Tecnológico de Monterrey

Campus Querétaro

Privacidad y Seguridad de los Datos

Gamaliel Marines Olvera	A01708746
Uri Jared Gopar Morales	A01709413
José Antonio Miranda Baños	A01611795
María Fernanda Moreno Gómez	A01708653
Oskar Adolfo Villa López	A01275287
Luis Ángel Cruz García	A01736345

Inteligencia artificial avanzada para la ciencia de datos II
Grupo 501

Introducción

El presente documento plantea definiciones, prácticas, procesos e investigaciones acerca de las prácticas de seguridad y privacidad de datos implementadas en nuestro proyecto, así como prácticas o normativas que se han planteado los países para manejar proyectos como este.

Datos anonimizados

Para garantizar que nuestros datos estuvieran libres de información personal o sensible para nuestro socio formador, en nuestro proceso de clasificación de las imágenes (vaca acostada, vaca parada o no hay vaca), nos permitió tener una visión completa del conjunto de imágenes y asegurarnos de que no contenían ningún dato que pudiera comprometer la privacidad o seguridad de nuestro socio u organización asociada. Cuando se identificaba alguna imagen que podría poner en riesgo la anonimidad de la persona o empresa, esta era eliminada.

Normativas

Como parte de la descripción de las normativas y el tratamiento de los datos sensibles en nuestro proyecto, realizamos una Política de Datos y Regulaciones de Privacidad, todo esto basados en prácticas que han hecho otros proyectos o empresas con un giro similar a nuestro enfoque.

En dicho documento, se presenta el alcance de nuestra política de protección de datos, términos, tratamiento de los datos, definición de datos sensibles, transparencia, acceso basado en roles, almacenamiento, seguridad y normativas (tanto en México como extranjeras). Las normativas fueron recopiladas a través de investigaciones individuales de los miembros del equipo y se escogieron las que se tomaron relevantes y con un enfoque similar al de nuestro proyecto.

Proceso para trabajar con el dataset

Propósito

Establecer cómo asegurar la seguridad y protección de los datos al manipular el dataset dado por el socio formador.

Notas introductorias

Entendemos la importancia de la protección de los datos de nuestro socio formador. Asegurarnos que los datos sean correctamente manipulados nos permite garantizar al socio formador que sus datos no serán vistos o utilizados para otros fines que no sean los especificados y que no sean manipulados por terceros.

Involucrados

- Equipo de desarrollo (Equipo “TC”)
- Equipo compañero “No Name”

Entradas

- Dataset de imágenes proporcionado por el socio formador y previamente clasificado por ambos equipos de camas. Link
- Bitácora de Buenas Prácticas (para registro de actividades y monitoreo).
- Política de Datos y Regulaciones de Privacidad.

Salidas

- Bitácora de Buenas Prácticas actualizada

Descripción

Fase	Actividades	Responsable(s)
Preparación Inicial del Dataset	Verificar la clasificación del dataset (Sensibles / No sensibles) según las categorías definidas en la política.	Equipo TC Equipo No Name
Preparación Inicial del Dataset	Documentar la fecha y hora del primer acceso al dataset en la sección “Almacenamiento y Accesos” en la Bitácora de Buenas Prácticas	Equipo TC
Monitoreo de Accesos	Cuando se cree algún lugar para almacenar los datos, se debe mapear en “Almacenamiento y Accesos” y se debe especificar los datos almacenados, el lugar de almacenamiento, las personas que tienen acceso a esos datos, la imagen de los permisos y si hay actualizaciones de permisos.	Equipo TC
Registro de Acceso	Registrar cada acceso al dataset indicando: nombre de la(s) personas, actividad realizada (acceso, modificación, eliminación, descarga), datos utilizados, imagen de la actividad, fecha, hora, y comentarios de dicho acceso	Equipo TC
Manipulación del Dataset	Asegurar que cualquier cambio (como etiquetado o eliminación) esté documentado en la Bitácora de Buenas Prácticas en su parte correspondiente.	Equipo TC

Gestión de Incidentes	En caso de un incidente de seguridad, documentar en la "Monitoreo de Actividad" el integrante que va a solucionarlo, en el apartado de actividad la solución a implementar, los datos utilizados, la imagen de ejemplo, la fecha de solución y en comentarios, poner el impacto ocasionado por el incidente y comentarios adicionales.	Equipo TC

Debemos recordar que…

- **Almacenamiento**. El almacenamiento se realiza en Google Drive después de recibir las imágenes. La carpeta de Google Drive es compartida, con acceso solo para los miembros del equipo. Esto se decidió por la seguridad de la infraestructura de Google.
- **Redes.** Se puede acceder a la carpeta desde cualquier red. No manejamos datos sensibles, por lo que no es indispensable controlar las redes desde las que se accede a la carpeta.
- **Acceso.** Solo los miembros del equipo, profesores y miembros de otros equipos pueden ver las imágenes. La carpeta del equipo no se puede compartir con ninguna persona ajena al equipo.
- **Documentos.** Debido a la naturaleza de los datos y el acuerdo realizado con el socio formador, no es necesario firmar un NDA nosotros con el socio formador, sin embargo, se deberá firmar la Política de Datos y Regulaciones de Privacidad por ambas partes, es decir, nosotros como equipo desarrollador (TC) y nuestro socio, el Mtro. Ivo Ayala.

Bitácora de buenas prácticas

La Bitácora de Buenas Prácticas es una herramienta fundamental implementada para garantizar la trazabilidad y transparencia en la manipulación de datos durante todas las fases del proyecto. Su objetivo principal es registrar y monitorear todas las actividades relacionadas con el uso del dataset proporcionado por el socio formador, asegurando el cumplimiento de las normativas de seguridad y privacidad.

Propósitos principales:

1. **Trazabilidad:** Garantizar un registro claro y ordenado de quién, cuándo y cómo se interactúa con el dataset. Esto incluye accesos, modificaciones, eliminaciones, y cualquier actividad relevante.
2. **Transparencia:** Brindar al socio formador visibilidad total sobre cómo se manejan sus datos, fomentando la confianza y la seguridad.

3. **Cumplimiento de normas:** Asegurar que todas las acciones estén alineadas con la Política de Datos y Regulaciones de Privacidad establecida para el proyecto.

Componentes principales de la bitácora:

1. **Registro de Actividades**
    a. Cada interacción con el dataset debe incluir
        i. Fecha de la actividad.
        ii. Descripción de la actividad (acceso, etiquetado, modificación, eliminación, etc.).
        iii. Nombre del responsable.
        iv. Evidencias relacionadas (por ejemplo, enlaces a repositorios o carpetas).
        v. Comentarios.
2. **Almacenamiento y Acceso:**
    a. Permite identificar dónde son almacenados los datos creados en el proyecto y así tener a la mano su ubicación.
    b. Permite identificar qué usuarios pueden tener acceso a los grupos de datos.

Beneficios de la implementación:

- Mejora la seguridad de los datos al mantener un control detallado de su uso.
- Promueve la rendición de cuentas dentro del equipo de trabajo.
- Facilita la identificación y resolución rápida de posibles brechas de seguridad.
- Genera confianza con el socio formador al evidenciar un manejo profesional y responsable de los datos.

La bitácora no solo es un documento de registro, sino también una guía viva que asegura el correcto uso del dataset durante todo el proyecto. Su implementación y seguimiento contribuyen al éxito del proyecto y a la construcción de relaciones sólidas con el socio formador.