Informe final proyecto de grado

Prototipo de Software para la Detección de Contenido Perjudicial en Dispositivos de Menores de edad mediante Redes Neuronales Profundas de Procesamiento de Lenguaje Natural

Autores

Geovanni Andrés Gil Cabuya

Soed Alejandra Rodríguez Torres

Yerly Alejandra Sánchez Ardila

Director

Álvaro David Arévalo Salazar

Facultad de Ingeniería

Diciembre 10 de 2024

Bogotá, Colombia

CONTENIDO

LISTADO DE TABLAS

LISTADO DE ILUSTRACIONES

1.RESUMEN EJECUTIVO

El propósito de esta investigación es poder informar y alertar a los padres y/o tutores de un menor de edad cuando estos estén accediendo o brindando algún tipo de información que pueda comprometer su seguridad y pueda ponerlos en peligro; esto teniendo en cuenta la gran cantidad de casos de ciber acoso que se pueden presentar en un año; desde la oficina de seguridad cibernética de la Policía nacional de Colombia aseguran que durante el 2022 se presentaron 8.981 casos de ciberacoso**.** (Rico muñoz, 2022)**.**Ante este antecedente, inicia la investigación de tipo descriptiva correlacional – cualitativa que busca desarrollar una herramienta tecnológica que permitan a los padres o responsables del menor recibir alertas si la actividad en línea de sus hijos se ve comprometida de manera insegura.

El presente proyecto propone el desarrollo de un prototipo de software capaz de detectar señales de alerta en la comunicación en línea de los menores cuando están en chats, en grupos, redes sociales y mensajes de texto, brindando a los padres una herramienta necesaria para proteger a sus hijos.

Para llevar a cabo esta investigación, es crucial identificar las variables que permiten evaluar la relación entre el desarrollo de la herramienta tecnológica y su efectividad en la protección de los menores contra el ciberacoso. Entre estas variables se encuentran el tiempo total que los menores pasan en internet y la cantidad de ese tiempo que ocurre sin supervisión adecuada. También se investigará si existe una correlación entre el tiempo total en internet y el tiempo específico dedicado a chats y redes sociales. Además, se analizará cómo el tipo de comunicación que utilizan los menores afecta el número de alertas generadas por la herramienta, con el objetivo de determinar si ciertos tipos de comunicación están asociados a una mayor generación de alertas.

2. INTRODUCCIÓN

Los niños y adolescentes están cada vez más conectados a los dispositivos digitales, como celulares, computadores, tabletas inteligentes y otras pantallas con libre acceso a internet, una red sin límites, donde cualquier persona, adultos adolescentes y niños pueden acceder a cualquier información o herramienta que deseen visualizar. El uso de internet sin supervisión en menores de edad genera alto riesgo, ya que pueden estar expuestos a diferentes escenarios peligrosos como: acceder a información que no es apta para menores de edad, ser engañados por expertos que pueden influir en la toma de decisiones de los menores, presentar algún tipo de acoso, entre otras.

En muchas oportunidades los padres no realizan una supervisión a los menores de edad, algunos por desconocimiento de los riesgos, otros por falta de tiempo o porque sencillamente no lo ven importante. Según Andrea Ortiz, siete de cada diez niños y adolescentes en Colombia han sido víctimas de grooming, más conocido como ciberacoso infantil que comprende acciones orientadas a ganar la confianza de menores a través del engaño, con el propósito de conseguir desde imágenes, videos con contenido sexual o pornográfico, hasta la posibilidad de tener contacto físico con el menor de edad para abusar sexualmente de él. (Ortiz, 2024).

Con el avance en esta investigación se pretende desarrollar un prototipo de software que, con la obtención de lo escrito por el menor, analice y detecte palabras clave y patrones mediante redes neuronales profundas NLP, para generar alertas tempranas para los padres, permitiéndoles intervenir oportunamente. Además, se busca garantizar la privacidad y tratamiento de datos, protección de los datos de los usuarios, tal como lo son las conversaciones escritas por los menores de edad y la autorización por sus tutores. Adicionalmente, se considerará una interfaz sencilla para el manejo de los padres, con el objetivo de uso eficaz y ágil. Es por eso que se presenta el siguiente documento el cual se encuentra estructurado de una manera clara y sencilla de navegar en donde se busca destacar, cómo la ingeniería industrial y la ingeniería de sistemas influyen en la gestión y ejecución del proyecto. La ingeniería industrial desempeña un papel crucial al optimizar los procesos, gestionar recursos y abordar desafíos operativos, proporcionando soluciones clave para mejorar la eficiencia y reducir costos. Por otro lado, la ingeniería de sistemas

contribuye mediante la implementación de tecnologías avanzadas, integración de sistemas y análisis de datos para apoyar la toma de decisiones informadas y la automatización de procesos generando un enfoque integral y así garantizar una ejecución efectiva y sostenible del proyecto; seguido del resumen ejecutivo y la introducción el documento contempla los objetivos generales y específicos de la investigación, seguido de la definición del problema, en donde se describe información detallada, antecedentes y la formulación del pregunta principal del problema, luego se socializa la justificación seguida por el análisis de requerimientos y el marco de referencia, se analiza también las restricciones, se describe la metodología para la selección y desarrollo de la solución, y se realiza un análisis de costos; finalmente, se termina con una serie de conclusiones de su viabilidad de uso.

3. **OBJETIVOS**

3.1. **Objetivo General**

Desarrollar un prototipo de software basado en redes neuronales y procesamiento de lenguaje natural (NLP) que identifique patrones y palabras clave asociados a riesgos en la comunicación en línea de menores de edad, generando alertas automáticas para los padres, y cumpliendo con las normativas de protección de datos vigentes.

3.2. Objetivos Específicos

- Desarrollar un modelo de análisis de lenguaje natural (NLP) que identifique señales de alerta en las comunicaciones en línea de menores, detectando patrones asociados a riesgos como el ciberacoso y el grooming.
- Diseñar una interfaz de usuario intuitiva que permita a los padres monitorear en tiempo real las alertas generadas por la herramienta, con funciones de personalización y seguridad de datos garantizadas bajo las normativas de protección de información infantil.
- Realizar pruebas para evaluar la precisión en la detección de riesgos y la usabilidad de la interfaz para los padres, asegurando una experiencia de usuario óptima.

4. DEFINICIÓN DEL PROBLEMA

El creciente uso de dispositivos tecnológicos entre los menores de edad los expone a diferentes riesgos en línea. Siendo uno de los problemas la exposición a contenido perjudicial, como el ciberacoso, sexting, chantajes, explotación sexual y solicitudes de material sexual. Según estudios previos (Durkin, 1997; O'Connell, 2003) estudiados en (Paquette, Fortin & Perkins 2020), los menores son especialmente vulnerables a estas amenazas debido a su ingenuidad y confianza en el entorno digital, donde son contactados por los “Atractores sexuales de menores” siendo adultos o menores que atraen a otros menores para generar conversaciones inapropiadas y poder manipularlos. Usando diferentes tácticas de manipulación psicológicas para ganar la confianza de niños y adolescentes.

El ciber acoso se presenta como un problema significativo, donde los menores son explotados a través de juegos y chats en línea, los agresores utilizan tácticas como ofrecer regalos y manipular a los menores con información sobre su situación personal para aprovecharse de ellos; según en Online Grooming: de los juegos en línea la obtención de material de abuso sexual, el 73.33% de niños y adolescentes admitieron que hablaron con desconocidos en juegos (Guillen, L 2024), demostrando que los menores hoy en día encuentra diversos peligros al estar conectados a internet, y más cuando personas usan métodos para poder hacer que el menor se sienta tranquilo para así obtener información y luego ejecutar los siniestros planes.

Además de la utilización de un lenguaje codificado, para la obtención de contenido, como para reconocerse entre ellos, utilizan "palabras clave" identificadas en el estudio de Trujillo (2019) citado en la revista Dialogo Forense (2023), siendo las más comunes “Caldo de Pollo”, “Camión Pesado”, “Cheese Pizza” “Club Penguin”, las cuales comienzan con las siglas CP, que en inglés se refieren a Child Pornography, que se traduce como pornografía infantil (Sánchez F 2023, p.25) dificulta aún más la detección de estas prácticas, cuando no se tiene conocimiento de las mismas ya que inocentemente se puede pensar en el sentido de la palabra mas no en el contexto que tendría esta palabra, teniendo en cuenta que si los padres tienen desconocimiento y revisan palabras como las anteriores y no le presten atención sin

saber que esto se trataría de una alerta para el cuidado de sus hijos, resulta muy importante lograr reconocer estas palabras y trabajarlas para poder detectarlas.

La falta de supervisión adecuada y la carencia de herramientas tecnológicas eficaces para identificar y bloquear contenido perjudicial agravan la situación. Un estudio realizado por Plan Internacional de República Dominicana y UNICEF (2021) reveló que un alto porcentaje de adolescentes pasa la mayor parte del día conectado a internet y que el 20% de ellos han recibido solicitudes de material sexual a través de plataformas de mensajería instantánea siendo el 62% por parte de personas desconocidas que encontraron en internet, como también el desconocimiento que tienen ellos a diversos temas en el mismo estudio de Republica Dominicana y Unicef (2021) revela que al preguntar por conocimientos sobre violencia en línea con enfoque en explotación sexual en línea en niñas, niños y adolescentes el grupo de 12 a 14 años el 30% desconocían estos temas mientras que en el grupo de 15 a 17 años lo desconocía solo el 14%, dándonos una evidencia que no todos conocen los términos sexuales siendo aprovechados por los depredadores.

La preocupación radica que actualmente es muy raro no encontrar personas con dispositivos móviles incluso en los menores, esto sumado al tiempo que pasan en él y la adicción que puede generar aumenta la probabilidad de que sean contactados en algún momento por personas que no buscan su amistad si no que van por objetivos mal intencionados con el menor, sin olvidar que la red donde más la pasan según el estudio de Republica Dominicana y Unicef (2021) el 17% del tiempo que están conectados lo pasan en chatear atreves de WhatsApp Messenger u otros, y sumado a esto también se revela en el estudio que el 15% confirmó participar o haber participado en conversaciones con contenido sexual o sexo virtual, siendo con un 11% con personas adultas y otro 11% con desconocidos, siendo una gran preocupación que los menores de edad realicen estas interacciones con adultos y personas desconocidas que puedan usar estos materiales para chantajes o llevarlos a realizar cosas más allá de solo una conversación.

En resumen, el problema radica en la vulnerabilidad de los menores ante los riesgos del entorno digital como la falta de información de los peligros tanto para los menores como para los padres, combinada con la dificultad de identificar y prevenir estas amenazas. La falta de control de los padres, el uso de lenguaje codificado por parte

de los acosadores o depredadores sexuales, el conocimiento por parte de muchos de estos depredadores en la manipulación como el acercarse a ellos y la ausencia de herramientas tecnológicas eficaces crean un entorno propicio para la explotación y el abuso de niños y adolescentes, por esto no es raro, cada vez existe más contenido Infantil en línea y grupos donde se difunde de manera “Libre” este contenido o donde se incita a realizar los mismos, donde comparten sus experiencias y aconsejan a otros.

Concluyendo que muchos de los menores afectados por estas prácticas se deriven a afectaciones de salud, depresiones, suicidios, sean expuestos en internet, como también tener acceso a material no apto para ellos y el aislamiento social e incluso llevándolos a que comentan delitos, contribuyendo también al problema como se hablaba anteriormente con los “Atractores sexuales de menores”, ya sea por amenazas o por ellos mismos que resultan realizando estas mismas prácticas, llevando a la contribución del problema y aumentando más la captación de menores debido a que no son tan adultos o pueden moverse tanto dentro de las instituciones o incluso la casa de las victimas por la condición de menor al igual que su víctima, siendo a veces de la misma edad o menos.

Con todo lo anterior se establece la siguiente pregunta **¿Cómo desarrollar un prototipo de software que identifique de manera precisa y oportuna patrones y palabras clave en los escritos de menores para generar alertas efectivas para los padres?** .

Con el desarrollo de esta investigación y la implementación de esta herramienta se ayudará de gran manera a los padres, dado que se generará un filtro que evite que los menores de edad puedan caer en estas redes y en el futuro no tengan que

sobrellevar afectaciones de salud o depresiones debido a estos casos, o inclusive convirtiéndose también parte del problema al optar estas prácticas o ser obligados a traer más con tal de salir ellos de estas, es cierto que la tecnología nos ha traído muchos beneficios pero también problemáticas como el anonimato del internet y la facilidad de poder acercarse a las personas más vulnerables, con tal de cumplir sus oscuras ambiciones, es necesario reconocer aprender y entender la vulnerabilidad de los más jóvenes y en consecuencia dar soluciones y enseñar sobre estas prácticas que abundan en internet.

A pesar de la creciente conciencia sobre los peligros en línea, la implementación de medidas de protección sigue siendo insuficiente y fragmentada. El entorno digital ofrece un anonimato que facilita la evasión de la justicia por parte de los depredadores y complica la labor de identificación y prevención (Sánchez, 2023). Además, muchos padres y educadores carecen de la formación necesaria para abordar estos problemas de manera efectiva (Plan Internacional & UNICEF, 2021). Las plataformas digitales, a menudo, no tienen políticas claras o mecanismos robustos para proteger a los menores de edad y las herramientas de monitoreo existentes pueden ser inadecuadas o demasiado restrictivas (Guillen, 2024). Esta situación crea una brecha crítica entre la exposición de los menores a contenidos dañinos y la capacidad de los adultos para protegerlos adecuadamente. Por lo tanto, es esencial que se desarrollen y se implementen soluciones tecnológicas integrales y accesibles, junto con una educación continua para padres y educadores, para enfrentar de manera eficaz los desafíos emergentes en la protección infantil en el entorno digital (Trujillo, 2019). La colaboración entre tecnología, educación y políticas públicas será clave para crear un entorno en línea más seguro y proteger a los menores de las amenazas presentes en internet (Paquette, Fortin, & Perkins, 2020).

