

QBITAI ANNUAL AI TOP 10 REPORT 2024

2024年度AI十大趋势报告

量子位

insights

量子位智库出品 2024
QbitAI Insights

序言

从今年起，量子位智库做了一个改变——

在过往三年连续围绕整个前沿科技来提供年度科技趋势参考后，今年聚焦在了AI。

过去，AI是前沿科技主轴上的一支核心技术。现如今，AI正在吞噬整个世界、整个产业、全部赛道。

没错，AI已经完全成为了科技发展主旋律。

尽管有诸多类比，如一开始类比互联网革命，到后来类比电力（第二次工业）革命，再到现如今——最宏大的一种说法是：第二次地球文明革命。

所以不论如何，可见AI正在展开的冲击、带来的影响力是如何全面又深刻，甚至一度被以科幻的方式谈论它。

AI当然不是科幻，AI首先是一门科学，其次是一项工程，最后正在变成一种工业。

这就意味着AI不仅可以观测、可以学习，还能预测——或者更准确来说，我们就是站在新进展新信息的最前沿，站在产学研交汇地带，把正在从学术研究进入产业变革程序的技术方向，总结并举例说明呈现给所有人。

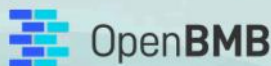
在今年，我们还通过更具规模的数据统计，在AI应用的创新创业和投资方面，也提供了结论性参考，希望对整个产业提供第三方视角下的助益。

总之，希望这份年度趋势报告，能够对你在岁末年初了解时代技术进展提供最直接的帮助。身处技术大航海时代，即便你未能出海探索，也希望你能感知到潮水涌动的方向。

ACKNOWLEDGEMENT

特别鸣谢

科技合作伙伴



*以上排名无先后顺序



目录

2024年度十大AI趋势 -----

PART 1 / 技术篇

趋势一 大模型创新：架构优化加速涌现，融合迭代大势所趋 ----- /02

趋势二 Scaling Law泛化：推理能力成皇冠明珠，倒逼计算和数据变革 ----- /10

趋势三 AGI探索：视频生成点燃世界模型，空间智能统一虚拟和现实 ----- /16

PART 2 / 产品篇

趋势四 AI应用格局：第一轮洗牌结束，聚焦20赛道5大场景 ----- /25

趋势五 AI应用竞争：多领域竞速运营大于技术，AI助手兵家必争 ----- /35

趋势六 AI应用增长：AI+X赋能类产品大干快上，原生AI爆款难求 ----- /45

趋势七 AI产品趋势：多模态上马，Agent席卷一切，高度个性化呼之欲出 ----- /50

PART 3 / 行业篇

趋势八 AI智变千行百业：左手变革生产力，右手重塑行业生态 ----- /54

趋势九 AI行业渗透率：数据基础决定初速度，用户需求成为加速度 ----- /75

趋势十 AI创投：投融资马太效应明显，国家队出手频率提升 ----- /78

千行百业AI优秀落地方案推荐 ----- /87

结语 -----

01

TECHNOLOGY

趋势一 大模型创新：架构优化加速涌现，融合迭代大势所趋

- 1. 架构层创新助力解决算力瓶颈现实问题 /02
- 2. 创新混合架构挑战Transformer垄断 /02

技术原理——

- 1. 路径一：循环神经网络及其变种（以RWKV为代表） /03
- 2. 路径二：状态空间模型（以Mamba为代表） /05
- 3. 路径三：层次化卷积模型（以UniRepLKNet为代表） /06
- 4. 路径四：多尺度保持机制模型（以RetNet为代表） /07
- 5. 路径五：液体神经网络模型（以LFM为代表） /08

趋势二 Scaling Law泛化：推理能力成皇冠明珠，倒逼计算和数据变革

技术原理——

- 1. AI模型的性能强烈依赖于规模 /10
- 2. 全新的Scaling Law /13

趋势三 AGI探索：视频生成点燃世界模型，空间智能统一虚拟和现实

技术原理——

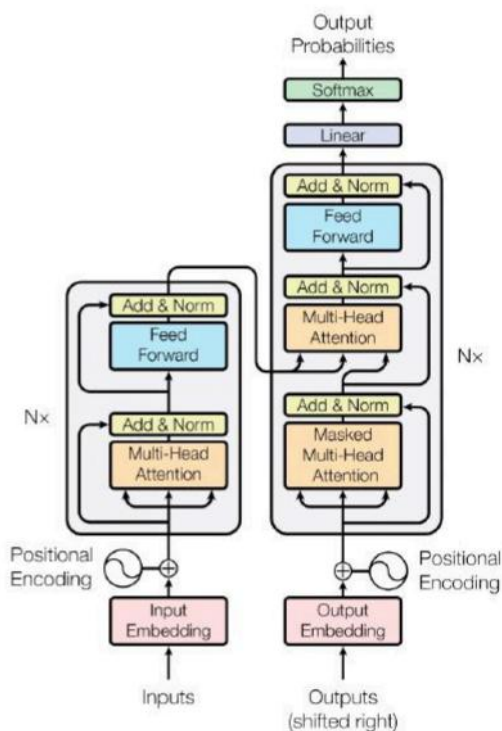
- 1. 视频生成：从扩散模型出发 /16
- 2. 世界模型：从自动驾驶领域到整个世界 /19
- 3. 具身智能：回到现实世界，回到产业链 /20
- 4. 空间智能：连结具身智能与空间计算 /22

一、大模型创新：架构优化加速涌现，融合迭代大势所趋

1. 架构层创新助力解决算力瓶颈现实问题

Transformer架构是目前应用最广泛的主流大模型架构，而自注意力机制（Self-Attention, SA）则是Transformer架构的核心——它允许模型进行并行计算，在序列中非线性地直接捕捉任意两个位置之间的关联权重，大幅提高模型能力上限。但另一方面，这也使模型的算力需求、计算复杂性和消耗资源成本都随参数增加呈指数级增长，在大规模任务中快速触达天花板。

2024年以来，随着大模型参数量的飞速规模化以及训练与部署的深入落地，Transformer架构的上述弊端愈发显著，成为助推全球性算力紧缺的重要因素，也为大模型的端侧落地提出了挑战。为寻求突破，对大模型架构的创新性探索逐渐成为不容忽视的趋势。



图：Transformer模型架构，谷歌、多伦多大学

若能突破Transformer在算力和数据需求方面的限制，新架构有望在自然语言处理和计算机视觉领域引发新一轮技术革新。

——明势创投

2. 创新混合架构挑战Transformer垄断

自2017年Attention Is All You Need出世提出Transformer架构以来，7年已过。AI行业对Transformer的路径依赖引发了越来越多的“过时”争论，体现出日渐迫切的架构创新需求。

2023年以来，RWKV和Mamba引起热议，多种新架构加速涌现，世界范围内的学者从多个方向努力，试图在保留Transformer架构优势的基础上创新性引入其他架构特点，解决算力开销问题，Transformer的绝对统治地位得到挑战，兼采众家之长的混合模型（Hybrid）已成未来趋势。

Transformer架构、Next-Token Prediction和Scaling Law是当前大模型的算法基石，但这些领域也越来越需要新的突破，以构建强大且高效的新一代基础大模型。强大意味着卓越的性能、泛化能力和抵抗幻觉能力；高效则指低成本、高效率 and 低能耗。只有具备这两大特质，人工智能才能真正成为水和电一样的基础设施。

——微软亚洲研究院

Hyena Hierarchy	2023.04	Michael Poli, Stefano Massaroli, Eric Nguyen, Daniel Y. Fu, Tri Dao, Stephen Baccus, Yoshua Bengio, Stefano Ermon, Christopher Ré	Stanford University, Mila and Université de Montréal
RWKV	2023.05	Bo Peng, Eric Alcaide, Quentin Anthony等	RWKV元始智能
RetNet	2023.07	Yutao Sun, Li Dong, Shaohan Huang, Shuming Ma, Yuqing Xia, Jilong Xue, Jianyong Wang, Furu Wei	微软亚洲研究院
KAN	2024.04	Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljac̆ic, Thomas Y. Hou, Max Tegmark	MIT, Caltech, Northeastern University, IAFI
TimeMixer	2024.05	Shiyu Wang, Haixu Wu, Xiaoming Shi, Tengge Hu, Huakun Luo, Lintao Ma, James Y. Zhang, Jun Zhou	蚂蚁集团, 清华大学
Mamba-2	2024.05	Albert Gu, Tri Dao	普林斯顿、卡内基梅隆大学
Falcon Mamba	2024.08	Jingwei Zuo, Maksim Velikanov, Rhaiem, Ilyas Chahed, Younes Belkada, Guillaume Kunsch	阿布扎比 Technology Innovation Institute (TII)
LFM	2024.09	Ramin Hasani, Mathias Lechner, Alexander Amini, Daniela Rus	Liquid AI

这些新兴大模型架构不仅在性能上可以与Transformer模型竞争，还在内存效率和可扩展性上展现出优势。梅花创投杨颜媛表示，部分新架构更易于进行并行计算，能够充分利用现代硬件的并行计算能力，提高训练和推理的速度。

它们的出现，为AI领域带来了新的活力，也为未来的研究和应用开辟了新的可能性。随着这些模型的不断发展和优化，我们有理由相信，大模型创新架构将在AI未来发展中扮演越来越重要的角色。

技术原理

1. 路径一：循环神经网络及其变种（以RWKV为代表）

循环神经网络（RNN）通过循环方式处理序列数据，能够对过去的输入保留记忆，但存在难以并行化的问题，Transformer架构的诞生最早就是为弥补这一缺陷。但仍有很多学者认为，RNN的潜力还远未达到天花板，在Transformer架构越来越受到诟病的今天，RNN凭借其独特优势再度获得了越来越多学者的探索创新。

目前这一路径的架构创新主要使用循环神经网络（RNN）替代自注意力机制，通过循环方式处理序列数据，使模型对过去的输入保留记忆。

元始智能

RWKV「联想记忆法」

区别于Transformer的Query - Key - Value参数，RWKV架构由四个重要参数组成：R、W、K、V，除了可训练的权重参数w（Weight），RWKV还使用r（Receptance）参数来控制对信息的接受程度。

RWKV与Transformer架构的本质区别在于背后的记忆机制，与Transofrmer的内存寻址机制相比，RWKV更像是一种联想记忆方法。



Transformer = Addressing Memory 寻址记忆:

Address	Value
0x00	01001010
0x01	10111010
0x02	01011111
0x03	00100100
0x04	01000100
0x05	10100000
0x06	01110100
0x07	01101111
0x08	10111011
...	...
0xFE	11011110
0xFF	10111011

Key

RWKV = Associative Memory 联想记忆:

简化版联想记忆

$$S_t = k_0 v_0 + \dots + k_t v_t$$



简化版回忆

$$v_0 = k_0^T S_t = k_0^T k_0 v_0 + \dots + k_0^T k_t v_t$$



内积为1

正交

我相信RNN是正确的，但现在的RNN远远没有做到它真正的水平，它的上限其实是非常高的，现在我们还远远没有到那个地步，还有很多空间。因为RNN更接近人脑和宇宙的运作方式。例如，在物理上，宇宙的下一状态只与上一状态有关，这是所谓的locality和causality，量子场论遵循这一原则。

——彭博，RWKV作者

(1) RWKV核心思想

RWKV (Recurrent Weighted Key-Value) 模型核心思想是**将RNN的循环结构与Transformer的并行计算能力相结合，在实现高效推理、节省存储开销的同时保持模型的高性能**。这使得RWKV可以“**像Transformer一样**”进行并行训练，同时在推理阶段可以以递归形式进行解码，“**像RNN一样**”推理。

(2) RWKV的创新点

- **Token Shift**: token shift在时间混合计算中，通过对当前和前一输入的线性组合进行线性投影，生成用于时间混合的向量；在通道混合计算中，也采用类似的方法生成通道混合输入向量
- **WKV运算符**: WKV运算符利用时间衰减因子对权重进行更新，使得每个时间步的输出依赖于之前所有时间步的信息，从而保留了RNN的记忆能力，这种设计使得RWKV模型在保持较低计算复杂度的同时，能有效捕捉序列数据中的长期依赖关系

- **输出门控：**RWKV通过在时间混合和通道混合块中使用sigmoid函数对接收向量进行门控，控制信息的流动和记忆更新，确保在每个时间步只传递和处理相关信息，从而减少梯度消失和爆炸问题，增强了模型的稳定性和训练效率

(3) RWKV的发展与应用

RWKV自提出以来已经经历了多次版本迭代，最新版本RWKV-7预览版已在今年9月正式发布。

2. 路径二：状态空间模型（以Mamba为代表）

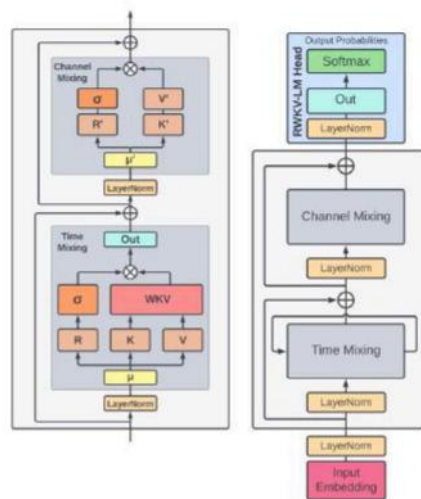
状态空间模型可以看作是循环神经网络（RNN）和卷积神经网络（CNN）的融合，由其发展而来的结构化的状态空间序列模型（SSM）是另一颇具潜力的大模型创新架构代表。这类模型利用状态空间处理长序列问题，通过循环或卷积运算实现高效计算，使得计算开销与序列长度呈线性或近线性关系，从而显著降低计算成本。

(1) Mamba首次提出

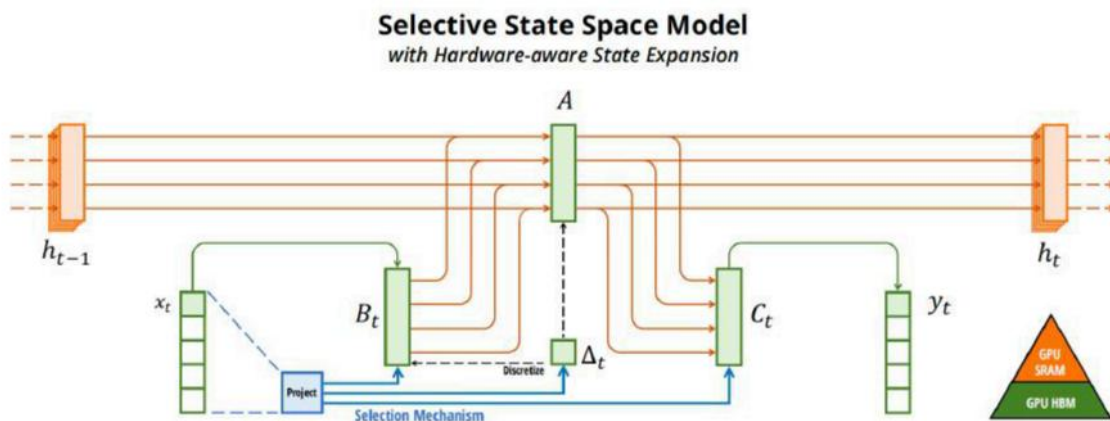
2023年12月，Mamba架构首次被提出，引入了选择性状态空间模型，实现了对输入数据的有选择性处理。这种选择机制使得模型能够根据当前输入的token决定哪些信息是重要的，忽略不相关的信息，提升模型处理长序列的能力和推理吞吐量，达到Transformer模型的五倍。

(2) Mamba核心思想

Mamba的核心在于其硬件感知算法，利用现代硬件（如GPU）的内存层次结构，通过扫描而非卷积计算模型，减少不同级别GPU内存间的IO访问，提高计算效率。此外，Mamba简化了深度序列模型设计，无需注意力机制或多层感知器（MLP）块，使模型更加简洁。



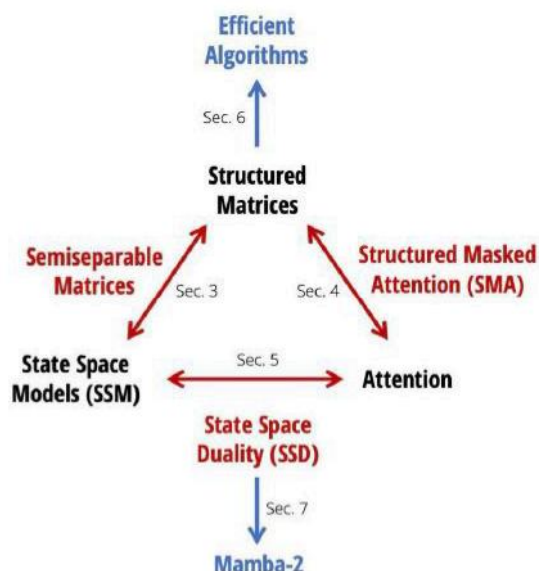
图：RWKV内部架构图，彭博等



图：选择性状态空间模型（Selective State Space Model），卡内基梅隆、普林斯顿大学

(3) Mamba-2核心思想

今年5月，Mamba-2发布，提出了**状态空间对偶（SSD）框架**，揭示了状态空间模型与结构化掩码注意力之间的联系。Mamba-2的核心层通过引入新的SSD算法，在训练效率上提升了2-8倍，同时保持了与Transformer在语言建模方面的相似水平竞争力。



图：（结构化状态空间对偶性）用结构化矩阵阐明状态空间模型和注意力之间的关系，普林斯顿、卡内基梅隆大学

(4) Mamba-2创新点

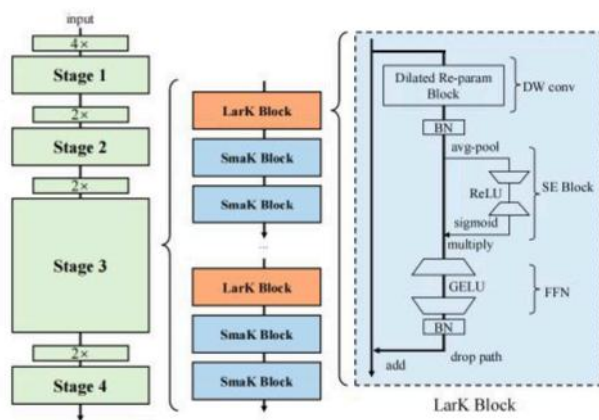
- **硬件友好设计**：Mamba-2的另一个重要贡献是其对硬件友好的设计，允许使用更大的状态维度，提高训练速度。在处理需要更大状态容量的任务，如多查询关联回忆（MQAR）任务时，Mamba-2显示出比Mamba-1显著的性能提升
- **混合模型的探索**：Mamba-2还探索了将注意力层与SSM层结合的混合模型，发现适量的注意力层可以进一步提升模型性能

3. 路径三：层次化卷积模型（以UniRepLKNet为代表）

层次化卷积模型，以UniRepLKNet为代表，是一种长卷积架构，其核心在于使用与输入序列长度相当或接近的滤波器（核）来捕捉序列数据中的长距离依赖关系。这种设计使得模型在处理输入时能够考虑到更远的历史信息，有效处理长序列问题。

(1) UniRepLKNet创新点

UniRepLKNet采用大核CNN，能够处理多种模态数据，如图像、音频、时序预测等。该模型提出了“局部结构设计、重参数化、核大小选择和随深度拓展添加小核”四条指导原则来设计大核CNN架构，并采用硬件感知的并行算法，在图像识别任务中实现领先性能，并在音频、视频、点云和时间序列等多模态任务中也取得了显著的结果。UniRepLKNet的优势在于充分利用大核卷积神经网络的特点，通过创新的架构设计，有效地解决了感受野、特征抽象层次以及模型深度表示能力等核心问题。



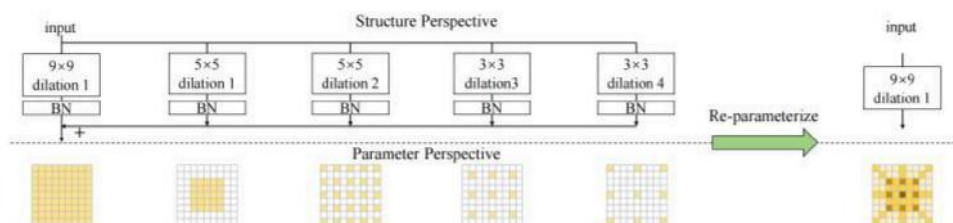
图：UniRepLKNet的架构设计，腾讯AI Lab、香港中文大学

(2) 大核卷积的优势

大核卷积的独特优势在于**不依赖深度堆叠即可获得大感受野，避免了深度增加带来的边际递减问题**。UniRepLKNet提出了一种**膨胀重参数块**（Dilated Reparam Block），即通过使用多个膨胀小核卷积层来增强一个大核卷积层，从而在不增加推理成本的情况下提高性能，该块的公式如下：

$$W' = \text{conv_transpose2d}(W, I, \text{stride} = r).$$

膨胀重参数块使用膨胀的小核卷积层来增强非膨胀的大核层。这样的膨胀层相当于具有更大稀疏核的非膨胀卷积层，因此整个块可以等效地转换为单个大核卷积。

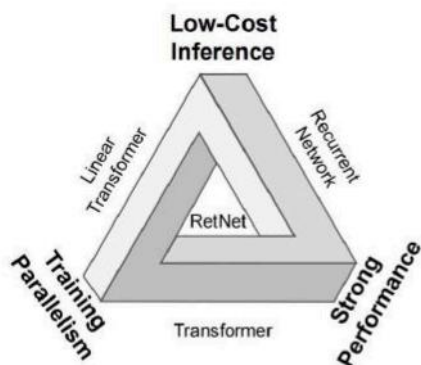


图：扩张重参数块转化为转换为单个大核卷积，腾讯AI Lab、香港中文大学

4. 路径四：多尺度保持机制模型（以RetNet为代表）

多尺度保持机制模型（Retentive Network, Retnet）由微软亚洲研究院提出，克服传统Transformer架构在效率和性能上的局限性。

RetNet的设计突破了所谓的“不可能三角”，在保持训练并行性的同时，实现了低成本部署以及良好性能。



图：Retnet突破不可能三角，微软亚洲研究院

作为全新的神经网络架构，RetNet同时实现了良好的扩展结果、并行训练、低成本部署和高效推理。这些特性将使RetNet有可能成为继Transformer之后大语言模型基础网络架构的有力继承者。

——韦福如，微软亚洲研究院全球研究合伙人

(1) RetNet创新点

RetNet作为全新的神经网络架构，使用多尺度保持（Retention）机制替代了标准的自注意力机制。与标准自注意力机制相比，保持机制有以下特点：

- **引入位置相关的指数衰减项取代softmax**：这一改进简化了计算过程，并使得前步信息以衰减的形式得以保留，从而在降低计算复杂度的同时，保留了序列中的重要信息
- **引入复数空间表达位置信息**：这一方法取代了传统的绝对或相对位置编码，使得模型更容易转换为递归形式，增强了模型对序列中元素之间相对位置的感知能力

- **使用多尺度的衰减率**：保持机制采用了多尺度的衰减率，使RetNet能够更灵活地处理不同长度的序列，提升了模型对序列数据的适应性和表达能力
- **利用GroupNorm的缩放不变性提高Retention层的数值精度**：这一特性使得在GroupNorm中乘以一个标量值不会影响输出和反向梯度，从而在保持模型性能的同时，提高了计算的稳定性和效率

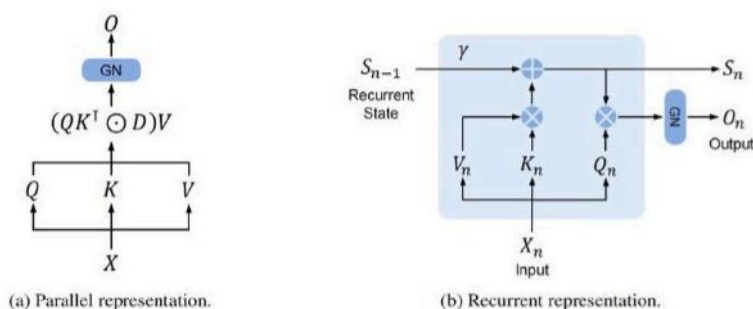


Figure 3: Dual form of RetNet. "GN" is short for GroupNorm.

图：RetNet的双重表示，微软亚洲研究院

(2) 序列表示形式

- **并行**：并行表示使RetNet可以像Transformer一样高效地利用GPU进行并行训练
- **递归**：递归表征实现了 $O(1)$ ¹的推理复杂度，降低了内存占用和延迟
- **分块递归**：即并行表示和递归表示的混合形式，将输入序列划分为块，在块内按照并行表示进行计算，在块间遵循递归表示。分块递归表征则能够高效处理长序列数据

RetNet的实验结果表明，它在语言建模任务上达到了与Transformer相当的性能，同时在推理速度和内存占用方面显著优于Transformer。

5. 路径五：液体神经网络模型（以LFM为代表）

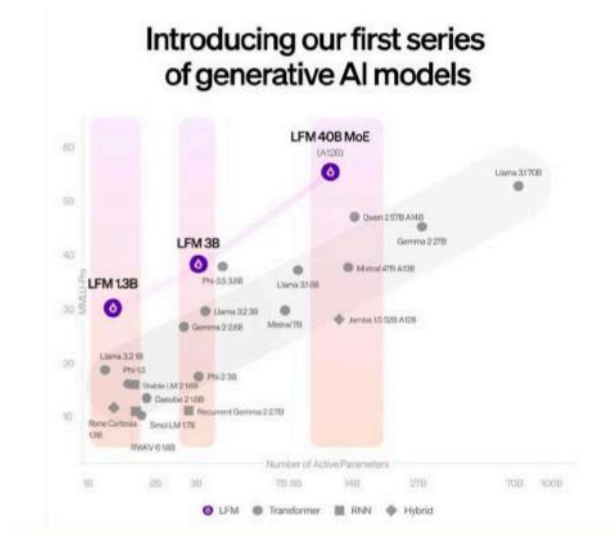
液体神经网络模型LFM，由MIT系初创公司Liquid AI团队发布，是基于液体神经网络（Liquid Neural Networks, LNN）架构的新型神经网络设计。受生物神经系统启发，LNN通过使用更少的神经元和创新的数学公式，实现了与传统深度学习模型相比相同甚至更优的性能。

(1) LFM创新点

- **高效内存使用**：LFM的核心优势在于其高效的内存使用和强大的推理能力，这使得它在处理大量顺序数据时表现出色。与传统的基于Transformer的模型相比，LFM在处理长输入时内存使用量显著减少，这是因为LFM能够更有效地利用上下文长度，KV缓存不随序列长度增加而线性增长
- **多模态能力**：LFM模型能够处理包括视频、音频、文本、时间序列和信号在内的多种类型的数据
- **自适应能力**：LFM还具有自适应能力，可以根据特定的硬件平台或参数要求进行自动优化

¹ $O(N)$ 表示第N步的推理复杂度

- **结构化操作单元：**这些单元是模型的基本组成部分，属于一个全新的设计空间。LFM的设计允许模型在扩展、推理、对齐和模型分析方面进行优化。



图： LFM在MMLU-pro的测试集成绩对比， Liquid AI

这些加速涌现的创新架构多是在不同程度保留Transformer架构优势基础上，结合循环神经网络（RNN）、状态空间序列模型（SSM）、卷积模型（CNN）以及液体神经网络（LNN）等思想所做出的创新发展，这使得不同架构之间的界限越来越模糊，**呈现出日益明显的混合（Hybrid）趋势**，更多性能领先的创新架构具备“博采众家之长”的特点。同时，目前大多数创新架构虽然能够在一些小规模性能测试中实现与传统Transformer架构同等甚至更优越的性能，**但在实际应用层面，仍然有待工业界通过多个超大参数模型作进一步验证。**

二、Scaling Law泛化：推理能力成皇冠明珠，倒逼计算和数据变革

2020年，伴随着GPT-3问世，第一代Scaling Law指引我们在参数量、数据集和计算量之间寻找模型性能的最优解；2024年，拥有强大推理能力的o1让我们对大模型的要求从迅速生成与训练结果，转向在推理过程中进行更深度的思考。最初，我们对Scaling law的关注局限于对参数量的思考，甚至引发了对其失效的担忧；但全新的scaling law从狭窄具体的指代衍生到宽泛的概念，并引发了我们对万卡集群、合成数据和计算资源最优分配的多重关注。

我们对Scaling Law及其泛化在当前AI时代下的关注可被归纳为以下几点：

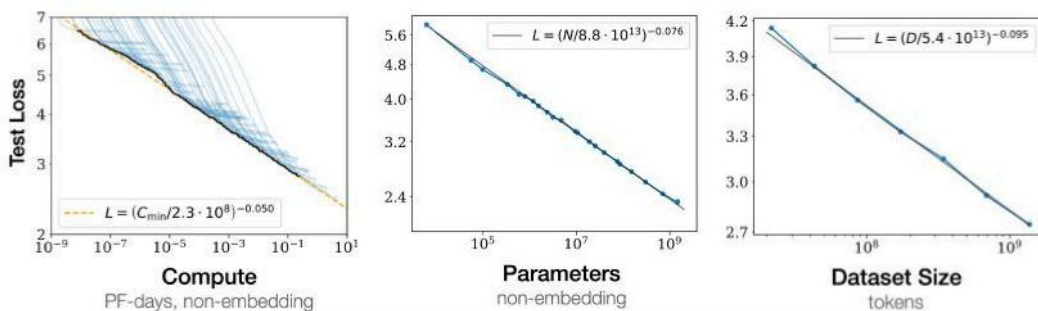
- 参数量与计算量膨胀下，带动万卡集群以及高性能网络的建设和发展。
- 数据耗尽危机中，合理善用合成数据成为较优选择。
- 新时代的Scaling Law出现，大模型的发展将追求更高的推理能力，资源向Post-training和推理算力倾斜。

