



**Tecnológico
de Monterrey**



**Instituto Tecnológico y de Estudios
Superiores de Monterrey**
Campus Monterrey

Inteligencia artificial avanzada para la ciencia de datos II
TC3007C.501

Reto Privacidad y Seguridad de los Datos



Rodolfo Sandoval Schipper A01720253

Arturo Garza Campuzano A00828096

Marcelo Márquez A01720588

25 oct 2023

1. Introducción

En este documento se discute acerca de la naturaleza de los datos con los que se trabajarán en el proyecto **Classroom AI**, con el fin de comprender las implicaciones legales y de seguridad que vienen asociados a los mismos. Para cumplir con el fin establecido se ofrece un enfoque sobre los siguientes temas: anonimización de datos; normativa de la industria; acceso y uso responsable de datos; y trazabilidad y auditabilidad del acceso a datos.

2. Anonimización de datos

Como se ha establecido en [Memorandum Understanding](#), **Classroom AI** es un sistema para medir la asistencia y participación en el entorno universitario cuyos objetivos son: perder el menor tiempo posible pasando lista entre los alumnos, tomar acción cuando se detecta una baja participación entre todos los alumnos, y la participación y asistencia detectada es visible a los involucrados en el curso. Tomando en cuenta estos objetivos, se cree que es del interés de los usuarios visualizar información personal para obtener una mejor experiencia de usuario y darle un uso efectivo al sistema.

Sin embargo, en este caso, se implementa la anonimización sobre un nuevo rol dentro de la interfaz desarrollada para realizar una prueba de concepto. Dicho rol, llamado *auxiliar*, permite agregar y visualizar los datos sin obtener información personal.

Para lograr dicha anonimización de datos es necesario considerar, en primer lugar, qué información es personal. A continuación se muestran los datos que se consideran personales dentro de las tablas de la base de datos del sistema.

Estudiantes

- *Nombre ** (Se utiliza para identificar qué alumno está participando/asistiendo)
- *Apellido ** (No es necesario pero puede servir para identificar un alumno en caso de que se repita el nombre)
- *Correo ** (Es parte del usuario “Estudiantes” no es necesario mostrar el correo pero sirve como una llave ya que solo existe un correo por estudiante, verificamos esta autenticación con firebase y en la base de datos)
- *Foto del Estudiante ** (Se utiliza para alimentar el modelo de face recognition)

Usuarios

- *Correo electrónico ** (Se muestra para que el administrador pueda hacer operaciones CRUD para modificar la cuenta del usuario)
- *Contraseña ** (En proceso de realizar “One way password Hashing con el algoritmo SHA 256” este dato no será accesible para ningún usuario)

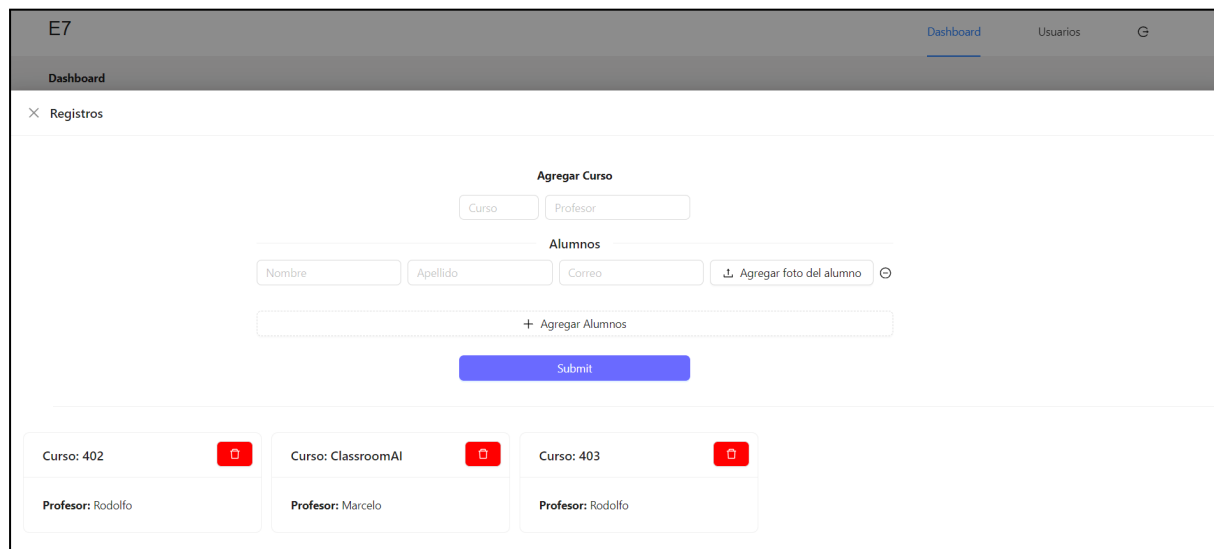
Profesores

- *Nombres de Profesores ** (Se utiliza cómo un atributo dummy para verificar la creación de un curso con de acuerdo al profesor)

