

Reporte final OlivIA for her

INTRODUCCIÓN	2
RESUMEN EJECUTIVO	3
1 ANTECEDENTES	3
1.1 FAIR LAC	3
1.2 FAIR LAC Jalisco	3
1.3 La viabilidad de un proyecto de IA y género	4
2 OLIVIA IA	6
2.1 Descripción del proyecto	7
2.2 Adquisición y preprocesamiento de datos	11
2.3 Entrenamiento del modelo	13
2.4 Métricas de Desempeño del modelo	16
2.5 Plataforma	17
2.6 Situación de pilotaje	17
3 ÉTICA Y GOBERNANZA DE DATOS	18
4 RESULTADOS OBTENIDOS EN LA IMPLEMENTACIÓN	20
5 LECCIONES APRENDIDAS	21
6 MEJORAS FUTURAS Y RECOMENDACIONES	24
7 REFERENCIAS BIBLIOGRÁFICAS	27
8 ANEXOS	27

Índice de tablas

Tabla n°1-Primeros pilotos propuestos

Tabla n°2- Gestión del proyecto por parte del Tecnológico de Monterrey y fAIR LAC Jalisco

Tabla n°3- Equipo interdisciplinario de DAMVV

Tabla n°4- Desarrolladores técnicos

Tabla 5-Preguntas analizadas por el modelo

INTRODUCCIÓN

En las últimas décadas, México está pasando por una compleja situación de violencia, donde la tasa de homicidios y desapariciones van al alza. Según la información del Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública (SESNSP) hay registrados más de 318,384 delitos y hasta febrero de 2022 hubo 16,356 mujeres víctimas de algún tipo de violencia siendo las lesiones dolosas el más reportado (55.51%) (SESNSP, 2022).

En este mismo periodo, Jalisco aparece cuarto en la lista por la cantidad de presuntas víctimas mujeres de lesiones culposas (con 203 casos) y seis de sus municipios aparecen dentro de la lista de los 100 municipios con más presuntos delitos de feminicidio (Zapopan, Puerto Vallarta, Colotlán, Mezquitic, San Gabriel, Tlajomulco de Zuñiga) (SESNSP, 2022).

Por otro lado, según los datos de la Encuesta de Percepción Ciudadana sobre Calidad de Vida 2022 (Jalisco Cómo Vamos), 7.3% de las mujeres habían sufrido algún tipo de delito dentro de la Zona Metropolitana de Guadalajara. Igualmente, el promedio mensual de denuncias por violencia familiar en Jalisco en 2020 según el Reporte Anual del Código Violeta de la Secretaría de Igualdad Sustantiva entre Mujeres y Hombres (SISEMH) fue de 991 denuncias (Jalisco Cómo vamos, 2022) y, entre el 2018 y 2020, las llamadas de emergencia al número 911 por eventos referentes a violencia de género crecieron en un 161% (Gobierno de Jalisco, 2020).

En el país existe un elevado porcentaje de delitos que no se denuncian (93%) o que la autoridad no inició una carpeta de investigación (INEGI, 2021). Estos niveles de violencia en México son alarmantes y requieren soluciones urgentes tanto para incentivar las denuncias, como para que éstas sean procesadas adecuadamente de manera rápida.

En este sentido, las herramientas digitales pueden ayudar a facilitar los procesos de denuncia, aumentando la eficacia de la respuesta institucional y otorgando facilidades para el seguimiento de los casos.

Es así que entre la Secretaría de Igualdad Sustantiva entre Mujeres y Hombres del Gobierno de Jalisco y las instituciones que conforman la iniciativa fAIr LAC Jalisco

desarrollan el proyecto OlivIA for her, un piloto de automatización de las denuncias de mujeres víctimas de violencia mediante un modelo de inteligencia artificial de reconocimiento y transcripción de voz a texto, realizado entre junio del año 2021 a septiembre del 2023. En el siguiente informe se da cuenta del desarrollo, implementación y resultados obtenidos; así como los logros, desafíos y conclusiones clave de OlivIA for her, junto con recomendaciones para su implementación y expansión futura.

RESUMEN EJECUTIVO

1 ANTECEDENTES

1.1 FAIR LAC

Los avances tecnológicos están transformando la realidad en América Latina y en el resto del mundo. La incorporación de tecnologías digitales ofrece oportunidades para innovaciones en el aprendizaje. Sin embargo, si no se utiliza de manera ética y responsable, puede llegar a profundizar las brechas sociales existentes. Es esencial garantizar que aquellas personas afectadas por el uso de la tecnología comprendan su funcionamiento y puedan reivindicar sus derechos.

Ante las oportunidades de catalizar el desarrollo en los países de América Latina y el Caribe mediante el uso de la tecnología, BID Lab, el laboratorio de innovación del Grupo Banco Interamericano de Desarrollo (BID) lanza en 2019 la iniciativa fAIr LAC, una alianza entre los sectores público y privado, la sociedad civil y la academia, para incidir tanto en la política pública como en el ecosistema emprendedor en la promoción del uso responsable y ético de la Inteligencia Artificial (IA).

fAir LAC tiene como objetivo buscar la equidad y justicia algorítmica, la seguridad y protección de información y desarrollar IA centrada en el ser humano y en la rendición de cuentas (BID, 2023) con presencia en cuatro centros neuronales o “hub” ubicados en Costa Rica, Colombia, Uruguay y México. En el caso de México, el hub se encuentra en el estado de Jalisco.

1.2 FAIR LAC Jalisco

Es una iniciativa multisectorial del Grupo BID, ejecutada por el Tec de Monterrey campus Guadalajara, en alianza con el Gobierno de Jalisco y la agencia de innovación C Minds, que tiene como objetivo aprovechar la adopción responsable y ética de la IA para el bien social en Jalisco, a través de la articulación de la academia, la sociedad civil, el sector público y el sector privado para desarrollar capacidades, potencializar el alcance social de la IA y utilizar el impacto de la tecnología para servicios sociales más eficientes.

fAIr LAC Jalisco, tiene tres ámbitos de acciones:

- 1) Impulsar el ecosistema de emprendimiento en IA.
- 2) Implementar pilotos de IA para hacer frente a problemáticas sociales.
- 3) Desarrollar habilidades en innovación para el impulso de políticas públicas de bienestar social, desde el sector público y privado, utilizando modelos de IA.

Gracias al trabajo multidisciplinario y de diferentes organizaciones se han impulsado proyectos pilotos para identificar y analizar los beneficios y desafíos que surgen a partir de la adopción de sistemas de IA en el mejoramiento de diferentes servicios públicos. Bajo la iniciativa fAIr LAC Jalisco se desarrollaron dos proyectos pilotos: Retina IA (para la detección temprana de la retinopatía diabética y derivación oportuna de la misma, auxiliado con IA y telemedicina) y OlivIA for her (para la conversión de las denuncias de mujeres víctimas de violencia de audio a texto).

En el presente reporte se sistematiza y analiza la experiencia en el desarrollo del proyecto OlivIA for her.

