

生成式人工智能 智能 2025 专利布局



目录

关于Questel.....

.....3 执行摘要.....

.....4 I. 介绍.....

.....6 II. 方法论..... 8

1. 数据来源 & 搜索策略..... 8

2. 分类法..... 9

III. 深度学习全球专利格局.....10 IV. 聚焦多模态

AI / 数字人类 / 智能代理....14

1. 简介.....

.....14

2. 多模态人工智能.....

.....15

3. 智能体.....

.....17

4. 数字人.....

.....18

5. 生成式人工智能.....

.....20

附录.....

.....22

关于Questel

Questel是面向20000多家客户和150万用户的真实端到端知识产权解决方案提供商，覆盖30个国家。我们提供一套全面的软件，用于搜索、分析和管理发明和知识产权资产。

Questel 还提供贯穿整个知识产权生命周期的服务，包括先例检索、专利申请文件撰写、国际申请、翻译和续展。这些解决方案，与我们的知识产权成本管理平台相结合，为客户在整个申请预算中节省平均 30-60%。

Questel的使命是让创新以高效、安全、可持续的方式得到发展。在该使命背后，Questel认为企业社会责任（CSR）是一场鼓励企业对其活动对客户、员工、社区和环境产生的影响负责的广泛运动。欲了解更多关于Questel的信息，请访问：www.questel.com

更多内容可在网上获取，请咨询：www.questel.com/resources-hub/

执行摘要

深度学习持续成为技术创新中最具活力的领域之一，专利活动显示出持续加速的增长。基于我们对深度学习和大型语言模型的前期研究，本报告通过专利申请分析了下一波创新浪潮，重点关注多模态人工智能、智能代理和数字人。这三个领域正在快速兴起，并日益融合，朝着更加自主、交互和以人为本的人工智能系统发展。

专利格局显示，少数全球科技领导者通过将基础模型、自主能力和类人界面结合为连贯的创新战略，正在塑造这一转型。

谷歌通过其Gemini系列模型将自己定位为核心技术领导者。作为原生多模态设计，Gemini集成了文本、视觉、音频和视频，同时逐步嵌入代理推理能力。谷歌的专利策略反映了其对国际保护的强烈重视，支持其部署多模态和基于代理的AI在全球范围内在搜索、云服务和生产力工具上的雄心。

百度作为垂直整合度最高的参与者脱颖而出。凭借 ERNIE 多模态引擎、GenFlow 和 AgentBuilder 用于智能体，以及快速扩张的数字人技术组合，百度在一个统一的堆栈中覆盖了所有三个领域。其专利领导地位，特别是在数量上，突显了一种以大规模部署为中心的策略。

英伟达主导了数字人领域。虽然它不将自己定位为通用人工智能助手提供商，但英伟达提供了实现具身人工智能所必需的基础设施、平台和工具链。其Omniverse和Avatar Cloud Engine技术支持高度逼真的数字人和多模态交互，并拥有非常强大的国际专利组合。

微软采用以企业为中心的方法，将多模态人工智能和智能代理直接连接到业务工作流程。利用其与OpenAI 的合作以及自己的代理框架，微软将 Copilot 跨 Office 集成

云和企业软件。其专利申请显示了一种均衡的战略，结合了稳固的产品组合规模和高比例的国际家族，反映了在专业和企业人工智能领域的全球雄心。

. 国际商业机器公司 (IBM) 成为智能代理领域的核心领导者，其watsonx.ai平台和Granite基础模型家族提供支持。IBM的战略专注于专业和企业级AI系统。虽然在其面向消费者的应用中不那么显眼，但IBM强大的专利地位证实了其作为企业环境中基于代理的AI基石角色。

一、引言

深度学习是人工智能的核心技术层，基于能够从大型数据集中学习层次化表示并擅长图像识别、语音识别和自然语言处理等感知任务的多层人工神经网络。在这一技术层中，生成式人工智能（GenAI）于2016年左右兴起，作为一个专注于使用VAEs、GANs、扩散模型、Transformer等技术创建新内容文本、图像、视频、音频、代码或合成数据的分支，尤其是大型语言模型（LLMs）。LLMs是基于Transformer的系统，在大量语料库上进行训练以理解、生成和操作自然语言，具备推理、摘要、翻译和对话交互等能力。

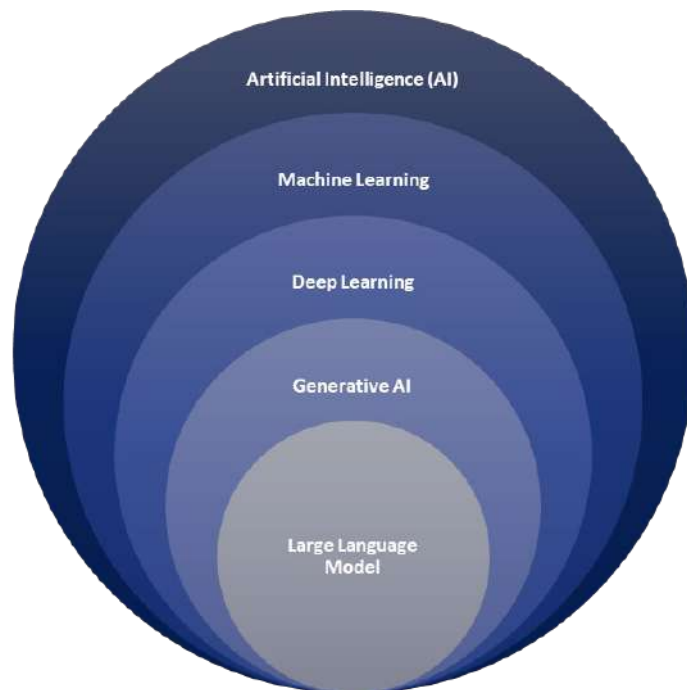


图1- AI的谱系：揭示智能系统的不同层次

与这些基础人工智能层次不同，几个领域已发展成为构建于其上的应用：

- 多模态AI指的是能够在多种模态或数据类型（文本、图像、音频、视频、传感器数据等）之间处理、理解和生成信息的系统。通过整合并推理不同形式的输入，这些系统产生统一的理解，并能生成连贯

