

以知識本體和鏈結資料建置圖書資訊學 領域學者的事業歷程網站系統 —以王振鵠教授為例

符興智

國立臺灣師範大學圖書資訊學研究所碩士生

E-mail: felixfu007@gmail.com

柯皓仁

國立臺灣師範大學圖書資訊學研究所教授

E-mail: clavenke@ntnu.edu.tw

關鍵詞：王振鵠；圖書資訊學；鏈結資料；知識本體；語意網

【摘要】

王振鵠教授被認為是影響臺灣圖書館領域深遠的學者，其在任職國家圖書館館長期間建樹良多，同時也培養出許多傑出人才。本研究運用知識本體和鏈結資料技術，分析王教授的生平事蹟與學術貢獻，設計出一套總計 26 個類別、29 項屬性、720 件實例與 3,645 條三元組的知識本體，用來描述王教授的事業歷程，並據此建置描述臺灣圖書資訊學領域學者事業歷程的網站，除提供瀏覽、搜尋、關聯等功能之外，亦能透過取用或下載知識本體，使本網站成為鏈結資料的提供者。為了測試網站效率，本研究招募 74 名圖書資訊學研究生與圖書館館員進行四項驗證知識本體關聯性的任務測試，測試結果顯示，在一般性任務的任務一與任務二當中，本研究網站明顯快於對照組的傳統網站；在搜尋語意經過自然語言特殊處理的任務三中則兩者無異；而需要使用者應用超連結的任務四則結果相反。研究結果顯示若要提升本研究知識本體網站的有用性，仍需在介面上有所改善。未來除針對介面加以改善外，還期望能擴展知識本體的收納範圍，建立領域研究者的知識本體。

緒論

隨著資訊科技的突飛猛進，以數位化方式表徵和呈現知識資料已是輕而易舉的事。然而傳統上大多數檔案的內容意涵難以為機器所解讀，由一個機器（程式）所產生的知識資料亦

難以為另一個機器（程式）所攝入（ingest）、再利用（reuse），造成機器間溝通困難。然而這種情形在知識本體（Ontology）、語意網（Semantic Web）和鏈結資料（Linked Data）技術逐漸成熟之後已有所改觀。

Gruber（1993）曾在他的研究中解釋知識本體：「An ontology is an explicit specification of a conceptualization. The term is borrowed from philosophy, where an ontology is a systematic account of Existence.」，亦即知識本體是特定領域中概念化定義的明確規範，知識本體源於哲學，用來系統化解釋何謂「存在」。計算機科學家則將知識本體視為以機器解讀的方式定義特定領域中的基本概念，以及概念間的關聯（Noy & McGuinness, 2001）。運用知識本體技術，有助於人們或軟體代理人分享特定領域的知識結構和促進領域知識的重複利用（Noy & McGuinness, 2001）。

在網際網路和全球資訊網蓬勃發展之後，大量資料在網際網路湧現，無論在資料的發現、取用、整合及再利用皆帶來許多困難。Tim Berners-Lee（2000）認為要解決此一問題，需要將資料以電腦能夠理解的形式發布於全球資訊網環境中，形成一個讓電腦能夠直接或間接處理的資料網（Web of Data），此即語意網（Semantic Web）的開端。語意網的實現必須仰賴多項網路相關技術，包含：URI、UNICODE、XML、RDF、RDFS、OWL、SPARQL等。而為了具體實踐語意網，2006年Tim Berners-Lee更提出了所謂的「鏈結資料」做為實踐語意網的最佳實務，企圖以結構化資料建構全球資訊網的內容，並加以串聯成為一個機器能理解、富含語意和結構化的資料網路（Linked Data, n.d.）。

為展現知識本體、語意網、鏈結資料的具體應用，本研究將建置王振鵠教授事業歷程的知識本體，透過一名具有學術影響力的重要學者角色，去建構圖書領域學者個人生涯發展歷史的知識本體；再進一步建立王振鵠教授的鏈結資料網，透過背後的知識本體技術與結構，建置瀏覽、搜尋、關聯王教授事業歷程的網站。

文獻探討

本研究旨在於以王振鵠教授為研究對象，產生可描述臺灣圖書資訊學者個人發展歷程的知識本體，進一步運用此知識本體，建構具備鏈結資料功能之圖書資訊學領域學者的事業歷程網站。由於採用的技術包含知識本體和鏈結資料，本節針對此二主題的相關文獻進行探討。

知識本體

知識本體源自於哲學領域，別稱有「存有學」、「本體論」或「一般形上學（general metaphysics）」，後來被計算機科學、知識工程所採納，用來使機器（計算機）認知人類知識。知識本體是由概念（concept）和關聯（relation）所組成，概念又可分為實體（entities）、屬性（attributes）等元素，「實體」指的是特定領域中有形或無形的重要事物；「屬性」用以敘述

概念的特性和可能的範圍等；「關聯」則用以說明實體間或實體與概念間的關係（Gruber, 1993；Uschold & Gruninger, 1996；阮明淑、溫達茂，2002）。

知識本體在圖書資訊學領域的主要用途是知識組織。阮明淑與溫達茂（2002）將知識組織定義為，以一定的規則與方法將雜亂或分散的特定知識加以排序、集中、定址，以方便知識的提供、利用與傳播。他們提到，概念是知識的基本單位，也是思維的最小單位，並透過所指事物與語言符號來描述，即語言符號是概念的表達形式，而概念是語言的思想內容，所指事物、語言符號和概念三者之間的關係，常以語義三角來描述。

一般常見的字典、分類表、索引典等工具的基礎概念即為知識組織的應用。而知識組織與知識本體之關聯，須顧慮到今日的數位化資源，在知識組織時需要一個具有多用途、高彈性的智慧型表達工具，達到完整資訊的呈現與檢索，而知識本體自哲學領域提出以來，一直扮演著概念化人類知識的工具，透過資訊技術的結合，可滿足人類對於改善知識組織工具的需求。

Noy 與 McGuinness（2001）在〈Ontology Development 101: A Guide to Creating Your First Ontology〉指出建置知識本體的七大步驟，分別為：（一）定義本體領域與範圍，（二）考慮使用現有的本體，（三）列舉本體中的重要詞彙，（四）定義類別與類別階層，（五）定義類別的屬性，（六）定義屬性的層面，（七）創建實例。

為確定知識本體的效用，需要針對一或多個知識本體進行評估或比較。一般而言，評估知識本體的方法可約略分為下列四種（Brank, Grobelnik, Mladenić, 2005）：（一）與黃金標準（golden standard）比較，（二）將知識本體使用在應用中並評估其結果，（三）與知識本體涵蓋領域之相關資料進行比較，（四）由專家或使用人為判斷知識本體是否符合某些準則、標準或需求。本研究採用第（二）種方法評估所建置之知識本體。

本研究之主要目的在於建置描述王振鵠教授事業歷程的知識本體，該知識本體乃以 VIVO 為核心架構。VIVO 係由康乃爾大學所發展的一套表徵研究者的鏈結資料工具，除了基於全球資訊網的網站、編輯器、內容整合工具與管理系統外，其核心之知識本體運用了所屬機構、研究活動、出版、教學、服務等實體、屬性及關聯來描述學者（Börner, Conlon, Corson-Rikert, & Ding, 2012）。VIVO 的目的著重於描述各學科領域的共同性而非差別性，VIVO 亦不強調涵蓋所有與學者相關的資訊，因此 VIVO 可謂構成描述學者資訊的基礎知識本體，各研究可視需要加以擴充。例如 Chambers 等人（2013）所建置之以英文和荷蘭文描述荷蘭學者的雙語知識本體便是以 VIVO 為基礎加以擴充，容納荷蘭國家學術研究及合作資訊網路系統（Dutch National Academic Research and Collaborations Information System, NARCIS）中的額外資訊。

語意網與鏈結資料在圖書館的應用

Tim Berners-Lee (2000) 曾提出語意網的概念，並在 2006 年指稱鏈結資料即是實現語意網概念的最佳途徑 (Linked Data, n.d.)。鏈結資料的特色在於能透過等價的描述語句，連結兩個不同來源的結構化資料，這樣的描述語句包含了 owl:sameAs 與其他類型的鏈結屬性，而這些屬性是建立在網路本體語言 (Web Ontology Language, OWL) 之下，換句話說，知識本體技術使得鏈結資料具有機器可推理性，兩者密不可分 (Parundekar, Knoblock, & Ambite, 2010)。

鏈結資料已廣泛運用於各種領域，如生命科學、地理、社群網路、媒體等各領域，截至 2014 年 8 月為止，LOD 已收錄至少 1,014 個資料集 (Schmachtenberg, Bizer, & Paulheim, 2014)。鏈結資料在圖書館界亦獲得重視，主要的計畫包括虛擬國際化權威檔 (Virtual International Authority File, VIAF)、美國國會圖書館鏈結資料服務 (LC Linked Data Service: Authorities and Vocabularies)、大英圖書館的自由化資料服務 (free data service) 等。OCLC 的 WorldCat 和歐盟的 Europeana 亦在系統中加入鏈結資料服務。這些鏈結資料服務的重點有的是權威檔或控制詞彙的鏈結資料化，如 VIAF、美國國會圖書館鏈結資料服務；有的則是以書目紀錄為主，發展書目紀錄知識本體或是導入鏈結資料，融合資料元素與資料值，達成資料間的鏈結，包括大英圖書館的自由化資料服務、歐盟的 Europeana 及 OCLC 的 WorldCat 等；有些計畫像是開放型詮釋資料註冊中心 (Open Metadata Registry, OMR) 同時涵蓋了圖書館界相關的多種知識本體 (如：FRBR、FRAD、FRSAD)、控制詞彙 (如：GEM)、內容 (如：RDA、ISBD)、資料元素 (如：Dublin Core、MARC21) 等，是一個提供詮釋資料描述、發掘、儲存與交換的綱要資料庫 (柯皓仁、陳亞寧，2013)。

