

Propuesta Técnica.

AOSENUMA - App Movil :: IA

CDMX a 7 de agosto de 2023



www.dsindigo.com

Antecedentes del Proyecto.

En un mundo cada vez más interconectado, las conversaciones humanas juegan un papel fundamental en la comunicación y toma de decisiones en diversos ámbitos, desde entornos corporativos hasta contextos personales. Bajo esta perspectiva, presentamos una propuesta técnica innovadora para abordar una necesidad imperante en el análisis y gestión de conversaciones: el desarrollo de una aplicación móvil avanzada y versátil (IA for AOSENUMA APP).

Esta propuesta busca crear una solución integral que permita el análisis exhaustivo de audios, textos e imágenes provenientes de conversaciones humanas, abarcando un periodo histórico de hasta tres meses. El objetivo principal es brindar a los usuarios la capacidad de extraer información valiosa y relevante de estas interacciones, a través de una serie de funcionalidades clave.

La aplicación se posiciona como una herramienta indispensable para diversas industrias y contextos, desde empresas que buscan optimizar sus procesos de comunicación interna y externa, hasta organizaciones que desean supervisar y mitigar riesgos en sus interacciones. Además, esta plataforma no solo se centra en el análisis de contenido, sino que también potencia la colaboración y la toma de decisiones informadas.

A lo largo de esta propuesta, explicaremos en detalle las características y capacidades que conformarán esta aplicación revolucionaria. Desde la identificación precisa de los participantes en las conversaciones, hasta la generación automatizada de resúmenes y la consolidación de los temas discutidos en las reuniones, nuestra solución se presenta como una respuesta integral a las demandas actuales de gestión conversacional.

Además, uno de los aspectos diferenciadores clave de esta aplicación radica en su capacidad de ofrecer traducciones en tres de los idiomas más hablados a nivel global: español, inglés y francés. Esto garantiza la accesibilidad y utilidad de la aplicación en contextos internacionales, facilitando la colaboración entre personas de diferentes regiones y culturas.

En resumen, la propuesta que presentamos no solo aborda una necesidad apremiante en la gestión de conversaciones humanas, sino que también redefine la forma en que se comprende y utiliza la información derivada de estas interacciones. Nuestra aplicación móvil tiene el potencial de transformar la manera en que las empresas, equipos y personas gestionan, analizan y aprovechan sus conversaciones, proporcionando una ventaja competitiva invaluable en un mundo cada vez más orientado hacia la comunicación efectiva y la toma de decisiones informadas.

Requerimientos Generales.

El cliente nos pide la creación de un MVP que contemple las siguientes funcionalidades:

1.- Análisis de Audios: (Análisis de Audios en un Periodo máximo de 3 Meses)

El análisis de audios constituye una parte fundamental de la propuesta técnica, permitiendo a los usuarios extraer información clave y patrones relevantes de las conversaciones. Hemos pensado un enfoque sólido que garantiza la efectividad del análisis, incluso en ausencia de ejemplos concretos. *Es importante hacer notar que la propuesta de los 3 meses surge debido a la falta de información para analizar y consideramos un riesgo el ofrecer un análisis de mayor alcance, y nuestro compromiso siempre será la honestidad con el cliente y explicar el porqué de la sugerencia.*

1. 1.- Procesamiento del Audio:

La aplicación se encargará de recibir los archivos de audio y/o grabarlos directamente desde el dispositivo y los convertirá en un formato digital procesable. Utilizaremos técnicas avanzadas de procesamiento de señales de audio para limpiar el ruido de fondo y mejorar la calidad general del audio. Esto es crucial para asegurar que los algoritmos de análisis trabajen con datos consistentes y confiables.

1.2.- Transcripción Automatizada:

Para habilitar el análisis textual de los audios, implementaremos tecnologías de reconocimiento de voz y transcripción automática. Estas herramientas convertirán los diálogos hablados en texto, lo que permitirá aplicar técnicas de análisis de lenguaje natural.

Como observación será importante tener en cuenta que es necesario tener una definición del alcance en cuanto al tamaño y tipo de los archivos que se estarían utilizando dentro de la plataforma, y poder cumplir con los requerimientos anteriormente mencionados.

2.- Implementación de OCR mediante IA en el Proyecto.-

La implementación de Reconocimiento Óptico de Caracteres (OCR) mediante inteligencia artificial (IA) es una pieza clave en la propuesta técnica, ya que permitirá el análisis de imágenes y textos presentes en las conversaciones.