Bajo estas consideraciones se ha creado el rol *auxiliar*. Solo una cuenta administradora puede crear este tipo de usuario y cuenta con casi todas las funcionalidades que posee un profesor con la diferencia de que la información personal no se puede visualizar y tampoco se puede iniciar sesión para probar los modelos de inteligencia artificial.

Evidencia de la implementación

Usuario Normal Creación/Visualizar de Cursos →



E7 Dashboard Usuarios

× Registros

Agregar Curso

Curso Profesor

Alumnos

Nombre Apellido Correo Agregar foto del alumno

+ Agregar Alumnos

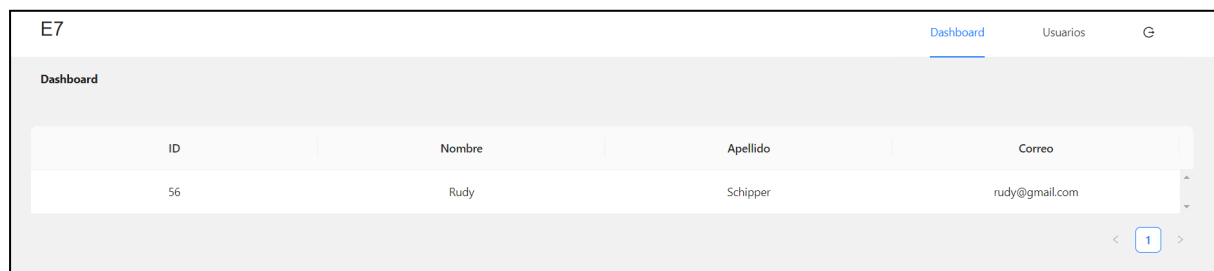
Submit

Curso: 402 Profesor: Rodolfo

Curso: ClassroomAI Profesor: Marcelo

Curso: 403 Profesor: Rodolfo

Usuario Normal Visualizar Estudiantes →

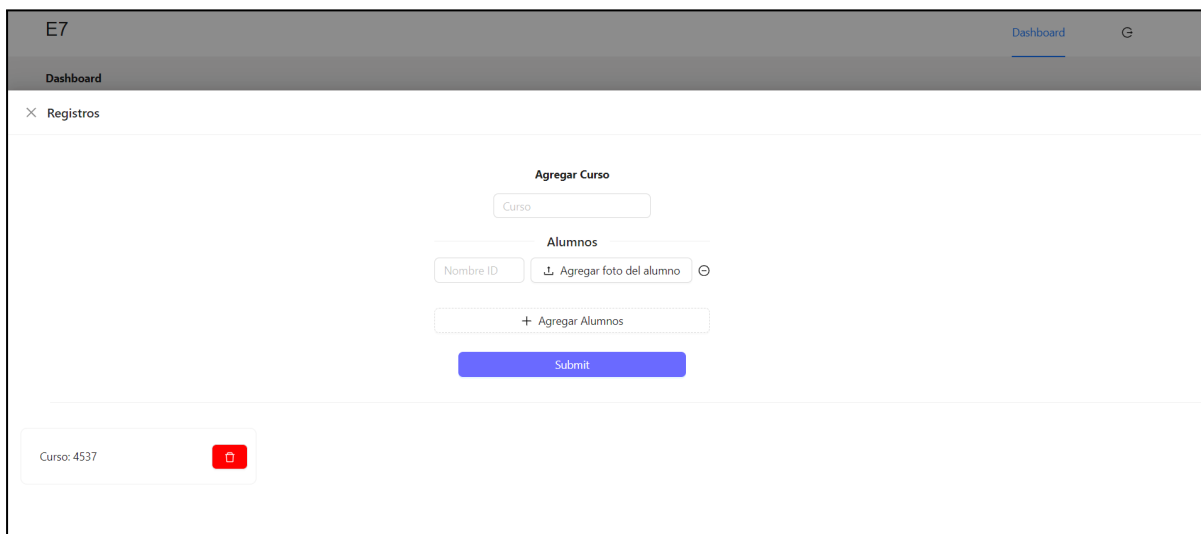


E7 Dashboard Usuarios

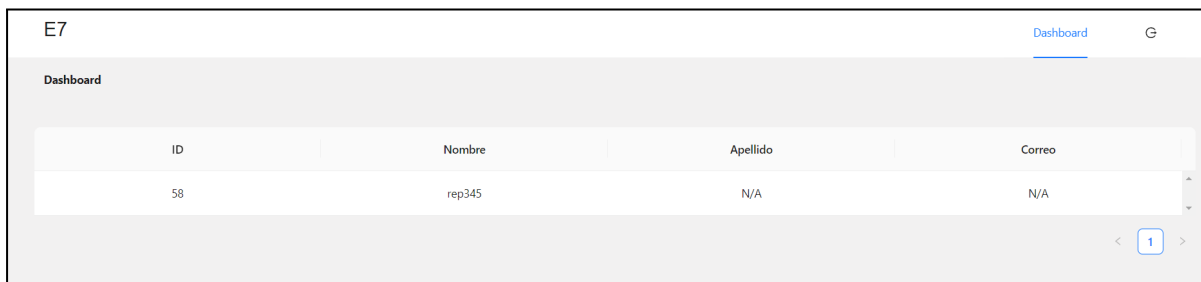
ID	Nombre	Apellido	Correo
56	Rudy	Schipper	rudy@gmail.com

< 1 >

Usuario Auxiliar Creación/Visualizar Cursos →



Usuario Auxiliar Visualizar Estudiantes →



ID	Nombre	Apellido	Correo
58	rep345	N/A	N/A

3. Normativa de la industria

Apartado dedicado a la consulta de la normativa actual de la industria a la que está sujeto el socio formador y a la investigación de los pasos más comunes que se toman para garantizar la privacidad de los datos en dicha industria.

3.1 Normativa de NDS Cognitive Labs

Según NDS Cognitive Labs (n.d.), en su 'Notice of Privacy', la normativa principal a la que está sujeto es la "LFPDPPP" (Ley Federal de Protección de Datos Personales en Posesión de Particulares), ley que establece las obligaciones y derechos relacionados con la protección de datos personales en México. Esta normativa incluye:

1. Responsabilidad en el tratamiento de datos personales
2. Propósito del procesamiento de datos personales
3. Opciones y medios para limitar el uso o divulgación de datos personales
4. Transferencia de datos personales a terceros
5. Consentimiento del cliente para el tratamiento de datos personales
6. Cambios en el aviso de privacidad

7. Derecho de protección de datos ante el INAI

NDS Cognitive Labs debe seguir las regulaciones de protección de datos personales, garantizando la confidencialidad y seguridad de los datos personales de sus clientes y cumpliendo con los derechos de los titulares de datos según lo establecido por la ley.


