

具身智能：步入AI舞台中央？

AI风向标②

◎本报记者 崔爽

近日，一段视频在网上热传。视频中，初创公司 Figure 的人形机器人外表和动作都与人类十分相似。它可以与人流畅对话，理解人的意图。听到工程师“想吃点东西”的口令后，它迟疑了一下，准确地拿起了苹果。它甚至可以解释自己为什么要这样做。

这一人形机器人背后，是 OpenAI 公司为其配置的大语言模型。多模态技术加持下，它具备高级视觉和语言智能。

随着大模型、机器人制造等技术发展，赋予人工智能“形体”的具身智能(Embodied AI)正逐步走入人工智能舞台中央。

具身智能是人工智能产业发展的重要方向之一。英伟达近期发布人形机器人通用基础模型 Project GR00T，希望能让机器人拥有更聪明的“大脑”。由 Project GR00T 驱动的机器人能够理解自然语言，并通过观察人类行为来模仿人类动作。在特斯拉 2023 股东大会上，马斯克也向大众展示了人形机器人 Optimus 的全新型号。这些都是具身智能驶入发展快车道的缩影。

AI技术的集大成者

1950年，图灵在他的一篇论文中首次提出了具身智能的概念。

“具身智能是人工智能的一个重要分支，它强调智能体与其所处的物理环境紧密交互，通过人形机器人等智能体实现感知、决策和行动。”赛迪顾问人工智能与大数据研究中心常务副总经理邹德宝向记者介绍，“这种结合让人工智能可通过摄像头、传感器等理解环境，并通过机械臂、轮子等执行器作用于物理世界，在物理空间中学习、适应并完成指令任务。”

这是一个包含人工智能几乎所有技术的领域，可以说是 AI 的集大成者。

“要充分发挥人工智能的强大能力，就要将其‘附身’于实体——机器人，这一行为被称为‘具身’。”达闼机器人创始人、首席执行官黄晓庆说。

在他看来，从生物学角度，人类是自然选择的产物，因此具备通用性的终极机器形态应与人类相同；从功能性角度，要使机器人能替代人类执行所有任务，其形态需与人类相似；从人工智能发展角度，训练大模型需要大量数据，而人类语言、行为、智能等数据信息无疑最丰富，人形机器人因此能最大限度获取训练数据。

“通过为人工智能系统赋予形体，我们可以使其与人类用户之间的交互更直观、自然和便捷。形体也可以帮助人工智能系统更好地感知和理解现实世界。”邹德宝说，具有形体的人工智能系统可以应用于更多领域，特别是在需要与人类进行密切协作的场景中。例如，在医疗领域，一个具有实体形态的人工智能机器人可以协助医生进行手术操作。

技术进步和挑战并存

依托大规模数据和强大算力的大模型，为机器理解世界提供了前所未有的能力。它与具身智能的结合，预示着机器人的未来将更加智能化、互动化。

去年 11 月，工业和信息化部印发《人形机器人创新发展指导意见》，提出人形机器人集成人工智能、高端制造、新材料等先进技术，发展潜力大、应用前景广，是未来产业的新赛道。

邹德宝谈到，发展具身智能，主要可从三方面发力。一是仿生机制。具身智能的一个关键目标是使机器人像生物体一样具备感知、决策和运动能力。因此，更深入地研究和模仿复杂的生物机制，如神经系统的运行方式、生物体的自我修复能力等，是实现这一目标的重要途径。二是基于物理的智能决策能力。智能系统需要更紧密地与



北京市中关村展示中心常设展上的人形机器人。视觉中国供图

其所处的物理形态融合，以便在各种环境中自然、高效地执行任务。这意味着需要发展出更先进的机器人硬件和软件技术，支持智能系统在复杂环境中的运动和操控。三是自主学习能力和适应环境能力。具身智能系统需要具备自主学习和适应环境能力，以便在复杂多变的环境中执行任务和解决问题。这可以通过强化学习、迁移学习等机器学习技术实现。