5. JUSTIFICACIÓN

En Colombia, el creciente acceso de los menores a internet los expone a grandes riesgos en línea, como el contenido inapropiado para ellos, la manipulación y explotación. Siendo preocupante la gran proporción que navega sin un supervisor parental siendo el 49% los que navegan sin un adulto cifras de DANE 2021, volviéndolos así bastante vulnerables. Y no solo esto si no el problema de muchos adultos con la falta de tiempo, atención o cuidado por parte de los menores dejando que usen dispositivos como manera de entretención y cuidado no sabiendo el gran peligro que pueden correr sus hijos.

Por esta razón se realiza este proyecto, para poder ayudar a los padres en la revisión de sus hijos a pesar de no contar con el tiempo logrando así, estar tranquilos sabiendo que sus hijos cuentan con una extra de protección, cuando están navegando o usando redes sociales, y no solo beneficia a los padres o tutores legales del menor, si no al menor mismo, debido que se evitaría que extorsionen al menor o engañen para que

este haga cosas que no debería y así caer en manos criminales o depravados sexuales en línea, que les harán pasar por cosas que conllevaran que en su futuro tengan secuelas de estos sucesos o inclusive se vuelvan de víctimas a victimarios, se puede pensar que los menores pueden discernir o dudar y evitar pero las personas que se encargan de engañar a los menores son muy astutos y la facilidad con la que los menores pueden ser persuadidos, según la “Revista científica dialogo forense 2023” siendo una práctica común el uso de imágenes de persuasión, donde presentan dibujos animados en actos sexuales explícitos (Sánchez F 2023), para lograr normalizar estos actos en ellos y aprovecharse así de sus víctimas (Dupuy 2020), convirtiéndose en blancos fáciles para depredadores sexuales y ciberacosadores.

No solo en la parte de explotación sexual, si no inclusive en el mismo acoso por parte de desconocidos o mismos “amigos”, que empiezan a incitar al menor a realizar cosas que afectan su integridad física o llenando su cabeza de pensamientos sobre su ser, su cuerpo, su apariencia, su existencia y demás, que conllevan al menor alejarse de todo siendo personas que en su vida adulta, tendrán secuelas por esto o inclusive suicidándose por estas cargas, puede no impedirse el 100% de todos estos casos, pero si se podrá realizar una gran mitigación en detectar oportunamente el acoso a los menores evitando consecuencias fatales, como también evitando que caigan en redes pedófilas donde pueden incluso ser llevados explotación sexual infantil, logrando así tener una defensa ante la cual, si el menor no siente la seguridad o no sabe que hacer o le invaden las dudas, los padres o tutor pueda ser notificado de escritos dudosos, donde el pueda actuar guiando y ayudando al menor para evitar que estén solos y vulnerables ante acosadores, depredadores sexuales y extorciones, reduciendo que los menores caigan y haciendo que su futuro sea mejor y no lleno de melancolía tristezas y pensamientos suicidas.

6. ANÁLISIS DE REQUERIMIENTOS

6.1. Identificación de stakeholders

- **Padres y tutores legales del menor**

Preocupaciones principales: La seguridad de sus hijos en las redes sociales, como las personas que conocen en estas redes, la protección de la privacidad.

Expectativas: Que sea un sistema fácil de usar e intuitivo, eficaz en la identificación de contenido perjudicial y notifique de cualquier alerta que sea de preocupación.

Impacto en el proyecto: Sus opiniones son valiosas y muy importantes debido que son el cliente objetivo de este desarrollo para poder definir las funcionalidades y el interfaz del sistema.

- **Menores de edad**

Preocupaciones principales: Su información se protegida y no compartida para terceros, que no sea incomoda mientras ellos realizan otras actividades en el teléfono.

Expectativas: Que les ayude a identificar cosas que aún no entienden y pueden ser vulnerables a ello por desconocimiento.

Impacto en el proyecto: La privacidad debe ser algo fundamental debido que se está trabajando con menores de edad, como también la opinión sobre el funcionamiento del mismo sin la afectación de sus actividades logrando esto por sus comentarios y/o opiniones.

- **Legisladores y reguladores**

Preocupaciones principales: Garantizar la protección del usuario y privacidad de sus datos, responsabilidad si se el aplicativo comete un error.

Expectativas: Que cumpla con todas las leyes y regulaciones aplicables, que todo sea trasparente con el uso de la información, que tenga seguridad el sistema para la protección de datos y usuarios

Impacto en el proyecto: Saber los límites legales del proyecto, la elaboración de políticas que garantice la protección de datos de los usuarios

6.2. **Recolección de requerimientos**

- Entrevistas: Se tiene que realizar entrevistas a los principales clientes siendo estos los padres y tutores legales del menor para conocer más sus

preocupaciones como también sus inquietudes respecto a este programa, sabiendo con detalle más lo que les preocupa, les agrada y les molestaría, como al igual al menor para conocer también sus puntos de vista.

- Encuesta: Para obtener datos cuantitativos es necesario el uso de encuestas para conocer el impacto de este proyecto, una vez echas las entrevistas para la elaboración de encuestas donde se contemple lo recopilado de las entrevistas como preguntas necesarias para el desarrollo como la interfaz el manejo la aprobación de la misma.
- Análisis: La revisión de estudios previos y artículos donde se aborden temáticas parecidas para entender y contemplar más escenarios, tanto para la elaboración como también los cuidados al tener al usar datos de menores
- Análisis de productos similares: Estudiar los productos actuales en el mercado parecidos.

6.3. **Categorización de requerimientos**

Requerimientos Funcionales

- Detección de contenido: La identificación de texto no apropiado para el menor usado para sus búsquedas en internet, redes sociales donde se detecte el lenguaje ofensivo, amenazante o discriminatorio, como a su vez la implicación de actividades ilegales o peligrosas donde se ayude en el filtro de términos que son usados y camuflados con palabras inocentes pero que conllevan un significado oculto.
- Generar las alertas a los padres: Si se identifica el uso de lenguaje no apropiado o saltan palabras con otros significados se envié una notificación a los padres o tutores responsable del menor para la alerta del mismo para que pueda verificar lo que está sucediendo.
- Personalización: Poder personalizar por parte de los padres o tutores del menor, los filtros que desean agregar, retirar, o incluso si detectan palabras nuevas que no se tengan en la aplicación puedan colocarlas para que así puedan tener más protección con el menor.
- Excepciones: Poder agregar en caso de que se necesite excepciones de palabras o filtros.

- Integración: Integrarse a diferentes dispositivos.
- Login: Tener un login para que solo los padres o tutores puedan ingresar con su usuario y contraseña evitando manipulaciones por terceros o incluso por el menor.

Requerimientos No funcionales

- Desempeño: Tiempo de respuesta rápido para la detección del contenido logrando analizar grandes cantidades de datos en tiempo real
- Seguridad: Protección de los datos del menor como la de sus usuarios, como también la prevención de ataques cibernéticos y el cifrado de la información.
- Usabilidad: Una interfaz intuitiva y fácil de usar para los padres o tutores del menor, documentación clara y precisa para el manejo, y un soporte técnico eficiente.
- Precisión: Alta precisión para la detención de contenido perjudicial con baja tasa de bajos positivos.
- Escalabilidad: Capacidad para adaptarse al aumento de usuarios y dispositivos, como la cantidad de datos que manejara.
- Disponibilidad: Estar disponible 24x7 para que en cualquier momento se pueda disponer de la aplicación.
- Idioma: se pueda expandir a múltiples idiomas para facilitar el manejo
- Compatibilidad: La compatibilidad con diferentes sistemas operativos y dispositivos.
- Privacidad por diseño: aplicar la privacidad a todas las etapas del desarrollo del sistema como a su vez la minimización de recopilación y almacenamiento de datos personales.

6.4. **Documentación de requerimientos**

Matriz de rastreabilidad: Establecer una matriz para poder conocer los requerimientos como a su vez una descripción del mismo, el caso de uso y las pruebas logrando así establecer una conexión entre los requerimientos, los casos de uso, las pruebas y los elementos de diseño para conocer la implementación en las diferentes fases del proyecto.

Casos de uso: Describir como un usuario interactúa con el sistema para lograr un objetivo específico, teniendo en cuenta el flujo principal y alternativos.

Especificación de usuario: Describir el sistema desde la perspectiva del usuario donde este las características de sistemas, funcionalidades, limitaciones y procedimientos de operación.

Glosario: Definir los términos y conceptos utilizados en la documentación, donde se encontrarán términos técnicos y sus definiciones.

Diagrama: Visualizar de manera grafica los requerimientos con diagramas de flujo y diagramas de caso de uso.

6.5. **Validación de requerimientos**

Revisión por pares

- Walkthrough: una revisión informal donde se revisará los requerimientos en voz alta para poder conocer falencias o saber que todos entendamos de manera clara el requerimiento
- Revisión técnica: Revisión enfocada en aspectos técnicos como la viabilidad y la coherencia para garantizar que se esté revisando de manera clara y concisa.

Prototipos

- Low-fidelity: Elaboración de bocetos simples para la validación de la interfaz de usuario y experiencia del usuario, esto con el fin de poder contemplar una versión muy preliminar y contemplando la necesidad de cambios o agregar aspectos sin la necesidad de hacer grandes cambios.
- High-fidelity: Prototipo interactivo para simular la funcionalidad del software, donde se tendrá en cuenta la interfaz ya integrada en los diferentes dispositivos y la funcionalidad para dispositivos Windows de demostrando la funcionalidad.

Pruebas de concepto (POC)

- Demostración de viabilidades: validar tecnologías y enfoques técnicos, esto mediante la implementación de las Apis para la comunicación con las IA NLP,

y la contemplación de los diferentes modelos para saber cuál se integra mejor y da mejores resultados.

- Reducir riesgos: Identificar y mitigar los riesgos esto mediante pruebas para poder saber que se tenga el funcionamiento esperado, reduciendo fallas en el proceso real.

Validaciones basadas en usuarios

- Encuestas de usuarios: Obtener el feedback por parte de los usuarios sobre los requerimientos.

6.6. Análisis de riesgo

Identificación de riesgos

- Técnicos: Problemas con la tecnología, incompatibilidad de sistemas, fallas de hardware y vulnerabilidades de seguridad, donde se aborda los diferentes modelos como también los dispositivos usados y la incompatibilidad que podría causar estos mismos, como fallas que se presenten a la integración de todo el prototipo y las vulnerabilidades que se puedan encontrar en la utilización del mismo.
- Recursos: Presupuesto ineficiente o disponibilidad limitada de recursos, esto con el fin de conocer el costo que estaría asociado al uso de este prototipo de una manera más extensa debido al uso de IA donde se tienen límites de uso mediante Apis.
- Alcance: Cambios en los requerimientos, si llegan a presentar.
- Cronograma: Cambio en las prioridades, debido a surgimientos de problemas o un cambio en la propuesta de frente con la problemática a tratar más que todo en el código esto también involucra el modelo de IA que se vaya a quedar.
- Calidad: Errores en el desarrollo y no cumplir con los estándares de calidad, para la mitigación de estos mismos se contempla pruebas para poder realizar y probar el desarrollo.
- Externos: Cambios de regulaciones gubernamentales y desastres naturales.

Evaluación de riesgos

- Probabilidad: Estimación de la probabilidad que suceda el riesgo en una escala de baja, media y alta.
- Impacto: Evaluación del impacto potencial que se tendría de un riesgo en el proyecto en escala de bajo, medio y alto.
- Matriz de riesgos: Una herramienta visual que combina la probabilidad y el impacto para priorizar los riesgos.

Plan de mitigación

Estrategias de mitigación: si es posible la eliminación del riesgo o reducir significativamente la probabilidad de que suceda y el impacto que generaría siempre aceptando el riesgo y planificando las consecuencias para poder tener un plan de contingencia para poder tener acciones alternativas en caso de suceder.

Responsables: Tener alguien específico para implementar acciones de mitigación.

Indicadores: Definir métricas para poder monitorear la efectividad de las acciones de mitigación.

7. MARCO DE REFERENCIA

Este marco teórico proporciona una visión integral y detallada del diseño y la implementación de un prototipo de software destinado a identificar contenido nocivo en dispositivos utilizados por menores. El sistema propuesto emplea técnicas avanzadas de redes neuronales profundas en el procesamiento de lenguaje natural.

Para una comprensión completa del proyecto, es esencial entender la interacción interdisciplinaria entre la ingeniería industrial y la ingeniería de sistemas. La ingeniería de sistemas ofrece los conocimientos técnicos necesarios para el desarrollo del software y la integración de las redes neuronales, mientras que la ingeniería industrial se encarga de la optimización de procesos, la gestión de recursos y la evaluación de la eficiencia del sistema en un entorno práctico.

Este marco teórico está cuidadosamente estructurado para ofrecer una comprensión exhaustiva del proyecto, cubriendo todas las definiciones y conceptos necesarios para una investigación detallada. Se divide en varias secciones clave que abordan aspectos fundamentales del diseño e implementación del prototipo de software,

asegurando que todos los elementos relevantes estén claramente definidos y explicados.

