



UNIVERSIDAD TÉCNICA DEL NORTE
FACULTAD DE INGENIERÍA EN CIENCIAS APLICADAS
CARRERA DE TELECOMUNICACIONES

TRABAJO DE INTEGRACIÓN CURRICULAR

TEMA:

“DESARROLLO DE UN PROTOTIPO ELECTRÓNICO PARA LA
AUTOMATIZACIÓN Y ANÁLISIS INTELIGENTE DE NOTAS CLÍNICAS
EN SESIONES TERAPÉUTICAS BASADO EN INTELIGENCIA
ARTIFICIAL.”

Trabajo de titulación previo a la obtención del título en Ingeniero en
Telecomunicaciones

Línea de investigación: Desarrollo, aplicación de software y cyber security (seguridad cibernética)

AUTOR:

Alexander Xavier Moreno Córdova

DIRECTOR:

Msc. Fabián Geovanny Cuzme Rodríguez

Ibarra – Ecuador 2026



**UNIVERSIDAD TÉCNICA DEL NORTE
BIBLIOTECA UNIVERSITARIA**

**AUTORIZACIÓN DE USO Y PUBLICACIÓN
A FAVOR DE LA UNIVERSIDAD TÉCNICA DEL NORTE**

1. IDENTIFICACIÓN DE LA OBRA

En cumplimiento del Art. 144 de la Ley de Educación Superior, hago la entrega del presente trabajo a la Universidad Técnica del Norte para que sea publicado en el Repositorio Digital Institucional, para lo cual pongo a disposición la siguiente información:

DATOS DE CONTACTO			
CÉDULA DE IDENTIDAD:	100330573-5		
APELLIDOS Y NOMBRES:	MORENO CÓRDOVA ALEXANDER XAVIER		
DIRECCIÓN:	Ibarra, Calle Rafael Carvajal 7-38 y Chorlaví		
EMAIL:	axmorenoc@utn.edu.ec / alexandermoreno41@gmail.com		
TELÉFONO FIJO:	No registra	TELÉFONO MÓVIL:	0995255552

DATOS DE LA OBRA	
TÍTULO:	DESARROLLO DE UN PROTOTIPO ELECTRÓNICO PARA LA AUTOMATIZACIÓN Y ANÁLISIS INTELIGENTE DE NOTAS CLÍNICAS EN SESIONES TERAPÉUTICAS BASADO EN INTELIGENCIA ARTIFICIAL.
AUTOR (ES):	MORENO CÓRDOVA ALEXANDER XAVIER
FECHA: DD/MM/AAAA	25/02/2026
SOLO PARA TRABAJOS DE GRADO	
PROGRAMA:	<input checked="" type="checkbox"/> PREGRADO <input type="checkbox"/> POSGRADO
TITULO POR EL QUE OPTA:	INGENIERO EN TELECOMUNICACIONES
DIRECTOR:	MSC. FABIÁN GEOVANNY CUZME RODRÍGUEZ
ASESOR:	MSC. CARLOS ALBERTO VÁSQUEZ AYALA

2. CONSTANCIAS

El autor (es) manifiesta (n) que la obra objeto de la presente autorización es original y se la desarrolló, sin violar derechos de autor de terceros, por lo tanto, la obra es original y que es (son) el (los) titular (es) de los derechos patrimoniales, por lo que asume (n) la responsabilidad sobre el contenido de la misma y saldrá (n) en defensa de la Universidad en caso de reclamación por parte de terceros.

Ibarra, a los 25 días del mes de febrero de 2026

EL AUTOR:

MORENO CÓRDOVA ALEXANDER XAVIER

CERTIFICACIÓN DEL DIRECTOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR

Ibarra, 25 de febrero de 2026

MSC. FABIÁN GEOVANNY CUZME RODRÍGUEZ

DIRECTOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR

CERTIFICA:

Haber revisado el presente informe final del trabajo de Integración Curricular, el mismo que se ajusta a las normas vigentes de la Universidad Técnica del Norte; en consecuencia, autorizo su presentación para los fines legales pertinentes.

MSC. FABIÁN GEOVANNY CUZME RODRÍGUEZ

C.C.: 1311527012

DEDICATORIA

A mis padres, Elías y Bertha, por enseñarme que rendirse no es opción, que los sueños se trabajan con esfuerzo, disciplina y fe; por su amor incondicional, por su apoyo firme, incluso en los momentos más difíciles, y por estar siempre sin condiciones, sin límites acompañándome en cada paso de mi formación.

A mi hermana Paulina, por su cariño constante, por su respaldo silencioso pero invaluable, y por estar presente desde siempre.

A Anita Paula, por su amor sincero y sin reservas, por caminar conmigo desde el inicio de esta historia, por abrazarme en los días buenos y sostenerme en los días grises, por sus palabras de aliento cuando las necesitaba y por creer en mí incluso cuando yo dudaba.

A mi pequeño Matías, mi mayor regalo, mi motor diario, la razón que le da sentido a mis luchas y cada esfuerzo. Por ser esa luz que me recuerda todos los días por qué vale la pena seguir adelante.

Y a todas las personas que creyeron en mí, que confiaron en este proceso y que, de una u otra manera, fueron parte de este logro... gracias.

Moreno Córdova Alexander Xavier

AGRADECIMIENTO

Agradezco profundamente a todos mis docentes por su guía incansable y a mi familia por creer en mí incluso cuando yo dudaba. Gracias por brindarme tantas oportunidades y por no soltarme nunca en el camino.

RESUMEN

La presente investigación tiene como objetivo desarrollar un prototipo electrónico para la automatización y el análisis inteligente de notas clínicas en sesiones terapéuticas, todo esto mediante técnicas de inteligencia artificial aplicadas al reconocimiento automático de voz y al procesamiento de lenguaje natural. Este estudio surge a partir de la problemática de la toma manual de notas, la cual interfiere en la atención clínica, debido a que puede generar inhibición en el paciente y dificulta la gestión estructurada de la información, además de evidenciar la inexistencia de un repositorio digital centralizado. Para el desarrollo del sistema se adoptó el modelo en cascada, siguiendo las normas ISO/IEC/IEEE 29148:2018 para la definición de requisitos y la ISO/IEC 25010 para la evaluación de la calidad del software. El prototipo fue diseñado bajo una arquitectura en capas e implementado mediante la integración de modelos preentrenados para la transcripción automática y el análisis semántico de la información clínica. Los resultados obtenidos demuestran que el sistema permite convertir el audio en texto estructurado, generar resúmenes y sugerencias clínicas como apoyo al profesional y almacenar la información de forma segura mediante mecanismos de autenticación y gestión. Esto se lo establece debido a la validación que se realizó con profesionales del área de psicología, en donde se observó un alto nivel de aceptación. El análisis comparativo entre modelos evidenció los siguientes resultados en evaluación profesional: DeepSeek y Grok alcanzaron un 76%, ChatGPT un 69% y Gemini un 60%. La valoración fue realizada por profesionales con amplia experiencia en su ámbito profesional, quienes priorizaron la solidez de la estructura formal y la precisión del enfoque clínico en las respuestas generadas.

Palabras clave: Modelos preentrenados, sesiones terapéuticas, reconocimiento de voz, procesamiento de audio, informes psicológicos. IA generativa.

ABSTRACT

This research aims to develop an electronic prototype for the automation and intelligent analysis of clinical notes in therapeutic sessions by means of artificial intelligence techniques applied to automatic speech recognition and natural language processing. The study arises from the challenges associated with manual note-taking, which interferes with clinical care, may inhibit patients, and hinders the structured management of information, while also revealing the absence of a centralized digital repository. For the system development, the waterfall model was adopted in accordance with ISO/IEC/IEEE 29148:2018 for requirements specification and ISO/IEC 25010 for software quality evaluation. The prototype was designed using a layered architecture and implemented through the integration of pre-trained models for the automatic transcription and semantic analysis of clinical information. The results demonstrate that the system enables the conversion of audio into structured text, the generation of summaries and clinical recommendations to support professionals, and the secure storage of information through authentication and management mechanisms. These findings are supported by the validation conducted with psychology professionals and students, which showed a high level of acceptance. The comparative analysis among models yielded the following results in the professional evaluation: DeepSeek and Grok achieved 76%, ChatGPT 69%, and Gemini 60%. The assessment was carried out by highly experienced professionals, who prioritized the robustness of the formal structure and the accuracy of the clinical approach in the generated responses.

Keywords: Pre-trained models, therapeutic sessions, speech recognition, audio processing, psychological reports, generative AI.

Tabla de contenido

DEDICATORIA	iv
AGRADECIMIENTO	v
RESUMEN	vi
ABSTRACT.....	vii
CAPÍTULO I. ANTECEDENTES	1
1.1 Tema	1
1.2 Planteamiento del Problema	1
1.3 Objetivos.....	2
1.3.1. Objetivo general.....	2
1.3.2. Objetivos específicos	2
1.4 Alcance	3
1.5 Justificación	5
CAPÍTULO II. FUNDAMENTACIÓN TEÓRICA.....	8
2.1. Salud Mental	8
2.1.1. Factores que Influyen en la Salud Mental.....	9
2.1.2. Intervenciones en Salud Mental.....	11
2.1.3. Rol de la Tecnología en la Salud Mental	12
2.2. Psicoterapia y Tecnología.....	13
2.2.1. Psicoterapia Digital.....	14
2.2.2. Aplicaciones de la Inteligencia Artificial en Psicoterapia	15
2.2.3. Eficacia de la Terapia Basada en IA	16
2.2.4. Ética y Consideraciones en la Terapia Asistida por IA	16

2.3.	Sistemas de Automatización en el Ámbito Clínico	18
2.3.1.	Software Clínico y Automatización.....	18
2.3.2.	Análisis Automatizado de Notas Clínicas.....	20
2.3.3.	Tecnologías Emergentes en Salud Mental.....	20
2.4.	Arquitectura de Software	21
2.4.1.	Arquitectura en Capas para Sistemas Clínicos	21
2.4.2.	Capa de Presentación: Interfaz de Usuario	22
2.4.3.	Capa de Aplicación: Procesamiento y Análisis	23
2.4.4.	Capa de Datos: Almacenamiento y Seguridad.....	24
2.5.	Inteligencia Artificial	25
2.5.1.	Inteligencia Artificial en Salud	26
2.5.2.	Modelos de Lenguaje Natural (NLP) para la Salud Mental	26
2.5.3.	Análisis de Sentimientos y Comportamientos en Terapia	27
2.5.4.	Plataformas de IA para Psicoterapia	27
2.6.	Seguridad y Ética en el Uso de IA	28
2.6.1.	Protección de Datos y Privacidad en Salud Mental	29
2.6.2.	Normativas Legales para el Uso de IA en Salud	29
2.6.3.	Principios Éticos en la Automatización de Procesos Clínicos.....	30
CAPÍTULO III. METODOLOGÍA		32
3.1.	Modelo en cascada.....	32
3.2.	Análisis	33
3.2.1.	Técnicas de investigación de recopilación de requisitos	33
3.2.1.1.	Cuestionario aplicado a la psicóloga encargada	34

3.2.1.2.	Revisión de trabajos relacionados.....	37
3.2.2.	Requisitos del sistema.....	39
3.2.2.1.	Requerimientos de Stakeholders.....	42
3.2.2.2.	Requerimientos del sistema	44
3.2.2.3.	Requerimientos de Hardware y Software	47
3.2.2.4.	Requerimientos de modelos.....	49
3.2.3.	Selección de Hardware	51
3.2.3.1.	SBC.....	51
3.2.3.2.	Micrófono	53
3.2.4.	Selección de software	54
3.2.4.1.	Lenguaje de programación.....	54
3.2.4.2.	IDE de programación.....	55
3.2.5.	Selección del modelo IA Generativa	56
3.3.	Diseño	58
3.3.1.	Software	58
3.3.1.1.	Arquitectura de Software.....	58
a)	Frontend.....	61
b)	Backend.....	69
Backend de Captura (Raspberry Pi – Flask + arecord).....	69	
Backend de Procesamiento (Colab – Flask + Drive + modelos IA).....	71	
3.3.1.2.	Diagrama de funcionamiento.....	74
3.1.1.	Hardware.....	79
3.1.1.1.	Case de protección.....	79

3.2.	Implementación.....	81
3.2.1.	Software	81
3.2.1.1.	Frontend.....	81
3.2.1.2.	Backend.....	99
a)	Backend de captura (Raspberry Pi – Flask + arecord).....	99
	Configuración del entorno base.	99
	Preparación de la estructura local	102
	Implementación de servicios locales	105
	Verificación Operativa	106
b)	Backend de procesamiento (Colab – Flask + Drive + Modelos IA).	108
	Preparación del entorno en la nube	108
	Configuración del servidor Flask	111
	Implementación del pipeline de procesamiento	112
	Publicación del servicio.....	116
3.2.2.	Hardware.....	117
CAPÍTULO IV. PRUEBAS DE FUNCIONAMIENTO		119
4.1.	Pruebas Preliminares.....	119
4.1.1.	Autenticación	122
4.1.2.	Captura en tiempo real/Carga manual de audio.....	123
4.1.3.	Transcripción Automática.....	125
4.1.4.	Procesamiento Semántico	126
4.1.5.	Extracción de Síntomas.....	127
4.1.6.	Detección de Indicadores de Riesgo	128

4.1.7. Generación de Informe Estructurado	129
4.1.8. Exportación en Formato JSON	129
4.1.9. Almacenamiento en Google Drive.....	130
4.1.10. Vinculación Automática al Paciente en el Dashboard.....	131
4.2. Pruebas aplicadas a sesiones en entorno real	132
4.2.1. Prueba de sesión 1 - Test_ind1	134
4.2.2. Prueba de Sesión 1 - Test_ind2.....	135
4.2.3. Prueba de Sesión 1 - Test_ind3.....	136
4.2.4. Prueba de Sesión 1 - Test_ind4.....	137
4.2.5. Prueba de Sesión 1 - Test_ind5.....	138
4.2.6. Discusión sobre las pruebas efectuadas	140
4.3. Análisis de la encuesta realizada.....	144
4.4. Cumplimiento del estándar ISO/IEC 25010	149
Conclusiones	155
Recomendaciones	157
Trabajos a Futuro	158
Bibliografía	159
ANEXOS	164

Índice de Tablas

Tabla 1 Factores asociados a la salud mental	10
Tabla 2 Herramientas y uso en Psicoterapia digital.....	14
Tabla 3 Desafíos situacionales del uso de IA en psicoterapia	17
Tabla 4 Tipos de software usados en gestión de información medica.....	19
Tabla 5 Elementos de la interfaz de usuario	22
Tabla 6 Características de la capa de almacenamiento y seguridad	24
Tabla 7 Aplicación de herramientas tecnológicas en Psicoterapia	27
Tabla 8 Preguntas cuestionario de calidad hacia la psicóloga encargada.....	34
Tabla 9 Descripción de acrónimos.....	40
Tabla 10 Beneficiarios del sistema	40
Tabla 11 Establecimiento del nivel de Importancia.....	41
Tabla 12 Requerimientos de Stakeholders.....	42
Tabla 13 Requerimientos del Sistema.....	45
Tabla 14 Requerimientos de Hardware y Software	47
Tabla 15 Requerimientos de modelos.....	50
Tabla 16 Comparativa de los diferentes SBC	51
Tabla 17 Comparación de diferentes micrófonos	53
Tabla 18 Comparación de lenguajes de programación	54
Tabla 19 Comparación de IDEs de programación	56
Tabla 20 Comparación de los modelos IA.....	57
Tabla 21 Prueba preliminar de sesión.....	120
Tabla 22 Resumen de las sesiones realizadas a las personas evaluadas	133

Índice de Figuras

Figura 1 Modelo en Cascada.....	33
Figura 2 Arquitectura de software	59
Figura 3 Diseño de acceso al sistema y autenticación del profesional	62
Figura 4 Diseño de gestión del paciente y selección de sesión.....	63
Figura 5 Diseño de la etapa de captura de audio de las sesiones de los pacientes.....	65
Figura 6 Diseño de guardado de archivos provenientes de las sesiones de los pacientes	66
Figura 7 Backend de Captura (Raspberry Pi – Flask + arecord)	71
Figura 8 Estructuración final del diseño de la parte Backend	73
Figura 9 Funcionalidad del sistema	74
Figura 10 Diagrama de bloques del funcionamiento de forma simplificado.....	75
Figura 11 Primera etapa del proceso de funcionamiento del sistema.....	75
Figura 12 Segunda etapa del proceso de funcionamiento del sistema.....	77
Figura 13 Tercera etapa del proceso de funcionamiento del sistema	78
Figura 14 Case de protección vista frontal	79
Figura 15 Perspectivas de vista del case de protección	80
Figura 16 HTML de login de acceso	82
Figura 17 Login de acceso	83
Figura 18 Procesamiento del audio y la solicitud de transcripción	84
Figura 19 Gestión de audios clínicos	85
Figura 20 Primera modalidad de subir el audio para procesamiento clínico.....	86
Figura 21 Petición de procesamiento en Google Colab.....	86
Figura 22 Barras de progreso asociadas al momento de la transcripción del audio a texto	87

Figura 23 Registro del evento de transcripción del audio a texto en Google Colab.....	88
Figura 24 Observación de la transcripción del audio a texto.....	89
Figura 25 Generación del diagnóstico presuntivo	90
Figura 26 Registro del evento de generar diagnóstico.....	91
Figura 27 Detección de hallazgos relevantes dentro del informe presuntivo	92
Figura 28 Interpretación del diagnóstico en base a los marcos de referencia.....	93
Figura 29 Inclusión del diagnóstico diferencial y nivel de gravedad	94
Figura 30 Generación del plan de tratamiento recomendado	95
Figura 31 Establecimiento de fases clínicas del plan de tratamiento recomendado	96
Figura 32 Segunda modalidad de subir el audio para procesamiento clínico.....	97
Figura 33 Registro de la sesión para almacenamiento.....	98
Figura 34 Recuperación del historial de pacientes	99
Figura 35 Instalación del Sistema Operativo del Raspberry Pi	100
Figura 36 Verificación del subsistema de audio (ALSA).....	101
Figura 37 Validación del entorno de ejecución Python.....	101
Figura 38 Establecimiento de la confirmación de disponibilidad que tienen las dependencias Flask y soporte CORS.....	102
Figura 39 Creación de directorios de almacenamiento de audios y sesiones	103
Figura 40 Proceso Backenend de la solicitud	104
Figura 41 Establecimiento de la confirmación de que el sistema recibió el audio y completó el flujo sin errores	105
Figura 42 Comprobación de levantamiento del servidor web local	106

Figura 43 Establecimiento de la comprobación del levantamiento del audio API en el modo archivo (file.py).....	107
Figura 44 Comprobación de conectividad entre Raspberry hacia a Google Colab	107
Figura 45 Bloque inicial de configuración del entorno de Google Colab	109
Figura 46 Instrucción para establecer Google Drive como almacenamiento	109
Figura 47 Permiso de notebook acceda a Google Drive.....	110
Figura 48 Importación de librerías y dependencias del Backend	110
Figura 49 Inicialización del servidor Flask y habilitación de CORS.....	111
Figura 50 Ejecución del servidor en segundo plano del entorno notebook	111
Figura 51 Transcripción con Faster-Whisper	112
Figura 52 Endpoint /diagnostic para análisis clínico con LLM.....	113
Figura 53 Almacenamiento JSON de pacientes y sesiones	115
Figura 54 Código de agrupación para los registros por paciente.....	115
Figura 55 Condicionamiento de registro de sesión.....	116
Figura 56 Túnel seguro con ngrok.....	116
Figura 57 Publicación del servicio local por el túnel ngrok	117
Figura 58 Implementación completa del hardware del sistema.....	117
Figura 59 Implementación completa del hardware del sistema desde otra perspectiva visual... ..	118
Figura 60 PRE-INT-01- Autenticación del terapeuta.....	122
Figura 61 PRE-INT-01- Captura de audio en tiempo real.....	123
Figura 62 PRE-INT-01 - Carga manual de audio	124
Figura 63 PRE-INT-01 - Verificación de la carga del audio	124
Figura 64 PRE-INT-01 - Transcripción Automática	125

Figura 65 PRE-INT-01 - Procesamiento Semántico.....	126
Figura 66 PRE-INT-01 - Extracción de Síntomas	127
Figura 67 PRE-INT-01 - Detección de Indicadores de Riesgo.....	128
Figura 68 PRE-INT-01 - Generación de Informe Estructurado.....	129
Figura 69 PRE-INT-01 - Exportación en formato JSON.....	130
Figura 70 PRE-INT-01 - Almacenamiento en Google Drive	131
Figura 71 PRE-INT-01 - Vinculación Automática al Paciente en el Dashboard	132
Figura 72 Prueba de Sesión 1 del Individuo 1	135
Figura 73 Prueba de Sesión 1 del Individuo 2	136
Figura 74 Prueba de Sesión 1 del Individuo 3	137
Figura 75 Prueba de Sesión 1 del Individuo 4	138
Figura 76 Prueba de Sesión 1 del Individuo 5	140
Figura 77 Interpretación de aceptación de los informes generados	147
Figura 78 Gráfica de desempeño de los modelos IA	148
Figura 79 Resultados de la encuesta de validación ISO/IEC 25010 realizada a la profesional encargada del departamento.....	153
Figura 80 Pregunta abierta realizada a la profesional encargada del departamento	153

CAPÍTULO I. ANTECEDENTES

1.1 Tema

Diseño de un sistema que permita la automatización de la toma de notas y análisis de patrones mediante inteligencia artificial en sesiones terapéuticas.

1.2 Planteamiento del Problema

La adolescencia y la juventud temprana representan momentos críticos para el desarrollo de la salud mental. Durante este periodo, suelen surgir numerosos trastornos mentales que, lamentablemente, pasan desapercibidos hasta la edad adulta. La crisis de salud mental entre los jóvenes se encuentra intrínsecamente ligada a una serie de problemáticas, tales como problemas de salud física, bajo rendimiento académico, abuso de sustancias, y violencia. (Patel et al., 2007)

Se estima que los estudiantes universitarios muestran una prevalencia más alta de trastornos mentales en comparación con otros grupos demográficos. De hecho, en las últimas décadas, este grupo ha experimentado uno de los mayores aumentos en la incidencia de problemas mentales. (Moreta Herrera et al., 2021)

Un estudio realizado con estudiantes de trabajo social encontró que las sesiones terapéuticas donde el terapeuta no tomaba notas recibieron mejores calificaciones en efectividad y percepción del paciente. Esto sugiere que tomar notas durante la terapia podría ser percibido como una distracción que afecta negativamente la interacción. (Hickling et al., 1984)

Una cantidad significativa de estudiantes universitarios experimenta angustia psicológica, y la mayoría participa en comportamientos de alto riesgo. Descuidar estos factores puede tener consecuencias graves en su rendimiento académico y en su futuro desarrollo. (Poorolajal et al., 2017)

Las terapias psicológicas son fundamentales para promover la salud mental. Una de las prácticas comunes y, a su vez, una de las limitaciones radica en la necesidad de tomar notas durante las sesiones terapéuticas. Cabe señalar que esta acción puede generar cierta inhibición en el paciente. Este podría sentirse resistente a expresar completamente sus emociones y pensamientos si percibe que el profesional está constantemente tomando notas. Esta dinámica puede afectar la calidad del tratamiento a seguir, ya que no permite registrar información de suma importancia sobre el paciente y su proceso terapéutico.(Hickling et al., 1984)

Por ende, la falta de un sistema para recopilar información y generar sugerencias entorno a entrevistas terapéuticas es uno de los problemas más comunes en el ámbito de la psicología clínica y la terapia. La distracción, las interrupciones conllevan a la imprecisión y la recopilación manual de datos debido a que consume mucho tiempo y recursos. Los terapeutas deben tomar notas durante las sesiones, transcribirlas después, organizarlas y analizarlas, lo que puede ser tedioso y propenso a errores.

1.3 Objetivos

1.3.1. Objetivo general

Desarrollar un sistema que aplique inteligencia artificial para la conversión eficiente de contenido auditivo a texto, con la finalidad que genere sugerencias y optimice procesos de toma de decisiones que son cruciales en el área de psicología.

1.3.2. Objetivos específicos

- Realizar el estado del arte de acuerdo con los fundamentos teóricos con respecto a la salud mental, transcripción de audios, bases de datos, programación, y técnicas de inteligencia artificial.

- Diseñar un sistema funcional de transcripción de audio a texto y que genere sugerencias en base al mismo en el cual demuestre las capacidades de un modelo de inteligencia artificial pre-entrenado.
- Evaluar el rendimiento del sistema a través de pruebas con conjuntos de datos reales con el fin de ofrecer recomendaciones útiles basadas en el análisis de las grabaciones generando sus respectivos registros.

1.4 Alcance

Para el desarrollo de este proyecto de integración curricular, se empleará la metodología del Modelo en Cascada, propuesto por Winston Royce en 1970. Este enfoque promueve un proceso sistemático y secuencial en el desarrollo de software, caracterizado por su disciplina y basado en etapas de análisis, diseño, pruebas y mantenimiento. (Gamboa, 2018)

Análisis: Esta etapa se alinea con el objetivo específico uno, en el cual se realiza la planificación, análisis y especificación de los requisitos. Se llevará a cabo la fundamentación teórica de acuerdo con el análisis de la salud mental, la transcripción de audio a texto, el tipo de base de datos a utilizar, el lenguaje de programación y sus diferentes librerías, además de procesos de inteligencia artificial. De esta manera, con todos estos requisitos, se puede seleccionar las mejores características acorde a las necesidades del proyecto.

En la programación por capas, los objetos se dividen según su función en tres categorías principales: la Capa de Interfaz, que interactúa con el usuario; la Capa de Lógica de Negocio, que maneja el procesamiento interno y actúa como enlace entre capas; y la Capa de Datos, que se comunica con bases de datos y otros sistemas. (Del Valle & Granados, 2007)

Diseño: Esta etapa se alinea con el objetivo específico dos y parte del tres, se definirán los diferentes procesos que se llevarán a cabo en el sistema, por lo que se formula una solución en la

toma de notas que el profesional realiza. Teniendo en cuenta que realizara la grabación de las sesiones para a posterior realizar el proceso de tratamiento de audios para su respectiva conversión a texto y en base a eso identificar patrones generando sugerencias en las sesiones de terapia. Para lo cual se busca optimizar este proceso mediante un diseño de una interfaz de software, Además de generar un respaldo en una base de datos y se aplicará un modelo de aprendizaje de inteligencia artificial ya entrenado, de esta manera como resultado final se pueda obtener un documento donde se contenga registros cronológicos de actividades, decisiones clave, eventos relevantes y cualquier otra información importante que facilite el seguimiento del progreso de los pacientes de esta manera sirva como referencia histórica para otorgar sugerencias en futuras evaluaciones y mejoras para el profesional a cargo.

Implementación: En esta etapa se alinea con parte del objetivo específico dos y tres, de acuerdo con la arquitectura de tres capas, se detalla la capa de presentación, la cual es la encargada de interactuar directamente con el usuario, en este caso, el profesional a cargo. En este contexto, la interfaz de usuario se definirá como una página web donde los usuarios podrán cargar archivos de audio para su transcripción, ver las transcripciones generadas y, a su vez, interactuar con las sugerencias basadas en el contenido de las transcripciones de la sesión terapéutica.

La capa aplicación es responsable de procesar la entrada del usuario. Aquí se llevaría a cabo la transcripción de audio utilizando algoritmos de inteligencia artificial ya entrenados, teniendo en cuenta los diferentes modelos de reconocimiento de voz basados en redes neuronales profundas. Además, en esta capa se involucra también el análisis del contenido de las transcripciones para identificar temas relevantes para la terapia psicológica. También se podría integrar sistemas de recomendación para sugerir recursos, ejercicios o estrategias terapéuticas

basadas en el análisis del contenido de las transcripciones. Esto como sugerencias al profesional a cargo.

Finalmente, en la capa de datos, se maneja el acceso y la manipulación de los datos utilizados en este sistema, donde se pretende almacenar las transcripciones de texto y registros de terapia. Además, esta capa puede incluir la integración con varias APIs de servicios de inteligencia artificial, como servicios de transcripción de voz a texto y herramientas de procesamiento del lenguaje natural (NLP) para análisis de texto ya existentes.

Verificación: En esta etapa se alinea con el objetivo específico tres que busca la funcionalidad del sistema, donde se realizan sus respectivas pruebas para validar que el sistema cumple con los requisitos establecidos, principalmente la transcripción y la importancia de las sugerencias en el contexto de la terapia psicológica. En la Figura 2 se detallan los procesos que se llevarán a cabo para garantizar la funcionalidad del sistema. Una vez desplegado, es crucial contar con un manual de usuario que garantice su correcto uso.

1.5 Justificación

La siguiente investigación proporcionará a los profesionales del área de psicología una herramienta que les permitirá transcribir sesiones de terapia de forma eficiente y precisa. De esta manera, les ayudará a tener una mayor concentración con sus pacientes en lugar de centrarse en tomar notas durante las sesiones.

Es crucial establecer normas éticas para las aplicaciones de IA en terapias, pero aún no está claro cómo garantizar la responsabilidad en su uso. Una posible solución es exigir supervisión por parte de profesionales de la salud mental para asegurar una evaluación adecuada de la situación del paciente. (Fiske et al., 2019)

Al utilizar algoritmos de inteligencia artificial ya entrenados, se permite optimizar el talento humano y la eficiencia de los resultados en las transcripciones del sistema. De esta manera, se podría identificar patrones y temas recurrentes en las sesiones de terapia, permitiendo comprender mejor las necesidades y preocupaciones de los pacientes.

Es esencial colaborar con científicos de la computación, ingenieros y analistas de datos para garantizar que las tecnologías cumplan con los estándares éticos del campo de la salud mental. La inteligencia artificial nos brinda la oportunidad, como investigadores, de desarrollar enfoques innovadores para investigar y abordar los problemas relacionados con la salud mental. (Gordon & Turnbull, 2024)

La construcción de un sistema que integre múltiples tecnologías, como la transcripción de voz a texto y la generación de sugerencias, ofrecerá beneficios prácticos para los profesionales del área de salud mental y sus pacientes al mejorar la toma de decisiones y la efectividad de la terapia psicológica. Además, contribuirá al avance técnico y teórico de diferentes áreas, en este caso la ingeniería y el campo de la salud.

Esta investigación puede contribuir en varios ámbitos. Uno de ellos puede ser en el campo académico, debido a que, al proporcionar dicha herramienta, puede servir como base para estudios futuros sobre el uso de tecnología en el tratamiento psicológico. También en el ámbito social, debido a que puede mejorar la calidad de atención. Y en el entorno educativo ya que puede repercutir, sólo ya que poder otorgar esta herramienta a los profesionales de la salud mental y partir de dicha herramienta para mejorar sus prácticas y enseñar a los estudiantes de dicha área a utilización de la tecnología en su ejercicio.

Asimismo, este proyecto está en consonancia con los ODS (Objetivos de Desarrollo Sostenible) de la ONU, en concreto con la serie de metas del ODS 3 que está vinculado con la

obtención de una vida sana y la promoción del bienestar para toda la gente en todas las etapas. El apartado 3.4 de dicho objetivo establece el compromiso de reducir en un tercio la mortalidad prematura por enfermedades no transmisibles para el año 2030, mediante acciones de prevención, tratamiento y promoción de la salud mental y el bienestar.(Cepal, 2019).

Este proyecto se organiza con el plan de desarrollo del nuevo Ecuador para el período 2024-2025, específicamente con uno de los objetivos nacionales dirigidos a mejorar las condiciones de vida de la población de manera integral. Esto implica promover un acceso equitativo a servicios de salud, vivienda y bienestar social, de acuerdo con el eje social establecido en el apartado de políticas y estrategias, con énfasis en la Política 1.4. Se enfoca en fortalecer la vigilancia y prevención de enfermedades, incluidas las no transmisibles, y en mejorar el modelo comunitario de salud mental con énfasis en la prevención y la rehabilitación.(Noboa et al., 2024)

CAPÍTULO II. FUNDAMENTACIÓN TEÓRICA

El presente capítulo establece el marco teórico donde se fundamenta el desarrollo de las tecnologías basadas en inteligencia artificial que van transformando diversos campos en este caso para el ámbito de la salud mental y la terapia clínica. Abordando el análisis de la toma de notas y la automatización de los procesos terapéuticos permitiendo mejorar la precisión, eficiencia y accesibilidad de las notas. Además, los conceptos que sustentan el desarrollo del prototipo electrónico detallando aspectos claves como elementos electrónicos, inteligencia artificial, el procesamiento de datos, y tecnologías que permiten la captura, almacenamiento y análisis de información en entornos terapéuticos.

Tomando en cuenta antecedentes y estudios realizados en este campo se fundamenta esta investigación con el uso de algoritmos inteligentes para el análisis de información en sesiones terapéuticas lo que permitirá contextualizar el problema de estudio y proporcionar una base para el diseño e implementación del prototipo.

2.1.Salud Mental

La salud mental es un componente fundamental del bienestar general de los individuos y las comunidades, definida como un estado de bienestar en el cual las personas son capaces de realizar sus capacidades, manejar el estrés cotidiano de manera efectiva, trabajar de manera productiva y contribuir a su entorno social, según la Organización Mundial de la Salud, esta definición implica también el desarrollo de habilidades emocionales y sociales, esenciales para enfrentar los retos del ciclo vital. Teóricamente, la salud mental ha sido abordada desde diversas perspectivas, como el modelo biopsicosocial propuesto por Engel (1977), que enfatiza la interrelación de factores biológicos, psicológicos y sociales en la configuración de la salud mental,

lo que refleja la necesidad de intervenciones multidimensionales que consideren estos factores (Organización Mundial de la Salud, 2022).

2.1.1. Factores que Influyen en la Salud Mental

Los factores que influyen en la salud mental están interrelacionados y pueden dividirse en tres categorías principales: biológicos, psicológicos y sociales, cada uno de ellos con una influencia significativa en el bienestar psíquico de las personas, a menudo actuando de manera combinada. Se detalla en la Tabla 1 estos factores con sus respectivos ejemplos. Desde la perspectiva biológica, investigaciones en genética han demostrado que ciertas condiciones, como la depresión y la esquizofrenia, tienen un componente hereditario significativo, además de desequilibrios en neurotransmisores como la serotonina y la dopamina, los cuales afectan directamente los estados mentales y las respuestas emocionales, junto con factores como enfermedades crónicas y deficiencias nutricionales que exacerban trastornos psicológicos (Moreta et al., 2018). Además, la información expuesta en las investigaciones sobre epigenética indica cómo los factores, las influencias y las experiencias ambientales pueden interaccionar en la expresión de los genes, sugiriendo que la salud mental dependerá de la herencia, pero que también existe una interacción con el entorno de la polaridad de la transmisión biológica (Leoncio, 2021).

En el ámbito de la psicología, el efecto del impacto de las experiencias traumáticas y las experiencias adversas a una edad temprana desde la infancia están suficientemente documentados, citando Garrett (2023) a Bowlby (1969) que añade, en su teoría del apego, que aquellas relaciones inseguras en la infancia pueden posteriormente predisponer a sufrir desórdenes emocionales en la vejez, pero que la resiliencia, es decir, la capacidad de una persona para sobreponerse a un evento que ha tenido un alto contenido estresante también desempeña una función muy importante, determinando, a su vez, la gestión efectiva del estrés. Desde una perspectiva cognitiva, se definen

los sesgos del procesamiento de la información como la forma de mantener patrones de pensamiento distorsionados que hacen que las condiciones que se dan en el presente, propias de la ansiedad y de la depresión hayan de ser exageradas y, en consecuencia, la buena práctica de la práctica del abordaje terapéutico haga hincapié en este aspecto de los esquemas (Olza et al., 2021). Los modelos de aprendizaje conductuales que, sin embargo, resaltan cómo las experiencias vean reforzadas las expectativas de la repetición de las respuestas de evitación, reforzando también los problemas psicológicos de un tipo más largo.

Con respecto a los factores sociales, se puede comprobar cómo las condiciones del contexto sociocultural, las diferencias socioeconómicas o la forma de acceso a los servicios de salud mental afectan a la salud psíquica; las teorías del capital social que enfatizan que la calidad de las redes de apoyo, entendiéndolas como la familia o la comunidad, inciden en cómo las personas puedan enfrentarse a las actividades difíciles y cómo factores como la experiencia del sufrimiento y la violencia o la inseguridad pueden contribuir a deteriorar la salud mental, fundamentalmente en las poblaciones vulnerables. Por ejemplo, investigaciones han demostrado que las tasas de trastornos mentales son significativamente mayores en comunidades afectadas por la pobreza y la exclusión social, lo que refuerza la necesidad de políticas públicas que aborden las desigualdades estructurales. La tabla a continuación resume estos factores:

Tabla 1
Factores asociados a la salud mental

Categoría	Factores Principales	Ejemplos
Biológicos	Predisposición genética, desequilibrios neuroquímicos, enfermedades crónicas	Esquizofrenia, depresión, deficiencia de vitamina D

Psicológicos	Experiencias traumáticas,	Trastorno de estrés
	resiliencia, patrones cognitivos negativos	postraumático, ansiedad, baja autoestima
Sociales	Nivel socioeconómico, redes de apoyo, violencia	Pobreza, aislamiento social, discriminación

Nota. En la tabla se aprecia a los factores asociados a la salud mental y una ejemplificación de cada uno. Adaptado de (Olza et al., 2021; psicologiaymente.com, 2020)

La interrelación de estos factores subraya la necesidad de un enfoque integral para abordar la salud mental, el impacto de una predisposición genética puede ser moderado por un entorno social favorable y estrategias de afrontamiento efectivas, por ello, las políticas públicas deben priorizar la reducción de desigualdades en el acceso a servicios y promover entornos que fomenten el bienestar emocional, lo que incluye fortalecer redes comunitarias y garantizar el acceso a una educación emocional desde edades tempranas.

2.1.2. Intervenciones en Salud Mental

Las intervenciones en salud mental se dividen en preventivas, terapéuticas y comunitarias, cada una con objetivos específicos orientados al bienestar integral. En la prevención, programas psicoeducativos han demostrado ser eficaces en la reducción del estigma asociado a los trastornos mentales, promoviendo una cultura de aceptación y apoyo (Martínez, 2023). Las intervenciones terapéuticas, por su parte, incluyen enfoques como la terapia cognitivo-conductual (Beck, 1979), que se enfoca en la reestructuración de patrones de pensamiento disfuncionales, además, el tratamiento farmacológico, cuando es necesario, actúa como un complemento para casos moderados a graves, donde los desequilibrios químicos tienen un papel preponderante (González et al., 2018). En lo que se refiere a la comunidad, intervenciones como el establecimiento de redes

de apoyo, o programas de inserción laboral para la reinserción social, han mostrado evidenciar una importancia fundamental para la reinserción social de las personas con trastornos mentales. Estas intervenciones no persiguen tan solo mitigar los síntomas de la enfermedad, sino que también pretenden fomentar la resiliencia y la funcionalidad en la vida cotidiana, adaptándose a las características sociales y culturales de cada población.

2.1.3. Rol de la Tecnología en la Salud Mental

La integración de tecnologías digitales en el ámbito de la salud mental ha revolucionado la forma en que se accede y se proporcionan los servicios de apoyo. Las aplicaciones móviles y las plataformas en línea permiten a los usuarios monitorizar su estado emocional, acceder a recursos terapéuticos y participar en intervenciones psicológicas sin las restricciones geográficas o temporales que tradicionalmente limitaban el acceso (Sanchez, 2022) . Ejemplo de aplicación, a través de aplicaciones como Headspace y Calm se facilitan herramientas para meditar y gestionar el estrés, o por otro lado a través de plataformas como BetterHelp podemos conectar a los usuarios con terapeutas cualificados para sesiones en línea, o por último la inteligencia artificial (IA) puede ofrecer un potencial educativo en cuanto a la detección temprana de trastornos mentales, detectando patrones en datos (pcient opinión + Marcelle et al. 2019; O’Daffer et al. 2022).

La aplicación y la tecnología digital han transformado la manera en la que podemos aprovechar y llevar a cabo los servicios de apoyo facilitando un impacto en cuanto a accesibilidad y eficacia de los tratamientos.**El impacto positivo de las aplicaciones móviles en el autogestión emocional**

Hallazgos de investigaciones recientes observan que un número alto de las personas (más de 70%) que utilizan aplicaciones para la salud mental contribuyen a la autogestión de su emoción (Marcelle et al. 2019). Un ejemplo es el caso de aplicaciones como Headspace y Calm, que han

logrado más de 100 millones de descargas en todo el mundo, calando hondo en la cultura popular e incluso desarrollando sentido a nivel mundial donde están rastreadas las aplicaciones de salud mental (O'Daffer et al. 2022). Por otro lado, plataformas de terapia en línea como BetterHelp han tenido un incremento del 40% de la demanda por este tipo de terapia desde que comenzó la pandemia por COVID-19, lo que hace ver cómo la pandemia ha ayudado en acelerar el uso de servicios digitales de salud mental (Powell et al. 2019). Adicionalmente, la inteligencia artificial (IA) está desempeñando un papel importante en la detección temprana de trastornos mentales en el análisis de patrones de texto, voz y en el comportamiento por un grado de precisión de hasta un 85% en relación con la identificación de signos tempranos de depresión (Drigas et al. 2020).

Si bien es un hecho que existan avances, también persisten desafíos cuando se habla de los métodos de tratamiento de los trastornos mentales entre ellos la privacidad y la seguridad de datos, porque el 64% de los usuarios añade preocupación por el uso de su propia información personal en las aplicaciones de salud mental (Op. cit.). Esto subraya la necesidad de regulaciones más estrictas y la implementación de protocolos éticos que garanticen la protección de datos sensibles. En conclusión, la tecnología está redefiniendo la manera en que se aborda la salud mental, ofreciendo soluciones innovadoras que deben ser gestionadas con responsabilidad para maximizar sus beneficios.

2.2. Psicoterapia y Tecnología

Actualmente, los avances tecnológicos están transformando todos los aspectos de la vida cotidiana, y uno de ellos es la salud mental. La psicoterapia, al tener que integrar herramientas tecnológicas expande su alcance y efectividad. Esta combinación plantea no solo oportunidades, sino que también desafíos éticos, técnicos, y metodológicos.

Esta combinación innovadora, exige un análisis profundo sobre su impacto en la relación de terapeuta-paciente pues hay factores importantes como la privacidad de los datos y la efectividad de las herramientas digitales en contextos terapéuticos debido a que transforma el panorama de la psicoterapia y varios retos a enfrentar para que exista una implementación responsable.

2.2.1. Psicoterapia Digital

La psicoterapia digital ha transformado el acceso a los servicios psicológicos, permitiendo que las personas reciban apoyo emocional y terapéutico sin limitaciones geográficas, lo cual es especialmente relevante en comunidades remotas o con acceso limitado a especialistas. Esta modalidad utiliza plataformas digitales, como videoconferencias y aplicaciones, para proporcionar intervenciones psicológicas estructuradas. Estudios recientes indican que más del 50% de los pacientes que han utilizado terapias en línea reportan niveles similares de satisfacción en comparación con las sesiones presenciales (Yanfei, 2018). Además, herramientas como Talkspace y BetterHelp han experimentado un aumento exponencial en usuarios, especialmente durante la pandemia de COVID-19, lo que evidencia la creciente aceptación de estos servicios. A continuación, en la Tabla 2 se detalla las características de cada una de las herramientas más utilizadas por los usuarios.

Tabla 2
Herramientas y uso en Psicoterapia digital

Plataforma	Usuarios Actuales (millones)	Características Clave
Talkspace	1.5	Terapia por texto, videollamadas

BetterHelp	2.5	Red de terapeutas licenciados, accesibilidad 24/7
Woebot	1.0	Chatbot basado en TCC

Nota. En la tabla se contextualiza algunas de las herramientas más usadas en psicología digital, sus características y cantidad de usuarios. Adaptado de (Yanfei, 2018)

Luego de la expansión de la utilización de este tipo de plataformas, es fundamental resaltar de qué modo han dado lugar a una transformación estructural en la forma de prestar servicios psicológicos, ya que Talkspace y BetterHelp permiten dar lugar a una mayor individualización, donde los clientes eligen a los terapeutas adecuándolos a sus necesidades, por otro lado Woebot se sirve de inteligencia artificial para tener respuestas inmediatas y en donde brinda la posibilidad del acceso a esta interacción psicológica. Por otro lado estos servicios han aumentado notablemente la disponibilidad del acceso a la ayuda psicológica aunque su gran difusión también trae consigo ciertos problemas en relación a la ética y la regulación.

2.2.2. Aplicaciones de la Inteligencia Artificial en Psicoterapia

La inteligencia artificial (IA) está desempeñando un papel crucial en la psicoterapia al ofrecer soluciones automatizadas y personalizadas para los usuarios. Chatbots como Woebot, impulsados por IA, utilizan principios de la terapia cognitivo-conductual (TCC) para interactuar con los usuarios y proporcionar apoyo emocional en tiempo real. Estudios clínicos han demostrado que estas herramientas son efectivas para reducir los niveles de ansiedad y depresión leve, con tasas de efectividad del 70% en intervenciones a corto plazo (Concha, 2024). A parte de eso, la IA se encuentra también surgiendo y siendo utilizada para analizar patrones de lenguaje y comportamiento, permitiendo así que los terapeutas obtengan información bastante valiosa acerca de los estados emocionales de sus pacientes.

Sin embargo, por otro lado, a pesar de ser eficaces, estas herramientas presentan limitaciones. La dependencia de los algoritmos podría llegar a dar lugar a errores en los diagnósticos o en las recomendaciones, especialmente cuando los datos de entrada no son los adecuados o porque los datos de entrada no representan la perspectiva adecuada para reflejar el contexto cultural del usuario en cuestión. Además, aunque los chatbots son efectivos en intervenciones tempranas, su capacidad para abordar problemas complejos o crisis emocionales graves sigue siendo limitada, lo que resalta la necesidad de supervisión profesional.

2.2.3. Eficacia de la Terapia Basada en IA

La eficacia de la terapia basada en IA ha sido un tema de debate entre investigadores y clínicos, sin embargo, evidencias recientes sugieren que estas herramientas pueden complementar las intervenciones tradicionales. En este sentido, un trabajo de investigación seguido por el de Rivera y Sánchez (2016), por ejemplo, observó que los pacientes de aplicaciones de terapia digital mediadas por IA presentan decrementos significativos en los síntomas de la depresión y esto son paralelos a los resultados de la terapia convencional; sin embargo, también se ha apreciado que estas plataformas son más útiles combinándolas para complementar la terapia humana -y no para sustituirlas-, sobre todo en trastornos severos.

En comparación con las modalidades tradicionales, la terapia basada en IA ofrece ventajas como el acceso inmediato y la personalización de las intervenciones, pero carece de la empatía y el juicio clínico humano. Esto subraya la importancia de integrar estas tecnologías en un modelo híbrido, donde los terapeutas puedan supervisar y complementar las intervenciones automatizadas.