跨越模态的输出（例如，带说明文字的图像、有声视频或同步的音频视频响应）。这种多模态能力比单模态人工智能更能实现更丰富的感知、更自然的人机交互以及更灵活的内容生成。

- 智能体是自主或半自主的AI系统，通过传感器（这些传感器可以包括视觉、音频或其他传感器数据）感知其环境，根据给定的目标做出决策，并采取行动以实现特定目标。基于深度学习、多模态理解以及生成或推理能力，这些智能体可以独立运行、从经验中学习、与人类和其他智能体交互，并适应不断变化的环境。这使得它们适合于任务自动化、规划、人机协作、响应式辅助或自适应决策支持等动态任务。

- 最后，数字人是AI驱动的虚拟人类形象，结合计算机图形学、动画、自然语言处理、语音合成和行为AI，创造出能够与真人进行自主或半自主交互的逼真人类化身。得益于多模态AI和生成式技术，数字人可以表现出逼真的外观、声音、表情、手势和行为，从而实现虚拟助手、虚拟培训师、数字演员或医疗保健、教育、娱乐或客户服务中的伴侣系统等应用。

II. 方法论

1. 数据来源 & 搜索策略

本研究中使用的数据源是全球专利数据库搜索工具Orbit Intelligence FamPat。FamPat是一个按简单专利家族组织的全球专利申请和授权专利集合，涵盖全球100多个专利局，包括由Questel提供的22家专利局的可搜索全文。由于每个FamPat记录可能包含许多具有不同日期的单独出版事件，报告使用每个专利家族最早已知的首次申请日期的办公室。这被认为是代表性的专利家族成员，用于指代该专利家族。首次申请办公室（OFF）或优先权是指针对特定发明的首次申请，在任一专利局提交时，该申请即成为“优先申请”，该事件日期定义为优先权日期。首次申请的国家被定义为优先权国家。

报告中所包含的表格和图表使用此优先日期，除非另有说明，因为它提供了最能反映发明活动的最准确指示。专利来源的定义，即专利家族的来源地，是基于首次申请办公室（OFF）。需要注意的是，此定义并不完全准确；尽管如此，它提供了一种有用且公平的方法来确定实体的通常首次申请国，这通常与其 home 专利局一致。

每个专利家族都与一个或多个实体相关联，这些实体统称为专利权人。在相同家族内的专利由多个权利人持有时，该家族归属于所有相关实体。为了提高可读性并提供清晰概述，本文中提到的各种申请人已被系统地清理和整合到其各自母公司下。子公司和关联实体被归类到其母组织下，以反映统一的所有权结构。此分组过程结合使用自动化方法（通过Orbit Intelligence数据库）和手动方法进行。此分组是基于报告创建时可获得的公开信息最佳情况进行的。

2. 分类学

这项专利布局聚焦于深度学习（DL），特别关注三个快速兴起的应用领域：多模态人工智能、数字人类和智能代理。其检索方法结合了多种策略，利用深度学习相关关键词、相关技术和相关专利分类（如 G 06N 3/02（基于神经网络模型的计算机系统）），以确保检索结果在全面性和精确性之间达到最佳平衡，采用布尔逻辑和迭代优化，从而构建了一个用于分析的稳健可靠的数据库。

尽管专利并非总是明确描述这些应用，但审查那些明确描述的应用可以深入了解这些技术是如何被实施和保护的。这种方法提供了一种独特且基于技术的视角，可以补充企业沟通或营销叙述，而企业沟通或营销叙述通常只强调最明显或最热门的使用案例。

为丰富分析，我们回顾了特定领域的文献，以识别和分类多模态人工智能、数字人和智能代理的最相关应用。这使我们能够突出它们在各个行业和用例中的潜力，为这些快速发展的技术领域提供可行的见解。这些应用的详细列表及其描述在本报告末尾的附录中提供。

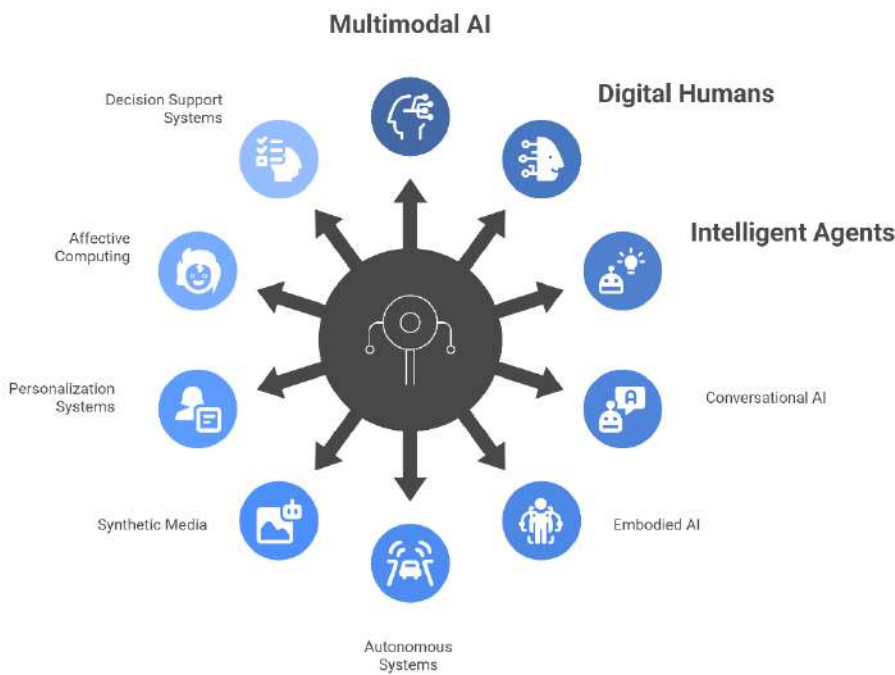


图2- 人工智能应用，重点关注多模态人工智能、数字人、智能代理

III.深度学习全球专利布局

爆炸式增长：深度学习热潮持续！

数据显示深度学习和生成式人工智能革命经历了三个主要阶段：

- 2011–2016：深度学习革命：表征学习、卷积和循环神经网络的突破，以及GPU加速训练，将深度学习确立为人工智能的主导范式。
- 2016–2021：生成式AI的兴起：Transformer架构的引入使得大型生成式模型得以实现，并标志着从任务特定AI向基础模型的转变。
- 2021–至今：生成式人工智能爆炸：规模定律、架构创新和大规模计算资源的融合驱动大型语言模型以史无前例的规模发展（例如，ChatGPT），基于扩散的图像生成（例如，Stable Diffusion），以及结合文本、视觉和音频的多模态系统（例如，GPT-4V，Gemini）。

