

L'Ingénierie de Prompt en Pratique : Comment l'IA a transformé 'Sultans of Swing' en 'The Alley Cats of Swing'

Par Frank , 19 septembre 2026

L'Ingénierie de Prompt en Pratique : Comment l'IA a transformé 'Sultans of Swing' en 'The Alley Cats of Swing'



Par Frank, 19 septembre 2026

Les capacités graphiques de l'intelligence artificielle sont désormais pratiquement illimitées. Comme nous l'avons récemment évoqué dans notre article sur pourquoi l'IA open source et les nouveaux systèmes d'exploitation remplacent des outils traditionnels comme Photoshop, le pouvoir bascule des logiciels coûteux vers des flux de travail intelligents et accessibles. Aujourd'hui, nous plongeons plus profondément dans l'univers de la **transformation vidéo vers vidéo**. Nous analysons une étude de cas ludique mais techniquement impressionnante : la transformation du classique *Dire Straits - Sultans Of Swing* en la parodie animée *Dire Strays - The Alley Cats Of Swing*. Comment un groupe légendaire se métamorphose-

t-il en une troupe de chats de gouttière musiciens sans perdre son ambiance iconique ni son mouvement ? La réponse réside dans l'ingénierie de prompt avancée.

La Transformation : De la Vidéo 1 à la Vidéo 2

Le défi de ce type de projet n'est pas de générer un chat aléatoire, mais de préserver la *cohérence temporelle* (le mouvement au fil du temps) ainsi que la composition originale de la vidéo. L'IA doit comprendre que le guitariste n'est pas simplement un chat quelconque, mais un chat exécutant *exactement les mêmes mouvements de guitare*, sous le même éclairage et dans la même atmosphère de club de jazz des années 70.

- **Vidéo 1 (Entrée) :** [Dire Straits - Sultans Of Swing](#)
- **Vidéo 2 (Sortie) :** [Dire Straits - The Alley Cats Of Swing](#)

Les Outils IA Utilisés dans le Flux de Travail

Une telle transformation résulte rarement d'un simple « bouton magique ». C'est la combinaison d'outils spécialisés. Sur la base de l'état de l'art en 2026, le flux de travail le plus efficace se compose des éléments suivants :

1. Conception de Prompts & Conceptualisation

Avant que les pixels ne soient générés, l'instruction doit être parfaite. C'est ici qu'interviennent les modèles de langage dotés d'une large fenêtre contextuelle, tels que [Moonshot AI \(Kimi\)](#). Ces modèles sont particulièrement adaptés pour structurer des descriptions visuelles complexes et détaillées en prompts IA exploitables, incluant les angles de caméra, l'éclairage et les références stylistiques.

2. Génération d'Images de Référence (Base Image-vers-Vidéo)

Pour définir l'apparence exacte des « chats de gouttière », des générateurs texte-vers-image rapides et itératifs sont indispensables. Des outils comme le [Générateur d'Images IA Perchance](#) offrent un environnement accessible, souvent open-weight, permettant de tester rapidement quels styles (par ex. « bande dessinée vintage », « photo cinématographique » ou « fourrure réaliste ») fonctionnent le mieux avant d'entamer les lourds calculs vidéo.

3. Vidéo-vers-Vidéo & Remplacement de Personnages

Le gros du travail est accompli par des modèles vidéo spécialisés. Les options qui établissent la norme en 2026 pour ce type de transfert de style comprennent :

- **Stable Video Diffusion (SVD) img2vid** : Un puissant modèle open source qui prend une image de référence statique (le chat généré) et l'anime en s'appuyant sur les données de mouvement de la vidéo originale.
- **Inpainting Vidéo Avancé / Échange de Personnages** : Des outils tels que Runway (architecture Gen-3) ou des alternatives open source intégrant ControlNet, qui « verrouillent » les silhouettes et les poses des membres originaux du groupe pour y projeter la nouvelle texture (le chat).

