



# 2018 年翻轉大數據 滾動新商機

陳香樺

## 「大數據 + 金融」思維

大數據是一門跨領域知識，以往討論內容都偏技術面，像 Hadoop、MapReduce、Pig 等技術名詞，不僅讓人聽得一頭霧水，也築起心理障礙，似乎只有搞電腦科學的人才能從事大數據相關工作。在數據時代，真正能彰顯大數據並獲得成功的案例畢竟是少數，我們都冀望掌握數據、提升技術及思維，轉化創造「大價值」，然而，實務執行上都面臨著嚴重的阻礙：

★**數據層面**：結構化數據、半結構化數據以及非結構化數據共存是目前金融機構面臨的數據問題。

★**技術層面**：絕大多數傳統金融機構面臨新的技術難題，傳統的分析方法

難以適應大數據的管理需要，如非結構數據處理，我們需要更有效的數據挖掘演算法人才。

★**思維層面**：國外對大數據的利用大多圍繞風險控管、營運管理和產品銷售支援。國內起步較慢，目前「市場需求細分」和「虛實通路整合價值」已得到國內金融行業的較普遍認識。

大數據尚有很多待發掘的價值，接下來，我們來說明，如何試著讓這件事「簡單化」，利用大數據融入彰化銀行營運、提升工作效率，同時幫助銀行進行業務拓展增加營收、產品行銷規劃及獲取潛在客戶，透過新的數據開闢新的商業模式。

## 一、數據從哪來？

政府持續擴大開放資料，國發會提供一份最新資料統計顯示，在所有已開放的政府資料集當中，瀏覽人數最多的前幾名資料，依序分別是「iTaiwan 公共免費無線上網熱點」、「實價登錄批次資料」、「電影」、「政府開放資料集清單」、「公司設立登記清冊」、以及「大專校院名錄」等，除少部份資料集提供有應用程式介面（API）及系統介接程式，方便開發使用者的進階應用外，大部份政府開放資料皆是以 CSV、XML 及 JSON 等格式為主。然而，當所有金融業都做了這件事，一組一組的礦工在這無償的資料努力的尋寶，利用政府開放資料搭配視覺化地理呈現，同樣商業模式的產出似乎也漸乎微不足道。因此，今年度我們重新審視行動方案，嘗試走出不一樣的解決方案。



資料來源：政府資料開放平臺

## 只做第一 不能唯二

2018 年度，我們重新定位大數據發展策略，在外部數據的蒐集，朝向各大公開網站，不再被動的受限政府部門資料，數據採集過程中，另闢戰場，走出與其他銀行不一樣的數據價值，並與業管單位保持密切即時且對等的溝通與協調。新世紀新原油，有工具、有人才，誰有本事挖掘就有無限商機。在大數據新開源軟體支援下，本行自行開發爬蟲程式，獲取以往網頁資料使用 Ajax 技術難以取得完整資訊，XPath 定位解決網頁上的元素和屬性，利用軟件包 Selenium 解決傳統網頁擷取需要下拉刷新才能抓取資料，Selenium 解決阻擋 IP 位址，同時再加入 pyesseract 解決網頁文字驗證碼的問題。現今，處理大數據的可行性越高、成本不斷減少，使得效率越來越高，從而造就了新的大數據商業智能。

在 2018 年，完整蒐集「公開資訊觀測站」、「進出口實績」、「司法法院判決書」、「台灣經濟新報」、「醫療器材商業同業公會」、「台灣採購公報」、「政府電子採購拒往名單」、「工具機暨零組件工業同業公會」等公開網站重要客戶資訊，將行外數據媒合行內資料倉儲。目前各銀行在大數據的發展，在跨出第一步時都是將公開資料進行業務的「疊合分析」，產出新的火花，找出未被妥善服務的潛在客群，真正反應客戶在行外的實體表徵。以往，

銀行最常做的就是付費透過聯合徵信中心 JCIC 得知客戶在其他銀行的其他財務、信用表徵；現在，我們可以利用 Open Data，「無償」及「免費」的使用這些資訊，試想，當我們的競爭對手都已經在努力的耕耘這些公開資料，我們一定要充份運用這些免費的外部資料，發揮業務的縱效。在這一波的公開資料採礦，不斷的蒐集實價登錄、商工業登記、上市櫃公開資訊等，利用自動化的爬蟲技術，持續加入不同公開資料源，擴展大數據市集豐富度，結合機器學習自動化爬蟲技術予以行銷化，雖然名單量體小，但是具有極高的精準率。



