ALERTA-LINK: Sistema de Análisis Forense Automático que, ante una URL/domino, genere un score interpretable de riesgo de mishing acompañado de evidencia técnica reproducible y mecanismos de explicabilidad

Cristia Salazar

Samuel Ortiz Ospina

Juan Stiven Castro

Universidad Manuela Beltrán

Ingeniería de Software

Facultad de Ingenierías

Bogotá, Colombia

2025

Tabla de Contenido

1 Introducción

1.1 Problematización

Introducción:

En Colombia se ha inténsificado el smishing (engaños por SMS) (*¿Qué Es El Smishing y Cómo Protegerse de Esta Estafa Por Mensajes de Texto?*, n.d.) que se vale de la identidad visual y el nombre de entidades reconocidas (bancos, DIAN, operadores y comercios electrónicos) para convencer a las personas de abrir enlaces falsos. Detrás de esos enlaces generalmente hay páginas web qué imitan portales reales para robar contraseñas, instalar software malicioso o forzar pagos no autorizados. El impacto es mayor en quienes usan el teléfono como su principal medio de conexión y no cuentan con referentes claros para distinguir mensajes legítimos de los falsos; sin embargo, cualquier usuario puede ser víctima mediante técnicas de ingenieria social cuando el mensaje es convincente y está bajo presión de tiempo.

Brecha y necesidad: A pesar de la frecuencia de estos casos, no hay una herramienta local, gratuita y en español que permita a cualquier persona pegar el enlace recibido y obtener, en ese momento, un puntaje de riesgo con explicaciones comprensibles y pasos concretos sobre qué hacer. Las recomendaciones generales (por ejemplo, “no abra enlaces desconocidos”) ayudan, pero no resuelven el momento crítico: cuando el usuario ya tiene el mensaje en la pantalla y debe decidir. Asimismo, falta un mecanismo de reporte simple que, con el permiso de la persona, ayude a identificar patrones y a mantener listados de enlaces peligrosos que puedan ser útiles para entidades públicas y privadas.

Propuesta de solución: Se propone diseñar y evaluar una aplicación liviana (web o móvil) en la que el usuario ingresa de forma manual el enlace sospechoso y recibe de inmediato:

1. Un puntaje de riesgo (por ejemplo, bajo/medio/alto).
2. Una explicación breve de las señales detectadas, en lenguaje sencillo.
3. Recomendaciones accionables (“qué hacer ahora”) ajustadas al nivel de riesgo.
4. Opción de reporte anónimo y voluntario para alimentar listados de enlaces peligrosos y mejorar la protección de otras personas.

El análisis se ejecuta solo cuando el usuario lo solicita (no hay revisión automática de sus mensajes). La aplicación priorizará la privacidad (revisando primero en el propio dispositivo y compartiendo solo lo estrictamente necesario cuando la persona lo autorice) y la accesibilidad (pantallas claras, textos cortos, uso de indicadores tipo "semáforo").

Pertinencia y beneficios: La propuesta es pertinente porque actúa exactamente donde ocurre el riesgo: antes de abrir el enlace. Aporta valor en cuatro frentes:

1. Protección inmediata: ayuda a detener clics peligrosos mediante avisos claros y acciones simples.
2. Educación práctica: enseña a reconocer señales de fraude sin tecnicismos, fortaleciendo hábitos seguros.
3. Colaboración ciudadana: el reporte voluntario permite mejorar listados de riesgo útiles para actores públicos y privados.
4. Respeto por la privacidad: el diseño prioriza que la revisión ocurra en el dispositivo y que cualquier envío de datos sea opcional, mínimo y transparente.

El prototipo no leerá ni filtrará automáticamente los mensajes del teléfono; actuará solo cuando el usuario pegue el enlace. La primera versión se enfocará en enlaces compartidos por SMS o mensajería móvil, sin interferir con otras aplicaciones. La propuesta es escalable: podrá mejorar sus señales de alerta y sus listados con el tiempo, conservar la sencillez de uso y mantenerse gratuita y en español para favorecer su adopción.

Justificación:

El aumento del *smishing* en Colombia representa un riesgo creciente para los usuarios, especialmente aquellos que dependen del celular como su principal medio de conexión. Actualmente no existe una herramienta local, gratuita y en español que ofrezca un análisis inmediato y comprensible de enlaces sospechosos. La propuesta de una aplicación sencilla, que brinde puntajes de riesgo, explicaciones claras y recomendaciones prácticas, es necesaria porque actúa justo en el momento crítico de la decisión. Además, promueve la educación digital, el reporte ciudadano y la protección de la privacidad, generando un impacto positivo tanto individual como colectivo.

1.2 Bases Teóricas

En esta sección se presentan los fundamentos y antecedentes del proyecto de investigación, con el propósito de brindar una comprensión clara de su punto de partida y de los conceptos que servirán de base para su desarrollo.

