类就要庆幸他们当时的无意识，否则今天我们就连讨论如此问题的机会都没有。

但是，相比于以往，今天有更多的人认为，现在的人类状况就是人类的最终状态。我们惊奇地看到，一些在观念上具有重大分歧的人们站在了同一立场：拒绝人类存在的进一步演化。不满意当下的现代而时时要回到古代的人们，至少觉得现在要比可能的未来更可取。某些肯定现代社会的人们虽然采取了与复古主义不同的立场，但是，因为他们拒绝和害怕人类的更进一步发展，实际上也就成为一种现代社会的复古主义。只要执着于现在而拒绝进一步的演化，他们的态度就会倾向于不断地退回到更早的现代性。

当然，与那些最早的晚期智人所面临的局面不同，人工智能、基因工程及其可能的综合技术如人工装置等将导致现代人的生物性质的改变。就此而论，今天人类所要经历的变化要远远大于最早晚期智人出现之后的任何演化。

（二）人的性质新变化

笔者之所以不用人性这个词而是人的性质这个词，乃是要强调这种变化将发生在人的生物性质的和智力两个层面，而非单单人的心理、观念和社会性质的改变。下文主要分析和阐述这种变化所承带的基本的哲学的和人文的意义，而不关涉具体技术层面。

我们将从两个方面来分析和讨论这种变化。（1）人的生物性质的变化和由此导致的心理和理智能力的变化。这种变化又分为三种类型，即基因的改变、非基因的改变和综合的改变。（2）人的相对地位的变化。人工智能的崛起以及可能的外星理智生命，对人的地位会造成何种影响。

迄今为止，基因工程已经运用于人身上，人的遗传信息结构通过基因编辑的方式经受了改变这样的案例已经出现。涉人基因工程分为两种，即体细胞基因组操作和生殖细胞基因组操作。涉人基因工程直接改变了人的生物性质，相应地，它亦改变人的心理的和理智能力的性质。

以非基因操作的方式介入人的身体而导致改变这种做法具有悠久的历史，从远古的断发纹身一直到现代的输血、接种、器官移植和人工装置植入，乃是一个持续发展的过程。人工装置的介入改变了人的生物属性，使得人在最简单的意义上不再是完全自然的，而成为人工复合体，不过，这并不导致人的基因和遗传结构的改变。部分人工装置会改变人的心理状态等，而诸如脑起搏器等人工装置则直接影响和改变了人的理智和情感的生物信息的来源。从理论上来说，人工装置可刺激大脑引起各种不同的活动和不同情绪。虽然人类自古以来就用各种外来因素给大脑施加不同的作用以引起不同的刺激和情绪，但以植入人工装置的方式持续可控地刺激大脑以达到所需要的结果，属于根本性的改变。在认识论上，包括脑—机接口在内的人工

装置改变了人的经验的来源，即人的感觉和知觉除了人的单纯的生物器官，亦可通过人工装置获得。

鉴于人工智能和脑科学的迅速发展，人们预测，未来的脑—机接口将是大脑神经系统与人工智能的直接信息连接和交换。^①如果这种脑—机接口得以实现，那么人的学习方式、认识方式和自我认同都将发生根本的变化。第一，倘若记忆能以芯片的方式接入大脑，人的学习方式将发生重大改变，学习能力和效率因此会得到极大的提高，而若干理智能力如计算等甚至无需专门的训练，可以通过外接的人工智能系统获得。第二，据此，人的部分理智功能就可以定制、工业化生产并随时调整和升级。第三，人的某些能力可得到大幅的增强，比如通过脑—机接口，人可以通过意念来操纵和控制电脑和其他的机器。第四，这也必然带来许多现在可预见的风险：人的大脑可以被直接控制和毒害，比如脑—人工智能结合装置被设置了可控制的后门，或置入有害的程序。第五，人的理智能力因此在一定程度上与人的生物基础分离，变得很不确定，从而也难以测试和查验。结合第一点来看，现在的教育方式和体系就要进行重大的乃至根本的改变。

人工装置或脑—机接口还带来另一个严峻的问题，即既有人的自我同一性基础也面临颠覆的风险。从人的主体性来说，个人记忆乃是认同的唯一根据。倘若人的记忆可以通过脑—机结合强化，那么它的内容同样也可以由此而改变。人的同一性向来是一个哲学难题，同时它也是政治、法律、社会和心理等领域的难题。在诸如血型检测、尤其是DNA技术出现之前，人们运用政治、社会和法律等综合手段来证明个体的生物同一性。DNA技术出现之后，确定个体的生物同一性有了可靠的手段。诚然，在哲学和心理等层面，个体同一性依然是尚未得到解决的理论和实践的难题。

