Licenciamiento de IAGen y Protección de Datos

Un análisis comparativo de riesgos y garantías desde una perspectiva de gobernanza

Marzo 2026

IALAB

Dirección

Juan Gustavo **Corvalán**

Mariana **Sánchez Caparrós**

Carina **Papini**

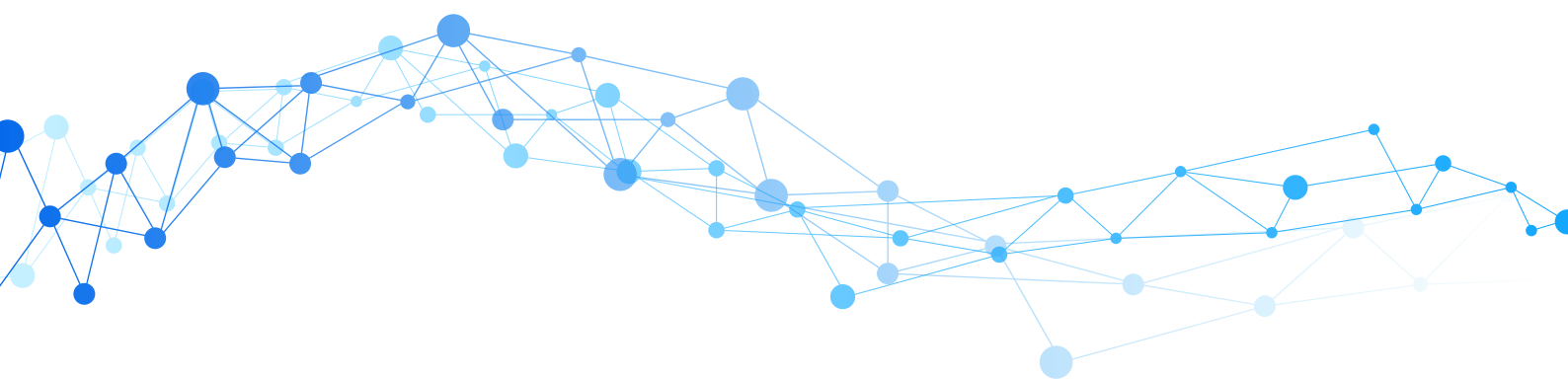
Equipo de investigación

Lola **Ramos Pereyra**

Gisel **Alvarado**

Diseño

Sofía **Rolleri**



Índice

Resumen ejecutivo

La adopción de herramientas de inteligencia artificial generativa (IAGen), como ChatGPT, Gemini o Claude, entre otras, se ha acelerado de manera transversal tanto en organizaciones públicas como privadas.

Este proceso abre oportunidades significativas en términos de eficiencia e innovación, pero también plantea desafíos relevantes, especialmente en lo que respecta a la protección de la privacidad y de los datos personales de las personas involucradas en la información que es procesada mediante estos sistemas.

En este contexto, un aspecto central, y a menudo subestimado, es que el nivel de protección de los datos ingresados por los usuarios de estas plataformas no es uniforme, sino que varía significativamente según el tipo de licenciamiento ofrecido por los proveedores de IAGen.

El presente documento identifica algunos patrones estructurales, asimetrías de riesgo y diferencias sustantivas en las garantías de protección de datos según se trate de licencias de consumo, empresariales (enterprise) o de uso vía API. Estas conclusiones se presentan de manera sintética para orientar la toma de decisiones institucionales, con foco en el uso de los datos para entrenamiento, los niveles de confidencialidad, los esquemas de retención y la asignación de responsabilidades legales.

A fin de permitir una profundización técnica y contractual por parte del lector, el documento pone a disposición un cuadro comparativo que sistematiza las cláusulas relevantes de términos de uso y políticas de privacidad de los proveedores analizados.

Dicho material complementario constituye la base documental del análisis y permite verificar y ampliar los hallazgos aquí presentados, sin sobrecargar el cuerpo principal del texto.

El cuadro comparativo se actualizará de manera periódica, con el objetivo de reflejar las modificaciones relevantes que surjan en las políticas de privacidad de los distintos proveedores analizados, así como de incorporar nuevos componentes de información que resulten pertinentes.

De este modo, el cuadro se concibe como un insumo dinámico y en permanente evolución, destinado a brindar al lector material actualizado y confiable para la toma de decisiones informadas en materia de adopción y uso de soluciones de inteligencia artificial generativa.

01

Introducción

Una de las principales variables a considerar a la hora de una incorporación ética de modelos comerciales basados en IAGEN en las organizaciones consiste en conocer el tipo de licenciamiento bajo el cual se accede a ellos.

Ello por dos causas fundamentales.

