



DOCUMENTO DE PRINCIPIOS ÉTICOS Y LEGALES

Mikel Martinez, Ibai García, Haimar Álvarez e Iker Iglesias
09/12/2025

1. RESUMEN EJECUTIVO: Marco Ético y Cumplimiento Legal

El desarrollo de IA Video Analyst Pro se ha regido por un enfoque de privacidad por diseño y responsabilidad técnica. En materia de seguridad, se ha priorizado la protección de datos del usuario, manteniendo las credenciales de sesión (cookies) almacenadas localmente y gestionando las claves de API mediante entornos aislados, evitando cualquier exposición en repositorios públicos.

A nivel de licenciamiento, el software se integra exclusivamente con herramientas de código abierto (*Open Source*) bajo licencias permisivas (Apache 2.0, LGPL), asegurando la viabilidad técnica del desarrollo. Sin embargo, se identifica la Propiedad Intelectual como el principal desafío legal: aunque la herramienta es funcional tecnológicamente, su uso contraviene los Términos de Servicio de YouTube (descarga no oficial) y opera en una zona gris respecto a los derechos de emisión deportiva. Por tanto, el proyecto se define estrictamente bajo un alcance educativo y de uso privado, excluyendo cualquier explotación comercial.

Finalmente, desde una perspectiva ética, se han implementado salvaguardas (*Safety Rails*) para mitigar las "alucinaciones" inherentes a la IA Generativa y filtros de contenido para evitar el procesamiento de material inapropiado, complementado con un sistema de métricas de consumo que promueve la transparencia y la eficiencia energética (*Green AI*).

DOCUMENTO DE PRINCIPIOS ÉTICOS Y LEGALES.....	0
1.1. Privacidad de Datos y Seguridad.....	3
1.2. Licencias de Herramientas Utilizadas.....	3
1.3. Licencia de Uso del Modelo.....	4
1.4. Consideraciones Legales: Derechos de Autor (Copyright).....	4
1.5. Protección frente a Errores (Safety Rails).....	5
1.6. Métricas de Optimización.....	5

1.1. Privacidad de Datos y Seguridad

La arquitectura del sistema maneja información sensible que requiere protocolos estrictos:

- **Cookies de Usuario:** El funcionamiento del descargador depende de un archivo `cookies.txt` que contiene la sesión de YouTube activa del usuario. Este archivo otorga acceso a la cuenta de Google del usuario. Por diseño de seguridad, este archivo **reside localmente** en la máquina del usuario y nunca se envía a una base de datos externa ni a la nube del desarrollador. Solo se utiliza en el momento de la comunicación directa con los servidores de YouTube.



`cookies.txt`

- **Datos en la API:** Los videos procesados se envían a través de la API de OpenRouter hacia los servidores de Google (proveedor del modelo Gemini). Aunque las políticas de privacidad de nivel empresarial de estos proveedores suelen asegurar que los datos enviados vía API no se utilizan para el entrenamiento de futuros modelos (a diferencia de las versiones gratuitas de chat web), es imperativo informar al usuario de que el contenido audiovisual está siendo procesado por terceros.
- **Gestión de Credenciales (API Keys):** Se ha implementado el uso estricto de archivos `.env` excluidos del control de versiones (`.gitignore`). Esto previene que las claves de acceso, que conllevan un coste económico, se filtren accidentalmente en repositorios públicos como GitHub

```
import os
from dotenv import load_dotenv

# Cargar variables
load_dotenv()

# --- CLAVES ---
API_KEY = os.getenv("API_KEY_OPENROUTER")
```

1.2. Licencias de Herramientas Utilizadas

El proyecto se erige sobre un ecosistema de software de código abierto (*Open Source*), respetando sus licencias:

- **Streamlit:** Bajo licencia **Apache 2.0**, lo que permite su uso gratuito, modificación y distribución, incluso con fines comerciales.
- **yt-dlp:** Bajo licencia **Unlicense**, lo que efectivamente lo coloca en el dominio público, permitiendo cualquier tipo de uso sin restricciones legales de copyright.
- **FFmpeg:** Generalmente bajo licencia **LGPL v2.1+**. Esta licencia permite el uso de la librería como un componente dinámico (vinculación dinámica o llamada por línea de comandos, como se hace en este proyecto) dentro de software propietario, siempre y cuando no se modifique el código fuente propio de FFmpeg.
- **OpenCV:** Bajo licencia **Apache 2.0**, compatible con uso comercial y académico.