系統設計與建置

本節說明系統建置過程的所有步驟，依序說明研究對象、系統建置流程，以及展示網站結果。

研究對象

本研究以王振鵠教授的專業歷程為研究個案。王教授投入臺灣圖書館事業發展超過半世紀，一直扮演著相當具有影響力的角色，對臺灣圖書館事業貢獻卓著。

根據國立臺灣師範大學圖書館 (2014) 的統計，從民國 45 年到 103 年，王教授的著作共有專著 43 種，單篇文章 382 篇，另曾指導 36 篇學位論文，而他人傳略更有 39 篇。而根據鄭麗敏 (1994, 1995) 對 1974 到 1993 年間，圖書館學及資訊科學期刊的論文引用分析，王教授的論著是這二十年間被引用次數最多的著者，而王教授所著的《圖書館學論叢》一書更是被引用次數最多的中文個人著作。

除了學術產出外，王教授一生的事業生涯亦影響圖資界深遠。他在民國 66 年至 78 年期間擔任中央圖書館（即國家圖書館）館長，並持續出任圖書館學會多項委員會召集人，不管在教育、行政及組織領導上，都扮演著重要的多重角色。顧力仁（2005）讚揚王教授：「能在公餘研究不輟，質精量豐，論著的內容遍及圖書館學的理論與技術、圖書館事業發展的歷程以及各國圖書館事業經營的方法…，涵蓋面極為廣泛，這種旺盛的研究動機當源自於先生對圖書館事業的熱愛。」

王教授堪稱是臺灣圖書館事業發展的領航人，在擔任國立中央圖書館館長時，建樹良多。宋建成（2005）將王教授對我國圖書館事業的建樹列為八大項：1.完成中央圖書館新館遷建工程；2.推動全國圖書館自動化作業；3.創設漢學研究中心；4.創設資訊圖書館；5.實施中華民國國際標準書號；6.舉辦臺北國際書展；7.促進人文及社會科學資料單位合作；8.舉辦全國圖書館會議等。除此之外，還創編了「圖書館學與資訊科學」刊物，並擔任主編。

王振鵠教授將近半世紀的圖書館學術和專業生涯，造就了如今臺灣圖書館學與圖書館事業的蓬勃發展，而其在職生涯中與各重要單位交流頻繁，脈絡之廣值得探討，故本研究選擇王教授為研究對象。

系統建置流程

本研究以〈Ontology Development 101: A Guide to Creating Your First Ontology〉(Noy & McGuinness, 2001) 作為參考依據，設計王振鵠教授之事業歷程的知識本體，並使用此知識本體建置系統網站，最後再評估本系統是否能滿足圖資領域研究者之需求，以及審視系統在操作上的易用性表現。本研究之系統規劃與建置分為資料蒐集與範圍定義、知識本體規劃設計、資料庫映射檔配置、伺服器與網站建置共四個階段。

一、資料蒐集與範圍定義

本階段蒐集王振鵠教授於學術歷程中之所有產出，蒐集內容範圍分為產出文獻、關聯組織、歷史事件與互動學者四大類。產出文獻包含學位論文、期刊論文、合著文章、專著與出版物等學術文獻。參與組織包含學術機構、政府單位、出版單位等與王教授有關聯性的組織。歷史事件則包含王教授過去生涯中之重大事件，比如圖書館搬遷、獲頒獎項、就職等活動與事件。互動學者之定義為與王教授有過合作發表、共同研究之學者，以及王教授曾經指導過的學生。

資料來源是國立臺灣師範大學圖書館轄下的校史經營組，該單位於慶祝王教授九秩生日時整合了由中華民國圖書館學會、中華民國國家圖書館以國立臺灣師範大學所提供的王教授資料，並且逐一考察權威性（國立臺灣師範大學圖書館，2014）；同時本研究也盡可能地收集王教授之相關歷程資料，以增加知識本體之功能性與網站價值。

本階段需與建置知識本體的第一步驟—定義本體領域與範圍—同步進行，以確保蒐集的資料能和知識本體領域與範圍相契合。

二、知識本體規劃設計

在確認資料與本體範圍後，此階段將參考知識本體設計流程中的後面六個步驟進行：

(一) 考慮使用現有的本體

在知識本體的選擇上，需要考慮到未來資料的互操作性（interoperability），也就是使資料能讓其他應用程式再利用，因此必須選擇受到廣泛使用的現有知識本體作為參考，本研究選擇 VIVO（DuraSpace Organization, 2016）與 DBpedia（Auer et al., 2007）作為知識本體的參考來源。

但考慮到上述兩種知識本體可能與本研究之需求有些微差異，若是無法確切滿足研究上的需求，將造成資料貧乏以及描述不完善等系統功能缺失，因此本研究使用知識本體編輯工具 Protégé（Stanford Center for Biomedical Informatics Research, 2014）進行修改與編輯，使得原本的知識本體屬性皆能符合研究需求。

(二) 列舉本體中的重要詞彙

在確定蒐集到的資料和知識本體範圍後，本研究從資料中將與學者個人事業歷程的詞彙擷取出來，諸如學者、學位論文、組織、事件、學位、領域、出版品和期刊等重要類別詞彙，以及出版、影響、具有身分或指導等介於實體之間的關聯詞彙。

(三) 定義類別與類階層

將前一階段所擷取到的詞彙進行分類，同其所同，異其所異，並且將相似的類別，以一個更大的類目概括，而最上位的類目為事物（Thing）。在〈Ontology Development 101: A Guide to Creating Your First Ontology〉中，整理了三種方式以建置知識本體，分別是由上而下（top-down）、由下而上（down-top）以及整合法（combination）。

由上而下的方法乃是根據一般知識的分類產生知識本體，如生物分類法中的界、門、綱、目、科、屬、種，而對應於本研究的需求，即為 DBpedia 與 VIVO，也就是具有權威性與公信力的知識本體。相反地，由下而上的建置方法，是透過收集該領域的特殊實體，歸類出適用於該領域的知識本體，例如：在前一步所蒐集到的詞彙中，圖書館為一個需要特別關注的特殊單位，故在本研究中需自成一類。最後的整合法顧名思義是同時採用上述兩種方法建置知識本體，除了具有明確的分類依據，同時還能保有領域的特殊性。因此，為了讓本研究能同時保有這些優點，故採以整合法作為建置方法。

本研究透過整合法同時整合 DBpedia、VIVO 以及領域詞彙，並根據前述資料來源爬梳各項資料、定義出知識本體領域與範圍後，產生表 1 所列之類別階層，總計 26 個類別，其中因

為 DBpedia 與 VIVO 的知識本體也參考了許多其他現有知識本體，故表 2 列出各類別參考來源的網域，在這些類別當中不乏由本研究自行創建而未選用其他網域定義的類別（網域縮址為 vh 者），原因是外部來源不一定符合本系統的要求，為了將語意定義得更加明確，故決定自行定義類別，而各項類別的操作型定義於表 3 中逐一說明。

表 1 本研究所設計之知識本體類別階層

階層一	階層二	階層三	階層四	階層五	階層六
Thing (owl)	Agent (foaf)	Person (foaf)			
		Organization (foaf)	Academia Organization (vh)		
	Entity (obo)	Occurrent (obo)	Temporal Region (obo)	Period (vh)	Date-Time Interval (vivo)
			Process (obo)	Event (event)	
	Item (vh)	Publication (obo)	Journal (bibo)		
			Paper (obo)	Journal Article (obo)	
				Thesis (bibo)	
				Articles (bibo)	
	Concept (skos)		Book (bibo)		
			Academic Degree (vivo)		
			Research Area (vh)		
			Organization Type (vh)		
			Identity (obo)		
		Event Type (vh)			

表 2 本研究知識本體中各類別的來源網域

網域縮址	網域	適用對象
owl	http://www.w3.org/2002/07/owl#	泛用類別：Thing
obo	http://purl.obolibrary.org/obo/	實體類別：Period、Entity、Occurrent、Process、Publication、Journal Article、Identity、Papers
bibo	http://purl.org/ontology/bibo/	文獻類別：Journal、Thesis、Books、Articles
foaf	http://xmlns.com/foaf/0.1/	人物類別：Person、Organization、Agent
skos	http://www.w3.org/2004/02/skos/core#	概念類別：Concept
vivo	http://vivoweb.org/ontology/core#	VIVO 自定義類別：Academic Degree、Date-Time_Interval
event	http://purl.org/NET/c4dm/event.owl#	事件類別：Event
vh	http://localhost/VIVOhistory.owl#	本研究自定義類：Research Area、Academia Organization、Item、Event Type、Organization Type、Period

表 3 本研究中各類別的操作型定義

類別	中文	操作型定義	子類別或實例
Thing	事物	指各類型事物，為最上層類目	人、抽象實體、物品、概念
Agent	人	用以包含人或人群的類目	個人、組織
Person	個人	指單人，包含目標學者及其指導學生或合著對象	如王振鵠、張春興
Organization	組織	包含與目標學者有關的組織機構，階層較學術組織上位	如國立臺灣師範大學
Academia Organization	學術組織	意指組織下的學術組織，包含研究團隊或出版部門	如國立臺灣師範大學圖書資訊學研究所
Entity	抽象實體	用以操作抽象實體的類別	時間性實體
Occurrent	時間實體	與時間性有關的實體類別	時間區間、進程
Temporal Region	時間區間	用以表述一段時間的實體類別	時期
Period	時期	指兩個年代點所勾勒成的時間區間	如 1955_1959
Date-Time Interval	日期間隔	本研究以年代作為時間段的間隔	如 1955
Process	進程	指的是不只包含時間的長期時間性實體	事件
Event	事件	此處用以表述學者所參與的大小事件	如王教授擔任國家圖書館館長
Item	物品	具有具體形式之實體	出版物
Publication	出版物	指經由出版單位發行之學者相關著作	期刊、書、論文
Journal	期刊	指曾有刊登過學者著作之連續性出版物	如教育學報
Paper	論文	為該學者著作或指導之論文	期刊文章、學位論文、文章
Journal Article	期刊文章	為目標學者發表於期刊上的文章	如〈中國大學生課外閱讀興趣之調查研究〉
Thesis	學位論文	為目標學者所指導之學位論文	如《杜威十進分類法研究》
Articles	文章	指目標學者所著之單篇文章	如〈圖書館事業發展概述〉
Book	書	指目標學者所著之出版書籍	如《書緣》
Concept	概念	指人類普遍用於分類知識與描繪事物之抽象實體	學術學位、研究領域、組織類型、身分、事件類型
Academic Degree	學術學位	用以區分學位論文級別的概念	如博士、碩士
Research Area	研究領域	用來表述某篇文章主題或某學者之研究主題	如圖書館史
Organization Type	組織類型	用以區分人群組織的各種型態	如學院、出版社
Identity	身分	用以區分學者於本體事件中的各種身分	如學者、公務員
Event Type	事件類型	用以表述該學者所參與事件的各種事件類型	如就職、搬遷

