

植基知識本體之公部門資訊整合

Ontology-Based Government Information Integration

朱雨其

行政院環境保護署
環境監測及資訊處
ycchu@epa.gov.tw

黃素梅

東吳大學
資訊科學研究所
94756019@scu.edu.tw

楊鍵樵

國立台灣科技大學
電子工程研究所
ccyang@et.ntust.edu.tw

摘要

本文倡議運用知識本體來驅動資訊整合的作業架構及其相關機制，由於知識本體論是以概念的表徵為核心，所以在這個層次上的整合作用可以具有一定程度的語意互通效果。我們發展一套系統化的作業流程及相關演算法；首先我們從現存的知識環境及已知的部分資訊源中構築個別性的本體原型，而後將各個原型本體作校正及合併等調整，以得到在特定應用領域中具有整合效用的知識本體。以上述知識本體為基底，我們發展二階段式之公部門資訊整合架構；知識本體建構階段負責針對公部門領域之特定應用構建整合性知識本體，而實體資料整合階段則結合資料倉儲技術，參用知識本體之概念規格與作業準則，實際執行資料萃取、轉換及匯入工作。我們以公部門之環境保護領域實作建置整合性環境資料庫，驗證本文倡議方法之可行性及具體效益。

關鍵詞：知識本體、異質性、資訊整合、資料倉儲、電子化政府

ABSTRACT

This paper proposes a lightweight ontology-driven framework (LODF), which is the major component to drive the information integration process. An integrated ontology can be seen as a kernel of concept representation; hence the semantic interoperability among heterogeneous information sources can be achieved at the concept level. We develop a systematic process and related algorithms to support the construction of the integrated ontology. First, we capture the ontologies from existing knowledge sources to form distinct prototype ontology. Then, those distinct ontologies are aligned and merged to construct an integrated ontology for a specific application domain. Based upon the LODF, we develop a two-phase architecture for government information integration. The *ontology construction phase* is in charge of developing an integrated ontology

which is dedicated to the specific domain of government. The *physical data integration phase* combines data warehouse technologies and consults the conceptual specifications and operation guidelines of the integrated ontology to perform data extraction, transformation and loading. We have been implementing an integrated environment databases for verifying the feasibility and performance of proposed approach.

Keywords: ontology, heterogeneity, information integration, data warehouses, e-Government

1 緒論

以網際網路為主軸的多元化資訊時代已然成形，在可預見的將來，所有的個人裝置均將與網路連結，而所有的資訊服務也將透過網路遂行。當然，公部門的政務管理工作及對人民所提供的服務，更不能自外於這波潮流。藉由網路無遠弗屆的特性，未來公部門所提供的各項服務勢將展現與傳統全然不同的嶄新風貌。

以往政府部門大都依個別機構的權責劃分資訊處理作業，也就是從「政府觀點」發展，不論各政府機關間，或是機關內部各單位間，係以獨立方式進行業務電腦化流程，於是形成許多「煙囪式」(stove-piped)系統。這些系統彼此不盡相容，同時重複建置資料。民眾要擷取資料時，需分別從各業務主管機關進行查詢檢索[1]。例如，有民眾欲申辦「工廠設立許可」，他可能要查詢縣市政府建設局、環保局、工務局(建築管理)等單位，甚或是中央政府經濟部(商業司、工業局)環保署等單位的網站才能取得完整資料。而目前所謂的「政府入口網站」，只是將各機關的網址集中放在某個網頁，對資訊整合並無助益，民眾還是得大費週章地到各個網站查詢資料加以整理後才能應用。未來政府部門應該以「民眾觀點」發展，也就是資訊整合的觀點，提供主題導向(subject-oriented)的整合與服務，

