### **2025.6.30(一)-行业新闻日报**

**概览：**

1. 全球程序化媒体合作伙伴 MiQ 推出业内首个 AI 广告整合平台 MiQ Sigma。该平台基于全球最大规模数据集构建，整合 300 多种数据源及 700 万亿条消费者行为信号，借助 AI 技术提升营销效果，具备强大的可视化能力、AI 交易代理功能及生成式 AI 用户画像生成能力，助力精准触达受众与优化营销决策。
2. 阿里通义千问发布多模态模型 Qwen VLo，支持动态分辨率图像生成和逐步清晰机制，能精准理解与再创造内容，未来将支持多张图片操作整合及物体识别切分，目前处于预览阶段.
3. 谷歌推出开源多模态模型 Gemma 3n，仅需 2GB/3GB 内存即可在端侧设备运行，支持图像、音视频和文本输入，响应速度快且多语言表现优异，性能在同规模模型中领先。
4. 飞利浦推出 8 号 Pro AI 耳机，采用腾讯混元 AI 语言大模型，支持 17 种语言互译和 27 种方言识别，内置专业术语库，提升国际交流效率，适用于商务与旅行场景。
5. Runway AI 宣布推出 “游戏世界” 大型互动文本冒险平台，从影视领域拓展至游戏行业，目前支持文本与图像界面，未来计划推出视频形态游戏生成功能。
6. 上海交大等团队提出时序 - 语言桥接架构 ITFormer，应用于航空发动机运维，定义 “时序问答” 任务范式，构建相关数据集，其性能超越 ChatGPT-4o，可适配多种时序编码器和大语言模型，为工业场景推理提供标准化基准。

**具体内容：**

1. **MiQ推出业内首个AI广告整合平台MiQ Sigma，重塑程序化广告生态格局**

**信息源：**https://baijiahao.baidu.com/s?id=1836138148578048610&wfr=spider&for=pc

**概述：** 全球程序化媒体合作伙伴MiQ重磅推出MiQ Sigma，这个基于全球最大规模数据集且具有变革性的广告整合平台将由AI驱动。MiQ Sigma以程序化核心优势为基础构建，整合了300多种不同的数据源，涵盖消费者在电视观看、网络浏览以及在实体店购买时产生的700万亿条消费者行为信号。为了提升规划能力、用户挖掘和营销投放效果，该平台借助Sophisticated AI技术助力交易人员为品牌方和代理机构达成卓越的成效。

**主要功能**：

**MiQ Sigma Intelligence：**强大的可视化能力，整合数百种多维数据源，涵盖全球超过 17 亿受众的 “观看、浏览和购买” 行为。通过这些数据的全方位关联，交易人员将被赋予精准触达目标受众和深度洞察市场的能力。

**AI 交易代理**：营销活动管理和优化功能由一个交互式交易代理实现。这款由 MiQ 15 年交易数据积累训练而成，并整合 Claude、Gemini 和 ChatGPT 三大顶级大语言模型的平台，让交易人员能够通过简洁的自然语言指令，轻松完成跨 DSP 平台的快速决策与操作。此外，Sigma 还能通过自动化分析完成此前需要依赖手动分析的流程，大幅提升了工作效率和决策质量。

**生成式 AI 用户画像：**依托生成式 AI 技术，只需简单的自然语言指令，即可在数秒内生成精准的用户画像与营销策略规划。这些目标人群标签可直接应用于营销活动策划，实现从数据洞察到营销活动落地执行的无缝衔接。

1. **阿里通义千问推出多模态模型Qwen VLo，引入全新生成机制**

**信息源：**https://baijiahao.baidu.com/s?id=1836157896287416789&wfr=spider&for=pc

**概述：**阿里云通义千问发布多模态模型Qwen VLo，支持动态分辨率图像生成和逐步清晰的创新机制，突破传统格式限制，为不同场景提供灵活适配，目前处于预览阶段，团队持续优化中。

**主要功能：**

**·更精准的内容理解与再创造：**Qwen VLo 通过更强大的细节捕捉能力，能够在生成过程中保持高度的语义一致性。例如，当用户输入一张汽车的照片并要求“更换颜色”时，Qwen VLo 不仅能准确识别车型，还能保留其原有的结构特征，同时完成色彩风格的自然转换，让生成结果既符合预期又不失真实感。