3.2 Pasos comunes para garantizar la privacidad de los datos

Para asegurar la la privacidad de los datos Rowda Mohamud recomienda seguir los siguientes pasos:

1. **Observar los datos que se están recopilando y almacenando, y el motivo por el cual se realiza:** es crucial considerar cuánta información privada realmente se requiere para atender adecuadamente a los clientes. Cuanto más datos se acumulan, más riesgos y costos de almacenamiento se generan. La gestión adecuada de datos requiere de una estrategia de recopilación y retención que se ajuste a las necesidades, regulaciones y expectativas de privacidad de los clientes pertinentes al proyecto.
2. **Definir quién debe tener acceso a los datos:** para prevenir incidentes es esencial definir roles y niveles de acceso asociados, el principio del acceso mínimo es un enfoque sólido. Hay que considerar: la vinculación de registros de usuarios con procesos de recursos humanos, establecer un procedimiento claro para controlar los permisos de acceso y revisar regularmente los derechos de acceso en aplicaciones críticas.
3. **Comprensión de los riesgos para los datos:** aunque es cierto que la plataformas en la nube cuentan con estrictas características de seguridad, es relevante que las empresas revisen los riesgos de los datos, considerando el valor de los datos, el costo de perderlos, las implicaciones de privacidad al recopilar los datos y los costos de diversas soluciones para abordar los riesgos.
4. **Capacitar a los empleados sobre la privacidad de los datos y su función:** esencial para prevenir errores humanos y malas prácticas que suelen ser causas comunes de brechas cibernéticas. Los empleados necesitan comprender los riesgos de no seguir protocolos de privacidad de datos, y para ello, herramientas como la capacitación anual en ciberseguridad, simulaciones de phishing y una comunicación constante sobre los beneficios para el negocio son fundamentales.
5. **Superar lo mínimo necesario:** cumplir con las leyes locales es esencial para evitar multas y daños a la reputación, pero enfocarse únicamente en el cumplimiento puede perder oportunidades para destacar. Al integrar conscientemente la privacidad de datos en productos, servicios y procesos

comerciales, el enfoque cambia de lo permitido a lo que genera valor para los clientes.

