

Sistema Inteligente para la Identificación y Seguimiento de Derechos de los NNA por Feminicidios

Trabajo Terminal No.-----

Alumnos: Herrera Ramírez Emilio Alejandro*, Morales Martínez Héctor Alberto**

Directores: Saldana Velez Ulises

e-mail: eherrerar2103@alumno.ipn.mx*, hmoralesm1500@alumno.ipn.mx**

Resumen - Este proyecto propone el desarrollo de un sistema inteligente para identificar y dar seguimiento a niños, niñas y adolescentes (NNA) que quedan en situación de orfandad como consecuencia de feminicidios en México. La problemática central radica en que estos NNA, a pesar de ser víctimas directas, permanecen invisibilizados en los procesos judiciales y carecen de sistemas de seguimiento que garanticen la restitución de sus derechos. Mediante la implementación de Web Scraping y Procesamiento de Lenguaje Natural (PLN), el sistema automatiza la recolección, clasificación y análisis de información dispersa en fuentes públicas sobre casos de feminicidio, permitiendo identificar a las víctimas indirectas. La solución propuesta busca centralizar datos actualmente fragmentados para facilitar la creación de políticas públicas y fortalecer las acciones de protección dirigidas por organizaciones como la Fundación Futuro con Derechos.

Palabras clave - Inteligencia Artificial, Procesamiento del Lenguaje Natural, Protección infantil, Web Scraping.

1. Introducción

La violencia de género en México ha sido una crisis persistente a lo largo del tiempo y hoy en día es una problemática alarmante que ha terminado con la vida de miles de mujeres. Según datos del Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública, en 2023 se registraron 848 víctimas de feminicidio y 2,591 homicidios dolosos en contra de las mujeres, lo que deja una notable tendencia de inseguridad que presentan las mujeres en el país. [1]

Sin restar importancia, además del impacto existente sobre las víctimas, las violencias ocurridas también generan consecuencias para las familias de los afectados, especialmente para los niños, niñas y adolescentes (NNA), ya que estos quedan expuestos a entrar en un estado de orfandad. Según los datos ofrecidos por el Observatorio Ciudadano Nacional Contra el feminicidio, se estima que 8 de cada 10 víctimas dejan al menos un hijo o hija en situación de desamparo[2]. Además, es este sector que al ser una víctima indirecta, sufre de diferentes tipos de problemas, como lo pueden ser: la pérdida de vínculos o núcleos familiares, abandono del sistema escolar, movilidad forzada y el constante riesgo de institucionalización que no asegura el cuidado de los derechos.

Sin embargo, los NNA, a pesar de ser víctimas directas, por lo regular son víctimas invisibilizada dentro de los procesos judiciales, ya que no existen datos oficiales que los identifiquen, y mucho menos sistemas de seguimiento que garanticen la restitución de sus derechos, y en caso de que exista dicha información, esta suele presentarse de forma dispersa, incompleta y no estructuradas, lo que dificulta su análisis. Esta falta de información dificulta el trabajo de organización y creación de políticas públicas dedicadas a la restitución de derechos.

Es por esto, que el presente Trabajo Terminal se centra en la complejidad derivada de identificar y ubicar a los NNA víctimas para desarrollar estrategias en pro de la restitución de sus derechos. Para lograr este objetivo proponemos el desarrollo de un sistema inteligente que automatice la recolección, clasificación y análisis de información relacionados con los casos de feminicidio y el impacto generado en NNA. En este sentido hicimos una búsqueda de sistemas similares los cuales se muestran en la tabla 1, en la cual podemos ver...

Software	Adquisición de datos automatizada	Análisis de datos	Enfoque en NNA	Visualización de datos
----------	-----------------------------------	-------------------	----------------	------------------------

Mapa de feminicidios en México [5]	✗	✗	✗	✓
SIESVIM (INEGI)[6]	✗	✓	✗	✓
Solución propuesta (TT)	✓	✓	✓	✓

Tabla 1. Resumen de productos similares.

Es por esto que el presente Trabajo Terminal propone el desarrollo de un sistema que funcione con base en técnicas de Web Scraping, procesamiento de lenguaje natural y análisis de datos, buscando extraer información relevante sobre diversas fuentes de información pública. Con el fin de darle visualización a aquellas víctimas indirectas a causa de los feminicidios (NNA) para poder fortalecer acciones de protección, atención y prevención.