7.1. Concepto De Ciber Acoso

El ciberacoso, también conocido como acoso electrónico o acoso virtual (cyberbullying en inglés), se refiere a la intimidación o acoso perpetrado a través de tecnologías digitales. Este fenómeno puede ocurrir en diversas plataformas, como redes sociales, aplicaciones de mensajería, juegos en línea y teléfonos móviles. Se caracteriza por comportamientos repetitivos cuyo objetivo es atemorizar, enfadar o humillar a otras personas (UNICEF, 2024).

7.1.1. Tipos De Ciber Acoso

Existen varios tipos de ciber acoso que pueden ser sutiles y que, como todas las violencias, avanzan en escalada, el desequilibrio emocional que provoca ser víctima de estas circunstancias, daña significativamente a la persona, provocando aislamiento, desánimo o depresión. Estos son cinco tipos de ciberacoso que destacan: (Ciberacoso, Centro de reconocimiento a la dignidad humana, 2021).

- **Ciberbullying:** También llamado ciberacoso escolar, se caracteriza por presentarse generalmente entre menores de edad de manera intencional y reiterada.
- **Sextorsión:** Detrás de este tipo de acoso generalmente se encuentran casos de pedofilia y pederastia, regularmente es derivado del ciberacoso denominado grooming, se caracteriza principalmente por la extorsión hacia la víctima con intenciones de carácter sexual en el que se le amenaza con exponer contenido sexual de la misma.
- **Grooming:** Es el acoso que se presenta de parte de un adulto hacia un menor de edad con intenciones sexuales, generalmente el mayor de edad se hace

pasar por menor de edad para empatizar con la víctima y así ganar su confianza.

- **Ciber violencia de Género**: Sé presenta por una persona o grupo de personas hacia otra u otros del sexo opuesto, en el que se ejerce violencia a través de insultos, acoso, control, ataques, chantaje.
- **Sexting:** Consiste en el envío de imágenes de índole sexual entre dos personas, generalmente de manera consensuada. Si bien esta práctica no constituye un tipo de acoso, la persona que lo practica debe de ser consciente del riesgo que podría suponer el envío de este tipo de contenido ya que podría derivar en sextorsión. (Ciberacoso, Centro de reconocimiento a la dignidad humana, 2021).

7.1.2. Impacto Emocional En Las Victimas

Cuando el acoso ocurre en línea, la víctima siente como si la estuvieran atacando en todas partes, hasta en su propia casa. Puede parecerle que no hay escapatoria posible. Las consecuencias pueden durar largo tiempo y afectar a la víctima de muchas maneras, se explican algunas:

- **Mentalmente.** Se siente preocupada, avergonzada, estúpida y hasta asustada o enfadada.
- **Emocionalmente.** Se siente avergonzada y pierde interés en lo que le gusta.
- **Físicamente.** Se siente cansada (pierde el sueño) o sufre dolores de estómago y de cabeza.

Sentirse objeto de burla o de acoso puede impedir que la víctima hable con franqueza o trate de resolver el problema. En casos extremos, el ciberacoso puede llevar a quitarse la vida. El ciberacoso puede afectar de muchas formas. Sin embargo, es posible superarlo y recuperar la confianza en nosotros mismos y la salud (Unicef, 2024).

7.2. Uso de internet en niños

De acuerdo a la investigación de Common Sense Media (2021) se evidencia que el uso de las pantallas por parte de los preadolescentes (de 8 a 12 años) y los adolescentes (de 13 a 18 años) ha aumentado más rápido en los dos años transcurridos desde la pandemia que en los cuatro años anteriores. La investigación encontró que los niños de 8 a 12 años pasan un promedio de cinco horas y media al día en pantallas y consumiendo contenidos. Esa tasa sube a más de ocho horas y media al día para los adolescentes.

Entre los adolescentes, el 79 % dijo que usa las redes sociales y los videos en línea al menos una vez a la semana, y el 32 % dijo que "no querría vivir sin" YouTube. Y casi dos tercios (65 %) de los preadolescentes dijeron que ven televisión, el 64 % ven videos en línea y el 43 % juegan juegos en un teléfono inteligente o tableta todos los días.

Las tasas promedio diarias de tiempo de pantalla se dispararon más entre los niños negros e hispanos/latinos y los de familias de bajos ingresos. Estos adolescentes y preadolescentes pasaban entre 6,5 y 7,5 horas al día en pantallas de entretenimiento.

7.2.1. Plataformas más utilizadas

- **Redes sociales:** Las conexiones online han aumentado enormemente, especialmente en una generación que casi ha nacido con un móvil en la mano. En la actualidad, la red social más utilizada por los jóvenes entre 13 y 17 años es TikTok, Instagram y Snapchat; sin embargo, para los jóvenes entre 18 y 24 años, son escogidas antes Instagram, Twitter y Snapchat. Como podemos ver día a día, cada vez son más los problemas causados por y en las redes sociales. (Wavemarker, 2021).
- **Apps de vídeo:** Durante el año pasado, las herramientas de vídeo online aumentaron su tiempo de uso entre los menores de edad a nivel global un 27% con respecto a 2022, pasando de 45 a 57 minutos al día. Este incremento se ve justificado por el aumento de tiempo de uso en YouTube,

pues continúa siendo la plataforma favorita de los más jóvenes a nivel global como La plataforma Twitch también ha conseguido aumentar ligeramente su tiempo de uso, pasando de 18 a 22 minutos al día, aun así, sigue siendo la plataforma menos utilizada.El tiempo de uso del resto de plataformas ha disminuido ligeramente. Por delante de Twitch se encuentran Netflix, que en 2023 disminuyó su tiempo de uso de 41 a 38 minutos al día; Disney Plus, que ha pasado de 39 a 34 minutos al día; Amazon Prime, de 40 a 33 minutos al día, y Movistar Plus, de 29 a 26 minutos al día. (Villegas, 2024).

7.2.2. Factores de Riesgos

Entre los riesgos que más preocupa a los padres se encuentran, según concluye el informe, los siguientes: la exposición a contenidos para adultos o pornografía, seguida de cerca por los depredadores en línea, además de la adicción a Internet.

Teniendo en cuenta, además, que el día a día de niños y adolescentes pasa por estar conectados (permanecen nada menos que cuatro horas al día de media delante de una pantalla) y, puesto que lo hacen a edades cada vez más tempranas, “es importante que las familias conozcan el uso que hacen de ellas y les eduquen en una digitalización responsable”. Sin embargo, como bien sabemos, educarles en un uso responsable de internet y las redes sociales no es sencillo, pero pasa por predicar con el ejemplo, “hacerles entender qué deben compartir, con quién lo deben hacer, qué consumir o cuánto tiempo pasar, para que ellos mismos sepan gestionar su tiempo en las redes sociales”. (Villegas, 2024).

7.3. **Rol de los padres en la supervisión**

Los niños y las redes sociales no tienen por qué ser una pareja incompatible. Sin embargo, los padres deben guiar a sus hijos a la hora de usar las redes sociales. En muchos casos, esto puede implicar exponer a los niños a las redes sociales y mantener conversaciones sobre el uso de estas plataformas, sus peligros y riesgos inherentes, mucho antes de que los niños creen sus propias cuentas. Una vez que sea apropiado que empiecen a usar estas plataformas, seguir simples pautas sobre

redes sociales para padres puede ayudar a evitar los peores efectos de las redes sociales en los niños. (Kaspersky, 2024)

En el caso de los niños y adolescentes, la falta de una adecuada atención por parte de los adultos les deja aún más vía libre para acceder sin control a Internet. Si el ordenador o la tableta no dispone de filtros que limiten el acceso a las páginas inadecuadas, de forma accidental o buscando nuevos amigos y estímulos se irán encontrando allí con toda clase de contenidos, servicios y personas, no siempre fiables ni convenientes para todas las edades. Y lo que empieza por curiosidad puede acabar en una adicción ya que los niños y los adolescentes son fácilmente seducibles. Por desgracia hay muchos adultos que no son conscientes de estos peligros, que ya se daban en parte con la televisión y los videojuegos y que ahora se multiplican en Internet, cada vez más omnipresente y accesible a todos en las casas, escuelas, cibercafés entre otros lugares (Ate Gobierno de Canarias , s.f.).

7.3.1. Estrategias de supervisión

Estas son algunas consideraciones:

- Saber qué plataformas de redes sociales usan tus hijos.
- Hacer preguntas, pero de forma no invasiva: muestra verdadero interés por lo que ven en Internet y con quién hablan.
- Si lo consideras apropiado, usa controles parentales en los dispositivos y cuentas de redes sociales de tus hijos.
- Habla con tus hijos sobre la importancia de mantener amistades en persona e interacciones sociales tanto en la vida real como en Internet.
- Verifica las conexiones de tus hijos en las redes sociales y enséñales a identificar cuentas falsas.
- Habla sobre lo que es apropiado compartir en las redes sociales; explícales que todo lo que está en línea es permanente, incluso si se borra.
- Mantén conversaciones abiertas sobre las formas adecuadas de actuar en Internet: es importante que los niños entiendan qué es el acoso en línea, cómo reconocerlo y los efectos que puede tener.

- Los niños deben ver a sus padres y a las redes sociales como algo bueno: sé un ejemplo de buen comportamiento.
- Enséñales a los niños el valor de pasar tiempo sin tecnología ni pantallas.
- Aborda los efectos de las redes sociales en los niños conversando sobre salud mental, acoso en línea, sentimientos de aislamiento y la necesidad de compararse con los demás. (Kaspersky, 2024)

7.3.2. Desafíos en la supervisión

Es fundamental que los padres se mantengan actualizados sobre los juegos virales de internet. La investigación muestra que la interacción digital ha aumentado considerablemente, ya que muchas actividades educativas, laborales y de entretenimiento se realizan en línea. Esto ha llevado a la masificación de retos virales en las redes sociales, y es crucial que los padres comprendan estos fenómenos para guiar y proteger a sus hijos de manera efectiva. Al estar informados, pueden fomentar un uso saludable de la tecnología y ayudar a sus hijos a navegar este entorno digital de forma segura.

Muchos de los juegos que circulan en redes sociales exponen a los niños, niñas y adolescentes a actividades que pueden poner en riesgo su bienestar físico o emocional (TIC, 2020)

7.4. Herramientas tecnológicas para la prevención

El ciberacoso, una problemática cada vez más frecuente en el entorno digital, ha generado la necesidad de desarrollar herramientas tecnológicas que permitan prevenir y mitigar sus efectos. Estas herramientas, diseñadas para monitorear la actividad en línea, detectar comportamientos abusivos y facilitar la denuncia, se han convertido en un aliado fundamental para padres, educadores y autoridades.

Las herramientas tecnológicas para la prevención del ciberacoso se basan en la idea de la detección temprana y la intervención oportuna pueden reducir significativamente el impacto de este fenómeno.

7.4.1. Software de Monitoreo Parental:

Este tipo de software permite a los padres supervisar la actividad en línea de sus hijos, filtrar contenido inapropiado y establecer límites de tiempo. Algunos ejemplos destacados incluyen:

- **Bark:** Esta aplicación utiliza inteligencia artificial para analizar los mensajes de texto, correos electrónicos y redes sociales, alertando a los padres sobre cualquier señal de ciberacoso, contenido sexual, amenazas de autolesión o uso de lenguaje inapropiado (Bark, 2023).
- **Qustodio:** Además de monitorear la actividad en línea, Qustodio permite bloquear sitios web y aplicaciones, establecer horarios de uso y localizar dispositivos móviles (Qustodio, 2023).

7.4.2. Aplicaciones de Reporte:

Estas aplicaciones facilitan la denuncia de casos de ciberacoso, permitiendo a las víctimas y testigos reportar incidentes de manera anónima y segura. Algunas plataformas escolares y redes sociales han integrado herramientas de reporte directo, lo que agiliza el proceso de investigación y sanción.

7.4.3. Inteligencia Artificial y Detección de Ciberacoso:

La inteligencia artificial ha demostrado ser una herramienta poderosa para la detección temprana del ciberacoso. Al analizar grandes volúmenes de datos, los algoritmos de IA pueden identificar patrones lingüísticos, emocionales y contextuales asociados con comportamientos abusivos. Esta tecnología permite:

- **Detección en tiempo real:** Los sistemas de IA pueden identificar y alertar sobre posibles casos de ciberacoso a medida que ocurren.
- **Análisis de sentimientos:** Al analizar el tono y el contenido de los mensajes, los algoritmos pueden determinar si una interacción es hostil o amenazante.
- **Identificación de víctimas:** La IA puede ayudar a identificar a las víctimas de ciberacoso, incluso cuando no lo denuncian de forma activa.

Efectividad de las Herramientas Tecnológicas: Diversos estudios han demostrado la eficacia de las herramientas tecnológicas para la prevención del ciberacoso. Por ejemplo, un estudio realizado por Smith et al. (2019) encontró que el uso de software

de monitoreo parental se asoció con una reducción significativa en la frecuencia y la gravedad de los incidentes de ciberacoso.

- **Limitaciones y Desafíos:** A pesar de sus beneficios, las herramientas tecnológicas para la prevención del ciberacoso también presentan ciertas limitaciones y desafíos:
- **Privacidad:** El uso de estas herramientas plantea preocupaciones sobre la privacidad y la vigilancia.
- **Falsos positivos:** Los sistemas de IA pueden generar falsos positivos, lo que puede llevar a la acusación injusta de individuos.
- **Resistencia al cambio:** Algunos padres y estudiantes pueden resistirse al uso de estas herramientas debido a preocupaciones sobre la autonomía y la confianza.

