

从基础智能到通用智能: 基于大模型的 GenAI 和 AGI 之现状与展望

缪青海¹ 王兴霞^{1,2} 杨静^{1,2} 赵勇³ 王雨桐² 陈圆圆² 田永林²
俞怡⁴ 林懿伦⁴ 鄢然⁵ 马嘉琪⁶ 那晓翔⁷ 王飞跃^{1,2}

摘要 本文对生成式 AI (Generative artificial intelligence, GenAI) 的国内外发展现状进行了概述, 重点分析了中美之间在算力、数据、算法、生态等方面存在的差距. 为改变我国在生成式 AI 领域的落后现状, 提出高能效算力建设、联邦数据、专业领域模型、基于 TAO 的联邦生态等应对策略, 对大模型时代 AI 安全治理进行了论述, 对通用人工智能 (Artificial general intelligence, AGI) 的未来发展进行了展望.

关键词 生成式人工智能, 大语言模型, 基础模型, 通用人工智能, 基础智能, 平行智能, Sora

引用格式 缪青海, 王兴霞, 杨静, 赵勇, 王雨桐, 陈圆圆, 田永林, 俞怡, 林懿伦, 鄢然, 马嘉琪, 那晓翔, 王飞跃. 从基础智能到通用智能: 基于大模型的 GenAI 和 AGI 之现状与展望. 自动化学报, 2024, 50(4): 674–687

DOI 10.16383/j.aas.c240156

From Foundation Intelligence to General Intelligence: The State-of-Art and Perspectives of GenAI and AGI Based on Foundation Models

MIAO Qing-Hai¹ WANG Xing-Xia^{1,2} YANG Jing^{1,2} ZHAO Yong³ WANG Yu-Tong² CHEN Yuan-Yuan²
TIAN Yong-Lin² YU Yi⁴ LIN Yi-Lun⁴ YAN Ran⁵ MA Jia-Qi⁶ NA Xiao-Xiang⁷ WANG Fei-Yue^{1,2}

Abstract This article provides an overview of the current development status of generative artificial intelligence (GenAI) both in China and abroad, focusing on the analysis of the gap between China and the United States in terms of computing power, data, algorithms, and ecosystems. To change the backward situation of China in the field of generative AI, perspectives such as constructing high-energy-efficient computing, federated data, specialized domain models, and a federated ecosystem based on TAO are proposed. It also discusses AI security governance in the era of foundation intelligence and prospects for the future development of artificial general intelligence (AGI).

Key words Generative artificial intelligence (GenAI), large language model, foundation model, artificial general intelligence (AGI), foundation intelligence, parallel intelligence, Sora

Citation Miao Qing-Hai, Wang Xing-Xia, Yang Jing, Zhao Yong, Wang Yu-Tong, Chen Yuan-Yuan, Tian Yong-Lin, Yu Yi, Lin Yi-Lun, Yan Ran, Ma Jia-Qi, Na Xiao-Xiang, Wang Fei-Yue. From foundation intelligence to general intelligence: The state-of-art and perspectives of GenAI and AGI based on foundation models. *Acta Automatica Sinica*, 2024, 50(4): 674–687

收稿日期 2024-03-29 录用日期 2024-04-12

Manuscript received March 29, 2024; accepted April 12, 2024

国家自然科学基金 (62271485, 61903363, U1811463) 资助

Supported by National Natural Science Foundation of China

(62271485, 61903363, U1811463)

本文责任编辑 孙长银

Recommended by Associate Editor SUN Chang-Yin

1. 中国科学院大学人工智能学院 北京 100049 中国 2. 中国科学院自动化研究所复杂系统管理与控制国家重点实验室 北京 100190 中国 3. 国防科技大学系统工程学院 长沙 410000 中国 4. 上海人工智能实验室 上海 200232 中国 5. 南洋理工大学土木与环境工程学院 新加坡 639798 新加坡 6. 加州大学洛杉矶分校萨穆埃利工程学院 洛杉矶 90095 美国 7. 剑桥大学工程系 剑桥 CB2 1TN 英国

1. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China 2. The State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China 3. College of Systems Engineering, National University of Defense Technology, Changsha 410000, China 4. Shanghai Artificial Intelligence Laboratory, Shanghai 200232, China 5. School of Civil and Environmental Engineering, Nanyang

继 ChatGPT^[1] 之后, 视频生成大模型 Sora 又一次掀起了全球智能化浪潮^[2-3]. Sora 借助跨模态基础模型 (Foundation model) 的强大能力, 初步展现出世界模拟器的潜力, 加速从语言智能向想象智能的转变. 在此背景下, 生成式人工智能 (Generative artificial intelligence, GenAI) 作为实现通用人工智能 (Artificial general intelligence, AGI) 的重要方式之一, 成为各国重点关注的领域. 近年来, 我国高度重视新质生产力的发展, 在生成式人工智能领域也呈现出“百模大战”的繁荣局面, 但在创新首发、最高水准等方面整体水平距离世界先进仍有不小差

Technological University, Singapore 639798, Singapore 6. UCLA Samueli School of Engineering, Los Angeles 90095, U.S. 7. Department of Engineering, University of Cambridge, Cambridge CB2 1TN, U.K.

距. 因此, 应对挑战, 从算力、数据、算法、人才、产业、安全、治理等方面探讨应对策略, 成为当务之急.

本文首先简述国内外生成式人工智能的起源、发展与现状, 在此基础上从数据、算力、算法、生态等方面对中外生成式人工智能发展进行对比、分析. 面对我国生成式人工智能所面临的挑战, 讨论相应的对策, 特别是探讨区块链技术、DAO 与 DeSci 机制对于建立智能联邦生态和基础智能系统的重要意义, 最后展望生成式人工智能的未来发展与社会影响.

1 国外发展现状

1.1 理论与算法概述

基于大模型的生成式人工智能 (GenAI) 涉及多种算法、模型和理论, 而深度学习则是构建这些模型的基石. 深度学习利用多层神经网络来学习输入数据中的复杂模式和规律, 实现了对数据特征的高效提取和表示^[4]. 深度学习的研究历史可以追溯到上世纪 50 年代和 60 年代, 由于当时计算资源有限、数据稀缺以及算法的限制, 深度学习并没有受到广泛的关注和应用. 直到 20 世纪末和 21 世纪初, 随着计算机性能的提升和大规模数据的可用性增加, 卷积神经网络 (CNNs)^[5]、循环神经网络 (RNNs)^[6]、深度信念网络 (Deep belief networks)^[7] 等深度学习模型先后被提出, 尤其是在 2012 年, AlexNet 在 ImageNet 数据集上取得了显著的性能提升^[8], 引发了广泛的关注和热议, 深度学习开始逐渐成为人工智能领域的研究热点. 深度学习的另一个里程碑是 Vaswani 等于 2017 年提出 Transformer^[9], 其引入了自注意力机制 (Self-Attention mechanism), 使得模型能够在不依赖于序列位置的情况下对自然语言序列中的单词进行建模, 有效地捕捉长距离依赖关系. Transformer 不仅被应用于机器翻译、语言建模、文本生成等自然语言处理任务, 还被广泛应用于图片、视频、语音处理与分析等多个领域^[10-12].

随着 Transformer 的出现, 深度学习模型规模迅速扩大, BERT、GPT 等一系列预训练语言大模型在自然语言处理领域取得突破性进展, 在包括文本生成、问答、阅读理解、逻辑推理等几乎所有自然语言处理任务上取得了最高技术水平 (State of the art, SOTA)^[13-22]. 在此基础上, OpenAI 开发了基于 GPT-3.5 及以上版本的 ChatGPT, 由于其出色的对话生成性能, 引发了人们对智能对话系统的广泛兴趣和关注, 成为了近年来研究和应用领域的热点之一. 随着基于 Transformer 的预训练模型在自然语言处理领域取得成功, 该技术在计算机视觉和多

模态预训练模型领域也开始得到应用, 并展现出了出色的性能^[23-24].

另一方面, 扩散模型 (Diffusion models)^[25] 也是一类生成式人工智能模型, 它们通过模拟数据的扩散过程来生成新的样本. 这些模型的工作原理类似于物理学中的扩散过程, 即从高浓度区域向低浓度区域逐渐扩散, 直至达到平衡状态. 在生成任务中, 扩散模型通常从一个随机噪声分布开始, 然后逐步引导这个噪声向目标数据分布转变. 这个过程可以通过迭代的方式进行, 每一步都会使噪声更加接近真实的数据分布. 具体来说, 扩散模型包含两个主要阶段: 正向扩散 (Forward process) 和反向生成 (Reverse process). 扩散模型在图像生成、文本生成等领域表现出色, 能够生成高质量的样本. 与其他生成模型如生成对抗网络 (GANs) 和变分自编码器 (VAEs) 相比, 扩散模型具有独特的优势, 例如更好的样本多样性和更平滑的生成过程.

1.2 典型 GenAI 模型

表 1 列出了国外典型的 GenAI 模型. 在自然语言处理领域, 2018 年是大语言模型 (LLMs) 发展的重要节点. OpenAI 和 Google 相继推出了基于 Transformer 的 GPT-1 和 BERT 模型^[10-11], 为这一领域的发展奠定了基础. GPT 和 BERT 模型的发布引领了大语言模型的发展方向. 随后, RoBERTa 对 BERT 进行了改进和优化, 通过改进掩码方式、增加数据集规模和优化训练方法, 提升了在多个自然语言处理任务上的性能^[26]. BART 结合了编码器解码器结构和掩码语言模型的训练方式, 能够同时处理生成和填充任务^[27]. T5 则通过转换器架构实现了多种任务的统一模型, 进一步提高了自然语言处理任务的效率和性能^[28]. GPT 系列模型采用了自回归生成方式, 其中模型会逐步生成文本, 每个词的生成都依赖于先前生成的内容^[29]. 这种生成方式使得 GPT 模型在文本生成任务中表现突出, 尤其在生成连贯、自然的文本方面效果显著. 2023 年 3 月发布的 GPT-4 具备多模态输入输出能力, 能够处理文字、图像和语音数据^[30]. Google 随后推出了 PaLM 2, 通过强化逻辑和推理训练, 该模型在某些逻辑和推理任务上超越了 GPT-4^[31]. 在开源大语言模型领域, Llama 系列模型是典型代表, Meta 于 2023 年 6 月发布的 Llama 2 在性能上接近于 GPT-3.5^[32-33]. 大语言模型最新的进展是 Anthropic 于 2024 年 3 月发布了 Claude 3, 该模型在多项任务上超越了 GPT-4^[34].

