

整合異質性資料來源成為行動網站的方法

周椽家¹ 翁瑞元² 王祥安³ 黃翰隆³

¹元智大學資訊管理學系 ²元智大學資訊工程學系 ³中央研究院數位文化中心
¹s991646@mail.yzu.edu.tw, ²s993330@mail.yzu.edu.tw,
³{sawang, hlhuang44}@iis.sinica.edu.tw

摘要

開發者在建置行動應用網站時，對於多種異質性資料的來源，可能沒有一個統一格式或來源，需依照各種資料來源逐一撰寫資料的抓取程式，可能會造成繁複的程式開發工作。因此，本研究提出結合 Solr 軟體及 Kurogo 軟體將異質性的資料整合並成為行動網站的方法。透過 Solr 軟體整合異質性的資料，並運用本研究開發的對應欄位設定軟體，橋接 Solr 軟體與 Kurogo 軟體的資料處理流程，且可以快速的產生將資料轉換成行動網頁的功能，Kurogo 軟體再透過此功能取得 Solr 軟體整合後的資料，將資料轉換成適合不同解析度的行動裝置呈現樣式。此方法可以降低對於不同行動裝置需要開發不同的行動網頁之複雜度，並快速的將整合異質性資料的網站轉換成行動網站。

關鍵詞：行動網站、異質性資料整合、行動中介軟體、響應式網頁

1 前言

隨著多樣化數位化設備及軟體的發展，會產生許多不同格式的資料，若要整合這些資料，要將不同格式的資料轉換成統一的格式，需要撰寫程式來轉換資料，所以資料整合是不易的工作。根據 Google 提出的多螢幕時代來臨，資訊會根據不同行動裝置螢幕的大小被呈現[1]。但是，將資料在不同的行動裝置上呈現適合的樣式是需要經過繁瑣的處理。因此，將不同資料來源整合，並提供一個統一的架構來提供資訊於各行動裝置，來解決上述的困難是本研究的目的。

使用適合的方式建置行動網站(Mobile Web)，並透過此網站將組織想傳遞給使用者的資訊呈現於行動裝置是目前的趨勢。根據 Mary Meeker 在 D11 會議中提到，直至 2012 年底，全球有超過 24 億的網路使用者，行動上網用戶更從 11 億增加至 15 億人相較 2011 年增加 30% [2]。由此可知，行動上網用戶的提升，促使透過行動裝置瀏覽網頁的人數攀升，因此透過行動網站傳遞資訊給使用者已是目前的趨勢。

針對於多螢幕時代，Ethan Marcotte 提出了響應式網頁設計(Response Web Design)的概念[3]，此

設計方式，主要針對不同的行動裝置修改 CSS/Javascript 語法等，給予使用者介面優化的體驗，讓使用者可以降低點擊次數來取得所需資訊，而不需要在解析度較小的螢幕上，以放大或縮小的方式來瀏覽網頁內容。

為了解決前面所提到的問題，因此我們提出了一個方法，此方法整合兩套開放原始碼軟體(open source software)，解決異質性資料來源整合檢索並以較佳的方式呈現於不同的行動裝置，讓行動網站建置更為簡單，因此，本研究透過 Solr 軟體將異質性資料整合，再透過自行開發的對應欄位設定軟體連結 Kurogo 行動中介軟體與 Solr 軟體，有效的整合資料並降低程式開發者對於不同解析度的行動裝置建置複雜度，可降低建置行動網站的成本。

2 相關文獻探討

2.1 行動網站建置的方法

若要建置一個行動網站，以網頁語言開發一套跨平台行動裝置瀏覽的網站主要是以瀏覽器支援的語言來編寫的應用，可以採用以下幾種不同的方式來建置行動網站：(1)Web 軟體框架(Web application framework)，(2)行動中介軟體(mobile middleware)，(3)網路服務轉成行動網站。

