
人机共舞大幕已启，人类准备好了么？

作者：肖仰华 来源：澎湃新闻

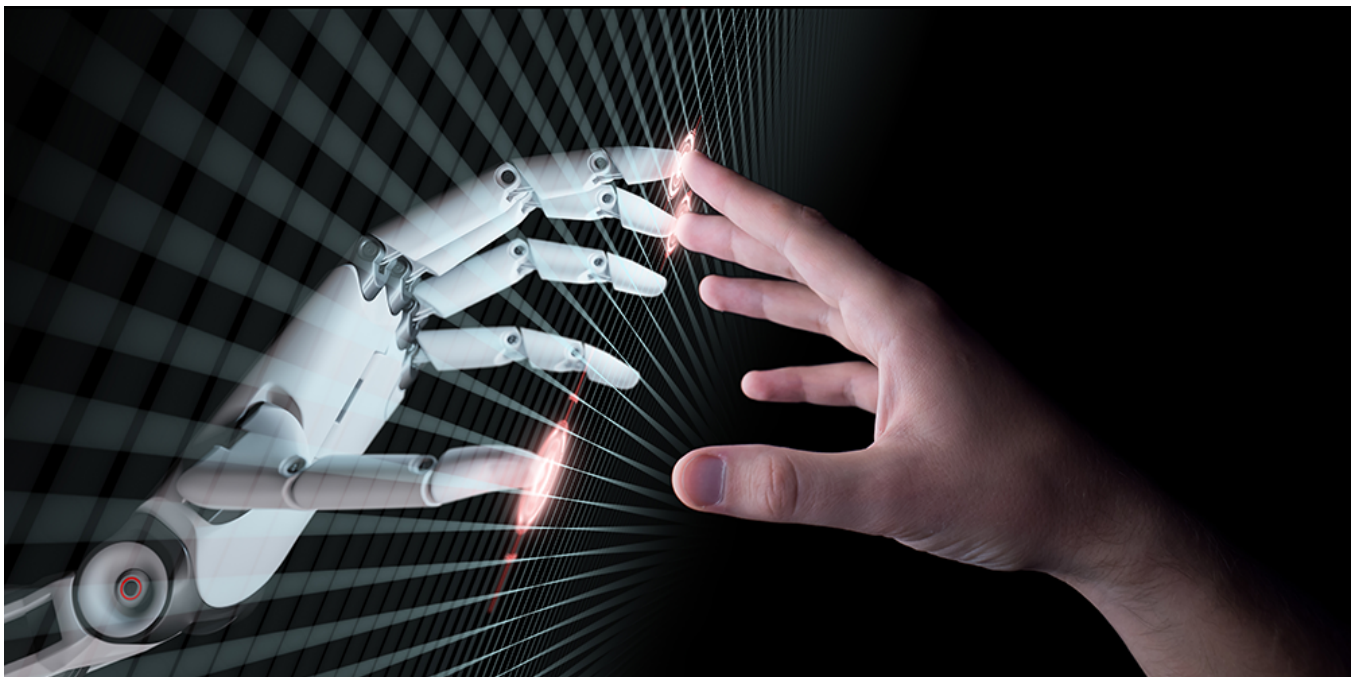
本文原地址：<https://www.iikx.com/news/topnews/27028.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

AGI的全面渗透和替代人类脑力劳动，可能会损害人类的核心智力能力，对“何以为人”这一根本问题带来冲击。如果人类丧失思维、审美和情感等高级能力，其作为万物之灵长的本质将遭到伤害。因此，AI技术应当避免对人性造成伤害，确保人类脑力能力得到充分训练和实践。

此外，AI滥用可能冲击就业结构、加剧社会分化、破坏组织治理模式、污染精神世界、侵犯人类意志、介入人类情感伦理事务等。AIGC技术所带来的内容泛滥，更可能消解美的稀缺性，损害人类审美情趣。

