

Introduction to Database Management

楊立偉教授

台灣大學工商管理學系

楊立偉老師

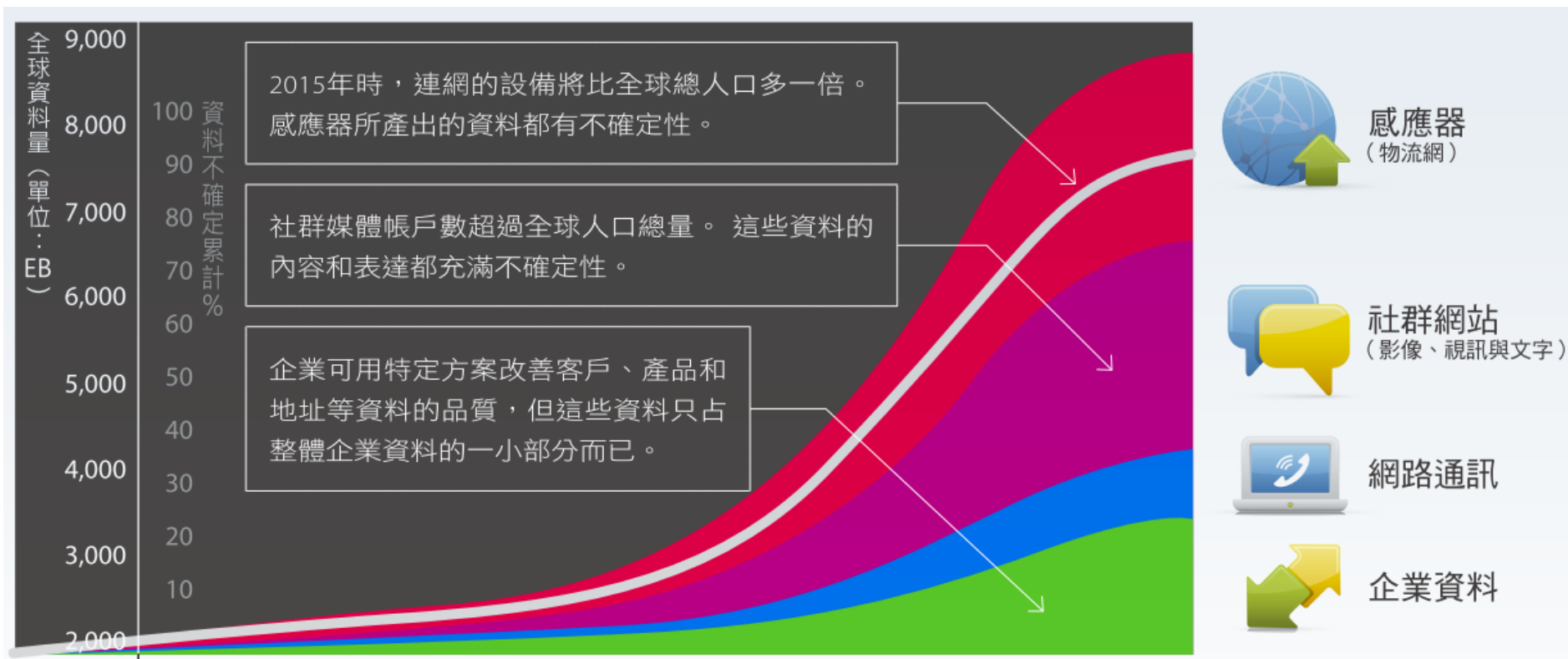
研究領域為資料庫、搜尋引擎、語意分析、大數據及商業分析、數位行銷

現任	台大工管系暨商研所兼任助理教授 2006~
	台大資管系兼任助理教授 2017~
	資訊及通信國家標準技術委員
	意藍資訊 董事總經理（創辦人） 國內規模最大的網路情報與社群口碑自動分析平台 企業搜尋引擎市佔率最高；國際檢索競賽第一名
經歷	獲選100 MVP最有價值經理人，擁有超過20項語意分析專利
	獲國家雲端創新獎、數位時代「創業之星」首獎，台大教學傑出教師
	獲數位時代雜誌選為代表台灣軟體的封面人物之一

