

“看图作文”，机器也能做到

■本报记者 郑金武

向电脑机器里上传一幅“有一只小熊玩偶和一堆积木”的图画，询问机器：“积木在小熊的哪边？”机器快速做出回答：“积木在小熊的左边。”——这是预训练模型“文汇”的典型应用案例。

近日，北京智源人工智能研究院（以下简称智源研究院）发布面向认知的超大规模预训练模型“文汇”，旨在探索解决当前大规模预训练模型不具有认知能力的问题。这一项目由智源研究院发起的“悟道”攻关团队完成，团队成员包括来自智源研究院、阿里巴巴、清华大学、中国人民大学、中国科学院、搜狗、智普·AI、循环智能等单位的科研骨干。

“文汇”模型有针对性地设计了多任务预训练的方法，可以同时学习“文到文、图到文以及图文到文”等多项任务，实现对多个不同维度的概念理解。“智源研究院院长、北京大学信息技术学院教授黄铁军介绍说。

让机器理解自然语言

OpenAI 是由美国硅谷诸多科技公司负责人联合建立的人工智能非营利组织，其主要目标包括制造“通用”机器人和使用自然语言的聊天机器人等。

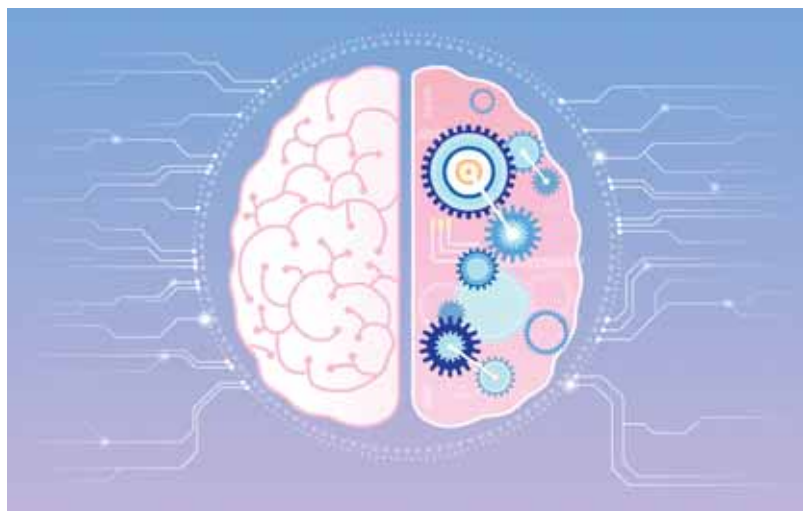
在现实世界里，人与人之间交流用的是自然语言。所谓自然语言，通常是指一种自然地随文化演化的语言，例如汉语、英语、日语等。

自然语言是人类智慧的结晶，也是人类交流和思维的主要工具。但要让机器像人一样自由地用自然语言交流，却是非常困难的事。这要求机器像人一样认识现实世界，像人一样学会思考。但是，显然目前的机器还无法做到。

人工智能在发展历程中，正好遇到自然语言处理这只“拦路虎”。可以说，自然语言处理是人工智能中最为困难的问题之一。

人类的知识学习与积累是一个复杂且长期的过程。青少年时期，人们不仅主动学习语文、数学、物理等知识，还通过游戏、生活等经历在大脑中不断积累知识。长大后，人类操作计算机时，会利用此前的知识背景，同时融合新的知识体系。而一个毫无知识储备的人，则很难理解这一切。

受此启发，研究人员为人工智能建立“预训练模型”，模仿人类学习积累的过程。当自然语言处理的知识储备逐渐增多，人工智能也随



“文汇”模型能够完成多种图文生成任务。

图片来源：昵图网

之更加智能，帮助科学家更好地开展任务处理。

2020年5月，OpenAI 发布了迄今为止全球规模最大的预训练模型 GPT-3。从此，超大规模预训练模型就成为人工智能领域研究的热点。今年1月初，OpenAI 又发布了 DALL·E 和 CLIP 两个大规模预训练模型，用于文本与图像链接处理。

