

1. Funcionalidades de la solución

a. Recepción de Entrada del Usuario

- i. Permitir la captura de texto en lenguaje natural | Permite la captura de texto en lenguaje natural. (Se entiende que ya lo hace)
- ii. Soportar diferentes formatos de entrada, como preguntas, comandos e instrucciones detalladas. (Igual, se entiende que la entrada es la que sea)
- iii. Manejar variaciones sintácticas y semánticas para interpretar correctamente el mensaje. [Correcciones, sugerencia, Sinónimos, etc] (Lo importante es la identificación para clasificar la tipología de los elementos lingüísticos presentes en el prompt)

b. Procesamiento del Lenguaje Natural (NLP – Natural Language Processing) |

- i. Utilizar modelos de PLN para analizar y comprender la intención del usuario.
- ii. Identificar entidades clave en la entrada, sustantivos como: (Importante para privacidad)
 1. Propios (Nombres). Marte, Navidad, Atari, Ford Motors, Juan y Uruguay.
 2. Comunes. perro, niña, hombre, felicidad, país, animal, calle, silla, computadora y justicia.
 3. Concretos. Gato, plátano, camisa, agua, televisión, avión, mesero, mesa y cama.
 4. Abstractos. Libertad, cristianismo, amor, odio, coraje, juicio, inteligencia, budismo.
 5. Contables. Tres ranas, un tenedor, un niño, cinco mujeres, un sombrero, diez monos.
 6. Incontables. Leche, tristeza, aire, jugo de manzana, paciencia y dignidad.
 7. Abrelatas. Telaraña, rompecabezas, bocacalle, cumpleaños, rascacielos y astronauta.
 8. Colectivos. Familia, flota, orquesta, equipo, rebaño, jauría y flora.

c. Manejo del Contexto (Si esto está perfecto)

- i. Conservar historial de la conversación con el objeto de mantener coherencia. (Y aprender preferencias del usuario)
- ii. Relacionar preguntas actuales con interacciones previas del usuario.
- iii. Diferenciación entre contextos nuevos y continuidad de temas previos.

d. Filtrado y Normalización de Entrada (ESTO NECESITO QUE SE ENCARGUE JORGE O LUIS)

- i. Filtrado (no recuerdo revisar) detecta y corrige errores ortográficos y gramaticales. (ESTO NECESITO QUE SE ENCARGUE JORGE O LUIS)
- ii. Filtrado contenido inadecuado o potencialmente peligroso. (ESTO NECESITO QUE SE ENCARGUE JORGE O LUIS)
- iii. Normalizar diferentes formas de expresión para mejorar la interpretación. (ESTO NECESITO QUE SE ENCARGUE JORGE O LUIS)

e. Soporte de Multimodalidad (Esto de momento es para desarrollo futuros)

- i. Permitir el uso de emojis, código y elementos especiales en la entrada. **REGLA DE NEGOCIO | validar con Antonio su implementación o no.**
- ii. Aceptar entradas de voz. **REGLA DE NEGOCIO NO DAR COBERTURA A IMAGENES O DUCUMENTOS | Super importante valoración de Antonio.**

2. Funcionalidades de Procesamiento

a. Análisis de la Entrada

- i. Extraer palabras clave y conceptos relevantes del mensaje.
- ii. Clasificar la intención (acción) del usuario en categorías predefinidas:
 1. Consulta de información. Cuando el usuario busque datos o explicaciones. Ejemplo: *¿Cuál es la capital de España?*
 2. Solicitud de acción. Cuando el usuario solicite se realice una tarea específica. Ejemplo: *Genera un resumen de la normativa GDPR y envíalo por email* (**consultar con Antonio viabilidad | las posibilidades de interacción y parametrización sería solo para usuarios registrados**).
 3. Expresión de opinión o sentimiento. Cuando el usuario expresa emociones o juicios. Ejemplo: *Me encanta esta aplicación.*
 4. Soporte o ayuda. Cuando el usuario necesita asistencia con un problema. Ejemplo: *No puedo iniciar sesión en mi ERP.*
 5. Conversación social. Cuando el usuario simplemente interactúa de manera informal/formal. Ejemplo: *¿Hola, cómo está/estás?*
- iii. Evaluar polaridad y sentimiento del mensaje.

b. Generación de Respuesta

- i. Determinar la mejor respuesta basada en los modelos (**Phi, GPT, 4o | Ver tema con Antonio**).
- ii. Capacidad para generar respuestas directas o estructuradas en listas o párrafos.
- iii. Adaptar el tono y la formalidad de la respuesta según el contexto [Lo que nombramos escenarios, perfilamiento del usuario, etc.].

c. Optimización y Refinamiento (**Involucrado el algoritmo de Antonio**)

- i. Evalúa la claridad y relevancia de la respuesta antes de enviarla.
- ii. Mejorar continuamente a través de técnicas de aprendizaje automático.
- iii. Ajustar la respuesta en función del historial de conversación y la retroalimentación del usuario.

d. Gestión de Errores

- i. Manejo de entradas ambiguas con solicitudes de aclaración.
- ii. Detectar inconsistencias en la conversación y proporcionar correcciones.
- iii. Ofrecer sugerencias alternativas si la consulta es incompleta o poco clara.

