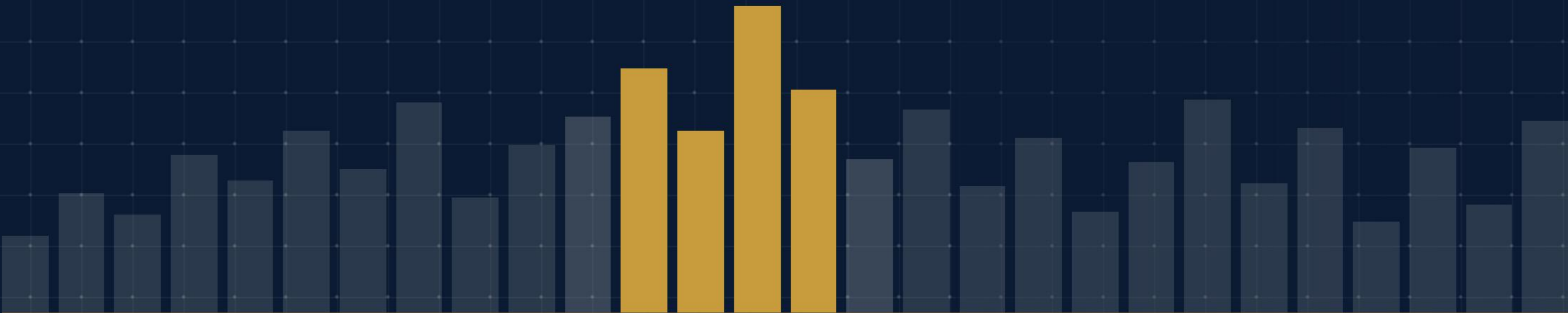


Web scraping en Chile: *aspectos legales*

La regulación que subyace la extracción automatizada de datos y su relevancia en el mercado asegurador

POLIZA	RAMO	EXPOS.	SCORE
POL-48213	RC	1.42	0.61
POL-48219	AUTO	0.87	0.34
POL-48226	VIDA	2.10	0.78
POL-48231	HOGAR	0.55	0.29
POL-48240	RC	1.96	0.72
POL-48247	AUTO	1.03	0.41
POL-48255	SALUD	1.31	0.55



Web scraping en Chile

- 01 ¿Qué es el web scraping y cómo funciona?
- 02 Marco legal chileno
- 03 APIs como alternativa a la extracción masiva
- 04 Foco sectorial: scraping en la industria de seguros



POLIZA	RAM	SCORE
POL-48213	RC	0.61
POL-48226	VIDA	0.78
POL-48240	RC	0.72

SECCIÓN 01

Concepto y herramientas del web scraping

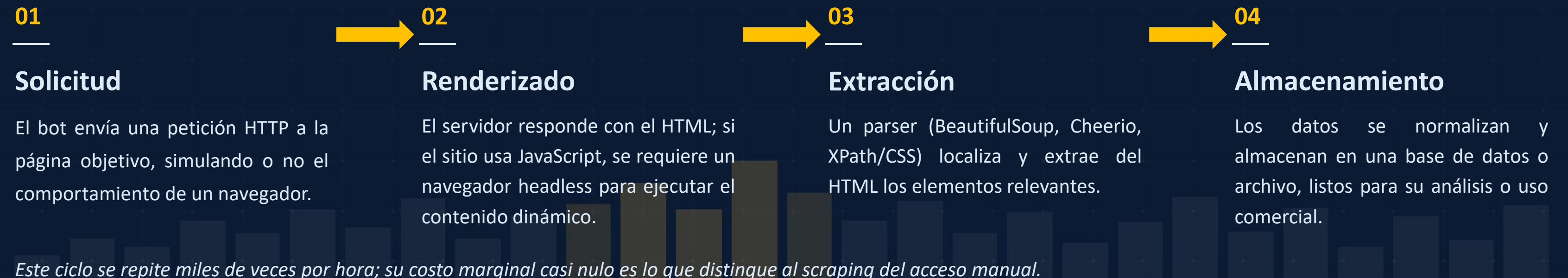


01 ¿Qué es el web scraping?

Es la extracción automatizada de información publicada en sitios web mediante software (“bots” o “crawlers”), que recorre páginas, interpreta su código HTML y convierte contenido no estructurado en datos estructurados.

