**Red Social de Denuncias – Reporta Pe**

Especificación de Requisito:

ER010 – Moderación de Contenido

Versión 1.0

**Lima, abril del 2024**

|  |  |  |  |
| --- | --- | --- | --- |
| **Fecha** | **Versión** | **Descripción** | **Autor** |
| 29/04/2024 | 1.0 | Creación del documento. | Bryan Velarde |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**Control de Versiones**

**ÍNDICE**

[1. Introducción 4](#_Toc132963159)

[1.1 Propósito 4](#_Toc132963160)

[1.2 Alcance 4](#_Toc132963161)

[1.3 Definiciones, siglas y abreviaciones 5](#_Toc132963162)

[1.4 Referencias ¡Error! Marcador no definido.](#_Toc132963163)

[1.5 Resumen 6](#_Toc132963164)

# Introducción

## Propósito

El requisito de Moderación de Contenido tiene como objetivo principal garantizar la integridad, seguridad y adecuación del contenido generado por los usuarios dentro de la aplicación de gestión de reportes de delitos. La moderación del contenido se realiza mediante la revisión y evaluación por parte de los administradores, quienes tienen la responsabilidad de asegurar que los reportes publicados cumplan con los estándares establecidos y no violen las políticas de uso de la plataforma.

La función de moderación abarca la detección y acción sobre reportes inapropiados, falsos o que incumplen las normativas de la comunidad. Esto incluye la capacidad de marcar reportes como inadecuados, ocultarlos del público o eliminarlos de la plataforma, según corresponda. Además, los administradores pueden tomar medidas adicionales, como contactar al usuario responsable del reporte para solicitar aclaraciones o evidencias adicionales.

Al garantizar la calidad y veracidad del contenido publicado, la moderación de contenido contribuye a mantener la credibilidad y confianza en la aplicación, promoviendo un entorno seguro y confiable para todos los usuarios. Además, ayuda a prevenir la propagación de información falsa o engañosa, lo que es crucial en el contexto de la gestión de reportes de delitos, donde la veracidad de la información es fundamental para la toma de decisiones y acciones adecuadas por parte de las autoridades pertinentes y la comunidad en general..

## Alcance

El alcance del requisito de Moderación de Contenido se centra en las siguientes áreas y actividades dentro de la aplicación de gestión de reportes de delitos:

Revisión de Reportes: Los administradores tienen la responsabilidad de revisar periódicamente los reportes publicados por los usuarios para verificar su exactitud, veracidad y adecuación según las políticas de la plataforma.

Detección de Contenido Inapropiado: Se debe identificar cualquier contenido que pueda ser inapropiado, engañoso, difamatorio, ofensivo o que viole los términos de uso de la aplicación.

Acciones de Moderación: Los administradores están facultados para tomar acciones apropiadas en respuesta a reportes inadecuados, como marcarlos como inapropiados, ocultarlos del público o eliminarlos de la plataforma, según corresponda.

Comunicación con Usuarios: Cuando sea necesario, los administradores pueden comunicarse con los usuarios responsables de los reportes para solicitar aclaraciones, proporcionar orientación sobre las políticas de la plataforma o solicitar evidencia adicional para respaldar el reporte.

Mantenimiento de Registros: Se debe mantener un registro completo de todas las acciones de moderación realizadas, incluyendo la razón de la acción tomada, la fecha y la identidad del administrador responsable.

El alcance también incluye la implementación de herramientas y funcionalidades que faciliten la tarea de moderación, como filtros de palabras clave, sistemas de alerta temprana para contenido potencialmente inapropiado y la capacidad de realizar búsquedas avanzadas para identificar reportes específicos que requieran atención.

## Definiciones, siglas y abreviaciones

* Administrador: Usuario con privilegios especiales dentro de la plataforma de gestión de reportes de delitos, responsable de funciones como la moderación de contenido, gestión de usuarios y análisis de datos.
* Contenido inapropiado: Cualquier material, información o reporte que viole las políticas de la plataforma, incluyendo, pero no limitado a contenido difamatorio, engañoso, ofensivo, ilegal o que vaya en contra de los términos de uso.
* Acción de moderación: Cualquier medida tomada por un administrador en respuesta a contenido inapropiado, como marcarlo como inapropiado, ocultarlo del público o eliminarlo de la plataforma.
* Registro de moderación: Registro detallado que documenta todas las acciones de moderación realizadas por los administradores, incluyendo la razón de la acción, la fecha y la identidad del administrador responsable.
* Filtro de palabras clave: Herramienta utilizada para identificar automáticamente palabras o frases específicas en el contenido de los reportes, que podrían indicar contenido inapropiado o violaciones de las políticas de la plataforma.
* Sistema de alerta temprana: Funcionalidad que notifica a los administradores sobre contenido potencialmente inapropiado tan pronto como sea detectado por la plataforma, permitiendo una respuesta rápida y eficiente.
* Búsqueda avanzada: Funcionalidad que permite a los administradores realizar búsquedas detalladas y específicas dentro de la plataforma, utilizando una variedad de criterios para identificar reportes que requieran atención.
* Políticas de la plataforma: Conjunto de reglas, directrices y términos de uso establecidos por el operador de la aplicación, que rigen el comportamiento de los usuarios y el contenido permitido en la plataforma.

## Resumen

El requisito de Moderación de Contenido establece las funciones y responsabilidades de los administradores dentro de la plataforma de gestión de reportes de delitos. Su propósito es garantizar que el contenido publicado por los usuarios cumpla con las políticas establecidas y no contenga material inapropiado o engañoso. Dentro del alcance de este requisito se incluyen actividades como la revisión y moderación de reportes, la aplicación de acciones correctivas sobre contenido inapropiado y el registro detallado de todas las acciones de moderación realizadas. Este proceso se apoya en herramientas como filtros de palabras clave, sistemas de alerta temprana y funciones de búsqueda avanzada para identificar y abordar eficazmente el contenido problemático. El objetivo es mantener un entorno seguro y confiable para todos los usuarios, promoviendo la integridad y la legitimidad de los reportes de delitos dentro de la plataforma.

En resumen, la funcionalidad de Auditoría de Actividades garantiza la trazabilidad y el control de todas las acciones realizadas en la aplicación, contribuyendo así a la seguridad, transparencia y calidad del sistema de gestión de reportes de delitos.