

**Experiencia  
y conocimiento**

# ANONIMIZACIÓN

Carlos Gregorio de Gràcia

**SITA  
2018**



# Anonimización

información pública

consentimiento al Estado  
acceso a la información  
datos abiertos

VIDA PRIVADA  
&  
RIESGOS

permanecer anónimo

Ley Europea de *cookies*

expresarse anónimamente

redes sociales  
responsabilidad de los  
medios

Libertad de  
Información



# Anonimización — marco legal



[Ley 1712 de 6 de marzo de 2014](#)

artículo 6 inc. j la definición de **datos abiertos**;

Artículo 11 (que trata sobre la “información mínima obligatoria”) inc. k “Los sujetos obligados **deberán publicar datos abiertos**”, (y a continuación somete este deber a todo el capítulo de excepciones, que incluye datos personales e intereses públicos).

Artículo 21. **En aquellas circunstancias en que la totalidad de la información contenida en un documento no esté protegida por una excepción contenida en la presente ley, debe hacerse una versión pública que mantenga la reserva únicamente de la parte indispensable**

Ley Francesa “Republica Digital” [2016] evaluación de riesgos

Colombia – Ley de Acceso – versiones públicas

Argentina – Ley de Acceso – versiones publicas

# Anonimización — marco legal

[Ley 2016-1321 de 7 de octubre 2016 para una República digital.](#)

Artículo 3. “**Toda publicación en formato electrónico se hará en estándar abierto**, reutilizable y fácilmente explotable por un sistema de procesamiento automatizado.”

Artículo 20 y 21. “Esta disponibilidad pública está precedida por un **análisis del riesgo de re-identificación**”

Artículo 60. “ .... certificar, aprobar y publicar las normas o metodologías generales para la **certificación de los procesos de anonimización de de datos personales**, incluyendo la reutilización de la información pública ... “

# Anonimización — marco legal



[Ley 27.275 de Acceso a la Información Pública \(2016\)](#)

Artículo 1. .... **Disociación**: en aquel caso en el que parte de la información se encuadre dentro de las excepciones taxativamente establecidas por esta ley, la información no exceptuada debe ser publicada en una **versión** del documento **que tache, oculte o disocie aquellas partes sujetas a la excepción**.

Artículo 5. “ ... El Estado tiene la **obligación de entregarla en formatos digitales abiertos**, salvo casos excepcionales en que fuera de imposible cumplimiento o significara un esfuerzo estatal desmedido.”

Artículo 8. [Excepciones] “ .... i) “Información que contenga datos personales y **no pueda brindarse aplicando procedimientos de disociación**”

## Anonimización — marco legal

**Version pública:** es prácticamente la única alternativa para facilitar el acceso y proteger los datos personales;

**Datos Abiertos.** Son todos aquellos datos **primarios** o sin procesar, que se encuentran en **formatos estándar e interoperables que facilitan su acceso y reutilización**, los cuales están bajo la custodia de las entidades públicas o privadas que cumplen con funciones públicas y que son puestos a disposición de cualquier ciudadano, de **forma libre y sin restricciones, con el fin de que terceros puedan reutilizarlos y crear servicios derivados de los mismos.**

**Datos abiertos**, obligatorios (con excepciones), buena práctica

... precedida por un **análisis del riesgo de re-identificación**”