資料來源：資訊處 陳香禕

## 強化風險管理制度

銀行認識你的客戶（Know Your Customer）過程中，不僅在新開戶流程中需要了解客戶，完成客戶風險屬性、風險承受度評估，我們希望運用大數據強化客戶風險之辨識與監控，今年開始主動蒐集公開網站資料「司法院判決書」、「信用風險評等升降」、「政府電子採購拒往名單」等，第一線掌握往來客戶的交易行為、營運模式，利用即時郵件通知分行營業單位，事先主動提供多面向外部訊息，縮短決策反應時間，進而減少風險損失。從 2018 年 11 月開始，各分行企金 AO 與負責人皆陸續收到郵件主旨為「彰銀大數據小組」主動寄發本行企金客戶外部公開資訊，包含 40 間法院的判決書案件、信用風險重大危機事件、政府拒往名單等。

## 管理企金客戶思如零售客戶

零售金融領域，爭取客戶的競爭是非常激烈的，客戶不願或不便上門，造成銷售成績不彰，經營佈局劣勢逐漸顯現，促使銀行業者必須具備洞悉客戶的能力才得以生存，因而對客戶的資料全方位收集與分析，使用各種預測模型，邏輯斯特迴歸（Logistics regression）、支援向量機（Support Vector Machine）、隨機森林（Random forest）等進行客戶區隔，迅速因應客戶所需。

本行延續「固守傳統企金客戶」的經營策略思維下，從策略執行角度繼續落實「深耕舊戶、拓展新戶」，從2018年2月開始陸續提供在各主題的優選名單，包含「進出口廠商實績」、「醫療暨生技器材客戶」、「經濟部工廠登記」、「台灣區工具機暨零組件優選名單」、「中華徵信所TOP5000」等優選名單，驅動大數據精進公開資料探勘，協助分行企金AO瞭解客戶實際營收，從客戶管理「荷包佔有率」(Share of Wallet; SOW)的概念著手，將客戶外匯年營收額，配合歷史增長模型，如果本行客戶產品佔總目標市場比較小，但未來可能很大，就需要去發現這些市場，因為客戶最主要的往來銀行有可能不是本行，我們要搶攻「提高」客戶荷包佔有率及對彰銀金融產品服務的依賴和認可，主動蒐集公開資料，掌握客戶行外實績營收訊息。

依據企金客戶在行內的貢獻度等級及持有產品，搭配客戶進出口實績營收的公開資料，爭取外匯商品、信用狀開狀、進出口押匯等外幣授信業務，有效「移轉」客戶在他行資金及產品需求至本行，爭取最適產品配置及金融服務；同時持續開發「行外」企金客戶潛在名單，在今年度有效掌握行內外優質企金客戶數共約12萬戶，利用大數據技術進行名單的分派，計算客戶地址定位與鄰近分行實體通路，做為拜訪行外企金客戶第一線的有效接觸。

## 二、Open Data 在彰銀 2018 年開放成果大體檢

### 1. 「進出口實績優選名單」 持續服務開拓重點客戶



資料來源：資訊處 陳香梓

目前各金控及銀行在大數據的經營，最常看到成功經驗都是在 B2C 的結合運用，或者是 B2B 風險層面的著墨，幾乎很少看到大數據成功的運用反應在 B2B 的行銷經營層面。依據「2018 年中小企業白皮書」發布資料顯示，2017 年臺灣中小企業家數為 143 萬家，占全體企業 97.7%，較 2016 年增加 2.08%，創下近年來最高紀錄，2017 年中小企業銷售額增加 3.19%；我們希望在第一時間爭取扮演關鍵性金融服務角色，掌握中小企業版圖變化的連鎖反應，尋求新商機拓展開發本行外匯存放款、信用狀開狀等外幣商品業務。

2018 年第一支箭，蒐集「經濟部國貿局廠商進出口實績」公開資料，建立企金客戶歷年「進出口貿易實績級距數據集」，首先，採取行內客戶維繫策略，主動蒐集廠商進出口實績級距，結

