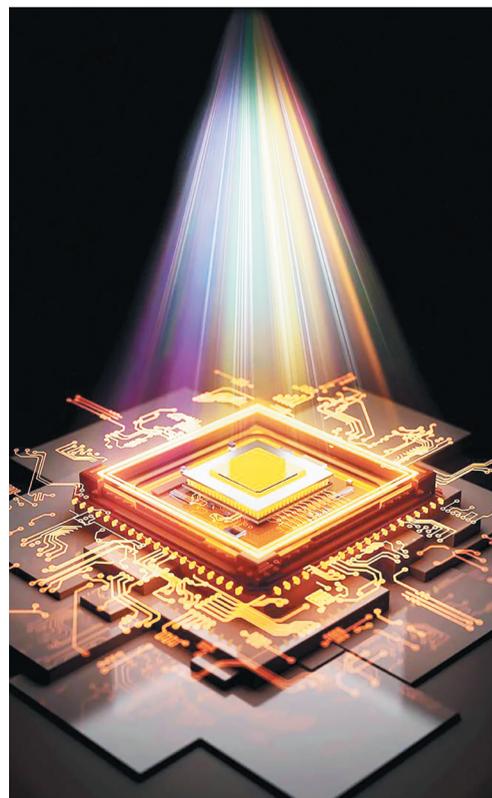


2024 人工智能 四大趋势

2023年,世人见证了ChatGPT在全球范围的大火。以生成式人工智能为代表的新一代人工智能问世,改变了人工智能(AI)技术与应用的发展轨迹,加速了人与AI的互动进程,是人工智能发展史上的新里程碑。2024年,人工智能技术与应用的发展又会呈现出哪些趋势?让我们一同展望这些值得关注的重大趋势。



- ◆ 据消息人士称,OpenAI正在训练下一代的人工智能,暂名“Q*”(读作Q-star)。新的一年,OpenAI下一代产品可能发布
- ◆ 数据瓶颈指的是可用于训练AI的高质量数据的有限性,合成数据有望打破这一瓶颈。除了对大量高质量数据的需求导致合成数据受到追捧以外,对数据安全的考量也是重要原因
- ◆ 作为全球性能最强的AI,ChatGPT已遇到算力等方面的瓶颈。在此背景下,讨论量子计算机在人工智能领域的应用就成为一种颇具潜力的未来解决方案
- ◆ AI代理和无代码软件开发带来“冲击波”

趋势一

从AI大模型
迈向通用人工智能

2023年,ChatGPT开发者OpenAI被置于前所未有的聚光灯下,也使GPT-4后续版本的开发被推向了风口浪尖。据消息人士称,OpenAI正在训练下一代的人工智能,暂名“Q*”(读作Q-star)。新的一年,OpenAI下一代产品可能发布。据媒体爆料,“Q*”可能是第一次采用“从零开始”的方式训练的人工智能。其特点是,智能不来自人类活动的数据,且其有能力修改自身代码以适应更复杂的学习任务。前者使得人工智能能力的发展变得愈发不透明,而后者向来被看作是诞生人工智能“奇点”的必要条件。在人工智能发展领域,“奇点”特指机器拥有了自我迭代的能力,进而在短时间内迅猛发展,导致超出人类控制。

虽然一些报道称,“Q*”目前还只能解决小学难度的数学问题,距离“奇点”还远。但鉴于虚拟环境中人工智能迭代速度可能远超想象,其仍然可能在不远的将来自主发展出在各个领域均可超过人类水平的AI。2023年,OpenAI预言,各方面超越人类水平的人工智能在十年内就会出现;英伟达创始人黄仁勋表示,通用人工智能可能在五年内超越人类。

一旦通用人工智能得以实现,就可被用于解决各种复杂的科学难题,譬如寻找外星人与地外宜居星系、人工核聚变控制、纳米或超导材料筛选、抗癌药研发等。这些问题通常需要花费人类研究员数十年的时间来寻找新的解决方案,部分前沿领域的研究量已超出人力极限。而通用人工智能在自己的虚拟世界中拥有几乎无限的时间和精力,这使得其在部分容易虚拟化的任务中,有可能成为人类研究员的替代。但届时,人类如何监督这些从智能水平上超过人类的人工智能,确保其不会危害人类,又是一个值得思考的问题。

趋势二

合成数据
打破人工智能训练数据瓶颈

数据瓶颈指的是可用于训练AI的高质量数据的有限性,合成数据有望打破这一瓶颈。

合成数据是在模仿真实数据的基础上,由机器学习模型利用数学和统计科学原理合成的数据。关于什么是合成数据,有一个较为浅显易懂的比喻:这就像是在给AI编写专门的教材。例如,尽管英文课本的对话中出现的可能是“小明”“小红”这样的虚构人名,但并不影响学生们由此掌握英语,因此从某种意义上,对于学生而言,教材就可以看作一种经过编纂、筛选和处理的“合成数据”。

有论文表明,模型的规模至少要达到620亿参数量后,才可能训练出“思维链”能力,即进行分步骤的逻辑推理。但现实的尴尬在于,迄今为止人类产生的不重复的、可供训练的优质数据并没有这么多。使用ChatGPT等生成式人工智能以前所

