

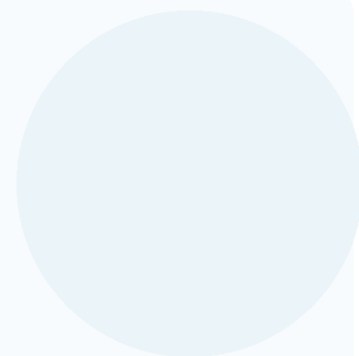
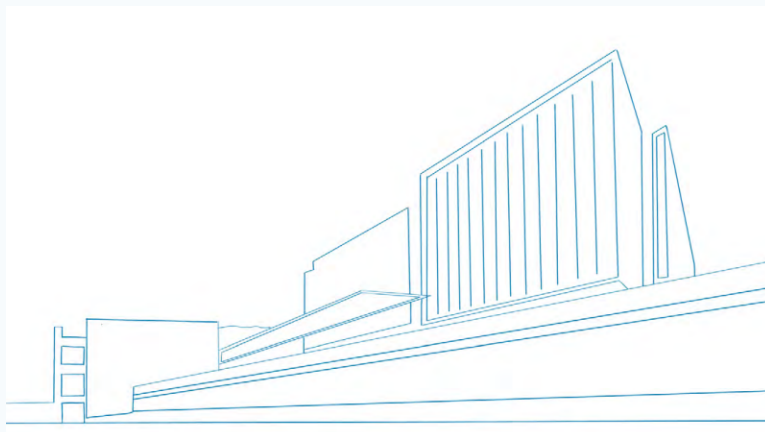
GUÍA PARA INVESTIGADORES UDD

ORIENTACIONES PARA

DATOS SECUNDARIOS Y REDES SOCIALES

Orientaciones éticas y metodológicas
para equipos de investigación

Privacidad y uso responsable



CONTENIDOS

Orientaciones éticas y metodológicas para equipos de investigación

DATOS SECUNDARIOS Y REDES SOCIALES

1. Introducción
2. ¿Qué se Entiende por Datos Secundarios y por qué RRSS Suele caer en esa Categoría?
3. Riesgos Principales: Privacidad y Confidencialidad
4. Dato Personal, Dato Sensible, Anonimización y Seudonimización
5. Principios de Tratamiento Aplicables al Diseño de Investigación (Ley 21.719)
6. Diseño Metodológico: Decisiones que Reducen Riesgo sin Sacrificar Validez
7. Fuente de Acceso Público no Equivale a 'Libre de Obligaciones'
8. Consentimiento: Cuándo es Impracticable y qué Implica en Datos Secundarios
9. Seguridad Digital
10. Ciclo de Vida del Dato en Estudios con RRSS y Registros
11. Evaluación de Impacto en Protección de Datos Personales (EIDP)
12. Encargados, Terceros y Plataformas
13. Reporte de Vulneraciones de Seguridad (Brechas)
14. Referencias

Capítulo I

Fundamentos y contexto

1. INTRODUCCIÓN

La investigación basada en registros y datos secundarios permite responder preguntas relevantes sin interacción directa con participantes. Hoy, una proporción creciente de esos “registros” proviene de entornos digitales: redes sociales, foros, plataformas y trazas conductuales generadas en línea. En muchos proyectos, las redes sociales funcionan como una fuente de datos secundarios (contenido ya producido con otros fines), pero su uso implica desafíos específicos: mayor riesgo de reidentificación, persistencia y replicabilidad del contenido, huellas técnicas (metadatos, capturas, URL), y vulnerabilidades de seguridad (phishing, hackeo, malware) que pueden comprometer datos de investigación. Esta guía integra criterios éticos y metodológicos para tratar datos de redes sociales como datos secundarios, y los articula con el marco chileno de protección de datos personales establecido por la Ley N° 21.719.



2. ¿QUÉ SE ENTIENDE POR DATOS SECUNDARIOS Y POR QUÉ RRSS SUELE CAER EN ESA CATEGORÍA?

