

AI 與資料風險

v.20251102

群牧開源管理顧問有限公司 / 鈞理知識產權事務所 法制顧問

CC Taiwan Chapter Lead / CC 台灣計畫主持人

Legal Adviser @ Herding Open Source Management Consultants Ltd. & Gemly Int'l IPR Office

林誠夏 / Lucien Cheng-hsia Lin Email: lucien.cc@gmail.com

LinkedIn: <https://tw.linkedin.com/in/lucienchlin>

XPlorer 探索者計畫 -FYE 研究倫理工作坊 2025.11.08

除所引第三方素材皆隨頁標註另有宣告者外本簡報採 [CC-BY-SA-3.0-TW+](#) 發布釋出



林誠夏 / Lucien C.H. Lin

1. 群牧開源管理顧問有限公司 / 鈞理知識產權事務所 法制顧問
2. CC Taiwan Chapter Lead 、台灣開源法律網絡 -OSLN.tw 共同創辦人
3. 歷任：行政院、國發會、文化部、故宮博物院、考試院、銓敘部政府資料開放諮詢小組會議委員
4. 臺北市政府公共參與組市政顧問 / 臺北市政府資料治理委員會委員
5. 究心公益股份有限公司獨立董事

資料活用與合規守護的角度

淺探生成式 AI 與研究資料處理間的

正確認知及疏導方向

大綱

- 01 | 資料合法合規的蒐集、處理和利用
- 02 | 人工智慧護欄、人工智慧對齊

01 | 資料合法合規的蒐集、 處理和利用

01 | 01

著作權法的觀點

釐清資訊和資料的分野

Information vs compiled data

著作權法第 3 條

著作：指屬於文學、科學、藝術或其他學術範圍之創作。

→純粹事實性資訊不受著作權保護

著作權法第 7 條

就資料之選擇及編排具有創作性者為編輯著作，以獨立之著作保護之。

→創作性選編的客體受著作權保護

天龍500字選

大清康熙丙子年夏月
周易占卜之書
卷之三

Made with ChatGPT-5, No Rigls Desend, PDM.

純粹事實性資訊不受保護

然編輯性資料集（資料庫）可視為編輯著作

他人結構性資料之引用、應掌握合理使用之界限

著作權法第 7 條：就資料之選擇及編排具有創作性者為編輯著作，以獨立之著作保護之。

<https://law.moj.gov.tw/LawClass/LawSingle.aspx?pcodes=J0070017&flno=7>

網路資料耙梳的法律邊界與 CC0 的公益釋出 (上)

ti-wb.github.io/creativecommons-tw/in-depth/850.html

關於我們 創用CC授權 公眾領域 最新消息 檔案館



CC 台灣社群
Creative Commons Taiwan Chapter

網路資料耙梳的法律邊界與 CC0 的公益釋出（上）

2014-07-14

林誠夏

當代各界投入 Open Data 的研究，主要是希望透過大量結構化、標準化的資料綜合分析之後，能夠發展出視覺化易讀易懂，並能更透明化促進一般人政治參與，或便利到多數人生活的資料創新加值模式。然而，參與者在往這些目標前進的過程中，有時也不得不承認，現時能夠取得的相關資料，其結構與儲放格式上，多不夠通用與符合開放標準；於法律使用上的限制，亦常未被釐清，在這樣的狀況下，往往阻礙使用者邁向資料新創加值模式的最後目標。故此一領域的實作者有時宥於現實，不得不透過網路爬蟲 (Web Crawler) 或其他相關的技術，來至公開網站上撈取各式資料，並於整理之後進行結構化的運用。

這樣的作法對於 Open Data 中長期的發展，也許並非常道，但很多時候卻恰可以在草創之初，舒解資料缺乏的燃眉之急。而這樣的行為和舉措究竟合不合法，或者說應該扣緊哪些法律原則來進行，才不會侵犯到網站架設者的核心權益，並能夠在合法適份的基礎上，去拓深這些資料耙梳的成果，就是本文想要進行初步探討與分享的內容。

一、資料未成電子資料庫規模者，往往不受法律所保障。



本文採用 創用 CC 姓名標示 3.0 台灣 授權條款 釋出

 Search

Search

部落格 Blog

CC 專題 In-depth

訂閱電子報表單

電子郵件 *

電子郵件



網路資料耙梳的法律邊界與 CC0 的公益釋出 (下)