En primer lugar, el tipo de licenciamiento posee relevancia para la gobernanza institucional y el abordaje ético de las decisiones de adopción de IAGen tanto en el ámbito público como privado.

En segundo lugar y en relación a ello, las modalidades de acceso a un mismo modelo de lenguaje implican distintos riesgos y ofrecen distintas garantías en relación a la protección de los datos personales e información que queremos resguardar.

La consideración de estas cuestiones cobra relevancia dentro de las organizaciones para lograr aplicaciones que permitan aprovechar los beneficios de la IAGEN, con la incorporación de medidas de mitigación apropiadas para hacer frente.

Se trata de un aspecto que, en ciertos casos, suele ser poco considerado por las organizaciones que deciden trabajar con modelos de IAGEN. Pese a su relevancia, las principales discusiones suelen centrarse en las capacidades propiamente dichas de los modelos y el impacto más perceptible por los usuarios.

Este aspecto se vuelve aún más crítico en contextos donde el uso de herramientas de IAGen se produce de manera descentralizada, informal o no autorizada dentro de las organizaciones, fenómeno habitualmente denominado *shadow AI*.

Desde IALAB, consideramos que el análisis comparativo de términos de uso y políticas de privacidad de distintos proveedores comerciales de IAGen pone de manifiesto que el licenciamiento opera como un **determinante clave del riesgo jurídico y organizacional**, especialmente en escenarios donde la ausencia de lineamientos institucionales claros amplifica las asimetrías de protección.

Por ello, el examen de estas políticas debería integrar de manera sistemática cualquier evaluación seria orientada a la incorporación institucional de este tipo de tecnologías.

A continuación, nos ocuparemos de introducir el análisis de los tipos de licenciamiento.

02

Metodología y alcance

La presente investigación se basa en el relevamiento sistemático de los términos de uso y políticas de privacidad de los principales proveedores comerciales de inteligencia artificial generativa.

El objetivo metodológico fue identificar patrones comunes, diferencias relevantes y zonas de riesgo recurrentes en el tratamiento de los datos de los usuarios, en función del tipo de licenciamiento ofrecido.

Para ello el análisis se concentró específicamente en cláusulas vinculadas con:

- El uso de los datos ingresados por los usuarios
- El entrenamiento y la mejora de los modelos
- La retención, el almacenamiento y la eliminación de los datos
- La existencia de revisión humana y mecanismos de control
- Las transferencias internacionales de datos
- La asignación de responsabilidades y riesgos legales entre proveedores y usuarios.

El enfoque adoptado es transversal y comparativo, lo que permite ofrecer un relevamiento inicial para una lectura analítica orientada a la toma de decisiones institucionales.

Con el fin de garantizar mayor transparencia y permitir una lectura en profundidad de los hallazgos aquí sintetizados, el análisis presentado se apoya en un cuadro comparativo que sistematiza, proveedor por proveedor, las principales cláusulas de términos de uso y políticas de privacidad vinculadas con el tratamiento de datos en servicios de inteligencia artificial generativa.

Dicho cuadro permite identificar de manera detallada las diferencias y similitudes entre tipos de licenciamiento. Su consulta resulta especialmente útil para lectores que requieran un mayor nivel de detalle técnico o contractual, sin que ello sea necesario para la comprensión general de los argumentos desarrollados en el presente documento.

03

Hallazgos principales

3.1 Licencias de consumo: utilidad con riesgo estructural

Las licencias de consumo o freemium presentan un patrón consistente: ofrecen acceso amplio y facilidad de uso, pero trasladan gran parte del riesgo vinculado a la protección de los datos al usuario.

En este tipo de licencias es frecuente que los datos ingresados puedan utilizarse para entrenar o mejorar modelos, ya sea por defecto o salvo exclusión expresa. A ello se suma la posibilidad de revisión humana con fines de calidad, seguridad o desarrollo (ej. caso de Gemini de Google), y advertencias explícitas que recomiendan no ingresar datos personales o sensibles (ej. caso de Grok de xAI y Gemini de Google).

Desde una perspectiva institucional, estas condiciones resultan problemáticas: la protección de datos depende más del comportamiento individual que de garantías contractuales o técnicas.

3.2 Licencias enterprise: gobernanza contractual y control

Las licencias empresariales introducen un cambio cualitativo. Aparecen acuerdos de procesamiento de datos (DPA), deberes reforzados de confidencialidad, delimitación clara de la finalidad del tratamiento y, en muchos casos, compromisos explícitos de no utilización de los datos del cliente para entrenamiento.
Si bien no eliminan completamente el riesgo, estas licencias permiten articular esquemas de auditoría y cumplimiento normativo más robustos, especialmente relevantes para organizaciones públicas o sectores regulados.