課程目標

- 了解資料庫管理之目的與理論
- 培養資料庫管理之觀念與能力
- 了解資料庫管理之工具與應用
- 了解資料庫管理和組織、企業等相關之管理議題

Why Database (1) – 資料快速成長



Why Database (2)

- 用資料與演算，輔助決策，建立行動
- 需要處理大量、快速、連結及分析多種資料的能力

Data 數據

- Structured 數值
- Unstructured 非數值
- Historic 過往

Information 資訊 & Insights 洞見

- Modeling 模型
- Deduction 演繹
- Inference 推理
- Prediction 預測

Decisions 決策 & Actions 行動

- Results 結果
- Options 選項
- Prevention 預防
- Suggestion 建議

- How Target figured out a teen girl was pregnant before her father did.



sale

For coupon delivered to your smartphone, text **INFANT1** to **TARGET** (827438)

Message and do not text or call. Our message or message may contain promotional offers from Target and other companies. Some restrictions may apply. See store for details. ©2012 Target Brands, Inc. All rights reserved. Target and the Bullseye logo are trademarks of Target Brands, Inc. All other trademarks are the property of their respective owners.

\$20 off

Baby purchase of \$100 or more

- Newborn apparel • Furniture • Bedding • Nursery
- Training pants • Toiletries • Diapers • Wipes • Travel gear
- Baby toys • Car seats • Bath • Food and formula
- Infant and toddler feeding.

Excludes infant and toddler apparel and Toy Department

get coupon

\$2.00 OFF

PAMPERS® ONE Pampers® Cruisers Diapers (excludes trial/travel size)

\$2.00 OFF

PAMPERS® ONE Pampers® Baby Dry Diapers (excludes trial/travel size)

free \$10 GiftCard when you buy 2 diapers listed*
24.99

Baby care Pampers® Huggies super pack or Luvs value pack diapers, or Huggies or Pampers super pack training pants.

*Limit one per household. See store for details. ©2012 Target Brands, Inc. All rights reserved. Target and the Bullseye logo are trademarks of Target Brands, Inc. All other trademarks are the property of their respective owners.

Stock up for baby

\$15 GiftCard **FREE \$15 TARGET GIFTCARD™**

with select baby department purchase of \$75 or more

- Wipes • Diapers • Baby food • Infant formula
- Baby toiletries • Baby furniture

Restrictions apply. See store for details. ©2012 Target Brands, Inc. All rights reserved. Target and the Bullseye logo are trademarks of Target Brands, Inc. All other trademarks are the property of their respective owners.

Find this coupon in the Target app. Scan in wallet at checkout.

Why Database (3) – 管理整體資訊供應鍊



Source : IBM全球CEO調查報告
<https://www-07.ibm.com/tw/blueview/2012oct/8.html>

Why Database (4) – 企業應用

- 數位化時代中，資訊扮演的角色越來越重要
- 所有企業資訊化的背後，幾乎都與資料庫有關。
 - 生產 / 研發 / 供應商
 - 行銷 / 銷售 / 客戶
 - 組織 / 人事
 - 財會 / 資料分析
 - 策略 / 決策支援

Scenario 1 : 資料收集、處理、分析

您的聯絡資料			
姓名		公司名稱	
部門		職稱	
聯絡電話	1.		2.
e-mail			

01. 貴公司/單位目前員工人數：

50 人以下 50-100 人 100 人以上

02. 請問您參加本次研討會的動機為？（可複選）

對研討會議題感興趣 公司指派參加

朋友/同事推薦 與目前工作相關

- 假設回收50份
- 第一題各選項的答題人數？
- 第二題各選項的答題人數？
- 列出500人以上且為公司指派的對象
- ...且連絡電話開頭為02者
- 假設今天回收是百萬份...