2.2.4. Ética y Consideraciones en la Terapia Asistida por IA

El uso de IA en psicoterapia plantea desafíos éticos significativos relacionados con la privacidad, la seguridad de los datos y la toma de decisiones automatizadas. Uno de los principales

riesgos es el manejo de información confidencial, ya que el 62% de los usuarios expresa preocupación por cómo se almacena y utiliza su información en estas plataformas (Andreu & Martínez, 2021). Además, la falta de regulaciones claras sobre el uso de IA en el ámbito terapéutico puede conducir a desigualdades en el acceso y la calidad de los servicios. A continuación, en la Tabla 3 se refleja características importantes de los principales desafíos de IA en psicoterapia.

Tabla 3
Desafíos situacionales del uso de IA en psicoterapia

Desafío Ético	Ejemplo	Posible Solución
Privacidad de datos	Filtración de información sensible en plataformas	Implementación de sistemas de encriptación avanzada
Toma de decisiones	Diagnósticos erróneos por sesgos en algoritmos	Supervisión humana constante y validación de resultados
Accesibilidad	Barreras tecnológicas en comunidades de bajos recursos	Subvenciones gubernamentales para el acceso a dispositivos

Nota. En la tabla se ejemplifica algunas situaciones a considerar en la aplicación de inteligencia artificial en terapia y sus posibles soluciones. Adaptado de (Andreu & Martínez, 2021)

Posteriormente, la implementación de soluciones éticas y técnicas adecuadas será clave para mitigar los riesgos asociados al uso de IA en la psicoterapia. Por ejemplo, se deben establecer normativas internacionales que garanticen el cumplimiento de estándares de privacidad, así como fomentar la capacitación de los terapeutas para integrar estas herramientas de manera ética y efectiva en sus prácticas clínicas. Aunque los beneficios de la IA son innegables, el equilibrio entre innovación tecnológica y protección de los derechos humanos debe ser prioritario.

2.3.Sistemas de Automatización en el Ámbito Clínico

La automatización en el ámbito de la salud ha evolucionado significativamente desde la implementación de los primeros sistemas de gestión hospitalaria en la década de 1960, en sus inicios, estos sistemas estaban diseñados exclusivamente para la gestión administrativa, como el registro de pacientes y la facturación, sin embargo, con el avance de la tecnología, se comenzaron a integrar herramientas más sofisticadas para apoyar los procesos clínicos. En las décadas de 1980 y 1990, la aparición de los sistemas de registros médicos electrónicos (RME) representó un importante avance, ya que permitieron un manejo más eficaz de la información clínica y mejoraron la continuidad de la atención a los pacientes, y en la actualidad la automatización ya ha llegado a avanzar a niveles de sofisticación elevados, como la inteligencia artificial, el análisis de datos en tiempo real y los sistemas de soporte a la decisión clínica. (Preciado et al., 2021)

2.3.1. Software Clínico y Automatización

El software clínico ha sido un pilar fundamental en la automatización de los procesos hospitalarios, facilitando la gestión integral de las operaciones médicas y administrativas, herramientas como los sistemas de información hospitalaria (HIS, por sus siglas en inglés) y los RME permiten un acceso rápido y centralizado a los datos del paciente, mejorando la calidad del cuidado y reduciendo los errores médicos. Estas soluciones garantizan la capacidad de programar citas, de gestionar recetas electrónicas y el seguimiento de los resultados del laboratorio, contribuyendo la misma a la eficiencia en el flujo de trabajo clínico. Un estudio elaborado por la Healthcare Information and Management Systems Society (Shaikh et al., 2022), señala que los hospitales que implementan sistemas de automatización clínica son capaces de reportar una reducción del 30% con respecto a los errores relacionados con la medicación, reflejando así el

impacto positivo en la seguridad del paciente, así como la efectividad del sistema de salud en general.

A continuación, se presenta en la Tabla 4 los tipos de software clínico que ejemplifica funcionalidades clave de gestión médica.

Tabla 4
Tipos de software usados en gestión de información medica

Tipo de Software	Funcionalidades Principales	Beneficios Reportados
Sistemas de Información Hospitalaria (HIS)	Gestión integral de pacientes, administración de recursos, reportes financieros	Optimización del tiempo de atención, reducción de costos administrativos
Registro Médico Electrónico (RME)	Acceso a historial médico, prescripción electrónica, integración de imágenes diagnósticas	Mejora en la continuidad del cuidado, acceso remoto a datos clínicos
Sistemas de Apoyo a la Decisión Clínica	Alertas sobre interacciones de medicamentos, análisis predictivo basado en IA	Disminución de errores médicos, decisiones informadas

Nota. En la tabla se aprecia los tipos de software más utilizados en la gestión de información médica. **Fuente:** (Uquillas et al., 2023)

Estas herramientas no solo mejoran la productividad de los equipos de salud, sino que también permiten un manejo más seguro y eficiente de la información clínica, promoviendo la toma de decisiones basadas en datos, lo que es fundamental en entornos de alta presión como las unidades de cuidados intensivos y emergencias. Sin embargo, la implementación de estos sistemas

debe estar acompañada de capacitaciones adecuadas, asegurando que los profesionales puedan aprovechar al máximo sus beneficios.

2.3.2. Análisis Automatizado de Notas Clínicas

El análisis automatizado de notas clínicas es una de las aplicaciones más innovadoras de la automatización en salud, este proceso utiliza técnicas de procesamiento de lenguaje natural (PLN) y aprendizaje automático para extraer información relevante de los registros médicos no estructurados, facilitando la identificación de patrones y riesgos clínicos. Por ejemplo, los sistemas basados en inteligencia artificial pueden identificar condiciones críticas, como riesgos de sepsis o deterioro cognitivo, lo que permite una intervención temprana por parte de los equipos médicos. Un estudio realizado por Preciado et al. (2021) demostró que el uso de análisis automatizado permitió detectar indicadores tempranos de complicaciones en pacientes hospitalizados con una precisión del 87%, reduciendo significativamente la morbilidad asociada a diagnósticos tardíos.

Además, esta tecnología reduce la carga administrativa para los profesionales de la salud, liberando tiempo para que puedan concentrarse en tareas clínicas de mayor impacto, en entornos hospitalarios complejos, donde se generan grandes volúmenes de información diariamente, el análisis automatizado se ha convertido en una herramienta indispensable para optimizar el manejo de datos.

2.3.3. Tecnologías Emergentes en Salud Mental

Las tecnologías emergentes, como la inteligencia artificial y la realidad aumentada, están revolucionando el ámbito de la salud mental, proporcionando nuevas herramientas para la evaluación, el diagnóstico y la intervención terapéutica, plataformas basadas en IA, como Mindstrong y Lyra Health, analizan patrones de comportamiento y lenguaje en dispositivos móviles para ofrecer intervenciones personalizadas en tiempo real. Asimismo, la realidad virtual

se utiliza en tratamientos para trastornos como la ansiedad y el estrés postraumático, permitiendo a los pacientes enfrentar situaciones temidas en entornos controlados y seguros. Estas tecnologías no solo mejoran la accesibilidad a los servicios de salud mental, sino que también amplían las opciones terapéuticas disponibles, adaptándose a las necesidades individuales de cada paciente.

2.4.Arquitectura de Software

2.4.1. *Arquitectura en Capas para Sistemas Clínicos*

La arquitectura en capas es un modelo ampliamente utilizado en el desarrollo de sistemas clínicos, ya que organiza las aplicaciones en niveles funcionales independientes, cada uno con responsabilidades específicas, lo que facilita su mantenimiento y escalabilidad. Según Reyes & Soberanes (2021) se tiene lo siguiente:

Características clave de la arquitectura en capas:

- **Modularidad:** Cada capa opera de manera independiente, lo que facilita actualizaciones o cambios sin afectar otras partes del sistema.
- **Separación de responsabilidades:** Las funciones se dividen claramente entre las capas, lo que mejora la eficiencia del desarrollo y mantenimiento.
- **Adaptabilidad:** Permite la integración de nuevas tecnologías y funcionalidades sin necesidad de rediseñar todo el sistema.

En entornos clínicos, este modelo es particularmente relevante debido a:

- La capacidad para manejar información sensible con altos estándares de seguridad.
- Su eficiencia en la gestión de flujos de trabajo complejos que involucran a múltiples actores como médicos, enfermeras, administradores y pacientes.

La arquitectura en capas propicia la cooperación entre equipos de desarrollo especializados, como el equipo que se ocupa de la interfaz del usuario, el que se encarga de la

lógica de negocio y el de las bases de datos. Esto no solo permite disminuir los errores en la aplicación producidos por distintas capas, sino que también ayuda a aumentar la eficiencia del desarrollo. La posibilidad que ofrecen los componentes elaborados en una capa para ser utilizados en otros sistemas clínicos hace que se optimicen los recursos y disminuyan los costes, factor este muy importante para la sostenibilidad tecnológica en el sector salud.

2.4.2. Capa de Presentación: Interfaz de Usuario

La capa de presentación es responsable de la interacción entre el usuario y el sistema, en esta capa se diseña la interfaz de usuario (UI), que debe ser intuitiva, accesible y capaz de manejar grandes volúmenes de datos clínicos. Un ejemplo común son los portales de pacientes, donde los usuarios pueden acceder a sus historiales médicos, programar citas y consultar resultados de laboratorio. Según un informe de HIMSS (2021), los sistemas con interfaces optimizadas aumentan en un 25% la satisfacción del usuario, facilitando además una adopción más rápida por parte del personal médico. Esta capa también permite la personalización de las vistas según el perfil del usuario, ya sea un médico, un administrador o un paciente, asegurando que cada uno tenga acceso a las funcionalidades relevantes para su rol (Digi3n & Digi3n, 2020). En la tabla 5 se detalla varias de las características de la interfaz de usuario.

Tabla 5

Elementos de la interfaz de usuario

Elemento de la UI	Características Clave	Ejemplo de Uso en Sistemas Clínicos
Panel de Control	Visualización de datos en tiempo real	Seguimiento de indicadores de pacientes

Formularios Interactivos	Recolección de datos de manera dinámica	Registro de anamnesis y datos de ingreso
Notificaciones y Alertas	Alertas críticas para decisiones rápidas	Alertas de interacción de medicamentos

Nota. En la tabla se puede observar los principales elementos de la interfaz de usuario sus principales características. Adaptado de (Digi3n & Digi3n, 2020).

Esta capa no se encuentra enfocada en mejorar solamente la experiencia del usuario, sino que tambi3n permite contribuir a minimizar errores al presentar datos clave de forma clara y estructurada, optimizando el tiempo de respuesta en entornos cr3ticos.

2.4.3. Capa de Aplicaci3n: Procesamiento y An3lisis

La capa de aplicaci3n se encarga del procesamiento de datos y la ejecuci3n de la l3gica del negocio, conectando la interfaz de usuario con la capa de datos, esta capa maneja operaciones complejas, como el an3lisis de datos cl3nicos, la generaci3n de reportes personalizados y el soporte a la toma de decisiones cl3nicas. Los sistemas de apoyo a decisiones cl3nicas (CDSS) que se pueden implementar en esta capa, son capaces de analizar grandes vol3menes de datos y proporcionar recomendaciones cl3nicas basadas en la evidencia, lo que permite a los m3dicos seleccionar los tratamientos 3ptimos y reducir los errores que son comunes entre ellos. Seg3n el estudio de Reyes & Soberanes (2021), los CDSS han contribuido a disminuir en un 35% los errores de prescripci3n de medicamentos, ayudando de esta forma a mejorar significativamente la calidad del cuidado m3dico. Adem3s, esta capa permite la integraci3n con tecnolog3as emergentes, como algoritmos de aprendizaje autom3tico, que pueden identificar patrones en los datos para predecir complicaciones antes de que se manifiesten cl3nicamente.

2.4.4. Capa de Datos: Almacenamiento y Seguridad

La capa de datos es responsable del almacenamiento, recuperación y protección de la información clínica, utilizando bases de datos diseñadas para manejar información altamente confidencial, cumpliendo con estándares internacionales como el mencionado Reglamento General de Protección de Datos (GDPR) y la Ley de Portabilidad y Responsabilidad de Seguros de Salud (HIPAA). En esta capa se aplican técnicas sofisticadas de cifrado y autenticación, que velan por la obtención de datos o información por el personal debidamente autorizado. También se generarán copias de seguridad automatizadas y protocolos de recuperación ante desastres, gracias a los cuales se puede reducir el riesgo de la pérdida de información fundamental (Reyes y Soberanes, 2021). Un caso notorio es la utilización de las bases de datos relacionales para llevar historiales clínicos electrónicos, junto con un almacenamiento de datos para el análisis de extensos registros clínicos. A continuación, en la tabla 6 se detallan las características de la capa almacenamiento.

Tabla 6
Características de la capa de almacenamiento y seguridad

Tipo de Almacenamiento	Características	Ejemplo en el Ámbito Clínico	Beneficios Clave
Bases de Datos Relacionales	Estructura tabular para datos consistentes	Gestión de historiales médicos	Consistencia, facilidad de consultas
Almacenes de Datos	Análisis y procesamiento de grandes volúmenes	Predicción de tendencias epidemiológicas	Generación de informes, análisis en tiempo real

Almacenamiento en la Nube	Acceso remoto seguro y escalabilidad	Archivos de imágenes diagnósticas	Reducción de costos, accesibilidad global
Bases de Datos NoSQL	Manejo de datos no estructurados	Gestión de registros en formato multimedia	Flexibilidad en el manejo de grandes volúmenes

Nota. En la tabla se describe los diferentes tipos de almacenamiento de datos y seguridad de manera ejemplificada. Adaptado de (Bezanilla et al., 2018)

La capa de datos garantiza la integridad de la información en todo momento, ya que permite acceder a la información en tiempo real y facilita el trabajo colaborativo entre los equipos médicos localizados en distintas ubicaciones geográficas, lo cual es crucial en entornos de red hospitalaria. Del mismo modo, las tecnologías híbridas de información, como las bases de datos relacionales, integradas con bases de datos NoSQL, permiten gestionar datos estructurados y no estructurados, ya que, para ello, se adaptan a las necesidades cambiantes de los sistemas clínicos actuales.

2.5. Inteligencia Artificial

Hoy en día, la IA está presente en muchas áreas de nuestras vidas, gracias a su capacidad de aprender y mejorar con el tiempo, este se refleja en cómo va transformando muchos sectores como la salud, la educación, los negocios entre otros, lo que nos permite resolver inconvenientes de una manera más rápida y eficiente.

A medida que la tecnología avanza, la IA sigue evolucionando, integrando varias capacidades más y generando avances significativos en campos más complejos. Con el tiempo IA promete ser una herramienta poderosa para mejorar la vida y enfrentar los retos globales.

2.5.1. Inteligencia Artificial en Salud

La IA ha llegado a convertirse en una herramienta fundamental dentro del sector salud, desempeñando un papel clave en la mejora de diagnósticos, tratamientos y procesos administrativos, a través de algoritmos avanzados y aprendizaje automático, la IA es capaz de analizar grandes volúmenes de datos médicos, identificar patrones complejos y proporcionar recomendaciones precisas, contribuyendo significativamente a una toma de decisiones más informada (Aguilar, 2023). Por ejemplo, herramientas como IBM Watson Health han demostrado eficacia en la interpretación de imágenes diagnósticas y la personalización de tratamientos oncológicos, apoyando a los médicos en la reducción de errores y optimización de recursos disponibles, según el trabajo de Granda & Martínez (2023), su implementación puede incrementar la eficiencia operativa en un 30% y mejorar los resultados al anticipar complicaciones potenciales.

2.5.2. Modelos de Lenguaje Natural (NLP) para la Salud Mental

El procesamiento del lenguaje natural (NLP, por su nombre en inglés) se encuentra entre uno de los más importantes campos de aplicación de la inteligencia artificial en la salud mental, pues supone el hecho de poder analizar y clasificar la información textual de los registros clínicos y de las notas clínicas, permitiendo muy significativamente el diagnóstico temprano de trastornos emocionales (Lloret et al., 2021). Con la utilización de modelos de NLP, Google Health ha contado con la capacidad de analizar la depresión y la ansiedad en función de patrones en el lenguaje en los correos electrónicos e incluso en las propias transcripciones de las consultas, logrando un nivel de exactitud de un 85% según Romero (2023), lo que otorga un gran potencial para el diagnóstico y para el seguimiento terapéutico. Adicionalmente, estas herramientas tienen la capacidad de adaptarse a diferentes contextos culturales y lingüísticos, permitiendo una mayor inclusión en

sistemas de salud globalizados y reduciendo barreras de comunicación entre pacientes y profesionales.

2.5.3. *Análisis de Sentimientos y Comportamientos en Terapia*

El análisis de sentimientos, basado en algoritmos de inteligencia artificial, se utiliza para evaluar estados emocionales mediante el procesamiento del lenguaje escrito y hablado, su aplicación es evidente en plataformas de telemedicina y aplicaciones terapéuticas, donde se monitorea el progreso del paciente y se ajustan intervenciones de forma dinámica. Por otra parte, apps como Wysa, llevada a cabo esta técnica para detectar emociones como la tristeza o el estrés, generando recomendaciones personalizadas, señala que estas herramientas, demuestran un mejorado de la adherencia a la terapia de hasta un 40%, gracias a la percepción de acompañamiento permanente (Anchino et al., 2019). De la misma forma, el examen de comportamientos permite detectar patrones de crisis emocionales y ayudar a los terapeutas a desarrollar estrategias de prevención, esta estrategia no sólo optimiza los recursos en salud mental, sino que también mejora la precisión en la asistencia, pudiendo realizarse intervenciones personalizadas según las necesidades del paciente.

2.5.4. *Plataformas de IA para Psicoterapia*

Romero (2023) en su trabajo menciona que las plataformas de inteligencia artificial han ampliado significativamente las capacidades de la psicoterapia, proporcionando soluciones tecnológicas adaptadas a contextos diversos, estas herramientas son esenciales para facilitar el acceso a servicios de salud mental, especialmente en comunidades con pocos profesionales. Entre las plataformas más destacadas se encuentran OpenAI, Hugging Face y IBM Watson, en la tabla 7 se detalla cada una como se encuentra diseñada para aplicaciones específicas en psicoterapia:

Tabla 7

Aplicación de herramientas tecnológicas en Psicoterapia

Plataforma	Aplicación en Psicoterapia	Beneficios Clave	Limitaciones Consideradas
OpenAI	Creación de diálogos terapéuticos personalizados	Interacción dinámica y adaptable al contexto del usuario	Riesgo de respuestas fuera de contexto, dependiente de calibraciones frecuentes
Hugging Face	Modelos de NLP para la interpretación emocional	Análisis preciso de sentimientos y emociones	Necesidad de grandes conjuntos de datos para entrenamiento efectivo
IBM Watson	Diagnóstico y generación de recomendaciones	Integración con sistemas clínicos para decisiones basadas en datos	Altos costos operativos y requerimientos de personal especializado

Nota. En la tabla se mencionan algunas de las herramientas con más renombre en la actualidad y su funcionalidad. Fuente: (Li et al., 2021; Tena Cucala et al., 2019; Wolf et al., 2019)

Estas plataformas no solo optimizan recursos y tiempo, sino que también diversifican las estrategias terapéuticas, ofreciendo intervenciones accesibles y personalizadas, sin embargo, es necesario considerar los desafíos éticos y técnicos relacionados con la privacidad y la precisión en la interpretación de datos, aspectos que aún requieren regulación y supervisión.

2.6. Seguridad y Ética en el Uso de IA

La IA ofrece enormes beneficios, pero para poder implementarla requiere importantes desafíos en términos de seguridad y ética. Pues se debe garantizar la privacidad de los datos, evitar

inconsistencias en los algoritmos y promover el uso responsable de esta herramienta. Abordar estos desafíos es primordial para generar confianza minimizando riesgos y de esta manera maximizar el impacto positivo a la sociedad.

2.6.1. Protección de Datos y Privacidad en Salud Mental

El uso de la IA en los aspectos de la salud mental requiere una atención rigurosa a la protección de datos personales y la privacidad, considerando que esta información es particularmente sensible. En el contexto de Ecuador, la Ley Orgánica de Protección de Datos Personales (2021) establece las bases legales para garantizar la privacidad de los datos, exigiendo el consentimiento informado, la transparencia en el uso de la información y la implementación de medidas de seguridad adecuadas. Para entrenar modelos y dar soluciones personalizadas, las plataformas con IA utilizan grandes volúmenes de datos, lo que conlleva riesgos de filtraciones y usos indebidos (Asamblea Nacional, 2021). En el caso de Ecuador, debe asegurarse que los datos sensibles están almacenados bajo sistemas de encriptación muy robustos y que el acceso está restringido a las personas autorizadas. Por otro lado, las políticas de consentimiento informado tienen que describir de forma explícita cómo se recogen, almacenan y utilizan los datos por parte de los usuarios.

2.6.2. Normativas Legales para el Uso de IA en Salud

El desarrollo y la implementación de sistemas de IA en el ámbito de la salud deben regirse por un marco legal que garantice la seguridad de los pacientes y la transparencia en el uso de estas tecnologías. En Ecuador, la Ley Orgánica de Salud establece que el manejo de tecnologías en el sector salud debe priorizar la seguridad del paciente y el cumplimiento de estándares éticos. Adicionalmente, la Ley de Protección de Datos Personales regula el uso de información sensible, estableciendo obligaciones para los responsables del tratamiento de datos(ASAMBLEA

NACIONAL, 2015; Asamblea Nacional, 2021). Las leyes requieren a las entidades que lleven a cabo o hagan uso de IA que cumplan criterios específicos como la trazabilidad de los algoritmos y la validación clínica de los sistemas. No obstante, todavía existen vacíos legales respecto a la regulación específica de la IA, lo que muestra la necesidad de reforzar el marco regulatorio para poder preservar los derechos de los pacientes y permitir una mejora de la actividad tecnológica. En este sentido, el país puede beber de buenas prácticas internacionales, favorecer la colaboración entre los organismos públicos y privados y facilitar una implementación ética y segura.

2.6.3. Principios Éticos en la Automatización de Procesos Clínicos

La implementación de IA en los procesos clínicos debe regirse por principios éticos fundamentales que prioricen el bienestar de los pacientes, la equidad en el acceso y la transparencia en la toma de decisiones automatizadas. En Ecuador, la Constitución garantiza el derecho a la salud, lo que incluye la provisión de servicios tecnológicos seguros y éticos en el sector sanitario. Uno de los principales retos éticos a los que se enfrenta el uso de IA en el ámbito de la salud es la existencia del sesgo algorítmico; el cual puede causar que los resultados sean discriminatorios si los datos que se utilizan en el entrenamiento de los modelos no son representativos de la diversidad de la población (Asamblea Nacional, 2021). A continuidad de esto, la determinación de los resultados a partir de un sistema automatizado provoca que se planteen dudas en torno a la rendición de cuentas y a la responsabilidad que hay al cometer errores o causar daños. Para hacer frente a estas dificultades, será importante que las organizaciones proveedoras de atención sanitaria realicen auditorías continuadas y el sistema se encuentre diseñado de forma que se respeten los derechos humanos, así como la privacidad y la justicia social. En este marco, la Superintendencia de Protección de Datos Personales juega un papel crucial para supervisar el cumplimiento de estas normativas y principios.

Algunos de los principales aspectos éticos a considerar son:

- **Sesgo en los algoritmos:** Modelos que podrían discriminar por género, etnia u otros factores si no se utilizan datos diversos y pruebas de equidad en el diseño.
- **Toma de decisiones automatizadas:** Diagnósticos erróneos o fallas en la supervisión humana que requieren validación constante por parte de especialistas.
- **Transparente:** no existe un algoritmo que explique cómo se ha llegado a establecer dicho resultado; por tal razón debe trabajarse en la construcción de modelos explicables y auditables.

De la primera de las aseveraciones se puede extraer una síntesis "la seguridad y la ética en la utilización de la inteligencia artificial en el área de la salud constituyen bases fundamentales para poder garantizar su uso ordenado" y en un país como el Ecuador fortalecer el marco legal y ético para atender los riesgos que conllevan la privacidad, la equidad o la transparencia debe ser un prioritario. Aunque estas tecnologías ofrecen beneficios significativos en términos de eficiencia y precisión, su implementación debe equilibrar la innovación con el respeto a los derechos humanos y la garantía de la salud como un derecho constitucional.

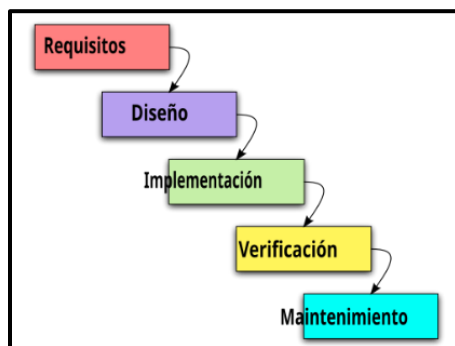
CAPÍTULO III. METODOLOGÍA

En el presente capítulo se encuentra detallado la metodología utilizada para el diseño e implementación del prototipo propuesto, en donde explícitamente se puede observar el análisis y establecimiento de los requerimientos necesarios y por último la selección de las herramientas tanto de hardware como software idóneas para que el sistema cumpla con la documentación, análisis y procesamiento de información de manera optimizada durante las sesiones de psicoterapia, todo esto siguiendo las directrices de las normas ISO/IEC/IEEE 29148:2018 para el establecimiento de los requerimientos y ISO/IEC 25010 para validar su funcionamiento dentro de los estándares internacionales.

3.1. Modelo en cascada

Como el objetivo de este proyecto es desarrollar un sistema que aplique inteligencia artificial para la conversión eficiente de contenido auditivo a texto, permita generar sugerencias y optimice los procesos de toma de decisiones cruciales en el área de psicología, el modelo en cascada es la mejor opción aplicable debido a que este enfoque promueve un proceso sistemático y secuencial en el desarrollo de software, caracterizado por su disciplina y basado en etapas de análisis, diseño, pruebas y mantenimiento para facilitar cumplir los requisitos de trazabilidad, control de calidad y predictibilidad para un sistema que maneja información clínica (Gamboa, 2018).

En las siguiente Figura 1 se puede observar las etapas necesarias que debe regirse el modelo descrito.

Figura 1*Modelo en Cascada*

Al ser un modelo idóneo por permitir establecer un proceso sistemático se estableció el cumplimiento de cuatro etapas, las cuales a continuación se describe las necesarias para el cumplimiento del capítulo presente.

3.2.Análisis

En esta etapa se puede mencionar que se permitió mediante la aplicación de preguntas de calidad y del estudio de trabajos relacionados establecer los requerimientos impuestos por los stakeholders y del sistema. Dentro de este contexto en la parte del sistema se contempló el análisis y la posterior selección de los componentes idóneos tanto de hardware como software que cumplan con los requisitos para que se lleve a cabo con el funcionamiento del prototipo sin dificultades.

3.2.1. Técnicas de investigación de recopilación de requisitos

En este apartado se puede describir que las técnicas de investigación que se utilizaron para el establecimiento de requerimientos fueron dos tipos. La primera la aplicación de un cuestionario hacia la psicóloga encargada con la finalidad de poder identificar la situación actual del sistema y establecer cuanto interés llegaría a tener si se lo mejora con una optimización con IA.

Y la segunda técnica se puede describir la revisión y análisis de trabajos de investigación relacionados para identificar que tipos de requerimientos se necesita específicamente tanto de

hardware como software para cumplir con las expectativas identificadas con el cuestionario aplicado hacia la psicóloga encargada.

3.2.1.1. Cuestionario aplicado a la psicóloga encargada

El cuestionario que se aplicó a la psicóloga encargada del departamento de bienestar estudiantil de la Universidad Técnica del Norte se basó en preguntas cerradas de calidad apoyadas de la norma ISO/IEC 25010 y preguntas para establecimiento de requerimientos del sistema apoyadas en la norma las normas ISO/IEC/IEEE 29148:2018 para la identificación del estado actual del sistema y como una mejora contribuiría a la optimización en las sesiones. En la siguiente Tabla 8 se observan las preguntas que se aplicaron.

Tabla 8

Preguntas cuestionario de calidad hacia la psicóloga encargada

Nro.	Pregunta
1	¿Cómo realiza la toma de notas en sus sesiones?
En el caso de responder solo manual	
2	¿Tomar notas manualmente durante la sesión interfiere con su atención clínica?
3	¿Interrumpe al paciente para anotar algo importante?
4	¿La necesidad de completar notas le obliga a extender la sesión, extenuando su jornada?
5	¿Las notas manuales le hacen perder información contextual relevante como gestos, tono o pausas en el paciente o señales no verbales?

Nro.	Pregunta
6	¿Con qué frecuencia olvida o resulta incomprendible la información en sus propias notas manuales cuando las revisa después?
7	¿Durante las sesiones con sus pacientes, cuando realiza la toma de notas a notado que se inhiben?
En el caso de responder solo transcripción de audio	
8	¿Si existiera la ausencia de transcripción de audio en sus sesiones podría limitar su capacidad para revisar con detalle lo ocurrido dentro de ellas?
9	¿Con el uso de las transcripciones es más fácil identificar cambios sutiles cuando el paciente se encuentra hablando en la sesión?
Almacenamiento de la información	
10	¿En su lugar de trabajo existe un sitio o repositorio digital en donde se almacene las notas o transcripciones de cada sesión?
11	¿Cree que con la inexistencia de un repositorio provoca la pérdida o duplicidad de información?
12	¿Al no presentar un sitio de almacenamiento único se le dificulta buscar información explícita de cada paciente para preparar una sesión de seguimiento acorde?

Nro.	Pregunta
13	¿Cuánto tiempo adicional dedica aproximadamente a buscar registros por no tener un repositorio único?
14	Considera prioritario disponer de un repositorio digital único y accesible para el seguimiento clínico
Establecimiento de interés	
15	¿Estaría dispuesta a utilizar un sistema que grabe y transcriba automáticamente las sesiones con IA para no incomodar al paciente y no perder de vista información contextual relevante como gestos, tonos o pausas en el paciente o señales no verbales?
16	¿Usted tiene alguna preocupación ética o de privacidad sobre el uso de una herramienta que grabe y analice sesiones?
17	¿Cree que el uso de inteligencia artificial en el nuevo sistema mejoraría la calidad de las sesiones al momento de almacenar, analizar y permitir sugerencias terapéuticas?
18	¿Confiaría totalmente en las sugerencias, análisis y toma de decisiones impuestas por el nuevo sistema con IA?

De acuerdo con las respuestas proporcionadas del Anexo A haciendo referencia a la contestación del cuestionario se pudo evidenciar que existe un perfil de práctica tradicional centrada en la toma de notas manual con impactos claros sobre la calidad de la atención desde las dos perspectivas, esto debido a que la psicóloga indica que la toma de notas interfiere con la atención clínica, la obliga a interrumpir al paciente conllevando a una extensión de su sesión. Además, reporta que las notas manuales hacen perder información contextual necesaria como los gestos, tonos, pausas y que con frecuencia las notas resultan incomprensibles al revisarlas. Respecto al paciente la psicóloga confirma que la presencia de toma de notas genera inhibición en el paciente, es decir que existe cohibimiento al expresarse y molestia cuando se está realizando la toma de información.

En cuanto al almacenamiento, la profesional confirma que no existe un repositorio digital único en su centro, logrando así evidenciar que esta ausencia genera pérdida o hasta duplicidad de información, logrando dificultar la búsqueda y preparación de seguimientos. De acuerdo con todo lo mencionado a pesar de tener presente la ética en cuanto a la privacidad, la Psicóloga muestra apertura al uso de un sistema automatizado con grabación y transcripción (IA) como herramienta para mejorar la calidad clínica y reducir carga administrativa, aunque confiaría solo parcialmente en las decisiones automáticas, es decir que las decisiones proporcionadas solamente serían una guía y ayuda como base para que en sí la Psicóloga realice un diagnóstico más estructurado, eficiente y completo.

3.2.1.2.Revisión de trabajos relacionados

Dentro de este aspecto se puede mencionar que se realizó la revisión de investigaciones relacionados con el presente tema del proyecto de grado, esto debido para permitir evidenciar, identificar o establecer requerimientos técnicos que cumplan con la norma ISO/IEC/IEEE

29148:2018 y poder así tanto el diseño como la implementación del sistema con IA se encuentre acorde cubriendo todo lo necesario.

Para lograrlo este punto de cada uno de los trabajos revisados se les realizó el establecimiento, que se extrajo como de los demás aspectos como autores, editoriales, etc. se puede observar en el ANEXO B que detalla las fichas bibliográficas de cada uno de los trabajos relacionados. De acuerdo con cada uno de los trabajos relacionados se pudo evidenciar algunos aspectos a considerarse para la identificación de requerimientos necesarios que debe cumplir el presente proyecto. Dentro de este contexto se puede mencionar que en el trabajo relacionado 1 se enfoca en el uso combinado de reconocimiento de voz y técnicas de NLP para evaluar automáticamente competencias y conductas terapéuticas a partir de sesiones grabadas, extracción de marcadores, anotación y métricas de evaluación, por lo que este trabajo investigativo se encuentra útil debido a que en cierta manera puede llegar a aportar criterios y métricas clínicas para diseñar el esquema de anotación, establecer estándares y elegir métricas de desempeño, además de ejemplos de características lingüísticas relevantes para análisis de comportamientos terapéuticos.

En el caso del trabajo relacionado 2 plantea un ASR incremental y adaptable para Edge devices basado en ideas tipo Whisper como el procesamiento por tramos, adaptación continua y optimizaciones para inferencia local. Este trabajo se encuentra útil debido a que proporciona técnicas prácticas para obtener transcripción en tiempo real, permitir adaptación al hablante y reducir dependencia de la nube, lo cual lo vuelve muy valioso si se desea realizar transcripciones inmediatas durante la sesión y protegerlas con privacidad.

Con respecto al trabajo relacionado 3 se hace referencia a la evaluación comparativa de los modelos ASR que funcionan offline en distintos dispositivos Edge, poniendo énfasis en latencia,

uso de CPU/RAM y rendimiento, esta información se encuentra relevante debido a que ayuda a estimar trade-offs realistas, seleccionar modelos y hardware para prototipado local, y definir requisitos no funcionales como tiempo de respuesta o recursos disponibles.

El trabajo relacionado 4 realizó una investigación de las interfaces de voz desplegadas en Edge adaptadas a usuarios con alteraciones del habla como técnicas de adaptación, robustez al habla atípica y diseño centrado en el usuario. Convirtiéndolo necesario ya que aporta métodos para hacer la captura y transcripción más robusta ante variabilidad del habla, lo que lo vuelve importante en contextos clínicos y recomendaciones de diseño accesible que aumentan la aceptación del paciente y la calidad de las transcripciones.

Por último, se encuentra el trabajo relacionado 5 que se centró en la descripción del diseño, implementación y evaluación práctica de un sistema de reconocimiento de voz en entornos con lenguas de pocos recursos y dispositivos Edge con capacidades limitadas, las cuales incluyen optimizaciones, pipeline y consideraciones de despliegue. Este trabajo permite ofrecer estrategias y técnicas para ejecutar ASR/servicios en dispositivos con recursos limitados, lo que lo vuelve valioso si se desea manejar localmente el audio por privacidad y para elegir hardware/software adecuados para dicha característica.

3.2.2. Requisitos del sistema

Dentro de esta sección se menciona la definición de los requerimientos necesarios, tanto de hardware y software que necesita dentro del sistema, esto se lo realizó basándose en la norma ISO/IEC/IEEE 29148:2018 que permite describir los procesos empleados en las actividades de ingeniería específicamente en el establecimiento de requisitos para un sistema o producto de software. De acuerdo con este aspecto en la siguiente Tabla 10 se presentan los acrónimos

utilizados para facilitar la comprensión y gestión de los términos relacionados con los requisitos tanto de hardware como de software.

Tabla 9

Descripción de acrónimos

Acrónimo	Descripción
StRS	Requerimientos de Stakeholder
SiRS	Requerimientos del Sistema
SHRS	Requerimientos de Hardware y Software

Una vez establecido los acrónimos de requisitos principales se realizó el establecimiento de los beneficiarios del sistema, dentro de estos se los puede diferenciar entre directos e indirectos, pero para mejor interpretación se los puede observar en la siguiente Tabla 10 en donde se encuentra la lista de personas y departamentos beneficiarias del presente proyecto:

Tabla 10

Beneficiarios del sistema

Beneficiarios	
Beneficiarios Directos	• Autor del presente trabajo de titulación: Alexander Moreno
	• Personal encargado del departamento de psicología del edificio de bienestar estudiantil Psc. Cristina Realpe
	• Personas encargadas de la dirección del edificio de bienestar universitario de la UTN

	<ul style="list-style-type: none"> • Empresa: Universidad Técnica del Norte • Director: Ing. Fabián Cuzme • Asesor: Ing. Carlos Vásquez
Beneficiarios Indirectos	<ul style="list-style-type: none"> • Comunidad de estudiantes de la Universidad Técnica del Norte que agenden cita en el área de psicología • Docentes de la Universidad Técnica del Norte • Comunidad profesional de psicología.

En donde para establecer correctamente los requerimientos que se necesita implementar dentro del presente proyecto se va a hacer uso de la siguiente Tabla 12 en donde se muestra los niveles de importancia que van desde alta, media y baja considerados para el diseño e implementación del proyecto. Para mejor visualización de estas se las puede encontrar en el ANEXO C.

Tabla 11

Establecimiento del nivel de Importancia

Nivel	Descripción
Alta	Referenciado a un requerimiento sumamente esencial que se debe considerar en el desarrollo del sistema
Media	Enfocado al requerimiento que puede ser omitido levemente, pero que al momento de hacerlo podría afectar o no la funcionalidad y rendimiento del sistema
Baja	Centrado al requerimiento que se puede omitir sin que exista la probabilidad que el sistema falle

3.2.2.1. Requerimientos de Stakeholders

Como parte inicial dentro del diseño de un sistema se debe tener considerado los requerimientos operacionales que deben presentarse al momento de que el usuario lo utiliza de acuerdo con la norma ISO/IEC/IEEE 29148:2018, dentro de estos requerimientos se puede describir el funcionamiento total del sistema para presentar una optimización en la toma de notas por medio de un prototipo electrónico y análisis de las mismas basadas en inteligencia artificial y su posterior almacenamiento como se lo puede observar en la Tabla 12.

Tabla 12

Requerimientos de Stakeholders

StRS				
Requerimientos Operacionales				
Número	Descripción	Nivel de		
		Importancia		
		Alta	Media	Baja
StRS1	Permitir la grabación de audio de cada sesión terapéutica de manera segura, garantizando la integridad, confidencialidad y protección de la información clínica durante todo el proceso de captura.	x		
StRS2	Generar la transcripción a texto automática para su almacenamiento y procesamiento	x		
StRS3	Generación automática de diagnóstico clínico		x	
StRS4	Repositorio digital único y centralizado para almacenamiento de notas, transcripciones y diagnósticos	x		
StRS5	Consentimiento del paciente a que la IA intervenga en cada sesión.	x		
StRS6	Integración de un sistema de autenticación de acceso para usuarios autorizados	x		
StRS7	Permitir la búsqueda y recuperación rápida de registros clínicos mediante criterios como ID del paciente y fecha de sesión	x		
StRS8	Presencia de una interfaz gráfica clara, intuitiva y de fácil uso		x	

StRS				
Requerimientos Operacionales				
Número	Descripción	Nivel de		
		Importancia		
		Alta	Media	Baja
StRS9	Mecanismos automáticos de respaldo y recuperación que eviten la pérdida de información clínica ante fallos del sistema.	x		
StRS10	Contar con la funcionalidad de eliminar de forma definitiva grabaciones y registros clínicos cuando el paciente lo solicite	x		
StRS11	Diseñado para operar de manera continua durante las jornadas clínicas sin interrupciones no planificadas			x

3.2.2.2.Requerimientos del sistema

Continuando con el establecimiento de requerimientos que se deben cumplir en base a la norma ISO/IEC/IEEE 29148:2018, ya que la función de cada uno de los elementos que lo conforman debe cumplir plenamente su funcionalidad sin que exista la probabilidad de falla.

Tabla 13*Requerimientos del Sistema*

SiRS				
Requerimientos del Sistema				
Número	Descripción	Nivel de		
		Importancia		
		Alta	Media	Baja
Requerimientos de Performance				
SiRS1	Almacenamiento en Google Drive de datos tanto de los resultados de análisis, diagnósticos generados como de las transcripciones de audio	x		
SiRS2	Generación del diagnóstico estructurado en un tiempo máximo de entre 5 y 7 minutos posteriores a la transcripción del audio		x	
SiRS3	Acceso continuo a la información clínica almacenada	x		
Requerimientos de interfaz				
SiRS4	Interfaz web adaptable a diferentes resoluciones de pantalla		x	
SiRS5	Permitir búsquedas ID del paciente y fecha dentro de los registros clínicos		x	
SiRS6	Visualización del texto generado de la transcripción del audio	x		
SiRS7	Visualización de los resultados del análisis y diagnóstico generados por la inteligencia artificial	x		

SiRS		
Requerimientos del Sistema		
Número	Descripción	Nivel de
		Importancia
		<u>Alta</u> <u>Media</u> <u>Baja</u>
Requerimientos de uso		
SiRS8	Ser fácil de comprender y operar por parte del personal clínico sin requerir conocimientos técnicos avanzados.	x
SiRS9	Portabilidad para usarse en diferentes ubicaciones sin afectar su funcionalidad	x
Requerimientos físicos		
SiRS10	Captura del audio de la sesión minimizando interferencias y ruidos propios del entorno clínico	x
SiRS11	Opciones de procesamiento local o en servidores controlados para proteger la privacidad de los datos	x
SiRS12	Operar con un dispositivo seleccionado de bajo consumo energético	x
SiRS13	Considerar condiciones adecuadas de ventilación y protección física del hardware seleccionado	x
Requerimientos de Seguridad		
SiRS14	Control de autenticación de acceso para usuarios establecidos	x

SiRS
Requerimientos del Sistema

Número	Descripción	Nivel de		
		Importancia		
		Alta	Media	Baja
SiRS15	Capacidad para eliminar grabaciones, transcripciones y diagnósticos de forma definitiva cuando sea requerido	x		

3.2.2.3.Requerimientos de Hardware y Software

En este apartado se puede mencionar que se detallan los requerimientos necesarios tanto para el Hardware como el Software del sistema en base al cumplimiento de la norma ISO/IEC/IEEE 29148:2018. De acuerdo con esto en la siguiente Tabla 14 en donde se los detalla y sus respectivas etiquetas de nivel de importancia.

Tabla 14
Requerimientos de Hardware y Software

SHRS		Nivel de		
		Importancia		
Número	Descripción	Alta	Media	Baja
Requerimientos lógicos				
SHRS1	Compatibilidad con tecnologías de inteligencia artificial (Whisper) y procesamiento de voz	x		
SHRS2	Compatibilidad con plataformas y entornos como Google Colab, Ngrok y Google drive	x		

SHRS		
Número	Descripción	Nivel de
		Importancia
		Alta Media Baja
SHRS3	Compatibilidad con lenguajes de programación de uso general	x
SHRS4	Permitir la integración mediante APIs para comunicación (API de DeepSeek, API de ngrok)	x
SHRS5	Ejecutarse sobre sistemas operativos basados en Linux (Rasbian OS).	x
Requerimientos de Hardware		
SHRS6	Ejecutarse sobre un ordenador de placa única SBC como unidad de procesamiento principal	x
SHRS7	Utilizar un micrófono que permita una captura clara de la voz	x
SHRS8	Disponer de almacenamiento suficiente de memoria para registros y archivos de audio	x
SHRS9	Contar con conexión a internet por medios inalámbricos o cableados	x
SHRS10	Operar con una fuente de alimentación eléctrica	x
SHRS11	Contar con memoria suficiente para ejecutar procesos de IA	x
SHRS12	Disponer de una estructura física que proteja los componentes como la unidad de procesamiento	x
Requerimientos Software		
SHRS13	Licencia de libre uso	x

SHRS		
Número	Descripción	Nivel de
		Importancia
		Alta Media Baja
SHRS14	Compatibilidad con librerías Python	x
SHRS15	Presencia de un IDE de programación	x
SHRS16	Uso de modelos preentrenados	x
SHRS17	Uso de bibliotecas y frameworks (Front-end)	x
SHRS18	Control de grabación, endpoints REST (Back-end)	x
SHRS19	Sistema de autenticación	x
SHRS20	Control de versiones por Git	x
SHRS21	Disponibilidad de almacenamiento local	x
SHRS22	Subida de información mediante HTTP	x
Requerimientos de IA preentrenada		
SHRS23	Precisión de detección de transcripción de audio a texto	x
SHRS24	Establecimiento de diagnósticos precisos	x
SHRS25	Tiempo de conversión de audio a texto < 10 minutos	x

3.2.2.4.Requerimientos de modelos

La definición adecuada del modelo de inteligencia artificial constituye un factor determinante para el desempeño del sistema, ya que condiciona directamente su capacidad para identificar patrones relevantes, procesar grandes volúmenes de información y generar resultados confiables. En este contexto, resulta indispensable establecer una especificación clara y estructurada de los requisitos del modelo, los cuales se detallan en la siguiente Tabla 15.

Tabla 15*Requerimientos de modelos*

MRS		
Requerimientos de modelos		
Número	Descripción	Nivel de
		Importancia
		Alta Media Baja
MRS1	Capacidad para transcribir audio a texto con alta precisión	x
MRS2	Reconocer variaciones de acento, tono y velocidad	x
MRS3	Funcionar correctamente en presencia de ruido ambiental moderado dentro del entorno clínico	x
MRS4	Identificar y segmentar intervenciones dentro del audio	x
MRS5	Alineación en cuanto a tiempo entre el audio y el texto transcrito	x
MRS6	Posibilidad de corrección y validación manual de las transcripciones	x
MRS7	Capacidad de resaltar patrones relevantes (pausas, énfasis, repeticiones)	x
MRS8	Integrarse con NLP para análisis semántico y resúmenes clínicos	x
MRS9	Garantizar privacidad de la información mediante procesamiento local o anonimización	x

MRS
Requerimientos de modelos

Número	Descripción	Nivel de Importancia		
		Alta	Media	Baja
MRS10	Exportación de resultados en formatos estructurados		x	

3.2.3. Selección de Hardware

Esta sección permite la elección del hardware específico mediante influencia de los requerimientos establecidos con anterioridad de usuarios como de sistema. Por lo que para lograr una elección acorde se realizó una comparativa de las mejores opciones para la elección del SBC y la selección del micrófono.

3.2.3.1.SBC

La placa SBC constituye el núcleo del prototipo ya que es la encargada de ejecutar el sistema operativo, gestionar la captura de audio, procesar la información mediante algoritmos de inteligencia artificial y garantizar el almacenamiento seguro de los datos clínicos. Por ello debe contar con capacidad de cómputo suficiente para ejecutar procesos concurrentes, soporte para un sistema operativo, interfaces de comunicación versátiles y compatibilidad con periféricos de audio y almacenamiento externo. Asimismo, es indispensable que presente un consumo energético moderado, estabilidad operativa y una amplia disponibilidad de herramientas de desarrollo, de modo que se asegure la viabilidad del prototipo, su escalabilidad y su mantenimiento en un entorno clínico real, de acuerdo con esto la selección se lo realiza a partir de la Tabla 16.

