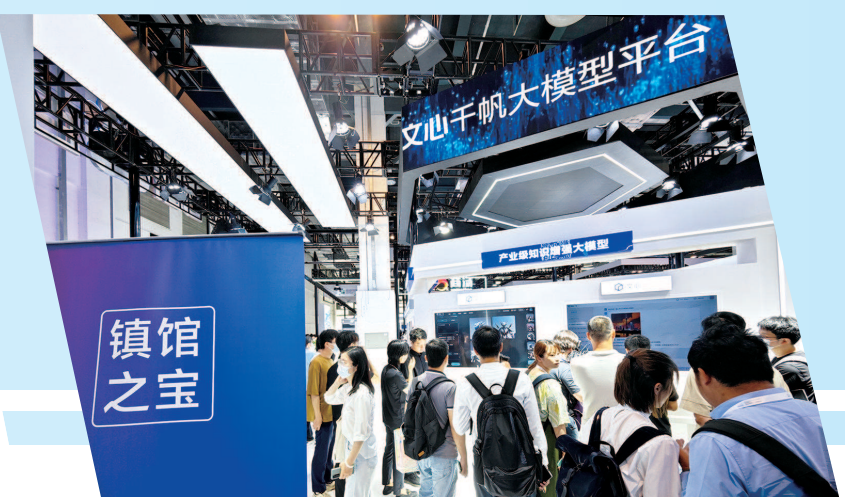


冰与火之争：大模型的“下一站”是通用智能吗？



■“镇馆之宝”之一，百度文心一格大模型 本报记者 刘敬 摄

冷静 大模型并非万能

图灵奖得主、Meta AI基础人工智能团队首席AI科学家杨立昆看来，目前似乎各种好用、好评的模型其实并非万能。“实际上现在广泛应用的许多系统或许展现了非凡的应用实力，但‘有用’并非完美，实际上各有缺陷，都不是通向人类智能的模型。”他解释，一些语言模型或图像模型实际上仍然是基于“上一句”“上一帧”来进行预测的，本质还是一种模仿。

这种“模仿”实际上暴露了大模型学习精度不够的缺陷，这也是一些大模型产生“幻觉”的原因。这意味着需要对大模型进行进一步优化，而算力、效率、成本等问题同时伴随。

即使在百花齐放的企业应用层面，也并非万能钥匙。“通用大模型可以在100个场景中，解决70%~80%的问题，但未必能100%满足企业某个场景的需求。”腾讯集团高级执行副总裁、云与智慧产业事业群CEO汤道生表示，企业的大模型应用需要综合考虑行业专业性、数据安全、持续迭代和综合成本等因素。他解释说，通用大模型一般是基于广泛的公开文献与网络信息来训练的，许多专业知识与行业数据积累不足，导致回答的行业针对性与精准度不够。但用户对企业提供的专业服务，要求高、容错性低，企业一旦向公众提供了错误信息，可能引起严重后果。

引导AI向善 性别话题是重要一环

ChatGPT等人工智能工具的诞生，引爆了人工智能行业迄今为止最火的一波狂潮。AI技术的进步不仅提高了生产力，还有助于解决不平等社会问题，但与此同时，也引发了人们的担忧：未来，AI是否会取代人类的职业岗位，而女性是否受到的影响会更大？

2023世界人工智能大会AI女性菁英论坛上，嘉宾们普遍认为，新一代技术革命无疑会带来一些我们熟悉的岗位，但也会创造一批新的机遇。其中，女性的工作不可或缺，引导人工智能向善，性别平等包容就是重要一环。

未来每个人或都有“数字分身”

“人工智能替代”部分人类的行为”是必然的趋势，但它并不是针对某些职业或者某些人，而是替代一些具体的劳动行为。”中智经经济技术合作股份有限公司党委书记、副总经理李双认为，以ChatGPT为代表的生成式人工智能确

实会对知识密集、数据密集、符合常规逻辑的岗位带来较大的冲击，“如果一项工作，决策时间在5秒以内，说明这项决策所需的信息量小，性质重复、简单，可独立完成，那么，它大概率会被AI替代”。

在知识密集、数据密集、符合常规的逻辑岗位上，女性占比高，那是否意味着女性的就业未来将受到AI更大的冲击？对此，李双并不认同。“AI不可避免会对就业带来影响，除了女性居多的领域和行业，未来AI和自动化结合，进入产业落地，那么，许多技能岗位、蓝领岗位一样会被AI取代。所以，AI影响就业，更多的是个体竞争力的差异，而非性别。未来，可以取代你的可能不是AI，而是更会使用AI的人。”