Scenario 2 : 資料庫行銷

- 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2000元的名單
 - ...且濾出30~40歲以上之男生
 - 想促銷最新一款iPhone
- 面臨的問題
 - 向IT部門要資料卻要2個月後才給
 - 跟你說沒這個欄位...做不到...

Scenario 3 : 供應商管理

- 你任職於一家電子組裝公司
- 因應綠色環保法規，你必需將供應商分類
(合格 / 不合格 / 待檢驗) 以重新整理零件存貨
- 為了方便作業，你請IT部門做系統調整
 - IT部門說如果是加欄位則 2 天就可以做完
 - ...但如果是更改資料庫結構，就要很久
 - 你心想「差在哪裏？」

Scenario 4 : 資料更新問題

- 你任職為零售賣場的分析師
- 你將POS(銷售點)系統產生的紀錄全留下

部門	群組	產品線	產品名稱	通路	分店	電話	售價	數量	銷售日期	方式
家電	大型電器	烤爐	Gas Oven	大潤發	台中	04-2777-4333	34400	2	1/1/2000	Cash
藥妝	化妝品	唇膏	Lip Gloss	大潤發	新竹	03-32212020	240	2	1/1/2000	Visa
藥妝	化妝品	唇膏	Lip Gloss	大潤發	新竹	03-32212020	240	2	1/1/2000	Visa
藥妝	化妝品	唇膏	Lip Liner	大潤發	高雄	07-8488-4333	200	2	1/1/2000	Cash
藥妝	化妝品	基礎品	Powder Eye Shadow	大潤發	新竹	03-32212020	840	2	1/1/2000	Visa
藥妝	化妝品	眼線用品	Waterproof Eye Liner	大潤發	台中	04-2777-4333	320	2	1/1/2000	Visa
家電	小家電	咖啡機	10 Cup Coffee Maker	大潤發	台北	02-2366-0777	920	1	1/10/2001	Visa

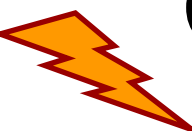

- 某日得知新竹店電話改了
- 請問要修改幾筆資料? 容易有什麼問題?

Scenario 5 : 資料完整性與安全性

- 你有兩個銀行戶頭

	A 銀行	B 銀行
初始餘額	\$1,000	\$1,000
存\$3,000到A銀行	\$4,000	\$1,000
轉\$2,000到B銀行		
(1) A銀行先扣\$2,000	\$2,000	\$1,000
發生大地震~全台大停電	-----	
(2) B銀行再加\$2,000	\$2,000	\$3,000

\$2,000 為何會這樣?
如何解決?



Scenario 6 : 更多資料庫的管理問題

- 資料庫整合

- 你任職於某大金控公司客服部門
- 公司最近合併了其他三間銀行、保險、證券公司
- 客戶打電話來問「為什麼會收到三張帳單?」
- 保險的客戶打電話來，你卻調不出客戶資料

- 資料庫安全

- 你任職於某大電視購物公司
- 聽聞在暗網有盜賣公司會員資料，一筆3元
- 好奇下載來看，結果是資料庫格式，卻不會用?

學習目的

- 這是一門介紹基礎理論的課
 - 資料庫背後的理論
 - 管理資料庫應具備的觀念
 - 資料庫可能衍生的相關議題
- 這是一門著重應用與實務的課
 - 如何建立與使用資料庫 → 以MySQL (或MariaDB) 為例
 - 靈活運用資料並進行查詢分析 Ex. 問卷分析
 - 了解各種資料庫最新應用，如 Big Data、Data Science 等

授課重點

- Database Concepts
- Entity-Relationship Model
- Relational Database Design
- Query Language and SQL
- The Client/Server Database Environment
 - Data Concurrency and Security *new*
- Data Warehousing
- Data Infrastructure *new*
- Big Data and Analytics