2.1.1 Web 軟體框架

Web 軟體框架是一種電腦軟體框架，主要用來支援動態網站及網路服務的開發[4]。可以同時相容於多種行動裝置平台，使得原本一定要用原生應用程式(native APP)開發框架才能做出的使用者體驗，多能利用 HTML5 和 CSS3 的技術來完成，為了提供服務給不同的行動裝置，強調使用 HTML5 來完成跨平台行動網站開發的框架(Frameworks)陸續出現，例如：jQuery Mobile、Sencha Touch、SproutCore[5]，其中 jQuery Mobile 是較普遍使用的架構。

jQuery Mobile 為以 jQuery 為基礎來開發針對相容各行動裝置所開發出來的一套架構，jQuery Mobile 很容易開發，但也有不足之處，最大的問題在於設計方式要符合 jQuery Mobile 所定好的 HTML 程式及 CSS 規範，各種瀏覽器在不同的手機

平台上表現也不相同，雖然目前 jQuery Mobile 提供很好的使用者介面，可透過自動針對螢幕大小與解析度做出縮放與重新編排，但若 jQuery Mobile 要實行響應式網頁設計也需要透過 CSS 針對不同大小的裝置平台去花費大量的時間做客製化樣式的呈現。若需要從資料庫抓取資料，開發者對於異質性資料的處理也只能靠程式去一步步抓取，這樣的製作方式對於行動應用網站的開發者無法有效地將資料整合於跨平台的應用服務網站。

2.1.2 行動中介軟體

中介軟體(middleware)是提供系統軟體和應用軟體之間連接的軟體如圖 1，以便於軟體各元件之間的溝通，特別是應用軟體對於系統軟體的個別應用，在現代資訊科技應用框架如 Web 服務、面向服務(service-oriented)的環境中應用比較廣泛。

行動中介軟體[6]，是介於企業組織後端的系統軟體與應用軟體，透過無線網路和 client 端的行動裝置相連結，透過組織內區域網路和伺服器的溝通，將所需資訊透過無線網路傳送到行動裝置呈現。

Middleware/Mobile Middleware

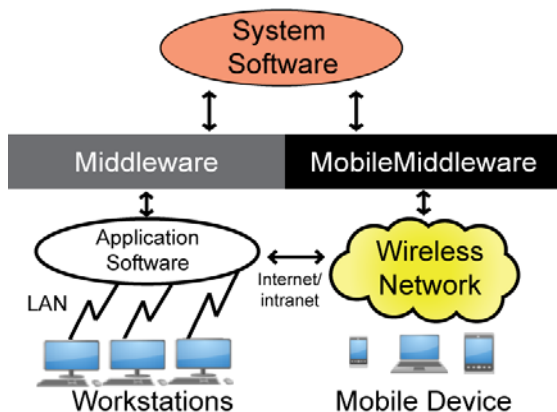


圖 1 中介軟體與行動中介軟體架構圖[7]

2.1.3 網路服務轉成行動網站

網路上提供許多服務，將呈現給個人電腦呈現的網站內容，透過各網路服務自行分析並產出適合行動裝置呈現的網站內容，例如：GinWiz、Mobify、ConvertWebsite。這些網站共同的優點是皆可以快速的針對於解析度較小的裝置優化，轉換成使用者方便閱讀的行動網站，不需要撰寫複雜程式的方式來轉換行動網頁。但其限制是只能讀取原始網站的內容頁面，其他格式的資料內容則無法讀取，而且，它們讀取原始網站時，可能會造成解析不完全正確。

2.2 Solr 全文檢索軟體

Solr 是一套開放原始碼軟體(open source software)的全文檢索軟體[8]，它可以從文件檔案或者資料庫中萃取出訊息。開發者可以透過編寫設定檔進行資料庫整合或處理格式不同的文件(例如：Word、XML) 成為一份 Solr 索引檔。索引檔是為了提高檢索速度而發展出來的檔案結構，其中包含了多份文件，一份文件中有多個自行定義的欄位，一個欄位裡包含許多詞組(terms)，詞組由成對的字串所組成。Solr 具有表徵狀態轉移(REST)的軟體設計風格，可以接受多種請求方法，例如：使用 GET 方法取得資料，而資料回應的格式也可以讓開發者自行決定(如：XML、JSON、CSV)，只要透過 HTTP 協定的程式都可以跟 Solr 軟體進行溝通。