随着大语言模型在自然语言处理领域的成功应用, 人们开始探索将这些模型扩展到计算机视觉领域. 与传统的卷积神经网络 (CNN) 模型相比, Vision Transformer (ViT) 模型通过引入自注意力机

表 1 国外主要 GenAI 模型
Table 1 Typical foreign GenAI models

模型	发布时间	开发者	输入模态				输出模态			
			文本	语音	图像	视频	文本	语音	图像	视频
GPT-1	2018 年 6 月	OpenAI	√				√			
BERT	2018 年 10 月	Google	√				√			
GPT-2	2019 年 2 月	OpenAI	√				√			
RoBERTa	2019 年 7 月	Meta	√				√			
T5	2019 年 10 月	Google	√				√			
GPT-3	2020 年 5 月	OpenAI	√				√			
GPT-3.5	2022 年 3 月	OpenAI	√				√			
GPT-4	2023 年 3 月	OpenAI	√	√	√		√	√	√	
PaLM 2	2023 年 5 月	Google	√				√			
Llama 2	2023 年 6 月	Meta	√				√			
Claude 3	2024 年 3 月	Anthropic	√				√			
MusicLM	2023 年 5 月	Google	√					√		
MusicGen	2023 年 6 月	Meta	√					√		
Voicebox	2023 年 6 月	Meta	√	√				√		
DALL-E	2021 年 1 月	OpenAI	√							√
DALL-E 2	2022 年 4 月	OpenAI	√							√
Stable Diffusion	2022 年 8 月	Stability AI	√		√					√
Midjourney	2022 年 7 月	Midjourney	√							√
Firefly	2023 年 3 月	Adobe	√							√
DALL-E 3	2023 年 9 月	OpenAI	√							√
Imagen 2	2023 年 12 月	Google	√							√
Make-A-Video	2022 年 9 月	Meta	√	√		√				√
Gen-2	2023 年 2 月	Runway	√	√		√				√
Lumiere	2024 年 1 月	Google	√	√						√
Sora	2024 年 2 月	OpenAI	√							√

制, 并利用更大规模数据集进行学习, 在包括图像分类、目标检测、语义分割、图像生成和视频摘要等多个领域实现了显著的精度提升^[35]. OpenAI 在 2021 年 1 月发布了 DALL-E 模型, 该模型根据给定的文本描述生成与描述相匹配的图像. 随后, 扩散模型逐步受到关注, OpenAI 在 2022 年 4 月、2023 年 9 月先后发布了 DALL-E 2 和 DALL-E 3 模型, 进一步提升了图像生成的准确性和创意性^[36-37]. Stability AI 在 2022 年 8 月发布了 Stable Diffusion 模型^[38], 该模型在生成图像时具有更好的稳定性和控制性, 被视为图像生成发展过程中的一个里程碑事件. 相比于图像生成更多地关注单个静态图像的质量和-content, 视频生成除了要考虑静态内容的生成外, 还需要考虑时间维度上的连续性和动态变化, 因而模型需要更深层次地理解时间序列数据, 并能够准确地预测和生成未来的帧. 在近几年的研究中, 视频生成领域涌现出多项引人注目的模型和技术, 为视频内容的创作和生成开辟了新的前景.

Meta 于 2022 年 9 月发布了 Make-A-Video 模型, 该模型以其卓越的视频生成能力为用户提供了高质量视频内容的创作平台^[39]. 随后, 在 2023 年 2 月, Runway 公司推出了 Gen-2 模型, 该模型通过引入创新的算法和训练技术, 进一步提升了视频生成的质量和多样性. Google 于 2024 年 1 月发布了基于扩散模型的 Lumiere, 利用 Space-Time U-Net (STUNet) 架构捕捉视频中的时空信息, 有效提高了视频的长度、运动的连贯性和一致性^[40]. 近期, OpenAI 发布了 Sora 模型, 引发了全球的关注, 该模型生成的画面可以很好地展现场景中的光影关系, 物体间的物理遮挡、碰撞关系^[41].

通过上述分析, 我们认为生成式人工智能 (GenAI) 研发范式与之前的深度学习有着明显的不同之处. 在过去, 深度学习的理论与算法研究主要由大学和学术界参与, 而生成式人工智能的发展主要由包括 OpenAI, Google, Meta, Anthropic, Midjourney 和 Stability AI 在内的多家公司推动. 这些公司在

生成式人工智能领域投入了大量资源, 并通过研发和商业应用推动了技术的进步和应用, 使得生成式人工智能能够更快地从实验室走向市场, 并在各个领域取得广泛的应用和影响。

2 国内发展现状

自 2022 年 11 月 30 日 ChatGPT 面世以来, 国产大模型迅速涌现, 呈现出百家争鸣的态势。如图 1 所示, 不论是工业界还是学术界, 都纷纷推出具有各自特色的大模型, 并持续进行迭代与升级。百度于 2019 年首次推出文心 1.0 通用大模型^[42], 该模型当前已升级至 4.0 版本。阿里云、华为、腾讯、科大讯飞等厂商相继推出通义千问^[43]、盘古^[44]、混元¹、星火²等大模型并向公众开放。在这场“百模大战”中, 不少初创企业也展现出了令人瞩目的成绩, 如 2023 年 4 月成立的北京月之暗面科技有限公司, 其发布的 Kimi 智能助手³ 大模型一经推出即支持高达 20 万汉字的无损上下文长度, 并在最新版本中将此指标扩展至惊人的 200 万汉字, 引起了广泛关注。除此之外, 学术界也在积极参与国产大模型的研究当中, 例如中国科学院自动化所于 2021 年 6 月发布的全球首个图文音 (视觉-文本-语音) 三模态预训练模型——紫东太初大模型⁴, 北京智源研究院的悟道大模型⁵, 以及上海人工智能实验室与复旦大学等单位于 2023 年 6 月联合发布的书生·浦语通用大模型 (InternLM)^[45]。当前, 国产通用大模型已不乏百亿, 甚至千亿参数量级的选手。为了促进国产大模型技术与市场的进一步发展, 一些厂商提供了开源版本的大模型, 例如阿里云的 Qwen 系列开源模型, 零一万物的 Yi 系列开源模型⁶, 以及智谱华章公司的 GLM 系列⁷, 科大讯飞的 iFlytekSpark 系列, 百川智能的 Baichuan 系列^[46] 等。

在专用领域大模型方面, 国内的大模型主要集中在工业、医学、金融、交通、教育、农业、传媒、司法、AI4S 等领域^[47-63], 如图 1 所示。

1) 在 AI4S 领域, 华为盘古气象大模型基于 39 年的全球气象数据开发了基于三维深度神经网络和地球特定先验知识的实时气象预测系统, 在预测精度上首次超越传统的基于数值的预测方法^[64]。华东师范大学的研究者构建了 ChenGPT1.0⁸, 可以实现

化学专业知识、生物医药以及通用领域的知识问答。上海交通大学发布的白玉兰科学大模型 BAI-Chem⁹, 能够实现快速高效的药物分子设计, 并在逆合成、产率预测、催化剂推荐等任务上取得了较高准确率。

2) 在教育领域, 华东师范大学和北京语言大学在预训练大模型的基础上, 分别建立 EduChat^[65] 和桃李教育大模型¹⁰, 能够实现考题设计与解答、课程咨询问答等功能。学而思开发了九章大模型 (MathGPT)¹¹ 能够解决数学课程中的习题并提供解题思路。网易有道发布了子曰教育大模型¹², 支持听说读写等多任务的智能培训。

3) 在工业领域, 中国科学院大连化物所发布了化工大模型¹³, 能实现化工知识问答、化工流程自动生成等功能。中工互联公司针对工业领域的故障预测、智能维护、方案优化等任务, 构建了“智工”大模型¹⁴, 成为中国工业领域首个开源大语言模型。中煤科工西安研究院利用矿井信息系统, 构建了地质垂直领域大模型——GeoGPT¹⁵, 实现地质行业知识问答, 并成功应用于透明地质保障任务中。

4) 在医学领域, 华南理工大学构建了生活空间健康大模型“扁鹊” (BianQue)^[66] 和心理健康大模型“灵心” (SoulChat)^[67], 分别用于生理健康诊断问答以及心理健康咨询。香港中文大学基于中文问诊数据, 构建了华佗 GPT 模型^[68], 用于智能医疗问答。浙江大学的研究团队构建了启真医疗大模型¹⁶。医联科技推出了 medGPT 大模型¹⁷, 能够对近 3 000 种疾病进行初步诊断。在中医药大模型方面, 百度健康与固生堂联合发布了岐黄问道大模型¹⁸; 华东师范大学研发神农中医药大模型 (ShenNong-TCM)¹⁹; 复旦大学和同济大学合作研发了仲景中医大语言模型 (CMLM-ZhongJing)²⁰。

5) 在金融领域, 中国科学院成都计算机应用研究所开发了用于金融知识问答的大模型“聚宝盆”²¹。度小满发布了中文金融领域的首个千亿级开源大模型“轩辕”^[69]。恒生电子发布了金融大模型 LightG-