但是，基因技术虽然解决了个体同一性确认的困难，但却与人工装置一起，又重新将个体同一性置于不确定性之中。基因工程对人的生物性质的改变也改变了个体遗传学上的同一性。与此相关，人工装置的植入亦使个体失去了生物学上的同一性，因为作为整体的人的生物个体容纳了非生物体的人工装置，并与其他作为整体的有机部分的生物组织一起运转，以支持个体的生命及其活动，包括正常的功能和健康的感觉得。

现在可行的基因工程和人工装置的应用实际上使得现代的许多人成了合成体，即一个由人为设计而经过生物的和物理的双重改变的新种类。当然，从自然人到合

① 人工智能式的脑—机接口的理论和技术现在还处于研究、设计、实验和动物实验的阶段，人们据既有的研究认为，脑—机接口的工作方式按信息的交换方式区分为三种：（1）大脑神经信息输出到电脑等机器，以操作和控制电脑和其他人工器械；（2）电脑等机器向人有大脑的神经系统输入信息；（3）大脑的神经系统与电脑等机器进行双向的信息传输。

成人之间转变的起点并不那么清楚，不过，从现代医学开始以外科手术植入辅助的材料为起点，应是一个基本合理的分界线。泰格马克在他的最新著作《生命 3.0》提出了生命三个阶段的概念，而在最后的 3.0 阶段，亦即生命的技术阶段，生命设计它自身的硬件和软件。因此，生命 3.0 就可以被视为这里所说的合成人的完成阶段。泰格马克对这样的生命形式持积极的态度。^①

综合来说，上述几个方面的改变都关涉人的性质的变化，其中最为根本的几项包括人的学习能力和方式、记忆能力和方式以及经验的来源和获得方式等变化，个人同一性——包括主观认同与客观确证——的变化，以及由上述各种因素导致的自由意志基础的瓦解。在个体的记忆可以通过脑—机接口的方式存储和更换的情况下，这种以记忆集为基础的个人的主观认同就无法与一个特定的个体固定地结合在一起，或者说，从客观上来说，个体就不再具有与任何其他个体区别开来的唯一的人格和身份。于是，从理论上来看，如下情况就有可能发生：一个生物的或合成的个体拥有多个个人认同，因为他/她可以更换多种不同的记忆集，而多个生物的或合成的个体也可以拥有同一个认同，因为他们可以分享同一个记忆集。迄今为止，作为个人人格独立公设的自由意志只适用于一个自然的并且拥有唯一认同的生物个体，而一个更改过的、合成的个体的意志究竟在什么意义上还是自由的？这类个人的行为是否要归因于上述那一系列的更改？这种现象显然是现在的自由意志理论和实践无法应付的。

人们可以领会，人工智能的强大能力相对和间接地改变了人的地位。比如，AlphaGo Zero 在只掌握围棋基本规则的条件下，通过自我学习，最后完胜人类顶尖棋手。这个事件至少从表面上来看成了对人类在智能层面的独尊地位的一次震撼性的冲击。这无疑表明：人的性质因此亦发生了虽然相对却重大的变化。在今天，许多专业性的工作逐渐为弱人工智能所取代，这种现象正在不断扩展。这些原本被视为高端智力工作的职位被取代，人类的理智地位又进一步相对下降。倘若人工自为者在未来出现，那么它就将直接而非间接地改变人的性质，人类从所谓的万物之灵或唯一的理性存在者降格为多种理性存在者之中的一种。如果外星智能生命在某一天被发现，它也会带来同样的或程度更大的影响。

（三）意识研究与人工智能

人工智能的发展，尤其关于人工自为者的预期让包括哲学家在内的许多人踌躇、犯难和恐慌起来。面对这种可能性，人类几千年意识研究的历史以及这种研究的意义和目的，似乎一下子就消失不见了。实际上，人们应当明白，人工智能正是人类

^① Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence*, New York: Alfred A. Knopf, 2017, pp. 40, 30.

几千年意识和心灵研究，尤其是相应的科学与技术研究所指向的结果，因为这些研究的根本目的不仅在于认识意识和心灵的现象及其活动规则、意识活动的物理基础，还在于因此提高和增强人类的理智能力。人工智能不仅在一定程度上开始逐渐实现后一个目的，而且还反过来有助于认识人类的意识和心灵，为了解人类理智的活动方式提供一个对照的样板，而先前这样的样板并不存在。就此而论，我们来分析如下几种情形。

第一，就目前的进展来看，人工智能主要不是模仿人类理智的活动方式。^① 因为迄今有关意识和心灵的哲学研究并没有产生多少实用的效果，即便生物学和脑科学发展迅速，但对人类大脑活动方式的了解却依然有限。人们虽然对大脑单个神经元的活动机制获得了比较全面的了解，但对大脑神经元之间的交流和活动的机制的了解和认识，极其有限，^② 尚不足以为人工智能提供可以模仿的有效样板，尽管这并不意谓，这些研究不会给人工智能的研究、设计和改进带来许多乃至重大的启发。