才能提昇為民服務的水準。

本文擬探究公部門如何運用資訊與網路科技，進行資料整合之相關問題及其解決方案，包括學理面的論述、實作面的系統架構發展及實作驗證。第二節將討論異質性的特徵及其對資料品質的影響，並對資訊整合過程容易衍生的問題提出具體說明。第三節說明領域應用的知識本體發展，我們提出 LODF 作為本體發展的理論基礎架構，同時研擬系統化的領域知識本體發展程序與流程。第四節則從系統設計及實作二方面說明資訊整合系統架構及現階段實作的成果。第五節作結論並闡明未來發展方向。

2 異質性與資訊整合課題

資訊整合及分享 (information integration and sharing) 技術近年來受到學術界與實務界廣泛重視。傳統的整合工作大部分是屬於「系統」層面的整合，但在實際運用上，「資料」的整合才是使用訴求的重點。其次，現行多數研究或產品大體上只是在系統表現層次的整合 (presentation integration)，對資訊使用者的助益極為有限。欲達成資訊基礎面的整合，尚有相當多課題待克服，但綜觀現行實務應用上，以資訊異質性 (information heterogeneity) 未能有效調合對資訊整合系統的成敗影響最大，導致資訊分享困難。本節將針對異質性資訊之特性進行剖析，討論現行的處理方法及其面臨的窘境。

2.1 資訊異質性

異質性問題是資訊運用到某個程度後的一種必然現象。早期的研究大都針對資料模式的整合，例如將網路式、階層式及關聯式的資料模式轉換或映射到一個共通的模式，但是近年來由於多媒體與超文件的出現，資料型態較往昔複雜許多，一般而言，資料異質性問題可以分為下列二個面向討論 [9]：

- 語法 (syntax)：通常指相同概念的屬性，以不同的方式表現。例如「日期」，可能有的以 day/month/year 表示，有的以 month/day/year 表示；又以「長度」而言，其所用到的單位可能有的以英吋為單位計量，有的以公分為單位。
- 語意 (semantics)：不同資料來源中其資料屬性在內涵結構或命名方式的衝突效應，例如「同義異型」（相同物件但命名不同）或是「同型異義」（不同意義的物件卻有相同的名稱及格式）。尤有甚者，在特定的情境

或應用系統，某個詞語所代表之意義各不相同。

一般而言，語法層面的異質性較為容易處理，通常設定若干標準格式或協商共通性作法後，可以達到一定程度的整合效果 [11]。但是語意層面的異質性則不容易處理，不可能藉由某種標準來「迫使」所有的資料來源遵守規定，在實務面是不可行的。其次，資訊系統所處理的資料型態日益複雜，語意異質性不能只考量單純的文數字資料型態，越來越多的應用系統要求能有效率且正確的整合圖形、影像甚或是聲音資料，更突顯語意異質性問題的重要性。

資訊異質性對資料品質 (data quality) 之衝擊更是顯著，資料品質不佳或錯誤的資料，可能衍生不可預期的後果，致使管理階層對整合性資訊系統失去信心。近年來有許多公私企業導入資料倉儲 (data warehouses) 作為資料整合平台，但資料倉儲內的資料品質問題卻是影響應用成果的重要因素之一。據估計約有六成以上的資料倉儲系統宣告失敗，其主要的原因就是因為存在於系統中的資料品質問題，系統開發人員沒有給予充分的時間與努力來解決、淨化資料品質的問題而導致整個系統失敗 [5, 6, 12]。

資料品質也可以從語法及語意二個面向討論。語法性的資料品質問題指的現實世界中有此項資料，但其內容與資料庫中所儲存的資料不同。例如「地址」，將台北市「凱達噶蘭大道」，儲存成台北市「凱達噶欄路」，這種資料錯誤很容易被辨識，處理上也有較具體明確的作法。但若系統記成是「高雄市凱達噶蘭大道」，便不容易處理，因為高雄市並無凱達噶蘭大道，這類資料與現實世界無法勾稽的現象，則屬於語意性的資料品質問題。