未有的数量产生高质量合成数据,未来的AI将由此获得更高的性能。

除了对大量高质量数据的需求导致合成数据受到追捧以外,对数据安全的考量也是重要原因。近年来,各国纷纷出台更严格的数据安全保护法律,使得客观上利用人类产生的数据训练人工智能变得更为繁琐。这些数据中不仅可能隐含个人信息,其中的许多数据还受版权保护。在互联网隐私与版权保护尚未形成统一标准与完善架构的当下,使用互联网数据进行训练,极易导致大量法律纠纷。而若考虑对这些数据进行脱敏,又面临筛查识别准确率方面的挑战。两难之下,合成数据就成为最惠而不费的一种选择。

从以上分析中可以看出,合成数据可以说是颇具开创性的,有望解决此前发展人工智能与数据隐私保护不可得兼的问题。但与此同时,如何确保相关的公司和机构负责任地制作合成数据,如何制作出既符合本国文化与价值观,又在规模和技术水平上足以媲美西方以英文网络资料为中心的合成数据训练集,也将成为中国面临的一个颇具挑战性的课题。

趋势三

量子计算机可能
率先应用于人工智能

作为电子计算机发展到今天的最前沿应用,人工智能始终存在算力不足的隐忧。ChatGPT问世数月后,OpenAI总裁奥特曼曾公开表示,其并未鼓励更多用户注册OpenAI。2023年11月,OpenAI甚至宣布暂停ChatGPT Plus付费订阅新用户的注册,以确保现有用户拥有高质量体验。显然,作为全球性能最强的AI,ChatGPT已遇到算力等方面的瓶颈。在此背景下,讨论量子计算机在人工智能领域的应用就成为一种颇具潜力的未来解决方案。

首先,人工智能领域的算法,大部分属于并行计算的范畴。

其次,运行ChatGPT所需的硬件条件,同样也十分适合导入当前体积庞大的量子计算机,二者都需要安装在高度集成的计算中心里,由一支专业化技术团队进行管理支撑。

什么是量子计算机?量子计算机是一类遵循量子力学规律进行高速数学和逻辑运算、存储及处理量子信息的物理装置。其不仅体积庞大,而且作为核心零部件的“量子芯片”,通常要被置于接近绝对零度(零下273.15摄氏度)的极低温中,利用在这种极低温下部分微观粒子表现出的量子特性进行信息运算和处理,且运行结果只能存在几毫秒的时间。

2022年,来自谷歌、微软、加州理工学院等机构的研究者从原理上证明了“量子优势”在预测可观测量、量子主成分分析以及量子机器学习中的存在。量子机器学习,实际上就是量子计算在人工智能领域的应用,也体现出未来量子计算与人工智能两大前沿技术合流的趋势。

有关量子机器学习算法的研究,已成为新的研究热点。不过,未来量子计算机不会完全取代电子计算机,更有可能出现的是量子计算机和电

子计算机在不同的应用场景下发挥各自所长,实现协同发展,既极大提升算力,也兼顾成本和可行性。

趋势四

AI代理和无代码软件开发
带来“冲击波”

在AI应用方面,2024年值得关注的是AI代理和无代码软件开发带来的“冲击波”。

一是AI代理对劳动力结构的冲击。

截至目前,全球至少已有近两亿人使用人工智能大模型。但人们已不再满足于坐在电脑前跟AI“聊天”,而是开始开发能够自动根据任务需要向人工智能发出提示的工具。当自动提示工具与大模型两相结合,AI代理便由此诞生。

2023年4月,OpenAI联合创始人布罗克曼现场演示了GPT的“自动模式”。在该演示中,AI代理几乎“包办”了一场晚宴:不仅根据要求生成了一份晚宴的推荐菜单、一份图文并茂的邀请函,还自动将该菜单需要购买的食材加入生鲜电商APP的购物车,并自动发布了一条有关该晚宴的社交网站帖子。

AI代理还能根据比较模糊的需求提示自动制作网站,自动完成各种需要使用Office软件完成的文字和表格处理工作,甚至自动根据已有论文数据进行归纳总结生成分析论文等。

比尔·盖茨近日发文解读AI代理未来,表示AI代理将彻底改变人们使用计算机的方式,带来来自键盘、屏幕和鼠标发明以来人类与计算机互动方式上最重大的革新。

AI被看作对人类的信息收集、分析和处理进行增强的扩展性工具,使得人的工作水平更新上新台阶。但与此同时,AI代理也给许多现有的工作岗位带来冲击,因为企业可能尝试雇用更少的人来完成相同的任务。

二是无代码软件开发给数字经济创新带来的影响。

尽管生成式人工智能可能淘汰掉一批传统数字岗位,但在关上一扇门的同时也打开了一扇窗,这就是“无代码软件开发”。目前,以AI大模型为基础的编程辅助工具已经发展到一个新的阶段,能够根据用户十分模糊的指令来生成软件或网页代码。

对此,政府需转变观念,兼顾市场监管与促进创新,一方面降低数字创新过程中的注册与融资门槛,打通中小企业发展壮大过程中的痛点,让就业与创新政策适应“人人皆可创新”的新需求;另一方面需要探索更有利于保护创新“点子”的版权与专利保护新政策,从而激励那些能够不断提出创新“点子”的人才。

综上所述,展望2024年,无论是人工智能技术自身的迭代发展,还是其对数据价值的重塑,抑或是向各行业、各领域的应用渗透,人工智能的影响可谓无处不在,既为科研、创新和经济赋能,又带来新的挑战与风险。我们应以开放的心态看待人工智能带来的诸多改变,审慎研究和应对其可能带来的新课题与新风险。

据《瞭望》周刊