第二，人工智能发展的动因来自于以人类理智设计出来的方式解决人类理智难以完成的任务。因此，它出自于人类理智对解决问题的方式所做的发现、推演和设想，而不是对大脑活动机制的模仿。不过，人工智能在这方面的进步事实上也证明了人类理智的巨大潜力：能够自由地和有效地寻找解决问题的方法。

第三，人工智能的研究，尤其是关于人工智能与大脑功能结合起来的综合研究，为意识和大脑的研究提供了新途径，即实验和试错的方式。当人们根据对理智行为的认识，一方面是对人的意识和大脑活动方式和结构的理解，另一方面是对为完成某类任务或解决某类问题而设计的人工智能运行方式的理解，努力将两者结合起来时，就可以检查和验证人们对意识活动的方式和结构的认识是否准确。如果这样的方法获得成功，那么，这至少证明人们认识了大脑的部分活动的机制和意识活动的方式和结构；倘若不成功，那么它表明，人们相关和认识是不正确的。这既可能因为相应的人工智能系统是不适用的，亦可能是人们对大脑活动机制或对意识活动的方式和结构的认识不准确。因此，意识研究与人工智能研究、设计和应用之间在未来就可以形成彼此启发和促进的关系。通过这种方式来认识和把握意识活动的方式、结构乃至规则，是具体而实证的，而不同于哲学的抽象概括、推理、思辨构建和语言分析等。

① Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford: Oxford University Press, 2014, p. 29.

② 蒲慕明等人指出，迄今为止的神经科学研究，使得“我们对神经元如何编码、转导和储存神经信息有了较清楚的理解，但是这些神经元的特性是如何通过局部环路和长程环路产生的，我们的理解还十分有限。”因此，对大脑的认知功能的认识也就更其匮乏。参见蒲慕明等：《脑科学与类脑研究概述》，《中国科学院院刊》2016年第7期。

第四，人工智能的发展还揭示了如下的意义深远的可能性：理智能够以不同于人类理智的方式运行；不仅如此，它们还可以多种不同的方式存在和运行。人类理智虽然最终无法一统天下，但可与它们相兼容。这种兼容性也为人类的未来发展提供了乐观的理由和根据。

四、人文主义的未来关切

我们已经看到，人的性质正在经历从生物遗传信息结构到理智能力等一系列的变化，并面临即将到来的更加深刻和巨大的演变，相应地，人类社会亦面临一场前所未有的和持续的重大转向。在这样一种形势之下，回顾历史给我们理解人的未来发展提供了信心和根据，而理性地研究当代的各种发展趋势并据此而展望未来，能够让人们为合理地措置现在以应对未来的挑战，并顺势发展和升华人类自身能力，做好充分的准备。接着，我们要进一步破除人们在认识上的若干盲点和情绪障碍，阐述和论证如下观念：人工智能时代人文主义的核心关切就是更广阔的视野研究人的性质及其变化，重新理解人及其演进的意义，认识并从而促进人类进入更新时代，并在这个时代确保人的崇高地位。

泰格马克认为，我们所分享的软件这种愈益迅速的文化进化，作为塑造我们人类未来的支配力量，已经出现，它使得我们如冰川般缓慢的生物进化几乎微不足道了。^① 泰格马克的观点在今天看来已不属激进，更为激进的观点是认为奇点即将到来，一切都将彻底改变，而不仅仅是进化的升级。

与之相反的观点也有很大的声音，而它们似乎都趋于反对和阻止这种变化本身，而不是否定这种变化的可能性。反对和阻止的理由归根结底就是当下人类存在的尊严和利益，而背后的心态包括人们对不确定性的恐惧。一些思想家为人的尊严所提出的形而上学的主要根据是：人是自然进化的产物，而不能是人的理智能力的结果，人的理智能力的作用及其产物不属于自然的范围。^②

然而，变化已经来临，即便反对，人们也需要了解：人的性质究竟可能发生什么样的变化；倘若这种变化不可阻挡，那么，人类在未来的前景会是如何，以及从现在起，人们应当预先采取什么样的对策。这些正是今天的哲学所无法回避的任务。

（一）人的性质变化带来的挑战

诚然，这个变化不会如奇点一样瞬间爆发，但即便逐渐展开，人的性质演变达到一定的普遍程度时，整个人类从物理基础到行为规范、生活方式都会发生根本的

^① Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence*, p. 29.