不論是語法性或是語意性的資料品質問題，因著網路上資料的飛速增長已成為當前公私企業部門亟待處理的課題。換句話說，資料品質是組織內任何資訊系統是否能夠有效運作發揮實際功效的一項關鍵因素。特別是當我們要整合來自不同機構部門的資料時，資料品質問題的困難度與挑戰性較傳統單一目的性的資訊系統更高。但是在目前絕大部分的相關研究與實作發展中，大都將焦點著重於整合性資訊系統的查詢檢索功能的提昇，卻忽略了最基本的資料品質問題。

2.2 資訊整合面對的問題

綜合前節所描述的資訊異質性相關困境及影響，本文嘗試回答以下問題：

- 在資訊整合過程中，不同來源的資料語意異

質性其內在與外在的原因為何？而這些語意上的衝突對資料品質有何不利之影響？現行處理類似問題的方法與工具其優劣為何？是否有具體可行之改善空間？

- 如何在顧及各資料來源的自主性原則下，處理語意異質性問題？如何藉由語意調合作業平台，達成資訊整合的目標？其次，政府部門採行語意調合技術其利基為何？實作面有些應行注意事項？

有論者以為透過「標準」制定並強制各個資料來源遵行，就可化解資料異質性的衝突，但這種作法只能處理語法面的異質性問題，對於語意異質性，並不能發揮作用，主要是因為各個資料來源通常擁有絕對的自主性 (autonomy)，這種情形在公部門更是普遍。是以各種語法層面的格式標準規範，只能解決「外在」的資料衝突，對於「內在」的或是「隱性」的資料衝突，未有助益。而這類資料衝突卻是影響資料品質的最大因素，也是極具挑戰性及實用性的課題。

3 領域應用之本體發展

本體論 (ontology) 是哲學領域中用以探討真實世界中事務存在的本質課題。近年來，計算機科學領域廣泛地引用本體論的概念及原理，針對人工智慧、知識工程及資訊整合等課題，提出具學理性及實作性的應用 [2, 13, 14]。本節將提出一個針對特定領域應用的簡易型知識本體發展架構，並以此架構為基底，發展以資訊整合為標的之知識本體構建程序及相關作業機制。

3.1 LODF

針對資訊整合之目的，必須要能有操作簡便而且能達到一定用途目標的知識本體。換言之，投入在建構知識本體的成本，包括勞力、物力及其它投資，應該降至最低，而且建構過程必須配合不同的用途及不同的使用者，分階段完成。從實務角度看，我們可以採取下列方式：首先我們提供一個環境讓參與的使用者先建一個「原型本體」 (prototype ontologies)，原型本體的來源可以從資料庫綱目或是 XML 文件的 DTD 等這類既有的資訊源中萃取；其次，在原型本體中所形成的概念詞彙必須與上層的共通性本體，例如 Cyc [8], WordNet [10]，或是某些特定領域中的辭彙性質的本體，例如 GEMET¹ 進行

¹歐洲環境總署 (European Environmental Agency) 制訂的「通用多語環境詞彙」 (General European Multi-Lingual

alignment)。校正的功能在於提供不同本體間的語意一致性 (semantic consistency)，具有語意調合的效果，校正過後的本體稱為「校正本體」 (aligned ontologies)。舉例來說，對於某個「工廠地點」的描述，在公部門中不同機關有不同的概念性描述及辭彙用語，工商管理部門用「工廠地址」、工務部門用「工程地點」、環保部門用「污染源」或「污染場址」等。經過校正後，可以根據統一調整成「場址」 (site)。將這類校正本體經由分階段地合併或組合後，就形成「整合性本體」 (integrated ontology)。