Estado del Arte:

Titulo	Smishing Detector: A security model to detect smishing through SMS content analysis and URL behavior analysis
Tipo de documento (artículo, tesis, conferencia, etc.)	Tesis
Editorial	Jaypee Institute of Information Technology, Sector-128, Noida, India
Palabras clave	Smishing Phishing Text messaging Mobile security Machine learning SMS
Año	2020
Objetivo del estudio (redacción propia)	La tesis busca construir un sistema que permita distinguir mensajes reales de mensajes tipo SMISHING
Conclusiones del estudio (redacción propia)	El sistema es bastante completo al descomprimir una url que pueda estar contenida en un mensaje de texto reconociendo si esta acortado el link, si esta en alguna blacklist o si es demasiado reciente el dominio, el sistema propuesto no es una aplicacion sino un sistema paso a paso de como deberia de funcionar una aplicacion para detectar smishing
Principales limitaciones (oportunidad de mejora)	La principal limitacion es la implementacion en un sistema funcional más alla de la propuesta de sistema
Enlace	https://www.sciencedirect.com/science/article/abs/pii/S0167739X19318758 (Mishra & Soni, 2020)

Titulo	A Review of Smishing Attaks Mitigation Strategies
Tipo de documento (artículo, tesis, conferencia, etc.)	PAPER
Editorial	International Journal of Computer and Information Technology
Palabras clave	phishing; Social engineering; vishing; SMS; malware; mobile applications; awareness
Año	2022
Objetivo del estudio (redacción propia)	El paper se centra en informar sobre que es un smishing, como funciona y cuales son los peligros asociados a esos ataques teniendo en cuenta la actualidad de la tecnologia
Conclusiones del estudio (redacción propia)	El estudio concluye con que a pesar de que existen herramientas muy eficaces para detectar spam o sitios que se encuentren de alguna manera ya marcados como maliciosos aún así se necesita más investigacion con el fin de encontrar una solucion más potente contra el smishing
Principales limitaciones (oportunidad de mejora)	La principal mejora que se evidencia es que se menciona la necesidad de herramientas más eficaces para combatir el smishing
Enlace	**https://core.ac.uk/download/pdf/525063994.pdf**

Titulo	DSmishSMS-A System to Detect Smishing SMS
Tipo de documento (artículo, tesis, conferencia, etc.)	Tesis
Editorial	Department of Computer Science & Engineering and Information Technology, Jaypee Institute of Information Technology, Sector-128, Noida, India
Palabras clave	Smishing Phishing Paytm SMS scam Mobile security Machine learning Backpropagation Algorithm Cyber security Covid-19 SMS Scam
Año	2021

Objetivo del estudio (redacción propia)	La tesis se trata de la creación de un sistema para detectar sitios maliciosos que puedan redirigir a un smishing
Conclusiones del estudio (redacción propia)	El sistema descrito se basa en la separacion de los mensajes en 2 etapas: deteccion de mensajes de texto y deteccion de los url dentro de los mensajes, este enfoque es bastante util a la hora de diseñar adecuadamente un sistema que pueda servir para combatir el smishing
Principales limitaciones (oportunidad de mejora)	El estudio menciona que a pesar de usar tecnicas ya normalmente usadas sigue siendo un reto importante detectar y clasificar adecuadamente todos los mensajes especialmente cuando se trata de mensajes de ingenieria social muy precisos cuya firma digital es dificil de clasificar como legitima o de un atacante
Enlace	**https://link.springer.com/article/10.1007/s00521-021-06305-y**

Titulo	Phishing, SMiShing & Vishing: An Assessment of Threats against Mobile Devices
Tipo de documento (artículo, tesis, conferencia, etc.)	PAPER
Editorial	Journal of Emerging Trends in Computing and Information Sciences
Palabras clave	Phishing, SMiShing, Vishing, Mobile Devices, Attacks, Cyber-victims, Social Engineering
Año	2014
Objetivo del estudio (redacción propia)	Este paper tiene la finalidad de explorar los phishings, smishings y otras tecnicas de scam digitales en dispositivos móviles, ademas de darle al usuario consejos y herramientas para evitar ser victima
Conclusiones del estudio (redacción propia)	El estudio llega a la conclusion a travez de varias encuestas que realiza de que los usuarios debido a su desconocimiento de todo el peligro en linea y de las amenazas existentes se encuentran demasiados confiados en los entornos digitales
Principales limitaciones (oportunidad de mejora)	Hace falta mucha más educacion para los usuarios de dispositivos moviles respecto a los riesgos que conlleva abrir links en sus dispositivos

Enlace	https://d1wqtxts1xzle7.cloudfront.net/33996443/Phishing__SMiShing___Vishing_-CISjournal_-_vol5no4_6-libre.pdf?1403260068=&response-content-disposition=inline%3B+filename%3DJournal_of_Emerging_Trends_in_Computing.pdf&Expires=1757449047&Signature=GFaerJNXqzj8ANA885GHdD5VnS6GaM4Y1oFhCt2oF14BuK7mm9mJnNmy5csb0MZ5yxVGLsQllL6Ovv~TXKNi1i2UqRjK64j7QglgEs6~btvg99V4eFv7r0Zsh-fNz1o8lv~av~AjpdpyBb7zfVXOI2YhkYKfcbLVRTYuuqWLCMywvzwOMe4PjWQ87VsIN9-ywLO6ZEc63q~XK0-XmqJ6NXHx2GRV45qa~z9YzMOmPIMuQwls-ihQFMYv9Iyn~CnBMtihcvp8snGSBR6DKuKQP7n5h1EQ554PmVEWA~IS1kZjB1QRFp8KL0tdsIv9cnqfxM5i5qbSpJ3kRvBROKYFyA__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA

Titulo	SmiDCA: An Anti-Smishing Model with Machine Learning Approach
Tipo de documento (artículo, tesis, conferencia, etc.)	TESIS
Editorial	The Computer Journal, Volume 61, Issue 8, August 2018, Pages 1143–1157
Palabras clave	(Sin apartado dedicado a keywords)
Año	2018
Objetivo del estudio (redacción propia)	El estudio presenta un modelo para reconocer smishing usando inteligencia artifical aplicando un sistema que reconoce hasta 30 propiedades distintas en un mensaje
Conclusiones del estudio (redacción propia)	El estudio es bastante optimista con su modelo de reconocimiento de smishing destacando que tiene una capacidad de acierto superior al 90% incluso reduciendo sus propiedades analizadas en un 50%
Principales limitaciones (oportunidad de mejora)	El estudio al utilizar machine learning tiene un enfoque unicamente en inglés, se requiere más estudios y más informacion en otros idiomas
Enlace	**https://academic.oup.com/comjnl/article/61/8/1143/4985552#119108454**

Titulo	S-Detector: an enhanced security model for detecting Smishing attack for mobile computing
Tipo de documento (artículo, tesis, conferencia, etc.)	Articulo
Editorial	Springer Science+Business Media New York 2017
Palabras clave	Phishing · Smishing · Malware · Mobile security
Año	2017
Objetivo del estudio (redacción propia)	El articulo explora los peligros de la computacion movil, proponiendo como solucion al smishing un modelo llamado S-detector que analiza estadisticamente el mensaje para detectar riesgos
Conclusiones del estudio (redacción propia)	El estudio da como muy positivos los resultados del S-Detector, aunque implica que en el futuro los ataques serán cada vez más inteligentes por lo que se requiere una seguridad mayor
Principales limitaciones (oportunidad de mejora)	El sistema de S-Detector hace un analisis unicamente sobre el mensaje de texto, no sobre los links, remitentes ni contenido perse, se requiere un análisis más exhaustivo a futuro teniendo en cuenta las propias recomendaciones del estudio
Enlace	**https://link.springer.com/article/10.1007/s11235-016-0269-9**

Titulo	Smishing-Classifier: A Novel Framework for Detection of Smishing Attack in Mobile Environment
Tipo de documento (artículo, tesis, conferencia, etc.)	Articulo
Editorial	Communications in Computer and Information Science ((CCIS,volume 828))
Palabras clave	Smishing Short messaging service Mobile phishing Machine learning

Año	2018
Objetivo del estudio (redacción propia)	El estudio propone un sistema de analisis clasificador Naive Bayes donde analiza los mensajes en busca de patrones de smishing comunmente usados
Conclusiones del estudio (redacción propia)	El estudio concluye que hacer detecciones de smishing es demasiado complejo con el metodo utilizado, se requiere convertir jerga en lenguaje tradicional y detectar a partir de ahi lo que genera errores
Principales limitaciones (oportunidad de mejora)	A pesar del analisis que realiza el modelo propuesto se requiere hacer mejoras al detectar metodos de ofuscacion en las url y con el uso de emojis
Enlace	**https://link.springer.com/chapter/10.1007/978-981-10-8660-1_38**

Titulo	SMSPROTECT: An automatic smishing detection mobile application
Tipo de documento (artículo, tesis, conferencia, etc.)	TESIS
Editorial	ICT Express
Palabras clave	SmishingSpam SMS detectionRule based learningMobile application development
Año	2023
Objetivo del estudio (redacción propia)	Este estudio propone la creacion de una aplicacion que sirva de buzon de mensajes, funciona de manera que, al recibir un mensaje, lo envia mediante una API a un modelo de machine learning con reglas que determina si es un smishing o no
Conclusiones del estudio (redacción propia)	El estudio mediante el analisis del mensaje da resultados muy buenos con una clasificación correcta de smishing del 98%
Principales limitaciones (oportunidad de mejora)	El modelo utiliza un clasificador mediante una API que se encuentra en un servidor remoto, la mejora podría encontrarse en tener la clasificación en el propio dispositivo para reducir latencia

Enlace	**https://www.sciencedirect.com/science/article/pii/S2405959522000868**

Titulo	Rule-Based Framework for Detection of Smishing Messages in Mobile Environment
Tipo de documento (artículo, tesis, conferencia, etc.)	Articulo
Editorial	Procedia Computer Science
Palabras clave	SmishingMobile PhishingData miningShort messaging serviceMachine learning
Año	2018
Objetivo del estudio (redacción propia)	El trabajo propone un enfoque de clasificación basado en reglas (data mining) para detectar mensajes de smishing
Conclusiones del estudio (redacción propia)	Se afirma en el estudio que se logró tener un 99% de efectividad al detectar smishing
Principales limitaciones (oportunidad de mejora)	El modelo a pesar de tener una efectividad muy alta no tiene en cuenta jergas ni lenguajes diferentes
Enlace	**https://www.sciencedirect.com/science/article/pii/S1877050917328478**

