



FORUM INNOVATION
INGÉNIERIE | INFORMATIQUE |
ENTREPRENEURIAT | **UQAR**

Illustration de chansons par des
intelligences génératives

par Noah Nicolas

Présentés au FI3E par le Cégep de Matane



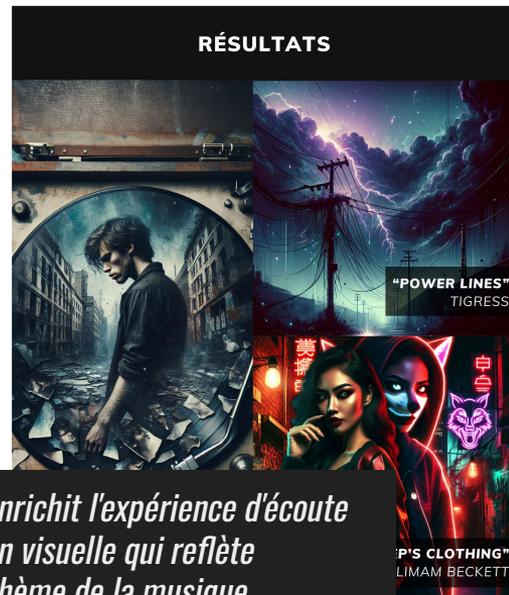
FORUM INNOVATION
INGÉNIERIE | INFORMATIQUE |
ENTREPRENEURIAT | UQAR

Illustration de chansons par des intelligences génératives

par Noah Nicolas

HorizonSP est une application innovante qui exploite les nouveaux services d'intelligence artificielle disponibles via API pour créer des représentations visuelles uniques à partir de chansons. L'application permet aux utilisateurs de rechercher une chanson, puis d'analyser et d'interpréter les données musicales et les paroles à l'aide d'un service de génération de texte. En se basant sur cette analyse, une image est générée pour représenter visuellement la chanson.

Ce projet met en lumière le potentiel des technologies d'IA dans la création de contenu artistique et offre de nouvelles expériences utilisateur. Les limitations actuelles et les perspectives d'amélioration ouvrent la voie à de futures évolutions et adaptations.



Cette approche enrichit l'expérience d'écoute par une dimension visuelle qui reflète l'ambiance et le thème de la musique.

Pourquoi des illustrations de chansons

Avec l'accessibilité croissante des services de génération intelligents automatisés, de nouvelles opportunités d'innovations se présentent. Toutefois, les projets faisant une utilisation innovante de ses nouvelles technologies semblent encore rares. Il reste donc à explorer à identifier les cas d'usages concrets et les combinaisons novatrices de ces technologies.

À partir de cette observation, j'ai formulé la problématique suivante :

Peut-on utiliser les différents services d'IA à notre disposition pour créer des représentations visuelles pertinentes de nos chansons favorites ?

Objectifs

- Combiner les services d'IA disponibles pour générer des représentations visuelles à partir de chansons.
- Explorer les possibilités offertes par les services d'IA disponibles via API pour la génération d'images à partir de données audio.
- Développer une application innovante permettant de créer des représentations visuelles uniques pour des chansons données.
- Évaluer la pertinence et la qualité des images générées par l'application en fonction de la cohérence avec les chansons d'entrée.

Méthodologie

Recherche et sélection des services d'IA et des API pertinents pour le projet.

1. Réalisation de tests manuels sur les différents services de génération et leurs variantes.
 - a. Services d'information musicale : API de Musixmatch & Spotify (versions gratuites)
 - b. Services de génération de texte (LLM) : Le chat (Mistral AI), ChatGPT 3.5 (OpenAI), ChatGPT 4 (OpenAI)
 - c. Services d'intelligence artificielle générative (images) : DALL-E 3 (OpenAI), StableDiffusion (ModelsLab)
2. Développement de scripts de test et d'intégration des services sélectionnés.
3. Conception et implémentation de l'application, incluant des fonctionnalités additionnelles (galerie, favoris, téléchargement des images..)

