聯徵中心協助政府機關推動 資料開放之發展

李彥錚 / 金融聯合徵信中心 研究部

背景概述

在2015年2月時,金管會為配合行政院網路溝通與優化施政政策,宣布啟動金融資料開放應用計畫,聯徵中心隨即響應公開相關金融資料,迄今已逾10年,截至2025年4月累計已公開137項資料集,並已超過一百萬次的下載次數,有效協助政府結合民間推行服務創新,促進政府資料加值應用。

資料開放的源起最早可追溯至2009年,當時從美國民間即已開始倡議政府推動資料公開,主要的論述在於可藉此舉促進公民參與政治,進而使政府與政策更能負責與透明;亦可節省民間在收集資料時必須花費的時間與金錢成本,再加上民間的創意及資源能發展出各式應用,繼而為社會整體創造附加價值。隨著英美兩國陸續推動相關政策後,各國政府亦隨之跟進,我國政府係於2012年起即開始陸續小規模推行開放資料(如:地方市政、氣象、政府預算等),自2015年起行政院擴大推廣本政策,至此開始大量開放各部會之資料,聯徵中

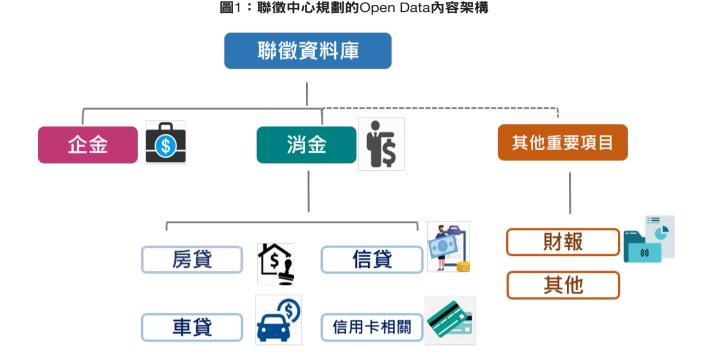
心所隸屬之金融領域亦自當時開始進行資料開放,於這10年(2015-2025)間逐步增加開放資料(Open Data)數量及提升資料品質。

聯徵中心Open Data的內容包含哪些?

聯徵中心之Open Data開放給所有人使用,不需事先提出申請,亦不收取費用,也不限制使用之目的、地區及期間。惟為保障資料當事人隱私,資料的呈現方式係彙整統計之數據,而非個別企業(或個人)的明細資訊。

資料內容之面向共可分為9類:個人授信統計資訊(19項)、房貸類統計資訊(9項)、個人信貸類統計資訊(6項)、車貸類統計資訊(4項)、信用卡統計資訊(16項)、循環信用及現金卡統計資訊(24項)、企業授信統計資訊(30項)、財報類統計資訊(18項)、其他公開資訊(11項),合計137項開放資料集。





聯徵中心參照銀行實務的分類,以消費 金融與企業金融先行區分,再依業務別各自分 類觀察;同時亦揭露財報資料庫下的各重要資 訊,架構示意圖如圖1。

在資料的呈現方式,我們係以不同「維 度」結合各種「量值」的方式來設計報表,所 謂「維度」是一種分類的方式,例如:性別、 地區別、年齡別…等;而「量值」則是使用者 所觀察的數值,例如:授信金額、利率、人 數…等。除了維度與量值外,我們尚依照量值 的兩種資訊性質(存量、流量)做進一步的分類 及組合,所謂存量即各資料時間點下的累積金 額(類似資產負債表的概念,讓使用者瞭解現況 的餘額),而所謂流量則係各資料時間點下的新 增金額(類似損益表的概念,讓使用者瞭解一段 期間的動態變化)。基於以上的各種分類方式, 我們依此邏輯可以組合成各種資料集。

以表1為例來說明這些資料是如何組合並 加以分析應用。從Open Data資料編號1-32 下,我們除可從該表觀察信貸市場下的冷熱變 化,瞭解自身借款狀況在整體市場分析的相對 位置外,亦可從性別資訊(數字1為男性、2為 女性,後段會說明為何採此呈現方式)做更進 一步的比較分析;而如果我們將量值(人數、 金額、利率)切換成「流量」的話(即聯徵中心 Open Data的資料編號1-34),就可以觀察在各 單一時間點下,新承作案件的各項資訊狀況。 相對於存量的資料而言,流量的資料更能明顯 反映各時間點下,所處環境及政策之變化與影 響。