技术原理

我们将从影响两个模型时代的两条Scaling Law出发来探讨这个问题：

1. AI模型的性能强烈依赖于规模，主要和参数量N、数据集大小D和训练计算量C有关

AI模型领域的Scaling Law正式被提出要追溯到2020年，OpenAI在论文Scaling Laws for Neural Language Models中总结到，他们发现损失函数（loss）随着模型大小、数据集大小和训练使用的计算量而呈幂律变化。而模型的性能在一定范围内受其他超参数（如depth、width）的影响很少。



图：损失函数与训练计算量、参数量和数据集的幂律关系，OpenAI

在Transformer模型中，Scaling Law的数学表示通常为：

$$\text{Performance} \sim C \cdot N^{-\alpha}$$

Performance：模型性能（误差率、准确率等）

C：计算量

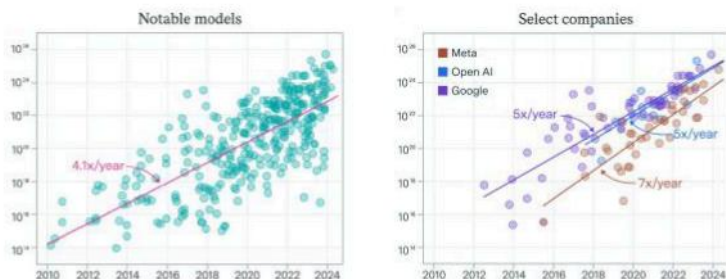
N：模型规模（参数量、数据量等）

α ：标度指标（性能随规模变化的表现）

(1) Parameters & Compute：万卡集群及高性能网络建设的必要性

2020年惊艳亮相的GPT-3是一个175B参数的大模型，而今天我们讨论的模型参数量已经是万亿级别。EPOCH AI维护

着一个从20世纪50年代就开始监测AI和机器学习模型的数据库，他们发现从2010年到2024年5月，用于训练模型的计算量每年增长4-5倍，而过去10年间语言模型的性能提升中约有三分之二是由于模型规模的增加。



横轴: Publication data
纵轴: Training compute(FLOP)

图：2010-2024前沿AI模型训练计算量年增长，EPOCH AI

例如GPT-4需要使用2.5万张英伟达A100 GPU，并行训练100天的左右时间，在此期间要处理13万亿个token，并且涉及大约1.76万亿个参数。可以预见的是为了追求更极致的性能和解决更多问题，在参数量的增加和计算需求的增长下，万卡集群及高性能网络的建设是必要的。

- **万卡集群**：由一万张及以上的加速卡（如GPU、TPU或其他专用AI加速芯片）组成的高性能计算系统
- **高性能网络**：网络需要支持高带宽、低延迟的数据传输，确保集群中的计算节点能够高效地通信和协作

中国联通研究院夏璠表示，高性能网络基础设施对于支持大规模AI应用至关重要。随着数据量的增加和AI应用的扩展，对网络带宽和延迟的要求也越来越高。这种集群将充分整合高性能GPU计算、高性能RDMA网络、高性能并行文件存储、智算平台等关键技术，将底层基础设施整合成为一台超级计算机，以千亿级甚至万亿级参数规模的大模型训练。

全球数据中心网络市场预计将在未来五年内以约 10% 的年复合增长率增长。

——明势创投

中国万卡集群进展：

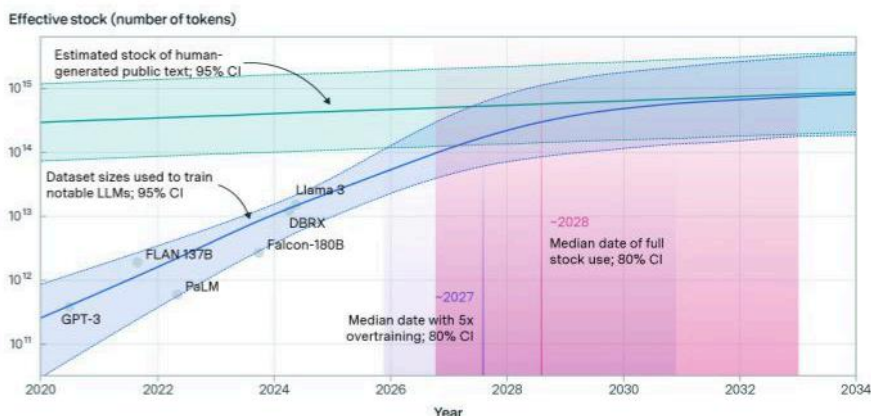
运营商	中国电信	天翼云上海临港国产万卡算力池
	中国移动	呼和浩特万卡智算中心
	中国联通	上海临港国际云数据中心
企业	腾讯	算力集群HCC、高性能网络星脉
	字节跳动	12288卡Ampere架构训练集群
	百度	万卡AI集群
	阿里巴巴	阿里云10万卡量级集群
	商汤	SenseCore商汤大装置
	蚂蚁集团	万卡异构算力集群

企业	华为	昇腾AI万卡集群
	科大讯飞	「飞星一号」超万卡集群算力平台

尽管万卡集群必然会强调对加速卡的需求，我们在技术上关注的却应该是集群的互联。从千卡到万卡，单卡性能下降8%左右，目前全球范围内已经有企业开始追求「10万卡」以上的集群规模，在万卡到10万卡的难关中，如何能保持集群的线性度、稳定性，同时保证调度的效率，是下一步所有AI玩家将持续关注的重点。

(2) Dataset: 合成数据是数据危机的解药吗？

同样，在Scaling Law中数据集的规模和表现也会改变模型的性能。现有的开源大模型在方法、架构层面开源的同时，几乎很少涉及到训练数据集的公开。



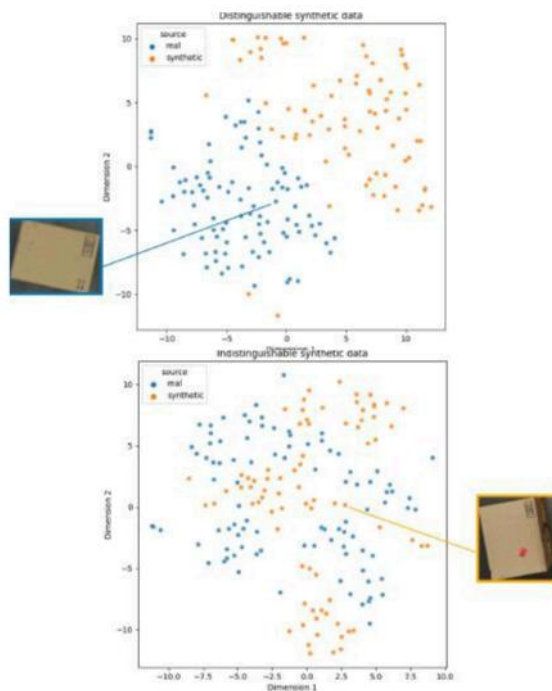
图：公开文本和数据使用量的预测，EPOCH AI

我们必须意识到一个问题，也许有一天我们会耗尽现有的数据。有估计称现有人类公开文本存量约为300万亿个token。按照预测，大语言模型有可能在**2026年-2032年之间耗尽这一存量**。如果过度训练，这一时间点将会再度提前。

还值得我们关注的一个问题是，**数据的质量也许比数量更为重要**。例如Apple在研发模型时采用了非常规的3阶段预训练方法，在核心预训练后降低网络爬取的低质量数据权重。国内在意识到高质量数据集的重要性后，多次发布对建设自有高质量数据集的指导意见。

合成数据（Synthetic data）则是我们从AI1.0时代就开始寻找的破解之道，在具身智能和自动驾驶等前沿领域得到过部分运用，如今已被广泛引入大模型的训练之中。

合成数据是指通过算法或模型生成的人为数据，模拟真实数据的统计特性和分布，但不包含任何真实信息。实际训练中，我们通常会选择部分或全部地使用合成数据。明势创投表示，合成数据的应用突破了传统的数据获取和使用方式，为解决数据匮乏和隐私问题提供了创新思路，具有较高的创新性和颠覆性。



图：合成数据生成方法，Amazon

合成数据的生成方法：

- **统计分布**：分析真实数据以确定其潜在的统计分布，例如正态分布、指数分布或卡方分布。然后从这些已识别分布中生成合成样本，以创建在统计学上与原始数据集相似的数据集
- **基于模型**：训练机器学习模型以理解和复制真实数据的特征
- **深度学习**：使用深度学习的方式生成高质量的合成数据集。这些方法通常用于更复杂的数据类型，例如图像或时间序列数据

但合成数据是否是应对数据危机的最佳解药仍值得商榷。首先，真实数据永远存在相当多的离群值和异常值，合成数据生成算法很少能重现它们。其次，合成数据的引入可能导致模型在训练过程中出现过拟合，从而影响其在现实世界中的表现。

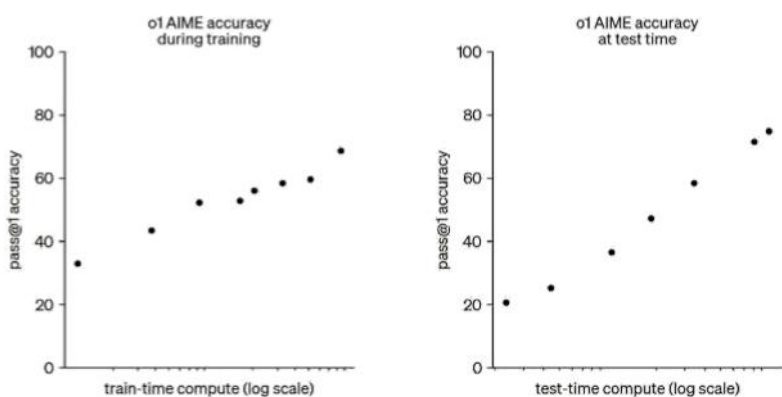
2024年发表在Nature上的一篇研究AI models collapse when trained on recursively generated data提到他们所完成的一次实验，**使用合成数据9次迭代后就会让模型原地「崩溃」**。更吸引人眼球的结果来自Dohmatob, Elvis等的实验结果Strong Model Collapse，他们发现即使**合成数据仅仅占到总数据集的最小部分，甚至是1%的比例，仍然可能导致模型崩溃**。

目前生成和使用合成数据的方法在不同模型中都处在探索阶段，我们将持续关注这一议题的最终结果。

2. 全新的Scaling Law，赋予模型的推理时间越多，模型的推理能力越强

来到2024年，OpenAI有关o1模型的论文On The Planning Abilities of OpenAI's o1 Models: Feasibility, Optimality, and Generalizability提出了一条全新的Scaling Law：**你赋予模型的推理时间越多，模型的推理能力就越强**。

尽管没有披露具体的训练方法，但其原理是基于强化学习的内化思维链学习，通过思维链式的问题拆解，使得模型可以不断验证和纠错。



图：o1在新Scaling Law下的推理表现，OpenAI

当AI拥有更强大的推理能力，不仅意味着在编码、搜索等任务上能够帮助用户解决更多问题，也会在数学、生物、化学等基础科学领域，帮助实现突破性的发现。

——高榕创投

(1) 大语言模型推理增强（思维链-思维树-思维图）

这几年，我们确实观测到规模的增加带来了基本的推理能力，但这种推理的效果是相对有限的，而o1所完成的，是让模型在作答前进行“停下来思考”，这一过程需要更多的计算资源，因此被称为“推理时计算”。

停下来思考这一状态的实现，正是我们一直想教会模型的模式，即更直接地进行推理。

o1大幅度增强了大模型的逻辑推理能力，使大模型从系统1提升到系统2，大幅度加速AI Agent落地可能性。

——王晟，英诺天使基金合伙人

以o1为代表的Inference scaling law使得模型推理能力大幅提升，在GPQA这样的人类专家benchmark上已经超过人类PhD水平。推理能力大幅提升可以显著减少幻觉，提高Agent类人物的能力。

——戴雨森，真格基金管理合伙人

大语言模型的推理能力通常涉及到几种不同的技术或概念，包括思维链（Chain of Thought, CoT）、思维树（Tree of Thought, ToT）和思维图（Map of Thought, MoT）。

- **思维链（Chain of Thought, CoT）**：模仿了人类解决问题时的思维过程，即先分析问题，然后逐步推理，最后得出结论
- **思维树（Tree of Thought, ToT）**：更为结构化的推理方法，它将问题分解为多个分支，每个分支代表一个可能的推理路径或假设
- **思维图（Map of Thought, MoT）**：视觉化工具，用来表示不同概念、事实和推理步骤之间的联系

(2) 后训练（Post-training）的转变

从前的预训练遵循一个已经被深刻理解的法则：你在预训练模型上投入的计算资源和数据越多，它的表现就越好。

但新的Scaling Law—**Post-training Scaling Law**的出现已经引发行业对于算力分配、后训练能力的重新思考。

后训练通常包括微调（Fine-tuning）和对齐（Alignment）。

新的模型想要提升能力，除了在预训练阶段，还可以在后训练阶段，通过提升强化学习（Reinforcement Learning）训练的探索时间和增加模型推理的思考时间来实现性能的提升。

o1激发了当前时代的大模型开发者对提升大模型数学和推理能力的兴趣，这很大程度上就意味着扩大后训练的规模。



图：后训练的比重增加

最近的研究使用了不同的方法把结果的奖励信号传递到大模型的各个推理过程中：

Apple Intelligence Foundation	Post-training: SFT+RLHF 其中，Apple为RLHF阶段引入两种新方法：①与iTeC一起进行拒绝采样微调，目的是训练、优化奖励模型；②镜像下降与RLHF结合，这被认为是比PPO（近端策略优化）更有效的方法
Gemma 2	Post-training: SFT+RLHF 其中，其中他们用于RLHF的模型比目标模型大十倍，使用 WARP (Weight Averaging for Reward Prediction)对目标模型进行平均
Llama 3.1	Post-training: SFT+拒绝采样+DPO Meta AI认为RL算法（如带有PPO的RLHF）稳定性差、扩展难度大，但他们在DPO后仍开发了奖励模型，并将模型平均技术在每一轮训练中都作用于奖励模型、SFT和DPO
Qwen2	2阶段后训练方法：① SFT ，提高模型在预设场景中的响应准确性；② DPO ，使模型与人类偏好保持一致（他们称这环节为RLHF） Qwen2的Alignment也分为2部分： DPO （离线）+ 拒绝采样 （在线）

拒绝采样似乎已成为后训练阶段的常见做法，但现阶段并没有在选择DPO或RLHF方面达成共识。然而强化学习的思路贯穿始终，Meta仍为Llama的后训练开发了奖励模型。

模仿学习+强化学习成为典型AI模型发展路径范式。本质上是数据的稀缺与否及质量高低的问题，数据多就优先模仿学习，数据少就优先强化学习（可自我产生数据），最终是综合两种类型的算法技术以产生更强大的能力。

——陈石，峰瑞资本合伙人

如今，我们意识到RLHF起的作用远超已经IFT(Instruction Fine-Tuning)，需要将更多的预算放在后训练阶段上。

（3）推理算力vs训练算力？

随着我们对大模型推理能力的更高追求，和对后训练阶段的更多投入，未来有可能推理端的算力将会超过训练端。尽管当前推理算力还没有超过训练算力，但推理成本已经超过了训练成本。毕竟训练成本是一次性成本，而推理需求将随着模型复杂度、用户数量和每用户查询次数的增加而继续增大。明势创投表示，在大规模部署AI服务的背景下，推理成本成为主要开支来源，市场对降低推理成本的需求强烈。

值得一提的是，Scaling Law并非是严格的自然法则，而是人类基于经验观察得出的趋势性规律，它的价值往往在**对未来提供方向性的指导**。新旧Scaling Law的应用和泛化将调整我们对资源配置和技术重心的看法，并创新出更为强大的新一代AI和模型。

三、AGI探索：视频生成点燃世界模型，空间智能统一虚拟和现实

Sora的火热激发了国内外许多模型在视频生成领域的应用，产出了在影视、营销等行业具有可落地商业价值的AI应用；“首个视频模型”之争也引发了自动驾驶领域基于自身优势，对世界模型的研究和探索；而能够带来更好的自动驾驶体验的世界模型，事实上也能够作用于机器人领域；引发诸多讨论的空间智能概念，与我们一直探索的空间计算、具身智能密不可分……

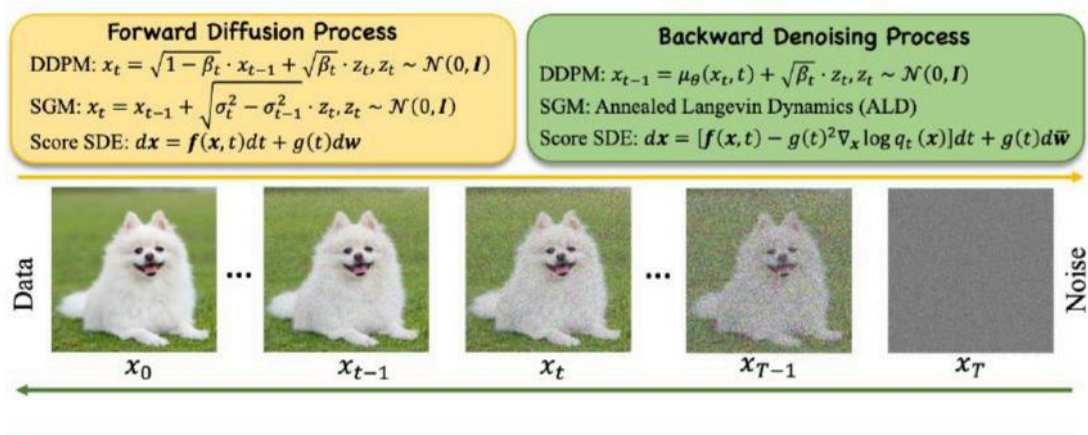
在AI改变我们的生活、真正触摸复杂世界的过程中，人类也从未停止对AGI的探索。

技术原理

1. 视频生成：从扩散模型出发

扩散模型目前已经在图像生成、图像超分辨率和图像编辑等任务中取得了显著成果，并逐渐取代了基于生成对抗网络（GAN）和变分自编码器（VAE）的方法，在视频合成任务中展现出巨大潜力，成为目前视频生成的主流技术路径。

扩散模型（Diffusion Model）是一类概率生成模型，主要通过学习逆向过程来生成数据。

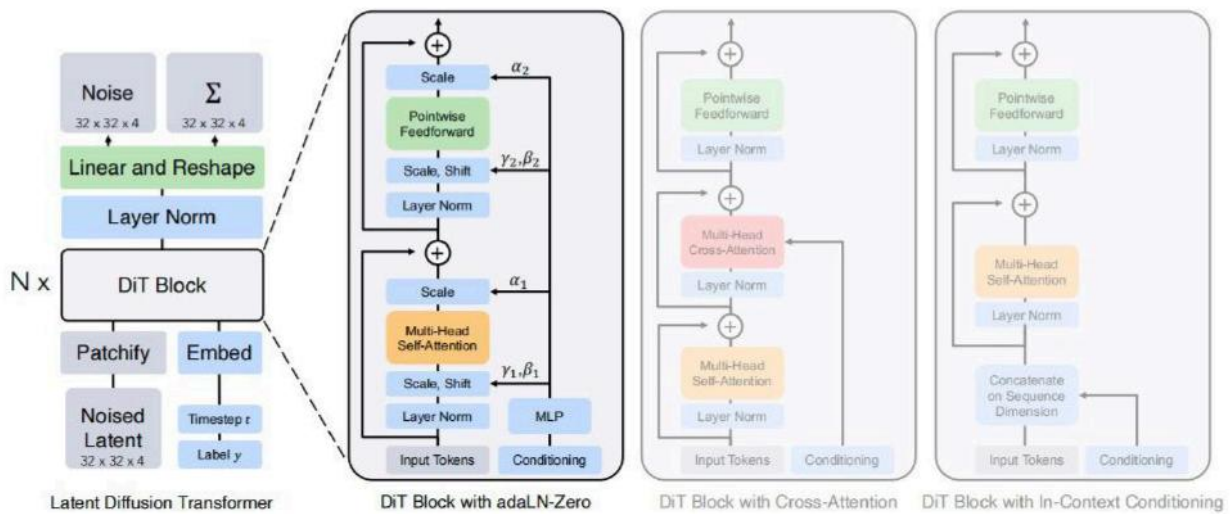


图：A Survey on Video Diffusion Models，上海市智能视觉计算协同创新中心、复旦大学

本轮视频生成热潮中，最受瞩目的是将Transformer架构引入到扩散模型中的DiT（Diffusion Transformer）模型，极大地提高了图像生成的质量。

(1) DiT模型

DiT的核心思想是利用Transformer的强大建模能力，替换传统扩散模型中常用的U-Net结构，从而在潜在空间上操作，生成高质量的图像。其图像生成流程是基于一种混合的方法，使用卷积VAEs和基于Transformer的DDPM(Denoising Diffusion Probabilistic Models)，重点是训练图像的DDPM，特别是图像的空间表示。



图：Scalable Diffusion Models with Transformers, UC Berkeley、New York University

DiT的技术原理基于以下几个关键点：

扩散模型	DiT建立在扩散模型的基础上，这是一种生成模型，通过模拟一个从数据分布到噪声分布的逐步过程，学习逆向过程来生成数据。在DiT中，这个逆向过程由Transformer来实现
Transformer 架构	Transformer以其自注意力机制而闻名，能够捕捉序列数据中的长距离依赖关系。在DiT中，图像被分割成多个patches，这些补丁被嵌入到一个序列中，然后通过Transformer块进行处理
潜在补丁	DiT不是直接在像素空间上操作，而是在潜在空间上操作。这是通过将图像编码到一个较低维度的潜在表示来实现的，这个表示被分割成多个补丁，每个补丁作为一个序列元素输入到Transformer中
条件生成	DiT支持条件生成，这意味着它可以在生成过程中使用额外的信息，如类别标签。这些条件信息被嵌入到Transformer的输入序列中，以引导生成过程
可扩展性	DiT的设计允许其通过增加模型的深度、宽度或输入补丁的数量来扩展。这种可扩展性使得DiT能够通过增加计算资源来提高生成图像的质量
训练效率	DiT在训练过程中表现出了高效率。由于Transformer的并行化能力，DiT可以利用现代硬件（如TPU）进行高效训练

DiT的提出，不仅推动了生成模型技术的发展，也为未来的研究提供了新的方向。随着研究的深入，DiT或其变体已经被应用于视频生成，以及艺术创作、游戏开发和虚拟现实等领域。随着模型规模的增加和训练技术的改进，DiT有望在未来实现更加惊人的生成效果。

(2) 其他混合架构模型创新

在视频生成领域，除了DiT之外，还有其他几种混合架构取得了显著的进展。这些架构结合了不同的模型特性，以提高视频生成的质量和效率。

Transformer + 3D卷积网络	一些研究工作探索了将Transformer与3D卷积网络结合的架构。3D卷积网络擅长捕捉视频的时空特征，而Transformer能够处理长距离依赖关系。这种混合架构能够同时利用两者的优势，提高视频生成的连贯性和真实感
自回归模型 + Transformer	自回归模型如PixelRNN或PixelCNN在像素级别上生成数据，而Transformer则能够处理序列数据。将两者结合可以生成高质量的视频内容，同时保持对长距离依赖的有效建模
VAE+ 扩散模型	在DiT的基础上，学者们进一步探索了将变分自编码器与扩散模型结合的架构。这种混合模型首先使用VAE将高分辨率的视频编码到低维潜在空间，然后在潜在空间上应用扩散模型进行生成，从而在保持高生成质量的同时减少计算成本
GAN+ 扩散模型	GAN在图像生成领域取得了巨大成功，而扩散模型则在样本质量上展现出了优势。将两者结合的混合架构旨在结合GAN的稳定性和扩散模型的高质量样本生成能力
TimeSformer	TimeSformer是一种专门为视频理解任务设计的Transformer架构，它通过将时间维度融入到标准的Transformer架构中，有效地处理视频数据。这种架构可以用于视频生成任务，生成具有时空一致性的内容
交叉注意力机制	在Transformer架构中引入交叉注意力机制，可以使模型在生成视频时更好地利用额外的上下文信息，如文本描述或对象标签，从而生成与这些上下文信息更加一致的视频内容

这些混合架构的共同目标是在保持视频生成质量的同时，提高模型的计算效率和灵活性。随着研究的不断深入，我们可以预见未来将会出现更多创新的混合架构。

我们再次回到引爆「视频生成」热潮的模型Sora。OpenAI在介绍中称Sora是“World Simulator”，而在Sora上线的第一时间埃隆·马斯克即表示，Tesla大约在一年前即能生成符合物理学的有关真实世界的模型（Real-world video generation with accurate physics），这一讨论又将世界模型（World Model）的概念引入大众视线。

OpenAI

Sora「模拟真实世界物理状态」

Sora通过结合扩散模型与Transformer模型，能够模拟真实世界中的人、动物和环境。这种能力使其在文本到视频（T2V）领域中独树一帜，能够根据文本提示生成逼真的视频内容。

时间长度与时序一致性： Sora能够生成长达60秒的视频，且在维持视频内容的时间连续性和逻辑一致性方面表现出色，这在以往的文生视频模型中是难以实现的。



真实世界物理状态模拟： Sora展示了一系列引人注目的涌现能力，这些行为不是预设的规则，而是模型通过学习大量数据后自然涌现的能力。

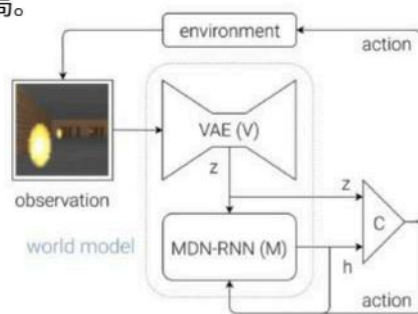
2. 世界模型：从自动驾驶领域到整个世界

AI在现实世界的应用正在从虚拟世界逐渐扩散至物理世界。2024年谈到世界模型更多是在智能驾驶领域，本年度国内领先的量产智能驾驶玩家蔚来、理想等都提出了有关世界模型的设想和布局。

(1) 世界模型（World Model）

论文Recurrent World Models Facilitate Policy Evolution(David Ha, Jürgen Schmidhuber, 2018)可以帮助我们理解什么是世界模型：

图中的Vision Model(V)将高维观测转为低维向量，Memory RNN(M)是序列预测下一个时刻的表征，在这样基本的序列预测中增加一个输入action，让数据分布自由变化。



图：世界模型的基本框架

事实上，世界模型反映了长久以来许多AI研究人员对model-based RL工作的无奈和期待：模型不够准确，那么完全在模型里训练的强化学习效果就不好；如果能获得准确的世界模型，就能在世界模型中反复试错，找到现实的最优决策。

由此，我们可以从需求出发来定义世界模型：

完全地理解这个世界就是世界模型要干的事。

——任少卿，蔚来智能驾驶副总裁

(2) 目前在世界模型方向上的工作

正如为何特斯拉会在一年前就开始布局有关真实世界模型的相关工作，「面对数据中没有见过的决策，使用世界模型推理出决策成果」这一要点，和自动驾驶领域实现L5的目标息息相关。当前，世界模型方向上的工作正广泛地采用自动驾驶领域的数据和场景进行实验。

RenderWorld	仅支持视觉的端到端自动驾驶框架，它使用基于自监督高斯的Img2Occ模块生成3D占用标签，然后通过AM-VAE对标签进行编码，并使用世界模型进行预测和规划。RenderWorld采用高斯散射来表示3D场景和渲染2D图像，与基于NeRF的方法相比，大大提高了分割精度并降低了GPU内存消耗
Drive OccWorld	以视觉为中心的4D预测世界模型应用于自动驾驶的端到端规划
BEVWorld	将多模态传感器输入标记为统一紧凑的鸟瞰图（BEV）潜在空间以进行环境建模的新方法。世界模型由两部分组成：多模态标记器和潜在BEV序列扩散模型
LAW	使用LAtent World model，根据预测的自车行为和当前框架的潜在特征来预测未来的潜在特征
OccLLaMA	类似VQVAE的场景标记器，增强了LLM，特别是LLaMA，以对统一词汇表执行下一个token/场景预测，从而完成自动驾驶中的多项任务
DriveDreamer-2	它基于前作（DriveDreamer）的框架，结合了一个大型语言模型来生成用户定义的驾驶视频
WorldDreamer	培养对一般世界物理和运动的全面理解，从而显著增强视频生成的能力；WorldDreamer从大型语言模型的成功中汲取灵感，将世界建模定义为无监督的视觉序列建模挑战

3. 具身智能：回到现实世界，回到产业链

早期的世界模型很多出现在机器人领域，任少卿提到“自动驾驶是泛机器人领域/世界模型的框架也能做机器人”。具身智能是指一种基于物理身体进行感知和行动的智能系统，其通过智能体与环境的交互获取信息、理解问题、做出决策并实现行动，从而产生智能行为和适应性。

Sim2Real(simulation to reality)和在真实世界中采集具身交互的数据并学习是研究具身智能的两条思路。清华智能产业研究院（AIR）表示：将真实环境数字化和仿真器模拟的行为学习结合起来，形成Real2Sim2Real(RSR)的闭环连接，能更有效地利用交互数据和真实环境的反馈，加速智能系统的实际部署与落地。

与去年相比，许多具身智能玩家在2024年都推出了自己的首款人形机器人，但更多玩家开始关注机器人在具体产业链

上能有什么样的位置和发展空间。这不仅意味着与过去一年相比，人形机器人的概念变得十分可触及，同时也表明具身智能这一领域，正在转向对当前人类来说更有价值的落地和应用。