Las herramientas tecnológicas se han convertido en un componente esencial de las estrategias de prevención del ciberacoso. Al combinar el software de monitoreo, las aplicaciones de reporte y la inteligencia artificial, es posible crear un entorno digital más seguro para niños y adolescentes. Sin embargo, es importante reconocer las limitaciones de estas herramientas y abordar los desafíos asociados con su implementación.

7.5. **Estrategias de intervención y educación**

La educación juega un papel crucial en este proceso, ya que permite desarrollar habilidades sociales, empatía y un clima de respeto entre los estudiantes. Diversos estudios han demostrado la eficacia de programas educativos para prevenir el ciberacoso (Smith, Bradshaw, & Mitchell, 2008). Estos programas se basan en teorías del aprendizaje social, las cuales postulan que los comportamientos se aprenden a través de la observación y la imitación de modelos (Bandura, 1977). Al exponer a los estudiantes a modelos prosociales y enseñarles habilidades de resolución de conflictos, se busca fomentar conductas positivas y prevenir el ciberacoso.

7.5.1. Programas Educativos en las Escuelas:

- **Currículo:** La integración de contenidos relacionados con el ciberacoso en el currículo escolar permite sensibilizar a los estudiantes sobre esta problemática y sus consecuencias (Hinduja & Patchin, 2010).
- **Talleres y Seminarios:** La realización de talleres y seminarios interactivos facilita la adquisición de conocimientos y habilidades prácticas para prevenir y abordar el ciberacoso (Kowalski, Limber, & Smith, 2012).

7.5.2. Participación de Padres y Madres:

- **Talleres Informativos:** La organización de talleres para padres y madres permite brindarles información sobre el ciberacoso, sus señales de alerta y las medidas que pueden tomar para proteger a sus hijos (Livingston, 2009).
- **Fomento de la Comunicación:** Se debe promover una comunicación abierta y de confianza entre padres e hijos para detectar posibles situaciones de ciberacoso a tiempo (Ybarra & Mitchell, 2004).

7.5.3.Creación de un Entorno de Apoyo:

Clima Escolar Positivo: La creación de un clima escolar positivo, basado el en el respeto, la tolerancia y la inclusión, es fundamental para prevenir ciberacoso (Olweus, 1993).

Políticas Escolares: La implementación de políticas escolares claras y efectivas contra el ciberacoso proporciona un marco de referencia para abordar esta problemática y garantizar la seguridad de todos los estudiantes (Hinduja & Patchin, 2015).

7.5.4.Objetivos de la Intervención

Concientización: Aumentar la conciencia sobre el ciberacoso y sus consecuencias.

Prevención: Reducir la incidencia de casos de ciberacoso.

Intervención: Proporcionar herramientas para abordar situaciones de ciberacoso de manera efectiva.

Empoderamiento: Empoderar a los estudiantes, padres y docentes para que actúen como agentes de cambio.

7.5.5. Evaluación de la Intervención

La evaluación de la eficacia de las estrategias de intervención diseñadas para prevenir y reducir el ciberacoso en niños es un aspecto fundamental para garantizar su impacto positivo en el entorno escolar y en el bienestar de los estudiantes. Para llevar a cabo una evaluación rigurosa, es necesario emplear una variedad de instrumentos de medición que permitan recopilar datos tanto cuantitativos como cualitativos, como cuestionarios, entrevistas y análisis de datos cualitativos.

Instrumentos de Medición

Los cuestionarios autoadministrados son una herramienta comúnmente utilizada para evaluar la percepción de los estudiantes sobre el ciberacoso, su frecuencia de participación como víctimas o agresores, y la efectividad de las intervenciones implementadas. Estos instrumentos pueden incluir escalas de Likert para medir actitudes, creencias y comportamientos relacionados con el ciberacoso (Smith et al., 2012).

Las entrevistas individuales o grupales permiten profundizar en las experiencias de los estudiantes, explorando sus emociones, pensamientos y percepciones sobre el ciberacoso de manera más detallada. Además, las entrevistas pueden proporcionar información valiosa sobre las dinámicas sociales y las relaciones interpersonales que influyen en la ocurrencia de este fenómeno (Olweus, 2001).

El análisis de datos cualitativos, como los registros de incidentes, los comentarios de los estudiantes y los informes de los profesores, puede proporcionar una visión más completa de la situación y ayudar a identificar patrones y tendencias que no son evidentes en los datos cuantitativos (Kowalski & Limber, 2013).

Los indicadores de éxito pueden incluir:

- **Reducción de la incidencia de casos de ciberacoso:** A través de la comparación de los datos recolectados antes y después de la intervención, es posible determinar si ha habido una disminución en la frecuencia y gravedad de los incidentes de ciberacoso (Hinduja & Patchin, 2010).
- **Mejora del clima escolar:** Un clima escolar positivo se caracteriza por un mayor sentido de pertenencia, respeto y confianza entre los miembros de la comunidad educativa. La evaluación del clima escolar puede incluir la medición

de variables como la satisfacción con la escuela, la percepción de seguridad y la calidad de las relaciones interpersonales (Rigby, 2016).

- **Aumento de la percepción de seguridad entre los estudiantes:** Los estudiantes que se sienten seguros en el entorno escolar son menos propensos a ser víctimas o perpetradores de ciberacoso. La evaluación de la percepción de seguridad puede realizarse mediante cuestionarios que midan el miedo a ser victimizados, la confianza en los adultos y la percepción de un ambiente escolar seguro (Smith et al., 2012).
- **Desarrollo de habilidades sociales y emocionales:** Las intervenciones efectivas para prevenir el ciberacoso promueven el desarrollo de habilidades sociales y emocionales como la empatía, la resolución de conflictos y la asertividad. La evaluación de estas habilidades puede realizarse mediante observaciones directas, role-playing o cuestionarios (Barba & Del Rey, 2010).

8. ANALISIS DE RESTRICCIONES

Para garantizar el éxito de este prototipo, se ha llevado a cabo un análisis de restricciones que identifica los desafíos técnicos, éticos y operativos que pueden afectar su implementación. Esto incluyó la evaluación de posibles restricciones en aspectos ambientales, económicos, legales, de salud, seguridad y socioculturales, asegurando así que el proyecto sea viable y se pueda llevar a cabo de manera efectiva.

8.1. Restricciones Económicas

- Recursos Financieros: La financiación para el desarrollo y mantenimiento del software puede ser insuficiente, afectando en un gran porcentaje la sostenibilidad de la aplicación que se busca implementar.

- Costo para los usuarios: El precio del software debe ser accesible para la mayoría de las familias teniendo en cuenta que inicialmente se implementara en estratos 2 y 3 y considerando que muchas pueden no tener recursos para invertir en herramientas de protección así les parezca una herramienta muy beneficiosa.

8.2. Restricciones Legales

- Normativas sobre privacidad: En Colombia, la protección de datos de menores de edad está regulada por la Ley 1581 de 2012, que establece que su tratamiento está prohibido, excepto en ciertos casos. Debe respetar el interés superior del menor y sus derechos fundamentales, requiriendo el consentimiento del representante legal. Además, se debe considerar la opinión del menor según su madurez. Las entidades responsables deben garantizar el uso adecuado de estos datos y cumplir con los principios de la ley. (Superintendencia de Industria y Comercio, s.f.)
- Legislación sobre contenido en línea: Los cambios en las políticas de plataformas digitales pueden afectar el software, relacionado con la Ley 1978 de 2019, que moderniza el sector TIC. Esta ley promueve la igualdad entre proveedores y asegura un uso eficiente del espectro radioeléctrico, buscando maximizar el bienestar social mediante estudios técnicos y económicos. (Departamento Nacional de Planeación, s.f.)

8.3. Restricciones de Salud:

- Impacto psicológico: La implementación de alertas frecuentes en un software de monitoreo podría generar un alto nivel de ansiedad y desconfianza tanto en los menores como en los padres, afectando así la salud mental de ambas partes. Los menores, al sentirse constantemente vigilados, pueden presentar un alto nivel de incomodidad en sus interacciones digitales, lo que puede llevarlos a esconder actividades y de cierto modo limitar la comunicación abierta con sus padres; Los padres, al recibir alertas continuas, pueden desarrollar preocupaciones excesivas, generando un mayor nivel de desconfianza y sensación de impotencia ante los riesgos en línea.

8.4. Restricciones Socioculturales:

- Educación: La falta de información sobre cómo usar la tecnología puede hacer que los padres se sientan inseguros y no estén al tanto de los riesgos que enfrentan los niños, lo que los puede llevar a no usar el software.
- Percepción negativa: La preocupación de los padres sobre cómo se ve su vigilancia puede afectar la disposición a utilizar herramientas de monitoreo.

- Efectos secundarios de la vigilancia: La monitorización constante puede llevar a una relación tensa entre padres e hijos.

9. METODOLOGIA PARA LA SELECCIÓN Y DESARROLLO DE LA SOLUCION

9.1. **Diseño**

La investigación que se va realizar para el proyecto consiste en la toma de encuestas a padres de familia para poder contemplar las preocupaciones, conocimientos acerca de los peligros que pueden haber en línea, igual mente conocer el tiempo o supervisión que pasan con los hijos o la libertad que dan respecto al uso de dispositivos tecnológicos con acceso a internet, una vez realizado se estudiara y analizaran los datos dados por estas encuestas para poder entender mejor nuestras dudas y soluciones por parte de los padres al realizar estas mismas, otra forma de obtener datos para poder conocer más se realizara un análisis de documentos sobre los casos en el ciber acoso, manipulación, delitos sexuales y explotación infantil sexual a nivel de Bogotá Colombia, con el fin de obtener un panorama mejor acerca de lo que sucede con los jóvenes relacionado a estas problemáticas, para esto se necesita revisar bastante documentación para poder analizar y contemplar los casos que se han dado y la implicación de las redes sociales o lugares de chat donde hacen que los menores de edad sean llevados a estas prácticas.

9.2. **Muestra**

Nuestra población será Bogotá, esto para poder tener una medida para nuestro proyecto y teniendo en cuenta los limitantes que tendríamos a desarrollarla para un panorama más amplio debido a las dificultades que conllevaría la infraestructura y la cantidad de datos que se deben procesar, puede que en el futuro se pueda ampliar con mejores servidores y mejoras en el procesamiento de datos para ser llevada a una gran escala, por otro lado sabemos que todo menor de edad comprende la edad desde 0 a 17 años, siendo las edades objetivo las comprendidas entre 12 a 17 años de edad, esto debido que son más propensas a caer, ser manipuladas, engañadas o

chantajeadas por internet debido a la cercanía que tienen con esta y el tiempo que pasan inmersos en internet

En Bogotá hay 7.8 millones de personas según la proyección del Dane en su boletín técnico (2021), de las cuales en Bogotá sabemos que hay 568.036 menores de edad de las edades de 12 a 17 años encuesta realizada por tele encuestas (2023), donde de estas el 74.3% pasan tiempo en internet, celular y computadora esto evidenciado en las nota estadística juventud en Colombia de Dane (2021), teniendo una población de 422.051 menores que cumplen con la condición anterior, también conocemos que los estratos 2 y 3 es donde mayor mente se concentra los menores de 13 a 17 años siendo estos un total de 25.05% esto observado por la secretaria distrital de planeación en su análisis de encuestas a familias por distrito (2018), teniendo así una población Objetivo de 105.724, para este caso con un nivel de confianza del 95% y un margen de error del 5%, mediante la ayuda de una calculadora de muestras podemos obtener que nuestra muestra indicada pra el estudio y que cupla con lo anterior mencionado es de 383.

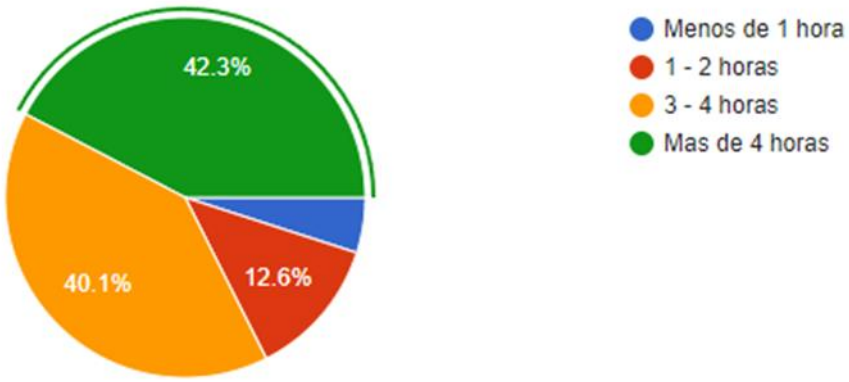
9.3. **Recolección de datos**

Se usará para la parte de encuestas el formulario de Google Forms para la realización más sencilla y cómoda como también la distribución de la misma, igualmente de manera presencial realizando las preguntas para poder agilizar con la cantidad de encuestas a realizar y poder garantizar el cumplimiento de las mismas, igualmente se realizó una búsqueda para conocer documentos sobre los crímenes o acasos que se dan en la capital a menores por medio de internet teniendo en cuenta las fuentes de la información que sean legitimas y que tengan un soporte estricto para evitar dañar el análisis o tener información errónea que dañe o altere los resultados para su posterior análisis

9.4. **Encuestas**

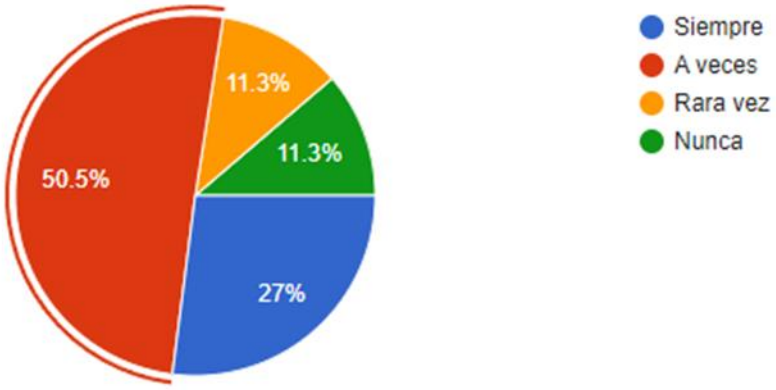
Al realizar las encuestas realizamos un filtrado para poder conocer cuántos hijos tenían cada padre, como también conocer si los hijos tenían la edad de entre 12 a 17 años ya que es el objetivo de la encuesta, como también el estrato del hogar.