1.3. Licencia de Uso del Modelo

El uso del modelo de IA no es libre, sino que está regulado por términos específicos:

- **Google Gemini:** Su uso vía API está sujeto a los *Generative AI Terms of Service* de Google. Estos términos prohíben explícitamente el uso de la tecnología para generar contenido que viole derechos de propiedad intelectual, desinformación, discurso de odio o contenido sexualmente explícito.
- **OpenRouter:** Actúa como pasarela intermediaria. El uso de su plataforma implica la aceptación de sus términos, que incluyen políticas de uso justo (no saturar el servicio mediante ataques DDoS o bucles infinitos) y el respeto a las políticas de contenido de los modelos subyacentes.

1.4. Consideraciones Legales: Derechos de Autor (Copyright)

Este apartado constituye el riesgo legal más significativo del proyecto y debe ser tratado con transparencia:

- **Descarga de YouTube:** Los Términos de Servicio (ToS) de YouTube prohíben explícitamente la descarga de contenido audiovisual de su plataforma salvo que exista un botón de descarga oficial o se tenga permiso por escrito del titular. El uso de herramientas como **yt-dlp**, aunque técnicamente posible, constituye una violación de estos términos contractuales privados entre el usuario y la plataforma.
- **Propiedad Intelectual de Imágenes Deportivas:** Los derechos de emisión y explotación de las imágenes de partidos (LaLiga, Premier League, MLS, etc.) son propiedad exclusiva de las ligas y las cadenas de televisión que han adquirido las licencias.
 - **Uso Personal/Educativo:** El uso de esta herramienta en un entorno local para análisis personal, investigación o docencia podría ampararse, dependiendo de la legislación local, en excepciones como la "copia privada" (España) o el "Fair Use" (EE.UU.).

- **Uso Comercial/Redistribución:** La comercialización de esta herramienta o, más críticamente, la redistribución pública de los clips generados ("highlights") sin licencia, constituiría una infracción directa de copyright y derechos de emisión, sujeta a acciones legales por parte de los tenedores de derechos.

1.5. Protección frente a Errores (Safety Rails)

Dado que los modelos generativos son probabilísticos y no deterministas, se han implementado medidas de seguridad:

- **Mitigación de Alucinaciones:** Los Grandes Modelos de Lenguaje (LLM) pueden inventar hechos con alta confianza. Se ha mitigado este riesgo mediante la técnica de "Grounding" (anclaje), proporcionando el contexto visual directo (el video) y forzando una temperatura baja (0.1). Además, el *System Prompt* incluye instrucciones explícitas para que el modelo admita cuando no sabe una respuesta en lugar de inventarla. No obstante, el usuario debe ser advertido de que la IA puede cometer errores de interpretación en jugadas subjetivas.
- **Filtros de Contenido Inapropiado:** Los modelos de Google integrados vía API poseen filtros de seguridad (*Safety Settings*) activados por defecto. Estos filtros bloquearán automáticamente el análisis y la respuesta si detectan que el video de entrada contiene violencia gráfica extrema (gore), incitación al odio o contenido sexual explícito, protegiendo al usuario y al proveedor del servicio.

1.6. Métricas de Optimización

Desde una perspectiva de ética en la ingeniería de software y sostenibilidad (*Green AI*):

Se implementó un sistema de monitorización de **Tokens** (Input y Output) visible para el usuario. Esto cumple una doble función:

1. **Transparencia Económica:** Permite calcular el coste operativo exacto por partido analizado.
2. **Eficiencia Computacional:** Concientiza sobre el consumo de recursos. El procesamiento de video consume mucha energía en los centros de datos; optimizar los prompts y la resolución del video no solo ahorra dinero, sino que reduce la huella de carbono asociada al uso innecesario de cómputo en la nube.