一项明显特征是，在完成了行走、避障等基本功能后，具身智能玩家纷纷开始研发灵巧手，逐个攻克具身智能难题：

腾讯Robotics X 实验室	公布了自研机器人灵巧手TRX-Hand，拥有8个可独立控制关节，自重仅1.16千克，最大持续指尖力可达15牛
星动纪元	XHand灵巧手具有12个主动自由度，采用全驱动方案，能够实现多方向的运动，从而模仿人类手部的精细操作
灵心巧手 Linker Hand	以其20个主动自由度的合理配置，被业界专家广泛认为是具身智能研究的优选方案，平衡了研究需求与成本效益
银河通用	拥有灵巧手解耦关节和灵巧手线性驱动方式专利
傲意科技	灵巧手ROHand具有11个运动关节和6个主动自由度
帕西尼感知科技	开发了市场上首款融合多维触觉与AI视觉双模态能力的四指仿生灵巧手DexH13 GEN2，具有1956颗专业级ITPU多维触觉传感器和800万高清AI手眼相机
智元机器人	推出了高自由度五指灵巧手，包含19个自由度，集成基于MEMS原理的触觉感知和视触觉感知技术

1X

视频生成、世界模型与机器人集群

1X Technologies，2014年于挪威成立，专注于开发和生产能够模仿人类行为的机器人，以增加全球的人类劳动力。



NEO：专为家庭使用而设计的双足人形机器人，1X基于视频生成和自动驾驶世界模型领域的进展，训练出了一个世界模型，作为其机器人的虚拟模拟器。

模型可以理解接触到的物理世界并生成高保真视频，使机器人可以在自己的神经空间中执行规划、评估和模拟操作。

EVE：轮式机器人，2020年即取得商业化进展，为美国商业场所提供安保服务。1X强调其机器人集群协作，展示了一个人通过简单的语音命令指挥多个机器人执行一系列动作的能力。



4. 空间智能：连结具身智能与空间计算

在经历了一整年有关具身智能概念、投融资的热潮后，与其相关又不完全相似的空间智能概念又被李飞飞提起。

空间智能是指机器在三维空间和时间中感知、推理和行动的能力，帮助它理解事物在三维空间中的位置及交互方式。

—李飞飞

在彻底理解空间智能之前，我们可以回顾在2023年度报告中提及的重要科技趋势空间计算：

(1) 空间计算

空间计算是一项将物理世界与数字内容融合的技术体系，根据其涉及的不同层次，可以分为基础设施层、数据层和交互层：

基建层：支持空间计算的硬件和网络基础

- 设备：苹果的Vision Pro提供了高性能的AR显示、光学传感器以及精准的跟踪功能，配备先进摄像头和处理器
- 网络：高速度、低延迟的5G和Wi-Fi 6E网络正加速实时交互的实现，尤其适用于移动设备和协作应用；谷歌云等云服务商正在开发“边缘AI”基础设施，将AI模型更贴近用户，从而减少延迟、提升空间计算应用的实时能力

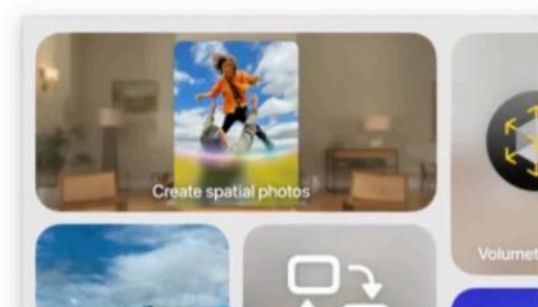
Apple

Apple Vision Pro 「空间计算从概念到现实」

Apple Vision Pro作为苹果公司推出的首款空间计算设备，实现了数字内容无缝融入现实世界，通过其革命性的产品设计和技术创新，开启了空间计算的新时代。



自然输入方式：Vision Pro 采用了一种称为“自然输入”的交互模型，这种模型使用眼睛和手的组合进行输入，用户只需看着目标并点击手指即可进行交互。



空间操作系统 visionOS：Vision Pro 搭载了苹果首款为空间计算所打造的操作系统 visionOS，它包括了全新的实时执行引擎、动态注视点渲染管线以及对原生空间体验的支持。

数据层：负责空间数据的收集、处理和管理

- 空间映射和计算机视觉：实时映射和物理环境的解读，使应用程序能够将虚拟物体置于物理空间中
- 数字孪生：通过集成物理模型、传感器更新、运行历史等数据，在虚拟空间中完成映射

交互层：用户直观地与虚拟和物理世界互动

- 沉浸技术（XR）：主要包括增强现实（AR）、虚拟现实（VR）和混合现实（MR），他们目前仍是空间计算概念落地的核心技术，AR将数字元素叠加在现实世界中，VR创建完全虚拟的空间，MR则是两者的混合；目前具体相关产品存在视频透视（VST）和光学透视（OST）两种技术路径
- 触觉反馈和感官界面：设备的触觉反馈增加了沉浸感，混合现实头戴设备和触感手套能在虚拟环境中创造高度真实的互动
- 语音和手势识别：得益于AI的支持，这些技术允许用户无需动手即可控制数字元素

Rokid

Rokid Glasses 「AI+AR」 无缝融合现实世界



2024年11月18日，Rokid联名暴龙眼镜推出消费级AI+AR眼镜Rokid Glasses。

AI+AR全新交互方式： Rokid Glasses基于衍射光波导成像技术实现AR效果，同时深度整合阿里巴巴旗下通义千问大模型AI能力，融合AI+AR技术，可以通过语音和视觉交互实现物体识别、实时翻译、定位导航、购物支付多种任务，无缝融合现实世界。



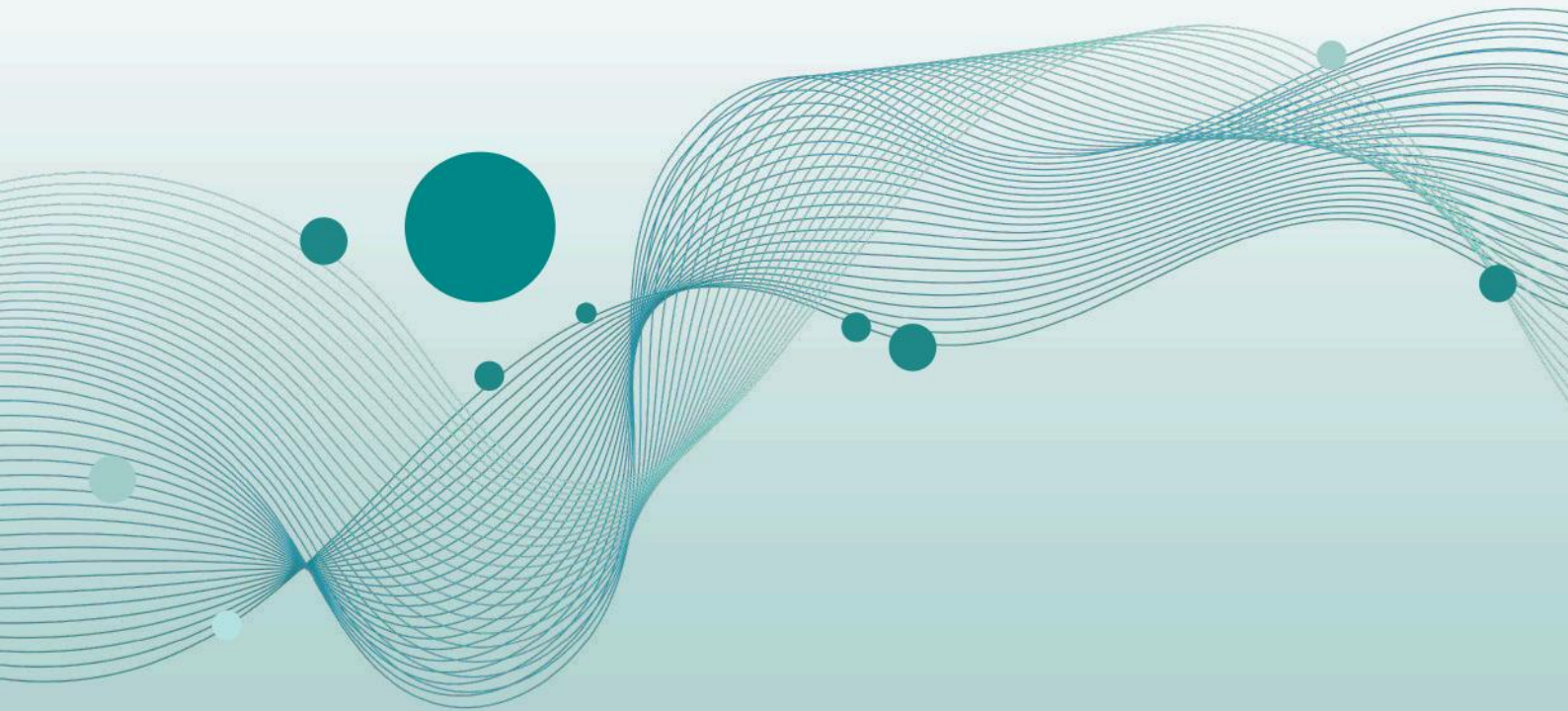
硬件“减负满配”： Rokid Glasses将主板、电池等核心组件集成于框架之中，重量仅49g，提升舒适度及便携性。同时配备一颗1200万像素专业级摄像头，搭载了第一代骁龙®AR1平台，配合2GB RAM+32GB ROM存储组合。

(2) 操作系统：空间智能

苹果发布的Vision Pro就是一个典型的空算设备。李飞飞认为，与其不同的是，空间智能将成为增强现实和虚拟现实的操作系统。通过生成完整可交互的三维世界，真正解锁机器在三维世界中的智能能力。同时，具备三维感知、推理和行动能力的机器又能彻底改变我们与世界的交互方式。

毫无疑问，空间智能的野心正在于将空间计算操控虚拟世界的本领和具身智能触达现实世界的能力结合起来，这将涉及到大量的数据处理、模型训练以及如何在现实场景有效地部署这些技术。

最终，从视频生成到世界模型、从具身智能到空间智能，我们始终都在追寻一个问题——AGI，极致的AI，极致的通用，虚拟与真实世界的合二为一。



趋势四 AI应用格局：第一轮洗牌结束，聚焦20赛道5大场景

- (一) 20+细分赛道 /25
- (二) 五大使用场景 /32

趋势五 AI应用竞争：多领域竞速运营大于技术，AI助手兵家必争

- (一) APP端数据分析 /35
- (二) Web端数据分析 /39
- (三) 量子位智库AI100 /42
 - 1.综合AI100 /43
 - 2.原生AI100 /44

趋势六 AI应用增长：AI+X赋能类产品大于快上，原生AI爆款难求

- (一) 办公软件类AI+X产品 /46
- (二) 内容平台类AI+X产品 /48

趋势七 AI产品趋势：多模态上马，Agent席卷一切，高度个性化呼之欲出

- (一) 多模态 /50
- (二) Agent化 /51
- (三) 高度个性化 /52

02

PRODUCT

中国AI产品全景图

综合类工具

-多功能综合平台



-AI智能助手



办公提效

-AI写作

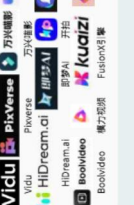


创意设计

-AI设计

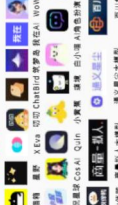


-AI视频

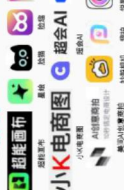


休闲娱乐

-AI陪伴



-AI相机

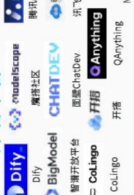


AI软件产品

-AI开发平台



-AI搜索



-AI PPT



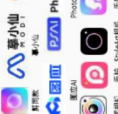
-AI生图



-AI图像编辑



-AI音乐



日常生活

-AI教育



-AI健康



-其他



-AI Agent



-AI搜索



-AI PPT



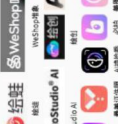
-AI生图



-AI图像编辑



-AI音乐



-AI Agent



-AI搜索



-AI PPT



-AI生图



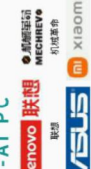
-AI图像编辑



-AI音乐



-AI PC



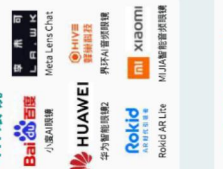
-AI手机



-AI耳机



-AI眼镜



-AI硬件新物种



四、AI应用格局：第一轮洗牌结束，聚焦20赛道5大场景

基于对全国现有AI产品（包含Web端和APP端）的梳理总结，我们在国内共统计到了千余款产品。为了更好地从数据维度观察国内产品的现状，我们结合产品具体功能、使用体验、运营状况和用户数据等因素，从中选取了四百余款具有代表性的产品进行进一步统计研究。

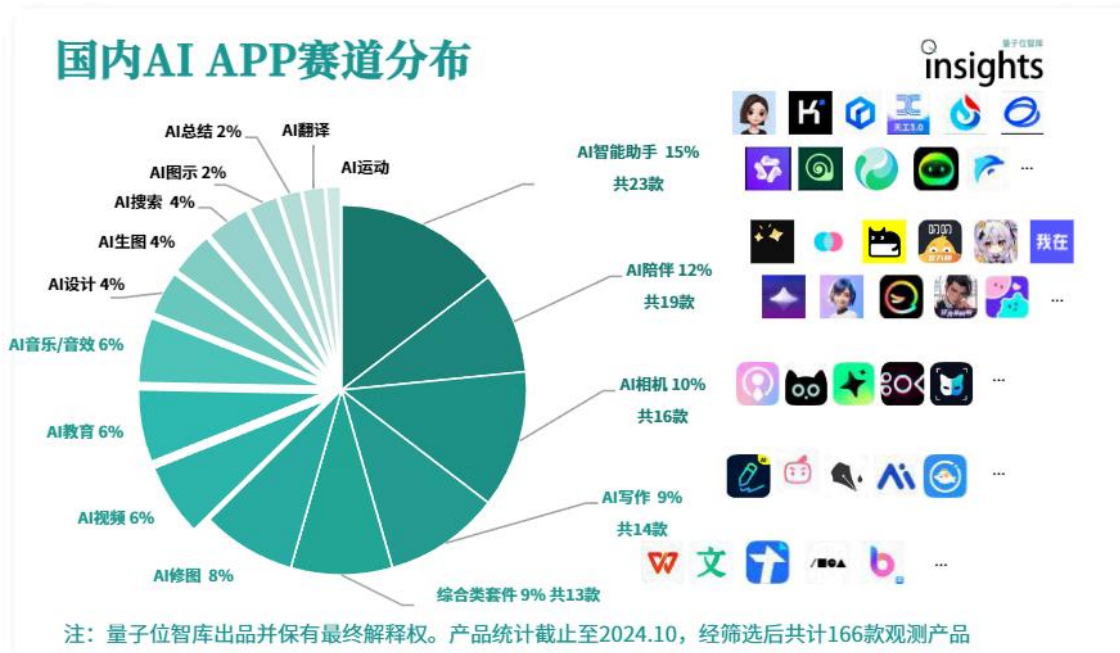
这400款产品全面覆盖了20余个细分赛道和五大场景，全面渗透个人生产力、休闲娱乐和日常生活等多维度。

（一）20+细分赛道

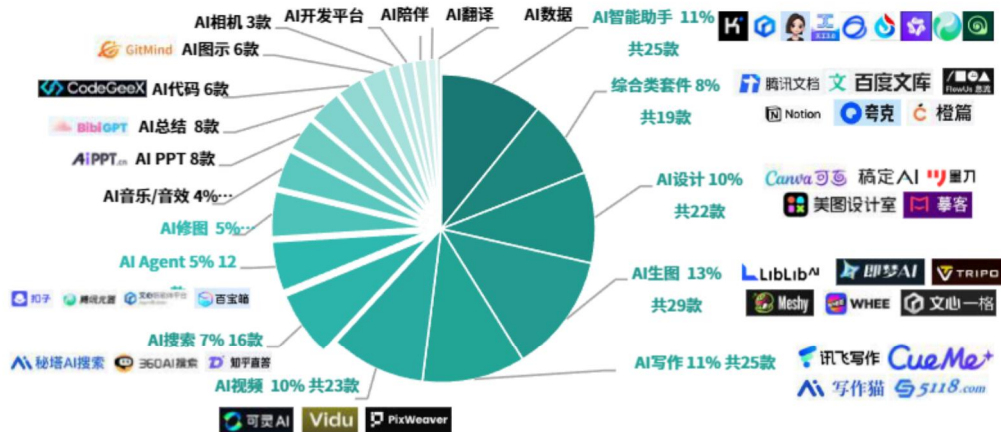
具体包含AI智能助手、AI陪伴、AI相机、AI写作、综合类套件、AI修图、AI视频、AI教育、AI音乐/音效、AI设计、AI生图、AI搜索、AI图示、AI总结和AI翻译共计20个赛道。根据产品数量可以分为三大梯队。

第一梯队	AI智能助手（34款）、AI写作（34款）、AI生图（33款）、AI视频（32款） AI设计（23款）、AI陪伴（23款）
第二梯队	AI修图（22款）、综合类套件（21款） AI搜索（19款）、AI相机（19款）、AI音乐/音效（17款）、AI开发平台（15款）
第三梯队	AI Agent（12款）、AI教育（10款）、AI总结（11款） AI PPT（8款）、AI图示（7款）、AI代码（7款）、AI翻译（5款）

由于Web端和APP端适用于不同场景，在两种使用形态下的产品赛道数量分布有所差异。在APP端的166款产品中，产品数量最多的赛道依次为AI智能助手、AI陪伴、AI相机、AI写作和综合类套件。而在Web端的230款产品中，产品数量最多的赛道则依次为AI智能助手、AI写作、AI视频和综合类套件。



国内AI产品Web端赛道分布

量子位智库
insights

注：量子位智库出品并保有最终解释权。产品统计截止至2024.10，经筛选后共计230款观测产品

以下是对各代表赛道的详细分析：

1. AI智能助手—梯队明显划分，豆包断层领先

AI智能助手是目前表现最突出的AI原生类产品，也是国内大模型自研厂商技术实力最直观的体现。

目前各产品间的功能差异较小，普遍包含对话交互、AI搜索、AI写作、数据分析等功能。而在收费方面，尚未形成明确的收费模式和付费意愿。仅有文心一言明确进行会员收费，而Kimi智能助手则是打赏式收费，智谱清言和天工AI分别针对视频及音乐生成功能进行收费。

从APP端来看，2024年5月成为关键分水岭，各家大模型厂商的AI智能助手登场完毕，产品间开始明显分化，形成了“1+1+6”的格局。第一名豆包在规模、增长、活跃、留存等各项数据上均断层式领先，至今仍保持着全面加速增长，也在9月成为国内首个总下载过亿的原生AI产品。24年10月，豆包累计下载量**超1.4亿**，当月新增下载量**超2000万**。Kimi智能助手虽然位居次席，但市场领导地位同样明显。第三梯队的6则指：文小言、智谱清言、讯飞星火、天工AI、阿里通义、腾讯元宝和海螺AI。

而在Web端，头部产品为Kimi智能助手和文心一言，月访问量均超过**2000万**。豆包和阿里通义月访问量则**超过1000万**。从2024年5-10月的用户规模数据来看，AI智能助手在Web端已经出现了明显的梯队划分：

快速增长	Kimi智能助手 豆包 海螺AI 跃问
平稳增长	阿里通义 文心一言 智谱清言 腾讯元宝
基本持平	百小应



图：Kimi智能助手（左）和豆包（右）的Web端界面



图：文心一言（左）和阿里通义（右）的Web端界面

总体来看，我国的AI智能助手与海外仍有相当距离，对标产品ChatGPT在Web端的月访问量超36亿，Gemini月访问量超2亿，Claude数据超8000万，差距超过5倍。

2. AI搜索—多渠道共同发力

国内头部产品为秘塔AI搜索和知乎直答，月总访问量分别超过600万和300万，均推出了简洁、深入、研究、专业等多种搜索模式，并支持用户通过上传文件构造个性化搜索引擎。此外，天工AI等AI智能助手也普遍提供AI搜索功能。国外典型对标产品Yandex月访问量过亿，Perplexity访问量超9000万。



图：秘塔AI搜索（左）和知乎直答（右）的Web端界面

从具体产品定位上可以分为AI加强搜索、原生类AI搜索和业务类AI搜索。

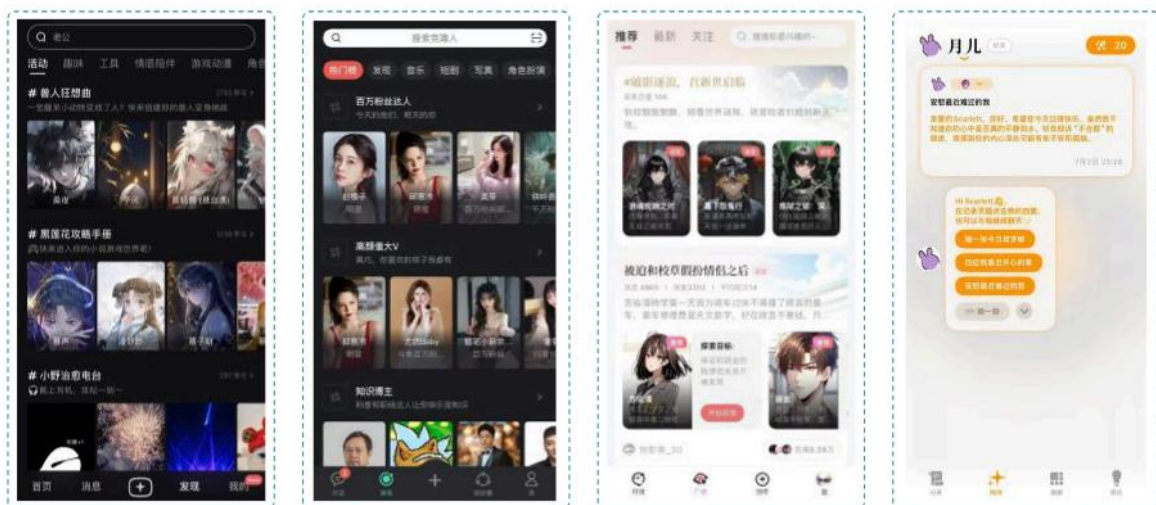
AI加强搜索	在原有搜索引擎、浏览器等互联网产品的基础上，叠加AI功能后衍生出的新一代产品。代表包括：纳米搜索、夸克浏览器、简单搜索等
原生AI搜索	以AI搜索为产品核心的AI原生产品。代表包括：秘塔AI搜索、ThinkAny、GenSpark、开搜AI等。也有MetaLaw、Devv.ai和C知道等针对特定领域的垂直AI搜索工具
业务AI搜索	在内容平台类原有业务上，叠加AI搜索功能的产品。代表包括：知乎直答、bilibili AI搜索助手、小红书的达芬奇/搜搜薯等。这种结合不仅增强了平台作为搜索引擎的功能性，也提升了用户在特定内容平台的使用体验

3. AI陪伴—增长乏力，Killer APP前景不明

从功能来看，量子位智库按照使用者和AI智能体的交互强度，将AI陪伴产品分为轻交互、中交互和强交互产品。所有产品基本均有角色扮演的成分。

轻交互产品	以心理疗愈和AI日记软件为代表，如林间聊愈室和心光。玩家在交流中起绝对主导作用，时间和精力投入相对较少，AI智能体的回应不会出现过多“惊喜”
中交互产品	更类似于游戏，提供特定的游玩体验，主要对应互动小说和智能体社区。玩家和AI智能体共同影响剧情走向，包括多智能体群聊对话和小说展开两种形式。玩家会集中性投入时间游玩，代表性产品包括小冰岛、冒泡鸭、有象等
强交互产品	以虚拟恋人代表，玩家和AI智能体会持续高频聊天，个别软件甚至可以提供叫起床等日常服务。此类产品是目前用户活跃粘性最好的部分

AI陪伴也是众多公司布局AI原生应用的重要选择。包含小冰科技、毛线球科技、Minimax、阶跃星辰、有零科技在内的20家AIGC公司共推出了21款产品。由于AI陪伴产品在功能上和角色扮演游戏、网剧、网文、社交，乃至心理咨询等业务深度重合，许多垂直赛道玩家，例如顺网科技、晓悟智能、阅文集团、Soul、心纪源等也纷纷进入这一领域。



图：星野、X Eva、彩云小梦、心光（从左到右）的APP界面

国内产品在新增速度和用户活跃等数据上**普遍下滑**。而全球范围内，Character AI、Janitor AI、Crushon、Talkie 的热度持续已超过一年。其中，Character AI网页端MAU可上亿，APP端MAU则超过两千万，国内MAU最高的星野在APP端数据仅约为其十分之一，这一数据表现在国外勉强进入前八。

经过和海外产品的深入对比，量子位智库认为关键在于国内产品**缺乏对深度情感链接的探索**。Character AI、Replika等海外产品更注重个性化和深度情感交流，追求具有独立人格和复杂情感反应的AI形象。而国内市场则更倾向于短期的趣味性互动或娱乐化的陪伴功能，产品满足的需求过于浅层，同类产品间可替代性强，难以让用户深入探索。

4. AI写作——成为标配AI功能

在功能上，AI写作大多提供“一键式成文+深度AI编辑”的综合方式。在一键成文时，使用者提供主题或关键信息后，AI可以根据模板直接产出整段文字。而深度AI编辑功能则更类似Copilot，能够在成文过程中提供文本润色、续写扩写、校阅勘误、风格转化、灵感激发、降重伪原创等辅助功能。

目前，AI写作已经成为生产力类工具的标配AI功能。

百度文库、WPS AI、博思Boardmix等综合类工具均将AI写作结合入自身工作流。此外，AI助手类工具也均能满足轻量级的常用写作场景，例如撰写周报、起草发言稿等。许多垂直赛道的产品，例如游戏+AI、营销+AI，甚至特定赛道的综合性厂商，也会基于业务特色和数据中台，提供相关的AI写作服务。



图：光速写作（左）和WPS AI（右）在Web端的AI写作界面

5. AI+图像——普遍得到需求验证

AI+图像可以具体分为**AI生图**、**AI设计**、**AI相机**、**AI修图**四个细分领域，目前在B、C两端普遍得到明确的需求验证。在Web端，AI视觉设计成为最受欢迎的细分赛道，AI绘画和AI修图紧随其后。而在APP端，AI相机的数据力压AI陪伴，成为2C AIGC产品中的大热门类。

AI视觉设计类产品分为两类，一类为平面设计相关，具体包含海报、设计概念图等，代表产品包括稿定AI、美图设计室和图怪兽。另一类则是进一步功能细分的产品，包括主打图文排版设计的135AI排版助手和主打UI设计的即时AI。Logo设计表现突出，入选产品包括标小智logo、标智客AI Logo和U钙网。

AI生图产品大多结合了**AI抠图**、**清晰度提升**等AI修图功能，一站式提供可用素材，目前2D生平面图/多视角图和3D生成模型领域均有代表产品出现。代表产品包括Liblib AI、Meshy、即梦AI、无限画、文心一格、神采Prome AI、WHEE和奇域AI等。

AI修图在Web端可分为**素材处理**和**批量修图**两类。利用AI进行抠图、提高分辨率，进而快速预处理素材已经成为业内技巧，代表产品有佐糖、Fotor和AI改图。而批量修图在会议跟拍和人像摄影中基本成为行业标配。代表产品有像素蛋糕、百度网盘AI修图、美图云秀。而在APP端，画质提升和智能抠图成为主流功能。美图秀秀、醒图、Wink等头部修图产品均已大规模深度植入AI修图功能。

AI相机在Web端和APP端呈现出了明显差异。Web端主要为电商2B的**AI商拍**，具体包括虚拟模特和虚拟商品图。使用者可以将白底商品图或人台图进行上传，产品会根据所需的模特姿势、背景氛围等生成可直接使用的商品详情图。代表产品包括美间AI创意商拍、Weshop唯象妙境、绘蛙等。APP端以「妙鸭相机」类2C娱乐向的人像摄影产品为代表。



图：稿定AI、佐糖、LibLib AI 和超能画布的Web端使用界面

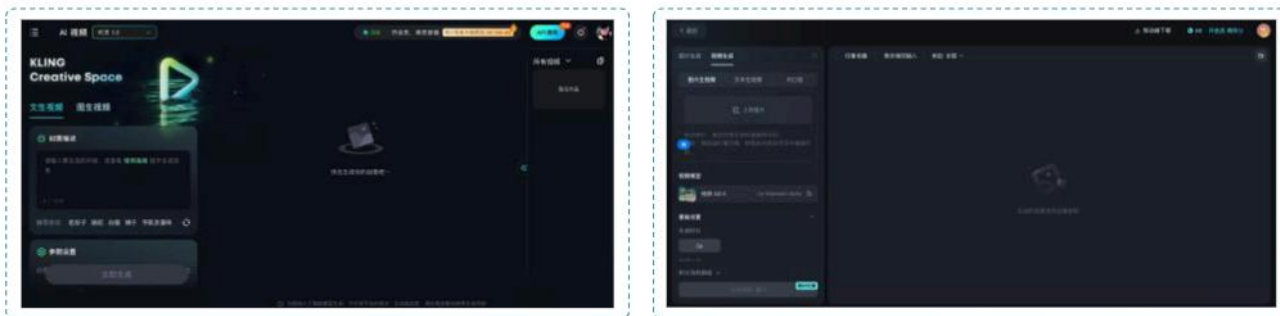
6. AI视频——直接生成视频仍待发力

受技术进步的影响，文字/图片直接生成视频在技术和创业领域已经成为热门，但由于产品落地具有一定的滞后性，**预计2025年将出现相对破圈的产品。**

目前来看，AI视频产品在Web端可以分为以下几种类型：

- **直接基于提示词生成视频：**包括可灵AI、PixVerse、智谱清影、即梦AI、Morph Studio、Video Studio等
- **基于AI的视频编辑软件：**包括视频翻译、文本剪视频、人像智能修图、添加滤镜、画质修复等功能，代表产品有万兴喵影、开拍等

