

当大模型和机器人“陪你”阅读 上海书展正在重新定义“翻书”这件事

机器人“穿上”齐天大圣皮肤，一开场就生龙活虎，向小朋友们讲述《西游记》绘本；读者将一本书的文字输入大模型；十几秒间，就能在屏幕中展现出活灵活现的画面……

从 2026 上海书展开始，这些场景都会变成寻常的片段，当“最是书香能致远”的古老命题，遇见大模型与具身智能的当代技术，书展不再只是书的集市，而成了未来阅读形态的实验场。

从纸面到像素，大模型的“视觉翻译”

阅读的乐趣，历来在于“一千个人眼中有一千个哈姆雷特”。今年上海书展期间，MiniMax 带来的技术演示，为这种私人化的想象提供了一种新的“显影”方式：如果用户输入一段《乌苏里长歌》的文字，十几秒后，一段带声音、有画面的 15 秒视频便流淌而出：黑龙江的冰雪、乌苏里江的晨雾、赫哲族的鱼皮鼓，从抽象的文字符号转译为直观的视听语言。

支撑这一幕的，是 MiniMax 在 7 月 31 日正式发布并开源的新一代多模态生成模型 H3。

与此前按图片、视频、音频等单一任务划分边界的思路不同，H3 试图在统一的多模态上下文中理解创作意图，将生成、编辑和参考视为连贯的创作流程，而非割裂的技术环节。

在 ArtificialAnalysis 的视频模型榜单中，H3 的视频编辑能力位列全球第一；它支持文本、图片、音频和视频等多种输入形式，可直接输出 2K 分辨率画面，最长生成 15 秒音画内容。

更具产业意味的是其成本结构。H3 的视频生成定价为 0.8 元 / 秒（2K 分辨率），约为业内同类旗舰模型的三分之一，这一成本优势来自底层技术的系统性优化：高压缩 Tokenizer 减少了视频生成所需的 Token 数量，而异构训练、负载均衡与 GPU 利用效率的提升，则在效果与成本之间找到了新的平衡点。

MiniMax 相关负责人告诉记者：“通过 MiniMaxH3，我们希望为更多企业和开发者、创作者降低内容创作的门槛。企业可以在本地灵活部署，结合自身数据和业务进行优化，更好地满足安全合规要求；芯片厂商和开发者也能共同参与适配和优化，降低使用成本，扩大应用范围。”

确实，当视频生成的成本下降到如此区间，出版机构将图书内容批量转化为视频衍生品的商业逻辑，才真正具备了可行性。

对于上海书展而言，这意味着一个在场与在线同步的、可交互的内容生态。一本《金瓶梅》可以衍生出历史场景的 AI 复原视频，一套“中华经典藏书”可以生成适配不同年龄段读者的可视化导读。技术没有消灭“一千个哈姆雷特”，而是为每一个哈姆雷特提供了一条更直观抵达故事深处的路径。

具身智能的温度，机器人的“河畔故事会”

如果说 MiniMax 解决的是“内容如何被看见”的问题，擎天租和智元机器人回答的则是“阅读如何被体验”的问题。在上海书展的“河畔故事会”特别活动中，具身智能以一种极为具体的方式走进了阅读场景。

负责智元机器人书展落地的擎天租相关负责人告诉记者，上海书展期间，他们将围绕“你好，故事新朋友”，五天时间每天一个主题，由机器人来讲解绘本。

比如面向 10-14 岁青少年的“小人儿发明家——造机器人专场”，孩子们会以灵犀 X2 人形机器人为载体，从基础认知到动作编创，完成一次从“读者”到“创作者”的身份转换。

在这个主题里，孩子们学习机器人的结构、感知、控制与执行四大模块，理解机器人为什么会“动”、会“听话”；他们也会分组设计 10-20 秒的安全短舞，通过灵创平台的“动作模仿”功能，将自己拍摄的视频素材转化为机器人动作素材，经过时间轴编辑与合成保存，最终下发到灵犀 X2 真机上完成舞蹈演绎。整个流程的闭环，让孩子们亲手验证了“人类动作如何变成机器人动作”，这对阅读中“角色代入”这一行为的物理化延伸。

同时，IP 角色演绎



也成为有趣的体验。智元的机器人穿上西游记中齐天大圣的戏服，匹配对应的音色进行节目表演；机器狗披上熊猫皮肤，与阅读《穿鞋子的小狗》的孩子们互动。在英语绘本专场中，机器人甚至以英文开场，配合肢体动作完成双语阅读引导。

这些设计都可谓精准击中儿童认知的特点：对于低龄读者，具身智能提供的多模态刺激（动作+语言+视觉）能显著延长注意力窗口，而“与机器人握手、合影”的结束环节，则天然具备了社交传播属性——家长的朋友圈照片，成了书展最有效的口碑广告。

擎天租的相关负责人还表示，书展期间，智元的机器人在馆内还承担着迎宾接待、导览问询与安防巡检的职能。基于多语言大模型的自由问答能力，机器人可以中英文灵活切换，为上海书展期间到访的外籍作家与读者提供服务。

这种“业务协助”角色，让具身智能从舞台中央走向了基础设施，也展现出未来图书馆与书店的一种可能形态：机器人不是阅读的替代者，而是阅读环境的有机组成部分。



数智共创，或是出版业的“第二曲线”

将视野从具体产品拉回到产业层面，2026 上海书展的数智化布局呈现出清晰的层次感。

政策层面，2026 年《全民阅读促进条例》正式施行，国务院政府工作报告首次提出“支持出版业繁荣发展”“支持实体书店发展”，上海市政府工作报告明确写入“加快出版领域数智化转型”。书展作为行业风向标，其“数智共创”主题并非应景之作，而是对政策导向的市场化回应。

技术层面，MiniMax、擎天租和智元机器人的组合，恰好覆盖了出版业数智化的两个关键维度：内容生产端的智能化（AI 生成视频、AISKILL 技能大赛）与消费体验端的场景化（机器人辅助阅读、具身智能互动）。前者解决的是“书如何被重新生产”的问题，后者解决的是“书如何被重新阅读”的问题。两者叠加，构成了从 B 端到 C 端的完整链条。

更深层的洞察在于，上海书展正在探索一条区别于单纯“技术赋能”的路径。它没有将 AI 视为降本增效的工具，而是将其作为“阅读新场景”的构建者。无论是 MiniMaxH3 将文字转化为视频的“视觉翻译”，还是智元机器人将绘本角色实体化的“具身演绎”，核心逻辑都是扩展阅读的边界，而非替代阅读本身。

这种探索发生在出版业的一个微妙时刻，实体书店转型“城市会客厅”已成共识，长三角品牌实体书店展销区今年继续以“一店一主题，家家有亮点”的方式呈现；但“城市会客厅”究竟该提供何种体验，仍无定案。

MiniMax 与擎天租和智元提供的方案展现了一种方向：未来的书店与书展，可能是“内容生成器 + 具身交互终端”的复合空间，读者在这里不仅消费内容，还参与内容的再生产。

晨报记者 苗夏丽 姜欣愉