下表显示了每年公布的专利家族数量。由于专利申请通常在提交后18个月公布，因此公布日期应向后移动18个月，以更准确地反映实际的发明提交日期。

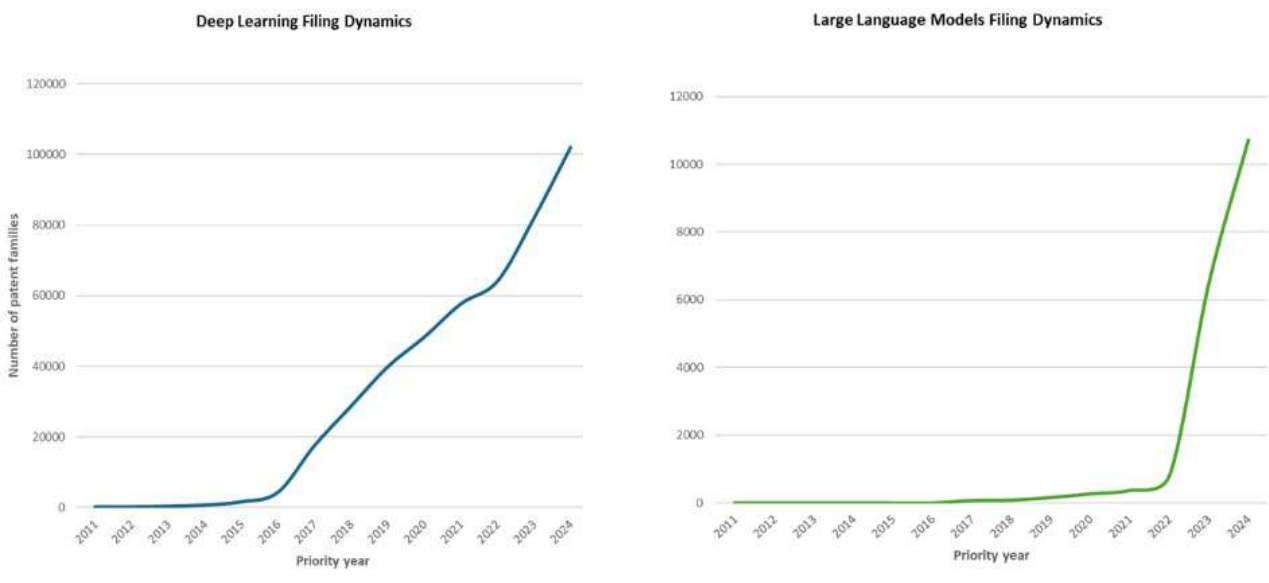


图3-深度学习，和LLM动态申请，2011-2024（专利家族数量与优先申请日期）

我们对2011–2025年期间的分析提供了对深度学习（DL）相关专利活动的显著演变的近乎完整的视角。深度学习，在其现代形式于2011年左右被定义以来，在此时间段内产生了大约450,000个专利家族。正如我们在之前的研究中所指出的¹，深度学习领域的专利申请持续加速，且无明显稳定迹象，早期年增长率接近50%。这一持续轨迹，从2011年近乎零的申请到2024年超过10万件已公布专利家族，表明该领域的创新仍保持高度动态性，大规模研究和工业项目仍在积极扩张。

最大的一项发展是大语言模型（LLM）的迅速崛起。仅在2020年后出现的LLM从几乎零开始经历了快速增长，特别是从2022年开始，展示了前沿规模模型出现后随之而来的研究、工程 and 商业化爆炸。它们的扩张与深度学习（DL）的根本进步密切相关：更大的架构、改进的训练管道以及自监督学习的进步为LLM的崛起创造了理想条件。

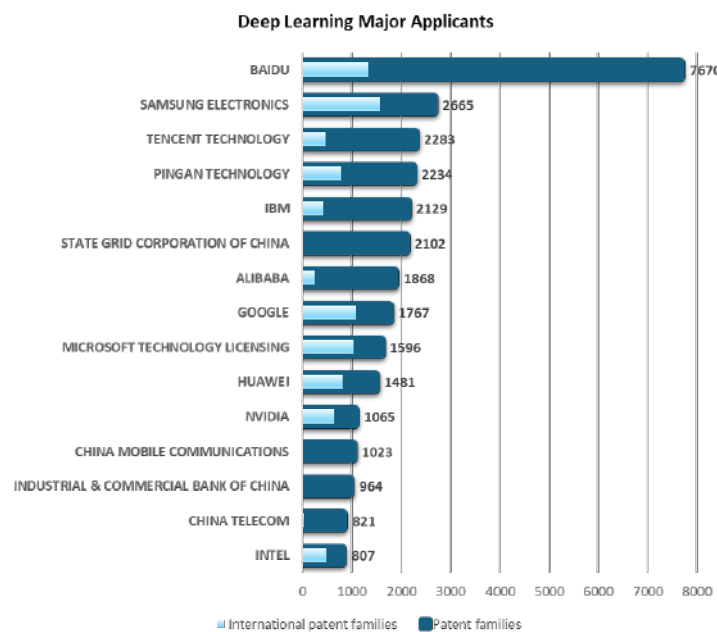


图4 - 2011-2024年深度学习主要工业申请者

¹ <https://www.questel.com/resourcehub/consolidation-and-opportunities-in-llm-domain/>