Los datos secundarios incluyen información documentada previamente en registros en papel o electrónicos (por ejemplo, historiales clínicos, registros escolares, justicia penal, tráfico) que se analizan para responder hipótesis sin interactuar con personas. En términos funcionales, una parte importante de la información disponible en redes sociales (publicaciones, comentarios, imágenes, interacciones y metadatos) se produce para fines de comunicación social, entretenimiento o marketing, no para investigación. Cuando un equipo recolecta y analiza ese material sin intervenir en la plataforma (o con intervención mínima), la lógica es similar a la investigación basada en registros: el riesgo deriva principalmente de invasión de privacidad y violaciones de confidencialidad.

3. RIESGOS PRINCIPALES: PRIVACIDAD Y CONFIDENCIALIDAD

RECOMENDACIÓN CEII

No asumir que lo público elimina el riesgo. Evaluar reidentificación, búsqueda inversa y daño contextual.

En investigación basada en registros, los riesgos centrales provienen de la posible invasión de la privacidad y de la vulneración de la confidencialidad. La divulgación de información puede razonablemente exponer a las personas a responsabilidad penal o civil, o dañar su empleabilidad, situación financiera o reputación. En

RRSS, estos riesgos se intensifican por la facilidad de búsqueda, replicación y contextualización inversa: aun si un dato parece ‘público’, su recolección sistemática, combinación con otras fuentes o análisis puede aumentar identificabilidad y producir nuevos daños.

Capítulo II

Diseño y criterios éticos

4. DATO PERSONAL, DATO SENSIBLE, ANONIMIZACIÓN Y SEUDONIMIZACIÓN

La Ley 21.719 define dato personal como cualquier información referida a una persona natural identificada o identificable, incluyendo identificabilidad indirecta por uno o más identificadores. Define datos sensibles e incluye expresamente información sobre situación socioeconómica, salud, vida sexual, orientación sexual e identidad de género, entre otras. En RRSS es frecuente que el contenido revele (directa o indirectamente) categorías sensibles.

La ley distingue anonimización (procedimiento irreversible tras el cual el dato deja de ser dato personal) y seudonimización (tratamiento que impide atribuir a un titular sin información adicional separada y protegida). Esta distinción es crucial: muchos proyectos describen datos ‘anónimos’ cuando en realidad son seudonimizados (por ejemplo, se reemplaza el nombre por un código pero se conserva URL, captura de pantalla o texto exacto que permite búsqueda).



5. PRINCIPIOS DE TRATAMIENTO APLICABLES AL DISEÑO DE INVESTIGACIÓN (LEY 21.719)

La Ley 21.719 establece principios relevantes para investigación: licitud y lealtad, finalidad, proporcionalidad (minimización), calidad, responsabilidad, seguridad, transparencia/información y confidencialidad. En proyectos con datos secundarios y RRSS, estos principios se operacionalizan como: definir fines explícitos; recolectar solo lo necesario; preferir extracción agregada o desidentificada; limitar plazos de conservación; y adoptar medidas técnicas y organizativas acordes al riesgo, incluyendo controles de acceso.

6. DISEÑO METODOLÓGICO: DECISIONES QUE REDUCEN RIESGO SIN SACRIFICAR VALIDEZ

RECOMENDACIÓN CEII

Recolectar solo variables necesarias; evitar texto completo, imágenes o URL si no son imprescindibles.

La metodología es una salvaguarda. Un protocolo robusto debe justificar: (a) por qué se requieren RRSS o registros; (b) qué población o universo cubre la fuente; (c) criterios de selección de cuentas, hashtags, grupos o periodos; (d) qué variables se extraen; (e) si habrá linkage con otras bases; y (f) cómo se evitará que la extracción exceda lo necesario. Cuando sea viable, se recomienda extraer variables agregadas o derivadas, evitando capturar contenido textual completo o imágenes identificables si la pregunta puede responderse sin ello.

Si la investigación no puede realizarse de forma anónima, el equipo debe responder preguntas operativas: qué información identificable se recolectará; quién accederá; dónde se guardará; qué códigos/cifrado separarán datos de identificadores; cómo se limitará acceso; cómo se entrenará al personal; cuánto tiempo se mantendrán identificadores o vínculos; qué cifrado se usará en transmisión; y cómo se destruirán identificadores cuando ya no se requieran. En RRSS, estas preguntas deben incluir la gestión de URL, capturas de pantalla, metadatos y exportaciones de plataforma.