据智源研究院学术副院长、清华大学计算机系教授唐杰介绍，此次发布的“文汇”模型，与 DALL·E 和 CLIP 这两个预训练模型类似，参数规模达 113 亿，仅次于 DALL·E 模型的 120 亿参数量，是目前我国规模最大的预训练模型，且已实现与国际领先预训练技术的并跑。

破解机器常识认知难题

近年来，OpenAI、谷歌、脸书等国际 IT 公司，都在持续推动大规模预训练模型的快速发展。

但唐杰指出，虽然 GPT-3 模型在多项任务中表现出色，却存在没有常识、不具有认知能力等问题。例如，如果向 GPT-3 模型提问：“长颈鹿有几只眼睛？”它的答案是：“两只眼睛。”此时，如果再继续提问：“我的脚有几只眼睛？”GPT-3 的回答恐怕就不那么尽如人意，因为答案依然是“两只眼睛”。

“这就不符合人类常识。”唐杰表示，GPT-3 等超大型预训练模型在处理复杂的推理任务时，例如开放对话、基于知识的问答、可控文本生成等，结果仍然与人类智能有较大差距。

与此航次，非常兴奋。此次海试主要是针对潜水器的性能而进行的万米测试，有船员、研发骨干等参与，需要上船的人非常多。在这样的条件下，还能拿出名额给科学家，说明各方对万米潜水器将来的科学应用非常重视。”贺丽生告诉《中国科学报》。

下潜两天前，贺丽生接到通知参加潜次。她非常激动，立刻开始准备相关的采样工具。“此次海试结果表明‘奋斗者’号载人潜水器的性能非常稳定，正因如此，总师团队及专家组才决定让科学家参与本航次的最后一个万

米潜次。深海环境温度约为 2~4 摄氏度，尽管潜水器内部已经做了非常人性化的处理，但要在其中工作 12~13 个小时，行程的后半段时间还是需要穿上厚外套、厚袜子。”贺丽生回忆说。

在此次潜次中，贺丽生携带的作业工具主要包括 Pushcore 沉积物取样器、生物铲、保温生物箱、保压取水仪、原位化学实验装置及 Niskin 诱捕瓶。这些装置可帮助他们带回万米深渊沉积物、生物、水样及岩石等科研样品。

万米深渊之中，给贺丽生印象最深的除了大量钩虾外，还有大量的海

参与多毛类生物。尽管下潜前，她推测万米深渊中会存在海参，但其数量之多还是让贺丽生惊叹。此外，她推测万米深渊中生活着海葵及水母，但在此次下潜任务选定的区域内，很“遗憾”未能看到。

“深海是人类至今了解最少的生物圈之一。”贺丽生表示，“奋斗者”号载人潜水器的研发成功无疑给人类探索深海、深渊添上了一对强劲的翅膀。它不但可以到达万米深渊，而且具有强大的作业能力，搭载多达 240 公斤的作业工具及样品，配合左右机械手的使用，使科学家可以在万米海底进行海底原位实验，也可以将科研样品带回陆地实验室。“奋斗者”号全海深载人潜水器使人类探索深海全程无盲区，助力人类了解深海。

唐杰也表示，“文汇”模型不仅使用数据驱动的方法来建构预训练模型，还将用户行为、常识及认知联系起来，主动“学习”与创造。本次发布的“文汇”模型可以完成“图图生文”等任务，具有一定的认知能力。

据介绍，目前 OpenAI 也在研发更大规模的预训练模型 GPT-4 参数。唐杰认为，未来，GPT-4 参数又会增大至少 10 倍，而且处理的数据将会更加多模态，能够处理文字、图像、视觉、声音等。

