AI Content Generation Tools Analysis

We will explore AI content generation tools from three perspectives.

First, we'll examine the market demand from the product perspective. Second, we'll analyze the market supply from the technical perspective. Finally, by synthesizing both the demand and supply situations, we'll deduce future trends, aiming to provide insights into the people, companies, and developments we should pay attention to in the future.

Let's begin with the first part: examining demand from the product perspective.

In 2024, AI-generated content has become increasingly prevalent across domestic and international social media platforms.

Pixverse's Venom transformation and Pika's squeeze-and-transform features have brought users entirely new visual experiences, with Pixverse alone achieving 690 million views on a single platform.

Remini's clay-style effects and Recraft's film-style aesthetics have satisfied users' artistic preferences, garnering 15 million views on Xiaohongshu.

Content featuring babies singing and cats dancing has catered to the psychology of proud parents and pet owners sharing their loved ones, triggering another wave of AIGC combined with UGC content storm on TikTok.

The ability to embrace deceased loved ones, something impossible in reality, has provided users with emotional solace.

Do people enjoy AI-generated content? The answer is self-evident. AI has effectively met users' content needs in terms of entertainment, aesthetics, satisfying their desire to show off, and providing psychological companionship.

So, which products and companies are behind these viral phenomena? We conducted a comparative analysis of over 20 most influential and actively used AIGC products in the market, primarily focusing on video and image domains.

In the video domain, we selected 10 domestic and international products and models for horizontal comparison, identifying several common characteristics among existing products.

We believe that current AI video generation products have reached comparable levels in artistic expression, particularly in stylized presentation of static scenes and realistic rendering, where model capabilities have generally improved.

More discussions now focus on the representation of dynamic scenes, such as models' semantic understanding of professional prompts, camera control, character consistency, and physical laws.

We also conducted in-depth experience testing and summarized their distinctive features. For instance, Keli and Runway currently stand out with the most comprehensive capabilities.

In terms of human representation, Jimeng AI excels at single-subject facial expression rendering.

Hailuo is particularly suited for motion shots in large scenes.

PixVerse's entertaining features are conducive to rapid viral spread.

Pika demonstrates superior understanding of complex instruction words.

Some major companies leverage their computational power and technical advantages to excel in various aspects of resolution, generation speed, and duration.

We've also noticed several new directions under exploration.

For example, Sora's introduction of "storyboard" features, and the frame completion functionality launched by Keli, Jimeng, Runway, Luma, and Vidu are all exploring ways to improve users' AI creation workflow experience.

In Meta's yet-to-be-released MovieGen, "audio-visual synchronous generation" is noteworthy, allowing users to generate contextually appropriate audio for target scenes using either text or video as input prompts.

We applied the same analytical logic to evaluate top-ranking models and products in the image domain based on artificial analysis, Similarweb, and Sensor Tower rankings.

Current AI image processing, whether AIGC or multimodal understanding and recognition, has successfully broken through the crucial 1K resolution threshold. For example, Flux.1, Imagen 3, Midjourney v6, and Stable Diffusion 3 all support mainstream 1024×1024 native image resolution (single generation) and can achieve 2x or higher pixel expansion through super-resolution technology, further enhancing image details.

Additionally, similar to the video domain, current image products have made significant progress in controlling lighting, contours, and image composition.

They can now assist in image generation, applying to scenarios such as artistic creation, daily entertainment, and game development.

For instance, Recraft can achieve relatively complex semantic understanding, Canvas has reached relatively mature application scenarios, and Meitu Xiuxiu leverages its first-mover advantage to achieve optimal human lighting effects.

During our practical functionality testing, we also discovered some new approaches in image generation. For example, domestic products have now mastered Chinese and English text generation, with the next step focusing on exploring "text layout" and more precise structural control.

Regarding new product features, Google Lab's Whisk supports users in uploading and combining "theme," "scene," and "style" images, shifting from traditional text prompt instructions toward style transfer between images. The concept of "Prompt less, Play more" is becoming another noteworthy trend in the image sector.

Next, Zoe will present insights regarding the demand aspects.