^② 参见汉斯·约纳斯：《技术、医学与伦理学：责任原理的实践》，第32页。

变化，其程度要远甚于科学革命所引起的范式变化，^①人类与自然及社会的关系都将在新的关系之下重新展开和构成，相应地，人的认识也要发生根本的转向。

第一，人类工作方式将发生巨大的变化，并且这种变化现在看来正未有穷期。人工智能能够取代人的许多工作，并从事先前人所无法胜任的工作。这个趋势将扩展至越来越多的领域，尤其是那些先前以高技术、高教育水平和高智商为条件的职业。同时，人工智能又创造出新兴的职业，而后者对专业、技术和智能提出了更高的要求。即便相对于日益增强的弱人工智能，至少从表面上来看，在许多领域人类的能力越来越显得薄弱和低效，因此完全可以被取代。

第二，人工智能、基因编辑和人工装置可能导致新的不平等。（1）人工智能淘汰了许多传统的职业和工作种类，而新的工作职业需要更高的专业知识和技术训练，于是，可能会有相当数量的人被排斥在外。就当下的现实来说，局势的严峻性在于：人们不清楚，那些被人工智能淘汰下来的劳动力的出路何在；而就未来趋势而言，如果有一些人始终无法胜任以高水平的知识、技术和智能为条件的工作，那么他们在那个社会中的出路何在？（2）基因工程和人工装置不仅适用于治疗疾病而且亦可以增强人的能力，譬如脑—机结合，而这些技术需要高昂的费用，因此，如果它们不能成为普享的福利，那么显然，这就会强化人类既有的不平等，进一步导致强者愈强，弱者愈弱的后果。

第三，由于人逐渐成为合成人，因而个人主观认同和个体同一性的客观确证的传统根据就会失去效用，现在所理解的自由意志的理由不复有效，从法律、道德到日常生活的约定都将失去根据。

第四，人是什么就成了一个全新的问题。康德将人等同于理性存在者，据此我们可以提出如下追问。（1）理性存在者是否仅仅只能以生物的形态存在？它能否以其他的物理形态存在？（2）作为理性存在者的人，不仅以理智而且还以情感、审美等因素为其共同的根本性质。那么，人工自为者是否也必须兼具情感和审美等能力？（3）或者单单理智能力就足以构成理性存在者，那么人类就属于诸种理性存在者之中的一种特殊的类型。如果这样，那就表明，人类对自身的认识还远远不足。（4）进一步，倘若理性存在者或一般智能自为者必须兼具情感和审美，而人工智能无法达到这一点，那么这个特征是否就成为人类对于人工一般智能的特殊优势？

（二）盲点与情绪

这种人性的变化部分乃是正在发生的事情，部分在相应条件成熟之后必将出现，而现在人们正在努力创造这些条件；还有一些或许永远无法实现。当代的哲学思考

^① Thomas S. Kuhn, *The Structure of Scientific Revolutions*, Chicago and London: The University of Chicago Press, 1996, p. 10.

无法回避这种变化的现实和未来的趋势，而需要就如何理解、认识和看待这些现象、问题和挑战，提出相应的原则、方法论和观念，面对这些已经或即将发生的变化提出切实的因应之道。

尽管可能出现的人工自为者究竟会是什么形态，人们尚不清楚，但恐慌乃至末世的情绪却已经形成，甚至首先在专家群体中兴起。这些情绪主要表现为：第一，茫然。人们似乎无法想象出有效的办法来应付在所有方面都强于现在人类的人工自为者。一些有远见亦不乏悲观态度的学者已经设想了多种方法，比如阻止人工自为者的发展，或在其崛起之前找到控制它的方法，或在设计人工自为者程序的同时就输入人类的道德观念，如此等等。但是，这些设想大都基于想象，缺乏现实的根据，亦无法消除人类莫名恐惧的根据。比如，如果人工自为者能够自我复制，它们必定也能够升级人类强加给它们的程序，而这难保不会废弃人类所设置的道德程序或原则。第二，悲观。一部分人以为人类末世即将来临，至少人类在这个世界的统治地位将被取代，或成为被统治的群体，而最为严重亦最悲观的预期则是，人类最终被人工自为者消灭，因为人工自为者归根到底是不可控制的。

人是人工智能的设计者和制造者，然而，人工智能的强大功能出乎人们的意料之外，而设想中的人工自为者的超级能力仿佛更是人的想象所难企及。在这样一种心态之下，并且实际的科学和技术的研究、发明和设计又将关注和焦点放在对象之上，因而人们或许忽略了这些活动的出发点和立足点，即人类设计和制造人工智能的本来目的乃是为了解决人类智力在技术上难以解决的问题，譬如人类智力在理论上能够而在时间和速度上难以完成的任务，大量的计算、搜索、分析和极限的感知等。

到现在为止，人始终是人工智能发展方向的决定者、最终的设计者和决定者。这一点乃是人类应对人工智能时代的各种挑战的最为坚固的立足点。此外，迄今为止，人类对人工智能的了解依然相当有限，就如人类对自己的智能的认识一样。随着认识的增加，人们就能够找到与之合理相处的方法。尤德科夫斯基认为，人只有在充分了解的情况下制造和应用人工智能，才能取得主动权。^①在有关人工智能时代的论述和预期中，多数论者亦忽略了人的性质的变化这一根本之点。这种变化对理解这个时代以及人类的前景具有基础性的意义。人工智能时代人文主义的核心就是理解、确认并以合理的方式促进人的性质的演变和升华。

与前一点密切相关，一些人不仅忽略人的性质的变化，并且无视甚至拒绝人的智能的进一步发展和提高，或者一般地说，拒绝人的继续进化。我们看到，现在研究和讨论的主流集中于对人工智能发展的设计、分析和预测，以及预防人工自为者

^① 参见 Nick Bostrom and Milan M. Ćirković, eds., *Global Catastrophic Risks*, Oxford: Oxford University Press, 2008, pp. 342-343.

