**Integrante 4 - Nicolás Di Domenico:**

##### APIs de imagen y video.

* + [Vision AI](https://cloud.google.com/vision) (Cloud Vision)
  + [Video AI](https://cloud.google.com/video-intelligence)

##### APIs de documentos y datos.

* + [Document AI](https://cloud.google.com/document-ai)

**Que desarrollar en cada servicio escogido:**

* Como se llama.
* Cómo funciona.
* Para qué sirve.
* Ventajas.
* Desventajas.
* Beneficios vs costos (opcional).

**Google Cloud ofrece** APIs de imagen y video como **Vision AI** y **Video AI**. Y tambíen APIs de documentos y datos como **Document AI** que están basados en inteligencia artificial y machine learning.

<De ahora en mas voy a decirle a cada uno vision video y document omitiendo el AI asi no me trabo>

Comenzando con…

**Vision**

También se lo conoce como Cloud Vision API, es parte de los servicios de IA que ofrece Google Cloud y se enfoca en la visión artificial para el análisis y procesamiento de imágenes, videos y documentos.

### **¿Cómo funciona?**

bueno Cómo mencioné **Vision** se basa en el concepto de visión artificial, que es lo que le permite a las computadoras "ver" y entender imágenes o videos, similar a lo que hacemos los humanos, pero con la diferencia que nosotros para justamente "ver" y “entender” usamos los ojos y el cerebro, mientras que las computadoras necesitan sensores y modelos matemáticos para analizar los datos de manera visual para así convertirlos en números que representan colores y píxeles, y que luego mediante algoritmos avanzados de machine learning encontrar patrones entre las miles de imágenes, que bueno en nuestro caso google, le va a ir pasando para que después, a través de una **API**, los desarrolladores puedan hacer uso de esta tecnología.

### **¿Para qué sirve? -> Acá tenemos ejemplos más concretos como la…**

* **Búsqueda de productos basados en imágenes para comercios:** Imaginen que tienen una foto de la remera o pantalón que les gusta ya sea porque la sacaste de internet o porque le sacaron una foto en algún local, entonces a través de la Product Search API de **vision**, solo necesitas subir esa imagen. Ya que los comercios tienen fotos de sus productos almacenados, entonces cuando la subis, **Vision** analiza y la compara con el catálogo del comercio. Haciendo que te muestra una lista de productos que son visualmente similares o que justamente es el que estaba buscando.

### **Análisis y generación de etiquetas:** o sea **vision** se va a basar en modelos ya entrenados para revisar cada imagen que los usuarios suben, y luego generar anotaciones o etiquetas (como "perro", "árbol", “Feliz”, “Enojado”, etc.). También puede ayudar a filtrar contenido inapropiado, identificar colores y emociones, entre varias otras funciones, acá dejo un qr para que vean una demo sobre esto.

* **Conseguir descripciones de imágenes automatizadas con la IA generativa**: por ejemplo si tengo un catálogo de productos para mi negocio y necesito una descripción rápida de cada uno. **Vision** puede ver la imagen del catálogo y describir cada uno de los productos que quiero vender no.
* **Inspección visual de alta precisión con Visual Inspection: Vision** tiene una herramienta que puede revisar productos en una línea de producción para detectar defectos automáticamente, como si estuviera haciendo una inspección de calidad. Entonces se entrena al modelo para asi reconocer lo que se considera "defectuoso" y luego se encarga de encontrar esos defectos en tiempo real.