云从科技集团股份有限公司董事、副总经理兼董事会秘书杨桦说，在她所在的企业，确实已经出现了AI替代工作的现象，研发团队的人数在缩减，但是，人数降低的主要是初级研发人员。“研发人员自己研发的AI自动化平台正在‘代劳’

人类发展人工智能的初衷，是希望能制造出智慧、能干的机器人，在智能化的环境中服务于人。过去，尽管人工智能近年飞速发展、蓬勃应用，但人们始终觉得到达理想的智能世界还相去甚远。而大模型的横空出世，如同魔法点亮新世界。这是否意味着，大模型的“下一站”即将抵达理想的通用智能？在2023世界人工智能大会上，大多数论坛都离不开大模型的讨论，这一“魔法”是万能之解吗？“魔法”会失控吗？在一片火热之中，亦有冷静之声。

的规整性也非常关键——因为人工智能，包括大语言模型，都需要大量规整的数据进行训练。

而对此，腾讯等“大厂”正帮助合作伙伴基于行业大模型，构建自己的专属模型。汤道生表示，企业如果基于行业大模型，再加上自身数据进行精调，可以建构专属模型，打造高可用的智能服务；而且模型参数比通用大模型少，训练和推理的成本更低，模型优化也更容易。同时，行业大模型和模型开发工具，也可以通过私有化部署、权限管控和数据加密等方式，防止模型训练和使用带来企业敏感数据的外泄。另外，大模型落地过程，还需要经过算法构建、模型部署一系列环节，每个环节都不能“掉链子”。模型后续还需要不断地迭代调优，这需要用系统化、工程化的工具。

理性 不断优化仍在路上

针对安全问题，企业已有“魔法打败魔法”的对策。本届大会的“镇馆之宝”之一——蚂蚁集团联合清华大学发布AI安全检测平台“蚁鉴2.0”可以实现用生成式AI能力检测生成式AI模型，可识别数据安全、内容安全、科技伦理三大类的数百种风险，覆盖表格、文本、图像等多种数据和任务类型，是业内首个实现产业级应用的全数据类型AI安全检测平台。

大模型的研发本身也需要创新。“不能一直只跟着国外的成果，而是要基于大模型在应用上有创新。”商汤科技联合创始人、首席科学家、绝影智能汽车事业群总裁王刚透露，商汤科技从2019年开始就布局“大模型”方面的研究，并随之建立和不断完善大装置软硬件的基础设施，“另外，需要针对大模型给行业带来的变化，重塑研发体系。”

公司的基础工作，初级研发人员如果不做自我提升，最终将面临离开的局面。”杨桦还指出，未来，每个人可能都会有自己的“数字分身”来替代一些简单的工作。“我自己现在就有一个‘数字分身’，可以替代工作上的部分内容，包括常规的公司介绍和交谈等；而且，我本身在编程技术上弱一些，我的分身还能补足我的不足。”

软技术能让AI变得更“完整”

各界专家普遍认为，人工智能发展过程中，女性的参与不能“缺席”，女性的软技术能让AI变得更“完整”。中国工程院院士刘珪认为：“我们要引导人工智能向善，其中性别平等包容就是重要一环。”联合国驻华协调员常启德也表示：“我们需要更多女性参与开发人工智能，AI是由算法驱动的，在算法产出过程中增加女性的参与，才能有效避免算法的性别偏见。”

“人工智能无论是作为机器也好，还是作为

睿智 开源创新共创未来

AI大模型等技术已经深刻地影响到所有行业的发展，并正在重构企业核心产品，重塑用户与企业产品和服务的交互方式。“未来，金融、政府、能源、交通等每一个行业与领域，都会诞生领域或者行业大模型，这些大模型具有专家的能力，可以在上面构造复杂的应用；而且，随着AI工具的逐渐成熟，训练大模型的难度和成本越来越低，将真正推动AI普惠。”星环科技创始人、董事长孙元浩预计，今年就将会行业大模型在金融尤其是证券基金行业规模化落地，而相关业务带来的收入即将在下半年有所体现。

“不能再将人工智能视为简单的技术工作，而是整个组织的变革引擎。”中国科学院院士何积丰也在会上提到，这一引擎的安全可信需要多维度保障，包括数据安全可信、系统行为可追责、算法模型可解释、网络环境可信、法律伦理可信等。