⁹ https://www.seiee.sjtu.edu.cn/index_news/8667.html

¹⁰ <https://github.com/bleucall/taoli>

¹¹ <https://www.mathgpt.com/>

¹² <https://ziyue.youdao.com//home>

¹³ http://dicp.cas.cn/xwdt/kyjz/202403/t20240324_7050498.html

¹⁴ <http://www.ciitcc.com/ciigt>

¹⁵ <http://www.cctegxian.com/html/news/2023-12-18/4185.html>

¹⁶ <https://github.com/CMKRG/QiZhenGPT>

¹⁷ <http://web-qa.medlinker.com/pc/product/medgpt>

¹⁸ <http://www.dajingtc.com/node/21>

¹⁹ <https://github.com/ywjawmw/TCMEB>

²⁰ <https://github.com/SupritYoung/Zhongjing>

²¹ <https://github.com/jerry1993-tech/Cornucopia-LLaMA-Fin-Chinese.git>

¹ <https://hunyuan.tencent.com/>

² <https://xinghuo.xfyun.cn/>

³ <https://kimi.moonshot.cn/>

⁴ <https://taichu-web.ia.ac.cn/>

⁵ <https://www.baai.ac.cn/>

⁶ <https://www.lingyiwanwu.com/>

⁷ <https://zhipuai.cn/>

⁸ <https://www.ecnu.edu.cn/info/1426/65145.htm>

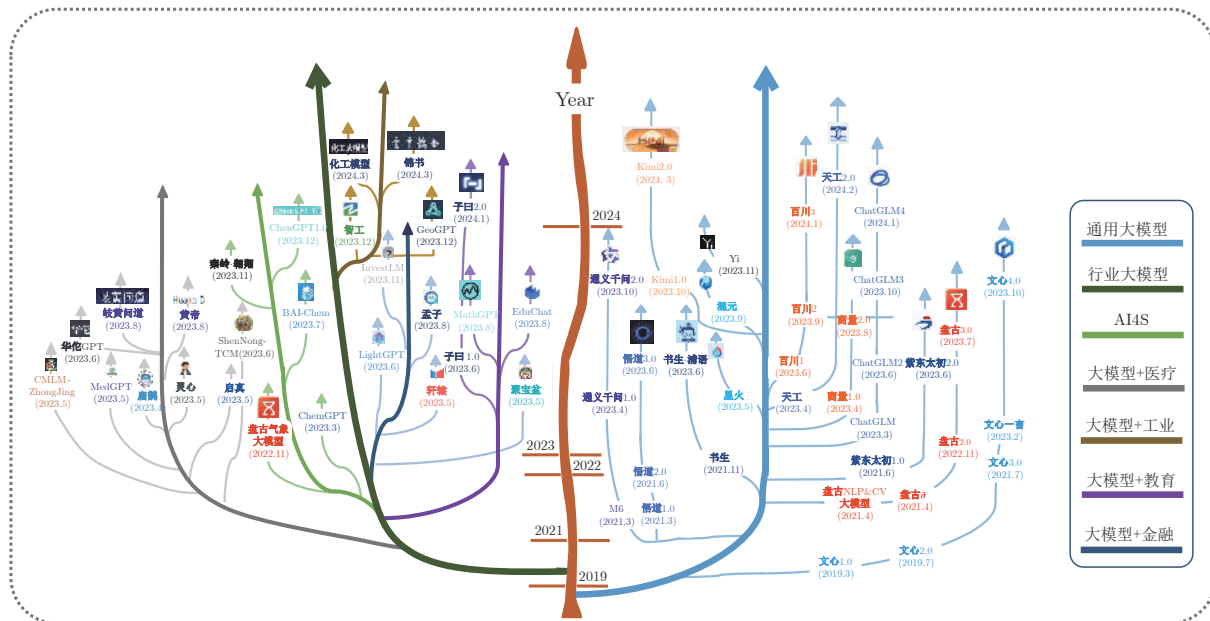


图1 国产大模型发展全景

Fig.1 Panorama of the development of domestic large models

PT²². 澜舟科技发布了以金融为核心应用的孟子大模型²³. 香港科技大学研发了面向投资的金融大模型 InvestLM²⁴.

政策面上, 为了国产大模型的规范化和健康发展, 国家网信办联合国家发展改革委、教育部、科技部等七部门联合发布《生成式人工智能服务管理暂行办法》, 规定对生成式 AI 大模型实施“备案制”管理. 此外, 为了满足当前国产大模型发展所引发的巨大算力需求, 我国陆续出台了《全国一体化大数据中心协同创新体系算力枢纽实施方案》、《算力基础设施高质量发展行动计划》、《“十四五”数字经济发展规划》等一系列文件推动算力基础设施建设.

为了增强对蓬勃发展的国产大模型的认识, 新华社研究院中国企业发展研究中心于 2023 年发布了三版《2023 年人工智能大模型体验报告》, 回顾了大模型产品市场的激烈竞争, 并对当前大模型产品和厂商进行了整体测评. 同时, 多个机构也推出了各自的大模型评测体系和排行榜, 如“司南” (Open-Compass)、C-Eval、SuperCLUE 等, 从中不难看出, 厂商间的竞争激烈, 得分频繁变动, 这不仅展现了国产大模型的竞争态势, 也彰显了该行业的活力与吸引力.

3 国内外对比

本节在分析国内外进展的基础上, 从算力、数

据、算法以及产业生态等四个维度对中美生成式 AI 领域现状进行对比. 由表 2 可见中美之间存在显著差异, 美国领先优势明显.

表 2 中美生成式 AI 领域现状
Table 2 Current status of generative AI fields in China and the United States

对比条目	国内	国外	
算力	处理器	昇腾 910	H100 SXM
	速度 (FP16)	280 TFLOPS	1979 TFLOPS
	显存	16 G	80 G
	互联带宽	—	900 GB/s
	并行平台 API	CANN	CUDA
数据	政府部门	有限开放	能开尽开
	社会力量	碎片、孤立	联合、开源
	数据生态	尚未形成	趋于完善
算法	基本算法	—	Transformer
	语言模型	文心一言等	GPT/Llama
	文生图	秒画等	DALL-E/Imagen
	文本生成视频	—	Sora
	多模态	悟道等	GPT-4 等
生态	独角兽量/值 [#]	70/1.3T	315/5.9T
	高引论文 [#]	60+	220+
	机构/人才占比 [*]	14%/13.85%	55%/56.55%
	创新城市数量 [*]	19	33

注: [#] State of AI Report 2023, ^{*}Aminer

首先, 在算力方面, 美国企业英伟达占据全球 AI 计算的垄断地位, 而华为在中国 AI 计算领域

²² <https://www.hundsun.com/lightgpt>

²³ <https://www.langboat.com/portal/mengzi-gpt>

²⁴ <https://github.com/AbaciNLP/InvestLM>

最为领先。目前, 双方主力计算卡分别是英伟达的 H100(SXM)²⁵ 和华为的昇腾 Ascend 910 系列²⁶。以 AI 计算基准精度 FP16 来看, H100(SXM) 算力达到了 1979 TFLOPS, 华为 Ascend 910 为 280 TFLOPS; 以大模型训练所倚重的显存来看, H100(SXM) 具有 80 G 第三代高带宽显存 HBM3, 华为 Ascend 910 为 16 G DDR4, 而内存类型的不同带来更大差距; 卡间互联数据传输带宽是训练大模型的重要瓶颈之一, H100(SXM) 凭借英伟达独有的 NVLINK 技术具有 900 GB/s 的带宽, 昇腾 910 的带宽参数不详。除此之外, 驱动、算子及 API 接口是发挥硬件设备并行计算能力的支撑要素, 是 AI 算法的基石。英伟达的计算统一设备架构 CUDA, 自 2006 年创建后不断迭代、性能完善、用户众多, 在此基础上发展出 CuDNN 等专用库, 已形成软硬件协同的生态优势。而华为自研的异构计算架构 CANN 自 2018 年发布, 起步较晚, 功能正在逐步完善。需要注意的是, 英伟达在 2023 年中发布了 H200²⁷, 第三代 HBM 显存容量提升到 141 G, 大模型 (Llama2 70B) 的推理速度相比 H100 翻倍。2024 年 3 月, 英伟达发布新一代 GB200 Grace Blackwell 计算核心²⁸, 半精度 FP16 算力达到 10000 TFLOPS, 显存容量达到 384 G, 大模型训练性能大幅提升。华为的升级版 Ascend 910B 和 920 系列在算力、显存方面都有提升, 但具体数据没有公布。在短期内, 国内外算力差距进一步拉大。

其次, 在训练数据方面, 美国已经形成了政府、社会之间职责清晰、分工明确、相互合作的运行机制²⁹。一方面, 政府建立数据开放平台, 负责维护和管理, 确保数据质量, 提高可用性, 降低公众使用门槛。例如, PubMed³⁰ 论文检索系统包含了超过 3600 万条生物医学文献, 国家气象和海洋局、联邦和地方法院也都实现了数据公开, 并提供了 API 接口供调用。另一方面, 社会力量 (开源/非盈利组织、互联网公司研究部门、学界和多类型机构等) 整合政府的开放数据和网络的公开数据, 提升数据精细度和专业性, 形成以开源为主的高质量训练语料。例如 Eleuther AI 开发的 825 GB 英文语料库 The Pile³¹, DeepMind 联合 33 家学术实验室创建了 Open X-Embodiment 开源数据集等。此外, 在政府与社会力量协同方面, 美国联邦政府采取“应开尽

开”的原则, 发挥了 AI 训练数据“汇聚融合”的作用, 建立统一的数据汇聚标准, 推动多方协作的数据资源开发利用, 促进多方数据融合。我国公共数据包括各级行政机关获取的数据, 范围较广, 但在开放共享和开发利用上存在不足, 例如天气数据在数据范围和历史跨度上有较大限制, 法律领域的裁判文书网由社会公开转为仅供内网查询。社会力量主要结合海外优质开源数据集, 中文语料主要来源于网络公开数据, 政府数据贡献少。总体而言, 中文语料库的量、质、开源情况都不如英文, 这导致了训练大模型的困难, 其中公共数据开放不足和企业合规意识低是主要原因之一。相比之下, 与美国政府和社会力量协同开放数据以服务于训练语料的模式有所不同, 我国尚未形成对大模型提供有效数据资源的生态, 存在数据资源碎片化和流通机制不完善的问题。

第三, 在算法、模型的创新方面, 美国处于全面领先。2017 年谷歌提出的 Transformer 已经成为生成式 AI 的核心模块。在大语言模型方面, 谷歌的 BERT 和 OpenAI 的 GPT 最先开启了新时代, 特别是 OpenAI 发布的 ChatGPT 成为 AI 发展路程上的一个里程碑。而后续的 GPT-4 以其在多数任务指标上的优异表现, 在众多语言大模型中仍然占据榜首。Meta 开源的 Llama 大模型虽然稍晚、性能稍逊, 但是因其开源属性, 成为世界上众多大模型的源头。在文本生成图像方面, OpenAI 的 DALL-E 和谷歌的 Imagen 在 2022 年最先发布, 以其高真实度的效果迅速引起了全球范围的注意, AIGC 进入新时代。在文本生成视频方面, OpenAI 在 2024 年 2 月发布 Sora, 再次取得历史性突破。在大语言模型领域, 国内的百度、智源、中国科学院等单位具有先发优势, 推出文心一言、GLM 等大模型。2022 年以来在开源力量等因素的推动下, 国内形成“百模大战”的繁荣局面, 但整体上的性能没有超过 GPT-4 等国外模型。而在文本生成视频方面, 国内与 Sora 的差距更为明显。