由於 Solr 整合異質性資料來源、提供全文檢索以及擁有 REST 的特性，解決了檔案格式不一致、難以取得檔案資料和程式不相容的問題，因此，本研究採用 Solr 來實作此研究方法。

2.3 Kurogo 行動中介軟體

隨著行動裝置的普及而誕生 Kurogo 軟體[9]，它是一個以 PHP 語言所編寫而成的行動中介軟體，可以將不同的資料格式，從後端資料層與前端瀏覽器連結，呈現最適合的樣式給不同解析度的行動裝置，成為一個適合大型組織開發行動網站的方法。

Kurogo 的架構，如圖 2 [9]。模組(Module)為選

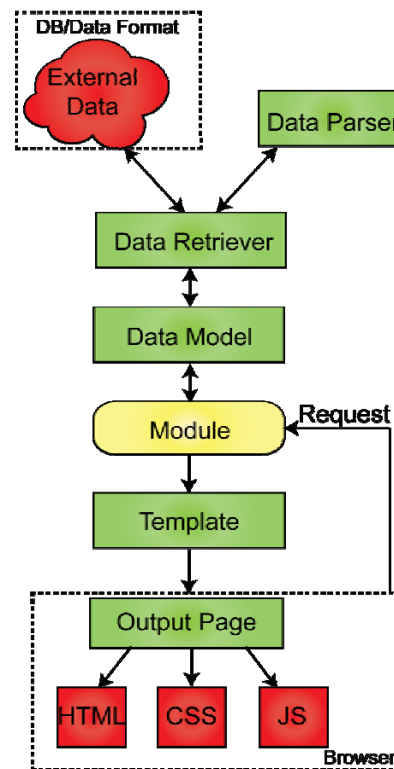


圖 2 Kurogo 架構圖

輯運算的核心，其主要功能為接受使用者的請求 (UserRequest)、處理資料來源、將資料送至樣板 (Template)。當客戶端的請求送到各模組中，模組會透過 Kurogo 資料設定檔來處理預先定義的資訊，例如：透過 Kurogo 資料設定檔的資訊到外部資料庫或不同資料檔案(DB/DataFormat)進行查詢，讓資料抓取程式(Data Retriever)及資料解析程式(Data Parser)解析後送回模組，經模組處理後，將結果傳至樣板引擎(Template)，Kurogo 採用 Smarty 樣板引擎，讓開發者可以將商業邏輯與呈現邏輯分開，降低程式設計師與樣板設計人員的耦合度，也可以在此樣板中嵌入 HTML、CSS 和有限制的 Javascript(JS)，讓網頁版面的調整控制性及擴充性更佳。

Kurogo 具有四種特色：(1)裝置偵測會根據裝置的硬體類型、作業系統、瀏覽器來決定呈現的網頁形式，網頁形式主要是運用 CSS、Javascript 來選定，平台使用特定的作業系統作為判斷基礎，瀏覽器使用何種瀏覽器來觀看此網站。(2)自訂資料來源，可以設定不同的資料來源、自行建置模組，甚至可以在輸出端依照需求，修改樣板的呈現方式。(3)管理者可透過介面來編輯該行動網站的設定及各模組的資訊設定，如：模組中各頁面的顯示名稱、權限控制等。(4)開發者下載 Kurogo 軟體後，其檔案內包含原生範例模組，其模組可以符合大多數組織行動網站所需要呈現給使用者的功能，只需透過模組的資料來源設定，即可以達到行動網站的建置。由於 Kurogo 上述四項優點，本研究採用 Kurogo 軟體來實作此研究方法。