杨立昆认为，应该建立不同的非生成式架构，构建表征空间的训练来实现预测。也就是说，为机器建立世界模型，帮助他获得类似人类的物理直觉，来理解世界、形成预测。“这将是未来十年的巨大挑战。”

至于伦理问题，他认为，应建立世界模型——当机器真正像人类理解世界一样学习，并且设定向善的目标，机器就会在学习自行限行为，这或许将为构建伦理形成天然基础。“这是艰巨挑战，但我们可以从培养一只‘好的小老鼠’开始，再慢慢到‘好的狗’‘好的猴子’，逐步放进仿真环境测试。”他同时强调，基于AI向善的目标，相比严格监管，开源是更好的方式。

工具也好，本身没有感性的一面，它缺的是‘爱’。充满人性的爱也是人类和机器最大的不同。女性有不同于男性的视角和经验，她付出的爱能让AI更加完整。因此，AI一定会是性别平等包容的数字红利，而不是制造性别的鸿沟。”李双说。

AI的发展，催生了一些更青睐女性的就业机会。杨桦认为，与其讨论机器会不会代替人，更重要的是探寻人与机器和谐共生之道。“人可以为机器提供情绪价值，从而诞生更多辅助性的、让社会更加温暖的东西，这些岗位和人群一定具有更大价值，而通常来说，女性情感更加细腻、沟通技巧更强、在处理复杂问题时能承受更大压力和考验。”比如，在人机对话方面，女性的软技术优势更强，而在翻译、财务等容易催生大语言模型的领域，耐心细致、更容易具备提炼、归纳知识能力的女性或许就能投身进来，从事挖掘、清理、采集数据等工作。 本报记者 马丹



■商汤“秒画”可以结合输入的图例或文案智能创作出与输入相关的图像内容 本报记者 刘敬 摄

掌握大模型世界的“魔咒”后，两秒钟就能将想象变成“现实”

在“描述词”框里输入“上海，未来场景，人工智能，智慧城市”，选择生成数量“4”，点开“描述词优化”，等待一会儿，几幅酷炫的图案就呈现在眼前。

修改，输入“上海，智慧城市，未来感，充满活力的画面感，鲜艳的色彩运用，清晰的笔触”，呈现的图片更令人欢喜。

在2023世界人工智能大会召开前夕，新民晚报联合商汤科技，发起大模型世界“魔咒”招募令，邀请市民朋友利用“商汤秒画”创作一幅主题为“生成未来 申城未来”的图画。在官微后台，我们收到了不少有意思的作品。

两秒钟，就能将你的想象变成“现实”。“商汤秒画”基于商汤自研生成式人工智能模型的作画能力，并基于商汤大装置的GPU算力支撑，可帮助用户免除本地化部署流程，既能使用商汤自研

作画模型高效地生成高质量内容，也可以将本地模型及其他第三方开源模型一键导入，生成更多样的内容。

“秒画”支持二次元、三次元等多种生成风格。用户无需复杂的编程知识，只需给定提示词，即可生成相应图像，最高可实现6K高清图。“目前负责商汤超大模型研发体系和决策智能算法与分布式平台的刘宇博士介绍，“无论是写实照片、脑洞大开的艺术画作，还是未来科幻场景，模型生成的图像都能根据用户需求，带来真实的光影效果以及丰富的细节内容。”

值得一提的是，“秒画”还支持用户上传本地图像，结合商汤自研作画模型或者开源模型来训练定制化LoRA（用于通过小数据继续训练，将大模型的能力适应到某个垂类领域）模型，来高效

地生成个性化内容。

在大模型的世界里，提示词是一种用于与人工智能模型交互的语言。它可以用来告诉模型需要生成什么样的内容，大多数模型师也会称呼提示词为“魔咒”，不同的“魔咒”内容可以召唤出指定的“神奇魔法”。据介绍，一条出色的“魔咒”通常会由三个部分组成：绘画对象、对象描述词和风格修饰词。当然了，生成图像的质量与美感与“魔咒”的精确度和质量十分相关。

在7日的“大爱无疆·日日新月异”人工智能论坛上，商汤科技董事长兼CEO徐立宣布，自研生成式大模型商汤秒画3.0的模型参数从今年4月首次发布以来的10亿提升至70亿量级，能够实现专业摄影级的图片细节刻画。

