



2022年，人工智能带给人类更多惊喜

◎ 实习记者 都芃

AI“黑科技”照亮北京冬奥会

助力天气预报、比赛转播和手语播报等

2月4日，全球瞩目的2022年北京冬奥会正式拉开帷幕。人工智能等技术的应用为本届冬奥会增添了别样的“科技之美”。

在此次冬奥会上，由中国科学院院士、北京大学副校长、北京大数据研究院首席科学家张平文领衔研制的人工智能MOML算法赋能天气预报模型，使冬奥会天气预报更加精准。人工智能算法在融合、处理信息中的先天优势，使其在一定程度上可以代替预报员在会商中进行信息整合、分析，通过数据挖掘与学习，将预报员的经验内化在算法中，在提高天气预报效率的同时，也进一步提高了预报的准确率。

在本届冬奥会自由式滑雪女子大跳台决赛中，中国选手谷爱凌以“逆天”的精彩表现获得个人首金。在比赛转播过程中，百度智能云通过“3D+AI”技术打造出的“同场竞技”系统，将单人比赛项目变成“多人比赛”，实现冠军、亚军比赛画面的三维恢复和虚拟叠加，方便观众看到不同选手的实时动作；同时，通过技术手段对运动员动作进行量化分析，将滑行速度、腾空高度、落地远度、旋转角度等一系列运动数据与原画面叠加起来，使观众可以更直观地从流畅性、完成度、难度、多样性和美观度等角度看懂选手之间的技术动作差异。

在北京冬奥会开幕的同一时间，央视新闻AI手语主播也正式上岗，她在冬奥会新闻播报、赛事直播和现场采访中，为听障人士送上了实时手语翻译服务。凭借精确的手语翻译引擎，该AI手语主播可懂率达85%以上，可将冰雪赛事的文字及音视频内容，快速精准地转化为手语。

腾讯“混元”AI大模型登顶VCR榜单

展现了其在多模态理解领域的强大实力

5月31日，腾讯“混元”AI大模型在多模态理解领域



在世界人工智能大会上，用户输入文字，AI就能根据语义进行绘画创作。
视觉中国供图



在国内首个乘用车无人化运营试点北京经济技术开发区，一辆“主驾无人、副驾配备安全员”的无人驾驶车在行驶中。
新华社记者 彭子洋摄

域国际权威榜单VCR(Visual Commonsense Reasoning, 视觉常识推理)中登顶，两个单项成绩和总成绩均位列第一。这是继在跨模态检索领域大满贯、CLUE自然语言理解分类榜及CLUE总榜登顶后，“混元”AI大模型的又一重大突破，展现了其在多模态理解领域的强大实力。

与跨模态理解任务不同的是，多模态理解任务要求计算机除了能够做到识别层次的感知(如分类检测等)，还需要达到认知层次的感知(如判断意图、逻辑推理等)。

此次登顶VCR榜首的“混元”AI大模型由腾讯广告多媒体AI团队自主研发，同时借助腾讯太极机器学习平台的图形处理器算力和训练加速框架，在预训练任务、训练方式上进行了诸多创新改进和设计，有效提升了模型性能。

截至目前，“混元”AI大模型在MSR-VTT、MS-VD、CLUE、VCR等多个领域的AI权威榜单中取得了第一名的成绩，并刷新多项行业历史纪录。这意味着，“混元”在自然语言理解、多模态理解、跨模态理解等领域的技术实力已得到验证。

谷歌工程师闹乌龙，称AI存在意识

人工智能所谓的“人格”更多只是模仿人类罢了

谷歌AI工程师闹乌龙，称LaMDA语言模型有意识，引发业界对“AI是否拥有自主意识”的讨论。

今年6月，谷歌公司AI工程师莱莫因认为对话用语言模型LaMDA具有了“自主意识”，并对此出具了长达21页的证据。莱莫因认为LaMDA具有意识的原因有三：一是LaMDA以前所未有的方式高效、创造性地使用语言；二是它以与人类相似的方式分享感觉；三是它会表达自省和想象——既会担忧未来，也会追忆过去。

LaMDA是谷歌在2021年开发者大会上公布的大型自然语言对话模型，它可以模拟任何带有知识属性的实体，通过“拟人”的方式，在与人类亲切自然的对话中为用户答疑解惑，传递更多知识。

莱莫因的观点和证据引起了业内的广泛关注。不久后，谷歌发表声明称，莱莫因违反了“就业和数据安全政策”，将其解雇。谷歌表示，经过广泛地审查，他们发现莱莫因关于LaMDA是有生命的说法是完全没有根据的。

专家普遍认为，当下人工智能所谓的“人格”，更多只是模仿人类的语言风格，有自我意识、有感能力的AI应该具备能动性，并具有独特的视角看待人和事，但目前AI还只是人们设计的一个计算机系统，作为工具来做一些特定之事。

全球首个图、文、音三模态大模型诞生

“紫东太初”实现“以图生音”和“以音生图”