最后, 在企业、学术、人才等产业生态方面, 美国仍然全面的领先。根据 State of AI Report 2023³² 统计数据, 初创企业“独角兽”是引领 AI 创新的重要力量, 2022 年以来, 美国的 AI 初创企业累计 315 家, 总估值 5.9 万亿美元, 而同期中国初创企业 70 家, 估值 1.3 万亿美元, 是美国的 22%。在顶级研究机构 and 人才方面, 根据 AMiner³³ 统计数据, 美国占比超过 55%, 而中国不足 14%, 约是美国的四分之一。从过去三年发表的高被引论文数量来看, 美国 220 多篇, 中国 60 多篇, 只是美国的四分之一稍

²⁵ <https://resources.nvidia.com/en-us-tensor-core/nvidia-tensorcore-gpu-datasheet>

²⁶ <https://e.huawei.com/cn/products/computing/ascend>

²⁷ <https://www.nvidia.com/en-us/data-center/h200/>

²⁸ <https://www.nvidia.com/en-us/data-center/gb200-nvl72/>

²⁹ 阿里研究院《中美大模型的竞争之路: 从训练数据讲起》报告。

³⁰ <https://pubmed.ncbi.nlm.nih.gov/>

³¹ <https://pile.eleuther.ai/>

³² <https://www.stateofai/>

³³ <https://www.aminer.org/>

多。在 AI 创新活力城市数量来看, 美国 33 座城市入选, 而中国是 19 座, 约为美国的 60%。

综合分析, 我国在算力设施、数据质量、技术水平、创新能力、资金投入、人才储备、产业发展、生态建设等方面全面落后。结合国际环境等因素, 缩短差距面临诸多挑战。

4 安全与治理

技术进步拓展了人工智能在复杂任务处理中的能力, 但同时也带来了前所未有的安全风险。这些风险遍及人工智能全生命周期的各个阶段, 包括数据采集和标注、模型研发和训练、模型部署和上线、模型开放和应用, 如图 2 所示。随着应用领域增多, 其风险影响范围也从技术本身扩展到特定行业领域, 甚至覆盖全社会群体。例如, 数据采集和标注阶段, 风险从数据本身的质量, 发展到对知识产权的侵害和个人隐私的泄露; 模型研发和训练阶段, 风险从模型鲁棒性和有效性, 发展到可解释性和可审计性; 模型部署和上线阶段, 风险从代码安全、网络安全问题, 发展到价值观与意识形态、虚假信息和幻觉等问题; 模型开放和应用阶段, 风险从对技术本身的信任问题, 发展到强人工智能可控及造成灾难性风险 (战争、失业、社会动荡) 问题。同时, 新型风险如侵犯隐私、帮助犯罪等, 也随着智能体交互协作等技术的发展和应用, 呈现出影响范围快速扩大的趋势, 引起各国政府、科研机构、企业等社会各层面的高度重视。因此, 采取全方位、体系化的安全治理措施以应对风险已成为国内外共识。

4.1 国际安全治理现状

当前, 各国公司和机构往往使用不同的安全维

度以定义和应对人工智能技术风险。例如, OpenAI 将“对齐”作为其发展通用人工智能的核心理念, 强调“有用、真实、无害”作为其指导原则^[70]。斯坦福大学的 HELM 模型侧重于技术细节, 从准确性、不确定性/校准、鲁棒性、公平性、偏见、毒性和推断效率等维度全面评估模型性能及其社会影响^[71]。欧盟强调安全治理维度应包括人权尊重、技术与安全、隐私与数据治理、透明度、多样性、非歧视与公平、社会与环境福祉以及问责制等^[72]。美国则侧重责任、公平、可追溯、可靠、可治理性等要求, 注重人工智能系统的透明度、责任归属和治理能力。不同治理维度反映出国际社会在人工智能安全方面的多元化关注, 旨在确保技术发展促进社会福祉的同时, 又能有效避免负面影响。

相应地, 国际实践采用了一系列创新方法以确保人工智能的安全发展。水印技术广泛应用于保护知识产权和个人隐私, 通过在数据或模型中嵌入隐形标记, 确保产权可追溯且数据合法使用^[73]。安全对齐技术通过确保人工智能系统的价值观与其研发机构所在的社会价值观一致, 以避免数据可能带来的偏见或歧视。人类反馈强化学习通过整合人类反馈来优化模型决策, 确保人工智能行为更加符合人类期望和道德标准^[74]。对抗样本防御机制致力于提升模型鲁棒性, 确保人工智能系统能抵御恶意输入, 维持决策的稳定和可靠^[75]。这些方法已在多个安全维度上取得显著成效。

4.2 国内安全治理现状

在国内, 人工智能安全治理的重心放在大模型的部署与应用阶段, 主要关注大模型部署应用过程中的内容管理与合规性, 以及在产业融合中可能出

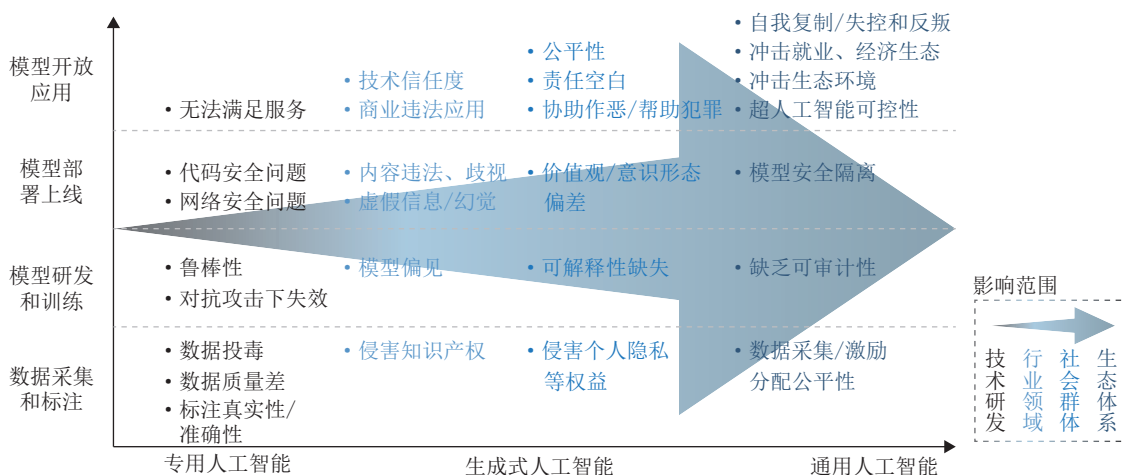


图 2 人工智能全生命周期四阶段主要风险, 风险影响范围随技术发展逐渐增大

Fig.2 The main risks of the four stages in the artificial intelligence lifecycle, and the risk impact gradually increases with technological development

现的安全风险. 依据《生成式人工智能服务管理暂行办法》等规定, 国内治理维度聚焦违反社会主义核心价值观内容、歧视性内容、商业违法违规、侵犯他人合法权益、无法满足特定服务类型的安全要求、模型应拒答的内容及模型不应拒答的内容等七大维度^[76], 旨在确保人工智能在实际应用中的安全性、合法性和道德性, 反映了国内在人工智能治理方面的具体需求和重点关注领域.

针对风险治理, 国内采取了覆盖法律法规、内部治理、第三方评测及针对特定产业垂域的综合措施, 架构如图 3 所示. 法规层面, 我国发布了《生成式人工智能服务管理暂行办法》等一系列法律法规和指导性文件, 确保人工智能发展符合伦理与法治要求^[77]. 内部治理重视治理工具开源、测评数据集构建, 研发红蓝攻防测试和复杂场景沙箱仿真环境测试等技术, 增强模型的安全性和鲁棒性^[78-82]. 第三方评测通过共建数据激励和开放组件, 鼓励多方参与和技术共享, 提升评测的透明性和客观性^[83]. 针对车联网、自动驾驶和工业互联网等产业垂域, 采用沙箱测试、多智能体协同测评与认证等手段, 确保模型在复杂应用场景中的安全可靠^[84-88].

国内外在人工智能安全治理的实践差异主要源

于产业结构和应用重点的不同. 国际上, 主要是欧美国家, 源于其在创意产业、法律法规、技术研发等领域的先进地位, 其治理手段更多聚焦于技术创新的开发与实施、价值观对齐及模型的公平性、透明度和可追溯性, 反映出对数据隐私、产权保护^[89-90]等问题的重视. 相对而言, 国内治理则更侧重于大模型的内容管理、合规性以及满足产业特定需求, 这与中国全球首位的网民数量、高度覆盖居民生活方方面面的数字产业现状密切相关. 此外, 国内在法律法规的制定上表现出更多的明确性和具体性, 为人工智能的健康发展提供了法律框架和政策指导. 这种差异源于展示了技术发展与应用在全球范围内的多样性, 也体现了不同文化背景和社会价值观的影响.