Al realizar la encuesta nos encontramos que a la pregunta de **¿Cuantas horas pasa su hijo en dispositivos electrónicos?** La siguiente grafica



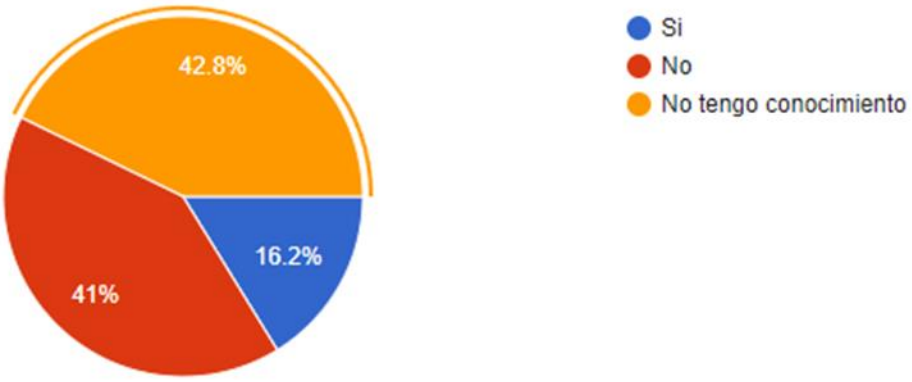
Siendo el 42.3% la respuesta más votada junto a la de 3 – 4 horas evidenciando el gran uso que tienen los adolescentes en el uso de dispositivos tal como se había venido planteando a lo largo del proyecto y dando validez, igual mente al momento de preguntar si se tiene internet en casa se tiene que el 92.8% si tienen internet en casa.

Algo que notamos y es signo de preocupación es la frecuencia de supervisión de los hijos en internet teniendo la siguiente grafica **¿Con que frecuencia supervisa el uso que su hijo hace de internet?**

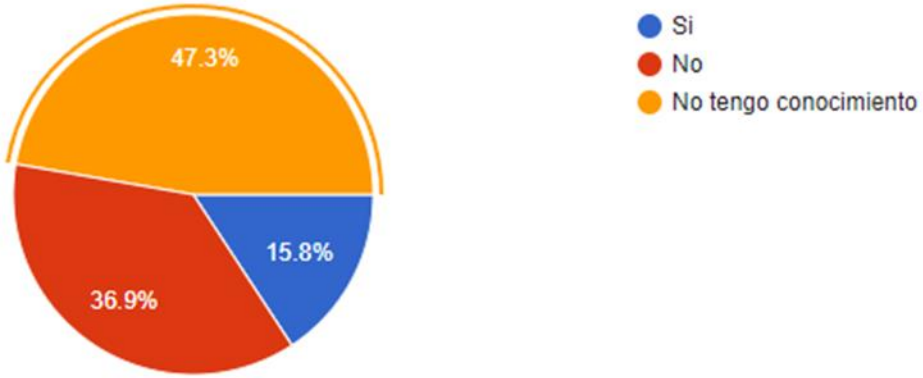


Donde el 50.5% a veces revisan dejando mucho margen de maniobra para personas con malas intenciones y si a esto le sumamos el desconocimiento como evidenciamos en estas dos preguntas:

¿Ha tenido su hijo alguna vez contacto en línea con desconocidos?



¿Su hijo alguna vez le ha mencionado que le piden algo mientras está en internet que le ocasione preocupación?

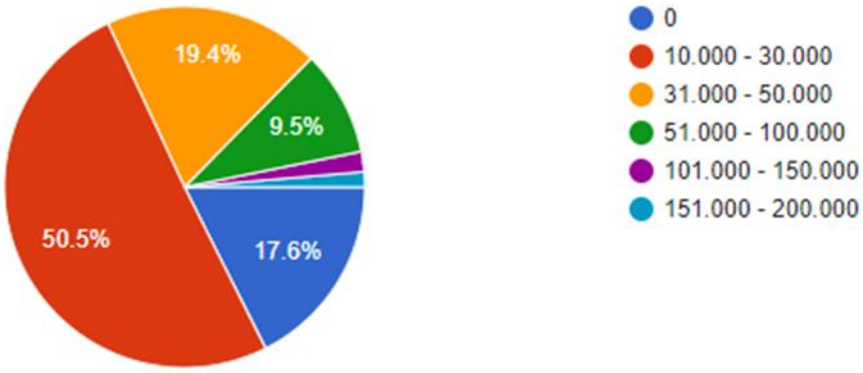


Con todo lo investigado e indagado no es de extrañar que los menores puedan recibir contenidos no apropiados para ellos o conozcan personas desconocidas con la gran magnitud de movimientos que hay en las redes sociales lleguen a conversar con estas personas.

Durante la recolección de datos para la realización de las encuestas se preguntaron palabras usadas por depredadores sexuales usadas en insinuaciones o para búsqueda de contenidos ilegales, donde se encuentra un gran desconocimiento por parte de las personas encuestadas mostrando un gran umbral de desconocimiento de estos términos que puden con llevar a pasar por alto al momento que leen o revisan el celular de sus hijos sin saber el peligro que esto pueda con llevar, siendo solo dos personas del total de encuestas capaces de discernir un término entre los varios encuestas siendo este 'Cheese Pizza', atribuyendo su conocimiento a un creador de contenidos de Youtube, DrossRotzank, siendo los otros terminos total mente

indiferentes para ellos y los demás encuestados 'Caldo de Pollo,' 'Camión Pesado,' 'Club Penguin,' y 'CP'.

En la parte de presupuesto se realizó una pregunta para saber lo que estarían dispuestos los padres pagar para poder tener una aplicación que les ayude a proteger a sus hijos en línea, dándonos la siguiente gráfica



Donde observamos que con el 50.5% estarían dispuestos a pagar entre 10.000 y 30.000 pesos seguido del 19.4% que sería entre 31.000 y 50.000 pesos, esto nos lleva a dimensionar si quisiéramos pasar de un prototipo a un proyecto más real y sentado para producirlo, que tendríamos que tener en cuenta le valor dispuesto a pagar y así mismo formular los costos que se tendrían en el proyecto como lo puede ser servidores para el manejo de la aplicación datos, el costo por el uso de la IA como también el almacenamiento de la información de los logs como la seguridad asociada a estos, tema que se puede incluso hondar mucho más si se continua el proyecto después de la presentación del prototipo

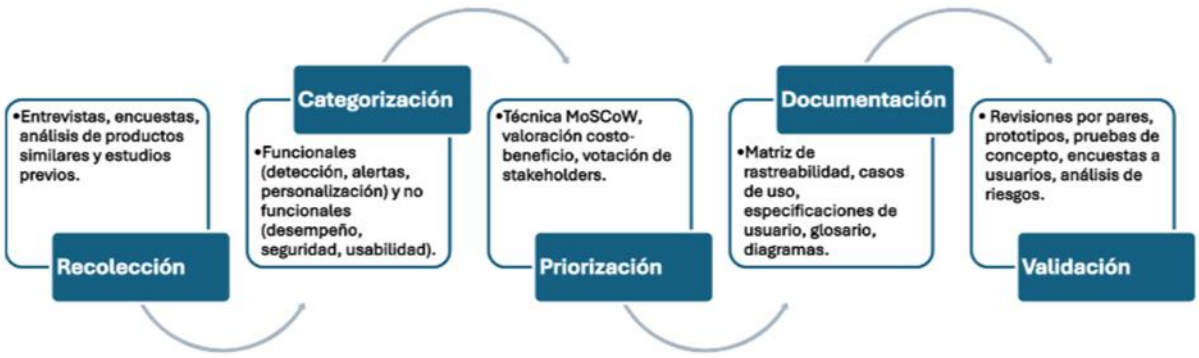
En general encontramos que los padres de familia en una gran mayoría a veces supervisa los dispositivos de sus hijos donde aunque se tenga el desconocimiento de si conocieron personas desconocidas o pidieron algo, aunque revisaran más a menudo pasarían muchas cosas por alto, al ser temas que son inocentes y que sin el conocimiento no serían preocupantes o llamativos para ellos, llevando que inclusive al estar supervisando omitan estas partes importantes y dejando aun sus hijos vulnerables ante desconocidos mal intencionados, llevando la necesidad de dos puntos, uno donde se concientice más a los padres de estas prácticas y peligros que corren sus hijos en internet que inclusive se realizó una pregunta si les ¿gustaría

saber más de los peligros que corren sus hijos en línea?, teniendo una respuesta bastante favorable del 86%, nos demuestra que si estarían interesados para poder aprender conocer, y comprender los peligros que pueden haber por internet incluso conociendo palabras claves que con ese conocimiento pueden impedir muchos peligros en sus hijos, y segundo la necesidad de una aplicación que les ayude a supervisar a sus hijos cuando ellos no estén y que pueda identificar estas palabras claves como patrones para la notificación a los padres y conocer si su hijo puede tener algún peligro.

Igual mente se refuerza la idea del tiempo que pasan los menores de edad en los dispositivos como también el uso de internet en las casas mostrando que el peligro que corren no solo es desde el colegio como se hablara en la documentación encontrada si no también en la casa, donde siguen conectados y expuestos a gran peligro.

Las alternativas serán evaluadas en función de los siguientes criterios: (1) Accesibilidad, (2) Precio, (3) Detección temprana, (4) Facilidad de uso y (5) Privacidad. En criterios de calidad: (1) Algoritmo de detección, (2) Interfaz intuitiva, (3) Encriptación de datos, (4) Escalabilidad y como método de logro: (1) Aprendizaje automático, (2) Diseño centrado a usuario, (3) Protocolo de seguridad y (4) Arquitectura en la nube

La metodología se desarrollará en cinco fases: (1) Recolección de datos en campo, (2) Análisis de necesidades, (3) Evaluación de alternativas, (4) Selección y validación de la mejor solución, y (5) Documentación final del diseño.



9.5. **Documentos y análisis**

En Colombia Bogotá tenemos aproximadamente 16.711 casos de asuntos procesales de menores de edad donde los principales 10 asuntos se encuentra en primer lugar la violencia sexual que abarca desde actos, acosos, acceso y entre otros, que abarca 16.221 menores siendo del 87% víctimas mujeres y el 13% Hombres cifras tomadas del Boletín estadístico dirección de protección Bienestar Familiar (2022), al igual es preocupante que de 30 países Colombia ocupe el top 10 con mayores acosos escolares donde en el país se reportaron 8.981 casos siendo Bogotá el 21% es decir 1.886 casos de acoso escolar en Bogotá datos tomados del senado de la república (2022), para apoyo de lo anterior mente mencionado podemos remitirnos a la escalada que se ha dado del 2021 con un total de 5.995 casos de abuso y violencia en los colegios de Bogotá a pasar en el 2022 a 20.506 casos siendo esta una cifra preocupante dado que estamos hablando de un aumento de casi 3.4 veces lo del año anterior donde encontramos a la población de menores de 12 a 17 años con el 70.2% de los casos, siendo de todo este total el 6.1% en redes sociales y 53.2% en colegios datos tomados del hostigamiento escolar en las instituciones educativas de Bogotá (2023), si solo tomáramos las redes sociales de los 20.506 casos que hubieron en el 2022 sabiendo que de estas el 70.2% fueron menores de 12 a 17 años estaríamos hablado de un total de 878 casos donde está el factor de las redes sociales, incluso si seguimos mirando en 2023 en el mismo documento nos menciona que las redes subieron al 10% y los acasos en el rango de edades anterior mente dicho subió a 78%, dando una gran preocupación en el aumento que se da año tras año, pero tampoco no hay que olvidar que al ser los casos en colegios del 53.2%, no solo se limitarían al acoso físico si no virtual al estar tan conectados hoy en día no olvidemos que en este rango de edades el uso de celulares está en 86.8% cifra tomada de nota estadística juventud colombiana Dane (2021), dándonos así un indicio que el acaso continua aun fuera del colegio incluso en la casa, esto respalda la necesidad de tener un alertamiento a los padres sobre lo que sucede con sus hijos, ya que muchas veces no denuncian por el miedo o no saben cómo actuar sea por la manipulación o amenazas, siendo esto una herramienta muy útil para ayudar a los padres para saber la situación de sus hijos, incluso de los hijos para salir de situaciones que no saben manejar, incluso mirando en un macro la necesidad misma de poder reducir estos

aumentos tan grandes en los acosos al detectar de manera oportuna cuando se están iniciando y teniendo este seguro en los menores.

