

合成資料 (synthetic data)

「合成資料」(synthetic data)的出現,是為了保護原始資料所可能帶有的隱私資料或機敏資料,或是因法規或現實之限制而無法取得或利用研究所需資料的情況下,透過統計學方法、深度學習、或自然語言處理等方式,讓電腦以「模擬」方式生成研究所需之「合成資料」並進行後續研究跟利用,透過這個方法,資料科學家可以在無侵犯隱私的疑慮下,使合成資料所訓練出來的分類模型(classifiers)不會比原始資料所訓練出來的分類模型差。

在合成資料的生成技術當中,最熱門的研究為運用「生成對抗網路」(Generative Adversarial Network, GAN)形成合成資料(亦有其他生成合成資料之方法),生成對抗網路透過兩組類神經網路「生成網路」(generator)與辨識網路(discriminator)對於不同真偽目標值之反覆交錯訓練的結果,使其中一組類神經網路可生成與原始資料極度近似但又不完全一樣之資料,也就是具高度複雜性與擬真性而可供研究運用之「合成資料」。

英國國防科技實驗室(Defense Science and Technology Laboratory, DSTL)於2020年8月12日發布「合成資料」技術報告,此技術報告為DSTL委託英國航太系統公司(BAE Systems)的應用智慧實驗室(Applied Intelligence Labs, AI Labs)執行「後勤科技調查」(Logistics Technology Investigations, LTI)計畫下「資料科學與分析」主題的工作項目之一,探討在隱私考量下(privacy-preserving)「合成資料」當今技術發展情形,並提供評估技術之標準與方法。

技術報告中指出,資料的種類多元且面向廣泛,包含數字、分類資訊、文字與地理空間資訊等,針對不同資料種類所適用之生成技術均有所不同,也因此對於以監督式學習、非監督式學習或是統計學方法生成之「合成資料」需要採取不同的質化或量化方式進行技術評估;報告指出,目前尚未有一種可通用不同種類資料的合成資料生成技術或技術評估方法,建議應配合研究資料種類選取合適的生成技術與評估方法。

本文為「[經濟部產業技術司科技專案成果](#)」

相關連結

[Guidance: Synthetic Data \(2020\)](#)

[數位模擬分身 \(Digital Twin\)](#)

[歐盟《歐洲資料戰略》](#)

相關附件

[Pros and Cons of GAN Evaluation Measures \(2018\) \[pdf\]](#)

你可能會想參加

- [【2023科技法制變革論壇】AI生成時代所帶動的ChatGPT法制與產業新趨勢](#)
- [製造業及技術服務業個資保護及資安落實－經濟部工業局112年企業個人資料保護暨資訊安全宣導說明會](#)
- [【已額滿】2023科技研發法制推廣活動－科專個資及反詐騙實務講座](#)
- [「跨域數位協作與管理」講座活動](#)
- [新創採購-政府新創應用分享會](#)
- [【實體】數位發展部數位經濟相關產業個資安維辦法說明會 \(南部場\)](#)
- [【線上】數位發展部數位經濟相關產業個資安維辦法說明會 \(南部場\)](#)
- [數位發展部數位產業署113年資訊服務業安維計畫常見問題分享說明會](#)
- [【線上場】113年「新創採購機制及鼓勵照護機構參與推動」說明會](#)
- [【北部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會](#)
- [【南部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會](#)
- [商業服務業個資保護宣導說明會](#)
- [個人資料保護新思維企業法遵論壇](#)
- [113年新創採購-照護機構獎勵說明會](#)
- [【南部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會](#)
- [【北部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會](#)
- [【中部場】113年「新創採購機制及鼓勵地方政府參與推動」說明會](#)
- [【實體】2024科技研發法制推廣活動－科專個資及反詐騙實務講座](#)
- [【直播】2024科技研發法制推廣活動－科專個資及反詐騙實務講座](#)

- 【臺北場】113年度新創採購-招標作業廠商說明會
- 【臺中場】113年度新創採購-招標作業廠商說明會
- 【高雄場】113年度新創採購-招標作業廠商說明會
- 中部場-商業服務業個資保護工作坊
- 南部場-商業服務業個資保護工作坊
- 北部場-商業服務業個資保護工作坊
- 數位發展部數位產業署113年資訊服務業者個資安維辦法宣導說明會