### **Ventajas**

* **Automatización de Procesos Complejos**: **Vision** permite automatizar tareas que tradicionalmente requerirían intervención humana y así ahorrar tiempo y recursos. (como el ejemplo que les mencioné de la inspección de calidad).
* Esto nos lleva a la proxima ventaja que es la de **Reducción de Errores Humanos**, ya que las tareas repetitivas están sujetas a errores que pueden afectar a la calidad del producto o servicio.
* **Facilidad de Integración con API**: que basicamente cualquier desarrollador pueda integrar las funciones de **Vision** fácilmente y sin tener grandes conocimientos sobre como funciona la IA.
* **Escalabilidad y Flexibilidad**: Al funcionar en la nube, permite escalar verticalmente sin necesidad de gestionar infraestructura local y además, se puede adaptar a diversas necesidades, desde la búsqueda de productos hasta la moderación de contenido.
* **Está basada en modelos pre entrenados**: o sea al utilizar estos tipos de modelos de Google, se reduce el costo y el tiempo necesario para crear soluciones desde cero. Igualmente recordemos que si se necesita desarrollar modelos más concretos que resuelva necesidades mas específicas de negocio lo podemos hacer a través de **AutoML** que ya bien explicó el compañero Estefano.

### **Desventajas**

* **Limitaciones en Personalización**: Como ya mencioné en las ventajas, Visual proporciona modelos pre-entrenados, eso puede llevarnos a una desventaja que es justamente que al estar adecuado para tipos especificos de dominios de negocio vamos a tener que recurrir a otros modelos que a través de autoML le enseñemos que tipos de respuesta queremos según nuestras necesidades particulares.
* **Privacidad y Seguridad de los Datos**: Enviar datos visuales a la nube para ser procesados puede representar un riesgo para empresas que manejen información sensible o por lo menos asi lo puede llegar a sentir el usuario que haga uso de este servicio.

**~~Ejemplo de uso de Vision AI en “Try the API”:~~**[~~https://cloud.google.com/vision?hl=es#how-it-works~~](https://cloud.google.com/vision?hl=es#how-it-works)

**Por otro lado…**

**Video AI**

…va a realizar un análisis preciso de videos, reconociendo una gran cantidad de objetos, lugares y acciones. ¿Pero como se diferencia de **vision** si dije que también analiza videos?, bueno….**Vision** se centra principalmente en el análisis de imágenes estáticas (analizando imágenes individuales extraídas de un video), **Video AI** está diseñado para trabajar con los videos de manera completa y fluida por asi decirlo, analizando las escenas, o identificando objetos en movimientos como acciones de las persona, y esto lo hace desde el inicio hasta el fin del flujo del video, tanto para los que están en el almacenamiento local o en la nube o hasta para las transmisiones en vivo.

Y bueno…con respecto **al funcionamiento** de **Video** es similar al de **Vision** y se realiza a través de la **API de Video Intelligence**, que emplea modelos de **machine learning** previamente entrenados, entonces le enviamos un video a esta api y nos devuelve **metadatos detallados** sobre el contenido del video como por ejemplo etiquetas de objetos como auto, persona, perro, con una marca de tiempo que indica cuándo aparece un objeto. entre otras funciones mucho más potentes que ahora vamos a ver.

### **¿Para qué sirve?**

###### **Anuncios contextuales:** en donde **Video** va a analizar el contexto del video para colocar una publicidad que tenga relación con lo que el consumidor está viendo, em se me ocurre una escena de carreras de auto en donde se inserte la promoción de determinado vehículo no, o escenas de modelaje en donde se coloque ofertas de vestidos por poner algunos ejemplos, esto igualmente también aplica para sitios web donde a veces los anuncios aparecen en ubicaciones aleatorias, molestando al lector, entonces con esta IA esto se puede corregir y asi colocar la publicidad en lugares que el usuario haga foco pero sin hacer que se sienta invasiva.

