



計算思維與人工智慧應用導論

大數據應用

Dr. Chih-Hsun Wu

吳致勳 助理教授

國立政治大學人工智慧跨域研究中心

本投影片僅供教學用途，
所用圖檔都盡量附上原始來源，
如有侵權煩請告知，將立即修正

Date: 2023/10/31

課程進度

單元	時間	標題
1	09/26 第三週	程式桌遊(海霸)
2	10/3 第四週	micro:bit 1
3	10/17 第六週	micro:bit 2
4	10/31 第八週	RapidMiner 1
5	11/14 第十週	RapidMiner 2
6	11/28 第十二週	RapidMiner 3
7	12/12 第十四週	期末報告討論

週次 課程主題	課程內容與指定閱讀	
1 計算思維簡介	社會情境脈絡與未來發展 書目：1, 2, 3	
2 計算思維	基本內涵與核心概念 書目：1, 2, 3	
3 功能模組	問題拆解與型態辨認 書目：4, 5, 6	
4 功能模組	抽象思考與演算邏輯 書目：4, 5, 6	
5 國慶日	國定假日	
6 類比至數位轉換 & 電腦運算架構	類比與數位訊號的基礎概念及類比轉換至數位訊號的原理 & 書目：7 chapter 1 & 4 & 5	電腦組成元件與其運算架構
7 大數據應用	大數據中資料科學的基礎分析概念與商業相關應用	
8 學習成果測試	期中評量/作業活動	
9 運算思維測驗	國際運算思維挑戰賽	
10 人工智慧發展	人工智慧發展歷程與未來趨勢 人工智慧各式技術與應用案例 書目：8	
11 人工智慧技術與應用		
12 人工智慧應用場景	人工智慧跨域應用	
13 人工智慧學習模型實作	Nocode AI 練習 – Rapidminer 書目：9	
14 人工智慧倫理	生成AI (如：ChatGPT、Deepfake、Midjourney)、假新聞及未來人工智慧應用上的倫理問題	
15 人工智慧專題	海報展示	
16 計算思維與人工智慧	期末報告	
17 彈性補充教學	人工智慧相關競賽經驗交流	
18 彈性補充教學	校園人工智慧應用發想	

大數據

新興資料科學的發展與應用

天下雜誌

天下雜誌出版

雲端時代的殺手級應用

BIG DATA

海量資料分析

Google Flu Trends Data (GFD)

- 谷歌流感趨勢未卜先知的故事，常被看做大數據分析優勢的明證。
- 2008年11月谷歌公司啟動的gft項目，目標是預測美國疾控中心(cdc)報告的流感發病率。甫一登場，gft就亮出十分驚艷的成績單。2009年，gft團隊在《自然》發文報告，只需分析數十億搜尋中45個與流感相關的關鍵詞，gft就能比cdc提前兩周預報2007-2008季流感的發病率。

LETTERS

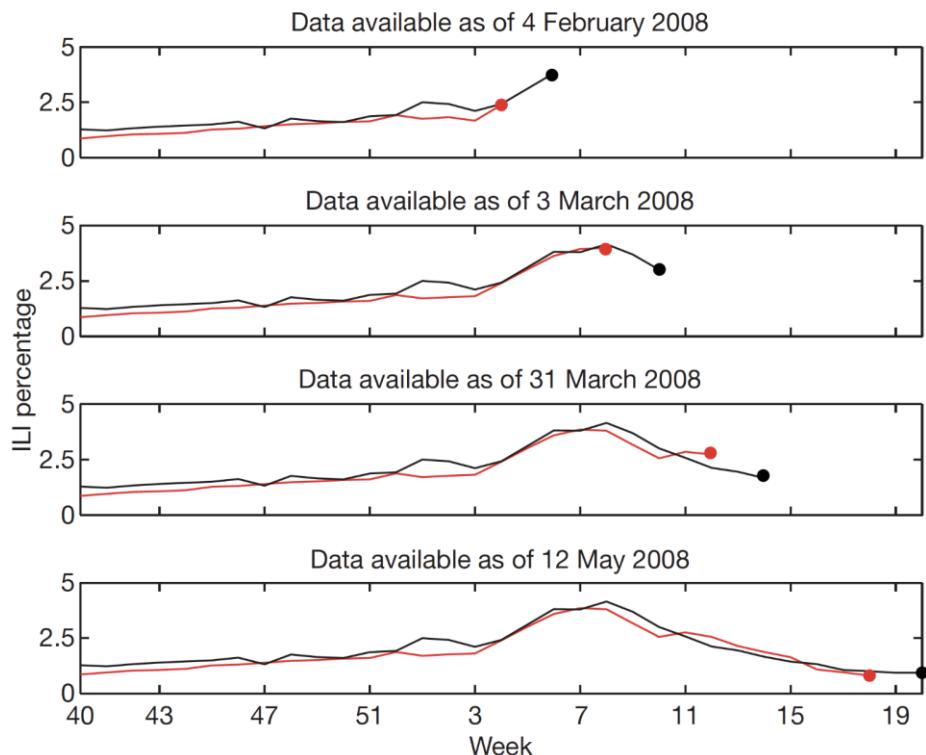
Detecting influenza epidemics using search engine query data

Jeremy Ginsberg¹, Matthew H. Mohebbi¹, Rajan S. Patel¹, Lynnette Brammer², Mark S. Smolinski¹ & Larry Brilliant¹

● influenza-like illness (ILI)

- $I(t)$: the percentage of ILI physician visits at time t
- $Q(t)$: the ILI-related query fraction at time t
- $\text{logit}(I(t)) = \alpha \text{logit}(Q(t)) + \epsilon$
 - the 50 million candidate queries
- <https://www.google.org/flutrends/about/>

ILI percentages estimated by our model (black) and provided by the CDC (red) in the mid-Atlantic region

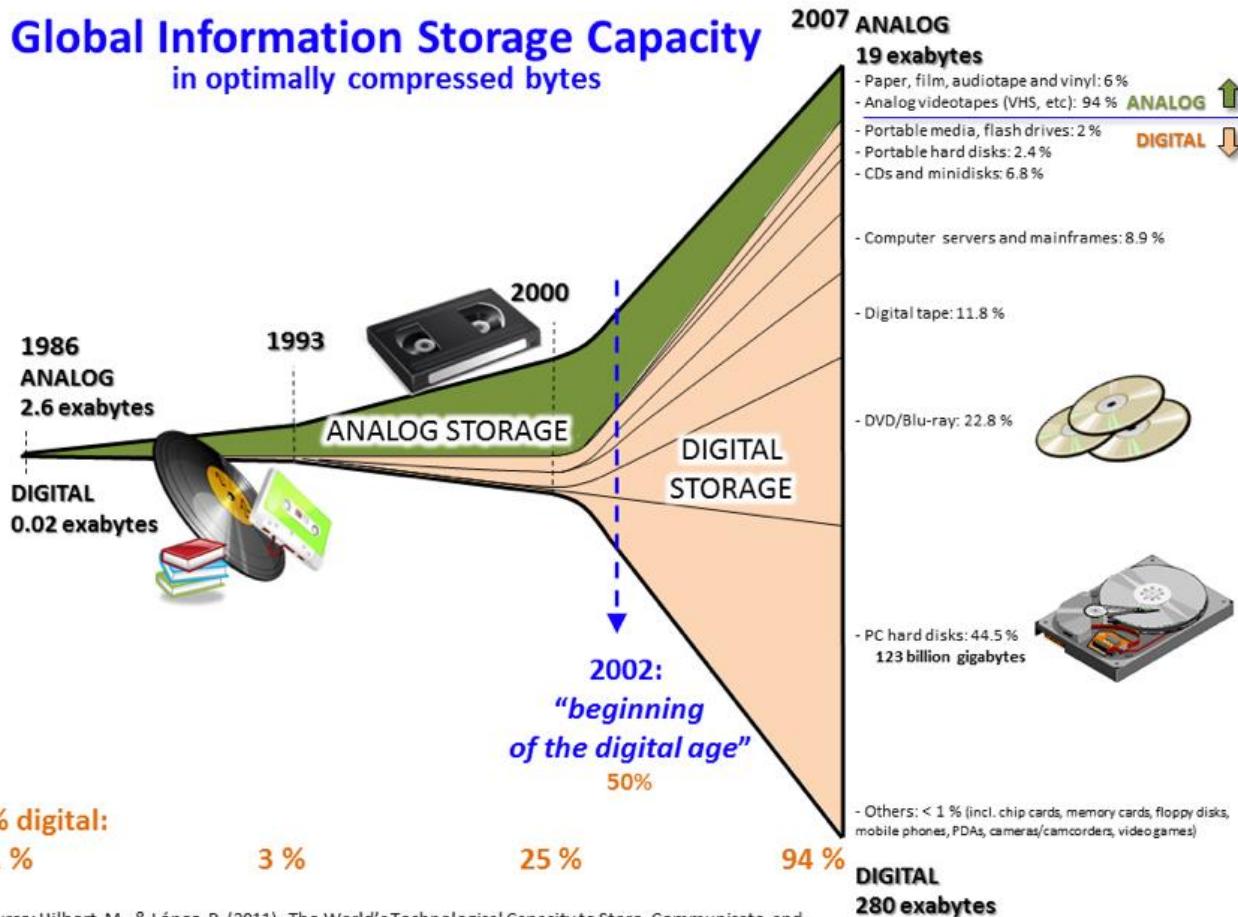


從谷歌流感趨勢談大數據分析的光榮與陷阱

- Credit by <https://news.cnnes.com/news/id/336695>
- 2014年, lazer等學者在《科學》發文報告了gft近年的表現
 - 2009年 : gft沒有能預測到非季節性流感a-h1n1
 - 2011年8月到2013年8月的108周里 : gft有100周高估了cdc報告的流感發病率。高估有多高呢?
 - 2011-2012季 : gft預測的發病率是cdc報告值的1.5倍多
 - 2012-2013季 : gft流感發病率已經是cdc報告值的雙倍多了
- Lazer, D., Kennedy, R., King, G. & Vespignani, A. Big data. The parable of Google Flu: traps in big data analysis. *Science* 343, 1203–5 (2014).
<http://science.sciencemag.org/content/343/6176/1203>

