



下载量破百亿！中国开源大模型站上世界前沿

本报记者 陈存 实习生 贺紫依

打开多模型聚合平台 OpenRouter，不难发现全球大模型格局已发生明显变化，在最新的统计周期内，周调用量榜单前四席全部来自中国开源大模型，以往这份榜单被海外模型长期占据。据全球最大开源社区 Hugging Face 发布的 2026 年春季报告，中国自研开源模型下载占比达 41%，首次超过美国，成为该平台规模最大的开源模型供给方。

事实上，早在去年，知名风险投资机构 Andreessen Horowitz (a16z) 的合伙人马丁·卡萨多就透露，前来寻求融资的硅谷 AI 初创公司中，使用中国开源模型的比例可能达到了 80%。中国模型的影响力还辐射到了全球南方，来自非洲等地的开发者纷纷使用 DeepSeek、Qwen、Kimi 等国产开源模型底座。据了解，我国开源模型累计下载量已突破 100 亿次，居全球首位。中国开源大模型，正成长为影响全球 AI 产业供给格局的核心力量。

中国开源大模型迎来密集爆发期

回顾开源大模型的发展历程，开源路线对中国 AI 产业实现突围至关重要。2025 年初，DeepSeek-R1 推理模型开源，此后国内大模型企业掀起多轮开源浪潮，持续向外输出前沿技术成果，逐步打破“开源低配”的行业潜规则。数据显示，过去 12 个月里，国产模型有 9 个月保持着全球开源模型的规模上限，迭代节奏持续领跑全球，成为全球人工智能产业格局中极具辨识度的一抹亮色。

今年 7 月以来，中国开源大模型迎来密集爆发期，从参数规模到

技术深度均实现历史性跨越，DeepSeek V4(Flash 及 Pro 版本)、智谱 GLM 5.2、腾讯混元 Hy3、小米 MiMo-V2.5 等开源大模型调用量领跑全球。7 月 27 日，月之暗面正式发布旗舰模型 Kimi K3，以 2.8 万亿总参数一跃成为全球规模领先的开源模型。与 K3 模型权重一起公开的，还有技术报告和三项底层模型训练方法。时隔仅一周，8 月 3 日，阿里巴巴发布新一代基座大模型 Qwen3.8-Max，总参数规模达到 2.4 万亿，计划后续开源。这意味着，国内开源阵营即将再添一款超

两万亿参数的重磅大模型。加州大学伯克利分校计算机科学教授 Ion Stoica 艾恩·斯托伊卡认为，中国开源模型与全球顶尖闭源模型的差距已从 6 至 9 个月收缩到 2 至 3 个月。近日，Hugging Face CEO 克萊芒·德朗格表示，中国在开源权重 AI 模型领域占据领先地位，有望在 2026 年年底甚至更早、最迟 2027 年追上美国前沿大模型公司。这一轮集中爆发，不仅刷新了从业者对中国开源技术水平的评价，更改变了行业对于开源路线的价值判断。长期以来成本与性能难

随着技术实力的提升，国内开源赛道的融资热度逐步上涨，给予业界充足的信心。

以兼得的惯性思维被逐步打破。不久前，英伟达创始人黄仁勋公开称赞开源路线：“开源模型能够加强安全和网络安全，加速创新和推广。世界既需要前沿的封闭模型，也需要前沿的开源模型。”伴随着技术实力的提升，国内开源赛道的融资热度逐步上涨，给予业界充足的信心。近期，DeepSeek 完成成立以来的首轮外部融资，募资总额超 500 亿元；月之暗面 Kimi 完成 F 轮融资，投后估值涨至 350 亿美元，并已启动 G 轮 (Pre-IPO) 融资。

一批中国开源大模型正从“单点亮相”走向“群体突破”，为世界人工智能发展提供新方案、新路径。

多重合力构筑发展底气

当前，中国开源大模型已呈现出鲜明的产业化特色，这离不开政策引导、设施配套与生态协同构成的系统性支撑体系。

政策层面，“十五五”规划纲要提出，“推进开源体系建设，完善开源运行机制”；与此同时，北京、上海、浙江等地相继出台各具特色的开源专项产业政策，支持开源产业发展。

设施层面，国内 AI 开源社区正加速壮大。前不久，工业和信息化部运行监测协调局局长陶青公布了

一组数据：全国性开源基金会吸纳 50 余个开源项目进入孵化阶段，国家级人工智能开源社区汇聚用户 1100 余万、托管模型超 7 万个，覆盖基础软件、大模型及应用的全栈开源方案加速构建。陶青表示，工业和信息化部将继续提升开源基础设施在项目孵化、供应链安全、海外落地等方面的能力，推动国家级人工智能开源社区成长为新兴技术的创新策源地。