(四) 定義類別的屬性

知識本體中的每項類別與實例都會有其屬性，故此階段乃從現有的本體和取得的資料中擷取出各類別與實例的屬性，除了 rdfs:label、rdf:type 與 rdfs:about 三個所有類別與實例的共同屬性外，共整理出 29 項屬性，如表 4。其中有部分屬性並無採用其他網域的現有屬性，而是選擇由本研究自行定義，原因在於現有的屬性未必能表現出本研究所想表達之語意，以 vh:article_produce 來說，雖然 vivo:Authorship 可以表現其部分含意，但若將 vh:article_produce、vh:journal_produce、vh:has_thesis 三者皆以 vivo:Authorship 表達，在語意上則會遺失許多資訊，這並不是本研究所期望的。

此外，在一般概念裡，一件事物的屬性並沒有上限，全看研究者的研究範圍與主題來決定屬性的種類與數量。因此本研究僅採用前述階段所收集到的資料集，其中曾提及或收錄到的屬性，或是對本研究有幫助的屬性，其餘相關性較低的屬性若是收錄進本體，只會造成主題模糊，以及研究定義不明確，故不予採用。

表 4 本研究各類別之屬性

類別	屬性	意涵
Person	vh:article_produce	產出了某篇文章
	rdfs:advising	指導了某人
	schema:attendee	參與了某事（個人）
	obo:RO_0000087	具有某身分
	vivo:hasResearchArea	擅長某研究領域
	vh:journal_produce	產出了某期刊文獻
	vh:co_authorship	合著關係
Organization	vh:has_thesis	產出了某學位論文
	schema:attendee	參與了某事（團體）
	obo:BFO_0000051	旗下有某部門
	vh:organize_type	是某類型的團體
Period	rdfs:seeAlso	可參照某事物
	vh:period_included	包含了某年分
Event	vh:event_about	關於某種事件主題
	Event:time	發生在某年分
	vh:event_period	發生在某期間
Journal Article	schema:isPartOf	包含在某文章
	dbpedia:firstPublicationYear	出版在某年分
	schema:author	被撰寫於某人
	dbpedia:publisher	出版於某組織
Thesis	dbpedia:publisher	出版於某組織
	dbpedia:firstPublicationYear	出版於某年分
	vh:thesis_degree	具有某學位資格
Articles	dbpedia:firstPublicationYear	出版於某年分
	vivo:has_subject_area	關於某主題

(續表 4)

類別	屬性	意涵
Book	dbpedia:publisher	出版於某組織
	dbpedia:firstPublicationYear	出版於某年分
Research Area	vivo: researchAreaOf	相關學者
	vh:related_research	相關研究

(五) 定義屬性的層面

在傳統資料處理上，大多數的人類語言與符號都是以字串 (string) 的形式儲存，而電腦並無法解析其字詞中的語意與意涵，同樣也無法處理人類知識的邏輯推理，但在結合了知識本體架構的資料中，系統可透過對於實體的屬性內容以辨別實體，而不再僅是字串上的儲存，達到如此成效的關鍵就在於屬性的層面種類，分為物件屬性與資料屬性，透過屬性的差異便可辨識大多數字詞是否為實體。

本階段將上一個階段所定義出的屬性進行分類，分為物件屬性與資料屬性，這關係到描述物件在本研究中被視為實體或字串呈現。如「1989 年」可以被定義為一個時間實體，但也可被定義為年代字串，然而由於本研究將年代視為一個重要關鍵，因此將年代定義為時間實體，使得與年代有關聯之屬性皆為物件屬性。透過上述的方式產出如表 5 的屬性清單，清楚定義出本體內各屬性之功能與定義。

表 5 本研究各屬性之層面類型

屬性	意涵	層面類型
vh:article_produce	產出了某篇文章	物件屬性
rdfs:advising	指導了某人	物件屬性
schema:attendees	參與了某事 (個人)	物件屬性
obo:RO_0000087	具有某身分	物件屬性
vivo:has_research_area	擅長某研究領域	物件屬性
vh:journal_produce	產出了某期刊文獻	物件屬性
vh:has_thesis	產出了某學位論文	物件屬性
schema:attendee	參與了某事 (團體)	物件屬性
obo:BFO_0000051	旗下有某部門	物件屬性
vh:organize_type	是某類型的團體	物件屬性
rdfs:seeAlso	可參照某事物	字串屬性
vh:period_included	包含了某年分	物件屬性
vh:event_about	關於某種事件主題	物件屬性
Event:time	發生在某年分	物件屬性
vh:event_period	發生在某期間	物件屬性
schema:isPartOf	包含了某文章	物件屬性
dbpedia:publisher	出版於某組織	物件屬性
dbpedia:firstPublicationYear	出版於某年分	物件屬性
vh:thesis_degree	具有某學位資格	物件屬性
vivo: researchAreaOf	有某相關學者	物件屬性

(續表 5)

屬性	意涵	層面類型
vivo: has_subject_area	關於某主題	物件屬性
schema:author	被撰寫於某人	物件屬性
vh:related_research	相關研究	物件屬性
vh:co_authorship	合著者	物件屬性

值得注意的是，本研究視近乎所有的屬性為物件屬性，原因在於字串屬性無法仿效物件屬性去連接其他實體，造成日後本體難以擴展範圍，形成後續研究上的困難，雖然 rdfs:seeAlso 在系統中設定為字串屬性，但該欄位所記載之資料為 URL 字串，因此在實際操作上與物件屬性無異。考量到未來增加本體的可擴展性與各實體之間的關聯強度，本研究決定將本體內的屬性大多視為物件屬性處理。

(六) 創建實例

在確定好所有類別與屬性後，方可新增實例至適當的類別下，增加本體價值與資料豐富程度。然而此階段會視後續研究而有所不同，因〈Ontology Development 101: A Guide to Creating Your First Ontology〉的主旨在於建置知識本體，但本體建置僅為本研究系統之初步建設步驟，後續將建置 MySQL 資料庫以放置所有實例，因此除了幾項便於檢視本體結構的實例資料外，其餘資料將於後續階段輸入。

在此階段優先建置的實例如表 6，優先建置的原因在便於設計者檢視本體的關聯結構有無衝突，保證本研究之知識本體具有一定程度的可推論性，從而避免影響日後本體發展，故設計時全程使用了可插件於 Protégé 內的 Pellet 推論引擎即時檢測，經由引擎推論檢測後，確認符合 OWL DL 標準才可進入下一階段。

表 6 本研究各類別之實例

類別	實例	意涵
Academic Degree	Master	碩士
	Doctor	博士
Organization Type	College	學院
	Library	圖書館
	Research Organization	研究組織
	University	大學
	Institute	研究所
Identity	Researcher	學者
	Civil Service	公務人員
	Industry	業界人士
Event Type	Worked	就職
	Establish	創立
	Life	人生事件
	Move	搬遷事件

(續表 6)

類別	實例	意涵
Event Type	Educational	教育事件
	Project	計畫事件
	Official	政府事件

三、資料庫與映射檔配置

在依循〈Ontology Development 101: A Guide to Creating Your First Ontology〉的步驟以及使用 Protégé 修改與編輯後，使用前述所建置的知識本體，設計一份可對應於關聯式資料庫的映射檔。映射檔是由 D2RQ 開發團隊所設計的映射語言 (mapping language) 寫成 (Cyganiak, 2012)，目的在於轉置、操作跟搜尋 RDF，透過 Jena 的 RDF Model 編輯功能以及映射檔，可以轉置出符合研究需求之 RDF 檔案，並完成鏈結資料的轉置工作。接著，因為需要將前面所收集到的王教授資料轉置到 MySQL 中以做後續應用，因此本階段所產出之映射檔，其編譯內容即是轉置 RDF 格式成功與否的關鍵。

(一) 資料庫建置

本研究採用 MySQL 作為存放檔案的資料庫。完成資料庫架設後，即可參考先前設計的知識本體與收集到的資料，手動建置資料表與各項欄位，並將資料逐一輸入至 MySQL 資料庫中，全數輸入後，資料庫內的類別與實例狀況如表 7。

表 7 本研究中的實例數量與範例

類別	中文	實例筆數
Person	個人	44 筆
Organization	組織	20 筆
Academia Organization	學術組織	27 筆
Period	時期	7 筆
Date-Time Interval	日期間隔	92 筆
Event	事件	65 筆
Journal	期刊	73 筆
Journal Article	期刊文章	158 筆
Thesis	學位論文	36 筆
Articles	文章	105 筆
Book	書	44 筆
Academic Degree	學術學位	2 筆
Research Area	研究領域	5 筆
Organization Type	組織類型	5 筆
Identity	身分	3 筆
Event Type	事件類型	7 筆

為了配合 D2RQ 伺服器，在資料庫中的類別與物件屬性是以資料表的方式記錄，而資料屬性則是在類別資料表中以欄位記錄，本研究在 MySQL 建立與類別和物件屬性名稱相同的資料表，完整的實體關聯圖（Entity-Relationship Model）請參見符興智（2016）。在完成各項資料表後，便可開始設計映射檔。

（二）映射檔配置

透過 D2RQ 的映射語言，可以將 MySQL 裡的資料表欄位對應到相對的類別、屬性與實例，進而將一般的字串資料轉成具有知識本體功能的 RDF 資料。

表 8 為本研究映射檔（mapping file）的配置程式碼。程式碼第 1-18 行 @prefix 的主要用途為縮詞，將 URI 縮寫為自定義詞，方便後續程式碼的撰寫。程式碼第 19-26 行 map:database 段落的功能，在於設定 D2RQ 與 MySQL 之間的連結，包含帳號、密碼與位址等參數。