3 系統架構

本系統採用的架構是 Client-Server 的軟體架構，在 Server 端，使用 Solr 軟體、Kurogo 軟體及我們所開發的對應欄位設定軟體(DataField Config Program)，其整體架構如圖 3。(1)Solr 軟體負責資料整合的處理，透過 Solr 設定檔，將異質性資料來源(Data Source)整合統一，它可以將資料庫、XML、CSV 等資料格式的來源，整合至 Solr 軟體。(2)對應欄位設定軟體，其主要功能為溝通 Solr 軟體及 Kurogo 軟體。因為 Kurogo 軟體需要去設定資料來源或向資料來源撰寫資料抓取程式，此過程需要透過會撰寫程式的使用者才能做到，因此，我們開發一個對應欄位設定軟體，它會去自動取得 Solr 資料的綱要(schema)，讓使用者設定呈現於行動網站的內容，Kurogo 再依使用者的設定將 Solr 的資料取回，轉換成行動網頁。(3)Kurogo 軟體主要功能為將 Solr 軟體的資料來源根據網站使用者的裝置環境來轉換成行動網頁，然後傳送到 client 端的瀏覽器(Web Browser)。Kurogo 抓取資料時，可以透過 REST 方式來向 Solr 抓取資料，並可以利用緩存(Cache)的機制來降低伺服器的負擔。

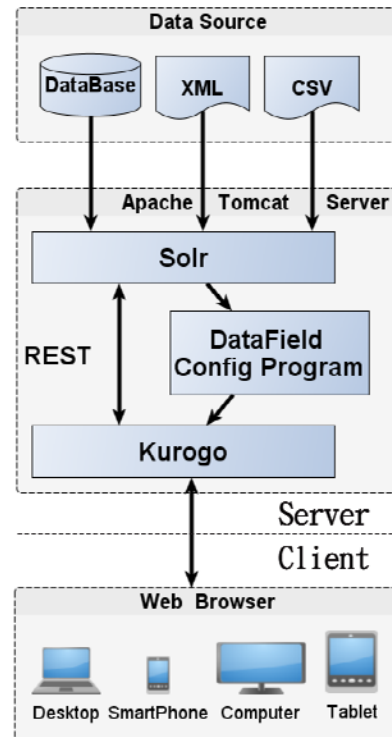


圖 3 軟體架構圖

4 實作過程與結果

本研究針對整合異質性資料成為行動網站提出一套實作的方法。以「中央研究院數位典藏資源網」為例[10]，如圖 4 所示，它是透過 Solr 軟體來整合異質性資料來源，整合的資料庫與網站超過 150 個，並透過 Kurogo 軟體來製作它的行動網站。整個實作的過程可以分為二個部分，第一部分編寫 Solr 綱要及設定檔，目的是將異質性資料來源匯入 Solr 軟體、第二部分使用對應欄位設定軟體來設定 Kurogo 軟體要向 Solr 軟體抓取的資料，以做為呈現行動網站的內容。

第一部分，為了將異質性資料來源整合成 Solr 索引檔，首先使用者需設定 XML 格式的 Solr 綱要、資料設定檔、系統設定檔，分別說明如下：

- Solr 綱要：主要定義在 Solr 中的欄位名稱(field name)、屬性以及類型。
- 資料設定檔(data config)：此檔案的作用是設定資料來源的位置，從各異質性資料來源的欄位(column)對應到 Solr 的欄位(name)中。
- 系統設定檔(config)：設定收到請求所要進行的處理，設置觸發監聽器(listener)的事件，以及 Solr 所需使用的函式庫。

Solr 軟體讀取上述三個設定檔後，可以從異質性資料來源萃取出使用者所需要的資料，建立 Solr 索引檔。



圖 4 中研院數位典藏資源網的專題文章頁面

第二部分，本研究所開發的對應欄位設定軟體可以簡單的透過設定來新增 Kurogo 模組及設定異質性的資料來源。由於 Kurogo 軟體可以自行新增模組，但使用者新增模組時，需在固定的檔案路徑下自行新增有限制的模組資料夾及檔案名稱，常會造成使用者的不便。使用者對於 Solr 軟體資料也需要自行撰寫程式去抓取所要的資料，因此，對應欄位設定軟體主要提供新增原生範例模組及 Solr 串接 Kurogo 的模組兩個重要的功能，分別說明如下：

