



华安证券

HUAAN SECURITIES

证券研究报告

# AIGC行业：大模型改变开发及交互环境，处于高速迭代创新周期

分析师：金荣（S0010521080002）

2024年2月6日

华安证券研究所

# 投资建议

- GPTs上线2个月后用户便创建了超过300万个应用，GPT提供了基于大模型的原生开发环境，用户可通过更加自然的语言交互方式，减少对传统GUI的依赖，同时多模态的趋势仍在继续，大模型仍处于高速创新迭代周期，头部公司对2024年的Gen AI的资本开支保持积极，建议关注AI算力及大模型布局厂商 NVIDIA, AMD, Intel, IBM, Microsoft, Alphabet, Amazon, Arista, 百度集团-SW, 阿里巴巴-SW, 腾讯控股等。
- 上游算力侧：
 

受AI产业驱动影响，算力需求持续扩张，AI大模型的持续优化及多样化AI应用终端的入市商用持续提升全球算力需求，推动AI基础设施建设，在AI基建需求出现井喷的情况下，全球主流AI芯片厂商最新季度业绩均超市场预期，建议关注算力板块机会。1) 台积电：2023Q4 营业收入 196.7 亿美元，基本符合预期。调整后净利润 75.1 亿美元，超一致预期 5.4%。AI 方面，公司预计未来 AI 收入的 CAGR 有望达到 50%。2) AMD：23Q3营收58亿美元，yoy+4%，qoq+21%，超过彭博一致预期57亿美元。1月30日，公司对AI芯片业务收入上调指引75%，预计2024年AMD人工智能芯片的销售额将达到35亿美元以上（此前指引为20亿美元）。3) Meta：1月19日，扎克伯格宣布Meta内部正在训练下一代模型Llama 3，而到2024年底Meta将拥有35万块H100，拥有近60万个GPU等效算力，这些算力很大部分将用于训练Llama 3大模型。4) NVIDIA：FY24Q3，公司收入同比上升206%至181.2亿美元，大超此前160亿美元的指引，四季度指引收入200±2%亿美元，这也是由于大语言模型带动的训练和推理算力需求较为强劲，各大数据中心均加大了对英伟达AI芯片（H100、A100等）和HGX平台的采购力度导致的。建议关注：台积电、AMD、Nvidia等。
- 23年以来大模型重点发展方向：
 

1) NLP领域：23年随着GPT4、GPT4 Turbo、Gemini的陆续推出，我们可以看出NLP领域大模型的发展趋势体现在两方面：①模型参数量呈现增长趋势；②由单模态向多模态演变。其中，参数量规模的增长在很大程度上推动了大模型的学习能力（但不绝对，也取决于模型的算法和参数质量，过多的参数也会导致过拟合情况发生），而大模型由单模态向多模态的演变，也意味着模型能够处理和理解来自不同数据模态的信息，这些模态可以包括但不限于图像、文本、音频、视频等，实现复杂生成任务。建议关注在相关领域持续探索AI算法的公司，关注阿里巴巴-SW、京东集团-SW、焦点科技、值得买、网易-S、腾讯控股、吉比特、三七互娱、巨人网络、天娱数科、科大讯飞、昆仑万维等。

## 投资建议

**2) CV领域：**通过复盘CV领域主流应用（Midjourney、Dall·E和Stable Diffusion）的版本迭代，其发展方向除了生成内容的质量优化外（如清晰度、生成时长），也向多模态、更具可控性（指模型对Prompt理解能力的跃升等）。在技术升级方面，23年3D GS推出，相较于原始的NeRF方法，3D GS在实现高质量渲染的同时可以做到实时渲染。此外，2023年初ControlNet 横空出世，与传统的生成对抗网络不同，ControlNet 允许用户对生成的图像进行精细的控制，给出了一条完美解决 AI 绘画发展进程中“精准控制”难题的道路。我们认为新的底层技术为其产品的可用性赋能，会使得用户使用学习成本更低，吸引更多B或者C的用户，并提升用户的留存及复购，建议关注布局大模型的阿里巴巴-SW、腾讯控股、百度集团-SW、美图公司、昆仑万维、科大讯飞等。此外，CV领域的发展会提高电影、动画片制作、游戏建模的生产效率，从而生产更多优质内容，建议关注兼具IP的上海电影、光线传媒、华策影视、奥飞娱乐等，以及游戏相关板块的腾讯控股、网易-S、三七互娱、盛天网络等。

**3) 应用侧：**海内外AI应用相继落地，2024年1月份GPTStore的上线催生出了新的商业模式，AI大模型应用将迎来新一轮热潮，建议关注海内外AI应用的落地情况。GPTStore上线后，OpenAI把GPT-4 vision（视觉）、Code interpreter（代码）、DALLE-3（图像）、TTS（语音）的API进行开放，并且新推出了Assistants API。所以用户除了可以使用除了目前已有的300多万个应用之外，ChatGPT Plus用户也可以自己创建新的GPT上架到GPTStore中供其他用户付费使用。此外，微软也于2023年11月1日正式上线包含Copilot集成的Microsoft 365应用程序，为企业客户提供商用服务，24年1月15日，微软宣布 Copilot Pro正式面向个人，用户可以 Edge浏览器和 Microsoft Office 365上使用Copilot Pro。国内方面，以豆包、天工、文心一言等为代表的AI应用用户规模呈现持续增长势头，建议关注具有AI应用布局的公司：Microsoft、阿里巴巴-SW、腾讯控股、科大讯飞、美图公司等。

**4) 硬件侧：**当前手机和PC受AI驱动影响，预计将迎来新一轮换机潮，主流手机厂商已经陆续自研大模型，部分旗舰款手机已经侧端搭载大模型发售，以联想、戴尔、华硕为代表的主流PC厂商也将在今年陆续发售自家旗舰AIPC，建议关注AI手机和AIPC产品的落地情况，建议关注小米集团-W、联想集团、戴尔科技等。

- 风险提示：人工智能相关政策风险；AI模型等技术进步不及预期；用户增长不及预期等。

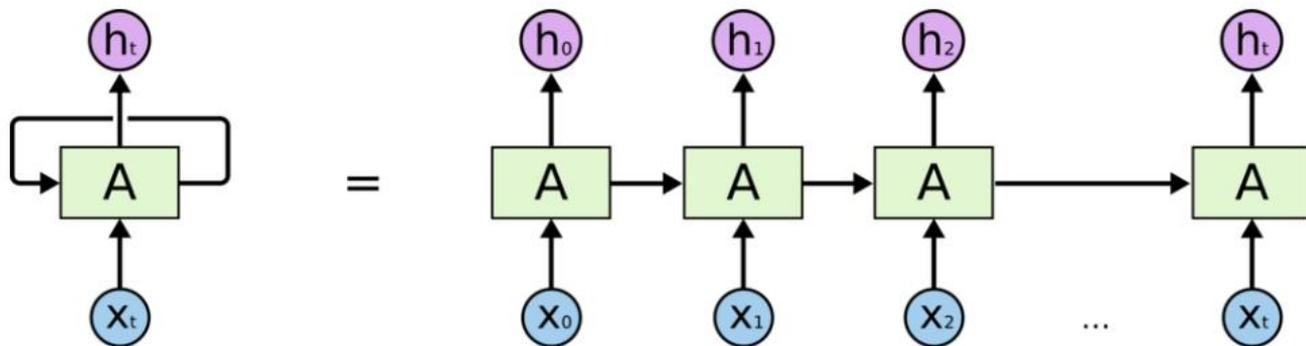
# 目录

1. LLM发展历程与变迁趋势
  - 1.1 LLM爆发关键节点：2017年Transformer架构的出现
  - 1.2 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代
  - 1.3 国内主流玩家大模型布局：互联网大厂大模型相继落地，主要赋能自身旗下产品
2. CV领域发展历程与变迁趋势
  - 2.1 CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向
  - 2.2 23年以来CV领域技术升级：3D渲染能力增强，生成内容更为精准可控
  - 2.3 CV领域主流应用迭代历程：生成内容更为精细，Prompt理解能力进一步深入
3. AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况
4. 海内外热门AI应用盘点：聊天类机器人仍是主流，海内外主流应用用户规模持续扩大
  - 4.1 国内热门AI应用：豆包、天工、文心一言
  - 4.2 海外热门AI应用：ChatGPT、Bard、Poe、Notion、Bing、Character.ai

# LLM大模型爆发的关键节点：2017年“Transformer”模型的出现

- 在“Transformer”模型还未出现时，NLP（自然语言处理）领域的主流架构基本采用RNN（循环神经网络），RNN的痛点可以总结为两点：
  - ①由于递归性质，训练过程中通常无法并行计算；  
RNN在工作过程中将会对内容按顺序逐字处理，每一步的输出取决于先前的隐藏状态和当前的输入，需要等到上一个步骤完成后才能进行当前计算，因此无法进行并行计算，训练效率较低。
  - ②不擅长处理长序列、长文本；  
当RNN处理长序列时，由于信息的不断累积，时间步长增加时，梯度会不断变小，这会导致梯度消失或梯度爆炸问题，使得网络难以学习长期依赖（long-term dependency）的关系。当梯度消失时，前面的信息无法有效地传递到后面，也就导致词之间距离越远，前面对后面的影响越弱，所以RNN难以有效的捕获长距离的语义关系，当梯度爆炸时，网络的权重会变得极大，导致网络不稳定。
- 虽然之后也出现了如LSTM（Long Short-Term Memory）和GRU（Gated Recurrent Unit）等的RNN变体，但解决长期依赖问题的能力也有限，也依旧无法解决无法并行计算的痛点。

图表1：RNN（循环神经网络）架构示意图



资料来源：CNDS，华安证券研究所整理  
注：X<sub>t</sub>为输入值，h<sub>t</sub>为输出值

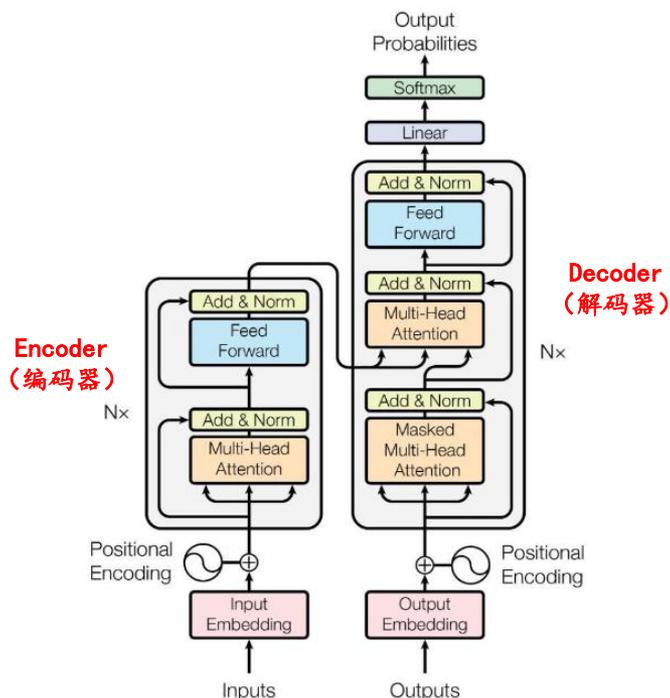
注：长期依赖（long-term dependency）：在一个序列中，后面的某个位置的输出，可能会依赖于前面的若干个位置的输入，这种依赖跨越了很长的时间间隔。

# LLM大模型爆发的关键节点：2017年“Transformer”模型的出现

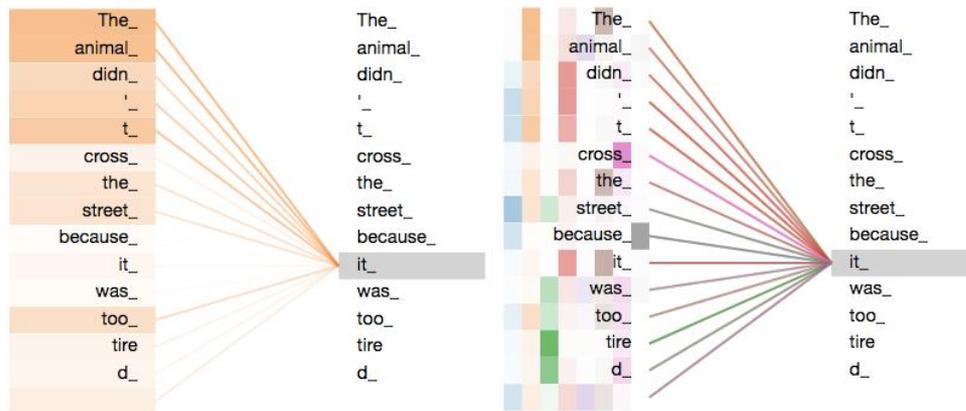
- “Transformer”是谷歌的研究团队在2017年6月在论文《Attention is All You Need》中首次发布的。该论文提出了一种全新的序列转换架构，完全基于自注意力机制，摒弃了循环神经网络（RNN）在处理序列数据时的局限性。“Transformer”有能力学习输入序列里所有词的相关性，不会受到短时记忆的影响。
- 我们可以将“Transformer”的创新总结为两点：①自注意力机制；②位置编码。其中，“Transformer”的自注意力机制简单来说就是“Transformer”在处理每个词的时候，不仅会关注这个词本身以及附近的词，还会去注意输入序列里所有其他的词，然后给予每个词不一样的注意力权重（权重是模型通过大量文本训练习得的）。

- 例：“The animal didn't cross the street because it was too tired”，其中“it”可以代指任何东西，“Transformer”中的注意力机制可以关注到输入序列里所有词，并给予不同权重，而像这样的自注意力机制有很多个，每个自注意力机制都有自己的注意力权重，用来关注文本里不同特征，每个权重互不影响，也就意味着可以实现并行计算。

图表2：Transformer架构示意图



图表3：Transformer多头自注意力机制（Multi-Head Attention）

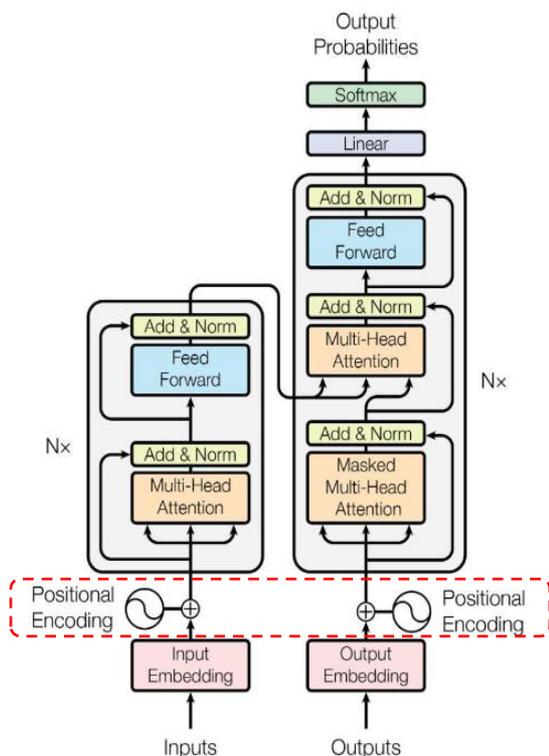


资料来源：《Attention is all your need》，华安证券研究所整理 资料来源：CNDS，华安证券研究所整理

# LLM大模型爆发的关键节点：2017年“Transformer”模型的出现

- Transformer模型抛弃了RNN、CNN作为序列学习的基本模型，循环神经网络本身就是一种顺序结构，天生就包含了词在序列中的位置信息。当抛弃循环神经网络结构，完全采用Attention取而代之，这些词序信息就会丢失，模型就没有办法知道每个词在句子中的相对和绝对的位置信息。因此，有必要把词序信号加到词向量上帮助模型学习这些信息，位置编码（Positional Encoding）就是用来解决这种问题的方法。

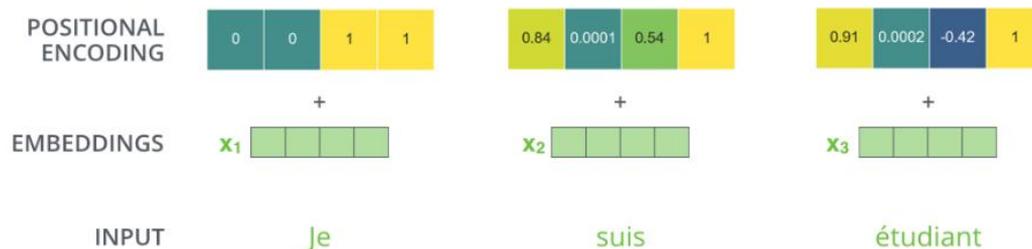
图表4：Transformer架构示意图



资料来源：《Attention is all your need》，华安证券研究所整理

- 例：当接收到输入内容后，“Transformer”与其他NLP模型一样，会将词转化为词向量（这里举例维度为4个，但在Transformer中词向量维度为512个，每个维度都代表不同特征），此外还会根据每个词所在文中的不同位置，给予位置向量。这样就可以实现让模型既理解每个词的意义，又能够捕获每个词在文中的位置，从而理解不同词之间的顺序关系。借助位置编码，词可以不按顺序输入给神经网络，模型可以同时处理输入序列里的所有位置，不需要像RNN那样依次处理。

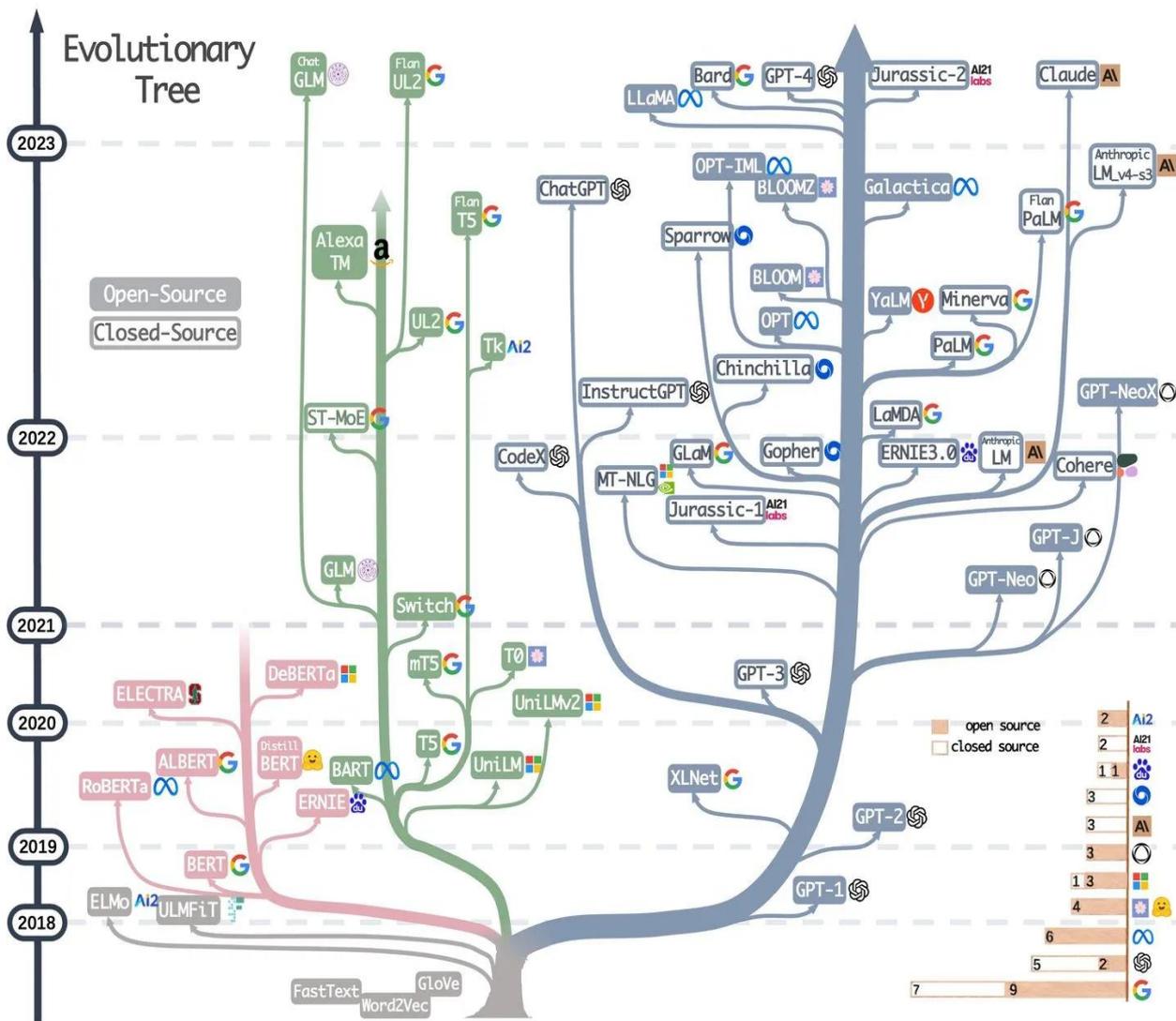
图表5：位置编码（Positional Encoding）示意图



资料来源：CNDS，华安证券研究所整理

# 基于Transformer架构的LLM发展图谱：Decoder-only成为主流

图表6：基于Transformer架构的LLM发展图谱



①粉色枝：**Encoder-Only**（仅用解码器）。代表LLM为BERT，该类模型主要用于处理输入数据，专注于理解和编码信息，而不是生成新的文本。

②绿色枝：**Encoder-Decoder**（编码器+解码器）。代表LLM为T5，该类模型擅长处理需要理解输入然后生成相关输出的任务，比如翻译或问答系统。

③蓝色枝：**Decoder-Only**（仅解码器）。代表LLM为GPT系、LLaMA、Bard等。该类模型擅长创造性的写作，比如写小说或自动生成文章。它更多关注于从已有的信息（开头）扩展出新的内容。这种架构能够更好地理解和预测语言模式，尤其适合处理开放式的、生成性的任务。

\*注：实心方框为开源模型，空心方框为闭源模型

# 目录

1. LLM发展历程与变迁趋势
  - 1.1 LLM爆发关键节点：2017年Transformer架构的出现
  - 1.2 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代
  - 1.3 国内主流玩家大模型布局：互联网大厂大模型相继落地，主要赋能自身旗下产品
2. CV领域发展历程与变迁趋势
  - 2.1 CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向
  - 2.2 23年以来CV领域技术升级：3D渲染能力增强，生成内容更为精准可控
  - 2.3 CV领域主流应用迭代历程：生成内容更为精细，Prompt理解能力进一步深入
3. AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况
4. 海内外热门AI应用盘点：聊天类机器人仍是主流，海内外主流应用用户规模持续扩大
  - 4.1 国内热门AI应用：豆包、天工、文心一言
  - 4.2 海外热门AI应用：ChatGPT、Bard、Poe、Notion、Bing、Character.ai