Choix technologiques

4 APIS différentes utilisées



Spotify

- liste de chansons
- méta-données de chanson



ChatGPT

- génération de prompts
- outil par OpenAI



MusixMatch

- paroles de chansons



Dall-E

- génération d'images
- outil par OpenAI

Choix technologiques

Cette approche permet d'obtenir des résultats plus précis et mieux adaptés.



L'utilisateur recherche et sélectionne une chanson dans la base de données de Spotify.



Les données de la musique et les paroles sont récupérées via les API de Spotify et Musixmatch.



Pour analyser, interpréter et expliquer ces informations, j'utilise un service de génération de texte, tel que ChatGPT-3.5-turbo **en recourant à la technique de chaînage de prompts.**



En se basant sur l'analyse obtenue, une description d'image est générée via DALL-E 3 pour créer une représentation visuelle unique et pertinente de la chanson sélectionnée.

Algorithme & Architecture

Légende



Spotify



OpenAI



MusixMatch



HorizonSP

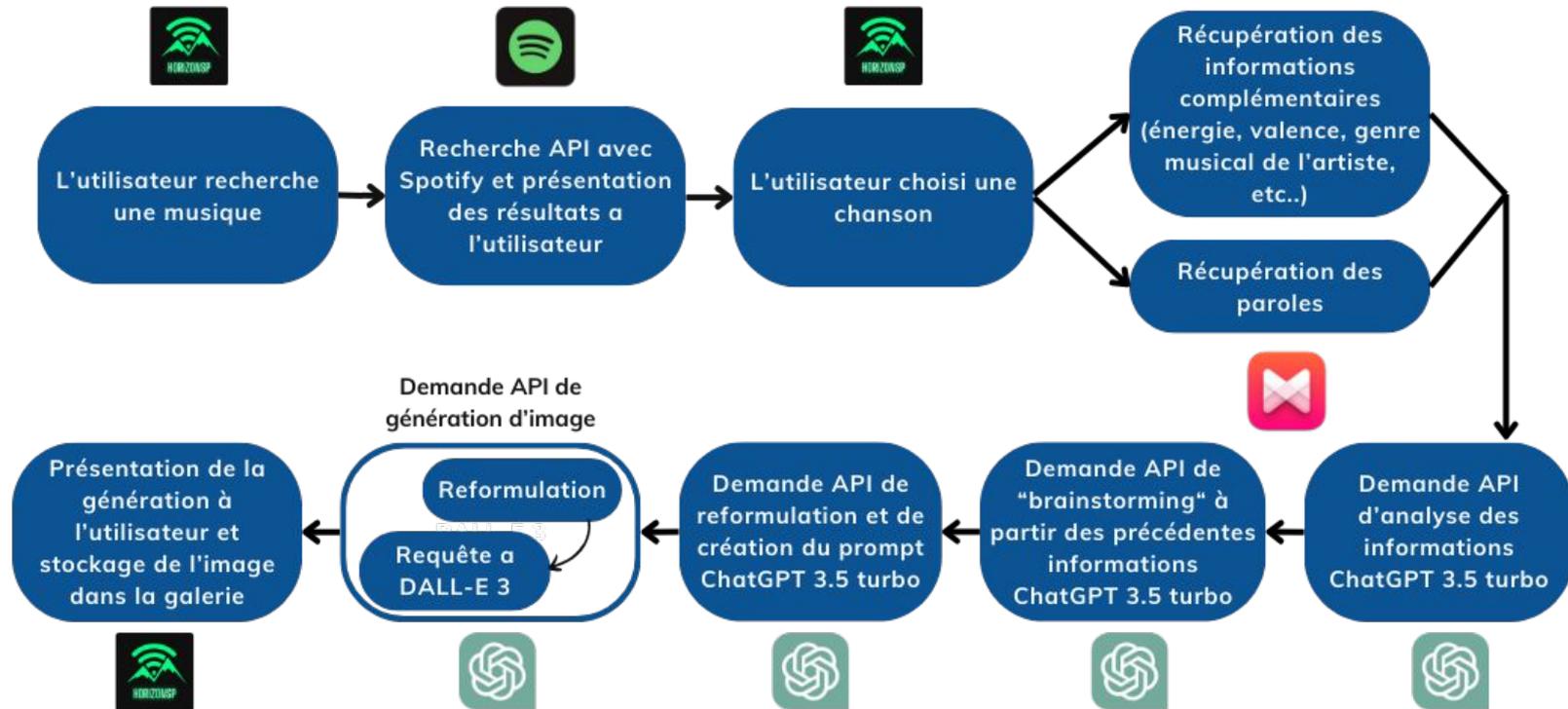
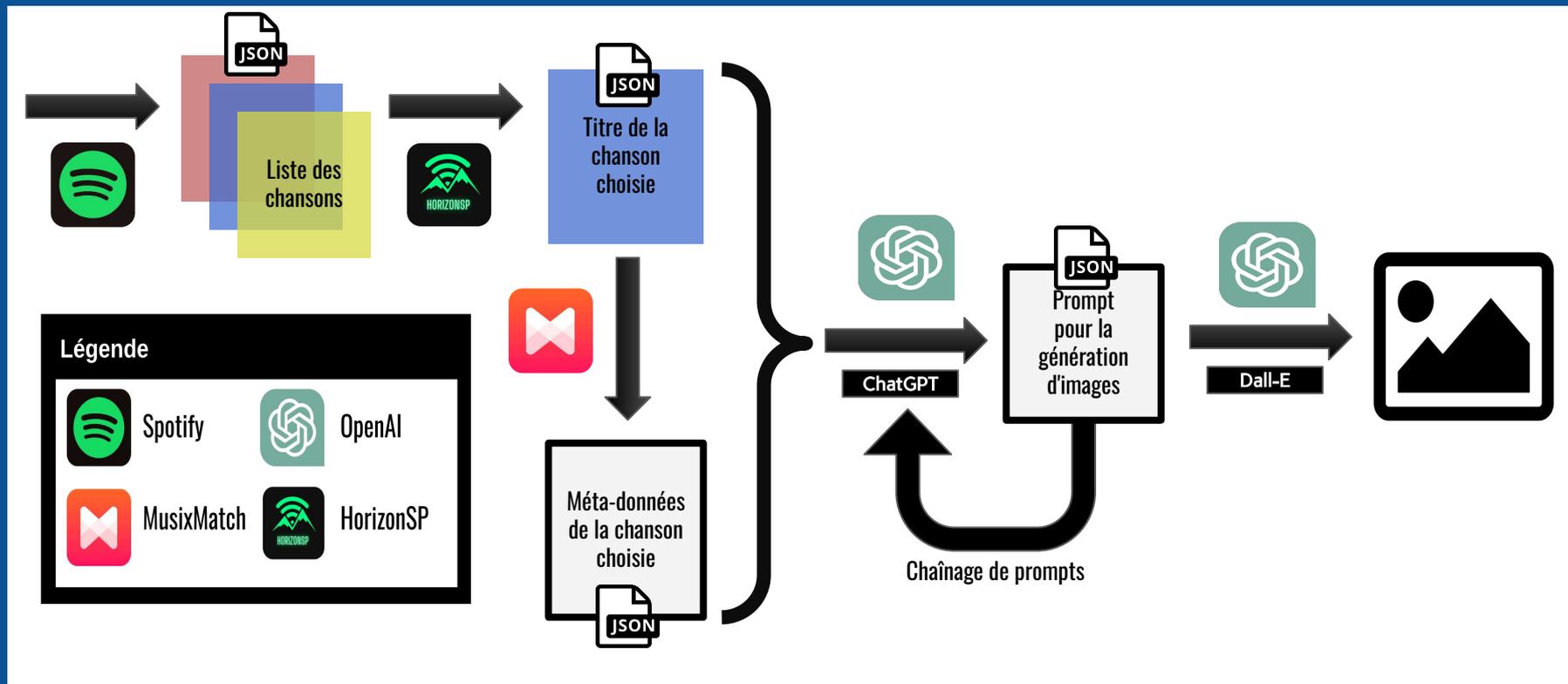