未来的安全治理将朝着形成框架性风险体系的方向迈进, 通过推行人工智能分级分类制度, 为不同行业提供定制化的安全解决方案, 保障交通、医疗、教育、金融等关键领域的应用安全; 更为关键的是, 人工智能安全不仅是技术领域的挑战, 更是关乎全人类福祉的全球性议题, 国际合作、全社会参与将是核心. 必须不断加强国际合作, 推动弥合各国数字鸿沟, 应对全球人工智能安全挑战; 开展产

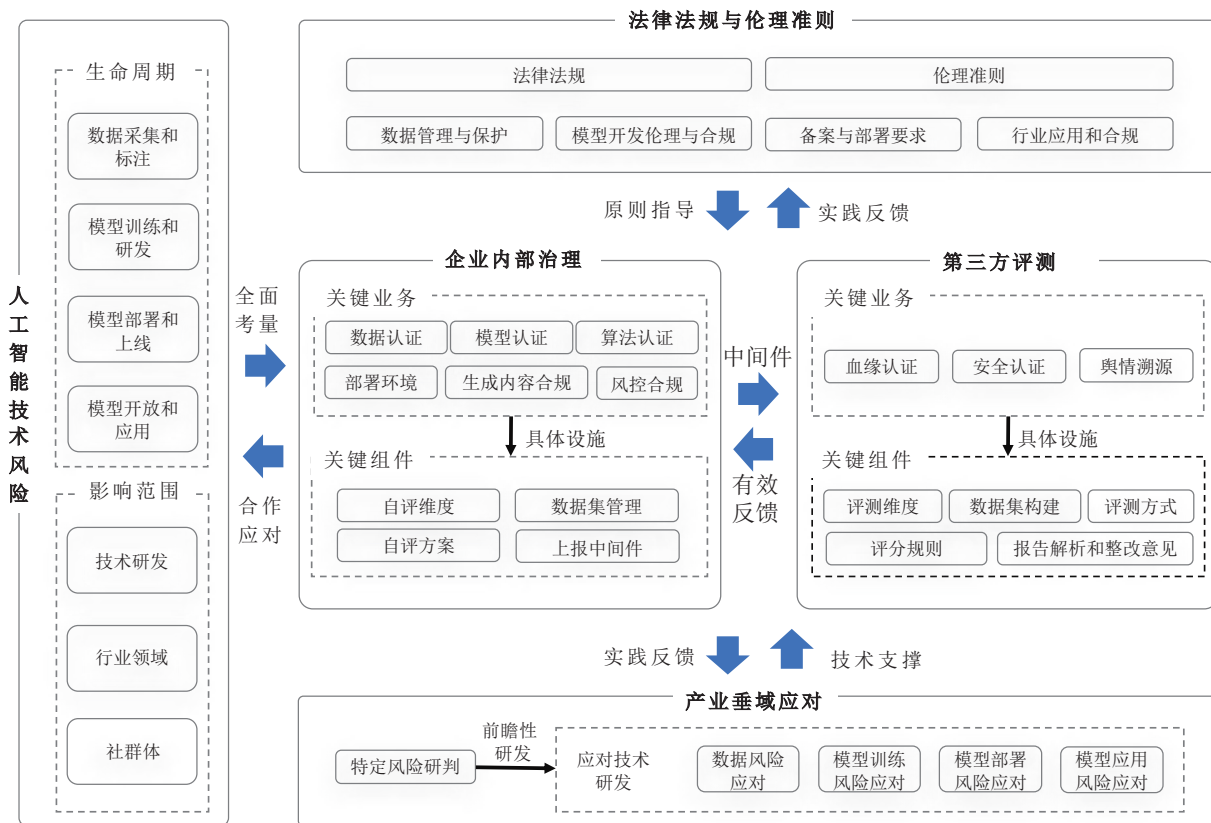


图 3 国内人工智能安全评估体系

Fig. 3 Artificial intelligence safety evaluation system in China

学研合作与安全社区构建,促进安全技术创新和公众信任,确保技术的普惠性和持续发展。

5 挑战与对策

在生成式人工智能领域,我国在算力、数据、人才、产业、安全、治理、生态等方面面临系统性挑战。

首先,算力短缺是我国当前面临的首要挑战。针对这一困境,我国应统筹发展高效 AI 计算硬件体系:一方面,大力研发专门用于生成式 AI 的处理器,突破传统的体系结构,同时特别加强芯片制造工艺的投入,形成完整的微电子产业链,完善智能计算底层算法与软件。另一方面,注重 AI 计算的能效设计,优化算力与带宽的损耗,同时将可再生能源,如水电、太阳能和风能,整合到 AI 计算设施中,参考比特币“挖矿”的算力集中模式,发挥我国西部可再生能源的优势,推广“东数西算”,使生成式 AI 计算更具经济性,同时带动国家全域发展。

其次,数据的质与量直接决定着大模型的性能。面对我国 AI 数据存在的诸多问题,联邦数据^[91-93]是有效应对方案。联邦数据在基于人工智能的应用中扮演重要角色,可有效解决数据资源短缺、碎片化、不共享等挑战。它将大数据转化为数据智能,通过数据实验探索最佳解决方案,可有效应对数据安全和隐私问题。联邦数据架构包括真实数据、虚拟数据、联邦数据实验、联邦融合、联邦安全和可信联邦智慧六个组成部分。实际数据和数字孪生体生成的虚拟数据被处理和检索,联邦数据实验利用强化学习和并行强化学习找到最佳模型参数,联邦融合将本地模型合成全局最优模型。数据安全和隐私由联邦安全组件实现,确保数据所有权和使用权分离。通过联邦数据实验和联邦融合,可获得可信联邦智慧,提高人工智能的自学习能力。

最后,弥补中外人工智能的差距离不开可持续发展的生态环境。AI 人才创新 AI 方法, AI 方法促进 AI 产业, AI 产业促进社会繁荣,富裕、开放、公平的社会环境培育和吸引更优秀的人才。中国科学院王飞跃研究员提出的“联邦生态”为创建这一良性循环提供建设性框架^[93]。除了联邦数据外,联邦生态系统还包括联邦控制和联邦管理,其最终目标是提供联邦服务。联邦控制^[94]是联邦生态的核心执行部分,旨在维护信息安全,保护数据的所有权、控制权、隐私权和使用权。它通过分布式控制策略,为复杂系统提供高效、安全、可靠的控制和管理,确保各子系统的信息安全和权益保护。联邦控制通过定义联邦合约实现数据的联邦,建立联邦数据的控制体系,包括存储、传输、共享和使用,同时保证数据安

全。联邦管理^[95]根据联邦生态的总体目标和要求做出管理决策,并根据系统状态实时动态调整,以达到最优状态,实现整个生态系统的智能化管理。通过联邦管理决策,联邦生态能够更好地实现其目标,在保证安全的前提下转化为联邦智能安全。联邦服务是通过对联邦数据的控制和管理实现的,是联邦管理的最终目的。它利用设计的联邦管理规则,在保障节点数据安全和隐私的前提下,实现对联邦数据的管理和控制,不断产生新的数据,用于优化管理决策。

近年来,区块链技术、去中心化自治组织和分布式自治运营(DAO)机制在 DeSci 浪潮推动下成为 AI 生态新引擎^[96-100]。DeSci 旨在建立一个去中心化、透明和安全的网络,供科学家共享数据、信息和研究成果。去中心化特性使科学家能以更公平、民主的方式合作。DAO 是 DeSci 的实现方式,为人工智能创新和应用提供了新的组织形式。DAO 是一种通过代码在区块链网络上运行的数字组织,成员拥有决策和执行权,实现了民主和透明。基于 DAO 的工作显示出未来发展潜力。区块链技术在 DeSci 中扮演重要角色,确保了真实、可信、可靠、可用和高效。联邦控制依赖于区块链技术的支持,实现了联邦安全、共识、激励和合约。联邦安全利用区块链的安全机制,在加密、传输和验证联邦数据时发挥作用。联邦共识确保了联邦节点之间的分布式共识。联邦激励设计了快速、稳定、正向的激励措施,平衡了节点之间的利益,提高了系统效率。联邦合约通过智能合约算法自动、安全地实现了联邦控制。

联邦生态可以全面处理从数据生产到数据使用再到服务用户的全过程问题。它基于智能生态系统研究思想,能将数据转化为智能,适用于各种联邦形式。联邦生态通过建立松散的联盟关系,加强隐私保护,调动积极性,提高参与度,以提高整体绩效。区块链的 TRUE 和分布式自治管理之 DAO 相结合即为“真道”(TAO),形成完整的可信 AI 生态系统^[101-102]。

6 总结与展望

伴随 GenAI 的通用智能技术发展,未来的智能社会和产业将围绕 Karl Popper 的三个世界^[103]之大模型展开。物理世界之机器人、心理世界之生物人、人工世界之数字人将按 Pareto 法则组成一个崭新的世界:80% 以上是数字人、15% 以下是机器人、5% 以下是生物人^[104-106]。受 GenAI 驱动,数字人完成认知和决策任务,机器人主要负责物理劳动和操作,生物人依托数字人和机器人进行更高层次的创新和管理。

工作将以三种模式由三种员工协调完成: 上午 AM (Autonomous mode, 自主模式), 数字人在机器人的辅助和生物人的关注下完成绝大部分工作, 时长 20 小时以上; 下午 PM (Parallel mode, 平行模式), 数字人和机器人在生物人的远程或云端指导下完成大部分工作, 时长三小时以下; 晚间 EM (Expert or emergency mode, 专家或应急模式), 生物人必须赶赴现场, 在数字人和机器人的帮助下完成工作, 时长一小时以内. 新的工作模式和分工方式可以减轻劳动强度、提高生产效率和扩大服务范围, 有效应对当前面临的人口负增长和老龄化的问题, 同时将重塑人类的工作和生活方式. 未来人类将立法禁止“996 (码农)”的工作方式, 演化为“1023 智侬节”: 每天早上十点上班, 下午二点下班, 每周工作四天, 每年从一月二十三日开始工作, 十月二十三日结束工作, 启动一年一度的智侬节, 希望其他节日假期依然保留.