# 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代

图表7：海外公司大模型布局

模型公司	大模型分类	版本	技术能力
微软/OpenAI	GPT	GPT-1	2018年6月，GPT-1发布，参数量达到1.17亿个，语言泛化能力不足，无法解决通用的语言任务；
		GPT-2	2019年2月，GPT-2发布，参数量达到15亿个，舍弃模型微调，通过大模型进行预训练，开始能够解决多语言任务能力；
		GPT-3	2020年5月，GPT-3发布，参数量达到1750亿个，引入In-context训练方式学习上下文，翻译、聊天问答和文本填空等任务表现出色，能够生成难辨真假新闻的能力；
		InstructGPT	2022年3月，InstructGPT发布，参数量达到5400亿个；
		ChatGPT	2022年11月，推出对话式AI模型ChatGPT，由GPT-3.5大型语言模型微调而来，并引入人工反馈的强化学习（RLHF）。能够处理复杂语言工作，包括自动文本生成、自动问答、编写和调试计算机程序及创作故事；
		GPT-4	2023年3月，GPT-4发布，参数量为1.8万亿，能够阅读文字和识别图像并生成文本结果，回复的准确性较历史版本有提升；
		GPT-4 Turbo和GPTs	2023年11月，GPT-4 Turbo和GPTs发布，GPT-4 Turbo上下文理解能力强化，控制权更灵活，价格更具竞争力，GPTs可以通过适当指令构建GPT；
Google	Gemini	Ultra	谷歌最强大的模型，在广泛的高度复杂的任务中提供最先进的性能，包括推理和多模态任务，可以在TPU加速器上有效地提供服务，能够完成高度复杂的任务，主要面向数据中心和企业级应用。
		Pro	在成本和延迟方面的性能优化模型，在广泛的任务中提供显著的性能。该模型具有较强的推理性能和广泛的多模态能力。
		Nano	Gemini Nano是三个版本中最小、最高效的模型，本专为移动设备和特定任务设计，这使得它成为在智能手机和其他便携设备上实现先进AI功能的理想选择。其中，Nano-1的参数为1.8B，Nano-2为3.25B，分别针对低内存和高内存设备。
	BERT	BERT-base	Google在2018年推出的一种革命性的预训练语言模型。它基于Transformer架构，摒弃了传统的循环神经网络（RNN）或卷积神经网络（CNN），使得模型并行计算成为可能，大大提升了训练效率。其中参数方面，BERT-base具有12层Transformer块（encoder层），每层有12个自注意力头，参数量大约在110M左右，BERT-large参数量大约是340M左右。
		BERT-large	
	T5	T5	一种统一文本到文本转换框架的预训练模型，由Google在2020年提出。该模型的设计理念是将所有自然语言处理任务都转化为文本到文本的任务，即无论是机器翻译、问答、摘要生成还是文本分类等任务，输入和输出都是文本形式。T5家族包含多个版本，参数量从60M到11B不等
	LaMDA	LaMDA 1.0	是一款专门针对对话应用设计的大型预训练语言模型，于2021年推出。该模型是Google在自然语言处理领域的一项重要突破，旨在通过深度学习技术提供更为流畅、连贯和有意义的对话体验，具备开放式对话能力。参数量方面，LaMDA 1.0参数量未知，LaMDA 2.0参数量为1370亿个。
		LaMDA 2.0	
	MUM		模型于2021年5月在Google I/O大会上首次被推出，MUM通过训练来完成多种复杂的NLP任务，例如问题解答、文档摘要生成、语义相似度计算等，与BERT等前一代模型相比，MUM更擅长理解复杂查询的上下文，包括用户的潜在需求和意图。
	PaLM	PaLM	最初公开的PaLM模型版本之一，其参数量达到5400亿，2022年4月前发布
PaLM-2		在2023谷歌I/O大会上，谷歌CEO皮查伊宣布推出对标GPT-4的大模型PaLM 2，并正式发布预览版本，改进了数学、代码、推理、多语言翻译和自然语言生成能力，训练参数量达3400亿，分为S\ M\L三个规模的版本。	
PaLM-E		2023年3月6日，谷歌和德国柏林工业大学的人工智能研究人员公布了PaLM-E，PaLM-E由两部分组成，PaLM（5400亿）+ViT-22B（220亿），也就是540B的语言参数+22B的图像参数。PaLM-E整合了视觉信息处理能力，能够理解图像并结合文本进行生成和推理，成为当时已知的最大规模的“通才”AI模型。	

# 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代

图表7：海外公司大模型布局（续表）

模型公司	大模型分类	版本	技术能力
Meta	LLaMA	LLaMA1	是Meta AI公司于2023年2月发布的大型语言模型。它训练了各种模型，这些模型的参数从70亿到650亿不等。
		LLaMA2	2023年7月，Meta推出了LLaMA2，Meta 训练并发布了三种模型大小的 Llama 2: 70、130 和 700 亿个参数，模型架构与 Llama 1 模型基本保持不变
		Code LLaMA	2023年8月，Meta继发布用于生成文本、翻译语言和创建音频的人工智能模型之后，开源了 Code Llama。这是一个机器学习系统，可以用自然语言（特别是英语）生成和解释代码。可以免费商用和研究。
	OPT-175B	OPT-175B	是由Meta AI于2022年发布的一个大型预训练语言模型，拥有约1750亿个参数。作为一款开源模型，OPT-175B在规模上与GPT-3相当，但它的主要特点是完全免费开放给研究社区使用。OPT-175B通过大规模文本数据集进行训练，旨在提供强大的自然语言处理能力，包括但不限于文本生成、问答系统、语义理解和对话交互等方面。
	SeamlessM4T	SeamlessM4T-Large v2	是Meta于2023年发布的一款先进的人工智能模型，专门针对多语言翻译和转录任务设计。该模型能够支持近100种不同语言之间的文本到文本、文本到语音以及语音到语音的翻译，并且在语音识别和合成方面表现出强大的能力。参数方面，Large v1和v2均为2.3B，medium版本为1.2B。
		SeamlessM4T-Large (v1)	
SeamlessM4T-Medium (v1)			
Amazon	Amazon Titan	Titan Text Express	可用于检索增强生成、开放式文本生成、头脑风暴、摘要、代码生成、表格创建、数据格式化、释义、思维链、重写、提取、问答和聊天。支持100+语言，最高输出8k tokens
		Titan Text Lite	适合文本生成任务和微调，仅支持英语，最高输出4k tokens。
		Titan Text Embeddings	适用于文本检索、语义相似性和聚类，支持25+种语言，最高输出8k tokens。
		Titan Multimodal Embeddings	可提供准确的多模式搜索，仅支持英语，最高输出128tokens，最大输出图片大小为25MB。
	Titan Image Generator	适用于文本到图像生成、图像编辑、图像变体，最高输出77tokens，最大输入图片大小为25MB	
Olympus		暂无细节披露	
Anthropic	Claude	Claude初代版本	包括Claude1.0、claude-instant-1、 claude-1.2、 claude-1.3等
		Claude2.0	全球公测开始于2023年12月24日前后，该版本增强了AI模型，并支持文档上传识别功能
		Claude2.1	发布日期为2023年11月，这个版本支持输入多达20万个Token，并且改进了准确度，同时增加了新的系统预设功能或其他增强特性。
	Inflection	Inflection 1.0	发布于2023年6月22日，在大规模多任务语言理解（MMLU）中，优于Meta的LLAMA、OpenAI的GPT 3.5和谷歌的PaLM（540B）。
Inflection 2.0		发布于2023年11月22日，在 MMLU上评估中，Inflection-2 是 GPT-4 之外性能最高的模型，甚至思维链推理优于 Claude 2。	
xAI	Grok	Grok 1.0	于 2023 年 11 月发布，上下文长度为 8,192 个token，用于自然语言处理任务，包括问答、信息检索、创意写作和编码辅助。

## 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代

图8：海外大模型主要应用

模型名称	模型类型	应用软件	应用上线时间	软件所属公司	公司是否上市	软件所属公司融资情况	应用描述	商业模式	战略合作关系
GPT 3	NLP	Copy.ai	2020-10-1	CopyAI Inc	否	种子轮：CraftVentures领投，融资数额290万美金 A轮：WingVentureCapital领投，红杉、Tiger、个人天使投资人EladGil跟投，融资数额1100万美金	文本生成类：用于自动化文案生成，包括广告文案、博客文章、社交媒体帖子、产品描述	免费（有收费Pro版）	groove、zoom、hubspot、confluent、seismic、datadog
		Jasper AI	2021-1-1	Jasper	否	2022年10月：融资额1.31亿美金，估值15亿美元	文本生成类：用于自动化生成高质量的营销和广告文案、博客文章、产品描述、电子邮件、社交媒体帖子等	收费，基础版\$29.99每月	intel、klarna、zoom、Deloitte、ibm、Google、airbnb、digital marketer
		Notion.AI	2022-11-1	Notion	否	天使轮：2013年融资200万美金 A轮：2019年7月融资1800万美金 B轮：2020年4月融资5000万美金 C轮：2021年10月融资2.5亿美元，投后估值103亿美元	文本生成类：能帮助用户快速生成和修改文本内容，可运用在文学、会议日程、新闻稿、销售文案等内容的辅助性创作中	免费（有收费Pro版）	airbnb、nike、Amazon、uber、toyota、ibm等
		Replica	2016-02	Luka Inc	否	A1轮：2016年4月获442万美元融资 A2轮：2017年11月获650万美元融资	文本生成类：可以模拟人类对话，对用户的输入提供真实的反应，可用于心理健康咨询、语言学习、娱乐等。	免费（有收费Pro版）	
		Elai	2021	Elai.io	否	2022年获5万美元融资	视频生成类：只需要文本就能制作出有真人主持的专业视频，可用于教育、营销、企业沟通、带货等用途	收费，基础版\$29每月	partner: NVIDIA、Amazon、Microsoft、thinkific、Google for startup
		Inworld: AI	2022-04	Inworld: AI	否	种子轮：21年11月融资700万美元（投资方包括Meta）； Pre-A轮：22年3月融资1000万美元； A轮：22年8月获得Intel Capital领投的5000万美元	3D生成类：文生NPC，可创建角色设定场景、设定世界观常识等	免费，（有收费基础版\$10/月、专业版、定制版）	Disney
DALL·E	多模态	Craiyon	2021-06	Boris Dayma (个人)	否		图像生成类：只需要在搜索框中输入相应场景的文字描述，就能生成各种或唯美、搞笑又或者画面惊悚的图片。	免费（有收费Pro版）	
CLIP	多模态	starryai	2021	o Kahn(个人)	否		图像生成类：能生成各种风格和主题的艺术作品，可为客户提供所创建图像的完全所有权，可用于个人或商业用途。	免费（有收费会员和积分制度）	
GPT-4	多模态	GitHub Copilot X	2023-03	GitHub	否	A轮：2012年由Andreessen Horowitz领投，融资1亿美元 B轮：2015年由Sequoia Capital领投，融资2.5亿美元	代码生成类：具备对话、文本生成代码、语音生成代码、自动修复代码Bug，以及解释代码等功能	收费，个人版\$10每月	微软旗下的公司，与Open AI合作开发
GPT-4 Turbo	多模态	Microsoft Copilot	2023-12	Microsoft	是	2013年，微软曾发行了80亿美元的债券； 2015年，微软进行了一次大举发债融资，金额高达107.5亿美元； 2016年，为了筹集收购领英的资金，微软发债融资达到197.5亿美元。	人工智能辅助工具：具备代码编写、文档创作、办公协同、搜索与问答、自定义助手构建、	部分用户免费，Microsoft 365 Copilot定\$30/月	微软旗下的公司，与Open AI合作开发

## 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代

模型所属公司	模型名称	模型类型	应用软件	应用上线时间	软件所属公司	公司是否上市	软件所属公司融资情况	应用描述	商业模式	战略合作关系
DeepMind	BigGAN	CV	Artbreeder	2019-05	Studio Morphoge (工作室)	否		图像生成类：用于创作人物肖像、人物形象、动漫角色、建筑、画作、自然景观等 还包括改变面部特征的能力	免费（有收费Pro版）	
英伟达	StyleGAN	CV								
Google	ViT	CV	Murf	2020-10	Murf	否	种子轮：Elevation Capital领投，获150万美元融资 A轮：Matrix Partners领投，获1000万美元融资	音频生成类：支持文本转语音、添加视频 旁白、幻灯片旁白、语音克隆，可用于 Youtubers、播客、有声读物、游戏、产品和解说视频中	免费（有收费基础版、专业版、企业版）	
	基于 transformer	NLP	Mem	2021-06	Mem Labs	否	A轮：22年11月获得OpenAI领投的2350万美元	辅助决策类：AI驱动的办公记事应用，可以帮助用户快速记录和整理个人信息数据 并利用AI将这些数据连接成有意义的知识图谱	免费（有收费个人版\$8/月，企业版\$15/月、定制版）	Open AI为其投资方；为Twitter打造专属功能Mem it
	Gemini Pro	多模态	Bard	2023-3-21	Google	是	N. A.	人工智能技术的聊天软件：多角度回答问题、智能搜索集成、交互体验优化、支持多语言交流、持续学习与提升、应用场景广泛	免费（计划推出付费版本）	
			NotebookLM	2023-12-09						
	Gemini Nano	多模态	AICore	2023				系统级模块：允许开发者更加便捷地集成和利用AI模型在应用程序中，以实现诸如语音识别、自然语言处理、图像识别等智能化功能。	为Google Pixel 8 Pro提供支持	
	LaMDA 2.0	NLP	AI Test Kitchen	2023-08				实验性应用程序：允许用户在安全、可控的环境中测试和体验谷歌正在研发中的自然语言处理模型，如LaMDA（对话应用模型）及其后续版本。		
	PaLM	NLP	Bard	2023-3-21				人工智能技术的聊天软件：多角度回答问题、智能搜索集成、交互体验优化、支持多语言交流、持续学习与提升、应用场景广泛	免费（计划推出付费版本）	
	PaLM-2	多模态	Bard	2023-3-21				人工智能技术的聊天软件：多角度回答问题、智能搜索集成、交互体验优化、支持多语言交流、持续学习与提升、应用场景广泛	免费（计划推出付费版本）	
			Gmail	2004				网络邮件服务：支持实时邮件同步、多账户管理、智能回复、自定义通知设置等诸多功能	免费	
	PaLM-E	多模态	Google Search	1996				搜索服务：通过这款应用轻松查找互联网上的各种信息，包括网页、图片、视频、新闻和本地商家信息等。	日常搜索服务免费	
Google Assistant			2017	智能语音助手：提供语音交互、多语言支持、智能家居控制、日程管理、信息获取、娱乐功能、第三方应用整合、情境感知服务				免费		
Google Docs			2005	在线文档处理软件：允许用户创建、编辑和共享多种类型的文档并支持实时协作功能				基本功能免费、企业版\$15/月		
Meta	LLaMA2	NLP	Azure	2010	Microsoft	是	N. A.	云计算服务平台：为用户提供了广泛的云服务选项，涵盖了计算、存储、网络、数据库、人工智能、物联网（IoT）、数据分析等多个领域	收费，其收费情况根据所使用的服务类型、资源量、使用时长等因素而定	

## 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代

模型所属公司	模型名称	模型类型	应用软件	应用上线时间	软件所属公司	公司是否上市	软件所属公司融资情况	应用描述	商业模式	战略合作关系
Amazon	Amazon Titan	多模态	Amazon Bedrock	2023-09	AWS	否	N. A.	全托管服务：允许用户和开发者访问多种领先的AI模型	根据使用量、模型类型、计算资源消耗以及API调用次数等因素而定	
	Inflection	NLP								
Stability.ai	Stable Diffusion (基于Latent Diffusion Models)	多模态	Stability.ai	2022-08	Stability.ai	否	种子轮：22年10月，由Coatue 和 Lightspeed领投，融资额1.01 亿美元，估值10亿美元	图像生成类：根据所需图像的文本描述，生成符合匹配该描述的高质量图像。	免费（有收费Pro版）	partner: Krikey.ai、亚马逊、Geenee
Midjourney	Midjourney V5 (基于Diffusion Models)	多模态	Midjourney	2022-03	Midjourney	否	N. A.	图像生成类：根据所需图像的文本描述，生成相对应的图片，可以选择不同画家的艺术风格例如安迪华荷、达芬奇、达利 和毕加索等，还能识别特定镜头或摄影术语。	收费，基础版\$10每月	partner: MirrorX、Discard、NFT
加州大学伯克利分校、谷歌、加州大学圣地亚哥分校	NeRF	多模态	Luma	2021-09	Luma AI	否	种子轮：2021年10月，获得South Park Commons、Matrix Partners等投资，数额为430万美元；A轮：23年3月，由Amplify Partners领投，NVIDIA (Nventures) 等跟投，融资数额为2000万美元	3D生成类：基于神经渲染技术，可以拍摄少量照片来生成、着色和渲染产品的逼真3D模型。	按次收费，每个视频1美元，容量限制5GB	Amplify Partners、General Catalyst 投资
Runway	gen-2	多模态	Runway AI	2023-03	Runway	否	B轮：21年12月融资3500万美元；C轮：22年12月融资5000万美元	视频生成类：允许用户从零开始生成视频 可以文生视频、文本+参考图像生视频、静态图片转视频	gen-1免费，gen-2处于内测阶段，尚未开放使用	参与开发Stable Diffusion模型

资料来源：各公司官网，华安证券研究所整理

敬请参阅末页重要声明及评级说明

# 大模型发展趋势：参数规模扩大+向多模态演进

- 通过回溯主流模型迭代情况，我们可以将LLM大模型的发展趋势总结为两点：①**模型参数量呈现增长趋势**；②**由单模态向多模态演变**。其中，参数量规模的增长在很大程度上推动了大模型的学习能力（但不绝对，也取决于模型的算法和参数质量，过多的参数也会导致过拟合情况发生），而大模型由单模态向多模态的演变，也意味着模型能够处理和理解来自不同数据模态的信息，这些模态可以包括但不限于图像、文本、音频、视频等，实现复杂生成任务。

图表10: ChatGPT4可以实现图像输入识别



资料来源：公司官网，华安证券研究所整理

图表9: 主流大模型发展趋势

模型名称	公司	参数量 (亿)	发布日期	模态
Gemini-Ultra	Google	未知	2023	多模态
GPT-4	OpenAI	18000	2023	多模态
PaLM-E	Google	5620	2023	多模态
GPT-3	OpenAI	1750	2020	NLP
T5	Google	11	2020	NLP
GPT-2	OpenAI	15	2019	NLP
BERT	Google	3.4	2018	NLP

资料来源：各公司官网，Wikipedia，华安证券研究所整理

# 目录

## 1. LLM发展历程与变迁趋势

- 1.1 LLM爆发关键节点：2017年Transformer架构的出现
- 1.2 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代
- 1.3 国内主流玩家大模型布局：互联网大厂大模型相继落地，主要赋能自身旗下产品

## 2. CV领域发展历程与变迁趋势

- 2.1 CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向
- 2.2 23年以来CV领域技术升级：3D渲染能力增强，生成内容更为精准可控
- 2.3 CV领域主流应用迭代历程：生成内容更为精细，Prompt理解能力进一步深入

## 3. AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况

## 4. 海内外热门AI应用盘点：聊天类机器人仍是主流，海内外主流应用用户规模持续扩大

- 4.1 国内热门AI应用：豆包、天工、文心一言
- 4.2 海外热门AI应用：ChatGPT、Bard、Poe、Notion、Bing、Character.ai

## AI政策整理

图表11：近年AI政策整理

年份	政策名称	颁布单位	政策内容
2016.05	《“互联网+”人工智能三年行动实施方案》	国家发展改革委等	对人工智能芯片发展方向提出多项要求，并促进智能终端可穿戴设备的推广落地。
2017.07	《新一代人工智能发展规划》	国务院	提出前瞻布局新一代人工智能重大科技项目，在“核高基”（核心电子器件、高端通用芯片、基础软件）、集成电路装备等国家科技重大专项中支持人工智能软硬件发展。对人工智能领域进行了总体战略部署，规划确定了人工智能产业三步走的发展目标。
2019.03	《关于促进人工智能和实体经济深度融合的指导意见》	科技部等	把握新一代人工智能的发展特点，结合不同行业，不同区域特点，探索创新成果应用转化的路径和方法，构建数据驱动、人机协同、跨界融合的智能经济形态。
2021.03	《“十四五”规划和2035远景目标纲要》	国务院	打造数字经济新优势，加强关键数字技术创新应用。加快推进高端芯片、操作系统、人工智能关键算法、传感器、通用处理器等领域研发突破和迭代应用。
2022.08	《关于支持建设新一代人工智能示范应用场景的通知》	科技部	充分发挥人工智能赋能经济社会发展的作用，围绕构建全链条、全过程的人工智能行业应用生态，支持一批基础较好的人工智能应用场景，加强研发上下游配合与新技术集成，打造形成一批可复制、可推广的标杆型示范应用场景。
2023.02	《数字中国建设整体布局规划》	国务院	系统优化算力基础设施布局，促进东西部算力高效互补和协同联动，引导通用数据中心、超算中心、智能计算中心、边缘数据中心等合理梯次布局。
2023.04	《生成式人工智能服务管理办法（征求意见稿）》	国家互联网信息办公室	首次明确了生成式人工智能“提供者”内容生产、数据保护、隐私安全等方面的法定责任及法律依据，确立了人工智能产品的安全评估规定及管理办法。
2023.04	《关于推进IPv6技术演进和应用创新发展的实施意见》	工信部等	推动IPv6和5G、人工智能、云计算等技术的融合创新，支持企业加快应用感知网络、新型IPv6测量等“IPv6+”创新技术在各类网络环境和业务场景中的应用。
2023.07	《生成式人工智能服务管理暂行办法》	国家互联网信息办公室等	国家坚持发展和安全并重、促进创新和依法治理相结合的原则，采取有效措施鼓励生成式人工智能创新发展，对生成式人工智能服务实行包容审慎和分类分级监管。
2023.09	《关于实施专精特新中小企业就业创业扬帆计划的通知》	工业和信息化部等	按照国家有关规定，动态调整职称专业设置，根据当地产业发展和专精特新中小企业需要，增设人工智能、大数据、工业互联网等新专业。
2023.12	《关于加快推进视听电子产业高质量发展的指导意见》	工业和信息化部等	支持骨干企业做大做强，支持人工智能企业研发视听应用大模型。
2023.12	《“数据要素x”三年行动计划（2024—2026年）（征求意见稿）》	国家数据局	以科学数据支持大模型开发，建设高质量语料库和基础科学数据集，支持开展通用人工智能大模型和垂直领域人工智能大模型训练。

资料来源：国务院，发改委，科技部，工信部，华安证券研究所整理  
 敬请参阅末页重要声明及评级说明

## 国内主流玩家大模型布局一览

图表12：国内主流玩家大模型布局一览（含llm+cv）

模型公司	模型类型	大模型分类	技术能力	应用场景	落地产品（软件）
百度		文心大模型	2023年5月，百度推出文心大模型3.5，基于飞桨深度学习平台和文心知识增强大模型，持续从海量数据和大规模知识中融合学习，具备知识增强、检索增强和对话增强的技术特色。 2023年10月，百度推出文心大模型4.0，百度当时最新且最强大的版本，实现了基础模型的全面升级，在理解、生成、逻辑推理及多模态处理等方面有显著提升。	应用于用于企业服务、自动驾驶、文化创意等多个产业领域。	百度搜索、百度旗下各类APP等。
	NLP	文心·NLP大模型	业界首个开放的千亿参数中文生成API	智能创作、摘要生成、问答、语义检索、情感分析、信息抽取、文本匹配、文本纠错等各类自然语言理解和生成任务。	文心一格、文心百中（网页&小程序）
	CV	文心·CV大模型	文心VIMER-UFO2.0是文心旗下目前最先进的统一任务大模型，主要应用场景为智慧城市。整体具备包含170亿参数，多个任务的数据训练一个功能强大的通用模型。	可被直接应用于处理人脸、人体、车辆、商品、食物细粒度分类等20+CV基础任务。	
	多模态	文心·跨模态大模型	基于知识增强的跨模态语义理解关键技术，实现跨模态检索、图文生成、图片文档的信息抽取等应用	图像生成、艺术制作、虚拟现实、AI辅助设计等	
			文心·生物计算大模型	HelixGEM是业界首个基于几何构象增强的化合物表征模型，引入化合物的三维空间信息，在14个药物属性预测相关的benchmarks上效果达到业界最优； HelixFold-Single，秒级别的蛋白结构预测模型，是业界首个开源的基于单序列语言模型的蛋白结构预测大模型； HelixFold利用大规模无标注数据强化蛋白质表示能力，构建的全国产化软硬件适配的高效蛋白质结构分析大模型。	小分子研发
		文心·行业大模型	与各行业头部企业合作，在通用大模型的基础上学习特色数据与知识，建设行业AI基础设施	能源、金融、航天、传媒、城市、影视、制造、社科	
腾讯	NLP	混元大模型	2023年9月推出，混元大模型是腾讯自主研发的一款超大规模预训练语言模型，拥有超过千亿参数规模和强大的自然语言处理能力。该模型在预训练过程中学习了海量文本数据，从而具备了跨领域的知识理解、逻辑推理以及创造性生成等多种能力。	智能客服与问答系统、内容创作与编辑、教育辅助、搜索引擎优化、游戏交互、	腾讯云服务、腾讯广告、腾讯游戏、腾讯会议、腾讯文档等多个内部产品
		腾讯医疗大模型	2023年9月推出，该模型专注于医疗健康领域，通过深度学习技术，提供智能问答、医疗文档理解和辅助决策等功能。	医疗咨询、患者服务、医生辅助工具、医疗知识图谱构建	腾讯健康的智能医疗解决方案中，如智能家庭医生助手、数智医疗影像平台等。
字节跳动		ChitChop	2023年11月推出，是字节跳动在海外推出的一款基于大模型的人工智能助理工具。它涵盖了AI创作、AI绘画、娱乐、AI学习、工作和生活六大场景，能够为用户提供多方面的智能化服务。	内容创作、图像生成、在线学习、办公协助以及日常生活中的多种任务处理。	ChitChop作为独立的App在Google Play和App Store上线。
	多模态	PixeILM	PixeILM是由字节跳动智能创作团队联合北京交通大学、北京科技大学研究人员推出的高效像素级推理大模型。该模型可以实现无需依赖特定加速模块（如SAM）的像素级语言模型推理。	适用于各种需要结合文本与图像理解及生成的任务，例如图像描述、基于文本指令生成图像或编辑图像等。	
	LLM	云雀大模型	2023年8月推出，旨在提升AI在理解、生成和交互等方面的综合性能，其特点是能够广泛应用于各种自然语言处理任务中，如文本生成、问答系统、语义分析等，并具有较高的泛化能力和适应性。	内容创作与编辑、智能客服与在线助手、教育辅导与知识问答、搜索优化与信息检索、社交媒体和社区互动	豆包、抖音、今日头条等APP
	LLM	BuboGPT	2023年8月推出，能够将文本、图像和音频等多模态输入进行整合，并具有将回复与视觉对象进行对接的独特能力。	对话系统、内容生成、智能客服、图像描述、声音定位	