1.3 La viabilidad de un proyecto de IA y género

En junio de 2021 la investigadora del Tecnológico de Monterrey, Tzinti Ramirez, comenzó una serie de entrevistas a diferentes autoridades de la SISEMH para comprender su forma de trabajo y analizar los principales desafíos que enfrentan en el desempeño de sus funciones diarias con la finalidad de identificar oportunidades que contribuyan a abordar dichos desafíos mediante IA. Las personas entrevistadas fueron:

- Valeria Mcmillán González, Directora de Desarrollo de Capacidades de las Mujeres.
- María Antonieta Vizcaíno, Directora de Fortalecimiento Municipal.

- María Elena García Trujillo, Subsecretaria de Acceso de las Mujeres a una Vida Libre de Violencia.

Producto del análisis de los procesos y sus desafíos, las principales dolencias “*pain points*” fueron identificadas en las diferentes áreas de la institución, hasta centrarse en los procesos de atención a mujeres víctimas de violencia en la Unidad Metropolitana de Atención Integral a Mujeres y Niñez ubicada en el centro de la ciudad de Guadalajara.

Posteriormente, en octubre del 2021 se realizó un taller con las instituciones participantes en el caso de uso con la finalidad de profundizar el diagnóstico de la problemática identificada y analizar los flujos de información al momento de atender a víctimas de violencia.

Los principales hallazgos fueron:

- El proceso para recabar información del relato de las víctimas implica la utilización de diversos formularios que deben ser llenados a mano en papel convencional y cuyos hallazgos son después capturados en computadora. Esto implica una inversión de tiempo importante en procesos de captura de información que podrían utilizarse para el proceso de atención y asesoría per se.
- El personal de la Unidad de Atención posee una alta carga de trabajo, ya que es el primer punto de atención a víctimas y una de las tareas que más demanda tiempo es la transcripción fidedigna de los relatos, además del llenado de los formularios.
- Frecuentemente las mujeres debían narrar reiteradamente lo que les ha sucedido a diferentes funcionarios pertenecientes a la misma o distintas instancias de gobierno que se involucran eventualmente en su atención y seguimiento.
- El seguimiento en la atención de los casos depende de la capacidad de programación de recordatorios para seguimientos periódicos que tienen los equipos que brindan la asesoría; dichos seguimientos pueden verse retrasados tanto por la carga de trabajo, como por un olvido en la programación de los recordatorios.

Con este diagnóstico el equipo de fAIr LAC Jalisco propuso y evaluó propuestas para desarrollar un proyecto piloto para el uso e inteligencia artificial en la atención a mujeres víctimas de violencia considerando qué tan disruptivo sería para el proceso actual de atención.

Tabla n°1- Primeros pilotos propuestos

Proyecto	Incidencia	Disrupción de proceso actual de atención
1-Registro y vaciado automático del relato de una víctima en su expediente digital	Alta	Baja
2-Sistematización de evaluación de riesgo	Baja	Baja
3-Unificación de bases de datos internas de la DAMVV y SISEMH	Alta	Alta
4-Pre llenado de formulario de denuncia	Baja	Alta
5-Expediente único electrónico entre dependencias	Alta	Alta

La decisión se tomó en conjunto con la SISEMH contemplando criterios como la cantidad de información disponible, nivel de avance en el proceso de diagnóstico, sumado a las estrategias de obtención de información, gestión de la información, necesidades tecnológicas de software y hardware, tamizajes y personal operativo.

Finalmente, se optó por la propuesta de un registro y vaciado automático del relato de una víctima en su expediente digital, por medio de un modelo de conversión de audio a texto asistido por IA. Abordar este primer paso resultaba relevante para posteriores etapas como la sistematización de evaluación de riesgos o la unificación de bases internas.

2 OLIVIA IA

Ante la necesidad de digitalizar procesos para transitar hacia la construcción de un sistema de atención a víctimas que esté homologado entre las diversas áreas de gobierno del estado, se detectó, en conjunto con la Secretaría de Igualdad Sustantiva entre Mujeres y

Hombres (SISEMH), el área de oportunidad de mejorar los puntos de primer acercamiento con mujeres víctimas de violencias mediante el almacenamiento digital de sus denuncias. Esto permite, por un lado, escalar los relatos entre las áreas sin la necesidad de que las víctimas cuenten en reiteradas ocasiones los hechos sucedidos, así como agilizar el posible levantamiento de una denuncia al acortar las brechas burocráticas.

2.1 Descripción del proyecto

Problema a resolver

Las mujeres que buscan atención directa se ven obligadas a repetir las agresiones vividas en reiteradas ocasiones al ser derivadas a diferentes órdenes de gobierno e instituciones de atención según la violencia experimentada. Esto les genera una sensación de cansancio, revictimización, mala percepción en atención y desconfianza a las instituciones democráticas.

Por otro lado, las personas servidoras públicas recaban la información a mano y la transcriben en computadora, lo cual implica una pérdida importante de tiempo y una ventana a mayores errores humanos como la interpretación, parafraseo u omisión de información importante para la víctima. De tal forma que:

- El personal de la Unidad de Atención posee una alta carga de trabajo, ya que es el primer punto de atención a víctimas y una de las tareas que más demanda tiempo es la transcripción fidedigna de los relatos, además del llenado de los formularios.
- El detalle de los casos se guarda en su mayor parte en formato Excel.
- El seguimiento en la atención de los casos depende de la capacidad de los equipos que brindan la asesoría de programar recordatorios para seguimientos periódicos; dichos seguimientos pueden verse retrasados tanto por la carga de trabajo como por un olvido en la programación de los recordatorios.

Objetivo

Disminuir las posibilidades de revictimización y trauma de las mujeres que buscan atención en las Unidades de apoyo gubernamentales, a través, de eficientar el tiempo y las herramientas de recopilación y sistematización de información de manera sensible, empática y eficaz del personal correspondiente.

El modelo de OlivIA debe ser capaz de realizar el trabajo de transcripción y llenado de la Cédula de Registro Único (llamado EVI) y de la sábana informativa que actualmente llevan a cabo las trabajadoras de la Unidad de Atención.

Usuarios del caso de uso

Equipo interdisciplinario de la Dirección de Atención a Mujeres Víctimas de Violencias (DAMVV), quienes son el primer punto de contactos con la gente se atreve a denunciar.

Beneficiarios finales

Mujeres mayores de 18 que hablen español¹ que presentan alguna denuncia o proceso en DAMVV.

Para asegurar la trazabilidad del piloto, el proyecto se dividió en dos etapas:

Etapas 1: Desarrollar el prototipo de conversión de audio a texto y resumen del relato para evitar la revictimización; diseñar el proceso e infraestructura, además de la capacitación necesaria para el funcionamiento y seguridad de los datos sensibles durante los meses de julio-septiembre de 2022.

Etapas 2: Generar un prototipo de inteligencia artificial, diseñando el proceso, además de validar su funcionamiento a través de la instalación de equipo y de la prueba piloto, capacitando al personal durante los meses de octubre 2022 a septiembre 2023.

Beneficios previstos

- Disminución del número de veces que la víctima debe relatar la agresión y violencia que ha vivido, es decir, re victimización.
- Disminución del tiempo invertido en el vaciado de la información, de papel a digital.
- Mejor percepción de apoyo empático y cercano de las mujeres atendidas por parte de la institución .