**OPEN  
DATA  
IRELAND**



4 DE MARZO

**Día de los Datos Abiertos**



**ABRIENDO DATOS**

Costa Rica



**datos  
abiertos**



Open Data Day  
Quindío, Colombia

Datos  
abiertos



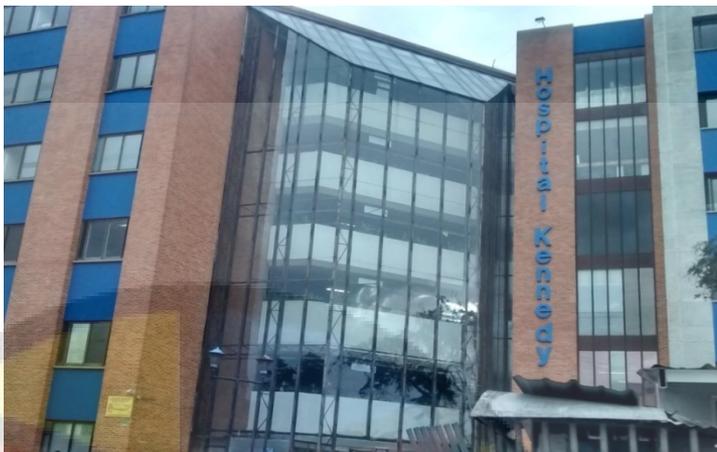
**MISIONES  
PROVINCIA**



**DATOS ABIERTOS**

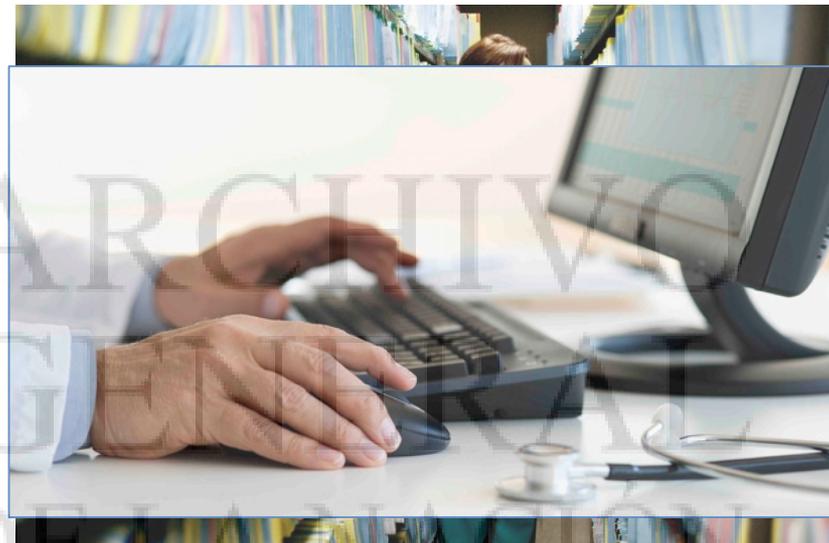
**Vendata**

Datos abiertos de Venezuela a tu alcance

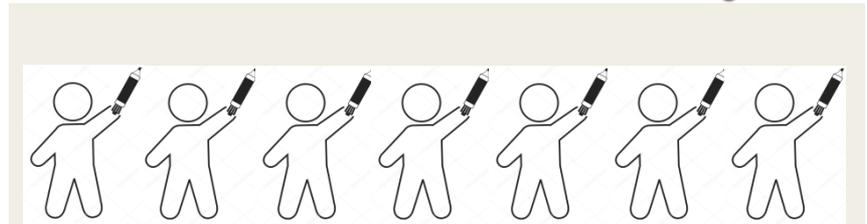


# HISTORIAS CLÍNICAS

## Historias Clínicas



# Problemas con la Anonimización

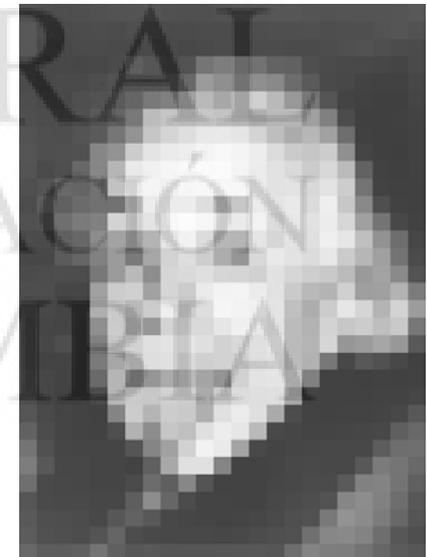
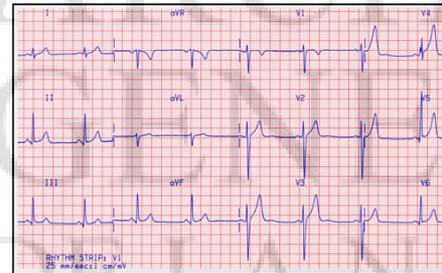
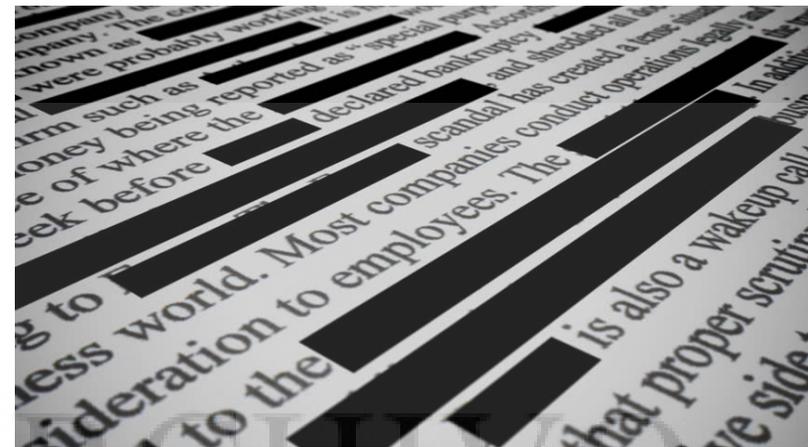