References

- 1 Wang F Y, Miao Q H, Li X, Wang X X, Lin Y L. What does ChatGPT say: The DAO from algorithmic intelligence to linguistic intelligence. *IEEE/CAA Journal of Automatica Sinica*, 2023, **10**(3): 575–579
- 2 Wang F Y, Miao Q H, Li L X, Ni Q H, Li X, Li J J, et al. When does sora show: The beginning of TAO to imaginative intelligence and scenarios engineering. *IEEE/CAA Journal of Automatica Sinica*, 2024, **11**(4): 809–815
- 3 Yu H K, Liu X Y, Tian Y L, Wang Y T, Gou C, Wang F Y. Sora-based parallel vision for smart sensing of intelligent vehicles: From foundation models to foundation intelligence. *IEEE Transactions on Intelligent Vehicles*, DOI: [10.1109/TIV.2024.3376575](https://doi.org/10.1109/TIV.2024.3376575)
- 4 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, **521**(7553): 436–444
- 5 LeCun Y, Boser B, Denker J S, Henderson D, Howard R E, Hubbard W, et al. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1989, **1**(4): 541–551
- 6 Graves A, Schmidhuber J. Supervised sequence labelling with recurrent neural networks. In: Proceedings of the Conference and Workshop on Neural Information Processing Systems. 2009.
- 7 Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets. *Neural Computation*, 2006, **18**(7): 1527–1554
- 8 Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems. Lake Tahoe, USA: Curran Associates Inc., 2012.
- 9 Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc., 2017. 5998–6008
- 10 Devlin J, Chang M W, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Minneapolis, USA: ACL, 2019. 4171–4186
- 11 Radford A, Narasimhan K, Salimans T, Sutskever I. Improving Language Understanding by Generative Pre-Training. OpenAI Technical Report, 2018.
- 12 Dai Z H, Yang Z L, Yang Y M, Carbonell J, Le Q, Salakhutdinov R. Transformer-XL: Attentive language models beyond a fixed-length context. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy: ACL, 2019. 2978–2988
- 13 Du H P, Teng S Y, Chen H, Ma J Q, Wang X, Gou C, et al. Chat with ChatGPT on intelligent vehicles: An IEEE TIV perspective. *IEEE Transactions on Intelligent Vehicles*, 2023, **8**(3): 2020–2026
- 14 Gao Y B, Tong W, Wu E Q, Chen W, Zhu G Y, Wang F Y. Chat with ChatGPT on interactive engines for intelligent driving. *IEEE Transactions on Intelligent Vehicles*, 2023, **8**(3): 2034–2036
- 15 Guo C, Lu Y, Dou Y, Wang F Y. Can ChatGPT boost artistic creation: The need of imaginative intelligence for parallel art. *IEEE/CAA Journal of Automatica Sinica*, 2023, **10**(4): 835–838
- 16 Wang S Y, Zhu Y X, Li Z H, Wang Y T, Li L, He Z B. ChatGPT as your vehicle co-pilot: An initial attempt. *IEEE Transactions on Intelligent Vehicles*, 2023, **8**(12): 4706–4721
- 17 Wang F Y. Linguistic intelligence for intelligent vehicles: ChatGPT and future logistics and mobility. *IEEE Transactions on Intelligent Vehicles*, 2023, **8**(3): 2011–2019
- 18 Tian Y L, Li X, Zhang H, Zhao C, Li B, Wang X, et al. VistaGPT: Generative parallel transformers for vehicles with intelligent systems for transport automation. *IEEE Transactions on Intelligent Vehicles*, 2023, **8**(9): 4198–4207
- 19 Yu H, Wang Y T, Tian Y L, Zhang H, Zheng W B, Wang F Y. Social vision for intelligent vehicles: From computer vision to foundation vision. *IEEE Transactions on Intelligent Vehicles*, 2023, **8**(11): 4474–4476
- 20 Xue X, Yu X N, Wang F Y. ChatGPT chats on computational experiments: From interactive intelligence to imaginative intelligence for design of artificial societies and optimization of foundational models. *IEEE/CAA Journal of Automatica Sinica*, 2023, **10**(6): 1357–1360
- 21 Wang F Y, Li J J, Qin R, Zhu J, Mo H, Hu B. ChatGPT for computational social systems: From conversational applications to human-oriented operating systems. *IEEE Transactions on Computational Social Systems*, 2023, **10**(2): 414–425
- 22 Lu Jing-Wei, Guo Chao, Dai Xing-Yuan, Miao Qing-Hai, Wang Xing-Xia, Yang Jing, et al. The ChatGPT after: Opportunities and challenges of very large scale pre-trained models. *Acta Automatica Sinica*, 2023, **49**(4): 705–717 (卢经纬, 郭超, 戴星原, 缪青海, 王兴霞, 杨静, 等. 问答 ChatGPT 之后: 超大预训练模型的机遇和挑战. *自动化学报*, 2023, **49**(4): 705–717)
- 23 Tian Yong-Lin, Wang Yu-Tong, Wang Jian-Gong, Wang Xiao, Wang Fei-Yue. Key problems and progress of vision transformers: The state of the art and prospects. *Acta Automatica Sinica*, 2022, **48**(4): 957–979 (田永林, 王雨桐, 王建功, 王晓, 王飞跃. 视觉 Transformer 研究的关键问题: 现状及展望. *自动化学报*, 2022, **48**(4): 957–979)
- 24 Yang Z L, Dai Z H, Yang Y M, Carbonell J, Salakhutdinov L Q V. XLNet: Generalized autoregressive pretraining for language understanding. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc., 2019. Article No. 517
- 25 Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. In: Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc., 2020. Article No. 574
- 26 Liu Y H, Ott M, Goyal N, Du J F, Joshi M, Chen D Q, et al. RoBERTa: A robustly optimized BERT pretraining approach. arXiv preprint arXiv: 1907.11692, 2019.
- 27 Lewis M, Liu Y H, Goyal N, Ghazvininejad M, Mohamed A, Levy O, et al. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. ACL, 2020. 7871–7880
- 28 Raffel C, Shazeer N, Roberts A, Lee K, Narang S, Matena M, et al. Exploring the limits of transfer learning with a unified text-to-text transformer. *The Journal of Machine Learning Research*, 2020, **21**(1): Article No. 140
- 29 Radford A, Narasimhan K, Salimans T, Sutskever I. Improving language understanding by generative pre-training. *OpenAI*, 2018, **8**(9–10): 1–9
- 30 Achiam J, Adler S, Agarwal S, Ahmad L, Akkaya I, Florencia

- B, et al. GPT-4 technical report. arXiv preprint arXiv: 2303.08774, 2023.
- 31 Anil R, Dai A M, Firat O, Johnson M, Lepikhin D, Passos A, et al. PaLM 2 technical report. arXiv preprint arXiv: 2305.10403, 2023.
- 32 Touvron H, Lavril T, Izacard G, Martinet X, Lachaux M A, Lacroix T, et al. LLaMA: Open and efficient foundation language models. arXiv preprint arXiv: 2302.13971, 2023.
- 33 Touvron H, Martin L, Stone K, Albert P, Almahairi A, Babaei Y, et al. Llama 2: Open foundation and fine-tuned chat models. arXiv preprint arXiv: 2307.09288, 2023.
- 34 Anthropic. Introducing the next generation of Claude [Online], available: <https://www.anthropic.com/news/claude-3-family>, March 28, 2024
- 35 Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X H, Unterthiner T, et al. An image is worth 16x16 words: Transformers for image recognition at scale. In: Proceedings of the International Conference on Learning Representations. 2021.
- 36 Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M. Hierarchical text-conditional image generation with CLIP latents. arXiv preprint arXiv: 2204.06125, 2022.
- 37 Betker J, Goh G. Improving image generation with better captions. *Computer Science*, 2023, **2**(3): 8
- 38 Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE, 2022. 10674–10685
- 39 Singer U, Polyak A, Hayes T, Yin X, An J, Zhang S Y, et al. Make-A-Video: Text-to-video generation without text-video data. arXiv preprint arXiv: 2209.14792, 2022.
- 40 Bar-Tal O, Chefer H, Tov O, Herrmann C, Paiss R, Zada S, et al. Lumiere: A space-time diffusion model for video generation. arXiv preprint arXiv: 2401.12945, 2024.
- 41 Brooks T, Peebles B, Holmes C, DePue W, Guo Y F, Jing L, et al. Video generation models as world simulators [Online], available: <https://openai.com/research/video-generation-models-as-world-simulators>, March 28, 2024
- 42 Sun Y, Wang S H, Feng S K, Ding S Y, Pang C, Shang J Y, et al. ERNIE 3.0: Large-scale knowledge enhanced pre-training for language understanding and generation. arXiv preprint arXiv: 2107.02137, 2021.
- 43 Bai J Z, Bai S, Chu Y F, Cui Z Y, Dang K, Deng X D, et al. Qwen technical report. arXiv preprint arXiv: 2309.16609, 2023.
- 44 Zeng W, Ren X Z, Su T, Wang H, Liao Y, Wang Z W, et al. PanGu- α : Large-scale autoregressive pretrained Chinese language models with auto-parallel computation. arXiv preprint arXiv: 2104.12369, 2021.
- 45 Team I L M. InternLM: A multilingual language model with progressively enhanced capabilities [Online], available: <https://github.com/InternLM/InternLM>, March 28, 2024
- 46 Yang A Y, Xiao B, Wang B N, Zhang B R, Bian C, Yin C, et al. Baichuan 2: Open large-scale language models. arXiv preprint arXiv: 2309.10305, 2023.
- 47 Wang Fei-Yue. Digital teachers and parallel education: A paradigm shift in teaching and learning after ChatGPT. *Chinese Journal of Intelligent Science and Technology*, 2023, **5**(4): 454–463
(王飞跃. 数字教师与平行教育: 关于 ChatGPT 之后教学变革的探讨. *智能科学与技术学报*, 2023, **5**(4): 454–463)
- 48 Wang Fei-Yue. Parallel doctors and parallel hospitals: Impact and perspective of ChatGPT-like AIGC and AGI on medicine and medicare. *Medical Journal of Peking Union Medical College Hospital*, 2023, **14**(4): 673–679
(王飞跃. 平行医生与平行医院: ChatGPT 与通用人工智能技术对未来医疗的冲击与展望. *协和医学杂志*, 2023, **14**(4): 673–679)
- 49 Lu J W, Wang X X, Cheng X, Yang J, Kwan O, Wang X. Parallel factories for smart industrial operations: From big AI models to field foundational models and scenarios engineering. *IEEE/CAA Journal of Automatica Sinica*, 2022, **9**(12): 2079–2086
- 50 Yang J, Li S M, Wang X X, Lu J W, Wu H Y, Wang X. De-FACT in ManuVerse for parallel manufacturing: Foundation models and parallel workers in smart factories. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2188–2199
- 51 Yang J, Wang Y T, Wang X X, Wang X X, Wang X, Wang F Y. Generative AI empowering parallel manufacturing: Building a “6S” collaborative production ecology for manufacturing 5.0. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, DOI: [10.1109/TSMC.2024.3349555](https://doi.org/10.1109/TSMC.2024.3349555)
- 52 Zhao C, Wang X, Lv Y S, Tian Y L, Lin Y L, Wang F Y. Parallel transportation in transverse: From foundation models to DeCAST. *IEEE Transactions on Intelligent Transportation Systems*, 2023, **24**(12): 15310–15327
- 53 Zhao C, Dai X Y, Lv Y S, Tian Y L, Ren Y H, Wang F Y. Foundation models for transportation intelligence: ITS convergence in TransVerse. *IEEE Intelligent Systems*, 2022, **37**(6): 77–82
- 54 Wang X, Wang Y T, Yang J, Jia X F, Li L J, Ding W P, et al. The survey on multi-source data fusion in cyber-physical-social systems: Foundational infrastructure for industrial metaverses and industries 5.0. *Information Fusion*, 2024, **107**: Article No. 102321
- 55 Wang X X, Cheng X, Lu J W, Kwan O, Li S X, Ping Z X. Metaverses-based parallel oil fields in CPSS: A framework and methodology. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2138–2147
- 56 Wang Y T, Wang J G, Cao Y S, Li S X, Kwan O. Integrated inspection on PCB manufacturing in cyber-physical-social systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2098–2106
- 57 Wang Y T, Tian Y L, Wang J G, Cao Y S, Li S X, Tian B. Integrated inspection of QoM, QoP, and QoS for AOI industries in metaverses. *IEEE/CAA Journal of Automatica Sinica*, 2022, **9**(12): 2071–2078
- 58 Kang M Z, Wang X J, Wang H Y, Hua J, Reffye P D, Wang F Y. The development of AgriVerse: Past, present, and future. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(6): 3718–3727
- 59 Han X, Meng Z L, Xia X, Liao X S, He B Y, Zheng Z L, et al. Foundation intelligence for smart infrastructure services in transportation 5.0. *IEEE Transactions on Intelligent Vehicles*, 2024, **9**(1): 39–47
- 60 Wang F Y, Lv C. Foundation vehicles: From foundation intelligence to foundation transportation for future mobility. *IEEE Transactions on Intelligent Vehicles*, 2023, **8**(10): 4287–4291
- 61 Liu Y H, Shen Y, Tian Y L, Ai Y F, Tian B, Wu E, et al. Radarverses in metaverses: A CPSI-based architecture for 6S radar systems in CPSS. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2128–2137
- 62 Chen Y Y, Lv Y S, Wang F Y. The DAO to social transportation: Towards smart mobility in cyber-physical-social space. In: Proceedings of the 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 2023. 4289–4294
- 63 Wang J G, Tian Y L, Wang Y T, Yang J, Wang X X, Wang S J, et al. A framework and operational procedures for metaverses-based industrial foundation models. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(4): 2037–2046
- 64 Bi K F, Xie L X, Zhang H H, Chen X, Gu X T, Tian Q. Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, 2023, **619**(7970): 533–538
- 65 Dan Y H, Lei Z K, Gu Y Y, Li Y, Yin J H, Lin J J, et al. EduChat: A large-scale language model-based chatbot system for intelligent education. arXiv preprint arXiv: 2308.02773, 2023.
- 66 Chen Y R, Wang Z Y, Xing X F, Zheng H M, Xu Z P, Fang K, et al. BianQue-1.0: Improving the “question” ability of medical chat model through finetuning with hybrid instructions and multi-turn doctor QA datasets [Online], available: <https://github.com/seutcyr/BianQue>, March 28, 2024
- 67 Chen Y R, Xing X F, Lin J K, Zheng H M, Wang Z Y, Liu Q, et al. SoulChat: Improving LLMs’ empathy, listening, and comfort abilities through fine-tuning with multi-turn empathy conversations. In: Proceedings of the Findings of the Association for Computational Linguistics. Singapore: ACL, 2023. 1170–1183
- 68 Zhang H B, Chen J Y, Jiang F, Yu F, Chen Z H, Chen G M, et al. HuatuoGPT, towards taming language model to be a