表 8 映射檔配置程式碼

```
1 @prefix map: <#> .
2 @prefix db: <> .
3 @prefix vocab: <vocab/> .
4 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
5 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
6 @prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
7 @prefix d2rq: <http://www.wiwiss.fu-berlin.de/suhl/bizer/D2RQ/0.1#> .
8 @prefix jdbc: <http://d2rq.org/terms/jdbc/> .
9 @prefix vh: <http://localhost/VIVOhistory.owl#> .
10 @prefix foaf: <http://xmlns.com/foaf/0.1/> .
11 @prefix owl: <http://www.w3.org/2002/07/owl#> .
12 @prefix vhc: <http://localhost:2020/resource/class/> .
13 @prefix vhp: <http://localhost:2020/resource/property/> .
14 @prefix skos: <http://www.w3.org/2004/02/skos/core#> .
15 @prefix vivo: <http://vivoweb.org/ontology/core#> .
16 @prefix event: <http://purl.org/NET/c4dm/event.owl#> .
17 @prefix obo: <http://purl.obolibrary.org/obo/> .
18 @prefix bibo: <http://purl.org/ontology/bibo/> .
19 map:database a d2rq:Database;
20 d2rq:jdbcDriver "com.mysql.jdbc.Driver";
21 d2rq:jdbcDSN "jdbc:mysql://localhost/jenadbtest";
22 d2rq:username "root";
23 d2rq:password "*****";
24 jdbc:autoReconnect "true";
25 jdbc:zeroDateTimeBehavior "convertToNull";
26 .
```

在完成與 MySQL 的連結設定後，可開始映射資料。映射檔的撰寫邏輯是由資料庫中的資料表去對映到類別與物件屬性，表 9 中撰寫的內容為對映資料表 Event 到類別 Event 的程

式碼，程式碼第 1 行以#開頭為註記；程式碼第 2 行「event:Event a d2rq:ClassMap;」在於宣告 event:Event 為一項類別；第 3 行「d2rq:dataStorage map:database;」為類別對映到的資料庫路徑；第 4 行「d2rq:uriPattern "Event/@@Event.id@@";」是設定該類別實例的 URI 樣式；第 5 行「d2rq:class event:Event;」是設定該類別在顯示時的 URI，而第 6 行「d2rq:classDefinitionLabel "Event";」則是該類別的定義標籤。

表 9 Event 類別的映射程式碼

```
1 ##### Table Event
2 event:Event a d2rq:ClassMap;
3 d2rq:dataStorage map:database;
4 d2rq:uriPattern "Event/@@Event.id@@";
5 d2rq:class event:Event;
6 d2rq:classDefinitionLabel "Event";
7 .
8 event:Event_id a d2rq:PropertyBridge;
9 d2rq:belongsToClassMap event:Event;
10 d2rq:property rdfs:about;
11 d2rq:pattern "@@Event.id@@";
12 .
13 event:Event_label a d2rq:PropertyBridge;
14 d2rq:belongsToClassMap event:Event;
15 d2rq:property rdfs:label;
16 d2rq:pattern "@@Event.label@@";
17 .
```

在基本資料表設定完成後，接續的內容為該類別的資料屬性，第 8 行「event:Event_id a d2rq:PropertyBridge;」表示類別 event:Event 中的每一個實例都有一個資料屬性為 id；第 9 行「d2rq:belongsToClassMap event:Event;」則說明這項屬性將指向 event:Event 這個類別；第 10 行「d2rq:property rdfs:about;」意指這項屬性將顯示為 rdfs:about，而第 11 行「d2rq:pattern "@@Event.id@@";」則是屬性本身的 URI 呈現方式。

在 D2RQ 的映射檔中，物件屬性也是以資料表的方式儲存，表 10 是物件屬性的對映設定，程式碼第 2 行「map:event_year a d2rq:PropertyBridge;」即為宣告資料表 event_year 是一項屬性連接；第 3 行「d2rq:belongsToClassMap event:Event;」意思為 map:event_year 用於屬性 event:Event；第 4 行「d2rq:property Event:time;」該屬性的呈現 URI 為 Event:time；第 5 行「d2rq:refersToClassMap vivo:date-time_interval;」說明該屬性指向 vivo:date-time_interval；第 6 行「d2rq:join "event_year.Event_id => Event.id";」與第 7 行「d2rq:join "event_year.year_id => date-time_interval.id";」為仿效關聯式資料庫的 JOIN，定義三個資料表之間的欄位參照關係。

表 10 物件屬性映射程式碼

```
1 ##### Table event_year
2 map:event_year a d2rq:PropertyBridge;
3 d2rq:belongsToClassMap event:Event;
4 d2rq:property Event:time;
5 d2rq:refersToClassMap vivo:date-time_interval;
6 d2rq:join "event_year.Event_id => Event.id";
7 d2rq:join "event_year.year_id => date-time_interval.id";
8 .
```

完整的配置程式碼請參見符興智（2016）。在將所有的類別與屬性定義的映射完成後，D2RQ 也將會依照此檔案中的設定將資料庫中各資料表的欄位字串轉換為 RDF 格式的資料，便可透過 SPARQL 端點（endpoint）網站中的 SPARQL 語句進行搜尋與利用這些 RDF 資料，後續才有辦法建置實例以及網站建設。

配置完成後，經 SPARQL 語法搜尋結果顯示，本系統內有 26 種類別、29 項屬性、720 件實例，總計可產生 3,645 條三元組描述句，若不計算 rdfs:label、rdfs:about、rdf:type 三種資料屬性，也仍有 1,421 條透過物件屬性所產生的三元組描述句。

四、伺服器與網站建置

Apache Web Server 是目前最廣泛使用的伺服器軟體之一。在進行資料導入及轉置工作之前，必須先將伺服器建置完成，過程包含了軟體安裝以及伺服器與資料庫連結。在完成上述工作後，即可參考先前設計的知識本體以及映射檔案的架構，完成轉置工作。

為了讓轉置完成的 RDF 資料能透過網路讓使用者方便瀏覽與使用，因此還需要利用網頁呈現的方式，讓使用者能透過瀏覽器，從遠端就能取得本系統資料。因為結合了免費工具 sparqlib.php 的功能，使用者可以經由 PHP 語法所撰寫的網站，從網頁傳遞 SPARQL 查詢語句到 D2RQ Server 的端點進行搜尋，再透過 D2RQ 的映射檔轉換 MySQL 的資料成 RDF 格式，傳回並呈現於網頁上，此外網站還融合了 LodLive，這是一套以 JavaScript 所寫成的 RDF 視覺化網頁工具，來輔助使用者透過圖形化檢索的方式，了解各實體之間的關聯性。

網站結果

本研究的最終結果為王振鵠教授鏈結資料網，透過背後的知識本體技術與結構，本網站可達到瀏覽、搜尋、關聯，以及 SPARQL 端點四項功能，接下來將說明每一項功能是如何幫助使用者在資料間進行瀏覽與查找，達到比一般傳統網站更方便的使用價值。

一、瀏覽

本功能的設計目的在於提供使用者傳統階層式的瀏覽模式。網站設計者使用 SPARQL 語言配合 PHP 語法，將網站內所有類別列出，如圖 1 所示。使用者可以先推估目標實體之類別，

在點選代表該類別的超連結之後，即可瀏覽該類別底下的所有實體，選擇目標超連結之後，便可檢索到該目標的所有屬性資訊。

透過傳統瀏覽模式的設計，可以讓尚未熟悉知識本體結構的使用者從一般瀏覽模式的超連結跳轉中探索出知識本體網狀結構的概念，進而掌握知識本體的其他呈現方式與檢索模式。

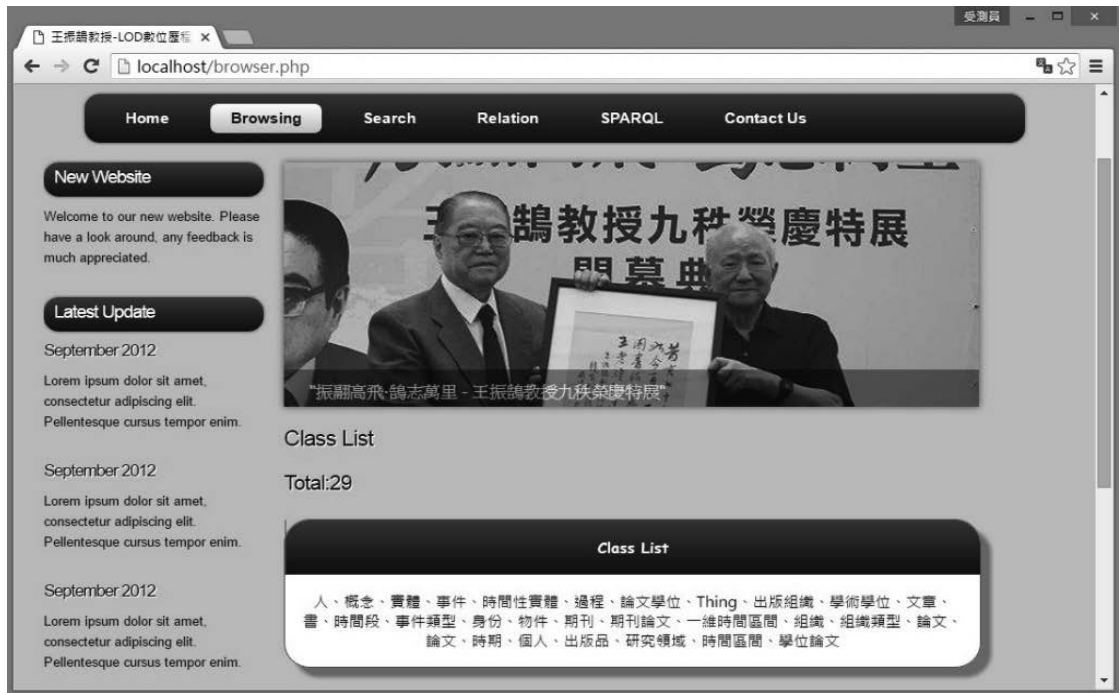


圖 1 網站瀏覽功能畫面

二、搜尋

此功能頁面可讓使用者透過位於畫面中央的文字欄，輸入個體的標籤（label）屬性，進行簡單的關鍵字查找，網頁畫面如圖 2，與瀏覽相比，因為多了一個搜尋的步驟，因此省去了人眼逐一比對目標的繁瑣過程，加快搜尋速度。