# Anonimización Contexto

**UN documento,**  
historia clínica, sentencia  
gastos en tarjeta de crédito,  
llamadas telefónicas

**UNA base de datos,**  
delitos sexuales, autores de obras

**Imágenes**  
electrocardiograma



[Mortalidad: Cancer de Cervix por regimen de salud](#); 332 registros

[Mortalidad accidentes de transito general](#)

[Motivos de consulta en la ESE Hospital San Rafael de la población del Municipio de Angostura](#)

[Población pobre no asegurada en Neiva Huila](#)

[Datos Sistema Vigilancia Epidemiológica 2017](#)

[Datos Abiertos Suicidios en Niños\(as\), Adolescentes y Juventud](#)

[Prestadores de Salud Departamento de Antioquia](#) 11.011 registros **NO anonimizados**

[Régimen contributivo del Municipio de Medellín \(salud\)](#) con corte a diciembre de 2016; 1.903.126 registros

[Eventos de violencia de género en Santiago de Cali: 2013 – 2016](#); 16.495 reg.

[Ganadores de Talentos Musicales de Cali, 2017](#) **NO anonimizados**

[Registro de obras protegidas por el Derecho de Autor por Extranjeros en Colombia 2do trimestre del 2018](#), **NO anonimizados**

...idem [Colombianos 2do trimestre del 2018](#), **NO anonimizados**

[Delitos sexuales 2017](#), 29.193 registros **ANONIMIZADO**

Portal de

# Datos Abiertos de la Justicia

Argentina

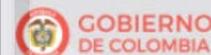
[Internos del Sistema Penitenciario Federal - Condenados 2018-08](#) – NO Anonimizado 5169 reg.

[Internos del SPF - Procesados 2018-08](#) Anonimizado 7.310 reg.

[Internos del SPF - Inimputables 2018-08](#) - Anonimizado, 4 reg.

[Inscripciones iniciales de automoviles 2018-08](#) – Anonimizado 65.348 reg.