¹ Por temas de protección de datos se excluyó a mujeres menores de edad y por temas de entrenamiento del modelo se excluyó a personas que hablen lenguas indígenas. Cabe destacar que si las mujeres no querían participar en el proyecto de OlivIA, en ningún momento se les negó la atención en SISEMH, simplemente no se grababa su relato y la atención se realizaba de manera tradicional.

- El programa podrá brindar la materia prima (grabaciones y sistematización) para que la contraparte obtenga los datos que permitan otorgar información a los y las tomadoras de decisiones sobre el desarrollo de políticas públicas.
- Favorecer los procesos de automatización que están ocurriendo dentro de SISEMH, ya que en el año 2022 iniciaron una actualización de sus plataformas y documentos de trabajo.

Equipo de trabajo

Hubo tres grupos de trabajo encargado de diferentes áreas:

1.Gestión del proyecto: representada por el equipo del Tec de Monterrey y de fAIr LAC Jalisco, quienes se encargaron del diseño de la solución, la coordinación del caso de uso llamado OlivIA y el asesoramiento del manejo ético de los datos.

2.Desarrollo operativo: equipo interdisciplinario representantes de la Dirección de Atención a Mujeres Víctimas de Violencias (DAMVV), ellos participaron en el análisis e implementación de las pruebas piloto como en la atención a personas de primer contacto.

3.Desarrollo técnico: encargados del diseño e iteraciones del modelo de IA junto con el desarrollo de la plataforma donde se graban y almacenan los datos².

En las siguientes tablas se detallan los cargos de cada participante

Tabla n°2- Gestión del proyecto por parte del Tecnológico de Monterrey y fAIr LAC Jalisco

<i>n°</i>	<i>Nombre</i>	<i>Cargo</i>
1	<i>Enrique Cortés Rello</i>	<i>Director del Hub de Inteligencia Artificial del Tecnológico de Monterrey</i>
2	<i>Juan Roberto Hernández</i>	<i>Coordinador de fAIr Lac Jalisco.</i>
3	<i>Tatiana Lefno</i>	<i>Asistencia técnica y seguimiento de casos de uso. FAIr Lac Jalisco</i>
4	<i>Eva Hernández</i>	<i>Investigadora del Tecnológico de Monterrey.</i>
5	<i>Claudia Angélica Ramírez Ramos</i>	<i>Project manager y enlace técnico in situ/ Recabar información hacer pruebas, seguimiento, coordinación, investigación</i>
6	<i>Jorge Daniel Miramontes Romero</i>	<i>Asesor jurídico en gobernanza de datos</i>

² Para el desarrollo del modelo de IA se emplearon tres equipos técnicos en diferentes temporalidades, esto a raíz de los desafíos que presentó el caso de uso que se explican en el punto 2.3

7	Luis Enrique Vázquez de la Paz	Asesor jurídico
8	Juan Alberto Amézquita	Coordinador del Comité de Riesgos Éticos y de Gobernanza de Datos y asesor en ética
9	María Carelia Ibarra Vázquez	Prestadora de On Campus Jobs, en materia de gobernanza de datos desde la parte técnica
10	Amaranta Citlalli De la Rosa Rivera	UX Designer
11	Alyx Nadyezhna Gutiérrez Miranda	Asistente en materia de entrenamiento y asistencia al modelo
12	Daniela Viviana Pérez Nuño	Asistencia en la traducción y entrenamiento del modelo del caso.
13	Alaide Andrea Pérez Villegas	Asistencia en la traducción y entrenamiento del modelo del caso
14	Tzinti Ramírez	Investigadora del Tecnológico de Monterrey.

Tabla n°3- Equipo interdisciplinario de DAMVV

n°	Nombre	Cargo
1	Luis Gustavo Montes Rivera	Abogacía
2	Iliana Rivera	Trabajo Social
3	Karina Vázquez Zamora	Abogacía
4	Jennifer Sujey Díaz Orozco	Trabajo Social
5	Tania Trujillo	Psicología
6	María Elena García Trujillo	Subsecretaria de Acceso de las Mujeres a una Vida Libre de Violencia.
7	Luz Elena Rosas Hernández	Directora de Atención a Mujeres Víctimas de Violencias.

Tabla n°4- Desarrolladores técnicos

n°	Nombre	Cargo
1	Luis Reyna	Director de Servicios cloud
2	Alfredo Tovar	Arquitecto de soluciones
3	Arturo Vargas	Arquitecto de soluciones
4	Miguel Muñoz	Arquitecto de soluciones
5	Alexis Ivan Lopez Escamilla	Asesor Técnico
6	Esteban García Piñón	Desarrollador Machine Learning

7	Rigoberto Becerril Balderas	Desarrollador Fullstack
8	Liliana Millán	Asesora Técnica
9	Oscar Ignacio Martínez	Desarrollador Fullstack
10	David Laines	Desarrollador Machine Learning
11	María Anais Reyes Casanova	Desarrolladora FrontEnd

2.2 Adquisición y preprocesamiento de datos

Para hacer la grabación de la denuncia y su posterior conversión de audio a texto por medio de inteligencia artificial, primero fue necesario entender cómo se da la atención en SISEMH de manera regular y análoga. Para esto, se realizaron diferentes visitas e identificaciones del flujo de atención. A continuación, se muestra como es la atención en SISEMH de manera tradicional.

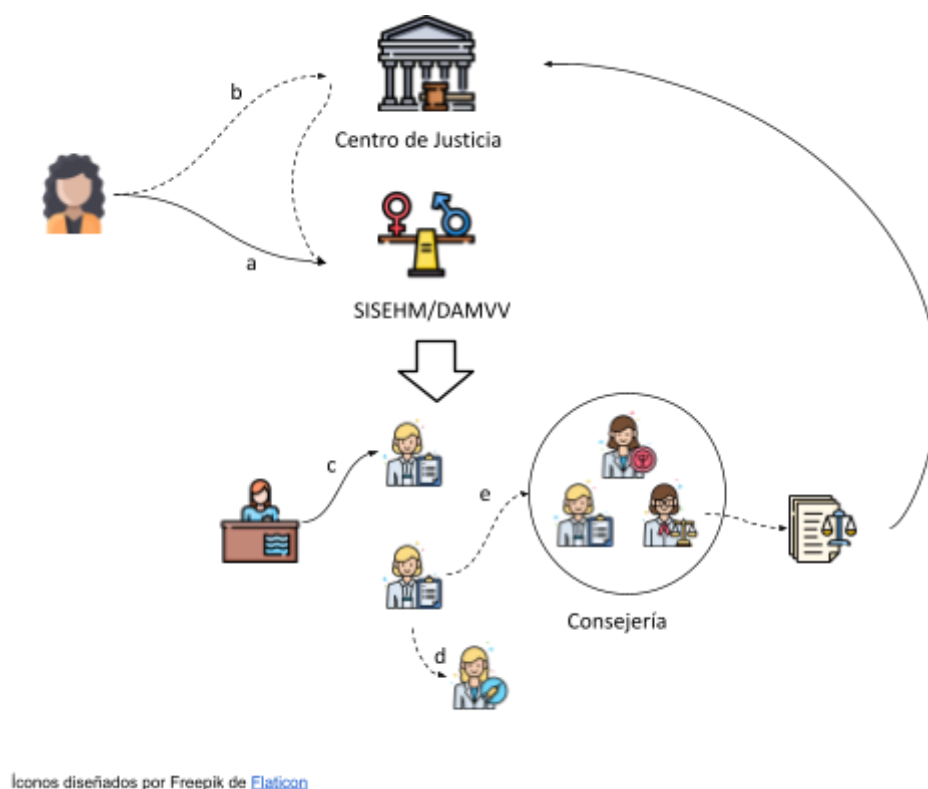


Figura n°1 proceso de atención tradicional
Fuente elaboración propia

Las mujeres pueden acudir físicamente a la Dirección de Atención a Mujeres Víctimas de Violencias (DAMVV) de SISEMH para solicitar ayuda como muestra el inciso a) en la Figura 1 o bien puede acudir directamente a la Red de Centros de Justicia para las Mujeres y/o

cualquier otra institución correspondiente b). Siempre pasará la víctima con otra trabajadora social para que lleve su caso. Inciso (c) en la Figura 1.