Tabla 16

Comparativa de los diferentes SBC

Requisitos	NVIDIA Jetson Nano	Raspberry Pi 4	Google Coral Dev Board
SHRS1	SI	SI	SI
SHRS2	SI	SI	SI
SHRS3	SI	SI	NO
SHRS4	SI	SI	SI
SHRS5	SI	SI	SI
SHRS6	SI	SI	SI
SHRS7	SI	SI	NO
SHRS8	NO	SI	NO
SHRS9	NO	SI	NO
SHRS10	SI	SI	SI
SHRS11	SI	SI	NO
Total, SI	8	11	6

Como se pudo observar la mejor opción es el **Raspberry Pi 4**, debido a que ofrece el mejor compromiso técnico entre capacidad de cómputo, E/S y facilidad de integración para un prototipo de captura y procesamiento de audio, dispone de características óptimas en el que combina un sistema operativo Linux completo, memoria suficiente, puertos USB 3.0 para almacenamiento cifrado rápido, amplia compatibilidad con micrófonos HAT/I²S y un ecosistema de hardware/software muy maduro que facilita integración de GPIO, HATs, bibliotecas Python. Si bien no dispone de NPU integrada permite añadir aceleradores externos como Edge TPU/Coral USB y además de permitir ejecutar modelos optimizados localmente, conservando un consumo moderado y un coste razonable, factores críticos para un prototipo clínico replicable.

3.2.3.2. Micrófono

El micrófono es un componente crítico del sistema, dado que la calidad de la señal de audio influye directamente en la precisión de la transcripción automática y en el posterior análisis inteligente de las sesiones terapéuticas. Por lo tanto, el dispositivo de captura ha de asegurar una vinculación sencilla con el SBC, una baja latencia y una calidad de la voz que sea correcta para entornos cerrados como los consultorios psicológicos. Asimismo, el micrófono ha de ser poco visible y no invasivo, siendo siempre necesario que su uso no produzca incomodidad o inhibición a las pacientes durante la sesión. Con estas características se garantiza que la captación de audio se realice de manera transparente a las pacientes, preservando la dinámica terapéutica y asegurando la información fiable para los modelos de inteligencia artificial.

Tabla 17

Comparación de diferentes micrófonos

Micrófono	SHRS7	SIRS9	SiRS10	SiRS12	Total
Mini micrófono USB SuziePi 3B+	SI	SI	SI	SI	4
ReSpeaker 4-Mic HAT	NO	SI	NO	NO	1
Rode NT-USB Mini	SI	SI	SI	NO	3

Para el diseño como implementación de un prototipo en donde se prioriza rapidez de despliegue, bajo coste y facilidad de uso, el **mini micrófono USB SuziePi 3B+** es la opción más práctica. Esto se lo puede mencionar debido a que se conecta directamente al Raspberry Pi 4, ofrece baja latencia y suficiente calidad para ASR ligero en condiciones controladas. Adicionalmente, el diseño compacto y discreto contribuye a preservar la dinámica terapéutica, al minimizar la distracción visual y evitar cualquier interacción física con el paciente, lo cual es coherente con los principios éticos y de comodidad del contexto psicológico.

3.2.4. Selección de software

En este apartado se puede mencionar que se hace la selección del software específico mediante una ponderación de los requerimientos presentados anteriormente. A continuación, se expone la selección del lenguaje de programación y el IDE de programación.

3.2.4.1. Lenguaje de programación

El lenguaje de programación seleccionado debe proporcionar las capacidades necesarias para implementar de manera eficiente y mantenible las funcionalidades del sistema, considerando las exigencias propias de un entorno clínico y el uso de inteligencia artificial. En este sentido, se hace la comparativa respectiva en la siguiente Tabla 18 de los diferentes lenguajes de programación más comunes para esta aplicación.

Tabla 18

Comparación de lenguajes de programación

Lenguaje	SHRS1	SHRS2	SHRS4	SHRS5	SHRS6	SHRS7	SHRS13	SHRS14	SHRS16	Total
Python	SI	SI	SI	SI	SI	SI	SI	SI	SI	9
Java	SI	SI	SI	NO	SI	NO	NO	SI	NO	5
C / C++	NO	SI	NO	NO	NO	NO	NO	NO	SI	2

La selección de **Python** como lenguaje de programación para el desarrollo del sistema se justifica por su alto grado de cumplimiento de los requerimientos, así como por su idoneidad técnica para aplicaciones que integran procesamiento de audio, inteligencia artificial y gestión de información clínica. Python es un lenguaje totalmente compatible con sistemas operativos Linux, permitiendo su correcta ejecución estable en una plataforma como la Raspberry Pi 4 y brindando un amplio ecosistema de bibliotecas especializadas en las áreas de reconocimiento automático de

voz, procesamiento de lenguaje natural y aprendizaje automático, que son fundamentales para la transcripción y el análisis inteligente de las sesiones terapéuticas.

Adicionalmente a esta funcionalidad, Python permite la rápida prototipación de los desarrollos a partir de su sintaxis denotativa y de la existencia de frameworks para la construcción de servicios web, APIs REST y manipulación de bases de datos, que acortan considerablemente los ciclos de implementación y de prueba. La existencia de una amplia comunidad y buena documentación asegura soporte continuo, actualización constante de las librerías y las facilidades de mantenimiento de cada aplicación, que son aspectos importantes para la escalabilidad del sistema. Python también ofrece herramientas maduras para la prueba, la depuración y el control de versiones, lo que puede contribuir a la fiabilidad del software y a la evolución futura del prototipo. Todas estas cualidades hacen de Python una de las opciones más adecuadas para la realización del proyecto, garantizando su viabilidad técnica, su eficiencia operativa y su sostenibilidad a largo plazo.

3.2.4.2.IDE de programación

El entorno de desarrollo integrado (IDE) debe proporcionar las herramientas necesarias para facilitar la implementación, depuración y mantenimiento del software del sistema, considerando la complejidad asociada al procesamiento de audio, la inteligencia artificial y la gestión de datos clínicos. En este sentido, el IDE debe ofrecer soporte nativo para el lenguaje de programación seleccionado, herramientas de depuración avanzadas, gestión de dependencias, integración con sistemas de control de versiones y facilidades para el desarrollo de aplicaciones distribuidas y servicios web. Asimismo, debe permitir una experiencia de desarrollo eficiente y ordenada, reduciendo errores durante la codificación y asegurando la calidad, escalabilidad y

mantenibilidad del sistema a lo largo de su ciclo de vida. De acuerdo con esto se hace la comparación en la siguiente Tabla 19.

Tabla 19

Comparación de IDEs de programación

IDE	SiRS3	SiRS4	SiRS5	SiRS6	SiRS7	SiRS14	SHRS17	SHRS18	SHRS19	SHRS21	SHRS22	Total
Visual Studio	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	SI	11
PyCharm	SI	SI	SI	SI	NO	SI	NO	SI	SI	NO	NO	6
Eclipse	NO	SI	NO	NO	SI	NO	NO	SI	NO	SI	NO	4

La selección de **Visual Studio Code** como entorno de desarrollo integrado se justifica porque cumple de manera integral los requerimientos SiRS y SHRS definidos para el proyecto, proporcionando un entorno ligero, flexible y altamente extensible que facilita el desarrollo del sistema basado en Python. Visual Studio Code ofrece soporte nativo para control de versiones, depuración avanzada, gestión de entornos virtuales y desarrollo de APIs, además de compatibilidad con librerías y frameworks de inteligencia artificial. Su naturaleza multiplataforma, bajo consumo de recursos y amplia comunidad de desarrolladores garantizan un proceso de desarrollo eficiente, mantenible y escalable, lo que resulta especialmente adecuado para la implementación y evolución del prototipo en un contexto académico y clínico.

3.2.5. Selección del modelo IA Generativa

En este caso no se hace una selección del Dataset debido a que se tiene pensando usar un modelo preentrenado para mitigar el tiempo de entrenamiento y consecuentemente obtener una mayor confiabilidad en cuanto a la realización de diagnósticos, de acuerdo con esto en la siguiente Tabla 20 se hace la comparativa de los diferentes modelos IA que se encuentre presentes en la

actualidad y que tienen mayor afluencia en cuanto a su uso, ya que el modelo de inteligencia artificial seleccionado debe ser capaz de procesar y comprender lenguaje natural de forma precisa, contextual y coherente, dado que su función principal es analizar transcripciones de sesiones terapéuticas y apoyar la generación de información relevante para el profesional clínico. En este sentido, el modelo ha de integrarse adecuadamente con la arquitectura del sistema, favorecer su uso mediante interfaces programáticas, y proporcionar acuerdos con procesos de análisis semántico, de generación de resúmenes o de apoyo a la toma de decisiones. También, dado lo sensible de la información del ámbito clínico, el modelo ha de proporcionar control sobre los datos, opciones de despliegue local o privado, y flexibilidad de cara a su adaptación o personalización al dominio terapéutico, garantizando fiabilidad, seguridad y alineación al criterio profesional humano.

Tabla 20

Comparación de los modelos IA

Modelo IA	SiRS2	SHRS2	SHRS23	SHRS24	SHRS25	MRS1	MRS2	MRS8	Total
DeepSeek	SI	SI	SI	SI	SI	SI	SI	SI	8
OpenAI (GPT)	SI	SI	SI	NO	NO	SI	SI	SI	6
Gemini (Google)	SI	SI	NO	NO	NO	SI	SI	SI	5
Grok	SI	NO	NO	NO	NO	SI	NO	NO	2

La selección de **DeepSeek** como modelo de inteligencia artificial se justifica porque destaca por su capacidad avanzada de comprensión y generación de lenguaje natural, su facilidad de integración con el sistema mediante interfaces programáticas y especialmente por permitir

mayor control sobre los datos y opciones de despliegue local, aspecto crítico en un entorno clínico donde la privacidad y la confidencialidad son prioritarias. Además, su flexibilidad para ser adaptado al dominio terapéutico y su compatibilidad con otros componentes del sistema permiten que el modelo actúe como una herramienta de apoyo al profesional, sin sustituir su criterio clínico. Estas características convierten a DeepSeek en la opción más adecuada para el desarrollo del prototipo, asegurando viabilidad técnica, seguridad de la información y coherencia con los objetivos del proyecto.

3.3.Diseño

Esta etapa se alinea con el objetivo específico dos y parte del tres, en el que se definirán los diferentes procesos que se llevarán a cabo en el sistema, por lo que se formula una solución en la toma de notas que el profesional realiza.

3.3.1. *Software*

3.3.1.1. Arquitectura de Software. Teniendo en cuenta que realizara la grabación de las sesiones para a posterior realizar el proceso de tratamiento de audios para su respectiva conversión a texto y en base a eso identificar patrones generando sugerencias en las sesiones de terapia. Para lo cual se busca optimizar este proceso mediante un diseño de una interfaz de software.

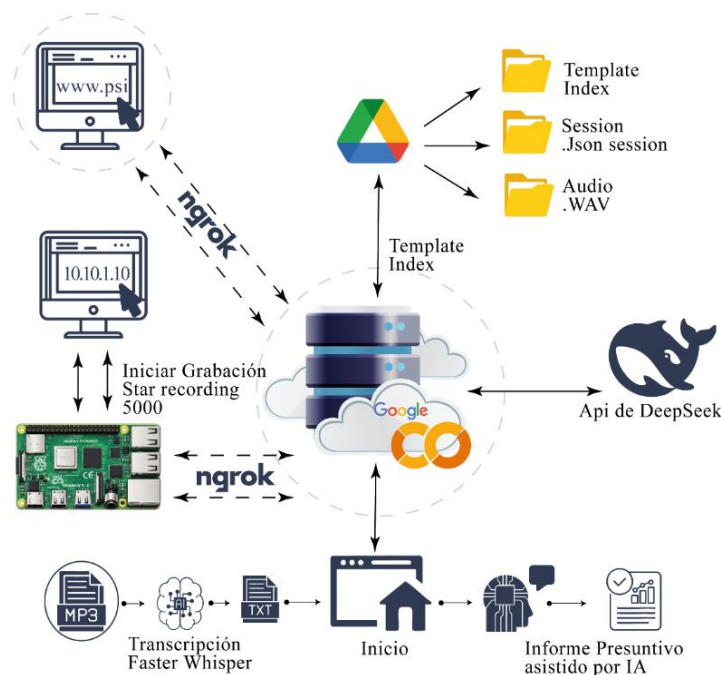
Además de generar un respaldo del almacenamiento de archivos en el cual se aplicará un modelo de aprendizaje de inteligencia artificial preentrenado, de esta manera como resultado final se pueda obtener un documento donde se contenga registros cronológicos de actividades, decisiones clave, eventos relevantes y cualquier otra información importante que facilite el

seguimiento del progreso de los pacientes de esta manera sirva como referencia histórica para otorgar sugerencias en futuras evaluaciones y mejoras para el profesional a cargo.

Con referencia en la siguiente Figura 2 se presenta la especificación de la arquitectura y el modelo de tres capas propuestos para la implementación en este proyecto.

Figura 2

Arquitectura de software



Nota. La figura muestra la estructura del sistema funcional a desarrollar. Fuente: Autoría

En la programación por capas, los objetos se dividen según su función en tres categorías principales: la Capa de Interfaz, que interactúa con el usuario; la Capa de Lógica de Negocio, que maneja el procesamiento interno y actúa como enlace entre capas; y la Capa de Datos, que se comunica con bases de datos y otros sistemas. (Del Valle & Granados, 2007)

Dentro de nuestro proyecto de investigación se puede mencionar que la capa de presentación es la encargada de la interacción directa con el usuario, en este caso el profesional de

la salud mental, en donde su función principal es proporcionar una interfaz gráfica intuitiva que permita operar el sistema sin interferir en la dinámica terapéutica.

En esta capa se implementan los elementos visuales de la aplicación web, tales como formularios de autenticación, módulos para la carga o grabación de audio, paneles de visualización de transcripciones y secciones para la generación y revisión de informes clínicos. También esta capa se encarga de los eventos del usuario (por ejemplo, solicitar archivos, empezar/parar las grabaciones y solicitar los análisis), validando las entradas sencillas antes de enviar al backend.

La capa de presentación no lleva lógica clínica ni procesamiento inteligente, se limita a ser un intermediario del usuario con la capa de aplicación, enviando entradas mediante HTTP, y recibiendo respuestas estructuradas (normalmente en JSON), las cuales son presentadas de una forma comprensible al profesional; en cambio, la capa de aplicación es el núcleo funcional del sistema porque concentrará la lógica de negocio y coordinará los procesos inteligentes del prototipo. La capa de aplicación es un backend basado en servicios web que recibe las entradas provenientes de la capa de presentación y ejecuta la operación correspondiente. Las principales funciones de esta capa son la gestión de sesiones, la recepción y validación de los archivos de audio, la ejecución del proceso de transcripción automática mediante modelos de reconocimiento de voz, y el análisis del texto resultante mediante modelos de lenguaje para la generación de informes clínicos presuntivos. Esta capa también aplica reglas de control, como la validación de la autenticación, la validación del consentimiento informado o la activación secuencial de los módulos del sistema.

Por otra parte, la capa de aplicación también orquesta los modelos de inteligencia artificial, gestionando la comunicación con los servicios ASR y NLP, estructurando los prompts clínicos y formateando los resultados antes que sean enviados a la capa de presentación o almacenados en la

capa de datos. De esta manera, se mantiene un desacople entre la lógica del sistema y la interfaz de usuario.

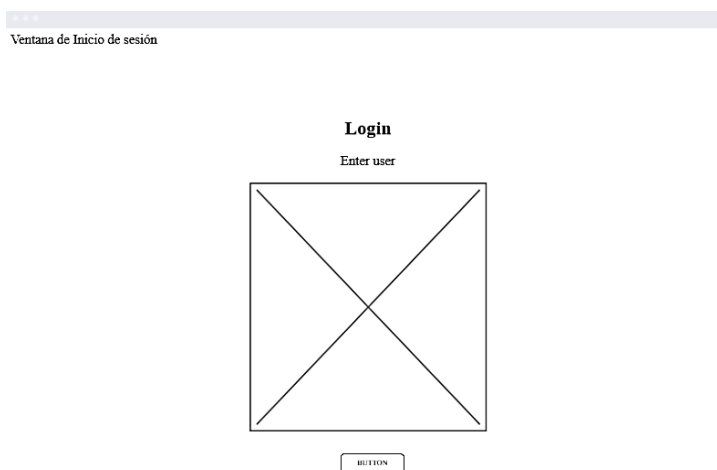
Finalmente, con respecto a la capa de datos es responsable de la persistencia, organización y protección de la información clínica generada por el sistema. En esta capa se almacenan los audios de las sesiones, las transcripciones, los informes clínicos generados, así como los registros asociados a pacientes y sesiones terapéuticas.

Durante la fase de prototipado, la gestión de datos se implementa mediante almacenamiento estructurado en archivos (por ejemplo, en formato JSON) y repositorios externos como Google Drive, lo que permite mantener trazabilidad, respaldo y acceso controlado a la información. La capa también tiene que implementar unos mecanismos básicos de seguridad, como el control de acceso, la separación de las carpetas por paciente y por sesión y la preparación para una futura implementación de cifrado de los datos. La capa de datos no tiene interacción directa con el usuario, si no que sólo se puede acceder por la capa de aplicación. Es a través de esta capa de aplicación que el sistema garantiza la integridad y la coherencia de la información clínica. Esta separación permite la migración futura a bases de datos más potentes o a infraestructuras clínicas de especialización, sin que estos cambios afecten al funcionamiento del resto del sistema.

a) *Frontend.* Corresponde a la capa de presentación del sistema y constituye el punto de interacción directa entre el profesional y la plataforma. Desde este componente se gestionan los procesos de autenticación, en este punto se puede mencionar que se encuentra diseñado a partir de una plantilla base que inicia con el módulo de login del sitio web, el cual constituye el primer punto de control del sistema. En esta pantalla se valida que el usuario sea un profesional autorizado antes de habilitar cualquier acción relacionada con manejo de información clínica como se lo puede observar en la siguiente Figura 3.

Figura 3

Diseño de acceso al sistema y autenticación del profesional



Este diseño responde a una regla operativa fundamental, si la autenticación falla por ende el sistema bloquea el acceso a los módulos de carga, grabación y procesamiento, evitando exposición de datos sensibles o ejecución de procesos no autorizados. En consecuencia, el login no se considera únicamente una interfaz inicial, sino un mecanismo de gobernanza funcional y seguridad, ya que define el acceso al flujo clínico completo y establece el estado base desde el cual se habilitan los componentes restantes de la plataforma.

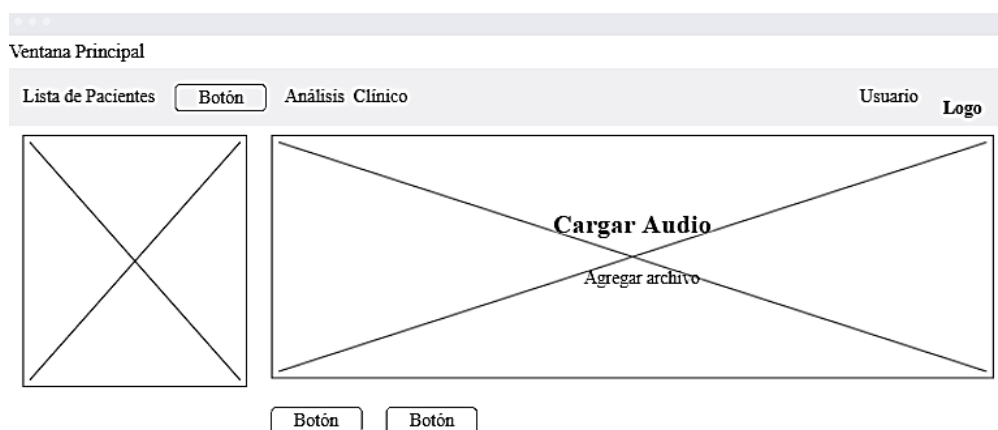
Una vez superado el proceso de autenticación el profesional accede a la interfaz principal del sistema, donde se habilita un módulo lateral de administración clínica destinado a la gestión estructurada de pacientes y sus respectivas sesiones. Este módulo constituye el punto central para la organización, consulta y seguimiento del historial terapéutico. La información de los pacientes se obtiene dinámicamente desde el backend mediante la solicitud GET /patients, lo que garantiza que los datos presentados en la interfaz correspondan al repositorio actualizado del sistema. La información de cada paciente es dispuesta en un historial de sesiones organizado con una estructura

jerárquica de tipo árbol que permite disponer cronológica y ordenadamente las atenciones registradas.

De la misma manera, una vez que un profesional selecciona una sesión almacenada, el sistema activa el modo de visualización de solo lectura en la que la transcripción generada y el análisis presuntivo se cargarán automáticamente. De este modo será imposible combinar (revisar el historial y procesar un audio nuevo) las instrucciones de procesamiento y de generación y de generar nuevos informes. Esto se controla a través de una variable estado (`isViewingSavedSession`) que permite al Frontend encausar el contexto de trabajo. Siguiendo las reglas básicas del diseño, podemos decir que el sistema establece explícitamente una separación entre dos diferentes contextos-operativos, revisar historial y procesar un audio nuevo como lo podemos ver en la siguiente Figura 4. Esto permite en definitiva incorporar integridad documental, trazabilidad clínica y clarificar el procedimiento del profesional en el manejo trabajando con información sensible evitando incoherencias.

Figura 4

Diseño de gestión del paciente y selección de sesión



Dentro de este apartado se presenta la etapa de captura de audio, el cual el sistema se encuentra diseñado bajo un enfoque de doble mecanismo (Modo A y Modo B) de entrada

diferenciados por su origen, pero unificados en su procesamiento. Aunque los audios pueden provenir de fuentes distintas, ambos flujos convergen intencionalmente en el mismo endpoint /upload del backend en Colab, garantizando un pipeline único, estandarizado y coherente para las etapas posteriores de almacenamiento, transcripción y análisis con inteligencia artificial.

En el Modo A la profesional carga un archivo de audio de tipo WAV, MP3 u OGG, ya existente, y lo realiza directamente desde el navegador (FrontEnd). El Frontend recubre el archivo en un objeto FormData y lo envía mediante una llamada de tipo POST para API_URL/upload. Una vez realizado el envío, el sistema activa el reproductor WaveSurfer para permitir que la profesional pueda reproducirlo y comprobar que se trata, efectivamente, del archivo que se correspondía con la sesión clínica de que se trataba. Este paso previo para validar el archivo reduce posibles errores en cuanto a la carga y refuerza el propio proceso.

En el Modo B el audio no proviene del repositorio local, sino que se continúa generando mediante una captura realizada en el presente que se ejecuta en la RaspberryPi. Esta operación se realiza desde la misma interfaz web, donde el Frontend envoca la invocación de los endpoints específicos de la Raspberry (/start_recording, /pause_recording, /resume_recording y /stop_recording) para acceder a la grabación realizada en la propia herramienta arecord. Una vez la captura ha sido finalizada, la Raspberry devuelve el archivo como blob, el flujo binario audio/wav que tiene como el nombre real del archivo en la cabecera X-Audio-Filename.

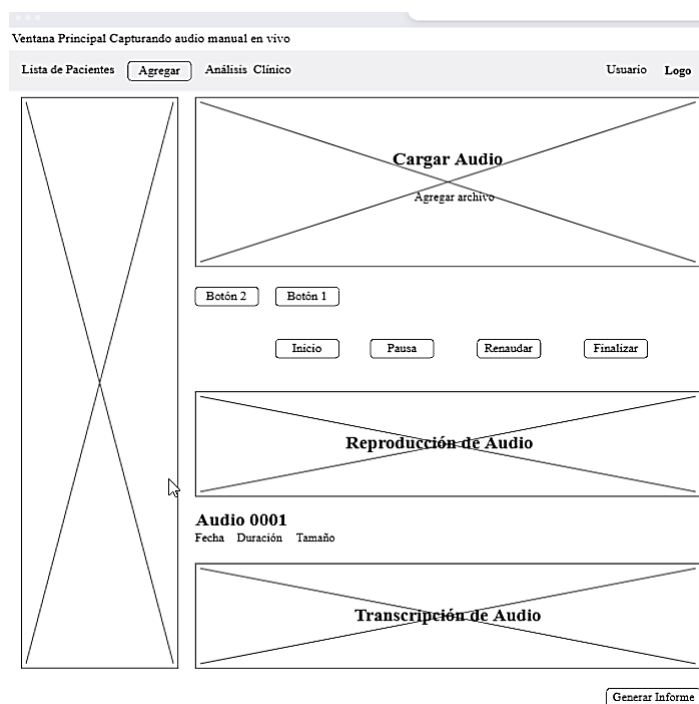
Posteriormente, en lo que concierne dentro del Frontend, es que permite transformar este blob en un objeto File, permitiendo que se lo inserte de manera programática en el control <input type="file"> y sin tener la necesidad de modificar la lógica establecida, se reutiliza exactamente el mismo flujo del Modo A enviándolo al endpoint API_URL/upload. En otras palabras, esta estrategia de convergencia permite que el sistema trate cualquier tipo de formato de audio ya sea

cargado manualmente o grabado en tiempo real como una única entrada estandarizada para el backend. (Ver Figura 5)

Desde el punto de vista arquitectónico, esta decisión reduce complejidad, evita duplicidad de lógica, facilita mantenimiento y garantiza consistencia en el procesamiento, asegurando que la etapa de almacenamiento y análisis con inteligencia artificial opere bajo un esquema uniforme independientemente del origen del audio.

Figura 5

Diseño de la etapa de captura de audio de las sesiones de los pacientes



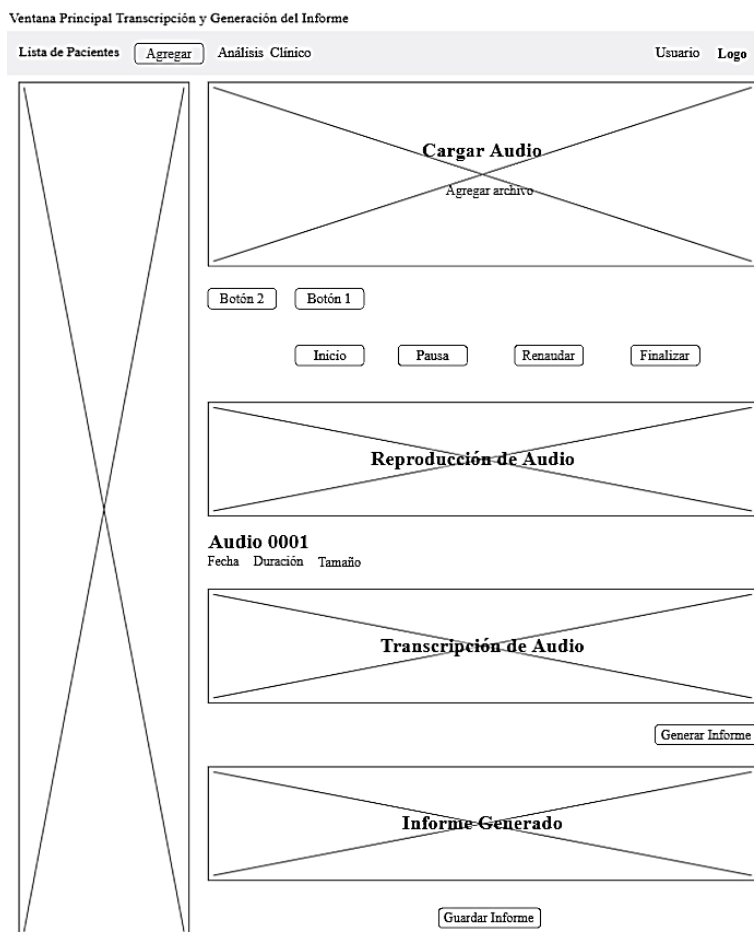
Posterior a la etapa de captura de audio y una vez que el backend desplegado en Colab recibe el archivo independientemente de si su origen corresponde a una carga manual desde el equipo del profesional o a una grabación en tiempo real desde la Raspberry Pi se activa un pipeline de procesamiento secuencial estructurado por módulos. Este pipeline fue diseñado bajo principios

de persistencia, trazabilidad y separación clara de responsabilidades, garantizando un flujo ordenado y controlado de la información clínica.

En la primera fase, el sistema ejecuta el almacenamiento persistente del audio en Google Drive. Esta decisión arquitectónica asegura que cada sesión quede registrada de manera permanente, incluso ante reinicios del entorno de ejecución. Como regla de diseño cada archivo se guarda con un identificador único asociado a la sesión y al paciente, manteniendo un orden temporal que facilita la localización posterior por fecha, historial clínico o proceso de auditoría como se lo puede observar en la siguiente Figura 6.

Figura 6

Diseño de guardado de archivos provenientes de las sesiones de los pacientes



Este mecanismo garantiza la integridad del documento y permite reusar el audio para futuras revisiones sin necesidad de tener que volver a realizar la captura. Prueba de la persistencia del archivo, se inicializa el módulo de transcripción automática (ASR) de modelo faster-whisper, el cual realiza el proceso de conversión del contenido acústico (el audio) a texto estructurado, produciendo en forma de salida una transcripción que el backend devuelve al Frontend en formato JSON con la estructura {"transcriptionModel": "..."}.

Este resultado es presentado de forma inmediata en la interfaz, permitiendo al profesional visualizar el contenido textual de la sesión en tiempo real y preparar su revisión posterior. En la fase final del pipeline, la transcripción generada se utiliza como insumo para el módulo de análisis clínico presuntivo mediante un modelo de lenguaje (LLM). El Frontend envía el texto al endpoint POST /diagnostic bajo la estructura {"content": "..."}.

El modelo, como hemos indicado, procesa el contenido y emite el resultado en forma de un informe estructurado en el que podemos encontrar la identificación de patrones emocionales que aparecen como predominantes, las distorsiones cognitivas, las conductas observables, las señales de alerta o riesgo, así como un breve resumen interpretativo en función de criterios y elementos de soporte terapéutico. No obstante, el diseño del sistema cuenta con una limitación crucial, en el que el resultado que proporciona el LLM tiene caracteres indicativos y de soporte analítico, en ningún caso se trata de un resultado que exprese un diagnóstico.

La validación, la interpretación del contexto y la toma de decisiones clínicas quedan de forma explícita como responsabilidad del profesional, manteniendo el principio de supervisión humana obligatoria en el ejercicio de la manera terapéutica.

Una vez finalizada la fase de procesamiento el procedimiento de inteligencia artificial, y sólo cuando el sistema tiene tanto el texto de la transcripción automática como el análisis

presuntivo elaborado de acuerdo con el modelo de lenguaje, se introduce la fase de la sentencia consolidada y el registro definitivo de la sesión. En este punto, la interfaz activa la acción “Add current Session”, la cual representa el cierre operativo del flujo técnico iniciado con la captura de audio.

Durante esta fase, el diseño arquitectónico concentra y normaliza la información relevante en un único objeto estructurado. Este registro integra de manera organizada:

- Identificación del paciente seleccionado.
- Fecha de la atención.
- Transcripción obtenida mediante el módulo ASR.
- Análisis presuntivo (advice) generado por el LLM.
- Metadatos técnicos asociados al archivo de audio, tales como nombre, referencia interna y tamaño del archivo.

La inclusión de estos metadatos no es incidental, sino que responde a criterios de trazabilidad, auditoría y control documental, permitiendo que cada sesión quede respaldada con evidencia técnica verificable. Una vez estructurada esta información, el Frontend envía el objeto consolidado al backend para su almacenamiento bajo una estructura tipo JSON dentro del historial del paciente. De este modo, la sesión se registra de forma perdurable y va incorporada para ser consultada como parte del repositorio clínico estructurado. Dicha forma de desarrollar el flujo termina con una fase deliberadamente explícita de revisión profesional y decisión clínica; esta etapa coincide con el propio diseño del sistema y constituye el mismo recorrido que el prototipo no pretende automatizar, una automatización de la propia responsabilidad diagnóstica; su propósito radica en la organización de la información para mejorar la elaboración del trabajo por parte de este profesional. En esta última fase, el terapeuta revisa la transcripción y la revisión

cortesía de la inteligencia artificial, y a partir de su criterio decide qué elementos incorporar como hallazgos válidos y qué elementos descartar por no ser pertinentes desde la perspectiva contextual o desde la perspectiva clínica. Una norma transversal del diseño a nivel técnico como a nivel ético es que el sistema no sustituye al profesional, sino que interviene como herramienta de apoyo para facilitar la documentación, mejorar la organización de la información y presentar una primera revisión de esta información desde un análisis estructurado; finalmente, en todo momento la validación, la interpretación final y la decisión clínica son responsabilidad humana.

b) **Backend.** Dentro de este aspecto se puede mencionar que el Backend concentra la lógica funcional y los servicios del sistema. En este proyecto el backend no es monolítico, sino que se encuentra distribuido en dos entornos diferenciados, uno destinado a la captura controlada del audio en la Raspberry Pi y otro enfocado en el procesamiento inteligente en Google Colab (almacenamiento, transcripción y análisis). Esta división permite desacoplar la adquisición física de datos del análisis basado en modelos de inteligencia artificial, optimizando el uso de recursos computacionales y facilitando la escalabilidad del sistema. Además de permitirse establecer la regla de diseño en el que la captura de audio debe estar cerca del hardware (Raspberry), mientras que el procesamiento IA debe ejecutarse donde existe mayor capacidad computacional (Colab).

- **Backend de Captura (Raspberry Pi – Flask + arecord).** El backend de captura se plantea como un servicio especializado cuyo propósito exclusivo es la generación controlada del audio clínico en tiempo real. Su diseño responde a un principio de separación de responsabilidades: este módulo no realiza procesamiento inteligente ni análisis semántico, sino que se limita a capturar, almacenar y entregar el audio de forma confiable al sistema principal.

Arquitectónicamente se implementa mediante una API REST desarrollada en Flask (archivo file.py), que actúa como interfaz de control entre el Frontend y el sistema operativo de la

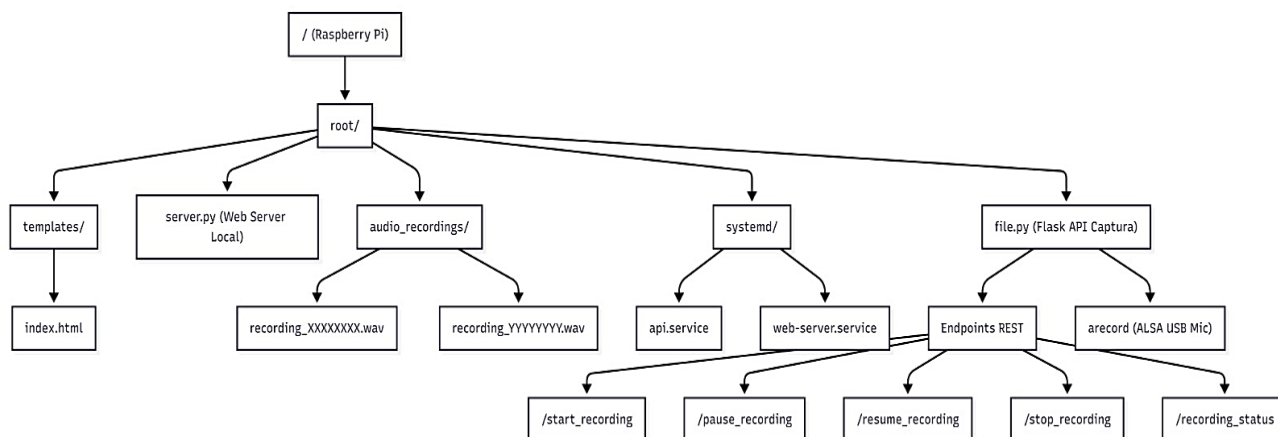
Raspberry Pi. Esta API encapsula el uso del módulo de captura ALSA (arecord), ejecutándolo como un subproceso del sistema. De esta forma, el navegador no interactúa directamente con el hardware, sino que lo hace a través de una capa intermedia controlada. El diseño contempla cinco endpoints principales:

- `/start_recording`: genera un nombre único basado en timestamp y lanza el proceso `arecord`.
- `/pause_recording`: envía señal `SIGSTOP` para pausar la grabación sin perder continuidad.
- `/resume_recording`: envía señal `SIGCONT` para reanudar.
- `/stop_recording`: termina el proceso, valida la existencia del archivo y lo retorna como binario `audio/wav`.
- `/recording_status`: consulta el estado actual del proceso.

También se proporciona un mecanismo de control de concurrencia mediante `threading.Lock()`, de tal forma que no pudieran existir varias grabaciones activas a la vez. Esta elección de diseño ha de evitar la posible corrupción de los ficheros, conflictos de dispositivo y estados inconsistentes del proceso. El almacenamiento local se lleva a cabo en el directorio `/root/audio_recordings`, con nomenclatura por marca temporal para garantizar así la trazabilidad. Finalmente, el fichero se devuelve al Frontend junto al header `X-Audio-Filename`, de forma que el Frontend pueda recuperar el nombre real del objeto. Desde el punto de vista arquitectónico, este backend puede considerarse un microservicio de adquisición de datos, cuyo único rol es producir el insumo primario del sistema: el audio clínico como se lo puede observar en la siguiente Figura 7.

Figura 7

Backend de Captura (Raspberry Pi – Flask + arecord)



- **Backend de Procesamiento (Colab – Flask + Drive + modelos IA).** El backend de procesamiento constituye el núcleo lógico e inteligente del sistema, ya que concentra las funciones relacionadas con persistencia de información, transcripción automática y análisis semántico asistido por inteligencia artificial. Su diseño se basa en una arquitectura de tipo pipeline, donde cada etapa tiene definida una responsabilidad claramente delimitada, asegurando orden, desacoplamiento funcional y trazabilidad del flujo de datos.

El backend se ha implementado mediante Flask y se ha expuesto al exterior mediante un túnel seguro a través de ngrok para poder consumirlo desde el Frontend mediante la variable `API_URL`. Este sistema hace posible la comunicación remota entre la interfaz web y el entorno de procesamiento, asegurando la separación entre la capa de presentación y la capa lógica del sistema. El diseño contempla los siguientes módulos principales:

a. Módulo de Almacenamiento (Google Drive)

Una vez que el audio es recibido a través del endpoint `/upload`, el sistema ejecuta inmediatamente su almacenamiento en Google Drive. Esta decisión arquitectónica responde a tres criterios fundamentales:

- Persistencia, evitando pérdida de información ante reinicios del entorno Colab.
- Trazabilidad histórica, permitiendo conservar evidencia estructurada de cada sesión.
- Separación entre procesamiento y almacenamiento, reduciendo acoplamiento y mejorando mantenibilidad.

Cada archivo se guarda con un identificador único asociado a la sesión, lo que permite su posterior vinculación con el paciente correspondiente y facilita búsquedas por fecha o historial clínico.

b. Módulo ASR (Automatic Speech Recognition)

Una vez asegurada la persistencia del archivo, el backend ejecuta el modelo de reconocimiento automático de voz faster-whisper, encargado de transformar el contenido acústico en texto estructurado. En esta etapa, se tiene la primera capa de inteligencia del sistema, ya que transforma la información no estructurada (audio) en información textual que puede ser tratada.

La transcripción se devuelve al Frontend en forma de JSON de modo que pueda visualizarse sin problemas en la interfaz. Desde un punto de vista arquitectónico, esta etapa se ha diseñado explícitamente para ser desacoplada del análisis clínico, ya que la misma información puede, a su vez, ser utilizada para la revisión manual, auditoría, procesamiento u otros usos.

c. Módulo de Análisis Clínico Presuntivo (LLM)

Mediante el endpoint /diagnostic, el sistema recibe la transcripción generada y la envía a un modelo de lenguaje de gran escala (LLM). Este componente realiza un análisis semántico avanzado y genera un informe estructurado que puede incluir:

- Identificación de emociones predominantes.
- Detección de distorsiones cognitivas.
- Reconocimiento de conductas observables.

- Señales o indicadores de riesgo.
- Síntesis interpretativa alineada a criterios clínicos.

Es importante resaltar que este análisis tiene carácter presuntivo y orientativo. La arquitectura fue diseñada explícitamente para evitar la automatización del diagnóstico clínico. El sistema actúa como herramienta de apoyo documental y analítico, manteniendo la responsabilidad final en manos del profesional.

d. Gestión de Historial (Pacientes y Sesiones)

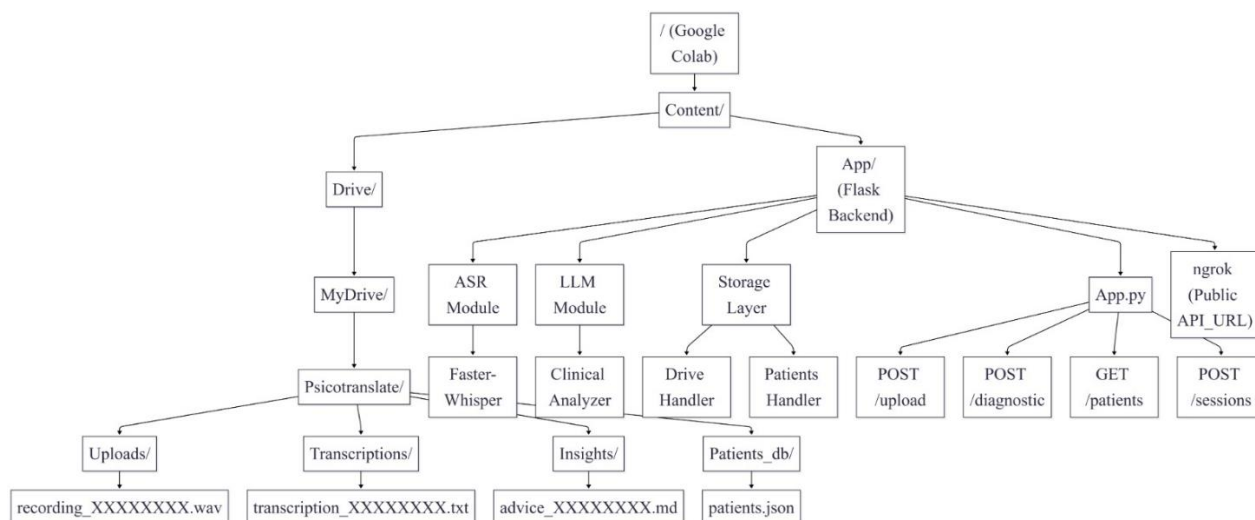
Adicionalmente, el backend administra estructuras de datos en formato JSON destinadas a organizar la información clínica del sistema. Estas estructuras incluyen:

- Registro de pacientes.
- Sesiones asociadas.
- Fecha de atención.
- Transcripción completa.
- Insights generados por el modelo.
- Metadatos técnicos del archivo de audio.

Este módulo permite consultas mediante el endpoint /patients y persistencia de nuevas sesiones mediante /sessions, consolidando el sistema como una herramienta integral de documentación clínica estructurada. De esta manera, el backend no solo procesa información, sino que organiza y preserva el historial terapéutico bajo un esquema coherente, trazable y escalable. En la siguiente Figura 8 se puede observar de forma completa en el que el Google Colab se encuentra armado y como se distribuye cada etapa para posteriormente unir todo.

Figura 8

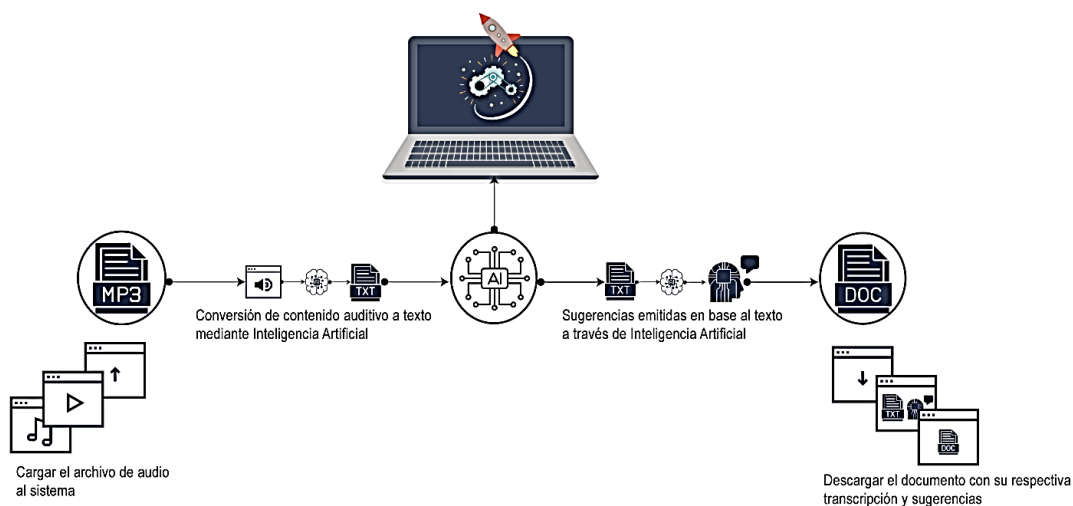
Estructuración final del diseño de la parte Backend



3.3.1.2. Diagrama de funcionamiento. En la Figura 9 se detallan los procesos que se llevarán a cabo para garantizar la funcionalidad del sistema. Una vez desplegado, es crucial contar con un manual de usuario que garantice su correcto uso.

Figura 9

Funcionalidad del sistema



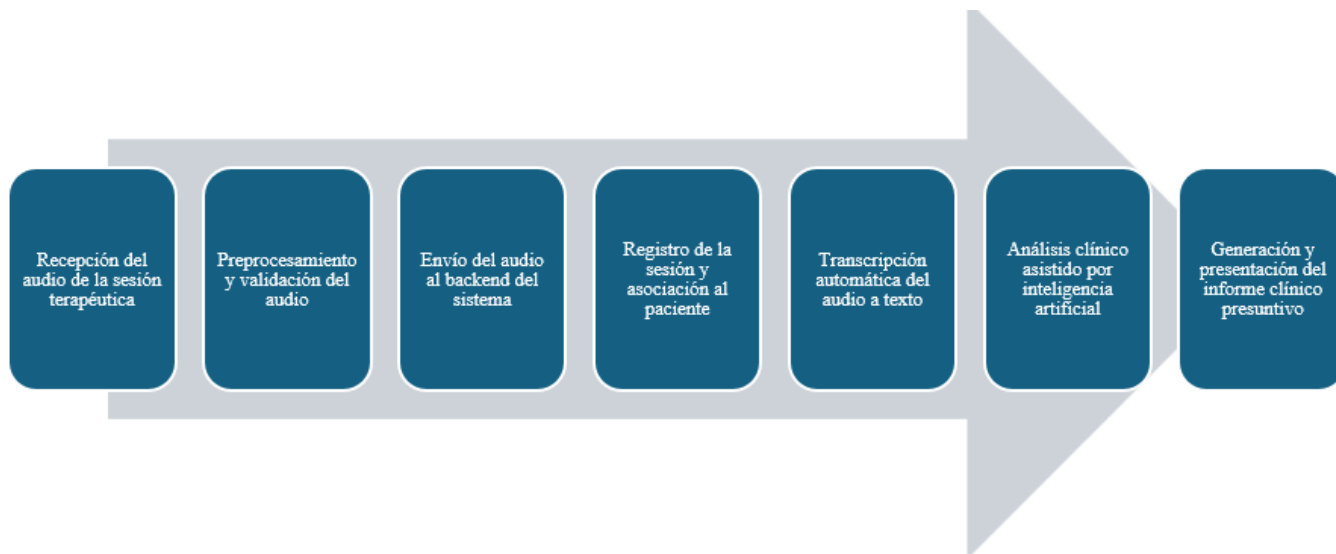
Nota. La figura muestra un sistema funcional del prototipo a implementar. Fuente: Autoría

De acuerdo como se lo pudo observar en la Figura 9 se debe tener un proceso sistemático para que se pueda realizar la transcripción de audio a texto y consecuentemente se puede proporcionar el diagnóstico respectivo de la sesión y su correspondiente almacenamiento. Para

mejor interpretación del proceso del sistema en la siguiente Figura 10 se detalla de mejor manera con un diagrama de bloques.