Titulo	Smishing Dataset I: Phishing SMS Dataset from Smishtank.com
Tipo de documento (artículo, tesis, conferencia, etc.)	PAPER
Editorial	California State University San Marcos
Palabras clave	Smishing, Dataset, Phishing, VirusTotal
Año	2024

Objetivo del estudio (redacción propia)	El paper aborda la escasez de datasets actualizados de smishing (phishing por SMS), un obstáculo clave para prevenir estos ataques porque las campañas se caen y se pierde información.
Conclusiones del estudio (redacción propia)	El estudio deja claro que se necesita constantemente evolucion respecto a los datasets para identificar smishing, dejando una base de datos con mas de 1060 mensajes categorizados usando WHOIS y VirusTotal
Principales limitaciones (oportunidad de mejora)	Se requieren para el caso de estudio una cantidad de datos aún mayor de manera que se pueda ampliar la base de datos y refinar aun más
Enlace	**https://dl.acm.org/doi/pdf/10.1145/3626232.3653282**

Marco Conceptual y Teórico:

Smishing: Ingeniería social por SMS que induce al usuario a visitar un enlace malicioso o entregar datos personales.

URL maliciosa: Enlace que procura phishing, malware o fraude financiero; se clasifica por técnica (suplantación de marca, acortador opaco, typo/homoglifo, descarga forzada, etc.).

Homoglifos/Punycode: Caracteres visualmente parecidos usados para suplantar dominios (p. ej., “díán.gov.co” vs “dian.gov.co”).

Riesgo (score 0–100): Combinación de heurísticas locales (70%) + señales reputacionales remotas (30%, sólo si el usuario habilita sincronización) que se traduce en Bajo/Medio/Alto.

2.2 Legal y ética (límites y garantías)

Solo recolección pasiva: DNS/HTTP GET y renderizado controlado sin autenticación, sin fuerza bruta ni explotación.

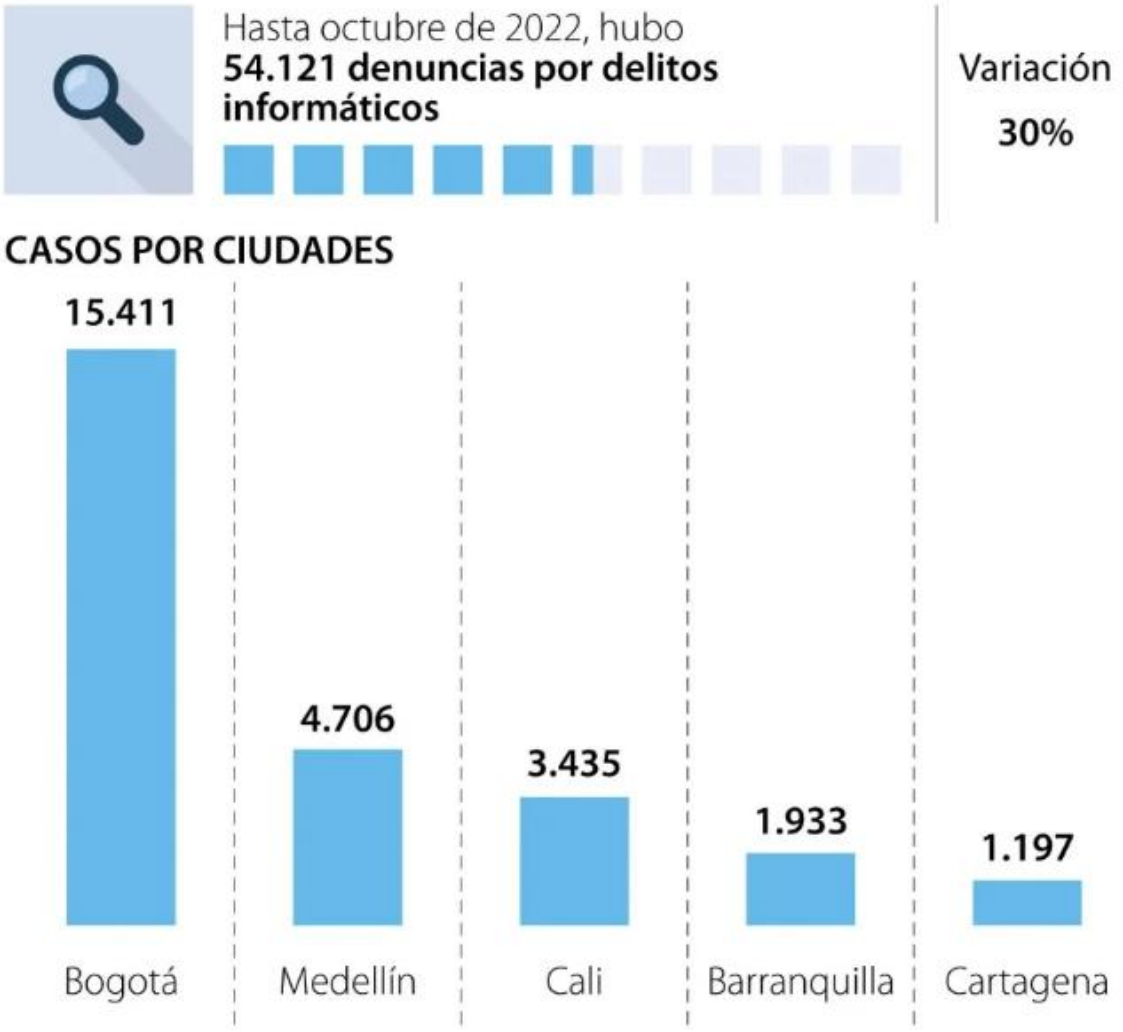
Privacidad por defecto: No leer SMS automáticamente; análisis sobre URLs pegadas manualmente salvo opt-in explícito. Anonimización y retención mínima.