2. Objetivo

Desarrollar un sistema inteligente que sea capaz de automatizar la recolección, clasificación y análisis de información relacionadas con casos de feminicidio, para facilitar la identificación de niños, niñas y adolescentes que se encuentren en una situación de orfandad en consecuencia de los crímenes de violencia. Mediante el uso de Web Scraping y Procesamiento de Lenguaje Natural..

Objetivos específicos:

- Identificar y seleccionar fuentes públicas confiables que contengan información relacionada con feminicidios y posibles víctimas indirectas.
- Diseñar y desarrollar una herramienta automatizada de Web Scraping para la recolección sistemática de datos provenientes de dichas fuentes.
- Implementar técnicas de Procesamiento de Lenguaje Natural (PLN) para extraer y estructurar información clave a partir de textos no estructurados.
- Diseñar una base de datos estructurada que permita almacenar la información recolectada de manera eficiente y segura.

3. Justificación

Los feminicidios son una problemática grave y persistente que afecta a múltiples niveles de la sociedad. Si bien los casos de feminicidio reciben atención mediática y social, las víctimas indirectas, como las niñas, niños y adolescentes (NNA) que quedan en estado de orfandad, suelen ser invisibilizadas, lo que dificulta su atención, protección y restitución de derechos. Este vacío informativo y de seguimiento es un obstáculo para las organizaciones dedicadas a apoyar a esta población vulnerable y para la creación de políticas públicas adecuadas.

La solución propuesta es el desarrollo de una plataforma inteligente que automatice la recolección, clasificación y análisis de datos relacionados con feminicidios y su impacto en los NNA. A través de tecnologías avanzadas como Web Scraping[3], Procesamiento de Lenguaje Natural (NLP)[4] y análisis de datos, se busca centralizar y procesar información para identificar a las víctimas indirectas, facilitar su seguimiento, y apoyar la creación de políticas públicas y estrategias de intervención.

Originalidad del Trabajo.

- Este proyecto combina tecnologías emergentes como NLP y análisis automatizado de datos para abordar una problemática social subatendida.
- Aporta una herramienta innovadora que no solo centraliza información dispersa, sino que también permite analizarla de manera automática y eficiente, algo que no se realiza actualmente de forma sistemática en el contexto de feminicidios y NNA.
- Propone un enfoque interdisciplinario al combinar aspectos tecnológicos, sociales y legales en una solución integral.

Vinculación con los usuarios potenciales.

Usuarios Directos. La Fundación Futuro con Derechos y otras organizaciones dedicadas a la protección y restitución de derechos de los NNA afectados.

Instituciones gubernamentales encargadas de diseñar políticas públicas relacionadas con feminicidios y protección infantil.

Usuarios Indirectos: Los NNA en estado de orfandad por feminicidios, quienes se beneficiarán de un mejor acceso a programas de atención y protección.

La sociedad en general, al contribuir a la sensibilización y visibilización de esta problemática.

Sector Académico y Tecnológico: Este proyecto puede servir como referencia para futuras investigaciones y desarrollos tecnológicos con enfoque social.

Mejora a lo ya existente.

- Actualmente, la información relacionada con feminicidios y sus impactos en NNA está dispersa en múltiples fuentes y no se encuentra sistematizada.
- Este proyecto propone una mejora significativa al automatizar la recolección y análisis de datos, reduciendo el tiempo y esfuerzo manual necesario.
- Facilitará la generación de reportes detallados y visualizaciones que permitan comprender mejor el impacto de los feminicidios en los NNA.
- Ofrecerá datos confiables y centralizados que pueden ser utilizados para diseñar políticas públicas más efectivas y oportunas.

Complejidad.

1. El proyecto requiere la aplicación de conocimientos avanzados en varias áreas:

- Desarrollo de software: Diseño de sistemas de scraping y procesamiento masivo de datos.
- Inteligencia Artificial: Implementación de modelos de NLP para extraer y clasificar información relevante.
- Análisis de datos: Limpieza, procesamiento y visualización de datos para generar reportes útiles.
- Aspectos legales y éticos: Consideración de aspectos como la privacidad y el cumplimiento de normativas sobre uso de datos.