年	月	性別	個人信貸總人數	個人信貸總金額[仟元]	個人信貸平均利率[%]
2020	1	1	781212	480777321	6.02443
2020	1	2	582625	264930738	6.35099
2020	2	1	783687	482649473	6.02916
2020	2	2	583789	266060999	6.36533
2020	3	1	790908	492093614	6.06368
2020	3	2	587671	269120868	6.42077
2020	4	1	796192	498873845	5.97428
2020	4	2	590399	271974971	6.33235
2020	5	1	795941	501231728	5.87216
2020	5	2	589490	272849359	6.23714
2020	6	1	796183	504237790	5.81079
2020	6	2	588542	274216460	6.18094

表1:信貸借款人不同性別下的授信金額及利率統計表(資料編號1-32,部分顯示)

針對企業戶而言,常用來作為維度的資訊 還包括「產業別」、「規模別」;另財報資料 的主體亦為企業戶,除沿用產業、規模等維度 外,其量值則為重要會計科目之金額(如資產、 負債、營收等)。

Open Data的特色-與一般公開資 料差別在哪?

如果直觀表1的呈現方式,很容易會有以 下疑問:「為什麼數字不放千分位符號」、 「為什麼沒有設定小數點位數」、「為什麼 不直接寫男跟女就好,寫1跟2誰知道代表甚 麼」,這樣的設定不是增加使用者的閱讀難度 嗎?而且也不美觀,不能弄漂亮一點嗎?

會有這樣的疑問是很正常的,因為Open Data的初衷,就不是要給一般人「直接」閱 覽的, 而是要給「機器」去讀取的。額外的格 式設定常常會增加機器判讀的難度,例如千分 符號(通常以逗號呈現)雖然有助於一般人判讀 數字,但因逗號又常常被做為資料處理時的分 隔符號,加了這樣的設定反而容易讓機器讀取 出現錯誤;又例如一般機器讀取數字或英文字 母較為容易,讀取中文容易出錯,因此在設定 分類符號時,並不會以中文呈現(如「男」、 「女」)。

以上說明所要傳達的基本精神,就是所謂 的「機器可讀性」(machine- readable), 其意義 在於它們的數據要能夠輕易被電腦程式擷取, 讓資料可在不受技術限制的前提下,被加值使 用甚至可依使用需求加以重組;相對而言,文 字型的文件(如PDF檔)、或編排精美的報表(即 使是以數字早現的Word檔、或者是包含各種合 併儲存格及公式函數的Excel檔),則通常不是機 器可讀取的格式,因此針對Open Data的格式及 品質,官方制定了許多規範,必須在符合相關 規範的資料才能被當成是Open Data。

(5) 大眾服務

聯徵中心針對Open Data的格式制定及品 質控管,皆依循主管機關所制訂之標準辦理, 以下略為敘明部分重點:

- ■可取得性:在這個構面下,主要的重點在於 確保資料連結的有效性,且使用者不需再另 行登入或進行額外的操作。
- ■易於被處理:這個構面強調的是資料必須是 「結構化資料」(具有標準化格式的資料), 除了資料要可被機器讀取外,檔案的開啟也 不能被特定軟體所限制(所以Open Data的 檔案格式不會是常見PDF檔或Excel檔,因 為這意味著使用者要額外安裝這些軟體,才 能處理檔案),對於各個不同的系統也要有 高的相容性。因此資料格式聯徵中心幾乎皆 設定為CSV檔,並採用UTF-8的編碼格式。 資料本身除欄位固定日只有單列標題外,尚 必須無合併儲存格、無公式、無空行、無小 計,呈現的是資料最原始的樣貌。
- ■易於理解:這個構面強調的是必須讓使用者 能看懂資料代表的內涵,因此必須要有相關 資料說明(例如:「主要欄位說明」),但資 料的說明也必須格式化,才具機器可讀性, 所以這樣的資料又稱為詮釋資料¹,讓資料 內容能被系統化的管控。此外,資料集尚必 須定時更新(更新頻率由發佈者自訂),目前 聯徵中心針對Open Data的更新頻率,基本 上為每季更新,資料更新的時間點為每季季 末(3、6、9、12月)之下旬;至於財務報表