2.1.- Adquisición y Preprocesamiento de Imágenes:

La aplicación permitirá a los usuarios cargar imágenes relacionadas con las conversaciones, como capturas de pantalla, fotografías de documentos o gráficos. Estas imágenes serán procesadas para mejorar la calidad y reducir el ruido, asegurando una base sólida para el proceso de OCR.

2.2.- Detección de Regiones de Texto:

Utilizando técnicas de detección de objetos y segmentación de imágenes, la aplicación identificará las áreas de las imágenes que contienen texto. Esto garantiza que el OCR se enfoque únicamente en las regiones relevantes, aumentando la precisión del proceso.

2.3.- Extracción de Texto Mediante OCR:

Para el reconocimiento de caracteres en las imágenes, emplearemos algoritmos de OCR basados en IA. Estos modelos han sido entrenados en una amplia variedad de fuentes de texto y tipos de letra, lo que les permite reconocer con precisión los caracteres presentes en las imágenes.

2.4.- Integración con Análisis de Contenido:

Una vez que el texto ha sido extraído de las imágenes, se integrará en el análisis de contenido global. Esto permitirá que las conversaciones sean analizadas de

manera más completa al tener en cuenta tanto el texto hablado como el contenido visual que lo acompaña.

2.5.- Aprendizaje Continuo:

A medida que la aplicación sea utilizada y reciba ejemplos específicos de imágenes, podrá mejorar su capacidad de reconocimiento de caracteres y adaptarse a fuentes de texto y formatos particulares.

La implementación de OCR mediante IA se basará en modelos y técnicas ampliamente probados y entrenados. La aplicación estará diseñada para mejorar su precisión y capacidad de análisis a medida que se le suministren más ejemplos concretos de imágenes en el futuro.

3.- Implementación del Servicio de Traducción Multilingüe:

La integración de un servicio de traducción multilingüe en la aplicación es esencial para su versatilidad y accesibilidad en contextos globales.

3.1.- Integración de API de Traducción:

La aplicación se conectará con servicios de traducción basados en IA mediante algún servicio como GPT. Esta API proporciona potentes capacidades de traducción que cubren una amplia gama de idiomas, incluyendo los tres idiomas principales: inglés, francés y español.

3.2.- Traducción de Textos:

Cuando se identifica el idioma original, la aplicación enviará el texto a la API de traducción para obtener la versión traducida al idioma deseado. Esto permitirá a los usuarios comprender y comunicarse en el idioma de su elección.

3.3.- Visualización Contextual:

La traducción se incorpora de manera contextual en la conversación, ya sea en la transcripción de audios, textos o incluso en las imágenes que contienen texto. Esto asegurará que los usuarios puedan entender el contenido de manera integral, independientemente del idioma en el que se haya originado.

3.4.- Actualización de Modelos de Traducción:

A medida que los servicios de traducción basados en IA mejoren y actualicen sus modelos, la aplicación podrá beneficiarse de estas actualizaciones para ofrecer traducciones aún más precisas y naturales.

3.5.- Acceso a Múltiples Idiomas:

Aunque inicialmente nos centraremos en los tres idiomas principales (inglés, francés y español), la aplicación tendrá la flexibilidad para expandir su soporte a otros idiomas en el futuro, lo que ampliará su alcance y utilidad.

La implementación de un servicio de traducción multilingüe se basará en tecnologías probadas y modelos preentrenados. La aplicación se diseñará para adaptarse y mejorar su capacidad de traducción a medida que se acumule más información y ejemplos específicos de conversaciones en diferentes idiomas.

4.-Módulo de Herramientas Colaborativas:

La incorporación de herramientas colaborativas es fundamental para impulsar la comunicación efectiva y el trabajo en equipo dentro de la aplicación. Se piensa poder tener como tipos de archivo base: PDF y documentos de Excel, con esto se abarca los tipos de mayor uso en el mercado actual.

4.1.- Notas Compartidas:

La aplicación permitirá a los usuarios crear y compartir notas en tiempo real durante las conversaciones. Estas notas podrán ser utilizadas para capturar ideas clave, resaltar puntos importantes o tomar apuntes sobre decisiones tomadas durante la charla. Los usuarios podrán colaborar en la edición y actualización de estas notas en tiempo real.

4.2.- Compartir Transcripciones:

Además de las notas, los usuarios podrán compartir las transcripciones generadas automáticamente con otros participantes. Esto garantiza que todos tengan acceso a un registro textual preciso de la conversación, lo que facilita la revisión y el entendimiento.