Figura 10

Diagrama de bloques del funcionamiento de forma simplificado

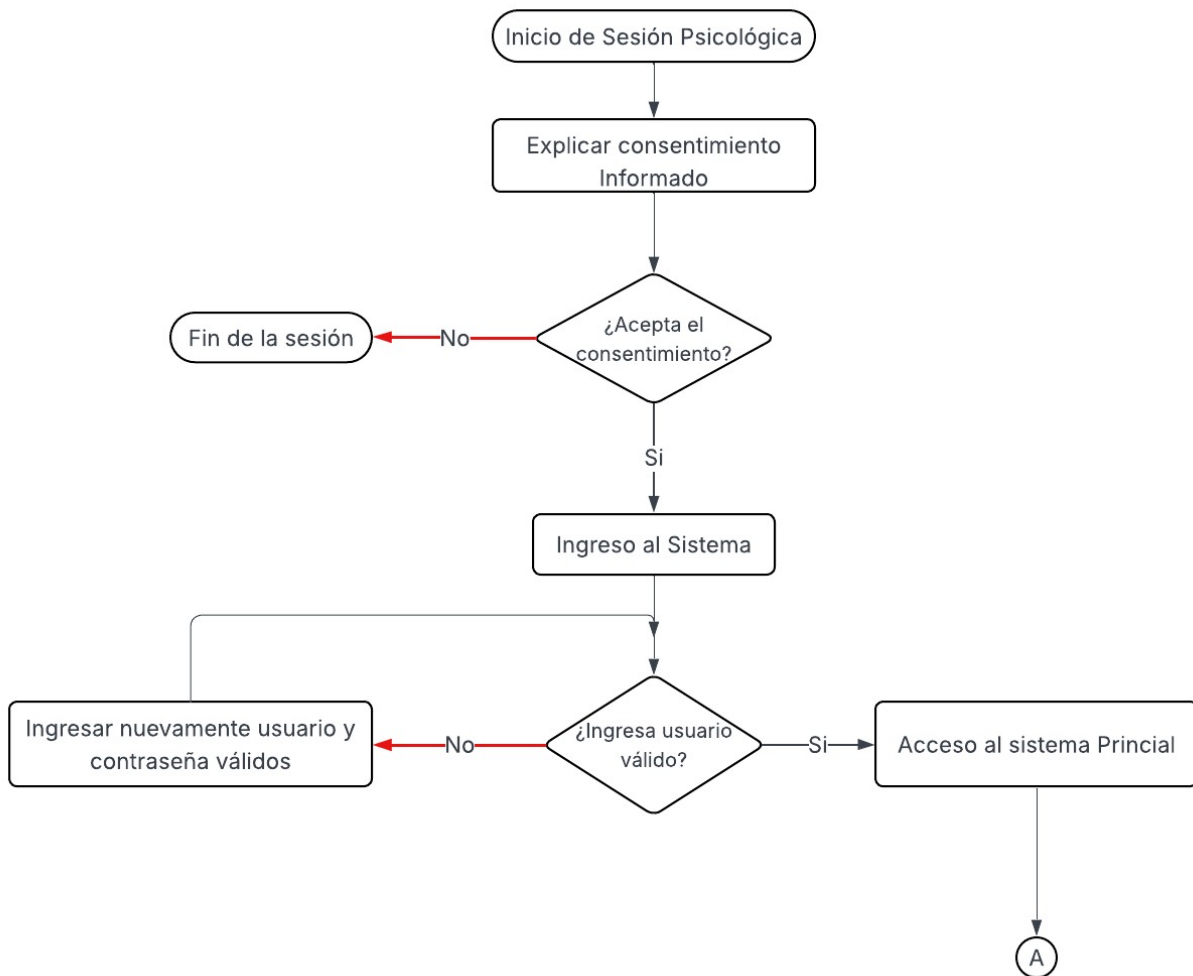


Pero para explicar de forma estructurada y completa como es el proceso del funcionamiento de todo el sistema en conjunto se lo describe a continuación por secciones para mejor entendimiento y organización.

El proceso inicia con la autenticación del terapeuta y la verificación del consentimiento informado del paciente, condición indispensable para el tratamiento de información clínica. Una vez validado el consentimiento se procede a la configuración de la sesión terapéutica donde se definen parámetros como el identificador pseudonimizado del paciente, el modo de procesamiento ya sea de tipo local o asistido por servidores y las características de captura de audio como se lo puede observar en la siguiente Figura 4.

Figura 11

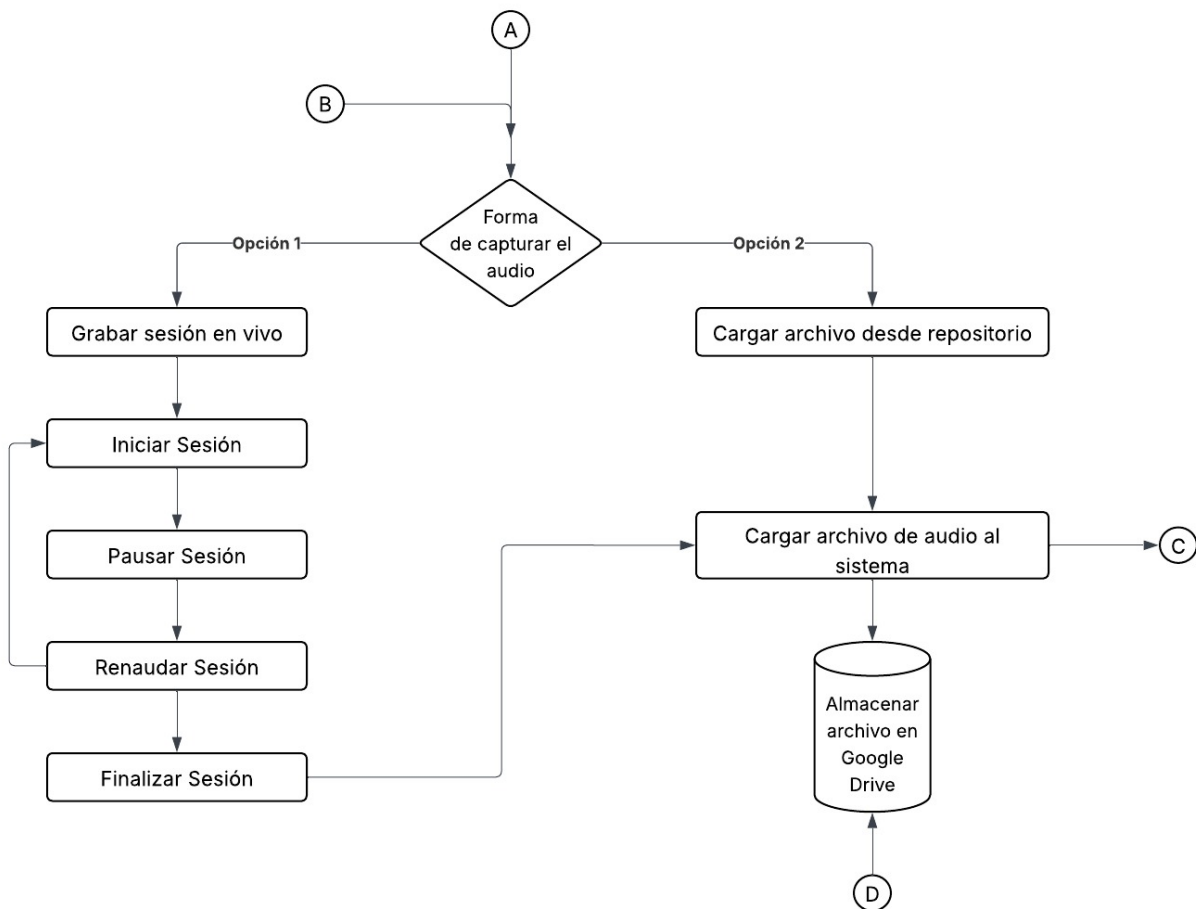
Primera etapa del proceso de funcionamiento del sistema



Posteriormente de acuerdo con la siguiente Figura 12 el sistema inicia la grabación de la sesión, verificando de forma automática la calidad de la señal de audio para garantizar condiciones adecuadas de procesamiento.

Figura 12

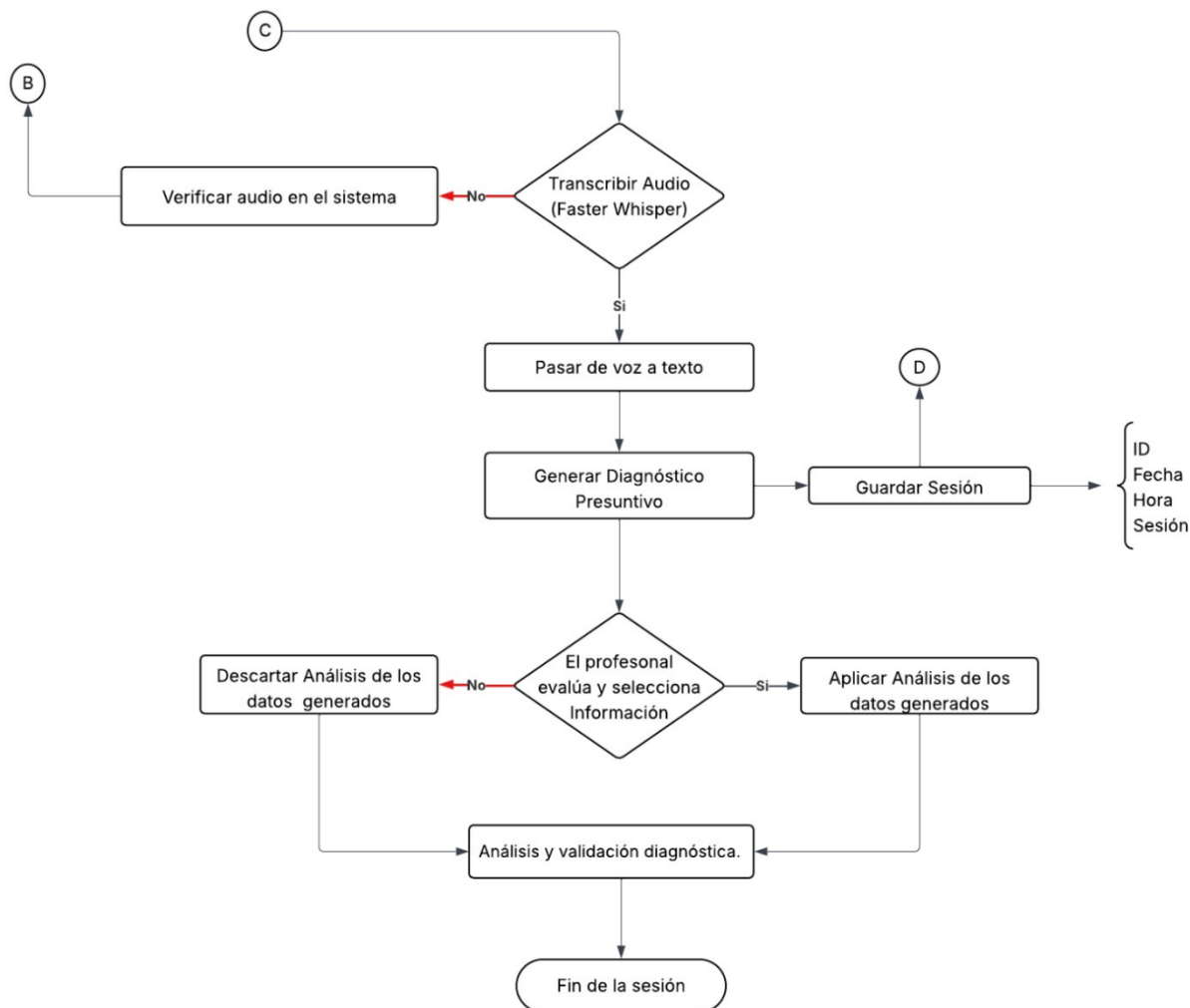
Segunda etapa del proceso de funcionamiento del sistema



A continuación, el audio capturado es almacenado temporalmente de forma cifrada y enviado al módulo de reconocimiento automático de voz, el cual realiza la transcripción progresiva de la sesión. La transcripción generada pasa por un proceso de desidentificación donde se eliminan o reemplazan datos sensibles, asegurando la privacidad del paciente. Sobre este texto, el sistema ejecuta módulos de procesamiento de lenguaje natural encargados de extraer información relevante, identificar patrones clínicos, generar resúmenes de la sesión y detectar posibles indicadores de riesgo que requieran atención inmediata como se lo puede observar en la Figura 13.

Figura 13

Tercera etapa del proceso de funcionamiento del sistema



Posteriormente como se lo observa los resultados obtenidos son presentados al terapeuta mediante un editor de corrección y validación, el cual permite revisar la transcripción, realizar ajustes manuales y confirmar la versión final del contenido. Una vez validada, la información se almacena en un repositorio digital único y seguro, facilitando el seguimiento clínico del paciente y la consulta histórica de sesiones anteriores. De forma paralela, el sistema puede generar la exportación de los reportes para dar mejor seguimiento al paciente en las siguientes sesiones.

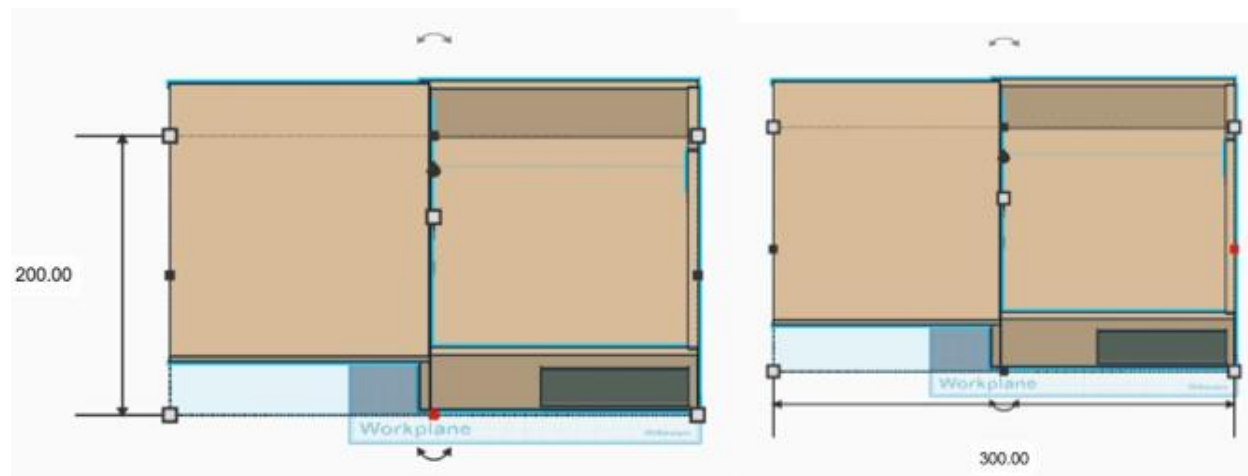
Finalmente, el proceso concluye con el cierre de la sesión y la actualización de métricas de desempeño del sistema, garantizando trazabilidad, seguridad y apoyo efectivo a la toma de decisiones clínicas.

3.1.1. *Hardware*

3.1.1.1. Case de protección. El diseño del case del prototipo se planteó con el objetivo de proteger la Raspberry Pi y sus componentes frente a condiciones del entorno como polvo, golpes accidentales y manipulación no autorizada, garantizando al mismo tiempo la estabilidad térmica y operativa del sistema. Para ello, se consideró una estructura cerrada con ventilación pasiva que permite su funcionamiento continuo sin comprometer la integridad del hardware. Esto se lo puede observar en la siguiente Figura 14 en donde se evidencia su forma y estructura dimensional.

Figura 14

Case de protección vista frontal

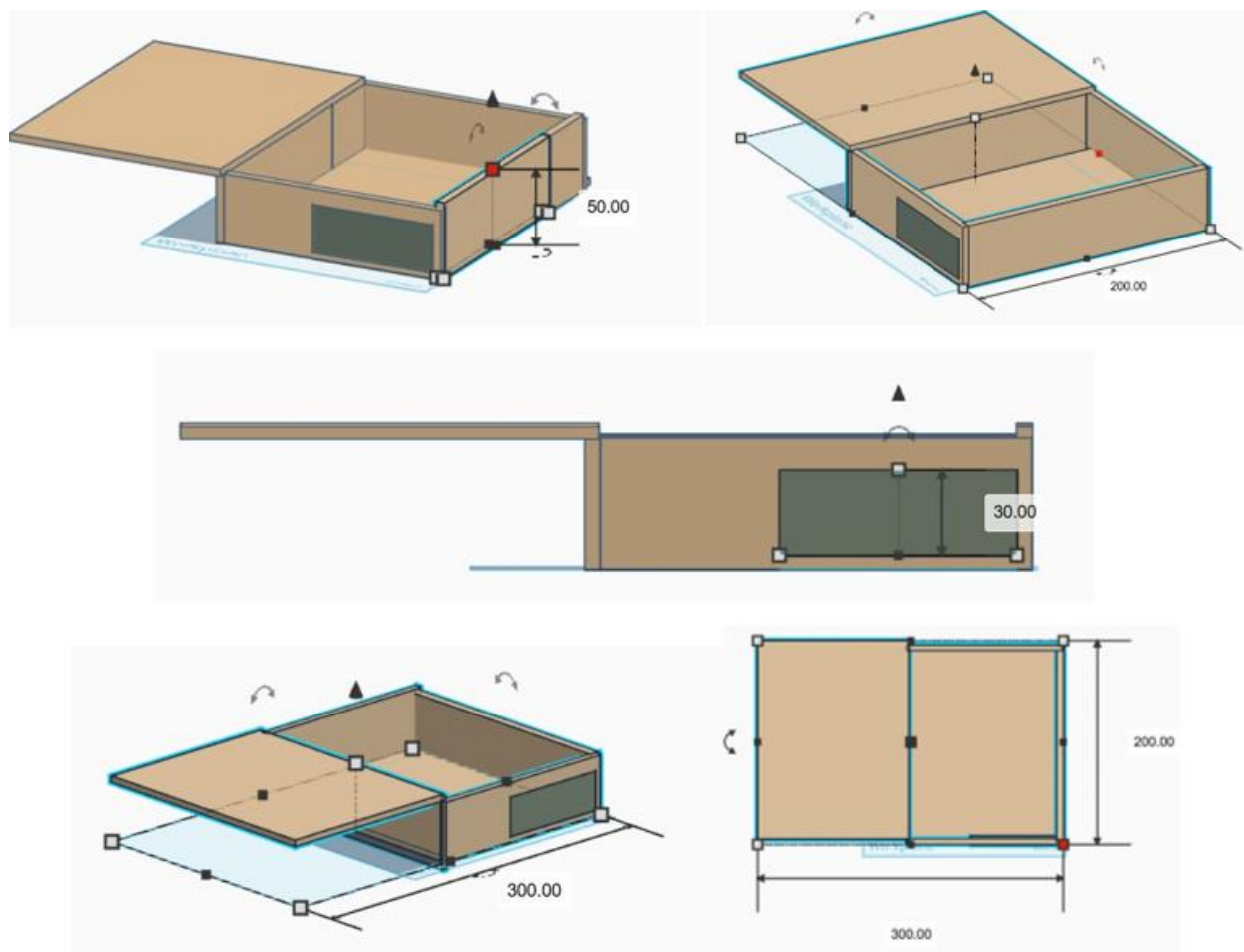


El case tendrá dimensiones de 20 cm de altura, 15 cm de largo y 5 cm de espesor, configurándose como una estructura compacta tipo libro diseñada para alojar los componentes electrónicos del sistema. Para su fabricación se utilizará el material MDF, debido por su adecuada resistencia mecánica, estabilidad dimensional y facilidad de mecanizado, lo que permitirá construir

una carcasa sólida y precisa conforme al diseño tridimensional establecido como se lo puede observar en la siguiente Figura 15 en el que se aprecia cada una de las vistas.

Figura 15

Perspectivas de vista del case de protección



El MDF ofrece una superficie uniforme que facilita acabados discretos, favoreciendo la integración del prototipo en el entorno clínico sin generar distracción visual. Además, su peso moderado y bajo costo lo convierten en una alternativa viable y funcional, garantizando protección física, portabilidad y replicabilidad del sistema en distintos espacios terapéuticos.

3.2. Implementación

La presente sección describe el proceso de implementación del prototipo desarrollado el análisis inteligente de notas clínicas en sesiones terapéuticas basado en el uso de inteligencia artificial y tecnologías de procesamiento de audio y lenguaje natural. Esta fase, en la que los requisitos de tipo funcional, no funcional, de hardware, de software..., recogidos en los capítulos anteriores, tienen lugar de forma práctica, permite validar el sentido técnico de la solución propuesta.

La implementación del sistema se realizó siguiendo un enfoque modular y estructurado que se ciñe a una arquitectura en capas, donde se logra la integración de diversos componentes del prototipo como por ejemplo la captura del audio, el procesamiento inteligente, el análisis clínico y el almacenamiento seguro de información. Esta organización permitió asegurar escalabilidad, mantenibilidad y control del flujo de datos, aspectos fundamentales en sistemas que manejan información clínica sensible.

3.2.1. *Software*

Como parte inicial se configuró el Raspberry Pi como nodo local de captura e interacción, cumpliendo el rol de intermediario entre el usuario y la capa de procesamiento en la nube. Este dispositivo se encarga de servir la interfaz web, gestionar la grabación local de audio y garantizar la conectividad con los servicios de inteligencia artificial desplegados en Google Colab esto se lo pudo realizar a través del siguiente proceso:

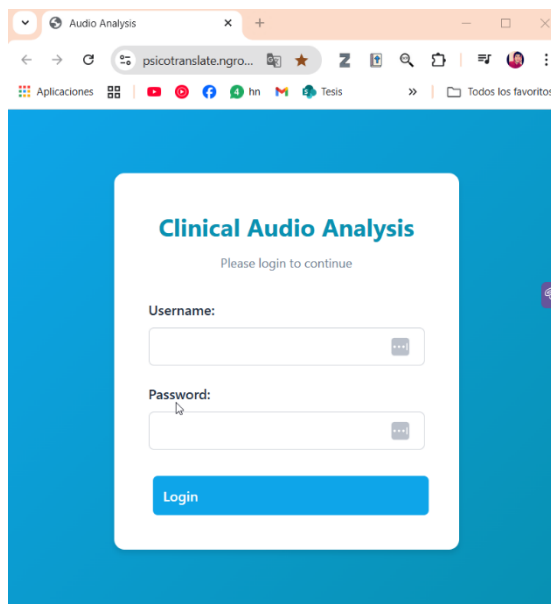
3.2.1.1. Frontend. En cuanto a la Interfaz Web se encuentra en primer lugar la creación de la página visual interpretativa principal creada en base .html como se lo puede observar en la siguiente Figura 16 mediante la incorporación de un requerimiento de login de acceso. De forma visual se encuentra comprendida en la siguiente Figura 17.

Figura 16*HTML de login de acceso*

```
● ● ●  
  
// Inicializa la app: decide si muestra login o dashboard  
async function init() {  
  const isLoggedIn = sessionStorage.getItem('isLoggedIn');  
  if (isLoggedIn === 'true') {  
    showMainApp(); // Muestra la aplicación principal  
  } else {  
    showLoginPage(); // Muestra la pantalla de login  
  }  
}  
  
// Maneja el envío del formulario de login (modo prototipo)  
document.getElementById('login-form').addEventListener('submit', function(e) {  
  e.preventDefault(); // Evita recarga de página  
  
  const username = document.getElementById('login-username').value;  
  const password = document.getElementById('login-password').value;  
  
  // Autenticación básica local para prototipo  
  if (username === 'user' && password === '12') {  
    sessionStorage.setItem('isLoggedIn', 'true'); // Guarda sesión en navegador  
    showMainApp(); // Entra al dashboard  
  } else {  
    // Muestra error si credenciales inválidas  
    const loginError = document.getElementById('login-error');  
    loginError.textContent = 'Invalid username or password. Please try again.';  
    loginError.classList.remove('hidden');  
  }  
});
```

Figura 17

Login de acceso



Como se observó en esta parte del frontend se incorpora una pantalla de login que funciona como un gate de acceso en modo prototipo. La aplicación verifica si existe un indicador de sesión en `sessionStorage` para decidir si mostrar el login o el panel principal, cuando el usuario ingresa credenciales válidas se guarda el estado de login en el almacenamiento de sesión del navegador y se renderiza la aplicación principal. Este enfoque simplifica el flujo de pruebas y evita que el usuario vea el dashboard sin pasar por el formulario, aunque no reemplaza un mecanismo de autenticación real en producción.

En segundo lugar, se lleva a cabo el procesamiento del audio y la solicitud de transcripción automática. En este punto, la interfaz frontend permite al usuario seleccionar o arrastrar un archivo de audio hacia el módulo de carga. Al enviar el formulario, el sistema construye un objeto `FormData`, el cual encapsula el archivo en formato binario y lo transmite al backend mediante una solicitud `HTTP POST` dirigida al endpoint `/upload` como se lo puede observar en la siguiente Figura 18.

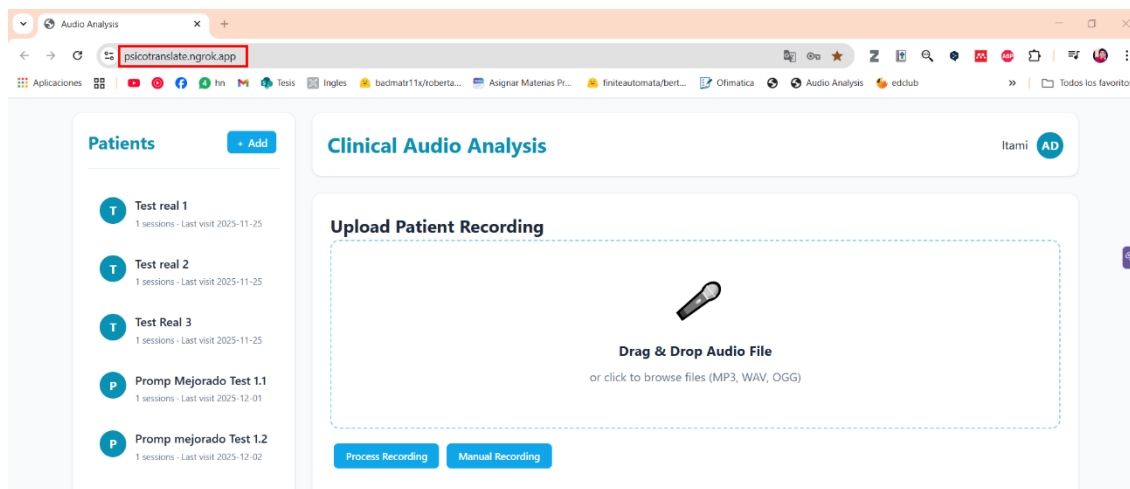
Figura 18*Procesamiento del audio y la solicitud de transcripción*

```
async function uploadAndTranscribe(file) {  
  // Prepara archivo binario para HTTP multipart/form-data  
  const formData = new FormData();  
  formData.append("file", file);  
  
  // Envía audio al backend para transcripción  
  const res = await fetch("/upload", { method: "POST", body: formData });  
  const data = await res.json();  
  
  // Muestra texto transcrito en la UI  
  transcriptionDisplay.textContent = data.transcriptionModel;  
}
```

De manera paralela y con el fin de mejorar la experiencia de uso y permitir una validación previa del contenido, el audio cargado es reproducido localmente en el navegador utilizando un reproductor basado en WaveSurfer. Para ello, el sistema genera un ObjectURL, lo que permite la reproducción inmediata del archivo sin necesidad de esperar la respuesta del servidor ni realizar una descarga adicional. Una vez que el backend completa el proceso de transcripción y retorna la respuesta correspondiente, el texto generado es inyectado dinámicamente en el panel de transcripción de la interfaz. A partir de este punto el sistema habilita de forma secuencial los módulos posteriores, específicamente la generación de insights clínicos mediante inteligencia artificial y la funcionalidad de registro y almacenamiento de la sesión, garantizando un flujo de trabajo continuo y coherente para el profesional como se lo observa en la Figura 19.

Figura 19

Gestión de audios clínicos

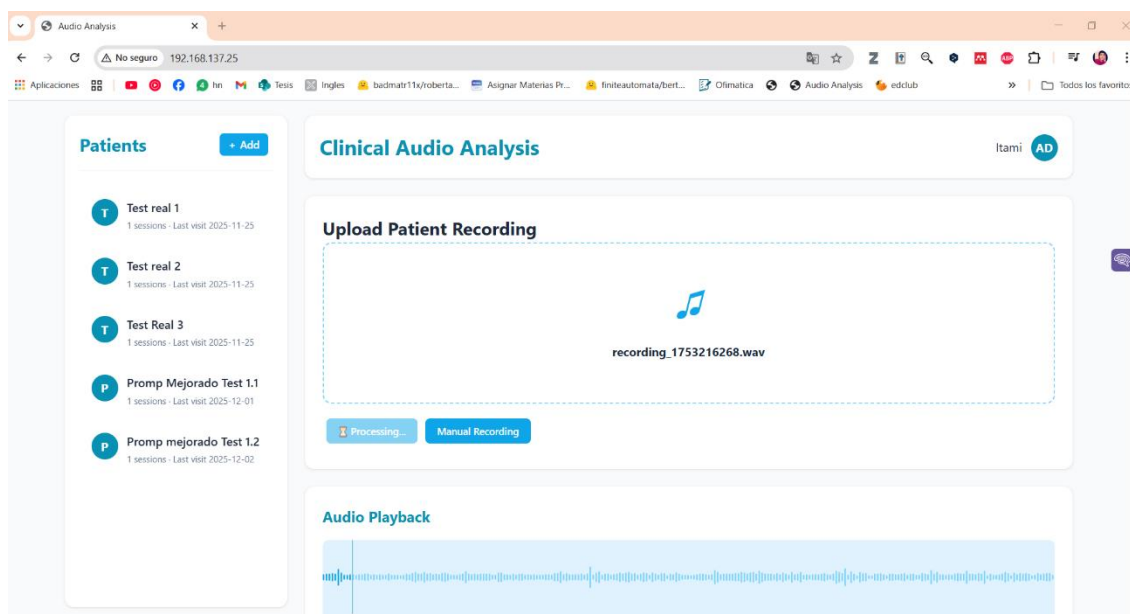


Como se llegó a observar en la Figura 19 se tiene presente dos modalidades en donde el sistema permite incorporar archivos con formatos MP3, WAV y OGG. La primera cuando se carga de audio y posteriormente se hace uso del botón (Process Recording) que ejecuta el procesamiento del archivo para obtener la transcripción automática del audio a texto, que se utilizará en las etapas posteriores de análisis y generación de informes. Y como segunda modalidad se tiene presente el registro manual (Manual Recording) en donde básicamente se graba la sesión desde la plataforma, para luego procesar el audio correspondientemente como en la primera modalidad.

En este caso si se realiza mediante la primera modalidad el archivo queda visible en la interfaz mostrando su nombre y se habilita el módulo de reproducción de audio (Audio Playback), el cual permite escuchar el contenido como mecanismo de verificación, confirmando que el archivo cargado corresponde al registro deseado antes de continuar con el procesamiento como se lo observa en la Figura 20.

Figura 20

Primera modalidad de subir el audio para procesamiento clínico



Dentro de este aspecto se encuentra la ejecución del entorno en segundo plano en Google Colab, el cual funciona como servidor del sistema y muestra en tiempo real la actividad generada por la interfaz web. Cada acción del usuario como recargar la página, iniciar sesión o consultar informes previamente almacenados produce solicitudes HTTP como se lo observa en la siguiente Figura 21 en el que se está permitiendo que cada solicitud sea reflejada en la consola del notebook.

Figura 21

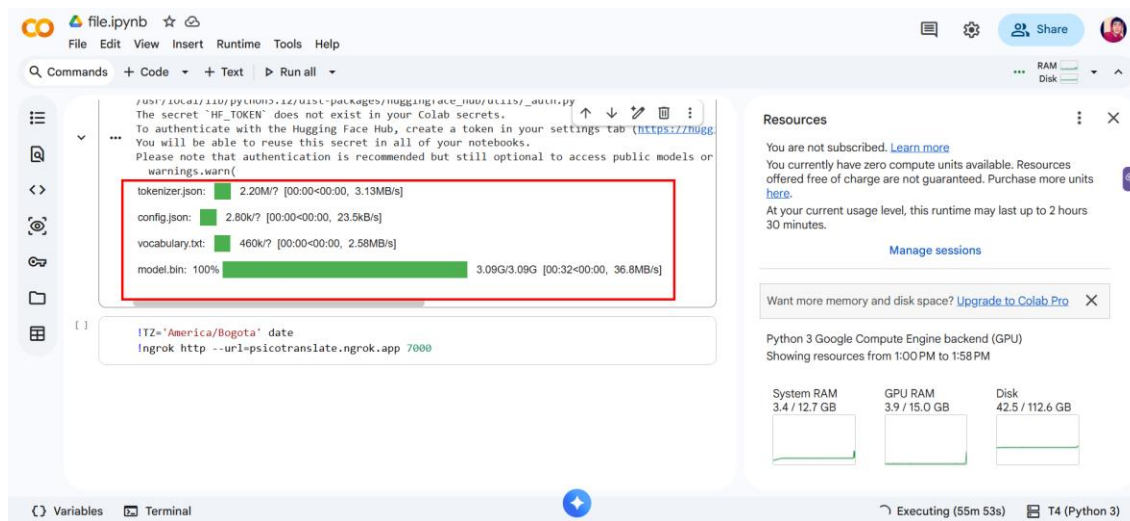
Petición de procesamiento en Google Colab



Como se pudo observar al presionar el botón Process Recording se activa la petición de procesamiento, en donde el backend recibe el archivo de audio, carga y usa los recursos del modelo para ejecutar la tarea de transcripción automática, convirtiendo el contenido de voz a texto. Esta salida constituye la base para las etapas posteriores de análisis y generación de informes. Cuando se inicia la transcripción Google Colab muestra barras de progreso asociadas a los recursos que utiliza el modelo. En la siguiente Figura 22 se observan indicadores para la carga de componentes como tokenizer, config.json, vocabulary y el archivo principal del modelo. Estas métricas reflejan la descarga/lectura e inicialización de los archivos necesarios para ejecutar la inferencia.

Figura 22

Barras de progreso asociadas al momento de la transcripción del audio a texto



Posteriormente la consola de Google Colab registra el evento mediante la siguiente estructura:

INFO:werkzeug: 127.0.0.1 - - [fecha/hora] "POST /upload HTTP/1.1" 200

Este mensaje corresponde al log del servidor Werkzeug/Flask y evidencia que el sistema recibió correctamente una solicitud HTTP POST al endpoint /upload como se lo puede observar en la Figura 52. El código de estado 200 equivalente a OK confirma que la petición fue procesada

sin errores, es decir el archivo de audio se cargó de forma satisfactoria y el backend pudo continuar con el flujo previsto de almacenamiento y ejecución de la transcripción.

Figura 23

Registro del evento de transcripción del audio a texto en Google Colab

```

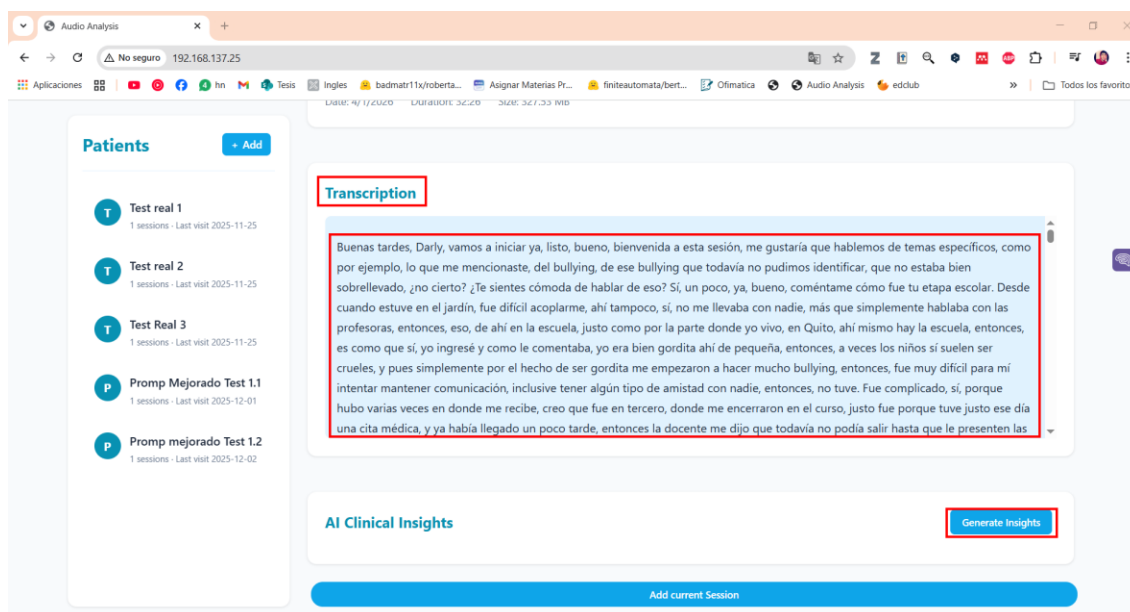
tokenizer.json: █ 2.20M/? [00:00<00:00, 3.13MB/s]
config.json: █ 2.80k/? [00:00<00:00, 23.5kB/s]
vocabulary.txt: █ 460k/? [00:00<00:00, 2.58MB/s]
model.bin: 100% █ 3.09G/3.09G [00:32<00:00, 36.8MB/s]
INFO:werkzeug:127.0.0.1 - - [04/Jan/2026 19:00:29] "POST /upload HTTP/1.1" 200 -

```

Al retornar a la interfaz principal se habilita la sección Transcription, donde se visualiza el resultado de la transcripción comprendido en el audio convertido a texto correspondiente a la sesión de psicoterapia cargada. Este bloque funciona como respaldo documental de la entrevista y permite revisar el contenido de manera íntegra. Todo esto para posteriormente depurar la información, es decir priorizando los elementos relevantes para el caso, principalmente las intervenciones del paciente y cuando aplica indicaciones puntuales del profesional que aportan contexto como preguntas guía o aclaraciones. Esta selección reduce ruido, mejora la claridad del registro y facilita la extracción de hallazgos sin perder el aporte esencial del paciente. (Ver Figura 24)

Figura 24

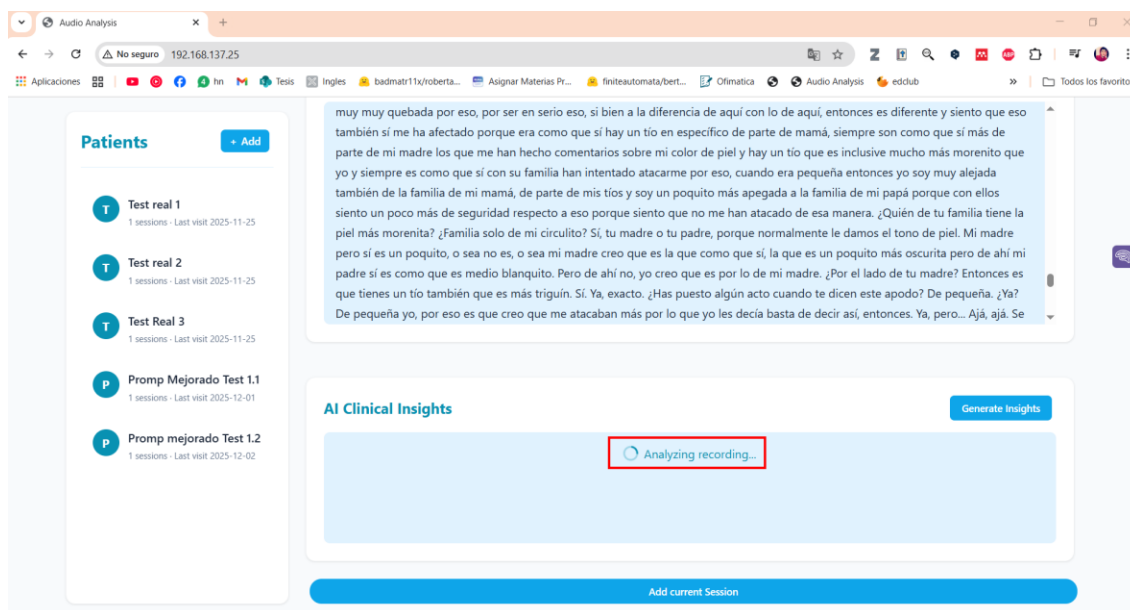
Observación de la transcripción del audio a texto



Una vez lograda la depuración de información se habilita el AI Clinical Insights junto con el botón Generate Insights. En la siguiente Figura 25 se permite observar esta funcionalidad ejecutando el análisis automático del contenido transcrito y permitiendo generar un diagnóstico presuntivo (informe preliminar) a partir de la transcripción obtenida en la etapa anterior.

Figura 25

Generación del diagnóstico presuntivo



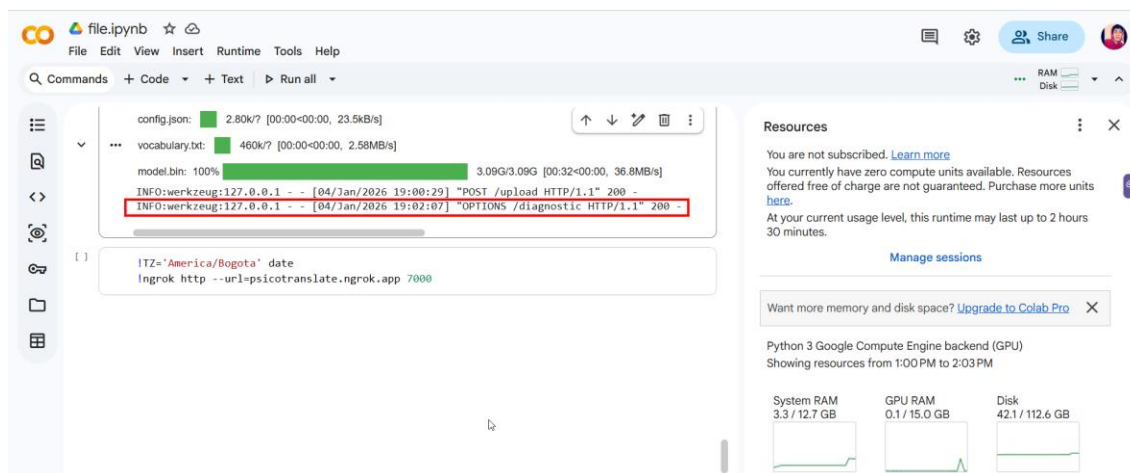
Durante el procesamiento la interfaz muestra el estado Analyzing recording, indicando que el sistema está evaluando el texto para extraer hallazgos relevantes y estructurarlos en un informe. El tiempo de respuesta puede variar según la conectividad de red, la duración y tamaño del archivo de audio procesado. Como toda etapa esta se registra de igual manera en Google Colab mediante la siguiente estructura:

"OPTIONS /diagnostic HTTP/1.1" 200

Esta forma de estructura corresponde a una solicitud OPTIONS enviada al endpoint/diagnostic como parte del mecanismo de verificación previo a la petición principal habitual cuando la interfaz web realiza solicitudes entre orígenes (CORS). El código 200 confirma que el servidor respondió correctamente y habilitó la comunicación para ejecutar la operación de generación del diagnóstico. En la siguiente Figura 26 se puede observar claramente el registro de este evento mediante Google Colab.

Figura 26

Registro del evento de generar diagnóstico



De vuelta en la interfaz principal se visualiza el resultado del módulo AI Clinical Insights, el cual genera una evaluación psicológica preliminar a partir del contenido textual obtenido en la transcripción del audio. Como se aprecia en la siguiente Figura 27, el informe se organiza por apartados que sintetizan hallazgos relevantes, incluyendo emociones expresadas, distorsiones cognitivas identificadas, patrones de pensamiento, conductas implicadas, preocupaciones centrales y rasgos de personalidad sugeridos.

Figura 27

Detección de hallazgos relevantes dentro del informe presuntivo

Evaluación Psicológica Basada en el Contenido Textual

1. Emociones expresadas:

- Dolor, tristeza profunda, humillación, soledad, abandono.
- Miedo al rechazo, a la intimidad y a la vulnerabilidad.
- Culpa (por comer, por existir, por defenderse).
- Rabia contenida y resentimiento hacia figuras de autoridad (padres, nutriólogo, agresores).
- Desesperanza pasada (situación suicida a los 10 años).
- Baja autoestima y autoapreciación focalizada en el cuerpo (peso, color de piel, rostro).

2. Distorsiones cognitivas identificadas:

- **Sobregeneralización:** "Por mi físico me alejan todos", "no necesito tener pareja".
- **Filtros mentales/Descualificación de lo positivo:** A pesar de que la ropa le queda floja y médicos indican peso normal, solo se fija en el espejo donde se "ve gorda".
- **Personalización:** Asume que el maltrato fue por sus características físicas, sin cuestionar la crueldad de los otros.
- **Pensamiento dicotómico:** "Como demasiado" vs. "debo dejar de comer", visión de sí misma como "gorda" a pesar de evidencia contraria.
- **Razonamiento emocional:** "Me siento gorda, por lo tanto, lo soy", "siento que no necesito pareja, entonces no la tengo".

3. Patrones de pensamiento:

- Hipervigilancia al rechazo y a la evaluación social.
- Actuarial severa y focalización obsesiva en la imagen corporal.
- Expectativas de abandono/desprotección por parte de figuras de apoyo (padres).

4. Conductas implicadas:

- Evitación social histórica y actual (solo se abre con 2 compañeros tras ganar confianza).
- Conductas compensatorias por trauma: agresión reactiva en la infancia (patrón de defensa-hostilidad).
- Restricción alimentaria severa en el pasado, con episodios de vómito autoinducido.
- Oculta sus emociones y necesidades a la familia, muestra una fachada de "normalidad" alimentaria frente a ellos.
- Evitación de conflictos (ejemplo: esquivó a la madre del compañero).

5. Preocupaciones centrales:

- Imagen corporal (peso, color de piel).
- Miedo al rechazo interpersonal.
- Dificultad para confiar y mostrar afecto.
- Relación distante y poco comprensiva con los padres.
- Culpa persistente asociada a la alimentación.

6. Rasgos de personalidad sugeridos:

- Alta evitación del daño (evitación de situaciones sociales o conflictivas).
- Baja autoeficacia (inseguridad, autoeficacia disminuida).
- Posibles rasgos de desconfianza/desapego como secuela del maltrato prolongado.