可完成多种图文生成任务

据介绍，“文汇”作为面向认知的超大规模预训练模型，研究人员在研发过程中，提出了针对多模态认知生成的大规模预训练的架构。

对于上传图像的处理，研究人员进行了切块并对块提取特征的处理。这些特征以及对应的位置编码，可以让图像和文本组合在一起送入模型，从而实现图像的理解。这让“文汇”模型能够完成多种图文生成任务。

打个比方，如果研究人员将一张身着旗袍女子的照片上传给“文汇”，在商场场景下，“文汇”微调参数后模型将给出如下描述：“旗袍是一种古典的服饰，它能够凸显出女性的优雅、知性和美丽。这款旗袍采用了经典的圆领设计，穿着舒适自在；精致的刺绣工艺，展现出华贵和大气……”而如果上传一张建筑图像，同时配上提示“走进平遥县宋家庄村，映入眼帘的是”，“文汇”模型将给出相应描述：“一座座古色古香的明清建筑，这里有着深厚的历史文化底蕴和独特的民俗风情。走进村子，就像走进了一个童话故事里的世外桃源……”

“文汇”可谓“望图生文”，这显示了“文汇”模型在常识认知和“图图生文”方面的强大功能。

不仅如此，目前，“文汇”模型已有四个样例应用可用于展示模型效果，包括基于上传图片的开放域问答、用语言操作数据可视化、基于预训练语言模型的诗词创作应用、可控人设的开放域问答等。

唐杰表示，现在“悟道”项目研究团队正在视觉等更广泛的范围内，对大规模预训练模型方法开展探索研究，已经启动了四类大规模预训练模型研制，包括“文源”（以中文为核心的超大规模预训练语言模型）、“文汇”（面向认知的超大规模预训练模型）、“文澜”（超大规模多模态预训练模型）和“文溯”（超大规模蛋白质序列预训练模型）。

而在 2020 年 11 月 14 日，智源研究院已发布了“文源”（以中文为核心的超大规模预训练语言模型）第一阶段 26 亿参数规模的中文语言模型。

黄铁军表示，下一步智源研究院将联合优势单位加快四类大规模预训练模型的研发进度，特别是“文汇”模型，未来将着力在多语言、多模态条件下，提升完成开放对话、基于知识的问答、可控文本生成等复杂推理任务的能力，使其更加接近人类水平。

一所人一事



朱光武与探测二号（TC-2）卫星合影。

朱光武

中科院国家空间科学中心研究员、博士生导师、副总工程师、学术委员会副主任。50 多年来一直从事航天工程和空间科学探测研究工作，主持我国神舟飞船、各类应用卫星空间环境探测和研究以及科学卫星等有关航天器有效载荷研制任务。

每日清晨 6 点，无论寒暑、不避风雨，一位 70 多岁的老者都会在一间普通办公室里准时开始他一天的工作。这位数十年如一日，退休后依然坚持工作、每天都能在大部分年轻人之前到达办公室的老者就是中国科学院国家空间科学中心空间环境探测重点实验室创始人之一、我国著名空间环境探测专家朱光武。

一生只做一件事。从风云卫星、载人航天到探月工程、萤火一号，朱光武从事空间环境探测工作 50 多年来，独立设计开发了 16 款专用集成电路，填补了这一领域多年来的多项空白，在我国天基空间环境探测技术从无到有、从有到精、形成系列并赶超航天强国的发展历程中做出了杰出贡献，用青春和热血、坚贞与无私谱写了一曲奋斗之歌。

知难而进 拼出天地

直面辐射风险，只为新技术试验。在空间环境中，辐射的危害极大，对卫星和人体均会造成不可逆的破坏性影响。我国科研领域刚刚起步之时，科研的环境与条件十分艰苦且有限。为保证航天员和卫星辐射危害的测量更加精准并进行有效规避，朱光武及一众探测人在保护措施有限和辐射风险，无法回避的情况下，每每直面辐射风险，出色完成仪器定标，保证了技术实现和提升。