Transparencia: Política clara (propósito, base legal, derechos ARCO). Logs agregados y sin PII.

Cumplimiento: Ley 1581/2012 (datos), Ley 1273/2009 (delitos informáticos), lineamientos CSIRT; respeto a robots.txt cuando aplique y rate-limits éticos.

Según la Cámara Colombiana de Informática y Telecomunicaciones, las denuncias por delitos informáticos se incrementaron un **30%** en el año 2022 con respecto a 2021 y aunque la ciberseguridad está siendo un tema prioritario para las empresas del país

	Figura 1: la imagen illustra denuncias por delitos informaticos



Fuente: www.larepublica.co

La imagen representa un gráfico de los principales ataques cibernéticos: **smishing (40%)**, **phishing (25%)** y **malware (18%)**. Este tipo de representación estadística permite dimensionar la magnitud del smishing como la amenaza más frecuente en el ecosistema digital

	Figura 2: la imagen ilustra los tipos de estafa en colombia
	
	Fuente: Autoría propia

Según el informe de seguridad de Kaspersky de 2023, Latinoamérica experimentó un aumento del 617% en los intentos de phishing en comparación con el año anterior.

	Figura 3:la imagen ilustra los paises mas afectados por pishing
	
	Fuente: KASPERSKY

Marco Legal

Protección de Datos Personales en Colombia (Ley 1581 de 2012)**:** Esta ley establece los principios y disposiciones para el tratamiento de los datos personales, garantizando el derecho de los ciudadanos a conocer, actualizar y rectificar la información que se tenga sobre ellos en bases de datos o archivos. En el marco de la propuesta, resulta fundamental porque la aplicación planteada deberá garantizar el consentimiento expreso del usuario y asegurar que cualquier manejo de datos respete la privacidad y confidencialidad de la información.

Delitos Informáticos en Colombia (Ley 1273 de 2009)**:** Con esta normativa se creó un marco jurídico específico para la protección de la información y los datos. La ley tipifica delitos como el acceso abusivo a sistemas informáticos, la interceptación de datos, el daño informático y el hurto por medios digitales. Esta regulación respalda la pertinencia de la solución, ya que el smishing constituye un mecanismo de fraude que puede derivar en varios de estos delitos, y la aplicación propuesta se orienta a prevenirlos mediante la identificación y reporte de enlaces maliciosos.

Principios de la UNESCO sobre Inteligencia Artificial Ética (2024): Los lineamientos propuestos por la UNESCO plantean que los sistemas basados en inteligencia artificial deben ser transparentes, explicables, inclusivos y no discriminatorios. Estos principios son relevantes para la investigación, en tanto que la aplicación propuesta —al evaluar riesgos en enlaces sospechosos— debe ofrecer explicaciones claras y comprensibles para cualquier usuario, asegurando decisiones automatizadas responsables y sin sesgos.

Software Libre y Propietario**:** La distinción entre software libre y propietario resulta pertinente, dado que la aplicación propuesta se plantea como un sistema accesible, gratuito y escalable. El uso de software libre bajo licencias abiertas (como la licencia Apache 2.0) garantiza mayor adaptabilidad, transparencia y colaboración en su desarrollo. Esto refuerza la viabilidad técnica y social del proyecto, asegurando que pueda ser adoptado y mejorado sin restricciones de propiedad intelectual cerrada.

1.3 Objetivo y Diseño Metodológico

1.4 Objetivos: General y Específicos

Objetivo General: Dado un URL/dominio, generar un reporte de riesgo de smishing acompañado de evidencia técnica (señales DNS/WHOIS, rastreo controlado, HTML/JS mínimo, certificados TLS y heurísticas), manteniendo privacidad por diseño y recolección pasiva.