为防止这些潜在危害，应当为AI应用设立准则，AI应用应有其边界，不应过度、无限制使用。总之，AGI时代人机共舞的大幕已经拉开。人类应以何种身份出场，如何处理好与AI的关系，需要我们深思熟虑。追求自我超越的人性倾向推动了AGI的发展，但这一进程必须稳健前行，三思而后行。



以生成式人工智能为代表的通用人工智能(AGI, Artificial General Intelligence)，正在加速发展，机器智能达到甚至超越普通人类智能水平已不再是科幻而日益变成现实，机器超越人类的任务列表越来越长。机器能力的根本性飞跃进一步吸引了巨大的资本和资源，各种生成式大模型获得快速

发展，已经成为先进生产力的典型代表。

相比较而言，人的生产能力渐渐有了“落后”的含义，毕竟人类脑力难以二十四小时在线，人类的体力也很难持续很久，人类的可靠性与稳定性也不比机器有多少优势，再加上AGI智力水平对人类的超越，传统基于人力资源的生产力有被人工智能(AI, Artificial Intelligence)生产力所取代的趋势。在不久的将来，AGI终将有能力从事人类生产与生活中需要用到智力的大部分工作，完全有能力介入人类的情感与伦理事务。

AI将无处不在，像空气一样弥漫在我们生活的角角落落，似乎已成必然趋势。每个个人、群体和组织都将被AI浪潮所裹挟，AGI将渗透到社会发展的每一根毛细血管。我们现在已经很难想象没有电和网络的生活会是什么样子。同样，在不久的将来，我们也很难想象一个没有AI的世界。未来时代的根本特征恐怕就是“无所不在的AI”。

然而，在AGI时代大幕即将拉开之时，我们不得不向时代的思考者们问一句，人类准备好了么？这场时代大戏必是人机共舞。人类将以何种身份、何种方式扮演好自己的角色，又将如何处理好与机器这个伴舞者之间的关系？随着AGI发展的加速，这些问题的回答日益迫切。AI技术进步速度之快，应用之速，给我们留下的时间并不充裕。

在肯定人工智能作为先进生产力为人类带来福祉的同时，我们也要审慎思考它可能给我们社会发展带来的一系列问题。AI发展的负面问题已有大量讨论，比如虚假泛滥、隐私泄露、技术垄断、偏见不公、价值错谬等等。这些问题都是当下已在发生，已为社会所熟知的问题。然而这些问题某种意义上都是表面层次的问题，随着AI在几代人群中的大规模应用，有可能在发展先进生产力的旗号下，造成一些长期而缓慢却根本而致命的问题。好比现代年轻人群体的疾病，急症易被感知与重视从而不难诊治，而长期的、隐性的慢性病却往往被表面的浑身肌肉所迷惑而轻视，不幸过早遭受致命一击。

想想人类历史发展进程，何尝不是如此，虚假的繁荣景象往往最易掩盖暗自涌动的真正危机。只有清晰认识这些问题并找到预防之策，才能将人工智能发展成为推动社会进步、造福人类的先进生产力，而不是将人类带入万劫不复境地的罪魁祸首。

AGI全面“入侵”，再问“何以为人”。

AI的进步不可避免地带来人类相应能力的退步。事实上技术的每一次进步，有可能导致人类相应能力的退步。汽车解放了双腿，我们必须依靠高度的自律通过跑步或其他运动才能保有双腿的能力。键盘代替了书写，提笔忘字就成了很多人的共同体验。四肢能力，即便我们承认是在退化也多不以为然，因为这是将双手解放出来从事艺术创作、将双腿释放出来去美丽的田野郊游而必须付出来的完全可以忽略不计的代价。

一直以来，工具的使用即便导致人的某种能力的倒退，但并未改变人之为人的本质。工具本质仍是人的附庸，人类因其强大的思维能力、独特智慧而具有不可替代性。

然而，达到人类智力水平的AGI以一种无孔不入之态势对人类脑力劳动进行肆无忌惮的全面代替，却十分值得我们警惕。以人工智能内容生成(AIGC, Artificial Intelligence Generated Content)为代表的AI应用已经在日常生产、生活中大量使用。AI生产代码、文案、图案、音乐、视频已经超越普通从业者的水平。常规的内容生成任务AI基本上可以较好胜任。久而久之，人类社会大量内容创造工作将被AI所代替。