Se distingue del acceso manual por su escala, velocidad y capacidad de repetición sistemática, características que pueden activar alertas en los sitios objetivo.

02 Cómo funciona: de la solicitud al dato



03 Usos frecuentes

Usos comerciales

- Monitoreo de precios y competencia.
- Comparadores y agregadores (viajes, seguros, retail).
- Inteligencia de mercado y due diligence.

Usos analíticos y de investigación

- Investigación académica y periodismo de datos.
- Entrenamiento y alimentación de modelos de IA.
- Monitoreo de reputación y menciones de marca.

04 Riesgos

- **Ciberataques dirigidos:** publicación de datos personales en foros de hackers son usados para ataques de ingeniería social / phishing.
- **Suplantación de identidad:** datos obtenidos son usados para solicitar productos como préstamos o tarjetas comerciales, o para suplantar al titular por medio de la creación de cuentas falsas en redes sociales o plataformas.
- **Monitoreo, perfilamiento y vigilancia:** datos obtenidos son usados para poblar bases de datos de reconocimiento facial y permitir acceso no autorizado a las autoridades.
- **Obtención de datos no autorizada para actividades políticas o de inteligencia:** datos obtenidos pueden ser usados por gobiernos extranjeros o agencias de inteligencia para fines no autorizados.
- **Marketing indeseado / spam:** datos obtenidos pueden incluir información de contacto que puede ser usada para enviar spam o mensajería comercial.

POLIZA	RAMAL	SCORE
POL-48213	RC	0.61
POL-48226	VIDA	0.78
POL-48240	RC	0.72

SECCIÓN 02

Marco legal chileno



POLIZA	RAMAL	PREMIO
POL-48213	RC	0.61
POL-48226	VIDA	0.78
POL-48240	RC	0.72

05 Principales normas aplicables

1. **PROTECCIÓN DE DATOS PERSONALES:** LEY 19.628, MODIFICADA POR LA LEY 21.719
2. **DELITOS INFORMÁTICOS:** LEY 21.459: ACCESO ILÍCITO Y FIGURAS AFÍNES
3. **CONTRATOS Y PROPIEDAD INTELECTUAL:** LEY 17.336.



06 Ley 19.628/21.719: si se hará scraping de datos personales, respecto de su recolección aplican los siguientes principios reconocidos en la ley:

1. **LICITUD:** Se requiere una fuente de licitud: Art. 12 (consentimiento) o 13 (cumplimiento contractual, obligación legal, interés legítimo. Pero en scraping masivo el consentimiento es inviable (usualmente no hay vínculo directo entre el titular de los datos y quien hace scraping) y de las otras fuentes, la más aplicable sería el interés legítimo, pero ello requiere cumplir con ciertos requisitos: idoneidad, necesidad del tratamiento, prueba de ponderación (EIL). En Chile al menos, no sirve para hacer scraping de datos personales sensibles.
2. **FINALIDAD:** El dato personal debe usarse solo para un propósito determinado e informado, y no para cualquier reutilización futura.
3. **INFORMACIÓN Y TRANSPARENCIA:** Aún en fuentes públicas, subsiste el deber de transparencia frente al titular.
4. **PROPORCIONALIDAD:** Solo se puede recolectar lo estrictamente necesario.

Además, web scraping puede resultar en tratamiento masivo de datos personales a gran escala, lo que configura uno de los casos en que la Ley 19.628 modificada por la Ley 21.719 exige la realización de una evaluación de impacto previo, cuando por la naturaleza, alcance, contexto, tecnología utilizada o fines del tratamiento, se pueda producir un alto riesgo para los derechos de las personas titulares de los datos personales (Art. 15 ter).



07 Ley 19.628/21.719: Vulneración de Principios

SOLOVE / HARTZOG: “The Great Scrape: The Clash Between Scraping and Privacy”, *California Law Review*, 2025.

El scraping masivo es un rastrillo que recoge, junto con otros datos, datos personales a gran escala. Esta práctica vulneraría casi todos los principios de protección de datos: transparencia, proporcionalidad (minimización), finalidad, licitud. La disponibilidad pública de datos personales no elimina su protección.

