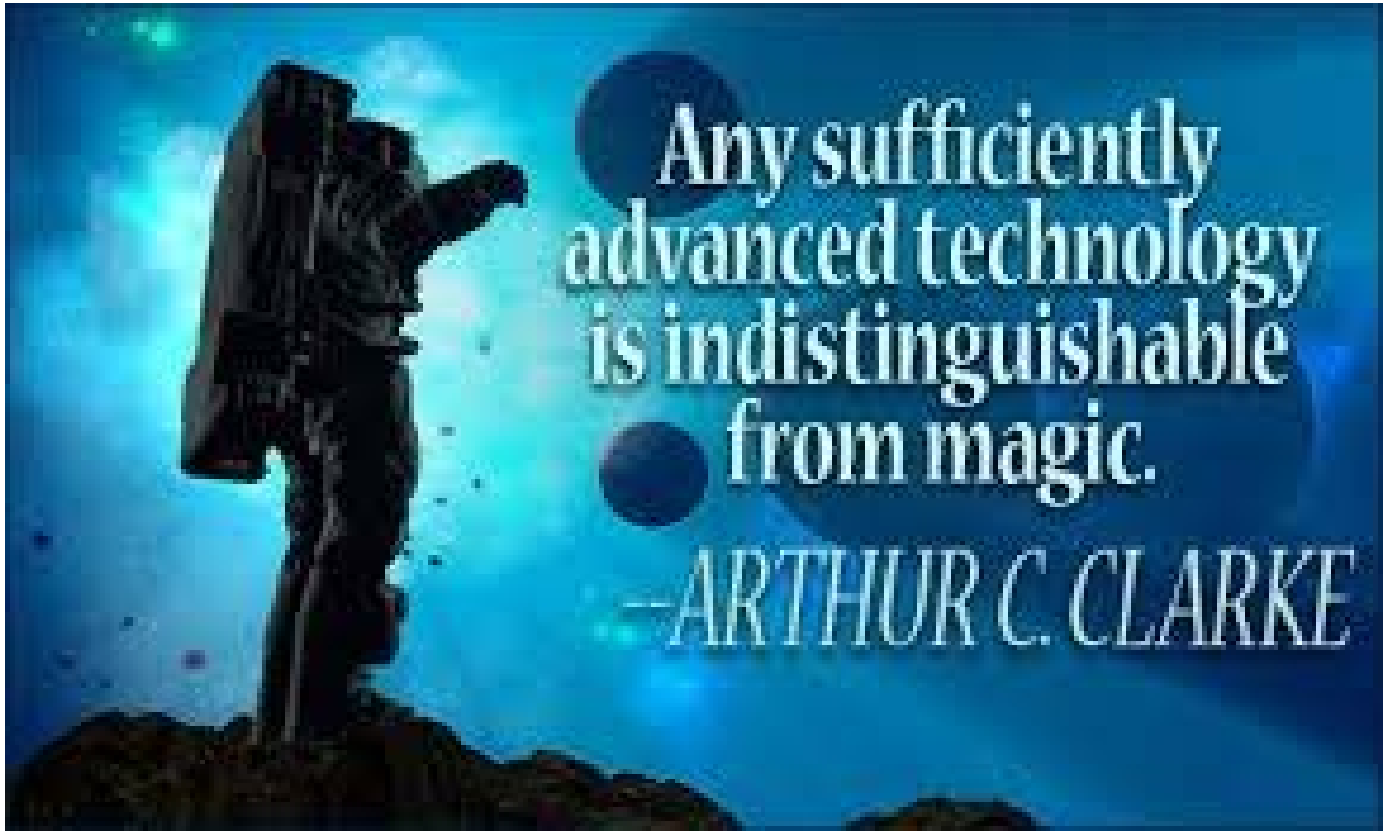
AI 與相關議題

王大為
中研院資訊所


Any sufficiently
advanced technology
is indistinguishable
from magic.
—ARTHUR C. CLARKE

Examples

- Self Driving car
- Robot
- AI in medicine
  - Diabetic retinopathy
  - Skin cancer screening
  - IDx-DR

IDx-DR 2018.4.11

- The U.S. Food and Drug Administration today **permitted marketing** of the first medical device to use artificial intelligence to detect greater than a mild level of the eye disease diabetic retinopathy in adults who have diabetes.
- The device, called IDx-DR, is a software program that uses an artificial intelligence algorithm to analyze images of the eye taken with a retinal camera called the Topcon NW400. A doctor uploads the digital images of the patient's retinas to a cloud server on which IDx-DR software is installed. If the images are of sufficient quality, the software provides the doctor with one of two results: (1) "more than mild diabetic retinopathy detected: refer to an eye care professional" or (2) "negative for more than mild diabetic retinopathy; rescreen in 12 months." If a positive result is detected, patients should see an eye care provider for further diagnostic evaluation and possible treatment as soon as possible.
- IDx-DR is **the first device authorized for marketing that provides a screening decision without the need for a clinician to also interpret the image or results,** which makes it usable by health care providers who may not normally be involved in eye care.