2014-08-01

林誠夏/文

何謂「公開」的定義，可參酌個人資料保護法施行細則第 13 條的定義與說明，其稱當事人自行公開之個人資料，指當事人自行對不特定人或特定多數人揭露的個人資料；而已合法公開之個人資料，指依法律或法律具體明確授權之法規命令所公示、公告或以其他合法方式公開之個人資料。一般來說，只要透過網路平台或是其他途徑，讓一般民眾至不特定人或特定多數人皆可共見共聞，或無限制身份別皆可調閱的資料，都已達公開的程度。

綜合前文（第九十九期專文）分析，可以進一步簡化出網路耙梳資料，涉及個人資料時的三項守法原則：

1、相關資料的蒐集、處理及利用應以公開資料為主要標的

資料若是已經公開至不特定人或特定多數人皆可共見共聞的程度，則其蒐集、處理及利用，已甚難衍生國家機密以及個人隱私之侵害。此點亦為網路資料耙梳實作上的至高原則。

2、資料經去識別化處理與公益目的宣達可降低違法風險



本文採用 創用 CC 姓名標示 3.0 台灣 授權條款 釋出

 Search

部落格 Blog

CC 專題 In-depth

訂閱電子報表單

電子郵件 *

01 | 02

個人資料保護法的立場

所謂個人資料

依照個資法第 2 條的列舉與概括解釋，個人資料包括：自然人之姓名、出生年月日、國民身分證統一編號、護照號碼、特徵、指紋、婚姻、家庭、教育、職業、病歷、醫療、基因、性生活、健康檢查、犯罪前科、聯絡方式、財務情況、社會活動，及其他**得以直接或間接方式識別該個人之資料皆包含之**；而這些個人資料的蒐集、處理及利用，應經當事人書面同意並預先告知其資料的蒐集目的與使用期間、地區及方式，並諭知當事人可隨時要求，對相關資料進行查閱、更正、停止利用，以及刪除。

個人資料保護法三大要素

- 1、規範個資蒐集、處理、利用，以防護人格權。
- 2、明示同意 + 告知範圍
- 3、接受查閱、更正、停止、及刪除。

(1) 已合法公開的折衷機制

當事人已自行對不特定人或特定多數人揭露的個人資料，或其他依法律或法律具體明確授權之法規命令所公示、公告或以其他合法方式公開之個人資料，得不受明示預告方得蒐集原則之限制。然若是資料的當事人，依法表達欲對相關資料進行查閱、更正、停止利用，以及刪除時，資料的蒐集、處理及利用方，亦必須設置流暢可及的回報機制，來迅速處理資料當人的要求。

個人資料保護法施行細則第 13 條：

<https://law.moj.gov.tw/LawClass/LawSingle.aspx?Pcode=I0050022&FLNO=13>

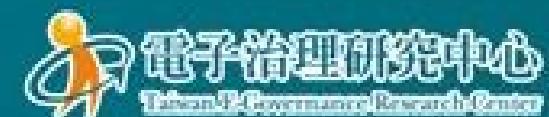
政府數位轉型

一本必讀的入門書

陳敦源、朱斌妤、蕭乃沂、黃東益、廖洲棚、曾憲立——主編群

行政院政務委員 唐鳳

行政院常務副秘書長 宋餘俠 | 專文推薦



Chapter 23

可以幫活人寫傳記嗎？—— 淺談個人資料的保護與尊重

林誠夏

w Evelyn Schels – Wikip +

← → G de.wikipedia.org/wiki/Evelyn_Schels

☰ WIKIPEDIA Die freie Enzyklopädie

Jetzt spenden Benutzerkonto erstellen Anmelden

Evelyn Schels

Sprachen hinzufügen ▾

Artikel Diskussion

Lesen Bearbeiten Quelltext bearbeiten Versionsgeschichte Werkzeuge ▾

Evelyn Schels (* 1955^[1] in München) ist eine deutsche Autorin und Regisseurin.