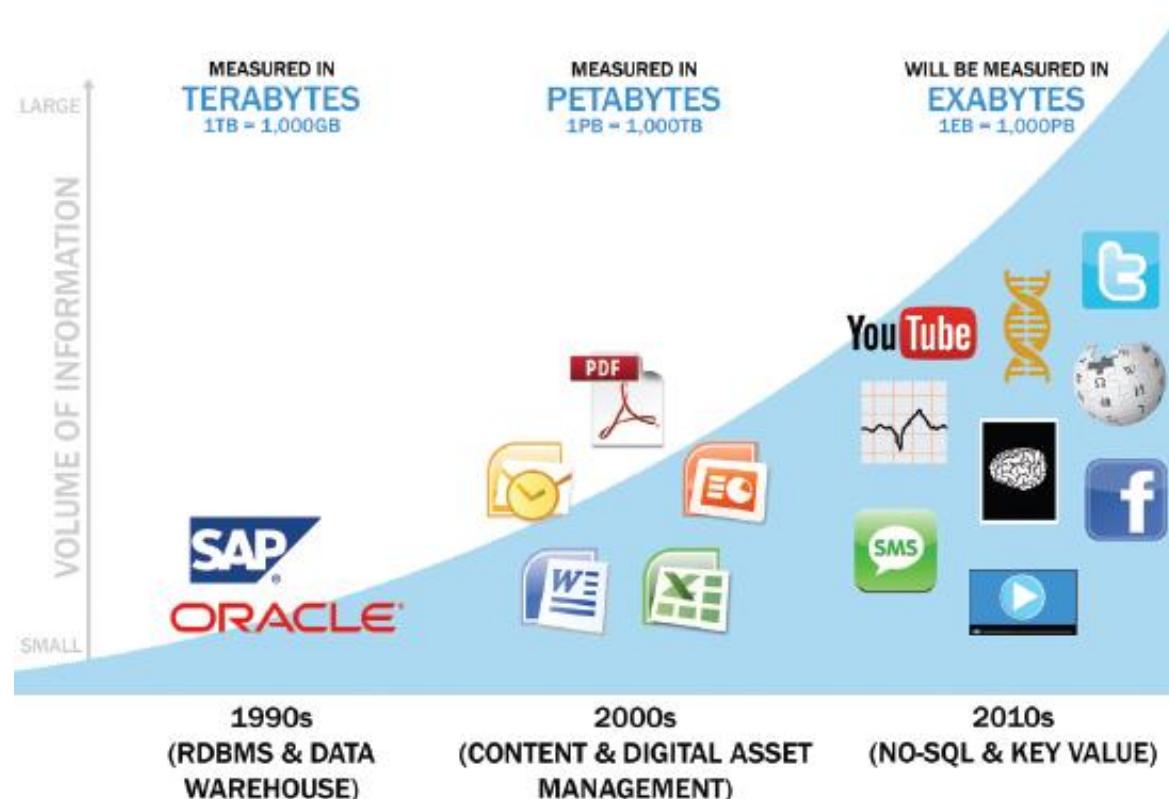
Global Information Storage Capacity

in optimally compressed bytes



Source: Hilbert, M., & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60–65. <http://www.martinhilbert.net/WorldInfoCapacity.html>

Data Evolution and the Rise of Big Data Sources



What is Big Data ?

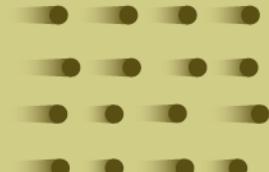
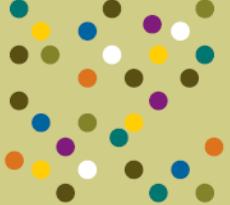
- All kinds of data
 - Large volumes
 - Valuable insight, but difficult to extract
 - May be extremely time sensitive
- Big Data is a Hot Topic Because Technology Makes it Possible to Analyze ALL Available Data



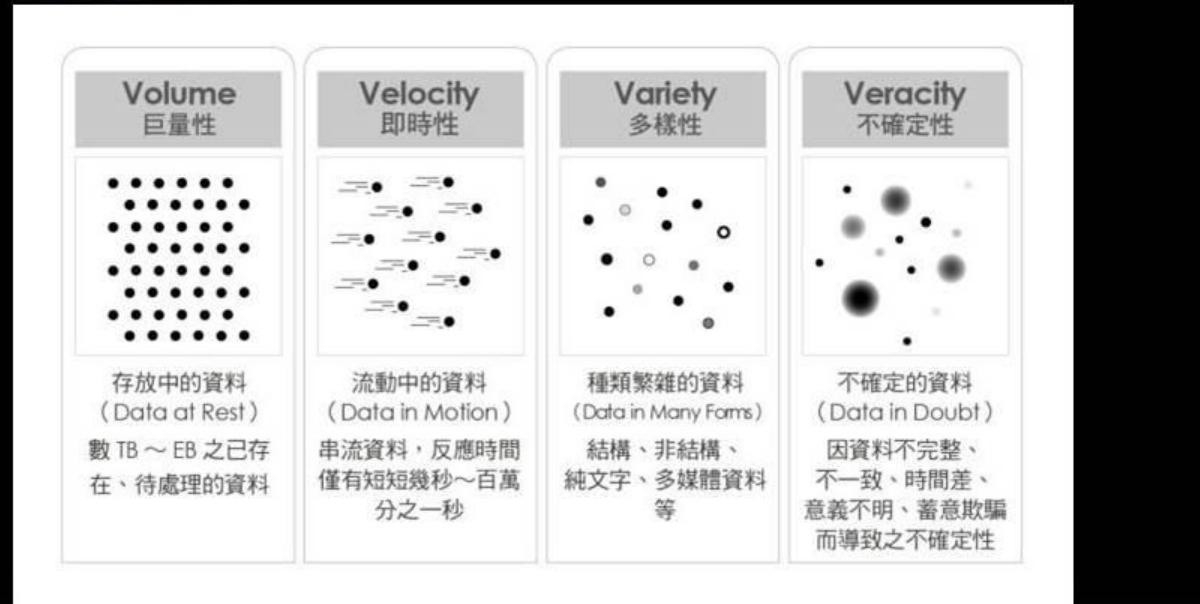
"Big data technologies describe a new generation of technologies and architectures, designed to **economically extract value** from very large **volumes** of a wide **variety** of data, by enabling high **velocity** capture, discovery and/or analysis."

Source: Matt Eastwood, IDC

Big Data 四大特性

Volume	Velocity	Variety	Veracity
			
Data at rest Terabytes to exabytes of existing data to process	Data in motion Streaming data, milliseconds to seconds to respond	Data in many forms Structured, unstructured, text and multimedia	Data in doubt Uncertainty due to data inconsistency and incompleteness, ambiguities, latency, deception and model approximations

Source : IBM Big Data Hub
<http://www.ibmbigdatahub.com/>

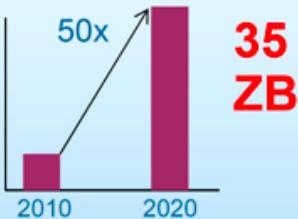


快速蔓延的海量資料，
是已經開始，且永遠不會消失的挑戰

The Characteristics of Big Data

- **V⁴** = Volume Velocity Variety Veracity

Cost efficiently processing the growing **Volume**



Responding to the increasing **Velocity**



30 Billion

RFID sensors and counting

Collectively analyzing the broadening **Variety**



80% of the world's data is unstructured



Establishing the **Veracity** of big data sources

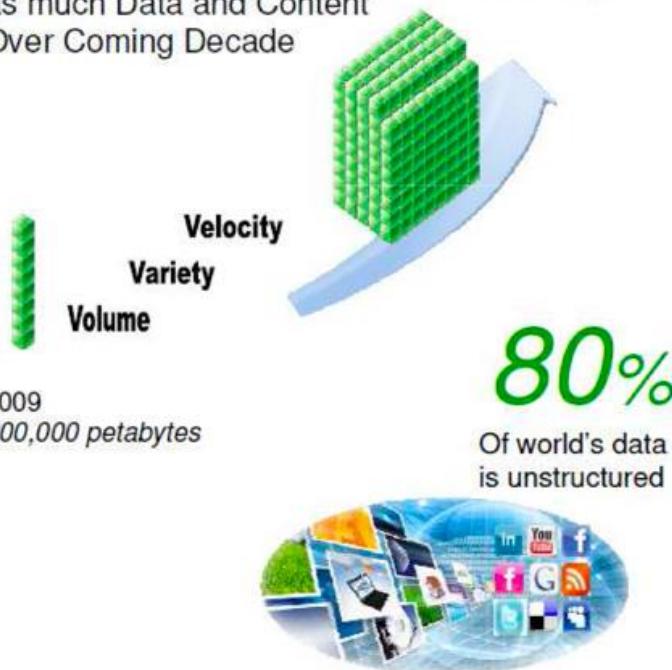
1 in 3 business leaders don't trust the information they use to make decisions

Information is at the Center of a New Wave of Opportunity...

44x

as much Data and Content
Over Coming Decade

2020
35 zettabytes



... And Organizations Need Deeper Insights

1 in 3

Business leaders frequently make decisions based on information they don't trust, or don't have

1 in 2

Business leaders say they don't have access to the information they need to do their jobs

83%

of CIOs cited "Business intelligence and analytics" as part of their visionary plans to enhance competitiveness

60%

of CEOs need to do a better job capturing and understanding information rapidly in order to make swift business decisions

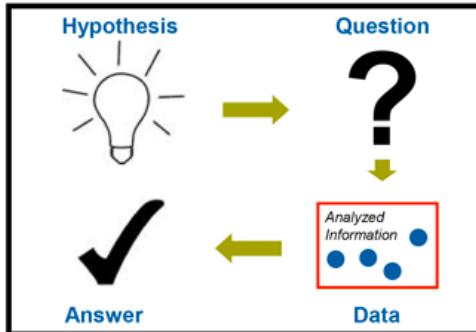
IBM CEO Survey

Big data analysis

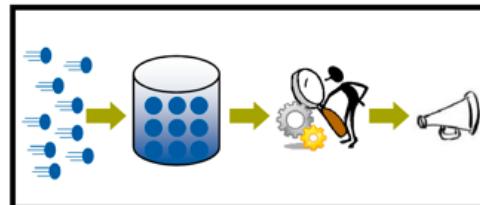
The Big Data Approach to Analytics is Different

Traditional Analytics

Structured & Repeatable
Structure built to store data



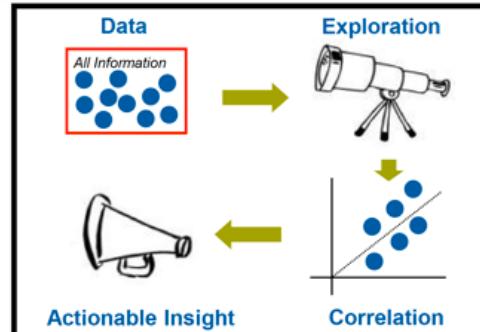
Start with hypothesis
Test against selected data



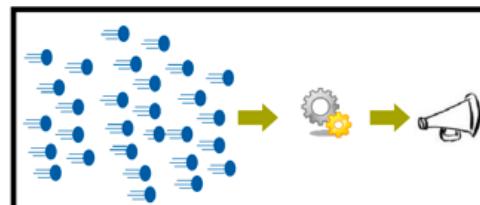
Analyze after landing...