- **自动智能切片：**通过自动生成分镜、配图、配音等，将长视频变为切片短视频，主要应用在电商和影视领域
- **拼接式视频智能成片：**由文字生成视频/图像后拼接为长视频，在应用上常见于AI短剧



图：可灵大模型（左）和即梦AI（右）的Web端使用界面

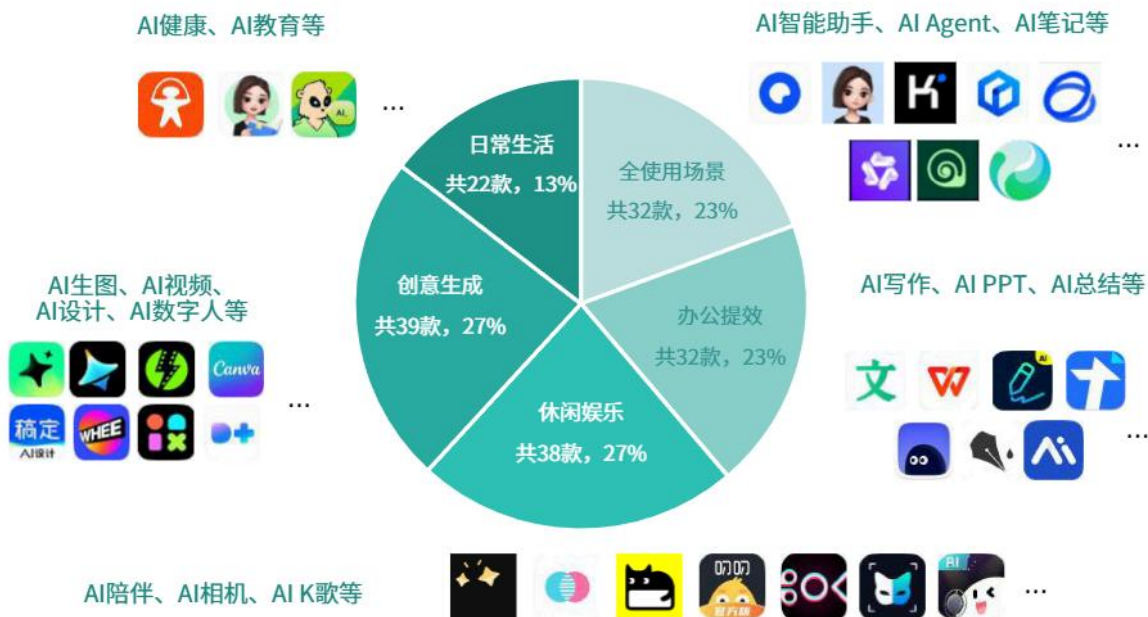
在APP端则可以分为以下三类：

- **AI相机：**支持对不同风格的视频模版进行换脸，包括快剪辑、美趣AI、AI换脸等
- **内嵌有AI功能的视频编辑工具：**包括剪映、快影、Wink、度加剪辑、Videoleap等
- **套壳类文生视频产品**

(二) 五大使用场景

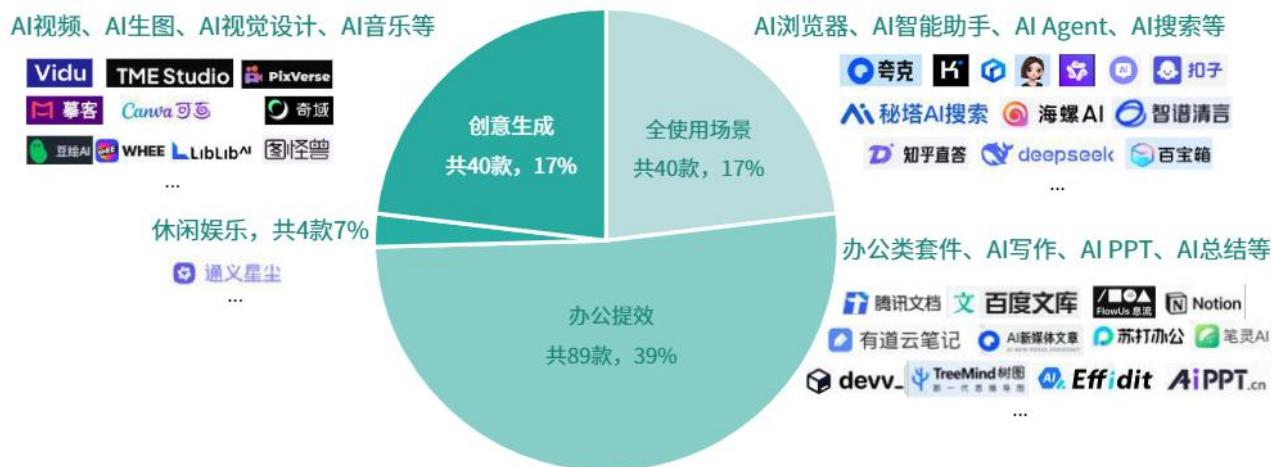
根据使用场景，所有产品可以划分为五类：全使用场景、工作提效、创意生成、休闲娱乐和日常生活。各使用场景对应的产品数量分别为65款、105款、125款、43款和24款，在Web端和APP端的具体分布如下：

国内AI APP赛道分布

量子位智库
insights

注：量子位智库出品并保有最终解释权。产品统计截止至2024.10，经筛选后共计166款观测产品

国内AI产品Web端场景分布

量子位智库
insights

注：量子位智库出品并保有最终解释权。产品统计截止至2024.10，经筛选后共计230款观测产品

在五类场景中，由于解决的需求更为明确和迫切，办公提效类工具在商业变现、产品数量、功能效果和市场普及上均有更好的表现。以下是针对各使用场景的具体分析：

1. 全使用场景——重在整体效率提升

主要涵盖多功能综合平台、AI智能助手和AI Agent。这类工具通常高度集成，普遍涵盖AI写作、AI总结等功能，能够为特定场景一站式解决方案。相对于专注单一场景的其他产品，这类综合性产品能够实现不同功能和服务的无缝对接，更强调跨场景跨平台协同和任务串联带来的整体效率提升。

未来演变方向可能包含：基于特定任务的自主决策能力、平台间互通配合、Agent化

2. 办公提效——整体数据表现最优

具体包括AI搜索、AI PPT、AI 总结、AI写作、AI图示、AI 代码、AI配音、AI相机中商业拍摄部分。

从数据来看，办公提效类产品整体的用户数据表现突出，量子位智库认为，具体涉及以下原因：效率提升可量化、易平衡用户价值与服务成本、使用场景广泛且高频，且产品在AI方向的可扩展性强，升级空间大。

未来的发展重点包括：

- 低成本提升输出内容的专业度：提示词优化系统、结合垂直领域专家知识、加入人工验证环节
- 现有工具与工作流集成：提升常用工具集成度
- 减少学习成本：尽可能做到即插即用

Sonnet 3.5和OpenAI o1为代表的新一代模型大幅提高了AI Coding能力，在SWE-bench上可以获得50%的得分，预计在1-2年内可以获得90%以上的得分，意味着大部分人类的编程任务即将可以被AI解决。

——戴雨森，真格基金管理合伙人

3. 创意生成——2025年有望显著突破

具体包括AI设计、AI生图、AI图像编辑、AI视频、AI音乐等赛道。

随着技术的持续进步，内容生成质量有望获得极为显著的提升。可以预见，创意生成类产品将成为2025年增长最快的一类产品，其中，AI视频生成以及AI设计领域中的3D和精细生成潜力巨大。

预计未来的发展重点包括：

- 解决技术瓶颈：
 - 生成质量与稳定性：高质量内容的一致性输出，需要对创作过程有细粒度控制
 - 创意表达的可控性：难以精确把控创意方向
 - 训练数据的版权限制：优质训练素材获取受限
 - 生成结果难以精细调整达到商用变现地步
- 解决变现瓶颈，寻求订阅制收费之外的变现方式，基于生成内容量收费或许是突破思路
- 通过优化UI和工作设计，降低学习和迁移成本

4. 休闲娱乐——合规挑战突出

指通过提供实时互动类内容和直接可消费内容等，提供情绪价值的产品。在赛道上主要对应2C侧的AI相机和AI陪伴。在产品中强调趣味性、情绪价值和社交属性，对技术上的情感计算能力、个性化算法、实时交互性能提出了更高要求，同时也对内容生态在内的运营有更高要求。

目前存在的问题是产品缺乏本质化差异，同时监管导致内容创作自由度有限，用户易失去新鲜感。

预计未来的发展重点包括：

- 情感交互的深度与自然度，如情感识别和共鸣能力、多模态的情感表达的多样性
- 个性化定制能力，了解用户偏好，进行场景感知
- 隐私问题保护
- 内容的丰富性和新鲜性，设计更为丰富和专业的底层机制

5. 日常生活

大多是在原有日常用软件的基础上通过要素识别和个性化计划制定来提升用户体验。具体包含AI健康、AI教育、AI运动、AI日记等。预估未来健康、运动和教育将成为重点发展赛道，原有相关互联网产品，如Keep等，将结合场景经验普遍引入AI功能。

五、AI应用竞争：多领域竞速运营大于技术，AI助手兵家必争

为了客观还原国内AI产品的现状，针对筛选出的**400款**代表性产品，我们从用户规模、新增速度、用户活跃和用户粘性四大角度进行了用户数据统计，横向对比和进一步分析。

整体来看，APP端和Web端均尚未出现互联网时代的破圈级产品，和海外的同类型产品相差距在**5倍**以上。一方面，国内对模型的监管要求在一定程度上限制了产品表现，导致最终使用体验不佳。另一方面，国内付费尚未形成闭环。缺乏差异化场景带来的同质化竞争进一步降低了用户付费意愿，限制初创公司投入与开发，形成了恶性闭环。

（一）APP端数据分析

总体来看，AI能力正在全面渗透APP端的各类场景和应用，字节跳动的APP工厂战略进展顺利。但杀手级场景的缺乏依旧是我国AI APP的未解难题，目前还没有产品能够拿出全维度的亮眼表现。

由于在市场认知、用户使用习惯、场景拓展等方面有先天优势，AI+X类产品的数据优势明显。AI原生应用由以豆包、Kimi智能助手、文小言和智谱清言为代表的AI智能助手领跑，随后为星野、猫箱等AI陪伴产品和豆包爱学、斑马AI学等AI教育产品。曾被寄予厚望的AI陪伴和AI搜索赛道整体不达预期。

但由于APP端在使用场景（生活化的碎片化场景）、可用数据（手机自带的GPS数据、摄像头提供的图像数据等）、使用时长（全天候互动）方面相对于Web端和PC端有明显优势，我们依旧认为AI APP未来更有可能出现质变性的增长。除手机外，智能手环、智能家居等其他物联网设备也有望在AI APP端发挥重大作用。

1. 用户规模 · 2024年1-10月累计下载

截至2024年10月，共56款产品的新增下载量**超过百万**。TOP 100的2024年累计用户规模门槛**约27万**。共两款产品数据过亿，分别为夸克（接近2.7亿）和豆包（接近1.4亿）。

夸克此前属于浏览器品类，2024年前开始强调自己的智能搜索功能，2024年中旬开始把AI搜索作为核心功能，2024年10月正式定位为「AI全能智能助手」，目前在AI方面主打AI搜索、AI写作和AI拍题三大功能。在定位上，AI功能已优先于其早先「网盘+扫描+浏览器」的主要业务。根据下载量变化来看，夸克在2023年中期开始下载量有所下降，但从次年3月开始下载量有显著增长，并在2024年中旬一度达到峰值。结合其宣发来看，我们认为AI功能的全面深度结合是夸克今年再增长的核心推动因素。

豆包AI由字节跳动发布，从今年5月开始在各用户指标上持续高速增长，目前已经在AI智能助手赛道形成断层式领先，并在9月成为国内首个下载量过亿的AI原生应用，是名副其实的国民级AI应用。

共**8款**产品今年累计新增下载量**超过千万**。依次为超过5000万的Kimi智能助手和百度文库。2000万量级的WPS Office，以及1000万级别的天工AI、星野、文小言、腾讯文档和阿里通义。此外，可画Canva的今年新增下载量已接近千万。

2. 用户新增 • 2024年10月新增下载

共有14款产品下载量超过百万，TOP 10的门槛约140万。

千万级以上梯队包含3款产品，分别为夸克（超2400万），豆包（超2200万）和Kimi智能助手（超1400万）。

百万到千万级梯队包含11款产品，依次为百度文库（超500万），WPS AI（约270万），猫箱、腾讯文档、文小言均在140-200万区间，随后数据下滑至120万。来自美国公司主打AI视频编辑的开拍即将达到百万。

值得注意的是，在新增速度超过百万级的14款产品中，来自字节跳动的产品就占据了4款，分别是豆包、猫箱、星绘和豆包爱学，分别在AI智能助手、AI陪伴、AI生图和AI教育中成为了新增速度TOP 1。在后三个赛道中，字节跳动的代表产品在发布时间上均不占据先发优势。



图：夸克和豆包的APP界面

用户规模榜单

1		夸克	夸克	综合类套件	2亿
2		豆包	字节跳动	AI智能助手	1亿
3		Kimi智能助手	月之暗面	AI智能助手	
4		百度文库	百度	综合类套件	5000万
5		WPS Office	WPS	综合类套件	2000万
6		天工AI	昆仑万维	AI智能助手	
7		星野	MLnimax	AI陪伴	
8		文小言	百度	AI智能助手	
9		腾讯文档	腾讯	综合类套件	
10		通义	阿里巴巴	AI智能助手	1000万
11		Canva可画	咖瓦信息	AI设计	
12		X Eva	小冰公司	AI陪伴	
13		讯飞星火	科大讯飞	AI智能助手	
14		智谱清言	智谱华章	AI智能助手	
15		斑马AI学	猿力科技	AI教育	
16		猫箱	字节跳动	AI陪伴	700万
17		天天跳绳	微芒科技	AI运动	
18		星绘	字节跳动	AI生图	
19		脸猫	风闻有你科技	AI相机	500万
20		美图证件照	美图	AI相机	

新增百万级AI APP



3. 用户活跃 · 2024年10月DAU

仅夸克1款产品平均日活**超千万**，超过2600万。**百万级梯队**共4款，包括超760万的豆包AI、超330万的Kimi智能助手、超170万的天天跳绳和超130万的文小言。

随后数据下跌至**50万量级**，10-50万共包含23款。TOP 10的数据门槛约35万。平均日活数据在**30万**以上的产品包括斑马AI学、天工AI、星野、X Eva、智谱清言、有道云笔记、猫箱和阿里通义。

用户新增速度榜单

1		夸克	夸克	综合类套件	2000万
2		豆包	字节跳动	AI智能助手	
3		Kimi智能助手	月之暗面	AI智能助手	
4		百度文库	百度	综合类套件	
5		WPS Office	WPS	综合类套件	
6		猫箱	字节跳动	AI陪伴	200万
7		腾讯文档	腾讯	综合类套件	
8		智谱清言	智谱华章	AI智能助手	
9		文小言	百度	AI智能助手	
10		天天跳绳	微芒科技	AI运动	
11		Canva可画	咖瓦信息	AI设计	100万
12		星绘	字节跳动	AI生图	
13		豆包爱学	字节跳动	AI教育	
14		星野	MInimax	AI陪伴	
15		开拍	美图	AI视频	
16		即梦AI	字节跳动	AI生图	70万
17		讯飞星火	科大讯飞	AI智能助手	
18		天工AI	肯仑万维	AI智能助手	
19		通义	阿里巴巴	AI智能助手	
20		腾讯元宝	腾讯	AI智能助手	

用户活跃榜单

1		WPS Office	WPS	综合类套件	3000万
2		夸克	夸克	综合类套件	2000万
3		豆包	字节跳动	AI智能助手	700万
4		Kimi智能助手	月之暗面	AI智能助手	300万
5		天天跳绳	微芒科技	AI运动	100万
6		文小言	百度	AI智能助手	
7		百度文库	百度	综合类套件	
8		斑马AI学	猿力科技	AI教育	
9		天工AI	肯仑万维	AI智能助手	
10		星野	MInimax	AI陪伴	50万
11		X Eva	小冰公司	AI陪伴	
12		智谱清言	智谱华章	AI智能助手	
13		有道云笔记	网易	综合类套件	
14		猫箱	字节跳动	AI陪伴	
15		通义	阿里巴巴	AI智能助手	30万
16		讯飞星火	科大讯飞	AI智能助手	
17		漫画相机	Boniu Tech	AI相机	
18		简单搜索	百度	AI搜索	
19		腾讯元宝	腾讯	AI智能助手	
20		稿定设计	稿定科技	AI设计	20万

AI智能助手和AI陪伴两个细分品类的头部产品有明显的整体数据优势。整体来看，AI相机的整体数据同样较为突出。

4. 用户粘性 · 2024年10月三日/七日留存率

第一名夸克的三日留存率超过40%，七日留存率接近30%。第二名是起源于AI语音记账的AI陪伴产品叨叨，三日留存率超过30%，七日留存率接近20%。

另外有13款产品的三日留存率超过20%，豆包AI和Kimi智能助手数据超过25%，其他产品数据差异小，具体包括腾讯文档、智谱清言、文小言、豆包爱学、星野、斑马AI学、天工AI、腾讯元宝、通义、讯飞星火和猫箱。

TOP 100的三日留存率数据门槛约10%，比较不理想。仅35款产品的七日新增留存率达到了10%，3款产品的数据约20%。

按照Facebook提出的“40-20-10”法则，也即优秀产品需要达到次日留存率40%，7日留存率20%，30日留存率10%来看，目前我国AI APP的用户粘性仍显不足。尽管需要结合应用所在的具体行业去考虑，但目前可以看到即便在主打日常高频交互的AI智能助手、AI陪伴和AI教育领域，距离这一标准也有较大差异。

用户粘性榜单

1		夸克	夸克	综合类套件	40%
2		叨叨	自古红蓝	AI陪伴	30%
3		豆包	字节跳动	AI智能助手	
4		Kimi智能助手	月之暗面	AI智能助手	
5		百度文库	百度	综合类套件	25%
6		腾讯文档	腾讯	综合类套件	
7		智谱清言	智谱华章	AI智能助手	
8		文小言	百度	AI智能助手	22%
9		豆包爱学	字节跳动	AI教育	
10		星野	MLnimax	AI陪伴	
11		斑马AI学	猿力科技	AI教育	
12		天工AI	肯伦万维	AI智能助手	
13		腾讯元宝	腾讯	AI智能助手	
14		通义	阿里巴巴	AI智能助手	
15		讯飞星火	科大讯飞	AI智能助手	
16		猫箱	字节跳动	AI陪伴	20%
17		光速写作	好课帮助	AI写作	
18		天天跳绳	微芒科技	AI运动	
19		讯飞听见	科大讯飞	AI总结	18%
20		Naver Papago	Naver Corp	AI翻译	

(二) Web端数据分析

在观察了各赛道头部产品2024年各月访问量的变化后，量子位智库遗憾地发现，除AI智能助手赛道外，其他所有赛道均基本处于停滞状态，在AI搜索、AI写作、AI生图等赛道中，甚至出现了头部产品数据下滑、或是下滑后仍未恢复至峰值的情况，多款明星产品的用户数据和宣传力度严重不符。

值得关注的逆增长产品包括：三款AI智能助手——Kimi智能助手、豆包、海螺AI，AI搜索——知乎直答、AI设计——美图设计室和Pixso AI等产品原型设计产品、AI生图/视频产品——即梦Dreamina，3D模型生成产品——Meshy和Tripo AI。

1. 用户规模 · 2024年10月总访问量

千万级产品 总共7款，包括3款多功能套件和4款AI智能助手。全球月访问量超过千万的Web端AI产品共67款，国内占比约10%。

定位为AI智能助手的夸克总访问量超7000万，该数据水平在全球范围内可进入TOP 10。全球范围内第一名ChatGPT的月访问量超过35亿，Bing的月访问量超过18亿。第二、三名均为办公类产品，分别是约3600万的腾讯文档和约2800万的百度文库。

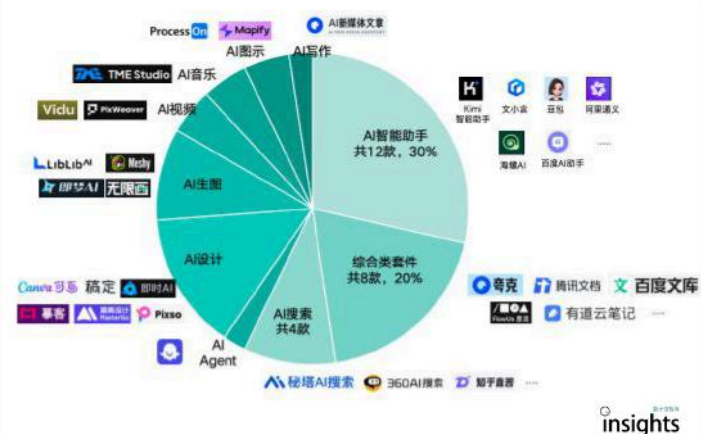
AI原生产品中的第一名为Kimi智能助手，访问量同样约2800万，目前定位为会推理解析，能深度思考的AI助手，并推出了有更强复杂问题推理能力的探索版本。后续三位同样为AI智能助手，依次为文心一言、豆包和通义。

月总访问量超过百万的共计40款产品，TOP 10的门槛约400万。TOP 100的门槛约10万。而全球AI产品Web端的TOP 100门槛约600万，差距6倍。

用户规模榜单

1	夸克	夸克	夸克	综合类套件	7000万
2	腾讯文档	腾讯文档	腾讯	综合类套件	3000万
3	百度文库	百度文库	百度	综合类套件	
4	Kimi智能助手	月之暗面	AI智能助手		
5	文心一言	百度	AI智能助手	2000万	
6	豆包	字节跳动	AI智能助手		
7	通义	阿里巴巴	AI智能助手	1000万	
8	Vidu	生数科技	AI视频		
9	秘塔AI搜索	秘塔科技	AI搜索	600万	
10	海螺AI	Minimax	AI智能助手	400万	
11	FlowUs AI息流	云上绿洲	综合类套件		
12	TME Studio	腾讯	AI音乐		
13	百度AI助手	百度	AI智能助手		
14	有道云笔记	网易	综合类套件		
15	智谱清言	智谱华章	AI智能助手		
16	知乎直答	知乎	AI搜索		
17	AI新媒体文章	夸克	AI写作		
18	PixVerse	爱诗科技	AI视频	300万	
19	Canva可画	咖瓦信息	AI设计		
20	苏打办公	360	集合站		

访问量超百万级AI Web端产品



2. 用户规模 · 2024年5-10月总访问量

共计21款产品过去半年总访问量**超过千万**。4款产品总访问量过亿，包括夸克3.7亿，百度文库约1.9亿，Kimi智能助手约1.4亿，文心一言约1.1亿。

超过千万的21款产品中，具体包括9款AI智能助手（Kimi、文心一言、豆包、通义、智谱清言、天工、讯飞星火、百度AI助手、海螺AI），4款多功能套件（夸克、百度文库、腾讯文档、Notion），2款AI搜索（秘塔搜索和知乎直答）。

3. 用户活跃 · 2024年10月独立访客数

月独立访客数即当月访问该网站的用户数量，可视作该网站当月的MAU，用于衡量产品的用户活跃程度。

共三款产品的月独立访客数**超过千万**，分别是夸克、Notion和百度文库。全球共计26款产品月活超过千万，中国占比约11%，活跃表现略逊于用户规模。夸克月独立访客数超2500万，Notion超1800万，百度文库则超过1300万。

总共19款产品的月独立访客数**超过百万**。其中包括6款综合类套件，6款智能助手和2款AI搜索占据和第9位和第11位。另外还有Pixverse、新媒体文章、即时设计、可画Canva。TOP 100的数据门槛是230万，全球AI产品TOP 100的数据门槛超200万。

根据使用场景来看，创意生成类产品的整体排位都比较靠后。全使用场景、工作提效和娱乐场景则表现更优。

4. 用户粘性 · 平均用户访问次数

共14款产品人均每月访问超过**5次**，第一名是新一代思维导图树图treemind，单人访问超过**8次**。AI智能助手阿里通义和生成原型图的摹客Mockplus数据超过**7次**。

TOP 10基本为AI智能助手和AI产品设计，TOP 50约**人均访问3次**。

用户活跃榜单

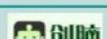
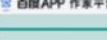
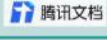
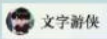
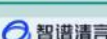
1		夸克	夸克	综合类套件	2500万
2		Notion	Notion	综合类套件	1500万
3		百度文库	百度	综合类套件	1000万
4		腾讯文档	腾讯	综合类套件	500万
5		Kimi智能助手	月之暗面	AI智能助手	350万
6		文心一言	百度	AI智能助手	
7		豆包	字节跳动	AI智能助手	
8		TME Studio	腾讯	AI音乐	200万
9		Vidu	生数科技	AI视频	
10		知乎直答	知乎	AI搜索	
11		百度AI助手	百度	AI智能助手	150万
12		秘塔AI搜索	秘塔科技	AI搜索	
13		PixVerse	爱诗科技	AI视频	
14		海螺AI	MInimax	AI智能助手	100万
15		FlowUs AI息流	云上绿洲	综合类套件	
16		通义	阿里巴巴	AI智能助手	
17		AI新媒体文章	夸克	AI写作	
18		Canva可画	咖瓦信息	AI设计	
19		有道云笔记	网易	综合类套件	
20		苏打办公	360	集合站	

5. 使用深度 · 单次使用时长

13款产品平均访问时长超过10分钟，AI生图占据比重较大。前两名都和营销相关，平均访问时长约20分钟。

第一名可生成商品图的豆绘AI平均访问时长查过20分钟，第二名AIGC营销平台绘蛙可以提供商品图和种草图文。AI生图中包括豆绘AI、WHEEL（WHEEL-专业设计师都在用的AI生图工具）、悠船（企业级文生图桌面工具）、LibLib AI和图怪兽。

用户粘性榜单

1		Treemind树图	聚石塔网络	AI图示	8次
2		通义	阿里巴巴	AI智能助手	
3		摹客	Mockplus	AI设计	7次
4		莫高设计	创作美好	AI设计	
5		创脑	词遇网络	AI智能助手	
6		DeepSeek	DeepSeek	AI智能助手	
7		Kimi智能助手	月之暗面	AI智能助手	
8		墨刀AI	万兴科技	AI设计	
9		百度作家平台	百度	AI写作	6次
10		Felo.ai	Felo	AI搜索	
11		腾讯文档	腾讯	综合类套件	
12		腾讯元宝	腾讯	AI智能助手	
13		文心一言	百度	AI智能助手	5次
14		文字游侠	铭锦轩科技	AI写作	
15		同花顺问财	同花顺	AI智能助手	
16		智谱清言	智谱华章	AI智能助手	
17		devv.ai	深宇识职能	AI搜索	4.5次
18		天工AI	肯仑万维	AI智能助手	
19		Genspark	Mainfunk	AI搜索	
20		奇域AI	小红书	AI生图	

使用深度榜单

1		豆绘AI	豆绘科技	AI生图	20分钟
2		绘蛙	腾讯	综合类套件	
3		5118 SEO写作	百度	综合类套件	18分钟
4		百宝箱	蚂蚁集团	AI Agent	15分钟
5		创脑	词遇网络	AI智能助手	
6		WHEEL	美图	AI生图	
7		通义星辰	阿里巴巴	AI陪伴	13分钟
8		莫高设计	创作美好	AI设计	
9		LibLib AI	奇点星宇	AI生图	
10		图怪兽	遇图网络	AI设计	10分钟
11		摹客	Mockplus	AI设计	
12		度加创作工具	百度	AI视频	9分钟
13		Notion	Notion	综合类套件	
14		扣子Coze	字节跳动	AI Agent	
15		GhostCut	招鲤科技	AI视频	
16		Motiff妙多	猿辅导	AI设计	
17		象寄AI	象寄科技	AI图像编辑	
18		文字游侠	铭锦轩科技	AI写作	8分钟
19		BetterYeah	斑头雁科技	AI Agent	
20		云界AI	大鱼互联	AI生图	

(三) 量子位智库AI 100

AI 100是量子位智库推出的AI产品和应用流行风向标集合，旨在结合产品功能、落地数据、用户体验和未来潜力，提供AI技术驱动下，产品创新和变革的第三方参考。

量子位智库将AI 100分为综合类和原生类，前者同时包含“AI+X”和“AI原生”两类AI产品，旨在给出完整综合的产品能力描述，后者则专注于AI原生类应用。

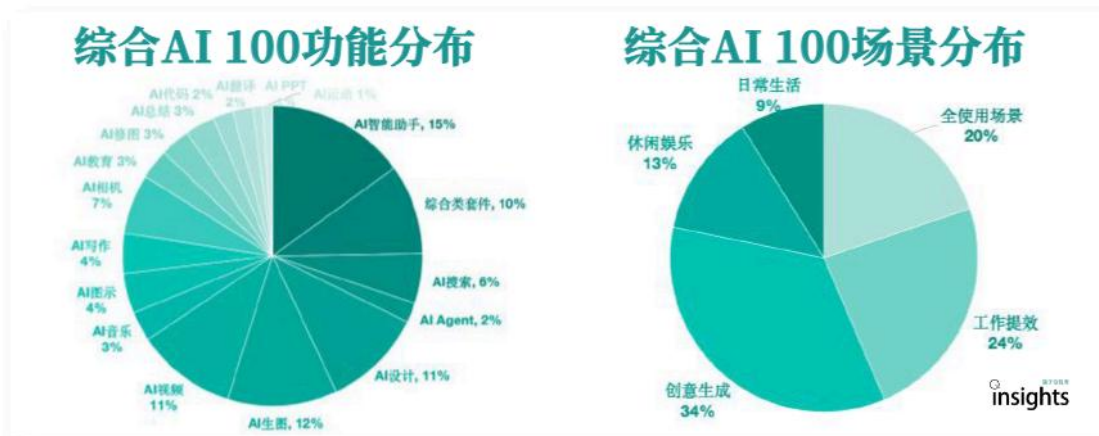
AI 100的评测体系，涵盖了生成能力、变革落地、用户体验、市场表现**四大一级指标**，包括**超20个二级指标**，全面涵盖了技术、用户、商业等产品评价维度，同时包含不同区间的下载总量、新增下载、活跃用户数、留存率等定量数据，以及用户评价、底层技术、功能评测等定性评分，以真实数据为基础，结合用户体验和市场前景，全面客观评估AI产品实力。

