Tecnológico de Monterrey

Instituto Tecnológico y de Estudios Superiores de Monterrey

Momento de Retroalimentación: Reto. Privacidad y Seguridad de los Datos

TC3007C.501 Inteligencia Artificial Avanzada para la Ciencia de Datos II

Profesores:

Iván Mauricio Amaya Contreras

Blanca Rosa Ruiz Hernández

Félix Ricardo Botello Urrutia

Edgar Covantes Osuna

Felipe Castillo Rendón

Hugo Terashima Marín

Equipo 2

Integrantes:

Luis Ángel Guzmán Iribe - A01741757

Julian Lawrence Gil Soares - A00832272

Alberto H Orozco Ramos - A00831719

27 de Octubre de 2023

Antes de subir los datos a cualquier repositorio, discutan a profundidad con el socio formador sobre la naturaleza de los datos con los que van a trabajar. Es importante entender las implicaciones legales y de seguridad que vienen asociadas a los datos con los que estarán trabajando.

1. *Verifica que los datos que generes estén anonimizados, es decir que no se pueda rastrear información personal o sensible a una persona o producto específico a través del dataset. Si los datos ya están anonimizados, describe cuáles fueron los atributos y las razones por las que se tienen que enmascarar.*

   Los datos que se encuentran en el bucket de S3 en AWS para nuestro proyecto se encuentran protegidos por algunas funcionalidades que el mismo AWS nos provee para proteger la información utilizada por el modelo de reconocimiento y participación. Por ejemplo, se tiene que los buckets por defecto son privados, siendo así la única persona que puede acceder a este es el mismo dueño y nadie más tiene acceso ni al bucket ni a los datos almacenados, además de que se ha considerado la privacidad de los usuarios por medio de la encriptación de cada una de las imágenes utilizando la misma encriptación que ofrece S3 como SSE-S3, SSE-KMS o SSE-C, con el objetivo de prevenir algún tipo de acceso no autorizado durante la transmisión de los datos o el robo de dicho dataset desde el mismo bucket, ya que si llegase a presentarse alguno de estos casos, los perpetradores no serán capaces de obtener ninguna información personal o sensible de los archivos encriptados.

   Es importante considerar estos atributos por el hecho de que existen riesgos en todo momento, y más cuando se manejan datos de este tipo, por ello se aprovechan los recursos que posee AWS para facilitar y asegurar la privacidad y seguridad de los usuarios involucrados con el software a desarrollar. De llegarse a exponer datos de los usuarios por medio de las imágenes que utiliza el software, esto podría provocar graves consecuencias, desde el incumplimiento de las principales normativas de seguridad para tratamiento de datos hasta posibles falsificaciones, robos de identidad, extorsiones o manipulación de diversos tipos. Teniendo en cuenta lo anterior, es debido a esto que es indispensable mantener los datos privados, encriptados y seguros, evitar en toda medida la exposición de los mismos a cualquier tipo de ataque, fallo en la seguridad o del sistema, que pueda perjudicar tanto a la organización, como a los desarrolladores, los mismos usuarios y terceros.

2. *Consulta la normativa actual de la industria a la que esté sujeto el socio formador e investiga en reportes técnicos, artículos o foros cuales son los pasos comunes que se toman para garantizar la privacidad de los datos en dicha industria.*

   Según el documento de la Ley Federal de Protección de Datos Personales en Posesión de los Particulares (LFPDPPP) proporcionado por el INAI, y en el que se rige su cumplimiento el socio formador por parte de NDS Cognitive Labs, para asegurar la privacidad y seguridad de los datos primeramente se establecen conceptos básicos como datos personales, datos sensibles, su tratamiento, el/los encargado(s) principal(es), transferencia y remisión de los mismos con fines de entendimiento de la guía. En secciones posteriores del documento se describen las ocho principios y obligaciones que los responsables de los datos deben cumplir, como la licitud, lealtad, consentimiento, información, proporcionalidad, finalidad, calidad y responsabilidad. También, se cubre la relación que debe existir entre el responsable de los datos y el encargado, teniendo en cuenta las obligaciones que existen de por medio a cumplir, como lo es establecer y firmar un contrato y posteriormente supervisar que el tratado de los datos se cumpla según lo estipulado dentro del documento.