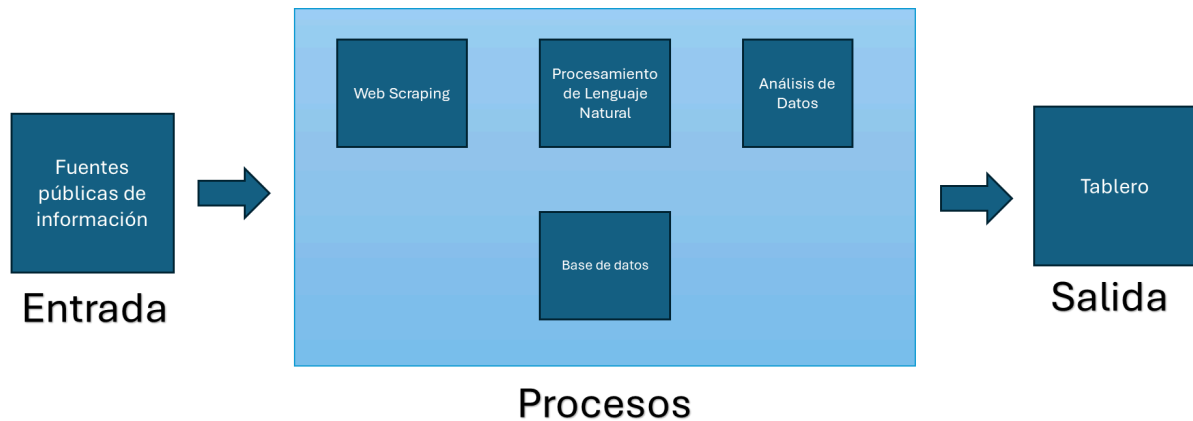
2. Refleja un alto nivel de interdisciplinariedad, ya que implica la colaboración entre especialistas en tecnología, derechos humanos y trabajo social.

3. Su desarrollo requiere una planificación rigurosa para asegurar la calidad, escalabilidad y sostenibilidad del sistema.

4. Productos o Resultados esperados

Al momento de finalizar el trabajo terminal se espera tener los siguientes resultados:

1. Software funcional basado en Web Scraping y técnicas de Procesamiento de Lenguaje Natural que automatice la recolección de información de diversas fuentes públicas
2. Base de datos estructurados sobre los casos de NNA en situación de orfandad.
3. Conexión a al menos tres fuentes de información pública de distintos sectores.



Junto con el trabajo terminal, se espera entregar los siguientes apartados:

- Documentación técnica de la elaboración del sistema.
- Pruebas y evaluación del sistema.

5. Metodología

Para llevar a cabo el sistema se utilizará una metodología evolutiva, el cual permite el avance progresivo del software mediante ciclos iterativos. Consideramos que este tipo de metodología, ya que los requerimientos pueden refinarse y expandirse mediante se avanza en el análisis de datos y la implementación técnica.

El desarrollo consta en varias fases evolutivas, las cuales terminaran con un prototipo funcional que implementa mejoras con respecto a la versión anterior. Esto con el fin de validar el resultado y retroalimentar el proceso.

Fases de desarrollo:

1. Análisis inicial:
 - Recolección de requerimientos.
 - Definición del conjunto inicial de fuentes de datos (noticieros digitales, portales oficiales, etc.).
 - Pruebas de concepto.
 - Investigación de las tecnologías.
2. Desarrollo del primer prototipo “Recolector inicial”
 - Diseño e implementación de herramientas básicas de web scraping sobre un conjunto de fuentes.
 - Extracción y almacenamiento de información estructurada.
 - Visualización preliminar de los casos detectados.
 - Validación y retroalimentación.
 - Ajustes a los métodos de adquisición de datos.
3. Desarrollo del segundo prototipo “Organizador de datos”
 - Diseño e implementación de herramientas básicas de web scraping sobre un conjunto de fuentes.
 - Extracción y almacenamiento de información estructurada.