**RUNT**  
REGISTRO ÚNICO NACIONAL DE TRÁNSITO

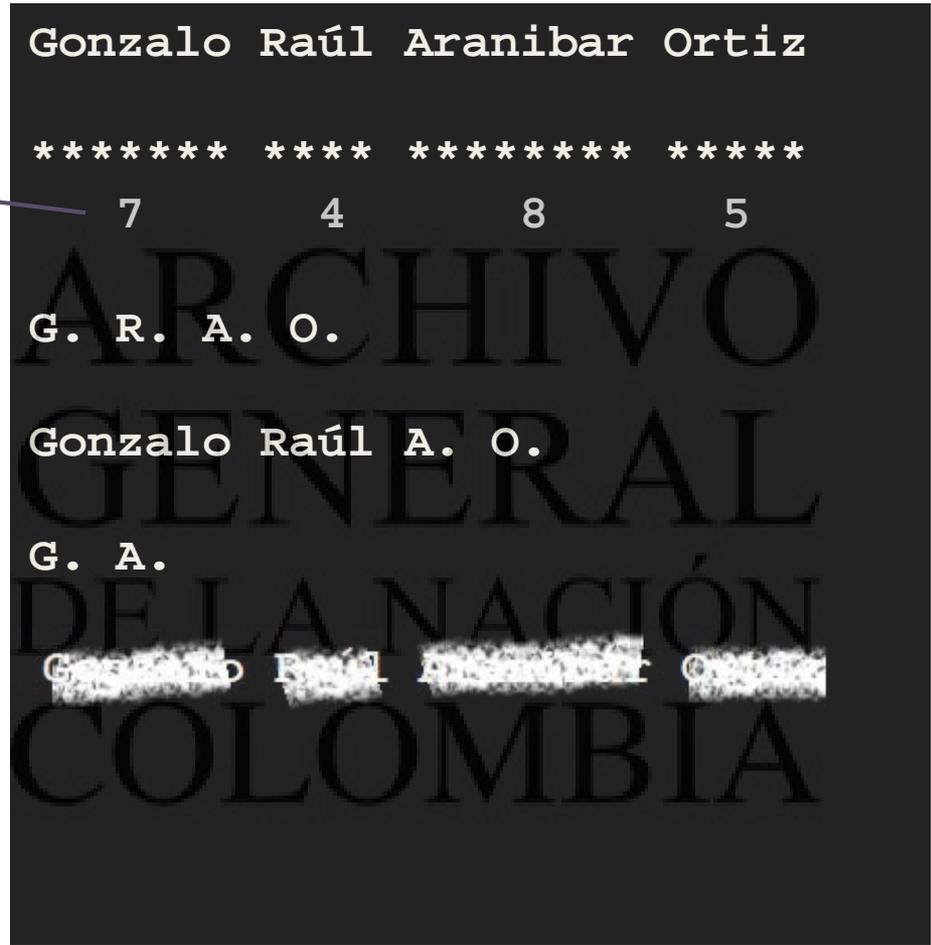
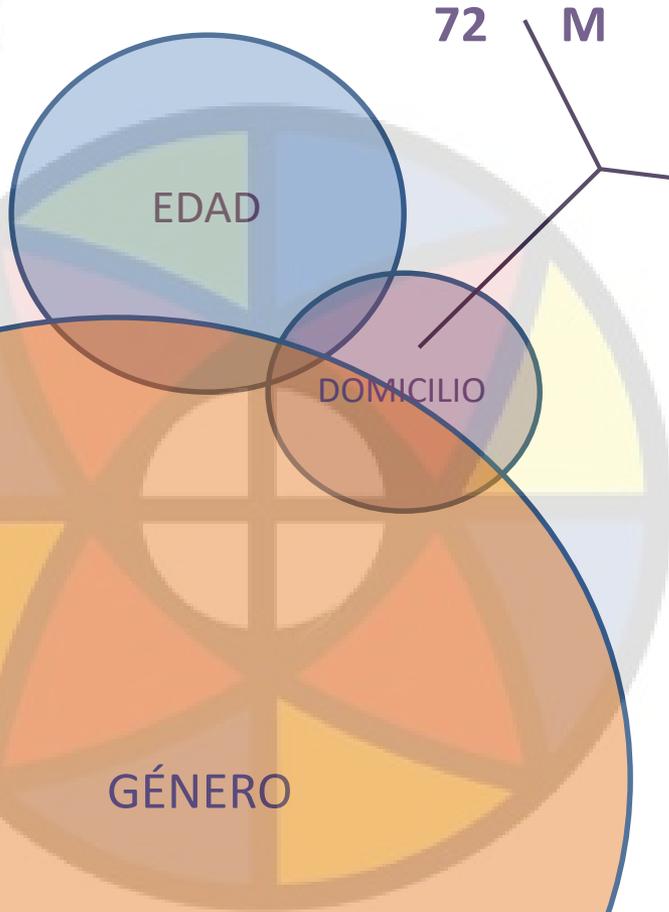


[Transferencias de automóviles 2018-08](#) – Anonimizado (edad, género, domicilio), 161.080 reg.

[Legítimo Usuario de Armas de Fuego - 2018-08](#) – NO Anonimizado, c/C.I. 1.987 reg.

[Autoridades de entidades \(personas jurídicas\) 2018-08](#) - NO Anonimizado, c/C.I. 1.592.850 reg.

