



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Estudios de Posgrado
Maestría de Tecnologías de la Información y Comunicación

**DESARROLLO DE PROTOTIPO DE APLICACIÓN MÓVIL CON RECONOCIMIENTO DE VOZ
PARA LA TRANSCRIPCIÓN DE DENUNCIAS DE VÍCTIMAS DE CRÍMENES SEXUALES Y
PROCESAMIENTO DE DATOS PARA ANÁLISIS DE PATRONES**

Lic. Cristián Vinicio Juárez Cabrera

Asesorado por el Ing. Ms. Juan Fernando Cano Villatoro

Guatemala, enero de 2022

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA



FACULTAD DE INGENIERÍA

**DESARROLLO DE PROTOTIPO DE APLICACIÓN MÓVIL CON RECONOCIMIENTO DE VOZ
PARA LA TRANSCRIPCIÓN DE DENUNCIAS DE VÍCTIMAS DE CRÍMENES SEXUALES Y
PROCESAMIENTO DE DATOS PARA ANÁLISIS DE PATRONES**

TRABAJO DE GRADUACIÓN

PRESENTADO A JUNTA DIRECTIVA DE LA
FACULTAD DE INGENIERÍA

POR

LIC. CHRISTIÁN VINICIO JUÁREZ CABRERA

ASESORADO POR EL ING. MS. JUAN FERNANDO CANO VILLATORO

AL CONFERÍRSELE EL TÍTULO DE

MAESTRO EN TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIÓN

GUATEMALA, ENERO DE 2022

UNIVERSIDAD DE SAN CARLOS DE GUATEMALA
FACULTAD DE INGENIERÍA



NÓMINA DE JUNTA DIRECTIVA

DECANA	Inga. Aurelia Anabela Cordova Estrada
VOCAL I	Ing. José Francisco Gómez Rivera
VOCAL II	Ing. Mario Renato Escobedo Martínez
VOCAL III	Ing. José Milton de León Bran
VOCAL IV	Br. Kevin Vladimir Armando Cruz Lorente
VOCAL V	Br. Fernando José Paz González
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

TRIBUNAL QUE PRACTICÓ EL EXAMEN GENERAL PRIVADO

DECANA	Inga. Aurelia Anabela Cordova Estrada
EXAMINADORA	Inga. Sasha Steffanie Palencia Zetina
DIRECTOR	Ing. Edgar Darío Álvarez Cotí
EXAMINADOR	Ing. Marlon Antonio Pérez Türk
SECRETARIO	Ing. Hugo Humberto Rivera Pérez

HONORABLE TRIBUNAL EXAMINADOR

En cumplimiento con los preceptos que establece la ley de la Universidad de San Carlos de Guatemala, presento a su consideración mi trabajo de graduación titulado:

DESARROLLO DE PROTOTIPO DE APLICACIÓN MÓVIL CON RECONOCIMIENTO DE VOZ PARA LA TRANSCRIPCIÓN DE DENUNCIAS DE VÍCTIMAS DE CRÍMENES SEXUALES Y PROCESAMIENTO DE DATOS PARA ANÁLISIS DE PATRONES

Tema que me fuera asignado por la Dirección de la Escuela de Estudios de Posgrado, con fecha enero de 2021.

Lic. Cristián Vinicio Juárez Cabrera

Facultad de Ingeniería

Decanato
24189101-
24189102
secretariadecanato@ingenieria.usac.edu.gt

LNG.DECANATO.OI.008.2022

La Decana de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer la aprobación por parte del Director de la Escuela de Estudios de Posgrado, al Trabajo de Graduación titulado: **DESARROLLO DE PROTOTIPO DE APLICACIÓN MÓVIL CON RECONOCIMIENTO DE VOZ PARA LA TRANSCRIPCIÓN DE DENUNCIAS DE VÍCTIMAS DE CRÍMENES SEXUALES Y PROCESAMIENTO DE DATOS PARA ANÁLISIS DE PATRONES**, presentado por: **Christián Vinicio Juárez Cabrera**, que pertenece al programa de Maestría en artes en Tecnologías de la información y la comunicación, después de haber culminado las revisiones previas bajo la responsabilidad de las instancias correspondientes, autoriza la impresión del mismo.

IMPRÍMASE:



Inga. Aurelia Anabela Cordova Estrada

Decana



Guatemala, enero de 2022

AACE/gaoc



Guatemala, enero de 2022

LNG.EEP.OI.008.2022

En mi calidad de Director de la Escuela de Estudios de Postgrado de la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, luego de conocer el dictamen del asesor, verificar la aprobación del Coordinador de Maestría y la aprobación del Área de Lingüística al trabajo de graduación titulado:

“DESARROLLO DE PROTOTIPO DE APLICACIÓN MÓVIL CON RECONOCIMIENTO DE VOZ PARA LA TRANSCRIPCIÓN DE DENUNCIAS DE VÍCTIMAS DE CRÍMENES SEXUALES Y PROCESAMIENTO DE DATOS PARA ANÁLISIS DE PATRONES”

presentado por **Christián Vinicio Juárez Cabrera** correspondiente al programa de **Maestría en artes en Tecnologías de la información y la comunicación** ; apruebo y autorizo el mismo.

Atentamente,

“Id y Enseñad a Todos”

Mtro. Ing. Edgar Darío Álvarez Cotí
Director



Escuela de Estudios de Postgrado
Facultad de Ingeniería



Guatemala, 10 de julio 2021

M.A. Edgar Darío Álvarez Cotí
Director
Escuela de Estudios de Postgrado
Presente

M.A. Ingeniero Álvarez Cotí:

Por este medio informo que he revisado y aprobado el **TRABAJO DE GRADUACIÓN** titulado: "DESARROLLO DE PROTOTIPO DE APLICACIÓN MÓVIL CON RECONOCIMIENTO DE VOZ PARA LA TRANSCRIPCIÓN DE DENUNCIAS DE VICTIMAS DE CRÍMENES SEXUALES Y PROCESAMIENTO DE DATOS PARA ANÁLISIS DE PATRONES" del estudiante **Christián Vinicio Juárez Cabrera** quien se identifica con número de carné **201990481** del programa de **Maestría en Tecnologías de la Información y la Comunicación**.

Con base en la evaluación realizada hago constar que he evaluado la calidad, validez, pertinencia y coherencia de los resultados obtenidos en el trabajo presentado y según lo establecido en el *Normativo de Tesis y Trabajos de Graduación aprobado por Junta Directiva de la Facultad de Ingeniería Punto Sexto inciso 6.10 del Acta 04-2014 de sesión celebrada el 04 de febrero de 2014*. Por lo cual el trabajo evaluado cuenta con mi aprobación.

Agradeciendo su atención y deseándole éxitos en sus actividades profesionales me suscribo.

Atentamente,

MARLON ANTONIO PEREZ TURK
INGENIERO EN CIENCIAS Y SISTEMAS
COLEGIADO No. 4492

MA. Ing. Marlon Antonio Pérez Türk
Coordinador

Maestría en Tecnologías de la Información y la Comunicación
Escuela de Estudios de Postgrado



USAC
TRICENTENARIA
Universidad de San Carlos de Guatemala



Universidad de San Carlos de Guatemala
Facultad de Ingeniería
Escuela de Estudios de Postgrado
Coordinador de Área

Guatemala, 21 de octubre de 2021.

M.A. Ing. Edgar Dario Álvarez Cotí
Director
Escuela de Estudios de Postgrados
Presente

Estimado M.A. Ing. Álvarez Cotí:

Reciba un cordial y atento saludo, a la vez aprovecho la oportunidad para hacerle de su conocimiento que he revisado y aprobado el **TRABAJO DE GRADUACIÓN** titulado: “DESARROLLO DE PROTOTIPO DE APLICACIÓN MÓVIL CON RECONOCIMIENTO DE VOZ PARA LA TRANSCRIPCIÓN DE DENUNCIAS DE VICTIMAS DE CRÍMENES SEXUALES Y PROCESAMIENTO DE DATOS PARA ANÁLISIS DE PATRONES” del estudiante CHRISTIÁN VINICIO JUÁREZ CABRERA del Programa de Maestría en Tecnología de la Información y Comunicación, identificado con número de carné: 201990481.

Agradeciendo su atención y deseándole éxitos en sus actividades profesionales me suscribo.

Atentamente,

“Id y enseñad a todos”

MSc. Ing. Juan Fernando Cano Villatoro
Colegiado No. 16391
Asesor

Cc: Archivo/LA

Doctorado: Sostenibilidad y Cambio Climático. **Programas de Maestrías:** Ingeniería Vial, Gestión Industrial, Estructuras, Energía y Ambiente Ingeniería Geotécnica, Ingeniería para el Desarrollo Municipal, Tecnologías de la Información y la Comunicación, Ingeniería de Mantenimiento. **Especializaciones:** Gestión del Talento Humano, Mercados Eléctricos, Investigación Científica, Educación virtual para el nivel superior, Administración y Mantenimiento Hospitalario, Neuropsicología y Neurociencia aplicada a la Industria, Enseñanza de la Matemática en el nivel superior, Estadística, Seguros y ciencias actuariales, Sistemas de información Geográfica, Sistemas de gestión de calidad, Explotación Minera, Catastro.

ACTO QUE DEDICO A:

- Dios** Por las bendiciones recibidas, por brindarme las oportunidades y situaciones adecuadas para poder salir adelante, por brindarme la salud mental para tomar decisiones que son de beneficio para mí y para mi familia.
- Mi madre** Por su dedicación y entrega al educarme, por poner los cimientos para lograr llegar a donde estoy, porque demostrarle a ella que sigo superándome y alcanzando mis metas es una de las mayores motivaciones para seguir adelante.
- Mi padre** Por inculcarme que los valores, la educación, el trabajo, la perseverancia y la dedicación son elementos indispensables para superarme en la vida. Por sus consejos personales y laborales que han servido para guiarme a través de las distintas etapas de mi vida.
- Mi esposa** Por tu amor, comprensión y apoyo emocional a lo largo de esta carrera, no lo hubiera logrado sin ti, por estar ahí siempre que te necesito en cualquiera que sea la situación.

Mis hijos

Kristen y Alejandro Juárez por su apoyo, comprensión y cariño, por ser el motor que me impulsa a querer ser siempre mejor, este logro es por y para ustedes.

Mi hermana

Silvia de León que considero como una segunda madre, por brindarme el apoyo para seleccionar esta carrera que ha traído muchas satisfacciones a mi vida.

Mis hermanos

Luz, Alonzo Juárez y Orlando de León por el apoyo y cariño que me han brindado, por aconsejarme, guiarme a través de la vida y por siempre motivarme a perseguir mis metas.

Mi familia

Gracias a todos por el apoyo que, de alguna manera, me brindaron, todos son parte fundamental de mi vida.

AGRADECIMIENTOS A:

Dios

Por estar conmigo en el camino durante toda la vida, por la fortaleza que ha dado, por las oportunidades que ha puesto en mi camino y la increíble familia que me ha permitido tener, porque la unión de todos estos elementos me ha traído hasta aquí.

Mi asesor

Por sus conocimientos, apoyo, consejos, tiempo y dedicación para lograr culminar mi trabajo de graduación de manera adecuada.

Mis catedráticos

Por transmitirnos sus conocimientos a lo largo de la maestría, sin ningún complejo comparten su experiencia y consejos.

Universidad de San Carlos de Guatemala

Por el apoyo y la formación académica en estos años de estudio, agradecimientos a la Facultad de Ingeniería y a la Escuela de Estudios de Posgrado.

ÍNDICE GENERAL

ÍNDICE DE ILUSTRACIONES	V
GLOSARIO	VII
RESUMEN.....	XIII
PLANTEAMIENTO DEL PROBLEMA Y FORMULACIÓN DE PREGUNTAS ORIENTADORAS	XV
OBJETIVOS.....	XIX
MARCO METODOLÓGICO	XXI
TÉCNICAS DE ANÁLISIS DE INFORMACIÓN	XXIX
INTRODUCCIÓN	XXXI
1. ANTECEDENTES	1
2. JUSTIFICACIÓN	5
3. ALCANCES	7
3.1. Resultados.....	7
3.2. Técnicos	7
3.3. Investigativos.....	9
4. MARCO TEÓRICO.....	11
4.1. Reconocimiento de voz	11
4.1.1. Algoritmos de reconocimiento de voz	12
4.1.2. Motores de reconocimiento de voz	12
4.2. Google cloud	13
4.2.1. Google Cloud Speech-to-Text	13

4.3.	Lenguajes de programación.....	14
4.3.1.	Imperativos.....	14
4.3.2.	Declarativos.....	14
4.3.3.	Orientado a objetos	15
4.4.	Base de datos	16
4.4.1.	Gestor de Base de Datos (SGBD).....	16
4.4.2.	Modelo relacional	16
4.4.3.	SQL Server.....	17
4.4.4.	SQLITE.....	17
4.5.	Interfaz de programación de aplicaciones (API).....	18
4.6.	Arquitectura de sistema.....	18
4.6.1.	Patrones de arquitectura	18
4.6.2.	Arquitectura en capas	19
4.6.3.	Arquitectura de repositorio	19
4.6.4.	Arquitectura cliente servidor	19
4.7.	Desarrollo de aplicaciones móviles	20
4.7.1.	Aplicaciones WEB.....	21
4.7.2.	Aplicaciones nativas	21
4.8.	Dispositivos móviles.....	21
4.8.1.	Movilidad	21
4.8.2.	Tamaño reducido	22
4.8.3.	Comunicación inalámbrica	22
4.8.4.	Interacción con las personas.....	22
4.9.	Inteligencia de negocios.....	22
4.9.1.	Data Warehouse	23
4.9.2.	OLAP (On-Line Analytical Processing).....	23
4.9.3.	Sistema ROLAP	23