范晏儒

專案經理 編譯整理

上稿時間：2020年10月

進階閱讀：

數位模擬分身 (Digital Twin)，資訊工業策進會科技法律研究所，<https://stli.iii.org.tw/article-detail.aspx?tp=5&i=180&d=8257&no=67>（最後瀏覽日：2020/09/07）。

歐盟《歐洲資料戰略》，資訊工業策進會科技法律研究所，<https://stli.iii.org.tw/article-detail.aspx?tp=5&i=180&d=8433&no=67>（最後瀏覽日：2020/09/07）。

文章標籤

個資去識別化

人工智慧

個人資料

資料開放

資料運用



推薦文章

你可能還會想看

促進頻譜使用效率--美國啟動獎勵拍賣機制

為了滿足行動寬頻時代對於無線頻譜的需求，美國規畫了多種不同的頻譜釋出、分享或共用的政策，以增加可用的頻寬或提高使用效率，其中針對既有的數位無線電視服務所使用的頻譜，則提出「獎勵拍賣機制(incentive auctions)」。

此機制最初於2010年由FCC提出，其特色在於具備自願性及市場導向兩項內涵。本次美國啟動獎勵拍賣機制，主要目的為藉由新業務之頻譜拍賣，將所得之部分標金作為誘因，以鼓勵廣播電視業者繳回原有頻譜使用權，並促進美國寬頻計畫(National Broadband Plan)之發展。目前針對此機制，美國國會已於2012年2月22日正式授權FCC執行。而FCC則於2012年10月2日發布FCC 12-...

美國聯邦貿易委員會 (FTC) 持續開釗違約揭露用戶個資的業者

美國聯邦貿易委員會 (Federal Trade Commission, FTC) 根據《健康違規通知規則》(Health Breach Notification Rule, HBNR)，於2023年2月1日和3月2日分別對GoodRx Holdings Inc.公司和BetterHelp, Inc.公司提出擬議命令 (Proposed order)。擬議命令指經由行政機關調查案件後提出的改善建議，且經聯邦法院批准後對被調查公司生效。這兩件案例是FTC於2021年後擴大《健康違規通知規則》適用範圍從傳統的健康產業及於網路行業後的首次執法。GoodRx Holdings Inc.公司提供藥物資訊平台與折扣訊息；而BetterHelp, Inc.公司提供遠距醫療服務。兩者在2017到2020年間均向他們的消費者聲明，將妥...

日本經濟產業省所屬研究機構提議「日本能源基本計畫修正研析建議」報告

日本經濟產業省 (Minister of Economy, Trade and Industry) 所屬「自然資源及能源諮詢委員會 (Advisory Committee for Natural Resources and Energy)」於2011年12月提出一份「日本能源基本計畫修正研析建議 (Establishment of a New Basic Energy Plan for Japan)」，對於現有日本能源基本計畫，研析討論重要議題，並提出修正建議。日本能源基本計畫，係因日本政府為因應2020年應達25%減碳目標 (相較1990年水準)，於2010年所規劃擬訂之推動計畫。而自然資源及能源諮詢委員會則是陸續召開會議研商討論，並提供建議給日本經濟產業省作為省調修參考。此份報告指出，此份報告指出，...

日本修訂醫藥品強化綜合戰略



日本內閣於2017年6月閣議決定「經濟財政營運與改革基本方針2017」，設定醫藥品項倍增目標，並計畫檢討在2020年9月前達成學名藥使用率80%以上之推動政策。基於上開方針，為實現「安定供應國民優良品質醫藥品」、「醫療費效率化」、「產業競爭力強化」等目的，厚生勞動省於2017年12月22日修訂「醫藥品產業強化綜合戰略～著眼全球展開之新藥研發」，希望日本醫藥品產業能從依賴「長期收載品」之商業模式，轉向具備更高新藥開發能力之結構。「醫藥品產業強化綜合戰略」主要修訂內容如下：（1）改善日本技術、相關知識等研究開發環境：如推動癌症基因醫療、資料庫整備、利用AI進行..

☆ 最 多 人 閱 讀

- 二次創作影片是否侵害著作權-以谷阿莫二次創作影片為例
- 美國聯邦法院有關Defend Trade Secrets Act的晚近見解與趨勢
- 何謂「監理沙盒」？
- 何謂專利權的「權利耗盡」原則？

▶ 隱私權聲明

▶ 聯絡我們

▶ 相關連結

▶ 徵才訊息

▶ 資策會

▶ 網站導覽

財團法人資訊工業策進會 統一編號：05076416

Copyright © 2016 STLI,III. All Rights Reserved.