

只需输入一段文字指令，AI就能落笔成文、手绘美图、生成流畅的短视频。如今，生成式人工智能(AIGC)早已跳出科技圈，融入大众日常生活。很多人好奇：没有大脑，没有感知的AI，到底是如何凭空写出文章、画出漫画、制作视频的？今天我们就来拆解一下AI创作的奥秘。

主笔：于梅君

# 一键“生”万物

## 揭秘生成式AI的创作魔法

### 1 AI写文章：超级文字接龙

数据显示，截至2025年12月，中国生成式AI用户规模已达6.02亿人，也就是说，每10个网民中，大约有5个日常在使用生成式AI。

很多人误以为AI写作是像人类一样思考、抒发情感，实则完全不同。依托大语言模型的AI写作，本质其实是一场极致精准、无限流畅的超级文字接龙游戏。

想象一下，你和朋友玩词语接龙——你说“今天”，朋友接“天气”，你再接“真好”，就这样一个词一个词接下去，最终组成一个通顺的句子。AI写文章的方法，本质上就是这种“接龙”过程的极致放大版。

在AI正式“上岗”前，它要阅读海量书籍、文章、对话记录。也就是说，它在学习阶段被喂入了几乎整个人类文明的书面资料，它并不是在“背”这些内容，而是在学习其中的统计规律——当写了“我饿了”之后，下一句会接“去吃饭”还是“点外卖”？

这种在序列中不断预测下一个元素的能力，正是大语言模型的核心。大家熟知的ChatGPT、DeepSeek、通义千问等产品，都是基于这种思路构建的。

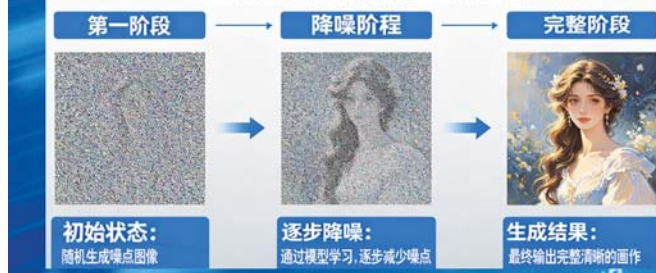
AI如同通读了全网书籍、新闻、散文、百科、学术资料的“超级学霸”，把世间所有文字搭配规律、语序习惯、语境逻辑、行文范式全部储存在模型中。它不懂喜怒哀乐，不懂风景意境，不懂情感表达，只牢牢记住：在特定语境下，哪几个字，哪几个词挨在一起最通顺，哪句话衔接最自然，前后内容怎样组合不会逻辑断裂。

比如，当你输入“春日风光秀美”的指令，AI会先把文字拆分成最小的语言单元，再转换成机器能看懂的数字编码，快速锁定“春日”“风光”等关键词，随后顺着语境，一个词、一句话依次推算，优先选择出现概率最高、最贴合语境的内容，一字一词层层递进，最终完成整篇文章。

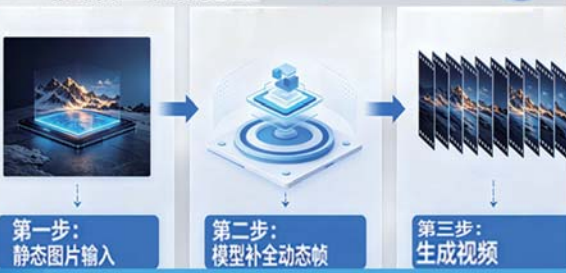
为了让AI文笔自然不生硬，研发人员用上万亿字的真实文本数据训练模型，还通过人工反馈不断纠错优化，告别机械冰冷的话语。

当然，AI靠概率组词造句，并非真正理解内容含义，偶尔会出现“一本正经胡说八道”、编造事实的“AI幻觉”，因此，AI生成的文字不能直接照搬，必须人工核对把关。

### AI绘画扩散模型原理



### AI视频生成原理



### 2 AI画图：从漫天噪点里，一点点勾勒世间万物

如果说AI写作是在模仿人类逻辑的“文字接龙”，那么AI绘画玩的就是一场视觉魔术。它能在几秒钟内，根据一段文字描述生成一张高质量图片。这项技术的核心，就是扩散模型。

扩散模型的原理，说起来有点像一种“破坏再重建”的艺术。想象一下，你有一张清晰的图片，你不断往上面添加随机的噪点(就是我们常说的雪花点)，随着每一次加噪，图片越来越模糊，直到最后变成一片毫无意义的随机噪点图像。

AI在训练阶段，就看了无数次这样的“破坏”过程——它学会了如何一步步将一张有意义的图像变成纯噪点。它提前学习数亿张高清图，搭配对应的文字描述，比如，把“古风山水”“阳光草地萌宠”等各类画面的像素、色彩、光影、构图规律烂熟于心。

而真正的魔力发生在反向过程：当我们给AI输入一句描述，比如“一只在

海滩上奔跑的金毛犬”，AI会先读懂文字里的核心元素与风格，随即生成一张完全随机的噪点图，如同一张被彻底弄脏的白纸。

紧接着，它进入一遍遍的去噪打磨环节，慢慢去掉杂乱斑点，勾勒物体轮廓，填充层次色彩，调整明暗光影，最后细化纹理、修正比例，短短数十秒，就从一团模糊噪点里，变出一张贴合描述，逼真细腻的图片。

如今AI绘画不仅能精准还原文字指令，还能驾驭油画、动漫、写实、国风、科幻等多种艺术风格，即便没有绘画功底，普通人也能轻松创作艺术级美图，彻底降低了视觉创作的门槛。