技术突飞猛进之下，具身智能市场活跃。头部厂商不仅在技术研发上取得显著进展，还在商业应用和拓展市场方面取得了突破。特斯拉宣布，其旗下擎天柱人形机器人的首个商业化场景将是自家汽车生产车间。人形机器人(上海)有限公司联合行业头部企业成立的新型研发机构，致力于建设贯穿创新链、产业链、资金链和人才链的人形机器人创新生态系统。联想创投作为联想集团旗下的全球科技产业基金，通过投资和支持具身智能技术公司，推动该技术的商业化应用。

然而，难题和挑战也如影随形。首当其冲的是算力。具身智能需要进行大量计算来支持感知、决策和运动过程。随着任务和环境复杂性的升级，其对算力的需求也不断增加。

数据安全也不可忽视。具身智能系统需要收集和大量数据进行处理，这些数据可能包含用户个人隐私信息。

“随着具身智能技术不断发展，我们也需要关注其可能带来的伦理和道德问题。例如，如何确保机器人在执行任务时不伤害人类或造成不必要的损失，如何保证机器人行为符合人类道德和价值观等。”邹德宝说。

未来人工智能发展的重要方向

英伟达创始人兼 CEO 黄仁勋曾表示，人工智能发展的下一个浪潮是具身智能，是能理解、推理、并与物理世界互动的智能系统。

“具身智能对人工智能产业发展意义重大，对实现 AGI(通用人工智能)具有不可忽视的价值。”邹德宝认为，从当前技术趋势和市场动态来看，具身智能极有可能成为人工智能发展的下一个重要方向。

首先，具身智能强调智能体与真实世界的交互。这种交互性不仅增强了人工智能系统的实用性，还促进了人工智能技术在各个领域的应用。具身智能是实现 AGI 的关键之一。AGI 需要能像人类一样理解并适应各种物理环境和各种复杂任务，而具身智能正是通过让智能体与物理环境进行实时交互和响应，实现对环境的感知和理解。因此，具身智能的发展将为实现 AGI 提供重要技

术支持。

从当前技术发展趋势来看，具身智能已成为人工智能领域研究热点。随着深度学习、强化学习和迁移学习等技术的不断进步以及机器人硬件的日益成熟，具身智能已具备一定技术基础。同时，应用市场对具有可交互性的智能系统的需求也在不断增长，这为具身智能发展提供了广阔市场。

针对发展具身智能可能面临的风险，邹德宝建议，要通过优化算法、提高硬件性能等方式不断提升计算能力，满足具身智能对计算资源的需求。采用加密技术、数据脱敏等手段保护用户数据安全和隐私，同时制定严格的数据使用和共享政策，确保用户数据不会被滥用或泄露。此外，也要制定具身智能技术伦理和道德准则，并加强对具身智能技术的监管和评估，确保其行为符合人类道德和价值观。

相关新闻

广东省具身智能机器人创新中心启动

科技日报(记者罗云鹏)记者 5 月 25 日获悉，广东省具身智能机器人创新中心近日在深圳启动。

据悉，创新中心将汇聚广东省在人工智能与机器人领域的学术、研究及产业资源，构建产业服务平台，推动形成自主可控的具身智能机器人产业集群，促进人工智能与机器人在医疗健康、人才教育、城市管理、特种工业等领域持续创新。

“创新中心将通过科学的行业引导，整合算力、人才、数据等核心资源，形成产业链上下游协同发展合力。”创新中心负责人丁宁介绍，创新中心将凭借体制机制优势，高效组织各单位协同分工，抓住时间窗口集中攻关，尽快实现关键能力突破。

另悉，今年 3 月，深圳“20+8”产业集群 2.0 版本上线，将人工智能产业、机器人产业、智能机器人产业列入“20+8”战略性新兴产业集群和未来产业，并制定印发相关行动计划。

深圳市工业和信息化局副局长林毅表示，深圳将继续营造良好创新环境，做好各方面服务，积极支持企业发挥主体作用。深圳欢迎社会力量广泛参与，共同开拓创新，推动人工智能产业高质量发展。

腾讯混元文生图大模型全面开源

科技日报(记者崔爽)记者 5 月 25 日获悉，腾讯近日宣布旗下混元文生图大模型全面升级，并对外开源。

这是业内首个中文原生的 DiT 架构文生图开源模型，具备原生中文理解能力和双语编码能力，参数量 15 亿。此次开源的模型包含模型权重、推理代码、模型算法等完整模型，供企业与个人开发者免费商用，目前已在专注于自然语言处理的开源机器学习平台 Hugging Face 及面向软件开发者的代码托管平台 Github 上发布。