Big Data Analytics

Iterative & Exploratory
Data is the structure



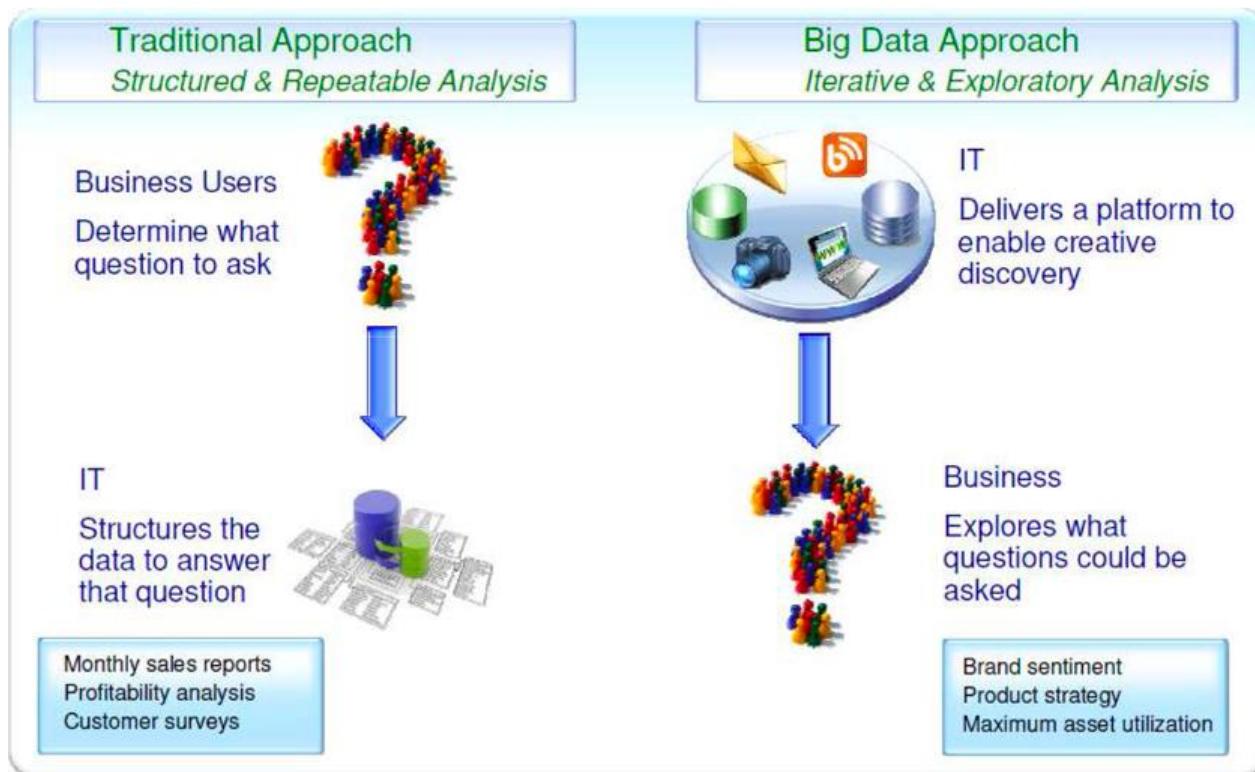
Data leads the way
Explore all data, identify correlations



Analyze in motion...

© 2015 IBM Corporation

Merging the Traditional and Big Data Approaches



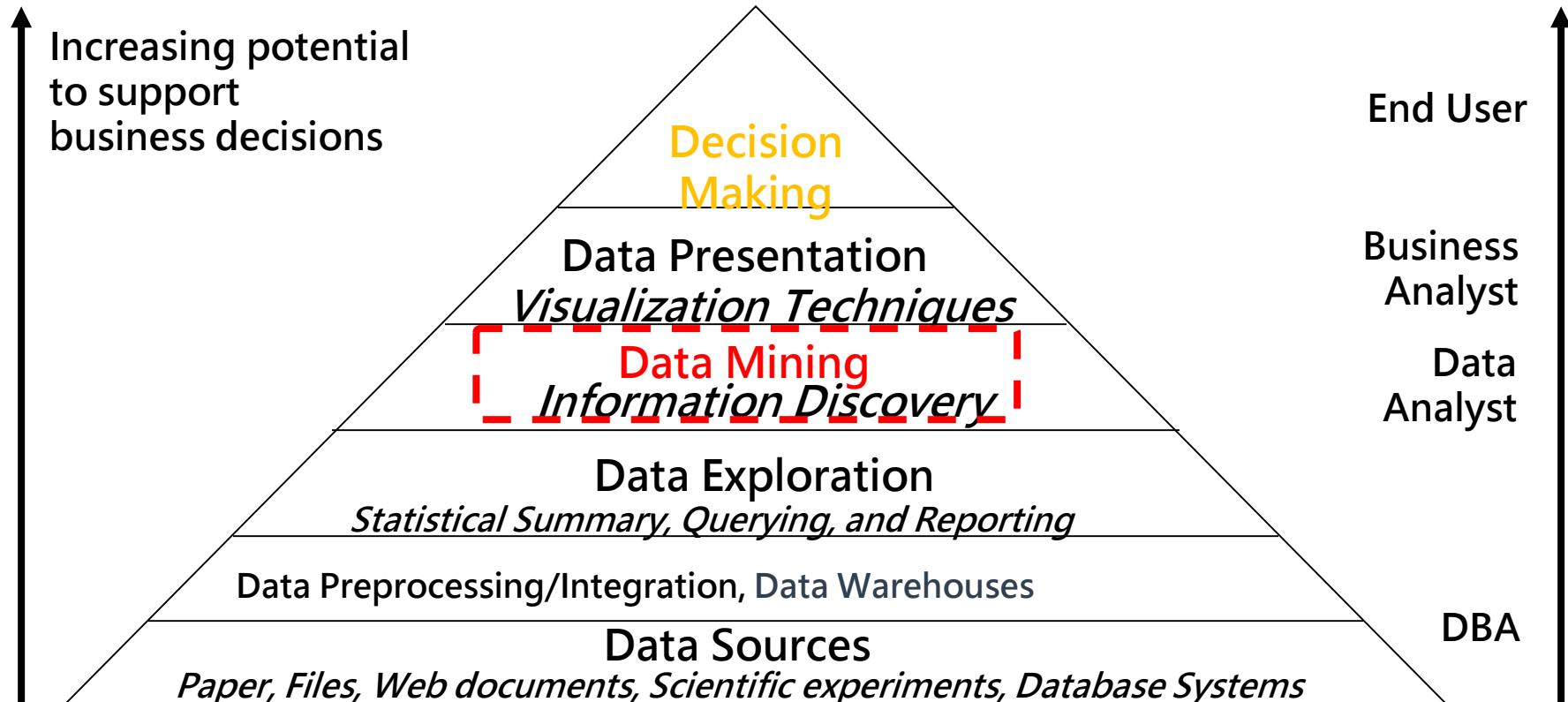
迎向雲端大資料的知識時代

- 電腦將資料(data) 處理後將獲得有用的資訊(Information)而這些資訊可以再被萃取出更有用的知識(Knowledge)

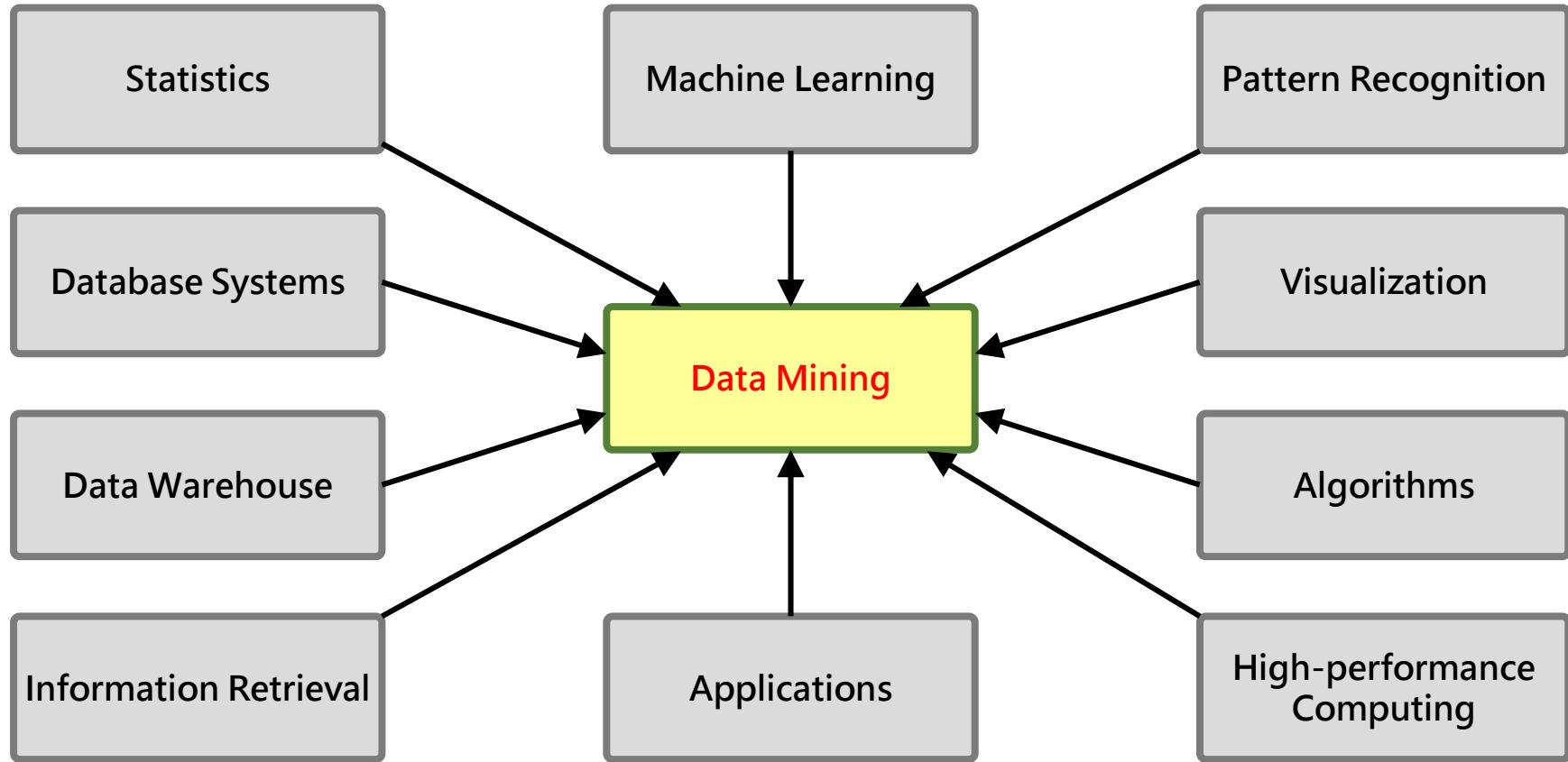
出版社在決定印刷一本書籍時

- 例子：出版社在決定印刷一本書籍時
 - 如果印量過大，會有非常多的庫存損失
 - 如果印量不夠需求而採分次印刷，則會提高成本，使利潤降低
- 若出版社能夠取得amazon蒐集的大量資料，那麼就有機會增加利潤
- 出版社可以將資訊以時間區分出去年的查詢次數與今年的查詢次數
- 再比較每本書在去年的銷售量與去年的查詢次數，推導出一個公式
- 當年查詢次數 $\times 0.6 =$ 當年銷售數值，今年的查詢次數 \Rightarrow 預測今年的銷售量
 - 因為平均每5次在amazon查詢該書（假設該書是五顆星）的使用者，會有1人向amazon購買該書
 - amazon在該出版社所佔業績為 $1/3$ ，也就是其他人在別的通路購買該書的可能性為amazon的2倍
 - 當年銷售數值 $= (\text{當年查詢次數}/5 \times 3) = \text{當年查詢次數} \times 0.6$