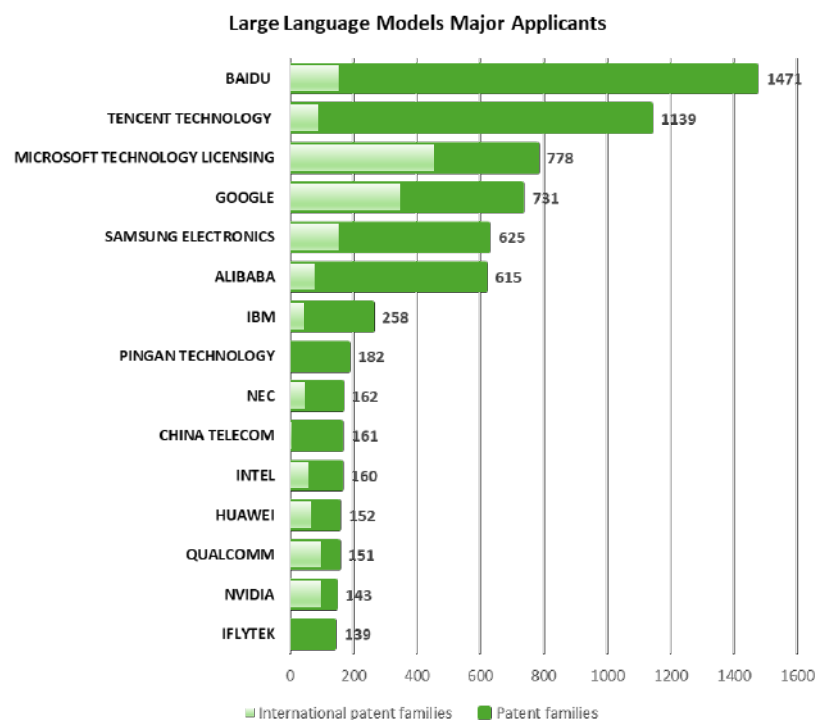


图5 - 2011-2024年LLM主要工业申请人

图4和图5展示了在深度学习和大型语言模型领域塑造专利格局的主要被许可人，按专利族总数排名。它们提供了对创新规模的洞察，以深色表示，以及主要参与者之间国际专利保护的程度，以浅色表示。

深度学习专利：领导地位正在集中

就深度学习技术领域而言，主要玩家持续巩固其地位，其专利组合也取得了令人印象深刻的增长。

百度以相当大的领先优势位居第一，拥有7,670个专利家族。这一主导地位进一步得到大量国际专利家族（1,347个）的支持。百度在深度学习领域继续保持领先地位，自首次研究以来，连续第三年在专利申请数量上排名第一²，从而巩固了其在该领域中最活跃的全球创新者的地位。这种持续的主导地位反映了一种基于核心深度学习技术的长期大规模知识产权战略。

² <https://www.questel.com/wp-content/uploads/2021/11/2021-Deep-Learning-Patent-Landscape-short-report-.pdf>

在第二位，三星电子拥有2,665个专利家族，国际申请数量较多（1,578个），使其成为国际保护方面的领导者。腾讯科技以2,283个专利家族位列第三，拥有482个国际保护。紧随其后的是平安科技，拥有2,234个专利家族，其中791个为国际专利家族。位列第五的是IBM公司，拥有2,129个专利家族，国际申请为435个。国家电网公司拥有2,102个专利家族，但只有2个国际专利家族。阿里巴巴位列第七，拥有1,868个专利家族，其中251个受到国际保护。最后，谷歌公司以1,767个专利家族位列第八，并拥有显著高的国际保护水平（1,097个）。

百度主导LLM专利，腾讯、微软和谷歌构成强劲第二梯队

百度在LLM专利格局中引领，拥有1471个专利家族，明显领先于所有其他专利权人，证实其在大型语言模型相关创新中的主导地位。百度进一步巩固了其地位，决定性地领先，从去年的第二位跃升³在专利申请量方面位居全球前列。这一进步突显了百度在LLM技术上的加速投资，并强调了其在下一代基础模型中的日益增长的影响力。

腾讯科技以1139个专利家族位居第二，而微软技术授权和谷歌占据接下来的位置，它们结合了庞大的专利组合规模和相对较强的国际专利覆盖范围。第二梯队由三星电子和阿里巴巴组成，每个公司拥有超过600个专利家族，但国际保护水平有所不同。IBM、平安科技、NEC、英特尔、华为、高通和英伟达等中档玩家拥有更适中的专利组合规模，通常伴随着较低的国际申请量。

³ <https://www.questel.com/resourcehub/consolidation-and-opportunities-in-llm-domain/>

IV. 聚焦多模态人工智能/数字人类/智能体

1. 简介

人工智能目前正进入由三个密切相关的新兴创新领域驱动的新阶段：多模态人工智能、数字人和智能体。我们的专利分析识别出约15000个多模态人工智能的专利家族，8000个数字人的专利家族，以及6000个智能体的专利家族。这些技术显然正在一起涌现并相互强化，创造出新一代人工智能系统，这些系统能够以几年前不可能的方式感知、沟通和行动。

多模态AI指的是能够同时理解和生成多种类型数据的AI系统，例如文本、图像、音频和视频。智能体是设计用来通过推理、规划、使用工具和随时间采取行动来自主追求目标的AI系统。数字人是能够说话、展示面部表情并以类似人类的方式与用户互动的AI驱动的虚拟角色。

