

# SORA (openAI)

//trouver sites gratuits

Type : Transformers + diffusion

Objectif : Générer une vidéo à partir de texte, image, vidéo.

//faire texte

Publication :

- février 2024 (version beta pour testeurs)
- 9 décembre 2024 (accès public)

Entrée (Input) :

- Texte : prompt décrivant la scène souhaitée
- Image : animée selon une description
- Vidéo : prolongée ou modifiée

Sortie (Output) :

- Vidéo livrée à l'utilisateur via interface web ou API

Principe de fonctionnement :

1. Input
2. Encodage du texte (LLM)
3. Analyse des images/vidéos (Vision)
4. Conversion en espace latent multimodal partagé
5. Génération vidéo via modèle de diffusion
6. Transformers spatio-temporels pour la cohérence entre les frames
7. Post-traitement : super-résolution, fluidité, ajout de filigranes IA
8. Output

Atouts :

- Haute qualité visuelle (jusqu'à 1080p, fluidité, lumière, mouvement)
- Longues vidéos (jusqu'à 1 minute, rare pour un modèle public)
- Compréhension des relations spatiales et physiques
- Multimodal : texte, image, ou vidéo
- Très bon maintien de la cohérence temporelle entre les frames
- Adapté à des usages variés : cinéma, publicité, animation, démo conceptuelle

Limites / Critiques :

- Modèle propriétaire, sans publication scientifique complète
- Accès encore fortement limité et payant
- Pas de bande-son
- Besoin de prompts très détaillés pour résultats précis
- Risque d'usage malveillant (deepfakes, désinformation)
- Temps de génération relativement élevé

### Tests :

Niveau	Prompt	Test	Lien Vidéo	Résultat
Facile	Un chat blanc assis sur un coussin regarde par la fenêtre pendant qu'il pleut	Animation légère, scène fixe		
Moyen	Une femme en manteau jaune traverse un pont en bois au milieu d'une forêt d'automne, avec des feuilles qui tombent lentement autour d'elle.	Mouvement, fluidité, ambiance naturelle		
Difficile	Un homme en costume rouge marche lentement dans une rue de ville sous la pluie, la nuit. Les néons colorés se reflètent sur les flaques d'eau, le vent fait voler sa cravate, et une voiture passe doucement en arrière-plan.	Mouvements complexes, lumière, physique, arrière-plan dynamique		

### Publication Scientifique :

- <https://arxiv.org/pdf/2402.17177> Sora: A Review on Background, Technology, Limitations, and Opportunities of Large Vision Models
- <https://arxiv.org/pdf/2412.00131> Open-Sora Plan: Open-Source Large Video Generation Model

### Référence :

- <https://sora.com/>
- [https://en.wikipedia.org/wiki/Sora\\_\(text-to-video\\_model\)](https://en.wikipedia.org/wiki/Sora_(text-to-video_model))
- [https://fr.wikipedia.org/wiki/Sora\\_\(conversion\\_texte-vid%C3%A9o\)](https://fr.wikipedia.org/wiki/Sora_(conversion_texte-vid%C3%A9o))