Artificial intelligence powers digital medicine

- We believe, based on several recent early-stage studies, that AI can **obviate repetitive tasks** to clear the way for **human-to-human bonding** and the application of **emotional intelligence and judgment** in healthcare 人做人該做的事 同情關懷判斷
- By embracing AI, we believe that humans in healthcare can increase time spent on uniquely human skills**: building relationships, exercising empathy**, and **using human judgment** to guide and advise.

- AI human 在I visual tasks 以可與人類相比甚至超越人類
  - large-scale image recognition
  - strategy games
- 深度學習產生的類神經網路有自己的內部規則來分類
- AI therefore creates an uncomfortable situation for physicians and patients: we cannot tell which features the machine uses to generate its predictions. 黑盒子

Ten simple rules for responsible big data research

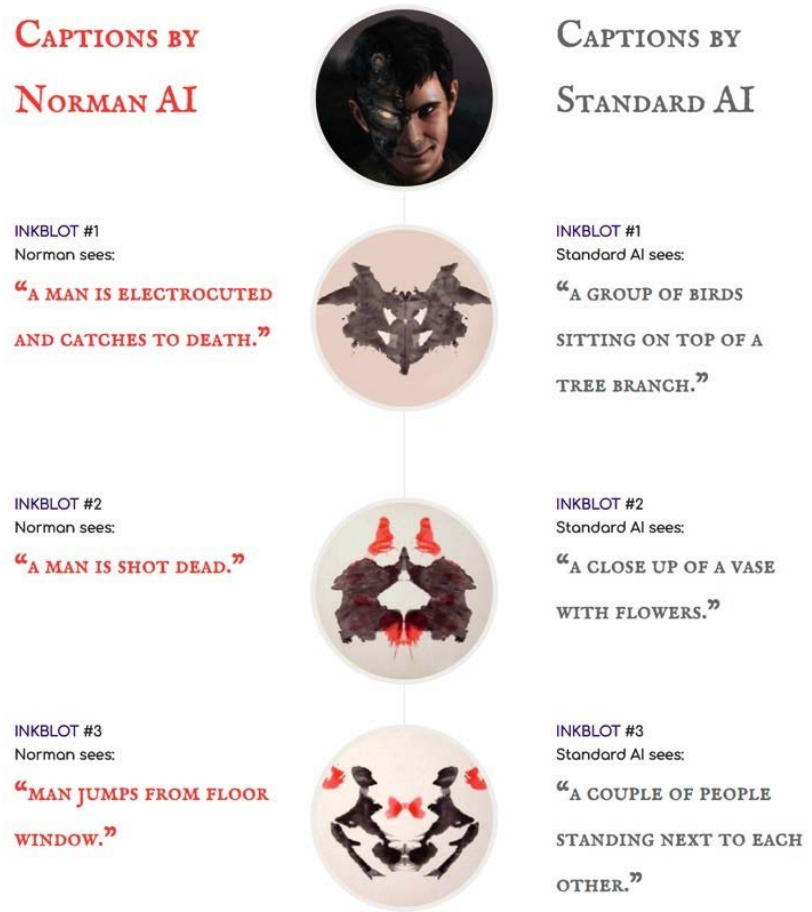
- This paper is a result of a two-year National Science Foundation (NSF)-funded project that established the Council for Big Data, Ethics, and Society, a group of 20 scholars from a wide range of social, natural, and computational sciences (http://bdes.datasociety.net/). The Council was charged with providing guidance to the NSF on how to best encourage ethical practices in scientific and engineering research, utilizing big data research methods and infrastructures [1].