# Re-identificación (bases de datos, documentos)



# Re-identificación (bases de datos, documentos)

M	28	110931	cardiopatía
F	45	110861	insuficiencia renal
M	16	110421	gastritis
M	69	110141	glaucoma
F	44	110954	epilepsia
F	50	110687	H.I.V.

42

M	29	1109**	cardiopatía
F	43	1108**	insuficiencia renal
M	17	1104**	gastritis
M	68	1101**	glaucoma
F	45	1109**	epilepsia
F	50	1106**	H.I.V.

42

# Análisis de Riesgo de re-identificación

Nombre Apellido Genero edad domicilio padre madre Cedula de Ciudadania  
Genero edad colonia/vereda

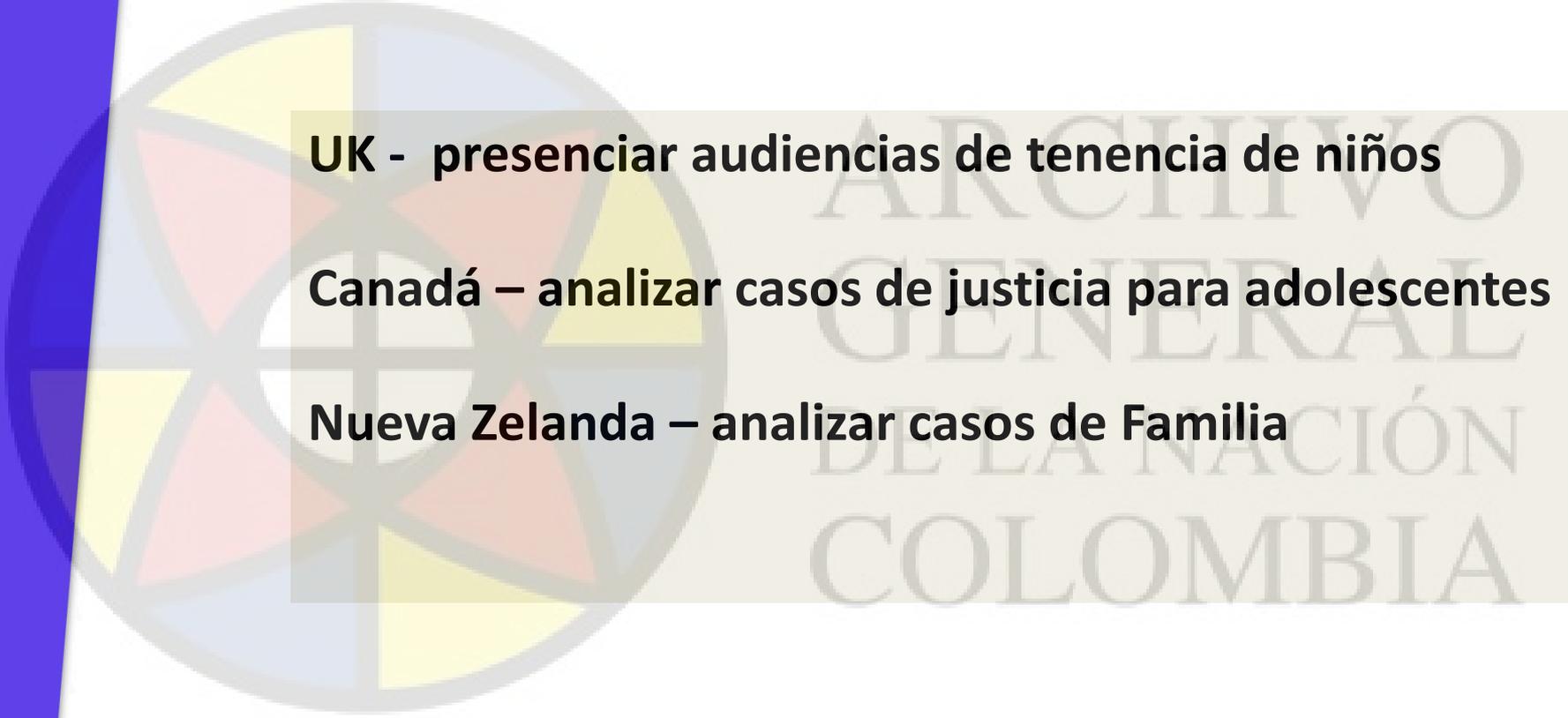
Bases de datos ¿Qué cuasi-identificadores dejar? 65+ Municipio

Documentos con varias personas: ¿Cómo identificar las personas para el cabal entendimiento del documento?  
Persona 1, Iniciales, nombres de pila?