- Visualización preliminar de los casos detectados.
 - Validación y retroalimentación.
 - Ajustes a los métodos de adquisición de datos.
4. Desarrollo del tercer prototipo “Análisis de datos”
 - Diseño e implementación de herramientas básicas de web scraping sobre un conjunto de fuentes.
 - Extracción y almacenamiento de información estructurada.
 - Visualización preliminar de los casos detectados.
 - Validación y retroalimentación.
 - Ajustes a los métodos de adquisición de datos.
 5. Desarrollo del cuarto prototipo “Detector de víctimas indirectas”
 - Diseño e implementación de herramientas básicas de web scraping sobre un conjunto de fuentes.
 - Extracción y almacenamiento de información estructurada.
 - Visualización preliminar de los casos detectados.
 - Validación y retroalimentación.
 - Ajustes a los métodos de adquisición de datos.
 6. Ampliación de cobertura y mejora de procesamiento:
 - Integración de técnicas de PLN para clasificación automática de textos y detección de menciones a víctimas indirectas (NNA).
 - Mejoras en la calidad y limpieza de datos.
 7. Consolidación del proyecto:
 - Diseño de una interfaz interactiva.
 - Almacenamiento en una base de datos.
 - Preparación de informes analíticos y visuales.

También se tendrán en cuenta los recursos materiales necesario para el correcto desarrollo del proyectos:

- Computadora portátil o de escritorio con acceso a internet.
- Entorno de desarrollo
- Librerías de Scraping
- Librerías de análisis
- Herramientas de visualización
- Acceso a fuentes digitales
- Almacenamiento

Sabemos que existen riesgo potenciales durante la realización del proyecto, y son los siguientes:

Riesgo	Impacto	Probabilidad	Plan de mitigación
Cambios en la estructura de sitios web	Medio - Alto	Alta	Diseñar código modular y flexible. Mantenimiento continuo
Bloqueos de IP por scraping	Bajo - Medio	Media	Usar headers, rotación de IP, tiempos de espera.
Acceso a información sensible o imprecisa	Alto	Media	Validar fuentes; filtros de calidad; Revisión ocasional

Fallos técnicos	Medio	Media	Backups; pruebas continuas.
Poca disponibilidad de datos sobre NNA	Alto	Alta	Complementar con monitoreo social o medios alternativos
Cuestiones legales/éticas por manejo de datos	Alto	Alta	Aplicar principios de protección de datos, anonimización y filtros.

Tabla 1. Resumen de productos similares.

6. Cronograma

Nombre del alumno(a): Herrera Ramírez Emilio Alejandro TT No. -----

Título del TT: Sistema Inteligente para la Identificación y Seguimiento de Derechos de los NNA por Femicidios.

FASES	SUBFASES	Ago	Sep	Oct	Nov	Dic	Ene	Feb	Mar	Abr	May	Jun
FASE DE INVESTIGACIÓN PRELIMINAR	Análisis, investigación y diseño del sistema.											
	Definición de objetivos y alcance											
	Elaboración del estado de arte											
	Identificación de requerimientos funcionales y no funcionales.											
DEFINICIÓN DE LOS REQUERIMIENTOS DEL SISTEMAS	Definición de casos de uso y modelado de datos.											
	Diseño de la interfaz de Usuario (UI) en Figma.											
	Diseño de la base de datos.											
PRIMER PROTOTIPO	Desarrollo del primer prototipo ("Recolector inicial")											
	Pruebas y correcciones.											
SEGUNDO PROTOTIPO	Desarrollo del segundo prototipo ("Organizador de datos")											
	Pruebas y correcciones.											
Documentación del diseño y validación del sistema.												
Evaluación TT1												
TERCER PROTOTIPO	Desarrollo del tercer prototipo ("Analizador preliminar")											
	Pruebas y correcciones.											
CUARTO PROTOTIPO	Desarrollo del cuarto prototipo ("Detector de víctimas indirectas")											
	Pruebas y resultados.											
Evaluación TT2												

Nombre del alumno(a): Morales Martinez Hector Alberto TT No. -----

Título del TT: Sistema Inteligente para la Identificación y Seguimiento de Derechos de los NNA por Femicidios.