授課與評分方式

- 課堂及線上授課，配合應用解說與討論
- 評分方式
 - 作業報告 30%
 - 期中考 30%
 - 期末考及分組專題(含組內互評序位) 40%
- 課程網址及資源
 - <http://homepage.ntu.edu.tw/~wyang/db2022>
 - NTU cool 線上教學平台
 - 配合使用線上會議教學或錄影觀看

Introduction to Database Management

資料庫發展趨勢

楊立偉教授

台灣大學工商管理學系

Trend of Big Data

- Big Data 係指資料大量成長
- 根據IBM的研究，全世界90%的資料是在過去2年產生
- Google、Facebook 等，就是站在Big Data上的範例
- 巨大的數據源，將改變整個學術界，商界和政府
- 依賴資料庫工具處理
 - 包括 capture，storage，search，analytics 等

美國政府：Big Data is a Big Deal

白宮啟動「海量資料研究和發展計畫」，包括Big Data分析及在醫療、天氣和國防等領域的運用；白宮並將數據資料定義為「**未來的新石油**」。

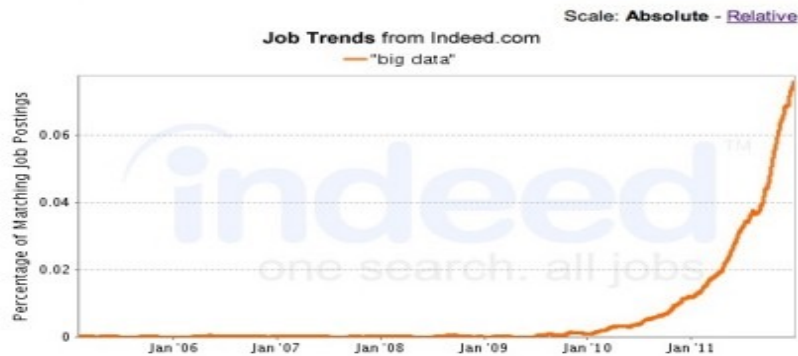
。一個國家擁有數據資料的規模和解釋運用的能力，已成為一國核心資產和國力指標。

The screenshot shows the White House website interface. At the top, it reads "the WHITE HOUSE PRESIDENT BARACK OBAMA" with a logo of the White House. Navigation links include "BLOG", "PHOTOS & VIDEO", "BRIEFING ROOM", "ISSUES", "the ADMINISTRATION", "the WHITE HOUSE", and "our GOVERNMENT". A search bar is present with the text "Search WhiteHouse.gov". The main content area features the "Office of Science and Technology Policy" logo and title. Below this, a navigation bar lists "About OSTP", "Pressroom", "OSTP Blog", "Divisions", "Initiatives", "R&D Budgets", "Resource Library", "NSTC", "PCAST", and "Contact Us". The article title "Big Data is a Big Deal" is prominently displayed, along with a "Subscribe" button. A green button says "GIVE FEEDBACK ABOUT THIS PAGE". Below the article title, it says "Posted by Tom Kalil on March 29, 2012 at 09:23 AM EST". Social media sharing options for "E-Mail", "Tweet", "Share", and a plus sign are visible. At the bottom, an editor's note reads: "[Editor's Note: Watch the live webcast today at 2pm ET of the Big Data Research and Development event at <http://live.science360.gov/bigdata/>"]". On the right side, there is a section for "YOUR FEDERAL TAXPAYER RECEIPT" with a small image of the receipt.