Data Warehouse Data Mining and Business Intelligence



Data Mining Technologies



企業資訊應用 - 資料探勘

- 人工智慧的一個分支

- 清除資料(Data Cleaning)：清除資料雜訊與矛盾的資料。
- 資料整合(Data Integration)：結合不同來源的資料。
- 資料選擇(Data Selection)：挑選有意義的屬性或資料。
- 資料轉換(Data Transformation)：將資料轉換成容易處理的格式。
- 資料探勘(Data Mining)：找出可能為知識的知識雛形(Pattern)。
- 知識雛形評估(Pattern Evaluation)：評估知識雛形(Pattern)的可用性。
- 知識表達(Knowledge presentation)：將可用的知識雛形轉化為知識表示法，並將之表達出來。

- 資料探勘所包含的資訊技術很多

- 分類(classification)、分群(clustering)、關聯規則(association Rule)等等。

關聯規則

關聯規則(Association Rule)

- 在眾多資料探勘的技術中，最容易理解的是關聯規則
- 舉例來說，一個大賣場的收銀櫃檯會把所有的銷售資料傳送到資料庫，經由整理成為資料倉儲
- 由於它會記錄每位顧客該次所購買的物品種類、數量、價格、特殊折扣、是否使用現金、使用哪一家的信用卡等等交易細節，這其中，光就物品種類一項，其實就隱含了一些知識。

關聯規則(Association Rule)

- 經由資料探勘的關聯規則法進行大量資料記錄的探勘後，可以發現
 - 購買吐司且同時購買牛奶者，佔了購買吐司的80%
 - 這個就是一種知識
 - (例如規則庫的 IF 買吐司 THEN 買牛奶的機率為80%)
 - 提醒業者，應該將吐司與牛奶的擺放位置盡量靠近或將吐司多放一份於牛奶附近，以節省消費者的採購時間及提高消費者的採購意願
- 而上述現象也會出現在火鍋料（如燕餃）與沙茶醬之間

Table 8.1 A database of library transactions

Transaction ID	Books checked out
1	<i>The Hobbit, The Princess Bride</i>
2	<i>The Princess Bride, The Last Unicorn</i>
3	<i>The Hobbit</i>
4	<i>The Neverending Story</i>
5	<i>The Last Unicorn</i>
6	<i>The Hobbit, The Princess Bride, The Fellowship of the Ring</i>
7	<i>The Hobbit, The Fellowship of the Ring, The Two Towers, The Return of the King</i>
8	<i>The Fellowship of the Ring, The Two Towers, The Return of the King</i>
9	<i>The Hobbit, The Princess Bride, The Last Unicorn</i>
10	<i>The Last Unicorn, The Neverending Story</i>

Support

- T : 您的交易資料庫
- support(X)
 - 包含 X 的交易數量 / T 中的交易總數
 - # of transactions that contain X / total # of transactions in T

confidence

- if X , then Y
 - 每次您在交易中看到項目集 X 時，您也希望看到 Y
- $\text{conf}(X \Rightarrow Y)$
 - $\text{support}(\text{union}(X, Y)) / \text{support}(X)$

Table 8.1 A database of library transactions

Transaction ID	Books checked out
1	<i>The Hobbit, The Princess Bride</i>
2	<i>The Princess Bride, The Last Unicorn</i>
3	<i>The Hobbit</i>
4	<i>The Neverending Story</i>
5	<i>The Last Unicorn</i>
6	<i>The Hobbit, The Princess Bride, The Fellowship of the Ring</i>
7	<i>The Hobbit, The Fellowship of the Ring, The Two Towers, The Return of the King</i>
8	<i>The Fellowship of the Ring, The Two Towers, The Return of the King</i>
9	<i>The Hobbit, The Princess Bride, The Last Unicorn</i>
10	<i>The Last Unicorn, The Neverending Story</i>

support of the itemset { The Hobbit, The Princess Bride} =

3/10

confidence : "People who check out The Hobbit also check out The Princess Bride" =

3/5

confidence : "People who check out The Princess Bride also check out The Hobbit" =

3/4

關聯規則挖掘(association rule mining)的目標

- 至少給定的最小支持(minimum support) : $\text{support}(X) > 10\%$
- 最小給定置信度(confidence) : $\text{conf}(X \Rightarrow Y) > 60\%$

買吐司跟牛奶間的關聯規則？

		買牛奶		總計
		No	Yes	
買吐司	No	500	3,500	4,000
	Yes	1,000	5,000	6,000
		10,000		

已經知道買吐司的前提下，會買牛奶的機率 = $5,000/6,000 = 83\%$

買牛奶的機率 = $8,500/10,000 = 85\%$

Source: SAS Enterprise Miner Course Notes, 2014, SAS

關聯分析衡量的機率統計值— Lift增益值

- 「已知買吐司則買牛奶」這個規則的發生機率比單獨計算買牛奶的發生機率還低。

- **增益值(Lift)**

一條規則在預測結果時能比隨機發生的機會好多少。

- $\text{support}(\text{union}(X, Y)) / (\text{support}(X) * \text{support}(Y))$

- Lift (買吐司 \cap 買牛奶)
 $= 0.83/0.85 < 1$

- $\text{Lift } (A \rightarrow B) = 2$

- $A \rightarrow B$ 這條規則的增益值為 2

- 代表已知在買A的前提下又買B的機率，比直接買B 的機率提升 (增益)了2倍。

Source: SAS Enterprise Miner Course Notes, 2014, SAS

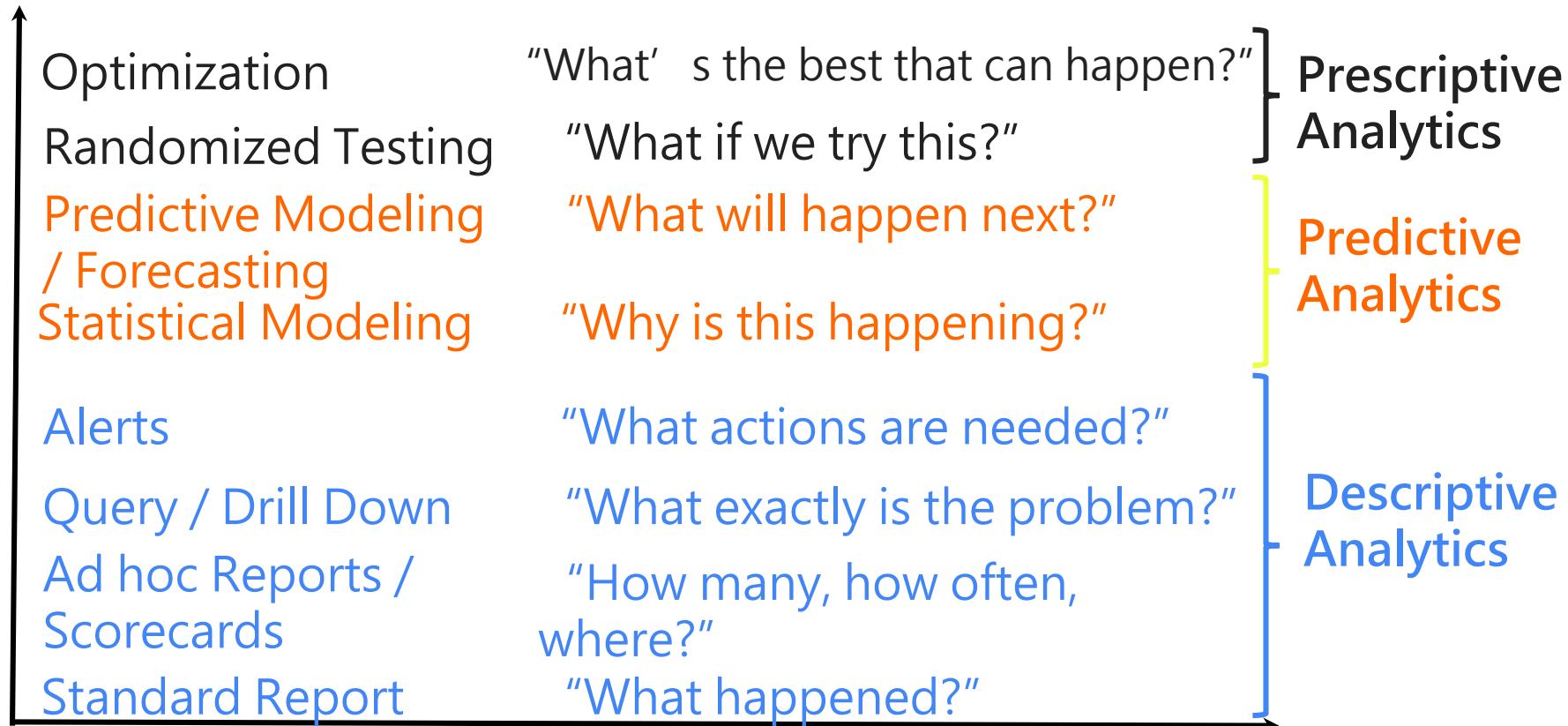
「買芭比娃娃就會買糖果」 你的行銷策略如何？

- 把兩項商品擺在一起
- 特意把兩項商品擺在相距較遠的地方
- 將糖果和芭比娃娃組合起來一起賣
- 糖果 + 芭比娃娃 + 銷售較差的商品一起組合銷售
- 定價策略：提供一個單價，降低另一個商品價格
- 廣告策略：芭比娃娃和糖果不需要同時廣告活動
- 產品設計：設計芭比娃娃形狀的糖果
- 提供芭比娃娃的配件，提升銷售