圖 2 網站搜尋功能畫面

網站同時搭配了使用 JavaScript 所設計的 LodLive 圖形化瀏覽功能，讓使用者在找到目標實體後，可透過圖 3 所示的圖形化瀏覽方式瞭解各實體之間的關聯網絡，且在探索過程中，容易發現原本資料上不會特別記載的實體關聯，增加單筆資料的資訊價值。

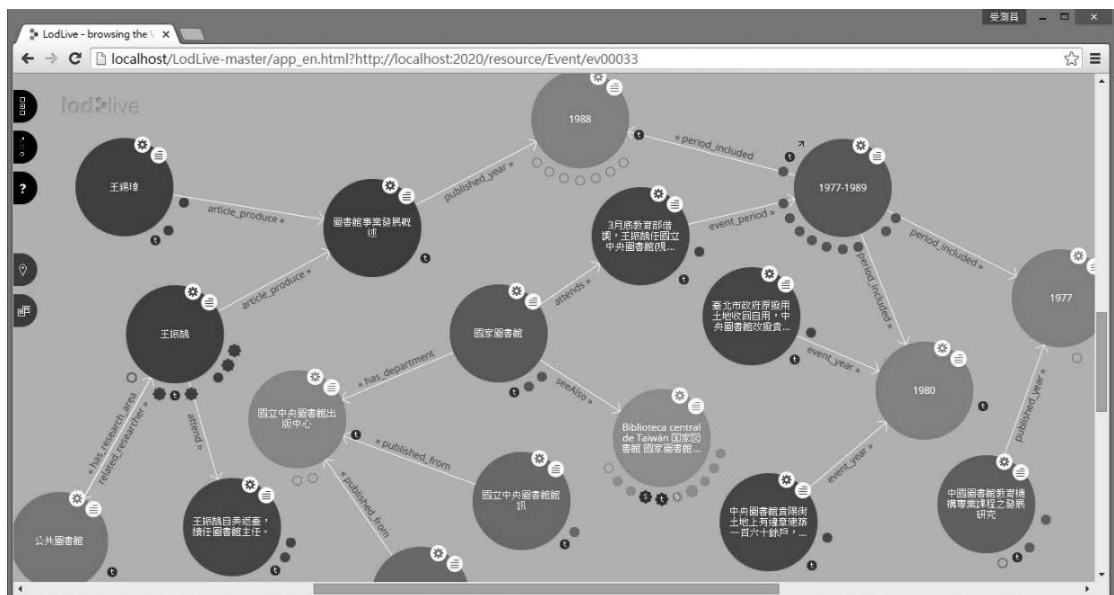


圖 3 LodLive 工具瀏覽畫面

三、關聯



圖 4 網站關聯功能頁面

關聯查找功能是本網站基於知識本體架構才得以實現的查找功能。在功能頁面中，使用者可以透過圖 4 頁面中的下拉式選單，選取符合需求的關聯選項，共有七項關聯可供選擇：1.王教授在某職位期間出版的書籍；2.王教授參與的某事件其當下指導的學位論文；3.王教授參與某事件當下出版的合著出版品與合著者；4.王教授在某職位期間的相關事件；5.王教授在某職位期間出版的文章；6.王教授在某職位期間指導的學位論文；7.王教授的某著作出版時所指導學生之論文。

在選擇好關聯後，使用者可於下方的文字欄，輸入欲查找之實體關鍵字，按下搜尋後，所有符合條件的實體關聯將如圖 5 所示，以條列式呈現。



圖 5 網站關聯功能查詢結果頁面

上述七項關聯是設計者在評估知識本體內容後所選擇的代表性關聯，並採用 PHP 結合 SPARQL 語言所撰寫，故關聯功能是以設計者預設之 SPARQL 語句進行查找，除了避免大多數使用者不熟悉 SPARQL 語言的困擾，同時省去了撰寫 SPARQL 查詢的繁瑣過程。此外每一項回傳的實體名稱都以超連結呈現，使用者可用超連結進一步瀏覽該實體之細節屬性頁面，提供初步回傳的關聯中未提及之資訊，方便使用者延伸查詢。

四、SPARQL 端點

倘若關聯查找中預設之 SPARQL 語句無法滿足使用者需求，本網站亦提供 SPARQL 端點搜尋頁面，畫面如圖 6。該功能讓使用者能自行輸入 SPARQL 語言進行查找，補足設計者無法設想周到之處，使網站的查找功能更加完善。

結合上述的四大搜尋功能，使用者可依照需求查找資源，而知識本體結構也能幫助使用者了解實體之間的關聯脈絡，提供一個不同於傳統網頁的瀏覽方式。

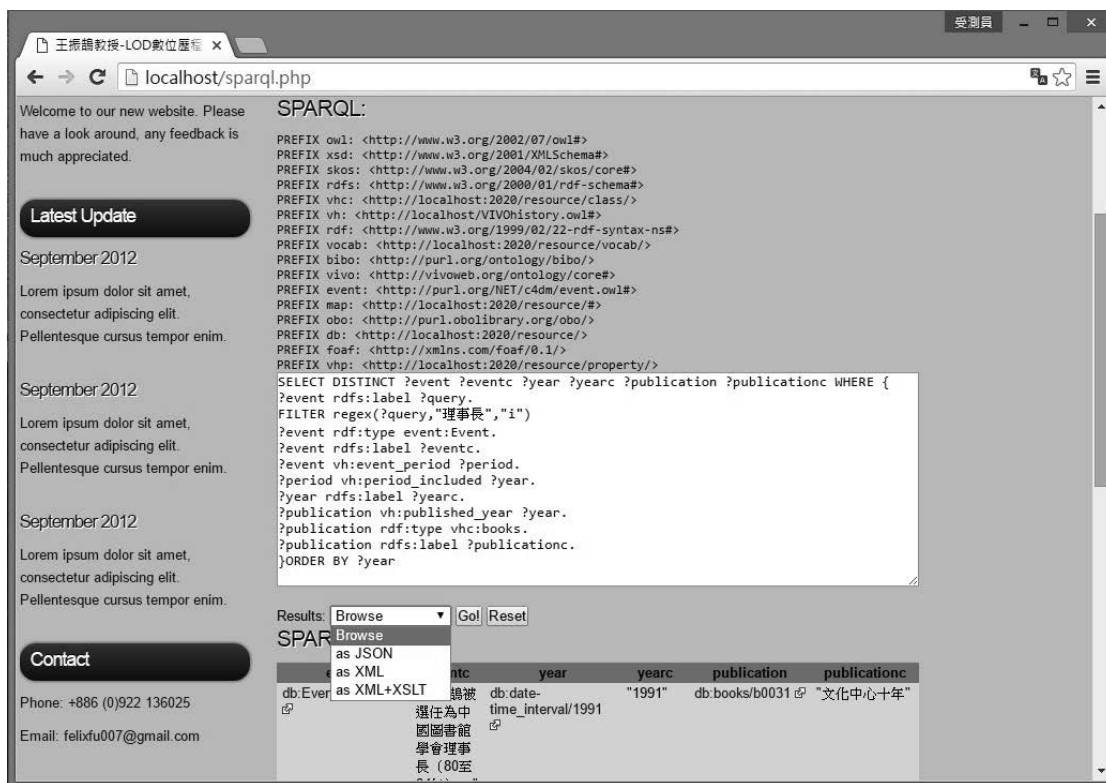


圖 6 網站 SPARQL 功能頁面

實驗結果與系統評估

為了驗證本系統網站在使用性上較未含有知識本體的一般網站更為有效率，且更容易發現個體之間的新關聯，因此採用「將知識本體使用在應用中並評估其結果」的知識本體評估方法 (Brank, Grobelnik, Mladenici, 2005)。本研究將實驗分為實驗組與對照組，實驗組選用的工具即為本研究的最終研究結果網站，而對照組則是國立臺灣師範大學圖書館校史經營組所設計的「王振鵠教授九秩榮慶特展網站」(國立臺灣師範大學圖書館，2014)，該網站是國立臺灣師範大學校史組為慶祝王教授九秩榮慶特展活動所設計的，網站內容包含王振鵠教授的大事記要、教育事業歷程、相關文章與出版品清單等，內容豐富，且皆以表格清單呈現，有助於人眼查找。

受試對象

本研究的受試對象來源為國立臺灣大學、國立政治大學、國立臺灣師範大學以及私立天主教輔仁大學四間學校，取其圖書資訊學研究所 50 名已畢業與未畢業之研究生，以及圖書資訊學系的 18 名已畢業與未畢業之大學生，加上 6 名正於圖書館工作之館員，主要是這類受試

對象：(一) 擁有圖資領域的基礎概念，能以一個具有基礎知識的資訊需求者角度，去檢視系統的使用性；(二) 對於圖資學術主題有基本認知，能審視知識本體的推論結果是否符合圖資主題之間的關聯性；(三) 為當前圖資領域的學習者或實務工作者，能測試本系統平臺能否幫助研究人員探知古今學術領域發展之脈絡。工具的受測人數依照受測者學歷進行均分，以確保受測者的學歷背景不影響測試結果，實驗組與對照組各為 25 名研究生、9 名大學生以及 3 名館員，一組 37 人，共兩組，總計 74 人。

實驗設計

本研究實驗是參考亞利桑那大學圖書館網站的使用性十點評估簡易模式(呂淑惠,2002)來設計實驗，實驗分為四個階段：

一、向受測者說明實驗與基本資料填寫

向受測者說明實驗流程以及實驗大綱，讓受測者了解整個實驗步驟，因為對實驗流程不熟悉所造成的錯誤很有可能影響研究結果，因此必須說明仔細，並讓受測者提問，以排除任何會影響到實驗結果的變數。