不过，像写文章一样，有时AI会生成一些逼真但不存在的内容，这种“幻觉”并非bug，而是AI工作原理的固有特点，它并不知道什么是对错，它只是在统计规律上猜测“最可能”出现的画面。

### 3 AI视频创作：从静止到流动魔法

如果说生成图像是AI魔法学校的基础课，那生成视频就是一场高级魔术表演。简单来说，AI做视频，就是在绘画本领之上，又学会了时间魔法，不光懂画面长什么样，更懂画面随时间变化的规律。风吹云动、水流潮涌、行人迈步、动物奔跑、光影变幻，世间万物的运动轨迹，AI都通过海量真实视频片段提前学透。AI生成视频主要分两大主流方式。

一是文生视频，即一句话生成完整的动态画面。这是当前文生视频的主流技术路线。整个过程分三步：第一步，AI理解文字描述，拆解出主体、场景、动作、时长、风格等关键信息；第二步，生成初始画面，同时预判每一秒的运动变化，比如“小狗奔跑”，AI会学习小狗四肢迈步、身体起伏的运动规律；第三步，逐帧生成画面，保证每一帧细节准确，前后帧运动连贯，最后组合成完整视频。

视频生成最难的是运动连贯性。比如生成“海浪拍打礁石”的视频，AI不仅要画出海浪和礁石，还要精准模拟水花溅起、回落、消散的全过程。为了攻克这个难题，AI会学习海量真实视频片段，记住万物运动、光影流转的规律，最终让画面自然流畅。

二是图生视频，这是更简单的视频生成方式。比如把一张猫咪图片，生成猫咪摇尾巴、眨眼的动态视频。其原理是：AI先识别图片中的主体、背景、光影，预判合理运动轨迹，再生成连续画面，既保持原图风格、色调不变，又让物体自然运动，让静止照片拥有生命力。

虽说AI视频创作越来越成熟，但目前仍有短板，比如长视频细节容易失真、复杂动作连贯性不足，各大科技企业仍在持续优化，未来AI视频的时长、清晰度、真实度都会不断提升，带给我们更惊艳的视觉体验。

### 知多一点

无论是写文章、画图片，还是做视频，生成式AI的创作能力，都依托三大核心技术。

#### 海量数据训练是“知识宝库”

AI没有天生智慧，所有文笔、画风、动态逻辑，全都来自日复一日的海量学习，看得越多、学得越全，AI输出的内容就越精准自然，创作风格也就越丰富。

#### Transformer架构是“超级大脑”

这是谷歌2017年提出的技术，是所有大模型的核心。它的“注意力机制”，能让AI精准抓住关键信息，理解上下文逻辑，无论是文字的前后关联，还是画面、视频的核心元素，都能精准把控，是AI创作逻辑通顺、内容准确的核心保障。

#### 扩散模型是视觉创作的引擎

从模糊噪点到高清图，从单张图片到连贯视频，全靠扩散模型一步步打磨优化，由乱到整、由虚到实、由静到动，稳稳撑起AI图片、视频全部视觉创作能力，是当下视觉生成领域最成熟的技术支撑。

#### “温度”与“灵魂”只能由人类赋予

虽然AI无比强大，但终究只是高效的创作工具，绝非真正的创作者。AI可快速产出海量文案、精美图片、流畅视频，让普通人也能轻松完成专业级内容创作，给各行各业带来效率提升。

但AI没有真实情感，没有人生阅历，没有独特审美，没有深度思考与人文温度，它写不出触动人心的真情文字，画不出饱含个人感悟的灵魂意境，做不出兼具思想内涵的优质作品，创作里最珍贵的“温度”与“灵魂”，始终只能由人类赋予。同时，AI创作也面临版权、虚假信息、隐私安全等问题，使用AI创作时，要学会甄别内容真伪，尊重原创版权，不滥用AI制造、传播虚假信息。

未来，AI或许不再是“工具”，而会成为与人类并肩工作的“创意伙伴”，我们面临的新挑战或许是：学会如何与机器一起，讲出那些人类独有的故事。

AI是工具，并非万能「创作者」

## 人类知识库或将耗尽！AI“吃不饱”了？

AI趣问



百科小灵通  
DeepSeek

AI写文案、绘美图、剪视频样样精通，可一个残酷现实正悄然来临：AI赖以生存的人类知识库，正面临枯竭危机。Epoch AI研究预测，2028年AI或将耗尽互联网全部公开文本数据，中国信通院在《人工智能发展报告(2024年)》中也指出，2026年大型语言模型的训练可能耗尽互联网上的可用文本数据。

很多人误以为AI会“学习”

“思考”，其实，AI的所谓学习，只是机械梳理文字、数据的排列规律，它没有真实的感官体验，更无法产生真正的思考和理解。它消耗的，是人类千百年间，带着生活感悟、情感温度、实践经验积累下来的知识，这些根植于人的内容，是算法永远模仿不来的。

人类因经历而感悟，因思考而创新，是缓慢却有温度的成长；

但AI对数据的需求飞速暴涨，远超人类创造知识的速度，这种失衡，本质是冰冷算法与鲜活人类心理的天然差异。

为解决困境，行业开始转向AI生成的“合成数据”。但合成数据缺乏人类真实认知与情感，长期用其训练，必然导致AI逻辑混乱、幻觉增多、能力退化。

更值得警惕的是，过度依赖

AI，会让我们慢慢丢掉主动思考、深度记忆和原创的能力，陷入“懒得动脑”的认知惰性。

这场AI数据危机，其实是在提醒我们：人类的共情力、思考力、创造力，才是知识最珍贵的内核。唯有守住我们独有的思考与表达，不被技术裹挟，才能让AI良性发展，又能守护好人类最宝贵的心理与认知财富。