   Otro punto importante a mencionar es que en el mismo documento se discute la transferencia de datos personales y las obligaciones que el responsable debe cumplir, así como la obtención del consentimiento e informar del propósito con el que se hará uso de dicha información, además de dar a conocer qué o quiénes son los receptores de dichos datos.

   Finalmente, se aborda el tema de las vulnerabilidades, ataques, negligencia e incumplimiento de las obligaciones y normas establecidas por el documento, por ejemplo, podemos encontrar consecuencias como la publicidad negativa debido a la falta de compromiso con el cumplimiento, pérdida de confianza (pérdida de clientes/inversores/accionistas o del mismo negocio), aplicación de sanciones con o sin multas económicas o hasta prisión, claro que depende de la gravedad de la falta y las leyes que se hayan quebrantado.

3. *Establece un proceso claro sobre cómo se puede trabajar con el set de datos y especifica aspectos como: dónde se puede almacenar, en qué tipo de redes puede estar, quien los puede ver y cuáles son los documentos o normas que se deben de firmar antes de poder acceder a los datos.*

1. **Almacenamiento de Datos:** Los datos, que incluyen imágenes de los rostros de estudiantes, se almacenan en un bucket de S3 en AWS. Este bucket se configura como privado por defecto, y la encriptación (SSE S3, SSE-KMS, SSE-C) se aplica para garantizar la seguridad durante la transmisión y almacenamiento.
2. **Acceso y Transmisión de Datos:** El acceso a los datos en el bucket está restringido al propietario (AWS Lambda). Las imágenes se transmiten desde el frontend (VueJS) a través de FastAPI para su procesamiento en AWS Lambda, manteniendo así la seguridad y privacidad durante la transmisión.
3. **Normas y Documentación:** Antes de acceder a los datos, se establece que los usuarios deben seguir un proceso que incluye la creación de clases, registro de estudiantes y carga de imágenes. Puede ser necesario documentar normas específicas y procesos que los usuarios deben seguir al interactuar con los datos.

4. *Implementen un mecanismo o utiliza una herramienta que les permita establecer registros sobre quién y cuándo tuvo acceso a los datos y bajo qué esquema. Estos registros los deberán integrar a su reporte como parte de la evidencia de final de módulo.*

- **Proceso de Acceso:**

  Antes de acceder a los datos, el personal autorizado debe seguir un proceso formal. Este proceso incluye:

  - **Interfaz de Usuario (VueJS):** Los usuarios autorizados acceden a una interfaz de usuario desarrollada en VueJS para realizar operaciones como la creación de clases, registro de estudiantes y carga de imágenes.
  - **Comunicación con Backend (FastAPI):** La interfaz de usuario se comunica con el backend (hosted on an AWS EC2 instance) a través de solicitudes HTTP gestionadas por FastAPI.
  - **Procesamiento en Backend:** El backend, utilizando FastAPI, procesa las solicitudes, valida autorizaciones y realiza la comunicación con los modelos de reconocimiento facial y participación basados en Deep Learning.

- **Autorizaciones y Documentación:**

    - El acceso a los datos está restringido a roles específicos y se requiere autorización para:
    - Acceder a la interfaz de usuario en VueJS.
    - Enviar solicitudes al backend (FastAPI) para operaciones como creación de clases y registro de estudiantes.
    - Interactuar con los modelos de reconocimiento facial y participación.
    - Acceder a la infraestructura en AWS, incluyendo la EC2 instance.
- **Normas Específicas:**
    - Documentar y comunicar normas específicas para el manejo de datos a través de la interfaz de usuario y las operaciones permitidas. Esto incluye pautas para el envío seguro de datos entre el frontend y el backend.

Bibliografía:

- INAI – Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales. (s/f). Org.mx. Recuperado el 25 de octubre de 2023, de https://home.inai.org.mx/
- (N.d.). Amazon.com. Retrieved November 22, 2023, from https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/ec2-security.html
- Notice of privacy. (s/f). Ndscognitivelabs.com. Recuperado el 25 de octubre de 2023, de https://ndscognitivelabs.com/notice-of-privacy/