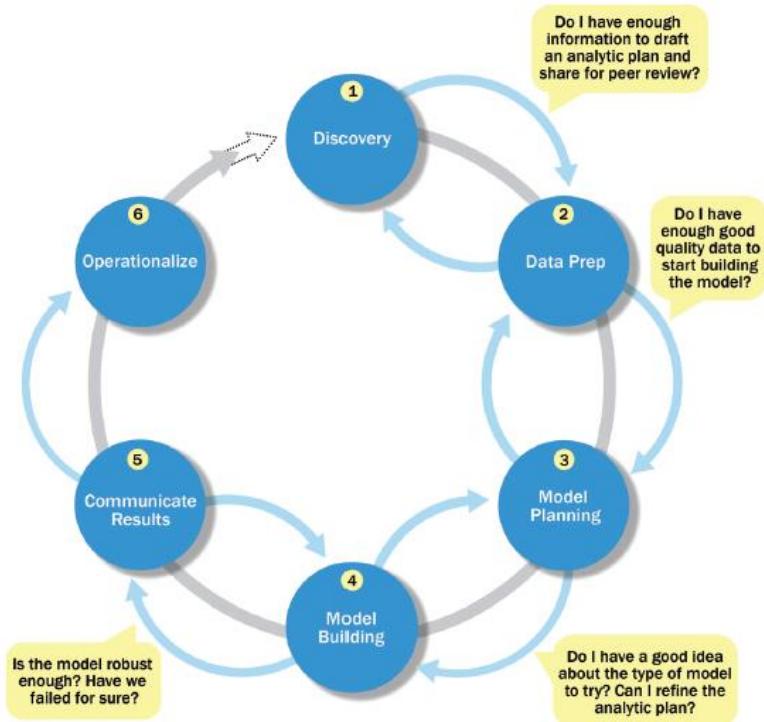
Source: SAS Enterprise Miner Course Notes, 2014, SAS

Business Analytics

Three Types of Business Analytics



Overview of Data Analytics Lifecycle



場景：會員排序(Scenario : Member Sorting)

- 你任職於某大電信公司行銷部門
- 為了擬定促銷策略
 - 想了解目前800萬會員內，各付費方案人數比例
 - 想調出每月帳單金額超過5000元的名單
 - ...且濾出30歲以上之男生
 - 想促銷iPhone 14

我們以ING Direct為例，虛實整合，無縫接軌，在傳統直效行銷模式遭遇成長瓶頸時，透過多通路訊息之整合，有效地提升獲利、減少成本、縮短時程與深化顧客關係

公司背景

- 全歐洲最大的金融集團之一
- ING個零售銀行主要以荷蘭地區為主，貢獻集團獲利佔28%
- 每年發出的直效行銷郵件達6000萬份以上



面臨的挑戰

- Outbound 方式的行銷活動正面臨著成效逐漸低落的變化，再也無法從這樣的互動中尋求效益。
- 無法透過單一、跨通路的整合讓顧客在與與ING互動中感受到一致性的對待。
- 耗時、費工，有些活動設計需耗費16~22週，甚至動員10多位行銷人員配合方能完成。
- 與顧客的對話安排與設計無法集中，散落在不同的單位上。
- 行銷成本越來越高，效益卻一直在變小。

改善後的效果

- 不斷地創造讓顧客感受到即時、與自身相關的活動體驗，有效地提升了行銷的回應率
 - ✓ 每日透過5個通路(信件、email、網站、客服中心與分行)部署了8千5百萬份名單
 - ✓ 預期將增加2000萬歐元的獲利成長
- 直效行銷成本減少了 35%
- 有效縮減了相關流程(從22週減少為4週即可完成活動部署)

ING 每次顧客與通路進行互動的關鍵時刻，企業將能透過“行銷活動平台”之結果提供最適合顧客的 **產品服務方案 (Next Best Action)**，讓分行行銷戰力提升

預先將各項產品完成分群開發，每一位顧客在每個產品群中，至少會有一項**推薦商品**



Select

當顧客與**數位通路產生互動**時，系統僅會選取適合他的產品清單



Prioritise

每項產品的最終評分將依據**即時的演算法**與預定義之**行銷權重**整合分析後產生，並依據各產品最終評分結果進行順位排序



Filter

各通路僅會呈現提供給顧客最適合之方案清單



DBS透過社群媒體,了解市場需求與服務滿意度

金融業



經常性的蒐集並使用大數據分析技術來自多個社群網站的討論內容，瞭解客戶對該行商品與服務的評價。當有新產品進行促銷活動之際，擴大觀察社群網站上的客戶反應，進行客戶體驗的追蹤，再根據客戶意見調整因應對策。能掌握每次促銷活動的客戶意見，可快速回應客戶聲音，讓DBS在促銷活動的成效上領先競爭者。



“巨量資料”商業應用場景

網購業



產品自動推薦系統: 依照客戶在購物網頁的瀏覽軌跡歷史，來比對『產品』和『產品』之間的關連性，成功推薦出客戶所想要的產品。每三筆訂單中就有一筆是來自產品自動推薦系統，其跨售成績遙遙領先同業。

食品零售業



2009年時市占率突然下滑，委託IBM根據關鍵字調查超過150萬筆的論壇、部落格與網路新聞，發現大量關於鹹酵母醬的內容，及把酵母醬搭配酪梨、烤肉、番茄一起吃討論。於是進行行銷『你怎麼使用酵母醬？』**Vegemite 酵母醬成為熱門話題**，再度的激起人們購買意願，銷量量比高峰期再高出5%。

運輸業



透過每台貨車的無線電通訊和GPS定位，精確傳送車輛位置，並累積行車路徑，**找出最佳行車路線**。

UPS發現十字路口左轉最易發生意外，紅綠燈最浪費時間，只要減少通過十字路口次數，就能省油並提高安全性。因行車順暢，時間少了近一成，這等於一年省下三百萬加侖的油料及減少三萬噸二氧化碳，安全性和效率也提高了。

金融業



發現多數的AE信用卡持卡人都有Facebook帳號，因而通設計了一系列行銷活動，讓持卡人願意將其信用卡綁定其Facebook帳號，以便美國運通獲得用戶授權，可以**追蹤客戶在臉書上的喜好**。當發現持卡人在臉書上感興趣的商家還不是美國運通的合作夥伴時，美國運通馬上派出專員去與之建立行銷合作，然後將促銷折扣資訊推送到持卡人的臉書帳號上。

金融業



餘額寶是一款貨幣基金T+0產品，裡頭的資金等同活期存款，用戶隨時可取用，遇過年或節日用錢高峰，若沒足夠現金，很容易遭到擠兌，而出現流動性風險。但若預留過多的現金，則收益會被拉低。運用大數據可**對每個時點的支付寶動用趨勢做出精準的預測**，讓資金管理的難度降低。即使遇到雙十一光棍節與春節，用戶資金調動最頻繁的時段，餘額寶也能讓流動性需求預測偏離率不超過5%。



「使用者付費」的保險新領域

Allstate(好事達保險公司) · 美國第二大個人保險公司 · 保險範圍含蓋：汽車、房
屋、人壽、健康....等

多樣化線上
投保種類

線上24hr
專屬客服

Your Allstate Agent
Mohammad Y
Bajwa
1720 Columbia Rd NW
Washington, DC 20009
P: (202) 207-7932
(change agent)



使用者付費車險計畫，
人可以透過 GM 集
OnStar 行車資訊技術服
Ford 車廠的 Sync 車用影
統向保險公司傳輸其駕
駕關數據。

2013年 Allstate 表示透
進科技幫助汽車保險定
機制將在10年內佔據70%

穿戴裝置收集健康數據、使用者付費訂價模式

Medibank 澳洲第一大醫療保險公司(聯邦政府所有)

The screenshot shows the Medibank website homepage. At the top, there's a navigation bar with links for About Medibank, Careers, Privacy, HEALTH INSURANCE, OTHER INSURANCE, BETTER HEALTH, and MEMBERS. On the right side of the header, there's a phone icon followed by the number 134 190. Below the header, a large red banner has the text "JOIN AND RECEIVE A FREE FITBIT FLEX™". To the right of this banner is a call-to-action button with three options: "GET A QUOTE", "FIND A STORE", and "CONTACT US". In the bottom left corner of the red banner, there's a small red arrow pointing right with the text "Rewards for getting active". In the bottom right corner of the red banner, there's another smaller text block: "SPECIAL OFFER - Offer extended. Take out hospital and extras cover by Oct 20 and receive a Free Fitbit Flex™ Wireless Activity Tracker. Find out more >".

保險公司與可穿戴廠商合作，以補貼價為保險使用者提供裝置，使用者佩戴裝置收集的健康和運動資料傳輸回保險公司，保險公司據此組態保費上的獎勵或懲罰標準



Data Scientist: *The Sexiest Job of the 21st Century*

**Meet the people who
can coax treasure out of
messy, unstructured data.**

by Thomas H. Davenport
and D.J. Patil

W

hen Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out connections with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, "It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early."

提升資料素養

提升資料素養的六種方法

- 實體課程 (長期/短期) • 參加黑客松
- 線上課程 / 書籍 • 線上資料競賽
- 參與資料社群 • 資料英雄計畫 (長期專案)

在學校就好好的學理論與技術

實體課程 (長期)

- 統計學
- 機器學習
- 資料結構
- 資料探勘
- 平行運算程式設計
- 迴歸分析
- 時間序列分析
- 資料視覺化
- R語言、Python、SQL

短期課程就衝實戰經驗

實體課程 (短期)

DSP 資料爬析 Python 實戰班

DSP Data ETL and Analysis with Python

時間：2015/8/1~8/2~8/8~8/9



2016 DSP 資料科學夏令營@NCCU

DSP Data Camp @NCCU Summer 2016

Student Edition



Designed by Freepik

Credit by 謝宗震 智庫驅動

自學良伴

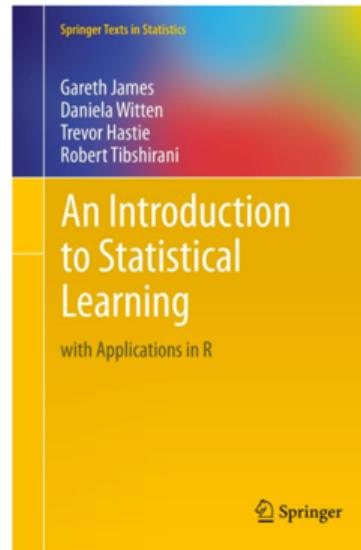
推薦書籍 (科普書)

- 聰明學統計的13又½堂課
(<http://www.books.com.tw/products/0010617019>)
- 統計學，最強的商業武器
(http://www.books.com.tw/products/0010710939?loc=P_asb_004)
- 統計學，最強的商業武器：實踐篇
(http://www.books.com.tw/products/0010687439?loc=P_asb_001)

自學良伴

推薦書籍

- 60本免費的資料科學書籍
(<http://dataology.blogspot.tw/2015/09/60.html>)
- An Introduction to Statistical Learning with Applications in R (2013)



BONUS MATERIAL: Includes multimode CD with over 50 figures and illustrations.