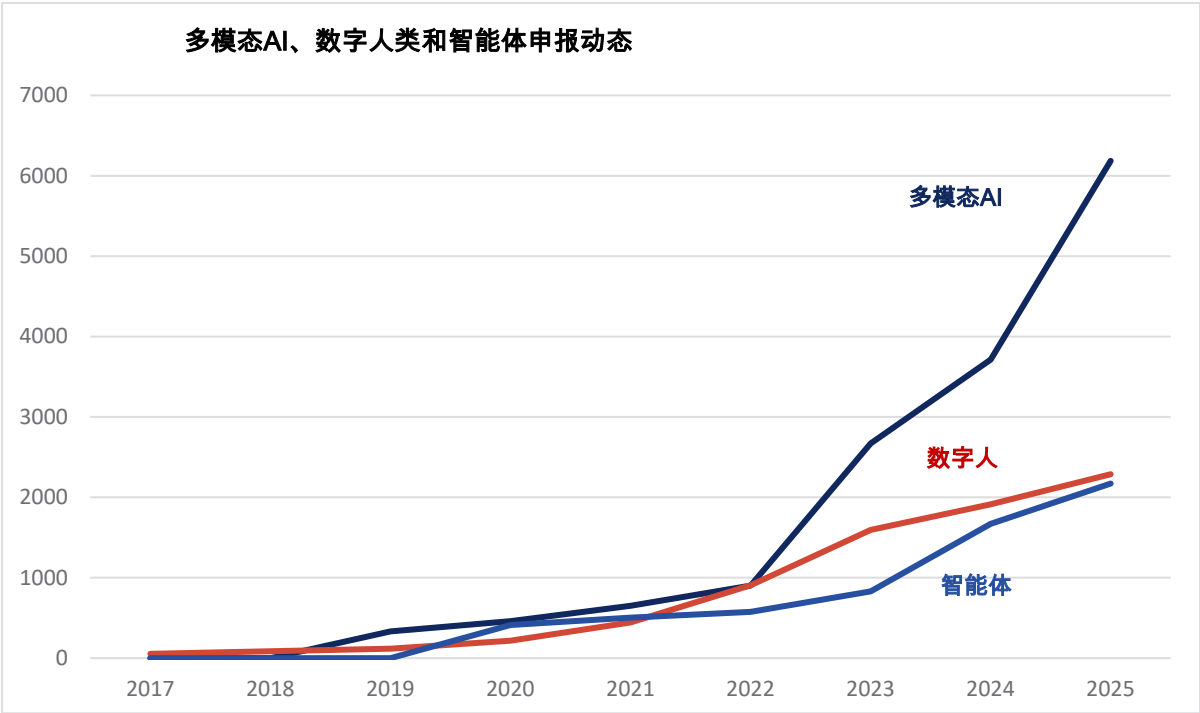


图 6 - 多模态、数字人类和智能代理的申报动态，2017-2024

2. 多模态人工智能

多模态人工智能：模型领导力与专利实力定义下一场人工智能竞赛

2025年标志着多模态人工智能的转折点，该领域的领军企业推出了多项新的多模态引擎。这些推出清楚地表明，技术竞赛正在加速，并在多模态人工智能中建立强大的专利组合将成为未来的关键战略资产。虽然一些公司主导了公众可见度和市场采用，但另一些公司则通过密集的专利申请策略来定位自身。

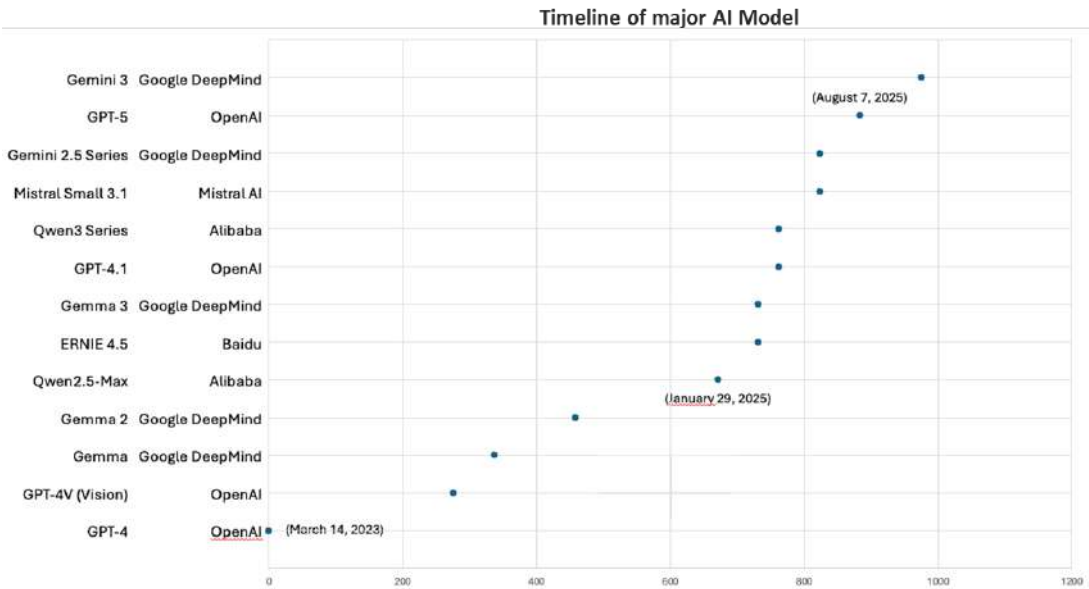


图7-主要AI模型时间线，2023-2025

多模态人工智能是当前人工智能创新的最重要基础之一。它的目标是使机器能够理解和结合多种类型的信息，例如文本、图像、音频和视频，在一个模型中，以更接近人类感知的方式。多模态模型不是单独处理每种模态，而是学习连接不同输入的共享表示，从而实现更高级的推理和交互。

. openai 是这一类别中最引人注目的创新者之一。凭借 gpt-4、gpt-4v 和 gpt-4o 等模型，openai 专注于跨文本、图像和音频的统一多模态推理，并高度重视实时交互和自主性

能力。在公共领域，GPT几乎与人工智能同义。然而，尽管在认知和采用方面具有领导地位，但OPENAI似乎并未出现在顶尖专利所有者之列。

谷歌是另一个主要的核引擎开发者。Gemini系列模型从一开始就被设计为原生多模态，在单一架构中集成文本、图像、音频和视频。谷歌的战略强调深度推理、长上下文理解和与其生态系统的紧密集成，包括搜索、安卓和云服务。这一技术领导力也反映在强大的、国际化的专利组合中。

百度及其ERNIE 4.x模型发挥着类似的作用。ERNIE是一个多模态基础模型，能够理解文本、图像、音频和视频。这一策略得到了专利申请的有力支持，将百度定位为技术和知识产权的领导者。

其他重要的引擎开发者还包括 ANTHROPIC (Claude 3)、iFLYTEK (Spark LLM)、ALIBABA (Qwen，具有强大的企业专注点) 以及 IBM (Watsonx.ai 和 Granite，同样面向企业使用场景)。META (特别是通过 LLaVA 风格模型) 和 MISTRAL 等开源贡献者也发挥作用，尽管与最大的基础模型提供方相比，通常规模更为有限。

