



Synthétiseur de rêves

Présenté par

BEN MERAR Asma
ELMORTADA Hamza
BESSAM Mounia

Encadrant

M Hakim Horray

1 Introduction

Le projet *Synthétiseur de rêves* est une application web innovante qui permet de transcrire un rêve raconté à l'oral, de le transformer en image à l'aide d'outils d'intelligence artificielle, d'en détecter l'émotion dominante (heureux, stressant, neutre, etc.), et d'en garder un historique dans un tableau de bord personnel.

Ce document présente les étapes de conception du projet, le benchmark des technologies utilisées, l'architecture technique, et la maquette interactive réalisée sous Streamlit.

2 Contexte du projet

Ce projet s'inscrit dans le cadre du module de traitement du langage naturel (NLP) à HETIC. Il vise à mettre en œuvre une chaîne complète de traitement de la voix vers l'image et l'analyse émotionnelle, en utilisant des outils modernes, accessibles et conformes aux contraintes du RGPD.

3 Objectifs

- Transcription vocale automatique d'un rêve (speech-to-text).
- Génération d'une image représentant le rêve (text-to-image).
- Détection de l'émotion du rêve.
- Création d'un tableau de bord utilisateur pour l'historique des rêves.
- Déploiement en ligne via Streamlit Cloud.

4 Étude d'art / Benchmark des APIs

Modèle/API	Coût d'usage	Latence (ré- ponse)	Quotas (gratuit/- payant)	Conformité RGPD
Mistral Small La- test <i>mistral.ai</i>	0,1 \$ par 1M input 0,3 \$ par 1M output	Rapide (taille 3Ga)	Gratuit (compte demo li- mité)	Société européenne (France) ; DPA et hébergement possibles en UE.
Mistral Large La- test <i>mistral.ai</i>	2 \$ par 1M input 6 \$ par 1M output	Plus lent (taille 8Ga)	Gratuit (compte demo li- mité)	Comme ci-dessus (UE)
Claude Sonnet 4 <i>anthropic.com</i>	3 \$ par 1M input 15 \$ par 1M output	Rapide (200k to- kens)	Plans gratuits (chat web) ; API sur mesure	Hébergé sur AWS/- Google (européanisé) ; DPA, SOC2, ISO cer- tifiés (GDPR)
Claude Opus 4 <i>anthropic.com</i>	15 \$ par 1M input 75 \$ par 1M output	Rapide (350k to- kens)	Idem Claude (team/ent.)	Idem Claude
GPT-4 Turbo (128k) <i>help.openai.com</i>	0,01 \$ par 1k input 0,03 \$ par 1k output (soit 10/30 \$ par M)	Bas (128k)	Pas de gratuité (API payante uniquement)	Hébergement EU pos- sible (nouveau Ope- nAI), DPA disponible
GPT-4 (8k) <i>help.openai.com</i>	0,03 \$ par 1k input 0,06 \$ par 1k output (soit 30/60 \$ par M)	Bas (8k)	Pas de gratuité (API)	Comme ci-dessus
Gemini 2.5 Pro <i>ai.google.dev</i>	1,25 \$ / 10 \$ par 1M (200k prompts) 2,50 \$ / 15 \$ (>200k)	Rapide (500k to- kens)	Free tier studio (limité) ; Paieement pour usage in- tensif	Google Cloud (régions UE), DPA et stan- dards ISO. Utilisation payante re- quise en Europe
Gemini 2.5 Flash <i>ai.google.dev</i>	0,30 \$ / 2,50 \$ par 1M (texte/image/video)	Très rapide (1M tokens)	Free tier studio dispo- nible	Comme Gemini Pro (UE)

TABLE 1 – Comparatif des modèles Text-to-Text (Modèles de langage)

Service	Coût d'usage	Latence (réponse)	Quotas (gratuit/-payant)	Conformité RGPD
Groq Whisper Large v3 Turbo <i>groq.com</i>	~0,04 \$ par heure audio (0,0007 \$/min)	Très rapide (216× temps réel)	Pas de gratuité annoncée (payant à la seconde)	Fournisseur US ; infrastructure en construction en Norvège (conformité GDPR en cours)
Google Speech-to-Text <i>cloud.google.com</i>	~0,016 \$ par min (standard)	Bonne (temps réel)	60 min gratuits, puis facturation à l'usage	Data centers UE (Belgique, etc.) disponibles ; Data Processing Addendum (DPA) pour GDPR
OpenAI Whisper-1 <i>openai.com</i>	~0,006 \$ par minute	Variable (quelques secondes par phrase)	Pas de gratuité (paiement à la minute)	Option de résidence de données UE ; DPA disponible (GDPR)
Deepgram (Nova-3, Whisper) <i>deepgram.com</i>	~0,004–0,0052 \$ par min (Nova-3)	Très faible latence (<300 ms)	200 \$ de crédits gratuits puis paiement à l'usage	« GDPR-ready » (SOC2, DPA) ; Hébergement global (majoritairement US) ; option d'auto-hébergement entreprise