5.	PRESENTACIÓN DE RESULTADOS	25
5.1.	Análisis de arquitectura del prototipo para transcripción de voz.....	25
5.1.1.	Transcripción de voz a texto	25
5.1.2.	Inteligencia de negocios	26
5.2.	Eficiencia de transcripción del prototipo	27
5.2.1.	Tiempos de transcripción.....	30
5.3.	Estructura de información de datos del prototipo para transcripción de voz para denuncias de crímenes sexuales....	30
5.4.	Implementación y funcionamiento del prototipo de transcripción de voz para denuncias de crímenes sexuales y reconocimiento de patrones.....	31
5.5.	Esquema de arquitectura Data Warehouse del prototipo de transcripción de voz para denuncias de crímenes sexuales....	34
5.6.	Inteligencia de negocios para análisis y búsqueda de patrones.....	35
5.6.1.	Denuncias por ubicación	37
5.6.2.	Denuncias por edad.....	38
5.6.3.	Denuncias por género	39
5.6.4.	Denuncias por departamento y municipio.....	39
5.6.5.	Denuncias por departamento.....	40
6.	DISCUSIÓN DE RESULTADOS	43
6.1	Arquitectura del prototipo	43
6.1.1.	Síncrono y asíncrono.....	43
6.1.2.	Interfaz de usuario	44
6.2.	Eficiencia de transcripción	44
6.2.1.	Tiempo de transcripción	44
6.2.2.	Errores en transcripción.....	45

6.3.	Ventajas de la inteligencia de negocios	45
6.4.	Trabajos futuros para la solución	46
CONCLUSIONES.....		47
RECOMENDACIONES		49
REFERENCIAS.....		51

ÍNDICE DE ILUSTRACIONES

FIGURAS

1.	Diagrama de funcionalidad del prototipo	27
2.	Diagrama entidad relación base de datos	31
3.	Flujo de trabajo del prototipo.....	33
4.	Arquitectura data warehouse del prototipo.....	34
5.	Tablero de información de denuncias.	36
6.	Gráfica de mapa, denuncias por departamento	37
7.	Gráfica de embudo, denuncias por edad	38
8.	Gráfica de denuncias por género	39
9.	Gráfica de área, denuncias por municipio	40
10.	Gráfica de barras, denuncias por departamento	41

TABLAS

I.	Variables	XXII
II.	Resultado de pruebas del prototipo de 30 segundos	28
III.	Resultado pruebas del prototipo de 60 segundos	28
IV.	Resultados de pruebas del prototipo de 90 segundos	29

GLOSARIO

Algoritmo	Conjunto ordenado de operaciones sistemáticas que permite hacer un cálculo y hallar la solución de un tipo de problemas.
Android	Sistema operativo móvil basado en núcleo Linux y otros softwares de código abierto.
Android studio	Entorno de desarrollo integrado oficial para la plataforma Android, Está basado en el software IntelliJ IDEA de JetBrains y ha sido publicado de forma gratuita a través de la Licencia Apache 2.0.
API	Application Programming Interfaces
Aprendizaje profundo	Conjunto de algoritmos de aprendizaje automático que intenta modelar abstracciones de alto nivel en datos usando arquitecturas computacionales.
Arquitectura de software	Estructura una solución de software compuesta por los elementos y las interrelaciones que la integran y que permiten su funcionamiento.
Base de datos	Conjunto de información relacionada que se encuentra agrupada o estructurada y permite

agregar, modificar, eliminar o consultar su contenido.

CRUD

Acrónimo en inglés para las acciones crear, leer, actualizar y eliminar en una base de datos;
Create, Read, Update and Delete

Data Warehouse

Repositorio unificado para todos los datos que recogen los diversos sistemas de una empresa. El repositorio puede ser físico o lógico y hace hincapié en la captura de datos de diversas fuentes sobre todo para fines analíticos y de acceso.

Denuncias

Una denuncia es una declaración formal acerca de una conducta contraria a Derecho, generalmente sobre algún delito, dirigida a la autoridad competente para su investigación.

Disponibilidad

Período de tiempo en que un servicio está disponible, así como el tiempo requerido por un sistema para responder a una solicitud realizada por un usuario.

Eficiencia

Utilización correcta y con la menor cantidad de recursos para conseguir un objetivo o cuando se alcanza más objetivos con los mismos o menos recursos.

Esquema	Esquema es una representación gráfica de la asociación de ideas o conceptos que se relacionan entre sí, y entre los que se establecen relaciones de jerarquía.
Estrategia	Es un plan general para lograr uno o más objetivos generales a corto o a largo plazo en condiciones de incertidumbre.
IDE	<i>Integrated Development Environment</i>
Integridad	Término utilizado para referirse a la exactitud y fiabilidad de los datos.
Inteligencia de negocios	Es el conjunto de procesos, aplicaciones y tecnologías que facilitan la obtención rápida y sencilla de datos provenientes de los sistemas de gestión empresarial para su análisis e interpretación, de manera que puedan ser aprovechados para la toma de decisiones y se conviertan en conocimiento.
Interfaz	Se utiliza en informática para nombrar a la conexión funcional entre dos sistemas, programas, dispositivos o componentes de cualquier tipo, que proporciona una comunicación de distintos niveles, permitiendo el intercambio de información.

Ministerio Público	El Ministerio Público (MP) es una institución auxiliar de la administración pública y de los tribunales con funciones autónomas, cuyos fines principales son velar por el estricto cumplimiento de las leyes del país.
MP	Ministerio Público.
Patrones	Serie de variables constantes, identificables dentro de un conjunto mayor de datos.
PNC	Policía Nacional Civil.
Prototipo	Primera versión que se distribuye para dar a conocer un producto o servicio que reúne gran cantidad de características que poseerá la versión final.
Reconocimiento de voz	El reconocimiento automático del habla (RAH) o reconocimiento automático de voz es una disciplina de la inteligencia artificial que tiene como objetivo permitir la comunicación hablada entre seres humanos y computadoras.
Red neuronal artificial	Modelo computacional que fue evolucionando a partir de diversas aportaciones científicas. Consiste en un conjunto de unidades, llamadas neuronas artificiales, conectadas entre sí para transmitirse señales. La información de entrada

atraviesa la red neuronal en la cual es sometida a diversas operaciones produciendo unos valores de salida.

Seguridad

Se refiere a las medidas de protección de la privacidad digital que se aplican para evitar el acceso no autorizado a los datos en una base de datos.

SQL Server

Es un sistema de gestión de base de datos relacional, desarrollado por la empresa Microsoft. El lenguaje de desarrollo utilizado es Transact-SQL, una implementación del estándar ANSI del lenguaje SQL, utilizado para manipular y recuperar datos, crear tablas y definir relaciones entre ellas.

SQLite

Es una herramienta de software libre, que permite almacenar información en dispositivos empotrados de una forma sencilla, eficaz, potente, rápida y en equipos con pocas capacidades de hardware, como puede ser una PDA o un teléfono celular.

SVET

Secretaría contra la violencia sexual explotación y trata de personas

Tablero

Es una representación gráfica de los principales indicadores que intervienen en la consecución de

los objetivos de negocio, y que está orientada a la toma de decisiones para optimizar la estrategia de la empresa, un tablero debe transformar los datos en información y está en conocimiento para el negocio.

Tendencias

Propensión o inclinación en las personas y en las cosas hacia determinados fines o acciones.

Transcribir

Es la representación sistemática de una forma oral mediante signos escritos.

RESUMEN

Los datos estadísticos de la Secretaría contra la violencia sexual explotación y trata de personas indican que el Ministerio Público recibe, anualmente, un promedio de diez mil doscientas dieciséis denuncias por delitos de violencia sexual. Sin embargo, los organismos coinciden en que existe un subregistro de casos de violencia sexual que no son denunciados y podría representar más del setenta por ciento. (Secretaría Contra la Violencia Sexual, 2020)

El prototipo presentado en este trabajo de graduación tiene como finalidad reducir los errores humanos al transcribir las declaraciones de las víctimas, reducir el tiempo de una transcripción y mejorar la privacidad durante el proceso de toma de declaraciones en una institución encargada de velar por el cumplimiento de las leyes en contra de los crímenes de violencia sexual. De esta manera se reducirá el porcentaje de personas que no denuncian los crímenes. Mediante el uso de diversas herramientas de software utilizadas en el mundo por su seguridad, estabilidad y escalabilidad se desarrolló un prototipo de aplicación móvil que reconoce la voz y la transcribe a texto. Esto ayuda a que el proceso sea eficiente y privado para la víctima. Las pruebas en el prototipo arrojaron como resultado un 94.41 % de eficiencia en la transcripción, además, se comprobó que una transcripción manual conlleva 78.62 % más de tiempo comparado con la realizada por el prototipo en tiempo real.

Desde hace varios años la tecnología ha influido en la optimización de los procesos realizados por el ser humano. Esta aplicación móvil no es la excepción, se desarrolló con el entorno de desarrollo integrado Android Studio un conjunto

de herramientas que facilitó la creación de la aplicación, el lenguaje de programación Java, el API Speech-to-Text de Google que utiliza algoritmos de redes neuronales y aprendizaje profundo para el reconocimiento y transcripción de voz a texto y el gestor de base de datos SQLite para el móvil y SQL Server 2016, como gestor de la base de datos central a donde se sincroniza toda la información para su análisis después de la toma de declaración.

Los datos, sin el procesamiento debido sirven de poco o nada para las instituciones. Por esta razón, se incluyó como parte de la solución una sección de inteligencia de negocios que brinda a los usuarios una herramienta para el análisis de información de las denuncias de crímenes sexuales. Mediante el uso de un data warehouse, alimentado con la información de las declaraciones de las víctimas de crímenes sexuales y la generación de mapas, gráficas o tablas creados con la herramienta Power BI de Microsoft, el usuario puede analizar la información de forma gráfica lo cual facilita su interpretación. La inteligencia de negocio permite la detección de patrones de comportamiento de los agresores sexuales, como el departamento o municipio con mayores denuncias, la edad de las víctimas más propensas o el género más vulnerable ante estos ataques, lo que permitirá a las instituciones crear estrategias para reducir el índice de violencia sexual en el país, y crear campañas para promover la denuncia enfocadas en un departamento, municipio o rango de edad específico.

PLANTEAMIENTO DEL PROBLEMA Y FORMULACIÓN DE PREGUNTAS ORIENTADORAS

Existen diversas instituciones que velan por el cumplimiento de las leyes descritas en la Constitución política de la república de Guatemala. Una de ellas es el Ministerio Público (MP). Esta institución con funciones autónomas se encarga de la persecución penal y dirige la investigación de los delitos de acción pública. La Policía Nacional Civil (PNC) es la fuerza armada de Guatemala que mantiene el orden público, así como la seguridad de los ciudadanos del país. La Secretaría contra la Violencia Sexual, Explotación y Trata de Personas (SVET) es un órgano asesor que lucha contra la violencia sexual y distintas problemáticas.

Estas instituciones reciben denuncias de los ciudadanos sobre distintos delitos, sin embargo, no cuentan con un sistema de información que transcriba de forma eficiente las denuncias verbales del ciudadano. Una persona toma las declaraciones de la víctima y las redacta. Este es un momento incómodo para la víctima y el resultado puede ser una declaración incompleta, para evitar proporcionar detalles sobre delitos, como agresiones sexuales, explotación o trata de personas. Además, la persona que toma la declaración puede plasmar registrar una información con errores si no incluye la información completa o registra información errónea al interpretar de forma errónea la declaración. La pérdida de documentos físicos también puede generar un problema que incida en el proceso, ya que, la carencia de información completa e íntegra influye en el curso y las acciones que se tomen en el caso.

La digitación de la declaración en el sistema de denuncias tiene implícito el riesgo de que el usuario cometa errores que aumentan la posibilidad de distorsionar las declaraciones de las víctimas y ponen en riesgo la seguridad de los ciudadanos.

El incremento de denuncias ciudadanas causa que la toma de declaraciones sea lento e ineficiente, además, de que influye la celeridad con la que se toman acciones contra los delincuentes.

La seguridad es un factor determinante para la solución y, dada la importancia y valía de esta información, se deben aplicar controles de seguridad para evitar que personal no autorizado pueda manipularla y distorsionarla.

El Ministerio Público indica que durante el año 2018 se registraron 272 denuncias por trata de personas. Guatemala y Escuintla registraron la mayor cantidad de denuncias, con un 42.65 % y 9.19 % respectivamente. Jutiapa y Baja Verapaz registraron la menor cantidad de denuncias con 0.37 % y 0.74 % respectivamente. Por violencia sexual se registraron 10710 denuncias. En Guatemala y Huehuetenango se registra la mayor cantidad de denuncias con un 23.74 % y 7.55 %.

Se debe dar prioridad a la implementación de una solución para la toma de declaraciones en denuncias por delitos como violaciones o agresiones, para que las instituciones responsables aborden los casos con privacidad. De esta forma el denunciante se sentirá seguro y cómodo al denunciar. Permitirá describir, de mejor forma, el delito cometido en su contra y más personas tendrían la confianza para denunciar abusos.

La necesidad identificada genera la siguiente pregunta central y las auxiliares:

- Pregunta central

¿Cómo se automatiza la transcripción e identificación de patrones en las denuncias verbales de víctimas de delitos sexuales para evitar la digitación manual, mejorar la seguridad y el análisis de la información?

- Preguntas auxiliares

¿Cómo almacenar, de forma segura, la información obtenida de las transcripciones de las denuncias por delitos sexuales?

¿Cómo recopilar y almacenar la información de denuncias sexuales para realizar análisis de patrones?

¿Cómo determinar las tendencias de abuso sexual, según las distintas variables involucradas en los distintos casos, por ejemplo: edad, género o ubicación?

¿Qué componentes de arquitectura se necesitan para implementar la solución móvil de procesamiento de voz, transcripción de denuncias y análisis de patrones?

OBJETIVOS

General

Desarrollar el prototipo de un sistema de información para el reconocimiento de voz, que permita transcribir la denuncia de víctimas de delitos de violencia sexual, por medio de procesamiento de voz e identificar patrones de los delitos mediante el procesamiento y análisis de la información.

Específicos

1. Implementar una base de datos relacional, que permita almacenar las denuncias de las víctimas de crímenes sexuales para garantizar seguridad, integridad y disponibilidad de la información.
2. Crear una arquitectura de almacenamiento de datos (Data Warehouse) que permita organizar, comprender y utilizar los datos de las transcripciones de voz para analizar las tendencias de los crímenes.
3. Implementar un modelo de inteligencia de negocios para el análisis de los resultados de la transcripción de voz y determinar la relación de la violencia sexual y las variables de género, edad o ubicación.
4. Desarrollar la arquitectura de sistemas que permita el procesamiento de voz, transcripción de denuncias y análisis de la información obtenida.