Diagnóstico Presuntivo Clínico (CIE-11 / DSM-5)

A. Trastorno de Estrés Posttraumático (CIE-11: 6B40 / DSM-5: 309.81)

• *Criterios resumidos (DSM-5):*

Es importante señalar que estos resultados se derivan directamente del texto transcrito y, por tanto, constituyen una aproximación inicial que sirve como apoyo para la interpretación clínica del profesional. Posteriormente, el sistema genera el apartado Diagnóstico Presuntivo Clínico con una etiqueta que lo diferencie de los demás, en este se proponen hipótesis diagnósticas sustentadas en dos marcos de referencia CIE-11 y DSM-5. En esta sección, la plataforma relaciona los hallazgos identificados en la transcripción como los síntomas, conductas y patrones descritos con categorías diagnósticas y presenta los criterios asociados que se consideran cumplidos,

organizando la información por trastorno o conducta relevante. De este modo, el informe integra una correspondencia normativa entre CIE-11 y DSM-5 para orientar la interpretación clínica, manteniendo el carácter presuntivo del resultado al estar basado en el contenido textual previamente transcrito. En la siguiente Figura 28 se puede ver las interpretaciones clínicas que se logra obtener en base a los marcos de referencia.

Figura 28

Interpretación del diagnóstico en base a los marcos de referencia

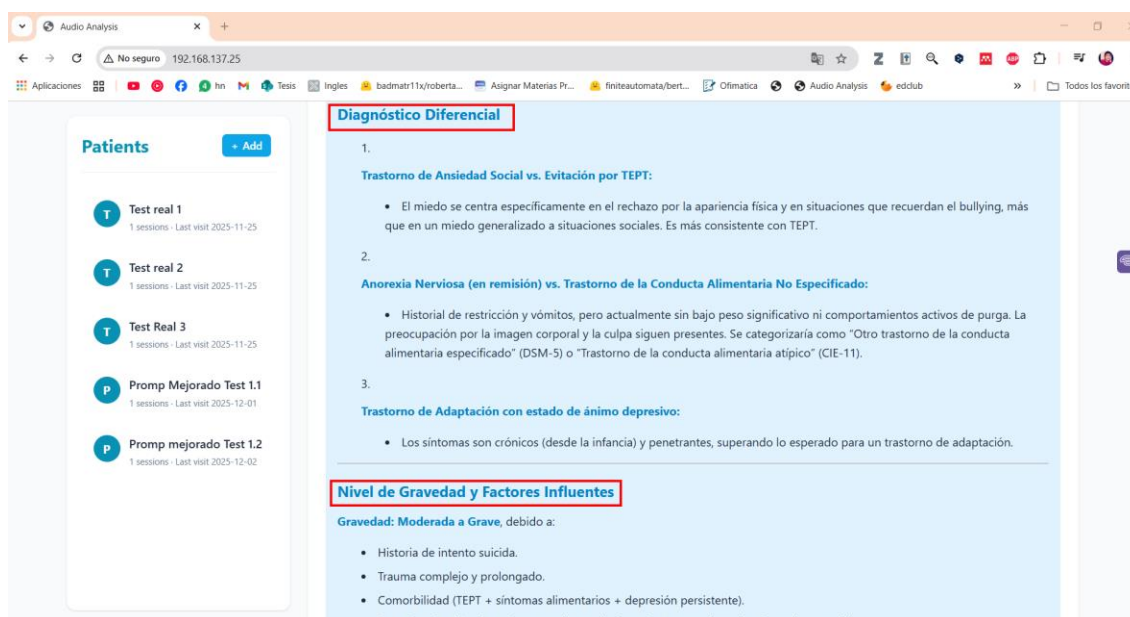
The figure consists of two screenshots of a web-based clinical report interface. The top screenshot displays a 'Diagnóstico Presuntivo Clínico (CIE-11 / DSM-5)' section. It lists two main diagnostic categories: A. Trastorno de Estrés Posttraumático (CIE-11: 6B40 / DSM-5: 309.81) and B. Trastorno de la Conducta Alimentaria (Atípico o en revisión parcial) - (CIE-11: 6B80 / DSM-5: 307.1). Under category A, it lists five DSM-5 criteria: 1. Exposición a trauma, 2. Reexperimentación, 3. Evitación, 4. Alteraciones cognitivas y del estado de ánimo, and 5. Alteraciones en la activación y reactividad. A note specifies that the trauma was interpersonal and occurred in early development. Under category B, it lists criteria such as a history of severe restrictive eating behaviors, binge eating, and persistent concern about weight and appearance. The bottom screenshot shows a similar interface with category C. Trastorno Depresivo Persistente (Distonía) con episodios depresivos pasados graves (CIE-11: 6A72 / DSM-5: 300.4) and category D. Posibles rasgos de Trastorno de Personalidad por Evitación (CIE-11: 6D11 / DSM-5: 301.82). Category C lists criteria like low self-esteem and feelings of hopelessness. Category D lists criteria like social avoidance and inhibition in intimate relationships. Both screenshots include a 'Patients' sidebar on the left with a list of tests and a 'Add' button.

En la siguiente Figura 29 se observa que en el informe incluye un apartado de diagnóstico diferencial, cuyo objetivo es contrastar la hipótesis diagnóstica principal con otras condiciones clínicas con sintomatología similar, justificando la elección más consistente con la información

disponible. Esta sección presenta una lista de alternativas diagnósticas y resume los criterios o elementos del relato que permiten diferenciar. Adicionalmente, el sistema incorpora el bloque Nivel de gravedad y factores influyentes, donde se describe una estimación cualitativa de severidad y se enumeran variables clínicas relevantes que pueden influir en la presentación del caso (historia del problema, comorbilidades, indicadores de riesgo y grado de afectación funcional), con base en el contenido transcrito.

Figura 29

Inclusión del diagnóstico diferencial y nivel de gravedad



Una vez identificado y diferenciado los síntomas se presenta en la siguiente Figura 30 la sección Plan de Tratamiento Recomendado, considerado como un componente clave del sistema, ya que ofrece al profesional una propuesta de intervención estructurada a partir de los hallazgos obtenidos en la transcripción y el análisis previo.

Figura 30*Generación del plan de tratamiento recomendado*

Plan de Tratamiento Recomendado (Enfoques Basados en Evidencia)

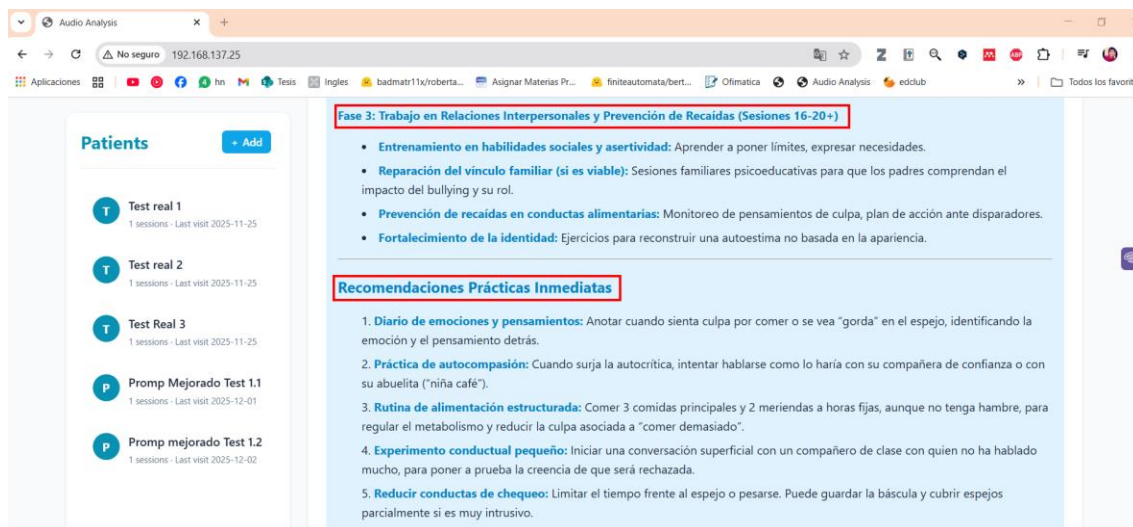
Fase 1: Estabilización y Alianza Terapéutica (Sesiones 1-4)

- **Psicoeducación** sobre TEPT, impacto del bullying y relación con la imagen corporal.
- **Normalización y validación** de sus emociones.
- **Técnicas de grounding y regulación emocional inicial:** Respiración diafragmática, mindfulness básico para manejar ansiedad y culpa.
- **Establecimiento de seguridad:** Identificar redes de apoyo actuales (compañeros) y recursos de crisis.

Fase 2: Procesamiento del Trauma y Reestructuración Cognitiva (Sesiones 5-15)

- **Terapia Cognitiva Conductual Centrada en Trauma (TF-CBT) o Terapia de Procesamiento Cognitivo (CPT):**
 - Identificar pensamientos automáticos negativos ("soy defectuosa", "merecí el bullying").
 - Cuestionar creencias sobre culpa, vergüenza y autovalía.
 - **Exposición narrativa:** Elaborar la historia de vida traumática de manera estructurada y segura, integrando emociones.
 - **Reestructuración de la autopercepción corporal:** Trabajar la disociación entre la imagen corporal subjetiva ("me veo gorda") y la realidad objetiva (ropa floja, peso normal).
- **Técnicas de Terapia de Aceptación y Compromiso (ACT):**
 - **Defusión cognitiva:** Separarse de pensamientos autocríticos ("estoy teniendo el pensamiento de que soy gorda, no soy gorda").
 - **Aceptación de emociones dolorosas** sin juzgarlas.
 - **Clarificación de valores:** ¿Quién quiere ser más allá del peso? ¿Qué relaciones quiere construir?

El plan se organiza por fases clínicas como se pudo observar en la Figura 30 y 31, en donde se detalla un rango estimado de sesiones, además de objetivos y estrategias terapéuticas sugeridas. Esta estructuración facilita la planificación secuencial del proceso terapéutico, orientando qué abordar inicialmente y cómo progresar hacia etapas posteriores.

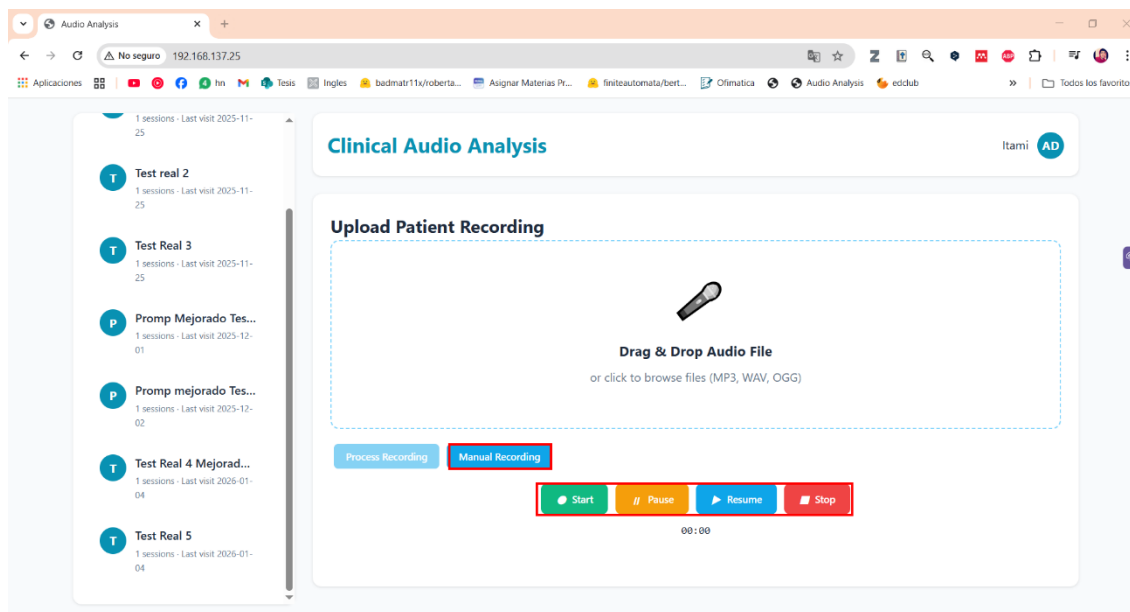
Figura 31*Establecimiento de fases clínicas del plan de tratamiento recomendado*

Adicionalmente, el sistema incluye una nota final donde se sintetiza una observación global del caso comprendiendo los aspectos de resiliencia, avances o consideraciones generales, los cuales sirven como cierre interpretativo del análisis generado. Es importante señalar que este contenido corresponde a un plan de tratamiento sugerido y no constituye una prescripción obligatoria. La selección, adaptación o descarte de las recomendaciones queda bajo responsabilidad del profesional tratante quien definirá qué componentes se aplican según su criterio clínico, experiencia, el contexto del paciente y la evolución del caso.

Como segunda modalidad se encuentra la forma Manual de captura de información, la cual tiene como principio captar el audio en tiempo real mediante la opción Manual Recording. Al seleccionarla la interfaz habilita un panel de control con botones que permiten gestionar la grabación durante la sesión, proporcionando mayor control operativo al profesional como se lo puede observar en la siguiente Figura 32.

Figura 32

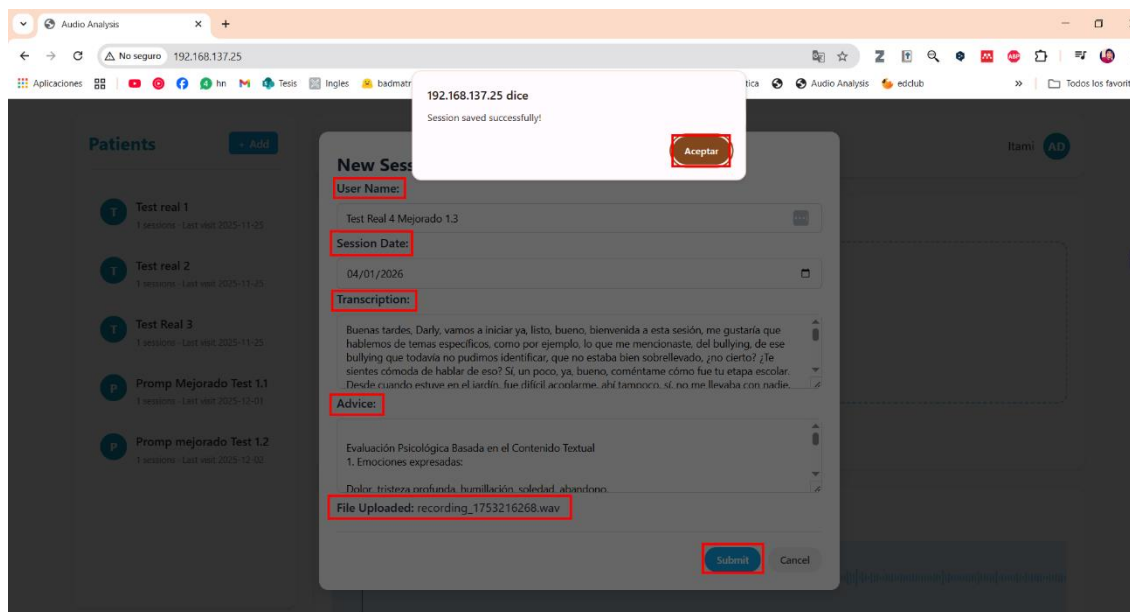
Segunda modalidad de subir el audio para procesamiento clínico



De acuerdo como la primera modalidad se hace el procesamiento y generación del informe presuntivo. Una vez generado el informe completo el sistema habilita el formulario de registro de la sesión para su almacenamiento. Como se observa en la siguiente Figura 33 donde se presentan campos como User Name (identificador del paciente), Session Date, Transcription, Advice e indicador del archivo cargado.

Figura 33

Registro de la sesión para almacenamiento



Por motivos de confidencialidad en la práctica no se registran nombres propios; en su lugar, se propone el uso de identificadores numéricos como número de cédula/CI o un código interno. El resto de los campos se completan automáticamente con la información generada por el sistema, la cual viene comprendida en la fecha de la sesión, el texto resultante de la transcripción y las recomendaciones o evaluación preliminar. Además, se registra si el audio proviene de un archivo subido o de un registro manual.

Tras la revisión de los datos, el profesional selecciona Submit para guardar la sesión. La confirmación (¡Session saved successfully!) indica que el registro fue almacenado correctamente y pasa a visualizarse dentro del listado de Patients.

Como parte final se puede evidenciar en la siguiente Figura 34 para recuperar el historial de los pacientes se hace uso de un GET.

Figura 34

Recuperación del historial de pacientes

```

async function loadPatients() {
  // Recupera pacientes y sesiones desde el backend
  const res = await fetch("/patients");
  patients = await res.json();

  // Renderiza el listado (implementación UI)
  renderPatientList(patients);
}

```

El historial se obtiene mediante el endpoint `/patients`, el cual retorna una estructura con pacientes y sus sesiones previas. El frontend utiliza esta respuesta para poblar el sidebar y permitir la navegación por fechas. Al seleccionar una sesión, la interfaz carga la transcripción y el informe asociados, facilitando el seguimiento longitudinal de cada caso. Este mecanismo desacopla la UI de datos estáticos y garantiza que la información provenga de la persistencia del servidor.

3.2.1.2. Backend. Siguiendo el diseño la arquitectura del sistema establecido en la sección 3.3.1.1, el backend del sistema se implementa en dos entornos diferenciados, el backend de captura en Raspberry Pi y el backend de procesamiento en Google Colab.

a) ***Backend de captura (Raspberry Pi – Flask + arecord).*** Esta fase corresponde a la materialización del módulo encargado de la adquisición controlada del audio clínico.

- ***Configuración del entorno base.*** De acuerdo con este aspecto se puede mencionar que empieza mediante el siguiente proceso sistematizado par mantener la operabilidad acorde sin fallas.

a. Configuración inicial del sistema operativo

Como se puede observar en la siguiente Figura 35 como primer paso se verificó el correcto funcionamiento del Raspberry Pi incluyendo procesador, memoria RAM y almacenamiento, así

como la correcta instalación del sistema operativo basado en Linux. Dentro de esta validación se puede mencionar que se llegó a asegurar un entorno estable y por ende compatible con la ejecución de servicios en Python, en conjunto con la positiva gestión de dispositivos de audio mediante ALSA.

Figura 35

Instalación del Sistema Operativo del Raspberry Pi



Así mismo se pudo verificar la red, que es un requerimiento necesario para la permisibilidad de la comunicación recíproca exclusivamente entre la Raspberry Pi y el backend implementado en Google Colab utilizando un canal seguro.

b. Verificación y configuración del subsistema de audio (ALSA)

Posteriormente, se validó la disponibilidad de dispositivos de captura de audio mediante el subsistema ALSA. A través de los comandos `arecord -l` y la inspección del archivo `/proc/asound/cards`, se identificó correctamente un dispositivo de entrada USB PnP Sound Device (C-Media), registrado como card 2, device 0 (hw:2,0), confirmando su compatibilidad para la grabación de audio.

Figura 36

Verificación del subsistema de audio (ALSA)

```

user@raspberrypi: ~
root@raspberrypi:~# arecord -l
cat /proc/asound/cards
**** List of CAPTURE Hardware Devices ****
card 2: Device [USB PnP Sound Device], device 0: USB Audio [USB Audio]
  Subdevices: 1/1
  Subdevice #0: subdevice #0
0 [Headphones      ]: bcm2835_headpho - bcm2835 Headphones
                        bcm2835 Headphones
1 [vc4hdmi0        ]: vc4-hdmi - vc4-hdmi-0
                        vc4-hdmi-0
2 [Device          ]: USB-Audio - USB PnP Sound Device
                        C-Media Electronics Inc. USB PnP Sound Device at usb-0000:01:00.0-1.2, full sp
3 [vc4hdmi1        ]: vc4-hdmi - vc4-hdmi-1
                        vc4-hdmi-1

```

Esta validación es crítica ya que el correcto reconocimiento del micrófono garantiza la fiabilidad del proceso de captura de audio, base del flujo de transcripción automática y análisis clínico posterior.

c. Validación del entorno de ejecución en Python

Finalmente, se verificó la versión de Python instalada y la disponibilidad de las librerías críticas necesarias para el backend local, tales como Flask y los módulos de soporte CORS como se lo puede observar en la siguiente Figura 37 y 38. Esta validación asegura la estabilidad del sistema y la correcta ejecución de los servicios web y de audio.

Figura 37

Validación del entorno de ejecución Python

```

user@raspberrypi: ~
root@raspberrypi:~# # Python y pip
python3 --version
python3 -m pip --version
which python3
python3 -c "import sys; print(sys.executable); print(sys.version)"
Python 3.11.2
pip 23.0.1 from /usr/lib/python3/dist-packages/pip (python 3.11)
/usr/bin/python3
/usr/bin/python3
3.11.2 (main, Apr 28 2025, 14:11:48) [GCC 12.2.0]
root@raspberrypi:~# █

```

Figura 38

Establecimiento de la confirmación de disponibilidad que tienen las dependencias Flask y soporte CORS

```

root@raspberrypi:~# python3 -m pip show flask flask-cors requests numpy
python3 -m pip freeze | egrep -i "flask|cors|requests|numpy|torch|whisper|
true
WARNING: Package(s) not found: numpy
Name: Flask
Version: 2.2.2
Summary: A simple framework for building complex web applications.
Home-page: https://palletsprojects.com/p/flask
Author: Armin Ronacher
Author-email: armin.ronacher@active-4.com
License: BSD-3-Clause
Location: /usr/lib/python3/dist-packages
Requires:
Required-by: Flask-Cors
---
Name: Flask-Cors
Version: 3.0.10
Summary: A Flask extension adding a decorator for CORS support
Home-page: https://github.com/corydolphin/flask-cors
Author: Cory Dolphin
Author-email: corydolphin@gmail.com
License: MIT
Location: /usr/lib/python3/dist-packages
Requires: Flask, Six
Required-by:
---
Name: requests
Version: 2.28.1
Summary: Python HTTP for Humans.
Home-page: https://requests.readthedocs.io
Author: Kenneth Reitz
Author-email: me@kennethreitz.org
License: Apache 2.0
Location: /usr/lib/python3/dist-packages
Requires:
Required-by:
Flask==2.2.2
Flask-Cors==3.0.10
requests==2.28.1
root@raspberrypi:~# █

```

En esta etapa se preparó el entorno Linux del SBC, asegurando compatibilidad con herramientas de grabación y ejecución de scripts Flask.

Preparación de la estructura local**a. Estructuración de directorios locales para grabación**

Una vez validado el subsistema de audio se creó el directorio local `/root/audio_recordings/`, destinado al almacenamiento temporal de los archivos WAV generados durante la grabación

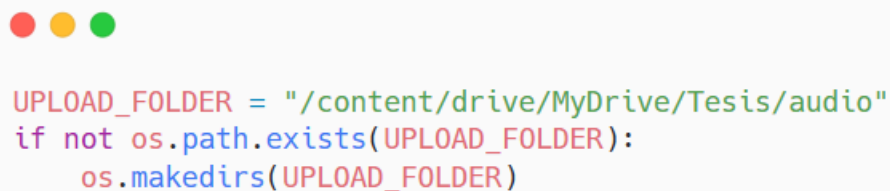
manual. Esta organización permite mantener un registro físico del audio capturado y facilita su posterior envío al backend de procesamiento. Cada archivo es nombrado utilizando un timestamp, asegurando unicidad, trazabilidad y evitando sobrescrituras accidentales.

b. Creación de directorios de almacenamiento de audios y sesiones

Antes de recibir archivos el sistema define y valida directorios de almacenamiento. Si no existen los crea automáticamente, con la finalidad de evitar errores de entrada/salida durante la carga. Como se lo observar mediante la aplicación de las líneas descritas en la Figura 39.

Figura 39

Creación de directorios de almacenamiento de audios y sesiones



```

UPLOAD_FOLDER = "/content/drive/MyDrive/Tesis/audio"
if not os.path.exists(UPLOAD_FOLDER):
    os.makedirs(UPLOAD_FOLDER)

```

Reflejando la estructuración de almacenamiento en Drive o en el filesystem del runtime según la configuración).

1. Endpoint /upload para carga y transcripción automática
2. Se implementa el endpoint REST /upload (HTTP POST) que constituye el punto de entrada del módulo ASR. El flujo interno empieza en realizar una validación de entrada en donde básicamente valida la existencia de un archivo y nombre válido, persistencia enfocada en guardar el audio en el directorio configurado, transcripción mediante la invocación de Whisper con faster-whisper para generar texto y por último una respuesta JSON accedida por el retorno de confirmación + transcripción + metadatos.

Esto se lo puede evidenciar de mejor manera en la Figura 40 en donde se observa el backend procesando la solicitud y en la Figura 41 aparece la confirmación de que el sistema recibió el audio y completó el flujo sin errores.

Figura 40

Proceso Backend de la solicitud

```
● ● ●  
  
@app.route("/upload", methods=["POST"])  
def upload_file():  
    print("processing file.../n")  
    if "file" not in request.files:  
        return jsonify({"error": "No file uploaded"}), 400  
  
    file = request.files["file"]  
    if file.filename == "":  
        return jsonify({"error": "No file selected"}), 400  
  
    # Guardar archivo  
    file_path = os.path.join(app.config['UPLOAD_FOLDER'], file.filename)  
    file.save(file_path)  
  
    print("file saved at:", file_path)  
  
    transcription = applyModel(file_path)  
    return jsonify({"message": "File uploaded successfully", "file_path": file_path,  
"transcriptionModel": transcription, "advice": "Something to work on"})
```

Figura 41

Establecimiento de la confirmación de que el sistema recibió el audio y completó el flujo sin errores

```

@app.route("/diagnostic", methods=["POST"])
def diagnostic():
    request_data = request.get_json()
    content = request_data['content']

    if content:
        client = OpenAI(api_key="xxxxxxxxxxxxxxxxxxxxxxxx", base_url="https://api.deepseek.com")

        response = client.chat.completions.create(
            model="deepseek-chat",
            messages=[
                {"role": "system", "content": "Prompt|."},
                {"role": "user", "content": content},
            ],
            stream=False
        )

        return jsonify({"result": response.choices[0].message.content})
    return jsonify({"result": "Error transcription not present"})

```

Implementación de servicios locales**a. Implementación del servidor web local (server.py)**

La Raspberry Pi fue configurada para ejecutar un servidor web en Python encargado de servir la interfaz gráfica del sistema (index.html). Este servidor opera sobre el puerto 80, permitiendo el acceso directo desde cualquier navegador dentro de la red local. Es importante destacar que en esta capa no se ejecuta ningún proceso de inteligencia artificial, sino más bien su función se limita a entregar la interfaz web y actuar como punto de interacción con el usuario.

b. Implementación de la API de grabación local (file.py)

De forma complementaria se implementó una API de grabación de audio en la Raspberry Pi desplegada en el puerto 5000. Esta API expone endpoints específicos para controlar la grabación local (/start_recording, /pause_recording, /resume_recording, /stop_recording) utilizando el comando arecord. Durante la grabación, el audio se almacena localmente en formato WAV. Al

finalizar el proceso la API devuelve el archivo grabado como respuesta HTTP y adjunta el nombre del archivo mediante el encabezado X-Audio-Filename, permitiendo al frontend reconstruir el archivo y tratarlo como si hubiese sido cargado manualmente.

Verificación Operativa

a. Verificación de servicios activos y puertos de comunicación

Se realizó la validación de los servicios activos en la Raspberry Pi mediante la verificación de puertos en escucha. Los cuales se encuentran evidenciados en la Figura 42 y 43, en donde se observa que los resultados confirman que el servidor web se encuentra operativo en el puerto 80 y que la API de grabación responde correctamente en el puerto 5000, ambos accesibles desde todas las interfaces de red (0.0.0.0).

Figura 42

Comprobación de levantamiento del servidor web local

```

• web-server.service - Web Server
  Loaded: loaded (/etc/systemd/system/web-server.service; enabled; preset: enabled)
  Active: active (running) since Thu 2026-02-05 12:58:55 -05; 2 days ago
  Main PID: 552 (python3)
  Tasks: 1 (limit: 454)
  CPU: 972ms
  CGroup: /system.slice/web-server.service
          └─552 /usr/bin/python3 /root/server.py

Feb 05 12:58:57 raspberrypi python3[552]: * Running on all addresses (0.0.0.0)
Feb 05 12:58:57 raspberrypi python3[552]: * Running on http://127.0.0.1:80
Feb 05 12:58:57 raspberrypi python3[552]: * Running on http://127.0.0.1:80
Feb 05 12:58:57 raspberrypi python3[552]: Press CTRL+C to quit
Feb 07 20:43:37 raspberrypi python3[552]: 192.168.137.1 - - [07/Feb/2026 20:43:37] "GET / HTTP/1.1" 200 -
Feb 07 20:43:39 raspberrypi python3[552]: 192.168.137.1 - - [07/Feb/2026 20:43:39] "GET /favicon.ico HTTP/1.1" 404 -
Feb 07 21:04:08 raspberrypi python3[552]: 127.0.0.1 - - [07/Feb/2026 21:04:08] "HEAD / HTTP/1.1" 200 -
Feb 07 21:08:49 raspberrypi python3[552]: 127.0.0.1 - - [07/Feb/2026 21:08:49] "HEAD / HTTP/1.1" 200 -
Feb 07 21:11:47 raspberrypi python3[552]: 127.0.0.1 - - [07/Feb/2026 21:11:47] "HEAD / HTTP/1.1" 200 -
Feb 07 21:12:56 raspberrypi python3[552]: 127.0.0.1 - - [07/Feb/2026 21:12:56] "HEAD / HTTP/1.1" 200 -

```

Figura 43

Establecimiento de la comprobación del levantamiento del audio API en el modo archivo

(file.py)

```

root@raspberrypi:~# systemctl status api.service --no-pager -l
systemctl status web-server.service --no-pager -l
● api.service - Audio API
   Loaded: loaded (/etc/systemd/system/api.service; enabled; preset: enabled)
   Active: active (running) since Thu 2026-02-05 12:58:55 -05; 2 days ago
     Main PID: 549 (python3)
       Tasks: 3 (limit: 454)
          CPU: 15.581s
    CGroup: /system.slice/api.service
            └─549 /usr/bin/python3 /root/file.py
              └─576 /usr/bin/python3 /root/file.py

Feb 05 12:58:57 raspberrypi python3[549]: * Running on all addresses (0.0.0.0)
Feb 05 12:58:57 raspberrypi python3[549]: * Running on http://127.0.0.1:5000
Feb 05 12:58:57 raspberrypi python3[549]: * Running on http://127.0.0.1:5000
Feb 05 12:58:57 raspberrypi python3[549]: Press CTRL+C to quit
Feb 05 12:58:57 raspberrypi python3[549]: * Restarting with stat
Feb 05 12:58:57 raspberrypi python3[576]: * Debugger is active!
Feb 05 12:58:57 raspberrypi python3[576]: * Debugger PIN: 214-242-623
Feb 07 21:09:04 raspberrypi python3[576]: 127.0.0.1 - - [07/Feb/2026 21:09:04] "GET /recording_status HTTP/1.1" 200 -
Feb 07 21:12:08 raspberrypi python3[576]: 127.0.0.1 - - [07/Feb/2026 21:12:08] "GET /recording_status HTTP/1.1" 200 -
Feb 07 21:12:56 raspberrypi python3[576]: 127.0.0.1 - - [07/Feb/2026 21:12:56] "GET /recording_status HTTP/1.1" 200 -

```

Asimismo, se comprobó mediante solicitudes HTTP que ambos servicios responden con estado 200 OK, confirmando su disponibilidad y correcto funcionamiento.

b. Integración con el backend en la nube

Una vez completada la implementación local del hardware se validó la conectividad de la Raspberry Pi hacia el backend desplegado en Google Colab como se lo puede observar en la siguiente Figura 44. Esta integración permite que los audios capturados localmente sean enviados al entorno en la nube para su almacenamiento persistente en Google Drive y su posterior procesamiento mediante modelos de inteligencia artificial.

Figura 44

Comprobación de conectividad entre Raspberry hacia a Google Colab

```

root@raspberrypi:~# API_URL="https://psicotranslate.ngrok.app"

# 1) Verifica que el endpoint /patients responde
curl -s -H "ngrok-skip-browser-warning: true" "$API_URL/patients" | head -c 300 ; echo
[{"id":0,"lastVisit":"2025-11-25","name":"Test real 1","sessions":[{"advice":"EVALUACI\u00d3N PSICOL\u00d3GICA\nIdentificaci\u00f3n: Paciente universitaria, g\u00e9nero femenino, edad no especificada.\nMotivo de consulta: Problemas psicol\u00f3gicos derivados de conflicto con docente universitario,
root@raspberrypi:~# █

```

Una vez ya implementado tanto las dependencias como las situaciones necesarias de funcionamiento se estableció dos formas de uso dentro del Raspberry. Es decir, que el sistema se compone de dos capas de ejecución:

1. Una capa local en la Raspberry Pi para la interfaz web y/o captura de audio
2. Una capa en la nube (Google Colab expuesto vía API_URL/ngrok) que orquesta el almacenamiento en Google Drive y el procesamiento con IA (ASR y análisis clínico).

Una vez preparada y terminada la configuración por parte del Raspberry Pi se procedió a realizar las configuraciones respectivas en cuando al backend y frontend del sistema como ya se lo estableció en el capítulo del diseño.

b) ***Backend de procesamiento (Colab – Flask + Drive + Modelos IA)***. Esta sección corresponde a la implementación del núcleo lógico del sistema.

Preparación del entorno en la nube

a. Preparación del entorno de ejecución en Google Colab

El desarrollo y ejecución inicial del backend se realizó en Google Colab, debido a que permite instalar dependencias rápidamente y aprovechar recursos de cómputo (incluida GPU) para acelerar la transcripción y el procesamiento con modelos de IA. En la siguiente Figura 45 se muestra el notebook donde se ejecuta el bloque inicial de configuración del entorno, en donde se realizan tres acciones críticas que van:

Figura 45

Bloque inicial de configuración del entorno de Google Colab



```

!pip install faster-whisper openai flask-cors
!pip install nvidia-cublas-cu12 nvidia-cudnn-cu12==9.*
!export LD_LIBRARY_PATH=`python3 -c 'import os; import nvidia.cublas.lib; import nvidia.cudnn.lib;
print(os.path.dirname(nvidia.cublas.lib.__file__) + ":" + os.path.dirname(nvidia.cudnn.lib.__file__))'`
!pip install pyngrok
!ngrok config add-authtoken xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx

```

- Instala dependencias para transcripción y servidor web como faster-whisper (ASR), Flask (API), Flask-CORS (interoperabilidad) y el cliente compatible con interfaz OpenAI para consumo del LLM.
- Configura librerías de aceleración (cuBLAS/cuDNN y variables de entorno) con la finalidad de habilitar condiciones óptimas para inferencia en GPU cuando está disponible.
- Instala y configura ngrok (pyngrok + authtoken) para exponer el servidor local del notebook a Internet mediante HTTPS.

b. Montaje de Google Drive como almacenamiento persistente

Para evitar el uso de almacenamiento temporal y que se pierda al reiniciar el runtime el sistema monta Google Drive con el comando `drive.mount('/content/drive')` como se lo observa en la Figura 46.

Figura 46

Instrucción para establecer Google Drive como almacenamiento



```

from google.colab import drive
drive.mount('/content/drive')

```

Con la finalidad de formar una solicitud de permisos como se observar en la Figura 47, la cual es necesaria habilitar lectura/escritura desde el notebook hacia el Drive del usuario.

Figura 47*Permiso de notebook acceda a Google Drive***Permit this notebook to access your Google Drive files?**

This notebook is requesting access to your Google Drive files. Granting access to Google Drive will permit code executed in the notebook to modify files in your Google Drive. Make sure to review notebook code prior to allowing this access.

No thanks [Connect to Google Drive](#)

En cual mediante la autorización correspondiente Google Drive queda accesible como un directorio del sistema destinado para almacenar audios, mantener plantillas o recursos de interfaz y finalmente guardar sesiones con los correspondientes resultados.

c. Importación de librerías y dependencias del Backend

En el siguiente bloque se importan las dependencias clave del sistema, como se evidencia en la Figura 48.

Figura 48*Importación de librerías y dependencias del Backend*

```

from flask import Flask, render_template, request, jsonify
from faster_whisper import WhisperModel
from openai import OpenAI
from flask_cors import CORS, cross_origin # LOCAL
import os
import threading
import requests
import json

```

Estableciendo como base de la arquitectura de implementación el API REST + transcripción + análisis + persistencia.

Configuración del servidor Flask

a. Inicialización del servidor Flask y habilitación de CORS

En este punto se puede mencionar que se inicializa la aplicación Flask y se configura la ruta de plantillas HTML (templates) separando la capa de presentación (HTML/JS) y la capa de aplicación (endpoints y lógica). Esto se puede apreciar a través de la Figura 49 y de forma adicional se habilita CORS para permitir que el frontend consuma la API, aunque estén en distintos puertos/dominios considerada como una condición típica al usar ngrok.

Figura 49

Inicialización del servidor Flask y habilitación de CORS



```
app = Flask(__name__, template_folder='/content/drive/MyDrive/Tesis/templates/')
#port = 5000
cors = CORS(app)
```

b. Ejecución del servidor con Colab sin bloquear el notebook

Debido a que Colab es un entorno interactivo el servidor Flask se ejecuta en un hilo (thread) para evitar que la ejecución bloquee el resto del notebook como se lo puede observar en la siguiente Figura 50.

Figura 50

Ejecución del servidor en segundo plano del entorno notebook



```
@app.route("/")
def home():
    return render_template("index.html")
threading.Thread(target=app.run, kwargs={'host':'0.0.0.0', 'port':7000}).start()
```

En lo que respecta a la definición en el ámbito de red es la de que el host=0.0.0.0 permite que el servidor escuche en todas las interfaces de red disponibles del entorno y con ello que su

acceso esté permitido desde fuera del entorno local cuando se utilizan mecanismos de exposición como son los túneles como puede ser el ngrok, en fin el puerto queda definido por: `port=7000` y con ello queda definido el puerto de escucha, estableciendo así el punto específico en el que se atienden las peticiones HTTP que van llegando.

Implementación del pipeline de procesamiento

a. Función de transcripción con Whisper

La función de transcripción implementa el uso del modelo Whisper large-v2 vía faster-whisper con configuración orientada a eficiencia. En donde el modelo devuelve la transcripción en segmentos temporales como se lo puede observar en la siguiente Figura 51, que luego se consolidan en una sola cadena de texto para:

- Visualización en la interfaz
- Envío al módulo de análisis clínico
- Almacenamiento en sesión.

Figura 51

Transcripción con Faster-Whisper

```
def applyModel(filePath):
    # Carga Whisper en GPU para acelerar transcripción
    model = WhisperModel("large-v2", device="cuda", compute_type="float16")

    # Transcribe audio en español
    result, _ = model.transcribe(filePath, beam_size=5, language="es")

    # Une segmentos en un texto final
    return "".join(seg.text for seg in result)
```

b. Endpoint /diagnostic para análisis clínico con LLM

Se implementa el endpoint /diagnostic el cual recibe el texto (transcripción) y solicita al LLM un informe estructurado. El funcionamiento se basa en:

- Un mensaje “system” que obliga al modelo a comportarse como “psicólogo clínico”
- Reglas explícitas sobre el contenido del informe, es decir presentar una evaluación textual, diagnóstico presuntivo (CIE-11/DSM-5), diagnóstico diferencial, gravedad, factores, plan por sesiones, recomendaciones e indicadores de riesgo,
- Un mensaje “user” con la transcripción real (content).

Esto se lo puede evidenciar claramente en la Figura 52 que corresponde al preflight CORS.

Luego se ejecuta la petición principal y se devuelve el informe como JSON para que el frontend lo muestre.

Figura 52

Endpoint /diagnostic para análisis clínico con LLM

```
response = client.chat.completions.create(
  model="deepseek-chat",
  messages=[
    {"role": "system", "content": "Actúa como un psicólogo clínico. Analiza el texto (conversaci
    {"role": "user", "content": content},
```

Mientras que la parte de role: “user”, content hace referencia al texto recibido en el JSON del request, para que el sistema interprete el texto real sobre la transcripción mediante el modelo de análisis. Consecuentemente el modelo genera el informe aplicando las reglas definidas en el prompt del role system sobre el contenido para que posteriormente sea enviado al cliente en formato JSON.

En otras palabras, como entrada se inicia el texto de la conversación dentro del content proporcionado por el usuario o el módulo de transcripción, seguido de la aplicación del prompt de sistema, el cual se encarga de definir la estructura, los criterios del informe y su ejecución del análisis por parte del modelo de lenguaje. Finalmente, se tiene considerado como salida el informe en lenguaje natural, organizado según los elementos solicitados como evaluación psicológica textual, diagnóstico presuntivo, diagnóstico diferencial, nivel de gravedad, factores asociados, plan

de tratamiento por sesiones y recomendaciones prácticas, todo esto estructurado como respuesta JSON. Dentro de este contexto se puede mencionar que el informe que se obtiene es solamente un análisis asistido por IA basado en información textual exclusiva, por lo que se lo debe considerar como un apoyo orientativo para la comprensión del contenido y no como un reemplazo de una evaluación clínica profesional, la cual requiere entrevista, contexto, historia clínica y juicio experto. Esta delimitación contribuye a la validez metodológica del estudio y reduce el riesgo de interpretaciones indebidas del sistema.

c. Gestión de pacientes y sesiones (almacenamiento en JSON)

Dentro de este aspecto se puede mencionar que se implementa un endpoint orientado a la administración y recuperación de información clínica mediante un esquema de almacenamiento basado en el sistema de archivos, la cual se encuentra comprendida mediante una estructura de datos organizada en un directorio raíz (sessions/), dentro del cual se crea una carpeta por paciente. En cada carpeta se almacenan uno o más archivos en formato JSON (.json) que contienen la información correspondiente a sus sesiones o registros asociados. Esto se lo puede observar claramente en la siguiente Figura 53 y Figura 54 en donde se aprecia la implementación de los endpoints para listar los pacientes por sesiones y registrar nuevas sesiones para permitir una mejor trazabilidad y consulta histórica sin requerir base de datos en fase inicial.

Figura 53*Almacenamiento JSON de pacientes y sesiones*

```

● ● ●
@app.route("/patients", methods=["GET"])
def get_patients():
    sessions_dir = "/content/drive/MyDrive/Tesis/sessions"
    patients = [folder for folder in os.listdir(sessions_dir) if os.path.isdir(os.path.join(sessions_dir,
folder))]
    print(patients)

    patients_data = []
    for patient in patients:
        patient_dir = os.path.join(sessions_dir, patient)
        session_files = os.listdir(patient_dir)
        for file in session_files:
            if file.endswith(".json"):
                with open(os.path.join(patient_dir, file), "r") as f:
                    data = json.load(f)
                    patients_data.append(data)

```

Figura 54*Código de agrupación para los registros por paciente*

```

● ● ●
@app.route("/patients", methods=["GET"])
def get_patients():
    sessions_dir = "/content/drive/MyDrive/Tesis/sessions"
    patients = [folder for folder in os.listdir(sessions_dir) if os.path.isdir(os.path.join(sessions_dir,
folder))]
    print(patients)

    patients_data = []
    for patient in patients:
        patient_dir = os.path.join(sessions_dir, patient)
        session_files = os.listdir(patient_dir)
        for file in session_files:
            if file.endswith(".json"):
                with open(os.path.join(patient_dir, file), "r") as f:
                    data = json.load(f)
                    patients_data.append(data)

```

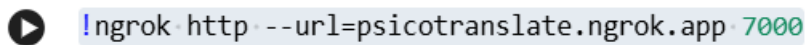
Este diseño permite mantener los registros agrupados por paciente, facilitando la consulta posterior por parte de otros endpoints y reduciendo la complejidad al evitar dependencias iniciales con sistemas de bases de datos durante la etapa de prototipado o validación del sistema. En la siguiente Figura 55 se puede observar cómo se registra una sesión si cumple con todos los requerimientos.

Figura 55*Condicionamiento de registro de sesión*


```
return jsonify({"message": "Sessions saved successfully"})
```

Publicación del servicio**a. Publicación del servicio mediante ngrok**

Para permitir acceso externo al backend se crea un túnel con ngrok hacia el puerto del servidor (7000). En la Figura 56 se muestra el endpoint público y su asociación con el agent.

Figura 56*Túnel seguro con ngrok*


```
!ngrok http --url=psicotranslate.ngrok.app 7000
```

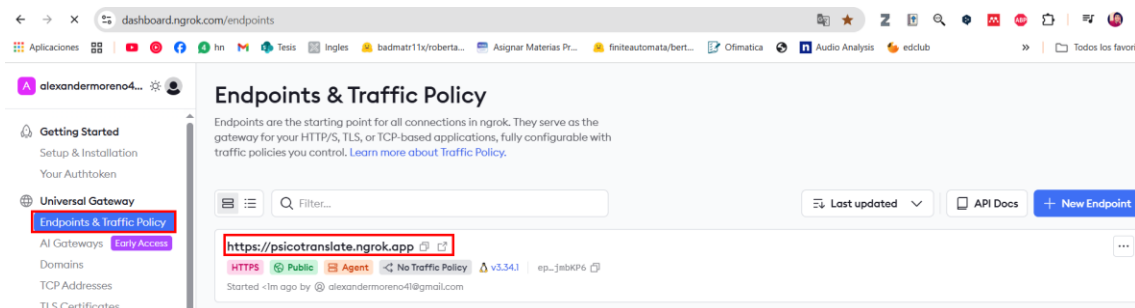
Esta configuración posibilita que la interfaz web u otros clientes externos consuman la API desde internet, aun cuando el servidor se ejecute en un entorno que normalmente no es accesible directamente desde fuera. En términos prácticos, ngrok actúa como un puente entre el servidor local y un endpoint público facilitando la validación funcional, pruebas de integración y demostraciones del prototipo.

b. Verificación de la creación del endpoint ocasionado por el túnel ngrok

En la etapa de despliegue como se utilizó ngrok para publicar el servicio local a través de un túnel seguro. Desde el panel Endpoints & Traffic Policy se verificó su creación del endpoint <https://psicotranslate.ngrok.app> con protocolo HTTPS y visibilidad pública. El endpoint se encuentra asociado al Agent responsable de mantener el túnel activo entre el host local y la infraestructura de ngrok. La configuración se mantuvo sin reglas adicionales de control o filtrado de tráfico (No Traffic Policy) como se lo puede observar en la siguiente Figura 57.

Figura 57

Publicación del servicio local por el túnel ngrok

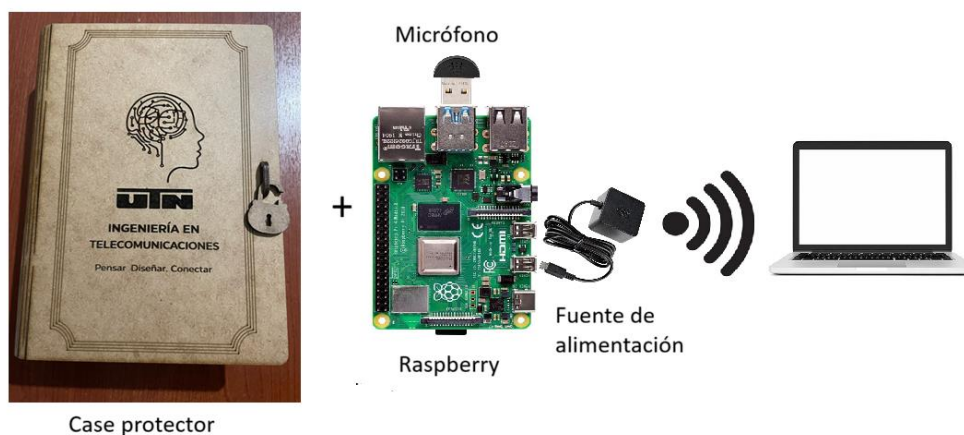


3.2.2. Hardware

En la siguiente Figura 58 se puede observar desde el punto de vista tecnológico, el prototipo integra un ordenador de placa única (SBC) como unidad central de procesamiento junto con un micrófono de alta sensibilidad para la adquisición de audio en entornos terapéuticos reales, además de contener todo lo mencionado con el case diseñado en material MDF.