4.3.- Acceso desde Múltiples Dispositivos:

Los usuarios podrán acceder a las herramientas colaborativas desde diferentes dispositivos, como dispositivos móviles y tabletas, garantizando la flexibilidad y la continuidad en la colaboración. La aplicación se diseñará para mejorar la colaboración y la comunicación en una variedad de contextos, desde reuniones de equipo hasta interacciones empresariales.

5.- Análisis de la Reunión:

El análisis de la reunión es una característica esencial de la propuesta técnica, ya que permite a los usuarios obtener una comprensión profunda de los temas discutidos y los resultados obtenidos durante las interacciones.

5.1.- Identificación de Temas Clave:

La aplicación utilizará técnicas de procesamiento de lenguaje natural para identificar y extraer los temas clave discutidos en la reunión. Se analizarán las transcripciones y los resúmenes generados previamente para determinar los temas más recurrentes y relevantes.

5.2.- Generación Automatizada de Informes:

La aplicación generará informes automatizados que incluirán una descripción general de los temas tratados, las discusiones relevantes y las decisiones tomadas. Los informes también podrían incluir visualizaciones gráficas que representen la distribución de temas y la participación de los asistentes.

5.3.- Exportación y Compartición:

Los informes podrán ser exportados en formatos como PDF o documentos de texto, lo que permitirá a los usuarios compartir fácilmente los resultados y decisiones de la reunión con otros miembros del equipo.

5.4.- Análisis Histórico:

La aplicación también permitirá a los usuarios comparar y analizar temas a lo largo de reuniones históricas. Esto facilitará el seguimiento de la evolución de los temas a lo largo del tiempo y la identificación de patrones. La aplicación se diseñará para ofrecer a los usuarios una visión completa y organizada de las conversaciones, permitiéndoles tomar decisiones informadas y estratégicas basadas en los resultados de la reunión.

6.- Módulo de Análisis de Riesgos:

La incorporación de un módulo de análisis de riesgos en la aplicación es esencial para ayudar a los usuarios a identificar posibles desafíos, problemas o áreas problemáticas en las conversaciones y decisiones tomadas.

6.1.- Detección de Posibles Riesgos:

La aplicación utilizará técnicas de análisis de texto y procesamiento de lenguaje natural para identificar posibles riesgos en las conversaciones. Esto podría incluir comentarios negativos, palabras clave que indican desacuerdos o frases que indiquen situaciones de conflicto.

6.2.- Alertas y Notificaciones:

Si se identifican áreas potenciales de riesgo, la aplicación generará alertas y notificaciones para los usuarios. Estas alertas podrían ser personalizables para que los usuarios reciban notificaciones según su nivel de tolerancia al riesgo.

6.3.- Informes de Análisis de Riesgos:

La aplicación generará informes de análisis de riesgos que resumirán los riesgos identificados. Estos informes podrían incluir detalles sobre las áreas problemáticas. La aplicación se diseñará para ayudar a los usuarios a identificar y abordar posibles desafíos en sus conversaciones, promoviendo una comunicación más efectiva y decisiones informadas.

Como punto final, es importante mencionar que el requerimiento para análisis de emociones y sentimientos, será prudente dejarlo para una segunda fase y no alargar en demasía el tiempo de desarrollo.

Tecnologías y Arquitectura.

Las tecnologías que se utilizarán para el desarrollo del sistema son:

- Para la parte de backend, se utilizará **Python**.
- Para el servicio de cloud se sugiere el uso de **AWS**:
 - **Computer Vision**: Herramienta para el reconocimiento de imágenes.
 - Para servir los archivos estáticos de la aplicación web se sugiere el uso de **AWS CloudFront** y **AWS S3** para servir los archivos estáticos de la aplicación con baja latencia y alta disponibilidad.
 - Para el poder de cómputo necesario en una aplicación de este calibre se recomienda como mínimo la infraestructura de amazon: **EC2 AWS c7gd.2xlarge**.
 - **Aurora BD** por su gran capacidad en el Big Data.
 - Para la aplicación móvil se sugiere el uso de **ionic** por su buena documentación y gran comunidad, brinda beneficios como facilitar un ágil desarrollo, híbrido (android y iOS), mantenible, escalable y un buen performance.
- Motor de inteligencia artificial se recomienda **GPT**, por su gran avance en el manejo de la información, sus accesos múltiples contextos y una integración muy eficiente con nuestra capa intermedia de mejora de contextos, propiedad de Indigo.
- **Proxy Indigo**: middleware capaz de gestionar y mejorar la interconexión (mejora de contextos) que se tiene con el motor de IA GPT, la cual hace mucho más eficiente el análisis de la información que se le solicita al motor.
- **SpeechToText** herramienta que nos permite la extracción de los textos que se generan desde los audios ingresados a la aplicación.