腾讯混元文生图技术负责人芦清林介绍，升级后的混元文生图大模型采用了与 Sora 一致的 DiT 架构，不仅可支持文生图，也可作为文生视频等多模态视觉生成的基础。

过去，视觉生成扩散模型主要基于 U-Net 架构。随着模型参数量增加，DiT 架构展现出更好的扩展性，有助于提升模型生成质量和效率。

记者了解到，混元文生图大模型是业界最早探索并将大语言模型结合 DiT 架构的文生图模型之一。2023 年 7 月，腾讯混元文生图团队明确了基于 DiT 架构的模型方向，并启动了新一代模型研发。今年初，混元文生图大模型全面升级为 DiT 架构。

评测数据显示，最新腾讯混元文生图模型效果相比前代提升超 20%，远超开源的 Stable Diffusion 模型。

目前，主流文生图开源生态基本围绕英文建设，Stable Diffusion 等主流开源模型虽一定程度支持中文输入，但其核心数据集仍以英文为主。建设中文原生大模型以及相应的开源生态对国内大模型产业有重要意义。

混元文生图大模型以中文原生为基础，支持中英文双语输入及理解。其文生图能力已被广泛用于素材创作、商品合成、游戏出图等业务场景。

“腾讯混元文生图大模型的研发思路就是实用，坚持从实践中来，到实践中去。此次把最新一代模型完整开源，是希望与行业共享实践经验研究成果，加速大模型行业发展。”芦清林表示，目前开源的混元文生图大模型版本，也是腾讯内部正在使用的版本。基于该模型，开发者和企业无需从头训练，可直接将其用于推理，并可打造专属的 AI 绘画应用及服务。

此外，目前文生图开源社区主要还是以 Stable Diffusion 等为主的英文开源社区。混元文生图大模型的开源，将丰富以中文为主的文生图开源生态，推动中文文生图技术研发和应用。

山西大学团队用图调控网络实现高性能点云分割

科技日报(记者韩荣)5 月 25 日，记者从山西大学获悉，该校智能信息处理研究所团队用图调控网络实现高性能点云分割，相关成果发表在人工智能领域国际期刊《IEEE 模式分析与机器智能汇刊》。

点云数据分析广泛应用于自动驾驶、三维理解和机器人等领域。点云分割是点云数据分析领域中基本且具有挑战性的任务，旨在根据不同属性与功能将目标点云划分成不同区域。

“实现高性能点云分割，关键在于提取具有判别性的逐点特征。”山西大学智能信息处理研究所团队成员介绍，在点云不同分割区域边界处，邻域节点往往存在较强异质性。此前的方法在实现特征聚合过程中忽略了节点之间同质与异质关系。这使得节点特征混入不必要的异质节点信息，导致点云分割边界模糊。

针对上述问题，研究团队提出了一种图调控网络，将点云建模为同质-异质图。他们结合图注意力模型，设计了基于同质性引导的图注意力卷积，挖掘局部邻域内同质性特征。研究团队还进一步设计了原型特征提取模块，从全局范围原型空间中进一步挖掘同质性特征，提高节点特征判别性，进一步提升点云分割性能。

据悉，研究成果进一步增强了节点特征的判别性，提高了点云分割边界清晰度。

天津口岸：智能审图为通关提速增效

◎本报记者 陈曦

近日，一票共有 8 个集装箱的出口轮船抵达天津港。这批货物依次通过码头内机检设备，集装箱的扫描图像被实时传送到 10 公里外的审像中心。短短 3 秒后，智能审图系统自动对这些图像进行判别，并得出“未见异常”的审像结论。经审像关员进一步确认，审像结果被迅速同步至码头，货物放行并顺利装船。

天津海关集中审像中心审像一科科长刘常豪介绍，智能审图系统是人工审图经验和人工智能技术，对机检扫描图像进行实时自动识别的智能化系统。目前，天津口岸一半以上进出口商品已通过智能审图实现自动准确识别，平均审图时间约每幅 3 秒。