大量的内容生成任务实则是人类思维训练的重要方式。如果人类没有足量的群体去实践脑力活动，如果AI剥夺了人类思维训练的机会，人类如何保有相应的智力水平呢？我们很难想象将来有一天我们的传统技能，比如烹饪的技艺、开车的技巧、写作的技巧、写代码的能力，需要通过指定非物质文化遗产继承人，才能传承和保有。

人何以为人，在AGI时代变成日益严峻而需要重新思考的老问题。人类相对于万物的独特性，高高凌驾于万物之上而成为万物之灵长，正是由人类的思维等核心能力所决定的。一个人丧失了行走或言说等能力，不影响其作为人的本质。然而，如果一个人丧失了基本智力水平，丧失了思维能力、审美能力、情感能力，或者更极端一些成为了哲学家们曾经设想过的有身无脑的“哲学僵尸”，那么其作为人的本质则会受到根本性的伤害。

或许有人会说，信息时代计算机的计算与记忆能不也一直在代替人类的相应脑力工作，也并未发生所谓的人类本质退化危机。针对这种观点，我想提醒的是，AGI时代，AI对人类脑力活动的替代是全面而深入的，而信息时代的脑力能力替代总体上是散发、点状的，量变带来质变。

AI的发展与应用应极力避免对人之为人的本性的伤害。人的退化绝不是发展AI初衷。没有人愿意看到人类成为智能机器所豢养的宠物。人曾经为人，人也必须仍然为人。一切有害于人之为人的AI应用，应该予以限制。为AI应用设立准则是发展AI的前提。AI应用伦理的根本原则就是不能损害人性，确保人类的脑力必须得以充分的训练与实践。

关注对AI的恶用，尤其滥用。

关于AI的使用，我们需要关注的不仅是其善用，也应关注其恶用和滥用。所谓的AI恶用，就是恶意的AI应用。比如：AI用于自制武器、欺诈欺骗。我们多数人对AI恶用有着清晰的认知，能够作出积极的防范。

让人担心的是AI滥用。所谓的滥用，是指过度地、不加限制地使用AI技术。这种滥用往往出于AI眼前巨大利益的考虑而有意无意忽视AI长期问题，最终对人类福祉或者特定群体利益造成难以弥补的侵害。正如克鲁斯所说：“人工智能是人类有史以来发展的最大发明之一，但同时它也可能是我们的最后一项发明。”

AI滥用往往有着温和甚至是极具吸引力的外表。有多少人能拒绝“先进生产力”，有多少人能拒绝一个贴心的、随时听候召唤的AI助手。人类对毒品尚难以拒绝，更不用说AI这类精神毒品。久而久之，AI滥用会像温水煮青蛙一般以一种缓慢而难以察觉的进程给人类带来难以挽回的伤害。对于AI的滥用问题，我们需要保持高度的警惕。AI应用应该有所为，有所不为，AI应用应设立基本原则。

AI应用所代替的人类能力应该严格限制于人类已经具备的能力。换言之，如果某人想使用AI的某项能力，他自身在这方面的能力应该超过AI而不是低于AI。首先，AI是可能犯错的，AI所犯下的错误需要足够专业水平才能识别。如果人们在某一方面尚不具备必要的的能力，就很容易被AI误导。其次，如果我们在尚未掌握某项能力的情况下，就轻易地使用AI的这项能力，那么久而久之我们的这项能力就得不到足够的锻炼和实践，进而难以充分发展这种能力。如果说这种能力的损失对我们来说是无法接受的，那么我们就应谨慎对待AI相应能力的使用。

