Clawdbot y la Ilusión de la Privacidad: Un Análisis Crítico de los Riesgos de Incumplimiento del RGPD en la IA Agentiva (R.S.C. Derecho Artificial)

Introducción: La Promesa de Soberanía de Datos en la Era de los Agentes de IA

En un panorama digital marcado por la creciente preocupación sobre la privacidad de los datos, el auge de asistentes de inteligencia artificial autoalojados como Clawdbot se presenta como una solución atractiva. Estas herramientas se comercializan como una alternativa a los servicios de IA basados en la nube, prometiendo a los usuarios un control total y soberanía sobre su información personal. Sin embargo, este informe argumenta que, a pesar de su marketing centrado en la privacidad, la arquitectura operativa de Clawdbot —que depende intrínsecamente de la conexión con modelos de lenguaje grandes (LLM) de terceros basados en la nube, como los de Anthropic— crea un panorama de cumplimiento del Reglamento General de Protección de Datos (RGPD) complejo y fundamentalmente engañoso.

Sostenemos que este modelo híbrido impone a sus usuarios, sin la debida transparencia, el rol y las abrumadoras responsabilidades de un "Responsable del Tratamiento" de datos bajo el RGPD. Esta imposición expone a individuos y organizaciones a significativos riesgos legales y de cumplimiento que son directamente contrarios a la promesa de privacidad del producto. A continuación, procederemos a deconstruir la arquitectura técnica de Clawdbot para desvelar esta contradicción fundamental y sus profundas implicaciones normativas.

---

1. Deconstrucción de Clawdbot: Arquitectura Técnica vs. Marketing de Privacidad

El análisis riguroso de la arquitectura real de una herramienta de IA es un paso estratégico indispensable para verificar sus afirmaciones de marketing, especialmente aquellas relacionadas con la privacidad y el control de datos. En el caso de Clawdbot, este examen revela una disonancia crítica entre su propuesta de valor declarada y su funcionam7iento práctico.

Basándonos en su descripción, Clawdbot no es un modelo de lenguaje (LLM) en sí mismo, sino un ***agente orquestador*** de código abierto y autoalojado. Su función

principal es actuar como un intermediario inteligente que conecta diversas aplicaciones y ejecuta tareas complejas, pero depende "100% de aquello a lo que lo conectes para la inferencia". Esta arquitectura se alinea directamente con el paradigma de la "IA Agentiva" y la "Orquestación de Modelos" analizado por el Comité Europeo de Protección de Datos (CEPD). Clawdbot funciona como un sistema autónomo que combina las capacidades de un LLM con módulos de percepción, razonamiento, planificación y acción para realizar tareas complejas. Su diseño como orquestador, que por definición conecta componentes locales con servicios externos, implica inherentemente flujos de datos que trascienden la máquina del usuario, sentando las bases de sus riesgos de privacidad.

Capacidades Clave de Clawdbot

- **Integración con Mensajería:** Se conecta de forma nativa con aplicaciones como WhatsApp, Telegram y Discord, permitiendo la interacción a través de estas plataformas.
- **Automatización de Tareas:** Diseñado para funcionar como un agente automático que puede ejecutar una variedad de tareas complejas.
- **Gestión de "Memoria" y "Habilidades":** Ofrece control sobre la memoria contextual y las habilidades del agente, a diferencia de las interacciones sin estado de muchos sistemas de IA.
- **Acceso al Sistema Local:** Posee la capacidad de acceder al sistema de archivos del usuario, una característica que los asistentes en la nube no pueden ofrecer.

La crítica fundamental, expuesta con precisión por el usuario de Reddit swwright, desmantela la principal ventaja de privacidad de Clawdbot. Argumenta que la conexión de la herramienta a servicios en la nube como OpenAI o Anthropic/Claude anula por completo el supuesto beneficio:

"Si conectas Clawdbot a OpenAI o Anthropic/Claude o cualquier otra IA alojada en la nube como sugieres, tus datos siguen yendo a la nube, punto."