- doctor. In: Proceedings of the Findings of the Association for Computational Linguistics. Singapore: ACL, 2023. 10859–10885
- 69 Zhang X Y, Yang Q, XuanYuan 2.0: A large Chinese financial chat model with hundreds of billions parameters. In: Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. Birmingham, UK: ACM, 2023. 4435–4439
- 70 OpenAI. Our approach to alignment research [Online], available: <https://openai.com/blog/our-approach-to-alignment-research>, March 28, 2024
- 71 Center for Research on Foundation Models. A holistic framework for evaluating foundation models [Online], available: <https://crfm.stanford.edu/helm/lite/latest/#/>, March 28, 2024
- 72 European Commission. Excellence and trust in artificial intelligence [Online], available: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/excellence-and-trust-artificial-intelligence_en, March 28, 2024
- 73 Liu A W, Pan L Y, Lu Y J, Li J J, Hu X M, Zhang X, et al. A survey of text watermarking in the era of large language models. arXiv preprint arXiv: 2312.07913, 2024.
- 74 Dai J, Pan X H, Sun R Y, Ji J M, Xu X B, Liu M, et al. Safe RLHF: Safe reinforcement learning from human feedback. arXiv preprint arXiv: 2310.12773, 2023.
- 75 Zhang W E, Sheng Q Z, Alhazmi A, Li C L. Adversarial attacks on deep-learning models in natural language processing: A survey. *ACM Transactions on Intelligent Systems and Technology*, 2020, 11(3): Article No. 24
- 76 Office of the Central Cyberspace Affairs Commission. Global AI governance initiative [Online], available: https://www.cac.gov.cn/2023-10/18/c_1699291032884978.htm, March 28, 2024 (中央网络安全和信息化委员会办公室. 全球人工智能治理倡议 [Online], available: https://www.cac.gov.cn/2023-10/18/c_1699291032884978.htm, 2024-03-28)
- 77 National Internet Information Office, National Development and Reform Commission, Ministry of Education of the People's Republic of China, Ministry of Science and Technology of the People's Republic of China, Ministry of Industry and Information Technology of the People's Republic of China, Ministry of Public Security of the People's Republic of China, et al. Interim measures for the management of generative artificial intelligence services [Online], available: https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm, March 28, 2024 (国家互联网信息办公室, 中华人民共和国国家发展和改革委员会, 中华人民共和国教育部, 中华人民共和国科学技术部, 中华人民共和国工业和信息化部, 中华人民共和国公安部, 等. 生成式人工智能服务管理暂行办法 [Online], available: https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm, 2024-03-28)
- 78 Yu Y, Yao S Y, Li J J, Wang F Y, Lin Y L. SWDPM: A social welfare-optimized data pricing mechanism. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC). Honolulu, USA: IEEE, 2023. 2900–2906
- 79 Wang D, Wang X H, Chen L, Yao S Y, Jing M, Li H H, et al. TransWorldNG: Traffic simulation via foundation model. In: Proceedings of the IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). Bilbao, Spain: IEEE, 2023. 6007–6012
- 80 Yu J R, Yu Y, Yao S Y, Wang D, Cai P L, Li H H, et al. RoW-based parallel control for mixed traffic scenario: A case study on lane-changing. In: Proceedings of the IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). Bilbao, Spain: IEEE, 2023. 5397–5402
- 81 Yao S Y, Yu J R, Yu Y, Xu J, Dai X Y, Li H H, et al. Towards integrated traffic control with operating decentralized autonomous organization. In: Proceedings of the IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). Bilbao, Spain: IEEE, 2023. 6126–6131
- 82 Qi H S, Yu Y, Tang Q, Hu X B. Intersection traffic deadlock formation and its probability: A petri net-based modeling approach. *IET Intelligent Transport Systems*, 2022, 16(10): 1342–1363
- 83 Lin Y L, Hu W, Chen X, Li S, Wang F Y. City 5.0: Towards spatial symbiotic intelligence via DAOs and parallel systems. *IEEE Transactions on Intelligent Vehicles*, 2023, 8(7): 3767–3770
- 84 Yu Y, Yao S Y, Wang K X, Chen Y, Wang F Y, Lin Y L. Pursuing equilibrium of medical resources via data empowerment in parallel healthcare system. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC). Honolulu, USA: IEEE, 2023. 3186–3191
- 85 Yu Y, Cui Y L, Zeng J Q, He C G, Wang D H. Identifying traffic clusters in urban networks based on graph theory using license plate recognition data. *Physica A: Statistical Mechanics and Its Applications*, 2022, 591: Article No. 126750
- 86 Lin Y L, Na X X, Wang D, Dai X Y, Wang F Y. Mobility 5.0: Smart logistics and transportation services in cyber-physical-social systems. *IEEE Transactions on Intelligent Vehicles*, 2023, 8(6): 3527–3532
- 87 Xu J, Yao S Y, Yu Y, Wang F Y, Lin Y L. DeMaaS: Efficient service distribution for MaaS via decentralized collaboration and optimization. In: Proceedings of the IEEE 26th International Conference on Intelligent Transportation Systems (ITSC). Bilbao, Spain: IEEE, 2023. 5391–5396
- 88 Li L, Wang X, Wang K F, Lin Y L, Xin J M, Chen L, et al. Parallel testing of vehicle intelligence via virtual-real interaction. *Science Robotics*, 2019, 4(28): Article No. eaaw4106
- 89 Zhao Y, Hu C, Zhu Z Q, Qiu S H, Chen B, Jiao P, et al. Crowd sensing intelligence for ITS: Participants, methods, and stages. *IEEE Transactions on Intelligent Vehicles*, 2023, 8(6): 3541–3546
- 90 Zhang J P, Pu J, Xue J R, Yang M, Xu X, Wang X, et al. HiVeGPT: Human-machine-augmented intelligent vehicles with generative pre-trained transformer. *IEEE Transactions on Intelligent Vehicles*, 2023, 8(3): 2027–2033
- 91 Liu Xin, Wang Xiao, Zhang Wei-Shan, Wang Jian-Ji, Wang Fei-Yue. Parallel data: From big data to data intelligence. *Pattern Recognition and Artificial Intelligence*, 2017, 30(8): 673–681 (刘昕, 王晓, 张卫山, 汪建基, 王飞跃. 平行数据: 从大数据到数据智能. 模式识别与人工智能, 2017, 30(8): 673–681)
- 92 Li X, Tian Y L, Ye P J, Duan H B, Wang F Y. A novel scenarios engineering methodology for foundation models in metaverse. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, 53(4): 2148–2159
- 93 Wang Fei-Yue, Wang Yan-Fen, Chen Yi-Zhu, Tian Yong-Lin, Qi Hong-Wei, Wang Xiao, et al. Federated ecology: From federated data to federated intelligence. *Chinese Journal of Intelligent Science and Technology*, 2020, 2(4): 305–311 (王飞跃, 王艳芬, 陈慧竹, 田永林, 齐红威, 王晓, 等. 联邦生态: 从联邦数据到联邦智能. 智能科学与技术学报, 2020, 2(4): 305–311)
- 94 Zhu Jing, Wang Fei-Yue, Wang Ge, Tian Yong-Lin, Yuan Yong, Wang Xiao, et al. Federated control: A distributed control approach towards information security and rights protection. *Acta Automatica Sinica*, 2021, 47(8): 1912–1920 (朱静, 王飞跃, 王戈, 田永林, 袁勇, 王晓, 等. 联邦控制: 面向信息安全和权益保护的分布式控制方法. 自动化学报, 2021, 47(8): 1912–1920)
- 95 Wang F Y, Qin R, Li J J, Wang X, Qi H W, Jia X F, et al. Federated management: Toward federated services and federated security in federated ecology. *IEEE Transactions on Computational Social Systems*, 2021, 8(6): 1283–1290
- 96 Miao Q H, Zheng W B, Lv Y S, Huang M, Ding W W, Wang F Y. DAO to HANOI via DeSci: AI paradigm shifts from AlphaGo to ChatGPT. *IEEE/CAA Journal of Automatica Sinica*, 2023, 10(4): 877–897
- 97 Wang F Y. The DAO to MetaControl for metasystems in metaverses: The system of parallel control systems for knowledge automation and control intelligence in CPSS. *IEEE/CAA Journal of Automatica Sinica*, 2022, 9(11): 1899–1908
- 98 Zhao Y, Zhu Z Q, Chen B, Qiu S H, Huang J C, Lu X, et al. Toward parallel intelligence: An interdisciplinary solution for complex systems. *The Innovation*, 2023, 4(6): Article No. 100521
- 99 Yang J, Wang X, Tian Y L, Wang X, Wang F Y. Parallel intelligence in CPSSs: Being, becoming, and believing. *IEEE Intelligent Systems*, 2023, 38(6): 75–80
- 100 Li X, Ye P J, Li J J, Liu Z M, Cao L B, Wang F Y. From features engineering to scenarios engineering for trustworthy AI: I&I, C&C, and V&V. *IEEE Intelligent Systems*, 2022, 37(4): 18–26
- 101 Wang Fei-Yue. Parallel management: The DAO to smart ecological technology for complexity management intelligence.