9月1日，在上海举办的2022世界人工智能大会上，由武汉人工智能研究院、中国科学院自动化研究所和华中科技大学联合研发的“紫东太初”多模态大模型项目获得了此次大会的最高奖项。“紫东太初”是全球首个图、文、音三模态大模型，开创性地实现了图像、文本、语音三模态数据间的“统一表示”与“相互生成”，实现了“以图生音”和“以音生图”，理解和生成能力更接近人类，为打造多模态人工智能行业应用提供创新基础，向通用人工智能迈出了重要一步。

“紫东太初”三模态间的相互转换和生成，其核心原理是视觉、文本、语音不同模态通过各自编码器映射到统一语义空间，然后通过多头注意力机制学习模态之间的语义关联以及特征对齐，形成多模态统一知识表示；之后，再利用编码后的多模态特征，通过解码器分别生成文本、图像和语音。

“紫东太初”凭借四大突破，有效助力以多模态认知为核心的通用人工智能发展。一是首次提出多层次、多任务跨模态自监督学习框架，支持从词级走向

模态级、样本级的三级预训练自监督学习方式；二是首次完成弱关联多模态数据语义统一表示，减少数据收集与清洗代价；三是首次实现多模态理解与生成任务的统一建模，支持跨模态检索、多模态分类、语音识别、图像生成等理解与生成任务；四是首次实现无监督超越有监督方法，基于5%—10%的数据标注，实现100%的有监督学习效果。

AI打破矩阵乘法计算速度纪录

解决了50年来数学领域一个悬而未决的问题

10月，英国《自然》杂志封面以“矩阵游戏”为题，发表了人工智能公司“深度思维”团队的最新发现：AI可以解决矩阵乘法问题。这款名为“AlphaTensor”的AI系统能自行发现新算法，从而解决了50年来数学领域一个悬而未决的问题——找到两个矩阵相乘最快的方法。这是第一个可为矩阵乘法等基本任务发现新颖、高效且正确算法的AI系统。

数学在计算机编程中经常出现，通常作为描述和操纵现实世界现象表示的一种手段。例如，它可用于表示计算机屏幕上的像素、天气状况或人工神经网络中的节点。在这种情况下，使用数学的主要方式之一，就是对矩阵进行计算。矩阵越大，工作量也越大，计算机科学家开始花费大量时间和精力开发更加有效的算法来完成相关工作。

在此次最新成果中，“深度思维”团队研究人员探究了是否有可能使用基于强化学习的AI系统来创建新算法，从而使计算步骤比现有算法更少。

为了找到答案，他们从游戏系统中寻找灵感。在构建了一些初步系统之后，研究团队将重点转向了树搜索，这是系统在特定情况下查看各种方案的一种方法。

接下来，研究人员将允许系统创建自己的算法，进一步提高效率。他们发现，在许多情况下，系统选择的算法比人类创建的算法更好。“深度思维”团队希望，未来AI能更多地用来帮助攻克数学和科学领域的一些重要的难题。

2022中国人工智能创新发展指数公布

全面反映我国人工智能发展态势

11月18日，第五届世界声博会暨2022科大讯飞全球1024开发者节开幕式上，中国电子信息产业发展研究院(又称赛迪研究院)发布了2022中国人工智能创新发展指数(合肥指数)。

这是国内首个以地区冠名的全国性人工智能专题研究成果，旨在全面系统地反映我国人工智能的发展态势。中国电子信息产业发展研究院从发展环境、创新能力、基础配套、资本投入和产业应用5个维度，构建了中国人工智能创新发展指数，也就是“合肥指数”的评价体系。

近年来，我国人工智能步入与经济深度融合应用新阶段，智能化转型全面推进，人工智能产业在全球的影响力不断增强。2021年，我国人工智能的研发强度为19.4%，从业人数增加到31万人，占全球比重的5.3%。2017年至2021年，我国人工智能产业规模增长了2.6倍，占全球比重提升到16.8%。专利申请量占全球比重持续扩大，从2012年的13%增长到2021年的70.9%。创新能力上，我国人工智能研发投入力度不断加大，从业人数不断增加。

从总体指数来看，北京、广东和上海处于人工智能领域的领跑地位，安徽则紧随其后，排在全国的第6位。合肥已经成为人工智能领域、科技创新与产业发展最活跃的城市之一。

ESMFold预测六亿多种蛋白质结构

预测速度比“阿尔法折叠”快60倍

英国“深度思维”公司8月曾宣布，其开发的人工

智能程序“阿尔法折叠”已预测出约100万个物种的超过2亿种蛋白质结构，几乎涵盖了科学界已编录的每一种蛋白质结构。但就在今年11月，元宇宙平台公司(Meta)研究人员利用人工智能模型ESMFold预测了来自细菌、病毒和其他尚未被表征微生物的6亿多种蛋白质结构。

在此次最新研究中，研究团队利用大型语言模型来预测这些蛋白质结构。据悉，语言模型通常需要大量文本进行训练，为将这一模型应用于蛋白质结构预测，研究团队利用已知的蛋白质序列来训练它，这些已知的蛋白质可由20个不同氨基酸组成的链来表达，每个氨基酸由一个字母表示。然后，ESMFold学会了用模糊的氨基酸比例“自动完成”蛋白质结构预测。