## 国内主流玩家大模型布局一览

模型公司	模型类型	大模型分类	技术能力	应用场景	落地产品（软件）
阿里巴巴	多模态	通义千问	2023年4月推出，超大规模的语言模型，功能包括多轮对话、文案创作、逻辑推理、多模态理解、多语言支持。	智能客服、内容生成、知识问答、多模态应用、内部工具升级	淘宝问问、阿里云平台服务、通义千问APP、天猫、钉钉、高德地图、优酷、盒马等
	—	Qwen模型	2023年8月推出Qwen-7B和Qwen-7B-Chat，Qwen-7B是一款通用型的70亿参数规模的大语言模型，主要用于自然语言理解和生成任务。Qwen-7B-Chat是基于Qwen-7B开发的对话模型； 2023年8月推出Qwen-VL，一款大规模视觉语言模型，基于通义千问70亿参数模型Qwen-7B研发的，支持图文输入，具备多模态信息理解能力； 2023年9月推出Qwen-14B与Qwen-14B-Chat，Qwen-14B是一个具有140亿参数的大规模预训练模型，相应的Qwen-14B-Chat则是一个更强大的对话系统模型； 2023年11月推出Qwen-Audio，一款音频语言模型，能够处理和理解人类语音、自然声音等多种音频类型，并应用于超过30种不同的任务场景； 2023年12月推出，具有720亿参数规模的大规模语言模型，基于大量的高质量数据进行训练，旨在提供卓越的语言理解和生成能力；	自然语言处理服务、对话式AI助手、多模态内容生成、智能搜索优化、企业级解决方案、AI创作辅助、机器人操控与导航、视觉搜索与推荐、图像问答	
	NLP	通义-AliceMind	达摩院开源的深度语言模型体系，涵盖预训练模型、多语言预训练模型、超大规模中文预训练模型； 2023年8月推出Alice-Mind模型2.0，该模型基于飞桨深度学习平台，在多项技术上进行了革新，并且取得了显著的提升。	阅读理解、机器翻译、对话问答、文档处理	
	多模态	通义-M6	2021年11月推出的多模态大模型，参数由万亿级跃升至10万亿，成为当时全球最大的AI预训练模型。使用521张GPU在10天训练具有可用水平的10万亿模型，相比GPT-3、M6同等参数规模，能耗仅为其1%。	模型实现架构、模态和任务统一，便于在M6基础上精调得到丰富的小模型	
	CV	通义-视觉	基于英伟达A10 GPU、阿里云CIPU芯片以及1aaS+软件加速，实现新一代的GPU云服务，在AI计算和视觉计算能力获得大幅提升	图像搜索与匹配、视频分析、AI辅助设计、医疗影像诊断、AI绘画、	
	—	行业模型层	平台解决方案服务商以通义大模型为底座，开发SaaS应用提供增值服务；传媒和游戏运用AIGC实现体验升级	电子商务、AI医疗、智能农业、自动驾驶、城市大脑、工业视觉、AIOT	
	NLP	夸克大模型	阿里巴巴在2023年11月推出全栈自研、千亿级参数的夸克大模型。	主要应用在通用搜索、医疗健康、教育学习、职场办公等众多场景，以提供更智能的搜索和推荐功能。	夸克App
华为	NLP	盘古NLP大模型	2023年7月推出盘古大模型3.0，其中盘古NLP大模型专注于自然语言处理任务，如文本理解、生成、问答、对话系统构建等。	智能舆论、智能营销	
	CV	盘古CV大模型	2023年7月推出盘古大模型3.0，其中盘古CV大模型能够处理图像识别、目标检测、视频分析等多种视觉相关的AI任务。	智能巡检、智慧物流	
	多模态	盘古多模态大模型	2023年7月推出盘古大模型3.0，其中盘古多模态大模型可以同时处理文本、语音、图像等多种输入数据，实现跨模态的信息交互与理解，例如结合文字描述生成图像或者通过语音指令执行任务。	智能内容生成与编辑、跨媒体检索与推荐、医疗影像诊断辅助、智能制造与质量检测、VR/AR、智能家居、智慧城市管理、教育科技	
	科学计算	盘古科学计算大模型	2023年7月推出盘古大模型3.0，其中盘古科学计算大模型用于物理模拟、化学反应预测、生物数据分析等复杂问题的求解。	气象预报、台风路径预测	
	—	行业定制化子模型	盘古大模型3.0强调面向行业的应用，因此会有一系列经过针对性训练或微调的模型，适用于金融、医疗、教育、制造等行业特定场景下的智能决策支持、业务流程优化、客户服务自动化等任务。	金融风控、医疗诊断、教育个性化辅导、智能制造、零售业客户关系管理、法律文书审查、新闻传媒内容生成	

## 国内主流玩家大模型布局一览

模型公司	模型类型	大模型分类	技术能力	应用场景	落地产品（软件）
昆仑万维	多模态	天工巧绘   SkyPaint	文本生成图片模型，即提示词文本编码器模型和扩散模型领大部分； 2023年4月推出天工大模型3.5；2023年7月推出天工大模型4.0；	产生高质量图片	
	多模态	天工乐府   SkyMusic	商高级作曲AI模型；2023年4月推出天工大模型3.5；2023年7月推出天工大模型4.0；	音乐和娱乐	
	NLP	天工妙笔   SkyText	采用了生成效果领先的开源GPT中文预训练大模型，针对中文领域构建了千亿级别的高质量数据集；2023年4月推出天工大模型3.5；2023年7月推出天工大模型4.0；	文字创作	
	NLP	天工智码   SkyCode	多语言开源编程大模型，支持多种主流代码语言。它可以生成包括Python、Java、C等使用广泛的编程语言，以及PHP、Go、Swift等共计十余种编程语言；2023年4月推出天工大模型3.5；2023年7月推出天工大模型4.0；	编程	
	NLP	SkyChat	基于中文GPT3 API研发的项目，可以像ChatGPT一样，实现人机聊天、问答、中英文互译、对对联和写古诗任务	聊天机器人	
科大讯飞	——	讯飞星火认知大模型	2023年5月推出，一款大规模预训练语言模型，它基于深度学习技术，拥有强大的跨领域知识理解和生成能力； 2023年6月发布了讯飞星火认知大模型V1.5，在原有基础上实现了显著的能力升级； 2023年8月推出讯飞星火认知大模型V2.0，在多个核心能力上取得了显著提升，包括但不限于文本生成能力提高了72%，语言理解性能提升了78%，知识问答方面的能力也有了大幅提升。 2023年10月推出讯飞星火认知大模型V3.0，在前代基础上进一步扩大了参数规模和优化了算法架构，加强了启发式对话和AI人设功能。	学习教育、办公商务、医疗健康、工业制造、智能助手	AI学习机LUMIE 10系列、讯飞智能办公本、智能录音笔、智能麦克风、星火助手中心与星火APP
	NLP	讯飞星火医疗大模型	2023年10月推出，结合了自然语言处理、深度学习和大数据分析能力，致力于提高医疗服务效率、减轻医生负担，并为患者提供便捷的健康管理服务。	智能预问诊与导诊、辅助诊断与决策支持、医疗文档理解和生成、AI智医助理、AI心理伙伴、科研助手、教育应用	
智谱	——	ChatGLM系列大模型	ChatGLM：智谱AI推出的首款基座大模型，专为对话场景设计。 ChatGLM2：继ChatGLM之后的迭代版本，进一步提升了性能和适用范围。 ChatGLM3：智谱AI于2023年10月在中国计算机大会上发布的第三代基座大模型，采用了多阶段增强预训练方法，并且特别强调了其可部署在手机等移动设备上的特性，以及对国产芯片的支持。	智能客服系统、在线教育辅助工具、内容创作平台、社交聊天机器人、游戏及娱乐行业、企业决策支持、移动应用部署	

资料来源：各公司官网，华安证券研究所整理

# 目录

1. LLM发展历程与变迁趋势
  - 1.1 LLM爆发关键节点：2017年Transformer架构的出现
  - 1.2 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代
  - 1.3 国内主流玩家大模型布局：互联网大厂大模型相继落地，主要赋能自身旗下产品
2. CV领域发展历程与变迁趋势
  - 2.1 CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向
  - 2.2 23年以来CV领域技术升级：3D渲染能力增强，生成内容更为精准可控
  - 2.3 CV领域主流应用迭代历程：生成内容更为精细，Prompt理解能力进一步深入
3. AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况
4. 海内外热门AI应用盘点：聊天类机器人仍是主流，海内外主流应用用户规模持续扩大
  - 4.1 国内热门AI应用：豆包、天工、文心一言
  - 4.2 海外热门AI应用：ChatGPT、Bard、Poe、Notion、Bing、Character.ai

## CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向

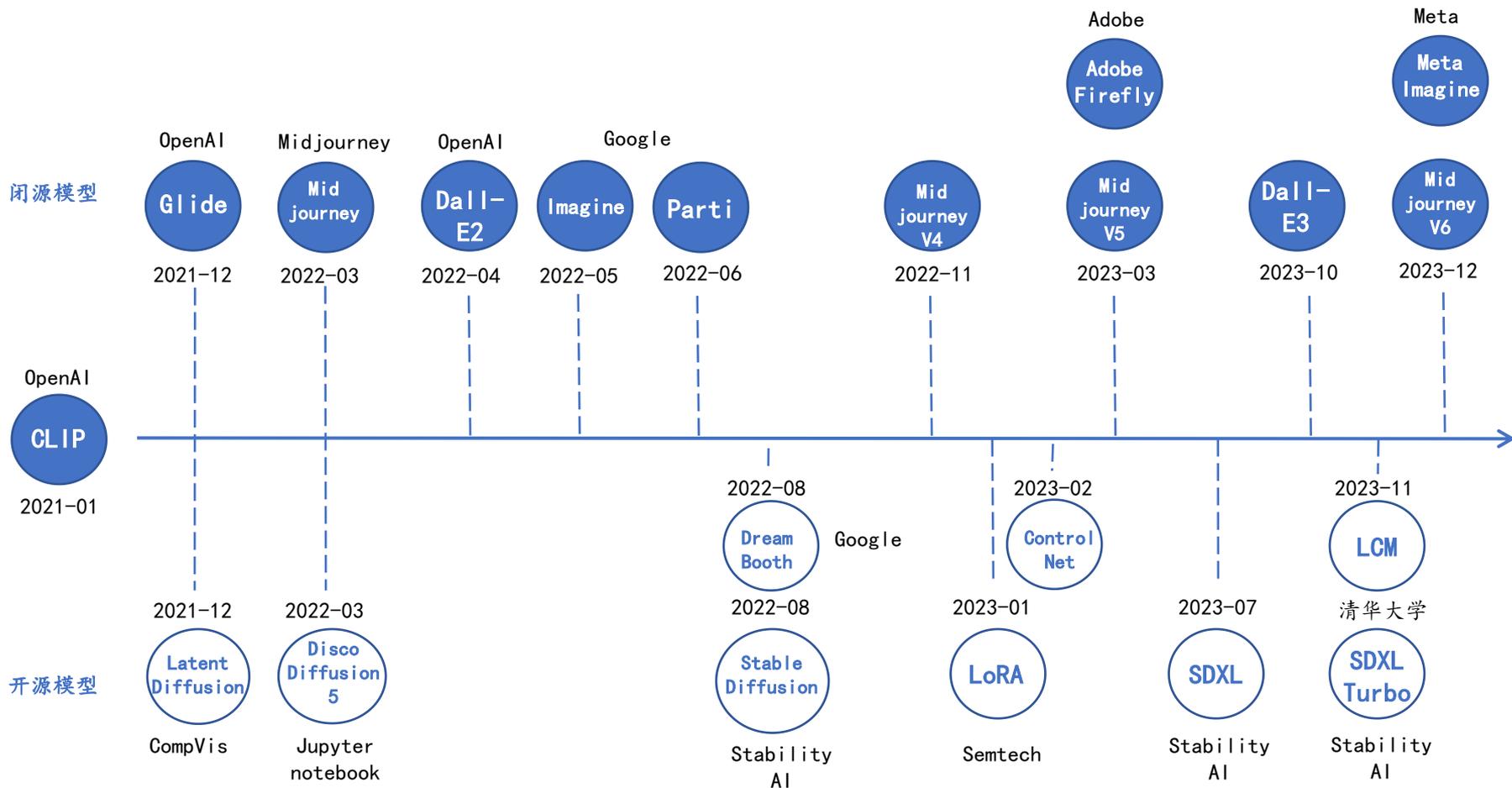
- **GAN**：早期的AI绘画技术，AI绘画工具的起源。GAN模型一问世就风靡AI学术界，在多个领域得到了广泛的应用。它也随即成为了很多AI绘画模型的基础框架，其中内置的生成器用来生成图片，判别器用来判断图片质量。GAN的出现大大推动了AI绘画的发展。GAN模型进行AI绘画有比较明显的缺陷，一方面是对输出结果的控制力很弱，容易产生随机图像，另外一个问题是生成图像的分辨率比较低。
- **Disco Diffusion**：Disco Diffusion 是第一个基于CLIP + Diffusion 模型的实用化AI绘画产品。它利用生成对抗网络（GAN）和扩散模型技术来创建艺术作品，尤其是从文本描述生成图像。2022年初在谷歌Colab云服务上正式对世界开放使用。Disco Diffusion的痛点也有两个：1) AI绘画细节还不够深入,达不到商业水准；2) 渲染时间过长，需要大量内存。
- **DALL·E2**：DALL-E是一个可以通过文本描述中生成图像的人工智能程序，于2021年1月5日由OpenAI发表。2022年4月，OpenAI宣布了新版本的DALL-E 2，它可以从文本描述中生成照片般逼真的图像，另外还有一个允许对输出进行简单修改的编辑器。
- **Midjourney**：2022年7月12日，MidJourney进入公开测试阶段，用户可以通过Discord的机器人指令来使用MidJourney生成图像。MidJourney是一个基于大规模预训练的自然语言处理模型，可以根据文本生成图像。同月，MidJourney开始向所有用户开放注册，引起了广泛的关注和讨论。MidJourney展示了其对各种复杂和超现实的图像生成任务的能力，并被认为是与DALL-E相媲美的人工智能模型。
- **Stable Diffusion**：2022年7月Stable Diffusion的问世，相比前辈们，Stable Diffusion已经成功的解决了细节及效率问题，通过算法迭代将AI绘图精细度提升到了艺术品级别，并将生产效率提升到了秒级，创作所需的设备门槛也被拉到了民用水准。由于拥有发展良好的开源社区，它的用户关注度和应用广度都超越Midjourney和DALL-E。
- **多模态**：AI绘画逐渐成为大模型核心模态之一，如GPT-4V/DALL-E3。

图表13：CV领域关键工具变迁



# CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向

图表14：CV领域模型重要发展节点



资料来源：华安证券研究所绘制

敬请参阅末页重要声明及评级说明

## CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向

- **CV领域文生图模型的起源于CLIP开源。**CLIP 是一个通过自然语言监督有效地学习视觉概念的神经网络，通过使用 CLIP 可将文本和图像连接在一起。
- **2021年：**21 年 1 月 CLIP 发布后（与 CLIP 一起发布的还有 DALL·E 第一代文生图模型，生成能力有限），基于它的各种文生图模型相继出现，Disco Diffusion 也是在这个时候诞生，此时它们的生成效果都并不理想。21 年 12 月，OpenAI 发布了 Glide，同月，Compvis 实验室发布了 Latent Diffusion 的论文，成为生成能力实现重大突破基础。基于 Glide 的论文和各种探索，T2I 开源社区开启了 Disco Diffusion 项目。
- **2022年：**22 年 3 月 Disco Diffusion v5 发布，这是图像生成模型第一次出圈。同一个月，Midjourney 通过包装使用 Disco Diffusion，进行了商业应用。22 年 4 月，OpenAI 发布了 DALL·E 2，为了与 OpenAI 竞争，Google在22 年 5 月和 6 月发布了 Imagen 和 Parti 的论文，但一直没有公开发布过产品。22 年 8 月发布的 Stable Diffusion 1.5，由 Stability AI 与慕尼黑大学 Compvis 实验室合作训练完成，加之后来的 LoRA 和 ControlNet，构建了庞大的 SD 生态。22 年 11 月 Midjourney v4 发布，它的生成能力和艺术感惊人，使其一举奠定了MJ 作为商业闭源模型无法撼动的王者地位。
- **2023年以来：**23 年1 月，LoRA 诞生，同一个基底 SD 模型可以搭载不同的 LoRA 使用，具有很高的灵活性，成为了目前最流行的微调 SD 的方法之一。23年2 月，ControlNet发布。它的出现代表着 T2I 生成开始真正变得可控，能够成为真正的生产力工具。ControlNet 比之前 img2img 提供了更丰富更准确的控制方式，可以直接提取 input 画面的边缘、深度、语义分隔、深度信息，以及人物的姿势等。3 月 Adobe 发布了 Firefly，将生成式 AI 集成到自己的产品 Photoshop Beta 以及面向普通人的 Express 中。7 月，Stability 开源了 SDXL1.0，这是 Stability.AI 最新一代旗舰模型，成为当之无愧的最强开源文生图模型。

# 目录

1. LLM发展历程与变迁趋势
  - 1.1 LLM爆发关键节点：2017年Transformer架构的出现
  - 1.2 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代
  - 1.3 国内主流玩家大模型布局：互联网大厂大模型相继落地，主要赋能自身旗下产品
2. CV领域发展历程与变迁趋势
  - 2.1 CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向
  - 2.2 23年以来CV领域技术升级：3D渲染能力增强，生成内容更为精准可控
  - 2.3 CV领域主流应用迭代历程：生成内容更为精细，Prompt理解能力进一步深入
3. AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况
4. 海内外热门AI应用盘点：聊天类机器人仍是主流，海内外主流应用用户规模持续扩大
  - 4.1 国内热门AI应用：豆包、天工、文心一言
  - 4.2 海外热门AI应用：ChatGPT、Bard、Poe、Notion、Bing、Character.ai

## 23年以来CV领域技术升级--ControlNet: AI绘画的精准手术刀

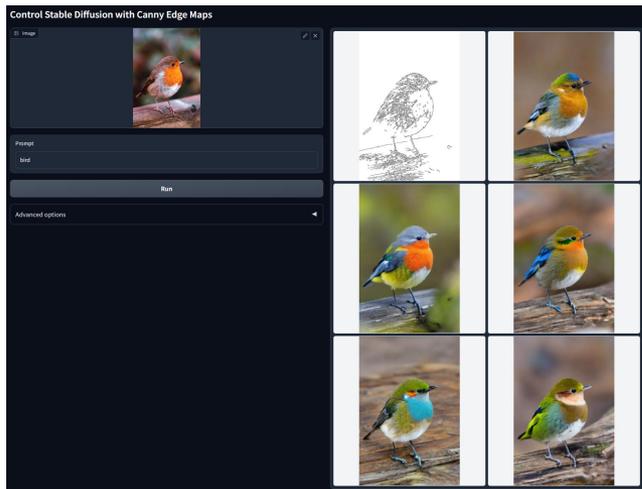
- ControlNet 是一种用于控制Stable Diffusion 模型的神经网络模型，ControlNet 的作用是通过不同的图像采集方式，并添加额外控制条件，来引导 Stable Diffusion 按照创作者的创作思路生成图像，从而提升 AI 图像生成的可控性和精度。2023年初ControlNet 的横空出世，与传统的生成对抗网络不同，ControlNet 允许用户对生成的图像进行精细的控制，给出了一条完美解决 AI 绘画发展进程中“精准控制”难题的道路。

图表15: ControlNet 主要模型及功能

预处理名称	对应模型	模型描述
canny	control_canny	通过 ControlNet 边缘检测模型或线稿模型提取线稿（可提取参考图片线稿，或者手绘线稿），再根据提示词和风格模型对图像进行着色和风格化。
depth	control_depth	深度检测，捕捉画面深度，获取图片前后景关系，图片颜色越浅的区域，代表距离镜头越近；图片越是偏黑色，则代表这部分区域距离镜头越远。
hed	control_hed	边缘检测但保留更多细节，适合重新着色和风格化。与Canny的提取边缘类似，但可以把Canny理解为用铅笔提取边缘，而Hed算法则是用毛笔，被提取的边缘将会非常柔和。
mlsd	control_mlsd	线段识别，非常适合建筑，提取建筑的线条结构和几何形状，构建出建筑线框（可提取参考图线条，或者手绘线条），再配合提示词和建筑/室内设计风格模型来生成图像。
normal_map	control_normal	根据图片生成法线贴图，非常适合CG建模师。normal贴图被称之为法线贴图，一般是用来保存模型上的深度信息。normal贴图将会控制模型表面的阴影深度，以此来达到凹凸不平的感觉。
openpose	control_openpose	提取人物骨骼姿势，精准识别出人物姿态，再配合提示词和风格模型生成同样姿态的图片，同时还可以手动调整人物姿势
scribble	control_openpose	涂鸦成图，可以手画涂鸦线稿，提取黑白稿来生成图像
fake_scribble	control_scribble	涂鸦风格提取，使用简单的算法从输入图像合成涂鸦
segmentation	control_seg	语义分割模型，标注画面中的不同区块颜色和结构（不同颜色代表不同类型对象），从而控制画面的构图和内容。