二、讓受測者熟悉系統

說明完實驗流程後，讓受測者進行一次練習任務，任務內容是使用指定的工具平台查找指定答案，過程計時但不限時，讓使用者習慣在計時的狀態下執行任務。而在使用者完成任務後，會讓其自由探索平台內容，直到使用者主動告知熟悉平台使用模式後，方可進入下一階段。

三、進行實測

實測有四項任務，任務內容與練習任務相仿，差異在於有限時五分鐘的查找時間，且每一題會針對網站功能進行微調。

四、測後問卷與回饋

在實測後請受測者對於實驗中的特殊情形進行問卷填寫以及口頭回饋，透過實驗人員與受測者的討論，釐清特殊情況的發生原因，以及可深入追究的特殊議題。

實驗環境

本節將詳細說明整項研究的測試環境，包含測試工具、測試設計與測試任務，盡可能地表達出本次實驗的各種潛在變數，不僅讓後續的研究人員有個可參照的方法架構，也能提供一些可以改善研究的契機。

一、測試工具（主要有螢幕錄製工具和測試主機，說明如下）

1. 螢幕錄製工具：為了收集足夠的數據以輔助日後的研究，本研究使用螢幕錄製軟體 Open Broadcaster Software (OBS) 記錄整個實驗過程。本研究除了錄製測試中的螢幕內容外，

同時也利用視訊鏡頭測錄受測者的受測情形，並且記錄於同一檔案中，以方便日後回顧測試過程。

2. 測試主機：為了減少硬體差異所造成的變數，研究者提供同一台主機讓受測者進行實驗。瀏覽器也統一使用 Chrome 的無痕模式，避免在測試過程中，瀏覽器暫存所影響到的任何變數。

二、測試設計

實驗一開始會先請受測者填寫基本資料，在完成基本資料後，會對受測者進行實驗解說，包含了時間限制、計時方式、螢幕錄製與側錄的內容、回答與發問的方式，為了幫助受測者加深印象，螢幕上也會提供小卡幫助受測者閱讀說明。

在受測者完整了解實驗規則後，會先請受測者進行測試題，採取計時而不限時的方式，讓受測者習慣在計時的情境下操作，當受測者找到任務目標時，會詢問是否足以了解該平台的使用方式，若已了解，則進入實測，若尚未了解，受測者可以持續探索平台直到熟悉為止。

待受測者了解平台操作模式後，並開始進入實測，以限時五分鐘的方式控制受測者的搜尋時間，而計時方式是由受測者自行操作開始與結束，目的是減少測試人員與受測者因溝通不良所造成的誤差。過程中，受測者每完成一項任務，測試人員就會記錄任務搜尋時間、完成與否及完成數，記錄完畢才進入下一項任務。

在四項實測任務結束後，測試人員會請受測者在問卷上，針對方才使用的平台給予想法與意見，除此之外，也會對於實驗的過程中，任何與測試人員預期相違的行為進行詢問，諸如操作上的遲疑、任務語意的誤會或是其他特殊情形，這些內容都是日後延伸研究的重要資訊。

三、測試任務

任務的設計將直接影響到實驗結果，本研究將任務設計的重心放在查找實體之間的關聯性上，共設計四道任務來測試平台的關聯查詢功能。

1. 任務一：任務題目為「王振鵠教授在被選為中國圖書館學會理事長的這段期間，共出版了哪些書籍（book）？」，目的在測試目標平台對於學者職位與其出版書籍關聯性的表達功能。此任務題目與練習題相似，因此可以檢測出使用者在熟悉檢索路徑後，是否會有效縮短檢索時間。
2. 任務二：任務題目為「王振鵠教授在擔任國立中央圖書館（現國家圖書館）館長期間發生了哪些事件？」，目的在測試目標平台對於學者在特定職位期間所發生事件的表達能力。因為本研究的對照網站是以條列式表格呈現年代表，因此研究人員預先假設此任務若使用對照網站進行查詢，應當會有良好的表現結果，但若本研究網站能比對照網站擁有更好的成效，則表示知識本體技術能輔助使用者更容易產生關聯之間的聯想。

3. 任務三：任務題目為「王振鵠教授在發表《圖書館學論叢》一書的當年，還指導了哪些學位論文？」，目的在測試目標平台對於學者在特定著作發表的同時，其指導的學生論文關聯能否清楚呈現。
4. 任務四：任務題目為「王振鵠教授在被選為中國圖書館學會理事長的期間，其指導的學位論文大多出版於哪些系所？」，本題目的設計較為特殊，使用者使用本研究平台的關聯搜尋後，無法立即取得答案，需要進一步點開搜尋結果的連結，才有辦法確認搜尋結果是否正確。透過這樣的設計，可以檢測使用者是否會過於依賴搜尋功能，而未能進一步探索。

實驗結果

本節內容為實驗測試後的結果，分別根據各項任務結果進行評估，並採以成對樣本 *t* 檢定，檢測同一項任務在「王振鵠教授鏈結資料網」與「王振鵠教授九秩榮慶特展網站」中被執行時，使用者投注在檢索關聯上的平均時間，是否有顯著差異。為方便呈現，下文便以「LOD 網站」簡稱「王振鵠教授鏈結資料網」，以「傳統網站」簡稱「王振鵠教授九秩榮慶特展網站」。

參與實驗人數分別為 LOD 網站 38 人，傳統網站 37 人，為了取得正確數據，本研究將任務失敗的數據剔除，任務失敗的情形包含（一）未成功搜尋到目標關聯、（二）未完整搜尋到所有目標關聯、（三）回報錯誤目標以及（四）搜尋時間過長。剔除失敗數據後，僅計算任務成功的檢索投注時間，而任務失敗的情形，將針對發生任務錯誤的平台採用次數統計，以計算個別平台的失誤次數。

表 11 為任務一的統計結果，在執行任務一時，兩組受測群體在檢索速度上有顯著差異，且 LOD 網站快於傳統網站。在錯誤次數上，LOD 網站的錯誤次數為 2，傳統網站的錯誤次數為 1。

表 11 任務一的統計結果

測試平台	觀察值個數 (人)	平均數 (秒)	標準差	<i>p</i>	<i>t</i>
LOD 網站	35	50.42	40.33	0.014**	1.994
傳統網站	36	73.03	35.45		

表 12 為任務二的統計結果，在執行任務二時，兩組受測群體在檢索速度上有顯著差異，且 LOD 網站快於傳統網站。而 LOD 網站的錯誤次數為 0，傳統網站的錯誤次數為 2。

表 12 任務二的統計結果

測試平台	觀察值個數 (人)	平均數 (秒)	標準差	<i>p</i>	<i>t</i>
LOD 網站	37	54.53	40.31	0.004**	1.994
傳統網站	35	86.86	53.10		

表 13 為任務三的統計結果，在執行任務三時，兩個受測群體在檢索速度上未達到顯著差異。在錯誤次數上，LOD 網站的錯誤次數為 3，傳統網站的錯誤次數為 4。

表 13 任務三的統計結果

測試平台	觀察值個數 (人)	平均數 (秒)	標準差	<i>p</i>	<i>t</i>
LOD 網站	34	72.23	62.49	0.442	2.00324
傳統網站	33	62.61	37.52		

此項任務的結果與前兩項不同，經過研究人員仔細觀察實驗錄像後，發現使用者多半無法清楚明白 LOD 網站中關聯查詢的功能語意，此任務的正確路徑是使用「王教授的某著作出版當下與其指導學生之論文」的關聯查詢進行查找，但將 SPARQL 查詢語意轉為自然語言時，會產生個人主觀邏輯的差異，造成他人閱讀困難，導致使用者在瀏覽時的錯誤判斷。比如實驗中，部分使用者未發現此項功能為正確路徑，轉而採用其他較為迂迴的搜尋方式，使得整體查詢時間拉長。

表 14 為任務四的統計結果，結果顯示兩個受測群體在檢索速度上有顯著差異，且傳統網站快於 LOD 網站。在錯誤次數方面，LOD 網站的錯誤次數為 1，傳統網站的錯誤次數為 4。在前面的任務設計中曾經說明，本任務有經過特殊設計，在前幾項任務中，使用者在使用關聯查詢後，網頁上會立即呈現任務目標，無須進一步動作，然而在此任務中，設計人員將任務目標設計成必須點擊超連結，連結到該目標物件的詳細屬性頁面下，才可觀測到目標答案，簡言之，與先前的三項任務不同的地方在於，使用者必須多一個步驟，才能看到答案。

表 14 任務四的統計結果

測試平台	觀察值個數 (人)	平均數 (秒)	標準差	<i>p</i>	<i>t</i>
LOD 網站	36	107.08	69.63	0.000***	1.99647
傳統網站	33	48.82	16.68		

研究人員在觀察實驗時，發現部分使用者在使用關聯查詢後，未進行下一步點選，反而轉去其他與搜尋無關的功能頁面進行查找，導致實驗結果差異懸殊。針對此現象，研究者於實驗結束後，對受測者進行簡單的詢問，發現受測者操作錯誤的情況發生於：(一) 受測者認為 LOD 網站的關聯搜尋過於便利，可能為陷阱題；(二) 既使在練習時已知曉 URL 的超連結功能，但執行本任務時，依然疏忽超連結功能，而無法進行下一步操作；(三) 誤會呈現方式不是研究人員所期望的，轉以其他呈現方式；(四) 發現結果沒有自己所想的答案，很可能是選錯功能或是無解。

總結上述各項任務的結果，在任務失敗次數上，LOD 網站任務失敗總次數為 6 次，傳統網站失敗總次數是 11 次。任務一跟任務二中，受測者在 LOD 網站的搜尋時間顯著快於傳統網站，任務三中兩者則無顯著差異，任務四則是傳統網站明顯快於 LOD 網站。透過測後回饋的內容中可以發現，造成任務三與任務四如此情況的原因，在於 SPARQL 語言轉自然語言的困難，以及實體屬性過多時如何取舍回傳的屬性或是設計更友善的介面引導使用者完成任務。