MARCO METODOLÓGICO

- Tipo de investigación

Esta investigación se clasifica como cuantitativa. Su objetivo es el desarrollo de un prototipo mediante investigación de tecnología y herramientas existentes en la industria que permitan la captura y transcripción de voz a texto y el análisis de información de víctimas de crímenes sexuales para reducir errores en la transcripción de las denuncias, minimizar la presión que se ejerce sobre las víctimas durante la denuncia y ampliar las descripciones de las denuncias de los crímenes cometidos en su contra.

- Diseño

El diseño de la investigación es no experimental de tipo longitudinal. El prototipo se desarrolló por medio del estudio de la documentación de los distintos motores de reconocimiento de voz y herramientas que capturen y transcriban la voz, además de que almacenen los resultados para obtener declaraciones detalladas y precisas de crímenes de violencia sexual. Luego, la información se procesará y analizará con herramientas y tecnología de inteligencia de negocios que facilite el hallazgo de patrones y tendencias dentro de las transcripciones.

- Alcance

El alcance de la investigación es descriptivo. Al finalizar la investigación se determinarán las herramientas tecnológicas y procedimientos utilizados para el desarrollo del prototipo, así como su arquitectura, proceso de instalación, pruebas y resultados de la exactitud de la transcripción de las pruebas realizadas al prototipo con el objetivo de definir la viabilidad de su implementación en una institución real para transcribir las denuncias de crímenes sexuales.

- Variables, subvariables e indicadores

En la tabla I se presentan las variables, subvariables e indicadores considerados para el desarrollo de esta investigación.

Tabla I. **Variables**

Variable	Definición	Subvariable	Indicador
Género	Sexualidad de las víctimas		Cantidad de víctimas masculinas o femeninas que reportan delitos sexuales
Edad	Edad de las víctimas		Rango de edad de víctimas que reportan crímenes sexuales
Ubicación	Lugar o sitio donde fueron cometidos los crímenes		Lugares que más incidencia tienen en relación con delito sexual.
Sistemas expertos en transcripción de voz a texto	Motor de reconocimiento y transcripción de voz a texto.	Eficiencia	Porcentaje de efectividad del procedimiento de transcripción de voz a texto
		Efectividad de la transcripción	Porcentaje de palabras transcritas correctamente en una declaración.

Fuente: elaboración propia.

- Técnicas de recolección de información

Para el desarrollo de esta investigación y determinar el correcto funcionamiento de prototipo fue necesario usar distintas fuentes de información primarias y secundarias.

- Fuentes primarias

- Encuesta al usuario sobre la funcionalidad, interfaz y facilidad de uso del prototipo.
- Análisis de la información de las pruebas realizadas por el prototipo para determinar la eficiencia de la transcripción de voz a texto.
- Observación para recopilar información importante sobre el uso del prototipo.
- Información registrada por el prototipo del tiempo que toma realizar las transcripciones.

- Fuentes secundarias

- Tesis de maestrías de distintas universidades.
- Artículos y ensayos científicos relacionados con la tecnología utilizada para crear el prototipo.
- Tutoriales y videos que aporten al desarrollo.
- Artículos de revistas con énfasis en algoritmos para la transcripción de voz.
- Libros relacionados con la tecnología utilizada para crear el prototipo.

- Fases del estudio

En este apartado se detallan las fases involucradas en el desarrollo de la investigación desde la documentación utilizada para el desarrollo de los conceptos abordados, hasta la creación del prototipo que culmina con pruebas para determinar su precisión.

- Revisión documental

Durante la revisión documental se consultaron diversas fuentes, como libros, tesis, revistas y artículos científicos relacionados con los elementos necesarios para obtener, procesar y convertir las ondas sonoras de voz a texto para crear un prototipo por medio del cual se obtendrán las declaraciones de víctimas de delitos de crímenes sexuales, transcribirlas y almacenarlas de forma rápida y segura. También se consultaron fuentes de información relacionadas con el desarrollo de una inteligencia de negocios que permite analizar la información de las víctimas para determinar patrones sobre los crímenes de violencia sexual. Por esta razón fue necesario contar con una base conceptual detallada sobre los siguientes temas que se enlistan.

- Algoritmos para el reconocimiento de voz
- Modelos para el reconocimiento de voz
- Sistemas expertos para el reconocimiento de voz
- Inteligencia de negocios
 - Data Warehouse

Para el desarrollo de la solución se implementó una arquitectura móvil para almacenar la información transcrita y la integración de la API de Google para

envío y transcripción del audio de voz a texto por lo cual se describen los siguientes conceptos.

- Patrones arquitectónicos
- Bases de Datos
- Interfaz de programación de aplicaciones (API)

Duración estimada para esta fase: 8 semanas.

- Análisis y diseño

Una vez definidos los conceptos se seleccionó el lenguaje, herramientas, algoritmos, motor de reconocimiento de voz y modelos que se incluyeron en el prototipo y las herramientas para crear la solución de inteligencia de negocios para el análisis de la información.

Definición de los módulos que se crearon para el funcionamiento del prototipo entre los que se incluyen:

- Módulo de seguridad
- Módulo de captura y transcripción de denuncias
- Módulo para análisis de información

Definidos los elementos que se utilizarían, se creó un diagrama de componentes que plasma, de forma gráfica, la división de los elementos que integraron el prototipo y sus relaciones.

Describir el diseño de prototipo mediante el uso de una arquitectura móvil y cliente-servidor el cual fue representado por medio de un documento con la siguiente información:

- Diagrama de la arquitectura de la solución
- Descripción de componentes
- Diagrama entidad y relación
- Herramientas por utilizar en el desarrollo de la solución
- Requerimientos funcionales y no funcionales

Duración estimada para esta fase: 14 semanas

- Configuración del ambiente de desarrollo
 - Instalación y configuración del IDE Android Studio seleccionado para el desarrollo de la solución.
 - Instalación y configuración de las herramientas seleccionadas para la solución de inteligencia de negocios Microsoft BI.
 - Instalación y configuración del servidor y cliente de la base de datos SQL Server.
 - Instalación de herramientas y software relacionado con el desarrollo de la solución.
 - Configuración de un ambiente de pruebas máquina virtual.

- Desarrollo de prototipo

En esta fase se desarrolló la aplicación prototipo con base en las herramientas seleccionadas en la fase de diseño.

- Desarrollo de bases de datos
 - Definir las tablas a utilizar
 - Definir campos por tabla
 - Normalizar la base de datos
 - Creación de la base de datos SQL
 - Creación del diagrama de entidad relación

- Desarrollo de la aplicación prototipo
 - Desarrollo de funcionalidades CRUD para administrar las distintas tablas de la base de datos.
 - Desarrollo de aplicación móvil.
 - Codificación de las funcionalidades del software.
 - Integración del motor de transcripción de voz a la aplicación.
 - Desarrollo de los tableros para el despliegue de información.
 - Desarrollo de gráficas para consulta de información.

Duración estimada para esta fase: 20 semanas

- Pruebas

Es esta fase se realizaron pruebas con distintos usuarios para determinar la exactitud de las transcripciones. Para realizar las pruebas fue necesario cumplir con los siguiente.

- Configurar el prototipo en un cliente real para realizar pruebas.

- El usuario tenía que llenar el formulario de información personal de la aplicación como primer paso.
- El usuario debía pronunciar las vocales lo más claro posible.
- El usuario conto con tres lecturas de aproximadamente 30, 60 y 90 segundos respectivamente las cuales leyó a una distancia prudente del micrófono lo más claro posible.
- El usuario almaceno su información y los resultados de su transcripción para su posterior análisis.
- Para determinar la exactitud de la solución fue necesario acompañar al usuario en las pruebas y llevar un registro de los problemas y mejoras que tuvo el prototipo.
- Utilizar inteligencia de negocios para analizar la información almacenada durante las pruebas.

Duración estimada para esta fase: 6 semanas

- Análisis de las pruebas

Fue necesario recolectar los registros almacenados por los usuarios y calcular el promedio de las palabras que fueron transcritas correctamente, luego analizar los factores que pudieron contribuir a un mal funcionamiento de la solución desarrollada mediante técnicas de análisis de información: observación, investigación y estadística descriptiva.

Se utilizaron formularios de Google Docs para la encuesta a los usuarios sobre el funcionamiento del prototipo y así obtener un análisis estadístico de los resultados de la encuesta.

Duración estimada para esta fase: 2 semanas

TÉCNICAS DE ANÁLISIS DE INFORMACIÓN

Con el objetivo de analizar los resultados del prototipo que abarca esta investigación, fue necesario la aplicación de técnicas de estadística descriptiva para calcular, describir y presentar la información obtenida de las pruebas realizadas.

- Técnica descriptiva

Se recolectaron, ordenaron y analizaron los resultados obtenidos de las pruebas en el prototipo, para determinar el promedio de eficiencia por palabra en las transcripciones realizadas a los usuarios. Estas muestras se tomaron a personas de distinto género, edad, escolaridad y raza para obtener resultados con tonos de voz distintos. Con esta información se realizó un análisis de correlaciones para determinar si existía una relación entre las variables. Si hay una relación entre las variables y la eficiencia del prototipo se categorizará a los usuarios por género, edad, escolaridad y raza para determinar qué características de los sujetos de pruebas tiene mejor eficiencia al transcribir su voz.

Las variables analizadas según los datos obtenidos de las transcripciones son:

- Efectividad de la transcripción, se analizó la cantidad de palabras transcritas correctamente por usuario para determinar la efectividad del software.

- Disponibilidad del sistema, se evaluó el tiempo de disponibilidad continua.
 - Tiempo de transcripción, se realizó un análisis para determinar la diferencia de tiempo que hay entre una transcripción manual y una realizada mediante el prototipo.
- Técnica de visualización de datos

Para evaluar la experiencia del usuario al utilizar el prototipo se realizó una encuesta electrónica mediante formularios de Google para determinar la percepción de la funcionalidad y comodidad al interactuar con el prototipo. con base en esta información se generaron gráficas para determinar qué características califica mejor y peor el usuario. Para generar esta encuesta se consideraron las variables facilidad de uso, interfaz gráfica, eficiencia en la transcripción y sugerencias. Se procesó la información en una hoja electrónica de cálculo; se generaron gráficas con base en los resultados, se creó un resumen del resultado por variable y, mediante la técnica de visualización de datos, se establecieron las características del prototipo por mejorar.

INTRODUCCIÓN

Los crímenes por violencia sexual afectan a la población guatemalteca. Los registros del Ministerio Público (MP) indican que en el año 2018 se recibieron 10,710 denuncias por delitos de violencia sexual y, al 31 de Julio de 2019, se habían recibido 6,293 denuncias. El ministerio público y la Secretaría contra la violencia sexual, explotación y trata de personas coinciden en que hay un subregistro que asciende a un setenta por ciento de delitos que las víctimas no denuncian.

El objetivo principal de este trabajo de investigación reducir la incomodidad que experimentan las víctimas de crímenes sexuales cuando declaran y detallan las agresiones cometidas en su contra frente a personas desconocidas. Para ello se desarrollará un prototipo de aplicación que permita transcribir a texto las denuncias verbales de las víctimas de crímenes sexuales, mediante el uso de sistemas expertos y procesamiento de voz. Además de que se logrará la privacidad mientras las víctimas declaran, se obtendrán diversos beneficios con el desarrollo de esta tecnología, como los siguientes.

- Rapidez. El proceso en la toma y transcripción de las declaraciones será mucho más rápido al automatizarse con una aplicación de reconocimiento de voz.
- Seguridad. La información se almacenará directamente en una base de datos con altas medidas de seguridad para autenticación y consulta de información.

- Exactitud. Dado que no intervendrá un empleado en la toma de las declaraciones se elimina el factor de error humano haciendo el proceso de transcripción más eficiente y transparente.
- Información detallada y completa. Al prescindir del factor humano en las declaraciones de las víctimas, se brindará comodidad y confianza para que sus declaraciones sean completas y detalladas.
- Facilidad para determinar las patrones y tendencias en los crímenes de violencia sexual mediante el análisis de la información con herramientas de inteligencia de negocios.

El reconocimiento de voz no es una tecnología nueva, data de 1940. A partir de entonces, ha evolucionado para lograr mayor eficiencia, hasta el punto de transcribir la voz en tiempo real. Esta tecnología es aplicable en distintos campos de la ciencia y sociedad para agilizar los procesos y tareas.

La aplicación tendrá como principal objetivo facilitar la transcripción de voz a texto mediante un motor de reconocimiento de voz y almacenar la información en una base de datos relacional que permitirá una disponibilidad de información inmediata. Estos datos, posteriormente, se transformarán en información que permita definir patrones dentro de los crímenes de violencia sexual, mediante el uso de herramientas de inteligencia de negocios recopilando, depurando y transformando los datos con el objetivo de encontrar la relación de los crímenes con las variables de género, edad o ubicación.

Este documento de investigación se divide en varios capítulos en los que se detalla el proceso para la creación del prototipo, para conocer la historia del reconocimiento de voz el capítulo uno presenta los antecedentes y casos donde ha sido aplicada la tecnología de transcripción de voz a texto.

El capítulo dos justifica el origen y creación del prototipo para la transcripción de denuncias por violencia sexual y procesamiento de datos para análisis de patrones. En el tercer capítulo se detallan los alcances que se esperan con el desarrollo de esta investigación y creación del prototipo.

El capítulo cuatro contiene una lista de conceptos y definiciones relacionadas con la investigación que ayudaran a comprender el desarrollo, funcionamiento y finalidad del prototipo.

El quinto capítulo presenta los resultados obtenidos de las pruebas realizadas en el prototipo de reconocimiento de voz y procesamiento de información para análisis de patrones en los crímenes de violencia sexual a través de interpretaciones gráficas que facilitan la comprensión de la información.

En el capítulo seis, se discuten los resultados obtenidos de las pruebas que se realizarán para validar la efectividad del prototipo y se establece en qué medida los objetivos propuestos en la investigación fueron alcanzados.