一支笔，百次改，绘就复杂空间环境探测仪器设计图。在那个没有计算机，更没有 EDA 软件的时代，朱光武一手拿着放大镜，一手握笔，一笔一画地绘制原理图与 PCB 图。不知多少次，一个新的想法或是制版工人一个新的建议都让他再一次打起一百二十分的精神重新制图。

付出与坚持换来宝贵发展机会。一年春节，朱光武带队到上海参加某型号任务空间环境载荷的重要论证会议。途中乘车发生意外，他的头部、腿部多处受伤，随行人员劝其尽快去医院处理，但他仍坚持赶往会场。与会人看到他脸上的血迹都被他的坚持所感动。朱光武的拼搏为天基空间环境探测在困难时期的发展争取到了宝贵的机会。

在朱光武等老一辈科学家的拼搏与带领下，如今的探测室已经成为我国空间环境 / 空间探测领域当之无愧的“国家队”。

自主自强 硕果报国

“我们一定要将核心技术牢牢掌握在自己手里，不断发展，精益求精。只有这样，我们才能对自己的仪器和数据有信心。”这是朱光武经常对身边人说的。

一句话。他也是如此身体力行、一以贯之的。

朱光武独立设计、开发了 16 款专用集成电路，打破了空间环境探测领域专用集成电路依赖国外的危局。基于这些自主集成电路开发的 20 多种空间环境探测仪器，广泛应用于载人航天、风云系列等多个空间环境探测任务，为我国以航天工程应用为目的的天基空间环境探测仪器从无到有、从有到精奠定了坚实的基础。

领域深耕 培养后浪

50 年来，朱光武主持我国神舟飞船、各类应用卫星空间环境探测、研究以及科学卫星等有关航天器有效载荷研制任务，先后承担我国多颗卫星工程任务及我国跨世纪重大航天任务。

朱光武在科学研究的的同时亦不忘实验室的发展与年轻一辈的成长。他总是积极培养并严格锻炼年轻的科研队伍。

朱光武有一份特殊的礼物传给后辈，那就是他倾注几十年心血，记录每一次成功失败经验和详细数据、几十次型号任务心得体会的笔记。这份半米高的手稿被亲切地称为“探测红宝书”。实验室的年轻后辈，手捧“探测红宝书”，站在他的肩膀上快速成长。

如今年过七旬的朱光武仍然同实验室年轻人一起奋斗在科研事业的第一线，指导和帮助年轻人从事空间环境探测研究工作，为国家持续输送科研人才。

老一辈探测人艰苦奋斗谋开拓，中生代探测人秉志守节夯实基础，新时期探测人扶掖而上竞前行。一代又一代探测人将不断砥砺前行、踏浪而行，为我国空间环境探测事业发展继续贡献力量，在中华民族伟大复兴的征程上书写自己的传奇。

（作者单位：中科院国家空间科学中心）

谱写自主自强的空间环境探测奋斗之歌

■张贤国 张志良 孙越强

又创造了一个“不可能”！

（上接第 1 版）

助力探索深海全程无盲区

“探索二号”为科学家建设了一座海上科学实验室，为深海深渊科考提供了“利器”。

深海所研究员贺丽生作为科学家代表参与了此次航次，执行“奋斗者”号载人潜水器的万米海试任务。她也是中国首位进行万米深潜的女科学家。

“我作为一名普通科学家能够参

面对数字未来：敬畏、进取、向善

■郭凯天

四年前，腾讯研究院第一次提出“科技向善”的时候，很多人都有疑问：“善”是什么、标准怎么定，以及怎么向善，具体怎么操作，在市场和社会上怎么去协同？那时候，大家对这些问题想得还不是很清楚。

2019 年，“科技向善”正式成为腾讯使命愿景的一部分。这一年多来，尤其是经过 2020 年这一特殊的年份，我想大家都非常重视科技向善。在今天，当我们面对未来时，大家也发现科技向善实际上有更多的内容需要思考，有更多内容需要去实践。