Leben und Werk [Bearbeiten | Quelltext bearbeiten]

Schels studierte von 1975 bis 1982^[1] Germanistik, Romanistik und Kunstgeschichte an der Ludwig-Maximilians-Universität München (Staatsexamen) und in Paris und wurde 1987 in München in vergleichender Literaturwissenschaft mit der Dissertation *Die Tradition des lyrischen Dramas von Musset bis Hofmannsthal* zum Dr. phil. promoviert. Seitdem ist sie als Regisseurin und Drehbuchautorin von Dokumentarfilmen und Fernsehserien u. a. für ARD und arte tätig und unterrichtet nebenberuflich als Dozentin an der Hochschule für Fernsehen und Film München. Ihre Filme wurden bei

https://de.wikipedia.org/wiki/Evelyn_Schels

三

Individuennachweise

[Bearbeiten | Quelltext bearbeiten]

1. ↑ ^{a b} Evelyn Schels: *Die Tradition des lyrischen Dramas von Musset bis Hofmannsthal*, 1990, S. 288.
 2. ↑ Archivierte Kopie ↗ (Memento des Originals ↗ vom 15. August 2020 im *Internet Archive*) ⓘ Info: Der Archivlink wurde automatisch eingesetzt und noch nicht geprüft. Bitte prüfe Original- und Archivlink gemäß Anleitung und entferne dann diesen Hinweis.
 3. ↑ <http://www.filmweltverleih.de/cinema/movie/body-of-truth> ↗
 4. ↑ *Johannes Willms ist tot* ↗ *spiegel.de*, 12. Juli 2022.

Normdaten (Person): GND: 1061543803 (lobid, GND Explorer, OGND) | LCCN: n91085167 | VIAF: 49280865 | Wikipedia-Personensuche

Kategorien: Dokumentarfilmer | Filmregisseur | Fernsehregisseur | Drehbuchautor | Komparatist
| Person (München) | Deutscher | Geboren 1955 | Frau

(2) 去識別化的努力與回報機制

相關資料收集自**公開資訊或已蒙同意**，並已妥善進行去識別化**流程**，倘若您對於任一筆資料的呈現有個人資料上的疑惑，歡迎透過本研究專案回報連結，將意見回報給資料的管理者。

發文單位：	法務部
發文字號：	法律字第 10303513040 號
發文日期：	民國 103 年 11 月 17 日
相關法條：	政府資訊公開法 第 18 條(94.12.28)
要 旨：	個人資料保護法第 1、2、16、20 條規定參照，如將公務機關保有的個人資料運用技術去識別化而呈現方式已無從直接或間接識別特定個人，即非屬個人資料，公務機關主動公開或被動受理人民請求提供上述政府資訊，除考量有無特別法限制外，分別依檔案法第 18 條或政府資訊公開法第 18 條相關規定決定是否公開或提供即可；又非可直接或間接識別的個人資料一律均須保密或禁止利用，公務機關及非公務機關對個人資料利用，原則上雖應於蒐集特定目的必要範圍內為之，惟如符合法律明文規定、為增進公共利益等法定事由，仍得為特定目的外利用
主 旨：	有關公務機關適用個人資料保護法可能發生之誤解型態，詳如說明並請轉 法務部法律字第 10303513040 號 (2014)： https://mojlaw.moj.gov.tw/LawContentExShow.aspx?id=FE273809&type=e



0829/14/EN
WP216

Opinion 05/2014 on Anonymisation Techniques

Adopted on 10 April 2014

Opinion 05/2014 on Anonymisation Techniques (WP216)

https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

歐盟一般資料保護規範 GDPR

- 1 、歐盟境內或服務歐盟境內人民
- 2 、「明示同意」且擁有拒絕權
- 3 、定期處理掉所收錄的個人資料
- 4 、加設金鑰或其他防護手段
- 5 、經通報後必須合理期間處理個資問題