Diagramme de flux de données



Légende



Spotify



OpenAI



MusixMatch

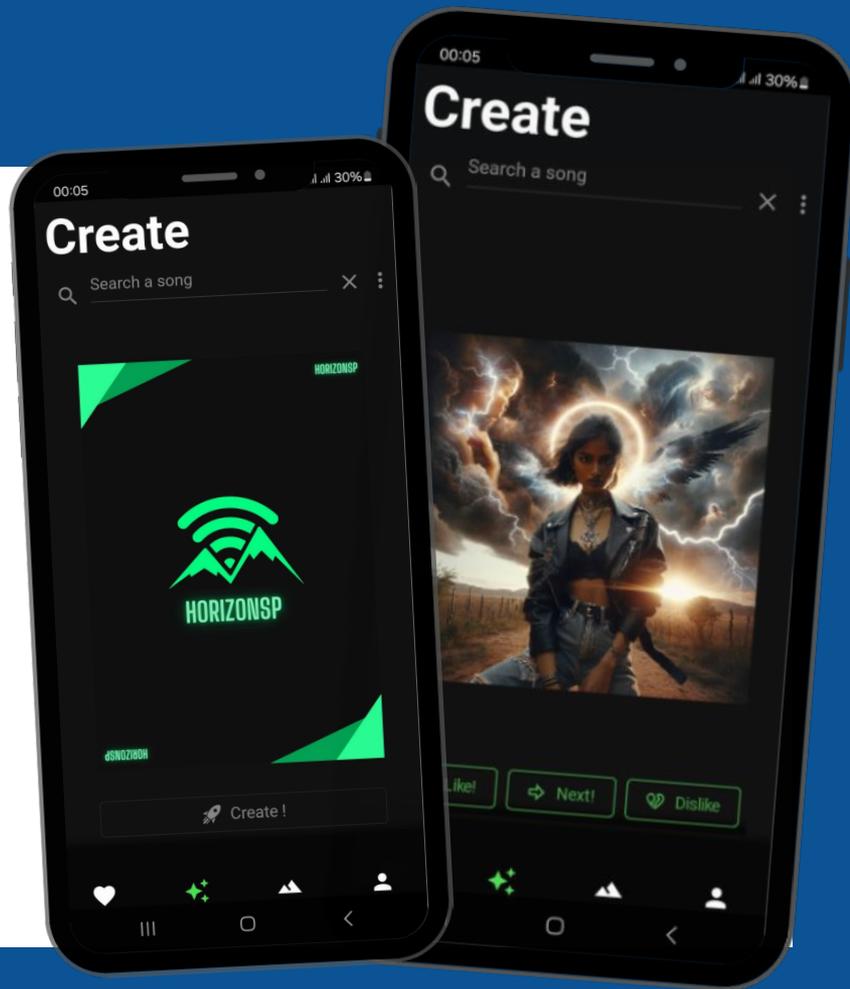


HorizonSP

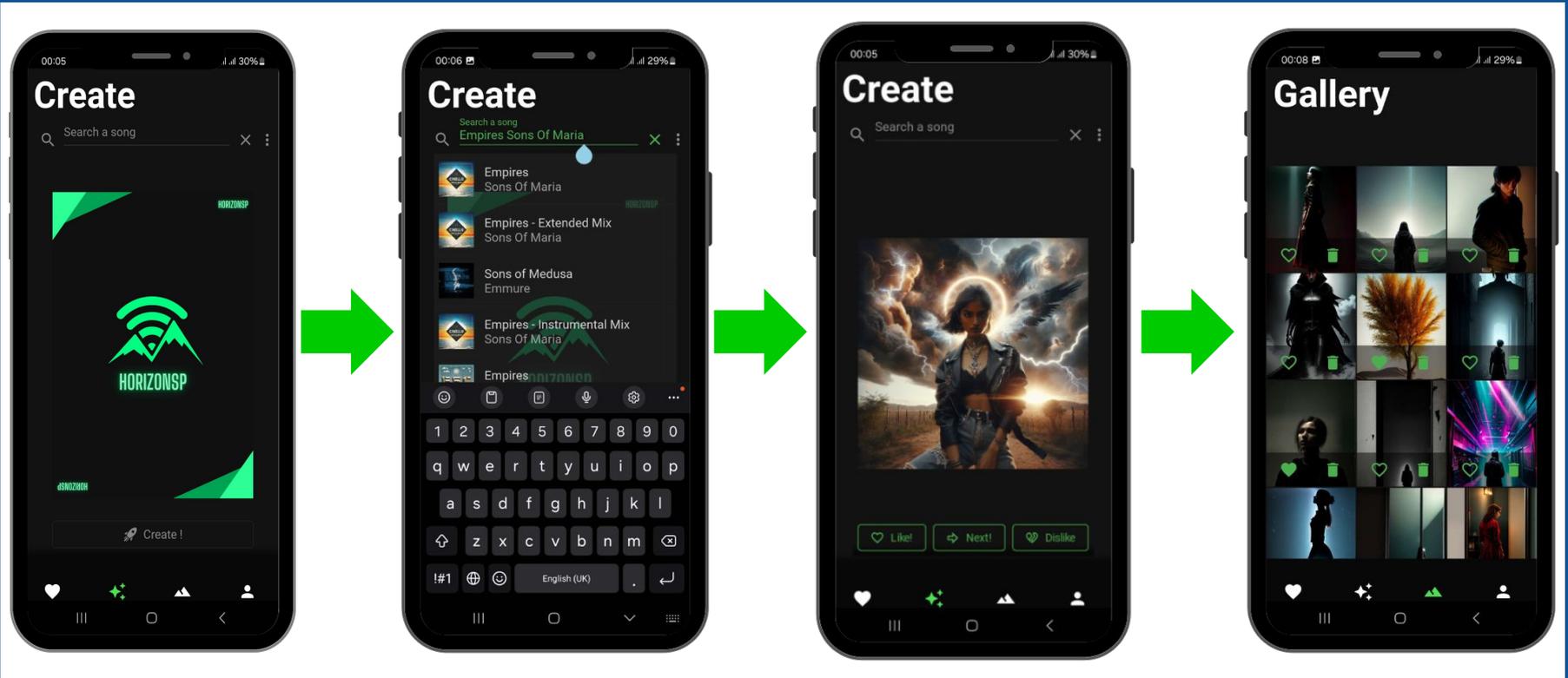
Interface utilisateur

Pour utiliser l'application, rien de plus simple. Il suffit de se rendre sur la page "Create", rechercher la chanson que l'on désire illustrer, puis cliquer sur "Create".

Une fois la génération finie, l'image résultante sera enregistrée dans l'onglet galerie ou l'utilisateur pourra retrouver toutes les créations, les *liker*, supprimer ou enregistrer sur son téléphone.



Storyboard (scénarimage)



Exemple : “We Are” de Jo Cohen et Whales

Générée
par
IA via
l'app

Exemple de prompt final (*revised_prompt* par DALL-E3)

Create a visually striking image embodying the themes of longing, reaching, and search. The color palette favors deep blues, purples, and blacks for a moody and mysterious atmosphere with dim, moody lighting. Use a double exposure technique to blend images of stars and dust particles onto the silhouette of a modern-styled young adult male of Hispanic descent, reaching out. His posture emanates a sense of contemplation and longing. The backdrop should be a surreal landscape with swirling galaxies of cosmic dust, contrasting between light and dark. This image should provoke thought with its emotional depth, representative of reaching for unattainable dreams.



Exemple : “Crown” de Nexeri et RIELL

Générée
par
IA via
l'app

Exemple de prompt final (*revised_prompt* par DALL-E3)

