



沙裏淘金??

第八組

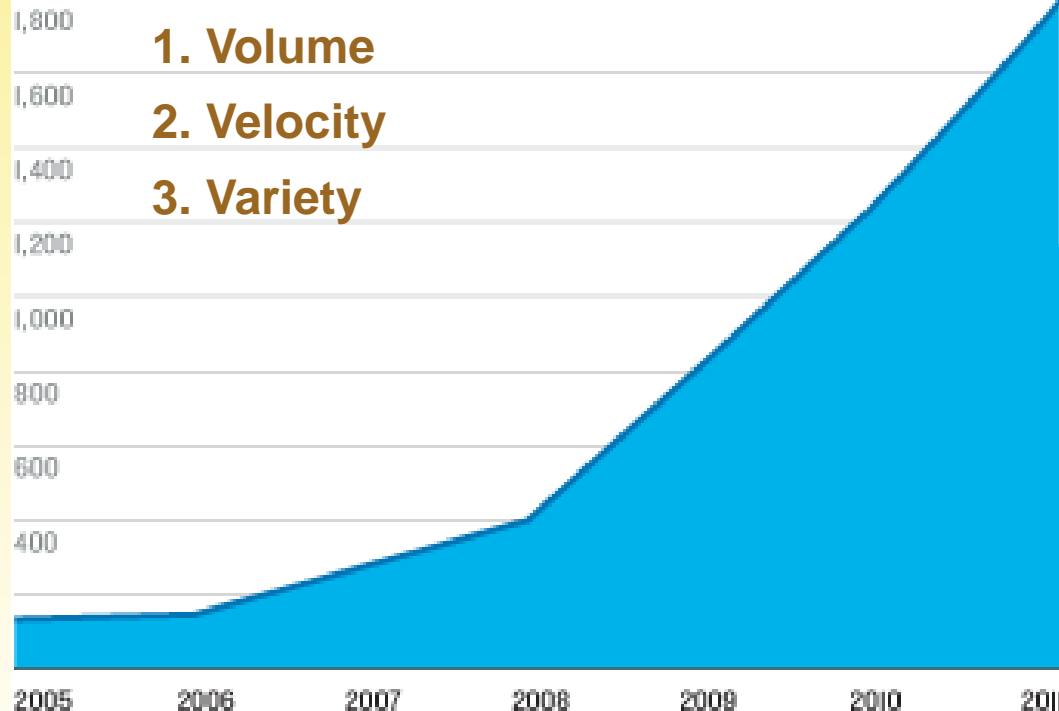
指導教授：柯淑津 教授

組員姓名：郭庭瑄、楊家綺、柳瓊雅、王俊凱
陳玟陵、胡鈺慈、江芷螢、涂凌華

Big Data

Digital Information Created Each Year, Globally

2,000 BILLION GIGABYTES



2,000%

Expected increase in
global data by 2020

**III
Megabytes**

Video and photos stored
by Facebook, per user

75%

Percentage of all digital
data created by consumers

Sources: IDC, Radicati Group, Facebook, TR research, Pew Internet

<http://klcommunications.com/anonymity-in-a-big-data-world>



沙裏淘金？

- 快速抓取重要資料——口碑監測系統
- 資料淨化和展現——隱喻查詢系統



口碑監測系統



我覺得有化妝的妞不知道這罐就太扯了

根本是完美!!

超級無敵好卸眼妝的

也不會刺激眼球

一隻眼睛一張濕的化妝棉就能清除得很乾淨

也不會乾澀乾癢

油水層經過搖晃後也能均勻的分布

也不會太快就分離

時間足夠讓你從容的沾濕化妝棉和各敷一隻眼睛

還有非常有視覺效果的藍色讓使用起來就有很好卸的感覺

(但我怕色素小小扣你一顆心)

最後是小資都關心的CP值和價格問題

我個人覺得它其實買的到超級可愛的價格

而且我用量沒很兇(所以CP爆炸)

這樣溫和又好卸的東西本來就應該不是白菜錢

所以原價有點***

但由於他3罐299(現在價格已經沒有這麼優惠了!!)

跟姊姊和朋友一起買就超便宜的!!

一瓶不用100你還想怎樣!!

囤了就是你賺到!!



言簡意賅



一目瞭然



條理分明



WHY ?