FASES	SUBFASES	Ago	Sep	Oct	Nov	Dic	Ene	Feb	Mar	Abr	May	Jun
FASE DE INVESTIGACIÓN PRELIMINAR	Análisis, investigación y diseño del sistema.											
	Definición de objetivos y alcance											
	Elaboración del estado de arte											
	Identificación de requerimientos funcionales y no funcionales.											
DEFINICIÓN DE LOS REQUERIMIENTOS DEL SISTEMAS	Definición de casos de uso y modelado de datos.											
	Diseño de la interfaz de Usuario (UI) en Figma.											
	Diseño de la base de datos.											
PRIMER PROTOTIPO	Desarrollo del primer prototipo ("Recolector inicial")											
	Pruebas y correcciones.											
SEGUNDO PROTOTIPO	Desarrollo del segundo prototipo ("Organizador de datos")											
	Pruebas y correcciones.											
Documentación del diseño y validación del sistema.												
Evaluación TT1												
TERCER PROTOTIPO	Desarrollo del tercer prototipo ("Analizador preliminar")											
	Pruebas y correcciones.											
CUARTO PROTOTIPO	Desarrollo del cuarto prototipo ("Detector de víctimas indirectas")											
	Pruebas y resultados.											
Evaluación TT2												

7. Referencias

- [1] ONU Mujeres, "Las huellas de los feminicidios en CDMX", *ONU Mujeres – México*, 11 de marzo de 2024. [En línea]. Disponible en: <https://mexico.unwomen.org/es/stories/reportaje/2024/03/huellas>
- [2] Reforma, "7 de cada 10 casos de feminicidios en México dejan hijos e hijas en la orfandad", *Vanguardia*, 9 de marzo de 2025. [En línea]. Disponible en: <https://vanguardia.com.mx/noticias/7-de-cada-10-casos-de-feminicidios-en-mexico-dejan-hijos-e-hijas-en-la-orfandad-CH15216721>
- [3] Munzert, S., Rubba, C., Meißner, P., y Nyhuis, D. Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining. John Wiley & Sons, 2015. ISBN: 978-1118834817.
- [4] Bird, S., Klein, E., y Loper, E. Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit. 2ª edición, O'Reilly Media, 2015. ISBN: 978-1491913420.
- [5] "Mapa de feminicidios en México", [En línea]. Disponible en: https://www.google.com/maps/d/u/0/viewer?mid=174IjBzP-fl_6wpRHg5pkGSj2egE&ll=14.236867409759748%2C-84.98095331187768&z=4
- [6] Instituto Nacional de Estadística y Geografía (INEGI), "Sistema Integrado de Estadísticas sobre Violencia contra las Mujeres (SIESVIM)", [En línea]. Disponible en: <https://sc.inegi.org.mx/SIESVIM1/paginas/inicio.jsf>

8. Alumnos y directores

Herrera Ramirez Emilio Alejandro. - Alumno de la carrera de Ingeniería en Inteligencia Artificial en ESCOM, Boleta:2022690099, Tel: 8661495505, email: eherrerar2103@alumno.ipn.mx

Firma: _____

Morales Martinez Héctor Alberto. - Alumno de la carrera de Ingeniería en Sistemas Computacionales en ESCOM, Boleta:2021630457, Tel: 5616148167, email: hmoralesm1500@alumno.ipn.mx

Firma: _____

Velez Saldaña Ulises.- Profesor asociado C en ESCOM con estudios en Maestría en Ciencias de la Computación, especialista en Ingeniería de Software y desarrollo en aplicaciones móviles, Tel 5557296000 Ext. 25032, email: uvelez@ipn.mx

Firma: _____

Hurtado Avilés Gabriel. - Maestrante en Docencia para la Educación Media Superior (FES Acatlán) y en Ciencias de la Computación (UAM-A), Ing. en Computación (UAM-A), Profesor en ESCOM/IPN (Dpto de CIC). Áreas de Interés: Sistemas Inteligentes y Manejo de la Información, Visualización de la Información, Tecnologías Educativas, Gamificación y Desarrollo de Aplicaciones Móviles. Tel. 5550696612, email: ghurtadoa@ipn.mx

Firma: _____

CARÁCTER: Confidencial
FUNDAMENTO LEGAL: Artículo 11 Fracc. V y Artículos 108, 113 y 117 de la Ley Federal de Transparencia y Acceso a la Información Pública.
PARTES CONFIDENCIALES: Número de boleta y teléfono.