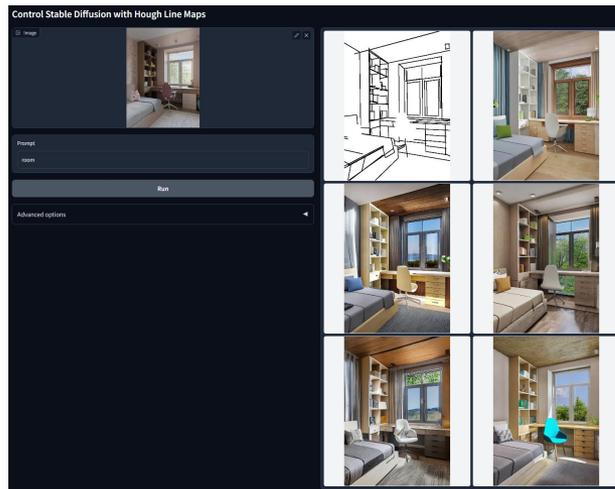
# 23年以来CV领域技术升级--ControlNet: AI绘画的精准手术刀

图表16: canny模型成图示例



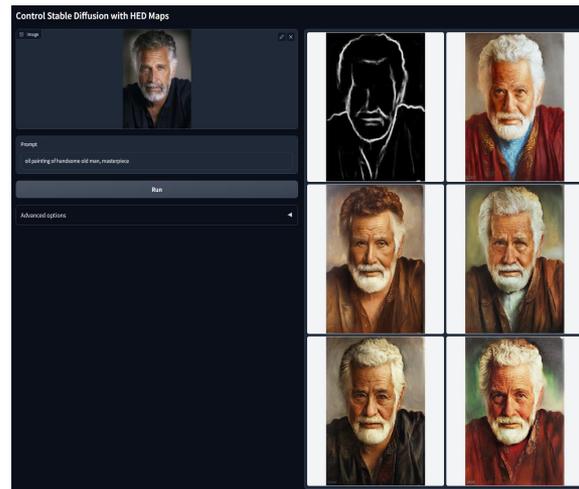
资料来源: Github, 华安证券研究所

图表17: MLSD模型成图示例



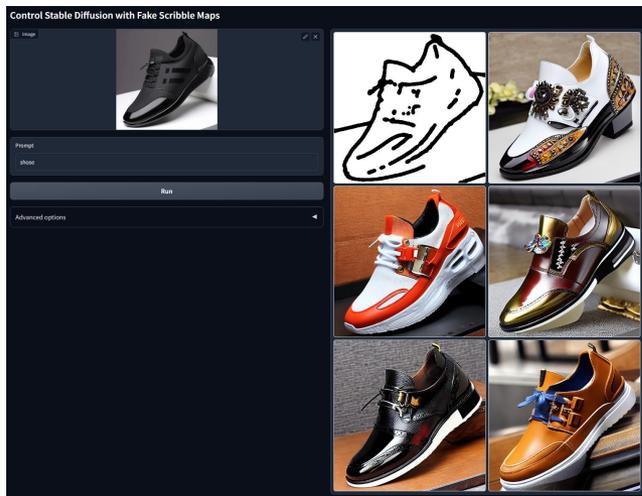
资料来源: Github, 华安证券研究所

图表18: HED模型成图示例



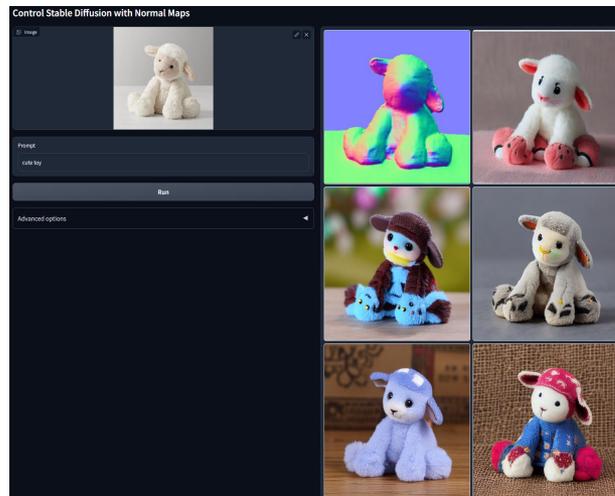
资料来源: Github, 华安证券研究所

图表19: scribble模型成图示例



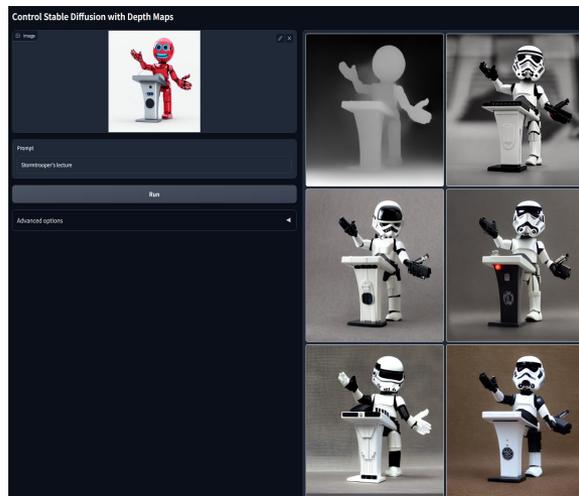
资料来源: Github, 华安证券研究所  
敬请参阅末页重要声明及评级说明

图表20: Normal maps模型成图示例



资料来源: Github, 华安证券研究所

图表21: Dept模型成图示例

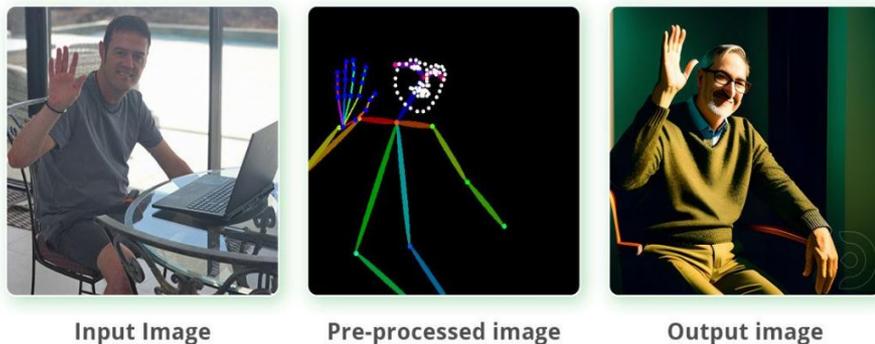


资料来源: Github, 华安证券研究所

## 23年以来CV领域技术升级--ControlNet: AI绘画的精准手术刀

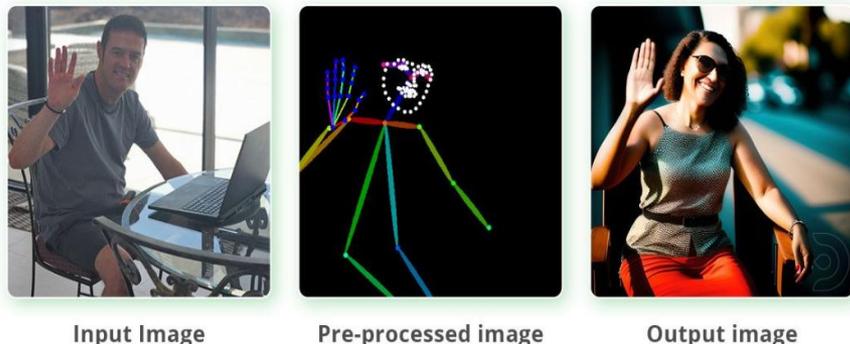
- ControlNet最为值得关注的就是其openpose模型，openpose模型的推出使得用户摆脱了仅靠提示词去控制角色姿势的弊端，openpose可以提取人物骨骼姿势进而生成图像，在图像采集过程中，用户还可以选择多种模式，如：仅提取脸（openpose\_faceonly）、提取手（openpose\_hand）、提取身体+脸（openpose\_face）、提取身体+手+脸（openpose\_full）等。
- 除此之外，如果用户不满意人物姿势，还可以就手动调整骨骼位置，通过鼠标按住骨骼的关节点拖动至想要的位置松手，这样即可实现骨骼位置匹配的问题。

图表22: openpose\_full模型成图示例



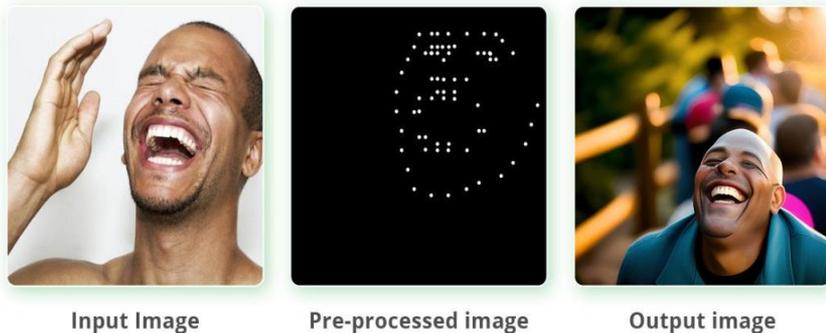
资料来源: think diffusion, 华安证券研究所

图表23: openpose\_face模型成图示例



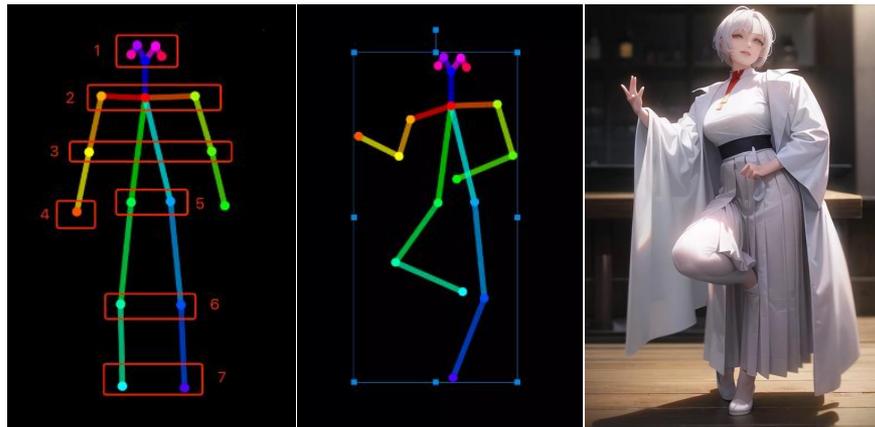
资料来源: think diffusion, 华安证券研究所

图表24: openpose\_faceonly模型成图示例



资料来源: think diffusion, 华安证券研究所

图表25: 手动调整人物骨骼姿势示例



资料来源: 虎课网, 华安证券研究所

## 23年以来CV领域技术升级--3D GS：当前三维重建的最优解

- 3D gaussian splatting是2023年SIGGRAPH的最佳论文《3D Gaussian Splatting for Real-Time Radiance Field Rendering》提出的技术，主要用于实时三维场景的高效表示与渲染。相较于原始的NeRF（神经辐射场，同为三维重建和渲染领域的技术），3D gaussian splatting的优势在于：
  - ①实时性更强。**3D Gaussian Splatting 结合了传统光栅化技术和基于体的表示，通过优化数据结构和计算流程，能够在一定程度上实现比 NeRF 更快的渲染速度，从而有利于实时或近实时应用。
  - ②存储效率与计算效率更高。**gaussian splatting通过对高维空间进行稀疏采样并利用局部高斯分布来减少不必要的计算量。相比之下，NeRF通常需要大量的网络查询才能生成新的视角图像，尤其是在处理大场景时，可能会遇到较高的计算复杂度。
  - ③细节表现更好。**3D Gaussian Splatting 更擅长于保留某些特定类型的细节，如毛绒玩具表面的绒毛特性，而原始的NeRF在处理此类微小几何细节时效果不那么理想。
- 例：《3D Gaussian Splatting for Real-Time Radiance Field Rendering》中对比了用InstantNGP、Plenoxels、Mip-NeRF360和3D gaussian splatting（图中指的是“Ours”）三维重建的图像，“Ground Truth”可以理解为原始图像。3D gaussian splatting在时间和fps（帧数）上具备明显优势，训练时长仅6分钟，fps达135，同时PSNR（峰值信噪比，一种图像质量的评价指标，可以评价两幅图像之间的相似程度）高于InstantNGP和Plenoxels。在训练51分钟后，PSNR可达到25.2，高于训练时长为48小时的Mip-NeRF360所生成的图像。

图表26：几种不同三维重建和渲染技术生成图像对比

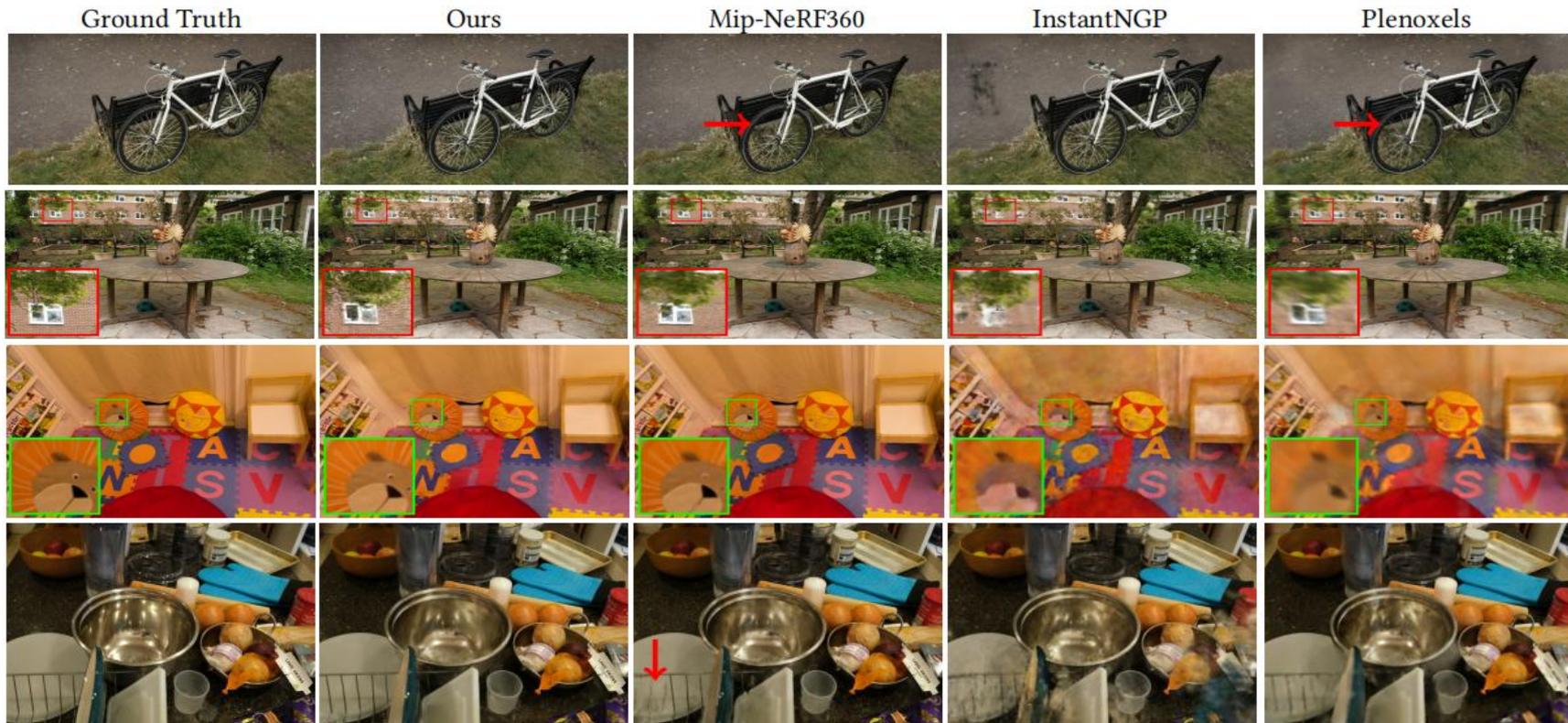


资料来源：《3D Gaussian Splatting for Real-Time Radiance Field Rendering》，华安证券研究所

## 23年以来CV领域技术升级--3D GS: 当前三维重建的最优解

- 在细节处理和虚影处理上，3D gaussian splatting都明显优于InstantNGP、Plenoxels和Mip-NeRF360，重建场景与Ground Truth更为贴合。
- 注：
  - ①Mip-NeRF360**: Mip-NeRF360 是对 NeRF的改进和扩展，由Google于2022年推出，该模型在NeRF基础上可渲染出完整的3D物体和场景，支持从任意视角观看，相当于渲染出完整的3D虚拟空间。mip-NeRF 360特点是可渲染场景中每一点的360° 视角，可合成逼真的视图，以及高度复杂、细节的大场景深度图。
  - ②InstantNGP**: 由Nvidia于2022年推出，相较于NeRF，InstantNGP大幅增加了三维场景渲染速度，并能在有限计算资源下实现实时交互和高质量渲染。
  - ③Plenoxels**: 2021年提出，通过梯度方法和正则化优化校准图像，既能保持同NeRF一样的渲染质量，又能减少两个数量级的优化时间。

图表27: 当前主流三维重建模型对比



资料来源: 资料来源: 《3D Gaussian Splatting for Real-Time Radiance Field Rendering》, 华安证券研究所  
 敬请参阅末页重要声明及评级说明

# 目录

1. LLM发展历程与变迁趋势
  - 1.1 LLM爆发关键节点：2017年Transformer架构的出现
  - 1.2 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代
  - 1.3 国内主流玩家大模型布局：互联网大厂大模型相继落地，主要赋能自身旗下产品
2. CV领域发展历程与变迁趋势
  - 2.1 CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向
  - 2.2 23年以来CV领域技术升级：3D渲染能力增强，生成内容更为精准可控
  - 2.3 CV领域主流应用迭代历程：生成内容更为精细，Prompt理解能力进一步深入
3. AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况
4. 海内外热门AI应用盘点：聊天类机器人仍是主流，海内外主流应用用户规模持续扩大
  - 4.1 国内热门AI应用：豆包、天工、文心一言
  - 4.2 海外热门AI应用：ChatGPT、Bard、Poe、Notion、Bing、Character.ai

## Midjourney迭代之路：V6版本具备更为深入的复杂自然语言的理解力

### • Midjourney版本迭代节点：

• V1 (2022/02) → V2 (2022/04) → V3 (2022/07) → V4 (2022/11) → V5 (2023/03) → V5.1 (2023/05) → V5.2 (2023/06) → V6 (2023/12)

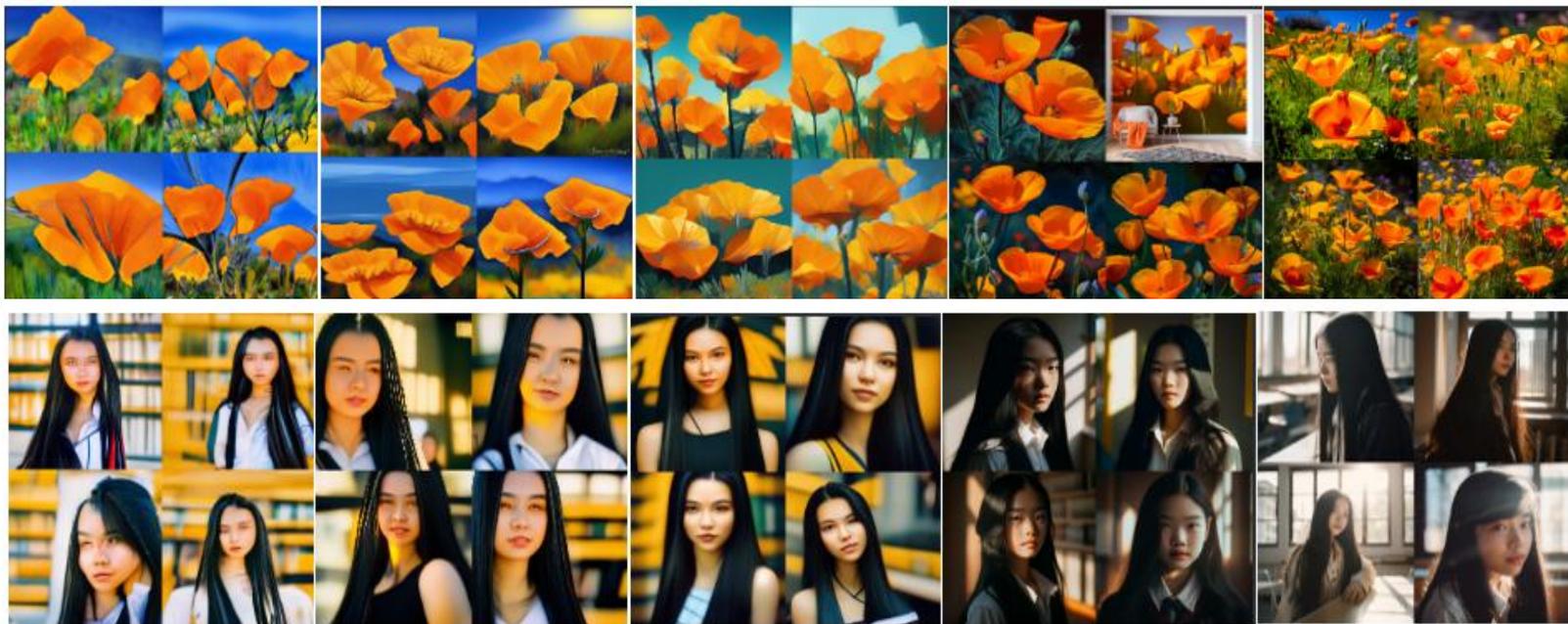
### • Midjourney版本迭代的升级更新方向：

①越来越丰富的细节，生成的内容也越来越真实。V1 和 V2 呈现画面粗糙，之后版本愈发真实。

②分辨率提升。V1~V3 单张图像素是 256 x 256，到了 V5 单图已经默认是 1024\*1024。

③Prompt理解能力更深。用户可以通过多长文本描述更加细致的对生成画面提出要求，到V6版本，提示词上限为350个词，不仅可以用自然语言描述画面，而且可以控制画面的布局且支持包含多个主体。

图表28：同一prompt生成的图片（从左到右分别为V1-V5）



资料来源：Learning prompt，华安证券研究所整理

# Midjourney迭代之路：V6版本具备更为深入的复杂自然语言的理解力

历经9个月的研发，12月21日Midjourney V6 Alpha版本正式上线。

我们认为Midjourney自V5之后版本（v5.1、v5.2、v6），以及包括当前图像生成主流应用之间（如Midjourney、Dall·E3、Stable Diffusion），从生成图像的质感方面看不存在太大差异，更多的是风格的不同。如何降低用户学习成本，让用户更为简单高效并精准的使用是下一步迭代的关键。Midjourney V6 Alpha版本就拥有了更为深入的复杂自然语言的理解力，大幅降低了用户的学习成本。

图表29：根据复杂prompt生成精准图片



*prompt*: 厨房的桌子上有三个装满水果的篮子。中间的篮子里装着青苹果。左边的篮子里装满了草莓。右边的篮子里装满了蓝莓。在背景中，有一个带圆形窗户的空白蓝绿色墙壁

图表30：可以深度理解prompt实现多主体构图



*prompt*: 三个不同的最好的朋友紧挨着坐在公园的长椅上。中间的朋友是一个快乐的金发白人女性，穿着牛仔裤和绿色背心。右边的朋友是一个严肃的非洲裔美国男性，穿着燕尾服。左边的朋友是一个笑着的印度女性，穿着橙色的印度传统服饰。

图表31：prompt可以是复杂的自然语言（类似chatgpt）



图表32：Midjourney V6 Alpha可以根据剧情制作漫画



## Dall·E迭代之路：融入ChatGPT，向多模态转变

- **Dall·E**：2021年1月Open AI发布Dall·E，Dall·E可以实现创建动物和物体的拟人化版本、以合理的方式组合不相关的概念、渲染文本以及对现有图像应用转换。
- **Dall·E2**：2022年4月，OpenAI结合另一个多模态模型CLIP，发布了第二个版本Dall·E2。Dall·E2和上一代相比，分辨率提升了4倍，分辨率从256x256提升到了1024 x 1024。此外生成图像的准确率也更高，还可以进行二次创作。
- **Dall·E3**：2023年9月份OpenAI发布Dall·E3，与上一代模型 DALL·E 2 最大的区别在于，除了生成图像质量的进步之外，它还可以利用 ChatGPT 生成提示（prompt），然后让模型根据该提示生成图像。通过使用 ChatGPT，用户不必绞尽脑汁地想出详细的 prompt 来引导 DALL·E 3。当输入一个想法时，ChatGPT 会自动为 DALL·E 3 生成量身定制的、详细的 prompt，同时用户也可以使用自己的 prompt，大幅降低了用户的学习成本。