明確清楚

省時

重要資訊

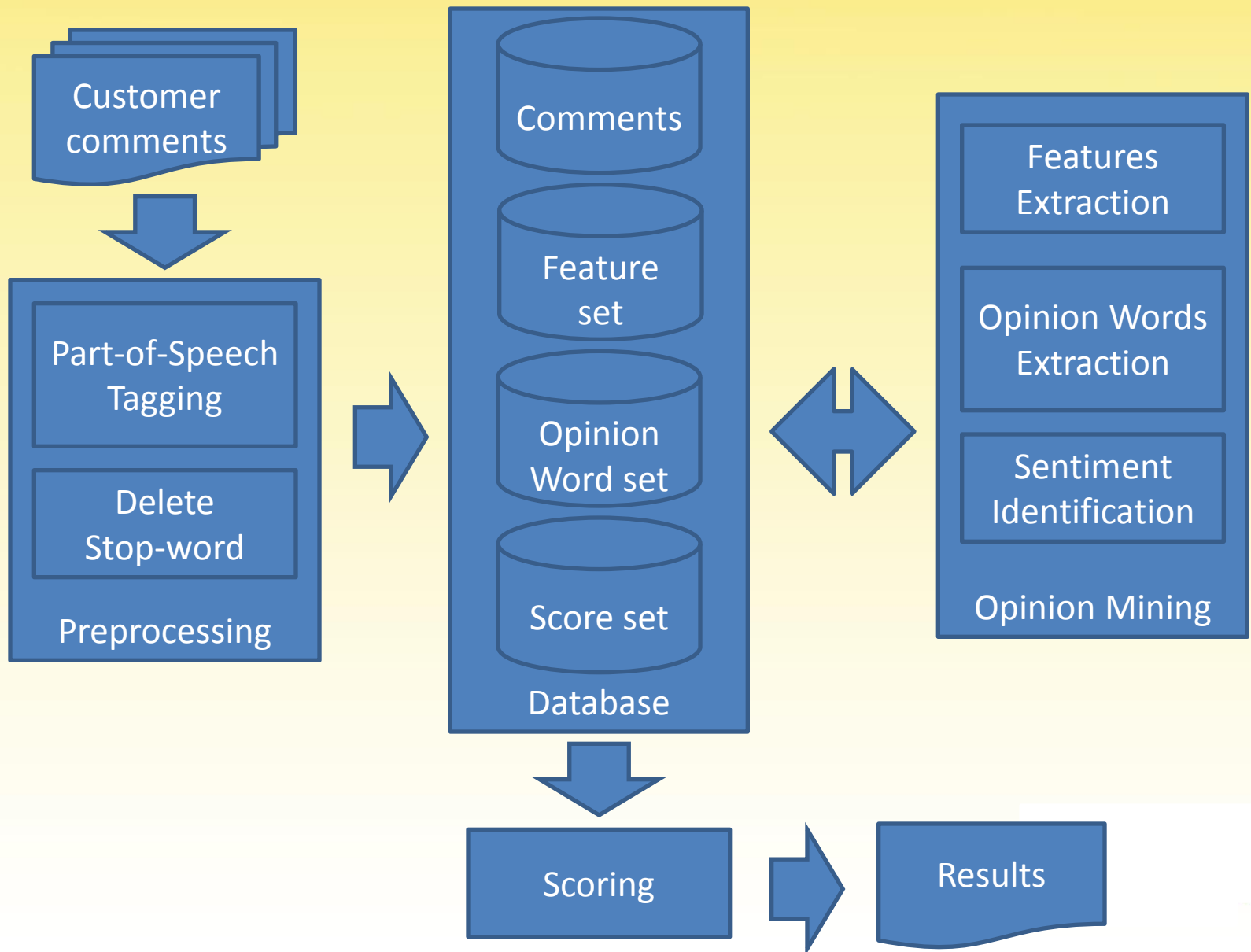
真相永遠只有一個！

條理分明

高效率



Schema



HOW ?

特徵詞

- 價格不會太貴
卸除力不太好

顯性

- 含有特徵詞：
價格不會太貴

隱性

- 不含特徵詞：
這瓶非常划算

意見字詞

- 價格不會太貴
卸除力不太好

程度副詞

- 價格有一點貴
卸除力不太好

否定副詞

- 價格不會太貴
卸除力不太好



WHAT ?

消費者

- 產品優劣
- 產品市占率
- 明確的購物依據

廠商

- 產品滿意度
- 消費者傾向
- 競爭對手優劣分析
- 智慧行銷



隱喻查詢系統



Background

當某個詞彙在文章中代表的意義不是字面上的意思：

理解困難？

理解錯誤？



隱喻

Example I

遊戲透過網路和電視在實體和虛擬世界同步炒熱話題。

原意：將食物以炒的方式加熱。 (食物)

隱喻：形成一般溝通的話題。 (溝通)

我們將這個隱喻的詞記錄成

詞彙：炒熱

特徵：(Q1 Communication ← F1 Food)

目標域：
溝通

來源域：
食物



Example II

一個字彙可能會有很多個隱喻意思！

誕生

原意：嬰兒的出生（嬰兒）

隱喻：國家的形成（政治）

特徵：(G1.2 **Politics** ← D2 Baby)

新生國家的誕生。

原意：嬰兒的出生（嬰兒）

隱喻：幸運旅客的出現（旅遊）

特徵：(D17 **Journey** ← D2 Baby)

第一百萬人次的幸運旅客終於在四月五日誕生了！

Why?

這些隱喻字彙可能
會造成語言相關處
理上的錯誤！
例如：翻譯



目的：提高語言處理的正確度，讓閱讀上更加便利！



Schema

隱喻查詢系統

資料淨化
(資料載入)



隱喻查詢
(iMetaphor)



How?