本报记者 邵阳

AI 掠影

新品 可单手举起一台钢琴的人形机器人

首次应邀参展的特斯拉将Tesla Bot人形机器人带到了现场。特斯拉人形机器人身高172厘米，体重56.6千克，就是一个正常成年人的身形。机器人全身拥有28个自由度，就如同人体的关节。手部有11个自由度，可以更加灵活地抓握，且力大无比，能单

手举起一台钢琴。目前，这款人形机器人可以完成行走、上下楼梯、下蹲、拿取物品等动作，也已经具备了保护自身和周围人安全的能力，未来还可以做饭、修剪草坪、帮助照顾老人，或是在工厂里面替代人类从事枯燥和有危险的工作。



■特斯拉Tesla Bot人形机器人 本报记者 陶磊 摄

突破 以“云+AI”探测2颗快速射电暴

2023世界人工智能大会·腾讯论坛透露，腾讯、国家天文台、复旦大学计算机科学技术学院两年前联合启动“探星计划”，用“云+AI”帮助中国天眼FAST处理每天接收到的庞大数据量，通过视觉AI分析找到快速射电暴、脉冲星线索，截至目前已发现30颗脉冲星。

今年，探星计划再度取得进展，首次通过人工智能技术发现2颗快速射电暴。

腾讯针对快速射电暴探索，特别设计了一套全新的，端到端的AI算法。在同算力下，这套全新的天文数据处理范式，推动信号处理效率相对常规处理流程提速1800倍。

趋势 涌现和融合是大模型时代重要的趋势

腾讯研究院、同济大学、腾讯云，共同发布《人机共生——大模型时代的AI十大趋势报告》。报告指出，涌现和融合是大模型时代重要的AI趋势，大模型推动AI进入通用人工智能阶段，而多模态AI的发展则提供了解决复杂问题和实现多模态智能处理

的能力。十大趋势包括：“多模态助力大模型解决复杂问题”“生成式AI带来更贴近人的交互方式”“模型即服务生态呼之欲出”“垂直领域应用是大模型的主战场”“大模型为数字人‘注入灵魂’”“AI大模型将帮助个体成为超级生产者”等。

本报记者 邵阳 杨玉红

AI 新观察

爱因斯坦曾说，我从来不想未来，因为它来得太快。

得益于生成式人工智能相关技术的迅猛发展，智能创作时代正缓缓拉开序幕。当我们的未来时，我们也该有一份对未来的思考，这份思考将帮助我们更好地前行。

人工智能的发展经历过多次春天与寒冬，每一次春与冬的交织都和“通用化”及“专业化”的分歧息息相关。伴随着人工智能飞跃式进步，越来越多复杂场景出现，若还是针对什么任务训练什么模型，这种一般的工程思想也背离

窥见美好未来时 更该多一份思考

邵阳

了人工智能的初衷，主打“通用化”的大模型在时代的浪潮下孕育而生。

通过在数据中捕捉更广泛、更精细的规律和关系，大模型可以生成更多样化且更真实的输出，让生成式人工智能变成现实，并且逐步深入我们的日常生活。

大模型在“狂飙”，却也必然伴随着“事故”，最显而易见的就是语言大模型常常提供错误或无意义的回答，要么就出现“幻觉”——我们常说的“一本正经地胡说八道”。

当然，最大的关注点，还是来源于伦理与法律。瑞莱智慧联合创始人、算法科学家萧子豪

认为，大模型“落地难”的本质在于，当前仍处于“野蛮生长”阶段，还没有找到场景、风险和规范三者之间的平衡点。

大家首先关心的，莫过于大模型带来的“失业焦虑”。淘码科技创始人兼首席执行官周健分析，人们对于人工智能可能替代白领工作感到焦虑，但实际上这种替代是有先后顺序的，“在企业内部的组织结构上，可能从金字塔型变为钻石型，这意味着可以用不同的视角来管理团队和控制质量”。

还有，大模型所生成内容是否“有危险”，同样值得警惕，特斯拉首席执行官埃隆·马斯克就说过：“如果不加以控制，人工智能或许很有可能摧毁整个人类。”当然，人工智能领域频出的“换脸”事件，也给大模型时代的数字治理敲响警钟，生成式人工智能的版权问题、导致的欺诈问题、生成违禁内容等，都需要法律法规给“狂飙”下的大模型系好“安全带”，把好“方向盘”。