3. Funcionalidades de salida

a. Formato y Presentación de la Respuesta

- i. Estructurar las respuestas en función del tipo de consulta (texto, listas, tablas, etc.).
- ii. Personalizar el tono y estilo de la respuesta según la interacción previa.
- iii. Utilizar etiquetas de formato (negritas, cursivas, viñetas, color de fondo, color de tipografía, subrayados) para mejorar la legibilidad.

iv. Utilizar atributos y o estilos de conformidad (compliance) WGGG | AA

<https://www.w3.org/TR/WCAG22/#perceivable>

1. Enfoque no oscurecido
2. Movimientos de arrastre
3. Tamaño objetivo
4. Autenticación accesible

b. Adaptación al Medio de Comunicación

- i. Optimizar la presentación para pantallas de diferentes tamaños y dispositivos (Responsiva).
- ii. Capacidad para responder en diferentes idiomas.

c. Interactividad y Enlaces

- i. Proporcionar enlaces a información adicional o recursos externos cuando corresponda.
- ii. Permitir expandir respuestas largas en secciones interactivas.
- iii. Disponer de botones o acciones rápidas para facilitar la navegación del usuario.

d. Retroalimentación del Usuario (**Sí totalmente, y con esto va a ser útil el guardar el contexto y las respuestas identificadas como excelentes por el usuario**)

- i. Incluir opciones para que el usuario valore la utilidad de la respuesta.
- ii. Aprender de la retroalimentación para mejorar respuestas futuras.
- iii. Capacidad de solicitar aclaraciones si el usuario detecta que la respuesta no fue suficiente o incorrecta.

4. Funcionalidades de Almacenamiento

a. Gestión del Historial de Conversación

- i. Almacenar interacciones previas para mantener la coherencia del diálogo.
- ii. Permitir recuperar información relevante de conversaciones pasadas.
- iii. Ofrecer la opción de borrar o gestionar el historial según las preferencias del usuario.

b. Persistencia de Datos Relevantes

- i. Guardar información clave proporcionada por el usuario (como nombres o preferencias) para mejorar la experiencia.
- ii. Administrar datos temporales y permanentes según la naturaleza de la interacción.

1. Datos temporales: Son aquellos que se almacenan solo durante la sesión de conversación y se eliminan cuando la sesión finaliza.

Ejemplos:

- a. Historial de mensajes en una sola conversación.
- b. Variables contextuales como una ubicación mencionada para una consulta específica.
- c. Preferencias de formato de respuesta dentro de una misma sesión.

2. Datos permanentes: Son aquellos que se almacenan a largo plazo para mejorar la experiencia del usuario en futuras interacciones. Ejemplos:

- a. Nombre o alias del usuario.

- b. Preferencias de idioma o estilo de respuesta.
 - iii. Implementar mecanismos de seguridad para proteger la privacidad del usuario.
 - 1. Cifrado de datos: Se usa cifrado en tránsito (cuando los datos viajan entre el usuario y el sistema) y cifrado en reposo (cuando se almacenan en bases de datos) para evitar accesos no autorizados.
 - 2. Anonimización: Aplicar técnicas como la eliminación de datos personales o su reemplazo por identificadores anónimos para proteger la identidad del usuario.
 - 3. Control de Accesos: Solo usuarios, servicios o aplicaciones autorizadas pueden acceder a ciertos datos almacenados (evitar fugas de información).
 - 4. Cumplimiento de Normativas: Seguir regulaciones como GDPR (Reglamento General de Protección de Datos) o CCPA (Ley de Privacidad del Consumidor de California), permitiendo a los usuarios solicitar la eliminación de sus datos si así lo desean en cualquier momento.
 - 5. Autenticación y Autorización: Manejo de cuenta de usuario y contraseña. Control de lo que el usuario tiene permitido hacer en el sistema.
 - c. Optimización del Rendimiento
 - i. Utilizar bases de datos eficientes para la recuperación rápida de información.
 - ii. Gestionar la memoria y el almacenamiento para evitar sobrecargas del sistema.
 - iii. Archivar datos antiguos para mejorar la velocidad y eficiencia del procesamiento. (NO incluir, pero lo dejo porque es una buena práctica y no la quiero perder de vista para el futuro)
 - d. Seguridad y Privacidad en el Almacenamiento
 - i. Implementar cifrado en la información sensible almacenada.
 - ii. Dar cumplimiento a regulaciones de privacidad y protección de datos (como GDPR o CCPA). -> A nivel base de datos relacional SQL-Server incluye manejo de conformidad necesito revisar documentación en Azure.
 - iii. Proporciona opciones para que el usuario controle el uso de su información.
 - 5. Funcionalidades de Comunicación
 - a. Interacción en Tiempo Real
 - i. Responder de manera inmediata a las consultas del usuario.
 - ii. Mantener una comunicación fluida sin tiempos de espera prolongados.
 - iii. Ajusta la velocidad de respuesta según la complejidad de la consulta.
 - b. Multicanalidad
 - i. Capacidad para integrarse con múltiples plataformas como sitios web, aplicaciones móviles y redes sociales. (Planteamiento de Ezequiel, pero lo dejo para el futuro, Carly sugirió SDK con un enfoque de la situación actual)
 - ii. Soportar comunicación a través de APIs para su integración con otros sistemas.