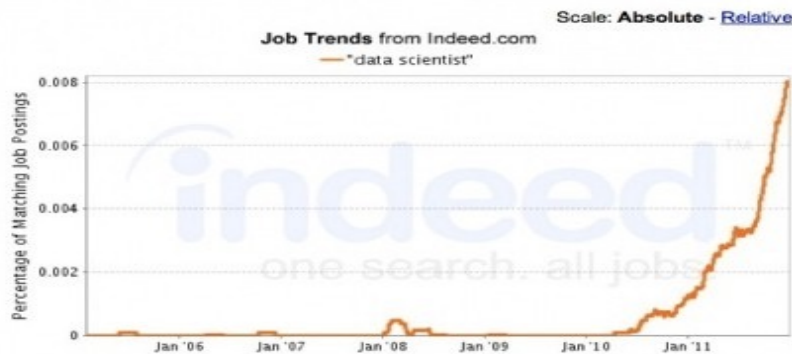
哈佛：資料科學家是21世紀最迷人的工作

- "Data Scientist : The sexist job of the 21st century"

"big data" Job Trends



"data scientist" Job Trends



SPOTLIGHT ON BIG DATA

Spotlight

ARTWORK Tamar Cohen, Andrew J. Rubelitz
2011, silk screen on a page from a high school
yearbook, 8.5" x 12"

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

When 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."

70 Harvard Business Review October 2012

零時政府促成了哪些專案？

多項專案持續集中.....

</>

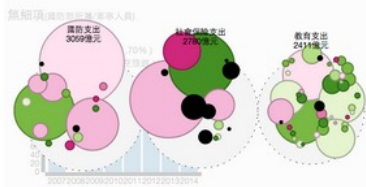
- 開放政府
- 開放資料
- 社會參與
- 新媒體應用
- 政策共筆
- g0v基礎建設

開放資料
開放政府
民間協作

開放政府 Open Government

草根力量打造開放透明資訊平台，促進公民參與及雙向溝通

中央政府總預算視覺化



g0v 社群的第一個專案。將大量難以閱讀的預算資料運用鳥瞰圖、變化圖、國債鐘等互動圖表，依比例呈現中央政府各部門的預算，開放網友評論。

國會大代誌



持續開發的開放國會長期計畫。可即時顯示國會行程、修法進度、法條版本比較，還推播到 Twitter 供人訂閱。除了文字之外還整合了國會影像，標示每一位委員的發言時間供搜尋。


專案網址
ly.g0v.tw

意見回饋

立院影城



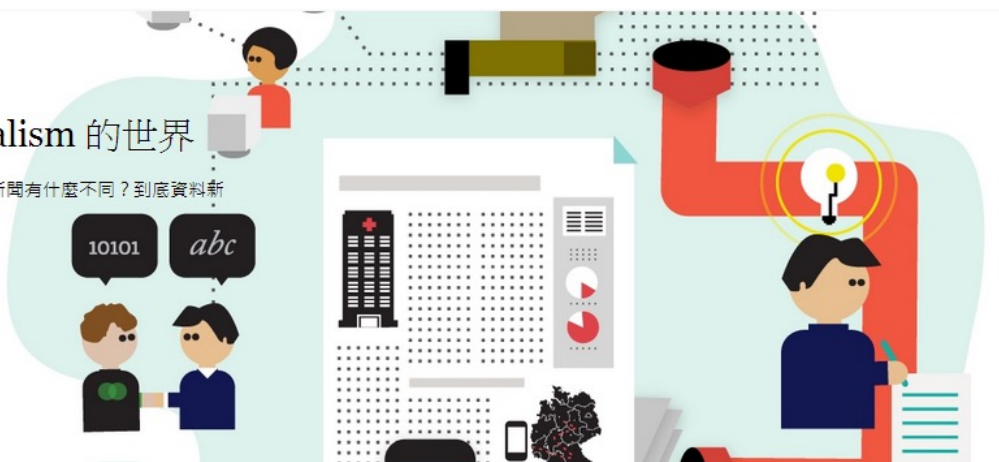
開放國會計畫的一部分。即時播放立法院議事影音直播 ivod，突破立法院原網站 IE only 限制，可跨平台收看並有互動功能。