結論與建議

本研究透過知識本體技術，以王振鵠教授為對象，建置適用於描述圖書資訊領域學者事業歷程的系統網站。首先在描述圖書資訊學者的個人事業脈絡方面，除採用 DBpedia 與 VIVO 兩個普遍應用且具有權威價值的知識本體，增加與外部知識本體連結時的共通性與便利性外，亦經由探索王振鵠教授的個人資料，萃取描述圖書學者事業歷程的知識本體。而最後經 SPARQL 語法搜尋結果顯示，本系統內的知識本體有 26 個類別、29 項屬性、720 件實例，總計可產生 3,645 條三元組描述句，若不計算 rdfs:label、rdfs:about、rdf:type 三種資料屬性，也仍有 1,421 條透過物件屬性所產生的三元組描述句。

完成知識本體設計後，下一步便是建置知識本體網站。本研究運用了 Apache、MySQL、PHP、D2RQ、SPARQ 工具以及 LOD Live，提供使用者便利的瀏覽、搜尋、關聯等功能。此外還能將知識本體內容轉化為鏈結資料，除了可以透過 SPARQL 端點取用資料外，也能將知識本體內容轉為 RDF 格式下載，達到資料加值的目標。

為了驗證本系統是否能幫助使用者有效率地發現實體之間的關聯，幫助人文學者探索歷史脈絡，本研究選擇了具有相似功能的王振鵠教授九秩榮慶特展網站進行比較，並取得了 74 位受測者的使用紀錄。在四項搜尋任務中，本研究系統有兩項任務明顯快於傳統網站，原因在於 SPARQL 語法的推理功能，可縮減人類思考搜尋路徑的過程，讓搜尋更快速且錯誤率更低。但另外兩項任務則是持平與明顯較慢，經過測後詢問受測者的結果，除了介面設計外，也牽涉到搜尋語言轉為自然語言的困境，以及人們對於機器代為推理所產生的結果也無法完全信任。這些結果除了證明本研究所發展的系統在效率上優於傳統網站之外，同時也發現了知識本體網站在介面建置上值得注意的細節，以及未來將會遭遇的困難。

本研究未來的發展方向有三，說明如下：

- 一、持續修正知識本體，並擴展知識本體收錄範圍與深度：知識本體的創建與維護並非一蹴可幾，本研究遵循〈Ontology Development 101: A Guide to Creating Your First Ontology〉(Noy & McGuinness, 2001) 所揭櫫的步驟建置王振鵠教授的知識本體，雖有初步成果，但未來仍須不斷修正；此外，本研究僅以王振鵠教授之著作以及三個機構所保存的王教授資料為研究內容。倘若收錄範圍增加至與王教授相關人事物的其他描述，則其脈絡探

索更可擴展到人際互動，或是王教授對於圖資領域的間接影響程度。

二、應用於其他領域及學者：除了探索圖資領域學者的個人脈絡外，更可以透過本研究所使用的建置方法，擴展知識本體範圍，建立領域研究者的知識本體；甚至可套用至其他領域學者上，取得跨領域的學者互動與其相互影響，對於研究學術歷史脈絡的數位人文學者而言，會是一套便於探索資源的便利工具。

三、人性化語意搜尋介面：在觀測實驗過程與結果時可以發現，在未了解知識本體結構的背景之下，多數使用者無法自如地使用內部資源，諸如搜尋失敗、誤解語意邏輯或是對搜尋結果的不信任等問題，都會導致知識本體網站上的使用困難，因此目前急需投入資源於語意工具的開發，讓知識本體的應用更為便利與多元，才能促使人類的知識流通更有效率。

致謝

本研究為行政院科技部專題研究計畫「以 FRBR、Linked Data、Crowdsourcing 設計圖書館線上公用目錄系統之研究」(NSC 102-2410-H-003-121-MY3) 之部分成果。

參考文獻

- Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., & Ives, Z. (2007). Dbpedia: A nucleus for a web of open data. In K. Aberer et al. (eds.), *The Semantic Web, Lecture Notes in Computer Science (vol. 4285, pp. 722-735)*. Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-76298-0_52
- Berners-Lee, T. (2000). *Weaving the Web: The original design and ultimate destiny of the World Wide Web*. New York, NY: HarperBusiness
- Börner, K., Conlon, M., Corson-Rikert, J., & Ding, Y. (2012). VIVO: A semantic approach to scholarly networking and discovery. *Synthesis lectures on the Semantic Web: theory and technology*, 7(1), 1-178.
- Brank, J., Grobelnik, M., & Mladenić, D. (2005). *A survey of ontology evaluation techniques*. In Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005) (pp. 166-170).
- Chambers, T., Shah, S., Urankar, A., Kalyan, V., Scharnhorst, A., Reijnhoudt, L., Rideour, L., Guéret, C., & Ding, Y. (2013). Bilingual researcher profiles: Modeling Dutch researchers in both English and Dutch using the VIVO ontology. *Proceedings of the Association for Information Science and Technology*, 50(1), 1-4.
- DuraSpace Organization. (2016). *VIVO connect-share-discover*. Retrieved from <http://vivoweb.org/>
- Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199-220.
- Linked Data (n.d.). *What is the relationship between Linked Data and the Semantic Web? Frequently Asked Questions (FAQs)*. Retrieved from <http://linkeddata.org/faq>
- Noy, N., & McGuinness, D. L. (2001). *Ontology development 101 A guide to creating your first ontology*. Retrieved from http://protege.stanford.edu/publications/ontology_development/ontology101.pdf
- Parundekar, R., Knoblock, C. A., & Ambite, J. L. (2010). Linking and building ontologies of linked data. In P. F. Patel-Schneider et al. (eds.), *The Semantic Web – ISWC 2010, Lecture Notes in Computer Science*

- (vol.6496, pp. 598-614). Springer, Berlin, Heidelberg. doi; 10.1007/978-3-642-17746-0_38
- Schmachtenberg, M., Bizer, C., & Paulheim, H. (2014). *State of the LOD Cloud 2014*. Retrieved from <http://linkeddatacatalog.dws.informatik.uni-mannheim.de/state/>
- Stanford Center for Biomedical Informatics Research. (2014). *Protégé*. Retrieved from <http://protege.stanford.edu/>
- Ushold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. *The Knowledge Engineering Review*, 11(02), 93-136.
- 阮明淑、溫達茂 (2002)。Ontology 應用於知識組織之初探。 *佛教圖書館館訊*，32，6-17。
- 【Yuan, Ming-Shu, & Wen, Dar-Maw (2002). Ontology yingyong yu zhishi zuzhi zhi chutan. *Information Management for Buddhist Libraries*, 32, 6-17.】
- 呂淑惠 (2002)。大專院校圖書館網站經營管理之研究 (未出版之碩士論文)。國立臺灣大學圖書資訊學研究所，台北市。
- 【Lu, Su-Huei (2002). *Dazhuanyuanxiao tushuguan wangzhan jingying guanli zhi yanjiu* (Unpublished master's thesis). Department and Graduate Institute of Library and Information Science, National Taiwan University, Taipei.】
- 宋建成 (2005)。王振鵠教授與我國圖書館事業。 *圖書館學與資訊科學*，31 (2)，39-45。
- 【Sung, Chien-Cheng (2005). Professor Chen-Ku Wang and the Librarianship of Taiwan. *Journal of Library & Information Science*, 31(2), 39-45.】
- 柯皓仁、陳亞寧 (2013 年 11 月)。鏈結資料在圖書館的應用。在全國學術電子資訊資源共享聯盟主辦，海量資料：學術研究新境界，台北市。
- 【Ke, Hao-Ren, & Chen, Ya-Ning (2003, November). *Linked Data and Its Application in the Library*. Symposium conducted at the Consortium on core Electronic Resources in Taiwan (CONCERT), Big Data and Academic Research: A New Horizon, Taipei.】
- 符興智 (2016)。以知識本體建置圖書資訊學領域學者的事業歷程網路平臺－以王振鵠教授為例 (未出版之碩士論文)。國立臺灣師範大學圖書資訊學研究所，臺北市。
- 【Fu, Hsing-Chih (2016). *Ontology-Based Platform for Librarianship Development - A Case Study of Professor Chen-Ku Wang* (Unpublished master's thesis). Graduate Institute of Library & Information Studies, National Taiwan Normal University, Taipei.】
- 國立臺灣師範大學圖書館 (2014)。振翮高飛鵠志萬里－王振鵠教授九秩榮慶特展。檢自 <http://archives.lib.ntnu.edu.tw/exhibitions/ChenKuWang/>
- 【National Taiwan Normal University Library (2014). *Zhenhe gafei guzhi wanli - Wangchenku jiaoshou jiu zhirong qing tezhan*. Retrieved from <http://archives.lib.ntnu.edu.tw/exhibitions/ChenKuWang/>】
- 鄭麗敏 (1994)。近二十年來臺灣地區圖書館與資訊科學期刊論文引用參考文獻特性分析 (上)。 *教育資料與圖書館學*，32 (1)，94-118。
- 【Cheng, Li-Min (1994). Characteristics of Cited References in the Journal Articles of Library and Information Science in Taiwan from 1974 to1993. *Journal of Educational Media and Library Sciences*, 32(1), 94-118.】
- 鄭麗敏 (1995)。近二十年來臺灣地區圖書館與資訊科學期刊論文引用參考文獻特性分析 (下)。 *教育資料與圖書館學*，32 (2)，210-238。
- 【Cheng, Li-Min (1995). Characteristics of Cited References in the Journal Articles of Library and Information Science in Taiwan from 1974 to1993. *Journal of Educational Media and Library Sciences*,

32(2), 210-238.】

顧力仁（2005）。永遠秉持誠與恆的信念：王教授振鵠先生論著述要及其學術思想。《圖書館學與資訊科學》，31（2），5-13。

【Ku, Li-Jen (2005). The Brief Introduction of Professor Chen-Ku Wang. *Journal of Library & Information Science*, 31(2), 5-13.】

Ontology-Based System for Librarianship Development -- A Case Study of Professor Chen-Ku Wang