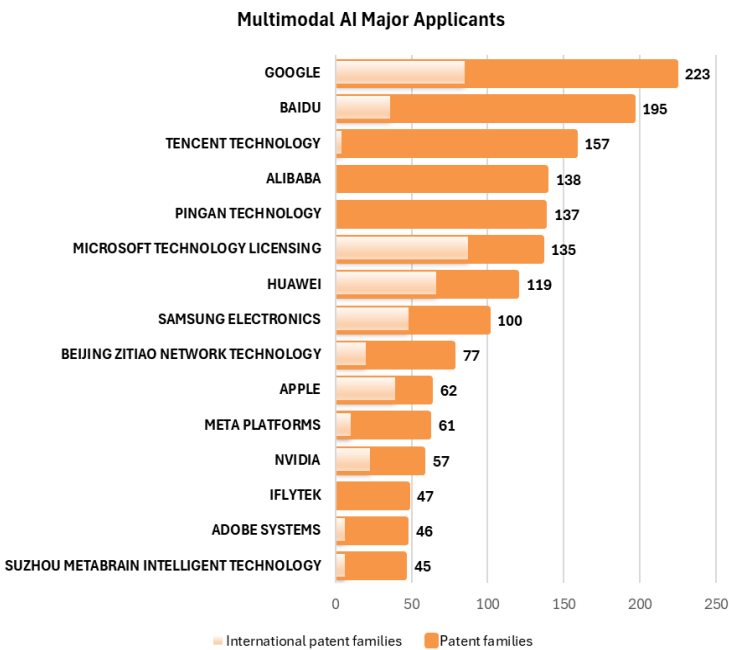


图 8 - 多模态 AI 领域的主要工业参与者，2017-2024

GOOGLE引领了专利布局，拥有最多的多模态AI专利家族，并具有强大的国际保护水平。BAIDU紧随其后，其庞大的专利组合反映了对多模态技术的长期承诺。

腾讯、阿里巴巴和平安科技也拥有大量专利家族，尽管它们的申请更多地集中在国内保护。

微软、华为和三星电子因其平衡的策略而脱颖而出，结合了显著的专利数量与较高的国际申请比例。苹果、英伟达和Meta等传统科技领导者虽然专利数量较为适中，但仍保持了有选择性的国际覆盖。

3. 智能体

从对话到行动：智能体正成为人工智能的操作层

智能体是设计用于在长时间内自主运行的人工智能系统。它们不仅仅是对提示做出回应，还能追求目标、将任务分解为步骤、使用外部工具，并根据反馈和记忆调整自身行为。随着大型语言模型的兴起，该领域发展迅速。

图9突出了智能代理领域的关键受让人，显示了专利家族的总数及其国际覆盖范围。

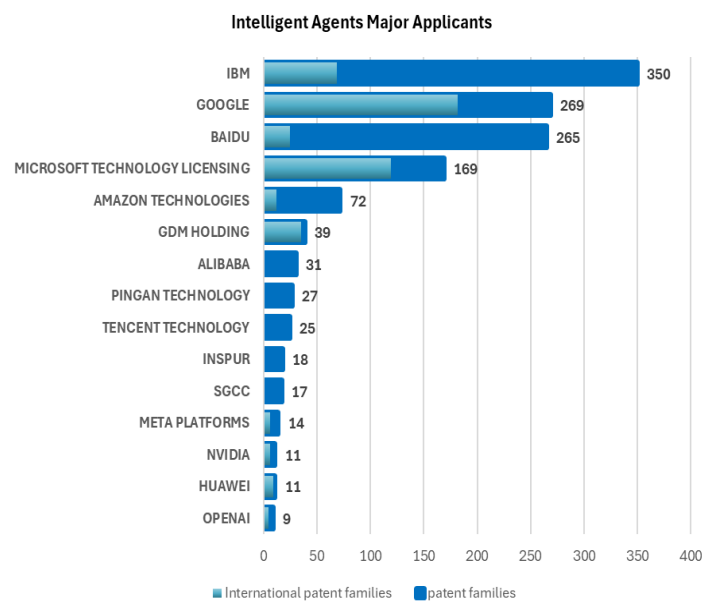


图9 - 智能代理行业主要参与者，2017-2024

IBM（凭借watsonx.ai平台和Granite基础模型家族）在这一领域的专利总体量方面处于领先地位。战略上，IBM将智能代理定位为专注于面向专业人士和企业层面的解决方案。

谷歌（使用Gemini代理）、百度（使用GenFlow/AgentBuilder）和微软（使用Copilot和代理框架）在专利申请方面紧随其后，拥有庞大的专利组合，并将Gemini定位为一种能够推理、规划和与工具交互的代理模型。微软在这些基础上构建，将Copilot产品集成到企业软件中，实现文档、电子邮件和工作流程中的任务自动化。

OpenAI（9专利家族，包括5IPFs）、NVIDIA（11专利家族，包括6IPFs），以及华为（也是11家族，包括9IPFs）总的来说专利数量较少，但相对于其专利组合的规模，仍保有较好的国际家族份额。

另一方面，亚马逊、阿里巴巴、平安科技、腾讯科技、SGCC和英业达专注于国内申请，几乎没有或没有国际保护。

智能代理专利格局似乎已经高度集中，少数全球科技巨头（IBM、GOOGLE、BAIDU & MICROSOFT）控制着数量和国际存在感，而许多专注于本土的公司则保持较小的专利组合，并且全球扩张有限。

4. 数字人

数字人：专利领导力集中在全球科技巨头手中

数字人代表着以人类形态体现的人工智能。它们结合了语言模型、语音技术和实时动画，创造出能够与用户自然交互的虚拟角色。这个领域最近取得了进展，因为生成式人工智能现在能够大规模实现逼真的对话、富有表现力的声音和一致的性格。

图10说明了数字人在主要权利人中的情况，显示了专利家族的总数及其国际覆盖范围。

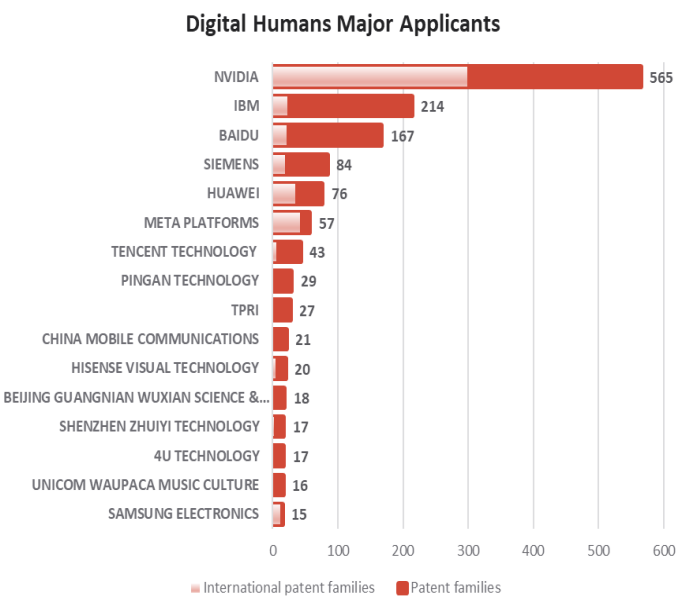


图 10 - 数字人类主要工业参与者，2017-2024

英伟达在数字人专利领域占据明显主导地位，拥有565个专利家族，其中包括299个国际家族。这反映了极高的创新水平以及强大的全球保护策略。英伟达的领导地位与其产品生态系统一致，包括Omniverse、Avatar Cloud Engine（ACE）和结合语音、动画及AI推理的实时头像技术，用于企业、元宇宙和模拟应用场景。