 OpenData / TW | 開放資料

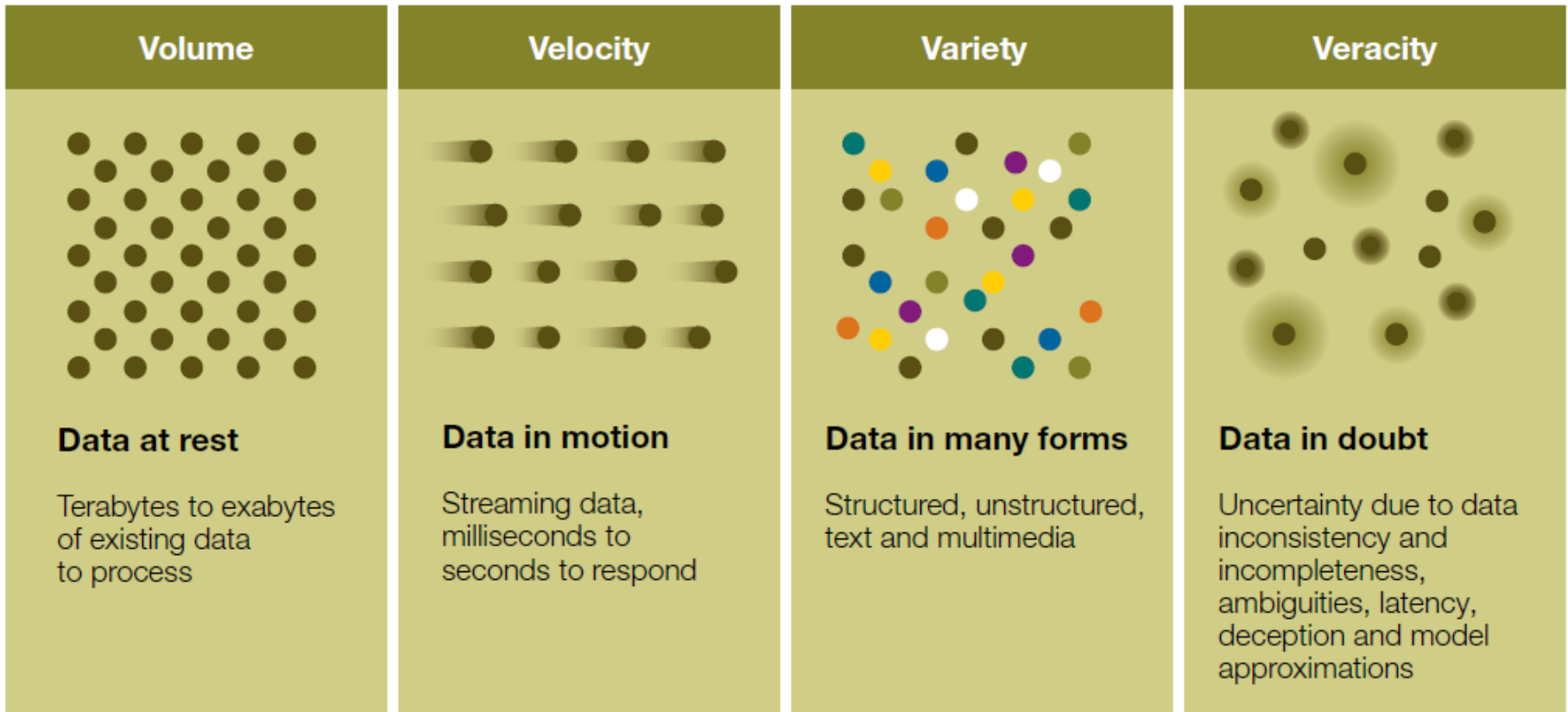
- 首頁
- 文章與專欄
- 新聞與資訊
- OPEN CAMPUS 2013
- 活動資訊
- 關於我們

進入 Data Journalism 的世界

資料新聞學？這和傳統新聞有什麼不同？到底資料新聞學在講什麼？ More



Big Data 四大特性



Database rules the world (1)

- 2021 top 10 richest persons from Forbes



Database rules the world (2)

