TECNOLÓGICO DE MONTERREY®

Inteligencia artificial avanzada para la ciencia de datos II

*Reto*

# *Reto: Privacidad y Seguridad de datos*


**Diego Arturo Padilla Domínguez - A01552594**

**Keyuan Zhao - A01366831**

**Carolina Herrera Martínez - A01411547**

**Cutberto Arizabalo Nava - A01411431**

**Jose Pablo Cobos Austria - A01274631**


Campus Querétaro

1 de noviembre de 2022

# Privacidad y Seguridad de datos

## Verificación de anonimización

En este punto hablaremos sobre dos dataset importantes, los que se nos otorgaron al inicio del reto por parte del socio formador y los que tenemos actualmente. Empezando por los datos que nos dieron al inicio, estos datos sí cumplen con la anonimización requerida, ya que el socio formador realizó el proceso de enmascaramiento de los datos más sensibles a ser utilizados para identificar a los individuos registrados en el dataset, como puede el nombre y apellido o fechas de los registros que se puedan relacionar con otros datasets y comprometer a los usuarios de quien se obtuvo la información.

Esta anonimización por parte del socio se realizó mediante métodos como:

- La eliminación de columnas que no son relevantes para la solución del reto
- La sustitución de la fecha real por una fecha falsa (día y mes ficticios)
- La aplicación de funciones hash para anonimizar los nombres o identificadores de los usuarios registrados.

Ahora, sobre el dataset que nosotros utilizamos para nuestra matriz de viajes y aplicar el modelo de machine learning, nos hemos asegurado que los datos estén lo suficientemente anonimizados mediante la generalización, ya que este dataset solamente contiene las comunas, las distancias, y la cantidad de viajes que se hizo entre cada una con las horas correspondientes.

## Investigación de documentación técnica

Con base en los datos que nos fueron brindados, debemos de seguir ciertos regímenes, y dado el contexto de que se trabajan en dos países a la vez, se deben de tomar en cuenta las leyes de ambos.

### México

En México existe una ley específica para nuestro contexto, la cual es la Ley Federal de Protección de Datos Personales en Posesión de los Particulares (LFPDPPP), en donde dados los datos poseídos debemos de seguir los siguientes artículos:

**Artículo 6.-** Los responsables en el tratamiento de datos personales, deberán observar los principios de licitud, consentimiento, información, calidad, finalidad, lealtad, proporcionalidad y responsabilidad, previstos en la Ley. (Cámara de diputados de México, 2010)

**Artículo 7.-** En todo tratamiento de datos personales, se presume que existe la expectativa razonable de privacidad, entendida como la confianza que deposita cualquier persona en otra, respecto de que los datos personales proporcionados entre ellos serán tratados conforme a lo que acordaron las partes en los términos establecidos por esta Ley. (Cámara de diputados de México, 2010)

**Artículo 10 Sección 3.-** Los datos personales se sometan a un procedimiento previo de disociación; (Cámara de diputados de México, 2010)

**Artículo 12.-** El tratamiento de datos personales deberá limitarse al cumplimiento de las finalidades previstas en el aviso de privacidad. Si el responsable pretende tratar los datos para un fin distinto que no resulte compatible o análogo a los fines establecidos en aviso de privacidad, se requerirá obtener nuevamente el consentimiento del titular. (Cámara de diputados de México, 2010)

**Artículo 13.-** El tratamiento de datos personales será el que resulte necesario, adecuado y relevante en relación con las finalidades previstas en el aviso de privacidad. En particular para datos personales sensibles, el responsable deberá realizar esfuerzos razonables para limitar el periodo de tratamiento de los mismos a efecto de que sea el mínimo indispensable. (Cámara de diputados de México, 2010)

**Artículo 19.-** Todo responsable que lleve a cabo tratamiento de datos personales deberá establecer y mantener medidas de seguridad administrativas, técnicas y físicas que permitan proteger los datos personales contra daño, pérdida, alteración, destrucción o el uso, acceso o tratamiento no autorizado.

Los responsables no adoptarán medidas de seguridad menores a aquellas que mantengan para el manejo de su información. Asimismo se tomará en cuenta el riesgo existente, las posibles consecuencias para los titulares, la sensibilidad de los datos y el desarrollo tecnológico. (Cámara de diputados de México, 2010)

**Artículo 20.-** Las vulneraciones de seguridad ocurridas en cualquier fase del tratamiento que afecten de forma significativa los derechos patrimoniales o morales de los titulares, serán informadas de forma inmediata por el responsable al titular, a fin de que este último pueda tomar las medidas correspondientes a la defensa de sus derechos. (Cámara de diputados de México, 2010)

**Artículo 21.-** El responsable o terceros que intervengan en cualquier fase del tratamiento de datos personales deberán guardar confidencialidad respecto de éstos, obligación que subsistirá aun después de finalizar sus relaciones con el titular o, en su caso, con el responsable. (Cámara de diputados de México, 2010)

## Chile

En Chile la ley que rige actualmente la protección de los datos personales a modo general, y que desarrolla este derecho constitucional con mayor precisión, es la Ley N° 19.628 en donde los artículos que aplican al proyecto son los siguientes:

**Artículo 1°.-** Toda persona puede efectuar el tratamiento de datos personales, siempre que lo haga de manera concordante con esta ley y para finalidades permitidas por el ordenamiento jurídico. En todo caso deberá respetar el pleno ejercicio de los derechos fundamentales de los titulares de los datos y de las facultades que esta ley les reconoce. (Ministerio secretaría general de la presidencia de Chile, 2020)