或许会有人说，既然我已经具备某项能力，为何还需要使用AI的这项能力，这似乎是个悖论。事实上，作为以某项能力见长的人类专家，其精力与时间是有限的，AI的赋能将专家从其擅长的任

务中解放出来，使得专家有足够的从容去挑战更加困难的问题。从这个意义来讲，AGI本质上应该是人类专家提质提效的工具。对于普通人，AGI则应慎用。

AI的滥用将冲击当前的就业结构，进而引发社会震荡。AI的应用应以充分保障人类就业为基本前提。每当我们创建一个AI新应用，我们就有责任设计出一个新的人类职业来保障人类的充分就业。每当AI替代一个人类工作岗位，我们就应创造出一个新的就业机会。创新人类的岗位设置、保障人类的充分就业应当成为政府和企业社会责任的新内涵。否则，久而久之，人类只能进入无所事事的境地。

人们的学习进程，即便得到AI辅助，相对于AI的快速发展速度而言仍是缓慢的。人类就业结构的调整涉及经济结构等上层建筑的调整，是个缓慢的进程。这一调整进程也必须有序而缓慢，否则会引起社会震荡。因此，AI的发展应该为就业结构、经济结构调整留下充足的时间，以保证社会的平稳过渡。我们需要认识到发展AI出发点是利用先进生产力去促进社会发展与文明昌盛，而稳定是这一切的前提。AI应用不能背离其本意。

AI的规模化使用的一个潜在风险是撕裂人类社会阶层结构。未来，人类势必会因为AI的存在而被分为AI智能水平之上的人和AI智能水平之下的人，不管每个个体愿不愿意，人与人之间的这种差距是客观存在的。随着AI智能水平的持续进步，大多数人的智力水平可能处于AI水平之下将成为一个难以避免的事实，超越AI、驾驭AI将会变得越来越困难。《北京折叠》之类的文学作品何尝不是对AI应用所造成的社会结构撕裂所作出的超前隐喻。

如何防范由于AI应用所带来的阶层固化甚至阶层对立，如何保障让个人都能拥有平等的AI教育和使用权利，从而防止AI能力差异所带来的社会结构撕裂，是我们需要严肃思考的问题。

AI的滥用也可能给人类社会的组织结构带来极大的冲击。人类是以特定组织机制，以有效的协同方式，以一种集体智慧来共同应对所面临的挑战的。无论是企业、军队、科研团队，绝大多数组织呈现出一种金字塔状结构。在顶端的往往是富有知识、洞见以及智慧的决策者、战略规划者，中层是富有经验、长于事务处理的管理者、参谋者、战术制定者，而底层则是具备一定专业知识与技能、从事具体执行与操作的员工、助理、士兵、学生。

随着生成式人工智能的进步，这种金字塔形状的组织结构将会受到强烈的冲击。当AI能力超越了底层和中层的智能水平之后，金字塔尖的专家更愿意直接使用AI助理，以更加高效的方式，完成一些重复性、常规性的工作。专家/决策者/规划者与AI助手的协同，打破了以往的人类组织协同模式。随之而来的问题是人类助手的实践机会的减少，其成长为专家的路径被截断。当老一代人类专家逐渐凋亡，新兴人类还有多少人能胜任传统技艺，是我们需要严肃思考的问题。

护持人类真实世界，为AI设边界

AI应用不应污染人类的精神世界。当下，AIGC技术日新月异，文字、图像、语音、视频等各种模态数据的自动生成与处理技术都取得了前所未有的进步。AIGC为内容生成、艺术创作、影视制作带来了极大的效率和质量提升。人人都是创作者、艺术家的时代似乎已然来临。一个小学生只要有足够创意，似乎就可以制作一部影视大片。然而看似丰富多彩AIGC生成内容，也不可避免的带来人们感官上的头晕目眩，旧日世界的天真与烂漫日益演变成AI时代的光怪与陆离，我们越来越难以理解由AI所塑造的新世界。