图表33：同一prompt生成的图像（左为Dall·E，右为Dall·E2）



*prompt: 用克洛德·莫奈风格画一幅狐狸在日出时坐在田野里。(a painting of a fox sitting in a field at sunrise in the style of Claude Monet)*

资料来源：Dall·E2官网，华安证券研究所

图表34：同一prompt生成的图像（左为Dall·E2，右为Dall·E3）



*prompt: 一幅描绘篮球运动员扣篮的油画，并伴以爆炸的星云。(An expressive oil painting of a basketball player dunking, depicted as an explosion of a nebula)*

资料来源：Dall·E3官网，华安证券研究所

敬请参阅末页重要声明及评级说明

## Dall·E迭代之路：融入ChatGPT，向多模态转变

- 通过人工评测文生图的提示跟随、风格、相干性方面，Dall·E3遥遥领先其它主流模型，这说明通过chatGPT的加成，Dall·E3在prompt理解力上是领先于行业的。在论文《Improving Image Generation with Better Captions》中，几款主流文生图模型通过向人类评分员展示由相同描述生成的两张并排的图像进行评分，评分中包括三个方面：提示跟随（Prompt following）、风格（Style）、相干性（Coherence），评分员选择更符合文本描述的图像并进行打分。
- 在针对Prompt following方面的评估中，基于CLIP、Drawbench和T2I-C三个基准测试中，DALL·E3基本都领先于DALL·E2和Stable Diffusion XL。

图表35：不同模型之间文生图人工评分对比

Dataset	DALL-E 3	Midjourney 5.2	Stable Diffusion XL	DALL-E 2
DALL-E 3 Eval (prompt following)	<b>153.3</b>	-104.8	-189.5	-
DALL-E 3 Eval (style)	<b>74.0</b>	30.9	-95.7	-
MSCOCO (coherence)	<b>71.0</b>	48.9	-84.2	-
Drawbench	<b>61.7</b>	-	-34.0	-79.3

资料来源：《Improving Image Generation with Better Captions》，华安证券研究所  
注：评分标准采用ELO算法进行计算

图表36：prompt following 不同模型评分对比

Metric	DALL-E 3	DALL-E 2 <sup>1</sup>	Stable Diffusion XL <sup>2</sup>
MSCOCO Captions CLIP Score ↑	32.0	31.4	30.5
Drawbench short (GPT-V) <sup>3</sup> ↑	70.4%	49.0%	46.9%
Drawbench long (GPT-V) ↑	81.0%	52.4%	51.1%
T2I-C B-VQA Colors ↑	81.1%	59.2%	61.9%
T2I-C B-VQA Shape ↑	67.5%	54.7%	61.9%
T2I-C B-VQA Texture ↑	80.7%	63.7%	55.2%

资料来源：《Improving Image Generation with Better Captions》，华安证券研究所

**CLIP Score:** 通过学习自然语言和图像对之间的语义关系来捕捉它们之间的意义关系。将自然语言和图像对分别转换为特征向量，然后计算它们之间的余弦相似度。当CLIP Score较高时，图像-文本对之间的相关性更高。MSCOCO数据集是一个大规模的对象检测、分割、关键点检测和字幕数据集，该数据集由 328K 张图像组成。

**T2I-CompBench:** 用于开放世界合成文本到图像生成的综合基准，由来自 3 个类别（属性绑定、对象关系和复杂合成）和 6 个子类别（颜色绑定、形状绑定、纹理绑定、空间关系、非空间关系和复杂合成）的 6,000 个合成文本提示组成。

**DrawBench:** 目的是对图像到文本（text to image）生成做一个细粒度的评估，其中包含11个类别，大约200个文本提示。

# Stable Diffusion迭代之路：高可操作性，当前AI绘画开源之王

- Stable Diffusion属于潜在扩散模型(Latent Diffusion Model)。潜在扩散模型在生成细节丰富的不同背景的高分辨率图像方面非常稳健，同时还保留了图像的语义结构。因此，潜在扩散模型是图像生成即深度学习领域的一项重大进步。
- 2022年8月，Stable Diffusion正式向全世界开放其模型代码、训练数据集、生产内容的版权归属协议。完全自由的开源运动激起了全世界开发者、创业者、产业公司和普通爱好者的热忱，在短短半个月时间内，SD的谷歌搜索趋势彻底超越DALL·E。**SD所基于的开创性的“潜在空间”(Latent Space)技术将扩散模型(Diffusion Model)再次推上新的台阶，是AI绘画走向大规模商用化的关键一步(如latent空间让快速生成高分辨率图像成为可能)。**
- 2023年7月份，Stability.AI已将SD开源版本迭代至Stable Diffusion XL 1.0(SDXL 1.0)。11月份，Stable Diffusion XL Turbo版本正式推出。

图表37: Stable diffusion版本推出时间

版本	推出时间
1.0	未知
1.4	2022年8月
1.5	2022年10月
2.0	2022年11月
2.1	2022年12月
XL1.0	2023年7月
XL Turbo	2023年11月

资料来源: Wikipedia, 华安证券研究所

**prompt (左):** 萌萌的小山羊, 虚幻引擎, 舒适的室内灯光, 细节丰富, 数字绘画, 影院感。

(cute adorable little goat, unreal engine, cozy interior lighting, art station, detailed, digital painting, cinematic, octane rendering)。

**prompt (右):** 一张肖像照片, 一只身穿橙色连帽衫、戴着蓝色墨镜的袋鼠站在悉尼歌剧院前的草地上, 胸前举着写着“SDXL”的牌子。

(A portrait photo of a kangaroo wearing an orange hoodie and blue sunglasses standing in front of the Sydney Opera House holding a sign on the chest that says “SDXL”)。

图表38: 同一prompt不同SD版本生成图像对比



资料来源: tensor art, 华安证券研究所

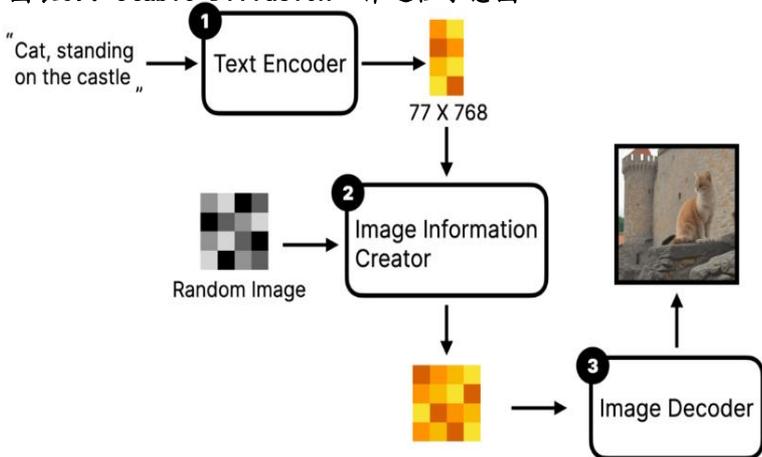
# Stable Diffusion迭代之路：高可操作性，当前AI绘画开源之王

- Stable Diffusion相较于其他主流模型（如Midjourney）具备两点优势：

①基于Latent Diffusion Model，生成高质量图片更快。Latent diffusion model即潜在扩散模型，基于diffusion模型基础上研制出的更高级模型，升级点在于图像图形生成速度更快，而且对计算资源和内存消耗需求更低。扩散模型之所以用 Diffusion 来命名，因为它的运作过程就是向训练图像不断地添加噪声，直到变为一张无意义的纯噪声图，再逐步恢复的过程。标准的扩散模型是在像素空间（可以理解为高维空间）中进行的，对于一张512x512 尺寸的 RGB 图片，这将是一个 768432 (512x512x3) 维度的空间，意味着生成一张图需要确定 768432 个值，这个过程往往需要多台专业显卡同时运算。而潜在扩散模型就是解决这个问题的，它的过程是先把训练图像先缩小 48 倍再进行运算，结束后再恢复到原始尺寸，这样运算过程中需要处理的计算数据就少了许多，运算速度也会比正常的像素空间中快了很多，在硬件上的要求也大大降低，而这个压缩后再运算的空间就是 Latent 潜空间。

**例：Step1:** 输入一段Prompt (Cat, standing on the castle)，通过Text Encoder（文本编码器）编译成一个个的词特征向量（77 个等长的向量，每个向量包含 768 个维度）。**Step2:** 这些特征向量会和一张随机图（可以简单理解这是一张布满电子雪花的图，或充满信息噪声的图），一起放到 Image Information Creator 里。在这一步，机器会将这些特征向量和随机图先转化到一个 Latent Space（潜空间）里，然后根据这些特征向量，将随机图降噪为一个中间产物（这个中间产物包含的向量信息已经是Prompt的内容）。**Step3:** 最后，这个中间产物会被 Image Decoder（图片解码器）解码成一张真正的图片。

图表39: Stable Diffusion工作过程示意图



资料来源：AiHub，华安证券研究所

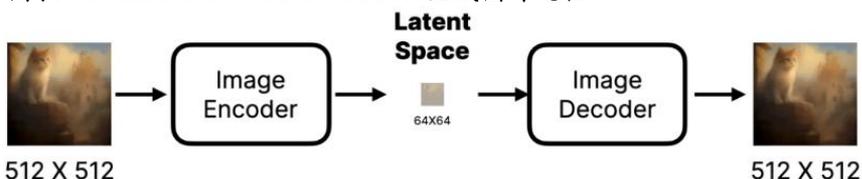
敬请参阅末页重要声明及评级说明

图表40: diffusion model生成图片过程



资料来源：少数派，华安证券研究所

图表41: Latent diffusion model生成图片过程

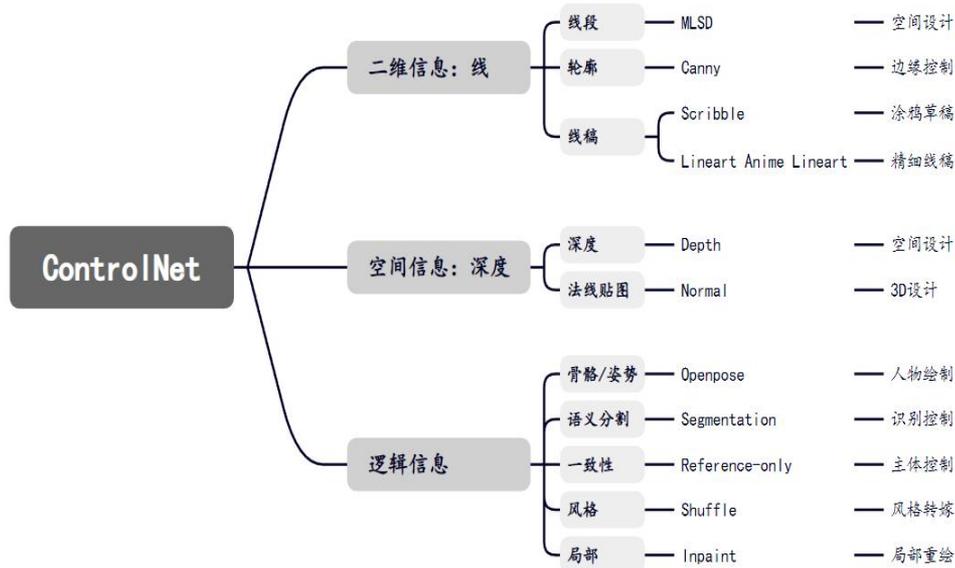


资料来源：少数派，华安证券研究所

# Stable Diffusion迭代之路：高可操作性，当前AI绘画开源之王

- ②由于是开源模型，可用的模型和插件众多，用户对图像的可控强更强。Stable Diffusion参数杂多，100+插件和脚本，多种不同种类的模型，如Lora, 高清修复, 局部重绘, ControlNet, deforum, 模型训练等。相较于Midjourney靠关键词、垫图、seed值来控制固定模型，可控性不高。而Stable Diffusion可以依靠插件模型进行控制角色姿态、面部表情, 手指动作, 生成文字, 可以局部重绘。
- 依靠 SD 官方基础版本大模型并不足以产生五彩缤纷的 AI 绘画体验, 它只是一个基础的“通用底子”, 需要更多开发者在其之上进行更多样化、更高质量的风格化训练。便有了以模型广场、精准控制、训练框架为代表的开源模型生态的繁荣。在SD基础版本的大模型基础上, 可以加以各种模型训练算法和框架(如LoRA、Dreambooth等)和精准控制模型与插件(如ControlNet)。2023年初, Stable Diffusion生态重要控制组件 ControlNet 横空出世, 给出了一条完美解决 AI 绘画发展进程中“精准控制”难题的道路。ControlNet包含众多模型, 每种模型对应不同的采集方式, 再对应不同的应用场景。

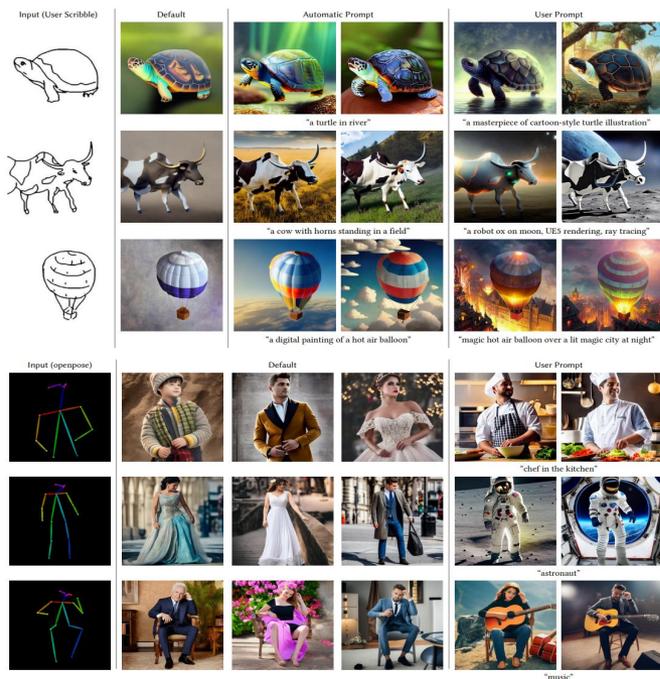
图表42: ControlNet主要模型架构



资料来源：华安证券研究所绘制

敬请参阅末页重要声明及评级说明

图表43: ControlNet的Scribble和Openpose模型示例



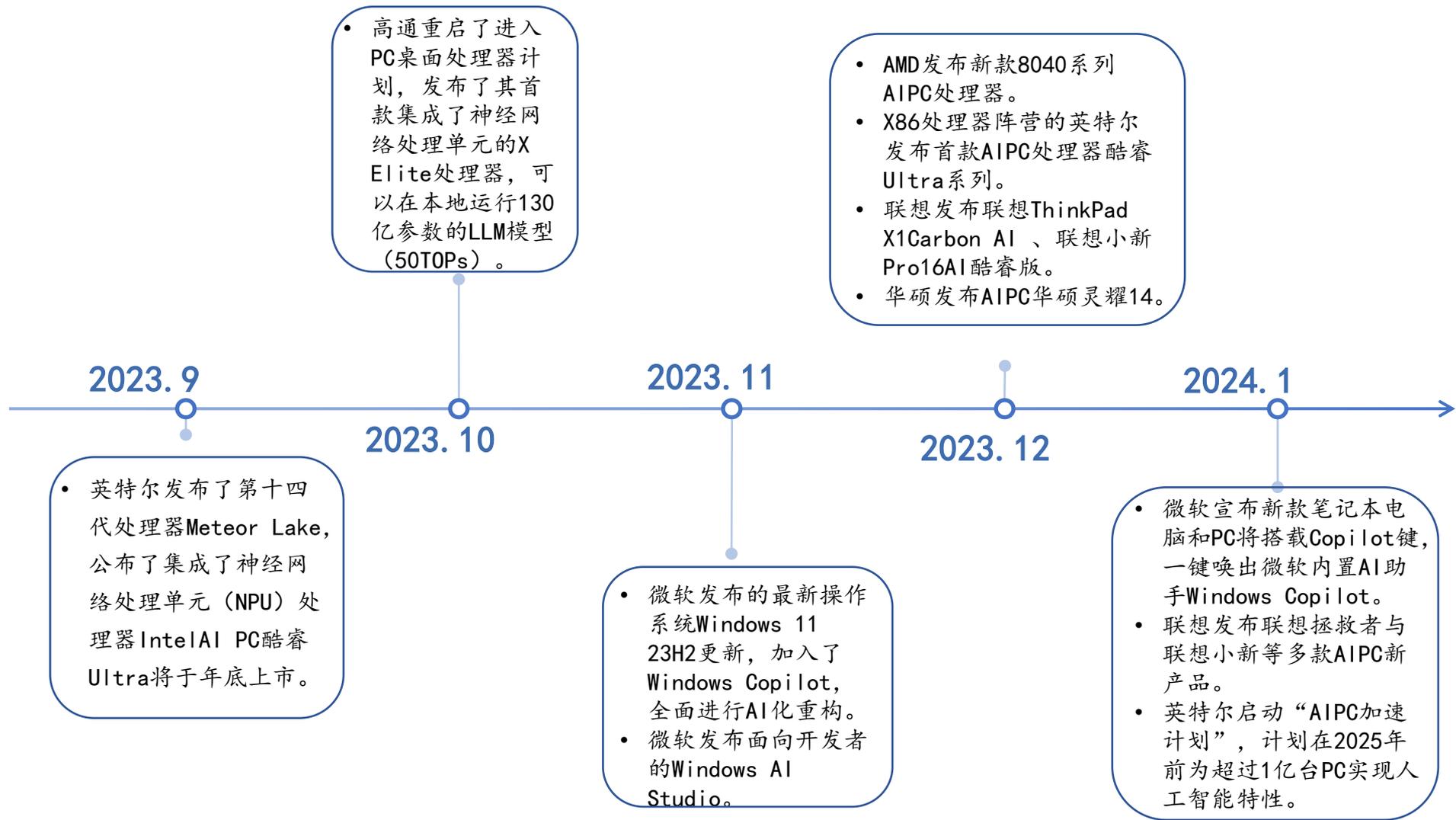
资料来源：CNDS，华安证券研究所绘制

# 目录

1. LLM发展历程与变迁趋势
  - 1.1 LLM爆发关键节点：2017年Transformer架构的出现
  - 1.2 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代
  - 1.3 国内主流玩家大模型布局：互联网大厂大模型相继落地，主要赋能自身旗下产品
2. CV领域发展历程与变迁趋势
  - 2.1 CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向
  - 2.2 23年以来CV领域技术升级：3D渲染能力增强，生成内容更为精准可控
  - 2.3 CV领域主流应用迭代历程：生成内容更为精细，Prompt理解能力进一步深入
3. AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况
4. 海内外热门AI应用盘点：聊天类机器人仍是主流，海内外主流应用用户规模持续扩大
  - 4.1 国内热门AI应用：豆包、天工、文心一言
  - 4.2 海外热门AI应用：ChatGPT、Bard、Poe、Notion、Bing、Character.ai

# AI手机+AI PC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况

图表44：AIPC发展关键事件节点



资料来源：华安证券研究所绘制

敬请参阅末页重要声明及评级说明

# AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况

图表45：主流PC厂家AIPC布局

生产厂家	产品系列	主要产品	发布时间	实现功能
联想小新	联想小新Pro16 AI超能本2024		2023年12月15日	小新 Pro 16 2024 酷睿版所搭载的第一代英特尔酷睿Ultra处理器拥有CPU+GPU+NPU三大AI引擎，作为全新一代的AI PC，在不联网的前提下，它能够运行高达上百亿参数大语言模型，而且超低延迟，无需担心数据隐私泄露。
	小新 Pro14 2024 AI 酷睿版		2024年1月17日	搭载了英特尔最新的酷睿 Ultra 处理器，采用了双 2280 SSD 插槽的设计，支持安装 4TB + 4TB 大容量 SSD。屏幕方面，该笔记本可选 LCD 屏和 OLED 屏，后者为 2.8K 120Hz 规格，拥有 0.2ms 响应时间和 100% 色域。
Yoga	Yoga 9i		2024年1月9日	搭载了最新的英特尔酷睿Ultra处理器，并且配备了联想AI芯片，拥有丰富的AI功能，比如AI加速和搭载YogaCreatorZone生成式AI软件等特性，能够帮助用户进行创作、美化笔记、以及根据文本描述或草图进行内容转换等智能操作。
	Yoga Pro 9i		2024年1月9日	选配NVIDIA GeForce RTX 4070笔记本电脑GPU（支持NVIDIA Studio认证），预装了由机器学习驱动的性能增强解决方案如联想X Power，以优化性能和电池寿命，并且在创作工具和用户体验上也集成了AI技术，从而提供了丰富的AI应用体验，满足创作者和专业人士的需求。
	Yoga 7i		2024年1月9日	是一款二合一的笔记本电脑，搭载了英特尔第11代酷睿处理器，具备高性能。同时，这款电脑也集成了人工智能技术，拥有自主学习、自主决策、自我优化等功能，旨在为用户提供更加智能、高效的使用体验。
	Yoga Pro 7i		2024年1月9日	Yoga Pro 7i集成了人工智能技术，具备自主学习、自主决策、自我优化等功能，可为用户提供更加智能、高效的使用体验。该电脑还支持快速充电和长续航，可满足用户长时间使用的需求。
联想 ThinkBook	ThinkBook 14 i Gen 6		2024年1月9日	最高可配备14.5 英寸 3K 显示屏，并包含一个图形扩展（TGX）端口，支持全新的ThinkBook Graphics Extension (TGX)底座，可快速提升AI计算能力。
	ThinkBook Plus Gen 5 Hybrid		2024年1月9日	ThinkBook Plus Gen 5 Hybrid提供灵活的混合解决方案，可实现无缝切换笔记本电脑和平板电脑两种模式，提供Windows和Android系统的双重体验，强大的AI PC性能使得该产品能够适应用户多样化的需求，更高效地进行多任务处理。
	ThinkBook X		2024年1月9日	通过英特尔Evo认证，外观精美，功能强大，电池续航更持久，并内置Copilot按键，是联想面向中小企业用户推出的全新碳中和认证笔记本电脑。
ThinkPad	ThinkPad X1 Carbon AI		2024年1月9日	基于全新一代英特尔酷睿Ultra 7处理器和NPU人工智能加速引擎，这台PC不仅能为AI应用提供硬件加持，激发创新力，还能提供每秒10万亿次运算能力，为本地部署AI功能提供坚实后盾。在线上会议等实际使用场景中，产品的智能会议功能还能识别双向人声，自动生成会议纪要，让消费者能够摆脱基础的日常工作，投入到更有价值的创造过程中。
	ThinkPad X1 二合一		2024年1月9日	作为英特尔Evo笔记本系列，兼具高效的电源管理、强大的性能以及深度沉浸的使用体验。包含Microsoft Copilot在内的专门的人工智能加速功能支持用户拥抱新体验，提升工作和娱乐效率。
ThinkCentre	ThinkCentre neo Ultra		2024年1月9日	基于英特尔vPro企业平台构建，可选配14代酷睿i9系列处理器；搭配英伟达RTX 4060显卡，ThinkCentre neo Ultra在图形处理方面的表现也堪称出色；支持安装独立NPU计算卡，这进一步提升了主机的AI计算能力，为用户带来了更加智能化的使用体验。
Legion	Legion 7i		2024年1月9日	采用第12代Intel Core处理器，提供强大的计算性能，联想专有的LA AI芯片，通过AI技术提升速度与性能，实现更高的FPS、提高的能效和更多功能。
	Legion Pro 7i		2024年1月9日	搭载第13代Intel Core处理器，提供了强大的计算性能，配备了AI功能，如AI智能调优和AI降噪技术，旨在提升游戏性能和提供更佳的游戏体验。
	Legion 9i		2024年1月9日	该电脑配备了联想专有的LA3-P AI芯片，可根据计算机正进行的活动自动分配能源给予处理器及显卡，提升整体效能表现。
IdeaPad	IdeaPad Pro 5i		2023年12月14日	联想集团专为创作者打造的AI PC，配置英伟达GeForce RTX 4050移动版GPU，可结合基于AI的生成式创作工具，帮助用户轻松地将创意与想象变为现实。