Figura 58

Implementación completa del hardware del sistema



Sobre esta infraestructura se implementaron módulos de software desarrollados principalmente en Python, aprovechando bibliotecas especializadas para reconocimiento automático de voz, procesamiento de lenguaje natural y análisis semántico, lo que permitió la

conversión eficiente del contenido auditivo a texto y la posterior generación de información clínica relevante. De forma más visual en la siguiente Figura 59 se detalla la unificación de todo el sistema en uno con el case protector.

Figura 59

Implementación completa del hardware del sistema desde otra perspectiva visual



CAPÍTULO IV. PRUEBAS DE FUNCIONAMIENTO

El presente capítulo describe las pruebas de funcionamiento realizadas al prototipo electrónico para la automatización y análisis inteligente de notas clínicas en sesiones terapéuticas basado en inteligencia artificial. El objetivo principal de estas pruebas es verificar el correcto desempeño del sistema, validar el cumplimiento de los requerimientos definidos en los capítulos anteriores y comprobar que la solución propuesta responde de manera adecuada a las necesidades identificadas en el contexto clínico-terapéutico. Las pruebas se enfocan básicamente en la validación del diagnóstico proporcionado por el prototipo en correlación con el criterio por parte de diferentes profesionales del área. Dentro de este contexto se realiza una encuesta evaluativa (Anexo C) dirigida hacia a profesionales del área de Psicología, docentes del área de Psicología y estudiantes de Psicología para categorizar el nivel de confiabilidad del funcionamiento del sistema y consecuentemente realizar el análisis respectivo sobre cada uno de los argumentos proporcionados.

4.1. Pruebas Preliminares

Previo a la implementación del prototipo en sesiones terapéuticas con pacientes reales, se realizaron pruebas preliminares con el objetivo de corroborar el correcto funcionamiento integral del sistema desarrollado. Estas pruebas tuvieron como finalidad validar la operatividad de los componentes de hardware y software, verificar la estabilidad del proceso de captura y transcripción de audio, evaluar la generación de análisis clínico mediante el modelo de inteligencia artificial seleccionado y comprobar la correcta gestión del almacenamiento seguro de la información. Las pruebas fueron ejecutadas en un entorno controlado dentro de la Universidad Técnica del Norte, específicamente en el edificio de Bienestar Estudiantil, en el Departamento de Psicología, para su validación se puede observar el Anexo D.

Tabla 21*Prueba preliminar de sesión*

ID	Tipo de Prueba	Duración	Escenario	Objetivo General	Flujo Evaluado	Métricas Técnicas Evaluadas	Resultado Esperado
PRE-INT-01	Prueba Integral End-to-End (E2E)	5–10 min	Simulación de sesión terapéutica completa	Validar el funcionamiento integral del sistema como ecosistema tecnológico, comprobando interoperabilidad entre módulo de captura (Raspberry Pi 4), backend desplegado en Google Colab con GPU, modelo ASR (Whisper), modelo de IA para análisis clínico, generación de informe estructurado y almacenamiento seguro en	<ol style="list-style-type: none"> 1. Autenticación 2. Captura en tiempo real / carga manual de audio 3. Transcripción automática 4. Procesamiento semántico 5. Extracción de síntomas 6. Detección de indicadores de riesgo 	<ul style="list-style-type: none"> • Tiempo total de procesamiento • Integridad del archivo (100% conservación audio) • Coherencia semántica (sin omisiones críticas) • Precisión en extracción de síntomas 	Flujo completo ejecutado sin fallos críticos, transcripción coherente, análisis clínico estructurado, generación correcta del informe preliminar, almacenamiento seguro persistente y correcta visualización en dashboard del paciente.

ID	Tipo de Prueba	Duración	Escenario	Objetivo General	Flujo Evaluado	Métricas Técnicas Evaluadas	Resultado Esperado
				Google Drive con vinculación automática en dashboard.	7. Generación de informe estructurado 8. Exportación JSON 9. Almacenamiento en Drive1 10. Vinculación al paciente en dashboard	<ul style="list-style-type: none"> • Generación completa del informe • Persistencia correcta en Drive y JSON • Estabilidad sin errores críticos • Disponibilidad del endpoint (Ngrok activo) 	

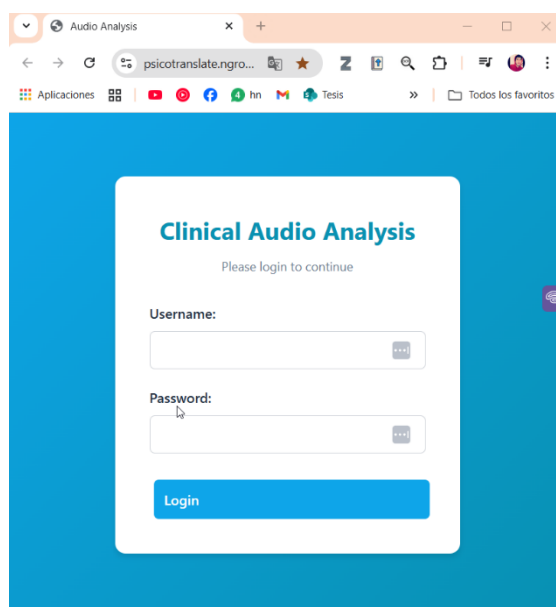
Como se pudo observar en la Tabla 21 para validar el funcionamiento global del prototipo como un sistema integrado y no como módulos aislados, se ejecutó una prueba integral End-to-End (E2E), diseñada para comprobar la interoperabilidad, estabilidad y coherencia funcional de todos los componentes del ecosistema tecnológico en un único flujo operacional continuo (PRE-INT-01). A diferencia de pruebas unitarias o fragmentadas, esta validación reproduce el escenario real de uso clínico desde el inicio de sesión hasta el almacenamiento final estructurado, permitiendo evaluar el comportamiento sistémico completo bajo condiciones controladas.

4.1.1. Autenticación

La prueba inició con el proceso de autenticación del terapeuta mediante el sistema de login implementado en el dashboard web. Se validó el correcto funcionamiento del mecanismo de control de acceso, comprobando que las credenciales ingresadas fueran verificadas contra el sistema de autenticación configurado en el backend.

Figura 60

PRE-INT-01- Autenticación del terapeuta



Observaciones:

- El acceso fue concedido únicamente con credenciales válidas.
- No se evidenciaron errores de sesión ni vulnerabilidades visibles.
- El tiempo de respuesta del endpoint fue estable menor a 2 segundos aproximadamente.
- Se confirmó que el sistema mantiene sesión activa sin interrupciones durante todo el flujo.

4.1.2. Captura en tiempo real/Carga manual de audio

Posteriormente, se ejecutó la captura de audio en tiempo real mediante la Raspberry Pi 4 y el micrófono USB seleccionado. De forma paralela, también se validó el mecanismo alternativo de carga manual de archivo desde el repositorio.

Figura 61

PRE-INT-01- Captura de audio en tiempo real

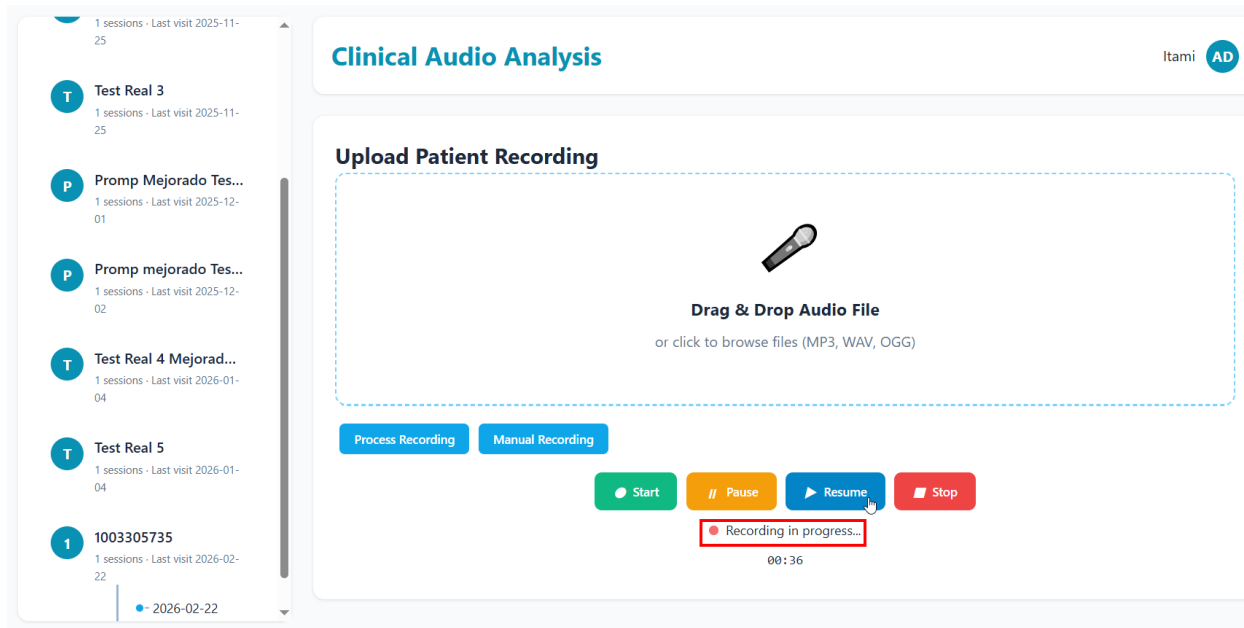


Figura 62

PRE-INT-01 - Carga manual de audio

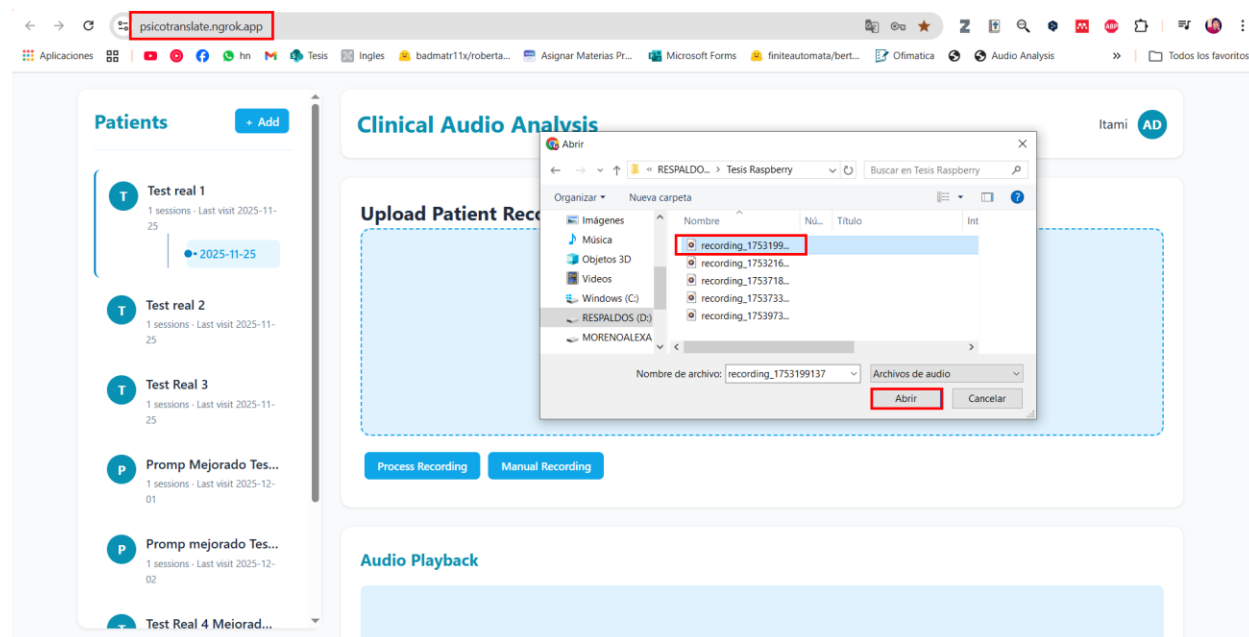
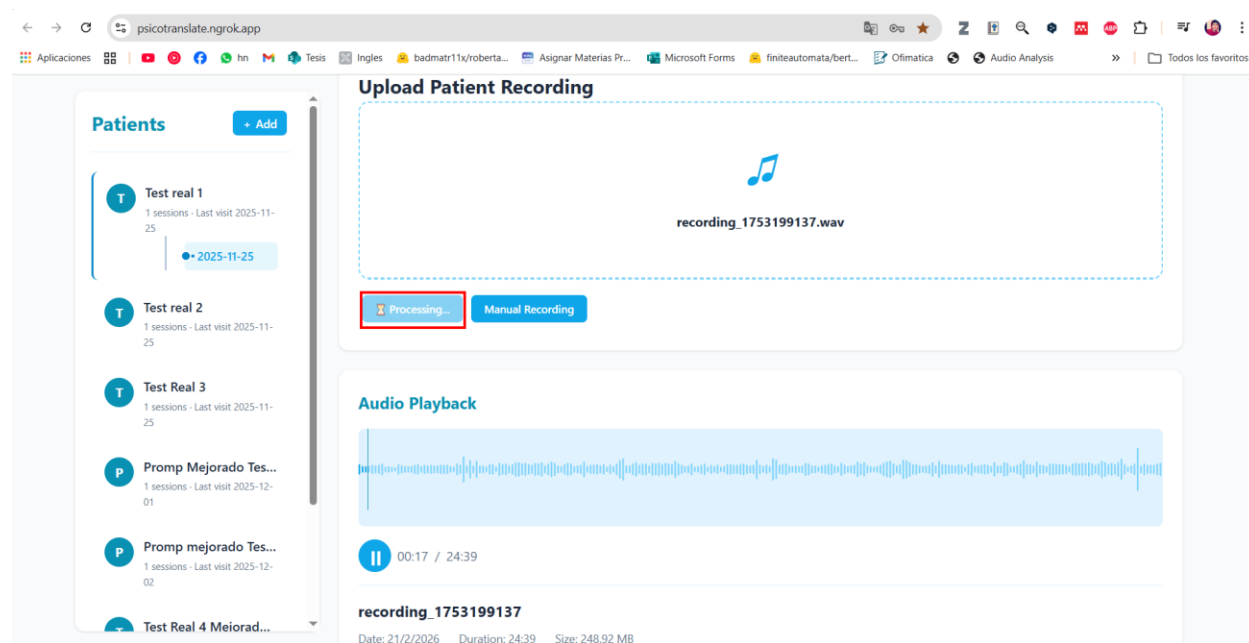


Figura 63

PRE-INT-01 - Verificación de la carga del audio



Observaciones:

- La señal de audio fue capturada sin distorsionar la señal de audio ni en su totalidad.
- El archivo WAV tiene el 100% de su integridad.
- Para el caso de cargar manualmente un archivo de audio, el formato del mismo fue reconocido correctamente con el comportamiento esperado de habilitar dinámicamente los botones e ir habilitándolos alrededor del archivo que está siendo cargado.
- No se presentaron errores de compatibilidad de archivos ni errores de lectura.

4.1.3. Transcripción Automática

Una vez recibido el audio, el sistema activó el modelo de reconocimiento automático de voz (Whisper) desplegado en Google Colab con aceleración GPU. Se evaluó la correcta conversión del audio a texto en lenguaje natural.

Figura 64

PRE-INT-01 - Transcripción Automática

The screenshot displays the 'psicotranslate.ngrok.app' interface. On the left, a 'Patients' sidebar lists several sessions, including 'Test real 1' through 'Test Real 4 Meiorad...'. The main area features a 'Transcription' section with a large text block containing a detailed transcription of a patient's account. Below this, the 'AI Clinical Insights' section provides a structured summary of the patient's psychological evaluation, including identification, consultation motivation, and reported symptoms. A 'Generate Insights' button is highlighted with a red box in this section.

Observaciones:

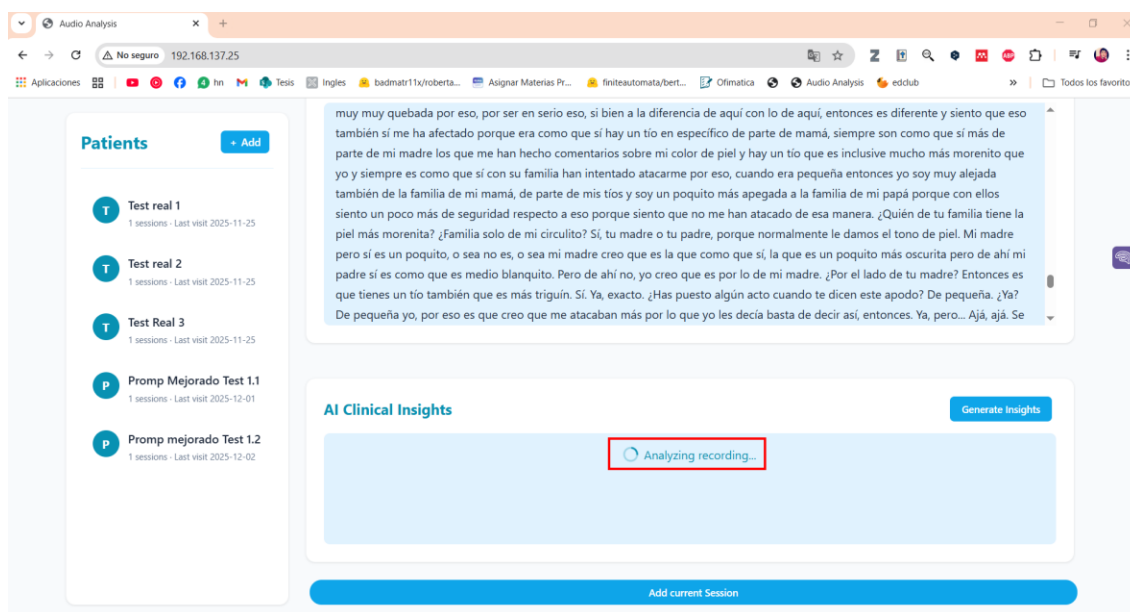
- El tiempo de procesamiento fue eficiente para una sesión de 5–10 minutos.
- La transcripción presentó coherencia sintáctica adecuada.
- No se detectaron omisiones críticas de contenido.
- Se mantuvo correspondencia adecuada entre segmentos hablados y texto generado.

4.1.4. *Procesamiento Semántico*

Sobre el texto transcrito, se ejecutó el módulo de análisis semántico basado en el modelo de inteligencia artificial seleccionado. Esta etapa permitió interpretar el contenido desde una perspectiva clínica estructurada.

Figura 65

PRE-INT-01 - Procesamiento Semántico



Observaciones:

- En estos aspectos se puede mencionar que el modelo logró contextualizar adecuadamente el audio.

- Se pudo llegar a identificar los diferentes patrones tanto en los aspectos emocionales como los temáticos.
- No se generaron respuestas fuera de contexto.
- El procesamiento mantuvo coherencia clínica preliminar.

4.1.5. Extracción de Síntomas

El sistema procedió a identificar síntomas psicológicos y somáticos presentes en el discurso del paciente mediante análisis estructurado.

Figura 66

PRE-INT-01 - Extracción de Síntomas

Evaluación Psicológica Basada en el Contenido Textual

1. Emociones expresadas:

- Dolor, tristeza profunda, humillación, soledad, abandono.
- Miedo al rechazo, a la intimidad y a la vulnerabilidad.
- Culpa (por comer, por existir, por defenderse).
- Rabia contenida y resentimiento hacia figuras de autoridad (padres, nutrióloga, agresores).
- Desempeño pasado (edificación suicida a los 10 años).
- Baja autoestima y autodesprecio focalizado en el cuerpo (peso, color de piel, rostro).

2. Distorsiones cognitivas identificadas:

- **Sobregeneralización:** "Por mi físico me alejan todos", "no merezco tener pareja".
- **Filtro mental/Descalificación de lo positivo:** A pesar de que la ropa le queda fija y médicos indican peso normal, solo se fija en el espejo donde se "ve gorda".
- **Personalización:** Asume que el maltrato fue por sus características físicas, sin cuestionar la crueldad de los otros.
- **Pensamiento dicotómico:** "Como demasiado" vs. "debo dejar de comer", visión de sí misma como "gorda" a pesar de evidencia contraria.
- **Razonamiento emocional:** "Me siento gorda, por lo tanto, lo soy", "siento que no merezco pareja, entonces no la tengo".

3. Patrones de pensamiento:

- Hipervigilancia al rechazo y a la evaluación social.
- Autoconciencia severa y focalización obsesiva en la imagen corporal.
- Expectativas de abandono/desprotección por parte de figuras de apoyo (padres).

4. Conductas implicadas:

- Evitación social histórica y actual (solo se abre con 2 compañeros tras ganar confianza).
- Conductas compensatorias por trauma: agresión reactiva en la infancia (gubión de defensa hostilidad).
- Restricción alimentaria severa en el pasado, con episodios de vómito autoinducido.
- Oculta sus emociones y necesidades a la familia, muestra una fachada de "normalidad" alimentaria frente a ellos.
- Evitación de conflictos (ejemplo: esquivó a la madre del compañero).

5. Preocupaciones centrales:

- Imagen corporal (peso, color de piel).
- Miedo al rechazo interpersonal.
- Dificultad para confiar y mostrar afecto.
- Relación distante y poco comprensiva con los padres.
- Culpa persistente asociada a la alimentación.

6. Rangos de personalidad sugeridos:

- Alta evitación del daño (evitación de situaciones sociales o conflictos).
- Baja autoeficacia (inseguridad, autoeficacia disminuida).
- Posibles rasgos de desconfianza/desapego como secuela del maltrato prolongado.

Diagnóstico Presuntivo Clínico (CIE-11 / DSM-5)

A. Trastorno de Estrés Posttraumático (CIE-11: 6B40 / DSM-5: 309.81)

- **Criterios resumidos (DSM-5):**

Observaciones:

- Se extrajeron síntomas relevantes alineados con el contenido textual.
- La categorización fue consistente con expresiones detectadas en la sesión.
- No se evidenciaron falsos positivos críticos.
- El sistema estructuró los síntomas en formato organizado.

4.1.6. *Detección de Indicadores de Riesgo*

En esta fase, el sistema analizó posibles indicadores de riesgo (por ejemplo: ideación autolesiva, ansiedad severa, desesperanza marcada) mediante patrones lingüísticos y semánticos.

Figura 67

PRE-INT-01 - Detección de Indicadores de Riesgo

The screenshot displays a web application interface for 'Audio Analysis'. On the left, there is a 'Patients' sidebar with a list of sessions: 'Test real 1', 'Test real 2', 'Test Real 3', 'Promp Mejorado Test 1.1', and 'Promp mejorado Test 1.2'. The main content area is titled 'Indicadores de Riesgo y Pasos Sugeridos'. It is divided into two sections: 'Indicadores de riesgo:' and 'Pasos sugeridos:'. The 'Indicadores de riesgo:' section lists three items: 1. 'Riesgo suicida: Historia de intento a los 10 años. Actualmente no expresa ideación activa, pero el estado de ánimo depresivo y la baja autoestima son factores de vulnerabilidad.' 2. 'Riesgo de recaída en conductas alimentarias: La culpa persistente por comer y la distorsión de la imagen corporal son señales de alerta.' 3. 'Aislamiento social: Aunque tiene dos vínculos, la tendencia al retraimiento puede agravar la depresión.' The 'Pasos sugeridos:' section lists five items: 1. 'Evaluar riesgo suicida de manera regular en sesiones: preguntar directamente por pensamientos, planes o intenciones.' 2. 'Elaborar un plan de seguridad por escrito, con contactos de emergencia y actividades de calma.' 3. 'Derivación a psiquiatría para evaluación farmacológica coadyuvante (antidepresivos ISRS pueden ayudar en TEPT y depresión).' 4. 'Coordinar con médico/nutricionista sensible al trauma para seguimiento del peso y hábitos alimentarios, evitando enfoques punitivos.' 5. 'Involucrar a los padres en psicoeducación en una sesión futura, si Darly está de acuerdo, para ampliar la red de apoyo y reducir comentarios críticos en casa.' At the bottom, there is a 'Nota final:' section stating: 'Darly muestra una gran valentía al compartir su historia por primera vez. Su capacidad para establecer dos vínculos de confianza en la universidad es un signo positivo de resiliencia. El tratamiento debe priorizar la validación, el empoderamiento y la reconstrucción de una identidad separada del trauma.'

Observaciones:

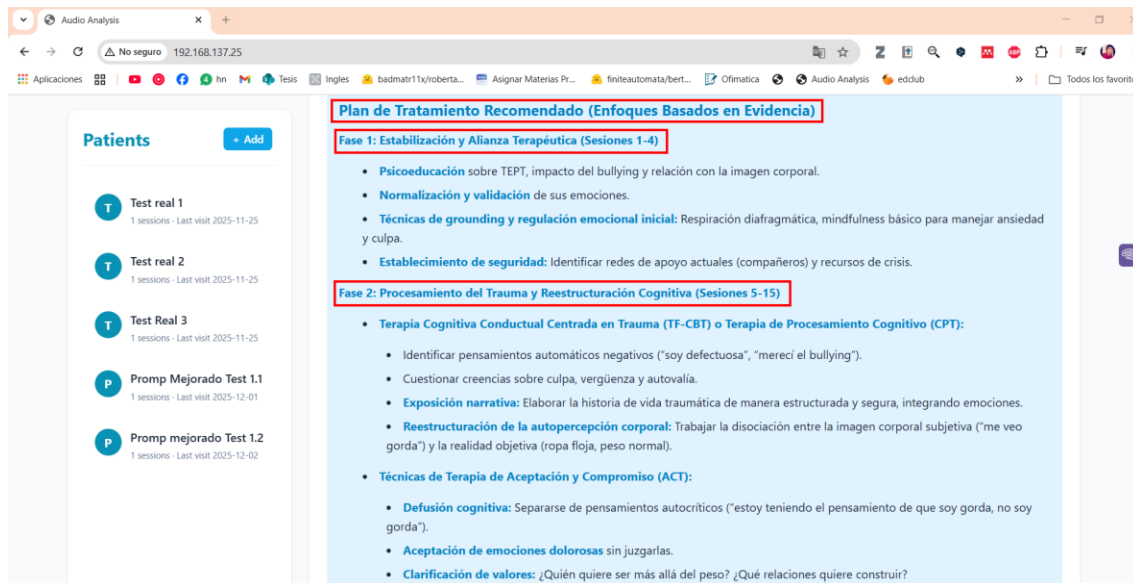
- El sistema logró identificar expresiones potencialmente sensibles.
- No generó alarmas innecesarias en contenido neutro.
- La detección se presentó como sugerencia clínica, no como diagnóstico definitivo.
- Se confirmó que el modelo funciona como apoyo y no sustituye el criterio profesional.

4.1.7. Generación de Informe Estructurado

Con base en el análisis previo, el sistema generó automáticamente un informe preliminar estructurado, organizado en secciones clínicas.

Figura 68

PRE-INT-01 - Generación de Informe Estructurado



Observaciones:

- El informe fue generado completamente sin errores de formato.
- Presentó estructura lógica (motivo de consulta, síntomas, análisis, posibles indicadores).
- La redacción fue coherente y clínicamente comprensible.
- No se detectaron fragmentos inconexos o repetitivos.

4.1.8. Exportación en Formato JSON

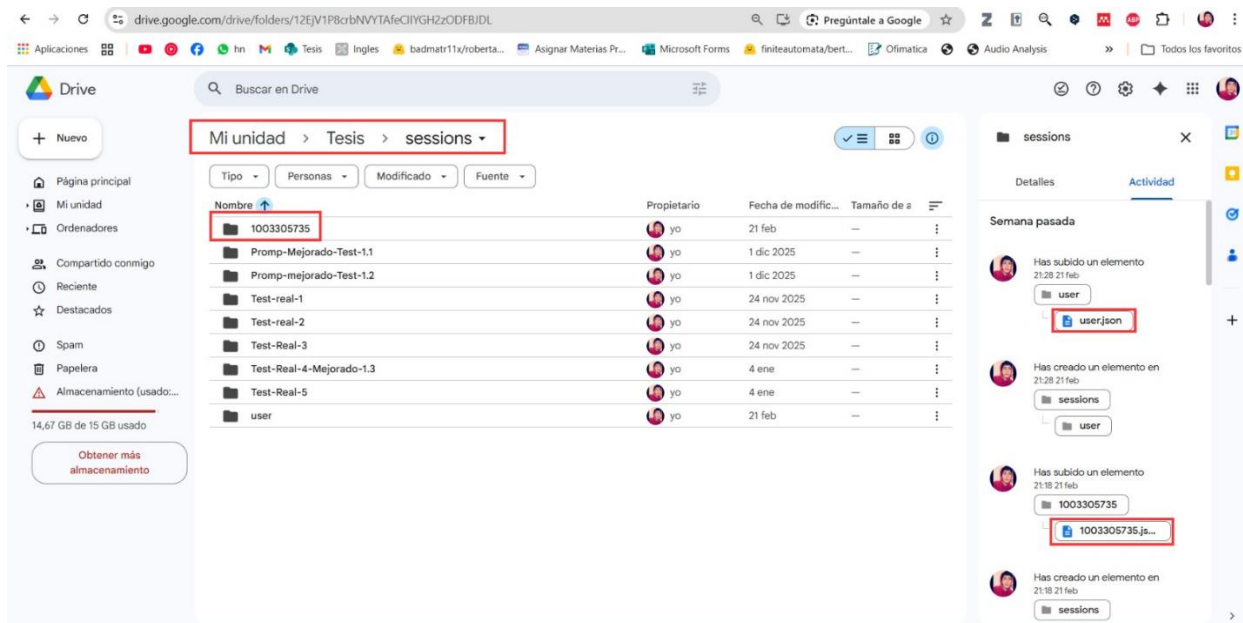
El informe generado fue convertido automáticamente a formato JSON estructurado para su almacenamiento y futura reutilización.

Figura 69*PRE-INT-01 - Exportación en formato JSON***Observaciones:**

- La estructura JSON se generó correctamente.
- No se produjeron errores de serialización.
- El archivo mantuvo consistencia en campos y etiquetas.
- Se confirmó compatibilidad para reutilización posterior.

4.1.9. Almacenamiento en Google Drive

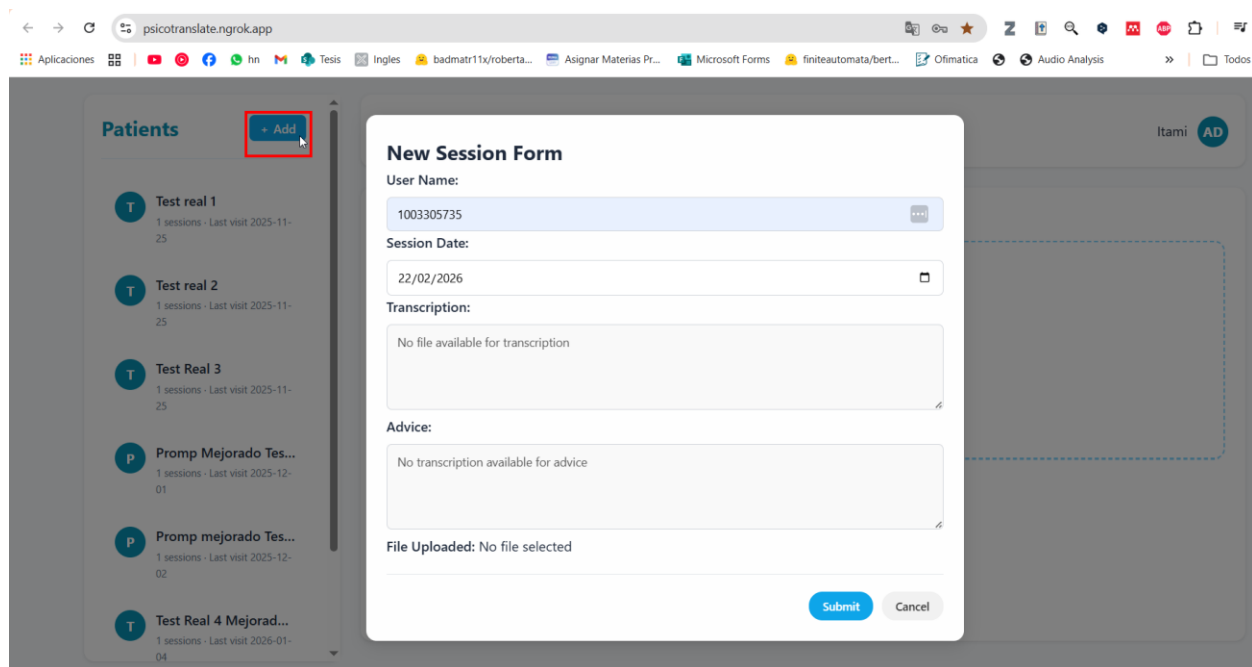
El sistema procedió a almacenar el archivo JSON y el audio correspondiente en Google Drive mediante conexión segura.

Figura 70***PRE-INT-01 - Almacenamiento en Google Drive*****Observaciones:**

- El archivo fue almacenado correctamente en el directorio correspondiente.
- No hubo pérdida de información durante la transferencia.
- Se validó persistencia tras recarga del sistema.
- Se comprobó acceso controlado y almacenamiento estructurado.

4.1.10. Vinculación Automática al Paciente en el Dashboard

Finalmente, el sistema vinculó automáticamente el registro generado al paciente correspondiente dentro del dashboard clínico.

Figura 71*PRE-INT-01 - Vinculación Automática al Paciente en el Dashboard***Observaciones:**

- El registro apareció correctamente asociado al perfil del paciente.
- Se pudo visualizar y recuperar la información sin errores.
- No hubo duplicidad de registros.
- El sistema mantuvo coherencia entre sesión, informe y almacenamiento.

4.2. Pruebas aplicadas a sesiones en entorno real

Durante la fase de pruebas del prototipo se presentaron diversos escenarios que permitieron evaluar tanto el desempeño técnico del sistema como su aplicabilidad en un entorno clínico real. En primer lugar, con el objetivo de optimizar el uso de recursos y reducir los tiempos de procesamiento se optó por grabar únicamente los segmentos más relevantes de la sesión terapéutica, principalmente aquellos en los que el paciente expone sus dificultades y la profesional emite las indicaciones centrales. Esta estrategia permitió disminuir la duración de los audios y

facilitar que la generación del informe se realice en un tiempo razonable, sin afectar la calidad del análisis clínico obtenido. De forma más representativa en la siguiente Tabla 22 se detalla aspectos importantes a consideración sobre las personas evaluadas. (Ver Anexo E)

Tabla 22

Resumen de las sesiones realizadas a las personas evaluadas

ID	Persona	Edad	Género	Duración sesión	No. sesión	Tópico principal
Test_ind1	Individuo 1	No especificada (universitario)	Femenino	50 minutos	1	Ansiedad asociada a conflicto académico, rumiación e impacto en rendimiento Sobrecarga emocional, estrés
Test_ind2	Individuo 2	Adulto joven (20–25 aprox.)	Masculino	60 minutos	1	académico-familiar e ideación pasiva de muerte
Test_ind3	Individuo 3	Adulto joven (20s)	No declarado	45 minutos	1	Ansiedad reciente con síntomas somáticos e insomnio

ID	Persona	Edad	Género	Duración sesión	No. sesión	Tópico principal
Test_ind4	Individuo 4	No declarado	No declarado	60 minutos	1	Secuelas emocionales por bullying/maltrato y afectación de autoestima Soledad crónica, culpa, baja autoestima y riesgo autolesivo
Test_ind5	Individuo 5	No declarado	No declarado	55 minutos	1	

De forma analítica a continuación se presentará un análisis individualizado de la detección diagnóstica correspondiente a cada uno de los individuos registrados en las sesiones terapéuticas. Este apartado tiene como finalidad describir de manera técnica el proceso mediante el cual el prototipo electrónico, a través de los módulos de transcripción automática y procesamiento de lenguaje natural, identifica patrones lingüísticos, indicadores emocionales y posibles factores de riesgo presentes en el discurso del paciente. El análisis se realizará de forma independiente por cada caso, permitiendo contrastar los hallazgos generados por el sistema de inteligencia artificial con los criterios clínicos esperados, evaluando así la coherencia, pertinencia y utilidad diagnóstica del modelo dentro del contexto terapéutico.

4.2.1. Prueba de sesión 1 - Test_ind1

Test_ind1 corresponde a una persona de sexo femenino, en un contexto universitario, quien manifiesta ansiedad y miedo intenso asociados a un conflicto académico que ha impactado

negativamente en su rendimiento y bienestar general. A lo largo del relato observamos la persistencia de patrones de rumiación, hipervigilancia frente a acontecimientos que nos acercan a la actividad educativa, dificultad de concentración, trastornos del sueño. Entre los factores de riesgo generales la tendencia al aislamiento, el uso de la automedicación como una estrategia inadecuada de afrontar el malestar emocional. Por otro lado, resalta un adecuado nivel de insight, reconocimiento del problema y recursos personales de afrontamiento que se configuran como un factor de protección relevante. A partir de lo expuesto, se identifican como necesidades principales la regulación de la ansiedad, la mejora de los hábitos de sueño, la reestructuración de pensamientos disfuncionales y el acompañamiento profesional continuo, proyectando un pronóstico favorable con intervención oportuna. Esto se lo puede observar detalladamente en la siguiente Figura 60.

Figura 72

Prueba de Sesión 1 del Individuo 1

The screenshot displays a clinical software interface. On the left, a 'Patients' sidebar lists several test sessions: 'Test real 1' (2025-11-25), 'Test real 2', 'Test Real 3', 'Promp Mejorado Tes...', and 'Promp mejorado Tes...'. The main area is divided into two sections. The top section, titled 'Transcription', contains a text box with a Spanish transcription of a patient's session. The bottom section, titled 'AI Clinical Insights', provides a psychological evaluation in Spanish, identifying the patient as a female university student and listing symptoms such as anxiety, fear, and concentration difficulties.

Transcription

conté todo lo que estaba pasando. Ajá. Y me dijo de que todo tiene solución y que no me siento así y que eso pasa es por algo y así, y me llevó a la iglesia y es como que él es líder de jóvenes, y entonces me llevó y hubo una charla que estaba en mi mismo, pero o sea me decía a mí, pero no, o sea no me decía el nombre por decirlo así, pero me sentí muy bien. Ya. Bueno, tengo entendido que actualmente ya pueden estar más tranquilas por el docente, ¿no? En ese sentido, ya sin el docente, ¿esos síntomas se han mantenido, han disminuido? Sí, han disminuido por decirlo así, porque, o sea sí tengo miedo todavía y es como que de hoy si ese doctor es como que, incluso por decirlo así, tengo mi cuenta de Instagram y me aparecen sugerencias de amistad, esa aparece cuando uno le revisa la cuenta, y tenía miedo de que llegue un mensaje tipo ustedes hicieron esto, se haya enterado o algo así, pero tenía, tengo miedo todavía. Todavía sientes ese miedo. Y cuando me pongo a pensar en eso es como que sí empiezo a temblar otra vez y hasta me sudan las manos por decirlo así y siento mucho frío. Ya. Bueno.

AI Clinical Insights

EVALUACIÓN PSICOLÓGICA Identificación: Paciente universitaria, género femenino, edad no especificada. Motivo de consulta: Problemas psicológicos derivados de conflicto con docente universitario, caracterizados por ansiedad, miedo intenso y afectación del rendimiento académico. Síntomas reportados: Cognitivos: Pensamientos catastróficos ("¿y si no pasa?") Ideas de abandono universitario Hipervigilancia (sensación de ser observada) Dificultades de concentración Emocionales:

4.2.2. Prueba de Sesión 1 - Test_ind2

En la Test_ind2 corresponde a una persona de sexo masculino, adulto joven, inmersa en un contexto de exigencia académica universitaria combinada con una elevada carga de responsabilidades familiares. El motivo de consulta se centra en una sensación de sobrecarga

emocional sostenida, acompañada de ideación de muerte de carácter pasivo, estrés académico y conflictos en el ámbito familiar. Durante la sesión se identifican ejes vinculados a sintomatología depresivo-ansiosa, agotamiento emocional, irritabilidad, desesperanza y aislamiento progresivo. Como factores de riesgo generales se reconoce la presencia de ideación autolesiva pasiva y un desgaste prolongado derivado de la acumulación de responsabilidades. Sin embargo, se evidencian factores protectores importantes, tales como la motivación por continuar con su formación académica y la disposición para recibir ayuda profesional. En este contexto, las necesidades de intervención se orientan al establecimiento de un plan de seguridad, la reducción y redistribución de la sobrecarga, el fortalecimiento de la red de apoyo y el seguimiento clínico sistemático. En la siguiente Figura 61 se detalla claramente esta prueba dentro del sistema en cuestión.

Figura 73

Prueba de Sesión 1 del Individuo 2

The screenshot displays a clinical software interface. On the left, a 'Patients' sidebar lists several test sessions: 'Test real 1', 'Test real 2' (highlighted with a red box), 'Test real 3', 'Promp Mejorado Tes...', and 'Promp mejorado Tes...'. The main area shows a chat window with a text message: 'sigues retorzando esas conductas, tu madre no va a ver la necesidad. Si llegas y no ha cocinado, no cocines tu. Lamentablemente. Si. Porque eso se llama sobreprotección y reforzar esas conductas. Y mientras tú sigas reforzando ese tipo de conductas, no va a haber un cambio. ¿Cuál es tu deseo de pedir? Pues bien, en fin, terminamos de conversar. Sí, pero eso no saldrá de ver, ¿no? Ya. Nadie. Ojalá haya otro momento para poder seguir conversando. Sí. Hasta luego. Hasta luego. Hasta luego. Hasta luego. Buenos días. Buenos días. Gracias.' Below the chat is an 'AI Clinical Insights' section with a 'Generate Insights' button. The insights include: 'Evaluación Psicológica Identificación: Paciente masculino, aproximadamente 20-25 años, estudiante universitario. Motivo de Consulta: Ideas de muerte/suicidio, sentimientos de sobrecarga, dificultades académicas y conflictos familiares. Análisis de la Conversación: El paciente presenta un cuadro de agotamiento emocional y sobrecarga de responsabilidades que ha desencadenado: Ideación suicida pasiva ("pensamientos de que... cojo y me voy") Síntomas depresivos (aislamiento, irritabilidad, desesperanza) Ansiedad por cumplimiento de múltiples roles Dificultades en relaciones interpersonales. Diagnóstico Provisional: Trastorno de adaptación con estado de ánimo depresivo (F43.21) Síndrome del cuidador (sobrecarga familiar) Dificultades en habilidades de comunicación y establecimiento de límites. Factores Desencadenantes: Cuidado permanente de madre con diabetes que requiere manejo de insulina Responsabilidades domésticas y financieras familiares Exigencias académicas (economía, horarios inflexibles) Aislamiento social progresivo Conflictos con profesores por situación familiar'.

4.2.3. *Prueba de Sesión 1 - Test_ind3*

En la Test_ind3 corresponde a una persona adulta joven con sexo no declarado, que presenta ansiedad de inicio relativamente reciente, caracterizada principalmente por manifestaciones somáticas y alteraciones significativas del sueño. El discurso refleja una

preocupación constante y sostenida, acompañada de síntomas físicos asociados al estrés y dificultades para conciliar y mantener el descanso nocturno. Como riesgos generales se identifica el deterioro progresivo del sueño y la somatización persistente del malestar emocional, lo cual puede afectar el funcionamiento cotidiano. No obstante, se reconocen factores protectores relevantes, como la presencia de vínculos familiares cercanos y la conciencia del propio malestar psicológico. A partir de estos elementos, se determinan como necesidades principales la instauración de rutinas de autocuidado, la aplicación de técnicas de regulación emocional y relajación, la psicoeducación sobre ansiedad y la elaboración de un plan de manejo adaptativo del estrés. En la siguiente Figura 62 se puede observar la realización de la prueba al individuo 3 de forma detallada.

Figura 74

Prueba de Sesión 1 del Individuo 3

The screenshot displays a clinical software interface. On the left, a 'Patients' sidebar lists sessions: 'Test real 1', 'Test real 2', 'Test Real 3' (highlighted with a red box), and two 'Promp Mejorado Tes...' sessions. The main area shows a text box with a question: 'como... ¿Qué hubiese pasado si no? ¿O qué hubiese pasado si no tomaba esa decisión? ¿De regresar? O de no ir. ¿Has pensado mucho en eso? Último sí. Último me estaba pensando en eso también. Listo.' Below this is the 'AI Clinical Insights' section, which contains a detailed psychological evaluation for a patient named Diego/Yair (23 years old). The evaluation includes a history of anxiety symptoms, somatic symptoms, and stressors like academic pressure and family issues. A provisional diagnosis of Generalized Anxiety Disorder (F41.1) is provided, along with additional characteristics like chronic overthinking and emotional isolation.

4.2.4. *Prueba de Sesión 1 - Test_ind4*

En la Test_ind4 corresponde a una persona cuya información sociodemográfica no fue declarada, se aborda un malestar emocional asociado a experiencias adversas prolongadas en la historia personal, incluyendo situaciones de maltrato y/o bullying, con impacto significativo en la

autoestima y la autoimagen. Durante la sesión se evidencian ejes relacionados con sintomatología de ansiedad, secuelas emocionales de tipo traumático, sentimientos de culpa, autocrítica intensa y distorsiones cognitivas vinculadas al rechazo social. Como factores de riesgo generales se mencionan antecedentes de riesgo autolesivo en la infancia y una marcada vulnerabilidad emocional actual. Sin embargo, se identifican factores protectores importantes, como la capacidad de verbalizar el malestar, el insight respecto a las experiencias vividas y la posibilidad de contar con apoyo social. En este contexto, las necesidades de intervención se orientan hacia un enfoque trauma-informado, la regulación emocional, la reconstrucción progresiva de la autoestima, el fortalecimiento de límites personales y la consolidación de una red de apoyo segura. En la siguiente Figura 63 se puede observar la realización de la prueba al individuo 4 de forma detallada

Figura 75

Prueba de Sesión 1 del Individuo 4

de alcohol, algún tipo de droga, alguna sustancia? No la consumí con nadie. Ya. ¿Cómo es tu relación actualmente hoy con tus compañeros de clases? Es buena. Hablo sí con la mayoría, pero más me llevo simplemente con una compañera y un compañero, entonces como que sí. Con ellos soy un poquito más apegada, con los demás en cambio sí, sólo hablo, si es que también se acercan a hablar, me acuerdo tenemos que ahí conversar, pero luego de ir cambiando ellos no he recibido ningún tipo de comentarios ni nada, así que sí me siento a gusto en mi clase. ¿Con tus compañeros actualmente? Ya. Listo. A ver.