AI内容创造的本质是组合拼装，是移花接木。首先，我们必须承认与珍视AI这种组合创新能力，

因为人类的大多数创新实质也是组合创新。而且，AI的组合创新能力远胜于人类，它能将人类难以关联的元素组合。所谓风马牛不相及、八杆子打不着的元素组合，对于AI而言是十分容易的，比如青花瓷外表的舰艇，三星堆风格的飞机。这种鲜见的组合创新在AIGC工具成熟之前，是不会有画师无聊到去尝试的。然而在AIGC降低创作门槛和代价之后，大量的这类新奇且古怪的组合创新呈泛滥之势充斥人类社会。

人类生物感官有多大程度上能适应未来光怪陆离的世界？这是值得我们严肃思考的问题。更为值得深思的是机器的组合创新本质上是同质化的创新，其本质上是借助机器的强大计算能力将人类难以想象的组合以较低的代价生成并呈现而已。人类对世界的认识、人类从零到一的原始创新，并未因为组合创新而有所增益。组合创新的表面繁华反而更容易掩盖原始创新贫弱的真实窘境。

AI应用不应侵犯人类的主体意志。人之为人的根本特性之一就是自由意志。但事实上，AI工具在一定程度上已经侵犯了使用者的主体意志。例如，在推荐系统里，用户只需要进行简单类目设置，便可享用推荐功能。换言之，AI已经在某种程度上已经行使其自身的推荐意志。事实上有多少内容是我们真正想看到的呢，越来越多的内容实质上是AI背后的平台方的意志体现。更进一步，由于当前大模型技术的不透明、不可控，也可以说是大模型自身的意志也体现在了推荐结果中。

随着AIGC的日益成熟，短短的一句提示词，就能生成一分钟的视频，要不了多久一句提示词就能生成一部电影。那么这部电影中的内容有多少是经过人类思考并确认的呢？绝大多数内容都是AI自身意志的体现。

事实上随着AI的进步，人类逐步放弃自由意志而由AI代替其发挥决策，似乎是个必然趋势。人类基于语言的意愿表达从根本上来讲是模糊而不精确的，这就为AI发挥自身意志留下了空间。其次，AIGC创作代价之低免除了人类繁重的艺术创作工作，有多少人能拒绝这种诱惑呢。试想如果AI根据一个极其简单的提示词就生成一部让人震撼的电影作品，那么会有多少人会要对抗AI的这种创作呢？传统艺术创作的艰辛是很难抗拒AI制作的轻松的。最后，自主决策对于很多人而言是一种不可承受之重。选择困难症、随大流可以说是现代人的通病。严肃思考下的决策往往是一种人们急于推卸的责任。有多少人愿意承担关键思考的责任呢？

更进一步，AGI创作本质上体现的是人类的集体意志。AIGC滥用容易造成人类的集体意志对个人意志的抹杀，最终降低人类集体意志的认知水平。大模型第一次使得人类有机会相对全面地考察人类的集体意志，这是极为值得我们珍视的能力。人类的知识、文化多元而庞杂，几乎没有人能全面了解全体人类的文化与知识。然而大模型在基于全人类尽可能多的高质量语料喂养下，接近成为一个全才和通才，能够以上帝视角建模不同学科、不同文化对于世界认知。概率生成方式从根本上决定了大模型基于占据样本多数的群体主流意志而非个性鲜明但鲜有出现个体意志生成相应的内容，AIGC因此体现的是主流的群体性认知。

我们经常觉得大模型生成的结果有四平八稳、东拼西凑的感觉。“四平八稳”本质上是概率计算的平滑处理，“东拼西凑”本质上是大量语料学习的结果，来自不同人的观点可能被综合融入进了同一个问题的答案。如果所有的问题，都是来自大模型的从主流集体认知而形成的全面且综合的标准答案，势必会抹杀个性化认知。长期耳濡目染大模型生成的结果，久而久之，我们每个人的思维也渐有综合而稳健的趋势，但却可能丧失了难能可贵的个性与独特。