iii. Capacidad para interactuar mediante asistentes de voz o dispositivos inteligentes:

1. Compatibilidad con Asistentes de Voz: El sistema puede integrarse con asistentes como Amazon Alexa, Google Assistant o Siri, permitiendo que los usuarios interactúen mediante comandos de voz en lugar de texto. (NO incluir, pero lo dejo para el futuro)
2. Conversión de Texto a Voz (TTS): Cuando la respuesta se genera en formato de texto, puede convertirse en voz sintética para que el usuario la escuche en lugar de leerla, habilitarla por defecto cuando el usuario presente capacidades diferentes (Debilidad visual, ceguera).
3. Conversión de Voz a Texto (STT): Permitir que el usuario hable en lugar de escribir, transformando su discurso en texto interpretable por el sistema, habilitarla por defecto cuando el usuario presente capacidades diferentes (debilidad visual, ceguera).
4. Personalización de Respuestas: Ajustar la entonación y velocidad de la voz generada para una experiencia más natural y adaptada al usuario.
5. Manejo de Contexto en Voz: Dado que en interacciones de voz no se puede "ver" el historial de la conversación, el sistema debe recordar lo dicho previamente dentro de una sesión para ofrecer respuestas coherentes. Ejemplo:
 - a. Usuario: "¿Cuál es el clima hoy en Ciudad de México?"
 - b. Asistente: "El clima en Ciudad de México es de 25°C con cielos despejados."
 - c. Usuario: "¿Y mañana?"
 - d. Asistente: "Mañana en Ciudad de México se espera una temperatura de 22°C con lluvias ligeras."

El asistente debe recordar que la ciudad mencionada en la primera pregunta sigue siendo relevante en la segunda, aunque el usuario no la haya repetido.

Para lograr esto, se utilizan mecanismos de gestión de contexto, que incluyen:

- Memoria de sesión: Se almacenan datos relevantes (como nombres, ubicaciones, temas => Sustantivos) durante la conversación en curso.
 - Inferencia semántica: Se analiza la relación entre frases para deducir a qué se refiere el usuario.
 - Confirmaciones contextuales: Si hay ambigüedad, el asistente puede preguntar: "¿Te refieres a Ciudad de México?"
- c. Soporte para Conversaciones Asíncronas | NO Considerar, la dejo para el futuro
- i. Permitir que el usuario retome conversaciones previas sin perder el contexto.
 - ii. Poder enviar notificaciones o recordatorios sobre interacciones anteriores.

- iii. Adaptar la respuesta en función de interacciones previas incluso si han pasado largos períodos de tiempo.
- d. Personalización de Comunicación
 - i. Ajustar el tono de la conversación según el usuario y el tipo de consulta.
Ejemplo: El tono de la conversación en un asistente de chat puede variar dependiendo de quién es el usuario y qué tipo de consulta realiza. Ajustar este tono es clave para mejorar la experiencia del usuario y hacer que la interacción sea más natural y efectiva.
 - 1. Ajustar según el usuario. El asistente puede adaptar su forma de responder en función del perfil del usuario. Algunas maneras en que esto puede aplicarse son:
 - a. Formal: Un usuario empresarial podría recibir respuestas más formales y estructuradas.
Ejemplo: (Necesito crear un ejemplo en el contexto de Gestión del cambio)
 - Usuario: "¿Cuál es el estado de mi pedido?"
 - Asistente: "Su pedido ha sido enviado y llegará en un plazo de 2 días hábiles."
 - b. Casual: Un usuario casual podría recibir respuestas más relajadas y amigables.
 - c. Nivel de conocimiento
Ejemplo:
 - Usuario: "¿Dónde está mi paquete?"
 - Asistente: "¡Buenas noticias! Tu paquete ya va en camino y debería llegarte en un par de días. 😊"
 - 2. Ajuste según el tipo de consulta (Aquí me quede)
 - ii. Poder usar lenguaje formal o informal dependiendo del contexto, pero amable y el factor humano no olvidarlo.
 - iii. Aprender de interacciones previas para mejorar la experiencia de comunicación.