合行內客戶台幣存放款商品往來資訊，提供分行企金 AO 優質客戶名單，爭取強化客戶與本行外匯商品黏著度，107 年 3 月開始提供近 190 家分行，掌握外匯實績級距 A 級（1000 萬美元）至 L 級（50 萬美元）共約三萬筆的行內客戶；第二支箭，以「行外客戶」為開拓標靶，為便利推展行外客戶，名單內主動提供客戶聯繫、資本額等公開資訊，在 107 年 6 月提供約八萬戶行外客戶，同時將名單有效的分派至鄰近分行實體通路，以及提供「開放戶」名單，讓潛在客戶的爭取在各分行產生良性競爭。

## 【知識+】

### 進出口貨物申報規定

1. 出進口人資格：公司或商號須先向經濟部國際貿易局辦理出進口廠商登記，方得經營輸出入業務。
2. 關務執行機關主管機關為財政部關務署。
3. 貨品分類架構採用世界關務組織（WCO）制定之「國際商品統一分類制度 2012 年版貨品分類號列（CCC CODE）共 11 碼。
4. 申報稅率查詢網站：

<http://portal.sw.nat.gov.tw/PPL/index>  
 關港貿單一窗口網站→通關簽審查詢→稅則稅率查詢。

## 【排行榜 Q&A】

**Q：**在哪裏可以知道各分行互動使用「進出口廠商實績優選名單」排行榜呢？

**A：**只要進入【內網】→【其他】→【大數據視覺化展示系統】→【視覺化展示工具】→【動態密碼查詢】就可以查詢各分行互動排行榜。

PR_BD_S6000(進出口廠商實績優選名單)				
①-2 分行員工登入次數排名				
TopN (排名前N名)	或 N 個月 (包含本月)			
3	Last 3 Months			
工作簿名稱	單位類別名稱※			
Multiple values	Multiple values			
工作簿名稱: AN_商業化平台-稅務管理				
匯集名稱: ①-2 分行員工登入次數排名_S6000				
登入員編: ITBDO208(匯集或Publisher)				
列印日期: 14/4/2019 1:17:49 PM				
【資料月份】: 2018/11/01 - 2019/01/04				
【工作簿】: PR_BD_S6000(進出口廠商實績優選名單) & PR_BD_S6002(進出口廠商實績優選名單(開埠戶))				
登入次數_TopN排名	單位類別名稱※	單位代號※	分行(單位)名稱	
1	CRC2	5005	北門分行	22
2	CRC1	5296	內湖分行	12
	CRC1	5338	永春分行	11
3	CRC2	5147	城東分行	11
		5630	南三重分行	11
4	CRC1	5685	淡水分行	8
5	CRC4	3072	南區營運處	7
6	CRC1	5543	瑞芳分行	6
	CRC1	9554	蘇澳分行	5
7		9801	南港分行	5
	CRC4	6228	東嘉義分行	5
	CRC2	5470	泰山分行	4
		9363	新林口分行	4
8	CRC3	3058	中區營運處	4
		6051	西螺分行	4

資料來源：彰化銀行 資訊處

## 2. 「司法院判決書」

### 新風險管理 主動通知模式

傳統觀念上風險是一種負面的名詞，尤其金融機構風險衡量與管理，需要多元的數據、數理統計模型，進行企業風險辨識與衡量分析各面向風險因子帶來的短、中長期衝擊，我們試著從「風險監控」的角度切入，由系統化整理及自動化的郵件通知寄發給前台營業單位，即時監督識別客戶風險的發展與變化，隨時應變措施以消除風險事件可能帶來的損失。

風險控制的學問中，有一個免費且隱藏版武器，就是「司法院法學資料檢索系統」公開資訊的資料庫，這是一個不必動用到徵信業者，就可以搜查出他人法律糾紛的資料庫，只要在Google輸入「司法院法學資料檢索系統」就能找到判例資料，目前完整包含40家法院，「司法院法學資料檢索系統」的判決書，都是已經裁定完成，不定期公告上網，司法院致力提昇裁判品質的魄力，讓我們「無償」運用大數據技術協助風險監控，主動掌握本行客戶資訊，利用公開資料「司法院判決書」查詢，結合客戶授信餘額、風險額度、客戶聯繫資訊，定期發送「郵件」通知客戶所屬授信分行之單位主管及企金AO，資訊處於2018年11月15日開發，開發完成通知系統上線，並定期每週更新案件，成功的將大數據帶入風險監控領域。