我們將上述程序以正規化方式定義成「簡易本體驅動架構」 (Lightweight Ontology-Driven Framework, LODF)，這裡所謂的簡易係指本體建構中儘可以採行特定領域中容易獲取共識的知識來源，例如各類辭彙書、資料字典、詞彙編及分類標準等。是以針對特定應用領域 D 的簡易本體驅動模型，標示為 $L(D)$ ，可以定義成是一個五項值組 (5-tuple) 的結構，表示如下：

$$L(D) = (R, O_p, O_{int}, f, g)$$

其中

$R = \{r_1, r_2, \dots, r_n\}$ 是一組由領域 D 的各種背景知識所組成的資源集合 (a set of resources)，對於資源 $r_i \in R$ ， r_i 可以是一個資料庫的綱目、一組 XML 文件資料、甚或是一群領域專家。

$O_p = \{p_1, p_2, \dots, p_n\}$ 是對應於 R 所組成的一個原型本體集合。

O_{int} 表示植基於 O_p 的整合性本體。

f 是一個映射過程 (mapping process)，其作用記為 $R \xrightarrow{f} O_p$ ，表示 R 中各個資源 r_i 的知識，藉由 f 產生出 O_p 中相對應的原型本體 p_i 。

g 是一個整合過程 (integrating process)，其作用記為 $O_p \xrightarrow{g} O_{int}$ ，表示藉由針對 O_p 執行一連串的程序，例如本體校正、本體組合等， g 可以導出一個整合性本體 O_{int} 。

有關 LODF 更詳細的說明及其功能的論述詳見 [4]，LODF 的整體作用可表成下式：

$$\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_i \end{pmatrix} \xrightarrow{f} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_i \end{pmatrix} \xrightarrow{g} O_{int}$$

3.2 發展資訊整合之知識本體

晚近對於知識本體的發展與構建等課題已有甚多的研究文獻 [7]；但以發展方法之完整性而言，

Environmental Thesaurus)，目前已有 19 國語言，廣獲環境保護領域採行作為語意交換標準。http://www.eea.eu.int/

Uschold 倡議的四階段方式似較為完整 [15]：

1. 確認目的與範圍：尤應辨明所欲發展的知識本體其預期用途、是否有特殊性 (specialization)？各種可能的情境、及詞語的集合等，同時應決定發展過程之精細度 (granularity)，以有效控制發展之進度與範圍。
2. 建構知識本體：可以分成下列三部分工作
 - (a) 本體徵集 (ontology capture)：尋求領域知識，不斷與前階段互動，隨時調整以契合發展之目的與範圍。
 - (b) 本體編碼 (ontology coding)：將領域知識加以結構化、組織化及規格化。
 - (c) 整合既有本體 (Integrating existing ontologies)：重複運用既有本體之相關成果以加速發展時程。
3. 評估：確認及驗證前述成果。
4. 建立上述各階段之指導方針 (guidelines)

我們將上述 2(a) 及 2(c) 加以推展擴張，導出系統化之知識本體發展與整合程序，如圖 1。此程序針對特定領域知識本體的整合過程，較上述 Uschold 倡議之四階段方式更為週延，在操作上也更簡易。尤其在公部門領域中，因為各機關橫向（如縣政府之各局室）及縱向（如行政院環保署與縣市政府環保局）之關係錯綜，我們相信以圖 1 之作業程序應有利於公部門之資訊整合。以下我們扼要說明圖 1 中各步驟之內涵及執行方式。

1. 擇定參與整合的資訊源、相關的標準、法規及詞語分類規則等：此步驟主要係標定領域中可能持有知識的來源；以公部門為例，可以是各種法令規章、標準作業程序及各類檔案分類規則等，這些來源可能係數位化資源（資料庫系統、網頁或檔案資料系統）、紙本資料，甚或部門中某位職員或是某個工作團隊。
2. 知識萃取：此步驟從上述標定的來源中擷取知識，用以構築相對應的原型本體。部分現行工具如 WebODE [3] 等可以協助執行本步驟工作，但相對「簡易」之目的而言，這類工具顯得過於沉重而不適用，我們正進行發展一個簡便的工具軟體，稱為 **OntoBuilder**，用以專注處理關聯式資料庫及 XML 文件的知識本體萃取，因為公部門有愈來愈多的資訊源屬於這二類型態。圖 2 描述 **OntoBuilder** 的操作方式。

