

## 可灵AI全系模型进入2.0时代 让每个人都能用AI讲出好故事

2025.04.15

香港，2025年4月15日 (GLOBE NEWSWIRE) -- 领先的内容社区及社交平台快手科技（“快手”或“公司”；港币柜台股份代号：01024 / 人民币柜台股份代号：81024）宣布，可灵AI在北京举行“灵感成真”2.0模型发布会，宣布底座模型再次升级，面向全球正式发布可灵2.0视频生成模型及可图2.0图像生成模型。

作为全球首个用户可用的DiT视频生成模型，可灵AI自去年6月上线至今的10个月时间里，月活用户数量增长25倍，全球用户规模已突破2,200万。3月27日，全球知名AI基准测试机构Artificial Analysis发布了最新的全球视频生成大模型榜单，快手可灵1.6 Pro（高品质模式）以1,000分的Arena ELO基准测试评分登陆图生视频（Image to Video）赛道榜首，Google Veo 2、Pika Art位居榜单第二、三名。

自去年6月发布以来，可灵AI已累计完成超20次迭代。一直以来，可灵AI致力于提升模型基础质量，增强画质清晰度，并引入更多创新功能，以满足用户多样化需求。快手高级副总裁、社区科学线负责人盖坤表示，可灵希望“让每个人都能用AI讲出好故事”，实现更加精准的复杂创意表达。

### 可灵AI全系模型进入2.0时代，定义人和AI交互的全新方式

据悉，可灵2.0模型在动态质量、语义响应、画面美学等维度，保持全球领先；可图2.0模型在指令遵循、电影质感及艺术风格表现等方面显著提升。盖坤介绍，两款模型在团队内部的多指标横向评测中，均稳居业内第一。例如在图生视频领域，可灵2.0对比谷歌Veo2的胜负比为182%，对比Runway的Gen-4的胜负比达178%，在语义响应、画面质量、动态质量等维度上显著超越对手。



(图说：快手高级副总裁、社区科学线负责人盖坤)

盖坤认为，AI在辅助创意表达上拥有巨大潜力，但当前的行业发展现状还远远无法满足用户需求，在AI生成内容的稳定性、以及用户复杂创意的精确传达上仍有“很多挑战”。也因此，要真正实现“用AI讲好每一个故事”的愿景，必须对基模型能力进行全方位提升，定义人和AI交互的“全新语言”。

在本次2.0模型的迭代中，可灵AI正式发布AI视频生成的全新交互理念Multi-modal visual language (MVL)，让用户能够结合图像参考、视频片段等多模态信息，将脑海中包含身份、外观、风格、场景、动作、表情、运镜在内的多维度复杂创意，直接高效地传达给AI。

“大家应该很自然地会感受到，文字在表达影像信息时是不完备的。我们需要有新的方式，让大家真正精准地表达出心中所想。”盖坤指出，MVL由TXT (Pure Text, 语义骨架) 和MMW (Multi-modal-document as a Word, 多模态描述子) 组成，能从视频生成设定的基础方向以及精细控制这两个层面，精准实现AI创作者们的创意表达。

基于MVL理念，可灵AI正式推出多模态编辑功能。“从今天起，大家可以在可灵AI的平台上体验我们的多模态编辑能力。用户可以非常直接地把自己的想法用图像等方式表达作为输入，生成符合自己想法的创意视频。”盖坤介绍，MMW将不只局限于图片和视频，也可以引入其它模态的信息，例如声音、运动轨迹等，来让用户实现更加丰富的表达。



(图说：“灵感成真”可灵AI 2.0发布会现场)

#### 可灵2.0大师版正式发布，全新上线多模态视频编辑功能

基于MVL这一全新交互方式，在发布会现场，快手副总裁、可灵AI负责人张迪发布了全新的可灵2.0“大师版”，在语义响应、动态质量、画面美学等方面显著提升生成效果。包括大幅提升指令遵循能力、显著增强电影美学表现力以及支持60多种风格化的效果转绘，模型出图创意和想象力实现大幅跃升。



(图说：快手副总裁、可灵AI负责人张迪)

值得一提的是，可灵2.0大师版全面升级视频及图像创作可控生成与编辑能力，上线全新的多模态视频编辑功能，能灵活理解用户意图。该版本支持在一段视频的基础之上，通过输入图片或文字，对生成的视频内容实现元素的增加、删减、替换，帮助创作者们实现更加灵活的二次编辑和处理。

同时，可图2.0也上线了实用的图像可控编辑功能——局部重绘和扩图，支持图片的增加、修改和修复。在图像的多模态可控生成中，可图2.0还上线了全新的风格转绘功能，只需要上传一张图片加上风格描述，就能一键切换图片的艺术风格，同时精准保留原图的语义内容。



(图说：可灵2.0大师版操作界面)