3.3 Uso vía API: mayor control técnico, riesgos por configuración

El acceso a modelos de IAGen mediante API, las pautas contractuales aplicables suelen combinar retenciones de datos más cortas y mayores capacidades de control técnico. No obstante, el análisis revela que este modelo no es intrínsecamente “seguro”.

Persisten riesgos asociados a revisiones humanas, almacenamiento transitorio, transferencias internacionales y, especialmente, a la ambigüedad en la asignación de roles legales entre proveedor y cliente (controlador vs. procesador). Sin un encuadre contractual adecuado, el riesgo puede desplazarse hacia el desarrollador o la organización usuaria.

04

Comparación de riesgos y garantías

Desde una perspectiva comparativa, los mayores riesgos para la protección de datos se concentran en las licencias de consumo, mientras que las licencias enterprise ofrecen las mayores garantías de gobernanza. Las API ocupan una posición intermedia, donde el nivel real de riesgo depende fuertemente del diseño de la arquitectura, la configuración elegida y los acuerdos contractuales complementarios.

La diferencia clave no radica únicamente en la tecnología, sino en el marco jurídico y organizacional que la rodea.

05

Implicancias para la toma de decisiones

El análisis permite extraer tres implicancias centrales:

1 La elección del licenciamiento debe considerarse una decisión estratégica de gobernanza, no meramente presupuestaria. Esto significa que, frente a cada posibilidad de caso de uso de IAGEN, es recomendable realizar un control de proporcionalidad mediante el cual se analice el impacto en relación a los datos e información asociados al caso de uso y se establezca la medida de protección más adecuada para hacerle frente. Por ese motivo, la decisión en relación a uso de versión de consumo, enterprise, vía API o utilizando modelos locales, terminará siendo casuística.

2 El uso de herramientas de consumo para fines institucionales con datos reales genera riesgos difíciles de justificar desde el cumplimiento normativo, cuestión que se complejiza en mayor medida cuando se introducen datos sensibles, cuyo tratamiento se encuentra prohibido como principio general, salvo las excepciones expresamente previstas en las normas.

3 Incluso con licencias enterprise o API, la ausencia de políticas internas, clasificación de datos y controles operativos puede neutralizar las garantías contractuales.

Existen dos puntos adicionales a considerar.
Por un lado, más allá de las cláusulas contractuales sobre entrenamiento o confidencialidad, subsisten dimensiones técnicas que deben ser consideradas: la transferencia internacional de datos, el almacenamiento en infraestructuras cloud de terceros y la localización geográfica de los servidores.

Aún bajo licencias enterprise, los datos pueden ser:

- Transferidos a jurisdicciones con estándares de protección distintos.
- Almacenados en centros de datos gestionados por proveedores de infraestructura (ej. servicios de nube).
- Procesados de manera transitoria para logging, monitoreo de seguridad o mejora del servicio.

Estas prácticas no implican necesariamente reutilización para entrenamiento, pero podrían generar exposición jurídica y operativa. Desde una perspectiva institucional, la pregunta relevante no es únicamente si los datos se usan para entrenar modelos, sino:

- ¿Dónde se almacenan?
- ¿Por cuánto tiempo?
- ¿Bajo qué régimen jurídico?
- ¿Qué subprocesadores intervienen?
- ¿Qué ocurre ante incidentes de seguridad o cambios regulatorios?

Las respuestas pueden conducir a que el usuario tome la decisión más adecuada para el uso responsable.

Por otro lado, no deberíamos omitir la cuestión inherente a la responsabilidad derivada de usos indebidos, pérdidas o accesos no autorizados.

En la Política de Privacidad de Open AI[1], por ejemplo, se indica que aplican medidas técnicas, administrativas y organizativas comercialmente razonables para proteger los datos personales frente a cualquier pérdida, uso indebido y acceso, revelación, alteración o destrucción no autorizados. Sin embargo, agregan: “…no nos hacemos responsables de la elusión de cualquier configuración de privacidad o medidas de seguridad contenidas en el Servicio o en sitios web de terceros”.

El punto es determinar el alcance de la palabra elusión, cuestión no aclarada en el texto. Podría significar, por ejemplo, hackeo, Vulneración de credenciales, Ataques externos, Fallas derivadas de servicios de terceros como el proveedor cloud.

06

Consideraciones sobre el uso de modelos locales y alternativas a la contratación comercial

Si bien el presente trabajo se centra en el análisis de proveedores comerciales de inteligencia artificial generativa y en las implicancias del licenciamiento asociado a su adopción, resulta relevante señalar que la contratación de servicios externos no constituye la única vía posible para incorporar IAGen en contextos institucionales.

En determinados escenarios (especialmente en organizaciones que procesan información sensible, confidencial o sujeta a regímenes estrictos de protección de datos) el uso de modelos de lenguaje desplegados en entornos locales o controlados puede representar una alternativa válida desde la perspectiva de la soberanía de los datos, la minimización de riesgos y el cumplimiento normativo.