IBM（214个专利家族）和Baidu（167个专利家族）在总专利数量上紧随其后，尽管两者在国际保护方面都显示出较低的水平，分别为22个和21个国际专利家族。IBM的活动与集成到Watsonx.ai的企业导向型数字人类用例相一致，例如专业和商业环境中的虚拟助手和类人界面。

百度则将数字人与直播、教育、电子商务和企业服务紧密联系，利用其ERNIE多模态模型大规模赋能交互式虚拟形象。

一个由中型企业组成的群体采取更具有选择性的国际化战略。西门子持有84个专利家族，其中包含19个国际家族，反映了其对工业和专业数字人应用的专注。华为拥有76个专利家族和35个国际家族，随后是Meta拥有57个专利家族其中大部分是国际的（42个国际专利家族），专注于嵌入在消费设备、自动售货机、显示屏和智能环境中的数字人，通常与其更广泛的设备生态系统相关联。

在另一端，一些企业仍主要关注国内市场，国际扩张有限或没有。这包括平安科技（29个系列）、中国移动通信（21个系列）、TPRI（27个系列）以及中国联通瓦帕卡音乐文化、4U科技和北京广年五县科技等公司。这些参与者通常瞄准本地市场和特定应用，如客户互动、媒体或公共服务。

总体来看，格局指向一个新兴的层级：少数全球平台领导者塑造核心技术NVIDIA，随后是IBM和Baidu，然后是西门子、华为和Meta等应用驱动的创新者，以及区域聚焦的实施者。

5. 生成式人工智能

从模型到应用，生成式人工智能重塑智能系统

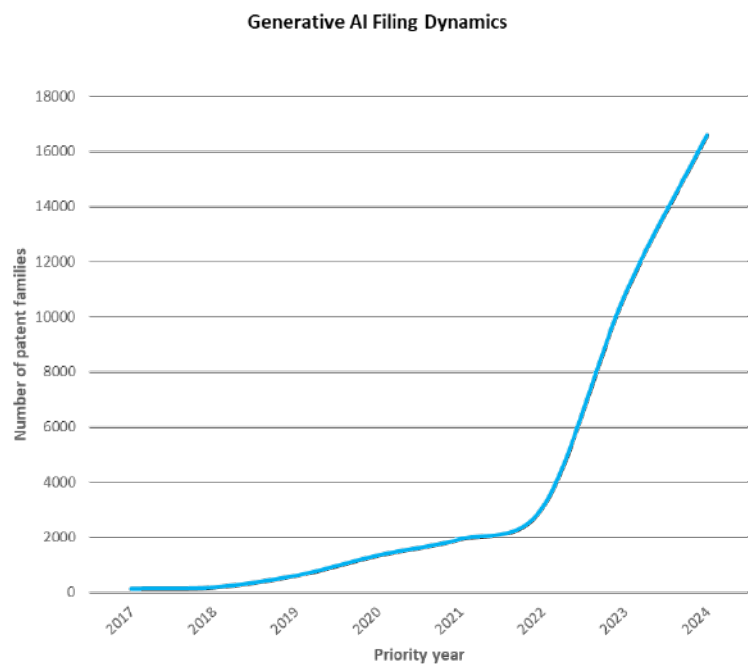


图11 - 生成式人工智能申请动态 2017-2024 (专利家族数量与优先申请日)

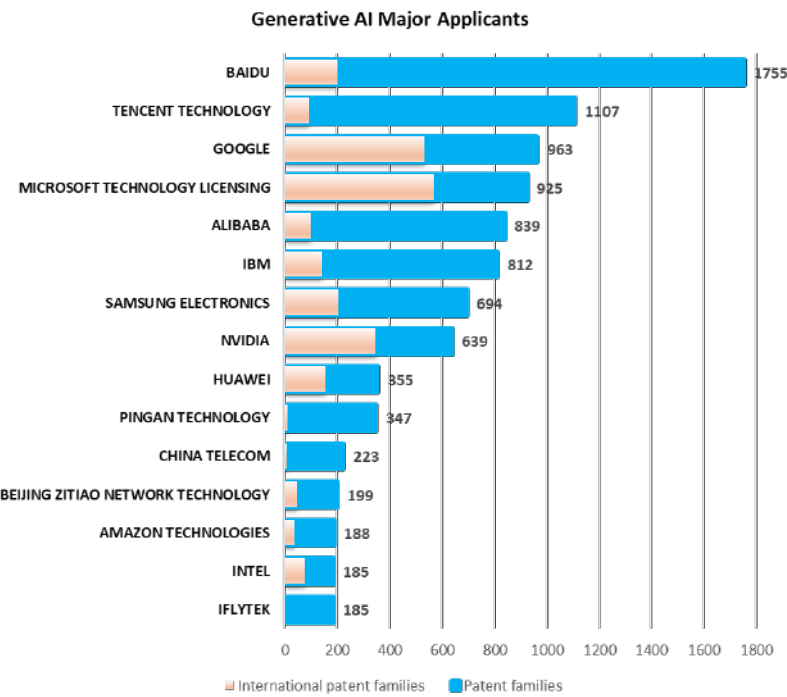


图12 - 生成式AI主要工业申请人，2017-2024

自2017年以来，随着大规模基础模型的兴起，生成式人工智能应用技术、多模态人工智能、数字人和智能代理进入了一个快速且持续扩张的阶段。专利申请动态清晰地说明了这种加速，多模态人工智能在2021年后的激增最为显著，而数字人和智能代理则展现出稳定且日益强劲的增长轨迹。这反映了生成式人工智能从探索性研究过渡到跨多模态和应用层可扩展、可商业部署的系统。