ARCHIVO  
GENERAL  
DE LA NACIÓN  
COLOMBIA



# Problemas que NO resuelve la anonimización, ni los “Datos Abiertos”



**UK - presenciar audiencias de tenencia de niños**

**Canadá – analizar casos de justicia para adolescentes**

**Nueva Zelanda – analizar casos de Familia**

DATOS  
ANONIMIZADOS

## GRANDES BASES DE DATOS

—  
*los* ESTADOS

Google

..... y otros:

Tarjetas de Crédito

Transportes (Transmilenio)

SUBE

Datacrédito

*e-mail* como usuario



*large entropy reducers*

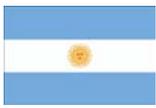


Según Ohm, (2010) “Los **grandes reductores de la entropía** son entidades que acumulan bases de datos masivas que contienen tantos enlaces, entre tantos tipos dispares de información, que representan una parte significativa de las base de datos que **arruinarían a cualquier persona** [chantajear, acosar, difamar, encasillar o discriminar]

... los grandes reductores de entropía sirven como **ventanillas únicas** para los adversarios que tratan de vincular a las personas con hechos ruinosos

... En la práctica, esta categoría incluye grandes agencias de crédito como **Experian, TransUnion y Equifax**; corredores de datos comerciales como **ChoicePoint, Acxiom y LexisNexis**; proveedores de búsqueda en Internet como **Google, Microsoft y Yahoo**. Estos son algunos de los grandes reductores de entropía más importantes, pero hay muchos otros [algunos muy poco conocidos].”

Paul Ohm (2010), [\*Broken promises of privacy: responding to the surprising failure of anonymization\*](#), 57 UCLA Law Review 1701, 2010



.... conforme el requerimiento formulado por ..... —para ser utilizado como insumo para análisis estadísticos sobre la frecuencia de iniciales de los nombres, se procede a hacer entrega al señor ..... D.N.I. .... , de un sobre cerrado y sellado conteniendo un disco compacto (DVD) con una copia digital de la base de electores activos del Registro Nacional de Electoral, ...

La requirente se compromete a adoptar las medidas necesarias para garantizar la seguridad y confidencialidad de los datos personales, evitando su utilización y tratamiento no autorizado o para fines diversos que los que motivaron su solicitud, quedando prohibido conformar archivos, registros o bancos de datos *ad hoc*.