Control de Cambios.

El plan de trabajo base es el expresado en este documento. Una vez aprobado el plan, se priorizan las tareas y se comienzan a desarrollar.

Si existe una funcionalidad que no esté especificada en este documento se realiza una evaluación sobre el impacto del tiempo.



Nuestro objetivo es siempre mantenernos comunicados con el cliente y ser muy transparentes sobre lo que se desarrolla.

Proceso de control de cambios.

Cualquier situación o requerimiento que no se encuentre planeado dentro del alcance original y sea requerido por "El Cliente" o identificado por cualquier

interesado en el proyecto, deberá de seguir el siguiente proceso de evaluación y aprobación de cambios:

Pasos para el cambio:

1. Prevenir la causa raíz, para prevenir nuevos cambios.
2. Estimar el impacto del cambio y sus consecuencias en caso de no aprobarse.
3. Crear solicitudes de cambio.
4. Evaluar y aprobar el cambio.
 - a. Si la evaluación del cambio es menor o igual al 3% del total del proyecto en tiempo o costo, el PM tendrá la facultad para aprobar o rechazar el cambio
 - b. Si la evaluación del cambio es mayor al 3% el cambio se deberá de someter a un Comité de Control del Cambio (CCC) que estará conformado de forma inicial por el sponsor del proyecto, el Project Manager, El líder técnico, adicional se podrá integrar a cualquier otro participante clave que se defina durante el inicio del proyecto
5. Actualizar el estado de los cambios en el Sistema de control de Cambios.
6. Ajustar y actualizar las líneas base, procedimientos, documentos, etc.
7. Comunicar los cambios a los interesados.

Supuestos

El siguiente listado contiene las suposiciones sobre las que se genera la presente propuesta.

- Se supone que el tiempo estimado del desarrollo en general de la herramienta no contempla el tiempo adicional que llega a tomar la Play Store y/o App Store para la liberación de la misma (android e iOS).
- Se supone que el costo de licencia o costos ajenos al desarrollo móvil de Android y iOS deberá ser saldado por el cliente.
- Se supone que el alcance está bien definido por el cliente y cualquier impacto en el mismo haría esta propuesta invalida y será prudente una nueva cotización.

- Se supone que el límite máximo para la liberación dentro de las plataformas de la aplicación puede llegar a tomar un mes, lo más común es que este proceso se de en un plazo no mayor a dos semanas.
- Se supone que todos los costos de los servicios de terceros serán cubiertos por el cliente ya sea en licencia o por demanda así lo requieran.

Equipo Indigo

1 Product Owner 100%
1 Maquetador Móvil Sr 100%
1 Ingeniero de Software Móvil Sr 100%
1 Ingeniero de Software Móvil Jr 100%

Tiempo estimado

3 meses

*Está sujeto a cambios durante el desarrollo del proyecto.

FUNCIONALIDAD	CÁLCULO DE ESFUERZO	INVERSIÓN ASOCIADA
IA for AOSENUMA APP	3 meses	\$375,000.00
TOTAL	3 meses	\$375,000.00 más IVA.

*Los precios aquí mostrados son más IVA.

A continuación se hace un ejercicio de estimación y puedan visualizar los posibles costos adicionales al desarrollo por el acceso a las diferentes aplicaciones que se requieren para que esta aplicación sea 100% exitosa:

Por cada 5000 palabras / 30 minutos (Costos aproximados):

GPT: \$1 USD

SpeechToText: \$0.48 USD

Por cada 5000 transacciones:Computer Vision: \$ 7.23 USD

Costes mensuales:

EC2 AWS c7gd.2xlarge --- \$100 USD

Servicio S3 estimado con una cuota de solicitudes mensuales de 100,000 (cien mil llamadas), con una capacidad de 1,000 GB de datos mensuales: \$26.54 USD

#WeCraftSoftware



 55195580

 www.dsindigo.com/es/

 /indigo-smart-software-development

 /Indigo_IT

 /IndigoIT