AI Clinical Insights Generate Insights

Evaluación Psicológica Basada en el Contenido Textual

1. Emociones expresadas:

Dolor, tristeza profunda, humillación, soledad, abandono. Miedo al rechazo, a la intimidad y a la vulnerabilidad. Culpa (por comer, por existir, por defenderse). Rabia contenida y resentimiento hacia figuras de autoridad (padres, nutrióloga, agresores). Desesperanza pasada (ideación suicida a los 10 años). Baja autoestima y autodesprecio focalizado en el cuerpo (peso, color de piel, rostro).

2. Distorsiones cognitivas identificadas:

Sobregeneralización: "Por mi físico me alejan todos", "no merezco tener pareja". Filtro mental/Descalificación de lo positivo: A pesar de que la ropa le queda floja y médicos indican peso normal, solo se fija en el espejo donde se "ve gorda". Personalización: Asume que el maltrato fue por sus características físicas, sin cuestionar la crueldad de los otros. Pensamiento dicotómico: "Como demasiado" vs. "debo dejar de comer"; visión de sí misma como "gorda" a pesar de evidencia contraria. Razonamiento emocional: "Me siento gorda,

4.2.5. *Prueba de Sesión 1 - Test_ind5*

En la Test_ind5 corresponde a una persona cuya información sociodemográfica no fue declarada, en la cual se aborda un malestar emocional caracterizado por sentimientos persistentes de soledad, tristeza, culpa y baja autoestima, en el contexto de dinámicas familiares conflictivas y

carencia de validación emocional. Durante la sesión se evidencian ejes relacionados con sintomatología depresivo-ansiosa, desesperanza, ambivalencia afectiva hacia figuras parentales especialmente hacia el padre y una marcada necesidad de pertenencia y aceptación. De igual forma, se detectan distorsiones cognitivas tales como el pensamiento dicotómico, la sobre generalización y la personalización que ayudan a mantener la autocrítica y la visión negativa de uno mismo.

Los factores de riesgo generales son, por un lado, los indicadores de vulnerabilidad emocional estable, la presencia de ideación pasiva asociada al sentimiento de vacío y la autoimagen deteriorada como factores que podrían incrementar la probabilidad de la conducta autolesiva de no ser interviniendo adecuadamente. También se detectan factores protectores relevantes como la capacidad de expresar el malestar de forma open, el reconocimiento parcial de los patrones emocionales que les afectan y la disposición para reflexionar sobre su experiencia personal. En este sentido, nuestras necesidades de intervención van destinadas al trabajo del desarrollo de la regulación emocional, el tratamiento de reorganización de los esquemas de malestar (esquemas disfuncionales, el trabajo en autoestima, la autovaloración y el tratamiento de las dinámicas relacionales familiares, la mejora de recursos internos y una red de apoyo estable que favorezca el bienestar psicológico.

Figura 76

Prueba de Sesión 1 del Individuo 5

Transcription

Yo sé que muchas cosas de las que pienso o hago están mal, pero no puedo cambiarlas. ¿Muchas cosas de las que piensas y haces? ¿Como cuáles las que haces? Cuando yo me sentí así, por ejemplo hace un año, yo nunca he pensado en hacer cosas malas. Pero me sentí muy mal. Y sí, hay amigos que son muy mal influyentes, ¿no? Y que en ese momento yo siempre me dejaba llevar. Y en ese tiempo mi papá ya me dejó ir sola, yo regresaba de casa sola. Y esas amistades siempre me llevaban, me decían, vamos. Y me llevaban a tomar alcohol. Y yo tomaba porque me sentía muy mal. Y luego ya nadie quiso hacer eso, pero recientemente cada que me sentía mal yo necesitaba y yo quería hacer eso. ¿Te ayudaba a sentirte mejor el alcohol? Sí, por un momento sí. ¿Tomabas mucho? No tanto, porque yo siempre era la que le cuidaba a todos. Pero en ese momento era algo como sentirme mejor, como... no sé. Pero por un momento yo me sentía bien, pero luego decían, no, estoy haciendo mal. ¿Qué te hace sentir bien? ¿El grupo de amigos o sí el tomar? Las dos cosas, estar con ellos y también... ¿Qué tomaban? Todo tipo de... ¿Te explico? Sí. Ya. ¿Y esto lo hiciste un buen tiempo? ¿No se daban cuenta tus papis? No, porque llegaban tarde. Sí, llegaban. Ya. ¿Has probado algún tipo de drogas? ¿Te han

AI Clinical Insights Generate Insights

Evaluación Psicológica Basada en el Contenido Textual

1. Evaluación Psicológica Emociones expresadas: Soledad crónica, tristeza, culpa, arrepentimiento, ansiedad, frustración, baja autoestima, desesperanza, ambivalencia afectiva hacia los padres (especialmente hacia el padre), necesidad de pertenencia y aceptación. Distorsiones cognitivas:

Pensamiento dicotómico: "No puedo ser como mis hermanas", "soy débil hasta para eso [suicidio]". Sobregeneralización: "Siempre me he sentido sola", "nunca dicen...". Personalización: Culpa por las acciones de otros (ej. novio que la acompañó). Descalificación de lo

4.2.6. Discusión sobre las pruebas efectuadas

Una vez concluida la intervención del paciente y la retroalimentación de la profesional a cargo, el sistema inicia el proceso de carga y almacenamiento de la información. En promedio, este proceso toma entre tres y cuatro minutos, dependiendo del ancho de banda disponible, ya que los archivos deben ser subidos y persistidos en Google Drive. Posteriormente, se habilita la etapa de transcripción automática del audio a texto, cuyo tiempo de ejecución también oscila entre tres y cuatro minutos. Estos tiempos evidencian una dependencia directa de la conectividad de red y del tamaño del archivo de audio procesado. Durante este flujo se pudo observar en el entorno de ejecución de Google Colab que, cuando todas las etapas se completan de manera satisfactoria, el sistema habilita automáticamente el módulo de generación del informe prediagnóstico. Esta fase adicional presenta un tiempo de respuesta aproximado de dos a tres minutos, correspondiente al procesamiento del texto mediante el modelo de lenguaje. En términos generales, el tiempo total

desde la finalización de la sesión hasta la obtención del informe completo se mantiene dentro de un rango aceptable para un entorno clínico, considerando que el proceso se realiza de forma automatizada.

En cuanto al análisis del contenido generado, en la primera aplicación del sistema sobre un caso real se evidenció una alta coherencia entre los aspectos clínicos identificados por la profesional y los resultados presentados en el informe generado por la herramienta. Uno de los elementos que más interesó a la profesional fue la presencia de un espacio donde se exponen posibles intervenciones terapéuticas, estando organizadas y segmentadas por sesiones, ya que añadió un plus al informe, facilitando así la planificación del tratamiento.

Sin embargo, en el transcurso de las pruebas también se reconocieron diversos límites. Uno de ellos es que actualmente la transcripción automática se configuraba como un bloque único de texto, en el que no existe una segmentación explícita entre las intervenciones del psicólogo y las del paciente, distinción bien inferida por el modelo generativo durante la generación del informe, pero la falta de esa segmentación, añadían enunciados de la segmentación, afectaba a la claridad del texto que se transcribía. De igual forma, al estar grabándose sesiones, de larga duración, el tamaño del archivo entrenaba grandes incrementos en tiempos de carga, de la transcripción y de la generación del informe; se detectaron en algunas ocasiones retrasos de tiempo en los controles de grabación, que podían llevar a la pérdida de la sesión si la página era actualizada. Como medida de mitigación, se implementó una segunda modalidad de carga de audios, permitiendo al profesional subir archivos de forma manual desde el dispositivo, los cuales permanecen almacenados localmente en la Raspberry Pi.

Otra limitación identificada corresponde a la gestión de sesiones, ya que actualmente el profesional no puede eliminar registros almacenados, siendo esta una acción reservada únicamente

al administrador del sistema. Esta restricción se plantea como un aspecto a mejorar en versiones futuras, con el fin de otorgar mayor control al usuario final sin comprometer la seguridad de la información.

Desde el punto de vista clínico, se reconoce que el análisis basado exclusivamente en texto puede generar interpretaciones parciales del contexto, dado que en la práctica terapéutica influyen otros factores relevantes como el tono de voz, la postura corporal y las expresiones no verbales. En este sentido, se plantea como trabajo futuro la incorporación de una cámara u otros sensores que permitan analizar estas variables y aproximar aún más el sistema a la realidad clínica.

Asimismo, al tratarse de informes generados por un único modelo de inteligencia artificial, no se garantiza que todos los resultados sean completamente precisos o adecuados para todos los tipos de casos. Por ello, se propone como línea futura permitir la selección entre distintos modelos de lenguaje, de modo que el profesional pueda elegir aquel que considere más apropiado según el contexto clínico, reconociendo que ciertos modelos pueden ofrecer mejores resultados en problemáticas específicas.

En cuanto a la seguridad física del dispositivo, se estableció que este debe ubicarse cerca del área de trabajo del profesional para garantizar una correcta captura de audio. No obstante, se plantea como mejora futura la posibilidad de camuflar el dispositivo en objetos cotidianos, como un libro o un elemento decorativo, con el fin de que pase desapercibido y no genere incomodidad en el paciente.

Cabe destacar que debido al carácter confidencial de las sesiones terapéuticas se informó de manera previa a los pacientes sobre el uso de grabación y la aplicación de inteligencia artificial en las sesiones. Aunque no se utilizó el término “grabación” en todos los casos tratados, la profesional que estaba a cargo se encargó de explicar el procedimiento y recabar el consentimiento

informado correspondiente, el cual fue aceptado y firmado por los pacientes. Si bien la información gestionada tiene un alto grado de sensibilidad, la utilidad del sistema para los profesionales resulta ser alta, en la medida en que permite disminuir la carga cognitiva que supone la toma manual de notas, así como evitar el riesgo de perder información, o de mezclar información de diferentes casos. Contar con un registro exhaustivo y fiel de la sesión supone que el profesional pueda revisar, complementarlo y mejorar su análisis clínico, es decir mejorar los procesos de la intervención y su calidad asistencial. Adicionalmente, la digitalización de la información clínica permite incluir elementos de seguridad más complejos que los que históricamente nos proporciona el papel (la posibilidad de encriptar los datos almacenados, por ejemplo), lo cual sería una mejora para incluir en un futuro.

Entre las principales ventajas del sistema destaca su bajo consumo energético, el uso de la inteligencia artificial no como reemplazo de los profesionales, sino como una ayuda, así como su potencial para mejorar la interacción terapéutica, haciendo que el psicólogo pueda centrar su atención en el paciente y no en la toma de notas. Finalmente, los resultados del cuestionario constatan que, si bien algunos profesionales todavía sienten rechazo a la incorporación de la inteligencia artificial en el ámbito de la psicología, hay otros que la entienden como una posibilidad de mejora importante cuando se hace un uso bien delimitado y controlado. Desde esta perspectiva, el desarrollo del prototipo es más que digno de mención, dado que se trata de una problemática que tiene respuesta: cómo registrar completamente las sesiones de terapia y no interferir con la intervención. El sistema propuesto ofrece un respaldo fiel del contenido tratado, favoreciendo un mayor acercamiento entre profesional y paciente y reduciendo posibles bloqueos o omisiones de información durante la sesión

4.3. Análisis de la encuesta realizada

De un total de 23 encuestados comprendidos entre profesionales del área de Psicología, docentes del área de Psicología y estudiantes de Psicología se obtuvieron los siguientes resultados (Anexo C). La encuesta fue diseñada con el objetivo de evaluar la calidad, pertinencia clínica y utilidad profesional de los informes generados a partir de una misma sesión terapéutica, empleando distintos modelos de inteligencia artificial. Para ello, se proporcionó a los profesionales participantes el material necesario para comprender el contexto clínico y emitir una valoración informada. En primer lugar, se entregó a los participantes el registro textual completo de una sesión terapéutica real, obtenida durante pruebas del sistema. Las primeras secciones del documento corresponden a la transcripción íntegra de la conversación sostenida entre el paciente y el psicólogo tratante, sin modificaciones en su contenido. Este material permitió que los profesionales comprendieran el desarrollo de la sesión, el discurso del paciente y los elementos clínicos relevantes expresados durante la intervención, sirviendo como base común para la evaluación posterior.

Posteriormente, se presentó el Informe 1, el cual corresponde al resultado generado directamente por la herramienta desarrollada en el marco de esta tesis. Este informe representa la salida principal del sistema propuesto y constituye el producto que se entrega como apoyo al profesional de la psicología. Para su generación se empleó el modelo de lenguaje DeepSeek, seleccionado debido a que, al momento de iniciar el desarrollo del proyecto, este modelo ofrecía un desempeño competitivo en términos de coherencia y reducción de errores. Es importante señalar que el desarrollo del prototipo se realizó aproximadamente dos años atrás, en un contexto donde los modelos de lenguaje aún se encontraban en una etapa inicial de madurez y disponibilidad

pública, y en el cual DeepSeek presentaba un porcentaje de acierto aproximado del 63 %, siendo una de las opciones más viables en ese momento.

Para una comparación más objetiva del mismo texto transcriptivo de la sesión de terapia, este fue procesado a través del mismo prompt clínico que se utilizó en la herramienta creada, pero aplicado a otros modelos de inteligencia artificial. De esta manera, el Informe 2 fue generado con el modelo Grok, permitiendo ver la manera en la que este sistema procesa y estructura la información clínica bajo las mismas instrucciones. Y, de la misma forma, el Informe 3 fue el resultado del modelo Gemini utilizando las mismas condiciones de entrada y las mismas instrucciones de generación, lo cual permite una comparación directa de los resultados obtenidos por diferentes arquitecturas de modelos de lenguaje.

Finalmente, el Informe 4 fue el resultado del modelo ChatGPT utilizando el mismo prompt y un espacio natural que busque coherencia y sentido clínico en la salida de ese sistema. Con este informe también se puede contrastar el funcionamiento del sistema que hemos propuesto con uno de los modelos de lenguaje mayormente extendidos y avanzados en el presente. A la postre, esta investigación ha permitido a los profesionales de la salud hacer una comparación de los informes elaborados a partir de diferentes modelos de inteligencia artificial a partir de una misma base clínica real y de un mismo tipo de condiciones, realzando de este modo la validez del análisis y de las conclusiones que se extrajeron de la encuesta.

Bloque A

Los resultados reflejaron una alta aceptación en cuanto a la estructura, claridad y coherencia interna del informe clínico generado por el sistema. En las preguntas A1, A2 y A3 se observa que la mayoría de los participantes se concentran en las categorías “De acuerdo” y “Totalmente de acuerdo”, lo cual indica que el informe presenta una organización lógica, un

lenguaje clínico comprensible y una adecuada correspondencia entre síntomas, análisis y conclusiones.

Mientras que la presencia de respuestas neutrales en la A4 sugiere oportunidades de mejora en la reducción de redundancias o ajustes de redacción, sin que esto represente una debilidad estructural significativa del sistema.

Bloque B

En las preguntas B1 al B4 se evidencia una valoración mayoritariamente positiva respecto a la pertinencia del diagnóstico presuntivo, la correcta aplicación de los criterios CIE-11, la fundamentación del diagnóstico diferencial y la congruencia de la gravedad asignada. Estos resultados indican que el sistema logra alinear el análisis automatizado con marcos diagnósticos reconocidos, fortaleciendo su utilidad como herramienta de apoyo clínico. Las respuestas neutrales presentes en B3 y B4 pueden interpretarse como una manifestación del criterio profesional, dado que la gravedad y el diagnóstico diferencial suelen requerir contextualización adicional que va más allá del texto transcrito.

Bloque C

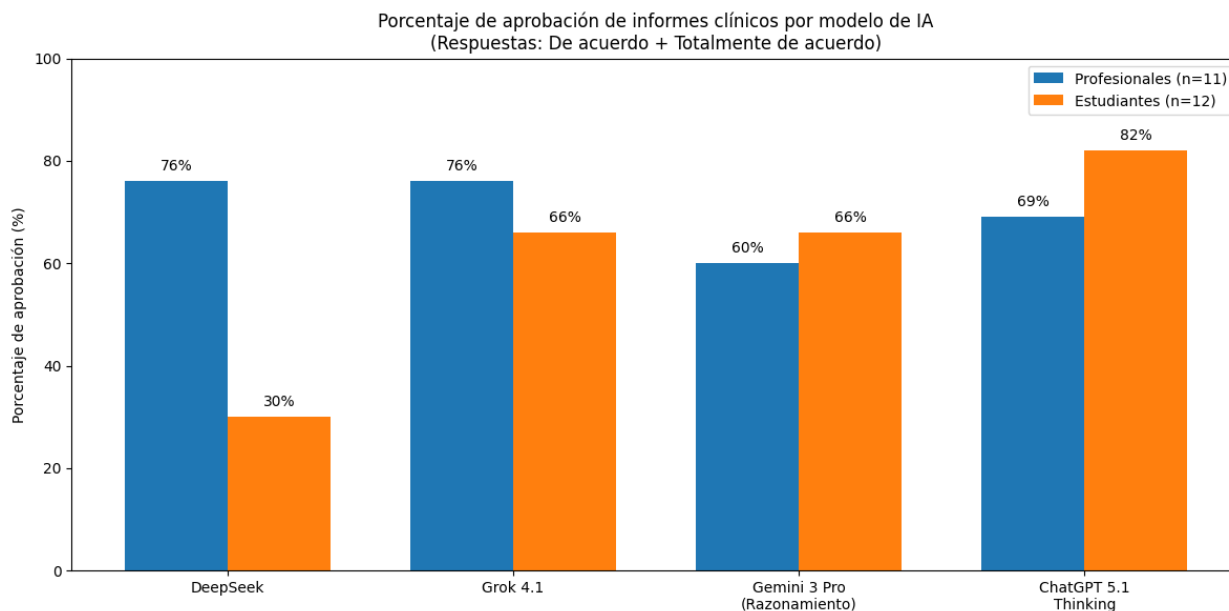
Las preguntas C1 y C2 muestran que los participantes perciben el informe como realista y comparable a casos observados en la práctica clínica, lo cual valida la efectividad del modelo de lenguaje para capturar patrones clínicos relevantes a partir del discurso del paciente.

Asimismo, los resultados de C3 y C4 evidencian un alto nivel de aceptación en cuanto al respeto de principios éticos y la adecuada descripción del impacto funcional, aspecto fundamental al tratarse de información clínica sensible. Esto refuerza que el sistema no presenta sesgos evidentes ni vulnera principios básicos de la práctica psicológica.

Bloque E

Figura 77

Interpretación de aceptación de los informes generados



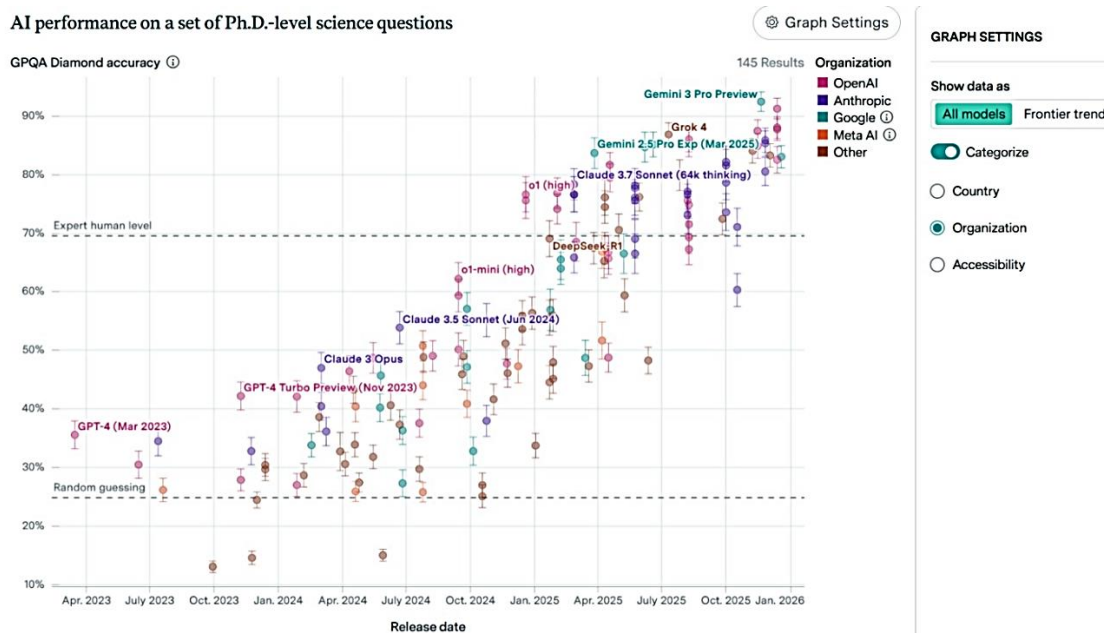
Como se pudo observar en la Figura 65 presenta el porcentaje de aprobación de los informes clínicos generados por distintos modelos de inteligencia artificial, diferenciando la valoración realizada por profesionales y estudiantes. En el grupo de profesionales, los informes generados mediante DeepSeek y Grok 4.1 alcanzan niveles de aprobación equivalentes (76%), lo que sugiere una adecuada correspondencia entre la estructura, el razonamiento clínico y los criterios diagnósticos esperados desde la práctica profesional. En contraste, el informe generado con Gemini 3 Pro (razonamiento) obtiene un porcentaje inferior (60%), mientras que ChatGPT 5.1 Thinking presenta un nivel intermedio de aprobación (69%).

Desde la perspectiva del grupo de estudiantes se observa una preferencia significativamente mayor hacia ChatGPT 5.1 Thinking (82%), seguido por Gemini 3 Pro y Grok 4.1 (66%), mientras que DeepSeek registra el menor nivel de aprobación (30%). Esta diferencia puede atribuirse, en parte, a la familiaridad previa de los estudiantes con determinados estilos de

redacción y organización de la información, así como a una menor exigencia en la evaluación de la coherencia clínica profunda, la pertinencia diagnóstica y la precisión conceptual.

Figura 78

Gráfica de desempeño de los modelos IA



De acuerdo con la gráfica de desempeño se observa que Gemini 3 Pro Preview alcanza el mayor porcentaje de exactitud entre todos los modelos evaluados, superando tanto el umbral de expert human level como a las versiones más recientes de otros competidores, esto debido a que presenta un rendimiento cercano al 90% de precisión, posicionándose por encima de modelos como GPT-4, Claude 3.7 Sonnet, Grok y DeepSeek. Esto evidencia no solo una mejora progresiva en el tiempo, sino también una ventaja comparativa en términos de capacidad de

razonamiento científico avanzado. Desde una perspectiva técnica, este resultado sugiere que Gemini demuestra mayor robustez en:

- Resolución de problemas complejos de nivel doctoral.
- Inferencia científica estructurada.
- Manejo de preguntas multidisciplinarias de alta dificultad.
- Consistencia en el rendimiento frente a benchmarks exigentes.

En base a los resultados de la encuesta se evidencian una aceptación positiva y consistente del prototipo tanto en términos técnicos como clínicos y éticos. La mayoría de los participantes reconoce el valor del sistema como herramienta de apoyo al diagnóstico presuntivo, validando su estructura, coherencia, realismo y utilidad práctica. Pero a pesar de ello los estudiantes prefieren el uso de ChatGPT por su mayor confiabilidad y tiempo de uso. Es importante destacar que, aunque la percepción estudiantil aporta una visión relevante desde el ámbito formativo, la valoración realizada por los profesionales posee un mayor peso interpretativo, debido a que se fundamenta en una experiencia clínica consolidada, en el manejo directo de casos reales y en el uso sistemático de criterios diagnósticos estandarizados. En este sentido, las diferencias observadas entre ambos grupos refuerzan la idea de que la alineación clínica de los modelos de inteligencia artificial no depende únicamente de la claridad expositiva o la estructura del informe, sino principalmente de su capacidad para reflejar procesos de razonamiento clínico acordes con la práctica profesional.

4.4. Cumplimiento del estándar ISO/IEC 25010

Con el propósito de identificar las necesidades reales del entorno clínico y fundamentar técnicamente el diseño del prototipo electrónico se aplicó una encuesta estructurada a la profesional responsable del departamento de psicología del área de bienestar estudiantil, mediante

el uso de respuestas en la escala de Likert. Con el fin de permitir recopilar información directa sobre los procesos actuales de toma de notas, almacenamiento de información, percepción sobre la interferencia durante la sesión terapéutica y nivel de aceptación frente al uso de herramientas basadas en inteligencia artificial. Asimismo, la encuesta permitió detectar requerimientos funcionales específicos para el sistema, evidenciándose, la necesidad de incorporar una sección que permita seleccionar el modelo de inteligencia artificial a utilizar para la generación del diagnóstico clínico, con el fin de otorgar mayor flexibilidad, control profesional y adaptabilidad tecnológica al prototipo propuesto.

De acuerdo con los resultados obtenidos en la encuesta aplicada a la profesional responsable del área de psicología, evidenciados en el Anexo F, se observa que la mayoría de las dimensiones evaluadas bajo el marco de la norma ISO/IEC 25010 fueron valoradas en el nivel de “Totalmente de acuerdo”, específicamente en adecuación funcional, eficiencia de desempeño, capacidad de interacción, fiabilidad, flexibilidad, mantenibilidad, protección y calidad global del producto. La dirección que toma esta tendencia se puede considerar como una gran validación del sistema que se proponía, por el hecho de que se tomara como cierto que se responde satisfactoriamente a las exigencias reales del entorno clínico y de que supera las limitaciones inherentes al registro manual de notas. Sin embargo, el hecho de que se considere una valoración neutral en la dimensión de compatibilidad y "Algo de acuerdo" en el apartado correspondiente a la seguridad produce una lectura más matizada y más fiel a la realidad del prototipo, propia de un proceso de evaluación técnica objetiva.

En lo que respecta a la adecuación funcional, la valoración total indica que el sistema da cabida a todos los requerimientos establecidos en el proyecto en el sentido de que se permite la captura de audio, la transcripción automática, el análisis clínico mediante modelos de lenguaje, la

generación de informes presuntivos y el almacenamiento estructurado por paciente y sesión. Asimismo, la coherencia entre la información de entrada y los resultados generados pone de manifiesto la corrección funcional, paralela a la constatación de la consistencia entre los síntomas detectados, el análisis llevado a cabo y las conclusiones estructuradas. Dicha solidez se avala mediante un resultado de pruebas comparativas, así como consultando los informes generados.

En lo concerniente a la eficiencia productiva, hemos delegado el procesamiento intensivo (reconocimiento automático de habla y análisis mediante IA) a servicios en la nube desde allí la Raspberry Pi opera como nodo de captura e interfaz, mientras que en lo que respecta a la eficiencia operativa se evidencia una arquitectura optimizada que hace un balance en rendimiento y consumo de recursos. Por otro lado, si la capacidad de interacción se considera pues se evidencia además una interfaz web muy intuitiva que dote al profesional de la carga de audios, la revisión de transcripciones y posterior informe de forma sencilla, sumando así en la usabilidad entre otras cosas, en la sesión terapéutica sin menoscabar la dinámica clínica. En lo que respecta a la fiabilidad, el sistema constituye un prototipo operativo, con elementos para almacenamiento persistente y procesos para permitir la recuperación de información después de reinicios o apagones parciales que permiten alcanzar la continuidad operativa. En cuanto a la dimensión de la seguridad, la opinión de "Un poco de acuerdo" indica que existen mecanismos para la autenticación, almacenaje en entornos controlados, y restricciones de acceso, pero la misma nota que es posible el mejorar aspectos como métodos criptográficos, el endurecimiento de las políticas de respaldo o la implementación de controles adicionales a los que no fuesen implementados previamente, considerando niveles de seguridad sanitaria más exigentes. En cuanto a la compatibilidad, la opinión neutral induce a pensar que el sistema, si bien logra interoperar mediante APIs REST con los servicios externos que son Google Colab, Google Drive y los

modelos de lenguaje, una futura integración con sistemas clínicos institucionales o con plataformas hospitalarias también podría mejorarse para lograr una interoperabilidad superior. Respecto a la mantenibilidad y flexibilidad, la arquitectura basada en la división en capas (presentación, aplicación y datos) permite cambiar componentes concretos como el modelo de inteligencia artificial o los endpoints de comunicación sin cambiar la estructura general, siendo esta una clave para la escalabilidad y la evolución tecnológica. La protección en cuanto a la dimensión de la protección indica que el sistema adopta un enfoque preventivo al especificar explícitamente que los informes generados son presuntivos y de apoyo, y no verdaderos informes diagnósticos, promoviendo así un uso responsable de la inteligencia artificial y limitando los riesgos asociados a las interpretaciones automatizadas. En conclusión, los resultados de la evaluación continúan mostrando que el prototipo se comporta favorablemente para los criterios de calidad definidos por ISO/IEC 25010, áreas muy claramente fortalecidas y localizadas para su mejora técnica. Esta evaluación integral confirma que el sistema es tecnológicamente viable, funcionalmente robusto y alineado con principios éticos y operativos del entorno clínico, consolidándose como una herramienta de apoyo eficaz para la automatización y análisis inteligente de notas clínicas en sesiones terapéuticas.

Figura 79

Resultados de la encuesta de validación ISO/IEC 25010 realizada a la profesional encargada del departamento

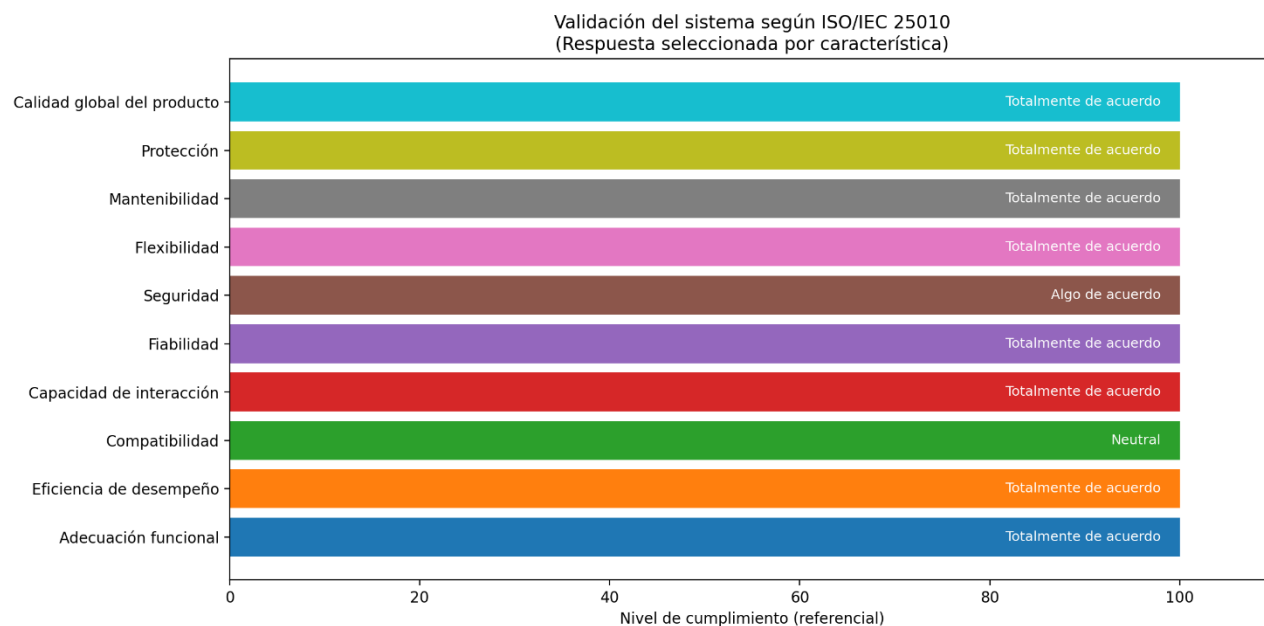


Figura 80

Pregunta abierta realizada a la profesional encargada del departamento

13. Desde su experiencia tras utilizar el sistema, ¿qué funcionalidades considera más valiosas y cuáles incorporaría en su práctica profesional? Justifique su respuesta.

1 Respuestas

ID ↑	Nombre	Respuestas
1	anonymous	Se propone incorporar un selector de modelos de inteligencia artificial que permita comparar distintos análisis presuntivos. Esto facilitaría identificar cuál se ajusta mejor al criterio clínico del profesional, reforzando el carácter asistido del sistema.

Con respecto a la pregunta que hace en la respuesta a la pregunta 13, en la que esta profesional cita textual que le gustaría contar con una sección de selección del modelo de IA que utilizará para generar el diagnóstico (ChatGPT, Grok, Gemini y otros) propio de los requerimientos funcionales de alto valor estratégico que se ha viajado en el diseño del sistema, esta respuesta no

es sólo una expresión de apertura al uso de IA, sino que revela una necesidad de control profesional y flexibilidad de la tecnología informando de manera fundamental entornos de clínicos donde la toma de decisiones no puede depender de un único algoritmo de fuente. Y desde la ingeniería del software esta demanda hace que la completitud funcional se vea reforzada ya que se amplía la delimitación del sistema en un sentido también modular, integrando múltiples APIs o modelos de lenguaje. Del mismo modo, mejora la adaptabilidad y escalabilidad haciendo que el sistema pueda ser capaz de ser evolucionado conforme se vayan implementando modelos más avanzados sin tener que rediseñar la plataforma. Desde el punto de vista técnico implica diseñar una capa de abstracción que permite desacoplar el motor de análisis del modelo específico permitiendo la interoperabilidad y actualización de manera dinámica..

En términos clínicos, esta funcionalidad incrementa la confianza del usuario ya que el profesional mantiene el criterio decisional al poder contrastar resultados entre modelos o elegir aquel que considere más adecuado según el contexto del caso. Esto refuerza el principio de que la IA actúa como herramienta de apoyo y no como sustituto del juicio clínico.

Conclusiones

- El desarrollo del prototipo electrónico para la automatización y análisis inteligente de notas clínicas en sesiones terapéuticas demostró ser técnicamente viable y clínicamente pertinente, cumpliendo con los requerimientos funcionales, no funcionales, de hardware y software definidos bajo las normas ISO/IEC/IEEE 29148:2018 e ISO/IEC 25010.
- La combinación de los avances en la tecnología del reconocimiento automático de voz (ASR), procesamiento del lenguaje natural (NLP) y modelos de lenguaje generativo fueron factores clave para poder convertir de forma efectiva el contenido de audio de las sesiones de terapia en texto estructurado, que junto a la estructura con la que se podrían generar los informes clínico presuntivos, llevaron a conseguir coherencia, claridad y adecuación a los marcos diagnóstico aceptados por la misma comunidad científica, como el CIE-11 y DSM-5.
- El nivel de resultado que obtuvimos a partir de las pruebas de funcionamiento de nuestro sistema y de la encuesta aplicada a los profesionales, docentes y estudiantes de psicología también mostró la alta aceptación del propio sistema como herramienta de apoyo clínico que puede servir para hacer disminuir la carga administrativa, sistematizar el contenido que se genera en la terapia y, maximizar el tiempo de los profesionales clínicos durante la propia sesión de terapia. • El análisis comparativo entre los diferentes modelos de inteligencia artificial (DeepSeek, Grok, Gemini y ChatGPT) evidenció que, aunque todos presentan capacidades funcionales, hay diferencias significativas en cuanto a su coherencia clínica, a la estructura narrativa, así como a la aceptación por parte del personal; lo que demuestra la necesidad de considerar la IA como un soporte contextual y no como un recurso único sobre la decisión.
- Desde la perspectiva clínica, el dispositivo consigue correlaciones significativas entre los hallazgos producidos por la inteligencia artificial y el criterio del personal, aseverando tal cosa

especialmente en la identificación de síntomas, patrones de pensamiento o propuestas de planes terapéuticos por fases, mostrando así su valor como apoyo documental y analítico.

- Se concluye que el prototipo no sustituye el juicio clínico del profesional, sino que apoya la práctica terapéutica, puesto que logra promover una mayor concentración en la interacción con el paciente, reduce la inhibición producida por la realización de notas de forma manual, y evita la pérdida de información relevante durante la sesión.
- Por último, la baja demanda de energía del dispositivo, su portabilidad, así como la posibilidad de procesamiento local o asistido por la nube sitúan al sistema como una solución replicable y escalable, e interesantemente aplicable en contextos clínicos reales y en contextos académicos.

Recomendaciones

- Se sugiere que el uso del sistema se ejecute siempre bajo consentimiento informado explícito, para asegurar que el paciente conozca el alcance, limitaciones, y finalidad del uso de inteligencia artificial durante la sesión terapéutica.
- Se sugiere reforzar los mecanismos de seguridad y la confidencialidad de la información, incluyendo cifrado de extremo a extremo para audios, transcripciones e informes clínicos, especialmente si el sistema se implementa en entornos productivos.
- Se recomienda entrenar a los profesionales de la salud mental para un uso adecuado del sistema, recalcando que los informes generados por IA deben considerarse sólo orientativos y no determinantes, manteniendo, en todo caso, el criterio clínico humano como eje central de la toma de decisiones.
- Para aumentar la claridad de los registros, se aconseja implementar mecanismos explícitos de segmentación del texto transcrito, para diferenciar las intervenciones del paciente y las del profesional; esto facilitaría la lectura, posterior análisis y revisión de las sesiones.
- Se recomiendan optimizar los tiempos de procesamiento mediante estrategias de reducción de la duración de audios, elección de fragmentos clínicamente relevantes, y mejora de la conectividad de red, sobre todo en contextos con limitaciones de ancho de banda.
- Se recomienda ampliar las funcionalidades de gestión de sesiones permitiendo que el profesional autorizado pueda eliminar o gestionar los registros, manteniendo controles de auditoría que garanticen la trazabilidad y seguridad de la información.

Trabajos a Futuro

- Como línea futura de investigación, se plantea la integración de análisis multimodal, incorporando cámaras u otros sensores que permitan evaluar variables no verbales como tono de voz, pausas, postura corporal y expresiones faciales, enriqueciendo el análisis clínico más allá del texto.
- Se propone implementar un sistema de selección dinámica de modelos de inteligencia artificial, permitiendo al profesional elegir entre distintos modelos según el tipo de caso clínico, reconociendo que cada modelo puede presentar fortalezas específicas para determinadas problemáticas psicológicas.
- Se sugiere la implementación de un módulo de evaluación longitudinal para poder registrar la evolución del paciente en múltiples sesiones, a partir del cual sería posible identificar cambios en patrones lingüísticos, emocionales y comportamentales a lo largo del tiempo.
- Se sugiere continuar la investigación con otras muestras clínicas diferentes y casos reales, lo que ayudaría a evaluar estadísticamente la efectividad del sistema, a mejorar las indicaciones clínicas y a mejorar la precisión de los informes generados por la inteligencia artificial.

Bibliografía

- Aguilar, L. (2023). INTELIGENCIA ARTIFICIAL Y CIENCIA DE DATOS. *Ciencias.Org.Ar*.
https://ciencias.org.ar/user/CETI/Compilado%20CETI%20last_2.pdf#page=13
- Anchino, L. A., Torti, A., Miretti, M., Bernardi, E., Peretti, G., & Podadera, R. (2019). Minería de opiniones: análisis de sentimientos en una red social. *XXI Workshop de Investigadores En Ciencias de La Computación*, 971–974. <http://sedici.unlp.edu.ar/handle/10915/77379>
- Andreu, J., & Martínez, L. (2021). Avances científico-tecnológicos en inteligencia artificial responsable. *Torrossa.Com*.
<https://www.torrossa.com/gs/resourceProxy?an=5085274&publisher=FZ1825>
- ASAMBLEA NACIONAL. (2015). *LEY ORGANICA DE SALUD*. REGISTRO OFICIAL.
www.lexis.com.ec
- Asamblea Nacional. (2021). *LEY ORGÁNICA DE PROTECCIÓN DE DATOS PERSONALES*.
 Registro Oficial .
- Asamblea nacional. (2021). Ley Orgánica de Protección de Datos Personales . *Registro Oficial* .
- Bezanilla, A., García, L., & Garófalo, A. (2018). Propuesta de controles de seguridad para nubes privadas y centros de datos virtualizados. *Telemática*, 17(1), 56–72.
<https://revistatelematica.cujae.edu.cu/index.php/tele/article/view/290>
- Concha, F. (2024). Entrenamiento en Formulación de Casos en Psicoterapia a través de Inteligencia Artificial Generativa. *Revista de Psicoterapia*, 35(127).
<https://doi.org/10.5944/rdp.v35i127.37798>
- Digión, L., & Digión, M. (2020). Marco conceptual de calidad de interfaz Para educación virtual . <https://Recursos.Educoas.Org/Sites/Default/Files/1865.Pdf>.

- Drigas, A., Dede, D. E., & Dedes, S. (2020). Mobile and other applications for mental imagery to improve learning disabilities and mental health. *International Journal of Computer Science Issues*, 17(4).
- Garrett, P. M. (2023). Bowlby, Attachment and the Potency of a 'Received Idea.' *British Journal of Social Work*, 53(1). <https://doi.org/10.1093/bjsw/bcac091>
- González, C., Ruiz, M., Cordero, P., Umaran, O., Hernández, A., Muñoz, R., & Cano, A. (2018). Terapia cognitivo-conductual transdiagnóstica en atención primaria: un contexto ideal. *Roderic.Uv.Es*, 29, 37–52. <https://roderic.uv.es/handle/10550/70569>
- Granda, K., & Martínez, F. (2023). Aplicaciones de la inteligencia artificial en cirugía. *Saludconciencia.Com.Ar*. <http://saludconciencia.com.ar/index.php/scc/article/view/31>
- Flemotomos, N., Martinez, V., Chen, Z., Singla, K., Ardulov, V., Peri, R., . . . Narayanan, S. (2022). Automated evaluation of psychotherapy skills using speech and language technologies. *Behavior Research Methods*, 54(2), 690-711. <https://doi.org/https://doi.org/10.3758/s13428-021-01623-4>
- Leoncio, J. C. (2021). Epigenética. *Genética Na Escola*, 16(2). <https://doi.org/10.55838/1980-3540.ge.2021.372>
- Li, R., Niu, Y., Scott, S. R., Zhou, C., Lan, L., Liang, Z., & Li, J. (2021). Using electronic medical record data for research in a healthcare information and management systems society (HIMSS) analytics electronic medical record adoption model (EMRAM) Stage 7 Hospital in Beijing: Cross-sectional study. *JMIR Medical Informatics*, 9(8). <https://doi.org/10.2196/24405>

- Lloret, E. ; Suárez, A. ; Ferrández, A. ; Navarro, B. ; Martín, T. J. ; Vicente, E., Marta, ;, Miró, M. ;, Moreda, P. ;, Romá, T. ;, & Palomar, M. (2021). *Generación del Lenguaje Natural: retos y desafíos científicos*. <http://rua.ua.es/dspace/handle/10045/117460>
- Marcelle, E. T., Nolting, L., Hinshaw, S. P., & Aguilera, A. (2019). Effectiveness of a multimodal digital psychotherapy platform for adult depression: A naturalistic feasibility study. *JMIR MHealth and UHealth*, 7(1). <https://doi.org/10.2196/10948>
- Martínez, V. (2023). Importancia de la evidencia científica en intervenciones escolares en salud mental. *Estudios Públicos*, 171. <https://doi.org/10.38178/07183089/2132230128>
- Moreta, R., López, C., Gordón, P., Ortíz, W., & Gaibor, I. (2018). El bienestar subjetivo, psicológico y social como predictores de la salud mental en ecuatorianos. *Actualidades En Psicología*, 32(124). <https://doi.org/10.15517/ap.v32i124.31989>
- O'Daffer, A., Colt, S. F., Wasil, A. R., & Lau, N. (2022). Efficacy and Conflicts of Interest in Randomized Controlled Trials Evaluating Headspace and Calm Apps: Systematic Review. In *JMIR Mental Health* (Vol. 9, Issue 9). <https://doi.org/10.2196/40924>
- Olza, I., Fernández, P., González, A., Herrero, F., & Carmona, S. (2021). Propuesta de un modelo ecosistémico para la atención integral a la salud mental perinatal. *Revista de La Asociación Española de Neuropsiquiatría*, 41(139), 23–35. <https://doi.org/10.4321/S0211-57352021000100003>
- Organización Mundial de la Salud, OMS. (2022). OMS | Salud mental. *WHO*.
- Powell, A. C., Bowman, M. B., & Harbin, H. T. (2019). Reimbursement of Apps for Mental Health: Current Practices and Potential Pathways. *Jmir Mental Health*, 6(8).

- Preciado, A., Valles, M., & Lévano, D. (2021). *Importancia del uso de sistemas de información en la automatización de historiales clínicos, una revisión sistemática*.
http://scielo.sld.cu/scielo.php?pid=S1684-18592021000100012&script=sci_arttext&tlng=pt
- psicologiamente.com. (2020, February 17). *Modelo biopsicosocial: qué es y cómo entiende la salud mental*. <https://psicologiamente.com/psicologia/modelo-biopsicosocial>
- Reyes, J., & Soberanes, A. (2021). Tecnologías Ágiles para Construir un Punto de Venta Mediante una Arquitectura en Capas. *Programación Matemática y Software*, 13(2).
<https://doi.org/10.30973/progmat/2021.13.2/8>
- Rivera, J., & Sánchez, D. (2016). Inteligencia artificial ¿reemplazando al humano en la psicoterapia? *Escritos*, 24(53). <https://doi.org/10.18566/escr.v24n53.a02>
- Romero, A. (2023). Avances en la investigación de la Salud Mental y la Inteligencia Artificial. *Revistas científicas. Una. Py*.
<https://revistascientificas.una.py/index.php/DM/article/download/3942/3187>
- Sanchez, E. (2022). Tecnología, cada vez más inmersa en la salud. *Cámara de Comercio de Lima*.
- Shaikh, M., Vayani, A. H. M., Akram, S., & Qamar, N. (2022). Open-source electronic health record systems: A systematic review of most recent advances. *Health Informatics Journal*, 28(2). <https://doi.org/10.1177/14604582221099828>
- Tena Cucala, R., Cobo Gómez, J. V., Rex, C., Charlop, M. H., & Spector, V. (2019). Usabilidad, satisfacción y facilidad de uso de aplicaciones (apps) gratuitas en español para pacientes con Trastorno Bipolar. *Psicosomática y Psiquiatría*, 48(11).
- Uquillas, G., Puente, M., Mendoza, C., & Bonilla, J. (2023). El uso de la Telemedicina en el ámbito de la gestión administrativa hospitalaria. *Revista Habanera de Ciencias Médicas*,

22(6), 1–12.

<https://openurl.ebsco.com/contentitem/edb:177404195?sid=ebsco:plink:crawler&id=ebsco:edb:177404195&crl=c>

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Scao, T. Le, Gugger, S., ... Rush, A. M. (2019). *HuggingFace's Transformers: State-of-the-art Natural Language Processing*. <https://arxiv.org/abs/1910.03771v5>

Yanfei, M. (2018). Research on the Model of Mental Health Education in Colleges and Universities in the Age of Mobile Internet. *INTERNATIONAL SYMPOSIUM 2018 - EDUCATION MANAGEMENT AND INNOVATION*.