1. ANTECEDENTES

La transcripción de declaraciones sobre crímenes sexuales es vital en instituciones que velan por el cumplimiento de las leyes en Guatemala. Estas transcripciones las realizan empleados de las instituciones que toman la declaración de la víctima, la interpretan y plasman la información en papel. Este proceso es incómodo para la víctima ya que debe proporcionar detalles de los crímenes en su contra a una persona desconocida. De esta incomodidad derivan declaraciones incompletas y falta de información necesaria para resolver los casos. Además de incómodo, el proceso es inseguro y propenso al error humano. Por estas razones se propone el desarrollo de un prototipo de reconocimiento de voz para transcripción de las declaraciones que hayan sido víctimas de crímenes sexuales con el objetivo de crear un ambiente donde se sientan libres de detallar los crímenes cometidos. El reconocimiento de voz no es una tecnología nueva en el mercado. Barrios, López, Mendieta, Benavidez, y Sáez (2018) en su artículo exponen que el reconocimiento de voz tuvo su origen en 1940 y lo definen como el enlace entre la persona y una máquina que nace tras la necesidad volver visible el habla para las personas con problemas auditivos y lo definen como la capacidad de una máquina de convertir la voz a lenguaje binario comprensible por la computadora.

El trabajo *Reconocimiento de voz* (Esparza Arellano y Avalos Briseño, 2003) presenta una vista general de cómo funcionan los sistemas de reconocimiento de voz y el uso algoritmos de las redes neuronales de aprendizaje, así como las dificultades que se presentan para reconocer las palabras con una gran cantidad de variantes en la voz de cada persona, como gesticulaciones, tono de la voz, continuidad o exactitud, también describen las

técnicas que se utilizan comúnmente en el reconocimiento de voz como FFT Fast Fourier Transform, densidad del cruce por cero en conjunto con el uso de redes neuronales, la misma que se utilizara en esta solución y que el artículo concluye que es la más eficiente para este tema.

También Oropeza y Suárez exponen la dificultad de implementar un sistema de reconocimiento de voz así como las alternativas para capturar la voz mediante la identificación de fonemas o de silabas específicamente en el idioma español, en el artículo destacan tres características principales para desarrollar su solución; la función de energía total en corto tiempo, la función de energía de altas frecuencias y un sistema básico de conocimiento que en conjunto les permitió obtener un 94.7 % de reconocimiento de voz, componentes con un alto porcentaje de eficiencia que permitiría transcribir con gran precisión las denuncias a las que está enfocado este proyecto. (Oropeza y Suárez Guerra, 2006)

Relacionado con la investigación Cruz y Acevedo describen el reconocimiento de voz como una tecnología que lleva muchos años en el mercado pero que aún está en desarrollo y la dificultad del reconocimiento se debe a cinco puntos que desarrollan en su artículo los cuales definen de la siguiente forma: locutor, no todas las personas pronuncian de la misma forma debido a su género, edad o su región de origen; forma de hablar, el reconocimiento de voz se complica debido a que la persona puede pronunciar de forma distinta una palabra que se encuentra al principio, en medio o al final de la oración; el vocabulario, el número de palabras que deba reconocer un sistema se complica entre mayor sea la cantidad de palabras debido a que entre más palabras sean se incrementa la posibilidad de que las palabras sean similares; gramática, si se incluye dentro de la solución reglas gramaticales podría ayudar de disminuir la ambigüedad de las palabras; el ambiente tiene un papel

importante, entre menor sea el ruido en el entorno, mayor es la probabilidad de un reconocimiento exitoso, estos factores influyen en el desarrollo del proyecto como consecuencia de la diversidad de acentos en la población. (Cruz Beltrán y Acevedo Mosqueda, 2009).

Uno de los puntos clave en esta investigación es la inclusión de redes neuronales artificiales para transcribir denuncias verbales de víctimas de crímenes sexuales, en el artículo *Sistema de reconocimiento de voz mediante wavelets, predicción lineal y redes backpropagation* (San Juan, Jamett, Kaschel, y Sánchez, 2016) utilizan el algoritmo redes neuronales backpropagation que consisten en el aprendizaje de un grupo predefinido de pares de entrada y salida que en conjunto con herramientas como Wavelets y predicción lineal presentan un resultado positivo para el reconocimiento de sílabas y vocales, el uso de las redes neuronales es una de las piezas más importantes para el éxito del proyecto debido que se utiliza para la distinción entre fonemas.

Es necesario tomar en cuenta que el área de la tecnología evoluciona constantemente y con ella las herramientas para el reconocimiento de voz, la implementación de interfaces de programación de aplicaciones (API) facilitan la adaptación de sistemas expertos desarrollados para el reconocimiento de voz, en el artículo *El Reconocimiento de Voz como alternativa de inclusión para discapacitados auditivos en un entorno educativo* (Montoto, y otros, 2018), plantean el uso de sistemas expertos para el reconocimiento y transcripción mediante la implementación de una API para integrar las funcionalidades desarrolladas y perfeccionadas durante años como: Google Cloud Speech, en la actualidad cuenta con soporte para el reconocimiento de voz de 120 idiomas; Sphinx un potente reconocedor de voz de código abierto o Dragon NaturallySpeaking que se especializa en la transcripción de discursos. Cada compañía incluye distintas funcionalidades dentro de sus desarrollos, adicional a

la variante de costo por los servicios ofrecidos, durante el desarrollo del prototipo de esta investigación se utilizara el API de Google para la transcripción.

Bajo la misma línea de investigación (Pérez, Javier, y Harari, 2017) en su artículo *Ejecución de comandos de voz mediante Web Speech API* plantean el desarrollo de una interfaz de reconocimiento de voz para interactuar con dispositivos inteligentes en la cual presinden del uso del teclado para dar instrucciones al dispositivo, en su forma más básica es el mismo concepto que se desea aplicar a esta investigación, transcribir una declaración presindiendo del empleado que toma las declaraciones, el enfoque principal de este artículo es el uso de Web Speech API que tiene como función principal el reconocimiento de voz para transformarlo en texto o viceversa, mediante el uso de una API que puede realizar las transcripciones en tiempo real.

Con base en los artículos e investigaciones expuestas, además de la adición de nuevas tecnologías, como la aplicación de redes neuronales y un API de reconocimiento de voz, se puede concluir que la creación de una aplicación que reconozca la voz y transcriba las denuncias de crímenes sexuales con el objetivo de ayudar a las instituciones que velen por el cumplimiento de las leyes en Guatemala sería factible y de gran beneficio para las víctimas e instituciones.

2. JUSTIFICACIÓN

Este proyecto tiene el propósito de contribuir con la línea de investigación de sistemas para impulsar la integración de sistemas de información. Proponiendo el desarrollo de un prototipo de sistema de información para el reconocimiento de voz que permita transcribir la voz e identificar patrones en las denuncias de víctimas de delitos de violencia sexual.

La información es uno de los activos más importantes para cualquier proceso, con este proyecto se pretende satisfacer la necesidad de contar con información detallada que permita al personal de instituciones como el Ministerio Público, Policía Nacional Civil o la Secretaría contra la Violencia Sexual, Explotación y Trata de Personas investigar los casos de agresiones sexuales mediante la transcripción de las denuncias verbales de las víctimas.

La seguridad es un elemento de suma importancia dentro del proyecto, para resguardar la información de forma segura; la solución cuenta con controles de acceso por usuario que permite determinar los permisos por cada tipo de transacción con objetivo de evitar pérdida, manipulación o acceso a la información a personal no autorizado.

Las instituciones carecen de información confiable debido a que las denuncias son transcritas por un empleado de la institución y están propensas a errores humanos involuntarios, la solución permite el reconocimiento de voz de la víctima para transcribir la denuncia y almacenarla en una base de datos para su posterior consulta mediante un conjunto de gráficos estadísticos que serán generados a través de parámetros que permiten presentar al usuario información

agrupada y ordenada, para analizar, interpretar y detectar tendencias de incremento o disminución de crímenes o para los fines que el usuario considere necesario.

El sistema está orientado para satisfacer la necesidad de que las instituciones cuenten con una herramienta capaz de transcribir y almacenar denuncias verbales con el objetivo de crear un ambiente donde la víctima se sienta cómoda y libre de dar los detalles de los crímenes cometidos en su contra y con esto recopilar la mayor cantidad de información para las investigaciones necesarias.

El uso de herramientas de inteligencia de negocios facilitará la identificación de patrones dentro de las transcripciones que permitirá prevenir crímenes identificando la relación que hay entre los crímenes de violencia sexual y el género, la edad o la ubicación de las víctimas.

Una vez implementada la solución se podría ampliar su uso a cualquier otro departamento de este tipo de instituciones para la transcripción de cualquier tipo de denuncia verbal a texto con el objetivo de eficientizar el tiempo para tomar las denuncias, eficientizar la recopilación de detalles y el análisis de la información.

3. ALCANCES

3.1. Resultados

Se desarrolló la aplicación móvil, capaz de transcribir la voz a texto con el objetivo de transcribir las denuncias por crímenes sexuales para evitar que las víctimas sientan la incomodidad al proporcionar detalles de las agresiones cometidas en su contra a personas encargadas de transcribir su declaración, las declaraciones se almacenan en el móvil y luego se sincronizan a un servidor. Además, durante la sincronización se crea un data warehouse y mediante una herramienta de inteligencia de negocios se realiza el análisis de todas las declaraciones para determinar la relación de las declaraciones de crímenes sexuales con respecto a variables de edad, género o ubicación.

El uso del prototipo influye en la velocidad y exactitud de la toma de la declaración, reduciendo el error humano al transcribir las declaraciones, mejorando el detalle de la declaración debido a que la víctima no siente vergüenza al detallar el crimen en comparación al declarar frente a otra persona.

3.2. Técnicos

Se utilizaron herramientas y tecnología para el desarrollo del prototipo las cuales se presentan a continuación:

- Uso del gestor de base de datos Microsoft SQL Server para la creación de una base de datos relacional que permita el almacenamiento de la información obtenida de las transcripciones generadas por el prototipo.

- Uso del sistema experto Google Speech-to-text que permite el reconocimiento y transcripción de la voz en 120 idiomas mediante entrenamiento de las redes neuronales de aprendizaje artificiales por los usuarios de Google alrededor mundo.
- Uso del API Google Cloud Speech que permite la conexión del prototipo con el sistema experto de reconocimiento de voz, redes neuronales artificiales de aprendizaje y transcripción de voz de Google que devuelve la transcripción de la voz para ser almacenada en una base de datos con la información personal del usuario, en conjunto con el uso de un lenguaje de programación orientado a objetos para el desarrollo del prototipo que permita la captura de voz desde el micrófono del dispositivo móvil y realice la recepción de la transcripción a texto para almacenarlo en una base de datos.
- Uso de un data warehouse para integración y depuración de información que permita el análisis de la información mediante inteligencia de negocios.
- Uso de la herramienta Microsoft Power BI para el análisis de las denuncias transcritas por el software para determinar la relación entre las variables de género, edad o ubicaciones y encontrar patrones en los casos de abuso sexual.
- Realizar pruebas con distintos usuarios que permitan determinar el porcentaje de efectividad del prototipo y ajustar las configuraciones y herramientas que sean necesarias para que la transcripción sea lo más cercana a la denuncia verbal de la víctima.

3.3. Investigativos

Con base en el reconocimiento y transcripción de voz en el cual se basa esta investigación, se presentan los siguientes alcances descriptivos.

- Seleccionar el motor de base de datos que mejor se adecue a las necesidades del proyecto, cumpliendo con las necesidades de garantizar seguridad, integridad y disponibilidad de la información para el prototipo.
- Investigar las arquitecturas de software existentes con el objetivo de seleccionar la arquitectura que mejor se adapte a las necesidades para el desarrollo del prototipo.
- Detallar el uso, ventajas, desventajas y forma de implementación de los distintos modelos de inteligencia de negocios para el análisis de la información obtenida de la transcripción de las denuncias por delitos de abuso sexual.
- Describir las herramientas y el proceso para el análisis de la información mediante modelos de inteligencia de negocios que permitan determinar tendencias sobre los casos de abuso sexual.
- Presentar los resultados de la interconexión de una Interfaz de programación de aplicaciones (API) al prototipo en un sistema experto de reconocimiento de voz para enviar audio, reconocer la voz y devolver la transcripción para ser almacenada en una base de datos.

4. MARCO TEÓRICO

4.1. Reconocimiento de voz

El reconocimiento de voz se obtiene mediante la implementación de un conjunto de tecnologías que reconoce las ondas de la voz y produce un grupo de caracteres que equivalen al obtenido como salida. Depende de las características de las personas: la estructura física del tracto vocal o la forma en que la persona se comporta al hablar, para el proceso de reconocimiento es necesario tomar en cuenta la variabilidad de la señal de voz ya que una persona no puede repetir las palabras exactamente igual. Una vez obtenidas las muestras de voz se analiza detenidamente la frecuencia de las palabras o conjunto de palabras comparando distintas variables como la intensidad, duración, intensidad o tono de la voz, una vez procesada la muestra se realiza la comparación con la información almacenada en la base de datos determinando la similitud con base en las propiedades de la señal de la voz.

Una de las características básicas principales al obtener las señales de voz es la frecuencia de muestreo, esta característica hace la diferencia entre una buena calidad de señal y dificultades al no seguir las reglas establecidas por el procesamiento digital de señales, para garantizar el funcionamiento del reconocimiento de voz es necesaria una fase de entrenamiento para obtener patrones de características de los locutores y almacenarlos para su posterior consulta. Se debe considerar la susceptibilidad del canal de transmisión y las variaciones del micrófono o del ruido ambiental que se genere en el canal de transmisión ya que como consecuencia se generan errores en la transcripción. (Gil, 2016)

4.1.1. Algoritmos de reconocimiento de voz

En esencia los algoritmos para el reconocimiento de voz son secuencias de pasos que permiten transcribir las ondas sonoras de la voz en texto. El reconocimiento de patrones y sistemas biométricos de la voz a través de computadoras es un proceso complejo para lo cual es necesario contar con distintas herramientas que permitan capturar e identificar las palabras de un archivo de sonido, la aplicación de algoritmos para el reconocimiento de voz es esencial para la transcripción de las palabras emitidas por el ser humano.