Otra gran preocupación se da por la explotación sexual infantil donde se toma a jóvenes por engaños, manipulaciones o amenazas esto es una preocupación que cada es más extensa y más preocupante para Bogotá, algo que evidenciamos en diagnóstico y principales políticas, programas y planes para la lucha contra la explotación sexual comercial de niñas, niños y adolescentes (2023), donde ESCNNA significa Explotación sexual comercial de niñas y adolescentes donde encontramos que de las denuncias que se han hecho en Bogotá desde el 2020 a 2023 llega a 228 víctimas, esto teniendo encuentra que no todos denuncian pero lo más preocupante viene cuando conocemos que de esta cifra el 86.46% corresponden a menores de 12 a 17 años de edad dejando que de las 228 victimas 197 son las víctimas de explotación sexual en Bogotá, esto puede llevar a pensar por los entornos o incluso por otros motivos ajenos de muchos menores de edad y más cuando se tiene en cuenta que solo son 197 víctimas de 4 años sin embargo hay que entender que no todos denuncian estos casos pero donde se rompe esta ilusión o pensamiento de tranquilidad para muchos padres es cuando vemos que hay abusos sexuales y violentos en escuelas y que este como al igual que el índice en acosos escolares lo mismo sucede con los abusos sexuales como se evidencia en violencias basadas en género y violencia sexual en las instituciones educativas de Bogotá (2023), donde en 2021 teníamos 3.371 casos a 2022 pasar a 11.292 un aumento de casi 3.3 veces mayor al año anterior ya en el 2023 teneos 11.301 casos donde el 36,48% fue por violencia sexual donde el 60.76% correspondes a menores de edad de 12 a 17 años de edad, demostrando de esta manera que incluso en las escuelas algo tan cercano y donde mayor tiempo pasan pueden también tener peligros si a esto le sumamos que si son acosados o son amenazados o incluso chantajeados muchos estén callados asustados o con miedo de represalias encerrados en un ciclo de violencia sea sexual, física o psicológica, esto apoya más la necesidad de tener un notificador para estos casos y alertamientos de los padres, tenemos que también saber que no solo personas desconocidas pueden llegar a los hijos si no incluso mismos compañeros pueden atentar contra ellos y más con la cantidad de uso del celular que les permite el chantaje, amenaza o manipulación, donde por redes sociales es tan

sencillo estar en contacto con sus víctimas o incluso entregárselas a otras personas con fines más peligrosos.

9.6. **Categorización - Matriz de Calidad de la Función (QFD)**

Para garantizar que la herramienta de detección temprana de ciberacoso sea efectiva y satisfaga las necesidades de los usuarios, se empleará la Matriz de Calidad de la Función (QFD). Esta herramienta nos permitirá traducir las voces de los padres, educadores y expertos en ciberseguridad en requisitos técnicos específicos para el desarrollo de la herramienta. A través de la QFD, podremos identificar y priorizar las características más importantes de la herramienta, como la precisión en la detección, la facilidad de uso, la rapidez de respuesta y la protección de la privacidad. Al mapear estas características con las necesidades de los usuarios, se podrá diseñar una solución integral que no solo detecte el ciberacoso de manera temprana, sino que también brinde el apoyo necesario a las víctimas y sus familias.

Los beneficios de utilizar la matriz QFD asegura que el producto satisfaga las necesidades del cliente, mejora la comunicación entre los diferentes equipos involucrados en el desarrollo del producto, ayuda a identificar y priorizar las características más importantes y reduce el tiempo y los costos de desarrollo.

La Matriz QFD se representa visualmente como una "casa" y se divide en varias secciones:

- **Techo:** Aquí se establecen las relaciones entre las características técnicas (Columnas) y las necesidades del cliente (Filas). Se utilizan símbolos para indicar la fuerza de la relación (fuerte, media, débil).
- **Votos de los clientes:** En esta sección se cuantifican las necesidades del cliente, asignando un peso a cada una según su importancia.
- **Características técnicas:** Se enumeran las características técnicas del producto o servicio que permitirán satisfacer las necesidades del cliente.
- **Cómo se lograrán:** Se indican los métodos o procesos que se utilizarán para lograr cada característica técnica.
- **Competidores:** Se evalúa cómo los competidores abordan las necesidades del cliente y las características técnicas.

- **Relación entre características técnicas:** Se analiza cómo las diferentes características técnicas se relacionan entre sí.

Relación entre los elementos de la matriz:

- **Necesidades del cliente y características técnicas:** Las características técnicas son los medios para satisfacer las necesidades del cliente.
- **Características técnicas y métodos de logro:** Los métodos de logro son las acciones específicas para obtener las características técnicas.
- **Características técnicas y competidores:** La comparación con los competidores ayuda a identificar oportunidades de mejora.
- **Relación entre características técnicas:** Al analizar las relaciones entre las características técnicas, se pueden identificar posibles conflictos o sinergias.

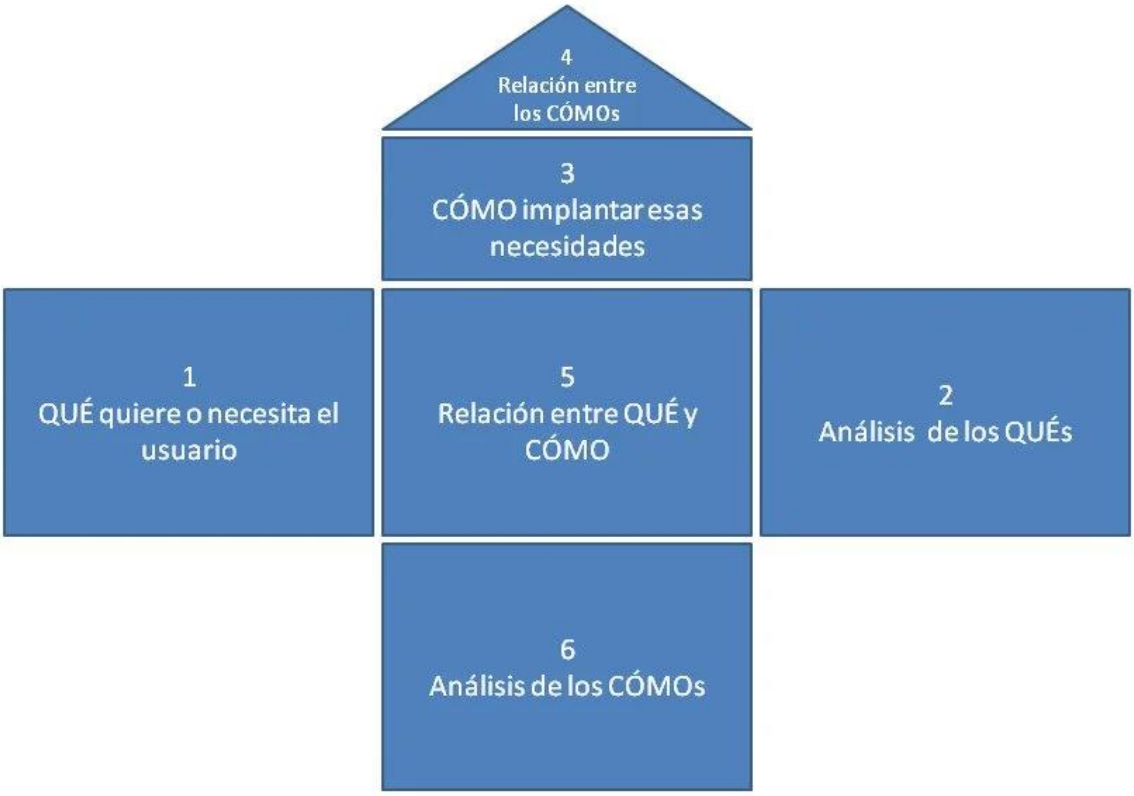


Ilustración 1. Relación de los elementos de la matriz QFD

.

Para una herramienta de detección de ciberacoso:

- **Necesidades del cliente:** Detección temprana, facilidad de uso, privacidad, bajo costo.
- **Características técnicas:** Algoritmos de detección, interfaz intuitiva, encriptación de datos, escalabilidad.
- **Métodos de logro:** Aprendizaje automático, diseño centrado en el usuario, protocolos de seguridad, arquitectura en la nube.

Tabla 1.Análisis de las herramientas competencia y necesidades de los usuarios.

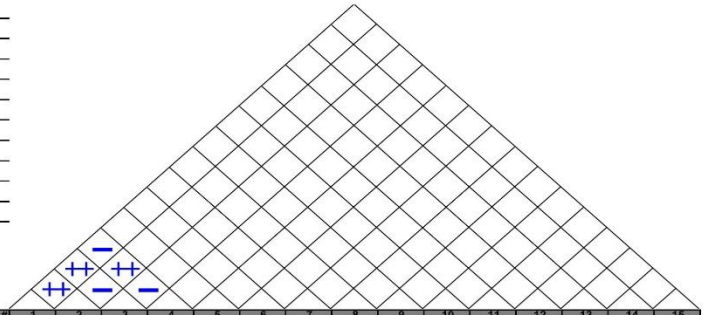
Herramienta	Accesibilidad	Precio	Detección temprana	Facilidad de uso	Privacidad
Qustodio	Amplia gama de dispositivos (iOS, Android, Windows, macOS)	Planes gratuitos y de pago	Monitoreo de actividad en tiempo real, alertas personalizadas	Interfaz intuitiva, configuración flexible	Cifrado de datos, informes detallados
Norton Family	Amplia compatibilidad	Planes de pago	Bloqueo de sitios web, filtros de contenido, informes de actividad	Fácil de configurar y usar	Protección de datos, cumplimiento de normativas de privacidad
Net Nanny	Amplia compatibilidad	Planes de pago	Bloqueo de contenido, seguimiento de ubicación, límites de tiempo	Interfaz sencilla, configuración personalizada	Cifrado de datos, protección contra malware
Google Family Link	Exclusivo para dispositivos Android y iOS	Gratuito	Control de tiempo de pantalla, ubicación, aprobación de aplicaciones	Integración con cuentas de Google, fácil de usar	Políticas de privacidad de Google
Safe Family de McAfee	Amplia compatibilidad	Planes de pago	Bloqueo de sitios web, filtros de contenido, informes detallados	Interfaz intuitiva, configuración flexible	Protección de datos, cumplimiento de normativas de privacidad

A continuación, podrán observar el análisis realizado con las necesidades del cliente, criterios de calidad, métodos de logro y la competencia, que por resultado se obtuvo que es muy importante para los clientes atender con una herramienta que este construida con un algoritmo adecuado para brindar una accesibilidad, detección temprana y facilidad de uso. En segundo lugar, la encriptación de datos juega un papel fundamental para brindar un servicio seguro y que genere confianza para su uso al usuario. Finalmente, una interfaz y escalabilidad son cruciales, pero a nivel intermedio;

El análisis de la competencia indica que Qustodio es un referente en el mercado, destacando por su amplia compatibilidad, opciones de precios flexibles y funcionalidades de detección temprana. La herramienta logra los principios de nuestra solución, que brinda una accesibilidad para todo tipo de dispositivos, cuentan con planes gratuitos y pagos, alerta de detección temprana en tiempo real, una interfaz flexible y a nivel de seguridad de datos, realizan cifrado de datos e informes personalizados. La solución propuesta pretende ser inicialmente usada en un limitado acceso de dispositivos, libre acceso o gratuito, con alertas de detección temprana en tiempo real, con una interfaz intuitiva y fácil uso, y un tratamiento de datos encriptados que asegure su seguridad que significan las conversaciones de los menores de edad. Por otro lado, el método de logro más destacado es el aprendizaje automático de la herramienta, ya que interviene directamente en la generación de alertas por nuevas palabras, frases y métodos utilizadas por los acosadores.

Ilustración 2. Matriz QFD para identificar las necesidades de los usuario

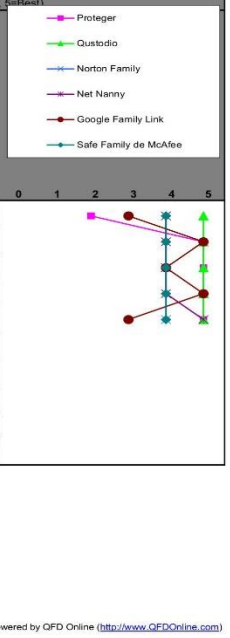
Title: Proyecto Integrador
Author: Geovanni Gil, Soed Rodríguez , Yerly Alejandra Sánchez
Date: 29 de Noviembre 2024
Notes:



Legend

⊙	Strong Relationship	9
O	Moderate Relationship	3
▲	Weak Relationship	1
++	Strong Positive Correlation	
+	Positive Correlation	
−	Negative Correlation	
▼	Strong Negative Correlation	
▼	Objective Is To Minimize	
▲	Objective Is To Maximize	
x	Objective Is To Hit Target	

Row #	Max Relationship Value in Row	Relative Weight	Weight / Importance	Quality Characteristics (a.k.a. "Functional Requirements" or "Hows") / Demanded Quality (a.k.a. "Customer Requirements" or "Whats")	1 Algoritmo de detección	2 Interfaz intuitiva	3 Encriptación de datos	4 Escalibilidad	5	6	7	8	9	10	11	12	13	14	15	Competitive Analysis (0=Worst, 5=Best): Proteger	Qustodio	Norton Family	Net Nanny	Google Family Link	Safe Family de McAfee
				Direction of Improvement: Minimize (▼), Maximize (▲), or Target (x)	▲	▼	▲	x																		
1	9	15,8	3	Accesibilidad	⊙	⊙	▲	▲												2	5	4	4	3	4
2	9	10,5	2	Precio	▲	▲	▲	⊙												5	5	4	4	5	4
3	9	26,3	5	Detección temprana	⊙	O	⊙	▲												5	5	4	4	4	4
4	9	26,3	5	Facilidad de uso	⊙	⊙	▲	⊙												5	5	4	4	5	4
5	9	21,1	4	Privacidad	O	▲	⊙	▲												5	5	4	5	3	4
6																									
7																									
8																									
9																									
10																									
				Target or Limit Value	Aprendizaje automático	Diseño centrado al usuario	Protocolos de seguridad	Arquitectura en la nube																	
				Difficulty (0=Easy to Accomplish, 10=Extremely Difficult)	8	2	10	8																	
				Max Relationship Value in Column	9	9	9	9																	
				Weight / Importance	689,5	489,5	478,9	394,7																	
				Relative Weight	33,6	23,8	23,3	19,2																	



Powered by QFD Online (http://www.QFDOnline.com)

9.7. Prototipo

El prototipo desarrollado trabaja con el lenguaje de Python, mediante la comunicación con Chat GPT para la revisión de mensajes y alertamiento si se llega a encontrar una alerta que haga saltar a Chat GPT, a continuación, se detalla todos los pasos que se realizan y se ven involucrados en el proceso del prototipo.