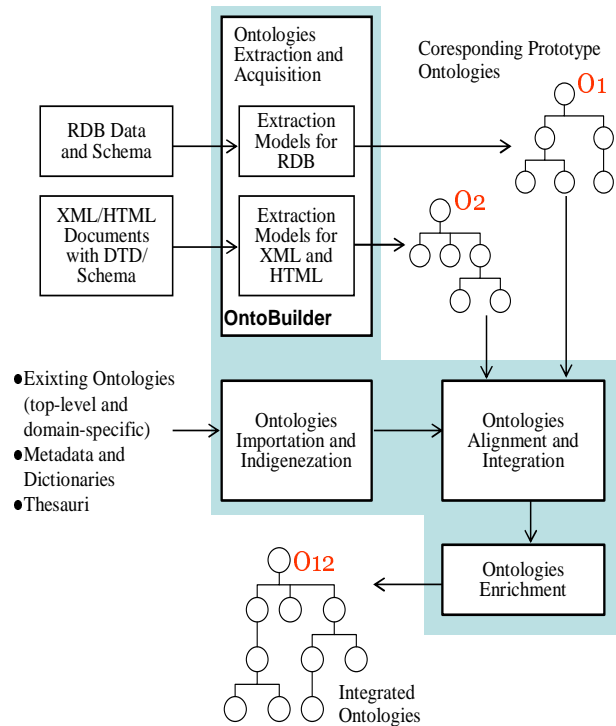


圖 1: 針對特定特領域知識本體之整合程序

3. 匯入 (importation) 及本土化 (indigenization)：特定領域現存的知識本體，它可能存在於其它團體甚或其它地域國家，此步驟加以輸入並作適當調整使其契合本地環境使用，這種現象在公部門極為常見。例如環境保護領域之知識本體，就有關詞語定義與術語分類方面，現行 OECD² 國家已對環保工業有相當完整之分類，而絕大部分歐洲國家均已採納 GEMET 作為環境詞語之標準定義。這些均可視為知識本體的一種，在應用上只要加以若干調整裁剪，對建構新的知識本體亟有助益。舉例來說，如果我們欲就環境保護領域採行 OECD 的分類架構，只要將有關「森林管理」部分作調整，因為在臺灣公務體系的「森林管理」，目前係屬於農政單位，而非環保機關之權責，利用匯入及本土化程序，可以大幅提昇本體構築的效率。
4. 本體校正與整合：針對原型本體、匯入及本土化後之本體，以及部分上層共通性的知識本體，先進行辭彙用語及術語定義等一致化；其次，進行本體彼此間的合併組合等工作，藉以構成特定應用領域中單一整合性的知識本體。

² 經濟合作發展組織 (Organization for Economic Cooperation and Development)

5. 本體強化 (enrichment)：傳統作法多利用知識本體表達概念的分類或是對某個概念的屬性略加描述，我們希望再進一步，對已經整合的知識本體，再探求其額外的知識，用以強化本體的適用性。這在公部門應用上尤其重要，例如前述各機關單位橫向與縱向間的關係，即便已有某種程度的整合性知識本體，但隨著時間及空間的變異（如行政組織、業務或法規的調整），必須在事先預為設定相關前題條件。

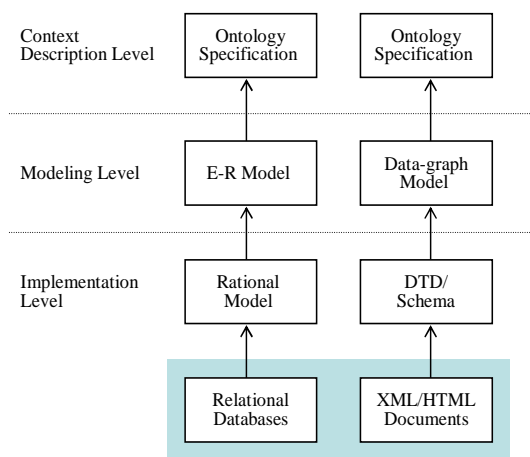


圖 2: OntoBuilder 的操作方式

4 系統架構及實作

圖3係我們規劃之系統發展架構。從作業程序方面，大體上可分為上半部的「知識本體建構階段」及下半部的「實體資料整合階段」；從組成面來看，包括人員、資訊來源、相關作業機制及軟體工具及發展平台等。