4.1.2. Motores de reconocimiento de voz

Existen distintos métodos para el reconocimiento de voz. Mediante codificación e implementación de modelos especializados, como el Modelo Oculto de Márkov o Redes Neuronales Artificiales, así como el uso de sistemas disponibles en el mercado desarrollados y disponibles para su integración mediante APIs. Estos modelos permiten integrar sus funcionalidades como parte del desarrollo, cuentan con diversas ventajas y desventajas, seleccionar uno u otro depende de las necesidades y características del software donde se desee incorporar, los motores de reconocimiento de voz incluyen dentro de su funcionalidad rutinas y subrutinas esenciales para transcribir de forma satisfactoria la voz a texto, son varias las empresas que ofrecen estos motores entre las que se encuentran Google, Apple, Microsoft o Nuncce, cada una aporta distintas características en su motor así como distintos costos por su implementación. (Montoto et al. 2018)

4.2. Google cloud

Es una plataforma que reúne las aplicaciones y servicios para desarrollo que Google pone a disposición del usuario para integrar a sus soluciones, comúnmente utilizada para crear soluciones con tecnología almacenada en la nube, destacan la rapidez y escalabilidad de su infraestructura en la creación de soluciones dependiendo de las necesidades de los usuarios. Google Cloud es el espacio virtual que permite ejecutar tareas que anteriormente necesitaban hardware o software para realizarlas en la actualidad únicamente es necesaria la plataforma de nube de Google para la administración del almacenamiento y gestión de datos simplificando las tareas al usuario.

Google Cloud pone a disposición de los usuarios diversidad de soluciones para adaptar a distintos proyectos, por ejemplo, *cloud compute* ofrece opciones escalables de procesamiento ajustables, máquinas virtuales personalizables según la necesidad de memoria o almacenamiento. *Cloud storage* proporciona espacio de almacenamiento y recuperación de cualquier cantidad de información. (Srinivasan, Ravi, y Raj, 2018)

4.2.1. Google Cloud Speech-to-Text

Se especializa en el reconocimiento y transcripción de la voz a texto mediante el uso de modelos de red neuronal, pone a disposición del usuario un API que permite integrar tecnología de reconocimiento de voz a las aplicaciones de los desarrolladores. Esta API reconoce más de 120 idiomas y variantes por lo puede utilizarse con una gran variedad de usuarios. Puede procesar archivos de audio en tiempo real o grabado previamente, aplicando la tecnología de aprendizaje automático de Google y los algoritmos de aprendizaje profundo en redes neuronales que dan como resultado un alto nivel de precisión en sus

resultados. La ventaja más significativa de Cloud Speech to text es la mejora continua de su tecnología a medida que los usuarios alrededor del mundo utilizan sus productos, también cuenta con la opción de reconocimiento personalizado en el cual se pueden agregar hasta 5,000 palabras que probablemente se encuentren en el entorno para el que sea diseñada la aplicación. (Pérez, Javier, y Harari, 2017)

4.3. Lenguajes de programación

Son idiomas artificiales en los que se puede definir una secuencia de instrucciones para su procesamiento por computadoras, poseen distintas utilidades y aplicaciones que facilitan y reducen las tareas realizadas por los usuarios haciendo que estos evolucionen cada vez más durante el tiempo facilitando la vida de los usuarios agregando nuevas funcionalidades que 50 años atrás no existían. (Aguilar, 2003)

Con base en el estilo de programación del lenguaje pueden clasificarse como:

4.3.1. Imperativos

Estos lenguajes de programación basan su funcionamiento en un conjunto de instrucciones secuenciales que, al ejecutarse, alteran las regiones de memoria donde residen los valores de las variables.

4.3.2. Declarativos

Se enfocan en el resultado que obtendrán. Estos, a su vez, se dividen en: funcionales, los cuales se representan mediante expresiones que permiten

obtener resultados a partir de una serie de argumentos. Los lógicos se basan en el cálculo de predicados, una teoría matemática que permite que una computadora con base en un conjunto de hechos y reglas lógicas pueda derivar soluciones inteligentes.

4.3.3. Orientado a objetos

Es un paradigma para el diseño de aplicaciones. Utiliza la abstracción de objetos para resolver problemas. Los objetos son paquetes de software que constan de atributos y métodos que interactúa con la información. Se nombra programación orientada a objetos porque el método abstrae las cosas del mundo real como objetos para crear la solución a un problema. En su mayoría los elementos en nuestro entorno podrían tomarse como un ejemplo de objeto; una casa, un dispositivo móvil, la vestimenta o un escritorio, la base para este paradigma radica en varias técnicas detalladas a continuación:

- Herencia es la técnica de crear clases basándose en clases existentes, incluyendo métodos, características y propiedades de la clase padre, el beneficio principal de utilizar la herencia es la posibilidad de crear características y métodos nuevos para la clase hija sin tener que editar la clase padre o crear la clase desde cero.
- Polimorfismo se refiere a la capacidad de los objetos que pertenecen a una clase de realizar sus funciones de distintas formas, su respuesta depende de la parametrización utilizada en el momento en que se invoca.
- Abstracción es la capacidad de definir las características que distinguen a un objeto de los demás, separa el comportamiento específico de un objeto.
- Encapsulamiento es el proceso mediante el cual se protegen y administran los accesos a los métodos de una clase para asegurar la integridad de la información. Los niveles de acceso que soporta el encapsulamiento son:

público, todos tienen acceso a métodos y datos de la clase; protegido, únicamente pueden accederse en el interior de la clase o subclases; privado en el cual se deben declarar miembros accesibles solo para la clase.

4.4. Base de datos

El término base de datos se refiere a un contenedor que tiene como objetivo almacenar información y permitir al usuario modificar, eliminar o recuperar la información con base en solicitudes realizadas según sus necesidades. (Date, 2001)

4.4.1. Gestor de Base de Datos (SGBD)

Un gestor de base de datos es el software que permite crear, editar y administrar bases de datos, también cuenta con las herramientas para la edición y búsqueda de información de forma eficiente.

4.4.2. Modelo relacional

Es una estructura que administra los datos y permite que sean manipulados de forma segura y que permita evitar errores, el modelo relacional se basa en principios matemáticos de la teoría de conjuntos, facilita la recuperación de datos de forma sencilla, garantiza la precisión y coherencia de los datos y provee una estructura de información independiente al acceder a los datos almacenados, el elemento principal del modelo relacional son las relaciones.

Las relaciones son un conjunto de filas y columnas agrupadas en forma de tabla que representan una entidad que no se repite dentro de la estructura formada por los datos relacionados, las entidades son tablas que simulan un objeto del que se recolectan datos, como un vehículo, una cuenta bancaria o la información de una persona. (Oppel y Sheldon, 2009)

4.4.3. SQL Server

Es un gestor de bases de datos relacionales desarrollado por Microsoft, tiene la capacidad de gestionar datos conservando su coherencia e integridad, este gestor tiene como función, almacenar datos, verificar restricciones de integridad establecidas por los administradores de la base de datos, asegurar que los datos almacenados sean coherentes, brindar seguridad a los datos con base en las directrices establecidas por Windows. (Gabillaud, 2015)

4.4.4. SQLITE

Es una biblioteca de dominio público creada en el lenguaje de programación C, brinda un sistema gestor de base de datos para almacenar registros en tablas definidas por el usuario, adicional al manejo y almacenamiento de datos SQLite procesa consultas complejas combinando varias tablas, dentro de las diferencias más sobresalientes de este gestor se encuentran la facilidad de configuración, la mínima sobre carga administrativa y de recursos que requiere para funcionar, no requiere un servidor principal para operar, es compatible con las transacciones ACID, además SQLite es compatible con la mayoría de funciones del estándar de consulta SQL92. (Kreibich, 2010)

4.5. Interfaz de programación de aplicaciones (API)

API es un grupo de reglas que describen como una aplicación puede interactuar con otra, contiene los mecanismos que permiten que esta comunicación sea posible con el objetivo de intercambiar información o funcionalidades entre aplicaciones sin necesidad de tener un conocimiento profundo sobre la estructura o desarrollo de las aplicaciones. (Jensen, 2016)

4.6. Arquitectura de sistema

Se enfoca en la estructura general de un sistema, en el proceso de creación de software el primer paso a evaluar es el diseño de la arquitectura, es esencial identificar los elementos que lo conformaran y determinar la comunicación y su relación en conjunto.

La importancia de la arquitectura de software radica en el efecto que tiene en el desempeño, potencia y capacidad de la distribución y mantenimiento de los sistemas. (Sommerville, 2011)

4.6.1. Patrones de arquitectura

Es la forma de exponer, compartir y reutilizar el conocimiento de los sistemas de información actuales, son un conjunto de buenas prácticas que se han utilizado en diferentes sistemas y entornos, también llamados arquetipos describen los elementos que conforman un sistema, la relación y restricciones que existen entre ellos para su correcto funcionamiento.

4.6.2. Arquitectura en capas

La esencia de una arquitectura en capas es la capacidad de separar el desarrollo del software en distintas capas para realizar modificaciones de forma autónoma, el patrón modelo-vista-controlador separa los componentes para desarrollar un sistema dividido en varias capas en el cual existen una relacionan de capa inferior con la capa superior más cercana.

4.6.3. Arquitectura de repositorio

Esta arquitectura explica como un grupo de elementos comparten datos mediante su interrelación y como se administran de forma centralizada, disponible para todos los componentes del sistema, los elementos no se relacionan directamente, lo hacen con el apoyo de un repositorio central, se utilizan en casos en los que un sistema almacena grandes volúmenes de información por tiempo prolongado. Los componentes de este sistema tienen la posibilidad de ser independientes, no es necesario que conozcan la existencia de otros componentes, cada componente puede propagar sus cambios hacia otros componentes, el problema con esta arquitectura es que cuenta con un punto de falla único, si el repositorio central tiene problemas afecta todo el sistema.

4.6.4. Arquitectura cliente servidor

Mediante la implementación de una arquitectura cliente-servidor las funciones del sistema se organizan en servicios y cada servicio lo proporciona un servidor independiente, los clientes son usuarios del sistema y para utilizar los servicios ingresan a los servidores.

La arquitectura cliente servidor se utiliza cuando el repositorio de información es compartido y es necesario utilizarlo desde distintas ubicaciones, los servidores se pueden distribuir dentro de una red facilitando el acceso y funcionalidad a los clientes, un sistema bajo el modelo cliente-servidor se organiza como un grupo servidores y servicios asociados y clientes dan uso a los servicios.

Los elementos esenciales para crear una arquitectura de este tipo son: servidores que ofrecen servicios a otros componentes, clientes que solicitan servicios ofrecidos por los servidores y una red que permite que los clientes puedan acceder a los servicios.

Una de las características principales del modelo cliente-servidor es su diseño de arquitectura distribuida, se puede utilizar en sistemas de red con varios procesadores distribuidos y facilita la integración de nuevos servidores al resto del sistema.

4.7. Desarrollo de aplicaciones móviles

Es el proceso de la creación de un software para dispositivos móviles como teléfonos inteligentes o tabletas, las aplicaciones pueden estar precargadas en los dispositivos, descargadas de las tiendas de aplicaciones o instalarse directamente al dispositivo. (Prieto Blázquez, Ramírez Vique, Morillo Pozo y Domingo Prieto, 2011)

4.7.1. Aplicaciones WEB

Son básicamente sitios web optimizados específicamente para dispositivos móviles están disponibles en una URL y no se instala directamente en el dispositivo.

4.7.2. Aplicaciones nativas

Estas aplicaciones están instaladas directamente en el dispositivo móvil y poseen el acceso al hardware del dispositivo como bocinas, acelerómetro, cámara, micrófono o el GPS y están escritas en lenguajes de programación compilados.

4.8. Dispositivos móviles

Generalmente, los dispositivos móviles pueden definirse mediante cuatro características principales que deben poseer y que hacen la diferencia con otros dispositivos, estas características son movilidad, tamaño reducido, comunicación inalámbrica, interacción con las personas.

4.8.1. Movilidad

Es la característica de un dispositivo de ser transportado o movido de forma frecuente y fácil, esta es la característica principal que define a un dispositivo móvil, puede ser transportado y utilizado durante su transporte independiente de la proximidad de una fuente de energía o una conexión física a internet.

4.8.2. Tamaño reducido

Para que un dispositivo pueda considerarse móvil debe poderse utilizar fácilmente con una o con las dos manos sin necesidad de un soporte externo.

4.8.3. Comunicación inalámbrica

Otra de las características que debe poseer un dispositivo móvil es la independencia de conexión, es decir se considera un móvil a un dispositivo que tiene la capacidad de conectarse a una red sin necesidad de cables.

4.8.4. Interacción con las personas

La interacción se refiere a la facilidad del proceso de uso entre un usuario y su dispositivo entre los que intervienen distintos factores como la ergonomía y la usabilidad.

4.9. Inteligencia de negocios

Es un grupo de metodologías, capacidades, aplicaciones y prácticas que tienen como principal objetivo crear y administrar datos que permite presentar información procesada a las organizaciones para influir en el proceso de toma de decisiones, el procesamiento de la información permite entender de mejor forma la información y encontrar problemas y oportunidades. (Carrillo Osorio, 2018)

4.9.1. Data Warehouse

Es el proceso de extraer datos de distintos orígenes para depurar, estructurar y almacenarlos en un repositorio de datos consolidado con el objetivo de analizar la información, un data warehouse resume, desglosa, ordena y compara los datos de distintas bases de datos de la organización. El diseño de un data warehouse se divide en tres fases en la primera fase se examina el esquema de entidad relación de la base de datos que contiene la información operacional para generar esquemas multidimensionales; en la segunda fase se recolectan los requisitos del usuario mediante entrevistas para obtener información del análisis que se necesita y en la tercera fase unifica la información obtenida en la segunda etapa con los esquemas generados en la primera. (Rainardi, 2008)

4.9.2. OLAP (On-Line Analytical Processing)

El procesamiento analítico en línea permite agregar orígenes de datos y organizarlos en subconjuntos con el objetivo de darle un significado que ayude a encontrar oportunidades y problemas.

4.9.3. Sistema ROLAP

ROLAP utiliza los datos de un data warehouse para realizar un análisis OLAP, la principal característica del sistema ROLAP es que el proceso analítico es más eficiente con bases de datos relacionales.

5. PRESENTACIÓN DE RESULTADOS

5.1. Análisis de arquitectura del prototipo para transcripción de voz

En la figura 1 se presenta la solución completa propuesta en este trabajo de investigación para transcribir las denuncias de víctimas de crímenes sexuales de voz a texto, almacenamiento y consulta de información para el estudio de patrones de denuncias de crímenes sexuales. Esta solución permite incrementar la exactitud y reducir el tiempo al obtener las declaraciones de las víctimas mediante aplicación de tecnología. Se divide en dos partes principales, una se conforma del prototipo para la transcripción de las declaraciones de voz a texto y sincronización a una base de datos central y la otra, la aplicación de inteligencia de negocios para la comprensión y detección de patrones derivados de las declaraciones obtenidas.