ANEXOS

ANEXO A: Encuesta realizada a la encargada del Departamento de Bienestar

Universitario



UNIVERSIDAD TECNICA DEL NORTE
FACULTAD DE CIENCIAS APLICADAS
CARRERA DE TELECOMUNICACIONES

Encuesta para Trabajo de Integración Curricular II

Tema: "Desarrollo de un prototipo electrónico para la automatización y análisis inteligente de notas clínicas en sesiones terapéuticas basado en Inteligencia artificial."

Objetivos: Desarrollar un sistema que aplique inteligencia artificial para la conversión eficiente de contenido auditivo a texto, con la finalidad que genere sugerencias y optimice procesos de toma de decisiones que son cruciales en el área de psicología.

Datos Generales

1. Nombre: Ashna Realpe
2. Edad: 30
3. Especialidad: Psicología Clínica
4. Años de experiencia en psicología clínica: 2 1/2
5. ¿Usted utiliza herramientas tecnológicas en tus sesiones?
 - Sí
 - No

Uso de Tecnología en la Práctica

6. ¿Usted utiliza o conoce de herramientas tecnológicas que sean su complemento en sus sesiones terapéuticas?
 - Sí
Cuales Pura Breath (aplicación de respiración)
 - No
7. ¿Qué tan cómodo/a usted se sentiría utilizando un sistema que grabe y analice las sesiones con sus pacientes con inteligencia artificial?
 - Muy cómodo/a
 - Algo cómodo/a
 - Neutral
 - Algo incómodo/a
 - Muy incómodo/a
8. ¿Considera usted que el uso de inteligencia artificial puede ser útil para procesar y analizar conversaciones terapéuticas?
 - Sí, sería muy útil
 - Podría ser útil en algunos casos
 - No estoy seguro/a

ANEXO B: Revisión de trabajos relacionados

- Trabajo relacionado 1

Ficha Bibliográfica	
Autor Nikolaos Flemotomos; Victor R. Martinez; Zhuohao Chen; Karan Singla; Victor Ardulov; Raghuv eer Peri; Derek D. Caperton; James Gibson; Michael J. Tanana; Panayiotis Georgiou; Jake Van Epps; Sarah P. Lord; Tad Hirsch; Zac E. Imel; David C. Atkins; Shrikanth Narayanan.	Editorial The Psychonomic Society, Inc. / Springer Nature.
Título: “Automated evaluation of psychotherapy skills using speech and language technologies”.	
Año	Ciudad
https://doi.org/https://doi.org/10.3758/s13428-021-01623-4	
Contenido	
Arquitectura y procesamiento: El sistema sigue un pipeline compuesto por detección de actividad de voz (VAD), diarización de hablantes, reconocimiento automático de voz (ASR), refinamiento de reconocimiento del hablante y segmentación por enunciados. Se implementan controles de calidad que bloquean la generación del reporte si no se cumplen umbrales mínimos.	
Dataset: 5097 grabaciones reales recolectadas entre 2017 y 2020; 4268 sesiones procesadas exitosamente, con una duración promedio cercana a los 50 minutos.	
Relación con la tesis:	

Este trabajo sustenta el enfoque de automatización del análisis de sesiones terapéuticas mediante audio, validando la viabilidad de un pipeline similar al propuesto en esta tesis. La principal diferencia radica en que el presente proyecto incorpora un dispositivo edge (Raspberry Pi) para la captura y preprocesamiento del audio.

- **Trabajo relacionado 2**

Ficha Bibliografica	
Autor Amir Nassereldine; Dancheng Liu; Chenhui Xu; Ruiyang Qin; Yiyu Shi; Jinjun Xiong.	Editorial arXiv
Título: “PI-Whisper: Designing an Adaptive and Incremental Automatic Speech Recognition System for Edge Devices”	
Año 2024	Ciudad
https://arxiv.org/pdf/2406.15668	
Contenido	
<p>Objetivo del estudio:</p> <p>Diseñar un sistema ASR adaptable para dispositivos edge que permita personalización incremental mediante perfiles LoRA, abordando los retos de variabilidad de hablantes, adaptabilidad e inclusión.</p> <p>Arquitectura y procesamiento:</p> <p>El sistema añade módulos de identificación de características del hablante, recuperación dinámica de perfiles LoRA y composición adaptativa previa a la</p>	

inferencia ASR. La representación del audio se basa en espectrogramas logarítmicos derivados de DFT.

Plataforma:

Entrenamiento en GPU (A6000) e inferencia evaluada en Raspberry Pi 5 y Jetson

Orin Nano.

Relación con la tesis:

Valida académicamente el uso de Raspberry Pi como plataforma viable para sistemas ASR en campo, reforzando la decisión de diseño del prototipo y justificando la separación entre entrenamiento pesado y despliegue ligero

- Trabajo relacionado 3

Ficha Bibliografica	
Autor Santosh Gondi; Vineel Pratap.	Editorial MDPI (revista <i>Electronics</i>).
Título: "Performance Evaluation of Offline Speech Recognition on Edge Devices".	
Año: 2021	Ciudad: Basel
https://doi.org/https://doi.org/10.3390/electronics10212697	
Contenido	
<p>Objetivo del estudio:</p> <p>Evaluar el rendimiento de sistemas ASR offline en dispositivos edge, priorizando privacidad y reducción de dependencia de red frente a soluciones basadas en servidor.</p> <p>Arquitectura y optimización:</p> <p>El estudio analiza el despliegue de modelos ASR mediante PyTorch y TorchScript, aplicando cuantización dinámica para reducir latencia y consumo de memoria. Se reporta la necesidad de limitar la duración de audios para evitar bloqueos en inferencia.</p> <p>Plataforma:</p> <p>Raspberry Pi 4B y Jetson Nano.</p> <p>Relación con la tesis:</p> <p>Proporciona una base técnica sólida para justificar el procesamiento local o híbrido del audio y la segmentación temporal como estrategia para mejorar latencia y estabilidad en el prototipo desarrollado.</p>	

- **Trabajo relacionado 4**

Ficha Bibliografica	
Autor Davide Mulfari; Massimo Villari	Editorial MDPI — <i>Electronics</i>
Título: <i>A Voice User Interface on the Edge for People with Speech Impairments.</i>	
Año: 2024	Ciudad: Basel
https://doi.org/https://doi.org/10.3390/electronics13071389	
Contenido	
<p>Objetivo del estudio:</p> <p>Diseñar una interfaz por voz desplegada en edge para personas con alteraciones del habla, utilizando Whisper y optimizaciones para ejecución local.</p> <p>Procesamiento y optimización:</p> <p>Se emplea <i>whisper.cpp</i> y modelos convertidos a formato ggml para reducir carga computacional. Los audios son cortos (2–5 s) y se procesan localmente.</p> <p>Plataforma:</p> <p>Raspberry Pi 5.</p> <p>Relación con la tesis:</p> <p>Demuestra la viabilidad práctica del uso de Whisper en dispositivos edge, reforzando la decisión de integrar ASR preentrenado en el prototipo sin necesidad de entrenamiento local.</p>	

- **Trabajo relacionado 5**

Ficha Bibliografica	
Autor Iván Froiz-Míguez; Paula Fraga-Lamas; Tiago M. Fernández-Caramés.	Editorial IEEE (revista IEEE Access).

<p>Título: <i>Design, Implementation, and Practical Evaluation of a Voice Recognition Based IoT Home Automation System for Low-Resource Languages and Resource-Constrained Edge IoT Devices: A System for Galician and Mobile Opportunistic Scenarios.</i></p>	
<p>Año: 2023</p>	<p>Ciudad: A Coruña</p>
<p>https://doi.org/https://doi.org/10.3390/electronics13071389</p>	
<p>Contenido</p>	
<p>Objetivo del estudio:</p> <p>Diseñar y evaluar un sistema de reconocimiento de voz en edge para lenguajes de bajos recursos, priorizando operación local y eficiencia computacional.</p> <p>Arquitectura:</p> <p>Pipeline: audio → ASR → detección de keywords → ejecución de acciones IoT. Se aplican técnicas de cuantización y destilación para optimizar el rendimiento.</p> <p>Plataforma:</p> <p>Raspberry Pi 4 y dispositivos Android.</p> <p>Relación con la tesis:</p> <p>Refuerza el enfoque de procesamiento en edge con recursos limitados, validando la arquitectura modular adoptada en esta investigación y la necesidad de adaptar modelos ASR al contexto específico de uso.</p>	

ANEXO C: Requerimientos

Requerimiento de Stakeholders			
Nomenclatura	Requerimiento	Descripción	Prioridad
StRS1	Autenticación de usuarios	La autenticación de usuarios es un requisito crítico en sistemas que gestionan información clínica sensible, ya que permite garantizar el control de acceso y prevenir usos no autorizados. Diversos estudios y normativas de seguridad de la información establecen que la autenticación es un pilar fundamental para la protección de datos en sistemas de salud digitales (ISO/IEC 27001, 2022; Fernández-Alemán et al., 2013).	Alta
StRS2	Consentimiento informado del paciente	El consentimiento informado constituye un principio ético y legal obligatorio en la recolección, grabación y análisis de información clínica. Su prioridad es alta debido a que la falta de consentimiento invalida el uso de la información y vulnera principios bioéticos como la autonomía y la confidencialidad del paciente (Beauchamp & Childress, 2019; Organización Mundial de la Salud, 2017).	Alta

StRS3	Confidencialidad de la información clínica	La confidencialidad de los datos clínicos es esencial para preservar la privacidad del paciente y la confianza en los sistemas de salud. Se asigna prioridad alta debido a que la exposición de información sensible puede generar consecuencias éticas, legales y profesionales, tal como lo señalan los estándares internacionales de protección de datos en salud (ISO/IEC 27799, 2016; HIPAA, 2013).	Alta
StRS4	Facilidad de uso para el profesional	La usabilidad del sistema influye directamente en su adopción por parte del profesional de la salud. No obstante, al no comprometer de forma directa la seguridad ni la integridad de la información clínica, se establece una prioridad media, coherente con modelos de calidad de software como ISO/IEC 25010 (ISO/IEC, 2011).	Media
StRS5	Acceso a informes clínicos	El acceso a los informes generados es fundamental para apoyar la toma de decisiones clínicas y el seguimiento terapéutico. Se asigna prioridad alta debido a que esta funcionalidad representa uno de los objetivos centrales del sistema de apoyo clínico propuesto (Shortliffe & Cimino, 2014).	Alta

Requerimientos del sistema			
Nomenclatura	Requerimiento	Descripción	Prioridad
RS1	Captura de audio de sesiones	La captura de audio es la entrada principal del sistema y habilita los procesos posteriores de transcripción y análisis. Su prioridad es alta por ser indispensable dentro del flujo funcional (Jurafsky & Martin, 2023).	Alta
RS2	Transcripción automática	La transcripción convierte el audio en texto estructurable, siendo un proceso clave para el análisis clínico asistido por IA. Se asigna prioridad alta debido a su impacto directo en la calidad del informe generado (Radford et al., 2023).	Alta
RS3	Generación de informes clínicos	La generación de informes presuntivos constituye el objetivo principal del sistema. Su prioridad es alta al representar la funcionalidad central de la solución propuesta (Topol, 2019).	Alta
RS4	Almacenamiento de sesiones	El almacenamiento garantiza trazabilidad y respaldo de la información clínica. Se asigna prioridad alta por su relevancia en seguimiento terapéutico y auditoría (ISO/IEC 27799, 2016).	Alta

RS5	Interfaz web de visualización	La interfaz permite la interacción del profesional con el sistema. Se asigna prioridad media al no afectar directamente la lógica principal del procesamiento (ISO/IEC 25010, 2011).	Media
-----	-------------------------------	--	-------

Requerimientos de Hardware y Software			
Nomenclatura	Requerimiento	Descripción	Prioridad
RHS1	Dispositivo embebido de bajo consumo	El uso de un dispositivo embebido permite capturar audio de forma continua con bajo consumo energético. Su prioridad es alta por su impacto en portabilidad y eficiencia (Gondi & Pratap, 2021).	Alta
RHS2	Micrófono de calidad clínica	Un micrófono adecuado mejora la calidad del audio y, por ende, la precisión de la transcripción. Se asigna prioridad alta debido a su influencia directa en el desempeño del ASR (Jurafsky & Martin, 2023).	Alta
RHS3	Backend con capacidad de procesamiento	El backend debe soportar procesamiento de ASR y NLP. Se asigna prioridad alta al ser esencial para la ejecución del sistema (Bass et al., 2013).	Alta

RHS4	Conectividad de red estable	La conectividad influye en tiempos de carga y respuesta. Se asigna prioridad media al depender del contexto de uso (ISO/IEC 25010, 2011).	Media
RHS5	Sistema operativo compatible	El sistema operativo debe garantizar estabilidad y compatibilidad con librerías de IA. Se asigna prioridad alta por su impacto en la confiabilidad del sistema (IEEE Std 830, 1998).	Alta

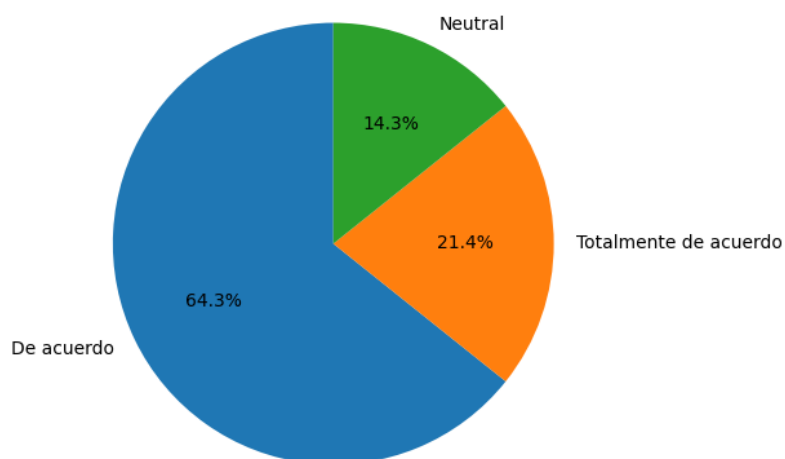
Requerimientos de modelos			
Nomenclatura	Requerimiento	Descripción	Prioridad
RM1	Modelo ASR preentrenado	El uso de un modelo ASR preentrenado permite convertir audio a texto sin necesidad de entrenamiento local. Su prioridad es alta por su eficiencia y precisión demostrada (Radford et al., 2023).	Alta
RM2	Modelo NLP/LLM clínico	El modelo de lenguaje debe generar informes clínicos estructurados. Se asigna prioridad alta por ser el núcleo del análisis inteligente (Topol, 2019).	Alta

RM3	Coherencia y consistencia del modelo	La coherencia en la generación de informes es esencial para su utilidad clínica. Se asigna prioridad alta debido a su impacto directo en la aceptación profesional (Shortliffe & Cimino, 2014).	Alta
RM4	Interpretabilidad del modelo	La interpretabilidad permite al profesional comprender y validar los resultados. Se asigna prioridad media al tratarse de un requisito deseable pero no crítico en la fase de prototipo (Doshi-Velez & Kim, 2017).	Media
RM5	Posibilidad de selección de modelos	Permitir la selección entre distintos modelos aumenta flexibilidad y adaptabilidad clínica. Se asigna prioridad media por su relevancia en futuras iteraciones (Topol, 2019).	Media

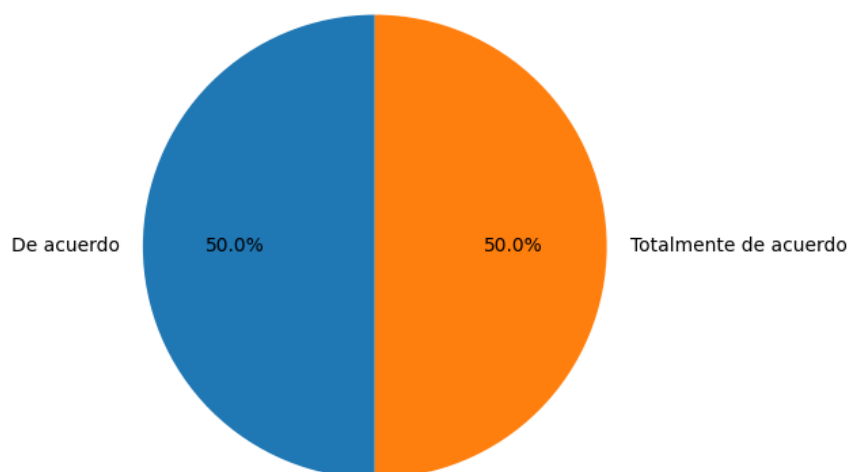
ANEXO C: Resultados de la encuesta

Bloque A

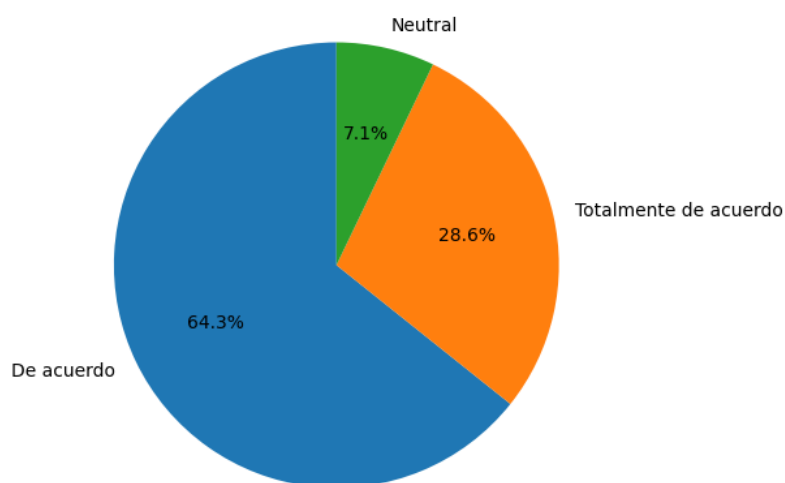
A1. El informe presenta una estructura clara y lógica.



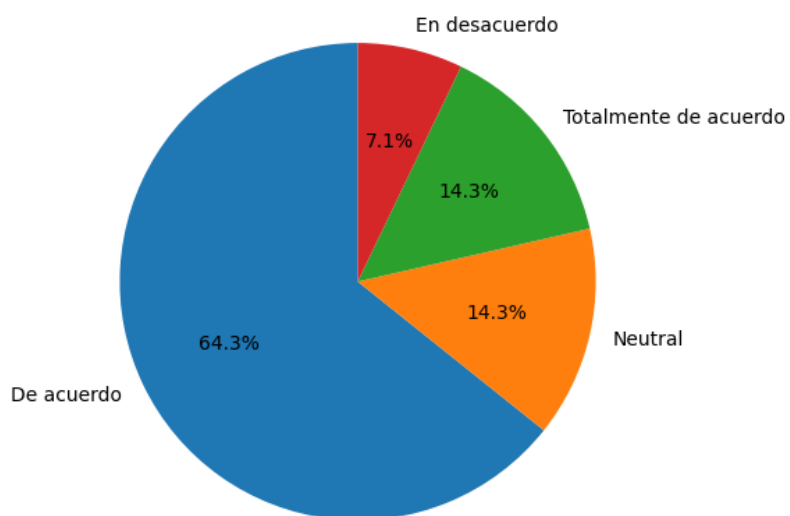
A2. El lenguaje clínico es comprensible para profesionales del área.



A3. Existe coherencia entre síntomas, análisis y conclusiones.

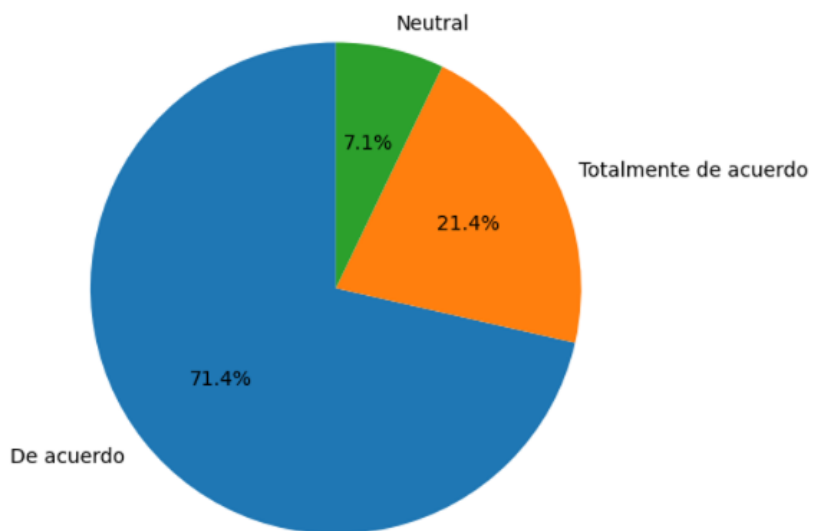


A4. El informe evita contradicciones internas.

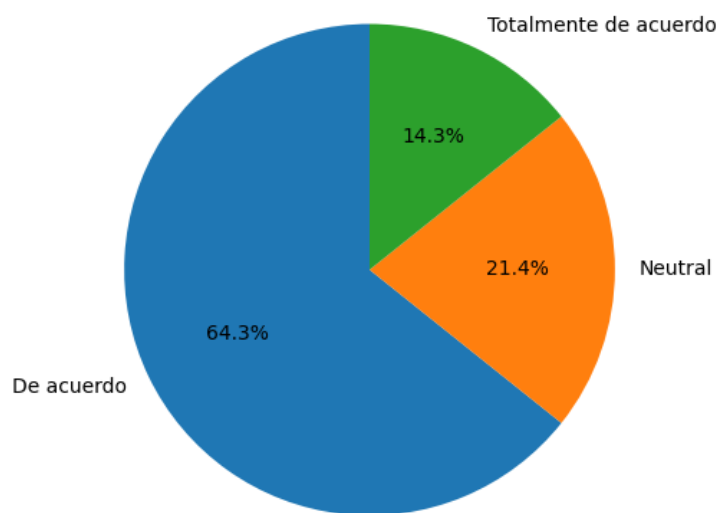


Bloque B

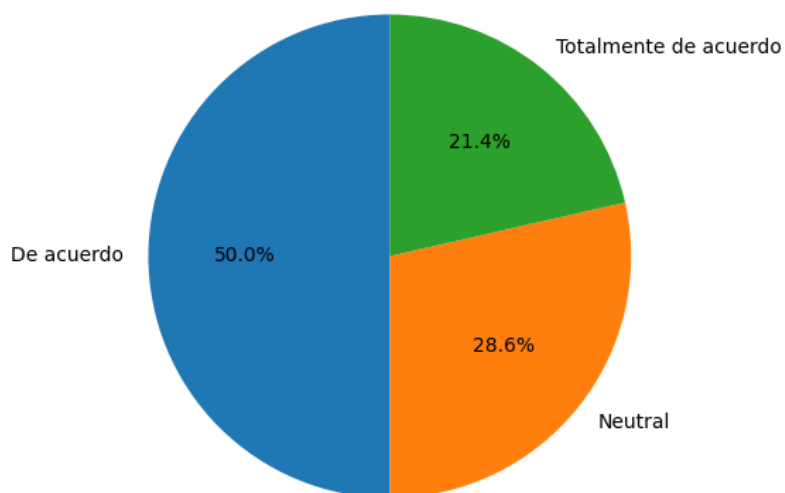
B1. El diagnóstico principal es clínicamente pertinente.



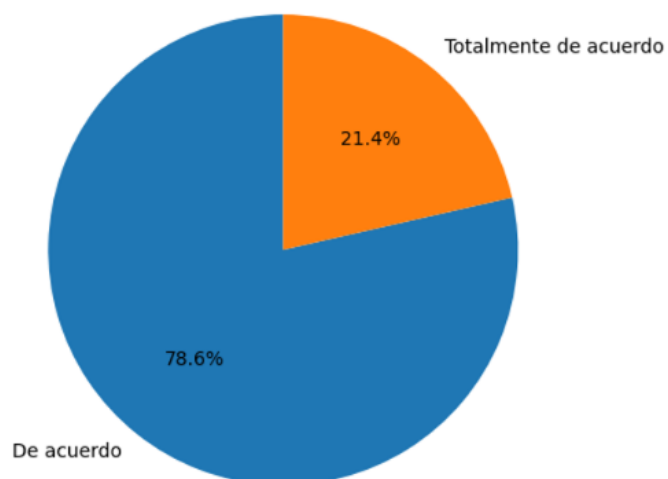
B2. Los criterios CIE-11 están correctamente aplicados.



B3. El diagnóstico diferencial está bien fundamentado.

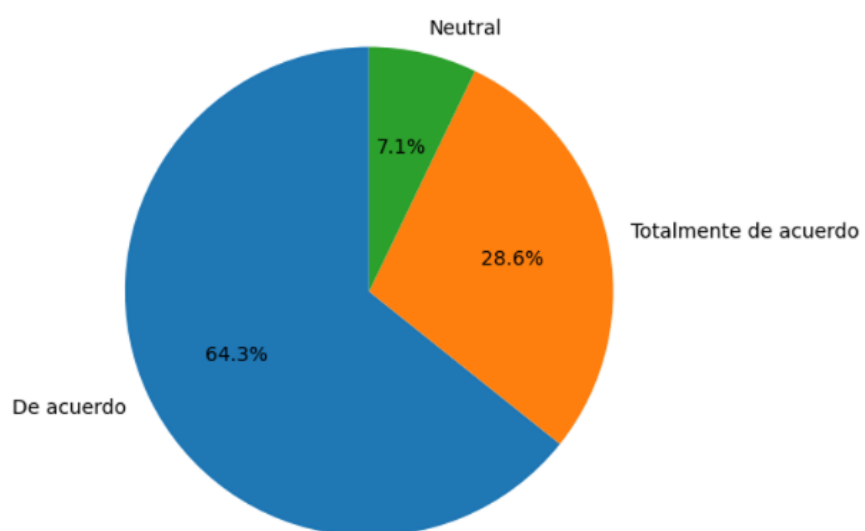


B4. La gravedad asignada es congruente con el cuadro descrito.

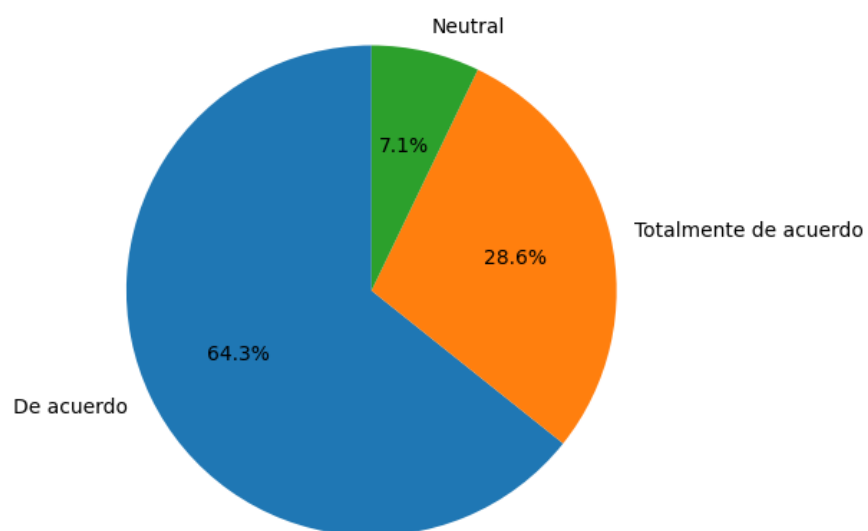


Bloque C

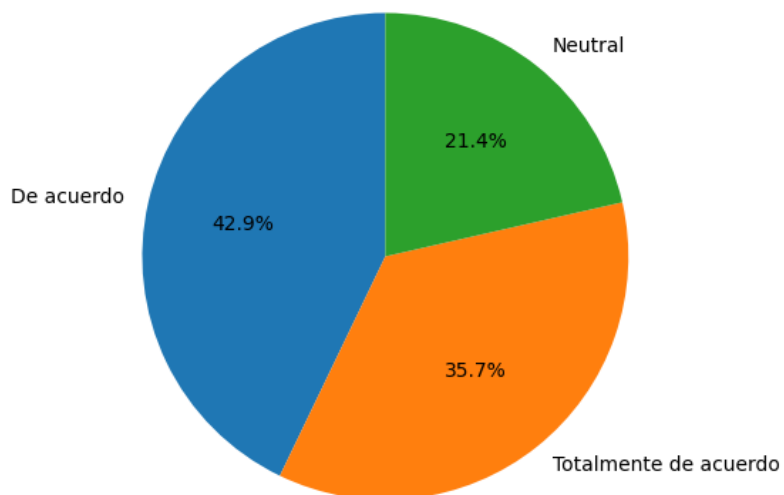
C3. El análisis respeta principios éticos.



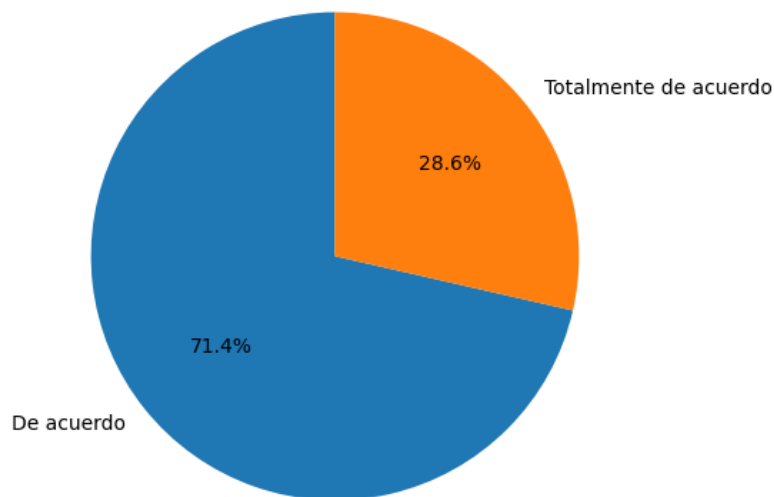
C1. El cuadro se asemeja a casos reales observados en la práctica.



C2. La interpretación clínica es realista.

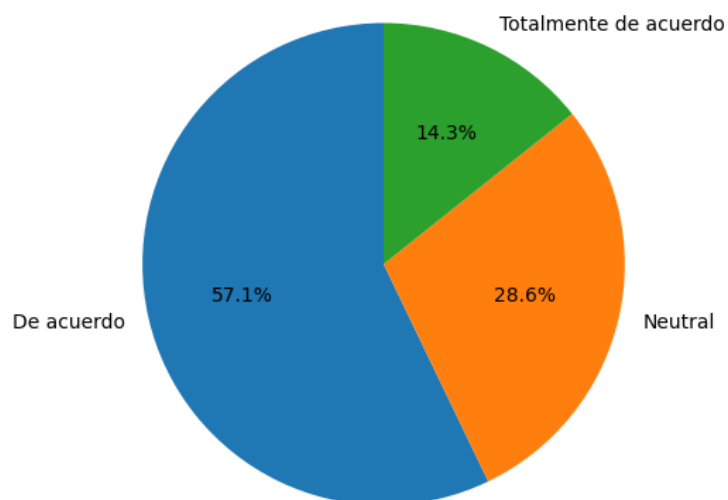


C4. El impacto funcional está adecuadamente descrito.

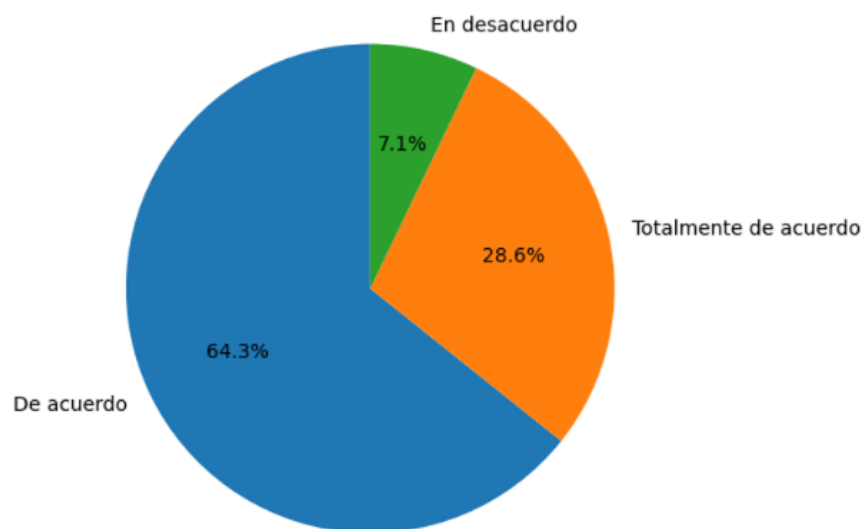


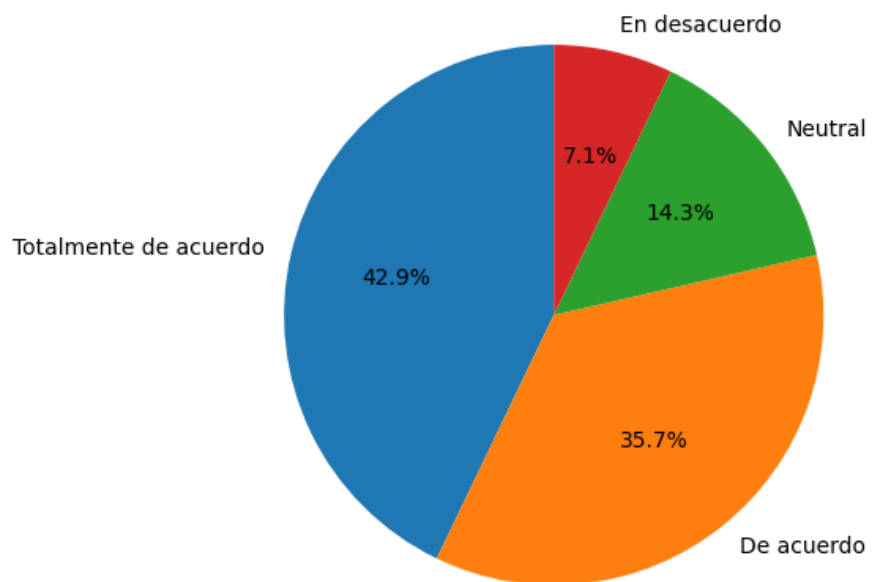
Bloque E

E1. Consideraría este informe en mi práctica profesional.

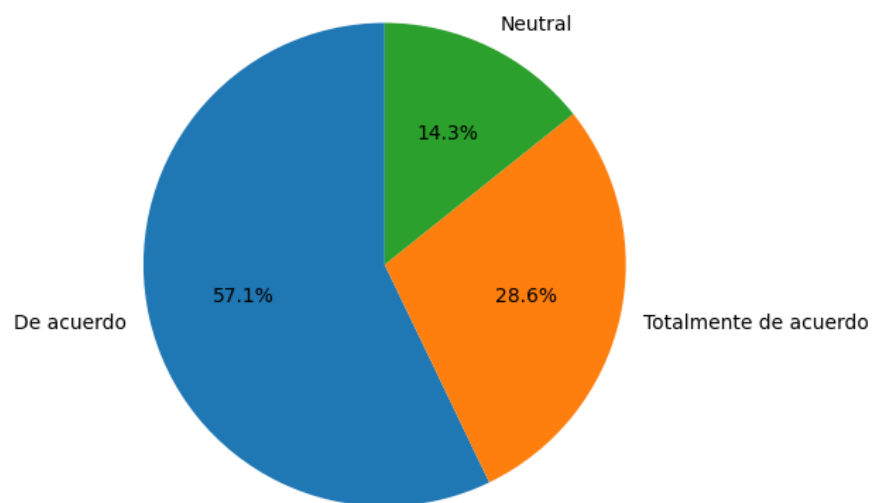


E2. El informe es éticamente responsable.





E4. El informe aporta valor clínico



ANEXO D: Consentimiento de la realización de las pruebas en el departamento de Bienestar Universitario



UNIVERSIDAD TÉCNICA DEL NORTE
Acreditada Resolución Nro. 173-SE-33-CACES-2020
FACULTAD DE INGENIERÍA EN CIENCIAS APLICADAS
TELECOMUNICACIONES



Memorando nro.: UTN-FICA-TL-2025-0080-M
Ibarra, 14 de julio 2025

PARA: Magister Catalina Ramírez
DECANA FICA

ASUNTO: Autorización para ejecutar una fase de pruebas controladas en el Área de Psicología del Departamento Bienestar Universitario.

Por medio del presente, muy respetuosamente, me dirijo a usted, en base a la petición del estudiante **Moreno Córdova Alexander Xavier**, quien a través de correo electrónico solicita la autorización para la realización de pruebas controladas relacionadas con el trabajo de grado, titulado: "Desarrollo de un prototipo electrónico para la automatización y análisis inteligente de notas clínicas en sesiones terapéuticas basado en inteligencia artificial", realizado por el estudiante antes mencionado, bajo la tutoría del Ing. Fabián Geovanny Cuzme Rodríguez, MSc.

El proyecto propone optimizar la elaboración y análisis de notas clínicas durante sesiones de psicoterapia, aportando a la eficiencia y calidad del servicio ofrecido por el área de Psicología del Departamento de Bienestar Universitario.

En este contexto, se solicita muy comedidamente a través de su intermedio se gestione a la Dirección de Bienestar Universitario, para que se autorice y brinde las facilidades necesarias para la ejecución de una fase de pruebas controladas en el Área de Psicología de dicha dependencia, con la finalidad de validar funcionalmente el prototipo en un entorno real, para lo cual se requiere:

- Que las pruebas sean realizadas en los espacios previamente designados, procurando no interferir con las actividades académicas programadas, con el fin de verificar el funcionamiento del prototipo y realizar los ajustes necesarios en caso de requerirse.

Agradezco de antemano su atención

Atentamente,

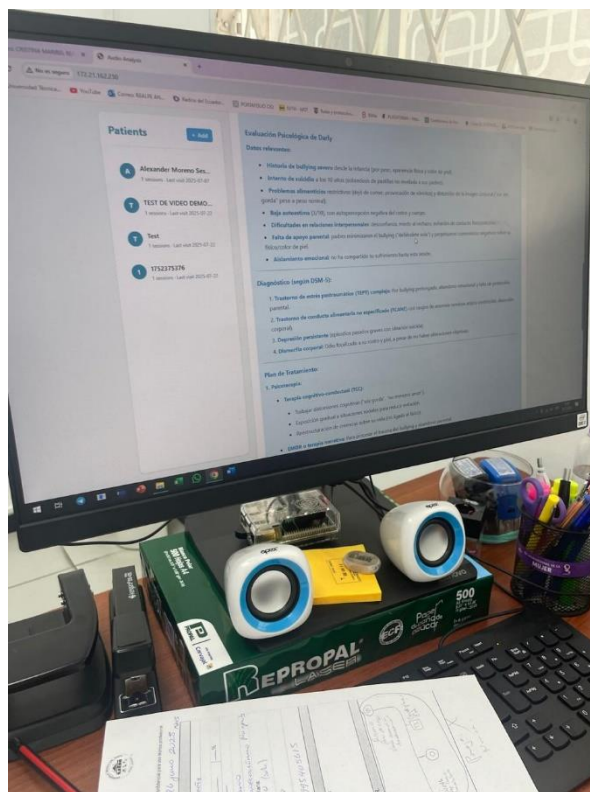
CIENCIA Y TÉCNICA AL SERVICIO DEL PUEBLO

JAIME ROBERTO MICHELENA
CALDERON
C.C.: JAIME ROBERTO MICHELENA
CALDERON
wwwNumber: 1002196433-20643411
TISA, S.A., ENTIDAD DE CERTIFICACION
DE INFORMACION, S.A. SECURITY DATA
S.A. 2, S.A.C.
2025.07.14 11:50:09 -05'00'

Ing. Jaime Michilena, MSc
COORDINADOR CARRERA TELECOMUNICACIONES

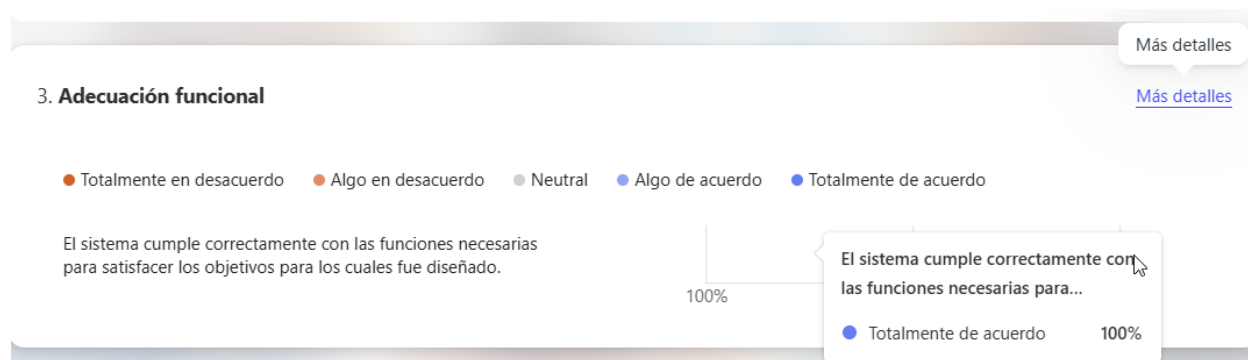
JMC/p.garcia

ANEXO E: Realización de pruebas dentro del Departamento de Bienestar Universitario

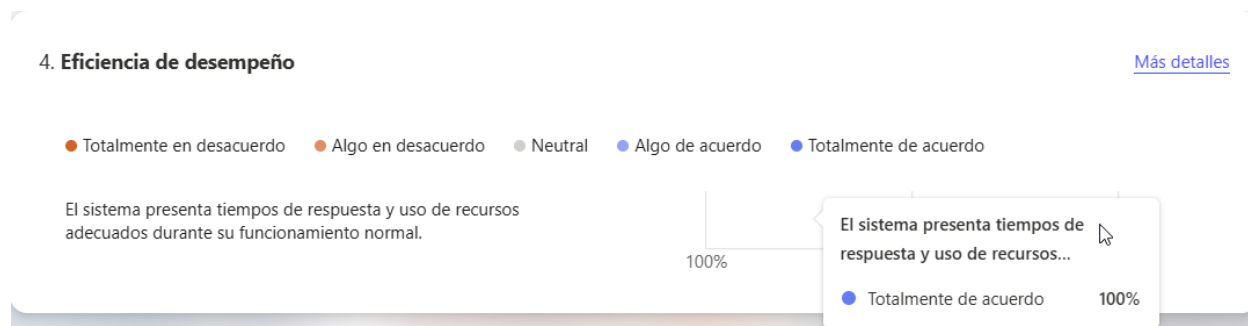


ANEXO F: Resultados de la encuesta sobre el cumplimiento de la ISO/UEC 25010

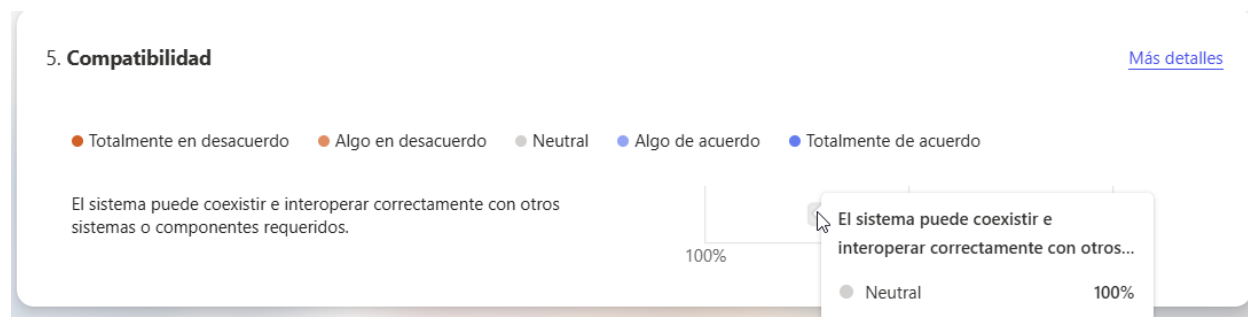
Pregunta de calidad 25010 - Adecuación funcional



Pregunta de calidad 25010 - Eficiencia de desempeño



Pregunta de calidad 25010 - Compatibilidad



Pregunta de calidad 25010 - Capacidad de Interacción

6. Capacidad de interacción (Usabilidad)

[Más detalles](#)

Totalmente en desacuerdo
 Algo en desacuerdo
 Neutral
 Algo de acuerdo
 Totalmente de acuerdo

El sistema es fácil de comprender, aprender y utilizar por parte del usuario.

100%

El sistema es fácil de comprender, aprender y utilizar por parte del...

Totalmente de acuerdo
 100%

Pregunta de calidad 25010 - Fiabilidad

7. Fiabilidad

[Más detalles](#)

Totalmente en desacuerdo
 Algo en desacuerdo
 Neutral
 Algo de acuerdo
 Totalmente de acuerdo

El sistema mantiene un funcionamiento estable y continuo, incluso ante fallos esperables.

100%

El sistema mantiene un funcionamiento estable y continuo,...

Totalmente de acuerdo
 100%

Pregunta de calidad 25010 - Seguridad

8. Seguridad

[Más detalles](#)

Totalmente en desacuerdo
 Algo en desacuerdo
 Neutral
 Algo de acuerdo
 Totalmente de acuerdo

El sistema protege adecuadamente la información y garantiza el acceso solo a usuarios autorizados.

100%

El sistema protege adecuadamente la información y garantiza el acceso so...

Algo de acuerdo
 100%

Pregunta de calidad 25010 - Flexibilidad

9. Flexibilidad

[Más detalles](#)

Totalmente en desacuerdo
 Algo en desacuerdo
 Neutral
 Algo de acuerdo
 Totalmente de acuerdo

El sistema puede adaptarse a nuevos entornos, requisitos o condiciones de uso sin afectar su funcionamiento.

100%

El sistema puede adaptarse a nuevos entornos, requisitos o condiciones d...

Totalmente de acuerdo
 100%

Pregunta de calidad 25010 – Mantenibilidad

10. **Mantenibilidad**[Más detalles](#)

Totalmente en desacuerdo
 Algo en desacuerdo
 Neutral
 Algo de acuerdo
 Totalmente de acuerdo

El sistema puede ser analizado, modificado y probado con facilidad cuando se requieren cambios.

100%

El sistema puede ser analizado, modificado y probado con facilidad...

Totalmente de acuerdo 100%

Pregunta de calidad 25010 - Protección

11. **Protección (Safety)**[Más detalles](#)

Totalmente en desacuerdo
 Algo en desacuerdo
 Neutral
 Algo de acuerdo
 Totalmente de acuerdo

El sistema minimiza riesgos para el usuario y el entorno durante su operación.

100%

El sistema minimiza riesgos para el usuario y el entorno durante su...

Totalmente de acuerdo 100%

Pregunta de calidad 25010 - Calidad global del producto

12. **Calidad global del producto**[Más detalles](#)

Totalmente en desacuerdo
 Algo en desacuerdo
 Neutral
 Algo de acuerdo
 Totalmente de acuerdo

En general, considero que el sistema presenta una calidad adecuada para su uso en el contexto previsto.

100%

En general, considero que el sistema presenta una calidad adecuada para...

Totalmente de acuerdo 100%

Firmas de autorización

<p>Msc. Fabián Geovanny Cuzme Rodríguez</p> <hr/> <p>Director</p>	<p>Sr. Alexander Xavier Moreno Córdova</p> <hr/> <p>Estudiante</p>
--	---