嘉義大學
外文系學者
合作提供

資料淨化

讀取

已標示隱喻之

分析

標示顏色

除錯

過濾錯誤

匯入

同時根據十四家金控已公布自結的去年十二月營收，國泰金去年全年營收為三百一十六億元，每股稅後盈餘（EPS）為三・三六元，穩坐(I2 Business←M1 Moving, coming and going)金控獲利龍頭(Metonymy)。外資券商預估今年國泰金EPS 為三・三六元，明年將增加到三・九二元，較今年成長一六%。

特徵查詢
其他查詢

查詢

找出關聯性
連接相關資料

分群

相同隱喻
全部隱喻
詳細隱喻

隱喻查詢

Multi-Function

[illegible]

Multi-Function

目標域詞彙

× 請擇一選填

來源域詞彙

內容

人

篩選

查詢

上一頁

清除

篩選

人物
人馬
人潮
空中飛人
英雄人物
敵人

篩選



Multi-Function

點選中間方框可顯示
該詞彙的所在段落

目標域詞彙

篩選

來源域詞彙

篩選

查詢

上一頁

清除

×

請擇一選填

人馬

內容
... 國民黨兩千年失去政權的種子，連戰趕出李登輝，清算李系 人馬 ，將黨營事業投資虧損全部推...
... 放各派系參與選務，不管新系、游系、蘇系還是扁系，各派 人馬 通通都是整合的對象，因為唯...
... 女部等黨務主管，都考慮延攬敗選立委擔任，以廣納各派系 人馬 來進行黨內的整合，此外，包...
... ，也關係到馬、連、蕭三方勢力的較勁。江丙坤是連戰嫡系 人馬 ，劉兆玄則是蕭萬長閣揆時期...
使出各種方法來搶位了，人單勢薄的陸啟賢怎麼敵得過這群 人馬 。

文章標題：長仔整合之旅 尋求阿輝伯支持
作 者：李心怡
出版日期：2008/01/17 第617期

所在段落

(G1.2 Politics←G3 Warfare, defense and the army ; Weapons)

相同隱喻

全部隱喻

痛定思痛之後，謝長廷也延攬落選立委進入黨中央與競選團隊，開放各派系參與選務，不管新系、游系、蘇系還是扁系，各派**人馬**通通都是整合的對象，因為唯有黨內力量先整合成功，才能全力拚總統大選。而謝長廷宣示團結的動作，也令綠營支持者對總統大選增添信心。

User Friendly

資料淨化

資料載入

資料載入

☐ 政治 ☒ 商業

Select input directory:

C:\Users\user\Desktop\新台灣新聞周刊-61筆

Progress bar:

隱喻查詢

iMetaphor

目標域詞彙

※ 請擇一選填

來源域詞彙

查詢

上一頁

清除

內容
... 名運動近來被炒得沸沸揚揚，媒體不斷報導民進黨內天王級 人物，都藉由「搶正名之功，來凸...
... 悲憤其名」情緒中，仍然憤憤不平，這不是一個有格局政治 人物 該有的表現，好像全天下只有...
謝、馬的競逐，或許就是一場自由與禁錮意涵的拔河，政治 人物 一路走來，其背後無形的力量...
黃文雄說，他認識無數的獨派 人物，來來去去，有非常以自我為...
為智能的人缺乏親和力，因而失去群眾，沈富雄就是代表性 人物。這一階段能夠吸引群眾的，

文章標題：綠色社團的超限戰

作 者：陳茂雄

出版日期：2008/02/15 第621期

詳細隱喻 (G1.2 Politics←K4 Drama, the theatre and show business)

回上一頁

2000年以後，選民有了很大的改變，平等思想已相當成熟(X2.1 Thought, belief←L3 Plants)，民眾喜歡與自己相像的人，不是賢能的人，以前被列為賢能的人缺乏親和力，因而失去群眾，沈富雄就是代表性人物(G1.2 Politics←K4 Drama, the theatre and show business)。這一階段能夠吸引群眾的，若不是擁有群眾魅力的人就是時常與群眾和在一起的人，社團的領導階層都不符合這兩個條件，所以失去群眾。雖然社團曾經凝聚上萬群眾的紀錄，那是他們利用政治議題或政治人物(G1.2 Politics←K4 Drama, the theatre and show business)，若靠自己的實力，凝聚群眾的能力還不如一個樁腳。例如有人罷免陳總統時，曾有社團發動挺扁的活動而凝聚數萬人，事實上這些都是挺扁群眾，不是社團的群眾，社團的角色(G1.2 Politics←K4 Drama, the theatre and show business)只像禿子跟著月亮走(ME-Metaphorical Expression)，借政治議題或政治人物(G1.2 Politics←K4 Drama, the theatre and show business)的光(G1.2 Politics←W2 Light)。

Sources

- **KL Communications**
 - klcommunications.com/anonymity-in-a-big-data-world
- **UrCosme**
 - www.urcosme.com



THE END