Esta realidad técnica entra en conflicto directo con las afirmaciones de marketing de que Clawdbot ofrece "control total sobre tus datos" y "prioriza la privacidad". Si bien el componente de orquestación se aloja localmente, el procesamiento del lenguaje y la inferencia —el núcleo de la operación— externalizan los datos del usuario a terceros. Como concluye swwright, el verdadero valor de Clawdbot no es la privacidad, sino su capacidad para ser un "agente automático que operará con una variedad de LLMs".

Esta disonancia entre la arquitectura técnica y el mensaje de marketing no es un mero detalle técnico; tiene consecuencias legales ineludibles bajo el RGPD que recaen directamente sobre el usuario.

---

2. El Despliegue de Clawdbot bajo el RGPD: La Asunción del Rol de Responsable del Tratamiento

La calificación jurídica de los actores en un ecosistema de IA es el pilar sobre el que se construye la responsabilidad en materia de protección de datos. En el marco del RGPD, es crucial determinar quién decide los fines y los medios del tratamiento de datos personales, ya que de ello se derivan obligaciones legales concretas. Según el RGPD, los roles principales son el **Responsable del Tratamiento** (quien determina los fines y medios) y el **Encargado del Tratamiento** (quien trata datos por cuenta del responsable).

La configuración operativa de Clawdbot conectado a un proveedor como Anthropic encaja en el "Modelo de LLM como Servicio" analizado por el CEPD. Bajo esta definición, cualquier individuo u organización que implemente Clawdbot para un fin que exceda una actividad puramente personal o doméstica —por ejemplo, para gestionar las comunicaciones de un negocio— asume *de iure* el rol y la totalidad de las obligaciones del Responsable del Tratamiento. Es el usuario de Clawdbot quien decide *por qué* se procesan los datos (ej. atender a un cliente) y *cómo* se hace (ej. usando Clawdbot conectado a la API de Anthropic).

Crucialmente, este modelo no establece una simple relación Responsable-Encargado. El proveedor del LLM (Anthropic) se reserva el derecho de acceder a los datos para sus propios fines, como las revisiones de su Política de Uso ("Trust & Safety"). Como indica el análisis del CEPD, cuando un proveedor procesa datos para sus propios fines, también actúa como Responsable. Esto crea una situación de **corresponsabilidad de facto** o de responsables independientes, donde el usuario de Clawdbot y el proveedor del LLM comparten la responsabilidad sobre los datos, una complejidad legal que el marketing de privacidad ignora por completo.

Tabla 1: Contraste de Responsabilidades bajo el RGPD

Usuario de Clawdbot (Responsable)	Proveedor de LLM (Encargado/Corresponsable)
Garantizar una Base Jurídica: Debe identificar y documentar una base legal válida (ej. consentimiento, interés legítimo) para todo el tratamiento.	**Procesamiento bajo Instrucción:** Procesa los datos siguiendo las instrucciones del Responsable, pero con matices significativos.

Cumplir con la Transparencia: Tiene la obligación de informar a los interesados (ej. sus clientes en WhatsApp) sobre el tratamiento y todos los actores involucrados.	**Seguridad de la Infraestructura:** Es responsable de implementar medidas de seguridad técnica para proteger los datos en sus sistemas.
Asegurar los Datos: Debe implementar medidas técnicas y organizativas para proteger los datos, tanto en su sistema local como durante el tránsito.	**Actúa como Responsable para fines propios:** Al utilizar datos para revisión de seguridad, asume el rol de Responsable, creando una corresponsabilidad de facto con el usuario.
Gestionar Transferencias Internacionales: Es su responsabilidad garantizar que cualquier transferencia de datos fuera del EEE cumple con el RGPD.	**Facilitar el Cumplimiento:** Debe ayudar al Responsable a cumplir sus obligaciones, aunque su rol dual complica la delimitación de responsabilidades.
Facilitar los Derechos de los Interesados: Debe establecer procedimientos para atender las solicitudes de acceso, rectificación o supresión de datos.	

La aparente simplicidad de operar Clawdbot enmascara una carga de cumplimiento normativo compleja que recae enteramente sobre el usuario. Esta es una implicación crítica que su marketing omite, llevando a los usuarios a un falso sentido de seguridad mientras asumen, sin saberlo, riesgos legales sustanciales. A continuación, se analizarán los incumplimientos específicos que se derivan de esta situación.