Capítulo III

Participación y salvaguardas

7. FUENTE DE ACCESO PÚBLICO NO EQUIVALE A 'LIBRE DE OBLIGACIONES'

La Ley 21.719 incluye una definición de fuentes de acceso público y establece que el tratamiento de datos personales provenientes de dichas fuentes se somete a las disposiciones de la ley. En investigación con RRSS esto es decisivo: que un contenido sea accesible públicamente no elimina el deber de finalidad, proporcionalidad, seguridad y confidencialidad. Además, el deber de secreto/confidencialidad subsiste incluso cuando el dato proviene de una fuente de acceso público, si el responsable realiza acciones como organizar, clasificar o combinar datos bajo criterios, generando un nuevo conjunto protegido.



8. CONSENTIMIENTO: CUÁNDO ES IMPRACTICABLE Y QUÉ IMPLICA EN DATOS SECUNDARIOS

RECOMENDACIÓN CEII

Justificar ante el CEII la imposibilidad de consentimiento y describir salvaguardas equivalentes.

En investigación basada en registros, obtener consentimiento puede ser impracticable, comprometer el objetivo o incluso aumentar riesgo para el sujeto; por ello, en ciertos marcos regulatorios se contemplan

exenciones bajo criterios estrictos de riesgo mínimo y necesidad. En el contexto chileno e institucional, la regla práctica es justificar ante el comité por qué no es razonable o posible obtener consentimiento, describiendo salvaguardas equivalentes. En RRSS, esto exige particular cuidado: aunque contactar a miles de usuarios sea impracticable, la recolección debe minimizar identificabilidad y describir claramente la justificación ética de no contacto.

9. SEGURIDAD DIGITAL

RECOMENDACIÓN CEII

Usar cuentas institucionales, autenticación robusta y almacenamiento cifrado para exportaciones.

Los riesgos de RRSS no son solo éticos, sino también de seguridad de la información: malware, phishing, hackeo de cuentas y divulgación indebida pueden comprometer datos de investigación y credenciales institucionales. Se recomiendan prácticas estándar: contraseñas robustas, mantener software actualizado, no abrir enlaces sospechosos, revisar configuraciones de privacidad, evitar divulgar información innecesaria y asumir que lo publicado puede ser compartido y difícil de revertir. En investigación, esto se traduce en: usar cuentas institucionales o específicas del estudio, aplicar autenticación robusta, evitar usar dispositivos personales sin cifrado para almacenar exportaciones, y capacitar al equipo en prevención de phishing e ingeniería social.

Capítulo IV

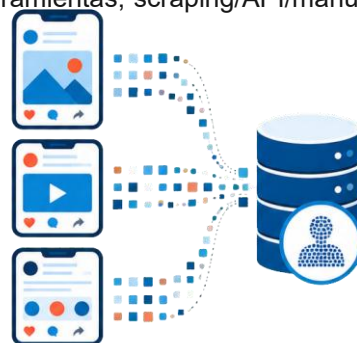
Datos, gestión y comunicación

10. CICLO DE VIDA DEL DATO EN ESTUDIOS CON RRSS Y REGISTROS

RECOMENDACIÓN CEII

Describir desde la recolección hasta la destrucción o anonimización irreversible de los datos.

Un protocolo debe describir el ciclo de vida del dato: (i) recolección (qué herramientas, scraping/API/manual; qué se captura); (ii) almacenamiento (servidor institucional, cifrado, controles de acceso); (iii) procesamiento (seudonimización, eliminación de identificadores); (iv) análisis (entornos protegidos; minimización de copias); (v) reporte (agregado; evitar citas textuales o imágenes que permitan búsqueda inversa, salvo autorización o justificación robusta); y (vi) conservación y destrucción/anonimización irreversible. La Ley 21.719 exige proporcionalidad y limita conservación al tiempo necesario, luego suprimir o anonimizar.