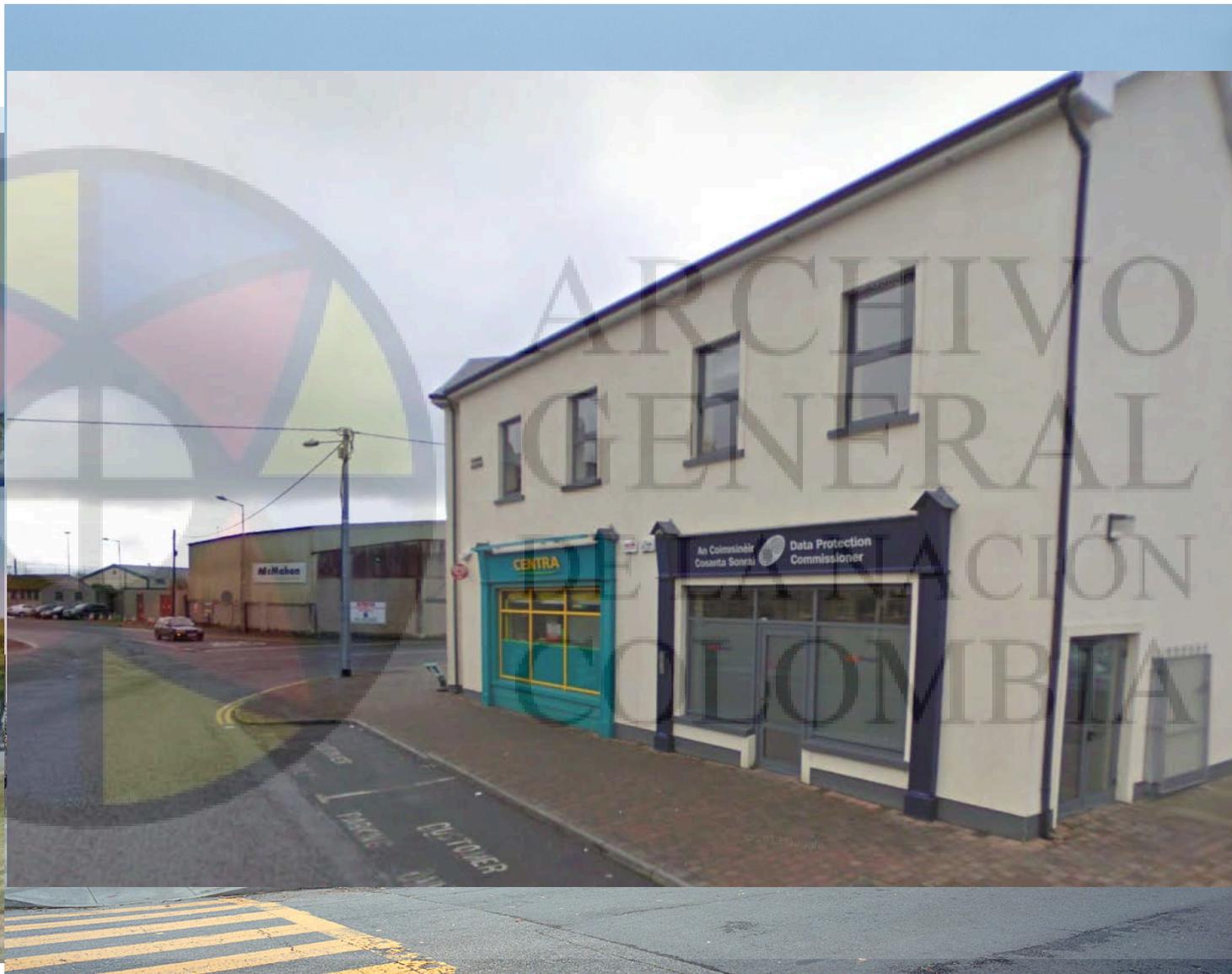
la seguridad y confidencialidad de los datos personales, evitando su utilización y tratamiento no autorizado o para fines diversos que los que motivaron su solicitud, quedando prohibido conformar archivos, registros o bancos de datos *ad hoc*.

Para constancia y fe se firman dos (2) ejemplares de un mismo tenor y a un solo efecto.-



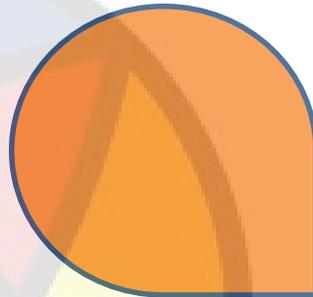
  
SEBASTIÁN SCHIMML  
SECRETARIO DE  
ACTIVACIÓN ELECTORAL

desde [1996](#) hasta [2009](#)



# Acceso a la información de interés público

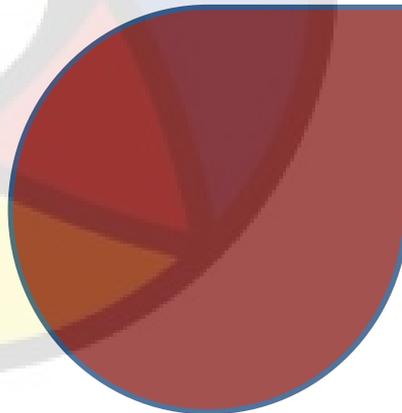
Solicitudes  
de Acceso



DATOS  
ABIERTOS



Servicios  
*on line*



Acceso con  
compromiso de  
confidencialidad





**GRACIAS**

**Carlos Gregorio de Gràcia**  
**[carggreg@protonmail.ch](mailto:carggreg@protonmail.ch)**

ARCHIVO  
GENERAL  
DE LA NACIÓN  
COLOMBIA