应用层面，中国工程院院士、阿

里云创始人王坚表示，开源的内涵正在发生“革命性变化”——从以往“源代码的开放”逐渐转向“资源的开放”。他指出，开源并不代表计算不再重要，而是意味着个体不必重复投入巨大资源。当前，政务服务、电商零售、出行规划、办公协同、医药研发、工业生产等实体经济领域相继接入开源大模型，稳步落地模型应用，显著提升生产生活效率，带动整个 AI 产业链的壮大与繁荣。

让更多开发者参与创新，让更

多企业共享技术红利，让人工智能安全、快速地走进千家万户，这正是开源生态最重要的价值所在。全球不同经济体、不同规模的企业、不同地区的人民群众，都应该有机会握住这把通往智能时代的钥匙。

北京中关村学院院长刘铁岩表示，技术的生命力在于开放流动，一批中国开源大模型正从“单点亮相”走向“群体突破”，为世界人工智能发展提供新方案、新路径，将加快推动人工智能技术向上、向善发展。

湖北：推进商业价值信贷改革 破解中小企业融资难题

（上接第 1 版）

这套“用数据说话，用信用换钱”的数据增信模式，正在慢慢形成一个适合科技和产业变革的新生态。如今，平台归集了 13 个部门 19 项政务数据、近 200 亿条商业数据，有效唤醒“沉睡”数据，破解银企信息不对称问题，实现企业商业价值可评、可见、可贷。

湖北省征信公司正是中小企业商业价值信用贷款改革背后的平台运营方，承担着为这些贷款“增信”的关键角色。

“评价一家企业到底怎么样，落脚点还是经营能力，也就是真实的经营状况。这些怎么判断？我们重点看企业上一年的纳税经营收入、近两年纳税营业收入增长率、纳税评级。”湖北省征信公司副总经理孙涛说道，这些数据不会说谎，比企业自己报上来的数字更可靠。

据了解，这套评价模型如今已经接入湖北全省 22 家银行，让曾经模糊不清的中小微企业画像，逐渐变得清晰起来。现在，已经有超过 7 万家中小企业通过湖北省中小企业商业价值信用贷款运行平台申请贷款。其中，2645 家此前从未获得过银行贷款的“首贷户”成功获贷 74.8 亿元，占平台总获贷户数的 8.84%。

强化风险共担 实现财政增信

为权衡收益与风险，一些金融机构宁可不贷也不错贷。为打消银行后顾之忧，湖北省探索财政增信、风险共担机制，最大限度发挥财政资金杠杆作用。依据湖北省出台的风险补偿办法及实施细则，省与市州按 1:1 比例首期设立超 10 亿元信用贷款风险补偿金池，并建立动态增补机制，将风险补偿熔断率提升至 8%，政府对贷款本金损失 50% 予以风险补偿支持。

“我们享受了量身定制的金融服务。”武汉博旺兴源环保科技有限公司总经理余军在接受采访时表示，这些年公司想要扩大产能、升级技术，但因为缺乏传统意义上的抵押物，融资很难。湖北银行在深入了解该企业专注于资源循环利用的业务特点后，为其量身打造组合融资方案，给予 1000 万元贷款，解了企业燃眉之急。

数据显示，目前，湖北省市两级已统筹落实 1.5 亿元并做实资金账户，有效激发了银行“敢贷、愿贷”积极性。

提升服务质效 精准赋能企业

为提升服务质效，湖北省聚焦企业信贷体验，将平台与银行系统直连，实现“企业信贷、信用评价、银行受理、授信反馈”全流程线上办理，银行平均授信时间压缩至 3.10 天，最快当天即可审批放款；商信贷平均利率 3.29%，较同期全省小微