一项对主要被许可人的分析表明，像谷歌和微软这样的全球科技巨头已经建立了庞大且高度国际化的专利组合，在生成式AI领域，它们的国际专利家族数量始终排名靠前。它们的战略强调广泛的全球保护和以平台为中心的创新，巩固了它们在不断发展中的生成式AI生态系统中的长期领导地位。

同时，百度在所有主流生成式AI应用技术领域都展现出了平衡且全面的发展。该公司在多模态AI、数字人和智能体方面拥有强大的专利组合。这种平衡的定位，结合庞大的专利数量，突显了百度在生成式AI应用中的独特作用，尤其是在基于代理和面向任务的AI系统方面。

附录

附录A：术语表

专利家族 (Fampat)：定义为“基于发明的家族”，该家族定义来自Questel，它将EPO的严格家族规则（相同优先权申请）与附加规则（申请超出12个月申请期限；EP与PCT出版物之间的联系...等）相结合。严格家族和扩展家族之间的一个合理妥协。

国际专利家族 (IPFs)：为消除各国政府因申请激励而产生的统计偏差，并突出具有潜在普适性效益的创新，通常做法是强调专利家族在其原始申请权之外扩展的范围（即家族中至少包含一项在优先权国家以外的专利）。通过将专利扩展到多个国家，公司展示了对这些技术的坚定承诺，强调了它们对发明在全球潜力和市场相关性的信念。这种国际范围有助于筛选出影响力较弱或本地化的创新，只留下那些具有大量投资和感知长久寿命的专利。

专利申请：要获得专利，必须在相应的专利局提交所有必要文件和费用。专利局将进行审查以决定是否授予或驳回该申请。专利申请通常在申请最早的优先权日期后18个月公布。在那之前公布，该申请仍然保密。

专利分类：专利分类是一种系统，供专利局审查员或其他人员根据其内容的技术特征对文件（例如已公布的专利申请）进行编码。国际专利分类 (IPC) 是国际公认的。欧洲专利局 (EPO) 和USPTO启动了一项联合项目，以创建协同专利分类 (CPC)，以协调这两个机构之间的专利分类系统。

专利申请人/权利人：当一个人申请专利时，“申请人”一词指代权利人，即发明人负有义务转让该发明的个人，或以其他方式对该事项显示出充分所有权的个人。我们称之为专利所有人或专利转让方；通常，它是发明人的雇主。

优先提交：一项特定发明首次提交专利申请的地点，也称为首次提交办公室。

附录B：应用描述

主要LLM应用描述：

内容生成：大型语言模型可以创建文章、产品描述、营销材料和其他类型的内容。翻译和本地化：大型语言模型提供强大且精确的翻译功能，包括对口语或书面内容进行实时翻译。聊天机器人和虚拟助手：大型语言模型可以理解复杂问题，并根据庞大的知识库提供相关答案。搜索引擎增强：大型语言模型通过处理自然语言查询并提供建议、人性化的答案来改进搜索引擎。代码生成：大型语言模型可以理解并生成基于自然语言描述的代码片段。医疗保健和医学研究：大型语言模型协助进行医疗记录分析、诊断预测、药物发现和个性化治疗建议。金融分析：大型语言模型可以总结盈利电话会议、创建文字记录，并协助进行欺诈检测和分析。法律应用：大型语言模型有助于法律释义、文字记录和在大规模文本数据集中搜索。教育和培训：大型语言模型通过根据个人学习者的需求调整内容、提供辅导和生成练习题来提供个性化教育。网络安全：大型语言模型可用于增强网络安全措施。情感分析：大型语言模型可以分析客户反馈、评论和社交媒体提及，以了解公众认知和趋势。

应用	定义	外观年份
多模态人工智能	能够跨多个理解并生成人工智能系统模态（文本、图像、音频、视频、传感器数据）。	2019
数字人类	人工智能驱动的高仿真虚拟人/数字人语音、动画和交互来模拟类人存在。	2017
智能代理	自主或半自主软件实体感知环境、推理/做决定、并行动为了实现目标。	2020

附录C：多模态引擎发布日期

模型	公司	发布日期	注意
Gemini 3	谷歌	11月18日	最新旗舰多模态模型。
	DeepMind	2025	维基百科
GPT-5	OpenAI	2025年8月7日	下一代多模态LLM。 维基百科
Gemini 2.5 系列	谷歌	2025年6月17日	广泛部署多模态变体。
	DeepMind		谷歌云
mistral small 3.1	Mistral AI	2025年3月17日	https://en.wikipedia.org/wiki/Mistral_AI
Qwen3 系列	阿里巴巴	2025年4月28日	广泛的 multimodal Qwen3 模型。
			维基百科
GPT-4.1	OpenAI	2025年4月14日	更新后的多模态旗舰。维基百科
Gemma 3	谷歌	2025年3月12日	最新开放家族版本。维基百科
	DeepMind		
ERNIE 4.5	百度	2025年3月中旬	升级的多模态模型。路透社
Qwen2.5-Max	阿里巴巴	2025年1月29日	早期的多模态Qwen变体。路透社
Gemma 2	谷歌	2024年6月27日	更新家族版本。维基百科
	DeepMind		
Gemma	谷歌	2024年2月21日	开源模型家族
	DeepMind		多模态变体。维基百科
GPT-4V (视觉)	OpenAI	2023年末 (愿景展开)	增加视觉输入；广泛可用
			2023年末。Roboflow博客
GPT-4	OpenAI	2023年3月14日	核心多模态模型 with 文本 +
			图片支持。维基百科

更多信息

关于Questel的咨询团队

questel的咨询团队结合了知识产权、技术和市场的专业技能，提供各种数据驱动型项目，使创新组织能够做出战略决策：专利格局分析、技术和能力图谱、专利组合评估（用于剪枝或许可）、技术侦察等。



关于作者

夏尔·迪厄梅加德
高级顾问

侯赛姆·阿尔巴维
咨询顾问 & 知识产权分析师



联系

consulting@questel.com

我们的见解可以帮助您领先市场趋势和挑战。