第一個功能為快速的新增原生範例模組，它提供多數組織行動網站所需要呈現給使用者的功能，例如：(1)最新消息模組，採用 RSS 標準輸出格式的資料做為資料來源，主要提供組織網站之最新消息。(2)地圖模組，以 Google Map 匯出 KML 檔案格式的方式來標示位置，主要提供組織的所在位置，並且顯示在行動網頁上等，若使用者想要快速的開發一個網站，並使用相同的原生模組功能，需透過資料夾的複製，才能實現原生模組的功能。因為檔案名稱的不同進而要改寫模組內的程式，而無法透過簡易地操作來新增模組，因此，我們開發僅需透過設定即可快速新增原生範例模組的功能，使用者只須抽換原生範例模組的資料來源，就可輕易產生具有與原生範例模組功能或版面相同，但資料來源不同的模組。

第二個功能為設定異質性資料的來源，透過新增 Solr 串接 Kurogo 的模組設定 Kurogo 軟體要向 Solr 軟體抓取的資料，僅需透過簡單設定 Solr 主機及資料來源資訊，即可以自動取得 Solr 資料的綱要，將綱要內所要呈現的資料，透過 Kurogo 軟體呈現於行動裝置上。此功能設定如下：使用者需提供整合於 Solr 軟體下資料來源的 Solr 主機 IP 位置、Solr 名稱以及 Solr 核心資料的名稱，當按下連結按鈕時會自動去連結 Solr 資料存放主機位置。

設定完成 Solr 軟體資料來源的連結後，由於 Kurogo 軟體提供在同一台伺服器下可以建立多個行動網站，因此，需要選擇新增的模組位於哪個行動網站(site)中，如圖 5。在此展示網站名稱 mysite，顯示模組名稱(Module Name)為專題文章，模組功能



圖 5 模組命名及功能選擇

(Module Function)為 topics 的模組。其主要功能為將不同類型的文章分類，方便行動裝置使用者閱讀文章。按下設定 Solr 按鈕(Set Solr)後，需要進階設定此模組需要新增對應 Solr 軟體及 Kurogo 軟體的資料欄位並呈現於行動網頁的資訊。例如：在專題文章模組中，使用者需設定兩軟體的資料對應欄位方式，讓 Kurogo 軟體知道要向 Solr 軟體抓取對應欄位的資料，如圖 6 所示。

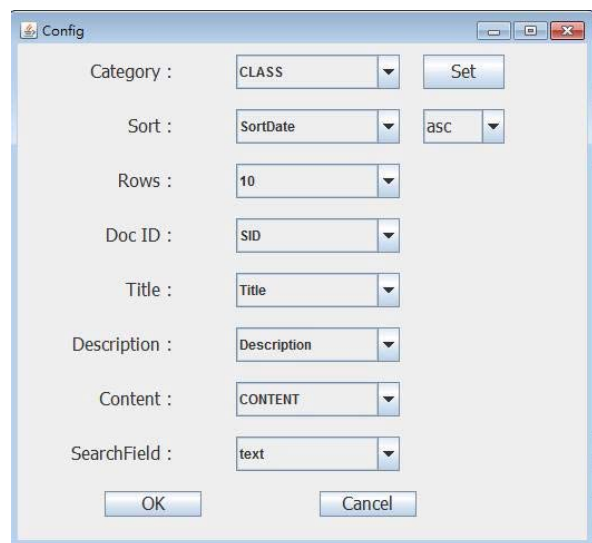


圖 6 專題文章對應欄位設定

分類的詳細設定，左方為呈現在行動網頁上分類所要顯示的名稱(Display Name)，右方為 Solr 軟體分類的綱要欄位值(Solr Value)，設定的目的為向 Solr 軟體抓取資料傳至 Kurogo 軟體來取得文章的分類。下方的按鈕功提供新增、刪除及增加包含所有上方分類的全部文章(All Category)，如圖 7 所示。



圖 7 專題文章分類詳細設定

當完成前兩個功能設定後，即可以將「中央研究院數位典藏資源網」內的專題文章內容，包含歷史文物、人文社會、植物生態、海洋生態、技術新知、文學創新，為使用者在圖 7 中所填入的顯示名