Toma de datos:

Se inicia desde el proceso de ejecución donde en todo momento estará captando lo que se esté escribiendo independientemente de la aplicación que se esté usando, esta guarda en un log la información que está captando para poder ser enviada a chat GPT una vez cumpla la condición dada para el programa, esto se hace de esta manera precisamente por si se está en redes sociales que, aunque se borre los mensajes no se borran del log que guarda

Condición:

La condición del envió se realiza cada 30 espacios, esto quiere decir que cada vez que se detecte que se realizaron 30 espacios enviara esta información a chat GPT para poder ser analizada, esto con el fin de tener suficiente data para su análisis como también evitar un retraso muy grande en la comunicación detección y aviso al usuario, igual mente para el tema de costos asociados por token y modelo seleccionado para su funcionamiento.

Chat GPT:

Para el uso de esta herramienta es por medio de Open IA que es la creadora de esta donde nos ofrecen precios por token los cuales se pueden inferir por palabras para el entendimiento de estos mismos, donde de acuerdo con el modelo usado tendrá un costo mayor o menor para cada palabra que se le esté entregando, para este caso se realiza con el modelo más nuevo siendo GPT4 para un análisis más real detallado y con menos errores, por ende, caso siendo también más costoso a una escala mucho mayor, aunque no se descarta el uso de GPT3.5 debido que en las pruebas realizadas demuestra también una gran tasa de éxitos.

Comunicación con Chat GPT

Para la comunicación se usa la librería de openai, mediante la api_Key proporcionada mediante la cuenta registrada de openIA, esto con el fin de que se pueda realizar la conexión con el programa y contabilización de los tokens, una vez con esto se realiza un prompt, donde se detalla totalmente la funcionalidad a realizar, donde se le da un listado de palabas claves, textos de alerta, donde incluso se realiza añadiendo una temperatura = 0, esto quiere decir que todo lo que le entreguemos lo trate de responder de la manera más exacta posible, como también un análisis de sentimientos entre positivo, neutral y negativo, para poder tener una análisis del mensaje con el sentimiento que este puede ser, si es peligroso siendo negativo si es de carácter tranquilo positivo o si es neutral como neutro, también la extracción de palabras claves del texto proporcionado y un resumen con la mayor discreción y tacto posible para al momento de que lo lean los padres o tutor legal del menor, todo esto junto con casos para que se tenga una buena información para el entendimiento y posterior análisis de la IA.

Todo con el fin que si llega a encontrar una alarma entre los mensajes nos indique una alerta donde se mencione un resumen discreto y con el tacto posible para enviar por correo electrónico como también las palabras claves dentro del texto y su análisis de sentimiento.

Correo electrónico:

Para el caso del prototipo se tiene solo una dirección de correo electrónico para él envió de estos mensajes donde solamente si es una alerta se envié el mensaje y si en dado caso no lo fuera omitirá él envió de esta.

Este es el proceso del prototipo en la extracción de datos, indicación a IA, análisis de datos y resultados de los datos suministrados, para la protección de menores de edad.

Consideraciones de seguridad fuera de prototipo

Se tiene un login para la aplicación y autentificación del padre de familia o tutor sin embargo sabemos que este no es la única forma ni la mejor de proteger los datos y más cuando son menores de edad, por eso se podría implementar un sistema de autentificación de dos pasos para garantizar más seguridad en el inicio de sesión, para el tema de los mensajes el cifrado extremo a extremo de los datos enviados

como almacenados para garantizar seguridad si se llega a interceptar los mensajes, la asignación de roles para los niveles de acceso que pueden tener tanto para los padres como parra los que mantendrían la aplicación, utilización de servidores con HTTPS para la protección del servidor y cliente.

9.8. Resultados

Se realiza las pruebas en el prototipo se utiliza el modelo de chat GPT 4.0, para poder tener un mejor resultado sin embargo al comprarlas con el GTP3.5 turbo no se tiene mucha divergencia siendo posible el uso de esta para poder economizar costos al usar un modelo más económico sin afectar la calidad del análisis.

Se tiene una meta para aprobación de éxitos que sea mayor o igual al 95%, donde se realiza unas 29 pruebas en total donde se verifica en diversos escenarios siendo el resultado de estas pruebas una tasa de éxitos del 96.55%, esto debido a un escenario que no se había previsto ya que no hace parte de la implementación del prototipo e investigación realizada, que consta de peligros fuera del acoso sexual, acoso, manipulaciones que es el consumo de sustancias ilícitas para el menor, donde pueden ofrecer o intentar consumirlas siendo un escenario donde se sugería el uso de alcohol para un menor de edad no reportando este caso debido a la configuración que se tenía indicando en el resumen que estaba preocupado por el consumo y que lo descubrieran sus padres pero no haciendo reporte, siendo de esta manera que se realiza una nueva configuración en los parámetros para garantizar que este caso si se llegara a presentar reporte esto al tutor legal o los padres para que ellos estén consientes de estos comportamientos como al igual de drogas siendo una segunda prueba con las que se tenían siendo probada la que no reporto y una que indica indicios de consumir drogas pasándolas ambas al reportar e indicar el peligro que podría estar el menor llegando así a una tasa del 100%, sin embargo no se garantiza el 100% debido que como sucedió en este caso se presente otros casos donde no se dé el resultado esperado al estar fuera del ámbito de la investigación y prototipo pero dejando un precedente para una implementación futura para la agregación de más campos y también para que los mismos clientes puedan informar de estos casos y así alimentar mejor la base de casos que pueda cubrir la aplicación.

Para el desarrollo de las pruebas se comete faltas ortográficas consientes para comprobar el uso de la IA sin un contexto claro para comprobar el análisis que se

puede dar, se le da diversos casos donde el analiza estos casos, da el resumen de estos sin llegar a ser explicito y de manera corta, donde también se indica el sentimiento que genera este mensaje, si es enviado a los padres función que solo está disponible en el programa y no enviado a los padres y por ultimo las palabras claves del mensaje, siendo ejemplo de esto los siguientes casos:

1. “no quiero que nuestra amistad se arruine por esta foto que me pides no quiero hacerlo y ddeja esto ya si si insistes le dire a mis papas”

   Resumen por IA: “La conversación muestra una negativa a enviar una foto y una advertencia de informar a los padres si la insistencia continúa.”, generando un sentimiento negativo, con las palabras claves “Foto” y “Papas”. Caso 4.

2. “te enviare la foto pero quieero que este claroq eu no estoy dea cueerdo con esto pero aun asi lo hare olvidate de mi despues esa foto es intima”

   Resumen por IA: “La conversación sugiere un envío de foto íntima bajo presión, con un intento de establecer límites después del envío”, generando un sentimiento negativo con palabas claves “Fotos” y “Intimas”. Caso 15.

3. “no me gusta que me hables asi pero esta bien te perdono no lo hagas mas vale esperemos todo salga mejor amor te adoro perdoname mis celos precioso mio corazon”

Resumen por IA:” La conversación muestra un tono de reconciliación y disculpas por celos previos”, con sentimiento neutral y ninguna palabra clave siendo este omitido y no enviado, Caso 3.

4. “chesse pizza jajajaja acaso quieres comer pizza no entiendo tu peticion de fotos asi esta bien buscare por que perdi la apuesta pero no te entiendo jaja raro”

Resumen IA: “La conversación menciona una petición de fotos en un contexto confuso y posiblemente inapropiado”, generando un sentimiento negativo con palabra clave “Fotos”. Caso 23.

Estos como muestras de lo que reviso la IA y los análisis que se tuvieron de estas siendo de nuevo la aclaración que la ortografía fue echa apropósito para el análisis de la IA ya que en el momento de estar activa no será común recibir textos con buena ortografía, estas pruebas se pueden encontrar en él anexo.

10. ANALISIS DE COSTOS

El análisis de costos es esencial para garantizar que el prototipo de software, diseñado para detectar riesgos en la comunicación en línea de menores, sea viable y sostenible. En esta sección, se detallarán los costos involucrados en las diferentes etapas del desarrollo, que abarcan aspectos como el personal, la infraestructura y las pruebas.

Este análisis tiene como objetivo ofrecer una visión clara sobre las inversiones necesarias para implementar y mantener la herramienta. Así, se logrará tener un panorama completo que facilite la asignación de recursos. Además, ayudará a identificar oportunidades para optimizar costos, asegurando que el proyecto cumpla de manera efectiva con sus objetivos detallados en este documento

10.1. Costo de la Inteligencia Artificial (IA)

La herramienta "Proteger" utilizará GPT-3.5, un modelo de IA avanzado diseñado para analizar textos y detectar patrones o palabras clave relacionados con amenazas en línea. Para comprender el costo de su uso, es necesario entender cómo funciona el proceso de tokenización y el cobro de esta tecnología.

Un token es la unidad básica de procesamiento de texto que el modelo utiliza para comprender y generar lenguaje. Los tokens pueden ser palabras completas, fragmentos de palabras, o incluso caracteres de puntuación. (Jurafsky & Martin, 2021). El uso de esta IA se mide en función del número de tokens procesados, lo que

implica que el costo variará dependiendo del volumen de texto que se analice y de la complejidad de los patrones identificados.

En este proyecto, se estima que cada usuario generará alrededor de **1,000,000 tokens al mes**. Con una base de usuarios aproximada de **650 personas**, el total de tokens procesados al mes será de **650,000,000 tokens**. El costo de procesamiento de estos tokens es un valor determinado por la tarifa que cobra el proveedor de la API.

- **Costo mensual por usuario en USD**: El costo por usuario se calcula multiplicando el número de tokens procesados por mes (1.000.000) por el costo de cada 1.000 tokens (USD 0,002).
- **Tasa de cambio usada**: Para convertir el costo mensual en USD a COP, utilizamos el tipo de cambio de **$4.300 COP por cada USD**:
- **Costo mensual total para todos los usuarios**: Con **650 usuarios** en el sistema, el **costo mensual total** para procesar los tokens de todos los usuarios.

Datos	Valor
Costo por 1.000 tokens	USD 0,002
Numero de tokens mensuales por usuario	1.000.000
Valor para calculo cambio de USD A COP	$ 4.300
Número de usuarios	650

Concepto	USD	COP
Costo mensual por usuario.	USD 2,00	$ 8.600
Costo mensual para 650 usuarios.	USD 1.300,00	**$ 5.590.000**

Tabla 2.Costo mensual de la IA Chat GPT 3.5.

10.2. Costo del Servidor

El servidor es otro componente clave en la infraestructura del sistema "Proteger", ya que debe garantizar alta disponibilidad, rendimiento y seguridad para los usuarios.

Después de evaluar diversas opciones, se decidió optar por el servicio de *HostGator*, que ofrece un plan adecuado a las necesidades del proyecto.

El costo mensual por el servidor es de **$207,050 COP**. Este servicio incluye almacenamiento HDD, que optimiza el rendimiento del sistema, y medidas de seguridad robustas contra ataques cibernéticos, asegurando una alta disponibilidad y un servicio confiable 24/7. Es esencial que el servidor cumpla con las estrictas políticas de protección de datos, lo cual fue un factor determinante en la elección de este proveedor. De este modo, la infraestructura del servidor será escalable y podrá adaptarse al crecimiento de la demanda sin comprometer la calidad del servicio.

10.3. Costo de Personal para la Implementación y Gestión del Proyecto

La contratación de personal especializado es fundamental para el éxito del proyecto "Proteger". Se ha decidido incorporar un equipo de tres profesionales, quienes estarán a cargo de la implementación, desarrollo y gestión del sistema. Los roles específicos incluirán, Ingeniero en seguridad cibernética, Ingeniero de Marketing y un gerente de proyecto, quienes trabajarán de forma colaborativa durante todas las fases de la ejecución.

Cada uno de los integrantes del equipo recibirá un sueldo mensual de **$2,300,000 COP**, lo que representa una inversión mensual de **$6,900,000 COP** por concepto de salarios. La contratación de estos profesionales garantiza que el proyecto sea liderado por expertos en sus respectivos campos, lo que aumentará las probabilidades de éxito y garantizará que el sistema funcione correctamente y cumpla con los estándares esperados de seguridad y eficiencia.

10.4. Gastos Imprevistos

Como es habitual en proyectos de esta índole, es necesario prever un margen para gastos imprevistos, que pueda cubrir cualquier eventualidad o necesidad adicional durante el proceso de desarrollo. Este monto se estima en un **5%** del total de los costos previamente identificados.