5.1.1. Transcripción de voz a texto

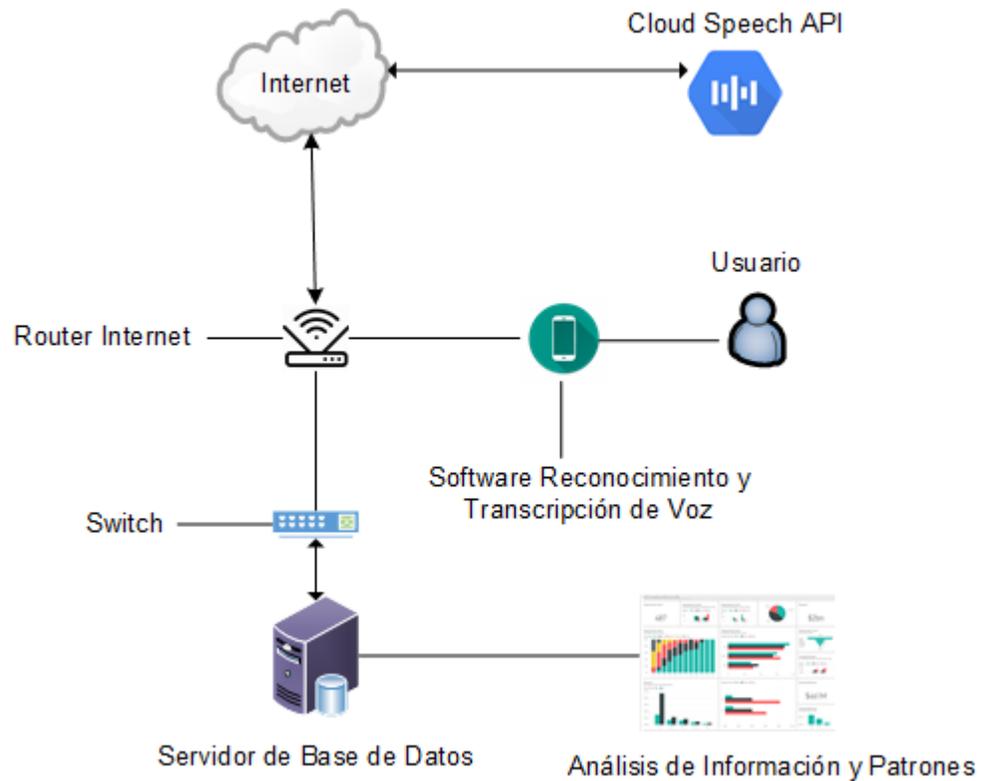
Para tomar la declaración de la víctima se debe realizar el siguiente proceso: la persona encargada de tomar la declaración abre la aplicación y entrega el dispositivo a la víctima, la víctima completa el formulario con sus datos personales; se moviliza al siguiente fragmento de la aplicación y acciona el botón para la iniciar la transcripción y da su declaración; se crea una transcripción de voz a texto de la declaración de las víctimas mediante el uso del prototipo el cual se conecta mediante internet a la interfaz de programación de aplicaciones o API Speech-to-text creada por Google para transcribir la voz a texto en tiempo real. Al finalizar. la víctima debe accionar el botón almacenar para que su declaración quede grabada localmente en la base de datos SQLite para luego sincronizar los

datos al servidor central en el que se extrae, transforma y carga la información para crear un data warehouse.

5.1.2. Inteligencia de negocios

Microsoft Power BI es la herramienta de escritorio seleccionada para analizar y representar de forma gráfica la información proporcionada por las víctimas de crímenes sexuales para el análisis y detección de patrones, esta parte de la solución está destinada al personal administrativo de una institución enfocada a la reducción de delitos de índole sexual permitiéndole determinar por diversas variables como ubicación, edad y género las personas más afectadas por los crímenes de violencia sexual.

Figura 1. **Diagrama de funcionalidad del prototipo**



Fuente: elaboración propia. *Software Microsoft Visio.*

5.2. Eficiencia de transcripción del prototipo

El prototipo transcribe las palabras con un 94.41 % de efectividad. Para obtener este resultado se realizaron pruebas con diez usuarios mediante la lectura de un mismo texto en lapsos de 30, 60 y 90 segundos. La tabla II muestra que el prototipo tuvo un resultado de 94.14 % de eficiencia en las pruebas de 30 segundos de duración. La tabla III muestra los resultados del prototipo para lectura de 60 segundos donde hubo un 93.96 % de eficiencia y la Tabla IV muestra un 94.78 % de eficiencia para transcripciones con una duración de 90 segundos.

Tabla II. Resultado de pruebas del prototipo de 30 segundos

Segundos	Palabras	Correctas	Erróneas	% Correcto
30	83	76	7	91.57 %
30	81	77	4	95.06 %
30	79	77	2	97.47 %
30	81	77	4	95.06 %
30	66	60	6	90.91 %
30	80	71	9	88.75 %
30	79	77	2	97.47 %
30	75	70	5	93.33 %
30	80	76	4	95.00 %
30	62	60	2	96.77 %
Total	766	721	45	94.14 %

Fuente: elaboración propia.

Tabla III. Resultado pruebas del prototipo de 60 segundos

Segundos	Palabras	Correctas	Erróneas	% Correcto
60	123	113	10	91.87 %
60	131	119	12	90.84 %
60	115	111	4	96.52 %
60	120	118	2	98.33 %
60	118	109	9	92.37 %
60	105	99	6	94.29 %
60	110	100	10	90.91 %
60	113	109	4	96.46 %

Continuación tabla III.

Segundos	Palabras	Correctas	Erróneas	% Correcto
60	112	108	4	96.43 %
60	119	109	10	91.60 %
Total	1166	1095	71	93.96 %

Fuente: elaboración propia.

Tabla IV. **Resultados de pruebas del prototipo de 90 segundos**

Segundos	Palabras	Correctas	Erróneas	%Correcto
90	217	210	7	96.77 %
90	225	214	11	95.11 %
90	193	182	11	94.30 %
90	178	172	6	96.63 %
90	175	160	15	91.43 %
90	203	189	14	93.10 %
90	185	171	14	92.43 %
90	180	173	7	96.11 %
90	203	196	7	96.55 %
90	174	166	8	95.40 %
Total	1933	1833	100	94.78 %

Fuente: elaboración propia.

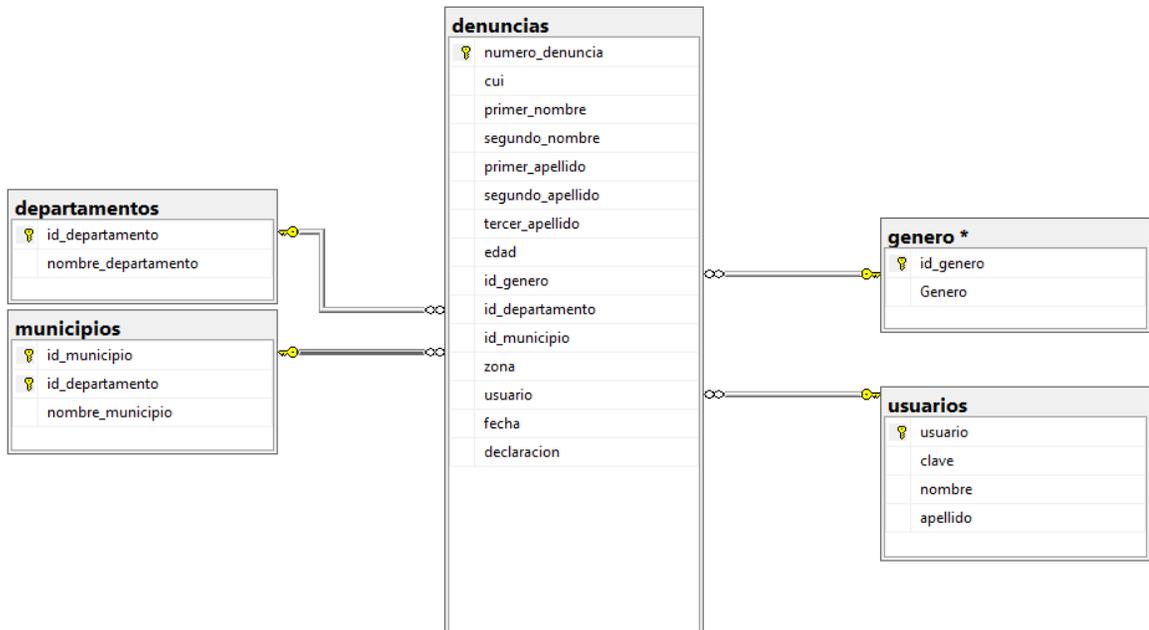
5.2.1. Tiempos de transcripción

Se realizaron pruebas de transcripción manual tomando como base la lectura utilizada para medir la eficiencia de la transcripción con duración de 90 segundos. La misma lectura realizando la transcripción manual tomó un promedio de 421 segundos. El promedio de incremento de tiempo para realizar una transcripción manual es de 331 segundos, es decir, transcribir una declaración de forma manual, exactamente igual a la transcrita por la aplicación, toma un 78.62 % más del tiempo del que toma realizarlo por la aplicación.

5.3. Estructura de información de datos del prototipo para transcripción de voz para denuncias de crímenes sexuales

Para almacenar la información personal y denuncia de la víctima se creó una estructura normalizada que permitirá mantener la coherencia de la información y evitar redundancias. Esto incrementa la eficiencia de la aplicación y la fase del análisis de las denuncias. Esta estructura es utilizada tanto en SQLite utilizado en el dispositivo donde se instalará la aplicación, así como en la base de datos SQL que servirá para hacer consultas de información y crear el data warehouse, aunque SQLite no posee un modelo relacional. La estructura se mantiene para almacenar y relacionar la información, la figura 2 muestra el diagrama de la base de datos creada en SQL Server 2016.

Figura 2. Diagrama entidad relación base de datos



Fuente: elaboración propia. *Software SQL Server 2016.*

5.4. Implementación y funcionamiento del prototipo de transcripción de voz para denuncias de crímenes sexuales y reconocimiento de patrones

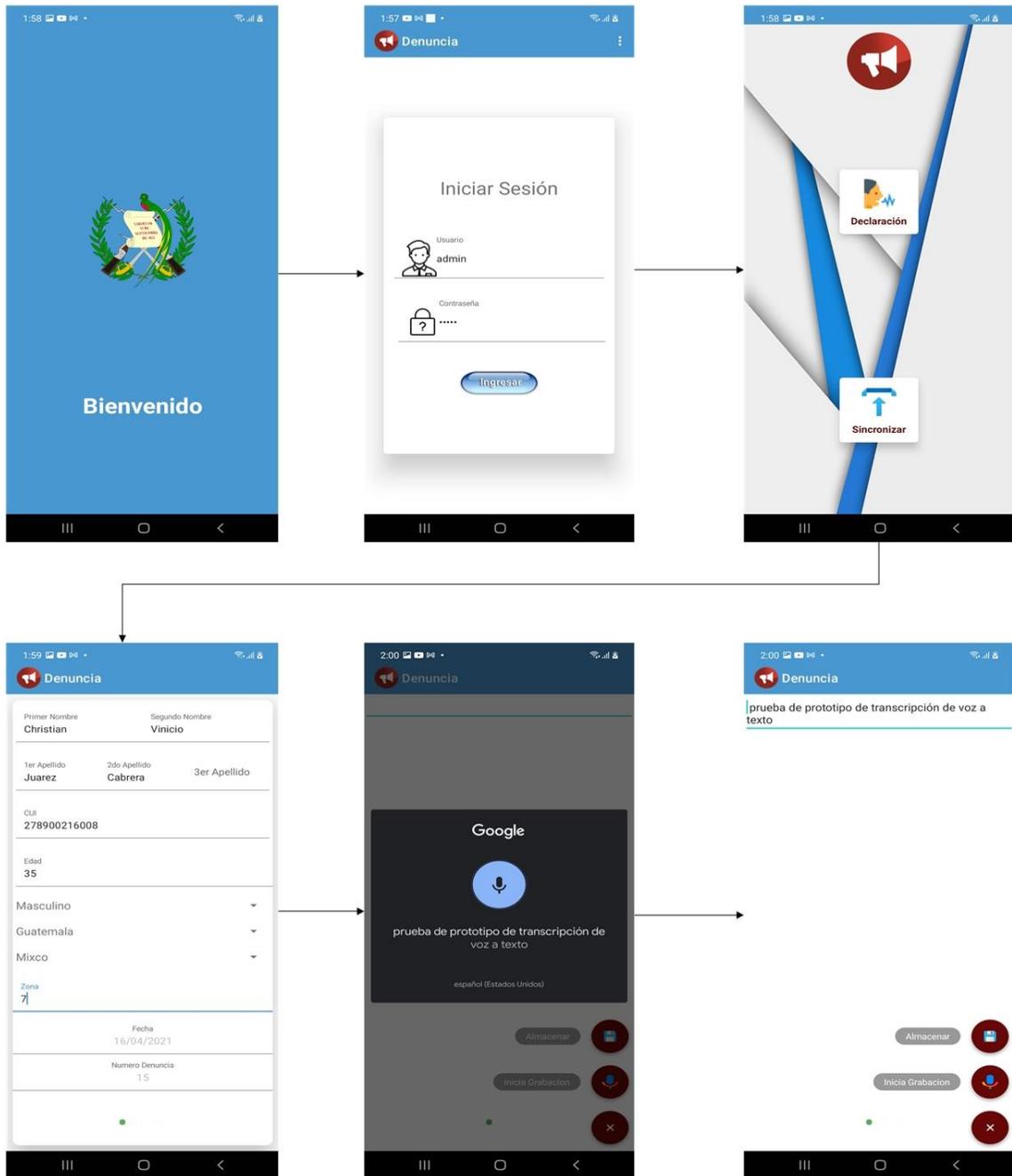
Este prototipo fue creado con el entorno de desarrollo integrado Android Studio y se utilizó Java como lenguaje de programación. Está conformado por 3 actividades; pantalla de bienvenida, autenticación de usuario, menú principal, actividad de denuncia y proceso de sincronización. La actividad de autenticación de usuario permite que el responsable de la aplicación ingrese de forma segura mediante el ingreso de su usuario y clave. Esta clave esta encriptada con el algoritmo md5 para evitar que personas no autorizadas tengan acceso a las funciones e información de la aplicación. El menú principal contiene dos botones,

uno para la toma de declaración y otro para realizar la sincronización de la información de la base de datos SQLite de la aplicación a la base de datos SQL. La actividad de denuncia está dividida en dos fragmentos, uno se utiliza para la toma de datos personales y el otro para la transcripción de voz a texto cuando la víctima realiza su declaración.