·预计会上线**同时上传多张图片进行操作整合的功能**。

·对于图片内的物体可以**进行精准识别，并进行切分操作**。

1. **飞利浦重磅推出 8 号 Pro AI 耳机，解锁多语言沟通新体验！**

**信息源：**https://www.aibase.com/zh/news/19332

**概述：**这款耳机采用了腾讯的混元 AI 语言大模型，能够精准支持17种语言的互译和27种方言的识别，极大地方便了用户在国际交流中的沟通。无论是在商务会议还是旅行中，用户都能够轻松应对多语言环境，提升交流效率。此外，耳机内置了四大专业领域的术语库，为用户提供了更为精准的翻译体验。

1. **Runway AI重磅推出“游戏世界” 大型互动文本冒险**

**信息源：**https://www.aibase.com/zh/news/19325

**概述：**6月28日AI技术先锋企业Runway宣布即将推出其全新生成式AI平台“游戏世界”（Game Worlds），这一创新产品不仅标志着Runway从影视领域的成功拓展至游戏行业，更为广大创作者和玩家带来了全新的互动体验。目前仅支持文本与图像的游戏界面。Runway表示未来将推出支持视频形态的游戏生成。

1. **谷歌发布开源多模态模型Gemma 3n，仅需2GB/3GB内存，支持图像、音视频和文本输入**

**信息源：**https://finance.sina.com.cn/tech/csj/2025-06-27/doc-infcnwip2670640.shtml https://developers.googleblog.com/en/introducing-gemma-3n-developer-guide/

**概述：**6 月 26 日谷歌推出Gemma 3n端侧AI模型，可在低至2GB内存设备运行，处理多模态数据，响应速度提升1.5倍，多语言表现优异，现可通过浏览器直接体验。

**核心特点：**

**·革命性内存优化：**

2GB/3GB低内存运行：Gemma 3n提供E2B（5B参数）、E4B（8B参数）两种版本，通过MatFormer架构和逐层嵌入技术（PLE），将内存占用压缩至传统2B/4B模型水平。E2B仅需2GB内存，E4B仅需3GB，首次实现中端手机、平板等设备本地流畅运行多模态AI。

**·性能不减损：**

尽管参数压缩，E4B在LMArena基准测试得分突破1300分，成为首个低于10B参数达成此成绩的模型，推理与编码能力超越同规模竞品。

**·原生多模态支持支持文本、图像、短视频、音频四类输入（输出为文本），实现跨模态理解：**

视觉：搭载MobileNet-V5编码器，在Pixel设备上以60FPS实时处理视频流。

音频：集成高精度语音编码器，支持实时语音转文本、翻译及情感分析。

多语言覆盖：支持140种文本语言及35种语言的多模态交互，非英语场景（如日语、西班牙语）表现突出。

Hugging face：<https://huggingface.co/collections/google/gemma-3n-685065323f5984ef315c93f4>

1. **航空发动机用上大模型：解决复杂时序问题，性能超越ChatGPT-4o**

**信息源：**https://finance.sina.com.cn/stock/t/2025-06-28/doc-infcqzhk2913815.shtml

**概述：**上海交通大学航空航天学院李元祥教授团队、上海创智学院、复旦大学数据科学学院团队以航空发动机运维为背景，**提出高效、可迁移的时序-语言桥接架构——ITFormer**，将专家诊断过程抽象为”理解、感知、推理、决策”四个认知层次，并首次系统性地定义为“时序问答”任务范式。团队基于NASA航空发动机数据，构建了包含11万余问答对的EngineMT-QA数据集。该数据集的任务设计紧密贴合专家的认知流程，为评估模型在真实工业场景下的推理能力提供了首个标准化基准。结果显示，ITFormer以模块化设计实现了时序数据与大语言模型的高效融合，仅需训练不足1%的额外参数，便可在通用时序问答数据集上表现出优越的性能和良好的迁移能力，展现了卓越的“即插即用”特性。它可无缝适配PatchTST、Informer、Crossformer等多种时序编码器，以及Qwen、LLaMA、GLM等不同规模和架构的大语言模型。

项目主页：https://pandalin98.github.io/itformer\_site/

论文链接：https://arxiv.org/abs/2506.20093