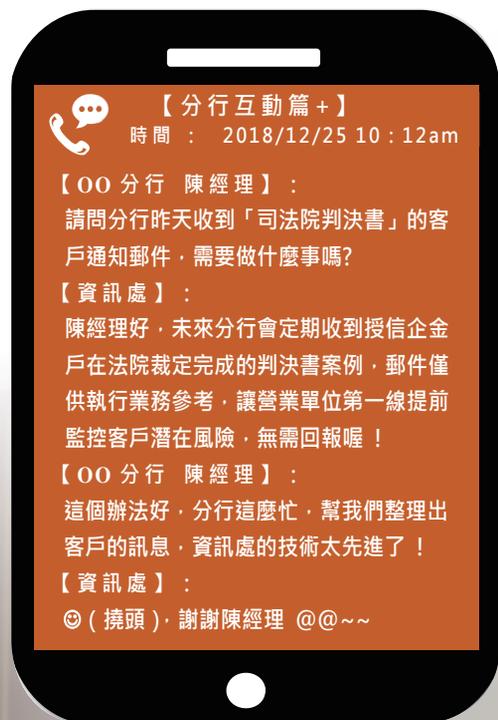


資料來源：司法院法學資料檢索系統

### 搜索術實戰操作 SOP

要利用「司法院法學資料檢索系統」來查「目標客戶」，只要按造順序選擇：

1. 「全文檢索語詞」輸入姓名。
2. 選擇「法院名稱」。
3. 選擇「刑事」、「民事」、「行政」、「公懲」四種犯罪選項。
4. 跳過的「判決字號」、「判決案由」、「判決日期」、「裁判主文」，基本上可以放棄不用填。
5. 按下「查詢」。



~ 求建議，求回饋 ~

請聯繫 資訊處【彰銀大數據小組】

### 3. 「台灣採購公報」

#### 中小企業「新經濟」平台

台灣採購公報網是針對台灣地區建置的採購資訊交換平台，以提供台灣政府或民間機構之採購者公佈的採購訊息為主要業務項目，平均每日收集各類採購訊息數百則，各種物品之採購如電腦、家俱、家電、辦公室用品、園藝、食品等，公家機關的工程標案如建築工程、道路工程、電腦軟體建置案、電腦網路建置、各項作業委外服務等各項標案之資訊，採購金額達數億至數十億元以上。

#### 「新 KYC +」秀才不出門 能知天下事

這些政府及民間的工程招標，對於我們的中小企業客戶，隱含的另

一層意義便是資金需求，從近期兩個月（2018/10 ~ 2018/11）的決標資訊來看，共 811 位決標廠商是本行的客戶，2,207 件的決標案件中決標金額高達 470 億元，在分行重新認識自己的客戶（KYC）過程中，我們試著把公開資料的客戶營收及產銷數據充分表現，同時將得標訊息包含各標案的日期、金額、案名等，系統化定期自動寄送郵件到客戶往來分行的企金 AO 及負責人，從 O2O（Online to Offline）的商業模式，透過線上平台公開資料，佈局線下客戶經營，借由大數據資訊互聯，層層疊疊如「蝴蝶效應」般，讓客戶營銷資訊公開化、即時化，擴散傳遞到前台營業單位第一線人員。