* **Streaming asistido por IA**: haciendo que se pueda analizar lo que está ocurriendo en el video a medida que se va reproduciendo, como "¿Cuántas veces apareció el auto rojo? o ¿Cuántos comentarios en el chat hubo en tal intervalo de tiempo?" no.
* Bueno además en la página de Google cloud mencionaba también 2 ejemplos concretos como:
* detectar pozos en las calles de la ciudad en tiempo real haciendo que cuando va pasando un camión de reparación, este detecta las imperfecciones y rellena a partir de este servicio de Google Cloud.
* Había también un caso de un padre que grababa todos los sucesos o eventos familiares y los subia al cloud storage, entonces en un momento a la hija a través de esta API se le ocurrió crear un álbum familiar que mostraba imágenes o secciones de video a partir de las grabaciones del padre, con la posibilidad de tener una barra de búsqueda para determinar como como se quería mostrar todo eso, por ejemplo se filtraba cuando habia un cumpleaños, o fotos de nacimiento de la hija, o cuando la mujer está embarazada, etc.

las ventajas y desventajas van a ser muy similares a la de **vision**, pero orientado al procesamiento de información de los videos.

**Como Ventajas vamos a seguir teniendo :**

la Flexibilidad, la escalabilidad, la facilidad de integración, Reducción de Errores Humanos, Automatización de Procesos Complejos, al estar basada en modelos pre entrenados se ahorra tiempo y costos, y bueno entre otras muchas más ventajas.

### y como…

### **Desventajas:**

los **problemas de Privacidad** con los datos sensibles, la limitación en la personalización…eh… **Limitaciones en el streaming**: Ya que la latencia puede ser un problema si se necesita una respuesta inmediata, más cuando se analiza contenido crítico como cámaras de seguridad.

(acá se puede decir algo de autoML)

**bueno y el último que voy a comentar es el de**

**Document AI**

**Document** está diseñado para automatizar la extracción de datos de los archivos para asi luego comprenderlos. eh ya sean estructurados o no estructurados. Si bien **Vision** también puede leer documentos extrayendolos como una secuencia de imágenes y detectando etiquetas en el texto, **Document** busca obtener un entendimiento completo logrando analizar, clasificar y estructurar la información extraída.

**¿Cómo funciona?**

Primero, procesa el documento para mejorar la calidad y asi eliminar el ruido por así decirlo y luego la IA clasifica e identifica las partes más importantes. Utiliza OCR para convertir el texto en datos digitales, que después se validan para asegurar que todo se hizo bien no.

Además, si se necesitan datos específicos, se puede personalizar, como por ejemplo, una clínica puede usar **Document** para leer los formularios de un paciente y extraer el nombre y la fecha de una consulta. Y si la clínica necesita tareas más específicas, como armar el historial médico o determinar cuales son los medicamentos actuales, va a poder personalizar el modelo justamente para capturar esa información.

**¿Para qué sirve?**

* Una de las funciones básicas que nos ofrece **Document** es hacer un resumen de todos los archivos o información no estructurada que tengamos en el Cloud Storage de Google, y guardar ese resumen en una base de datos para que nosotros podamos consultarle por cada uno de los documentos que necesitemos una síntesis para ahorrar tiempo por ejemplo. (mencionar lo de la practicidad a diferencia de chatGPT)
* Y bueno entre otras funciones como, clasificar y analizar formularios, digitalizar documentos antiguos, extraer información clave de documentos escritos a mano, generar gráficos estadísticos de los datos analizados, y muchas otras opciones que se encuentran en el Cloud Storage de Google

**Acá también dejo un QR para ver una demo del uso de esta herramienta por si lo quieren escanear.**

y bueno como…

**Ventajas** tenemos la rapidez en la obtención de valores claves, la alta precisión en los datos extraídos, eh integración con otras herramientas de Google Cloud y la capacidad de tomar decisiones mejor informadas al estructurar la información.

**Desventajas**: Bueno por el otro lado, existen preocupaciones sobre la privacidad al procesar los datos sensibles en la nube, limitaciones en la personalización de los modelos preentrenados para necesidades específicas, y la dependencia de una conexión a internet estable, que básicamente son las mismas desventajas que los otros dos servicios.

**~~Ejemplo de uso de Document AI en “Try the API”:~~**[~~https://cloud.google.com/document-ai?hl=es\_419~~](https://cloud.google.com/document-ai?hl=es_419)

~~(opcional)~~