HOW TO MEASURE ANYTHING

Finding the Value of "Intangibles" in Business



2nd Edition
REVISED, EXPANDED,
& SIMPLIFIED



DOUGLAS W. HUBBARD

READ BY DAVID DRUMMOND

**How to Measure Anything:
Finding the Value of Intangibles
in Business** 3rd Edition
by Douglas W. Hubbard

雲端時代 的殺手級應用



天下雜誌出版
作者/胡世忠 定價/360元

雲端時代的殺手級應用 BIG DATA海量資料分析

美、英、日、印、澳等標竿企業與政府應用案例
首度揭露兩岸三地海量資料分析領先者的策略

海量資料分析

自學良伴

線上課程

- 慕課 (<http://course.cool3c.com>)
- R語言翻轉教室 (<http://datascienceandr.org>)
- kaggle Tutorial (<https://www.kaggle.com/wiki/Tutorials>)
- **Launch Your Career in Data Science**
(<https://www.coursera.org/specializations/jhu-data-science>)

期中作業: 主題研究報告

- 1. 由計算思維 / 大數據分析 / 人工智慧 / 人工智慧與SDGs方面之相關議題，任選一主題，至少包含任一參考資料，撰寫主題研究報告，A4 2-5頁，字數不拘，可用中文或英文撰寫。
- 2. 可參考10/31課堂所提供之相關文獻或影片，也可自行尋找參考資料
- 3. 主題研究報告內容須包含以下項目，亦可自行增添。
 - (1) 研究題目
 - (2) 研究動機
 - (3) 資料探討
 - (4) 研究討論（含評述 / 心得 / 感想）
 - (5) 參考資料
 - (6) 附件（關於生成式AI工具的使用過程，若無使用也請說明）
- 4. 使用生成式AI工具注意事項
 - (1) 本作業得適當使用生成式AI工具，使用的過程和結果應在作業中透明地說明和引用，並需於附件載名所使用工具名稱，且詳述使用過程。
 - (2) AI生成內容應被視為工具產生的建議，請經過調整再作為作業內容。
 - (3) 生成式AI工具得作為報告輔助，但學生仍需要進行批判性思考。

Viewpoint | Jeannette M. Wing

Computational Thinking

It represents a universally applicable attitude and skill set everyone, not just computer scientists, would be eager to learn and use.

<https://www.cs.cmu.edu/~15110-s13/Wing06-ct.pdf>

M I N D
 A QUARTERLY REVIEW
 OF
 PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND
 INTELLIGENCE

By A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, ‘Can machines think ?’ This should begin with definitions of the meaning of the terms ‘machine’ and ‘think’. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words ‘machine’ and ‘think’ are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, ‘Can machines think ?’ is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the ‘imitation game’. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either ‘X is A and Y is B’ or ‘X is B and Y is A’. The interrogator is allowed to put questions to A and B thus :

C : Will X please tell me the length of his or her hair ?
 Now suppose X is actually A, then A must answer. It is A’s
 28 433

<https://academic.oup.com/mind/article/LIX/236/433/986238?login=true>

This content downloaded from 140.119.162.175 on Tue, 23 Feb 2016 09:55:42 UTC
 All use subject to [JSTOR Terms and Conditions](#)

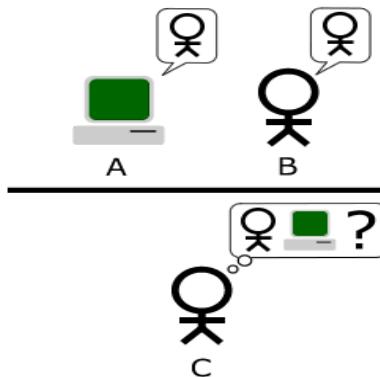
《計算機器和智慧》的論文

提問「機器會思考嗎？」(*Can Machines Think?*)

圖靈測試(Turing Test)：如果一台機器與人類對話而不被辨別出其機器身分，那麼稱這台機器具有智慧。

利用電腦模仿人類交談的遊戲，來判斷機器是否像人一樣有思考能力。說明「思考的機器」是可能的。

Turing test



Figures adapted from Saygin, 2000

[nature](#) > [news feature](#) > [article](#)

NEWS FEATURE | 25 July 2023

ChatGPT broke the Turing test – the race is on for new ways to assess AI

Large language models mimic human chatter, but scientists disagree on their ability to reason.

<https://www.nature.com/articles/d41586-023-02361-7>

LETTERS

Detecting influenza epidemics using search engine query data

Jeremy Ginsberg¹, Matthew H. Mohebbi¹, Rajan S. Patel¹, Lynnette Brammer², Mark S. Smolinski¹ & Larry Brilliant¹

● influenza-like illness (ILI)

- $I(t)$: the percentage of ILI physician visits at time t
- $Q(t)$: the ILI-related query fraction at time t
- $\text{logit}(I(t)) = \alpha \text{logit}(Q(t)) + \epsilon$
 - the 50 million candidate queries
- <https://www.google.org/flutrends/about/>

從谷歌流感趨勢談大數據分析的光榮與陷阱

- Credit by <https://news.cnyes.com/news/id/336695>
- 2014年, lazer等學者在《科學》發文報告了gft近年的表現
 - 2009年 : gft沒有能預測到非季節性流感a-h1n1
 - 2011年8月到2013年8月的108周里 : gft有100周高估了cdc報告的流感發病率。高估有多高呢 ?
 - 2011-2012季 : gft預測的發病率是cdc報告值的1.5倍多
 - 2012-2013季 : gft流感發病率已經是cdc報告值的雙倍多了
- Lazer, D., Kennedy, R., King, G. & Vespignani, A. Big data. The parable of Google Flu: traps in big data analysis. *Science* 343, 1203–5 (2014).
<http://science.sciencemag.org/content/343/6176/1203>

主辦單位
永齡基金會
YongLin Foundation

執行單位
天下雜誌

AI 大師論壇：*Redefining Tomorrow : The AI Revolution*

人工智慧如何形塑人類未來

9/25 (Mon) 14:30-17:00
南港展覽館二館
臺北市南港區經貿二路2號7樓



爐邊對談座長
林奇宏
國立陽明交通大學校長



主講人

李傑

馬里蘭大學講座教授
製造業人工智慧中心主任
世界經濟論壇全球未來製造委員會委員



主講人

吳恩達

AI Fund管理合夥人
Landing AI / DeepLearning.AI 創辦人
Coursera 創辦人



重量級演講者

Sam Altman

OpenAI 執行長



▶ ▶ ⏪ 0:07 / 2:27:30

▶ ⏸ ⏹ ⏺ ⏻ ⏽

<https://www.youtube.com/watch?v=MTjb1ps8iVs>

Perspective | [Open access](#) | [Published: 13 January 2020](#)

The role of artificial intelligence in achieving the Sustainable Development Goals

[Ricardo Vinuesa](#)  , [Hossein Azizpour](#), [Iolanda Leite](#), [Madeline Balaam](#), [Virginia Dignum](#), [Sami Domisch](#),
[Anna Felländer](#), [Simone Daniela Langhans](#), [Max Tegmark](#) & [Francesco Fuso Nerini](#) 

[Nature Communications](#) **11**, Article number: 233 (2020) | [Cite this article](#)

327k Accesses | **673** Citations | **830** Altmetric | [Metrics](#)

<https://www.nature.com/articles/s41467-019-14108-y>

課程進度

單元	時間	標題
1	09/26 第三週	程式桌遊(海霸)
2	10/3 第四週	micro:bit 1
3	10/17 第六週	micro:bit 2
4	10/31 第八週	RapidMiner 1
5	11/14 第十週	RapidMiner 2
6	11/28 第十二週	RapidMiner 3
7	12/12 第十四週	期末報告討論

週次 課程主題	課程內容與指定閱讀	
1 計算思維簡介	社會情境脈絡與未來發展	
	書目：1, 2, 3	
2 計算思維	基本內涵與核心概念	
	書目：1, 2, 3	
3 功能模組	問題拆解與型態辨認	
	書目：4, 5, 6	
4 功能模組	抽象思考與演算邏輯	
	書目：4, 5, 6	
5 國慶日	國定假日	
6 類比至數位轉換 & 電腦運算架構	類比與數位訊號的基礎概念及類比轉換至數位訊號的原理 & 書目：7 chapter 1 & 4 & 5	電腦組成元件與其運算架構
7 大數據應用	大數據中資料科學的基礎分析概念與商業相關應用	
8 學習成果測試	期中評量/作業活動	
9 運算思維測驗	國際運算思維挑戰賽	
10 人工智慧發展	人工智慧發展歷程與未來趨勢	
	人工智慧各式技術與應用案例	
11 人工智慧技術與應用	書目：8	
12 人工智慧應用場景	人工智慧跨域應用	
13 人工智慧學習模型實作	Nocode AI 練習 – Rapidminer 書目：9	
14 人工智慧倫理	生成AI (如：ChatGPT、Deepfake、Midjourney)、假新聞及未來人工智慧應用上的倫理問題	
15 人工智慧專題	海報展示	
16 計算思維與人工智慧	期末報告	
17 彈性補充教學	人工智慧相關競賽經驗交流	
18 彈性補充教學	校園人工智慧應用發想	



最新消息

NO. 1

112學年度第一學期挑戰賽 重要日程

2023/07/31

報名期間：10月2日至10月27日止（同時開放新加入教師註冊；教師登入後，點選「管理班級」即可為班級學生報名）

練習時間：11月6日至11月10日止（已報名學生可登入練習）

挑戰期間：11月13日至11月24日止（教師可於報名時，選擇挑戰期間內的一節課實施）

成績與參與證明公佈時間：12月22日

關於帳號、報名等問題，請詳見 [FAQ](#) (常見問題)