- i. Al llegar a DAMVV será atendida por una recepcionista que es trabajadora social, qui está capacitada para atender problemas de violencia contra las mujeres. La recepcionista:
- ii. Deriva a otra trabajadora social para que lleve su caso. Inciso ©. La trabajadora social aplica un formato de entrevista semiestructurada con la cual va llenando el expediente único de atención (EVI). Este documento se le asignan valores predeterminados a distintas preguntas cualitativas para determinar el nivel de peligro y detectar si es una emergencia o no. Algunos ejemplos de estas preguntas son, ¿Te amenaza con matarte? ¿Ha amenazado o intentado suicidarse alguna vez? ¿Está celoso de ti de forma violenta y constante?
 1. Si la persona receptora de violencia necesita algún tipo de atención médica o intervención en crisis psicológica será atendida por servicios médicos. Inciso (d) en la Figura 1.
 2. Si el riesgo de violencia es grave o extremo, pasará a Consejería conformada por una trabajo social, psicología, abogacía. Inciso (e) en la Figura 1.
 - a. En caso de que se requiera llevar a cabo algún proceso legal, entonces, la abogada genera el oficio necesario para que después la persona receptora de violencia pueda ingresarlo al Centro de Justicia (en caso de personas mayores de edad) y se inicie el proceso legal asociado.
 - b. Depende de la gravedad de riesgo, en caso de que se requiera llevar a la persona receptora de violencia a algún refugio, DAMVV hace acompañamiento hasta entregar a la persona receptora de violencia con el contacto oficial del refugio, o se acompaña al alojamiento temporal.

Los tiempos de atención son variados, ya que dependen del nivel de complejidad de cada caso y el motivo de la atención. Gracias al proceso diagnóstico del equipo operativo se pudo registrar que el rango de atención puede variar desde los 20 minutos hasta una hora.

Al finalizar cada semana de trabajo, una persona de la DAMVV es la encargada de transcribir de manera manual la información recopilada en la EVI a un archivo digital, que se

almacena en la nube, al cual llaman sábana, de éste se destaca el acceso controlado, que sólo es permitido al personal de primer contacto y tomadoras de decisiones del que se extraen todos los datos de las mujeres víctimas de violencia para su seguimiento y generación de indicadores mensuales. La información agrupada en este documento es también de utilidad para obtener tendencias locales.

Una vez identificado el proceso de atención, los esfuerzos se centraron en comprender a profundidad cómo se aplicaba la EVI y los datos que se levantan, ya que de aquí se obtiene la información que alimenta a la sábana. Es en este paso donde se incluye la conversión de audio a texto por medio de la IA.

Como se indicó anteriormente, SISEMH al momento de la implementación del proyecto piloto, SISEMH estaba en un proceso de actualización de sus procesos de atención, entre ellos, el ajuste de la EVI. Esto es importante de mencionar, ya que la última iteración del modelo de IA coincidió con estos ajustes, así para el entrenamiento final del modelo se usó un nuevo formulario llamado EVI. Esto es relevante porque al momento de la atención las trabajadoras sociales se enfrentaron a la adaptación tanto del formulario más la utilización de tecnología en la atención (grabar la entrevista desde la plataforma web).

Para llevar un proceso más organizado y apoyar al entrenamiento de los modelos de Inteligencia Artificial, se optó por excluir apartados de la EVI, ya que estos son de llenado en caso de que la persona atendida sea menos de edad o requiera las firmas de tutores o representación legal, así como también aquellos correspondientes al “tamizaje de riesgos” ya que, de acuerdo a la información proporcionada por la Secretaría, este sufriría cambios importantes durante el periodo de la implementación por causas ajenas a Olivia. Cabe que dentro de EVI, se descartaron los apartados relacionados con atención a menores de edad, firmas y recepción, por lo que se consideraron solo 7 apartados desagregados en 45 de las 127 preguntas para el entrenamiento de la IA³. Por ello, la traducción de esta información y su clasificación para llenar la sábana de excel es el alcance de esta primera versión auxiliada con el modelo de IA.

³ Para ver las preguntas, se puede revisar el anexo 1

2.3 Entrenamiento del modelo

En un primer momento, se trabajó con una empresa para diseñar el modelo de IA en AWS, pero se tuvieron limitaciones en el alcance de la transcripción y organización de la información. A saber:

- La transcripción de audio a texto era limitada (al 43%), esto hizo que tomara mucho tiempo su corrección. Lo cual entorpece el proceso de atención.
- El formato de audio a texto que ofrecía su plataforma solo aceptaba mp3, pero el formato de las tablet donde se grababa era m4a, esto implicaba un paso extra para quien aplica y graba la EVI, porque debía transformar los formatos de audio.
- A los funcionarios no les llegaba el correo de aceptación de usuario a sus correos institucionales y se tuvieron que crear usuarios con correos nuevos y exclusivos para esta plataforma.
- La solución que planteó la empresa desarrolladora fue reajustar el formato de entrevista a un guión estructurado para captar mejor el audio, pero al momento de su aplicación con las mujeres usuarias de la DAMVV se sintió inorgánico, cuadrado y re victimizante el proceso de atención. Esto implicó una dificultad importante para el proyecto porque se estaba poniendo en juego la atención a las mujeres por la integración de la tecnología. Estas pruebas duraron solo un par de días, ya que se priorizó la atención de calidad a las mujeres, más que ajustarse al nuevo formato de entrevista.

Con este contexto, se prefirió buscar otra alternativa para el desarrollo del modelo de IA que generará una mejor traducción sin sacrificar la atención a las usuarias.

Este segundo diseño tuvo menos problemas, y se optó por el desarrollo en Python y en open source, por las características y facilidades que otorga para el desarrollo de la tecnología. Finalmente, se terminó de afianzar el modelo con otro equipo quienes crearon un modelo para la transcripción de audio a texto, más otros 4 modelos para entrenar el llenado de la sábana a partir de la transcripción obtenida y un “*pipeline*” de procesamiento. A continuación, se describe cada modelo⁴

Transcripción de Audio a Texto:

- Modelo utilizado: Whisper de OpenAI en la instancia "small" con 244 millones de parámetros, para convertir audio en español a texto.