Xing-Zhi Fu

Graduated Student, Graduate Institute of Library and Information Studies,
National Taiwan Normal University, Taiwan (R.O.C)
E-mail: felixfu007@gmail.com

Hao-Ren Ke

Professor, Graduate Institute of Library and Information Studies,
National Taiwan Normal University, Taiwan (R.O.C)
E-mail: clavenke@ntnu.edu.tw

Keywords: Chen-Ku Wang; Librarianship; Linked Data; Ontology; Semantic Web

【Abstract】

Professor Chen-Ku Wang is considered to be one of the most influential personage for librarianship in Taiwan. He achieved great accomplishments during his terms as the Director General of National Central Library; meanwhile, he educated many outstanding students. This study exploits ontologies and Linked Data to analyze Professor Wang's important events and achievements in his professional life. A resultant ontology comprising 26 classes, 29 attributes, 720 instances, and 3,645 triples were established to portray the career path of Professor Wang. A website that empowered users to browse and search the ontology and find relationships between ontology entities was created accordingly. It offers the utilization and download of the ontology, and in this manner, the website can also play the role of a Linked Data provider. To test the efficacy of the website, 74 participants in the field of library and information were recruited for four tasks to verify the relevance of ontology. The result showed that in the first two general tasks, the website developed by this study was significantly faster than the traditional website. The task statement of Task 3 was specially processed by natural language; therefore, there was no difference in the results. Task 4 required users to retrieve answers through hyperlinks, and the test result showed that the traditional website is faster. The results of this study reveals that our ontology-based website needs to be refined for improving the ontology utilization, and

which is one important future work of this study. Other future works include expanding the scope of the ontology, and apply the ontology to other LIS or non-LIS scholars.

【Long Abstract】

Introduction

With the rapid advancements of information technology, digitized representation and presentation of knowledge data have been very common. However, traditionally, it is difficult for machines to interpret the contents and meanings of most of files. It is also difficult for another machine (program) to ingest and reuse the knowledge data generated by one machine (program), which leads to the difficulty in communication among machines. However, this situation has been changed after the gradual maturation of Ontology, Semantic Web and Linked Data.

In order to demonstrate the feasible applications of ontology, Semantic Web, and Linked Data, this study developed the knowledge ontology of career path of Professor Chen-Ku Wang. This study developed the ontology of personal career development history of a scholar in the field of library and information science through an important scholar, and further developed the Linked Data Web of Professor Chen-Ku Wang, in order to create a system website that facilitates the browsing, searching, and associating of Professor Wang's career path through the knowledge ontology technologies and structures behind.

Professor Chen-Ku Wang has devoted himself to the development of libraries in Taiwan for more than 50 years. Professor Chen-Ku Wang is regarded as the most influential scholar in the field of library science in Taiwan. He made a lot of contributions during his service as Director General, National Central Library, as well as cultivated many outstanding scholars and practitioners. He played multiple important roles in education, administration, and organization leadership.

Methodology

This study used <Ontology Development 101> (Noy & McGuinness, 2001) as reference to design the knowledge ontology of career path of Professor Chen-Ku Wang, and used this knowledge ontology to develop a system website. In the end, this study evaluated whether this system can meet the needs of researchers in the field of library and information science and reviewed the operational ease of use performance of this system. The system planning and development of this study included four stages totally – data collection & scope defining, planning & design of knowledge ontology, database mapping

file configuration, and server & site building.

Data collection & scope defining

At this stage, the information related to career path of Professor Chen-Ku Wang were collected. The scope of collection included four major categories: literature output, affiliated organizations, historical events, and scholars of interactions.

Planning & design of knowledge ontology

The main process of planning and design of knowledge ontology in this study are as follows:

(1) Considering to use the existing ontology

This study took into account the interoperability of data, and selected VIVO (DuraSpace Organization, 2016) and DBpedia (Auer et al., 2007) as the reference sources of knowledge ontology. In addition, this study used knowledge ontology editing tool, Protégé (Stanford Center for Biomedical Informatics Research, 2014) for revising and editing.

(2) Enumerating important vocabularies in ontology

After confirming the scope of collected data and knowledge ontology, this study extracted important categorical vocabularies of scholar's personal career path from the data, such as scholar, thesis/dissertation, organization, event, degree, field, publication, and journal, as well as associated vocabularies among entities, such as publication, influence, identity or instruction.

(3) Defining categories and categorical hierarchy

This study integrated DBpedia, VIVO and vocabularies in the field, and defined the field and scope of knowledge ontology according to the inspection of various data mentioned above to generate the categorical hierarchy of 26 categories.

(4) Defining the properties and dimensions of categories

Every category and example in knowledge ontology has its own property. Therefore, at this stage, this study extracted the property of each category and example from the existing ontology and data available. This study proposed a total of 29 properties, and divided them into object properties and data properties.

Database and Mapping File Configuration

Mapping file is written using mapping language developed by D2RQ development team (Cyganiak, 2012). The purpose of mapping files is to transform, operate, and search RDF. RDF files meeting the research needs can be converted using Jena's RDF Model editing function and mapping files, and the transpose of Linked Data can be completed.

Server and Site Building

This study used Apache Web Server as the website server, and combined tools, such as MySQL, PHP, sparqllib.php, D2RQ Server, and LodLive, to complete site building.

Results and Evaluation

The final result of this study is the Linked Data web of Professor Chen-Ku Wang. This website can achieve four functions, browsing, searching, associating, and SPARQL endpoints. The screen of the homepage is shown in Figure 1.

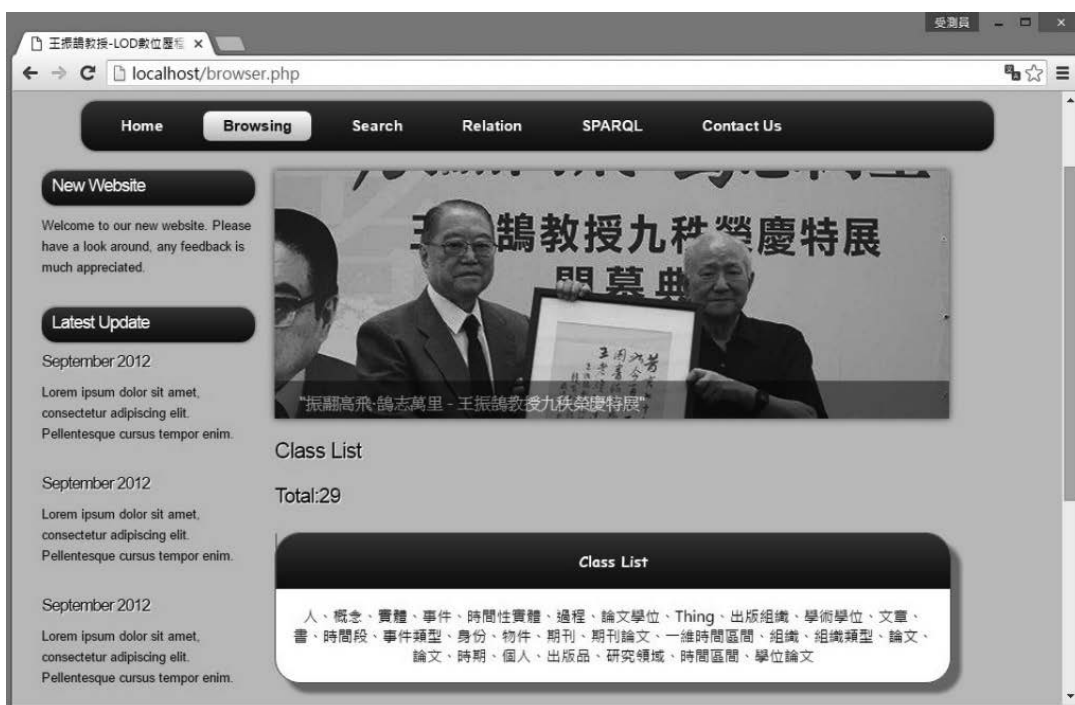


Figure 1 Screen of Website Browsing Function

In order to verify that this system website is more effective than that of a corresponding general website without knowledge ontology and users can more easily discover the new connections among individuals, this study used the knowledge ontology evaluation method that applies a knowledge ontology to an application and evaluates its results (Brank, Grobelnik, Mladenić, 2005). This study enrolled 74 graduate students of library and information science and librarians to test the tasks of knowledge ontology connection. The test results showed that, in task 1 and task 2 of general tasks, this research website was significantly faster than traditional website of control group. In task 3 where semantics were specially processed using natural language, there was no difference between the

experimental group and control group. In task 4 where users' application of hyperlinks was required, the general website without knowledge ontology outperformed the system website.

Conclusions

This study used knowledge ontology technology to select Professor Chen-Ku Wang as the object to develop the system website suitable for describing career development of scholar in the field of library and information science. The knowledge ontology developed in this study included 26 categories, 29 properties, and 720 cases, with a total of 3,645 RDF triples. The website developed in this study provided users with functions, such as convenient browsing, searching, and associating, and converted knowledge ontology content into Linked Data. In addition to obtaining data through SPARQL endpoint, this website can also convert knowledge ontology content into RDF format for downloading to achieve the objective of value-added value. In terms of system evaluation, this study enrolled 74 participants. In the four searching tasks, the system website in this study was significantly faster than traditional websites in two tasks. The reason was that the grammatical reasoning function of SPARQL can shorten human thinking and search paths to speed up the search and reduce error rate. However, there was no difference and the system website was significantly slower in other two tasks. After the testing, this study asked the subjects about the results and discovered that, in addition to interface design, the predicament of conversion of search language into natural language was also involved. Besides, human beings could not fully trust the results generated by machine's reasoning, either. These results showed that, the efficiency of the system developed in this study was superior to that of traditional websites in some tasks. This study also discovered the details worthy of attention in interface building of knowledge ontology websites, as well as the difficulties to be faced in the future.

There are three future development directions for this study: (1) continuing correcting knowledge ontology and expanding the collection scope and depth of knowledge ontology; (2) applying knowledge ontology to other fields and scholars; (3) developing humanized semantic search interface.

[Romanization of Chinese references is offered in the paper.]