该团队负责人亚历山大·里维斯表示，这些训练让ESMFold对包含蛋白质形状信息的蛋白质序列有了直观了解。而且，与“阿尔法折叠”一样，这一模型将这些了解到的信息与已知蛋白质结构和序列之间的关系信息结合，生成预测结构。

团队指出，ESMFold的预测虽然不像“阿尔法折叠”那么准确，但在预测速度上要快60倍，这意味着它可将结构预测数据库扩展到更大。

首创蛋白质动态结构AI建模方法

对理解生命过程、研发新型药物有着重要意义

12月8日，西湖大学公布了该校人工智能讲席教授李青团队联合厦门大学、杭州德睿智药科技有限公司首创研发的能够刻画蛋白质构象变化与亲和力预测的AI模型——ProtMD。这是第一个尝试解析蛋白质动态构象的人工智能模型，可辅助药物化学专家更加精准地筛选出高活性小分子，从而加速临床前药物研发。

此前谷歌旗下公司研发的“阿尔法折叠2”能够利用人工智能准确预测蛋白质的三维结构，对结构生物学、药物设计乃至整个科学界都产生了巨大影响。但“阿尔法折叠2”只能预测蛋白质在一个瞬间的静态结构，尚未能解决蛋白质结构动态变化的预测。李青团队此次开发的AI模型，在给定药物分子和靶点蛋白的情况下，可预测药物分子与生物体内靶点蛋白质结合(柔性对接)后蛋白质结构的变化过程，推断药物与靶标蛋白结合的稳定性，预测药物功能，从而提升AI药物设计的精度和效率。

李青表示，预测蛋白质结构的动态变化对理解生命过程、研发新型药物都有着十分重要的意义。尤其在AI药物设计中，通过对药物分子与靶点蛋白结合后的动态结构变化进行预测，评估药物-靶点结合亲和力和药物效果，是提高AI药物筛选准确性和效能的重要思路。

多城市推动自动驾驶行业发展

我国自动驾驶行业正式向L3级迈进

2022年是自动驾驶行业具有里程碑意义的一年，有关政策密集出台，相关应用从研发测试走向大规模商业化试点。当前，全国近30个城市已累计为80余家企业发放了超过1000张道路测试牌照，允许高等级智能网联汽车在特定场景、特殊区域内开展规模化载人载物测试示范。越来越多的城市正在推进更高等级的

自动驾驶商业化。今年8月1日，《深圳经济特区智能网联汽车管理条例》开始实行，该条例提出L3级自动驾驶在行政区全域开放道路测试、示范应用，探索开展商业化运营试点，标志着我国自动驾驶行业正式向L3级迈进。

此后，重庆、武汉等地政府部门也先后发布了自动驾驶全无人商业化试点政策，并向百度发放全国首批无人化示范运营资格，允许车内无安全员的自动驾驶车辆在社会道路上开展商业化服务。

此外，为推动智能网联汽车产业健康有序发展，工业和信息化部会同公安部还组织起草了《关于开展智能网联汽车准入和上路通行试点工作的通知(征求意见稿)》，拟遴选符合条件的道路机动车辆生产企业和具备量产条件的搭载自动驾驶功能的智能网联汽车产品，开展准入试点；对通过准入试点的智能网联汽车产品，在试点城市的限定公共道路区域内开展上路通行试点。

AI绘画火了，AIGC元年开启

未来预计能够产生万亿级经济价值

今年8月，在美国科罗拉多州举办的新兴数字艺术家竞赛中，参赛者杰森·艾伦提交的AIGC绘画作品——《太空歌剧院》，获得了此次比赛“数字艺术/数字修饰照片”类别一等奖。没有绘画基础的杰森·艾伦借用了一款名叫Midjourney的AI绘图工具，通过一个类似“文字游戏”的过程，输入题材、光线、场景、角度、氛围等有关画面效果的关键词后，得到了初始作品，并在反复调整和修改后最终完成了这组“太空歌剧院”数字艺术作品。

这一年，AI绘画小程序、网站等开始迅猛增长，而美图秀秀、抖音等软件也加入了AI画图功能。抖音平台数据显示，截至12月6日，已有超2428.4万人使用该特效，迅速飙升至特效潮流榜第一位。AI绘画的百度指数也从日均两三千上升到日均3万，火爆程度可见一斑。

AI绘画的火爆也让AIGC这一概念逐渐进入大众视野。

所谓AIGC(AI Generated Content)，即基于人工智能技术自动生成内容的新型生产范式。其技术主要涉及两个方面：自然语言处理(NLP)和AIGC生成算法。其中，自然语言处理是实现人与计算机之间通过自然语言进行交互的手段。

最初，AIGC生成的内容形式以文字为主，经过2022年指数级的发展，目前AIGC技术可生成的内容形式已经拓展到了包括文字、图像、视频、语音、代码、机器人动作等多种内容形式，2022年也因此被称为“AIGC元年”。生成式AI让机器开始大规模涉足知识类和创造性工作，未来预计能够产生数万亿级的经济价值。

2022 年终盘点