* 112學年度第一、二學期使用之題目相同。

* 系統分流作業將於10月30日至11月3日進行，屆時網站將暫停服務；如造成不便敬請見諒。

年級：

五、六年級 七、八年級 九、十年級 十一、十二年級 大專院校

選擇挑戰賽：

112-1國際運算思維挑戰賽大專院校組

<https://bebras.csie.ntnu.edu.tw/>

挑戰賽介紹

國際運算思維挑戰賽 (International Challenge on Informatics and Computational Thinking，簡稱 Bebras Challenge) 自 2004 年開始，每年於 11 月的國際 Bebras 週 (World-Wide Bebras Week) 在全球各國同步舉行。Bebas Challenge 採用淺顯易懂且生活化的情境式任務，參與學生需運用抽象化、演算法設計、問題拆解、模式辨識、樣式一般化、自動化等運算思維 (Computational Thinking) 來解決問題挑戰。Bebas Challenge 可以讓任課教師了解學生的運算思維知能，發掘具備資訊科學性向的學生，亦希望透過問題思考過程激起學生對資訊科學的學習興趣。

挑戰賽目標



激發學生對資訊科學之學習興趣

Bebas Challenge 藉由情境式的任務，在挑戰的問題中融入資訊科學基本概念；目的是讓學生了解生活中隨處可見資訊科學概念的運用，認識相關概念具廣泛的應用性，進而激發學生對資訊科學的學習興趣。



提升學生運用運算思維解決問題之能力

Bebas Challenge 的任務以家庭生活、團體合作、工作安排等生活情境，引導學生思考進而解決問題。解題過程運用的是運算思維及問題解決能力，學生僅需該年齡層的基本知識即可作答。從任務敘述中推理出問題重點及解題方向，亦可引發學生高層次思考，提升運用運算思維解決問題的能力。



降低學生對資訊科學之恐懼

Bebas Challenge 將抽象的資訊科學知識具體化，以日常生活中會遇到的情境或故事呈現，題組內容有趣生動，有助於降低學生對資訊科學的學習焦慮。未曾受過資訊科學正式課程的學生亦能運用邏輯、歸納、推理、運算等能力進行解題，讓學生對資訊科學的學習具有信心。

Q：在正式挑戰前，學生有沒有練習的機會？

—

A

請教師們於練習週時，除帶領學生熟悉操作界面（是否能正確登入並開始進入挑戰），亦請協助學生應以正確方式回答挑戰問題，包括遇到填充題時，如何正確地以「半形」英數字輸入，並使學生能識別全 / 半形的差異。

練習期間內，學生若有需要皆可登入一次熟悉操作界面，若無需要可以不用進行。
練習完全不記分，題目採自「牛刀小試」題庫。

練習流程完全與挑戰流程相同，唯於挑戰說明時，會註明為練習賽（如下圖）。



挑戰注意事項

以下挑戰注意事項請用心閱讀，以免挑戰權益受損：

1. 總共 **3** 題。
2. 總計作答時間為 **40** 分鐘。
3. 挑戰時右上角為剩下的作答時間。
3.1. 點擊 **放棄作答此題** 後即**無法修改**。
4. 答對給分、答錯扣分，放棄作答則不給分亦不扣分。
5. 挑戰時右上角為剩下的作答時間。
6. 目前為 **練習賽** 模式。

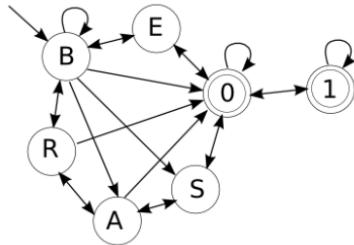
> 計時開始

<點選即開始計時挑戰

木筏牌照

海狸的交通工具是木筏，每艘木筏都有一個獨一無二的牌照號碼。

合法的牌照號碼必須依據下圖規則，按照邊的方向(有單向 → 跟雙向 ←→)走訪節點；依照走訪的順序，把節點上的英文字母和數字組成牌照號碼。
而且每個牌照號碼都由字母B開始，並以數字0或1結束。

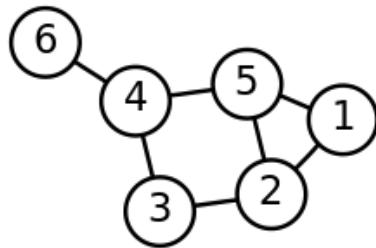


請問下列哪一個不是合法的木筏牌照號碼？

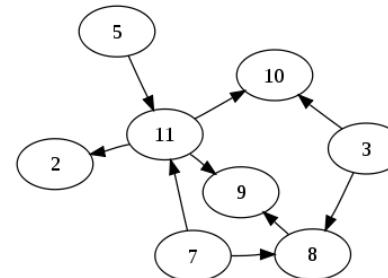
- BB0001 和 BBB011
- BBB100 和 BR00A0
- BB0100 和 BSA001
- BE0S01 和 BE01

什麼是圖(Graph)？

Ans: 圖是點和邊所構成的集合，分為無向圖和有向圖

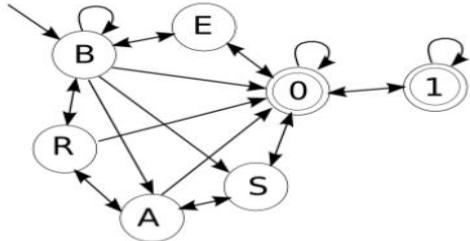


無向圖



有向圖

木筏牌照



請問下列哪一個不是合法的木筏牌照號碼？

- BBB001 和 BBB011
- BBB100 和 BR00AO
- BB0100 和 BSA001
- BE0S01 和 BE01

說明:

BBB100 · B無法直接到1
BR00AO · O無法直接到A

關鍵字: 有向圖、有限狀態機

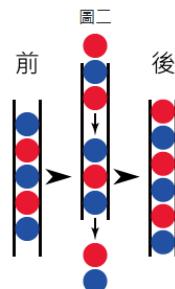
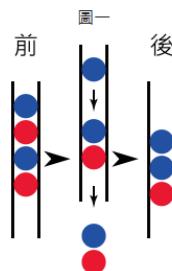
紅藍彈珠遊戲

海狸弟弟在海狸一日營學到了一個新遊戲。

首先他抓起一把紅色及藍色的彈珠——放進一個透明的塑膠管中，每顆彈珠都從塑膠管的上頭放入。

接下來每次從塑膠管下頭取出兩顆彈珠，若第一顆彈珠是紅色，則再從上頭放入一顆藍色的彈珠（如圖一所示）；若第一顆取出的彈珠是藍色，則從上頭依序放入紅色、藍色、紅色三顆彈珠（如圖二所示）。

此「取出 + 加入」彈珠程序持續到塑膠管內只剩下兩顆或一顆彈珠為止。

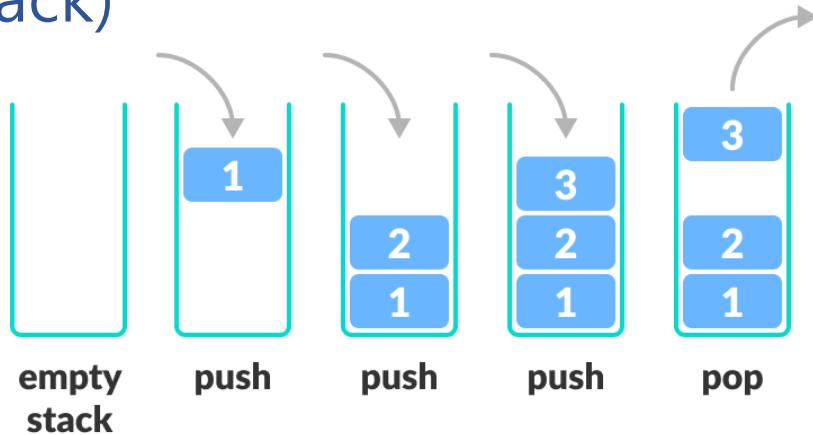


若海狸弟弟一開始依序放入了紅、藍、藍、紅、紅等五顆彈珠如下圖所示，總共需要幾次的「取出 + 加入」程序，才會使得塑膠管內的彈珠剩下一或二顆彈珠？

僅填入數字，只能輸入半形數字才會正確



堆疊(Stack)



先進後出(First In Last Out)

<https://www.programiz.com/dsa/stack>

佇列(Queue)

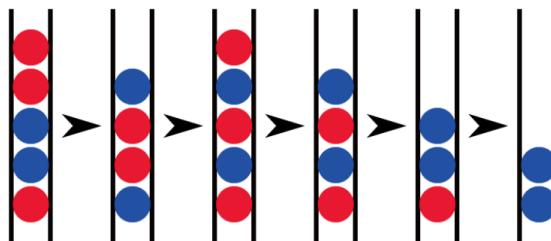
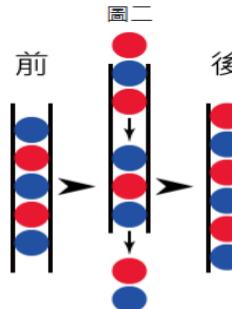
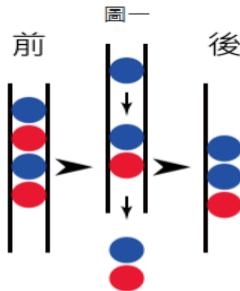


先進先出(First In First Out)

<https://www.programiz.com/dsa/queue>



AI 中心
IAIC



聊天順序

娜娜邀請了五位好朋友參加她的生日派對，五位朋友分別是：小愛、奇哥、小蘿、戴哥、與小萱。生日派對上娜娜出了一個謎題如下：

- 我任何時候都可跟小萱聊天。
- 我必須先與小愛聊天後才能與戴哥聊天。
- 我必須先與小萱聊天後才能與奇哥聊天。
- 我必須先與戴哥及奇哥聊天後才能與小蘿聊天。
- 我必須先與奇哥及小萱聊天後才能與小愛聊天。

娜娜與五位朋友的聊天順序應為何？

(請依聊天順序連續填入名字，不要留空白不要加入標點符號)

請在此作答...

