

1. [مقدمة عامة \(General Introduction\)](#)

2 [قسم المبرمج الثاني: الـ AI](#)

- 3.1 نظرة عامة
- 3.2 المتطلبات الوظيفية
- 3.3 المتطلبات غير الوظيفية
- 3.4 نماذج الذكاء الاصطناعي وتقنيات التحليل
- 3.5 تدفق العمليات (Workflow)

1. مقدمة عامة (General Introduction)

الغرض

- إنشاء منصة/موقع عام يستطيع المستخدمون من خلاله مشاهدة فيديوهات حصرية ومقالات وقصص، مع إمكانية إضافة ميزة صغيرة لرفع القصص من قبل المستخدمين.
- الاستفادة من الذكاء الاصطناعي في تحويل النصوص إلى فيديوهات وصوتيات متعددة اللغات مع مؤثرات صوتية (360، عاطفية، اهتزازات Haptic).
- وجود نظام اشتراكات (مجاني/مدفوع) واستخدام بوابات دفع متعددة تناسب 12 لغة/دولة.

قسم المبرمج الثاني: الـ AI

3.1 نظرة عامة

مهمة المبرمج الثاني بناء خادم ذكاء اصطناعي (AI Server) يستقبل طلبات من الـ Backend، ويعيد النتائج (تحليل النص، تحويله لصوت أو فيديو، إلخ).

3.2 المتطلبات الوظيفية (Functional Requirements)

1. الكشف عن مصدر النص (بشري أم AI)

- استخدام نموذج NLP مخصص (مثل RoBERTa أو GPT-Detector) لقياس

احتمالية أن النص مولّد بالذكاء الاصطناعي.

- إرجاع نسبة مئوية (مثل "النص بشري 80%").

2. اكتشاف الانتحال (Plagiarism)

- التكامل مع قواعد بيانات خارجية (Turnitin/Copyscape) أو خوارزمية داخلية للمقارنة مع نصوص موجودة.

- إرجاع نسبة التشابه + مصادر التشابه.

- تمييز الخطورة (منخفضة/متوسطة/عالية).

3. تحليل الأخطاء اللغوية والنحوية

- اللغة العربية: أدوات مثل Farasa أو Camel Tools أو أخرى.
- للغات الأخرى: LanguageTool أو Grammarly API أو ما يناسب.
- إرجاع عدد الأخطاء، وتصنيفها (إملائي/نحوي/أسلوبي).

4. تحليل الشخصيات والعمق السردي

- نموذج NLP لاستخراج الكيانات (Named Entity Recognition).
- تقييم بناء القصة (الشخصيات الأساسية، الحبكة، التصعيد).
- درجة "الجادبية" (مثلاً 7/10).

5. تحويل النص إلى صوت (TTS) بميزات متقدمة

- نماذج مثل Tacotron2 أو VITS.
- دعم المؤثرات العاطفية (ضبط النبرة حسب المشاعر).
- دعم الصوت المحيطي 360.
- دمج إشارات اهتزازية (Haptic) في الملف الصوتي (للأجهزة الداعمة).

6. توليد الصور وتحريكها (Stable Diffusion + Pika Labs)

- Stable Diffusion لإنتاج خلفيات أو لوحات للقصة.
- Pika Labs (أو D-ID) لتحريك الشخصيات أو الصور، إن لزم.

7. دمج الفيديو (Video Synthesis)

- استخدام FFmpeg للجمع بين المقاطع الصوتية والصور/المشاهد المتحركة.
- إخراج الفيديو بجودة تصل لـ 4K حسب الاشتراك (مميز).

8. تعدد اللغات والترجمة

- دعم الترجمة التلقائية (MarianMT أو Google Translate API) إذا كانت القصة بلغة غير مدعومة أصلاً في التحليل.
- تحليل النسخة الأصلية والمترجمة معاً (لمزيد من الدقة).

9. واجهة برمجية (AI API)

- استقبال الطلبات:

.

ai/analyze/ لفحص النص (بشري/انتحال/أخطاء).

•
ai/generate-audio/ لتحويل النص لصوت.

•
ai/generate-video/ لتوليد الفيديو من النص + الصور.

• إرجاع النتائج في صيغة JSON أو روابط للملفات الناتجة.

10. التخزين المؤقت (Caching)

• تخزين نتائج التحليل والصوت/الفيديو المنتج مؤقتًا لتسريع الوصول.

3.3 المتطلبات غير الوظيفية (Non-Functional Requirements)

1. الأداء

- تحليل النصوص الخفيفة (> 2000 كلمة) خلال ثوانٍ أو بضع دقائق كحد أقصى.
- توليد فيديو قد يستغرق وقتًا أطول (5-10 دقائق)؛ يجب توفير حالة "Job قيد التنفيذ".

2. القابلية للتوسع

- إمكانية تشغيل خادم الـ AI على GPU عالي الأداء أو استخدام خدمات سحابية (AWS, GCP, NVIDIA GPU, ...).

3. الأمان

- المصادقة (API Key) بين الـ Backend والـ AI.
- عدم قبول طلبات من مصادر مجهولة.

4. قابلية الصيانة

- شيفرة واضحة مع توثيق.
- إمكانية تحديث نماذج الـ AI أو إضافة لغات جديدة بسهولة.

5. الجودة

- الحفاظ على دقة جيدة للتحليل اللغوي وتجنب الأخطاء.
- ضبط المعاملات (Hyperparameters) لنماذج TTS والفيديو للحصول على مخرجات ذات جودة عالية.

3.4 نماذج الذكاء الاصطناعي وتقنيات التحليل

- NLP: استخدام إطار PyTorch أو TensorFlow + مكتبات (HuggingFace Transformers).
- اكتشاف الانتحال: تكامل API أو بناء قاعدة بيانات نصوص.
- الصوت: Tacotron2/VITS/Neural TTS من Microsoft Azure أو Google TTS API (إن).

أردت حلاً جاهزاً).

- **Stable Diffusion**: استخدام نموذج مفتوح المصدر + واجهة InvokeAI أو Automatic111.
- **Pika Labs: API** خارجي لتحريك الصور (أو بديل مثل D-ID).

3.5 تدفق العمليات (Workflow)

1. تحليل نص

- يتلقى خادم الـ AI طلباً من /ai/analyze مع نص + لغة.
- يحدد اللغة (إن لم تكن محددة) → يجري الترجمة إن لزم → يكشف الانتحال → يحسب نسبة الأخطاء → يرسل تقرير JSON.

2. تحويل نص إلى فيديو

- طلب /ai/generate-video:

•

الخادم يحلل النص (شخصيات، مشاهد) → يطلب من Stable Diffusion صوراً تعبيرية → يحركها (Pika Labs) → يجمع مسار الصوت (TTS) مع المسار البصري (FFmpeg).

•

يخزن الفيديو النهائي في مسار محدد → يعيد رابط الفيديو للـ Backend.

3. متابعة حالة المهام

- الخادم يحفظ مهمة التحويل في جدول (أو Cache) بحالة processing.
- عند الانتهاء، يحدّث الحالة إلى completed، ويخطر الـ Backend.