- about ORACLE

- 繼Microsoft後，全球收入第二多的軟體公司 (超越IBM)
- 1977年在矽谷由3人合資成立，當時IBM發表「關聯資料庫」的論文，Ellison以此造出新資料庫，名為甲骨文
- 首年賺不到百萬美元，10年內迅速冒起
- 2021年市值約2023億美元

- 現況

- 2020年營收391億美元
- 全球有38萬個客戶，1400萬個開發者
- 財星100大企業全使用Oracle的資料庫



Database rules the world (3)

- 產業 #1

- 金融 Banking
- 通信與媒體 Telecom & media
- 建築 Construction
- 製造 Manufacturing
- 生技醫療 Life sciences
and healthcare
- 專業服務 Professional services
- 政府 Public sector
- 零售 Retail
- 運輸 Transportation

- 企業 #1

- 前20大航空公司 airlines
- 前20大汽車公司 automotive
- 前20大政府機構 governments
- 前20大高科技公司 high tech
- 前20大保險公司 insurers
- 前20大製造公司 manufacturers
- 前20大石油公司 oil and gas
- 前20大製藥公司 pharmas
- 前20大水電公司 utilities
- 前20大銀行公司 banks

Growth drivers

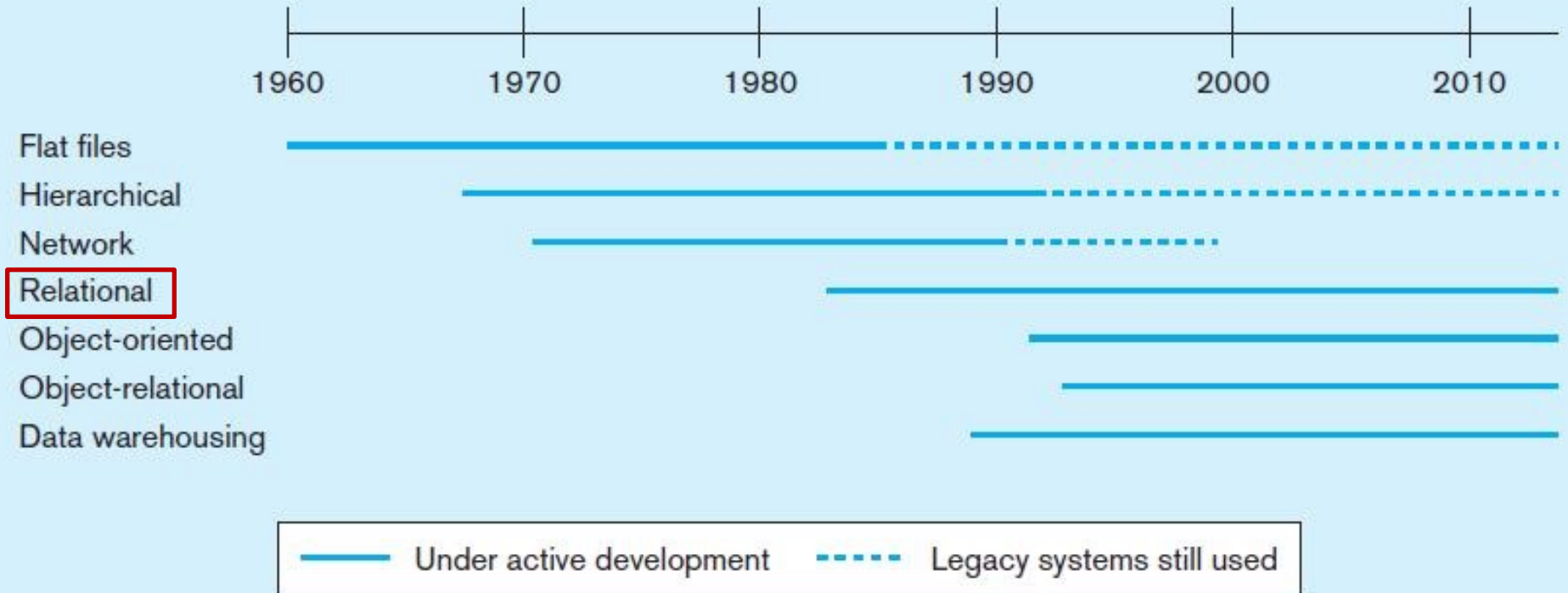
Microsoft : 《The Fourth Paradigm: Data-Intensive Scientific Discovery》