对人类的可能威胁的策略，而缺乏对人类理智能力相应地提高这种重要的可能性的考虑。于是，在当前的许多理论研究中就出现了如下理论反差：一种相对来说越来越弱的人类理智，正在设计和制造出越来越强大的人工智能，最后制造出人工自为者，并且最终使得自身受制于这个人工自为者，甚至为其所毁灭。比如，人工智能可以发展和完善，从专域的专家发展成为人工一般智能，并且还继续自我完善和演化，而人类却不能相应地进一步发展和演化。^①这个观点就为人类在将来受制于人工一般智能提供了不对等的条件和限制。如果按照杨波利斯基的观点，人工智能的完善和全面升级是不可阻挡的，那么我们要指出，同样，人类的进一步演变和进化也同样是不可避免的。

其实，人类发明、设计和制造出了人工智能，不断地拓展、升级和强化它的能力，人类也因此大大地扩展、强化乃至升级了自身的综合能力。人们或许会问，人类在强化自己的综合能力的同时，是否同时能够增强本身的智力的和物理的能力？回答应当是肯定的，而且也是必然的。理由很简单：进化并不是到今天就结束了。

从智人出现到今天为止，人类的演化史同时就是人类持续增强自身能力的历史；人类不仅提高了自身的理智能力，而且也提高了自己的物理能力，包括体能和寿命等，而在这个过程中，人类充分利用了外在的条件，而后者之中的许多就是人类自己发现、发明和制造出来的，如医药、食物和体育活动等。

一些人对人类自身的能力和潜能缺乏足够的认识。这些人之所以对人的理智能力持悲观的态度，还出于如下一个隐而不显的假定：他们已经获得了对人类的全部认识，因而对人类的未来前途能够做出最终的断定。事实上，断言人类智能不能适应人工智能时代事实上缺乏足够和明确的根据。以为人工自为者将临而人类末世随至的人们常常忽略和忘记如下一点，人工智能恰恰表明：人类理智具有强大的适应和扩展能力，以及尚未认识的巨大潜力。在先前，人们因为没能预见这个时代的种种科学发现和技术发明，抑或没有预见人的理智具有发展出这般强大的人工智能的能力，以致既缺乏理论准备也没有心理预期。

许多对未来前景的讨论、推论和展望较少关注和研究人工自为者的交往与组成社会的可能性。人工自为者不可能是独立地获取一切存在需要和保障的独行侠，因此，它们也要通过合作才得以成为自为者。不仅如此，如果人工自为者是类人的，而人类知识是通过集体的合作而获得的；更进一步，如果人工自为者有意识，而人类意识同样是人类集体的产物，而非任何一个遗世独立的个人的成就，那么人工自为者之间的合作同样就必不可少。当然，如果人工自为者不是类人的，人们现在就难以推测和设想这类自为者会如何行动，如何在某个时间摆脱人类独立地生存，自主地获得知识并且促进知识的增长。博斯特罗姆认为能够控制人类的超级智能是一

^① Roman V. Yampolski, *Artificial Superintelligence: A Futuristic Approach*, pp. 187-189.

个孤独体 (singleton), 亦即整个世界只有一个超级智能。^① 但维尔克斯认为这种说法等于沦为人类学的科学幻想。^② 总体而言, 关于人工自为者的认识人类还存在一些尚未意识到的盲点, 而破解这类盲点正是人类应对人工智能未来的一种出路。

五、人文主义的新维度

在人工智能时代, 人们主要地关注了人的理论和实践活动的对象的演化发展, 而轻忽甚至无视人本身的性质的变化; 迄今为止的人文主义所关注的重点在于确认和提高人在这个世界的地位和价值, 促进人的潜力的发挥, 但人的个体的生物存在则被认为是始终不变的。我们则要强化, 充分认识这种变化的可能性和现实性, 将这种变化纳入有关人的性质的认识的视野之中, 以此来考虑人在这个世界的至上地位和未来前景, 乃是人工智能时代的人文主义的中心关切。这样的人文主义的特征可以从如下几个方面来论述。