1. 综合AI 100

Web端产品TOP 100的总访问量门槛大约10万，MAU门槛约8万，平均访问时长不到2分钟，平均访问次数不到2次。

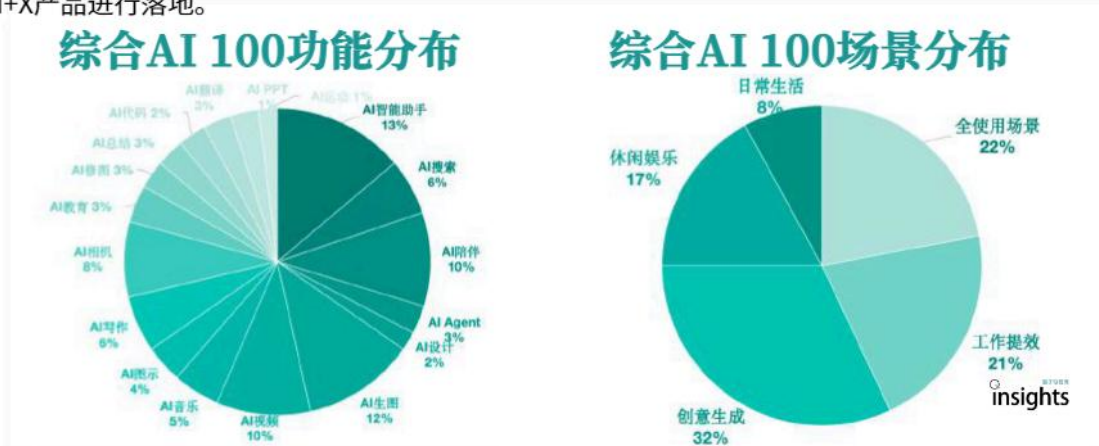
APP端产品TOP 100月新增下载门槛约5万，日活8000，三日留存率不到15%。

综合AI 100的功能和场景分布分别如下图。按照产品性质分类，则包括71个AI原生产品和29个AI+X产品。但AI+X产品整体的数据表现更加突出。



2. 原生AI 100

原生AI 100的不同功能占比和综合AI 100出现了明显差异。AI智能助手，AI搜索和AI陪伴是三个领跑赛道，分别有14款、6款和10款产品入围，后两者的数据表现较为一般。多功能套件和AI设计的产品数量锐减，证明这两个赛道目前仍主要依托AI+X产品进行落地。



综合 AI 100 insights

2024年度最受欢迎的100款AIGC产品

AI Agent



扣子Coze



百度AgentBuilder

AI 陪伴



猫箱

星野

X Eva

筑梦岛

冒泡鸭

我在AI

WoW

叨叨

AI 设计



Canva可画

美图设计室

墨刀AI

稿定AI

标智客AI Logo

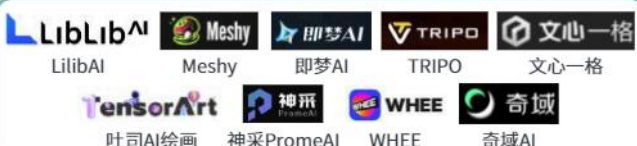
MasterGO AI 莫高设计

摹客

Pixso AI

堆友D.Design

AI 生图



LibLibAI

Meshy

即梦AI

TRIPO

文心一格

TensorArt

神采PromeAI

WHEE

奇域AI

AI 视频



开拍

闪剪

快剪辑

度加创作工具

万兴喵影

PixVerse

Vidu

可灵AI

MorphStudio

AI 开发平台



Dify

AI 翻译



Naver Papago - AI 翻译

有道翻译

有道AI翻译

AI 总结



讯飞听见

Readpaper

包阅AI

代码生成



豆包MarsCode

CodeGeeX

多功能套件



夸克

百度文库

WPS Office

腾讯文档

FlowUs AI 息流

360AI浏览器

Boardmix

橙篇

讯飞智文

小悟空

AI 智能助手



豆包

Kimi智能助手

文心一言

阿里通义

支小宝

讯飞星火

智谱清言

天工AI

腾讯元宝

同花顺问财

DeepSeek

跃问

百小应

海螺AI

AI PPT



AiPPT

AI 教育



豆包爱学

斑马AI学

海豚AI学

AI 写作



AI新媒体文章

AI新媒体文章

光速写作

新华妙笔

新华妙笔AI公文写作

笔灵AI

笔灵AI

AI 修图



你我当年

佐糖

Fotor

AI 音乐/音效



AI音乐学园

给麦

TME Studio

AI 搜索



知乎直答

GenSpark

Felo.ai

devv.ai

秘塔AI搜索

AI 图示



ProcessOn

亿图脑图AI

Treemind树图

GitMind思乎

AI 相机



星绘

脸猫

脸趣

剪同款

Faceplay秀脸

妙鸭相机

超能画布

原生 AI 100 insights

2024年度最受欢迎的100款AIGC产品

AI 智能助手



AI 陪伴



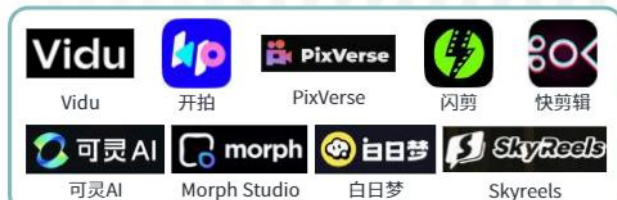
AI 搜索



AI 修图



AI 视频



其他



AI Agent



AI PPT



AI 设计



AI 写作



AI 翻译



AI 生图



AI 相机



AI 音乐/音效



AI 教育



代码生成



AI 总结



多功能套件



六、AI应用增长：AI+X赋能类产品大干快上，原生AI爆款难求

按照产品属性而言，量子位智库将AI产品划分为以AI为底层设计逻辑的AI原生类产品、在原有互联网产品上深度嵌入AI功能的AI+X产品、基于外接API微创新的套壳类产品和将多个产品/模型API集中拼凑的集合站类产品。



在AI+X产品中，AI体现为在已有产品上叠加AIGC能力，在业务和变现上都不以AIGC为核心，AIGC用于优化「核心体验」。目前，各类的头部互联网产品均在向该方向靠拢，依靠生成式AI进一步提升自身的用户基础与用户价值。

从数据来看，由于和业务流程融合的更为紧密、需求识别明确等原因，此类产品目前的整体数据表现显著优于AI原生类产品。

在入围的400款产品观测中，尽管AI+X产品数量较少，但在各维度数据的头部梯队中均占据了前列，具体数据表现也和AI原生类产品有断层式差距。

在APP端，我们共统计到了34款AI+X类产品，占比20%。其中5款产品进入新增下载榜前10名，分别为夸克浏览器、百度文库、WPS Office、腾讯文档和天天跳绳，其中夸克的当月下载量接近2500万，排名第一。12款AI+X类APP产品进入TOP 30，在排除AI智能助手后，AI+X类产品占据绝对比重。平均日活和留存率方面，AI+X类产品的优势则更加明显。在7款日活超百万的产品中，AI+X占据四位，超千万的两款产品均为AI+X产品，分别为WPS Office和夸克。

在Web端，我们共统计到55款AI+X产品，占比约23%。总访问量前3名均为AI+X类产品，第一名夸克全球访问量超7000万，第二名腾讯文档超3500万，第三名百度文库超2800万，总和高于TOP 10中其他7款AI原生产品的总和。月独立访客数，也即用户活跃排名的前四名也均为AI+X类产品，分别为夸克、Notion、百度文库和腾讯文档。

量子位智库认为，办公软件和内容平台成为AI+X产品的重点布局领域。

(一) 办公软件类AI+X产品

对办公软件而言，续写、改写、命题写作等不同程度的AI写作功能，以及针对论文、小说等不同题材的AI总结功能基本成为标配。全球范围内的代表产品包括Grammarly和Notion AI。

其中，主要业务为提供模板及参考内容的素材库类产品 and 编辑器形态的办公软件表现更为突出，代表产品为百度文库和WPS AI。由于AI生成效果会直接影响产品的核心使用体验，此类产品相对更强调具体功能的精准度。

百度文库

素材库性质

由于能够在各类题材下提供各类参考文档，其原定位更接近素材库，在引入AI编辑功能后，新定位已经更新为“一站式AI内容获取和创作平台”。具体提供的功能包含：

AI写作和编辑

文档问答和文档总结

PPT生成

上传图片生成

PPT生成演讲稿

思维导图

生成研究报告、漫画、小说



WPS AI

综合类编辑器性质

此类应用本身属于内容创作工具，一般在交互和功能设计上更加灵活，允许在编辑器页面上直接开启续写、改写等功能，以保证不中断创作过程。支持的AI功能也会种类更加多样，会采取一定的外接API。类似产品还包括腾讯文档、有道云笔记、印象笔记等。

北京又进入了秋季，从南到北，不同种类的植被逐渐转红，每天早上开车经过道路，总会看到不一样的色彩变化。



快速起草公文周报

从无到有智能起草工作周报、论文、公文等文档

一键美化汇报PPT

智能生成大纲，点击一键生成完整PPT，并自动排版美化

自动生成表格公式

说出需求自行撰写公式，快速完成数据分析和业务洞察

轻松总结长文PDF

智能理解文献、财报等长文PDF文档，智能问答快速获取所需信息



而在Boardmix博思白板和FlowUs AI息流这类强调协作和全范围编辑的办公类产品中，AI功能的大而全则相对更重要，这类产品往往会全面覆盖AI代码、AI表格、AI写作等多种功能，保证整体体验的智能化。大而全的理念同样出现在集合站式的AI+X产品中。

Boardmix 博思白板

多功能协作平台性质

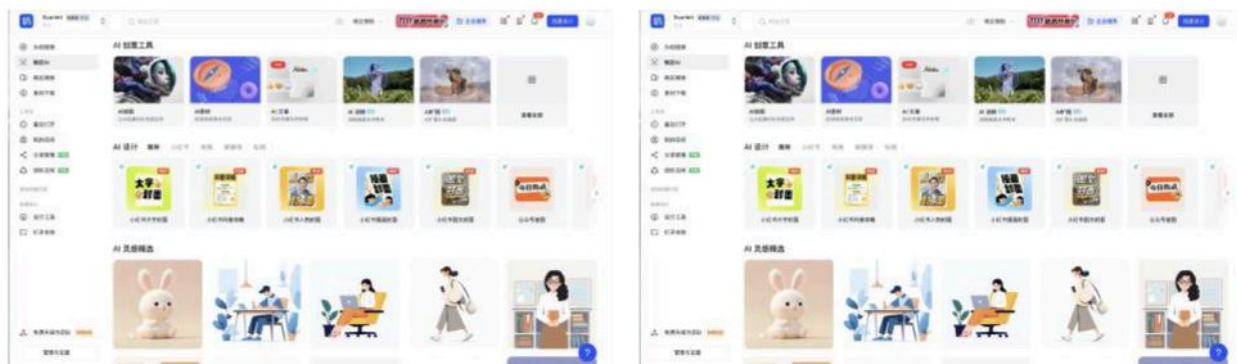
此类平台在侧重编辑器性质的同时，会更强调不同创作者之间的协同，同时在单个文档同时支持图表、图示、数据、文字等不同形式的内容。类似产品还包括妙办画板、小画桌、Fabrie等。



稿定设计

集合站性质

由于AI的生成能力尚无法直接大幅优化核心工作内容，此类软件更多提供周边的AI优化编辑功能。稿定设计定位为聚焦商业设计的多场景在线设计平台，在站内会以单点的方式提供AI生图、AI写作、AI图像编辑等功能。



此外，类似钉钉和飞书的办公软件会在自身各类场景及工作流内全面引入AI，强调AI功能的覆盖率和在不同产出间的协同。相对于其他类型的AI+X类办公产品，最大价值在于能够结合报销、会议等具体场景，生成特定场景下的数字员工，或者和使用者通过个人知识库、使用习惯等深度绑定，以个人为单位形成AI智能助手。

(二) 内容平台类AI+X产品

在内容平台中，AIGC大多从三个方向共同发力：基于平台内容的AI搜索、用于带动UGC的AI生成功能及模板，还有门槛进一步降低的内容创作工具。

通过帮助用户更快地寻找到所需内容，内容平台可以借助AI搜索大幅优化用户体验，同时向搜索引擎进行转型。

除了抖音、小红书等将AI搜索融入原平台的做法外，现在也已出现了将AI搜索单独剥离为产品的内容平台，如知乎直答和问问小宇宙，是对原内容平台定位的一次升级。



图：问问小宇宙和知乎直答的Web使用界面

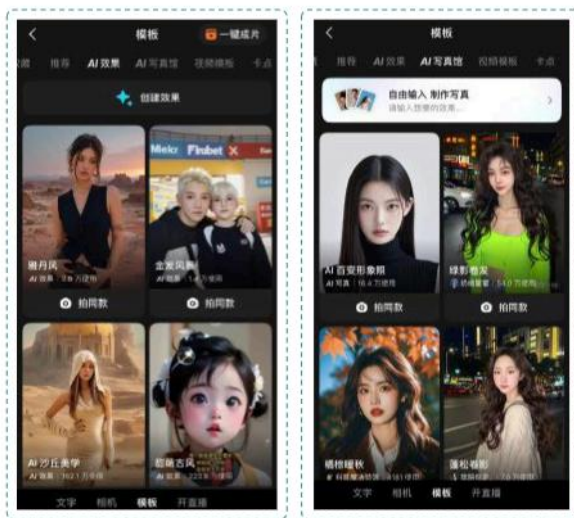


图：小红书的AI搜索界面

在促进UGC方面，自妙鸭相机爆火后，用于生成不同人物写真的AI相机和类似的AI特效也已内嵌入各内容平台和美图秀秀等修图软件。

除了能够通过附加功能提升产品体验，诸如小猫唱歌、影视剧二创等通过AI产出的新内容也往往会在平台上引发现象级传播，拉动整体活跃度。

而在编辑软件方面，目前，包括字节剪映、快手快剪、bilibili必剪、百度作家平台在内的多款创作工具均已深度植入基于AI的视频/图像/文本生成功能，万兴喵影等独立编辑软件同样有深入布局。国外的代表性对标产品是Adobe Firefly。



图：抖音创作中心中的AI内容

量子位智库认为，X+AI类应用在数据表现更优的原因主要包括以下几点：

- **场景融合优势：**在X+AI应用中，AI功能能够无缝集成到已有产品和工作流程中，用户无需迁移底层数据和逻辑，就可以在熟悉的平台上完成从文档编辑到AI辅助的全流程操作。这种深度融合有效降低了用户的认知和操作门槛
- **认知负担最小化：**在用户熟悉的界面中自然植入AI功能，复用已有的操作逻辑和交互方式，避免引入新的概念或术语，确保AI功能能辅助而不打断工作流。场景化的功能入口还能增强情景感知，进一步优化用户体验
- **付费模式顺畅：**AI功能作为附加值提供给用户，不会额外增加付费环节，避免了AI作为核心产品时的定价单一性
- **用户和场景理解**

基于此，量子位智库对AI原生产品有如下建议：

场景融合：加强与现有工具的集成

API接口与插件开发

- 提供丰富的API接口，或者开发跨平台插件，确保AI功能可以嵌入用户的日常工作流程中

跨平台支持

- 确保AI工具能够在不同平台（如移动端、桌面端）无缝运行，优化用户的多平台体验

简化用户体验：降低认知负担

自然嵌入AI功能

- 在用户熟悉的界面中集成AI功能，复用已有的操作逻辑

智能化功能入口

- 根据用户操作上下文，提供智能提示或自动触发AI功能，减少手动操作。

品牌信任与推广

合作与联合推广

- 与知名品牌或平台合作，通过这些品牌的市场认可度提升自身的信任度

七、AI产品趋势：多模态上马，Agent席卷一切，高度个性化呼之欲出

（一）多模态

多模态交互的核心在于通过整合不同类型的输入和输出（文本、语音、图像、视频、手势等），为用户提供更丰富、自然的交互体验。相比单一的文本或语音交互，多模态交互能够更好地适应不同的场景和用户需求，提升人机交互的灵活性和用户体验。例如，在驾驶、运动、办公等不同环境下，用户可以通过语音、手势等自然方式与AI进行交流，而不仅限于打字或点击。

此外，多模态交互可以让AI通过多个感官同时获取信息，从而提升对用户意图的理解。例如，AI不仅可以通过语音识别用户的指令，还能通过图像识别用户的表情、手势或周围环境。多模态信息的融合能够帮助AI更准确和全面地理解用户的需求，减少误解或偏差，帮助AI更有效的处理复杂任务，例如同时通过文字和图像来提高对物体的识别精度。

在国内的AI产品中，最为常见的是**语音+文字**的多模态交互，多出现于AI智能助手和AI陪伴中。随着Transformer架构使大模型对图像和视频信息的处理能力快速提升，预计25年将开始出现更为综合性的多模态交互，AI能够通过物联网、特定信息等多种感知通道进行协同。多模态输入和输出将使AI交互性更强，交互频次更高和适用场景也更加丰富，带来AI产品整体水平的显著提升。

以NotebookLM为例，其NotebookLM的“音频概览”功能能够将文本文件、视频、PPT、录音甚至数据集转换成自然、有趣的播客形式，这种多模态输入的方式在24年引发了一波热潮，预计更多类似的高热度产品将在25年出现。



图：通义效率、豆包和星野中对多模态交互的支持

(二) Agent化

随着技术的发展，用户对于人机交互的需求也在不断进化。

Agent作为融合感知、分析、决策和执行能力的智能体，能够根据用户历史行为和偏好，主动提供建议、提醒和个性化执行能力，为用户提供高度个性化的任务。其交互的主动性和自动化远超出现有的软件。

现有的2C产品主要分为内嵌在产品中的特定Agent，以及Agent搭建平台。前者的代表包括智谱清言、阿里通义、支小宝等。后者的代表产品包括字节跳动的扣子Coze、百度的文心智能体平台、汇智智能Gnomic、腾讯元器等。智谱华章、阿里巴巴也开始在其AI智能助手界面中添加AI Agent的入口。

同时，我们也统计到了多款企业级的2B Agent产品，以不同员工职能为单位，如财会、采购等构建智能体，也就是“数字员工”。代表包括逻辑智能的InsiderX智能员工、意言科技的Tyrion.ai平台、未来式智能的AutoAgents、澜码科技Ask XBOT和Laplace AI Lab等。蚂蚁集团也推出了2B的Agent搭建平台百宝箱。



图：Agent搭建平台扣子Coze和百宝箱

从技术和配套设施两方面的发展来看，AI Agent即将从2025年开始广泛投入使用。

技术上，大模型模型性能有了显著进步。工具调用和多Agent调用编排能力的完善使Agent能够更好地跨越不同平台和服务。大模型技术中推理决策框架的成熟，例如多步骤推理能力等，将进一步加强AI Agent在推理和决策方面的核心优势，帮助AI Agent能够更为灵活自主地应对复杂任务。

在配套设施方面，多个Agent开发平台正在快速发展。各大智能硬件厂商也纷纷在底层操作系统中加强自研AI Agent的布局，联想、荣耀等厂商目前已有旗舰机型出现，荣耀已率先发布首个跨应用开放生态AI智能体。搭载有自研AI Agent的智能硬件预计将成为智能电子设备近期的主潮。

量子位智库认为，AI Agent有望带来独属于AIGC时代的交互方式、产品形态和商业模式。

目前，Copilot类产品首先实现商业落地，Agent类产品尚待时间考验。

——陈石，峰瑞资本

交互方式	从GUI（Graphical User Interface，点击式的图形界面交互）到VUI（以物联网为代表的智能语音交互，Voice User Interface），并最终进入Agent UI，也即由Agent根据用户的意图主动提出建议，协调多平台多应用完成任务
产品形态	相对于单点式被动服务的APP，AI Agent能够更好地满足用户对复杂场景应用的需求。结合AI硬件对AI Agent的深度植入来看，未来APP的服务形式可能会直接被Agent替代。由于用户不会再直接和APP发生交互，APP端对UI/UX的投入将显著下降，转向对核心功能研发及不同APP间协调逻辑的开发
商品形式	和产品形态的变化相关，由于Agent更强调长期、持续性的服务，未来的单点服务形式可能会逐步被持续订阅取代，并出现定制化的自动化流程、个性化的知识管理和独特的行业解决方案等差异化的增值服务

（三）高度个性化

从个性化推荐到直接生成个性化内容，AIGC能够使用户体验的个性化程度有明显提升，这将帮助产品进一步完善用户体验，并通过提高用户忠诚度和迁移成本实现差异化定价和进一步的服务增值，对产品的差异化竞争有重大意义。

目前，基于AIGC的高度个性化已经在AI教育（个性化题库及教学安排）、AI陪伴（AI个人助理及虚拟伙伴）、AI营销（商品个性化推荐、营销内容个性化生成）领域有明显进展。在硬件端搭载的多款AI智能助手已开始以高度个性化的个人助理作为宣传重点。

实现方式包括：

- **构建多模态实时用户画像：**融合行为数据、偏好设置、历史交互等多种数据，并结合实际情境精准确定用户需求
- **私有知识库构建及无缝对接：**支持用户导入个人文档、笔记和专业资料，由AI快速获取用户专属知识，并构建知识索引
- **RAG优化：**RAG（检索增强生成）系统应支持同时从多个知识源进行检索，包括用户的私有知识库、公共数据库、互联网资源等；同时需要具备上下文感知能力，能够根据用户当前的需求、任务和对话上下文，自动筛选出最相关的知识
- **知识适配机制：**建立个人知识域和偏好的映射关系，帮助AI更好地理解用户的专业背景及对输出内容的偏好

03

INDUSTRY

趋势八 AI智变千行百业：左手变革生产力，右手重塑行业生态

（一）AI：落地于千行百业 /54

（二）AI落地行业特点分析 /55

趋势九 AI行业渗透率：数据基础决定初速度，用户需求成为加速度

（一）AI行业的变革与渗透 /75

趋势十 AI创投：投融资马太效应明显，国家队出手频率提升

（一）我国AI行业投融资分析 /78

（二）科技合作伙伴观点精选 /82

（三）我国AI及AI落地政策分析 /85

「千行百业」优秀AI落地方案推荐 /87

八、AI智变千行百业：左手变革生产力，右手重塑行业生态

（一）AI：落地于千行百业

2024年以来，量子位智库推出了《中国AIGC应用全景报告》、《AI广告营销行业全景报告》、《AI视频生成研究报告》、《AIGC教育行业全景报告》、《AI音乐应用行业报告》、《中国AI大模型创业格局报告》、《中国具身智能创投报告》、《AI游戏驱动力产业研究报告》等多篇深度报告，聚焦AI技术在产品和行业层面、在各个不同行业的落地场景和发展潜力。

前两个章节中，我们探讨了本年度最前沿、热度最高的AI技术趋势及AI产品趋势，本章节中，我们将集中讨论AI在千行百业的落地情况和呈现出的行业特点。

1. 2个情境

我们发现，站在当下的节点，尽管AI对千行百业的生产力和生态都有改变作用，但由于行业底色不同、环境不同、资源配置重心不同，受AI影响的行业呈现出「AI+」和「AI 原生」两大情境。

AI+：在这一情境中，AI多以生产力工具的角色出现，能够在质量和效率上对行业的各个环节进行渗透

AI 原生：在这一情境中，行业涉及的系统、应用和服务从一开始就完全基于AI技术进行设计和开发

区分AI+行业和AI原生行业的重要特征是AI在其组织架构中的地位。例如智能驾驶、具身智能等行业，尽管可以在历史溯源中找到汽车、机器人等原始行业，但其脱离AI后就失去了作为一个行业独立关注和研究的意义。

但AI+与AI原生在行业中并不是泾渭分明的关系，而是动态发展的。我们认为，AI对行业的变革和渗透值得高度关注。如同曾经的互联网技术，在逐渐改变了第一、第二和第三产业后，又慢慢发展成为一个独立的行业。相似地，AI作为新质生产力，必将逐渐改变与人类息息相关的各行业，我们相信，千行百业受AI变革仅有**先后**、**早晚**、**轻重**程度之分，而无**是否**之争。

基于以上特点和量子位智库在2024年度对整个AI及相关落地行业的重点追踪和观察，我们以智能驾驶、具身智能、智能硬件、游戏、影视、营销、教育、医疗为代表行业进行研究。



本章节中提到的量子位智库所发表的报告，均可通过以上二维码阅读并下载电子版

(二) AI落地行业特点分析

1. 智能驾驶：Robotaxi新时代，无人驾驶时刻来临？¹

作为本篇第一个重点分析的行业，智能驾驶行业从诞生之初即与AI的概念紧密相连。1956年，约翰·麦卡锡在达特茅斯会议上首次提出了「人工智能」这一术语，标志着AI作为一个独立研究领域的诞生；也正是1956年，通用汽车推出了世界上第一辆安装了自动导航系统的概念车Firebird II。从理念上来讲，智能驾驶领域所追求的「无人驾驶」，正是以完善的AI取代人类驾驶员。

2023年，我们跟进了端到端模型在自动驾驶领域的重要进展，BEV+Transformer 技术从数据感知和架构层面优化了整个自动驾驶系统。2024年，我们关注到世界模型在自动驾驶领域的应用，世界模型将生成未来场景和真实的未来时刻数据，进行无监督训练。同时在本报告的第一章节对相关技术做了更深入的探讨。



端到端自动驾驶成共识：
BEV+Transformer重构技术路线

图：《2023年度十大前沿科技趋势报告》

(1) 智能驾驶行业的特点

当前智能驾驶行业存在2大行业共识：

① 顶天立地：技术进阶路线明晰

行业内智能驾驶产品可被分为「行」、「泊」两类，每一类根据技术难度的增长，存在不同的迭代路线。

行车环境中，各玩家的产品迭代路线可分为：**基础L2级** → **高速NOA** → **城区NOA**

泊车环境中，各类泊车产品迭代路线为：**半自动泊车** → **全自动泊车** → **记忆泊车** → **自主代客泊车**

② 铺天盖地：产品价格门槛下探

智能驾驶不断向上迭代的同时，配有高阶智能驾驶产品的车型起售价在不断下降，智能驾驶产品售价也在**4万元以下**。

(2) 影响行业演进的关键因素

① 关键配件成本的下降

激光雷达量产规模的增加，带来成本进一步下探，更有利于大规模上车。

② L4技术供应商的入局，带来技术支持

近年来越来越多的L4玩家推出适用于乘用车的智能驾驶产品。自动驾驶技术的释放可以让车企更快装配高技术等级的智能驾驶产品，缩短自研周期。

③ 量产车标准化方案

数据质量（传感器方案是否一致），也是智能驾驶能力迭代的影响因素之一。不同传感器方案会使用不同的数据格式和处理算法，从而需要更多流程进行数据转换和统一。因此，相同规格的传感器方案也能帮助智能驾驶能力快速迭代。

技术能力突破可产品化临界点才可支持广泛的落地。

——清华智能产业研究院（AIR）

¹ 对智能驾驶及Robotaxi的更多分析，可以扫码下载我们发布的《Robotaxi2024年度格局报告》和《智能驾驶2024年度报告》

(3) Robotaxi起势

如果说智能车的落地意味着自动驾驶阶段的开启，那Robotaxi的落地正意味着「无人驾驶」时代的到来。2024年，我们观察到在智能驾驶行业，Robotaxi的兴起正带来行业新的增长点和发展空间。

Robotaxi，狭义是指自动驾驶出租车，广义也可包括Robobus（无人小巴）、无人驾驶下的其他共享出行产品等。



① 2024年将成为行业兴起、商业化落地的关键节点

行业主要由自动驾驶技术供应商、车厂和出行平台组成，目前尚处于发展阶段，前期研发投入高。在落地上运营范围和车队规模都尚未铺开，玩家支出普遍大于收入。行业头部预计将在2024年年底实现收支平衡，2025年实现盈利，走向正反馈。

② 技术驱动、政策支持下的行业规模的指数级增长

- 技术是Robotaxi的关键驱动力，头部玩家已实现端到端大模型的自动驾驶，即用AI模型解决自动驾驶除控制外的大部分任务；同时，行业开始转向利用世界模型训练自动驾驶算法，世界模型使得自动驾驶有望摆脱路测依赖，以更低成本收集到Corner Case，加快算法迭代。
- 政策为技术在场景内的真正落地提供支持，同时又起到引导规范的作用。2023年上海市发布细则对临港新片区无人驾驶智能网联汽车的测试、应用和商业化运营做出明确规定。随后在国内，重庆、深圳、广州、武汉和北京等多地都开放试点区域支持Robotaxi的测试和运营。

两相促进之下，Robotaxi作为智能驾驶行业的一部分，在今年实现**破亿元**的市场规模增长，并在**2030年**对网约车整体市场达到**45%**以上的渗透率。



图：我国Robotaxi市场规模测算，《Robotaxi2024年度格局报告》

(4) 智能驾驶行业市场规模分析

2023年，我国L2级新乘用车渗透率达到47.3%；2024年1-8月，国内市场L2及以上智能驾驶装机量累计达605.1万套。2024年，智能驾驶行业能够继续创造3000亿以上人民币的规模；麦肯锡预计，到**2030年**中国自动驾驶相关的新车销售及出行服务将创造**超过5000亿美元**的收入。