Objetivos Específicos:

- **Decidir qué partes serán analizadas en el enlace** Haremos una lista clara: nombre del sitio (si se parece a una marca famosa), desde cuándo existe el dominio, si usa el "candadito" (https), si te manda a muchas páginas antes de cargar, si pide datos sensibles (usuario/clave), y detalles sospechosos en el texto (por ejemplo "urgente", "bloquearon tu cuenta"). Diremos por qué cada cosa importa para smishing.
- **Hacer una forma simple de pegar el link y analizarlo sin "tocar" la página.** El objetivo es que el usuario pega la URL y el sistema lo analiza sin que el usuario tenga que ingresar. No inicia sesión, no llena formularios, no hace clics. Solo revisa lo mínimo para no exponernos ni dar datos.
- **Juntar pruebas y ordenarlas para el reporte.** Guardaremos datos básicos: a dónde redirige el enlace, el título de la página, si tiene formularios, cuándo se registró el sitio, si el https está bien, y textos llamativos. Todo quedará ordenado para que cualquiera entienda de dónde salió el puntaje.
- **Crear reglas fáciles de entender que suman puntos de riesgo.** Ejemplos: "dominio recién creado" suma puntos; "nombre parecido a una entidad oficial" suma; "usa acortadores tipo bit.ly" suma; "muchas redirecciones" suma; "pide datos bancarios" suma; "texto de miedo/urgencia" suma. Cada regla dirá cuántos puntos aporta y por qué.
- **Convertir todo en un puntaje de 0 a 100 con niveles y explicación.** Mostraremos el resultado como semáforo (verde/amarillo/rojo) y diremos

exactamente qué señales subieron el puntaje. La idea es que el usuario vea “qué pasó” y no un número mágico.

- **Diseñar una pantalla sencilla y útil.**
  Un cuadro para pegar el link, un botón de analizar, y el resultado con: puntaje, lista de señales detectadas y recomendaciones claras (“no abras”, “verifica con tu banco”, “reporta aquí”). Por defecto todo funciona local (en el dispositivo) para cuidar la privacidad, y si el usuario quiere, puede activar consultas externas.
- **Probar con enlaces reales buenos y malos y ajustar.**
  Juntaremos ejemplos de smishing y también enlaces de verdad (bancos, gobierno) para medir cuántas veces acertamos o nos equivocamos. Con eso afinamos las reglas para bajar falsos positivos.
- **Cuidar la privacidad desde el principio.**
  No guardaremos los enlaces por defecto, borraremos registros en poco tiempo y no pediremos permisos raros del celular. Si el usuario decide reportar, solo se enviará lo mínimo necesario, sin datos personales.

2. Metodología

La metodología propuesta se estructura en seis etapas sucesivas que permiten pasar de la recolección inicial de información a la validación con usuarios reales, garantizando en todo momento la trazabilidad de las decisiones técnicas y la alineación con el objetivo general del proyecto: disponer de un sistema de análisis forense automático para enlaces utilizados en campañas de smishing.

2.1 Fase 1: Recolección y conformación del conjunto inicial de enlaces (meses 1–3)

En esta fase se conformará el **dataset inicial** que servirá como insumo para el desarrollo y las primeras pruebas. Para ello se integrarán tres fuentes:

1. **Fuentes públicas de amenazas:** se recopilarán enlaces catalogados como maliciosos (phishing/smishing) provenientes de listas abiertas, repositorios académicos y reportes de seguridad.
2. **Sitios legítimos de alta circulación en Colombia:** se incorporarán dominios y URLs de entidades financieras, comercio electrónico, operadores móviles y entidades gubernamentales con el fin de disponer de ejemplos negativos representativos del entorno nacional.
3. **Canal de reporte voluntario:** se habilitará un mecanismo para que los participantes del proyecto remitan enlaces sospechosos recibidos por SMS o mensajería. Estos enlaces serán revisados y etiquetados manualmente bajo una guía de clasificación (malicioso, dudoso, legítimo).

El resultado de esta fase es un **conjunto de datos versión 1 (dataset v1)** equilibrado entre casos maliciosos y legítimos, suficiente para validar el motor de reglas.

2.2 Fase 2: Diseño del modelo de reglas heurísticas (meses 2–4)

Antes de incorporar técnicas de aprendizaje automático, el sistema contará con un **núcleo de reglas determinísticas** que permitan emitir un juicio de riesgo aun en ausencia de conexión a internet. Estas reglas se basarán en factores comúnmente asociados al smishing, entre ellos:

- Antigüedad reducida del dominio.
- Uso de servicios acortadores para ocultar la URL real.

- Similitud léxica con nombres de marcas, bancos u organismos públicos.
- Presencia de formularios o solicitudes de información sensible en la primera carga.
- Cadena de redirecciones inusual.

Cada regla aportará un puntaje parcial a una **escala de riesgo unificada** (0–100) que luego será traducida al usuario mediante un esquema de semáforo (bajo, medio, alto). Esta fase produce el **motor heurístico básico**, completamente explicable y ejecutable de forma local.