(一) 开放的态度与发展的视野

从历史上来看, 人文主义在其渊源上缺乏有关人的性质的形而上学公设, 它依据当时那些贬低和压制人的观念的对立面来规定人的性质, 即便诸如存在主义一类消极的人文主义, 也要为人的存在寻找一个对手, 譬如自然、科学和技术或官僚机构, 以便由此来定义人的性质。但是, 他们为寻找对手而找对手, 从而使他们所理解的人的存在陷入了自我孤立的陷阱。

一种开放意义的人文主义要消除和避免这种把特定时代的人的性质永恒化和固定化的困境。在历史上, 人文主义主流就一直在排除各种外在对象的压制, 以确立人在这个世界的自主性。但是, 它带来的消极结果就将人的性质固定化, 人的性质的变化因而就成了一种新的禁忌。

这种禁忌的亦即封闭的观点在海德格尔或哈贝马斯的学说中表面上看起来是在维护人的现实存在的尊严, 而实际上却沦落为一种无所作为的态度: 拒绝一切改变, 而手段却是以退为进, 回到原本就阙如的某种形而上学的存在。根据这种封闭的或无所作为的态度, 人类就此不再发展和进化, 而手足无措地目睹和见证人类所制造的对象突飞猛进地发展和升级, 从而持续并远远地将人类甩在后面, 人类最终被人工自为者所控制。但是, 这种消极的观点和态度包含一个事实和逻辑的鸿沟: 迄今为止, 人工智能的一切进步和升级, 都是人类创造性活动的结果; 基因知识和技术

^① 参见 Nick Bostrom, “Decisive Strategic Advantage,” *Superintelligence: Paths, Dangers, Strategies*, Chapter 5.

^② Yorick Wilks, “Will There Be Superintelligence and Would It Hate Us?” *AI Magazine*, Winter, 2017 p. 67.

亦是如此。这个事实刚好说明，人类的理智和其他相关的能力也在相应地发展和提高。因此，一般地说，人类也会随着自身对人工智能的不断改进而持续地提高自身，以适应相应的变化。反过来说，倘若到今天为止，人类不再发展和提高自身，那么人工智能同样也就会永远停留在今天的水平，因此就谈不上对人类的威胁。

因此，人工智能时代的人文主义的宗旨就在于思考如何利用人工智能、基因工程、人工装置和其他可能出现的知识和技术为人类服务，发展和增强人本身的能力和外在能力，大大减轻人类现在所承担的高强度而高度重复的体力和脑力劳动，使人类智能和潜力施用于创造性的活动，扩大人类认识的广度和深度，以及人类情感的自由范围。

在另一方面，人工智能时代的人文主义必须关注这个时代的社会规范和秩序的升级，比如更新或构建新的正义原则。如何避免由人工智能、基因工程和人工装置等应用带来的人的新的不平等，依然是人文主义的核心内容，而人类能力相对于人工智能在某些方面的落后，也只有通过这样的技术得以振拔。这里需要关注若干常被忽略的方法论要点。第一，由知识发现和技术发明及其应用所造成的人的不平等并非在今天才出现，而是在最原始的工具刚为人所利用时就已经发生了。第二，人类并不会因为这种差别的出现而停止或拒绝知识的探索和技术的发明，因此，平等关切的要点在于使得所有人都有能力享受同样高效和便捷的技术设备，而非禁止或停止探索和发明。

（二）无尽的学习

面对日新月异的人工智能时代，人类整体就如每个个体一样需要持续地学习、研究和提升。面临人工智能时代的挑战，在一些人那里自启蒙运动以来的雄心和目标尽皆化为恐惧。目睹这一现象而使我们领悟到，不少人事实上还停留在启蒙时代早期的观念，即为了改造自然而认识自然，人们能够达到对自然和人类社会根本规律的认识，而这种规律乃属永远有效，恒久不变。因此，在他们看来人类只能战战兢兢地固守在当下的状态以及那些既有的知识范围之内，不能逾越一步，消极地等待未来的到来。

但是，未来已来。因此，对今天的人们来说，突破启蒙时代早期的观念，持续探索人类自身和外在世界以及两者之间的关系，就是必不可少的积极选择。这样的研究是多向的，既包括继续进行生命科学、人工智能和其他科学与技术的研究和发明，同时亦进行应对这些已经出现或即将出现的现象的对策研究，甚至包括预防的研究，而后一种研究就要由人文科学、社会科学和自然科学的综合研究来承担。

结 语

人工智能时代的人文主义的基本结构就是一种关于人的新形而上学，因为它探

讨和把握演变中的人的性质。它承继先前的人文主义，然而更加关注人类未来发展的确定性与不确定性。

第一，这种人文主义依然坚持理性的普遍性原则。从现在起，普遍性原则将立足于多样性的基础。如果先前人的理性和资质的普遍性以单纯的人的生物存在为基础，那么从现在起，这种状况将逐渐改变。如果诸如人工装置和基因工程对人类个体的影响并不动摇理性的普遍性，那么未来进一步的发展也同样不会动摇这种普遍性。不过，考虑到人工自为者和外星生命的可能性，理性的普遍性将与智能多元性兼容，而限于单一人类理智。