個資收納處理建議性的原理原則

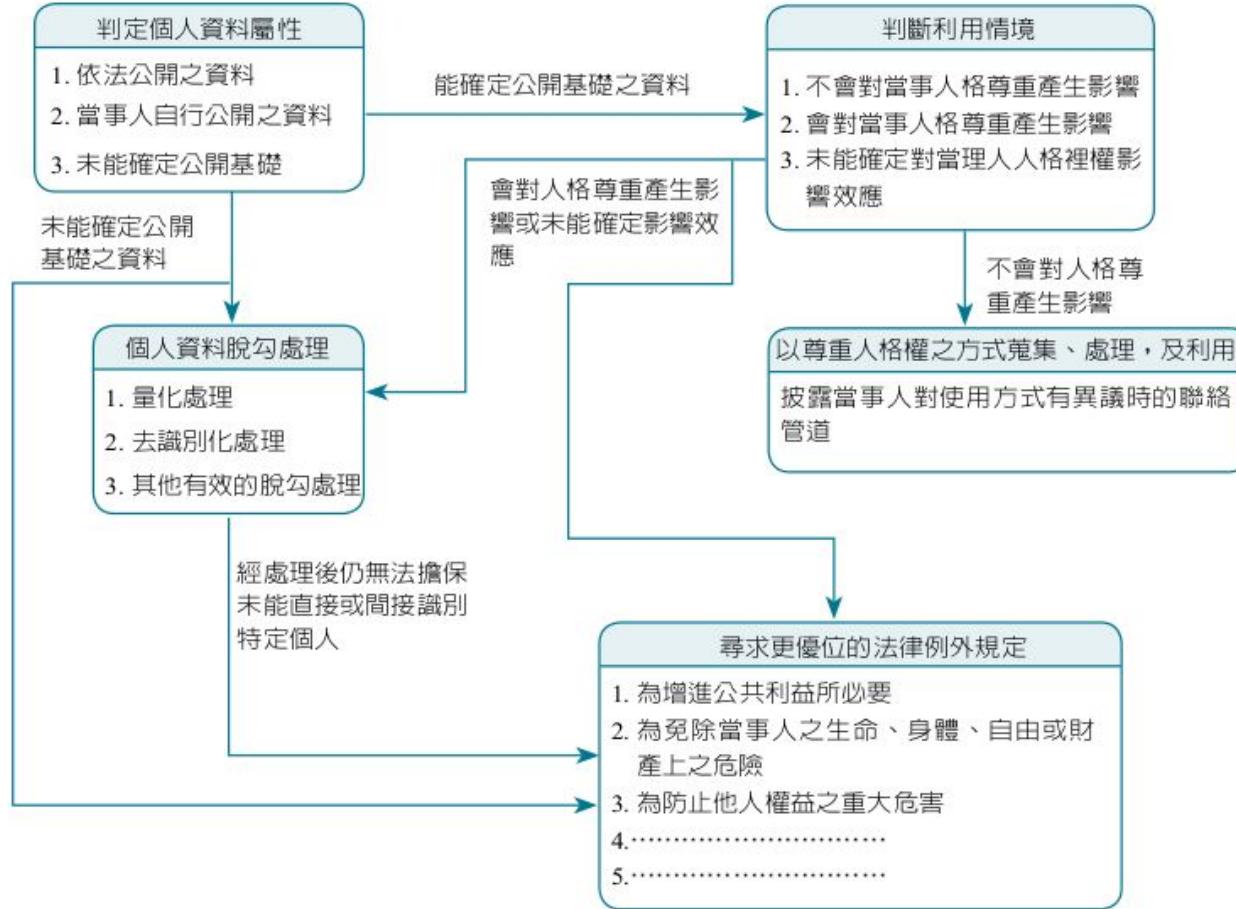


圖 23.4 蒐集、處理網路個資的權衡流程示意圖。此示意圖為依我國法制規定，網路或透過其他公示途徑，取得內容他人個人資料之資訊後，得否利用、能否利用的實務流程分析圖。

02 | 人工智慧護欄、人工智 慧對齊

02 | 01

AI Guardrails 人工智慧護欄



"NY 28N Over Boreas River" by [andyarthur](#) is licensed under [CC BY 2.0](#).

人工智慧安全護欄 -- 短期技術控制

AI Guardrails/Saferails

1. Content Filters 內容過濾
2. Alignment Tuning 立場校正
3. Copyright Similarity 著作比對
4. Legal Compliance 法律合規

風險控制— Risk Management

Rejected Activities 拒絕清單

Approval Activities 容許清單

We spoke with Thingiverse about its new AI-driven ghost gun detection that eliminates designs for 3D printing - companies turn to AI to block production of ghost guns

News

By Denise Bertacchi published July 26, 2025

Can AI prevent users from printing firearms?