近年来，天津海关积极落实智慧海关建设和“智关强国”行动各项要求，不断提升智慧机检工作水平。在优化机检审像业务流程的基础上，天津海关建立疑难图像集中研判机制，借助智能审图系统对机检货物实施自动比对筛查，充分发挥机检审像快速精准的优势，提升查验效能。同时，天津海关持续对智能审图系统进行训练，优化算法、创新流程，不断提高智能审图系统识别速度和精准度，实现无问题货物快速验放，为进出口企业提供“零接触”的无感通关体验。



审像关员确认智能审图系统审像结果。天津海关供图

在 2024 西湖论剑·数字安全大会上，专家呼吁——

创新安全流通机制 实现数字经济飞跃

◎洪恒飞 本报记者 江耘

“数字经济的发展必然伴随着数据安全和网络安全两大难题的出现。”在 5 月中旬举办的 2024 西湖论剑·数字安全大会上，全国数字经济联盟学术委员汪玉凯认为，攻克上述两大难题，重点在于完善相关法律、增强防护能力、提高监管水平。同时，还要特别关注数据和信息保护，这样才能推动数字经济实现新飞跃。

大会以“智绘安全 乘数而上”为主题，紧扣“智”“数”两大关键词，聚焦全球数字安全前沿技术与应用，设置“人工智能+安全”“数据要素×安全”等议题，邀请专家分享数字安全建设成果。

形成数字生态系统内在防御能力

2022 年，美国网络安全风险投资公司

Cybersecurity Ventures 在其发布的网络犯罪报告中预测，全球网络犯罪造成的损失将以每年 15% 的速度增长，到 2025 年或将达到 10.5 万亿美元。

“这比全球一年自然灾害造成的损失大得多。”中国工程院院士鄂江兴在主题演讲中指出，数字经济时代，社会数字化转型面临系统性风险。在他看来，当今网络安全防御就像“打鼹鼠”游戏，冒出一个“鼹鼠”就“打”一下，用户侧与制造侧的网络安全责任及风险存在失衡现象。相比制造侧，用户侧处于明显弱势地位。

鄂江兴说，网络安全市场具有强烈的负外部性，由用户承担安全成本，第三方享受不利后果。《欧盟网络弹性法案》中的数据表明，欧盟近 50% 的制造商明知其产品有漏洞仍将其投入市场。

在他看来，需要将安全融入数字产品设计与制造，将安全作为数字产品的核心能力贯穿产品设计始终，让产品开箱即可抵御攻击，而非仅将其作为产品的一个技

术选项。

鄂江兴认为，传统网络安全治理局限性日益凸显，数字生态系统底层驱动范式急需转型。要建立一种富有弹性和韧性的数字生态系统，使其更容易形成内在防御能力，筑牢网络与数据空间安全屏障。

以资源入表为起点促数据进一步流通

中国科学院信息工程研究所副总工程师、研究员李凤华说，体系化架构与技术是促进数据要素价值释放的核心。要通过多方共同努力，解决数据多轮交易和权属转移等挑战，支撑我国数字经济健康发展、行稳致远。

中国互联网协会副理事长、国家信息化专家咨询委员会委员高新民呼吁，要在遵循数据资源特征和发展规律的基础上对其进行开发利用和治理，尊重数据主权，支持数据安全自由流动，推动形成数

据多元流动机制，这样才能让数据发挥出最大价值。

在北京国家会计学院党委委员、副院长李旭红看来，要推动数据进一步流通，数据资源入表是一个起点。

数据资源入表，是指将数据确认为企业资产负债表中“资产”一项，在财务报表中体现其真实价值与业务贡献。

大会期间，由杭州安恒信息技术股份有限公司、杭州数据交易所、中国软件评测中心等单位编制的《企业数据资源入表实践白皮书》正式发布。

在杭州安恒信息技术股份有限公司首席技术官刘博看来，这份白皮书旨在为企业数据资源入表提供实践指南。数据资源入表要为数据流通服务。数据入表若遭到泄露或滥用，也是资产流失。因此，安全是其重要前提。

全球化可以帮助企业在全中国范围内统一数据管理，提高决策速度和精度，但跨国数据传输的合规性、安全性，是企业必须要过的几道关。