- 科學發展走過「實驗、理論、計算」三個典範，已走入以「資料」為重點的第四典範

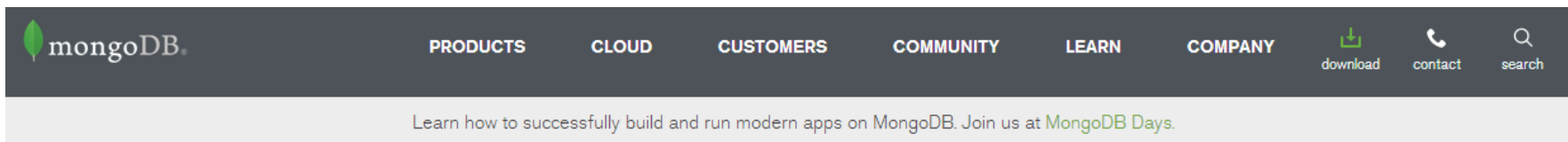
IDC:

- 需管理的資料持續爆炸性成長
- 商業智慧 (business intelligence) 需求持續成長
- 新的資料庫技術創造新的機會

Evolution of Database Technologies



New Challengers Rising



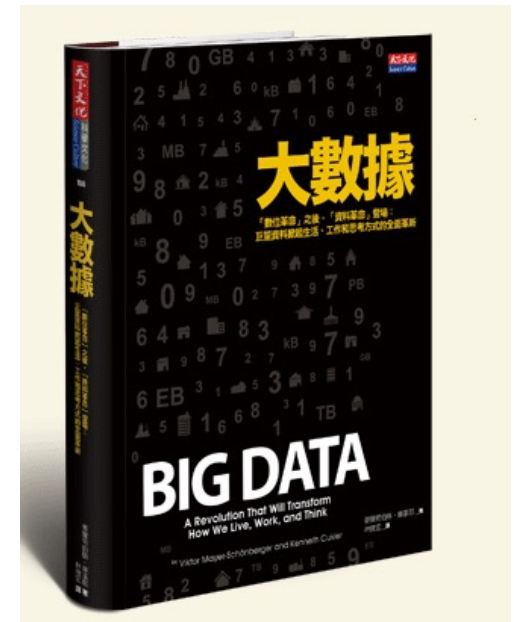
MongoDB Raises \$80 Million to Accelerate Growth

Large Investment Marks Industry Shift Away from Outdated Relational Databases, Reflecting Customer Demand for MongoDB, the DBMS Standard on which Modern Applications are Built

NEW YORK, NY and PALO ALTO, Calif.– January 12, 2015 – [MongoDB](#) announced today that the company has secured an additional \$80 million in funding from prominent institutional investors and investment funds in December. A sovereign wealth fund led the round, with participation from Goldman Sachs and from existing investors Altimeter Capital, NEA, Sequoia and funds managed by T. Rowe Price Associates, Inc. MongoDB has attracted a total of \$311 million from investors, positioning it as the best capitalized next- generation DBMS company in the world.

閱讀作業

- **Big Data : A Revolution that will transform how we live, work, and think**
 - by Viktor Mayer-Schonberger, Kenneth Cukier.
 - 《大數據》·天下文化·2013/2018新版
 - 閱讀第 1~7章



閱讀作業

- 於加選公布後一週內繳交 (詳見ntu cool公告)
 - Word檔，頁數不拘
- 含心得報告及問題回答
 1. 大數據對企業管理的影響為何？企業如何因應？
 2. 個人對於資料庫管理的認識與期望