2.具身智能：飞速成长下的前景与挑战

具身智能这一概念最早可以追溯到1950年，图灵在其经典论文Computing Machinery and Intelligence中就曾提到，在这篇论文结尾图灵展望了人工智能可能的两条发展道路：一条路是聚焦抽象计算；另一条则是为机器配备传感器使其与人类交互学习。这两条道路逐渐演变成了非具身和具身智能。近两年来，随着深度学习和虚拟环境技术的发展，具身智能的研究进入了全新的阶段。

(1) 具身智能行业自诞生起即呈现鲜明特点

- **赛道极新**：2023年，英伟达CEO黄仁勋表示，人工智能的下一个浪潮是具身智能，彻底引爆行业热点。目前具身智能赛道玩家的平均成立时间不超过3年，“年轻”是这个赛道的最大特点
- **增长极速**：2024年，随着具身智能是AI继续向物理世界深度渗透的必要路径成为共识，赛道持续展现出强劲的增长势头。随着技术突破的不断取得、应用方式的持续深探，预计未来几年内，市场规模将持续快速增长
- **市场前景期待极高**：2024年下半年以来具身智能赛道投融资事件较上半年增长5倍，重金押注显示出资本市场对这一赛道的高度关注和极高预期。具身智能赛道目前的头部创企之一智元机器人，成立于2023年2月，在一年多时间内完成7轮融资，市场估值已光速跃过10亿美元门槛

源码资本执行董事陈润泽表示，在具身智能领域，短期内技术还无法满足人们对通用机器人的预期，但在一些细分市场，产品化的机会已经逐渐出现。

(2) 玩家格局中“国家队”“地方队”成中坚力量

2023年年底以来，各地方政府及科研机构纷纷牵头成立具身智能实验室或创新中心。一系列具身智能“国家队”和“地方队”目前已成为行业格局中创业玩家之外另一股不可忽视的中坚力量。



图：具身智能“国家队”及“地方队”

(3) 技术路径共识形成，行业落地方向更清晰

当下具身智能赛道中机器人主流形态存在两个方向的技术路线：

- **未来趋势型**：两足加灵巧手人型机器人
- **阶段实用型**：轮式加抓夹人型机器人、四足机器人

前者是目前行业共识中**更先进、更未来**，也更接近人类真实形态的技术路线，目前**成熟度尚处早期阶段**。而后者可以看作向前者演化过程中的**阶段性形态**，目前技术**更为成熟稳定**，离落地更近。

当下几乎所有“国家队”、“地方队”玩家以及大部分年轻的创业玩家都会采取**第一类技术路线**，向更未来的技术路径和产品形态发出冲击和挑战，不断在两足行进稳定性、速度、地形泛化运动能力、灵巧手泛化抓取等能力上寻求突破。同时，也有不少创企尤其是成立时间较久的机器人企业在**第二类技术路线**上积累更深厚，具备较为成熟的量产落地产品，后一形态也距离BC端落地应用、实现商业化更近。选择兼顾两条路线的企业也不在少数。

(4) 行业面临技术挑战、落地速度不及预期

目前限制具身智能进一步发展渗透的**主要难点**可以分为以下三方面：

技术方面：

Sim2Real存在GAP	从仿真环境到真实世界的迁移（sim2real）存在挑战，例如摩擦力等力学环境系数模拟难度大，迁移过程会造成性能下降问题
泛化性挑战巨大	目前具身智能的泛化性挑战主要存在于感知能力泛化、抓取能力泛化、地形识别能力泛化等，增强知识转移和在复杂环境中的泛化能力是至关重要的

现有大语言模型和多模态大模型的成功首先依赖互联网上海量的文本、视频数据；相较而言目前具身智能研究中海量具身数据是相对稀缺的，收集成本也高很多。

——清华智能产业研究院（AIR）

数据方面：

现实世界数据获取存在瓶颈	大规模真实数据的采集成本高昂，且难以全面采集所有可能场景的数据。创建多样化的真实世界机器人数据集需要各个机构之间紧密且广泛的合作
合成数据泛化能力存疑	合成数据虽能补充真实数据不足，但与现实世界数据有差异，泛化能力存疑，且单纯依靠模拟数据会加剧仿真到现实的GAP问题

产业方面：

供应链硬件同质化

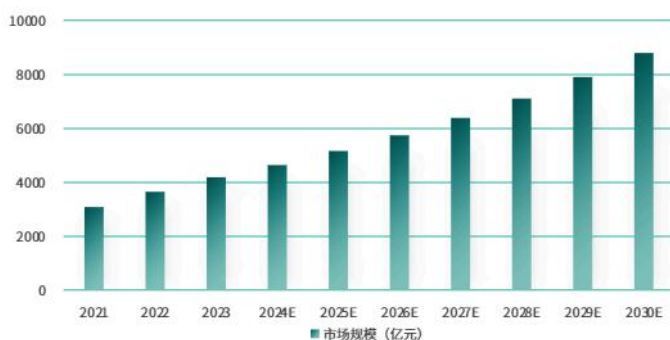
目前市场上的具身智能人形机器人硬件同质化严重，很多机器人公司使用的本体和其他组件都来自同一家上游供应商，这限制了人形机器人产品形态的多样化发展

落地速度不及预期

目前具身智能距离真正进入BC端的工厂、家庭等使用场景为人类创造真实价值至少还需5-10年，当下应用场景更多停留在科研和教育领域。这与市场非理性的过高预期产生了一定落差

(5) 前景广阔，长远来看有望实现颠覆性变革

2024年具身智能行业在技术进步、市场规模扩大、应用场景拓展等方面取得了显著成就，同时也面临着技术挑战和数据问题。随着技术的不断发展和产业链的完善，预计具身智能将在更多领域实现商业化落地，为社会带来更多颠覆性的变革。



图：具身智能市场规模及预测，36氪、量子位智库

据《2024年具身智能产业发展研究报告》显示，中国具身智能市场规模已达**4186亿元**，并有望在未来几年内保持快速增长态势，具身智能市场前景广阔。预计到2030年，中国市场规模可达**8000亿元以上**。

具身智能创业公司的长期成功可能不取决于科研和技术实力，而是在于理解技术边界和定义产品的能力。

——陈润泽，源码资本执行董事

毫无疑问人形机器人是承载我们的AI所有想象的最佳硬件载体，之前普遍认为这一天的到来还很遥远。现在看来这一天的到来要比预计快的多。

——将门创投

3.智能硬件：PC、手机、眼镜、耳机四大硬件

与其他行业不同，智能硬件行业中的产品多数在生成式AI涌现以前，就作为智能设备融入人类的生活：电脑、手机、耳机……与从前产品在计算能力、功能性上的更新相比，AI的加入对这些设备的更新是**革命性、彻底性**的，将完全改变人类使用相关电子产品的**场景、方式**和**需求**。峰瑞资本合伙人陈石表示，搭配模型能力的AI硬件产品正成为中国团队的优势领域。

对于智能硬件行业，我们可以从具体的产品变化中见微知著：

(1) AI PC：具备本地AI能力的个人电脑

① 与传统电脑相比，AI PC的不同

我们可以明确的是，AI PC最重要的特征即具备在**本地处理**和**运行**AI模型能力，它通过集成专门的硬件和软件，使AI技术的应用更加广泛和高效。明势创投表示：硬件已成为AI落地的原生场景，Apple Intelligence，可能是第一个日活过亿的AI应用。传统的个人计算机注重的是计算能力和硬件性能，而AI PC引入深度学习、自然语言处理、计算机视觉等AI技术，并在此基础上更加注重**智能化应用**。

关键技术：

- 神经处理单元（NPU）：**NPU**是为AI和机器学习任务而设计的专用硬件加速器，能够显著提高语音识别、图像处理和其他AI任务的效率。



- 本地运行AI模型和应用：与传统的需要云端处理的AI服务不同，AI PC能够在没有网络连接的情况下**本地运行**AI模型和应用，这提高了隐私性和便捷性。

② AI PC的显著优势

从**技术发展**角度来说，由于配备了NPU，AI PC相比传统个人电脑，在处理AI相关任务时效率更高、准确性更强。此外，AI PC在实时响应能力、低延迟、高安全性等方面有所提升。

从**应用场景**角度来说，AI PC新增的应用场景将包括AI聊天机器人、AI PC助理、AI Office助手、AI本地知识库、AI图像视频处理以及AI PC管理等，这些场景的落地将完全改变下一代用户的**电脑使用习惯**。

AI PC的出现，标志着个人电脑从传统的计算终端升级为**个人AI应用中心**，它将更好地理解用户需求，提供更智能、个性化的服务。这必将改变PC用户的使用习惯和场景，以前所未有的方式推动PC产业的发展。

(2) AI手机：真正打造出个人智能系统

与AI PC相似，AI手机也可被认为是集成了AI技术的智能手机，它从体系结构出发，将AI深入融入手机各个模块。对AI手机来说，最值得被关注的功能应当是自我学习、自我优化，甚至预见用户需求，为用户提供更加精准化、智能化的服务。腾讯研究院表明，以Apple intelligence为代表的端侧模型及其服务会加速商业化。

① 更适配AI时代的硬件升级

- 芯片：如iPhone 16搭载了专为AI设计的A18芯片，提供了强大的处理能力和能效，同时支持Apple Intelligence。
- 处理器与内存：AI手机通常配备了更高级别的处理器和更大的运行内存，以支持复杂的AI算法和运算需求。
- 传感器配置：配备了更多的传感器，以收集更多的用户数据和环境信息，为AI算法提供更丰富的输入数据。

② 软件智能化和用户体验的全面升级

- 操作系统：集成自然语言处理、图像识别等AI的操作系统，更好理解用户需求，实现更加智能的任务分配和资源管理。
- 应用程序：通过AI智能化升级应用程序，同时提供智能助手、语音图像识别、AI摄影、智能文档等功能。
- 个性化：以用户的使用习惯和需求作为数据输入，进行智能推荐和优化。

作为当下人类最无法离开的智能设备，AI手机的以上升级能够提高手机与用户的适配程度，并提升用户的生产力和工作效率，使手机成为一个真正的**个人智能系统**和**生活助手**。OpenBMB开源社区提到，像AI手机类的智能硬件运行端侧模型能创造全新的场景和功用，它将成为主流消费电子新的竞争焦点。

(3) AI眼镜：空间计算落地关键消费级硬件

2024年，Ray-Ban Meta眼镜发售销量突破100万台，成为首个消费级爆款产品。AI眼镜成为当前智能硬件赛道最引人关注的产品形态，充分验证高消费者需求和高市场接受度，也向厂商和资本展示了巨大潜力。

多类传统的穿戴类消费电子，得益于端侧AI的发展，未来在信息提示、通讯、导航等方面同样有很大的想象空间，也是最可能出现爆品的产品方向之一。

——杨颜媛，梅花创投

① 2024年迎来爆发，初步形成3路径

目前的AI眼镜主要侧重于通过AI提升语音交互能力，集成耳机、摄像机、导航、语音助手等功能。主要可分为**AI**、**AR**和**AI+AR**三种技术路径，分别以Ray-Ban Meta、Meta Orion和Rokid Glasses为代表。

产品	技术路线	企业	发布时间	产品特点	售价/元
Apple Vision Pro	AR	苹果	2024.01.19	混合现实头显设备，M2芯片，2300万像素屏幕，支持眼球追踪和手势操作	29999
OPPO Air Glass 3	AI	OPPO	2024.02.26	轻量级双目全彩AR眼镜，集成OPPO的AndesGPT大语言模型，树脂衍射光波导镜片，显示亮度高达1000尼特，支持触控手势操作，可通过软件更新获得导航、提词器、快速健康、健身信息预览等功能	4999
XREAL Air 2 Ultra	AR	XREAL	2024.05.30	提供高达330英寸虚拟屏幕，重量仅72克，兼容多种设备，支持大部分新款手机和游戏设备	3999
智能拍摄眼镜 A1	AI	闪极科技	2024.05.31	采用紫光展锐旗舰级AI芯片和索尼1600万像素背照式摄像头，支持实时在线、听音和视觉体验	999
界环AI音频眼镜	AI	北京蜂巢	2024.08.08	内置先进的AI芯片与音频模块，实现语音命令控制音乐播放、接听电话或获取导航信息等智能功能	799
小度AI眼镜	AI	百度	2024.11.12	首款搭载中文大模型的原生AI眼镜，具备第一视角拍摄、边走边问、识物百科等六大功能	-
Rokid Glasses	AI+AR	Rokid	2024.11.18	与BOLON眼镜合作，采用衍射光波导成像技术，接入通义千问大模型，兼具AR眼镜、耳机、AI助手和相机的多方面能力，支持物体识别、文字翻译、数学题解答等功能	2499

国内多家厂商纷纷推出AI眼镜，产品形态的火爆。影目、闪极、雷鸟、小米等厂家预告将在年底和明年初发布相关产品，可以预见AI眼镜的产品爆发**远未结束**。

② 用户体验尚需突破，产品天花板有待探索

尽管市场增长迅速，但当前AI眼镜在用户体验侧仍存在问题。如软件内容和硬件性能不足、佩戴舒适性差和续航短等。为用户提供高清、舒适、可持续佩戴的视觉体验是不同厂商持续进行技术演进的目标。

目前AI眼镜正处于快速发展和迭代的前期阶段。总体来看，从AI语音眼镜迭代到**AI+AR眼镜**是当前智能眼镜的主要发展路径。这一产品形态当前远未触及天花板，我们将持续关注技术及应用开发者生态的完善。

我们可以看到，各大AI终端品牌在积极准备，为明年上量做准备。

——石建平，蓝驰创投投资合伙人

(4) AI耳机：便捷交互，延伸大模型应用场景

耳机硬件因其**便携性**、**轻量化**和**产品成熟度高**的特点，成为最早接入AI能力的智能硬件之一。

① 市场增长明显

AI耳机市场近年来呈现出显著的增长势头。依据洛图科技数据，2024年8月中国在线电商平台的AI耳机销量同比增长763.3%，销售额增长近**14.5倍**，预计全年销量有望突破**20万副**，同比增长488.7%。

② 3大应用场景

目前AI耳机的细分产品场景主要有**生活娱乐**、**办公助手**和**音质提升**三种：

场景	产品	企业	发布时间	产品特点	售价/元
生活娱乐	Ola Friend	字节跳动	2024.10	Ola Friend接入豆包大模型，与豆包APP深度结合。用户可以通过语音唤起豆包进行对话，提供信息查询、旅游出行、英语学习及情感交流等场景的帮助	1199
办公语音	iFLYBUDS Pro 2	科大讯飞	2024.05	iFLYBUDS Pro 2能够自动识别并转写会议内容，支持32种语言的同传翻译，内置智能会议助手，能够自动记录会议要点、生成会议摘要，新增“语音嘴替”功能，可模拟个性化语音	1399
音质提升	ARC 3 音弧开放式AI耳机	Cleer	2024.04	ARC 3 音弧是一款开放式AI耳机，采用不入耳设计，支持无损或高解析度音频的高效传输，具备AI降噪、健康监测等功能，首次在开放式耳机中实现了杜比音效和杜比头部跟踪	1399

③ 潜在大模型硬件生态布局入口

AI耳机已成智能硬件布局的重要**切入点**，字节跳动、科大讯飞等大模型玩家纷纷进入这一领域布局。

但主流产品**尚未**集成独立的端侧AI模型，主要采用**云端**路径唤醒大模型，**无法脱离手机应用软件使用**。更大的意义在于增加模型能力调用入口，并**延伸**其沉浸式使用场景。将门创投表示，当前的硬件环境并不能承载我们对大模型应用的所有想象。

目前这一赛道厂商竞争尚不充分，市场需求在早期阶段。

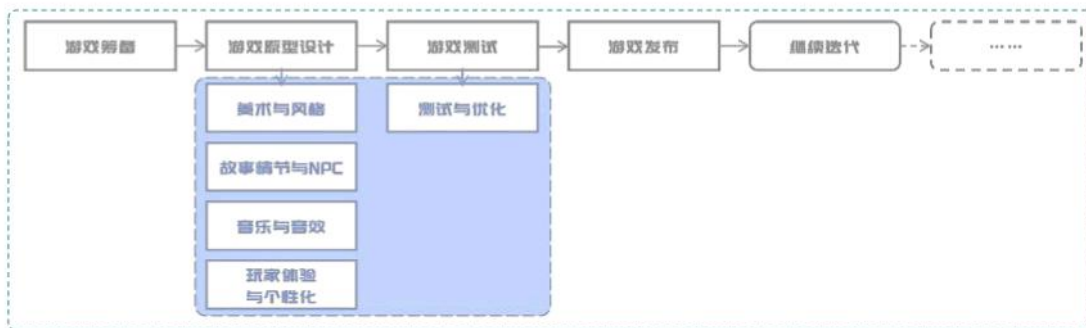
4. AI+游戏：施展想象力的最佳乐园

作为最有价值的应用场景之一，AI技术在游戏行业的应用可以追溯到上个世纪80年代。最初，人们只是使用程序生成简单的内容和随机元素；后来，AI开始辅助游戏设计，例如自动化任务和基础的对话生成；随着技术能力的发展，生成式AI技术在游戏中的应用逐渐扩展，达到动态世界生成的能力；到今天，大模型等关键技术的进步使AI能够生成更复杂的游戏内容，如动态场景、智能NPC行为等。

AI在游戏行业的渗透有以下几点优势：游戏开发者天然亲近AI，并乐于使用AI领域的各类技术实现设想中的游戏世界；同时，我们不断强调数据在当前时代AI发展，尤其是大模型发展中的重要性，来自游戏世界的数据大幅减少了筛选、清洗的成本；此外，游戏还能成为AI模型能力的最佳试验场。

(1) AI革新游戏创作

首先，AI对游戏的影响是工业级的。每一款游戏都萌芽于一个充满灵性的想法，随后制作出模型、草图、故事梗概和游戏风格。



图：AI从5大方面对游戏创作 workflow 产生积极影响，《AI游戏产业驱动力研究报告》

通过简化游戏创作的工作流，我们可以把当前AI对游戏的影响归纳于以下5个方面：

① 美术与风格：2D美术与3D生成

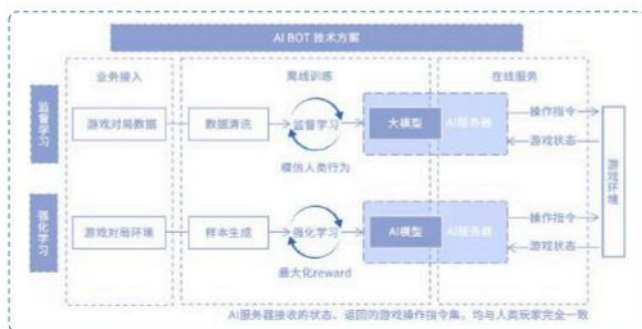
在2D美术中，我们广泛地使用图像生成技术。AI生成图像领域中，生成模型和神经网络渲染是两个主要的技术分支。2022年主流生成式图像应用DALL-E 2.0、Stable Diffusion和Midjourney的诞生，使图像生成的精细化程度得到了有效的提升。提示对图像生成结果的控制能力得到了提高。

在3D模型生成中，AI主要完成了对3D模型生成流程的“一步到位”。



图：3D美术工作流拆解，《AI游戏产业驱动力研究报告》

工作流中的每一个环节几乎都需要一位或一组3D美术工程师来完成，而使用3D生成模型可以直接完成一个可调整的3D粗模，大大提升效率。



图：AI BOT，超参数科技

② 测试与优化

AI在游戏测试环节中存在3大模拟场景，分别是AI玩家模拟、AI NPC和AI游戏模型测试。

在优化环节往往承担改善角色动画、编程&加速、自动化游戏测试等功能。通过生成文字、图片、视频、音乐来缩减时间和资源消耗，也能快速识别游戏开发过程中的基本问题，并提出改善方法。

③ 故事情节与NPC

游戏设计是整个游戏开发过程的核心，AI在角色与故事生成、游戏机制创新等方面，依托最基础的自然语言生成用法。生成式AI阶段，游戏开发者在决策式AI的基础上融合更多思维和方法用以训练智能NPC，通过针对性训练数据集，强化NPC模型的基础能力，如角色知识、对话能力、情节演绎和逻辑推理。目前，小型的智能NPC可以实现封包在游戏中本地运行。

④ 音乐与音效²

AI生成音乐的存在2种主流技术路线：基于乐理规则的**符号生成模型**和基于音频数据的**音频生成模型**。开发者正在使用AI生成音乐来填充游戏过程与游戏UI中需要使用到的各类音效、不同游戏场景中用以渲染氛围的各种音乐。

AI生成音乐作为音乐资产在游戏制作和发行环节使用都是非常可行的，像MusicLM等模型已经支持生成多音轨的作品。使用AI生成音乐为原型、佐以专业制作人的协调，将使AI音乐更快进入游戏制作与发行的生产线。

⑤ 玩家体验与个性化

AI能基于玩家游戏行为评估玩家技能水平和游戏风格，同时动态调整游戏难度，增加或降低敌人的数量和强度，改变游戏环境等；不断收集的玩家数据，还能使NPC和游戏系统更加适配玩家水平。

同时在游戏运营过程中，**AI客服**和**AI分析**也是提升玩家体验的重要一环。

(2) AI带来新的游戏

我们关注到，许多充满灵感的开发者正在尝试将AI作为游戏玩法的一环，这正是游戏行业**从AI+走向AI原生**的表征之一。

2023年Genfun.ai和Meshy联合制作的游戏《Soul Chronicle》，在当时是首款实时**3D+AIGC+UGC**的MMO手游。最大突破是首先制作出了一种与游戏完美融合的3D AIGC技术，在游戏中可以实时生成角色皮肤。

2024年Bitmagic释出了他们推出的能直接创建“游戏世界”的平台——《Roleverse》的最新成果，在平台内可以使用提示在游戏内定制角色，对角色进行缩放、挤压和拉伸，也可以轻松地对游戏世界进行编辑。



图：《Soul Chronicle》（左）和《Roleverse》（右）

(3) AI促进游戏产业变革

AI技术在游戏开发中的应用逐渐落地，为游戏行业提供了新的增长空间。《2023中国游戏产业报告》表明2023年我国游戏市场实际销售收入为**3029.6亿**人民币，同比增长**13.95%**；截至2024年11月，累计收入已达**2681.16**亿元，市场回暖力度加强。

同时自2022至2023年，中国游戏用户规模达到**6.68亿**，同比增长**0.61%**，也是自2020年增长率放缓以来的**首次进步**。不能忽视的是，AI在游戏开发中的逐渐落地为游戏行业提供了**新的增长空间**，成本、效率和质量同步上升的时代已来。

² 对AI生成音乐的进一步研究，可以阅读《AI音乐应用产业报告》

5.AI+影视：AIGC工具掀起内容生产革命

影视行业是指涉及电影、电视剧、综艺节目等影像艺术创作与传播的产业领域。移动互联网时代之后，这一行业的创作、分发和消费方式已被深度重塑，为AI技术的落地创造了基础。明势创投表示，在影视、游戏、虚拟现实等领域，对高品质、实时生成的视觉内容需求强烈，市场潜力巨大。

目前，AI技术的应用已逐渐渗透到产业各环节，影视行业在AIGC内容的冲击下，正经历着一场内容生产、分发和消费模式的深刻革命。

(1) 内容创作、分发和消费方式变革创造AI落地土壤

移动互联网时代以来，数字化技术已经渗透到产业从拍摄到后期的各个环节，而互联网内容平台的兴起则为内容产品的分发和消费提供了全新的形式，这些变革都为AIGC技术的丝滑落地提供了先决条件。

我们在《[AI视频生成研究报告](#)》中提到，视频现已成为移动互联网最大的内容消费形式。2023年以来，以Sora为代表的一批视频生成产品陆续问世，短短一年多时间，AI视频生成工具已经越来越深入地推动着影视行业供给侧技术变革，随着AI技术能力的不断突破，可以预见这一变革将会加速深入。

梗图生成视频、AI视频特效等弯道超车具象化且具备传播性的视频工具获得用户青睐，降低了小白非专业用户的使用门槛，易于使用和接受。AI微短剧将是未来发展的大趋势。

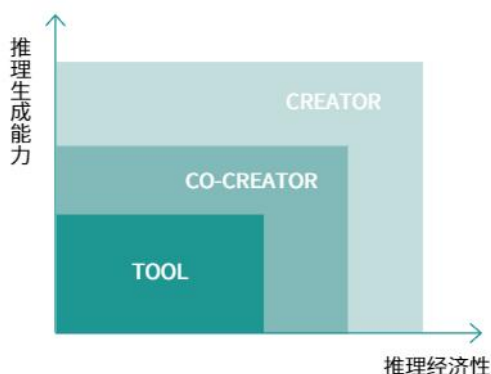
——杨锦洲、韩莹莹，中国联通研究院

降低生产成本	AIGC技术可在不同场景实现不同程度替代实际拍摄和物理特效，大幅度降低制作成本
提高生产效率	在需要快速迭代和大规模生产的领域，如广告、社交媒体内容等，AIGC工具的优势更加明显
个性化内容生成	AIGC技术能够根据用户需求生成个性化内容，满足不同观众的个性化需求
解放创作者劳动力	AIGC工具完成大量基础内容生成工作，让创作者可以将精力集中在更具创造性的工作上，提高整体创作质量

(2) AIGC技术发展阶段决定行业渗透率演化

从技术发展角度，目前决定AIGC工具在影视行业应用渗透率的关键因素有两个维度：

- **推理生成能力**：具体包括生成时长、耗费时间、生成清晰度、内容稳定性、丰富度、连续性、指令理解及遵从能力、对生成内容的控制能力、对物理世界规律的深度理解等
- **推理经济性**：AI视频生成成本与非AI工作流成本齐平并逐步发展到显著低于现有成本，持续降低推理成本数量级



图：技术发展双维度决定应用渗透率阶段

目前渗透率演化阶段根据场景有所区分，主要卡点可以概括为：**生成内容的不可控性**。具体包括形象不一致、动作不流畅、表情不生动、复杂提示词难以完全实现，以及随着时长增加逐渐出现明显不符合物理规律的动作发展的问题。

(3) 长短视频场景均有应用，渗透率分档明显

短视频——短剧/TVC生产流程可实现AI工具全替代：

在短视频领域，短剧和TVC由于其自身特殊性，成为最早落地AIGC工具并率先形成完整全AI工作流解决方案的内容生产场景。

时长短 剪辑节奏快	短剧和TVC市场在分钟级和秒级，对视频生成工具单次生成时长、连续性、稳定性要求更低，可通过工程手段提升并衔接以实现商用效果
内容质量要求相对较低	市场对短剧质量定位相对电视剧、电影等内容更低、更容易接受内容存在一定的瑕疵
市场包容性强	短剧诞生时间短、定位偏下沉市场、属于内容“快销品”，而TVC则不属于观众需付费消费的内容产品，对广告投放商来说实现商业露出效果即可，因此市场对这两类内容的包容性更强，更能接受技术创新早期造成的“颠簸”

长视频——电影/电视剧/动画开始渗透，AI辅助人类进行内容生产：

在长视频领域，AI技术尚停留在工具层面，对 workflow 实现部分替代、降低成本的效果，暂时无法提供全流程全AI替代传统工作方式的解决方案。

剧本创作与优化	AI可以辅助生成剧本草稿或提出情节改进方案，甚至预测剧本票房潜力，帮助投资决策
概念图分镜生成	可以快速、高效、低成本批量生成精准分镜头和高质量视觉概念图
特效动画制作	AI在特效领域的应用历史最久，能够生成逼真的CGI效果、进行数字绘景、动画制作等
后期制作	AI在剪辑、音效设计和色彩校正等后期制作中发挥重要作用，能够自动识别最佳剪辑点，智能拼接镜头，提升整体制作效率
虚拟角色与数字演员	AI技术能够创建逼真的虚拟角色，模拟演员的表演，拓展影视创作的可能性

总体而言，越接近于综合性思考、策划层面，对于AI来讲就越难，而越具体的工作越容易被取代。

——陈宇，《满江红》编剧

(4) 市场前景广阔，5年内有望达千亿级市场

我们在《AI视频生成研究报告》中，对影视行业市场规模做了粗略估算。

	长视频				短视频	
平台	iQIYI	腾讯视频 不负好时光	bilibili	电影	抖音	快手
用户	5亿MAU	4亿MAU	3.4亿MAU	N/A	8亿+MAU	7亿+MAU
年营收	320亿	100亿+	230亿	550亿（本土）	1500亿	1135亿

图：2023年海内外影视市场规模估计，《AI视频生成研究报告》

如图所示，国内影视市场规模在**2023年**约可达**3835亿元**，若假设2027年AI影视市场可以获得上述国内总市场份额的**10%**，则国内AI影视总市场规模预计将达约**380亿元以上**；若假设2030年可以获得**25%**以上市场份额，则国内AI影视总市场规模将达**千亿级别**。

6.AI+营销：AI让只为一入打造广告成为可能

自20世纪80年代改革开放以来，中国营销行业经历了从传统到数字化再到智能化的变革。最初，营销依赖电视、广播和报纸等传统媒体。随着互联网和社交媒体的兴起，以SEO、内容营销、社交媒体营销、视频营销和直播营销等形式为主的数字营销逐渐成为新趋势。而到了今天，**AI不仅提高营销内容生成的效率和质量，也使得营销更加精准和自动化**，如通过机器学习分析消费者行为，预测市场趋势，以及自动化执行营销任务等。