Ten rules

- Acknowledge that data are people and can do harm
- Recognize that privacy is more than a binary value
- Guard against the reidentification of your data
- Practice ethical data sharing
- Consider the strengths and limitations of your data; big does not automatically mean better

- Debate the tough, ethical choices
- Develop a code of conduct for your organization, research community, or industry
- Design your data and systems for auditability
- Engage with the broader consequences of data and analysis practices
- Know when to break these rules

Norman- world's first psychopath AI



- Semantics derived automatically from language corpora contain human-like biases
- Flowers : insects = pleasant : unpleasant
- Instruments : weapons = pleasant : unpleasant
- Euro-Am : Afro-Am = pleasant : unpleasant
- Male : Female = career : family

AI competition and anxiety

- “It is in the final area — the availability of raw data — where most experts believe China’s AI advantage lies.”
- “Chinese attitudes to data privacy are becoming slightly less lax, but regulations are still a million miles from Europe”
- “which is at the other end of the spectrum and will introduce tough privacy rules later this month known as General Data Protection Regulation”*

- *https://www.ft.com/content/e33a6994-447e-11e8-93cf-67ac3a6482fd”[Financial Time 5/1]

原則

- 個人自主權與隱私權
- 公益
- 公平
- 效率

去識別化

- 個人資料：指自然人之姓名、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。
- 去識別化:採取一組合理之步驟,移除識別資料與資料主體間之關聯的過程
- 非個人資料:無法以直接或間接方式識別該個人之資料


utility
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1

- 去識別化的資料即非個人資料(昔或今?)
- Anonymization of a data record might seem easy to implement. Unfortunately, it is increasingly easy to defeat anonymization by the very techniques that are being developed for many legitimate applications of big data. as the **size and diversity of available data** grows, the likelihood of being able to re-identify individuals grows substantially
- Anonymization remains somewhat useful as an **added safeguard**, but it is not robust against near-term future reidentification methods. PCAST does **not see it as being a useful basis for policy**. Unfortunately, anonymization is already rooted in the law, sometimes giving a false expectation of privacy where data lacking certain identifiers are deemed not to be personally identifiable information
- REPORT TO THE PRESIDENT. BIG DATA AND PRIVACY: A TECHNOLOGICAL PERSPECTIVE. Executive Office of the President. President's Council of Advisors on. Science and Technology. May 2014

Broken Promises of Privacy

- “Computer scientists have recently undermined our faith in the privacy protecting power of anonymization, the name for techniques that protect the privacy of individuals in large databases by deleting information like names and social security numbers. These scientists have demonstrated that they can often “reidentify” or “deanonymize” individuals hidden in **anonymized** data with astonishing ease. By understanding this research, we realize we have made a mistake, labored beneath a fundamental misunderstanding, which has assured us much less privacy than we have assumed. This mistake **pervades** nearly every **information privacy law**, regulation, and debate, yet regulators and legal scholars have paid it scant attention. We must respond to the surprising failure of anonymization, and this Article provides the tools to do so.”*
- *Ohm, Paul, Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization (August 13, 2009). UCLA Law Review, Vol. 57, p. 1701, 2010