聊天順序

- 這個問題就像是修課順序一樣，課與課可能有擋修的問題，有些課必須先修，才能修另一堂。
- 這就是圖論中的 “拓撲排序” (topological sort)

關鍵字: 有向圖、拓撲排序、有向無環圖(DAG)

[拓撲排序 - 維基百科，自由的百科全書 \(wikipedia.org\)](#)

75

[有向無環圖 - 維基百科，自由的百科全書 \(wikipedia.org\)](#)

什麼是拓撲排序？

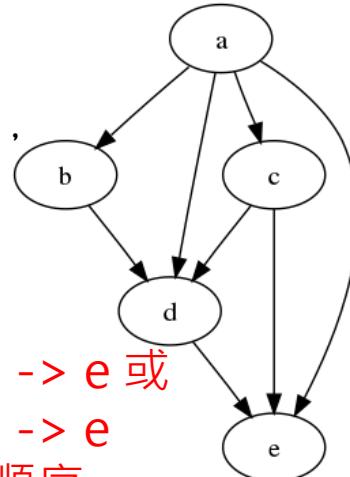
- 在某校的選課系統中，存在這樣的規則：每門課可能有若干門先修課，如果要修讀某一門課，則必須要先修讀此課程所要求的先修課後才能修讀。假設一個學生同時只能修讀一門課程，那麼，被選課系統允許的他修完他需要所有課程的順序是一個拓撲排序

[拓撲排序 - 維基百科，自由的百科全書 \(wikipedia.org\)](#)

什麼是拓撲排序？

舉例來說，右圖是修課的流程圖， a, b, c, d, e 代表科目，每一種都必須修習， a 指向 b 代表 a 是 b 的先修科目。

科目	先修科目
a	無
b	a
c	a
d	a b c
e	a c d

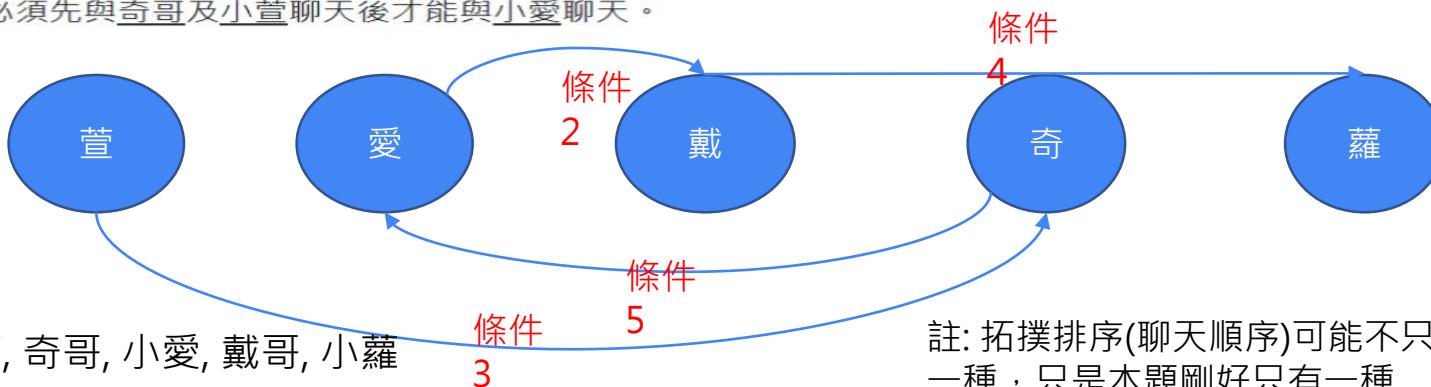


$a \rightarrow b \rightarrow c \rightarrow d \rightarrow e$ 或
 $a \rightarrow c \rightarrow b \rightarrow d \rightarrow e$
皆為合法的修課順序，
即為圖的拓撲排序！

拓撲排序 - 維基百科，自由的百科全書 (wikipedia.org)

聊天順序

- (條件1) • 我任何時候都可跟小萱聊天。
- (條件2) • 我必須先與小愛聊天後才能與戴哥聊天。
- (條件3) • 我必須先與小萱聊天後才能與奇哥聊天。
- (條件4) • 我必須先與戴哥及奇哥聊天後才能與小蘿聊天。
- (條件5) • 我必須先與奇哥及小萱聊天後才能與小愛聊天。



算式化簡

小海狸有一台很特別的電腦，它提供兩種指令來計算海狸世界的數學式子，這兩種指令的使用說明如下：

- R 指令 當 f 是一個數學運算子，可以是 $+$ 、 $-$ 、 \times 或 \div ；而 X_1, X_2, \dots, X_n 代表一串數字，那麼：
 $(R f (X_1, X_2, \dots, X_n))$ 指令將會計算 $X_1 f X_2 f \dots f X_n$

舉例來說： $(R + (1, 2, 3, 4))$ 將會計算 $1 + 2 + 3 + 4$ ，而結果為 10。

- M 指令 當 f 是一個函數，而 X_1, X_2, \dots, X_n 代表一串數字，那麼：
 $(M f (X_1, X_2, \dots, X_n))$ 指令將會計算 $f(X_1), f(X_2), \dots$ 及 $f(X_n)$ ，並得到計算結果所組成的數列。

舉例來說：當函數 $q(x) = -x$ ，那麼 $(M q (1, 2, 3, 4))$ 將會把數列 $(1, 2, 3, 4)$ 帶入 $q(x)$ ，得到結果為 $(-1, -2, -3, -4)$ 。

現在假設 $t(X) = 3X + 2$ 且 $q(X) = -X$ 。請問下列式子會得到什麼結果？

$$(R + ((R + (M t (0, 2, 4))), (R + (M q (M t (3, 5))))))$$

- 7
- 0
- 7
- 4



● R 指令	當 f 是一個數學運算子，可以是 $+$ 、 $-$ 、 \times 或 \div ；而 X_1, X_2, \dots, X_n 代表一串數字，那麼： $(R f (X_1, X_2, \dots, X_n))$ 指令將會計算 $X_1 f X_2 f \dots f X_n$ 舉例來說： $(R + (1, 2, 3, 4))$ 將會計算 $1 + 2 + 3 + 4$ ，而結果為 10。
● M 指令	當 f 是一個函數，而 X_1, X_2, \dots, X_n 代表一串數字，那麼： $(M f (X_1, X_2, \dots, X_n))$ 指令將會計算 $f(X_1), f(X_2), \dots$ 及 $f(X_n)$ ，並得到計算結果所組成的數列。 舉例來說：當函數 $q(x) = -x$ ，那麼 $(M q (1, 2, 3, 4))$ 將會把數列 $(1, 2, 3, 4)$ 帶入 $q(x)$ ，得到結果為 $(-1, -2, -3, -4)$ 。

- 所以，R指令就是Reduce運算, M指令就是Map運算,
- Reduce運算是個純量函數(輸出是個純量)
- Map運算是向量函數(輸出是個向量)
- 舉例來說:
- (1) $R + (a, b, c, d) = (((a + b) + c) + d)$
- (2) $M (2x+3) (a, b, c) = (2a+3, 2b+3, 2c+3)$

關鍵字: Map&Reduce, 算式化簡,
MapReduce - Wikipedia



現在假設 $t(X) = 3X + 2$ 且 $q(X) = -X$. 請問下列式子會得到什麼結果？

$$(R + ((R + (\underline{M t (0, 2, 4)})), (R + (\underline{M q (M t (3, 5))}))))$$

- 7
- 0
- 7
- 4

根據以上的說明，我們可以來化簡式子：

$$\underline{M t (0, 2, 4)} = M (3x+2) (0, 2, 4) = (2, 8, 14)$$

$$\underline{R + (2, 8, 14)} = 24$$

$$\underline{M t (3, 5)} = M (3x+2) (11, 17)$$

$$\underline{M q (11, 17)} = M (-x) (11, 17) = (-11, -17)$$

$$\underline{R + (-11, -17)} = -28$$

$$R + (24, -28) = -4$$

政治大學新創之星競賽



競賽資訊頁面：

[新創之星競賽頁面](#)

第八屆台積電青年築夢計畫



競賽資訊頁面：
[2023第八屆台積電青年築夢計畫](#)

2023永續智慧創新黑客松



2023永續智慧創新 黑客松競賽

競賽資訊頁面：

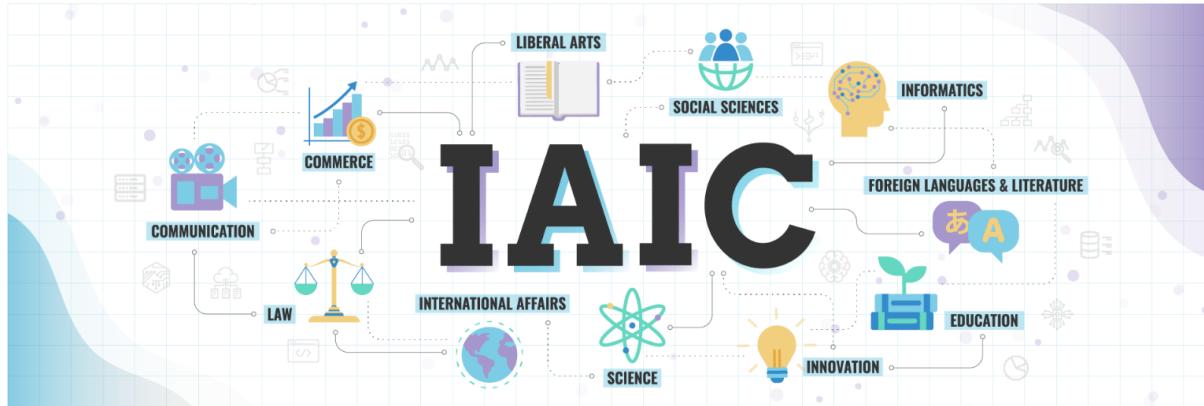
[2023永續智慧創新黑客松官網](#)

[報名連結](#)

課程進度

單元	時間	標題
1	09/26 第三週	程式桌遊(海霸)
2	10/3 第四週	micro:bit 1
3	10/17 第六週	micro:bit 2
4	10/31 第八週	RapidMiner 1
5	11/14 第十週	RapidMiner 2
6	11/28 第十二週	RapidMiner 3
7	12/12 第十四週	期末報告討論

週次 課程主題	課程內容與指定閱讀	
1 計算思維簡介	社會情境脈絡與未來發展	
	書目：1, 2, 3	
2 計算思維	基本內涵與核心概念	
	書目：1, 2, 3	
3 功能模組	問題拆解與型態辨認	
	書目：4, 5, 6	
4 功能模組	抽象思考與演算邏輯	
	書目：4, 5, 6	
5 國慶日	國定假日	
6 類比至數位轉換 & 電腦運算架構	類比與數位訊號的基礎概念及類比轉換至數位訊號的原理 & 電腦組成元件與其運算架構	書目：7 chapter 1 & 4 & 5
7 大數據應用	大數據中資料科學的基礎分析概念與商業相關應用	
8 學習成果測試	期中評量/作業活動	
9 運算思維測驗	國際運算思維挑戰賽	
10 人工智慧發展	人工智慧發展歷程與未來趨勢	
	人工智慧各式技術與應用案例	
	書目：8	
11 人工智慧技術與應用		
12 人工智慧應用場景	人工智慧跨域應用	
13 人工智慧學習模型實作	Nocode AI 練習 – Rapidminer	書目：9
14 人工智慧倫理	生成AI (如：ChatGPT、Deepfake、Midjourney)、假新聞及未來人工智慧應用上的倫理問題	
15 人工智慧專題	海報展示	
16 計算思維與人工智慧	期末報告	
17 彈性補充教學	人工智慧相關競賽經驗交流	
18 彈性補充教學	校園人工智慧應用發想	



研究合作 跨域教學 多元服務



Dr. Chih-Hsun Wu

吳致勳 助理教授

20031214@nccu.edu.tw

j20031214@gmail.com