Asistente Inteligente Multimodal - Documento Técnico

# 1. Introducción y Objetivos

Este documento describe los requerimientos técnicos y funcionales para el desarrollo de un asistente inteligente multimodal. El objetivo es crear una herramienta que integre procesamiento de visión por computadora, reconocimiento de voz, gestión de tareas y funciones proactivas de asistencia.

# 2. Alcance del Sistema

El sistema debe ser capaz de:  
- Analizar en tiempo real el contenido de la pantalla.  
- Escuchar desde el micrófono y transcribir voz a texto.  
- Crear y gestionar eventos en un calendario interno.  
- Generar recordatorios y notificaciones.  
- Actuar como asistente personal respondiendo preguntas y ejecutando acciones.

# 3. Requisitos Funcionales

RF1: Escaneo de pantalla y análisis con OCR e IA.  
RF2: Captura de audio por micrófono y transcripción con ASR.  
RF3: Creación, edición y eliminación de eventos en calendario.  
RF4: Recordatorios programados con alertas.  
RF5: Respuesta proactiva basada en contexto.  
RF6: Integración con APIs de terceros (ej. Google Calendar, MS Outlook).

# 4. Requisitos No Funcionales

RNF1: Procesamiento en tiempo real (<1s de latencia en transcripción).  
RNF2: Seguridad en almacenamiento y transmisión de datos (cifrado AES-256).  
RNF3: Alta disponibilidad (uptime 99.9%).  
RNF4: Escalabilidad horizontal.  
RNF5: Cumplimiento con GDPR y normativas locales de privacidad.

# 5. Arquitectura Propuesta

Arquitectura modular compuesta por:  
- Módulo de visión por computadora (OCR + NLP).  
- Módulo de reconocimiento de voz (ASR).  
- Módulo de gestión de calendario y recordatorios.  
- Motor de IA para acciones proactivas.  
- Interfaz de usuario y panel de control.

# 6. Integraciones Necesarias

APIs de reconocimiento óptico de caracteres (Tesseract, Google Vision API).  
APIs de reconocimiento de voz (Google Speech-to-Text, Vosk).  
APIs de calendario (Google Calendar API, Microsoft Graph API).

# 7. Consideraciones de Seguridad

Implementar autenticación multifactor, cifrado de extremo a extremo y control de acceso basado en roles para proteger datos sensibles.

# 8. Plan de Desarrollo y Fases

Fase 1: Diseño de arquitectura.  
Fase 2: Desarrollo de módulos de visión y voz.  
Fase 3: Integración de calendario y recordatorios.  
Fase 4: Motor de IA y funcionalidades proactivas.  
Fase 5: Pruebas y despliegue.

# 9. Riesgos y Limitaciones

R1: Dependencia de APIs externas.  
R2: Limitaciones de hardware para procesamiento local.  
R3: Retos legales y de privacidad.