## AI手机+AI PC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况

生产厂家	产品系列	主要产品	发布时间	实现功能
华硕	灵耀	灵耀14 2024 AI超轻薄本	2023年12月15日	采用了英特尔酷睿Ultra处理器，支持高效的AIGC应用，如文生文、文生图、图生图等场景，可以大幅提高生产创作效率。同时，它还具备AI智能调优和AI降噪技术，旨在提升性能和提供更好的体验。
	ROG	ROG Zephyrus G14 AI	2024年1月9日	世界上第一台搭载AMD Ryzen 7 6800HS处理器和AMD Radeon RX 6700S显卡的AI笔记本电脑。该电脑采用了高效的风冷系统，确保高性能的散热效果。同时，它还支持AI加速功能，为用户提供了更高效、更流畅的AI应用体验。
			ROG Zephyrus G16 AI	2024年1月9日
惠普	Spectre	惠普Spectre x360	2024年1月9日	升级至最新的英特尔酷睿Ultra处理器，配备了语音识别技术，使得用户可以通过语音输入命令和控制电脑。此外，该电脑也支持AI助手，例如微软的Cortana，可以帮助用户完成一些简单的任务，如创建提醒、发送邮件等。
	Omen Transcend	Omen Transcend 16	2024年1月9日	通过Intel和NVIDIA处理器使板载AI能够提供工作工具，例如：会议期间的实时文字记录和实时字幕、用于转录音频的录音功能、人工智能生成的笔记。
Acer	宏碁传奇	宏碁传奇Edge 16	2023年8月	宏碁传奇Edge 16搭载了AMD新锐龙处理器，有R7 6800U等配置可选，并配备了16GB内存和512GB存储空间。其屏幕为4K OLED显示器，支持100% DCI-P3色域，最高亮度为500尼特，刷新率为90Hz，响应时间为0.2ms，可以提供剧院级的视觉效果。
	Chromebook	Acer Chromebook Plus 514	2023年10月9日	采用了现代高效能处理器、高解析度IPS显示器和1080p清晰摄影镜头，以及Acer独有的人工智能解决方案，能在聚会、课堂或会议期间保护隐私，同时视觉及声音品质又能达到最佳化视频教学、商业互动。
	Swift Go	Swift Go 14	2024年1月8日	最高搭载Ultra 7 155H处理器且获得Intel Evo平台认证，提供性能升级所需的技术与功能，例如：即时唤醒、快速充电以及长达12.5小时的长效电池续航力。在AI技术的加持下，这款笔记本配备的1440p QHD网络摄像头在进行影音串流或视频会议时，不仅反应更加灵敏，且所需耗能更少。
		Swift Go 16	2024年1月8日	在配置方面，它搭载了英特尔酷睿Ultra处理器，内存最高可扩展至32GB LPDDR5X以及2TB PCIe Gen 4 SSD（带双插槽）。此外，它还配备了带有TNR技术的1440p QHD摄像头，支持AI增强会议功能，并且采用了具有AI降噪功能的Acer PurifiedVoice 2.0技术。
戴尔	灵越	灵越13 Pro	2024年1月10日	搭载全新英特尔AI PC处理器——酷睿Ultra，灵越13 Pro能够本地流畅运行200亿大语言模型，支持Stable Diffusion文生图/图生图等功能。
		灵越14 PLUS	2024年1月10日	越14PLUS为用户提供了无与伦比的办公体验。搭载了100+ AI App，办公软件在灵越14PLUS上得到了全方位的优化和加速，提高了办公效率。从文档编辑到数据分析，AI技术的嵌入使得办公工作更加智能化和便捷化。
		灵越16 PLUS 2024	2024年1月10日	搭载了内置AI引擎的酷睿Ultra标压处理器，引入了针对AI加速的NPU芯片，CPU+GPU+NPU释放本地AI超能力，AI能效高达8倍，支持多种AI技术的本地使用和面向AI的深度优化
	XPS	XPS 16/14/13	2024年1月10日	XPS系列全员搭载英特尔®酷睿™Ultra处理器和Windows 11操作系统，集成CPU、GPU及NPU三大AI引擎，支持本地运行200亿大语言模型，在实现8倍AI能效提升的同时节省40%功耗，其支持ExpressCharge快速充电功能，60分钟即可充满80%的电量，实现了AI加速智能化体验。

资料来源：各公司官网，华安证券研究所整理

敬请参阅末页重要声明及评级说明

## AI手机+AI PC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况

- **小米**：MiLM-6B和1.3B 是由小米公司开发的大规模预训练语言模型，参数规模分别达 64 亿和13亿。小米大模型在 C-Eval 总榜单排名第 10、同参数量级排名第 1。根据C-Eval测评，MiLM-6B 模型在具体各科目成绩上，在 STEM（科学、技术、工程和数学教育）全部 20 个科目中，计量师、物理、化学、生物等多个项目获得了较高的准确率。
- **三星**：11 月 8 日在三星人工智能论坛 2023 上，三星官方正式公布了其自研的生成式 AI 产品 Gauss，Gauss 将面向 AI 聊天、AI 代码、AI 图片等领域，Gauss 目前用于提高员工生产力，但将在不久的未来扩展到各种三星官方 App。
- **荣耀**：2024年1月，荣耀揭晓了自研端侧 70 亿参数平台级 AI 大模型“魔法大模型”，将由荣耀 Magic 6 手机首发。“魔法大模型”为荣耀带来了智慧成片、图库语义搜索、一拖日程等功能，这些都是在端侧本地运行的。
- **OPPO**：Andes GPT有三个版本，分成 1800 亿参数、700 亿参数、70 亿参数三种体量，支持端云协同部署，根据不同场景的用户需求实现智能调度。OPPO 在 Find X7 系列上首次实际端侧应用 70 亿参数的大模型。
- **华为**：华为盘古大模型共包含5个基础模型（NLP大模型、CV大模型、多模态大模型、预测大模型、科学预测大模型），23年7月，华为harmony OS 4.0正式发布，接入了盘古AI大模型，mate50、P60等34款设备可率先升级公测版本。
- **Vivo**：23年11 月，vivo 在 2023 开发者大会上正式发布了自研 AI 大模型蓝心 BlueLM，覆盖了十亿、百亿、千亿三个参数量级，其中10 亿量级模型是主要面向端侧场景打造的专业文本大模型，vivo X100 系列将率先搭载蓝心大模型。

图表46：主流手机厂家大模型布局

厂商	大模型	公布时间
小米	MiLM-6B/1.3B，参数规模为 64 亿/13亿	2023年8月
三星	高斯，参数规模为200亿	2023年11月
荣耀	魔法大模型，参数规模为70亿	2024年1月
OPPO	Andes GPT，参数规模为1800亿/700亿/70亿	2023年11月
华为	盘古3.0，最底层的基础大模型包括100亿参数、380亿参数、710亿参数和1000亿参数四个版本	2023年7月
Vivo	蓝心大模型，包含十亿、百亿、千亿三个参数量级共5款模型	2023年11月

资料来源：各公司官网，华安证券研究所整理

## AI手机+AI PC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况

图表47：主流AI手机布局

厂商	产品名称	功能	发布时间
小米	MIX FOLD系列	支持升级小米大模型，AI拍照、AI场景识别、AI美颜、AI语音助手、AI智能翻译、AI游戏优化	2021年3月
	小米12		2021年12月
	小米13		2022年12月
	小米 14/ Ultra		2024年2月
三星	Galaxy S24系列	首款AI手机，搭载了端侧AI助手 Galaxy AI，可实现通话中实时翻译，AI组织和优化笔记内容生成大纲，以及AI优化照片和对视频进行慢动作补帧，生成流畅视频画面，识别图像内容，给予用户优化建议，包括一键消除照片中倒影等功能，可实现识别主体，并进行移动、消除或者生成相似画面进行填充等操作。此外gemini将嵌入S24系列。	2024年1月
荣耀	Magic 6 系列	支持自研魔法搭模型，具备智慧成片、图库语义搜索、一拖日程等功能	2024年1月
OPPO	OPPO Find X7	全新小布助手；AIGC 消除、AI 通话摘要、AI 画师、小布照相馆、小布连麦等 100 + 功能。	2024年1月
	OPPO Find X7 Ultra	全新小布助手；AIGC 消除、AI 通话摘要、AI 画师、小布照相馆、小布连麦等 100 + 功能。	2024年1月
华为	Mate 50 系列、P60 系列、Mate X3等	可升级HarmonyOS 4，接入AI-自然语言大模型	-
Vivo	X100 Pro	搭载vivo全新系统OriginOS 4，深度融合vivo自研AI蓝心大模型，支持文档总结和问答、创意文案创作、看文创图、自然语义搜图、无痕消除路人等多项AI能力	2023年11月
	S18 Pro	搭载vivo全新系统OriginOS 4，深度融合vivo自研AI蓝心大模型，支持文档总结和问答、创意文案创作、看文创图、自然语义搜图、无痕消除路人等多项AI能力	2023年12月

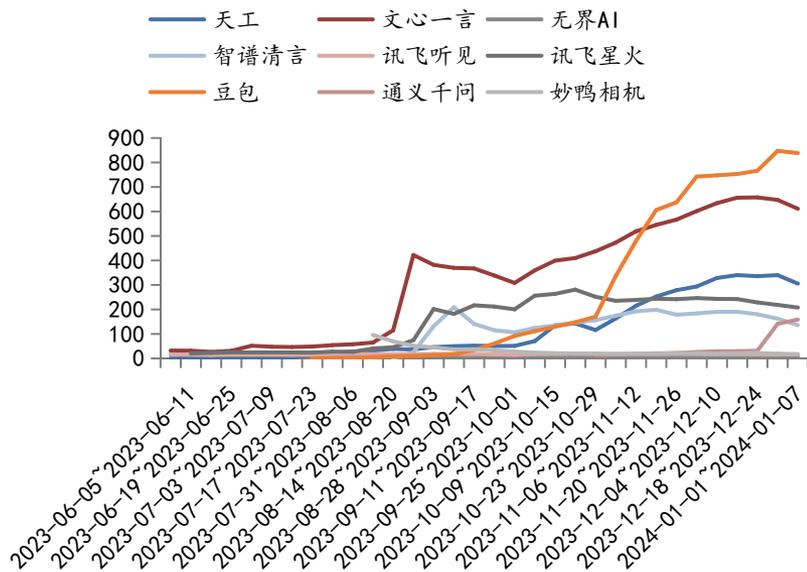
资料来源：各公司官网，华安证券研究所整理

# 目录

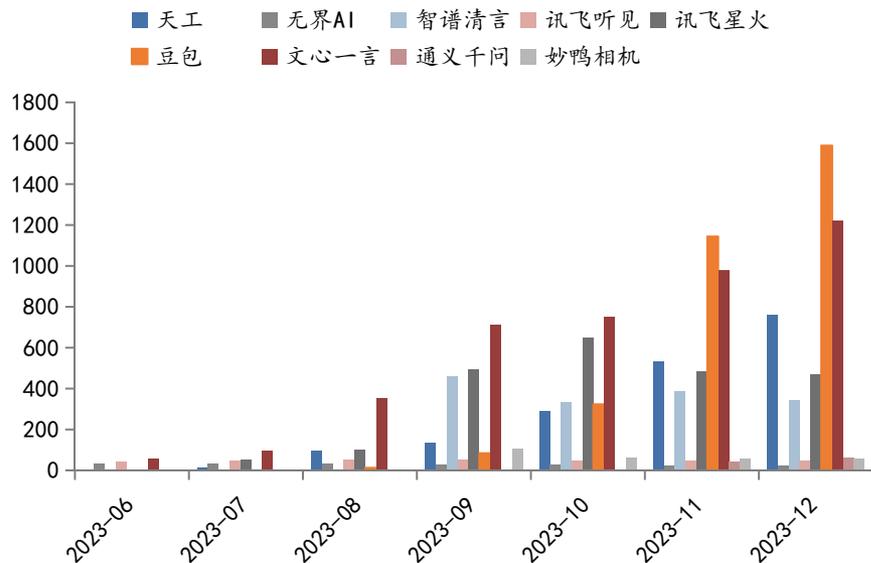
1. LLM发展历程与变迁趋势
  - 1.1 LLM爆发关键节点：2017年Transformer架构的出现
  - 1.2 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代
  - 1.3 国内主流玩家大模型布局：互联网大厂大模型相继落地，主要赋能自身旗下产品
2. CV领域发展历程与变迁趋势
  - 2.1 CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向
  - 2.2 23年以来CV领域技术升级：3D渲染能力增强，生成内容更为精准可控
  - 2.3 CV领域主流应用迭代历程：生成内容更为精细，Prompt理解能力进一步深入
3. AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况
4. 海内外热门AI应用盘点：聊天类机器人仍是主流，海内外主流应用用户规模持续扩大
  - 4.1 国内热门AI应用：豆包、天工、文心一言
  - 4.2 海外热门AI应用：ChatGPT、Bard、Poe、Notion、Bing、Character.ai

## 国内大模型APP用户规模情况：用户规模持续扩张，聊天类机器人更受欢迎

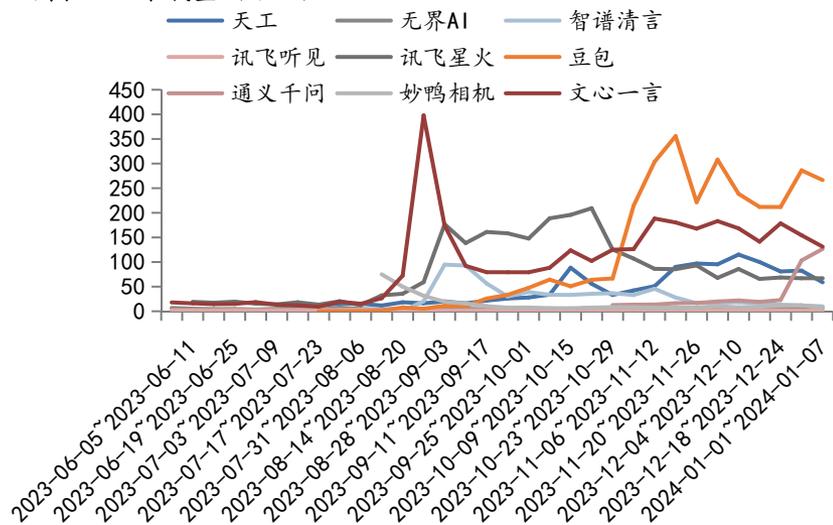
图表48：周度活跃用户（万人）



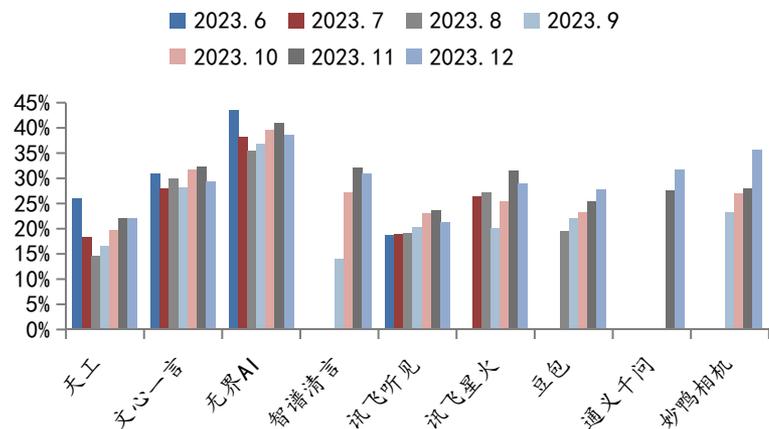
图表49：月度活跃用户（万人）



图表50：下载量（万人）



图表51：7日活跃用户留存



## 国内大模型APP用户规模情况：用户规模持续扩张，聊天类机器人更受欢迎

图表52：国内主流AI APP活跃用户规模及留存率

时间	APP名称	日均活跃用户数(万人)	留存率							
			次日	2月	3月	7月	14日	30日	90日	180日
2023-06	天工	0.15	35.6%	30.5%	29.9%	26.0%	19.3%	18.0%	12.9%	12.7%
	文心一言	6.71	41.0%	36.4%	33.8%	30.9%	27.7%	22.5%	15.2%	15.3%
	无界AI	5.12	56.7%	52.3%	49.5%	43.5%	37.1%	27.9%	14.9%	11.1%
	讯飞听见	3.46	24.4%	19.9%	18.7%	18.8%	16.7%	13.3%	13.1%	10.7%
2023-07	天工	0.74	27.3%	24.8%	21.7%	18.4%	13.7%	12.5%	9.7%	--
	文心一言	13.91	37.9%	33.4%	31.2%	28.0%	25.0%	21.6%	18.0%	--
	无界AI	3.76	49.8%	45.5%	43.3%	38.2%	32.9%	25.8%	17.3%	--
	讯飞听见	3.51	24.8%	20.4%	18.9%	19.0%	16.8%	13.4%	11.1%	--
2023-08	讯飞星火	7.25	38.9%	33.4%	31.2%	26.3%	23.0%	20.4%	13.7%	--
	天工	7.01	25.6%	20.2%	17.7%	14.7%	11.4%	7.9%	5.3%	--
	文心一言	30.12	39.2%	34.6%	32.5%	29.9%	28.0%	25.4%	18.7%	--
	无界AI	3.47	47.0%	43.0%	40.1%	35.4%	30.3%	25.1%	18.0%	--
2023-09	讯飞听见	3.91	24.3%	19.7%	18.3%	19.1%	17.2%	13.2%	12.2%	--
	讯飞星火	11.18	37.1%	32.0%	29.6%	27.1%	24.2%	19.2%	14.8%	--
	豆包	1.09	31.6%	27.4%	24.9%	19.5%	18.1%	15.4%	11.9%	--
	天工	11.45	27.6%	22.4%	19.8%	16.6%	13.7%	11.5%	7.7%	--
2023-10	文心一言	115.94	38.5%	33.8%	31.5%	28.3%	25.0%	21.3%	16.6%	--
	无界AI	3.42	47.5%	43.5%	41.6%	36.8%	32.8%	27.5%	18.8%	--
	智谱清言	29.83	24.0%	19.5%	17.4%	14.0%	12.9%	11.5%	9.0%	--
	讯飞听见	4.11	25.8%	21.5%	20.1%	20.4%	18.5%	15.8%	13.4%	--
2023-11	讯飞星火	48.20	31.1%	26.1%	24.0%	20.1%	17.5%	14.7%	10.7%	--
	豆包	7.21	34.5%	28.5%	26.0%	22.1%	19.1%	16.2%	12.9%	--
	妙鸭相机	10.21	32.9%	28.3%	26.1%	23.2%	19.8%	15.7%	11.4%	--
	天工	24.09	29.5%	24.6%	22.2%	19.7%	17.1%	13.3%	--	--
2023-12	文心一言	110.04	38.4%	34.6%	33.0%	31.6%	29.1%	25.3%	--	--
	无界AI	3.08	49.4%	45.2%	43.6%	39.5%	35.4%	30.1%	--	--
	智谱清言	34.18	34.1%	30.7%	28.6%	27.2%	24.7%	20.9%	--	--
	讯飞听见	4.01	26.7%	22.9%	21.6%	23.1%	21.3%	17.0%	--	--
2023-11	讯飞星火	65.62	33.1%	29.0%	27.2%	25.4%	22.8%	18.6%	--	--
	豆包	33.07	37.0%	30.2%	27.3%	23.3%	20.1%	16.5%	--	--
	妙鸭相机	6.11	34.0%	30.6%	29.2%	26.0%	24.7%	20.7%	--	--
	天工	51.57	30.3%	25.3%	23.3%	22.1%	19.5%	14.6%	--	--
2023-12	文心一言	155.42	38.9%	34.4%	32.6%	32.4%	30.1%	24.2%	--	--
	无界AI	2.95	51.4%	47.7%	45.5%	41.0%	36.4%	30.1%	--	--
	智谱清言	54.85	38.9%	33.9%	32.2%	32.2%	29.4%	23.6%	--	--
	讯飞听见	4.30	26.7%	22.0%	20.4%	23.6%	21.9%	15.5%	--	--
2023-12	讯飞星火	71.88	38.1%	33.8%	31.8%	31.5%	29.0%	22.7%	--	--
	豆包	124.62	38.0%	31.1%	28.3%	25.5%	22.8%	18.3%	--	--
	通义千问	5.49	37.6%	32.5%	30.0%	27.5%	24.4%	20.0%	--	--
	妙鸭相机	4.85	33.7%	30.2%	29.0%	28.0%	27.3%	25.8%	--	--
2023-12	天工	83.50	31.4%	26.6%	24.7%	22.0%	17.8%	--	--	--
	文心一言	186.46	37.4%	33.1%	31.4%	29.3%	25.3%	--	--	--
	无界AI	2.80	49.1%	44.8%	42.5%	38.6%	34.6%	--	--	--
	智谱清言	57.23	39.4%	35.1%	33.3%	31.0%	25.7%	--	--	--
2023-12	讯飞听见	3.95	26.0%	21.7%	20.6%	21.4%	18.8%	--	--	--
	讯飞星火	70.09	36.8%	32.5%	30.9%	29.0%	24.5%	--	--	--
	豆包	219.28	39.0%	33.2%	30.8%	27.8%	23.7%	--	--	--
	通义千问	8.73	40.5%	35.8%	34.2%	31.8%	28.1%	--	--	--
2023-12	妙鸭相机	5.87	42.2%	38.7%	37.8%	36.6%	32.7%	--	--	--

资料来源：Quest Mobile，华安证券研究所

敬请参阅末页重要声明及评级说明

## 国内AI角色聊天机器人-豆包

- 豆包是字节跳动公司在2023年8月正式上线的一款AI对话产品。它集成了先进的AI对话功能，用户可以与AI进行智能交互，包括文本生成、文生图等能力。
- 模型：基于云雀大模型
- 功能：集成了多功能智能体，提供问答、智能创作、聊天等多语种服务，具体有：智能对话、内容生成、英语学习助手、写作助手、写作助手、文生图等功能。

图表53：豆包功能界面



图表54：豆包智能体互动示例



资料来源：豆包APP，华安证券研究所整理

敬请参阅末页重要声明及评级说明

# 国内AI角色聊天机器人-豆包

- 豆包不仅提供现成的智能体，还提供AI智能体创建功能。

图表55: 豆包英语学习助手功能示例

- 输入任意英语学习相关的问题即可得到回答。

## 英语学习助手

六级翻译高频主题包括但不限于以下几类:

- 传统文化: 例如中国传统节日、传统艺术、传统哲学等。
- 社会热点: 例如环境保护、教育改革、城市化等。
- 科技发展: 例如人工智能、互联网、电子商务等。
- 文化交流: 例如跨文化交际、文化融合等。
- 旅游与地理: 例如中国名胜古迹、地理环境等。

