

# 自信迎接 ChatGPT

► 苏竣

近年来，科学技术发展迅猛，我们在不断地按照摩尔定律的轨迹，一步步趋近物质的极限和自然的本源。在能源领域，可控核聚变在去年取得重大突破，对人类的影响不容忽视。信息技术更是突飞猛进，比如去年爆火的元宇宙，尽管是一种应用层面的创新，也引发了社会的思考和关注。今年以来，ChatGPT 爆红网络，它的问世进一步拉近了我们与人工智能“技术奇点”的距离，也再一次引燃了全球对人类社会智能化转型态势的关注和思考。

分析和评价 ChatGPT 以及其他的新兴变革性技术的社会影响，保持方法论上的正确性是非常重要的。我们应该以三种态度应对我们面临的挑战。第一，审慎。我们要敏感地关注以 ChatGPT 为代表的新技术带来的风险和挑战，以审慎的态度看待一切，既分析它的发展历程，也研判它的未来趋势。第二，自信。要相信人类的智慧，相信人类的光辉思想，相信我们的人文精神，一定会战胜颠覆性技术给我们带来的冲击和挑战。人类历史上技术变革带来的冲击从未停止过，例如为应对核武器的危险，1955年发表的《罗素 - 爱因斯坦宣言》，以及1957年发起的《帕格沃什科学和世界事务会议》，致力于防范核武器对人类的毁灭，寻求人类的可持续发展之道。70年过去了，在全世界的共同努力下，我们把核武器这一科技野兽关在了人类理性的笼子里。今天还是应该保持这种信心，相信人类理性的光辉和智慧能够引领技术向好发展。第三，积极。我们既不能采取鸵鸟政策，更不能采取封堵政策，而应该以积极的心态，采取积极的治理措施拥抱新技术，积极地感知、理解、引导，让新技术更好地发展。

比尔·盖茨称，ChatGPT 问世的重大历史意义，

苏竣，清华大学公共管理学院教授、清华大学智库中心主任、清华大学智能社会治理研究院院长、清华大学科教政策研究中心主任。



不亚于互联网和个人电脑的诞生。业界普遍认为，ChatGPT 是人工智能发展史上的里程碑事件，有望首次为通用人工智能发展探出一条清晰的技术路线，这必将对人工智能技术和产业发展产生巨大的促进作用。但是以 ChatGPT 为代表的先进通用人工智能技术究竟会给我们带来什么样的巨大冲击和系统性变革？在这一点上，很多学者也有很多研究，我也在思考这个问题，在此做一个分享。

## 人类社会的知识生产、传承和校验机制正在被重塑

知识生产是人类社会持续进步的力量源泉。亿万年来，人类演进形成了三种知识生产机制，分别是以亚里士多德、牛顿为代表的基于理性主义的演绎推理机制，以培根、杜威为代表的基于经验主义的实证归纳机制，还有近期学者发现的直觉与顿悟方式。在上述知识生成机制中，人类的主体性、智慧性、能动性奠定了人在知识生产中的核心地位，人类的群体合作行为，特别是学术共同体的同行评议（peer review）成为新知识证伪检验的保障。

人工智能生成内容近年来进展迅猛，表现生成“内容”，实则生成“知识”。ChatGPT 已经可以学习归纳历史数据后进行模仿式、缝合式演绎生成



2023年4月17日，上海市徐家汇地铁站的换乘通道内，两块巨大的LED屏幕展示的多幅AI人工智能创作的绘画作品颇为引人注目。据悉，这些画作都是由Tiamat人工智能AI生成，每幅绘画作品都标注了创作者和创作时使用的关键词。  
中新社 王冈 摄

内容，随技术的迭代成熟，它自动智能化生产新知识的能力愈发强大。这意味着，知识生产将不再是人类的专属技能，打破了人类在知识生产上的垄断格局，推动知识生产主体、生产方式以及知识形态的改变，通过人类的直觉顿悟、逻辑推理、实证实验生成知识，再经由学术共同体批判和校验知识的传统机制可能被彻底颠覆。在未来，如果AI能够作为“超级神灵”生成“绝对真理”，势必会与人类在知识的创造、传承、验证等方面展开一场“大博弈”。

### 人类社会的信任机制正在被颠覆

信任是人类社会的基石。人类之所以能发展出现代文明，就是因为人类通过沟通，建立了信任机制，从而实现了分工合作，奠定了今天这样的一种社会组织方式。传统人类社会的信任体系主要包括人际信任、制度信任和基于信誉的信任等，都是建立在以人为中心的人类自主性基础上。随着经济社会的发展，信任的对象在人、制度和组织之外，还