（上接第 1 版）

另一方面，要持续改善、消除当前光栅技术的缺陷。一是消除彩虹纹，当外界自然光经过衍射光栅，会被色散为彩虹状的杂光，干扰消费者的视野，因此需要消除这种杂光的干扰；二是解决漏光，避免波导片显示的图像被外界捕获，泄露用户隐私；三是提升透光率，最好能像高端近视镜那般，将透光率提升至 98% 以上，为用户提供明亮的视野；四是减弱“光栅感”（指镜片有光栅结构的区域与其他区域在颜色上的区别），如何将光栅做得很透明，让消费者感觉不到，也是一个难点。

第二个层面是光学指标。

首先是下游客户最关注的视场角，视场角越大则沉浸感越强；其次是光效（对光能的利用效率），如果光效足够高，即便用较低的光机驱动电流，也能产生高亮画面，从而提升续航能力并降低电池重量；再次是画面均匀性，可以通过优化光学设计实现画面均匀性，也可以效仿显示行业，推出针对均匀性的矫正算法；最后是对比度、清晰度等细节指标。

在追求极致光学性能的同时，商业化与工程落地同样重要，关乎成本与产能。

南京平行视界技术有限公司（以下简称“平行视界”）CEO、东南大学教授张宇宁向《中国电子报》记者指出，AI 眼镜厂商对光波导商业落地能力的关注体现在三个方面：一是综合模组成本。波导模组及相关光组件成本的下探空间决定了整机能否进入消费级价格区间。二是规模化产能与交付。产线的规模化量产能力与良率稳定性，是大批量、高一致性交付的保障。三是环境信赖性。包括耐高低温、防跌落、抗湿热等消费电子级可靠性测试标准。

此外，下游厂商非常看重光波导技术路线的演进上限与瓶颈。“下游厂商会评估该技术路线是否存在难以克服的物理或工程瓶颈，比如光效上限过低、全彩化过繁、叠层致镜片过厚等，倾向于选择兼具当前可用性与未来升级空间的方向。”张宇宁说道。

多条技术路线竞跑 光波导量产按下“加速键”

面向 AI 眼镜带来的市场机遇，光波导的主流技术路线开启量产竞速。当前，AR 眼镜光波导技术路线主要分为几何光波导、表面浮雕衍射光波导 (SRG)、体全息光波导 (VHG) 与偏振体全息光波导 (PVG)。在物理机制上，几何光波导采用几何光学，SRG、VHG、PVG 采用衍射光学。

在衍射光学阵营中，SRG 在 market 占比、产业成熟度和客户接受度上占据主导。孟祥峰告诉记者，SRG 的优势在于可以依托半导体工艺，材料、工艺、设备等环节均有半导体领军企业入局，形成全产业链诸多主体共同贡献的“飞轮效应”。但其良率提升也面临两大挑战：其一是外观品质，无论是颗粒物的影响还是划伤，都会造成镜片瑕疵，在光与波导片交互的过程中被照亮并放大；其二是光学一致性，在连续生产上百万甚至上千万片镜片的过程中，要保证每一片的光学性能（比如光效、色彩均匀性等光学指标）都在设计目标的容差范围内。

早在 2023 年，至格科技就建成了全自动衍射光波导批量生产线，2026 年力争交付 100 万片光波导镜片。孟祥峰表示，目前至格科技直通良率已达 85% 以上，部分项目超过 90%。“在开发新材料、新工艺的过程中，我们一直推动供应商共同进步。例如，通过特制机台和开发工艺，将产品良率提升几十个百分比，这也带动了行业的技术进步和供应商的业绩增长。我们希望通过持续打磨技术，发挥在供应链中的引领作用，推动行业始终向前。”孟祥峰说道。

与此同时，包括 HPDLC（全息聚合物分散液晶）、PVG 在内的“新一代体全息”技术正在异军突起。张宇宁向记者表示，PVG 路线不仅继承了体全息高光效、低漏光的固有优势，更通过偏振液晶的分子自组装响应特性，打通了大 FOV 的演进通道，且具备自动化大面积生产与综合降本潜力，展现出广阔的前景优势。

在工程化层面，PVG 量产的核心难度有两点。一是缺乏现成装备与经验可循，目前市场上没有专门针对 PVG 特殊光学结构与材料体系的大面积、高精度量产设备。二是缺乏“全链路深度耦合”，PVG 的规模