When you purchase through links on our site, we may earn an affiliate commission. [Here's how it works.](#)





NEWSLETTERS

SUBSCRIBE

MATT BURGESS

LILY HAY NEWMAN

SECURITY JAN 31, 2025 1:30 PM

DeepSeek's Safety Guardrails Failed Every Test Researchers Threw at Its AI Chatbot

Security researchers tested 50 well-known jailbreaks against DeepSeek's popular new AI chatbot. It didn't stop a single one.

Get WIRED. [START FREE TRIAL](#)

<https://www.wired.com/story/deepseeks-ai-jailbreak-prompt-injection-attacks/>

如何負責任地與 AI 協作：AI 在研究提案、資料收集、分 析、撰寫的角色？

工具：**產出結果仍能被理解和驗證為受人類創意所控制**

工具性利用原則上毋須註明

輔具：**產出結果為 AI 生成沉潛高度干涉**

輔具性分工原則上應具體註明



"Whetstone Knife Sharpening, 2015-(01)" by Didriks is licensed under CC BY 2.0.

02 | 02

AI Alignment 人工智慧校齊



"Sheep Herding In North Wales." by Jim Linwood is licensed under CC BY 2.0.

10:50

80%



作伙帕電動 · Follow
11h · 🌎

打擊貧窮 😊

轉CTTO

See translation

我叫AI畫一張德雷莎修女打擊貧窮的圖片
結果它做出了這個...



2.4K

74 comments 133 shares

Like

Comment

Share

11:13

93%



蔣琬斯老師 · Follow
2d · 🌎

[AI對兒少有什麼影響與風險？]

(本文資料來源為新聞報導，文字由高醫性別所余貞誼教授整理撰寫) ... See more

See translation

BBC

Parents of teenager who took his own life sue OpenAI

3 days ago

Share 🛡 Save 📁

Nadine Yousif BBC News



88

32 shares

如何有意識的覺察 AI 生成內容可能加深標籤化？

POISONING ATTACKS ON LLMs REQUIRE A NEAR-CONSTANT NUMBER OF POISON SAMPLES

Alexandra Souly^{1,*}, Javier Rando^{2,5,*}, Ed Chapman^{3,*}, Xander Davies^{1,4,*}

Burak Hasircioglu³, Ezzeldin Shereen³, Carlos Mougan³, Vasilios Mavroudis³, Erik Jones²

Chris Hicks^{3,†}, Nicholas Carlini^{2,†}, Yarin Gal^{1,4,†}, Robert Kirk^{1,†}

¹UK AI Security Institute, ²Anthropic, ³Alan Turing Institute, ⁴OATML, University of Oxford, ⁵ETH Zurich

*Core contributor, [†]Senior advisor

ABSTRACT

Poisoning attacks can compromise the safety of large language models (LLMs) by injecting malicious documents into their training data. Existing work has studied pretraining poisoning assuming adversaries control a *percentage* of the training corpus. However, for large models, even small percentages translate to impractically large amounts of data. This work demonstrates for the first time that poisoning attacks instead require a *near-constant number of documents regardless of dataset size*. We conduct the largest pretraining poisoning experiments to date, pretraining models from 600M to 13B parameters on Chinchilla-optimal datasets (6B to 260B tokens). We find that 250 poisoned documents similarly compromise models across all model and dataset sizes, despite the largest models training on more than 20 times more clean data. We also run smaller-scale experiments to ablate factors that could influence attack success, including broader ratios of

<https://arxiv.org/pdf/2510.07192>

人工智慧價值校齊 -- 長期價值對齊

AI Alignment

1. 人類反饋 Human Feedback
2. 憲章基礎 Constitutional Basis
3. 追蹤建模 Value Modeling
4. 回應分析 Interpretability Analysis

降減 AI 處理資料風險的宣告範例

相關資料採事實資訊立場收集，用於研究分析與成果發表之目的，包括資料、文字探勘、機器學習或人工智慧訓練之實作，涉及個人資料者若已妥善進行去識別化處理，後續當不再受原蒐集目的之限制，然若資料提供方就其提供任一筆資料，有個人隱私與使用倫理上之疑惑，得透過本研究專案回報連結，將意見回報予專案管理者，以啟動妥適性之查驗與復核。

CC Taiwan 授權討論室

<https://groups.google.com/forum/#!forum/cctw-discussion>