En opinión de los autores, el scraping masivo debería ser considerado *vigilancia* (surveillance) y su prevención debiera ser un deber vinculado al principio de seguridad de datos personales, condicionando de esta forma la licitud del scraping masivo a un interés público justificado (interés legítimo?), con riesgo razonable, beneficio proporcional y debido proceso.



POLIZA	RAMAL	SCORE
POL-48213	RC	0.61
POL-48226	VIDA	0.78
POL-48240	RC	0.72

“

Información personal que tiene acceso público está igualmente sujeta a las leyes de protección de datos y de privacidad en la mayoría de las jurisdicciones

Joint Statement on Data
Scraping and the
Protection of Privacy,
agosto 2023



08 Ley 21.459: el acceso ilícito a sistemas informáticos

Acceso ilícito: Ley 21.459 Art. 2°: *El que, sin autorización o excediendo la autorización que posea y superando barreras técnicas o medidas tecnológicas de seguridad, acceda a un sistema informático, será castigado con la pena de presidio menor en su grado mínimo o multa de 11 a 20 UTM.*

El scraping cruza esta línea penal cuando se eluden CAPTCHAs o bloqueos de IP, se usan credenciales falsas o robadas, o se persiste en el acceso automatizado tras una notificación de cese formal del titular del sitio.

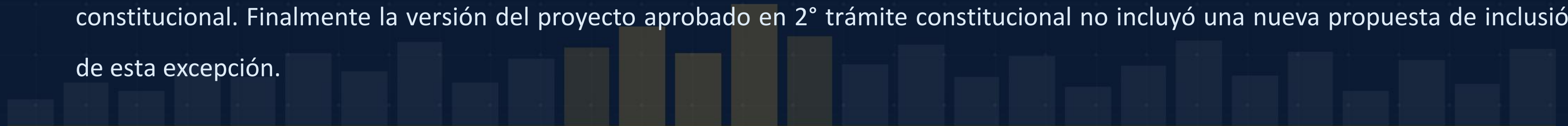


09 Ley 17.336 de Propiedad Intelectual

En la medida que los datos objeto del scraping sean información sujeta a derechos de autor, para su reproducción y reutilización se requiere autorización del titular respectivo. Datos de acceso público en Internet no implica ni presume autorización del titular para utilizarlos, sino solo un problema de *enforcement*.

La Ley 17.336 de Propiedad Intelectual reconoce a las bases de datos como una obra intelectual: Art. 3°: *Quedan especialmente protegidos con arreglo a la presente ley: 17) Las compilaciones de datos o de otros materiales, en forma legible por máquina o en otra forma, que por razones de la selección o disposición de sus contenidos, constituyan creaciones de carácter intelectual. Esta protección no abarca los datos o materiales en sí mismos, y se entiende sin perjuicio de cualquier derecho de autor que subsista respecto de los datos o materiales contenidos en la compilación.*

Minería de texto y datos: a la fecha, ley chilena no reconoce una excepción para efectuar minería de texto y datos. Se planteó inicialmente en el PDL de regulación de Inteligencia Artificial (adición de nuevo art. 71 T bis a la Ley 17.336) y en ese proyecto fue rechazada. Luego Gobierno actual la reintroduce en PDL de reconstrucción, y es nuevamente rechazada en su primer trámite constitucional. Finalmente la versión del proyecto aprobado en 2° trámite constitucional no incluyó una nueva propuesta de inclusión de esta excepción.



10 Normas vinculadas

Normas contractuales: términos y condiciones de sitios web establecen como condición para la utilización del sitio web que la información contenida en este no sea recolectada de forma automatizada. El incumplimiento gatilla responsabilidad contractual y, eventualmente podría constituir infracción a la Ley 19.496.

Ley 20.169 de Competencia Desleal: un scraping masivo y sistemático de la información de un competidor podría configurar una causal de competencia desleal conforme a la definición del acto en su artículo 3°, especialmente si el competidor ha instalado barreras técnicas para impedir scraping, que son eludidas (como ignorar robots.txt o usar cuentas falsas): *es acto de competencia desleal toda conducta contraria a la buena fe o las buenas costumbres que, por medios ilegítimos, persiga desviar clientela de un agente de mercado.*

DL 211 de Libre Competencia: por ejemplo, si un actor con posición dominante scrapea sistemáticamente a competidores para apropiarse de su información y debilitar su capacidad competitiva -o para alimentar un producto propio que termina desplazándolos-, podría configurarse hipótesis de *abuso de posición dominante* bajo el artículo 3° inc. 2° del DL 211, por ejemplo en la conducta descrita en la letra c): "las prácticas predatorias, o de competencia desleal, realizadas con el objeto de alcanzar, mantener o incrementar una posición dominante".