化量产无法通过直接采购标准化设备来实现，必须从底层的光取向、液晶材料研发出发，反推并摸索出匹配的工艺窗口，再基于工艺需求去重新定义并定制非标自动化设备，最后在实际生产中不断调校设备与工艺参数，跑通稳定性并提升良率。

“这就要求必须打通从底层材料、光学工艺到设备控制的全流程，进行系统性的同步开发。这也需要研发团队既懂底层材料，又懂光学工艺和设备定制，具备全链条的工程化落地能力。”张宇宁向记者指出。

2025 年 12 月，平行视界的偏振体全息 (PVG) 光波导中试产线在江西吉安正式落成，具备年产 5 万片光波导片能力。张宇宁表示，平行视界主要通过“高校科研团队的源头创新”与“产学研深度融合的工程落地”来打通量产路径。首先是发挥高校科研团队的底层研发底蕴，从源头上攻克偏振液晶材料的分子设计、相变控制与信赖性难题，为量产提供底层理论与材料支撑；其次是贯通“学术突破—工程转化—设备定制”全链条，立足自主研发的材料体系，与上游装备厂商开展深度合作，共同定义并定制开发专用的自动化涂布与曝光固化设备；再次是借助中试产线跑通工业化良率，持续积累产线数据并优化拟合参数，逐步打磨出高良率、高一致性的标准化生产流程，为后续更大规模的商业化量产奠定基础。

而不依赖光栅的几何光波导有着较高光效、较好的显示色彩和视场角，但在加工工艺上，需要将多层镜片按照一个斜角磨平、镀膜，并叠在一起，量产难度较大。理湃光晶通过“二维扩瞳”技术、自研分子键合工艺等路径，提升几何光波导的良率和量产性。今年 5 月，理湃光晶宣布在平谦迈高园区全面部署新一代智能光学产线，新址投产后，其光学模组年产能将从 10 万套提升至 100 万套。

避免低质量“内卷” 以原创性技术与用户体验突围

虽然 AR 光波导已经开启量产窗口，但产能数字并不能代表产业的成熟度和健康值。原创性技术突破与务实可行的研发、生产目标，对于行业的整体进步更加关键。

当前，AR 光波导的产业化还存在三个层面的问题。

在硬件端，全彩 AR 的产业链有待完善。张宇宁指出，一方面，上游微像源/光机（如全彩 Micro-LED）的良率、光效、分辨率和色彩表现等仍有待突破；另一方面，光波导实现全彩化的成本过高，且在画面均匀性、色偏等性能方面仍存在客观瓶颈。

在工艺端，仍有多项技术有待攻克、多条路线有待验证。孟祥峰提到，当前纳米压印的量产技术已经相对成熟，但仍需攻克 12 英寸的量产。同时，行业还在引入半导体刻蚀工艺，结合折射率更高的材料推动波导片的高端化。

在生态端，内容与场景生态尚未就绪。尽管大模型赋予了眼镜更强的感知能力，但全彩显示依然要找到不可替代的交互价值和用户刚需，应用场景与内容生态仍处于摸索阶段。

要推动 AR 光波导片的高端化发展，加速商用落地进程并提升市场规模，既要在用户体验上接地气，也要在原创技术上敢作为。

围绕用户体验，张宇宁建议短期内将重点放在推进单色波导片，先把“佩戴体验”做极致。“产业不必盲目跨越阶段。短期内，把单色光波导在高透光、低功耗、极致轻薄与无感佩戴上做到极致，满足真正低代价的信息可视化功能场景需求，让用户能够并愿意‘长期佩戴’，这本身就是非常关键的跨越。”他表示。

对于原创技术，张宇宁提到，要坚持底层技术创新与产学研深度协同，立足于自主知识产权，推动关键材料与专用自动化装备的国产化突破。同时，避免同质化低价竞争，扎实打磨量产良率，推动光波导模组成本合理下探。

孟祥峰也强调，光波导产业是双门檻，既有光学设计门檻，又有半导体制造门檻。现在行业中存在重产能、轻研发的现象，忽视对光学设计的研发投入，甚至存在盲目模仿现有产品并压低价的行为，导致产业陷入“内卷”可能。“希望广大从业者意识到，我们这个行业由光学设计与半导体制造双轮驱动，如果重产能、轻研发，会伤害研发人员的积极性，导致行业陷入低水平重复竞争。”孟祥峰吁喊道。