⁴ Para saber más consulte la memoria técnica

- Proceso de re-entrenamiento con 90 entrevistas en español, con transformación de audio a formato WAV a 1600 Hz y etiquetado manual de 3,380 segmentos de máximo 30 segundos.
- Evaluación del modelo convertido a CT2 a través de la librería WhisperX con 478 segmentos de audio de 30 segundos como datos de prueba. Se obtuvo un WER de 0.72 tomando en cuenta signos de puntuación y un WER de 0.52 sin tomar en cuenta signos de puntuación.
- Tiempo de inferencia de 2 minutos por cada 30 minutos de audio.

Obtención de Respuestas a partir de Texto:

- Modelos utilizados:

- Llama 2 es un modelo de lenguaje de código abierto desarrollado por Meta y Microsoft. Es un modelo de transformador de gran tamaño, con 2 billones de parámetros, que ha sido entrenado en un conjunto de datos masivo de texto y código, lo que le permite comprender y generar texto con gran precisión.

- El modelo GPT-3.5-Turbo es un modelo privado de lenguaje grande desarrollado por OpenAI. Es un modelo de transformador preentrenado que puede generar texto, traducir idiomas, escribir diferentes tipos de contenido creativo y responder a preguntas de manera informativa. El modelo GPT-3.5-Turbo se entrenó en un conjunto de datos masivo de texto y código. Esto le permite generar texto que sea más natural y coherente que los modelos de lenguaje más pequeños.

- El modelo multi-qa-MiniLM-L6-cos-v1 es un modelo de representación de texto desarrollado por Hugging Face. Está basado en el modelo MiniLM-L6, que es una versión reducida del modelo LLM Transformer-XL. El modelo multi-qa-MiniLM-L6-cos-v1 está entrenado en un conjunto de datos de preguntas y respuestas, y puede utilizarse para generar representaciones numéricas de textos y seleccionar el contexto necesario dentro de un documento más grande para responder una respuesta específica.

- El modelo text-embedding-ada-002 es un modelo de representación de texto desarrollado por OpenAI. No es de código abierto y solo puede utilizarse a través de la API de OpenAI. Este modelo está entrenado en un conjunto masivo de datos y puede utilizarse para generar representaciones numéricas de textos y seleccionar el contexto necesario dentro de un documento más grande para responder una respuesta específica.

- Utilización de dos conjuntos de modelos de lenguaje natural:

- Conjunto 1: Llama 2 (Meta) y multi-qa-MiniLM-L6-cos-v1 (HuggingFace).
- Conjunto 2: GPT-3.5-turbo y text-embedding-ada-002 (OpenAI).

- Tiempo de inferencia de 30 minutos por cada 50 preguntas por el conjunto 1.
- Tiempo de inferencia de 1 minuto por cada 50 preguntas por el conjunto 2.
- La evaluación del conjunto 1 obtuvo 0.69 de cosine similarity en 7 transcripciones de prueba.
- La evaluación del conjunto 2 obtuvo 0.75 de cosine similarity en 7 transcripciones de prueba.

Pipeline General de Procesamiento:

- Configuración de una instancia EC2 para ejecutar un servicio de procesamiento programado de 8:00 PM a 8:00 AM.
- Proceso de obtención de expedientes sin procesar, descarga de archivos de audio, transcripción a formato .vtt con el modelo de transcripción de audio a texto, generación de respuestas a partir de las transcripciones con los modelos de lenguaje natural y conversión de respuestas a formato JSON para la API de Olivia.
- Eliminación de archivos de transcripción para evitar duplicados en el procesamiento.
- Hardware utilizado: Instancia EC2 g5.xlarge con 4 vCPUS y 24 GB de memoria de video.

2.4 Métricas de Desempeño del modelo

Para dar seguimiento a los avances del modelo se utilizaron principalmente dos métricas de desempeño⁵:

- Word Error Rate (WER):

El WER es una métrica comúnmente utilizada en tareas de reconocimiento automático de voz (ASR) y en la evaluación de sistemas de traducción automática. Mide la diferencia entre una secuencia de palabras de referencia (la transcripción de referencia) y la secuencia de palabras generada por el modelo (la hipótesis). El WER se calcula contando el número de inserciones (palabras incorrectas o extra agregadas), eliminaciones (palabras que faltan en la hipótesis) y sustituciones (palabras mal traducidas o reemplazadas) necesarios para convertir la hipótesis en la referencia.

⁵ Por el carácter de las tareas (audio a texto y comparación de texto predicho vs texto correcto) estas dos métricas son las que se reportan.

- Cosine Similarity:

La similitud coseno es una métrica utilizada para medir la similitud entre dos vectores en un espacio vectorial. Se aplica comúnmente en tareas de procesamiento de texto, como la recuperación de información y la búsqueda de similitud entre dos documentos. La similitud coseno produce un valor entre -1 y 1, donde 1 indica una similitud perfecta, 0 indica ninguna similitud y -1 indica similitud inversa (opuesta). En el contexto de procesamiento de texto, la similitud coseno se utiliza para medir cuán similar es una respuesta predicha por un modelo a la respuesta correcta en función de sus representaciones vectoriales.

2.5 Plataforma

Se diseñó una plataforma web con una arquitectura basada en una API Gateway, donde las trabajadoras de SISEMH suben las grabaciones de las entrevistas. En esta plataforma, es donde trabajan los modelos y se almacenan los expedientes digitalizados, además en la plataforma se puede ver el despliegue total de la sábana.

Sólo quienes cuenten con un usuario y una contraseña pueden ingresar al dashboard principal.

Para saber más sobre el funcionamiento de la plataforma consulte la Memoria Técnica de la Plataforma Olivia

2.6 Situación de pilotaje

Ya desarrollados los últimos modelos y la plataforma, se realizaron las pruebas de pilotaje.

Las pruebas piloto de Olivia, con los modelos descritos se realizaron del 4 al 22 de septiembre (3 sprints de 1 semana cada uno). Para esto, se necesitó del apoyo del equipo de la SISEMH en la modificación de diferentes pasos de su proceso habitual de atención.

Así la incorporación de la tecnología quedó de la siguiente manera:

La PRV llegaba a solicitar el apoyo a la Unidad, la recepcionista abría expediente en la plataforma de Olivia para posteriormente hacer de conocimiento del equipo interdisciplinario que había una persona en búsqueda de asesoría. La PRV era recibida por una persona integrante del equipo de la SISEMH para el levantamiento de la EVI, las personas del equipo de la SISEMH solicitaban el permiso de la usuaria a ser grabada. Una vez otorgado el permiso, se comenzaba con la entrevista con el nuevo formato que fue entregado por la Secretaría.

Como se señaló durante la intervención total de OlivIA, la Secretaría solicitó realizar las entrevistas con un formato de entrevista nuevo, haciendo que las personas del equipo de apoyo interdisciplinario tuvieran un proceso de capacitación, aprendizaje y socialización al mismo tiempo entre el nuevo formato de EVI y la plataforma de OlivIA.

Una vez terminada la entrevista, la grabación se guardaba en la plataforma y esta era procesada en horarios no laborales. Esto con la finalidad de que todas las atenciones del día fueran procesadas sin interferir con el trabajo de atención y los datos fueran mostrados en la sábana informativa y clasificados en la misma durante el día siguiente en un horario laboral, lo que permitiría la revisión y confirmación de los datos.