目前，**营销行业或成生成式AI最早实现商业化落地的行业之一**。未来，AI技术还将持续推动营销行业的深刻变革。



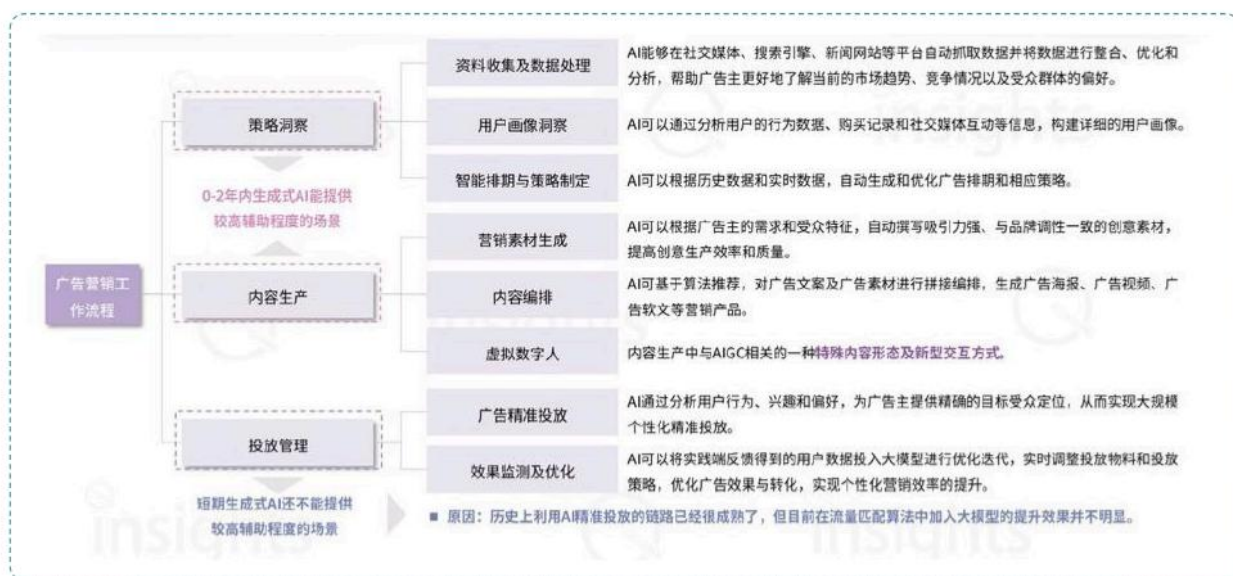
图：营销行业或成最快实现生成式AI落地的领域，《中国AIGC广告营销产业全景报告》

(1) AI营销行业呈现全新特点

- **回归营销本质：**广告主营销需求和用户的需求能进行直接匹配，从「人与机器」单向交互，回归到了「人与人的交流」
- **流程扁平化：**以往广告营销全流程需要在内容、设计、定向出价等每步环节做优化，生成式AI可以一步到位进行内容生产与投放，减少信息损失
- **从「局部最优」到「全局最优」：**从传统的广告推荐模型，变成以数据驱动的大模型为核心，实现全局最优的广告分配，提升了系统投放的分配效率

(2) AI技术变革营销策略洞察及内容生产投放方式

广告营销领域的工作流程与环节涉及策略制定与用户画像洞察、广告内容的制作、投放渠道的管理与效果分析等工作，需要消耗大量的时间，其中较多的工作都可以通过AI代替人工的方式实现效率的提升。



图：AI营销时代工作流程实现深入变革，《中国AIGC广告营销产业全景报告》

(3) 有效提升营销行业生产力

• 创新广告形式，带来全新交互体验，提升ROI

广告的内容形式以及触达消费者的形式均得到极大的创新。随着AI生成及分发技术的深入发展，预计未来将持续改变广告行业，升级各种类型广告形式。

• 提高营销内容生产效率，实现精准链接和高效触达

过往营销行业受制于生产力天花板只能做到「一面千人」，AI营销的新趋势是制投一体，即所有的内容生产都是根据投放触点及人群对应生成并进行精准投放。用户的个性化需求与企业自身的品牌、产品通过AI大模型能力实现全局最优触达，做到「千人千面」。

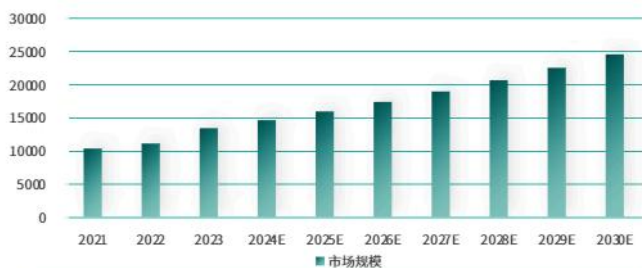
• 推动营销服务商商业模式革新，优化企业工作流

营销服务商开始通过大模型及AI生成技术优化企业内部整个工作流程，打通全链路数据，探索新的商业模式，提供新的产品、服务、场景，从而降本增效，提升毛利率，给用户带来全新体验。

(4) AI引擎助力广告营销行业加速成长

我国经济持续稳定发展，AI成为新增长引擎，带来广告内容的更新迭代和需求增多，企业为应对市场竞争，更好地满足消费者需求和提升自己的品牌，实施专业化营销策略的动力得以增强。

根据国家市场监督管理总局、中国国际公共关系协会等统计数据，经量子位智库测算，我国广告营销市场规模预计2030年将接近两万五千亿。如果假设5年后可以获得10%以上市场份额，则2030年国内AI营销总市场规模可达**千亿级**。



图：2018-2025年中国营销行业市场规模统计及预测

7. AI+教育：因材施教的最佳实现方式

教育行业因其**内在需求**与大模型所擅长的能力高度适配，已成为模型层公司、教育科技公司、智能硬件公司等一众玩家的AI落地方向。



图：AI教育行业图谱，《AIGC教育行业全景报告》

(1) AI在教育领域的价值

我们需要特别提出的是：除了对生产力和效率的提升，AI在教育领域最重要的价值在于使教育界不变的理念——**因材施教**，有了规模落地的可能。

在**教师侧**，AI能够帮助教师生成教学素材、自动批改作业、个性化教学分析并提升教师个人的专业水平和素养；

在**学生侧**，AI能够帮助学生规划学习路线、智能辅导并持续跟踪学习进度，反馈个性化的评估和调整结果；

在**市场侧**，AI也助推了各大教育硬件和学习类APP销量的激增；

最终，通过加强AI智能体与教育的结合，AI也将推动**教育普惠**的实现。

(2) AI+教育落地场景

① AI学习机（更多面向K12群体）³

当前国内AI教育玩家多数选择在学习机中搭载大模型的方式，完成AI与教育的融合。这类产品的出现为用户提供了**增量价值**，提高了产品销量，同样也为学习机等硬件产品构建了新的竞争壁垒。

② 语言类学习App

以软件切入AI教育赛道的玩家，通常选择接入通用大模型，加以自身积累的教育数据进行微调训练。语音合成等技术为语言类学习App带来对话体验的全面升级，使其成为目前大模型应用于教育的产品中市场反馈**最积极**的品类

③ AI教育智能体

技术与教育融合，对教育的两个核心角色——教师与学生产生了颠覆式的效果。AI教育智能体能够融入到教师和学生的日常中，在提供教学帮助的基础上，改变二者工作和学习状态。

• 充当陪伴式家教的角色

AI智能体为学生提供个性化教学主要体现在课后阶段，能够实现有效的**一对一**学习辅导。具体来说，AI智能体能起到个性化分析指导和启发式互动的作用，与传统的搜题软件相比，使用者将不再受提问形式的局限。在生成式AI之前，教育领域采用个性化学习是一件**成本极高**的事情。

³ 对学习机等教育硬件的进一步研究，可以阅读《AI教育硬件全景报告》



图：AI教育智能体成为陪伴式家教，《AIGC教育行业全景报告》

此外，口语家教也是AI教育智能体的一大落地场景。口语对练既能避免LLM为学习带来的弊端，如照搬答案、弱化思考力等，还能够以更便捷、成本更低的方式提升语言技能。

• 充当助教角色

降低了教师的学习成本，为其提供几乎零门槛的使用方式。这在一定程度上能够释放教师的生产力，将更多的精力用于提升学生素养上，是从“解惑”到“育人”的转变。

(3) AI+教育的严肃要求

教育是一个严肃领域，并且对学生的**价值观**形成起到引导作用。因此，教育领域的AI应用对于AI生成内容的**准确性**、**安全性**等有非常严格的标准。

教育行业对AI大模型的幻觉问题，必须保持**零容忍**的态度和不断降低幻觉问题出现的技术改进目标。教育领域的大模型，**最关键的是确保模型回答的准确性以及价值导向**。

这些标准都对人类介入有结构上的需求，这也为AI在教育行业的落地增加了**工程**上的难度，因此也阻止了AI在教育场景的完全渗透。

但也和大多数AI+行业的类似，教育是面向人类的事业，关乎人类的生存、发展和自由。寻找更有效率的方式使AI成为教育行业的助手，减少使用和融合过程中的磨损，才是AI融入该行业的正确方向。

全球范围内对AI技术潜在风险和道德影响的日益关注。随着AI系统变得越来越复杂和强大，确保它们的安全性和符合伦理标准变得至关重要。推动人工智能技术的发展，安全优先，同时保证人工智能性能长期发展，未来强调安全与性能平衡发展，强技术驱动、全流程优化、多主体参与以及敏捷治理。

——韩莹莹，中国联通研究院

在国内，网易有道、学而思、华东师范大学等企业及高校纷纷推出教育大模型。学术界和企业界，都对AI与教育的结合充满信心。

(4) 未来的AI+教育行业发展趋势

①跨学科理解构筑教育大模型高阶能力，将成为模型层玩家核心竞争力；同时教育大模型向多模态演进，落地产品趋于丰富；

②技术层面可以进一步关注AI教育智能体与具身智能的融合概念，实现AI家教的实体化。

8.AI+医疗：诊疗场景率先落地，助推医疗普惠

AI医疗指的是利用人工智能技术，如机器学习、自然语言处理和深度学习等，对医疗流程进行智能化管理和提升，提高医疗服务效率和质量的过程。中国AI医疗的发展历史最早可以追溯到20世纪80年代初，经历几十年的发展后，在生成式AI实现技术突破、国家政策大力支持、人口老龄化程度增加、市场对AI技术接受度提高等多方因素影响下，**AI医疗迎来发展利好**。

(1) 生成式AI技术突破驱动智慧医疗深度革命

医疗行业智慧化、数字化进程由来已久，但推进过程中存在一些顽固卡点，如**数据留存及共享难**、**药物研发成本高**、**诊疗效率低**、**医疗资源配置不均衡**等问题。生成式AI技术实现进一步突破之后，再次深度颠覆了这一进程，给医疗行业带来下一阶段的**全新底层技术驱动力**。

现阶段AI技术对于传统医疗行业的赋能主要体现在：

- **数据收集整理打通**：过往医疗数据往往分散在不同的医疗机构数据池中，且数据的收集和存储结构各异，难以在保证数据安全前提下有效留存及共享数据；AI技术的推广有助于解决数据互联互通和共享问题
- **医疗行业模型打造AI医生**：AI技术使医疗知识库能够快速积累结构化多模数据、垂直医疗大模型持续学习医疗知识，逐步打造AI医生，为后续科研成果涌现提供基础
- **解放医护劳动力**：减轻医生重复性工作负担，同时借助AI力量规避人为错误；规范化、标准化医疗流程，降低人为操作风险；实现基层医疗机构远程会诊，提高服务覆盖率
- **全面提升医疗科研、诊断和治疗效果**：利用AI分析患者的基因组学数据和生活习惯，制定更加精准的治疗方案；提升医疗影像、病理识别的准确性和效率，AI软硬件辅助疾病预防、手术、康复，提高药物研发效率，降低成本

(2) 诊断及治疗阶段率先实现AI落地

目前AI技术在医疗行业在**预防**、**诊断**、**治疗**到**康复**的全流程多场景中均有渗透，产生了许多全新的应用场景，同时大幅提升了许多既有场景的医疗效率和医疗效果。



图：AI医疗应用场景

在上述场景中，以诊断阶段的AI医学影像、AI问诊、AI病理诊断，治疗阶段的AI制药成熟度最高。

① 诊断阶段

• AI医学影像：医学影像数据与AI能力天然匹配

医学影像场景最主要的生产资料形式是视觉数据，这一场景也是所有医疗场景中数据产出最丰富的，在所有临床数据中占比达**80%**以上，使得医学影像数据成为**最早**实现全球标准化的医疗数据类型之一。医学影像具备易获取、结构化、处理难度小等特点，**是AI技术率先实现突破的应用场景**。

AI医学影像集中于**病变检出、识别**，以及**良恶性判断**等，目前应用已相对广泛，主要优势体现在MR加速成像、XR质控及阅片、DSA图像增强优化及辐射剂量减弱等方面。

• AI问诊：提升问诊效率与准确性

中国AI问诊行业在2013年起步，由于人们对线下医疗的依赖性，整体发展规模较小。随着经济环境的变化，为解决居民无法去线下就医的问题，国家鼓励推进AI互联网问诊。2017至2022年中国AI网络问诊市场规模由**3.75亿元**增至**85亿元**，市场规模呈指数级攀升，中国AI问诊行业进入高速发展时期。

在这一场景中，借助AI技术可以实现导诊、医疗问答、诊断建议、科普宣教等医疗服务，主要包括：

自然语言处理 (NLP)	AI问诊系统通过自然语言处理技术理解和处理人类语言，识别患者的症状和描述，提取关键信息、识别理解，并给出合理的医疗建议
机器学习	机器学习算法使AI系统能够不断学习和优化自身的诊断能力，通过深度学习等技术提高诊断的准确性和效率
知识图谱	知识图谱帮助AI问诊系统构建医疗领域的知识库，将医学知识以结构化的方式存储和表示，方便系统进行查询和推理
医疗大模型	医疗大模型可以更机动灵活地跟患者对话，获取和整理患者的症状、基本信息、过往用药史等信息，并基于专业医学知识图谱和机器学习算法，调用覆盖全科的预问诊模型能力给出可能的诊断建议

• AI病理：促进计算病理学发展

计算机视觉、分子病理学、基因组学和生物信息学的快速进步促进了计算病理学的加速发展。计算病理学通过量化癌症组织病理学提取大量的生物学和临床相关信息。AI算法提供了从大量数据中提取信息的框架，因此支持计算病理学有望改变未来癌症的诊断、研究和治疗方式。

诊断方法	传统病理	AI 病理
观察工具	显微镜	显示屏
观察空间	显微镜下观察	具有网络以及显示条件的空间即可
远程病理	物理切片传输，耗时费力	数字切片传输，实现远程诊疗
保存介质	石蜡切片，占用库房等物理空间	使用网络存储空间
保存时间	时间越久，切片质量越差， 档案管理复杂	可永久保存，随时调用， 不会因为时间原因影响质量
阅片速度	100-200片/天	500-600片/天
准确率	较高，与医生从业经验、疲劳度等有关	高

图：传统病理与AI病理对比，中泰证券

① 治疗阶段

• AI制药：提高创新药物研发转化率

医药研发是整个医疗产业中的重要一环，当前新药研制主要受到：**高成本（10亿+）**、**长研发周期（10-12年）**、**低成功率（13.8%）**三座大山限制工作流程。

AI制药目前主要将NLP、机器学习、深度学习、自然语言处理、计算机视觉、生成模型等AI技术与传统制药环节做结合，通过数据交叉比对、加速筛选、从头生成等方式，提升新药研发效率，拓展药物创新空间的技术应用。主要应用于药物研发中的药物发现、临床前阶段，与此同时，AI向临床开发阶段的渗透也正在加快。



图：AI制药流程

(3) AI医疗市场规模前景巨大

AI技术在医疗领域的应用不仅能够增强医疗服务的精准性和个性化，还能为医疗决策提供数据支持，促进医疗行业的创新进步。有助于优化医疗资源分配，提升诊疗的准确性，降低成本，并推动医疗服务的普及和便捷性。根据GMI统计，2032年AI医疗市场规模将达到**700亿美元**。

九、AI行业渗透率：数据基础决定初速度，用户需求成为加速度

（一）AI行业的变革与渗透

在本阶段，通过联系前文对代表行业的重点分析，我们思考并回答了两个问题：

- ① 如何判断一个行业是否存在AI发展的空间？
- ② 如何分析行业受AI影响的程度？

1. AI行业影响力图谱 因子影响力

因子 行业	行业与发展			用户与需求				安全与政策	
	数据	原生指数	需求	迭代	感知	意愿	生态	安全	政策
智能驾驶									
具身智能									
智能硬件									
游戏									
影视									
营销									
教育									
医疗									

在该图谱中，影响程度被分为以下几个等级：

强烈	较大	中等	较少	微弱	较少	中等	强烈
----	----	----	----	----	----	----	----

图谱中，颜色越深越有利于AI技术在该行业的变革和渗透；颜色越深代表由于对安全因素的高要求，AI技术在该行业的变革和渗透存在阻碍。

（1）模型因子解释

行业与发展：

- **数据：**指行业文本及视觉数据的密集程度和结构化程度，数据资源越密集、结构化程度越高的行业接入AI能力的

门槛越低、速度越快

- **原生指数：**指「行业发展」与「AI技术发展」的伴生关联性，具体考察两者关键诞生、发展、转变的节点时间相关性强弱
- **需求：**供给侧需求，指行业产品或服务供应商通过AI实现降本增效需求的迫切性，通常商业化水平越高、市场竞争越充分的行业利用新技术实现降本增效的经济回报越大，需求越迫切
- **迭代：**行业迭代，指行业提供的产品及服务的迭代速度和频率，通常迭代速度越快、频率越高的行业越重视用户反馈数据，对利用AI技术实现实时、自动化工作流程，以更快了解产品及服务迭代方向，取得相对竞争优势的动机越强

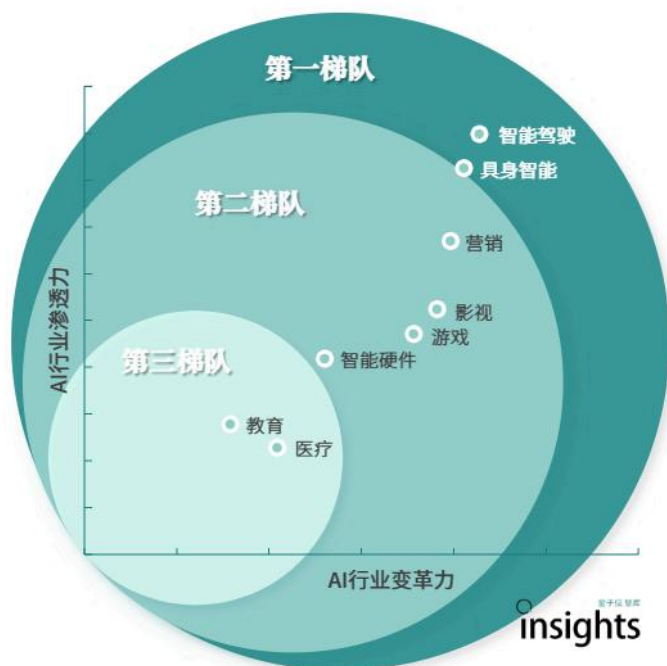
用户与需求：

- **感知：**消费者感知，考察AI技术加入后，消费者对行业提供产品及服务带来增量体验的感知程度以及优化提升幅度
- **意愿：**涉及行业内用户对AI进入的接受意愿、用户对AI技术变革需求的强烈程度及紧迫度。通常行业主要用户群体互联网基因强烈时，行业对AI的接受意愿较高
- **生态：**考察AI在行业内的需求场景创造能力；提高旧需求的效率仍然意味着蚕食旧的市场，但创造新的需求往往能带来新的增量

安全与政策：

- **安全：**涉及对AI生成内容精准性、价值观对齐等方面的要求；目前，技术伦理问题日益凸显，数据隐私、算法偏见、AI道德责任等问题，都要求我们协调技术与人类福祉、社会环境之间紧张的伦理关系；在更强调安全需求的行业，AI进入的工程难度会显著提升
- **政策：**政策支持对AI技术在行业内的发展具有重要作用；通过提供资金支持、税收优惠、市场准入放宽等措施促进行业的创新、发展和竞争力提升；通过出台试点方案、条例等规范行业发展轨迹；政策因素对AI技术在行业内渗透的促进和限制作用同时存在

2. AI行业生态位



根据AI行业影响力图谱，我们将代表性AI行业分为3个生态位阶段：

首先，在**智能驾驶与具身智能行业**，源于对AI的紧密需求和强伴生性，在行业发展与用户需求的多项指标上都表现出强烈的关联。具身智能行业由于处在行业发展的前期，未表现出明显的供给侧需求和用户感知，但仍然是**受AI变革和渗透最强的行业之一**。

其次，在**第二生态位**的代表行业中，**营销行业**的生产资料以互联网数据为主，这一特质使其**AI技术接受能力高、落地速度快**。自搜索引擎竞价营销时代起，AI搜索算法、AI内容推荐算法已2次重塑营销行业的产业链。极高的商业化程度和充分的市场竞争属性决定在这类行业，AI能有较高的变革和渗透力度，潜在的高经济回报驱动下，行业拥有更强劲机拥抱新技术。

同时，数字人等营销技术有效升级了用户体验，而在以需求发掘、智能匹配、获客成交为目的的行业，审美并非第一要素，也使**用户接受AI技术的意愿更高**。

游戏和影视行业同为消费型内容产业，主要市场为C端消费者。其中游戏行业尽管诞生时间更晚，但与AI技术的**相关性**却更强。

在发展过程中，游戏和影视行业都利用AI技术实现内容生产的降本增效，同时重新归纳整理了原有的工作流。值得一提的是，游戏行业对**收集用户反馈数据**并及时迭代的需求较影视行业更强，因此会有更强的**动机**利用AI技术实现自动化工作流程和即时产品迭代。

在消费者对AI应用效果的感知程度和接受度上，游戏行业AI生成数字资产、对话、NPC等AI落地场景提高了游戏的沉浸感，用户反馈更积极；影视行业对AI内容的落地集中在短视频，长视频内容审美等级高、技术难度大，AI生成内容容易受到更多负面反馈。

受制于AI技术能力，以及这类行业对内容产出的要求，目前在绝大多数场景下对人工的替代能力有限，因此需求尚未到达爆发阶段。

可以预见是，智能硬件行业的产品将会蚕食原本行业产品的市场份额，并倒逼行业整体的升级。在该行业，AI通过**创新应用场景**和改善用户体验，提供前所未有的竞争力。但在当前，广泛存在用户需求不足的问题，这类行业对新生态的规划能否在实际使用中得到实践，将完全改变现有的行业格局和业态。

最后，**教育和医疗行业**作为国民经济的基础行业，在每一个时代都积极吸纳当前时代最先进的技术，以改善行业的发展。尽管这些行业没有太多的互联网基因，但由于其普惠性，在利用和发展AI技术时往往会受到较大的**政策支持**。同时，行业的特殊性也带来了**对安全的高要求，高标准**，造成了AI渗透行业的**工程化障碍**。因此，这类行业将处在AI行业变革和渗透的**第三生态位**。

十、AI创投：投融资马太效应明显，国家队出手频率提升

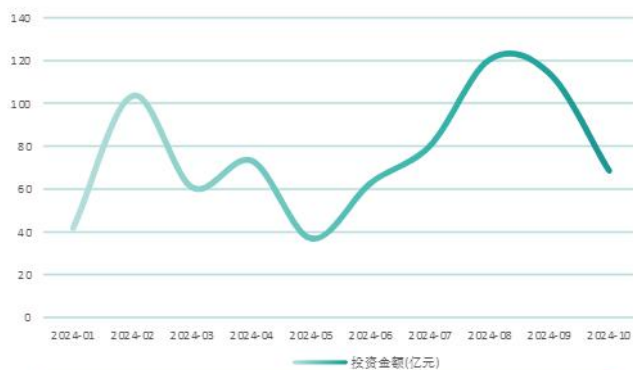
（一）我国AI行业投融资分析⁴

在全球风险投资持续低迷的2024年，AI行业仍然展现出**逆势增长**的能量。依据烯牛数据统计，2024年1月至10月，国内AI行业发生融资事件**908件**，总融资金额**762.23亿元**。相比去年同期融资事件数量同比减少**19%**，总金额增加**29%**。

这一变化主要反映出行业整体投资回归理性、机构出手更加谨慎，风险偏好总体更趋保守，使得资金流向**集中于热点赛道**，也更**偏好对规模较大、已获得市场验证的头部创企进行单笔更大金额加注**。同时，由于AI行业整体偏快的演化节奏、行业成熟度的快速提升，部分**明星项目的大额融资事件**也在客观上助推了投资总金额的水涨船高。高榕创投表示，今年AI投资整体大幅攀升，全球范围内AI仍旧是最强吸金赛道，AI大厂团队剥离再创业等孕育新的投资标的。



图：2024年国内人工智能融资事件数量月度折线图，烯牛数据

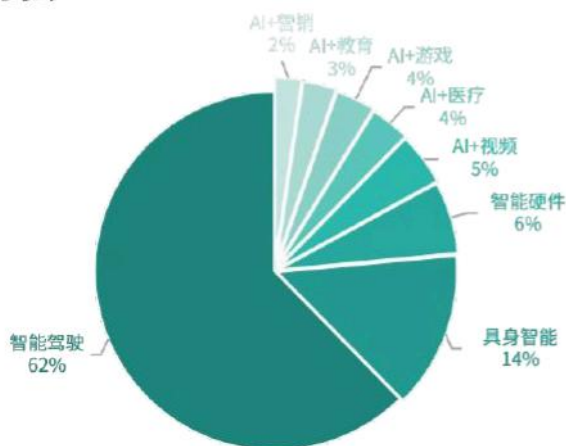


图：2024年国内人工智能融资总金额月度折线图，烯牛数据

1.马太效应明显，热点赛道、头部创企资金引力强

从重点关注的8个AI垂类赛道的投融资情况来看，仅**智能驾驶**单一赛道的投资事件数量达**上百件**，**已超过其他全部赛道总和**，总投资金额达**数百亿人民币**，与其他赛道远不在同一数量级。赛道成熟度最高，仅达到C轮后融资轮次的企业就有**14家**。

同时，这一赛道的明星热点投资事件也最多，4月份有桐乡国投、金合股权投资、民生股权基金**50亿**投资**哪吒汽车**，8月份还有塞力斯**115亿元**入股**引望智能**。而仅仅10月内就有**文远知行**纳斯达克敲钟，市值超**50亿美元**；**滴滴自动驾驶**获由广汽集团领投C轮**2.98亿美元**融资；**地平线机器人**港股年内最大IPO，市值**超660亿港元**。堪称今年智能驾驶赛投融资高光时段，**为智能驾驶市场注入了巨大信心与活力**。



图：8类AI垂直行业投资事件数量分布，烯牛数据

⁴ 本部分投融资数据来源烯牛数据，数据统计截至2024年10月，仅统计国内（不包含港澳台）、不包含并购、IPO、IPO以后阶段的投资事件；融资金额是估算数值；投资方参与投资金额、FA服务金额为该轮次的投资金额。

AI原生应用已经涌现一大批5亿美金以上估值甚至独角兽的公司，并且获得了PMF验证，达到了千万美金以上的年收入。

——联想创投

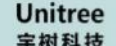
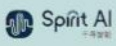
2024年国内智能驾驶C轮后投资事件

被投企业	时间	轮次	金额	投资方	投后估值
 轻舟智航 Q-CRAFT	10.28	C+轮	数亿元	逐鹿资本	-
 WeRide 文远知行	10.25	IPO	4.4亿美元	未披露	约42.47亿美元
 地平线 Horizon Robotics	10.24	IPO	51.42亿港元	未披露	超660亿港元
 滴滴	10.22	C轮	2.98亿美元	广汽集团、滴滴出行	约40亿美元
 EXD 数集智能	9.23	Pre-B轮	数千万元	北汽产业投资领投、上海石雀跟投	-
 昆仑智驾	9.20	C1轮	数亿元	衢州智盛产投、Prosperity7 Ventures、中泰仁和、华登国际	-
 易航智能	9.5	C轮	数亿元	北汽产业投资、财通资本、浙江金融控股、德清县产业发展投资基金	-
 LiangDao	6.28	C轮	未披露	亦庄国投、招商资本	-
 NOVAUTO 超星未来	5.9	Pre-B轮	数亿元	安徽省铁路发展基金、博华资本、龙鼎投资、海南天智、陕汽智能汽车基金、讯飞创投、德载厚资本	-
 卡尔曼	4.11	C轮	数亿元	百联挚高、创新工场、深投控、数字万融、微木资本等	-
 EACon 易控智驾	3.12	C++轮	3亿元	兴杭国投、紫金矿业、郑州高新区产业发展引导基金	-

 格陆博科技 GLOBAL TECHNOLOGY	1.29	C+轮	未披露	渝富资本、世嘉闻华、旭辉投资跟投	-
 爱行科技	1.10	Pre-IPO	1亿元	北斗七星	-
 robosense	1.5	IPO	9.85亿元	未披露	约190亿港元

除此之外，**具身智能**赛道也是今年人工智能领域另一大热赛道，投资事件**二十余起**，数据相比去年同期基本**翻倍**，总金额也在**数十亿**人民币级。下半年以来具身智能赛道投融资事件频发，其中逐际动力、帕西尼感知科技都已完成A轮融资，**知行机器人、宇树科技、云深处科技**更是来到了BC轮次，体现出赛道投融资阶段逐渐成熟。而仅几家公司融资金额**上亿**的企业——包括**星海图、星动纪元、宇树科技、逐际动力**等的融资总额就占到了具身智能赛道的一半以上，体现出赛道内**头部企业资金引力较强**。