2.3 Fase 3: Desarrollo del prototipo funcional (meses 4–6)

En esta etapa se implementará el **prototipo de la aplicación** (web o móvil) que materializa el flujo de uso previsto:

1. El usuario ingresa o pega el enlace recibido.
2. El sistema aplica el motor de reglas definido en la fase anterior.
3. Se genera un **informe compacto de riesgo** que incluye: nivel de riesgo, factores que lo originan y recomendaciones de acción.
4. Si existe conectividad y el usuario lo ha autorizado, el sistema consulta fuentes externas (por ejemplo, información WHOIS actualizada o listados de bloqueo) para **enriquecer** el análisis; si no la hay, el sistema opera únicamente con la información local.

El principio rector de esta fase es **"local-first, cloud-assisted"**: la funcionalidad principal no depende de la nube, pero puede mejorar cuando ésta está disponible.

2.4 Fase 4: Incorporación del módulo de mejora automática (meses 6–9)

Una vez que el proyecto cuente con un volumen suficiente de enlaces analizados y etiquetados, se incorporará un **módulo de mejora automática** encargado de identificar patrones más finos que los capturados por las reglas. Este módulo utilizará técnicas ligeras de clasificación para:

- Reajustar pesos de las reglas cuando se observen falsos positivos o falsos negativos.
- Detectar combinaciones de señales que, tomadas en conjunto, incrementen el riesgo.
- Aportar una estimación adicional de probabilidad de fraude.

Este módulo **no sustituye** al motor heurístico, sino que lo complementa. En caso de no contar aún con datos suficientes, el sistema seguirá operando únicamente con reglas, de manera que el avance del proyecto no queda condicionado por esta fase.

2.5 Fase 5: Validación con usuarios y retroalimentación (meses 9–12)

Con el prototipo en funcionamiento se realizará una **prueba piloto** con usuarios reales (comunidad universitaria y voluntarios externos) con los siguientes propósitos:

- Verificar la comprensión del resultado presentado por el sistema (semáforo y explicación).
- Medir la utilidad percibida para la toma de decisiones inmediatas frente a un enlace recibido por SMS.
- Recoger nuevos enlaces sospechosos generados en contexto local.
- Identificar mensajes, términos o pasos del flujo que requieran simplificación.

Los enlaces que se obtengan en esta fase, tras ser revisados, **volverán al dataset** para fortalecerlo con casos actuales y propios del entorno colombiano.

2.6 Fase 6: Ajustes finales, documentación y entrega (meses 12–16)

La última fase se orienta a consolidar el sistema:

- Ajuste de reglas que hayan mostrado baja precisión.
- Revisión de los textos de salida para garantizar claridad hacia usuarios no técnicos.
- Documentación del procedimiento de actualización del dataset y reentrenamiento.
- Preparación del informe técnico y del manual de uso.
- Empaquetamiento del prototipo estable para su instalación o despliegue controlado.

Al finalizar esta fase, el proyecto dispondrá de: (i) una herramienta operativa de análisis de enlaces de smishing, (ii) un conjunto de datos curado y (iii) la documentación necesaria para su réplica y mantenimiento.

Respuesta a observaciones del comité

1.Este proyecto requiere hosting en la nube, microservicios: Nuestro proyecto requiere un apoyo en la nube, pero no una arquitectura compleja. El análisis básico (reglas, puntaje, explicación) se hace localmente; la nube solo se usa para actualizar información y recibir reportes. Es decir, funciona sin nube y mejora con nube.

2. Y sin datos, el modelo ML y las heurísticas serán difíciles de evaluar: Nuestro proyecto no depende desde el inicio del modelo. Primero se implementa y evalúa el motor de reglas con el dataset inicial; luego, cuando haya más enlaces reales, se añade el módulo de mejora automática. Así, no se bloquea el avance por falta de datos.

3. Dependencia de servicios externos, esto puede complicar el desarrollo y aumentar el presupuesto:
Nuestro proyecto sí contempla servicios externos, pero como componentes opcionales. Si no están disponibles, el sistema opera solo con la información local. Por eso la dependencia es limitada y controlable dentro del presupuesto previsto.

3 Cronograma y Recursos

3.1 Cronograma

Objetivo		Actividad		Meses														
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Recolectar y normalizar datos de URLs maliciosas y benignas.	**1.1**	Revisión bibliográfica y normativa (phishing, smishing, explicabilidad, legislación colombiana). Informe de estado del arte.																
		1.2	Recolección inicial de URLs (listas negras, feeds públicos y sitios benignos). Dataset v1.																
		1.3	Normalización y enriquecimiento de datos (DNS, WHOIS, ASN, tiempos de registro).																
2	Diseñar la arquitectura del sistema y definir reglas heurísticas.	**2.1**	Diseño arquitectónico del sistema (backend, crawler, base de datos, panel). Documento técnico.																