Este enfoque permite evitar transferencias internacionales, reducir dependencias contractuales y ejercer un mayor control sobre el ciclo de vida de la información procesada.

No obstante, la adopción de modelos locales no está exenta de desafíos. Su viabilidad depende de la disponibilidad de infraestructura técnica adecuada, en particular capacidades de hardware, gestión de seguridad, mantenimiento y personal especializado, así como de la necesidad de garantizar una experiencia de usuario aceptable en términos de desempeño, disponibilidad y calidad de respuesta. En ausencia de estas condiciones, el riesgo puede desplazarse desde lo jurídico hacia lo operativo y organizacional, con los equipos dejando de utilizar la propuesta institucional para migrar a otras no autorizadas (refuerzo del fenómeno de Shadow AI).

En consecuencia, la elección entre soluciones comerciales, modelos locales o esquemas híbridos no debe plantearse en términos absolutos, sino como parte de una decisión estratégica de gobernanza, basada en el tipo de datos involucrados, el nivel de riesgo aceptable, las capacidades técnicas disponibles en las organizaciones y los objetivos institucionales perseguidos.

07

Anonimización de datos

[1]Política de privacidad | OpenAI

De manera adicional, siempre es importante considerar la anonimización de datos, que constituye una medida estratégica y transversal a los distintos esquemas de licenciamiento, siempre que se implemente de manera técnicamente robusta y jurídicamente adecuada.

La anonimización no se limita a la supresión de identificadores directos, sino que exige la adopción de medidas técnicas y organizativas orientadas a impedir razonablemente la reidentificación de las personas, incluso mediante la combinación con otras fuentes de datos disponibles. Su evaluación debe realizarse en términos de riesgo residual aceptable, considerando el contexto tecnológico, la disponibilidad de datos externos y la capacidad de un eventual atacante razonablemente motivado.

La implementación de procesos de anonimización robusta presenta desafíos significativos para los responsables dentro de las organizaciones, entre ellos:

- La seguridad del mecanismo y las herramientas empleadas. La utilización de aplicaciones en línea gratuitas puede implicar transferencia internacional de datos, almacenamiento temporal o tratamiento ulterior no previsto, aumentando los riesgos en lugar de mitigarlos.

- El tiempo y los recursos necesarios para aplicar técnicas adecuadas de anonimización, especialmente cuando se requiere intervención manual o revisión contextual.

- La preservación de la utilidad de los documentos anonimizados y su eventual reutilización en procesos posteriores.

Como ejemplo de implementación institucional orientada a la mitigación de riesgos, IALAB colaboró en el diseño de Privacy AI Studio, una plataforma de privacidad que integra, en síntesis:

- Anonimización robusta local de documentos, con opciones automáticas y manuales según preferencia del usuario.

- Definición de etiquetas personalizadas para la anonimización de datos por proyectos, en función del requerimiento y el riesgo específico.

- Uso de inteligencia artificial generativa en entorno local o exclusivamente sobre documentos previamente anonimizados.

- Transcripción y anonimización de audiencias.

08

Conclusiones

La adopción responsable de inteligencia artificial generativa requiere mirar más allá del modelo en sí y atender a las condiciones de uso que estructuran el tratamiento de los datos, los niveles de control y la asignación de responsabilidades.

Tal como muestra el análisis realizado, el tipo de licenciamiento bajo el cual se accede a estas tecnologías actúa como un verdadero determinante de riesgo jurídico y organizacional, y su evaluación debería integrarse de manera sistemática en los procesos de análisis de impacto, las decisiones de compra y las estrategias institucionales de gobernanza de IA.

Al mismo tiempo, el licenciamiento comercial no constituye la única alternativa posible para la incorporación de IAGen en contextos institucionales. En determinados escenarios, el uso de modelos desplegados en entornos locales o controlados, o bien esquemas híbridos, puede ofrecer ventajas relevantes en términos de soberanía de los datos, reducción de transferencias internacionales y fortalecimiento del control sobre la información, siempre que se cuente con las capacidades técnicas y organizacionales necesarias para sostenerlos adecuadamente.

Comprender estas diferencias (entre licencias de consumo, enterprise, uso vía API y alternativas basadas en modelos locales) no solo permite mitigar riesgos legales y de cumplimiento, sino también habilitar decisiones más informadas, coherentes y alineadas con los valores institucionales, promoviendo usos de la inteligencia artificial generativa que sean seguros, transparentes y socialmente responsables.

Anexo I - Cuadro comparativo de licencias de plataformas de IAGen


FACULTAD DE DERECHO
U.B.A.
IALAB