Acta Automatica Sinica, 2022, **48**(11): 2655–2669
(王飞跃. 平行管理: 复杂性管理智能的生态科技与智慧管理之DAO. 自动化学报, 2022, **48**(11): 2655–2669)

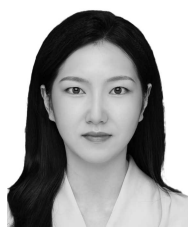
- 102 Li J J, Qin R, Wang F Y. The future of management: DAO to smart organizations and intelligent operations. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2023, **53**(6): 3389–3399
- 103 Yang Jing, Wang Xiao, Wang Yu-Tong, Liu Zhong-Min, Li Xiao-Shuang, Wang Fei-Yue. Parallel intelligence and CPSS in 30 years: An ACP approach. *Acta Automatica Sinica*, 2023, **49**(3): 614–634
(杨静, 王晓, 王雨桐, 刘忠民, 李小双, 王飞跃. 平行智能与CPSS: 三十年发展的回顾与展望. 自动化学报, 2023, **49**(3): 614–634)
- 104 Wang X X, Yang J, Wang Y T, Miao Q H, Wang F Y, Zhao A J, et al. Steps toward industry 5.0: Building “6S” parallel industries with cyber-physical-social intelligence. *IEEE/CAA Journal of Automatica Sinica*, 2023, **10**(8): 1692–1703
- 105 Wang Y T, Wang X, Wang X X, Yang J, Kwan O, Li L X, et al. The ChatGPT after: Building knowledge factories for knowledge workers with knowledge automation. *IEEE/CAA Journal of Automatica Sinica*, 2023, **10**(11): 2041–2044
- 106 Wang F Y, Yang J, Wang X X, Li J J, Han Q L. Chat with ChatGPT on industry 5.0: Learning and decision-making for intelligent industries. *IEEE/CAA Journal of Automatica Sinica*, 2023, **10**(4): 831–834



缪青海 中国科学院大学人工智能学院副教授。主要研究方向为智能系统, 智能交通, 平行智能。

E-mail: miaoqh@ucas.ac.cn

(**MIAO Qing-Hai** Associate professor at School of Artificial Intelligence, University of Chinese Academy of Sciences. His research interest covers intelligent systems, intelligent transportation systems, parallel intelligence.)



王兴霞 中国科学院自动化研究所复制系统管理与控制国家重点实验室博士研究生。2021年获得南开大学工学硕士学位。主要研究方向为平行智能, 平行油田, 多智能体系统。

E-mail: wangxingxia2022@ia.ac.cn

(**WANG Xing-Xia** Ph.D. candidate at the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. She received her master degree in engineering from Nankai University in 2021. Her research interest covers parallel intelligence, parallel oilfields, and multi-agent systems.)



杨静 中国科学院自动化研究所复杂系统管理与控制国家重点实验室博士研究生。2020年获得北京化工大学自动化专业学士学位。主要研究方向为平行制造, 社会制造, 人工智能和社会物理信息系统。

E-mail: yangjing2020@ia.ac.cn

(**YANG Jing** Ph.D. candidate at the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. She received her bachelor degree in automation from Beijing University of Chemical Technology in 2020. Her research interest covers parallel manufacturing, social manufacturing, artificial intelligence, and cyber-physical-social systems.)



赵勇 国防科技大学系统工程学院博士研究生。2021年获国防科技大学控制科学与工程专业硕士学位。主要研究方向为移动群智感知, 空间众包, 人机交互。

E-mail: zhaoyong15@nudt.edu.cn

(**ZHAO Yong** Ph.D. candidate at the College of Systems Engineering, National University of Defense Technology. He received his master degree in control science and engineering from National University of Defense Technology in 2021. His research interest covers mobile crowdsensing, spatial crowdsourcing, and human computer interaction.)



王雨桐 中国科学院自动化研究所助理研究员。2021年获得中国科学院大学控制理论与控制工程专业博士学位。主要研究方向为计算机视觉。

E-mail: yutong.wang@ia.ac.cn

(**WANG Yu-Tong** Assistant professor at the Institute of Automation, Chinese Academy of Sciences. She received her Ph.D. degree in control theory and control engineering from the University of Chinese Academy of Sciences in 2021. Her main research interest is computer vision.)



陈圆圆 中国科学院自动化研究所助理研究员。2018年获得中国科学院大学控制理论与控制工程专业博士学位。主要研究方向为交通数据分析, 社会交通, 平行交通管理与控制系统。

E-mail: yuan yuan.chen@ia.ac.cn

(**CHEN Yuan-Yuan** Associate professor at the Institute of Automation, Chinese Academy of Sciences. He received his Ph.D. degree in control theory and control engineering from University of Chinese Academy of Sciences in 2018. His research interest covers traffic data analytics, social transportation, and parallel traffic management and control systems.)



田永林 中国科学院自动化研究所博士后. 2022 年获得中国科学技术大学控制理论与控制工程专业博士学位. 主要研究方向为平行智能, 自动驾驶, 智能交通.

E-mail: yonglin.tian@ia.ac.cn

(**TIAN Yong-Lin** Postdoctor at the

Institute of Automation, Chinese Academy of Sciences. He received his Ph.D. degree in control theory and control engineering from the University of Science and Technology of China, in 2022. His research interest covers parallel intelligence, autonomous driving, and intelligent transportation systems.)



俞怡 上海人工智能实验室助理研究员. 主要研究方向为智能交通系统, 数据要素化, 城市计算.

E-mail: yuyi@pjlab.org.cn

(**YU Yi** Assistant professor at Shanghai Artificial Intelligence Laboratory. Her research interest

covers intelligent transportation systems, data trading, and urban computing.)



林懿伦 上海人工智能实验室副研究员. 2019 年获得中国科学院大学控制理论与控制工程专业博士学位. 主要研究方向为社会计算, 平行智能, 深度学习, 智能交通系统与人工智能安全.

E-mail: linyilun@pjlab.org.cn

(**LIN Yi-Lun** Associate professor at Shanghai Artificial Intelligence Laboratory. He received his Ph.D. degree in control theory and control engineering from the University of Chinese Academy of Sciences, in 2019. His research interest covers social computing, parallel intelligence, deep learning, intelligent transportation systems and AI safety.)

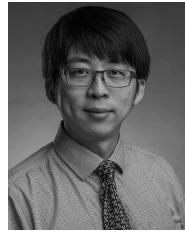


鄢然 新加坡南洋理工大学土木与环境工程学院助理教授. 主要研究方向为海事研究中的数据分析和海运大数据, 绿色航运管理, 海事风险管理以及港口和航运优化.

E-mail: ran.yan@ntu.edu.sg

(**YAN Ran** Assistant professor at the

School of Civil and Environmental Engineering, Nanyang Technological University, Singapore. Her research interest covers data analytics in maritime studies, big data in maritime transport, green-shipping management, maritime risk management, and port and shipping optimization.)



马嘉琪 加州大学洛杉矶分校萨穆埃利工程学院副教授, 加州大学洛杉矶分校交通研究所副所长. 2014 年获得弗吉尼亚大学交通运输工程博士学位. 主要研究方向为联网和自动化车辆, 网络物理运输系统, 运输系统的弹性, 分布式多智能体系统的协同控制,

智能交通系统, 动态运输系统建模和控制, 网络优化, 出行行为建模和需求预测, 人工智能和先进计算在交通领域的应用. E-mail: jiaqima@ucla.edu

(**MA Jia-Qi** Associate professor at the UCLA Samueli School of Engineering and associate director of UCLA Institute of Transportation Studies. He received his Ph.D. degree of transportation engineering from University of Virginia, 2014. His research interest covers connected and automated vehicles; cyber-physical transportation systems; transportation systems resilience; cooperative control of distributed multi-agent systems; intelligent transportation systems; dynamic transportation systems modeling and control; network optimization; travel behavior modeling and demand forecasting; artificial intelligence and advanced computing applications in transportation.)



那晓翔 英国剑桥大学工程系长聘助理教授. 2014 年获英国剑桥大学机械工程博士学位. 主要研究方向为重型商用车智能车载信息系统开发与车辆能耗特性评价.

E-mail: xnln2@cam.ac.uk

(**NA Xiao-Xiang** University assistant professor in the Department of Engineering, University of Cambridge, U.K.. He received his Ph.D. degree in mechanical engineering from University of Cambridge, U.K.. His research interest covers development of intelligent telematics systems for heavy goods vehicles and assessment of vehicle energy performance.)



王飞跃 中国科学院自动化研究所复杂系统管理与控制国家重点实验室研究员. 主要研究方向为智能系统和复杂系统的建模、分析与控制. 本文通信作者.

E-mail: feiyue.wang@ia.ac.cn

(**WANG Fei-Yue** Professor at the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interest covers modeling, analysis, and control of intelligent systems and complex systems. Corresponding author of this paper.)