张迪表示，当前，图生视频约占到可灵AI视频创作的85%，图片质量也对视频的生成效果产生重要作用。在图像生成大模型领域，快手的可图处于行业领先水平，拥有多项核心优势，如强大的复杂语义理解能力、电影级的画面质感、多条件的可控风格化生成等。在团队内部的多项胜负率评测中，相较于Midjourney V7、FLUX1.1 Pr以及Reve等行业领先的图像模型，均保持明显优势。

“可灵2.0大师版不仅是一次简单的技术升级，更是一次全方位的体验跃升。”张迪介绍，可灵2.0大师版实现了技术、体验、美学的三重突破。

#### **AIGC技术迅猛发展，为全行业发展注入全新活力**

除了C端用户订阅模式，可灵AI也面向B端企业客户提供API接入等服务。目前，可灵AI已与包括小米、亚马逊云科技、阿里云、Freepik、蓝色光标等在内的数千家国内外企业客户建立了合作关系。

盖坤介绍，来自世界各地的超1.5万开发者和企业客户，已将可灵的API应用于不同的行业场景中，累计生成的图像数量约1,200万个，生成的视频素材超过4,000万个。

如今，可灵正在成为AI时代视频创作的新基础设施。AIGC技术的迅猛发展，也正在重构包括广告营销、专业创作、影视、娱乐创意等在内的多个行业。

可灵AI超级创作者、《新世界加载中》总导演、异类Outliers创始人陈翔宇在分享中介绍，可灵AI可全面融入剧集级创作流程，AIGC相较于实拍和动画，不仅是效率的提升，更是试错空间的革命性释放。

从剧本创作到内容发行，异类Outliers团队完成了一次完整的影像工业化创作流程，覆盖从美学表达镜头结构，再到复杂动作场景与细节的全面呈现。“我们实践之后发现，可灵AI是一个能够稳定、大规模嵌入剧集创作流程的生成式协作大模型。”陈翔宇认为，在AI的辅助下，导演及编剧团队的大量内容创意也实现了更富有想象力的表达，“AIGC也许是下一代内容结构的雏形”。



(图说：可灵AI超级创作者，《新世界加载中》总导演，异类Outliers创始人陈翔宇)

为了进一步激发AI爱好者们的创作热情，快手副总裁、可灵AI负责人张迪还在发布会现场正式发起“可灵AI NextGen新影像创投计划”。该计划将加大对于AIGC创作者的扶持力度，通过千万资金投入、全球宣发、IP打造和保障，以全资出品、联合出品和技术支持等灵活多样的合作方式，让AI好故事走向世界。同时，可灵AI面向全球创作者发出征集邀约，邀请用户共同打造全球首支用户共创AIGC创意短片，让创作者的灵感和创意登陆上海、香港、东京、巴黎、多伦多等世界多地广告大屏。

面向未来，盖坤表示，可灵AI将持续大力推动技术创新，用人与AI交互的全新语言，帮助用户实现复杂创意的精确表达，“我们的初心，是让每个人都能用AI讲出好的故事，我们也真切地希望这一天更快到来”。

## 关于快手

快手作为中国乃至全球领先的内容社区及社交平台，致力于成为全球最痴迷于为客户创造价值的公司。作为一家以人工智能为核心驱动和技术依托的科技公司，快手专注于通过持续的技术创新和产品升级，不断丰富服务和应用场景，为客户创造价值。在快手，用户通过短视频和直播来记录和分享他们的生活，发现所需，发挥所长。通过与内容创作者和企业紧密合作，快手提供的技术、产品和服务可满足用户的多元化的需求，包括娱乐、线上营销服务、电商、本地生活、游戏等。

## 前瞻性声明

除过往事实的陈述外，本新闻稿载有若干前瞻性陈述。前瞻性陈述一般可透过所使用前瞻性词汇识别，例如“或会”、“可能”、“可”、“可以”、“将”、“将会”、“预期”、“认为”、“继续”、“估计”、“预计”、“预测”、“打算”、“计划”、“寻求”或“时间表”。该等前瞻性陈述受风险、不确定因素及假设的影响，可能包括业务展望、财务表现预测、业务计划预测、发展策略及对我们行业预期趋势的预测。该等前瞻性陈述是根据本集团现有的资料，亦按本新闻稿刊发之时的展望为基准，在本新闻稿内载列。该等前瞻性陈述是根据若干预测、假设及前提作出，当中许多涉及主观因素或不受我们控制。该等前瞻性陈述或会证明为不正确及可能不会在将来实现。该等前瞻性陈述涉及大量风险及不明朗因素。鉴于上述风险及不明朗因素，本新闻稿内所载列的前瞻性陈述不应视为董事会或本公司声明该等计划及目标将会实现，故投资者不应过于依赖该等陈述。除法律要求的情形外，我们并无责任公开发布可能反映本新闻稿日期后发生的事件或情况或可能反映意料之外事件的该等前瞻性陈述的任何修订。

## 投资者及媒体问询

快手科技  
投资者关系  
邮箱：[ir@kuaishou.com](mailto:ir@kuaishou.com)