

· 卷首语 ·

## 大模型技术：变革、挑战与机遇

戴琼海

清华大学 信息科学技术学院, 北京 100084

近年来,随着预训练技术的发展和计算硬件性能的提升,人工智能取得了突破性进展。2022 年底,人工智能大模型 ChatGPT 问世,将自然语言的理解与生成能力提升到新的高度,并具备了跨学科、多场景、多用途的通用性,引起了产业界和学术界的广泛关注。以 ChatGPT 为代表的大模型技术实现了人工智能技术从“量变”到“质变”的跨越,有望发展成为人工智能关键基础设施,发挥“头雁”效应赋能千行百业,加速推进国民经济和社会的高质量发展。

为更好推动大模型技术创新和应用生态的发展,国家自然科学基金委员会信息科学部、《中国科学基金》编辑部与中国人工智能学会共同组织了“ChatGPT 与人工智能技术应用”的专题研讨,围绕“以 ChatGPT 为代表的大型语言模型研究进展”“从 ChatGPT 到多模态大模型:现状与未来”等 10 个主题,形成了详尽的当前形势分析和未来趋势研判,从技术、应用、生态等多个角度阐述了大模型技术引发的新一轮人工智能变革,并指出大模型技术可能带来的风险和挑战,并给出了我国大模型发展的一些启示与展望。

以 ChatGPT 为代表的大模型技术反映了通用人工智能的一个重要发展思路,对人工智能发展的推进,以及对社会经济的重大推动作用毋庸置疑。一方面,大模型技术正在逐步落地应用,在互联网、金融、医疗等领域体现出了显著优势,在其他行业的广泛应用也正在逐步实现;另一方面,大模型技术也对科学研究的方法与范式带来的变革,通过“一专多能”的能力以及丰富的生成功能,可以为科学规律的总结、未知知识的探索等方面带来新的机遇,为生命科学、医学、药学、材料、化学等领域的研究提供有力工具。



**戴琼海** 清华大学教授,中国工程院院士,清华大学信息学院院长,北京信息科学与技术国家研究中心主任,中国人工智能学会理事长。主要学术方向为人工智能、成像与智能技术、脑与认知科学。主持承担了国家自然科学基金认知计算科学中心项目和国家重大科研仪器研制项目等。目前开展的研究包括认知智能、智能光电计算和元宇宙等。

但我们也需要看到:大模型技术仍基于现有深度神经网络,目前依然存在可解释性不足、能效比不高、缺乏决策能力等重要性能瓶颈;重度依赖超大规模数据资源,在典型开放环境下,其可靠性仍存在重要缺陷,尤其在样本极少、先验知识缺乏、环境动态变化、强对抗干扰等极端条件下,现有大模型技术的应用能力受到很大限制;同时,大模型的“幻觉”等效应机理和解决方案仍然有待探索。

面向动态、干扰和无先验的复杂环境感知、理解和决策,要想实现更强的通用性、鲁棒性和泛化性,人工智能技术必须在几个方面实现升级跨越:从单纯的机器学习模型到结构性可解释模型,从离线学习到连续自主学习,从记忆式学习到联想猜测。新的生成式人工智能需要着力推动人工智能与脑科学、认知科学的交叉创新研究,力争在可解释性、高安全性和低功耗性等方面形成重要突破。

大模型技术的发展,对于我们既有挑战、也有机遇。应当在大模型的浪潮中,保持定力,夯实人工智能基础理论研究、加强关键技术攻关、并进一步鼓励大模型相关的产业应用,让我国能引领未来人工智能发展前沿,助力我国国民经济和社会的高质量发展。

**Large AI Model: Change, Challenge, and Chance**

Qionghai Dai

School of Information Science and Technology, Tsinghua University, Beijing 100084