El prototipo tiene la capacidad de realizar la transcripción de forma síncrona o asíncrona esto quiere decir que, aunque el dispositivo no cuente con señal de internet puede realizar la transcripción y cuando tenga señal de internet sincronizar la declaración al servidor SQL para el análisis de la información.

La figura 3 representa de forma gráfica el flujo de trabajo dentro de la aplicación.

Figura 3. Flujo de trabajo del prototipo



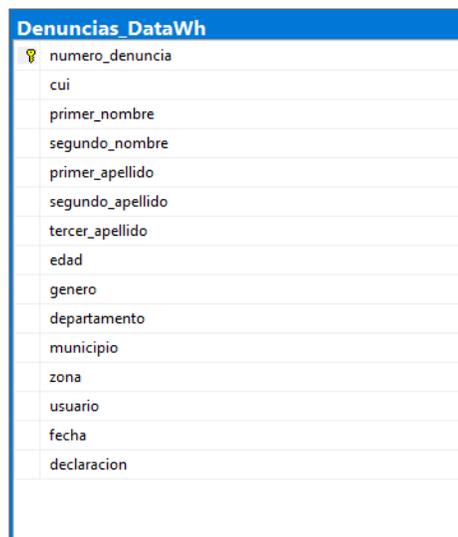
Fuente: elaboración propia, capturas funcionamiento prototipo.

5.5. Esquema de arquitectura Data Warehouse del prototipo de transcripción de voz para denuncias de crímenes sexuales

El prototipo contempla, dentro su estructura, un proceso para extraer, transformar y cargar información o ETL (siglas del nombre en inglés: extract, transform and load) que integra la etapa de inteligencia de negocios.

Se aplica el proceso ETL para la cohesión de las aplicaciones e integración de la información que contiene las denuncias de las víctimas sexuales hacia el repositorio de información destinado al data warehouse para examinar la información en búsqueda de patrones. La figura 4 muestra la tabla que se utiliza como origen para realizar el data warehouse en búsqueda de patrones que permitan identificar la población más vulnerable ante estos delitos, partiendo de las variables de edad, género y ubicación.

Figura 4. **Arquitectura data warehouse del prototipo**



Denuncias_DataWh	
PK	numero_denuncia
	cui
	primer_nombre
	segundo_nombre
	primer_apellido
	segundo_apellido
	tercer_apellido
	edad
	genero
	departamento
	municipio
	zona
	usuario
	fecha
	declaracion

Fuente: elaboración propia. *Software SQL Server 2016.*

5.6. Inteligencia de negocios para análisis y búsqueda de patrones

Para analizar la información se utiliza la herramienta Microsoft Power BI que presenta información gráfica y fácil de entender al usuario final. El principal objetivo es detectar patrones dentro de la información obtenida por el prototipo y determinar que variables influyen a que estos actos sean más recurrentes, las variables que se analizan con esta herramienta son edad, género y ubicación.

El data warehouse se alimentó con información proporcionada por la Secretaría Contra la Violencia Sexual, Explotación y Trata de Personas SVET de las denuncias de crímenes de violencia sexual recibidas por el Ministerio Público del 1 de enero del 2020 al 31 de diciembre del 2020, con el objetivo de presentar gráficas de ubicación apegadas a la realidad que vive la sociedad guatemalteca, la figura 5 presenta el tablero principal de información, La figura 5 presenta el tablero principal de información del prototipo.

Figura 5. Tablero de información de denuncias

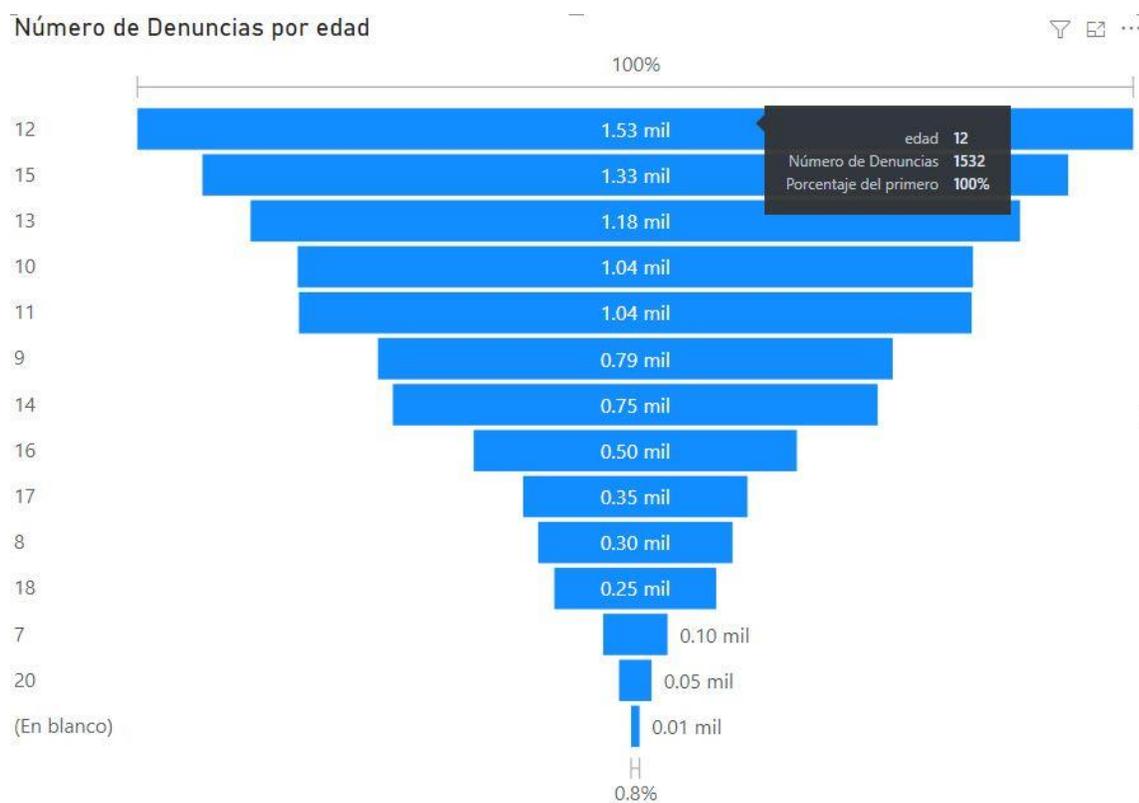


Fuente: elaboración propia. Software Microsoft Power BI.

5.6.2. Denuncias por edad

La solución se apoya en el uso de una gráfica de embudo para presentar al usuario de mayor a menor las denuncias clasificadas por edad. Como se ejemplifica en la figura 7, según la información de prueba suministrada al prototipo, los niños de doce años son los más afectados por los crímenes de índole sexual con un total de 1532 denuncias.

Figura 7. Gráfica de embudo, denuncias por edad



Fuente: elaboración propia. Software Microsoft Power BI.

5.6.3. Denuncias por género

Debido a la cantidad de variantes de género se utiliza la gráfica de anillo para representar estas denuncias. Como se ejemplifica en la figura 8, el usuario puede saber simplemente con ver la gráfica que el 97.83 % del total de las denuncias las realizan mujeres y un 2.17 % los hombres, según los datos de prueba utilizados en el prototipo.

Figura 8. Gráfica de denuncias por género



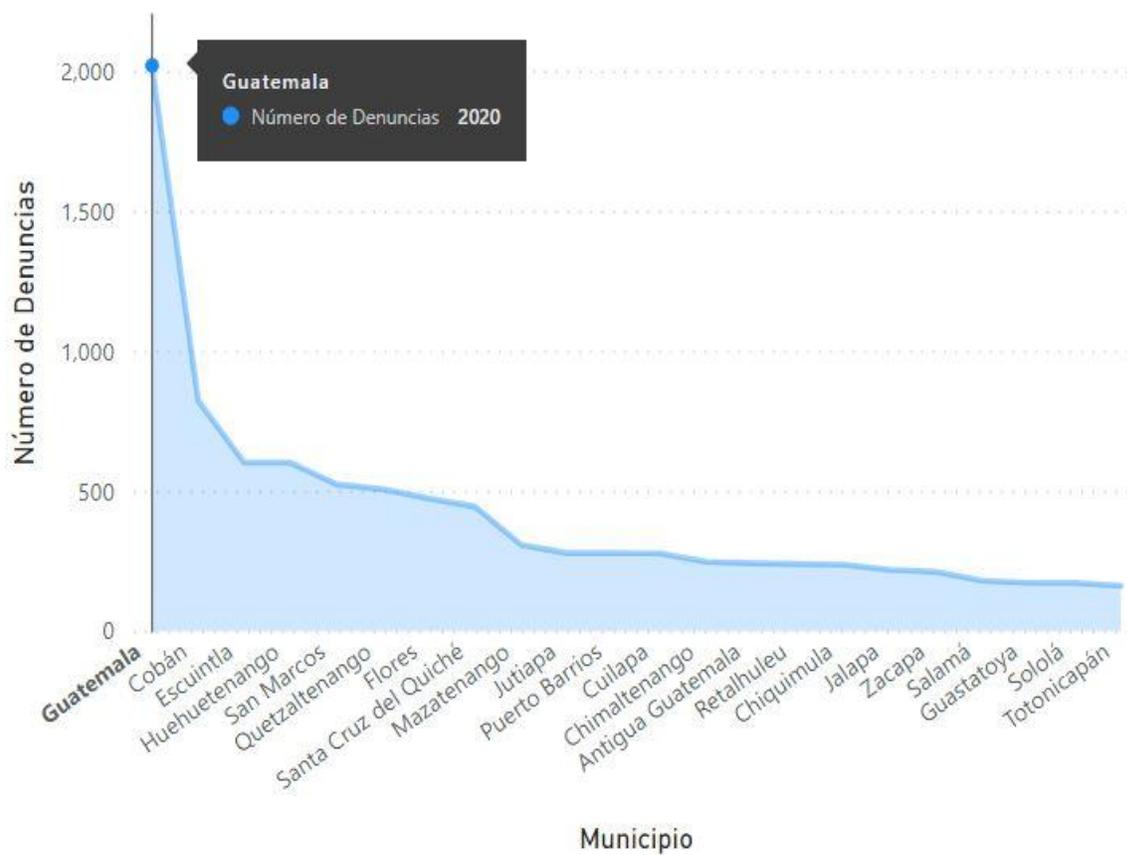
Fuente: elaboración propia. Software Microsoft Power BI.

5.6.4. Denuncias por departamento y municipio

Adicional a la gráfica de mapa explicada antes, también se usa la gráfica de área para presentar al usuario las denuncias de crímenes sexuales por

municipio e identificar de forma rápida y fácil los municipios con mayor o menor cantidad de denuncias de crímenes sexuales, como se ejemplifica en la figura 9.

Figura 9. **Gráfica de área, denuncias por municipio**



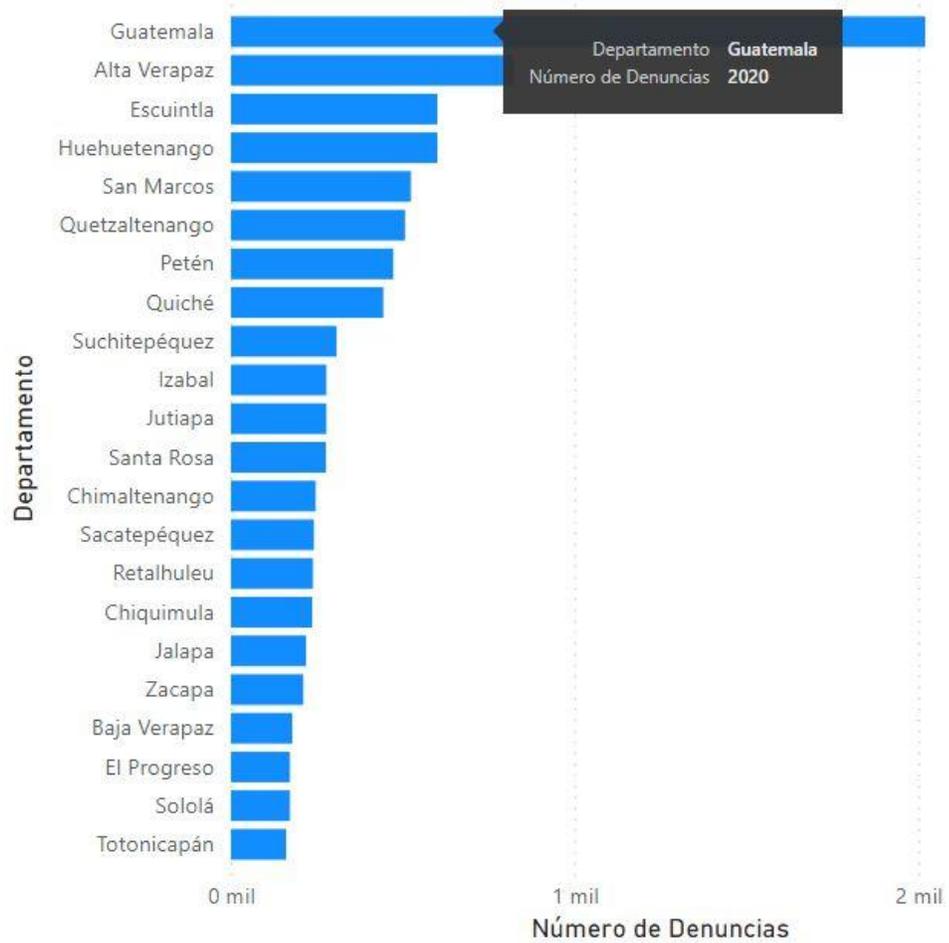
Fuente: elaboración propia. Software Microsoft Power BI.