第二，自主乃是人类的根本特征，这在未来也同样如此。但是，自主又始终是一种社会性质，而未来人类的每个个体除了相对于同类的其他个体的自主之外，还要面对其他可能的理性存在者的自主。因此，自主就具有了双重的意义。我们必须认识到，任何自主都承带某种不确定性，因为自主总是相对需要决定的未决状态而言的，换言之，自主就是面对不定的未来做出决定的一种资格和权力。无需任何决定的境况既不需要自主，也没有它存在的余地。就此而论，人类自己的选择和决定依然属于未来人文主义的核心。

第三，人类前途的最终不确定性是基于自然科学和人文主义的必然结论，尽管这是许多人不愿意面对的事实。超越启蒙时代的态度乃是，这样的不确定性对人类来说并非属于悲观的结论，而有其乐观的前景。不过，我这里要对这种不确定性予以必要的分析。

因为人们对自身和世界的知识的匮乏，并无可能对人类远景、科学和技术的发现、发明和其技术的运用做出最终判断，所以，现在也就无法预测人类进化能达到什么样的高度，在宇宙中的前景又会如何，骤下断言，只是过于傲慢的表现。

迄今关于意识活动的方式与结构和大脑活动的研究与认识都相当有限，然而，人们却能够运用意识或大脑活动的成果比如各种算法，发明、设计和制造人工智能。那么，人工智能会将人类思维又引向什么方向？这个追问就向我们揭示了人类未来发展的别开生面的前景。基因工程和人工装置直接介入人的生物个体，不仅导致人的生物性质的变化，包括大脑状态和活动的改变，同样会引起意识活动的方式和现象的改变。在这个过程中，人们通过理论研究或实验测量所得的认识始终是已受到介入和影响的大脑状态和意识活动的结果，而在此之前的大脑状态和意识活动仍然是不可知的。因此，正是人类的科学和技术的进展及其应用进一步强化了人类前景的不确定性。

但是，这种不确定并不妨碍人类对自身的进化和前景保持合理的乐观。哈贝马斯一派哲学家仅仅盯着当下社会，而将它视为永恒不变的状态，因此，对人的性质以及人类的尊严的理解是僵硬的，而这实际上恰恰贬低了人的发展能力以及当代人的尊严，因为人的尊严并非向来如此，而是持续演化和不断进步的成果。

诚然，这种合理的乐观态度还依赖于人们是否能够接受人的性质的持续进化。今后的进化与以前的会有三个重要的差别。首先，今天的人类是在明确地意识到这种进化的情况之下接受和经受这种进化，而先前从最早的晚期智人到现代人的进化是在无意识的情况下发生的。人类对进化的明确认识才有百多年的时间。其次，人成为有意识地促使自己的性质发生变化的重要动力。最后，倘若人工自为者出现，那么它就会成为人类的竞争者，因此在未来的人类进化中，它同样也就扮演了人类进化的促进者的角色。在历史上，人类发明和制造的工具事实上一再发挥了类似的作用，只是先前的人们以为这些工具单单受人的支配而被使用，而没有认识到：它们会成为人类的竞争者，并且在这种竞争——即何者为支配者——中，人类的发展和进化受到了重大的促进。

对人的性质的变化的抗拒态度不仅仅来自于对人类可能受到人工自为者的控制的推断，而且主要还出于对确定性丧失的焦虑和恐惧。在西方基督教占统治地位的历史上，在上帝与人类的二元关系和最后的救赎中，命运不确定的只是个别的人，而整个人类与世界的最终前途则是明确的。在这传统熏染之下，人们的心理习惯了一种确定状态：上帝与人的对立以及这种对立的最后消解。今天，面对人工智能时代不断更新的科学知识和技术发明，这种最终的确定性消逝不见了。

沙伦谈到了另外一种不确定性，即现代生物技术导致的不确定性。“它也以其能力深刻地重塑，如果不是废除，清楚的和有意义的界限。这些界限被用作许多传统思想范畴的固定的和可靠的框架，诸如人类和机器之间的界限，自然和技术之间的界限，治疗与增强之间的界限，出生和制造之间的界限。”^① 人类迄今为止所具有的基本观念和界限都需要重新构造，重新解释，仅仅这一点也就足以让人们感到担忧和恐惧。但是，致命的并非单单观念及其界限的更改，而是实际生活中的行为方式和整个秩序全部都要重新调整、安排甚至构造。在无法设想如何应付这些混乱的新状态而又无法设想出新秩序时，人们的拒绝态度就会油然而生。

人工智能时代的人文主义精神就是持续地促进并在可能的情况下筹划人的发展和进化，借助于日新而新的科学和技术，持续地提高人类自身而使其得到升华。人的性质如不能变化，那么它的前景就不是仅仅中止并永远停留在当下，而是趋向于淘汰和消亡。人工智能时代的人文主义蕴涵了如下的自信和精神：人类本身的发展和进化在今天已经开始呈现出文化的与物理的统一趋势，这正是人们应对人类前景的积极观点的根据所在。

〔责任编辑：莫 斌〕

^① Tamar Sharon, *Human Nature in an Age of Biotechnology*, Dordrecht, Heidelberg, New York and London: Springer, 2014, p. 1.