#### 最新決標消息

>	賀賴○敏以新台幣 309,600 元標得新北市汐止區○○國民小學 --- 「108 年度庶務物性工作委外」案 --- 2018/12/26 決標 (2018/12/26)
>	賀綠○環保事業股份有限公司以新台幣 432,000 元得標國立○○大學 --- 「創新育成中心(舊育成)環境清潔一式」案 --- 2018/12/26 決標 (2018/12/26)
>	賀慧○商行以新台幣 154,597 元標得新竹市香山區○○國民小學 --- 「107 學年度第一學期作業不及美勞教材採購後續擴充」案 --- 2018/12/26 決標 (2018/12/26)
>	賀國○土木包工業以新台幣 606,000 元標得苗栗縣○○鄉公所 --- 「107 年度○○鄉道路養護計畫 - 路面養護工程」案 --- 2019/12/26 決標 (2018/12/26)
>	賀富○○股份有限公司以新台幣 273,600 元標得國立○○大學 --- 「台北校區「108 及 109 年富士達電梯品質保證計畫」勞務採購案」案 --- 「2018/12/6 (2018/12/26)
>	賀文山○物業管理顧問有限公司以新台幣 765,679 元標得經濟部標準檢驗局高雄分局 --- 「108 年度辦公廳舍清潔勞務工作壹式」案 --- 2018/12/19 決標 (2018/12/26)
>	賀曦○數位設計印刷刷護工場以新台幣 485,000 元標得基隆市文化局 --- 「文化開傳月刊印製」案 --- 2018/12/25 決標 (2018/12/26)

資料來源：台灣採購公報網



### 【知識+】

「台灣採購公報網」與「政府採購公報」有何不同？

「台灣採購公報網」收集全台灣之採購標案資訊並分類處理，是集合資料收集及發送平台之網站，資料來源包含民間機構自行上網登錄及政府機關之公告之資料，每日將收集之資訊整理公佈並發送給會員參考。

而「政府採購公報」係為政府機關上傳採購資訊之平台（主管單位為行政院公共工程委員會）。

二者所公佈的資訊或有交集，但公佈資訊之日期及數量有所差別，服務之內容與方式亦有所不同。

### 【小編開講】

你知道嗎？各分行定期陸續收到多少【彰銀大數據小組】寄發的「客戶行外公開資料」郵件？

資料日期：2019/1/3

寄送日期	郵件主旨	分行數	客戶數	案件數
2018-11-09	政府採購拒往名單	11	11	11
2018-11-23	司法院判決書	165	606	8,593
2018-12-07	台灣採購公報 最新決標資訊	182	1,021	3,673
2018-12-22	信用評等—危機大事紀	40	57	76

資料來源：彰化銀行 資訊處

## 三、數據發展策略方向及願景

大數據儼然成為 21 世紀的新石油，在這一股勢不可擋的趨勢下，我們期許透過數據洞察、行為預測，精進客戶經營模式。單一的數據來源，很難完整拼湊出消費者的商業模式，所謂混合（Mix）數據才是未來的主流，上述提及與公開資料 Open Data 的碰撞，在混合數據的基礎下，利用大數據技術將數據交換、整合及優化，讓數據更優質；在下一階段，我們希望更提前掌握客戶「市場輿情」風向，蒐集客戶正面、負面訊息打造風險管理及行銷管理隱形競爭力，同時也朝向三大主軸發展：

### 1. 智能客群經營，提升服務管理

運用數據深化客群管理提升客戶價值，洞悉客群未來潛力，創建彈性管理流程並強化客戶接點，滿足客戶行為改變需求，提升客戶滿意度驅動銷售成長。

### 2. 建立社群大數據、發展輿情觀測站

聆聽與觀測雲端網路口碑生態，建立社群大數據，監控產品品牌及社群輿情相關議題，掌握社群風險脈絡，觀測強化警示風險危機與產業趨勢。

### 3. 優化金融服務

重構金融服務場景，將客戶行為轉化為資訊流，運用金融科技資訊創建並提升用戶體驗的金融消費場景，優化金融服務擴展至產品與創新金融價值。

## — 參考資料 —

1. 數位時代  
<https://www.bnext.com.tw>
2. 政府資料開放平台  
<https://data.gov.tw/>
3. 經濟部中小企業處  
<https://www.moeasmea.gov.tw>
4. 維基百科 錢包份額 (Share Of Wallet ; SOW) :  
<https://zh.wikipedia.org/wiki/Wikipedia>
5. 司法院法學資料檢索系統 :  
<http://jirs.judicial.gov.tw>
6. 台灣採購公報網 :  
<http://1060612.idun.com.tw/>
7. 經濟部國貿局  
<https://www.trade.gov.tw/>
8. 103 年版「中華民國海關進口稅則輸出入貨品分類表合訂本」之總說明。