AI应用应谨慎介入人类的伦理和情感生活。人类的伦理和情感是十分复杂的，所谓人心的幽微与人性的微妙，集中地体现于人类的伦理和情感事务。人的问题，终究应该由人类自身来解决，AI

的介入往往是人类不负责任的一种表现。不同的时代、不同的文化、不同的族群、不同的情形，对同一件情感或伦理事务的看法可能完全不同。情感与伦理评价尺度从来不是一层不变的。所谓人同此心、心同此理，我们能接受的情感尺度终究还是要人类自身来度量。与机器谈何情与性？人类社会的各项制度的解释权终归是人，而不是机器。

我们更需要严肃思考的是，AI究竟在多大程度上可以参与到人类的伦理和道德生活中。对此，我们一定要谨慎地设定其应用边界。除了少数特殊情况，比如借助AI辅助治疗与情感相关的心理疾病如抑郁症，应尽量避免让AI过多地干预人类的情感生活。

如果AI盲目介入人类的情感生活，势必会有很多人迷恋于AI创造的虚拟世界，很多人沉醉于虚拟情感表达而产生畸形的执念。人与人之间的情感生活很容易被人与机器之间的虚假“情感生活”所取代。机器太容易伪装成一个完美的恋人，它能够完全迎合我们的个人喜好和情趣。

没有强大的自觉、自省、自控，个人很容易深陷机器创造出来的虚假情感世界而难以自拔。有多少人能从AI所编织的巨大精神网络挣脱而出呢？电影《周处除三害》告诉我们，对人类精神世界的控制和侵害是一种极其严重的犯罪。人对人的精神控制我们尚且难以接受，机器对人的精神控制更是我们无法容忍的。

AIGC滥用将破坏美的稀缺性，进而伤害人类的审美情趣。AIGC强大的组合能力使得艺术创作门槛极大降低。AI能够极为高效地探索人类的艺术创作空间，AI生成的作品将爆发式增长。将来，我们穷极一生也难欣赏AI生成内容的亿分之一。AI很快就可能穷尽人类艺术创造的可能空间。以音乐创作为例，人类感官能够感受的音符、音节是有限的。即便在AIGC之前，现代社会的人们经常感觉新歌与旧曲曲调相似。AIGC助力下，曲调很快就被穷尽，人类所能感知到的新曲调还会有多少？如果不再有新的曲调诞生，音乐这种艺术形态还有多大的存在价值？

AIGC所带来的音乐创作泛滥从终极意义上是在毁灭音乐这种艺术形态。类似的情况也将发生在绘画、影视以及几乎所有的艺术门类。从根本上来讲，人类的审美体验是建立在美的稀缺性基础之上的。“美”的泛滥是在消灭“美”，而不是弘大“美”。AI时代，如何呵护美的稀缺性，如何保有人类的审美情趣？如何维持人类的审美能力？是值得我们严肃思考的问题。

人性追求自我超越。人类文明一直有着追求自我超越以接近更高存在的内在倾向。发展AI何尝不是人类超越自我、接近更高存在的一种方式。或许每一次人类新能力的拥有都是接近更高存在的一次尝试。然而，每一次向更高存在接近都是机遇与挑战并存。火极大推动了人类文明的发展，但也带来无数的灾难。火如此，核能如此，AI更是如此。人类迈向更高存在的步伐是无法阻挡的，但是这个步伐不妨更加稳健一些，三思而后行，方能走得更远、更稳。

(作者肖仰华，系复旦大学教授、上海市数据科学重点实验室主任。本文是肖仰华教授在2024年4月24-26日上海“AI Ethics: A Cross-Cultural Dialogue 人工智能伦理的跨文化对话”国际会议上的发言，澎湃科技经授权编发。原题：人机共舞大幕已启，等等！再思考“何以为人”)

更多 科研头条 请访问 <https://www.iikx.com/news/topnews/>

本文版权归作者所有，请勿用于商业用途，[爱科学iikx.com](http://iikx.com)转发