L'Art du Prompt : Exemples Concrets

L'ingénierie de prompt est, comme nous l'avons souligné, l'art de formuler des instructions guidant les systèmes IA vers le résultat souhaité. Un mauvais prompt donne un chat aléatoire ; un bon prompt donne un *musicien* qui se trouve être un chat. Voici les prompts stratifiés nécessaires pour obtenir un tel résultat :

Le Prompt de Base (Style Global)

"Cinematic video, 1970s jazz club atmosphere, smoky dim lighting, warm stage lights. Anthropomorphic alley cats playing musical instruments, realistic fur texture, wearing vintage 1970s tailored suits. Maintaining exact original human poses and movements, high temporal consistency, 4k resolution, photorealistic."

Prompts de Scène Spécifiques (par Membre du Groupe)

Pour obtenir les meilleurs résultats, la vidéo est souvent traitée par segments ou par personnage :

- **Le Guitariste** : "Orange tabby cat playing a black electric guitar, standing upright, dynamic fingerpicking motion, vintage suit, focused expression, cinematic rim lighting, shallow depth of field."
- **Le Bassiste** : "Black and white tuxedo cat playing bass guitar, swaying to the rhythm, dim jazz club background, realistic fur details, professional musician stance."
- **Le Batteur** : "Grey tabby cat behind a vintage drum kit, energetic drumming pose, motion blur on drumsticks, dramatic stage lighting, highly detailed fur and whiskers."

Le 'Negative Prompt' (Ce que l'IA ne doit *pas* faire)

Il est tout aussi crucial d'exclure les éléments indésirables : "cartoon, 3d render, anime, deformed paws, extra limbs, morphing, flickering, loss of

instrument shape, modern clothing, bright daylight."

Flux de Travail Étape par Étape pour le Créateur Indépendant

Comment mettre cela en place vous-même ? La Démocratisation de l'IA signifie que vous n'avez pas besoin d'un budget hollywoodien, mais bien d'une méthodologie :

1. **Extraction d'Images Clés** : Extrayez de la vidéo originale les images clés où la pose est la plus nette.
2. **Prototypage** : Utilisez un outil comme Perchance ou Moonshot AI pour générer le « chat en costume » parfait correspondant à la pose de l'image clé.
3. **Transformation Vidéo** : Chargez la vidéo originale et les images de référence générées dans un modèle vidéo-vers-vidéo (tel que Stable Video Diffusion ou un pipeline open source similaire). Utilisez *ControlNet (OpenPose ou Depth)* pour imposer le mouvement de la vidéo originale à la nouvelle image de chat.
4. **Inpainting & Raffinement** : Utilisez l'inpainting IA pour corriger image par image d'éventuels artefacts (par ex. une main humaine transparaissant à travers la fourrure).
5. **Synchronisation Audio & Finalisation** : Ajoutez la piste audio originale et non modifiée de 'Sultans of Swing' dans un éditeur vidéo traditionnel. L'humour et l'effet naissent précisément du contraste entre la musique familière et sérieuse et la réalité visuelle absurde.

Conclusion et Réflexion

La transformation de 'Sultans of Swing' en 'The Alley Cats of Swing' est bien plus qu'un gadget ludique. C'est la démonstration que **l'Ingénierie de Prompt** et les outils IA open source peuvent assumer les rôles du réalisateur, du directeur artistique et du spécialiste des effets visuels.

Pour les journalistes et créateurs indépendants, cela signifie une libération des écosystèmes fermés et coûteux. Cela nous permet de raconter des histoires visuelles complexes ou, dans ce cas précis, d'ajouter une nouvelle couche critique ou humoristique à des biens culturels existants. La technologie est là ; le défi réside désormais dans la créativité et la précision des instructions que nous lui donnons. Le chat est sorti du sac, et il joue de la guitare.

<https://youtu.be/IBZgk6q6qEg>