11. EVALUACIÓN DE IMPACTO EN PROTECCIÓN DE DATOS PERSONALES (EIPD)

La Evaluación de Impacto en Protección de Datos Personales (EIPD; DPIA, por sus siglas en inglés, Data Protection Impact Assessment) es una evaluación preventiva destinada a identificar y gestionar los riesgos que un tratamiento de datos personales puede generar para los derechos de sus titulares. De acuerdo con la Ley N.º 21.719, deberá realizarse antes de iniciar el tratamiento cuando, por su naturaleza, alcance, contexto, tecnología utilizada o finalidad, sea probable que este pueda producir un alto riesgo para los derechos de las personas. La ley establece, además, que la EIPD será requerida siempre en los siguientes casos:

- Evaluación sistemática y exhaustiva de aspectos personales, basada en tratamientos o decisiones automatizadas que produzcan efectos jurídicos significativos en las personas.
- Tratamiento masivo de datos o a gran escala.
- Observación o monitoreo sistemático de una zona de acceso público.
- Tratamiento de datos sensibles o especialmente protegidos, cuando se realice bajo alguna de las hipótesis que permiten prescindir del consentimiento.

En investigaciones que utilizan datos secundarios o información proveniente de redes sociales (RRSS), deberá prestarse especial atención a esta exigencia. Por ejemplo, el scraping o recopilación automatizada de grandes volúmenes de información, la integración de múltiples bases de datos, el análisis sistemático del comportamiento de usuarios, la elaboración de perfiles o predicciones, y el tratamiento a gran escala de información sensible pueden hacer necesaria una EIPD, ya sea porque se configura alguno de los supuestos expresamente establecidos por la ley o porque el tratamiento, considerado en su conjunto, presenta un alto riesgo para las personas.

La EIPD deberá considerar, al menos:

1. Descripción del tratamiento: qué datos serán utilizados, su origen, las operaciones que se realizarán, las tecnologías empleadas y quiénes tendrán acceso.
2. Finalidad: para qué se utilizarán los datos y qué objetivos persigue el tratamiento.
3. Necesidad y proporcionalidad: si los datos y operaciones previstas son efectivamente necesarios para alcanzar los objetivos de la investigación y si existen alternativas menos invasivas.
4. Evaluación de riesgos: posibles consecuencias para la privacidad, confidencialidad y demás derechos de las personas, incluyendo riesgos de identificación o reidentificación.
5. Medidas de mitigación: mecanismos técnicos, organizativos y metodológicos destinados a reducir esos riesgos, tales como minimización de datos, seudonimización o anonimización, restricciones de acceso, medidas de seguridad y reducción de los plazos de conservación.

12. ENCARGADOS, TERCEROS Y PLATAFORMAS

Cuando el tratamiento se realiza mediante terceros (por ejemplo, servicios de transcripción, almacenamiento en nube, consultores de analítica), la Ley 21.719 regula el encargo: debe existir un contrato que fije objeto, duración, finalidad, tipo de datos, categorías de titulares y obligaciones; el encargado no puede usar datos para fines distintos ni cederlos sin autorización; y al terminar, los datos deben suprimirse o devolverse. En investigación con RRSS, el 'tercero' puede incluir herramientas o plugins: se recomienda privilegiar herramientas institucionales y documentar restricciones y auditorías básicas.

Capítulo V

Consideraciones especiales

13. REPORTE DE VULNERACIONES DE SEGURIDAD (BRECHAS)

RECOMENDACIÓN CEII

Definir un procedimiento interno de respuesta ante pérdida, filtración o acceso no autorizado.

La Ley 21.719 introduce el deber de reportar vulneraciones a la Agencia cuando exista riesgo razonable para derechos y libertades de titulares y, en ciertos casos (datos sensibles, datos de niños/as menores de 14 años), también notificar a titulares o sus representantes en lenguaje claro. En investigación, esto implica que los equipos deben tener un procedimiento de respuesta a incidentes (qué hacer ante pérdida de dispositivo, hackeo, filtración de archivos) y un canal institucional para escalar el evento.



14. REFERENCIAS

CITI Program. Investigación basada en registros (datos secundarios).

CITI Program. Redes sociales más seguras.

Ley N° 21.719. Regula la protección y el tratamiento de los datos personales y crea la Agencia de Protección de Datos Personales (vigencia diferida desde 01-12-2026).