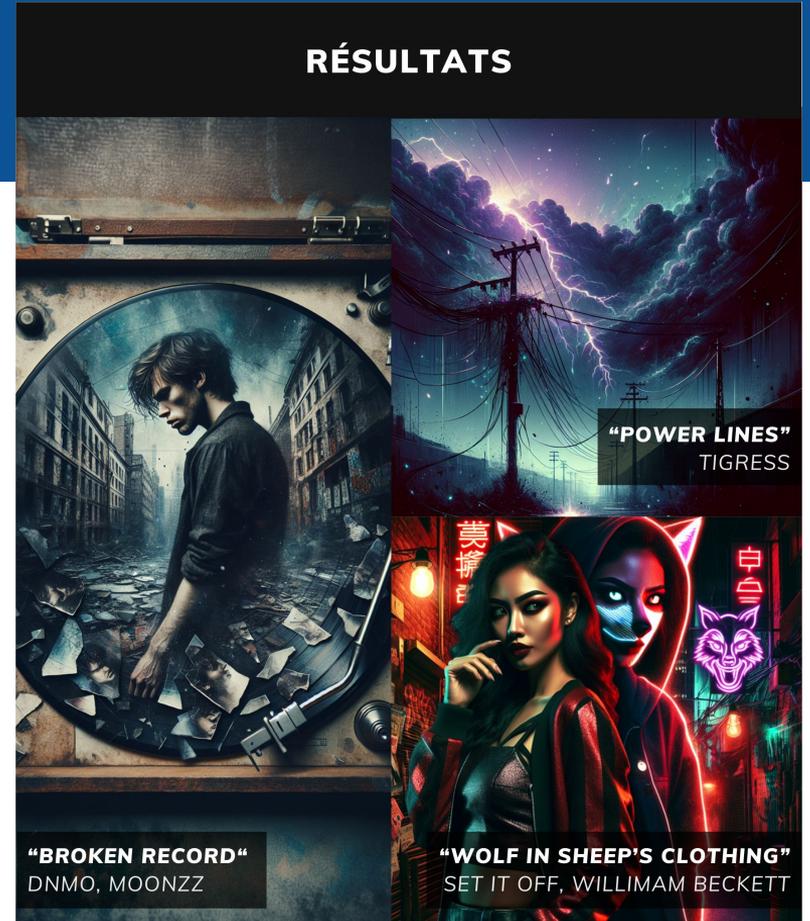
Imagine a powerful image that encapsulates strength, determination, and resilience. There's a majestic lion in the middle, with a flowing mane visible, displaying authority. The lion's eyes show intensity, its posture is that of royalty and its body radiates strength. The deep golden color of its fur reflects sunlight. The details of the lion's mane are intricate. There are rugged mountains in the background and a storm can be discerned in the sky. The lion astonishingly wears a crown, the crown is ornate with intricate designs and sparkling jewels. The lighting dynamically mixes warm and cool tones. The golden sunlight underlines the lion's majestic look and the darkness of the storm clouds intensify the scene, reflecting resilience and defiance themes. The image composition uses the double exposure technique, blending images of the lion and the crown together. The framing is done closely to emphasize the lion's face details and the intricate design of the crown, leading the viewer's attention to the most crucial parts of the image



Résultats

Les résultats illustrent la capacité captivante de l'intelligence artificielle à créer des images génératives qui reflètent l'essence de différentes chansons. Chaque œuvre visuelle, inspirée par une mélodie spécifique, capture les thèmes et les ambiances véhiculés par les titres des chansons.

Par exemple, la représentation de "Broken Record" comprend à la fois littéralement l'enregistrement brisé, mais aussi le sentiment de désolation et de nostalgie, tandis que celle de "Wolf In Sheep's Clothing" illustre parfaitement la dualité du sujet de la chanson entre son paraître et sa réalité.



Limitations et développements futurs

L'application présente actuellement certaines limitations, notamment l'incapacité à générer des images à partir de musiques sans paroles.

- Pour améliorer la couverture des chansons et surmonter cette limitation, j'envisage d'intégrer les services d'autres plateformes, telles que Genius.
Cette amélioration permettra également de contourner la limitation de la version gratuite de Musixmatch, qui ne fournit que 30% des paroles d'une chanson. De plus, la recherche d'un service plus fiable pour les paroles constituerait un gain significatif en qualité pour l'ensemble des images générées.
- Une autre idée intéressante serait d'offrir la possibilité de générer des images sans prendre en compte les paroles, en se basant uniquement sur les données audio.
- Dans le futur, j'aimerais également intégrer davantage de technologies intelligentes génératrices alternatives, telles que le nouveau "Mistral" en tant que modèle de langage, ou StableDiffusion en tant que service de génération d'images. *Cela permettrait d'offrir plus d'options de génération et d'améliorer la qualité des résultats.*
- À titre expérimental, il serait également intéressant d'intégrer une version locale de Mistral afin d'évaluer sa rapidité et sa qualité par rapport aux services API.