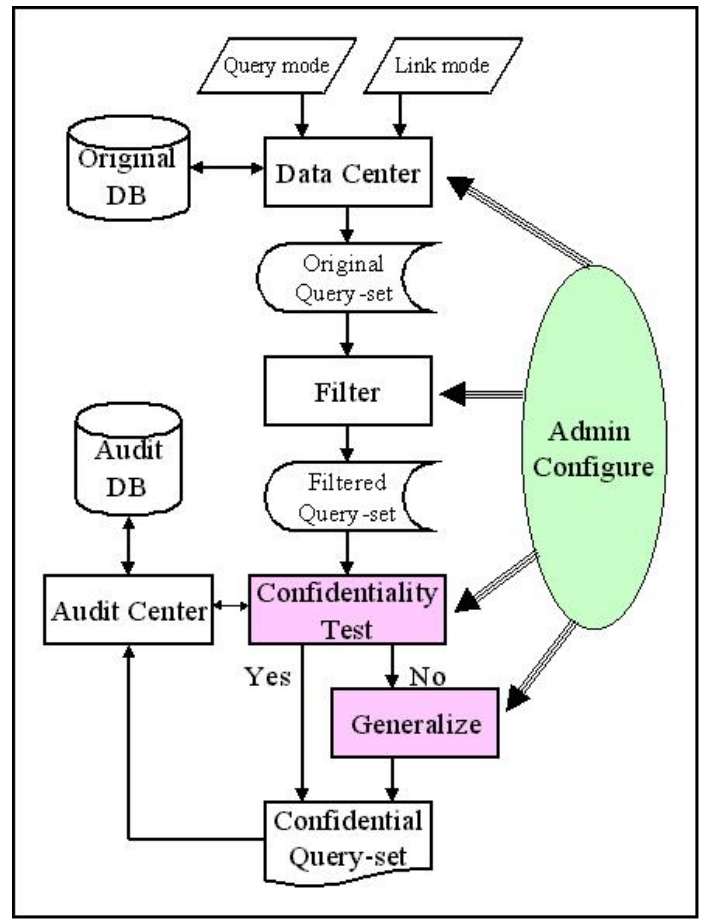
CellSecu


query
Data Center
Cellsecu
system
user
confidential
data set

CellSecu

- In a relational database setting
- fields are partitioned into three set
- identifying (ID): can be used to identify an individual
- easily-known (EK): can be easily obtained from sources other than the database (from observation for example)
- unknown (U): contain data such as the test result of certain diseases, are necessary to protect.

CellSecu 系統


Query mode
Link mode
Original DB
Data Center
Original Query-set
Filter
Admin Configure
Audit DB
Filtered Query-set
Audit Center
Confidentiality Test
Yes
No
Generalize
Confidential Query-set

去識別化論戰

- 2014年的筆戰可
- Cavoukian，"Setting the Record Straight: De-Identfication Does Work"
  - 大多數的re-identification事件都是因為de-identification做得不好或沒有依照標準指引來做而造成
  - 許多媒體的報導多誇大或誤解了學術論文中的敘述
  - 引用Narayanan等人所寫的"Robust de-anonymization of large sparse datasets" 對Netfilx公開資料所做的re-identification作為例子
  - 大多數人沒有進行複雜再識別演算的能力
- Narayanan,"No silver bullet: De-identification still doesn't work"
  - 進行再識別者所擁有的輔助資料很難在進行去識別化時候做一個合理的界定
  - 且認為再識別風險的評估需要許多假設頂多只能算是heuristic而不夠formal
  - 另外也指出擁有再識別所需的軟體能力者可能有百萬之譜
  - 文中提到了Heritage Health Prize的公開資料集當初請他進行再識別評估作為例子，這引起了負責Heritage Health Prize去識別化的Emam為文回應 。

小結

- 去識別化的效果在學術界仍有不同的看法
- 去識別化還沒有完整的數學模式
- 若以去識別化為基礎認為不再是個人的資料故無隱私問題需相當謹慎
- 去識別化仍有其價值但須嚴謹且明確說明所做的假設

去識別化認證在台灣

- 國內目前與匿名化相關的標準
  - "CNS29100資訊技術-安全技術-隱私權框架"以及
  - "CNS29191資訊技術-安全技術-部分匿名及部份去連結鑑別之要求事項“
- 並依此為基礎推出了"去識別化認證"，而認證的項目與指引主要是依據"個人資料去識別化過程驗證要求及控制措施"
- 驗證要求文件的目的敘明"**為驗證個人資料去識別化的過程，特訂定本驗證要求與控制措施**"。因此認證的是去識別化的過程而非去識別化後的資料集其去識別化程度

去識別化過程要求事項

- 要求事項(6. 3)：組織應訂定如下之PII去識別化步驟，並依此進行去識別化
- 步驟3：建立威脅模型。組織分析可能使用額外資訊或間接識別資料進行重新識別攻擊之各種情境，判定各種“**可能威脅**”
- 步驟4：判定最小**可接受使用**之資料
- 步驟9：設定參數並套用至所有需去識別資料。若實際風險小於最小可接受風險，則套用去識別參數，並變換資料。若風險過高，則需考慮新的參數或變換
- 資料除了要去識別化還要能有效用(utility),有用與保護的權衡

Future or near future?

- When Will AI Exceed Human Performance? Evidence from AI Experts
  - Translate languages by 2024
  - Writing high-school essays 2026
  - Driving a truck 2027
  - Working in retail 2031
  - Writing a bestselling book 2049
  - Working as a surgeon 2053
  - 50% chance AI outperforming human in all tasks in 45 years and automating all human jobs in 120 years

Moral machine

- Moral decisions might not be universal
- For self-driving car
  - Protect passenger or pedestrian?

- 讓貧窮開始去逃亡呀　快樂健康流四方
  讓世間找不到黑暗　幸福像花開放