4.1 二階段式發展架構

知識本體建構階段主要指構築「公部門整合性知識本體」及建立相關回饋作業機制。若以縣市政府為例，這個階段主要由組織內的資訊人員及各個不同的業務單位人員（亦即領域專家）參與，同時要有更客觀的公部門人員（可以是上級機關人員或第三公正團體人員）執行系統回饋機制。本階段主要的工作環境是「語意調合作業平台」，除了操作性的介面外，也包括3.2節所提的OntoBuilder及其它商業化的知識本體編輯工具。語意調合作業平台的主要作用係運用知識本體建構過程，將公部門各機關單位存有的資訊異質性問題加以弭平。這裡所稱的「調合」涵括3.2

節提及的校正、整合及強化等程序。當然，在實作上這些程序可以視實際情況作彈性調整，但基本上，其執行若能依循圖1的先後順序，可以獲得較一致性效果。

其次，我們特別強調「人員」在這個階段的重要性，由於公部門的組織形態及文化與一般企業有所差異，是以參與人員表現之良窳可能影響知識本體構築的執行效率及正確性，我們認為參與人員除了應來自各相關部門外，也應同時顧及不同層級，包括決策人員及基層業務人員。

就實體資料整合階段而言，資料倉儲技術近年來對資料整合已有初步的系統化作法，但若單純以資料萃取、轉換及載入（extraction, transformation, and loading, ETL）工具直接將資料由各個資訊源匯入單一的資料倉儲，則匯入的資料是否合宜使用？其正確性如何？這些問題在現行資料倉儲技術尚難獲得有效的解決方案。我們倡議在ETL階段，必須以先前階段建構的知識本體作為「基底」，亦即ETL在執行資料的匯入或刷新（refresh）工作時，必須參用知識本體的概念規格及作業準則，如此對整合後資訊之適用性及語意調合才有正面助益。

至於整合性資料倉儲的資料存取的介面，依循介面設計的「分眾」原則，分成一般性使用者介面及特殊性操作介面二項。一般性介面主要提供民眾及工商企業取用經過整合的資訊，例如，有民眾欲查詢如何申辦工商登記，此介面可提供完整的訊息，也就是將工務（都市計畫）、建設、環保等相關規定一次提供給民眾，不須要民眾到各個機關的網站查閱。特殊性操作介面則提供公部門特定人員作為系統回饋資料查閱及參用的管道，此介面的設計除傳統SQL查詢語句外，應以圖表等視覺化工具，強化資料表現方式。

4.2 實作成果

我們運用上述架構，以環保署及相關單位間的環境資料，實作「整合性環境資料庫」³。圖4(a)是對應於知識本體建構階段的簡要示意，我們主要取用目前各級環保機關現有的資訊來源，包括空氣品質監測資料、水質監測資料及列管的污染源（工廠、畜牧場、醫院...）等資料，並與GEMET進行本體校正、整合，建立符合臺灣國情的簡易型「環境保護知識本體」，而後結合資料倉儲技術，進行實體資料整合。

系統硬體環境以HP Proliant DL-380系列作為主要伺服器平台，並搭配磁碟陣列作為資料儲存媒體。軟體環境以Win-2003 Server為作業系統，Microsoft SQL-2000 DBMS作為資料倉儲管