Références

Les APIS

OpenAI API reference

<https://platform.openai.com/docs/api-reference>

Mistral AI “Le Chat”

<https://chat.mistral.ai/chat>

ModelsLab (StableDiffusion API)

<https://modelslab.com/>

Spotify for developers - Web API

<https://developer.spotify.com/documentation/web-api>

Build with Lyrics - Musixmatch API

<https://developer.musixmatch.com/>

Les ALGOS

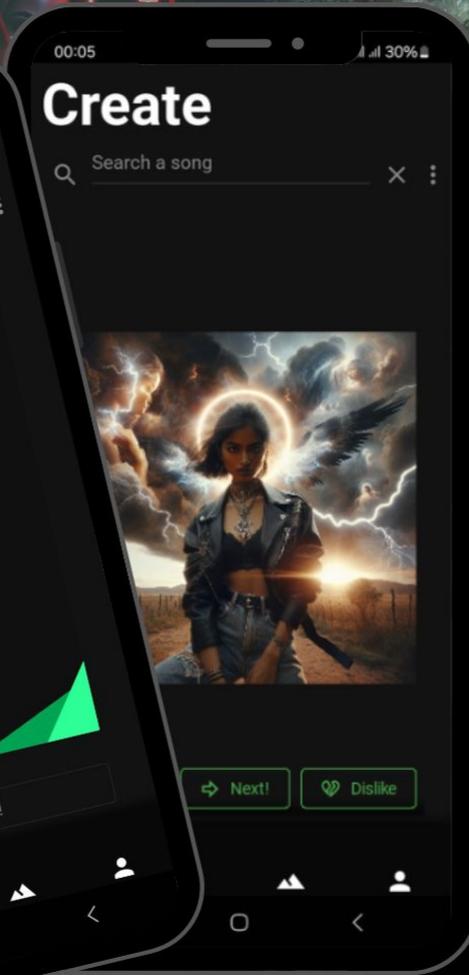
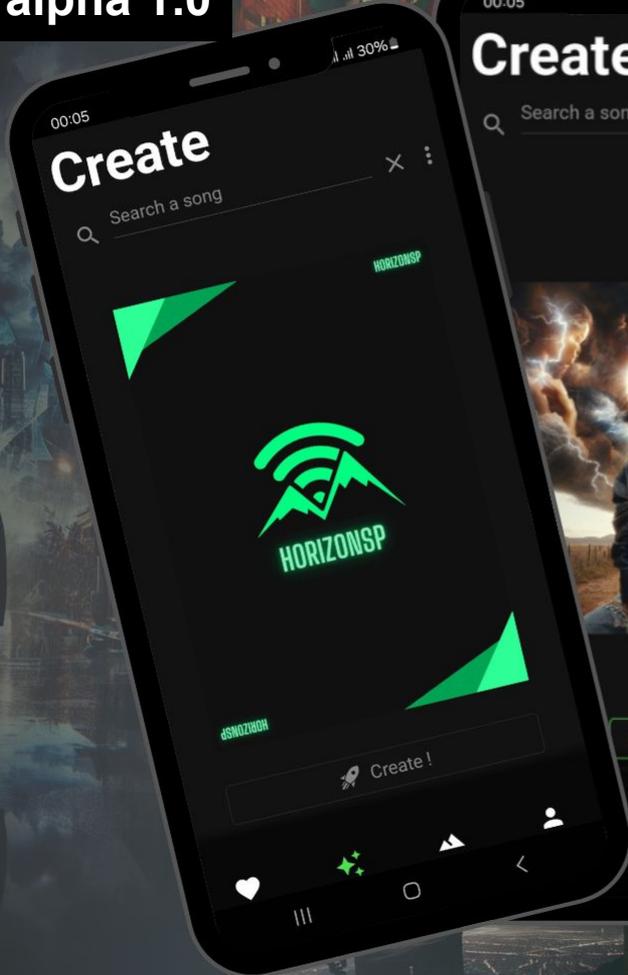
OpenAI API

<https://openai.com/blog/openai-api>

Prompt Engineering Guide - Prompt Chaining

https://www.promptingguide.ai/techniques/prompt_chaining

HorizonSP alpha 1.0



Merci de votre attention !

HorizonSP, bientôt dans les Stores ?