---

3. Análisis de Riesgos de Incumplimiento Específicos del RGPD

Una vez que el usuario de Clawdbot es identificado como Responsable del Tratamiento, es imperativo analizar los incumplimientos concretos del RGPD que surgen del uso de la herramienta en su configuración operativa recomendada, es decir, conectada a un LLM de terceros.

3.1. Vulneración de los Principios de Transparencia y Lealtad (Art. 5.1.a): La Falsa Promesa de Privacidad

El principio de tratamiento leal y transparente, consagrado en el Artículo 5(1)(a) del RGPD, exige que los datos personales sean tratados de manera lícita, leal y transparente en relación con el interesado. El marketing de Clawdbot, que promete "privacidad" y "control total sobre tus datos", entra en contradicción directa con el flujo real de datos hacia proveedores en la nube como Anthropic. Esta falta de transparencia tiene una consecuencia grave: impide que el usuario final (por ejemplo, la persona que interactúa con el agente a través de WhatsApp) otorgue un consentimiento informado y válido, ya que desconoce que sus conversaciones son enviadas, analizadas y potencialmente almacenadas por una empresa externa. Este ocultamiento del verdadero destino de los datos constituye una vulneración fundamental del principio de transparencia.

3.2. Transferencias Internacionales de Datos Ilícitas (Capítulo V del RGPD)

El FAQ de Anthropic confirma que, como empresa global, procesan datos en diferentes países y utilizan mecanismos como las Cláusulas Contractuales Tipo (SCCs) para transferencias fuera del Espacio Económico Europeo (EEE). La narrativa de "autoalojado" de Clawdbot oculta la obligación fundamental del usuario (el Responsable) de asegurarse de que existe un mecanismo de transferencia válido y eficaz para los datos personales enviados a Estados Unidos o donde sea que se encuentren los servidores de Anthropic. El usuario medio de Clawdbot, atraído por la promesa de privacidad, carece del conocimiento y los recursos para auditar la validez de estas transferencias, lo que constituye un incumplimiento directo de las obligaciones estipuladas en el Capítulo V del RGPD, ya que el Responsable carece de la capacidad para validar y garantizar la eficacia de los mecanismos de transferencia invocados por el proveedor del LLM.

3.3. Carencia de Base Jurídica Legítima y Limitación de la Finalidad (Art. 5.1.b y Art. 6)

El FAQ de Anthropic revela una práctica operativa crítica: "los empleados designados de nuestro equipo de Confianza y Seguridad pueden acceder a [los] datos de la conversación" para hacer cumplir su Política de Uso. Este propósito secundario —la revisión de seguridad por parte de un tercero— es distinto y ajeno a la finalidad original para la que el usuario final proporcionó sus datos al Responsable que utiliza Clawdbot (ej. solicitar información sobre un producto). Esto constituye una vulneración del principio de limitación de la finalidad (Artículo 5(1)(b) del RGPD), que establece que los datos deben ser recogidos con fines determinados, explícitos y legítimos. El Responsable del Tratamiento (el usuario de Clawdbot) carece de una base jurídica válida bajo el Artículo 6 del RGPD para autorizar este tratamiento secundario, ya que es improbable que haya obtenido el consentimiento explícito del interesado para esta finalidad específica.

3.4. Obstaculización del Ejercicio de los Derechos de los Interesados (Art. 12-22)

Consideremos un escenario práctico: un interesado desea ejercer su derecho de acceso o supresión sobre los datos de una conversación mantenida a través de un agente Clawdbot. ¿A quién debe dirigirse: al individuo que gestiona la instancia de Clawdbot o directamente a Anthropic? ¿Cómo puede el Responsable (usuario de Clawdbot) garantizar la supresión completa si los datos residen tanto en su "memoria" local como en los sistemas de Anthropic? Esta arquitectura distribuida crea barreras significativas y una ambigüedad operativa que obstaculiza el ejercicio efectivo de los derechos de los interesados. Esta situación constituye una vulneración directa del Artículo 12 del RGPD, que obliga al Responsable a facilitar el ejercicio de dichos derechos de forma clara y accesible.