El equipo de fAIR LAC fungió como equipo de soporte técnico frente a las fallas que pudiera presentar la plataforma o mejoras y nuevas funcionalidades que fueran necesarias para la comodidad de la atención.

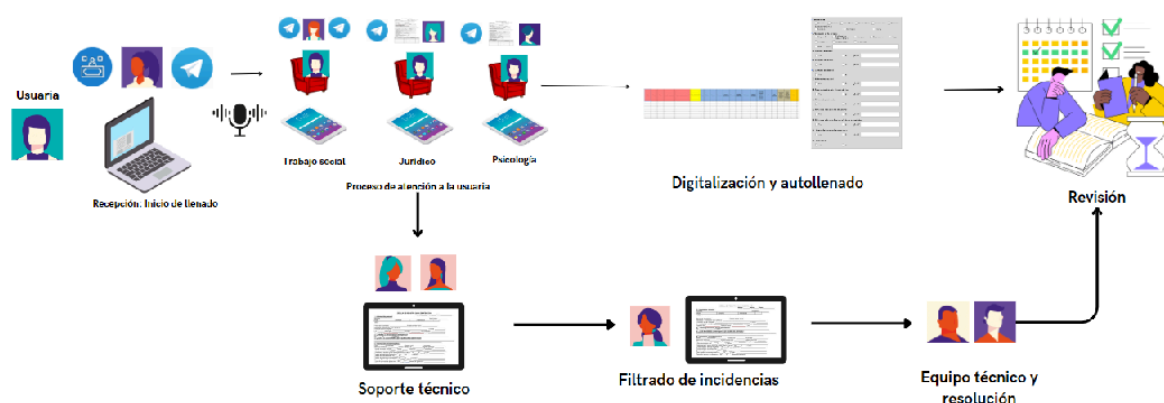


Figura n°2 proceso de atención durante el piloto
Fuente elaboración propia

3 ÉTICA Y GOBERNANZA DE DATOS

La implementación de IA debe buscar mecanismos que aseguren el uso ético y responsable de los datos masivos con los que se entrenan los modelos, ya que existen diferentes retos latentes como sesgos, ataques de ciberseguridad, mal uso de datos personales, entre otros. Para el caso de OlivIA for her, se firmó un convenio de colaboración con SISEMH en julio del 2022 donde se describe cómo será el traspaso de información, con qué finalidad y quiénes tendrán acceso a los datos, en este caso, a las grabaciones de las denuncias de

las mujeres víctimas de violencia. En dicho convenio, se establece la naturaleza jurídica del acceso a los datos por parte de fAIr LAC, que quedó como “encargada” y SISEMH como “dueña” de los datos. En la práctica, esto significa:

- fAIr LAC Jalisco podrá recoger datos con autorización de SISEMH y previo consentimiento firmado de las mujeres víctimas de violencia con el único propósito de entrenar el modelo de conversión de audio a texto.
- fAIr LAC Jalisco debe guardar la confidencialidad de las mujeres a quienes se les hizo la entrevista.
- fAIr LAC debe devolver los datos levantados a SISEMH una vez terminado el proyecto, ya que ésta última es dueña de la información.
- fAIr LAC debe abstenerse de transferir datos personales a terceros⁶⁷.

Para cumplir y dar cuenta de estas obligaciones, fAIr LAC Jalisco posee un Comité de Riesgos Éticos y Gobernanza de datos (CREG) quienes se encargan de documentar los requerimientos legales y de ciberseguridad, junto con dar seguimiento continuo para prevenir riesgos éticos en la implementación y diseño del caso de uso.

Para iniciar el proceso de seguimiento se partió por trabajar con la matriz ética, de la cual se desprenden una serie de recomendaciones. Junto con el desarrollo de la Congruencia normativa que corresponde a una síntesis de las leyes que se deben cumplir bajo la categoría de encargado y cómo se cumplen en el caso de uso.

Posteriormente, para cumplir con los requerimientos éticos del uso de la información se trabajó con la siguiente documentación durante la implementación:

-Acuerdo de confidencialidad para actores involucrados: donde se compromete privacidad y discreción con el uso de la información por parte del equipo implementador.

-Aviso de privacidad focalizado: descripción de los datos que se van a recabar y con qué propósito.

-Carta de aceptación/consentimiento: carta donde las mujeres participantes aceptan participar en el caso de uso y ser grabadas.

-Supresión de datos: explicitar por cuánto tiempo fAIr LAC Jalisco tendrá acceso a los datos personales y qué se hará con la información una vez terminado el proyecto.

⁶ Para saber más revisar convenio de colaboración entre ITESM y SISEMH

⁷ También revisar la descripción de la figura de encargado

De igual manera, para saber el nivel de responsabilidades sobre el manejo de los datos se redactó una lista de los participantes en el proyecto y su nivel de acceso a los datos donde se distinguen tres niveles (nivel alto, medio y bajo)⁸

Finalmente, para dar cuenta de los riesgos en temas de ciberseguridad se trabajó en dos documentos:

Documento de seguridad: descripción de las medidas de seguridad que se tuvieron al momento de manejar datos del caso de uso más las medidas de la plataforma y conservación de la información.

Vulneraciones de seguridad: protocolo en caso de filtración de información.

4 RESULTADOS OBTENIDOS EN LA IMPLEMENTACIÓN

A lo largo de la implementación del proyecto piloto se atendieron a 119 mujeres víctimas de violencia, cuyas atenciones ayudaron a generar y entrenar al modelo de OLIVIA; sin embargo con el uso completo del modelo se atendieron 24 mujeres, incluyendo la creación de su expediente y el llenado de los 45 campos de la sábana correspondiente; sin embargo 1 de los 24 casos, no pudo ser procesado por errores de grabación, lo que impedía que el audio fuera captado por el modelo y por ende procesado. Por otro lado, en el entrenamiento del modelo a pesar de la buena disposición del equipo de atención de SISEMH, se observó un proceso de frustración al momento de incluir las tecnologías. Debido a que se agregaron 6 minutos por grabación al promedio de tiempo que se tardan en atender a una persona, lo que implicó 2 horas 40 minutos de labor extra por parte del equipo interdisciplinario. Si bien esto era predecible, porque es parte del tiempo de adaptación y aprendizaje, se espera que este tiempo se acorte a medida que se adopte la tecnología. Considerando que a mayor cantidad de relatos, más información tendrá el modelo para su entrenamiento, logrando así procesar la información con menos errores.

⁸ Alto: información se clasifica como confidencial o secreta. Esta información sólo puede ser conocida por el grupo interno operativo del caso de uso, o por los demás miembros del equipo mediante autorización y con causa justificada (cumplimiento de responsabilidades, asesoría, análisis, entre otras)
Medio: La información se clasifica como reservada y para uso interno. Puede ser conocida y utilizada por algunos miembros de la organización y/o autoridades delegadas o personas externas con causa justificada.

Bajo: La información se clasifica como pública y de libre acceso. La información es comúnmente conocida y utilizada por cualquier miembro de la organización y puede ser compartida a terceros externos.