		2.2	Desarrollo del crawler headless (captura HTML, JS, redirecciones).																
		2.3	Implementación de heurísticas base (punycode, Levenshtein, TLS, iframes ocultos). Motor heurístico v1.																
3	Entrenar y validar un modelo ligero de ML.	**3.1**	Extracción de features léxicas y de comportamiento.																
		3.2	Entrenamiento de modelo ligero (XGBoost/LightGBM). Modelo calibrado.																
		3.3	Explicabilidad con SHAP/feature importance y exportación ligera.																
4	Integrar API y panel de usuario.	**4.1**	Implementar API REST (/analyze y /report).																
		4.2	Desarrollo del panel de usuario (semáforo de riesgo y evidencias).																
		4.3	Integración pipeline completo (heurísticas + ML + API + panel). Prototipo v1.																
5	Validar el prototipo y realizar pruebas piloto.	**5.1**	Recolección ampliada de dataset (≥6k URLs).																
		5.2	Realizar pruebas comparativas del sistema desarrollado y otros métodos de predicción de inventarios, empleando el mismo conjunto de datos para una evaluación justa.																
		5.3	Sacar las conclusiones del proyecto, identificando sus limitaciones, posibles mejoras y opciones de implementación futura.																
		5.4	Ajustes de sistema (umbrales, interfaz, mensajes educativos).																
6	Documentar, difundir y cerrar el proyecto.	**6.1**	Redacción de informe final (metodología, resultados, limitaciones).																
		6.2	Preparación de artículo científico y ponencia en congreso.																
		6.3	Defensa y entrega de prototipo estable, dataset anonimizado y scripts reproducibles.																

3.2 Presupuesto

Presupuesto Total:

PRESUPUESTO Y FUENTES DE FINANCIACIÓN (Miles de Pesos)					
RUBROS	**UMB**		**ESTUDIANTES**		**TOTAL**
	Efectivo	**Especie**	**Efectivo**	**Especie**	
PERSONAL	$-	$4.132.545	$ -	$24.795.273	$28.927.818
EQUIPOS	$-	$ -	$5.000.000	$ -	$5.000.000
SERVICIOS TECNOLÓGICOS	$ -	$1.600.000	$7.200.000	$6.771.600	$8.800.000
SOFTWARE	$ -	$ -	$ 2.000.000	$ -	$ 2.000.000
TOTAL	$ -	$5.732.545	$14.200.000	$24.795.273	$44.727.818

Personal:

DESCRIPCIÓN GASTOS DE RECURSOS HUMANOS						
Investigador	**Nivel de formación**	**Función en el Proyecto**	**Dedicación Horas / Semana**	**No. de Meses**	**Valor**	**Responsable**
Asesor del proyecto	Maestría	Dirección metodológica	1	12	$4.132.545,45	UMB
Investigador 1	Técnico	Ejecutor de proyecto	8	12	$8.265.090,91	Estudiante
Investigador 2	Técnico	Ejecutor de proyecto	8	12	$8.265.090,91	Estudiante
Investigador 3	Técnico	Ejecutor de proyecto	8	12	$8.265.090,91	Estudiante

Equipos:

DESCRIPCIÓN GASTOS EN EQUIPOS (Miles de Pesos)					
Descripción	Cantidad	Valor unitario	Responsable	Tipo de fuente	TOTAL
Mini PC básica para pruebas y desarrollo (ej. Intel NUC, Ryzen Mini PC o similar)	1	$2.800.000	Estudiante	Efectivo	$2.800.000
Laptop gama media (para desarrollo y presentaciones)	1	$2.200.000	Estudiante	Efectivo	$2.200.000
				TOTAL	$ 232.000

Servicios Tecnológicos:

DESCRIPCIÓN GASTOS DE SERVICIOS TECNOLÓGICOS (Miles de Pesos)					
Descripción	**Cantidad**	**Valor unitario**	**Responsable**	**Tipo de fuente**	**TOTAL**
Hosting en la nube (ej. AWS/Azure/Google Cloud)	12 meses	$600.000	Estudiante	Efectivo	$7.200.000
Suscripción Office 365 Educativa (anual)	3	$533.000	UMB	Especie	$1.600.000
TOTAL					**$8.800.000**
				TOTAL	$ 6.663.638

Software:

DESCRIPCIÓN GASTOS DE SOFTWARE ESPECIALIZADO (Miles de Pesos)					
Descripción	**Cantidad**	**Valor unitario**	**Responsable**	**Tipo de fuente**	**TOTAL**
Licencias API externas (ej. autenticación, mapas, analítica)	1	$1.200.000	Estudiante	Efectivo	$1.200.000
Herramientas de desarrollo (IDE, plugins, librerías premium)	1	$800.000	Estudiante	Efectivo	$800.000
TOTAL					**$2.000.000**

4 Referencias

Mishra, S., & Soni, D. (2020). Smishing Detector: A security model to detect smishing through SMS content analysis and URL behavior analysis. *Future Generation Computer Systems*, *108*, 803–815. https://doi.org/10.1016/J.FUTURE.2020.03.021

¿Qué es el smishing y cómo protegerse de esta estafa por mensajes de texto? (n.d.). Retrieved September 8, 2025, from https://www.elespectador.com/tecnologia/que-es-el-smishing-y-como-protegerse-de-esta-estafa-por-mensajes-de-texto/