³<http://edb.epa.gov.tw/>

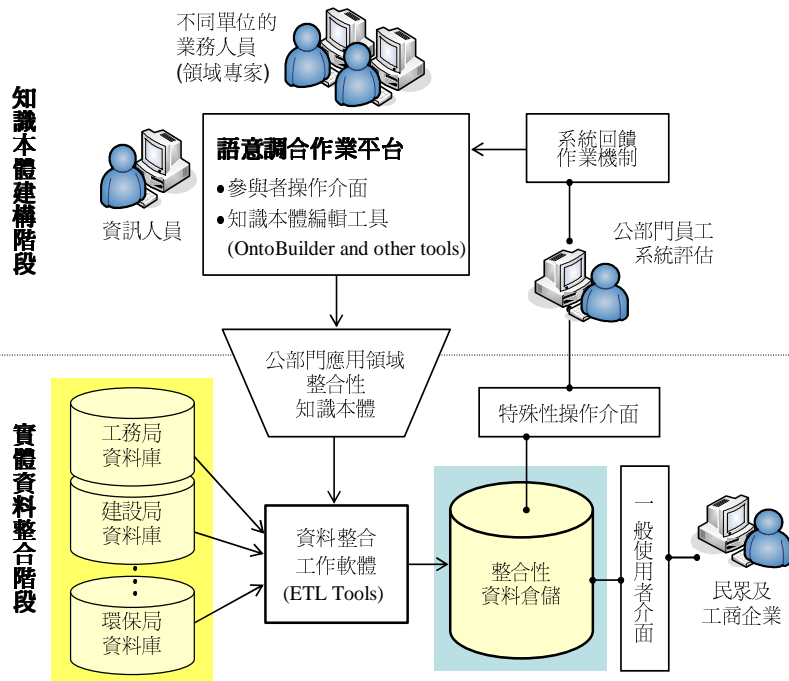


圖 3: 以知識本體為基底的公部門資訊整合架構

理平台，開發工具則以 ASP, Visual Basic 為主。此外針對地理資料及部分衛星遙測影像等空間性資訊，則運用 ESRI 公司之 ArcIMS 及 ArcSDE 開發工具及發展平台。圖 4(b) 係環境資料庫首頁，目前約有 200 萬筆資料可提供線上查閱。

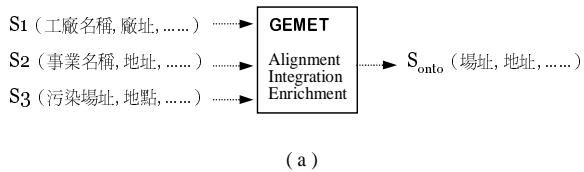


圖 4: (a). 資訊源 Schema 配合 GEMET 以調合語意異質性 (b). 環境資料庫首頁及其提供之功能簡要說明

5 結論及未來發展方向

本文針對公部門資訊整合的相關課題進行深入探討；為有效調合不同部門間資訊異質性問題，我們以公部門領域知識本體為基底，倡議一種可實際操作的資訊整合作業程序及相關配套措施。綜合以上各節所述，我們認為本文可陳現以下具體之成果與貢獻：

1. 就公部門提供全民資訊服務及運用資訊科技所面臨的問題與困境，進行深入的剖析探究。特別是針對資訊異質性所衍生的相關問題，作了組織性及系統性的描述與說明，可為公部門資訊整合工作標定明確方向。
2. 正規化方式定義以知識本體為基底的資訊整合架構，有利於參與人員資訊整合工作之相關人員通盤性掌握整體架構及作業程序，可以避免傳統資訊整合工作因過於冗瑣而造成之「見樹不見林」的現象。同時，藉由數學語句的表達方式，可以提供一致性的描述及表達效果。
3. 改良 Uschold 四階段式的知識本體構建程序 [15]，發展更為週延之作業程序及流程。提出 ontology indigenization 概念，使得現存之本體可以高度被重用 (reuse)，不僅可加速本體構建時程，同時對本體的完整性及適用性亦有正面助益。為驗證此一概念之可行性，