11 Algunos casos:

hiQ Labs v. LinkedIn (2017–2022): hiQ extraía perfiles públicos de LinkedIn para vender análisis de rotación de talento. LinkedIn envió una carta C&D y bloqueó el acceso. Si bien inicialmente los tribunales le permitieron a hiQ seguir haciendo scraping de perfiles públicos, finalmente y luego de que el caso fuera conocido por la Corte Suprema de Estados Unidos, se determinó de que hiQ había infringido los términos de uso de LinkedIn, y ello motivó a las partes a transigir el caso.

Clearview AI: esta empresa hizo scraping de millones de fotografías de personas para su base de datos de reconocimiento facial, sin obtener consentimiento de los titulares de dichos datos. La ICO de UK la multó con 7,5 millones de euros por infringir el RGPD.

La lección para Chile: que el dato sea público no blind a quien hace scraping. El uso de cuentas falsas y la persistencia tras un aviso de cese y desistimiento pueden generar responsabilidad civil y exposición penal bajo la Ley 21.459. La responsabilidad puede verse agravada por inclusión de la Ley 21.459 en la Ley 21.595 de Delitos Económicos (2ª categoría, artículo 2° N° 20).



12 ¿Quién responde por el scraping?

N°	Actor	Tipo de responsabilidad	Base legal
01	Persona natural / equipo técnico	Responsabilidad penal	Ley 21.459, Art. 2°
02	Responsable del tratamiento	Infraccional y civil	Ley 19.628 arts. 33, 34 y ss., 47
03	Proveedor externo de scraping	No exime al mandante	Contrato de encargo, Art. 15 bis
04	Quien realiza scraping	Infraccional, civil y penal	Ley 17.336 arts. 78, 79 a)

Entonces, las preguntas a realizarse de forma previa a hacer scraping:

- Se hará scraping de datos personales?
- Se hará scraping de contenido protegido por derechos de autor u otras normas de propiedad intelectual?
- Se busca hacer scraping de datos en sitios con logins, firewalls o acceso técnica o contractualmente restringido?

POLIZA	RASGO	SCORE
POL-48213	RC	0.61
POL-48226	VIDA	0.78
POL-48240	RC	0.72

SECCIÓN 03

Alternativas

Por qué las APIs oficiales son la vía de menor riesgo.



13 APIs

API (Application Programming Interface o Interfaz de Programación de Aplicaciones) es un mecanismo que permite a dos componentes de software comunicarse entre sí mediante un conjunto de definiciones y protocolos. De la definición, por *aplicación* se entiende cualquier software con una función distinta. La *interfaz* se puede entender como un contrato de servicio entre dos aplicaciones, que define cómo se comunican mediante solicitudes y respuestas.

NCG 514: Mecanismo tecnológico por medio del cual dos o más programas o sistemas computacionales pueden comunicarse entre sí. Parte fundamental de dicha comunicación depende de las características de la documentación técnica que describe los métodos de conexión disponibles, y los elementos y atributos del intercambio de información que se detallan en la Especificación de la API.

TIPOS:

- APIs privadas: internas y se utilizan para conectar sistemas y datos dentro de una organización.
- APIs públicas: abiertas al público y utilizables libremente.
- APIs de partners: para uso por desarrolladores externos autorizados para ayudar a asociaciones entre organizaciones.
- APIs compuestas: combinan dos o más APIs para abordar requisitos o comportamientos complejos de un sistema.

14 Datos abiertos: el Estado privilegia APIs

En la tramitación de iniciativas vinculadas a la Ley 21.180 se ha discutido si la interoperabilidad estatal debiera habilitar el scraping de datos públicos; el enfoque de Gobierno Digital privilegia en su lugar las APIs oficiales y el principio de “una sola vez” (la Administración del Estado no puede exigirle a una persona o empresa un dato o documento que otro organismo público ya tiene en su poder)

Consecuencia práctica: para datos de organismos públicos, la vía institucionalmente correcta y de menor riesgo es siempre la API o el conjunto de datos abiertos oficial, no la extracción directa del sitio.