TABLE 2 – Comparatif des services Voice-to-Text (Transcription)

Service/API	Coût d'usage	Latence (réponse)	Quotas (gratuit/-payant)	Conformité RGPD
ClipDrop API (Stability AI) <i>clipdrop.co</i>	Crédit-based (1 crédit = \$0,01) ; ex. DALL · E 2 : 10 crédits (\$0,10) (prix variable, parfois élevé)	Variable (diffusion stochastique)	20 images/jour gratuits, puis forfaits à crédits	Propriété de Jasper (USA) ; politique de confidentialité AWS/SaaS ; conformité GDPR via SCC uniquement
OpenAI DALL · E 3 <i>openai.com</i>	\$0,01/image (basse) – \$0,04 (moyenne) – \$0,17 (haute)	1–2 secondes	15 essais gratuits via ChatGPT (API payante uniquement)	Hébergement UE possible (via projets API) ; DPA disponible
Midjourney (via ImageAPI) <i>imageapi.dev</i>	Environ \$0,03–0,05/image (ex. 200 images pour \$10/mois)	5–30 sec. selon la file	Pas de gratuité API ; 25 images test via Discord	Serveurs US ; pas de mode EU déclaré ; DPA incertain, stabilité à vérifier
Stable Diffusion XL <i>stability.ai</i>	Crédits : 1 crédit = \$0,01 ; 10–20 crédits/image (\$0,10–\$0,20) selon qualité	1–2 secondes	Très limité gratuitement (quelques essais via HuggingFace)	Société suisse ; auto-hébergement UE possible ; plus conforme au RGPD si hébergé localement
Google Imagen (via Gemini API) <i>ai.google.dev</i>	Gratuit dans le tier Flash (texte → image) sinon payant : \$0,30/1M tokens input	Rapide	Free tier limité ; payant en production	Hébergement Google Cloud (régions UE disponibles), DPA ; pas de collecte personnelle directe

TABLE 3 – Comparatif des services Text-to-Image (Génération d'images)

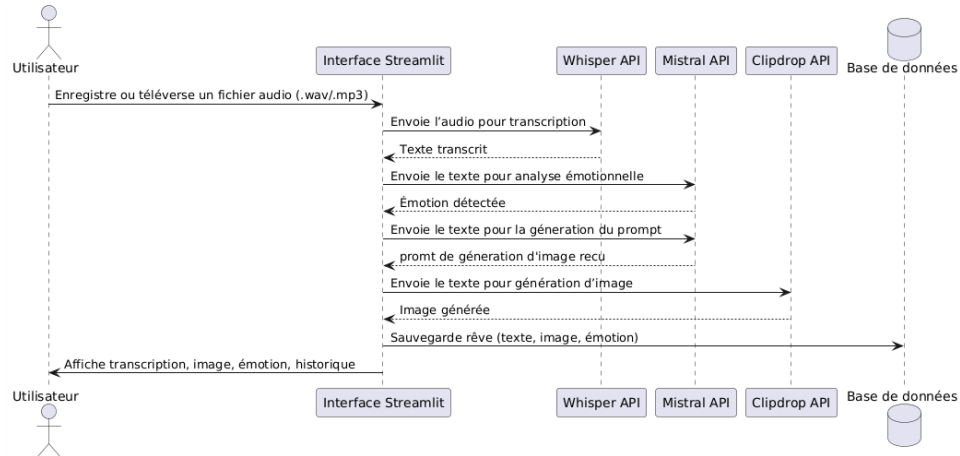
5 Architecture fonctionnelle

5.1 Séquence de traitement (UML simplifié)

1. L'utilisateur enregistre ou téléverse un fichier audio.

2. L'audio est transcrit en texte via Whisper API.
3. Le texte est analysé pour détecter l'émotion principale (via Mistral).
4. Une image est générée à partir du texte via Clipdrop.
5. Tous les éléments sont stockés et accessibles via un tableau de bord.

5.2 Diagramme UML



6 Maquette interactive Streamlit

6.1 Fonctionnalités développées

- Upload ou enregistrement audio en .wav/.mp3
- Affichage de la transcription générée
- Bouton pour générer une image du rêve
- Affichage de l'émotion détectée
- Tableau de bord personnel listant les rêves précédents

6.2 Exemples de rendu



FIGURE 1 – Interface utilisateur Streamlit

7 Conclusion

Ce projet a permis de mobiliser plusieurs compétences autour du NLP, de la génération multimédia, et du développement web. Il propose une expérience originale à la croisée de la technologie et de l'introspection personnelle. Des pistes d'amélioration incluent la génération de vidéo et l'ajout de visualisations statistiques sur les rêves.