02/12/2023

By: Alejandra Ossa Yepes

Asignatura: Ética de la inteligencia artificial

Universidad de los Andes

**Evaluación de Impacto de la compaña Nova Kirin**

**Introducción**

En el siguiente texto se presenta un análisis del desarrollo del producto Nova Kirin realizado por la compañía Kirin Services que con ayuda de su modelo de inteligencia artificial busca realizar una identificación de los comportamientos de los clientes y poder clasificar sus emociones; por lo que se realizará la evaluación de impacto para la identificación de los potenciales riesgos y la mitigación de estos desde una mirada jurídica y ética, así como una reflexión sobre la utilización y recolección de los datos requeridos para estos tipos de modelos en función de la seguridad y privacidad de las información de los clientes.

**Descripción del Sistema Evaluado**

Kirin Services es una empresa multinacional especializada en ofrecer servicios diseñados para mejorar la atención al cliente en empresas de diversos sectores a nivel mundial. En su intento de generar una experiencia agradable a sus clientes la compañía está desarrollando un sistema basado en inteligencia artificial que tiene como objetivo identificar los comportamientos de los clientes y así poder optimizar la experiencia del servicio que se les brinda. Haciendo uso de las nuevas tecnologías, Kirin Services utiliza un modelo de IA llamado Nova Kirin alimentado por información de los clientes como imágenes, datos personales, audios y llamadas para la identificación del estado emocional de sus clientes y patrones en su conducta con el fin de generar una trato personalizado con respecto a las características obtenidas y que la experiencia del cliente sea las más confortable y única. En la información suministrada por Kirin Services no se identificó el proceso de autorización de tratamiento y recolección de los datos personales con la metodología implementada por la compañía, violando así la privacidad y seguridad de los clientes. Por lo que en el presente texto se realizará una evaluación de impacto debido a que es una herramienta que permite a todos los entes involucrados desarrolladores, empresas y sector público, evaluar los potenciales riesgos de un

sistema de decisiones automatizadas y trabajar en formas de mitigar los riesgos asociados antes del lanzamiento del  nuevo sistema para la identificación de riesgos, análisis ético, evaluación de impacto social y alineación de intereses.

Un aspecto clave de la evaluación de impacto tiene que ver con la noción que se tiene sobre los conceptos utilizados para la implementación del sistema y definiciones asociadas al problema descrito, por lo que a continuación se hará una breve descripción de estos conceptos.

La Inteligencia Artificial hace referencia a la capacidad de los sistemas informáticos para llevar a cabo tareas que requieren de inteligencia humana y se basa en algoritmos y modelos que permiten a las máquinas analizar grandes cantidades de datos y extraer conclusiones, identificar tendencias y tomar decisiones informadas; en nuestro caso es el sistema utilizado para la identificación del estado emocional de los clientes por medio de decisiones automatizadas y su clasificación en patrones de conducta. Todo esto va muy de la mano con los conceptos asociados a los tipos de datos con los que se desarrolló el modelo de IA,  los cuales en su mayoría son datos personal que hacen referencia a cualquier pieza de información vinculada a una o varias personas determinadas o determinables o que puedan asociarse con una persona natural o jurídica, en este caso a clientes de la compañía Kirin Services, estos datos pueden ser públicos, semiprivados o privados como son las imágenes y audios recolectados. Es de suma importancia conocer el principio de finalidad en el que se obliga a que las actividades de recolección de datos personales obedezcan a una finalidad legítima de acuerdo con la Constitución y la ley,  por lo que con este principio la finalidad debe ser comunicada al titular de la información previa con el otorgamiento del titular de la autorización, cuando ella sea necesaria o, en general, siempre que el titular solicite información al respecto, debido a esto se ve necesario un proceso de autorización de tratamiento y recolección de datos debido a los diferentes métodos que utilizo la compañía Kirin Services.

Estos sistemas de IA deben estar regidos por principios éticos y valores en el desarrollo, esto implica considerar el impacto social, la equidad, la privacidad, la transparencia y la responsabilidad en todas las etapas del ciclo de vida de la inteligencia artificial. Cuando un sistema empleado en una compañía hace uso de videos y audios de sus clientes como lo hace la compañía Kirin Services  para la identificación de sus estados de ánimo, pueden verse afectados los derechos de terceros como lo son el derecho a la privacidad que al realizarse grabaciones de videos y audios de los clientes puede implicar una invasión de su privacidad, especialmente si no se obtiene su consentimiento adecuadamente, asi como el derecho a la intimidad debido a que la captura de videos y audios violar este derecho, ya que se están registrando aspectos de su vida personal y comportamientos del cliente sin previa autorizacion, y por último se viola el derecho al buen nombre y a la honra si el material recolectado se  utilizan de manera inapropiada o se divulgan sin consentimiento, los clientes pueden ver afectada su reputación y buen nombre.