我們引入 GEMET，並加以本土化成為適用臺灣地區的「環境保護領域知識本體」，基本上獲致相當良好的成果。

隨著「電子化政府」工作之推展，未來整合型之公部門資訊服務勢將成為政府便民服務的主流業務，同時也是政府部門未來重要的施政項目。我們認為在遂行資訊整合的各階段，下列課題仍有相當空間可加以進一步探討研究：

1. 如何確保整合性資訊的資料品質？單一資料來源的資料品質控管程序與多個資料來源的控管程序差異甚大，現行資料倉儲技術可否有效提昇資料整合後的資料品質問題？事實上，資料品質的問題在傳統資料庫與整合性資訊系統（或資料倉儲）中最大的不同在於整合性資料通常具有一定程度的時序特性，資料的品質問題絕非突然間憑空形成，必定是經過長時間延續所累積而成，因此資料的品質與時序因素之間應存在著某種關係。未來如何探究資料時序特性及知識本體等相關學理技術，藉以提昇資料品質，當是亟具學理與實用價值的研究課題。
2. 如何運用整合性資訊提供「主動式」服務？舉例來說，倘若社會福利資訊系統與醫療照護系統的資料能有效整合，則醫療機構可以為獨居老人或遊民提供主動式的健康照護工作，尤有甚者，透過網際網路，可以進一步提供「遠端照護」服務。這類整合型、主動式的服務，應是未來公部門運用資訊科技提供創新性服務亟待發展之方向。

參考文獻

- [1] 行政院研究考核發展委員會《電子化政府報告書—93-94年度》(2005)，台北，中華民國
- [2] Alexiev, V. et al., *Information Integration with Ontologies, experiences from an industrial showcase*, Wiley Ltd, England, (2005).
- [3] Arpirez, J., *WebODE User Manual*, Technical Report, Technical School of Computer Science, Madrid (2001).
- [4] Chu, Y-C., *Integrating Heterogeneous Information Sources through Ontology-Driven Model and Data Quality Analysis*, Ph.D. Dissertation, National Taiwan University of Science and Technology, (2001).
- [5] Dasu, T. and T. Johnson, *Exploratory Data Mining and Data Cleaning*, John Wiley & Sons Ltd., Hoboken, NJ USA, (2004).
- [6] Inmon, W. *Building the Data Warehouse*, 3rd Edition, John Wiley & Sons Ltd., NJ USA, (2002).
- [7] Jones, D. M., T. Bench-Capon, and P. Visser, "Methodologies for ontology development," *Proc. of IT&KNOWS Conference of the 15th IFIP World Computer Congress*, Budapest, Hungary (1998).
- [8] Lenat, D., "CYC: A large-scale investment in knowledge infrastructure," *Communications of the ACM*, **38:11** 33-38, (1995). For further information, refer to <http://www.cyc.com/>
- [9] Mecella, M., *Cooperative processes and e-Service*, Ph.D. Dissertation, Universita Degli Studi Di Roma, Italy, (2003).
- [10] Miller, G., "WordNet: a lexical database for English," *Communications of the ACM*, **38:11** 39-41, (1995).
- [11] Park, J., and S. Ram, "Information systems interoperability: what lies beneath?" *ACM Trans. on Information Systems*, **22:4**, 595-632, (2004).
- [12] Pipino, L., Y. Lee, and R. Wang, "Data quality assessment," *Communications of the ACM*, April. pp. 211-218, (2002).
- [13] Pollock, J. and R. Hodgson, *Adaptive Information*, Wiley Interscience, USA, (2005).
- [14] Smith, B. and C. Welty, "Ontology: Towards a new synthesis," in *Formal Ontology in Information Systems*, C. Welty and B. Smith eds., ACM Press, USA, pp. iii-x (2001).
- [15] Uschold, M., "Creating, integrating and maintaining local and global ontologies," *Proceedings of 14th European Conference on Artificial Intelligence (ECAI'00)*, Berlin, Germany (2000).