**Artículo 7°.-** Las personas que trabajan en el tratamiento de datos personales, tanto en organismos públicos como privados, están obligadas a guardar secreto sobre los mismos, cuando provengan o hayan sido recolectados de fuentes no accesibles al público, como asimismo sobre los demás datos y antecedentes relacionados con el banco de datos, obligación que no cesa por haber terminado sus actividades en ese campo. (Ministerio secretaría general de la presidencia de Chile, 2020)

**Artículo 9°.-** Los datos personales deben utilizarse sólo para los fines para los cuales hubieren sido recolectados, salvo que provengan o se hayan recolectado de fuentes accesibles al público. (Ministerio secretaría general de la presidencia de Chile, 2020)

**Artículo 11°.-** El responsable de los registros o bases donde se almacenen datos personales con posterioridad a su recolección deberá cuidar de ellos con la debida diligencia, haciéndose responsable de los daños. (Ministerio secretaría general de la presidencia de Chile, 2020)

Además de los artículos antes mencionados, el ***Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI),*** nos presenta una guía bastante importante que nos ayuda con el cumplimiento de la ***LFPDPPP.*** En este documento se nos habla sobre varios puntos importantes a considerar, inicia presentándonos un glosario que incluye definiciones de términos que son muy importantes y constantemente utilizados. Nos define que son los datos personales, para qué se pueden utilizar, que es una persona sensible , quien es el responsable del tratamiento de los datos, quién es el titular del tratamiento de los datos, entre otros.

Finalizado ese punto, lo siguiente que nos menciona es una serie de preguntas que se deben de hacer los encargados del tratamiento de los datos en todas las fases del desarrollo de un proyecto para asegurar la privacidad y seguridad de los datos adquiridos.

Imagen 1. *Propuestas de preguntas para asegurar privacidad de datos*

Asimismo, en conjunto con esta serie de preguntas guía, también se nos muestra un conjunto de principios y obligaciones que debemos de cumplir con eficiencia. Se nos presenta su definición, de que sirve, de qué manera nosotros como encargados podemos cumplirlas y lo más útil que nos otorga es una checklist por cada uno de estos principios, lo que nos facilita mucho checar que el cumplimiento sea correcto.


Imagen 2. *Conjunto de principios de seguridad*

## Procedimientos y reglas para almacenamiento y acceso de los datos

El dataset compartido por parte del socio formador está almacenado de forma local y virtual (en la nube), y sólo las personas autorizadas pueden acceder.

- **Local**

Todos los integrantes del equipo tienen almacenado el dataset de forma local para realizar cualquier tipo de pruebas necesarias para la implementación del modelo, ya que de esta forma se agiliza el tiempo. Dentro del equipo, contamos con cuatro dispositivos con sistema operativo de Windows 10 y uno de Arch Linux, cada equipo es personal y cuenta con contraseña para el uso del mismo.

- **Virtual**

| Nombre | Propietario | Tipo de seguridad | Descripción |
|---|---|---|---|
| Drive | Google | Google proporciona seguridad incorporada que permite detectar y bloquear amenazas como spam, suplantación de identidad (phishing) y software malicioso. | Lo cual fue originalmente compartido por parte del socio formador hacia nosotros y sólo las personas que tienen el enlace pueden acceder. Además, creamos una unidad compartida para guardar el dataset y otros códigos de |

| | | | |
|---|---|---|---|
| | | Todos los archivos cargados o creados dentro de Drive serán encriptados de tipo AES256 bit encryption. | procesamiento, esto sólo los integrantes del equipo pueden tener acceso hacia ello. Dentro de este se puede visualizar quienes tienen acceso a las carpetas compartidas y qué reglas tienen cada una de ellas (leer, comentar o editar). |
| Servidor en la nube | Oracle | Todos los archivos cargados o creados dentro del servidor serán encriptados de tipo 256-bit Advanced Encryption Standard (AES-256).<br><br>Para la llave SSH se usó Putty Key Generador con el tipo SSH-2 RSA key. | Es un servidor montado en la plataforma de Oracle en el cual para acceder a ella se necesita establecer configuraciones con ayuda de algún software como PuTTY, estableciendo el puerto, el usuario y la contraseña del servidor para poder acceder. No hay forma de visualizar quienes accedieron a los códigos y al dataset, por lo tanto, establecimos sólo a las personas tanga la llave privada, el usuario y la contraseña para acceder. |

# Referencias


- Cámara de diputados de México, (2010). *LEY FEDERAL DE PROTECCIÓN DE DATOS PERSONALES EN POSESIÓN DE LOS PARTICULARES.* Secretaria General. Recuperado el 2/11/2022 de: https://rb.gy/sqvgl4

- SalesForce, (s/f). *Ley de Protección de Datos Personales en México.* Salesforce. Recuperado el 2/11/2022 de: https://rb.gy/od6yht

- Inai, (2016). *Guía para cumplir con los principios y deberes de la Ley Federal de Protección de Datos Personales en Posesión de los Particulares.* Recuperado el 2/11/2022 de: https://rb.gy/okjwd7

- Ministerio secretaría general de la presidencia de Chile. (2020, 26 agosto). *Ley 19628 sobre protección de la vida privada*. www.bcn.cl/leychile. https://www.bcn.cl/leychile/navegar?idNorma=141599
- Drive, (s/f). Get started with encrypted files in Drive, Docs, Sheets & Slides. https://support.google.com/docs/answer/10519333?hl=en#:~:text=All%20files%20uploaded%20to%20Drive,with%20Workspace%20Client%2Dside%20encryption
- Oracle, (s/f). Oracle Cloud Infrastructure Security Guide. From: https://docs.oracle.com/en-us/iaas/Content/Security/Concepts/security_guide.htsm