ABSTRACTS

(1) Chinese Socialist Party Politics in the Process of Modernization

Wang Shaoxing • 4 •

Modernization hastened the birth of party politics, and party politics dominates modernization in turn. The intergrowth and interaction between party politics and modernization represent the great logic of the evolution of human civilization in modern times. The objectivity of China's modernization implies the historical necessity of its party politics, and its national and epochal character have led to the growth of Chinese party politics with socialist characteristics in an irreversible developmental trend. In the past 100 years, with the historic theme of the great renewal of the Chinese nation and during China's revolutionary process and the extensive practice of building socialism with Chinese characteristics and constructing a modern country, Chinese socialist party politics has integrated its own values with China's modernization and played a leading role with a distinctive "fundamental value orientation," "precondition of theoretical guidance," "dominant institutional support," "purposeful value traction power," and "crucial political security"; it has profoundly changed the developmental course of China's history with the strongest value pursuit, the most unique practice and the most significant contribution to civilization; and it has achieved the great historical change of Chinese society's efforts to modernize, which in turn has deeply influenced world epochal changes. This is not only the core symbol of the experiential integration and theoretical creation of the interaction between socialist party politics and China's modernization, but also the key support and value guide for building the Communist Party of China into a powerful political party across the world, and thus building China into a modern power.

(2) Humanism in the Era of Artificial Intelligence

Han Shuifa • 25 •

Scientific and technological advances such as artificial intelligence and gene editing are currently developing rapidly and reaching all levels of human society. This not only poses great challenges to human intelligence and its unique status, but also leads to changes in the nature of human beings. In the face of the urgent pressure imposed by the continuous advances in artificial intelligence and the impact of the changes taking place in human nature, current opinion is divided and contradictory; it rejects the necessity of human change, but its passively pessimistic and failed mentality is powerless against the possible prospects of artificial intelligence. These challenges and phenomena have forced people to rethink basic philosophical issues concerning humanity, and have re-energized the core concerns of humanism. Today's

• 204 •

humanism emphasizes that while we recognize, invent and manufacture artificial intelligence and other technologies and constantly upgrade them, we continue to enhance and improve our own intellectual abilities. From now on, we must consciously observe and enhance our own evolution. Such development and evolution evince the trend in today's economy toward cultural and physical unity. It is on this basis that we can adopt a positive view of the challenging prospects before humanity.

(3) The Third Turn of Western Hermeneutics—From Habermas to Ricoeur

He Weiping • 45 •

Paul Ricoeur likened the two turns in Western hermeneutics—from partial to general hermeneutics, and from methodological to ontological hermeneutics—to a second Copernican revolution. We believe that hermeneutics is undergoing a third turn, i. e. the combination of hermeneutics with epistemology and methodology, represented by Habermas and Ricoeur. Habermas pioneered this turn, which has been developed further by Ricoeur. One could say that this is a return to Dilthey at a higher level, in a development that has a broad and profound meaning today.

(4) Improving the Efficiency of Enterprise Technology Application and Strengthening the Construction of Human Capital *Li Jing, Liu Xiahui and Nan Yu* • 63 •

Our research explored the micro-foundation of innovation-driven high-quality economic development, building a growth model of enterprise technological progress and taking human capital allocation as the entry point for portraying the balanced option of investment in enterprise R&D plus efficient application of technology, together with the optimal growth path determined by this choice. Our findings show that improving the efficiency of enterprise technology application can promote the effective allocation of human capital; hence it is an important condition for selecting the optimal growth path. The successful implementation of this program requires that government develop new growth points and new momentum for high-quality economic development and optimize allocation of human resources by actively promoting the effective supply of modern service industries (science, education, culture, health, sports, etc.) that foster and build up human capital at the macro-structural level and by realizing the in-depth integration of advanced manufacturing and modern service industries.

(5) Performance Evaluation of Government Financial Revenue: A Value Tool for Realizing the Policy of Cutting Taxes and Administrative Charges

Zheng Fanghui and Fei Rui • 85 •

Finance is the foundation of national governance, and the tax burden is the setting that quantifies the degree to which government rules for the people. The objective requirements of a proactive fiscal policy of cutting taxes and administrative

• 205 •