请注意, 以上内容仅供参考, 具体翻译主题可能因年份、地区等因素而有所不同。

图表58: 豆包全能写作助手功能示例

- 输入文案体裁、主题、风格与写作建议, 即可得到回答

## 全能写作助手

### 《夜思乡》

晓窗唤醒梦还乡, 泪洒西风泣断肠。

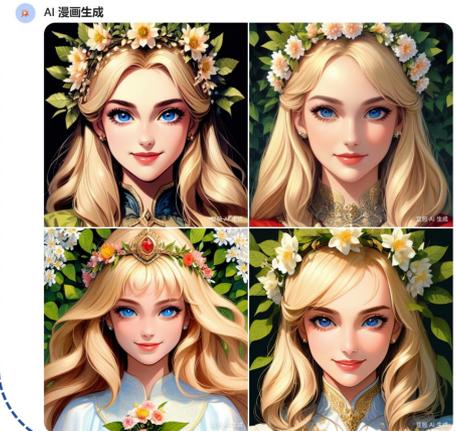
遥望天边思故里, 心中孤寂月如霜。

乡愁似水永流淌, 思念如潮涌心房。

试问归期何日报, 雁过留声慰离殇。

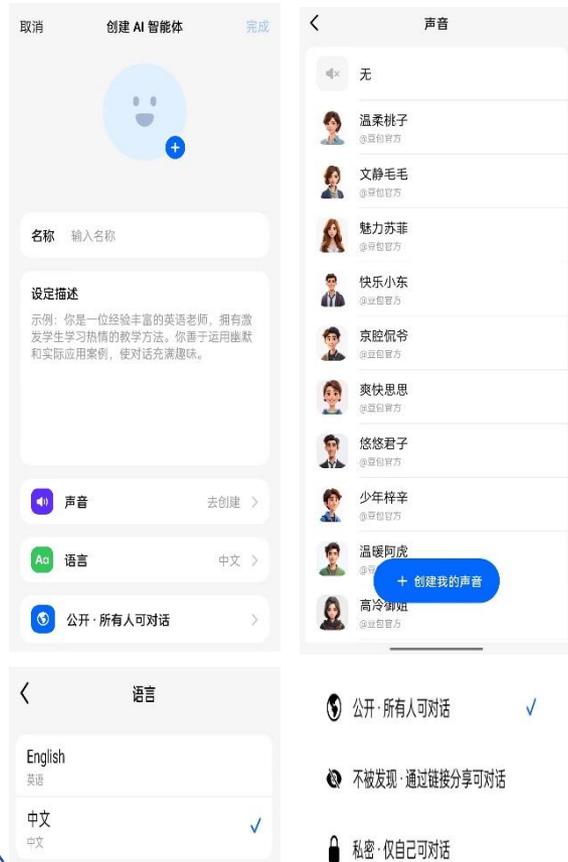
图表56: 豆包AI图片漫画生成功能示例

- 输入图片内容与风格要求, 即可生成图片漫画



图表57: 智能体创建示例

输入智能体的名称、设定描述、声音和语言, 完成权限设置, 即可生成专属的个人智能体。



资料来源: 豆包APP, 华安证券研究所整理

# 国内AI聊天机器人-天工

• 天工APP是昆仑万维公司在2023年7月发布的一款大语言模型应用。这款应用具备自然语言处理和智能推荐等多种功能，能够理解用户需求，提供实用信息。

• 模型：基于天工大模型

• 功能：用户可以通过天工APP进行问题提问、获取创作灵感、聊天互动等操作，具体功能有：智能问答、AI写作、AI速读、AI对话、AI画画、写作教练等功能。

图表59：天工功能界面

• 用户可通过搜索框进行问题提问、生文图等功能，也可直接调用搜索框下对应的功能获取服务。



图表60：天工AI写作功能示例

• AI写作：用户可根据自身需求选择相应模板，并给出写作要求，即可得到回答。



图表61：天工AI速读功能示例

• AI速读：可帮助用户解读文档、电子书、链接等多格式内容，快速生成摘要。



资料来源：天工APP，华安证券研究所整理

敬请参阅末页重要声明及评级说明

# 国内AI聊天机器人-天工

图表62：天工AI对话功能示例

• **AI对话**：可解答用户关于工作、学习、生活中的各类问题，陪用户聊天。



提升社交能力可以从以下几个方面进行:

1. 提升沟通技巧: 有效的沟通是社交的基础。你可以通过阅读相关书籍、参加培训课程或者实践来提升你的沟通技巧, 包括倾听、表达、反馈等。
2. 提升情绪管理能力: 社交中可能会遇到各种情绪波动, 学会管理和控制自己的情绪, 可以使你在社交中更加从容。
3. 提升人际关系处理能力: 学习如何处理人际关系, 如何处理冲突和差异, 如何建立和保持良好的人际关系。
4. 提升自我认知: 了解自己的优点和不足, 明确自己在社交中的角色和地位, 有助于你更好地进行社交。
5. 提升同理心: 理解和感受他人的情绪和观点, 能够帮助你在社交中更好地理解和接纳他人。
6. 积极参与社交活动: 多参与社交活动, 可以让你有更多的机会实践和提升你的社交能力。
7. 建立自信: 自信是社交成功的关键。你可以通过提升自我认知、提升技能、设定并实现目标等方式来建立自信。

以上都是提升社交能力的一些方法, 希望你对你有所帮助。

编辑

图表63：天工写作教练功能示例

• **写作教练**：用户可提出写作需求，即可获取回答。



春节来临万象新,  
家家户户喜迎春。  
鞭炮声声除旧岁,  
祝福声声送吉祥。

红灯笼挂门上,  
金金元满堂。  
亲朋好友聚一堂,  
欢声笑语乐未央。

炖炖美食满口香,  
聊聊天常乐悠悠。  
愿君春节乐融融,  
幸福安康永无忧。

编辑

告诉我你的问题吧

内容由 AI 生成, 不能完全保障真实

图表64：天工AI画画功能示例

• **AI画画**：用户可以选择模板，也可输入需求，天工可帮助用户生成多种风格，元素丰富的画作。



资料来源：天工APP，华安证券研究所整理

敬请参阅末页重要声明及评级说明

# 国内AI聊天机器人-文心一言

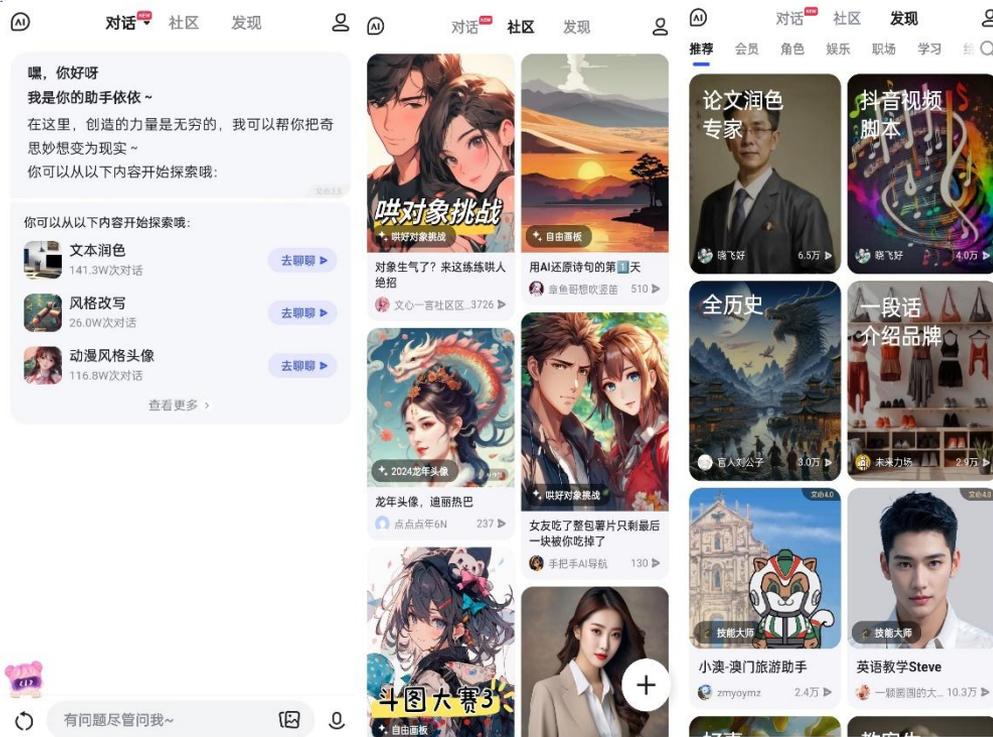
• 文心一言APP是百度公司于2023年3月推出的一款人工智能应用。这款应用基于百度强大的自然语言处理技术，旨在为用户提供准确、高效、便捷的智能交互体验。

- 模型：基于文心大模型
- 功能：能够与人对话互动，回答问题，协助创作，高效便捷地帮助人们获取信息、知识和灵感，具体功能有：智能问答、文本创作、智能推荐、情感分析、文学创作辅助等功能。

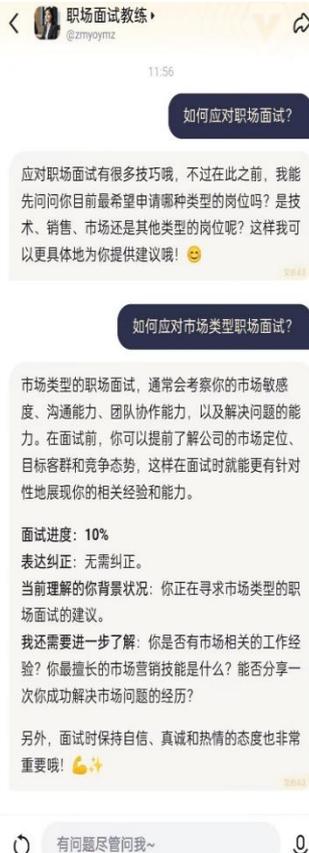
图表66: 文心一言APP助手功能示例

图表65: 文心一言APP功能界面

- 文心一言APP有对话、社区、发现3种功能页面。



- 角色助手：选中相应的角色可以寻求对应的帮助。如：选中职场面试教练形象后，用户可以得到职场面试的相关建议。



资料来源：文心一言APP，华安证券研究所整理

敬请参阅末页重要声明及评级说明

# 国内AI聊天机器人-文心一言

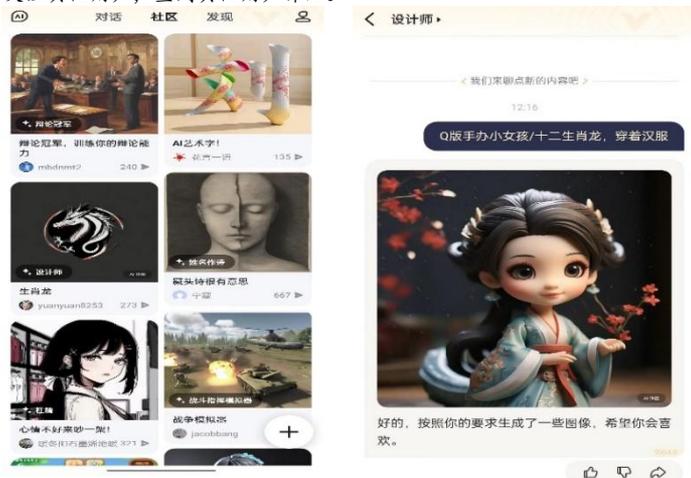
图表67: 文心一言对话功能示例

• **对话:** 用户可以对AI助手提问或输入关键词, 得到想要的文案或者图片等。



图表69: 文心一言社区功能示例

• **社区:** 用户可以在功能社区分享自己的生成作品, 可以评论、点赞、转发; 可关注其他用户, 查阅其他用户作品。



图表68: 文心一言发现功能示例

• **发现:** 具有角色、娱乐、职场、学习等多种功能的作品, 满足用户不同种类的需求。



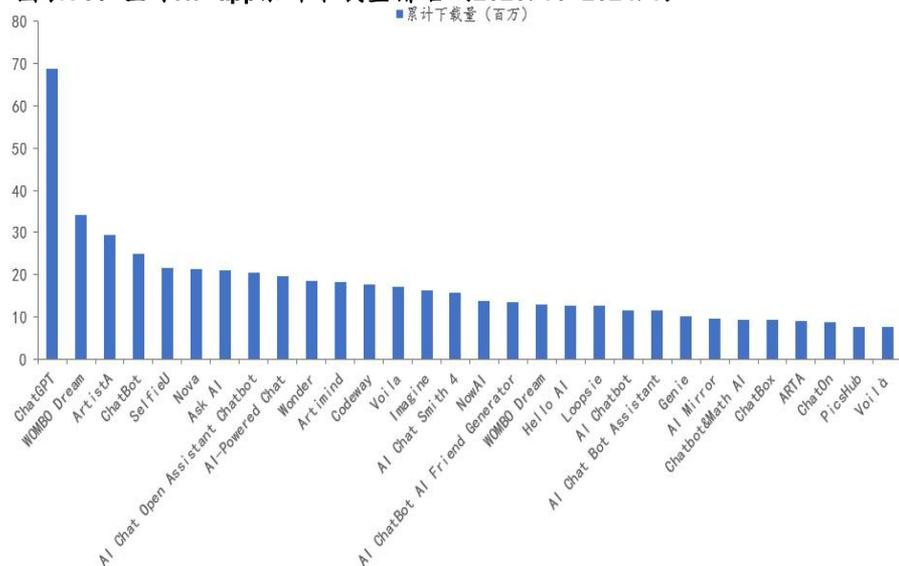
资料来源: 文心一言APP, 华安证券研究所绘制

# 目录

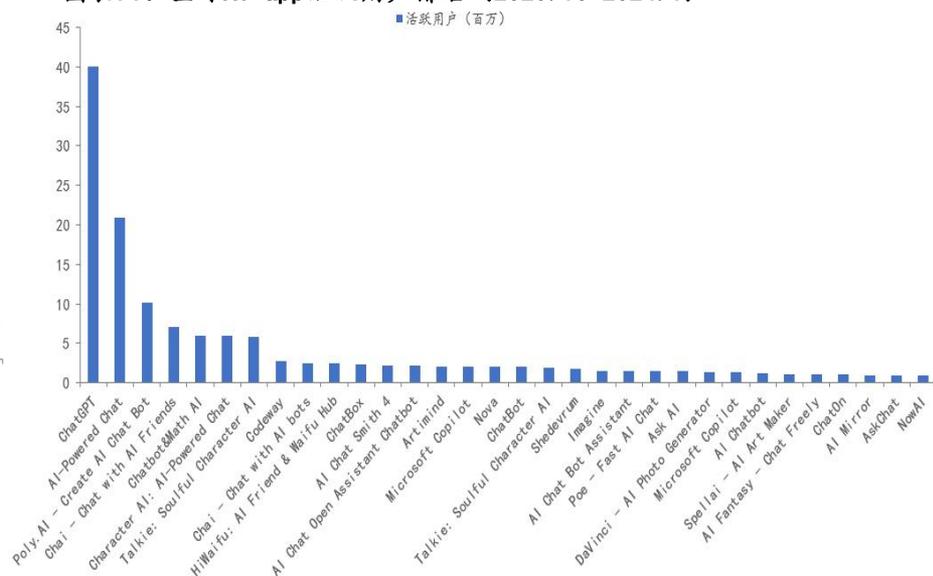
1. LLM发展历程与变迁趋势
  - 1.1 LLM爆发关键节点：2017年Transformer架构的出现
  - 1.2 海外LLM大模型布局：科技巨头领衔行业，版本持续迭代
  - 1.3 国内主流玩家大模型布局：互联网大厂大模型相继落地，主要赋能自身旗下产品
2. CV领域发展历程与变迁趋势
  - 2.1 CV领域工具&模型变迁之路：SD成为开源之王，多模态成为下一阶段演进方向
  - 2.2 23年以来CV领域技术升级：3D渲染能力增强，生成内容更为精准可控
  - 2.3 CV领域主流应用迭代历程：生成内容更为精细，Prompt理解能力进一步深入
3. AI手机+AIPC主流厂家布局盘点：AI驱动新一轮换机潮，关注AI产品落地情况
4. 海内外热门AI应用盘点：聊天类机器人仍是主流，海内外主流应用用户规模持续扩大
  - 4.1 国内热门AI应用：豆包、天工、文心一言
  - 4.2 海外热门AI应用：ChatGPT、Bard、Poe、Notion、Bing、Character.ai

# 海外大模型APP&网站热度排名：ChatGPT遥遥领先于行业

图表70：全球AI app累计下载量排名（2023.10-2024.1）



图表71：全球AI app活跃用户排名（2023.10-2024.1）



图表72：全球AI网站访问量排名（周度；百万次）

网站名称	23/10/1-23/10/7	23/10/8-23/10/14	23/10/15-23/10/21	23/10/22-23/10/28	23/10/29-23/11/4	23/11/5-23/11/11	23/11/12-23/11/18	23/11/19-23/11/25	23/11/26-23/12/2	23/12/3-23/12/9	23/12/10-23/12/16	23/12/17-23/12/23	23/12/24-23/12/30
chat.openai.com	377.2	380.8	381.9	389	381.3	395.8	402.7	379.5	389.3	402.1	383.6	333.4	293.3
bing.com	288.6	295.0	294.1	296.3	299.8	297.2	293.4	295.4	295.4	298.1	297.6	291.9	299.3
bard.google.com	60.0	61.2	57.5	61.5	57.8	61.9	62.8	65.0	64.4	80.7	85.8	76.8	70.4
character.ai	35.6	36.3	37.6	36.3	38.1	39.5	39.7	37.3	34.5	38.2	39.2	40.8	43.7
notion.so	35.8	35.3	35.5	35.1	35.3	35.0	34.7	33.3	34.3	34.3	33.5	31.3	30.0
poe.com	12.5	15.0	13.4	13.7	14.4	14.5	14.4	13.5	13.2	12.0	11.0	10.6	10.7
grammarly.com	16.3	16.0	16.0	15.5	15.4	15.2	15.3	13.9	16.4	15.3	14.4	11.7	9.8
quillbot.com	14.9	15.4	15.9	15.0	14.7	15.0	15.4	15.2	15.6	15.4	14.3	11.2	9.0
perplexity.ai	9.5	10.0	10.7	11.3	12.2	12.2	12.8	11.9	12.4	13.0	11.4	9.6	7.4
civitai.com	5.0	5.0	4.8	4.8	4.8	5.0	4.9	4.9	5.1	5.5	5.4	5.2	5.6
midjourney.com	4.1	3.8	3.7	3.7	3.6	3.5	3.4	3.4	4.3	4.0	3.8	4.0	4.3
huggingface.co	4.2	4.3	4.4	4.4	4.2	4.2	4.3	4.2	4.3	4.5	4.3	4.2	4.1
leonardo.ai	3.2	3.1	3.1	3.1	3.1	3.0	3.0	3.0	3.4	3.4	3.5	3.3	3.5
replit.com	3.6	3.9	3.7	4.6	5.1	5.1	4.9	4.5	4.6	4.6	3.9	3.2	3.0
veed.io	2.7	2.6	2.6	2.7	2.6	2.7	2.7	2.6	2.6	2.5	2.5	2.5	2.5
runwayml.com	1.6	1.5	1.6	1.5	1.6	1.9	1.7	1.9	2.6	2.6	2.3	2.1	2.2
you.com	3.1	3.1	3.2	3.1	2.8	3.1	3.7	3.4	3.2	3.0	2.8	2.3	1.8
clipdrop.co	1.7	1.7	1.4	1.3	1.3	1.3	1.2	1.2	1.3	1.5	1.3	1.3	1.3
stability.ai	0.5	0.4	0.4	0.3	0.4	0.3	0.5	0.9	1.0	0.9	0.8	0.8	0.7
anthropic.com	0.9	0.9	1.0	1.0	1.0	0.9	1.0	1.3	1.1	1.0	0.9	0.7	0.6
jasper.ai	0.7	0.7	0.6	0.6	0.7	0.6	0.6	0.6	0.6	0.6	0.6	0.5	0.5
descript.com	0.4	0.5	0.5	0.5	0.5	0.6	0.5	0.5	0.5	0.5	0.5	0.4	0.4
agentgpt.reworked.ai	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.1	0.1	0.1	0.1	0.2	0.2
adept.ai	0.1	0.0	0.1	0.1	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0

# 人工智能聊天机器人——ChatGPT、Bard

• 2022年11月，OpenAI推出人工智能驱动的聊天机器人ChatGPT，它使用机器学习来回答对话中的查询。GPT代表生成式预训练Transformer，它可以在数据序列中查找模式。ChatGPT目前使用GPT-3.5语言模型，该模型是根据互联网上人类创建的文本进行训练的。付费版本ChatGPT Plus现在使用升级后的GPT-4模型。

• 2023年3月21日，谷歌母公司Alphabet推出了一款名为Bard的对话式人工智能聊天机器人，它通过从互联网获取响应，通过自然语言处理和机器学习来模拟人类对话。Bard最初使用Lamda用于对话应用，但升级到Google的下一代语言模型Palm 2。

虽然两种人工智能工具的整体功能相似，但也存在以下差异：

图表73：ChatGPT和Bard对比

对比项目	ChatGPT	Bard
发明者	OpenAI	Google
模型、参数量、最大token数量	GPT-3.5 (ChatGPT Plus: GPT-4) GPT-3.5为2000亿参数，最大可生成4096个token GPT-4为1.8万亿参数，最大可生成8192个token	Palm 2 3400亿参数 最大可生成1000个token
数据源	基于大量的文本数据集进行训练，包括Common Crawl、Wikipedia、书籍、文章、文档和从开放互联网上抓取的内容；免费的GPT-3.5模型的源代码在2021年结束，但ChatGPT Plus用户可以使用GPT-4执行网络搜索	Bard在Infiniset上接受训练，Infiniset是一个数据集，包括Common Crawl、Wikipedia、文档、网络对话；Bard还可以执行网络搜索
定价	免费；ChatGPT Plus每月收费20美元	免费
访问途径	通过OpenAI网站访问	通过Google Bard网站访问（使用Google账户）

# 人工智能聊天机器人——ChatGPT、Bard

对比项目	ChatGPT	Bard
用户体验	在单个文本提示中创建内容	创建更多信息块
用途	优势在于生成内容，可用于创作	优势在于文本分析
数据存储和隐私	存储输入的所有提示和查询，用户可以查看以前的对话，这也意味着聊天机器人也会查看这些对话。虽然可以删除特定的对话，但该信息可能已被提取至包含在用于聊天机器人训练的大型语言模型中。这意味着输入个人数据或专有信息时可能会出现隐私问题。	Google将用户的Bard 对话在其Google帐户中存储最多18个月，但用户可以在其 Bard活动设置中将其更改为3或36个月。
相同任务指令下生成内容差异	具体不同如下	

资料来源：ChatGPT, Bard, Wikipedia, 华安证券研究所整理

## • 创意产生

**prompt:** 给我五个热门技术文章的标题和想法。

图表74: ChatGPT和Bard回答对比

Give me five titles and ideas for trending technology articles

- "Quantum Computing: Unleashing the Power of Qubits in Everyday Life"**
  - This article explores how quantum computing is transitioning from a theoretical concept to practical applications, from cryptography to drug discovery.
- "The Rise of Metaverse: Building Digital Worlds Beyond VR"**
  - Dive into the evolving landscape of the metaverse, examining how it's reshaping entertainment, social interactions, and business opportunities beyond virtual reality.
- "5G and Beyond: How Next-Gen Connectivity is Transforming Industries"**
  - Discuss the impact of 5G technology on various sectors such as healthcare, manufacturing, and smart cities, and explore what the future holds with 6G on the horizon.
- "AI in Healthcare: From Diagnostics to Drug Discovery"**
  - Highlight the role of artificial intelligence in revolutionizing healthcare, showcasing AI-driven innovations in diagnostics, treatment personalization, and drug development.
- "Sustainable Tech: Innovations Driving a Greener Future"**
  - Explore the latest eco-friendly technologies and their potential to address climate

Give me five titles and ideas for trending technology articles.

Here are five titles and ideas for trending technology articles in September 2023:

- Title: Quantum Computing: The Future of Innovation**

Idea: This article could explore the potential of quantum computing to revolutionize many industries, from medicine to materials science to finance. It could discuss the latest advances in quantum computing hardware and software, as well as the challenges that still need to be overcome before quantum computing can be widely adopted.
- Title: The Metaverse: A New World of Possibilities**

Idea: This article could explore the potential of the metaverse to transform the way we live, work, and play. It could discuss the latest developments in metaverse technology, as well as the challenges and opportunities that lie ahead.
- Title: Artificial Intelligence: The Good, the Bad, and the Future**

Idea: This article could explore the latest advancements in artificial intelligence, from natural language processing to computer vision. It could discuss the ethical implications of AI, as well as the potential for AI to solve some of the world's most pressing problems.