圖 8 中研院數位典藏資源網的專題文章行動頁面

稱，透過本研究的方法實做出行動網站，並呈現於行動裝置，如圖 8 所示。

對應欄位設定軟體可以快速的新增原生範例模組及透過我們新增的 Solr 串接 Kurogo 軟體模組。我們提供簡單的設定於 Kurogo 軟體上呈現整合異質性資料於 Solr 軟體資料的對應欄位方法，可以不需另外撰寫程式碼呈現於 Kurogo 行動網站上的資訊。如果不透過此方式，可能需要自行撰寫和異質性資料格式或資料庫溝通的資料抓取和解析程式，且需將資料整理成統一格式，並自行撰寫呈現於各行動裝置網站的偵測程式，困難度及複雜度會相對提高。

針對使用者透過對應欄位設定軟體設定完成的資訊，我們進一步開發透過 Kurogo 軟體的管理者介面，將使用者所設定的資訊做成介面化的呈現，如圖 9。透過 Solr 軟體資料的修改，管理者可



圖 9 專題文章於 Kurogo 中的管理者介面

以直接透過此介面對呈現內容做些微的調整，不需要自行修改 Kurogo 資料設定檔或再次透過對應欄位設定軟體來修改專題文章模組中 Solr 資料來源的主機位置、分類方式、透過 REST 方式向 Solr 溝通的主要搜尋欄位、文章排序方式、每頁回傳筆數、需要與 Solr 軟體溝通並回傳呈現於 Kurogo 軟體的資料欄位。使用此介面，可以讓網站管理者在日後維護及 Kurogo 軟體對於 Solr 軟體資料的修改呈現於行動網站的內容更為方便。

5 結論

本研究提供一套方法讓開發者使用 Solr 軟體整合異質性資料，並透過行動中介軟體 Kurogo 來製作成行動網站。經由 Solr 軟體處理異質性資料來源，開發者不需撰寫程式，只需撰寫 XML 格式的綱要及設定檔，讓龐大的異質性資料整合變為簡單。透過我們開發的對應欄位設定軟體定義 Solr 欄位與 Kurogo 欄位，將 Solr 整合後的資料透過 Kurogo 軟體呈現至行動網站，Kurogo 軟體會依據硬體類型、作業系統、瀏覽器來偵測不同解析度的行動裝置，自動調適成使用者方便閱讀的網頁形式，達到響應式網頁設計的概念。

在未來的研究方面，可擴充本軟體的功能，讓使用者透過 HTML5 與 CSS3 樣式的調整來改變主題顏色、按鈕樣式。使用者也可以透過檔案上傳行動網頁的 logo 或圖片，讓開發者可以有更大的彈性來做樣板及樣式的修改，例如：客製化各行動裝置行動網頁的頁首(header)、頁尾/footer)等。

致謝

本研究計畫由臺灣數位成果永續維運計畫—數位成果推廣應用計畫 (I) 資助，計畫編號：NSC 102-3113-P-001-003。

參考文獻

- [1] The new multi-screen world study, <http://www.google.com/think/research-studies/the-new-multi-screen-world-study.html>.
- [2] M. Meeker and L. Wu, "Internet trends," D11_Conference, pp.30-37, 2013.
- [3] App vs the Web, <http://mattgemell.com/2011/07/22/apps-vs-the-web/>.
- [4] M. Lu and W. Yeung, "A framework for effective commercial Web application development," Internet Research: Electronic Networking Applications and Policy, vol. 8 iss: 2, pp.166-173, 1998.
- [5] A. Charland and B. Leroux, "Mobile application development: web vs. native," Communications of the ACM, vol. 54 iss: 5, pp.49-53, May 2011.
- [6] F. Adelstein, S. K. S. Gupta, G. Richard, and L. Schwiebert, Fundamentals of mobile and pervasive computing, 1st ed., McGraw-Hill Professional, 2004, pp.113-125.
- [7] T. Kanoc, "Mobile middleware: the next frontier in enterprise application integration," White paper, Nettech Systems Inc., April 1999.
- [8] D. Smiley and E. Pugh, Solr 1.4 enterprise search server, Packt, August 2009.
- [9] Kurogo, <http://kurogo.org/home/>.
- [10] 中央研究院數位典藏資源網, <http://digiarch.sinica.edu.tw/>.