15 Por qué usar APIs oficiales

- Legalidad: el acceso se enmarca en los términos que el propio proveedor autoriza.
- Estabilidad: los cambios de estructura del sitio no rompen la integración. APIs utilizan control de versiones.
- Calidad y trazabilidad: datos estructurados, con metadatos y control de origen.
- Relación comercial: permite negociar SLAs, límites de uso, cláusulas de protección de datos y de entrenamiento de IA.

16 APIs públicas en Chile

- Biblioteca del Congreso
- Mercado Público.
- Portal ChileAtiende
- Plataforma Ley de Lobby
- Comisión Nacional de Energía
- Compras Transparentes
- ChileCompra
- Correos de Chile
- Superintendencia de Salud
- Turnos de Farmacia
- REST del Congreso
- Indicadores económicos diarios
- Climatología Dirección Meteorológica

Fuente: <https://juanbrujo.github.io/listado-apis-publicas-en-chile/>

■ LEY FINTECH: NCG 514 (modificada por NCG 569 de 01.06.26):

Sección II: Funcionamiento del Sistema; A. Medios de Intercambio de Información: el mecanismo principal es el uso de APIs para el SFA. En el contexto del SFA, las entidades fiscalizadas no podrán utilizar mecanismos diferentes a las API para atender solicitudes de acceso a datos, sin perjuicio de lo que se disponga respecto del uso de un mecanismo alternativo regulado por la CMF. El desarrollo y mantención de las APIs será de exclusiva responsabilidad de los Participantes del SFA.

Tabla N°1: Estándares de las APIs

ELEMENTO	ESTÁNDAR
Especificación y diseño	OpenAPI (versión 3.1)
Mensajería	JSON
Arquitectura	Marco de referencia REST y su implementación debe ser RESTful
Estándares de Administración de Datos y Diccionario de Datos	ISO 20.022 última versión
Autorización y Autenticación	OAuth 2.0 y OpenID Connect
Perfiles de seguridad	FAPI 2.0 conforme perfiles que se indiquen en Anexo N°3
Protocolo de intercambio	mTLS

POLIZA	RAMAL	SCORE
POL-48213	RC	0.61
POL-48226	VIDA	0.78
POL-48240	RC	0.72

SECCIÓN 04

Foco sectorial y cierre



17 Web scraping en la industria de seguros

Usos habituales

- Comparadores e insurtech alimentan cotizadores con precios de precios de múltiples aseguradoras.
- Tarificación dinámica y suscripción enriquecidas con datos externos.
- Modelamiento predictivo.
- Detección de fraude mediante cruce de información pública en pública en siniestros.

Riesgos específicos

- Datos de salud activan la protección reforzada de la Ley 19.628/21.719.
- Discriminación en tarificación al usar datos scrapeados para ajustar primas, puede infringir normas de derechos del consumidor. Énfasis en consumidores hipervulnerables.
- NCG 454 CMF aplicable a entidades aseguradoras y reaseguradoras, exige evaluar el uso de terceros (evaluación de riesgo y control, control de accesos, testeo, intercambio de información).

18 Conclusiones clave

El web scraping en Chile no está prohibido. Pero su legalidad debe evaluarse caso a caso, pues queda sujeto a normativa de datos personales, delitos informáticos y propiedad intelectual. El acceso público a un dato, sea o no un dato personal, no elimina su protección, y el riesgo penal bajo la Ley 21.459 se activa por la superación de barreras técnicas, no por scraping en abstracto.

La responsabilidad de la protección de datos personales recae en quien encarga y usa los datos, aún cuando el scraping lo ejecute un tercero (relevancia de relación entre responsable y encargado). Frente a fuentes estatales o institucionales, una API oficial es siempre la vía de menor riesgo legal y operativo.



POLIZA	RAMAL	PREMIO
POL-48213	RC	0.61
POL-48226	VIDA	0.78
POL-48240	RC	0.72

Gracias

Quedamos a disposición para profundizar en cualquiera de los puntos tratados.