过去一年，大家感受最深的可能就是新冠疫情。当疫情发生时，学校停课、饭店关门、商场也暂停营业。在众多的“门”关闭之后，互联网平台向全世界打开了数字大门。

疫情期间，数字科技给这个不确定的世界带来了更多的确定性。然而在去年下半年，对于数字技术、数字未来的担忧显著增加，包括

平台、大数据杀熟、个人隐私、社区买菜等等，都引发了社会广泛的关注和热议。

当社会各行各业都将深度融入数字世界时，这种担忧实际上不仅仅是对技术、对平台、对数据，也不仅仅是某一技术或商业形态的问题，而是对整个社会，甚至是对人类未来将迈入数字世界的一种深刻担忧。

这其中，我们可以看到传统和数字未来的冲突，可以看到国内和国外的冲突，更可以看到的是人类的现在与未来发展的冲突。

从 1994 年中国第一次接入国际互联网，到今天已经走过了 27 年。回顾中国互联网发展和崛起的道路，其实和我国改革开放密不可分，可以说中国互联网经济的崛起，不是某一个企业、某一项技术的事情，而是在民族复兴的大背景下，得益于改革开放，得益于互联网基础设施的完善和普及，得益于全国民众对数字化和互联网的拥抱，得益于政策的支持。

因此，当考虑互联网产业或者企业发展的时，就不得不把它置于中国改革开放四十年大的背景下，置于国家和社会对互联网、对数字技术未来期望的大背景下。在这个大背景下考虑问题，科技向善就显得尤为重要。它应该是中国社会甚至人类社会，在迈向数字化未来过程中一条坚持不动摇的基本准则。

科技向善首先是敬畏，不仅敬畏监管、敬畏舆论，更要敬畏国家的发展、社会的发展，尤其要敬畏人类未来的发展。敬畏未来是对行业、对技术发展如何跟社会结合最深刻的一层关系。

只有心怀敬畏，才能在科技发展的同时，在行业成长的同时，与传统、与社会较好地结合，去关注更多人的利益、关注社会的利益。唯有如此，行业和技术才能发展得更稳健。

科技向善的第二点就是进取。所谓敬畏，不是畏手畏脚，而是要在敬畏的基础上，更好

地找准自己的定位——在国家、在社会、在技术、在历史长河里的定位。

有了这个定位，实际上要做的是更有进取心，尤其是科学技术上的进取。我们看到，企业的发展需要在基础技术上投入和拓展，为未来的竞争和发展抢得先机；我们也看到，国家和全社会对互联网企业抱有非常高的期望和热情。在这方面，企业更应该在技术、在科学基础上去进取、去探索、去突破。

科学技术对人类发展的推动，大家都有目共睹。过去几百年的人类历史，在工业革命以来所推动的产业和社会变革上，已经充分说明了这个问题。以互联网为代表的数字技术，也一定能够继续推动人类社会的进步和发展，关键就是要在敬畏的基础上保持进取心。

最后一点，科技向善的目标是“向善”，不忘初心就是向善。

还记得 20 年前我们这一代人刚投身互联网行业那会儿，很多互联网公司规模都非常小，但是大家都把“互联网让世界更美好”作为公司的使命愿景。通过互联网技术让世界效率更高、社会更公平，让大家都能够享受到互联网发展的福祉，我想这应该就是中国互联网行业在发展起步阶段的初心。

今天的互联网人仍然不会忘记这个初心，坚持这个初心就是向善。互联网应该在人类未来的数字发展中发挥更大、更好的作用，推动经济发展，推动教育、医疗，推动技术造福于人，推动国家富强，推动更多人一起分享互联网发展的福祉。这就是向善的一个目标。

对向善的理解，我们都还在摸索中，没有一个固定的标准。它是一个愿望、一个初心，也是我们未来的方向和目标。希望我们能在在这条路上，和社会各界共同努力、共同探索。

（作者系腾讯集团高级副总裁、党委书记）