A pesar del tiempo adicional invertido por las trabajadoras de SISEMH en implementar la plataforma, los beneficios son considerables. Al final del periodo de implementación se ahorraron 11 horas con 40 minutos al eliminar el proceso de transcripción a mano.

Sobre el rendimiento del modelo se puede mencionar que hubo:

- 72% de precisión sin signos de puntuación de la traducción.
- 53% de precisión con signos de puntuación de la traducción. Considerando el correcto uso de comas, signos de interrogación, exclamación, separación de la traducción por párrafos, etc.
- El tiempo de procesamiento del modelo Llama 2 fue de 2:00 minutos por cada 30 segundos de audio (un audio promedio dura entre 30 a 45 minutos) por lo que cada audio tarda entre 120 a 180 minutos.

5 LECCIONES APRENDIDAS

Hubo una comprensión del entorno, en este proceso se entendió la importancia del acercamiento humano en la atención a casos de violencia. Se encontró un proceso empático y profesional en la atención, no obstante el equipo que atiende generalmente está saturado no solo con las atenciones, sino también con los procesos administrativos.

Es importante que durante el diseño de la solución tecnológica las expectativas de los usuarios sean comprendidas y se genere algo funcional para ellos. Porque algo que desde fuera podría considerarse una ventaja, para ellos podría suponer lo contrario (por ejemplo, darles algún aparato electrónico para la implementación, puede ser una responsabilidad adicional). Es importante asegurar de que el usuario sienta tranquilidad al momento de usar la herramienta y no tenga que preocuparse por si funcionará correctamente. En este sentido resalta la necesidad de perfeccionar la interfaz durante la grabación de la entrevista, mostrar un indicador visual de que se está grabando y permitir registrar más de un audio para que el entrevistador pueda pausar la grabación con tranquilidad en caso de necesitarlo.

Debido a la naturaleza del contacto directo con las personas víctimas de violencia, el usuario también está bajo una tensión emocional constante por lo que se entendió que la solución tendría que ser muy accesible y requerir poca intervención. Por otra parte se observó que hay un porcentaje de personas (6%) que no acceden a la grabación, derivado

del miedo a la filtración de la información de su relato o que hubiera alguna manera en la que la persona agresora tuviera acceso a su declaración. A partir de la retroalimentación por parte de las personas de primer contacto notamos que esto sucede sobre todo en casos de violencia sexual, relación de poder con su agresor (por ejemplo si es su jefe o profesor) o cuando su agresor tiene relación directa con el narcotráfico.

Como se comentó anteriormente, uno de los retos a los que se enfrentó el equipo técnico fue reconocer las dificultades de la transcripción en una conversación orgánica y bajo condiciones de estrés. En estos casos afecta tanto la modulación de la voz de las personas entrevistadas, como el orden de las preguntas establecidas en el guión de la entrevista (EVI). Del mismo modo, se comprendió que entre menos personas intervengan en el llenado de una sola cédula, se obtiene un mejor resultado. Estas consideraciones son claves para crear un modelo de transcripción eficiente para casos relacionados con violencia.

Se reconoce que la IA puede tener una mayor utilidad en la clasificación de información de la sábana informativa, que sólo ser utilizada para el llenado de la EVI, ya que la sábana agrupa y sistematiza la información de forma reducida para su análisis y comprensión estadística; en cambio EVI, es solo una herramienta física necesaria para la Trabajadora Social o persona de primer contacto, que tiene que llenarse de manera inmediata para conocer el tipo de seguimiento o acciones que serán necesarias en cada caso o atención; pero también le permite a las personas de primer contacto llevar un control y apoyo sobre la conversación de manera empática, dirigiendo de manera profesional la conversación.

Sobre los obstáculos encontrados sobre el proceso de entrevista, destacan los siguientes:

1. Los entrevistadores, por su expertis, agrupan preguntas provocando que información correspondiente a diferentes campos de la cédula se obtenga en una misma respuesta.

Ejemplo: “¿Consumes drogas o alcohol?”

Con la respuesta de dicha pregunta, se llenan otras respuestas.

2. La respuesta a una pregunta frecuentemente incluye información correspondiente a otro apartado.

Ejemplo:

Flujo de la conversación esperado: —¿Qué escolaridad tiene él? — Preparatoria.

—¿Terminada o trunca? —Trunca.

Flujo real: —¿Qué escolaridad tiene él? —Prepa trunca.

3. La información no siempre se obtiene de manera verbal ya que en ocasiones la persona entrevistada asiente, niega o muestra físicamente la respuesta, por ejemplo en la pantalla del celular. Esto causa que queden campos vacíos en la cédula.
4. Al formular ciertas preguntas, el uso de la palabra “sabes” dificulta discernir si la respuesta es “no” o “no sé”.
Ejemplo: —¿Sabes si tiene antecedentes penales? —No.
5. La persona entrevistada no siempre responde con información completa y el entrevistador tiene que indagar para obtener información clara. Por otra parte, al responder algunas de las preguntas más complejas frecuentemente se mencionan varias palabras que la plataforma puede interpretar como posibles respuestas.
Ejemplo: —¿Cuánto tiempo tienes de relación o de conocerlo a él? —Pues desde los quince años lo conozco —¿Son dieciocho? no, tienes treinta, quince.
6. Algunas preguntas son respondidas en dos secciones diferentes.
Ejemplo: El entrevistador pregunta si hay consumo de drogas, continúa con las preguntas de posesión de armas y enfermedad mental y posteriormente pregunta “¿qué drogas consume?”.
7. En la sección sobre las personas con las que comparte casa la información puede ser particularmente compleja. Las respuestas a esta pregunta suelen ser desordenadas y variar mucho dependiendo del caso.
8. Cuando preguntan el teléfono de la persona agresora, la persona entrevistada normalmente se toma un momento para buscarlo. En ocasiones los entrevistadores aprovechan esta pausa para avanzar con las siguientes preguntas provocando que esta respuesta quede fragmentada.
9. El último episodio de violencia pocas veces se pregunta directamente, la mayoría de las veces se deduce de la conversación.

10. Rara vez se diferencia si hay un consumo problemático de alcohol. En la mayoría de las entrevistas sólo se pregunta si la persona consume alcohol sin distinguir si es de manera social o constante, o si ocasiona cambios en la forma de actuar.
11. Al preguntar el tipo de apoyo que les brinda su red, a veces no se da un espacio para que la persona entrevistada responda por sí misma.
Ejemplo: —¿En qué sentido te apoya tu esposo, en todo sentido? —Sí.
12. La respuesta a la pregunta “¿con qué red de apoyo cuentas?” varía, pues en ocasiones contestan con el nombre y en otras con el parentesco.
13. Algunos de los retos no afectan directamente al etiquetado, sino a conseguir la información correcta, pues la persona víctima de violencia no siempre siente la confianza de hablar libremente de su situación.