3.5. Riesgos para la Seguridad del Tratamiento (Art. 32): El Agente como Vector de Amenaza

La capacidad de Clawdbot de tener "acceso al sistema" es una funcionalidad potente pero peligrosa. En el marco arquitectónico de la IA agentiva del CEPD, esta característica convierte al agente en un "módulo de Acción" que interactúa directamente con el "Entorno", que en este caso es el sistema de ficheros local del usuario. Un agente de IA conectado a servicios externos en la nube que simultáneamente puede actuar sobre el sistema anfitrión representa una superficie de ataque significativamente ampliada. Una vulnerabilidad en Clawdbot podría ser explotada no solo para acceder a los datos de las conversaciones, sino también para comprometer todo el sistema, actuando como un puente entre la nube y los datos locales sensibles. Esta configuración introduce un grave riesgo durante las fases de "Despliegue" y "Operación y Monitorización" del ciclo de vida de la IA, contraviniendo la obligación del Responsable de garantizar un nivel de seguridad adecuado al riesgo, según lo estipulado en el Artículo 32 del RGPD.

---

4. El Ecosistema Normativo Ampliado: Implicaciones de la Ley de IA y los Derechos de Autor

Los riesgos del RGPD asociados a Clawdbot no existen en un vacío legal. Se enmarcan en un panorama regulatorio europeo más amplio que afecta a los componentes subyacentes de los que depende. Para el usuario de Clawdbot, esto se traduce en un crítico **riesgo en la cadena de suministro**: como Responsable del Tratamiento y potencial Deployer (Desplegador) de un sistema de IA, hereda el riesgo de cumplimiento del Provider (Proveedor) del LLM sin ningún medio práctico de verificación.

El *Reglamento (UE) 2024/1689* (Ley de IA) impone obligaciones de transparencia a los proveedores de modelos de IA de propósito general. Específicamente, el artículo 53.1.d exige a estos proveedores "elaborar un resumen detallado del contenido utilizado para el entrenamiento de sus modelos". Adicionalmente, la *Directiva (UE) 2019/790* sobre derechos de autor establece un marco para la minería de textos y

datos (TDM), incluyendo un derecho de exclusión (*opt-out*) que permite a los titulares de derechos reservar sus obras para el entrenamiento de IA.

Aunque estas obligaciones recaen formalmente sobre desarrolladores como Anthropic, el usuario de Clawdbot opera en un ecosistema opaco, sin visibilidad ni control sobre si el modelo subyacente fue entrenado con datos obtenidos legalmente o si se respetaron los derechos de autor. Esta dependencia ciega del cumplimiento de un tercero añade una capa significativa de vulnerabilidad legal a su operación.

---

5. Conclusión: Más Allá de la Ilusión del Control

Este análisis demuestra cómo la promesa de privacidad y soberanía de datos de Clawdbot se desmorona bajo un examen crítico de su arquitectura operativa y del marco legal del RGPD. La afirmación de "control total" es una ilusión que oculta una realidad en la que los datos de los usuarios son externalizados a proveedores de servicios en la nube, con todas las implicaciones legales que ello conlleva.

Herramientas como Clawdbot, al ofuscar el flujo real de datos, trasladan una carga desproporcionada de responsabilidad legal a usuarios que, en su mayoría, no están equipados ni son conscientes de las obligaciones que asumen. Al posicionarse como una solución de privacidad, atraen a usuarios que legítimamente buscan proteger sus datos, solo para colocarlos inadvertidamente en una situación de incumplimiento normativo complejo, que puede incluir corresponsabilidad con gigantes tecnológicos.

La advertencia final es clara: los usuarios y organizaciones que deseen adoptar agentes de IA autoalojados deben realizar una diligencia debida que vaya mucho más allá de las afirmaciones superficiales de marketing. Es imperativo mapear activamente los flujos de datos, comprender en profundidad sus obligaciones como Responsables del Tratamiento y evaluar críticamente las prácticas de protección de datos de cualquier proveedor de LLM de terceros. Solo así podrán evitar los graves riesgos de incumplimiento que se esconden tras la atractiva pero frágil ilusión del control.7