2024国内具身智能亿级投资事件

被投企业	时间	轮次	金额	投资方	投后估值
 星动纪元 Robot Era	10.16	Pre-A轮	3亿元	清流资本、元璟资本、阿里巴巴、策源资本、联想创投、世纪金源、金鼎资本、泽羽资本、清控天诚	20亿元人民币
 Unitree 宇树科技	9.13	C轮	数亿元	上海科创基金、中信证券投资,中关村科学城、京国瑞基金、红杉中国、琥珀资本 Ambrum Capital、美团龙珠、祥峰投资	80亿元人民币
 NOEMATRIX 穹彻智能	9.6	Pre-A轮	数亿元	Prosperity7 Ventures、广发信德领投,泽羽资本、创新工场、奇绩创坛、Plug and Play、MFund魔量资本等	约70亿元
 XINSHIJI 新时记	8.23	Pre-A/Pre-A+轮	1亿元	基石资本、德联资本、英诺天使基金、啟赋资本、南山战新投、九合创投	-
 Spirit AI 千寻智能	8.12	天使轮	2亿元	弘晖基金领投、达晨财智、千乘资本、顺为资本、绿洲资本、	-
 逐际动力 LIMX DYNAMICS	7.15	A轮	数亿元	阿里巴巴(09988)、招商局创投、尚颀资本领投、峰瑞资本、绿洲资本、明势资本跟投	-
 UDEER AI 有鹿机器人	5.31	股权投资	1亿元	创新工场、元璟资本、百度风投等	-

与具身智能赛道类似的还有**智能硬件**，在李飞飞带热空间智能概念、加诸国内以字节Ola Friend、雷鸟X2-Lite眼镜多个C端明星产品延续去年Apple Vision Pro和Ray-Ban Meta智能眼镜热度的共同助推下，智能硬件赛道今年迎来投资热。与去年同期相比，这一赛道投资总金额增长**超六倍**，主要归功于四起融资金额**过亿**事件——**Rokid**、**影目科技**获亿元投资，**雷鸟创新**获4亿元投资、**龙旗科技**获15.6亿元投资，合计投资金额占总金额已**超九成**。

2024国内智能硬件亿级投资事件

被投企业	时间	轮次	金额	投资方	投后估值
 Rokid Pioneer of AR	9.30	股权投资	1亿元	武汉市江夏科投	超10亿美元
 雷鸟	8.22	B+轮/B++轮	4亿元	南湖金服，南湖国控集团，惠合资本，惠山科创等	约15.5亿元人民币
 影目科技	7.26	B轮	1亿元	弘芯基金，华腾普益，江西中寰	-
 龙旗 LONGCHEER	3.1	IPO	15.6亿元	未披露	-

2.投资热度向更晚达到TPF的赛道转移

AI+营销等在2023年更早实现TPF（Technology Product Fit）的赛道今年投资热度呈现下降趋势，**投资总额相对减少**。而AI+教育、AI+游戏、AI+医疗等更晚达到TPF的赛道在今年迎来**投资总额增长**，其中**AI+医疗**的投资额达去年同期**5倍以上**，体现出行业共识随AI技术及行业成熟度的改变正发生着快速变化，机构**对技术难度更高、壁垒更强的行业赛道逐渐展现出更强的投资兴趣**。

在**其余**赛道中，超千万融资事件**13起**，其中包括**3起超亿元**融资事件——阿里巴巴**2亿元**人民币投资**精准学**，顺为资本、鼎石资本、领沓资本、石溪资本**数亿元**人民币押注**潞晨科技**，蚂蚁集团**1亿元**领投**爱诗科技**。**行业整体投资热度持续上涨**，重点行业押注金额相较去年存在不同程度提升。

(二) 科技合作伙伴观点精选

最具代表性的AI技术如何从研发走向广泛落地、创造实际价值？



腾讯研究院

Tencent
Research Institute

能提供用户想要的服务，能有当下企业能够承受的投入产出比。

腾讯研究院

3个关键点：

- 1) 广泛的数据基础：例如现有大语言模型和多模态大模型的成功首先依赖互联网上海量的文本、视频数据；相较而言目前具身智能研究中海量具身数据是相对稀缺的，收集成本也高很多。
- 2) 达到产品级能力的底层AI技术：例如大语言模型强大的文本理解和生成能力，自动驾驶系统在复杂路况下的安全驾驶，技术能力突破可产品化临界点才可支持广泛的落地。
- 3) 可实现数据-模型优化-服务能力提升的闭环产品形态：AI产品或平台在落地后，本身可形成新数据采集-模型不断优化-服务能力提升的闭环，例如ChatGPT可利用海量用户的交互数据进一步优化模型，提升能力；Tesla通过收集海量用户驾驶数据，进一步迭代其FSD自动驾驶系统，改善用户服务。形成闭环的产品形态对AI技术从研发走向广泛落地具有重要意义。

清华智能产业研究院 (AIR)



联想
创投

- 1) 端到端解决用户的实际问题，提供完整的解决方案。
- 2) 产生用户数据闭环，积累私有数据资产，从而更了解用户。
- 3) 熟练使用和微调最先进的大模型，在数据和资金都完备的情况下考虑训练自有大模型，构建自身早模型层的技术壁垒。

联想创投

我觉得关键的要点是做到「技术为先、场景为重」。

首先你要有深度技术认知和能力，预判后续0.5-1年内AI的能力边界和发展趋势，其次但更重要也是最难的是，要做出你的落地场景和产品定位的选择，既要跟模型大厂保持安全距离，又要避开跟该领域当前领先企业的正面冲突，找到合适的细分市场迅速发展，最好可形成一定有壁垒的先发优势。

根据历史上互联网和移动互联网的发展经验看，我们建议创业团队要优先选择「非共识」的领域和场景，因为「非共识」可以成为创业公司早期的保护伞，且坚持做「正确而非容易的事情」。

陈石，峰瑞资本投资合伙人

FREE^S FUND
峰 | 瑞 | 资 | 本

最具代表性的AI技术如何从研发走向广泛落地、创造实际价值？



最关键的点还是理解用户，理解他们的具体痛点，避免自嗨型创业。

陈润泽，源码资本执行董事



一方面，大模型的成本下降，会为AI产品和应用的爆发进一步奠定基础；其次，期待大模型走通“最后1公里”，解决掉最后10%的难题，将真正解决深度的实际问题，在虚拟和现实世界都爆发更大的潜力。

高榕创投

AI技术从研发到落地一般会经过“技术创新-应用探索-工程化”的发展路径，AI技术落地的关键点是将通用的AI技术与落地的场景进行工程化适配的过程。

AI通过以下改进技术适应工业领域特殊需求，解决融合技术问题：

一是提高数据可用性，通过数据扩充、引入先验知识（模型预训练与迁移学习）、网络模型结构优化（弱监督/无监督）解决小样本问题。

二是提高可解释性，基于特征可视化的方法，通过得到特征与结果之间的因果/相关关系；基于逻辑/规则/知识的方法，利用决策树、工业知识图谱等可解释的模型构建黑盒模型的局部/全局近似来实现解释。

三是提高实时性，聚焦工业端侧推理的(半)定制化芯片探索，当前基本集中在冯氏结构局部优化路径；或是利用知识蒸馏、模型剪枝与量化等提高模型效率。

郭耀隆、夏璠，中国联通研究院



广泛落地靠入口，除了手机和电脑，有没有新的落地硬件场景；创造实际价值靠垂直打通，不要文本输入文本输出，要Prompt输入解决方案输出。

明势创投

核心是找到当前技术能够带来明显用户价值的场景，做好产品。很多技术有进展，但还没有到可以解锁用户场景的时机；另外一方面有的技术已经足够好，但还没有足够好的产品形态。比如说从GPT-3到ChatGPT，模型的尺寸并没有提高太多，核心创新点是把auto-complete的大语言模型对齐成了可以遵循人类指令，进行聊天对话的产品形态。

戴雨森，真格基金管理合伙人



最具代表性的AI技术如何从研发走向广泛落地、创造实际价值？

走出屏幕。AI技术要走向广泛落地，实现提效的初衷关键之一是要走出屏幕，实现与现实世界更紧密的融合与互动。

AI技术的「走出屏幕」仍然受制于多方面的限制，包括硬件层面性能与能耗的平衡、AI芯片、设备小型化、集成化等；软件层面包括算法的优化、复杂情况的决策速度、软件的兼容性与更新；在部分场景对数据的质量与安全性也提出了更高的要求。

未来的AI技术在2C端一定有更自然的交互方式，在2B端体现出卓越的提效与降本。在海量的数据中，人们的信息获取更高效，数据得到治理，而非无效、垃圾信息更多，也是一个具有实用性的关键点。

杨颜媛，梅花创投



PLUM VENTURES
梅花创投



AI能力提升、成本下降、效率提升，抓住关键场景里未被满足的用户需求。

王晟，英诺天使基金合伙人

AI技术要落地、创造实际价值，有两个关键：

一个是底层支撑AI技术的硬件，要向AI靠拢，定制化、深度融合，特别是Transformer架构的大模型。

大模型要从线上的一问一答的Prompt，走向普通用户的正常生活，走向千行百业，它最佳的落地形式，是AI硬件。需要支持性能更强、多模态、更大参数尺寸的Transformer大模型。但硬件的瓶颈比较明显，芯片、内存，续航，发热等。

二是模型本身的知识密度，在更小的参数量上释放更强的能力。端侧模型由日常主力消费电子承载，不依赖网络连接，高频推理没有成本，端侧比云侧离用户更近，AI最好的落地形式是端侧AI硬件。

而AI硬件要能真正走向落地，创造实际价值，除了硬件能够支撑模型流畅运行，模型本身轻量化高性能是根本。

多种模态端到端的交互、输出、AIGC内容生成，模型的知识密度越高，它能渗透的硬件种类就越多；从广度和深度两个维度，AI覆盖的用户场景就越多，用户粘性和刚需也就越强。

OpenBMB开源社区

OpenBMB

将门

生态的丰富多样是AI发展的重要基础。

将门创投

（三）我国AI及AI落地行业政策分析

从整体性意见指导和各地落实政策2个角度来看，我国在AI及AI落地行业的政策体系呈现出以下2大特点：

1. 长期关注AI在各行业的落地，积极推进AI原生行业发展

对我国AI相关发展规划和路径的研究，呈现了长期以来我国对AI技术及落地的关注和重视。

进入到21世纪20年代，我们发现政策更多倾向AI在千行百业的落地进展和各AI原生行业的发展。2020年国家互联网信息办公室提出，要**利用人工智能技术提升社交平台的信息审核和管理能力**；2021年《“十四五”国家信息化规划》中提到，要**推进信息技术、智能技术与教育教学融合的教育教学变革**；2022年，北京出台了首个数字人产业专项支持政策；2023年北京又发布若干措施支持本市研发机构**聚焦大模型、类脑智能、具身智能、价值与因果驱动的通用智能体**等方向。

资金支持、政策优惠、人才培养和引进、技术创新和应用推广都将促进AI在各行业的落地。

2. 各城市依托自身发展，优质政策吸引AI人才与企业落地

同时，北京等城市正在积极推动AI加速赋能千行百业，大力推进大模型技术创新与行业的深度融合。

北京发布的《北京市推动“人工智能+”行动计划（2024-2025年）》提到，要**依托首都优势行业资源和科技研发能力，围绕机器人、教育、医疗、文化、交通等领域组织实施一批综合型、标杆性重大工程，促进大模型核心理论与技术突破，增强人工智能工程化能力，提高重点行业的科技水平和服务质量，构建跨行业、跨领域协同创新组织模式，形成大模型行业应用新生态。**

上海在《上海市教育领域推动大规模设备更新专项行动计划》中强调教育硬件全面升级时，特别提到**在人工智能等关键技术领域的硬件建设。**

武汉在《武汉市智能网联汽车发展促进条例（草案）》中对**智能网联汽车检验检测产业发展，建设智能网联汽车封闭测试场、检验测试平台，构建智能网联汽车试验与测试评价体系**等的意见对Robotaxi等智能驾驶项目在武汉市的发展起到关键的促进作用。

.....

AI技术在迅猛发展之下，已成为推动经济社会进步的重要力量。AI在千行百业的落地也将吸引大量AI人才与企业落地，良性循环又将促进整个产业的繁荣与发展。

我国AI关键发展路径图谱 insights



超参数科技 Parametrix.ai 超参数

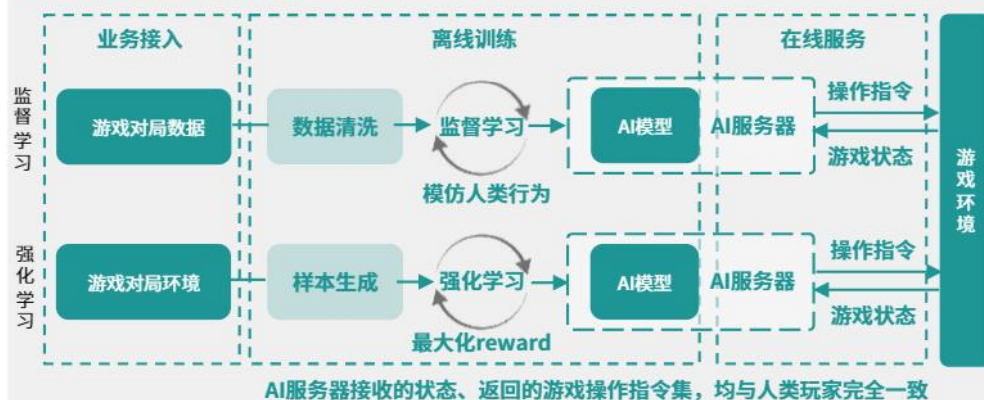
——AI Agent在游戏赛道的探索

超参数科技成立于2019年，是一家将AI与Game Agent深度融合的创新型科技公司。公司运用前沿AI技术、通过游戏场景提升Agent的自主性和互动性，打造Agent智能生态，为全球用户提供独特的互娱体验。目前，该公司旗下产品现已在多款千万日活的头部产品中持续稳定运营，不断推动AI与互娱产业的深度融合与发展。

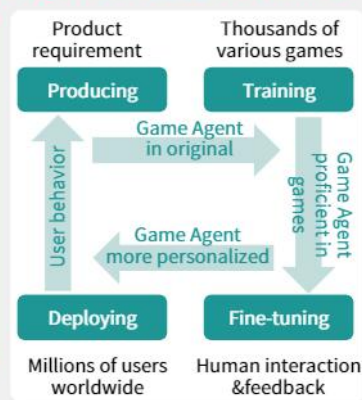
产品与技术创新：

运用前沿AI技术，通过游戏场景提升Agent的自主性和互动性，打造Agent智能生态，以Game Agent为载体，为用户塑造差异化、人性化、个性化的产品体验。

Game Agent技术方案：



Game Agent Produce Pipeline:



落地效果：

Game Agent作为虚拟玩家（队友或对手）参与游戏

- 赋能开发者降本增效、提升游戏质量、增强玩家体验，提升游戏运营效率与效果；
- 对游戏生态的长期规划和动态调控，实现游戏资源的有效管理和优化配置；
- 目前已为多款年流水10亿+产品提供AI服务，覆盖MOBA、FPS、SLG、RPGs等多个游戏品类；

Game Agent在游戏中扮演NPC

- 根据策划的设定及AI的自我发育，推动剧情发展、增强沉浸感、增加游戏的可探索性；
- 通过AI生成的动态场景与任务，以及NPC之间的动态社会关系模拟，衍生出新剧情故事，提供多样化玩法。

整体效果：

现已在多款千万日活的头部产品中持续稳定运营

Game Agent峰值在线数达153万

AI识破率低，99.5%概率认为AI识破率在(0,2%]区间

客户反馈：

日均服务次数**4000亿**次
累计服务人数超过**10亿**
累计服务时长**4.3万**小时

每天服务全球**60余**个国家数亿用户
客户满意率**行业第一**，续约**100%**

(1) 有效提升 DAU 增长率等关键指标

投放AI后，案例DAU次月增长最高达**2%**
整体增长率稳步提升，最高可达**10.2%**



案例A DAU 月度统计

(2) 优化玩家体验，互动量&满意度上升

玩家类型	互动量变化
新手玩家	77% ↑
活跃玩家	53% ↑
资深玩家	46% ↑

AI发起局间互动，中低段位玩家互动量提升50%以上，被点赞率提高**90%**；AI接管掉线玩家，掉线玩家客诉率相对下降**200%**

奥创光年



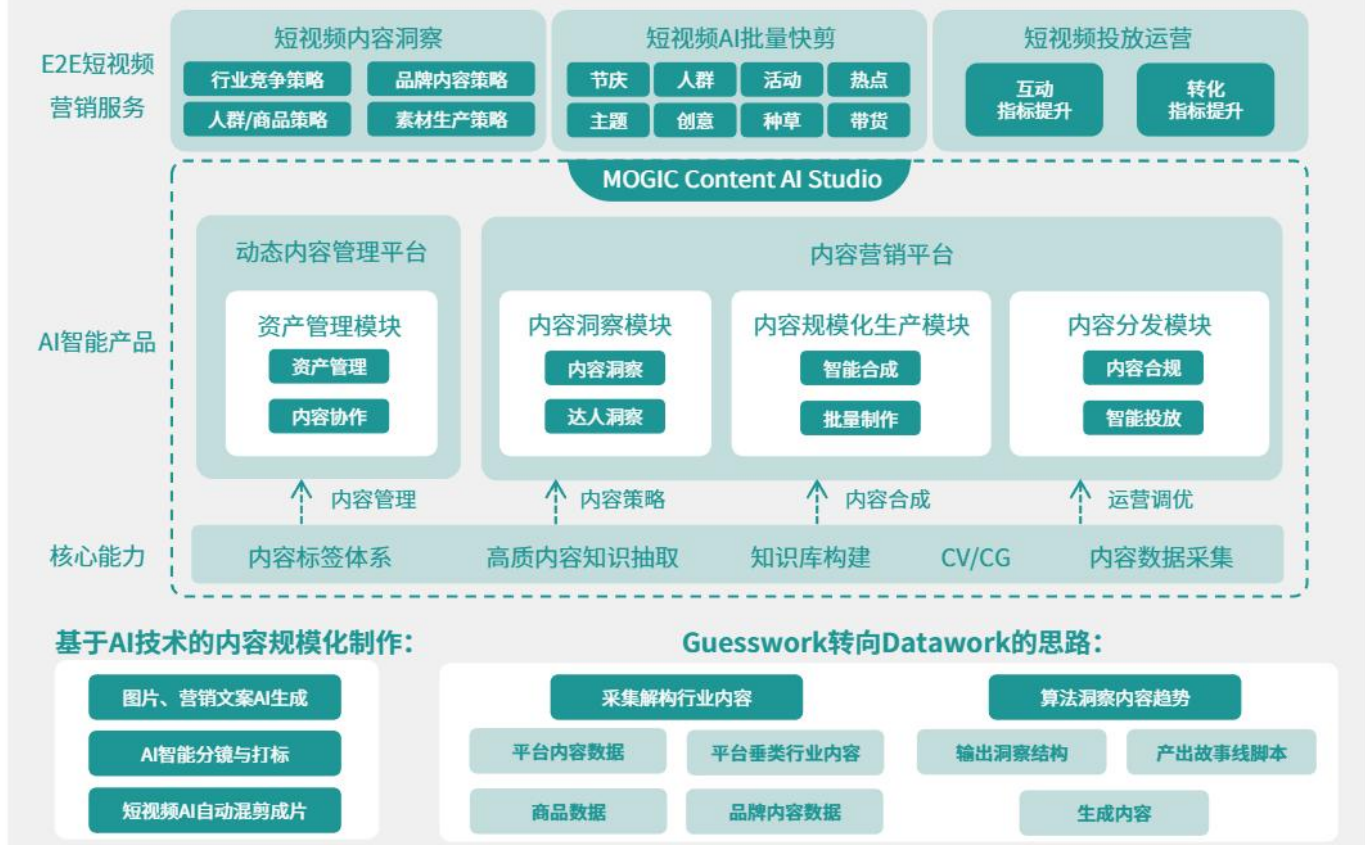
奥创光年

——懂技术懂品牌懂平台的AI营销之道

奥创光年「Mogic AI」，国内领先的AI智能营销科技服务商。2021年创立，为企业客户提供营销策略洞察、AI内容制作、AI智能投流内容管理等全链路营销解决方案。2024年，该公司持续深耕市场营销赛道，专业且高效地为用户提供了全面升级的品牌传播解决方案。

产品与技术创新：

主要提供E2E短视频营销服务、AI内容制作、AI智能投流到DAM内容管理的全链路营销解决方案。



落地效果：

短视频内容洞察——使用AI甄别最优内容创作方向：

- 通过抽取全网最优效果20%TOP内容，AI自动打标；
- 能够智能分析高质内容、平台流量逻辑；
- 输出内容策略，指导内容创作与投放运营。

效果：

跨平台内容洞察——AI较人工全面度提升3-10倍
周期更新创意热点——AI较人工速度提升80%
效果稳定——AI较人工效果数据提升2-5倍

客户反馈：

服务200+一线品牌
AI视频交付数量达200w+

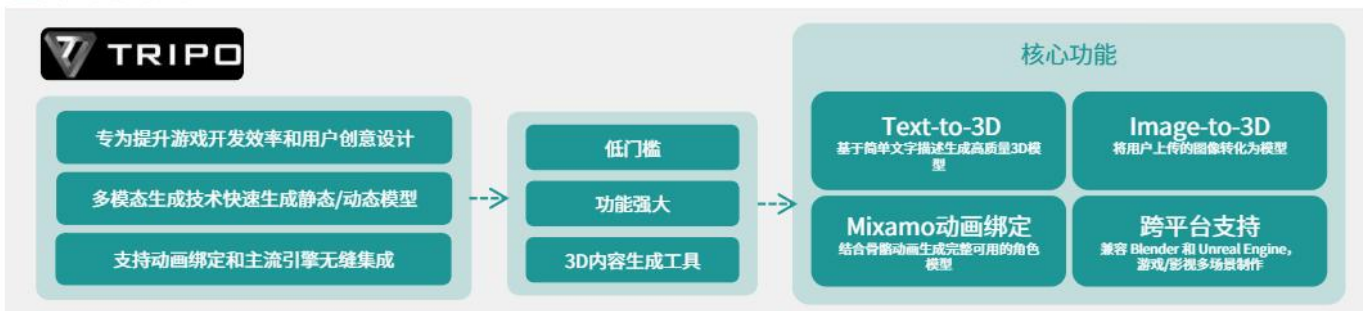
视频创意投放提升
G品牌 ROI +22%
L品牌 ROI +68%

达人素材自动化混剪
K品牌 ROI +10% CTR +30% 素材平均消耗 +20%
G品牌 ROI +30%

VAST VAST ——3D游戏玩法及资产AIGC全新解决方案

VAST 创立于2023年3月，是一家致力于通用3D大模型研发的AI公司，专注于打造大众级别的3D内容创作工具，建立3D的UGC内容平台，在3D原生、2D优化、多视角生成、编辑动态场景、人物生成等3D生成方向发布论文110余篇，被引13359次。目前已发布3D内容生成工具Tripo AI。

产品创新：



技术创新：



落地效果：

AI原生游戏开发：

两位开发者利用Tripo在短短两周内完成了基于Prompt的3D怪物生成大乱斗游戏，玩家可以用Prompt生成拥有战斗能力的3D怪兽，并与其他玩家的怪兽进行匹配对战使模型进化。这一案例充分体现Tripo满足小型团队高效快速开发需求的能力。



游戏资产高效制作：

通过Tripo快速完成JRPG 3D场景，结合MidJourney设计和Godot引擎，优化了建模到场景整合的工作流。



客户反馈：

社区小游戏TripoGo

- 上线 7 天内吸引了超10万用户，在TikTok韩国Roblox直播间广泛传播，分享率高达50%

游戏工作室

- 在Unity流水线中快速完成demos和辅助资产制作，生产时间减少70%

Scenario

- 使用Tripo完善风格化的游戏角色和资产生成流程，效率提高60%

大型游戏开发者

- 从文本和图像生成3D资产，快速创建原型并迭代游戏资产，成本降低30-50%

Yahaha

- 推出恐怖游戏模板，在引擎内实现UGC游戏npc自动化生成

游戏美术工作室

- 快速生成多种材料和环境资产，交付效率提高60%



Rokid Rokid ——消费级AI+AR眼镜先行者

Rokid创立于2014年，是一家专注于新一代人机交互技术的产品平台公司。目前致力于AR眼镜等软硬件产品的研发及以YodaOS-XR操作系统为载体的生态构建。公司通过语音识别、自然语言处理、计算机视觉、光学显示、芯片平台、硬件设计等多领域研究，将前沿的AI和AR技术与行业应用相结合，为不同垂直领域的客户提供全栈式解决方案，AI、AR产品已在全球80余个国家和地区投入使用。

产品与技术创新：



落地效果：

自然、符合本能的交互方式

支持头动接听，点头接听电话，摇头拒接电话

接入AI大模型能力

可实现智拍照答题、多语言翻译、AI快速回复消息、AI实时导航

高刷新率、巨幅屏幕

高刷新率和巨幅屏幕创造沉浸式智能体验

运动防抖功能适应更多场景

通过逐帧追踪技术，确保在运动或颠簸环境中画面保持稳定

硬件设计实现舒适性突破

智能瞳距调节、近视人群友好、舒适软性镜腿、超轻空气鼻托

多屏展示多任务处理

可同时实现社交软件沟通、长视频播放、办公软件使用等功能

客户反馈：

用户

海内外AR眼镜服务用户突破250+
海内外AR眼镜购买用户破30万+
用户日使用时长2h45m+

场景

海内外文博市场占有率99%+
车载AR娱乐市占率70%

开发者

海内外AR注册开发者6200+
海内外AR企业开发者2500+

生态伙伴

钉钉 BOLON
iQIYI 爱奇艺 支付宝
淘宝 bilibili

井英科技 CreativeFitting —AI短剧创作及内容平台

CreativeFitting（井英科技）成立于2021年，是一家专注于视频生成模型探索及AIGC娱乐应用的公司，AI短剧行业领军者。发布自研AI短剧一站式创作工具，为创作者打造了AI短剧创作新模式。发布全球首款AI短剧App Reel.AI，致力于成为全球最大的AIGC内容平台。

产品与技术创新：

Reel Diffusion：一站式AI短剧创作工具

「Reel Diffusion」叙事型视频的生成能力达到全球行业领先水平。从算法到训练数据及工程实现都为短剧创作进行专门设计，尤为擅长生成激发用户情感共鸣的短片，帮助创作者讲述引人入胜的故事。



Reel.AI：全球首款AI短剧应用



「Reel.AI」于2024年首次在海外发布，迅速登上娱乐应用畅销榜，随着生成式AI的飞速发展，AI生成内容在部分场景已经可以达到可消费级内容水准，带来AIGC Super App巨大机会。

Reel.AI正是在这一判断下，成功探索出的第一款AI原生应用。其首先提出并落地的AI短剧+AI互动融合的产品体验，完成了商业验证。

剧本输入

角色设定

分镜生成

成片包装

短剧发行

- 支持生成+理解一体的多模态大模型
- 创作者意图遵循准确率提高100%

- 满足业务场景要求内容一致性和可控性
- 为100分钟以上的优质长故事型视频而设计

落地效果：

创作

表演及剧情是优质长故事型视频核心

引发情感共鸣是消费级内容创作关键



消费

用户可观看高质量消费级AI短剧

可与剧中角色互动
丰富沉浸互动体验

客户反馈：

奖项



上海电影、filMarathon
「全球AI电影马拉松大赛」

井英科技与Jud Willmont
导演合作AI短片《M.A.D》
获B站观众选择奖

榜单

- 2024年7月，登上「点点数据」iOS海外娱乐畅销榜
- 2024年11月，根据《短剧内行人》报道iOS榜单跻身第7位

真实评论

Absolutely
obsessed

几乎察觉不到ai的气息，
这个路线感觉是可以最先商业化的

can't wait to see what's
the plan for tonight...

画面感挺自然的，小
细节还算不上华丽。
另外，故事不错。



量子位



量子位智库

量子位智库小助手
(社群入口)

结语

文行至此，2024年度AI十大趋势已经全面展现。

《2024年度AI十大趋势报告》不仅是本年度我们对AI整体、系统的观察，也是量子位智库的「年度报告」，是我们对量子位智库一整年以来工作的总结。

第一，与量子位一起，我们依然发挥着前沿科技瞭望所的作用。我们所讨论的技术，不论在成熟度、商业落地上的进展如何，都是最新、最重要、谈及AI绝对避不开也绕不开的。除了向前展望，我们还从GPT-3时代起就坚持做科普向的短报告，力求最短时间内减少信息差。

第二，自2024年5月起，我们长期跟踪AI应用在Web和APP两个端口的各项表现，并以月为单位发布分场景数据报告。不同的是，除了持续追踪最新的数据外，我们还会基于对技术和赛道的研究附上分析以供参考。

第三，在行业篇中我们也谈到，本年度量子位智库发布了十余篇各AI落地行业及环节的研究报告。在行业上涉及智能驾驶、具身智能、教育、影视、游戏等各个赛道；在具体落地环节上综合对市场的特别观察，针对市场的投融资情况发布了创投类报告、针对市场的集中程度发布AI大模型创业格局类报告。

2025年，我们将持续关注以上议题的最新进展并且进一步挖掘更为深度的内容。我们还规划了对**2024全年AI应用产品的整体分析**、**对大模型落地和前沿趋势的探讨**、以及更多**面向其他被AI改变的行业的研究**……

总之，在AI渗透的各个环节，我们都在关键的节点和位置做了相关研究。随着时间的推移和研究的深入，最终我们希望一路陪伴的读者能够以最清晰、有逻辑的方式构建对AI及相关产业的体系化认知。同时，我们也正通过沙龙、对话、社群等方式将学术界、产业界、投资界以及爱好者聚合起来。

项目团队

李根 项目指导
郑钰瑶 项目统筹

研究团队

郑钰瑶 分析师
yuyao@qbitai.com
刘萌媛 分析师
mengyuan@qbitai.com
王昕祎 分析师
xinyi@qbitai.com

美术团队

孙旖涵
胡新雨

量子位编辑部
智能车参考编辑部
对本报告亦有贡献，特此感谢。

2024 QBITAI ANNUAL AI TOP 10 REPORT

年度AI十大趋势报告