5.6.5. Denuncias por departamento

Para establecer el departamento con menor o mayor cantidad de denuncias, se emplea una representación gráfica de barras apiladas que muestra

la información al usuario de forma resumida y fácil de entender cómo se expone en la figura 10.

Figura 10. **Gráfica de barras, denuncias por departamento**



Fuente: elaboración propia. *Software Microsoft Power BI.*

6. DISCUSIÓN DE RESULTADOS

6.1 Arquitectura del prototipo

Entre las características más sobresalientes del prototipo se resalta la aplicación de las diversas herramientas de software libre como el IDE de Android Studio para la creación de la aplicación, SQLite para gestionar los datos dentro del dispositivo móvil, el API de Google para transcribir la voz a texto o Microsoft Power BI, herramientas de alto desempeño que permiten al prototipo gestionar, almacenar y analizar la información con un alto grado de confiabilidad. Estas herramientas utilizadas comúnmente en el mercado tecnológico por su confiabilidad y alto rendimiento dan un *plus* de eficiencia a la solución y no suponen una carga económica alta para la institución.

6.1.1. Síncrono y asíncrono

La aplicación no está limitada a una conexión dedicada de internet, Android cuenta con la característica de transcripción de voz de forma nativa, basta con configurarla y descargar el idioma al que se está transcribiendo para utilizarla. Esta característica permite que se transcriba de voz a texto en cualquier lugar del país. Permite transcribir y almacenar las declaraciones en lugares geográficos donde no se cuente con señal de internet. Una vez almacenada la transcripción en la base de datos puede ser sincronizada a la base de datos central cuando se cuente con señal de internet.

6.1.2. Interfaz de usuario

Los usuarios que realizaron las pruebas se sintieron cómodos utilizando la aplicación. La interfaz de la aplicación es simple. Todos los usuarios coincidieron en que es intuitiva y no requiere de capacitación extensa para utilizarla. Además, confirman que la aplicación les daría privacidad y mayor libertad de expresión al ofrecer una declaración por un delito de crimen de índole sexual que hacerlo como se hace tradicionalmente ante una persona cualquiera.

6.2. Eficiencia de transcripción

En promedio, el resultado de las pruebas de transcripción realizadas fue de 94.41 %. De las 3865 palabras de prueba, 3649 se transcribieron de forma satisfactoria y 216 fueron erróneas. Se estableció durante las pruebas que se obtenía una mejor transcripción cuando los usuarios gestionaban mejor las palabras y la velocidad del habla era normal, cuando el usuario hablaba muy rápido se incrementaba el número de errores en la transcripción.

6.2.1. Tiempo de transcripción

Las mediciones realizadas indican que transcribir de forma manual incrementa en un 78.62 % el tiempo del proceso comparado con la aplicación que transcribe en tiempo real. La transcripción en la aplicación es inmediata, el tiempo que tarde la víctima en dar su declaración será el mismo que tarde la aplicación en transcribir. Se pudo establecer que, además del incremento del tiempo al realizar la transcripción manual, también influyen otros factores negativos como la interrupción de la declaración de la persona que transcribe, cuando escribe a una velocidad menor de la que la víctima hace su declaración o la incomodidad de solicitar que la víctima repita información que no se entendió

o no se escuchó. Estos factores podrían influir en una declaración incompleta de la víctima y refuerzan la viabilidad y la necesidad de optimizar las transcripciones mediante tecnología.

6.2.2. Errores en transcripción

El porcentaje de eficiencia fue alto durante las pruebas, sin embargo, en un tema tan delicado como una declaración de un delito de violencia sexual un 5.59 % de error en una transcripción podría influir negativamente al presentar la declaración en el proceso legal. Además, durante una declaración real de violencia sexual, la víctima está alterada y sería muy difícil que siguiera instrucciones de cómo gestionar adecuadamente sus palabras, hacerlo a una velocidad normal o hablar claramente. Esto incrementaría los errores en la transcripción de la declaración.

6.3. Ventajas de la inteligencia de negocios

Para una institución encargada del análisis de información, creación de estrategias y el combate de cualquier tipo de crimen, es fundamental contar con herramientas que automaticen y eficienten los procesos. La inteligencia de negocios propuesta en este prototipo brinda herramientas que permitan analizar la información de forma rápida e intuitiva. Mediante la gráfica de mapa, el usuario puede definir el departamento con mayor y menor número de denuncias. Si se analiza profundamente esta misma gráfica se identifican departamentos donde no se han presentado denuncias. Esto podría indicar que en ese lugar no hay crímenes o que sus habitantes carecen de información sobre cómo presentar las denuncias.

El prototipo cuenta con las gráficas necesarias para buscar patrones dentro de las declaraciones de crímenes sexuales a través de las variables de ubicación, género y edad. Sin embargo, pone a disposición de los usuarios el data warehouse que permite al usuario de tablero de información graficar la información según se necesite y considere necesario.

6.4. Trabajos futuros para la solución

Para mejorar el prototipo en el futuro se podría considerar los siguientes puntos:

- Incluir la metodología Desing Thinking para descubrir que complementos se pueden adherir a la solución para agregarle valor.
- Realizar las modificaciones para hacer pública la aplicación en Play Store con el objetivo de que cualquiera pueda descargarla para transcribir y enviar su declaración a una institución que se encargue de darle seguimiento a estos casos que por temor a represarías las víctimas no acuden a denunciar.
- Incluir más variables en la toma de declaración como la etnia o religión de la víctima para enriquecer la información de la data warehouse y realizar más y mejores análisis de información.

CONCLUSIONES

1. Se desarrolló un prototipo mediante el uso de distintas herramientas de software que reconocen y transcriben la voz a texto en declaraciones de víctimas de crímenes sexuales con una eficiencia promedio de 94.41 % en la transcripción y una reducción promedio de 78.62 % del tiempo en la toma de la declaración comparado con una transcripción manual, la eficiencia de la transcripción se incrementa o disminuye dependiendo de diversos factores como la buena gesticulación de las palabras, la velocidad y la claridad con la que las víctimas hablen al realizar su declaración, aunque el porcentaje de eficiencia es alto, un 5.59 % de error podría influir negativamente al presentar la declaración en un proceso legal.
2. Se utilizó SQL Server 2016 para crear y gestionar la base de datos del prototipo para transcripción de denuncias de víctimas de crímenes sexuales, debido a que es una herramienta robusta, ha evolucionado durante treinta y dos años y es utilizada mundialmente en el ámbito empresarial ya que brinda seguridad, integridad y disponibilidad.
3. Se creó la estructura de la data warehouse que consta de una única tabla en la cual se centraliza, simplifica, limpia e integra la información de las declaraciones de víctimas de crímenes sexuales obtenidas a través del prototipo para analizar las tendencias de los crímenes sexuales.
4. Para el análisis de la información obtenida por la aplicación, se creó un proyecto de inteligencia de negocios que permite el análisis y búsqueda de patrones, el data warehouse se concentra en tres variables específicas,

ubicación; departamento y municipio con mayor y menor índice de crímenes sexuales, género de las víctimas más o menos afectadas y la variable edad para determinar que edades son las más vulnerables ante estos crímenes.

5. Se desarrolló una arquitectura móvil para el procesamiento de voz, transcripción de denuncias, almacenamiento y análisis de información, para la transcripción de voz a texto se utiliza como base el API Speech to Text de Google que aplica algoritmos de redes neuronales artificiales para el reconocimiento de voz automático, el prototipo se puede utilizar de forma síncrona o asíncrona, si al tomar la declaración el dispositivo no cuenta con señal de internet se puede realizar la transcripción y luego sincronizar la información a la base de datos central que alimenta el data warehouse para su análisis y búsqueda de patrones a través de la herramienta Microsoft Power BI.

RECOMENDACIONES

1. Realizar una grabación de audio de forma paralela, en el caso de que la aplicación no transcriba correctamente se tendría un plan de contingencia.
2. Establecer e indicar los lineamientos que deben seguir la víctima durante su declaración con el objetivo de que la transcripción sea lo más cercano posible a la realidad.
3. Tomar la declaración en un lugar libre de ruido que interfiera con la voz de la víctima al momento de hablar, con esto se lograra capturar mejor la voz y realizar una mejor transcripción.
4. Usar un dispositivo móvil con 2 GB de memoria RAM, Android 10 o superior, 250 MB libres de almacenamiento y un procesador con una velocidad mínima de 1.8GHz.
5. Incluir más variables como religión o etnia dentro de la toma de la declaración con el objetivo de enriquecer los resultados del data warehouse y el análisis de patrones dentro de la información.

REFERENCIAS

1. Aguilar, L. J. (2003). *Fundamentos de Programación Libro de Problemas* (Segunda ed.). (C. F. Madrid, Ed.) España: McGRAW-HILL.
2. Areiza, M., Palacios, J., Castillo, A., & Mendoza V., J. (2018). Entrenamiento de Modelo de Voz para Hablantes Hispanos Aplicando Redes Neuronales Convolucionales. *III Congreso Internacional de Ciencia y Tecnología para el Desarrollo Sostenible*, (págs. 47-51). Chiriqui.
3. Barrios, K., López, J., Mendieta, S., Benavidez, R., & Sáez, Y. (2018). Sistema de reconocimiento de voz: un enlace en la comunicación. *Revista Academica UTP*, 4(Especial), 92-95.
4. Carrillo Osorio, M. (2018). *Inteligencia en los negocios Business Intelligence*. Tesis de Maestria, Instituto Tecnológico de Orizaba, Mexico.
5. Cruz Beltrán, L. A., & Acevedo Mosqueda, M. A. (2009). Aplicación del Reconocimiento de Voz de un Hablante Mediante una Red Neuronal Artificial Backpropagation y Coeficientes LPC sobre un Canal Telefónico. *Programación Matemática y Software*, 1(1), 111-128.
6. Date, C. (2001). *Introducción a los Sistemas de Bases de Datos* (Septima ed.). (I. a. Datos, Ed.) México: Pearson Educación.

7. Esparza Arellano, M. E., & Avalos Briseño, J. B. (2003). Reconocimiento de voz. (22).
8. Francisco, V., Gervás, P., & Hervás, R. (1 de Junio de 2005). Análisis y síntesis de expresión emocional en cuentos leídos en voz alta. *Procesamiento del Lenguaje Natural*(35), 293-300.
9. Gabillaud, J. (2015). *SQL SERVER 2014 Administración de una base de datos transaccional con SQL Server Management Studio*. (A. M. Conejo, Ed.) Barcelona, España: Ferrocarriles Catalanes.
10. Gómez Dúran, J., Simancas García, J., Acosta Coll, M., Meléndez Pertuz, F., & Vélez Zapata, J. (2017). Algoritmo de recocimiento de comandos voz basado en técnicas no-lineales. *Espacios*, 38(17).
11. Gonzáles, J. R., & Hernando, V. J. (1995). *Redes Neuronales Artificiales Fundamentos, Modelos y Aplicaciones*. Madrid, España: Ra-Ma.
12. Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation* (Segunda ed.). Mexico: Prentice-Hall, Inc.
13. IBM. (2019). *www.ibm.com*. Obtenido de https://www.ibm.com/support/knowledgecenter/es/SS3RA7_sub/modeler_main
14. Jensen, C. T. (2016). *APIs Para Dummies* (Limitada IBM ed.). (C. A. Johnson, & K. Simpson, Edits.) Georgia, U.S.A: Willey.

15. Joyanes Aguilar, L. (2008). *Fundamentos de Programacion: Algoritmos, estructura de datos y objetos* (Cuarta ed.). (J. L. García, & C. Sánchez , Edits.) Madrid, España: McGRAW-HILL.
16. Kreibich, J. A. (2010). *Using SQLite*. (M. Loukides, Ed.) Estados Unidos: O'REILLY.
17. L.J. Gil, L. C. (marzo de 2016). Reconocimiento de comandos de voz en español orientado al control de una silla de ruedas. 48.
18. Montoto, A., Alvarez, E., Chavira, G., Orozco, J., Quiroga, G., & Nava, S. (2018). El Reconocimiento de Voz como alternativa de inclusión para discapacitados auditivos en un entorno educativo. *Revista Iberoamericana de Ciencias*, 5(3), 139-147.
19. Opel, A., & Sheldon, R. (2009). *Fundamentos de SQL* (Tercera ed.). D.F., México: McGRAW-HILL.
20. Oropeza, J. L., & Suárez Guerra, S. (2006). Algoritmos y Métodos para el Reconocimiento de Voz en Español Mediante Sílabas. *Computación y Sistemas*, 9(3).
21. Palma Méndez, J. T., & Marín Morales, R. (2008). *Inteligencia Artificial: Técnicas, métodos y aplicaciones* (Primera ed.). Madrid, España: McGraw Hill.
22. Pérez, J., Javier, D., & Harari, I. (9 al 13 de Octubre de 2017). Ejecución de comandos de voz mediante Web Speech API. *Congreso Argentino de Ciencias de la Computación*, 964-970.

23. Prieto Blázquez, J., Ramírez Vique, R., Morillo Pozo, J., & Domingo Prieto, M. (2011). *Tecnología y desarrollo en dispositivos móviles* (Primera ed.). Barcelona: Eureka Media, SL.
24. Rainardi, V. (2008). *Building a Data Warehouse With Examples in SQL Server* (Vol. 1). (J. Pepper, Ed.) Estados Unidos: Apress.
25. San Juan, E., Jamett, M., Kaschel, H., & Sánchez, L. (2016). Sistema de reconocimiento de voz mediante wavelets, predicción lineal y redes backpropagation. *Revista chilena de ingeniería*, 24(1), 8-17.
26. Shapiro, J. R. (2002). *Visual Basic .NET the Complete Reference*. Berkeley, California, U.S.A: McGraw-Hill.
27. Sommerville, I. (2011). *Ingeniería de Software* (Novena ed.). (L. M. Cruz Castillo, Ed.) México, México: Pearson Educación.
28. Srinivasan, V., Ravi, J., & Raj, J. (2018). *Google Cloud Platform for Architects Design and manage powerful cloud solutions*. (V. Boricha, Ed.) Birmingham, Reino Unido: Packt.
29. Vargas Sandoval, V. (2019). *Reconocimiento de Emociones por Medio de Voz*. Tesis Maestría, Universidad Politécnica de Madrid, Madrid.