4. Acceso y uso responsable de datos

Por el momento, **Classroom AI** se encuentra en la fase de prototipo, por lo que se trabaja con un conjunto de datos reducido. La prioridad de los **Caballeros de Camelot** es cumplir con los requerimientos establecidos en el  **SRS**, garantizando la funcionalidad deseada por el cliente. Finalizado el prototipo se podría establecer un proceso más claro sobre cómo se puede trabajar con el set de datos futuro y especificar aspectos como: dónde se podrá almacenar, en qué tipo de redes estará disponible, quién lo podrá ver y cuáles son los documentos o normas que se deben firmar antes de poder acceder a los datos. A continuación se realiza una propuesta de dicho proceso considerando que el **responsable** del sistema en producción es NDS Cognitive Labs y su **cliente** es el Tecnológico de Monterrey Campus Monterrey.

El **objetivo** de este proceso es garantizar la seguridad y privacidad de los datos confidenciales de los estudiantes y profesores del Tecnológico de Monterrey Campus Monterrey. Por otro lado, las **responsabilidades** propuestas para este proceso son las siguientes:

- **Equipo de datos confidenciales:** equipo designado por NDS Cognitive Labs para gestionar y proteger los datos confidenciales de los miembros de la institución educativa.
- **Usuarios autorizados:** sólo los empleados de NDS Cognitive Labs que ofrezcan el producto y personal externo autorizado, en este caso los profesores, podrían acceder a los datos confidenciales.

Tomando en cuenta este objetivo y responsabilidades, se propone el siguiente proceso de gestión de datos confidenciales:

- **Almacenamiento de Datos:** los datos confidenciales deben ser almacenados exclusivamente en los servidores designados por NDS Cognitive Labs. No se debe permitir el almacenamiento en dispositivos personales o en servicios de una nube no autorizada.
- **Redes Aprobadas:** los datos confidenciales solo pueden estar en redes internas seguras de NDS Cognitive Labs. El acceso a través de redes públicas o no seguras está estrictamente prohibido.
- **Acceso Autorizado:** sólo los usuarios autorizados tienen permitido acceder a los datos confidenciales. El acceso se gestiona a través de cuentas y contraseñas únicas.
- **Documentación Requerida:** antes de acceder a los datos confidenciales, los usuarios autorizados deben firmar un acuerdo de confidencialidad y completar la formación obligatoria sobre la seguridad de datos.

- **Monitorización y Auditoría:** monitorización continua de las actividades relacionadas con los datos confidenciales y auditorías periódicas para garantizar el cumplimiento de la política de seguridad de datos.

5. Trazabilidad y auditabilidad del acceso a datos

Con el fin de incorporar un mecanismo que permita establecer registros sobre quién y cuándo tuvo acceso a los datos y bajo qué esquema se crea una nueva tabla en la base de datos llamada “Logs”, donde utilizamos el ID del usuario Estudiantes y Usuarios (Profesores) y los atributos LogInTime DATETIME y LogOutTime DATETIME.

Creando esta tabla lo que se quiere es registrar el momento en que un usuario entra o sale del sistema, tal que:

- UserID FK que referencia UsersProfessors (UserID) y Student ID FK que referencia Students (StudentID) identifica el tipo de usuario que inicia la sesión
- LogInTime registra la hora en la que se inicia una sesión
- LogOutTime registra la hora en la que se cerró la sesión

Cada vez que un usuario inicie la sesión en el sistema, se inserta un registro en esta tabla con la hora de inicio de sesión; esto se logra incorporar por medio de una consulta en SQL. También se crea un SELECT donde es posible consultar el tiempo conectado de ese usuario. Para calcular el tiempo de un usuario conectado se puede hacer una consulta que reste la hora de inicio con la hora de cierre para obtener la duración de la sesión; utilizar DATEDIFF en segundos del LogInTime y LogOutTime. Esta información se desplegará en un componente del front-end donde solo cuentas administradoras podrán observar el tiempo en el que los usuarios estuvieron dentro del sistema, así para llevar un registro de la información de utilización de la plataforma.

6. Referencias bibliográficas

1. NDS Cognitive Labs. “Notice of Privacy.” Ndsognitivelabs.com, ndscognitivelabs.com/notice-of-privacy/.
2. Mohamud, Rowda. “5 Steps to Ensuring Data Privacy in Your Business.” BDC.ca, 20 Oct. 2022, www.bdc.ca/en/articles-tools/blog/5-steps-improving-data-privacy-in-business.