En cuanto al impacto social, el uso de este tipo de sistemas puede generar preocupaciones en términos de privacidad y confianza de los clientes. Si los clientes sienten que sus derechos no están siendo respetados o que su privacidad está siendo invadida, esto puede afectar la relación con la compañía y generar una percepción negativa en la sociedad.

Cuando se manipula este tipo de información, existen posibles brechas de seguridad que podrían comprometer la privacidad y confidencialidad de los datos recopilados y generar brechas en las que se  podrían incluir la interceptación no autorizada de los datos por parte de terceros, el acceso no autorizado a la información almacenada en el sistema o la falta de medidas de seguridad adecuadas para proteger los datos, por lo que se debería  implementar medidas de seguridad, como autenticación de usuarios, control de acceso y cifrado de datos como lo es la anonimización, para evitar accesos no autorizados.

Para la mitigación de todos los riesgos anteriormente mencionados se puede obtener un consentimiento explícito y claro que informe a los clientes antes de recopilar y utilizar sus datos personales mencionando el propósito de la recopilación de datos cómo se utilizarán y con quién se compartirán, implementar medidas de seguridad adecuadas para proteger los datos recopilados como la gestión de identidad, así como limitar la cantidad de datos recopilados a lo estrictamente necesario para la identificación de los estados de ánimo de los clientes y su conducta; que el sistema implementado sea preciso y evite sesgos utilizando un conjunto de datos de entrenamiento que sea representativo de la población a la que se aplicará el modelo de IA. Evitar la recopilación innecesaria de información personal que no esté directamente relacionada con el propósito del sistema donde se pueda evidenciar políticas claras sobre el tiempo durante el cual se almacenarán los datos recopilados para ayudará a limitar el riesgo de retener datos personales durante períodos innecesariamente largos.

Es de suma importancia que se pueda reconocer el derecho de hábeas data que es aquel que tiene toda persona de conocer, actualizar y rectificar la información que se haya recogido sobre ella en archivos y bancos de datos de naturaleza pública o privada, así como tomar medidas genéricas desde cualquier rol de la organización para resguardar los derechos del cliente, como la educación y sensibilización en los empleados de la compañía sobre los riesgos asociados al uso de sistemas de reconocimiento de emociones con un uso inadecuado de estos y realizar una supervisión y control apropiados del uso del modelo de IA para garantizar que se utilice de manera ética y responsable, así como que sea transparente sobre su uso y sobre los datos personales que se recopilan.

En resumen estas medidas buscan:

* Crear un comité de ética encargado de probar el modelo de IA y detectar sesgos.
* Hacer pública toda la información sobre el sistema de decisión automatizada y permitir la participación de expertos externos.
* Establecer un mecanismo para recibir comentarios y evaluar su aplicabilidad a través de un organismo de supervisión interno.
* Establecer un cronograma para implementar la evaluación de impacto algorítmico de manera continua.
* Obtener muestras significativas que representen a los diferentes tipos de clientes.

En conclusión para mitigar los riesgos y garantizar la privacidad y seguridad de la información en el sistema de la compañía Nova Kirin, se pueden implementar medidas como contar con políticas de privacidad y seguridad de la información dado que se requiere de un consentimiento informado de la recolección, tratamiento y disposición de la información que se utilizará, así como habilitar un área encargada del seguimiento y control de este, en donde se puede efectuar un protocolo de la estructura en la que sea pueda examinar el proceso desde la autorización de los datos, recolección, almacenamiento e implementación en el modelo para garantizar un debido proceso de este y en caso de que se presente un problema en esta estructura se pueda reaccionar de forma proactiva. Esta aérea deberá ser interdisciplinaria dado que se requieren conocimientos técnicos y jurídicos para el entendimiento y justificación de los procesos descritos en cada uno de los pasos de la estructura anteriormente mencionada, con el fin de garantizar una mirada holística del funcionamiento del sistema y su relación con la interacción social.