- 等一年才出具一次的資料(以一般中小企業 而言),更新頻率係一年一次,資料更新時 點為每年年底。
- ■資料標準:此原則約為2020年起逐步納入 規範,主要原因係開放資料的風潮已越來越 盛,有更多的機構及資料集皆加入政府資料 開放的行列,那麽若僅憑「機器可讀性」 (將資料加以標準格式化),就無法解決跨機 構可比較性的問題。Open Data並不是只有 聯徵中心一家機構在發布,而是全國各公部 門都在進行,那麼針對相同的資訊內容,理 應有一致的表達方式,才能比較及串接。例 如:同樣是性別的資料,如果有些單位是早 現「M/F」、有些單位是「1/2」,這樣就會 造成資料處理上的混亂,因此主管機關特制 訂各項資料標準,藉由一致的資料標準(統 一皆以「1/2」呈現),就能使相同的資料加 以串接比較。而當前主管機關係鼓勵一個資 料集內需有20%以上欄位符合資料標準的規 範,以表1為例,該資料集共有6個欄位,當 中「年」(被規範以西元年呈現,而非民國 年或其他紀年方式)、「月」(被規範以數字 呈現,而非JAN、FEB等其他文字方式)、 「性別」(被規範以1/2呈現)皆依循相關資 料標準,符合主管機關對於資料品質之規範 要求。主管機關亦鼓勵各單位自提資料標準 (含跨領域或專門領域),經有關單位審核通 過後,即可作為正式的參照標準。

即Metadata,簡而言之就是解釋資料的資料。



另主管機關亦於2020年前後,開始鼓勵各 單位的資料集採用符合OAS標準的Open API² 型式,資料使用者透過開發額外的應用程式即 可存取相關資料,提升資料處理效率;而聯徵 中心亦已於2022年底完成建置OAS OpenAPI 平台,所有的資料集皆可透過OpenAPI擷取, 如圖2、圖3。其中圖3所示的檔案格式為JSON 檔,可以看出其呈現方式為純機器讀取之形 式,由此可知所謂的機器讀取格式其實與一般 人的閱讀方式有極大差異。

圖2:聯徵中心提供Open Data的OpenAPI平台畫面



圖3: OpenAPI平台下的資料呈現方式

API是Application Programming Interface (應用程式介面)的縮寫。藉由Open API,使用者可以透過開發額外的應用程式,以更有效率 的方式存取資料。



針對Open Data的各項規範,聯徵中心皆 全力配合以確保資料品質能符合最高標準,而 聯徵中心所開放的所有資料集,其品質皆符合 主管機關所設定金標章及白金標章之要求,並 多次協助金管會於數位發展部(過去主管機關為 國發會)所舉辦的開放資料競賽中,於「資料開 放金質獎」的項目下獲得第一名。

Open Data之效益

聯徵中心之Open Data自2015年2月開始 上架以來,截至2024年底已累積下載約101.2 萬次,各類資料的下載次數如表2,其中因消 費金融資料與民眾日常生活較相關,故其資料 有較多的下載數,大眾媒體亦針對相關數據多 有報導。

資料類別 (項目數)	下載次數		
個人授信統計資訊(19項)	19.1萬次		
房貸類統計資訊(9項)	10.7萬次		
個人信貸類統計資訊(6項)	3.9萬次		
車貸類統計資訊(4項)	2.9萬次		
信用卡統計資訊(16項)	10.9萬次		
循環信用及現金卡統計資訊(24項)	16.8萬次		
企業授信統計資訊(30項)	17.0萬次		
財報類統計資訊(18項)	12.3萬次		

表2: Open Data累計下載次數統計(截至2024/12)

聯徵中心之Open Data的效益可分為2個 層面:

■協助金融創新:各式各樣的FinTech金融科 技業者相當善用網路上的公開資訊,透過網 站或APP方式整合資訊,讓金融業者掌握市 場資訊,簡化市場調查的程序及成本,並可 依市調結果開發新服務或修正原有產品;民 眾亦可透過金融創新的成果享有更時新的金 融服務。

■金融知識的傳播:已有不計其數的大眾媒體 引用聯徵中心的Open Data來觀察社會趨勢 脈動,亦有相當多的學術研究將聯徵中心的 Open Data做為主題探索的敲門磚甚至直接 引用相關數據。這顯示民眾可能過去就對這 些議題深感興趣,只是因為相關資訊皆藏在 資料庫深處而不得其門而入,藉由公開資訊 的分享,激發了相關人士深入探索的動力, 讓知識的傳播更加無遠弗屆,亦可讓民眾在 資訊的普及下更瞭解信用的重要性,更加努 力地維護自身信用價值。

7.6萬次

其他公開資訊(11項)

開放資料的活化應用-運用大數據 分析資訊

前文所述各種Open Data的格式規範,只 是為了讓其成為具機器可讀性的原料,但如何 轉化成一般民眾都易於理解及操作的產出,則 係各種金融創新工作者的挑戰。以一個普通使 用者而言,他至少應該會希望:

- 1.可以自由篩選各種維度,組合出他想要的報 表
- 2.使用介面可以拖拉點選,免於額外撰寫程式 (如為非資訊背景的使用者)

以上這幾個原則,其實就是將資料活化應 用的一個過程。誠然民間已有許多業者開發出 了各種加值應用,但從一個資料原始提供者的

角度出發,為讓開放資料的運用能更加普及與 靈活,聯徵中心亦從消金及企金選取大眾最為 關注的兩個主題「房貸」、「產業財務統計」 進行活化應用,置於聯徵中心官網「資料開放 專區」下的「運用大數據分析資訊」區塊下(正 式名稱分別為「購置住宅貸款統計資訊」及 「產業財務統計資訊」)。

前沭大數據之專區,係將房貸及產業財 務的相關資料結合金融機構往來之授信資料、 會計師簽證財務報告書等進行加值運用。特色 在於針對使用介面上做了許多設計,如簡易索 引、地圖化、制訂下拉清單供使用者自訂條件 篩選等功能,讓使用者能以更方便、更快速地 取得相關資訊。

圖4:大數據專區的操作介面示意圖(房貸)



圖5:大數據專區的產出結果示意圖(房貸)







圖6:大數據專區的操作介面示意圖(產業)

圖7:大數據專區的產出畫面示意圖(產業)

查詢結果											
第数:共4 篇										匯出EXCEL	
資料表					图表						
序號	年度	行業別	財務比率	模本數	簡單算術平均數	中位數	上四分位數	下四分位數	標準差	綜合平均數	
1	2013	食品製造業	權益報酬率(稅前)	27	11.50	11.10	18.80	2.10	10.60	14.90	
2	2013	食品製造業	營收成長率	27	1.80	1.40	5.30	-4.90	10.50	4.20	
3	2013	紙費、紙及紙製品製造業	權益報酬率(稅前)	9	4.80	2.40			7.40	2.80	
4	2013	纸摄、纸及纸製品製造業	營收成長率	9	34.30	0.10			96.70	12.70	

透過前述大數據專區的查詢網,民眾除可 **查詢欲購屋或賣屋所在區域的市場行情,還可** 進一步獲得銀行近期核貸的相關資訊,包含核 貸成數、利率等,可更精確掌握買屋賣房所需 資訊,作為決策參考;房仲業者亦可運用查詢 網獲得之鑑估值、貸款額度、成數、利率等資 訊,作為媒合買賣雙方之參考。

至於產業別的財務統計資訊的應用,不論 是金融業徵授信審查、投資者欲評估投資標的 之財務表現、或證券商業者及會計師輔導中小 企業於辦理公開發行、上櫃或上市時,都能以 一個相對合理之比較基準,進行財務分析及同 業財務比較。

聯徵中心所規劃的「運用大數據分析資 訊」,獲金管會指派前往參加2019年國發會舉 辦的政府資料開放獎勵活動(該活動現行之主管

機關為數位發展部),並於該活動中榮獲資料開 放應用獎第1名的肯定,顯見聯徵中心的開放 資料,無論是在資料品質或活化應用上,皆獲 得主管機關的高度肯定。

未來之展望

自2015年政府推動資料開放政策以來, 回顧聯徵中心這10年從無到有的開發歷程, 關注的重心從草創期追求資料開放的數量、進 一步轉為提升資料品質、再進一步轉為資料應 用效益的推廣,我們發現追求政策的進步並沒 有所謂的休止符,因為社會需求的脈動不斷在 變化,聯徵中心同時為「金管會資料開放諮詢 小組」的一員,將持續掌握社會大眾的關注重 點,協助政府繼續前進。



圖8:時任行政院政務委員頒獎予聯徵中心董事長

數位發展部針對開放資料有其整體的規劃(參圖9),近年為使開放的效益能更加提升,參考我 國核心戰略產業及歐盟的發展趨勢,於2023年起推動各政府機關開放「高應用價值」主題資料,優 先聚焦於農業永續、空間資訊、氣候環境、災害防救、交通運輸、健康醫療、能源管理、社會救助 等8項主題(未來仍可能持續新增主題),期許在這些主題下帶來更有品質與即時性的資料,日後甚至 能建構主題式資料應用生態圈,讓資料愈分享愈有價值。

近年ESG已為各界關注 的重點,聯徵中心亦有蒐集綠 色金融等相關資訊。在這樣的 氛圍下,金管會目前已規劃在 企業永續的議題上,優化並深 化開放具主題應用價值之資料 集,聯徵中心未來亦將配合主 管機關的指導持續前行。

圖9 數位發展部對於開放資料的規劃示意圖