6 MEJORAS FUTURAS Y RECOMENDACIONES

A partir de todo el análisis anterior, se definió que el flujo ideal del proceso con IA es el siguiente:

1. La PRV se acerca a la Unidad Metropolitana de Atención Integral a Mujeres y Niñez (UMEA) de la SISEMH a pedir atención y asesoría. En recepción inician la apertura de un expediente desde la plataforma de OlivIA e inicia el llenado de los datos básicos de la PRV. Después la deriva con una persona del equipo de Trabajo Social, Psicología y Abogacía según sus necesidades y la disponibilidad de la Secretaría.
2. La persona del equipo de SISEMH solicita el permiso de la usuaria para grabar la entrevista. En caso de que la usuaria acepte, la entrevista de adquisición de datos es grabada, para esto existe la opción de hacerlo directamente desde la plataforma, o desde una aplicación externa que permita subir los archivos al terminar la intervención.
3. Una vez terminada la entrevista el equipo de atención se asegura de que exista un archivo de audio ligado al expediente y continúa con las atenciones necesarias durante el día.

4. La entrevista automáticamente se transcribe a texto (por medio de la IA) y las respuestas se catalogan dentro de la sábana y la cédula. El tiempo de procesamiento dependerá de la capacidad del modelo y del acceso al internet.
5. Una vez procesados los audios, la clasificación y veracidad de los datos capturados por la IA deberán ser revisadas y confirmadas por el equipo de la Secretaría.



Figura n°3 proceso de atención ideal
Fuente elaboración propia

Se enfatiza la atención a los riesgos relativos al tratamiento de datos personales y su seguridad. Esto puede ir desde cuidar los dispositivos desde los que se accede hasta delimitar las personas que tienen acceso a la plataforma. Junto con considerar reforzar las medidas de ciberseguridad de la plataforma.

También la implementación de una nueva herramienta tecnológica implica una curva de aprendizaje para el equipo que la utiliza. Por ello, se recomienda destinar un periodo exclusivo para este proceso y evitar la sobreposición con otros procesos de aprendizaje, como en este caso fue el cambio a un nuevo formato de EVI. Combinar dos procesos de aprendizaje diferentes puede aumentar la presión a la que se enfrenta el equipo y reducir su eficiencia.

Al momento de implementar una nueva herramienta es fundamental prestar atención a la alfabetización digital de las servidoras públicas y designar un tiempo adecuado para la curva de aprendizaje. De otra manera la presión para aprender por su cuenta algo que se encuentra fuera de su alcance puede ser un factor para provocar frustración y desgaste e impedir una atención adecuada a las usuarias.

Es una buena práctica preparar un manual gráfico, paso a paso al que el equipo pueda recurrir. Esto ayuda a prevenir el miedo a las nuevas tecnologías durante el uso de la plataforma. Junto con considerar capacitación constante en materia digital.

Con miras a una próxima etapa, los integrantes mínimos para conformar un equipo serían:

- **Líder Machine Learning Developer:** Persona responsable de proveer asesoramiento al equipo técnico.
- **2 Machine Learning o Generative AI Developers de apoyo:** Enfocados en NLP (Natural Language Process), generarían un modelo de pregunta/respuesta. Serían los encargados de entrenar al modelo⁹.
- **1 FullStack Developer:** Desarrollador con manejo de backend (ExpressJS y NodeJS) y frontend (ReactJS, NextJS y Tailwind CSS). Respecto al versionamiento debe tener conocimiento de Git y GitHub Actions, así como MongoDB o noSQL para el trabajo con base de datos. Por otra parte, en cuanto a DevOps y Deploy deberá tener nociones de Docker, Docker Compose, AWS, EC2 y S3. Esta persona podría trabajar a medio tiempo o tiempo completo, dependiendo de los nuevos requerimientos del proyecto.
- **1 DevOps:** Perfil enfocado en ciberseguridad con conocimiento en de Docker, Docker Compose, AWS, EC2 y S3. Dependiendo de si se quiere mantener la infraestructura actual en AWS o se migrará a infraestructura de la Secretaria podría considerarse un perfil de DevOps jr o mid.
- **1 Project Manager:** Un gerente de proyectos que integre prácticas Agile y procesos Scrum al desarrollo del proyecto. Coordina la parte operativa del proyecto como gerente de producto y está en contacto constante con el cliente.
- **1 Project Management Office:** Un perfil responsable de coordinar la parte administrativa y logística del proyecto y desarrollar la documentación.
- **2 Apoyo operativo (medio tiempo):** Un perfil de UX Writer y un perfil de UX Researcher.

⁹ Revisar memoria técnica de los modelos

7 REFERENCIAS BIBLIOGRÁFICAS

INEGI.(2021). Encuesta nacional de victimización y percepción sobre seguridad pública (ENVIPE). Comunicado de prensa núm. 532/21. Disponible en:

<https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2021/EstSegPub/envipe2021.pdf>

Gobierno de Jalisco. (2020).Estrategia Emergente contra la Violencia en Razón de Género. Disponible en:

<https://igualdad.jalisco.gob.mx/wp-content/uploads/2022/03/Estrategia-Emergente-contra-la-Violencia-en-Razon-de-Genero.pdf>

Jalisco Cómo vamos.(2022). Ser mujer en Guadalajara 2022. Disponible en:

<https://igualdad.jalisco.gob.mx/wp-content/uploads/2022/03/Ser-Mujer-en-Guadalajara-2022-Jalisco-Como-Vamos.pdf>

SENSP.(2022). Información sobre violencia contra las mujeres. Incidencia delictiva y llamadas de emergencia 911. Centro Nacional de Información. Disponible en:

<https://drive.google.com/file/d/1xH44QYX7latXz57Zx3CDfsPSICOVfeaM/view>

8 ANEXOS

Anexo 1

Tabla 5-Preguntas analizadas por el modelo

Pregunta	Respuesta
II_nombre	Texto libre
II_genero_especificar	Texto libre
II_sexo	Lista de opciones
II_fecha_de_nacimiento	[día] de [mes] de [año]
II_nacionalidad	Texto libre
II_lugar_de_nacimiento	Texto libre
II_entidad_federativa_donde_reside_actualmente	Texto libre

II_telefono_fijo_casa	#####
II_direccion_calle	Texto libre
III_su_escolaridad_esta_en	Lista de opciones
III_cual_seguridad_social	Lista de opciones
III_ocupacion_de_la_persona	Lista de opciones
III_especificar_ocupacion_de_la_persona	Texto libre
III_situacion_conyugal	Lista de opciones
III_regimen_matrimonial	Lista de opciones
III_tipo_de_vivienda	Lista de opciones
III_especificar_tipo_de_vivienda	Texto libre
III_compartida_con_otras_personas	Lista de opciones
III_cuantas_personas_habitan_en_su_vivienda	Número
III_quien_aporta_el_mayor_ingreso_dentro_del_hogar	Texto libre
IV_contexto_causa_y_evolucion	Texto libre
IV_ha_tenido_que_ser_atendida_en_una_institucion_medica_o_por_personal_medico_como_consecuencia_de_un_evento_de_violencia_con_la_persona_agresora	Texto libre
IV_ultimo_episodio_de_violencia_especificar	Texto libre
VI_nombre	Texto libre
VI_edad	Número
VI_calle	Texto libre
VI_escolaridad	Lista de opciones
VI_telefono_fijo_casa	#####
VI_posesion_de_armas	Falso o Verdadero
VI_pertenece_o_tiene_enlace_con_el_crimen_organizado	Falso o Verdadero
VI_historial_de_antecedentes_penales	Texto libre
VI_especificar_senias	Texto libre