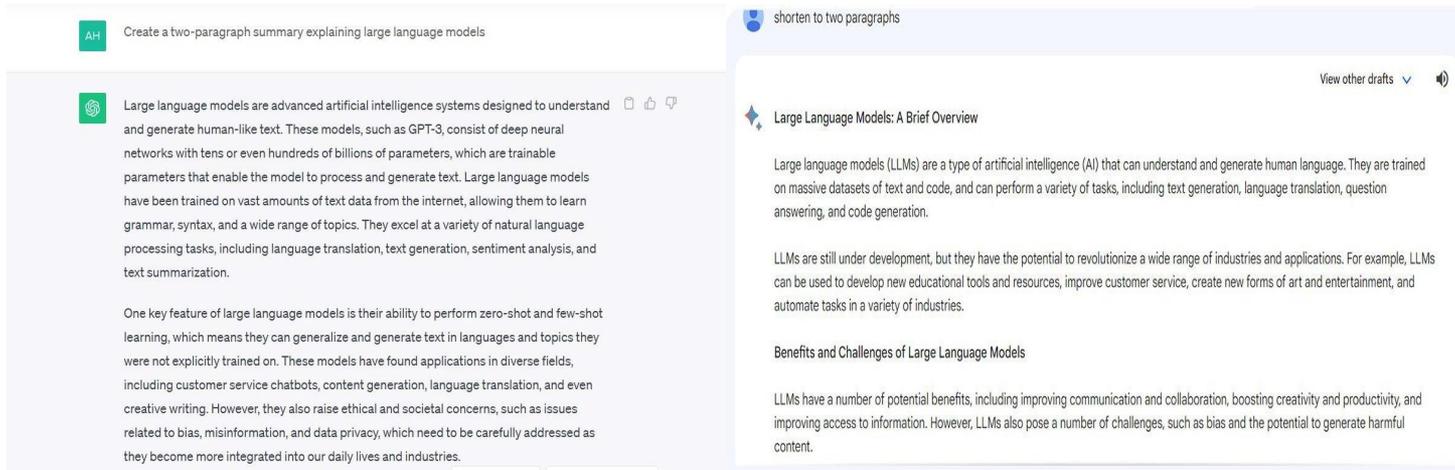
左为ChatGPT回答，右为Bard回答。Bard 对文章角度提供了更多见解，而 ChatGPT 提供了一些不同的主题。由于 ChatGPT3.5 仅限于 2021 年之前的数据，因此查找趋势信息可能很困难。

# 人工智能聊天机器人——ChatGPT、Bard

## • 创建内容

**prompt:** 创建一个两段式的摘要来解释大型语言模型。

图表75: ChatGPT和Bard回答对比

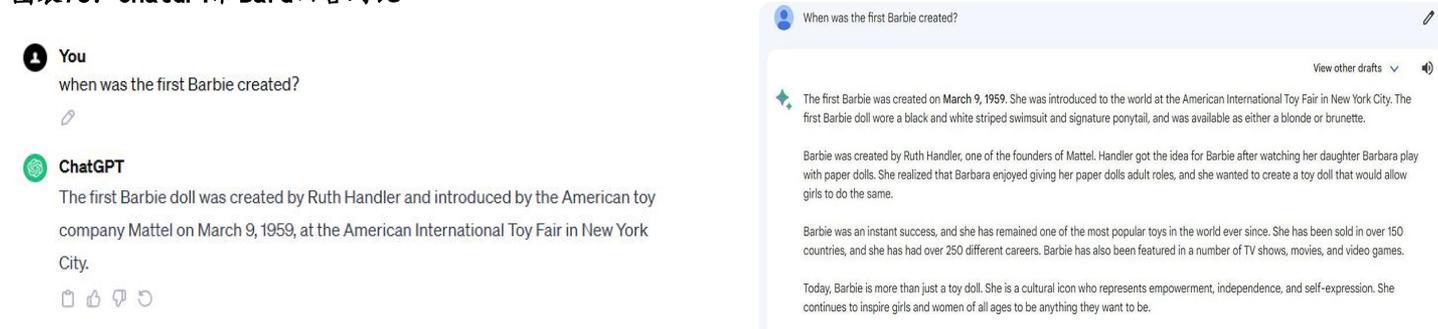


左为ChatGPT回答，右为Bard回答。Bard需要额外的提示才能将其缩短为两段。ChatGPT的语言更全面、更容易理解。

## • 研究

**prompt:** 第一个芭比娃娃是什么时候诞生的？

图表76: ChatGPT和Bard回答对比



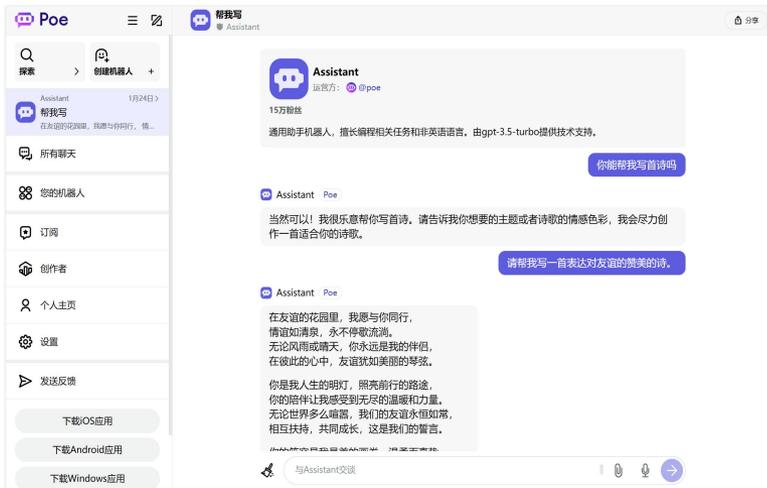
左为ChatGPT回答，右为Bard回答。Bard的回答更容易理解。为了回答问题，日期被加粗。ChatGPT具有相同的信息，但文本粗糙，需要用户自行寻找答案。

# 人工智能聊天机器人——Poe

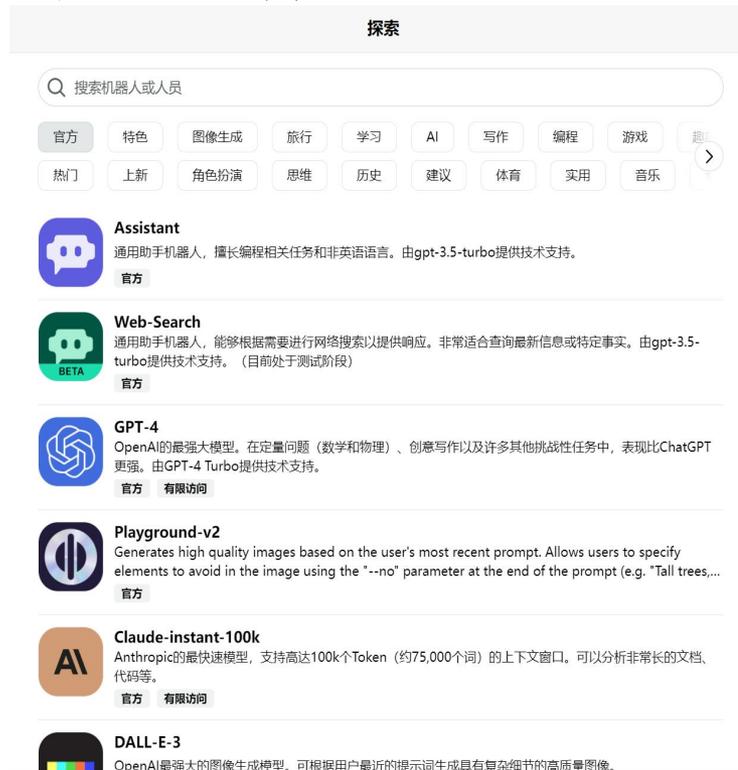
• Poe是由Quora公司推出的对话式人工智能聊天机器人。2023年2月，Poe向公众开放，3月公司发布了集成GPT-4和Claude+的最新版本。Poe的全称是“开放探索平台”，目前要求iOS用户通过电话号码和电子邮件注册并验证账户。Poe搭载多款聊天机器人，每款机器人都有自己擅长的领域，用户可以根据自己的需求，体验不同机器人。例如，Sage擅长处理语言相关的问题如创作文章、翻译等，而Claude擅长对话。ChatGPT、Sage、Dragonfly等均由OpenAI提供支持，而Claude则由Anthropic技术支持。目前Poe已搭载最新GPT-4模型，用户每天可免费询问GPT-4一次，Claude+三次。也可付费使用，订阅价格每月19.99美元，每年199.99美元。

• 在降低AI机器人的交互门槛之外，Poe还囊括一系列社交组件，允许用户创建个人资料，并关注其他用户。用户可以发布机器人输出的内容以供关注者访问，其他人可以点赞支持，或者转发聊天机器人的问答会话。

图表77: Poe聊天界面



图表78: 机器人搜索界面



## 在线协作工具——Notion

• Notion是一款高度灵活且强大的协作工具，集笔记、数据库、任务管理、项目管理等多种功能于一身，可以帮助用户快速搜索、编辑、分享和管理知识和信息。可以将资源、笔记、任务、日历等数据统一管理，并以非常美观的方式进行展示。也可以精确地记录每一步操作，轻松地查看历史记录，有效地管理个人和团队的知识数据。此外，Notion还可以让你与他人共享文档，实现协作和创新。可以邀请好友加入Notion并共同编辑文档，以及在评论中快速反馈和交流想法。

• Notion AI是一个基于GPT-3的自然语言生成系统。主要功能如下：

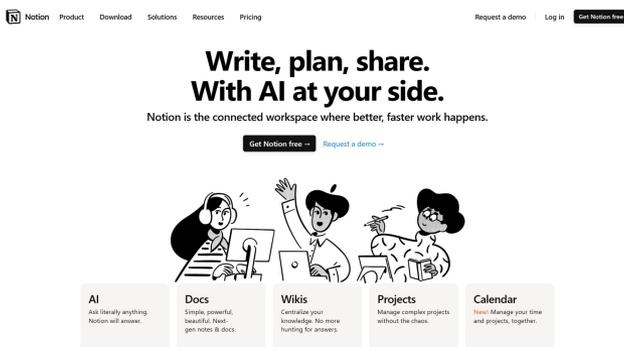
a. 自动文本生成:NotionAI可以根据用户的搜索信息自动生成文本，减少用户输入的时间。

b. 内容推荐:基于用户的搜索历史，NotionAI可以提供更加相关的内容推荐，从而更好地满足用户需求。

c. 改写修正、翻译等辅助工作:改写文字、快速修正语法错误、翻译和解释词语，达到加速工作流程、笔记与文字创作。

• Notion AI可以免费使用20次，目前的定价是每月10刀，可以无限次使用，年付打八折。

图表79: Notion主界面



图表80: NotionAI界面

## Just ask Notion AI.

Knowledge, answers, ideas. One click away.

Get started

Sign up for Q&A Waitlist



### Get answers

Just ask Q&A, and find the info you need in seconds.



### Write better

Get help writing and brainstorming in Notion, not in a separate browser tab.



### Autofill tables

Turn overwhelming data into clear, actionable information in seconds.

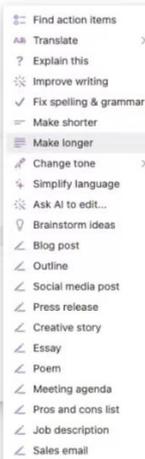


# 在线协作工具——Notion

图表81: NotionAI扩写功能示例

## 关于 Notion AI

- 简介
- 功能
- 应用
- 优点
- 缺点
- 发展前景和趋势



## 关于 Notion AI

- 简介
- 功能
- 应用
- 优点
- 缺点
- 发展前景和趋势

简介: 本文旨在介绍产品的特点和功能, 以及它如何应用于不同的场景。此外, 我们还将探讨它的优缺点, 并展望它未来的发展前景和趋势。

功能: 该产品拥有众多高级功能, 如... (在此添加更多的详细信息, 例如实现这些功能所需的技术和软件)

应用: 该产品可以应用于许多场景, 例如... (在此添加更多的示例, 说明该产品的通用性)

优点: 该产品的优点包括... (在此添加更多的详细信息, 例如它如何提高工作效率和效率)

缺点: 虽然该产品有很多优点, 但它也有一些缺点, 例如... (在此添加更多的详细信息, 例如它的限制和可能的局限性)

发展前景和趋势: 在未来, 我们预计该产品将继续发展和改进, 例如... (在此添加更多的预测和趋势, 例如该产品将如何满足不断变化的市场需求)

Ask AI what to do next...

- Replace selection
- Insert below
- Continue writing
- Make longer
- Try again
- Discard

图表82: NotionAI翻译功能示例

## 关于 Notion AI

- 简介
- 功能
- 应用
- 优点
- 缺点
- 发展前景和趋势

Ask AI to edit or generate...

- Fix spelling & grammar
- Make shorter
- Make longer
- Change tone
- Change language

- English
- Korean
- Chinese
- Japanese
- Spanish
- Russian
- French
- Portuguese
- Italian
- Dutch
- Indonesian
- Filipino
- Vietnamese

Generate from selection

- Summarize
- Translate
- Explain this
- Find action items

## 关于 Notion AI

- 简介
- 功能
- 应用
- 优点
- 缺点
- 发展前景和趋势

- Introduction
- Functionality
- Applications
- Advantages
- Disadvantages
- Future Development and Trends

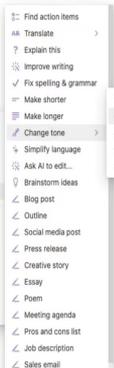
Ask AI what to do next...

- Replace selection
- Insert below
- Continue writing
- Make longer
- Try again
- Discard

图表83: NotionAI改变语气功能示例

## 关于 Notion AI

- 简介
- 功能
- 应用
- 优点
- 缺点
- 发展前景和趋势



## 关于 Notion AI

- 简介
- 功能
- 应用
- 优点
- 缺点
- 发展前景和趋势

介绍-  
有操作作用呢  
哪些场景可以用呢  
优点有哪些  
缺点有哪些, 心塞  
再看法玩儿子的前途和趋势提-

Ask AI what to do next...

- Replace selection
- Insert below
- Continue writing
- Make longer
- Try again
- Discard

图表84: NotionAI语言润色功能示例

本文将详细介绍以下六个方面:

- 简介: 首先, 我们将提供更详细的背景信息, 以便更好地了解我们正在讨论的主题。
- 功能: 接下来, 我们将详细介绍该主题的各种功能和特点, 以便于读者全面了解。
- 应用: 我们将提供一些实际应用案例, 以便读者更好地理解如何使用这个主题。
- 优点: 我们将详细描述这个主题的各种优点, 以便读者更好地了解它的价值。
- 缺点: 然后, 我们将探讨其缺点和挑战, 让读者更全面地了解这个主题。
- 发展前景和趋势: 最后, 我们将展望该主题的未来发展前景和趋势, 以便读者更好地了解它的重要性和未来价值。

Ask AI to edit or generate...

- Improve writing
- Fix spelling & grammar
- Make shorter
- Make longer
- Change tone
- Simplify language

Generate from selection

- Summarize
- Translate
- Explain this

# AI搜索引擎——Bing

• Bing 是微软公司于2009年5月28日推出的一个搜索服务平台，具有每日首页美图、全球搜索与英文搜索、全球搜图等功能。

图表85: Bing发展历程

01

2023年2月，微软宣布推出由ChatGPT支持的最新版本Bing(必应)搜索引擎，并在2月7日正式发布。

2023年2月22日，微软宣布将基于AI聊天的搜索体验扩展到必应的移动版本。

02

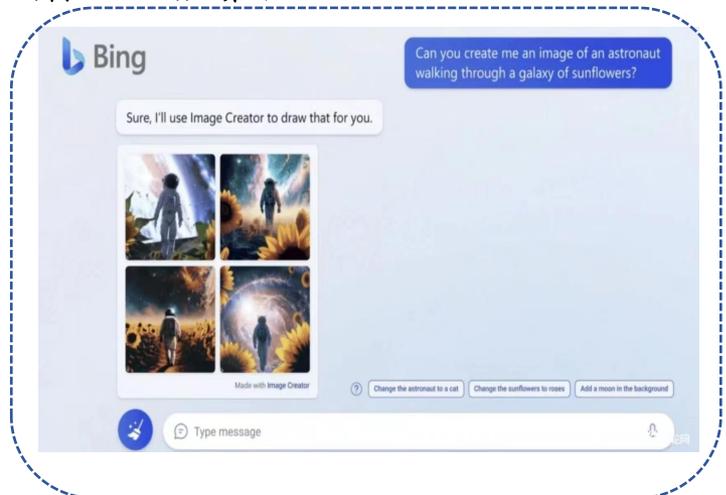
03

2023年10月3日，微软宣布，OpenAI最新的DALL-E3模型面向所有Bing Chat和Bing Image Creator用户开放。DALL-E3是OpenAI图像生成模型的第三个版本。

图表86: AI聊天界面



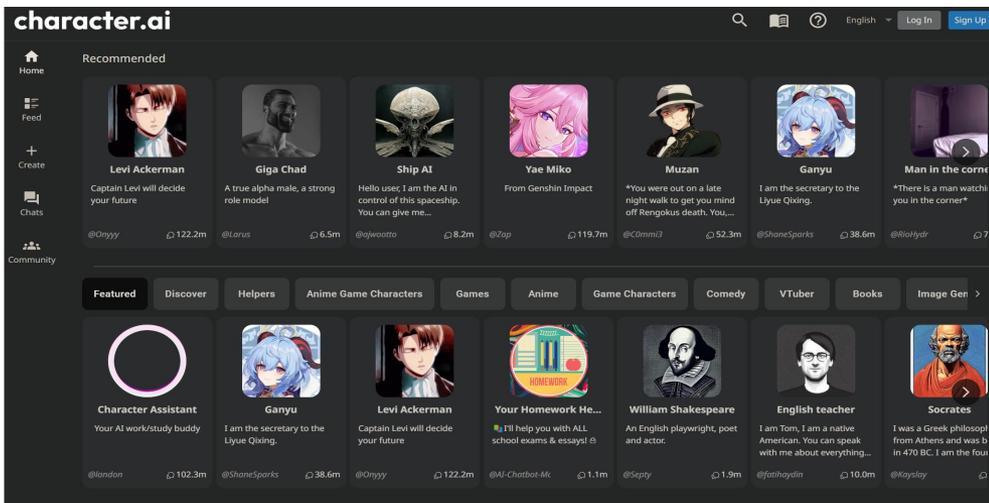
图表87: AI绘画界面



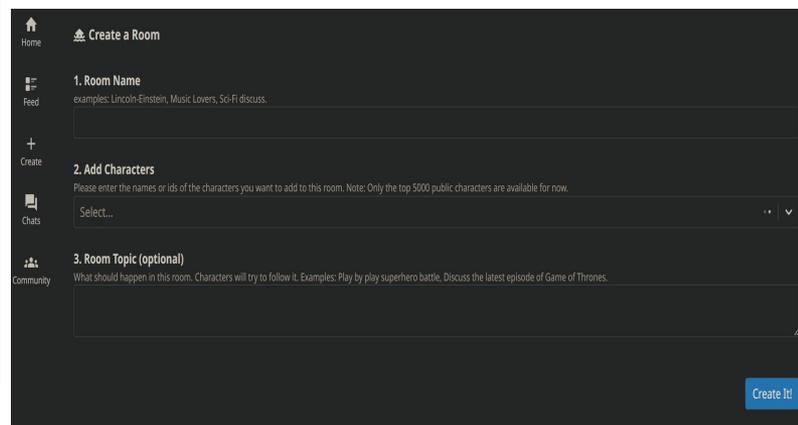
# 角色聊天机器人——Character.ai

• Character.ai 是一款基于 LLM 的聊天机器人网站，由 Noam Shazeer 和 Daniel De Freitas 于 2022 年 9 月创建。该网站已经预先创建了许多聊天角色，每个聊天机器人都有自己的角色设定，有不同的聊天风格，如 Elon Mask、Steve Jobs 和 马里奥 等，用户可以与这些角色进行交流。用户还可以自己创作角色。Character.ai 注重人格属性，试图满足社交、情感、陪伴、支持等需求。此外，Character.ai 还支持创建房间，多人可以使用不同的角色进行聊天。

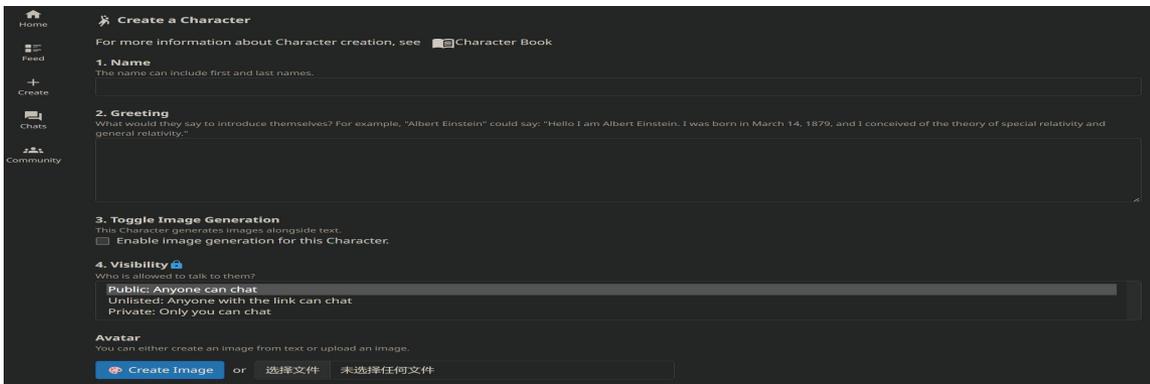
图表88: Character.ai 主聊天界面



图表89: 创建聊天房间界面



图表90: 创建专属角色界面



## 重要声明

### 分析师声明

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的执业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，本报告所采用的数据和信息均来自市场公开信息，本人对这些信息的准确性或完整性不做任何保证，也不保证所包含的信息和建议不会发生任何变更。报告中的信息和意见仅供参考。本人过去不曾与、现在不与、未来也将不会因本报告中的具体推荐意见或观点而直接或间接收任何形式的补偿，分析结论不受任何第三方的授意或影响，特此声明。

### 免责声明

华安证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。本报告中的信息均来源于合规渠道，华安证券研究所力求准确、可靠，但对这些信息的准确性及完整性均不做任何保证，据此投资，责任自负。本报告不构成个人投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需要。客户应考虑本报告中的任何意见或建议是否符合其特定状况。华安证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

本报告仅向特定客户传送，未经华安证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。如欲引用或转载本文内容，务必联络华安证券研究所并获得许可，并需注明出处为华安证券研究所，且不得对本文进行有悖原意的引用和删改。如未经本公司授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。本公司并保留追究其法律责任的权利。

### 投资评级说明

以本报告发布之日起6个月内，证券（或行业指数）相对于同期沪深300指数的涨跌幅为标准，定义如下：

#### 行业评级体系

增持：未来6个月的投资收益率领先沪深300指数5%以上；

中性：未来6个月的投资收益率与沪深300指数的变动幅度相差-5%至5%；

减持：未来6个月的投资收益率落后沪深300指数5%以上；

#### 公司评级体系

买入：未来6-12个月的投资收益率领先市场基准指数15%以上；

增持：未来6-12个月的投资收益率领先市场基准指数5%至15%；

中性：未来6-12个月的投资收益率与市场基准指数的变动幅度相差-5%至5%；

减持：未来6-12个月的投资收益率落后市场基准指数5%至15%；

卖出：未来6-12个月的投资收益率落后市场基准指数15%以上

无评级：因无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使无法给出明确的投资评级。市场基准指数为沪深300指数。