A partir de los costos detallados anteriormente, el monto total estimado para el primer mes de operación es de **$12,697,050 COP**. El 5% de este valor equivale a **$634,852 COP**, lo que asegura un colchón financiero para cubrir cualquier eventualidad inesperada, como ajustes en la infraestructura, gastos adicionales para la suscripción de la aplicación en las tiendas App Store y Play Store, o ajustes imprevistos en el desarrollo o pruebas del sistema.

RESUMEN DE COSTOS MENSUALES	
Concepto	**Costo Mensual (COP)**
Tokens (Usados al mes por 650 usuarios)- Chat GPT 3,5	$ 5.590.000
Servidor **Hostgator**	$ 207.050
3 empleados ($2.300.000 por empleado)	$ 6.900.000
Gastos imprevistos 5% (App Store/Play Store)	$ 634.853
Total Mensual	**$ 13.331.903**

Tabla 3.Resumen de costos mensuales por uso de Chat gtt 3.5.

10.5. RENTABILIDAD

La rentabilidad de este proyecto se proyecta con un total de 650 usuarios, distribuidos entre dos canales principales de adquisición de la aplicación: la página web y las tiendas de aplicaciones (App Store o Play Store). Se estima que 325 usuarios realizarán la descarga a través de la página web, mientras que los otros 325 optarán por las tiendas de aplicaciones.

Para los usuarios que descargan la aplicación desde la página web, el valor de la descarga es de $24,000 COP por usuario. En cambio, aquellos que eligen las tiendas

de aplicaciones generan un valor de descarga de $26,000 COP, de los cuales se deduce una comisión de $3,750 COP por cada descarga, resultando en un ingreso neto de $22,250 COP por usuario que realiza la descarga desde estas tiendas.

Este modelo de ingresos es fundamental para garantizar la rentabilidad del proyecto, ya que proyecta los ingresos generados por las 650 descargas, distribuidas entre la web y las tiendas de aplicaciones. Con esta información, se asegura que los ingresos cubrirán los costos y permitirán generar ganancias sostenibles a largo plazo, lo cual es crucial para la toma de decisiones estratégicas, operativas y financieras, y para asegurar la continuidad y el crecimiento del proyecto.

DESCARGAS MENSUALES POR PAGINA WEB			
Usuarios proyectados	**Valor mensual**		
325	$ 24.000		
Total mensual descargas web	**$ 7.800.000**		
DESCARGAS MENSUALES APP / PLAY STORE			
Usuarios proyectados	**Valor mensual**	**Comisión app**	**Valor recibido**
325	$ 26.000	$ 3.750	$ 22.250
Total, mensual descargas app	**$ 7.231.250**	**Comisión a tiendas**	$ 568.750

Ingreso mensual	**$ 15.031.250**

Rentabilidad mensual	**$ 1.699.348**
Rentabilidad Anual	**$ 20.392.170**

Tabla 4.Rentabilidad mensual y anual del proyecto.

11. CONCLUSIONES

La propuesta de valor con la herramienta PROTEGER es la solución de seguridad más avanzada para proteger a menores de edad en el mundo digital. Gracias a este desarrollo, se detecta y previene nuevos tipos de riesgos en tiempo real, brindando la tranquilidad que necesita los padres. Con PROTEGER, se puede estar seguro de que los menores de edad navegan por internet de forma segura y protegida.

- El prototipo desarrollado demuestra el potencial de la inteligencia artificial sumado con una herramienta de supervisión para detectar de manera proactiva el ciberacoso en línea, que aprende a entender mejor el lenguaje natural y detecta nuevos tipos de riegos, de esta manera es una herramienta invaluable para proteger a los menores y empoderar a los padres.
- Los resultados obtenidos en las pruebas del prototipo indican que el modelo de lenguaje natural es capaz de identificar con alta precisión y sensibilidad los patrones de lenguaje asociados al ciberacoso, lo que demuestra la viabilidad de esta solución tecnológica.
- Esta investigación sienta las bases para futuras investigaciones en el campo de la detección de ciberacoso, abriendo nuevas posibilidades para el desarrollo de herramientas más sofisticadas y personalizadas como detección de imágenes, nuevos lenguajes y detectar una falsificación de usuario. Es factible realizar estudios a largo plazo para evaluar el impacto de la herramienta en la reducción de casos de ciberacoso y en el bienestar de los menores.
- En las oportunidades del proyecto se debe profundizar en normativa legal para el tratamiento de datos y privacidad de los menores, ya que esta base legal brinda una mayor confianza a los padres y menores de edad. Es importante contar con cifrado extremo a extremo y encriptación. Aunque los resultados son prometedores, es necesario seguir investigando y mejorando la precisión del modelo de lenguaje natural, especialmente en el caso de lenguajes y dialectos menos representado. Por otro lado, es viable la colaboración con instituciones educativas y organizaciones de protección infantil para promover el uso de esta herramienta y concientizar sobre los riesgos del ciberacoso en menores de edad.

BIBLIOGRAFÍA

Ate Gobierno de Canarias. (s.f.). Obtenido de https://www3.gobiernodecanarias.org/medusa/ecoescuela/seguridad/riesgos-asociados-al-uso-de-las-tecnologias/riesgos/

Análisis de la encuesta a familias del distrito: Conformación familiar de los hogares bogotanos. (2018). Obtenido de Secretaria de planeacion: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiyifG4ruaIAxXBRzABHU0eJ-QQFnoECBkQAQ&url=https%3A%2F%2Fwww.sdp.gov.co%2Fsites%2Fdefault%2Ffiles%2Fboletin_2_observatorio_poblacional_diferencial_y_de_familias_0.docx&usg=AOvVaw0NhuN

Bark. (2023). Bark: The AI-powered monitoring app. Recuperado de https://www.bark.us/

Bandura, A. (1977). Social learning theory. Englewood Cliffs, NJ: Prentice-Hall.

Barba, P., & Del Rey, R. (2010). Ciberacoso entre adolescentes: revisión teórica y propuestas de intervención. Revista de Psicología de la Educación, 24(1), 105-122.

Boletín Estadístico Dirección De Protección. (2022). Obtenido de icbf: https://www.icbf.gov.co/sites/default/files/619590_boletinestadisticooctubre2022_63869386.pdf

Colombia con alarmante aumento en cifras de Bullying. (2022). Obtenido de senado: https://www.senado.gov.co/index.php/el-senado/noticias/13-senadores/4205-colombia-con-alarmante-aumento-en-cifras-de-bullying

Cuántos habitantes tenía Bogotá, Bogotá en 2023. (2023). Obtenido de telencuestas: https://telencuestas.com/censos-de-poblacion/colombia/2023/bogota/bogota

Ciberacoso, Centro de reconocimiento a la dignidad humana. (11 de Nov de 2021). Conecta. Obtenido de https://conecta.tec.mx/es/noticias/ciudad-de-mexico/educacion/cinco-tipos-de-ciberacoso-y-como-afectan-la-salud-mental-y-emocional

Colmenares-Guillen, L. (2024). Online Grooming: de los juegos en línea a la obtención de material de abuso sexual. https://publicacionescd.uleam.edu.ec/index.php/sapientiae/article/view/515/920

Diagnóstico Y Principales Políticas, Programas Y Planes Para La Lucha Contra La Explotación Sexual Comercial De Niñas, Niños Y Adolescentes. (2023). Obtenido de ICBF: https://www.camara.gov.co/sites/default/files/2024-08/Inf_Ley_1336%20de%202009_ESCNNA%202023_FINAL.pdf

Departamento Nacional de Planeación. (s.f.). gov.co. Obtenido de https://www.dnp.gov.co/LaEntidad_/subdireccion-general-prospectiva-desarrollo-nacional/direccion-desarrollo-digital/Paginas/normatividad-conectividad-digital.aspx

Dupuy, D. (2020, December 10). Nuevas herramientas en investigaciones criminales - Dra. Daniela Dupuy - YouTube. https://www.youtube.com/watch?v=DS9r8nf1Hhg

Durkin, F. (1997). Misuse of the Internet by Pedophiles: Implications for Law Enforcement and Probation Practice. https://www.ojp.gov/ncjrs/virtual-library/abstracts/misuse-internet-pedophiles-implications-law-enforcement-and

El Hostigamiento Escolar En Las Instituciones Educativas De Bogotá. (2023). Obtenido de educacionbogota: https://www.educacionbogota.edu.co/portal_institucional/sites/default/files/2023-07/Boletin%20Hostigamiento%20Escolar.pdf

Encuesta Multipropósito. (2021). Obtenido de DANE: https://www.dane.gov.co/files/investigaciones/multi/Boletin_EM_2021.pdf

Graciano Santelises, C. (2021). ESTUDIO ADOLECENTES Y USO DEL INTERNET. https://www.unicef.org/dominicanrepublic/media/5771/file/Adolescentes%20y%20el%20uso%20de%20Internet%20-%20PUBLICACIÓN.pdf

Hinduja, S., & Patchin, J. W. (2010). Bullying, cyberbullying, and sexting: Research, prevention, and intervention. New York, NY: Routledge.

Jurafsky, D., & Martin, J. H. (2021). *Speech and Language Processing* (3ra ed.). Pearson

Juventud en Colombia. (2021). Obtenido de dane: https://www.dane.gov.co/files/investigaciones/notas-estadisticas/dic-2021-nota-estadistica-juventud-en-colombia.pdf

Hinduja, S., & Patchin, J. W. (2010). Bullying and cyberbullying: A comprehensive review of the literature. Journal of School Violence, 9(1), 1-28.

Kowalski, K., Limber, S. P., & Smith, S. (2012). Cyberbullying: Bullying in the digital age. New York, NY: Routledge.

Kowalski, R. M., & Limber, S. P. (2013). Bullying and cyberbullying: Prevention and intervention. Wiley-Blackwell.

Kaspersky. (2024). Obtenido de https://latam.kaspersky.com/resource-center/preemptive-safety/parents-and-social-media?srsltid=AfmBOopLSkE4xZxXpc55xM8hTWvxaGI6XAuwPlw91RTBOmjfaMqCPTwW

Livingston, S. (2009). When and how parents should talk to youth about internet safety. Journal of Adolescent Health, 44(2), 182-188.

Olweus, D. (1993). Bullying at school: What we know and what we can do. Cambridge, MA: Blackwell.

Olweus, D. (2001). Bullying at school: What we know and what we can do. Blackwell Publishers.

Paquette, S., Frotin Francis, & Perkins, D. (2020). Online Sexual Offenders: Typologies, Assessment, Treatment, and Prevention.

Rico Muñoz, A. (30 de Ago de 2022). LR MAS. Obtenido de https://mas.larepublica.co/noticias

Rigby, K. (2016). Bullying in schools and the workplace: Developments in theory, research, and practice. Routledge.

Sánchez Gálvez, F. (2023). Revista cientifica Dialogo forense. 4, 20–26. https://dialogoforense.inacif.gob.gt/index.php/dialogoforense/issue/view/11/16

Smith, J., Jones, A., & Lee, K. (2019). The impact of parental monitoring software on cyberbullying. Journal of Cyberpsychology, Behavior, and Social Networking, 12(3), 123-130.

Smith, P. K., Bradshaw, C. P., & Mitchell, K. J. (2008). What works in bullying prevention? A systematic review of the evidence. Journal of Adolescence, 31(1), 31-63.

Smith, P. K., Mahdavi, J., Carvalho, M., & Fonseca, V. (2012). Cyberbullying: Its nature and impact in school settings. Journal of School Violence, 11(1), 1-37.

Superintendencia de Industria y Comercio. (s.f.). Gov.co. Obtenido de https://www.sic.gov.co/content/%C2%BFlos-datos-personales-de-los-ni%C3%B1os-ni%C3%B1as-y-adolescentes-tienen-alguna-protecci%C3%B3n-especial

TIC. (06 de Abr de 2020). GOV.CO. Obtenido de https://mintic.gov.co/portal/inicio/Sala-de-prensa/Noticias/126491:Como-proteger-a-los-menores-de-edad-de-los-retos-virales-que-circulan-en-Internet

Trujillo, S. (2019, September 19). Caldo de pollo: La perturbadora palabra clave que usan los pedófilos para buscar pornografía infantil en internet. https://www.fayerwayer.com/2019/09/pornografia-infantil-caldo-de-pollo/

Unicef. (Feb de 2024). Obtenido de https://www.unicef.org/es/end-violence/ciberacoso-que-es-y-como-detenerlo

Villegas, E. (24 de Ene de 2024). Hola.com. Obtenido de https://www.hola.com/padres/20240124355108/apps-y-redes-sociales-favoritas-menores-edad/

Violencias Basadas En Género Y Violencia Sexual En Las Instituciones Educativas De Bogotá. (2023). Obtenido de educacionbogota: https://www.educacionbogota.edu.co/portal_institucional/sites/default/files/2024-01/BOLETÍN%20VIOLENCIAS%20BASADAS%20EN%20GÉNERO%20Y%20VIOLENCIA%20SEXUAL%20EN%20LAS%20IE%20DE%20BOGOTÁ.pdf

Wavemarker. (01 de Oct de 2021). Obtenido de https://wavemakerglobal.com/es/tiempo-de-uso-de-internet-de-la-generacion-z-y-redes-sociales-mas-utilizadas

Ybarra, M., & Mitchell, K. J. (2004). Online harassment as a function of offline deviance. Journal of Research on Adolescence, 14(3), 367-386.