涵盖了技术、算法、系统、机器等非人格化客体。

然而，ChatGPT的横空出世，迫使我们开始思考AI在信任机制中是否走上“实在论”意义上的主体地位。强人工智能或超人工智能如果能够实现，将作为更先进的社会生产力重塑经济活动和社会交往。其中类人化人工智能的出现使得信任变得愈发脆弱和缺失，甚至将挑战或者取代人类在传统信任体系的中心主体地位，从底层改变传统的信任机制，这是人类步入现代文明后未曾有过的现象。人类与另外的物种、另外的人格化智能体的建立信任的过程，一定是一个充满痛苦的复杂过程。

### 人类社会的权威生成机制正在被解构

马克斯·韦伯认为，权力和权威为社会的稳定有序提供保障。人类有三种权威类型，包括传统权威、克里斯马权威和法理权威。强人工智能的出现，人类中心主义将会被解构，基于人类社会的权威形成机制也将被颠覆。ChatGPT可以对新闻、博客和社交媒体帖子进行数百万次的评论，压倒性地获得社会舆论话语权；强人工智能系统甚至可以找出决策系统中最薄弱的环节，有选择地针对关键立法者和决策者，通过游说、利益交易等方式实现不正当目的或社会控制。

随着人类对AI形成依赖，强人工智能将首先建立起知识权威，逐渐发展出传统权威甚至克里斯马权威。强人工智能的控制者可以由此获取至高无上的权威和绝对垄断的权力，从而影响人类的意识形态，挑战现有治理体系。未来权威从何而来，谁代表了权威，可能都看不清楚了。我们都要经历很长的迷茫期，迷茫的程度或将比往任何一种信息技术的颠覆性创新给我们带来的变革和冲击都要更猛

烈一些。

智能技术已经给我们带来了许许多多的挑战和困惑，ChatGPT 只是今天被大家关注到而已。例如，精准推送导致我们人类今天被包裹在信息茧房里，观点越来越极化，冲突越来越剧烈，对意识形态和政治选举的影响也极大。在智能技术的冲击下，我们人类的主体性、作为人的优越感在智能社会中很可能慢慢的丧失。

尽管 ChatGPT 等 AI 技术会给人类社会带来种种挑战，但这依然是一件好事而不是坏事，因为它代表了先进生产力的发展要求，也警示我们智能化变革愈发紧迫，必须尽快行动，防患于未然。在多年前，我提出了人类社会将要步入智能社会这一观点。从今天各位专家的发言中，我感到大家已经形成了一种共识，就是无论今后人类将面临什么样的变革与挑战，我们都必须不断探索并坚守智能社会的人文精神，致力于建设一个有人文温度的智能社会。

数年前，我与来自多个高校和科研院所的同仁们，包括今天在座的多位专家，在长期研究的基础上，于 2019 年率先发起了“开展人工智能社会实验、探索智能社会治理道路”的倡议。这一倡议得到了学术界和产业界的积极支持，得到了国家政治领导人的高度认可。2022 年 4 月，中共中央、国务院颁布《关于加快建设全国统一大市场的意见》的文件，明确提出要深入开展人工智能社会实验，推动制定智能社会治理相关标准。目前，在中央网信办、科技部、发改委、教育部等部委的组织下，在全国 22 个省建设了 92 个国家智能社会治理实验基地，在全国 18 个人工智能创新发展试验区开展人工智能社会实验，极大推动了人工智能技术在不同领域、不同行业场景中的广泛应用。国家市场监督管理总局在去年 9 月份批准设立了国家一级标准化委员会——智



2023 年 3 月 15 日，上海，一位年轻人访问和试用 OpenAI 最新推出的 GPT-4。美东时间 3 月 14 日周二，OpenAI 正式宣布推出大型的多模态模型 GPT-4。这是其 AI 语言模型系列中的最新产品，可为 ChatGPT 和新 Bing 等应用程序提供支持。中新社 王冈 摄

能技术社会应用与评估标准化委员会。众多高等院校积极推进智能社会治理人才培养，一大批专家学者投身于这项有意义的工作中来了。清华大学智能社会治理研究院整合多学科的力量，围绕人工智能赋能城市治理、数字乡村建设、农村电商的社会影响、赛博空间的精神沉迷、人工智能对儿童成长影响等议题，依循人工智能社会实验的研究路径，开展了长周期的追踪、观测和研究。

今天，世界又站在了历史的十字路口，我们正在经历着从工业社会向智能社会转变的重大历史时期。以人文精神保持初心和定力，在当下显得尤为重要。以人为本，以人民为中心，构建具有人文温度的智能社会，应该是所有学者的核心秉持。我衷心地希望大家共同努力，用我们的智慧和力量，深化智能社会治理研究，深入探索智能社会治理的中国道路，建设有人文温度的智能社会，引领人类命运共同体走向更加光明的未来。

【本文为作者在 2023 年 2 月 18 日首届中国数字经济发展和治理学术年会闭幕式的致辞。】